Topology of stimulus space via directed network persistent homology

Samir Chowdhury, Bowen Dai, and Facundo Mémoli

THE OHIO STATE UNIVERSITY

Introduction

Causality analysis methods are often applied to spike train data to recover information about the excitatory/inhibitory nature of the neurons being observed, and also of hidden neurons that exert influence on recorded neurons. The causality maps which are outputs of such methods are asymmetric matrices that we view as directed networks. We consider the problem of postprocessing such asymmetric networks via a tool that has recently gained popularity for its ability to detect the organization of data—the **persistent homology (PH)** method. For example, prior results in the literature have established a "learning time model" that uses PH methods to recover topological information (i.e. connectivity of locations) about the stimulus space of rodent hippocampal place cells from their spike trains. By preprocessing with a causality analysis method before applying PH, we expect to decrease the error rate in such a model. However, standard PH requires input data to be symmetric, and cannot be applied directly to asymmetric causal interaction data. We explore a recently-developed tool called Dowker Persistent Homology (DPH) that accepts any asymmetric, non-metric data as input. Using this tool, we recover topological information from simulated spike train data with higher accuracy than the learning time model.

Contributions

We simulate a database of 3000 hippocampal spike train rasters from random walks of a rodent in arenas with different numbers of randomly placed obstacles. We process the rasters into directed, weighted networks, and then apply a recently developed topological data analysis technique called *Dowker Persistent Homology (DPH)*. We conclude that DPH is successful in recovering the ground truth topological information (i.e. the number of obstacles) from such spike train data, and that it exhibits higher accuracy in a classification task than existing approaches.

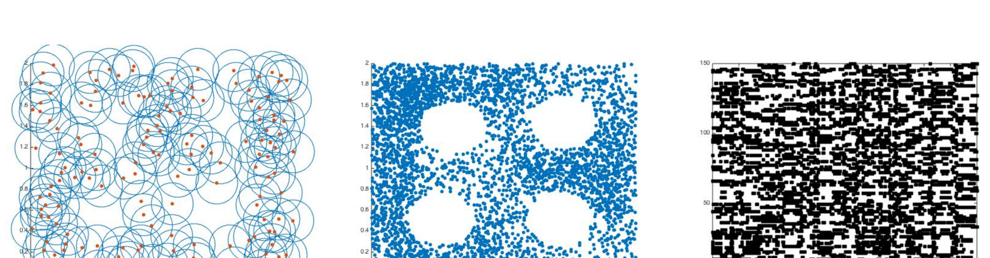

Figure 1: Sample place fields, place cells, and spike train

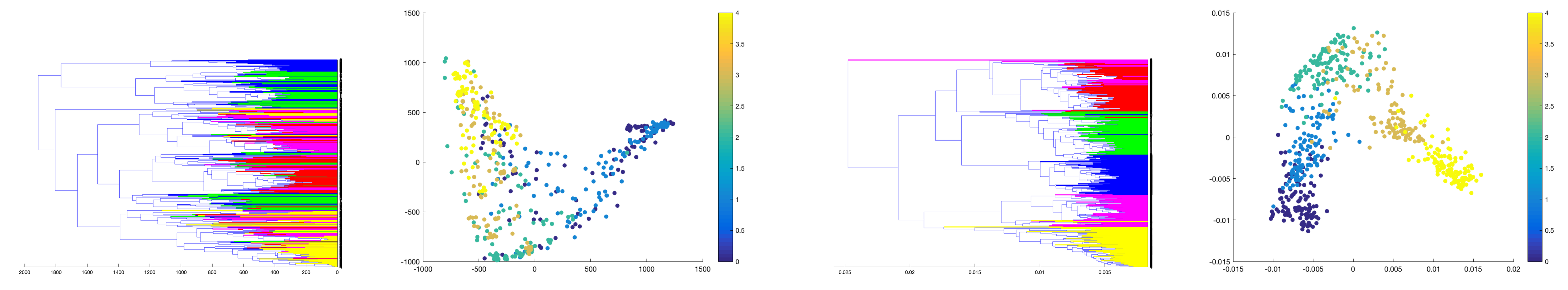

Figure 2: Single linkage dendrograms (0:blue, 1:green, 2:red, 3:magenta, 4:yellow) and 2 dimensional MDS plots of persistence diagrams produced by the learning time model (left) and the network DPH model (right)

Simulation Procedure

The first step was to generate 5 arenas by randomly choosing 5 locations to place obstacles, such that the obstacles do not overlap with each other or the arena's boundary. For each arena we chose 150 locations uniformly at random for place field centers such that the place fields covered the whole arena, and a corresponding random walk trajectory $R(t)$ recorded as a function of time. The spike train for each place cell was generated by an inhomogenous Poisson process. The mean μ_c of the Poisson firing model of each place cell c at step t is defined as follows:

$$\mu_c := f \times \exp\left(\frac{1}{2}\left(\frac{dist(c,R(t))}{1.2r}\right)^2\right) \tag{1}$$

Here f is generated by a log normal distribution whose mean is the max firing rate $r = 20Hz$ and standard deviation is $1.2r$, and $dist(c, R(t))$ is the distance between c and position $R(t)$ of the rat at step t. We first computed 20 spike train rasters using the Poisson firing model. Then we repeated the experiment by choosing a new trajectory and list of place field centers. In total, for each arena and choice of obstacle centers, we chose 6 trajectory-place field pairs, giving a total of 120 rasters. We repeated the process of choosing obstacle centers a total of 5 times, thus giving 600 rasters for each of the 5 arenas for a total of 3000 rasters.

Directed Network Transformation and Dowker Persistent Homology

A network is a set of nodes with pairwise edge relations given by real numbers. For each spike train, a network (X, ω_X) was constructed as follows: X consisted of 150 nodes (one per place cell), and for each $1 \leq i, j \leq 150$, the weight of edge (x_i, x_j) was given by $\omega_X(x_i, x_j) = 1 - \frac{N_{i,j}(5)}{\sum_j N_{i,j}(5)}$, where

$N_{i,j}(5) = \#$ time cell x_j spiked at least 3 times in a window of 5 steps after cell x_i spiked at least 3 times.

The DPH method proceeds by defining a nested sequence of *simplicial complexes*. Persistent homology methods are then applied to this intermediate construction to obtain an output called a *persistence barcode*. Examples are provided below. The long bars correspond to the number of obstacles in the arena, and short bars correspond to topological noise in the data.

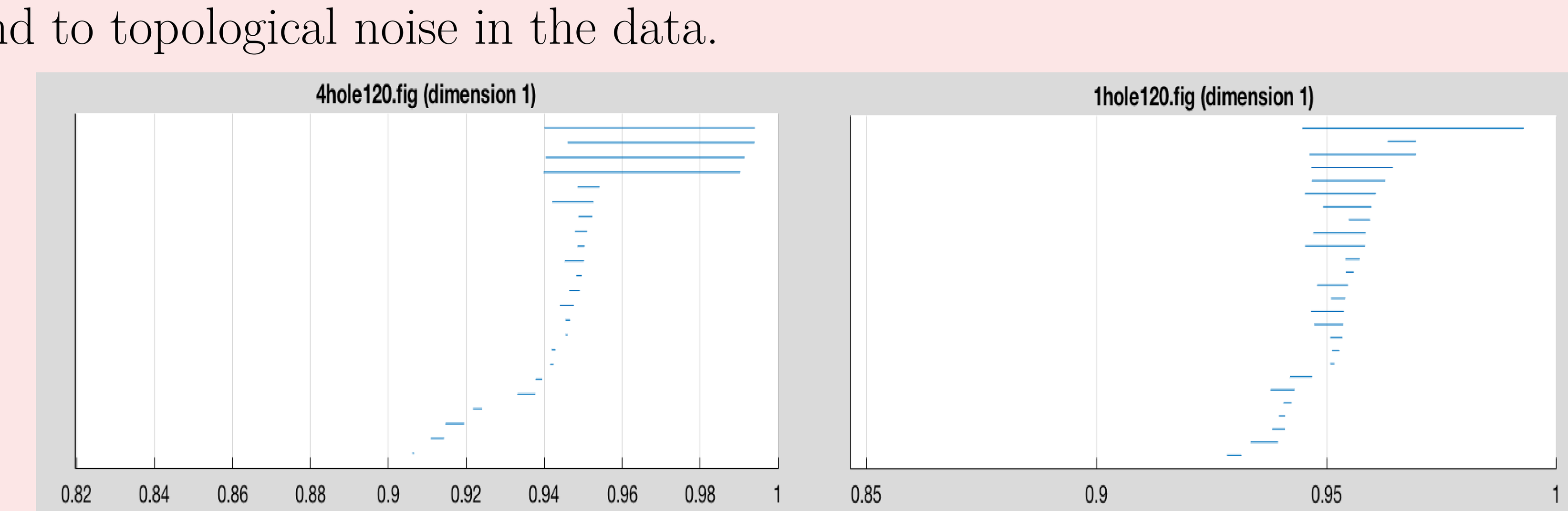

Figure 3: Dowker persistence diagrams of arenas with 1 and 4 obstacles

Variables	
Arena size	$2m \times 2m$
Radius of obstacle	$25cm$
Number of place fields	150
Max radius of place fields	$15cm$
Max firing rate	$20Hz$
Number of Steps	5000
Distance per step	15cm
1-nn error (600 rasters; network model)	0.19607
1-nn error (600 rasters; learning time)	0.5546

Table 1: Simulation variables and error rates.

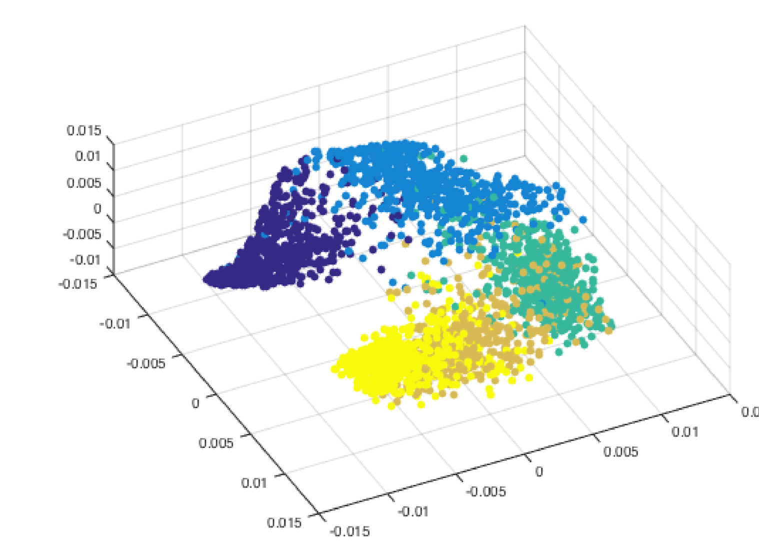

Figure 4: Slice of 3D MDS plot of Dowker persistence diagrams. Video on `research.math.osu/networks/Datasets.html`

Comparison to prior approaches

The DPH method was compared to an existing approach, the *learning time model* [1], for obtaining topological information from spike train data. This older model does not incorporate the asymmetric structure of the time series data in the spike trains, and its performance is worse than our method of applying DPH after first transforming to a directed network. On a restricted dataset of 600 rasters, the 1-nearest neighbor classification error rates for the learning time model and the network model were **0.5546** and **0.19607** respectively. The error for the network model on the full dataset was **0.1995**.

References

[1] Y. Dabaghian, F. Mémoli, L. Frank, and G. Carlsson, "A topological paradigm for hippocampal spatial map formation using persistent homology," *PLoS Comput Biol*, vol. 8, no. 8, p. e1002581, 2012.

[2] S. Chowdhury and F. Mémoli, "Persistent homology of asymmetric networks: An approach based on dowker filtrations," *arXiv preprint arXiv:1608.05432*, 2016.

[3] H. Edelsbrunner and J. Harer, *Computational topology: an introduction.* American Mathematical Soc, 2010.

[4] P. Andersen (editor), *The hippocampus book.* Oxford university press, 2007.

Acknowledgements

Supported by NSF DMS-1547357, CCF-1526513, IIS-1422400.