# Topological Characteristics of Neural Manifolds

Kosio Beshkov Supervisor: Paul Tiesinga

#### Abstract

In recent years, neural population activity has been analysed by treating it as a point cloud supported on a manifold whose structure gives information for the the type of computation that the network can perform and the features it can represent. Simultaneously a data focused approach to topology, which is a fundamental property of manifolds, known as *topological data analysis* (TDA), has also emerged. We use a method from that toolbox called *persistent homology*, it essentially finds the holes of different dimensions and sizes in point clouds and helps us understand the underlying manifold. We study the topology of neural populations by creating theoretical models capable of recreating a particular manifold's topology in their activity and also analysing the topological structure of neural activity during spontaneous and stimulus induced states in mouse cortex. We find significant differences between the topological structure of neural manifolds for different stimulus conditions across the brain.

# 1 Introduction

One of the fundamental goals of Neuroscience is to explain how our brains transform sensory inputs into useful and efficient behavior. However, the transformations of these inputs via brain circuits at the single-cell level are not known in sufficient detail. Furthermore, sensory inputs don't always influence observable behavior and a truly satisfying theory of cognition has to explain not only observable behavior but also the internal states that occur independent of said behavior.

There have been several studies that have looked to explain internal representations by showing that the activity in neural populations is constrained to low-dimensional manifolds, whose structure reflects the relevant variables of the representation (Gao and Ganguli (2015); Gallego et al. (2017); Chaudhuri et al. (2019); Saxena and Cunningham (2019); Mastrogiuseppe and Ostojic (2018)). A natural next step is to ask what it would take to construct a theory of neural coding purely in terms of manifolds. In the machine learning literature there similarly is increasing interest in the "manifold hypothesis", formalized in (Fefferman et al. (2016)). The basis of this research is that data itself lies on a low-dimensional manifold, which is consistent with the low dimensionality found in neural activity. Another condition for the manifold hypothesis to make sense in the context of cognitive science is that slight perturbations of a stimulus preserve its perceptual identity and that the relationships between stimuli are preserved in some manner in the relationships between the corresponding neural responses. This idea is also known as second order isomorphism (Shepard and Chipman (1970); Kriegeskorte et al. (2008)).

With these assumptions in mind, the remaining question we should ask is how do the manifolds emerge as a result of the transformations that neural networks apply to a set of inputs (the stimulus manifold). If we commit to the idea that networks in the brain extract particular features of a stimulus, then it is to be expected that the neural manifold representing the given features will have a different shape than the stimulus manifold, but the same shape as the one on which the stimulus features themselves lie. These transformations don't have to be continuous, which complicates their analysis. Nevertheless, we can study them by looking at the structure of the manifolds.

There are many ways to classify manifolds by their global structure, however we find that the most suitable tool for this type of analysis is *topological data analysis* (TDA) (Wasserman (2018)). Topology studies the shapes of objects under continuous transformations (homeomorphisms), a useful intuition for topology is to think of it as a geometry in which the objects are made from infinitely stretchable rubber. An example of a topological invariant which can be used to classify manifolds are the *homology* groups, which count how many holes of different dimensions there are in an object. To calculate these groups, a procedure called *persistent homology* (Carlsson (2009)) is used (**Methods**). While it is briefly described in the methods and appendix sections, for a more in depth treatment of algebraic topology see (Hatcher (2002); Munkres (2018); Edelsbrunner and Harer (2010)).

It is important to note that there are some previous works that use topolog*ical data analysis* to explore neural recordings. In (Curto (2017); Curto et al. (2017)) the authors show how cells with convex receptive fields can be used to cover feature spaces and how their coding activity can be mathematically formalized. There are works exploring the topological features in visual cortex (Singh et al. (2008)), but most of this work has focused on the hippocampus (Curto and Itskov (2008); Babichev et al. (2016)), where the authors look at how the overlapping structure of receptive fields in a population of place cells can be interpreted as a cover of the space which the rat is navigating. A similar line of theoretical research comes from (Babichev and Dabaghian (2018)) in which the authors explore how topology can be used to explain memory spaces in hippocampal neurons. Another more statistical approach is described in (Sengupta et al. (2018)), where the authors show that neurons in optimal similarity preserving networks (networks whose optimization function requires that similar inputs produce similar outputs) lead to the emergence of receptive fields which cover stimulus manifolds.

We theoretically study the topological properties of neural manifolds by presenting neural field models (Wilson and Cowan (1972); Wilson (1999); Bressloff (2011)) whose receptive fields cover feature spaces with smooth manifold structure and by inheriting a metric with which to specify their connectivity matrix, they respond in a stable and localized manner to each feature. In this case covering a feature is equivalent to having a receptive field responsive to particular external inputs, internal states, task rules or some combination of these (Fusi et al. (2016)). This is useful as it could allow us to model networks responsive to an arbitrary feature space, as long as it allows for the construction of a metric. After that, we tackle the question of how we can implement transformations on the feature space. We do this in a similar manner, by specifying two networks connected in a feedforward manner, such that the postsynaptic connections are determined by inheriting the metric structure of the transformed manifold.

We also analyze the topological structure of the activity manifolds of several regions in the Visual Coding Neuropixel mouse dataset from the Allen Brain Institute (Siegle et al. (2019)). We use the data to see how the transformations on the stimulus manifolds of static gratings and natural scenes change along the visual stream. Similar work has been done in (Singh et al. (2008)) but for a much smaller number of neurons in macaque visual cortex. We find that the natural scene and static grating manifolds have significantly different topological properties for each visual region that was analysed. We also find that the transformed stimulus manifolds are topologically complex and cannot be described by a few large topological features. After that we compare the complexity of topological structure in spontaneous activity across multiple hippocampal, thalamic and visual regions. We find that the hippocampus has a significantly more complicated topology than the visual cortex. Lastly we run simulations with which to see how functional connectivity impacts the topology emerging from a network.

# 2 Methods

There are two types of manifolds that can be extracted from neural recordings, we will refer to them as the neural and the activity manifold. In the neural manifold case we obtain a *neurons*×*neurons* distance matrix by calculating the euclidean distance between all pairs of vectors describing the neural activity of all pairs of neurons. This approach is taken when analysing the neural field models and is informative of the coding structure of a network. As our models cover manifolds with receptive fields, the distance between neurons is proportional to the overlap of their receptive fields. On the other hand, the activity manifold case is calculated by transposing the *neurons*×*times* matrix and calculating the distance between each pair of time points, after which we end up with a *times*×*times* distance matrix. This type of analysis is informative of the structure of neural activity and we use it when analyzing the data. This choice is consistent with the manifold representation of neural activity in past works on neural manifolds and allows us to work with more samples.

### 2.1 Simulations

### 2.1.1 Computational Model

The approach we take is to start with a particular feature manifold and cover it with receptive fields. This way a feature invokes activity only in the manifold region covered by the neurons tuned to it and is silent otherwise. Here we use adaptive neural field models which can exhibit complex patterns of activity, but any model in which the neurons cover the manifold and similar inputs lead to similar responses should lead to the correct topology.

In order to keep the activity localized to a particular region of the feature space, we use Mexican hat connectivity between neurons, which operates on the distance between features.

$$w[d(x,y)] = \omega_E e^{\frac{-d(x,y)^2}{\sigma_E}} - \omega_I e^{\frac{-d(x,y)^2}{\sigma_I}}.$$

In all of our simulations we set  $\omega_E = 5$ ,  $\omega_I = 3$ ,  $\sigma_E = 0.1$  and  $\sigma_I = 0.5$  which are constants that determine the width and amplitude of excitatory and inhibitory gaussians and d(x, y) denotes the distance between the positions of two features on the manifold. For the distance function we will use the geodesic distance on the manifold which is explained in more detail below.



Figure 1: Schematic drawing of our network setup.  $\mathcal{M}$  denotes a manifold covered by receptive fields indicated by the circles. Each receptive field is related to a neuron and the connections between the neurons are determined by applying a weight function w(x) to the length of the minimal geodesic curve  $||\gamma(t)||$  between the centers of the receptive fields on the manifold.

Once we know the distances between neurons in feature space, we can define the equations for the neural dynamics. Here we will use the neural field model with adaptation similar to the one presented in (Coombes et al. (2014)) except for the fact that the activation function is outside of the integral, which is equivalent to applying a non-linear integration of the inputs in the dendrites of the neuron. This makes our model an activity-based adaptive neural field model.

$$\tau \frac{\partial u(x,t)}{\partial t} = -u(x,t) - \beta v(x,t) + F[\int w[d(x,y)]u(x,t)dy + I(x,t)],$$
$$\frac{1}{\alpha} \frac{\partial v(x,t)}{\partial t} = u(x,t) - v(x,t).$$

Where  $\alpha = 1$  denotes the adaptation rate,  $\beta = 1$  denotes the adaptation

strenght,  $F(x) = \frac{1}{1+e^{-ax+b}}$  is a sigmoidal function with a = 5 and b = 10 and I(x,t) is the input at particular position and time.

### 2.1.2 Calculating the geodesic distance

In traditional flat space the shortest paths are straight lines and so the distance between two points is a straight line. However when a space is curved, the shortest path between two points also has to be curved, so it is necessary to find the straightest curves (also called geodesics) between two points and take the shortest one to be the distance between the points. The geodesic distance is defined as the infimum of a functional in the following manner:

$$d(x,y) = \inf \int_{a}^{b} \sqrt{g_{\gamma(t)}[\dot{\gamma}(t), \dot{\gamma}(t)]} dt$$

Where  $g_{\gamma}(t)$  describes the metric tensor and  $\gamma(t)$  describes a curve parametrized by  $t \in [a, b]$  which starts at x and ends at y.

The explicit form of the distance can be found by solving the geodesic equations:

$$\frac{d^2 x^m}{dt^2} = \Gamma^m_{\sigma\rho} \frac{dx^\sigma}{dt} \frac{dx^\rho}{dt}.$$

Where the  $\Gamma_{\sigma\rho}^m$  is the Christoffel symbol and the superscripted x's are the coordinates in which we choose to describe the manifold, the choice doesn't matter (for more detail see Kreyszig (1991); Carroll (2019); Tu (2011)).

All of the geodesic equations that were used in our simulations were calculated analytically before the numerical simulation. For the circular model we used a circle with radius R = 0.25 and for the spherical model we used R = 0.2. In each case we sampled the surfaces uniformly.

### 2.1.3 Stimulation and Numerical methods

For all simulations on manifolds, we defined our inputs so that they uniformly activate every position on the manifold. If this isn't done then there is no reason to expect that the manifold structure will be reproduced. This is because sampling only a subset of the manifold is equivalent to covering a submanifold which doesn't necessarily preserve the topological properties of the original manifold.

The simulations were done in Python where we used a fourth order Runge-Kutta method, with a step size of  $\mu = 0.05$ . For the circle and the sphere case we used N = 1000 neurons, for the network implementing a transformation we used N = 2000 neurons, 1000 for each manifold.

# 2.1.4 Integrate and fire networks with varying sparsity and weight strength

The integrate-and-fire network model (Gerstner et al. (2014)) can be written as:

$$\frac{du}{dt} = u_{rest} - u(t) + W \cdot \Delta(t - t^f) + I(t).$$

Where W is the weight matrix,  $\Delta(t - t^f)$  is a vector containing a Dirac delta function representing the spikes  $120\delta(t - t^f)$  of each neuron. The resting potential was set to  $u_{rest} = -60$ , the spike threshold -  $\theta = 0$  and the reset potential  $u_{reset} = -70$ . I(t) is the varying external input driving the network, we chose to use Poisson spikes of size 180 with 20% probability. The step size we used was  $\mu = 0.05$ . We simulated a number of N = 80 neurons to approximately match the neurons analysed in the datasets. The excitatory and inhibitory weights were drawn as max(0, s) where s denotes a sample from a normal distribution centered at 5(excitatory) or -5(inhibitory) with varying standard deviation ranging from 0.5 to 6. The inhibitory weights were multiplied by 4 and were assigned to 25% percent of the neurons. The sparsity was varied between 0 and 0.9. A sparsity value of 0.5 (for example) means that for each entry in the weight matrix there is a 50% probability of that entry being set to 0, regardless of whether it is excitatory or inhibitory.

### 2.2 Data analysis

All data analysis is performed on the Allen Institute Visual Coding - Neuropixels dataset (Siegle et al. (2019)). Cells were selected based on their signal to noise ratio (SNR > 2) and firing rate (> 1Hz). This choice was made so that we only work with cells which are active and reliable. We compared both the visual stimulus induced and spontaneous activity in 4 mice. For the stimulated conditions we analysed the neural manifold response to static gratings and natural scenes. To compare them we used a persistence image representation of their topology (Adams et al. (2017)). Since our goal is to quantify not only the manifold structure of activity in different regions, but also to use that information to explore the type of transformations that different regions could be doing, we define a scalar statistic with which the regions can be compared. We use this measure when analysing the spontaneous activity.

### 2.2.1 Data preprocessing

In the stimulated recordings, we use 0.025s bins for all stimuli (120 static gratings and 119 natural scenes) and all 50 trials (each trial has a duration of 0.25s) per stimulus. Once we have a *neurons*×*times* matrix for each stimulus presentation, we average each stimulus across its trials and concatenate the resulting matrices for all stimuli. We do this separately for the static gratings and the natural scenes. Since this is done for each region separately, we end up with two large point clouds per region representing the activity manifold for the two stimulus types.

In the spontaneous recordings, for every mouse, the spike count during 3 intervals without stimulation with a duration of approximately 300s were binned into 0.25s windows, which provided us with a stable spike count. Consequently, we split the long trials into 20 second long subtrials. The point of creating the 20 second subtrials was to use them as subsamples on which to calculate statistics.

The spontaneous activity dataset contains recordings from a vast amount of regions, however some of them have too few cells to analyze further. To avoid such problems the analysis was done on regions in which at least 15 cells, which pass the aforementioned SNR and firing rate constraints, were recorded.

### 2.2.2 Dimensionality Reduction

Each region contains a different number of neurons (N) passing the selection criteria. We tried to see whether a 15 dimensional embedding explained more than 90% of the variance in every case. This turned out to not be true, so we decided to not use dimensionality reduction in our analysis. However we still used it in order to visualize the manifolds of our target regions and those generated from the neural field models.

The 3 dimensional embedding was done with the Uniform Manifold Approximation and Projection (UMAP) algorithm (McInnes et al. (2018)), which is a relatively recent nonlinear dimensionality reduction algorithm. I takes two parameters - number of neighbors and minimal distance. We set the number of neighbors to 40 and the minimal distance to 0. As the goal of the study is to see if the topology of neural activity varies across regions, for the visualization it makes the most sense to choose parameters which are sensitive to both local and global topological features. Embedding with a large number of neighbors captures the global structure of the manifold, whereas the low minimal distance allows for the discovery of fine topological details.

### 2.2.3 Distance matrix computation

The *persistent homology* algorithm (explained below) which is the essential tool in our analysis needs to be supplied with a distance matrix. The standard euclidean metric was used on each trial.

Persistent homology takes a distance matrix and varies a parameter to get a set of simplicial complexes, such that by increasing the parameter value, the previous complex is always a subset of the next one. In formal notation, where  $\subset$  denotes the subset relation, this can be written as:  $S_0 \subset S_1 \subset ... \subset S_N$ . This means that as long as the ordering of inclusion of two simplicial complexes is preserved, their topological features, but not necessarily the scale of their persistence, will be the same. This implies that the ordering is preserved under nonlinear monotonic transformations (Giusti et al. (2015)).

Sometimes it is also useful to focus on the purely topological features of a point cloud, ignoring geometric features like scale. A transformation that fulfills both of these requirements is an ordering function which takes a matrix and outputs another matrix in which each entry is assigned a number indicating its rank. After that, all entries are divided by the maximal entry to normalize the matrix. We use this approach when analysing the stimulated trials.

#### 2.2.4 Persistent Homology

*Persistent homology* calculates the *homology* groups at differing scales and outputs a birth-death diagram. The distance matrices indicate a set of points in a metric space, then the Vietoris-Rips complex (Zomorodian (2010)) is calculated by choosing a neighborhood radius around each such point and adding a (k-1)-simplex whenever the distance between k points is less or equal to the radius. Increasing the radius parameter induces a sequence of simplicial complexes:

$$S_0 \subset S_1 \subset \ldots \subset S_N.$$

*Persistent homology* then calculates the *persistent homology* groups for each such simplicial complex (This is described in more detail in **Appendix**), so we end up with a set of topological features for each parameter value:

$$PD = \{H(S_0), H(S_1), ..., H(S_N)\}.$$

Every topological feature appears in one of the simplicial complexes and disappears in a later one, as a result the persistence of a feature is defined as the time at which the feature disappears (the death) minus the time at which it first appears (the birth). For the calculation of the *persistent (co)homology* we used the Python package Ripser (Tralie et al. (2018)).

### 2.2.5 Persistence Complexity and Persistence Images

Even when normalized, persistence diagrams are difficult to compare as they are sets with differing number of potentially identical elements (multisets), because of this they are not suited to statistical analysis. To escape this problem it is necessary to either define a measure which can be extracted from a diagram and analysed separately, or map the diagrams to a space suited for statistical analysis.

The former is achieved by defining the *persistence complexity*, which is just the sum of persistences of all features in a diagram (A similar strategy is used for analyzing deep neural nets in Naitzat et al. (2020)). In other words we can define the persistence of a feature  $\psi_i^j$  as  $p(\psi_i^j)$ , where *i* denotes the dimension and *j* - the index of the feature under some ordering  $j \in J$ . The persistence complexity can then be written in its general form as:

$$\sum_{i=0}^{\infty} \sum_{j} p(\psi_j^i).$$

Here we only calculate the 1 and 2 dimensional features for computational reasons. The persistence of 0 dimensional holes was ignored as it doesn't provide information about interesting features. This measure was only used to compare regions in the spontaneous recordings.

There are several strategies to map diagrams to a statistically meaningful space (Bubenik (2015); Hofer et al. (2017); Adams et al. (2017)). One of these

strategies is to map the persistence diagram to a vector space in which the distance between two vectors can be used as a test statistic. Here, we do this with *persistence images* (Adams et al. (2017)) which rotate the persistence diagram to (birth+death)×(death-birth) coordinates and apply weighted 2D Gaussian kernel functions to every feature. Once that is done the diagram is transformed into a 2D image which can be flattened into a vector and used to calculate statistics.

The persistence diagrams can also have a weighting function, kernel width and a pixel size. For the stimulated recordings we use a pixel size of 0.01 and a kernel width of 0.001. For the weighting function we used the persistence identity function.

### 2.2.6 Statistical tests

To analyze the persistence complexity across regions, histograms of each region's persistence complexity were created. After that a Kolmogorov-Smirnov two sample test was performed for each pair of regions. Also a bootstrap method with 10000 resamples was used to determine the confidence intervals of the means of the distributions.

To analyze the persistence images, a Monte-Carlo based permutation test with 1000 permutations was used. The test statistic was defined as  $S_{test} = \sum_{ij} \sum |X_i - Y_j|$ , where  $X_i$  and  $Y_j$  are two flattened persistence images (vectors), the inner sum goes over the entries of the resulting vector and the outer sum goes over all pairs of vectors in the two samples. Each region contained one point cloud per stimulus type, so in order to get more samples we subsampled the points before running the permutation tests.

# **3** Neural Fields on Manifolds

Numerical simulations of neural field models on curved geometries have been explored in (Martin et al. (2018)) and the solutions on a spherical geometry are studied in (Bressloff (2019)). There are two main differences in our approach, first we don't include the volume element of the manifold in the integral and second we study the topology preserving properties of such networks instead of their specific mathematical solutions. As a result, our analysis does not aim to provide new insights as to the precise dynamical effect of varying the parameters of the model over different manifolds. With regards to the former, our end goal is to cover feature spaces and including the volume element means biasing some neurons simply because of their position in feature space, but every feature is a priori equally important. Not doing this can result in some bumps over regions of feature space, where the volume element is high, becoming more stable than others. This can ruin the topological structure in the network, as activity concentrates over the highly biased regions. As a result the manifold structure of feature space is reflected purely in the geodesic distances between neurons in feature space. With regards to the latter, the neural field equations are simply a stepping stone to study the topological structure of neural activity and that can be done with many different models, our results were also confirmed for spike counts coming from integrate and fire models. This procedure also gave worked appropriately for our goals.

### 3.1 Covering a single feature space

To confirm that our models correctly cover feature manifolds, we ran simulations for both circular and spherical manifolds. The circular neural field model has been used in the past to model head direction circuits (Redish et al. (1996)). Here we show that this neural field model reproduces the topology of the circular feature space. This is expected since due to the choice of our connectivity function, nearby (in feature space) neurons will have correlated activity and as a result will also be closer in neural activity space. The result for this network is shown in figure 2, panel **B** shows a barcode plot (Ghrist (2008)) of the neural manifold. The way to interpret barcode plots is to consider each line as a topological feature of some dimension and the length of the line as the persistence or the "importance" of the feature (for an in depth description of *persistent homol*ogy look in **Appendix 7.1**). Here we will show at most 3 topological features, denoting the number of connected components ( $\beta_0$  - blue), 1 dimensional holes which intuitively are like the holes inside a ring ( $\beta_1$  - red) and 2 dimensional holes intuitively analogous to the cavity inside a football ( $\beta_2$  - green).



Persistent Homology of the Circular Neural Field Model

Figure 2: Visualization of the topology of neural activity in the circular neural field manifold model. **A**) the left panel shows the manifold on which the model is defined and the right panel shows a UMAP visualization (**Methods**) of the position of each neuron. **B**) shows a barcode plot indicating the topological features  $\psi^0$  and  $\psi^1$  characteristic of the network.  $\epsilon$  denotes the scale of the radius parameter used in persistent homology.

In the circular case it is very clear that the neural field model successfully reproduces the topological structure of the feature space. This is a very simple example of a neural manifold, but our strategy generalizes to more complex manifolds as well (**Appendix 7.3** contains a more detailed motivation for this statement). Figure 3 shows that it also works in the case of the sphere where we have used the great circle distance to compute the geodesics. It is clear that the green bar denoting the cavity in the sphere is the most persistent component. The red components denoting one dimensional holes are not very persistent, but there is still the question of what their meaning is. In this case their presence is simply due to the fact that we can only use a finite amount of neurons with which to cover feature space and there is variability in the position of the neurons in neuron activity space. Hence when the radius parameter is not large enough, holes will appear in the sphere. This type of 1 dimensional cluster of holes can be characteristic for randomly sampled points in space (Giusti et al. (2015)), the only difference in the spherical manifold case is that the randomness occurs

on the surface of a sphere. While the most persistent features are considered to be the most important, *persistent homology* also provides information in the features that are not as persistent. For example there are mathematical analyses showing it contains information about the curvature of a surface (Bubenik et al. (2020)). While this problem is difficult, in the appendix we document what these other smaller components could be showing us. More importantly it is interesting to think what exactly they could tell us about neural activity.



Persistent Homology of the Spherical Neural Field Model

Figure 3: Same as figure 2 but in the case of the sphere

These example networks cover incredibly simple feature spaces and the problem of covering more realistic spaces isn't as simple, since the neurons have to be distributed uniformly on the manifold and proper metric tensor has to be chosen to calculate the geodesics. However, as long as we can compute the distances between enough features, our strategy should lead to a good covering of the underlying manifold. This can be useful in several cases, to model the activity of networks coding for certain features and to potentially think about the specific transformations between layers, regions or other biologically interesting categories which might play a role in neural coding.

### **3.2** Networks realizing transformations

To give an idea of how we might be able to study the maps between networks we first provide a neural field example in which one network covering a feature space connects to a second network in such a way as to implement a transformation of the first feature space. We can think of networks in the brain as a set of transformations on feature spaces and this approach allows us to specify the connectivity structure necessary to realize an explicit transformation. As a result every subnetwork implements a transformation on the feature space from which its inputs come. As we will show in the data analysis section this is harder said than done, but it still seems worthwhile to outline one way to implement transformations with neural field models.

The question is how do we connect the first to the second network in such a way as to achieve a specific transformation. To do this it is helpful to think about the issue in a slightly more abstract manner. Every transformation between two manifolds  $F: \mathcal{M} \to \mathcal{M}'$  can be used to define a distance metric using the geodesic distance on  $\mathcal{M}'$  - expressed as  $d_{induced\mathcal{M}'}(F(x), F(y))$  (also known as the pullback of the metric). This equation provides a clear instruction about how to connect the networks, namely the weighting function can be applied to the distances between points in the second manifold. In practice we sample coordinates for the neurons in the second network and calculate the distances between them and the projections coming from the first network. In other words if we denote the sampled coordinates of a neuron in the second network as  $\mu$ and the projected neuron coordinates as F(x), we can calculate the distance matrix by computing  $d_{\mathcal{M}'}(\mu, F(x))$ . The same calculation is done in the case in which two neurons come from the second network -  $d_{\mathcal{M}'}(\mu_1, \mu_2)$ . The example we have used is a network defined over a sphere connecting to a second network in such a way as to contract the sphere onto a circle. What this should result in, is a transformation of the original spherical feature space into a circular feature space. The explicit form of the transformation can be written as the function  $F: (R\cos\theta\sin\phi, R\sin\theta\sin\phi, R\cos\phi) \to (R\cos\theta, R\sin\theta, 0)$ . In our example we stick to networks with an equal amount of neurons and the metric on the second network is the same as the one on the circle. Figure 4 shows the results, both the UMAP embedding and the barcode plots show that the mapping successfully preserves the topological structure which we would expect in both networks.



Persistent Homology of two networks implementing a transformation on a manifold

Figure 4: The topology of two networks where the first excites the second in such a way as to map features on a spherical surface to a circle.

# 4 Topological Structure of Neural Activity Manifolds

To get an idea of how complex the topological structure of neural activity is, we analysed data from the Visual Coding Neuropixel dataset from the Allen Brain Institute (Siegle et al. (2019)). This dataset is mainly focused on the activity in the visual stream during stimulus processing, however it also contains recordings from thalamic and hippocampal subregions during spontaneous activity. Calcium recordings for the same stimulation protocols have been analysed in (de Vries et al. (2020)) The topological features in the visual cortex have previously been explored in (Singh et al. (2008)), however the data that they analysed had only a limited number of neurons which were located solely in V1. The Allen Institute dataset contains stimulated recordings from hundreds of neurons spread across several areas, which allowed us to study neural activity during different stimulation protocols across a larger number of areas. In their study, the authors find many persistent topological components, but we see that the topological structure is much more complex and single large persistent features don't appear as often in the larger populations of neurons that we analysed. A recurring theme in this analysis is that even when there is no clear big topological features, there could still be significant differences between two activity manifolds. However the difference becomes very difficult to interpret and for the moment we will use these features only as a marker signifying that there exists a difference between the transformations which two networks (or one network for different stimuli) implement on a stimulus manifold and these differences can be indicative of either complicating or simplifying the structure of the activity manifold.

We begin by analysing the topology of neural activity during the visual stimulus processing of static gratings and natural scenes. After that we move on to study region specific topological structure during spontaneous behavior.

### 4.1 Topological structure of neural activity during stimulus processing

We analyse the topological features of five regions related to visual stimulus processing. The only thalamic region included in this analysis is the Lateral posterior nucleus of the thalamus (LP) which makes up the visual pulvinar and is implicated as a high order region in a functional hierarchy of visual processing (Siegle et al. (2019)). The remaining four regions are the primary visual area (VISp), the rostrolateral visual area (VISrl), the lateral visual area (VISp), the rostrolateral visual area (VISrl), the lateral visual area (VISI) and the anteromedial visual area (VISam). In order to compare the persistence diagrams (also known as barcode plots), we used a tool known as persistence images (**Methods**). The idea behind it is that persistence diagrams can be represented as a picture by putting a 2 dimensional gaussian at each point in a persistence diagram, after that the image can be flattened and treated as a vector, which allows us to use it to compute statistics. Panel B of figure 5 shows the flattened persistence images for the concatenated 0th, 1st and 2nd Betti numbers.

By inspecting the barcode plots, there seems to be several consistent patterns in which the barcodes describing neural activity during static grating presentation differ from those during natural scene presentation. These differences were confirmed by performing a permutation test on the flattened persistence images.



Topological Analysis of Visual regions during stimulus processing

Figure 5: **A)** Barcode plots for all visual regions during processing of static gratings or natural stimuli. **B)** Flattened persistence images (histogram) which were used to verify the statistical significance of the difference between the topology in for the two stimuli types. The three bumps reflect the three different topological components that were analyzed. The \* symbol indicates that a Bonferonni correction for 5 tests was used.

For instance, in the static grating case the 0th components disappear earlier

than those in the natural scenes. Since the set of 0th components describe the connectivity of a space, we can interpret this result as saying that the activity of networks participating in visual processing, tend to be more compactly clustered when presented with the simpler static grating stimuli. A possible explanation for this could be that static gratings are both more similar to each other and also explore a smaller subset of the manifold of possible visual stimuli. This could be reflected in the structure of the responses. A second prominent difference between the two stimulus types is the later birth of both the one and two dimensional holes. This difference is harder to interpret, one reason could be the same as that for the 0th components, the point cloud covers more space so as the components start to connect later, the holes also start to appear later. It is also plausible that the activity manifold responding to natural scenes simply contains larger scale features. Understanding the manifold structure of natural scenes is a complex problem, so far only the space of 3 by 3 patches is known to live on a Klein bottle (Carlsson et al. (2008)).

It does seem like *persistent homology* is a method capable of differentiating between the responses to a stimulus manifold. However it is generally a difficult problem to interpret the output, since there are too many features which could appear for a large number of reasons (in the appendix we provide an intuition for how to interpret these components, by looking at how they change under continuous deformations of a point cloud).

## 4.2 Topological structure of different regions during spontaneous behavior

To explore the differences between the topological features across regions we defined a scalar metric which we call the persistence complexity (**Methods**), it is essentially the sum of the persistence of every topological feature in this case we chose to focus on features of dimension 1 and 2. Another way to think of this is as a measure of how many and how big the hollow circles and spheres in a space are. Due to its scalar nature, this measure contains less information than the persistence image representation, however it is easier to interpret - the higher it is the more complex components there are in the space.

In figure 6 we show the differences in persistence complexity across regions for 4 mice. To see if the differences are significant we have applied a Kolmogorov-Smirnov 2 sample test with a Bonferonni correction to the significance threshold (Wasserman (2013)).



## Topological Analysis of brain regions during spontaneous activity

Figure 6: The results for 4 mice from the Kolmogorov-Smirnov test. A) matrices showing the K-S statistic and the associated Bonferonni corrected p values. B) barplots showing the mean persistence complexity of each region.

There is some inconsistency in the regions being analysed, since not every region was recorded in each mouse. This is why some bars are missing in panel

- B. The target regions that we analyse are split into three groups:
  - 1. Hippocampal (HIP) including the CA1, the dentate gyrus (DG) and the subiculum (SUB)
  - 2. Thalamic (THA) including the dorsal (LGd) and ventral (LGv) lateral geniculate complex and the lateral posterior nucleus
  - 3. Visual cortical (VISc) including the same regions as in the analysis of the stimulated recordings.

The last three entries of the matrices and the last three bars in the bar plots show the differences between the three groups of regions we determined. In general there are many differences between regions and not all of them are consistent, however one consistent result is that hippocampal neurons seem to form more topologically complex spaces than visual ones. This is consistent with recent results showing the presence of strong topological features in the hippocampal head direction system (Chaudhuri et al. (2019)) and theories as to the function of hippocampus (Babichev and Dabaghian (2018)). The specific value of the persistence complexity seems to also be relatively consistent for the CA1 and hippocampus, which can be used as a characteristic marker for these regions. However this is just an observation and confirming its consistence across mice will involve analysing many more subjects.

One cause for the high topological persistence found in Hippocampus, could be due to functional connectivity patterns similar to those found in Hopfield networks (Hopfield (1982)). This inspired us to check what is the impact of functional connectivity on the topological complexity of the activity manifold. We simulated balanced integrate and fire networks (Brunel (2000)) with varying sparsity and weight strength. In figure 7, the persistence complexity as a function of sparsity and weight strength is shown.



# Persistence complexity surface of balanced integrate and fire networks

Figure 7: Persistence complexity as a function of weight standard deviation and sparsity.

As sparsity and weight strength increase so does the persistence complexity up until a certain point at which increasing the variance starts to decrease the persistence complexity. This result indicates that balanced sparse networks with strong, but not too variable, connections are more likely to generate complicated activity manifolds.

# 5 Discussion

We have presented both theoretical and empirical results, based on *topological data analysis*, about the spaces underlying neural activity and their role in understanding the neural code. Under the proposed view neural populations have the role of applying transformations of differing complexity to stimulus manifolds. Stacking these transformations can produce representations of varying complexity. We have also proposed an explicit strategy with which to create network models that realize a particular manifold and can also be used to define transformations between manifolds. While interesting, there are several limitations to our approach. For one it involves explicit specification of the metric, which is not easy to generalize to more complex and realistic scenarios where there is no way to determine a metric for the manifold of the feature space which a network is coding for. Another complication comes from the fact that these transformations are easy to define only when it is possible to separate connectivity in such a way as to identify two (or more) disjoint networks that

have feedforward connectivity among themselves.

On the empirical side we have analysed both stimulated and spontaneous recordings of mouse data. The domains in which topological data analysis has had the most success are those in which the information being coded for is simple enough to be mapped to a simple space like the circle in head direction cells (Chaudhuri et al. (2019)) or the torus for grid cells (Moser et al. (2008); Guanella et al. (2007)). In this work we have argued that, while easier to interpret, by only looking for components with large persistence values we might miss important information about the manifold structure at smaller scales. For this reason we have used persistence images and have defined a global scalar measure called persistence complexity as a summary of the topology of an activity manifold. We have shown that the persistence images of neural activity differ between the static grating and natural scene stimulus manifolds. With the natural scene activity manifold showing components with a later birth and a longer persistence. We speculated that this is due to the fact that natural scenes are more spread out in possible visual stimulation space, so the components start to connect later. The analysis of the spontaneous recordings showed that there are significant differences between the persistence complexity across regions. A consistent pattern showing that hippocampus has a larger persistence complexity value than the visual cortex emerged across all mice. We simulated integrate and fire networks to check how functional connectivity relates to persistence complexity. We found out that sparse connectivity with synapses with large (but not too large) variability leads to a high persistence complexity. This result does not support the assertion that Hopfield-like structure leads to high persistence complexity. However, this is not the only reason why the persistence complexity can be high, as there are many more connectivity patterns which weren't explored in this work. An interesting further analysis is to explore whether the functional connectivity of a model approximating the dynamics of a region's activity can be varied as to generate different persistence complexity values. Another future approach to study the same question is to simulate more complex connectivity patterns, like small world and scale free networks and see what the impact of these architectures is on the persistence complexity. While our analysis finds many significant differences between regions and provides potential reasons for their different persistence complexity values. It is still difficult to really understand what the components mean, especially when we try to focus on the whole set of features, not only on the really big and obvious ones.

In conclusion the results show that there are interesting features which a topological approach to neural activity can reveal, but extracting and interpreting them comes with many challenges. It is a matter of future theoretical research to see whether there is a particular computational role of topological features of a given scale. We hope that our results can be used as an inspiration for a further, more focused and systematic exploration of manifold structure underlying neural activity.

# 6 Appendix

### 6.1 Persistent Homology and its prerequisites

Topological data analysis (TDA) is a relatively novel approach to data science which applies tools from algebraic and differential topology to study the structure of data. Several clear introductions specifically suited for neuroscientists are presented in (Sizemore et al. (2019); ), there are also more mathematically focused resources presenting algebraic topology in more detail (Hatcher (2002); Munkres (2018)) and more specifically computational topology (Edelsbrunner and Harer (2010)). Here we will give a short but self-contained introduction to TDA.

The field of Topology studies the shapes of objects, by only considering those properties that remain unchanged after the object is deformed continuously (for example by being locally or globally stretched or shrunken). In other words, it studies properties which are invariant under continuous transformations. This way one can draw an equivalence between objects that share such an invariant but have differing geometric features. The invariant that we will consider in this text is the number of holes. The notion of a hole will be made more precise and generalized to higher dimensions later, but from this point of view, a circle and a cylinder or the plane and a cone are the same object.

To find the holes in a point cloud we need to calculate its *Homology* groups, this is done by associating something called a *Chain Complex* to a given point cloud. While there are several variants of *Homology*, we will work with *Simplicial Homology* in which we associate a simplicial complex made up of n-simplices to a given space.

### 6.1.1 Simplicial and Chain complexes

As already mentioned one of the most basic objects in TDA is the simplicial complex. It is a more general version of the already familiar notion of a graph/network. First we will need to look at graphs in a more abstract manner, not as a picture in which vertices/nodes are connected by edges/connections but as a set of relations. Vertices/nodes describe the relation between only one object  $(v_1)$  and edges/connections describe binary relations between two vertices/nodes  $(v_1v_2)$ . Following this line of logic, we can also write higher order relations as  $(v_1v_2...v_{n-1}v_n)$ , we will refer to these as (n-1) simplices. In this formalism vertices/nodes are called 0-simplices and edges/connections -1-simplices (from now on we will refer to them as simplices instead of vertices or edges). There is also a more geometric way to look at things, namely a 2simplex can be visualized as a filled in triangle and a 3-simplex - as a filled in tetrahedron (this is also known as their geometric realization). Sadly, our geometric intuition breaks down for higher order simplices so it is easier to think of the underlying abstract structure. Having an intuition behind the structure of simplicial complexes we can start describing them in a formal manner. The geometric realization of an n-simplex can be written as a set:

$$\Delta^{n} = \{(t_{0}, ..., t_{n}) \in \mathbb{R}^{n+1} | \sum_{i} t_{i} = 1 \text{ and } t_{i} \ge 0 \text{ for all } i\}.$$

This definition puts a simplex in the n+1 dimensional plane, by essentially identifying its vertices with the unit vectors on each axis of n-dimensional Euclidean space.

The abstract simplicial complex is a collection X of simplices, with the property that,  $\sigma \in X$  and  $\tau \subset \sigma$  imply that  $\tau \in X$ .

As previously mentioned we can write an n-simplex as:  $\sigma = [v_0, v_1, ..., v_n]$ . Every n-simplex has a set of (n-1)-simplices which are called its faces and can be written as the set:  $\{\tau_i | \tau_i = [v_0, v_1, ..., \hat{v_i}, ..., v_n]\}$  where the hat operator deletes the element over which it appears.

For a simplicial complex  $\Delta$  on a *Topological Space* X, we can associate a *Chain complex* C(X). The chain complex identifies an algebraic structure with the simplicial complexes of a given dimension. Furthermore it includes a directed sequence of maps (homomorphisms) between every group (Bhattacharya et al. (1994) for a reference on groups) of n and n-1 simplices, such that the image of each map lands in the kernel (null space) of the next map.

The chain complex contains groups (vector spaces can also be used to explain this) for the simplices present in every dimension, these groups are called the chain groups and are denoted by  $C_n(X)$ . The elements in each such group are called n-chains and are written as a formal sum  $c_n = \sum_i a_i \sigma_i^n$ , where  $a_i$ are coefficients coming from some Abelian group (a group with a commutative operation). This group (or vector space) structure essentially allows us to do a type of algebra with the different n-simplices by combining them into chains. As we already observed simplices of different dimensions are related by their faces. This fact allows us to define the boundary operators, which map any n-dimensional chain to an (n-1)-dimensional chain in the following way:

$$\partial_n : C_n \to C_{n-1},$$
  
 $\partial_n(\sigma^n) = \sum_{i=0}^n (-1)^i [v_0, v_1, ..., \hat{v_i}, ..., v_n]$ 

One last important property of the boundary operator is that  $\partial_{n+1} \circ \partial_n = 0$ . This property implies that the image of  $\partial_{n+1}$  is contained in the kernel of  $\partial_n$  or  $\operatorname{Im}(\partial_{n+1}) \subset \operatorname{Ker}(\partial_n)$ . We will call the chains  $c \in \operatorname{Ker}(\partial_n)$  cycles and  $b \in \operatorname{Im}(\partial_{n+1})$  boundaries.

### 6.1.2 Homology

Now that all of the necessary concepts have been laid out, it is time to give an intuition and precisely define the concept of a hole. The first thing to realize is that cycles can either surround something or have an empty interior. Imagine a trampoline, it has an elastic interior and a solid outer frame surrounding the

elastic part. There is no hole in this configuration, however if we take out the elastic part the outer ring now surrounds an empty space or in other words - a 1-hole. The same concept can be generalized to higher dimensions, the only difference is that the surface surrounding empty space is higher dimensional (a higher dimensional example is a football which is a 2 dimensional surface surrounding an empty interior - a 2-hole).

From this intuitive example it is clear that for there to be a hole we need a cycle that surrounds an empty space, so the definition we need has to consider all cycles with an empty interior and exclude those that are filled. This gives us the definition of the *homology* groups.

$$H_n(X,G) = \operatorname{Ker}(\partial_n) / \operatorname{Im}(\partial_{n+1}).$$

The terms in the kernel are cycles and we exclude the cycles that are filled in by quotienting by the elements in the image of the higher dimensional boundary map. Thus for every dimension we have an Abelian group (or vector space) and its rank tells us the number of holes present in the Topological Space we started with. The rank of each *homology* group is called the *n*-th Betti number  $\beta_n = \operatorname{rank}(H_n(X;G))$ . As a result any space can be specified by the sequence of its Betti numbers. Figure 8 shows the Betti numbers of some surfaces, it should be noted that there are surfaces which are not homeomorphic but have the same sequence of Betti numbers, as in the case of the torus and a sphere with two circles glued to it (the 3rd and 4th examples in the figure).



Figure 8: Examples of manifolds and their Betti numbers.

### 6.1.3 Filtrations

The main question that remains is how to determine the simplicial complex of the activity of a neural network. The way to do this is to represent the activity as a point cloud embedded in some metric space. As mentioned previously, we can either consider every neuron as a point in firing rate/spike space (neural manifold) or we can consider the activity in every time bin as a point in some N dimensional space (activity manifold). In either case we will end up with a point cloud, to which we can attach a Vietoris-Rips complex which is constructed by attaching n-dimensional balls of a given radius to each point and creates an k-simplex when the distances between k points are all larger or equal to the specified radius. An example for a random point cloud in the plane is given in figure 9. In this case the radius is the filtration parameter, the fact that this parameter varies across a range of values means *persistent homology* provides information about the point cloud at multiple scales.

Whenever we use point clouds there is an assumption that they are coming from an underlying manifold. To say that these points really sample the manifold in enough detail, the simplicial complexes that we get as a result of the filtration need to be topologically equivalent to the underlying manifold. Of course that is not true for all filtration parameters, but what is required is that the obvious features in the manifold remain stable throughout the filtration. In the range of parameter values which preserve the true topological features, the simplicial complex (constructed from a Čech complex which is well approximated by the Vietoris-Rips complex, see Edelsbrunner and Harer (2010)) is a truthful representation of the topology of the underlying manifold. This is guaranteed by the Nerve theorem (see 4G.3 of Hatcher (2002)), which states that a simplicial complex constructed from the covering of a countable number of contractible open sets is topologically equivalent (homotopy equivalent to be more precise) to the manifold on which the covering is defined.

### Persistent Homology barcode example



Figure 9: Persistent Homology barcode of a random point cloud in the plane. The bottom panel shows the Vietoris-Rips complex for increasing values of the neighborhood radius. Only 0,1 and 2 simplices are shown for better visibility.

### 6.1.4 Persistent Homology

Persistent homology is nothing more than a combination of the last two sections, figure 9 also gives an intuition for what is happening. For any point cloud Xsampled from a manifold  $\mathcal{M}$ , we assign a ball of radius  $\epsilon$  to each point, using the Vietoris-Rips construction - this provides us with a simplicial complex  $V(X, \epsilon)$ . The radius of the balls is chosen from some interval in the real numbers, in theory it can go to infinity so  $\epsilon \in [0, \infty)$  and as result this defines a continuous family of simplicial complexes. This continuity can be seen in the barcode plots, where the start and end of each line reflects the birth and death of a topological feature  $\psi$ . Since all that we are doing is adding more simplices as  $\epsilon$  increases, the simplicial complexes form a family of inclusions in which each abstract simplicial complex is a subset of the ones constructed with a larger radius. In other words, taking the values from the figure as an example,  $V(X, \epsilon = 0) \subset V(X, \epsilon = 0.0681) \subset ... \subset V(X, \epsilon = 0.2723).$ 

Once the family of inclusions of simplicial complexes has been calculated, all that is left to do is calculate the *homology* for each one. Then for each radius value there is a set of Betti numbers describing the number of holes in the given simplicial complex. In practice, the general construction can be summarized in the following diagram, where the subscripts of the epsilon's denote an increasing sequence of  $\epsilon$  values:

Here each  $H(V(X, \epsilon), G)$  denotes the set of all homology groups

 ${H_0(V(X), G), H_1(V(X), G), ..., H_N(V(X), G)}.$ 

for the given epsilon value. Taking the rank of the *homology* group of each order gives a sequence of Betti numbers describing the homological structure at the current epsilon value.

Let's say that as the epsilon value grows at some point a new topological feature (call it  $\psi_i^j$ , *i* indexes the dimension of the feature and *j* - the feature number under some ordering  $j \in J$ ) appears, we call that epsilon value - the birth of the feature  $b(\psi_i^j)$ . On the other hand, if the radius grows too much, topological features start to disappear as every set of points becomes connected. The value of epsilon at which a topological feature disappears is called the death of the feature  $d(\psi_i^j)$ . Then every feature has a birth and a death, the persistence is calculated by subtracting the birth from the death  $p(\psi_i^j) = d(\psi_i^j) - b(\psi_i^j)$ . The persistence of each feature is equal to the length of the lines in figure 9, the birth is where the lines begin and the death where they end.

### 6.2 How to interpret persistence.

The biggest difficulty with what the results are telling us is how to interpret the topological features extracted with *persistent homology*. There are just too many components which might be significant but their meaning remains unintuitive. The most obvious interpretation of a component is that it is a hole of some size, determined by its birth-death values (an example of this is shown in the panel A of figure 10). Under this view, we can imagine the one-holes in a 2 dimensional neural manifold as a punctured flat surface with punctures of different sizes. This intuitive understanding can be generalized to any surface, but the holes can come from other sources. One example is curvature, like in the example of one circle deformed into an ellipse with the middle lines pushed close to each other (panel  $\mathbf{B}$  of figure 10). While this surface is topologically (also homologous) equivalent to a circle, *persistent homology* finds additional structure, which makes it identify the surface as two circles. For this reason we study how we can manipulate the birth-death coordinates of a feature with two homotopic maps f(x), q(x) parameterized by  $t \in [0, 1]$ . This gives the general equation F(x) = tf(x) + (1-t)g(x) which can be used to smoothly vary the birth death coordinates of a topological component. The idea is that when t = 0- q(x) is the only map being calculated and when t = 1 - f(x) is the only map. In between there are combinations of the two functions. If we start from a circle given by a map q(x), we can smoothly deform it into the ellipse squeezed in the middle given by f(x), which lets us inspect the transition between a one hole surface and a two hole surface.



Figure 10: A) shows an example in which the death of a feature is controlled by changing the size of one of the holes. B) shows how a new feature is born during the transition from one circle to a ellipse squeezed in the middle

This approach has the downside that there is an infinite set of maps which we might want to explore. We only use this to illustrate some non trivial reasons for the appearance of topological features. Interesting future research would be to look at how these smaller features develop on top of different surfaces and one way to do that is to analyse the *persistent homology* of stochastic processes on said surfaces. Different stochastic processes could also have a different persistent homological signature due to the way they sample the space, which could also be interesting to explore within the context of neural coding.

# 6.3 Motivation for topology preservation of neural field models on manifolds

We wish to show that the activity of a neural network which realizes unstable bump solutions over a well behaved compact manifold preserves the topology of the manifold. This proof strategy assumes that bumps appear and shrink one at a time. Furthermore it is assumed that every bump on the manifold is sampled with the same frequency.

**Theorem:** For a compact topological manifold  $\mathcal{M}$  with a topology induced by the geodesic distance and a set of functions  $\Phi = \{\phi_{\mu}(x,t) : x, \mu \in \mathcal{M} \& t \in [0,1]\}$  with a topology induced by the L2 distance

$$d(\phi_{\mu}(x,t),\phi_{\mu'}(x,t)) = \left(\int_{\mathcal{M}\times[0,1]} (\phi_{\mu}(x,t) - \phi_{\mu'}(x,t))^2 dx dt\right)^{\frac{1}{2}}.$$

The Vietoris-Rips complexes of  $\mathcal{M}$  and  $\Phi$  are homologous.

**Motivation:** We can take  $\phi_{\mu}(x,t)$  to be piecewise defined scalar fields of the form:

$$\phi_{\mu}(x,t) = \begin{cases} (1-t)e^{-(x-\mu)^2} & \text{if } x \in U\\ 0 & \text{otherwise} \end{cases}$$

Where U denotes a simply connected neighborhood on which the bump is centered (a ball of some radius centered on  $\mu$ ). From now on we will write  $\mathcal{M} \times [0, 1]$ as  $\mathcal{M}$  and the phi's as  $\phi_{\mu}$  for clarity. Then the distance between two identical bumps defined on two non-intersecting neighborhoods  $U \cap V = \emptyset$  can be written as the sum of two integrals on the respective domains, since everywhere else the integral will be 0, the form of the distance is then:

$$d_{\emptyset}(\phi_{\mu},\phi_{\mu'}) = (\int_{U} \phi_{\mu}^{2} dx dt + \int_{V} \phi_{\mu'}^{2} dx dt)^{\frac{1}{2}} = (2 \int_{\mathcal{M}} \phi_{\mu}^{2} dx dt)^{\frac{1}{2}}.$$

When there is overlap in the neighborhoods the distance can be written in a more general form, by also subtracting the contribution coming from the overlap of the two bumps in the following way:

$$\begin{aligned} d(\phi_{\mu},\phi_{\mu'}) &= (\int_{U/V} \phi_{\mu}^{2} dx dt + \int_{U\cap V} (\phi_{\mu} - \phi_{\mu'})^{2} dx dt + \int_{V/U} \phi_{\mu'}^{2} dx dt)^{\frac{1}{2}} = \\ &= (\int_{U/V} \phi_{\mu}^{2} dx dt + \int_{U\cap V} \phi_{\mu}^{2} dx dt - 2 \int_{U\cap V} \phi_{\mu} \phi_{\mu'} dx dt + \int_{U\cap V} \phi_{\mu'}^{2} dx dt + \int_{V/U} \phi_{\mu'}^{2} dx dt)^{\frac{1}{2}} = \\ &= (\int_{U} \phi_{\mu}^{2} dx dt - 2 \int_{U\cap V} \phi_{\mu} \phi_{\mu'} dx dt + \int_{V} \phi_{\mu'}^{2} dx dt)^{\frac{1}{2}} = \\ &= (2 \int_{\mathcal{M}} \phi_{\mu}^{2} dx dt - 2 \int_{U\cap V} \phi_{\mu} \phi_{\mu'} dx dt - 2 \int_{U\cap V} \phi_{\mu} \phi_{\mu'} dx dt)^{\frac{1}{2}}. \end{aligned}$$

From this definition of distance it is clear that because both terms are strictly positive - the larger the overlap between two bumps the smaller the distance, which implies that all local (meaning within the radius of the ball on which the bump is defined) relationships are the same as those of the points on the manifold themselves. Then for uniform and dense enough sampling of  $\mathcal{M}$ , as long as the radius used to construct the Vietoris-Rips complex of  $\Phi$  doesn't exceed the radius of the domain on which bumps are defined, it will have the same topology as the Vietoris-Rips complex of the sampled points under the geodesic distance. While there might be scaling issues since the distances don't have to have the same scale, a normalized *persistent homology* should give the same birth death diagram for point clouds sampled from both the manifold and the set of bumps.

## References

- Adams, H., Emerson, T., Kirby, M., Neville, R., Peterson, C., Shipman, P., Chepushtanova, S., Hanson, E., Motta, F., and Ziegelmeier, L. (2017). Persistence images: A stable vector representation of persistent homology. *The Journal of Machine Learning Research*, 18(1):218–252.
- Babichev, A. and Dabaghian, Y. A. (2018). Topological schemas of memory spaces. Frontiers in Computational Neuroscience, 12:27.
- Babichev, A., Ji, D., Mémoli, F., and Dabaghian, Y. A. (2016). A topological model of the hippocampal cell assembly network. *Frontiers in computational neuroscience*, 10:50.
- Bhattacharya, P. B., Jain, S. K., and Nagpaul, S. (1994). *Basic abstract algebra*. Cambridge University Press.
- Bressloff, P. C. (2011). Spatiotemporal dynamics of continuum neural fields. Journal of Physics A: Mathematical and Theoretical, 45(3):033001.
- Bressloff, P. C. (2019). Stochastic neural field theory of wandering bumps on a sphere. *Physica D: Nonlinear Phenomena*, 399:138–152.
- Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of computational neuroscience*, 8(3):183–208.
- Bubenik, P. (2015). Statistical topological data analysis using persistence landscapes. The Journal of Machine Learning Research, 16(1):77–102.
- Bubenik, P., Hull, M., Patel, D., and Whittle, B. (2020). Persistent homology detects curvature. *Inverse Problems*, 36(2):025008.
- Carlsson, G. (2009). Topology and data. Bulletin of the American Mathematical Society, 46(2):255–308.
- Carlsson, G., Ishkhanov, T., De Silva, V., and Zomorodian, A. (2008). On the local behavior of spaces of natural images. *International journal of computer* vision, 76(1):1–12.
- Carroll, S. M. (2019). Spacetime and geometry. Cambridge University Press.
- Chaudhuri, R., Gercek, B., Pandey, B., Peyrache, A., and Fiete, I. (2019). The intrinsic attractor manifold and population dynamics of a canonical cognitive circuit across waking and sleep. *Nature neuroscience*, 22(9):1512–1520.

- Coombes, S., beim Graben, P., Potthast, R., and Wright, J. (2014). *Neural fields: theory and applications.* Springer.
- Curto, C. (2017). What can topology tell us about the neural code? Bulletin of the American Mathematical Society, 54(1):63–78.
- Curto, C., Gross, E., Jeffries, J., Morrison, K., Omar, M., Rosen, Z., Shiu, A., and Youngs, N. (2017). What makes a neural code convex? SIAM Journal on Applied Algebra and Geometry, 1(1):222–238.
- Curto, C. and Itskov, V. (2008). Cell groups reveal structure of stimulus space. PLoS Comput Biol, 4(10):e1000205.
- Curto, C. P. and Youngs, N. (2020). Neural ring homomorphisms and maps between neural codes. In *Topological Data Analysis*, pages 163–180. Springer.
- Dabaghian, Y., Mémoli, F., Frank, L., and Carlsson, G. (2012). A topological paradigm for hippocampal spatial map formation using persistent homology. *PLoS Comput Biol*, 8(8):e1002581.
- de Vries, S. E., Lecoq, J. A., Buice, M. A., Groblewski, P. A., Ocker, G. K., Oliver, M., Feng, D., Cain, N., Ledochowitsch, P., Millman, D., et al. (2020). A large-scale standardized physiological survey reveals functional organization of the mouse visual cortex. *Nature Neuroscience*, 23(1):138–151.
- Edelsbrunner, H. and Harer, J. (2010). *Computational topology: an introduction*. American Mathematical Soc.
- Fefferman, C., Mitter, S., and Narayanan, H. (2016). Testing the manifold hypothesis. *Journal of the American Mathematical Society*, 29(4):983–1049.
- Fusi, S., Miller, E. K., and Rigotti, M. (2016). Why neurons mix: high dimensionality for higher cognition. *Current opinion in neurobiology*, 37:66–74.
- Gallego, J. A., Perich, M. G., Miller, L. E., and Solla, S. A. (2017). Neural manifolds for the control of movement. *Neuron*, 94(5):978–984.
- Gao, P. and Ganguli, S. (2015). On simplicity and complexity in the brave new world of large-scale neuroscience. *Current opinion in neurobiology*, 32:148– 155.
- Gerstner, W., Kistler, W. M., Naud, R., and Paninski, L. (2014). Neuronal dynamics: From single neurons to networks and models of cognition. Cambridge University Press.
- Ghrist, R. (2008). Barcodes: the persistent topology of data. Bulletin of the American Mathematical Society, 45(1):61–75.
- Giusti, C., Ghrist, R., and Bassett, D. S. (2016). Two's company, three (or more) is a simplex. *Journal of computational neuroscience*, 41(1):1–14.

- Giusti, C., Pastalkova, E., Curto, C., and Itskov, V. (2015). Clique topology reveals intrinsic geometric structure in neural correlations. *Proceedings of the National Academy of Sciences*, 112(44):13455–13460.
- Guanella, A., Kiper, D., and Verschure, P. (2007). A model of grid cells based on a twisted torus topology. *International journal of neural systems*, 17(04):231– 240.
- Hatcher, A. (2002). Algebraic Topology. Cambridge University Press.
- Hofer, C., Kwitt, R., Niethammer, M., and Uhl, A. (2017). Deep learning with topological signatures. In Advances in Neural Information Processing Systems, pages 1634–1644.
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of* sciences, 79(8):2554–2558.
- Kreyszig, E. (1991). Differential geometry.
- Kriegeskorte, N., Mur, M., and Bandettini, P. A. (2008). Representational similarity analysis-connecting the branches of systems neuroscience. Frontiers in systems neuroscience, 2:4.
- Levi, R. A short course on algebraic topology geared towards applications in neuroscience.
- Martin, R., Chappell, D., Chuzhanova, N., and Crofts, J. J. (2018). A numerical simulation of neural fields on curved geometries. *Journal of computational neuroscience*, 45(2):133–145.
- Mastrogiuseppe, F. and Ostojic, S. (2018). Linking connectivity, dynamics, and computations in low-rank recurrent neural networks. *Neuron*, 99(3):609–623.
- McInnes, L., Healy, J., and Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426.
- Moser, E. I., Kropff, E., and Moser, M.-B. (2008). Place cells, grid cells, and the brain's spatial representation system. *Annu. Rev. Neurosci.*, 31:69–89.
- Munkres, J. R. (2018). Elements of algebraic topology. CRC Press.
- Naitzat, G., Zhitnikov, A., and Lim, L.-H. (2020). Topology of deep neural networks. arXiv preprint arXiv:2004.06093.
- Redish, A. D., Elga, A. N., and Touretzky, D. S. (1996). A coupled attractor model of the rodent head direction system. *Network: Computation in Neural* Systems, 7(4):671–685.

- Reimann, M. W., Nolte, M., Scolamiero, M., Turner, K., Perin, R., Chindemi, G., Dłotko, P., Levi, R., Hess, K., and Markram, H. (2017). Cliques of neurons bound into cavities provide a missing link between structure and function. *Frontiers in computational neuroscience*, 11:48.
- Saxena, S. and Cunningham, J. P. (2019). Towards the neural population doctrine. Current opinion in neurobiology, 55:103–111.
- Sengupta, A., Pehlevan, C., Tepper, M., Genkin, A., and Chklovskii, D. (2018). Manifold-tiling localized receptive fields are optimal in similarity-preserving neural networks. In Advances in Neural Information Processing Systems, pages 7080–7090.
- Shepard, R. N. and Chipman, S. (1970). Second-order isomorphism of internal representations: Shapes of states. *Cognitive psychology*, 1(1):1–17.
- Siegle, J. H., Jia, X., Durand, S., Gale, S., Bennett, C., Graddis, N., Heller, G., Ramirez, T. K., Choi, H., Luviano, J. A., et al. (2019). A survey of spiking activity reveals a functional hierarchy of mouse corticothalamic visual areas. *Biorxiv*, page 805010. https://allensdk.readthedocs.io/en/latest/visual\_coding\_neuropixels.html.
- Singh, G., Memoli, F., Ishkhanov, T., Sapiro, G., Carlsson, G., and Ringach, D. L. (2008). Topological analysis of population activity in visual cortex. *Journal of vision*, 8(8):11–11.
- Sizemore, A. E., Phillips-Cremins, J. E., Ghrist, R., and Bassett, D. S. (2019). The importance of the whole: topological data analysis for the network neuroscientist. *Network Neuroscience*, 3(3):656–673.
- Stringer, C., Pachitariu, M., Steinmetz, N., Carandini, M., and Harris, K. D. (2019). High-dimensional geometry of population responses in visual cortex. *Nature*, 571(7765):361–365.
- Tralie, C., Saul, N., and Bar-On, R. (2018). Ripser.py: A lean persistent homology library for python. The Journal of Open Source Software, 3(29):925.
- Tu, L. W. (2011). An introduction to manifolds. Springer.
- Wasserman, L. (2013). All of statistics: a concise course in statistical inference. Springer Science & Business Media.
- Wasserman, L. (2018). Topological data analysis. Annual Review of Statistics and Its Application, 5:501–532.
- Wilson, H. R. (1999). Spikes, decisions, and actions: the dynamical foundations of neurosciences. pages 103–112.
- Wilson, H. R. and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical journal*, 12(1):1–24.
- Zomorodian, A. (2010). Fast construction of the vietoris-rips complex. Computers & Graphics, 34(3):263–271.