*Annual Review of Neuroscience*

# Topological Neuroscience: Linking Circuits to Function

Carina Curto and Nicole Sanderson

Division of Applied Mathematics and Carney Institute for Brain Science, Brown University, Providence, Rhode Island, USA; email: carina_curto@brown.edu



Annu. Rev. Neurosci. 2025. 48:491–518

First published as a Review in Advance on April 15, 2025

The *Annual Review of Neuroscience* is online at neuro.annualreviews.org

https://doi.org/10.1146/annurev-neuro-112723-034315



## Keywords

topology, torus, grid cells, graph, simplicial complex, persistent homology, head direction cells, neural circuits

## Abstract

We review recent developments of the use of topology in neuroscience. From grid cells and head direction cells to the geometry of olfactory space, modern applied topology methods such as persistent homology are increasingly being used to study neural circuits and perception. In addition to outlining the big picture and reviewing various applications of topological data analysis (TDA) to neuroscience, we take a deep dive into the basic homology computation to make the underlying mathematics more accessible to neuroscientists. A discussion of practical considerations and pointers to TDA software are also included.

## Contents





# INTRODUCTION

## A Torus in the Rat's Brain

In February 2021, a paper was posted on *bioRxiv* titled "Toroidal Topology of Population Activity in Grid Cells" (Gardner et al. 2021). It was a tour de force collaboration between a Nobel Prize–winning lab, a couple of computational neuroscientists, and three mathematicians. The results were spectacular, if not entirely unexpected. Using Neuropixels silicon probes, the Moser lab had recorded more than 7,500 single units in the medial entorhinal cortex of freely moving rats. The recordings were dense enough to identify six grid cell modules across four recording sessions. And once they limited their analyses to "pure" grid cells, clustered by module, they were able to see what many of us were expecting them to see: a torus in the rat's brain. Or, to be more precise, several tori—one per grid cell module.

What does this mean? Here's the basic idea. Grid cells are neurons that encode an animal's position in space via overlapping grid fields. These grid fields are similar to place fields for hippocampal place cells with one important difference: they each have multiple peaks that are organized in a hexagonal grid. Two grid cells belong to the same *module* if their grid fields have the same scaling and orientation (see **Figure 1**). Within each module, grid fields differ by their spatial phase. One grid field has the peak in the center of each hexagon, another in the upper right corner, and so on, and collectively they fully cover the animal's environment. In this way, the grid fields within

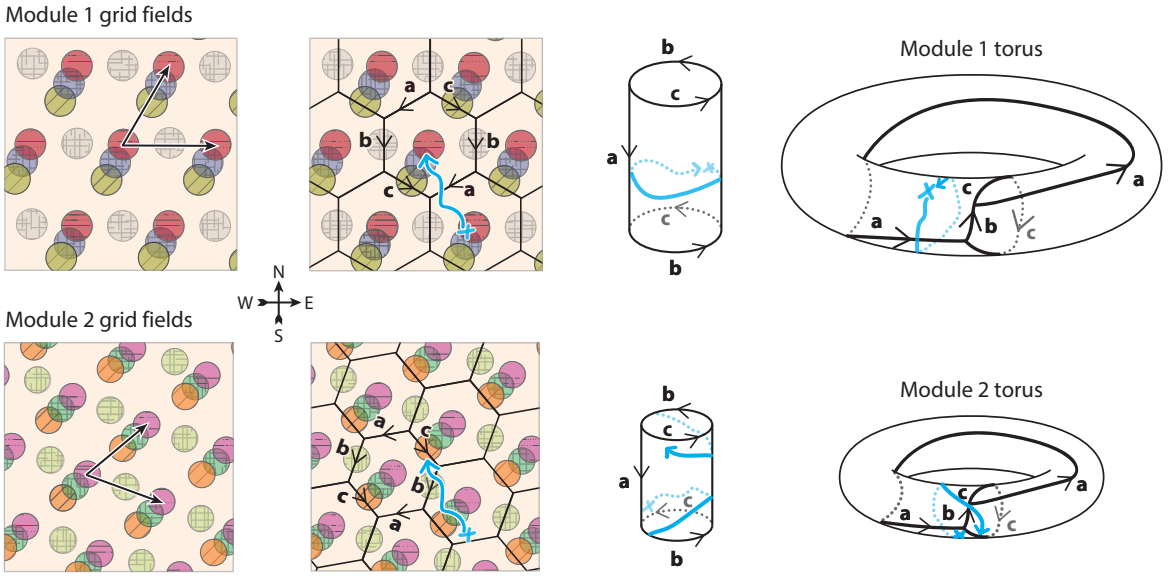



**Figure 1**

Grid fields from a single grid cell module. (*Top*, *left*) Grid fields from a single module have the same scaling and orientation. Grid fields for four neurons are shown (one per color); each field has the periodic pattern of a hexagonal lattice. The grid field pattern repeats in a hexagonal tiling of space. A single hexagon forms a "fundamental domain" represented by the grid cell activity. (*Top*, *right*) Identifying one pair of opposite sides of the hexagon yields a cylinder, and identifying the other sides results in a twisted torus. An example trajectory (*blue*) that crosses from one tile to another results in a repetition of the same grid cell activity at the beginning and end. (*Bottom*) Same as above, but for a module with a different orientation and a smaller scaling of the grid fields, resulting in a smaller torus. The same trajectory that looped back to a similar position on the module 1 torus does not show repeating population activity in module 2.

a module correspond to the same hexagonal tiling of space. [These details are explained in many places by now; a good place to start is Edvard Moser's Nobel Prize lecture (Moser 2014).] But the position information from the grid cell activity of neurons in a single module only nails down the animal's location relative to the hexagon.[1] In other words, it can tell you that the animal is located in the lower left corner of a hexagonal tile, but not which tile the animal is in! Moreover, because of the repeating arrangement of hexagonal tiles, positions on opposite sides along the boundary of the hexagon get identified—they give rise to the same grid cell activity. The resulting shape, topologically, is a twisted torus. You can verify this yourself: imagine a large rubbery sheet in the form of a hexagon, and label the sides clockwise as a, b, c, a, b, c (see **Figure 1**). Glue the top and bottom "a" sides together, and you get a cylinder where each boundary circle consists of b and c sides. Now twist the cylinder, as if you are opening a can of biscuits, and glue the two ends together so that you match b to b and c to c. Voilà: this is your torus.

But this torus is formed from a hexagonal tile in two-dimensional space. What does it mean to see the torus "in the brain," arising from the population activity of grid cells? If you collect the activity of $n$ grid cells from a single module into population vectors that evolve over time, they form trajectories in an $n$-dimensional Euclidean space. However, as the animal moves through space from one tile to another, the same (or very similar) population vectors can get repeated, as nearby positions on the hexagon are repeatedly visited, albeit on far-away tiles. The trajectories

[1]Note that a parallelogram can also be used as the fundamental domain; in this case, the torus emerges when opposite sides are identified. See Langdon et al. (2023) for a nice visualization.

thus loop back on themselves, and the population activity ends up taking the shape of points sampled from a two-dimensional torus embedded in the $n$-dimensional activity space.

This idea, that data from a covering of space by local "fields" can reflect the underlying topology of that space, is the neuroscience instantiation of the famous *Nerve theorem* from algebraic topology (Curto et al. 2013, Schenck 2022). In the case of place cells in the hippocampus, the represented space is the animal's environment (Curto & Itskov 2008, Curto 2017). In the case of a module of grid cells, the represented space is a hexagonal tile with opposite sides identified—topologically a torus (Curto 2017). These are the tori that were observed by Gardner et al. (2021, 2022) using an important method of topological data analysis (TDA) called *persistent homology*.

Roughly speaking, persistent homology is a data analysis method that was designed to analyze point cloud data (e.g., a bunch of points in $\mathbb{R}^n$) by estimating topological features of the underlying manifold from which the points appear to have been sampled (Niyogi et al. 2008, Kang et al. 2021). In the case of grid cell activity from a single module, the population vectors trace out points on a torus in $\mathbb{R}^n$, where $n$ is the number of grid cells in the population. In the case of population activity representing a circular variable, such as heading direction, we expect the population vectors that appear to be sampled from a circle in $\mathbb{R}^n$ (Kim et al. 2017, Chaudhuri et al. 2019, Petrucco et al. 2023). But this is just one of the ways topology appears in neuroscience.

### Roadmap

The rest of this article is organized as follows. First, we give an overview of some important examples where topology has played a role in studying neural circuits and the structure of neural coding. Back in 2017, when topology was still virtually unknown in neuroscience, the first author attempted to give a fairly comprehensive review of all the work that had been done in this area (Curto 2017). Eight years later, a comprehensive review is impossible. Nevertheless, we flesh out some examples that reflect the variety of ways in which these techniques have been used and have been influential in spurring future research.

Next, we take a bit of a deep dive into the core mathematical ideas underlying algebraic topology, persistent homology, and so on and explain the basic homology computation by working out a detailed example. The philosophy behind including this example comes from the belief that many neuroscientists (and perhaps all computational neuroscientists) already have the necessary background in linear algebra to be able to understand the essence of what homology is computing, and that having a more concrete understanding of the algebraic topology will enable better and more honest interpretation of topological analyses. We then give an intuitive explanation of how persistent homology works and describe the various outputs of persistent homology software: barcodes, persistence diagrams, and Betti curves.

Finally, we come back to neuroscience applications and explain how barcodes and Betti curves have been used in practice to detect topological and geometric structure in the neural activity of grid cells, head direction cells, place cells, and in olfactory space.

## TOPOLOGY IN NEUROSCIENCE: WHAT, WHEN, WHERE, AND WHY?

In a soundbite, topology is the study of shape without angles or distances—a kind of geometry that is agnostic to the choice of metric. The mathematical field of topology makes these ideas rigorous and is still an active area of research motivated by many open questions. Most topologists use tools from *algebraic topology*, a theory that defines and studies topological invariants of various spaces using techniques from homological algebra. If you have heard of Betti numbers, homology groups, homotopy groups, and the fundamental group—these are all basic invariants that are computed in algebraic topology (Munkres 2000, Hatcher 2002).

TDA can be thought of as an applied math spin-off from algebraic topology. At its origin, the field was motivated by two main questions: (*a*) how can we make large-scale homology computations on a computer more efficient, and (*b*) how do we associate topological features to point cloud data? In other words, how can we detect whether a set of points, such as population vectors in $\mathbb{R}^n$, have an underlying shape—as if they had been sampled from a torus or a sphere? These questions were not at all addressed in traditional algebraic topology, but several topologists got to work developing the new theory underlying TDA. This work began more than 25 years ago (Robins et al. 1998; Robins 1999, 2000, 2002; Edelsbrunner et al. 2002; Kacynski et al. 2004; Zomorodian & Carlsson 2005; Edelsbrunner & Harer 2009; Zomorodian 2009), and much progress has been made (Edelsbrunner & Morozov 2013, Ghrist 2014, Turner et al. 2014, Zomorodian 2014, Otter et al. 2017, Boissonnat et al. 2018, Kanari et al. 2020, Carlsson & Vejdemo-Johansson 2022, Dey & Wang 2022, Schenck 2022, Kahle et al. 2023, Curry et al. 2024). The tools of TDA are now well-developed and readily available for the use of any computational scientist (see the Software section).

Within neuroscience, the first uses of these modern topological methods appeared a little over 15 years ago. One of the very first papers using these techniques involved the collaboration of a neuroscientist, Dario Ringach, and a topologist, Gunnar Carlsson—one of the original founders of TDA. Their paper, "Topological Analysis of Population Activity in Visual Cortex" (Singh et al. 2008), used persistent homology to detect interesting topological structure in both spontaneous and stimulus-evoked activity in primary visual cortex (V1). That same year, the first author, together with Vladimir Itskov, published a paper showing that the shape of an animal's environment could be inferred via topological analysis of hippocampal place cell activity—without knowledge of the actual place fields (Curto & Itskov 2008). Very similar observations were being made at the time by Yuri Dabaghian (Dabaghian et al. 2012, 2014) in the lab of Loren Frank. This early work on applications of topology in neuroscience attracted considerable interest from within the math community and was reviewed by Ghrist (2007), Curto (2017), and Levi (2017). On the neuroscience side, however, these ideas were initially somewhat slow to catch on, until recently.

There are at least three ways in which ideas and structures from topology may be useful for understanding the brain. These are summarized in **Figure 2*a***. First, topological methods can be used to understand and classify features of the *physical brain*. This can include neuron morphology (Kanari et al. 2017, 2019, 2022; Li et al. 2017), the physical structure of neural networks (Kim et al. 2017), or even the structure of nonneuronal networks such as the brain's vasculature controlling blood flow (Bendich et al. 2016, Haft-Javaherian et al. 2020, Yao et al. 2024).

Second, *neural activity* may have interesting topology, as we saw in the case of grid cells. Such structure has also been observed in systems where the underlying circuits are believed to have a ring-like structure (Kim et al. 2017, Chaudhuri et al. 2019, Petrucco et al. 2023). In addition to position coding and vision systems, topological methods have also been used to analyze spike train coactivity in songbird auditory cortex (Theilman et al. 2021), to study auditory coding more generally (Reyes 2021, Tudoras & Reyes 2021), and to study the flow of neural activity through a cortical circuit (Riemann et al. 2017, Bardin et al. 2019). Furthermore, topological methods have been used fruitfully to study the structure of neural activity in human functional MRI (fMRI) (Giusti et al. 2016; Sizemore et al. 2016, 2018; Anderson et al. 2018; Catanzaro et al. 2024), and topological invariants for fMRI have been shown to have diagnostic value (Stolz et al. 2021).

Third, the *represented stimulus spaces* may have interesting topology. For example, the head direction system represents a circular variable tracking head direction. Similarly, orientation-tuned neurons in V1 represent angles on a circle. We would thus expect population activity vectors in either of these systems to appear as if they were sampled from a circle in $\mathbb{R}^n$. But the structure of stimulus spaces can also be inferred directly, using perceptual measures (Waraich & Victor 2024)

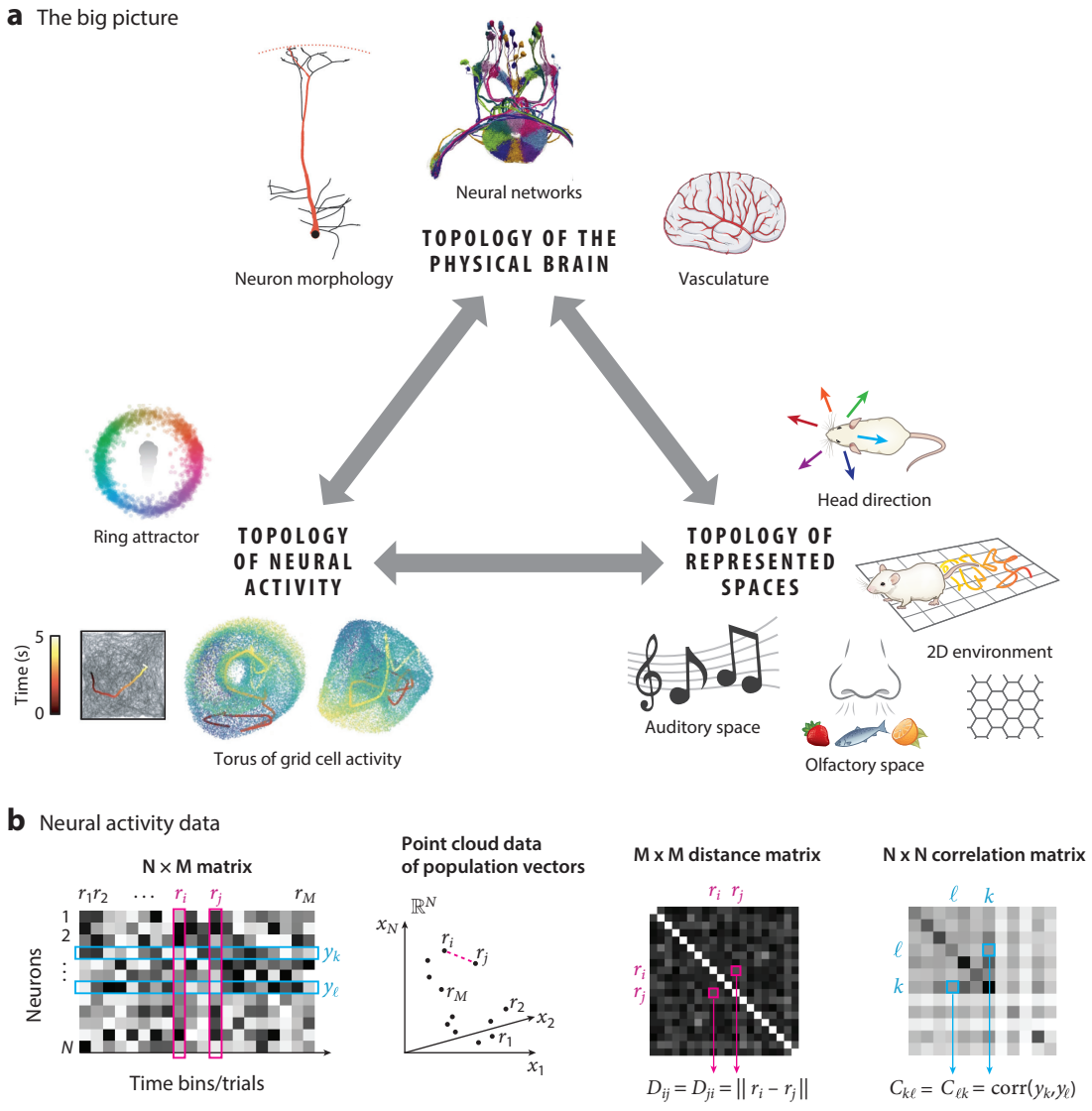



**Figure 2**

The big picture. (*a*) Three ways in which topology can provide insights into neuroscience. (*b*) Neural activity data are amenable to topological analyses in two ways: from the columns, one gets a distance matrix; from the rows, a correlation matrix. Both are symmetric matrices that can be analyzed using persistent homology. Several images in panel *a* adapted from Kanari et al. (2017) (CC BY 4.0), Noorman et al. (2024) (CC BY 4.0), Petrucco et al. (2023) (CC BY 4.0), Gardner et al. (2022) (CC BY 4.0), and Servier Medical Art (**https://smart.servier.com/smart_image/brain-circulation/**) (CC BY 4.0).

or other measures of stimulus distance/similarity (Zhou et al. 2018, Chen et al. 2024), and the underlying geometry of such spaces may be hyperbolic rather than Euclidean (Zhou et al. 2018, Sharpee 2019, Zhang et al. 2023, Waraich & Victor 2024).

Going back to the case of neural activity, there are two important (and in some sense, dual) perspectives that can be used for analyzing neural activity through a topological lens. They have to do with focusing on the columns versus rows of a neural activity matrix (**Figure 2*b***, left). Consider

an $N \times M$ matrix, where $N$ is the number of neurons and $M$ is the number of time bins (or stimuli). Each column of this matrix, $r_i$, represents a population vector, and the collection of population vectors $r_1, \ldots, r_M$ is a set of point cloud data in $\mathbb{R}^N$ (**Figure 2*b***, middle). From these data one can then compute pairwise distances, yielding a distance matrix with entries $D_{ij} = \|r_i - r_j\|$. Such a distance matrix can then be the input to a persistent homology computation that is used to analyze the shape of the point cloud. This is the perspective that has been used to see a torus in the population activity of a grid cell module, albeit with some important preprocessing steps (Gardner et al. 2022). It is also the perspective underlying the earlier V1 analysis by Singh et al. (2008).

Instead of focusing on the columns, however, one can also look at the rows of a neural activity matrix. Each row, $y_k$, collects the response of neuron $k$ across time and/or stimulus presentations. One may use these row vectors to compute an $N \times N$ correlation matrix (**Figure 2*b***, right), whose entries $C_{k\ell}$ represent the correlation between neurons $k$ and $\ell$. Although such a matrix is not a distance matrix, it turns out that it, too, can be taken as the input to a persistent homology computation. In fact, TDA on correlation or similarity matrices has proven to be quite fruitful, yielding interesting results about the geometry of hippocampal place cell coding (Giusti et al. 2015, Zhang et al. 2023) and olfactory space (Zhou et al. 2018). The duality suggested by looking at information in the columns versus rows of the neural activity matrix echoes a deep duality in topology known as Dowker's theorem (Dowker 1952; Chowdhury & Mémoli 2018; Wu & Itskov 2022; Vaupel & Dunn 2023; Yoon et al. 2023, 2024).

Finally, independent of the topological interpretation, persistent homology also provides a novel method for generating matrix invariants that may be more appropriate than traditional spectral techniques for neural data analysis. More on that later.

## TOPOLOGY OF GRAPHS AND SIMPLICIAL COMPLEXES

In all of the above examples, from grid cells to bird song, the basic topological tools that were used involve persistent homology and the underlying combinatorial objects: graphs, cliques, and simplicial complexes. In this section, we take a deep dive into how these objects come together in the basic homology computation. Providing mathematical details requires some standard notation and terminology that may be unfamiliar to neuroscientists; for completeness, we include a table of notation and a glossary of terminology at the end of this article (see the Appendix).

Before we get into specific topological invariants, like homology groups, it is important to understand what topologists mean when they say that two spaces are topologically "the same." There are some technical definitions here, but we illustrate what they mean with some easy examples. Strictly speaking, we can only say that two spaces are topologically equivalent if they are *homeomorphic*. That is, $X \cong Y$ if there is a continuous map $f: X \to Y$ that is a bijection, and for which $f^{-1}$ is also continuous. Such a map is called a *homeomorphism*. This definition captures that idea that two spaces should be considered topologically equivalent if you can deform one into the other by stretching or shrinking different regions of the space in a continuous and invertible manner, without closing or creating any holes. It is the strongest sense of topological equivalence.

On the other hand, there is also the weaker notion of *homotopy equivalence*. Two topological spaces $X$ and $Y$ are *homotopy equivalent* if there exist continuous maps $f: X \to Y$ and $g: Y \to X$ such that $g \circ f$ is *homotopic* to the identity map $\mathrm{id}_X$ and $f \circ g$ is homotopic to the identity map $\mathrm{id}_Y$ (see the Appendix). In other words, you can continuously deform $X$ into $Y$ by operations that are allowed to expand or contract parts of the space without creating (or destroying) any new connected components or holes, but the map $f$ need not be invertible. For example, consider the cylinder and the circle in **Figure 3*a***. There is a noninvertible projection map that sends each point on the cylinder to the corresponding point below it on the bottom circle. If the cylinder is $X = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 = 1, z \in [0, 2]\}$ and the circle is $Y = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$, then a

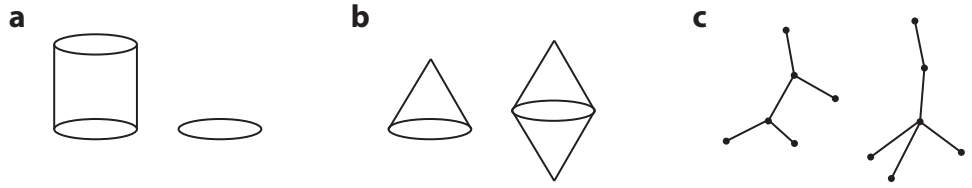



**Figure 3**

Homotopy equivalent versus homeomorphic. (*a*) A cylinder and a circle are homotopy equivalent spaces, but they are not homeomorphic. (*b*) A cone is homotopy equivalent to a point, while gluing two cones together along their circular boundary creates a space homeomorphic to a sphere. (*c*) All trees are homotopy equivalent to each other. However, the tree on the left is not homeomorphic to the tree on the right (e.g., by removing a single point one can disconnect the tree on the right into four components; this is not possible for the tree on the left). All the standard invariants from algebraic topology, such as homology groups, cannot detect the difference between homotopy equivalent spaces that are not homeomorphic.

homotopy equivalence is given by the functions $f$ and $g$ where $f : X \to Y$ maps $(x, y, z) \mapsto (x, y)$, and $g : Y \to X$ maps $(x, y) \mapsto (x, y, 0)$. Notice that $f \circ g : Y \to Y$ sends $(x, y) \mapsto (x, y)$, so $f \circ g = \mathrm{id}_Y$; in particular, $f \circ g$ is homotopic to $\mathrm{id}_Y$. On the other hand, $g \circ f : X \to X$ sends $(x, y, z) \mapsto (x, y, 0)$ and is thus a noninvertible projection not equal to $\mathrm{id}_X$. However, the family of continuous maps $h_t : X \to X$, where $(x, y, z) \mapsto (x, y, z - tz)$, connects $h_1 = g \circ f$ to $h_0 = \mathrm{id}_X$ in a continuous manner, showing that $g \circ f$ is homotopic to $\mathrm{id}_X$ (see the Appendix). We can conclude that the cylinder and the circle are homotopy equivalent. They are not, however, homeomorphic.

Homotopy equivalence is an important notion in topology because the vast majority of common topological invariants, such as Betti numbers and homology groups, are invariant under homotopy equivalence. These invariants cannot tell apart $X$ and $Y$ if they are homotopy equivalent, even if they have different dimensions! In particular, they cannot distinguish between the cylinder and the circle in **Figure 3*a***. On the other hand, they can distinguish between the cone and the double-cone in **Figure 3*b***, as they are not homotopy equivalent (the double-cone has a two-dimensional cavity that the cone does not).

## Do Graphs Have Interesting Topology Even if They Are Trees?

A *tree* is a graph with no cycles, and it is common to hear people say "trees have trivial topology." What is meant by this is that a tree is homotopy equivalent to a single point (also referred to as *contractible*). This means that traditional homological invariants cannot distinguish between any two trees (**Figure 3*c***).

But trees are not actually topologically equivalent to points or balls. Thinking of a tree as a one-dimensional space, it is clear that the removal of a single point can disconnect the tree into two or more components. This is not true for a point or ball. Moreover, pairs of trees are seldom homeomorphic to each other. For example, the two trees in **Figure 3*c*** are not homeomorphic: the one on the right can be disconnected into four components by removing a single point (the vertex of degree 4). This is not possible for the tree on the left, an indication that there can be no homeomorphism between the two trees. On the other hand, two trees with the same combinatorial pattern of vertices and edges are topologically identical (i.e., homeomorphic), no matter how long the edges are or how the vertices are positioned in space.

While traditional homological invariants are too weak to distinguish between trees (they are all homotopy equivalent), with the advent of persistent homology there are additional topological invariants to consider. And these invariants do capture interesting structure about trees and how they are embedded in space (Turner et al. 2014). This makes trees surprisingly interesting

to analyze using modern topological methods. Such trees may reflect the structure of individual neurons, temporal branching processes, or vasculature in the brain (Kanari et al. 2017, 2019; Haft-Javaherian et al. 2020; Yao et al. 2024). On the neural coding side, trees provide models of hierarchical structures from spatial and sensory representations (Zhang et al. 2023, Waraich & Victor 2024).

The more salient aspects of topology, however, emerge when considering spaces that are more complex than graphs or trees. Next, we discuss simplicial complexes and segue into the basic homology computation underlying the standard homological invariants, as well as persistent homology.

## Simplicial Complexes

A simplicial complex $\Delta$ is a generalization of a graph, composed of simplices. Depending on the dimension, a simplex may be a vertex, an edge, a triangle, a tetrahedron, or a higher-dimensional analog. Geometrically, one may picture a $d$-dimensional simplex as the convex hull of $d+1$ points in general position in a Euclidean space, $\mathbb{R}^d$. Combinatorially, simplices are represented by single vertices, pairs of vertices, triples of vertices, quadruples of vertices, and so on (see **Figure 4*a***). In the same way, an undirected graph $G$ is typically given as a collection of vertices and pairs of vertices (the edges). The defining property that makes a simplicial complex special (as compared to, say, a hypergraph) is that the collection of simplices must be closed under subsets (see the Appendix). In other words, if $\sigma \in \Delta$ and $\tau \subseteq \sigma$, then $\tau \in \Delta$.

Simplicial complexes can also be viewed as topological spaces and are the basis for what is arguably the simplest theory of homological invariants: *simplicial homology*. If one can model a topological space via a simplicial complex, the homology computations become a straightforward matter of linear algebra. (If you have ever seen a picture of the triangulation of a surface, this is precisely a model of the surface as a two-dimensional simplicial complex whose simplices consist of vertices, edges, and triangles.) In the neuroscience setting, there are various ways to turn data

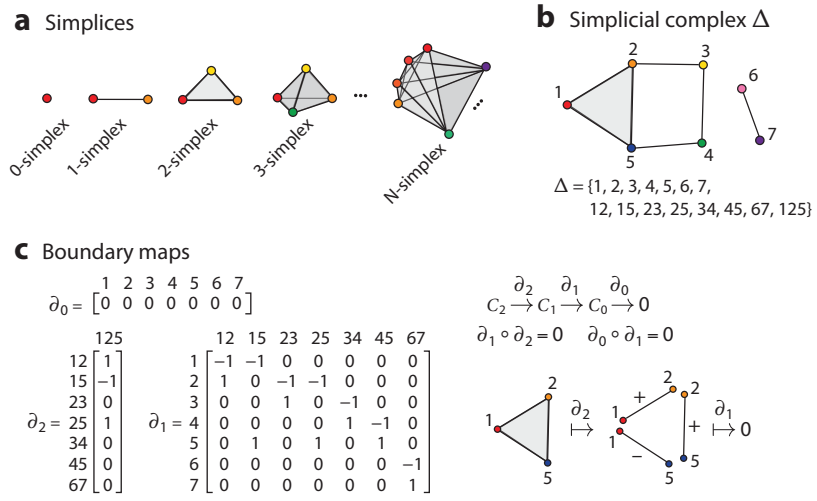



**Figure 4**

The basic homology computation. (*a*) Simplices in various dimensions form the building blocks for simplicial complexes. (*b*) A simplicial complex composed of one 2-simplex, seven 1-simplices, and seven 0-simplices. (*c*) Boundary maps $\partial_k$ connect linear combinations of simplices (chains) in dimension $k$ to their boundaries in dimension $k-1$.

into simplicial complexes. Finding a good simplicial complex, or a good family of them, is the first step toward computing meaningful topological invariants.

More on that later. But first, what exactly is homology? And how does one compute it for a simplicial complex?

## The Basic Homology Computation, in the Simplest Setting

You may have heard the idea that topology is about studying topological features such as "holes" in a geometric object, while ignoring angles, distances, and other geometric details. If $X$ is a topological space, the 0th homology group $H_0(X)$ captures the connected components of $X$; the 1st homology group $H_1(X)$ encodes one-dimensional holes, like an empty region bounded by a circle; the 2nd homology group $H_2(X)$ encodes two-dimensional holes, like the hollow inside of a sphere; and so on. However, these "holes" are represented carefully, identifying homological cycles that are considered equivalent. This is the part that does not feel very transparent. The best way to demystify what exactly is being computed by simplicial homology is to do an example. The only prerequisite needed to follow these computations is basic linear algebra.

What follows is a bit of a deep dive into an explicit example of a homology computation, writing down the boundary maps and all. The reader who already knows basic algebraic topology can skip it, as can the reader who is not yet interested in all the mathematical details.

So, let us work through an example in detail. Consider the simplicial complex $\Delta$ given in **Figure 4*b***. It consists of seven vertices, seven edges, and one triangle. Combinatorially, it is the following collection of subsets of the vertex set $\{1, \ldots, 7\}$:

$$\Delta = \{1, 2, 3, 4, 5, 6, 7, 12, 15, 23, 25, 34, 45, 67, 125\},$$

where we are slightly abusing notation and writing things like "45" and "125" to denote the subsets $\{4, 5\}$ and $\{1, 2, 5\}$. Note that because $\Delta$ is closed under subsets, it can also be characterized by its maximal simplices: 23, 34, 45, 67, and 125. Adding in all subsets of these simplices, such as the edges 12, 15, and 25, as well as the vertices $1, \ldots, 7$, recovers the full simplicial complex $\Delta$.

In order to compute the homology groups of $\Delta$, the first thing we must do is create vector spaces $C_0$, $C_1$, and $C_2$ that encode linear combinations of simplices in each dimension. What we are doing here is very simple: to each simplex $\sigma \in \Delta$, we associate a basis vector $e_\sigma$, and then group these into different vector spaces according to the simplex dimension. The elements of the vector space $C_k$ are thus linear combinations of the $k$-simplices, with coefficients in a prescribed field $\mathbf{k}$.[2] You can think of $\mathbf{k} = \mathbb{R}$ for purposes of this example.[3] (We avoid using rings like $\mathbb{Z}$ for the coefficients so that the $C_k$ are vector spaces.) Specifically,

$$C_0 = \left\{ \sum_{i=1}^{7} c_i e_i \mid c_i \in \mathbf{k} \right\},$$

$$C_1 = \left\{ \sum_{\sigma} c_\sigma e_\sigma \mid c_\sigma \in \mathbf{k} \text{ and } \sigma \in \{12, 15, 23, 25, 34, 45, 67\} \right\},$$

$$C_2 = \{ c_{125} e_{125} \mid c_{125} \in \mathbf{k} \}.$$

[2] Unfortunately, but unavoidably, we have used the word "field" in three different ways so far. In neuroscience, there is the "field" of receptive fields, as in grid fields and place fields. In mathematics, a number system like $\mathbb{R}$ with additive and multiplicative inverses is called a field. And, finally, there is "field" as in an area of scientific research. We have not discussed corn fields, though.

[3] In software, the tendency is to use finite fields such as $\mathbb{Z}/p\mathbb{Z}$. However, this can yield a subtle phenomenon called *torsion* that does not happen over $\mathbb{R}$ and which is beyond the scope of this review.

The vectors in each of these vector spaces are called *chains*: 0-chains in $C_0$, 1-chains in $C_1$, and 2-chains in $C_2$. The coefficients, $c_\sigma$, are simply numbers in the field $\mathbf{k}$, just as with any other vector space over $\mathbf{k}$.

Another important convention concerns the *orientations* of the simplices. Each simplex, represented by $e_\sigma$, has two possible orientations: positive or negative. We will assume the positive orientation is the one where the vertices are given in increasing numerical order, or with any *even permutation* of this order (see the Appendix). Thus, $e_{125} = e_{251} = e_{512}$ all represent the positively oriented 125 simplex, while $e_{215} = e_{152} = e_{521}$ represent the negatively oriented 125 simplex; and we declare $e_{215} = -e_{125}$ in the vector space. In our definitions of $C_0$, $C_1$, and $C_2$ above, we have chosen the positively oriented representatives as the basis vectors. However, the negatively oriented elements are in there, too, as $e_{21} = -e_{12} \in C_1$ and $e_{125} = -e_{215} \in C_2$, and so on.

What does a $k$-chain with multiple terms represent? In general, it is just a formal linear combination of $k$-simplices, and that is it. However, some of these linear combinations are special. For example, the 1-chain $e_{23} + e_{34} + e_{45} - e_{25}$ represents a one-dimensional cycle in the graph of vertices and edges of $\Delta$ (the "$-$" sign on $e_{25}$ arises because going around the cycle actually yields $e_{23} + e_{34} + e_{45} + e_{52}$, and $e_{52} = -e_{25}$). Similarly, the chain $e_{12} + e_{25} - e_{15}$ yields another cycle. In the context of $\Delta$, however, these cycles are fundamentally different: the first represents a "hole" in the simplicial complex, while the second one does not because there is a triangle that has "filled in" this hole. Homology is a method that uses linear algebra to algorithmically detect this difference. Here is how it works.

First, we need *boundary maps* between the vectors spaces $\{C_k\}$. These are linear transformations, $\partial_k : C_k \to C_{k-1}$, that take each $k$-chain to a $(k-1)$-chain. Since $\partial_k$ is a linear transformation, we can define it by specifying where it sends each basis element. The general rule for a basis vector $e_\sigma \in C_k$, where $\sigma = \{v_1, \ldots, v_{k+1}\}$ is an oriented $k$-simplex in $\Delta$, is as follows:

$$\partial_k \left(e_{\{v_1,\ldots,v_{k+1}\}}\right) = \sum_{i=1}^{k+1} (-1)^{i+1} e_{\{v_1,\ldots,\widehat{v_i},\ldots,v_{k+1}\}}, \qquad 1.$$

where $\widehat{v_i}$ indicates that this element is removed from the ordered list of vertices, so that $e_{\{v_1,\ldots,\widehat{v_i},\ldots,v_{k+1}\}} \in C_{k-1}$. For example,

$$\partial_2(e_{125}) = e_{25} - e_{15} + e_{12},$$
$$\partial_1(e_{12}) = e_2 - e_1.$$

Note that $\partial_0 : C_0 \to 0$ is defined to be the zero map—that is, the map that sends all inputs to 0. Because $\partial_k : C_k \to C_{k-1}$ and $\partial_{k-1} : C_{k-1} \to C_{k-2}$, adjacent boundary maps can be composed, such as $\partial_{k-1} \circ \partial_k$, and fit together into what is called a *chain complex*:

$$C_2 \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} 0.$$

Now let us go back to the two cycles we had considered in $C_1$ and apply the boundary map. Since $\partial_1$ is a linear transformation, we have

$$\begin{aligned}
\partial_1(e_{23} + e_{34} + e_{45} - e_{25}) &= \partial_1(e_{23}) + \partial_1(e_{34}) + \partial_1(e_{45}) - \partial_1(e_{25}) \\
&= e_3 - e_2 + e_4 - e_3 + e_5 - e_4 - (e_5 - e_2) = 0, \\
\partial_1(e_{12} + e_{25} - e_{15}) &= \partial_1(e_{12}) + \partial_1(e_{25}) - \partial_1(e_{15}) \\
&= e_2 - e_1 + e_5 - e_2 - (e_5 - e_1) = 0.
\end{aligned}$$

Both cycles map to zero! Initially, we called these 1-chains *cycles* because they corresponded to cycles in the underlying graph of $\Delta$ (see **Figure 4*b***). In homology, however, we now take this "mapping-to-zero" property as the definition of a cycle. Namely, a $k$-cycle is an element of $C_k$ that is in the *kernel* of the boundary map $\partial_k : C_k \to C_{k-1}$.

The idea of homology is to count such cycles (up to equivalence), but only the ones that are not "filled in" by higher-order simplices. It turns out that those are also easily detected, because they will arise in the *image* of the previous boundary map. Recall our earlier computation, $\partial_2(e_{125}) = e_{25} - e_{15} + e_{12}$. This gave us precisely the cycle bounding the 125 triangle. Combining computations, we obtain

$$\partial_1 \circ \partial_2(e_{125}) = \partial_1(e_{25} - e_{15} + e_{12}) = (e_5 - e_2) - (e_5 - e_1) + (e_2 - e_1) = 0.$$

In fact, it is true in general that

$$\partial_{k-1} \circ \partial_k = 0$$

for any $k$. This very important observation can be proven via direct computation using the general formula for $\partial_k$ given above in Equation 1, and it is usually assigned as a homework exercise to students in an algebraic topology course. What it means in terms of the usual notions of kernel and image for linear transformations is that

$$\operatorname{im} \partial_k \subseteq \ker \partial_{k-1} \subseteq C_{k-1}.$$

In particular, the *boundary* of a $k$-simplex in $C_k$ is always a cycle of $(k-1)$-simplices in $C_{k-1}$. But not all cycles are boundaries. Our other 1-cycle, $e_{23} + e_{34} + e_{45} - e_{25}$, is not the boundary of any element in $C_2$. The goal of homology is to detect the cycles that are *not* boundaries—that is, the genuine "holes."

We are finally ready to define *homology groups*. (Note that, in our case, these groups are actually vector spaces since the coefficients come from a field **k**.) The $k$th homology group of a simplicial complex $\Delta$ with coefficients in a field **k** is the *quotient vector space*:

$$H_k(\Delta) = \frac{\ker \partial_k}{\operatorname{im} \partial_{k+1}}.$$

The *Betti numbers*, $\beta_k$, are simply the dimensions of these quotient spaces:

$$\beta_k = \beta_k(\Delta) = \dim H_k(\Delta) = \dim(\ker \partial_k) - \dim(\operatorname{im} \partial_{k+1}).$$

Since the boundary maps are linear transformations, we can also write

$$\beta_k = \operatorname{nullity} \partial_k - \operatorname{rank} \partial_{k+1}.$$

These Betti numbers are integers, and they provide a simple summary of the homology of a topological space.

What are the Betti numbers in our example? Going back to the simplicial complex $\Delta$ in **Figure 4*b***, we see that the corresponding boundary maps are given quite explicitly, as matrices, in **Figure 4*c***. By multiplying matrices, it is straightforward to verify that $\partial_1 \circ \partial_2 = 0$ and $\partial_0 \circ \partial_1 = 0$. Furthermore, we easily see that $\operatorname{rank} \partial_2 = 1$, $\operatorname{nullity} \partial_2 = 0$, $\operatorname{rank} \partial_0 = 0$, $\operatorname{nullity} \partial_0 = 7$, and we can work out that $\operatorname{rank} \partial_1 = 5$ and $\operatorname{nullity} \partial_1 = 2$. We thus compute:

$$\beta_0 = \operatorname{nullity} \partial_0 - \operatorname{rank} \partial_1 = 7 - 5 = 2,$$

$$\beta_1 = \operatorname{nullity} \partial_1 - \operatorname{rank} \partial_2 = 2 - 1 = 1,$$

$$\beta_2 = \operatorname{nullity} \partial_2 - \operatorname{rank} \partial_3 = 0 - 0 = 0,$$

where $\partial_3 : C_3 \to C_2$ is a trivial map, whose rank is 0, since there are no 3-chains. Notice that $\beta_0 = 2$ is the number of connected components of $\Delta$, while $\beta_1 = 1$ is the number of 1-cycles that were not "filled in." And $\beta_2 = 0$ corresponds to the fact that there are no two-dimensional cavities. These patterns are not just a coincidence in our example. By design, $\beta_0$ always counts the number of connected components of a topological space, $\beta_1$ counts 1-cycles that are not boundaries (up to equivalence), $\beta_2$ counts 2-cycles, like the surface of a sphere, that are not "filled in," and so on.

The "up to equivalence" bit has so far been swept under the rug, but the idea is relatively simple. In the quotient spaces $\ker \partial_k / \operatorname{im} \partial_{k+1}$, two $k$-cycles in $\ker \partial_k$ get identified if they differ by a boundary in $\operatorname{im} \partial_{k+1}$. So, for example, if our simplicial complex were the shape of a cylinder, with a 1-cycle at the top and another at the bottom, the two 1-cycles would end up identified. The intuition that the linear algebra captures is that a loop of string around the top of a cylinder could be pulled across the surface to the bottom cycle. The two cycles are thus considered to be *homologous*, and only get counted once by $\beta_1$.

In our **Figure 4*b*** example, something similar happens if you consider the long 1-cycle $e_{12} + e_{23} + e_{34} + e_{45} - e_{15}$, which goes around both the filled-in triangle and the empty square in $\Delta$. Although this cycle is different from the two we considered before, and it is not itself a boundary, it differs from $e_{23} + e_{34} + e_{45} - e_{25}$ by precisely the boundary cycle $e_{12} + e_{25} - e_{15}$ (literally, add these two cycles and you will get the long one). So they get identified, and this identification is implicit in the computation we did where we found that $\beta_1 = 1$.

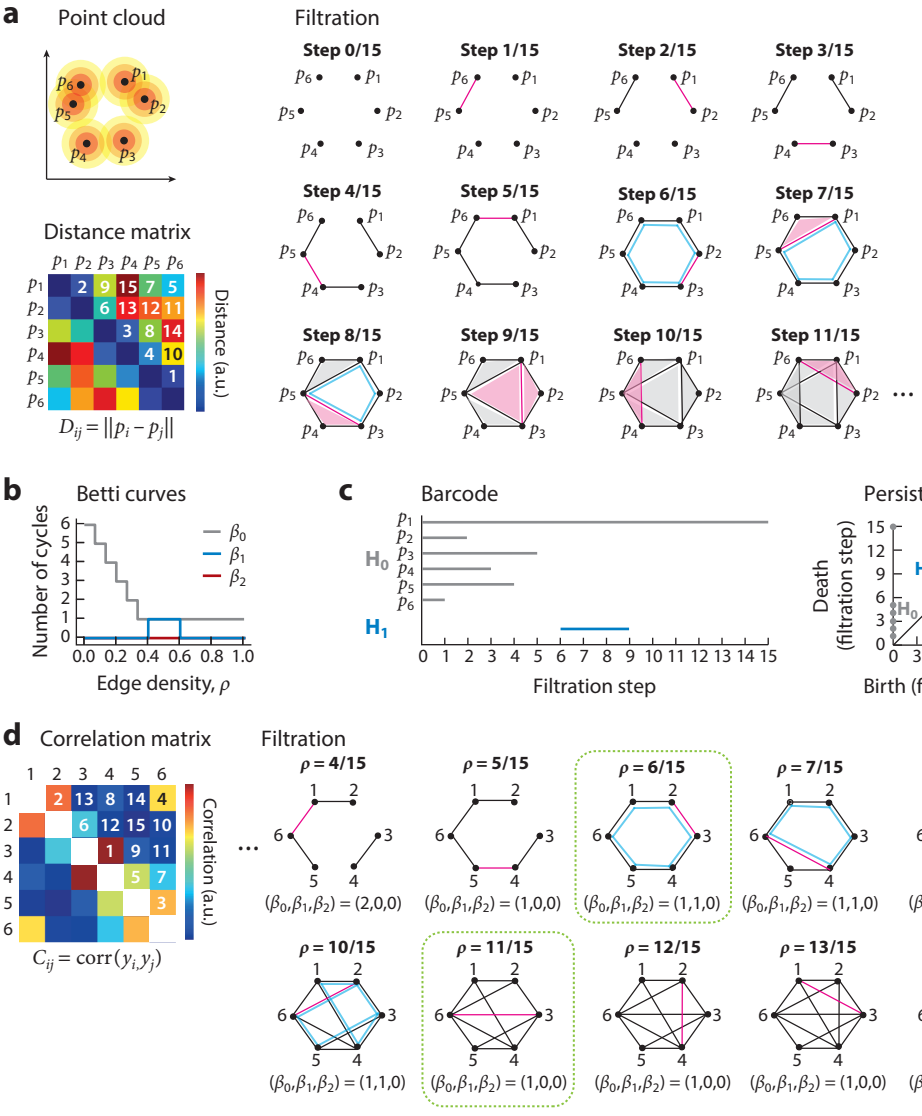
## PERSISTENT HOMOLOGY

### The Idea of Persistent Homology

Persistent homology was originally designed as a tool for detecting the underlying shape of point cloud data. **Figure 5*a*** illustrates this idea with $n = 6$ points in $\mathbb{R}^2$, but the goal of the technique is to detect structure in a large number of points living in a high-dimensional space $\mathbb{R}^d$. In our small example, the points appear to be noisy samples from a circle. But how can we "see" the circle?

Strictly speaking, the points themselves are topologically just a finite, disconnected set of points. The fact that they seem to come from a circle has to do with the particular pattern of distances between them. Intuitively, if we could somehow "fatten" the points, they would coalesce into a ring. Here is a simple strategy to make this intuition precise: imagine surrounding each point by a ball of radius $r$ and keeping track of which balls intersect as the radius $r$ grows. In this way, we can build a sequence of simplicial complexes—one for each value of $r$—that tracks the intersection data of the growing balls. The simplicial complex $\Delta(r)$ has a vertex for every point and an edge between two vertices if their balls of radius $r$ intersect. Higher-order simplices are then automatically included in $\Delta(r)$ when all pairs of vertices corresponding to the simplex are connected by an edge. (This kind of simplicial complex is often called a *clique complex* because there is a $(k-1)$-dimensional simplex for each $k$-clique in the underlying graph.)

Although in our small **Figure 5*a*** example we can visualize the growing balls and see by eye when pairs intersect, this is obviously not the way to build simplicial complexes in general. The good news is that all the information is actually given in a single distance matrix $D$, whose entries $D_{ij} = d(p_i, p_j)$ give the pairwise distances between points. The way to detect whether the balls centered at $p_i$ and $p_j$ intersect is simple: check whether or not $d(p_i, p_j) < 2r$. In fact, as you increase $r$ from 0 to $\infty$, you can imagine a sliding threshold on the distance matrix whereby at each $r$ you obtain a binary matrix $B(r)$ with $B_{ij}(r) = 1$ if $d(p_i, p_j) < 2r$, and 0 otherwise. Interpreting these binary matrices as adjacency matrices gives a nested sequence of graphs, $G(r)$, which form the basis for the simplicial complexes $\Delta(r)$. The distance matrix is all you need. In fact, the *ordering* of entries in the distance matrix is all you need.




*(Caption appears on following page)*

**Figure 5** (*Figure appears on preceding page*)

Examples of persistent homology. (*a*) Point cloud data yield a distance matrix that can then be input into a persistent homology computation. The first step is to translate the distance matrix into a filtration of graphs, which then gets turned into a filtration of simplicial complexes. (*b*) Betti curves track the Betti numbers $\beta_0$, $\beta_1$, $\beta_2$ across each step of the filtration. (*c*) Barcodes track individual homology cycles across the filtration and thus carry more information than Betti curves. Persistence diagrams encode the start (birth) and end (death) times of each bar in the barcode. (*d–e*) Topological data analysis can just as easily be applied to a pairwise correlation matrix. (*d*, *right*) In the filtration, a homology 1-cycle (*cyan*) appears at step 6 and disappears at step 11. Note that although there appear to be two 1-cycles at step 10, they only count as one because they are homologous. In step 11, the 1-cycle disappears because it becomes "coned off" by node 6. (*f*, *left*) Betti curves for random independent and identically distributed (i.i.d.) matrices are highly stereotyped (averages are *dashed lines*, 95% confidence intervals shown). (*f*, *right*) Betti curves for random Euclidean distance matrices are also stereotyped and vary by dimension. Panel *f* adapted from Giusti et al. (2015).



In **Figure 5*a*** (right), we see the first 12 steps of the filtration corresponding to the point cloud $p_1, \ldots, p_6$. The edges are added in order of increasing distance, so that the first added edge connects $p_5$ and $p_6$, the second connects $p_1$ and $p_2$, and so on. Notice that the first few steps serve only to connect various components of the graph. Then a homology 1-cycle emerges at step 6, and persists (albeit with a different shape) at steps 7 and 8, where the triangles with vertices $p_1, p_5, p_6$ and $p_3, p_4, p_5$ get "filled in" as 2-simplices because they are cliques in the graph. At step 9, the 1-cycle disappears. After building $G(r)$ and $\Delta(r)$, we can now compute homology groups for the simplicial complexes at every step of the filtration. The results of these computations are shown in **Figure 5*b*,*c***.

*Betti curves* are functions that simply track the Betti numbers, $\beta_k$, for each homology group at every step of the filtration. The Betti curves in **Figure 5*b*** track the number of connected components, $\beta_0(\rho)$, and the number of homology 1-cycles, $\beta_1(\rho)$, as a function of edge density $\rho$ as we move through the filtration. In **Figure 5*a***, we began with six disconnected points, yielding $\beta_0 = 6$ and $\beta_1 = \beta_2 = 0$ at step 0 of the filtration. As we progress, components get joined and $\beta_0$ gradually decreases to 1 by step 5 and then stays at 1. In contrast, $\beta_1(\rho)$ starts at 0, goes up to 1 at $\rho = 6/15$ (step 6 of the filtration), and goes back to 0 at $\rho = 9/15$ (step 9). Note that $\beta_2(\rho)$ is identically 0 in this example, as no homology 2-cycle ever emerges.

The *barcode* is perhaps the most common method for depicting how homology cycles emerge, persist, and disappear as we move from one step to the next in the filtration. At the beginning of the filtration, each point is in its own connected component, so we start with six gray bars in $H_0$, which we label by the points $p_1, \ldots, p_6$. At step 1, we add an edge joining $p_5$ and $p_6$; now there are only five connected components. Using the convention that the labels with higher index "die" first, the $p_6$ component gets merged into the $p_5$ component and the $p_6$ bar in $H_0$ ends. Similarly, at step 2 of the filtration we add an edge joining $p_1$ and $p_2$, so the $p_2$ component gets merged into the $p_1$ component and the $p_2$ bar ends. As we progress through the early steps of the filtration, each added edge connects two previously disconnected components, and another $H_0$ bar ends. Finally, at step 5 we are left with only one connected component, and this corresponds to the long gray bar that persists for the entire filtration. At step 6, when the homology 1-cycle emerges, we see a blue $H_1$ bar emerge. It persists until step 9. Note that all the information about the bars is encoded by the set of "birth" and "death" times defining the intervals over which the bars occur. In a *persistence diagram*, each bar is represented by a point whose coordinates are the ordered pair (birth time, death time).

While the barcode and the persistence diagram encode exactly the same information, the Betti curves lose information. Namely, they lose track of which homology cycles persist across several steps of the filtration. It is, in fact, nontrivial to identify homology cycles from one simplicial complex to the next in a filtration. Much of the early theory of persistent homology was making precise mathematical sense of this idea and developing rigorous tools to track cycles across a sequence of

homology computations. Betti curves do not require this tracking and can in principle be computed in parallel and for sequences of simplicial complexes that do not necessarily fit into a nested filtration. In practice, however, we use the same persistent homology software to compute Betti curves (Giusti et al. 2015, Sanderson 2023).

## How Long Does a Bar Need to Be?

When is a bar in a barcode long enough to indicate a significant topological feature of an underlying manifold? This is an important question taken seriously by the TDA community, which has been addressed with both theory and heuristics (see Cohen-Steiner et al. 2007, Bendich et al. 2013, Chazal et al. 2014, Bubenik 2015, Munch et al. 2015, Robins & Turner 2015, Adams et al. 2017, Kahle et al. 2023). One approach to determine the significance of a feature in a barcode is to shuffle the data in some way and see how the feature is affected. If shuffling destroys meaningful structure in the underlying manifold, we would expect the corresponding features in the barcode to disappear. This is the heuristic used by Gardner et al. (2021, 2022), where the data were shuffled to get rid of the relationship between grid cells with neighboring grid fields by shifting each cell's time series by a random independent amount. The authors then recomputed the pairwise distances between population vectors of grid cells and the corresponding persistent homology. Repeating this process 1,000 times produced a distribution of longest bar lengths for the shuffled control. Comparing the bars of the original barcode to this distribution allowed them to identify two significantly long bars in $H_1$ and one significantly long bar in $H_2$ (Gardner et al. 2021, 2022) (see **Figure 6*c***).

## Do You Need to Have a Distance Matrix?

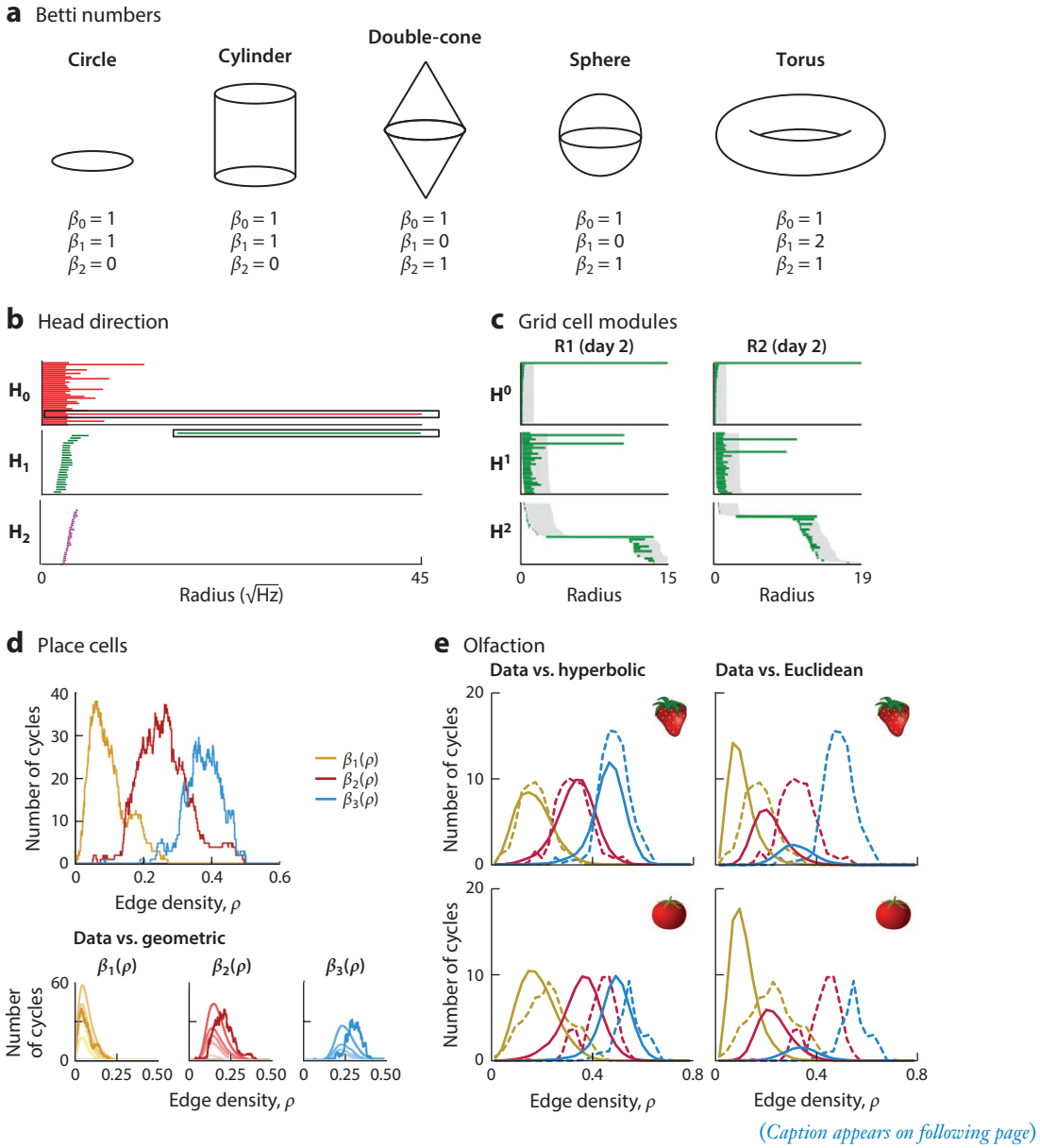
There is considerable confusion in the literature about whether or not using the techniques of persistent homology requires beginning with a distance matrix. In fact, the method works for any symmetric matrix. In particular, what is used as the "distance" between points need not satisfy a triangle inequality, and need not stem from an actual embedding of the data points in $\mathbb{R}^n$. One can just as easily compute persistent homology from a correlation matrix (**Figure 5*d*,*e***). The main caveat is that in the case of correlation or similarity matrices, it makes the most sense to run the filtration backwards—from high to low matrix values—so that the most similar points are connected first (**Figure 5*d***, right). This is analogous to connecting the "closest" points first when the input is a distance matrix and the filtration is run from low to high (**Figure 5*a***). Flipping the direction of the filtration can be easily accomplished by simply negating the input matrix; some software even gives an option to specify the direction (Sanderson 2023). It does not matter if the input is positive or negative—all symmetric matrices are equally valid inputs to the persistent homology pipeline. This is because Betti curves, barcodes, and persistence diagrams can all be computed with respect to a filtration step, as we did in **Figure 5*b*,*c***, rather than as a function of the radius of balls in Euclidean space.[4] What matters, then, is only the ordering with which edges are added in the filtration, and this comes from the ordering of entries in a symmetric matrix. These ideas are explained by Giusti et al. (2015) and underlie the flexibility we have

[4]That said, when the matrix entries are meaningful distances, it can be useful to track persistent homology features—and study their significance—as a function of ball radius. Moreover, because the point cloud setup was the original motivation, many software packages for persistent homology are superficially set up to expect a distance matrix, and the output barcodes and persistence diagrams are often parameterized by this radius. Translating to parameterize by filtration step amounts to a nonlinear rescaling of the $x$-axis for Betti curves and barcodes.

in neuroscience to apply TDA techniques to correlation and similarity matrices, as well as to "distance" matrices that may not satisfy the triangle inequality.

## What Kind of Structure Can Be Seen with Betti Curves?

While long bars in a barcode can help identify the shape of a neural manifold underlying a set of population vectors, Betti curves lose information about the persistence of individual cycles. However, they can provide valuable insight into the underlying geometry from correlations. This is because we can compare Betti curves from data to Betti curves from null models with




(*Caption appears on following page*)

**Figure 6** (*Figure appears on preceding page*)

Betti numbers, barcodes, and Betti curves in neuroscience applications. (*a*) Betti numbers for the circle, cylinder, double-cone, sphere, and torus. The circle and cylinder have the same Betti numbers because they are homotopy equivalent, while the double-cone and sphere Betti numbers match because they are homeomorphic. (*b*) Mammalian head direction cells during foraging in an open 2D environment produce population vectors whose persistent homology barcodes show one long bar in $H_0$, one long bar in $H_1$, and no long bars in $H_2$, indicating a Betti signature of ($\beta_0 = 1, \beta_1 = 1, \beta_2 = 0$) that is indicative of a connected, circular structure in the underlying neural manifold (Chaudhuri et al. 2019). This topological structure was also found during REM and non-REM sleep. (*c*) Persistent cohomology barcodes for population activity vectors from mammalian grid cells belonging to two distinct modules (R1 and R2) during 2D open field exploration. The barcodes show one long bar in $H^0$, two long bars in $H^1$, and one long bar in $H^2$, suggesting a toroidal structure for each module's neural activity (Gardner et al. 2022). (*d*) Hippocampal place cells during open-field exploration have pairwise spike-train correlations whose Betti curves reflect underlying geometric structure (Giusti et al. 2015). Betti curves $\beta_1(\rho)$, $\beta_2(\rho)$, and $\beta_3(\rho)$ for a place cell correlation matrix are shown in yellow, red, and blue, respectively. (*Bottom*) The same three Betti curves for the place cell data are shown separately, overlayed on top of average Betti curves for random Euclidean geometric controls in various dimensions (*smooth curves*, lighter shades correspond to lower dimensions). (*e*) Betti curves for the correlations of monomolecular odor concentrations of strawberries and tomatoes (*dashed lines*) match those of an underlying low-dimensional hyperbolic geometry (*solid lines*, *left*) and differ markedly from Betti curves for low-dimensional Euclidean structure (*solid lines*, *right*). Panel *b* adapted with permission from Chaudhuri et al. (2019), panel *c* adapted from Gardner et al. (2022) (CC BY 4.0), panel *d* adapted from Giusti et al. (2015), and panel *e* adapted with permission from Zhou et al. (2018) (CC BY-NC 4.0).

known geometric (or random) structure. In particular, Betti curves for different families of matrices are strikingly stereotyped (Kahle 2009, 2011; Kahle & Meckes 2013; Giusti et al. 2015; Zhou et al. 2018). For example, random symmetric matrices with independent and identically distributed (i.i.d.) entries chosen uniformly from [0,1] have $\max(\beta_1) < \max(\beta_2) < \max(\beta_3)$, whereas random (Euclidean) geometric matrices whose entries are pairwise distances between random points in $[0, 1]^d$ have $\max(\beta_1) > \max(\beta_2) > \max(\beta_3)$ (Giusti et al. 2015) (see **Figure 5*f***). Moreover, the Betti curves for random i.i.d. matrices are orders of magnitude larger than Betti curves for random geometric matrices. The utility of Betti curves to detect underlying geometric structure is not altogether surprising, as Robins & Turner (2015) showed that certain summary statistics of persistence diagrams can distinguish a variety of subtly distinct random spatial point processes.

## Persistent Homology as a Tool for Matrix Analysis

More generally, we see that Betti curves, barcodes, and other features that we can compute via persistent homology can be thought of as methods for associating topological invariants to symmetric matrices. In other words, they are new tools for matrix analysis (Giusti et al. 2015). But what are they invariants of? Recall that in traditional matrix analysis we typically compute spectral features, such as eigenvalues and singular values; these are invariants to change-of-basis transformations from linear algebra. In contrast, the topological features computed via persistent homology can produce invariants under nonlinear rescaling of the matrix entries, where the nonlinearity is *monotonic* and thus order preserving. Since barcodes and Betti curves can be parameterized by filtration step, as in **Figure 5*a–f***, any order-preserving transformation of the matrix entries results in a new matrix with exactly the same barcodes and Betti curves. TDA thus provides an opportunity to move beyond traditional matrix invariants, based on linear algebra, to invariants that may be more appropriate for neural data—which often contain hidden monotonic nonlinearities that can lead to artifacts when using spectral techniques (Giusti et al. 2015). By design, many modern TDA tools are agnostic to these hidden nonlinearities. In particular, persistent homology may be useful for detecting underlying features of matrices that are obscured by such nonlinearities, including low-rank or low-dimensional structure (Curto et al. 2021; E. Hansen, N. Sanderson, S. Nourin, V. Candat, C. Curto & G. Sumbre, unpublished manuscript).

## Software

Early software packages to compute homology and persistent homology include CHomP (Harker & Mischaikow 2014, Pilarczyk et al. 2014), JavaPlex (Tausz et al. 2014), Perseus (Mischaikow & Nanda 2013), Dionysus (Morozov 2023a), PHAT (Bauer et al. 2017), GUDHI (Boissonat et al. 2025, GUDHI Proj. 2025), and the R-package TDA (Fasy et al. 2025).[5]

More recent software includes Ripser (Tralie et al. 2018), SimBa (Dey et al. 2019), Persim (Saul 2019), Eirene (Henselman & Ghrist 2016, Henselman-Petrusek 2021), Dionysus 2 (Morozov 2023b), giotto-tda (Tauzin et al. 2021), DREiMac (Perea et al. 2023), and open applied topology (OAT) (Henselman-Pestrusek et al. 2024). The Python package Ripser is perhaps the most widely used because of its accessibility and speed. Ripser takes as input a distance matrix and outputs a persistence diagram (equivalently barcode) as a list of (birth, death) pairs (see **Figure 5*c***).

Relevant to many neuroscience applications, the same persistent homology computations can also take as input a correlation matrix, although most persistent homology software packages do not make this explicit. Computation of Betti curves is also not standard in many TDA packages, although a sophisticated user can easily compute Betti curves from a barcode or persistence diagram. Because of this, a couple of "wrapper" software packages have been devised for easier use by neuroscientists. An older wrapper, CliqueTop (Giusti 2015), was written for use by Giusti et al. (2015) and was subsequently used by Zhou et al. (2018), Zhou & Sharpee (2022), and Zhang et al. (2023). CliqueTop is MATLAB code that calls Perseus to compute persistent homology and translates the results into Betti curves. It can take any symmetric matrix (e.g., distance matrices or correlation matrices) as input. We recently introduced an updated wrapper in Python, PyCliqueTop_2023, that instead calls Ripser for faster, more memory-efficient persistent homology computations (Sanderson 2023). PyCliqueTop_2023 is designed to work in exactly the same way as CliqueTop, taking any symmetric matrix as input and providing Betti curves as output. It was first used to analyze neural correlations in zebrafish calcium imaging data (E. Hansen, N. Sanderson, S. Nourin, V. Candat, C. Curto & G. Sumbre, unpublished manuscript).

# BACK TO NEUROSCIENCE: TOPOLOGY AS A WINDOW INTO CIRCUIT FUNCTION

Now that we have seen more of the mathematical details of topology, Betti numbers, and persistent homology, we return to a few of the key neuroscience examples from before. In a nutshell, there are three main types of applications, all represented in **Figure 2*a*** (counterclockwise from bottom left):

1. Persistent homology can be used to detect topological features of an underlying *neural manifold*. The neural manifold may represent the various states of a continuous or dynamic attractor or a collection of population vectors that arise as stimulus responses (or both).
2. Persistent homology can be used to understand the topological structure of a *stimulus or representation space*. The input can be stimulus-dependent neural correlations or another type of stimulus similarity matrix, as in the olfaction study (Zhou et al. 2018).
3. Persistent homology and its variations, such as the persistent homology transform (Turner et al. 2014), may be used to classify various graphical and tree-like structures that arise in the *physical brain* (e.g., neuron morphology, vasculature, and connectomes).

It is important to note, however, that these applications are far from exhaustive. As previously mentioned, topological tools have also been used in a "black box" manner to provide new

[5]Note that final publication dates often come much later than the original software release dates.

data-driven features from, say, fMRI or other imaging data (Giusti et al. 2016; Sizemore et al. 2016, 2018; Anderson et al. 2018). In this context, topological features may be more difficult to interpret, but they appear to have diagnostic value (Stolz et al. 2021).

## Connecting Neural Circuits to Function

Perhaps the most exciting aspect about using topology in neuroscience, however, is the potential to connect structures that are echoed across different areas of analysis, such as the ones represented in **Figure 2*a***. In particular, topology can provide a natural approach to unifying the neural manifold and neural circuit perspectives (Langdon et al. 2023), as well as connecting neural circuits and activity to the structure of represented spaces (Curto 2017). For example, the circular structure of a stimulus/representation space, such as the space of possible heading directions, can be reflected in the topology of the corresponding neural circuit—for example, the ring of "compass neurons" in the fly's neural compass system (Kim et al. 2017). In this case, the topological structure of the represented space and that of the neural circuit are more or less obvious. However, the fact that this same circular structure is also echoed in high-dimensional *neural activity* is much less transparent, especially as the flies' mental maps can be modified by experience (Fisher et al. 2019, Kim et al. 2019). In contrast, in the mammalian head direction circuit or grid cell circuits, circular or toroidal topologies are expected from the representations, but it is the physical structure of the network that is least transparent. In these cases, topological tools have been successfully used to show that the structure of neural activity echoes the topology expected from representations (Chaudhuri et al. 2019, Gardner et al. 2022). We can now explain how.

**Figure 6*a*** reviews the Betti numbers associated to various basic topological structures, including the circle, cylinder, sphere, and torus. Shapes that are homotopy equivalent or homeomorphic have identical Betti numbers. When persistent homology is applied to population activity vectors that lie along a lower-dimensional neural manifold, the barcode is expected to contain long bars corresponding to the nonzero Betti numbers of the manifold. In the case of mammalian head direction cells, the barcodes revealed a Betti signature of ($\beta_0 = 1, \beta_1 = 1, \beta_2 = 0$), consistent with a circle or a cylinder (Chaudhuri et al. 2019) (**Figure 6*b***). For grid cells, on the other hand, the barcodes for each grid cell module produced the topological signature of a torus: ($\beta_0 = 1, \beta_1 = 2, \beta_2 = 1$) (Gardner et al. 2022) (**Figure 6*c***). In both cases, population activity vectors were analyzed as point cloud data, following the first paradigm of **Figure 2*b***.

However, it is important to note that the data can be sufficiently high-dimensional and noisy that various dimensional reduction methods are often applied before persistent homology computations are performed (Kang et al. 2021). For example, the data preprocessing by Gardner et al. (2022) involved first clustering grid fields to yield distinct grid cell modules, projecting the population activity vectors for each module to a 6D space using principal component analysis (PCA), and then downsampling in time to yield a $6 \times 1{,}200$ matrix reflecting 1,200 reduced activity vectors in $\mathbb{R}^6$. From here, a $1{,}200 \times 1{,}200$ distance matrix was computed, and this was input into the Ripser software. The outputs were barcodes, as seen in **Figure 6*c***.[6]

In the framework of **Figure 2*b***, the above analyses focused on point clouds of population activity vectors given by the columns of the neural activity matrix. **Figure 6*d*,*e*** shows results using the complementary approach, focusing on correlation structure computed from the rows. The goal of

[6]The topology computation reported by Gardner et al. (2022) was technically persistent cohomology rather than persistent homology. Homology and cohomology are dual topological theories, although cohomology typically contains additional information. The distinction is not important for understanding the barcode results presented here and is beyond the scope of this review.

these analyses was to infer the *geometric* structure of the represented space via topological signatures of neural or stimulus correlations. Using the CliqueTop software (Giusti 2015), Betti curves were computed from place cell correlations and compared to those arising from various control matrices (Giusti et al. 2015). The place cell Betti curves revealed a low-dimensional Euclidean structure (**Figure 6*d***). In contrast, a very similar Betti curve analysis in the case of olfactory correlations revealed low-dimensional hyperbolic geometry (Zhou et al. 2018) (**Figure 6*e***). Although the interpretation of persistent homology is less straightforward when the input is a matrix of similarities or correlations—as opposed to distances—this approach allows one to gain interesting geometric insights even when the underlying space being represented is topologically trivial, such as the place field representation of an open field environment (Giusti et al. 2015, Curto 2017).

## Circling Back to Grid Cells

We end by going back to the beginning. In the case of grid cells, it was expected that neural activity would take the shape of a torus because the structure of the space represented by grid fields is a torus (Curto 2017, Kang et al. 2021). However, this expectation only held in the context of an animal wandering around a two-dimensional "open field" environment. What would happen when the interaction between grid fields and the represented space was not so transparent? Fortunately, Gardner et al. (2021, 2022) also addressed this question. They recorded grid cells while the animal was in a wagon-wheel environment, which itself had multiple holes, as well as while the animal was running in place (running wheel), and during sleep. In all cases, they found the neural population activity conformed to a torus in $\mathbb{R}^n$. This remarkable fact provides evidence that the toroidal topology of grid cell activity is emerging from within the entorhinal cortical circuit itself, rather than being inherited from external sensory inputs. In other words, the topological findings reveal an important fact about circuit structure and function. Similarly, Giusti et al. (2015) found that the geometric structure of place cell coding detected by Betti curves was also present during wheel running and sleep.

Finding the tori in grid cell activity required first finding the modules in a large population of grid cells, and this is a step that involved considerable preprocessing (Gardner et al. 2022). Grid fields needed to be computed for each neuron, and these in turn had to be classified according to their scale and orientation so that the grid cells could be separated by module (see **Figure 1**). What if we did not have access to the grid fields, but only the grid cell activity? This is analogous to the question asked by Curto & Itskov (2008), who argued that place cell activity should reflect the topology of an animal's environment—even when the place fields are unknown. What shape is represented by the full population of grid cells when the individual grid fields—and corresponding modules—are unknown? This remains an open question.

# APPENDIX

Here we provide a table (**Table 1**) with some standard mathematical notation, as well as a glossary of terms that were used in our explanations of topology and persistent homology. All notation and definitions are standard in the mathematics literature and can be found in textbooks (e.g., see Hatcher 2002). We include them for completeness, as many of these terms may be unfamiliar to neuroscientists.

## Functions

A function $f$ is typically denoted as $f : X \to Y$, where $X$ is the domain (input space) and $Y$ is the codomain (output space). A subset of elements in $X$ satisfying a certain property $\mathcal{P}$ is denoted with the set-theoretic notation $\{x \in X \mid x \text{ satisfies } \mathcal{P}\}$, where "|" is read as "such that."

**Table 1 Some common mathematical notation**

| Notation | Meaning |
|---|---|
| $\mathbf{k}$ | A field of numbers such as $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{C}$, or $\mathbb{Z}/p\mathbb{Z}$ for a prime $p$ |
| $\mathbb{R}$ | The real numbers |
| $\mathbb{Q}$ | The rational numbers |
| $\mathbb{C}$ | The complex numbers |
| $\mathbb{Z}$ | The integers (a ring but not a field because of a lack of multiplicative inverses) |
| $[a, b]$ | The closed interval of all real numbers between $a$ and $b$, including end points |
| $(a, b)$ | The open interval of all real numbers between $a$ and $b$, excluding end points |
| $\emptyset$ | The empty set |
| $f^{-1}$ | The inverse of a function $f$ |
| $\mathrm{id}_X$ | The identity map (function) on $X$ that sends every element of $X$ to itself |
| $g \circ f$ | The composition of functions $f$ and $g$; if $f: X \to Y$ and $g: Y \to X$, then $g \circ f: X \to X$ sends $x \mapsto g(f(x))$, and $f \circ g: Y \to Y$ sends $y \mapsto f(g(y))$ |
| $\in$ | An element of (e.g., $\sigma \in \Delta$ means $\sigma$ is an element of the set $\Delta$) |
| $\subseteq$ | A subset of (e.g., $\tau \subseteq \sigma$ means all elements of $\tau$ are also in $\sigma$) |
| $\beta_k$ | The $k$th Betti number, an integer quantifying a topological feature |

- **Preimage**: Given a function $f: X \to Y$ and an element $y \in Y$, the *preimage* $f^{-1}(y)$ is the set of all elements of $X$ that map to $y$ under $f$. In other words, $f^{-1}(y) = \{x \in X \mid f(x) = y\}$.
- **Inverse $f^{-1}$**: If $f: X \to Y$, the *inverse* $f^{-1}$ assigns to each $y \in Y$ its preimage $f^{-1}(y)$. Note that the set $f^{-1}(y)$ may have more than one element, or be empty, so the inverse is not necessarily a function.
- **Bijection**: A function $f: X \to Y$ is a *bijection* if it provides a one-to-one correspondence between elements of $X$ and $Y$. That is, for each $y \in Y$ there is a unique $x \in X$ such that $f(x) = y$. In this case, $f^{-1}$ is also a bijection.
- **Monotonic**: A function $f: \mathbb{R} \to \mathbb{R}$ is *monotonically increasing* (*decreasing*) if for all $x < y$, we have $f(x) < f(y)$ [$f(x) > f(y)$ for decreasing].

## Combinatorics

The following combinatorial objects are essential for understanding the basics of simplicial and persistent homology.

- **Graph**: A *graph* $G = (V, E)$ is a set of vertices, $V = \{1, 2, \ldots, n\}$, and a set of edges between vertices, $E \subseteq \{(i, j) \mid i, j \in V\}$. In a *directed graph*, the edges are viewed as ordered pairs, so that $(i, j) \in E$ indicates $i \to j$ in $G$.
- **Clique**: A *clique* is a complete graph (or subgraph) where all possible edges between vertices are included.
- **Hypergraph**: A *hypergraph* is a generalization of a graph that allows for higher-dimensional "edges" consisting of subsets with more than two vertices. Equivalently, a hypergraph is simply a set of subsets of $\{1, \ldots, n\}$.
- **Simplicial complex**: A *simplicial complex* $\Delta$ on $n$ vertices (nodes) is a set of subsets of $\{1, \ldots, n\}$ that is closed under subsets of its elements. In other words, if $\sigma \in \Delta$ and $\tau \subseteq \sigma$, then $\tau \in \Delta$. The elements of $\Delta$ are called *simplices* and are often thought of geometrically as in **Figure 4*a***.
- **Permutation**: A *permutation* is a bijection $\pi: \{1, \ldots, n\} \to \{1, \ldots, n\}$ that reorders the elements in a sequence $12 \cdots n$. For example, $12345 \to 15324$ corresponds to the permutation with $\pi(1) = 1$, $\pi(2) = 5$, $\pi(3) = 3$, $\pi(4) = 2$, and $\pi(5) = 4$.

- **Transposition**: A *transposition* is a permutation that interchanges exactly two elements of a sequence and fixes all others. For example, $12345 \to 32145$ corresponds to the transposition $1 \leftrightarrow 3$ with $\pi(1) = 3$, $\pi(3) = 1$, and all other elements sent to themselves.
- **Even/odd permutation**: A permutation is called *even* if it can be achieved as the composition of an even number of transpositions. Otherwise, it is *odd*. For example, $12345 \to 23145$ is an even permutation, as it can be achieved by composing $1 \leftrightarrow 2$ with $1 \leftrightarrow 3$. In contrast, $12345 \to 42315$ is odd.

## Topology

The following concepts take some work to get used to. For more intuition and details, we recommend Chapter 0 of Hatcher (2002).

- **Topological space**: A *topological space* is a set $X$ equipped with a "topology" $\mathcal{U}$ (see below).
- **A topology**: A *topology* on a set $X$ is a collection $\mathcal{U}$ of subsets of $X$, called the *open sets* of $X$, that satisfy the following three criteria: (*a*) $\emptyset \in \mathcal{U}$ and $X \in \mathcal{U}$; (*b*) any union of open sets, possibly infinite, is also an open set in $\mathcal{U}$; and (*c*) any intersection of finitely many open sets is also an open set in $\mathcal{U}$.[7]
- **Continuous function (or "map")**: A function $f : X \to Y$ is *continuous* if for all open sets $U \subseteq Y$ the preimage $f^{-1}(U)$ is an open set in $X$. Note that this definition requires a designation of what are the "open sets" (the topology) for both $X$ and $Y$. This matches the familiar notion of continuity for real-valued functions $f : \mathbb{R} \to \mathbb{R}$, when $\mathbb{R}$ is equipped with the standard topology based on open intervals of the form $(a, b)$.
- **Homeomorphism**: A *homeomorphism* is a continuous bijective function $f : X \to Y$ such that $f^{-1}$ is also continuous.
- **Homeomorphic**: Two topological spaces $X$ and $Y$ are said to be *homeomorphic* if there exists a homeomorphism $f : X \to Y$. This is the strongest notion of "topological equivalence" that two spaces can have, and we write $X \cong Y$.
- **Homotopy**: A family of continuous functions $h_t : X \to Y$, for $t \in [0, 1]$, is a *homotopy* if the associated map $H : X \times [0, 1] \to Y$, given by $H(x, t) = h_t(x)$, is continuous.
- **Homotopic**: We say that two functions $f : X \to Y$ and $g : X \to Y$ are *homotopic*, and write $f \simeq g$, if there exists a homotopy $h_t$ connecting them so that $f = h_0$ and $g = h_1$.
- **Homotopy equivalent**: Two topological spaces $X$ and $Y$ are *homotopy equivalent* if there exist maps $f : X \to Y$ and $g : Y \to X$ such that $g \circ f \simeq \mathrm{id}_X$ and $f \circ g \simeq \mathrm{id}_Y$.

## Linear Algebra

Here we review standard terms from linear algebra that can be found in any introductory textbook. We assume familiarity with vector spaces and omit the definition with its lengthy list of axioms. The concept of a quotient vector space is probably less familiar, but essential for defining homology groups.

- **Linear transformation**: A *linear transformation* is a function $T : X \to Y$ between two vector spaces, $X$ and $Y$, satisfying $T(ax + by) = aT(x) + bT(y)$ for all vectors $x, y \in X$ and all scalars $a, b$.
- **Kernel**: The *kernel*, $\ker T$, of a linear transformation $T : X \to Y$ is the set of all vectors $x \in X$ such that $T(x) = 0$.

[7] Why only finite intersections? Consider open sets of the form $(-\frac{1}{n}, \frac{1}{n}) \subseteq \mathbb{R}$, for $n = 1, 2, 3, \ldots$. Any finite intersection yields the open interval $(-\frac{1}{N}, \frac{1}{N})$ where $N$ is the largest $n$ among the intersected sets. If, however, we take the infinite intersection $\bigcap_{n=1}^{\infty}(-\frac{1}{n}, \frac{1}{n})$, we obtain the set $\{0\}$, which is not open.

- **Image**: The *image*, $\operatorname{im} T$, of a linear transformation $T : X \to Y$ is the set of all vectors $y \in Y$ such that $y = T(x)$ for some $x \in X$.
- **Rank**: The *rank* of a linear transformation $T$, denoted $\operatorname{rank} T$, is the dimension of its image. If $T$ is given as a matrix, then $\operatorname{rank} T$ is the dimension of the column space; equivalently, it is the largest number of linearly independent column vectors of $T$.
- **Nullity**: The *nullity* of $T$, denoted $\operatorname{nullity} T$, is the dimension of its kernel. The rank-nullity theorem says that for $T : X \to Y$, we always have $\operatorname{rank} T + \operatorname{nullity} T = \dim X$.
- **Quotient vector space**: If $X$ and $Y$ are vector spaces, and $X \subseteq Y$, then the *quotient space* $Y/X$ is the vector space obtained by identifying all elements of $Y$ that differ by an element of $X$. We can represent the elements of $Y/X$ as equivalence classes $[y]$, for $y \in Y$, where $[y] = \{y + x \mid x \in X\}$. With this notation, $Y/X = \{[y] \mid y \in Y\}$.

## DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

## ACKNOWLEDGMENTS

This work was supported by National Institutes of Health R01 NS120581. We would like to thank Nikolas Schonsheck for helpful feedback on the manuscript. We would also like to thank an anonymous reviewer for many useful comments and for encouraging us to include a glossary of mathematical terms.

## LITERATURE CITED

Adams H, Emerson T, Kirby M, Neville R, Peterson C, et al. 2017. Persistence images: a stable vector representation of persistent homology. *J. Mach. Learn. Res.* 18(1):218–52

Anderson KL, Anderson JS, Palande S, Wang B. 2018. Topological data analysis of functional MRI connectivity in time and space domains. *Connect Neuroimaging* 11083:67–77

Bardin JB, Spreemann G, Hess K. 2019. Topological exploration of artificial neuronal network dynamics. *Netw. Neurosci.* 3(3):725–43

Bauer U, Kerber M, Reininghaus J, Wagner H, Skraba P. 2017. PHAT—persistent homology algorithms toolbox. *J. Symb. Comput.* 78:76–90

Bendich P, Edelsbrunner H, Morozov D, Patel A. 2013. Homology and robustness of level and interlevel sets. *Homol. Homotopy Appl.* 15(1):51–72

Bendich P, Marron J, Miller E, Pieloch A, Skwerer S. 2016. Persistent homology analysis of brain artery trees. *Ann. Appl. Stat.* 10(1):198–218

Boissonnat JD, Chazal F, Yvinec M. 2018. *Geometric and Topological Inference*. Cambridge Texts in Applied Mathematics 57. Cambridge Univ. Press. 1st ed.

Boissonat JD, Loiseaux D, Glisse M, Rouvreau V, Jamin C, et al. 2025. GUDHI. *Software*. **https://github.com/GUDHI**

Bubenik P. 2015. Statistical topological data analysis using persistence landscapes. *J. Mach. Learn. Res.* 16:77–102

Carlsson G, Vejdemo-Johansson M. 2022. *Topological Data Analysis with Applications*. Cambridge Univ. Press

Catanzaro M, Rizzo S, Kopchick J, Chowdury A, Rosenberg D, et al. 2024. Topological data analysis captures task-driven fMRI profiles in individual participants: a classification pipeline based on persistence. *Neuroinformatics* 22(1):45–62

Chaudhuri R, Gerçek B, Pandey B, Peyrache A, Fiete I. 2019. The intrinsic attractor manifold and population dynamics of a canonical cognitive circuit across waking and sleeping. *Nat. Neurosci.* 22:1512–20

Chazal F, Fasy B, Lecci F, Rinaldo A, Wasserman L. 2014. Stochastic convergence of persistence landscapes and silhouettes. In *SOCG'14: Proceedings of the Thirtieth Annual Symposium on Computational Geometry*. ACM. **https://doi.org/10.1145/2582112.2582128**

Chen S, Thielk M, Gentner TQ. 2024. Auditory feature-based perceptual distance. Preprint, bioRxiv. **https://www.biorxiv.org/content/10.1101/2024.02.28.582631v1**

Chowdhury S, Mémoli F. 2018. A functorial Dowker theorem and persistent homology of asymmetric networks. *J. Appl. Comput. Topol.* 2:115–75

Cohen-Steiner D, Edelsbrunner H, Harer J. 2007. Stability of persistence diagrams. *Discrete Comput. Geom.* 37(1):103–20

Curry J, DeSha J, Kanari L, Mallery B, et al. 2024. From trees to barcodes and back again II: combinatorial and probabilistic aspects of a topological inverse problem. *Comput. Geom.* 116:102031

Curto C. 2017. What can topology tell us about the neural code? *Bull. AMS* 54:63–78

Curto C, Itskov V. 2008. Cell groups reveal structure of stimulus space. *PLOS Comput. Biol.* 4:e1000205

Curto C, Itskov V, Veliz-Cuba A, Youngs N. 2013. The neural ring: an algebraic tool for analyzing the intrinsic structure of neural codes. *Bull. Math. Biol.* 75(9):1571–611

Curto C, Paik J, Rivin I. 2021. Betti curves of rank one symmetric matrices. In *Geometric Science of Information, 5th International Conference, GSI 2021, Paris, France, July 21–23, Proceedings*. GSI. **https://dokumen.pub/geometric-science-of-information-5th-international-conference-gsi-2021-paris-france-july-2123-2021-proceedings-lecture-notes-in-computer-science-12829-1st-ed-2021-3030802086-9783030802080.html**

Dabaghian Y, Brandt V, Frank L. 2014. Reconceiving the hippocampal map as a topological template. *eLife* 3:e03476

Dabaghian Y, Mémoli F, Frank L, Carlsson G. 2012. A topological paradigm for hippocampal spatial map formation using persistent homology. *PLOS Comput. Biol.* 8:e1002581

Dey TK, Shi D, Wang Y. 2019. SimBa: an efficient tool for approximating rips-filtration persistence via simplicial batch collapse. *J. Exp. Algorithmics* 24:1–16

Dey TK, Wang Y. 2022. *Computational Topology for Data Analysis*. Cambridge Univ. Press

Dowker C. 1952. Homology groups of relations. *Ann. Math.* 56:84–95

Edelsbrunner H, Harer J. 2009. *Computational Topology: An Introduction*. Am. Math. Soc.

Edelsbrunner H, Letscher D, Zomorodian A. 2002. Topological persistence and simplification. *Discrete Comput. Geom.* 28:511–33

Edelsbrunner H, Morozov D. 2013. Persistent homology: theory and practice. In *European Congress of Mathematics, Krakow, 2–7 July, 2012*. ESM Press. **https://ems.press/books/standalone/229/4399**

Fasy B, Kim J, Lecci F, Maria C, Millman D, Rouvreau V. 2025. Statistical tools for topological data analysis. *Statistical Software*. **https://cran.r-project.org/web/packages/TDA/TDA.pdf**

Fisher YE, Lu J, D'Alessandro I, Wilson R. 2019. Sensorimotor experience remaps visual input to a heading direction network. *Nature* 576:121–25

Gardner R, Hermansen E, Pachitariu M, Burak Y, Baas N, et al. 2021. Toroidal topology of population activity in grid cells. Preprint, bioRxiv. **https://www.biorxiv.org/content/10.1101/2021.02.25.432776v1**

Gardner R, Hermansen E, Pachitariu M, Burak Y, Baas N, et al. 2022. Toroidal topology of population activity in grid cells. *Nature* 602:123–28

Ghrist R. 2007. Barcodes: the persistent topology of data. *Bull. Am. Math. Soc.* 45(1):61–75

Ghrist R. 2014. *Elementary Applied Topology*. Createspace. 1st ed.

Giusti C. 2015. CliqueTop: Matlab package for clique topology of symmetric matrices. *Software*. **https://github.com/nebneuron/clique-top**

Giusti C, Ghrist R, Bassett D. 2016. Two's company, three (or more) is a simplex: algebraic-topological tools for understanding higher-order structure in neural data. *J. Comput. Neurosci.* 41(1):1–14

Giusti C, Pastalkova E, Curto C, Itskov V. 2015. Clique topology reveals intrinsic geometric structure in neural correlations. *PNAS* 112(44):13455–60

GUDHI Proj. 2025. *GUDHI User and Reference Manual*. GUDHI Editor. Board. **https://gudhi.inria.fr/doc/3.11.0/**

Haft-Javaherian M, Villiger M, Schaffer CB, Nishimra N, Golland P, Bouma BE. 2020. A topological encoding convolutional neural network for segmentation of 3D multiphoton images of brain vasculature

using persistent homology. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE. **https://ieeexplore.ieee.org/document/9150930**

Harker S, Mischaikow K. 2014. Using CHomP and conley-morse-database. Paper presented at the Lorentz Center, Leiden Univ., Aug. **https://chomp.rutgers.edu/Projects/Databases_for_the_Global_Dynamics/software/LorentzCenterAugust2014.pdf**

Hatcher A. 2002. *Algebraic Topology*. Cambridge Univ. Press

Henselman G, Ghrist R. 2016. Matroid filtrations and computational persistent homology. Preprint, arXiv:1606.00199 [math.AT]

Henselman-Petrusek G. 2021. Eirene. *Software*. **https://github.com/henselman-petrusek/Eirene.jl**

Henselman-Pestrusek G, Ziegelmeier L, Giusti C. 2024. Open applied topology. *Software*. **https://github.com/OpenAppliedTopology**

Kacynski T, Mischaikow K, Mrozek M. 2004. *Computational Homology*. Springer

Kahle M. 2009. Topology of random clique complexes. *Discrete Math*. 309(6):1658–71

Kahle M. 2011. Random geometric complexes. *Discrete Comput. Geom.* 45:553–73

Kahle M, Meckes E. 2013. Limit theorems for Betti numbers of random simplicial complexes. *Homol. Homot. Appl.* 15(1):343–74

Kahle M, Tian M, Wang Y. 2023. On the clique number of noisy random geometric graphs. *Random Struct. Algorithms* 63(1):242–79

Kanari L, Dictus H, Hess K, Markram H, et al. 2022. Computational synthesis of cortical dendritic morphologies. *Cell Rep*. 39(1):110586

Kanari L, Dłotko P, Scolamiero M, Levi R, Shillcock J, et al. 2017. A topological representation of branching neuronal morphologies. *Neuroinformatics* 16:3–13

Kanari L, Garin A, Hess K. 2020. From trees to barcodes and back again: theoretical and statistical perspectives. *Algorithms* 13(12):335

Kanari L, Ramaswamy S, Shi Y, Morand S, Meystre J, et al. 2019. Objective morphological classification of neocortical pyramidal cells. *Cereb. Cortex* 29(4):1719–35

Kang L, Xu B, Morozov D. 2021. Evaluating state space discovery by persistent cohomology in the spatial representation system. *Front. Comput. Neurosci.* 15:616748

Kim S, Hermundstad A, Romani S, Abbott L, Jayaraman V. 2019. The generation of stable heading representations in diverse visual scenes. *Nature* 576:126–31

Kim S, Roualt H, Druckmann S, Jayaraman V. 2017. Ring attractor dynamics in the *Drosophila* central brain. *Science* 356(6340):849–43

Langdon C, Genkin M, Engel T. 2023. A unifying perspective on neural manifolds and circuits for cognition. *Nat. Rev. Neurosci.* 24:363–77

Levi R. 2017. *A short course on algebraic topology geared towards applications to neuroscience*. Lecture, Kyoto University, July. **https://doi.org/10.13140/RG.2.2.29375.82082**

Li Y, Wang D, Ascoli G, Mitra P, Wang Y. 2017. Metrics for comparing neuronal tree shapes based on persistent homology. *PLOS ONE* 12(8):e0182184

Mischaikow K, Nanda V. 2013. Morse theory for filtrations and efficient computation of persistent homology. *Discrete Comput. Geom.* 50(2):330–53

Morozov D. 2023a. Dionysus. *Software*. **https://www.mrzv.org/software/dionysus/**

Morozov D. 2023b. Dionysus 2. *Software*. **https://mrzv.org/software/dionysus2/**

Moser EI. 2014. *Grid cells and the entorhinal map of space*. Nobel Prize Lecture Physiol. Med., Dec. 7. **https://www.nobelprize.org/uploads/2018/06/edvard-moser-lecture.pdf**

Munch E, Turner K, Bendich P, Mukherjee S, Mattingly J, Harer J. 2015. Probabilistic Fréchet means for time varying persistence diagrams. *Electron. J. Stat.* 9(1):1173–204

Munkres J. 2000. *Topology*. Prentice Hall. 2nd ed.

Niyogi P, Smale S, Weinberger S. 2008. Finding the homology of submanifolds with high confidence from random samples. *Discrete Comput. Geom.* 39:419–41

Noorman M, Hulse B, Jayaraman V, Romani S, Hermundstad A. 2024. Maintaining and updating accurate internal representations of continuous variables with a handful of neurons. *Nat. Neurosci.* 27:2207–17

Otter N, Porter M, Tillman U, Grindrod P, Harrington H. 2017. A roadmap for the computation of persistent homology. *EPJ Data Sci.* 6:17

Perea J, Scoccola L, Tralie C. 2023. DREiMac: dimensionality reduction with Eilenberg-MacLane coordinates. *J. Open Source Softw.* 8(91):5791

Petrucco L, Lavian H, Wu Y, Svara F, Stih V, Portugues R. 2023. Neural dynamics and architecture of the heading direction circuit in zebrafish. *Nat. Neurosci.* 26:765–73

Pilarczyk P, Mrozek M. Kalies W, Harker S. 2014. CHomP: computational homology project. *Software*. **https://chomp.rutgers.edu/**

Reyes A. 2021. Mathematical framework for place coding in the auditory system. *PLOS Comput. Biol.* 17(8):e1009251

Riemann M, Nolte M, Scolamiero M, Turner K, Perin R, et al. 2017. Cliques of neurons bound into cavities provide a missing link between structure and function. *Front. Comput. Neurosci.* 11:266051

Robins V. 1999. Towards computing homology from approximations. *Topol. Proc.* 24:503–32

Robins V. 2000. *Computational topology at multiple resolutions: foundations and applications to fractals and dynamics*. PhD Diss., University of Colorado at Boulder

Robins V. 2002. Computational topology for point data: Betti numbers of $\alpha$-shapes. In *Morphology of Condensed Matter*, ed. K Mecke, D Stoyan. Lecture Notes in Physics, Vol. 600. Springer. **https://doi.org/10.1007/3-540-45782-8_11**

Robins V, Meiss J, Bradley E. 1998. Computing connectedness: an exercise in computational topology. *Nonlinearity* 11(4):913

Robins V, Turner K. 2015. Principal component analysis of persistent homology rank functions with case studies of spatial point patterns, sphere packing and colloids. *Phys. D Nonlinear Phenom*. 334:99–117

Sanderson N. 2023. PyCliqueTop_2023. *Software*. **https://github.com/nerdnik/PyCliqueTop_2023**

Saul N. 2019. Persim: persistence landscapes and machine learning. *Software*. **https://persim.scikit-tda.org/en/latest/**

Schenck H. 2022. *Algebraic Foundations for Applied Topology and Data Analysis*. Springer

Sharpee T. 2019. An argument for hyperbolic geometry in neural circuits. *Curr. Opin. Neurobiol.* 58:101–4

Singh G, Memoli F, Ishkhanov T, Sapiro G, Carlsson G, Ringach D. 2008. Topological analysis of population activity in visual cortex. *J. Vis.* 8(8):11

Sizemore A, Giusti C, Bassett D. 2016. Classification of weighted networks through mesoscale homological features. *J. Complex Netw.* 5(2):245–73

Sizemore A, Giusti C, Kahn A, Vettel JM, Betzel RF, Bassett DS. 2018. Cliques and cavities in the human connectome. *J. Comput. Neurosci.* 44(1):115–45

Stolz B, Emerson T, Nakuri S, Porter M, Harrington H. 2021. Topological data analysis of task-based fMRI data from experiments on schizophrenia. *J. Phys. Complex.* 2(3):035006

Tausz A, Vejdemo-Johansson M, Adams H. 2014. JavaPlex. *Software*. **http://appliedtopology.github.io/javaplex/**

Tauzin G, Lupo U, Tunstall L, Burella Pérez J, Caorsi M, et al. 2021. giotto-tda: a topological data analysis toolkit for machine learning and data exploration. *J. Mach. Learn. Res.* 22(39):1–6

Theilman B, Perks K, Gentner T. 2021. Spike train coactivity encodes learned natural stimulus invariances in songbird auditory cortex. *J. Neurosci.* 41(1):73–88

Tralie C, Saul N, Bar-On R. 2018. Ripser.py: a lean persistent homology library for python. *J. Open Source Softw.* 3(29):925

Tudoras A, Reyes A. 2021. Topological conditions for propagation of spatially-distributed neural activity. Preprint, bioRxiv. **https://www.biorxiv.org/content/10.1101/2021.11.30.470616v1**

Turner K, Mukherjee S, Boyer D. 2014. Persistent homology transform for modeling shapes and surfaces. *Inform. Inference* 3(4):310–44

Vaupel M, Dunn B. 2023. The bifiltration of a relation and extended Dowker duality. Preprint, arXiv:2310.11529 [math.AT]

Waraich S, Victor J. 2024. The geometry of low- and high-level perceptual spaces. *J. Neurosci.* 44:e1460232023

Wu MC, Itskov V. 2022. A topological approach to inferring the intrinsic dimension of convex sensing data. *J. Appl. Comput. Topol.* 6(1):127–76

Yao J, Hagermann N, Xiong Q, Chen J, Hermann D, Chen C. 2024. Topological analysis of mouse brain vasculature via 3D light-sheet microscopy images. Preprint, arXiv:2402.16894 [q-bio.NC]

Yoon H, Ghrist R, Giusti C. 2023. Persistent extensions and analogous bars: data-induced relations between persistence barcodes. *J. Appl. Comput. Topol.* 7:571–617
Yoon I, Henselman-Petrusek G, Yu Y, Giusti C. 2024. Tracking the topology of neural manifolds across populations. *PNAS* 121(46):e2407997121
Zhang H, Rich P, Lee A, Sharpee T. 2023. Hippocampal spatial representations exhibit a hyperbolic geometry that expands with experience. *Nat. Neurosci.* 26:131–39
Zhou Y, Sharpee T. 2022. Using global t-SNE to preserve intercluster data structure. *Neural Comput.* 34:1637–51
Zhou Y, Smith B, Sharpee T. 2018. Hyperbolic geometry of the olfactory space. *Sci. Adv.* 4(8):eaaq1458
Zomorodian A. 2009. *Topology for Computing*. Cambridge Univ. Press
Zomorodian A. 2014. *A Short Course in Computational Geometry and Topology*. Springer
Zomorodian A, Carlsson G. 2005. Computing persistent homology. *Discrete Comput. Geom.* 33(2):249–74