WHAT CAN TOPOLOGY TELL US
ABOUT THE NEURAL CODE?

CARINA CURTO

ABSTRACT. Neuroscience is undergoing a period of rapid experimental progress and expansion. New mathematical tools, previously unknown in the neuroscience community, are now being used to tackle fundamental questions and analyze emerging data sets. Consistent with this trend, the last decade has seen an uptick in the use of topological ideas and methods in neuroscience. In this paper I will survey recent applications of topology in neuroscience, and explain why topology is an especially natural tool for understanding neural codes.

1. INTRODUCTION

Applications of topology to scientific domains outside of pure mathematics are becoming increasingly common. Neuroscience, a field undergoing a golden age of progress in its own right, is no exception. The first reason for this is perhaps obvious—at least to anyone familiar with topological data analysis. Like other areas of biology, neuroscience is generating a lot of new data, and some of these data can be better understood with the help of topological methods. A second reason is that a significant portion of neuroscience research involves studying networks, and networks are particularly amenable to topological tools. Although this paper will touch on a variety of such applications, most of my attention will be devoted to a third reason—namely, that many interesting problems in neuroscience contain topological questions in disguise. This is especially true when it comes to understanding *neural codes* and questions such as, How do the collective activities of neurons represent information about the outside world?

I will begin with some well-known examples of neural codes and then use them to illustrate how topological ideas naturally arise in this context. Next, I'll take a brief detour to describe other uses of topology in neuroscience. Finally, I will return to neural codes and explain why topological methods are helpful for studying their intrinsic properties. Taken together, these developments suggest that topology is not only useful for analyzing neuroscience data, but it may also play a fundamental role in the theory of how the brain works.

2. NEURONS: NODES IN A NETWORK OR AUTONOMOUS SENSORS?

It has been known for more than a century, since the time of Golgi and Ramon y Cajal, that the neurons in our brains are connected to each other in vast,

intricate networks. Neurons are electrically active cells. They communicate with each other by firing action potentials (spikes)—tiny messages that are only received by neighboring (synaptically connected) neurons in the network. Suppose we were eavesdropping on a single neuron, carefully recording its electrical activity at each point in time. What governs the neuron's behavior? The obvious answer—it's the network, of course! If we could monitor the activity of all the other neurons, and we knew exactly the pattern of connections between them, and were blessed with an excellent model describing all relevant dynamics, then (maybe?) we would be able to predict when our neuron will fire. If this seems hopeless now, imagine how unpredictable the activity of a single neuron in a large cortical network must have seemed in the 1950s when Hodgkin and Huxley had just finished working out the complex nonlinear dynamics of action potentials for a simple, isolated cell [32].

And yet, around 1959, a miracle happened. It started when Hubel and Wiesel inserted a microelectrode into the primary visual cortex of an anesthetized cat and eavesdropped on a single neuron. They could neither monitor nor control the activity of any other neurons in the network; they could only listen to one neuron at a time. What they *could* control was the visual stimulus. In an attempt to get the neuron to fire, they projected black and white patterns on a screen in front of the open-eyed cat. Remarkably, they found that the neuron they were listening to fired rapidly when the screen showed a black bar at a certain angle, say, 45°. Other neurons responded to different angles. It was as though each neuron was a sensor for a particular feature of the visual scene. Its activity could be predicted without knowing anything about the network, but by simply looking *outside* the cat's brain—at the stimulus on the screen.

Hubel and Wiesel had discovered orientation-tuned neurons [21], whose collective activity comprises a *neural code* for angles in the visual field (see Figure 1B). Although they inhabit a large, densely connected cortical network, these neurons do not behave as unpredictable units governed by complicated dynamics. Instead, they appear to be responding directly to stimuli in the outside world. Their activity has *meaning*.

A decade later, O'Keefe made a similar discovery, this time involving neurons in a different area of the brain—the hippocampus. Unlike the visual cortex, there is no obvious sensory pathway to the hippocampus. This made it all the more mysterious when O'Keefe reported that his neurons were responding selectively to different locations in the animal's physical environment [28]. These neurons, dubbed *place cells*, act as position sensors in space. When an animal is exploring a particular environment, a place cell increases its firing rate as the animal passes through its corresponding *place field*—that is, the localized region to which the neuron preferentially responds (see Figure 1C).

Like Hubel and Wiesel, who received a Nobel prize for their work in 1981 [1], O'Keefe's discovery of place cells had an enormous impact in neuroscience. In 2014, he shared the Nobel prize with Edvard and May-Britt Moser [5], former postdocs of his who went on to discover an even stranger class of neurons that encode position in a neighboring area of hippocampus called the entorhinal cortex. These neurons, called *grid cells*, display periodic place fields that are arranged in a hexagonal lattice. We'll come back to grid cells in the next section.
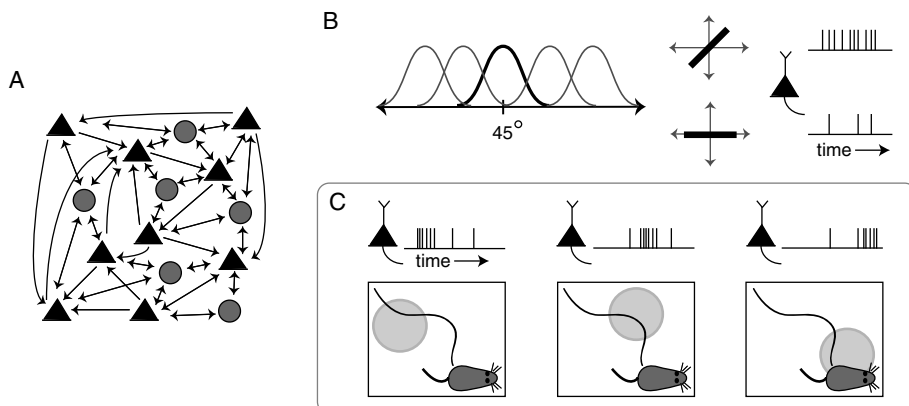
FIGURE 1. The neural network and neural coding pictures. (A) Pyramidal neurons (triangles) are embedded in a recurrent network together with inhibitory interneurons (circles). (B) An orientation-tuned neuron in primary visual cortex with a preferred angle of 45°. The neuron fires many spikes in response to a bar at a 45° angle in the animal's visual field, but few spikes in response to a horizontal bar. (C) Place cells in the hippocampus fire when the animal passes through the corresponding place field. The activity of three different neurons is shown (top), while the animal traces a trajectory starting at the top left corner of its environment (bottom). Each neuron's activity is highest when the animal passes through the corresponding place field (shaded disc).

So, are neurons nodes in a network or autonomous sensors of the outside world? Both pictures are valid, and yet they lead to very different models of neural behavior. Neural network theory deals with the first picture, and seeks to understand how the activity of neurons emerges from properties of the network. In contrast, neural coding theory often treats the network as a black box, focusing instead on the relationship between neural activity and external stimuli. Many of the most interesting problems in neuroscience are about *understanding the neural code.* This includes, but is not limited to, figuring out the basic principles by which neural activity represents sensory inputs to the eyes, nose, ears, whiskers, and tongue. Because of the discoveries of Hubel and Wiesel, O'Keefe, and many others, we often know more about the coding properties of single neurons than we do about the networks to which they belong. But many open questions remain. And topology, as it turns out, is a natural tool for understanding the neural code.

3. Topology of hippocampal place cell codes

The term *hippocampal place cell code* refers to the neural code used by place cells in the hippocampus to encode the animal's position in space. Most of the research about place cells, including O'Keefe's original discovery, has been performed in rodents (typically rats), and the experiments typically involve an animal moving around in a restricted environment (see Figure 1C). It was immediately understood that a population of place cells, each having a different place field, could collectivity
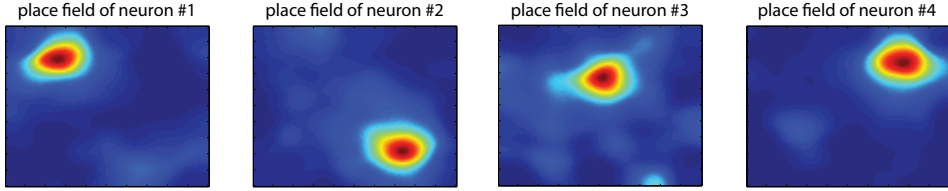
FIGURE 2. Place fields for four place cells, recorded while a rat explored a two-dimensional square box environment. Place fields were computed from data provided by the Pastalkova lab.

encode the animal's position in space [29], even though for a long time electrophysiologists could only monitor one neuron at a time. When simultaneous recordings of place cells became possible, it was shown via statistical inference (using previously measured place fields) that the animal's position could indeed be inferred from population place cell activity [3]. Figure 2 shows four place fields corresponding to simultaneously recorded place cells in area CA1 of rat hippocampus.

The role of topology in place cell codes begins with a simple observation, which is perhaps obvious to anyone familiar with both place fields in neuroscience and elementary topology. First, let's recall the standard definitions of an open cover and a good cover. A collection of open sets $\mathcal{U} = \{U_1, \ldots, U_n\}$ is an *open cover* of a topological space X if $X = \bigcup_{i=1}^{n} U_i$. We say that $\mathcal{U}$ is a *good cover* if every nonempty intersection $\bigcap_{i \in \sigma} U_i$, for $\sigma \subseteq \{1, \ldots, n\}$, is contractible. Now, observe that a collection of place fields in a particular environment looks strikingly like an open cover, with each U_i corresponding to a place field. Figure 3 displays three different environments, typical of what is used in hippocampal experiments with rodents, together with schematic arrangements of place fields in each.

Moreover, since place fields are approximately convex (see Figure 2) it is not unreasonable to assume that they form a good cover of the underlying space. This means the Nerve Lemma applies. Recall the notion of the *nerve*[1] of a cover:

$$\mathcal{N}(\mathcal{U}) \overset{\text{def}}{=} \{\sigma \subset [n] \mid \bigcap_{i \in \sigma} U_i \neq \emptyset\},$$
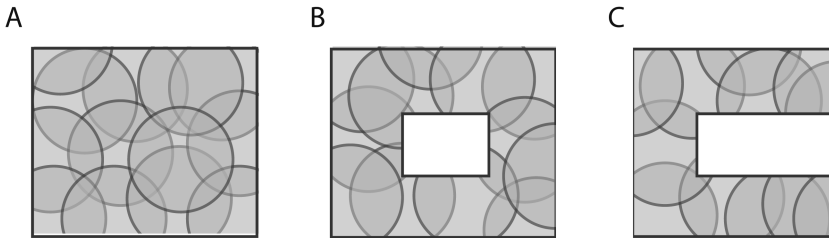

FIGURE 3. Three environments for a rat: (A) a square box environment, also known as an "open field"; (B) an environment with a hole or obstacle in the center; and (C) a maze with two arms. Each environment displays a collection of place fields (shaded discs) that fully cover the underlying space.

---

[1] Note that the name "nerve" here predated any connection to neuroscience!

where $[n] = \{1, \ldots, n\}$. Clearly, if $\sigma \in \mathcal{N}(\mathcal{U})$ and $\tau \subset \sigma$, then $\tau \in \mathcal{N}(\mathcal{U})$. This property shows that $\mathcal{N}(\mathcal{U})$ is an abstract *simplicial complex* on the vertex set $[n]$— that is, it is a set of subsets of $[n]$ that is closed under taking further subsets. If X is a sufficiently "nice" topological space, then the following well-known lemma holds.

Lemma 3.1 (Nerve Lemma). *Let $\mathcal{U}$ be a good cover of X. Then $\mathcal{N}(\mathcal{U})$ is homotopy-equivalent to X. In particular, $\mathcal{N}(\mathcal{U})$ and X have exactly the same homology groups.*[2]

It is important to note that the Nerve Lemma fails if the good cover assumption does not hold. Figure 4A depicts a good cover of an annulus by three open sets. The corresponding nerve (right) exhibits the topology of a circle, which is indeed homotopy-equivalent to the covered space. In Figure 4B, however, the cover is *not* good, because the intersection $U_1 \cap U_2$ consists of two disconnected components and is thus not contractible. Here the nerve (right) is homotopy-equivalent to a point, in contradiction to the topology of the covered annulus.
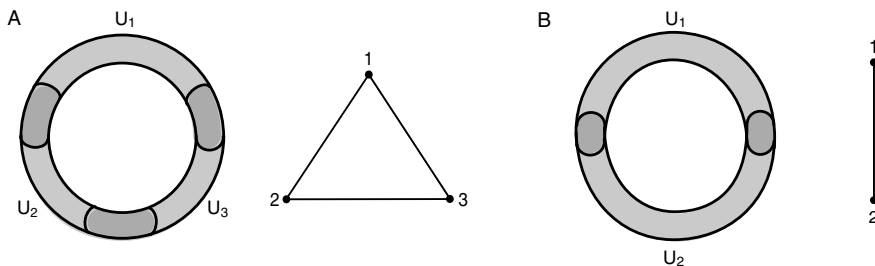

FIGURE 4. Good and bad covers.   (A) A good cover $\mathcal{U} = \{U_1, U_2, U_3\}$ of an annulus (left), and the corresponding nerve $\mathcal{N}(\mathcal{U})$ (right). (B) A "bad" cover of the annulus (left), and the corresponding nerve (right). Only the nerve of the good cover accurately reflects the topology of the annulus.

The wonderful thing about the Nerve Lemma, when interpreted in the context of hippocampal place cells, is that $\mathcal{N}(\mathcal{U})$ can be inferred from the activity of place cells alone—without actually knowing the place fields $\{U_i\}$. This is because the concurrent activity of a group of place cells, indexed by $\sigma \subset [n]$, indicates that the corresponding place fields have a nonempty intersection, $\bigcap_{i \in \sigma} U_i \neq \emptyset$. In other words, if we were eavesdropping on the activity of a population of place cells as the animal fully explored its environment, then by finding which subsets of neurons co-fire (see Figure 5) we could, in principle, estimate $\mathcal{N}(\mathcal{U})$, even if the place fields themselves were unknown. Lemma 3.1 tells us that the homology of the simplicial complex $\mathcal{N}(\mathcal{U})$ precisely matches the homology of the environment X. The place cell code thus naturally reflects the topology of the represented space.[3]

---

[2]If we are only interested in matching homology, the good cover assumption can be relaxed [15].

[3]In particular, place cell activity from the environment in Figure 3B could be used to detect the nontrivial first homology group of the underlying space and, thus, distinguish this environment from that of Figure 3A or 3C.
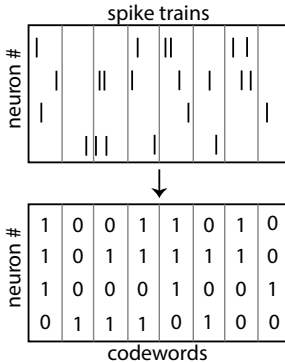
spike trains

1	0	0	1	1	0	1	0
1	0	1	1	1	1	1	0
1	0	0	0	1	0	0	1
0	1	1	1	0	1	0	0

codewords

FIGURE 5. By binning spike trains for a population of simulta-neously recorded neurons, one can infer subsets of neurons that co-fire. If these neurons were place cells, then the first codeword 1110 indicates that $U_1 \cap U_2 \cap U_3 \neq \emptyset$, while the third codeword 0101 tells us $U_2 \cap U_4 \neq \emptyset$.

These and related observations have led some researchers to speculate that the hippocampal place cell code is fundamentally topological in nature [6, 14], while others (including this author) have argued that considerable geometric information is also present and can be extracted using topological methods [11, 20]. In order to disambiguate topological and geometric features, Dabaghian et al. performed an elegant experiment using linear tracks with flexible joints [13]. This allowed them to alter geometric features of the environment while preserving the topological structure as reflected by the animal's place fields. They found that place fields recorded from an animal running along the morphing track moved together with the track, preserving the relative sequence of locations despite changes in angles and movement direction. In other words, the place fields respected topological aspects of the environment more than metric features [13].

What about the entorhinal grid cells? These neurons have firing fields with multiple disconnected components, forming a hexagonal grid (see Figure 6A). This means that grid fields violate the good cover assumption of the Nerve Lemma—if we consider them as an open cover for the entire two-dimensional environment. If, instead, we restrict our attention to a fundamental domain for these firing fields, as illustrated in Figure 6B, then each grid field has just one (convex) component, and the Nerve Lemma applies. From the spiking activity of grid cells we could thus infer the topology of this fundamental domain. The reader familiar with the topological classification of surfaces may recognize that this hexagonal domain, with the identification of opposite edges, is precisely a torus. This indicates that the space represented by grid cells is not the full environment, but a torus.

4. TOPOLOGY IN NEUROSCIENCE: A BIRD'S-EYE VIEW

The examples from the previous section are by no means the only way that topology is being used in neuroscience. Before plunging into further details about what topology can tell us about neural codes, we now pause for a moment to acknowledge some other interesting applications. The main thing they all have in common is their recency. This is no doubt due to the rise of computational and
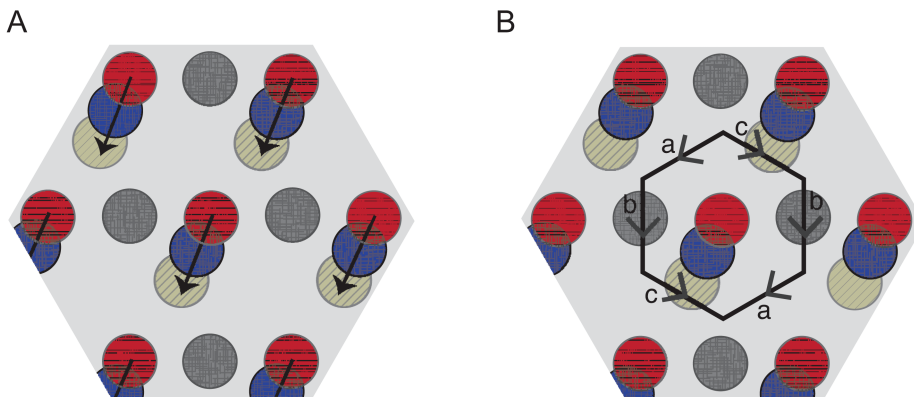
A
B

FIGURE 6. Firing fields for grid cells. (A) Firing fields for four entorhinal grid cells, each in a different color. A single grid field consists of multiple disconnected regions and forms a hexagonal grid in the animal's two-dimensional environment. (B) A hexagonal fundamental domain contains just one disc-like region per grid cell. Pairs of edges with the same label (a, b, or c) are identified with orientations specified by the arrows.

applied algebraic topology, a relatively new development in applied mathematics that was highlighted in the *Bulletin of the American Mathematical Society* nearly a decade ago [17].

Roughly speaking, the uses of topology in neuroscience can be categorized into three (overlapping) themes: (i) "traditional" topological data analysis applied to neuroscience; (ii) an upgrade to network science; and (iii) understanding the neural code. Here we briefly summarize work belonging to (i) and (ii). In the next section we'll return to (iii), which is the main focus of this talk.

4.1. "Traditional" TDA applied to neuroscience data sets. The earliest and most familiar applications of topological data analysis (TDA) focused on the problem of estimating the "shape" of point-cloud data. This kind of data set is simply a collection of points, $x_1, \ldots, x_\ell \in \mathbb{R}^n$, where n is the dimensionality of the data. A question one could ask is, Do these points appear to have been sampled from a lower-dimensional manifold, such as a torus or a sphere?

The basic strategy is to consider open balls $B_\varepsilon(x_i)$ of radius ε around each data point, and then to construct a simplicial complex $\mathcal{K}_\varepsilon$ that captures information about how the balls intersect. This simplicial complex can either be the *Čech complex*, which is the nerve of the open cover defined by the balls, or the *Vietoris–Rips complex*, which is the clique complex[4] of the graph obtained from pairwise intersections of the balls. By varying ε, one obtains a sequence of nested simplicial complexes $\{\mathcal{K}_\varepsilon\}$ together with natural inclusion maps. *Persistent homology* tracks homology cycles across these simplicial complexes, and allows one to determine

---

[4] Recall that the *clique complex* of a graph G, denoted $X(G)$, is simply the collection of all cliques (all-to-all connected subgraphs) in G, viewed as a simplicial complex:

$$X(G) = \{\sigma \subset [n] \mid (ij) \in G \text{ for all } i, j \in \sigma\}.$$

whether there were homology classes that "persisted" for a long time. For example, if the data points were sampled from a 3-sphere, one would see a persistent 3-cycle.

There are many excellent reviews of persistent homology, including [17], so I will not go into further detail here. Instead, it is interesting to note that one of the early applications of these techniques was in neuroscience—to analyze population activity in the primary visual cortex [33]. Here it was found that the topological structure of activity patterns is similar between spontaneous and evoked activity, and is consistent with the topology of a 2-sphere. Moreover, the results of this analysis were interpreted in the context of neural coding, making this work exemplary of both themes (i) and (iii). Another application of persistent homology to point cloud data in neuroscience was the analysis of the spatial structure of afferent neuron terminals in crickets [4]. Again, the results were interpreted in terms of the coding properties of the corresponding neurons, which are sensitive to air motion detected by thin mechanosensory hairs on the cricket. Finally, it is worth mentioning that these types of analyses are not confined to neural activity. For example, in [2] the statistics of persistent cycles were used to study brain artery trees.

4.2. **An upgrade to network science.** There are many ways of constructing networks in neuroscience, but the basic model that has been used for all of them is the graph. The vertices of a graph can represent neurons, cell types, brain regions, or fMRI voxels, while the edges reflect interactions between these units. Often, the graph is weighted and the edge weights correspond to correlations between adjacent nodes. For example, one can model a functional brain network from fMRI data as a weighted graph where the edge weights correspond to activity correlations between pairs of voxels. At the other extreme, a network where the vertices correspond to neurons could have edge weights that reflect either pairwise correlations in neural activity or synaptic connections.

Network science is a relatively young discipline that focuses on analyzing networks, primarily using tools derived from graph theory. The results of a particular analysis could range from determining the structure of a network to identifying important subgraphs and/or graph-theoretic statistics (the distribution of in-degree or out-degree across nodes, number of cycles, etc.) that carry meaning for the network at hand. Sometimes, graph-theoretic features do not carry obvious meaning, but are nevertheless useful for distinguishing networks that belong to distinct classes. For example, a feature could be characteristic of functional brain networks derived from a subgroup of subjects, distinguishing them from a "control" group. In this way graph features may be a useful diagnostic tool for distinguishing diseased states, pharmacologically induced states, cognitive abilities, or for uncovering systematic differences based on gender or age.

The recent emergence of topological methods in network science stems from the following "upgrade" to the network model: instead of a graph, one considers a simplicial complex. Sometimes this simplicial complex reflects higher-order interactions that are obtained from the data, and sometimes it is just the clique complex of the graph G (see footnote 4); in other words, the higher-order simplices correspond to cliques of G. Figure 7A shows a graph (top) and the corresponding clique complex (bottom), with shaded simplices corresponding to two 3-cliques and a 4-clique. The clique complex fills in many of the 1-cycles in the original graph, but some 1-cycles remain (see the gold 4-gon), and higher-dimensional cycles may emerge.
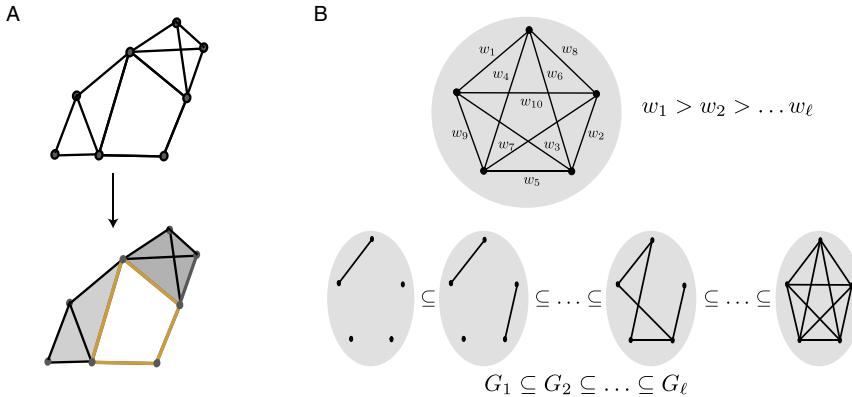
FIGURE 7. Network science models: from graphs to clique complexes and filtrations.

Computing homology groups of a clique complex $X(G)$ is then a natural way to detect topological features that are determined by the underlying graph G. In the case of a *weighted* graph (the weights often represent interaction strengths in a network), one can obtain a sequence of clique complexes by considering a related sequence of simple graphs, $\{G_i\}$, where each graph is obtained from the previous one by adding the edge corresponding to the next-highest weight (see Figure 7B). The corresponding nested sequence of clique complexes,

$$X(G_1) \subseteq X(G_2) \subseteq \ldots \subseteq X(G_\ell),$$

can then be analyzed using persistent homology.

A major advantage of this approach for analyzing weighted graphs/networks is that any topological invariant of $\{X(G_i)\}$ is automatically an invariant of the underlying weighted graph under nonlinear transformations of the weights $\{w_{ij}\}$, provided that the total ordering of weights is preserved. This feature is particularly important in neuroscience applications, because interaction strengths are often measured in somewhat arbitrary units (such as image brightness or signal strength). While one can usually assume a larger weight means a stronger interaction, the precise values of the weights may not be meaningful. Topological tools discard this information, while revealing a surprising amount of biologically relevant network structure. In particular, the statistics of *Betti curves* have been used to infer hidden geometric structure in the correlations of hippocampal neurons [20].

For a more thorough survey of topological methods in network science, I recommend the recent review article [19]. Here I will only mention that topological network analyses have already been used in a variety of neuroscience applications, many of them medically motivated: fMRI networks in patients with ADHD [16]; FDG-PET based networks in children with autism and ADHD [25]; morphological networks in deaf adults [24]; metabolic connectivity in epileptic rats [7]; and functional EEG connections in depressed mice [23]. Other applications to fMRI data include human brain networks during learning [35] and drug-induced states [30]. At a finer scale, recordings of neural activity can also give rise to functional connectivity networks among neurons (which are not the same as the neural networks defined by synaptic connections). These networks have also been analyzed with topological methods [20, 31, 34].

5. The code of an open cover, and its algebraic description

We now return to neural codes. We have already seen how the hippocampal place cell code reflects the topology of the underlying space, via the nerve $\mathcal{N}(\mathcal{U})$ of a place field cover. In this section, we will associate a binary code to an open cover. This notion is closer in spirit to a combinatorial neural code (see Figure 5), and carries more detailed information than the nerve. There is also a nice algebraic description of these codes, which keeps track of the essential features of the cover. In the next section, we'll see how topology is being used to determine intrinsic properties of neural codes, such as convexity and dimension.

First, a few definitions. A *binary pattern* on n neurons is a string of 0's and 1's, with a 1 for each active neuron and a 0 denoting silence; equivalently, it is a subset of (active) neurons $\sigma \subset [n]$. (Recall that $[n] = \{1, \ldots, n\}$.) We use both notations interchangeably. For example, 10110 and $\sigma = \{1, 3, 4\}$ refer to the same pattern, or codeword, on $n = 5$ neurons. A *combinatorial neural code* on n neurons is a collection of binary patterns $\mathcal{C} \subset 2^{[n]}$. In other words, it is a binary code of length n, where we interpret each binary digit as the "on" or "off" state of a neuron. The *simplicial complex of a code*, $\Delta(\mathcal{C})$, is the smallest abstract simplicial complex on $[n]$ that contains all elements of $\mathcal{C}$. In keeping with the hippocampal place cell example, we are interested in codes that correspond to open covers of some topological space.

Definition 5.1. Given an open cover $\mathcal{U}$, the *code of the cover* is the binary code

$$\mathcal{C}(\mathcal{U}) \overset{\text{def}}{=} \{\sigma \subset [n] \mid \bigcap_{i \in \sigma} U_i \setminus \bigcup_{j \in [n] \setminus \sigma} U_j \neq \emptyset\}.$$

Each codeword in $\mathcal{C}(\mathcal{U})$ corresponds to a region that is defined by the intersections of the open sets in $\mathcal{U}$ (Figure 8A). Note that the code $\mathcal{C}(\mathcal{U})$ is not the same as the nerve $\mathcal{N}(\mathcal{U})$. Figures 8B and 8C display the code and the nerve of the open cover in Figure 8A. While the nerve encodes which subsets of the U_i's have nonempty intersections, the code also carries information about set containments. For example, $U_2 \subseteq U_1 \cup U_3$ can be inferred from $\mathcal{C}(\mathcal{U})$ because each codeword of the form $*1**$ has an additional 1 in position 1 or 3, indicating that if neuron 2 is firing, then so is neuron 1 or 3. Similarly, $U_2 \cap U_4 \subseteq U_3$ can be inferred from the code because any word of the form $*1*1$ necessarily has a 1 in position 3 as
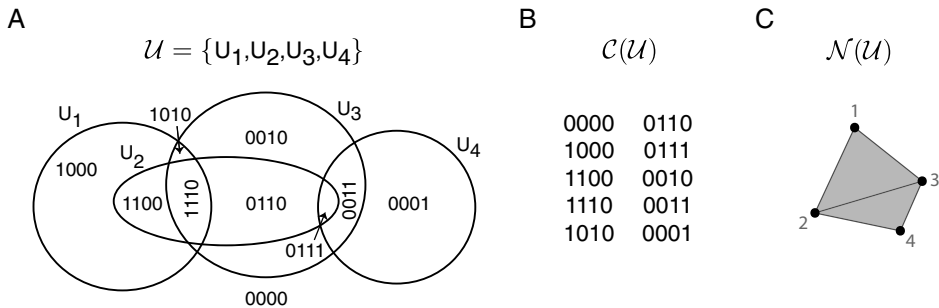

FIGURE 8. Codes and nerves of open covers. (A) An open cover $\mathcal{U}$, with each region carved out by the cover labeled by its corresponding codeword. (B) The code $\mathcal{C}(\mathcal{U})$. (C) The nerve $\mathcal{N}(\mathcal{U})$.

well. These containment relationships go beyond simple intersection data and cannot be obtained from the nerve $\mathcal{N}(\mathcal{U})$. On the other hand, the nerve can easily be recovered from the code since $\mathcal{N}(\mathcal{U})$ is the smallest simplicial complex that contains it—that is,

$$\mathcal{N}(\mathcal{U}) = \Delta(\mathcal{C}(\mathcal{U})).$$

$\mathcal{C}(\mathcal{U})$ thus carries more detailed information than what is available in $\mathcal{N}(\mathcal{U})$.

The combinatorial data in $\mathcal{C}(\mathcal{U})$ can be encoded algebraically via the *neural ideal* [12], much as simplicial complexes are algebraically encoded by Stanley–Reisner ideals [27]. In order to represent a code algebraically, it is useful to consider binary patterns of length n as elements of $\mathbb{F}_2^n$, where $\mathbb{F}_2$ is the finite field of two elements: 0 and 1. Given a code $\mathcal{C} \subseteq \mathbb{F}_2^n$, it is natural to consider the ideal of polynomials with coefficients in $\mathbb{F}_2$ that vanish on all codewords:

$$I_\mathcal{C} \stackrel{\text{def}}{=} \{f \in \mathbb{F}_2[x_1, \ldots, x_n] \mid f(c) = 0 \text{ for all } c \in \mathcal{C}\}.$$

Note that a polynomial $f \in \mathbb{F}_2[x_1, \ldots, x_n]$ can be evaluated on a binary pattern of length n by simply replacing each indeterminate x_i with the 0/1 value of the ith neuron. For example, if $f = x_1 x_3 (1 - x_4) \in \mathbb{F}_2[x_1, \ldots, x_4]$, then $f(1010) = 1$ and $f(1100) = 0$.

The ideal $I_\mathcal{C}$ and the closely related neural ideal $J_\mathcal{C}$ [5] algebraically package the combinatorial data of a code in a way that makes the set-theoretic relationships of corresponding open covers easy to extract.

Lemma 5.2 ([12]). *Let $\mathcal{C}$ be a binary code of length n (i.e., on n neurons), and let $\mathcal{U} = \{U_1, \ldots, U_n\}$ be any collection of open sets such that $\mathcal{C} = \mathcal{C}(\mathcal{U})$. For any $\sigma, \tau \subseteq [n]$,*

$$\prod_{i \in \sigma} x_i \prod_{j \in \tau} (1 - x_j) \in I_\mathcal{C} \quad \Leftrightarrow \quad \bigcap_{i \in \sigma} U_i \subseteq \bigcup_{j \in \tau} U_j.$$

Note that it is easy to see that any binary code can be realized as the code of an open cover.[6] It is not true, however, that any code can arise from a good cover or a *convex cover*—that is, an open cover consisting of convex open sets. The algebraic framework of the neural ideal can be used to determine, in many cases, whether or not a given code has a convex realization [10].

6. Using topology to study intrinsic properties of neural codes

In our previous examples from neuroscience, the place cell and grid cell codes can be thought of as arising from convex sets covering an underlying space. Because the spatial correlates of these neurons are already known, it is not difficult to infer what space is being represented by these codes. What could we say if we were

---

[5]In fact, $I_\mathcal{C}$ can be expressed as the sum $I_\mathcal{C} = J_\mathcal{C} + \mathcal{B}$, where $\mathcal{B} = \langle x_1^2 - x_1, \ldots, x_n^2 - x_n \rangle$ and $J_\mathcal{C} \stackrel{\text{def}}{=} \langle \chi_v \mid v \in \mathbb{F}_2^n \setminus \mathcal{C} \rangle$ is the *neural ideal*, generated by the characteristic functions

$$\chi_v \stackrel{\text{def}}{=} \prod_{\{i \mid v_i = 1\}} x_i \prod_{\{j \mid v_j = 0\}} (1 - x_j),$$

corresponding to noncodewords [12].

[6]For example, if the size of the code is $|\mathcal{C}| = \ell$, we could choose disjoint open intervals $B_1, \ldots, B_\ell \subset \mathbb{R}$, one for each codeword, and define the open sets $U_1, \ldots, U_n$ such that U_i is the union of all open intervals B_j corresponding to codewords in which neuron i is "on" (that is, there is a 1 in position i of the codeword).

given just a code, $\mathcal{C} \subset \{0,1\}^n$, without *a priori* knowledge of what the neurons were encoding? Could we tell whether such a code can be realized via a convex cover?

6.1. **What can go wrong.** The following lemma illustrates the simplest example of what can go wrong.

Lemma 6.1. *Let $\mathcal{C} \subset \{0,1\}^3$ be a code that contains the codewords* 110 *and* 101 *but does <u>not</u> contain* 100 *and* 111. *Then $\mathcal{C}$ is not the code of a good or convex cover.*

To see why the lemma holds, suppose $\mathcal{U} = \{U_1, U_2, U_3\}$ is a cover such that $\mathcal{C} = \mathcal{C}(\mathcal{U})$. Because neuron 2 or 3 is "on" in any codeword for which neuron 1 is "on", we must have that $U_1 \subset U_2 \cup U_3$. Moreover, we see from the code that $U_1 \cap U_2 \neq \emptyset$ and $U_1 \cap U_3 \neq \emptyset$, while $U_1 \cap U_2 \cap U_3 = \emptyset$. This means we can write U_1 as a disjoint union of two nonempty sets, $U_1 \cap U_2$ and $U_1 \cap U_3$, showing that U_1 is disconnected, and thus $\mathcal{U}$ can be neither a good nor a convex cover.

This simple example shows that not all codes can arise from convex covers. Moreover, the problem that prevents the code in Lemma 6.1 from being convex is topological in nature. Specifically, the code dictates that there must be a set containment,

$$U_\sigma \subseteq \bigcup_{j \in \tau} U_j,$$

where $U_\sigma = \bigcap_{i \in \sigma} U_i$, but the nerve of the resulting cover of U_σ by the sets $\{U_\sigma \cap U_j\}_{j \in \tau}$ is not contractible. This leads to a contradiction if the sets U_i are all assumed to be convex, because the sets $\{U_\sigma \cap U_j\}_{j \in \tau}$ are then also convex and thus form a good cover of U_σ (see Figure 9). Since U_σ itself is convex, it follows from the Nerve Lemma that $\mathcal{N}(\{U_\sigma \cap U_j\}_{j \in \tau})$ must be contractible, contradicting the data of the code.

These observations lead to the notion of a *local obstruction* to convexity [18], which captures the topological problem that arises if certain codes are assumed to have convex covers. The proof of the following lemma is essentially the argument outlined above.
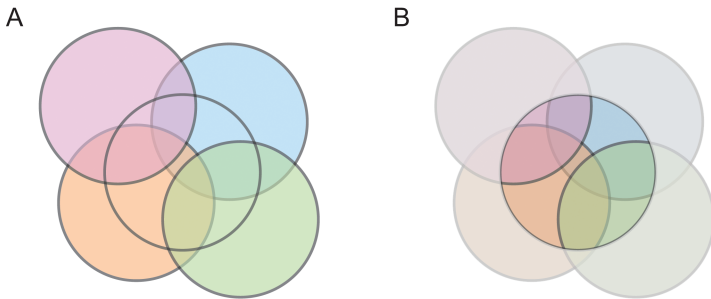

FIGURE 9. (A) A convex cover $\{U_i\}_{i=1}^5$. U_5 is the middle set, covered by the others. (B) The convex cover restricted to U_σ, for $\sigma = \{5\}$. Note that the restriction of a convex cover to a covered intersection, $U_\sigma \subseteq \bigcup_{j \in \tau} U_j$, yields another convex cover, $\{U_\sigma \cap U_j\}_{j \in \tau}$. (In contrast, restricting a good cover to a covered intersection does not guarantee a new good cover.)

Lemma 6.2 ([18])**.** *If $\mathcal{C}$ can be realized by a convex cover, then $\mathcal{C}$ has no local obstructions.*

The idea of using local obstructions to determine whether or not a neural code has a convex realization has been recently followed up in a series of papers [8–10, 26]. In particular, local obstructions have been characterized in terms of links,

$$\mathrm{Lk}_\Delta(\sigma) \overset{\text{def}}{=} \{\omega \in \Delta \mid \sigma \cap \omega = \emptyset \text{ and } \sigma \cup \omega \in \Delta\}.$$

The following theorem implies that convex codes must include as codewords any $\sigma \in \Delta(\mathcal{C})$ for which $\mathrm{Lk}_\Delta(\sigma)$ is not contractible.

Theorem 6.3 ([9])**.** *Let $\mathcal{C}$ be a neural code, and let $\Delta = \Delta(\mathcal{C})$. Then $\mathcal{C}$ has no local obstructions if and only if $\mathrm{Lk}_\Delta(\sigma)$ is contractible for all $\sigma \in \Delta \setminus \mathcal{C}$.*

A nice consequence of this result is that for any simplicial complex Δ, one can use a minimal free resolution of the Stanley–Reisner ideal I_Δ, together with Hochster's formula from commutative algebra, to compute a set of "mandatory" codewords. These codewords can be read off of the Betti numbers of the free resolution, and they correspond to faces of Δ that must be included in any convex code $\mathcal{C}$ whose simplicial complex is Δ [9].

Unfortunately, a code with no local obstructions need not be a convex code— that is, it may not be realizable as the code of a convex open cover. In [26], the following counterexample to the converse of Lemma 6.2 was discovered.

Example 6.4 ([26])**.** The code $\mathcal{C} = \{2345, 123, 134, 145, 13, 14, 23, 34, 45, 3, 4\}$ is not a convex code, despite the fact that it has no local obstructions.

That this code has no local obstructions can be easily seen using Theorem 6.3. The fact that there is no convex open cover, however, relies on geometric arguments that are not obviously topological. Moreover, this code does have a good cover [26], suggesting the existence of a new class of obstructions to convexity which may or may not be topological in nature.

6.2. **What can go right.** Finally, it has been shown that several classes of neural codes are guaranteed to have convex realizations. *Intersection-complete codes* satisfy the property that for any $\sigma, \tau \in \mathcal{C}$, the intersection $\sigma \cap \tau \in \mathcal{C}$. These codes, as well as the more general class of *max intersection-complete codes*, were shown constructively to have convex covers in [8]. Additional classes of codes with convex realizations have been described in [9].

Despite these developments, a complete characterization of convex codes is still lacking. Finding the minimum dimension needed for a convex realization is also an open question.

7. Codes from networks

We end by coming back to the beginning. Even if neural codes give us the illusion that neurons in cortical and hippocampal areas are directly sensing the outside world, we know that of course they are not. Their activity patterns are shaped by the networks in which they reside. What can we learn about the architecture of a network by studying its neural code? This question requires an improved understanding of neural networks, not just neural codes. While many candidate architectures have been proposed to explain, say, orientation-tuning in the visual cortex, the interplay of neural network theory and neural coding is still in early stages of development.
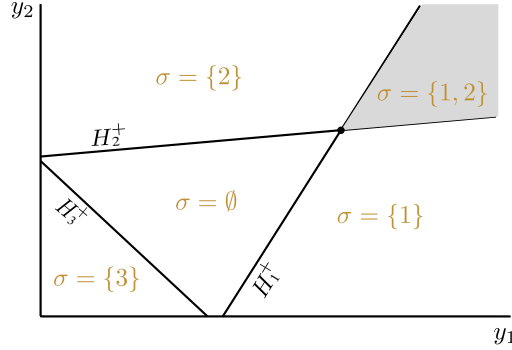
FIGURE 10. A hyperplane arrangement in the positive orthant, and the corresponding feedforward code.

Perhaps the simplest example of how the structure of a network can constrain the neural code is the case of simple feedforward networks. These networks have a single input layer of neurons and a single output layer. The resulting codes are derived from hyperplane arrangements in the positive orthant of $\mathbb{R}^k$, where k is the number of neurons in the input layer and each hyperplane corresponds to a neuron in the output layer (see Figure 10). Every codeword in a *feedforward code* corresponds to a chamber in such a hyperplane arrangement.

It is not difficult to see from this picture that all feedforward codes are realizable by convex covers—specifically, they arise from overlapping half-spaces [18]. On the other hand, not every convex code is the code of a feedforward network [22]. Moreover, the discrepancy between feedforward codes and convex codes is not due to restrictions on their simplicial complexes. As was shown in [18], every simplicial complex can arise as $\Delta(\mathcal{C})$ for a feedforward code. As with convex codes, a complete characterization of feedforward codes is still unknown. It seems clear, however, that topological tools will play an essential role.

ACKNOWLEDGMENTS

ABOUT THE AUTHOR

Carina Curto is associate professor of mathematics at the Pennsylvania State University. Her research interests lie at the interface between neuroscience and mathematics.

REFERENCES

[1] *Physiology or Medicine 1981—Press Release*, 2014, `Nobelprize.org` Nobel Media AB. http://www.nobelprize.org/nobel_prizes/medicine/laureates/1981/press.html.

[2] P. Bendich, J. S. Marron, E. Miller, A. Pieloch, and S. Skwerer, *Persistent homology analysis of brain artery trees*, Ann. Appl. Stat. **10** (2016), no. 1, 198–218, DOI 10.1214/15-AOAS886. MR3480493

[3] E. N. Brown, L. M. Frank, D. Tang, M. C. Quirk, and M. A. Wilson, *A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells*, J. Neurosci. **18** (1998), 7411–7425.

[4] J. Brown and T. Gedeon, *Structure of the afferent terminals in terminal ganglion of a cricket and persistent homology*, PLoS ONE **7** (2012), no. 5.

[5] N. Burgess, *The 2014 Nobel Prize in Physiology or Medicine: A Spatial Model for Cognitive Neuroscience*, Neuron **84** (2014), no. 6, 1120–1125.

[6] Z. Chen, S. N. Gomperts, J. Yamamoto, and M. A. Wilson, *Neural representation of spatial topology in the rodent hippocampus*, Neural Comput. **26** (2014), no. 1, 1–39, DOI 10.1162/NECO_a_00538. MR3155578

[7] H. Choi, Y. K. Kim, H. Kang, H. Lee, H.-J. Im, E. Edmund Kim, J.-K. Chung, D. S. Lee, et al., *Abnormal metabolic connectivity in the pilocarpine-induced epilepsy rat model: a multiscale network analysis based on persistent homology*, NeuroImage **99** (2014), 226–236.

[8] J. Cruz, C. Giusti, V. Itskov, and W. Kronholm, *On open and closed convex codes*, `arXiv: 1609.03502v1[math.CO]`, 2016.

[9] C. Curto, E. Gross, J. Jeffries, K. Morrison, M. Omar, Z. Rosen, A. Shiu, and N. Youngs, *What makes a neural code convex?*, Available online at http://arxiv.org/abs/1508.00150, 2016.

[10] C. Curto, E. Gross, J. Jeffries, K. Morrison, Z. Rosen, A. Shiu, and N. Youngs, *Algebraic signatures of convex and non-convex codes*, In preparation, 2016.

[11] C. Curto and V. Itskov, *Cell groups reveal structure of stimulus space*, PLoS Comput. Biol. **4** (2008), no. 10, e1000205, 13, DOI 10.1371/journal.pcbi.1000205. MR2457124

[12] C. Curto, V. Itskov, A. Veliz-Cuba, and N. Youngs, *The neural ring: an algebraic tool for analyzing the intrinsic structure of neural codes*, Bull. Math. Biol. **75** (2013), no. 9, 1571–1611, DOI 10.1007/s11538-013-9860-3. MR3105524

[13] Y. Dabaghian, V. L. Brandt, and L. M. Frank, *Reconceiving the hippocampal map as a topological template*, Elife **3** (2014), e03476.

[14] Y. Dabaghian, F. Mémoli, L. Frank, and G. Carlsson, *A topological paradigm for hippocampal spatial map formation using persistent homology*, PLoS Comp. Bio. **8** (2012), no. 8, e1002581.

[15] E. Colin de Verdiere, G. Ginot, and X. Goaoc, *Multinerves and Helly Numbers of Acyclic Families*, Symposium on Computational Geometry - SoCG '12 (2012).

[16] S. P. Ellis and A. Klein, *Describing high-order statistical dependence using "concurrence topology," with application to functional MRI brain data*, Homology Homotopy Appl. **16** (2014), no. 1, 245–264, DOI 10.4310/HHA.2014.v16.n1.a14. MR3211745

[17] R. Ghrist, *Barcodes: the persistent topology of data*, Bull. Amer. Math. Soc. (N.S.) **45** (2008), no. 1, 61–75, DOI 10.1090/S0273-0979-07-01191-3. MR2358377

[18] C. Giusti and V. Itskov, *A no-go theorem for one-layer feedforward networks*, Neural Comput. **26** (2014), no. 11, 2527–2540, DOI 10.1162/NECO_a_00657. MR3243436

[19] C. Giusti, R. Ghrist, and D. S. Bassett, *Two's company, three (or more) is a simplex*, J. Comput. Neurosci. **41** (2016), no. 1, 1–14, DOI 10.1007/s10827-016-0608-6. MR3517602

[20] C. Giusti, E. Pastalkova, C. Curto, and V. Itskov, *Clique topology reveals intrinsic geometric structure in neural correlations*, Proc. Natl. Acad. Sci. USA **112** (2015), no. 44, 13455–13460, DOI 10.1073/pnas.1506407112. MR3429279

[21] D. H. Hubel and T. N. Wiesel, *Receptive fields of single neurons in the cat's striate cortex*, J. Physiol. **148** (1959), no. 3, 574–591.

[22] V. Itskov, Personal communication, 2015.

[23] A. Khalid, B. S. Kim, M. K. Chung, J. C. Ye, and D. Jeon, *Tracing the evolution of multiscale functional networks in a mouse model of depression using persistent brain network homology*, Neuroimage **101** (2014), 351–363.

[24] E. Kim, H. Kang, H. Lee, H.-J. Lee, M.-W. Suh, J.-J. Song, S.-H. Oh, and D. S. Lee, *Morphological brain network assessed using graph theory and network filtration in deaf adults*, Hear. Res. **315** (2014), 88–98.

[25] H. Lee, M. K. Chung, H. Kang, B.-N. Kim, and D. S. Lee, *Discriminative persistent homology of brain networks*, Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on, IEEE, 2011, pp. 841–844.

[26] C. Lienkaemper, A. Shiu, and Z. Woodstock, *Obstructions to convexity in neural codes*, Available online at http://arxiv.org/abs/1509.03328.

[27] E. Miller and B. Sturmfels, *Combinatorial commutative algebra*, Graduate Texts in Mathematics, vol. 227, Springer-Verlag, New York, 2005. MR2110098

[28] J. O'Keefe and J. Dostrovsky, *The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat*, Brain Res. **34** (1971), no. 1, 171–175.

[29] J. O'Keefe and L. Nadel, *The hippocampus as a cognitive map*, Clarendon Press Oxford, 1978.

[30] G. Petri, P. Expert, F. Turkheimer, R. Carhart-Harris, D. Nutt, P. J. Hellyer, and F. Vaccarino, *Homological scaffolds of brain functional networks*, J. Roy. Soc. Int. **11** (2014), no. 101, 20140873.

[31] V. Pirino, E. Riccomagno, S. Martinoia, and P. Massobrio, *A topological study of repetitive co-activation networks in in vitro cortical assemblies.*, Phys. Bio. **12** (2014), no. 1, 016007–016007.

[32] J. Rinzel, *Discussion: Electrical excitability of cells, theory and experiment: Review of the Hodgkin–Huxley foundation and update*, Bull. Math. Biol. **52** (1990), no. 1/2, 5–23.

[33] G. Singh, F. Memoli, T. Ishkhanov, G. Sapiro, G. Carlsson, and D. L. Ringach, *Topological analysis of population activity in visual cortex*, J. Vis. **8** (2008), no. 8, 11.

[34] G. Spreemann, B. Dunn, M. B. Botnan, and N. A. Baas, *Using persistent homology to reveal hidden information in neural data*, arXiv:1510.06629 [q-bio.NC] (2015).

[35] B. Stolz, *Computational topology in neuroscience*, Master's thesis, University of Oxford, 2014.

DEPARTMENT OF MATHEMATICS, THE PENNSYLVANIA STATE UNIVERSITY
E-mail address: ccurto@psu.edu