An Introduction to Applied Topology

Henry Adams
Colorado State University
An Introduction to Applied Topology

Part I: Topology applied to data analysis
Part II: Real examples, machine learning, and dynamical systems
Part III: Topology applied to sensor networks
Datasets have shapes
Example: Diabetes study
145 points in 5-dimensional space

An attempt to define the nature of chemical diabetes using a multidimensional analysis by G. M. Reaven and R. G. Miller, 1979
Datasets have shapes

Example: Cyclo-Octane (C₈H₁₆) data
1,000,000+ points in 24-dimensional space

Non-Manifold Surface Reconstruction from High Dimensional Point Cloud Data
by Shawn Martin and Jean-Paul Watson, 2010.
Datasets have shapes
Example: Cyclo-Octane (C\textsubscript{8}H\textsubscript{16}) data
1,000,000+ points in 24-dimensional space

Non-Manifold Surface Reconstruction from High Dimensional Point Cloud Data
by Shawn Martin and Jean-Paul Watson, 2010.
Datasets have shapes
Persistent homology

Analysis of Kolmogorov flow and Rayleigh–Bénard convection using persistent homology by Miroslav Kramár, Rachel Levanger, Jeffrey Tithof, Balachandra Suri, Mu Xu, Mark Paul, Michael F Schatz, Konstantin Mischaikow
Persistent homology

Analysis of Kolmogorov flow and Rayleigh–Bénard convection using persistent homology by Miroslav Kramár, Rachel Levanger, Jeffrey Tithof, Balachandra Suri, Mu Xu, Mark Paul, Michael F Schatz, Konstantin Mischaikow
A donut and coffee mug are “homotopy equivalent”, and considered to be the same shape. You can bend and stretch (but not tear) one to get the other.
The torus has a Betti sequence (1, 2, 1, 0, 1), since it has a single connected component, two different loops that cannot be deformed into a point (shown in red in the bottom panel of Figure 2c), and there is a two-dimensional surface that cannot be deformed into a point (shown in orange in Figure 2c). The Klein bottle has the same sequence as the torus (1, 2, 1, 0, 1). This shows that while two objects that are equivalent must have the same Betti sequences, two objects that are not equivalent do not necessarily have different sequences. Finally, a sphere has a sequence (1, 0, 1, 0, 1), as any one-dimensional loop on its surface can be deformed into a point. The Betti sequence therefore provides a signature (albeit not unique) of the underlying topology of the object.

These definitions work for smooth continuous objects. But suppose now that instead of a continuous rubbery object we are faced with a finite set of (noisy) points sampled from it, which may represent actual experimental data. How can one estimate the Betti numbers of the original object from these samples? The proposed method...
Topology studies shapes

Torus
Topology studies shapes

Klein bottle
Topology studies shapes

Klein bottle

Image credit: https://plus.maths.org/content/imaging-maths-inside-kliebottle
Topology studies shapes

Klein bottle

Image credit: https://plus.maths.org/content/imaging-maths-inside-klein-bottle
Homology

- $i$-dimensional homology $H_i$ “counts the number of $i$-dimensional holes”
- $i$-dimensional homology $H_i$ actually has the structure of a vector space!

0-dimensional homology $H_0$: rank 6
1-dimensional homology $H_1$: rank 0

0-dimensional homology $H_0$: rank 1
1-dimensional homology $H_1$: rank 3

0-dimensional homology $H_0$: rank 1
1-dimensional homology $H_1$: rank 6
Homology

- $i$-dimensional homology “counts the number of $i$-dimensional holes”
- $i$-dimensional homology actually has the structure of a vector space!

0-dimensional homology $H_0$: rank 1
1-dimensional homology $H_1$: rank 0
2-dimensional homology $H_2$: rank 1

0-dimensional homology $H_0$: rank 1
1-dimensional homology $H_1$: rank 2
2-dimensional homology $H_2$: rank 1

Be careful! (Same as torus over $\mathbb{Z}/2\mathbb{Z}$)

Image credit: https://plus.maths.org/content/imaging-maths-inside-klein-bottle
torus has a Betti sequence \((1, 2, 1, 0, 1)\), since it has a single connected component, two different loops that cannot be deformed into a point (shown in red in the bottom panel of Figure 2c), and there is a two-dimensional surface that cannot be deformed into a point (shown in orange in Figure 2c). The Klein bottle has the same sequence as the torus \((1, 2, 1, 0, 1)\). This shows that while two objects that are equivalent must have the same Betti sequences, two objects that are not equivalent do not necessarily have different sequences. Finally, a sphere has a sequence \((1, 0, 1, 0, 1)\), as any one-dimensional loop on its surface can be deformed into a point. The Betti sequence therefore provides a signature (albeit not unique) of the underlying topology of the object.

These definitions work for smooth continuous objects. But suppose now that instead of a continuous rubbery object we are faced with a finite set of (noisy) points sampled from it, which may represent actual experimental data. How can one estimate the Betti numbers of the original object from these samples? The proposed method...
Homology

0-simplex 1-simplex 2-simplex 3-simplex

• — ▲ ▣
Homology

0-simplex

1-simplex

2-simplex

3-simplex

Simplicial complexes
Homology

0-simplex  1-simplex  2-simplex  3-simplex
Homology

0-simplices

1-simplices

2-simplices
Homology

0-simplices

1-simplices

2-simplices
Homology

A cycle has no boundary.
A cycle has no boundary.
Fundamental Lemma of Homology.

\[ \partial p \partial_{p+1} + d = 0 \]

for every integer \( p \) and every \((p+1)\)-chain \( d \).

Proof. We just need to show that \( \partial p \partial_{p+1} \tau = 0 \) for a \((p+1)\)-simplex \( \tau \).

The boundary, \( \partial_{p+1} \tau \), consists of all \( p \)-faces of \( \tau \). Every \((p-1)\)-face of \( \tau \) belongs to exactly two \( p \)-faces, so \( \partial p (\partial_{p+1} \tau) = 0 \).

It follows that every \( p \)-boundary is also a \( p \)-cycle or, equivalently, that \( B_p \) is a subgroup of \( \mathbb{Z}_p \).

Figure IV.4 illustrates the subgroup relations among the three types of groups and their connection across dimensions through boundary homomorphisms.

\[ \begin{array}{c}
\mathcal{B}_p \\
\mathcal{Z}_p \\
\mathcal{C}_p
\end{array} \]

Hence, the coset is obtained by adding each \( p \)-boundary to a given \( p \)-cycle, \( c \), \( \in \mathbb{Z}_p \).

More formally, this collection is called a coset.

Any two cycles in the same coset are said to be homologous, which is denoted as \( c_1 \sim c_2 \); see Figure IV.5.

We may take \( c_1 \) as the representative of this coset.

Figure IV.5: A torus with three closed curves, each an 1-cycle. Only one 1-boundary and it is homologous to the sum of the other two. The sum of the three curves is therefore a 1-boundary, namely of the pair of pants between them.

Two cycles are equivalent if they differ by a boundary. \( H_i \) measures equivalence classes of \( i \)-cycles.
Homology

- $i$-dimensional homology “counts the number of $i$-dimensional holes”
- $i$-dimensional homology actually has the structure of a vector space!

\[\begin{align*}
0\text{-dimensional homology } H_0 &: \text{ rank } 1 \\
1\text{-dimensional homology } H_1 &: \text{ rank } 0 \\
2\text{-dimensional homology } H_2 &: \text{ rank } 1 \\
\end{align*}\]

\[\begin{align*}
0\text{-dimensional homology } H_0 &: \text{ rank } 1 \\
1\text{-dimensional homology } H_1 &: \text{ rank } 2 \\
2\text{-dimensional homology } H_2 &: \text{ rank } 1 \\
\end{align*}\]

Be careful! (Same as torus over $\mathbb{Z}/2\mathbb{Z}$)

Image credit: https://plus.maths.org/content/imaging-maths-inside-klein-bottle
Topology studies shapes

What shape is this?
Definition

For metric space $X$ and scale $r \geq 0$, the Vietoris–Rips simplicial complex $\text{VR}(X; r)$ has vertex set $X$ and finite simplex when $\text{diam}(X) \leq r$. 

---

Diagram: A set of points arranged in a circular pattern, representing a Vietoris–Rips simplicial complex.
Definition

For metric space $X$ and scale $r \geq 0$, the Vietoris–Rips simplicial complex $\text{VR}(X; r)$ has vertex set $X$ and finite simplex when $\text{diam}(X) \leq r$. 
Definition

For metric space $X$ and scale $r \geq 0$, the Vietoris–Rips simplicial complex $\text{VR}(X; r)$ has a finite simplex when $\text{diam}(X) \leq r$. 
Definition

For metric space $X$ and scale $r \geq 0$, the Vietoris–Rips simplicial complex $\text{VR}(X; r)$ has vertex set $X$ and finite simplex when $\text{diam}(X) \leq r$. 
Definition

For metric space $X$ and scale $r \geq 0$, the Vietoris–Rips simplicial complex $\operatorname{VR}(X; r)$ has vertex set $X$ and contains a finite simplex when $\text{diam}(X) \leq r$. 
Definition

For metric space $X$ and scale $r \geq 0$, the Vietoris–Rips simplicial complex $\text{VR}(X; r)$ has vertex set $X$ and a finite simplex when $\text{diam}(X) \leq r$. 
For metric space $X$ and scale $r \geq 0$, the Vietoris–Rips simplicial complex $\text{VR}(X; r)$ has vertex set $X$ by finite simplex when $\text{diam}(X) \leq r$. 

Diagram: A cloud of points with circles of radius $r$ drawn around each point, indicating proximity and potential simplicial connections.
**Definition**

For a data set $X \subseteq \mathbb{R}^n$ and scale $r \geq 0$, the Čech simplicial complex $\check{\text{Č}}\text{ech}(X; r)$ has

- vertex set $X$
- finite simplex $\{x_0, x_1, \ldots, x_k\}$ when $\cap_{i=0}^k B(x_i, r) \neq \emptyset$. 

---

**Diagram**

A diagram showing a set of points in a two-dimensional space, illustrating the concept of a Čech complex.
For a data set $X \subseteq \mathbb{R}^n$ and scale $r \geq 0$, the Čech simplicial complex $\check{\text{C}}ech(X; r)$ has

- vertex set $X$
- finite simplex $\{x_0, x_1, \ldots, x_k\}$ when $\bigcap_{i=0}^k B(x_i, r) \neq \emptyset$. 
Definition

For a data set $X \subseteq \mathbb{R}^n$ and scale $r \geq 0$, the Čech simplicial complex $\check{\text{Č}}ech(X; r)$ has

- vertex set $X$
- finite simplex \{\(x_0, x_1, \ldots, x_k\)\} when $\bigcap_{i=0}^{k} B(x_i, r) \neq \emptyset$. 

\[ \text{Definition} \]

For a data set $X \subseteq \mathbb{R}^n$ and scale $r \geq 0$, the Čech simplicial complex $\check{\text{Č}}ech(X; r)$ has

- vertex set $X$
- finite simplex $\{x_0, x_1, \ldots, x_k\}$ when $\bigcap_{i=0}^{k} B(x_i, r) \neq \emptyset$. 

\[ �� \]
Definition

For a data set $X \subseteq \mathbb{R}^n$ and scale $r \geq 0$, the Čech simplicial complex $\check{\text{Čech}}(X; r)$ has

- vertex set $X$
- finite simplex $\{x_0, x_1, \ldots, x_k\}$ when $\cap_{i=0}^k B(x_i, r) \neq \emptyset$. 
Definition

For a data set $X \subseteq \mathbb{R}^n$ and scale $r \geq 0$, the Čech simplicial complex $\check{\text{Čech}}(X; r)$ has

- vertex set $X$
- finite simplex $\{x_0, x_1, \ldots, x_k\}$ when $\bigcap_{i=0}^{k} B(x_i, r) \neq \emptyset$. 
Definition

For a data set $X \subseteq \mathbb{R}^n$ and scale $r \geq 0$, the Čech simplicial complex $\check{\text{Cech}}(X; r)$ has

- vertex set $X$
- finite simplex $\{x_0, x_1, \ldots, x_k\}$ when $\bigcap_{i=0}^{k} B(x_i, r) \neq \emptyset$. 
Definition

For a data set $X \subseteq \mathbb{R}^n$ and scale $r \geq 0$, the Čech simplicial complex $\check{C}\text{ech}(X; r)$ has

- vertex set $X$
- finite simplex $\{x_0, x_1, \ldots, x_k\}$ when $\cap_{i=0}^{k} B(x_i, r) \neq \emptyset$. 

[Image of a geometric structure]
Nerve Lemma. \( \check{\text{Čech}}(X; r) \simeq \text{union of balls} \)

**Definition**

For a data set \( X \subseteq \mathbb{R}^n \) and scale \( r \geq 0 \), the \( \check{\text{Čech simplicial complex}} \ \check{\text{Čech}}(X; r) \) has

- vertex set \( X \)
- finite simplex \( \{x_0, x_1, \ldots, x_k\} \) when \( \bigcap_{i=0}^{k} B(x_i, r) \neq \emptyset \).
Definition

For a metric space $X$ and scale $r \geq 0$, the Vietoris–Rips simplicial complex $VR(X; r)$ has

- vertex set $X$
- finite simplex $\{x_0, x_1, \ldots, x_k\}$ when $d(x_i, x_j) \leq r$ for all $i, j$. 
Definition

For a metric space $X$ and scale $r \geq 0$, the *Vietoris–Rips simplicial complex* $\text{VR}(X; r)$ has

- vertex set $X$
- finite simplex $\{x_0, x_1, \ldots, x_k\}$ when $d(x_i, x_j) \leq r$ for all $i, j$. 

\[
\text{Cech simplicial complex} \\
\text{Appearance} \\
draw one simplex \\
draw Cech complex \\
draw Rips complex \\
\text{Filtration parameter} t = 0.14
\]
Definition

For a metric space \( X \) and scale \( r \geq 0 \), the
Vietoris–Rips simplicial complex \( \text{VR}(X; r) \) has

- vertex set \( X \)
- finite simplex \( \{x_0, x_1, \ldots, x_k\} \) when \( d(x_i, x_j) \leq r \) for all \( i, j \).
Definition
For a metric space $X$ and scale $r \geq 0$, the Vietoris–Rips simplicial complex $\text{VR}(X; r)$ has

- vertex set $X$
- finite simplex $\{x_0, x_1, \ldots, x_k\}$ when $d(x_i, x_j) \leq r$ for all $i, j$. 

Definition

For a metric space $X$ and scale $r \geq 0$, the **Vietoris–Rips simplicial complex** $\text{VR}(X; r)$ has

- vertex set $X$
- finite simplex $\{x_0, x_1, \ldots, x_k\}$ when $d(x_i, x_j) \leq r$ for all $i, j$. 
**Definition**

For a metric space $X$ and scale $r \geq 0$, the *Vietoris–Rips simplicial complex* $VR(X; r)$ has

- vertex set $X$
- finite simplex $\{x_0, x_1, \ldots, x_k\}$ when $d(x_i, x_j) \leq r$ for all $i, j$. 
Definition

For a metric space $X$ and scale $r \geq 0$, the
*Vietoris–Rips simplicial complex* $\text{VR}(X; r)$ has

- vertex set $X$
- finite simplex $\{x_0, x_1, \ldots, x_k\}$ when $d(x_i, x_j) \leq r$ for all $i, j$. 
In persistent homology, given a large dataset, the topology of logical spaces is arbitrary, as opposed to the unidirectional sequence of maps of zigzag persistence. The direction of maps along a sequence of topological spaces is important. However, without knowing the underlying space, we disregard short intervals as noise. Hence, this barcode diagram shows the intervals in the persistence barcodes correspond to real topological features of the underlying space.

- **Input**: Increasing spaces. **Output**: barcode.
- **Significant features persist.**
- **Cubic computation time in the number of simplices.**
Persistent homology

- Significant features persist.
- Cubic computation time in the number of simplices.
Persistent homology

- Significant features persist.
- Cubic computation time in the number of simplices.
Persistent homology

- Significant features persist.
- Cubic computation time in the number of simplices.
**Persistent homology**

**Sublevelset persistence**

- **Input**: Increasing spaces. **Output**: barcode.
- **Significant features persist.**
- **Cubic computation time in the number of simplices.**

*Figure 5.2: (a) A snapshot of the $z$-component of the vorticity field for Kolmogorov flow from the stable relative periodic orbit found at $Re = 25.43$. (b) A snapshot of the renormalized 8-bit mid-plane temperature field $T^*$ for Rayleigh-Bénard convection from the stable almost-periodic orbit found at $Ra = 3000$ and $Pr = 1$.***

*Figure 5.3: (a-d) Sublevel sets $\mathcal{C}(\omega \rightarrow \mathcal{X}) = \{x \in \mathcal{D}_\omega : \omega(x) \leq \mathcal{X}\}$ of the vorticity field, shown in Figure 5.2(a), for different values of $\mathcal{X}$, depicted in black. (e) $PD_0(\omega)$ and (f) $PD_1(\omega)$ persistence diagrams of the vorticity field indicate the values of $\mathcal{X}$ at which the connected components and loops appear and disappear (merge together). Video 1 of the supplementary materials of [1] provides an animation.*

*Analysis of Kolmogorov flow and Rayleigh–Bénard convection using persistent homology* by Miroslav Kramár, Rachel Levanger, Jeffrey Tithof, Balachandra Suri, Mu Xu, Mark Paul, Michael F Schatz, Konstantin Mischaikow
Persistent homology

Sublevelset persistence

Analysis of Kolmogorov flow and Rayleigh–Bénard convection using persistent homology by Miroslav Kramár, Rachel Levanger, Jeffrey Tithof, Balachandra Suri, Mu Xu, Mark Paul, Michael F Schatz, Konstantin Mischaikow

- Significant features persist.
- Cubic computation time in the number of simplices.
Persistent homology applied to data

Example: Cyclo-Octane (C₈H₁₆) data

1,000,000+ points in 72-dimensional space

Non-Manifold Surface Reconstruction from High Dimensional Point Cloud Data
by Shawn Martin and Jean-Paul Watson, 2010.
Persistent homology applied to data

Example: Cyclo-Octane (C₈H₁₆) data
1,000,000+ points in 72-dimensional space

Non-Manifold Surface Reconstruction from High Dimensional Point Cloud Data
by Shawn Martin and Jean-Paul Watson, 2010.
**Table 2**

Example run times. Here we show the run times obtained for the different examples investigated in this section. For each example we provide the number of points \( n \), number of landmarks \( L \), neighborhood size \( k \), time in seconds for pre-processing, and time in seconds for reconstruction.

<table>
<thead>
<tr>
<th>Example</th>
<th>( n )</th>
<th>( L )</th>
<th>( k )</th>
<th>Pre-proc.</th>
<th>Recon.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sphere</td>
<td>10,000</td>
<td>886</td>
<td></td>
<td>1.7</td>
<td>368.2</td>
</tr>
<tr>
<td>Torus</td>
<td>10,000</td>
<td>667</td>
<td>28</td>
<td>1.61</td>
<td>220.5</td>
</tr>
<tr>
<td>Double torus</td>
<td>20,000</td>
<td>813</td>
<td>26</td>
<td>3.7</td>
<td>263.1</td>
</tr>
<tr>
<td>Mobius strip</td>
<td>10,000</td>
<td>416</td>
<td>23</td>
<td>0.9</td>
<td>123.8</td>
</tr>
<tr>
<td>Klein figure</td>
<td>8</td>
<td>1940</td>
<td>33</td>
<td>3.8</td>
<td>778.0</td>
</tr>
<tr>
<td>( \mathbb{R}P^2 )</td>
<td>100,000</td>
<td>753</td>
<td>35</td>
<td>1.73</td>
<td>111.5</td>
</tr>
<tr>
<td>Two spheres</td>
<td>83,646</td>
<td>1588</td>
<td>13</td>
<td>446.4</td>
<td>344.4</td>
</tr>
<tr>
<td>Klein immersion</td>
<td>61,440</td>
<td>4566</td>
<td>14</td>
<td>295.7</td>
<td>1183.3</td>
</tr>
<tr>
<td>Henneberg</td>
<td>13,637</td>
<td>1463</td>
<td>39</td>
<td>40.9</td>
<td>723.4</td>
</tr>
</tbody>
</table>

**Fig. 7.** Conformation space of cyclo-octane. Here we show how the set of conformations of cyclo-octane can be represented as a surface in a high-dimensional space. On the left, we show various conformations of cyclo-octane as drawn by PyMol ([www.pymol.org](http://www.pymol.org)). In the center, these conformations are represented by the 3D coordinates of their atoms. The coordinates are concatenated into vectors and shown as columns of a data matrix. As an example, the entry \( c_{1,1,x} \) of the matrix denotes the \( x \)-coordinate of the first carbon atom in the first molecule. On the right, the Isomap method is used to obtain a lower-dimensional visualization of the data.

**3.5. Run times**
The run times for the nine examples we have investigated are shown in Table 2. These times were obtained on a 2.26 GHz Intel Xeon dual quadcore workstation with 16 GB of RAM. The algorithm was implemented in Matlab ([www.mathworks.com](http://www.mathworks.com)) using the optimization toolbox to solve the linear program in (6).

Table 2 shows that pre-processing is negligible except for the non-manifold examples. In the case of the non-manifold examples, the pre-processing is generally faster than the triangulation.

**4. Application**
Cyclo-octane is a saturated eight-member cyclic compound with chemical formula \( \text{C}_8\text{H}_{16} \). Cyclo-octane has received attention in computational chemistry because it has multiple conformations of similar energy, a complex potential energy surface, and significant (steric) influence from the hydrogen atoms on preferred conformations [32–34]. Cyclo-octane is also interesting because there are enumerative algorithms available which can provide a dense sampling of the conformation space [35,36]. These algorithms show from first principles that the resulting conformation space has two degrees of freedom, suggesting that the space is a surface (but not necessarily a manifold).

Using dimension reduction methods, we have previously analyzed the cyclo-octane conformation space [16]. In our analysis, we used a dataset of 1,031,644 cyclo-octane conformations, enumerated using the triaxial loop closure algorithm of Coutsias et al. [35]. Each conformation is placed in Cartesian space via the 3D position coordinates of each atom in the molecule. The conformations are then aligned to a reference conformation such that the Eckart conditions are satisfied [37]. The final positions of a given conformation are concatenated to obtain a vector in \( \mathbb{R}^{72} \). The resulting collection is a dataset \( \{x_i\}_{i=1}^{1,031,644} \subset \mathbb{R}^{72} \) which is presumed to describe a surface. In Brown et al. [16] we applied a variety of dimension reduction methods to the cyclo-octane dataset, one of which was Isomap [38]. A summary of our analysis using the Isomap reduction is shown in Fig. 7.

Beyond dimension reduction, the next step in our analysis is surface reconstruction. Unfortunately, the Isomap representation of the cyclo-octane conformation space is only a visualization, and is not accurate enough for use with a 3D surface reconstruction methods. Therefore we applied Freedman’s algorithm for surface reconstruction in the original high-dimensional conformation space. Freedman’s method failed because the surface had self-intersections of the type discussed in this paper. Thus we developed our method for non-manifold surface reconstruction and applied it to the cyclo-octane dataset.

**Persistent homology applied to data**

**Non-Manifold Surface Reconstruction from High Dimensional Point Cloud Data** by Shawn Martin and Jean-Paul Watson, 2010.
Persistent homology applied to data

Example: Cyclo-Octane (C₈H₁₆) data

1,000,000+ points in 72-dimensional space

Non-Manifold Surface Reconstruction from High Dimensional Point Cloud Data
by Shawn Martin and Jean-Paul Watson, 2010.
Persistent homology applied to data

Example: Cyclo-Octane (C₈H₁₆) data

1,000,000+ points in 72-dimensional space

Non-Manifold Surface Reconstruction from High Dimensional Point Cloud Data
by Shawn Martin and Jean-Paul Watson, 2010.
Persistent homology applied to data

Example: Cyclo-Octane ($C_8H_{16}$) data

1,000,000+ points in 72-dimensional space

Non-Manifold Surface Reconstruction from High Dimensional Point Cloud Data
by Shawn Martin and Jean-Paul Watson, 2010.
Persistent homology applied to data

Example: Cyclo-Octane (C₈H₁₆) data

1,000,000+ points in 72-dimensional space

Non-Manifold Surface Reconstruction from High Dimensional Point Cloud Data
by Shawn Martin and Jean-Paul Watson, 2010.
Persistent homology applied to data

Example: Equilateral pentagons in the plane

Image credit: Clayton Shonkwiler
Persistent homology applied to data

Example: Equilateral pentagons in the plane

Image credit: Clayton Shonkwiler
Persistent homology applied to data

Image credit: Clayton Shonkwiler
Persistent homology applied to data

- **Stability Theorem.**
  
  If $X$ and $Y$ are metric spaces, then

  $$d_b(\text{PH}(\check{\text{Cech}}(X)), \text{PH}(\check{\text{Cech}}(Y))) \leq 2d_{\text{GH}}(X, Y)$$
Conclusions for Part I

• Datasets have shape, which are reflective of patterns within.
• Persistent homology is a way to measure some of the local geometry and global topology of a dataset.
Topology applied to image data
The receptive fields of cells in our primary visual cortex (V1) are related to the statistics of natural images.

*Independent component filters of natural images compared with simple cells in primary visual cortex* by JH van Hateren and A van der Schaaf, 1997
Persistent homology applied to data

3x3 high-contrast patches from images
Points in 9-dimensional space, normalized to have average color gray and contrast norm one (on 7-sphere).

Persistent homology applied to data

1. Densest patches according to a global estimate
Persistent homology applied to data

1. Densest patches according to a global estimate

Interpretation: nature prefers linearity
Persistent homology applied to data

2. Densest patches according to an intermediate estimate
Persistent homology applied to data

2. Densest patches according to an intermediate estimate

Interpretation: nature prefers horizontal and vertical directions
Persistent homology applied to data

3. Densest patches according to a local estimate
Persistent homology applied to data

3. Densest patches according to a local estimate
Persistent homology applied to data

3. Densest patches according to a local estimate
Persistent homology applied to data

3. Densest patches according to a local estimate

Interpretation: nature prefers linear and quadratic patches at all angles

Image credit: https://plus.maths.org/content/imaging-maths-inside-klein-bottle
Persistent homology applied to data

3. Densest patches according to a local estimate

Interpretation: nature prefers linear and quadratic patches at all angles
Persistent homology applied to data

Range Images

Fig. 4.4. Linear step edges in the top row and their binary approximations beneath a manifold model. In Figures 4.5 and 4.6 are sample Betti barcode plots for the core subsets $X_{v300, 30y}$ and $X_{v300, 30y}$. $0$ $0.05$ $0.1$ $0.15$ $0.2$ $0.25$ $0.3$ $0.35$ $0.4$ $0$ $10$ $20$ $30$ $40$ $50$ Betti

Fig. 4.5. Barcodes for $X_{v300, 30y}$. $0$ $0.05$ $0.1$ $0.15$ $0.2$ $0.25$ $0.3$ $0.35$ $0$ $10$ $20$ $30$ $40$ $50$ Betti $0$ $1$ $0$ $0.05$ $0.1$ $0.15$ $0.2$ $0.25$ $0.3$ $0.35$ $0$ $10$ $20$ $30$ $40$ $50$ Betti

Fig. 4.6. Barcodes for $X_{v300, 30y}$. $0$ $0.05$ $0.1$ $0.15$ $0.2$ $0.25$ $0.3$ $0.35$ $0$ $10$ $20$ $30$ $40$ $50$ Betti

Both the $X_{v300, 30y}$ plot and the $X_{v300, 30y}$ plot contain a single long Betti $0$ interval and a single long Betti $1$ interval, evidence of circular topology. The homology cycle producing the Betti $1$ interval is the primary circle, visible in Figures 4.7 and 4.8.

We ran 25 trials each on $X_{v300, 30y}$ and $X_{v300, 30y}$, selecting different landmark points. In each trial, the circular profile $Betti_0 = Betti_1 = 1$ is obtained for a range of $R$ values of length greater than 0.25, and no other Betti plot interval has length greater than 0.05.
Optical flow is a vector field representing the apparent motion (or projected motion) in a video.

On the nonlinear statistics of optical flow by HA, Johnathan Bush, Brittany Carr, Lara Kassab, and Joshua Mirth, 2018
Optical flow is a vector field representing the apparent motion (or projected motion) in a video.

*On the nonlinear statistics of optical flow* by HA, Johnathan Bush, Brittany Carr, Lara Kassab, and Joshua Mirth, 2018
Persistent homology applied to data

Optical Flow

On the nonlinear statistics of optical flow by HA, Johnathan Bush, Brittany Carr, Lara Kassab, and Joshua Mirth, 2018
Why is applied topology popular when few datasets have Klein bottles?

- Many datasets have clusters & flares (as in the diabetes example)
- Motivates interesting questions in many pure disciplines: mathematics, computer science (computational geometry), statistics
- Interest from domain experts in biology, neuroscience, computer vision, dynamical systems, sensor networks, ...
- Materials science, pattern formation
- Machine learning: small features matter
- Agent-based modeling (swarming)

---

Why is applied topology popular when few datasets have Klein bottles?

- Many datasets have clusters & flares (as in the diabetes example)
- Motivates interesting questions in many pure disciplines: mathematics, computer science (computational geometry), statistics
- Interest from domain experts in biology, neuroscience, computer vision, dynamical systems, sensor networks, ...
- Materials science, pattern formation
- Machine learning: small features matter
- Agent-based modeling (swarming)

Answer: \( \text{From left: } 1.75, 2, 1.75, 2, 2 \).
Why is applied topology popular when few datasets have Klein bottles?

- Many datasets have clusters & flares (as in the diabetes example)
- Motivates interesting questions in many pure disciplines: mathematics, computer science (computational geometry), statistics
- Interest from domain experts in biology, neuroscience, computer vision, dynamical systems, sensor networks, ...
- Materials science, pattern formation
- Machine learning: small features matter
- Agent-based modeling (swarming)
Agent-Based Modeling

Collective phenomenon, self-organization
Dutch Starling murmuration filmed by Roald van Stijn

https://www.youtube.com/watch?v=YjDYE5CUb7Q
The following slides and images are largely from: Chad Topaz, Lu Xuan, Lori Ziegelmeier

**Biological Aggregations**

- A system in which agents interact with each other
- Formed through social interaction and coordinated behaviors like attraction, repulsion, and/or alignment
- Self-organization

Video retrieved from [1], images from [4] [3] [5]
In many natural systems, particles, organisms, or agents interact locally according to rules that produce aggregate behavior.
Vicsek Model

*Topological data analysis of biological aggregation models*
by Chad M Topaz, Lori Ziegelmeier, Tom Halverson, 2015.
Vicsek Model

\[ \theta_i(t + \Delta t) = \frac{1}{N} \left( \sum_{|x_i - x_j| \leq R} \theta_j(t) \right) + U\left(-\frac{\eta}{2}, \frac{\eta}{2}\right) \]

Topological data analysis of biological aggregation models
by Chad M Topaz, Lori Ziegelmeier, Tom Halverson, 2015.
Vicsek Model

\[ \theta_i(t + \Delta t) = \frac{1}{N} \left( \sum_{|x_i - x_j| \leq R} \theta_j(t) \right) + U\left(-\frac{\eta}{2}, \frac{\eta}{2}\right) \]

Topological data analysis of biological aggregation models
by Chad M Topaz, Lori Ziegelmeier, Tom Halverson, 2015.
Fig 7. Simulation snapshots of the Vicsek model (9). These simulations are analogous to Fig. 1 in [11]. Circles indicate particle positions and line segments represent heading. For all simulations, \( N = 300 \) particles, the particle speed is \( v_0 = 0.03 \), and the initial state consists of uniform random positions and headings. We vary box size \( \ell \) and noise \( \eta \). Dotted lines indicate the bounds of the periodic domain. (A) Groups moving in different directions with \( \ell = 25 \), \( \eta = 0.1 \), \( t = 3000 \). (B) Random movement with some correlation with \( \ell = 7 \), \( \eta = 2 \), \( t = 600 \). (C) Highly polarized motion with \( \ell = 5 \), \( \eta = 0.1 \), \( t = 300 \).
Persistent homology is a reasonable summary of local geometry and global topology.

Can you reject the null hypothesis that your model fits the data?
# Aphid Motion

## Assessing Model Validity

Assessing Biological Models Using Topological Data Analysis

### Summaries of statistical tests comparing models of aphid motion using order parameters.

<table>
<thead>
<tr>
<th>Exp</th>
<th>$D$</th>
<th>$R_{95%}$</th>
<th>$M_{avg}$</th>
<th>$D$</th>
<th>$R_{95%}$</th>
<th>$M_{abs}$</th>
<th>$d_a$</th>
<th>$R_{95%}$</th>
<th>$D$</th>
<th>$R_{95%}$</th>
<th>$Mov%$</th>
<th>$R_{95%}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.17</td>
<td>0.13</td>
<td>1.50</td>
<td>0.16</td>
<td>-2.00</td>
<td>0.15</td>
<td>0.26</td>
<td>0.05</td>
<td>20.2</td>
<td>0.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-0.25</td>
<td>0.10</td>
<td>-0.97</td>
<td>0.11</td>
<td>-1.66</td>
<td>0.09</td>
<td>0.25</td>
<td>0.02</td>
<td>0.78</td>
<td>0.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.09</td>
<td>0.10</td>
<td>0.40</td>
<td>0.11</td>
<td>-1.71</td>
<td>0.10</td>
<td>0.54</td>
<td>0.04</td>
<td>25.1</td>
<td>0.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.70</td>
<td>0.08</td>
<td>-0.78</td>
<td>0.09</td>
<td>-1.24</td>
<td>0.07</td>
<td>0.16</td>
<td>0.02</td>
<td>1.68</td>
<td>0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-2.22</td>
<td>0.09</td>
<td>-0.63</td>
<td>0.09</td>
<td>-1.10</td>
<td>0.07</td>
<td>0.31</td>
<td>0.03</td>
<td>7.68</td>
<td>0.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.04</td>
<td>0.09</td>
<td>-0.49</td>
<td>0.10</td>
<td>-1.68</td>
<td>0.09</td>
<td>0.53</td>
<td>0.04</td>
<td>24.0</td>
<td>0.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>-0.97</td>
<td>0.11</td>
<td>-1.28</td>
<td>0.13</td>
<td>-1.99</td>
<td>0.11</td>
<td>0.57</td>
<td>0.04</td>
<td>23.9</td>
<td>0.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>-3.30</td>
<td>0.25</td>
<td>-2.11</td>
<td>0.22</td>
<td>-2.87</td>
<td>0.26</td>
<td>0.57</td>
<td>0.25</td>
<td>13.2</td>
<td>0.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-4.69</td>
<td>0.25</td>
<td>-3.14</td>
<td>0.22</td>
<td>-2.81</td>
<td>0.20</td>
<td>0.48</td>
<td>0.15</td>
<td>11.7</td>
<td>0.34</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Summaries of statistical tests comparing models of aphid motion using topology.

<table>
<thead>
<tr>
<th>Exp</th>
<th>$b_0(\text{pose})$</th>
<th>$D^*$</th>
<th>$R_{95%}$</th>
<th>$b_1(\text{pose})$</th>
<th>$D^*$</th>
<th>$R_{95%}$</th>
<th>$b_0(\text{pose/vel})$</th>
<th>$D^*$</th>
<th>$R_{95%}$</th>
<th>$b_1(\text{pose/vel})$</th>
<th>$D^*$</th>
<th>$R_{95%}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1147</td>
<td>38.73</td>
<td>135.1</td>
<td>5.19</td>
<td>831.9</td>
<td>42.81</td>
<td>124.2</td>
<td>5.640</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1703</td>
<td>23.98</td>
<td>297.0</td>
<td>6.50</td>
<td>858.9</td>
<td>16.88</td>
<td>261.1</td>
<td>7.258</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1893</td>
<td>37.36</td>
<td>290.4</td>
<td>6.83</td>
<td>1347</td>
<td>35.60</td>
<td>271.3</td>
<td>7.218</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1825</td>
<td>18.36</td>
<td>359.0</td>
<td>5.74</td>
<td>552.3</td>
<td>19.04</td>
<td>332.5</td>
<td>6.301</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1370</td>
<td>15.81</td>
<td>249.2</td>
<td>6.03</td>
<td>270.8</td>
<td>15.25</td>
<td>225.1</td>
<td>6.401</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2066</td>
<td>40.53</td>
<td>322.9</td>
<td>7.06</td>
<td>1484</td>
<td>37.35</td>
<td>307.7</td>
<td>7.116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1747</td>
<td>34.32</td>
<td>264.8</td>
<td>6.23</td>
<td>1298</td>
<td>33.69</td>
<td>247.5</td>
<td>6.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>298.5</td>
<td>20.81</td>
<td>10.28</td>
<td>1.92</td>
<td>147.9</td>
<td>19.81</td>
<td>3.573</td>
<td>1.641</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>467.5</td>
<td>19.66</td>
<td>22.34</td>
<td>2.14</td>
<td>245.3</td>
<td>21.10</td>
<td>10.93</td>
<td>2.067</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Ulmer, Z., Topaz 2018) Assessing Biological Models Using Topological Data Analysis
ASSESSING MODEL VALIDITY

Fig 6. Examples of average $b_0(\text{pos})$ crocker plots over 100 runs of the two models simulated using initial conditions from experiments. The left column corresponds to position data simulated with the control model and the right from the interactive model. The first row corresponds to Experiment 2, the second from Experiment 3, and the third from Experiment 7. In general, crocker plots arising from experiments with more tightly clustered initial configurations exhibit upward trending of the contours over time in contrast to experiments where aphids are more dispersed initially; refer to Fig 1.
D’Orsogna Model

Repulsion on short length scales $L_r$, attraction on large length scales $L_a$

\[
\dot{x}_i = v_i,
\]

\[
m\dot{v}_i = (\alpha - \beta |v_i|^2)v_i - \nabla Q_i,
\]

\[
Q_i = \sum_{j\neq i} C_re^{-|x_i - x_j|/L_r} - C_a e^{-|x_i - x_j|/L_a}
\]

Topological data analysis of biological aggregation models
by Chad M Topaz, Lori Ziegelmeier, Tom Halverson, 2015.
Applied Mathematical Modeling with Topological Techniques
Aug 5 - 9, 2019

Organizing Committee

- Henry Adams
  Colorado State University
- Jose Perea
  Michigan State University
- Maria D'Orsogna
  California State University, Northridge
- Chad Topaz
  Williams College
- Rachel Neville
  University of Arizona
Takens’ Theorem

Roughly speaking: Let $M$ be a $d$-dimensional compact manifold, let $\phi: M \to M$ be a flow, and let $f: M \to \mathbb{R}$ be a measurement. Then \textit{generically},

$$m \mapsto (f(m), f(\phi(m)), f(\phi^2(m)), \ldots, f(\phi^{2d}(m)))$$

is an embedding $M \hookrightarrow \mathbb{R}^{2d+1}$.

Detecting strange attractors in turbulence by Floris Takens, 1982
Takens’ Theorem

Roughly speaking: Let $M$ be a $d$-dimensional compact manifold, let $\phi: M \to M$ be a flow, and let $f: M \to \mathbb{R}$ be a measurement. Then generically,

$$m \mapsto (f(m), f(\phi(m), f(\phi^2(m), \ldots, f(\phi^{2d}(m))))$$

is an embedding $M \hookrightarrow \mathbb{R}^{2d+1}$.

---

Sliding windows and persistence: An application of topological methods to signal analysis by Jose Perea and John Harer, 2014
### Sliding windows and persistence: An application of topological methods to signal analysis

**Table 1:** Ranking of signals by periodicity. For each algorithm we provide the score and a normalized plot of the signal. The ranking goes from top (highest score) to bottom (lowest score).

**7.2 Classification Rates**

We compare the different algorithms by their ability to separate periodic from nonperiodic signals. The performance of this type of binary classification can be visualized using a receiver operating characteristic (ROC) plot, which compares the true positive rate (TPR) to the false positive rate (FPR) as a cutoff on the scores is varied. Here the TPR is the proportion of correctly identified positive cases out of all positives, and FPR is the proportion of negative cases incorrectly identified as positives out of all the negatives. The line $TPR = FPR$ is the performance of random guessing; the higher the ROC curve is above this line, the better its classification performance. An algorithm that is able to perfectly separate all positive from negative test cases would have a ROC curve that passes through the point $TPR = 1$ and $FPR = 0$. It follows that a reasonable measure of classification success for a particular method is the area under its ROC curves.

The synthetic data are generated as follows: the periodic signals (positive cases) span two periods and include a cosine, cosine with trending, cosine with damping, and cosine with increased peak steepness. The nonperiodic signals (negative cases) include a constant and a linear function. We generate 100 profiles from each shape by adjusting its phase. For instance, in the case of the cosine shape we let $f_i(t) = \cos(2\pi t - j\pi/50)$, $j = 0, \ldots, 99$. 

**Sliding windows**

**Takens’ Theorem**

![Takens' Theorem Image](Wikipedia)
Takens’ Theorem

Roughly speaking: Let $M$ be a $d$-dimensional compact manifold, let $\phi : M \to M$ be a flow, and let $f : M \to \mathbb{R}$ be a measurement. Then generically,

$$m \mapsto (f(m), f(\phi(m)), f(\phi^2(m)), \ldots, f(\phi^{2d}(m)))$$

is an embedding $M \hookrightarrow \mathbb{R}^{2d+1}$.

Detecting strange attractors in turbulence by Floris Takens, 1982
Sublevel set filtrations of dynamical system data

Figure 5.2: (a) A snapshot of the z-component of the vorticity field \( \omega \) for Kolmogorov flow from the stable relative periodic orbit found at \( \text{Re} = 25.43 \). (b) A snapshot of the renormalized 8-bit mid-plane temperature field \( T^\star \) for Rayleigh-Bénard convection from the stable almost-periodic orbit found at \( \text{Ra} = 3000 \) and \( \text{Pr} = 1 \).

\[ T^\star = \text{mean} \left( t + T^\star \right) \equiv \text{mean} \left( t \right), \] where \( T^\star \) is the period of the PO that is guiding the dynamics of the turbulent trajectory. The turbulent trajectory depicted in Figure 5.1(b) passes close to unstable EQ and REQ solutions which are indicated by the red dots.

5.1.2 Rayleigh-Bénard Convection

Rayleigh-Bénard convection is a canonical pattern forming system that has been used to gain many new fundamental insights into the spatiotemporal dynamics of systems that are driven far-from-equilibrium [43, 44]. Rayleigh-Bénard convection is the buoyancy driven fluid flow that occurs when a shallow layer of fluid is heated uniformly from below in a gravitational field. The dynamics are governed by the Boussinesq equations:

\[ \begin{align*}
\frac{\partial u}{\partial t} + u \cdot \nabla u &= -\nabla p + \nabla^2 u + \frac{\text{Ra}T^\star \hat{z}}{(5.3)} \\
\frac{\partial T}{\partial t} + u \cdot \nabla T &= \nabla^2 T \quad \text{\( (5.4) \)} \\
\nabla \cdot u &= 0 \quad \text{\( (5.5) \)}
\end{align*} \]

where \( u(x, y, z, t) \) is a vector field of the fluid velocity, \( p(x, y, z, t) \) is the pressure field, and \( T(x, y, z, t) \) is the temperature field. In our notation, the origin of the Cartesian coordinates \((x, y, z)\) at the center of the domain are at the lower heated plate where \( \hat{z} \) is:

- \( T^\star \leq 25 \) (a)
- \( T^\star \leq 100 \) (b)
- \( T^\star \leq 215 \) (c)
- \( T^\star \leq 230 \) (d)

Figure 5.4: (a-d) Sublevel sets \( C(T^\star \mapsto \mathcal{F}) = \{ x \in D \} \) of the renormalized 8-bit temperature field \( T^\star \), shown in Figure 5.2(b), for different values of \( \mathcal{F} \), depicted in black. As in Figure 5.3, the persistence diagrams \( \text{PD}_0(T^\star) \) and \( \text{PD}_1(T^\star) \) indicate the values of \( \mathcal{F} \) at which the connected components and loops appear and disappear (merge together). Video 2 of the supplementary materials of [1] provides an animation.
Conley index theory
“Topology! The stratosphere of human thought! In the twenty-fourth century it might possibly be of use to someone ...”

-Aleksandr Solzhenitsyn, *The First Circle*
Conclusions for Part II

- Data density matters
- Persistent homology is now being used as a summary statistic of local geometry and global topology, for use both in machine learning and in dynamical systems.
Evasion Paths in Mobile Sensor Networks
Topology applied to sensor networks

- Sensors move in a ball-shaped domain $B \subset \mathbb{R}^d$ over time interval $I = [0, 1]$. Fixed sensors cover the boundary.
- Measure only the Čech complex.
- Is there an evasion path?

Coordinate-free Coverage in Sensor Networks with Controlled Boundaries via Homology by V. de Silva and R. Ghrist
Topology applied to sensor networks

- Sensors move in a ball-shaped domain $B \subset \mathbb{R}^d$ over time interval $I = [0, 1]$. Fixed sensors cover the boundary.
- Measure only the Čech complex.
- Is there an evasion path?

Coordinate-free Coverage in Sensor Networks with Controlled Boundaries via Homology by V. de Silva and R. Ghrist
Topology applied to sensor networks

- Sensors move in a ball-shaped domain $B \subset \mathbb{R}^d$ over time interval $I = [0, 1]$. Fixed sensors cover the boundary.
- Measure only the Čech complex.
- Is there an evasion path?
Topology applied to sensor networks

- Sensors move in a ball-shaped domain $B \subset \mathbb{R}^d$ over time interval $I = [0, 1]$. Fixed sensors cover the boundary.
- Measure only the Čech complex.
- Is there an evasion path?
Topology applied to sensor networks

- Sensors move in a ball-shaped domain $B \subset \mathbb{R}^d$ over time interval $I = [0, 1]$. Fixed sensors cover the boundary.
- Measure only the Čech complex.
- Is there an evasion path?
Topology applied to sensor networks

Čech complex

- One vertex for each ball
- Edges when 2 balls overlap
- Triangles when 3 balls overlap

Coordinate-free Coverage in Sensor Networks with Controlled Boundaries via Homology by V. de Silva and R. Ghrist
Topology applied to sensor networks

Čech complex
- One vertex for each ball
- Edges when 2 balls overlap
- Triangles when 3 balls overlap

Vietoris-Rips complex
- One vertex for each ball
- Edges when 2 balls overlap
- All possible triangles

Coordinate-free Coverage in Sensor Networks with Controlled Boundaries via Homology by V. de Silva and R. Ghrist
Topology applied to sensor networks

- Let $X \subset B \times I$ be the covered region.
- An *evasion path* is a time-preserving map from $I$ to the uncovered region.

---

Coordinate-free Coverage in Sensor Networks with Controlled Boundaries via Homology by V. de Silva and R. Ghrist
Topology applied to sensor networks

- Let $X \subset B \times I$ be the covered region.
- An evasion path is a time-preserving map from $I$ to the uncovered region.
- Evasion Problem. Using the time-varying Čech complex, can we determine if an evasion path exists?
Topology applied to sensor networks

- Let $X \subset B \times I$ be the covered region.
- An evasion path is a time-preserving map from $I$ to the uncovered region.
- Evasion Problem. Using the time-varying Čech complex, can we determine if an evasion path exists?
Topology applied to sensor networks

- Let $X \subset B \times I$ be the covered region.
- An evasion path is a time-preserving map from $I$ to the uncovered region.
- Evasion Problem. Using the time-varying Čech complex, can we determine if an evasion path exists?
Topology applied to sensor networks

- The stacked complex $SC \simeq X$ encodes all Čech complexes.
Topology applied to sensor networks

- The stacked complex $SC \simeq X$ encodes all Čech complexes.

Coordinate-free Coverage in Sensor Networks with Controlled Boundaries via Homology by V. de Silva and R. Ghrist
Topology applied to sensor networks

- The stacked complex $\mathcal{SC} \simeq X$ encodes all Čech complexes.
Topology applied to sensor networks

- Theorem (de Silva, Ghrist).

Coordinate-free Coverage in Sensor Networks with Controlled Boundaries via Homology by V. de Silva and R. Ghrist
Topology applied to sensor networks

- Theorem (de Silva, Ghrist).
Topology applied to sensor networks

- Theorem (de Silva, Ghrist).

Coordinate-free Coverage in Sensor Networks with Controlled Boundaries via Homology by V. de Silva and R. Ghrist
Topology applied to sensor networks

- Theorem (de Silva, Ghrist).
  If there is an $\alpha \in H_d(SC, \partial B \times I)$ with
  $0 \neq \partial \alpha \in H_{d-1}(\partial B \times I)$, then no evasion path exists.
Theorem (de Silva, Ghrist). If there is an \( \alpha \in H_d(SC, \partial B \times I) \) with \( 0 \neq \partial \alpha \in H_{d-1}(\partial B \times I) \), then no evasion path exists.

Coordinate-free.
Topology applied to sensor networks

- **Theorem (de Silva, Ghrist).**
  If there is an $\alpha \in H_d(SC, \partial B \times I)$ with $0 \neq \partial \alpha \in H_{d-1}(\partial B \times I)$, then no evasion path exists.
- Coordinate-free.
- Not sharp. Can it be sharpened?

*Coordinate-free Coverage in Sensor Networks with Controlled Boundaries via Homology* by V. de Silva and R. Ghrist
Topology applied to sensor networks

• Theorem (de Silva, Ghrist). If there is an $\alpha \in H_d(SC, \partial B \times I)$ with $0 \neq \partial \alpha \in H_{d-1}(\partial B \times I)$, then no evasion path exists.

• Coordinate-free.

• Not sharp. Can it be sharpened?
Zigzag persistent homology

Form zigzag module for $X \rightarrow I$ with $(d - 1)$–dimensional homology.
Zigzag persistent homology

Form zigzag module for $X \to I$ with $(d-1)$–dimensional homology.
Zigzag persistent homology

Form zigzag module for $X \to I$ with $(d - 1)$–dimensional homology.
Zigzag persistent homology

Form zigzag module for $X \to I$ with $(d - 1)$–dimensional homology.
Zigzag persistent homology

Form zigzag module for $X \rightarrow I$ with $(d-1)$-dimensional homology.
Zigzag persistent homology

Form zigzag module for $X \rightarrow I$ with $(d - 1)$–dimensional homology.
Zigzag persistent homology

Form zigzag module for $X \rightarrow I$ with $(d-1)$-dimensional homology.
Zigzag persistent homology

Form zigzag module for $X \to I$ with $(d-1)$-dimensional homology.
Zigzag persistent homology

Form zigzag module for \( SC \to I \) with \((d - 1)\)-dimensional homology.
Zigzag persistent homology

Form zigzag module for $SC \rightarrow I$ with $(d - 1)$-dimensional homology.
Zigzag persistent homology

Form zigzag module for $SC \to I$ with $(d - 1)$-dimensional homology.
Zigzag persistent homology

Form zigzag module for $SC \to I$ with $(d - 1)$–dimensional homology.
Zigzag persistent homology

Form zigzag module for $SC \rightarrow I$ with $(d - 1)$-dimensional homology.

- **Theorem.**
  
  If there is an evasion path then there is a full-length bar.
Zigzag persistent homology

Form zigzag module for $SC \rightarrow I$ with $(d-1)$-dimensional homology.

- **Theorem.**
  
  If there is an evasion path then there is a full-length bar.
Zigzag persistent homology

Form zigzag module for $SC \rightarrow I$ with $(d-1)$-dimensional homology.

- **Theorem.**
  
  If there is an evasion path then there is a full-length bar.

- Streaming computation.
Dependence on embedding $X \hookrightarrow B \times I$

- The time-varying Čech complex of $X$ does not determine if an evasion path exists!
Dependence on embedding $X \hookrightarrow B \times I$

- The time-varying Čech complex of $X$ does not determine if an evasion path exists!
Dependence on embedding $X \hookrightarrow B \times I$

- The time-varying Čech complex of $X$ does not determine if an evasion path exists!
Dependence on embedding $X \hookrightarrow B \times I$

- The time-varying Čech complex of $X$ does not determine if an evasion path exists!
Dependence on embedding $X \hookrightarrow B \times I$

- The time-varying Čech complex of $X$ does not determine if an evasion path exists!
Dependence on embedding $X \hookrightarrow B \times I$
Dependence on embedding $X \hookrightarrow B \times I$

- The two covered regions are “topologically indistinguishable in a time-preserving way”, but the uncovered regions are not!
Zigzag persistence

- Caution 2.9 of *Zigzag Persistence*. Not every submodule isomorphic to an interval corresponds to a summand.
Zigzag persistence

- Caution 2.9 of *Zigzag Persistence*. Not every submodule isomorphic to an interval corresponds to a summand.
Dependence on embedding $X \hookrightarrow B \times I$

- The two covered regions are “topologically indistinguishable in a time-preserving way”, but the uncovered regions are not!
Fat graphs

- A fat graph structure specifies a cyclic ordering of edges about each vertex (left).
- Equivalent to a set of boundary cycles (right).
Planar sensors measuring cyclic orders

- **Theorem.** In a planar sensor network that remains connected, the time-varying alpha complex with rotation information determines if an evasion path exists.
Planar sensors measuring cyclic orders

• Theorem. In a planar sensor network that remains connected, the time-varying alpha complex with rotation information determines if an evasion path exists.
Planar sensors measuring cyclic orders

• Theorem. In a planar sensor network that remains connected, the time-varying alpha complex with rotation information determines if an evasion path exists.
Planar sensors measuring cyclic orders

- **Theorem.** In a planar sensor network that remains connected, the time-varying alpha complex with rotation information determines if an evasion path exists.

![Diagram showing planar sensors and evasion paths]
Planar sensors measuring cyclic orders

- **Theorem.** In a planar sensor network that remains connected, the time-varying alpha complex with rotation information determines if an evasion path exists.
Planar sensors measuring cyclic orders

• **Theorem.** In a planar sensor network that remains connected, the time-varying alpha complex with rotation information determines if an evasion path exists.
Planar sensors measuring cyclic orders

- **Theorem.** In a planar sensor network that remains connected, the time-varying alpha complex with rotation information determines if an evasion path exists.
Planar sensors measuring cyclic orders

- Theorem. In a planar sensor network that remains connected, the time-varying alpha complex with rotation information determines if an evasion path exists.
Planar sensors measuring cyclic orders

- Theorem. In a planar sensor network that remains connected, the time-varying alpha complex with rotation information determines if an evasion path exists.
Planar sensors measuring cyclic orders

- **Theorem.** In a planar sensor network that remains connected, the time-varying alpha complex with rotation information determines if an evasion path exists.
Planar sensors measuring cyclic orders

- **Theorem.** In a planar sensor network that remains connected, the time-varying alpha complex with rotation information determines if an evasion path exists.
Planar sensors measuring cyclic orders

- **Theorem.** In a planar sensor network that remains connected, the time-varying alpha complex with rotation information determines if an evasion path exists.
Planar sensors measuring cyclic orders

- **Theorem.** In a planar sensor network that remains connected, the time-varying alpha complex with rotation information determines if an evasion path exists.
Planar sensors measuring cyclic orders

- **Theorem.** In a planar sensor network that remains connected, the time-varying alpha complex with rotation information determines if an evasion path exists.
- **Open question.** Is the Čech complex with rotation information sufficient?
Planar sensors measuring cyclic orders

- **Theorem.** In a planar sensor network that remains connected, the time-varying alpha complex with rotation information determines if an evasion path exists.

- **Open question.** Is the Čech complex with rotation information sufficient?
Planar sensors measuring cyclic orders

- **Theorem.** In a planar sensor network that remains connected, the time-varying alpha complex with rotation information determines if an evasion path exists.

- **Open question.** Is the Čech complex with rotation information sufficient?
Conclusions for Part III

• There is a streaming one-sided criterion for the evasion problem using zigzag persistence.

• Čech complex insufficient.  
  Alpha complex with rotation information suffices.  
  What about the Čech complex with rotation information?
Where can I find resources if I am interested in applied topology?

- You may be interested in the Applied Algebraic Topology Research Network. Become a member to receive email invites to the online research seminars. Recorded talks are available at the YouTube Channel. There is also a forum.
- Another source of applied topology news is appliedtopology.org.
- A second online research seminar is GEOTOP-A: Applications of Geometry and Topology.
- Mailing lists with announcements in applied topology include WinCompTop and ALGTOP-L.

https://www.math.colostate.edu/~adams/advising
Persistent homology software tutorial

https://github.com/henryadams/Leiden-PersistentHomology

• Examples using Ripser-live in your html browser.
• Cyclo-octane molecule, optical images
Cyclo-Octane (C$_8$H$_{16}$) data

Figure 7: Conformation Space of Cyclo-Octane. Here we show how the set of conformations of cyclo-octane can be represented as a surface in a high dimensional space. On the left, we show various conformations of cyclo-octane as drawn by PyMol (www.pymol.org). In the center, these conformations are represented by the 3D coordinates of their atoms. The coordinates are concatenated into vectors and shown as columns of a data matrix. As an example, the entry $c_{1,1,x}$ of the matrix denotes the $x$-coordinate of the first carbon atom in the first molecule. On the right, the Isomap method is used to obtain a lower dimensional visualization of the data.

To reduce complexity and avoid potential error due to hydrogen placement, we used only ring atoms to obtain a dataset \{x_i\}$_{i=1}^{1,031,644}$ ⊂ $\mathbb{R}^{24}$. We applied our algorithm to this dataset using parameters $\epsilon = 0.23$, $dt = 0.05$, $dp = 0.01$, and $\epsilon_p = 0.02$. We used five different values of $dt$, given by 0.08, 0.09, 0.10, 0.11, and 0.12. We produced five different triangulations with 6,040; 7,114; 8,577; 10,503; and 13,194 vertices. We used the Plex and Linbox toolboxes to check the accuracy of the triangulations. For each of the five triangulations, we verified that every neighborhood $B_i$ (before decomposition) had Betti numbers 1,0,0. This is an accuracy check because any neighborhood $B_i$ should be homotopic to a point and should therefore have Betti numbers 1,0,0. We also computed Betti numbers for each of the five full triangulations. In all cases we found the Betti numbers to be 1,1,2. This consistency strongly suggests that the triangulations are all representative of the actual conformation space. A visualization of the triangulation with 6,044 vertices using the Isomap coordinate representation.

Non-Manifold Surface Reconstruction from High Dimensional Point Cloud Data by Shawn Martin and Jean-Paul Watson, 2010.
Cyclo-Octane ($C_8H_{16}$) data

Persistence intervals in dimension 0:

Persistence intervals in dimension 1:

Persistence intervals in dimension 2:

Non-Manifold Surface Reconstruction from High Dimensional Point Cloud Data
by Shawn Martin and Jean-Paul Watson, 2010.
3x3 High-contrast patches from images

3x3 High-contrast patches from images

Interpretation: nature prefers horizontal and vertical directions
Devanagari Example

The following image of Devanagari script, which is used in India and Nepal, is from wikipedia (https://en.wikipedia.org/wiki/Devanagari). We will use a dataset with 36 different characters.
What applied topology software options do I have?

This list undoubtably has unintentional omissions; please email Henry with updates to this list!

https://www.math.colostate.edu/~adams/advising
Persistent homology software tutorial

https://github.com/henryadams/Leiden-PersistentHomology

• Examples using Ripser-live in your html browser.
• Cyclo-octane molecule, optical images