Research Statement

Henry Adams (henrya@math.stanford.edu)

I am interested in applied and computational topology, and in particular, applications to sensor networks and to data analysis.

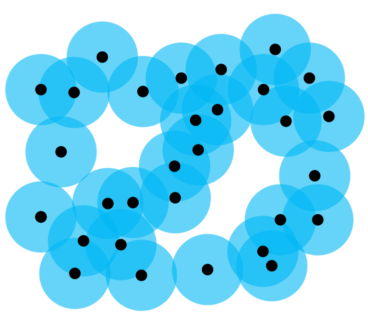
In sensor network problems one combines local measurements from individual sensors to answer a global question, and topology can be useful for this passage from local to global. For example, imagine that disk-shaped sensors wander in a planar domain. A sensor can't measure its location but does know when it overlaps a nearby sensor. We say that an evasion path exists if a moving evader can avoid being detected by the sensors. Can we determine if an evasion path exists? This is the evasion problem (§1). De Silva and Ghrist give a necessary homological condition for an evasion path to exist [dSG06], but it turns out that homology alone is insufficient. I provide necessary and sufficient conditions for the existence of an evasion path using sensors with stronger capabilities.

The evasion problem motivates a natural extension: can we describe the entire space of evasion paths? This question is interesting from a theoretical point of view, though it isn't as important for applications to real-world sensor networks. To study the space of evasion paths I use an analogue of the J. F. Adams spectral sequence for diagrams of spaces (§2).

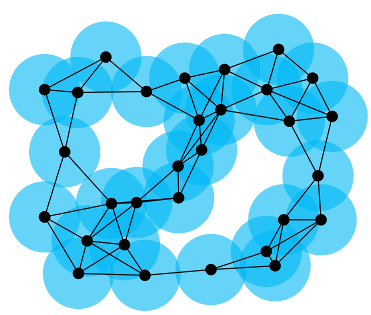
A second application of topology is to data analysis. Large sets of high-dimensional data are common in most branches of science and engineering, and their shapes reflect important patterns within. Sometimes these shapes are nonlinear and difficult to detect with traditional tools. The goal of topological data analysis is to produce simple combinatorial descriptors for data sets with interesting shapes [Car09]. In §3 I discuss an approach, inspired by Morse theory, to model the dense regions of a data set with cell complexes. I test this method on data arising in social networks, in image processing, and in microarray analysis.

1. The Evasion Problem

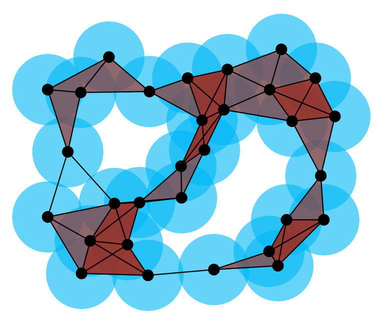
Suppose that disk-shaped sensors wander in a simply-connected planar domain D over time interval $I = [0, 1]$. The sensors don't know their locations but do measure their time-varying connectivity graph. Assume that immobile sensors cover the boundary ∂D and that the network remains connected. Let $X \subset D \times I$ be the subset of spacetime covered by sensors, and let $X^c = (D \times I) \setminus X$ be the uncovered region. Both X and X^c are *fibrewise spaces*, that is, spaces equipped with projection maps $X \to I$ and $X^c \to I$ to time. An evasion path is a section $I \to X^c$ of the projection map, and an evasion path exists precisely when a moving evader can avoid being seen by the sensors.



Sensor network at fixed time.



Connectivity graph.



Simplicial complex model.

The Evasion Problem. *The time-varying connectivity graph of a sensor network determines, to a close approximation, the fibrewise homotopy type of covered region X. Using only this input, is it possible to determine if an evasion path exists?*

De Silva and Ghrist prove the following theorem in [dSG06].

Theorem (de Silva, Ghrist). *If there is an evasion path in the sensor network, then every $[\alpha] \in H_2(X, \partial D \times I)$ satisfies $0 = [\partial\alpha] \in H_1(\partial D \times I)$.*

Their homological condition is necessary but not sufficient: the bottom network in Figure 1 satisfies the condition but contains no evasion path.

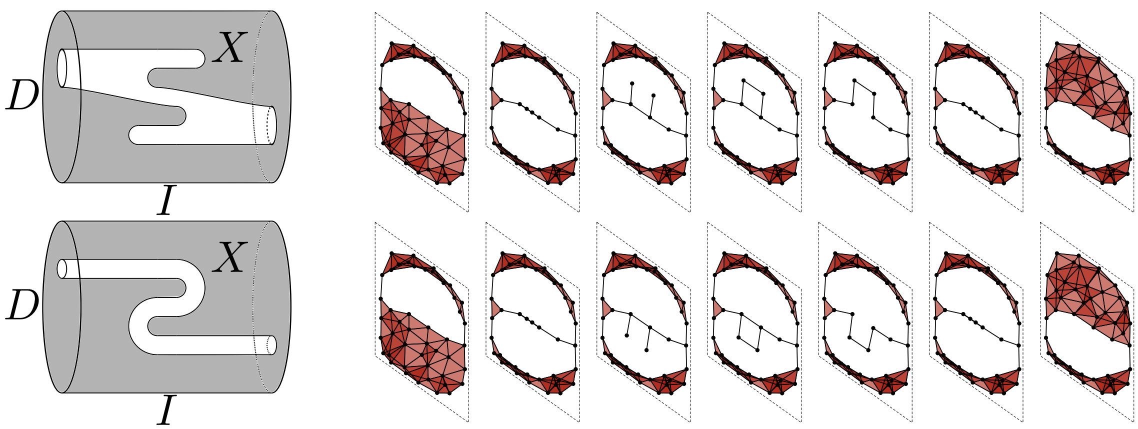


FIGURE 1. Each row is a sensor network, represented both as a covered region X in spacetime $D \times I$ and by seven sequential connectivity graphs. The top network contains an evasion path but the bottom one does not since the evader cannot travel backwards in time.

I began working on the evasion problem with the goal of finding an if-and-only-if criterion based on zigzag persistence, a generalization of persistent homology (the maps can go either direction) and a special case of quiver theory [CdS10]. A zigzag diagram of spaces is

$$Y_1 \to Y_2 \leftarrow Y_3 \to \ldots \leftarrow Y_{n-1} \to Y_n$$

with each Y_i a space and each arrow a continuous map. One can think of zigzag persistence as a homology functor that instead accepts a zigzag diagram of spaces as input. We can build a zigzag diagram of spaces that models fibrewise space $X \to I$, and in our setting zigzag persistence describes how the first-dimensional homology of the region covered by sensors changes with time. It is an invariant of the fibrewise homotopy type of X.

I discovered that the answer to the evasion problem is no: neither the fibrewise homotopy type of covered region X nor any invariants thereof (such as zigzag persistence) determine if an evasion path exists. The fibrewise embedding of X in $D \times I$ also matters. This is demonstrated by the two networks in Figure 1. Their covered regions X are fibrewise homotopic and their time-varying connectivity graphs are equal, but the top network contains an evasion path while the bottom one does not.

What minimal sensing capabilities might we add to get necessary and sufficient conditions? It is reasonable for real-world applications to assume that a sensor can measure the cyclic order of its neighbors [GLPS06], and local distances may be approximated by time-of-flight. I prove the following result.

Theorem. *If each sensor measures the cyclic order of its neighbors and the lengths of its adjacent edges in the sensor network's connectivity graph, then we can determine if an evasion path exists.*

The proof uses local distances to reconstruct the time-varying alpha simplicial complex, a subcomplex of the Delaunay triangulation that is homotopy equivalent to the nerve of the sensor disks. The cyclic order data gives the 1-skeleton of the alpha complex the structure of a fat graph or ribbon graph, whose boundary cycles track how the connected components of the uncovered region merge, split, appear, and disappear. It is an open question (§4.1) if the local distance measurements are necessary.

2. The Space of Evasion Paths

The evasion problem from §1 motivates a natural extension: can we describe the space of evasion paths? That is, what information must we measure about covered region X and its fibrewise embedding in spacetime $D \times I$ to describe the space of sections $I \to X^c$? This extension is not as important in applications but is interesting from a theoretical point of view. In this section I describe how to study the space of sections using a J. F. Adams spectral sequence for diagrams of spaces. In §4.2 on future work I discuss the remaining step of obtaining the input to our spectral sequence (as unstable invariants of uncovered region X^c) from embedding invariants of covered region X.

The unstable Adams spectral sequence for a space Y has as its E_2-term an unstable Ext depending only on $H_*(Y, \mathbb{Z}/p\mathbb{Z})$ as a coalgebra over the Steenrod algebra [BK72b], and often converges to $\pi_*(Y)$ modulo torsion prime to p. We study an analogous spectral sequence for a zigzag diagram of spaces $\mathcal{Y}$:

$$\mathcal{Y} = \quad Y_1 \to Y_2 \leftarrow Y_3 \to \ldots \leftarrow Y_{n-1} \to Y_n.$$

In particular, pick $\mathcal{Y}$ to model uncovered region $X^c \to I$. Under favorable circumstances the spectral sequence for $\mathcal{Y}$ converges to information about the space of sections $I \to X^c$.

We construct our Adams spectral sequence for diagrams as follows. The forgetful functor that discards all maps in a zigzag diagram of spaces has a right adjoint, and this adjunction defines a monad (or triple) T. Let Z be the monad that maps a based simplicial set Y to the $\mathbb{Z}/p\mathbb{Z}$–module generated by the simplices of Y [BK72b]. We combine monads T and Z to get a monad on the category of zigzag diagrams of spaces, and our Adams spectral sequence for diagrams is the homotopy spectral sequence of a cosimplicial object [BK72a] built from this resulting monad. We identify the E_2-term algebraically using a derived functor on the category of diagrams of unstable coalgebras.

3. Morse Theory in Topological Data Analysis

The analysis of high-dimensional data is a fundamental problem in many branches of science and engineering. In [AAC] we introduce a method, inspired by Morse theory, for data sets that are genuinely nonlinear and difficult to study with traditional tools. We adapt Morse theory to the setting of point clouds, i.e. finite sets of points in Euclidean space, using a kernel density estimator as the analogue of the Morse function. We sample cells from the skeleton of the Morse complex with the nudged elastic band method from computational chemistry [JMJ98]. The result is an increasing sequence of cell complexes modeling the dense regions of the data. In accordance with the idea of topological persistence, this output gives a more accurate representation of the data than the choice of any single complex.

We test our Morse-based approach on a variety of data sets, including sets arising in social networks, image processing, and microarray analysis. We find compact complexes that reveal important nonlinear patterns and assist in our qualitative understanding of the data, as in Figure 2.

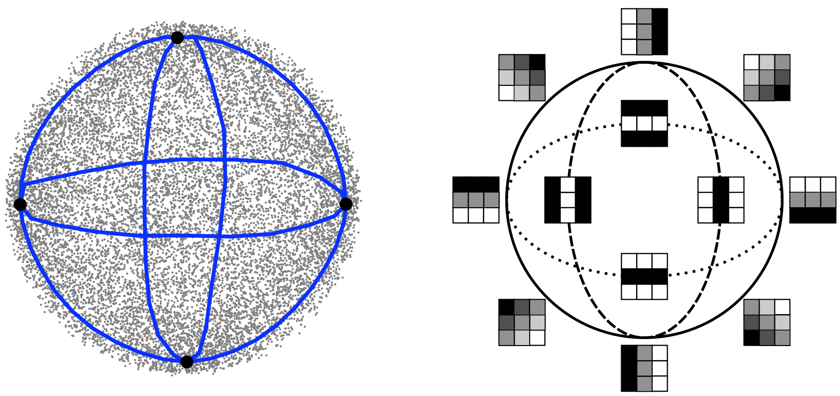


FIGURE 2. *(Left)* An image processing data set of 3×3 pixel patches, and our cellular model containing four 0-cells and eight 1-cells. *(Right)* The model's interpretation. The most common non-constant 3×3 patches are linear gradients at all angles and quadratic gradients in the preferred horizontal and vertical directions.

4. Future Research

I enjoy not only applied problems in computational topology but also the theoretical questions they motivate. Below I discuss three avenues for future research.

4.1. **The Evasion Problem.** At the end of §1 on the evasion problem I ask whether local distances are necessary for the theorem. This question is rephrased below.

Open Question. *The time-varying connectivity graph of a sensor network determines, up to close approximation, the fibrewise homotopy type of covered region X. Using only this input and the cyclic order of the edges about each sensor, is it possible to determine if an evasion path exists?*

An answer would fill the gap between the theorem of [dSG06], which is not sharp but uses only minimal sensor capabilities, and my result, which is sharp but requires more advanced sensors that measure local distances. This question seems to depend only on the computational geometry of planar disks, but it has proven challenging so far.

4.2. **The Space of Evasion Paths.** In §2 I study the space of evasion paths with a spectral sequence for diagrams that depends on unstable invariants of uncovered region X^c. It remains to obtain these unstable invariants from embedding invariants of covered region X. One idea is to use the tools of embedding calculus [Wei99] in a fibrewise setting. In particular, let

$\mathrm{Emb}(X, D \times I)$ be the space of all fibrewise embeddings of X in spacetime $D \times I$. Note that a point in $\mathrm{Emb}(X \coprod I, D \times I)$ encodes both the data of an embedding $X \hookrightarrow D \times I$ and an evasion path $I \to X^c$. Hence I am interested in studying the fibre of the restriction map

$$\mathrm{Emb}(X \coprod I, D \times I) \to \mathrm{Emb}(X, D \times I).$$

4.3. **Persistent Homology and Data Analysis.** I am a coauthor of a tutorial on javaPlex, a software package for persistent homology [TVJA11], and I also have experience with persistent homology in data analysis, in particular range images [AC09]. As a result, I serve as a contact person for users of javaPlex who have questions about how to use the software with their data. I have received questions from scientists in a wide range of disciplines, including mathematics, computer science, biology, medicine, neuroscience, chemical engineering, electrical engineering, physics, and economics. I enjoy this role because it benefits my research community and introduces me to potential future collaborators.

5. Undergraduate Research Mentoring

Applied and computational topology is an attractive field for undergraduate research. Applied topology projects show students that math can be useful for other branches of science. Moreover, such projects help students understand topology better, since abstract topics like functoriality play key roles in the more concrete and applied settings. I am prepared to mentor projects in topological data analysis, and in particular, I am a coauthor of a tutorial on javaPlex, a sotware package for persistent homology. The tutorial is written at a level appropriate for advanced undergraduates and contains examples, exercises, and solutions. In addition, I think that there are many sensor network problems that are easy to state and interesting to explore from a topological point of view.

References

[AAC] H. Adams, A. Atanasov, and G. Carlsson. Morse theory in topological data analysis. Submitted December 2011, `arXiv:1112.1993`.

[AC09] H. Adams and G. Carlsson. On the nonlinear statistics of range image patches. *SIAM J. Imag. Sci.*, 2:110–117, 2009.

[BK72a] A. K. Bousfield and D. M. Kan. *Homotopy Limits, Completions and Localizations*, volume 304 of *Lecture Notes in Mathematics*. Springer, Berlin, 1972.

[BK72b] A. K. Bousfield and D. M. Kan. The homotopy spectral sequence of a space with coefficients in a ring. *Topology*, 11:79–106, 1972.

[Car09] G. Carlsson. Topology and data. *Bull. Amer. Math. Soc.*, 46:255–308, 2009.

[CdS10] G. Carlsson and V. de Silva. Zigzag persistence. *Found. Comput. Math.*, 10:367–405, 2010.

[dSG06] V. de Silva and R. Ghrist. Coordinate-free coverage in sensor networks with controlled boundaries via homology. *Int. J. Robot. Res.*, 25:1205–1222, 2006.

[GLPS06] R. Ghrist, D. Lipsky, S. Poduri, and G. Sukhatme. Surrounding nodes in coordinate-free networks. In *In Workshop in Algorithmic Foundations of Robotics*, 2006.

[JMJ98] H. Jónsson, G. Mills, and K. W. Jacobsen. Nudged elastic band method for finding minimum energy paths of transitions. In B.J. Berne, G. Ciccotti, and D.F. Coker, editors, *Classical and Quantum Dynamics in Condensed Phase Systems*, pages 385–404. World Scientific, Singapore, 1998.

[TVJA11] A. Tausz, M. Vejdemo-Johansson, and H. Adams. Javaplex: A research software package for persistent (co)homology. Software available at `http://code.google.com/javaplex`, 2011.

[Wei99] M. Weiss. Embeddings from the point of view of immersion theory: Part I. *Geometry and Topology*, 3:67–101, 1999.