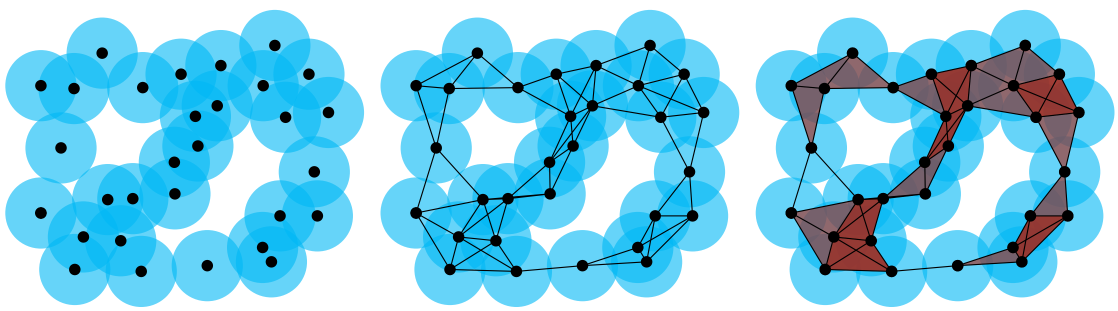
Research Statement

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.

1. The Evasion Problem in Mobile Sensor Networks

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, 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.

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 equivalent, 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 reconstructs a parameterized 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 if the local distance measurements are necessary.

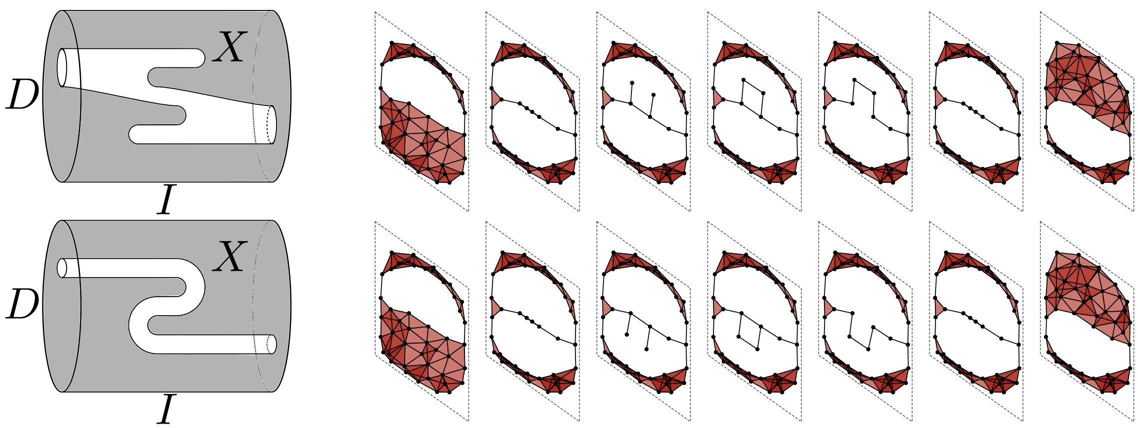


FIGURE 1. Each row is a sensor network, represented both as a covered region X in spacetime $D \times I$ and by seven sequential connectivity graphs. The top network contains an evasion path but the bottom one does not since the evader cannot travel backwards in time.

2. The Space of Evasion Paths

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 of the evasion problem from §1 is not as important for applications to real-world sensor networks, but it 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.

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

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 [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 approach on a variety of data sets, including sets arising in social networks, image processing, and microarray analysis, and find compact complexes that reveal important nonlinear patterns and assist in our qualitative understanding of the data.

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.

[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.