RESEARCH STATEMENT

Rachel Levanger (rachel@math.rutgers.edu)

I specialize in the study of high-dimensional spatiotemporal data using tools from the field of topological data analysis (TDA). For the past three years, I have studied data generated both by experiments and simulations of fluid flows (convection, combustion, and fully-developed turbulence), and have used these applications to push the development of new theoretical results in TDA. My methods primarily rely on a tool called persistent homology, which provides dimensionality reduction while retaining important topological features of the original data. Each data point is transformed into an object called a persistence diagram, which encodes the topological features of the data as points in the Euclidean plane. The persistence diagrams, which are computationally less expensive to work with than the original high-dimensional data, are then studied using established techniques in TDA or used as features for machine learning algorithms.

Completed work

As a PhD candidate at Rutgers University under the direction of Konstantin Mischaikow, one of my first theoretical contributions was to solve an open problem in persistent homology concerning the foundations of numerical analysis on persistence diagrams. Up until my work, numerical errors introduced during the generation of the original data or the computation of the persistence diagrams were stated only in terms of a uniform maximum error. The framework I developed provides a way to state local errors in a persistence diagram in addition to providing an easy method for computing the errors resulting from multiple stages of approximation.

Our research group was then able to use this framework to give a more precise statement of the persistence diagram generated from almost-periodic dynamics in Rayleigh-Bénard convection flow. Our dataset concerned 70K data points, which was down-sampled due to computational constraints, and thus we were only able to compute an approximate persistence diagram. Had we used the uniform maximum error, we would not have been able to rigorously separate the feature representing the periodic orbit from computational noise. However, using the precise bounds from our framework made it clear that the topological features encoded in our persistence diagram were clearly separated from any computational artifacts introduced by the approximation.

Current research

I am actively working on numerous applied problems in TDA.

- Classifying defect patterns in Spiral Defect Chaos. Collaboration with the Schatz Labs at Georgia Tech and the Paul Research Group at Virginia Tech.
- Development of an interactive tool for studying persistence diagrams generated by image data or any scalar field on a 2D domain.

- Studying the vorticity fields of fully-developed turbulence using persistent homology and diffusion maps. Using persistent homology to study the low-temperature oxidation regime of a large-scale simulation of 33 chemical species undergoing homogeneous charge compression ignition. Collaboration with Takashi Ishihara at Nagoya University.

On the theoretical side, I am working toward a framework that will enable the rigorous study of temporal data in the space of persistence diagrams. Presently, even though the transformation from an input dataset to the space of persistence diagrams is continuous, it is not possible to mathematically track temporal changes in the space of persistence diagrams. That is, combinatorial tracking at the level of the diagrams, while certainly possible, is not guaranteed to match corresponding features in the underlying data. To date, I have developed a heuristic algorithm that, in the case of image data that varies continuously with respect to time and that is sufficiently well sampled in time, gives a correct matching for tracking topological features. I am working to find an underlying algebraic framework that will give rigorous, mathematical guarantees for when this type of tracking is possible, and also algorithms for computing accurate topological trajectories in the space of persistence diagrams.

Future work

When I first entered graduate school, my ultimate goal was to combine TDA with statistical methods and machine learning algorithms to find complex patterns in large datasets (to this end, the minor topic in my oral exam was probability). When I applied to graduate schools, this aspect of TDA was still on the horizon, and while I've concentrated the last four or five years on building my knowledge of persistent homology and its applications, the larger community has continued to make in-roads on the statistical and machine learning aspects of the field.

To jump-start the next phase of my research, I am studying advanced techniques in Bayesian inference for latent variable models. I am currently working on developing a statistical model for classifying topological trajectories in the space of persistence diagrams for a simulated dataset containing 1500 frames of Rayleigh-Bénard convection. In the future, I will extend this model to experimental data in an effort to develop a predictive model for convection patterns—models for small-scale weather phenomena.

I hope to continue to combine techniques in TDA with machine learning approaches to build predictive models for large, high-dimensional datasets. While I have focused primarily on applications to fluid flows, there are many other datasets that are natural to study using persistent homology (e.g. connectivity matrices and point cloud data). Once the topological transformation to the space of persistence diagrams has taken place, most of the theory I have developed or worked on is immediately transferrable to these other settings. I also look forward to continuing to push the theory of persistent homology to solve theoretical problems presented by these other domains.