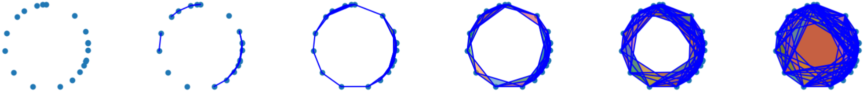
Differentiability and Optimization of Multiparameter Persistent Homology

Luis Scoccola, Siddharth Setlur, David Loiseaux, Mathieu Carrière, Steve Oudot

ICML, July 2024

Single Parameter Persistence



- A tool to analyze discrete data by studying the underlying shape of data.
- Construct a *family of topological spaces* (usually a filtration) and then use tools from algebraic topology to track how topological features evolve across the family.
- This information can be completely represented by a *persistence barcode*, a collection of points $(b, d) \in \mathbb{R}^2$ each of which represents the lifespan of a topological feature.

Single Parameter Persistence is Differentiable

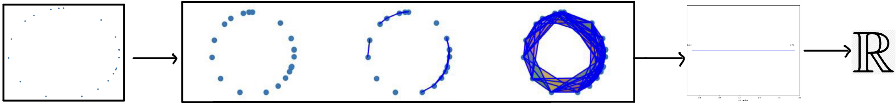
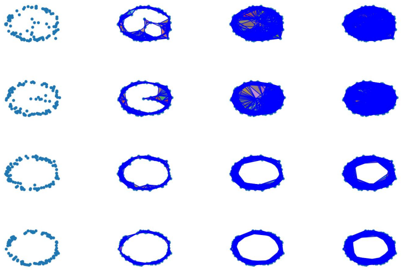


Figure: The single parameter persistence pipeline

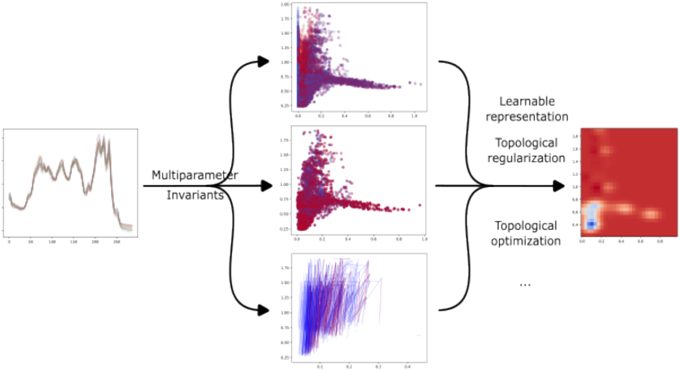
- [Car+21] proved that the above pipeline is differentiable almost everywhere and that stochastic gradient descent (SGD) converges almost surely to a critical point.
- Allows for persistence based loss functions to be incorporated into the machine learning pipeline.
- Many successful applications in fields such as drug discovery [Dem+22], topology preserving autoencoders [Moo+20] among others.

Multiparameter Persistence



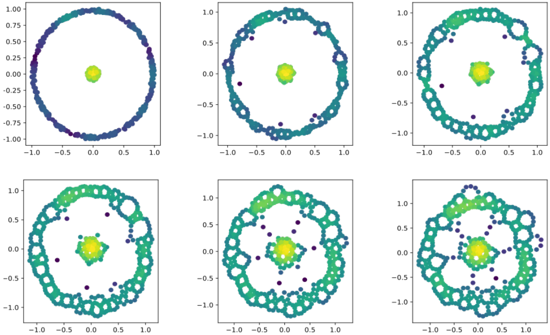
- There is often more than one option to filter the topology of data. This leads to multifiltrations.
- Due to representation theoretic reasons, there is no analog to the barcode that can completely characterize the topology of multifiltrations.
- Multiple incomplete invariants have been proposed - Hilbert decompositions, signed barcodes, and Euler characteristics among others.

Our Main Result



- Our framework generalizes the differentiability results from the single parameter case to the multiparameter pipeline.
- Crucially, results hold for a *large class of invariants* of multiparameter persistence. Multiparameter persistence based loss functions can now be used in machine learning!

Demonstrative Example - Point Cloud Optimization



- Goal - Use 2-parameter (radius, density) persistence to create cycles using the dense points in the center of the main circle, while maintaining the main outer circle.
- In contrast, the 1-parameter analog of the experiment (see paper) destroys the outer circles and creates many similar sized circles.

Summary

- A novel, general framework for differentiating and optimizing loss functions based on multiparameter persistence.
- Proof of concept on applications such as graph classification, time series classification, topological autoencoders, and point cloud optimization.

Thank you!

Drop by - Hall C 4-9 #1112

Wed 24 July 1:30 p.m. - 3 p.m. CEST

Paper

Code