

Union College Math Conference: Applied Topology and Geometry

June 3–5, 2022

FRIDAY PROGRAM

4:30–5:30pm: Reception (Olin Rotunda)

5:30–6:30pm: Plenary talk: Tai Melcher *Title TBA* (Olin 115)

SATURDAY PROGRAM

8:00–9:00am: Coffee & pastries, registration (Olin Rotunda)

9:00–10:30am: Session II (ISEC 218)

- 9:00–9:25 Chad Giusti: *Are these topological features analogous?*
- 9:30–9:55 Melinda Kleczynski: *Detecting Related Topological Features Across Samples of a Time-Varying System*
- 10:00–10:25 Chris Tralie: *Discovering And Perfecting Periodicity in Videos using Topological Autoencoders*

10:30–11:00am: Coffee break (Olin Rotunda)

11:00am–12:00pm: Plenary talk: Yusu Wang, *Weisfeiler-Lehman Meets Gromov-Wasserstein* (Olin 115)

12:00–2:00pm: Lunch break

2:00–4:00pm: Session III (ISEC 218)

- 2:00–2:25 Nick Scoville: *The homotopy type of the Morse complex for some collections of trees*
- 2:30–2:55 Youjia Zhou: *Topological Simplifications of Hypergraphs*
- 3:00–3:25 Jerome Roehm: *Anti-Geometric Persistence*
- 3:30–3:55 Tung Lam: *ℓ^p metric on merge trees*

4:00–4:30pm: Coffee break (Olin Rotunda)

4:30–5:30pm: Plenary talk: Claude LeBrun, *Four-Manifolds, Conformal Curvature, and Differential Topology* (Olin 115)

SUNDAY PROGRAM

8:00–8:30am: Coffee & pastries (Olin Rotunda)

8:30–10:30am: Session IV (ISEC 218)

- 9:00–9:25 Mikael Vejdemo-Johansson: *Multiple Hypothesis Testing with Persistent Homology*
- 9:30–9:55 Michael Lesnick: *Dual functional inequalities and optimal transport*
- 10:00–10:25 Sarah Percival: *Efficient Computation of a Semi-Algebraic Basis of the First Homology Group of a Semi-Algebraic Set*
- 10:00–10:25 Sarah Percival: *Statistical Inference and Parallel Transport of Probability*
- 10:30–10:55 Yu Qin: *A Domain-Oblivious Approach for Learning Concise Representations of Filtered Topological Spaces*

11:00–11:00am: Coffee break (Olin Rotunda)

11:00am–12:00pm: Plenary talk: Sergio Lopez-Permouth (Olin 115)

ABSTRACTS

Yusu Wang (University of California, San Diego)

Plenary Talk: *Weisfeiler-Lehman Meets Gromov-Wasserstein*

The Weisfeiler-Lehman (WL) test is a classical procedure for graph isomorphism testing. The WL test has also been widely used both for designing graph kernels and for analyzing graph neural networks. In this talk, I will describe the so-called Weisfeiler-Lehman (WL) distance we recently introduced, which is a new notion of distance between labeled measure Markov chains (LMMCs), of which labeled graphs are special cases. The WL distance extends the WL test (in the sense that the former is positive if and only if the WL test can distinguish the two involved graphs) while at the same time it is polynomial time computable. It is also more discriminating than the distance between graphs used for defining the Wasserstein Weisfeiler-Lehman graph kernel. Inspired by the structure of the WL distance we identify a neural network architecture on LMMCs which turns out to be universal w.r.t. continuous functions defined on the space of all LMMCs (which includes all graphs) endowed with the WL distance. Furthermore, the WL distance turns out to be stable w.r.t. a natural variant of the Gromov-Wasserstein (GW) distance for comparing metric Markov chains that we identify. Hence, the WL distance can also be construed as a polynomial time lower bound for the GW distance which is in general NP-hard to compute. This is joint work with Samantha Chen, Sunhyuk Lim, Facundo Memoli and Zhengchao Wan.

Chad Giusti (University of Delaware)

Are these topological features analogous?

The most common analysis pipeline in applied topology involves computing persistent homology of some filtered simplicial complex derived from data, then using the resulting persistence diagram as a feature for statistical or ML techniques. This is the default in part because it is very difficult to compare persistence modules computed from different data sets directly. In an ideal world, we would try to approximate a function between the systems underlying data sets and apply functoriality to build induced maps. However, in many applications it is unreasonable to assume the existence of such functions. Even if we have full control over and knowledge of one the state of one of the systems, hidden variables and noise in the other will usually result in a range of resulting behaviors and states in the other. Thus, the best we can hope for is to observe a relation between system states, in the guise of correlation or other cross-similarity. In this talk, we describe a method for leveraging such observations to compare persistent homology classes between related persistence diagrams. This is joint work with Iris Yoon and Robert Ghrist.

Melinda Kleczynski (University of Delaware)

Detecting Related Topological Features Across Samples of a Time-Varying System

Topological data analysis of successive sampling of a time-varying system generates a sequence of barcodes. Given two bars, each from a different barcode, it is reasonable to ask whether these bars represent two samples of the same feature. The barcodes alone don't provide a solution. It's possible, for example, to sample a topological feature in two ways such that the point clouds both visually capture the same feature, but the corresponding bars do not overlap at all. Maps between successive point clouds also don't provide a solution; such maps are generally not known for complex systems. Complementing recent work on analogous bars

by Yoon, Ghrist, and Giusti, we explore data-focused techniques for creating a filtration to search for similar features in images of a time-varying system. We discuss what it means for features to be similar and how to detect them using techniques including zigzag persistence. Potential applications include detection of certain aggregations of birds on weather surveillance radar. This is joint work with Dr. Chad Giusti and Dr. Jeffrey Buler through the University of Delaware Applied Topology Group and the University of Delaware Aeroecology Lab.

Tung Lam (UAlbany)

ℓ^p Metric on Merge Trees

We introduce an ℓ^p -type distance as an extension for the interleaving distance on merge trees by adapting the definition in Bjerkevik and Lesnick’s recent work for multiparameter persistence modules. Our distance coincides with the interleaving distance for $p = \infty$. We show that this distance is a metric on isomorphism classes of merge trees, and upper-bounds the p -Wasserstein distance on barcodes of associated merge trees. Building on an idea of Skraba and Turner on Cellular Wasserstein stability, we show that our distance is stable and in fact, universal for any $p \in [1, \infty]$. For the case $p = \infty$, this gives a novel proof of universality for the interleaving on merge trees. This is joint work with R. Cardona, J. Curry, and M. Lesnick.

Michael Lesnick (UAlbany)

Stability of 2-Parameter Persistent Homology

We show that the standard stability results for union-of-balls, ϵ -ech, and Rips persistent homology have close analogues in the 2-parameter setting, formulated in terms of the multicover bifiltration and Sheehy’s subdivision bifiltrations. Our results imply that these bifiltrations are robust, i.e., stable to outliers, in a strong sense. We also give similar stability results for degree bifiltrations; these results are weaker, but tight. I’ll also discuss computation of the bifiltrations we study and the implication of our results for the practical development of 2-parameter persistence. This is joint work with Andrew Blumberg.

Sarah Percival (Michigan State University)

Efficient Computation of a Semi-Algebraic Basis of the First Homology Group of a Semi-Algebraic Set

We give an algorithm for computing a semi-algebraic basis for the first homology group, $H_1(S, \mathbb{F})$, with coefficients in a field \mathbb{F} , of any given semi-algebraic set $S \subset \mathbb{R}^k$ defined by a closed formula, where \mathbb{R} is a real closed field. The complexity of the algorithm is bounded singly exponentially, that is, if the given quantifier-free formula involves s polynomials whose degrees are bounded by d , the complexity of the algorithm is bounded by $(sd)^{k^{O(1)}}$. This algorithm generalizes well known algorithms having singly exponential complexity for computing a semi-algebraic basis of the zero-th homology group of semi-algebraic sets, which is equivalent to the problem of computing a set of points meeting every semi-algebraically connected component of the given semi-algebraic set at a unique point.

Yu Qin (Tulane University)

A Domain-Oblivious Approach for Learning Concise Representations of Filtered Topological Spaces

Persistence diagrams have been widely used to quantify the underlying features of filtered topological spaces in data visualization. In many applications, computing distances between diagrams is essential; however, computing these distances has been challenging due to the

computational cost. In this paper, we propose a persistence diagram hashing framework that learns a binary code representation of persistence diagrams, which allows for fast computation of distances. This framework is built upon a generative adversarial network (GAN) with a diagram distance loss function to steer the learning process. Instead of using standard representations, we hash diagrams into binary codes, which have natural advantages in large-scale tasks. The training of this model is domain-oblivious in that it can be computed purely from synthetic, randomly created diagrams. As a consequence, our proposed method is directly applicable to various datasets without the need for retraining the model. These binary codes, when compared using fast Hamming distance, better maintain topological similarity properties between datasets than other vectorized representations. To evaluate this method, we apply our framework to the problem of diagram clustering and we compare the quality and performance of our approach to the state-of-the-art. In addition, we show the scalability of our approach on a dataset with 10k persistence diagrams, which is not possible with current techniques. Moreover, our experimental results demonstrate that our method is significantly faster with the potential of less memory usage, while retaining comparable or better quality comparisons.

Jerome Roehm (University of Delaware)

Anti-Geometric Persistence

The input to many persistent homology computation algorithms is a square symmetric ?dissimilarity? matrix, where the rows and columns correspond to vertices and the matrix entries describe the dissimilarity between the vertices. However, many data analysis pipelines produce similarity measures, and it is a common experience among TDA practitioners to forget to transform these into dissimilarity before computing persistence. If this reverse step is omitted from the workflow, the persistence diagrams resulting from the computations are often wildly different from the expectation. In this talk, we explore the persistence diagrams obtained by reversing the distance matrix of uniformly distributed points on a circle and discuss other questions that arise from reverse persistence.

Nick Scoville (Ursinus College)

The Homotopy Type of the Morse Complex for Some Collections of Trees

The Morse complex is the simplicial complex consisting of all gradient vector fields on a fixed simplicial complex K . The homotopy type of this complex is in general not well understood. In this talk, we give some recent results on the homotopy type of the Morse complex for certain collections of trees. We show that if K contains two leaves that share a common vertex, then the Morse complex is strongly collapsible and hence has the homotopy type of a point. We also use the result that the Morse complex of a disjoint union $K \sqcup L$ is the Morse complex of the join $K * L$ to compute the homotopy type of the Morse complex of some families of graphs, including Caterpillar graphs.

Chris Tralie (Ursinus College)

Discovering And Perfecting Periodicity in Videos using Topological Autoencoders

The Takens "sliding window" framework and TDA have formed a potent combination for analyzing periodic signals and videos. However, relatively little work has explored this framework for the purpose of *modifying* signals to promote periodic content, which could help with visualizing, denoising, and amplifying periodic motions. In this preliminary work, we leverage recent work on topological autoencoders and differentiable TDA pipelines to address these problems. We use a topological autoencoder consisting of a deep convolutional network with a Vietoris-Rips-based sliding window video embedding regularizer to promote periodic

fluctuations in our videos, while cutting down on unrelated background fluctuations and noise. We show some preliminary results on synthetic videos and real videos of hidden periodic color changes and motion, including ordinary webcam videos that pick up on a heartbeat via color fluctuations, and we conclude with some future directions. This is joint work with Umberto Lupo from EPFL.

Mikael Vejdemo-Johansson (CUNY College of Staten Island)

Multiple Hypothesis Testing with Persistent Homology

We propose a computationally efficient multiple hypothesis testing procedure for persistent homology. Most current approaches to hypothesis testing using persistent homology is based on resampling or permutation procedures, this is because there exists little distribution theory that provides a natural null model for persistence diagrams. In this paper we propose a null model based approach to testing for acyclicity, coupled with a Family-Wise Error Rate (FWER) control method that does not suffer from the computational costs associated to resampling and permutation based approaches. We adapt standard False Discovery Rate (FDR) control procedures to the topological setting, and use our null model for hypothesis testing using persistent homology. A key idea in our paper is that one can compute an empirical distribution under a null hypothesis model that can be used for hypothesis testing in the same way that a normal or t-distribution is used in classical statistics. An argument for the use of this empirical null is based on simulations and limit theorems for persistent homology for point processes.

Youjia Zhou (University of Utah)

Topological Simplifications of Hypergraphs

We study hypergraph visualization via its topological simplification. We explore both vertex simplification and hyperedge simplification of hypergraphs using tools from topological data analysis. In particular, we transform a hypergraph into its graph representations, known as the line graph and clique expansion. A topological simplification of such a graph representation induces a simplification of the hypergraph. In simplifying a hypergraph, we allow vertices to be combined if they belong to almost the same set of hyperedges, and hyperedges to be merged if they share almost the same set of vertices. Our proposed approaches are general and mathematically justifiable, and put vertex simplification and hyperedge simplification in a unifying framework.