Applied homological algebra beyond persistence diagrams

organized by
Chad Giusti, Gregory Henselman-Petrusek, and Lori Ziegelmeier

Workshop Summary

Applied Homological Algebra Beyond Persistence Diagrams

Chad Giusti, Gregory Henselman-Petrusek, and Lori Ziegelmeier

The unifying theme of this workshop was applied homological algebra, with a particular emphasis on moving away from summary statistics such as persistence diagrams to instead provide more detailed information about and explicit connections among complex data. The aim of this workshop was to stimulate new connections and new research collaborations, with the goal of facilitating the transfer of ideas, and hence long-term growth in applied topology. The workshop placed special emphasis on the interaction between makers and users of theory, algorithms, and code.

Each morning there were two long discussions that were intended to provide background and motivation of particular topics to stimulate ideas and open questions around the theme of the workshop and presentations. The speakers and topics are listed below:

- Monday:
  - Iris Yoon, cycle registration and analogous bars
  - Matthew Wright and Michael Lesnick, multiparameter persistence
- Tuesday:
  - Dmitriy Morozov, topological optimization and efficient updates
  - Barbara Giunti, parametrized chain complexes and other algebraic structures
- Wednesday:
  - Emilie Purvine, topology of hypergraphs
  - Brief Software Demos
    - Bei Wang, hypergraph visualization TopoAct and TopoBERT which applies mapper graphs to generic high-dimensional data
    - Michael Robinson, software library to teach students about chain complexes and persistence modules
    - Luis Scoccola, DREiMac data coordinatization with persistent cohomology and Persistable to cluster data with multiparameter persistence
    - Yossi Bleile, Correa for looking at structure of simple closed curves
    - Primoz Skraba, significance testing of individual cycles
    - Barbara Giunti, DONUT database of TDA applications in the form of a search engine
    - Arnur Nigmetov, Oineus parallel algorithm to compute persistence using shared memory
- Thursday:

  - Ling Zhou, ephemeral bars
  - Gregory Henselman-Petrusek, U-match factorization and persistence computations including a demo of the OAT software
- Friday:
  - Luis Scoccola, resolutions, decompositions, signed barcodes, and Mobius inversion
  - Panel to discuss structural barriers to entry into the field of applied topology, what resources are needed to make our field more accessible, how do researchers get credit for various things like software, and how can results get found?
    * Panelists: Bei Wang, Michael Robinson, Ines Garcia-Redondo, Vin de Silva, Gregory Henselman-Petrusek

On Monday afternoon, we generated a list of open problems centered around the workshop theme, compiling a list of 20 topics. The organizers synthesized these topics and settled on a set of nine questions, which individuals voted on to indicate interest. Six working groups ultimately formed. A high-level overview of each working group and what they accomplished during the week follows:

- **Topic/Question:** Given the persistence diagram of a filtered chain complex, how do we interpret the points on the diagonal?
  Participants: Barbara Giunti, Arnur Nigmetov, Bei Wang, Ling Zhou, Alyson Bittner, and Emilie Purvine
  Overview: In a persistence theory pipeline, input data are processed via filtered chain complexes, which are then simultaneously transformed into persistence modules and decomposed into barcodes. Traditionally, barcodes consist of bars of non-zero length. However, in practice, one handles filtered chain complexes instead of persistence modules. The current persistence algorithms actually output a decomposition of the filtered chain complexes, which include zero-length bars. Several existing works point toward the fact that these zero-length bars carry (geometric) information, but at the moment we do not have a precise mathematical formula to decode it.

  This working group aims to address the following question: What (geometric) information is encoded in the zero-length bars? They now understand that zero-length bars encode information about (local) diameter. Moving forward, they will explore a few directions. First, what notions of curvature/convexity of the space could be linked to the zero-length bars? Second, what are the practical applications of zero-length bars, for instance, in hypothesis testing of detecting outliers in a dataset? Third, what is the relation between zero-length bars and the doubling dimension of the underlying space?
- **Topic/Question:** Is there a higher-order Dowker theory for high-order products?
  Participants: Iris Yoon, Vladimir Itskov, Nikolas Schonsheck, Chad Giusti, Michael Robinson, Vin de Silva, Melvin Vaupel, and Radmila Sazdanovic
  Overview: This group generalized the famous Dowker theorem from binary relations to multi-way relations. They showed that there are three equivalent constructions, each of which yield natural "Dowker" theorems. One may derive the generalized Dowker complex via (1) a cuboid complex, (2) the cosections of a cosheaf, and via (3) homotopy colimits of a pushout diagram. While apparently quite distinct, they carry the same information, albeit structured in a different way. Specifically, both

the global cosections of the cosheaf and the homotopy colimit of the pushout diagram recover the cuboid complex.

- **Topic/Question:** Hypergraph structures using homology
  Participants: Emilie Purvine, Alyson Bittner, Helen Jenne, Peter Bubenik, and Vladimir Itskov
  Overview: This group convened to think about methods to understand hypergraph structures using homology. Since a hypergraph is not a topological space, simplicial complex, or other structure that homology can apply to directly a transformation must be made to turn it into one of these objects. There can be many ways to do this, each resulting in different homological features. There are two directions one can take on this topic. The first is noting which hypergraph structural features are of interest and designing a construction such that homology will capture those features. The second is coming up with interesting constructions and then understanding what they capture. This group focused on the latter direction. They started by surveying some of the prior work that has been done in this space. Later, Vladimir educated the group about his work on combinatorial codes, which are hypergraphs in disguise, and the polar complex, and Peter introduced a simplicial complex filtration based on hyperedge size. The group spent much of their time playing with these methods on small examples to gain intuition.
- **Topic/Question:** Cycle representatives and computable summaries
  Participants: Peter Bubenik, Dmitriy Morozov, Gregory Henselman-Petrusek, and Vladimir Itskov
  Overview A group met to discuss cycle representatives of a persistent homology class and computable and informative summaries of this set. Given a degree k persistent homology class (with coefficients in the field of two elements), Peter suggested the following problem. For each k-cell, compute the proportion of cycle representatives of the homology class containing it. Dmitriy showed that this question has a simple answer. If a k-cell is (part of) the boundary of a $(k+1)$-cell then the answer is $1/2$. If not, then it is either 1 or 0, which is determined by checking the representative cycle provided by the persistence algorithm. The group proceeded to consider deriving more informative summaries by considering Markov chain Monte Carlo methods (MCMC) by taking random walks on the set of cycle representative by adding boundaries using the matrix decomposition provided by the persistence algorithm. Gregory coded up some software to use this idea to provide a visualization of this summary. It seemed that the space of cycles may be too high dimensional for MCMC. Finally, the group discussed harmonic cycle representatives.
- **Topic/Question:** What does multiparameter persistence of a time-varying function tell us about the dynamics?
  Participants: Michael Lesnick, Lori Ziegelmeier, Yossi Bokor, John Emanuello, Dmitriy Morozov, Ben Cassidy, and Helen Jenne
  Overview: One working group focused on the study of time-varying data (i.e., "dynamic data") using multiparameter persistence. The group members have very diverse backgrounds and research focuses. In order to get everyone on the same page, our initial discussions surveyed
  (1) several types of time-varying data studied with TDA and
  (2) several natural constructions of bifiltrations and trifiltrations from dynamic data.

In the second discussion, Dmitriy Morozov proposed a novel and appealing definition of a bifiltration built from dynamic data, which makes sense when the data satisfies a Lipschitz property. He calls this the "vineyard bifiltration." The group was very interested in the vineyard bifiltration, and the remaining discussions focused on determining whether its homology (in all degrees) determines the homology of a certain more naive trifiltration built from the same data. While the group is still discussing this, the emerging picture is that the homology of the bifiltration does not fully determine the homology of the trifiltration. Even so, the group sees the vineyard bifiltration as a promising object in practical applications, and there are several potential directions for research involving this bifiltration, including computational questions.

- **Topic/Question:** What are typical properties of bifiltrations for a given probabilistic model?

  Participants: Matthew Wright, Luis Scoccola, Ines Garcia-Redondo, Primoz Skraba
  Overview: This group studied bifiltrations that arise from random geometric models, specifically Erdős–Rényi random graphs and geometric point processes. They considered a two-parameter Erdős–Rényi model on a complete graph K_n. Each edge of the graph is associated with two independent random variables, which is used as filtration functions f and g. They note that the one-parameter filtration given by the diagonal $f = g$ recovers a (reparameterization of a) standard Erdős–Rényi model. The group wanted to determine the existence of large indecomposables in the H_0 bipersistence module arising from this bifiltration. They proved that this module contains a non-interval indecomposable whose size can be quantified, with high probability, in terms of n.

  The group also considered function-Rips bifiltrations arising from a Poisson point process. They observed that these modules contain an interval indecomposable with infinite support, which is generated by the connected component associated with the point with the lowest function value.