Exemplars of Sheaf Theory in TDA

Justin M. Curry + the hard work of many others, cited in due time

May 18, 2022

University at Albany, SUNY

First, a Word from our Sponsors...





Many thanks to the NSF for CCF Award #1850052 and NASA GRC for TIMAEUS.

Plan for the Talk

This talk will focus on four case studies of (co)sheaves in TDA.

Two examples in the small:

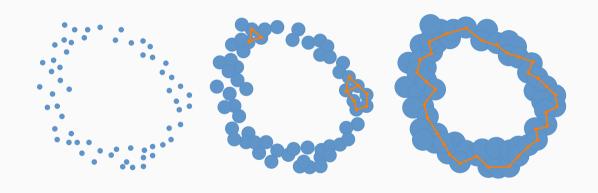
- 1. Persistent Homology
- 2. **Decorated Merge Trees**

And two examples in the large:

- 3. The Moduli Space of Merge Trees
- 4. The Persistent Homology Transform Sheaf

Formalizing Persistent Homology

Point Cloud and Sub-level Set Filtrations



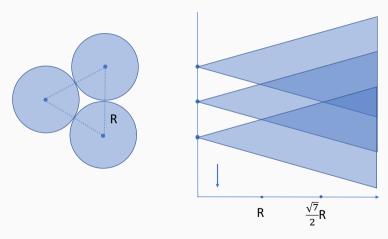
Traditional Persistence

Mantra for Traditional Persistence

Study inclusions from "lower" to "higher" parameters via functoriality.

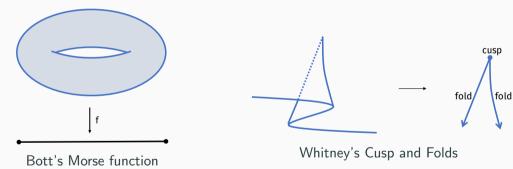
The "Algebraic Geometry" of TDA

TDA studies maps to metric spaces $f: X \to S$ by algebratizing $f^{-1}(s)$.



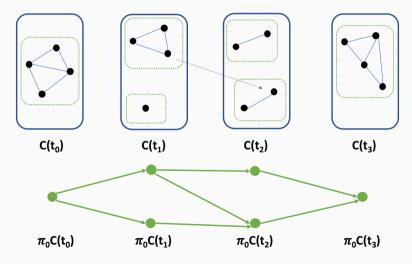
Challenges of Level Set Persistence

How to organize the homology of the fibers $f^{-1}(s)$ for a general map $f: X \to S$?



Challenges of Time-Evolving Persistence

What if we wanted to study a time-evolving point cloud or network?



The Answer

- Sheaves and Cosheaves provide a unifying language for:
 - Merge Trees \leftrightarrow Coshv($\mathbb{R}_{\mathsf{Alex}}$; Set) and Reeb graphs \leftrightarrow Coshv($\mathbb{R}_{\mathsf{Eucl}}$; Set)
 - Sublevel Set Persistence \leftrightarrow **Shv**(\mathbb{R}_{Alex})
 - Level Set Persistence \leftrightarrow **Shv**($\mathbb{R}_{\mathsf{Eucl}}$)
 - Time-varying Persistence \leftrightarrow **Shv**($\mathbb{R}_{\mathsf{Eucl}} \times \mathbb{R}_{\mathsf{Alex}}$)
- (Co)Sheaves serve as a calculus of TDA in the sense that
 - there is an existing toolbox of results and theory, where
 - one can write down back-of-the-envelope calculations, and
 - there are computable discrete models called cellular (co)sheaves.

Cellular Sheaves and Cosheaves

Definition (MacPherson, Shepard, et al)

Let X be a cell complex and \mathbf{D} a category, e.g. **Set** or **Vect**.

A **cellular sheaf** F assigns to every

- cell $\sigma \subseteq X$ an object $F(\sigma)$, and to every
- pair $\sigma \subseteq \overline{\tau}$, written $\sigma \le \tau$, a morphism $F(\sigma) \to F(\tau)$, such that

whenever $\sigma \leq \gamma \leq \tau$, the morphism $F(\sigma) \to F(\tau)$ is equal to the composition $F(\sigma) \to F(\gamma) \to F(\tau)$.

In other words, a cellular sheaf is just a functor $F : \mathbf{Cell}(X) \to \mathbf{D}$. We can turn arrows around to define a cellular cosheaf $\hat{F} : \mathbf{Cell}(X)^{\mathrm{op}} \to \mathbf{D}$.

9

"Continuous" Sheaves and Cosheaves

Pre-Cosheaf and Cosheaf

Let X be a topological space and \mathbf{D} a category, e.g. **Set** or **Vect**.

A pre-cosheaf F

- assigns to every open set $U \subseteq X$ an object F(U)
- ullet assigns to every pair $U\subseteq V$ a morphism F(U) o F(V)

If the object F(U) can be determined as a colimit of objects assigned to elements of any cover $\{U_i\}$ of U, then F is a **cosheaf**.

Pre-Sheaves and Sheaves

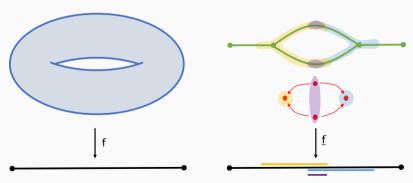
By turning around the arrows, so that F restricts from larger open sets to smaller ones, one obtains the notion of a pre-sheaf and sheaf.

Fundamental Example I: Reeb Cosheaf

Given a map $f: Y \to X$, the **Reeb cosheaf**, which is *the* fundamental cosheaf,

$$\mathcal{R}_f \colon U \rightsquigarrow \pi_0(f^{-1}(U))$$

tracks components and plays the same role as the sheaf of sections does in sheaf theory.

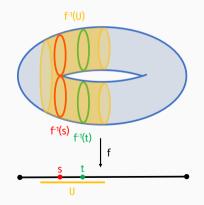


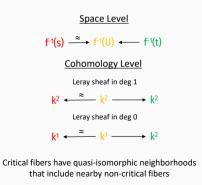
Fundamental Example II: Leray Sheaves

To a map $f: Y \to X$ the **Leray sheaf in degree** n is the sheafification of the pre-sheaf

$$F^n: U \rightsquigarrow H^n(f^{-1}(U); \mathbb{k}),$$

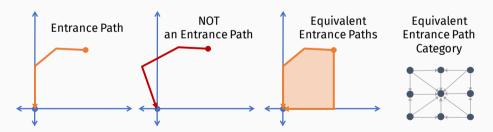
written as \mathcal{F}^n or $R^n f_* \mathbb{K}_Y$; the n^{th} right derived pushforward of the constant sheaf.





MacPherson's Entrance Path Category

Let X be stratified, i.e. partitioned into manifolds called strata. An **entrance path** is a path that only leaves a stratum by entering a lower dimensional one. Two entrance paths are **equivalent** if they are homotopic, rel endpoints, through entrance paths. The **entrance path category Ent**(X) has points of X for objects and equivalence classes of entrance paths for morphisms.

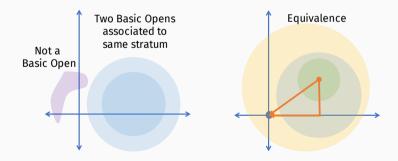


For cell complexes the entrance path category is equivalent to the face relation poset.

Classification of Constructible Cosheaves

Cosheaf $F : \mathbf{Open}(X) \to \mathbf{D}$ is **constructible** if whenever $U \subseteq V$ are basic opens associated to the same stratum, the morphism $F(U) \to F(V)$ is invertible.

Theorem (w/ Amit Patel) $\mathsf{Coshv}_c(X;\mathsf{D}) \simeq [\mathsf{Ent}(X);\mathsf{D}]$



Constructible Cosheaves are Finite Descriptors

Corollary

If $X = \mathbb{R}$ is stratified into finitely many pieces, the entrance path category is equivalent to

$$(-\infty, t_1) \to \{t_1\} \leftarrow \cdots \to \{t_n\} \leftarrow (t_n, \infty)$$

and thus the study of constructible cosheaves is equivalent to the study of zig-zag diagrams:

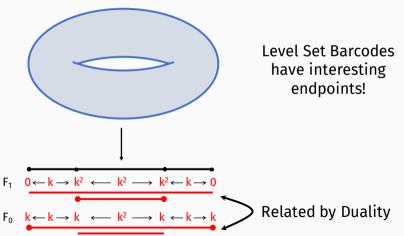
$$F(-\infty, t_1) \to F(t_1) \leftarrow \cdots \to F(t_n) \leftarrow F(t_n, \infty)$$

For D = Set, this says constructible cosheaves are equivalent to Reeb graphs.

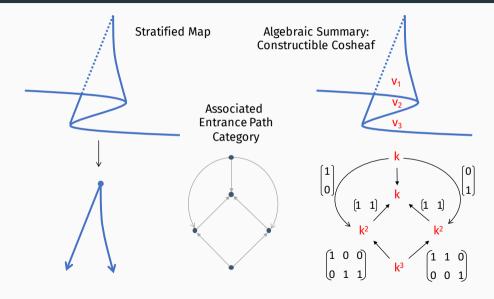
For $\mathbf{D} = \mathbf{Vect}$, this says that constructible cosheaves are equivalent to representations of an A_{2n+1} -type quiver, which has a **barcode decomposition**.

Bott's Torus Revisited

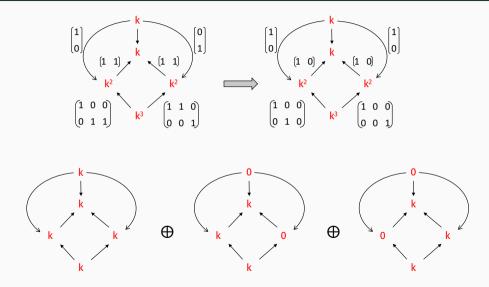
Leray cosheaves are the canonical zig-zag modules, adapted to the stratification by critical values.



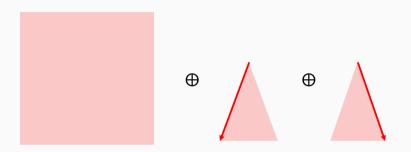
Whitney's Cusp Revisited



"Unbreakable Summands" = Indecomposables



Higher-Dimensional Barcodes?



PROBLEM: Not every constructible cosheaf decomposes "nicely".

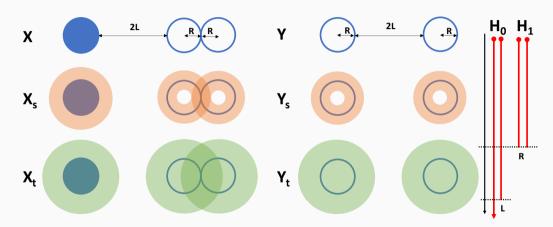
SOLUTION: Abandon All Indecomposables, Ye Who Enter Here

One should look at "stable" invariants or generalized rank invariants. Consider their Möbius Inversions à la MacPherson and Patel; Gulen and McCleary.

Decorated Merge Trees

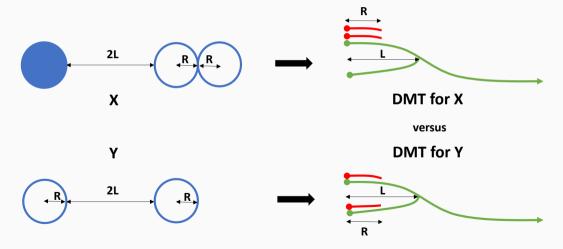
The Challenge

We want to distinguish the offset filtrations of X and Y using a minimal data structure.



The Cartoon Solution

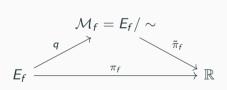
A decorated merge tree should distinguishes these.

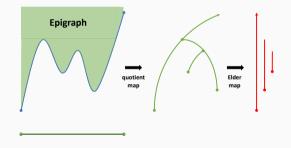


The Sheaf-Theoretic Solution

Sublevel set persistent homology is just the study of the Leray sheaves of the map π_f .

(Concrete) Decorated Merge Trees are defined by the Leray sheaves of the map q.





Point Cloud Example

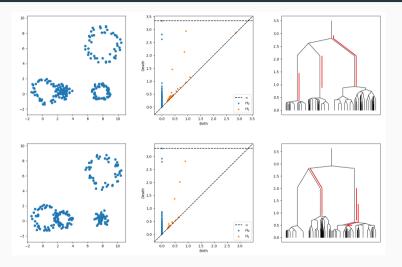


Figure Credit: Tom Needham (FSU)

Cycle Localization

Our DMT algorithm provides cycle localization for free!

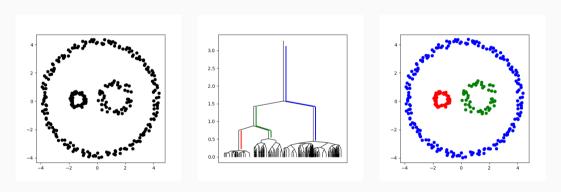


Figure Credit: Tom Needham (FSU)

Cycle Localization in Image Data

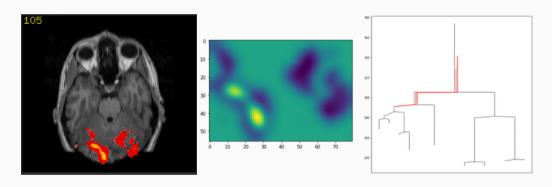


Figure Credit: Tom Needham (FSU)

Discrimination for Takens Embedding

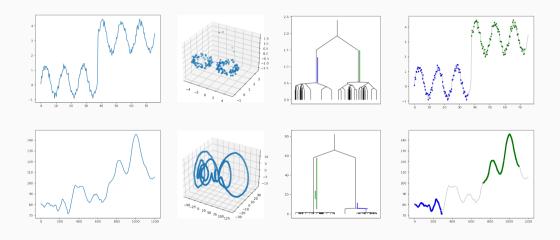
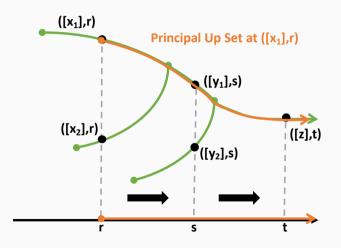


Figure Credit: Tom Needham (FSU)

Restricting the Concrete DMT

In practice, we don't work with $\mathcal{F}:\mathcal{M}_f \to \textbf{Vect}$ directly.



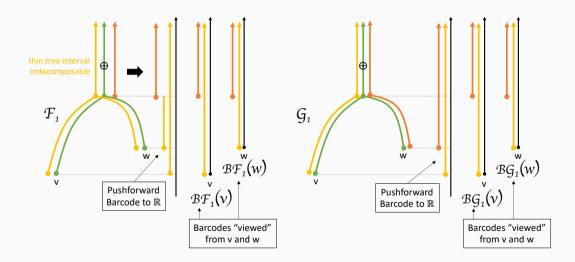
The Barcode DMT

Barcode Decorated Merge Tree

Associated to $\mathcal{F}_n: \mathcal{M}_F \to \mathbf{Vect}$, the **barcode DMT** assigns to each $p = ([x], r) \in \mathcal{M}_F$ the barcode of the restricted Leray (co)sheaf, i.e. $\mathcal{F}_n|_{U_p}$. This defines

$$\mathcal{BF}_n: \mathcal{M}_F \to \textbf{Barcodes}.$$

The Barcode Transform is Not Injective



Three Approaches to DMTs

Two perspectives on the Leray (co)sheaf:

$$\mbox{(Concrete DMT)} \quad \mathcal{F}_n: \mathcal{M}_f \to \mbox{Vect} \quad \Leftrightarrow \quad \widetilde{F}_n: \mathbb{R} \to \mbox{pVect} \quad \mbox{(Categorical DMT)}$$

In practice we use the

(Barcode DMT)
$$\mathcal{BF}_n : \mathcal{M}_f \to \mathsf{Barcodes},$$

https://github.com/trneedham/Decorated-Merge-Trees

Some Comments

- There is an obvious generalization to Reeb graphs and decorated Mapper graphs.
- Also easy to prove convergence results between these.
- Tractable invariants from DRGs is a challenge.

Upshot: Sheaves on graphs are interesting objects of study!

Plug: Graph Neural Nets as Cellular Sheaves



Figure Credit: Bodnar, Bronstein, di Giovanni, et al

Sheaf Theory in The Large

Bird's Eye Perspective

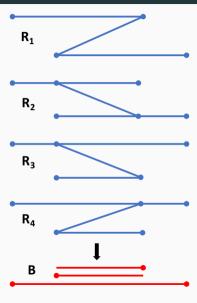
We have considered the TDA pipeline by studying individual inputs and outputs.

Most of these steps can be summarized using sheaf theory.

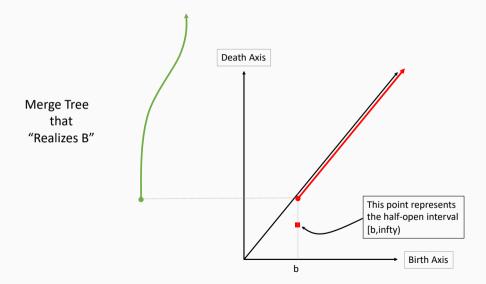


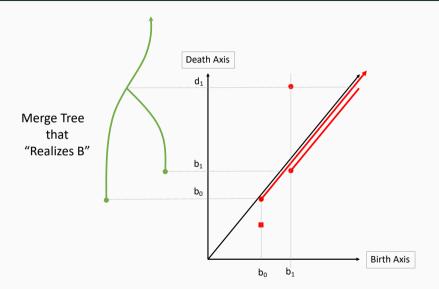
We now recurse and study the TDA pipeline as a map, algebratizing it accordingly.

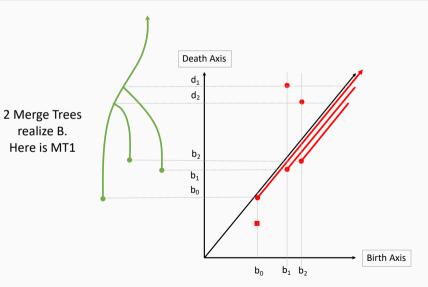
Open Problem! (cf. Jordan deSha's thesis)

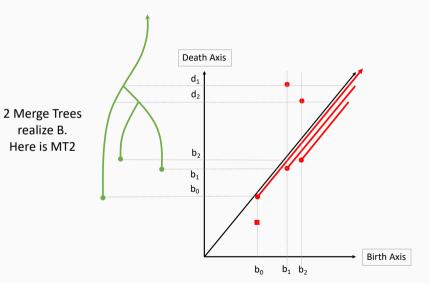


From Trees to Barcodes

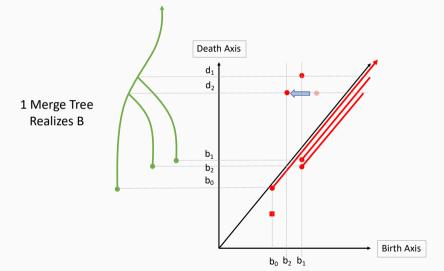




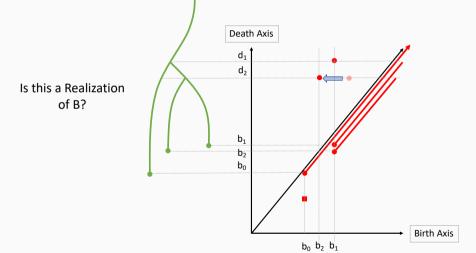




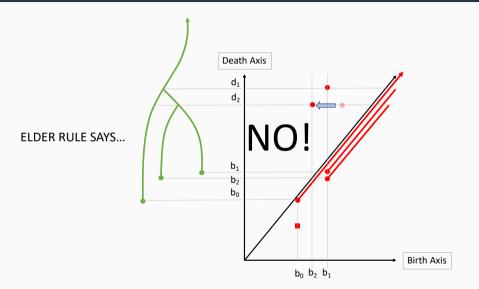
Change of Birth Order



Possible Realization?



Elder Rule



Enumerating the Fiber of the Elder Map

Theorem (C. '17; Garin, Hess, Kanari '20)

Fix a barcode B where every endpoint is distinct:

$$B = \{[b_0, \infty); [b_1, d_1); \ldots; [b_n, d_n)\}$$

Further, we assume $\mathit{I}_0 = [\mathit{b}_0, \infty)$ and $\mathit{I}_j := [\mathit{b}_j, \mathit{d}_j)$ satisfy

- (Containment) $I_j \subset I_0$ for all $j \geq 1$ and
- (Increasing Birth Times) $b_1 < b_2 < \cdots < b_n$.

Set $\mu_B(I_j) = \#\{k < j \mid d_j < d_k\}$. The number of merge trees realizing B is

$$TRN(B) := R(B) := \prod_{j=1}^{n} \mu_B(I_j)$$

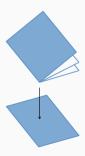
Significance: Stratified Covering Spaces

Containment Poset

The barcode $B = \{I_j\}$ forms a poset, ordered by containment of intervals.

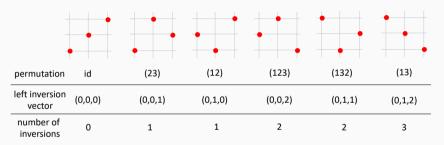
The space of persistence diagrams/barcodes is *stratified* by the containment poset.

The Elder Map defines a *stratified covering space* or a **Set**-valued constructible cosheaf over barcode space. The number of sheets over each top-dimensional stratum is R(B).



Inversion Vectors

Adélie Garin, Kathryn Hess, and Lida Kanari observed that *generic* persistence diagrams can be viewed as permutations, i.e. elements of the symmetric group.



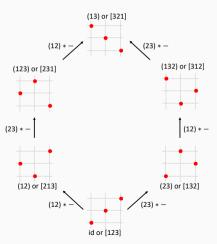
Observation by GHK + C. + Brendan Mallery and Jordan DeSha

Let σ be the permutation type of a generic B and let $\ell(\sigma)$ denote the left-inversion vector of σ , then $R(B) = \prod_{i=1}^{n} (\ell_i(B) + 1)$

Tree Realization Number is Bruhat Order Preserving

Lemma (CDGHKM '21)

If $\sigma, \sigma' \in S_n$ are such that $\sigma < \sigma'$ in the (left) Bruhat order, then $R(\sigma) < R(\sigma')$.



TRN on the Cayley Graph

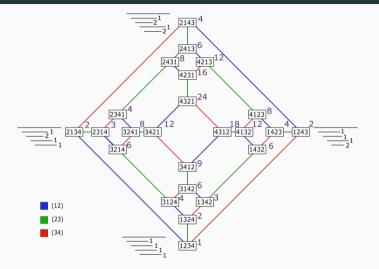


Figure Credit: Adélie Garin (EPFL)

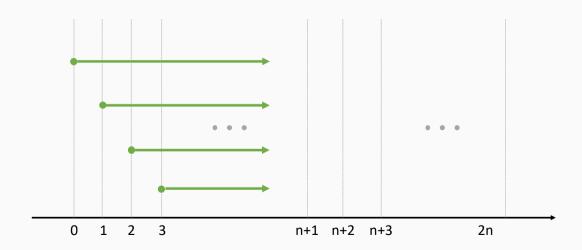
Counting Combinatorial Merge Trees

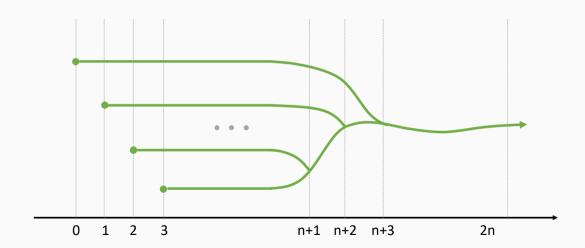
- 1. Top 2*n*-dimensional strata of barcode space $\mapsto S_n$, so there are n! many.
- 2. Merge trees form a (stratified) covering space of barcode space, so the sum

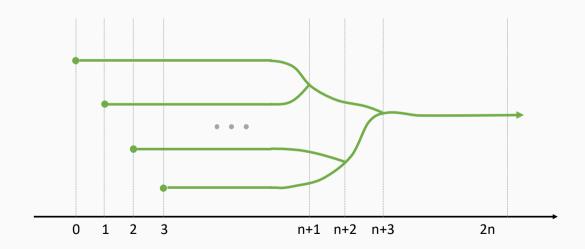
$$\sum_{\sigma \in S_n} R(B_\sigma)$$

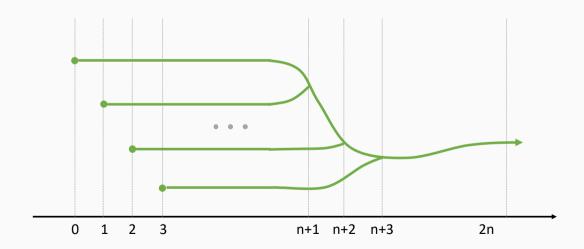
counts top 2n-dimensional strata.

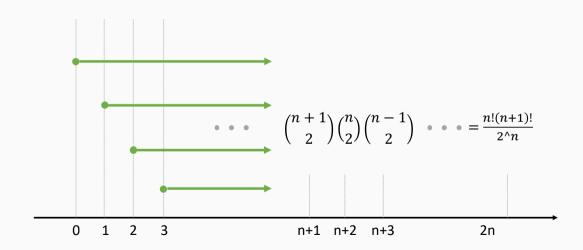
3. Both the strata of barcode space and merge tree space are **convex.**











Maximal Chains in the Lattice of Partitions

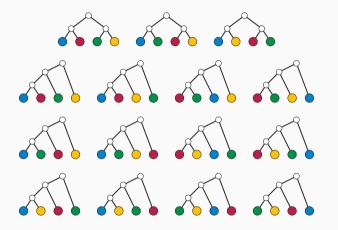
Theorem (CDGHKM '21)

The Elder Map stratifies MT Space so that top-dimensional strata are in bijection with maximal chains in the lattice of partitions. In summary,

$$\sum_{\sigma \in S_n} R(B_{\sigma}) = \sum_{\sigma \in S_n} \prod_{i=1}^n (I_i(\sigma) + 1) = \frac{(n+1)! \, n!}{2^n}.$$

cf. BHV space, where orthants are counted by (2n-1)!!, where n=#leaves -1

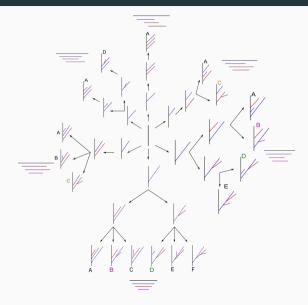
A Lattice-Theoretic Perspective on the Persistence Map by Mallery, Garin, C.



For 4 leaves, there are 15 BHV topologies.

Figure Credit: Wikipedia

18 MT Trees (Figure Credit: Adélie Garin)



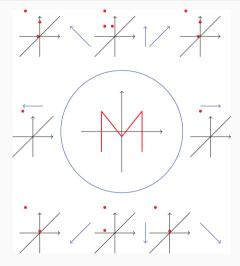
Remarks on MT Space versus BHV Space

Merge Tree space is very different from BHV space.

- Points in MT Space correspond to *isomorphism* classes of merge trees.
- Generically we can label leaf nodes by birth time, but this is not continuous.
- Moreover, a given orthant of BHV space (split topology type) does not have a
 uniquely associated permutation type of barcode, but we can bound this
 difference precisely. See Prop 3.23 of https://arxiv.org/abs/2107.11212
- Understanding the stratified space structure is critical for doing good statistics.

The Big PHT Sheaf

Persistent Homology Transform



Copyright held by Katharine Turner, Sayan Mukherjee, Doug Boyer

PHT for Shape Discrimination

The PHT is sufficient for describing shapes.

Theorem (TMB '14)

If $K, K' \subseteq \mathbb{R}^3$ are two PL-embedded simplicial complexes and $\mathbf{PHT}(K) = \mathbf{PHT}(K')$, then K = K'.

This was generalized by using Schapira's Inversion Theorem for the Radon transform.

Theorem (CMT '18/21; Ghrist, Levanger, Mai '18)

If $M, M' \subseteq \mathbb{R}^d$ are two constructible subsets and $\mathbf{PHT}(M) = \mathbf{PHT}(M')$, then M = M'.

Sheaf Interpretation

Definition

Given $M \subseteq \mathbb{R}^d$ we have $Z_M := \{(x, v, t) \in M \times S^{d-1} \times \mathbb{R} \mid x \cdot v \leq t\}.$

The **derived PHT** of M is the Leray sheaf of the map $\pi_M: Z_M \to S^{d-1} \times \mathbb{R}$.

By restricting $\mathbf{PHT}(M) := R\pi_{M*} \Bbbk_{Z_M}$ to $\{v\} \times \mathbb{R}$ we obtain a sheaf on a totally ordered subset and hence barcodes in each degree.

A Pre-Sheaf of PHT Sheaves

Notice that if we include a subset $A \hookrightarrow M$ then we have an associated inclusion

$$Z_A \hookrightarrow Z_M$$

Restricting cohomology of sublevel sets in M to those in A induces a sheaf morphism

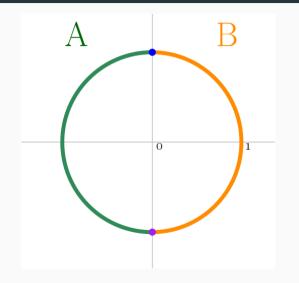
$$PHT(M) \Rightarrow PHT(A)$$

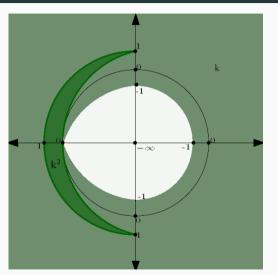
Lemma (Arya, C., Mukherjee '21)

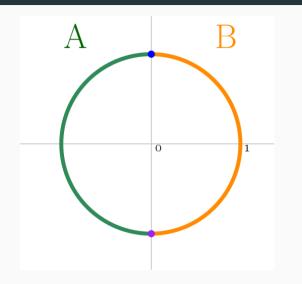
The following assignment is a pre-sheaf

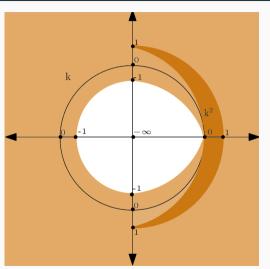
$$\mathcal{F}:\mathcal{CS}(\mathbb{R}^d)^{\mathrm{op}} o\mathcal{D}^b(\mathsf{Shv}(S^{d-1} imes\mathbb{R}))\qquad M\mapsto\mathsf{PHT}(M)$$

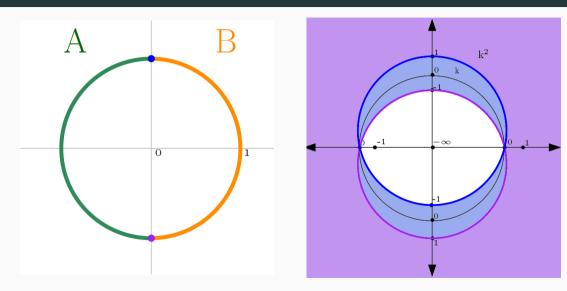
where PHT(M) is the derived sheaf version of the PHT.

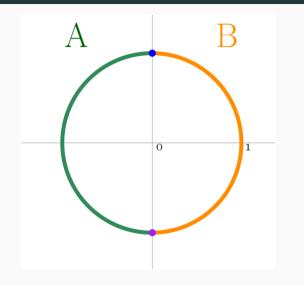


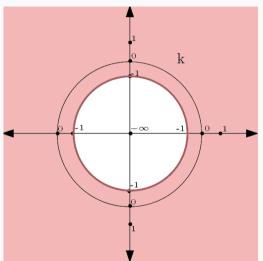












PHT for PL shapes

- For a convex subset $A \subseteq \mathbb{R}^d$, all of the PHT is concentrated in degree 0.
- If we view a polyhedron $M \subseteq \mathbb{R}^d$ as glued together convex shapes, then we can recover $\mathbf{PHT}(M)$ completely in terms of $\mathbf{PHT}^0(M_i)$, where $\{M_i\}$ is a locally finite convex cover of M.

Theorem (ACM '22)

For the simplicial complex $M \in \mathbb{R}^d$ and cover $\mathcal{V} = \{M_i\}_{i \in I}$ of M, $\mathbf{PHT}^n(M)$ is the n-th cohomology of the following complex of *sheaves*:

$$0 \to \bigoplus_{i \in I} \mathsf{PHT}^0(M_i) \to \bigoplus_{i < j} \mathsf{PHT}^0(M_i \cap M_j) \to \cdots$$

where the \cdots represents the higher intersection terms.

Cech Descent for the PHT

Definition

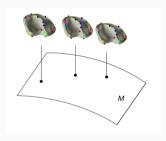
A pre-sheaf $\mathcal{F}: \mathcal{C}^{op} \to \mathcal{D}^b(\mathbf{A})$ is a **homotopy sheaf** (satisfies **Čech descent**) if for every object $U \in \mathcal{C}$ and cover $\mathcal{U} = \{U_i \to U\}$ the following map is a quasi-iso:

$$\mathcal{F}(U)\stackrel{\simeq}{ o} \mathsf{holim}\left[\prod_i \mathcal{F}(U_i)
ightrightarrowtail \prod_{i,j} \mathcal{F}(U_{ij}) \cdots
ight]$$

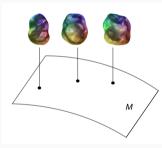
Theorem (ACM '22)

PHT is a homotopy sheaf on the o-minimal site of constructible sets $\mathcal{CS}(\mathbb{R}^d)$

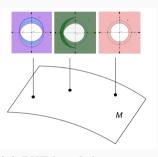
A Sheaf-Theoretic Construction of Shape Space



(A) Kendall's shape space.



(B) Grenander's shape space.



(c) PHT-based shape space.

- Sheaves, Cosheaves and TDA:
 - 1. "Sheaves, Cosheaves and Applications" by C. 1303.3255
 - 2. "Discrete Morse theory for computing cellular sheaf cohomology" by C., Ghrist, Nanda 1312.6454
 - 3. "Topological Data Analysis and Cosheaves" by C. 1411.0613
 - 4. "Categorified Reeb Graphs" by de Silva, Munch, Patel 1501.04147
 - 5. "Classification of Constructible Cosheaves" by C., Patel 1603.01587
 - 6. "Persistent Homology and Microlocal Sheaf Theory" by **Kashiwara, Schapira** 1705.00955
 - 7. "Persistent Local Systems" by MacPherson, Patel 1805.02539
 - 8. "A derived isometry theorem for constructible sheaves on R" by **Berkouk, Ginot** 1805.09694
 - 9. "Canonical stratifications along bisheaves" by Nanda, Patel 1812.05593
 - 10. "Zig-Zag Modules: Cosheaves and K-Theory" by Grady, Schenfish 2110.04591
 - 11. "Neural Sheaf Diffusion" by **Bodnar, et al** 2202.04579

- Enriched Topological Summaries
 - 1. "The Fiber of the Persistence Map for Functions on the Interval" by C. 1706.06059
 - 2. "Time Series, Persistent Homology and Chirality" by Baryshnikov 1909.09846
 - 3. "Moduli Spaces of Morse Functions for Persistence" by Catanzaro, C., Fasy, Lazovskis, Malen, Riess, B. Wang, Zabka 1909. 10623
 - 4. "Decorated Merge Trees for Persistent Topology" by C., Hang, Mio, Needham, Okutan 2103.15804
 - 5. "Counting Embedded 2-Spheres with the Same Persistence" by **C.,deSha** in.prep

- Studying the Persistence Map
 - "Continuation of Point Clouds via Persistence Diagrams" by Gameiro, Hiraoka, Obayashi 1506.03147
 - "A Framework for Differential Calculus on Persistence Barcodes" Leygonie, Oudot, Tillman 1910.00960
 - 3. "From trees to barcodes and back again: theoretical and statistical perspectives" by **Kanari, Garin, Hess** 2010.11620
 - 4. "The Fiber of Persistent Homology for simplicial complexes" by **Leygonie, Tillman** 2104.01372
 - 5. "Inverse Problems for Topological Summaries in TDA" by deSha (PhD Thesis)
 - 6. "From Trees to Barcodes and Back Again II" by CDGHKM 2107.11212
 - 7. "Fiber of Persistent Homology on Morse functions" Leygonie, Beers 2108.07512
 - 8. "A Gradient Sampling Algorithm for Stratified Maps with Applications to TDA" by Leygonie, Carrière, Lacombe, Oudot 2109.00530
 - "A Lattice-Theoretic Perspective on the Persistence Map" by Mallery, Garin, C. 2203, 00643

- The Persistent Homology Transform
 - "Persistent Homology Transform for Modeling Shapes and Surfaces" by Turner, Mukherjee, Boyeri 1310.1030
 - 2. "Predicting Clinical Outcomes in Glioblastoma" **Crawford, Monod, Chen, Mukherjee, Rabadán** https://arxiv.org/abs/1611.06818
 - "Persistent Homology and Euler Integral Transforms" by Ghrist, Levanger, Mai 1804.04740
 - 4. "How Many Directions Determine a Shape?" by **C., Mukherjee, Turner** 1805.09782
 - 5. "A Faithful Discretization of the PHT" by **Fasy, Micka, Millman, Schenfisch, Williams** 1912.12759
 - "A TDA Approach for Discovering Biophysical Signatures in Protein Dynamics" https://doi.org/10.1101/2021.07.28.454240
 - 7. "A Sheaf Theoretic Construction of Shape Space" by Arya, C., Mukherjee https://arxiv.org/abs/2204.09020

Thank You!