Persistent 1-Cycles: Definition, Computation, and Some Applications

Tao Hou

Department of Computer Science and Engineering
The Ohio State University

In collaboration with Tamal K. Dey and Sayan Mandal

Persistence diagram

Birth-death and persistence diagrams (barcodes)

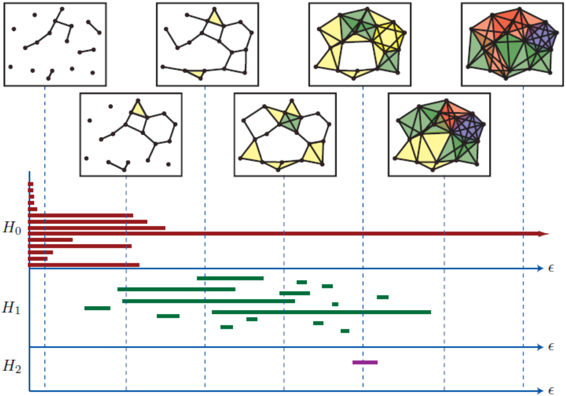


Figure: Courtesy of [Ghrist, 2008]

Persistence module

- Simplicial filtration ($\mathcal{F}$):

$$\varnothing = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n = K$$

- Persistence module ($\mathcal{P}_d^{\mathcal{F}}$):

$$\begin{tikzcd}
0 \arrow[r] & \mathrm{H}_d(K_1) \arrow[r] & \mathrm{H}_d(K_2) \arrow[r] & \cdots \arrow[r] & \mathrm{H}_d(K_n) \\
0 \arrow[r] \arrow[u] & 1 \arrow[r] \arrow[u] & 2 \arrow[r] \arrow[u] & \cdots \arrow[r] \arrow[u] & n \arrow[u]
\end{tikzcd}$$

$$\mathcal{P}_d^{\mathcal{F}} : \mathbb{Z}^+ \to \textbf{Vec}$$

Interval module decomposition

- Interval module ($\mathcal{I}^{[\beta,\delta)}$):

$$\mathcal{I}^{[\beta,\delta)}(i) = \begin{cases} \mathbb{Z}_2 & i \in [\beta,\delta) \\ 0 & otherwise \end{cases}$$

- Interval module decomposition

$$\mathcal{P}_d^{\mathcal{F}} = \bigoplus_{j \in J} \mathcal{I}^{[\beta_j,\delta_j)}$$

- Persistence diagram

$$\mathrm{Dgm}(\mathcal{P}_d^{\mathcal{F}}) = \{[\beta_j,\delta_j) \mid j \in J\}$$

Definition (Persistent d-basis)

An indexed set of d-cycles $\{c_j \mid j \in J\}$ for $\mathcal{F}$ *s.t.*

- $\mathcal{P}_d^{\mathcal{F}} = \bigoplus_{j \in J} \mathcal{I}^{[\beta_j, \delta_j)}$
- For each $j \in J$ and $\beta_j \leq k < \delta_j$, $\mathcal{I}^{[\beta_j, \delta_j)}(k) = \{0, [c_j]\}$.

Definition (Persistent d-cycle)

For $[\beta, \delta) \in \mathrm{Dgm}(\mathcal{P}_d^{\mathcal{F}})$, it is a d-cycle c for $[\beta, \delta)$ *s.t.*

- $\delta = +\infty$ (infinite interval): c is a cycle in K_β containing $\sigma_\beta^{\mathcal{F}}$
- $\delta \neq +\infty$ (finite interval): c is a cycle in K_β containing $\sigma_\beta^{\mathcal{F}}$ & c is not a boundary in $K_{\delta-1}$ but becomes boundary in K_δ

Theorem

An indexed set of d-cycles $\{c_j \mid j \in J\}$ is a persistent d-basis of $\mathcal{F}$ $\iff$ c_j is a persistent d-cycle for every $[\beta_j, \delta_j) \in \mathrm{Dgm}(\mathcal{P}_d^{\mathcal{F}})$

Definition (Minimal persistent d-cycle)

Persistent d-cycle for the interval with the minimal weight

Definition (Minimal persistent d-basis)

- A persistent d-basis $\{c_j \mid j \in J\}$
- $\forall j \in J$, c_j is a minimal persistent d-cycle for $[\beta_j, \delta_j)$

Problem (PCYC-FIN$_d$)

Given: Simplicial complex K, filtration $\mathcal{F}$, finite interval $[\beta, \delta) \in \mathrm{Dgm}_d(\mathcal{F})$
Compute: A minimal persistent d-cycle for the interval

Problem (PCYC-INF$_d$)

Similar to PCYC-FIN$_d$, only the interval $[\beta, +\infty)$ becomes an infinite one

Theorem

PCYC-FIN$_1$ is NP-hard (reduce from MAX-2SAT)

Remark

PCYC-INF$_1$ can be computed in polynomial time by finding the shortest path

Algorithm 1: A framework for persistent 1-basis

Maintains a basis $\mathcal{B}_i$ for $\mathsf{H}_1(K_i)$ $\forall\, i = 0, \ldots, n$ ($\mathcal{B}_0 \leftarrow \varnothing$):

- for $i = 1, \ldots, n$:
  - $\sigma_i^{\mathcal{F}}$ is a positive 1-simplex:
    - find a 1-cycle c_i in K_i containing $\sigma_i^{\mathcal{F}}$
    - $\mathcal{B}_i \leftarrow \mathcal{B}_{i-1} \cup c_i$
  - $\sigma_i^{\mathcal{F}}$ is a negative 2-simplex:
    - find a $G \subseteq \mathbb{Z}$ *s.t.* $\forall\, g \in G, c_g \in \mathcal{B}_{i-1}$ and $\sum_{g \in G} [c_g] = 0$ in K_i
    - assign $\sum_{g \in G} c_g$ to $[g^*, i)$ as a persistent cycle ($g^* = \max G$)
    - $\mathcal{B}_i \leftarrow \mathcal{B}_{i-1} \setminus c_{g^*}$
  - otherwise: $\mathcal{B}_i \leftarrow \mathcal{B}_{i-1}$
- at the end, for each cycle $c_j \in \mathcal{B}_n$, assign c_j to $[j, +\infty)$

Finding G can be done in $O(n^\omega)$ [Dey et al., 2014]

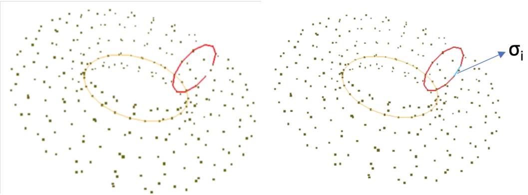
Algorithm 2: Meaningful persistent 1-cycles

Modifying Algorithm 1:

- $\sigma_i^{\mathcal{F}}$ positive: let c_i be shortest cycle containing $\sigma_i^{\mathcal{F}}$

Constructing c_i:

- add $\sigma_i^{\mathcal{F}}$ to the shortest path between vertices of $\sigma_i^{\mathcal{F}}$ in K_{i-1}
- Dijkstra's algorithm on 1-skeleton of K_{i-1}



Algorithm 2: Meaningful persistent 1-cycles

Algorithm 2 produces meaningful persistent 1-cycles:

- Shortest cycle: Good representation $\Longrightarrow$
  Sum of shortest cycles: Good representative for interval

Proposition

In Algorithm 2, when $\sigma_i^{\mathcal{F}}$ is a negative 2-simplex, if $|G| = 1$, then $\sum_{g \in G} c_g$ is a minimal persistent 1-cycle

Above scenarios are quite common in practice

Problem of Algorithm 2:

- Cycles of all intervals (including noises) are computed
- Often, user only cares about some large intervals

We introduce Algorithm 3 ...

Problem of Algorithm 2:

- Cycles of all intervals (including noises) are computed
- Often, user only cares about some large intervals

We introduce Algorithm 3 ...

Proposition

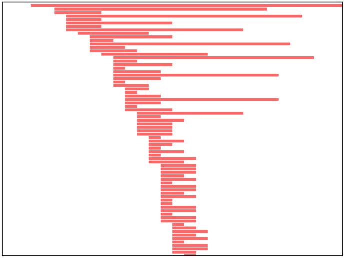
In Algorithm 1 and 2, when $\sigma_i^{\mathcal{F}}$ is negative, for any $g \in G$, its corresponding interval satisfies: $\beta_g \leq g^*$ and $\delta_g \geq i$.

Recall that $\{c_g \mid g \in G\} \subseteq \mathcal{B}_{i-1}$ and $\sum_{g \in G}[c_g] = 0$

Algorithm 3: Optimization for a single interval

- Compute persistent cycle for an interval $[\beta, \delta)$: Compute shortest cycles at birth indices whose corresponding intervals contain $[\beta, \delta)$

- Long input interval:
intervals containing it are a small subset



- Much faster than Algorithm 2:
Less shortest path computations
in practice

- Worst case of both:
$O(n^\omega + n^2 \log n) = O(n^\omega)$

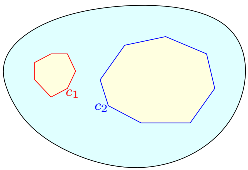
Algorithm 3: Optimization for a single interval

Input: The input of Algorithm 2 plus an interval $[\beta, \delta)$
Output: A persistent 1-cycle for $[\beta, \delta)$ output by Algorithm 2

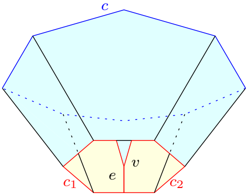
$G' \leftarrow \varnothing$
for $i \leftarrow 1, \dots, \beta$ **do**
    if $\sigma_i^{\mathcal{F}}$ is positive **and** ($\sigma_i^{\mathcal{F}}$ is paired with a $\sigma_j^{\mathcal{F}}$ $s.t\ j \geq \delta$ **or** $\sigma_i^{\mathcal{F}}$ never gets paired) **then**
        $c_i \leftarrow$ the shortest cycle containing $\sigma_i^{\mathcal{F}}$ in K_i
        $G' \leftarrow G' \cup \{i\}$
find a $G \subseteq G'$ s.t. $\sum_{g \in G} [c_g] = 0$ in K_δ
output $\sum_{g \in G} c_g$ as the persistent 1-cycle for $[\beta, \delta)$

Stability of persistent 1-cycles

Instablili by: Minimal cycles & cycles by our algorithm



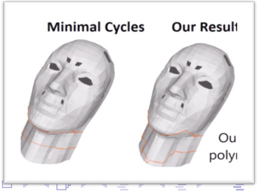
(a) Minimal cycles



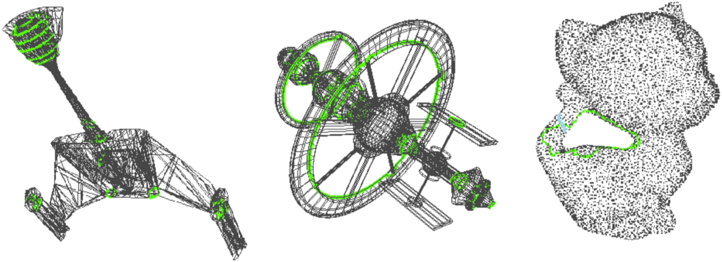
(b) Our cycles

Our algorithm:

- Works well in practice
- Has nice properties

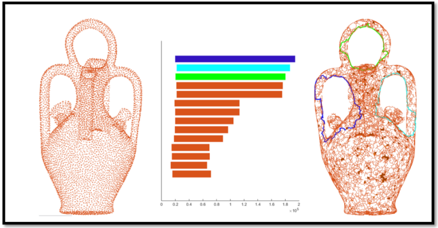


Persistent 1-cycles for 3D point clouds



Persistent 1-cycles for 3D point clouds

Snapshot from our software



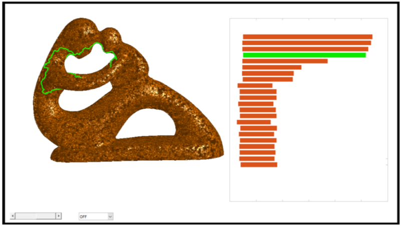
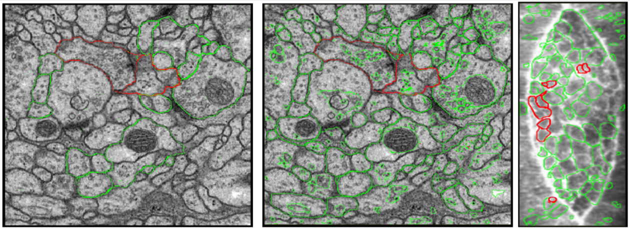


Image segmentation and characterization

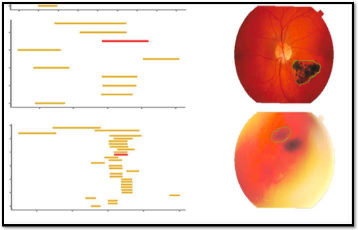
Sections from Drosophila larva tissues and embryos

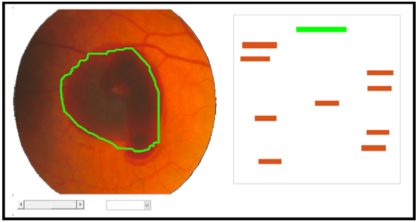


Left: Top 20 intervals, Center: Top 350 intervals, Right: Top 200 intervals

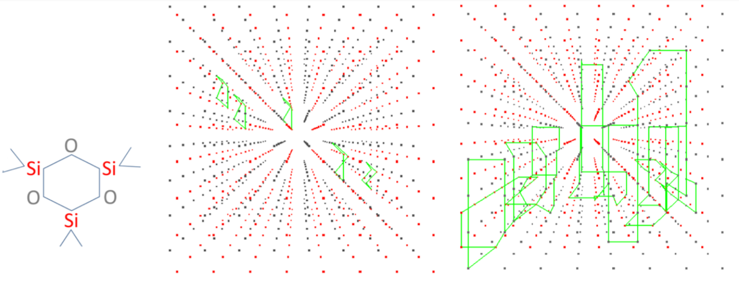
Image segmentation and characterization

STARE dataset for hemorrhage detection





Hexagonal structure of crystalline solids



Left: Hexagonal cyclic structure of silicate glass,
Center: 1-cycles, Red dots: Silicon, Grey dots: Oxygen.
Right: Persistent 1-cycles computed for the long range order

References

Dey, T. K., Fan, F., and Wang, Y. (2014).
Computing topological persistence for simplicial maps.
In *Proceedings of the thirtieth annual symposium on Computational geometry*, page 345. ACM.

Ghrist, R. (2008).
Barcodes: the persistent topology of data.
Bulletin of the American Mathematical Society, 45(1):61–75.

Thank You