

Towards a Persistence Diagram that is Robust to Noise and Varied Densities

ICML 2023

Hang Zhang*, Kaifeng Zhang*, Kai Ming Ting, Ye Zhu

2023.6.29

南

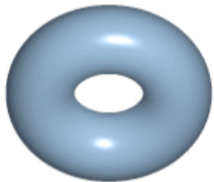
京

大

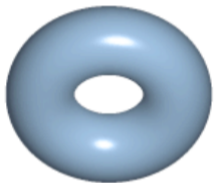
学

- 1 Topology and Persistent Homology
- 2 How to Construct a Correct PD
- 3 Λ -filter
- 4 Experiments
- 5 Conclusions

Topology



Topology



A-DNA



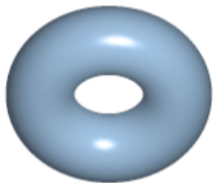
B-DNA



Z-DNA



Topology



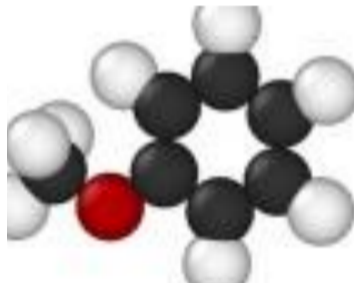
A-DNA



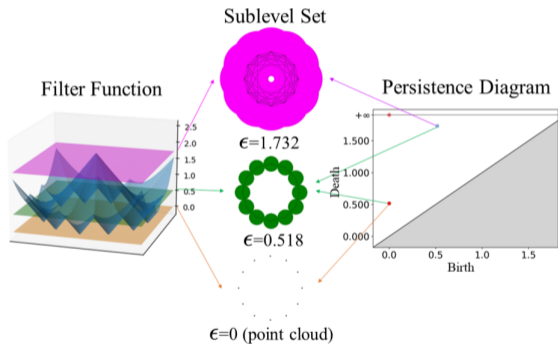
B-DNA



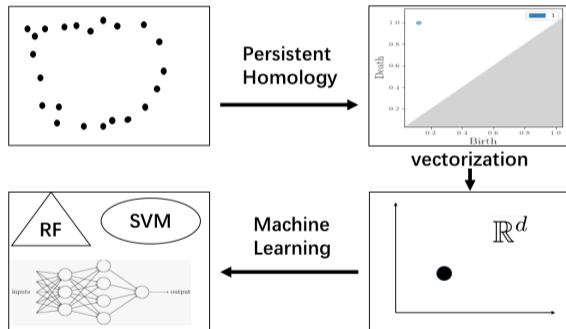
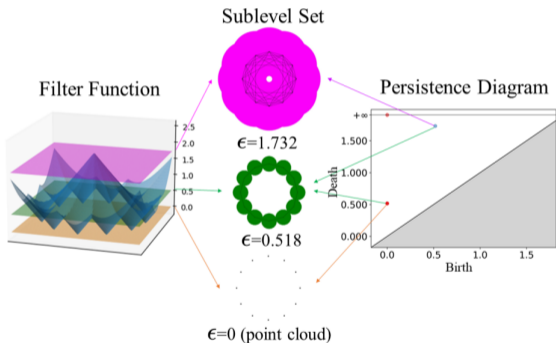
Z-DNA



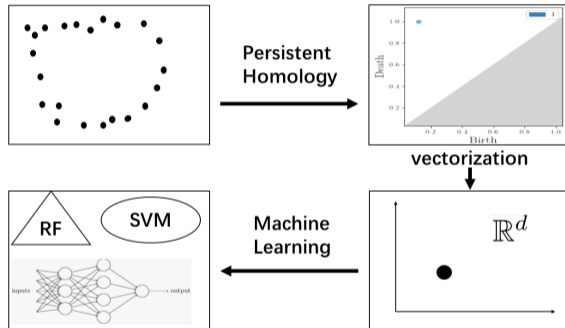
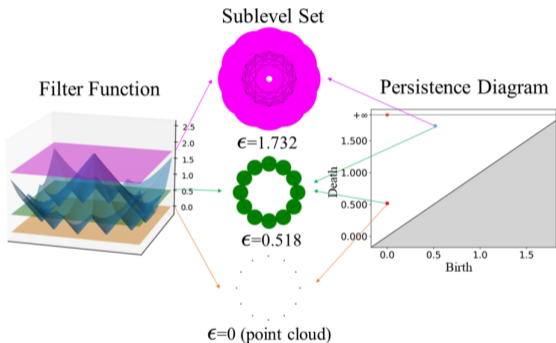
Persistent Homology



Persistent Homology



Persistent Homology



Do existing methods produce a correct and good PD?

- ① Topology and Persistent Homology
- ② How to Construct a Correct PD
- ③ Λ -filter
- ④ Experiments
- ⑤ Conclusions

How to Construct a Correct PD

- A correct PD must be robust to noise

How to Construct a Correct PD

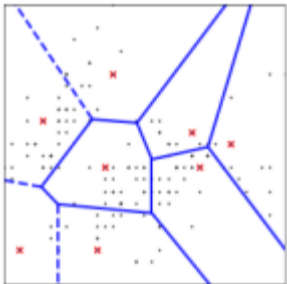
- A correct PD must be robust to noise
- A correct PD must be robust to varied densities

How to Construct a Correct PD

- A correct PD must be robust to noise
- A correct PD must be robust to varied densities

Approach	Filter Function $f(x)$	Robust to	
		noise	varied densities
Rips	$2 \min_{y \in X} \ell(x, y)$	×	×
DTM	$\sqrt{\frac{1}{k} \sum_{i=1}^k \ell_i^2(x)}$	✓	×
CkNN ³	$\min_{y \in X} \frac{\ell(x, y)}{2\ell_k(y)}$	×	✓
Λ -filter	$4 \min_{y \in X} (1 - \kappa(x, y))$	✓	✓

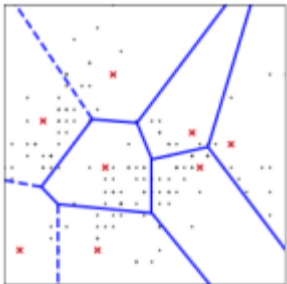
- ① Topology and Persistent Homology
- ② How to Construct a Correct PD
- ③ Λ -filter
- ④ Experiments
- ⑤ Conclusions



$$\Phi(x | \Lambda) = \left[\frac{e^{-\eta \ell(x, v_1)}}{\Upsilon_\ell}, \dots, \frac{e^{-\eta \ell(x, v_\psi)}}{\Upsilon_\ell} \right]^\top$$

$$s(x, y | \Lambda) = \langle \Phi(x | \Lambda), \Phi(y | \Lambda) \rangle$$

$$\hat{\kappa}(x, y | X) = \frac{1}{t} \sum_{i=1}^t s(x, y | \Lambda_i)$$



$$\Phi(x | \Lambda) = \left[\frac{e^{-\eta \ell(x, v_1)}}{\Upsilon_\ell}, \dots, \frac{e^{-\eta \ell(x, v_\psi)}}{\Upsilon_\ell} \right]^\top$$

$$s(x, y | \Lambda) = \langle \Phi(x | \Lambda), \Phi(y | \Lambda) \rangle$$

$$\hat{\kappa}(x, y | X) = \frac{1}{t} \sum_{i=1}^t s(x, y | \Lambda_i)$$

$$\hat{f}_\Lambda(x) = 4 \min_{y \in X} (1 - \hat{\kappa}(x, y | X))$$

- ① Topology and Persistent Homology
- ② How to Construct a Correct PD
- ③ Λ -filter
- ④ Experiments
- ⑤ Conclusions

Noise and Varied Densities Coexist



Noise and Varied Densities Coexist

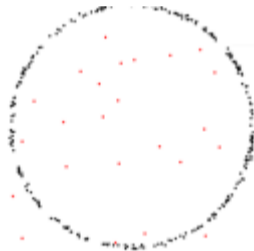
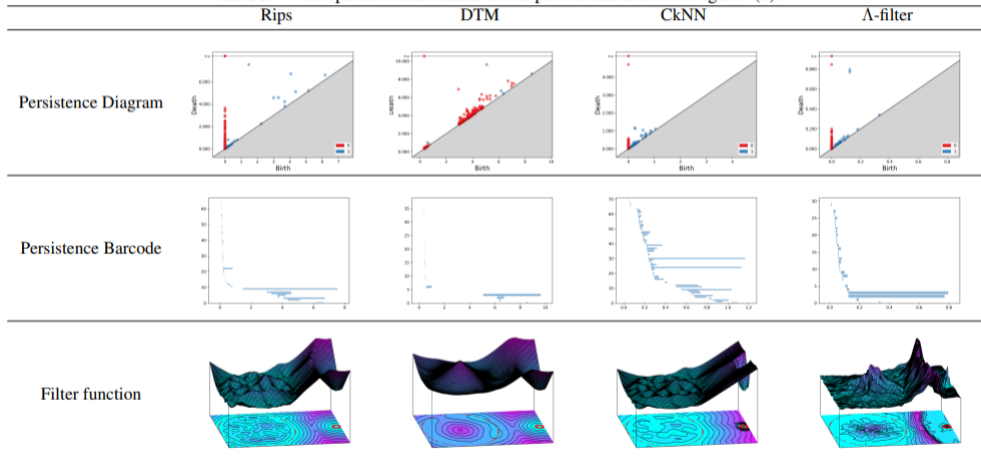
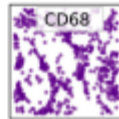
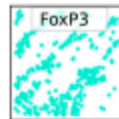
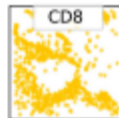
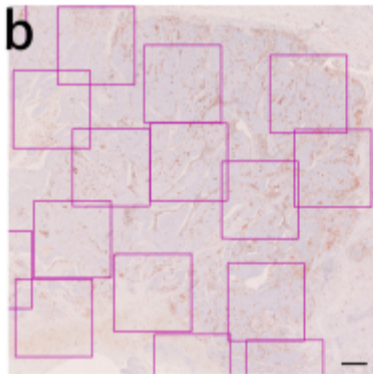
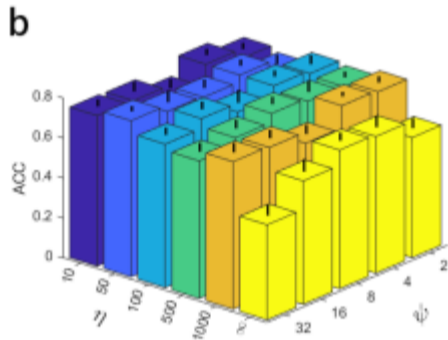
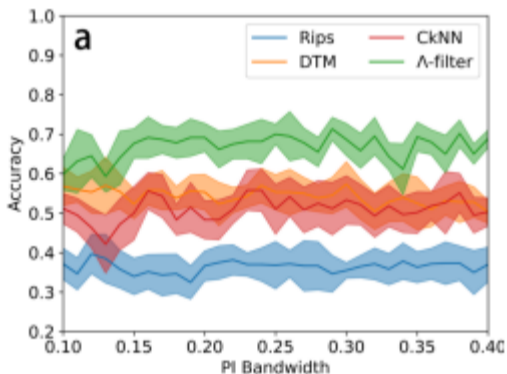


Table 3. A comparison result based on the point cloud shown in Figure 4(b).

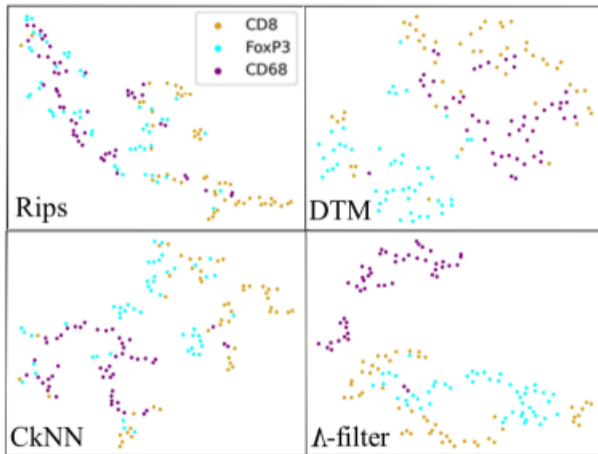




Classification of Bone Scripts



Visualization of Immune Cells



- ① Topology and Persistent Homology
- ② How to Construct a Correct PD
- ③ Λ -filter
- ④ Experiments
- ⑤ Conclusions

- Investigating the **shortcomings of existing methods** to produce correct and good PDs.
- Highlighting the importance of the problem of **the lack of robustness** of PD when noise and varied densities coexist in a point cloud.
- Proposing **a new data-dependent kernel and a new method** to produce a PD in order to address the above problem.
- Showing that using a good PD **significantly improves** the visualization outcome of t-Distributed Stochastic Neighbor Embedding and the classification accuracy of Support Vector Machines.