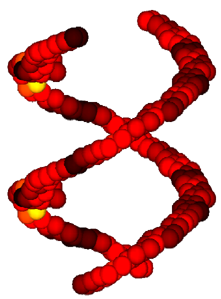
A Barcode Shape Descriptor for Curve Point Cloud Data

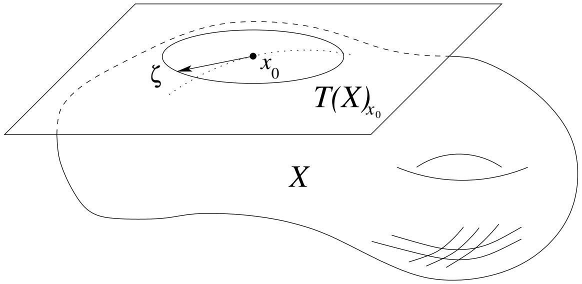


Anne Collins
Centre College

Afra Zomorodian, Gunnar Carlsson, and Leonidas Guibas
Departments of Mathematics and Computer Science
Stanford University

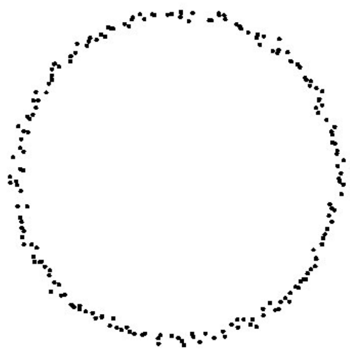
Plan

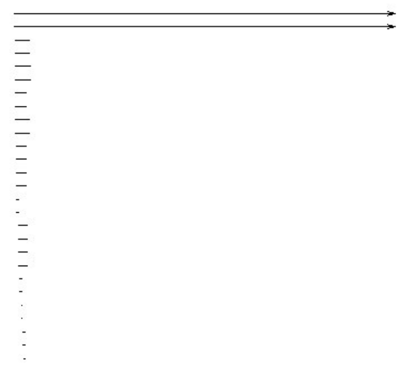
- Yesterday: *Theory*
  - Persistence Barcodes for Shapes by Afra Zomorodian



- Today: *Practice*
  - Point cloud data
  - Algorithms for construction
  - Detailed study

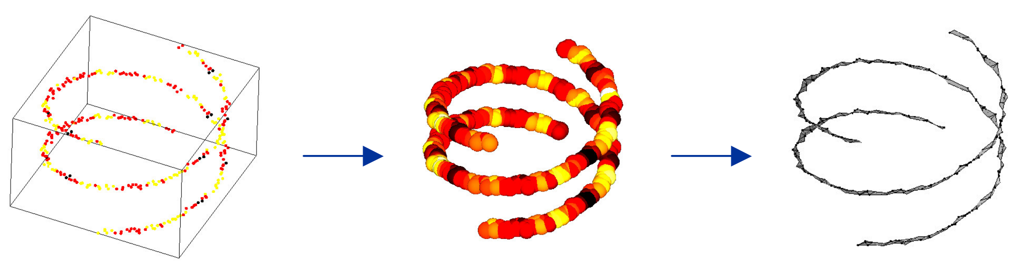
Overview





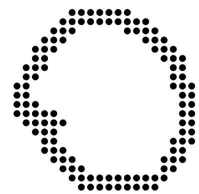
Input: Shape

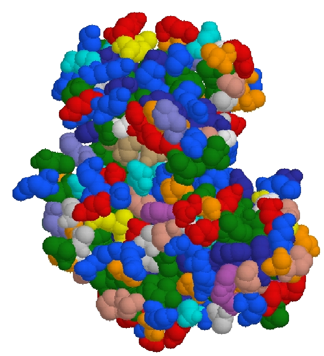
Output: Descriptor

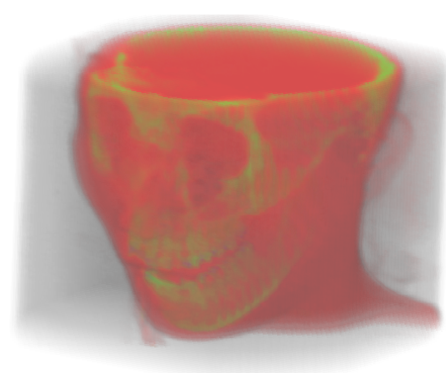


Point Cloud Data

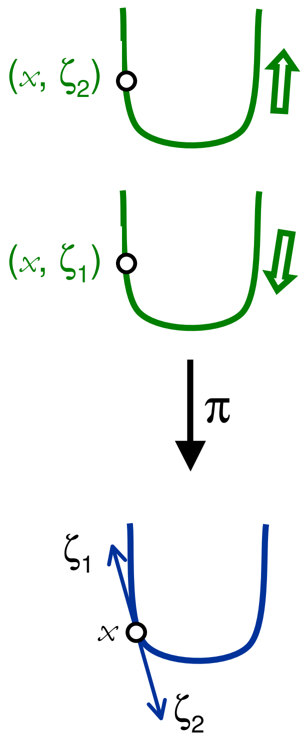
- Nearly all shape data
  - 2D: images
  - 3D: surface data, molecules
  - *n*D: shape spaces
- Traditional approach: construct useful structure
  - Vision: OCR
  - Graphics: surface reconstruction
  - CAD: conversion to B-spline surfaces
- Recent approach: first class object
  - Shape representation
  - Rendering
  - Modeling







Review of Tangent Complex

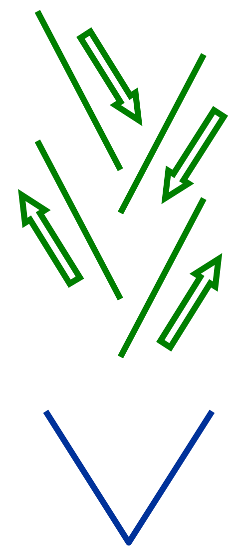


T(*X*) has two components: $\beta_0(T(X)) = 2$

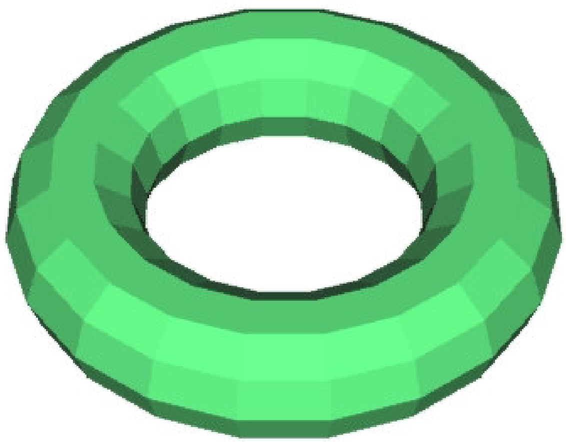
There are two points in its fiber $\pi^{-1}(x)$

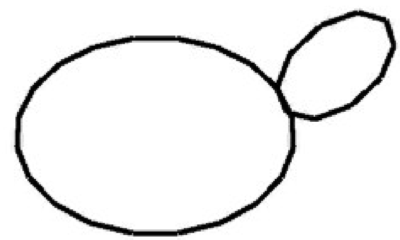
Every point *x* on a smooth curve *X* has two tangent directions.

A corner point has four tangent directions: $\beta_0(T(X)) = 4$

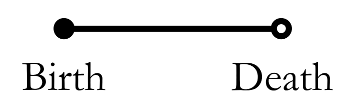


Review of Persistence



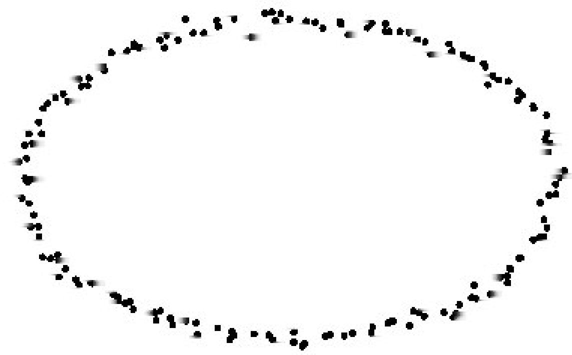


- As the space evolves, topological attributes are created and destroyed, each having a lifetime of existence.



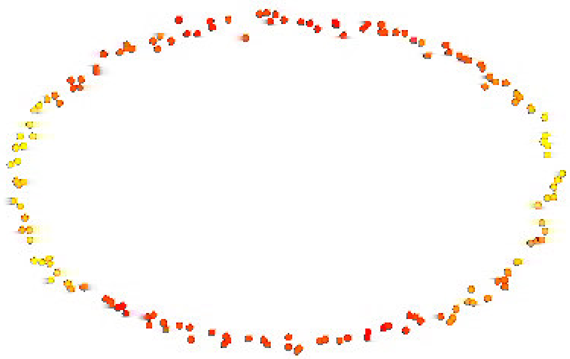
- Persistent Homology gives us these lifetimes as a collection of intervals, a persistence barcode.

Fibers



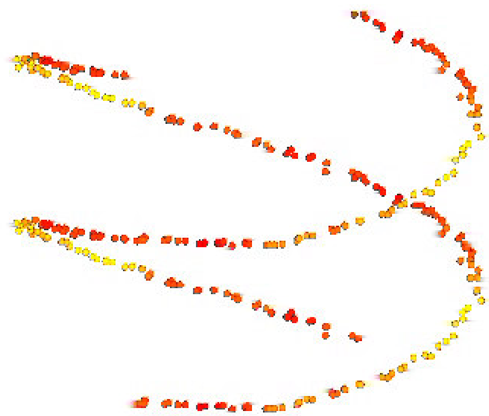
- PCD $P \subset X$, sampled from smooth closed 1-manifold
- We compute fibers $\pi^{-1}(P)$ by Total Least Squares using k-nearest neighbors (What is k?)

Filtering by Curvature



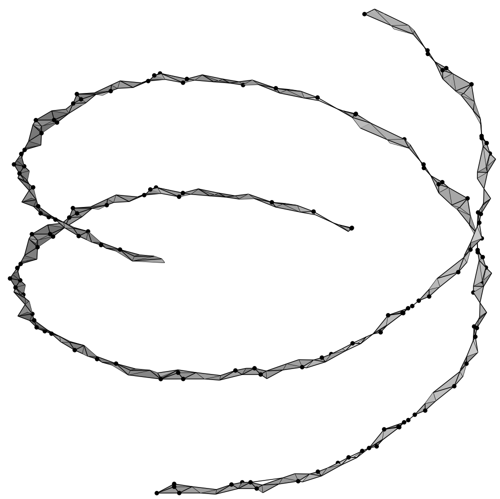
- Construct tangent complex incrementally
- Transform points to coordinate frame provided by tangent computation
- Fit osculating parabola to estimate curvature

Approximating $T(X)$



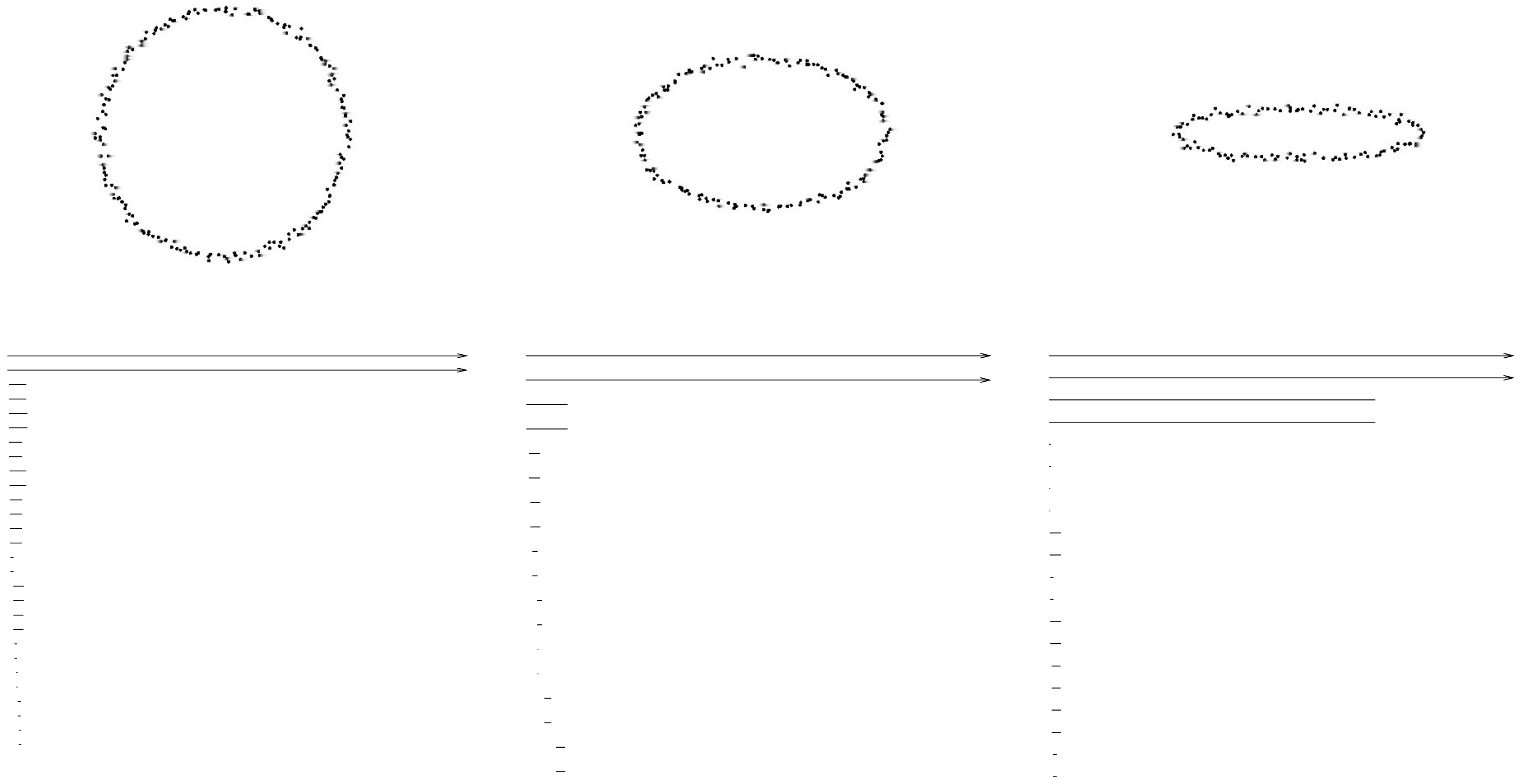
- $\mathbb{R}^n \times \mathbb{S}^{n-1}$ with $ds^2 = dx^2 + \boldsymbol{\omega}^2\, d\boldsymbol{\zeta}^2$
- $T(X) \approx \bigcup_{p \in \boldsymbol{\pi}^{-1}(P)} \mathrm{B}_{\boldsymbol{\varepsilon}}(p)$

Complex

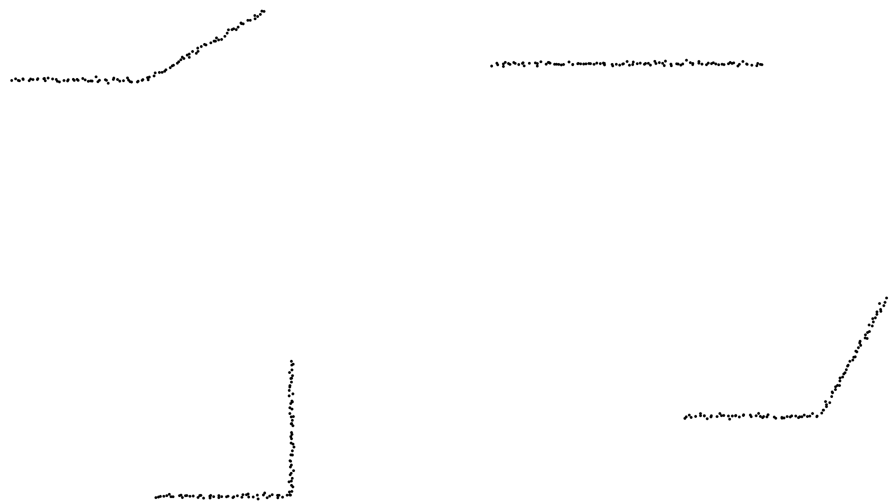


- Rips complex
- $R_\varepsilon(M) = \{\text{conv } T \mid T \subseteq M, \text{d}(s, t) \leq \varepsilon, s,t \in T\}$

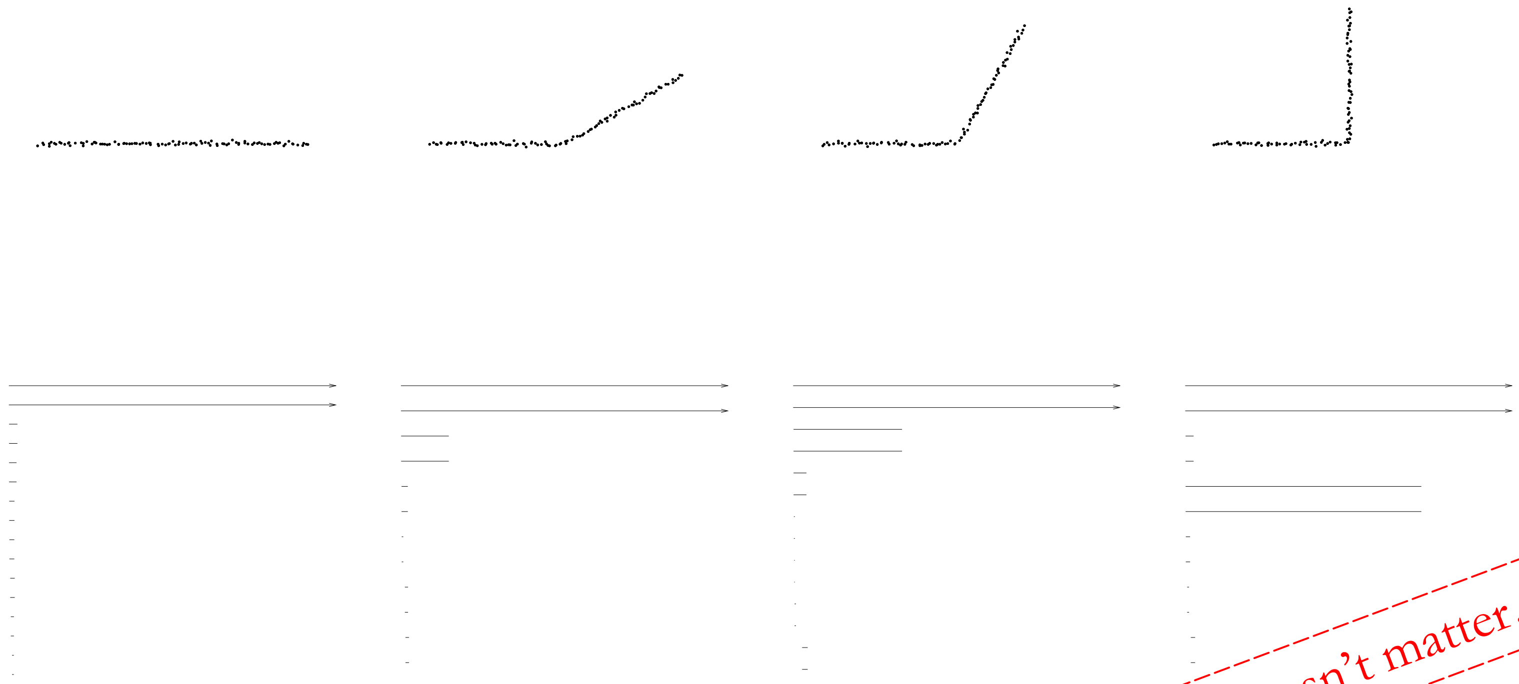
Family of Ellipses



Articulated Arm Family



Parameterization



Order doesn't matter…

Barcode Metric Algorithm

Three steps: $D = D_1 + D_2 + D_3$

1. Count infinite intervals

$$D_1(B_1, B_2) = \begin{cases} 0 & \text{if equal,} \\ \infty & \text{otherwise} \end{cases}$$

2. Match infinite intervals (same number)
   - Sort by birth times
   - Match in order

$$D_2(B_1, B_2) = \text{sum of paired length differences}$$

Finite Interval Distance

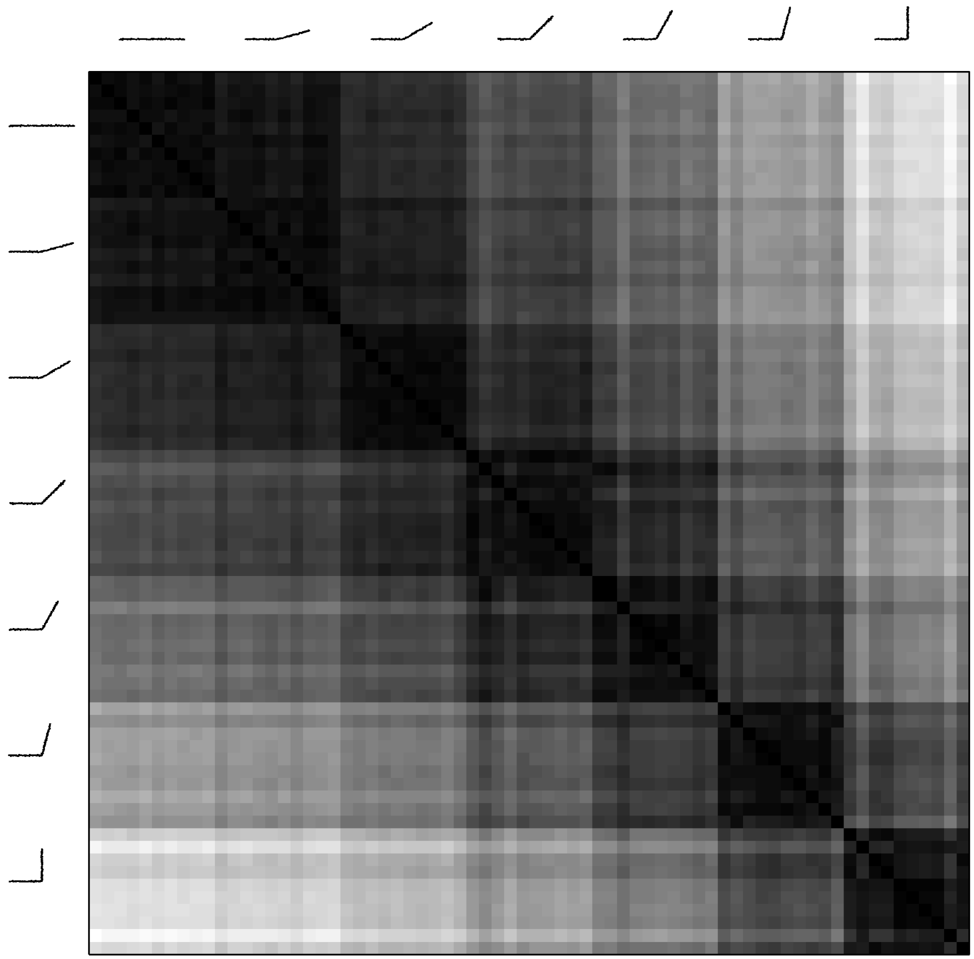
3. Match finite intervals

   - Matching M = pairing of intervals
   - D_M = dissimilarity between matched pairs
     + lengths of unmatched intervals
   - $D_3(B_1, B_2) = \min_M D_M(B_1, B_2)$
   - Minimizing Dissimilarity = Maximizing Similarity
   - Maximum Weighted Bipartite Matching (LEDA)

$\delta(I, J) = |I \cup J - I \cap J|$

Barcodes — Bipartite graph

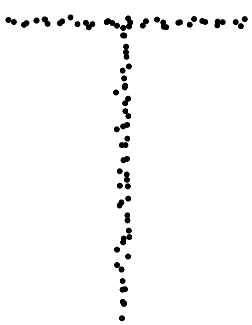
Distance Matrix

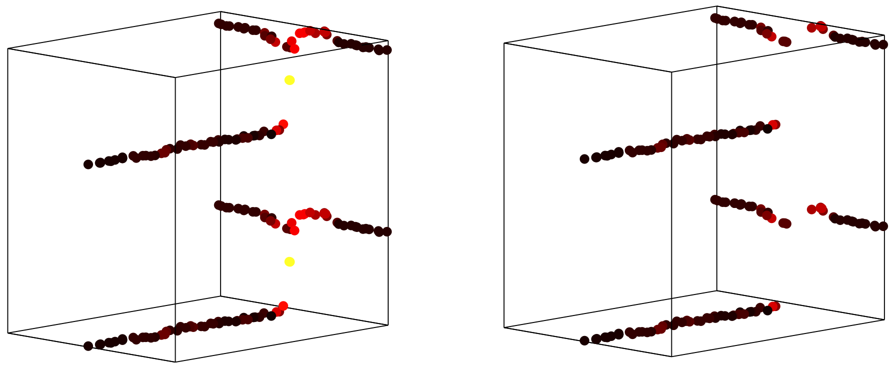


Extensions

- Boundary points
- Non-manifold points
- Singularities

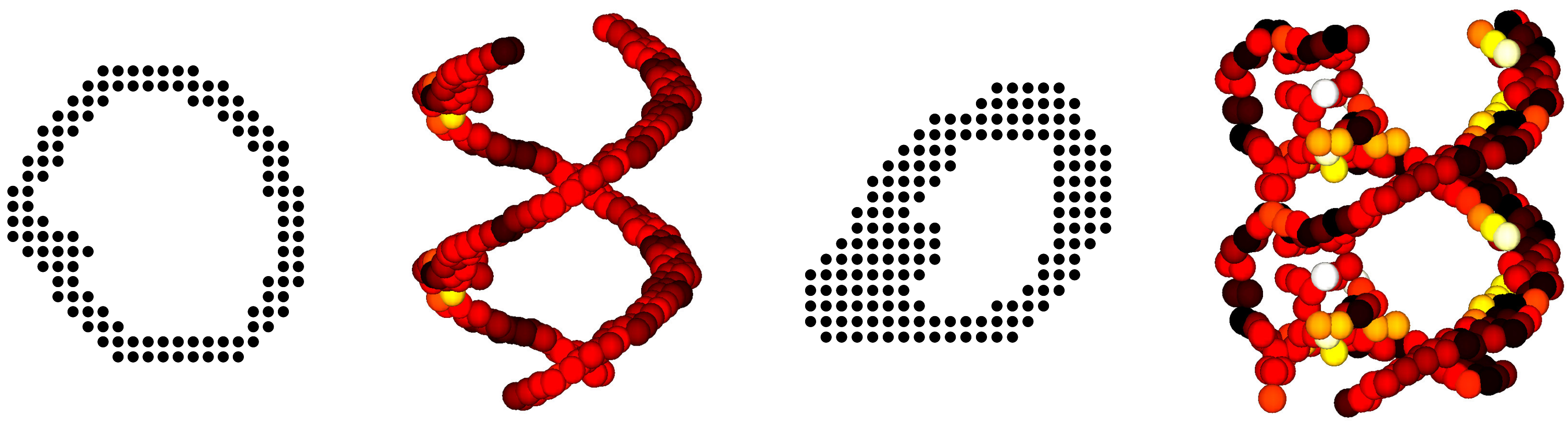
} Remove points near singularities



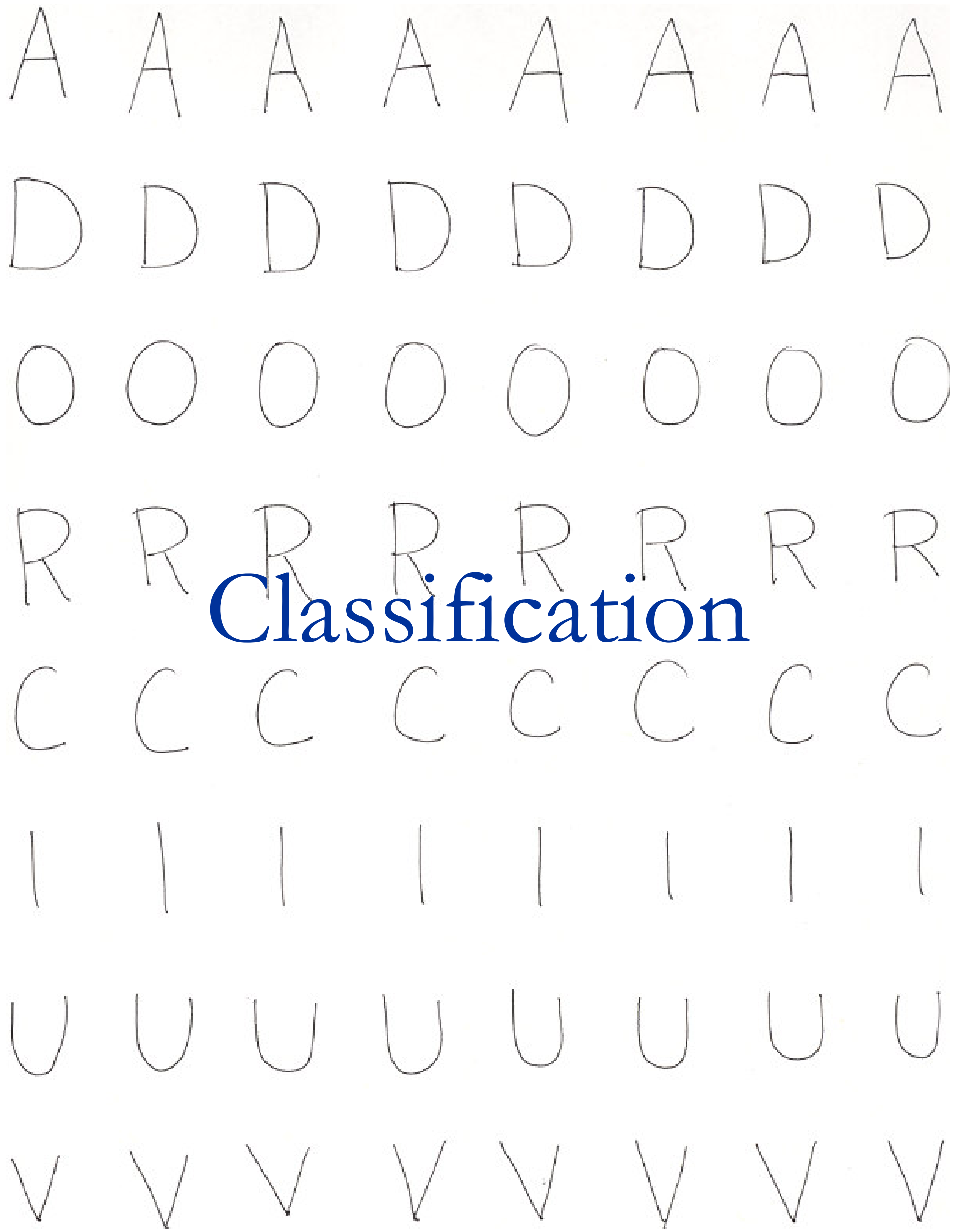


Noise

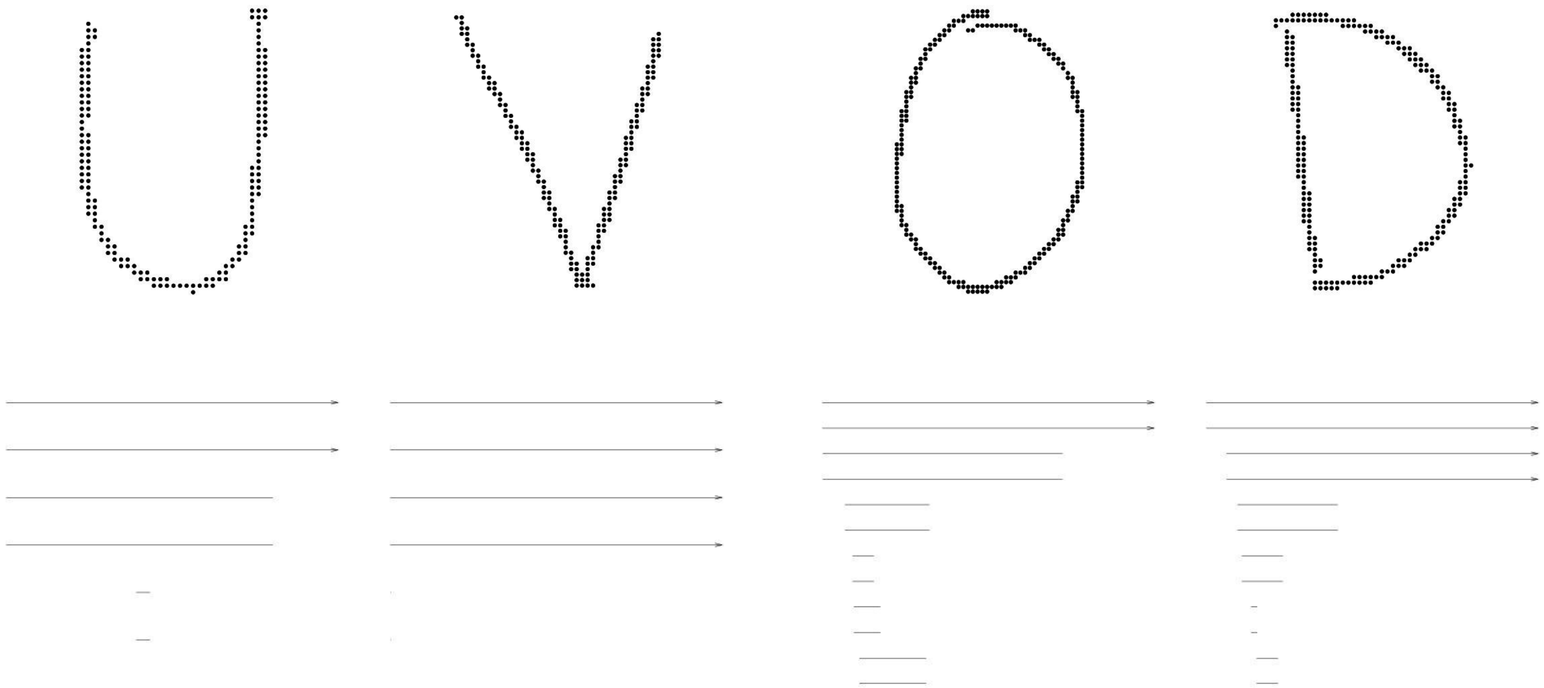
- Outliers introduce spurious components
- Thick curves affect tangent/curvature estimates
- MNIST database of handwritten digits



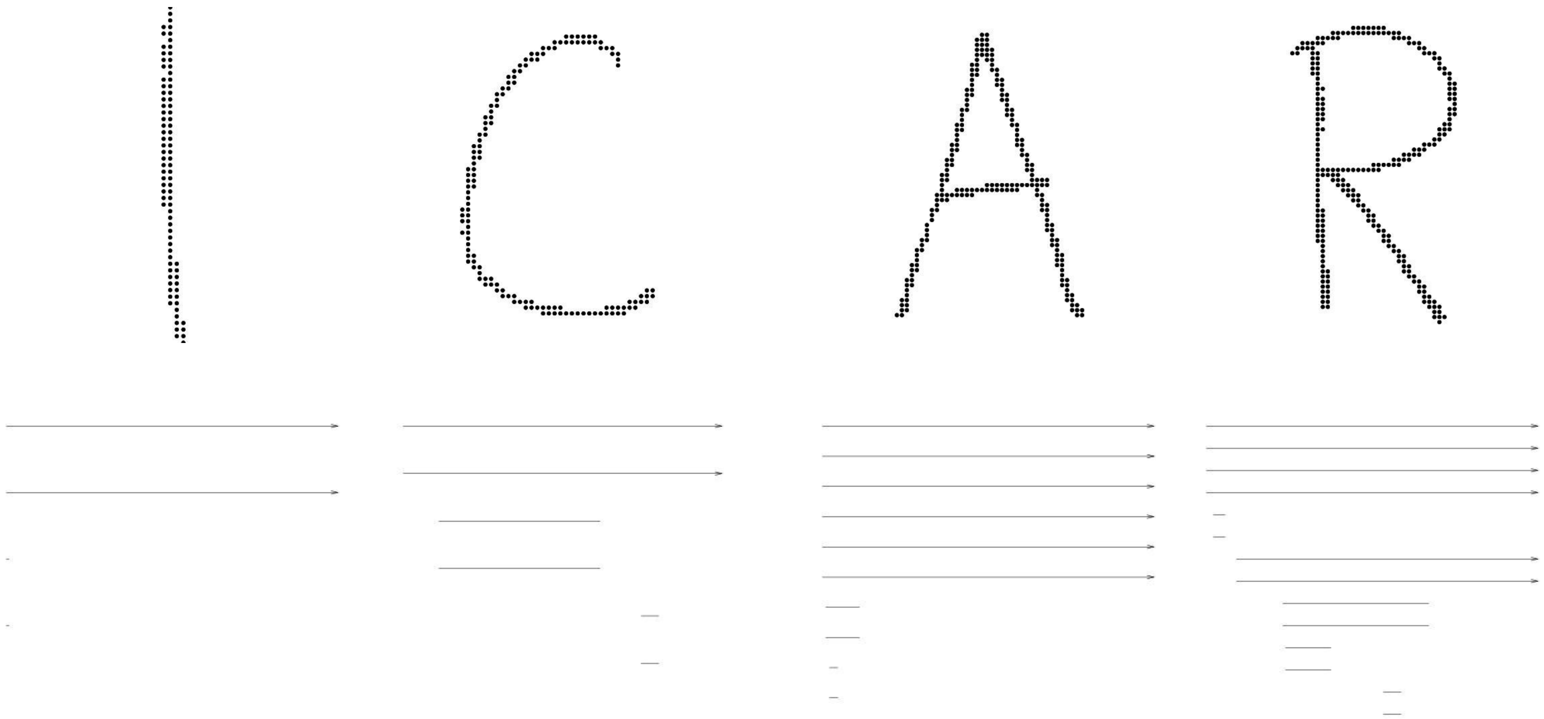
Classification



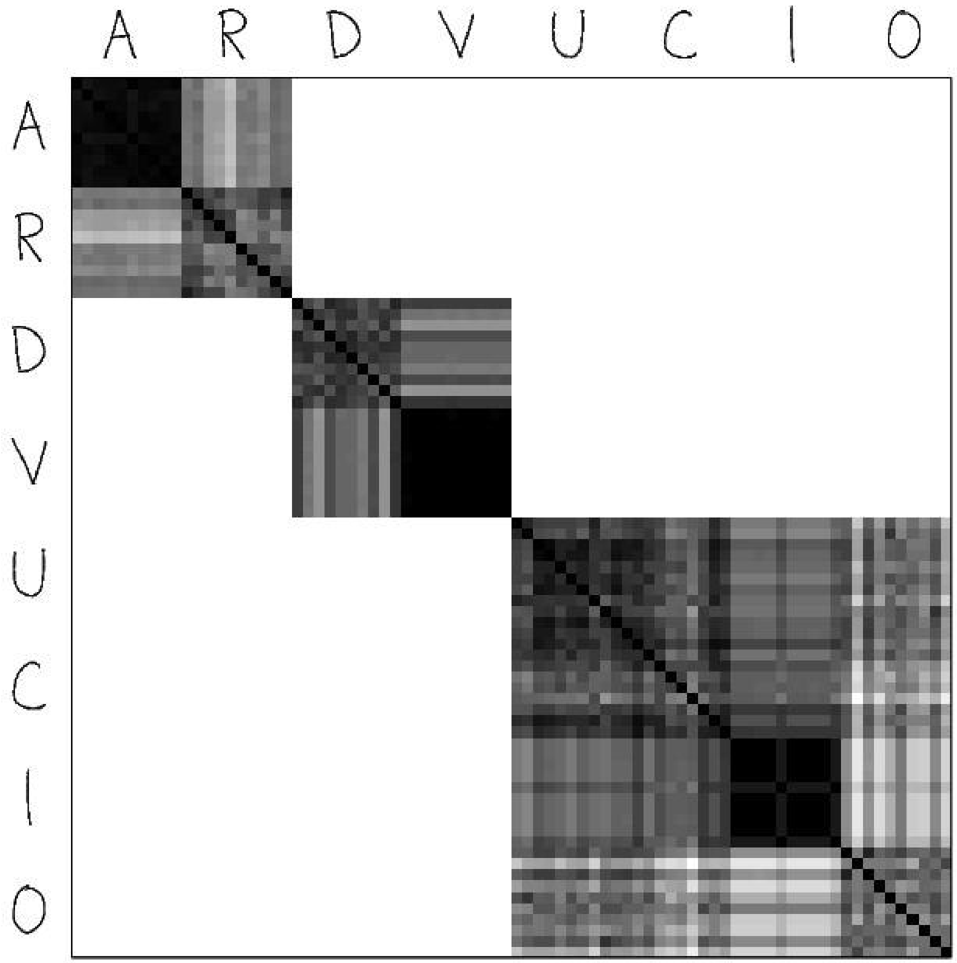
Singularities



Curvature

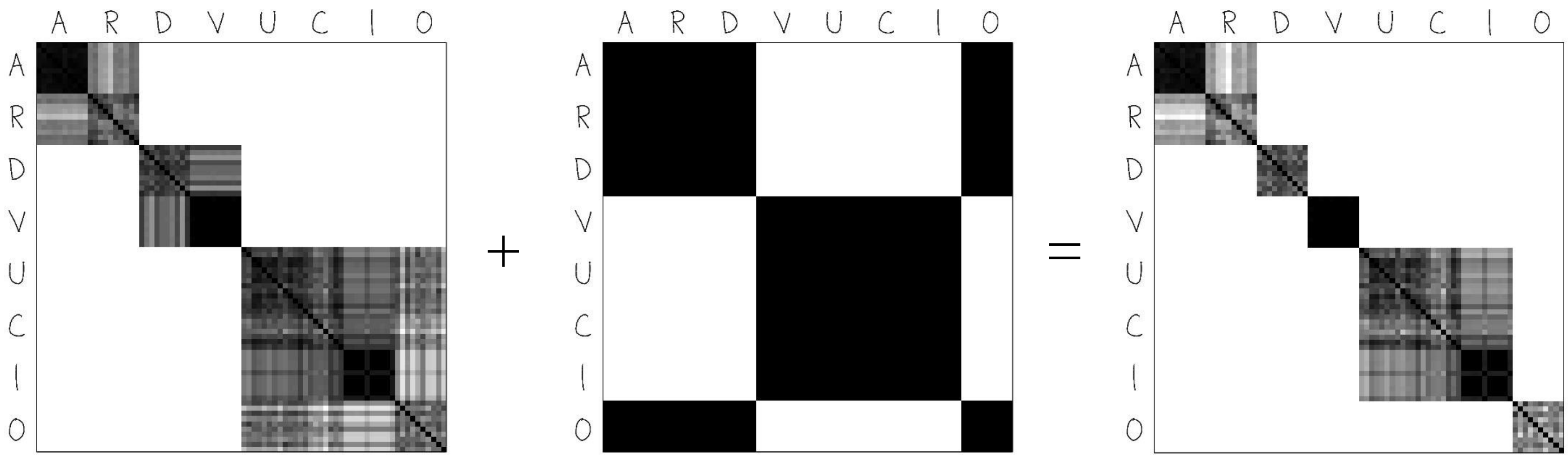


Curvature Signature



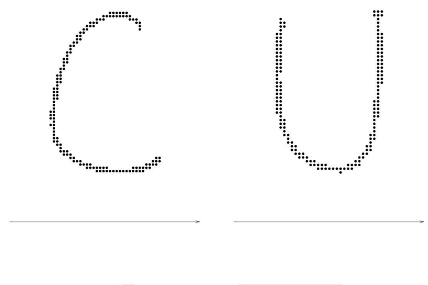
Multiple Signatures

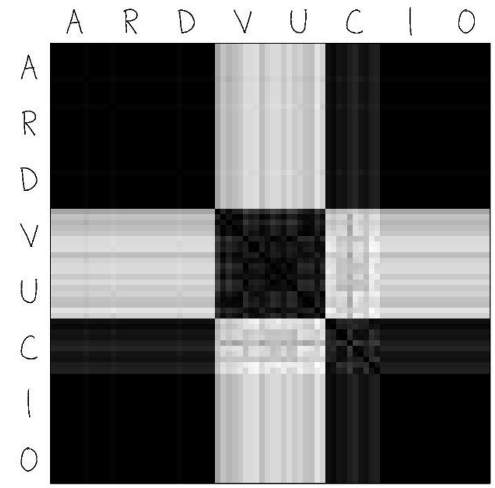
- Improve classification with additional information
  - New filtration
  - New derived space
  - New metric
- eg: include β_1 of PCDs



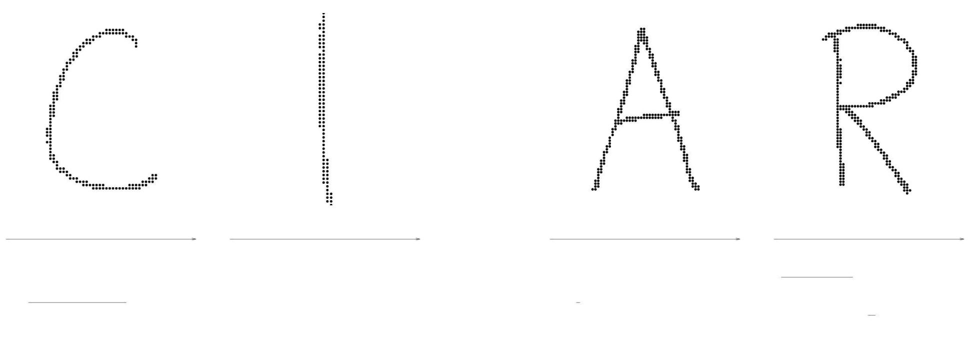
Directional Filtrations

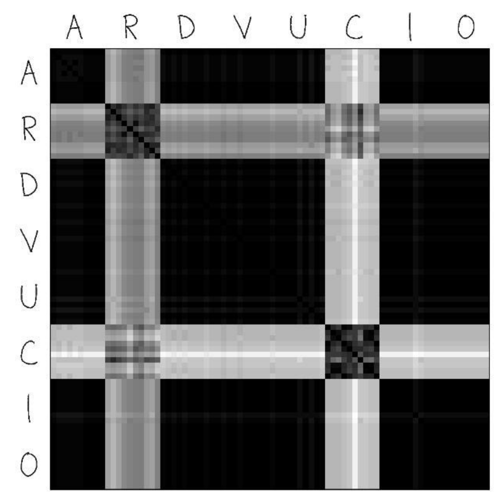
- Top-Down β_0 of PCDs



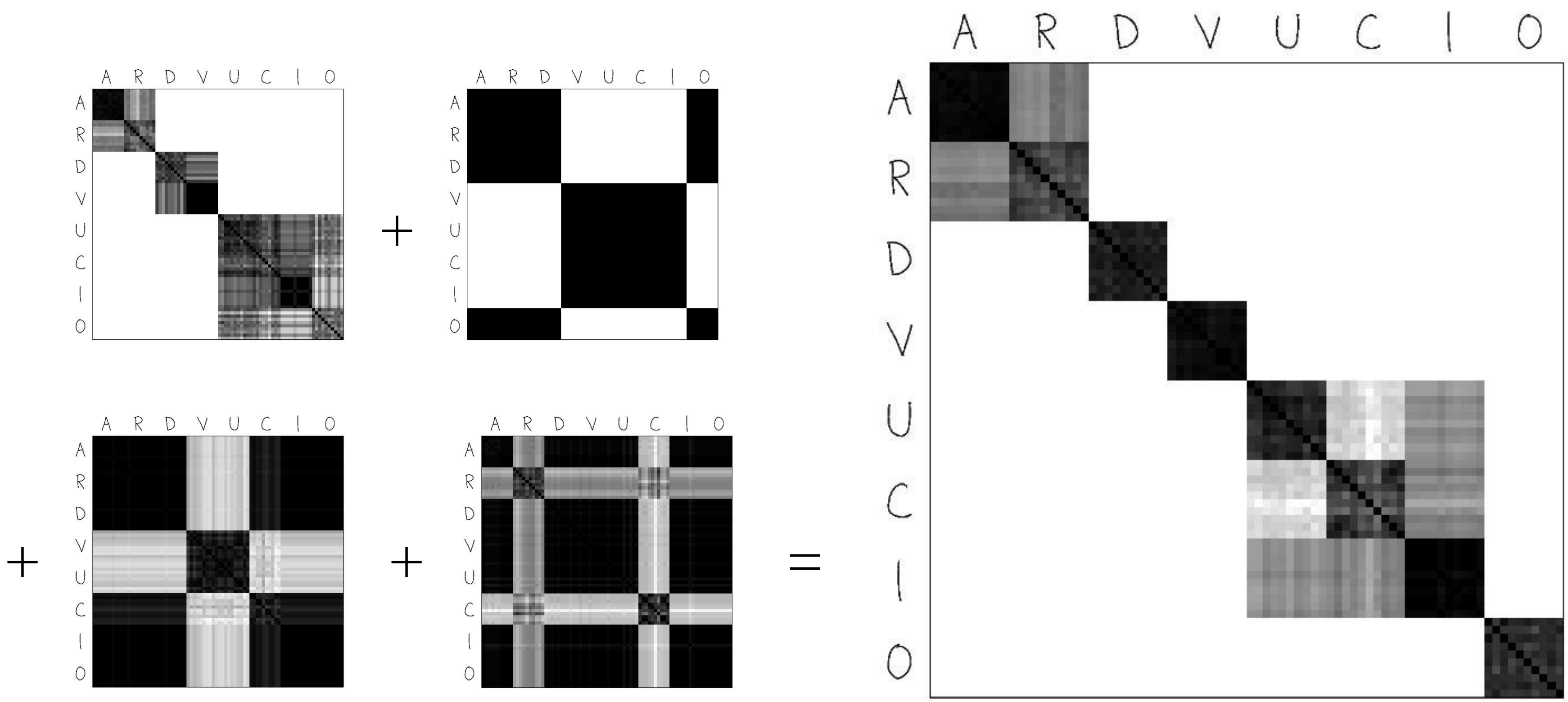


- Right-Left β_1 of PCDs





Final Distance Matrix



Conclusion

- Apply persistent homology to geometry-rich derived space to get a compact shape descriptor called barcode
- Compute tangent complexes for Curve Point Cloud Data
- Improve classification with multiple signatures
- Future Work:
  - Better handling of non-manifold points, outliers, curve thickness, etc.
  - Surface PCD
  - Complexes with lower intrinsic dimension
  - Learning Algorithms