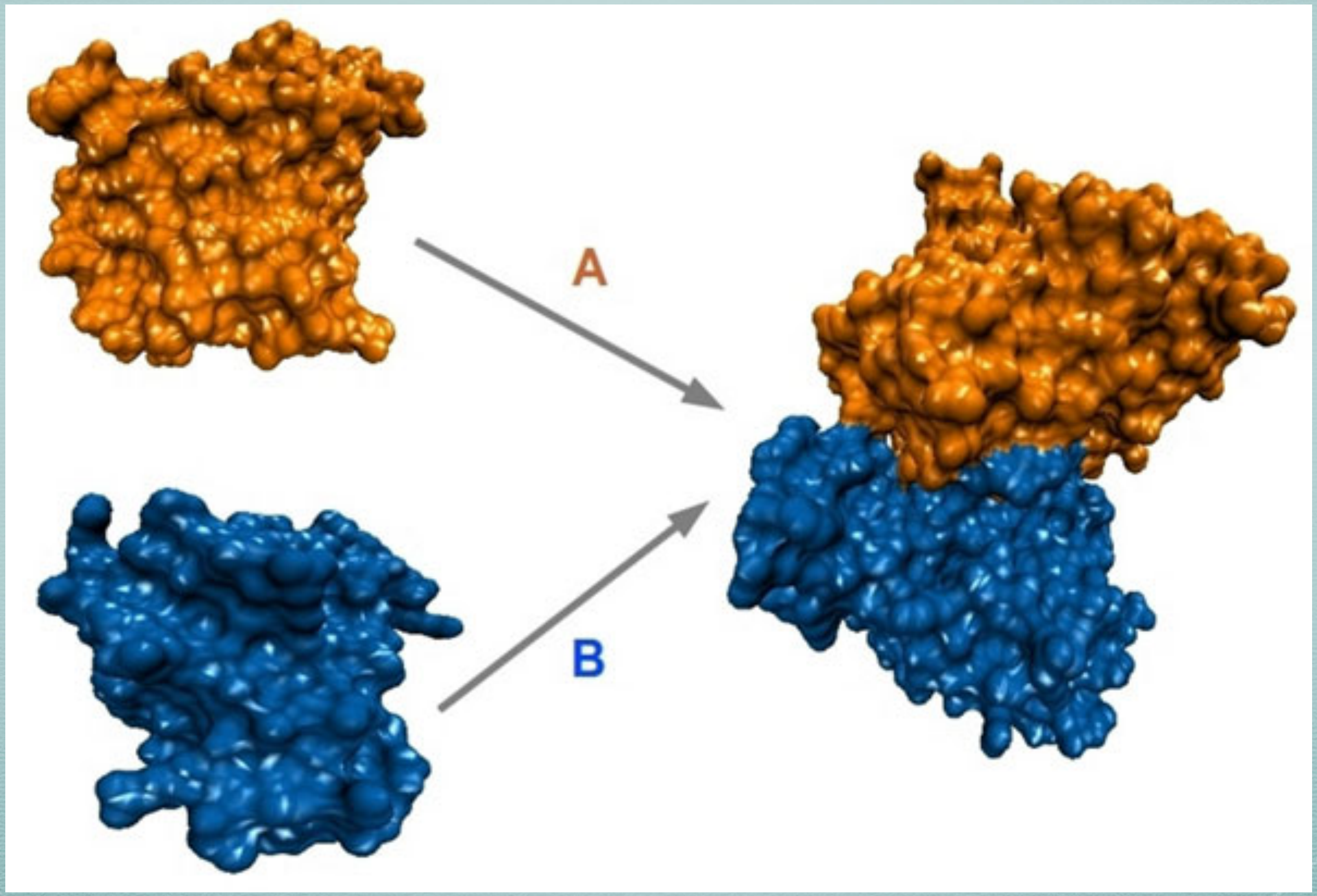
Persistent homology

Stability and applications

Inception

Edelsbrunner's motivation: Protein docking

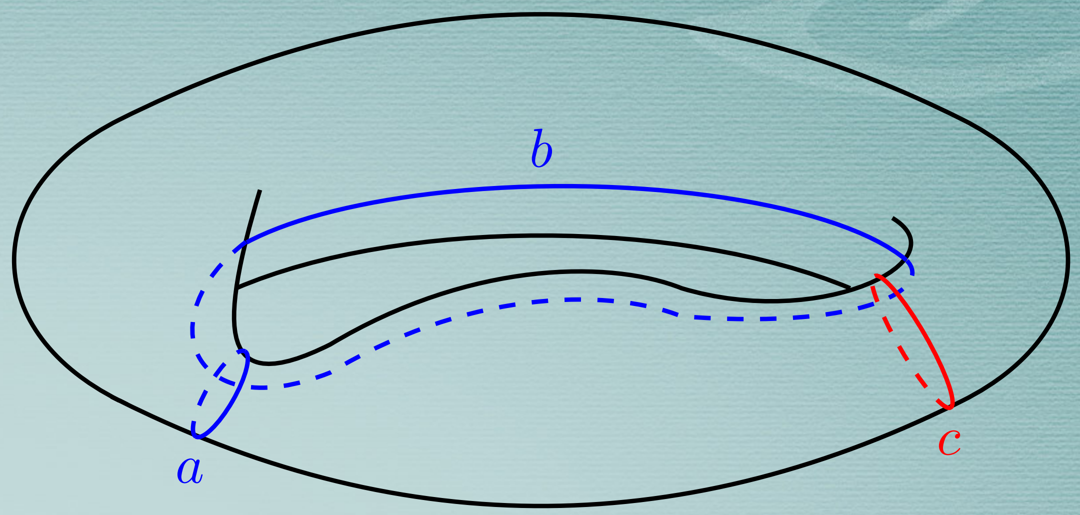


Irina Hashmi and Amarda Shehu: HopDock: A probabilistic search algorithm for decoy sampling in protein-protein docking

Homology

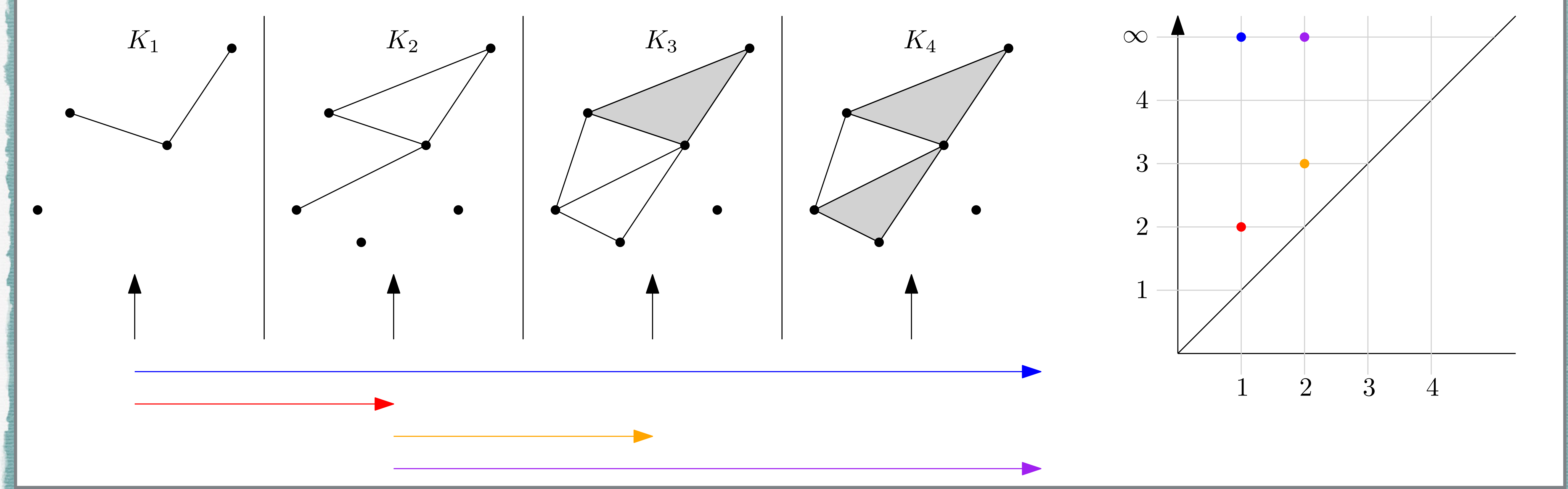
Counts holes:

$\beta_0 = 0; \quad \beta_1 = 2; \quad \beta_2 = 1$



Persistent homology

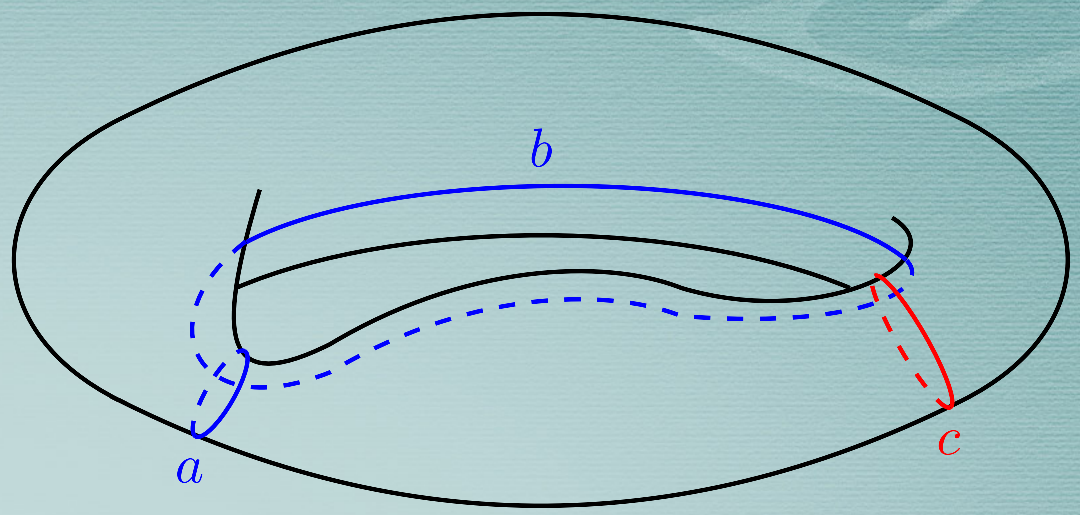
Describes evolution of holes:



Homology

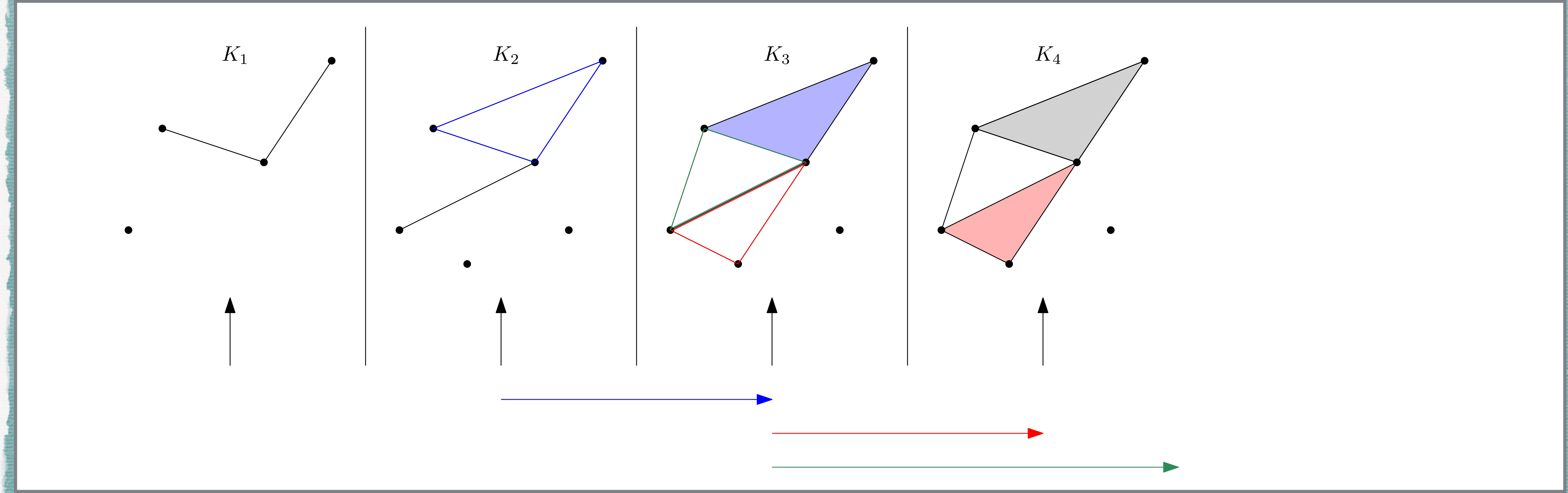
Counts holes:

$\beta_0 = 0; \quad \beta_1 = 2; \quad \beta_2 = 1$



Persistent homology

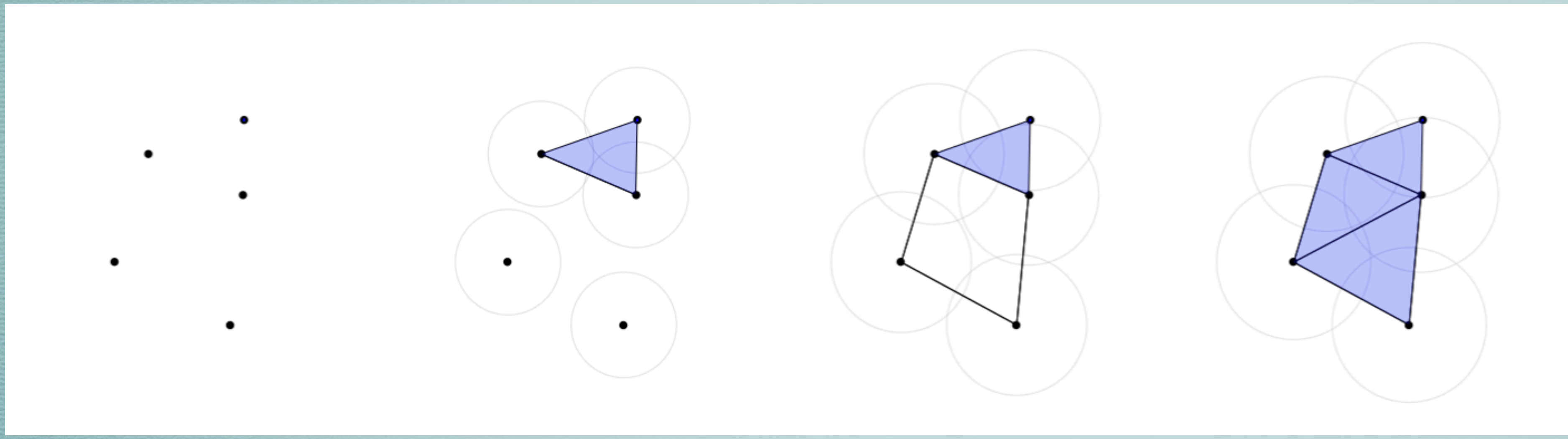
Describes evolution of holes



TDA pipeline

TDA pipeline

Space or pointcloud → Filtration



Rips complex of a metric space X at $r > 0$:

$$\sigma \subset X \in \mathrm{Rips}(X, r) \Leftrightarrow \mathrm{diam}(\sigma) \leq r$$

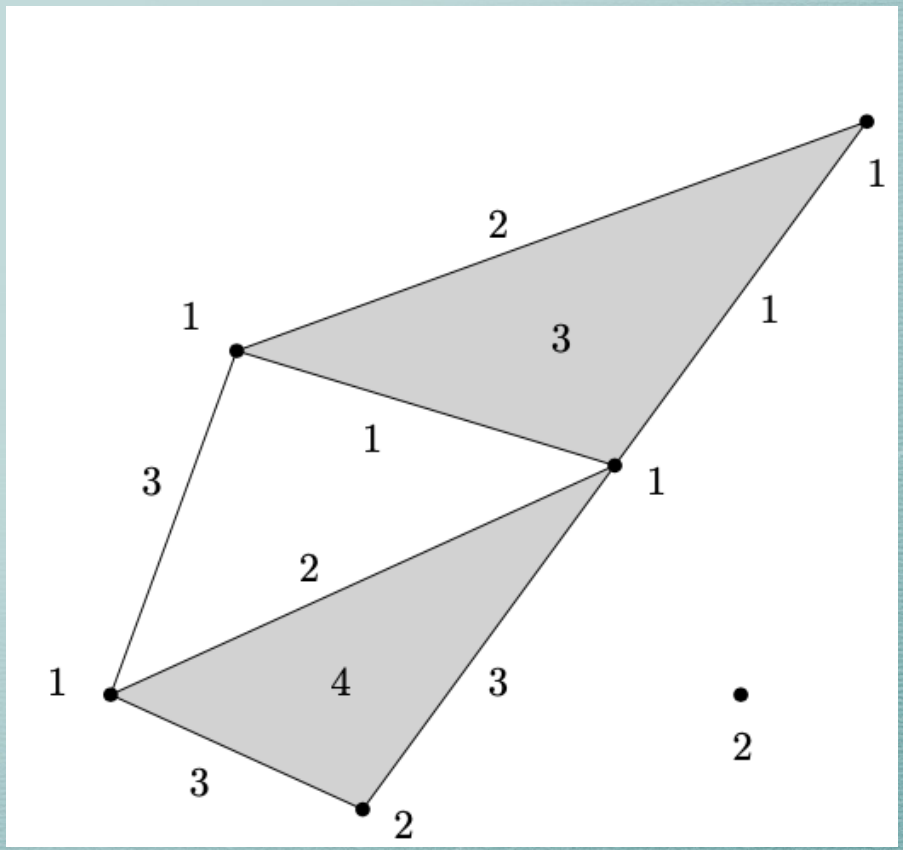
TDA pipeline

Space or pointcloud → Filtration → Computing persistent homology

$$
\begin{array}{c|cc|cc|ccc}
 & \langle b,c\rangle & \langle b,d\rangle & \langle a,b\rangle & \langle c,d\rangle & \langle a,c\rangle & \langle a,e\rangle & \langle b,e\rangle \\
\hline
\langle a\rangle & & & -1 & & -1 & -1 & \\
\langle b\rangle & -1 & -1 & 1 & & & & -1 \\
\langle c\rangle & 1 & & & -1 & 1 & & \\
\langle d\rangle & & 1 & & 1 & & & \\
\hline
\langle e\rangle & & & & & & 1 & 1 \\
\langle f\rangle & & & & & & &
\end{array}
$$

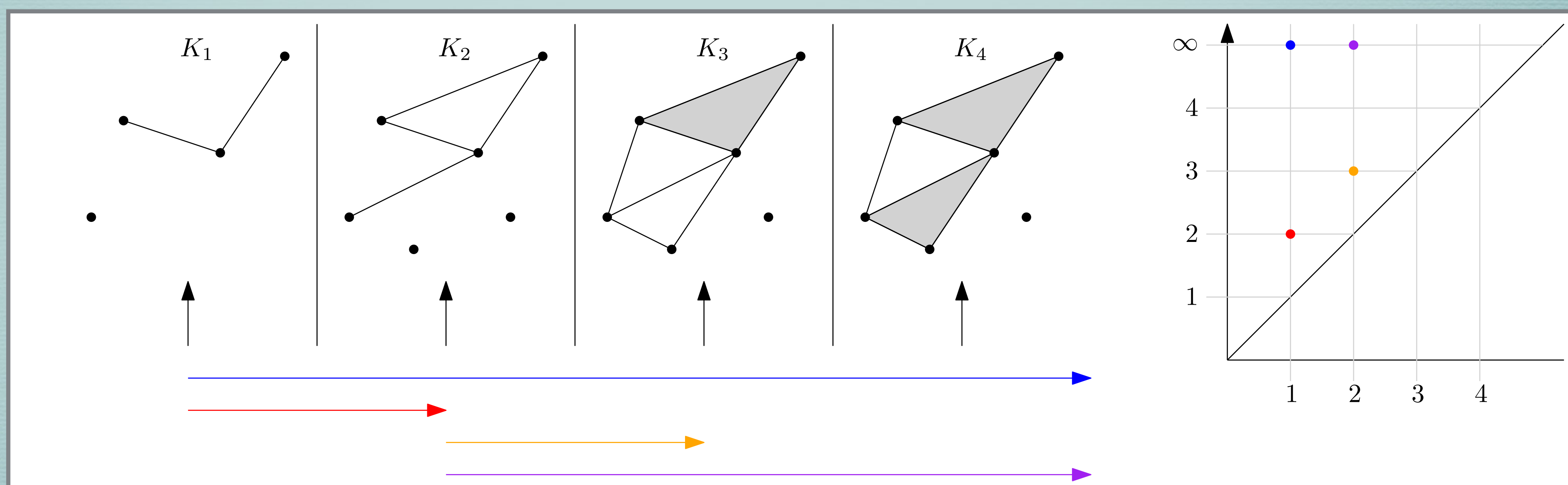
↓

$$
\begin{array}{c|cc|cc|ccc}
 & \langle b,c\rangle & \langle b,d\rangle & \langle a,b\rangle & \langle c,d\rangle & \langle a,c\rangle & \langle a,e\rangle & \langle b,e\rangle \\
\hline
\langle a\rangle & & & -1 & & & -1 & \\
\langle b\rangle & -1 & -1 & \boxed{1} & & & & \\
\langle c\rangle & \boxed{1} & & & & & & \\
\langle d\rangle & & \boxed{1} & & & & & \\
\hline
\langle e\rangle & & & & & & \boxed{1} & \\
\langle f\rangle & & & & & & &
\end{array}
$$



TDA pipeline

Space or pointcloud → Filtration → Computing persistent homology → Barcode



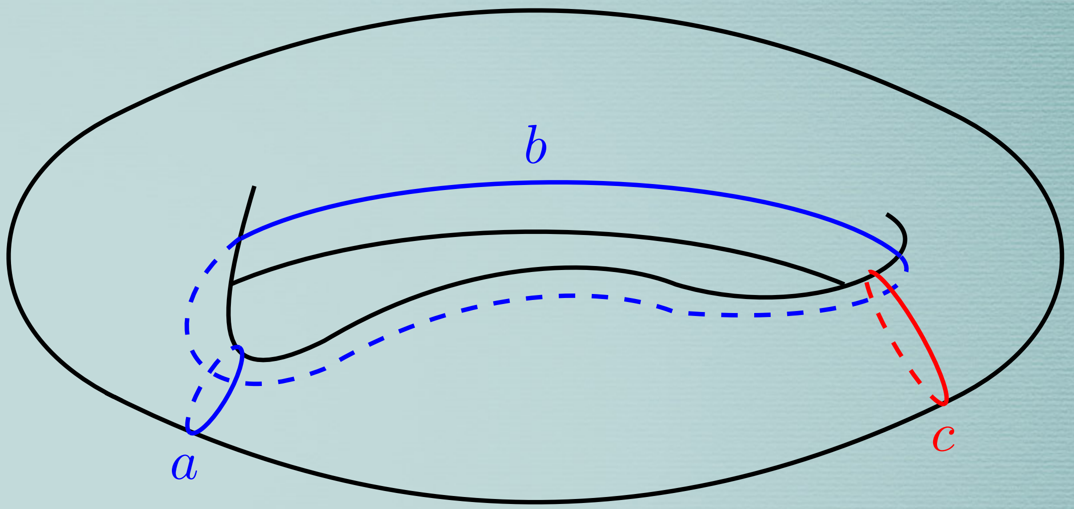
	$\langle b,c\rangle$	$\langle b,d\rangle$	$\langle a,b\rangle$	$\langle c,d\rangle$	$\langle a,c\rangle$	$\langle a,e\rangle$	$\langle b,e\rangle$
$\langle a\rangle$			-1			-1	
$\langle b\rangle$	-1	-1	1				
$\langle c\rangle$	1						
$\langle d\rangle$		1					
$\langle e\rangle$						1	
$\langle f\rangle$							

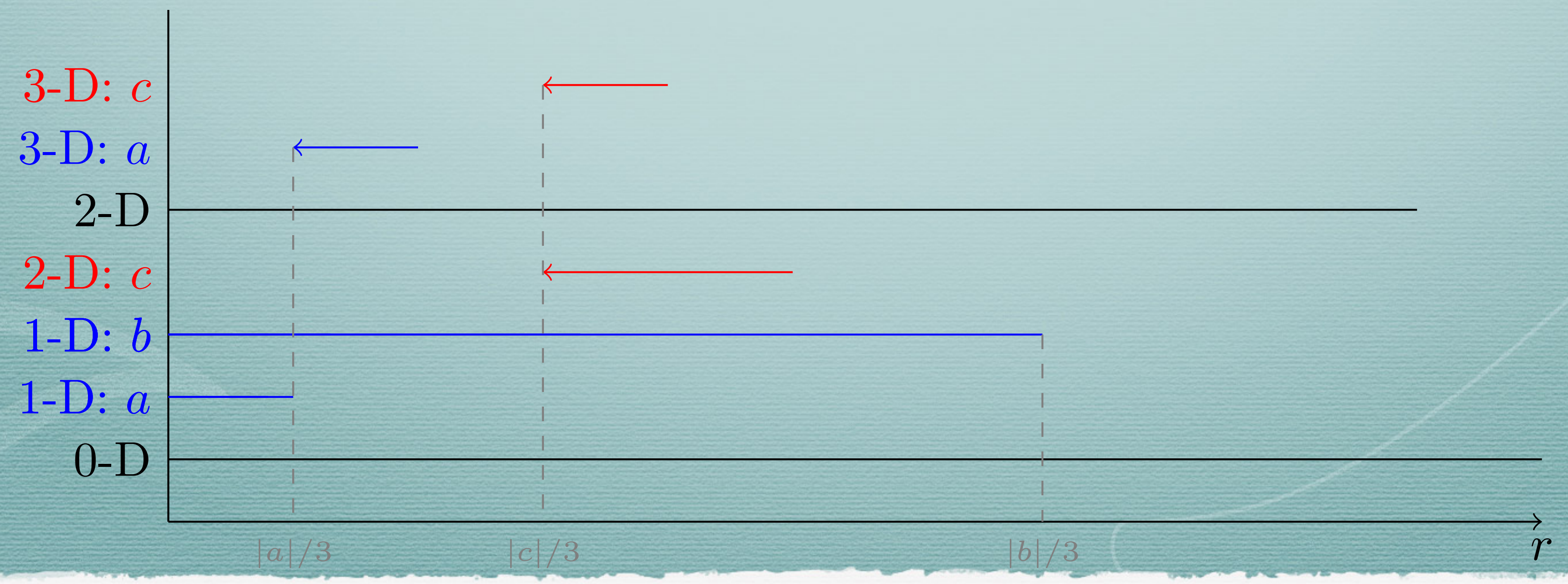
What to expect from persistence diagrams?

What to expect from a PD?

Expectations: Given a space *X*, PH via Rips or Cech complexes looks as follows:

- Initially homology of *X* at certain intermediate scales
- Only one 0-bar at infinity (contractibility)
- Mysterious homology in between.





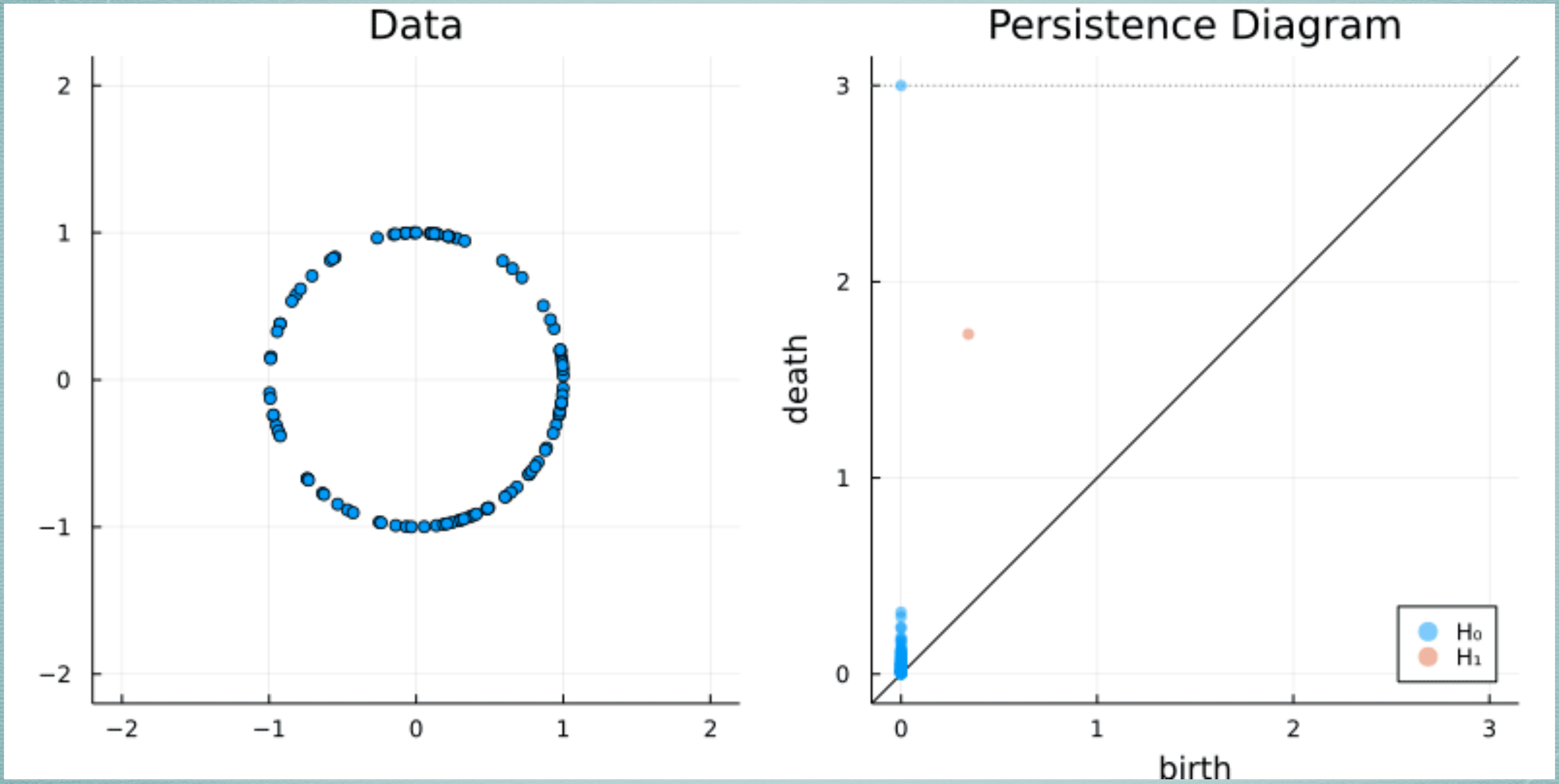
What to expect from a PD?

Expectations: Given a space *X* and a fairly good approximation *S*, PH via Rips or Cech complexes looks as follows:

- Initially a lot of 0-bars (discrete set) for very small scales
- Homology of *X* at certain intermediate scales
- Only one 0-bar at infinity (contractibility)
- Mysterious homology in between.

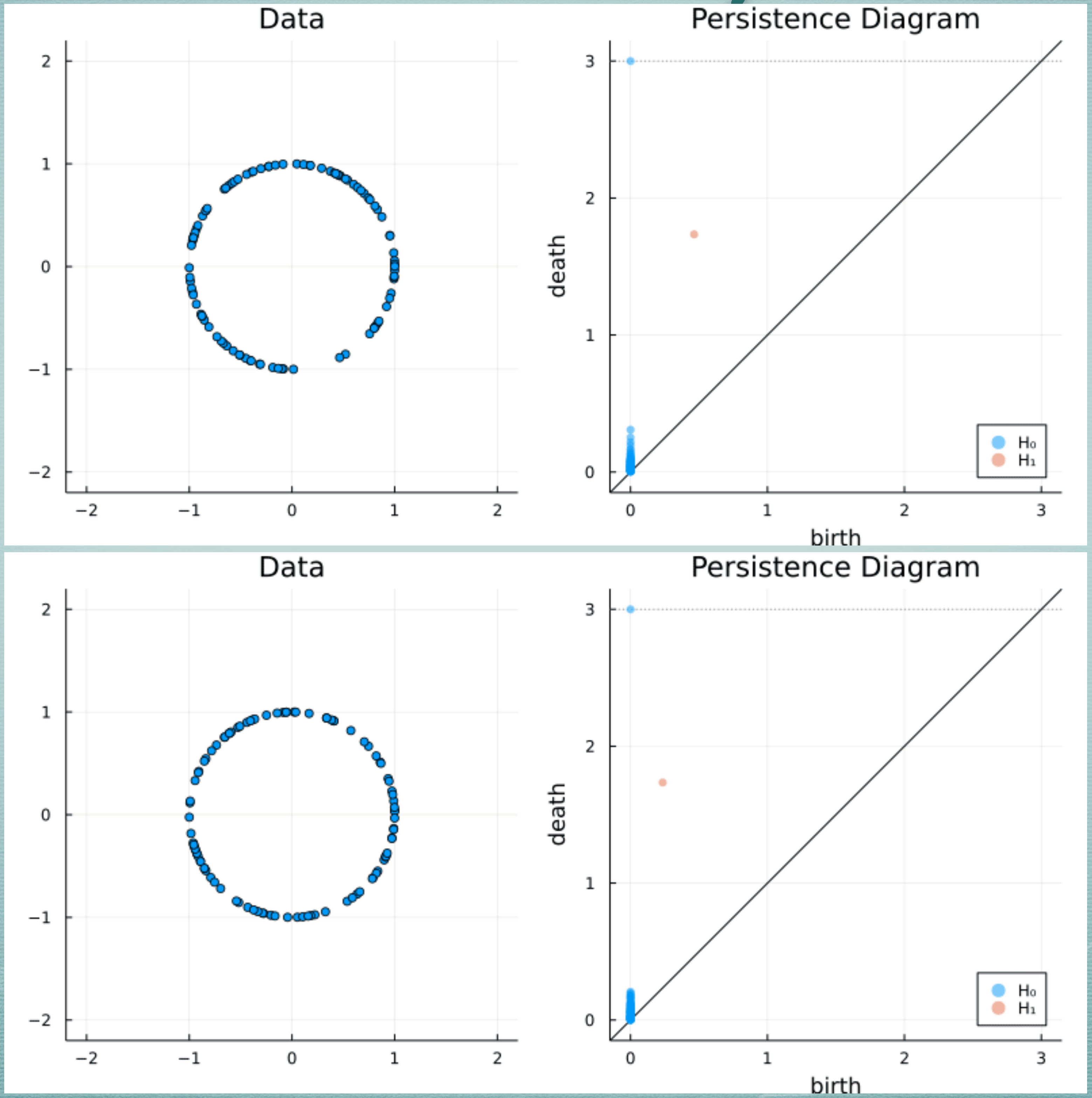
Stability

Stability

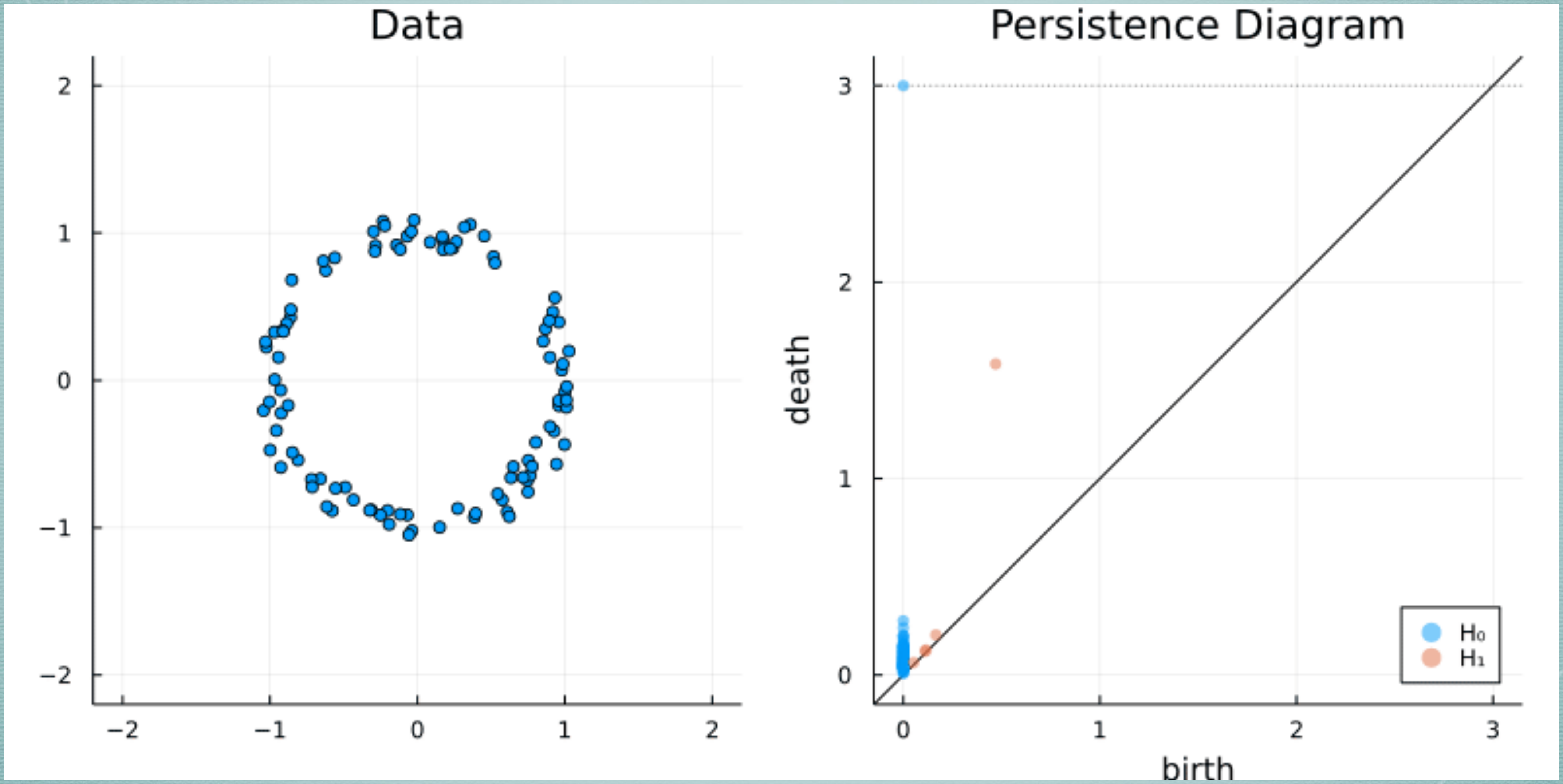


Generated using Ripserer.jl by Matija Čufar

Stability

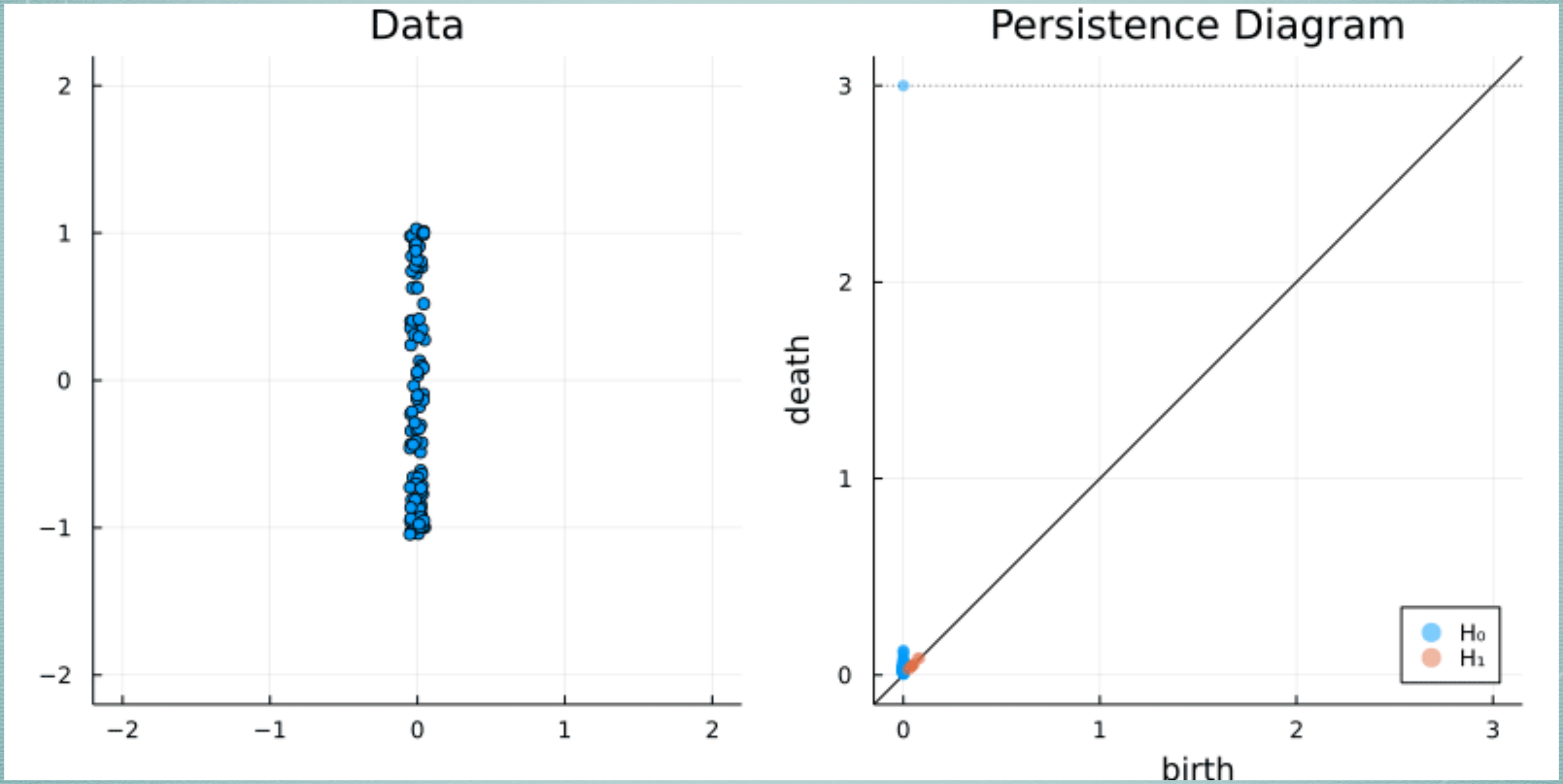

Data
Persistence Diagram
death
birth
H0
H1
Data
Persistence Diagram
death
birth
H0
H1

Stability



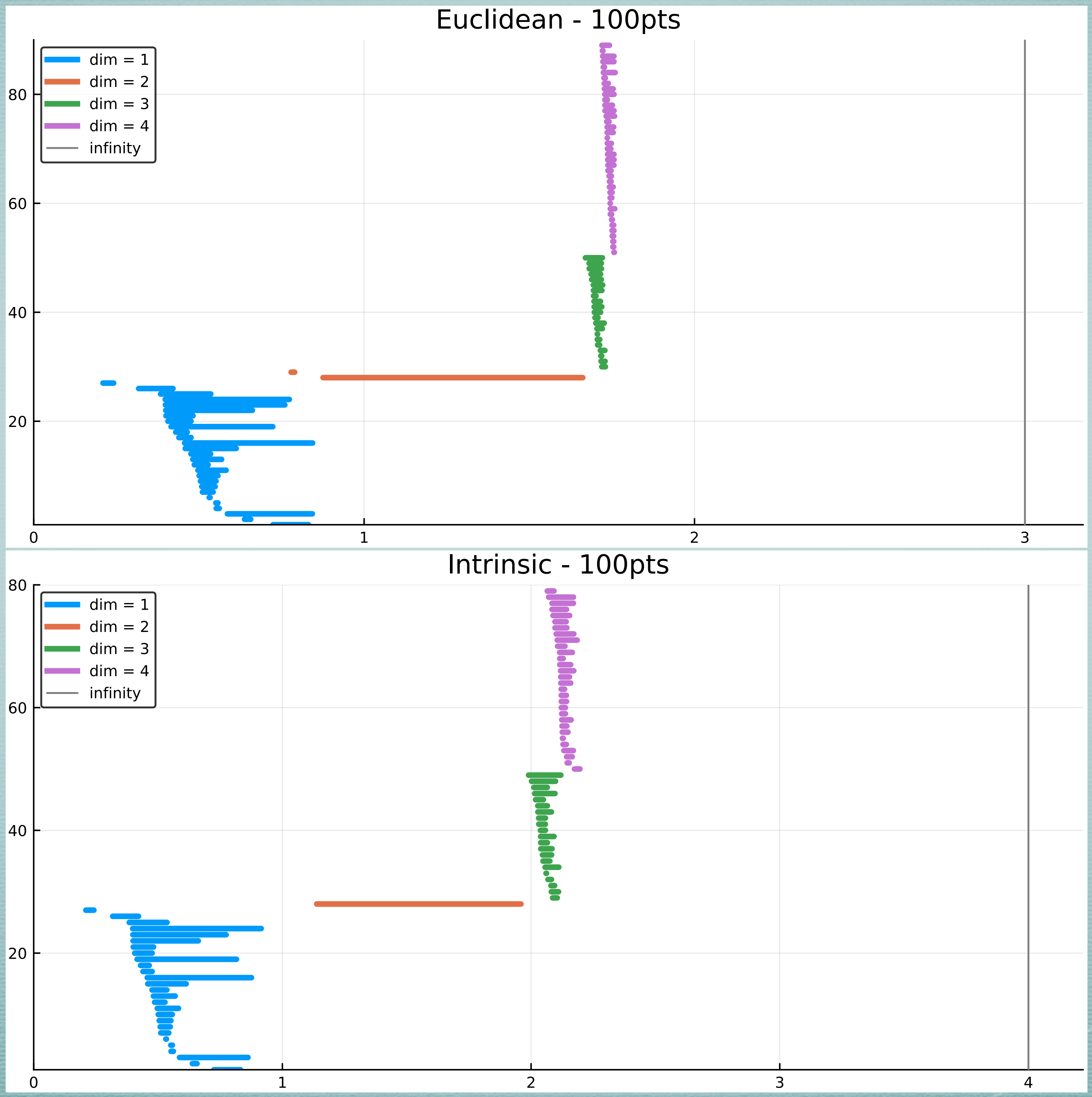
Data
2
1
0
−1
−2
−2
−1
0
1
2
Persistence Diagram
death
birth
3
2
1
0
0
1
2
3
H0
H1

Stability

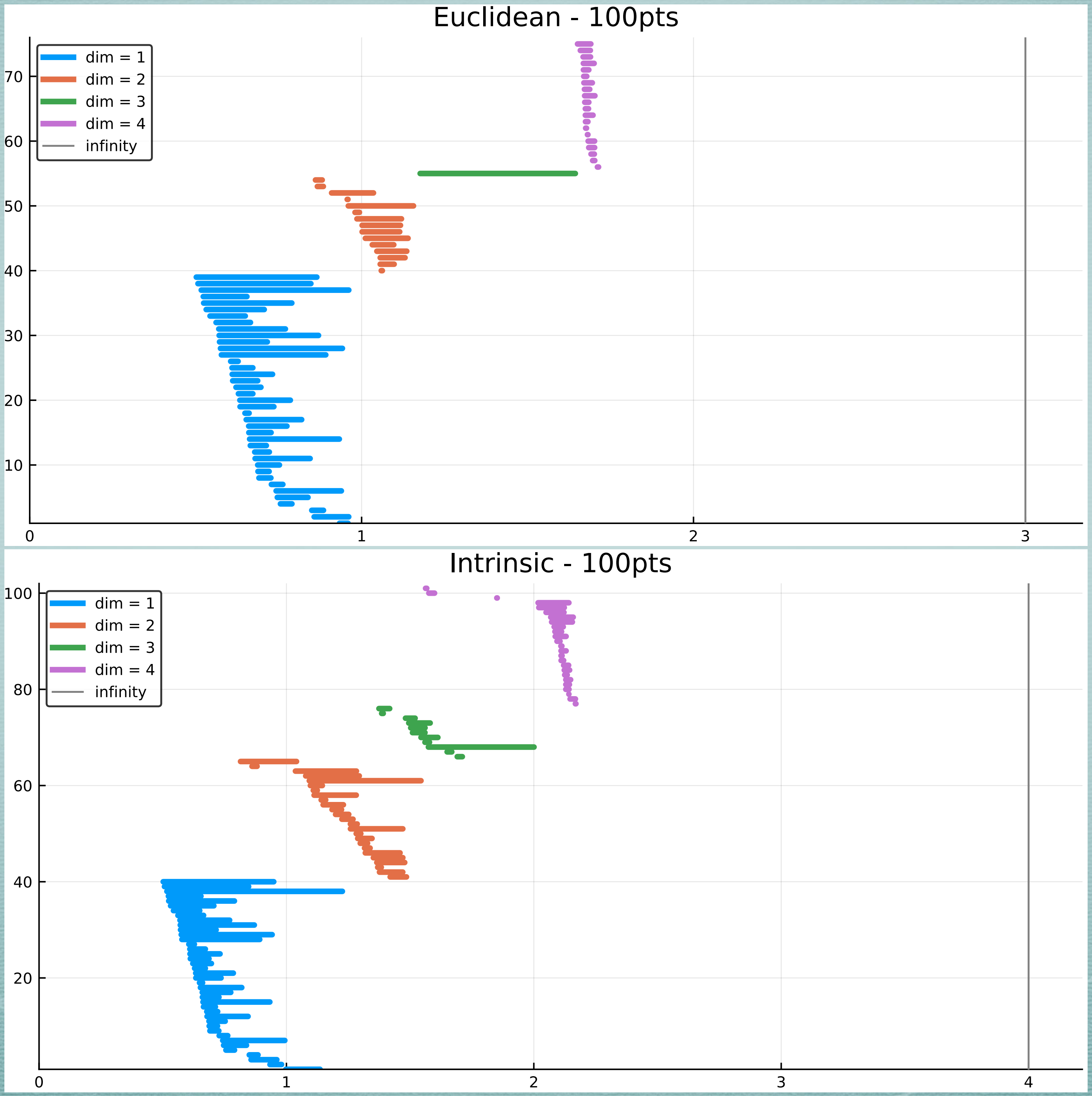


Examples

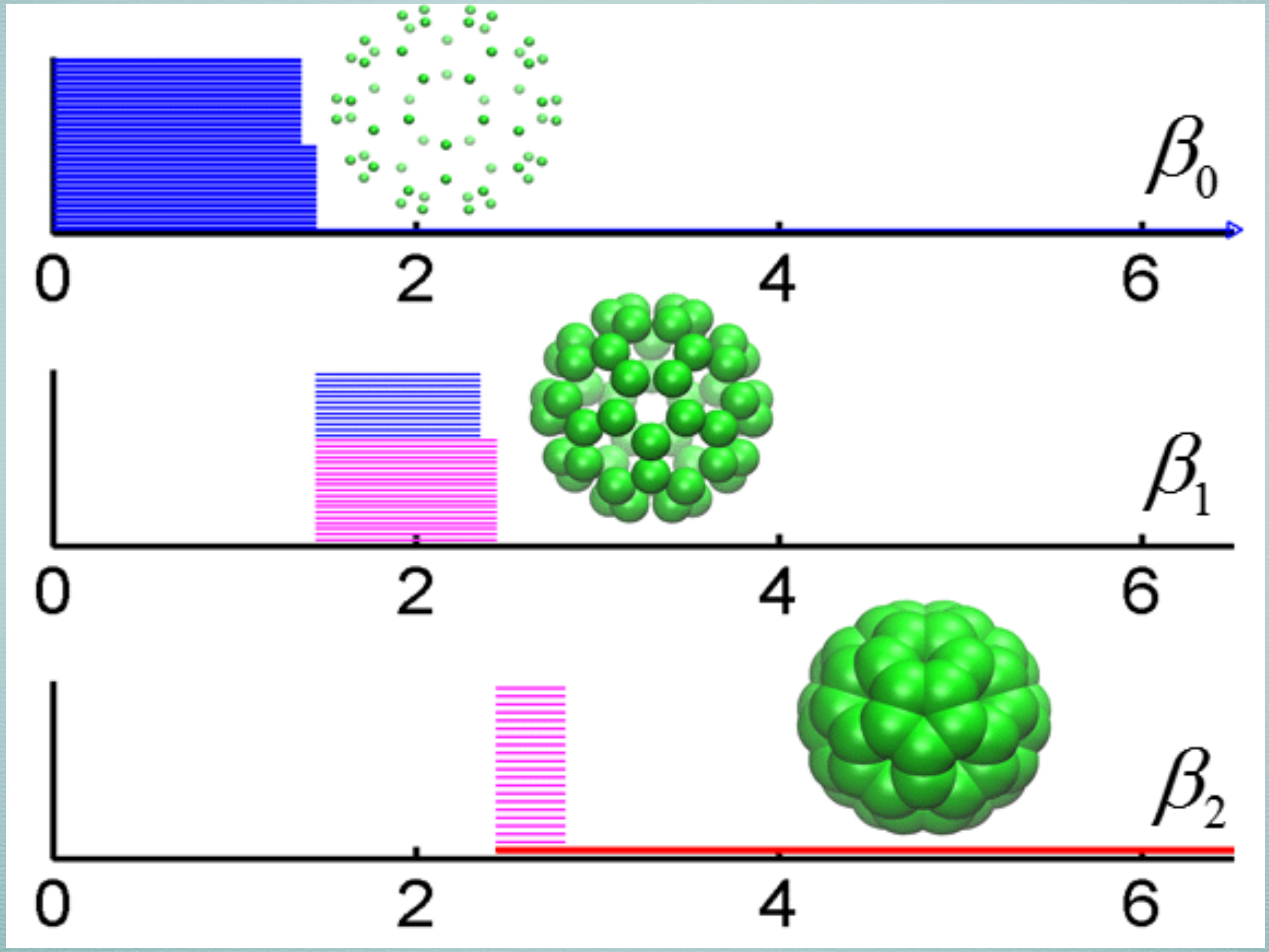
Examples of PH: S^2



Examples of PH: S^3



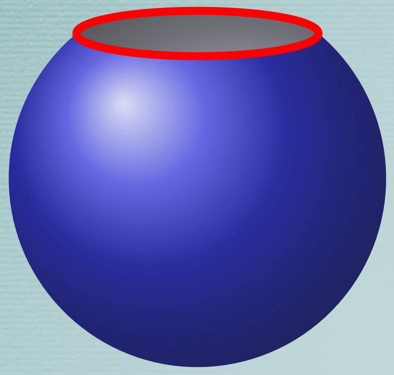
Persistence diagrams

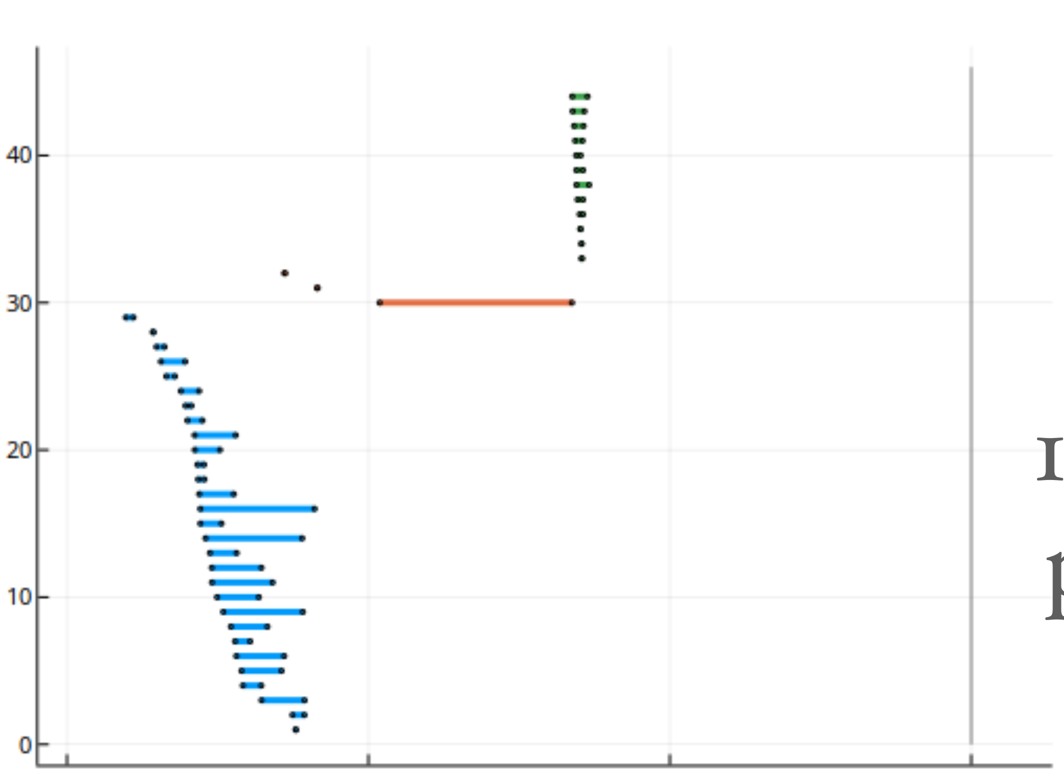


Fluoren C_{60}

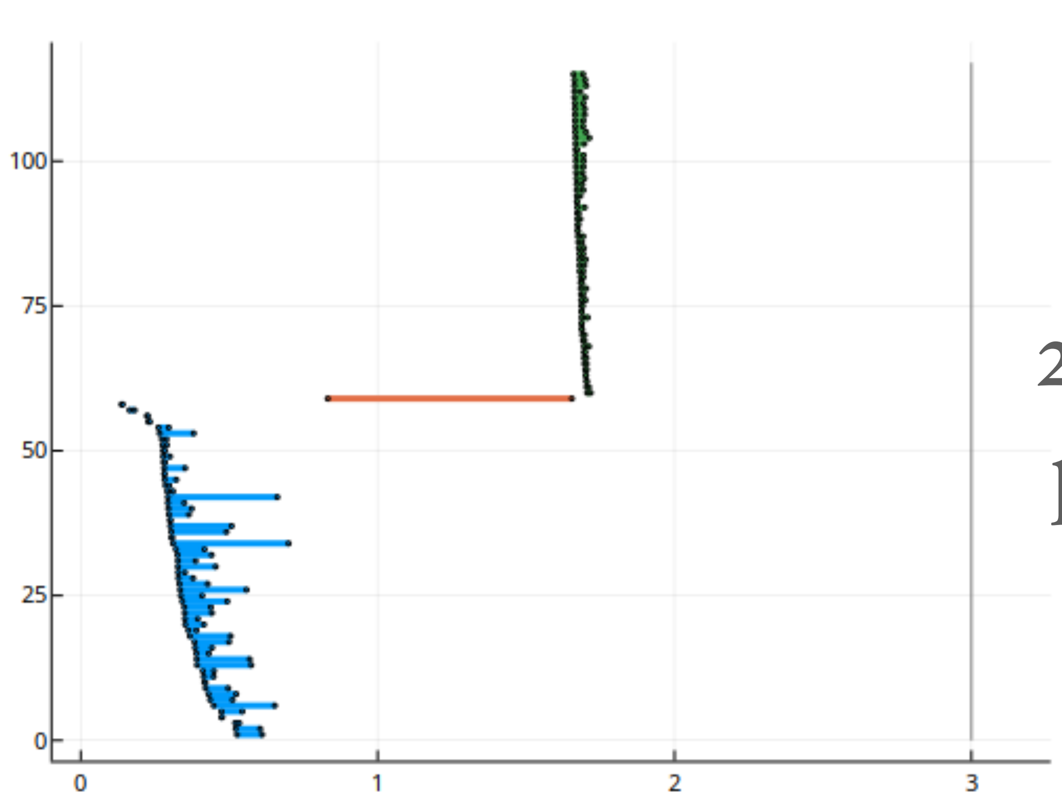
Xia, Wei: A review of geometric, topological and graph theory apparatuses for the modeling and analysis of biomolecular data

Effects of density

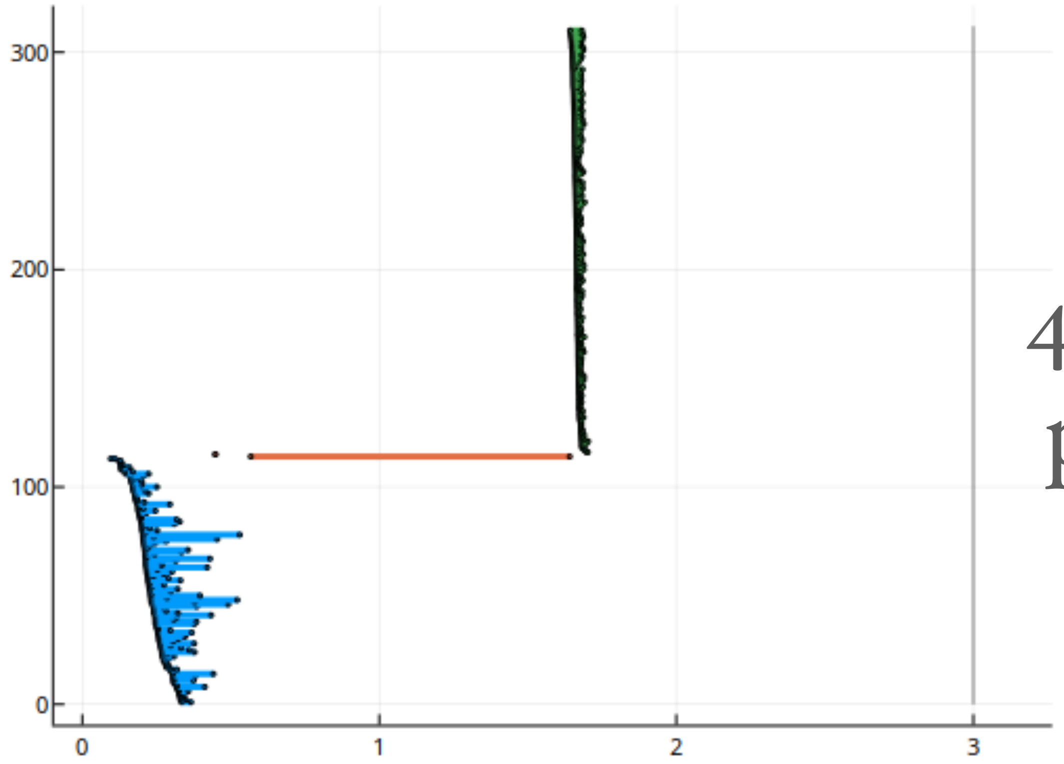




100 pts

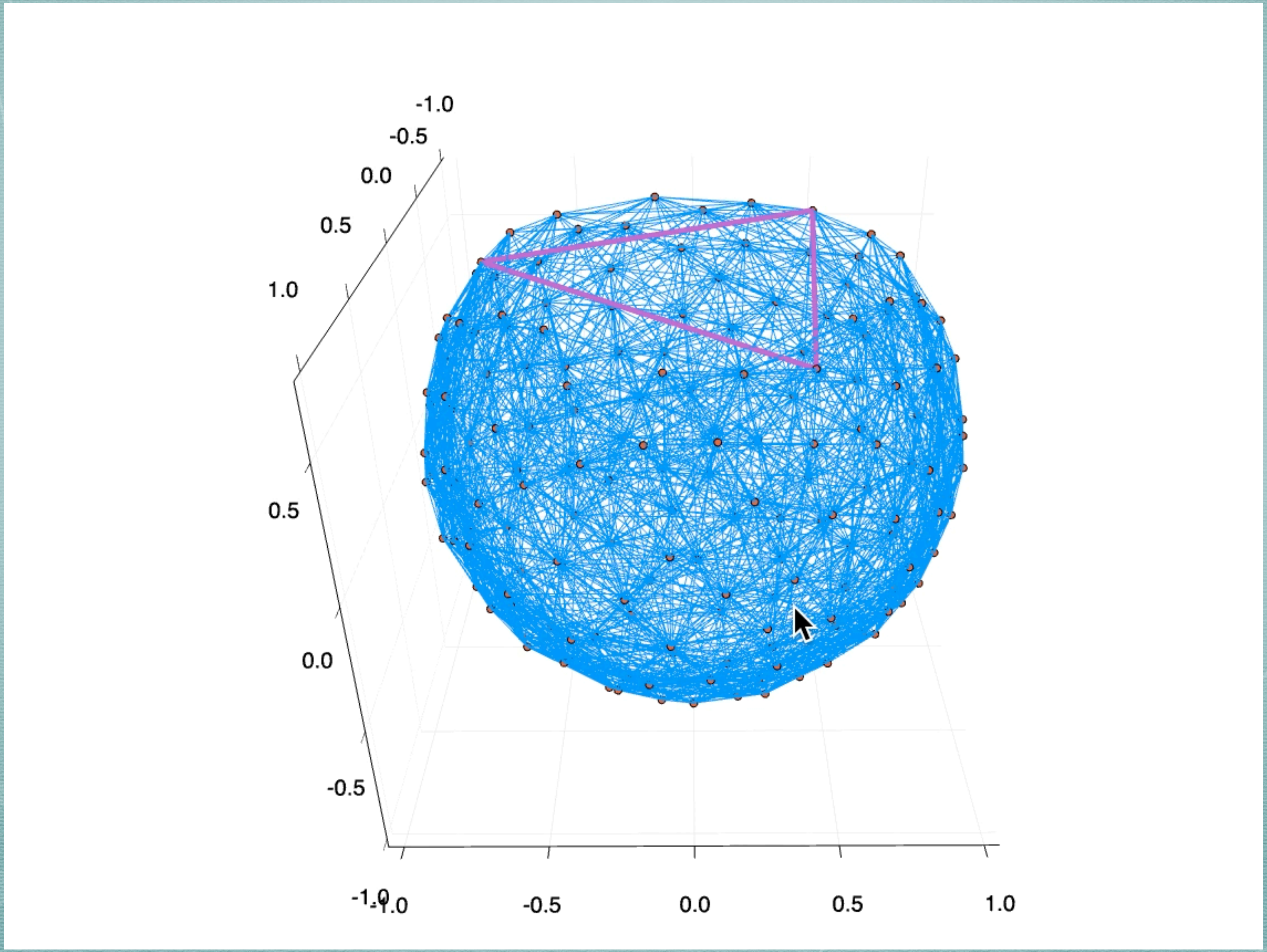


200 pts



400 pts

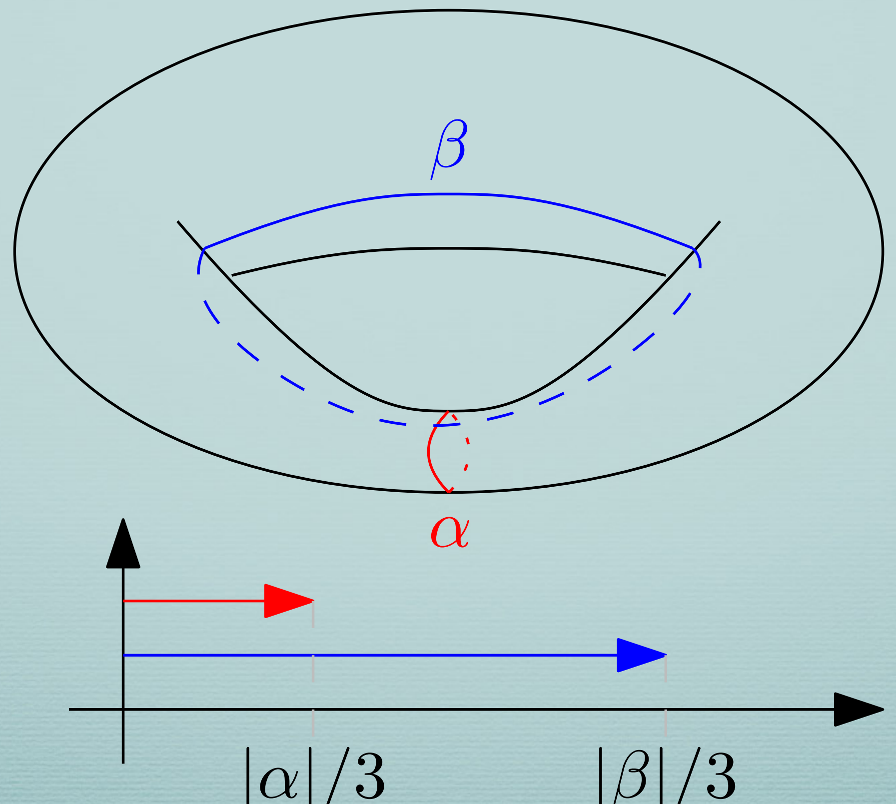
Critical simplex of a 2-dim class



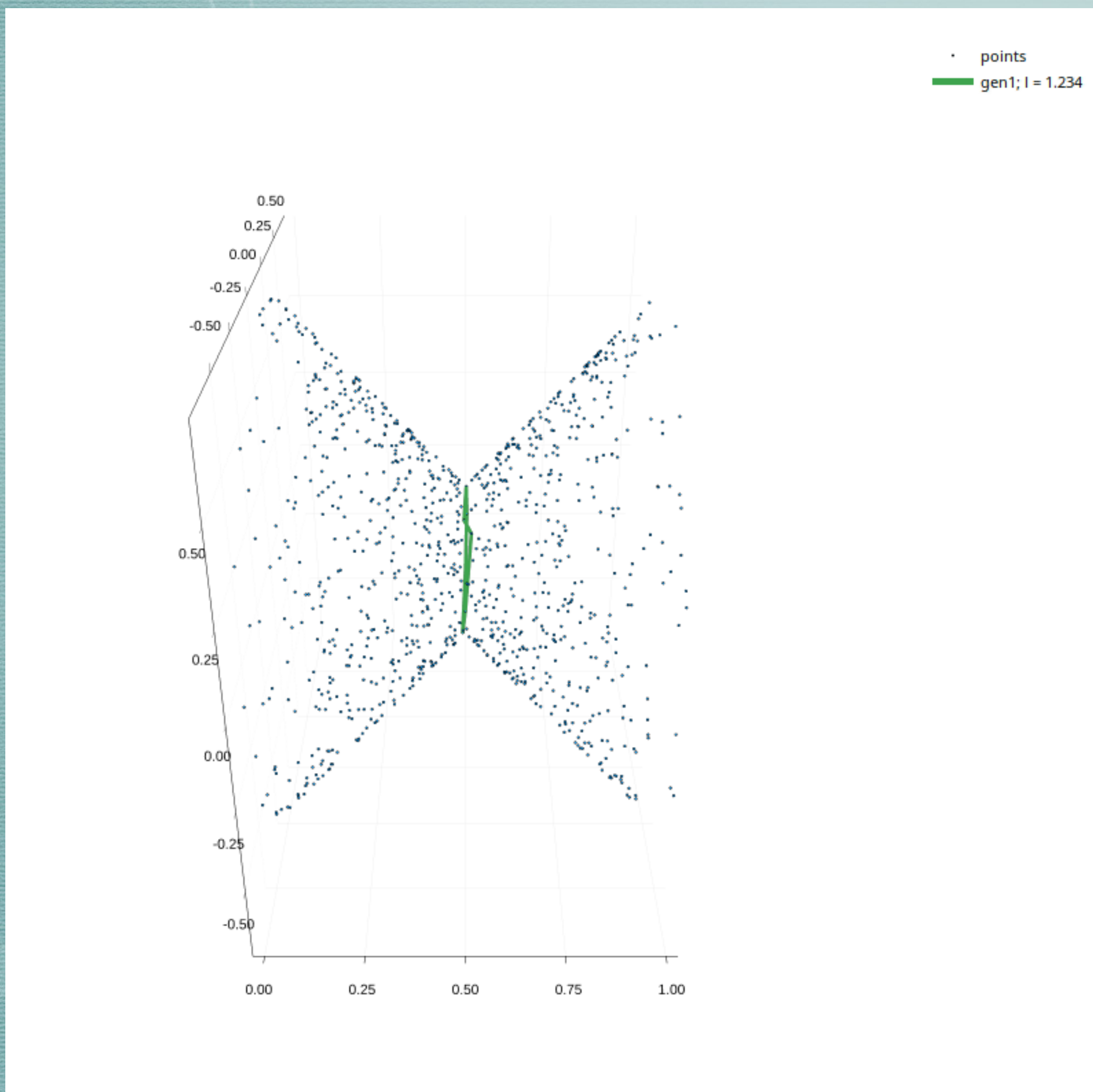
1-dimensional PH of geodesic spaces

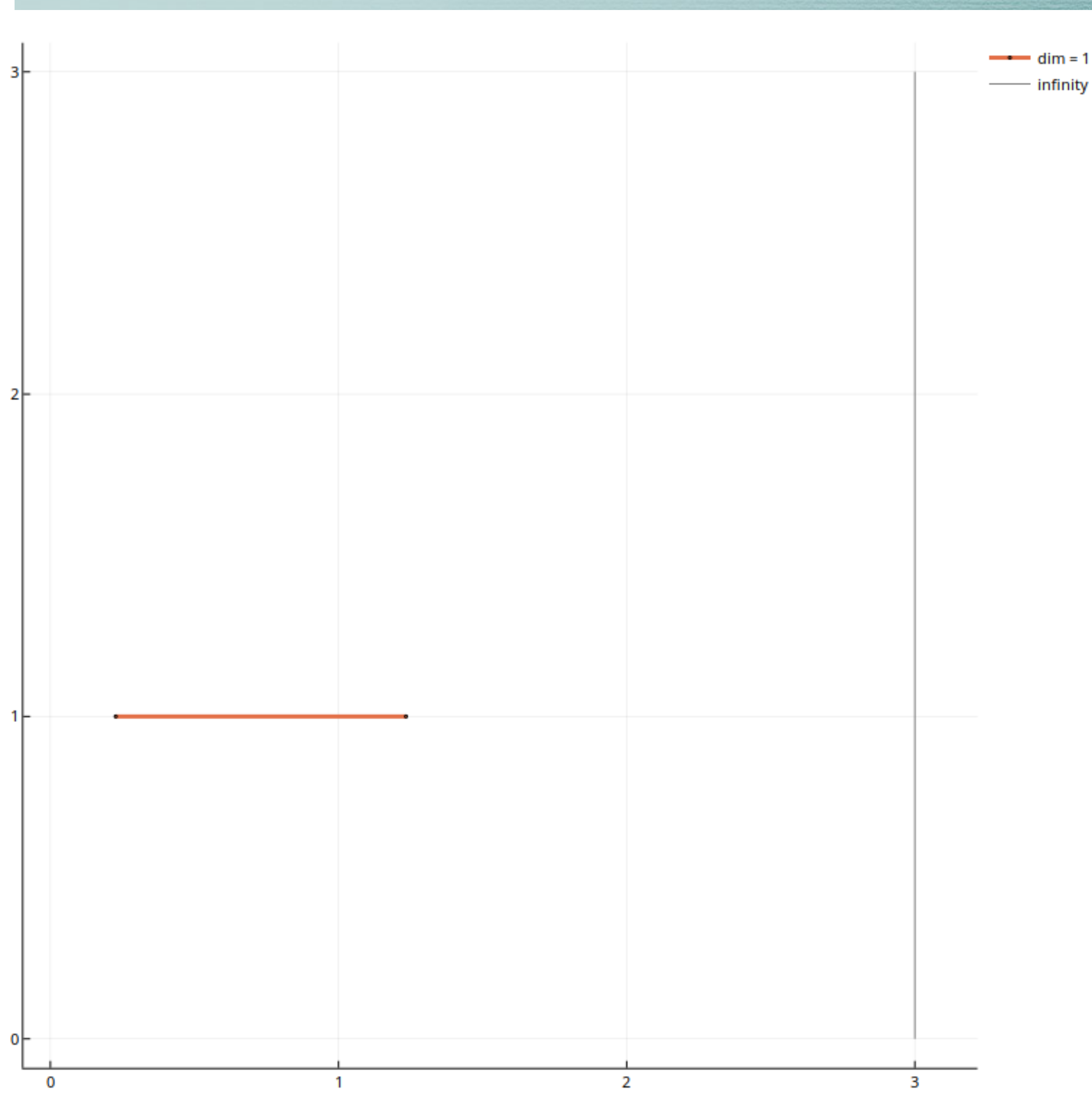
X a geodesic metric space. 1-dim PH looks as follows:

- For each element α of the shortest homology base we obtain a bar.
- In Rips filtration these bars terminate at $length(\alpha)/3$.

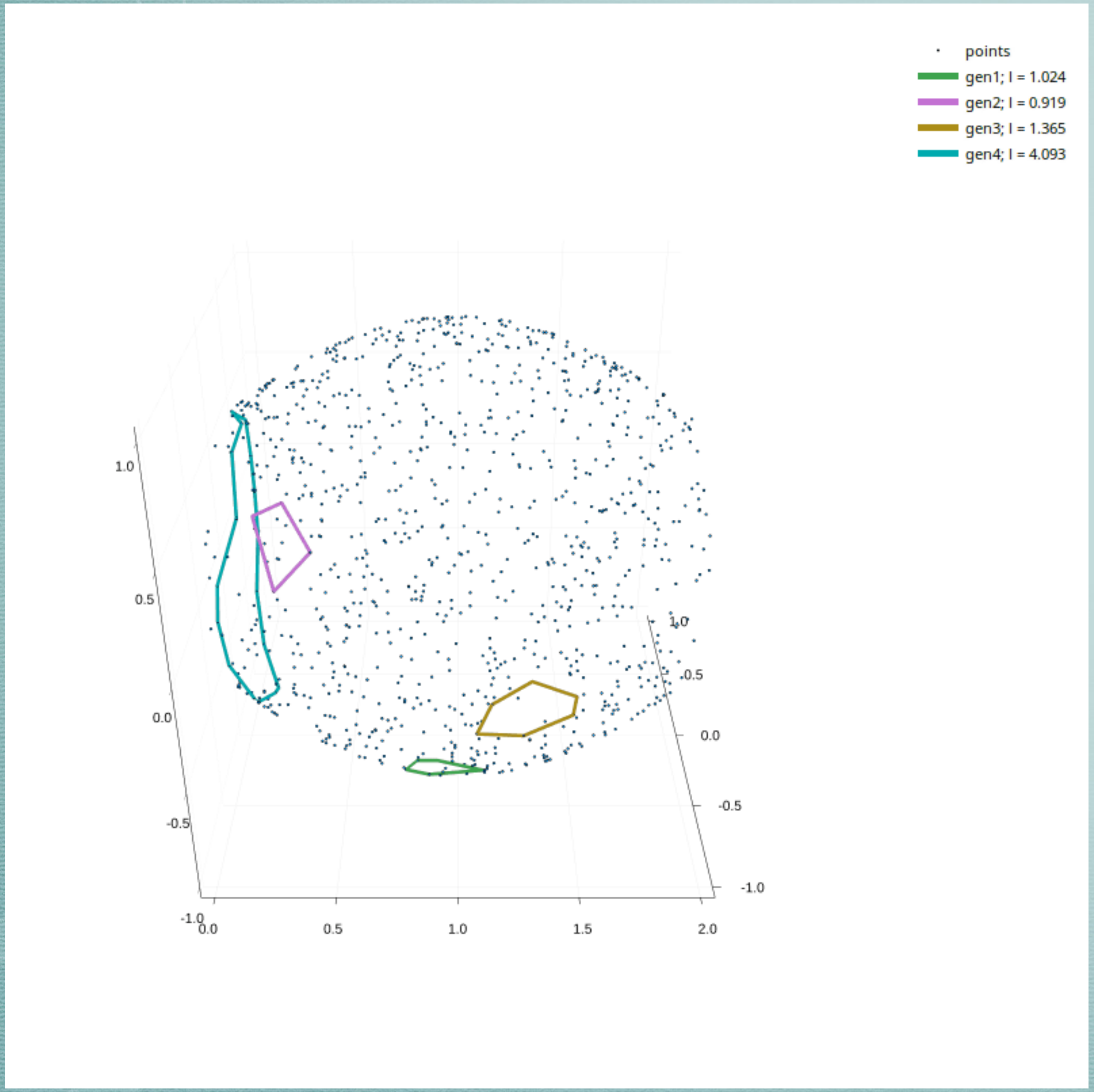


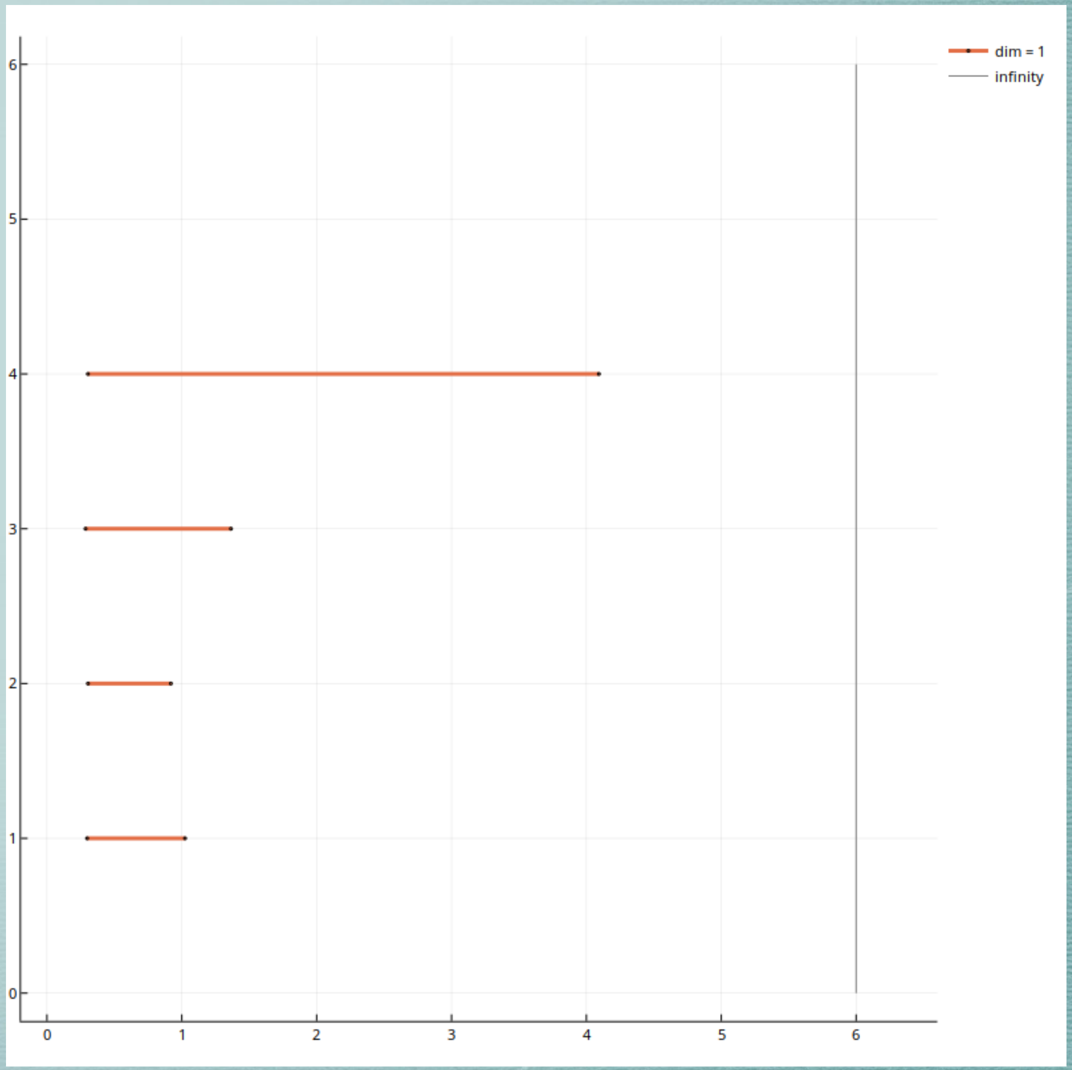
1-dimensional PH of geodesic spaces



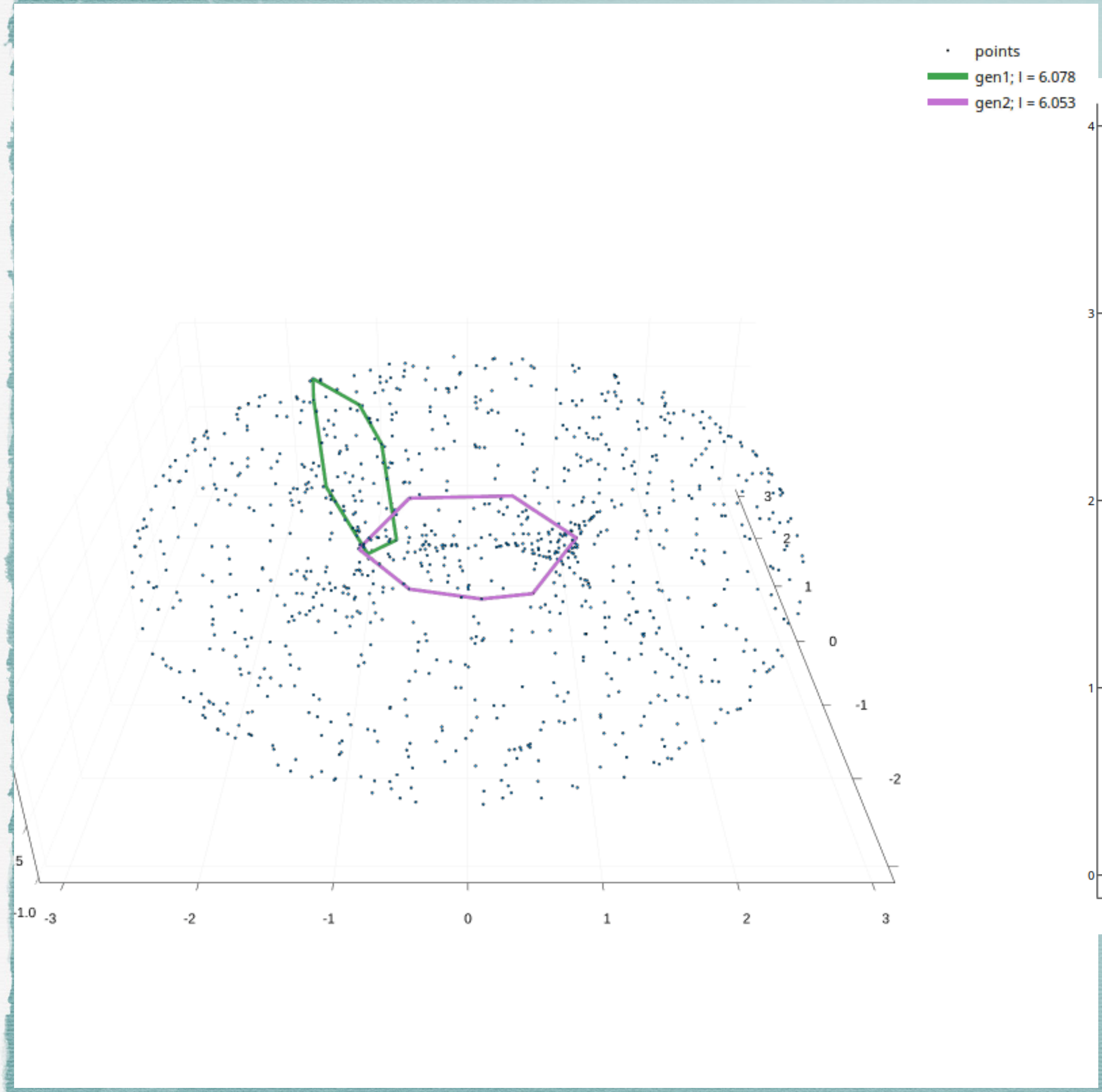


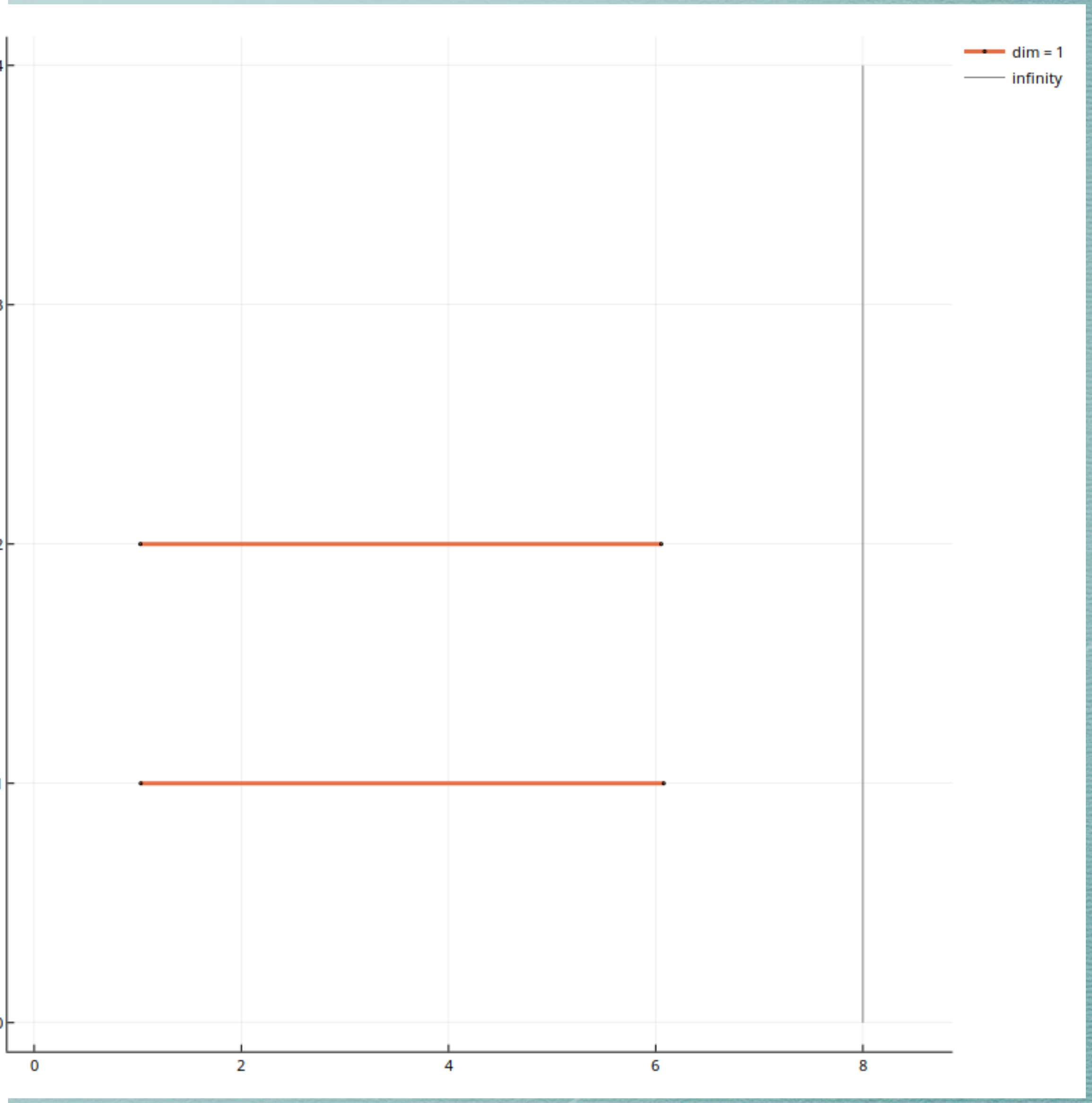
1-dimensional PH of geodesic spaces



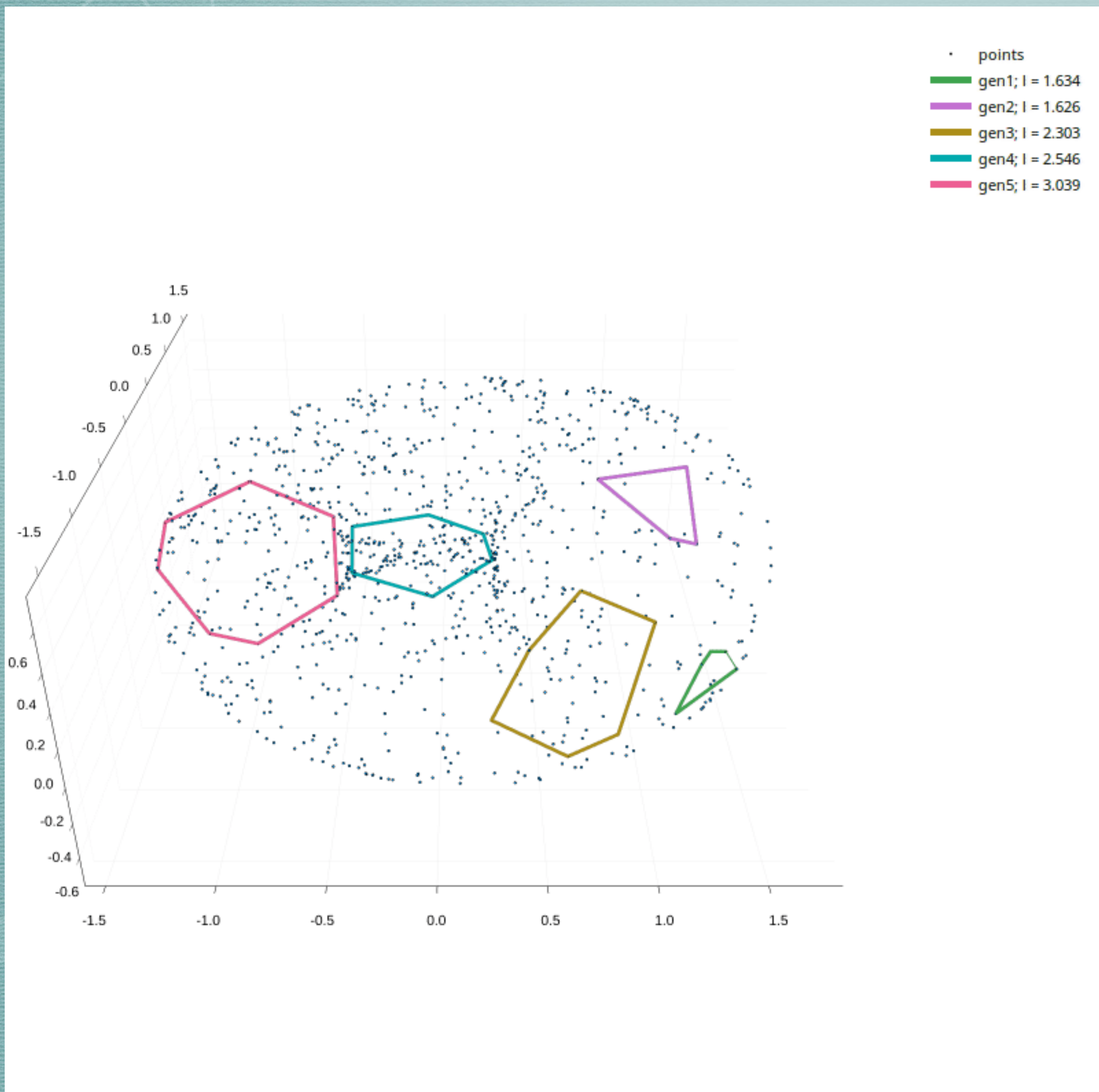


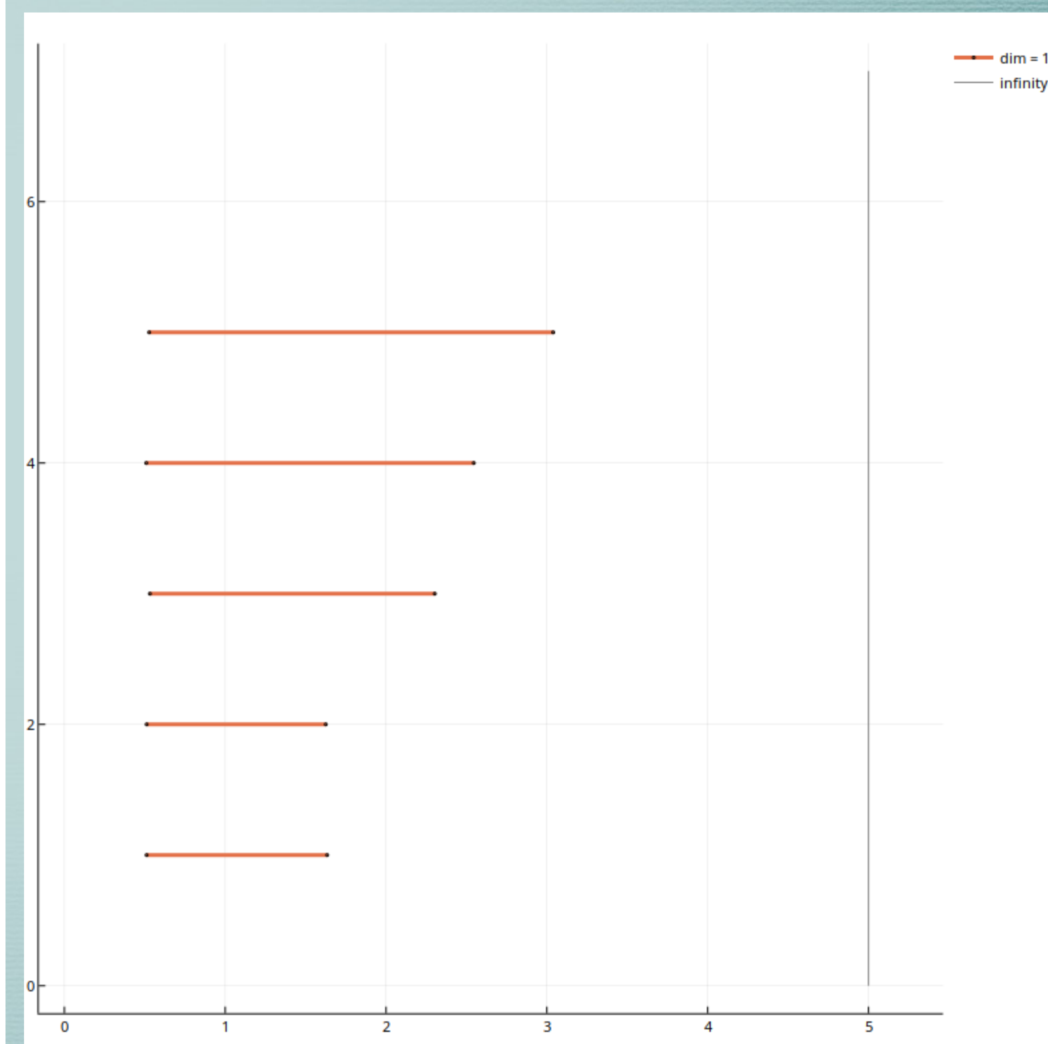
1-dimensional PH of geodesic spaces



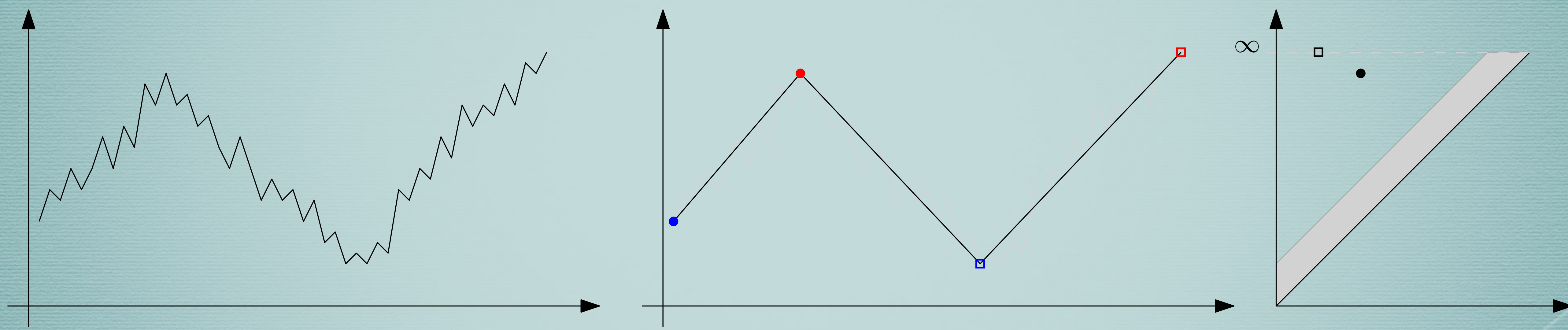


1-dimensional PH of geodesic spaces





Denoising a function



An application of sublevel 0-dimensional persistence.

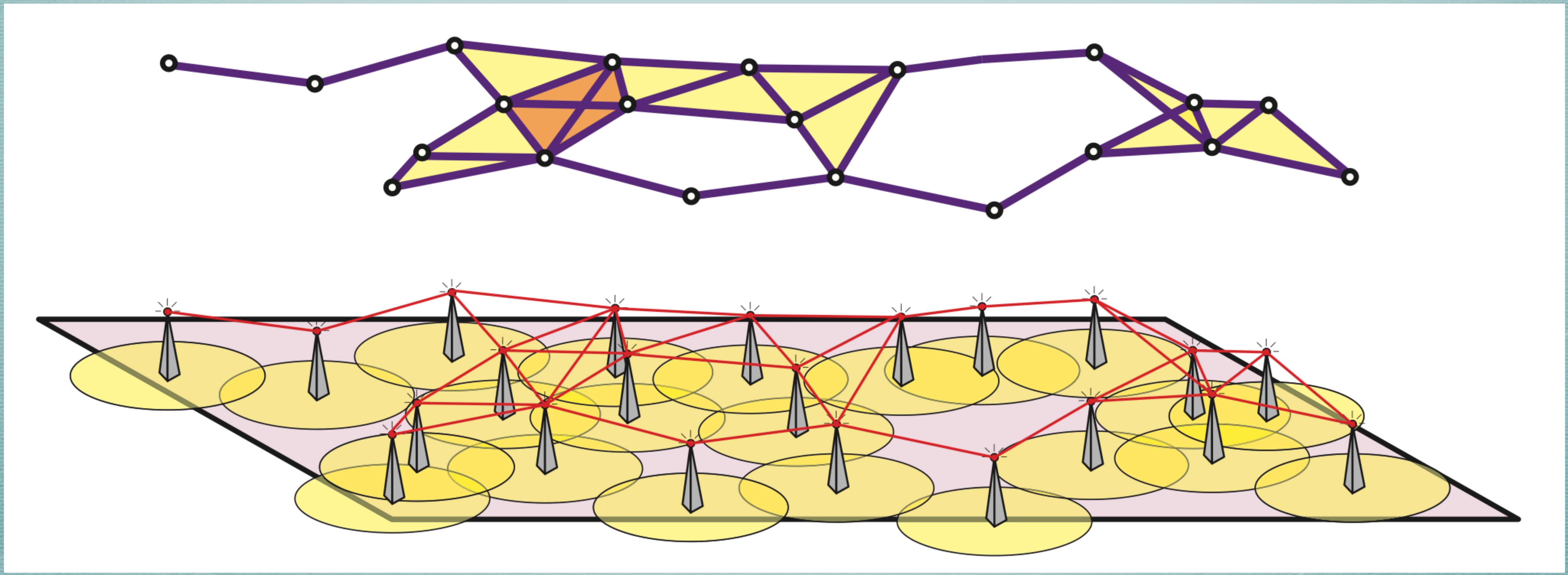
Applications to data

Sensor coverage

Coordinate free sensors.

Sensors only detect neighbors in broadcast radius r_B.

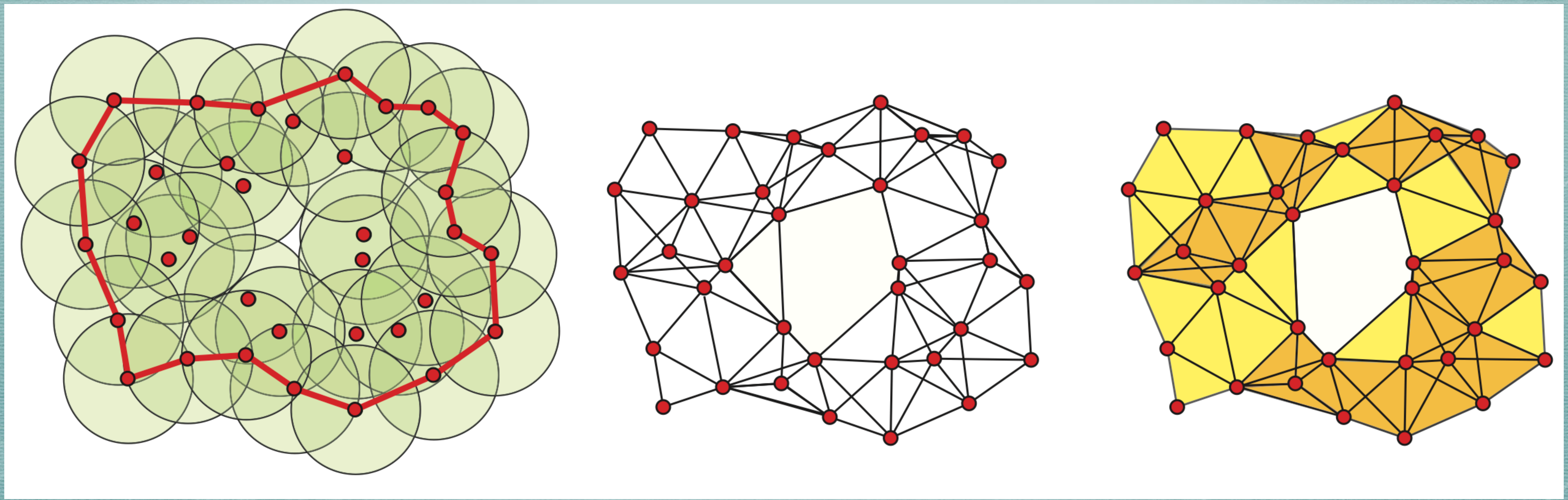
Sensors cover ball of covering radius $r_c > r_b/\sqrt{3}$.



V. De Silva and R. Ghrist, Homological sensor networks.

Sensor coverage

If sensors S cover the boundary (fence), then they cover the interior if $H_1(Rips(S, r)) = 0$.

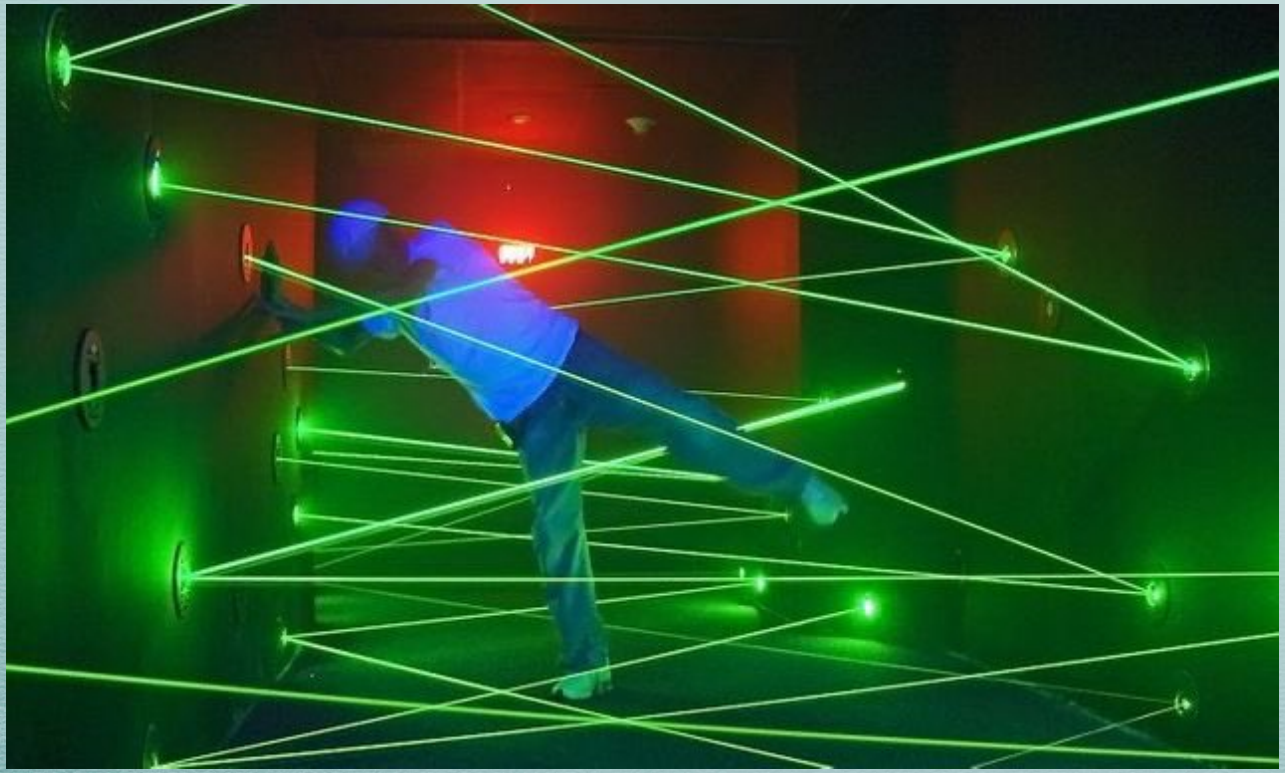


V. De Silva and R. Ghrist, Homological sensor networks.

Sensor coverage

(Persistent) homological criteria also exist for:

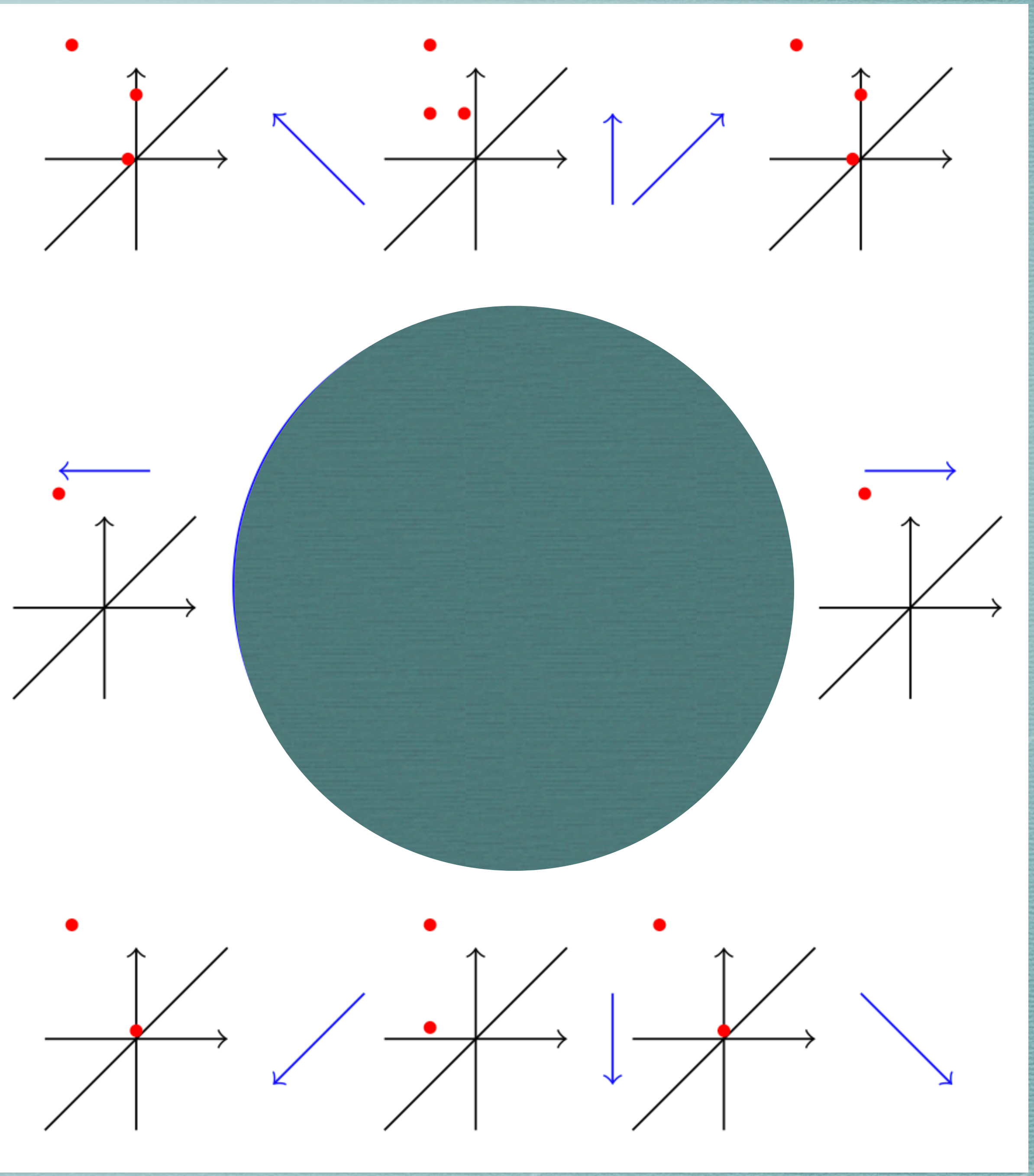
- Coverage with dynamic sensor network
- Intruder detection (weaker than coverage).



V. De Silva and R. Ghrist, Homological sensor networks.

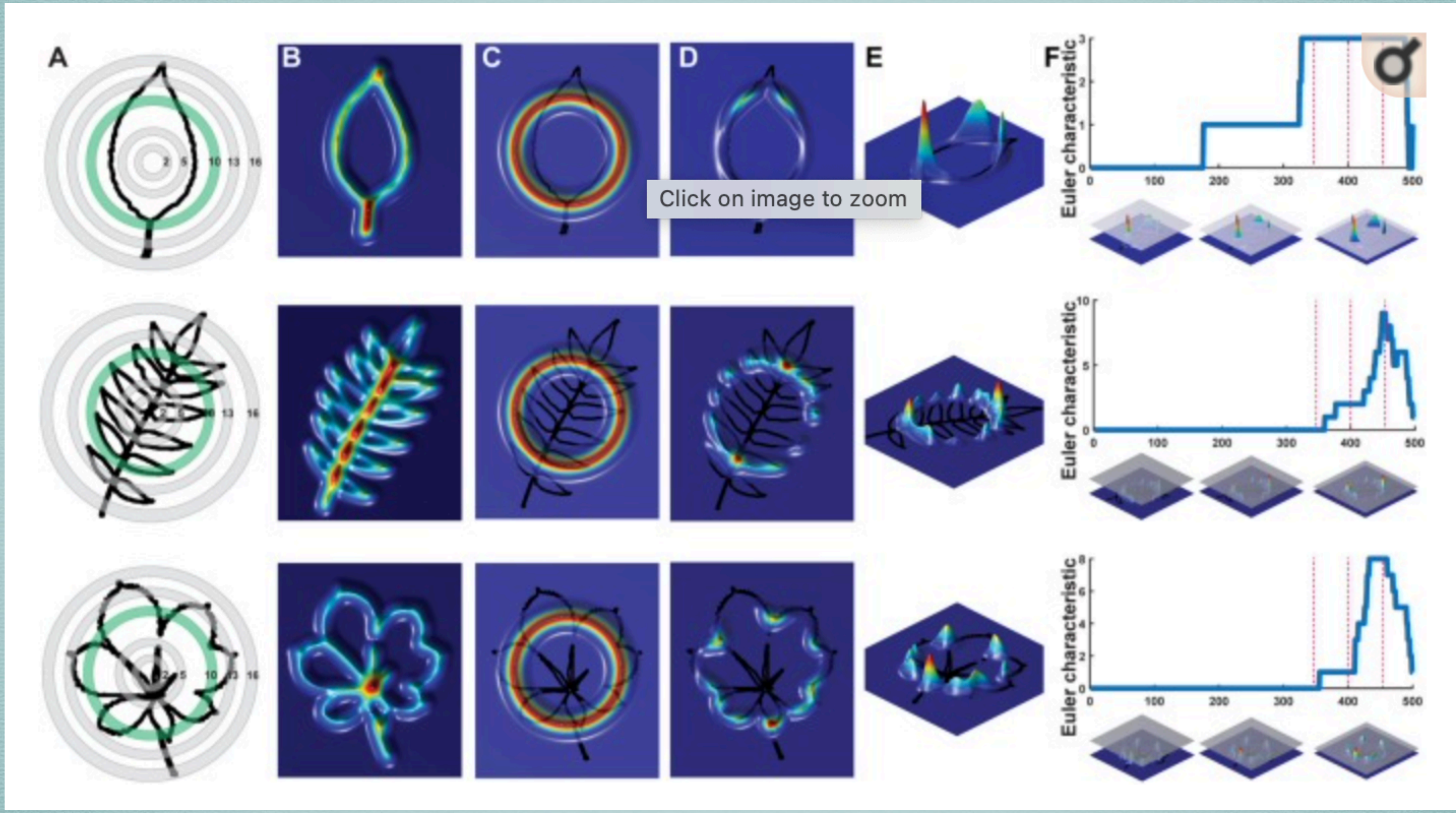
Directional scanning

- $K \subset \mathbb{R}^2$ be "nice", compact.
- "Scan" K along each direction.
- For each $v \in S^2$ obtain the sublevel diagram PD(K, v).
- These scans uniquely determine K .
- Map $v \mapsto PD(K, v)$ is called persistent transform.
- Euler transform maps a direction to Euler curve…it is also injective.



Katharine Turner, Sayan Mukherjee, Doug M Boyer: Persistent Homology Transform for Modeling Shapes and Surfaces

Classification of leaves



Li M, An H, Angelovici R, et al. Topological Data Analysis as a Morphometric Method: Using Persistent Homology to Demarcate a Leaf Morphospace.

Preprocessing of medical data

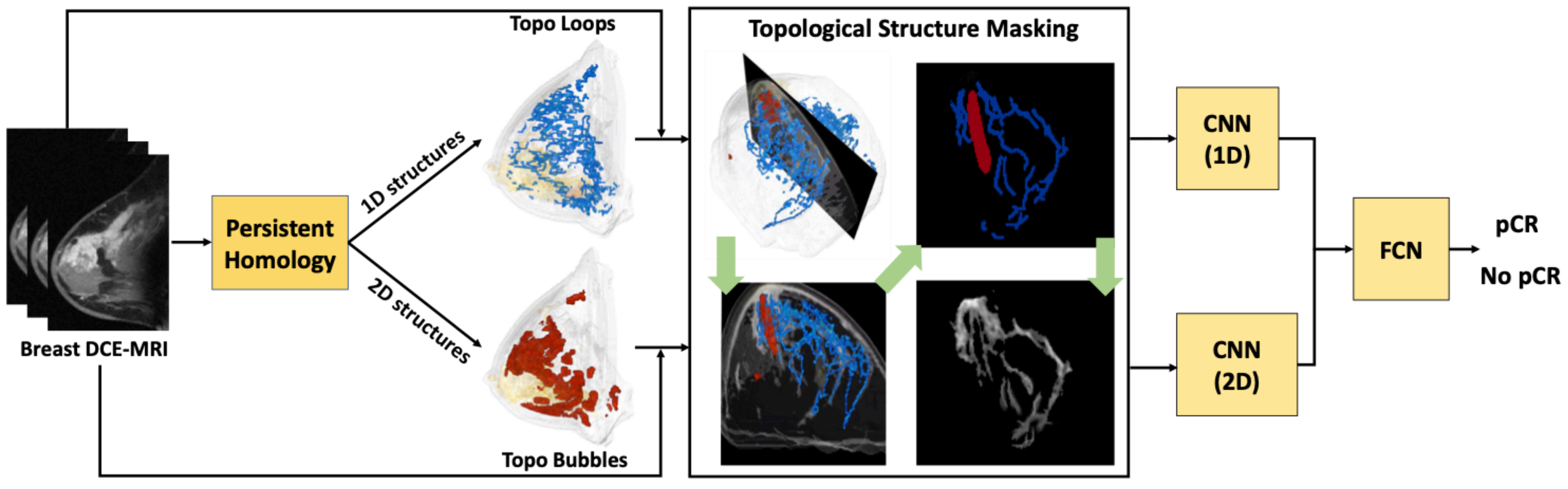
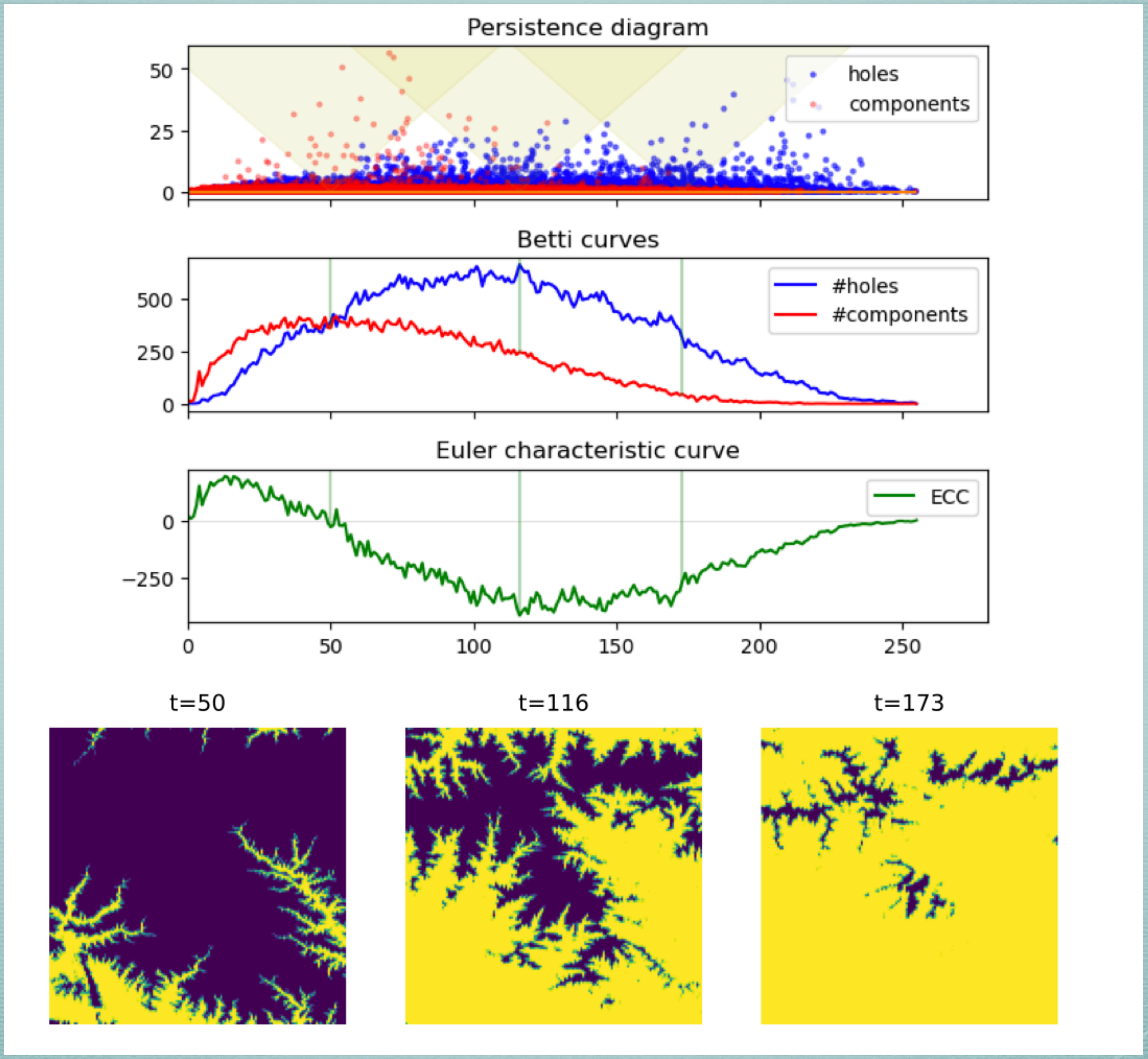


Fig. 2. Our proposed TopoTxR pipeline. We extract 1D and 2D topological structures from breast MRI with persistent homology. We create new images in which voxels belonging to these structures have their intensity value from the original breast MRI, and zero otherwise. We create two images corresponding to 1D and 2D topological structures. We use two 3D CNNs and a fully connected network for pCR prediction.

Fan Wang, Huidong Liu, Dimitris Samaras, Chao Chen: A topological biomarker for predicting treatment response in breast cancer

Derived curves for medical data



Fan Wang, Hubert Wagner, Chao Chen: GPU Computation of the Euler Characteristic Curve for Imaging Data

Topology of cosmic web

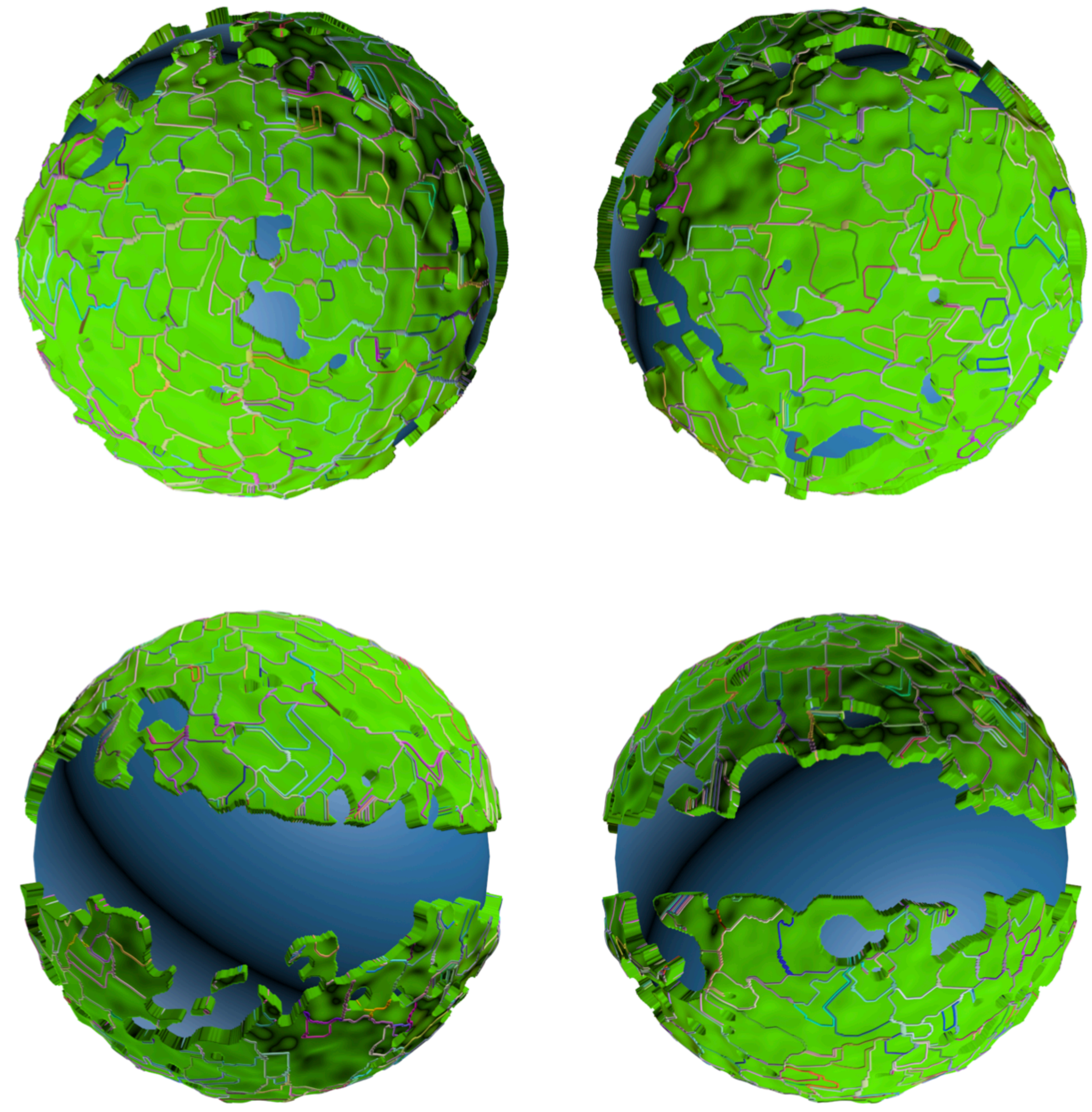
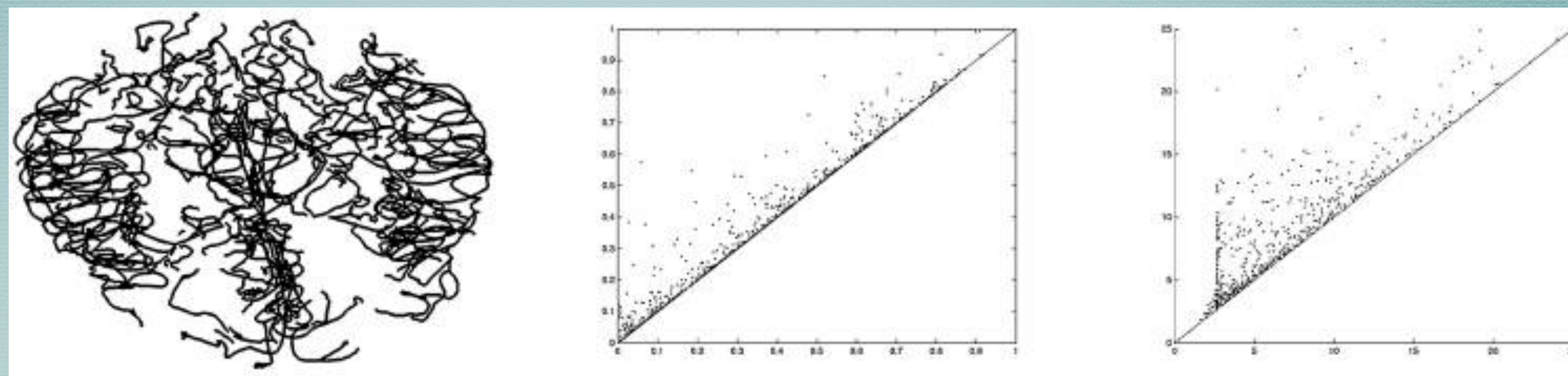


Fig. 12. Visualization of the loops for the largest excursion set, which consists of the entire sphere minus the mask. To improve the visualization, the temperature field has been smoothed by a small amount, and we do not draw very short loops. *From left to right*: the sphere from the top, the bottom, the left, and the right views.

Pranav et al: Unexpected topology of the temperature fluctuations in the cosmic microwave background

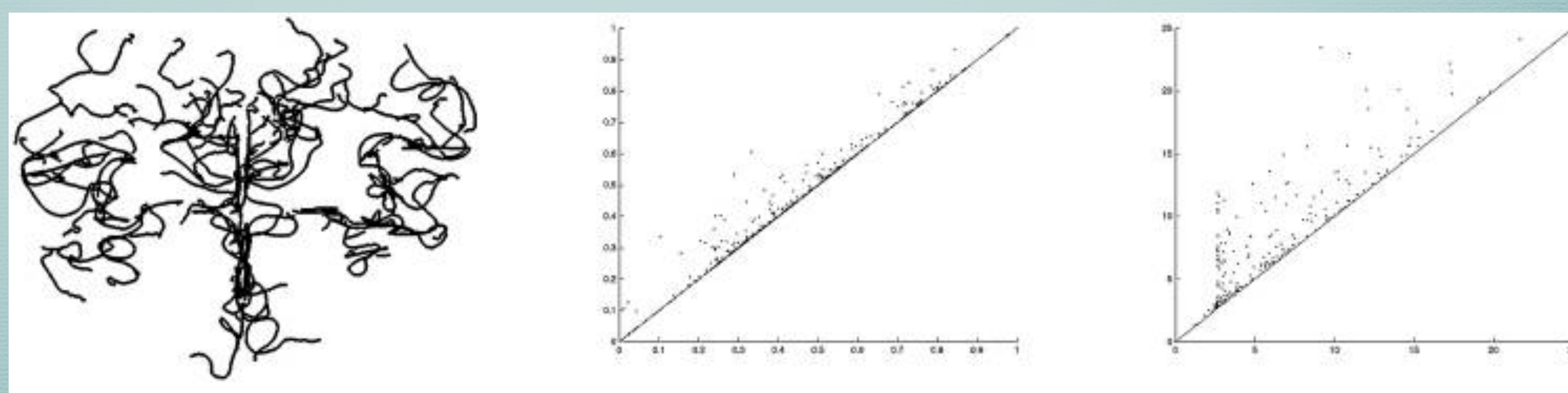
Topology of artery trees

Persistent homology data objects from a 24-year-old.



(a) Brain tree (b) Dgm_0 (c) Dgm_1

Persistent homology data objects from a 68-year-old.



(a) Brain tree (b) Dgm_0 (c) Dgm_1

"Novel approaches to the statistical analysis, through various summaries of the persistence diagrams, lead to heightened correlations with covariates such as age and sex, relative to earlier analyses of this data set."

Bendich et al: Persistent Homology Analysis of Brain Artery Trees

Persistence in material science

Silica (a mineral in earth) simulated

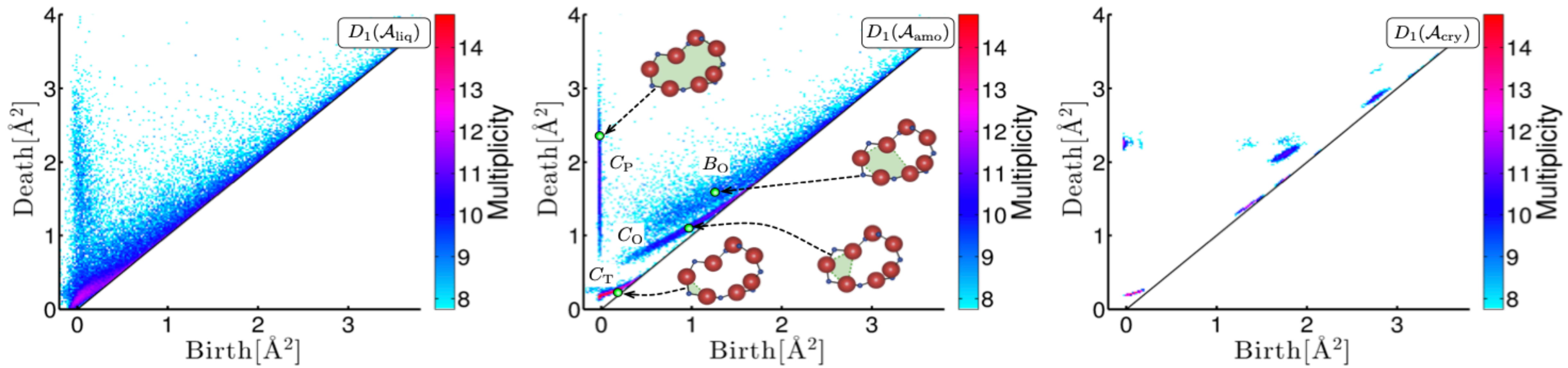


Fig. 2. PDs of the liquid (left), amorphous (middle), and crystalline (right) states with the multiplicity on the logarithmic scale. In the amorphous state, the three characteristic curves and one band region are labeled $C_{\mathrm{P}}, C_{\mathrm{T}}, C_{\mathrm{O}}$, and B_{O}, respectively. The insets in $D_1(\mathcal{A}_{\mathrm{amo}})$ show rings in the hierarchical relationship, where the red and blue spheres represent oxygen and silicon atoms, respectively.

Yasuaki Hiraoka, Takenobu Nakamura, Akihiko Hirata, Emerson G. Escolar, Kaname Matsue, Yasumasa Nishiura: Hierarchical structures of amorphous solids characterized by persistent homology

Persistence in material science

Silica (a mineral in earth) simulated

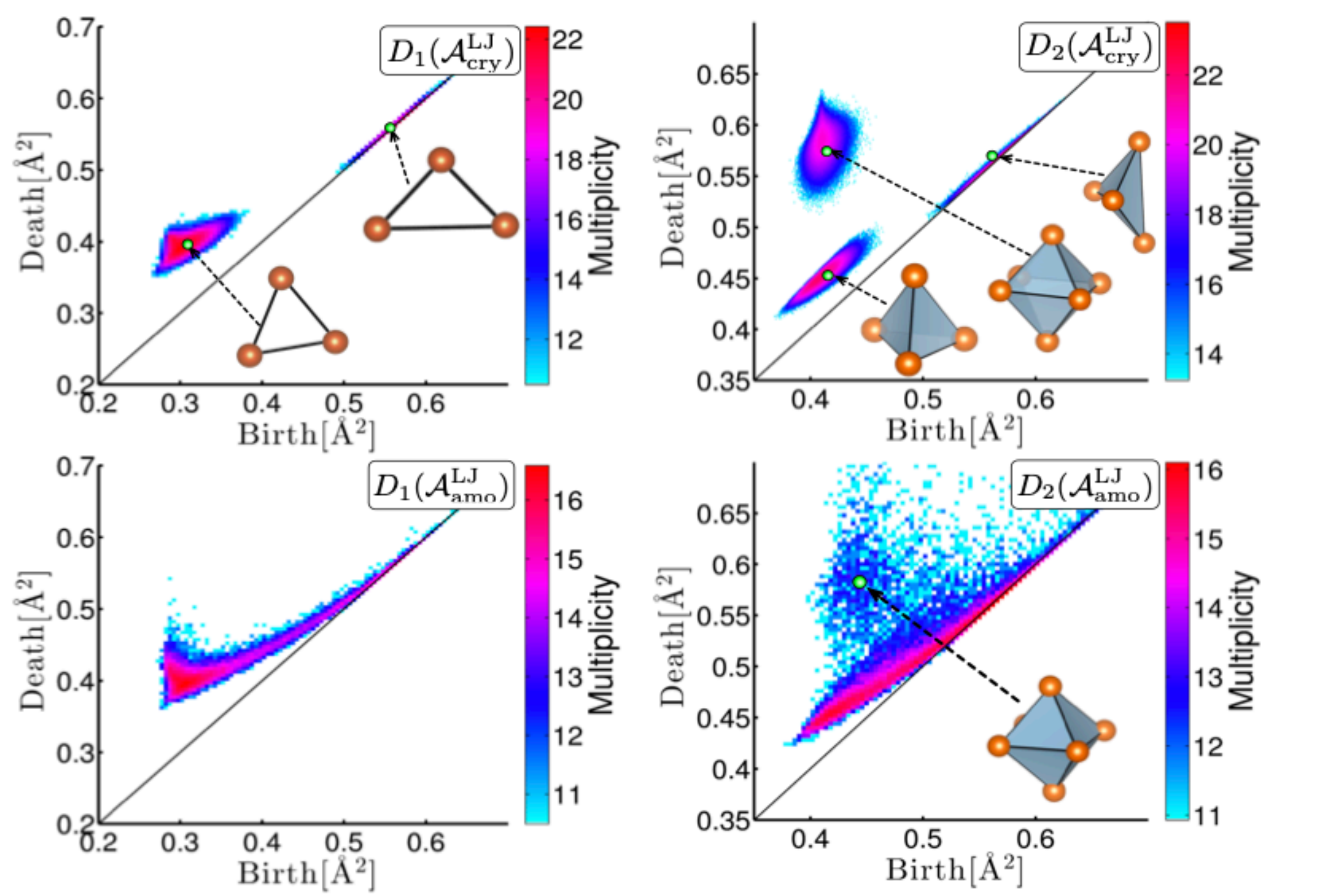
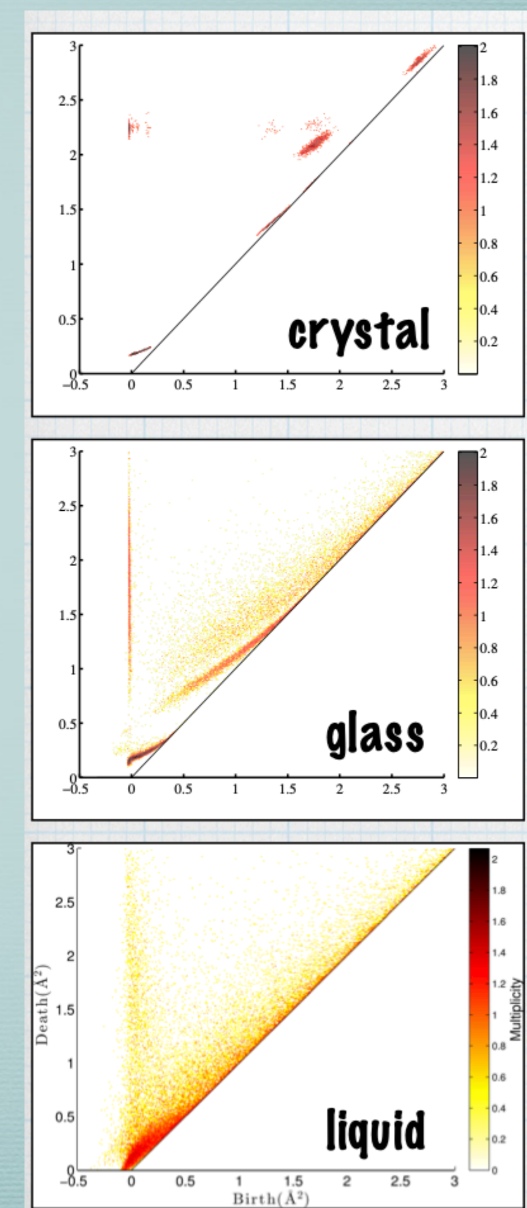


Fig. 6. PDs for the Lennard-Jones system with the multiplicity on the logarithmic scale. The left and right panels correspond to D_1 and D_2 respectively. Top panels correspond to FCC crystal at $T = 0.1$ and bottom panels correspond to amorphous state at $T = 10^{-3}$.

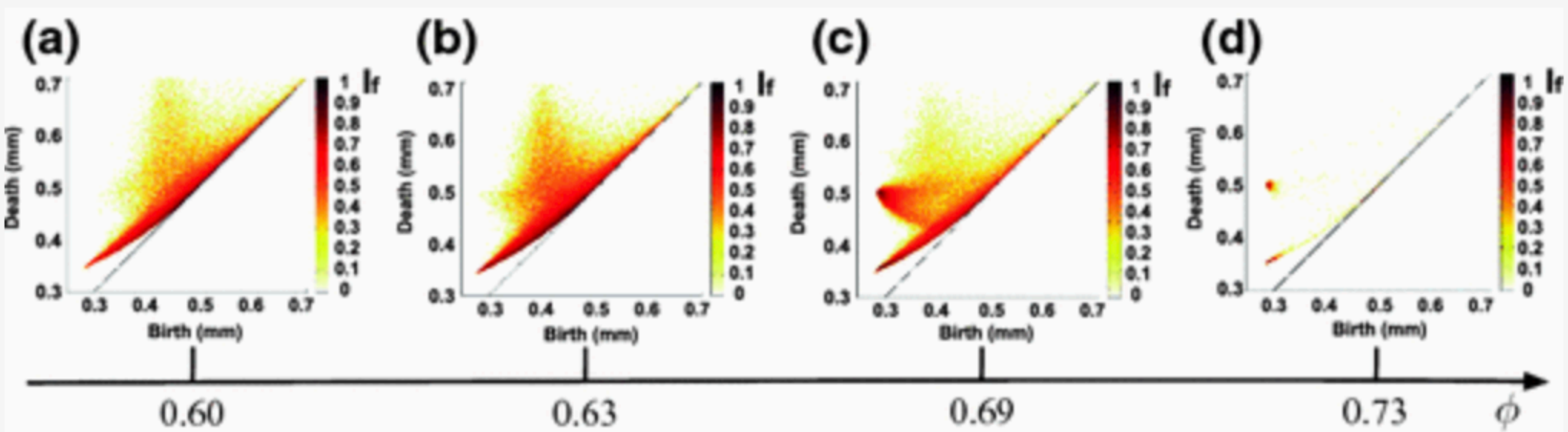


1-dim PD of sample obtained from MDI simulation (on right)

Yasuaki Hiraoka, Takenobu Nakamura, Akihiko Hirata, Emerson G. Escolar, Kaname Matsue, Yasumasa Nishiura: Hierarchical structures of amorphous solids characterized by persistent homology

Persistence in material science

"crystallization mechanism of three-dimensional granular packings of frictional spheres is studied at the grain-scale using Xray tomography...three-dimensional images of granular packings with several packing ratio are obtained by using XCT"



2*D*-diagrams

Fig. 5.14

Persistence diagrams of grain configurations for different packing ratios (Reproduced from [5])

M. Saadatfar, H. Takeuchi, V. Robins, N. Francois, Y. Hiraoka, Pore configuration landscape of granular crystallization

Persistence in material science

"crystallization mechanism of three-dimensional granular packings of frictional spheres is studied at the grain-scale using Xray tomography...three-dimensional images of granular packings with several packing ratio are obtained by using XCT"

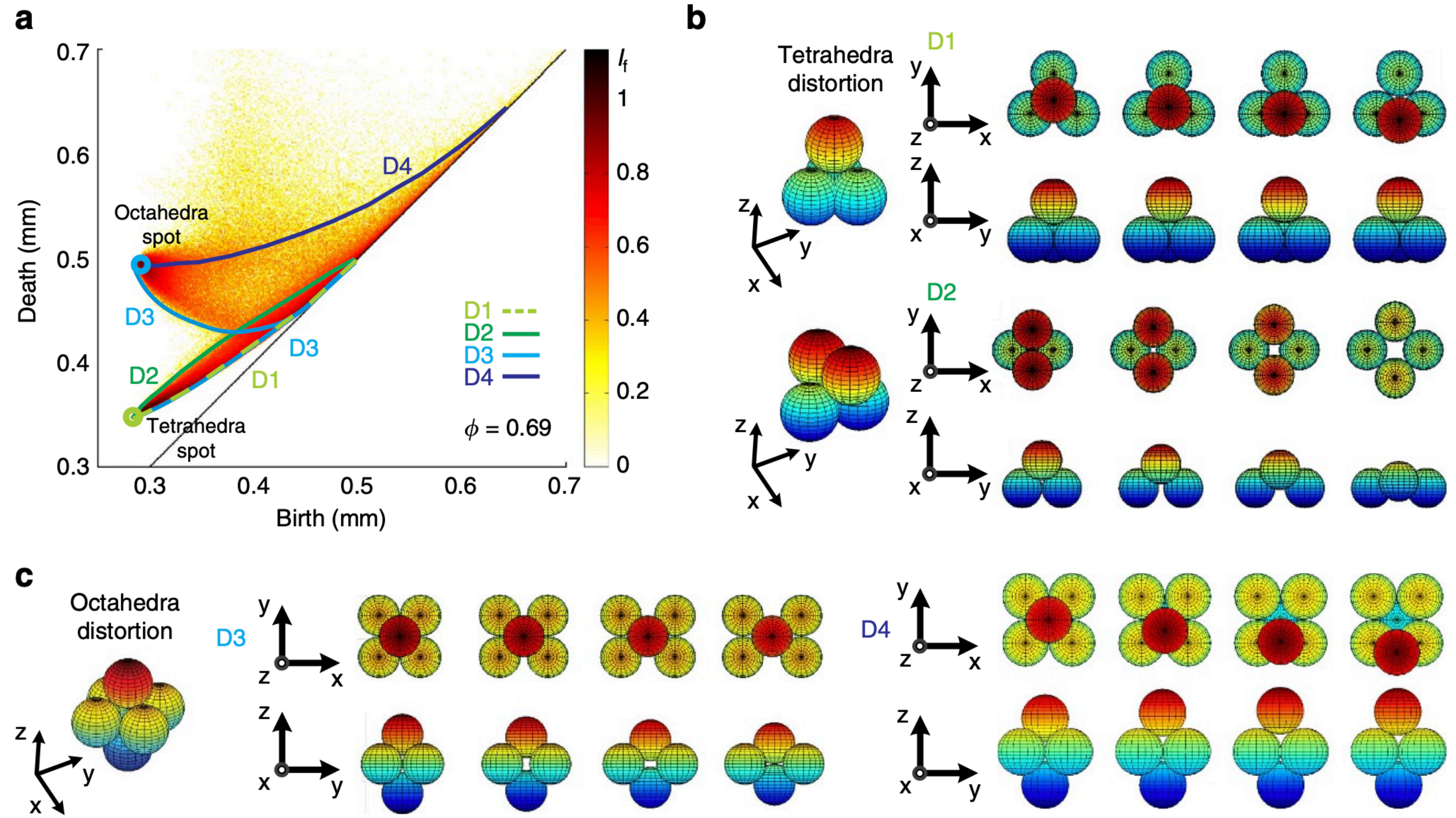


Figure 6 | Grain-scale tetrahedral and octahedral formation/deformation scenarios. (**a**) PD_2 of a partially crystallized sphere packing with density $\phi = 0.685$. The superimposed curves correspond to analytically computed birth-death curves of the deformation scenarios shown in panels (**b**,**c**). (**b**) Top and side views of D1 and D2 deformations scenarios of a tetrahedral cavity. (**c**) Top and side views of D3 and D4 deformation scenarios of an octahedral cavity. The colour code indicates the relative height of sections of the grain with respect to the horizontal median plane.

M. Saadatfar, H. Takeuchi, V. Robins, N. Francois, Y. Hiraoka, Pore configuration landscape of granular crystallization

Connecting with statistics

Into a vector space

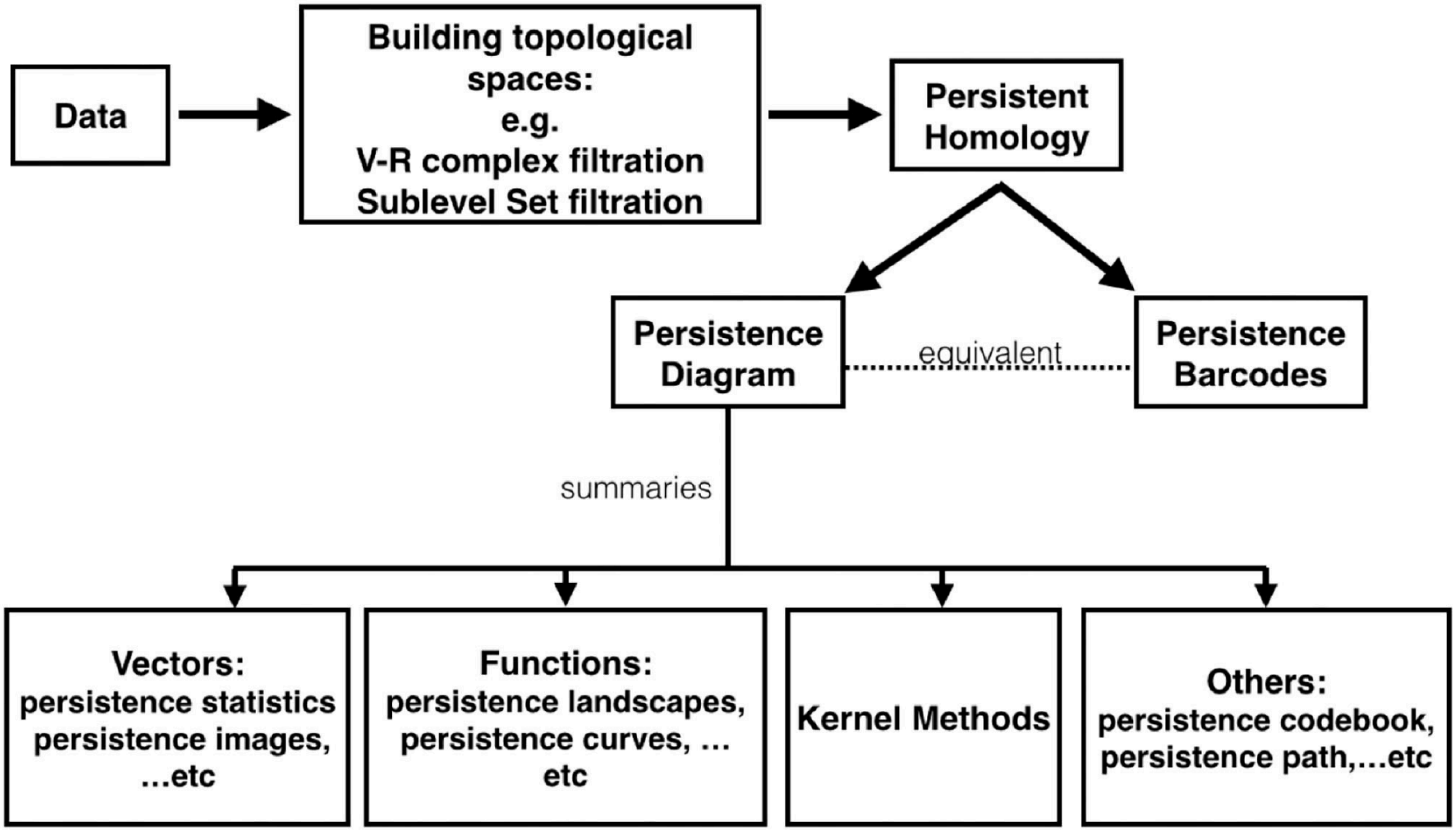


FIGURE 5 | A chart depicting relationship among existing TDA tools.

A Persistent Homology Approach to Heart Rate Variability Analysis With an Application to Sleep-Wake Classification

Topological signature

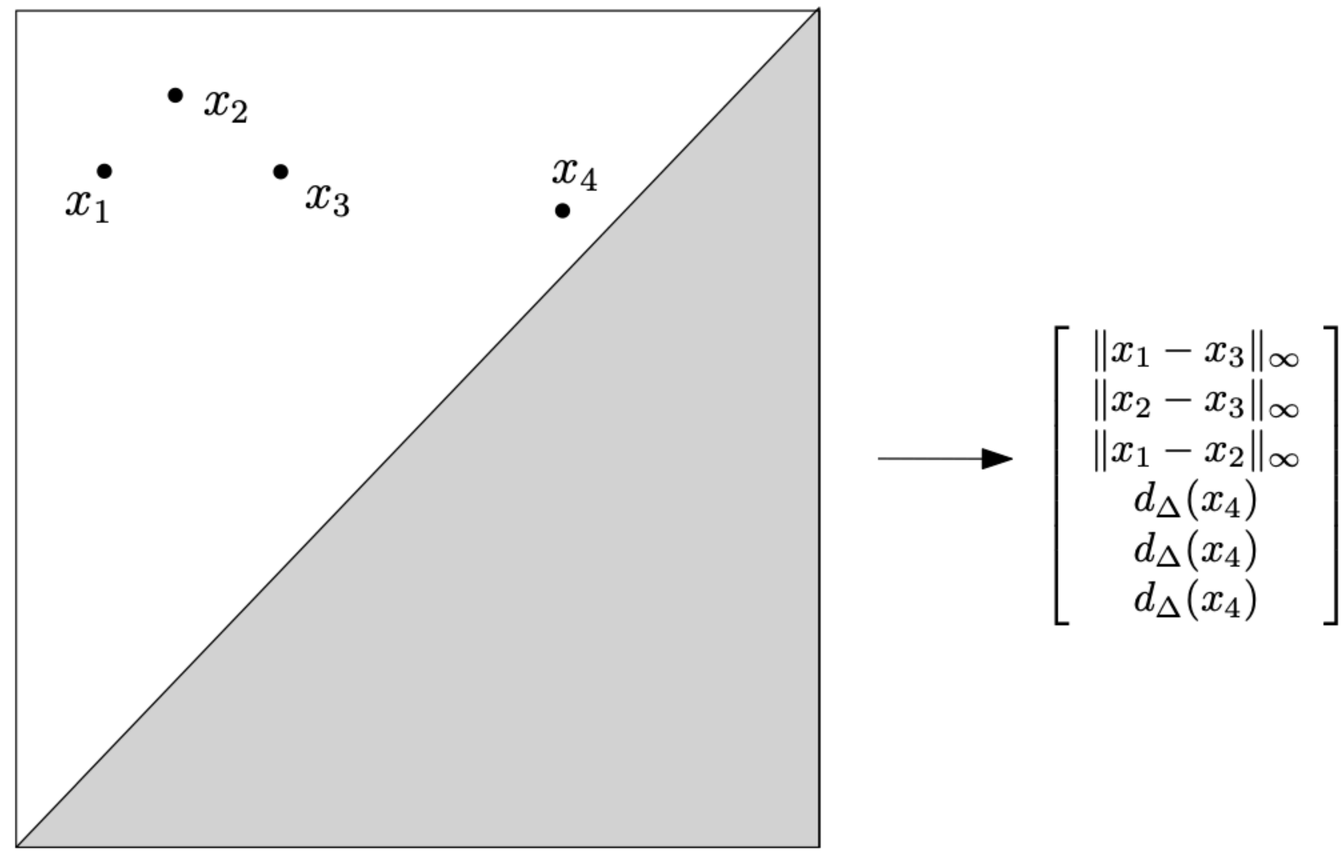
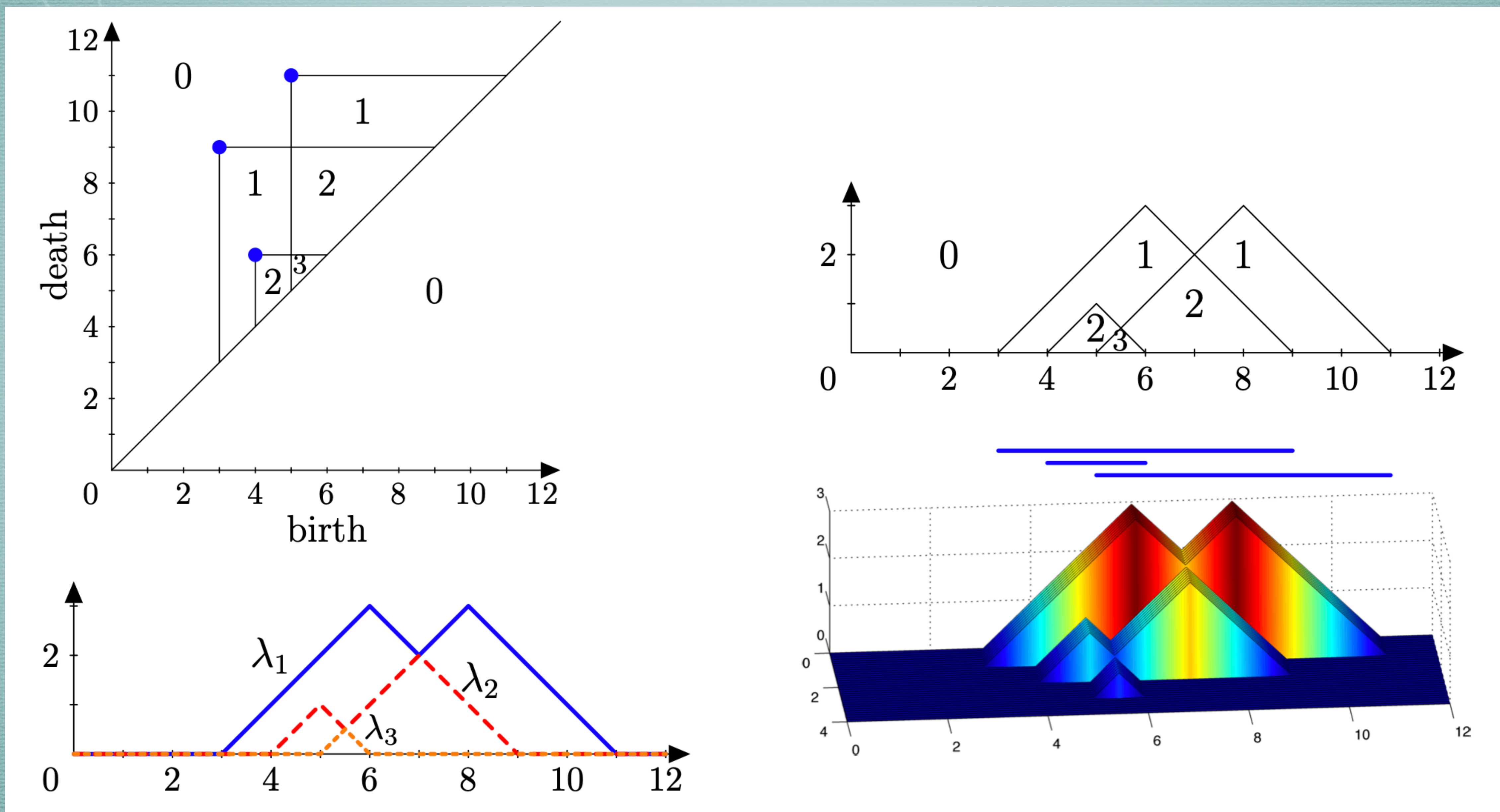


Figure 6: *Mapping of a PD to a vector, where $d_\Delta(\cdot)$ denotes the distance to the diagonal Δ.*

Mathieu Carrière, Steve Y. Oudot, and Maks Ovsjanikov. Stable topological signatures for points on 3d shapes.

Persistence landscape



P. Bubenik. Statistical topological data analysis using persistence landscapes.

Persistence landscape: means

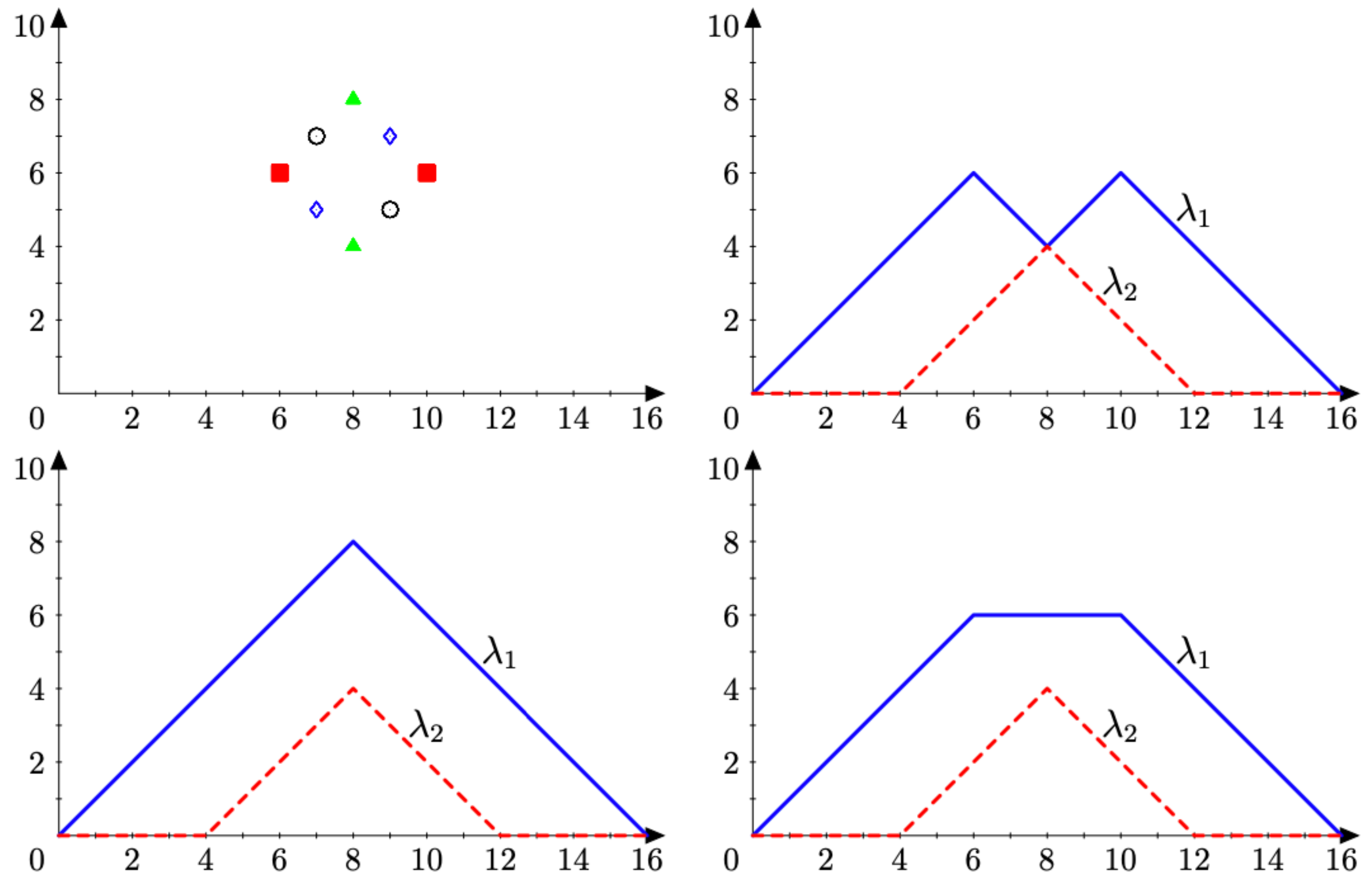


FIGURE 3. Means of persistence diagrams and persistence landscapes. Top left: the rescaled persistence diagrams $\{(6,6),(10,6)\}$ and $\{(8,4),(8,8)\}$ have two (Fréchet) means: $\{(7,5),(9,7)\}$ and $\{(7,7),(9,5)\}$. In contrast their corresponding persistence landscapes (top right and bottom left) have a unique mean (bottom right).

P. Bubenik. Statistical topological data analysis using persistence landscapes.

Persistence landscape: means

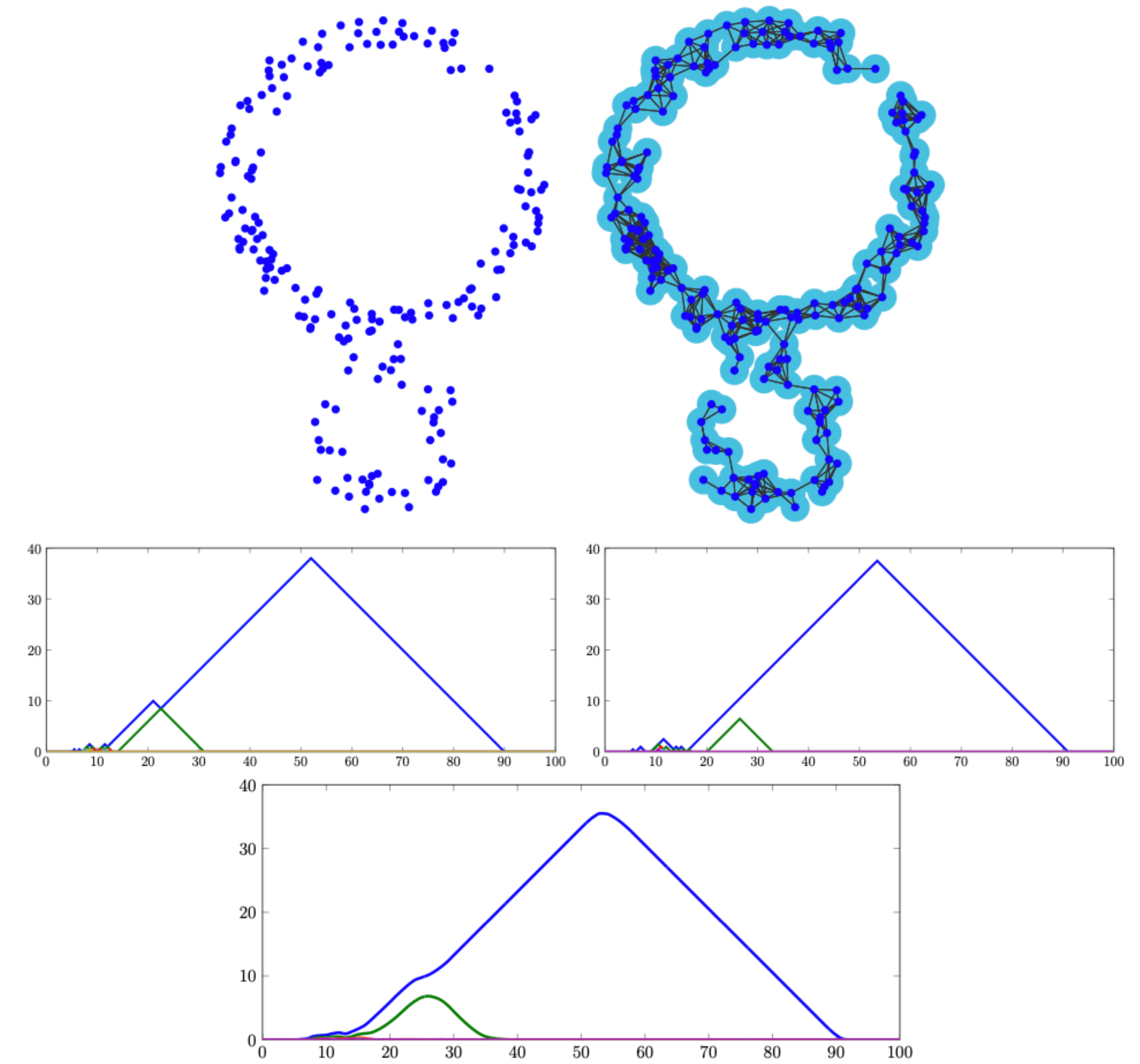
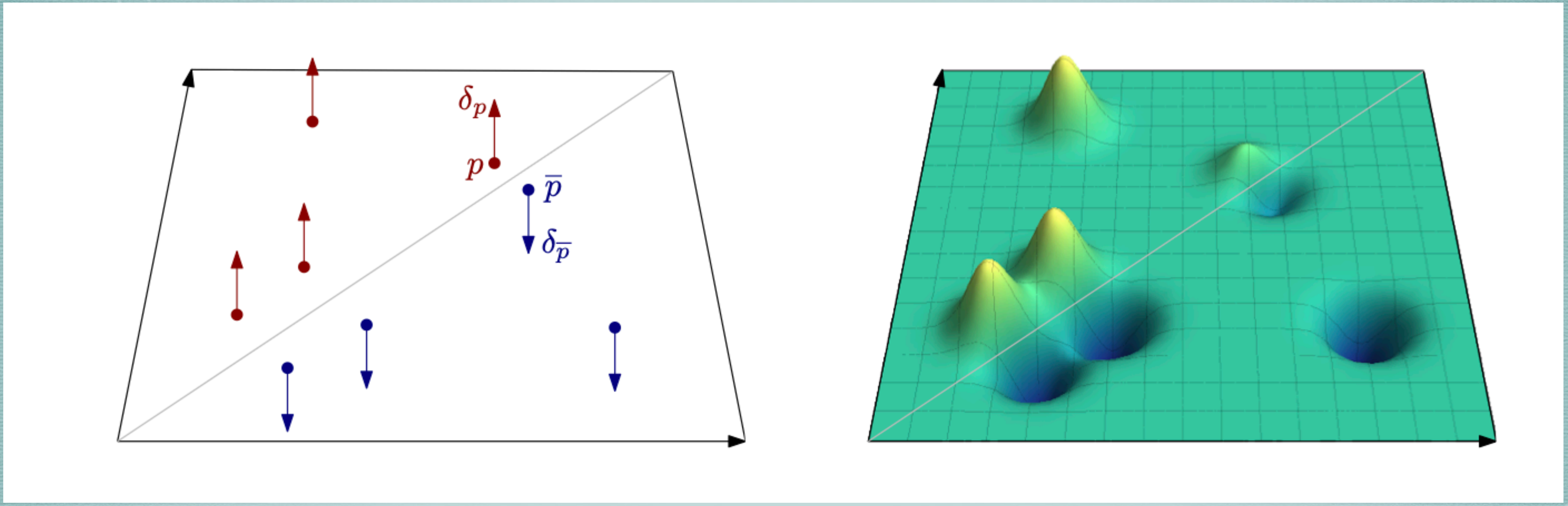


FIGURE 4. 200 points were sampled from a pair of linked annuli. Here we show the points and a corresponding union of balls and 1-skeleton of the Čech complex. This was repeated 100 times. Next we show two of the degree one persistence landscapes and the mean degree one persistence landscape.

P. Bubenik. Statistical topological data analysis using persistence landscapes.

Multi-scale kernel



Jan Reininghaus, Stefan Huber, Ulrich Bauer, and Roland Kwitt. A stable multi-scale kernel for topological machine learning.

Persistence images

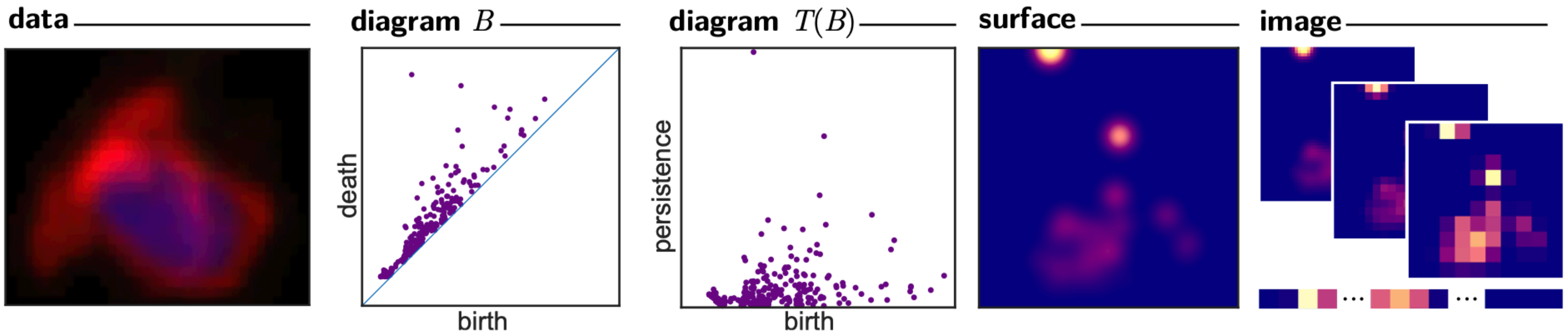


Figure 1: Algorithm pipeline to transform data into a persistence image.

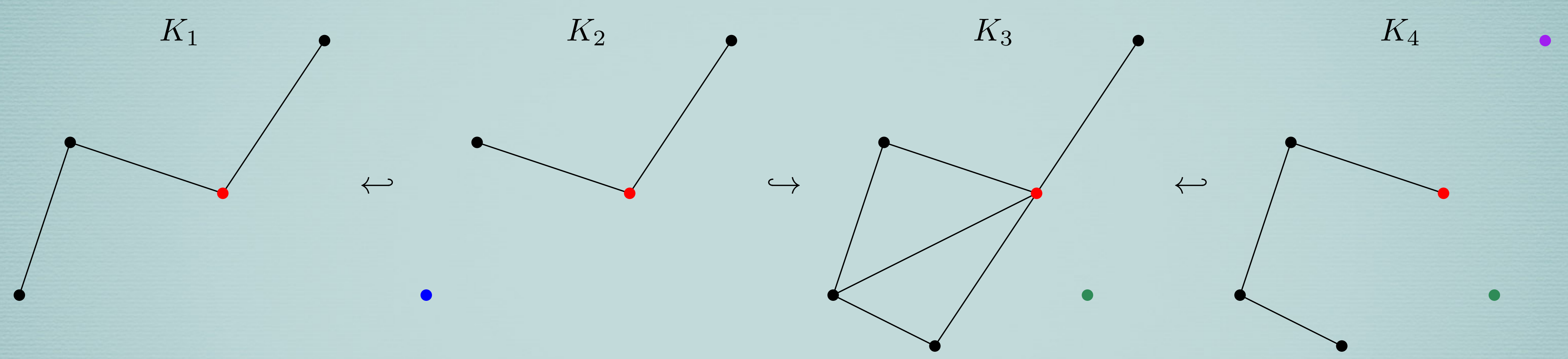
H. Adams et. al: Persistence images: A stable vector representation of persistent homology

Further analysis and examples: Danielle Barnes, Luis Polanco and Jose A. Perea: A Comparative Study of Machine Learning Methods for Persistence Diagrams

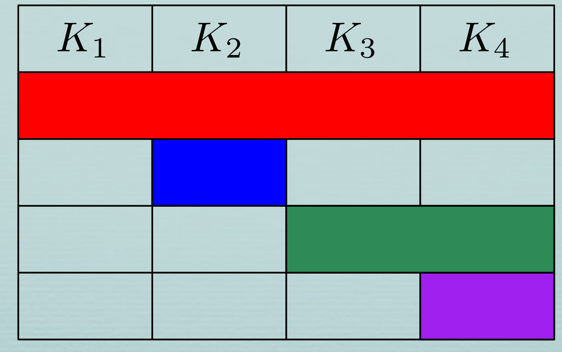
Generalizations

Zig-Zag persistent homology

Instead of “linear” filtration with inclusions, we have a sequence of complexes with inclusions (or maps) in **some** direction



They still deompose as bars.



G. Carlsson, V. de Silva, and D. Morozov. Zigzag persistent homology and real-valued functions.

Multi-parameter persistence

$$
\begin{array}{ccccccccc}
K_{1,m} & \hookrightarrow & K_{2,m} & \hookrightarrow & K_{3,m} & \hookrightarrow & \cdots & \hookrightarrow & K_{m,m} \\
\uparrow & & \uparrow & & \uparrow & & & & \uparrow \\
\vdots & & \vdots & & \vdots & & & & \vdots \\
\uparrow & & \uparrow & & \uparrow & & & & \uparrow \\
K_{1,2} & \hookrightarrow & K_{2,2} & \hookrightarrow & K_{3,2} & \hookrightarrow & \cdots & \hookrightarrow & K_{m,2} \\
\uparrow & & \uparrow & & \uparrow & & & & \uparrow \\
K_{1,1} & \hookrightarrow & K_{2,1} & \hookrightarrow & K_{3,1} & \hookrightarrow & \cdots & \hookrightarrow & K_{m,1}
\end{array}
$$

Decomposition problematic, other summaries are used.

Other settings

- Stochastic topology
- Sliding window (persistence for time series)
- ...