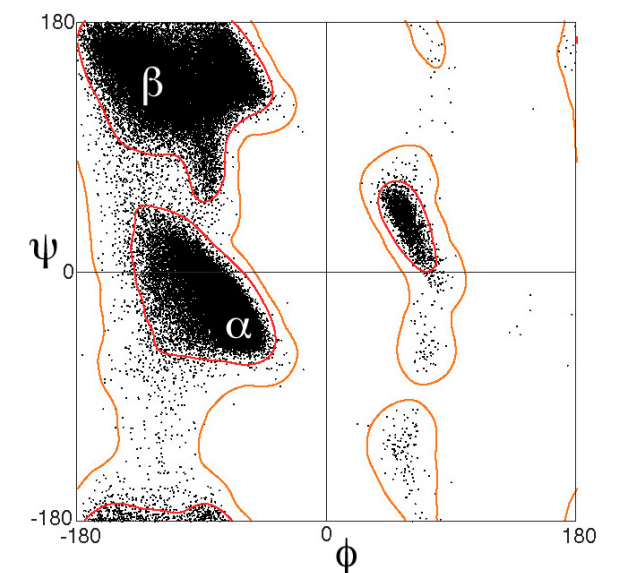
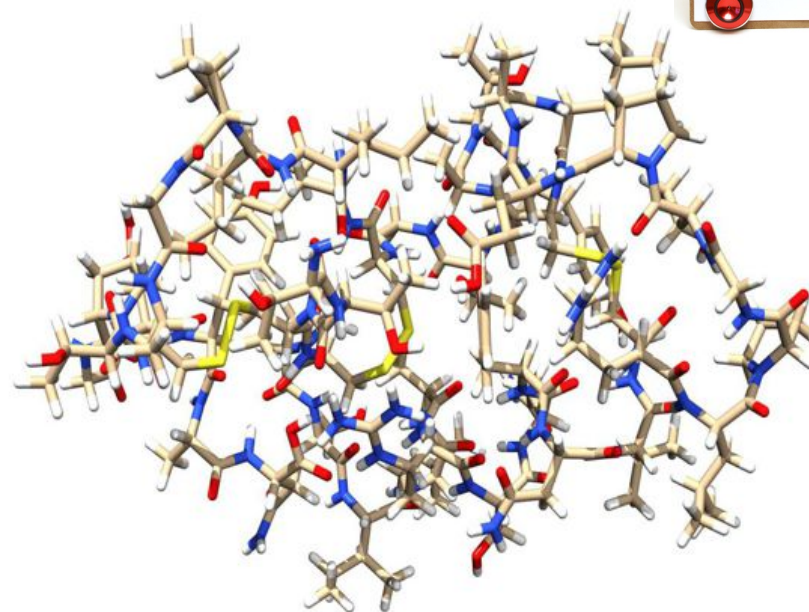
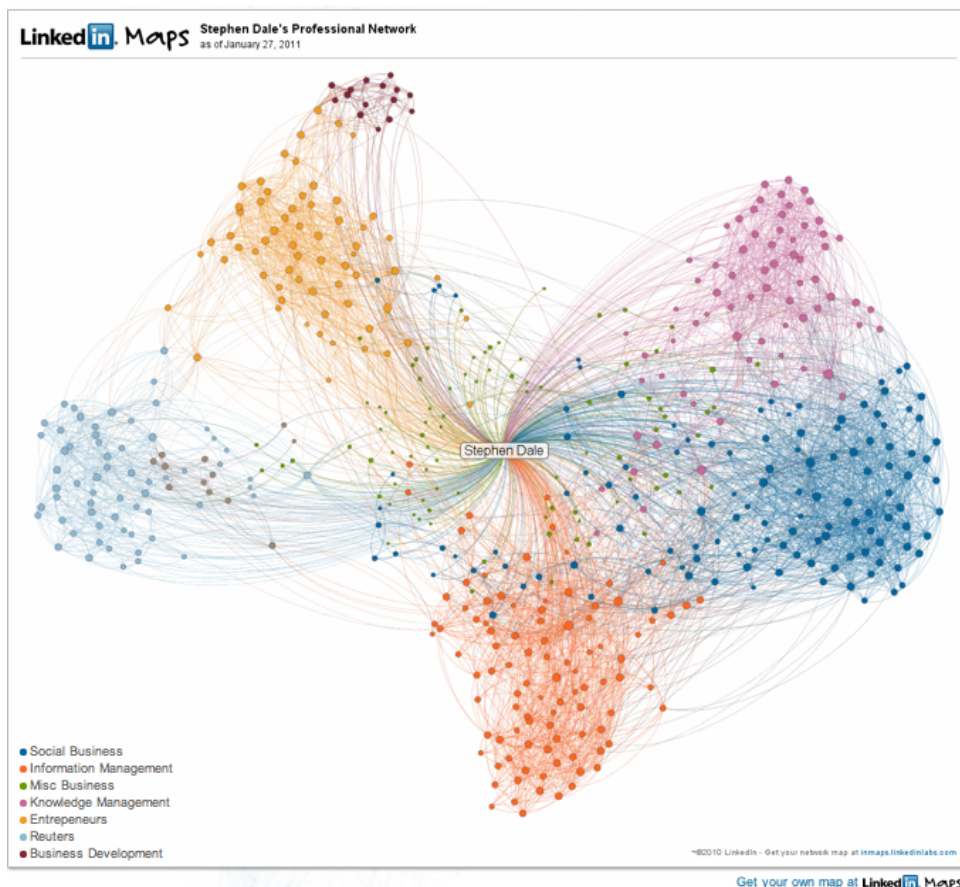
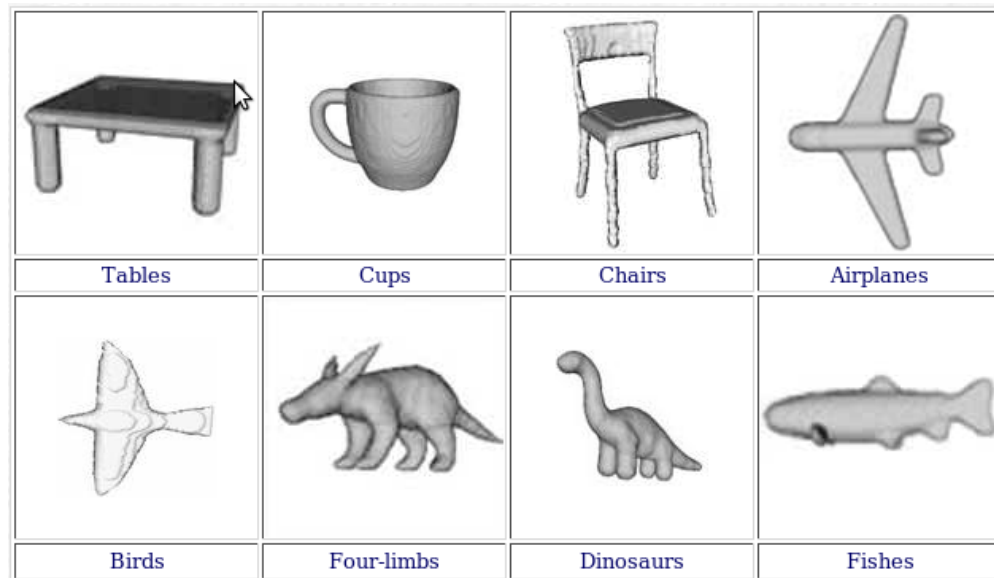


Geometric Data

Input: point cloud equipped with a metric or (dis-)similarity measure

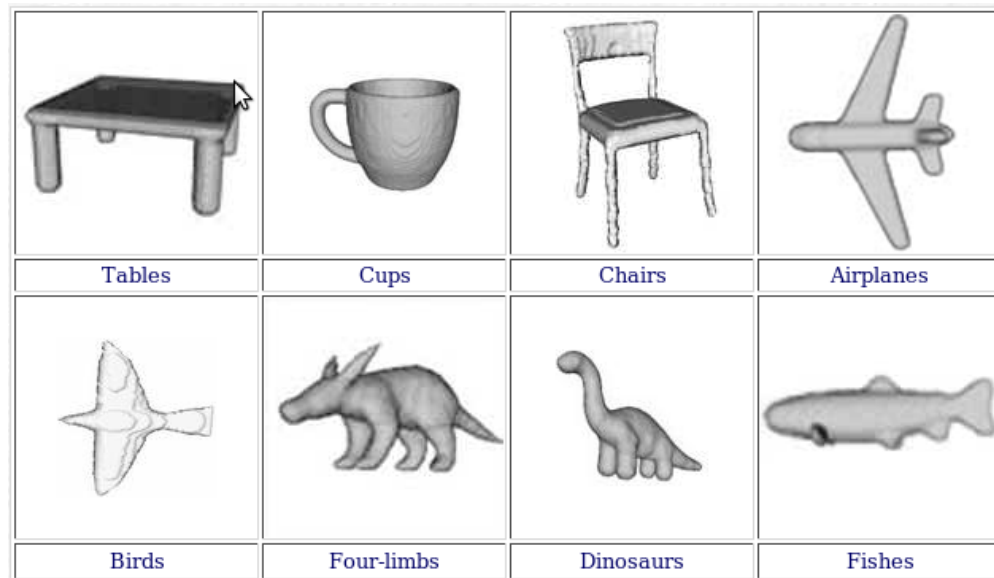
data point $\equiv$ image/patch, geometric shape, protein conformation, patient, LinkedIn user...



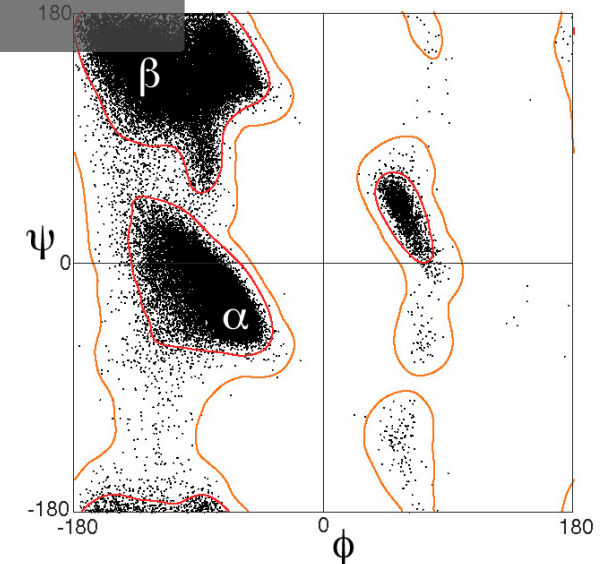
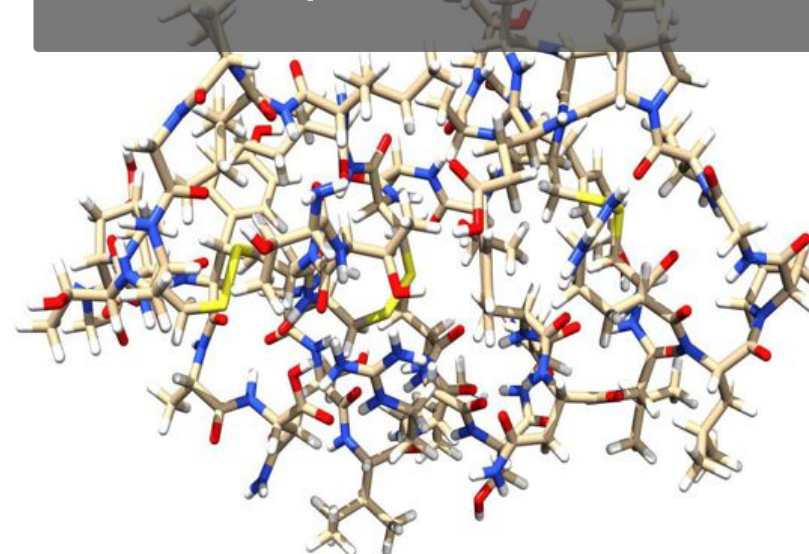
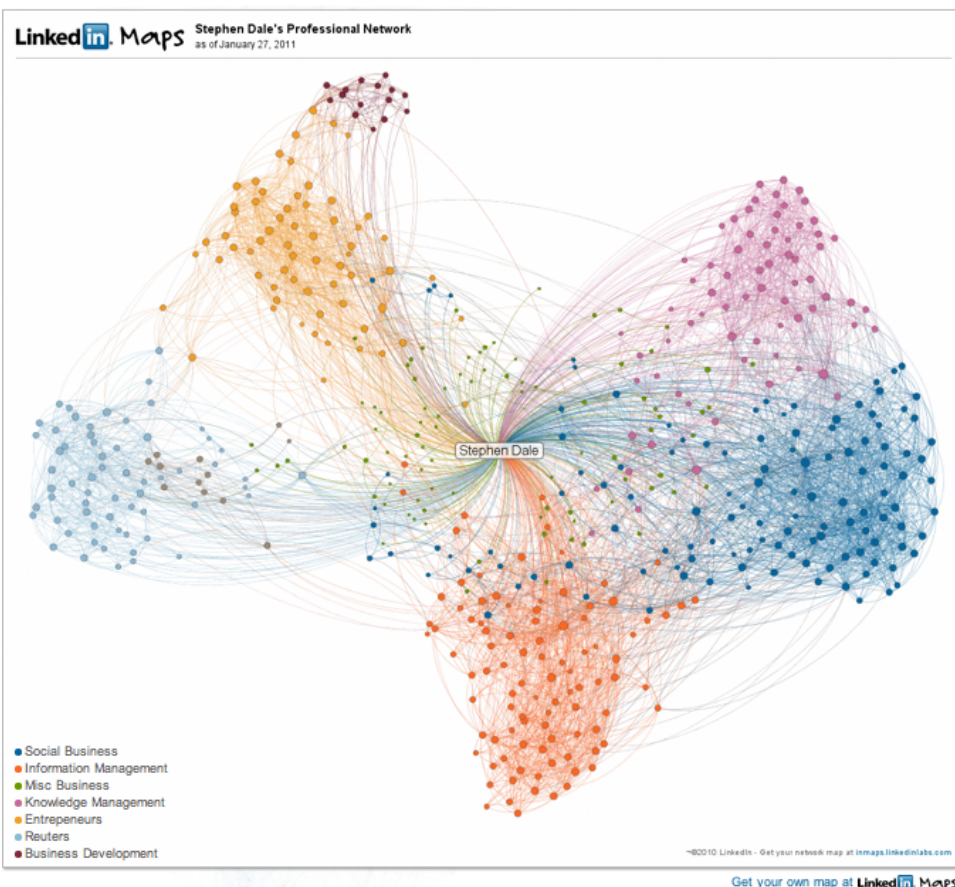
Geometric Data

Input: point cloud equipped with a metric or (dis-)similarity measure

data point $\equiv$ image/patch, geometric shape, protein conformation, patient, LinkedIn user...



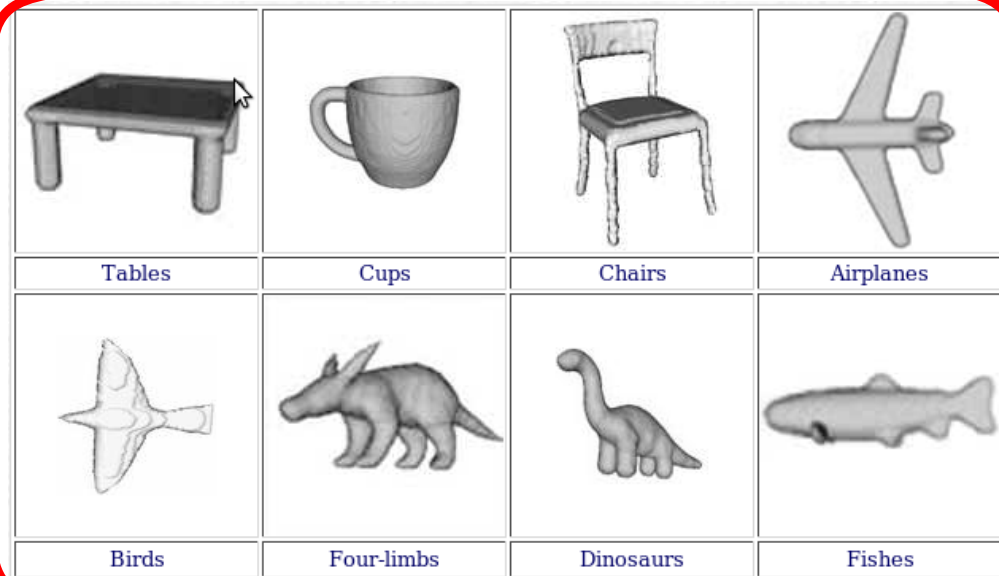
Goal: describe the structure of the geometry underlying the data, for interpretation or summary



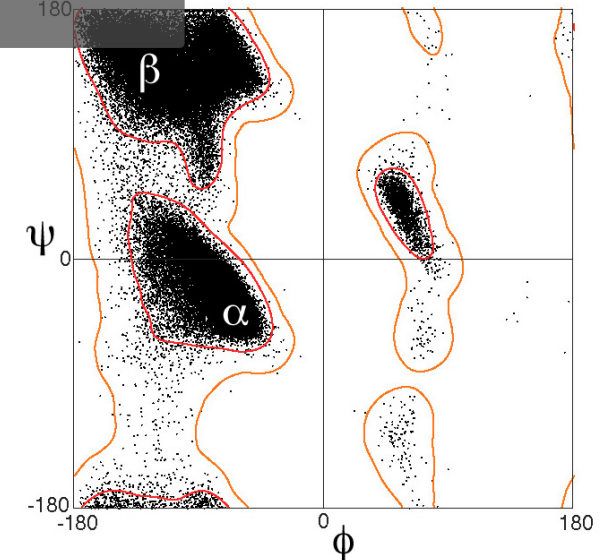
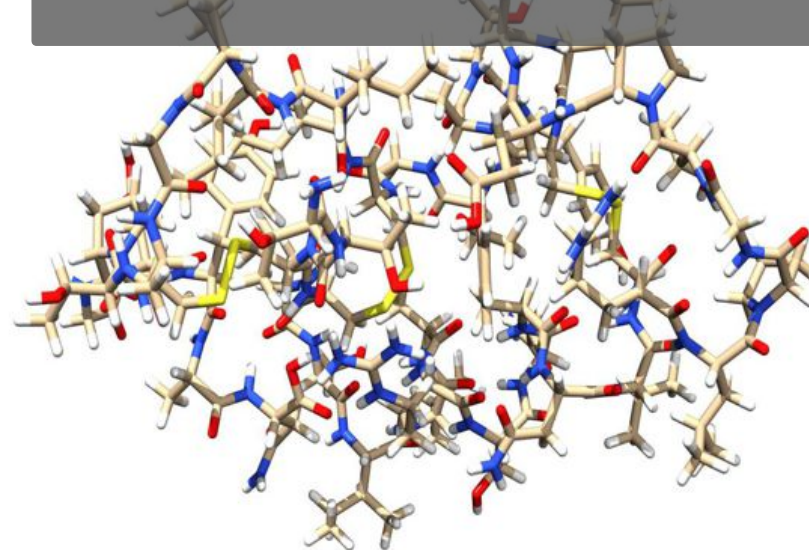
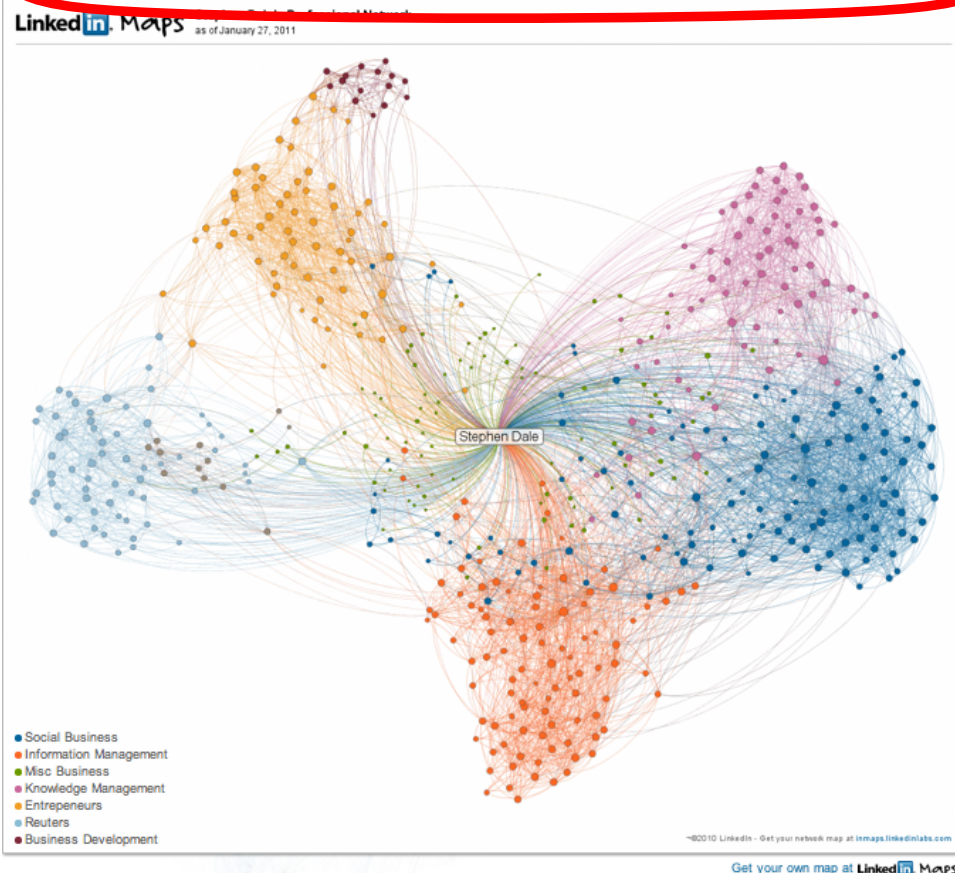
Geometric Data

Input: point cloud equipped with a **metric** or (dis-)similarity measure

data point $\equiv$ image/patch, geometric shape, protein conformation, patient, LinkedIn user...

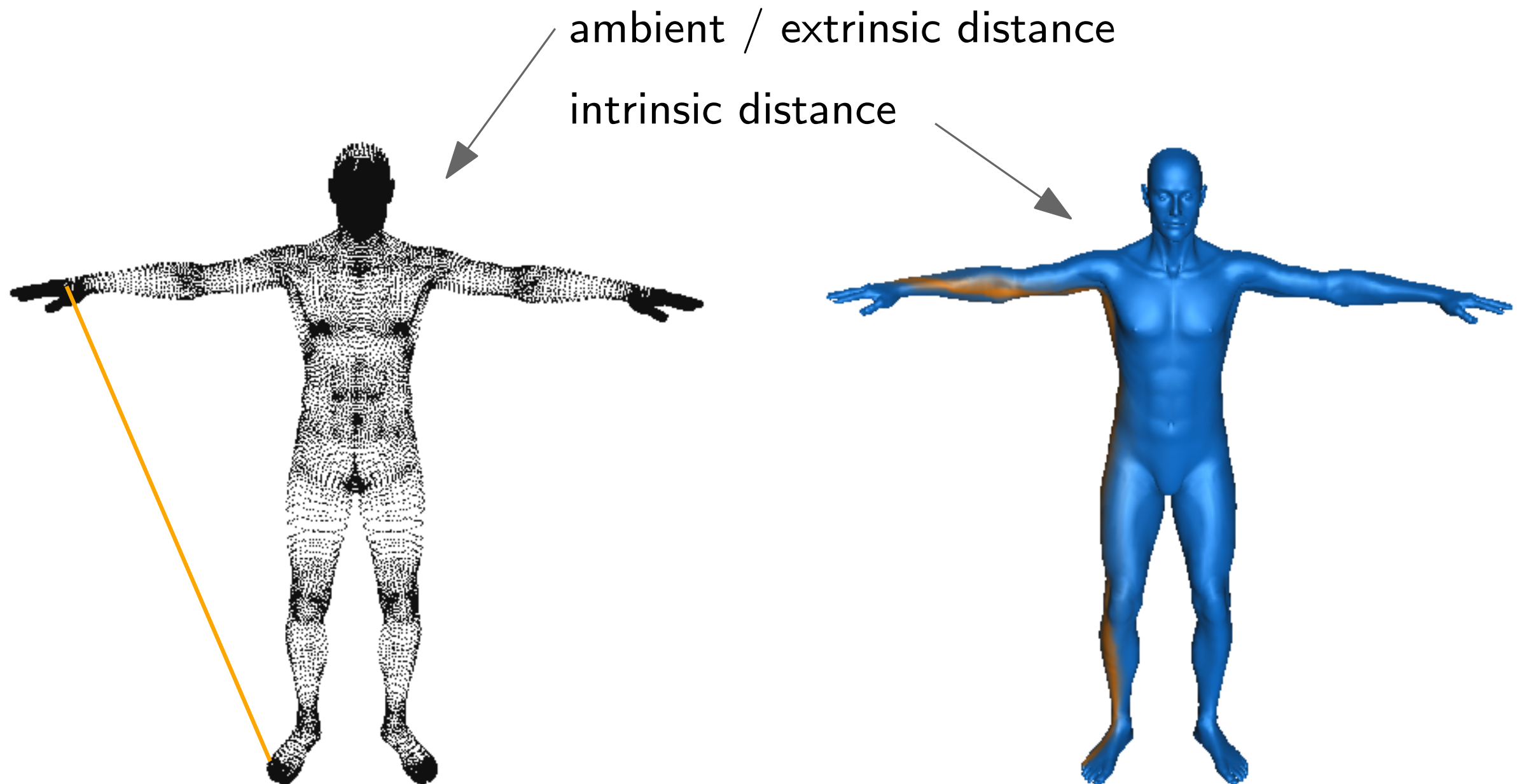


Goal: describe the structure of the geometry underlying the data, for interpretation or summary



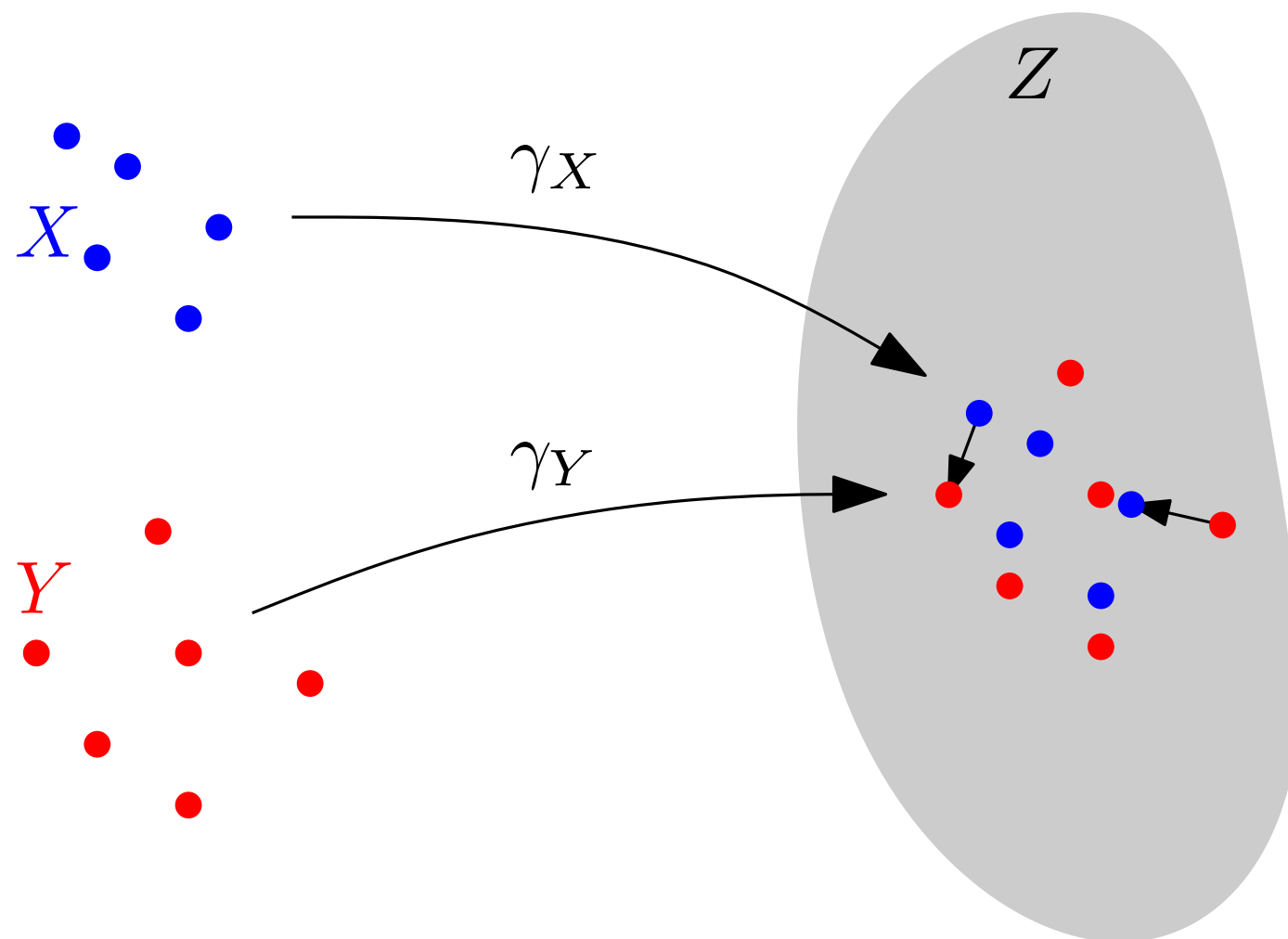
Mathematical framework

- geometric data set / underlying space $\equiv$ compact metric space



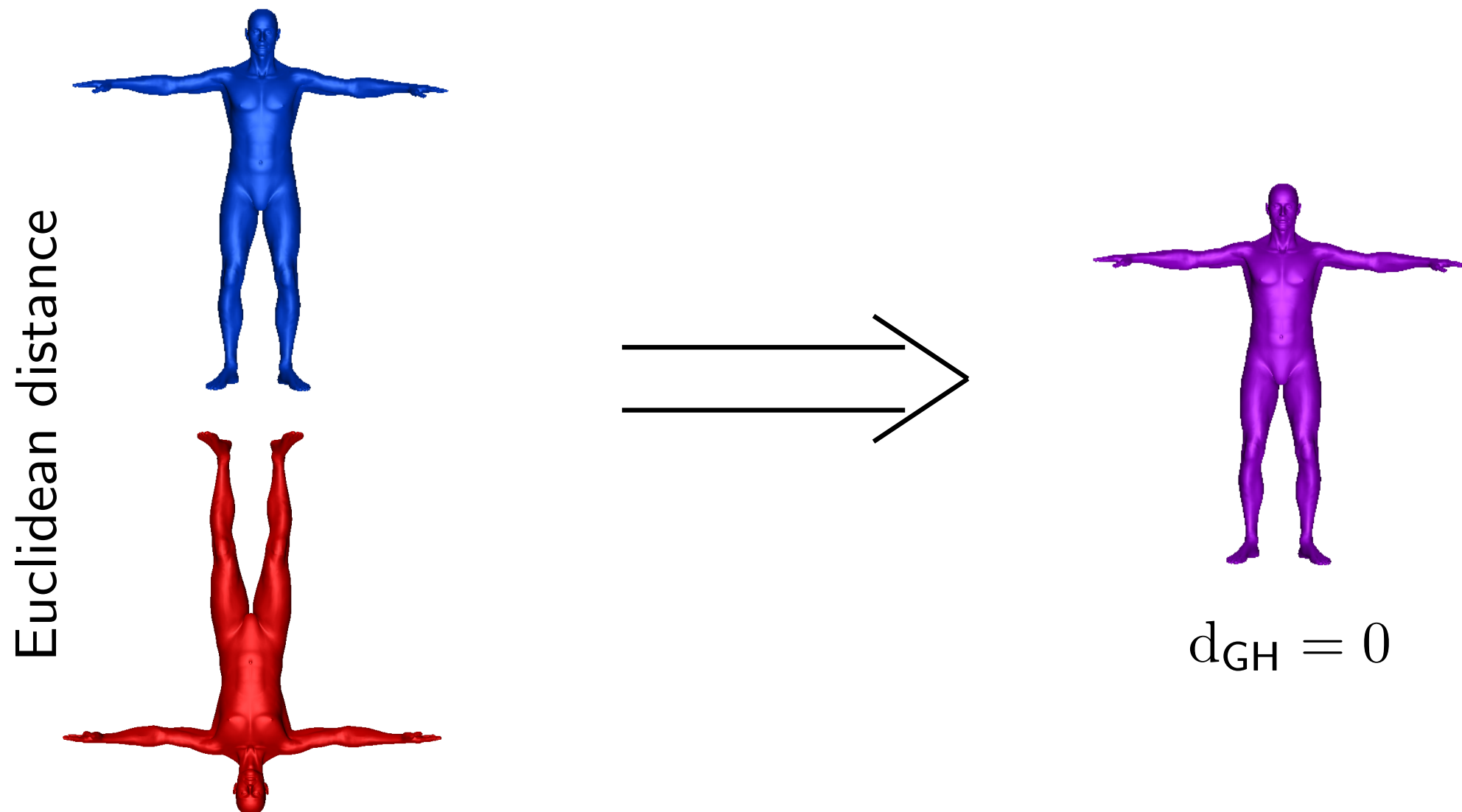
Mathematical framework

- geometric data set / underlying space $\equiv$ compact metric space
- distance between compact metric spaces $\equiv$ Gromov-Hausdorff (GH) distance



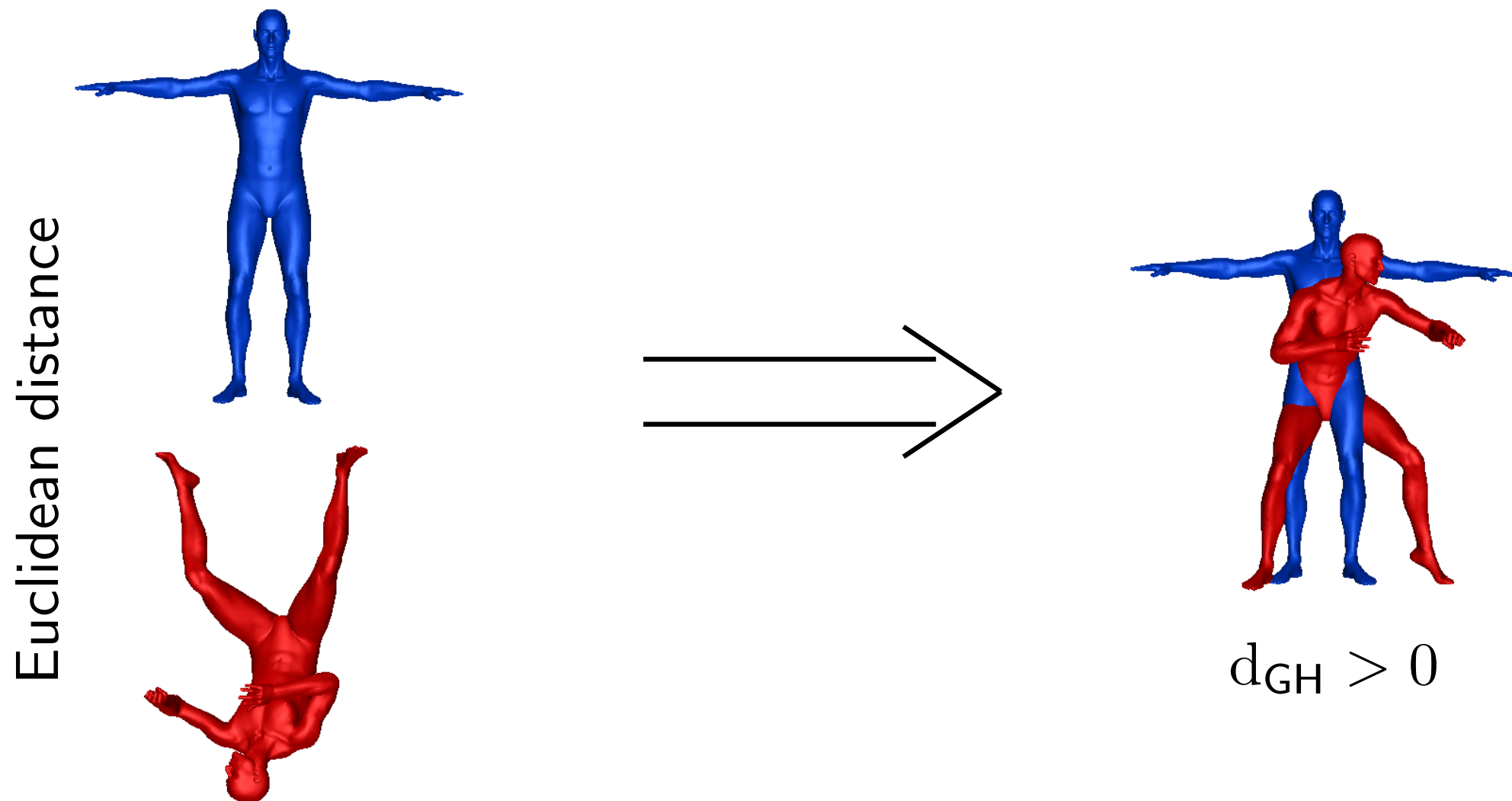
Mathematical framework

- geometric data set / underlying space $\equiv$ compact metric space
- distance between compact metric spaces $\equiv$ Gromov-Hausdorff (GH) distance



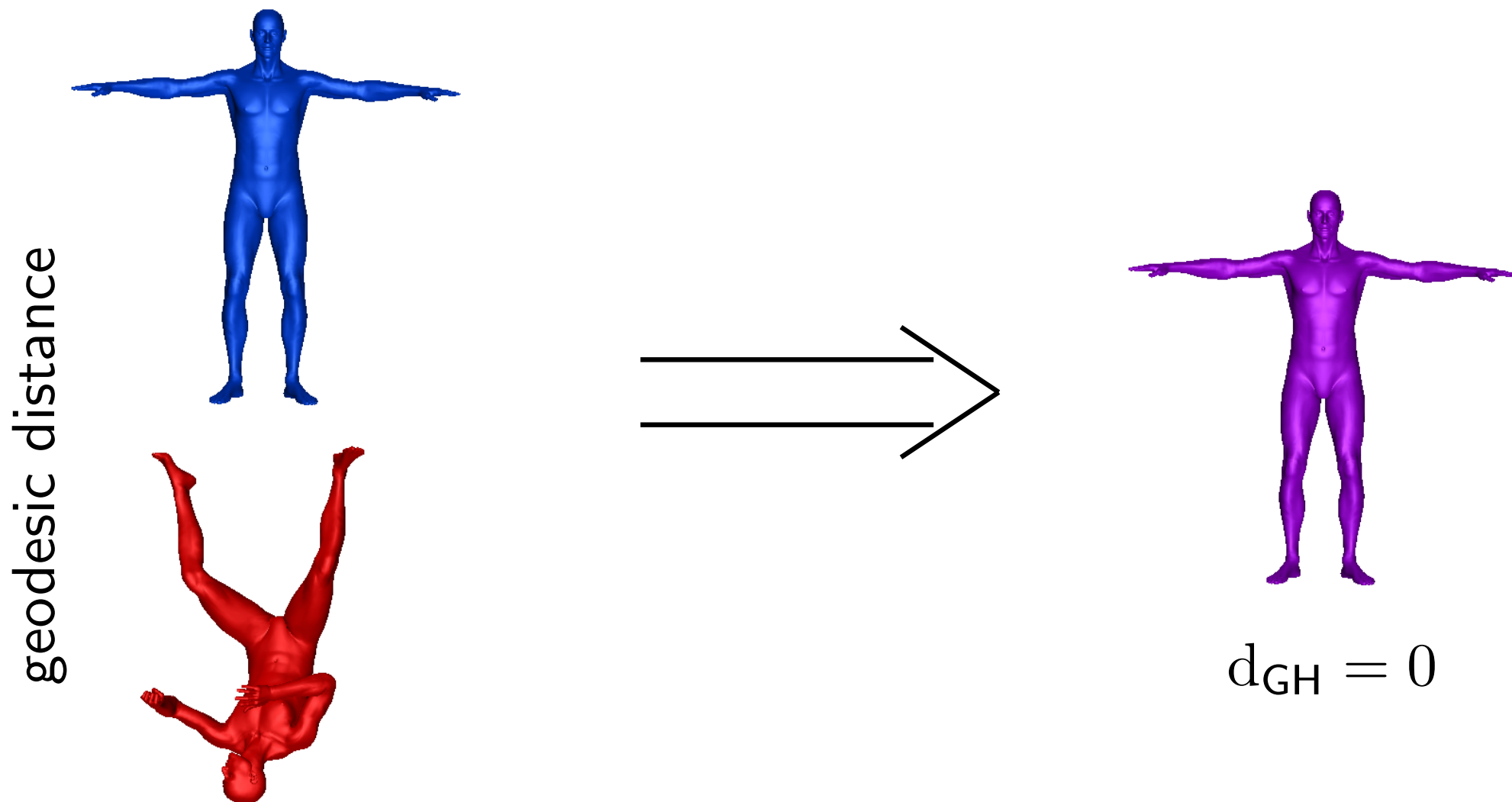
Mathematical framework

- geometric data set / underlying space $\equiv$ compact metric space
- distance between compact metric spaces $\equiv$ Gromov-Hausdorff (GH) distance



Mathematical framework

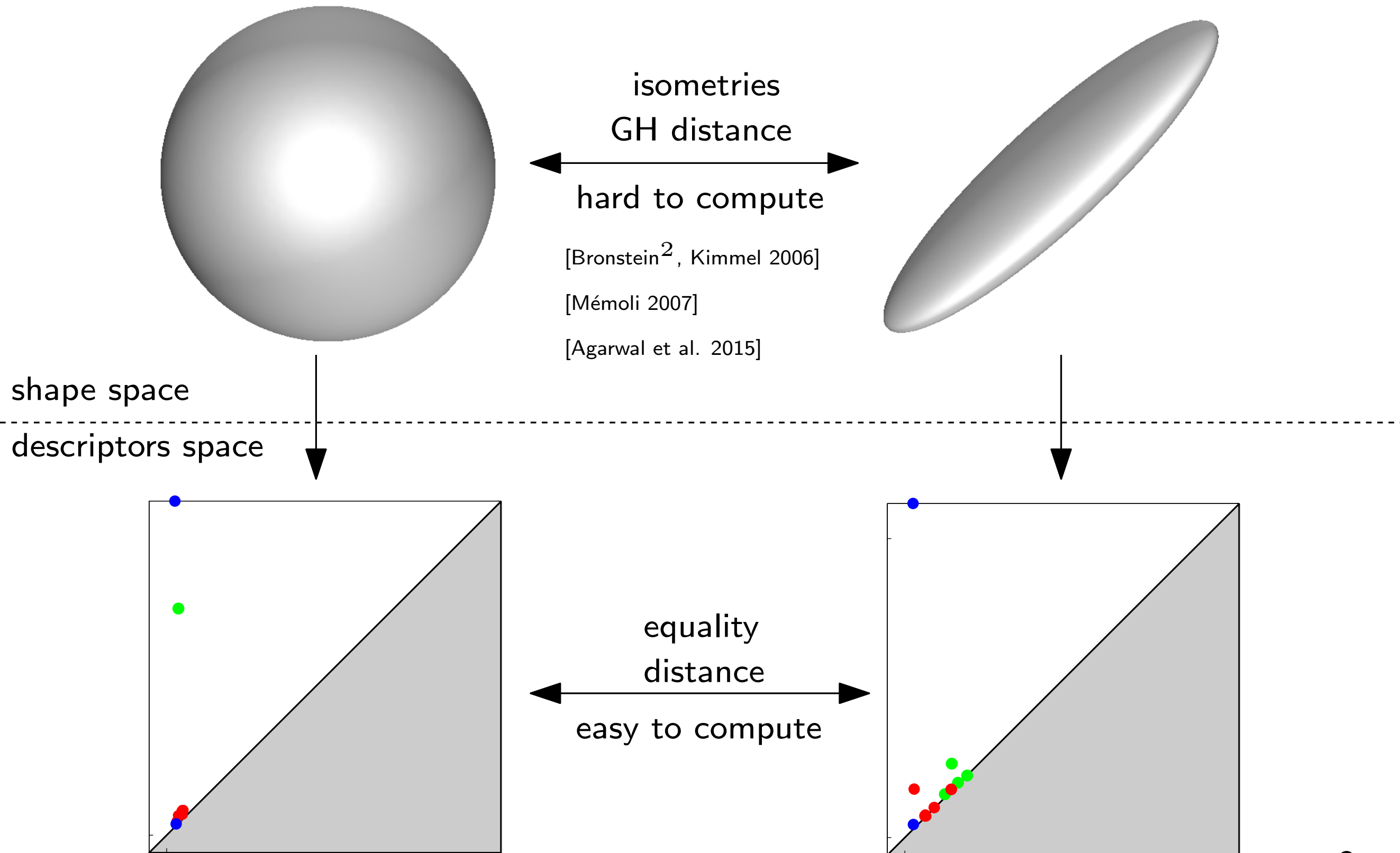
- geometric data set / underlying space $\equiv$ compact metric space
- distance between compact metric spaces $\equiv$ Gromov-Hausdorff (GH) distance



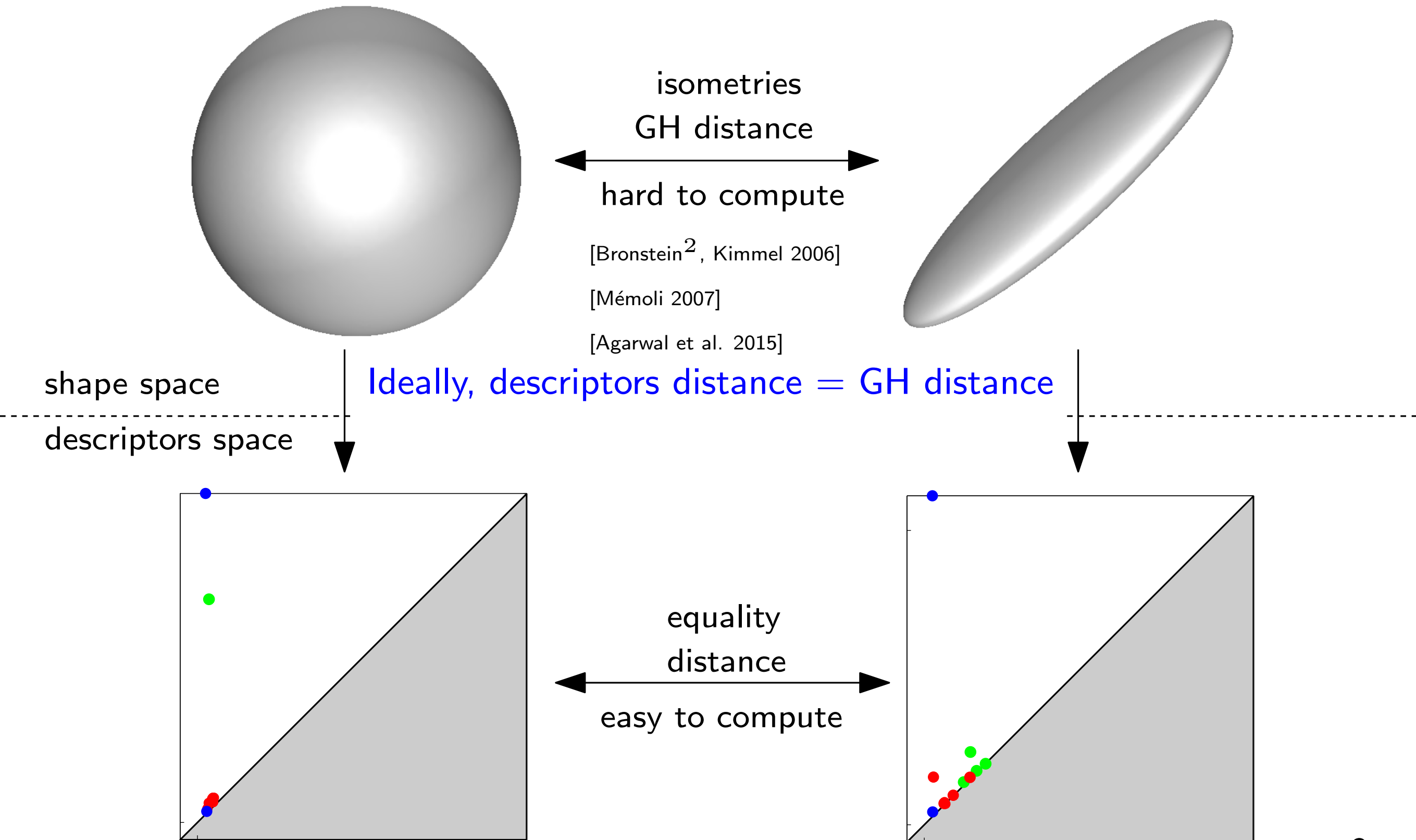
Mathematical framework

- geometric data set / underlying space $\equiv$ compact metric space
- distance between compact metric spaces $\equiv$ Gromov-Hausdorff (GH) distance
- *descriptor* $\equiv$ persistence diagram (choose the filtration)
 - **multi-scale** $\equiv$ reflects the structure of the shape across scales
 - **global/local** $\equiv$ attached to the whole shape / to a base point(s)
 - **stable** $\equiv$ variations with GH-distance and base point location are controlled

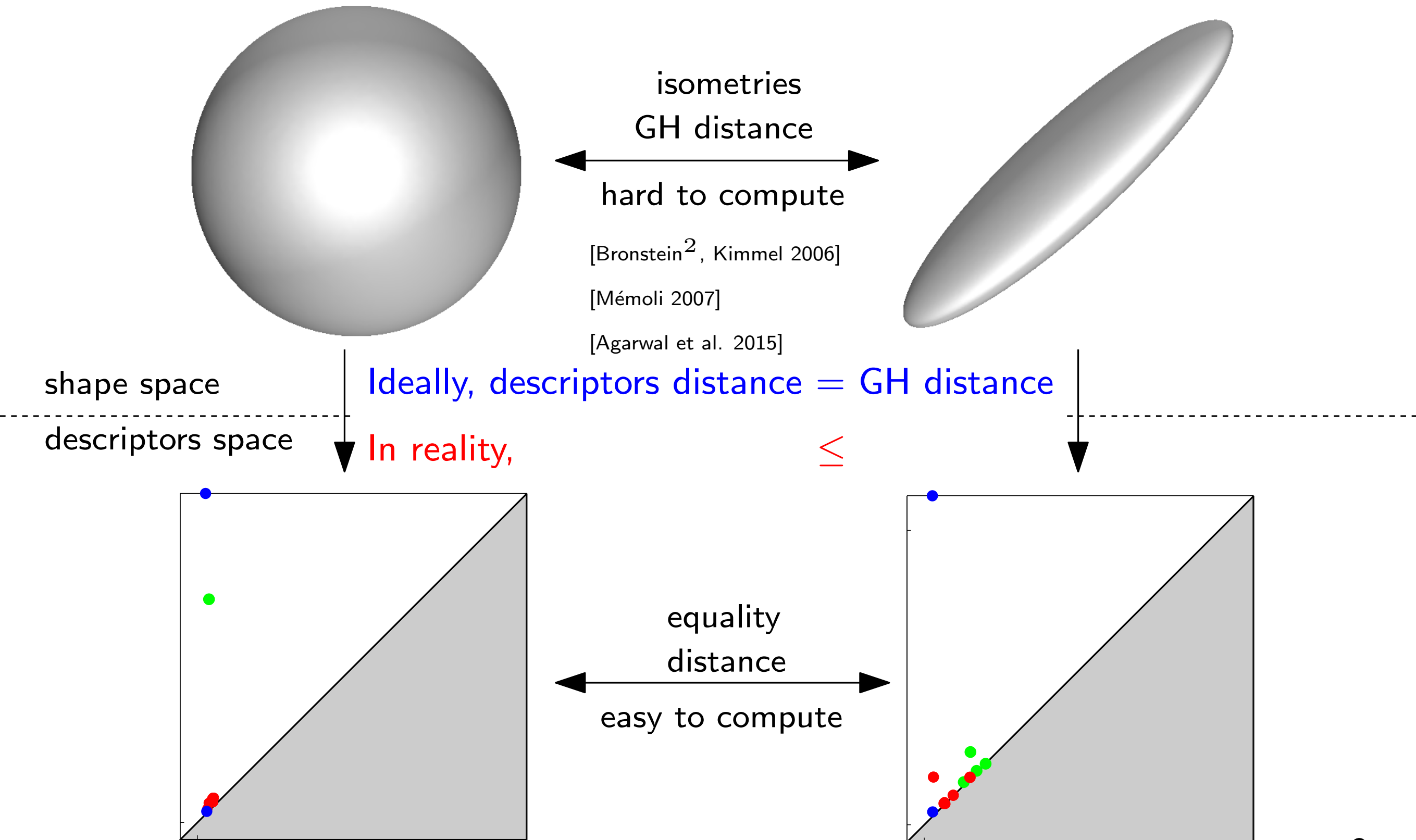
Why use descriptors



Why use descriptors



Why use descriptors



Why use descriptors

Some descriptors for images / 3d shapes / metric spaces:

- diameter
- curvature (mean, Gaussian, sectional)
- shape context (distribution of distances)
- heat kernel signature (heat diffusion)
- wave kernel signature (Maxwell's equations)
- spin image (local neighborhood parametrization)
- SIFT features (local distribution of gradient orientations)
- etc.

Why use descriptors

Some descriptors for images / 3d shapes / metric spaces:

- diameter
- curvature (mean, Gaussian, sectional)
- shape context (distribution of distances)
- heat kernel signature (heat diffusion)
- wave kernel signature (Maxwell's equations)
- spin image (local neighborhood parametrization)
- SIFT features (local distribution of gradient orientations)
- etc.

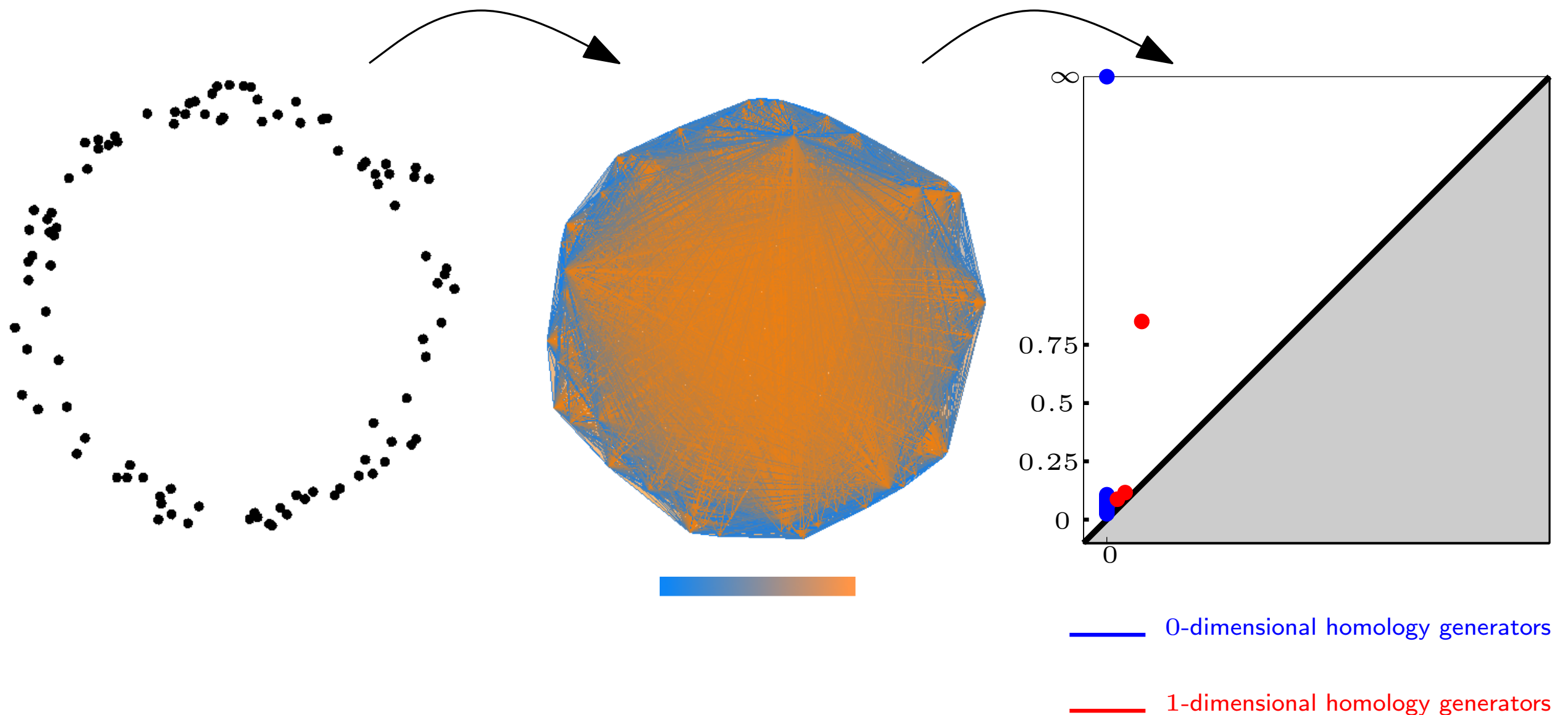
geometry
statistics

1. Global topological descriptors
2. Local topological descriptors

Global topological descriptors

Input: a compact metric space (X, d_X)

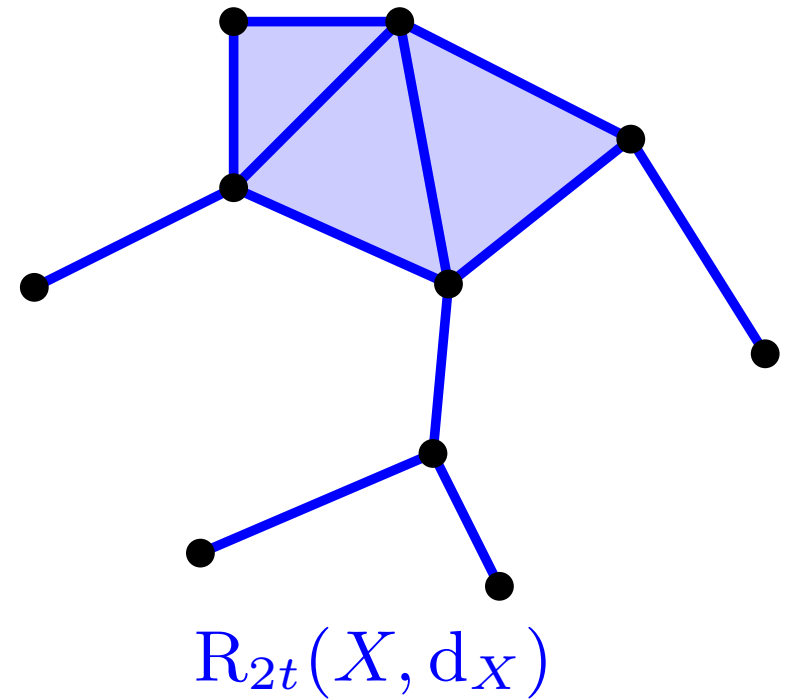
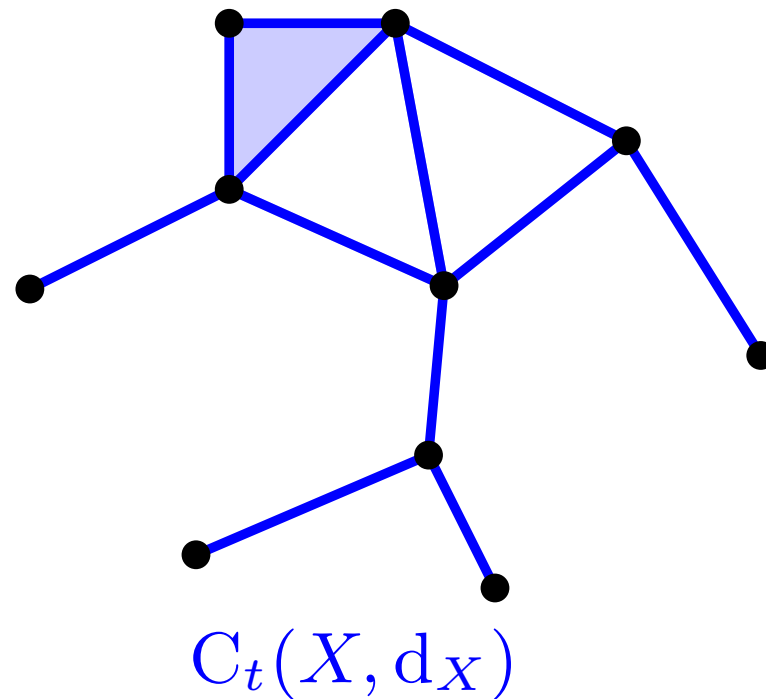
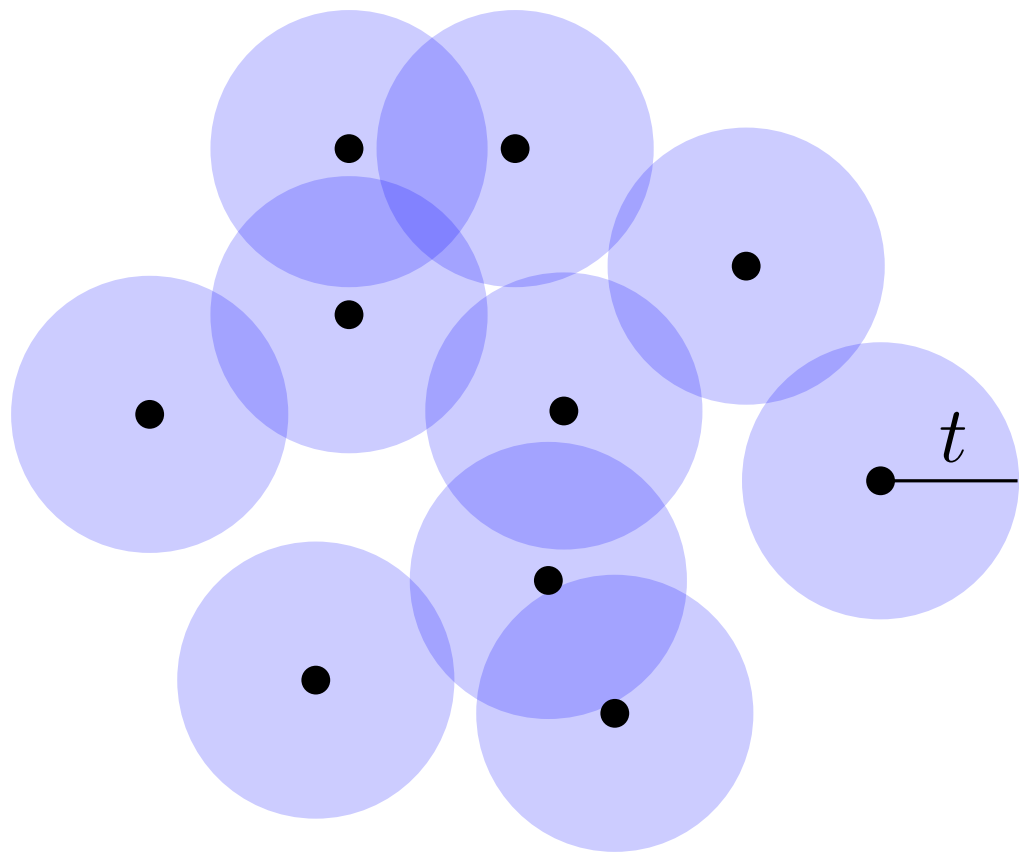
Descriptor: $\text{dgm } \mathcal{F}(X, d_X)$, where $\mathcal{F}(X, d_X)$ is some simplicial filtration over X derived from d_X (proxy for union of balls)



Global topological descriptors

Input: a compact metric space (X, d_X)

Descriptor: $\text{dgm } \mathcal{F}(X, d_X)$, where $\mathcal{F}(X, d_X)$ is some simplicial filtration over X derived from d_X (proxy for union of balls)



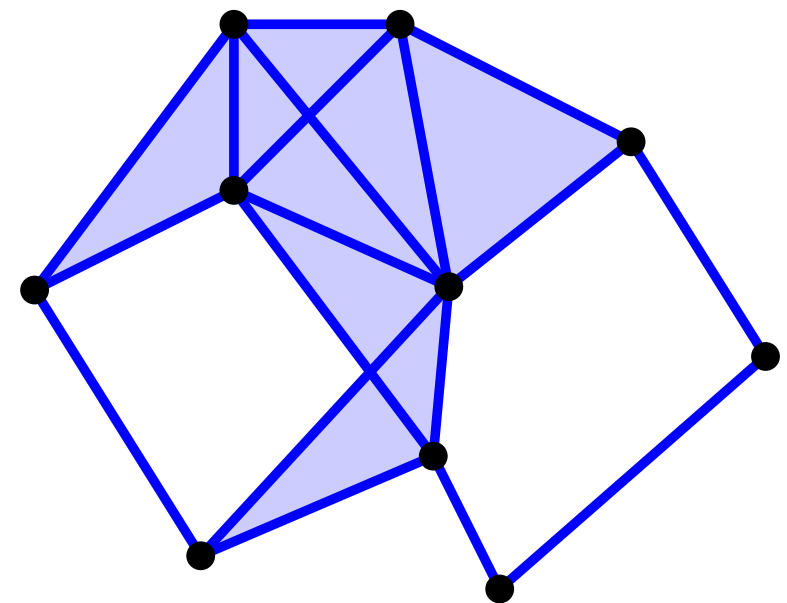
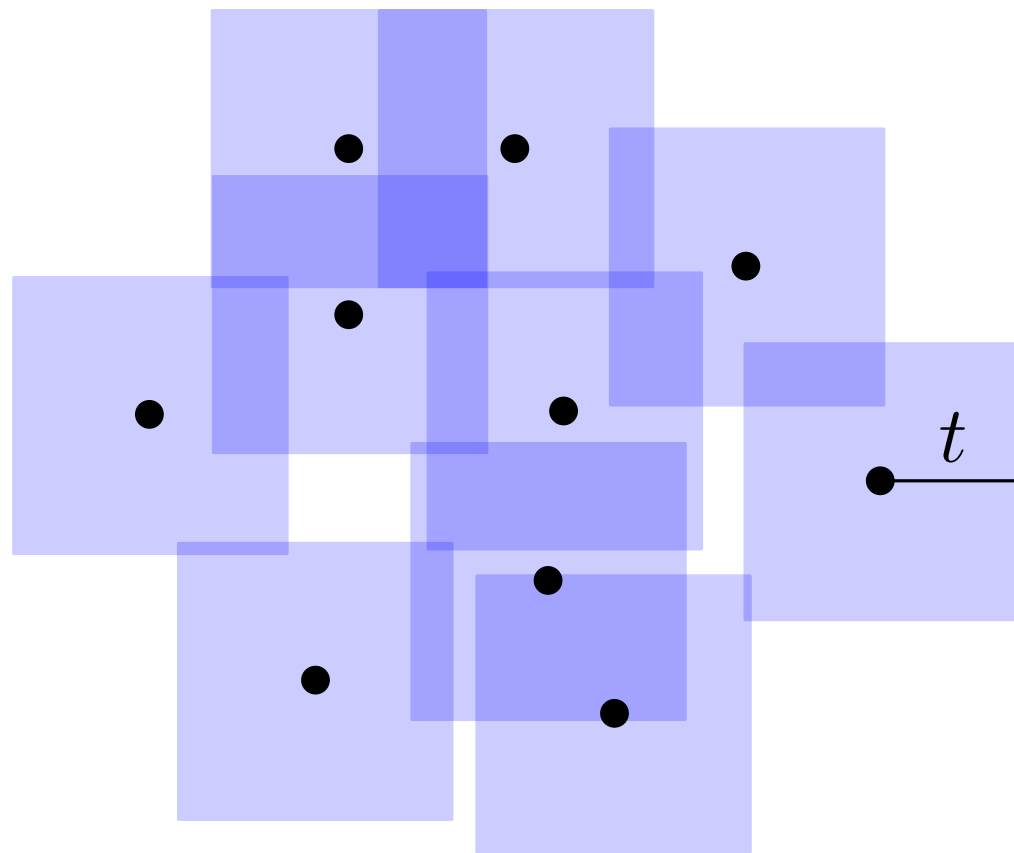
→ popular choices:

- Čech/Nerve filtration $\mathcal{C}(X, d_X)$
- (Vietoris)-Rips filtration $\mathcal{R}(X, d_X)$

Global topological descriptors

Input: a compact metric space (X, d_X)

Descriptor: $\text{dgm } \mathcal{F}(X, d_X)$, where $\mathcal{F}(X, d_X)$ is some simplicial filtration over X derived from d_X (proxy for union of balls)



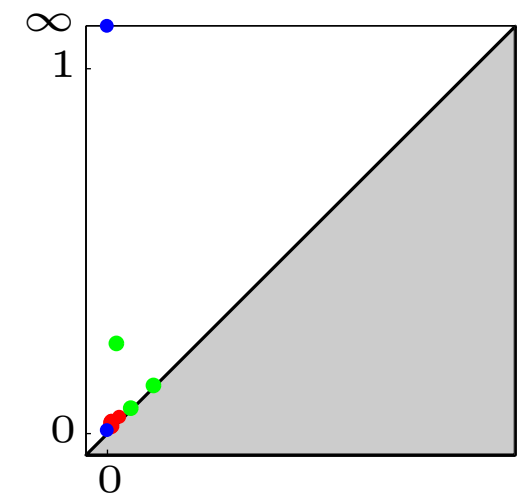
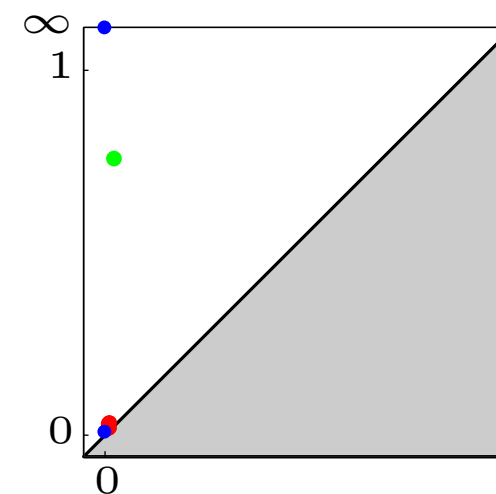
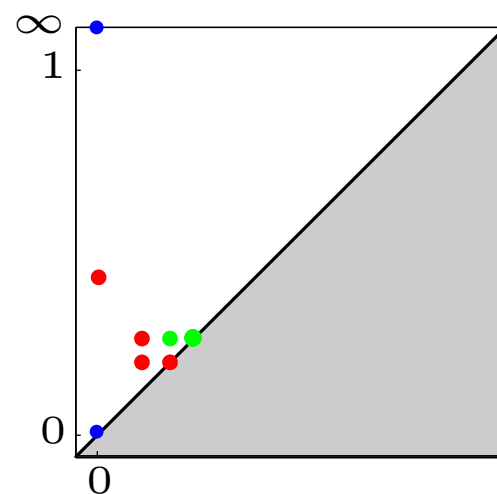
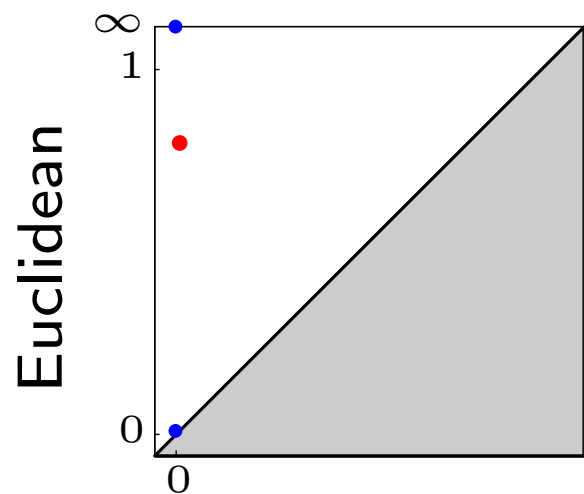
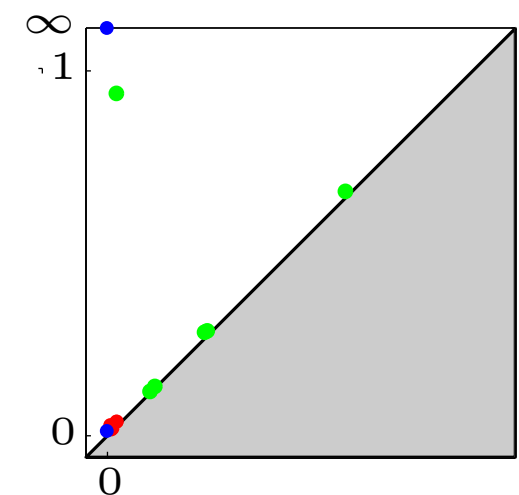
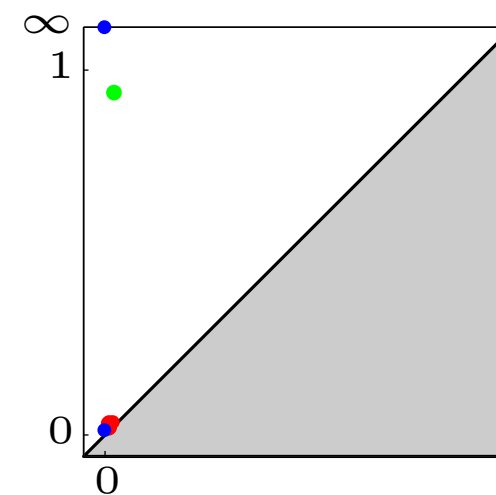
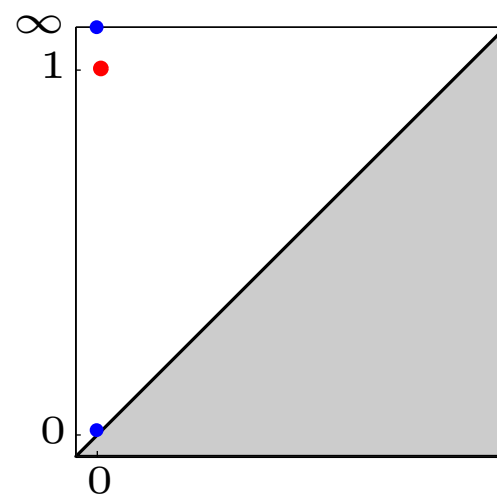
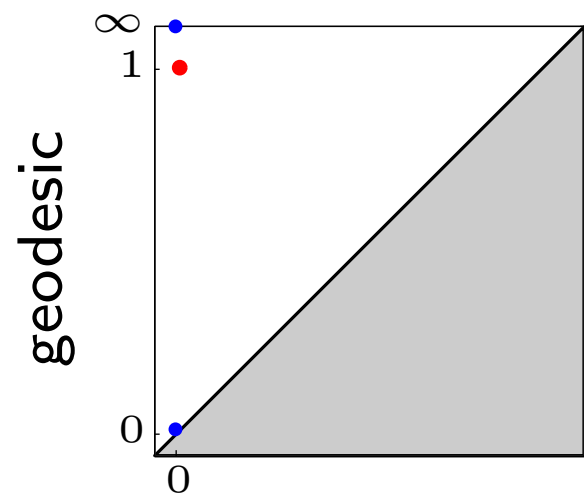
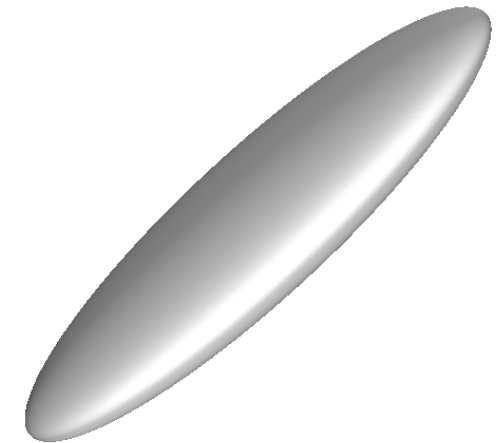
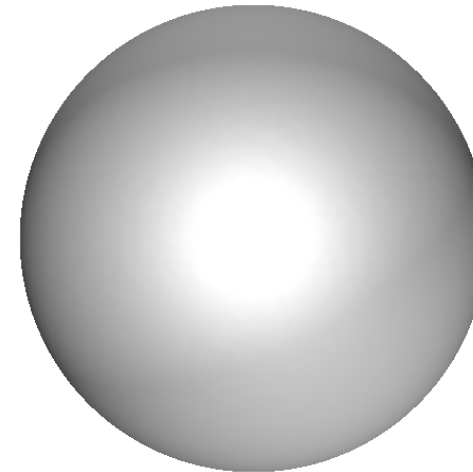
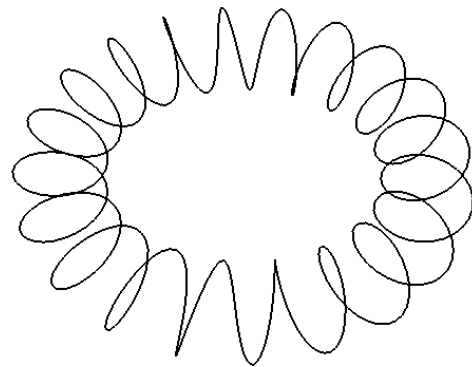
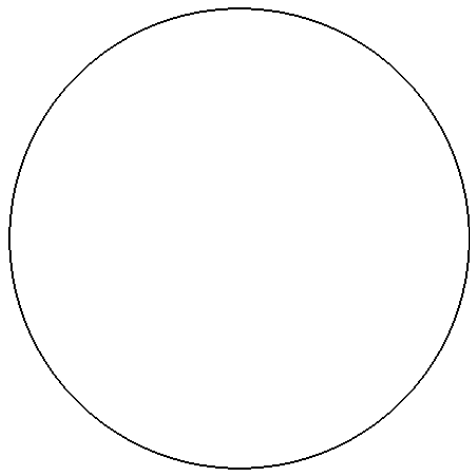
$$C_t(X, d_X) = R_{2t}(X, d_X)$$

→ popular choices:

- Čech/Nerve filtration $\mathcal{C}(X, d_X)$
- (Vietoris)-Rips filtration $\mathcal{R}(X, d_X)$

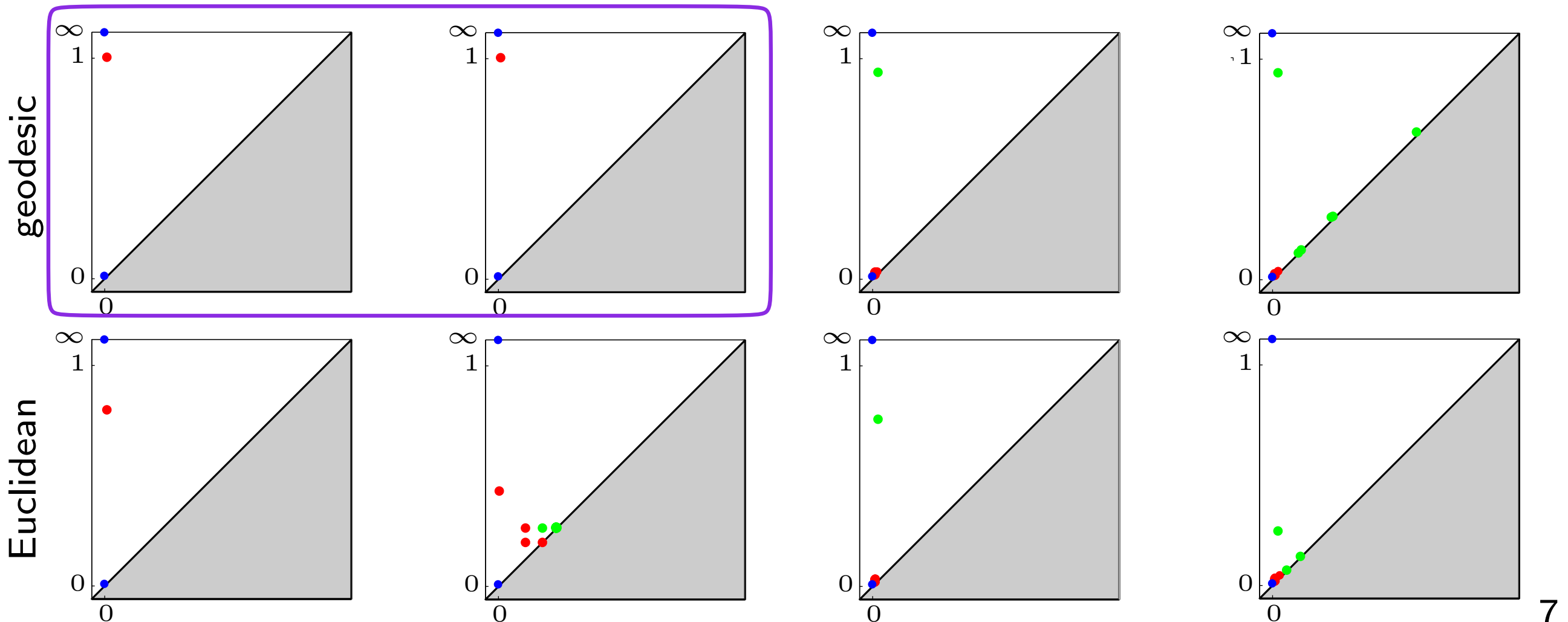
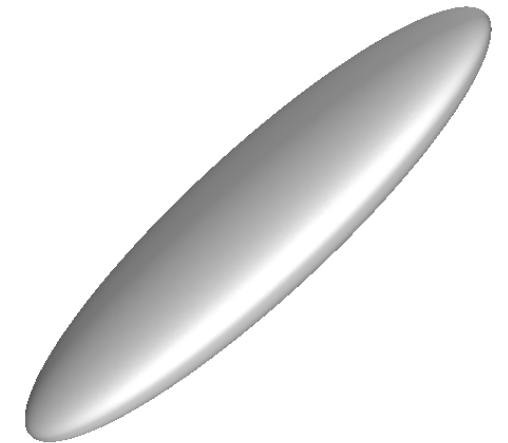
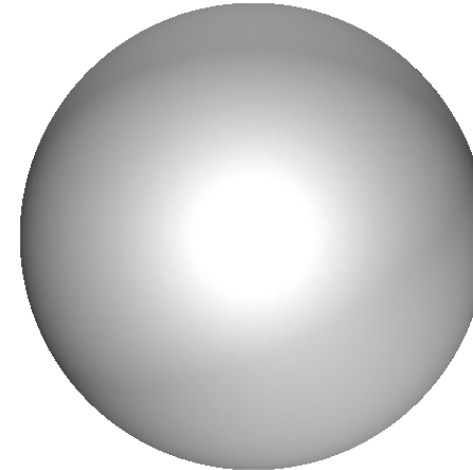
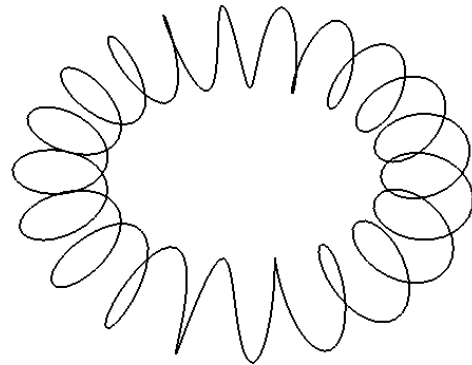
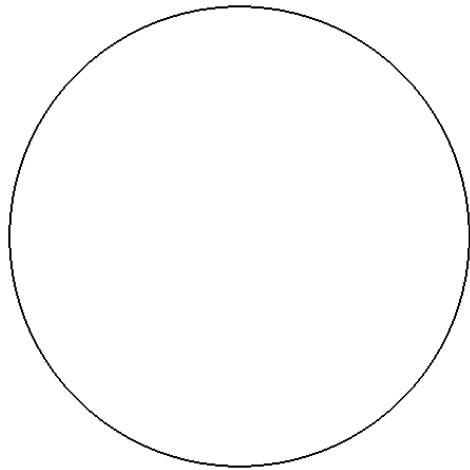
Some examples

Descriptors of some elementary shapes (approximated from finite samples):



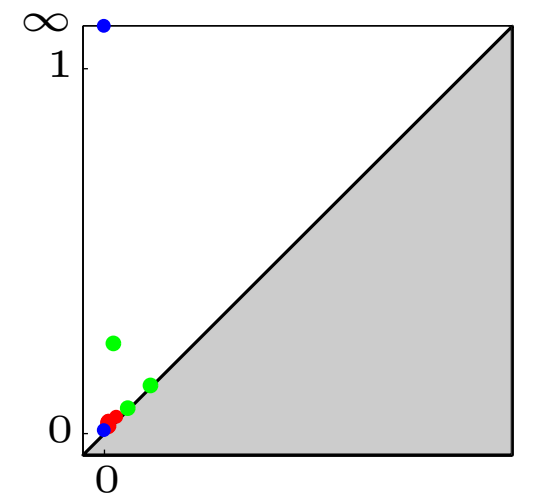
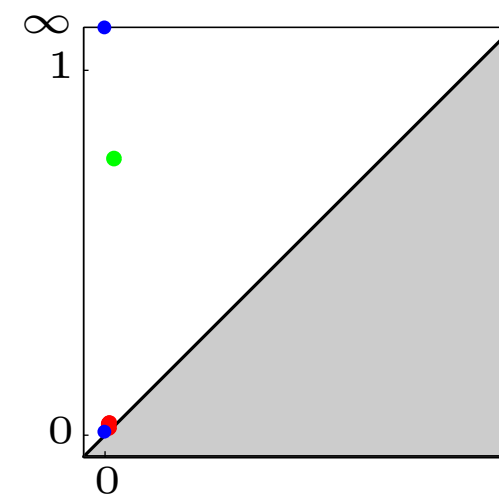
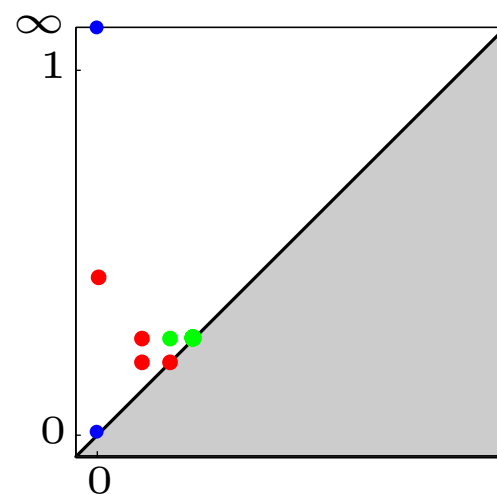
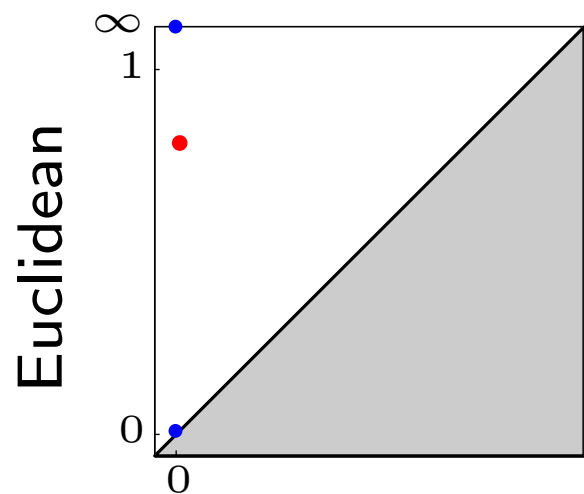
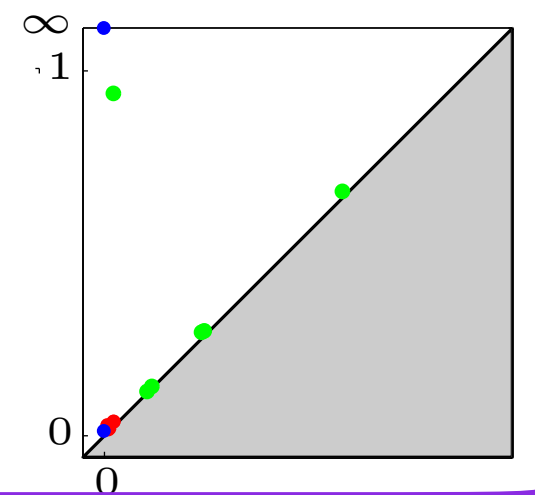
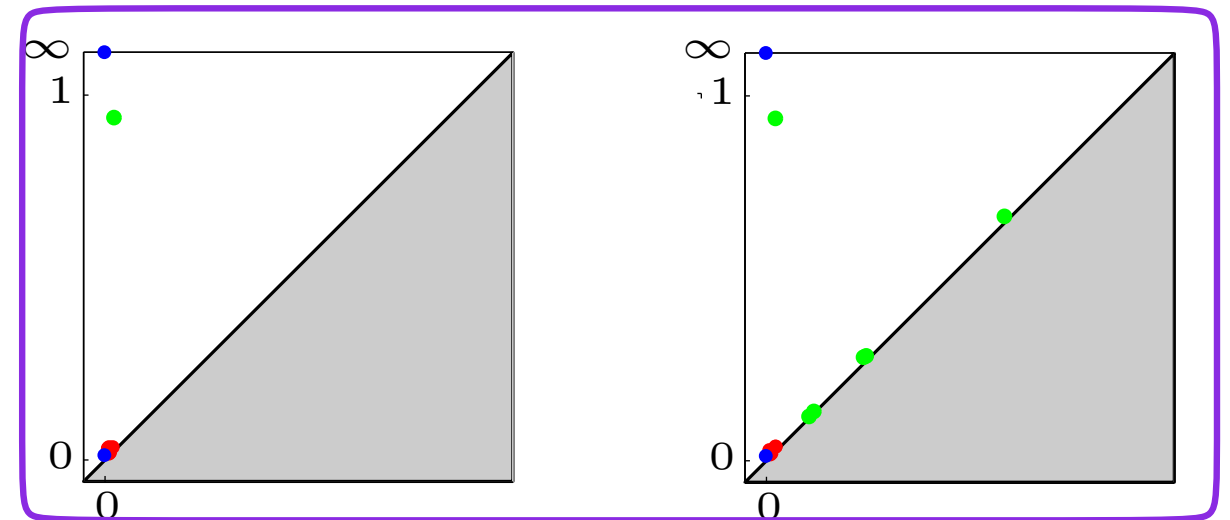
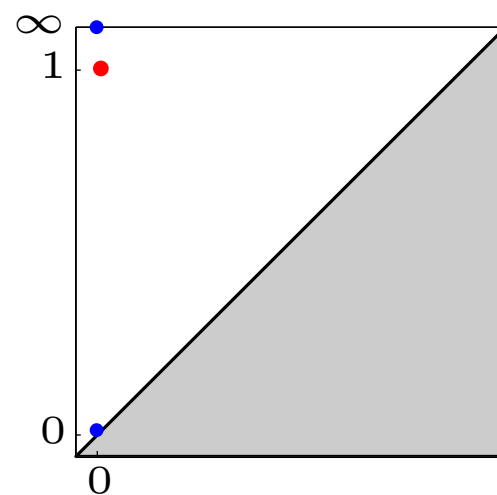
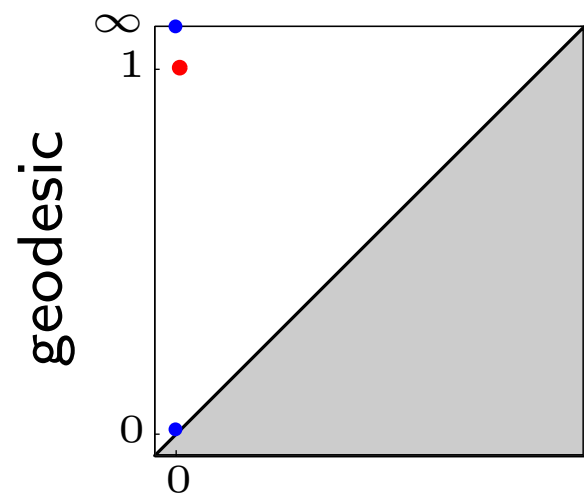
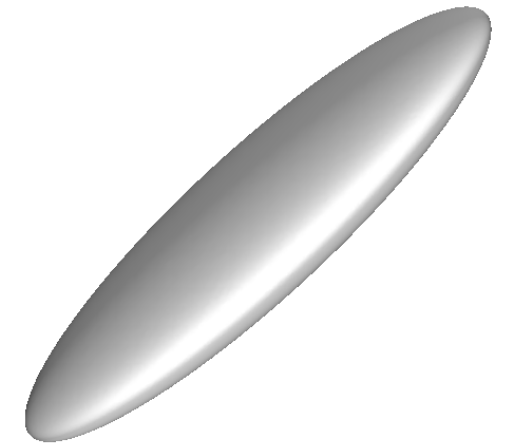
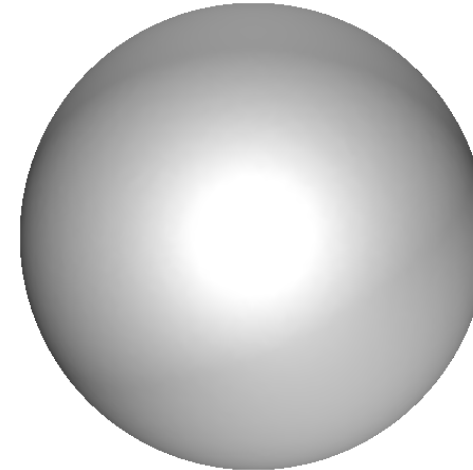
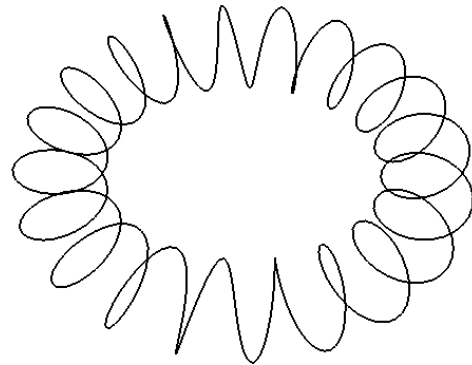
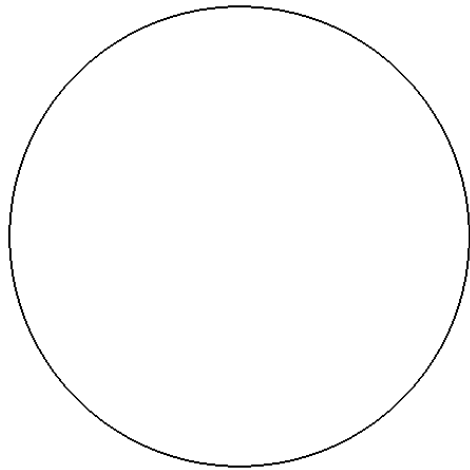
Some examples

Descriptors of some elementary shapes (approximated from finite samples):



Some examples

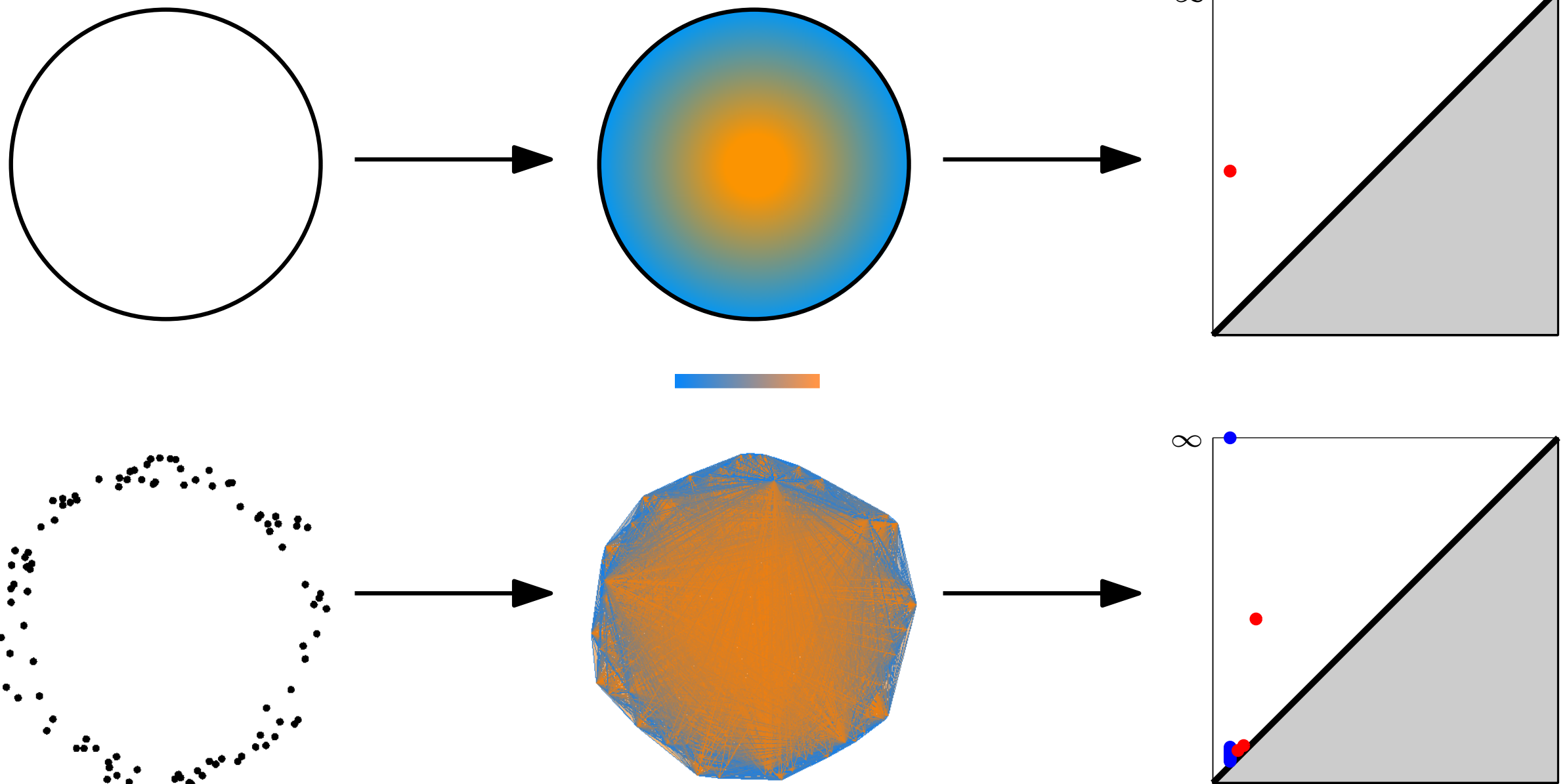
Descriptors of some elementary shapes (approximated from finite samples):



Stability

Theorem: [Chazal, de Silva, O. 2013]

For any compact metric spaces (X, d_X) and (Y, d_Y) ,
 $d_B^\infty(\text{dgm } \mathcal{R}(X, d_X), \text{dgm } \mathcal{R}(Y, d_Y)) \leq 2d_{\text{GH}}(X, Y)$.

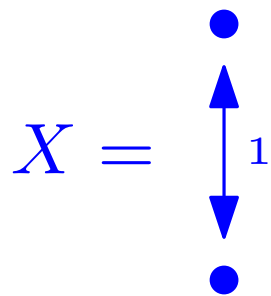


Stability

Theorem: [Chazal, de Silva, O. 2013]

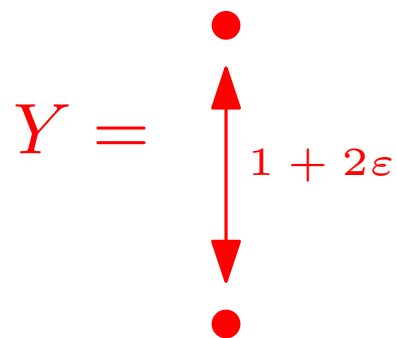
For any compact metric spaces (X, d_X) and (Y, d_Y) ,
 $d_B^\infty(\text{dgm } \mathcal{R}(X, d_X), \text{dgm } \mathcal{R}(Y, d_Y)) \leq 2d_{\text{GH}}(X, Y)$.

The bound is worst-case tight...



$$d_{\text{GH}}(X, Y) = \varepsilon$$

$$\text{dgm } \mathcal{R}(X, d_X) = \{(0, \infty), (0, 1)\}$$



$$\text{dgm } \mathcal{R}(Y, d_Y) = \{(0, \infty), (0, 1 + 2\varepsilon)\}$$

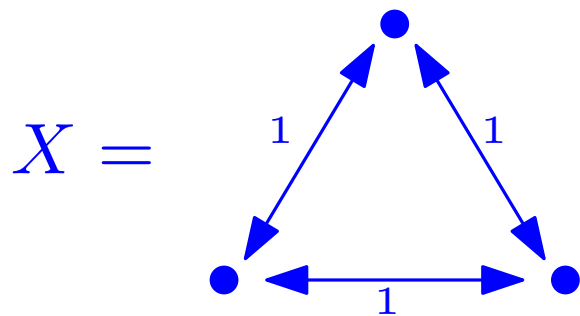
$$\Rightarrow d_B^\infty(\text{dgm } \mathcal{R}(X, d_X), \text{dgm } \mathcal{R}(Y, d_Y)) = 2\varepsilon$$

Stability

Theorem: [Chazal, de Silva, O. 2013]

For any compact metric spaces (X, d_X) and (Y, d_Y) ,
 $d_B^\infty(\text{dgm } \mathcal{R}(X, d_X), \text{dgm } \mathcal{R}(Y, d_Y)) \leq 2d_{\text{GH}}(X, Y)$.

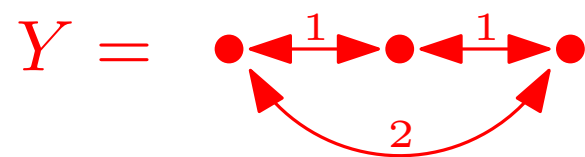
The bound is worst-case tight... but it is still only an upper bound



$$d_{\text{GH}}(X, Y) = \frac{1}{2}$$

$$\text{dgm } \mathcal{R}(X, d_X) = \{(0, \infty), (0, 1), (0, 1)\}$$

$$\text{dgm } \mathcal{R}(Y, d_Y) = \{(0, \infty), (0, 1), (0, 1)\}$$



$$\Rightarrow d_B^\infty(\text{dgm } \mathcal{R}(X, d_X), \text{dgm } \mathcal{R}(Y, d_Y)) = 0$$

Stability

Theorem: [Chazal, de Silva, O. 2013]

For any compact metric spaces (X, d_X) and (Y, d_Y) ,
 $d_B^\infty(\text{dgm } \mathcal{R}(X, d_X), \text{dgm } \mathcal{R}(Y, d_Y)) \leq 2d_{GH}(X, Y).$

Variants and extensions:

- other filtrations: Čech / Nerve, witness complex, etc.
- larger classes of metric spaces: precompact, totally bounded, etc.
- (dis-)similarity measures

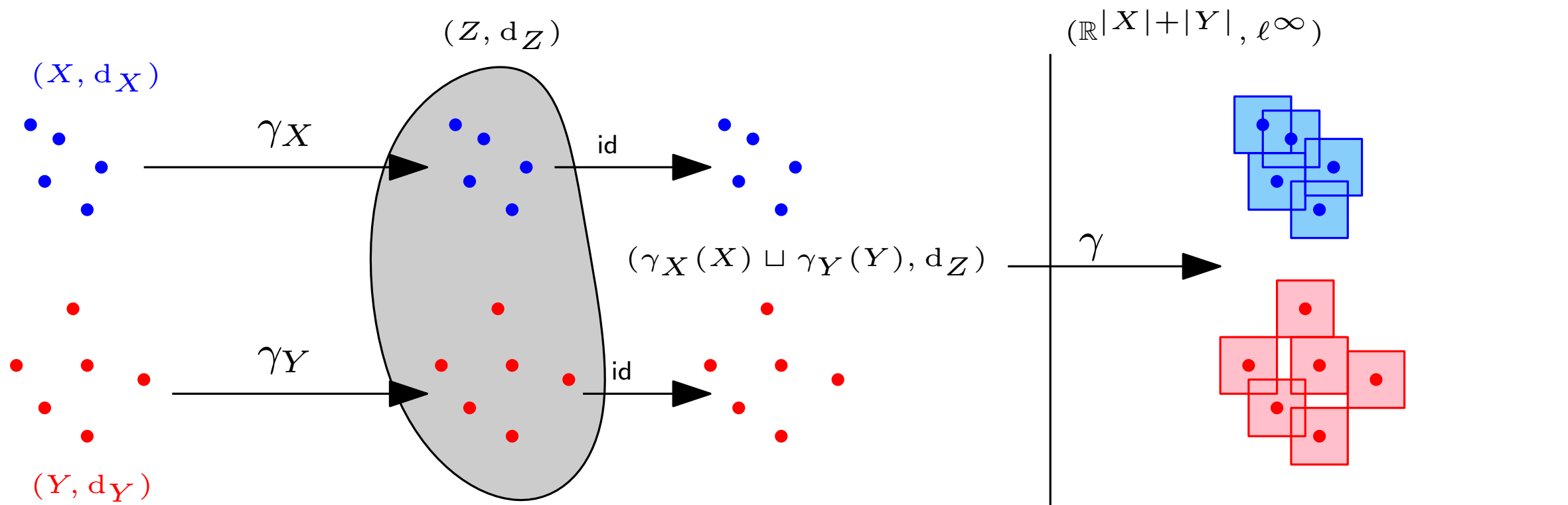
Stability

finite

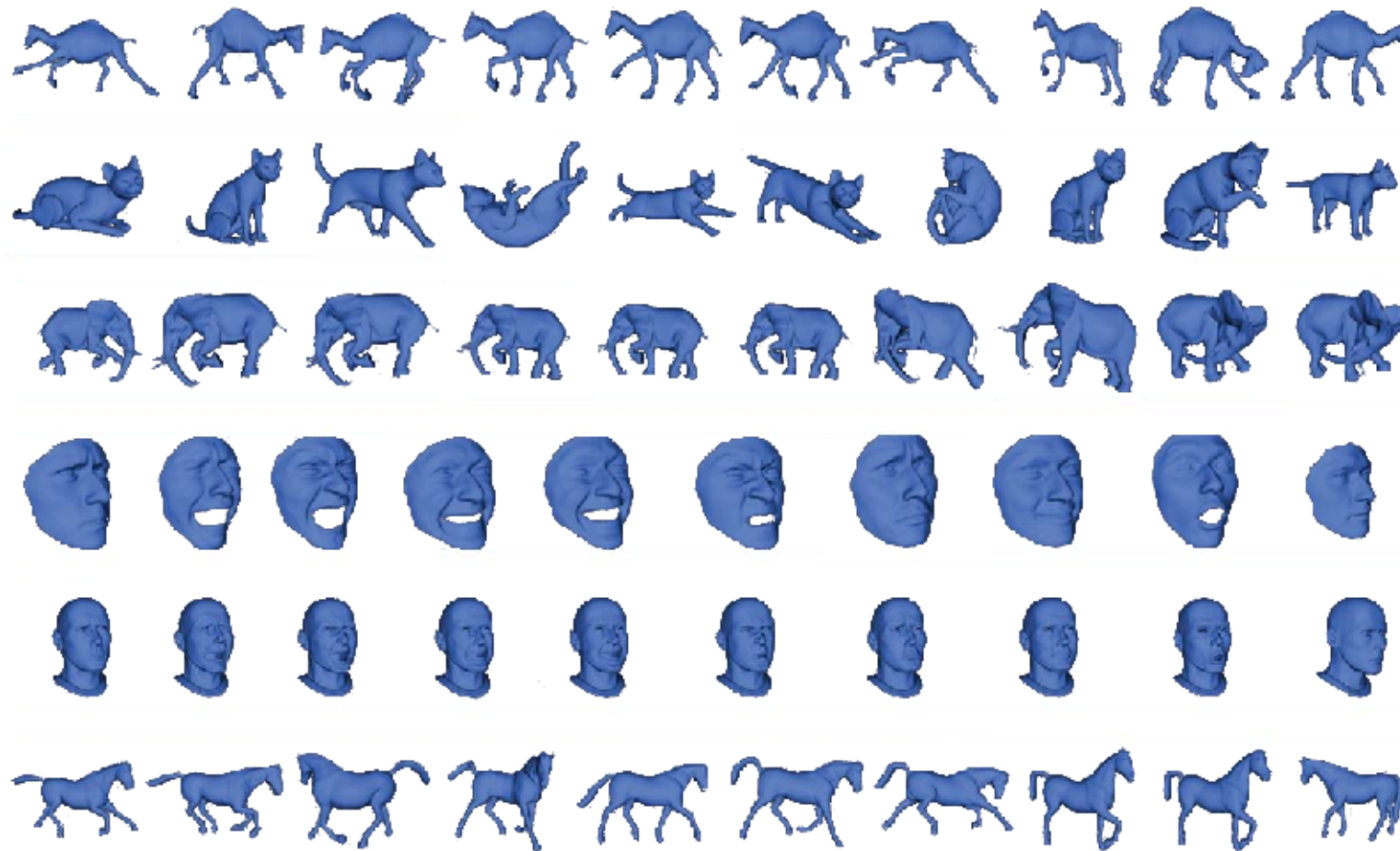
Theorem: [Chazal, de Silva, O. 2013]

For any ~~compact~~ metric spaces (X, d_X) and (Y, d_Y) ,
 $d_B^\infty(\text{dgm } \mathcal{R}(X, d_X), \text{dgm } \mathcal{R}(Y, d_Y)) \leq 2d_{\text{GH}}(X, Y).$

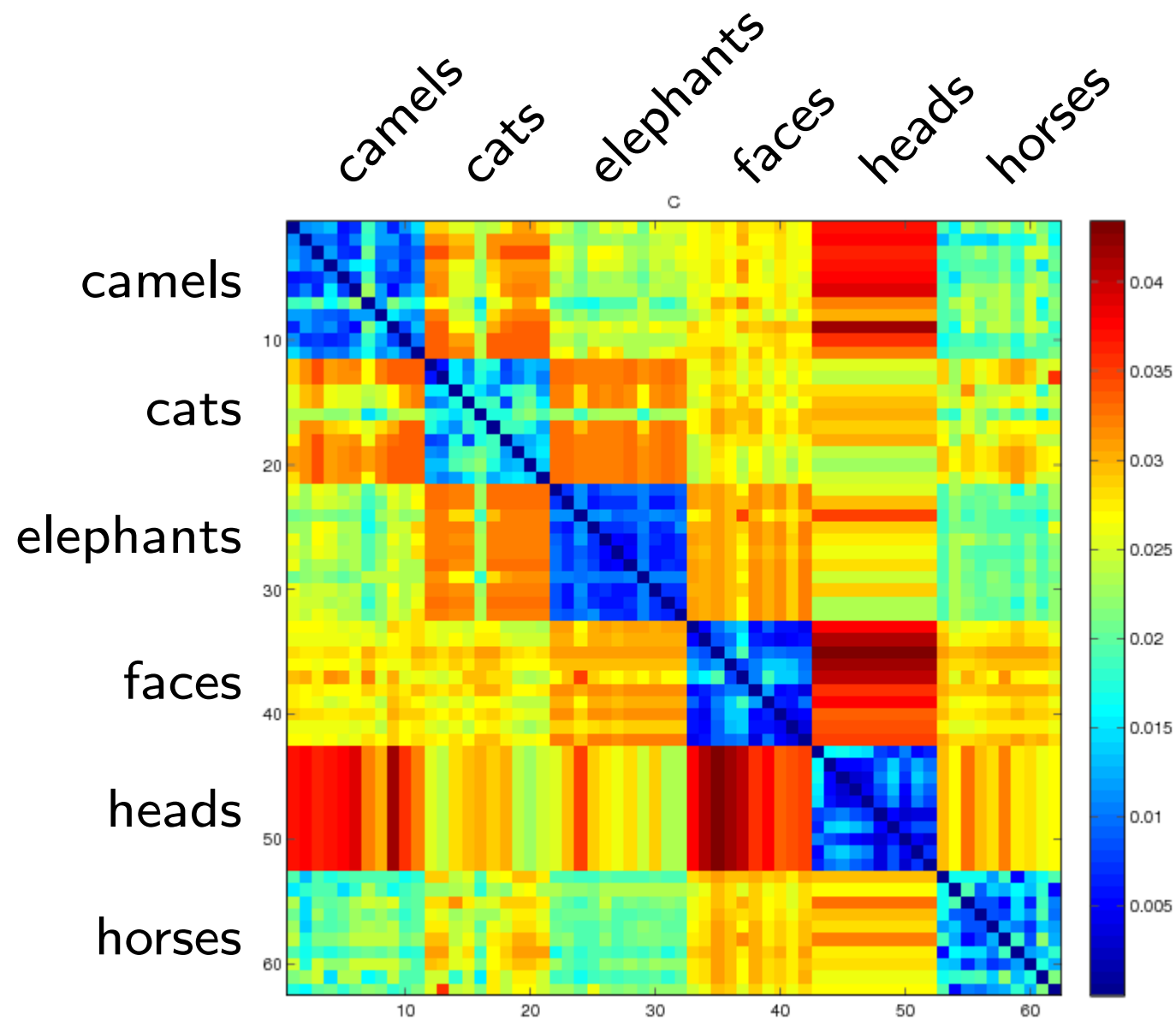
Proof outline:



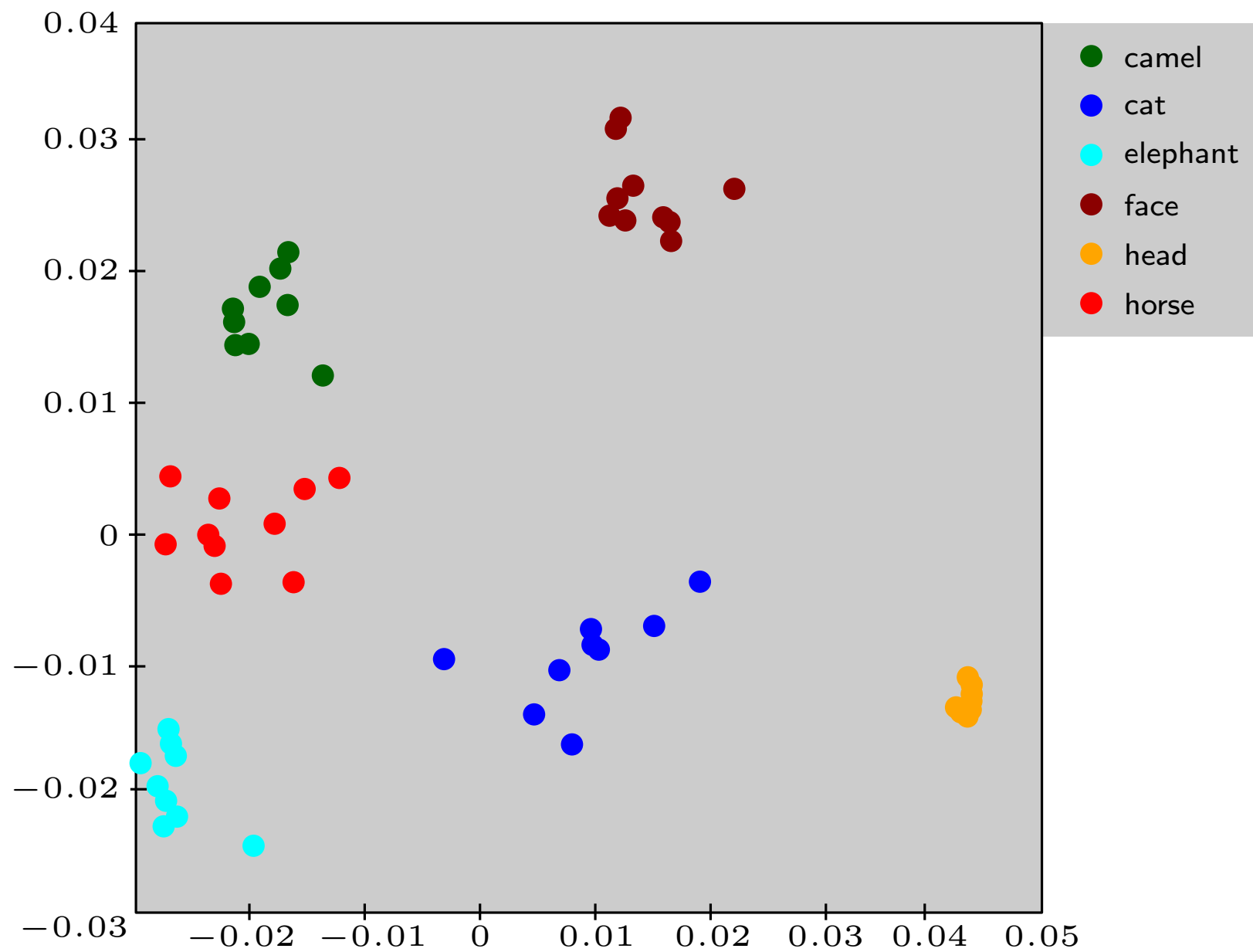
Toy application (unsupervised shape classification)



Toy application (unsupervised shape classification)



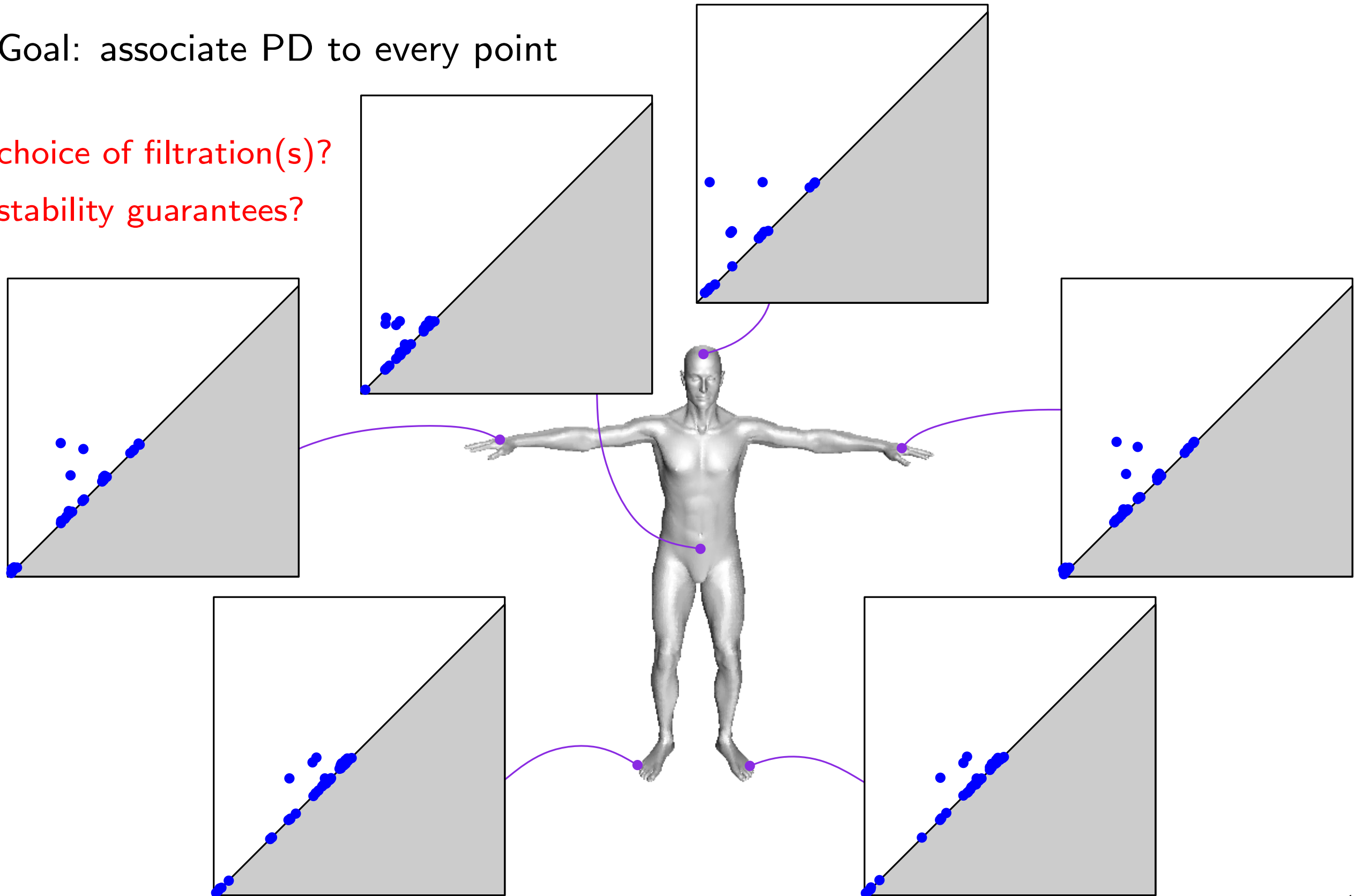
Toy application (unsupervised shape classification)



Local topological descriptors

Goal: associate PD to every point

choice of filtration(s)?
stability guarantees?



Local topological descriptors

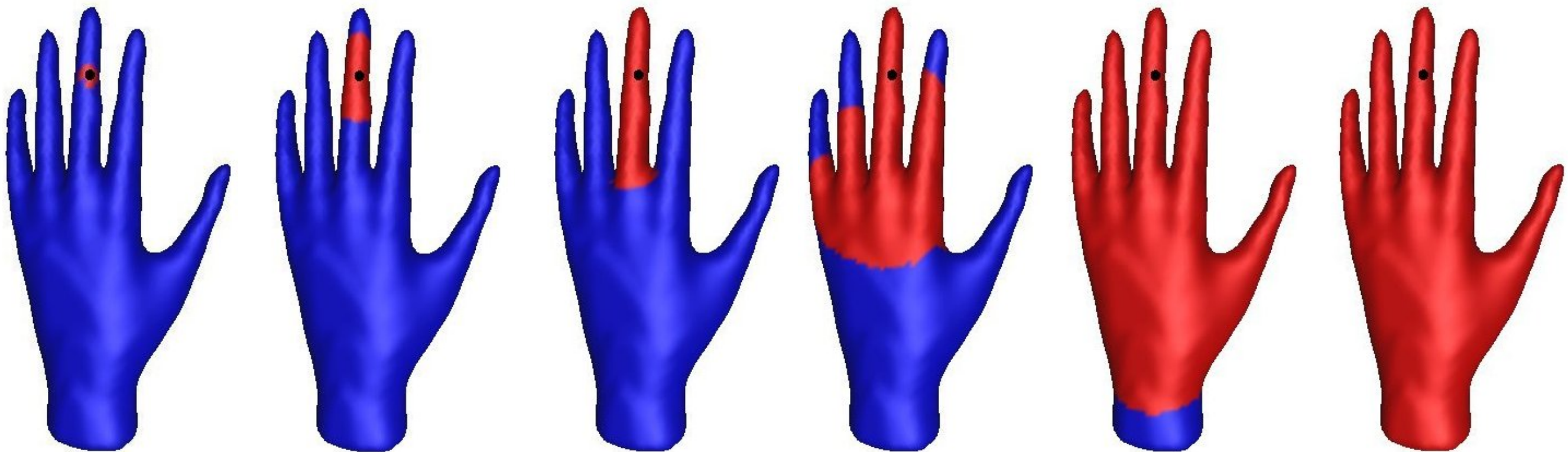
Input: a compact *intrinsic* metric space (X, d_X) , a basepoint $x_0 \in X$

Construction: filtration of the sublevel sets of $d_{x_0}(\cdot) = d_X(x_0, \cdot)$

Descriptor: persistence diagram of the filtration, denoted $\text{dgm } d_{x_0}$

In practice: compute descriptor from point cloud using a pair of Rips complexes

[Chazal et al. 2009]



Stability

Theorem: [Carrière, O., Ovsjanikov 2015]

Let (X, d_X) and (Y, d_Y) be compact intrinsic metric spaces with positive *convexity radius* ($\varrho(X), \varrho(Y) > 0$). Let $x_0 \in X$ and $y_0 \in Y$. If $d_{\text{GH}}((X, x_0), (Y, y_0)) \leq \frac{1}{20} \min\{\varrho(X), \varrho(Y)\}$, then $d_B^\infty(\text{dgm } d_{x_0}, \text{dgm } d_{y_0}) \leq 20 d_{\text{GH}}((X, x_0), (Y, y_0))$.

Stability

Theorem: [Carrière, O., Ovsjanikov 2015]

Let (X, d_X) and (Y, d_Y) be compact intrinsic metric spaces with positive *convexity radius* ($\varrho(X), \varrho(Y) > 0$). Let $x_0 \in X$ and $y_0 \in Y$. If $d_{\text{GH}}((X, x_0), (Y, y_0)) \leq \frac{1}{20} \min\{\varrho(X), \varrho(Y)\}$, then $d_B^\infty(\text{dgm } d_{x_0}, \text{dgm } d_{y_0}) \leq 20 d_{\text{GH}}((X, x_0), (Y, y_0))$.

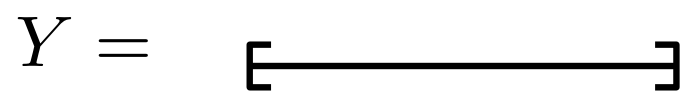
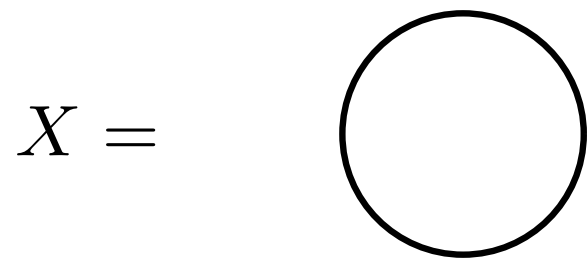
Definitions:

- convexity radius
- Gromov-Hausdorff distance between pointed spaces

Stability

Theorem: [Carrière, O., Ovsjanikov 2015]

Let (X, d_X) and (Y, d_Y) be compact intrinsic metric spaces with positive *convexity radius* ($\varrho(X), \varrho(Y) > 0$). Let $x_0 \in X$ and $y_0 \in Y$. If $d_{\text{GH}}((X, x_0), (Y, y_0)) \leq \frac{1}{20} \min\{\varrho(X), \varrho(Y)\}$, then $d_{\text{B}}^{\infty}(\text{dgm } d_{x_0}, \text{dgm } d_{y_0}) \leq 20 d_{\text{GH}}((X, x_0), (Y, y_0))$.



Prerequisite: $d_{\text{GH}}(X, Y) < \frac{1}{20} \min\{\varrho(X), \varrho(Y)\}$

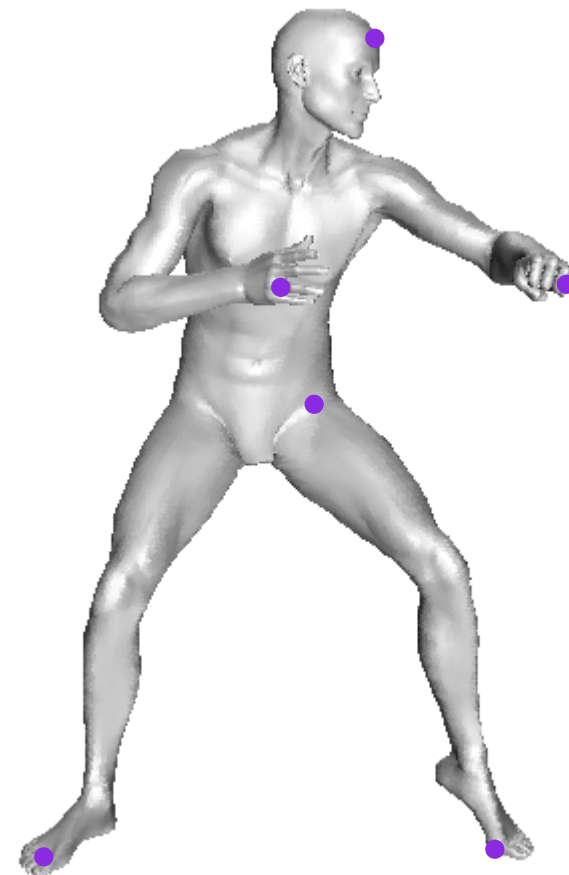
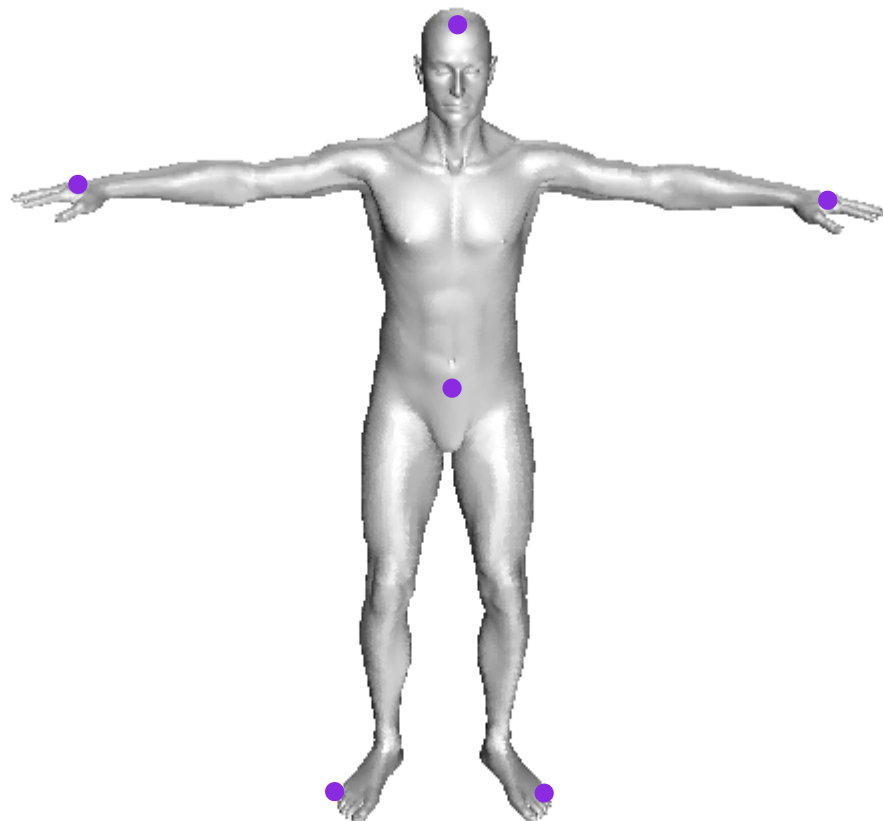
$$d_{\text{GH}}(X, Y) < \infty$$

$$d_{\text{B}}^{\infty}(\text{dgm } f, \text{dgm } g) = \infty$$

Toy application (unsupervised shape segmentation)

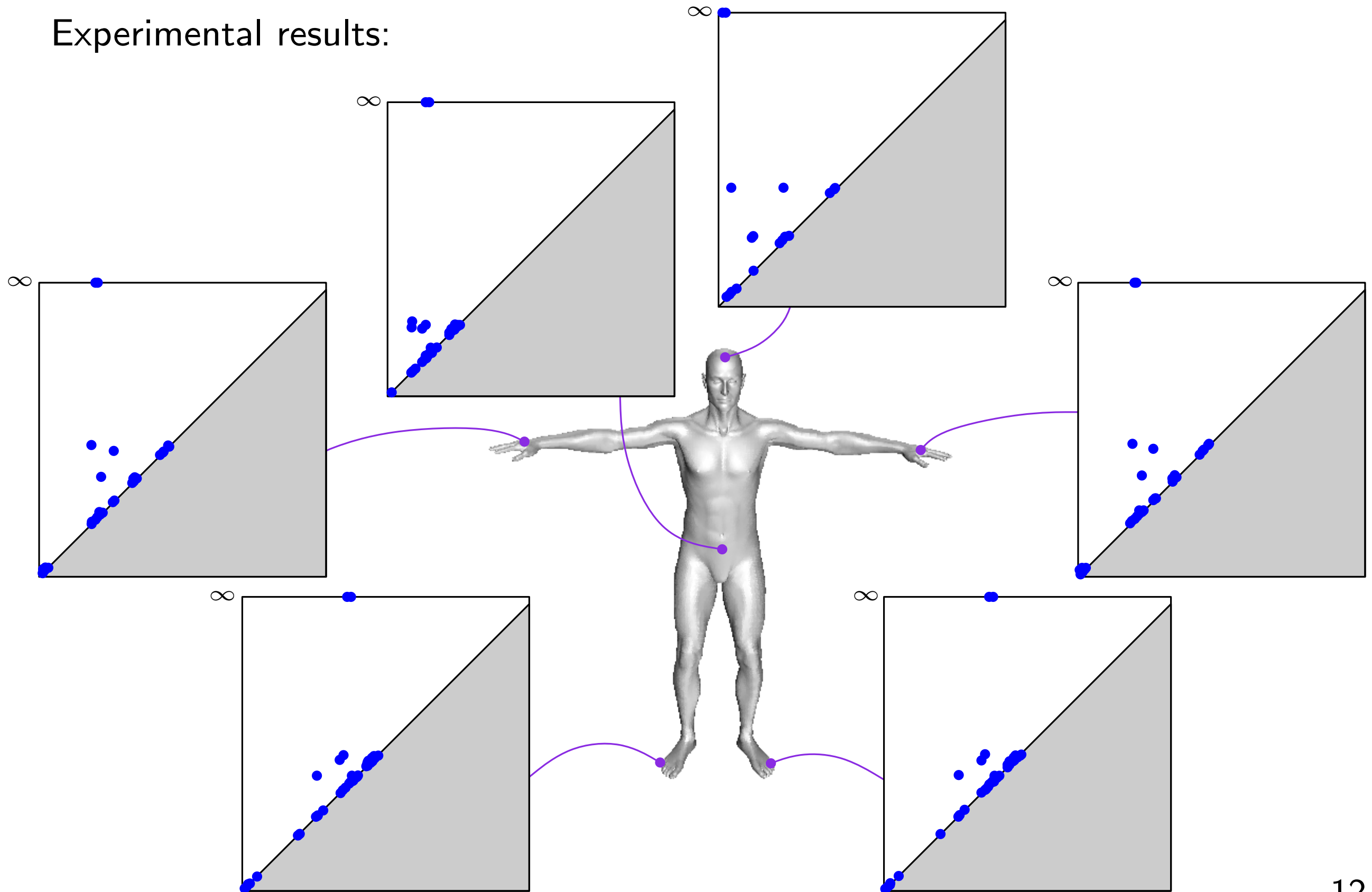
Experimental results:

- input: shapes from the TOSCA database, in *mesh* form (triangulated)
- select a few base points by hand on each shape
- approximate geodesic distances to base points using the 1-skeleton graph
- use the PDs of the PL interpolations over the meshes as descriptors



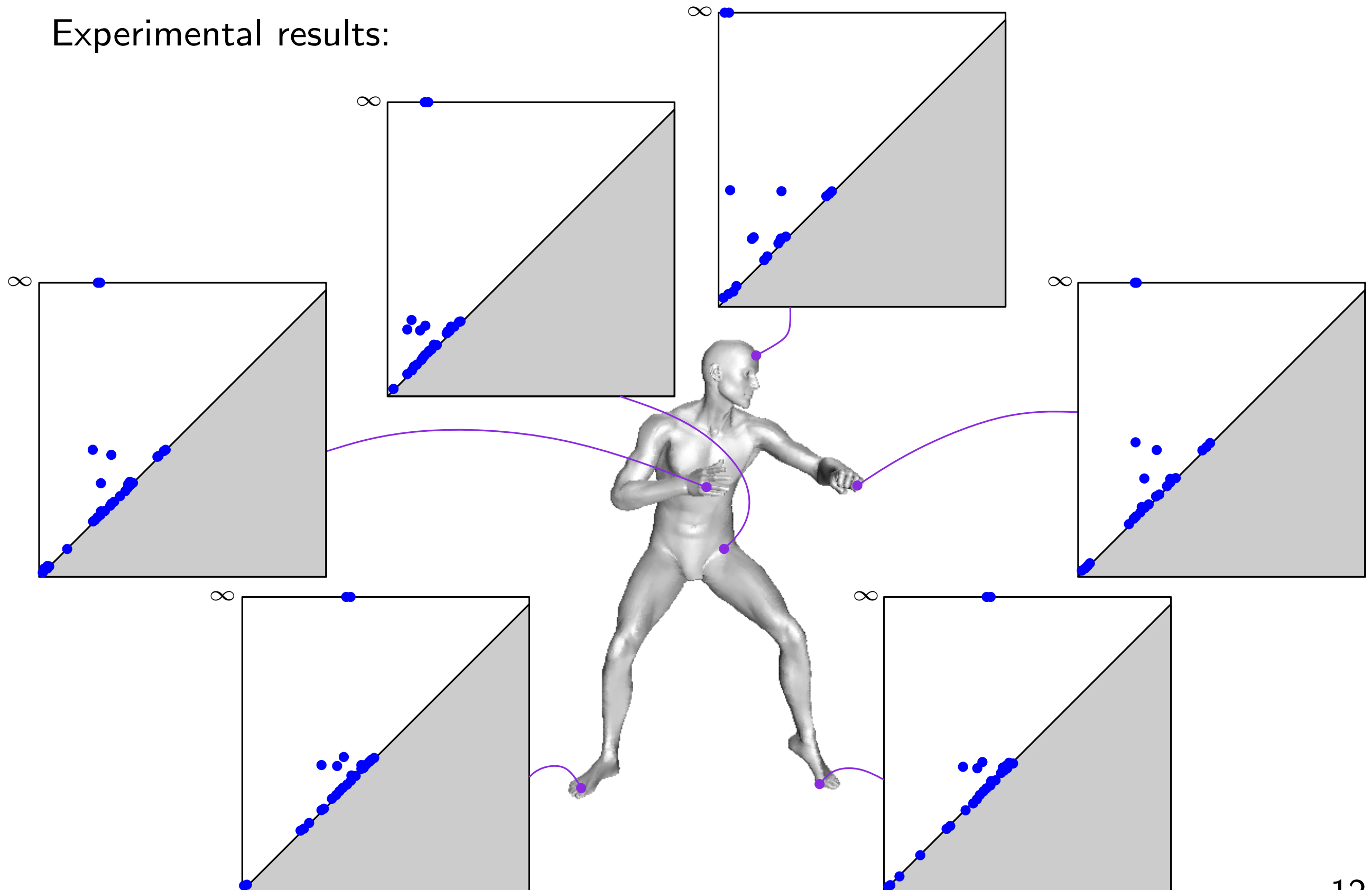
Toy application (unsupervised shape segmentation)

Experimental results:



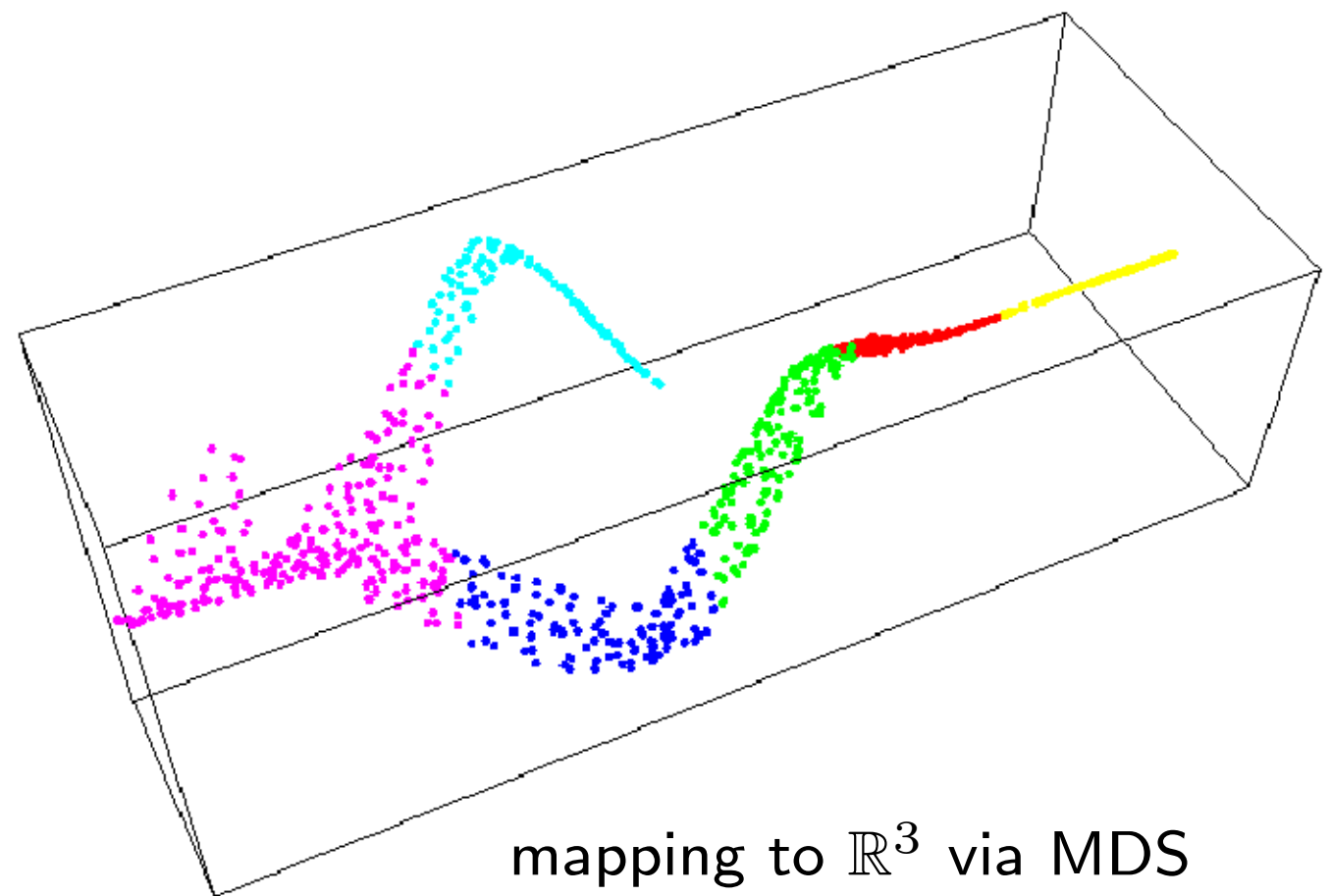
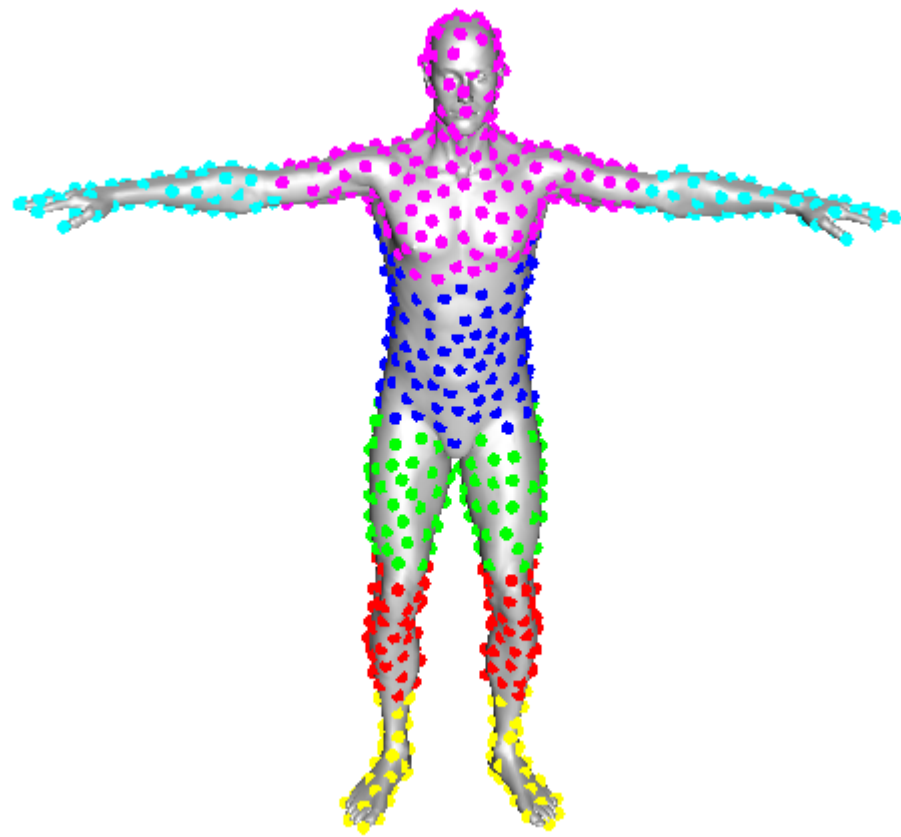
Toy application (unsupervised shape segmentation)

Experimental results:



Toy application (unsupervised shape segmentation)

Experimental results:

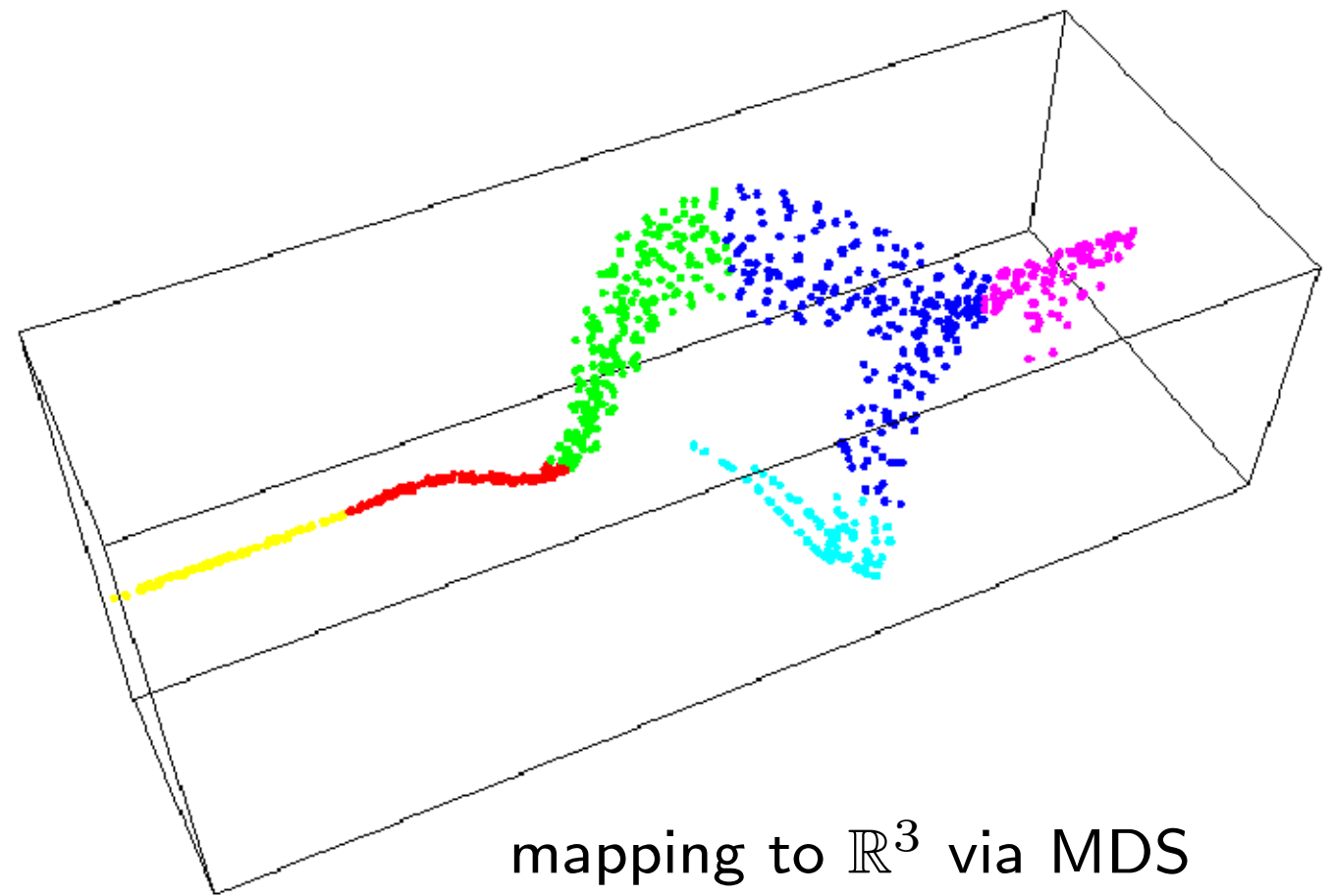
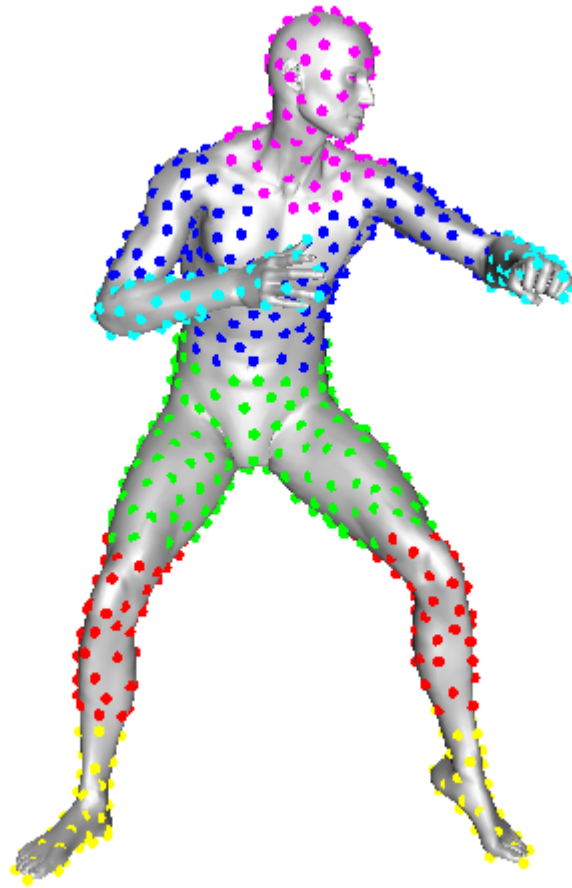


mapping to $\mathbb{R}^3$ via MDS

k -means in $\mathbb{R}^3$

Toy application (unsupervised shape segmentation)

Experimental results:



mapping to $\mathbb{R}^3$ via MDS

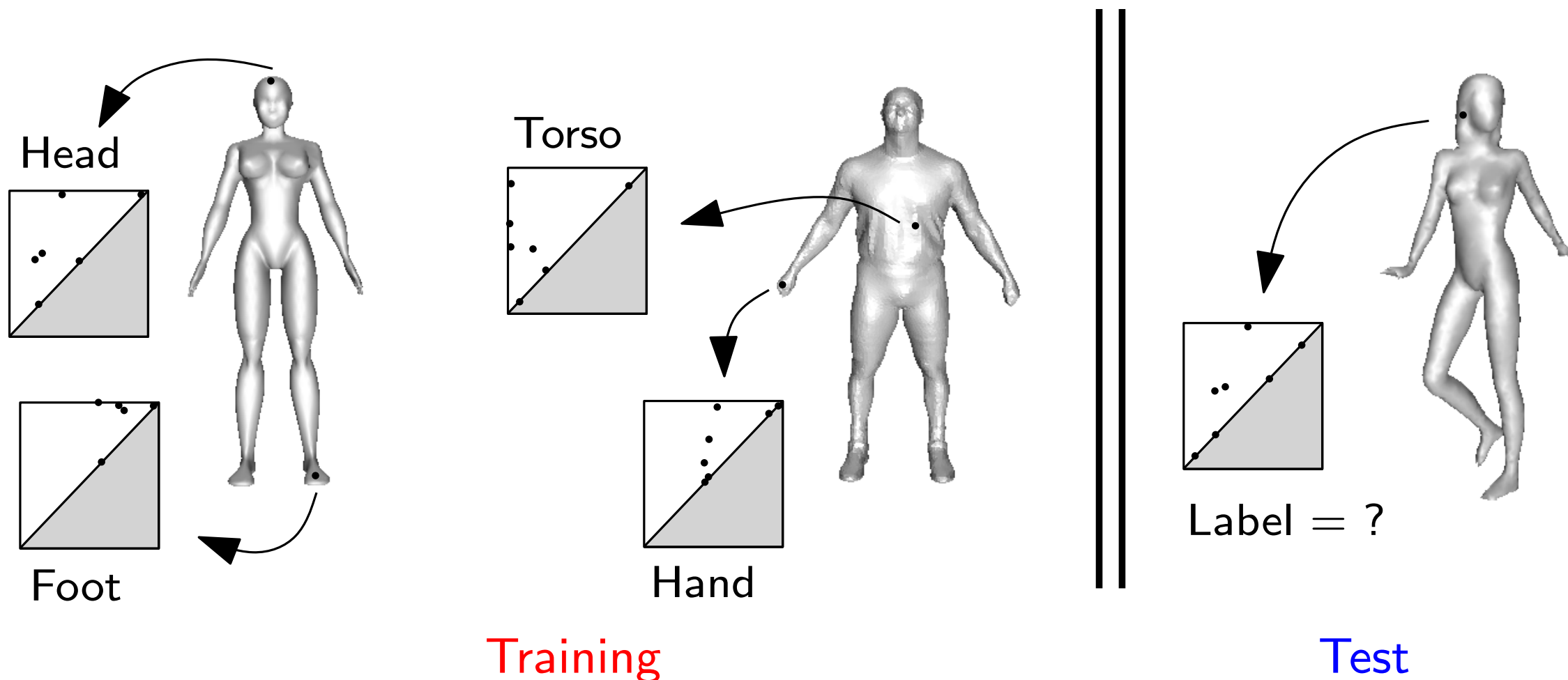
k -means in $\mathbb{R}^3$

Toy application (supervised shape segmentation)

Goal: segment 3d shapes based on examples

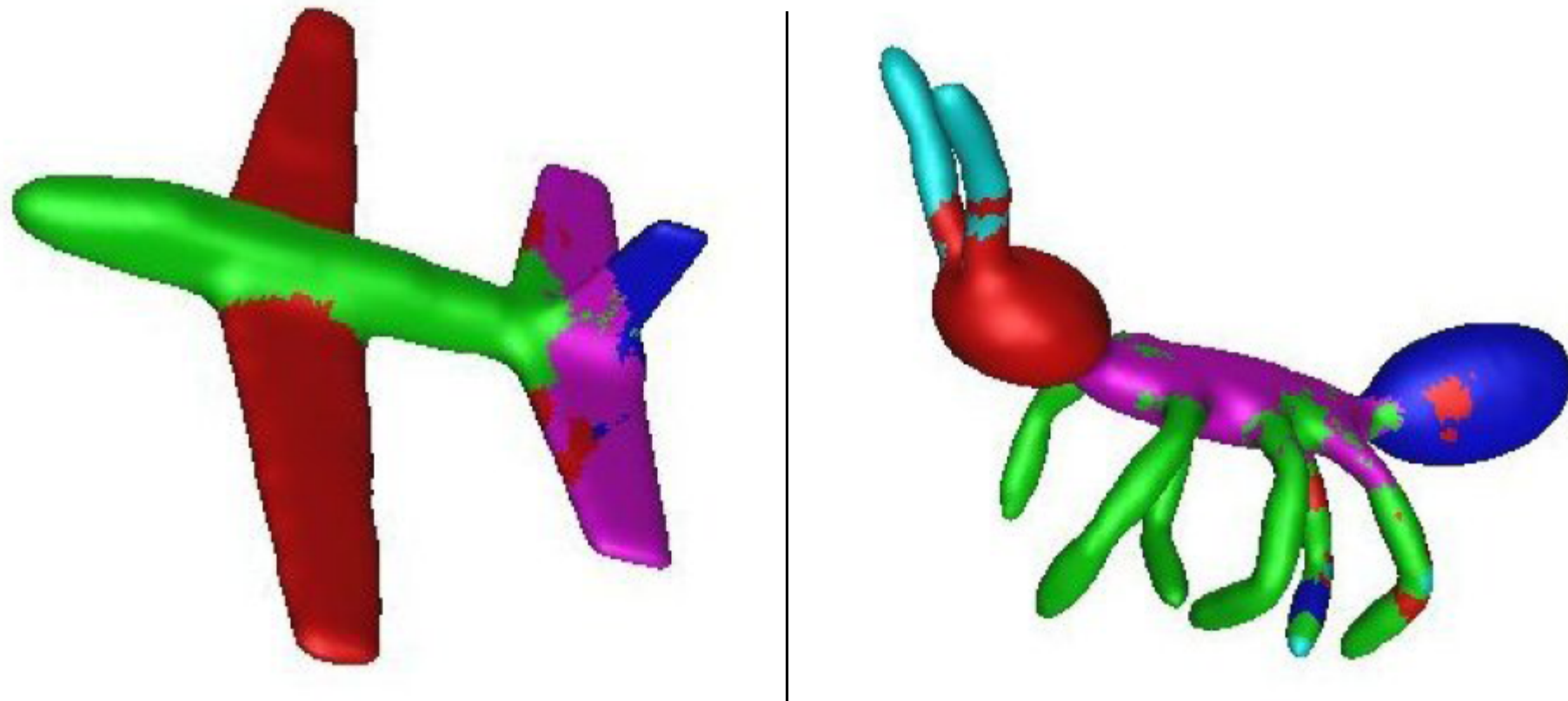
Approach:

- train a (multiclass) classifier on PDs extracted from the training shapes
- apply classifier to PDs extracted from each test shape



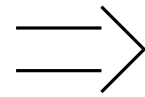
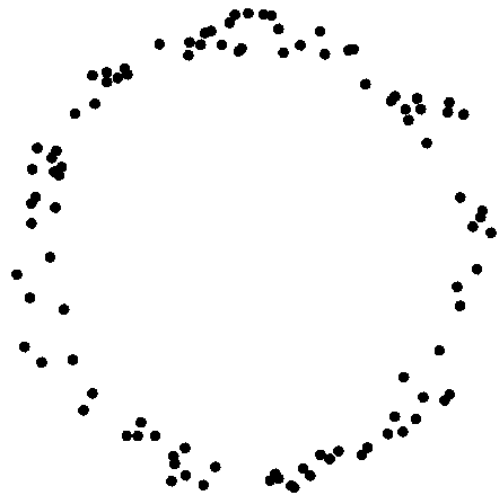
Toy application (supervised shape segmentation)

Strategy: use k-NN classifier in diagram space (equipped with d_B^∞)

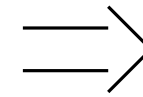
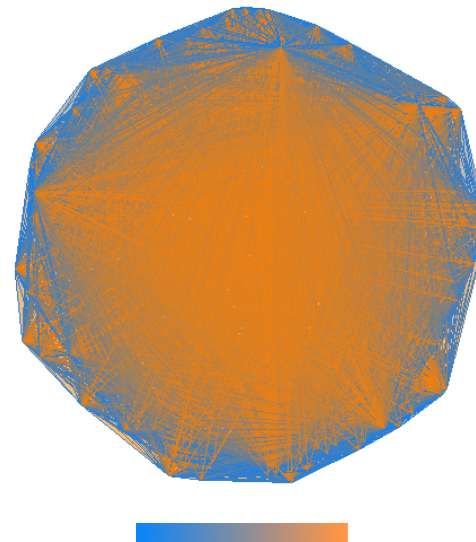


Recap'

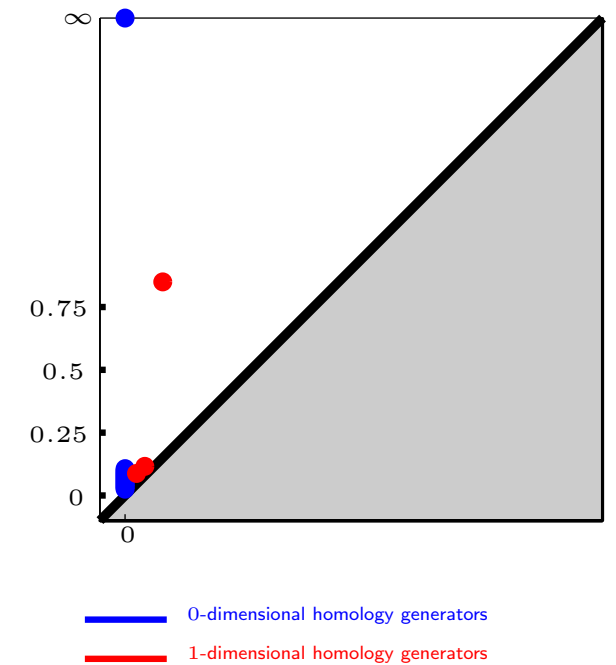
finite metric space / basepoint



filtration



persistence diagram

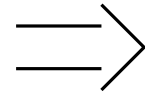
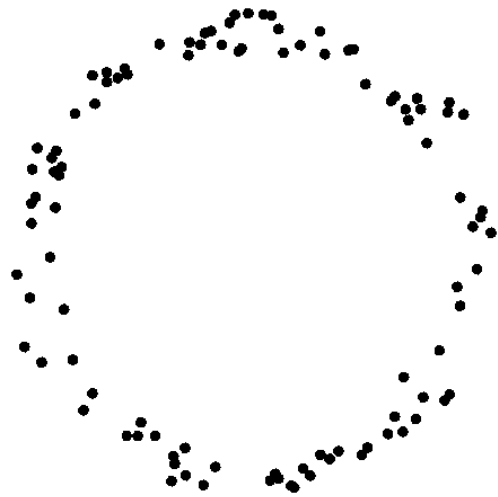


Pros:

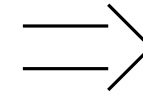
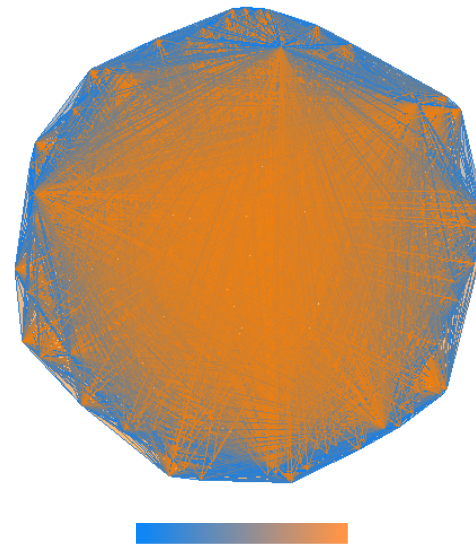
- topological descriptors carry information of a different nature
- they enjoy stability properties, e.g. $d_B^\infty(\mathcal{R}(X), \mathcal{R}(Y)) \leq 2d_{GH}(X, Y)$

Recap'

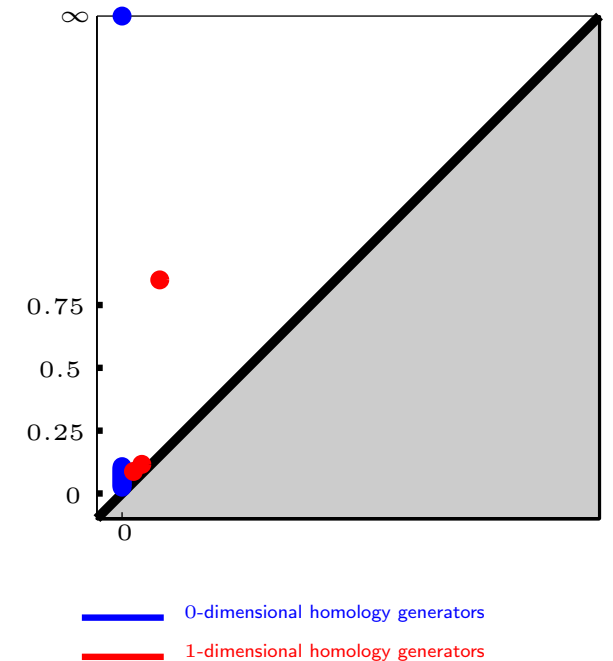
finite metric space / basepoint



filtration



persistence diagram



Pros:

- topological descriptors carry information of a different nature
- they enjoy stability properties, e.g. $d_B^\infty(\mathcal{R}(X), \mathcal{R}(Y)) \leq 2d_{GH}(X, Y)$

Cons:

- the space of persistence diagrams is not a vector/Hilbert space
→ bad for supervised learning and statistics
- descriptors can be slow to compute and (more importantly) to compare
→ bad for applications