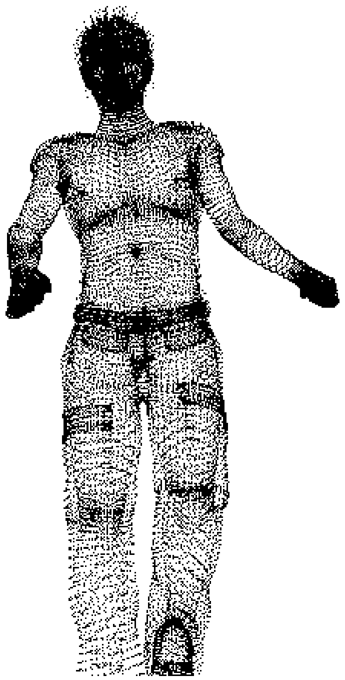
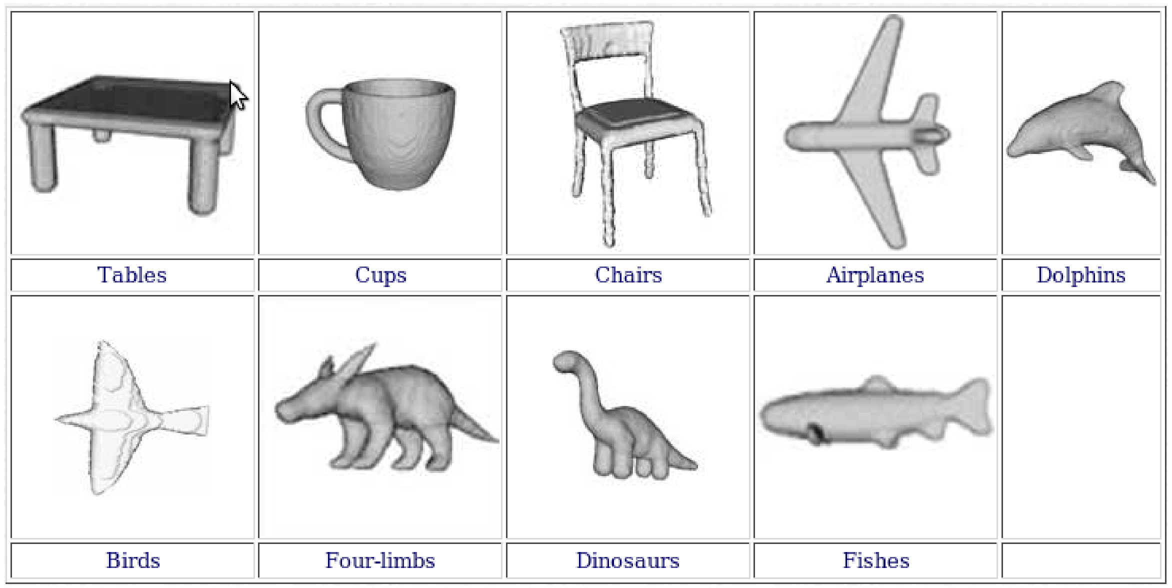
Topological Signatures

Geometric Data

Input: set of data points with metric or (dis-)similarity measure

data point $\equiv$ point on 3d shape, image patch, atom/site in protein, Facebook user, etc.

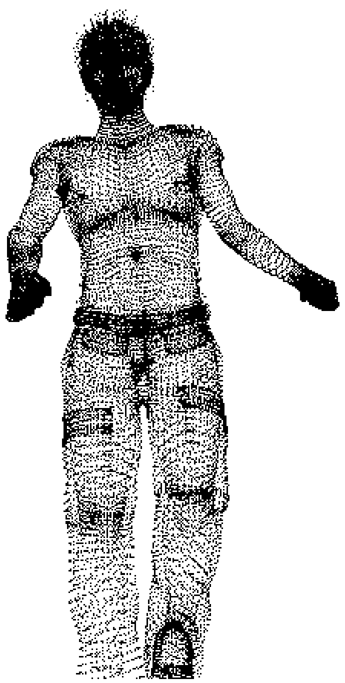


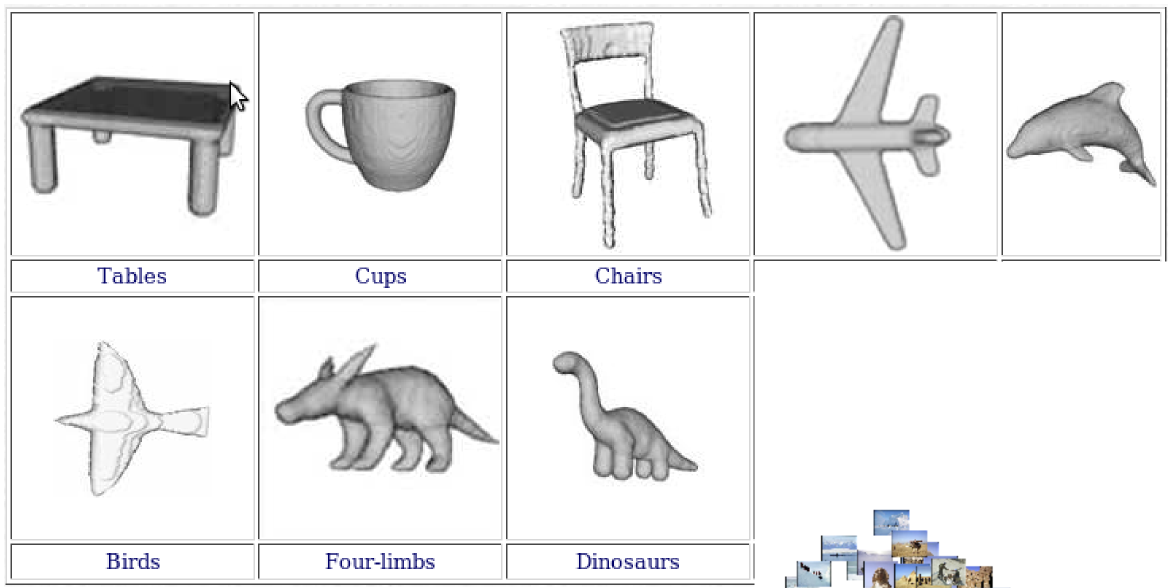


Geometric Data

Input: set of data points with metric or (dis-)similarity measure

data point $\equiv$ point on 3d shape, image patch, atom/site in protein, Facebook user, etc.

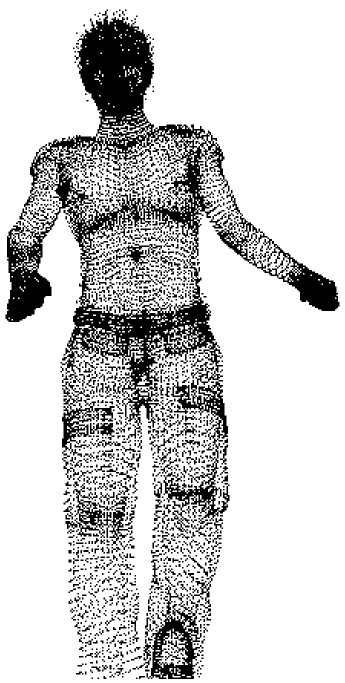


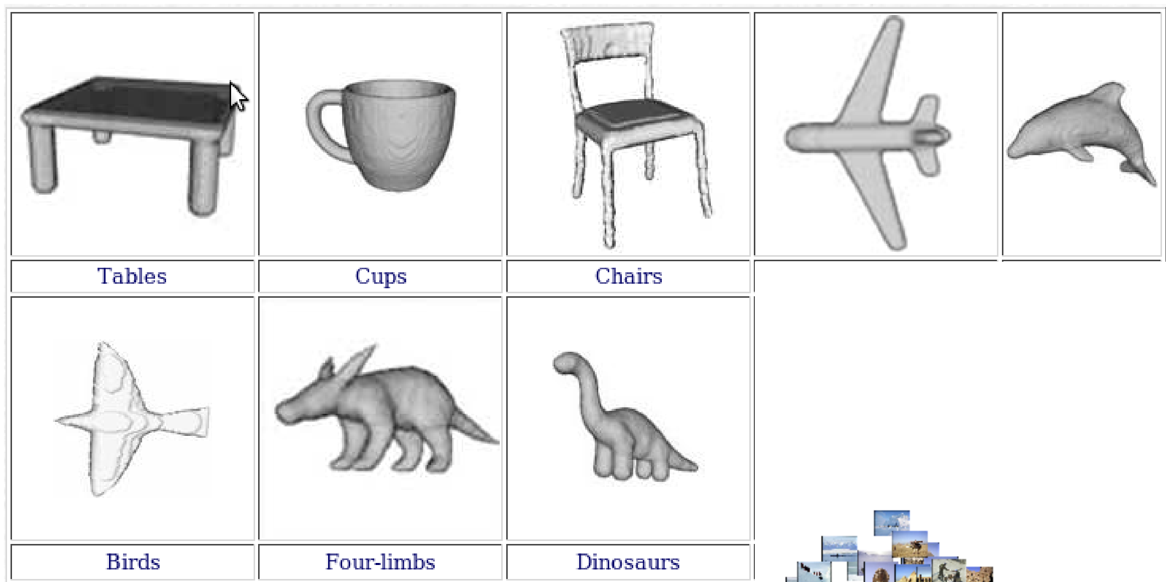


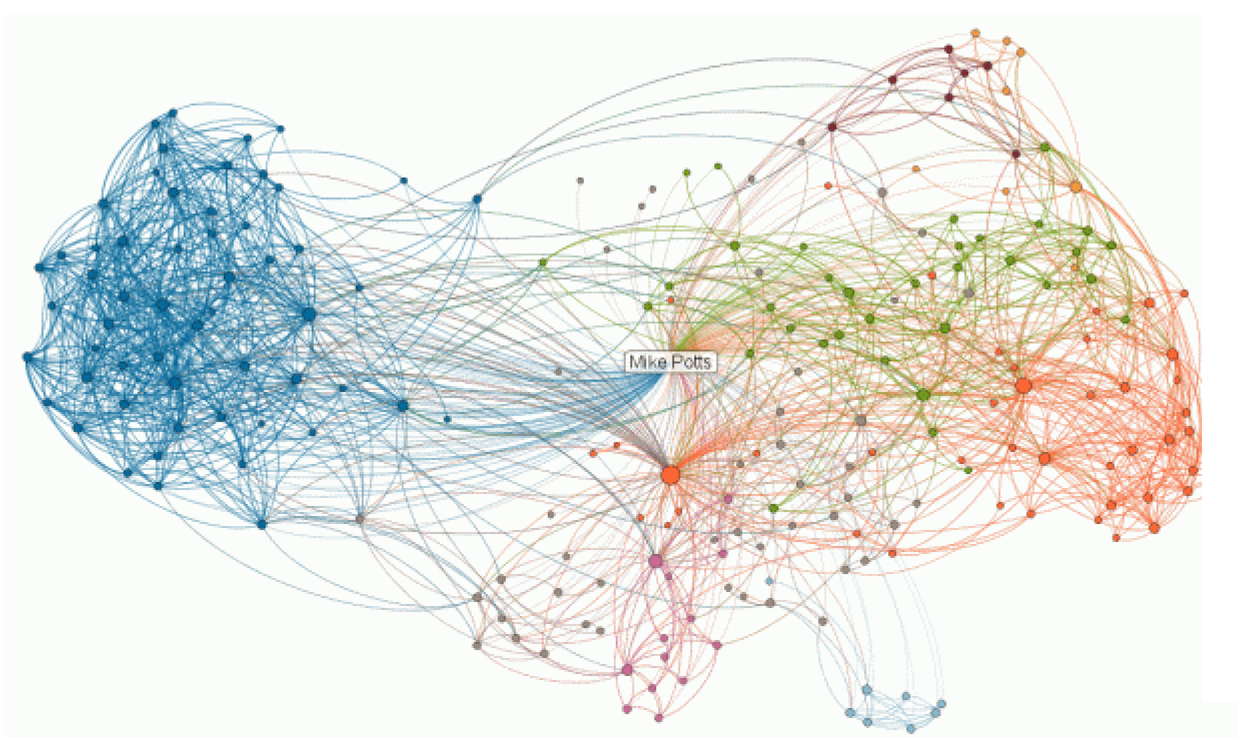
Geometric Data

Input: set of data points with metric or (dis-)similarity measure

data point $\equiv$ point on 3d shape, image patch, atom/site in protein, Facebook user, etc.



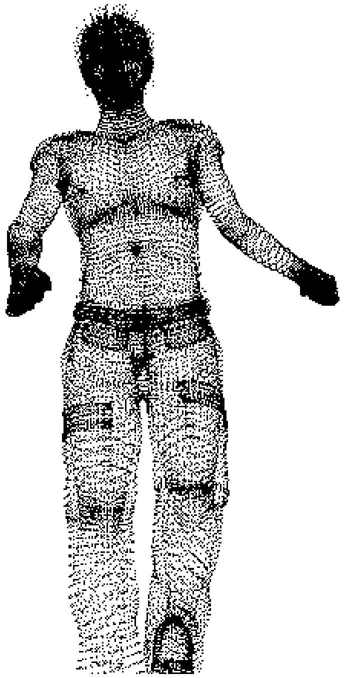


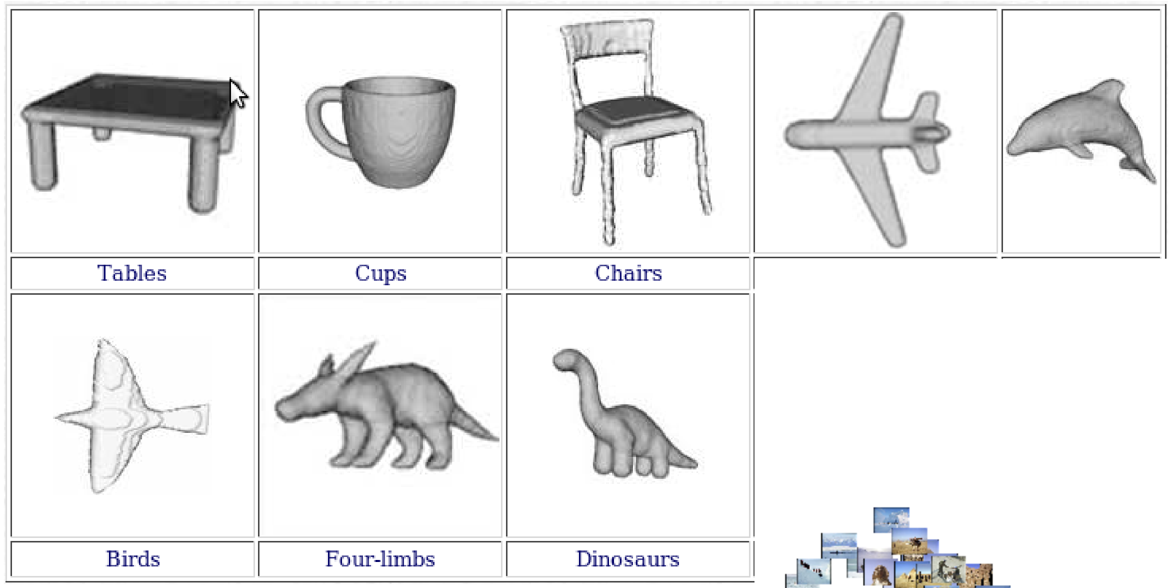


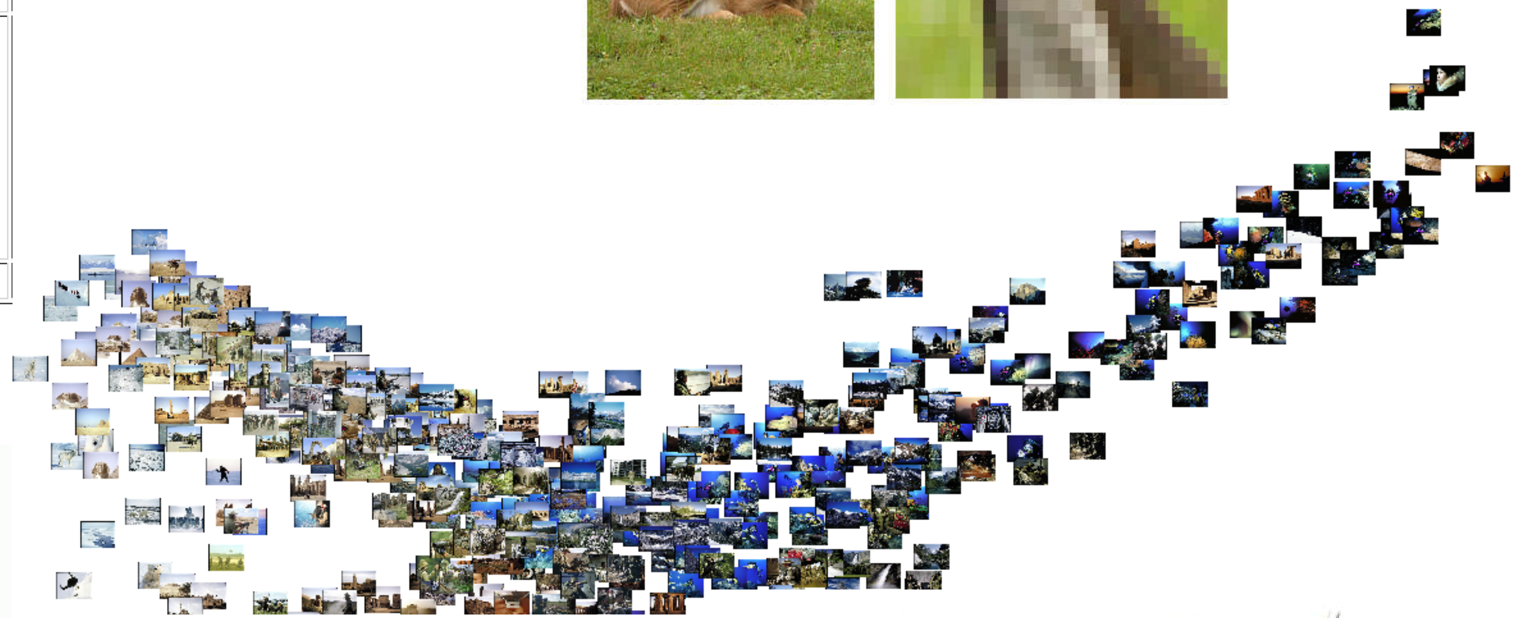
Geometric Data

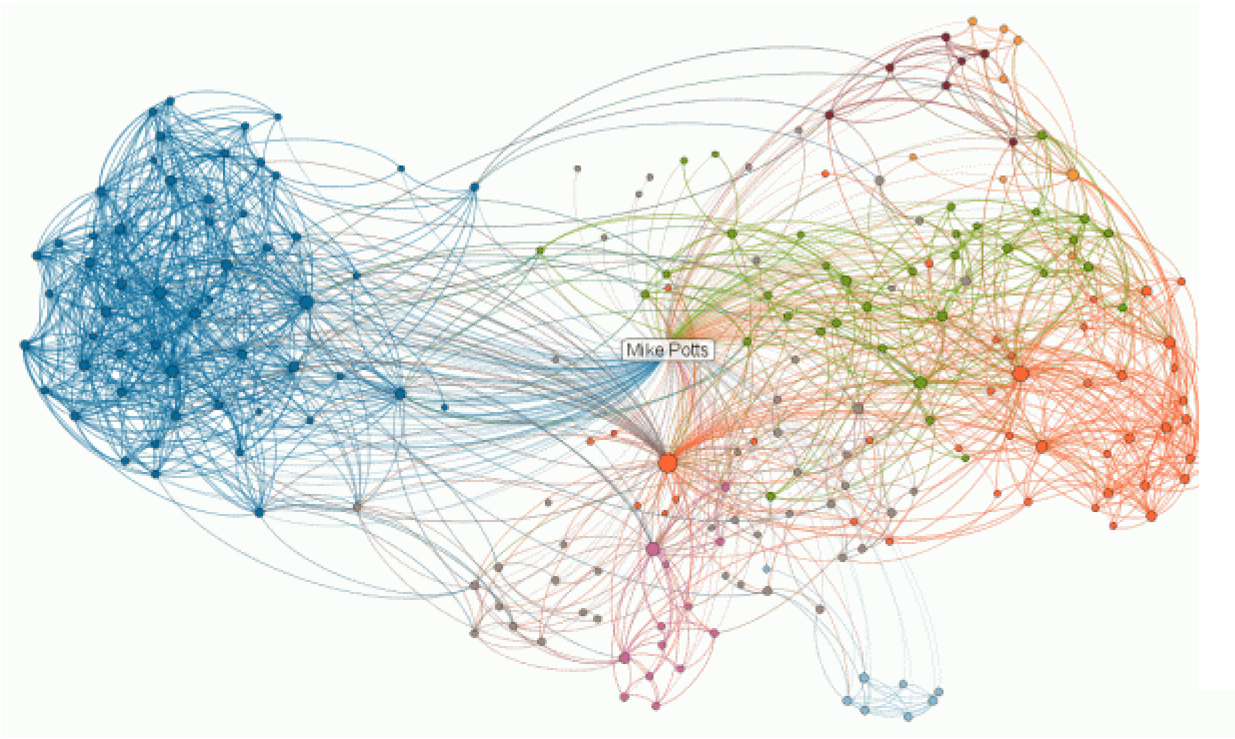
Input: set of data points with metric or (dis-)similarity measure

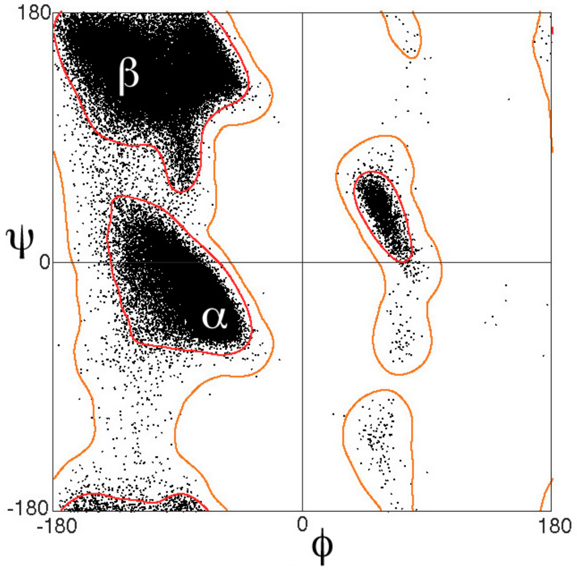
data point $\equiv$ point on 3d shape, image patch, atom/site in protein, Facebook user, etc.

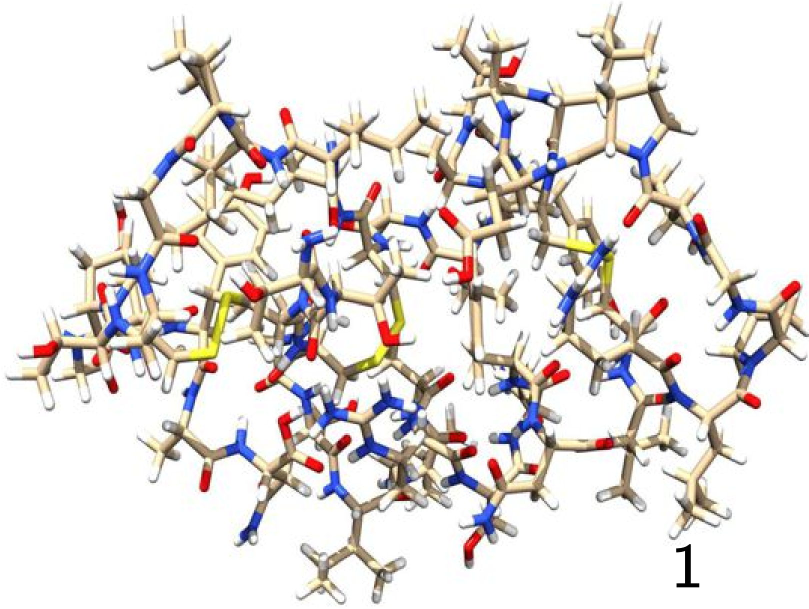










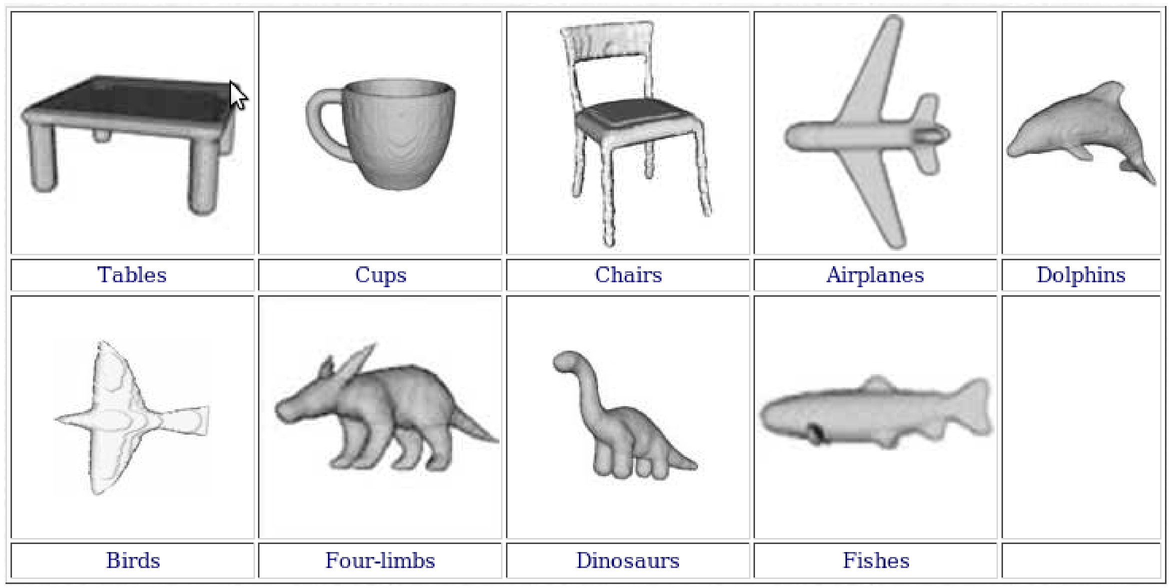


Why Compare Geometric Data

Comparisons between geometric data sets or parts thereof occur in:

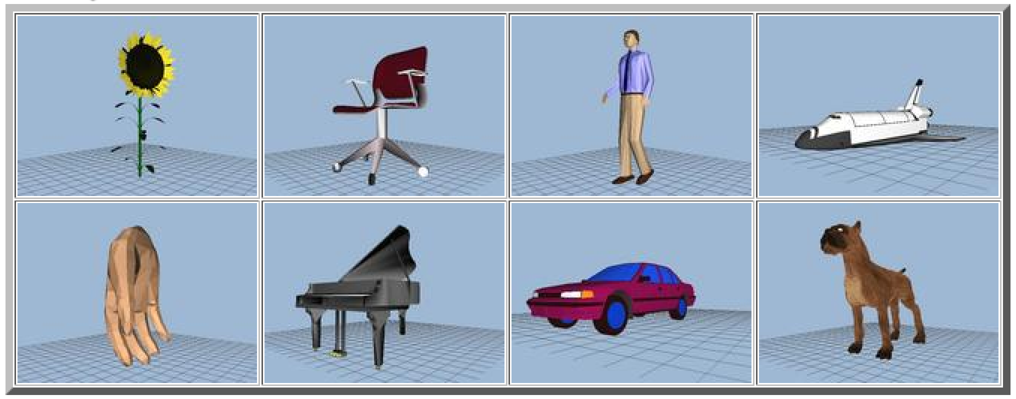
- classification (organizing large databases)

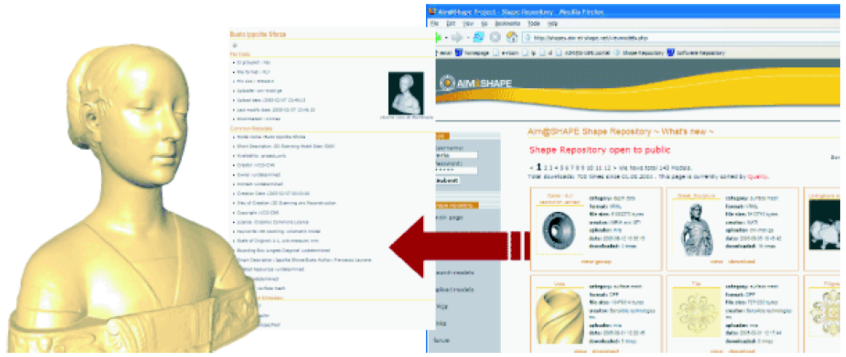
TOSCA



Princeton Shape Retrieval and Analysis Group
Princeton Shape Benchmark

McGill Shape Benchmark



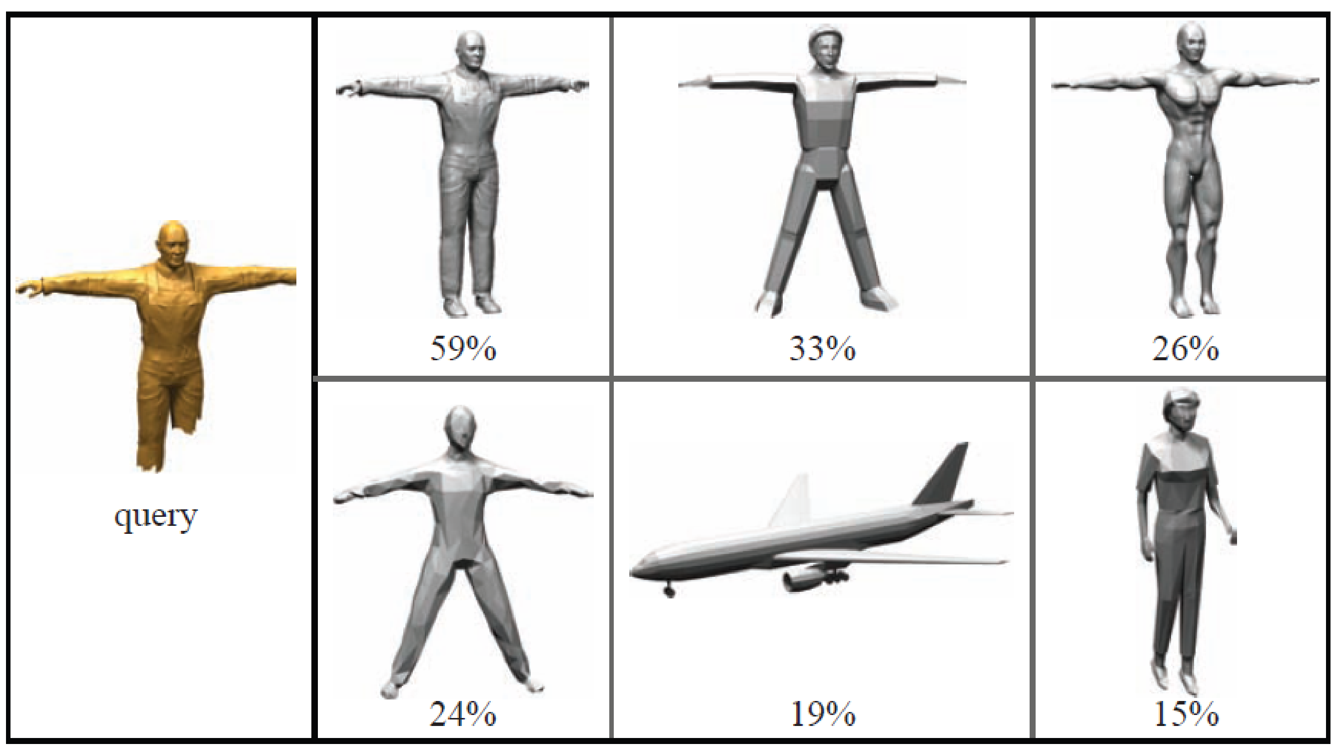


AIM@SHAPE
Digital Shape WorkBench

Why Compare Geometric Data

Comparisons between geometric data sets or parts thereof occur in:

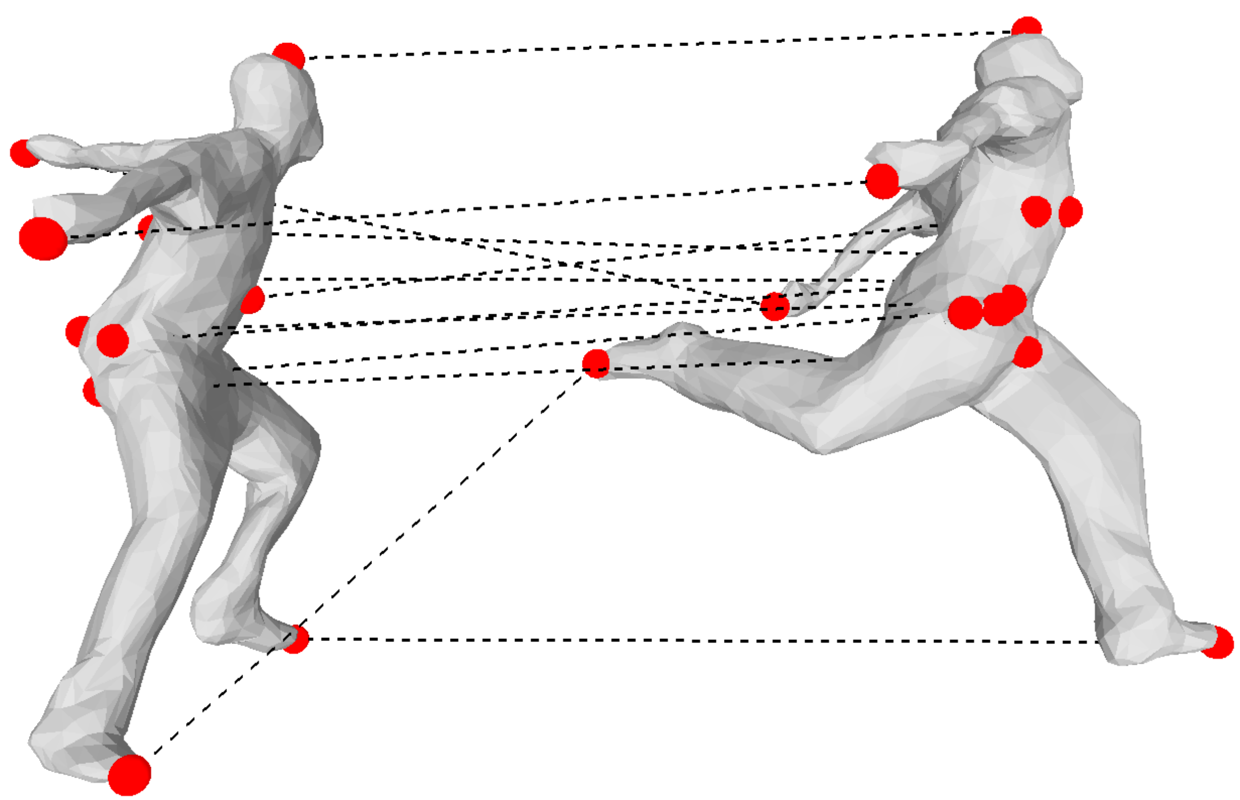
- classification (organizing large databases)
- retrieval (searching in databases)



Why Compare Geometric Data

Comparisons between geometric data sets or parts thereof occur in:

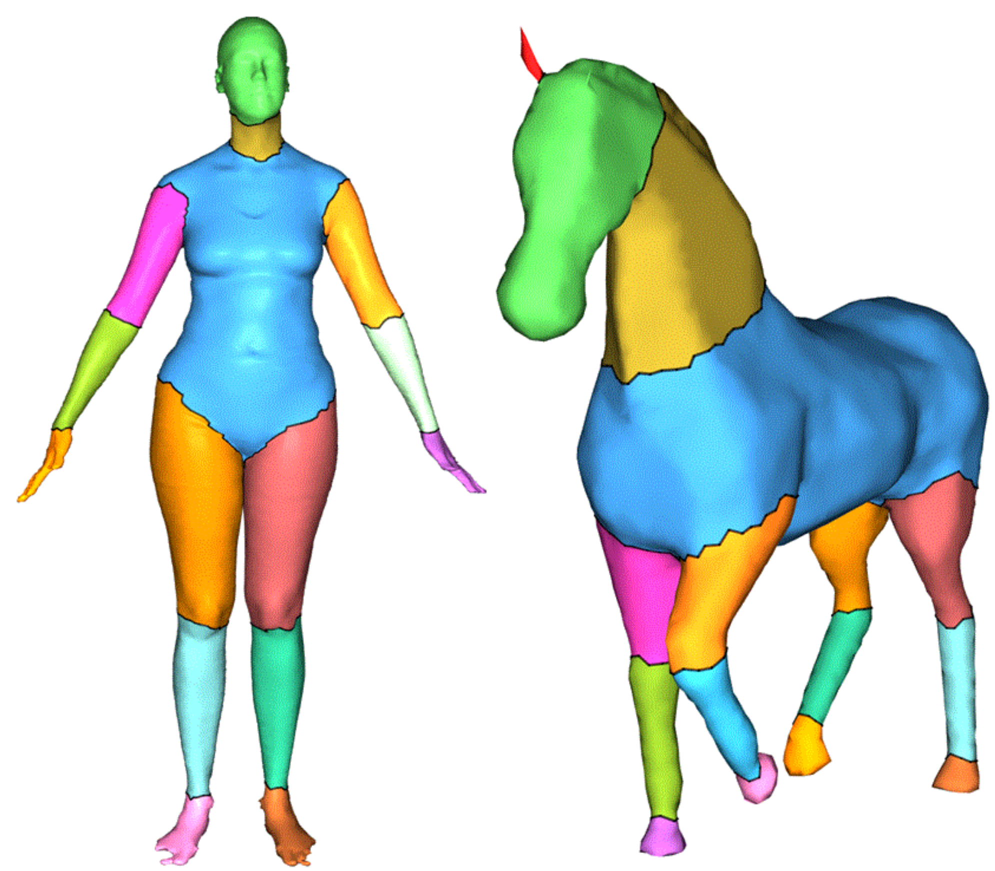
- classification (organizing large databases)
- retrieval (searching in databases)
- partial/global matching (finding the *best* mapping between data sets)



Why Compare Geometric Data

Comparisons between geometric data sets or parts thereof occur in:

- classification (organizing large databases)
- retrieval (searching in databases)
- partial/global matching (finding the *best* mapping between data sets)
- segmentation and labelling

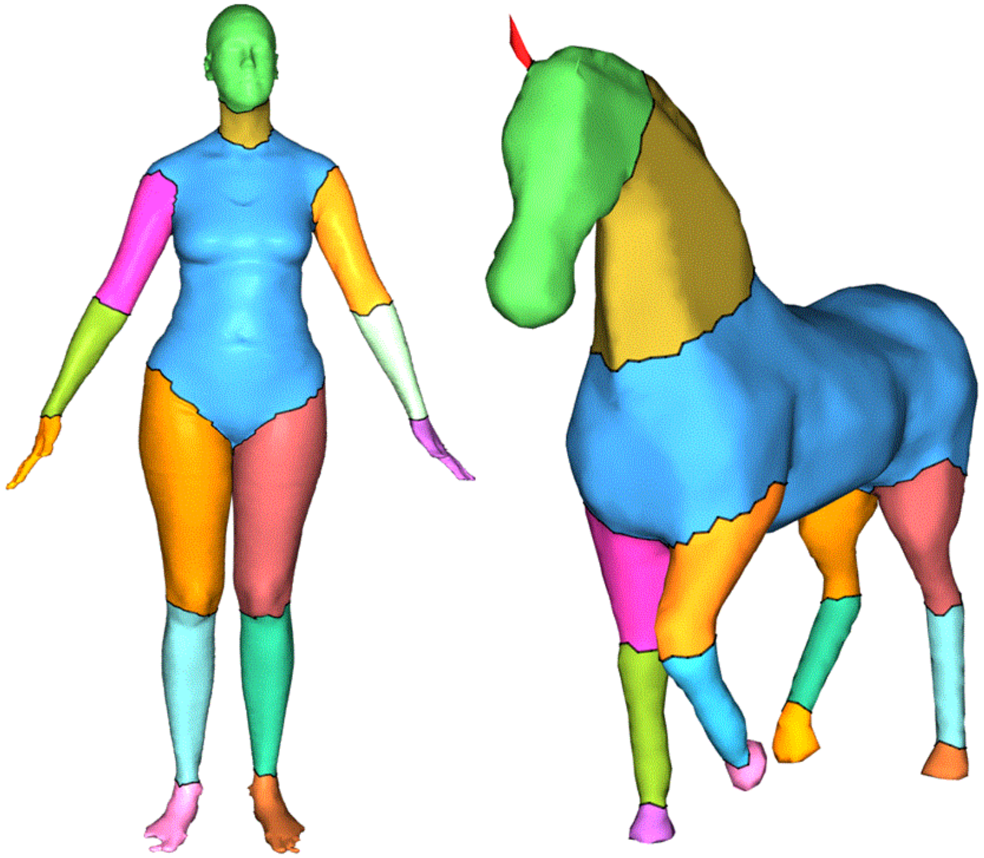


Why Compare Geometric Data

Comparisons between geometric data sets or parts thereof occur in:

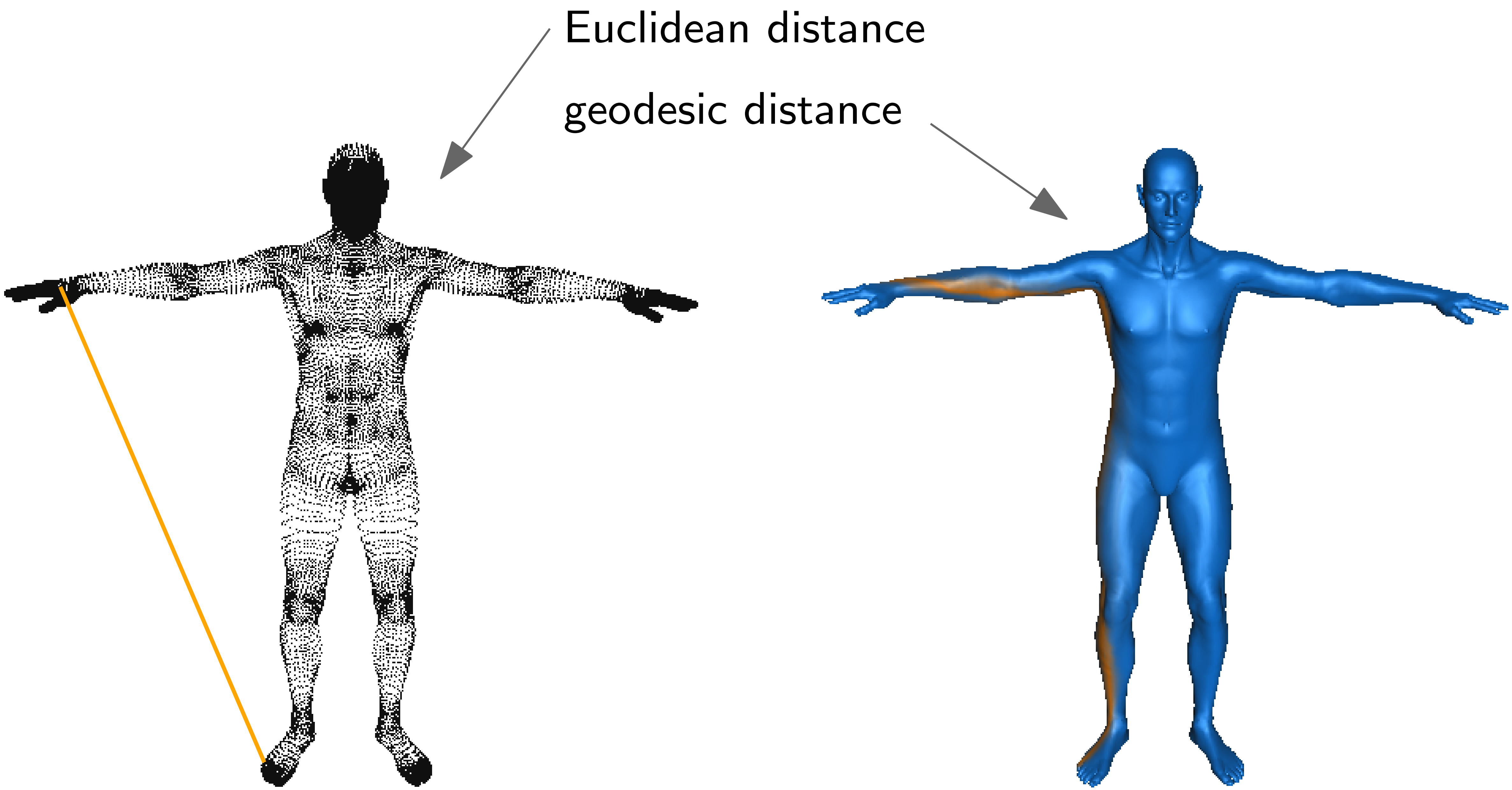
- classification (organizing large databases)
- retrieval (searching in databases)
- partial/global matching (finding the *best* mapping between data sets)
- segmentation and labelling

data comparison is
the basic building block



Mathematical Framework

- geometric data set $\equiv$ compact metric space



Mathematical Framework

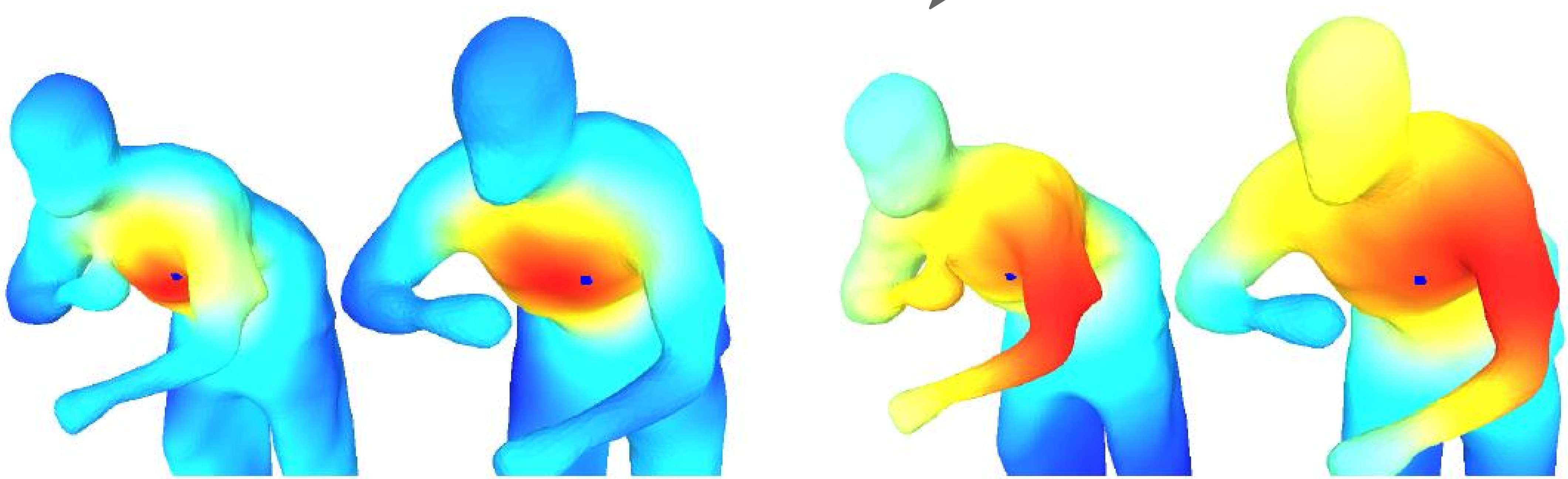
- geometric data set $\equiv$ compact metric space

Euclidean distance

geodesic distance

diffusion distance

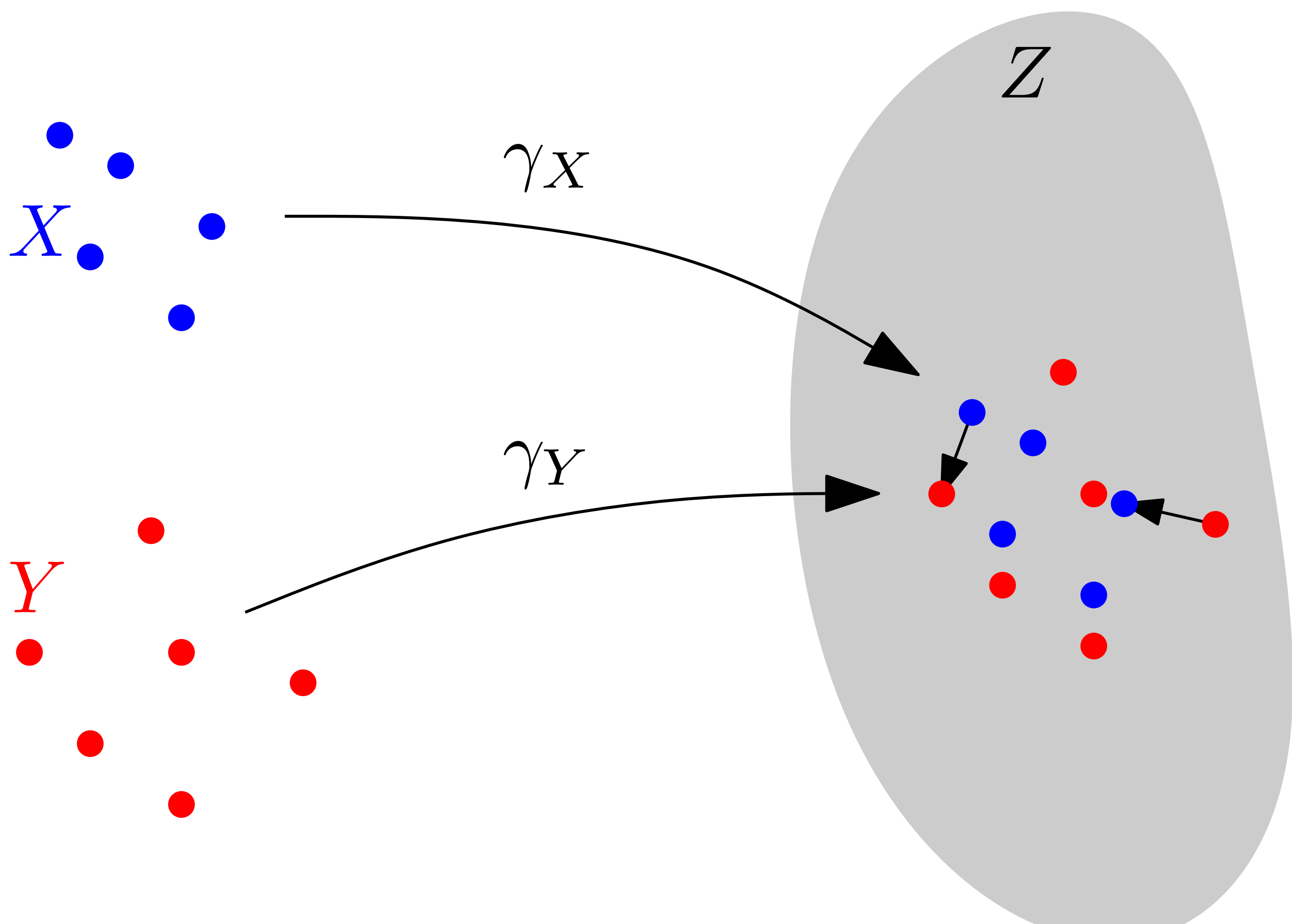
...



Mathematical Framework

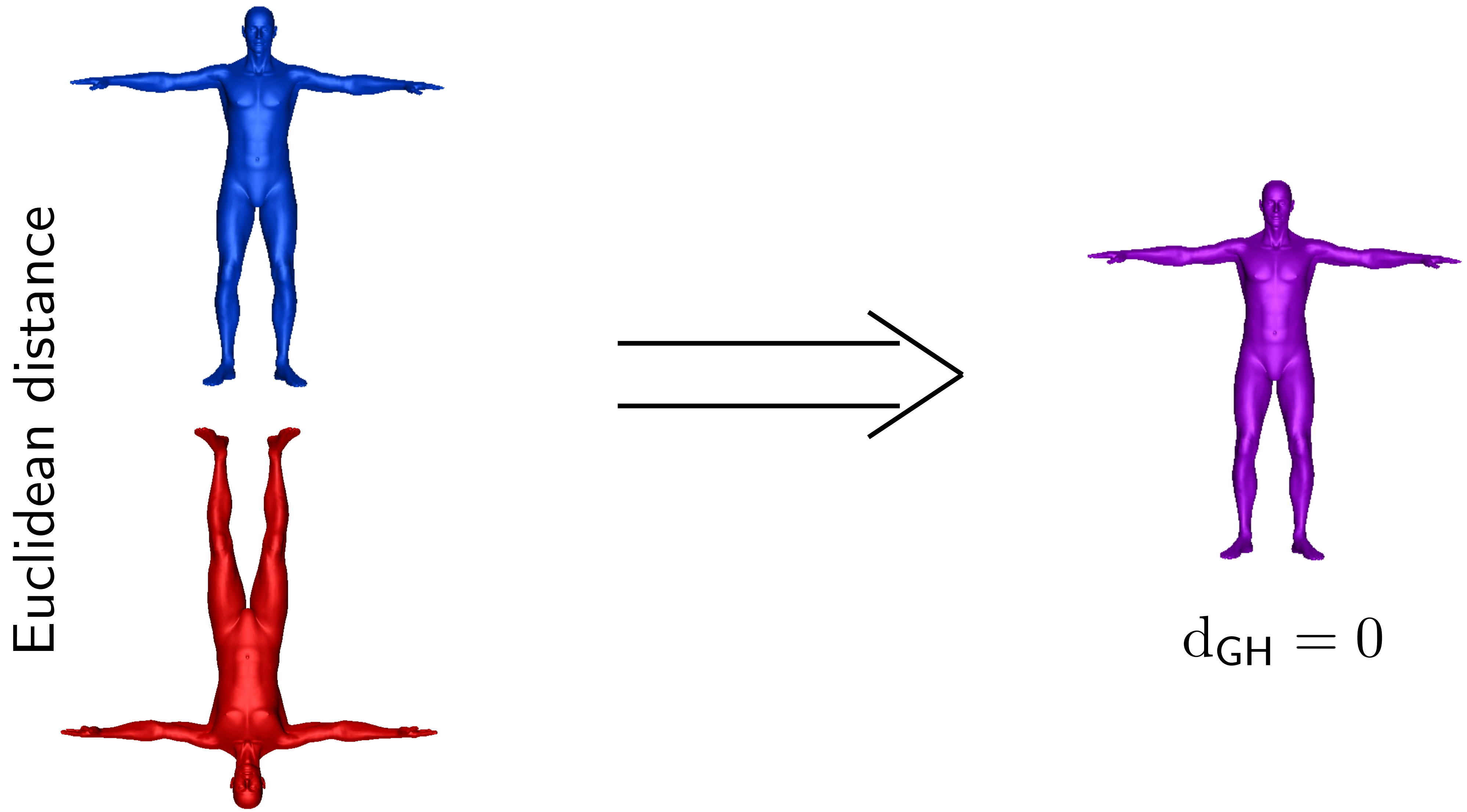
- geometric data set $\equiv$ compact metric space

- distance between data sets $\equiv$ Gromov-Hausdorff (GH) distance



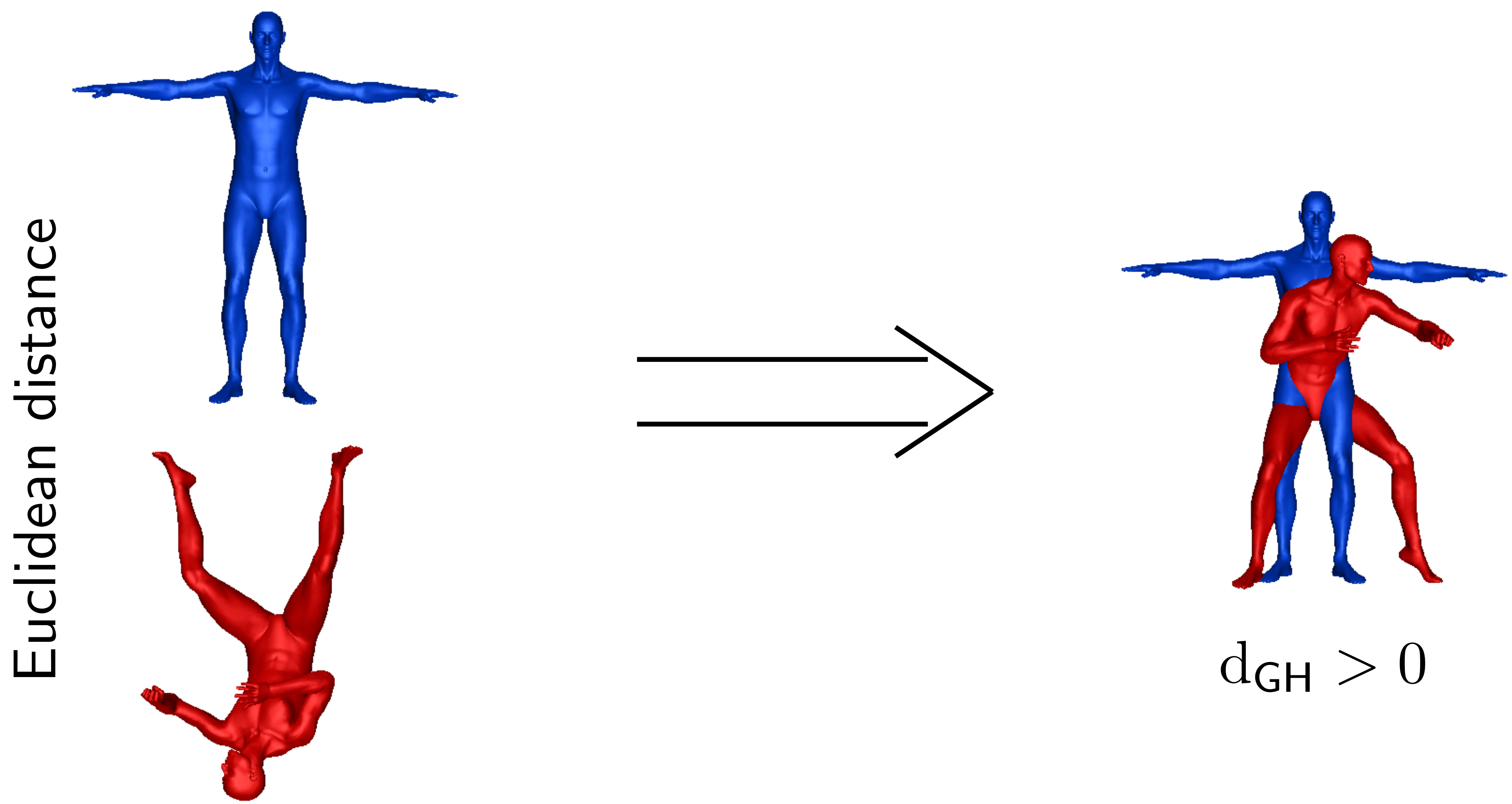
Mathematical Framework

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



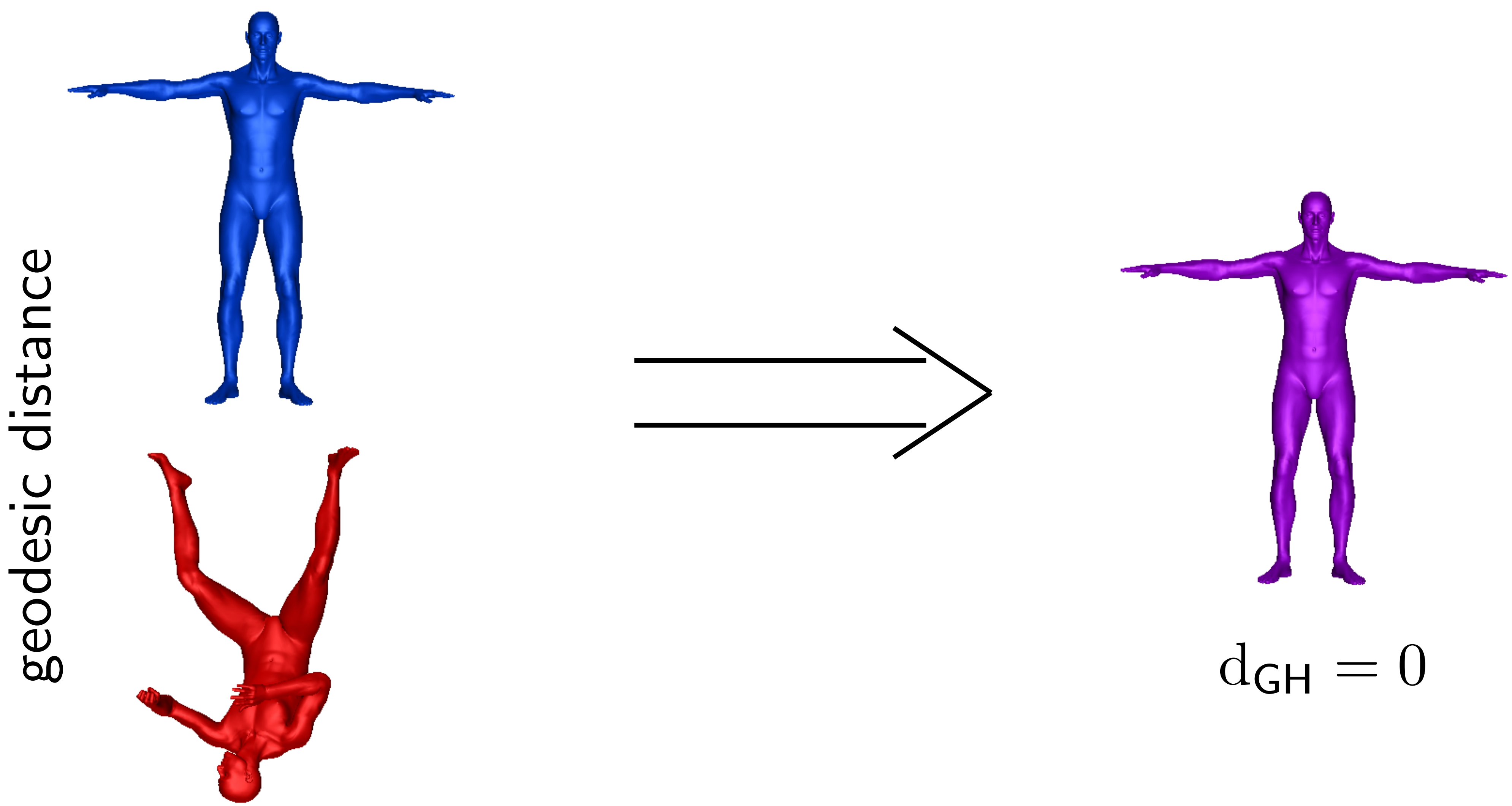
Mathematical Framework

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



Mathematical Framework

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

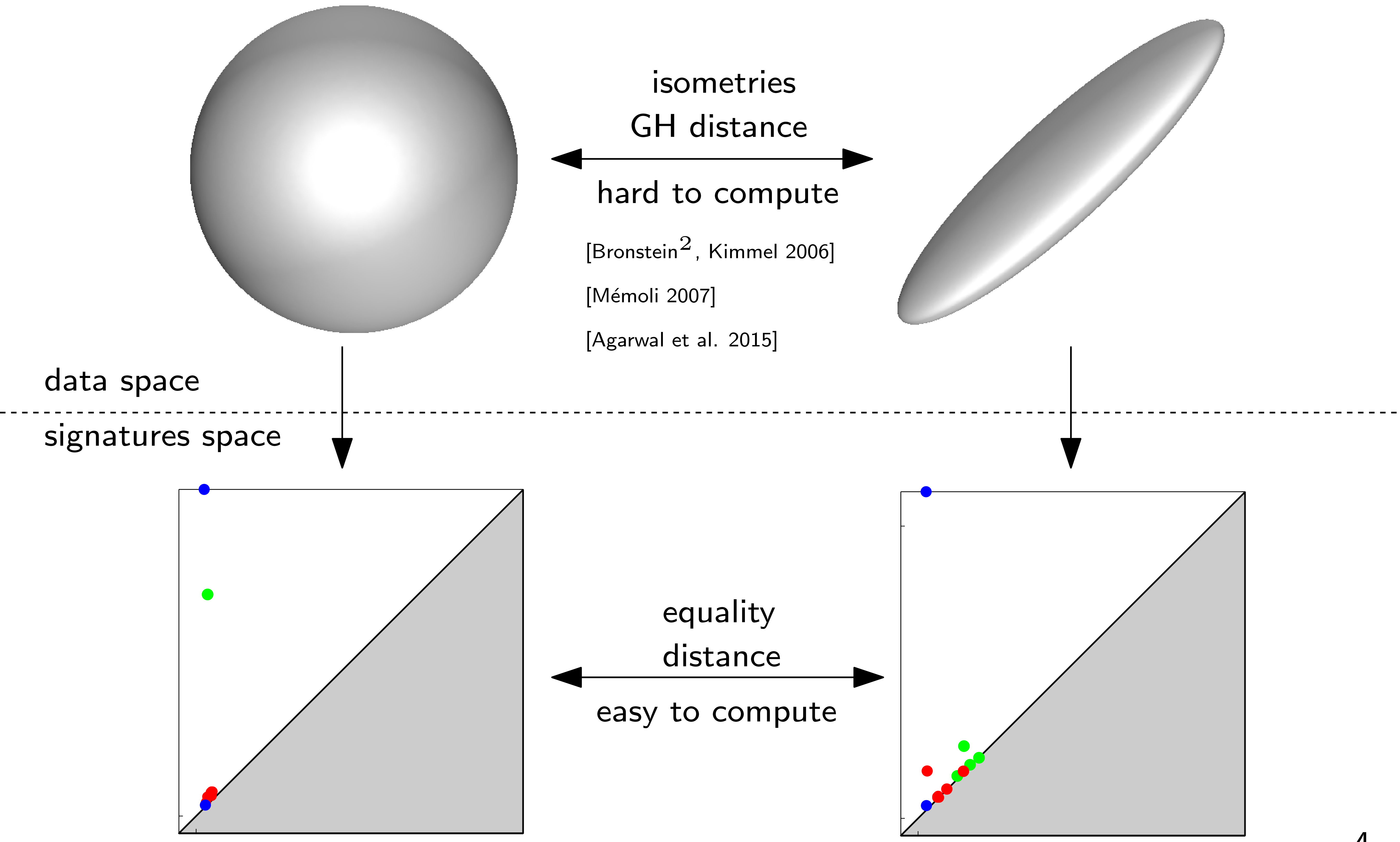


Mathematical Framework

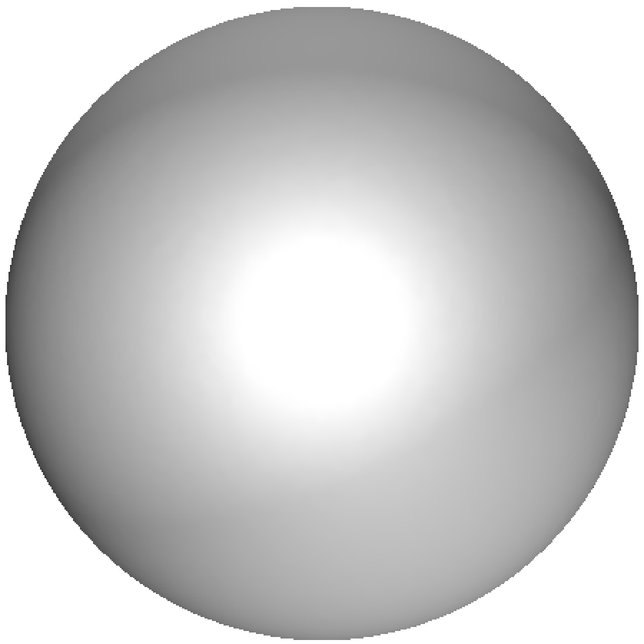
- geometric data set $\equiv$ compact metric space
- distance between data sets $\equiv$ Gromov-Hausdorff (GH) distance
- *signature* $\equiv$ persistence diagram (choose the filtration)

  - **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 Signatures



Why use Signatures



isometries

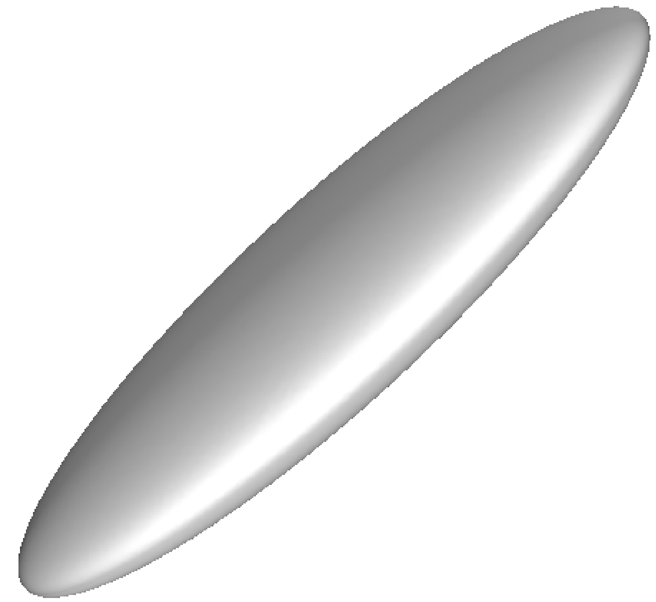
GH distance

hard to compute

[Bronstein^2, Kimmel 2006]

[Mémoli 2007]

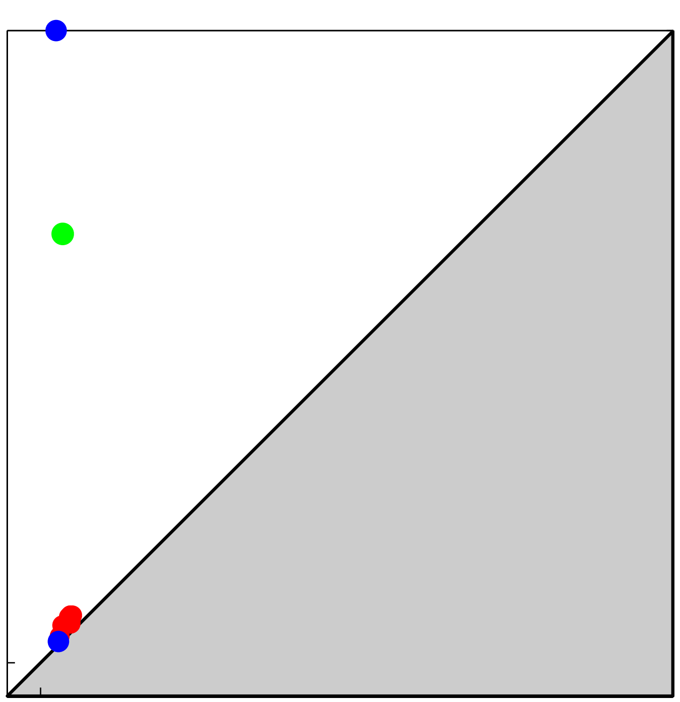
[Agarwal et al. 2015]



data space

signatures space

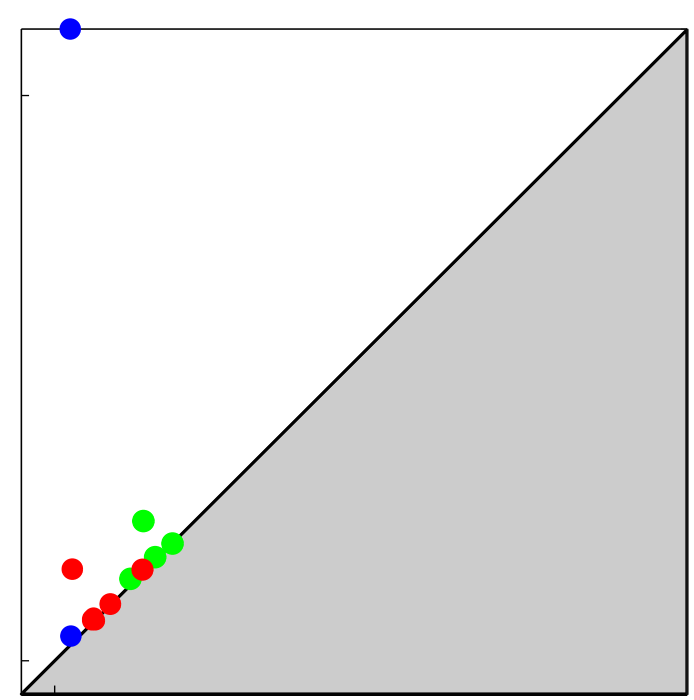
Ideally, signatures distance = GH distance



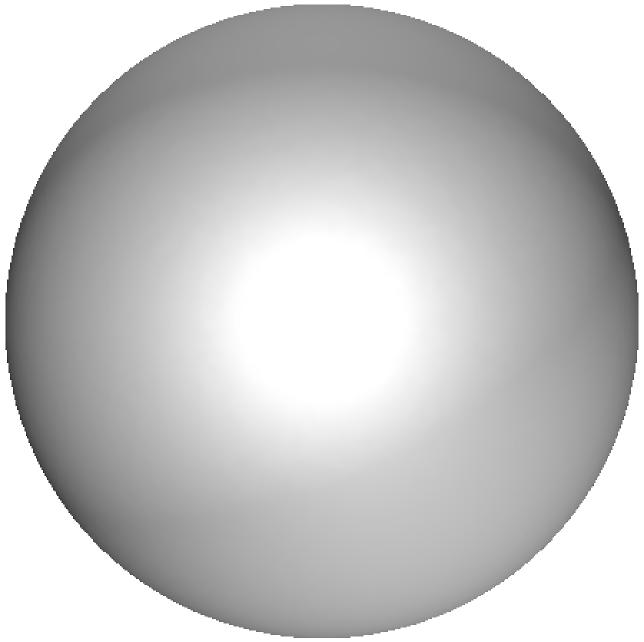
equality

distance

easy to compute



Why use Signatures



isometries

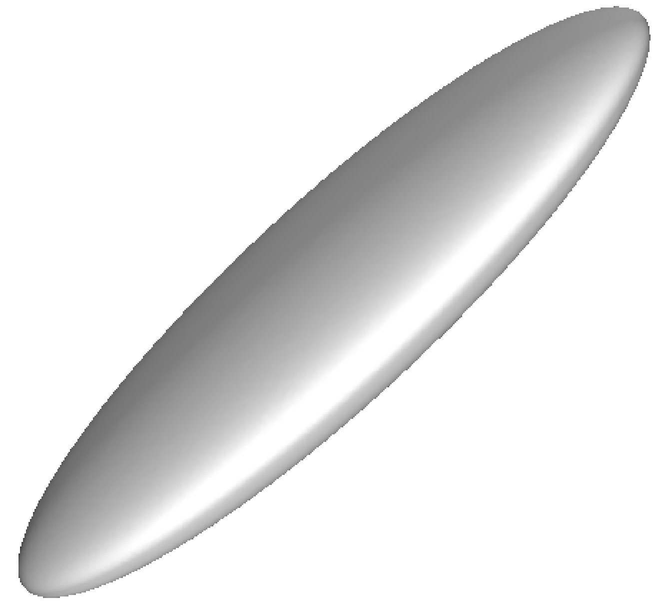
GH distance

hard to compute

[Bronstein^2, Kimmel 2006]

[Mémoli 2007]

[Agarwal et al. 2015]

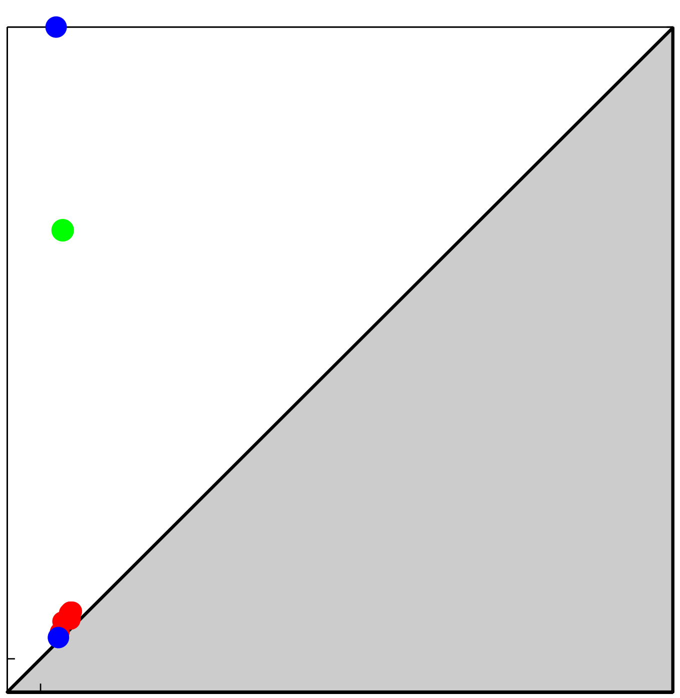


data space

signatures space

Ideally, signatures distance $=$ GH distance

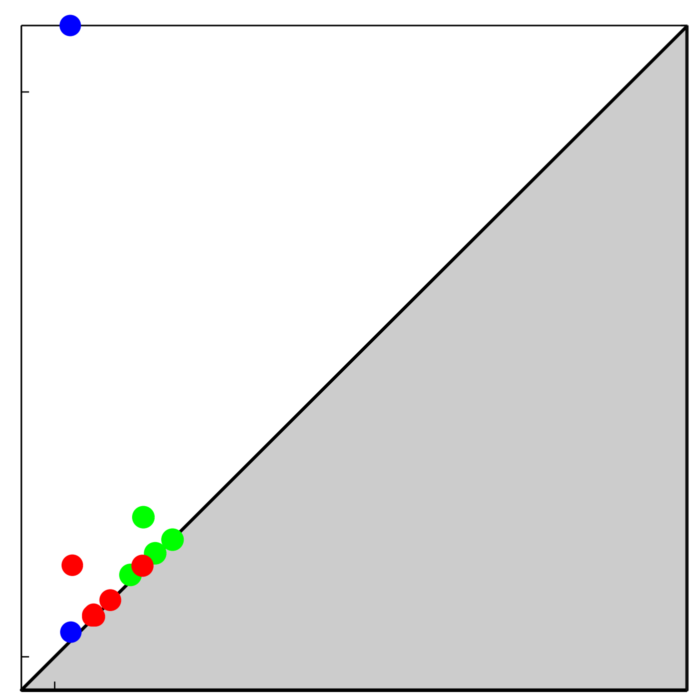
In reality, $\leq$



equality

distance

easy to compute



Why use Signatures

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.

Outline

1. Global topological signatures

2. Local topological signatures

3. Kernels for topological signatures

Outline

1. Global topological signatures

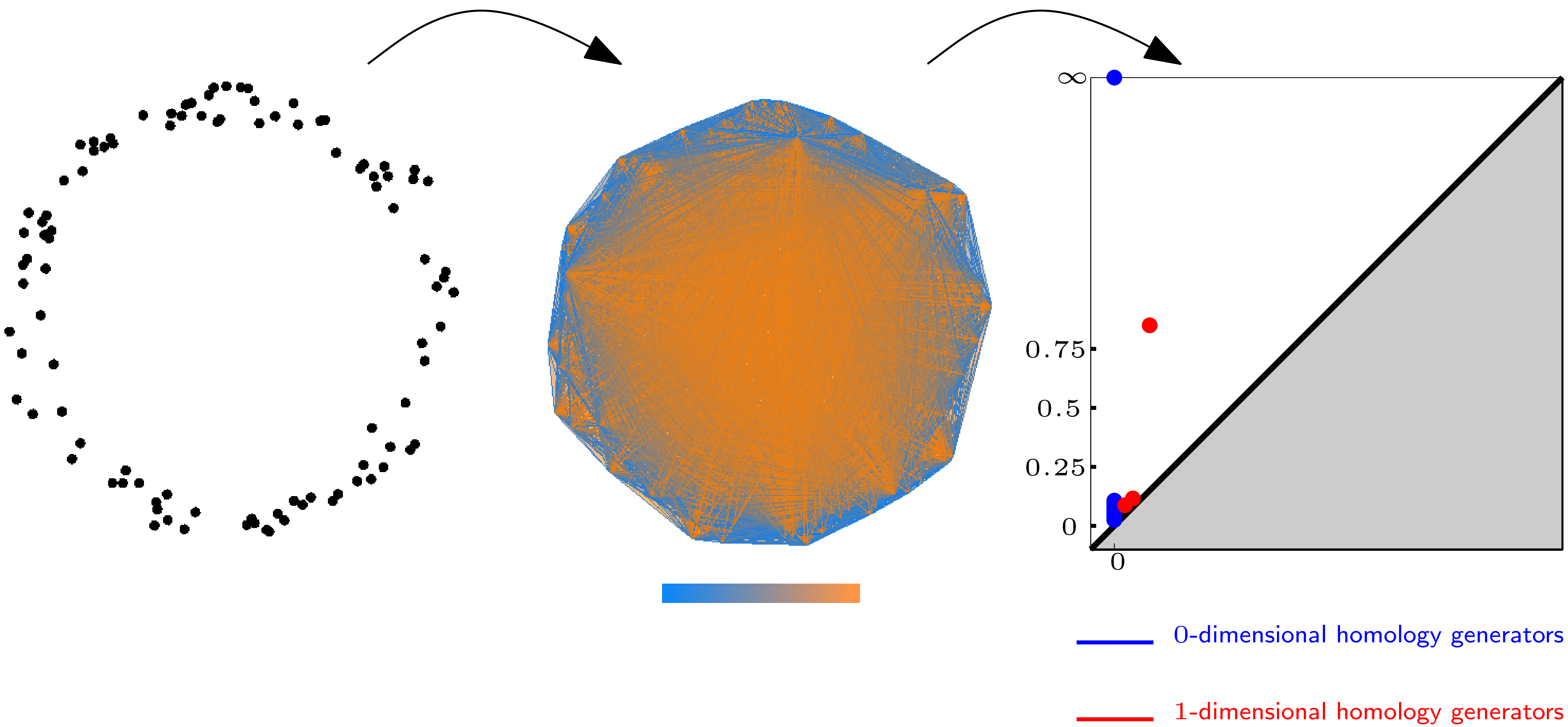
2. Local topological signatures

3. Kernels for topological signatures

Global Topological Signatures

Input: a compact metric space (X, d_X)

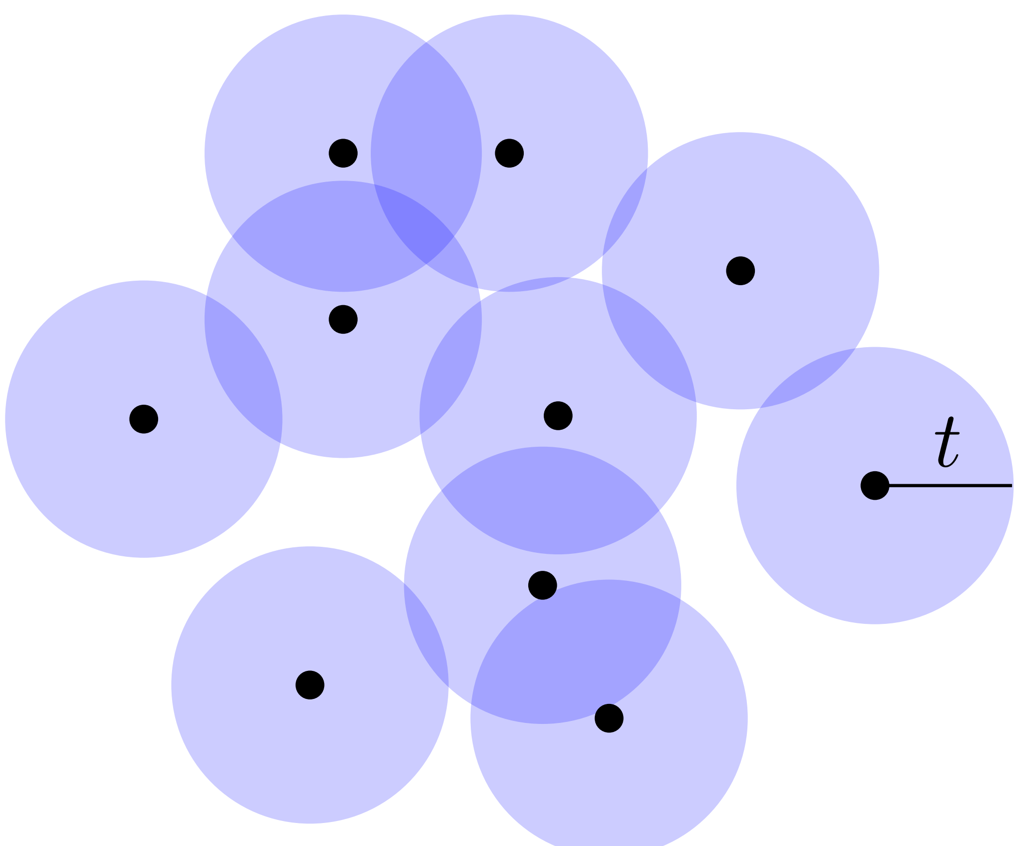
Signature: $\mathrm{dgm}\, \mathcal{F}(X, \mathrm{d}_X)$, where $\mathcal{F}(X, \mathrm{d}_X)$ is some simplicial filtration over X derived from d_X (proxy for union of balls)

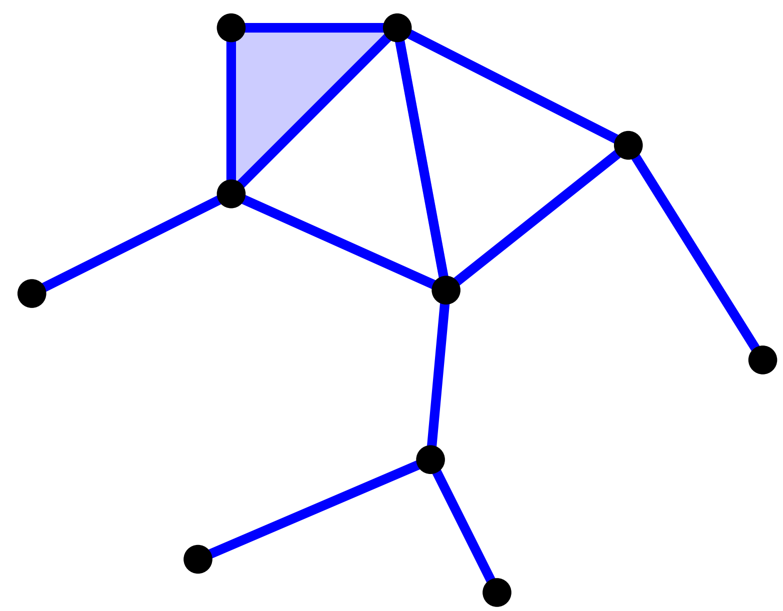


Global Topological Signatures

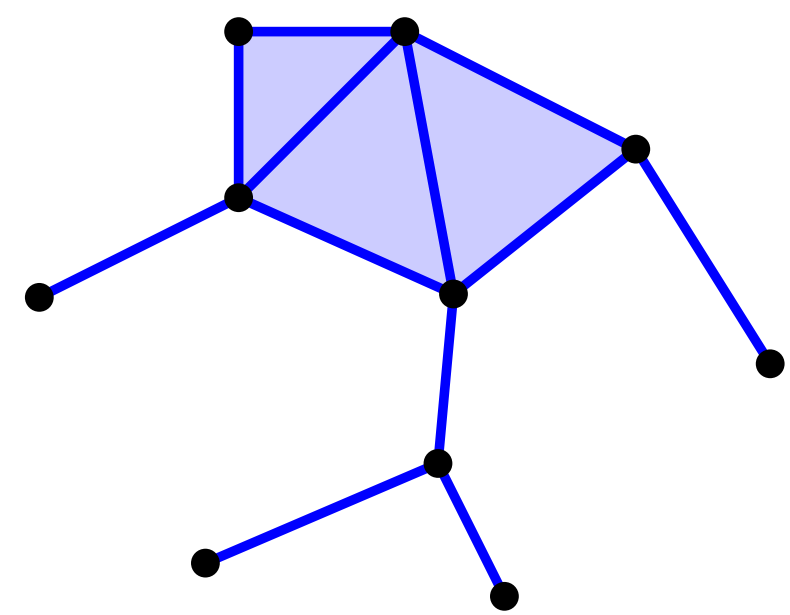
Input: a compact metric space (X, d_X)

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





$\mathrm{C}_t(X, \mathrm{d}_X)$



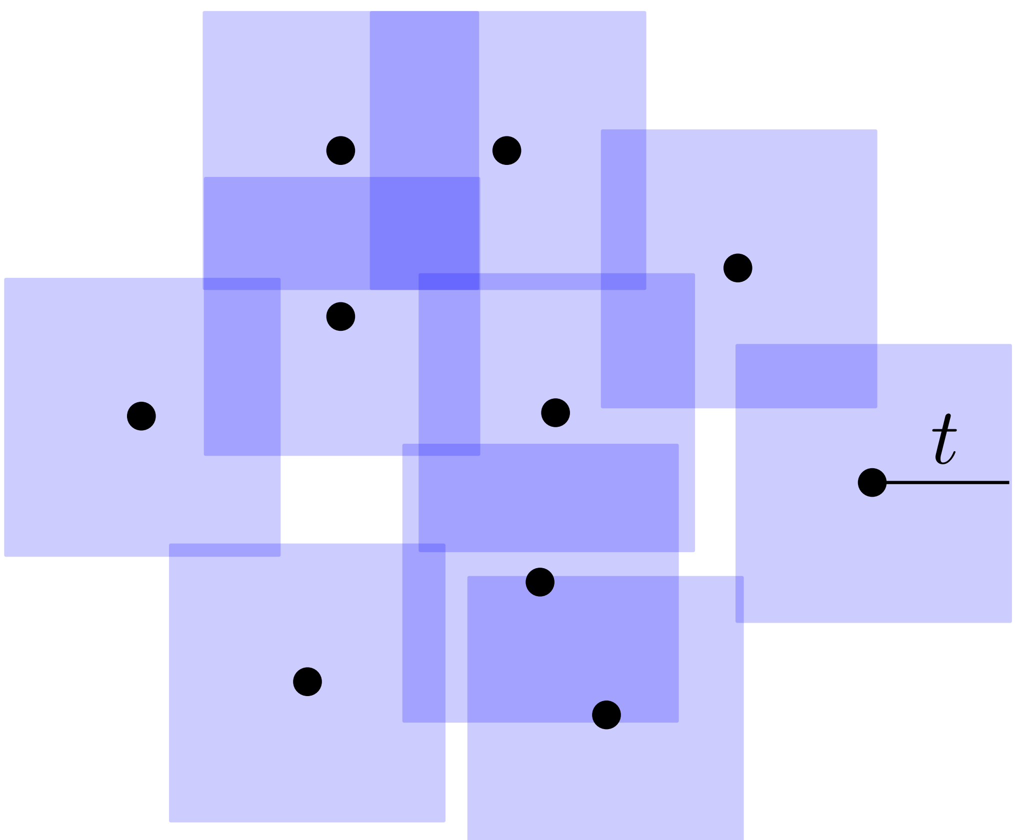
$\mathrm{R}_{2t}(X, \mathrm{d}_X)$

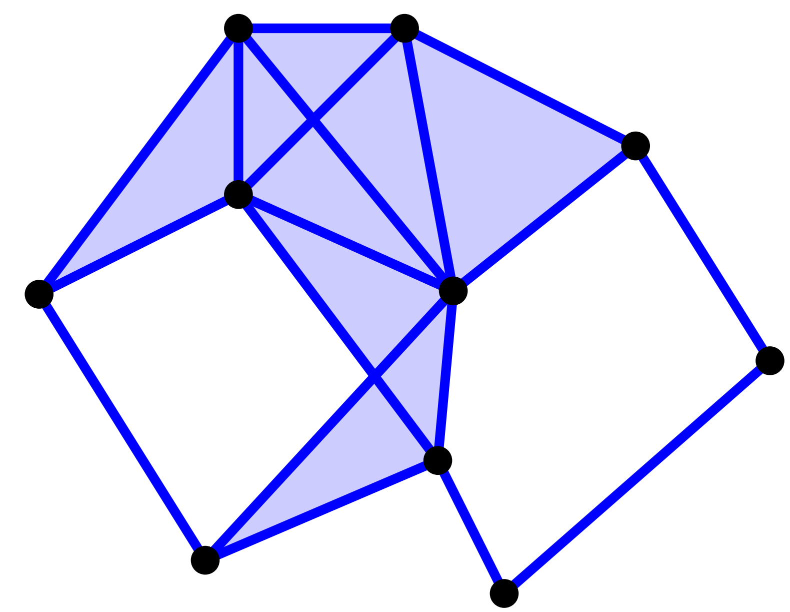
$\rightarrow$ popular choices:
- Čech/Nerve filtration $\mathcal{C}(X, \mathrm{d}_X)$
- (Vietoris)-Rips filtration $\mathcal{R}(X, \mathrm{d}_X)$

Global Topological Signatures

Input: a compact metric space (X, d_X)

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



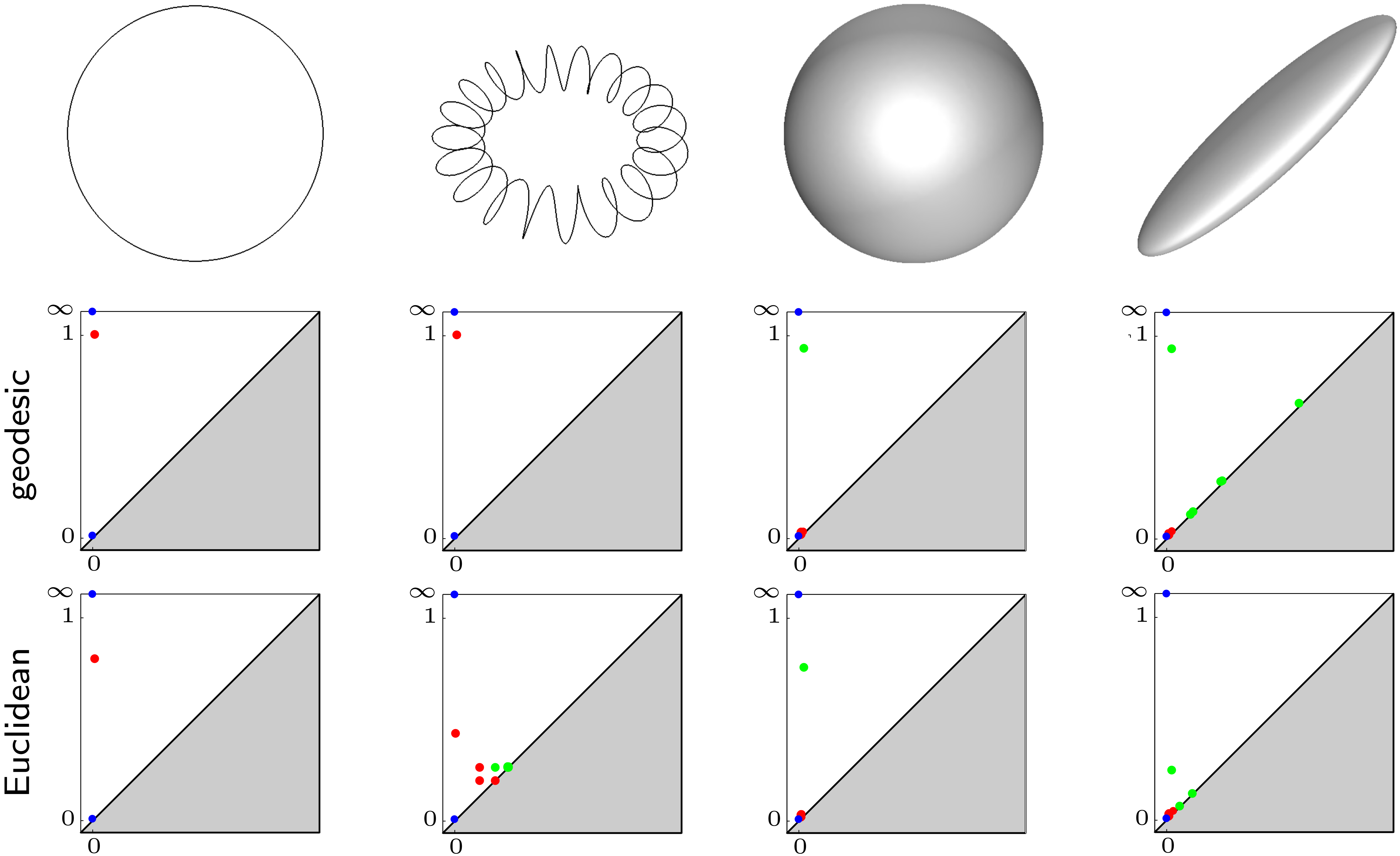


$\mathrm{C}_t(X, \mathrm{d}_X) = \mathrm{R}_{2t}(X, \mathrm{d}_X)$

$\rightarrow$ popular choices:
- Čech/Nerve filtration $\mathcal{C}(X, \mathrm{d}_X)$
- (Vietoris)-Rips filtration $\mathcal{R}(X, \mathrm{d}_X)$

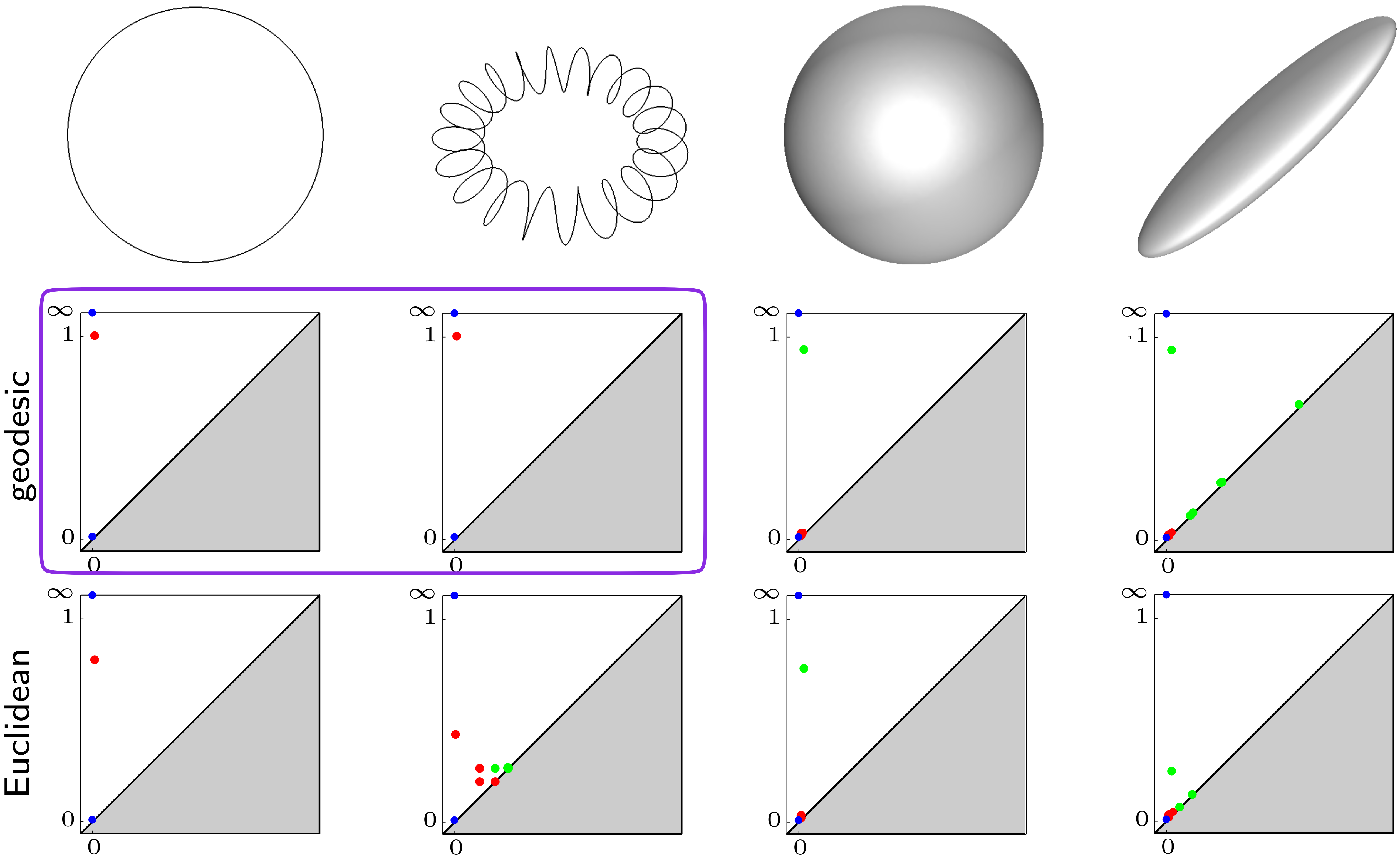
Examples

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



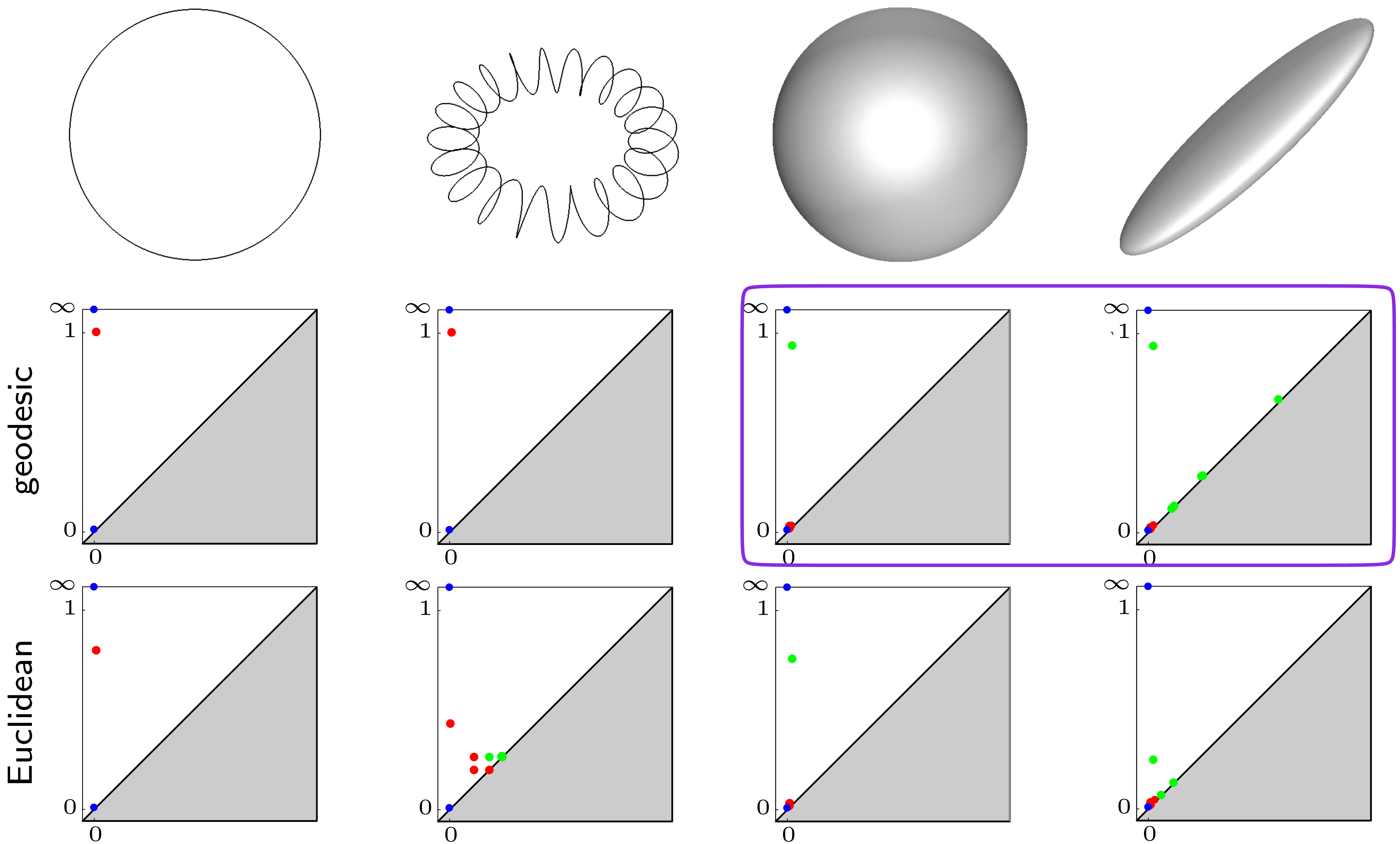
Examples

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



Examples

Signatures 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),
$\mathrm{d}_{\mathrm{B}}^{\infty}(\mathrm{dgm}\,\mathcal{R}(X, \mathrm{d}_X),\ \mathrm{dgm}\,\mathcal{R}(Y, \mathrm{d}_Y)) \leq 2\mathrm{d}_{\mathrm{GH}}(X, Y)$.

Variants and extensions:

- Čech / Nerve filtrations

- Witness complex filtrations (landmarks fixed)

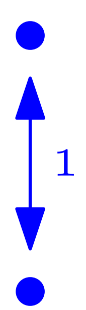
- precompact metric spaces

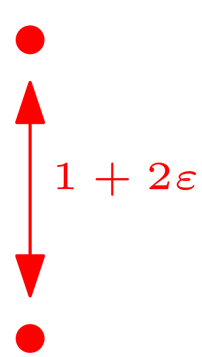
- (dis-)similarity measures

Stability

Theorem: [Chazal, de Silva, O. 2013]
For any compact metric spaces (X, d_X) and (Y, d_Y),
$\mathrm{d}_\mathrm{B}^\infty(\operatorname{dgm}\mathcal{R}(X, \mathrm{d}_X),\ \operatorname{dgm}\mathcal{R}(Y, \mathrm{d}_Y)) \leq 2\mathrm{d}_\mathrm{GH}(X, Y)$.

The bound is worst-case tight...

$X =$ 

$Y =$ 

$$\mathrm{d}_\mathrm{GH}(X, Y) = \varepsilon$$

$$\operatorname{dgm}\mathcal{R}(X, \mathrm{d}_X) = \{(0, \infty),\ (0, 1)\}$$

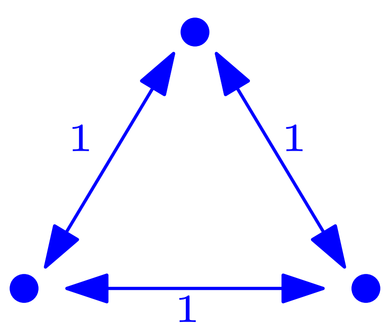
$$\operatorname{dgm}\mathcal{R}(Y, \mathrm{d}_Y) = \{(0, \infty),\ (0, 1 + 2\varepsilon)\}$$

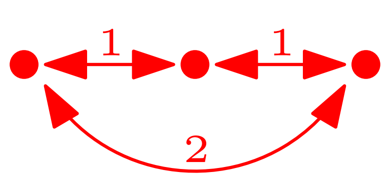
$$\Rightarrow \mathrm{d}_\mathrm{B}^\infty(\operatorname{dgm}\mathcal{R}(X, \mathrm{d}_X),\ \operatorname{dgm}\mathcal{R}(Y, \mathrm{d}_Y)) = 2\varepsilon$$

Stability

Theorem: [Chazal, de Silva, O. 2013]
For any compact metric spaces (X, d_X) and (Y, d_Y),
$\mathrm{d}_{\mathrm{B}}^{\infty}(\mathrm{dgm}\,\mathcal{R}(X, \mathrm{d}_X),\ \mathrm{dgm}\,\mathcal{R}(Y, \mathrm{d}_Y)) \leq 2\mathrm{d}_{\mathrm{GH}}(X, Y)$.

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

$X =$ 

$Y =$ 

$$\mathrm{d}_{\mathrm{GH}}(X, Y) = \tfrac{1}{2}$$

$$\mathrm{dgm}\,\mathcal{R}(X, \mathrm{d}_X) = \{(0, \infty),\ (0, 1),\ (0.1)\}$$

$$\mathrm{dgm}\,\mathcal{R}(Y, \mathrm{d}_Y) = \{(0, \infty),\ (0, 1),\ (0, 1)\}$$

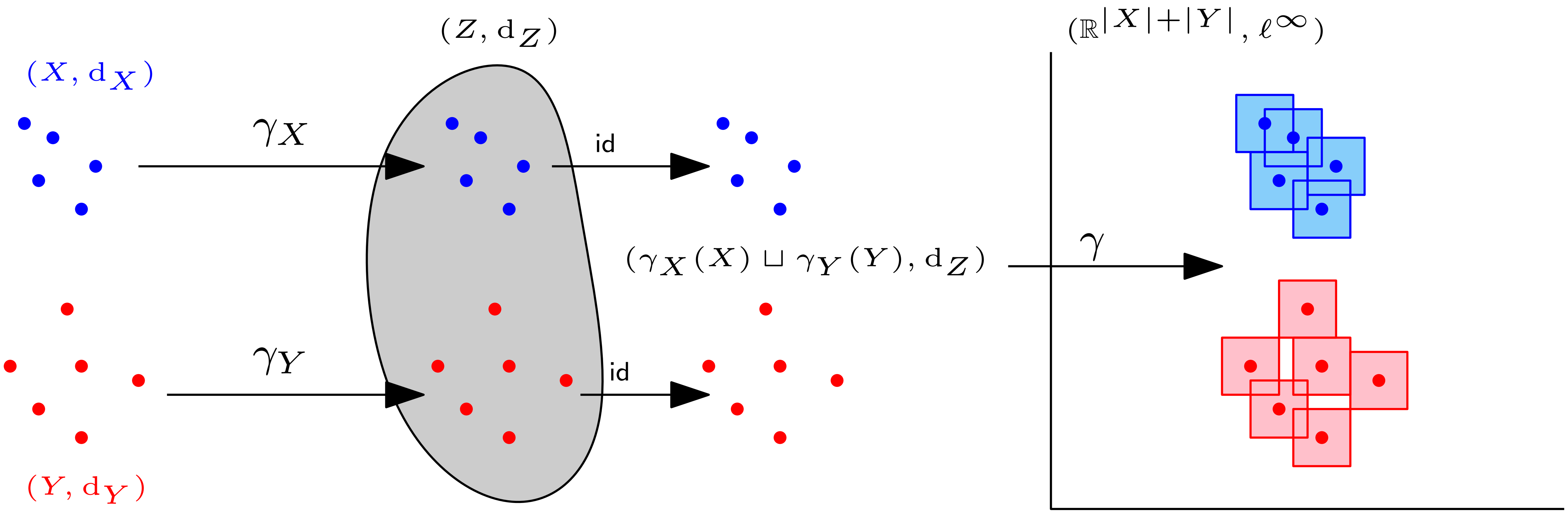
$$\Rightarrow \mathrm{d}_{\mathrm{B}}^{\infty}(\mathrm{dgm}\,\mathcal{R}(X, \mathrm{d}_X),\ \mathrm{dgm}\,\mathcal{R}(Y, \mathrm{d}_Y)) = 0$$

Stability

finite

Theorem: [Chazal, de Silva, O. 2013]
For any ~~compact~~ metric spaces (X, d_X) and (Y, d_Y),
$\mathrm{d}_\mathrm{B}^\infty(\mathrm{dgm}\,\mathcal{R}(X, \mathrm{d}_X),\ \mathrm{dgm}\,\mathcal{R}(Y, \mathrm{d}_Y)) \leq 2\mathrm{d}_{\mathrm{GH}}(X, Y)$.

Proof outline:



Convergence Rates

(X, d_X): compact metric space

$\mathcal{P}$: proba. measures μ on X satisfying the *(a,b)-standard* condition:

$$\forall x \in \operatorname{supp} \mu,\ \forall r > 0,\ \mu(B(x, r) \geq \min\{1, ar^b\}.$$

Convergence Rates

(X, d_X): compact metric space

$\mathcal{P}$: proba. measures μ on X satisfying the *(a,b)-standard* condition:

$$\forall x \in \operatorname{supp}\mu,\ \forall r > 0,\ \mu(B(x,r) \geq \min\{1, ar^b\}.$$

Given $\mu \in \mathcal{P}$, let $\hat{X}_n = \{X_1, \cdots, X_n\}$ be sampled i.i.d. according to μ.

Theorem. [Chazal, Glisse, Labruère, Michel 2014]

$$\sup_{\mu \in \mathcal{P}} \mathbb{E}\left[\mathrm{d_B}\left(\operatorname{dgm}\mathcal{R}(\hat{X}_n),\ \operatorname{dgm}\mathcal{R}(\operatorname{supp}\mu)\right)\right] \leq C\ \left(\frac{\log n}{n}\right)^{1/b},$$

where C depends only on a, b. Moreover, the estimator $\operatorname{dgm}\mathcal{R}(\hat{X}_n)$ is minimax optimal on the space $\mathcal{P}$ up to a $\log n$ factor.

Convergence Rates

(X, d_X): compact metric space

$\mathcal{P}$: proba. measures μ on X satisfying the *(a,b)-standard* condition:

$$\forall x \in \operatorname{supp} \mu,\ \forall r > 0,\ \mu(B(x,r) \geq \min\{1, ar^b\}.$$

Given $\mu \in \mathcal{P}$, let $\hat{X}_n = \{X_1, \cdots, X_n\}$ be sampled i.i.d. according to μ.

Theorem. [Chazal, Glisse, Labruère, Michel 2014]

$$\sup_{\mu \in \mathcal{P}} \mathbb{E}\left[\mathrm{d_B}\left(\operatorname{dgm} \mathcal{R}(\hat{X}_n),\ \operatorname{dgm} \mathcal{R}(\operatorname{supp} \mu)\right)\right] \leq C\ \left(\frac{\log n}{n}\right)^{1/b},$$

where C depends only on a, b. Moreover, the estimator $\operatorname{dgm} \mathcal{R}(\hat{X}_n)$ is minimax optimal on the space $\mathcal{P}$ up to a $\log n$ factor.

Proof:

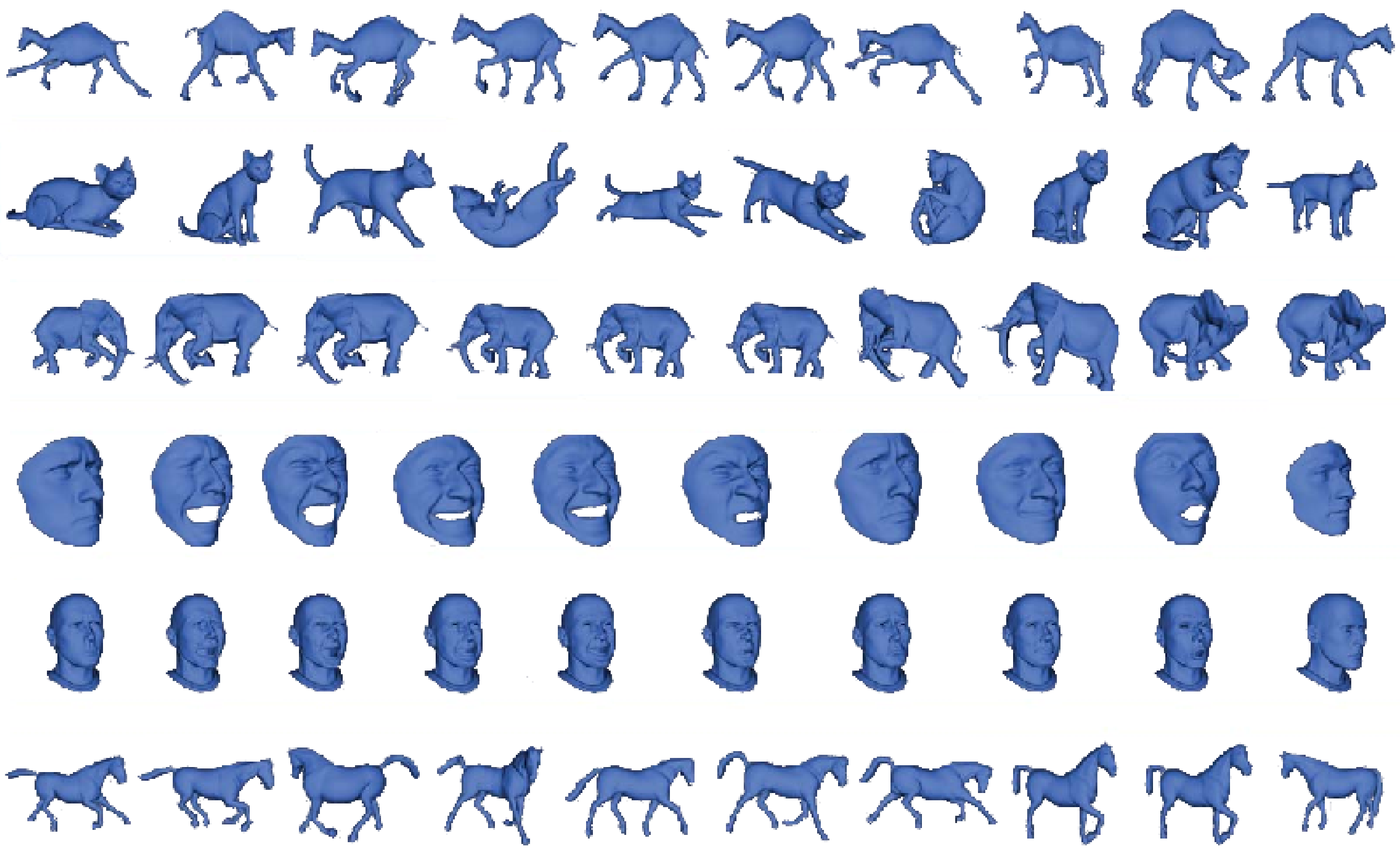
- upper bound: Hausdorff estimation of $\operatorname{supp} \mu$ + stability

- lower bound: Le Cam's lemma

$\square$

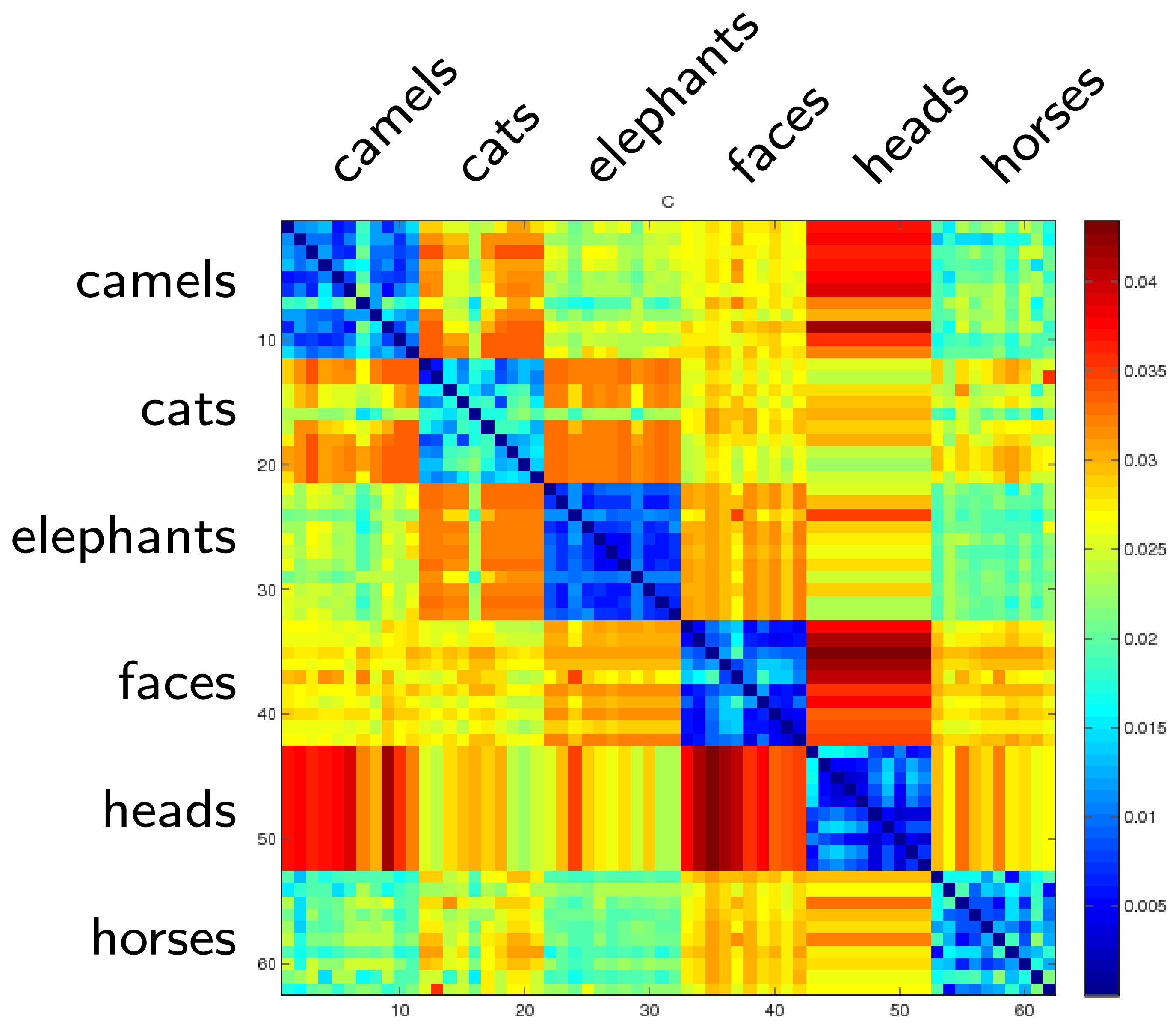
Application: Unsupervised Classification

Experimental results:



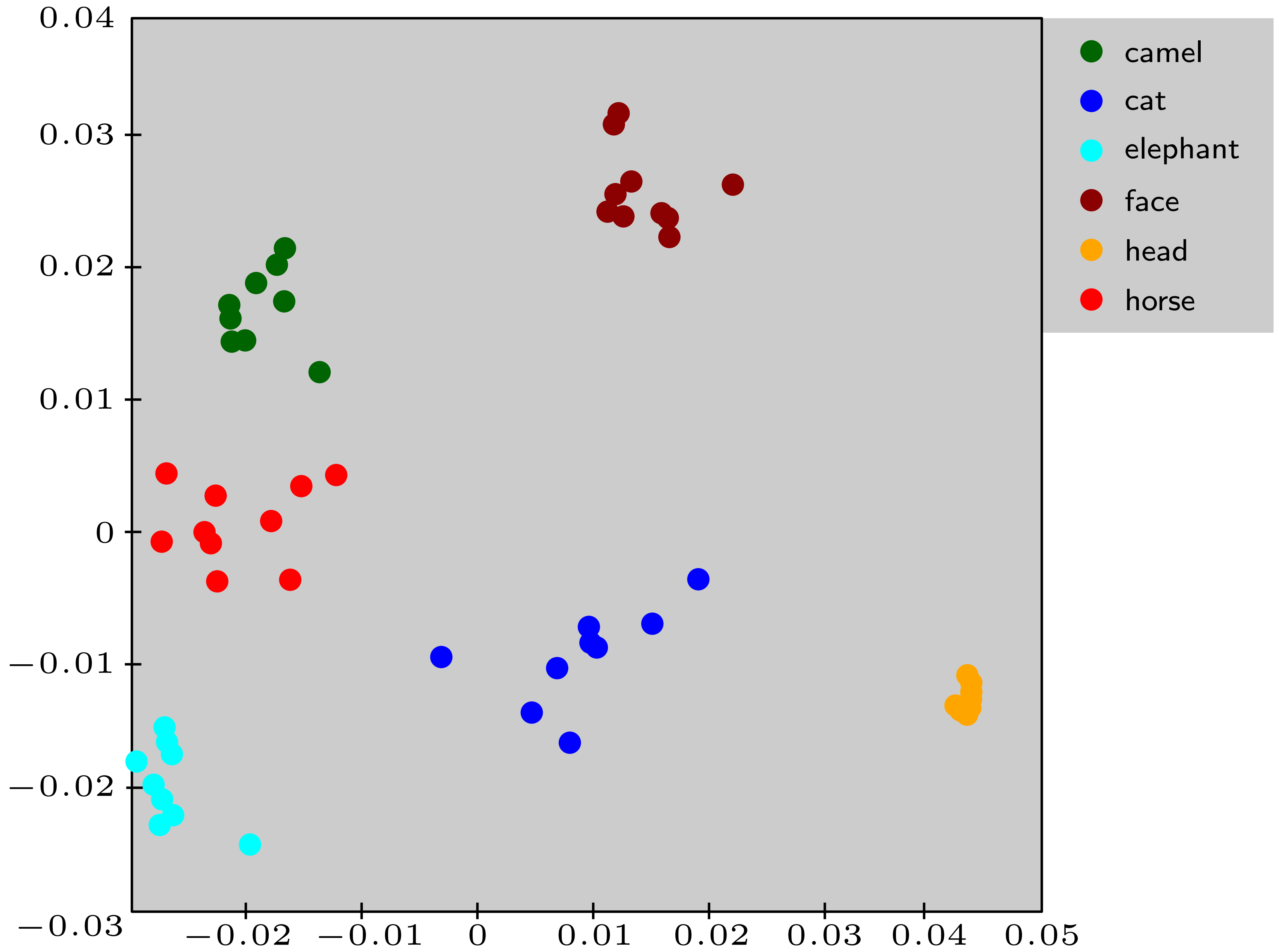
Application: Unsupervised Classification

Experimental results:



Application: Unsupervised Classification

Experimental results:



Outline

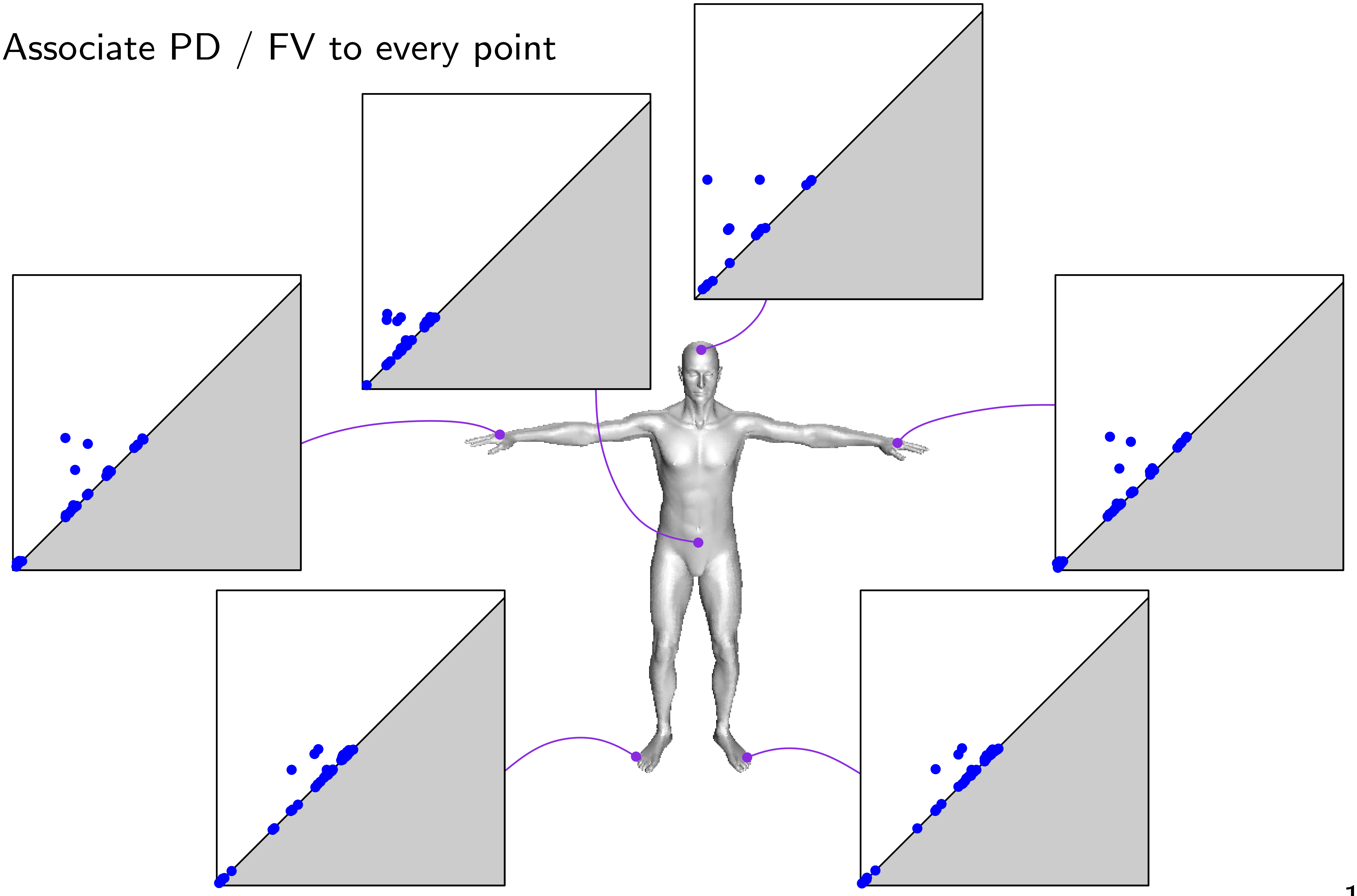
1. Global topological signatures

2. Local topological signatures

3. Kernels for topological signatures

Local Topological Signatures

Associate PD / FV to every point

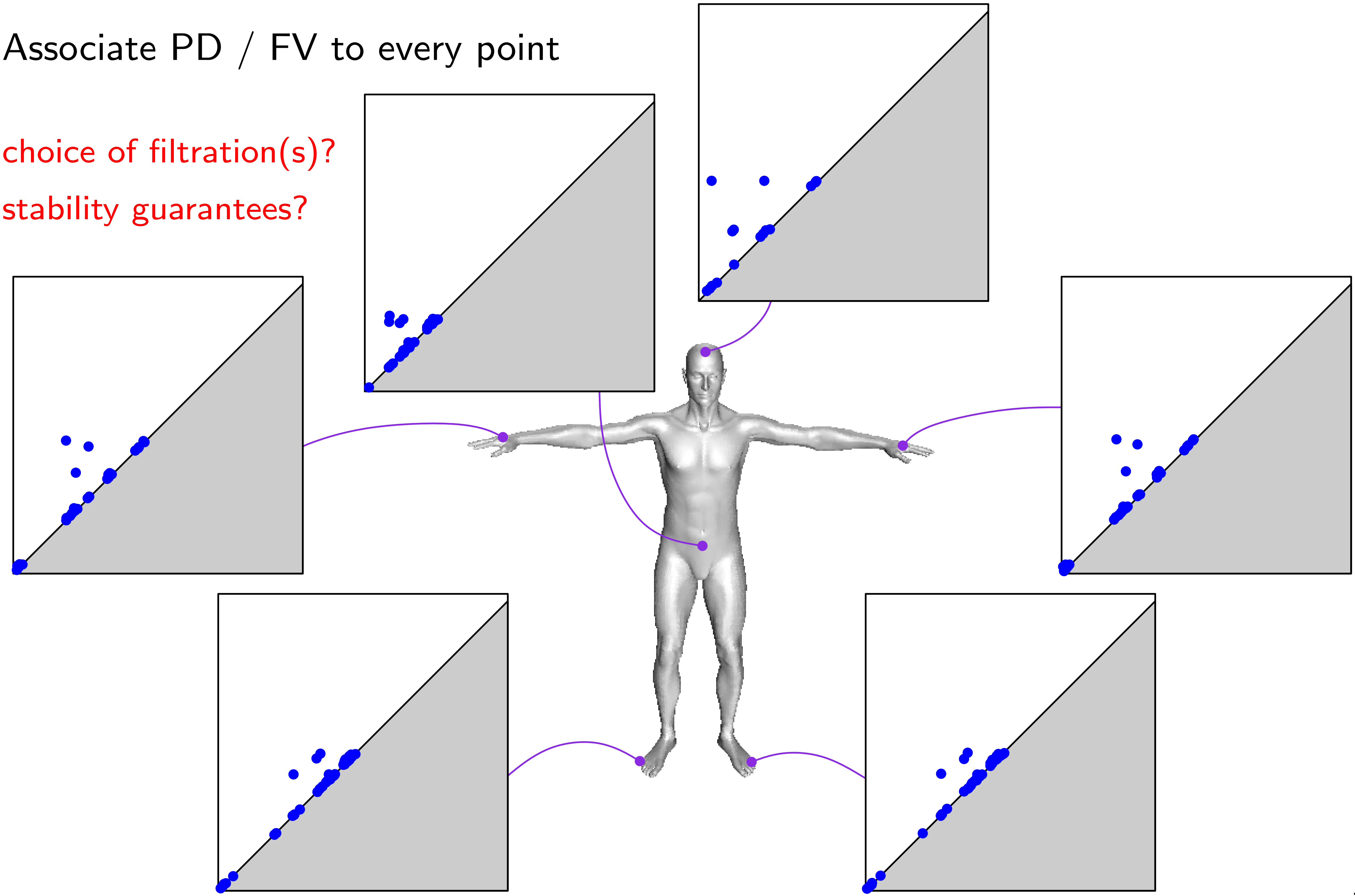


Local Topological Signatures

Associate PD / FV to every point

choice of filtration(s)?

stability guarantees?

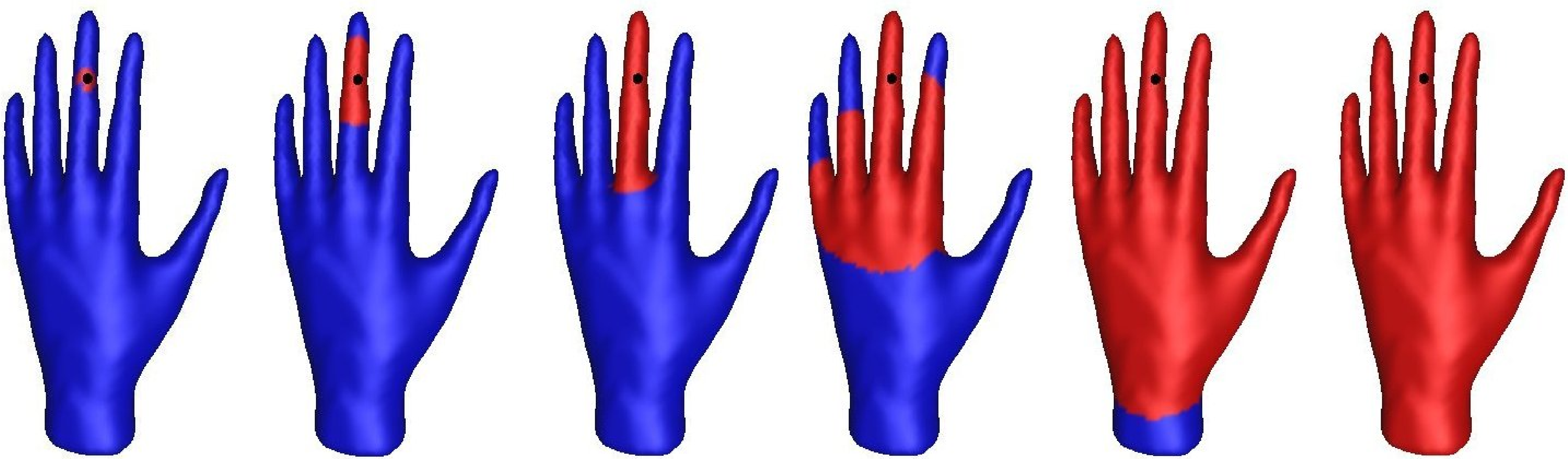


Local Topological Signatures

Input: a compact Riemannian manifold (X, d_X), a basepoint $x \in X$

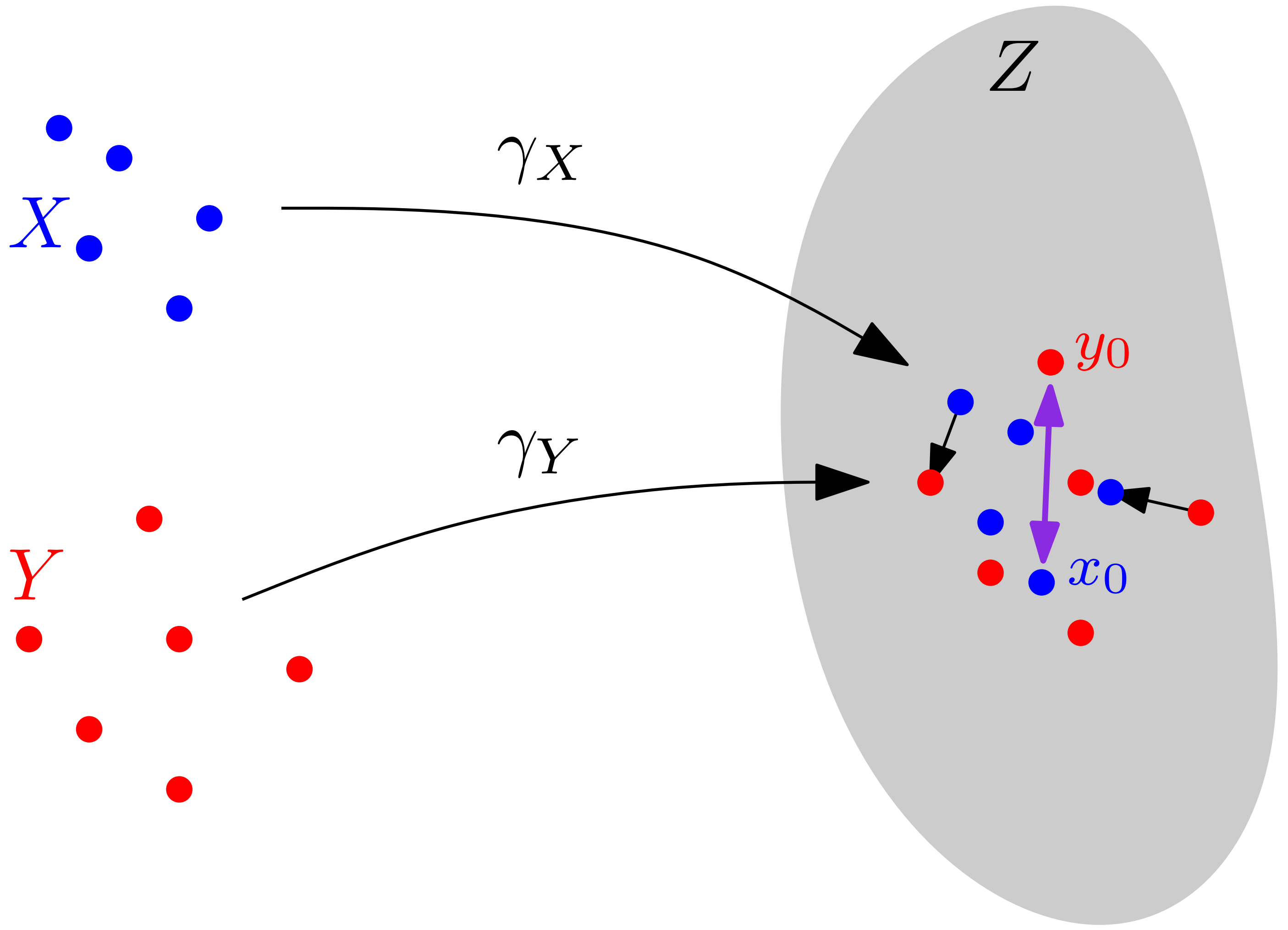
Construction: filtration of the sublevel sets of $\mathrm{d}_{x_0}(\cdot) = \mathrm{d}_X(x_0, \cdot)$

Signature: the persistence diagram of the filtration, denoted $\operatorname{dgm} \mathrm{d}_{x_0}$



Stability

Theorem: [Carrière, O., Ovsjanikov 2015]
Let (X, d_X) and (Y, d_Y) be compact Riemannian manifolds. Let $x_0 \in X$ and $y_0 \in Y$. If $\mathrm{d}_{\mathrm{GH}}((X, x_0), (Y, y_0)) \leq \frac{1}{20} \min\{\varrho(X), \underbrace{\ }_{\text{(convexity radius)}} \varrho(Y)\}$, then $\mathrm{d}_{\mathrm{B}}^{\infty}(\operatorname{dgm} \mathrm{d}_{x_0}, \operatorname{dgm} \mathrm{d}_{y_0}) \leq 20\, \underbrace{\mathrm{d}_{\mathrm{GH}}((X, x_0), (Y, y_0))}_{\text{(adaptation of } \mathrm{d}_{\mathrm{GH}} \text{ to pointed spaces)}}$.

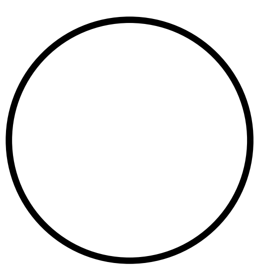


Stability

Theorem: [Carrière, O., Ovsjanikov 2015]
Let (X, d_X) and (Y, d_Y) be compact Riemannian manifolds. Let $x_0 \in X$ and $y_0 \in Y$. If $\mathrm{d}_{\mathrm{GH}}((X, x_0), (Y, y_0)) \leq \frac{1}{20} \min\{\varrho(X), \ \varrho(Y)\}$, then $\mathrm{d}_{\mathrm{B}}^{\infty}(\operatorname{dgm} \mathrm{d}_{x_0}, \operatorname{dgm} \mathrm{d}_{y_0}) \leq 20 \ \mathrm{d}_{\mathrm{GH}}((X, x_0), (Y, y_0))$.

Prerequisite: $\mathrm{d}_{\mathrm{GH}}(X, Y) < \frac{1}{20} \ \min\{\varrho(X), \ \varrho(Y)\}$

$X =$



$Y =$

$$\mathrm{d}_{\mathrm{GH}}(X, Y) < \infty = \varrho(Y)$$

$$\mathrm{d}_{\mathrm{B}}^{\infty}(\operatorname{dgm} f, \operatorname{dgm} g) = \infty$$

Application: Unsupervised Segmentation

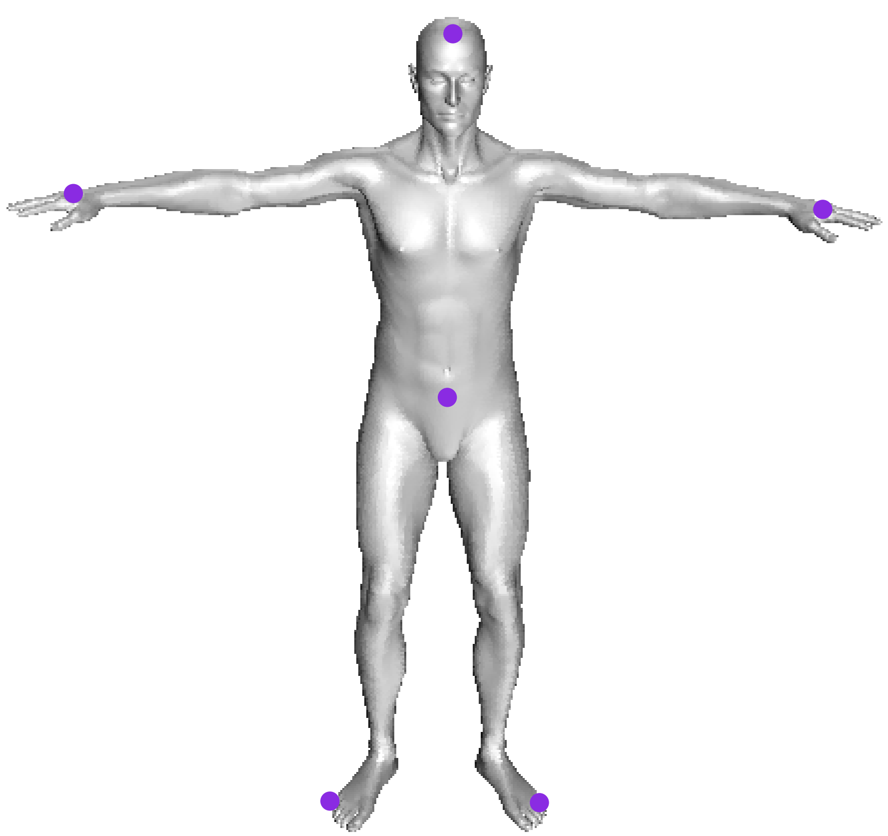
Experimental results:

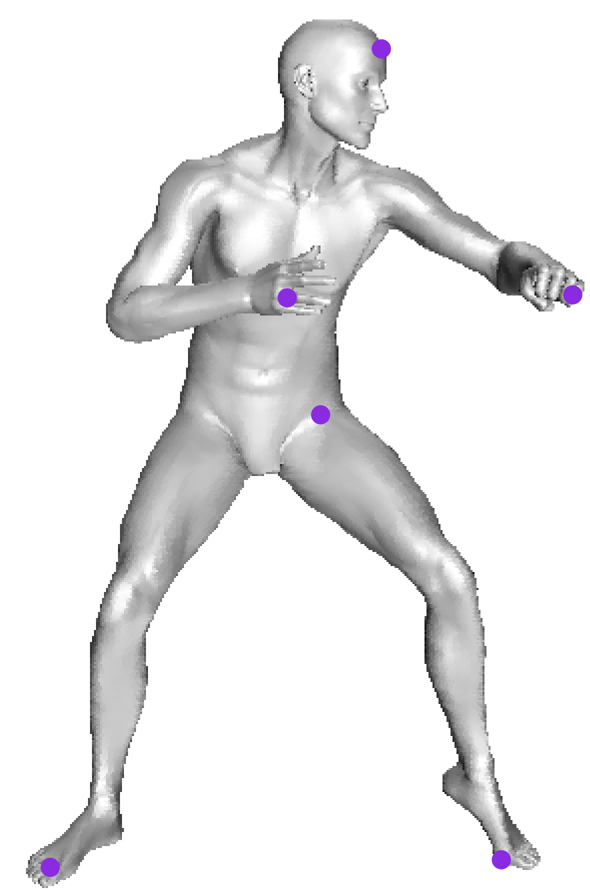
- input: shapes from the TOSCA database, in *mesh* form

- select a few base points by hand on each shape

- approximate geodesic distances to base points using the 1-skeleton graph

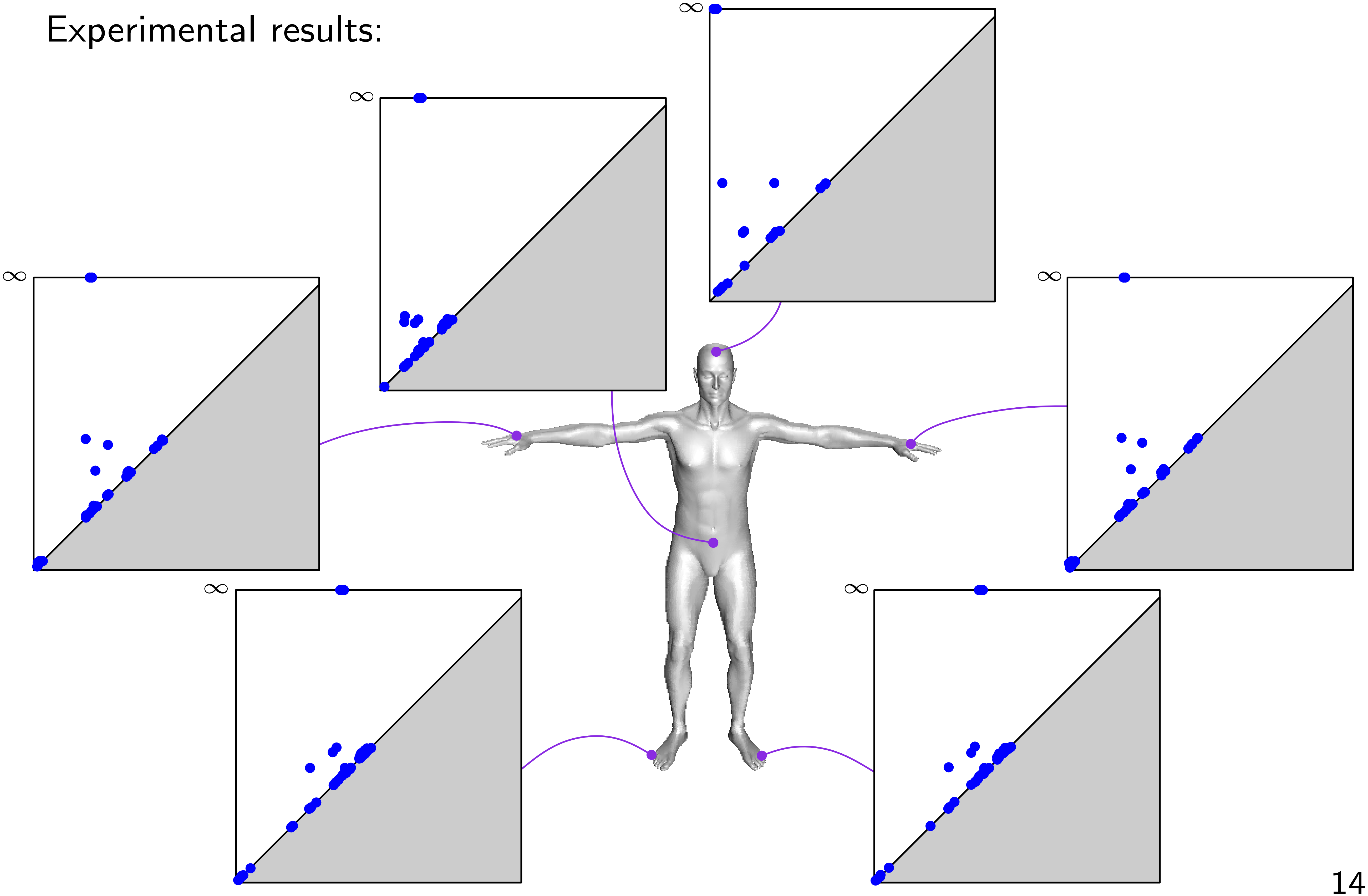
- use the PDs of the PL interpolations over the meshes as signatures





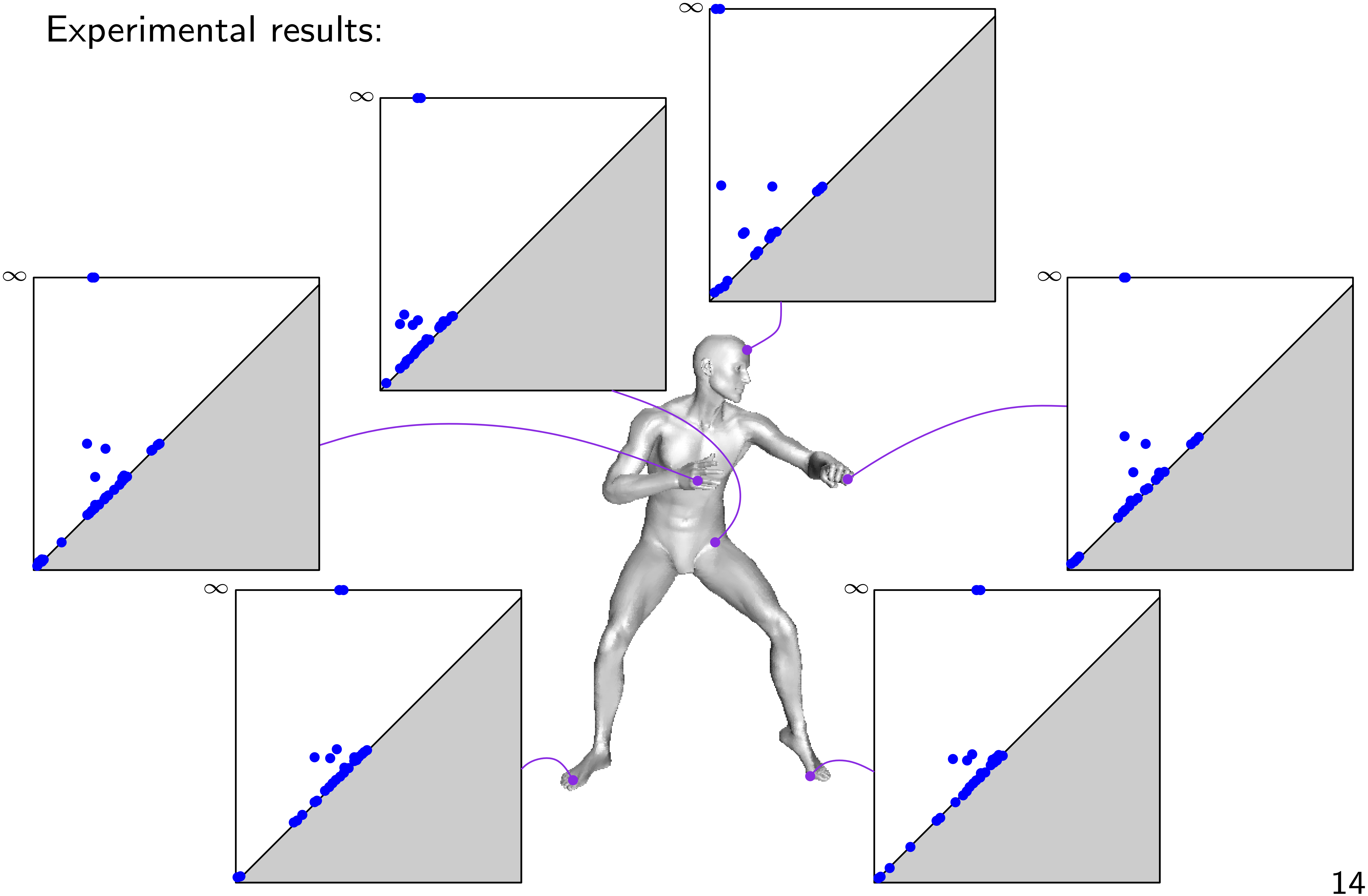
Application: Unsupervised Segmentation

Experimental results:



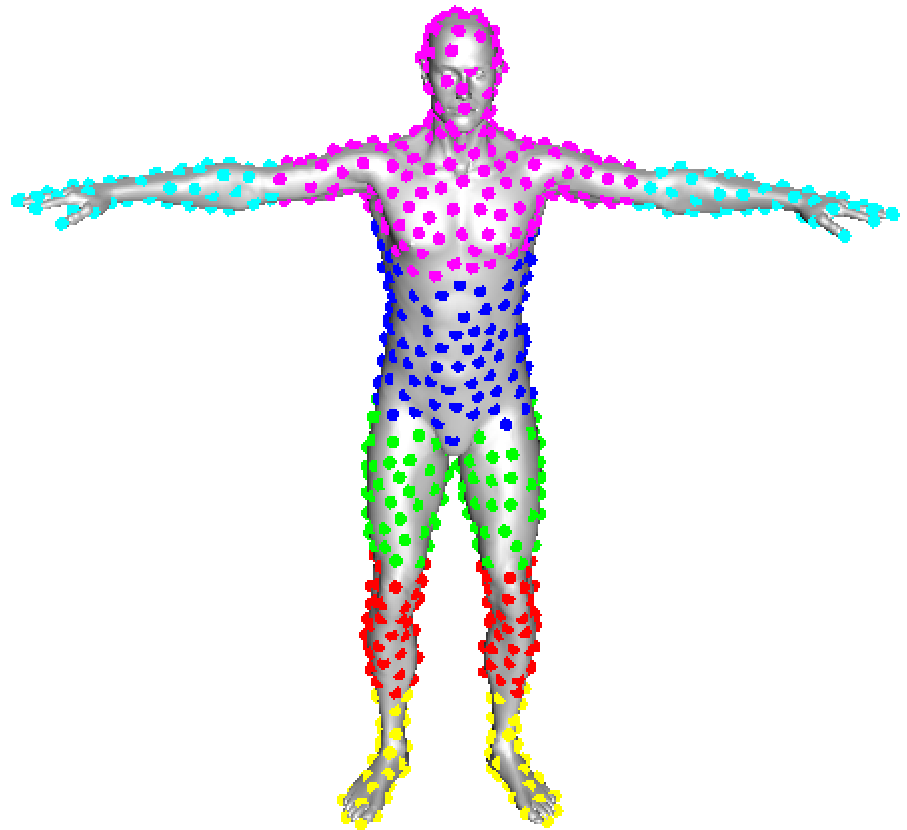
Application: Unsupervised Segmentation

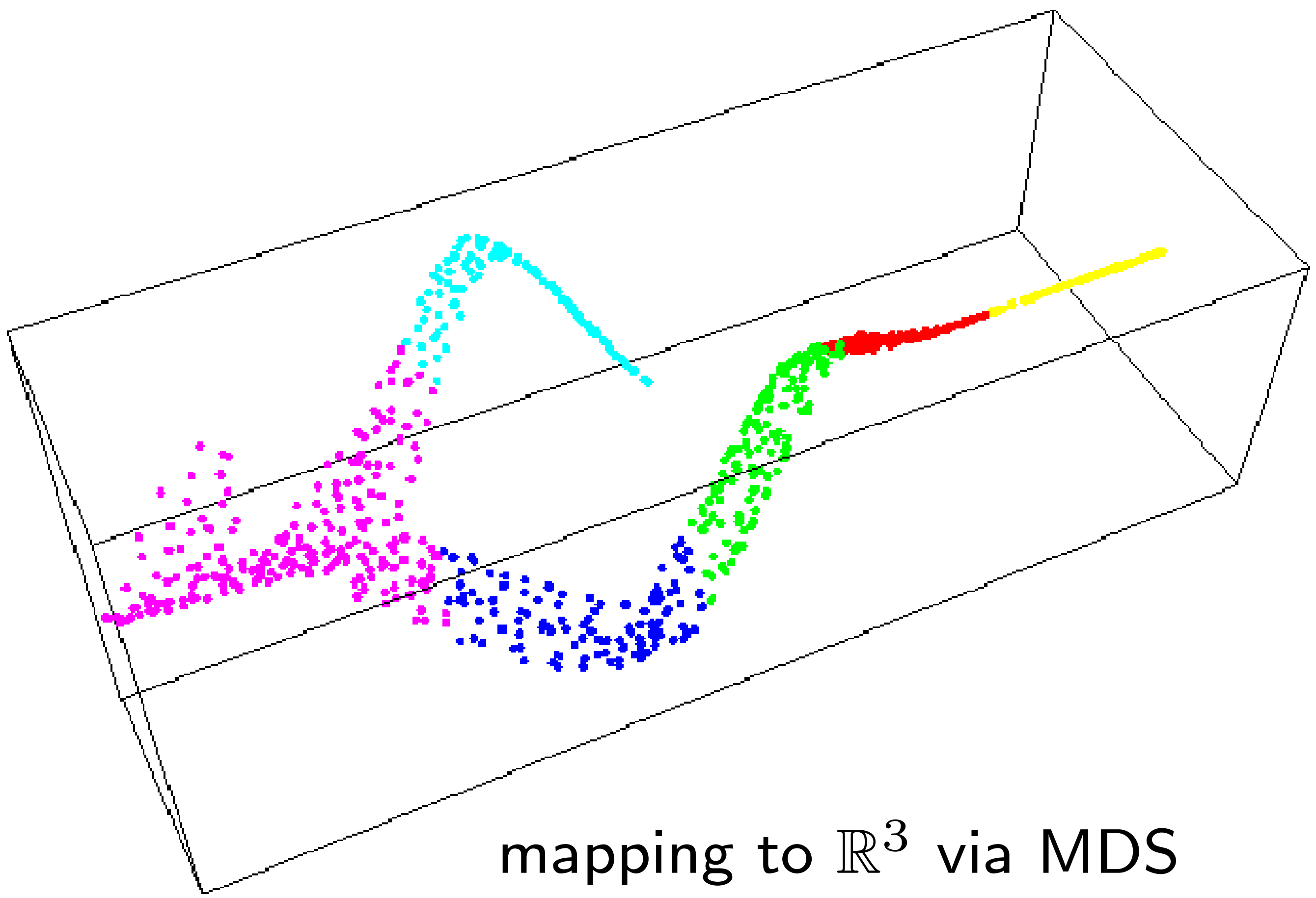
Experimental results:



Application: Unsupervised Segmentation

Experimental results:



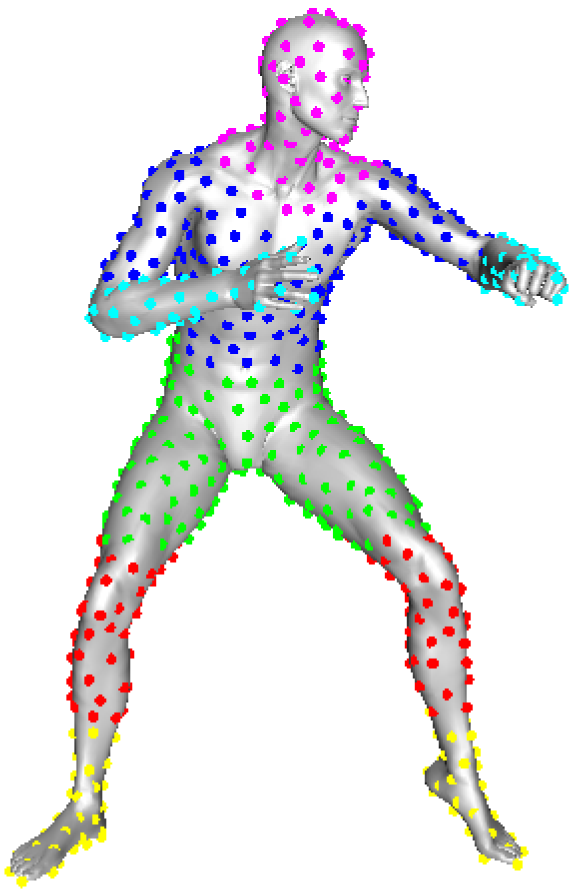


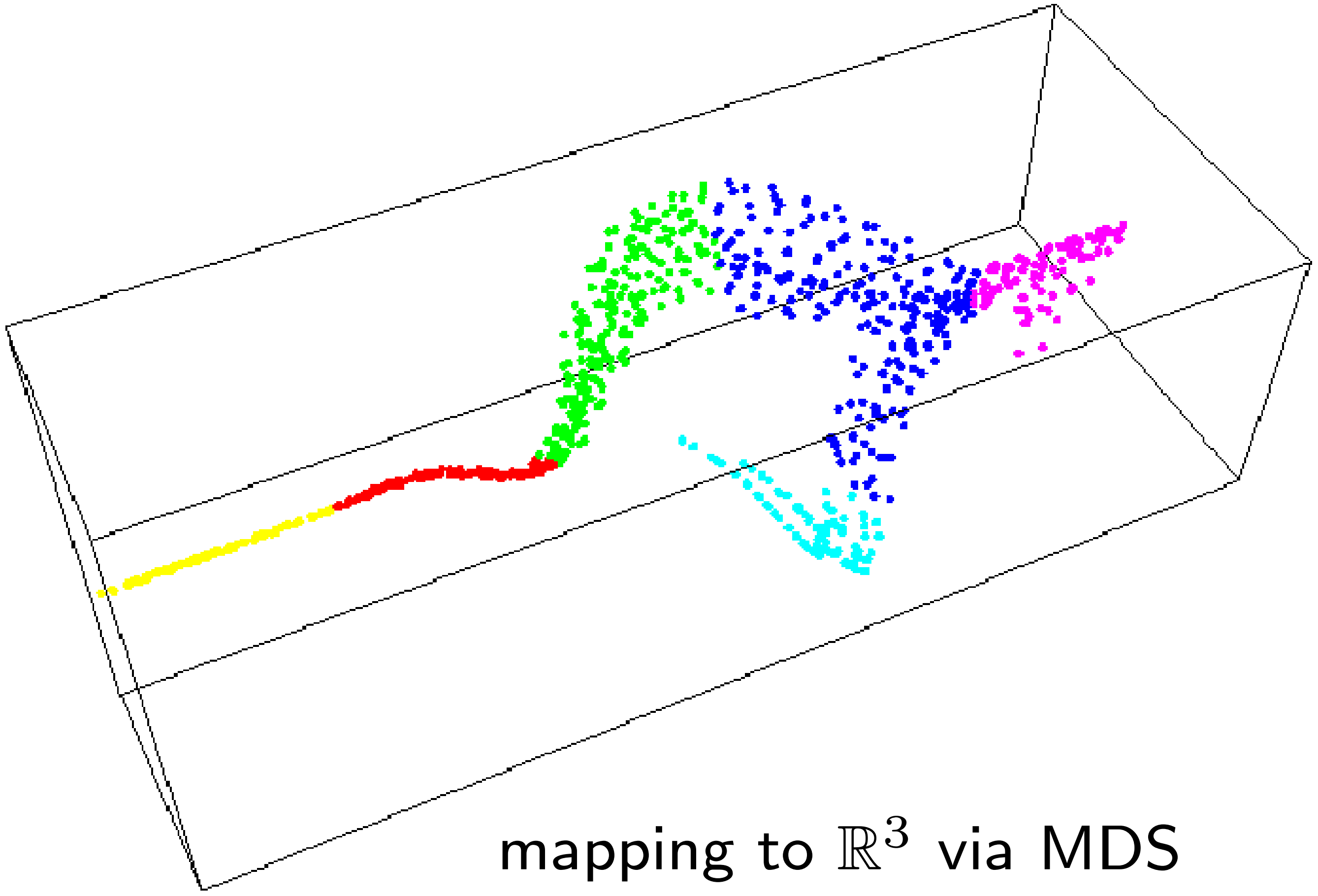
mapping to $\mathbb{R}^3$ via MDS

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

Application: Unsupervised Segmentation

Experimental results:





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

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

Outline

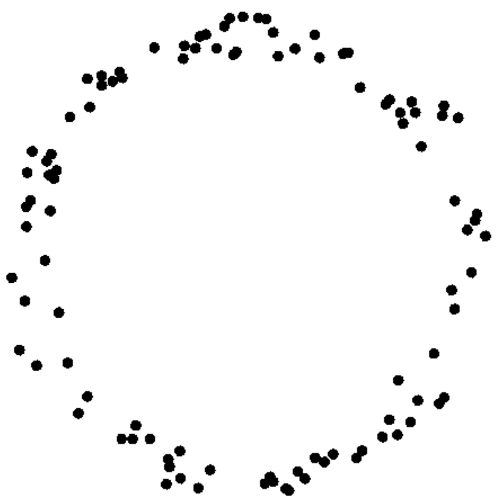
1. Global topological signatures

2. Local topological signatures

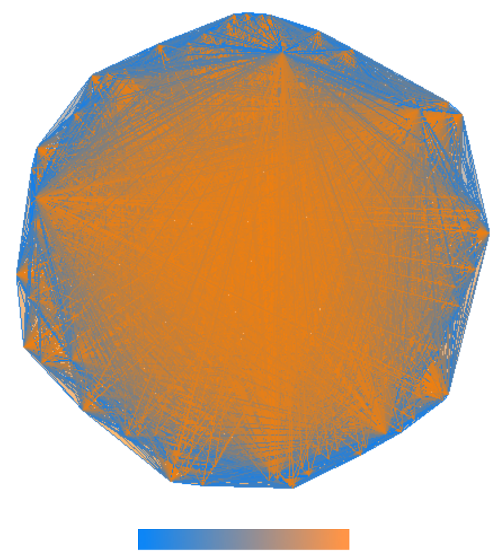
3. Kernels for topological signatures

Persistence Diagrams as Signatures

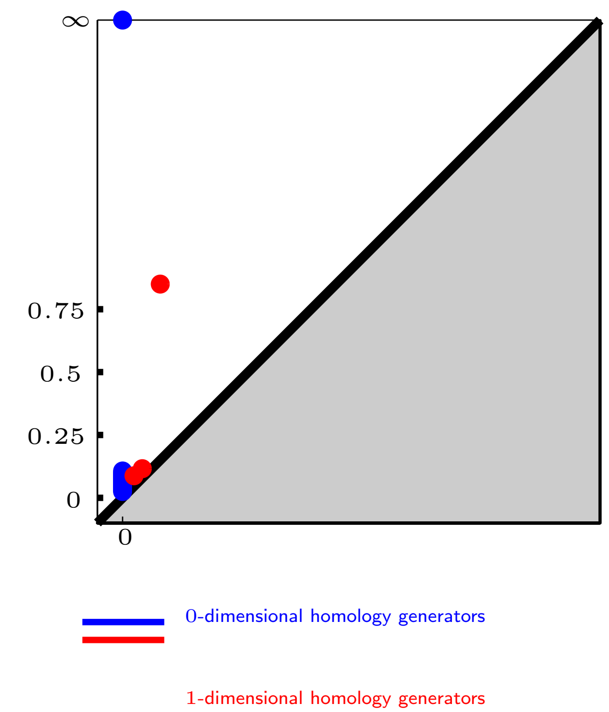
finite metric space / basepoint



filtration



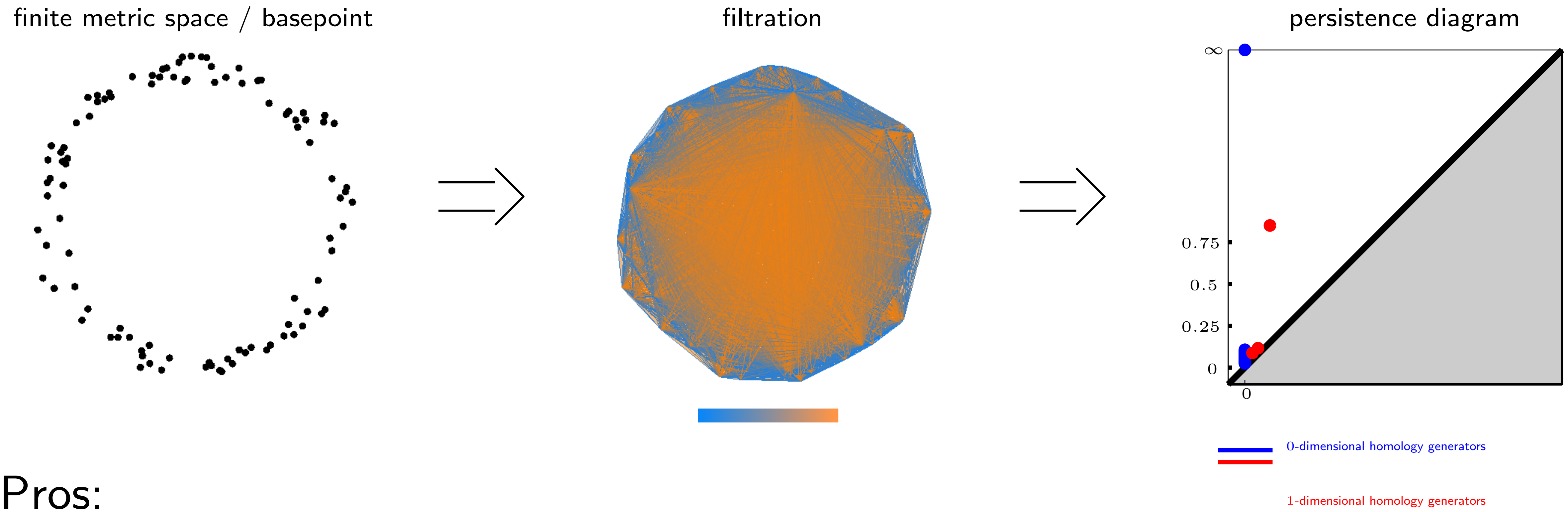
persistence diagram



Pros:

- topological signatures carry complementary information
- stability properties, e.g. $\mathrm{d}_{\mathrm{B}}^{\infty}(\mathcal{R}(X), \mathcal{R}(Y)) \leq 2\mathrm{d}_{\mathrm{GH}}(X, Y)$

Persistence Diagrams as Signatures



Pros:

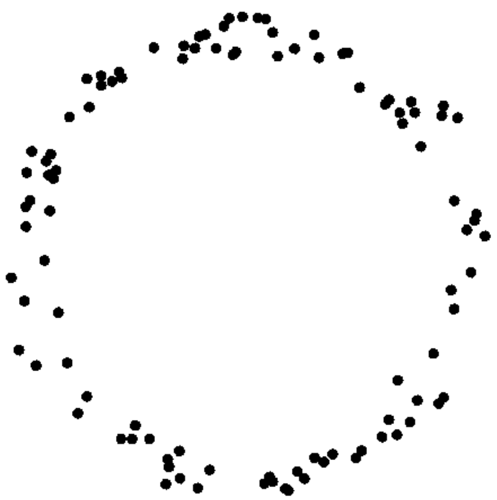
- topological signatures carry complementary information
- stability properties, e.g. $\mathrm{d}_{\mathrm{B}}^{\infty}(\mathcal{R}(X), \mathcal{R}(Y)) \leq 2\mathrm{d}_{\mathrm{GH}}(X, Y)$

Cons:

- the space of persistence diagrams is not a Hilbert space
- signatures are slow to compute and (more importantly) to compare

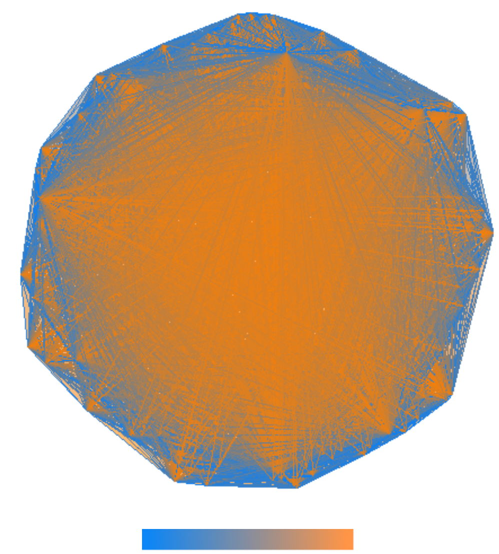
Persistence Diagrams as Signatures

finite metric space / basepoint



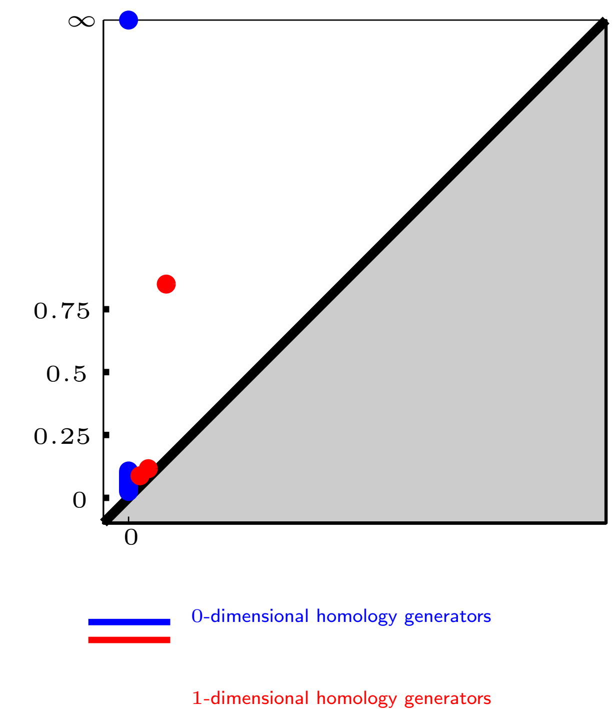
$\Longrightarrow$

filtration



$\Longrightarrow$

persistence diagram



Pros:

- topological signatures carry complementary information

- stability properties, e.g. $\mathrm{d}_{\mathrm{B}}^{\infty}(\mathcal{R}(X), \mathcal{R}(Y)) \leq 2\mathrm{d}_{\mathrm{GH}}(X, Y)$

Cons:

- the space of persistence diagrams is not a Hilbert space

  $\rightarrow$ define kernels on the space of diagrams

- signatures are slow to compute and (more importantly) to compare

  $\rightarrow$ explicit mapping to feature space

The Kernel Trick

$\mathcal{X}$: be a space in which we want to compare/classify elements

- feature map $\phi : \mathcal{X} \to \mathcal{H}$ equipped with inner product $< \cdot, \cdot >_{\mathcal{H}}$

- lift training/testing data to $\mathcal{H}$ through ϕ then solve learning problem

The Kernel Trick

$\mathcal{X}$: be a space in which we want to compare/classify elements

- feature map $\phi : \mathcal{X} \to \mathcal{H}$ equipped with inner product $< \cdot, \cdot >_{\mathcal{H}}$
- lift training/testing data to $\mathcal{H}$ through ϕ then solve learning problem
- observation: many learning methods use only inner product

  $\to$ do not lift the data, instead compute the $k(x, y) =< \phi(x), \phi(y) >_{\mathcal{H}}$

The Kernel Trick

$\mathcal{X}$: be a space in which we want to compare/classify elements

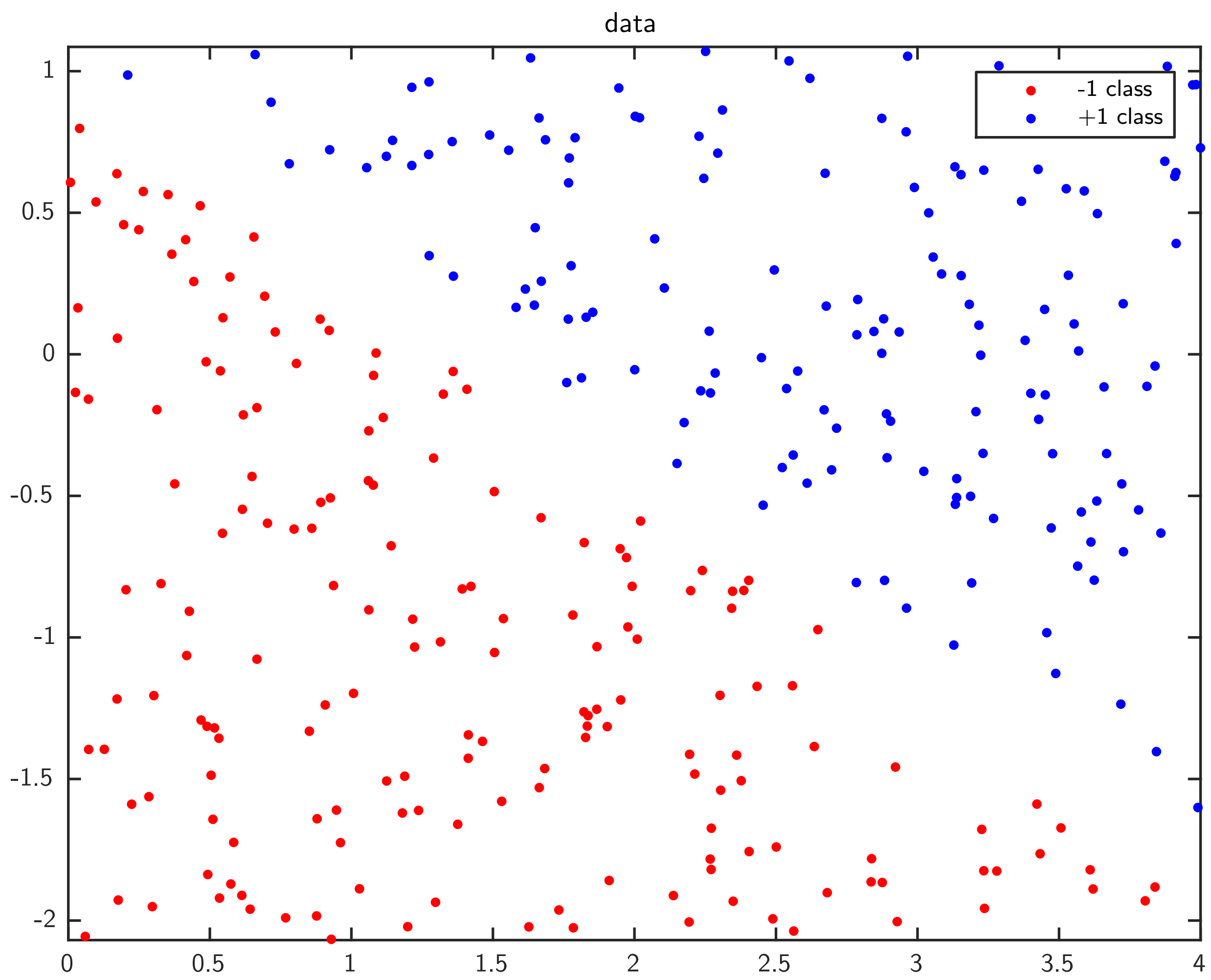
- feature map $\phi : \mathcal{X} \to \mathcal{H}$ equipped with inner product $< \cdot, \cdot >_{\mathcal{H}}$
- lift training/testing data to $\mathcal{H}$ through ϕ then solve learning problem
- observation: many learning methods use only inner product

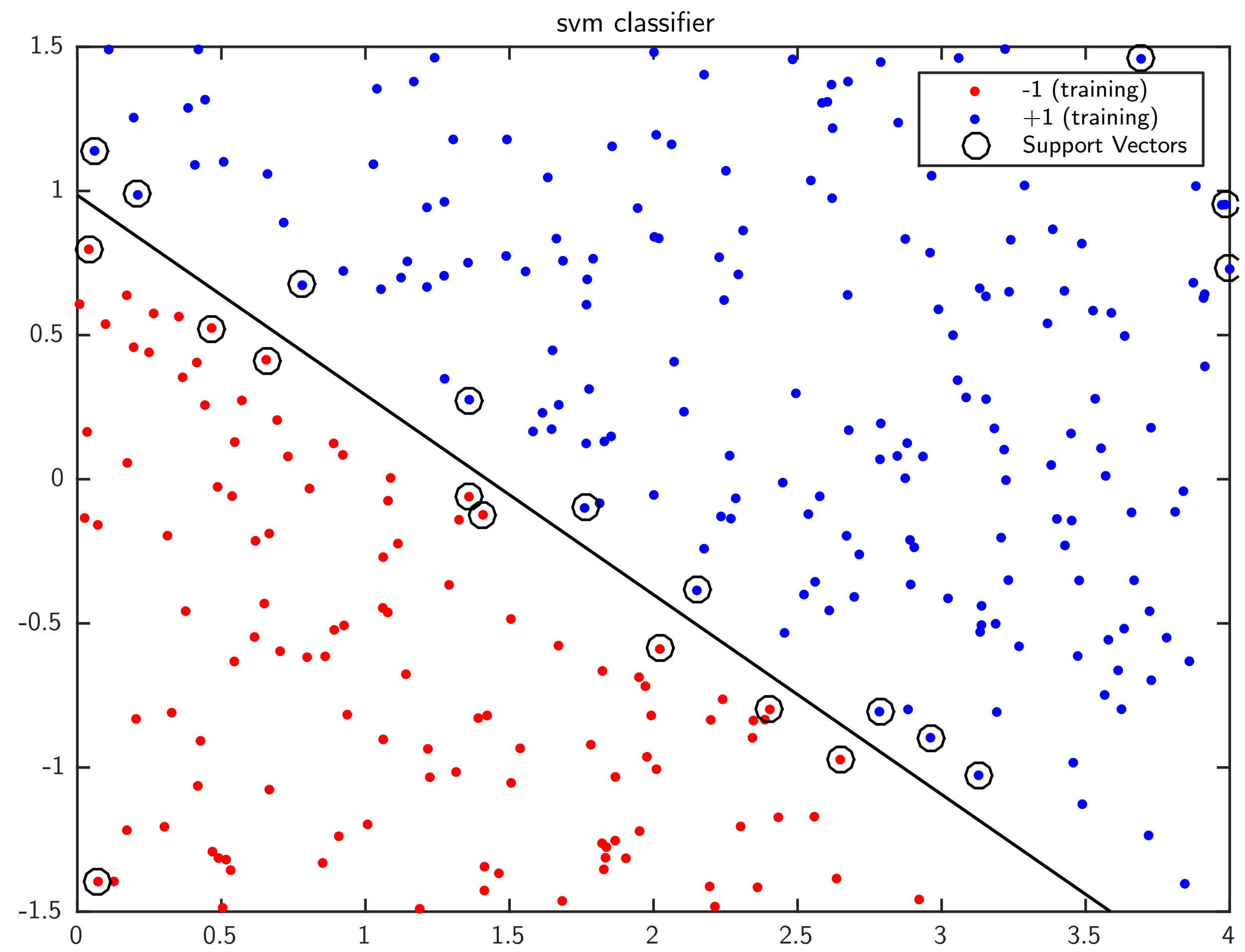
    $\to$ do not lift the data, instead compute the $k(x, y) =< \phi(x), \phi(y) >_{\mathcal{H}}$

Def.: A *reproducing kernel* is a map $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ such that $k(\cdot, \cdot) =< \phi(\cdot), \phi(\cdot) >_{\mathcal{H}}$ for some pair $(\phi, \mathcal{H})$.

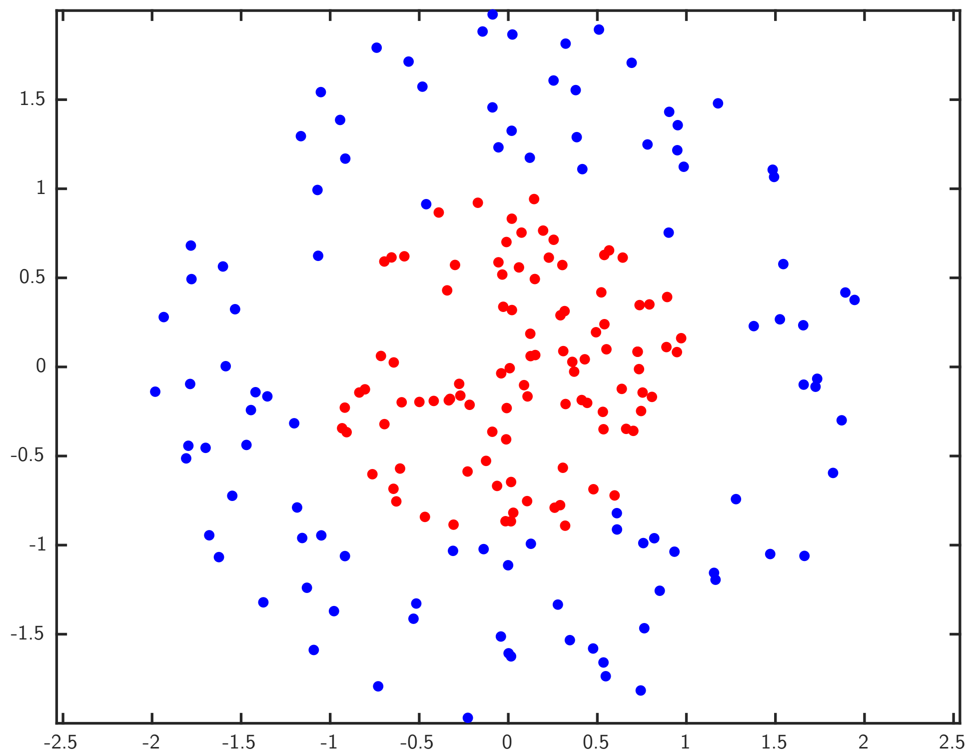
Thm.: [Moore, Aronszajn]
A pair $(\phi, \mathcal{H})$ exists whenever k is *positive semidefinite*, i.e.
$\sum_{i,j=1}^{n} c_i c_j \, k(x_i, x_j) \geq 0$ for all $n \in \mathbb{N}$, $c_1, \cdots, c_n \in \mathbb{R}$, and $x_1, \cdots, x_n \in \mathcal{X}$.

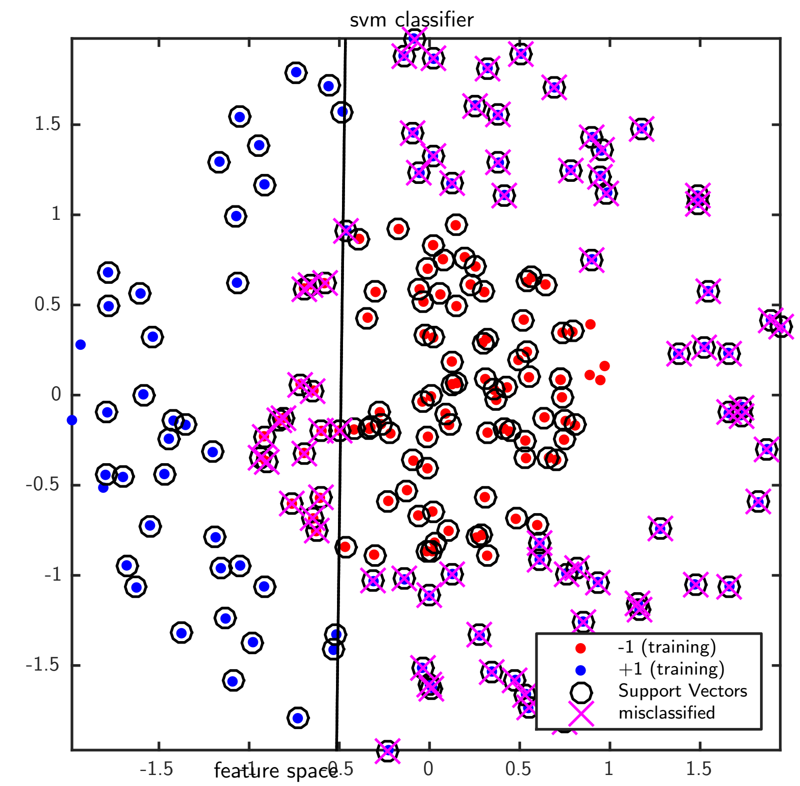
The Kernel Trick

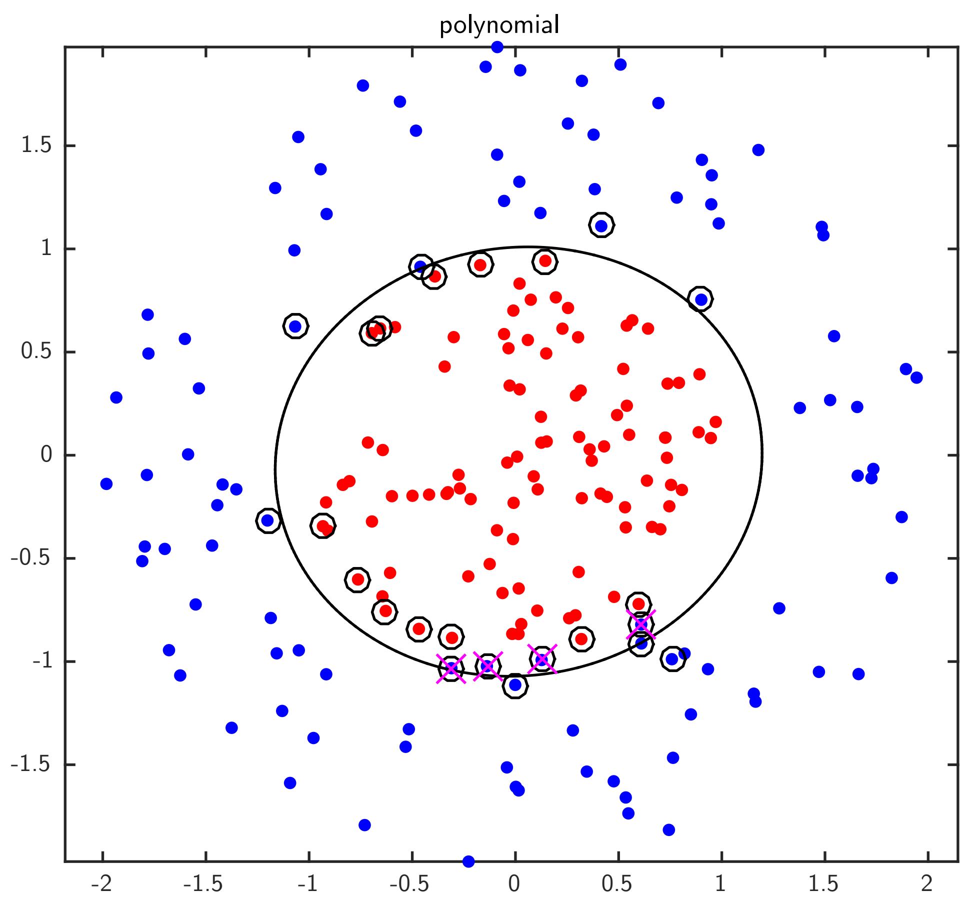


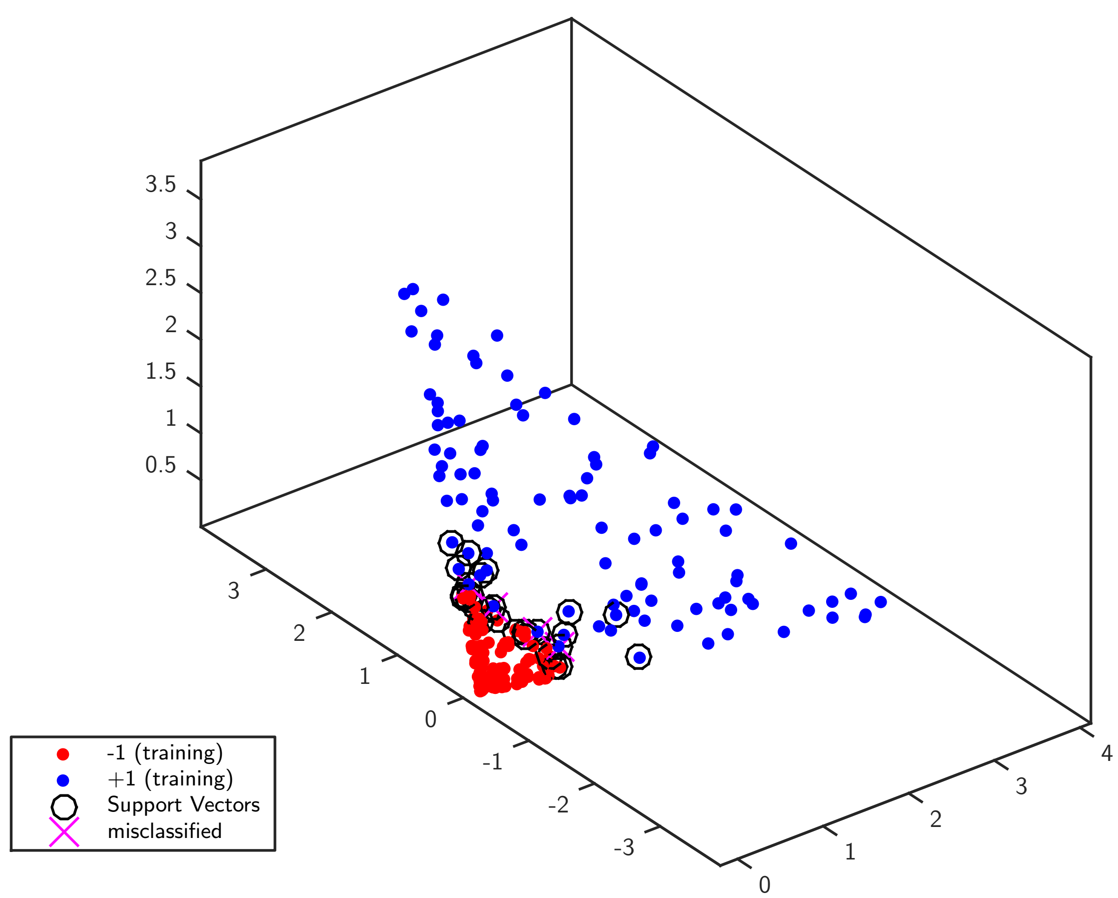


The Kernel Trick









Kernels for Persistence Diagrams

View persistence diagrams as:

- landscapes [Bubenik 2012] [Bubenik, Dłotko 2015]
- empirical measures:
  - $\rightarrow$ histogram [Bendich et al. 2014]
  - $\rightarrow$ density estimator [Chepushtanova et al. 2015]
  - $\rightarrow$ heat diffusion [Bauer et al. 2015]
- metric spaces [Carrière, O., Ovsjanikov 2015]
- roots of polynomials [Di Fabio, Ferri 2015]

Side-by-side Comparison

landscapes

[Bubenik 2012]

feature space: $L^2(\mathbb{N} \times \mathbb{R}) \to L^2(\mathbb{R}^2)$

feature map: explicit (comb. construction)

kernel(s): linear, Gaussian, etc.

complexity on n-points diagrams:

- feature map: $O(n^2)$
- kernel: $O(n^2)$

stability: $\|\cdot\|_\infty \leq O(\mathrm{d}_{\mathrm{B}}^{\infty})$

$\|\cdot\|_p \leq O(\mathrm{Pers}\ \mathrm{d}_{\mathrm{B}}^{\infty}(\cdot))$

injective feature map

$\to$ pd kernel?

kernels on diagrams are not additive

empirical measures

[Bauer et al. 2015]

feature space: $L^2(\mathbb{R}^2)$

feature map: explicit (closed form solution)

kernel(s): k_σ

complexity on n-points diagrams:

- feature map: N/A (or discretize...)
- kernel: $O(n^2)$

stability: $\|\cdot\|_2 \leq O(\mathrm{d}_{\mathrm{B}}^{1}(\cdot))$

injective feature map

$\to$ pd kernel?

kernel is additive

metric spaces

[Carrière, O. , Ovsjanikov 2015]

feature space: $\ell^p \to (\mathbb{R}^D, \ell^p)$

feature map: explicit (comb. construction)

kernel(s): linear Gaussian, etc.

complexity on n-points diagrams:

- feature map: $O(n^2)$
- kernel: $O(D)$

stability: $\|\cdot\|_\infty \leq O(\mathrm{d}_{\mathrm{B}}^{\infty})$

$\|\cdot\|_p \leq O(D^{1/p}\ \mathrm{d}_{\mathrm{B}}^{\infty}(\cdot))$

non-injective feature map

$\to$ psd kernel

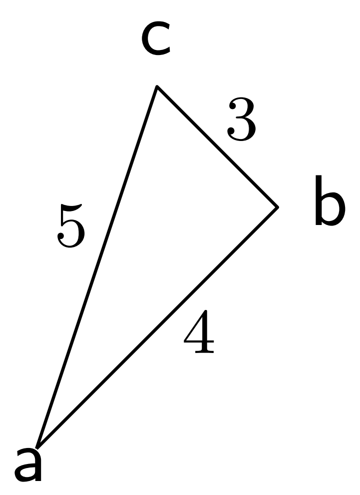
kernels on diagrams are not additive

Side-by-side Comparison

landscapes	empirical measures	metric spaces
[Bubenik 2012]	[Bauer et al. 2015]	[Carrière, O. , Ovsjanikov 2015]
feature space: $L^2(\mathbb{N} \times \mathbb{R}) \to L^2(\mathbb{R}^2)$	feature space: $L^2(\mathbb{R}^2)$	feature space: $\ell^p \to (\mathbb{R}^D, \ell^p)$
feature map: explicit (comb. construction)	feature map: explicit (closed form solution)	feature map: explicit (comb. construction)
kernel(s): linear, Gaussian, etc.	kernel(s): k_σ	kernel(s): linear Gaussian, etc.
complexity on n-points diagrams: - feature map: $O(n^2)$ - kernel: $O(n^2)$	complexity on n-points diagrams: - feature map: N/A (or discretize...) - kernel: $O(n^2)$	complexity on n-points diagrams: - feature map: $O(n^2)$ - kernel: $O(D)$
stability: $\\|\cdot\\|_\infty \leq O(\mathrm{d}_\mathrm{B}^\infty)$ $\\|\cdot\\|_p \leq O(\mathrm{Pers}\ \mathrm{d}_\mathrm{B}^\infty(\cdot))$	stability: $\\|\cdot\\|_2 \leq O(\mathrm{d}_\mathrm{B}^1(\cdot))$	stability: $\\|\cdot\\|_\infty \leq O(\mathrm{d}_\mathrm{B}^\infty)$ $\\|\cdot\\|_p \leq O(D^{1/p}\ \mathrm{d}_\mathrm{B}^\infty(\cdot))$
injective feature map $\longrightarrow$ pd kernel?	injective feature map $\longrightarrow$ pd kernel?	non-injective feature map $\longrightarrow$ psd kernel
kernels on diagrams are not additive	kernel is additive	kernels on diagrams are not additive

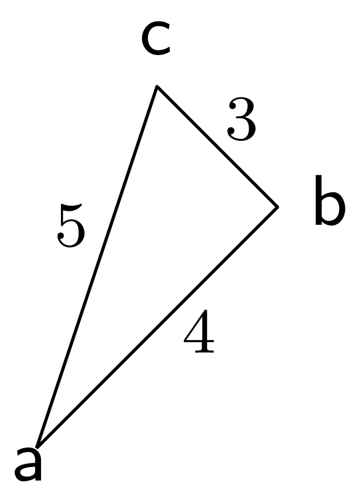
Feature Map

finite metric space



Feature Map

finite metric space



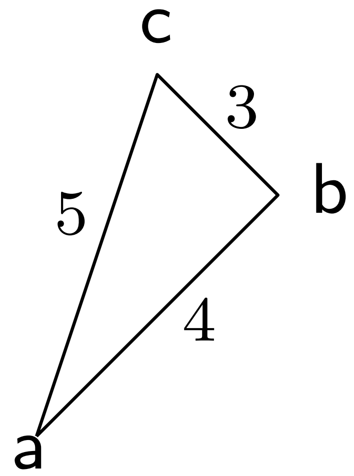
$\xrightarrow{\phi_1}$

distance matrix

$$\begin{array}{c c} & \begin{array}{ccc} \mathsf{a} & \mathsf{b} & \mathsf{c} \end{array} \\ \begin{array}{c} \mathsf{a} \\ \mathsf{b} \\ \mathsf{c} \end{array} & \begin{bmatrix} 0 & 4 & 5 \\ 4 & 0 & 3 \\ 5 & 3 & 0 \end{bmatrix} \end{array}$$

Feature Map

finite metric space

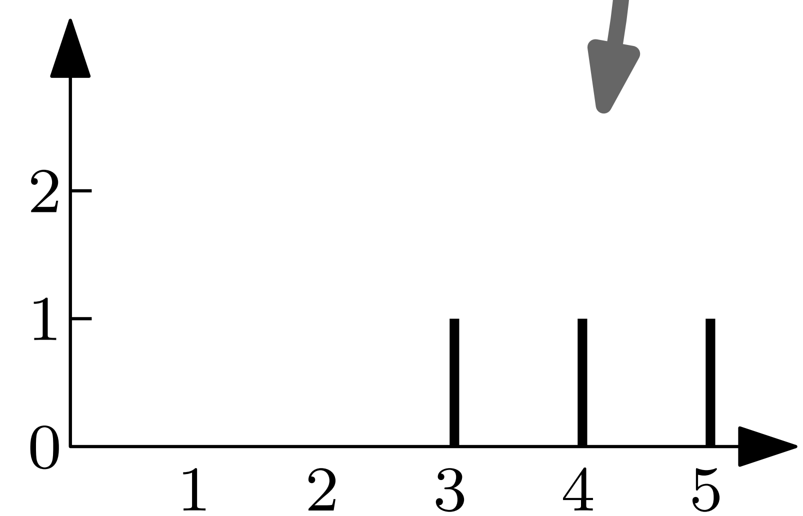


ϕ_1

distance matrix

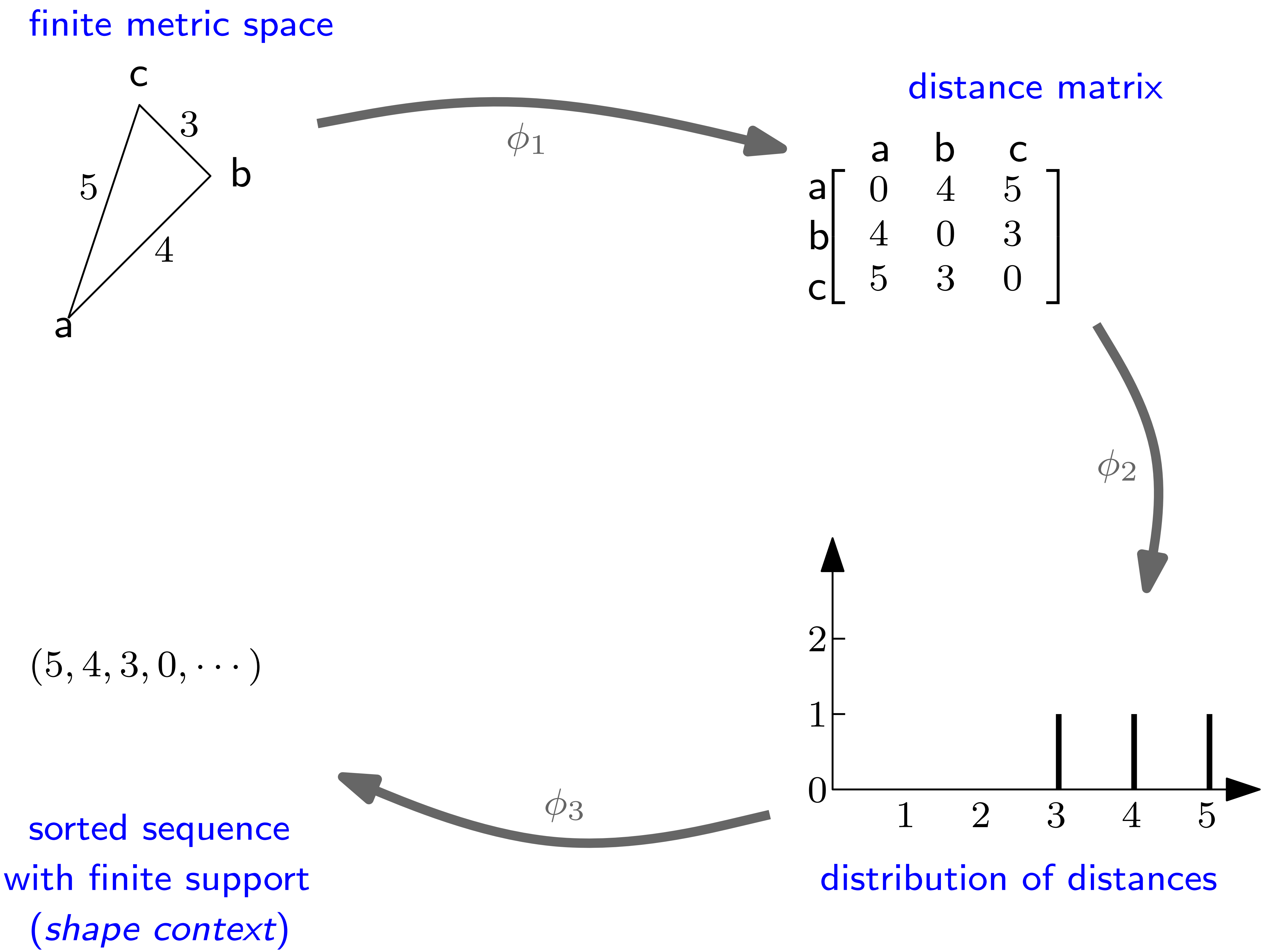
$$\begin{array}{c|ccc} & \text{a} & \text{b} & \text{c} \\ \hline \text{a} & 0 & 4 & 5 \\ \text{b} & 4 & 0 & 3 \\ \text{c} & 5 & 3 & 0 \end{array}$$

ϕ_2



distribution of distances

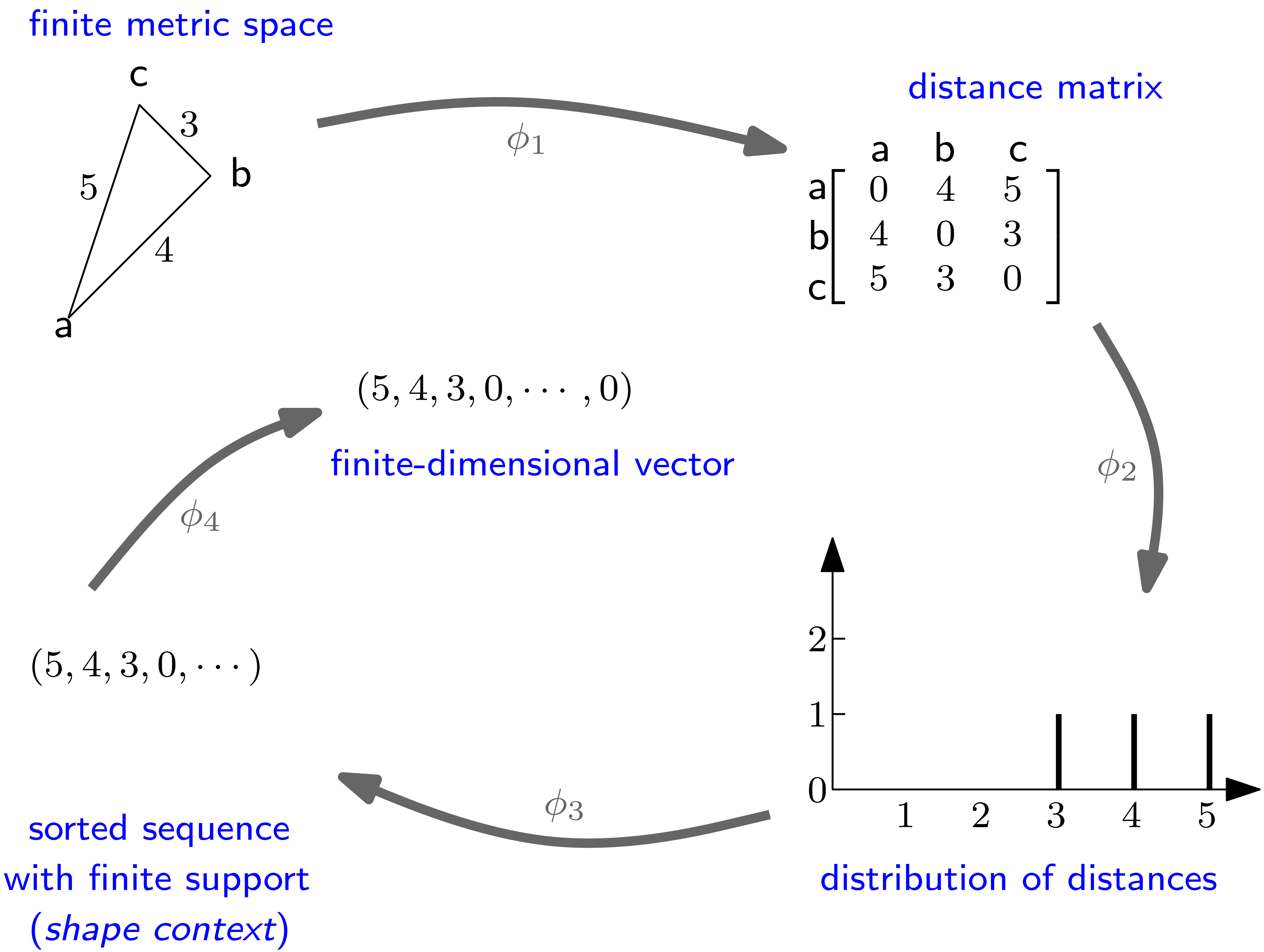
Feature Map



finite metric space
c
3
b
5
4
a
ϕ_1
distance matrix
a b c
a 0 4 5
b 4 0 3
c 5 3 0
ϕ_2
2
1
0
1 2 3 4 5
distribution of distances
ϕ_3
$(5, 4, 3, 0, \cdots)$
sorted sequence
with finite support
(*shape context*)

Feature Map

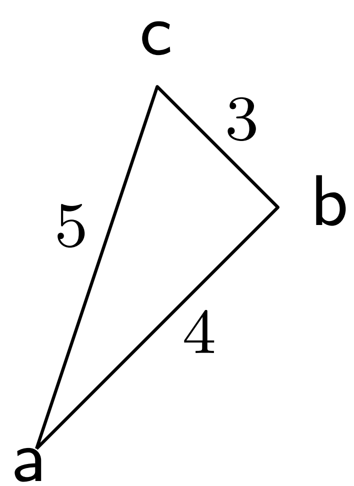
$$\Phi = \phi_4 \circ \phi_3 \circ \phi_2 \circ \phi_1$$



Stability Properties

$\Phi = \phi_4 \circ \phi_3 \circ \phi_2 \circ \phi_1$

finite metric space $\in \mathbf{P}_\infty(\mathbb{R}^2)$

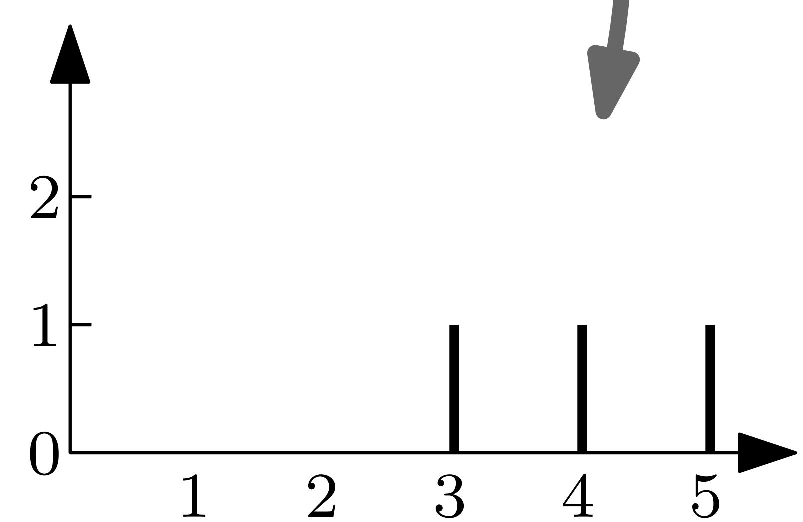


$\xrightarrow{\phi_1}$

distance matrix

$$
\begin{array}{c|ccc}
 & \text{a} & \text{b} & \text{c} \\
\hline
\text{a} & 0 & 4 & 5 \\
\text{b} & 4 & 0 & 3 \\
\text{c} & 5 & 3 & 0
\end{array}
$$

$\xrightarrow{\phi_2}$



distribution of distances $\in \mathbf{P}_\infty(\mathbb{R})$

$W_\infty(P,Q) = \inf_{m:P\to Q} \sup_{p\in P} \|p - m(p)\|_\infty$

$\xrightarrow{\phi_3}$

$(5, 4, 3, 0, \cdots)$ $\in \ell^\infty$

sorted sequence
with finite support
(*shape context*)

$\xrightarrow{\phi_4}$

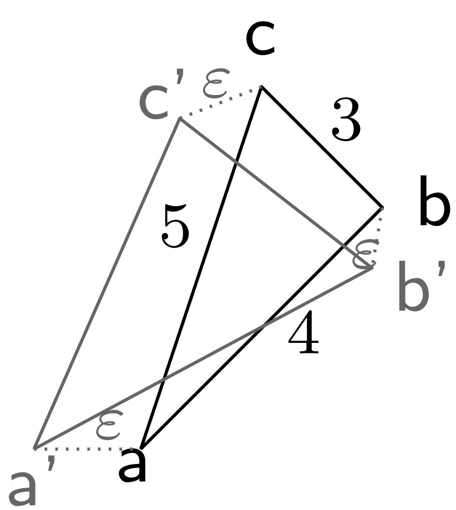
$(5, 4, 3, 0, \cdots, 0)$ $\in (\mathbb{R}^D, \ell^\infty)$

finite-dimensional vector

Stability Properties

$\Phi = \phi_4 \circ \phi_3 \circ \phi_2 \circ \phi_1$

finite metric space $\in \mathbf{P}_\infty(\mathbb{R}^2)$



ϕ_1

distance matrix

$$\begin{array}{c} \\ a \\ b \\ c \end{array}\begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{bmatrix} 0 & 4 & 5 \\ 4 & 0 & 3 \\ 5 & 3 & 0 \end{bmatrix} \end{array} + \begin{bmatrix} \varepsilon_{aa} & \varepsilon_{ab} & \varepsilon_{ac} \\ \varepsilon_{ba} & \varepsilon_{bb} & \varepsilon_{bc} \\ \varepsilon_{ca} & \varepsilon_{cb} & \varepsilon_{cc} \end{bmatrix}$$

$\varepsilon_{xy} \in [-2\varepsilon, +2\varepsilon]$

ϕ_2

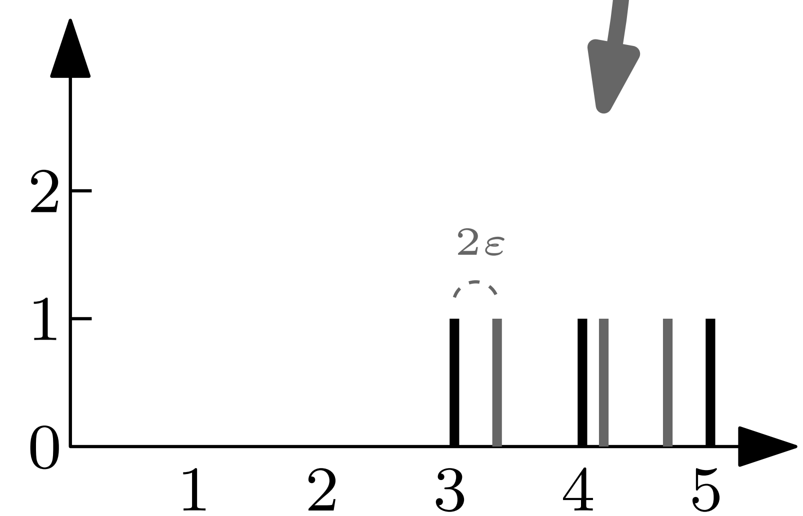
?

$(5 \pm 2\varepsilon, 4 \pm 2\varepsilon, 3 \pm 2\varepsilon, 0 \cdots, 0)$

$(5, 4, 3, 0, \cdots, 0) \quad \in (\mathbb{R}^D, \ell^\infty)$

finite-dimensional vector

ϕ_4



$(5, 4, 3, 0, \cdots) \quad \in \ell^\infty$

$(5 \pm 2\varepsilon, 4 \pm 2\varepsilon, 3 \pm 2\varepsilon, 0, \cdots)$

?

ϕ_3

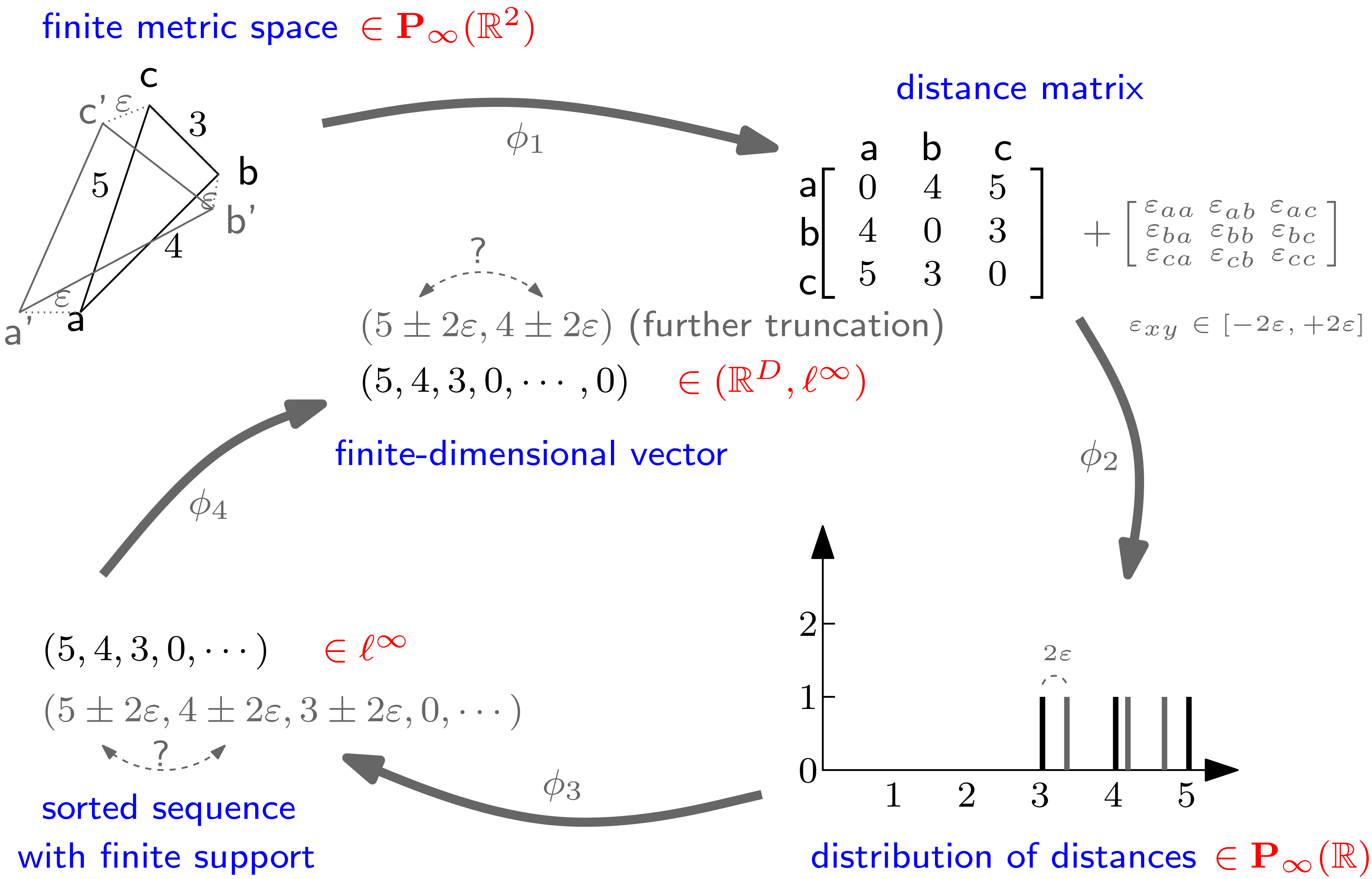
sorted sequence
with finite support
(*shape context*)

distribution of distances $\in \mathbf{P}_\infty(\mathbb{R})$

$W_\infty(P, Q) = \inf_{m : P \to Q} \sup_{p \in P} \|p - m(p)\|_\infty$

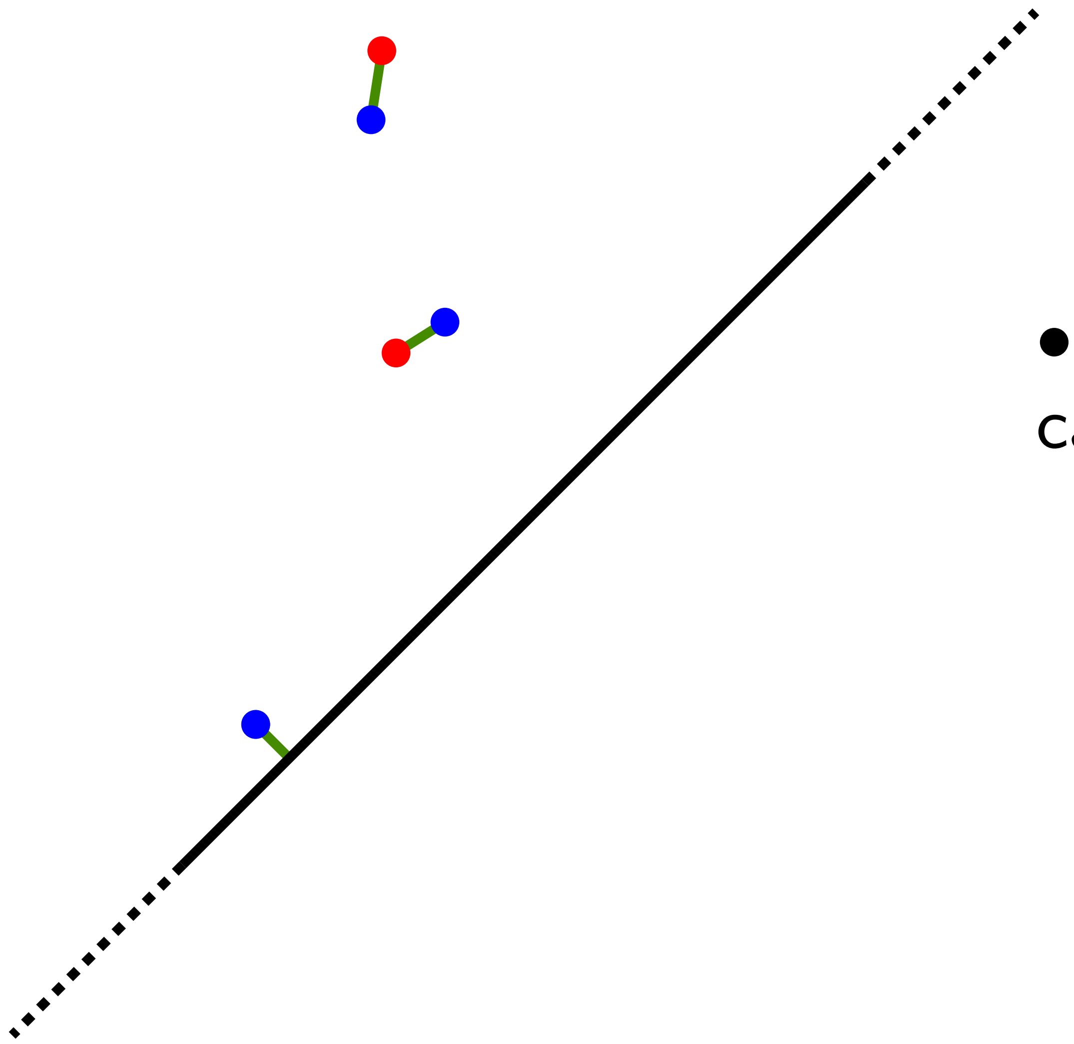
Stability Properties

$\Phi = \phi_4 \circ \phi_3 \circ \phi_2 \circ \phi_1$



$W_\infty(P, Q) = \inf_{m: P \to Q} \sup_{p \in P} \|p - m(p)\|_\infty$

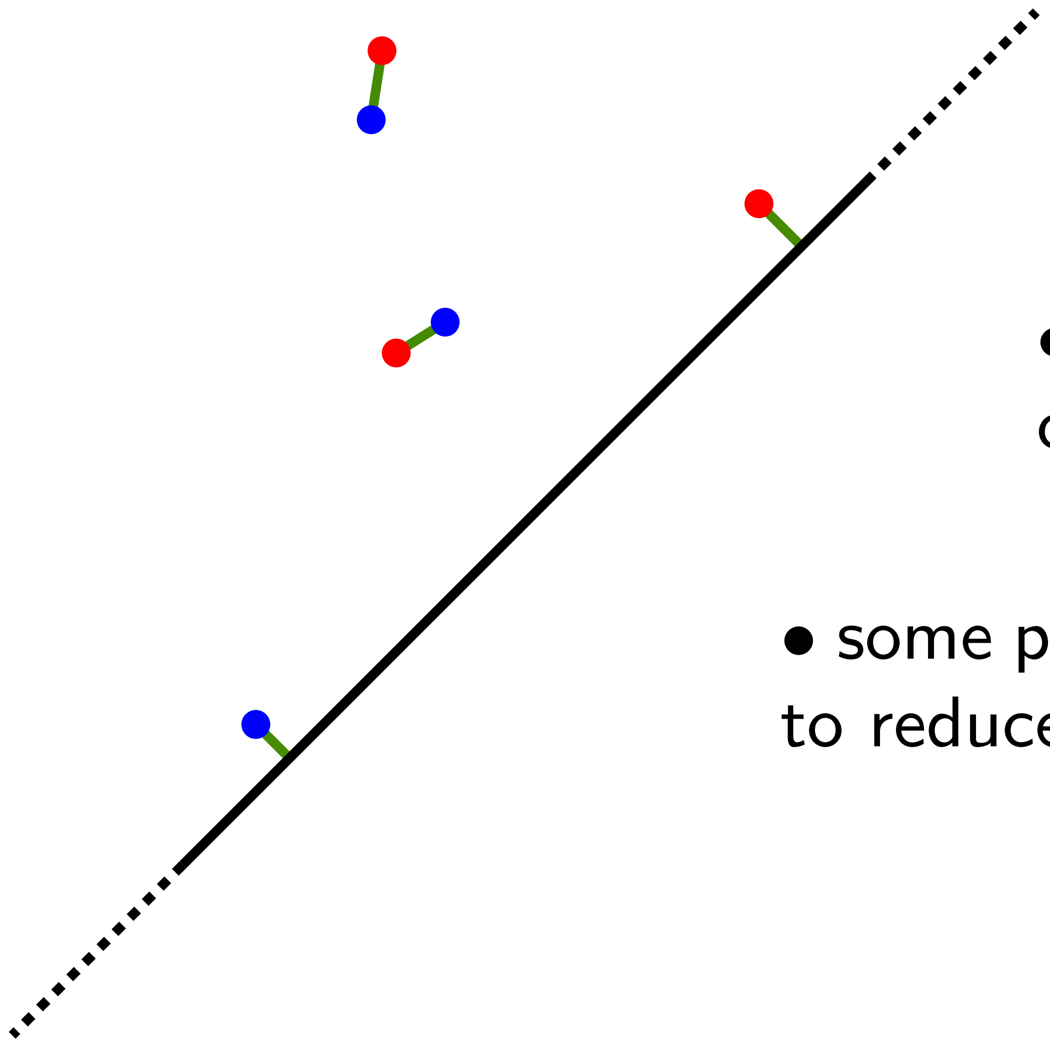
Adding the diagonal



- diagonal has infinite multiplicity

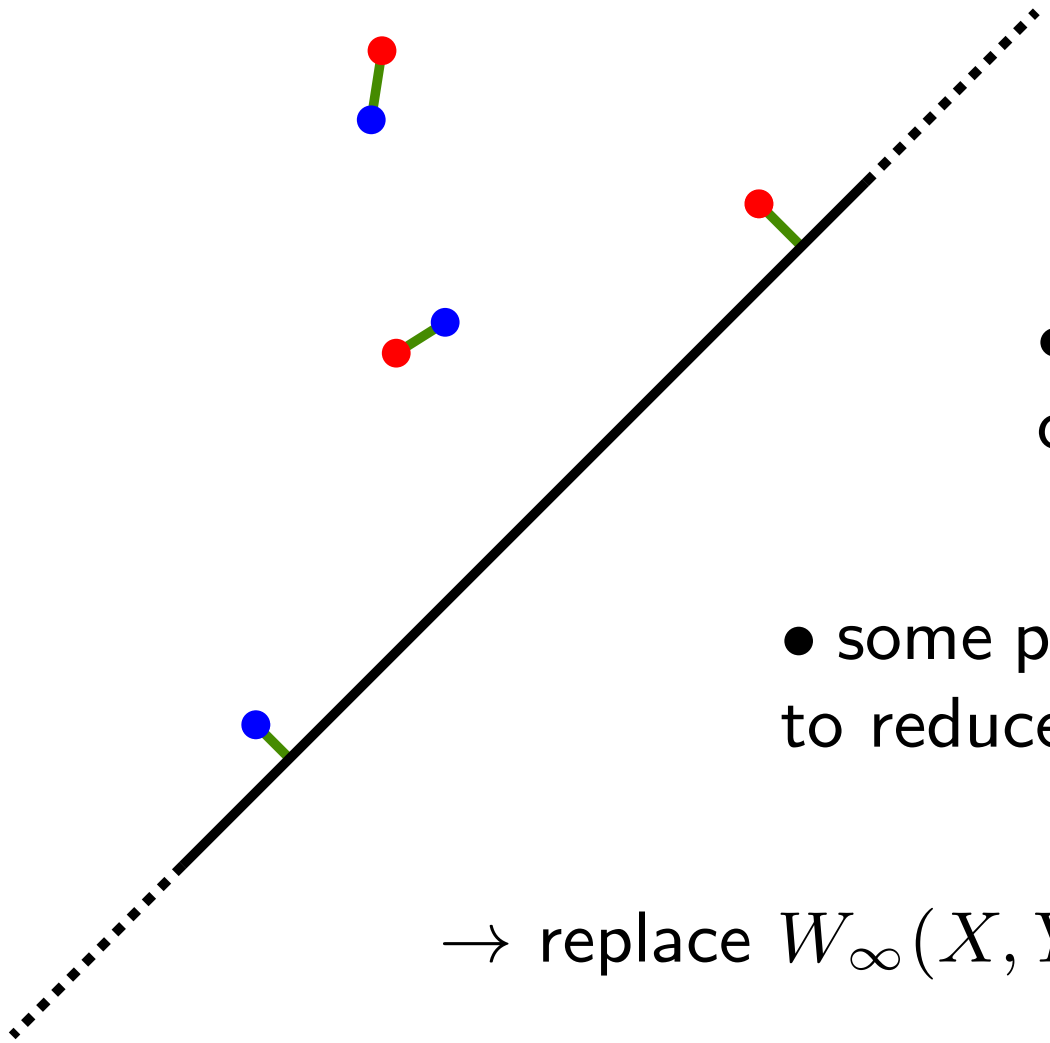
- useful for when point clouds have different cardinalities

Adding the diagonal



- diagonal has infinite multiplicity
- useful for when point clouds have different cardinalities
- some points may prefer the diagonal to other points to reduce the cost of the matching

Adding the diagonal



- diagonal has infinite multiplicity

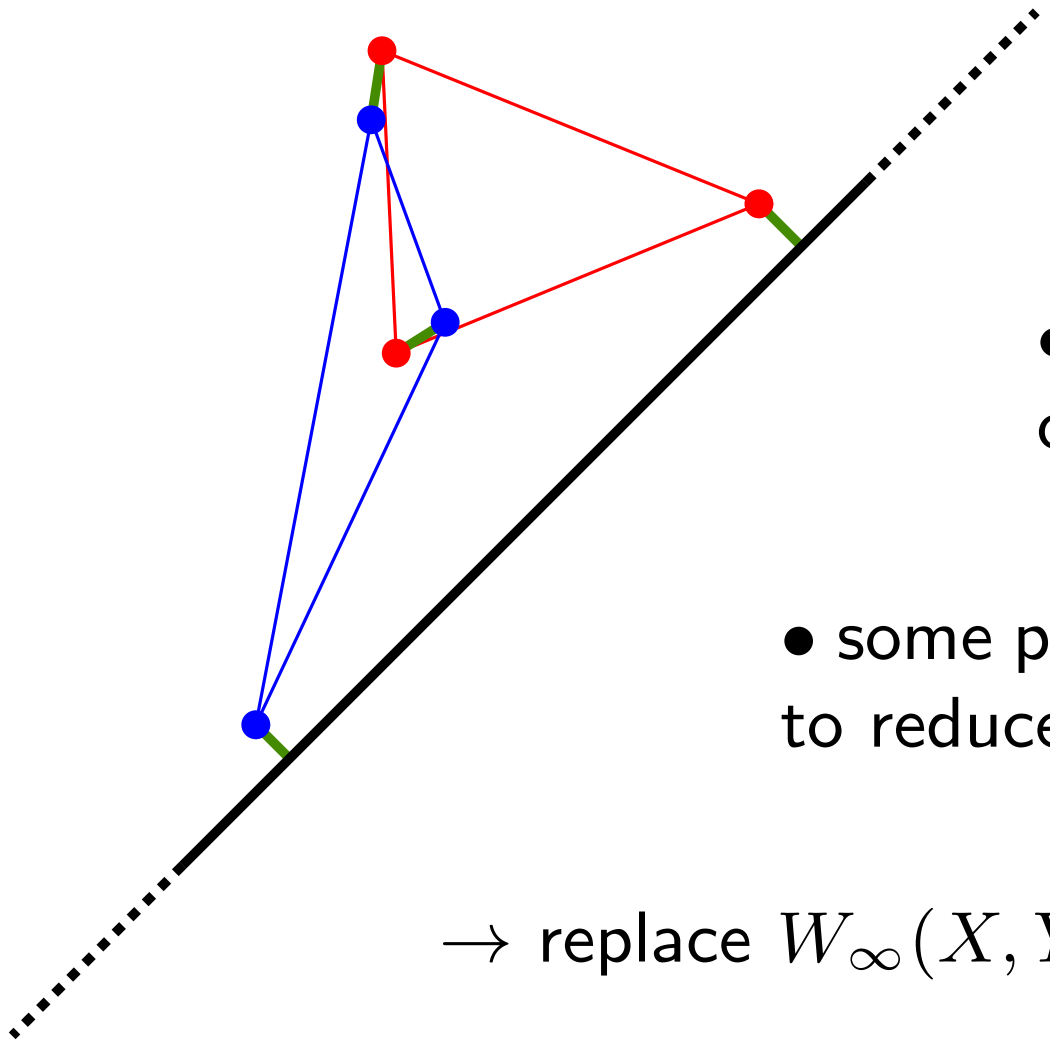
- useful for when point clouds have different cardinalities

- some points may prefer the diagonal to other points to reduce the cost of the matching

$\rightarrow$ replace $W_\infty(X,Y)$ with

$$\mathrm{d}_{\mathrm{B}}^{\infty} = \inf_{m : X \leftrightarrow Y} \max \left\{ \sup_{p \text{ matched}} \|p - m(p)\|_\infty, \sup_{\text{p unmatched}} \|p - \bar{p}\|_\infty \right\}$$

Adding the diagonal



- diagonal has infinite multiplicity
- useful for when point clouds have different cardinalities
- some points may prefer the diagonal to other points to reduce the cost of the matching

$\rightarrow$ replace $W_\infty(X,Y)$ with

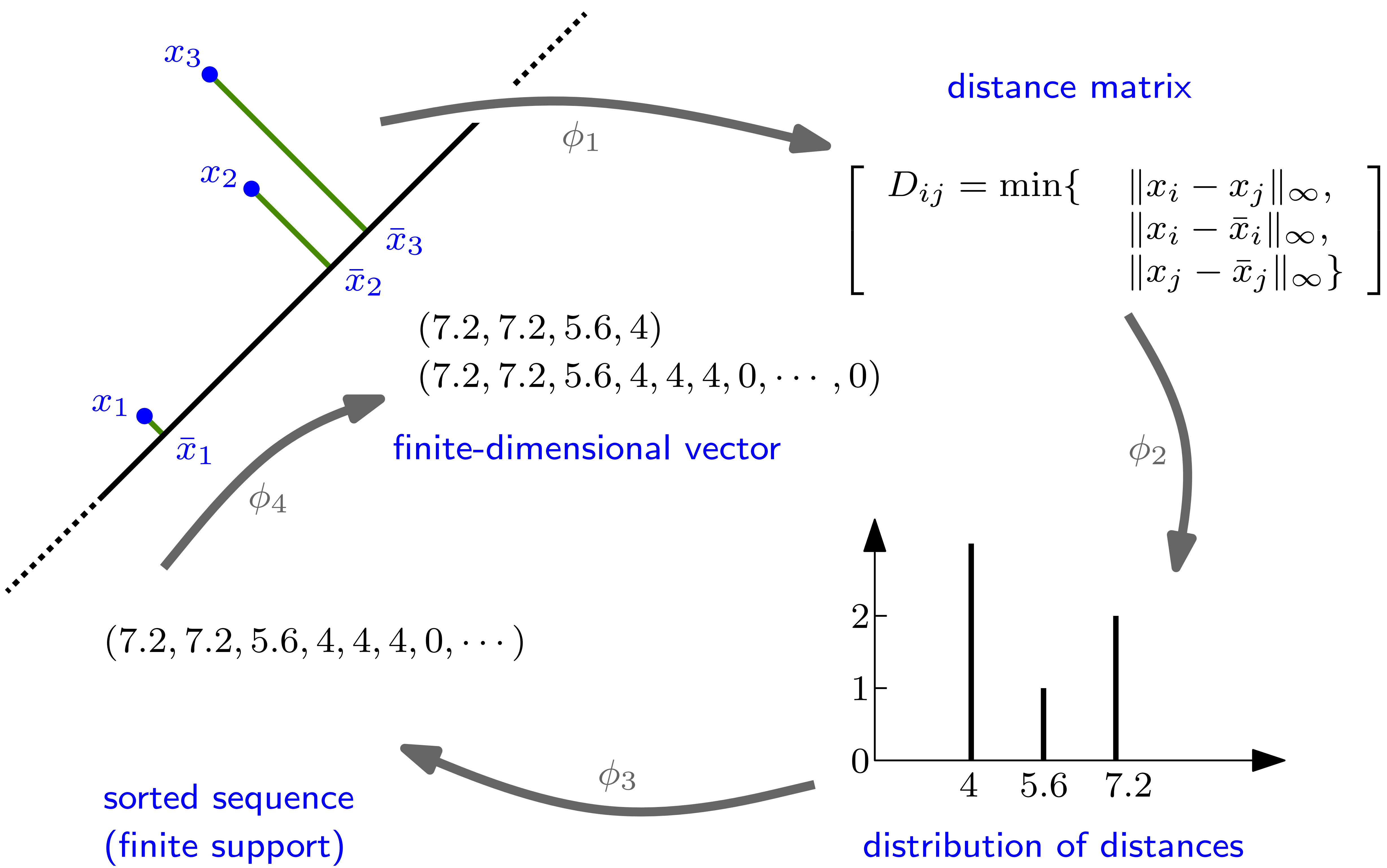
$$\mathrm{d}_\mathrm{B}^\infty = \inf_{m:X\leftrightarrow Y} \max\left\{ \sup_{p \text{ matched}} \|p - m(p)\|_\infty,\ \sup_{\text{p unmatched}} \|p - \bar{p}\|_\infty \right\}$$

Problem: generates instability in distance matrix ($\mathrm{d}_\mathrm{B}^\infty \ll W_\infty$)

Solution: change the metric

Adding the diagonal

$\Phi = \phi_4 \circ \phi_3 \circ \phi_2 \circ \phi_1$



Stability

Theorem: [Carrière et al. 2015] For any persistence diagrams X, Y, for any feature space dimension D, $\|\Phi(X) - \Phi(Y)\|_\infty \leq 2\mathrm{d}_{\mathrm{B}}^{\infty}(X, Y)$

Stability

Theorem: [Carrière et al. 2015] For any persistence diagrams X, Y, for any feature space dimension D, $\|\Phi(X) - \Phi(Y)\|_\infty \leq 2\mathrm{d}_\mathrm{B}^\infty(X, Y)$

$$\forall p \geq 1,\ \|\Phi(X) - \Phi(Y)\|_p \leq 2D^{1/p}\mathrm{d}_\mathrm{B}^\infty(X, Y)$$

Stability

Theorem: [Carrière et al. 2015] For any persistence diagrams X, Y, for any feature space dimension D, $\|\Phi(X) - \Phi(Y)\|_\infty \leq 2\mathrm{d}_\mathrm{B}^\infty(X, Y)$

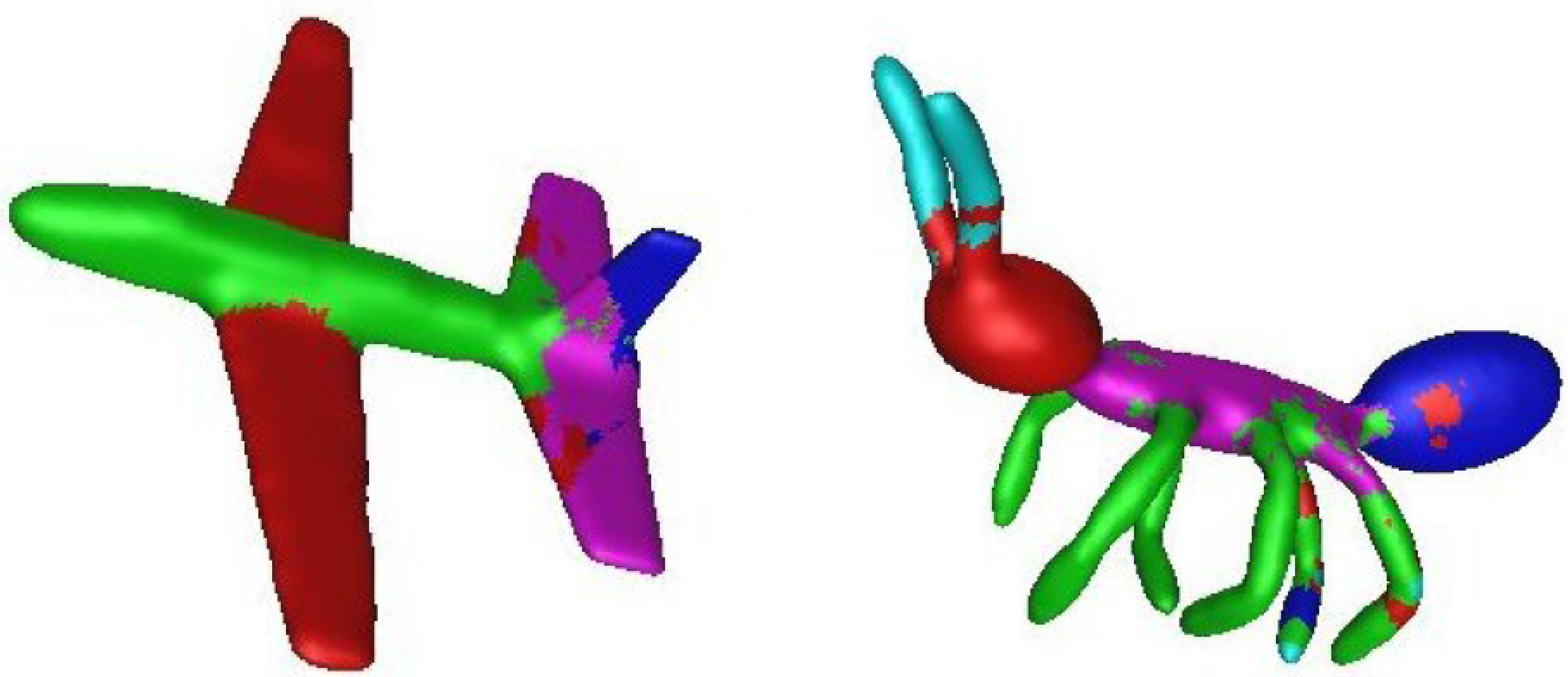
$$\forall p \geq 1,\ \|\Phi(X) - \Phi(Y)\|_p \leq 2D^{1/p}\mathrm{d}_\mathrm{B}^\infty(X, Y)$$

- case $p = \infty$ useful for retrieval and NN-classifiers (fast proximity queries)

- case $p = 2$ useful for linear / kernel-based classifiers (scalar product)

Application: supervised segmentation

Approach 1: use k-NN classifier in feature space $(\mathbb{R}^D, \ell^\infty)$

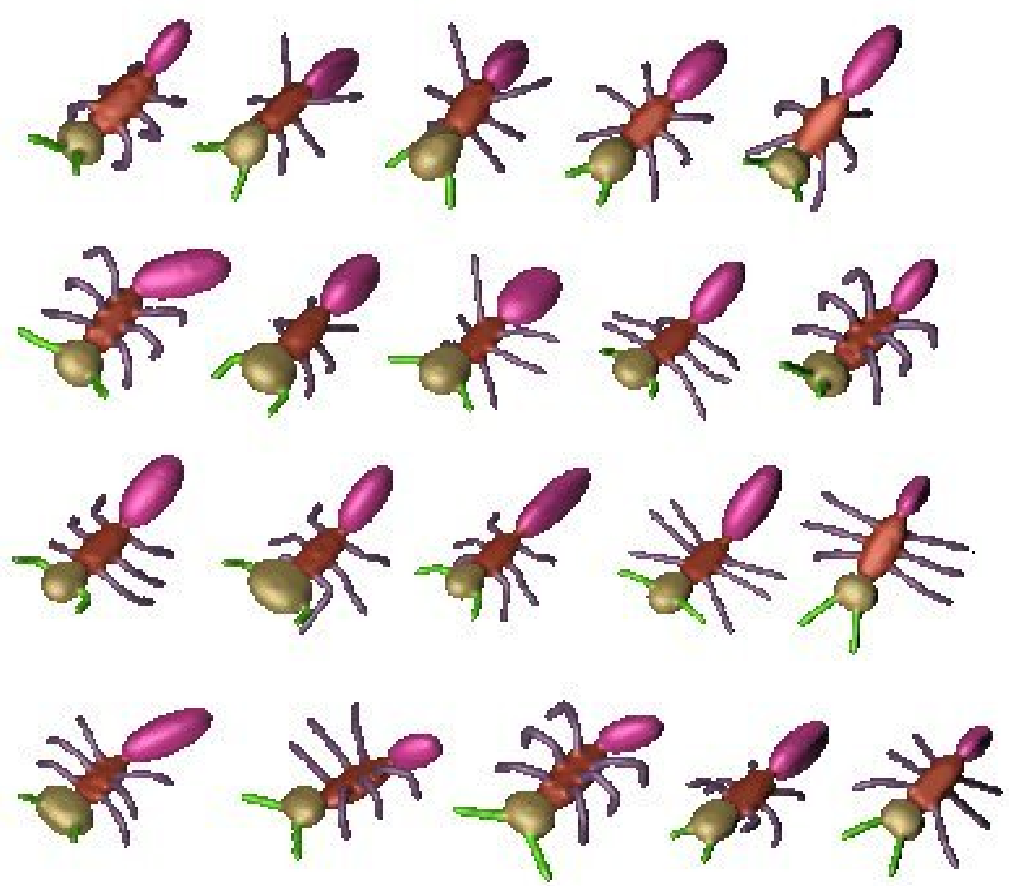


Application: supervised segmentation

Approach 1: use k-NN classifier in feature space $(\mathbb{R}^D, \ell^\infty)$

Approach 2: use linear classifier (SVM) in feature space $(\mathbb{R}^D, \ell^2)$

+ graph cut [Kalogerakis et al. 2010]



Application: supervised segmentation

Approach 1: use k-NN classifier in feature space $(\mathbb{R}^D, \ell^\infty)$

Approach 2: use linear classifier (SVM) in feature space $(\mathbb{R}^D, \ell^2)$

$+$ graph cut [Kalogerakis et al. 2010]

	SB5	SB5+PDs
Human	21.3	**11.3**
Cup	10.6	**10.1**
Glasses	21.8	**25.0**
Airplane	18.7	**9.3**
Ant	9.7	**1.5**
Chair	15.1	**7.3**
Octopus	5.5	**3.4**
Table	7.4	**2.5**
Teddy	6.0	**3.5**
Hand	21.1	**12.0**

	SB5	SB5+PDs
Plier	12.3	**9.2**
Fish	20.9	**7.7**
Bird	24.8	**13.5**
Armadillo	18.4	**8.3**
Bust	35.4	**22.0**
Mech	22.7	**17.0**
Bearing	25.0	**11.2**
Vase	26.4	**17.8**
FourLeg	25.6	**15.8**

percentage of mislabelling ($100-$rand index)

Application: non-rigid shape matching

Approach: use framework of *functional maps* [Ovsjanikov et al. 2012]

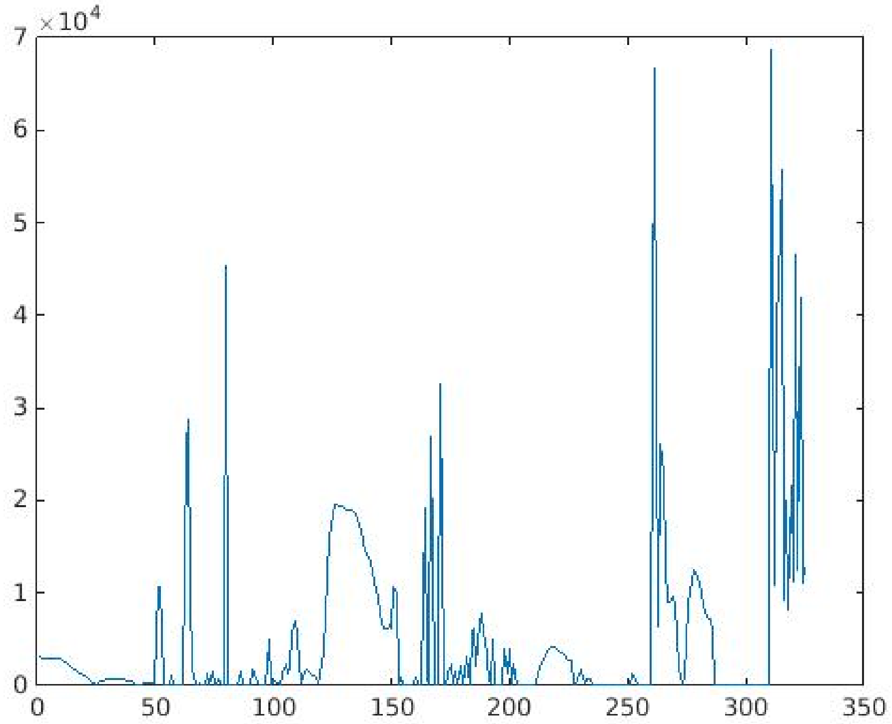
- compute an optimal linear map that best preserves a set of signatures (vectors)

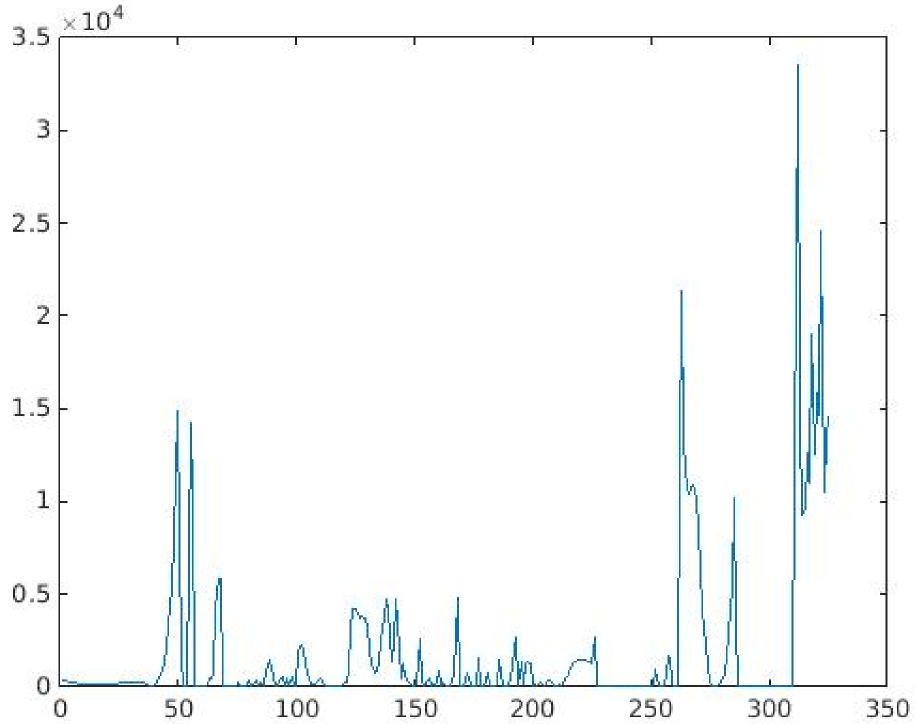
- derive a point-to-point correspondence from this map (via indicator functions)

- evaluate the quality of the correspondence

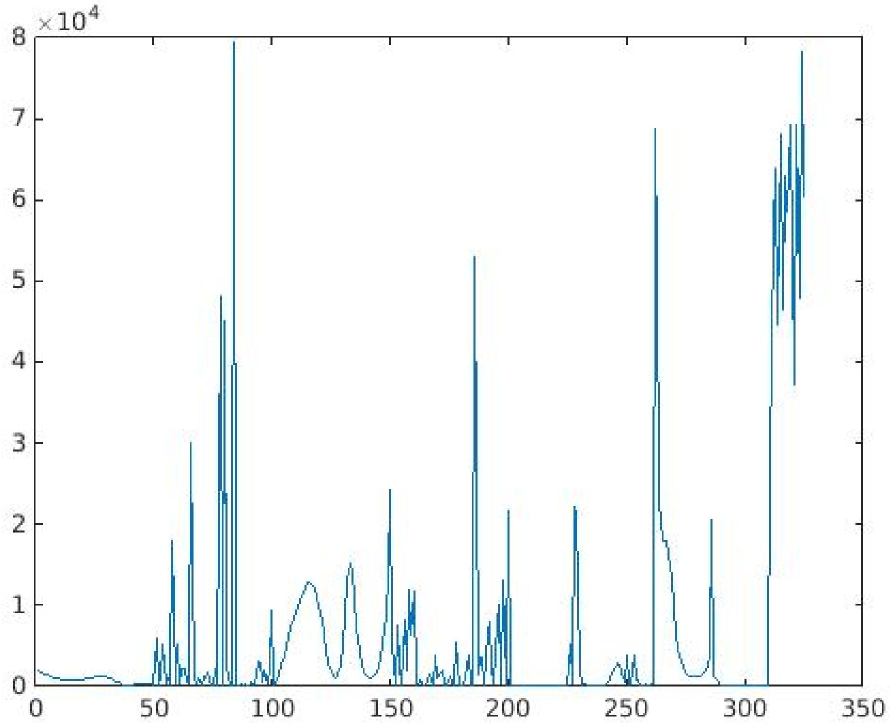
Application: non-rigid shape matching

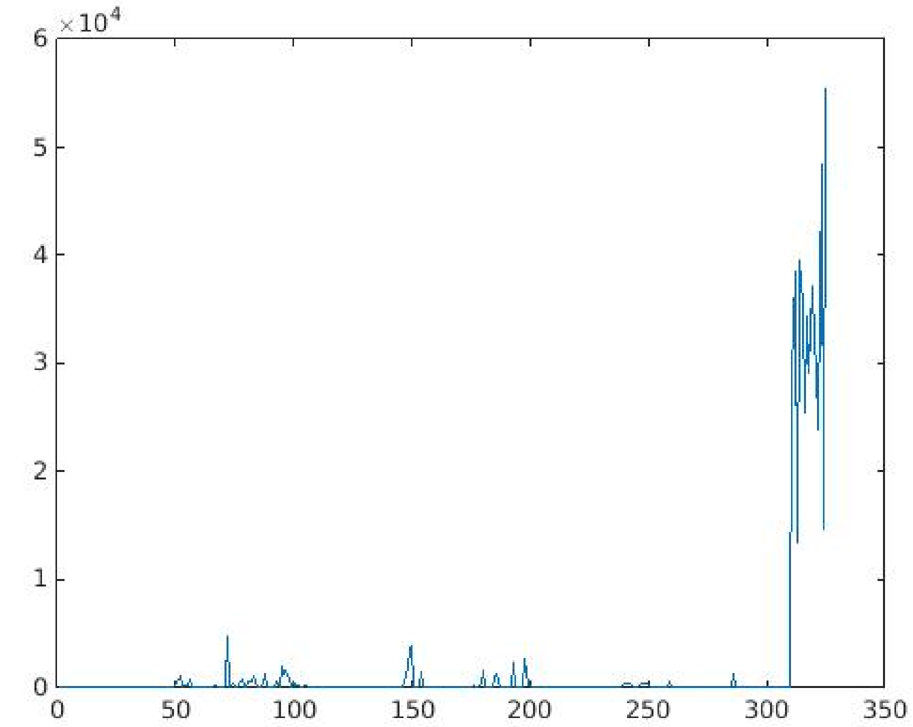
Approach: use framework of *functional maps* [Ovsjanikov et al. 2012]





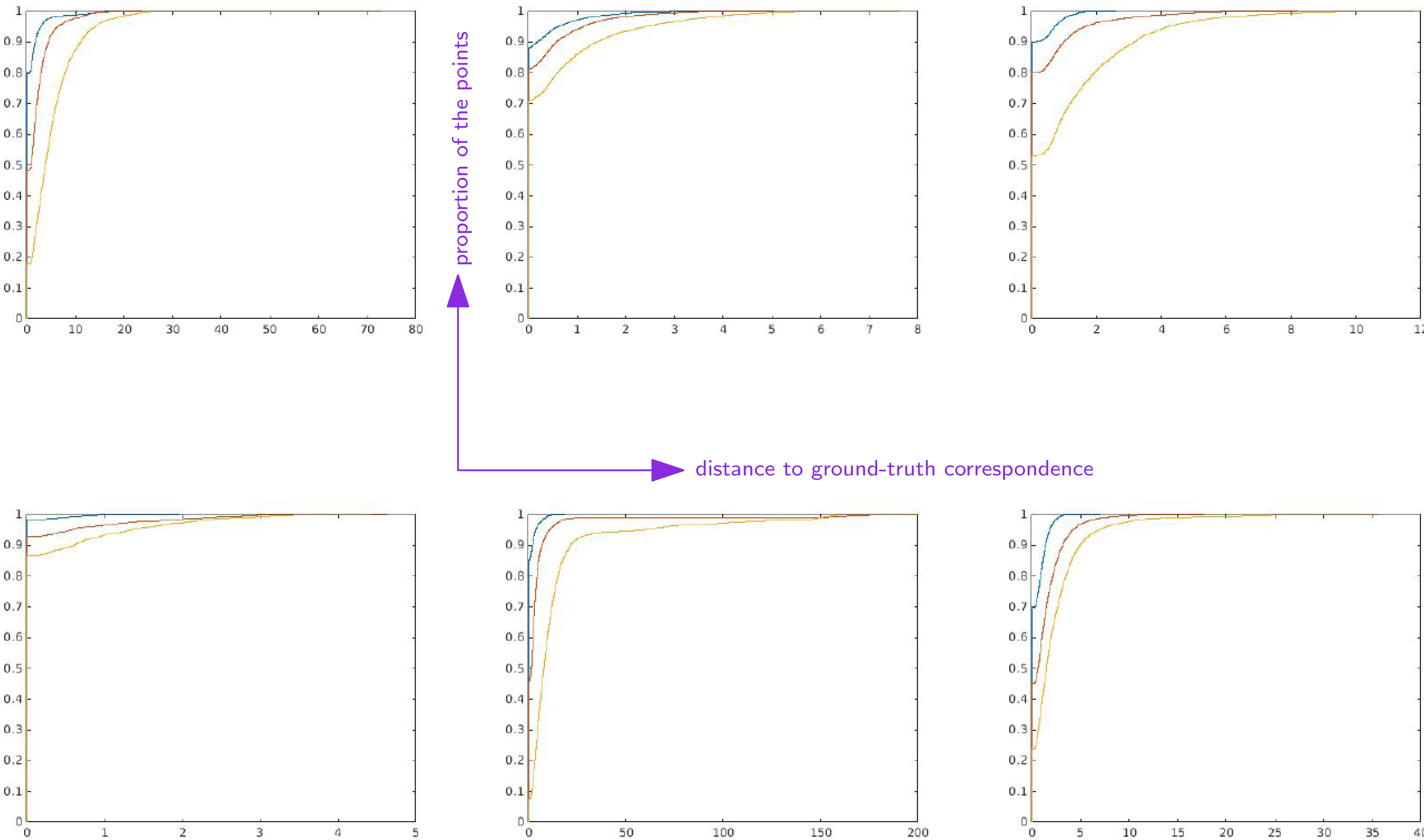
topological signatures (last 30 indices) have a high influence on the choice of optimal map





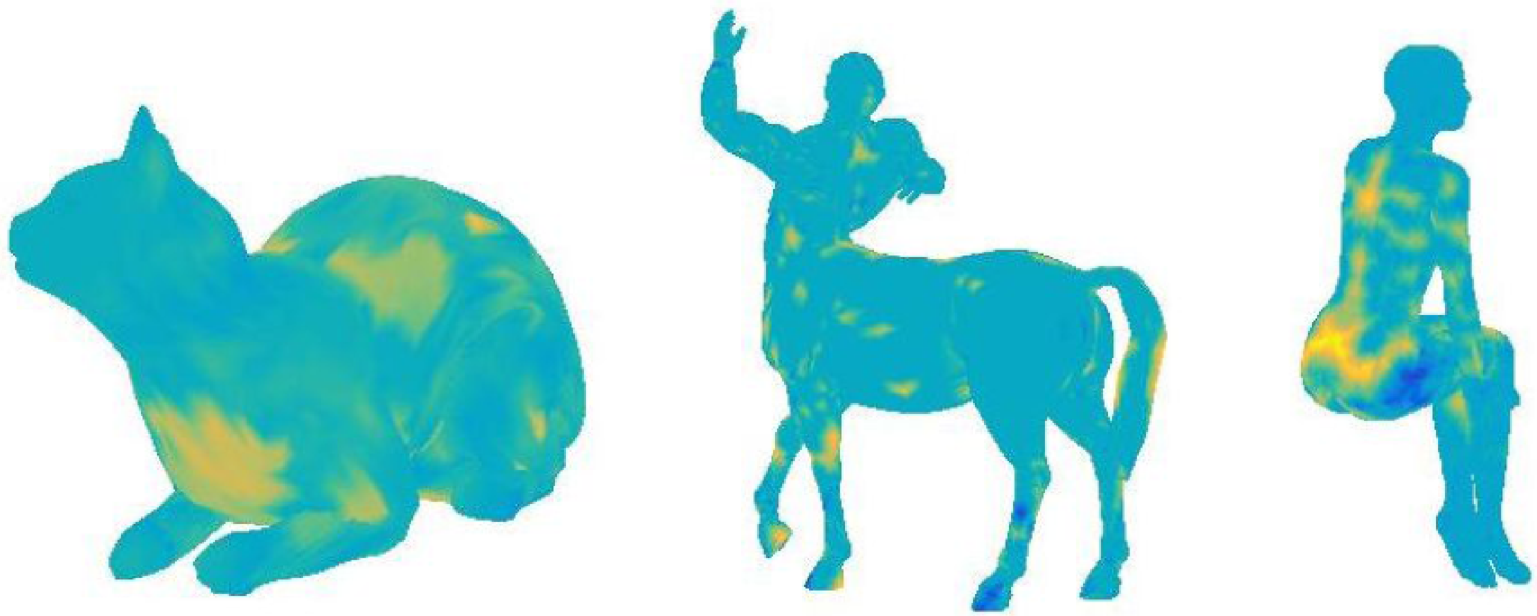
Application: non-rigid shape matching

Approach: use framework of *functional maps* [Ovsjanikov et al. 2012]

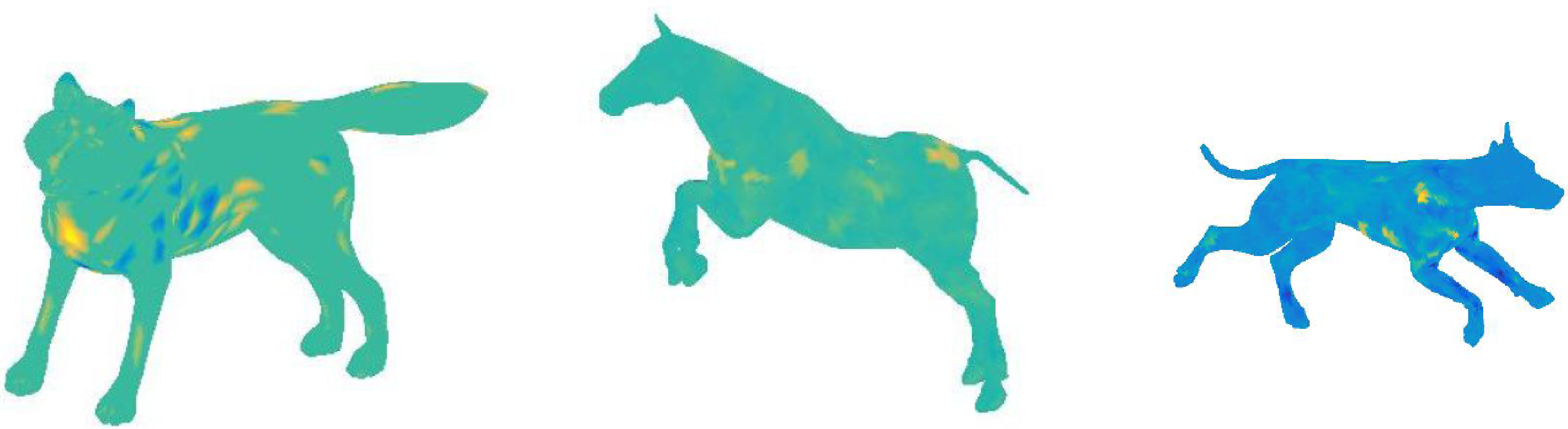


Application: non-rigid shape matching

Approach: use framework of *functional maps* [Ovsjanikov et al. 2012]

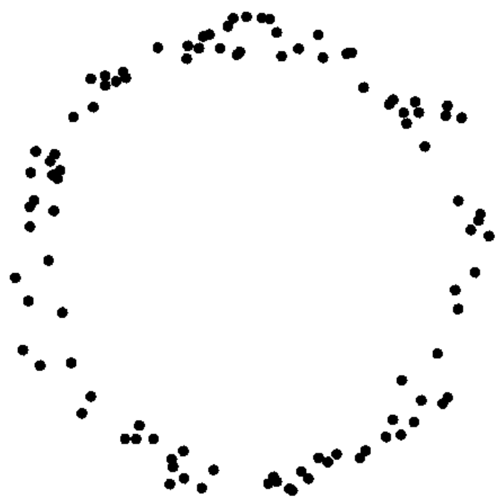


correspondences in flat regions are improved by topological signatures



Wrap-Up

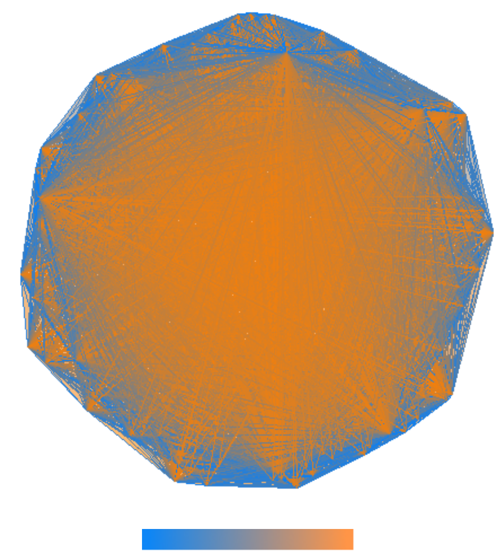
finite metric space / basepoint



$$\begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ & & \cdots & \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{bmatrix}$$

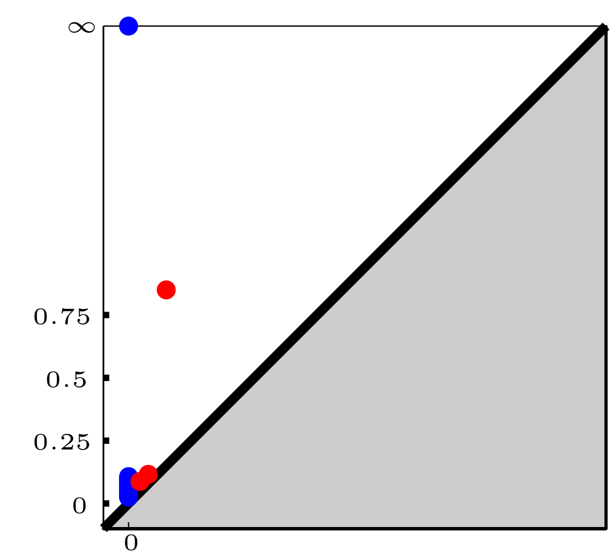
$\Longrightarrow$

Filtration



$\Longrightarrow$

persistence diagram



$\Downarrow$

$$(7.2, 5.3, 5.3, 3.4, 0, \cdots, 0)$$

feature vector

- topological descriptors are **provably stable**
- they provide **complementary information**
- they can be computed and mapped to feature spaces **efficiently**