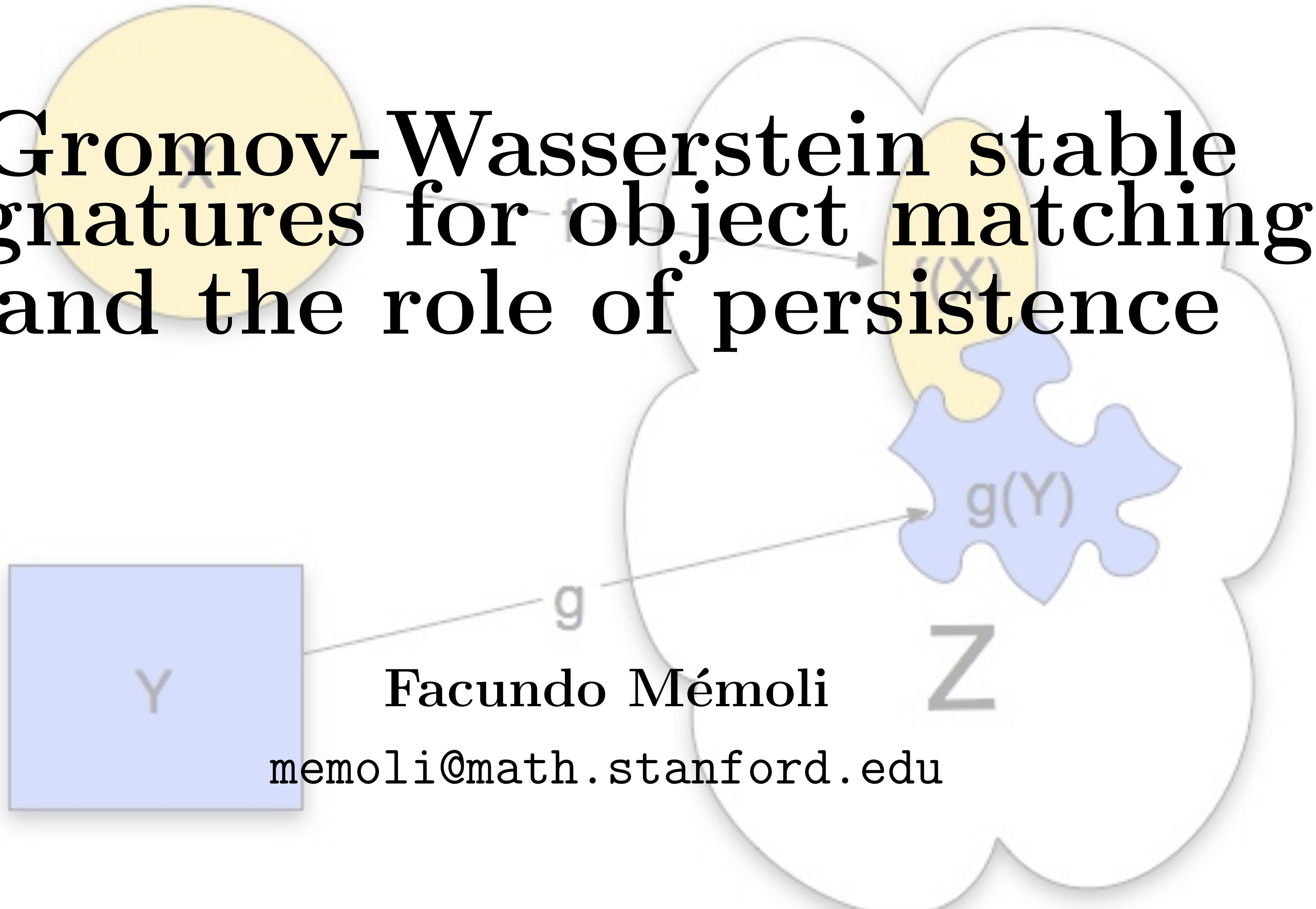
Gromov-Wasserstein stable signatures for object matching and the role of persistence



Facundo Mémoli

memoli@math.stanford.edu

Some definitions..

- Let (Z, d) be a compact metric space.
- For closed $A, B \subset Z$, we define the **Hausdorff distance** as:

$$d_{\mathcal{H}}^{Z}(A, B) = \max(\max_{b \in B} \min_{a \in A} d(a, b), \max_{a \in A} \min_{b \in B} d(a, b)).$$

- For **probability measures** μ_A and μ_B on Z and $p \geq 1$, we define the **Wasserstein distance** as:

$$d_{\mathcal{W},p}^{Z}(\mu_A, \mu_B) = \min_{\mu} \left(\iint_{Z \times Z} d^p(x, y)\, \mu(dx, dy) \right)^{1/p},$$

where $\mu \in \mathcal{M}(\mu_A, \mu_B)$, the collection of all **measure couplings** between μ_A and μ_B: probability measures on $Z \times Z$ with marginals μ_A and μ_B, respectively.

Comparison of objects

- Given a **compact** metric space (Z, d_Z), called the **ambient space**, one can define **objects** to be either
  - **compact subsets of** Z: $\mathcal{C}(Z)$, or
  - **probability measures on** Z: $\mathcal{C}^w(Z)$.

  I will be redundant and say that objects in $\mathcal{C}^w(Z)$ are pairs (A, μ_A) where A is the **support** of the probability measure μ_A.

- In each case, one can put a metric on objects and regard the collection of all objects as a metric space in itself.

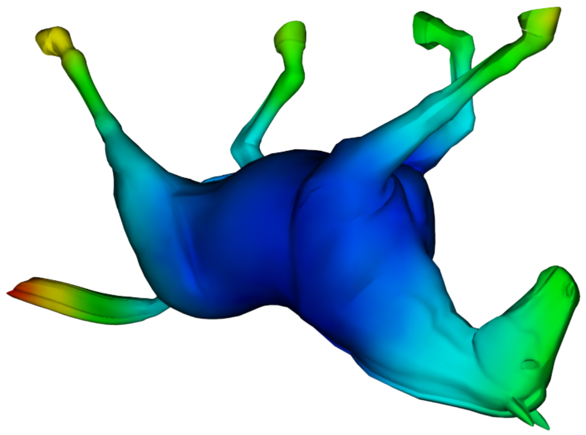
- In the case of $\mathcal{C}(Z)$, this metric is the **Hausdorff** metric $d^Z_{\mathcal{H}}$. One has:

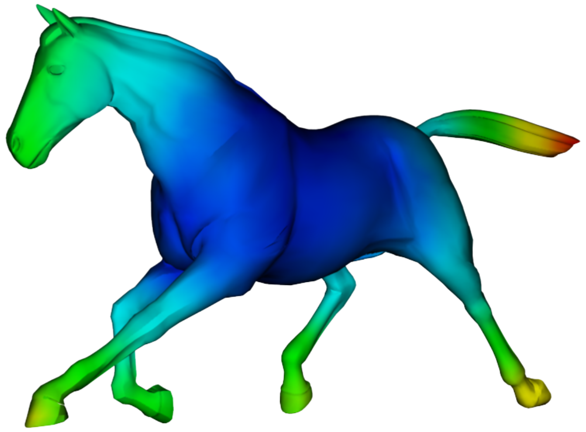
  Theorem (Blaschke). *For (Z, d_Z) compact, $(\mathcal{C}(Z), d^Z_{\mathcal{H}})$ is also a compact metric space.*

- In the case of $\mathcal{C}^w(Z)$, this metric is the **Wasserstein** metric $d^Z_{\mathcal{W},p}$. One has:

  Theorem (Prokhorov). *For (Z, d_Z) compact, $(\mathcal{C}^w(Z), d^Z_{\mathcal{W},p})$ is also a compact metric space.*

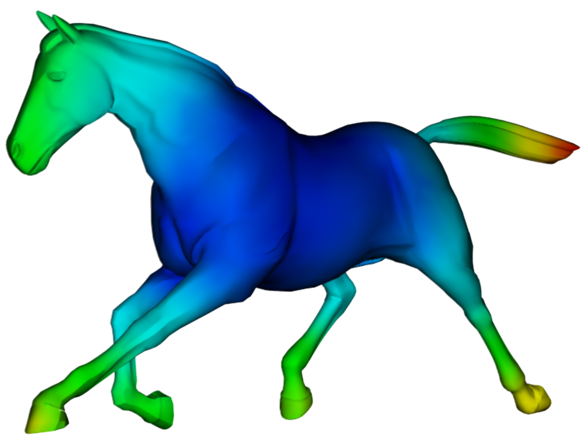
- What if one wants to consider "**invariances**"? consider for example objects in $\mathbb{R}^d$: you may want to factor out all rigid isometries.





T acts on sets in the usual way: $T(A) = \{T(a),\, a \in A\}$. On measures it acts by push-forward: If C is measurable, then $T(\mu)(C) = T_{\#}\mu(C) = \mu\big(T^{-1}(C)\big)$.

- What if one wants to consider "**invariances**"? consider for example objects in $\mathbb{R}^d$: you may want to factor out all rigid isometries.



T acts on sets in the usual way: $T(A) = \{T(a),\, a \in A\}$. On measures it acts by push-forward: If C is measurable, then $T(\mu)(C) = T_{\#}\mu(C) = \mu\big(T^{-1}(C)\big)$.

- This can be incorporated into our formulation: let $I(Z)$ be the *isometry group* on Z, and define

  – For $A, B \in \mathcal{C}(Z)$,

$$d_{\mathcal{H}}^{Z,\text{iso}}(A,B) := \inf_{T \in I(Z)} d_{\mathcal{H}}^{Z}(A, T(B)).$$

  – For $A, B \in \mathcal{C}^w(Z)$,

$$d_{\mathcal{W},p}^{Z,\text{iso}}(A,B) := \inf_{T \in I(Z)} d_{\mathcal{W},p}^{Z}(A, T(B)).$$

- These two constructions provide metrics on the (isometry classes of) objects in Z.

- This is what one could call the **extrinsic approach** to object matching: there is an **ambient space**.

The intrinsic approach, briefly.

- What if we regard objects as metric spaces? This may make sense since we are actually trying to get rid of **ambient space isometries**.

- For example, given $A \in \mathcal{C}(Z)$, upgrade this object to the metric space (A, d_A) where d_A is the **restriction** of d_Z to $A \times A$.

- Then, given two objects A, B, one could attempt to compute some notion of distance between the metric spaces:

$$d_{\mathcal{GH}}((A, d_A), (B, d_B)).$$

  Here, GH stands for **Gromov-Hausdorff**.

- Similarly, for $(A, \mu_A), (B, \mu_B) \in \mathcal{C}^w(Z)$ one constructs the **measure metric spaces** (mm-spaces: metric spaces enriched with a probability measure) (A, d_A, μ_A) and (B, d_B, μ_B). Then, one would compute some distance on mm-spaces:

$$d_{\mathcal{GW},p}(A, d_A, \mu_A)(B, d_B, \mu_B).$$

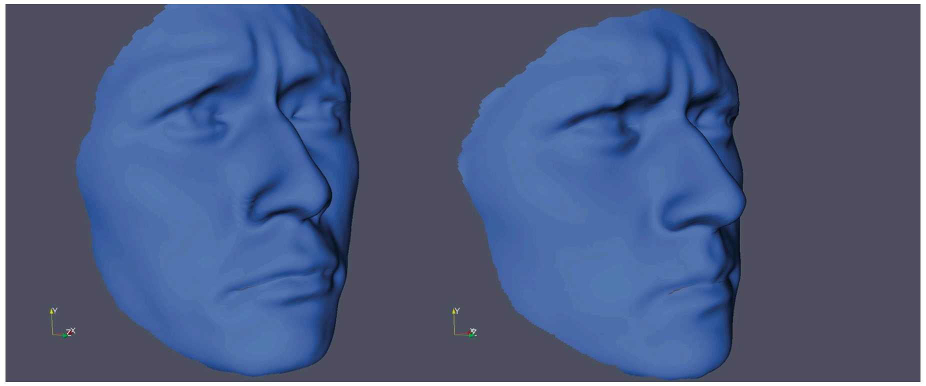
  GW stands for **Gromov-Wasserstein**.

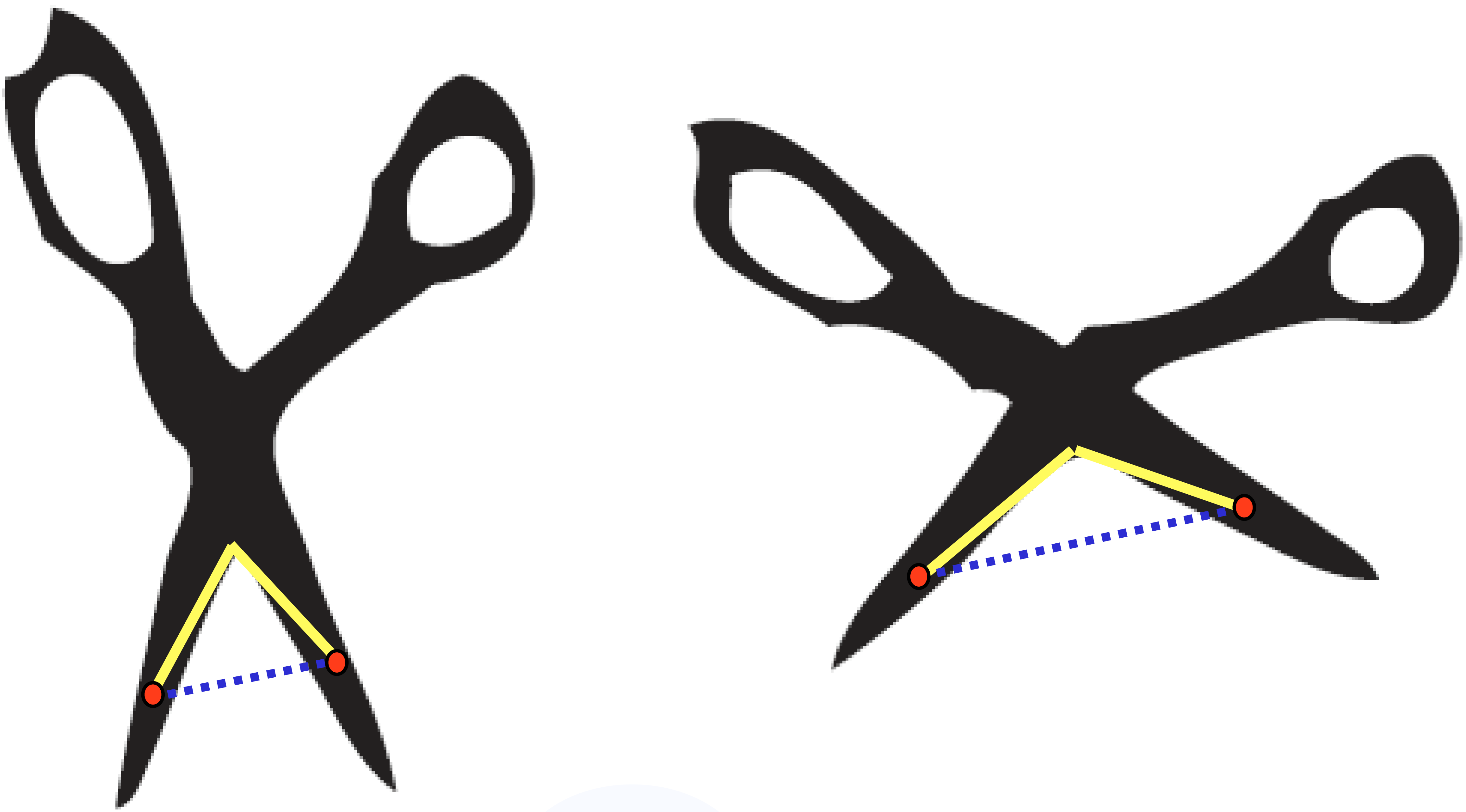
- There are practical examples that motivate pursuing the intrinsic approach.
- Consider for example invariance to **bends**, **articulations** or **poses**: the geodesic distance is (approximately) preserved– but there is no ambient space isometry that maps one shape to a vicinity of the other.

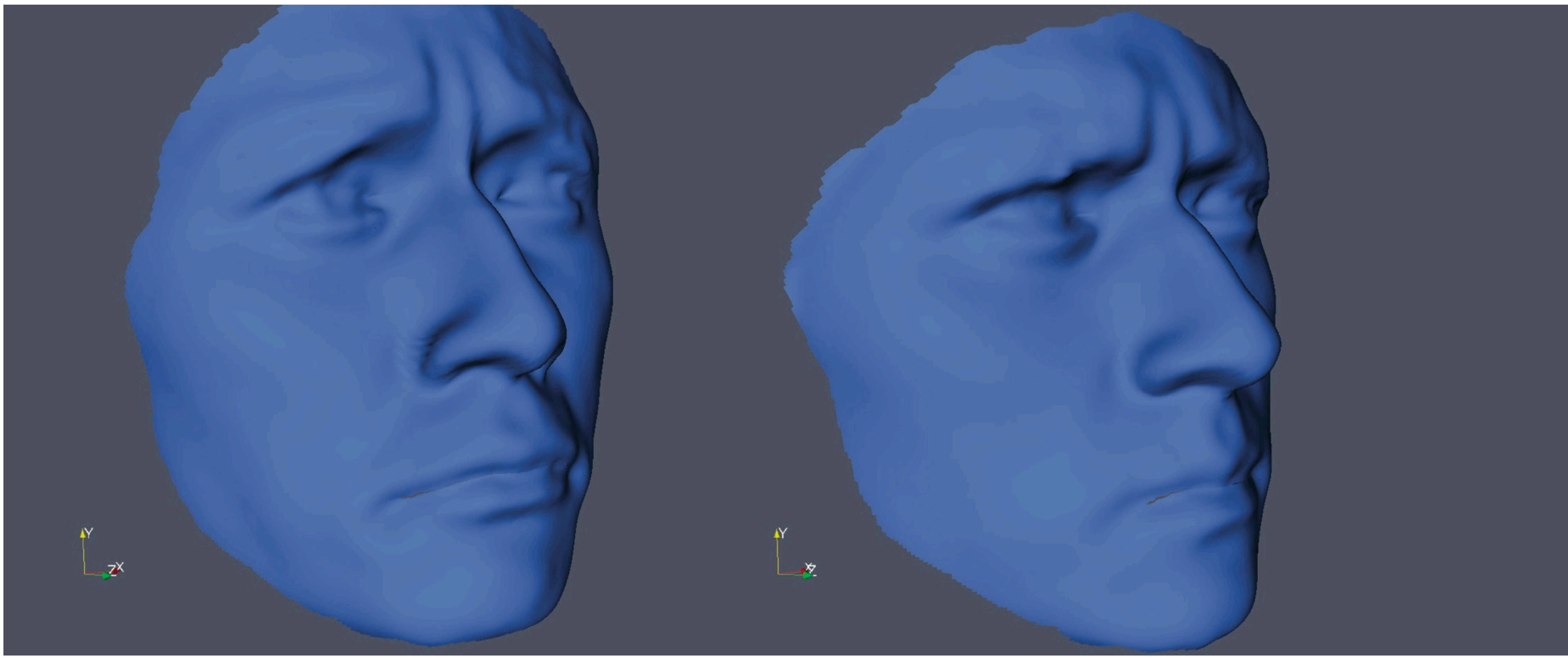
- There are practical examples that motivate pursuing the intrinsic approach.
- Consider for example invariance to **bends**, **articulations** or **poses**: the geodesic distance is (approximately) preserved– but there is no ambient space isometry that maps one shape to a vicinity of the other.

- There are practical examples that motivate pursuing the intrinsic approach.

- Consider for example invariance to **bends**, **articulations** or **poses**: the geodesic distance is (approximately) preserved– but there is no ambient space isometry that maps one shape to a vicinity of the other.

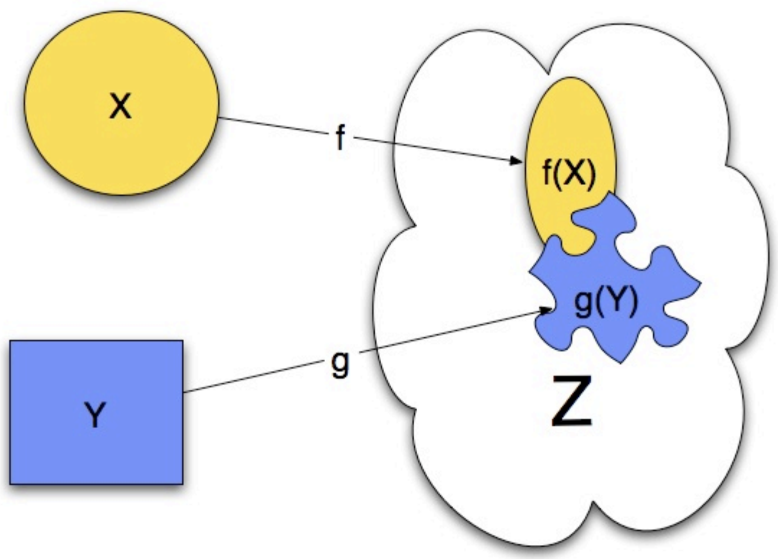






The Gromov construction: a distance between compact metric spaces.

- Let $\mathcal{X}$ denote the collection of all compact metric spaces.
- Let $(X, d_X), (Y, d_Y) \in \mathcal{X}$ and consider all metric spaces (Z, d) s.t. there exist maps $f : X \to Z$ and $g : Y \to Z$, **isometric embeddings** of X and Y into Z, respectively.
- Inside Z, one can compute the Hausdorff distance between the isometric copies $f(X)$ of X and $g(Y)$ of Y.
- Then, take infimum over all possible choices of Z, f and g. The result is known as the **Gromov-Hausdorff distance**.



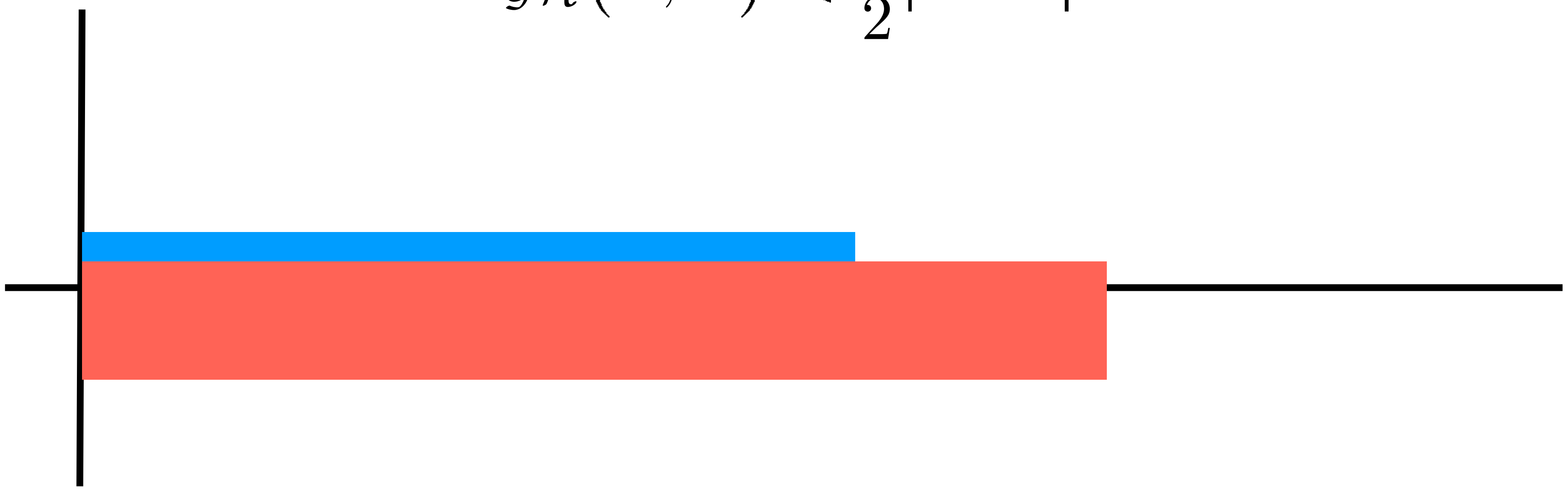
$$d_{\mathcal{GH}}((X, d_X), (Y, d_Y)) := \inf_{Z,f,g} d_{\mathcal{H}}^Z(f(X), g(Y)).$$

Example. *Let compact $A, B \subset \mathbb{R}$ be endowed with the Euclidean metric. Then,*

$$d_{\mathcal{GH}}\big(A, B\big) \leqslant \inf_{\gamma \in \mathbb{R}} d_{\mathcal{H}}^{\mathbb{R}}(A, B+\gamma).$$

If $A = [0, a]$ and $B = [0, b]$ for some $a, b \geqslant 0$, then

$$d_{\mathcal{GH}}\big(A, B\big) \leqslant \frac{1}{2}|a-b|$$

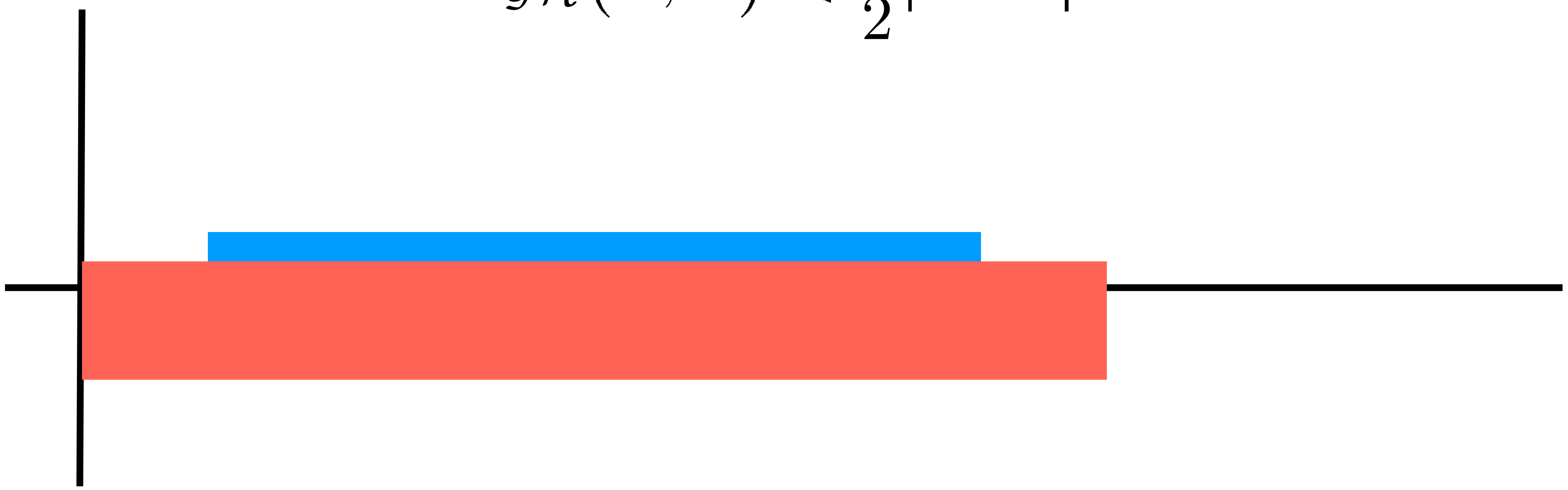


Example. *Let compact $A, B \subset \mathbb{R}$ be endowed with the Euclidean metric. Then,*

$$d_{\mathcal{GH}}\big(A, B\big) \leqslant \inf_{\gamma \in \mathbb{R}} d_{\mathcal{H}}^{\mathbb{R}}(A, B + \gamma).$$

If $A = [0, a]$ and $B = [0, b]$ for some $a, b \geqslant 0$, then

$$d_{\mathcal{GH}}\big(A, B\big) \leqslant \frac{1}{2}|a - b|$$



A bit more background: correspondences

Definition [Correspondences]

For sets A and B, a subset $C \subset A \times B$ is a *correspondence* (between A and B) if and and only if

- $\forall\ a \in A$, there exists $b \in B$ s.t. $(a, b) \in R$
- $\forall\ b \in B$, there exists $a \in A$ s.t. $(a, b) \in R$

Let $\mathcal{C}(A, B)$ denote the set of all possible correspondences between sets A and B.

Note that in the case $n_A = n_B$, correspondences are larger than bijections.

correspondences

Note that when A and B are finite, $C \in \mathcal{C}(A,B)$ can be represented by a matrix $((r_{a,b})) \in \{0,1\}^{n_A \times n_B}$ s.t.

$$\sum_{a \in A} r_{ab} \geqslant 1 \;\; \forall b \in B$$

$$\sum_{b \in B} r_{ab} \geqslant 1 \;\; \forall a \in A$$

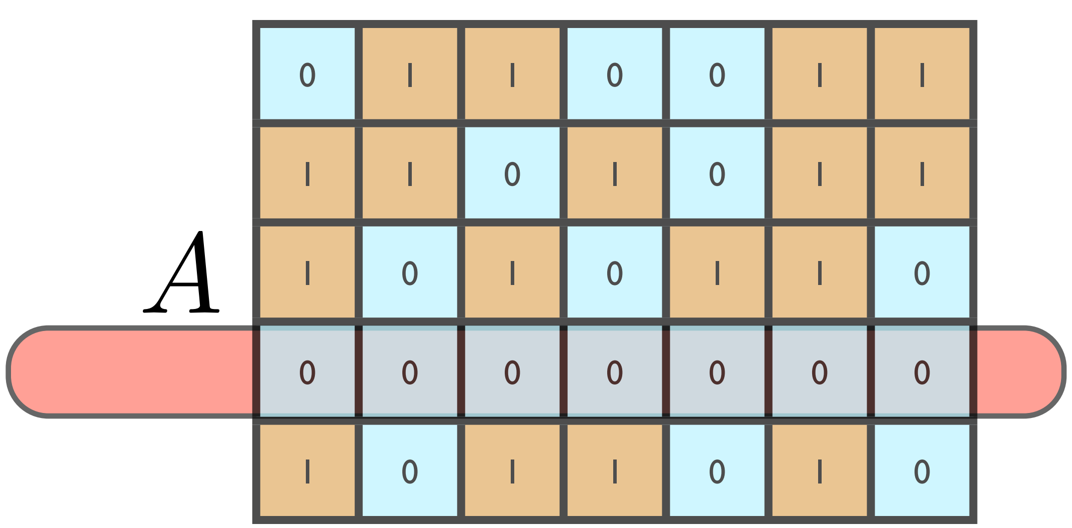
$$A \quad \begin{array}{|c|c|c|c|c|c|c|} \hline 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ \hline 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ \hline 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ \hline \end{array}$$

correspondences

Note that when A and B are finite, $C \in \mathcal{C}(A,B)$ can be represented by a matrix $((r_{a,b})) \in \{0,1\}^{n_A \times n_B}$ s.t.

$$\sum_{a \in A} r_{ab} \geqslant 1 \ \ \forall b \in B$$

$$\sum_{b \in B} r_{ab} \geqslant 1 \ \ \forall a \in A$$

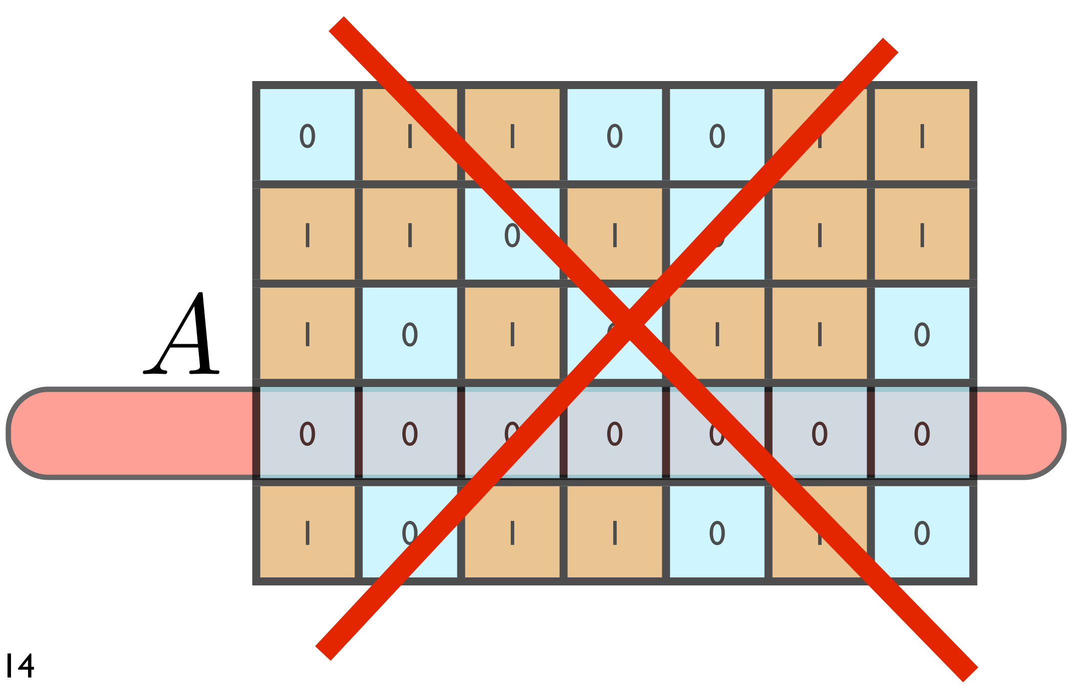


correspondences

Note that when A and B are finite, $C \in \mathcal{C}(A, B)$ can be represented by a matrix $((r_{a,b})) \in \{0,1\}^{n_A \times n_B}$ s.t.

$$\sum_{a \in A} r_{ab} \geqslant 1 \;\; \forall b \in B$$

$$\sum_{b \in B} r_{ab} \geqslant 1 \;\; \forall a \in A$$

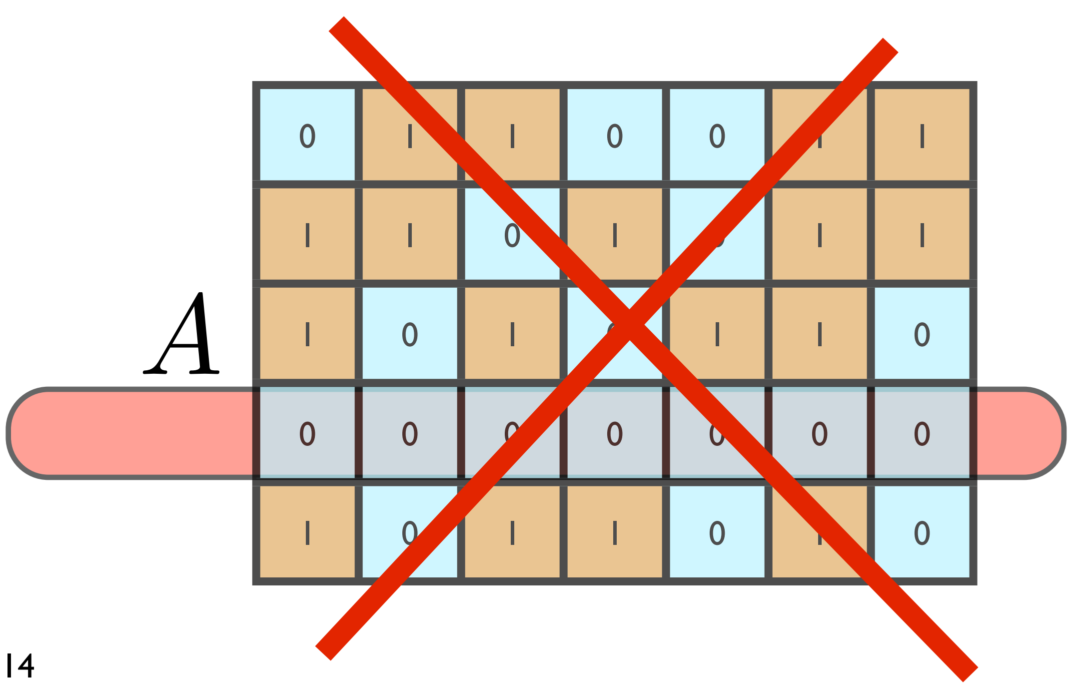


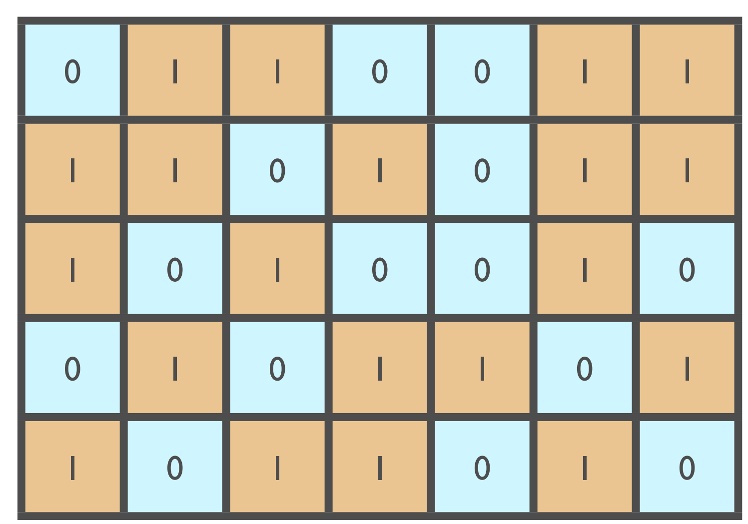
correspondences

Note that when A and B are finite, $C \in \mathcal{C}(A, B)$ can be represented by a matrix $(\!(r_{a,b})\!) \in \{0,1\}^{n_A \times n_B}$ s.t.

$$\sum_{a \in A} r_{ab} \geqslant 1 \;\; \forall b \in B$$

$$\sum_{b \in B} r_{ab} \geqslant 1 \;\; \forall a \in A$$



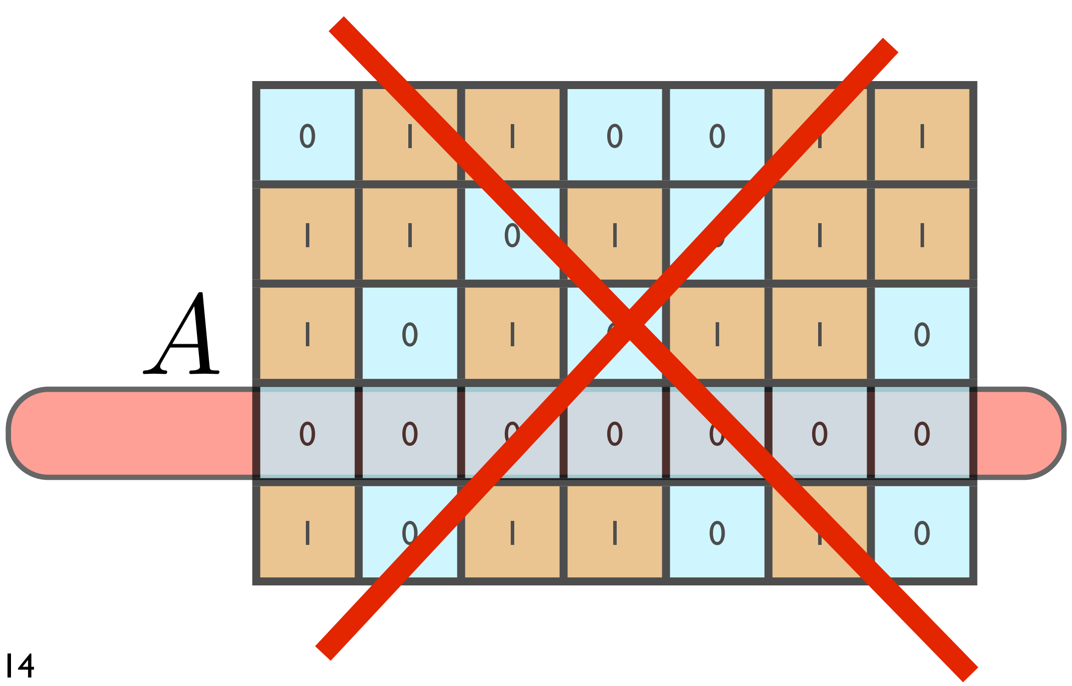


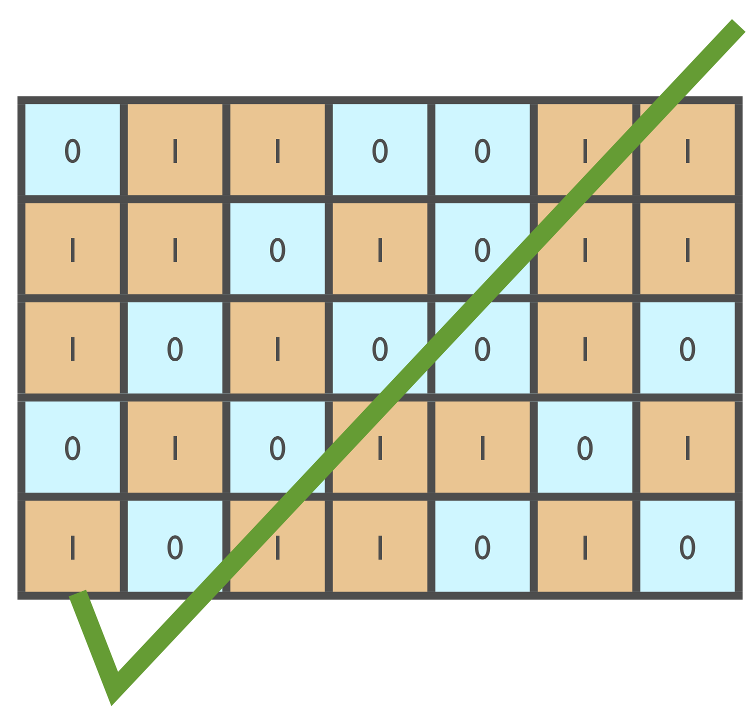
correspondences

Note that when A and B are finite, $C \in \mathcal{C}(A,B)$ can be represented by a matrix $(\!(r_{a,b})\!) \in \{0,1\}^{n_A \times n_B}$ s.t.

$$\sum_{a \in A} r_{ab} \geqslant 1 \;\; \forall b \in B$$

$$\sum_{b \in B} r_{ab} \geqslant 1 \;\; \forall a \in A$$



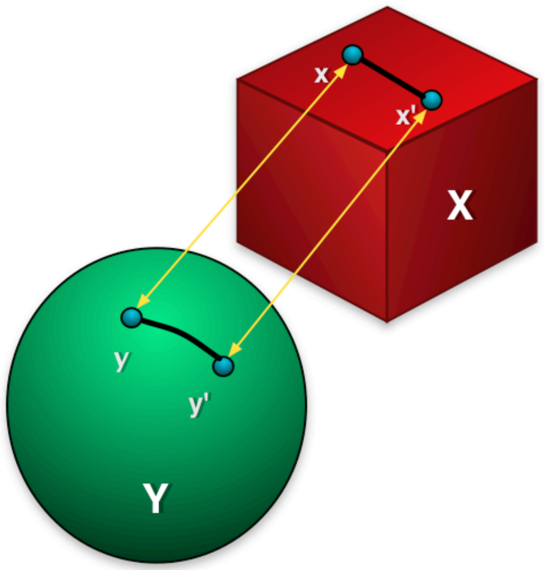


Another expression for the GH distance

Theorem. *[BBI] For compact metric spaces* (X, d_X) *and* (Y, d_Y)*,*

$$d_{\mathcal{GH}}(X,Y) = \frac{1}{2} \inf_C \max_{(x,y),(x',y')\in C} |d_X(x,x') - d_Y(y,y')|$$

Remark. *Let* $\Gamma_{X,Y}(x,y,x',y') = |d_X(x,x') - d_Y(y,y')|$*. We write, compactly,*



$$d_{\mathcal{GH}}(X,Y) = \frac{1}{2} \inf_C \|\Gamma_{X,Y}\|_{L^\infty(C\times C)}$$

Properties of the GH distance.

Theorem ([**BBI**]). *1. Let (X, d_X), (Y, d_Y) and (Z, d_Z) be metric spaces then*

$$d_{\mathcal{GH}}(X, Y) \leqslant d_{\mathcal{GH}}(X, Z) + d_{\mathcal{GH}}(Y, Z).$$

2. If $d_{\mathcal{GH}}(X, Y) = 0$ and (X, d_X), (Y, d_Y) are compact metric spaces, then (X, d_X) and (Y, d_Y) are isometric.

3. Let $\mathbb{X} \subset X$ be a closed subset of the compact metric space (X, d_X). Then,

$$d_{\mathcal{GH}}(X, \mathbb{X}) \leqslant d_{\mathcal{H}}^{X}(X, \mathbb{X}).$$

4. For compact metric spaces (X, d_X) and (Y, d_Y):

$$\begin{aligned} \frac{1}{2}\left|\mathbf{diam}\,(X) - \mathbf{diam}\,(Y)\right| &\leqslant d_{\mathcal{GH}}(X, Y) \\ &\leqslant \frac{1}{2}\max\left(\mathbf{diam}\,(X), \mathbf{diam}\,(Y)\right) \end{aligned}$$

*5. **Compact families**: Let $L > 0$ and $N : \mathbb{R}^{+} \to \mathbb{N}$. Define $\mathcal{F}(L, N) \subset \mathcal{X}$ to be s.t. any $X \in \mathcal{F}$ has $\mathbf{diam}\,(X) \leqslant L$ and for all $\varepsilon > 0$, X admits an ε-net with at most $N(\varepsilon)$ points. Then $(\mathcal{F}(L, N), d_{\mathcal{GH}})$ is pre-compact.*

Properties of the GH distance.

Theorem (**[BBI]**). *1. Let (X, d_X), (Y, d_Y) and (Z, d_Z) be metric spaces then*

$$d_{\mathcal{GH}}(X,Y) \leqslant d_{\mathcal{GH}}(X,Z) + d_{\mathcal{GH}}(Y,Z).$$

2. *If $d_{\mathcal{GH}}(X,Y) = 0$ and (X, d_X), (Y, d_Y) are compact metric spaces, then (X, d_X) and (Y, d_Y) are isometric.*

3. *Let $\mathbb{X} \subset X$ be a closed subset of the compact metric space (X, d_X). Then,*

$$d_{\mathcal{GH}}(X, \mathbb{X}) \leqslant d_{\mathcal{H}}^{X}(X, \mathbb{X}).$$

4. *For compact metric spaces (X, d_X) and (Y, d_Y):*

$$\begin{aligned} \frac{1}{2} \left|\mathbf{diam}\,(X) - \mathbf{diam}\,(Y)\right| &\leqslant d_{\mathcal{GH}}(X,Y) \\ &\leqslant \frac{1}{2} \max\left(\mathbf{diam}\,(X), \mathbf{diam}\,(Y)\right) \end{aligned}$$

5. ***Compact families**: Let $L > 0$ and $N : \mathbb{R}^+ \to \mathbb{N}$. Define $\mathcal{F}(L,N) \subset \mathcal{X}$ to be s.t. any $X \in \mathcal{F}$ has $\mathbf{diam}\,(X) \leqslant L$ and for all $\varepsilon > 0$, X admits an ε-net with at most $N(\varepsilon)$ points. Then $(\mathcal{F}(L,N), d_{\mathcal{GH}})$ is pre-compact.*

Example. *From item 4, since* $\mathbf{diam}\,(A) = a$ *and* $\mathbf{diam}\,(B) = b$*, then* $d_{\mathcal{GH}}\big(A, B\big) \geqslant \frac{1}{2}|a - b|$*. Then by previous computation,*

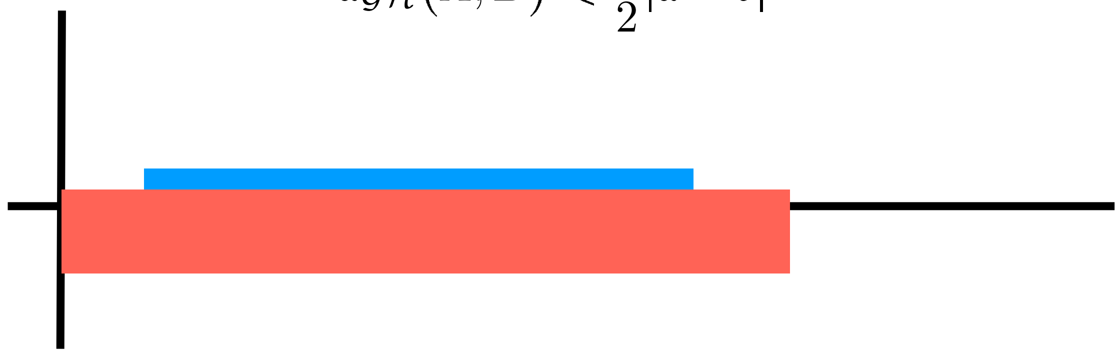
$$d_{\mathcal{GH}}\big([0, a], [0, b]\big) = \frac{1}{2}|a - b|.$$

Example. *Let compact* $A, B \subset \mathbb{R}$ *be endowed with the Euclidean metric. Then,*

$$d_{\mathcal{GH}}\big(A, B\big) \leqslant \inf_{\gamma \in \mathbb{R}} d_{\mathcal{H}}^{\mathbb{R}}(A, B + \gamma).$$

If $A = [0, a]$ *and* $B = [0, b]$ *for some* $a, b \geqslant 0$*, then*

$$d_{\mathcal{GH}}\big(A, B\big) \leqslant \frac{1}{2}|a - b|$$



Comments...

- The GH distance has been used in the applied object matching literature for a few years now [**MS04,MS05,BBK06,M07,M08,..**].

- It provides a useful set of ideas for reasoning about desirable properties of matching algorithms.

- Without further assumptions on the underlying metric spaces, it leads **combinatorial optimization problems**, more precisely, **Bottleneck Quadratic Assignment problems**, which are NP hard.

- Haven't been able to explain or relate to too many pre-existing practical approaches to object matching. There's a plethora of methods: it would be nice to understand inter-relation between them.

- Furthermore, the GH distance is a "pessimistic" measure of similarity: it is based on L^∞ dissimilarities: sensitivity to errors.

- From now on, we'll talk about the **Gromov-Wasserstein** distance, which yields **continuous optimization problems** directly and admits lower bounds based on easily computable and previously reported metric invariants.

Construction of the Gromov-Wasserstein distance(s)

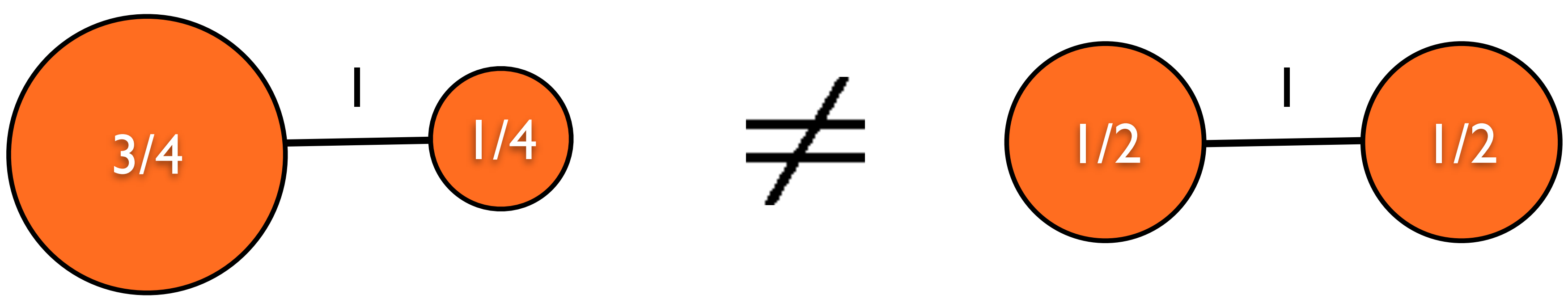
mm-spaces and their invariants

Definition.

- *The support* $supp[\mu]$ *of a probability measure* μ *on a compact metric space* (X, d_X) *is the minimal closed set outside of which there is zero mass.*

- *An* ***mm-space*** *is a triple* (X, d_X, μ_X) *where* (X, d_X) *is a compact metric space and* μ_X *is probability measure on* X *with* ***full support****:* $supp[\mu_X] = X$*. Let* $\mathcal{X}^w$ *denote the collection of all mm-spaces.*

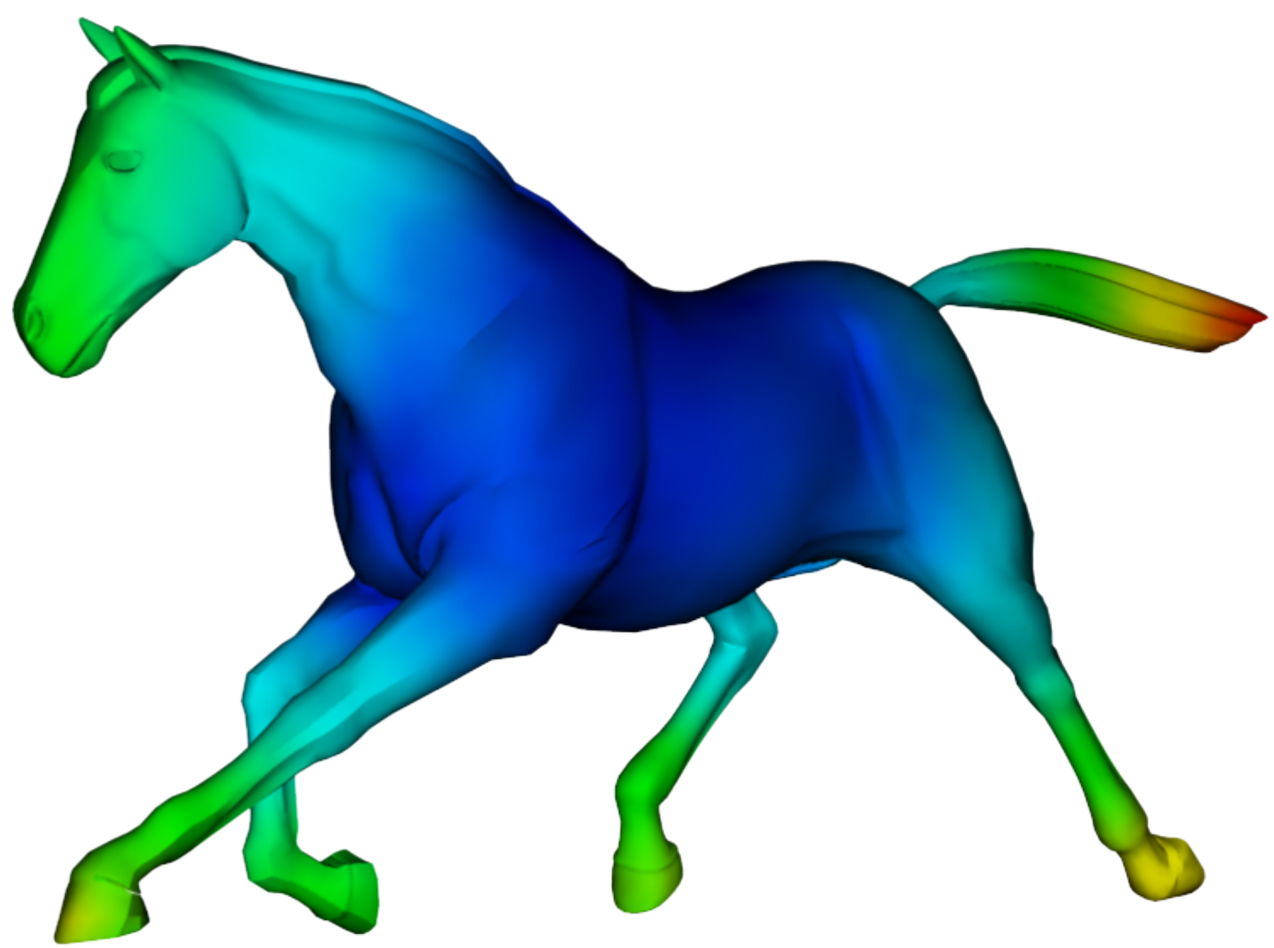
- *An* ***isomorphism*** *of mm-spaces is an isometry* $\Phi : X \to Y$ *s.t.* $\Phi_{\#}\mu_X = \mu_Y$*.*



Some invariants of mm-spaces.

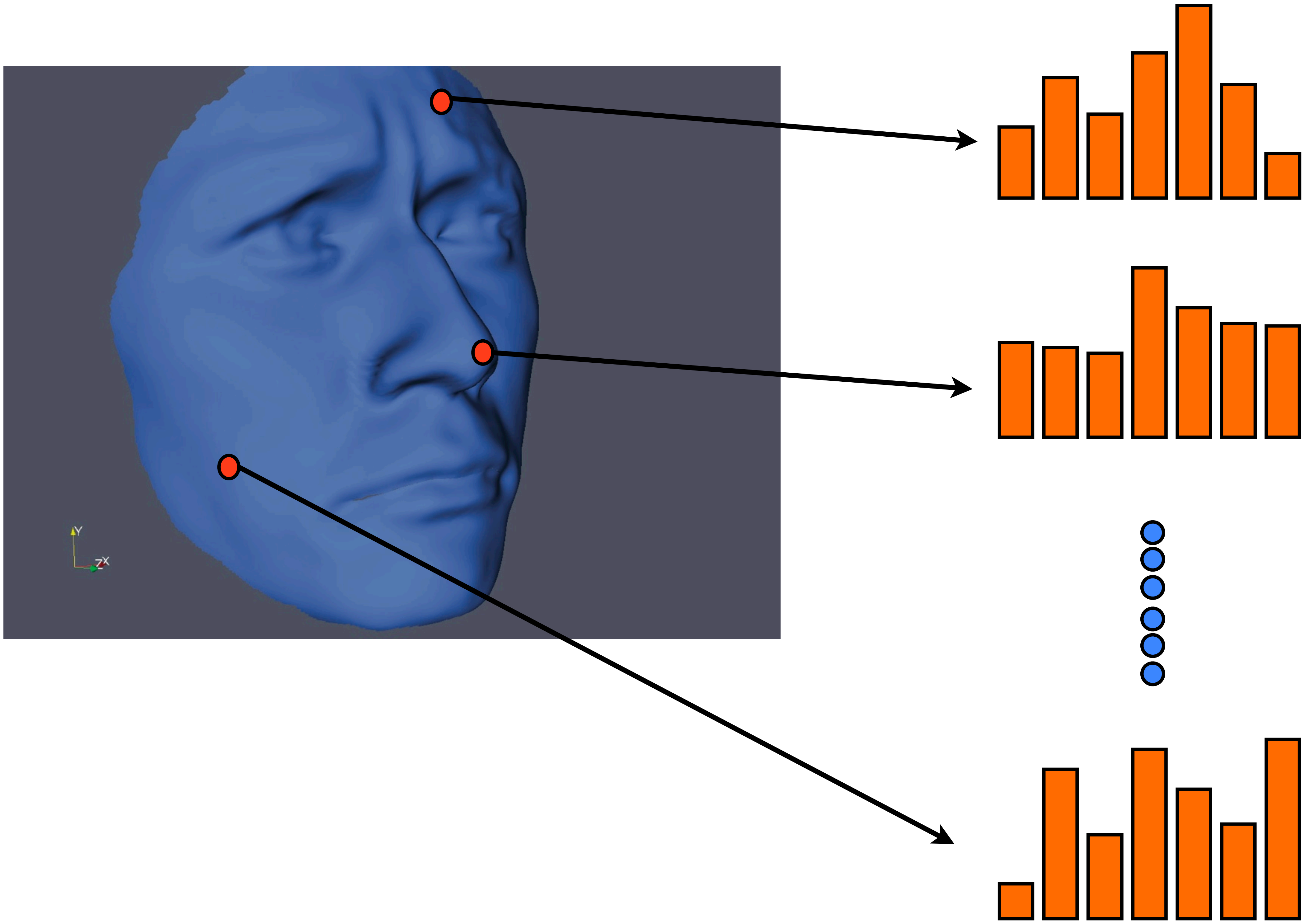
Example (Eccentricity on real shapes.). *This is three dimensional model of a horse. The metric is estimated from the mesh using Dijkstra, the measure is the uniform one. Red means high, blue means low. Notice how extremities get high values of the eccentricity ($p = 1$).*

$$s_{X,p}(x) = \|d_X(x, \cdot)\|_{L^p(\mu_X)} = \left(\sum_{x' \in X} \big(d_X(x, x')\big)^p \mu_X(x') \right)^{1/p}.$$



Local distributions of distances

$$h_X(x,t) = \mu_X\left(\overline{B(x,t)}\right)$$



Some invariants of mm-spaces.

- Given a mm-space (X, d_X, μ_X) define
  - p-**eccentricity**: $s_{X,p} : X \to \mathbb{R}^+$, $x \mapsto \|d_X(x, \cdot)\|_{L^p(\mu_X)}$.
  - **Local distribution of distances:**

$$h_X : X \times \mathbb{R}^+ \to [0,1], \quad (x,t) \mapsto \mu_X\left(\overline{B(x,t)}\right).$$

- Invariants similar to these have been used in the CS/EE literature. In particular, the eccentricity was explored by Hamza and Krim in 2002. The distribution of distances underlies a very famous work by the Princeton shape retrieval group. The Local shape distributions is similar to the integral invariants used by Manay-Soatto et al and the *Shape Contexts* of Bengio and Malik.

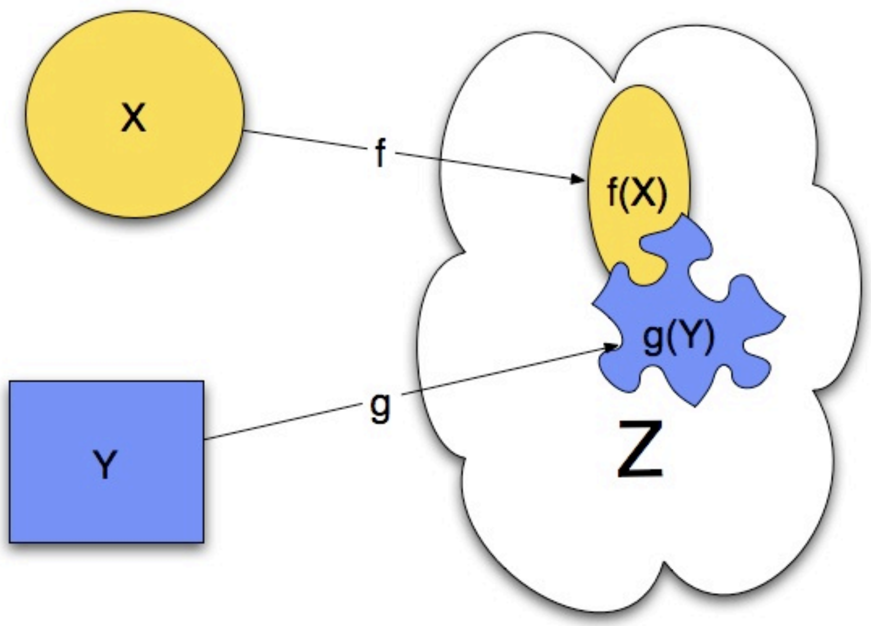
Construction of the GW distance

The Gromov construction: same thing for mm-spaces!

Fix $p \geqslant 1$. We may now define the **Gromov-Wasserstein** distance between $X, Y \in \mathcal{X}^w$ as

$$d_{\mathcal{GW},p}((X, d_X, \mu_X), (Y, d_Y, \mu_Y)) = \inf_{Z,f,g} d^Z_{\mathcal{W},p}(f_\# \mu_X, g_\# \mu_Y),$$

where f, g are isometric embeddings into Z.



Remark. *This definition is due to K.T. Sturm **[Sturm06]**.*

*This metric **does not** seem computationally appealing. In **[M07]** we constructed a closely related distance that is more suitable for practical computations.*

- Have lower bounds for $d_{\mathcal{GW},p}(X,Y)$ involving the invariants I described, **[M07,M08]**.

- These invariants have been reported in the literature and have been shown to provide good discrimination over databases of objects. Therefore the interest in inter-relating them and in finding these lbs.

- These invariants **cannot** be controlled by the GH distance alone: a notion of weight of a point is involved and therefore GW distances are natural here.

- Computation of these lower bounds leads to simpler problems than solving the GH or GW distaces.

- The question arises as to whether one could obtain lower bounds for the GH or GW distances of a completely different nature: how about **persistent topology** type of invariants?

- Have lower bounds for $d_{\mathcal{GW},p}(X,Y)$ involving the invariants I described, [**M07,M08**].

- These invariants have been reported in the literature and have been shown to provide good discrimination over databases of objects. Therefore the interest in inter-relating them and in finding these lbs.

- These invariants **cannot** be controlled by the GH distance alone: a notion of weight of a point is involved and therefore GW distances are natural here.

- Computation of these lower bounds leads to simpler problems than solving the GH or GW distaces.

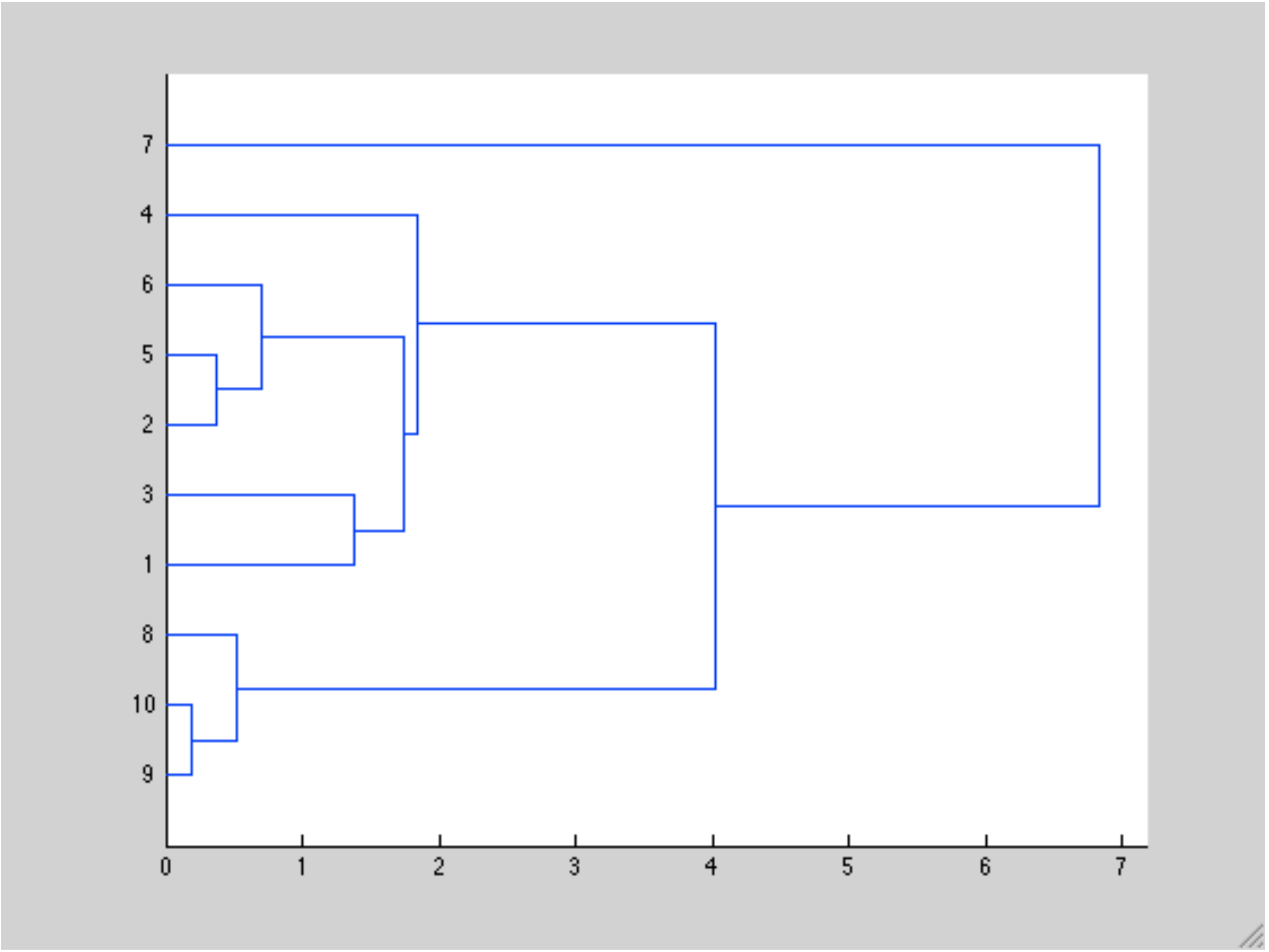
- The question arises as to whether one could obtain lower bounds for the GH or GW distances of a completely different nature: how about **persistent topology** type of invariants?

Lower bounds using persistence

Joint work with F. Chazal, D. Cohen-Steiner, L. Guibas and S. Oudot, [**CCGMO09**].

Motivation: Clustering

- Imagine you have underlying metric space (X, d_X) from which you can take only finitely many samples.
- Let $(\mathbb{X}, d_{\mathbb{X}})$ be a finite sampling from X where we assume $d_{\mathbb{X}}$ is the restriction metric.
- Apply a **hierarchical clustering method** $\mathfrak{C}$ to $(\mathbb{X}, d_{\mathbb{X}})$ and obtain a **dendrogram**:



- Can ask the question: how sensitive is $\mathfrak{S}\big(\mathbb{X}, d_{\mathbb{X}}\big)$ to $(\mathbb{X}, d_{\mathbb{X}})$?

- Can I guarantee that the answers I get from two different $\mathbb{X}_1$ and $\mathbb{X}_2$ are **similar** in some way when samplings become **denser** and denser in X?

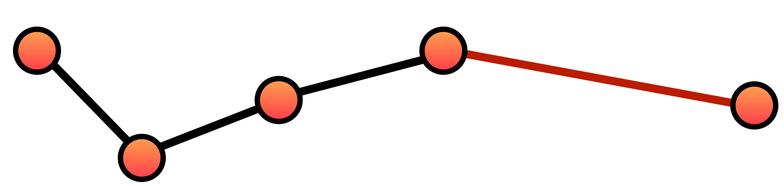
- Dendrograms are **rooted trees** and therefore equivalent to **ultrametrics**. Then can regard $\mathfrak{S}$ as a map from $\mathcal{M}$ to $\mathcal{U}$, where $\mathcal{M}$ (resp. $\mathcal{U}$) is collection of all finite metric (resp. ultrametric) spaces s.t. $\mathfrak{S} : \mathcal{M}_n \to \mathcal{U}_n$ for $n \in \mathbb{N}$.

- Let's assume that $\mathfrak{S}$ corresponds to **single linkage clustering**.

- Fix a finite metric space (Z, d_Z). For each $\varepsilon \geqslant 0$ consider the equivalence relation $\sim_\varepsilon$ on Z given by $z \sim_\varepsilon z'$ if and only if there exist $z_0, z_1, \ldots, z_n$ in Z s.t. $z_0 = z$, $z_n = z'$ and $d_Z(z_i, z_{i+1}) \leqslant \varepsilon$. We define

$$u_Z(z, z') := \min\{\varepsilon \geqslant 0 \;\; s.t. \; z \sim_\varepsilon z'\}.$$

- It turns out that u_Z is an ultrametric on Z and that $\mathfrak{S}\big((Z, d_Z)\big) = (Z, u_Z)$ **[CM07]**



Stability

- So, can regard a hierarchical clustering procedure as a map from metric spaces to metric spaces.

- What about the question we set out to investigate?

  Theorem (**[CM07]**). *For all $X, Y \in \mathcal{M}$,*

$$d_{\mathcal{GH}}\big(\mathfrak{S}(X), \mathfrak{S}(Y)\big) \leqslant d_{\mathcal{GH}}\big(X, Y\big).$$

Stability

- So, can regard a hierarchical clustering procedure as a map from metric spaces to metric spaces.

- What about the question we set out to investigate?

  Theorem (**[CM07]**). *For all $X, Y \in \mathcal{M}$,*

  $$d_{\mathcal{GH}}\big(\mathfrak{S}(X), \mathfrak{S}(Y)\big) \leqslant d_{\mathcal{GH}}\big(X, Y\big).$$

Proof. Let $\eta = d_{\mathcal{GH}}\big(X, Y\big)$ and C be a correspondence between X and Y s.t.

$$|d_X(x, x') - d_Y(y, y')| \leqslant 2\eta \text{ for all } (x, y), (x', y') \in C.$$

Fix $(x, y), (x', y') \in C$ and let $x = x_0, x_1, \ldots, x_n = x' \in X$ be s.t. $u_X(x, x') = \max_i d_X(x_i, x_{i+1})$. For each $i \in \{1, \ldots, n-1\}$ pick $y_i \in Y$ s.t. $(x_i, y_i) \in C$ and let $y_0 = y$, $y_n = y'$. Then, it follows that

$$u_Y(y, y') \leqslant \max_i d_Y(y_i, y_{i+1}) \leqslant \max_i d_X(x_i, x_{i+1}) + 2\eta = u_X(x, x') + 2\eta.$$

$\square$

Stability – lower bounds

- Can view stability results for invariants as providing **lower bounds** for the GH distance.

- how good is the Lower bound given by the previous Theorem?

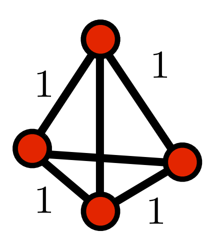
  Remark. *The bound is tight. Indeed, pick X_a to be two points at distance $a > 0$. Then, $\mathfrak{S}(X_a) = X_a$. Hence, the equality holds for X_a and X_b, $a, b > 0$; i.e.:*

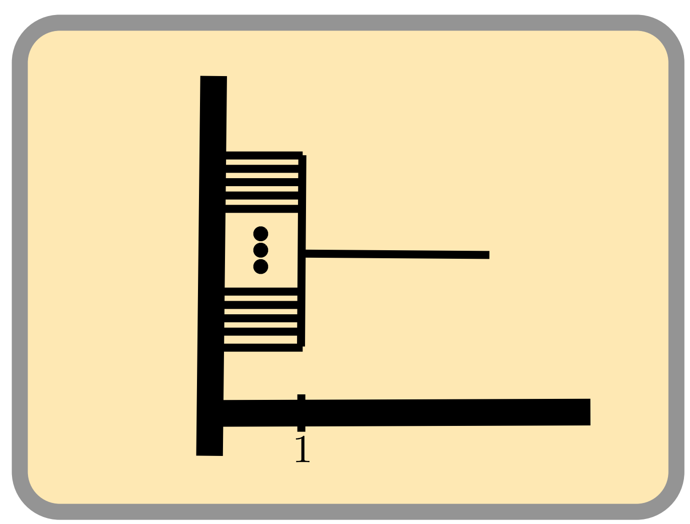
$$d_{\mathcal{GH}}\big(\mathfrak{S}(X_a), \mathfrak{S}(X_b)\big) = d_{\mathcal{GH}}\big(X_a, X_b\big).$$

- But there are cases that suggest one should hope for more.

L_n

Δ_n





what's next:

- go beyond 0-th Homology
- use functions to **probe** the data/shapes.

Simplicial complexes and friends

Definition. • *Given a set of points X and $k = 0, 1, 2, \ldots$, a k-**simplex** is a an unordered list $\{x_0, x_1, \ldots, x_k\}$ of different points in X. The **faces** of this simplex are all the (k-1)-simplices of the form $\{x_0, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k\}$ for some $i \in \{0, 1, \ldots, k\}$.*

- *A **simplicial complex** K is a finite collection of simplices such that every face of a simplex of K is also in K and the intersection of any two simplices is either empty or a common face of each of them.*

- *A **filtration** $\mathcal{K}$ of a simplicial complex K is a nested sequence of subcomplexes $\emptyset = K_{\alpha_0} \subseteq K_{\alpha_1} \subseteq \cdots \subseteq K_{\alpha_m} = K$, where $\alpha_0 < \alpha_1 < \cdots < \alpha_m$ is an ordered sequence of real numbers.*

- *Given a simplex $\sigma \in K$, the **filtration value** $F(\sigma)$ of σ is given by*

$$F(\sigma) = \alpha_{i(\sigma)}$$

where $i(\sigma) = \min\{i \ s.t.\ \sigma \in K_{\alpha_i}\} - 1.$

Simplicial complexes and friends

Definition. • *Given a set of points X and $k = 0, 1, 2, \ldots$, a k-**simplex** is a an unordered list $\{x_0, x_1, \ldots, x_k\}$ of different points in X. The **faces** of this simplex are all the (k-1)-simplices of the form $\{x_0, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k\}$ for some $i \in \{0, 1, \ldots, k\}$.*

- *A **simplicial complex** K is a finite collection of simplices such that every face of a simplex of K is also in K and the intersection of any two simplices is either empty or a common face of each of them.*

- *A **filtration** $\mathcal{K}$ of a simplicial complex K is a nested sequence of subcomplexes $\emptyset = K_{\alpha_0} \subseteq K_{\alpha_1} \subseteq \cdots \subseteq K_{\alpha_m} = K$, where $\alpha_0 < \alpha_1 < \cdots < \alpha_m$ is an ordered sequence of real numbers.*

- *Given a simplex $\sigma \in K$, the **filtration value** $F(\sigma)$ of σ is given by*

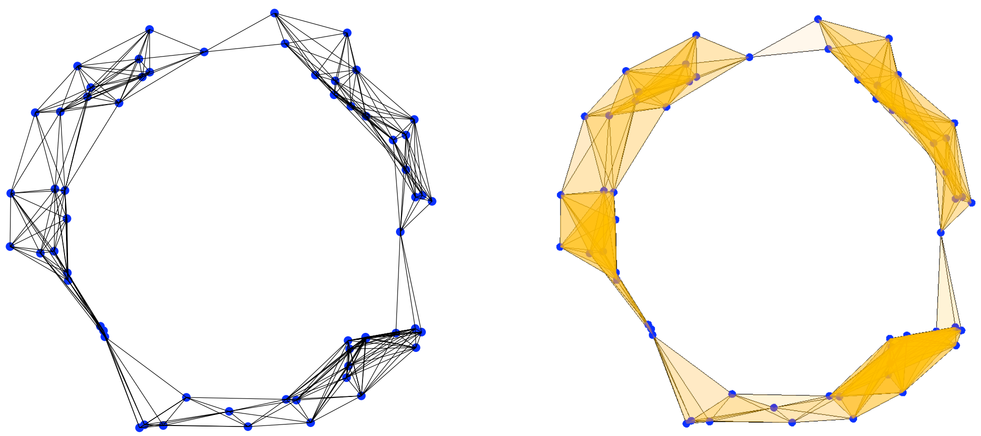
$$F(\sigma) = \alpha_{i(\sigma)}$$

where $i(\sigma) = \min\{i \text{ s.t. } \sigma \in K_{\alpha_i}\} - 1$.

Remark. *Now we see a construction of Simplicial complexes and filtrations that arises from finite metric spaces.*

Rips Simplicial complexes and filtrations

Definition. *Given* $(X, d_X) \in \mathcal{M}$ *and a parameter* $\alpha > 0$*, the* **Rips complex** $R_\alpha(X, d_X)$ *is the abstract simplicial complex of vertex set* X*, whose simplices are those* $\sigma \subset X$ *s.t.*



- $\sigma \neq \emptyset$
- $\mathbf{diam}\,(\sigma) < 2\alpha$.

The **Rips filtration** *of* (X, d_X)*, noted* $\mathcal{R}(X, d_X)$*, is the nested family of Rips complexes obtained by varying parameter* α *from* 0 *to* $+\infty$.

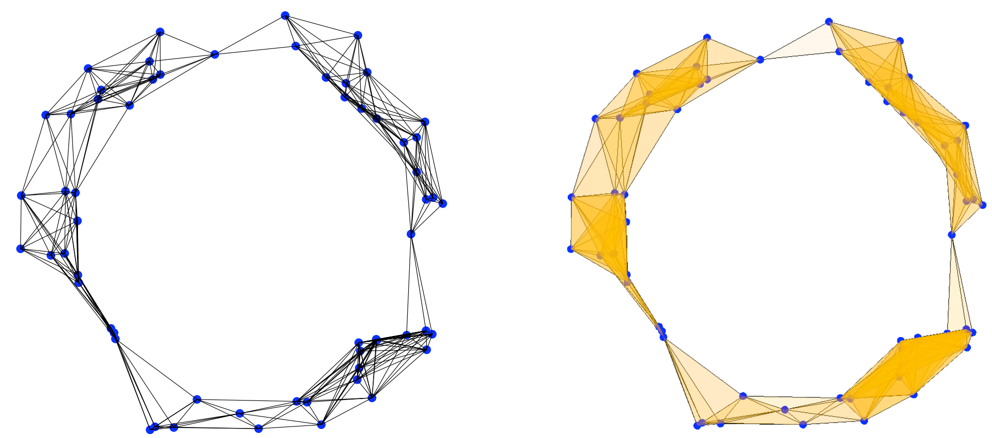
Note that underlying simplicial complex over which the Rips filtration is defined is $K(X)$ (collection of non-empty subsets of X). Also, given any $\sigma \in K(X)$, $F(\sigma) = \frac{1}{2}\mathbf{diam}\,(\sigma)$.

Remark. *Now we see a construction of Simplicial complexes and filtrations that arises from finite metric spaces.*

Rips Simplicial complexes and filtrations

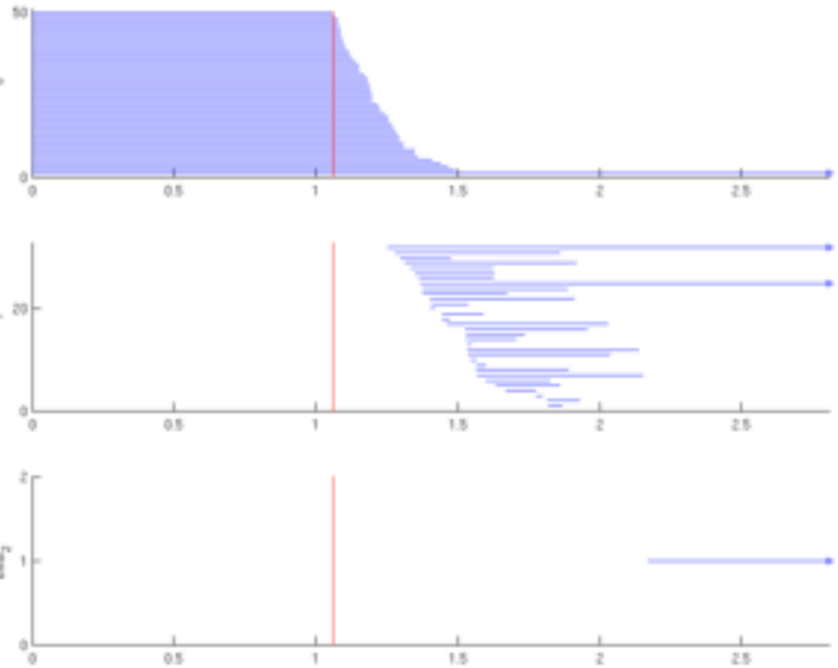
Definition. *Given* $(X, d_X) \in \mathcal{M}$ *and a parameter* $\alpha > 0$*, the* **Rips complex** $R_\alpha(X, d_X)$ *is the abstract simplicial complex of vertex set* X*, whose simplices are those* $\sigma \subset X$ *s.t.*

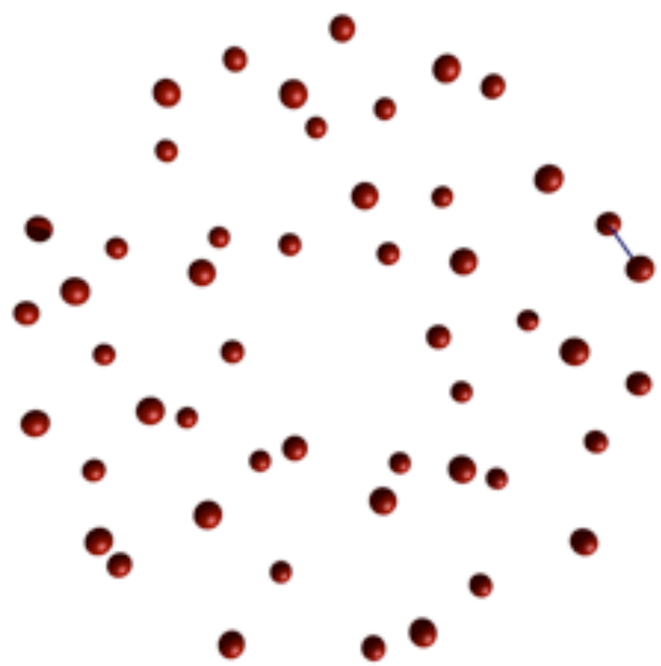
- $\sigma \neq \emptyset$
- $\mathbf{diam}\,(\sigma) < 2\alpha$.

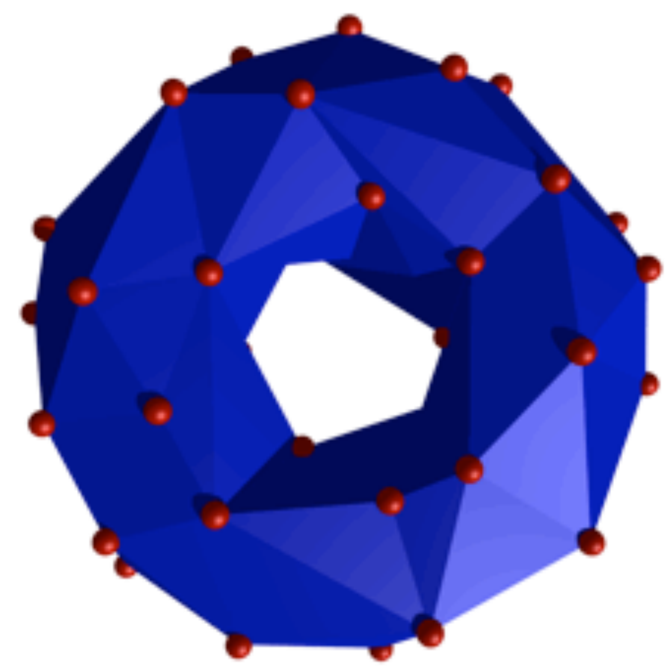


The **Rips filtration** *of* (X, d_X)*, noted* $\mathcal{R}(X, d_X)$*, is the nested family of Rips complexes obtained by varying parameter* α *from* 0 *to* $+\infty$*.*

Note that underlying simplicial complex over which the Rips filtration is defined is $K(X)$ (collection of non-empty subsets of X). Also, given any $\sigma \in K(X)$, $\boxed{F(\sigma) = \frac{1}{2}\mathbf{diam}\,(\sigma).}$







Remark. *We now want to compute certain invariants out of the filtrations. These will be* ***analogues to the dendrograms*** *we discussed in the situation of clustering.*

Persistence diagrams

- Recall that filtration $\mathcal{K}$ of a simplicial complex K is a nested sequence of subcomplexes $\varnothing = K_{\alpha_0} \subseteq K_{\alpha_1} \subseteq \cdots \subseteq K_{\alpha_m} = K$, where $\alpha_0 < \alpha_1 < \cdots < \alpha_m$ are in $\mathbb{R}$.

- The inclusion maps induce a **persistence module**, involving their k-dimensional homology groups:

$$H_k(K_{\alpha_0}) \xrightarrow{\phi_0^1} H_k(K_{\alpha_1}) \xrightarrow{\phi_1^2} \cdots \xrightarrow{\phi_{m-1}^m} H_k(K_{\alpha_m}). \tag{1}$$

- The structure of this persistence module can be encoded as a multi-set of points $\mathrm{D}_k\mathcal{K}$, called the *k-th persistence diagram* of $\mathcal{K}$

Let $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ and $\Delta = \{(x, x) : x \in \overline{\mathbb{R}}\}$.

Definition. • *The k-th **persistence diagram** $\mathrm{D}_k\mathcal{K}$ of the filtration $\mathcal{K}$ is a multi-subset of the extended plane $\overline{\mathbb{R}}^2$ contained in*

$$\Delta \cup \{\alpha_0, \cdots, \alpha_m\} \times \{\alpha_0, \cdots, \alpha_m, \alpha_\infty = +\infty\}.$$

- *The multiplicity of all points in Δ is set to $+\infty$, while the multiplicities of the points of the form (α_i, α_j), $0 \leqslant i < j \leqslant +\infty$, are defined in terms of the ranks of the homomorphisms $\phi_i^j = \phi_{j-1}^j \circ \cdots \circ \phi_i^{i+1}$.*

Remark. *Two persistence diagrams can be compared using the **bottleneck distance**.*

Definition. *The bottleneck distance $d_{\mathrm{B}}^{\infty}(A, B)$ between two multi-sets in $(\overline{\mathbb{R}}^2, l^\infty)$ is the quantity $\min_\gamma \max_{p \in A} \|p - \gamma(p)\|_\infty$, where γ ranges over all bijections from A to B.*

shapes/spaces

$$\left(\mathcal{M}, d_{\mathcal{GH}}\right)$$

$\longrightarrow$

signatures (persistence diagrams)

$$\left(\mathcal{D}, d_B^{\infty}\right)$$

X, Y

$D_k\mathcal{R}(X),\ D_k\mathcal{R}(Y)$

Stability results

Theorem (I, [CCGMO09]). *For any* ***finite*** *metric spaces* (X, d_X) *and* (Y, d_Y)*, and any* $k \in \mathbb{N}$,

$$d_{\mathrm{B}}^{\infty}\big(\mathrm{D}_k\mathcal{R}(X, d_X), \mathrm{D}_k\mathcal{R}(Y, d_Y)\big) \leqslant d_{\mathcal{GH}}\big((X, d_X), (Y, d_Y)\big).$$

Stability results

Theorem (I, [CCGMO09]). *For any* ***finite*** *metric spaces* (X, d_X) *and* (Y, d_Y), *and any* $k \in \mathbb{N}$,

$$d_{\mathrm{B}}^{\infty}\big(\mathrm{D}_k\mathcal{R}(X, d_X), \mathrm{D}_k\mathcal{R}(Y, d_Y)\big) \leqslant d_{\mathcal{GH}}\big((X, d_X), (Y, d_Y)\big).$$

Proof. Let $\eta = d_{\mathcal{GH}}(X, Y)$ and d be a metric on $Z = X \sqcup Y$ s.t. $d_{\mathcal{H}}^{(Z,d)}(X, Y) = \eta$. Then (Z, d) is a finite metric space of cardinality $n = \#X + \#Y$. Hence, it can be embedded *isometrically* into $(\mathbb{R}^n, \ell^{\infty})$. Let Z', X', Y' be subsets of $\mathbb{R}^n$ s.t. $Z' = X' \cup Y'$, $X \sim_{\text{isom}} X'$, $Y \sim_{\text{isom}} Y'$ and $Z \sim_{\text{isom}} Z'$. Then, $d_{\mathcal{H}}^{(\mathbb{R}^n, \ell^{\infty})}(X', Y') = \eta$. This means that $\delta_{X'}, \delta_{Y'} : \mathbb{R}^n \to \mathbb{R}^+$ defined by

$$p \mapsto \min_{x' \in X'} \|x' - p\|_{\ell^{\infty}}$$

and

$$p \mapsto \min_{y' \in Y'} \|y' - p\|_{\ell^{\infty}}$$

are s.t. $\|\delta_{X'} - \delta_{Y'}\|_{L^{\infty}} \leqslant \eta$. One can see that $\delta_{X'}$ and $\delta_{Y'}$ are *tame* and hence their persistence diagrams are η-close in the bottleneck distance according to standard stability theorem. But, then these persistence diagrams agree with the persistence diagrams of the Čech filtrations of X' and Y'. Finally, since the underlying metric is ℓ^{∞}, the Čech filtrations agree with the Rips filtrations. $\square$

Remark. *The bound is tight. Indeed, fix* $\varepsilon > 0$ *and let* X *be a set of two points at distance* 2 *and* Y *a set of two points at distance* $2+2\varepsilon$. *Then,*

$$d_{\mathcal{GH}}\big((X,d_X),(Y,d_Y)\big) = \varepsilon,$$

and

$$D_0\mathcal{R}(X,d_X) = \{(0,+\infty),(0,1)\},\ D_0\mathcal{R}(Y,d_Y) = \{(0,+\infty),(0,1+\varepsilon)\}.$$

- This theorem states GH-**stability** of persistence diagrams arising from Rips filtrations.
- Another way of saying this: it provides a lower bound for the GH distance! Can I use it for object recognition?
- Not very discriminative– can do better: use functions!

Metric spaces endowed with functions

We now consider triples (X, d_X, f_X) where $(X, d_X) \in \mathcal{M}$ and $f_X : X \to \mathbb{R}$. Let $\mathcal{M}_1$ denote the collection of all such triples. We declare $X, Y \in \mathcal{M}_1$ to be **isomorphic** whenever there exist an isometry $\Phi : X \to Y$ s.t. $f_Y \circ \Phi = f_X$.

We put a metric on the collection of all isomorphism classes of $\mathcal{M}_1$ by suitably extending the GH distance:

$$d^1_{\mathcal{GH}}(X, Y) = \inf_C \max\left(\frac{1}{2}\|\Gamma_{X,Y}\|_{L^\infty(C\times C)}, \|f_X - f_Y\|_{L^\infty(C)}\right).$$

Metric spaces endowed with functions

We now consider triples (X, d_X, f_X) where $(X, d_X) \in \mathcal{M}$ and $f_X : X \to \mathbb{R}$. Let $\mathcal{M}_1$ denote the collection of all such triples. We declare $X, Y \in \mathcal{M}_1$ to be **isomorphic** whenever there exist an isometry $\Phi : X \to Y$ s.t. $f_Y \circ \Phi = f_X$.

We put a metric on the collection of all isomorphism classes of $\mathcal{M}_1$ by suitably extending the GH distance:

$$d^1_{\mathcal{GH}}(X, Y) = \inf_{C} \max\left(\frac{1}{2}\|\Gamma_{X,Y}\|_{L^\infty(C\times C)}, \|f_X - f_Y\|_{L^\infty(C)}\right).$$

Definition. *Given $(X, d_X, f_X) \in \mathcal{M}_1$ and a parameter $\alpha > 0$, the* **modified Rips complex** *$R_\alpha(X, d_X, f_X)$ is the abstract simplicial complex of vertex set $X_\alpha := f_X^{-1}((-\infty, \alpha])$, whose simplices are those $\sigma \subset X_\alpha$ s.t.*

- $\sigma \neq \emptyset$
- $\mathbf{diam}\,(\sigma) < 2\alpha$.

The **modified Rips filtration** *of (X, d_X, f_X), noted $\mathcal{R}(X, d_X, f_X)$, is the nested family of modified Rips complexes obtained by varying parameter α from 0 to $+\infty$.*

Given any $\sigma \in K(X)$, $\boxed{F(\sigma) = \max\left(\frac{1}{2}\mathbf{diam}\,(\sigma), \max_{x\in\sigma} f_X(x)\right).}$

Stability results with functions

Theorem (II, [CCGMO09]). *For any **finite** metric spaces endowed with functions (X, d_X, f_X) and (Y, d_Y, f_Y), and any $k \in \mathbb{N}$,*

$$d_{\mathrm{B}}^{\infty}\big(\mathrm{D}_k\mathcal{R}(X, d_X, f_X), \mathrm{D}_k\mathcal{R}(Y, d_Y, f_Y)\big) \leqslant d_{\mathcal{GH}}^{1}\big((X, d_X, f_X), (Y, d_Y, f_Y)\big).$$

Stability results with functions

Theorem (**II**, [**CCGMO09**]). *For any **finite** metric spaces endowed with functions (X, d_X, f_X) and (Y, d_Y, f_Y), and any $k \in \mathbb{N}$,*

$$d_{\mathrm{B}}^{\infty}\big(\mathrm{D}_k\mathcal{R}(X, d_X, f_X), \mathrm{D}_k\mathcal{R}(Y, d_Y, f_Y)\big) \leqslant d_{\mathcal{GH}}^{1}\big((X, d_X, f_X), (Y, d_Y, f_Y)\big).$$

Proof. Similar arguments, need to invoke stronger stability result of [**CCSGGO09**]. □

Remark. *When $f_X = f_Y = 0$ we recover Theorem (I).*

Remark.

- *Goal is to obtain lower bounds for GH distance.*
- ***Idea**: let functions f_X and f_Y depend on the metric!*
- *Need **canonical constructions**: methods for constructing a function out of a metric that can be applied to any metric space.*
- *Example: **eccentricity**. Given a metric space (X, d_X), one can form the triple (X, d_X, ecc^X) where $ecc^X(x) = \max_{x' \in X} d_X(x, x')$.*
- *So the idea is to try to come up with a **rich** family of maps $h : \mathcal{M} \to \mathcal{M}_1$.*

Definition. *For each* $L > 0$ *let* $\mathcal{H}_L$ *denote the class of maps* $h : \mathcal{M} \to \mathcal{M}_1$ *s.t.*

$$\|f_X - f_Y\|_{L^\infty(C)} \leqslant L \cdot \frac{1}{2} \|\Gamma_{X,Y}\|_{L^\infty(C \times C)}$$

for all $X, Y \in \mathcal{M}$, $C \in \mathcal{C}(X, Y)$, *where* $h(X, d_X) = (X, d_X, f_X)$ *and* $h(Y, d_Y) = (Y, d_Y, f_Y)$.

(some kind of Lipschitz continuity across different metric spaces)

Definition. *For each $L > 0$ let $\mathcal{H}_L$ denote the class of maps $h : \mathcal{M} \to \mathcal{M}_1$ s.t.*

$$\|f_X - f_Y\|_{L^\infty(C)} \leqslant L \cdot \frac{1}{2}\|\Gamma_{X,Y}\|_{L^\infty(C\times C)}$$

for all *$X, Y \in \mathcal{M}$, $C \in \mathcal{C}(X, Y)$, where $h(X, d_X) = (X, d_X, f_X)$ and $h(Y, d_Y) = (Y, d_Y, f_Y)$.*

(some kind of Lipschitz continuity across different metric spaces)

Remark. *Note that if $h \in \mathcal{H}_L$, then*

$$d^1_{\mathcal{GH}}\big(h(X), h(Y)\big) \leqslant \max(1, L) \cdot d_{\mathcal{GH}}\big(X, Y\big).$$

Example. *Let h_{ecc} be the map that assigns to each metric space (X, d_X) the triple (X, d_X, ecc^X). Then, $h_{ecc} \in \mathcal{H}_2$.*

Proof. *Let C be s.t. $\|\Gamma_{X,Y}\|_{L^\infty(C\times C)} \leqslant 2\eta$. Then,*

$$ecc^X(x) = \max_{x' \in X} d_X(x, x') \geqslant d_X(x, x') \geqslant d_Y(y, y') - 2\eta$$

for all $(x, y), (x', y') \in C$. Then, by symmetry, $|ecc^X(x) - ecc^Y(y)| \leqslant 2\eta$ and the conclusion follows. $\square$

Corollary. *For all* $X, Y \in \mathcal{M}$, *and* $k \in \mathbb{N}$

$$\frac{1}{\max(1, L)} \sup_{h \in \mathcal{H}_L} d_{\mathrm{B}}^{\infty}\big(\mathrm{D}_k \mathcal{R}(h(X)), \mathrm{D}_k \mathcal{R}(h(Y))\big) \leqslant d_{\mathcal{GH}}\big(X, Y\big).$$

Corollary. *For all* $X, Y \in \mathcal{M}$, *and* $k \in \mathbb{N}$

$$\frac{1}{\max(1,L)} \sup_{h \in \mathcal{H}_L} d_{\mathrm{B}}^{\infty}\big(\mathrm{D}_k\mathcal{R}(h(X)), \mathrm{D}_k\mathcal{R}(h(Y))\big) \leqslant d_{\mathcal{GH}}\big(X, Y\big).$$

Remark (Aggregation properties)**.** *Let* $h, h' \in \mathcal{H}_L$. *Then,*

- $\sup(h) : \mathcal{M} \to \mathcal{M}_1$ *given by* $(X, d_X) \mapsto (X, d_X, \max(f_X))$ *is in* $\mathcal{H}_L$.
- $\max(h, h') : \mathcal{M} \to \mathcal{M}_1$ *given by* $(X, d_X) \mapsto (X, d_X, \max(f_X, f'_X))$ *is in* $\mathcal{H}_L$.
- $(h + h') : \mathcal{M} \to \mathcal{M}_1$ *given by* $(X, d_X) \mapsto (X, d_X, f_X + f'_X)$ *is in* $\mathcal{H}_{2L}$.
- *For any* $\alpha, \beta \in \mathbb{R}$, $\alpha \cdot h + \beta$ *given by* $(X, d_X) \mapsto (X, d_X, \alpha \cdot f_X + \beta)$ *is in* $\mathcal{H}_{|\lambda| L}$.

Note that $\mathrm{D}_k\mathcal{R}(h(X))$ *and* $\mathrm{D}_k\mathcal{R}\big((\alpha \cdot h + \beta)(X)\big)$ *are not related by a simple transformation.*

Remark (Critique)**.**

- *It is difficult to find many functions in $\mathcal{H}_L$ that are easily computable.*
- *For example, $N-$ ecc eccentricities are too expensive:*

$$x \mapsto \max_{x_1,\ldots,x_N} \min_{i\neq j} d_X(x_i, x_j)$$

is in $\mathcal{H}_2$ but for N large is not an option. (complexity is $O\big((\#X)^{N+1}\big)$.

- *Also, functions such as L^p eccentricites are beyond $\mathcal{H}_L$. For example, one may consider*

$$ecc_p^X(x) = \left(\frac{1}{\#X}\sum_{x'\in X} d_X^p(x,x')\right)^{1/p},\ p \geqslant 1.$$

But there's a choice of ***probability measure*** *implicit in this.. Need to make this explicit in the formulation.*

mm-spaces: more admissible functions

Definition. *An mm-space (measure metric space) will be a finite metric space endowed with a* ***probability measure****: a triple* (X, d_X, μ_X)*, where* $\mu_X(x) > 0$ *for all* $x \in X$ *and* $\sum_{x \in X} \mu_X(x) = 1$*. We say that two mm-spaces are isomorphic if there is an isometry which also respects the weights. Let* $\mathcal{M}^w$ *denote the collection of all (finite) mm-spaces. Similarly, we may define* $\mathcal{M}^w_1$*, the collection of all cuadruples* (X, d_X, μ_X, f_X) *where* $(X, d_X, \mu_X) \in \mathcal{M}^w$ *and* $f_X : X \to \mathbb{R}$*.*

mm-spaces: more admissible functions

Definition. *An mm-space (measure metric space) will be a finite metric space endowed with a* ***probability measure****: a triple* (X, d_X, μ_X)*, where* $\mu_X(x) > 0$ *for all* $x \in X$ *and* $\sum_{x \in X} \mu_X(x) = 1$*. We say that two mm-spaces are isomorphic if there is an isometry which also respects the weights. Let* $\mathcal{M}^w$ *denote the collection of all (finite) mm-spaces. Similarly, we may define* $\mathcal{M}_1^w$*, the collection of all cuadruples* (X, d_X, μ_X, f_X) *where* $(X, d_X, \mu_X) \in \mathcal{M}^w$ *and* $f_X : X \to \mathbb{R}$*.*

Definition (Coupling)**.** *Given two mm-spaces* X, Y*, a* ***coupling*** μ *of* X *and* Y *is a probability measure on* $X \times Y$ *marginals* X *and* Y*. Since one regard* μ *as a matrix of size* $\#X$ *times*$\#Y$*, the conditions are that*

- $\mu(x, y) \geqslant 0$
- $\sum_x \mu(x, y) = \mu_Y(y)$ *for all* $y \in Y$
- $\sum_y \mu(x, y) = \mu_X(x)$ *for all* $x \in X$

mm-spaces: more admissible functions

Definition. *An mm-space (measure metric space) will be a finite metric space endowed with a* ***probability measure****: a triple* (X, d_X, μ_X)*, where* $\mu_X(x) > 0$ *for all* $x \in X$ *and* $\sum_{x \in X} \mu_X(x) = 1$*. We say that two mm-spaces are isomorphic if there is an isometry which also respects the weights. Let* $\mathcal{M}^w$ *denote the collection of all (finite) mm-spaces. Similarly, we may define* $\mathcal{M}_1^w$*, the collection of all cuadruples* (X, d_X, μ_X, f_X) *where* $(X, d_X, \mu_X) \in \mathcal{M}^w$ *and* $f_X : X \to \mathbb{R}$*.*

Definition (Coupling)**.** *Given two mm-spaces* X, Y*, a* ***coupling*** μ *of* X *and* Y *is a probability measure on* $X \times Y$ *marginals* X *and* Y*. Since one regard* μ *as a matrix of size* $\#X$ *times*$\#Y$*, the conditions are that*

- $\mu(x, y) \geqslant 0$
- $\sum_x \mu(x, y) = \mu_Y(y)$ *for all* $y \in Y$
- $\sum_y \mu(x, y) = \mu_X(x)$ *for all* $x \in X$

Definition. *For all* $X, Y \in \mathcal{M}_1^w$*, we also define the distance*

$$d^1_{\mathcal{GW},\infty}(X, Y) := \inf_{\mu \in \mathcal{M}(\mu_X, \mu_Y)} \max\left(\frac{1}{2}\|\Gamma_{X,Y}\|_{L^\infty(C(\mu)\times C(\mu))}, \|f_X - f_Y\|_{L^\infty(C(\mu))}\right).$$

Definition. *For each* $L > 0$ *let* $\mathcal{H}_L^w$ *denote the class of maps* $h : \mathcal{M}^w \to \mathcal{M}_1^w$ *s.t.*

$$\|f_X - f_Y\|_{L^\infty} \leqslant L \cdot \frac{1}{2}\|\Gamma_{X,Y}\|_{L^\infty(R(\mu)\times R(\mu))}$$

for all $X, Y \in \mathcal{M}$, $\mu \in \mathcal{M}(\mu_X, \mu_Y)$, *where* $h(X, d_X, \mu_X) = (X, d_X, \mu_X, f_X)$ *and* $h(Y, d_Y, \mu_Y) = (Y, d_Y, \mu_Y, f_Y)$.

(some kind of Lipschitz continuity across different mm-spaces)

Definition. *For each* $L > 0$ *let* $\mathcal{H}_L^w$ *denote the class of maps* $h : \mathcal{M}^w \to \mathcal{M}_1^w$ *s.t.*

$$\|f_X - f_Y\|_{L^\infty} \leqslant L \cdot \frac{1}{2}\|\Gamma_{X,Y}\|_{L^\infty(R(\mu)\times R(\mu))}$$

for all $X, Y \in \mathcal{M}$, $\mu \in \mathcal{M}(\mu_X, \mu_Y)$, *where* $h(X, d_X, \mu_X) = (X, d_X, \mu_X, f_X)$ *and* $h(Y, d_Y, \mu_Y) = (Y, d_Y, \mu_Y, f_Y)$.

(some kind of Lipschitz continuity across different mm-spaces)

Remark. *Note that if* $h \in \mathcal{H}_L^w$, *then*

$$d^1_{\mathcal{GW},\infty}(h(X), h(Y)) \leqslant \max(1, L) \cdot d_{\mathcal{GW},\infty}(X, Y).$$

Example. *For* $p \in [1, \infty]$ *let* h_{ecc_p} *be the map that assigns to each mm-space* (X, d_X, μ_X) *the cuadruple* (X, d_X, μ_X, ecc_p^X). *Then,* $h_{ecc_p} \in \mathcal{H}_2^w$.

Proof. *Let* μ *be s.t.* $\|\Gamma_{X,Y}\|_{L^\infty(C(\mu)\times C(\mu))} \leqslant 2\eta$. *Then, also,*

$$\left\|d_X(x, \cdot) - d_Y(y, \cdot)\right\|_{L^\infty(C(\mu))} \text{ for all } (x, y) \in C(\mu).$$

Write, $\left|ecc_p^X(x) - ecc_p^X(x)\right| = \left|\|d_X(x,\cdot)\|_{\ell^p(\mu_X)} - \|d_Y(y,\cdot)\|_{\ell^p(\mu_Y)}\right| \leqslant \|d_X(x,\cdot) - d_Y(y,\cdot)\|_{\ell^p(\mu)} \leqslant 2\eta$ *for all* $(x, y) \in C(\mu)$. $\square$

Corollary. *For all* $X, Y \in \mathcal{M}^w$, *and* $k \in \mathbb{N}$

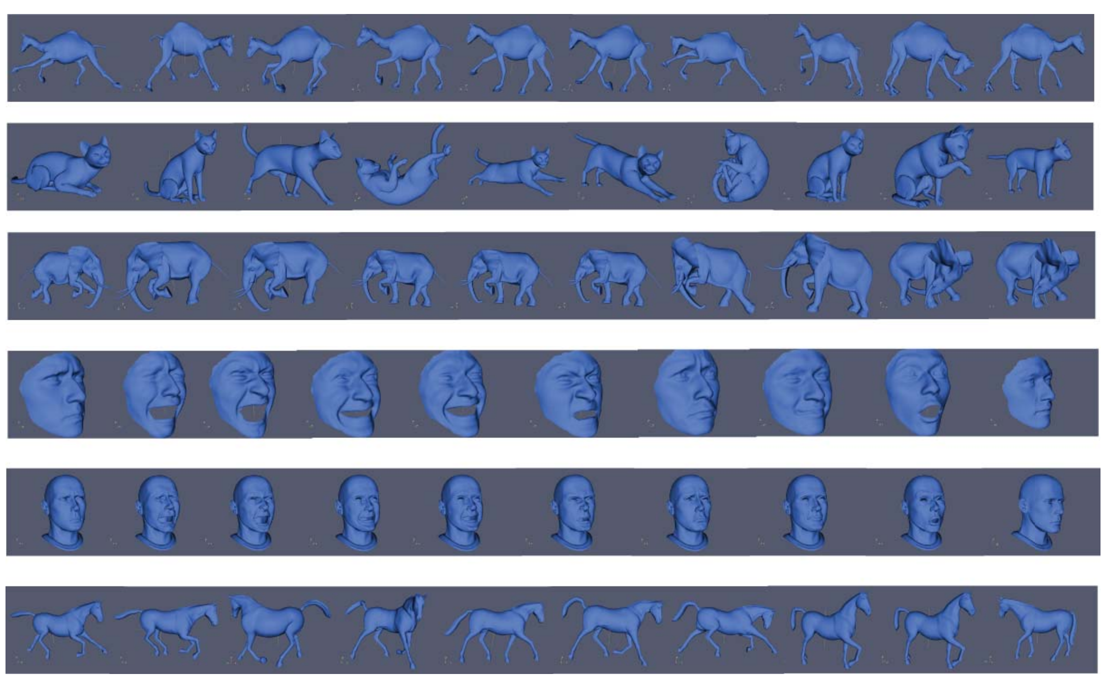
$$\frac{1}{\max(1, L)} \sup_{h \in \mathcal{H}_L^w} d_{\mathrm{B}}^{\infty}\big(\mathrm{D}_k \mathcal{R}(h(X)), \mathrm{D}_k \mathcal{R}(h(Y))\big) \leqslant d_{\mathcal{GW},\infty}(X, Y).$$

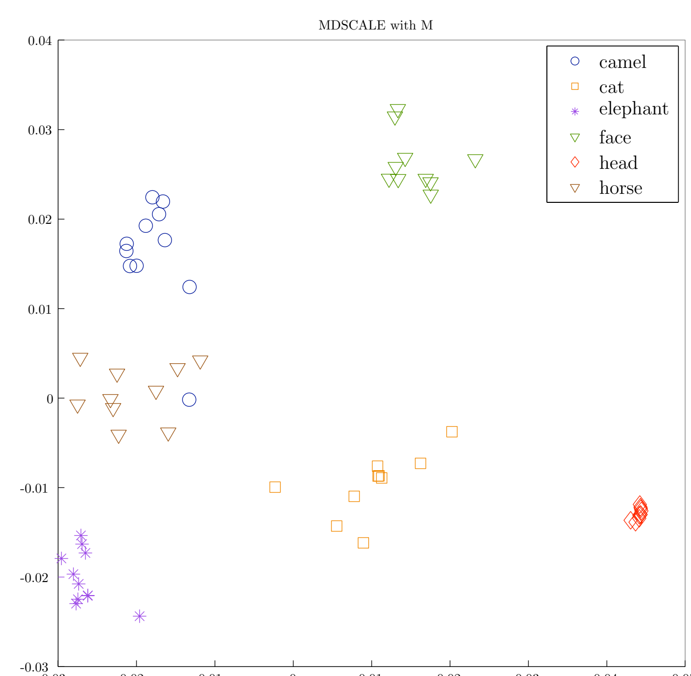
Remark (Aggregation properties)**.**

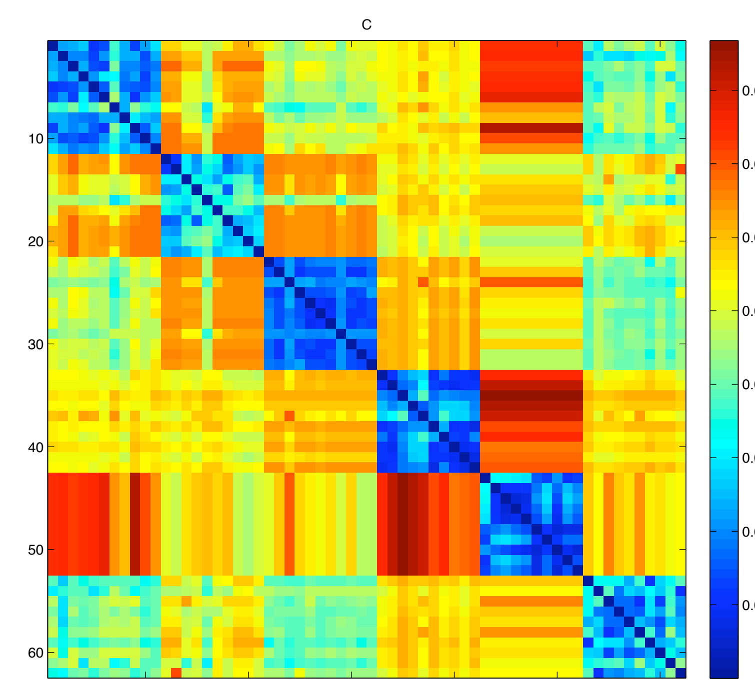
- *The collection* $\mathcal{H}_L^w$ *has similar aggregation properties as* $\mathcal{H}_L$*.*
- *There is a sense in which* $\mathcal{H}_L^w$ *contains all maps in* $\mathcal{H}_L$*.*

Discussion

- The computation of these lower bounds lead to solving **bottleneck assignment problems**. Standard problems.
- Ran these on a database of shapes. There are interesting details about the implementation.
- The stability theorem is very interesting as it permits to define the notion of a limit Rips persistent diagram of a compact metric space.







Bibliography

[**BBI**] Burago, Burago and Ivanov. A course on Metric Geometry. AMS, 2001

[**HK-03**] A. Ben Hamza, Hamid Krim: Probabilistic shape descriptor for triangulated surfaces. ICIP (1) 2005: 1041-1044

[**Osada-02**] Robert Osada, Thomas A. Funkhouser, Bernard Chazelle, David P. Dobkin: Matching 3D Models with Shape Distributions. Shape Modeling International 2001: 154-166

[**SC-00**] S. Belongie and J. Malik (2000). "Matching with Shape Contexts". IEEE Workshop on Contentbased Access of Image and Video Libraries (CBAIVL-2000).

[**Gromov-80**] M. Gromov. Metric structures for Riemannian and non- Riemannian spaces, volume 152 of Progress in Mathematics. Birkhäuser Boston Inc., Boston, MA, 1999.

[**CCGMO09**] Chazal, Cohen-Steiner, Guibas, Memoli, Oudot. Gromov-Hausdorff stable signatures for shapes using persistence. SGP 2009.

[**M08-euclidean**]F. Mémoli. Gromov-Hausdorff distances in Euclidean spaces. NORDIA-CVPR-2008.

[**M07**] F.Mémoli. On the use of gromov-hausdorff distances for shape comparison. In Proceedings of PBG 2007, Prague, Czech Republic, 2007.

[**MS04**] F. Mémoli and G. Sapiro. Comparing point clouds. In SGP '04: Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing, pages 32–40, New York, NY, USA, 2004. ACM.

[**MS05**] F. Mémoli and G. Sapiro. A theoretical and computational framework for isometry invariant recognition of point cloud data. Found. Comput. Math., 5(3):313–347, 2005.

[**CM07**] G. Carlsson and F. Mémoli. Persistent clustering and a theorem of J. Kleinberg. *preprint*, 2007. 2005.

http://math.stanford.edu/~memoli