Coarse geometry of the Gromov-Hausdorff space and applications to computational topology

Nicolò Zava
Institute of Science and Technology Austria (ISTA)

Abstract

In this paper, we motivate the applications of the Gromov-Hausdorff distance in computational topology and topological data analysis. We present how dimension theory and metric geometry can provide theoretical limits to the precision of metric invariants. We conclude by discussing generalisations of the Gromov-Hausdorff distance and further research directions.

The paper's title is that of the talk delivered at the RIMS symposium 'General topology and related fields', but this presentation has a different focus. A more suitable title is 'The role of the Gromov-Hausdorff distance and embeddability results in computational topology'.

1 Introduction

Computational topology is a field of mathematics at the intersection of computational geometry, computer science and topology. The central question that is investigated is the following: how can we compute if two spaces (e.g., metric spaces, triangulated manifolds, simplicial complexes) are the same? In topology, to distinguish between topological spaces, we usually rely on topological invariants. Topological dimensions (small inductive, large inductive, covering, Hausdorff, etc.) are classical examples: if two topological spaces have different dimensions, then they are not homeomorphic ([24, 32]). In computational topology, a similar approach is pursued, but there are two crucial differences:

(a) these invariants have to be efficiently computable and comparable, and
(b) they have to be stable under small perturbation to limit the impact of possible errors in the representation of the spaces.

A central example of such invariants is persistent homology. In the past decades, successful applications of these invariants to compare real-world datasets and extrapolate relevant patterns lead to the establishment of a self-standing subject known as topological data analysis (TDA). We refer the interested reader to [20] and to [28] for a database of real-world applications of TDA.

As for the requirement (b), we expect that an invariant returns similar values if it receives in input either a continuous object, e.g., a manifold, representing the ground truth or a finite *point cloud* (i.e., a finite metric space) obtained by sampling it (see Figure 1).

Introduced by Edwards ([23]) and rediscovered and generalised by Gromov ([29]), the Gromov-Hausdorff distance provides a theoretical tool to measure how two metric spaces resemble each other. In particular, we can formally discuss how well a point cloud,

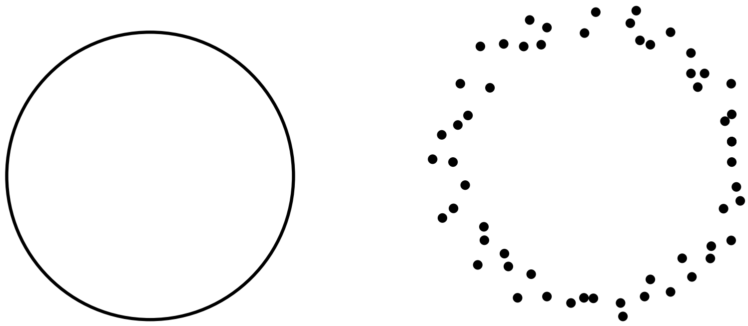

Figure 1: In the applications, we need a distance notion showing that a circle and a finite sample of it are similar even though the latter may contain small errors.

obtained by sampling a continuous object, approximates the ground truth. Its role in computational topology has grown in the last decades ([43, 38, 37, 39, 40, 12, 13, 1, 2, 3, 35]).

In this survey paper, after recalling some basic notions and properties of the Gromov-Hausdorff distance, we discuss its role in computational topology (§2). After mentioning its computational limits, we describe how stable invariants can be used to approximate it (§3). We then show how notions coming from dimension theory and coarse geometry can be applied to demonstrate theoretical limits to the metric distortion created by stable invariants (§4). We conclude with a list of further research directions requiring suitable generalisations or modification of the Gromov-Hausdorff distance (§5).

2 Distances between metric spaces and the Gromov-Hausdorff distance

Let us recall some notions and results in metric geometry. We refer the reader to [10, 47, 52].

Definition 2.1. For a pair (X, d) consisting of a set X and a map $d\colon X \times X \to \mathbb{R}$, we define the following properties:
(M1) for every $x, y \in X$, $d(x, y) \geq 0$ and $d(x, x) = 0$;
(M2) for every $x, y \in X$, $d(x, y) = d(y, x) = 0$ if and only if $x = y$;
(M3) for every $x, y \in X$, $d(x, y) = d(y, x)$ (*symmetry*);
(M4) for every $x, y, z \in X$, $d(x, y) \leq d(x, z) + d(z, y)$ (*triangle inequality*).
The pair (X, d) is called:
(a) a *pseudo-metric space* if it satisfies (M1), (M3) and (M4);
(b) a *metric space* if it is an extended pseudo-metric space satisfying (M2).

The path metric on the vertex set of an undirected graph, which associates to a pair of vertices the length of the shortest path connecting them, is a classical example of a metric. If the edges are weighted with positive values, the path metric can be similarly defined as the length of a path is the sum of weights that it follows.

For every metric space (X, d), the *Hausdorff distance* between two subsets Y and Z of X is defined as

$$d_H(Y, Z) = \inf\{\varepsilon > 0 \mid Y \subseteq B_\varepsilon(Z),\ \text{and } Z \subseteq B_\varepsilon(Y)\},$$

where, for a subset A of Z, $B_\varepsilon(A) = \bigcup_{a\in A} B_\varepsilon(a)$ and $B_\varepsilon(a)$ denotes the open ball centred in a with radius ε.

The Hausdorff distance can be equivalently characterised using the notion of correspondence, which will be important in the sequel. A relation $\mathcal{R}$ between two sets A and B is a *correspondence* if, for every $a \in A$, there exists $b_a \in B$ such that $(a, b_a) \in \mathcal{R}$ and, vice versa, for every $b \in B$, there is $a_b \in A$ satisfying $(a_b, b) \in \mathcal{R}$. Equivalently, if the maps obtained by restricting the two canonical projections of the product $A \times B$ to $\mathcal{R}$ are surjective. Then, if Y and Z are subsets of a metric space X,

$$d_H(Y, Z) = \inf_{\mathcal{R}\subseteq Y\times Z \text{ correspondence}} \sup_{(y,z)\in\mathcal{R}} d(y, z). \tag{1}$$

It is known that d_H is a metric on the family of non-empty compact subsets of Z. Since they are bounded, the distance between each pair of objects is finite, and two subsets are at distance 0 if and only if they coincide since they are closed. This space is sometimes referred as the *hyperspace of* Z and denoted by 2^Z.

The Hausdorff distance is the stepping stone to define other, less embedding-dependent distances. First, we modify the Hausdorff distance to identify two subspaces that are isometric even though they may be located in different areas of the ambient space. Let us consider this definition in the particular case where Z is $\mathbb{R}^d$, but it can be straightforwardly extended.

Let X and Y be two subsets of $\mathbb{R}^d$. Their *Euclidean-Hausdorff distance* d_{EH} (see, for example, [5]) is defined as follows:

$$d_{EH}(X, Y) = \inf\{d_H(X, f(Y)) \mid f \in \mathrm{Isom}(\mathbb{R}^d)\},$$

where $\mathrm{Isom}(\mathbb{R}^d)$ denotes the set of all isometries of $\mathbb{R}^d$. This notion is a metric on the space of isometry classes of compact subsets of $\mathbb{R}^d$.

The Euclidean-Hausdorff distance can be conveniently differently characterised.

Theorem 2.2 (see [5]). *Let X and Y be two subsets of $\mathbb{R}^d$. Consider them as metric spaces. Then,*

$$d_{EH}(X, Y) = \inf_{\substack{i_X : X\to\mathbb{R}^d \text{ and} \\ i_Y : Y\to\mathbb{R}^d \\ \text{isometric embeddings}}} d_H(i_X(X), i_Y(Y)).$$

The characterisation of the Euclidean-Hausdorff distance described in Theorem 2.2 inspires the definition of the Gromov-Hausdorff distance. If X and Y are two arbitrary metric spaces, we look for all possible isometric embeddings into some ambient metric space where they are compared with the Hausdorff distance.

Definition 2.3 ([23, 29])**.** Given two metric spaces X and Y, their *Gromov-Hausdorff distance* d_{GH} is the value

$$d_{GH}(X,Y) = \inf_{Z \text{ metric space}} \inf_{\substack{i_X : X \to Z \text{ and} \\ i_X : X \to Z \\ \text{isometric embeddings}}} d_H(i_X(X), i_Y(Y))$$

(see Figure 2 for a representation).

The previous definition is not precise since the first infimum is not well-defined since all metric spaces form a proper class. However, it is enough to investigate the pseudo-metric spaces whose support is the disjoint union of the two metric spaces we are comparing. This fact will follow from the construction provided in the proof of Theorem 2.4.

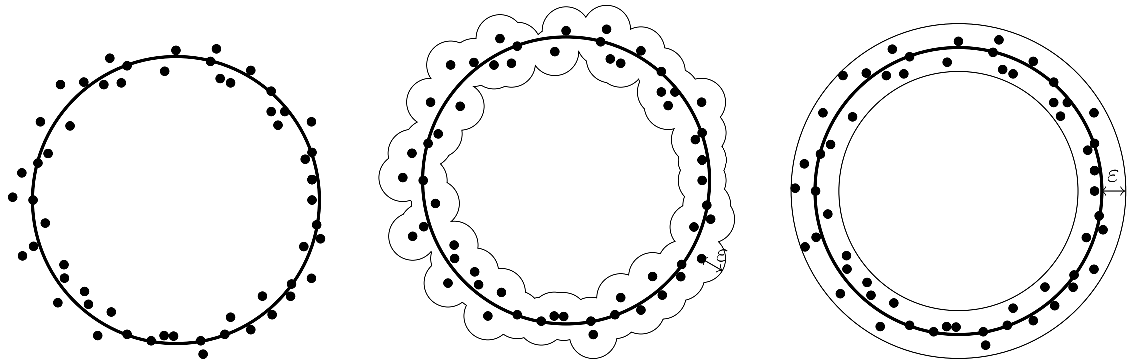

Figure 2: A representation of the fact that the spaces illustrated in Figure 1 are ε-close with respect to the Euclidean-Hausdorff and Gromov-Hausdorff distances.

Given Theorem 2.2, it is trivial that, pointwisely, $d_{GH} \leq d_{EH}$. In the Remark 2.6, we show that they do not coincide even in simple cases.

Although intuitive, the definition of the Gromov-Hausdorff distance provided is not practical for computations since it relies on a third, unknown metric space. A change of perspective allows for a more manageable characterisation.

If (X, d_X) and (Y, d_Y) are two metric spaces, and $\mathcal{R} \subseteq X \times Y$ is a correspondence between them, the *distortion of* $\mathcal{R}$ is the value

$$\operatorname{dis} \mathcal{R} = \sup_{(x_1,y_1),(x_2,y_2) \in \mathcal{R}} |d_X(x_1,x_2) - d_Y(y_1,y_2)|.$$

Theorem 2.4 ([10])**.** *Let X and Y be two metric spaces, then*

$$d_{GH}(X,Y) = \frac{1}{2} \inf_{\mathcal{R} \subseteq X \times Y \textit{ correspondence}} \operatorname{dis} \mathcal{R}.$$

Proof. For a fixed $\varepsilon > 0$, let Z be a metric space, $i_X : X \to Z$ and $i_Y : Y \to Z$ two isometric embeddings showing that $d_H(i_X(X), i_Y(Y)) \leq d_{GH}(X,Y) + \varepsilon$ and $\mathcal{S} \subseteq X \times Y$ be a correspondence such that $\sup_{(x,y)\in\mathcal{S}} d(i_X(x), i_Y(y)) \leq d_H(i_X(X), i_Y(Y)) + \varepsilon \leq d_{GH}(X,Y) + 2\varepsilon$. Note that $\operatorname{dis} \mathcal{S} \leq 2 \sup_{(x,y)\in\mathcal{S}} d(i_X(x), i_Y(y))$, and so $\operatorname{dis} \mathcal{S} \leq 2(d_{GH}(X,Y) + \varepsilon)$.

Vice versa, let $\mathcal{R} \subseteq X \times Y$ be a correspondence. We construct a new metric space $Z_{\mathcal{R}}$ as follows. Consider a weighted undirected graph whose vertices are $X \sqcup Y$, and whose edges consist of all the pairs $(x,x') \in X \times X$ weighted $d_X(x,x')$, $(y,y') \in Y \times Y$ weighted $d_Y(y,y')$, and $(x,y) \in \mathcal{R}$ weighted by $\operatorname{dis}\mathcal{R}/2$. Then, $Z_{\mathcal{R}}$ is the $X \sqcup Y$ endowed with the path metric. Consider the canonical inclusions i_X and i_Y of X and Y into Z. Trivially, $d_H(i_X(X), i_Y(Y)) \leq \operatorname{dis}\mathcal{R}/2$. Using the distortion's definition, it can be shown that i_X and i_Y are isometric embeddings. Given that $\mathcal{R}$ is arbitrary, then $d_{GH}(X,Y) \leq \frac{1}{2}\inf_{\mathcal{R}\subseteq X\times Y \text{ correspondence}} \operatorname{dis}\mathcal{R}$. □

As an easy application of the previous characterisation, we obtain the following bounds for the Gromov-Hausdorff distance between two bounded metric spaces. We denote by $\operatorname{diam} X = \sup_{x,y\in X} d(x,y)$ the *diameter* of a metric space X.

Proposition 2.5. *Let X and Y be two bounded metric spaces. Then,*

$$\frac{1}{2}|\operatorname{diam} X - \operatorname{diam} Y| \leq d_{GH}(X,Y) \leq \frac{1}{2}\max\{\operatorname{diam} X, \operatorname{diam} Y\}.$$

Proof. The upper bound is obtained by considering the trivial correspondence $\mathcal{R} = X \times Y$ whose distortion is $\max\{\operatorname{diam} X, \operatorname{diam} Y\}$. As for the lower bound, if $\operatorname{diam} X = \operatorname{diam} Y$, there is nothing to prove. Without loss of generality, we can now suppose that $\operatorname{diam} X > \operatorname{diam} Y + \varepsilon$. For every $\varepsilon > \delta > 0$, there are $x_1, x_2 \in X$ such that $d_X(x_1,x_2) \geq \operatorname{diam} X - \delta$. Let us take an arbitrary correspondence $\mathcal{R} \subseteq X \times Y$. If $y_1, y_2 \in Y$ satisfy $(x_1,y_1),(x_2,y_2) \in \mathcal{R}$,

$$\operatorname{dis}\mathcal{R} \geq |d_X(x_1,x_2) - d_Y(y_1,y_2)| = d_X(x_1,x_2) - d_Y(y_1,y_2) \geq \operatorname{diam} X - \delta - \operatorname{diam} Y.$$

We can conclude since δ can be arbitrarily taken. □

Remark 2.6. Given two point clouds in $\mathbb{R}^d$, the reader may wonder for which reasons the Gromov-Hausdorff distance should be preferred to the Euclidean-Hausdorff distance. The essential advantage of the first is that it does not depend on embedding the point clouds in $\mathbb{R}^d$, and it considers them as metric spaces on their own. Already small examples show the difference in the two approaches (see [38]).

Let $X_2 = \{x_0, x_1, x_2\}$ be the vertices of an equilateral triangle in the plane of edge length 2, and Y be a singleton. The isometry minimizing the Hausdorff distance between X_2 and Y places Y as the centre c of X_2, which achieves $d_{EH}(X_2,Y) = d_H(X_2,\{c\}) = 2\sqrt{3}/3$. However, if we embed X_2 and Y into a different space, we can lower their distance. Indeed, consider the tree T consisting of three leaves $\{a_0,a_1,a_2\}$ and another vertex b that is connected to the three leaves through edges of length 1. Endow it with the path distance. We can isometrically embed X_2 into the leaves of T and Y as its centre b. Then,

$$d_{GH}(X_2,Y) \leq d_H(\{a_1,a_2,a_3\},\{b\}) = 1 < \frac{2\sqrt{3}}{3} = d_{EH}(X_2,Y).$$

According to Proposition 2.5, $d_{GH}(X_2,Y) = 1$.

We can readily extend this construction to every dimension by replacing X_2 with a d-dimensional simplex $X_d = \{x_0,\ldots,x_d\}$ embedded in $\mathbb{R}^d$ with pairwise distances equal to 2 ([38]).

A lower bound on the Gromov-Hausdorff distance depending on the Euclidean-Hausdorff distance was provided in [38].

Theorem 2.7. *For every pair of compact subsets X and Y of $\mathbb{R}^d$,*

$$d_{GH}(X,Y) \le d_{EH}(X,Y) \le c_d\sqrt{M \cdot d_{GH}(X,Y)},$$

where $M = \max\{\operatorname{diam} X, \operatorname{diam} Y\}$ and c_d is a constant depending only on the dimension d.

However, if X and Y are two finite subsets of $\mathbb{R}$, a linear lower bound to d_{GH} depending on d_{EH} can be proved.

Theorem 2.8. ([36, Theorem 3.2]) *For every pair X and Y of compact subsets of $\mathbb{R}$,*

$$\frac{4}{5} d_{EH}(X,Y) \le d_{GH}(X,Y) \le d_{EH}(X,Y).$$

2.1 Properties of the Gromov-Hausdorff distance

The Gromov-Hausdorff distance has nice metric properties. A metric space (X,d) is *geodesic* if, for every pair of points $x, y \in X$, there is an isometric embedding $\gamma\colon [0, d(x,y)] \to X$ such that $\gamma(0) = x$ and $\gamma(1) = y$. Such map γ is called a *geodesic* connecting x and y.

Proposition 2.9. *(a) Given two compact metric spaces X and Y, $d_{GH}(X,Y) = 0$ if and only if they are isometric.*
(b) The space $\mathcal{GH}$, sometimes called Gromov-Hausdorff space*, of isometry classes of compact metric spaces endowed with the Gromov-Hausdorff distance is a metric space.*
(c) The space $\mathcal{GH}$ is separable.
(d) The space $\mathcal{GH}$ is complete.
(e) The space $\mathcal{GH}$ is geodesic.

Proof. We refer to [10] for items (a) and (b), and to [47] for items (c) and (d). Because of its importance, let us hint at the proof of item (e) in the simpler case of finite metric spaces.

Let X and Y be two finite metric spaces, and $\mathcal{R} \subseteq X \times Y$ be a correspondence such that $\operatorname{dis} \mathcal{R} = 2d_{GH}(X,Y)$. Let $T = d_{GH}(X,Y)$. For every $0 < t < T$, define the metric d_t on $\mathcal{R}$ as follows:

$$d_t((x,y),(x',y')) = \left(1 - \frac{t}{T}\right) d_X(x,x') + \frac{t}{T} d_Y(y,y').$$

For each such t, d_t is a metric. Furthermore, $(\mathcal{R}, d_t)$ is a finite metric space.

Define $\gamma\colon [0,T] \to \mathcal{GH}$ as follows:

$$\gamma(t) = \begin{cases} (X, d_X) & \text{if } t = 0, \\ (Y, d_Y) & \text{if } t = T, \\ (\mathcal{R}, d_t) & \text{otherwise.} \end{cases}$$

Let us estimate $d_{GH}(\gamma(s), \gamma(t))$ for every $s, t \in [0, T]$. With simple computations, the inequality $d_{GH}(\gamma(s), \gamma(t)) \leq |s - t|$ can be shown. The triangular inequality implies that this is enough to show that γ is a geodesic ([14]). □

The proof of item (e) in the compact case follows the same structure. The additional ingredients are the followings: the distortion-minimising correspondence still exists if the spaces are compact and the spaces $(\mathcal{R}, d_t)$ are compact.

The geodesics defined in the previous proof are called *straight-line geodesics*. Already in [33], Proposition 2.9(e) was proved, but no explicit geodesics were provided. The construction given in Proposition 2.9(e) shows that we have access to an optimal transformation between two compact metric spaces as soon as a correspondence with optimal distortion is provided.

Geodesics in $\mathcal{GH}$ are not unique. There may be more optimal correspondences that induce distinct geodesics (Example 2.10).

Example 2.10. There may be more than one straight-line geodesics between two (finite) metric spaces. Consider the subsets $X = \{1, 2, 3, 4\}$ and $Y = \{1, 2, 4\}$ of $\mathbb{R}$, and two correspondences $\mathcal{R}_i \subseteq X \times Y$, for $i = 1, 2$, defined as follows:

$$\mathcal{R}_1 = \{(1, 1), (2, 2), (3, 2), (4, 4)\}, \text{ and } \mathcal{R}_2 = \{(1, 1), (2, 2), (3, 4), (4, 4)\}$$

(see Figure 3 for their representation).

It is easy to see that the distortion of both correspondences is 1. Hence, $d_{GH}(X, Y) \leq 1/2$. We have that $d_{GH}(X, Y) = 1/2$ since every correspondence $\mathcal{R}$ between X and Y has to associate two distinct points of X to the same point of Y for the pigeon hole principle. Hence, its distortion is at least 1.

Consider now the straight-line geodesics γ_1 and γ_2 between X and Y that $\mathcal{R}_1$ and $\mathcal{R}_2$ induce. They are distinct since $\gamma_1(t)$ and $\gamma_2(t)$ are not isometric for every $t \in (0, 1/2)$. In fact, for every $t \in (0, 1/2)$, consider the point $(3, 3)$ in $\gamma_1(t)$. For every $(3, 3) \neq z \in \gamma_1(t)$, $d_t((3, 3), z) > 1$. However, for every $t' \in (0, 1/2)$ and $y \in \gamma_2(t')$, there is $y \neq y' \in \gamma_2(t')$ with $d_{t'}(y, y') \leq 1$. Hence, $\gamma_1(t)$ and $\gamma_2(t')$ are not isometric.

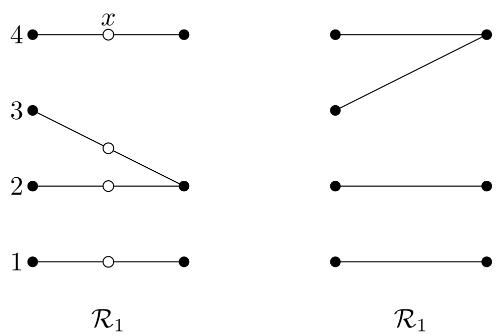

Figure 3: A representation of the two correspondences $\mathcal{R}_1$ and $\mathcal{R}_2$. The metric space $\gamma_1(1/4)$ consists of the hollow dots. The set of distances from the top-most point x is $\{0, 3/2, 2, 3\}$, and it is easy to see that none of the points in $\gamma_2(t)$ has the same set. More precisely, for every point $y \in \gamma_2(t)$, there is $z \in \gamma_2(t)$ such that $0 < d_t(y, z) \leq 1$. Hence, the two geodesics are distinct.

Not all geodesics are straight-line. In [15], the authors provided continuously many distinct geodesics between finite metric spaces.

2.2 Computational limitations of the Gromov-Hausdorff distance

We have discussed properties of the Gromov-Hausdorff distance that are relevant for the applications. In addition to those enlisted in Proposition 2.9, the Gromov-Hausdorff distance does not depend on the particular embeddings of the spaces and can be used to compare discrete and continuous objects. However, there is a fundamental limitation to the practical employment of the Gromov-Hausdorff distance in computer science, which is its computational complexity.

Remark 2.11. Suppose that we have an algorithm to compute the Gromov-Hausdorff distance between two finite metric spaces.

The *graph isomorphism problem* is a decision problem that is not known to be solvable in polynomial time nor to be NP-complete. Let $X = (V, E)$ and $X' = (V', E')$ be two graphs. We can compute in polynomial time their two path metrics (e.g., using Dijkstra's algorithm formulated in 1956). Then, we compare the two metric spaces using the Gromov-Hausdorff distance and obtain a solution for the graph isomorphism problem since $d_{GH}(X, X') = 0$ if and only X and X' are isomorphic.

More precisely, it can be proved the following result, which shows that even approximating it lead to computational bottlenecks.

Theorem 2.12 ([4, 50])**.** *The following decision problem is NP-hard:*
3-Gromov-Hausdorff. *Given two metric trees T_1 and T_2 with unit length edges, determine whether $d_{GH}(T_1, T_2) < 3$.*

Given this computational bottleneck, developing techniques to compute lower and upper bounds to the Gromov-Hausdorff distance between to metric spaces is a fertile research direction.

On the other hand, the Gromov-Hausdorff distance provides a theoretical framework to evaluate computable invariants of metric spaces, formally describing how precisely they reflect the geometry. These aspects will be described in more detail in the following section.

Before concluding the section, let us describe one more layer of complexity. Given a compact metric space M—typically a manifold—and a (finite) sample X of it, a natural question is the following: how dense X in M has to be in order to correctly capture the geometry of M? Intuitively, we would expect that for some parameter $\varepsilon > 0$, if X is *ε-dense* in M, i.e., $d_H(X, M) < \varepsilon$, then $d_{GH}(X, M) = d_H(X, M)$. However, in general this is not the case.

Theorem 2.13 ([2])**.** *For every $\varepsilon > 0$, there are a finite metric space X and a finite subset $X' \subseteq X$ such that*

$$d_{GH}(X, X') \leq \varepsilon d_H(X, X').$$

In [2], the authors provide conditions on manifolds and samples ensuring bounds on the Gromov-Hausdorff distance depending on the Hausdorff distance. In [3], this study is further developed. Particular attention is paid to the case of metric graphs.

3 Stable invariants

In the previous section, we mentioned that computing the Gromov-Hausdorff distance is mostly unfeasible. To practically compare two metric spaces, we often rely on (*metric*) *invariants*. These associate to every metric space X a value $\psi(X)$ of some other metric space in such a way that $\psi(X) = \psi(Y)$ provided that X and Y are two isometric metric spaces. More formally, an invariant is a map $\psi\colon \mathcal{GH} \to \mathcal{X}$ where $\mathcal{X}$ is a metric space. In practical settings, we require that these invariants can be efficiently computed and compared.

Invariants assuming values into a (finite-dimensional) Hilbert space are called *vectorisation methods* and they are particularly relevant for the applications since they can be exploited in machine learning pipelines.

The following property of invariants is essential.

Definition 3.1. Following [39], an invariant $\psi\colon \mathcal{GH} \to \mathcal{X}$ is *stable* if there exists a non-decreasing function $\rho_+\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that

$$d_{\mathcal{X}}(\psi(X), \psi(Y)) \leq \rho_+(d_{GH}(X, Y))$$

for every pair of compact metric spaces X and Y.

Why is stability important? Consider the following decision problems. We simplify the questions for expository reasons.

$\mathcal{P}_{\mathcal{GH}}$: Given two finite metric spaces, are they 'close' in the Gromov-Hausdorff distance?

We have discussed before why this problem is hard to tackle. Let now $\psi\colon \mathcal{GH} \to \mathcal{X}$ be an invariant, and define the following decision problem.

$\mathcal{P}_{\mathcal{X}}$: Given two points in $\mathcal{X}$, are they 'close' relatively to the distance $d_{\mathcal{X}}$?

We can look for a metric space $\mathcal{X}$ in which the problem $\mathcal{P}_{\mathcal{X}}$ is easy to solve. Then, to study whether two spaces $X, Y \in \mathcal{GH}$ are close, we can consider their images $\psi(X)$ and $\psi(Y)$ and solve the problem $\mathcal{P}_{\mathcal{X}}$ relatively to this instance. We would have the following possible situations.

		$\mathcal{P}_{\mathcal{GH}}$	
		Yes	No
$\mathcal{P}_{\mathcal{X}}$	Yes	true positive	false positive
	No	false negative	true negative

Suppose that ψ is stable. If $\mathcal{P}_{\mathcal{X}}$ returns no, $\mathcal{P}_{\mathcal{GH}}$ cannot return yes. This is due to the fact that, intuitively, $d_{\mathcal{X}}(\psi(X), \psi(Y))$ has to be small provided that $d_{GH}(X, Y)$ is small. Hence, false negatives, which are particularly undesirable in applications, are avoided. Indeed, if small changes in the spaces considered, maybe due to some unavoidable experimental errors retrieving or classifying the data, lead to very different outputs, then such method is not reliable.

Furthermore, stable invariants can be used to provide lower bounds to the Gromov-Hausdorff distance. If ψ is a stable invariant and ρ_+ is invertible, then, for every pair of compact metric spaces X and Y,

$$d_{GH}(X,Y) \geq \rho_+^{-1}(d_{\mathcal{X}}(\psi(X),\psi(Y))).$$

Proposition 2.5 shows that the diameter of a metric space is a stable invariant. The following, further examples are taken from [39].

Example 3.2. Let (X,d) be a metric space. We define the following stable invariants:
- its *circumradius* as the value $\mathrm{rad}(X) = \min_{x\in X}\max_{x'\in X} d(x,x')$;
- its *circumradius set* as the subset $\mathcal{C}(X) = \{\max_{x'\in X} d(x,x') \mid x \in X\}$;
- its *eccentricity function* as the map $\mathrm{ecc}_X\colon X \to \mathbb{R}_{\geq 0}$, $\mathrm{ecc}_X(x) = \max_{x'\in X} d_X(x,x')$;
- its *distance set* as the subset $\mathcal{D}(X) = \{d(x,x') \mid x,x' \in X\}$;
- its *local distance set* as the map $\mathcal{L}_X\colon X \to \mathcal{P}(\mathbb{R}_{\geq 0})$, $\mathcal{L}_X(x) = \{d(x,x') \mid x' \in X\}$.

It is easy to see that the invariants introduced in Example 3.2, can be polynomially computed. Less trivial, but still easy is showing that the distance between those invariants can also be polynomially computed. More refined arguments can be found in [39, Remark 3.7].

These invariants are clearly related one another, and they have different 'discerning power'. Namely, if ψ and φ are two invariants such that, for every compact metric space X, $\psi(X) \leq \varphi(X)$, φ has a higher chance than ψ to distinguish more pairs of non-isometric metric spaces. Typically though, a more discerning invariant is more costly to compute.

We announced in Example 3.2 that all the invariants presented are stable. Let us phrase this property more precisely in the following result, which is easy to show (see [39]).

Proposition 3.3. *The circumradius, the circumradius set, the eccentricity function, the distance set and the local distance set are stable invariants. More precisely, for every pair of metric spaces X and Y,*

$$\begin{aligned}\max\{|\mathrm{rad}(X)-\mathrm{rad}(Y)|, |\mathrm{diam}\, X - \mathrm{diam}\, Y|\} &\leq d_H(\mathcal{C}(X),\mathcal{C}(Y)) \leq \\ &\leq \inf_{\substack{\mathcal{R}\subseteq X\times Y\\ \textit{correspondence}}} \sup_{(x,y)\in\mathcal{R}} |\mathrm{ecc}_X(x)-\mathrm{ecc}_Y(y)| \leq \\ &\leq \inf_{\substack{\mathcal{R}\subseteq X\times Y\\ \textit{correspondence}}} \sup_{(x,y)\in\mathcal{R}} d_H(\mathcal{L}_X(x),\mathcal{L}_Y(y)) \leq \\ &\leq 2d_{GH}(X,Y),\end{aligned}$$

and

$$\begin{aligned}|\mathrm{diam}\, X - \mathrm{diam}\, Y| &\leq d_H(\mathcal{D}(X),\mathcal{D}(Y)) \leq \\ &\leq \inf_{\substack{\mathcal{R}\subseteq X\times Y\\ \textit{correspondence}}} \sup_{(x,y)\in\mathcal{R}} d_H(\mathcal{L}_X(x),\mathcal{L}_Y(y)) \leq \\ &\leq 2d_{GH}(X,Y).\end{aligned}$$

More refined invariants, which focus on retrieving specific features, are used in the computational topology community. A cornerstone example is given by persistent homology ([20]), which intuitively capture the size of the homological features of a metric space. In [12, 13] (see also [14]), it is shown that the persistence homology of the Vietoris-Rips and Dowker filtrations is stable.

All the examples of invariants we have discussed are not injective and, intuitively, return false positives. We refer the reader to [39] for a more detailed discussion. In the next section we discuss conditions ensuring that these false positives can or cannot be avoided.

4 Applications of dimension theory and coarse geometry to invariants' precision

Miming the definition of stability, we say that an invariant $\psi\colon \mathcal{GH} \to \mathcal{X}$ *avoids false positives* if there exists a non-decreasing function $\rho_-\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that $\rho_- \to \infty$ and

$$\rho_-(d_{GH}(X,Y)) \leq d_{\mathcal{X}}(\psi(X), \psi(Y)) \tag{2}$$

for every pair of compact metric spaces X and Y.

A lower bound as in (2) prevents false positives since the larger the Gromov-Hausdorff distance, the larger the distance between the two associated invariants.

We investigate the following question.

Question 4.1. *Does there exist an invariant ψ with values into a Hilbert space which is stable and avoids false positives?*

A map $\psi\colon (X, d_X) \to (Y, d_Y)$ between two metric spaces is said to be a *coarse embedding* (see [46]) if there exist two maps $\rho_-, \rho_+\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, called *control functions*, such that $\rho_- \to \infty$ and, for every $x, y \in X$,

$$\rho_-(d_X(x,y)) \leq d_Y(\psi(x), \psi(y)) \leq \rho_+(d_X(x,y)). \tag{3}$$

In particular, an invariant that is stable and avoids false positives is a coarse embedding.

Coarse embeddings have been introduced by Gromov and extensively studied in the field of metric geometry known as coarse geometry. A crucial application of this theory is due to Yu, who proved in [57] that those metric spaces that can be coarsely embedded into a Hilbert space satisfy the Novikov and the coarse Baum-Connes conjectures generalising results contained in [56].

Immediate examples of coarse embeddings are isometric and bi-Lipschitz embeddings, where the control functions are identities and linear maps, respectively.

By studying the geometry of the Gromov-Hausdorff space, we can show that sufficiently regular embeddings cannot exist. Let us denote by $\mathcal{GH}^{<\omega}$ the subspace of $\mathcal{GH}$ consisting of all finite metric spaces.

Fact 4.2. *The spaces $\mathcal{GH}$ and $\mathcal{GH}^{<\omega}$ cannot be isometrically embedded into any Hilbert space.*

Proof. It is easy to show that a Hilbert space has unique geodesics. However, in $\mathcal{GH}$ and $\mathcal{GH}^{<\omega}$ there are pair of spaces with several distinct geodesics between them (see Example 2.10 and [15]). □

To prove the (non-)existence of coarse and bi-Lipschitz embeddings of the Gromov-Hausdorff space into some Hilbert space, more advanced tools are needed.

The result contained in [57] and mentioned above motivated two research directions. On one hand, since an explicit coarse embedding can be hard to construct, a plethora of conditions ensuring its existence have been defined and investigated (see [46]). Among these properties, if a space has finite asymptotic dimension (a dimension notion introduced by Gromov, [30]), then it can be coarsely embedded into a Hilbert space ([57, 31]). On the other hand, examples of metric spaces that cannot be coarsely embedded were constructed for example in [19, 34]. Showing that one of those pathological examples can be coarsely embedded into a metric space X proves that X itself cannot be coarsely embedded into any Hilbert space. Using this approach and modifying constructions described in [55], the following result was proved.

Theorem 4.3 ([59])**.** *The spaces $\mathcal{GH}^{<\omega}$ and $\mathcal{GH}$ cannot be coarsely embedded into any Hilbert space.*

In the same paper, by computing the asymptotic dimension, it is shown that a positive result can be obtained if considering metric spaces with a uniform bound on their cardinality.

Theorem 4.4 ([59])**.** *The space $\mathcal{GH}^{\leq n}$ consisting of all metric spaces with at most n points can be coarsely embedded into any Hilbert space.*

More precisely, it is shown that the asymptotic dimension of $\mathcal{GH}^{\leq n}$ is $n(n-1)/2$.

Let us now consider bi-Lipschitz embeddability. To prove the non-existence of certain bi-Lipschitz embeddings, Assouad defined in his PhD thesis the following dimension notion (see [6, 7]), now named after him. The same notion was already studied in [8]. We refer the interested reader to [49, 26], where the proofs of the following facts can be found.

Let X be a metric space, $A \subseteq X$, and $\rho > 0$. Denote by $N(A, \rho)$ the minimal number of closed balls with radius ρ needed to cover A.

Definition 4.5. Given a metric space X, a subset A is (M, s)-*homogeneous* if any ball of radius r centred in a point of A and intersected with A can be covered by at most $M(r/\rho)^s$ balls of radius $\rho \leq r$. In formula, for every $x \in A$, $N(B(x, r) \cap A, \rho) \leq M(r/\rho)^s$.

Definition 4.6. Let X be a metric space and A be a subset of it. The *Assouad dimension* $\dim_A A$ of A is the infimum of all s such that A is (M, s)-homogeneous for some $M \geq 1$.

Proposition 4.7. *Let X and Y be metric spaces, and $\varphi\colon X \to Y$ be a bi-Lipschitz embedding. Then,* $\dim_A \varphi(X) = \dim_A X$.

Example 4.8. For every subset Y of $\mathbb{R}^n$, $\dim_A Y \leq n$.

As an immediate consequence of Proposition 4.7 and Example 4.8, if a metric space X satisfies $\dim_A X = \infty$, then it cannot be bi-Lipschitz embedded into any $\mathbb{R}^n$.

Theorem 4.9 ([59]). *The subspace $\mathcal{GH}^{<\omega}_{\leq R} \subseteq \mathcal{GH}^{<\omega}$ consisting of finite metric spaces with diameter bounded by R has infinite Assouad dimension. Therefore, it cannot be bi-Lipschitz embedded into any $\mathbb{R}^d$.*

Coarse geometry and dimension theory have already been applied in computational topology to study spaces of persistence diagrams, summaries of the persistent homology of spaces. Here we mention [11, 44, 45, 48, 9, 54].

5 Generalisations of the Gromov-Hausdorff distance and further research directions

Let us conclude this paper with a discussion about different research directions. We have discussed the role of the Gromov-Hausdorff distance in computational topology. The crucial assumption is that datasets and shapes can be represented as metric spaces. However, this assumption is too restrictive in many real-world settings and these structures fail to capture essential properties. In the sequel, we discuss some of these situations.

Example 5.1. Let X be a finite set, and $w\colon X \times X \to \mathbb{R}$ be a map. Following [14], we call (X, w) a *network*. For every pair of points $x, y \in X$, we can interpret $w(x, y)$ as their distance. The further axioms that need to be fulfilled to make w a metric may not be satisfied in general. Consider for example a directed graph $G = (V, E)$. For the sake of simplicity, let us assume that it is *strongly connected*, i.e., for every pair of vertices $x, y \in V$, there are directed paths going from x to y and from y to x. We can equip V with its *path quasi-metric* w, obtained by assigning to the pair $(x, y) \in V \times V$ the length of the shortest path going from x to y. Then, w has all the desired properties of a metric but for symmetry. Such a distance function is an example of a *quasi-metric*, i.e., it satisfies axioms (M1), (M2) and (M4) in Definition 2.1. By enforcing symmetry on a quasi-metric, we are loosing essential characteristics of the system.

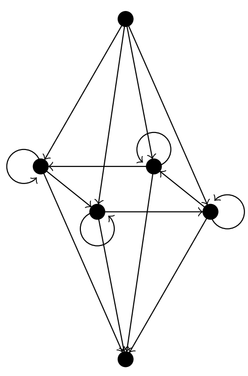

Figure 4: An example of a directed graph.

In the previous example, w still retains a distance-like flavour, but there are other datasets where this intuition cannot be straightforwardly found. In biology, the three-dimensional structure of DNA strands inside the nucleus of a cell is often captured by collecting the frequency of bases' proximity (for example, see Hi-C data). The result of several of these detections is a network (X, w), where X is a string of bases and, for every $x, y \in X$, $w(x, y)$ is defined as the number of occurrences (experiments) in which x and y were found close to each other. The analysis of such data is a frontier direction at the intersection of mathematics and biology. In these kinds of networks, the triangular inequality cannot be enforced in general.

We define the *network distance* between two networks X and Y ([14]) as the value

$$d_{\mathcal{N}}(X, Y) = \frac{1}{2} \inf_{\substack{\mathcal{R} \subseteq X \times Y \\ \text{correspondence}}} \operatorname{dis} \mathcal{R}.$$

According to Theorem 2.4, the network distance generalises the Gromov-Hausdorff distance. Furthermore, and maybe surprisingly, the network distance is symmetric and satisfies the triangular inequality.

A characterisation of the network distance using embeddings into a common space, as in the definition of the Gromov-Hausdorff distance (Definition 2.3), cannot be provided in general. However, this characterisation can be given for the network distance between quasi-metric spaces ([58]).

Another example of intrinsically asymmetric objects is given by directed spaces, where a collection of directions or paths is privileged. A different approach to generalise the Gromov-Hausdorff distance to those spaces is proposed in [25].

Example 5.2. Another situation where metric spaces are too restrictive comes from higher-order interactions, where relationships are not limited to pairwise connections, but involve more complex interactions. The importance of studying those aspects is getting more and more attention in several fields, such as social network analysis, biology and neuroscience. More suitable structures to model that information are (weighted) hypergraphs, simplicial complexes and filtrations.

A *hypergraph* is a pair $H = (V, E)$, where V is a set and E is a set of *hyperedges*, i.e., finite non-empty subsets $e \subseteq V$. A hypergraph is an (*abstract*) *simplicial complex* if E is closed under taking subsets. In this case, its elements are called *simplices*. A *weighted hypergraph* is a triple (V, E, w) where (V, E) is a hypergraph and $w \colon E \to \mathbb{R}$. Similarly, a *filtration* is a triple (V, E, w) where (V, E) is a simplicial complex and $w \colon E \to \mathbb{R}$ is monotonous (i.e., if $e \subseteq e'$ are two simplices, then $w(e) \leq w(e')$).

Motivated by the applications, there is a growing interest in computational topology around those objects [16, 60]. In [41], the author defines a distance notion between two finite filtrations (X, w_X) and (Y, w_Y). A *tripod* (Z, f, g) between X and Y consists of a finite set Z and two surjective maps

$$X \xleftarrow{\;f\;} Z \xrightarrow{\;g\;} Y,$$

Then,

$$d_{\mathcal{F}}(X, Y) = \inf_{(Z,f,g) \text{ tripod}} \sup_{\emptyset \neq A \subseteq Z \text{ finite}} |w_X(f(A)) - w_Y(g(A))|.$$

The author showed that it is symmetric and satisfies the triangular inequality. In [41], the stability of persistence homology with respect to this distance is shown.

Example 5.3. In crystallography and material science, lattices and periodic point sets are central objects used to represent periodic crystal structures. In those applications, each point—representing atoms or molecules—is relevant and so a distance notion between two such structures needs to preserve them. Let us present two definitions.

For every $X, Y \subseteq \mathbb{R}^d$, we define

$$d_B(X,Y) = \inf_{f\colon X\to Y \text{ bijection}} \sup_{x\in X} ||x - f(x)||, \quad \text{and}$$
$$d_{EB}(X,Y) = \inf_{\psi\in \mathrm{Isom}(\mathbb{R}^d)} d_B(X, \psi(Y)).$$

In both cases the infimum is assumed to be infinite if no bijection exists. We call d_B the *bottleneck distance* and d_{EB} the *Euclidean bottleneck distance* ([22]).

We refer to [22, 21] for constructions of topological invariants to study lattices and periodic point sets. Furthermore, we mention that the authors of [27] discussed embeddability results of the bottleneck and Euclidean-bottleneck distances and deduced theoretical limits of similar invariants.

Example 5.4. We have already discussed how the Gromov-Hausdorff distance is stable under small perturbations; if experimental errors occur while sampling a metric space, the Gromov-Hausdorff distance does not change sensibly provided that the errors are small. However, even a single outlier—i.e., a point that differs significantly from the expectation—can heavily impact the distance (Figure 5).

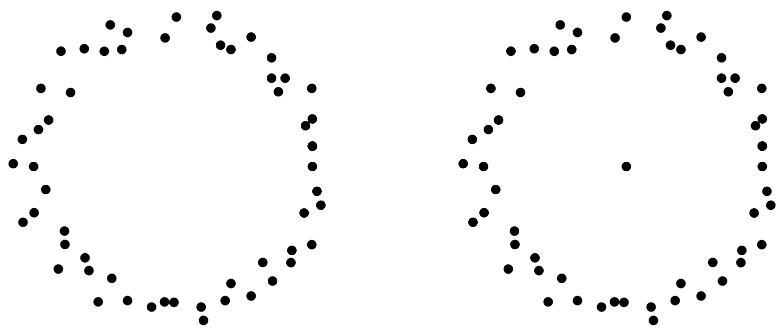

Figure 5: An example showing that an unwanted single outlier in the centre of the second point cloud drastically increases the Gromov-Hausdorff distance.

This sensitivity to outliers is an undesirable feature in most applications. Indeed, collecting data, their creation cannot be ruled out in general, and constructing topological tools to reduce their impact is an active research direction in the computational topology community.

A key-idea to tackle this issue and, at the same time, bypass some of the computational bottlenecks associated with the Gromov-Hausdorff distance, is representing the data as metric measure spaces, i.e., metric spaces equipped with a compatible probability measure. With this viewpoint change, tools from optimal transport [53] have been successfully applied. Several metrics have been proposed and studied to compare metric measure spaces. Examples range from the classical box and Gromov-Prohorov distances [29] to the recently introduced Gromov-Wasserstein [40] and Gromov-Monge distances [42]. These notions allow for more efficient comparisons while preserving essential geometric properties. In addition, probability measures on the point sets can encode the relevance or trustworthiness of the sample points.

Example 5.5. In several applications, not all the points in a point cloud are of the same kind, and considering their spacial interactions is important. As a motivating example, advances in biology enable collecting data describing cells' positions and types. Consistently with the motivation, we can represent these datasets as coloured point clouds (see Figure 6). Defining computable invariants and theoretical approaches to study these datasets is an active research direction. Here we mention, for example, [18, 17, 51].

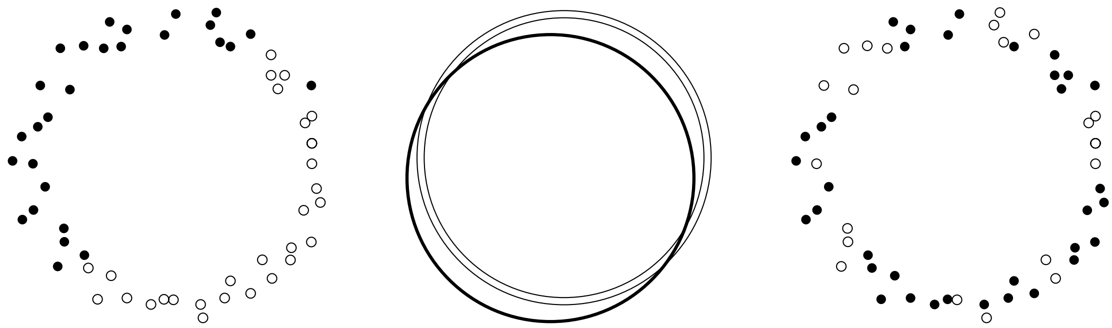

Figure 6: Three different datasets coloured in black and white showing different spatial configurations.

References

[1] H. Adams, J. Bush, N. Clause, F. Frick, M. Gómez, M. Harrison, R. A. Jeffs, E. Lagoda, S. Lim, F. Mémoli, M. Moy, N. Sadovek, M. Superdock, D. Vargas, Q. Wang, L. Zhou, *Gromov-Hausdorff distances, Borsuk-Ulam theorems, and Vietoris-Rips complexes*, to appear, 2022.

[2] H. Adams, F. Frick, S. Majhi, N. McBride, *Hausdorff vs Gromov-Hausdorff distances*, `arXiv:2309.16648v1`.

[3] H. Adams, S. Majhi, F. Manin, Ž. Virk, N. Zava, *Lower Bounding the Gromov-Hausdorff distance in Metric Graphs*, `arXiv:2411.09182`.

[4] P.K. Agarwal, K. Fox, A. Nath, A. Sidiropoulos, Y. Wang, *Computing the Gromov-Hausdorff Distance for Metric Trees*, ACM Transactions on Algorithms 14 (2), 1–20 (2018).

[5] S.A. Antonyan, *The Gromov-Hausdorff hyperspace of a Euclidean space*, Advances in Mathematics, Volume 363, 2020, 106977.

[6] P. Assouad, *Étude d'une dimension métrique liée à la possibilité de plongements dans* $\mathbb{R}^n$ (French with English summary), C. R. Acad. Sci. Paris Sér. A–B 28 8(1979), no.15, A731–A734.

[7] A. Assouad, *Plongements lipschitziens dans* $\mathbb{R}^n$, Bull. Soc. Math. France 111 (1983), 429–448.

[8] M.G. Bouligand, *Ensembles impropres et nombre dimensionnel*, Bull. Sci. Math. 52 (1928), 320–344 and 361–376.

[9] P. Bubenik, A. Wagner, *Embeddings of persistence diagrams into Hilbert spaces*, J. Appl. Comput. Topol. 4 (2020), no. 3, 339–351.

[10] D. Burago, Y. Burago, S. Ivanov, *A Course in Metric Geometry*, AMS, 2001.

[11] M. Carriére, U. Bauer, *On the metric distorsion of embedding persistence diagrams into separable hilbert spaces*, SoCg 2019.

[12] F. Chazal, D. Cohen-Steiner, L. J. Guibas, F. Mémoli, S. Y. Oudot, *Gromov-Hausdorff stable signatures for shapes using persistence*, in Computer Graphics Forum, volume 28, 1393–1403, Wiley Online Library (2009).

[13] F. Chazal, V. De Silva, S. Oudot, *Persistence stability for geometric complexes*, Geometriae Dedicata, 173(1):193–214 (2014).

[14] S. Chowdhury, F. Mémoli, *A functorial Dowker theorem and persistent homology of asymmetric networks*, Journal of Applied and Computational Topology 2 (1) (2018), 115–175.

[15] S. Chowdhury, F. Mémoli, *Explicit geodesics in Gromov-Hausdorff space*. Electron. Res. Announc. 25, 48–59 (2018).

[16] S. Chowdhury, T. Needham, E. Semrad, B. Wang, Y. Zhou, *Hypergraph co-optimal transport: metric and categorical properties*, J Appl. and Comput. Topology (2023).

[17] S. Cultrera di Montesano, O. Draganov, H. Edelsbrunner, M. Saghafian, *Chromatic Topological Data Analysis*, preprint `arXiv:2406.04102` (2024).

[18] S. Cultrera di Montesano, O. Draganov, H. Edelsbrunner, M. Saghafian, *Chromatic Alpha Complexes*, preprint `arXiv:2212.03128` (2022).

[19] A. N. Dranishnikov, G. Gong, V. Lafforgue, G. Yu, *Uniform embeddings into Hilbert space and a question of Gromov*, Canad. Math. Bull. 45 (2002), no. 1, 60–70.

[20] H. Edelsbrunner, J. Harer, *Computational topology: an introduction*, American Mathematical Soc., 2010.

[21] H. Edelsbrunner, T. Heiss, *Merge Trees of Periodic Filtrations*, preprint, `arXiv:2408.16575` (2024).

[22] H. Edelsbrunner, T. Heiss, V. Kurlin, P. Smith, M. Wintraecken, *The density fingerprint of a periodic point set*, SoCG 2021.

[23] D. A. Edwards, *The Structure of Superspace*, published in: Studies in Topology, Academic Press, 1975.

[24] R. Engelking, *Dimension Theory*, North Holland Publishing Company, 1978.

[25] L. Fajstrup, B.T. Fasy, W. Li, L. Mezrag, T. Rask, F. Tombari, Ž Urbančič, *Gromov-Hausdorff distance for directed spaces*, preprint, `arXiv:2408.14394` (2024).

[26] J.M. Fraser, *Assouad Dimension and Fractal Geometry*, Cambridge University Press (2020).

[27] A. Garber, Z. Virk, N. Zava, *Geometry of spaces of lattices and periodic point sets*, `arXiv:2310.07594`.

[28] B. Giunti, J. Lazovskis, B. Rieck, *DONUT: Database of Original and Non-theoretical Uses of Topology*, database, `https://donut.topology.rocks/` (2022).

[29] M. Gromov, *Structures Métriques pour les Variétés Riemanniennes* (J. Lafontaine and P. Pansu, eds.), Textes Mathématiques, 1, CEDIC, Paris, 1981.

[30] M. Gromov, *Asymptotic invariants of infinite groups*, in "Geometric group theory, Vol. 2" (Sussex, 1991), 1–295, London Math. Soc. Lecture Note Ser., 182, Cambridge Univ. Press, Cambridge, 1993.

[31] N. Higson, J. Roe, *Amenable group actions and the Novikov conjecture*, J. Reine Angew. Math. 519 (2000), 143–153.

[32] W. Hurewicz, H. Wallman, *Dimension Theory*, Princeton University Press, 1941.

[33] A.O. Ivanov, N.K. Nikolaeva, A.A. Tuzhilin, *The Gromov-Hausdorff metric on the space of compact metric spaces is strictly intrinsic*, Math Notes 100, 883–885 (2016).

[34] V. Lafforgue, *Un renforcement de la propriété (T)*, Duke Math. J. 143 (2008), no. 3, 559–602.

[35] S. Lim, F. Mémoli, Z. Smith, *The Gromov-Hausdorff distance between spheres*, Geom. Topol. 27 (2023) 3733–3800.

[36] S. Majhi, J. Vitter, C. Wenk, *Approximating Gromov-Hausdorff Distance in Euclidean Space*, Comput. Geom. 116 (2024), Paper No. 102034, 32 pp.

[37] F. Mémoli, *On the use of Gromov-Hausdorff distances for shape comparison*, Point Based Graphics, 2 (2007).

[38] F. Mémoli, *Gromov-Hausdorff distances in Euclidean spaces*, in IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, June (2008), pp. 1–8.

[39] F. Mémoli, *Some properties of Gromov-Hausdorff distances*, Discrete & Computational Geometry 48 (2) (2012), 416–440.

[40] F. Mémoli, *Wasserstein Distances and the Metric Approach to Object Matching*, Found. Comput. Math. 11, 417–487 (2011).

[41] F. Mémoli, *A distance between filtered spaces via tripods*, preprint `arXiv:1704.03965` (2017).

[42] F. Mémoli, T. Needham, *Distance distributions and inverse problems for metric measure spaces*, Studies in Applied Mathematics, 149(4):943–1001, 2022.

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

[44] A. Mitra, Ž. Virk, *The space of persistence diagrams on n points coarsely embeds into Hilbert space*, Proc. Amer. Math. Soc. 149 (2021), no. 6, 2693–2703.

[45] A. Mitra, Ž. Virk, *Geometric embeddings of spaces of persistence diagrams with explicit distortions*, preprint, `arXiv:2401.05298`.

[46] P. W. Nowak, G. Yu, *Large scale geometry*, European Mathematical Society, Zürich, 2012.

[47] P. Petersen, *Riemannian Geometry*, Springer, New York (1998).

[48] N. Pritchard, T. Weighill, *Coarse embeddability of Wasserstein space and the space of persistence diagrams*, `arXiv:2307.12884v2`.

[49] J. C. Robinson, *Dimensions, Embeddings, and Attractors*, Cambridge University Press, 2011.

[50] F. Schmiedl, *Computational aspects of the Gromov-Hausdorff distance and its application in non-rigid shape matching*, Discrete Comput. Geom. 57, 4 (2017), 854–880.

[51] B.J. Stolz, J. Dhesi, J.A. Bull, H.A. Harrington, H.M. Byrne, I.H.R. Yoon, *Relational Persistent Homology for Multispecies Data with Application to the Tumor Microenvironment*, Bull Math Biol 86, 128 (2024).

[52] A. A. Tuzhilin, *Lectures on Hausdorff and Gromov-Hausdorff Distance Geometry*, The course was given at Peking University, Fall 2019, `arXiv:2012.00756v1`.

[53] C. Villani, *Topics in Optimal Transportation*, Graduate Studies in Mathematics, vol. 58 (American Mathematical Society, Providence, 2003).

[54] A. Wagner, *Nonembeddability of persistence diagrams with $p > 2$ Wasserstein metric*, Proc. Amer. Math. Soc. 149 (2021), no. 6, 2673–2677.

[55] T. Weighill, T. Yamauchi, N. Zava, *Coarse infinite-dimensionality of hyperspaces of finite subsets*, Eur. J. Math. 8 (2022), no. 1, 335–355.

[56] G. Yu, *The Novikov conjecture for groups with finite asymptotic dimension*, Ann. of Math. (2) 147 (1998), no. 2, 325–355.

[57] G. Yu, *The coarse Baum-Connes conjecture for spaces which admit a uniform embedding into Hilbert space*, Invent.Math. (1) 139 (2000), 201–240.

[58] N. Zava, *The stabilty of the q-hyperconvex hull of a quasi-metric space*, preprint, `arXiv:2208.10619`.

[59] N. Zava, *Coarse and bi-Lipschitz embeddability of subspaces of the Gromov-Hausdorff space into Hilbert spaces*, Algeb. Geom. Topol., to appear, arXiv:2303.04730.

[60] S. Zhang, F. Lan, Y. Zhou, A. Barbensi, M.P.H. Stumpf, B. Wang, T. Needham, *Geometry of the Space of Partitioned Networks: A Unified Theoretical and Computational Framework*, preprint, arXiv:2409.06302v1.

Institute of Science and Technology Austria (ISTA)
Am Campus 1,
3400, Klosterneuburg
AUSTRIA
E-mail address: nicolo.zava@gmail.com