Interleaving Persistent Distances

Xiabing Ruan

June 2022

This thesis is a summary of my Master 1 study at the end of the semester. We follow the work of Ginot and Leray [5] and together with other references, give a summary of it. The work is done in Paris-Saclay University under the supervision of professor Pierre Pansu.

Contents

1 Introduction

There are many cases when we need to detect the "shape" of data points. For example, taking 100 points on the unit circle randomly, one can recongnize the shape of circle by eyes. In dynamical systems, circular attractors imply periodicity. What if the circle is embedded in $\mathbb{R}^4$ or higher dimensional Euclidean space? Or even worse, what if the data is taken from a non-Euclidean topological space? This is when topological method being used. Recovering the underlying space of discrete data points is what persistent homology do.

The theory of persistent homology is well-formulated. Given point clouds, we track the change of homology classes when changing the scale of observation.

Classic persistent homology determine the homotopy type of data by considering homology group. In this thesis, we take additional structures of the homology groups into consideration, for example the cup product, the Massey products and other higher structures. These structures make the homology group a differential graded ring, an $\mathcal{A}_\infty$-algebra and so on. In algebraic topology, these structures of homology enable us to distinguish between more spaces that cannot be distinguished when we only consider the group structure. It is the same case for persistent homology.

Chapter 2 is a brief review of persistent homology. Chapter 3 gives a kind of distance between persistent homology structures. Chapter 4 introduces the additional structures on the homology groups and access the stability of the model based on the distance defined earlier. Chapter 5 states and proves the stability theorem.

2 A review of Persistent Homology

If we take 100 arbitrary points from a circle, we can recognize the shape of the circle by our eyes, as in figure 1. If the points are taken from a torus in $\mathbb{R}^3$, we might still be able to recover the shape of the torus in our mind. What if the points are taken from a Klein bottle or a projective plane embedded in $\mathbb{R}^4$, or S^n that lies in $\mathbb{R}^{n+1}$? In many cases we have similar problems, having to recover the original space (for example, an attractor of a dynamical system) from which a discrete point cloud is taken out. One may think of calculating the homology to find its homotopy type, but the topology on a finite set is discrete, giving almost trivial homology. A strategy is to connect close points, fill in the blank spaces near each point to get a non-discrete space and then calculate the homology of that space. This might work, but the new question is the definition of "close". If the points come from a large circle, we may have to connect points whenever their distance is less than 100. But if they come from a small torus with diameter 1, the parameter 100 will connect every two pints and simply yield a contractible space. Similarly, if we let the threshold of closeness to be 1 for the large circle, we would get nothing new but the discrete set again. Different parameters change the homotopy type and homology groups, while it turns out that there is no systematic method to choose the parameter. The solution of this problem is to try all parameters from small to large and to calculate the homology groups for every parameter. After that, compare all the homology groups and pick the homology classes that persist longer than others, which are more possible to be the homology class of the original space, and regard them as the homology classes for the real space. This is what persistent homology do. In this section, we will talk about persistent homology in detail. The main reference of this chapter is [3].

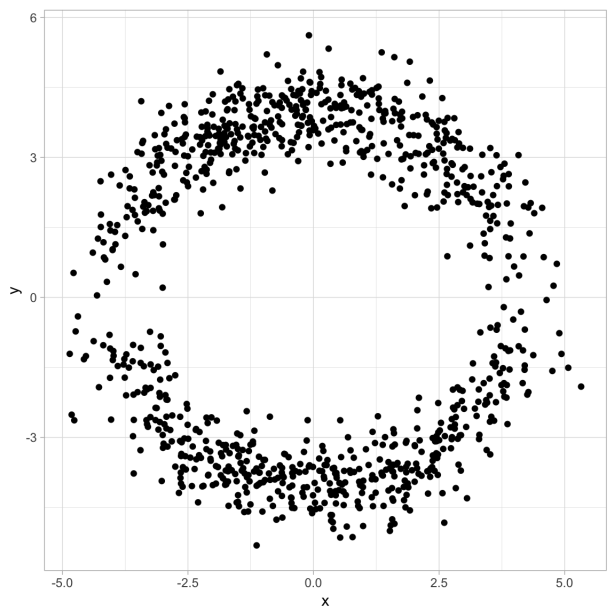


Figure 1: A statistical circle

2.1 Complex Constructions

The method to get a non-discrete space from a point cloud is to construct a complex by joining certain points together. Here we introduce two kind of complexes: Vietoris-Rips complex and Čech complex.

Definition 1 (Čech Complex)**.** Let $X = \{x_1, \ldots, x_n\}$ be a point cloud in a metric space (M, d). The **Čech complex at scale** α is the set $\check{C}_\alpha(X) := \{\sigma \subseteq X \mid \bigcap_{x\in\sigma} B(x,\alpha) \neq \varnothing\}$, where $B(x, \alpha)$ is the ball centered at x with radius α.

$\check{C}_\alpha(X)$ can be though of as a Δ-complex as follows: let $|\sigma|$ be the convex hull spanned by all points of $\sigma = \{x_{k_1}, \ldots, x_{k_m}\}$ with $k_1 < \cdots < k_m$, that is, $|\sigma| := \{\sum_{i=1}^{m} a_i x_{k_i} \mid a_i \geq 0, \sum a_i = 1\}$. Then identify σ with the only linear map $\Delta_m \to |\sigma|$ that preserves the order of the vertices.

$|\check{C}_\alpha(X)| := \bigcup_{\sigma\in\check{C}} |\sigma| \subseteq \mathbb{R}^d$ is called the **geometric realization** of $\check{C}_\alpha(X)$, whose Δ-complex structure is given by maps in $\check{C}_\alpha(X)$. In this case, we call $\check{C}_\alpha(X)$ an **abstract Δ-complex**.

When no ambiguity is caused, we identify an abstract Δ-complex and its geometric realization.

Remark. Once regarding $\check{C}_\alpha(X)$ as a Δ-complex, we can talk about its n-cycles, n-boundary chains and homology groups. It is clear that $\check{C}_\alpha(X) \subseteq \check{C}_{\alpha'}(X)$ when

$\alpha < \alpha'$. So there is a natural inclusion $C_n^\Delta(\check{C}_\alpha(X)) \hookrightarrow C_n^\Delta(\check{C}_{\alpha'}(X))$ for every $\alpha < \alpha'$ and $n \geq 0$.

Definition 2 (Vietoris-Rips Complex)**.** Let X be the same as in 1. The **Vietoris-Rips complex at scale** α (or simply Rips complex) is the set $\mathcal{V}_\alpha(X) := \{\sigma \subseteq X \mid B(x_i, \alpha) \cap B(x_j, \alpha) \neq, \forall x_i, x_j \in \sigma\}$.

$\mathcal{V}_\alpha(X)$ can be thought of as a Δ-complex in the same way as $\check{C}_\alpha(X)$. The homology groups can still be computed and the natural inclusion for every $\alpha < \alpha'$ still exists.

Remark. The Rips Complex is a **flag complex**, or a **full complex**: That is, it is the maximal element in all Δ-complexes with the given 1-skeleton. This is easy to check by the definition. Therefore, the 0-skeleton (vertices) and 1-skeleton completely determine a Rips complex, making the Rips complex less expensive than other complexes in computation.

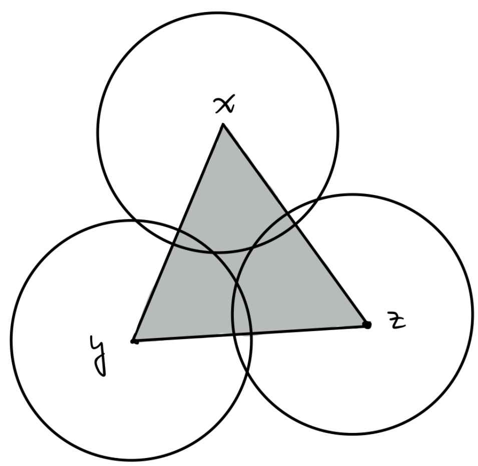


Figure 2: Complex construction for the point cloud $X = \{x, y, z\}$. Each circle has radius 1. Then $\check{C}_1(X) = \{\{x\},\{y\},\{z\},\{xy\},\{xz\},\{yz\}\}$ and $\mathcal{V}_1(X) = \{\{x\},\{y\},\{z\},\{xy\},\{xz\},\{yz\},\{xyz\}\}$. The shaded triangle is not contained in $\check{C}_1(X)$ but is in $\mathcal{V}_1(X)$.

For both Čech and Rips complex, if the parameter α is small enough, $\check{C}_\alpha(X)$ and $\mathcal{V}_\alpha(X)$ contain only discrete points; if α is larger than the diameter of X, then both $\check{C}_\alpha(X)$ and $\mathcal{V}_\alpha(X)$ contain the full complex spanned by points as vertices in X (that is, all subsets of X are simplices in $\check{C}_\alpha(X)$ and $\mathcal{V}_\alpha(X)$. For certain medium α, $\check{C}_\alpha(X)$ and $\mathcal{V}_\alpha(X)$ may very similar to the the space where X is taken from, topologically.

It is clear that for the same cale α, $\check{C}_\alpha(X) \subseteq \mathcal{V}_\alpha(X)$. Fig 2 illustrates a situation when $\check{C}_\alpha(X)$ contains strictly less simplices than $\mathcal{V}_\alpha(X)$. Conversely, the

Čech complex also contains the Rips complex with smaller radius, as illustrated in proposition 1.

Proposition 1. *For any point cloud X and $\alpha > 0$, $\mathcal{V}_{\alpha/2}(X) \subseteq \check{C}_\alpha(X) \subseteq \mathcal{V}_\alpha(X)$.*

Proof. The second inclusion is obvious. For the first inclusion, suppose $\sigma = \{x_1, \dots, x_m\} \in \mathcal{V}_{\alpha/2}(X)$, then $B(x_i, \alpha) \cap B(x_j, \alpha) \neq, \forall x_i, x_j \in \sigma$. That is, $d(x_i, x_j) < \alpha$, $\forall x_i, x_j \in \sigma$. Then the diameter of the set σ is less than α, so we can find a ball $B(y, \alpha)$ with radius α, centered at some $y \in \mathbb{R}^d$, such that $\sigma \subseteq B(y, \alpha)$. Then y is in the intersection of all $B(x_i, \alpha)$, $x_i \in \sigma$. Thus $\sigma \in \mathcal{V}_\alpha(X)$. □

This proposition allows us to compare the homology computed using the two kind of complexes.

2.2 Persistence

In this section we will deal with the varying scale, and extract important homology classes from each $H_n^\Delta(\check{C}_\alpha(X))$ and $H_n^\Delta(\mathcal{V}_\alpha(X))$.

Definition 3. Let K be an abstract Δ-complex with geometric realization $|K|$. A **filtration** of K is a sequence of subcomplexes $= K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n = K$.

A function $f : K \to \mathbb{R}$ is called **monotonic** if $f(\sigma) \leq f(\tau)$ whenever σ is a face of τ.

Example 1. (1) A filtration can be generated using a monotonic function. Let $a_1 < \cdots < a_n$ be the distinct function values of f and $K_i := f^{-1}(-\infty, a_i]$. By monotonicity of f, each K_i is a subcomplex of K and $K_1 \subseteq K_1 \subseteq \cdots \subseteq K_n = K$.

In particular, one may have $K_i - K_{i-1} = \{\sigma_i\}$ for some p-simplex σ_i. In this case, the filtration illustrates the construction of K by adding one simplex at a time. See fig 3.

(2) If we take values $0 < \alpha_1 < \cdots < \alpha_n$, then the Vietoris-Rips complex or Čech complex give a filtration $\mathcal{V}_{\alpha_1}(X) \subseteq \cdots \subseteq \mathcal{V}_{\alpha_n}$ or $\check{C}_{\alpha_1}(X) \subseteq \cdots \subseteq \check{C}_{\alpha_n}(X)$.

Let $= K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n = K$ be a filtration. For every $i \leq j$, we have $K_i \hookrightarrow K_j$ the inclusion. This induces a map between homology groups, which we denoted by $f_p^{i,j} : H_p(K_i) \to H_p(K_j)$ for all dimension p. So we obtain a sequence of homology groups connected by homomorphisms:

$$\cdots \to H_p(K_{i-1}) \to H_p(K_i) \to H_p(K_{i+1}) \to \cdots$$

where $f_p^{j,k} \circ f_p^{i,j} = f_p^{i,k}$ for every $i \leq j \leq k$. We can actually define this structure more generally, although this formal definition is not so important in this article:

Definition 4. A **persistent vector space** is a set of vector spaces $\{V_\alpha\}_{\alpha \in P}$ indexed by a total ordered set P together with linear maps $f_{\alpha,\alpha'} : V_\alpha \to V_{\alpha'}$ for every $\alpha < \alpha'$ such that $f_{\alpha,\alpha} = \mathrm{id}_{V_\alpha}$, $\forall \alpha$ and $f_{\alpha',\alpha''} \circ f_{\alpha,\alpha'} = f_{\alpha,\alpha''}$ for every $\alpha \leq \alpha' \leq \alpha''$.

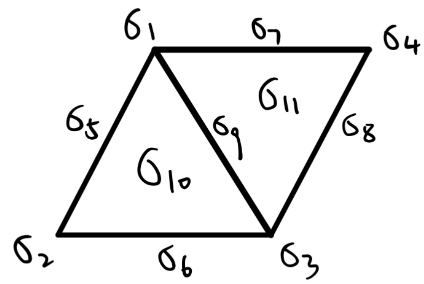


Figure 3: A filtration with $K_i = \{\sigma_1, \dots, \sigma_i\}$. Here $\sigma_1, \dots, \sigma_4$ are vertices, $\sigma_5, \dots, \sigma_9$ are edges and $\sigma_1 0, \sigma_1 1$ are faces.

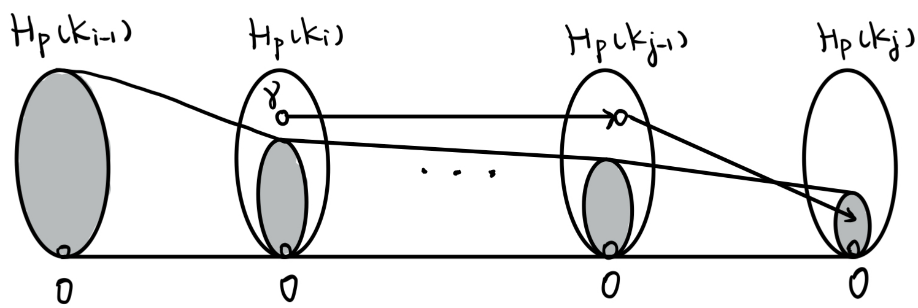


Figure 4: A homology class γ born at K_i and dies entering K_j.

Easy to see that a filtration induces a persistent vector space indexed by a finite set $\{0, 1, \dots, n\}$.

Definition 5. The p-th persistent homology group of the filtration $\emptyset = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n = K$ (with respect to i, j is $H_p^{i,j} := \mathrm{im}(f_p^{i,j}) \subseteq H_p(K_j)$. The corresponding p-th persistence betti number is $\beta_p^{i,j} := \mathrm{rank}(H_p^{i,j})$.

We say a homology class $\gamma \in H_p(K_i)$ **born at** K_i if $\gamma \notin H_p^{i-1,i}$; say γ **dies entering** K_j if $f_p^{i,j-1}(\gamma) \notin H_p^{i-1,j-1}$ but $f_p^{i,j}(\gamma) \in H_p^{i-1,j}$. Define the **persistence** of γ to be $\mathrm{pers}(\gamma) := j - i$. If γ is born at K_i and never dies, set $\mathrm{pers}(\gamma) = \infty$. See figure 4

From K_i to K_j, new cycles may be born (for example, when the simplex σ_8 is added in figure 3, creating new 1-cycle). Existing cycles may vanish, or become homologous to older cycles (for example, σ_9, σ_{11} kill 1-cycles in figure 3). The persistent homology group $H_p^{i,j}$ contains the homology classes in K_i that are still alive independently in K_j. The betti number $\beta_p^{i,j}$ indicates the number of such classes. The persistent of a homology class γ illustrates how long does this class exist. We should of course take the classes who persistent longer into consideration. To visualize the persistence, we use the following diagram:

Definition 6. The **p-th persistent diagram** of the filtration $\emptyset = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n = K$ is the multiset of points in the extended real plane $\bar{\mathbb{R}}^2 := (\mathbb{R} \cup \{\pm\infty\})^2$ such that (i, j) appears $\mu_p^{i,j}$ times, where $\mu_p^{i,j}$ is the number of p-dimensional homology classes born at K_i and dying entering K_j.

If the filtration is given by a monotonic function $f : K \to \mathbb{R}$, then we denote its persistent diagram by $\mathrm{Dgm}_p(f)$.

Example 2. In figure 3, additions of $\sigma_1, \sigma_2, \sigma_3$ and σ_4 create new 0-cycles, that is, new connected components. σ_5, σ_6 and σ_8 kill the components created by σ_2, σ_3 and σ_4 respectively, making them homologous to the 0-class $[\sigma_1]$. σ_7 creates a new 1-cycle $(\sigma_5\sigma_6\sigma_8\sigma_7)$. σ_9 further creates another 1-cycle, making the second hole. σ_{10} kills the class generated by σ_9, making it homologous to the 1-cycle from σ_8. Finally, σ_{11} kills the remaining 1-cycle from σ_8. The persistent diagram of this filtration is shown in fig

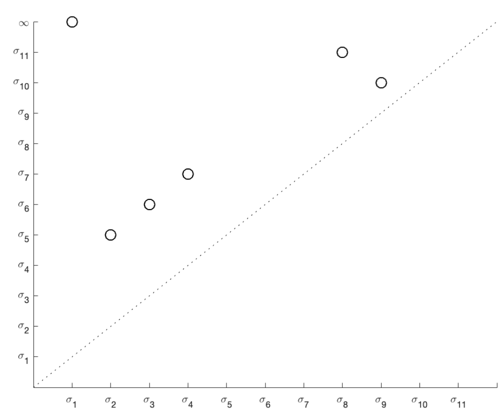


Figure 5: The persistence diagram in example 2.

The next lemma tells us that the persistent diagram contains all the information about the persistence homology groups, and vise versa:

Theorem 1 (Fundamental Lemma of Persistent Homology)**.** *Let* $= K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n = K$ *be a filtration. Then for every* $p \geq 0$*,*

(1) $\mu_p^{i,j} = (\beta_p^{i,j-1} - \beta_p^{i,j}) - (\beta_p^{i-1,j-1} - \beta_p^{i-1,j})$

(2) $\beta_p^{k,l} = \sum_{i \leq k} \sum_{j > l} \mu_p^{i,j}$.

Proof. (1) Indeed, the difference in the first bracket is the number of independent classes born at or before K_i and dying entering K_j; the difference in the second bracket is the number of independent classes born at or before K_{i-1} and dying entering K_j. By definition, their difference is $\mu_p^{i,j}$, which is the number of independent classes that are born at K_i and dies entering K_j.

(2) $\beta_p^{k,l}$ is the number of independent classes in K_i that are still alive in K_j, which is the number of independent classes that are born before K_i and dies after K_j. So $\beta_p^{k,l}$ is the number of points (counted with multiplicity) in the left upper quadrant with corner point (k, l). This is the summation in the right hand side. □

To calculate the persistent diagram explicitly, we use matrix reduction. The detailed technique of calculation is omitted here. Interested readers may refer to

There is an alternative representation of the persistent diagram, called the **barcode**. Specifically, the barcode *bcd* of a persistence diagram D is a multiset of intervals $\{[a_i, a_j] \mid (a_i, a_j) \in D\}$. The two things contain equivalent information, namely the birth and death of the homology classes. However, the barcode is more intuitive, since the length of the intervals indicate the persistence of homology classes. We can visualize the persistence homology by drawing the barcodes. We will see examples of barcodes in the next chapters.

Therefore, given a filtration, we can systematically reduce its boundary matrix and get the persistent diagram or the barcode. We are done with the technical details of persistence homology. In the next chapters, we will see how it combine with other techniques and apply to real-world problems.

3 Distances between the results

As discussed in the previous chapter, we can use the method of persistent homology to distinguish different point cloud data. A natural question is how to distinguish the results. Will the resulting persistent vector spaces be different if we change the structural complex from Čech complex to Rips complex? What if we change the point cloud? If we disturb the point cloud data slightly, will the resulting persistent object vary a lot, and if so, how to measure it? This gives us the urge to define a distance on persistent objects. We will first define persistent objects using the language of category, then define the interleaving distance on the abstract persistent objects. The relation between the interleaving distance and the Gromov-Hausdorff distance on point clouds and the stability are discussed in chapter 4.

Definition 7 ((Co)persistent Object)**.** Let $\mathbf{C}$ be a category and $(\mathbb{R}, \geq)$ be the category of real numbers considered as a totally ordered set. A **persistent object** in $\mathbf{C}$ is a functor $F : \mathbb{R} \to \mathbf{C}$. Denote F_r for $F(r)$. For every $r \leq s$, there is a natural morphism $F_r \to F_s$, the image of the structural morphism $r \leq s$.

A copersistent object is a contravariant functor $F : \mathbb{R}^{op} \to \mathcal{C}$.

Example 3. (1) Definition 4 is equivalent to definition 7 when $\mathbf{C}$ is the category of vector spaces.

(2) Given a point cloud data X, the Čech complex $\check{C}_*(X) : \mathbb{R} \to \Delta\mathrm{Cpx}$ and Rips complex $\mathcal{V}_*(X) : \mathbb{R} \to \Delta\mathrm{Cpx}$ are persistent Δ-complexes.

Definition 8 (Shifting)**.** Let $F : \mathbb{R} \to \mathbf{C}$ be a persistent object in the category $\mathbf{C}$ and $\epsilon \in \mathbb{R}$ be a real number. The ϵ**-shifting** $F[\epsilon]$ **of** F is the persistent object $F[\epsilon] : \mathbb{R} \to \mathbf{C}$ defined by $F[\epsilon]_r := F_{r+\epsilon}$ for every $r \in \mathbb{R}$.

If $F : \mathbb{R}^{op} \to \mathbf{C}$ is a copersistent object, the ϵ**-shifting** $F[\epsilon]$ is defined similarly but with $F[\epsilon]^r := F^{r-\epsilon}$.

Remark. There is a canonical natural transformation $\iota_\epsilon : F \to F[\epsilon]$ with $(\eta_\epsilon)_r : F_r \to F[\epsilon]_r$ being the structural morphism $F_r \to F_{r+\epsilon}$ for each $r \in \mathbb{R}$.

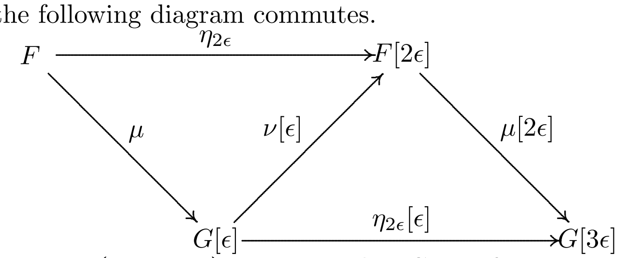
Definition 9 (ϵ-morphism)**.** Let $F, G : \mathbb{R} \to \mathbf{C}$ be two (co)persistent objects in the category $\mathbf{C}$. An ϵ**-morphism** from F to G is a natural transformation $F \to G[\epsilon]$.

Remark. If $\epsilon = 0$, we recover the notion of morphism between functors F, G, that is a natural transformation $F \to G$.

Definition 10 (The Interleaving Distance)**.** Let F, G be two persistent (co)objects. F, G are called ϵ**-interleaved** if there exists two ϵ-morphisms

$$\mu : F \to G[\epsilon], \nu : G \to F[\epsilon]$$

such that the following diagram commutes.



The **interleaving (pseudo-)distance of** F, G is defined by

$$d_{\mathbf{C}}(F, G) := \inf\{\epsilon \geq 0 \mid F \text{ and } G \text{ are } \epsilon\text{-interleaved}\}$$

Remark. If we have $F, G : \mathbb{R} \to C$ and a functor $H : \mathbf{C} \to \mathbf{D}$, then we can define a distance given by $d_{\mathbf{D}}(HF, HG)$.

Intuitively, the interleaving distance measures the smallest shift of two functors so that they can "react" with each other. Applying a third functor H as in the above remark usually makes the distance smaller, as the following lemma say.

Lemma 1. *([5], lemma 8) Let $F, G : \mathbb{R} \to \mathbf{C}$ be two persistent objects in $\mathbf{C}$ and $H : \mathbf{C} \to \mathbf{D}$ be a functor. Then $d_{\mathbf{D}}(HF, HG) \leq d_{\mathbf{C}}(F, G)$.*

Proof. Suppose there is an ϵ-interleaving between F and G, then there is a diagram as in definition 10. A functor $H : \mathbf{C} \to \mathbf{D}$ can apply on the whole diagram and translate it to an ϵ-interleaving from HF to HG in the category $\mathbf{D}$. By definition $d_{\mathbf{D}}(HF, HG) \leq d_{\mathbf{C}}(F, G)$. □

The following version of stability is easy to prove under this definition of distance.

Theorem 2 (Morse-type Stability)**.** *([5], theorem 9) Let X be a topological space and $f, g : X \to \mathbb{R}$ be two continuous maps. Then $X^f : \mathbb{R} \to \mathbf{Top}, r \mapsto f^{-1}(-\infty, r]$ is a persistent topological space and X^g is defined similarly. We have*

$$d_{\mathbf{Top}}(X^f, X^g) \leq \|f - g\|_\infty$$

where $\|f - g\|_\infty := \sup_{x \in X} |f(x) - g(x)|$.

Proof. If $\|f - g\|_\infty = \infty$ the conclusion is trivial. Otherwise, $\forall \epsilon \geq \|f - g\|_\infty$, $r \in \mathbb{R}$, $x \in X^f_r$, we have $f(x) \leq r \implies g(x) \leq f(x) + \epsilon \leq r + \epsilon \implies x \in X^g_{r+\epsilon}$. Hence $X^f_r \subseteq X^g_{r+\epsilon}$. Similarly, $X^f_r \subseteq X^g_{r+\epsilon} \subseteq X^f_{r+2\epsilon}$, $X^g_r \subseteq X^f_{r+2\epsilon} \subseteq X^g_{r+3\epsilon}$. The inclusion maps give an ϵ-interleaving from X^f to X^g, thus $d_{\mathbf{Top}}(X^f, X^g) \leq \epsilon$. Since this is true for any $\epsilon \geq \|f - g\|_\infty$, we obtain the inequality. □

By theorem 2, the interleaving distance of the persistent sublevel sets X^f and X^g is controlled by the infinity norm of f and g. In practice, X is usually a smooth manifold and f, g are Morse functions on it. This is why it's called Morse-type stability.

4 Additional Structures on Cohomology

In this chapter we introduce the additional structures such as the cup product and the Massey product on homology groups. We will recall the definitions and basic properties of the structures, and then discuss the corresponding interleaving distance. We will give several examples, including concrete calculations.

4.1 The Singular (Co)chain and Graded Vector Space

Definition 11 (Singular (co)chain functor)**.** Let k be a field. The functor $C_* : \mathbf{Top} \to \mathbf{Ch}_k$, mapping a topological space to its singular chain complex, is called the **singular chain functor**.

$C^* : \mathbf{Top}^o p \to \mathbf{Ch}_k$ given by $C^* = \mathrm{Hom}(-, k) \circ C_*$ is called the **singular cochain functor**.

Compositing with the (co)homology functor $H_*(H^*) : \mathbf{Ch}_k \to \mathbf{grVect}$, the resulting homology has the structure of graded k-vector spaces, where the grading is the natural dimension of homology.

Definition 12. Let $X, Y : \mathbb{R} \to \mathbf{Top}$ be two persistent topological spaces. Define the **gr-Vect interleaving distance**, or **classical interleaving distance** as $d_{\mathbf{grVect}}(X, Y) := d_{\mathbf{grVect}}(H^*(X), H^*(Y))$.

We shall see that the graded vector space is the coarsest structure on cohomology, therefore, the gr-Vect interleaving distance is the shortest interleaving distance among others.

4.2 Cup Product and Dg-Algebra

On cohomology there is a binary product called the cup product. We give an explicit definition of it on the level of cochains, and then on cohomology by lifting.

Definition 13 (Cup product on cochains)**.** Let X be a topological space and $C^*(X)$ be the singular cochain of X. There is a product $-\cup- : C^i(X) \times C^j(X) \to C^{i+j}(X)$ is given by

$$(f \cup g)(\sigma) := f(\sigma|_{[v_0,\dots,v_i]})g(\sigma|_{[v_i,\dots,v_{i+j}]})$$

where $\sigma \in C_{i+j}(X)$ is any singular simplex on X, i.e., a map $\sigma : \Delta^{i+j} \to X$ from the standard $(i+j)$-simplex to X.

For cohomology classes $[f] \in H^i(X)$, $[g] \in H^j(X)$, one can lift them to the level of cochain, do cup product, and then pass back to the cohomology. It turns out that this process is well-defined.

Proposition 2. *(1) For $f \in C^i(X), g \in C^j(X)$, we have $\delta(f \cup g) = \delta f \cup g + (-1)^i f \cup \delta g$ where δ is the coboundary map. This is the graded Leibniz rule.*

(2) The cup product $-\cup- : C^*(X) \times C^*(X) \to C^*(X)$ maps a couple of cocycle to a cocycle, and a cocycle and a coboundary in either order to a coboundary. Thus there is an induced cup product $-\cup- : H^*(X) \times H^*(X) \to H^*(X)$ on cohomology.

(3) The cup product is associative, (anti-)commutative on the level of cohomology.

These properties make the cochain $C^*(X)$ a *differential graded assotiative algebra*, as defined below.

Definition 14 (dg-Algebra)**.** A **differential graded associative algebra** is a graded algebra A equipped with a differential map $d : A \to A$ of degree 1 such that

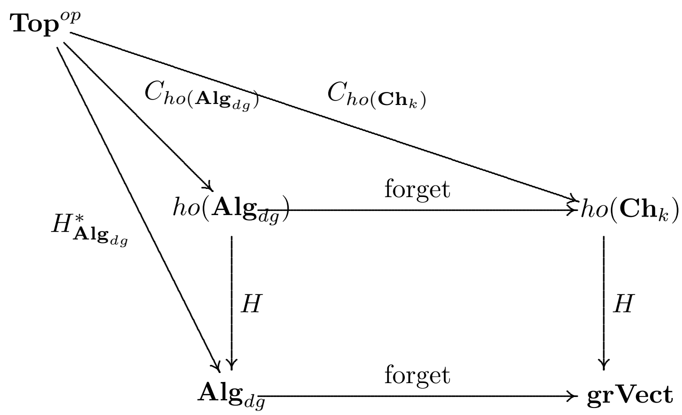
(1) $d \circ d = 0$

(2) $d(a \cdot b) = da \cdot b + (-1)^{\deg(a)} a \cdot db$

A differential graded associative algebra is called a dg-algebra for short. Denote $\mathbf{Alg}_{dg}$ the category of dg-algebras.

Therefore, we obtain a dg-algebra structure on $C^*(X)$, with the products being the cup products and the differentials being the coboundary maps. $H^*(X)$ can also be regarded as a dg-algebra with the trivial differential. Both $\mathbf{Ch}_k$ and $\mathbf{Alg}_{dg}$ have natural notion of homotopy, thus we can pass to the homotopy categories $ho(\mathbf{Ch_k})$ and $ho(\mathbf{Alg}_{dg})$ respectively. The cohomology functor H^* obviously factors through the homotopy categories.

A dg-algebra is in particular a graded vector space. There are two different cohomology functors $H^*_{\mathbf{grVect}} : \mathbf{Top}^{op} \to \mathbf{grVect}$ and $H^*_{\mathbf{Alg}_{dg}} : \mathbf{Top}^{op} \to \mathbf{Alg}_{dg}$, related by a forgetful functor $\mathbf{Alg}_{dg} \to \mathbf{grVect}$. We have, in conclusion, the following commutative diagram:



Definition 15. Let $X, Y : \mathbb{R} \to \mathbf{Top}$ be two persistent topological spaces. Define the **dg-algebra interleaving distance** of X, Y as $d_{dg}(X, Y) := d_{\mathbf{Alg}_{dg}}(H^*(X), H^*(Y))$ where the subscript of H^* is omitted.

Proposition 3. *([5], proposition 19) For any persistent topological spaces X, Y, we have the following inequalities:*

$$d_{\mathbf{grVect}}(X, Y) \leq d_{\mathbf{Alg}_{dg}}(X, Y) \leq d_{ho(\mathbf{Alg}_{dg})}(C^*(X), C^*(Y))$$
$$d_{\mathbf{grVect}}(X, Y) \leq d_{ho(\mathbf{Ch}_k)}(C^*(X), C^*(Y)) \leq d_{ho(\mathbf{Alg}_{dg})}(C^*(X), C^*(Y))$$

Proof. By lemma 1, this is the two rectangular paths of the above diagram. □

Remark. When defining the interleaving distances on the level of cochains, we passed to the homotopy category. The distance $d_{\mathbf{Alg}_{dg}}(C^*(X), C^*(Y))$ in the usual category is of course also well-defined. However, if we consider a simplicial complex X and its triangulation T, it turns out that there doesn't exist small $\epsilon > 0$ such that $C^*(X)$, $C^*(T)$ are ϵ-interleaved. Hence $d_{\mathbf{Alg}_{dg}}(C^*(X), C^(T)) \gg 0$ although X and T have the same underlying space. This is very unnatural and leads directly to the disobedience of the stability theorem in chapter 5. The problem disappears after passing to the homotopy category.

The computation of cup product is usually not easy but there are plenty of ways to this goal. In the following examples we will use classic results in algebraic topology without proof. Interested readers may refer to [6]

Example 4. ([5], example 21) In this example we will see the case for which the dg-algebra distance is strictly greater than the gr-vect distance, so the later can distinguish the two spaces while the former cannot.

Let $X := \{(x, y, 0) \in \mathbb{R}^3 \mid x^2 + (y-2)^2 = 1\} \cup \{x^2 + y^2 + z^2 = 1\} \cup \{(x, y, 0) \mid x^2 + (y+2)^2 = 1\}$ be the union of a sphere and two circles, and $Y := \{(x, y, z) \in \mathbb{R}^3 \mid (x^2 + y^2 + z^2 + 3)^2 = 16(x^2 + y^2)\}$ be the torus, see figure 6.

For a given field k, the cohomology of X is given by $H^0(X) \cong k$, $H^1(X) \cong k^2$, $H^2(X) \cong k$ (as graded vector space) and all higher dimensional cohomologies vanish. Let f, g be a basis of $H^1(X)$.

$H^*(Y)$ is isomorphic to $H^*(X)$ as graded vector spaces. However, the product structure is different. The cup product structure is trivial on $H^*(X)$ while nontrivial on $H^*(Y)$. Intuitively, the classes f, g represent the two circles on X which have no "interaction" with each other. On the contrary, f', g' correspond to the meridian and parallel of the torus, which entangle with each other and give a nontrivial cup product.

More precisely, consider the simplicial complex structure of X and Y, shown in figure 7, where the small letters represent one-dimensional simplices and the capital letters represent two-dimensional simplices. We abuse the notations of (co)simplex and (co)homology classes when no confusion is caused. For X, $H^1(X)$ has a basis $\{u, v\}$, then f, g are the dual of u, v respectively, i.e., $f(u) = 1$ and $f(x) = 0$ for all $x \neq u$, and similar for g. By definition $(f \cup g)(A) = (f \cup g)(B) = \pm u(a)v(c) = 0$. Hence $f \cup g = 0 \in H^2(X)$. For Y, $\{a, b\}$ is a basis of $H^*(Y)$ and f', g' is the dual basis. On the cochain level, choose the representative $f'(c) = g'(c) = 1$ in order to make it a cocycle. $A + B$ is a nontrivial 2-cycle. By definition $(f' \cup g')(A+B) = (f' \cup g')(A) + (f' \cup g')(B) = \pm(f'(a)g'(c) + f'(c)g'(b) = \pm 1 \neq 0$. Therefore $f' \cup g' \neq 0$.

Next, we will discretize the spaces X and Y, generate simplicial complexes and compute their interleaving distances. For $\epsilon \geq 0$, define the thickened spaces $X_\epsilon := \bigcup_{x \in X} B(x, \epsilon)$, $Y_\epsilon := \bigcup_{y \in Y} B(y, \epsilon)$. Since the radii of the circle, sphere and torus are 1, it is easy to verify that $X_\epsilon \cong X$, $Y_\epsilon \cong Y$ when $\epsilon < 1$ and $X_\epsilon \cong Y_\epsilon \cong \{*\}$ when $\epsilon \geq 1$. This process is to imitate the random errors.

Then we choose finitely many points from the thickened spaces. Let $0 < \alpha \ll \frac{1}{2}$ be fixed, there exists finitely many points $D_X^\alpha := \{x_i\}_{i \in I} \subseteq X_\alpha$ such that $\{B(x_i, \alpha)\}_{i \in I}$ covers X_α and similarly for Y, let $D_Y^\alpha := \{y_j\}_{j \in J}$. Hence we get finite samples $\{x_i\}_{i \in I}$, $\{y_j\}_{j \in J}$ of X, Y respectively. From the discrete spaces we are able to generate complexes.

Consider the Čech complex $\check{C}(D_X^\alpha)_*, \check{C}(D_Y^\alpha)_* : \mathbb{R} \to \mathbf{\Delta Cpx}$. We will compute the interleving distances of these two persistent Δ-complexes. Note that for $\alpha < r < 1 - \alpha$, i.e., when the radii of the balls are big enough to cover the original space but not too big to fill the holes, we have $\check{C}(D_X^\alpha)_r \cong X_\alpha \cong X$ and $\check{C}(D_Y^\alpha)_r \cong Y_\alpha \cong Y$. When $r > 1 + \alpha$, the balls will fill the holes no matter how the points are sampled, so $\check{C}(D_X^\alpha)_r = \check{C}(D_Y^\alpha)_r \cong \{*\}$. See figure 8.

Proposition 4. *([5], proposition 22) Keeping the notations of example 4, we have*

$$d_{\mathbf{grVect}}(\check{C}(D_X^\alpha)_*, \check{C}(D_Y^\alpha)_*) \leq 2\alpha$$

and

$$\frac{1-2\alpha}{2} \leq d_{\mathbf{Alg}_{dg}}(\check{C}(D_X^\alpha), \check{C}(D_Y^\alpha)_*) \leq \frac{1}{2}$$

In particular, when $\alpha < \frac{1}{4}$, $d_{\mathbf{grVect}}(\check{C}(D_X^\alpha, \check{C}(D_Y^\alpha)) < d_{\mathbf{Alg}_{dg}}(\check{C}(D_X^\alpha), \check{C}(D_Y^\alpha))$. As $\alpha \to 0$, $d_{\mathbf{grVect}} \to 0$ and $d_{\mathbf{Alg}_{dg}} \to \frac{1}{2}$.

Figure 6: The spaces X and Y

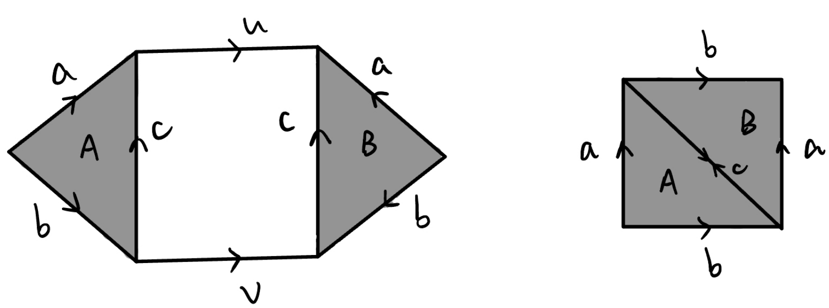


Figure 7: The Delta complex structure on the space X (left) and Y (right)

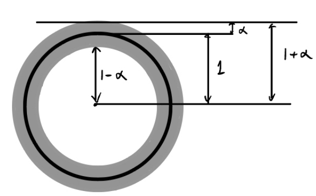


Figure 8: Thickening the spaces X and Y

Proof. As previously mentioned, $\forall r \in [\alpha, 1-\alpha) \cup (1+\alpha, \infty)$, $H^*(\check{C}(D^\alpha_X)_r) \cong H^*(\check{C}(D^\alpha_Y)_r)$ as graded vector spaces. In other words, there is only a finite interval of length 2α in which the homology graded vector spaces of $\check{C}(D^\alpha_X), \check{C}(D^\alpha_Y)$ are not necessarily isomorphic. Thus we have $\forall \epsilon > 0$, there exist a $(2\alpha+\epsilon)$-interleaving from $H^*(\check{C}(D^\alpha_X))$ to $H^*(\check{C}(D^\alpha_Y))$. The first inequality follows.

For the second inequality, first note that for every $\epsilon \in \mathbb{R}$, either $H^*(\check{C}(D^\alpha_X))_\epsilon = 0$ or $H^*(\check{C}(D^\alpha_X))_{\epsilon+1} = 0$, so the zero map always gives an $\frac{1}{2}$-interleaving from $\check{C}(D^\alpha_X)$ to $\check{C}(D^\alpha_Y)$, thus $d_{\mathbf{Alg}_{dg}}(\check{C}(D^\alpha_X), \check{C}(D^\alpha_Y)) \leq \frac{1}{2}$.

Now suppose there is $\epsilon < \frac{1-2\alpha}{2}$ such that there is an ϵ-interleaving from $H^*(\check{(}D^\alpha_X))$ to $H^*(\check{(}D^\alpha_Y))$, then there will be a commutative diagram

$$\begin{array}{ccccc} & & H^*(\check{C}(D^\alpha_X)_{\alpha+\epsilon} & & \\ & \overset{\mu}{\nearrow} & & \overset{\nu}{\searrow} & \\ H^*(\check{C}(D^\alpha_Y))_\alpha & & \xrightarrow{\quad\cong\quad} & & H^*(\check{C}(D^\alpha_Y))_{\alpha+2\epsilon} \end{array}$$

Since $\epsilon < \frac{1-2\alpha}{2}$, $\alpha+\epsilon < \alpha+2\epsilon < 1-\alpha$. By the previous discussions, $\check{C}(D^\alpha_X)_r \simeq X$, $\check{C}(D^\alpha_Y)_r \simeq Y$ whenever $\alpha < r \leq 1-\alpha$. Thus the map $H^*(\check{C}(D^\alpha_Y))_\alpha \to H^*(\check{C}(D^\alpha_Y))_{\alpha+2\epsilon}$ induced by the inclusion is an isomorphism. I.e., $\nu \circ \mu$ is an isomorphism, hence μ is injective. This gives an injective graded algebra homomorphism $H^*(Y) \to H^*(X)$. However, we know that $f' \cup g' \neq 0 \in H^*(Y)$ but $f \cup g = 0 \in H^*(X)$, so such injective morphism does not exist. This is a contradiction. So there is no such ϵ-interleaving, we have the last inequality.

□

Remark. The Čech complex in this example can be replaced by Rips or Alpha complex as one wished.

Example 4 tells us that while the gr-Vect intereleaving distance cannot distinguish the two data sets D^α_X and D^α_Y, the dg-Alg distance can. This is of course because of taking the ring structure into consideration.

Example 5. ([5], example 27) We can create similar examples on all spaces that can be distinguished using cup product but cannot with only the structure of graded vector space. For example, consider the trivial link T and the Hopf link L as shown in figure 9. Let $X := \mathbb{R}^3 - T$, $Y := \mathbb{R}^3 - S$ be the complements in $\mathbb{R}^3$. Then we have $H^*(X) \cong H^*(Y)$ as graded vector spaces. $H^2(X) \cong H^2(Y) \cong k^2$ are two-dimensional, whose generators represent the two loops around the two circles of the links of X and Y, respectively. However, the cup product of the dg-Alg $H^*(X)$ is trivial while that of $H^*(Y)$ is not. In fact, trivial cup product illustrates the possibility to untangle the links, while the nontrivial one indicates that the Hopf link cannot be untangled.

If the link includes three or more loops, the cup product may not be enough. We need higher structures to distinguish them, as discussed in the next section.

Figure 9: The trivial link T (left) and the Hopf link L (right)

5 Stability Theorem

After a method of prediction is invented (e.g., persistent homology to predict the shape of data), the stability is essential to measure the reliability of the method. Intuitively, when the original data is disturbed slightly, the results should not change a lot. The topology properties of a manifold, for example, is of course stable under local perturbation. To measure the "slight change" of data sets, we should define a distance on the data sets.

5.1 Gromov-Hausdorff distance

Definition 16. Let X, Y be two sets. (1) A **multi-valued map** from X to Y is a subset C of $X \times Y$ such that $\pi_X|_C : C \to X$ is surjective, where $\pi_X : X \times Y \to X$ is the canonical projection. Denote a multi-valued map by $C : X \rightrightarrows Y$.

(2) The image of a subset $S \subseteq X$ is defined to be $C(S) := \pi_Y(\pi_X^{-1}(S) \cap C)$.

(3) A map $f : X \to Y$ is called **subordinate to** C if $\forall x \in X$, $(x, f(x)) \in C$. Write $f : X \xrightarrow{C} Y$.

(4) Given two multi-valued maps $C : X \rightrightarrows Y$, $D : Y \rightrightarrows Z$, the composition $D \circ C : X \rightrightarrows Z$ is the subset $\{(x, z) \in X \times Z \mid \exists y \in Y$ such that $(x, y) \in C$ and $(y, z) \in D\}$.

(5) Given $C : X \rightrightarrows Y$, if $\pi_Y|_C : C \to Y$ is surjective, then C is called a **correspondence**. The **transpose** C^T of a correspondence C is the multi-valued map from Y to X such that $C^T := \{(y, x) \in Y \times X \mid (x, y) \in C\}$.

Remark. For any correspondence $C \rightrightarrows Y$, we have id_X is subordinate to $C^T \circ C$ and id_Y is subordinate to $C \circ C^T$.

Definition 17 (Gromov-Hausdorff Distance)**.** Let (X, d_X), (Y, d_Y) be metric spaces and $C : X \rightrightarrows Y$ be a correspondence. The **distortion** of C is defined to be

$$\mathrm{dist}(C) := \sup_{(x,y),(x',y') \in C} |d_X(x, x') - d_Y(y, y')|.$$

The **Gromov-Hausdorff distance** between X and Y is defined to be

$$d_{GH}(X,Y) := \frac{1}{2} \inf_{X \rightrightarrows Y} \mathrm{dist}(C)$$

.

Remark. One should check that this is actually a distance. To understand this definition, consider, for example C being a isometry, then $\mathrm{dist}(C) = 0$. If C is a correspondence distorts the metrics a lot, then the distortion will increase.

In fact, we have

$$d_{GH}(X,Y) = \inf_{(\eta,\tau)\in I} \min\{\epsilon \geq 0 \mid \eta(X) \subseteq \tau(Y)_\epsilon \text{ and } \tau(Y) \subseteq \eta(X)_\epsilon\}$$

where $I := \{(\eta : X \to Z, \tau : Y \to Z) \mid (Z, d_Z)$ is a metric space and η, τ are isometrical embeddings $\}$ and $\eta(X)_\epsilon := \bigcup_{x\in\eta(X)} B_Z(x,\epsilon)$. Intuitively, this distance describes the smallest amount of thickening of X and Y so that they can be embedded to each other in an arbitrary space.

5.2 The Theorem

Definition 18. Let X, Y be two sets and $\mathcal{S}, \mathcal{T} : \mathbb{R} \to \Delta Cpx$ be two persistent delta complexes whose vectex sets are X, Y respectively, for all $r \in \mathbb{R}$. A multivalued map $C : X \rightrightarrows Y$ is called **ϵ-simplicial** for $\mathcal{S}$ and $\mathcal{T}$ if $\forall r \in \mathbb{R}$, $\forall \sigma \in \mathcal{S}_r$ a simplex, every finite subset of $C(\sigma)$ is a simplex of $\mathcal{T}_{r+\epsilon}$.

Definition 19 (Contiguous simplicial maps)**.** Let K and L be two simplicial complexes. Two simplicial maps $f, g : K \to L$ are called **contiguous** if for any simplex $[v_0, \dots, v_n]$ of K, the points $f(v_0), \dots, f(v_n), g(v_0), \dots, g(v_n)$ span a simplex of L.

We use without proving the following lemma:

Lemma 2. *([7], proposition 10.20) Keep the notions in definition 19. If f, g are continuous, then their geometric realizations $|f|, |g| : |K| \to |L|$ are homotopic as continuous maps.*

Lemma 3. *([4], proposition 3.3) Let $\mathcal{S}, \mathcal{T} : \mathbb{R} \to \Delta Cpx$ with vertex sets X, Y respectively be two persistent Δ-complexes. Suppose there exists a ϵ-simplicial multi-valued map $C : X \to Y$ from $\mathcal{S}$ to $\mathcal{T}$. Then any two maps $f, g : X \xrightarrow{C} Y$ subordinating to C induce contiguous simplicial maps $\mathcal{S}_r \to \mathcal{T}_{r+\epsilon}$ for all $r \in \mathbb{R}$. Moreover, all the geometric realizations $|f|, |g| : |\mathcal{S}_r| \to |\mathcal{T}_{r+\epsilon}|$ are contiguous.*

Proof. First suppose that $f : X \xrightarrow{C} Y$ is a map subordinate to C. Then $\forall r \in \mathbb{R}, \sigma \in \mathcal{S}_r$, $f(\sigma)$ is a finite subset of $C(\sigma)$, hence a simplex of $\mathcal{T}_{r+\epsilon}$ since C is ϵ-simplicial from $\mathcal{S}$ to $\mathcal{T}$. Hence f, g induce simplicial maps $\mathcal{S}_r \to \mathcal{T}_{r+\epsilon}$.

To show that the induced maps are contiguous, note that for any simplex σ, $f(\sigma) \cup g(\sigma)$ is also a finite subset of $C(\sigma)$, hence spans a simplex in $\mathcal{T}_{r+\epsilon}$. By definition, f and g induce contiguous simplicial maps at every point. □

Lemma 4. *([4], lemma 4.3 and 4.4) Let (X, d_X) and (Y, d_Y) be two metric spaces. Suppose there is a correspondence $C : X \rightrightarrows Y$ with distortion ϵ. Consider the persistent Δ-complexes $\check{C}(X)$, $\mathcal{V}(X)$ and $\check{C}(Y)$, $\mathcal{V}(Y)$. Then we have the correspondence C is*

(1) ϵ-simplicial from $\mathcal{V}(X)$ to $\mathcal{V}(Y)$

(2) ϵ-simplicial from $\check{C}(X)$ to $\check{C}(Y)$.

Proof. (1) Let C be such a correspondence. Fix $r \in \mathbb{R}$, $\forall \sigma \in \mathcal{R}(X)_r$, $\tau \subseteq C(\sigma)$ finite subset, we want to show that τ is a simplex of $\mathcal{R}(Y)_{r+\epsilon}$. $\forall y, y' \in \tau$, there exists $x, x' \in \sigma$ such that $y \in C(x)$, $y' \in C(x')$. By the definition of Rips complex, we have $d_X(x, x') \leq r$. By the definition of the distortion, we have $|d_X(x, x') - d_Y(y, y')| \leq \epsilon \implies d_Y(y, y') \leq d_X(x, x') + \epsilon \leq r + \epsilon$. This is true for all $y, y' \in \tau$, hence τ is a simplex of $\mathcal{R}(Y)_{r+\epsilon}$. Thus C is ϵ-simplicial from $\mathcal{R}(X) \to \mathcal{R}(Y)$.

(2) If σ is a simplex of $\check{C}(X, d_X)_r$, let $\bar{x}$ be an r-center of σ, then $\forall x \in \sigma$, we have $d_X(x, \bar{x}) \leq r$. Then rest of the arguments are exactly similar to (1). □

Now we are able to prove the stability theorem for the interleaving distances. For the original version of stability theorems, see [1].

Theorem 3 (Stability Theorem for the Interleaving Distances)**.** *([5], theorem 66) Let (X, d_X), (Y, d_Y) be two metric spaces. There are the following inequalities:*

(1)

$$d_{ho(\mathbf{Alg}_{dg})}(C^*(\mathcal{R}(X)), C^*(\mathcal{R}(Y))) \leq 2d_{GH}(X, Y)$$

$$d_{ho(\mathbf{Alg}_{dg})}(C^*(\check{C}(X)), C^*(\check{C}(Y))) \leq 2d_{GH}(X, Y)$$

(2)

$$d_{\mathbf{Alg}_{dg}}(H^*(\mathcal{R}(X)), H^*(\mathcal{R}(Y))) \leq 2d_{GH}(X, Y)$$

$$d_{\mathbf{Alg}_{dg}}(H^*(\check{C}(X)), H^*(\check{C}(Y))) \leq 2d_{GH}(X, Y)$$

Proof. Let $d_{GH}(X, Y) := \frac{d}{2}$, then $\forall \epsilon > 0$, there is a correspondence $C : X \rightrightarrows Y$ of distortion at most $d + \epsilon$. Then by lemma 4, C is $(d + \epsilon)$-simplicial from $C^*(\mathcal{R}(X))$ to $C^*(\mathcal{R}(Y))$. Choose a map $f : X \to Y$ subordinating to C, then by lemma 3, f induces an $(d + \epsilon)$-morphism from $C^*(\mathcal{R}(X))$ to $C^*(\mathcal{R}(Y))$. By lemma 3 again, f can be passed to the homotopy category $ho(\mathbf{Alg}_{dg})$.

Similarly, $C^T : Y \rightrightarrows X$ induces a $(d + \epsilon)$-morphism $g : C^*(\mathcal{R}(Y)) \to C^*(\mathcal{R}(X))$. Recall that $\mathrm{id}_X : X \xrightarrow{C^T \circ C} X$, $\mathrm{id}_Y : X \xrightarrow{C \circ C^T} X$, and $g \circ f \xrightarrow{C^T \circ C}$, $f \circ g \xrightarrow{C \circ C^T}$. So the maps $g \circ f : C^*(\mathcal{R}(X)) \to C^*(\mathcal{R}(Y))$ is a $2(d+\epsilon)$-morphism which is exactly the structural morphism of $C^*(\mathcal{R}(X))$. Similarly for $f \circ g$. By definition f, g define a $(d+\epsilon)$-interleaving from $C^*(\mathcal{R}(X))$ to $C^*(\mathcal{R}(Y))$. Since ϵ is arbitrary, we have $d_{ho(\mathbf{Alg}_{dg})}(C^*(\check{C}(X)), C^*(\check{C}(Y))) \leq d = 2d_{GH}(X, Y)$. □

Therefore, the dg-Alg and $\mathcal{A}_\infty$ interleaving distances are controlled by the Gromov-Hausdorff distance of the data sets. The interleaving distance will not be large for similar spaces, and will not change a lot as long as the disturbance of data points is small.

References

[1] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. In *Proceedings of the twenty-first annual symposium on Computational geometry*, pages 263–271, 2005.

[2] David S. Dummit and Richard M. Foote. *Abstract algebra.* Wiley, 3rd ed edition, 2004.

[3] Herbert Edelsbrunner and John Harer. *Computational Topology - an Introduction.* American Mathematical Society, 2010.

[4] Chazal Frederic, Vin de Silva, and Oudot Steve. Persistence stability for geometric complexes, 2012. 10.48550/ArXiv.1207.3885.

[5] Ginot Grégory and Leray Johan. Multiplicative persistent distances, 2019. 10.48550/ArXiv.1905.12307.

[6] Allen Hatcher. *Algebraic topology.* Cambridge Univ. Press, Cambridge, 2000.

[7] J. McCleary. *A First Course in Topology: Continuity and Dimension.* Student mathematical library. American Mathematical Soc.

[8] J.R. Munkres. *Topology.* Featured Titles for Topology. Prentice Hall, Incorporated, 2000.