



**HAL**  
open science

## Multiplicative persistent distances

Grégory Ginot, Johan Leray

► **To cite this version:**

| Grégory Ginot, Johan Leray. Multiplicative persistent distances. 2019. hal-02141757v1

**HAL Id: hal-02141757**

**<https://hal.science/hal-02141757v1>**

Preprint submitted on 28 May 2019 (v1), last revised 19 Jul 2019 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# MULTIPLICATIVE PERSISTENT DISTANCES

GRÉGORIE GINOT AND JOHAN LERAY

*"Bats-toi, signe et persiste" – France Gall*

ABSTRACT. We define and study several new interleaving distances for persistent cohomology which take into account the algebraic structures of the cohomology of a space, for instance the cup product or the action of the Steenrod algebra. In particular, we prove that there exists a persistent  $\mathcal{A}_\infty$ -structure associated to data sets and we define the associated distance. We prove the stability of these new distances for Čech or Vietoris Rips complexes with respect to the Gromov-Hausdorff distance, and we compare these new distances with each other and the classical one, building some examples which prove that they are not equal in general and refine effectively the classical bottleneck distance.

## INTRODUCTION

Persistent homology arised as a successful attempt to make invariants of algebraic topology computable in practice in various contexts. A prominent example being to study data sets and their topology, which have become increasingly important in many area of sciences. In particular, to be able to discriminate and compare large data sets, it is natural to associate invariants to each of them in order to be able to say if they are similar and describe similar phenomenon or not. The latter operation is obtained by considering a metric on the invariants associated to the data which, classically, is the interleaving or bottleneck distance on the persistent homology of the data. The interested reader may consult [Oud15a, EH08] for an extended discussion of the theory and of its many applications. Our goal is to study and compare several refinements of those distances obtained by considering more structure, inspired by homotopical algebra, on the persistent cohomology which discriminate more data sets.

**Topological data analysis.** Associating algebraic invariants to *shapes* is a main apparatus of algebraic topology. *Topological data analysis* (TDA for short) associates and studies the topology of data sets through the help of algebraic topology invariants characterizing as finely as possible the data. Roughly, a main idea of TDA is to associate to, a potentially large, set  $X$  of  $N$  points a family of spaces  $X_\varepsilon$  given by the union of balls centered on each point with radius given by the parameter  $\varepsilon$ . Now we can consider the invariants of each space but, even better, we can study the set  $\{X_\varepsilon\}$  as a continuous family of spaces, called a *persistent space*, and considering the evolution of these invariants when  $\varepsilon$  grows. The more accessible topological invariant is the homology of these spaces also known as *persistent homology*.

**Persistent homology.** The homology of a persistent space gives us a parametrized family of graded vector space. To such object, one associates a *barcode*, which represents the evolution of the dimension of each homology group when the parameter varies. For instance, a  $i^{\text{th}}$ -homology class can be born at the time  $\varepsilon_1$  in the  $i$ -th group and dies at time  $\varepsilon_2$ . This class is associated to a bar of length  $\varepsilon_2 - \varepsilon_1$  and the collection of those is the *barcode* of the persistent homology groups. This barcode defines a invariant of the persistent space  $\{X_\varepsilon\}$ .

---

*Date:* May 28, 2019.

The authors acknowledge support from the project ANR-16-CE40-0003 ChroK. The second author is financed by a postdoctoral allocation of the DIM Math Innov - Région Île de France.

To compare two barcodes  $B_X$  and  $B_Y$  associated to datasets  $X$  and  $Y$ , Cohen *et al.* defined in [CSEH07] the *bottleneck distance* which is a (pseudo-)distance. A more intrinsic notion of distance, directly defined on the persistent (co)homology, is given by the *interleaving distance* introduced by Chazal *et al.* in [CDSO14].

**Applications.** These techniques of TDA have been applied in many areas: for example in reconstruction of shapes (see [CO08, CZCG05]), gene expression analysis (see for example [JCR<sup>+</sup>17]), or in neurosciences (see [KDS<sup>+</sup>18]). It should be noted that the bottleneck distance is only slightly or not sensitive to data noise.

**Content of the paper.** Algebraic topologists have constructed several other invariants finer than homology of the space: for instance the cohomology has a natural graded *algebra* structure induced by the cup product. This structure is itself a shadow of the differential graded algebra structure carried by the *cochains* which is a better invariant, see Example 29. These refined invariants thus encode in a much more effective way the homotopy type of a space compared to mere homology and in fact, the homotopy of a (nilpotent finite type) space is completely encoded by this dg-algebra structure together with higher homotopies for its commutativity, i.e., its  $\mathcal{E}_\infty$ -algebra structure [Man06].

In this paper, we give the theoretical framework to construct and compare new interleaving distances (on data sets) which take into account these extra algebraic structures in the persistence setting, in a systematic way. They refine the classical interleaving distance  $d_{\text{gr-Vect}}$  but are of different computational difficulties. In fact, we have the following commutative diagrams of functors:

- for  $p$  a prime, we have

$$\begin{array}{ccccc}
 & & \text{Top}^{\text{op}} & & \\
 & \swarrow & \downarrow & \searrow & \\
 \mathcal{H}^*(-, \mathbb{F}_p) & & & & \\
 \downarrow & & & & \\
 \mathcal{A}_p\text{-Alg} & \xrightarrow{\text{forget}} & \text{Alg}_{\mathcal{A}_p} & \xrightarrow{\text{forget}} & \text{gr-Vect}
 \end{array}$$

where  $\mathcal{A}_p\text{-Alg}$  is the category of algebras over the Steenrod algebra  $\mathcal{A}_p$  see Section 4.2;

- for  $k$ , a field of characteristic zero, we have

$$\begin{array}{ccccc}
 & & \text{Top}^{\text{op}} & & \\
 & \swarrow & \downarrow & \searrow & \\
 \mathcal{C}^*(-, k) & & & & \\
 \downarrow & & & & \\
 \text{Alg}_{\mathcal{A}_p} & \xrightarrow{\mathcal{H}^*(-)} & \text{Alg}_{\mathcal{A}_p} & \xrightarrow{\text{forget}} & \text{gr-Vect}
 \end{array}$$

*Each of the categories in the lines of the diagram gives rise to an interleaving distance.*

Several of the more refined interleaving distances introduced above are computed in the homotopy categories of cochains with some extra structure and not a cohomological level for which we have the barcode decomposition. The homotopy category of such cochain algebra is hard to study and cochain algebras are too big to be easily used on a computer at the moment.

We bypass this problem in characteristic zero by using the *homotopy transfer theorem* for  $\mathcal{A}_\infty$ -algebras which allows to encode the cochain algebra on the cohomology groups *without losing information*. Indeed, an  $\mathcal{A}_\infty$ -algebra is an associative algebra up to homotopy (see Definition 22) and, for all topological spaces  $X$ , the singular cochain complex of  $X$  is equivalent of the cohomology of  $X$  as  $\mathcal{A}_\infty$ -algebras (see Theorem 27). Further, unlike quasi-isomorphisms of dg-algebras,  $\mathcal{A}_\infty$ -quasi-isomorphisms have inverses which simplifies greatly the study of interleaving in these category. However transfer theorems are not very functorial and therefore it is unlikely that they can be applied to abstract persistence spaces in general (Remark 35).

Nevertheless, for persistent simplicial sets  $X : \mathbb{R} \rightarrow \text{sSet}$  satisfying some mild finiteness assumption (that we call *finite filtered data*), we prove the following theorem.

**Theorem A** ( $\mathcal{A}_\infty$  interleaving distance (see Theorem 37)). *There exists a persistent  $\mathcal{A}_\infty$ -structure on the persistent cohomology of finite filtered data and an interleaving distance  $d_{\mathcal{A}_\infty}$ , which refine the cohomology algebra and takes into account higher Massey products of singular cohomology.*

Some approaches to use  $\mathcal{A}_\infty$ -structures for persistence have already been considered in the literature, notably in the work of F. Belchí *et al.* (see [GM14, BM15, Bel17, BS19]) and Herscovich (cf. [Her18]). In both cases, they consider transferred structures but do not consider the full either  $\mathcal{A}_\infty$  or persistent structure and in particular do not define an associated interleaving distance.

The homology of a space is sensitive to the characteristic of its coefficient and this reflects on the additional algebraic structure of the cochains. Therefore, we define (see Section 4.2) two new distances: the first  $d_{p,\infty}$  given by the maximum between the  $\mathcal{A}_\infty$ -distance and the one defined by the structure of Steenrod algebra of the cohomology with coefficient in  $\mathbb{F}_p$  (Section 4.2) where  $p$  is a fixed prime; the second  $d_{\mathbb{P},\infty}$  is given by the supremum of  $d_{p,\infty}$  over the set  $\mathbb{P}$  of primes. Our second main contribution is the comparison of these distances.

**Theorem B** (see Section 5). *All distances defined for finite filtered data in this paper satisfy the following inequalities:*

$$\begin{array}{ccccccc}
 & & & & d_{\mathcal{A}_\infty} & \geq & d_{\mathcal{A}_S, \mathbb{Q}} & \geq & d_{\text{gr-Vect}, \mathbb{Q}} \\
 & & & \cong & & & & & \\
 d_{\mathbb{P}, \infty} & \geq & d_{p, \infty} & & & & & & \\
 & & & \cong & & & & & \\
 & & & & d_{\mathcal{A}_p} & \geq & d_{\mathcal{A}_S, \mathbb{F}_p} & \geq & d_{\text{gr-Vect}, \mathbb{F}_p}
 \end{array}$$

*which are not equalities in general.*

We believe that these refined distances are reasonable approximation of the most refined of all, that is the interleaving distance associated to the  $\mathcal{E}_\infty$ -structure of cochains, see Section 4.1. Unlike this latter one, they are defined on the underlying persistent cohomology groups and therefore they seem much more “computerizable”. Further, we prove refined stability theorems for those distances. Indeed all these distances satisfy a property of stability for Čech or Vietoris Rips complexes (see Section 1.1.2) with respect to the Gromov-Hausdorff distance.

**Theorem C** (Stability results (see Theorem 53, Theorem 65)). *Let  $X$  and  $Y$  be two finite set of points of  $\mathbb{R}^n$ . We have the following inequality:*

$$\begin{aligned}
 d_+(\mathcal{R}(X), \mathcal{R}(Y)) &\leq 2d_{\mathcal{GH}}(X, Y) ; \\
 d_+(\check{\mathcal{C}}(X), \check{\mathcal{C}}(Y)) &\leq 2d_{\mathcal{GH}}(X, Y) ,
 \end{aligned}$$

*where  $d_+$  is one of the distances of Theorem B.*

Such a theorem is very important for applications since, for set of points  $X$  and  $Y$  representing the same data up to some noise, this theorem implies that if the noise is small then the distance  $d_+(\mathcal{R}(X), \mathcal{R}(Y))$  is also small.

## CONTENTS

Introduction	1
1. Multiplicative distances	4
2. The $\mathcal{A}_\infty$ -interleaving distance	10
3. A stability theorem for multiplicative distances	17
4. Homotopy commutativity preserving distances	20

**Notations.** We introduce some notations used throughout the paper:

- $k$  is a field unless otherwise specified;
- If  $\mathcal{C}, \mathcal{D}$  are categories,  $\mathcal{D}^{\mathcal{C}}$  stands for the category of functors  $\mathcal{C} \rightarrow \mathcal{D}$  ;
- $(\mathbb{R}, \leq)$  is the poset of real numbers  $(\mathbb{R}, \leq)$ , viewed as a category. For all  $r < s$  in  $\mathbb{R}$ , the unique corresponding morphism in the category  $\mathbb{R}$  is denoted by  $(r \leq s)$  ;  $\mathbb{R}^{\text{op}}$  corresponds to the poset  $(\mathbb{R}, \geq)$  viewed as a category: we denote by  $(r \geq s)$  its morphisms. Functors  $X : \mathbb{R} \rightarrow \mathcal{C}$  and  $Y : \mathbb{R}^{\text{op}} \rightarrow \mathcal{C}$  are denoted by  $X_{\bullet}$  and  $Y^{\bullet}$  respectively;
- $\text{gr-Vect}$  stands for the category of graded vector space;  $\text{Ch}_k$  the category of  $\mathbb{Z}$ -graded chain complexes and  $\text{coCh}_k$  the category of  $\mathbb{Z}$ -graded cochain complexes. We when needed, we will denote the degree of objects in  $\text{Ch}_k$  (resp.  $\text{coCh}_k$ ) by a lower index  $C_*$  (resp. an upper index  $C^*$ ).
- $\text{Alg}_{\mathcal{O}}$  is the category of  $\mathcal{O}$ -algebras in the category of cochain complex over  $k$ , with  $\mathcal{O}$  an operad (see [LV12, Chapter 5]) (and strict morphisms);
- when we define an interleaving distance  $d_{\dagger}$  by using a functor of cohomology  $H^*(-) : \text{Top}^{\text{op}} \rightarrow \mathcal{C}$ , we denote it by  $d_{\mathcal{C}}(-_1, -_2) := d_{\mathcal{C}}(H^*(-_1), H^*(-_2))$ .
- we denote by  $\text{For}$ , "the" forgetfull functor (which occurs in several contexts).

## 1. MULTIPLICATIVE DISTANCES

**1.1. Persistent objects and interleaving.** In this section, we recall the notion of persistent objects and interleaving distances in generic categories.

### 1.1.1. Definitions.

**Definition 1** (Persistent object – Shifting). Consider a category  $\mathcal{C}$ . A *persistent object* in  $\mathcal{C}$  is a functor  $F_{\bullet} : \mathbb{R} \rightarrow \mathcal{C}$ . The image  $F_{r \leq s}$  of the morphism  $(r \leq s)$  by the functor  $F$  is called a *structural morphism of  $F$* . Let  $\varepsilon$  be an object of  $\mathbb{R}$ , we have the natural shifting operation  $(-)[\varepsilon]$  which sends a persistent object  $F_{\bullet}$  to  $F[\varepsilon]_{\bullet}$  which is defined, for all  $r$  in  $\mathbb{R}$ , by  $F[\varepsilon]_r := F_{r+\varepsilon}$ . For each  $\varepsilon$  in  $\mathbb{R}$ , we have the natural transformation  $\eta_{\varepsilon} : F \rightarrow F[\varepsilon]$ , given, for all  $r$  in  $\mathbb{R}$ , by the structural morphism  $(r \leq r + \varepsilon)$ .

**Definition 2** (Copersistent object – Shifting). A *copersistent object* in  $\mathcal{C}$  is a functor  $F^{\bullet} : \mathbb{R}^{\text{op}} \rightarrow \mathcal{C}$ . Let  $\varepsilon$  be an object of  $\mathbb{R}$ , we have the natural shifting  $(-)[\varepsilon]$  defined, for all copersistent object  $F^{\bullet}$  and all  $r$  in  $\mathbb{R}$ , by  $F[\varepsilon]^r = F^{r-\varepsilon}$ . We also have the natural transformation  $\eta_{\varepsilon} : F \rightarrow F[\varepsilon]$ , given, for all  $r$  in  $\mathbb{R}$ , by the structural morphism  $(r \geq r - \varepsilon)$ .

**Definition 3** ( $\varepsilon$ -morphism). Let  $F$  and  $G$  be two (co)persistent object in the category  $\mathcal{C}$ , and  $\varepsilon$  in  $\mathbb{R}$ . A morphism  $F \rightarrow G$  is a natural transformation and a  *$\varepsilon$ -morphism from  $F$  to  $G$*  is a natural transformation  $F \rightarrow G[\varepsilon]$ .

**Definition 4** (Interleaving pseudo-distance). Let  $F$  and  $G$  be two functor in  $\mathcal{C}^{\mathbb{R}}$  and let  $\varepsilon$  be in  $\mathbb{R}$ . The persistent objects  $F$  and  $G$  are  *$\varepsilon$ -interleaved* if there exists two  $\varepsilon$ -morphisms

$$\mu : F \longrightarrow G[\varepsilon] \quad \text{and} \quad \nu : G \longrightarrow F[\varepsilon]$$

such that, the following diagram commutes

$$\begin{array}{ccc}
 F & \xrightarrow{\eta_{2\varepsilon}} & F[2\varepsilon] \\
 \searrow \mu & & \nearrow \mu[2\varepsilon] \\
 & G[\varepsilon] & \xrightarrow{\eta_{2\varepsilon}[\varepsilon]} G[3\varepsilon] \\
 & \nearrow v[\varepsilon] & 
 \end{array}$$

We define the *interleaving (pseudo-)distance* by:

$$d_C(F, G) := \inf \{ \varepsilon \geq 0 \mid F \text{ and } G \text{ are } \varepsilon\text{-interleaved} \} .$$

**Remark 5.** The definitions of  $\varepsilon$ -interleaving and distance of interleaving between copersistent objects are completely similar.

**Remark 6.** Given an interleaving distance  $d_D$  for persistent objects in the category  $D$  and a functor  $H : C \rightarrow D$ , then we obtain a (pseudo-)distance defined by

$$d_D(H(-_1), H(-_2))$$

for persistent objects in  $C$ .

**Remark 7** (About our notations for the distances). In this paper, we introduce several distances between persistent spaces. We make the following choice of notations: when such a distance is defined by using a functor of cohomology  $H^*(-) : \text{Top}^{\text{op}} \rightarrow C$ , we decide to denote it as follows

$$d_{\dagger}(-_1, -_2) := d_C(H^*(-_1), H^*(-_2))$$

with a suitable choice of notation to replace the symbol  $\dagger$ . Furthermore, whenever we use (co)chain complexes of (co)homology functors to compute the distance, the result depends on the base ground field of coefficients. But we will usually *not* include it in the notation. However, when there might be some confusion or we want to put emphasis on the correct coefficient we will write

$$d_{\dagger, k}(-_1, -_2)$$

for the distance computed with coefficients being  $k$ .

We will often use tacitly the following easy lemma.

**Lemma 8.** *Let  $F$  and  $G$  be two (co)persistent objects in  $C$  and let  $H : C \rightarrow D$  be a functor. We have*

$$d_D(HF, HG) \leq d_C(F, G).$$

*Proof.* Let  $\varepsilon \geq 0$  such that  $d_C(F, G) \leq \varepsilon$ , then there exists an  $\varepsilon$ -interleaving between  $F$  and  $G$ . As  $H$  is a functor, it preserves the diagram of  $\varepsilon$ -interleaving, so  $d_D(HF, HG) \leq \varepsilon$ .  $\square$

This lemma implies readily stability for persistent spaces constructed out of a Morse-type function stability thanks to the following initial stability result:

**Theorem 9** (Stability). *Let  $X$  be a topological space and let  $f, g : X \rightarrow \mathbb{R}$  be two continuous maps. Denote*

$$\begin{array}{ccc}
 X_{\bullet}^f : \mathbb{R} & \longrightarrow & \text{Top} \\
 r & \longmapsto & f^{-1}((-\infty, r])
 \end{array}
 \quad \text{and} \quad
 \begin{array}{ccc}
 X_{\bullet}^g : \mathbb{R} & \longrightarrow & \text{Top} \\
 r & \longmapsto & g^{-1}((-\infty, r])
 \end{array}$$

*the two filtered spaces associated to  $f$  and  $g$ . We have*

$$d_{\text{Top}}(X_{\bullet}^f, X_{\bullet}^g) \leq \|f - g\|_{\infty}.$$

*Proof.* Let  $\varepsilon \geq 0$  such that  $\|f - g\|_\infty \leq \varepsilon$  and fix  $r$  in  $\mathbb{R}$ . Let  $x$  be a point in  $X_r^f$ : there exists  $s \leq r$  such that  $f(x) = s$ . By assumption, we have  $|g(x) - f(x)| \leq \varepsilon$ , then  $g(x) \leq s + \varepsilon$  so  $x$  is in  $X_{r+\varepsilon}^g$ . So we have  $X_r^f \subseteq X_{r+\varepsilon}^g$ . By the argument, we have, for all  $r$  in  $\mathbb{R}$ , the following inclusions

$$X_r^f \subseteq X_{r+\varepsilon}^g \subseteq X_{r+2\varepsilon}^f \text{ and } X_r^g \subseteq X_{r+\varepsilon}^f \subseteq X_{r+2\varepsilon}^g,$$

which define an  $\varepsilon$ -interleaving between  $X_\bullet^f$  and  $X_\bullet^g$ .  $\square$

1.1.2. *Two canonical examples of persistent spaces.* Many interesting persistent objects arise from simplicial complexes/sets constructions. We denote by  $\text{sSet}$ , the category of simplicial sets. A simplicial set has a canonical associated topological space, called its geometric realization. We denote

$$|-| : \text{sSet} \longrightarrow \text{Top}$$

the associated functor. It is the left adjoint of the *singular simplicial complex functor*  $\text{Sing} : \text{Top} \rightarrow \text{sSet}$  which is defined, for  $X$  a topological space, by  $\text{Sing}_n(X) := \text{Hom}_{\text{Top}}(\Delta_{\text{Top}}^n, X)$ , where  $\Delta_{\text{Top}}^n$  is the  $n$ -th topological standard simplex. An important property is that they have the same homotopy theories (and the functors actually form a Quillen equivalence). In the rest of the paper, we often identify simplicial sets and topological spaces when defining (co)chains and (co)homology type functors/constructions.

**Definition 10** (Vietoris-Rips complex). Let  $X$  be a metric space. For  $a$  in  $\mathbb{R}$ , we define a simplicial set  $\mathcal{R}_a(X, d_X)$  on the vertex set  $X$  by the following condition:

$$\sigma \in \mathcal{R}_a(X, d_X) \iff d_X(x, y) \leq a \text{ for all } x, y \text{ in } \sigma.$$

The collection of these simplices is the *Vietoris-Rips filtered complex* of  $X$  denoted

$$\mathcal{R}(X, d_X) : \mathbb{R} \longrightarrow \text{sSet}.$$

**Definition 11** (Intrinsic Čech complex). Let  $X$  be a metric space. For  $a$  in  $\mathbb{R}$ , we define a simplicial set  $\check{\mathcal{C}}_a(X, d_X)$  on the vertex set  $X$  by the following condition:

$$[x_0, x_1, \dots, x_k] \in \check{\mathcal{C}}_a(X, d_X) \iff \bigcap_{i=0}^k B(x_i, a) \neq \emptyset.$$

We denote the *intrinsic Čech filtered complex* of  $X$  by

$$\check{\mathcal{C}}(X, d_X) : \mathbb{R} \longrightarrow \text{sSet}.$$

For  $a$ , a real number, and  $\sigma = [x_0, \dots, x_n]$ , a simplex of  $\check{\mathcal{C}}_a(X, d_X)$ , an element  $\bar{x}$  of  $\bigcap_i B(x_i, a)$  is called a *a-center* of  $\sigma$ .

1.2. **Distances for persistent algebras.** We start by reviewing of some classical objects in algebraic topology.

**Definition/Proposition 12** (Singular (co)chain functor). The *singular chain functor*, denoted by  $C_*^{\text{Sing}} : \text{Top} \rightarrow \text{Ch}_k$ , is defined by  $C_*^{\text{Sing}}(X) := k[\text{Sing}(X)]$ , where the differential is given by the signed sum of the maps induced by the face maps. The *singular cochain functor*  $C_{\text{sIs}}^* : \text{Top}^{\text{op}} \rightarrow \text{Alg}_{\text{sIs}}$  is defined, for all topological space  $X$ , by the cochain complex  $C_{\text{sIs}}^*(X) := \text{Hom}_k(C_*^{\text{Sing}}(X), k)$ . Equipped with the *cup product* defined as the composite

$$- \cup - : C_{\text{sIs}}^*(X) \otimes C_{\text{sIs}}^*(X) \longrightarrow C_{\text{sIs}}^*(X \times X) \xrightarrow{\Delta^*} C_{\text{sIs}}^*(X)$$

where the first map is the Künneth map and the second is the map induced by the diagonal, the singular cochain complex is a *differential graded associative algebra* (dg-algebra for short).

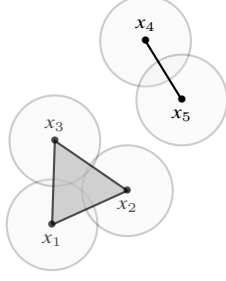


FIGURE 1. The Vietoris Rips complex of a set  $X$

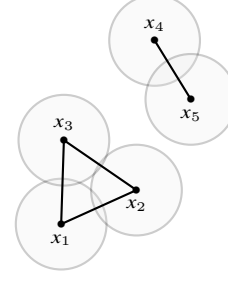


FIGURE 2. The Čech complex of the same set  $X$

Therefore we have in particular a notion of persistent cochain complex and a notion of persistent *dg-algebra* which are naturally induced by persistent spaces  $X : \mathbb{R} \rightarrow \text{Top}$ . The natural notion of equivalences for these are given by persistent quasi-isomorphisms, that is natural transformations  $A^* \rightarrow B^*$  which induced isomorphisms in cohomology; here  $A^*, B^*$  are either persistent dg-algebras  $\mathbb{R} \rightarrow \text{Alg}_{\mathcal{S}S}$  or persistent (co)chain complexes  $\mathbb{R} \rightarrow \text{gr-Vect}$ . Both chain complexes and dg-algebras have natural notions of homotopies and we can pass to their homotopy categories. The cohomology functors from dg-algebras (resp. cochain complexes) factors through the respective homotopy categories.

**Definition 13.** Let

- $\text{ho}(\text{Alg}_{\mathcal{S}S}) := \text{Alg}_{\mathcal{S}S}[\text{qiso}^{-1}]$  be the homotopy category of dg-associative algebras where the quasi-isomorphisms are formally inverted;
- $\text{ho}(\text{gr-Vect}) := \text{Ch}[\text{qiso}^{-1}]$  be the homotopy category of cochain complexes where the quasi-isomorphisms are formally inverted.

We denote  $C_{\text{ho}(\mathcal{S}S)} : \text{Top}^{\text{op}} \rightarrow \text{ho}(\text{Alg}_{\mathcal{S}S})$ , the composition of  $C_{\mathcal{S}S}$  with the canonical quotient functor  $\text{Alg}_{\mathcal{S}S} \rightarrow \text{ho}(\text{Alg}_{\mathcal{S}S})$ .

**Notation 14.** We consider the following functors associated to spaces (and simplicial sets):

- (1)  $C_{\text{ho}(\text{Ch})}^* : \text{Top}^{\text{op}} \rightarrow \text{ho}(\text{Ch}_k)$  with  $C_{\text{ho}(\text{Ch})}^* = \text{For} \circ C_{\text{ho}(\mathcal{S}S)}^*$ ;
- (2)  $C_{\text{Ch}}^* : \text{Top}^{\text{op}} \rightarrow \text{Ch}_k$  with  $C_{\text{Ch}}^* = \text{For} \circ C_{\mathcal{S}S}^*$ ;
- (3)  $H_{\mathcal{C}om}^* : \text{Top}^{\text{op}} \rightarrow \text{Alg}_{\mathcal{C}om}$ ;
- (4)  $H_{\mathcal{S}S}^* : \text{Top}^{\text{op}} \rightarrow \text{Alg}_{\mathcal{S}S}$  with  $H_{\mathcal{S}S}^* = \text{For} \circ H_{\mathcal{C}om}^*$  or  $H_{\mathcal{S}S}^* = H \circ C_{\mathcal{S}S}^*$ ;
- (5)  $H_{\text{gr-Vect}}^* : \text{Top}^{\text{op}} \rightarrow \text{gr-Vect}$  with  $H_{\text{gr-Vect}}^* = \text{For} \circ H_{\mathcal{S}S}^*$ .

We can summarize all these functors in the following commutative diagram:

$$\begin{array}{ccccc}
 & & & & C_{\text{ho}(\text{Ch})} \\
 & & & & \curvearrowright \\
 \text{Top}^{\text{op}} & & & & \text{ho}(\text{Ch}_k) \\
 \downarrow H_{\mathcal{C}om} & \searrow C_{\text{ho}(\mathcal{S}S)} & \searrow \text{For} & \searrow H & \\
 & & \text{ho}(\text{Alg}_{\mathcal{S}S}) & \xrightarrow{\text{For}} & \text{ho}(\text{Ch}_k) \\
 & \searrow H_{\mathcal{S}S} & \downarrow H & \downarrow H & \\
 & & \text{Alg}_{\mathcal{S}S} & \xrightarrow{\text{For}} & \text{gr-Vect} \\
 & \downarrow H_{\mathcal{C}om} & \downarrow H & \downarrow H & \\
 \text{Alg}_{\mathcal{C}om} & \xrightarrow{\text{For}} & \text{Alg}_{\mathcal{S}S} & \xrightarrow{\text{For}} & \text{gr-Vect}
 \end{array}$$

Associated to each of these functors, we can define their interleaving distances according to Definition 4.



**Notation 15.** Let  $X : \mathbb{R} \rightarrow \text{Top}$  and  $Y : \mathbb{R} \rightarrow \text{Top}$  be two persistent topological spaces. According to Remark 7, some of these distance are denoted as follows:

$$d_{\text{gr-Vect}}(X, Y) := d_{\text{gr-Vect}}(H^*(X), H^*(Y)), \quad d_{\text{dls}}(X, Y) := d_{\text{Alg}_{\text{dls}}}(H^*(X), H^*(Y)).$$

Remark that the distance  $d_{\text{gr-Vect}}$  is the *classical* interleaving distance (see [Oud15b]).

**Remark 16.** Considering the homotopy category of persistent cochain complexes, that is objects of the type  $C_{\text{ho(Ch)}}^*(X)$ , is an analogue of considering the derived category of sheaves over  $\mathbb{R}$  which is a recent and promising nice approach to persistent homology, see [KS18]. Indeed, the associated interleaving distance  $d_{\text{ho(Ch}_k)}$  has been considered in [BP19] and seen to agree with the convolution distance of [KS18] for sheaves.

**Remark 17.** We can of course also consider the interleaving distance on the persistent cochain complexes  $C^*(-)$  of persistent spaces  $X, Y : \mathbb{R} \rightarrow \text{Top}$  both taken in *dg-algebras*. This is however not a very pertinent notion to look at. Indeed, if  $T$  is a triangulation of simplicial complex  $X$ , there is no natural dg-algebra map  $C^*(X) \rightarrow C^*(T)$  though they are homotopy equivalent simplicial complexes and, in fact, have the same underlying space.

In particular, given a persistent space and a persistent triangulation of it,  $C^*(X)$  and  $C^*(T)$  are not  $\varepsilon$ -interleaved in  $\text{Alg}_{\text{dls}}$  for small  $\varepsilon$  and therefore  $d_{\text{Alg}_{\text{dls}}}(C^*(X), C^*(T)) \gg 0$  in general event though they represent the same topological space. In particular this distance will not satisfy stability (as in Theorem 53). This problem of course disappear precisely in the homotopy category of algebras.

Note that, unlike for dg-algebras, over a field, one can always find a highly non-natural inverse for a cochain complex map  $C^*(X) \rightarrow C^*(T)$  for fixed spaces so that the situation looks better. But however, in general, one can not find such an inverse in a persistent way. The problem is similar to Remark 35.

These distances satisfy the following inequalities.

**Proposition 18.** *Let  $X : \mathbb{R} \rightarrow \text{Top}$  and  $Y : \mathbb{R} \rightarrow \text{Top}$  be two persistent topological spaces. We have the following inequalities:*

$$\begin{aligned} d_{\text{gr-Vect}}(X, Y) &\leq d_{\text{dls}}(X, Y) = d_{\text{Alg}_{\mathbb{Z}\text{-om}}}(H^*(X), H^*(Y)), \\ d_{\text{gr-Vect}}(X, Y) &\leq d_{\text{ho(Ch)}}(C^*(X), C^*(Y)) \leq d_{\text{ho(Alg}_{\text{dls}})}(C^*(X), C^*(Y)), \\ d_{\text{gr-Vect}}(X, Y) &\leq d_{\text{dls}}(X, Y) \leq d_{\text{ho(Alg}_{\text{dls}})}(C^*(X), C^*(Y)). \end{aligned}$$

*Proof.* It follows from Lemma 8. □

**Example 19.** We exhibit two data sets  $D_X$  and  $D_Y$  in  $\mathbb{R}^3$  such that the interleaving distance between the cohomology of their Čech complex as graded vector space is strictly smaller than their interleaving distance as copersistent graded algebras. Consider the following two topological spaces  $X$  and  $Y$  living in  $\mathbb{R}^3$ :

$$\begin{aligned} X &= \{(x, y, 0) \mid x^2 + (y - 2)^2 = 1\} \cup \{x^2 + y^2 + z^2 = 1\} \cup \{(x, y, 0) \mid x^2 + (y + 2)^2 = 1\}; \\ Y &= \{(x, y, z) \mid (x^2 + y^2 + z^2 + 3)^2 = 16(x^2 + y^2)\}, \end{aligned}$$

which are represented in Figure 3 and Figure 4 respectively. For  $\varepsilon \geq 0$ , we denote

$$X_\varepsilon := \bigcup_{x \in X} B(x, \varepsilon) \quad \text{and} \quad Y_\varepsilon := \bigcup_{y \in Y} B(y, \varepsilon).$$

Note that we have

$$H^*(X) \underset{\text{gr-vect}}{\cong} H^*(Y) \quad \text{but} \quad H^*(X) \underset{\text{Alg}_{\text{dls}}}{\not\cong} H^*(Y)$$



FIGURE 3. The space  $X$

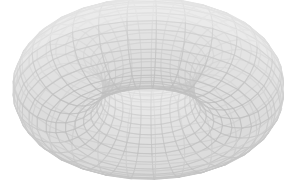


FIGURE 4. The space  $Y$

since the cup product of any non-zero cohomology classes of  $X$  is trivial. Now remark that, for  $\varepsilon < 1$ , we have homotopy equivalences  $X_\varepsilon \simeq X$  and  $Y_\varepsilon \simeq Y$  and, for  $\varepsilon \geq 1$ ,  $X_\varepsilon$  and  $Y_\varepsilon$  are contractile. We fix  $0 < \alpha \ll \frac{1}{2}$  and we choose a finite cover by open balls of radius  $\alpha$  of  $X_\alpha$  and  $Y_\alpha$ :

$$X_\alpha \subset \bigcup_{i \in I} B(x_i, \alpha) \quad \text{and} \quad Y_\alpha \subset \bigcup_{j \in J} B(y_j, \alpha)$$

where  $I$  and  $J$  are finite. We can then define the discrete spaces

$$D_X^\alpha = \bigcup_{i \in I} \{x_i\} \quad \text{and} \quad D_Y^\alpha = \bigcup_{j \in J} \{y_j\}.$$

We can think these spaces as noisy discretisations of our spaces  $X$  and  $Y$ . Remark that, for  $\alpha \leq \varepsilon' < 1 - \alpha$ , we have

$$\check{\mathcal{C}}(D_X^\alpha)_{\varepsilon'} \simeq X_\alpha \simeq X \quad \text{and} \quad \check{\mathcal{C}}(D_Y^\alpha)_{\varepsilon'} \simeq Y_\alpha \simeq Y.$$

For all  $\varepsilon \geq \alpha$ , we have  $H^*(\check{\mathcal{C}}(D_X^\alpha)) \cong_{\text{gr-Vect}} H^*(\check{\mathcal{C}}(D_Y^\alpha))$ , therefore

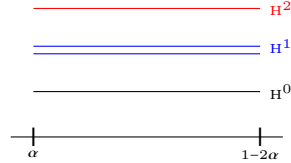


FIGURE 5. A part of the barcode of  $H^*(\check{\mathcal{C}}(D_X^\alpha)_\bullet)$  and  $H^*(\check{\mathcal{C}}(D_Y^\alpha)_\bullet)$

$$d_{\text{gr-Vect}} \left( \check{\mathcal{C}}(D_X^\alpha)_\bullet, \check{\mathcal{C}}(D_Y^\alpha)_\bullet \right) \leq \alpha.$$

Suppose that there is  $\varepsilon < \frac{1-2\alpha}{2}$  such that there exists a  $\varepsilon$ -interleaving in the category  $\text{Alg}_{\text{Sis}}$  between  $H^*(\check{\mathcal{C}}(D_X^\alpha)_\bullet)$  and  $H^*(\check{\mathcal{C}}(D_Y^\alpha)_\bullet)$ : then we will have the following diagram

$$\begin{array}{ccc} & H^*(\check{\mathcal{C}}(D_Y^\alpha)_{\alpha+\varepsilon}) & \\ \mu \nearrow & & \searrow \nu \\ H^*(\check{\mathcal{C}}(D_X^\alpha)_\alpha) & \xrightarrow{\cong} & H^*(\check{\mathcal{C}}(D_X^\alpha)_{\alpha+2\varepsilon}) \end{array},$$

that implies that  $\mu$  and  $\nu$  are isomorphisms: this is a contradiction because  $H^*(\check{\mathcal{C}}(D_X^\alpha)_\alpha)$  and  $H^*(\check{\mathcal{C}}(D_Y^\alpha)_{\alpha+\varepsilon})$  have not the same product. So, we have the following inequality:

$$d_{\text{Sis}} \left( \check{\mathcal{C}}(D_X^\alpha)_\bullet, \check{\mathcal{C}}(D_Y^\alpha)_\bullet \right) \geq \frac{1-2\alpha}{2}.$$

Then, as soon as  $\alpha < \frac{1}{4}$ , we have

$$d_{\text{gr-Vect}}\left(\check{\mathcal{C}}(D_X^\alpha), \check{\mathcal{C}}(D_Y^\alpha)\right) < d_{\mathcal{A}s}\left(\check{\mathcal{C}}(D_X^\alpha), \check{\mathcal{C}}(D_Y^\alpha)\right).$$

and the smaller is  $\alpha$ , the bigger is the difference between the two distances.

**Remark 20.** The previous example can be done for the Rips complex of the same data sets  $D_X^\alpha, D_Y^\alpha$  by considering similar balls but for the euclidean norm replaced by the  $\| - \|_\infty$  norm.

**Remark 21.** It is easily to construct several examples along the line of Example 19.

## 2. THE $\mathcal{A}_\infty$ -INTERLEAVING DISTANCE

One big drawback of working with the homotopy category  $\text{ho}(\text{Alg}_{\mathcal{A}s})$  of dg-associative algebras is that it is hard to study, namely because we cannot, in general, invert a quasi-isomorphisms of dg-algebras (that is find a morphism in the opposite direction inducing the inverse in cohomology). One is forced to work with zigzags of morphisms (in other words to consider maps from  $X$  to  $Y$ , one has to consider to pass through any other object  $Z$ ) and to put a complicated equivalence relation on zigzags. One classical way to avoid that in algebraic topology, which will also relate the structure to the barcode of a persistence space, is to replace dg-algebras by  $A_\infty$ -algebras. We investigate the associated distance in this section.

**2.1. Review on associative algebras up to homotopy.** This part is a rapid overview of the theory of  $A_\infty$ -algebras: the interested reader can consult [LV12, Chapter 9].

**Definition 22** ( $A_\infty$ -algebra). An  $A_\infty$ -algebra is a graded vector space  $A = \{A_k\}_{k \in \mathbb{Z}}$ , equipped with an  $n$ -ary operation

$$m_n : A^{\otimes n} \longrightarrow A \quad \text{of degree } n - 2 \text{ for all } n \geq 1,$$

which satisfy, for all  $n \geq 1$ , the following relation

$$(\text{rel}_n) \quad \sum_{p+q+r=n} (-1)^{p+qr} m_{p+r+1} \circ (\text{id}_A^{\otimes p} \otimes m_q \otimes \text{id}_A^{\otimes r}) = 0.$$

**Remark 23.** (1) By the relation  $(\text{rel}_1)$ , the map  $m_1$  is a differential, and the relation  $(\text{rel}_2)$  implies that the binary product  $m_2$  is a chain complex map.

(2) A differential graded associative algebra  $(A, \mu, d_A)$  is an  $A_\infty$ -algebra where  $m_1 = d_A$ ,  $m_2 = \mu$  and, for  $n \geq 3$ ,  $m_n = 0$ .

The notion of morphisms of  $A_\infty$ -algebras is too rigid to encode the homotopy theory of  $A_\infty$ -algebras in a practical way, so we use a more flexible notion of morphisms, called  $\infty$ -morphisms. This notion of  $\infty$ -morphism allows to define the category  $\infty\text{-Alg}_{\mathcal{A}s_\infty}$  (see Definition/Proposition 26): this category has a notion of  $\infty$ -quasi-isomorphism which we use to define its homotopy category. The point is that each  $\infty$ -quasi-isomorphism has an inverse in homology in the strict category  $\infty\text{-Alg}_{\mathcal{A}s_\infty}$ . So, we do *not* need to consider zigzags of morphisms to understand what is a morphism in the homotopy category of  $\mathcal{A}_\infty$ -algebras. This makes the notion of interleaving completely straightforward.

**Definition 24** ( $\infty$ -morphisms between  $A_\infty$ -algebras). Let  $(A, m_*^A)$  and  $(B, m_*^B)$  be two  $A_\infty$ -algebras. An  $\infty$ -morphism  $f : A \rightsquigarrow B$  of  $A_\infty$ -algebras is a family of maps  $\{f_n : A^{\otimes n} \rightarrow B\}_{n \geq 1}$  of degree  $n - 1$  such that  $f_1$  is a

morphism of chain complexes and, for  $n \geq 2$ ,  $f_n$  satisfies the relation

$$\begin{aligned} \partial(f_n) = & \sum_{p+q+r=n} (-1)^{p+qr} f_{p+r+1} \circ (\text{id}_A, \dots, \text{id}_A, m_q^A, \text{id}_A, \dots, \text{id}_A) \\ & - \sum_{\substack{k \geq 2 \\ i_1 + \dots + i_k = n}} (-1)^{\sum_{j=i}^{k-1} (k-j)(i_j-1)} m_k^B \circ (f_{i_1}, \dots, f_{i_k}) . \end{aligned}$$

The set of  $\infty$ -morphism from  $A$  to  $B$  is denoted by  $\text{Hom}_{\infty-A_\infty}(A, B)$ .

**Remark 25.** A (strict) morphism  $f : A \rightarrow B$  of  $A_\infty$ -algebras, i.e. which satisfies

$$f \circ m_n^A = m_n^B \circ f^{\otimes n}$$

for all  $n \geq 1$ , is canonically a  $\infty$ -morphism given by the family of maps  $(f, 0, 0 \dots)$ .

**Definition/Proposition 26** (The category  $\infty\text{-Alg}_{\mathcal{S}l_\infty}$ ). There exists an associative composition of  $\infty$ -morphisms:

$$\text{Hom}_{\infty-A_\infty}(B, C) \times \text{Hom}_{\infty-A_\infty}(A, B) \longrightarrow \text{Hom}_{\infty-A_\infty}(A, C)$$

such that, for all  $A_\infty$ -algebra  $A$ , the identity  $A \rightsquigarrow A$  is given by the strict classical one. The  $A_\infty$ -algebras and the  $\infty$ -morphisms form a category denoted by  $\infty\text{-Alg}_{\mathcal{S}l_\infty}$ .

An important tool we will use is the fact that a chain complex which is quasi-isomorphic to an  $A_\infty$ -algebra automatically inherits a transferred quasi-isomorphic  $A_\infty$ -structure. In particular, we have the following theorem, which gives us a canonical  $A_\infty$ -structure on the homology of an  $A_\infty$ -algebra.

**Theorem 27** (Homotopy Transfer Theorem [LV12, Theorem 10.3.7 and 10.3.10]). *Let  $k$  be a field of characteristic 0. Let  $A$  be an  $A_\infty$ -algebra,  $i, p$  two morphisms of chain complexes and  $h$  a map of degree 1:*

$$h \circlearrowleft (A, d_A) \begin{array}{c} \xrightarrow{p} \\ \xleftarrow{i} \end{array} (H(A), 0) ,$$

such that  $ip - \text{id}_A = d_A h + h d_A$ .

- (1) *There is a  $A_\infty$ -structure on the homology  $H(A)$  of the underlying chain complex of  $A$ , which extends its associative algebra structure.*
- (2) *The embedding  $i$  and the projection  $p$ , associated to the choice of sections for the homology, extend to  $\infty$ -quasi-isomorphisms of  $A_\infty$ -algebras.*
- (3) *The  $A_\infty$ -structure on the homology  $H(A)$  is independent of the choice of sections of  $H(A)$  into  $A$  in the following sense: any two such transferred structures are related by an  $\infty$ -isomorphism, whose first map is the identity on  $H(A)$ .*

**Remark 28.** The characteristic 0 assumption in the previous theorem is used to guarantee the existence of the  $\infty$ -morphism  $p_\infty$ .

**Example 29** (Higher Massey products). Let  $X$  be a connected topological space,  $p, i, h$  be a choice of contraction between  $C^*(X)$  and  $H^*(X)$  as follows:

$$h \circlearrowleft (C^*(X), \cup, \partial) \begin{array}{c} \xrightarrow{p} \\ \xleftarrow{i} \end{array} (H^*(X), 0) ,$$

where  $\cup$  is the associative cup product on  $C^*(X)$ . Then, the transferred  $A_\infty$ -structure on  $H^*(X)$  gives us the higher Massey products. These products allow to make the difference between some spaces which have the same cohomology as graded algebra. For instance, we can consider the complement in the 3-sphere of the

Borromean rings (see Figure 6) and the trivial entanglement of three rings (see Figure 7). These two spaces have the same cohomology as graded algebra but they have different  $m_3$  Massey product (see [LV12, Section 9.4.5]).

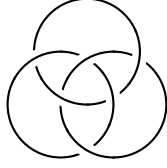


FIGURE 6. Borromean rings

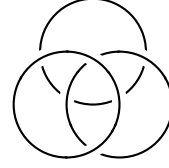


FIGURE 7. Trivial entanglement

Other standard examples of non-trivial Massey products are obtained by Kodaira-Thurston manifolds (see [RT00] for example).

The notion of  $A_\infty$ -algebra and  $A_\infty$ -morphisms gives a nice and practical model for the homotopy category of dg-algebras.

**Theorem 30** (Equivalence of homotopy category (see [LH02] or [LV12, Theorem 11.4.8])). *The homotopy category of differential graded associative algebras and the homotopy category of  $A_\infty$ -algebras with the  $\infty$ -morphisms are equivalent:*

$$\text{ho}(\text{Alg}_{\mathcal{A}S}) \cong \text{ho}(\infty\text{-Alg}_{\mathcal{A}S}) \cong \text{ho}(\infty\text{-Alg}_{\mathcal{A}\infty}).$$

**2.2. The problem of persistent transfer datum.** For a persistent topological space  $X : \mathbb{R} \rightarrow \text{Top}$  and  $a > b$ , we want to associate canonically contractions datum:

$$\begin{array}{ccc} h_a \hookrightarrow C^*(X_a) & \begin{array}{c} \xrightarrow{p_a} \\ \xleftarrow{i_a} \end{array} & H^*(X_a) \\ C^*(\varphi_{ab}) \downarrow & & \\ h_b \hookrightarrow C^*(X_b) & \begin{array}{c} \xrightarrow{p_b} \\ \xleftarrow{i_b} \end{array} & H^*(X_b) \end{array} ;$$

This is not possible in general, because we have to make choices, which have no reason to be compatible with the persistence structure maps (see Remark 35). Therefore we restrict our category of persistent topological spaces to a subcategory of objects satisfying some finiteness conditions. These conditions are automatically satisfied for a set of data such as those arising in applications.

Let us denote by  $\Delta\text{Cpx}$ , the category of delta complexes as in [Hat02]. Note that for Čech and Rips complexes arising from a data set, it is sufficient to consider simplicial complexes and not general delta complexes.

**Definition 31** (Finite filtered data). We consider the full subcategory of the category of persistent delta-complex  $\text{Func}(\mathbb{R}, \Delta\text{Cpx})$  with objects  $X : \mathbb{R} \rightarrow \Delta\text{Cpx}$  satisfying:

- (1) for all  $r$  in  $\mathbb{R}$ , the delta complex  $X_r$  is finite;
- (2) for all  $a < b$  in  $\mathbb{R}$ , the morphism  $X_a \hookrightarrow X_b$  is an injection;
- (3) the set  $\{X_r \mid r \in \mathbb{R}\} / \sim$ , where  $X_a \sim X_b$  if the structural morphism  $X_{a < b}$  is an isomorphism, is finite ;
- (4) for each  $a$  in  $\mathbb{R}$ , we have a total order  $\leq_a$  on  $X_a$  such that
  - for  $\alpha$  and  $\beta$  in  $X_a$  such that  $\alpha \subset \beta$  then  $\alpha \leq_a \beta$  ;
  - for all  $a < b$  in  $\mathbb{R}$ , for all  $\beta$  in  $X_b \setminus X_a$ , and all  $\alpha$  in  $X_a$ , then  $\alpha \leq_a \beta$ .

We call this category *the category of finite filtered data* and we denote it by  $\text{fData}^{\mathbb{R}}$ .

Our motivating example for this definition is the following.

**Example 32.** Let  $X$  be a finite set of points in a metric space. Then,  $\tilde{\mathcal{C}}(X)_\bullet$  and  $\mathcal{R}(X)_\bullet$  are objects in  $\text{fData}^{\mathbb{R}}$  (see Definition 10 and Definition 11).

The conditions of a finite filtered data ensures that the interleaving distances are closed on this subcategory:

**Lemma 33.** *Let  $X$  and  $Y$  be two finite filtered data and  $F : \text{Top} \rightarrow \mathcal{C}$  be a functor. We have that  $d_{\mathcal{C}}(F(X), F(Y)) = 0$  if, and only if, the persistent objects  $F(X)$  and  $F(Y)$  are isomorphic in  $\mathcal{C}^{\mathbb{R}}$ .*

*Proof.* We denote by  $r_1, \dots, r_n$ , the objects of the category  $\mathbb{R}$  such that, for all  $1 \leq i \leq n$  and for all  $\varepsilon > 0$  in  $\mathbb{R}$ ,  $F(X)_{r_i - \varepsilon} \not\cong F(X)_{r_i}$  or  $F(Y)_{r_i - \varepsilon} \not\cong F(Y)_{r_i}$ .

Suppose that  $d_{\mathcal{C}}(F(X), F(Y)) = 0$ . Let  $t < r_n$  be in  $\mathbb{R}$ : there exists  $1 \leq i \leq n$  such that  $r_i \leq t < r_{i+1}$ . We consider  $\varepsilon = \frac{r_{i+1} - t}{3}$ . By assumption, there exists an  $\varepsilon$ -interleaving between  $F(X)$  and  $F(Y)$ :

$$F(X)_t \longrightarrow F(Y)_{t+\varepsilon} \longrightarrow F(X)_{t+2\varepsilon};$$

as  $t + 2\varepsilon < r_{i+1}$ , then  $F(X)_t \cong F(X)_{t+2\varepsilon}$ , so  $F(X)_t \cong F(Y)_{t+\varepsilon} \cong F(Y)_t$ . By the same argument, for  $t \geq r_n$ ,  $F(X)_t \cong F(Y)_t$ .  $\square$

**2.3. Contractions in family.** In order to apply the homotopy transfer theorem for a persistent space given by a finite filtered data, we need to encode some part of the data allowing to obtain the transferred structure in an explicit way. This is the role of the following of definition.

**Definition 34** (Category of transfer data). The category  $\text{Trans}_{\text{Ch}}$  (resp.  $\text{Trans}_{\text{coCh}}$ ) is defined as follows: its objects are  $(A, H, i, p, h)$  such that

$$h \circlearrowleft (A, d_A) \xrightleftharpoons[i]{p} (H, 0) ,$$

where  $A$  is a chain complex (resp. cochain complex),  $H$  is a graded vector space,  $i$  and  $p$  are morphisms of chain complexes,  $h$  is a linear map of homological degree 1 (resp. homological degree  $-1$ ) such that:

$$ip - \text{id}_A = d_A h + h d_A, \quad pi = \text{id}_H, \quad hi = 0, \quad ph = 0 \quad \text{and} \quad h^2 = 0.$$

The morphisms  $(A, H, i, p, h) \rightarrow (A', H', i', p', h')$  in the category  $\text{Trans}_{\text{Ch}}$  are defined as the morphisms  $A \rightarrow A'$  of chain complexes.

We also define the category  $\text{Trans}_{\text{Ch}, \text{sls}}$  to be the subcategory of  $\text{Trans}_{\text{Ch}}$  whose objects are  $(A, H, i, p, h)$  such that  $A$  is a differential graded associative algebra, and the morphisms are morphisms of differential graded associative algebras.

**Remark 35.** Let  $\varphi : (A, H, i, p, h) \rightarrow (A', H', i', p', h')$  be a morphism in  $\text{Trans}_{\text{Ch}}$ . We do not suppose any compatibility between the morphism  $\varphi$  and the structural morphisms  $i, p, h$  and  $i', p', h'$ , contrarily to the classical notion of [MS17]. The reason is that we do not have a persistent inclusion of the homology of a chain complex in it, even in the simplest cases, as the following example shows. Let  $X$  be the union of two points  $(0, 0)$  and  $(1, 0)$  in  $\mathbb{R}^2$ , and consider the associate Vietoris-Rips complex  $\mathcal{R}(X)_\bullet$ . We have the table of Figure 8.

Suppose that we have a persistent inclusion  $i_\bullet : H_*(\mathcal{R}(X))_\bullet \rightarrow C_*(\mathcal{R}(X))_\bullet$ . Then, we obtain, for all  $\varepsilon < \frac{1}{2}$ , the following diagram

$$\begin{array}{ccc} kx_1 \oplus kx_2 & \xrightarrow[\cong]{i_\varepsilon} & kx_1 \oplus kx_2 \\ (\varepsilon \leq \varepsilon + \frac{1}{2}) \downarrow & & \cong \downarrow (\varepsilon \leq \varepsilon + \frac{1}{2}) \\ kx_1 & \xrightarrow{i_{\varepsilon + \frac{1}{2}}} & kx_1 \oplus kx_2 \end{array}$$

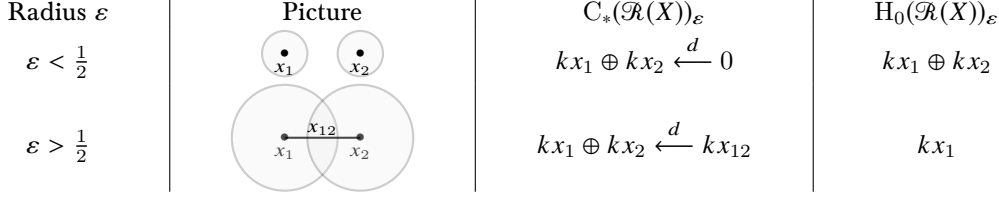


FIGURE 8. Simplicial chain complex and simplicial homology of the Vietoris-Rips complex associated to  $\{(0, 0), (1, 0)\}$

which is *not* commutative. Therefore, it cannot exist such a persistent inclusion. By the same argument, it cannot exist a persistent projection  $p^\bullet : H^*(\mathcal{R}(X))^\bullet \rightarrow C^*(\mathcal{R}(X))^\bullet$  from the persistent cohomology to the persistent cochain complex.

**Notation 36.** We denote by  $\text{Trans}_{\text{Ch}}^{\mathbb{R}}$  the category of functors  $\mathbb{R} \rightarrow \text{Trans}_{\text{Ch}}$ , that is the category of persistent transfer data. We also denote by  $\text{fTrans}_{\text{Ch}}^{\mathbb{R}}$ , the full subcategory of  $\text{Trans}_{\text{Ch}}^{\mathbb{R}}$  of objects  $A_\bullet$  such that the set  $\{A_r \mid r \in \mathbb{R}\} / \sim$ , where  $A_\alpha \sim A_\beta$  if the structural morphism  $A_{\alpha < \beta}$  is an isomorphism, is finite.

The functor of simplicial complex  $C_* : \Delta\text{Cpx} \rightarrow \text{Ch}_k$  induces a functor

$$\text{TC}_* : \text{fData}^{\mathbb{R}} \longrightarrow \text{fTrans}_{\text{Ch}}^{\mathbb{R}}$$

where the contraction are given by [RMA09, Algorithm 1]: consider a  $\mathbb{R}$ -filtered data  $X : \mathbb{R} \rightarrow \Delta\text{Cpx}$ . For all  $r$  in  $\mathbb{R}$ ,  $X_r$  is totally ordered, and, as  $X_a \hookrightarrow X_b$  for all  $a < b$  in  $\mathbb{R}$ , we denote  $X_0 = \{c_0, \dots, c_{i_0}\}$ ,  $X_1 = \{c_0, \dots, c_{i_0}, c_{i_0+1} \dots c_{i_1}\}$  and for all  $r$  in  $\mathbb{R}$ ,

$$X_r = \{c_0, c_1, \dots, c_{i_0}, c_{i_0+1}, \dots, c_{i_1}, \dots, c_{i_r}\}.$$

Fix  $r$  in  $\mathbb{R}$ , denote by  $m$ , the cardinal of  $X_r$ , we construct algorithmically a linear map  $h_m$  of homological degree one on  $(\bigoplus_{i=1}^m kc_i, \partial_m)$  as follows:

```

 $C_0 := \{c_0\}, \partial_0, h_0(c_0) := 0$ 
For  $i = 1$  to  $m$  do
   $C_i := \{C_{i-1} \cup \{c_i\}, \partial_i\}$ 
  If  $(\partial_i - \partial_{i-1}h_{i-1}\partial_i)(c_i) = 0$ , then
     $h_i(c_i) := 0$ 
    For  $j = 0$  to  $i - 1$  do
       $h_i(c_j) = h_{i-1}(c_j)$ 
  If  $(\partial_i - \partial_{i-1}h_{i-1}\partial_i)(c_i) = \sum_{k=1}^n \lambda_k u_k \neq 0$  with  $u_1 < \dots < u_k < \dots < u_n \in C_{i-1}$ , then
     $\bar{\varphi}(u_1) := -\lambda_1^{-1}c_i$  and  $\bar{\varphi}(u_k) := 0$  otherwise
    For  $j = 0$  to  $i$  do
       $h_i(c_j) = (h_{i-1} + \bar{\varphi} - \bar{\varphi}h_{i-1}\partial_{i-1}\bar{\varphi}\partial_{i-1}h_{i-1})(c_j)$ 
Output:  $h := h_m$ 

```

Then, for each  $r$  in  $\mathbb{R}$ , the functor  $C_*$  sends  $X_r$  on the contraction  $((C_*(X_r), \partial), \text{Im}(\pi), \iota, \pi, h)$  where  $\pi := \text{id} - \partial h - h \partial$ ,  $\iota$  is the inclusion of  $\text{Im}(\pi)$  in  $C_*(X_r)$ , and  $h$  given by the previous algorithm. Finally, we have define a functor

$$\text{TC}_* : \text{fData}^{\mathbb{R}} \longrightarrow \text{fTrans}_{\text{Ch}}^{\mathbb{R}}.$$

Composition with the functor  $\text{Hom}_k(-, k)$  gives us a functor

$$(1) \quad \text{TC}^* : \text{fData}^{\mathbb{R}} \longrightarrow \text{fTrans}_{\text{coCh}, \mathcal{A}_S}^{\mathbb{R}^{\text{op}}}$$

which we call the *persistent transfer data dg-algebra functor*.

#### 2.4. $\mathcal{A}_\infty$ -interleaving distance.

**Theorem 37.** *There is a functor*

$$\begin{aligned} H_* : \text{fTrans}_{\text{coCh}, \mathcal{A}_S}^{\mathbb{R}^{\text{op}}} &\longrightarrow \text{ho}(\infty\text{-Alg}_{\mathcal{A}_\infty})^{\mathbb{R}^{\text{op}}} \\ (A, H, i, p, h)^\bullet &\longmapsto (H, \{\mu_i\}_{i \in \mathbb{N}})^\bullet \end{aligned}$$

whose composition with the forgetful functor  $\text{ho}(\infty\text{-Alg}_{\mathcal{A}_\infty})^{\mathbb{R}^{\text{op}}} \rightarrow (\text{gr-Vect})^{\mathbb{R}^{\text{op}}}$  is the underlying persistent cohomology. We call  $H_*$  the transferred  $\mathcal{A}_\infty$ -structure functor.

*Proof.* Let  $(A, H, i, p, h)^\bullet$  be an object in  $\text{fTrans}_{\text{coCh}, \mathcal{A}_S}^{\mathbb{R}^{\text{op}}}$ , and consider  $t_0, \dots, t_n$  in  $\mathbb{R}$  such that, for all  $1 \leq j \leq n$ , for all  $\varepsilon > 0$ ,  $A_{t_j - \varepsilon} \not\cong A_{t_j}$  as differential graded algebras. By the Homotopy Transfer Theorem (see Theorem 27), for each  $t_j$ ,  $H_{t_j}$  has an  $\mathcal{A}_\infty$ -structure, and for all  $t_i \leq t < t_{j+1}$ , as  $H_{t_j}$  and  $H_t$  are isomorphic as graded vector spaces. We can thus put on  $H_t$  the same  $\mathcal{A}_\infty$ -structure as on  $H_{t_j}$  and take the identity as the structural morphism between them:

$$(t_j \leq t) : H_{t_j} \xrightarrow{\text{id}} H_t.$$

We need to construct the structural morphism  $\eta_i^H := (t_j \leq t_{j+1})_H : H_{t_j} \rightarrow H_{t_{j+1}}$ . If we denote by  $\eta_i^A$ , the structural morphism  $A_{t_j} \rightarrow A_{t_{j+1}}$ , we define  $\eta_i^H$  by the following composition:

$$\begin{array}{ccc} H_{t_j} & \xrightarrow{\eta_i^H} & H_{t_{j+1}} \\ i_j \downarrow & & \uparrow p_{j+1} \\ A_{t_j} & \xrightarrow{\eta_i^A} & A_{t_{j+1}}. \end{array}$$

Then, by Theorem 27,  $i_{j+1}$  is a  $\infty$ -quasi-isomorphism and  $p_{j+1}$  is its inverse, so the following diagram is homotopy commutative

$$\begin{array}{ccc} H_{t_j} & \xrightarrow{\eta_i^H} & H_{t_{j+1}} \\ i_j \downarrow & & \downarrow i_{j+1} \\ A_{t_j} & \xrightarrow{\eta_i^A} & A_{t_{j+1}} \end{array}.$$

So we have proved that the datum of  $H_\bullet$  with transferred structure and morphisms defined previously is a persistent object in the category  $\text{ho}(\infty\text{-Alg}_{\mathcal{A}_\infty})$ .  $\square$

Combining the persistent transfer data dg-algebra functor (1) and the transferring  $\mathcal{A}_\infty$ -structure functor of Theorem 37, we obtain the following definition.

**Definition 38.** We define the  *$\mathcal{A}_\infty$ -algebra homology functor* as the composition

$$H_* \circ \text{TC}^* : \text{fData}^{\mathbb{R}} \longrightarrow \text{ho}(\infty\text{-Alg}_{\mathcal{A}_\infty})^{\mathbb{R}^{\text{op}}}.$$

It is therefore a functor from finite filtered data to the (homotopy) category  $\text{fo} \mathcal{A}_\infty$ -algebras. This functor is the  $\mathcal{A}_\infty$  analogue of the cochain algebra functor from persistent spaces to (homotopy classes of) dg-algebras (see Definition/Proposition 12 and Notation 14).



**Definition 39** ( $\mathcal{A}_\infty$ -interleaving distance). Let  $X$  and  $Y$  be two filtered data. We define the  $\mathcal{A}_\infty$ -interleaving distance by

$$d_{\mathcal{A}_\infty}(X, Y) := d_{\text{ho}(\infty\text{-Alg}_{\mathcal{A}_\infty})}(\mathbb{H}_*(\text{TC}^*(X)), \mathbb{H}_*(\text{TC}^*(Y))).$$

where the functor  $\mathbb{H}_* \circ \text{TC}^*$  is given by Definition 38.

The  $\mathcal{A}_\infty$ -interleaving distance realizes the interleaving distance in the homotopy category of dg-associative algebras (in characteristic 0), see Proposition 41 below. The point of the  $\mathcal{A}_\infty$ -distance is that we only need to consider actual  $\mathcal{A}_\infty$ -morphisms (instead of zigzags passing to arbitrary intermediate dg-algebras) to study interleavings.

**Remark 40.** F. Belchí *et al.* give a definition of  $\mathcal{A}_\infty$ -barcode, and consider the associated bottleneck distance (see [GM14, BM15, Bel17, BS19]). They work with the transferred  $\mathcal{A}_\infty$ -coalgebra structure  $\{\Delta_n\}$  of the homology of space and construct the  $\mathcal{A}_\infty$ -barcode using the kernel of the coproducts  $\Delta_n$ . However, this definition of barcode has a drawback: they only consider the kernel of the first coproduct  $\Delta_1$  which is not trivial because the kernel of the higher coproducts depend highly on the transfer data (see [Bel17, Section 3]). Therefore this definition lose a large part of the  $\mathcal{A}_\infty$ -structure in general and thus some reachable topological information.

**Proposition 41.** *Let  $X$  and  $Y$  be two filtered data.*

- (1) *We have the inequality  $d_{\mathcal{A}_\infty}(X, Y) \leq d_{\text{Alg}_{\mathcal{A}_\infty}}(C^*(X), C^*(Y))$ .*
- (2) *There is the equality  $d_{\mathcal{A}_\infty}(X, Y) = d_{\text{ho}(\text{Alg}_{\mathcal{A}_\infty})}(C^*(X), C^*(Y))$ .*

What we are really interested about is point (2) of the proposition.

*Proof.* (1) It is an immediate consequence of Lemma 8 and Theorem 37.

- (2) By the Homotopy Transfer Theorem (see Theorem 27), for each  $r$  in  $\mathbb{R}$ , we have the following two  $\infty$ -quasi-isomorphisms

$$C^*(X)^r \begin{array}{c} \xrightarrow{p_\infty^r} \\ \xleftrightarrow{\sim} \\ \xleftarrow{i_\infty^r} \end{array} H^*(X)^r$$

which are quasi-inverse of each others. Therefore, for each  $r$  in  $\mathbb{R}$ ,  $H^*(X)^r$  and  $C^*(X)^r$  are isomorphic in the homotopy category of  $\mathcal{A}_\infty$ -algebras so that we have

$$d_{\mathcal{A}_\infty}(X, Y) = d_{\text{ho}(\infty\text{-Alg}_{\mathcal{A}_\infty})}(C^*(X), C^*(Y)).$$

By the Theorem 30, we have  $\text{ho}(\infty\text{-Alg}_{\mathcal{A}_\infty}) \cong \text{ho}(\text{Alg}_{\mathcal{A}_\infty})$ , so we deduce the result. □

**Remark 42.** Recall that to define the distance  $d_{\text{ho}(\text{Alg}_{\mathcal{A}_\infty})}(C^*(X), C^*(Y))$ , we do not need any finiteness assumption on the functors  $X, Y : \mathbb{R} \rightarrow \text{Top}$ .

**Example 43.** We now gives a finite filtered data version of Example 29. Consider the space  $X$  (respectively the space  $Y$ ) defined as the complementary of an (open)  $\beta$ -thickening of the borromean ring (resp. trivial entanglement of three circles) in  $S^3$ , both embedded in  $\mathbb{R}^4$ . As in Example 19,  $X$  and  $Y$  are compact, so for  $\alpha \ll \beta/2$ , we can construct finite discretisations of  $X$  and  $Y$ , denoted  $\tilde{X}$  and  $\tilde{Y}$  respectively such that  $\bigcup_{x \in \tilde{X}} B(x, \alpha)$ ,  $\bigcup_{y \in \tilde{Y}} B(y, \alpha)$  are covers of  $X$  and  $Y$  respectively. Similarly to Example 19, we have, for  $\alpha < \varepsilon < \beta - \alpha$ , that

$$\mathcal{R}(\tilde{X})_\varepsilon \simeq \bigcup_{x \in \tilde{X}} B(x, \alpha) \simeq X$$

and similarly for  $\mathcal{R}(\widetilde{Y})_\varepsilon$ . Therefore, since  $X$  and  $Y$  have the same cohomology algebras, but different  $\mathcal{A}_\infty$ -algebra structures, we have that

$$d_{\mathcal{A}_S}(\mathcal{R}(\widetilde{X}), \mathcal{R}(\widetilde{Y})) \leq \alpha \quad \text{and} \quad d_{\mathcal{A}_\infty}(\mathcal{R}(\widetilde{X}), \mathcal{R}(\widetilde{Y})) \geq \frac{\beta - 2\alpha}{2}.$$

Recall that  $d_{\mathcal{A}_S}(\mathcal{R}(\widetilde{X}), \mathcal{R}(\widetilde{Y}))$  is the distance associated to the cohomology algebras  $H^*(\mathcal{R}(\widetilde{\quad}))$ . It follows that, as soon as  $\alpha < \frac{\beta}{4}$ , we have

$$d_{\mathcal{A}_S}(\mathcal{R}(\widetilde{X}), \mathcal{R}(\widetilde{Y})) < d_{\mathcal{A}_\infty}(\mathcal{R}(\widetilde{X}), \mathcal{R}(\widetilde{Y})).$$

Note further that the smaller the parameter  $\alpha$  used in the discretisation is, the larger the difference between the two metrics is.

### 3. A STABILITY THEOREM FOR MULTIPLICATIVE DISTANCES

Consider  $X$  and  $Y$ , two compact topological spaces. Assume  $X$  and  $Y$  are almost the same, i.e. the Gromov-Hausdorff distance between  $X$  and  $Y$  is small (for example, if  $Y$  is a small perturbation of  $X$ ), then, we want the interleaving distances between  $C^*(X)$  and  $C^*(Y)$  or  $H^*(X)$  and  $H^*(Y)$  (defined in Section 1.2 and Section 2.4) to be small as well. We obtain this result (see Theorem 53) by adapting the proof of its classical version (cf. [CDSO14]).

**3.1. Gromov-Hausdorff distance.** We first review the basic definitions of Gromov-Hausdorff distance.

**Definition 44** (Multivalued map – Correspondence). Let  $X$  and  $Y$  be two sets. A *multivalued map* from  $X$  to  $Y$  is a subset  $C$  of  $X \times Y$  such that the canonical projection restrict to  $C$   $\pi_X|_C : C \rightarrow X$  is surjective. We denote a multivalued map  $C$  from  $X$  to  $Y$  by  $C : X \rightrightarrows Y$ . The *image*  $C(\sigma)$  of a subset  $\sigma$  of  $X$  is the canonical projection onto  $Y$  of the preimage of  $\sigma$  through  $\pi_X$ . A map  $f : X \rightarrow Y$  is *subordinate* to  $C$  if, for all  $x$  in  $X$ , the pair  $(x, f(x))$  is in  $C$ . In that case, we write  $f : X \xrightarrow{C} Y$ . The *composition* of two multivalued maps  $C : X \rightrightarrows Y$  and  $D : Y \rightrightarrows Z$  is the multivalued map  $D \circ C : X \rightrightarrows Z$ , defined by:

$$(x, z) \in D \circ C \iff \text{there exists } y \in Y \text{ such that } (x, y) \in C \text{ and } (y, z) \in D.$$

A multivalued map  $C : X \rightrightarrows Y$  such that the canonical projection  $C$   $\pi_Y|_C : C \rightarrow Y$  is surjective, is called a *correspondence*. The *transpose* of a correspondence  $C$ , denoted  $C^T$ , is the correspondence defined by the image of  $C$  through the symmetry  $(x, y) \mapsto (y, x)$ .

**Remark 45.** Consider a correspondence  $C : X \rightrightarrows Y$ . Then we have

$$\text{id}_X : X \xrightarrow{C^T \circ C} X \quad \text{and} \quad \text{id}_Y : Y \xrightarrow{C \circ C^T} Y.$$

To a correspondence  $C : X \rightrightarrows Y$ , we associate a quantity called the distortion metric, and we define the Gromov Hausdorff distance.

**Definition 46** (Distortion of a correspondence – Gromov-Hausdorff distance). Let  $(X, d_X)$  and  $(Y, d_Y)$  be two metric spaces. The *distortion* of a correspondence  $C : X \rightrightarrows Y$  is defined as follows:

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

The *Gromov-Hausdorff distance* between the metric spaces  $X$  and  $Y$  is defined as follows:

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

**Remark 47.** Let  $X$  and  $Y$  be two metric spaces. The Gromov-Hausdorff distance between  $X$  and  $Y$  is equal to the following one:

$$d_{\mathcal{GH}}(X, Y) = \inf_{(\gamma, \eta) \in \Gamma} \min\{\varepsilon \geq 0 \mid \gamma(X) \subset \eta(Y)_\varepsilon \text{ and } \eta(X) \subset \gamma(Y)_\varepsilon\}$$

where  $\Gamma = \{X \xrightarrow{\gamma} Z \xleftarrow{\eta} Y \mid (Z, d_Z) \text{ metric space and } \gamma \text{ and } \eta \text{ are isometrical embeddings}\}$ , and

$$(\gamma(X))_\varepsilon := \bigcup_{x \in \gamma(X)} B_Z(x, \varepsilon).$$

### 3.2. Simplicial multivalued map.

**Definition 48** ( $\varepsilon$ -simplicial multivalued map). Let  $\mathcal{S}$  and  $\mathcal{T}$  be two persistent delta complexes such that, for all  $r$  in  $\mathbb{R}$ , the vertex sets of  $\mathcal{S}_r$  and  $\mathcal{T}$  are  $X$  and  $Y$  respectively. A multivalued map  $C : X \rightrightarrows Y$  is  $\varepsilon$ -simplicial for  $\mathcal{S}$  and  $\mathcal{T}$  if, for any  $r$  in  $\mathbb{R}$  and any simplex  $\sigma$  in  $\mathcal{S}_r$ , every finite subset of  $C(\sigma)$  is a simplex of  $\mathcal{T}_{r+\varepsilon}$ .

**Definition 49** (Contiguous maps). Let  $K$  and  $L$  be two simplicial complexes. Two simplicial maps  $f, g : K \rightarrow L$  are *contiguous* if, for each 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  $\tau$  of  $L$ .

**Lemma 50** ([CDSO14, Proposition 3.3]). *Let  $\mathcal{S}, \mathcal{T} : \mathbb{R} \rightarrow \Delta\text{Cpx}$  with vertex sets  $X$  and  $Y$  respectively and let  $C : X \rightrightarrows Y$  be a  $\varepsilon$ -simplicial multivalued map from  $\mathcal{S}$  to  $\mathcal{T}$ . Then, any two subordinate maps  $f_1, f_2 : X \xrightarrow{C} Y$  induce simplicial maps  $\mathcal{S}_a \rightarrow \mathcal{T}_{a+\varepsilon}$  which are contiguous. Also, the maps  $|f_1|$  and  $|f_2|$  are homotopic.*

*Proof.* Let  $\mathcal{S}, \mathcal{T} : \mathbb{R} \rightarrow \Delta\text{Cpx}$  with vertex sets  $X$  and  $Y$  respectively and let  $C : X \rightrightarrows Y$  be a  $\varepsilon$ -simplicial multivalued map from  $\mathcal{S}$  to  $\mathcal{T}$ . Any subordinate map  $f : X \xrightarrow{C} Y$  induces a simplicial map  $\mathcal{S}_r \rightarrow \mathcal{T}_{r+\varepsilon}$  for each  $r$  in  $\mathbb{R}$  by definition of an  $\varepsilon$ -simplicial multivalued map.

Consider two subordinate maps  $f_1, f_2 : X \xrightarrow{C} Y$ , and let  $\sigma = [v_0, \dots, v_n]$  be a simplex in  $\mathcal{S}_r$ . As  $C$  is a  $\varepsilon$ -simplicial multivalued map, then, by definition, every subset of  $C(\sigma)$  is a simplicial set of  $\mathcal{T}_{r+\varepsilon}$ . Therefore, since every  $f_i(v_j)$  is in  $C(\sigma)$ , the set  $f_1(v_0), \dots, f_1(v_n), f_2(v_0), \dots, f_2(v_n)$  is a simplex of  $\mathcal{T}_{r+\varepsilon}$ . Consequently, the simplicial maps induced by  $f_1$  and  $f_2$  are contiguous. By [McC06, Proposition 10.20], contiguous simplicial maps have homotopic realisations; hence, the realisations of  $f_1$  and  $f_2$  are homotopic.  $\square$

**Proposition 51** ([Jam95, Chapter 16, Theorem 3.8.]). *The functor  $C^* : \text{Top}^{\text{op}} \rightarrow \text{Alg}_{\text{dgs}}$  converts weak homotopy equivalences to quasi-isomorphisms and homotopy classes of maps to homotopy classes of morphisms of dg-associative algebras.*

**3.3. Case of Čech and Vietoris-Rips complexes.** In this section, we prove the stability of the distances that we introduced before for the Čech and Vietoris-Rips complexes.

**Lemma 52** ([CDSO14, Lemmas 4.3 and 4.4]). *Let  $(X, d_X)$  and  $(Y, d_Y)$  be two metric spaces, and let  $C : X \rightrightarrows Y$  be a correspondence with distortion at most  $\varepsilon$ . Then*

- (1) *the correspondence  $C$  is  $\varepsilon$ -simplicial from  $\mathcal{R}(X, d_X)$  to  $\mathcal{R}(Y, d_Y)$ ;*
- (2) *the correspondence  $C$  is  $\varepsilon$ -simplicial from  $\mathcal{C}(X, d_X)$  to  $\mathcal{C}(Y, d_Y)$ .*

*Proof.* Let  $C : X \rightrightarrows Y$  be a correspondence with distortion at most  $\varepsilon$ .

- (1) If  $\sigma$  is a simplex of  $\mathcal{R}(X, d_X)_r$ , then  $d_X(x, x') \leq r$  for all  $x, x'$  in  $\sigma$ . Let  $\tau$  be any subset of  $C(\sigma)$ : for any  $y, y'$  in  $\tau$ , there exist  $x$  and  $x'$  in  $\sigma$  such that  $y \in C(x)$  and  $y' \in C(x')$ , and therefore:

$$d_Y(y, y') \leq d_X(x, x') + \varepsilon \leq r + \varepsilon.$$

Therefore  $\tau$  is a simplex of  $\mathcal{R}(Y, d_Y)_{r+\varepsilon}$ . We have thus shown that  $C$  is  $\varepsilon$ -simplicial from  $\mathcal{R}(X, d_X)$  to  $\mathcal{R}(Y, d_Y)$ .

- (2) Let  $\sigma$  be a simplex of  $\check{\mathcal{C}}(X, d_X)_r$ , and let  $\bar{x}$  be an  $r$ -center of  $\sigma$ , so, for all  $x$  in  $\sigma$ , we have  $d_X(x, \bar{x}) \leq r$ . Take an element  $\bar{y}$  in  $C(\bar{x})$ . For any  $y$  in  $C(\sigma)$ , we have  $y$  in  $C(x)$  for some  $x$  in  $\sigma$ , and therefore

$$d_Y(\bar{y}, y) \leq d_X(\bar{x}, x) + \varepsilon \leq r + \varepsilon ,$$

Let  $\tau$  be a subset of  $C(\sigma)$ ;  $\bar{y}$  is an  $(r + \varepsilon)$ -center for  $\tau$  and hence  $\tau$  is a simplex of  $\mathcal{C}(Y, d_Y)_{r+\varepsilon}$ . We have thus shown that  $C$  is  $\varepsilon$ -simplicial from  $\check{\mathcal{C}}(X, d_X)$  to  $\check{\mathcal{C}}(Y, d_Y)$ . □

**Theorem 53** (Stability theorem - Associative version). *Let  $(X, d_X)$  and  $(Y, d_Y)$  be two metric spaces. We have the following inequalities:*

- (1) in the homotopy category  $\text{ho}(\text{Alg}_{\mathcal{A}S})$ :

$$d_{\text{ho}(\text{Alg}_{\mathcal{A}S})}(C^*(\mathcal{R}(X, d_X)), C^*(\mathcal{R}(Y, d_Y))) \leq 2d_{\mathcal{GH}}((X, d_X), (Y, d_Y)) ;$$

$$d_{\text{ho}(\text{Alg}_{\mathcal{A}S})}(C^*(\check{\mathcal{C}}(X, d_X)), C^*(\check{\mathcal{C}}(Y, d_Y))) \leq 2d_{\mathcal{GH}}((X, d_X), (Y, d_Y)) ;$$

- (2) in the strict category  $\text{Alg}_{\mathcal{A}S}$  :

$$d_{\mathcal{A}S}(\mathcal{R}(X, d_X), \mathcal{R}(Y, d_Y)) \leq 2d_{\mathcal{GH}}((X, d_X), (Y, d_Y)) ;$$

$$d_{\mathcal{A}S}(\check{\mathcal{C}}(X, d_X), \check{\mathcal{C}}(Y, d_Y)) \leq 2d_{\mathcal{GH}}((X, d_X), (Y, d_Y)) .$$

Recall (Remark 7) that in the inequalities (2), we are considering the associative algebra structures given by the cup product on the *cohomology* algebras  $H^*(\mathcal{R}(X, d_X))$ ,  $H^*(\mathcal{R}(Y, d_Y))$  (and not algebra structures at the cochain level).

*Proof.* Let  $X$  and  $Y$  be two metric spaces and let  $C : X \rightrightarrows Y$  be a correspondence with distortion at most  $\varepsilon$ . By Lemma 52,  $C$  is  $\varepsilon$ -simplicial from  $\mathcal{R}(X, d_X)$  to  $\mathcal{R}(Y, d_Y)$ . Thus, by Lemma 50, any two subordinate maps  $f, g : X \xrightarrow{C} Y$  induce  $\varepsilon$ -morphisms of copersistent differential graded algebras  $C^*(\mathcal{R}(X, d_X)) \rightarrow C^*(\mathcal{R}(Y, d_Y))$  which are homotopic in the category  $\text{Alg}_{\mathcal{A}S}$ . Therefore, the correspondence  $C$  induces a  $\varepsilon$ -morphism  $\varphi : C^*(\mathcal{R}(X, d_X)) \rightarrow C^*(\mathcal{R}(Y, d_Y))$  in the homotopy category  $\text{ho}(\text{Alg}_{\mathcal{A}S})$  thanks to Proposition 51.

By the same argument, the correspondence  $C^T : Y \rightrightarrows X$  gives us an  $\varepsilon$ -morphism  $\psi : C^*(\mathcal{R}(Y, d_Y)) \rightarrow C^*(\mathcal{R}(X, d_X))$ . The correspondence  $C^T \circ C$  (respectively  $C \circ C^T$ ) gives us the  $2\varepsilon$ -morphism  $\psi \circ \varphi$  (resp.  $\varphi \circ \psi$ ), which is the canonical  $2\varepsilon$ -endomorphism of  $C^*(\mathcal{R}(X, d_X))$  (resp.  $C^*(\mathcal{R}(Y, d_Y))$ ) given by the structural morphisms of  $C^*(\mathcal{R}(X, d_X))$  (resp.  $C^*(\mathcal{R}(Y, d_Y))$ ). Therefore the  $\varepsilon$ -morphisms  $\varphi$  and  $\psi$  define a  $\varepsilon$ -interleaving between  $C^*(\mathcal{R}(X, d_X))$  and  $C^*(\mathcal{R}(Y, d_Y))$  .

The same argument applies *verbatim* for all other cases. □

**Corollary 54** (Stability theorem -  $\mathcal{A}_\infty$  version). *Let  $X$  and  $Y$  be two finite set of points of  $\mathbb{R}^n$ . We have the following inequalities:*

$$d_{\mathcal{A}_\infty}(\mathcal{R}(X), \mathcal{R}(Y)) \leq 2d_{\mathcal{GH}}(X, Y) ;$$

$$d_{\mathcal{A}_\infty}(\check{\mathcal{C}}(X), \check{\mathcal{C}}(Y)) \leq 2d_{\mathcal{GH}}(X, Y) .$$

*Proof.* It follows from Proposition 41 (2) and Theorem 53. □

#### 4. HOMOTOPY COMMUTATIVITY PRESERVING DISTANCES

In this section we also consider the homotopy *commutativity* of the cup-product in cohomology. The cochain algebra is not commutative on the nose though the cohomology is commutative (in the graded sense). Indeed, there are structures (encoded in higher homotopies) yielding the commutativity of the cup-product after passing to cohomology. Thus commutativity is *additional structures* on the cochains (or their cohomology) which can be used to distinguish homotopy types. We start by studying distance associated to the most homotopical additional structure in Section 4.1 before moving to more tractable but useful ones in Section 4.2.

In this section, whenever needed, we adopt the operadic language (see [LV12, Section 5.2] for the definition of an (algebraic symmetric) operad and the definition of an algebra over an operad). However, we try to make the statement and constructions understandable without knowledge of operadic methods as much as we can.

**4.1. A theoretical construction:  $\mathcal{E}_\infty$ -structures and  $\mathcal{E}_\infty$ -distances.** In this section, we introduce a new distance, based on the  $\mathcal{E}_\infty$ -algebra structure of cochain complexes, which dominate all the other ones one can build on persistence cohomology associated to a space (or data set); see Remark 59. The  $\mathcal{E}_\infty$ -algebra structures are (differential graded) *homotopy commutative* and associative structures which are *functorially* carried by cochain complexes associated to spaces or more generally simplicial sets or complexes. We first need the following standard construction.

**Definition 55** (Normalized (co)chain complex). Let  $X$  be a simplicial set. The *normalized chain complex*  $N_*(X)$  is the quotient of the dg-module  $C_*(X)$  by the degeneracies:

$$N_d(X) = \frac{C_d(X)}{s_0 C_{d-1}(X) + \dots + s_{d-1} C_{d-1}(X)}.$$

We consider also the dual cochain complex  $N^*(X) = \text{Hom}_k(N_*(X), k)$ . If  $Y$  is a topological space, we define  $N_*(Y) := N_*(\text{Sing}_\bullet(Y))$  to be the normalized complex of the singular simplicial set associated to  $Y$  (see Section 1.1.2).

**Remark 56.** The normalized (co)chain complexes  $N^*(-)$  and standard simplicial (co)chain complexes  $C^*(-)$  functors are canonically quasi-isomorphic [Wei94]. In particular, for any topological space the canonical map  $N^*(Y) \rightarrow C^*(Y)$  to the singular cochains of  $Y$  is a natural quasi-isomorphism, i.e., induces an isomorphism in cohomology. Further, if  $X$  is a simplicial complex, then the natural cochain complex associated to  $X$  is isomorphic to  $N^*(X)$ , the normalized complex of the simplicial sets associated to  $X$  viewed as a simplicial sets, that is where we have added all degeneracies freely. Said in simpler terms, the normalized cochain complex precisely computes the cochains of a simplicial complex.

The data of an  $\mathcal{E}_\infty$ -algebra involves infinitely many homotopies and there are several equivalent models (meaning models which yields the same homotopy categories of  $\mathcal{E}_\infty$ -algebras) for them. We refer to [Man02, Man06] for details on their homotopy theories. A nice explicit and combinatorial model<sup>1</sup> for  $\mathcal{E}_\infty$ -algebras was given in [BF04]. It is an explicit operad, called the *surjection operad*, which we denote by  $\mathcal{E}_\infty$  in this paper, which is a cofibrant resolution of the commutative operad  $\mathcal{C}om$ . The important point is that this operad encodes the structure of associative commutative product up to homotopy and in particular the category of algebras over the surjection operads models  $\mathcal{E}_\infty$ -algebras.

---

<sup>1</sup>Other popular models are given by the algebras over the Barrat-Eccles operad or the linear isometry operad

**Theorem 57** (see [BF04, Theorem 2.1.1.]). *For any simplicial set  $X$ , we have evaluation products  $\mathcal{E}_\infty(r) \otimes N^*(X)^{\otimes r} \rightarrow N^*(X)$ , functorial in  $X$ , which give the normalized cochain complex  $N^*(X)$  the structure of an  $\mathcal{E}_\infty$ -algebra. In particular, the classical cup-product of cochains is an operation  $\mu_0 : N^*(X)^{\otimes 2} \rightarrow N^*(X)$  associated to an element  $\mu_0$  in  $\mathcal{E}_\infty(2)_0$ .*

**Remark 58.** The above theorem (and Definition 55) are valid with any coefficient ring; in particular over  $\mathbb{Z}$ ,  $\mathbb{Q}$  or  $\mathbb{F}_p$ .

**Remark 59.** By [Man06, Main Theorem], we know that the  $\mathcal{E}_\infty$ -structure on the cochain complex of a topological space  $X$  (under some finiteness and nilpotence assumptions) is a faithful invariant of the homotopy type of  $X$  and essentially encodes it.

**Remark 60** (see [BF04, Section 1.1.1.]). As we have a factorisation of operad morphisms

$$\mathcal{A}s \longrightarrow \mathcal{E}_\infty \longrightarrow \mathcal{C}om ,$$

we have a forgetful functor  $\text{Alg}_{\mathcal{E}_\infty} \xrightarrow{\text{forget}} \text{Alg}_{\mathcal{A}s}$ . Applied to the normalized cochain complex, this functor recovers the cup-product structure, that is, the usual differential graded algebra structure on cochain.

**Definition 61.** We denote  $N_{\mathcal{E}_\infty}^* : \text{Top}^{\text{op}} \rightarrow \text{Alg}_{\mathcal{E}_\infty}$  the functor induced by Theorem 57 and call it the *cochain  $\mathcal{E}_\infty$ -algebra* functor. We denote in the same way its composition with the canonical functor  $\text{Alg}_{\mathcal{E}_\infty} \rightarrow \text{ho}(\text{Alg}_{\mathcal{E}_\infty})$ .

As for dg-associative algebras in Section 1, we will only consider the homotopy category of  $\mathcal{E}_\infty$ -algebras.

**Remark 62.** The Remark 60 implies that the composition of functors

$$\text{Top}^{\text{op}} \xrightarrow{N_{\mathcal{E}_\infty}^*} \text{Alg}_{\mathcal{E}_\infty} \xrightarrow{\text{forget}} \text{Alg}_{\mathcal{A}s}$$

is equal to  $N_{\mathcal{A}s}^*$ .

The functoriality of the  $\mathcal{E}_\infty$ -structure on the normalized cochain complex given by Theorem 57 justifies the following refined interleaving distance.

**Definition 63** ( $\mathcal{E}_\infty$ -interleaving distance). Let  $X_\bullet, Y_\bullet : \mathbb{R} \rightarrow \text{Top}$  be two persistent spaces. The  $\mathcal{E}_\infty$  *interleaving distance* is defined by

$$d_{\mathcal{E}_\infty}(X, Y) := d_{\text{ho}(\text{Alg}_{\mathcal{E}_\infty})}(N_{\mathcal{E}_\infty}^*(X), N_{\mathcal{E}_\infty}^*(Y))$$

where the right hand side is the interleaving distance in the *homotopy category of  $\mathcal{E}_\infty$ -algebras*.

**Remark 64.** The  $\mathcal{E}$ -interleaving distance depends on the choice of the ground field  $k$ . In particular, as we will see, it behaves very differently in characteristic 0 than in characteristic  $p$ . When we need to be explicit on the ground ring, we will use the notation

$$d_{\mathcal{E}_\infty, k}(X, Y) := d_{\text{ho}(\text{Alg}_{\mathcal{E}_\infty})}(N_{\mathcal{E}_\infty}^*(X, k), N_{\mathcal{E}_\infty}^*(Y, k))$$

for the  $\mathcal{E}$ -interleaving distance computed with coefficient in  $k$ .

The  $\mathcal{E}_\infty$ -interleaving distance is the more refined distance we can put on the persistence cochain complex of a space (or simplicial set). Indeed every other ones we consider are smaller, see Theorem 83.

**Theorem 65** (Stability theorem -  $\mathcal{E}_\infty$  version). *Let  $(X, d_X)$  and  $(Y, d_Y)$  be two metric spaces. We have the following inequalities:*

$$\begin{aligned} d_{\mathcal{E}_\infty}(\mathcal{R}(X, d_X), \mathcal{R}(Y, d_Y)) &\leq 2d_{\mathcal{E}\mathcal{H}}((X, d_X), (Y, d_Y)) ; \\ d_{\mathcal{E}_\infty}(\check{\mathcal{C}}(X, d_X), \check{\mathcal{C}}(Y, d_Y)) &\leq 2d_{\mathcal{E}\mathcal{H}}((X, d_X), (Y, d_Y)) . \end{aligned}$$

*Proof.* It is the same proof that for Theorem 53 using the analogue (see Proposition 66) of Proposition 51. In positive characteristic the proposition follows from [Man02, Proposition 4.2 and 4.3] and this result is extended over  $\mathbb{Z}$  (and therefore any coefficient ring) in [Man06, Section 1]; the case of  $\mathbb{N}^*$  follows from that and the main construction of [BF04]. Here notions of homotopy of algebras are with respect to the standard model structures, see [Man06] for instance.  $\square$

**Proposition 66.** *The functor  $C^* : \text{Top}^{\text{op}} \rightarrow \text{Alg}_{\mathcal{E}_\infty}$  (respectively  $N_{\mathcal{E}_\infty}^* : \text{sSet}^{\text{op}} \rightarrow \text{Alg}_{\mathcal{E}_\infty}$ ) converts weak homotopy equivalences of spaces (resp. simplicial sets) to quasi-isomorphisms and homotopy classes of maps (resp. simplicial sets morphisms) to homotopy classes of morphisms of  $\mathcal{E}_\infty$ -algebras.*

Despite being the most interesting distance from a purely theoretical point of view, the  $\mathcal{E}_\infty$ -interleaving distance is not really easily computable for the moment, therefore we now introduce coarsest ones, which are more computer friendly.

**4.2. Positive characteristic and the Steenrod interleaving distance.** In this subsection, we fix  $k = \mathbb{F}_p$ , where  $p$  is a prime.

**Definition 67** (Steenrod algebra  $\mathcal{A}_p$  (see [Bau06, Section 1.1])). Let  $p$  be a prime number. The *mod  $p$ -Steenrod algebra*, denoted by  $\mathcal{A}_p$ , is the graded commutative algebra over  $\mathbb{F}_p$  which is

- for  $p = 2$ , generated by elements denoted  $\text{Sq}^n$  and called the *Steenrod squares*, for  $n \geq 1$ , with cohomological degree  $n$ ;
- for  $p > 2$ , generated by elements denoted  $\beta$ , called the *Bockstein*, of degree 1, and  $P^n$  for  $n \geq 1$  of degree  $2n(p-1)$ ;

whose product satisfy the following relations, called the *Ádem relations*:

- for  $p = 2$  and for  $0 < h < 2k$ , then

$$\text{Sq}^h \text{Sq}^k = \sum_{i=0}^{\lfloor \frac{h}{2} \rfloor} \binom{k-i-1}{h-2i} \text{Sq}^{h+k-i} \text{Sq}^i,$$

- for  $p > 2$  and for  $0 < h < pk$ , then

$$P^h P^k = \sum_{i=0}^{\lfloor \frac{h}{p} \rfloor} (-1)^{h+i} \binom{(p-1)(k-i)-1}{h-pi} P^{h+k-i} P^i,$$

and

$$P^h \beta P^k = \sum_{i=0}^{\lfloor \frac{h}{p} \rfloor} (-1)^{h+i} \binom{(p-1)(k-i)}{h-pi} \beta P^{h+k-i} P^i + \sum_{i=0}^{\lfloor \frac{h-1}{p} \rfloor} (-1)^{h+i-1} \binom{(p-1)(k-i)-1}{h-pi-1} P^{h+k-i} \beta P^i$$

We denote by  $\mathcal{A}_p\text{-Alg}$ , the category of commutative algebras over  $\mathcal{A}_p$ .

**Remark 68.** The Steenrod algebra also have a structure of Hopf algebra. Further, an important formula is given, for  $x$  in  $H^*(x)$ , by

$$\text{Sq}^{|x|}(x) = x \cup x.$$

The Steenrod algebra is the algebra of cohomological operations, i.e. all the natural transformations of degree  $d$  for all  $d$  in  $\mathbb{N}$ :

$$H^*(-, \mathbb{F}_p) \longrightarrow H^{*+d}(-, \mathbb{F}_p),$$

as illustrated by the following theorem.

**Theorem 69** (Steenrod, Ádem). *Let  $X$  be a topological space. The singular cohomology with coefficient in  $\mathbb{F}_p$  is a commutative algebra over the  $\mathcal{A}_p$  Steenrod algebra, so we have the functor*

$$H_{\mathcal{A}_p}^*(-, \mathbb{F}_p) : \text{Top}^{\text{op}} \longrightarrow \mathcal{A}_p\text{-Alg.}$$

**Remark 70.** The Steenrod algebra operations are the trace on the cohomology of the non-commutativity of the cup-product. In particular, they are determined by the  $\mathcal{C}_{\infty}$ -structure on the cochains.

Let us now define a new interleaving distance in positive characteristic.

**Definition 71.** Let  $X$  and  $Y$  be two persistent spaces and let  $p$  be a prime. The  $\mathcal{A}_p$ -interleaving distance is defined by

$$d_{\mathcal{A}_p}(X, Y) := d_{\mathcal{A}_p\text{-Alg}}(H^*(X, \mathbb{F}_p), H^*(Y, \mathbb{F}_p))$$

that it is the interleaving distance (see Definition 4) computed in the category of  $\mathcal{A}_p$ -algebras. Then we do not wish to specify a particular  $p$ , we will simply refer to this distance as *Steenrod interleaving distance*.

We have the following commutative diagram of functors

$$(2) \quad \begin{array}{ccccccc} & & \text{Top}^{\text{op}} & & & & \\ & \swarrow N_{\mathcal{C}_{\infty}}^* & \downarrow & \searrow N_{\mathcal{A}_p}^* & & & \\ & \text{Alg}_{\mathcal{C}_{\infty}} & H_{\mathcal{C}_{\infty}}^* & \text{Alg}_{\mathcal{A}_p} & \xrightarrow{\text{For}} & \text{Ch}_k & \\ \downarrow H_{\mathcal{A}_p}^* & \searrow H^* & \downarrow & \downarrow H & & \downarrow H & \\ \mathcal{A}_p\text{-Alg} & \xrightarrow{\text{For}} & \text{Alg}_{\mathcal{C}_{\infty}} & \xrightarrow{\text{For}} & \text{Alg}_{\mathcal{A}_p} & \xrightarrow{\text{For}} & \text{gr-Vect} \end{array}$$

Using Lemma 8 in characteristic  $p$ , we will deduce the following inequalities between the distances.

**Proposition 72.** *Let  $X, Y : \mathbb{R} \rightarrow \text{Top}$  be two persistent spaces. We have*

$$d_{\text{Alg}_{\mathcal{C}_{\infty}}} (H^*(X, \mathbb{F}_p), H^*(Y, \mathbb{F}_p)) \leq d_{\mathcal{A}_p}(X, Y) \leq d_{\mathcal{C}_{\infty}, \mathbb{F}_p}(X, Y)$$

and for a commutative ring  $k$  in any characteristic

$$d_{\text{Alg}_{\mathcal{C}_{\infty}}} (H^*(X, k), H^*(Y, k)) \leq d_{\text{ho}(\text{Alg}_{\mathcal{A}_p})} (C^*(X, k), C^*(Y, k)) \leq d_{\mathcal{C}_{\infty}, k}(X, Y) .$$

*Proof.* Propositions 51 and 66 implies that the functors  $N_{\mathcal{C}_{\infty}}^*$  and  $N_{\mathcal{A}_p}^*$  passes to the homotopy category and thus diagram (2) induces a commutative diagram:

$$(3) \quad \begin{array}{ccccccc} & & \text{Top}^{\text{op}} & & & & \\ & \swarrow N_{\mathcal{C}_{\infty}}^* & \downarrow & \searrow N_{\mathcal{A}_p}^* & & & \\ & \text{ho}(\text{Alg}_{\mathcal{C}_{\infty}}) & H_{\mathcal{C}_{\infty}}^* & \text{ho}(\text{Alg}_{\mathcal{A}_p}) & \xrightarrow{\text{For}} & \text{ho}(\text{Ch}_k) & \\ \downarrow H_{\mathcal{A}_p}^* & \searrow H^* & \downarrow & \downarrow H & & \downarrow H & \\ \mathcal{A}_p\text{-Alg} & \xrightarrow{\text{For}} & \text{Alg}_{\mathcal{C}_{\infty}} & \xrightarrow{\text{For}} & \text{Alg}_{\mathcal{A}_p} & \xrightarrow{\text{For}} & \text{gr-Vect} ; \end{array}$$

The two claimed string of inequalities then follows from Lemma 8 applied to diagram (3).  $\square$

**Remark 73** (Effective computability of Steenrod distance). An important practical fact that makes the Steenrod distance appealing is that there exists *algorithms* to compute the persistent Steenrod squares with coefficient in  $\mathbb{F}_2$  (see [Aub11, MM18]).

We also have the following theorem of stability:



**Theorem 74** (Stability theorem - Commutative mod  $p$  version). *Let  $(X, d_X)$  and  $(Y, d_Y)$  be two metric spaces. We have the following inequality:*

$$\begin{aligned} d_{\mathcal{A}_p}(\mathcal{R}(X, d_X), \mathcal{R}(Y, d_Y)) &\leq 2d_{\mathcal{E}\mathcal{H}}((X, d_X), (Y, d_Y)) ; \\ d_{\mathcal{A}_p}(\check{\mathcal{C}}(X, d_X), \check{\mathcal{C}}(Y, d_Y)) &\leq 2d_{\mathcal{E}\mathcal{H}}((X, d_X), (Y, d_Y)) . \end{aligned}$$

*Proof.* It follows from Theorem 65 and the first inequality in Proposition 72.  $\square$

**Example 75.** Consider the compact spaces  $X = S^3 \vee S^5$  and  $Y = \Sigma CP^2$ , the suspension of  $CP^2$ . The spaces  $X$  and  $Y$  have the same cohomology as a graded vector space. As  $H^*(X)$  is a sub-algebra of  $H^*(S^3) \times H^*(S^5)$ , and  $Y$  is a suspension, both spaces have trivial cup products in cohomology. In particular, their cohomology are the same as associative algebras.

But, if we consider their cohomology with coefficient in  $\mathbb{F}_2$  as module under the Steenrod algebra, then  $H^*(X)$  has only trivial Steenrod square while  $H^*(Y)$  has non-trivial ones. Note also that this difference is also detected by the  $\mathcal{E}_\infty$ -structure on the cochain level. Therefore, we can proceed as in Example 19. We can embed  $X$  and  $Y$  in  $\mathbb{R}^7$  and take discretisations of  $X$  and  $Y$ , depending of a small parameter  $\alpha \ll 1$ . We denote by  $\tilde{X}$  and  $\tilde{Y}$  respectively those discretisation, which are thus finite sets. We then get

$$d_{\text{Alg}_{\mathcal{E}_\infty}}(H^*(\mathcal{R}(\tilde{X})), H^*(\mathcal{R}(\tilde{Y}))) < d_{\mathcal{A}_p}(\mathcal{R}(\tilde{X}), \mathcal{R}(\tilde{Y})).$$

Similarly, Proposition 72 also implies

$$d_{\text{hoAlg}_{\mathcal{E}_\infty}}(N^*(\mathcal{R}(\tilde{X})), N^*(\mathcal{R}(\tilde{Y}))) < d_{\mathcal{E}_\infty}(\mathcal{R}(\tilde{X}), \mathcal{R}(\tilde{Y})) .$$

**Example 76.** Let us return again to Example 43, that is the complements of  $\beta$ -thickenings of the borromean links and trivial entanglements of three circles in  $S^3$ . Since the cohomology of these spaces is concentrated in degree less than 2, we obtain that the only non-zero Steenrod squares are given by  $Sq^1 : H^1(X) \rightarrow H^2(X)$ . The latter is given by the self cup-product  $Sq^1(x) = x \cup x$ , see Definition 67 and Remark 68. In other words, the Steenrod squares of  $X$  and  $Y$  are the same. Similarly, the other Steenrod powers coincide for  $X$  and  $Y$ . In particular, for the discretisation  $\tilde{X}$  and  $\tilde{Y}$  we obtain

$$d_{\mathcal{A}_p}(\mathcal{R}(\tilde{X}), \mathcal{R}(\tilde{Y})) < d_{\mathcal{A}_\infty}(\mathcal{R}(\tilde{X}), \mathcal{R}(\tilde{Y})).$$

**4.3. Distances in characteristic zero.** In this section, we fix  $k$  a field of characteristic 0. In this case, the homotopy theory of  $\mathcal{E}_\infty$ -algebras is equivalent to that of algebras over another operad which encodes the structure of commutative associative algebra up to homotopy, called  $\mathcal{E}_{\text{om}_\infty}$ , given by given by Koszul duality (see [LV12]). This operad is smaller than the operad of surjection (Theorem 57). One of the advantage of  $\mathcal{E}_{\text{om}_\infty}$  is that we have a homotopy transfer theorem as for  $\mathcal{A}_\infty$  in Section 2. Therefore, as in Section 2.3, we have the functor

$$\begin{aligned} H_* : \text{fTrans}_{\text{coCh}, \mathcal{E}_{\text{om}}}^{\mathbb{R}^{\text{op}}} &\longrightarrow \text{ho}(\infty\text{-Alg}_{\mathcal{E}_{\text{om}_\infty}})^{\mathbb{R}^{\text{op}}} \\ (A, H, i, p, h)^\bullet &\longmapsto (H, \{\mu_i\}_{i \in \mathbb{N}})^\bullet \end{aligned}$$

In particular, given  $X_\bullet, Y_\bullet : \mathbb{R} \rightarrow \Delta\text{Cpx}$  two finite filtered data, we can define their  $\mathcal{E}_{\text{om}_\infty}$ -interleaving distance:

$$d_{\text{ho}(\infty\text{-Alg}_{\mathcal{E}_{\text{om}_\infty}})}(H^*(X), H^*(Y)) .$$

The structure of  $\mathcal{E}_{\text{om}_\infty}$ -algebra encodes more homotopy structure than that of  $\mathcal{A}_\infty$ -algebra. However, by Lemma 77, this new distance does not necessarily permit to differentiate more persistent spaces than the  $\mathcal{A}_\infty$  one.

**Lemma 77.** Let  $X_\bullet, Y_\bullet : \mathbb{R} \rightarrow \Delta\text{Cpx}$  be two finite filtered data. We have

$$d_{\text{ho}(\infty\text{-Alg}_{\mathcal{E}_{\text{om}\infty})}(H^*(X), H^*(Y)) = 0 \iff d_{\mathcal{A}_\infty}(X, Y) = 0 .$$

*Proof.* Let  $X_\bullet, Y_\bullet : \mathbb{R} \rightarrow \Delta\text{Cpx}$  be two finite filtered data such that  $d_{\text{ho}(\infty\text{-Alg}_{\mathcal{E}_{\text{om}\infty})}(H^*(X), H^*(Y)) = 0$ . As in Theorem 30, by [LV12, Theorem 11.4.8], the homotopy category of differential graded commutative associative algebras and the homotopy category of  $\mathcal{E}_{\text{om}\infty}$ -algebras with the  $\infty$ -morphisms are equivalent. Therefore, by Lemma 33,  $H^*(X)$  and  $H^*(Y)$  are quasi-isomorphic as homotopy commutative algebras. By [CPRNW19, Theorem A], this is equivalent to the fact that  $H^*(X)$  and  $H^*(Y)$  are quasi-isomorphic in the homotopy category of dg-associative algebras. Then by Theorem 30 and Lemma 33

$$d_{\text{ho}(\infty\text{-Alg}_{\mathcal{E}_{\text{om}\infty})}(H^*(X), H^*(Y)) = 0 \iff d_{\mathcal{A}_\infty}(X, Y) = 0 .$$

□

**Remark 78.** Since a  $\mathcal{E}_{\text{om}\infty}$ -algebra structure has a canonical underlying  $\mathcal{A}_\infty$ -algebra structure, we automatically have an inequality

$$d_{\text{ho}(\infty\text{-Alg}_{\mathcal{E}_{\text{om}\infty})}(H^*(X), H^*(Y)) \geq d_{\mathcal{A}_\infty}(X, Y)$$

where the left hand side is the interleaving distance in the category of  $\mathcal{E}_{\text{om}\infty}$ -algebras. The Lemma 77 suggests that the inequality might actually be an equality in many practical cases.

**Remark 79** (The functor  $A_{PL}$ ). In characteristic zero, there is a functor

$$A_{PL} : \text{Top} \longrightarrow \text{Alg}_{\mathcal{E}_{\text{om}}} ,$$

(see [FHT12] for the definition) such that, for any topological space  $X$ , there exists two natural quasi-isomorphisms of differential graded associative algebras

$$C^*(X) \xrightarrow{\sim} \bullet \xleftarrow{\sim} A_{PL}(X)$$

(see [FHT12, Corollary 10.10]). As the functor  $A_{PL}$  (or the equivalent combinatorial model given by Felix *et al.* in [FJP09]) is more computable, we expect that these constructions can be used to compute the  $\mathcal{A}_\infty$ -interleaving distance for finite filtered data (see Proposition 41).

**4.4. The best of both worlds.** We have seen in Proposition 72 that the  $\mathcal{E}_\infty$ -interleaving distance is one of the finer distances that we can define, but this distance seems difficult to calculate because of the intricate structure of an algebra of the operad of surjection and its homotopy category. As the operad  $\mathcal{E}_\infty$  encodes in particular the cup-product and the Steenrod operations we can restrict to the following distances.

**Definition/Proposition 80.** Let  $X_\bullet, Y_\bullet : \mathbb{R} \rightarrow \Delta\text{Cpx}$  be two finite filtered data and let  $p$  be a prime number. We define two distances given by

$$\begin{aligned} d_{p,\infty}(X, Y) &:= \max(d_{\mathcal{A}_\infty}(X, Y), d_{\mathcal{A}_p}(X, Y)) , \\ d_{\mathbb{P},\infty}(X, Y) &:= \max(d_{\mathcal{A}_\infty}(X, Y), \sup_{p \in \mathbb{P}} d_{\mathcal{A}_p}(X, Y)) , \end{aligned}$$

where  $\mathbb{P}$  is the set of prime numbers. By definition we have

$$d_{p,\infty}(X, Y) \leq d_{\mathbb{P},\infty}(X, Y) .$$

**Example 81.** Fix  $p$  an odd prime and consider the real projective plane  $\mathbb{R}P^2$ : it has the following cohomology

$$H^0(\mathbb{R}P^2, \mathbb{Z}) = \mathbb{Z} , H^1(\mathbb{R}P^2, \mathbb{Z}) = 0 , H^2(\mathbb{R}P^2, \mathbb{Z}) = \mathbb{Z}/2\mathbb{Z} ,$$

and

$$H^0(\mathbb{R}P^2, \mathbb{Z}/2\mathbb{Z}) = \mathbb{Z}/2\mathbb{Z}, \quad H^1(\mathbb{R}P^2, \mathbb{Z}/2\mathbb{Z}) = \mathbb{Z}/2\mathbb{Z}, \quad H^2(\mathbb{R}P^2, \mathbb{Z}/2\mathbb{Z}) = \mathbb{Z}/2\mathbb{Z}.$$

Let  $X = \Sigma\mathbb{R}P^2$  be the suspension of  $\mathbb{R}P^2$  and  $Y$  be a closed ball in  $\mathbb{R}^3$ . The spaces  $X$  and  $Y$  have trivial cup product, but the real projective space has non trivial Steenrod square, and its suspension too. As in Example 19, we can construct two discretisations  $\tilde{X}$  and  $\tilde{Y}$  of  $X$  and  $Y$  respectively, depending of a parameter  $\alpha$  such that

$$\begin{aligned} d_{p,\infty}(\check{\mathcal{C}}(\tilde{X}), \check{\mathcal{C}}(\tilde{Y})) &< d_{p,\infty}(\check{\mathcal{C}}(\tilde{X}), \check{\mathcal{C}}(\tilde{Y})); \\ d_{p,\infty}(\mathcal{R}(\tilde{X}), \mathcal{R}(\tilde{Y})) &< d_{p,\infty}(\mathcal{R}(\tilde{X}), \mathcal{R}(\tilde{Y})). \end{aligned}$$

Similar examples can be obtained in other characteristic using lens spaces in place of projective spaces.

**Remark 82.** For many practical applications, it seems that the  $\mathcal{A}_\infty$ , the  $\mathcal{A}_2$  and the  $\mathcal{A}_3$  interleaving distances will be useful: in fact, the spaces which are not differentiated by these distances but which are differentiated by other refined ones will be complicated to compute algorithmically, at least for the moment.

## 5. RÉSUMÉ AND PROOF OF THEOREM B

In this paper we have defined several distances between finite filtered data, using the cohomology of their associated persistent spaces. In this section we finish to compare them all and in particular prove Theorem B. We fix  $p$  a prime and we recall that the  $\mathcal{A}_\infty$  interleaving distance is defined for the cohomology of persistent data with coefficient in a field of characteristic zero. Further we consider the  $\mathcal{E}_\infty$ -interleaving distance of Section 4.1 with value in  $\mathbb{Z}$  for coefficient. We can summarize all the interleaving distances constructed in this paper in the following diagram of distances.

**Theorem 83.** *Consider the various interleaving distances introduced in the paper.*

(1) *There is a string of inequalities*

$$\begin{array}{ccccccc} & & & & d_{\mathcal{A}_\infty} & \geq & d_{\mathcal{A}_2, \mathbb{Q}} & \geq & d_{\text{gr-Vect}, \mathbb{Q}} \\ & & & & \geq & (2) & & (1) & \\ & & & & \geq & (4) & & & \\ d_{\mathbb{P}, \infty} & \geq & d_{p, \infty} & & & & & & \\ & (6) & & & & & & & \\ & & & & \geq & (5) & & & \\ & & & & d_{\mathcal{A}_p} & \geq & d_{\mathcal{A}_2, \mathbb{F}_p} & \geq & d_{\text{gr-Vect}, \mathbb{F}_p} \\ & & & & & (3) & & (1) & \end{array}$$

for finite filtered data (Definition 31). None of these inequalities are equalities in general.

(2) *More generally, for arbitrary persistent spaces, there is a string of inequalities*

$$\begin{array}{ccccccc} & & & & d_{\text{hoAlg}_{\mathcal{A}_2}} & \geq & d_{\mathcal{A}_2, \mathbb{Q}} & \geq & d_{\text{gr-Vect}, \mathbb{Q}} \\ & & & & \geq & (2) & & (1) & \\ & & & & \geq & (4) & & & \\ d_{\mathcal{E}_\infty} & \geq & d_{\mathbb{P}, \infty} & \geq & d_{p, \infty} & & & & \\ & & (6) & & & & & & \\ & & & & \geq & (5) & & & \\ & & & & d_{\mathcal{A}_p} & \geq & d_{\mathcal{A}_2, \mathbb{F}_p} & \geq & d_{\text{gr-Vect}, \mathbb{F}_p} \\ & & & & & (3) & & (1) & \end{array}.$$

Further, these inequalities are not equalities in general.

In particular, those distances are not equal for Rips or Čech complex associated to discretisations of spaces.

**Remark 84.** Note that the Main Theorem B is nothing but a special case of Theorem 83.

*Proof of Theorem 83.* The first string of inequalities as well as the fact that they are strict in general for finite data sets follows from

- (1) Proposition 18 and Example 19
- (2) Proposition 41 and Example 43
- (3) Proposition 72 and Example 75
- (4) forgetting the Steenrod power operations and Example 75
- (5) forgetting the  $\mathcal{A}_\infty$ -structure and Example 76
- (6) Definition/Proposition 80 and Example 81.

The same argument and part (2) of Proposition 41 yields the inequalities of the second part as well as the fact that they are strict in general. By the universal coefficient theorem [Wei94], any  $\varepsilon$ -interleaving between normalized chain complex  $N^*(X, \mathbb{Z})$  and  $N^*(Y, \mathbb{Z})$  induces a  $\varepsilon$ -interleaving between  $N^*(X, k)$  and  $N^*(Y, k)$  for any field  $k$ . This proves the inequality  $d_{\mathbb{Z}, \infty} \geq d_{\mathbb{P}, \infty}$ .

Note that for general persistent spaces the counter-examples are easier to produce than for those arising from Rips or Čech complexes. Indeed, it suffices to take  $X$  and  $Y$  two topological spaces such that  $N^*(X)$  and  $N^*(Y)$  are equivalent in the category  $\mathcal{C}$  but distinct in the category  $\mathcal{D}$  (here the categories are any of the one we consider for interleavings) to obtain two persistent spaces such that  $d_{\mathcal{C}} < d_{\mathcal{D}}$ . Indeed one can consider the constant functions  $f : X \rightarrow \mathbb{R}$  and  $g : Y \rightarrow \mathbb{R}$  that sends the spaces to the point  $0 \in \mathbb{R}$ . Then the persistent spaces associated to the function satisfy the strict inequalities.  $\square$

**Remark 85.** This work is a first theoretical step towards new distances taking into account more topological information in Topological Data Analysis. For the moment, many of these distances are hard to compute. In future work we will study how to define bottleneck distances taking the multiplicative structure into account in the same way, in order to get more computer friendly distances. The existence of a bottleneck distance in the derived category of sheaves [BG18] is an evidence for those.

## REFERENCES

- [Aub11] HB Aubrey. *Persistent Cohomology Operations*. Duke University, 2011. 23
- [Bau06] Hans-Joachim Baues. *The algebra of secondary cohomology operations*. Springer, 2006. 22
- [Bel17] Francisco Belchí. Optimising the topological information of the  $A_\infty$ -persistence groups. *arXiv preprint arXiv:1706.06019*, 2017. 3, 16
- [BF04] Clemens Berger and Benoit Fresse. Combinatorial operad actions on cochains. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 137, pages 135–174. Cambridge University Press, 2004. 20, 21, 22
- [BG18] Nicolas Berkouk and Grégory Ginot. A derived isometry theorem for constructible sheaves on  $\mathbb{R}$ . *arXiv preprint arXiv:1805.09694*, 2018. 27
- [BM15] Francisco Belchí and Aniceto Murillo.  $A_\infty$ -persistence. *Applicable Algebra in Engineering, Communication and Computing*, 26(1-2):121–139, 2015. 3, 16
- [BP19] Nicolas Berkouk and François Petit. Ephemeral persistence modules and distance comparison. *arXiv preprint arXiv:1902.09933*, 2019. 8
- [BS19] Francisco Belchí and Anastasios Stefanou.  $A_\infty$  persistent homology estimates the topology from pointcloud datasets. *arXiv preprint arXiv:1902.09138*, 2019. 3, 16
- [CDSO14] Frédéric Chazal, Vin De Silva, and Steve Oudot. Persistence stability for geometric complexes. *Geometriae Dedicata*, 173(1):193–214, 2014. 2, 17, 18
- [CO08] Frédéric Chazal and Steve Yann Oudot. Towards persistence-based reconstruction in euclidean spaces. In *Proceedings of the twenty-fourth annual symposium on Computational geometry*, pages 232–241. ACM, 2008. 2
- [CPRNW19] Ricardo Campos, Dan Petersen, Daniel Robert-Nicoud, and Felix Wierstra. Lie, associative and commutative quasi-isomorphism. *arXiv preprint arXiv:1904.03585*, 2019. 25
- [CSEH07] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. *Discrete & Computational Geometry*, 37(1):103–120, 2007. 2

- [CZCG05] Gunnar Carlsson, Afra Zomorodian, Anne Collins, and Leonidas J. Guibas. Persistent barcodes for shapes. *International Journal of Shape Modeling*, 11(02):149–187, 2005. 2
- [EH08] Herbert Edelsbrunner and John Harer. Persistent homology—a survey. *Contemporary mathematics*, 453:257–282, 2008. 1
- [FHT12] Yves Félix, Stephen Halperin, and J-C Thomas. *Rational homotopy theory*, volume 205. Springer Science & Business Media, 2012. 25
- [FJP09] Yves Felix, Barry Jessup, and Paul-Eugène Parent. The combinatorial model for the Sullivan functor on simplicial sets. *Journal of Pure and Applied Algebra*, 213(2):231–240, 2009. 25
- [GM14] Francisco Belchí Guillamón and Aniceto Murillo Mas.  $A_\infty$  persistence. *arXiv preprint arXiv:1403.2395*, 2014. 3, 16
- [Hat02] Allen Hatcher. Algebraic topology. 2002. *Cambridge UP, Cambridge*, 606(9), 2002. 12
- [Her18] Estanislao Herscovich. A higher homotopic extension of persistent (co)homology. *J. Homotopy Relat. Struct.*, 13(3):599–633, 2018. 3
- [Jam95] Ioan Mackenzie James. *Handbook of algebraic topology*. Elsevier, 1995. 18
- [JCR<sup>+</sup>17] Rachel Jeitiner, Mathieu Carrière, Jacques Rougemont, Steve Oudot, Kathryn Hess, and Cathrin Brisken. Two-Tier Mapper: a user-independent clustering method for global gene expression analysis based on topology. *arXiv preprint arXiv:1801.01841*, 2017. 2
- [KDS<sup>+</sup>18] Lida Kanari, Paweł Dłotko, Martina Scolamiero, Ran Levi, Julian Shillcock, Kathryn Hess, and Henry Markram. A topological representation of branching neuronal morphologies. *Neuroinformatics*, 16(1):3–13, 2018. 2
- [KS18] Masaki Kashiwara and Pierre Schapira. Persistent homology and microlocal sheaf theory. *Journal of Applied and Computational Topology*, 2018. 8
- [LH02] Kenji Lefevre-Hasegawa. *Sur les  $A_\infty$ -catégories*. PhD thesis, Ph. D. thesis, Université Paris 7, UFR de Mathématiques, 2003, math. CT/0310337, 2002. 12
- [LV12] Jean-Louis Loday and Bruno Vallette. *Algebraic operads*, volume 346. Springer Science & Business Media, 2012. 4, 10, 11, 12, 20, 24, 25
- [Man02] Michael A. Mandell. Equivariant  $p$ -adic homotopy theory. *Topology Appl.*, 122(3):637–651, 2002. 20, 22
- [Man06] Michael A Mandell. Cochains and homotopy type. *Publications Mathématiques de l’Institut des Hautes Études Scientifiques*, 103(1):213–246, 2006. 2, 20, 21, 22
- [McC06] John McCleary. *A first course in topology: Continuity and dimension*, volume 31. American Mathematical Soc., 2006. 18
- [MM18] AM Medina-Mardones. Persistence steenrod modules. *arXiv preprint arXiv:1812.05031*, 2018. 23
- [MS17] Marco Manetti and Chiara Spagnoli. The projective model structure on contractions. *arXiv preprint arXiv:1703.03231*, 2017. 13
- [Oud15a] Steve Y Oudot. *Persistence theory: from quiver representations to data analysis*, volume 209. American Mathematical Society Providence, RI, 2015. 1
- [Oud15b] Steve Y Oudot. *Persistence theory: from quiver representations to data analysis*, volume 209. American Mathematical Society Providence, RI, 2015. 8
- [RMA09] Pedro Real and Helena Molina-Abril. Cell AT-models for digital volumes. In *International Workshop on Graph-Based Representations in Pattern Recognition*, pages 314–323. Springer, 2009. 14
- [RT00] Yuli Rudyak and Aleksy Tralle. On Thom spaces, Massey products, and nonformal symplectic manifolds. *International Mathematics Research Notices*, 2000(10):495–513, 2000. 12
- [Wei94] Charles A. Weibel. *An introduction to homological algebra*, volume 38 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1994. 20, 27

LAGA, UNIVERSITÉ PARIS 13, 99 AVENUE JEAN BAPTISTE CLÉMENT 93430, VILLETANEUSE, FRANCE

*E-mail address:* ginot@math.univ-paris13.fr, leray@math.univ-paris13.fr