



**HAL**  
open science

# Computing persistent Stiefel-Whitney classes of line bundles

Raphaël Tinarrage

► **To cite this version:**

Raphaël Tinarrage. Computing persistent Stiefel-Whitney classes of line bundles. *Journal of Applied and Computational Topology*, 2021, 10.1007/s41468-021-00080-4 . hal-02619607v2

**HAL Id: hal-02619607**

**<https://hal.science/hal-02619607v2>**

Submitted on 11 Nov 2021

**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.

# Computing persistent Stiefel-Whitney classes of line bundles

Raphaël Tinarrage

<https://raphaeltinarrage.github.io/>

**Abstract.** We propose a definition of persistent Stiefel-Whitney classes of vector bundle filtrations. It relies on seeing vector bundles as subsets of some Euclidean spaces. The usual Čech filtration of such a subset can be endowed with a vector bundle structure, that we call a Čech bundle filtration. We show that this construction is stable and consistent. When the dataset is a finite sample of a line bundle, we implement an effective algorithm to compute its first persistent Stiefel-Whitney class. In order to use simplicial approximation techniques in practice, we develop a notion of weak simplicial approximation. As a theoretical example, we give an in-depth study of the normal bundle of the circle, which reduces to understanding the persistent cohomology of the torus knot (1,2). We illustrate our method on several datasets inspired by image analysis.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Background</b>	<b>3</b>
2.1	Stiefel-Whitney classes . . . . .	3
2.2	Simplicial approximation . . . . .	5
2.3	Persistent cohomology . . . . .	7
2.4	Notations . . . . .	8
<b>3</b>	<b>Persistent Stiefel-Whitney classes</b>	<b>9</b>
3.1	Definition . . . . .	9
3.2	Čech bundle filtrations . . . . .	10
3.3	Stability . . . . .	14
3.4	Consistency . . . . .	16
<b>4</b>	<b>Computation of persistent Stiefel-Whitney classes</b>	<b>19</b>
4.1	Simplicial approximation to Čech bundle filtrations . . . . .	19
4.2	A sketch of algorithm . . . . .	21
<b>5</b>	<b>An algorithm when <math>d = 1</math></b>	<b>22</b>
5.1	The star condition in practice . . . . .	22
5.2	Triangulating the projective spaces . . . . .	25
5.3	Vietoris-Rips version of the Čech bundle filtration . . . . .	27
5.4	Choice of the parameter $\gamma$ . . . . .	28
<b>6</b>	<b>Numerical experiments</b>	<b>30</b>
6.1	Datasets with symmetries . . . . .	30
6.2	Datasets with intrinsic (non-)orientability . . . . .	36
<b>7</b>	<b>Conclusion</b>	<b>39</b>
<b>A</b>	<b>Supplementary material for Sect. 5</b>	<b>40</b>
A.1	Study of Example 5.9 . . . . .	40
A.2	Study of Example 5.10 . . . . .	41

# 1 Introduction

The inference of relevant topological properties of data represented as point clouds in Euclidean spaces is a central challenge in Topological Data Analysis (TDA). Given a (finite) set of points  $X$  in  $\mathbb{R}^n$ , persistent homology provides a now classical and powerful tool to construct persistence diagrams whose points can be interpreted as homological features of  $X$  at different scales.

In this work, we aim at developing a similar theoretical framework for another topological invariant: the Stiefel-Whitney classes. These classes, and more generally characteristic classes, are a powerful tool from algebraic topology, that contains additional information to the cohomology groups. For the Stiefel-Whitney classes to be defined, the input topological space has to be endowed with an additional structure: a real vector bundle. They have been widely used in differential topology, for instance in the problem of deciding orientability of manifolds, of immersing manifolds in low-dimensional spaces, or in cobordism problems (Milnor and Stasheff, 2016). Our work is motivated by introducing this tool to the TDA community.

**Previous work.** To our knowledge, the problem of estimating Stiefel-Whitney classes from a point cloud observation has received little attention. In the work of Aubrey (2011), one finds an algorithm to compute the Stiefel-Whitney classes in the particular case of the tangent bundle of a Euler mod-2 space (that is, a simplicial complex for which the link of each simplex has even Euler characteristic). Close to the subject, Perea (2018) proposes a dimensionality reduction algorithm, based on the choice of a Stiefel-Whitney class, seen as a persistent cohomology class.

Recently, Scoccola and Perea (2021) developed several notions of vector bundle adapted to finite simplicial complexes, one of which is used in this paper. They propose algorithms to compute the first two Stiefel-Whitney classes, which are conceptually different than the one presented here.

**Our contributions.** Just as persistent homology allows to extract homological features from filtrations of topological spaces, we propose a framework that allows to extract Stiefel-Whitney classes features from filtrations of vector bundles. It is briefly motivated here.

In general, if  $X$  is a topological space endowed with a vector bundle  $\xi$  of dimension  $d$ , there exists a collection of cohomology classes  $w_1(\xi), \dots, w_d(\xi)$ , the Stiefel-Whitney classes, such that  $w_i(\xi)$  is an element of the cohomology group  $H^i(X)$  over  $\mathbb{Z}_2$  for  $i \in \llbracket 1, d \rrbracket$ . In order to define Stiefel-Whitney classes in a persistent-theoretic framework, we will use a convenient definition of vector bundles: defining a vector bundle over a compact space  $X$  is equivalent (up to isomorphism of vector bundles) to defining a continuous map  $p: X \rightarrow \mathcal{G}_d(\mathbb{R}^m)$  for  $m$  large enough, where  $\mathcal{G}_d(\mathbb{R}^m)$  is the Grassmann manifold of  $d$ -planes in  $\mathbb{R}^m$ . Such a map is called a *classifying map* for  $\xi$ .

Given a classifying map  $p: X \rightarrow \mathcal{G}_d(\mathbb{R}^m)$  of a vector bundle  $\xi$ , the Stiefel-Whitney classes  $w_1(\xi), \dots, w_d(\xi)$  can be defined by pushing forward some particular classes of the Grassmannian via the induced map in cohomology  $p^*: H^*(X) \leftarrow H^*(\mathcal{G}_d(\mathbb{R}^m))$ . If  $w_i$  denotes the  $i^{\text{th}}$  Stiefel-Whitney class of the Grassmannian, then the  $i^{\text{th}}$  Stiefel-Whitney class of the vector bundle  $\xi$  is

$$w_i(\xi) = p^*(w_i). \tag{1}$$

In order to translate these considerations in a persistent-theoretic setting, suppose that we are given a dataset of the form  $(X, p)$ , where  $X$  is a finite subset of  $\mathbb{R}^n$ , and  $p$  is a map  $p: X \rightarrow \mathcal{G}_d(\mathbb{R}^m)$ . Denote by  $(X^t)_{t \geq 0}$  the Čech filtration of  $X$ , that is, the collection of the  $t$ -thickenings  $X^t$  of  $X$  in the ambient space  $\mathbb{R}^n$ . It is also known as the *offset filtration* of  $X$ . In order to define some *persistent* Stiefel-Whitney classes, one would try to extend the map  $p: X \rightarrow \mathcal{G}_d(\mathbb{R}^m)$  to  $p^t: X^t \rightarrow \mathcal{G}_d(\mathbb{R}^m)$ . However, we did not find any interesting way to extend this map. To overcome this issue, we propose to look at the dataset in a different way. Transform the vector bundle  $(X, p)$  into a subset of  $\mathbb{R}^n \times \mathcal{G}_d(\mathbb{R}^m)$  via

$$\tilde{X} = \{(x, p(x)), x \in X\}.$$

The Grassmann manifold  $\mathcal{G}_d(\mathbb{R}^m)$  can be naturally embedded in  $M(\mathbb{R}^m)$ , the space of  $m \times m$  matrices. From this viewpoint,  $\check{X}$  can be seen as a subset of  $\mathbb{R}^n \times M(\mathbb{R}^m)$ . Let  $(\check{X}^t)_{t \geq 0}$  denotes the Čech filtration of  $\check{X}$  in the ambient space  $\mathbb{R}^n \times M(\mathbb{R}^m)$ . A natural map  $p^t: \check{X}^t \rightarrow \mathcal{G}_d(\mathbb{R}^m)$  can be defined: map a point  $(x, A) \in \check{X}^t$  to the projection of  $A$  on  $\mathcal{G}_d(\mathbb{R}^m)$ , seen as a subset of  $M(\mathbb{R}^m)$ :

$$p^t: (x, A) \in \mathbb{R}^n \times M(\mathbb{R}^m) \mapsto \text{proj}(A, \mathcal{G}_d(\mathbb{R}^m)).$$

The projection is well-defined if  $A$  does not belong to the medial axis of  $\mathcal{G}_d(\mathbb{R}^m)$ . We show that this condition can be verified in practice (Lemma 3.1). The Čech filtration of  $\check{X}$ , endowed with the extended projection maps  $(p^t: \check{X}^t \rightarrow \mathcal{G}_d(\mathbb{R}^m))_t$ , is called the *Čech bundle filtration*. Now we can define the  $i^{\text{th}}$  persistent Stiefel-Whitney class as the collection of classes  $w_i(X) = (w_i^t(X))_t$ , where  $w_i^t(X)$  is the push-forward

$$w_i^t(X) = (p^t)^*(w_i),$$

and where  $w_i$  is the  $i^{\text{th}}$  Stiefel-Whitney class of the Grassmann manifold (compare with Equation (1)). We summarize the information given by a persistent Stiefel-Whitney class in a diagram, that we call a *lifebar*.

The construction we propose is defined for any subset of  $\mathbb{R}^n \times M(\mathbb{R}^m)$ . We prove that this construction is stable, a result reminiscent of the usual stability theorem of persistent homology (Corollary 3.4). We also show that the persistent Stiefel-Whitney classes are consistent estimators of Stiefel-Whitney classes (Corollary 3.7).

Moreover, we propose a concrete algorithm to compute the persistent Stiefel-Whitney classes. This algorithm is based on several ingredients, including the triangulation of projective spaces, and the simplicial approximation method. The simplicial approximation, widely used in theory, can be applied only if the simplicial complex is refined enough, a property that is attested by the star condition. However, this condition cannot be verified in practice. We circumvent this problem by introducing the *weak star condition*, a variant that only depends on the combinatorial structure of the simplicial complex. When the simplicial complex is fine enough, the star condition and the weak star condition turn out to be equivalent notions (Proposition 5.5).

**Numerical experiments.** A Python notebook, containing a concise demonstration of our method, can be found at <https://github.com/raphaeltinarrage/PersistentCharacteristicClasses/blob/master/Demo.ipynb>. Another notebook, containing experiments on datasets inspired by image analysis, can be found at <https://github.com/raphaeltinarrage/PersistentCharacteristicClasses/blob/master/Experiments.ipynb>.

**Outline.** The rest of the paper is as follows. Sect. 2 gathers usual definitions related to vector bundles, Stiefel-Whitney classes, simplicial approximation and persistent cohomology. The definitions of vector bundle filtrations and persistent Stiefel-Whitney classes are given in Sect. 3, where their stability and consistency properties are established. In Sect. 4, we propose a sketch of algorithm to compute these classes, based on simplicial approximation techniques. In Sect. 5 we give a particular attention to some technical details needed to implement this algorithm. In Sect. 6 we apply our algorithm on concrete datasets. For the clarity of the paper, the proofs of some results have been postponed to the appendices.

## 2 Background

### 2.1 Stiefel-Whitney classes

In this subsection, we define vector bundles and Stiefel-Whitney classes. The reader may refer to Milnor and Stasheff (2016) for an extended presentation. Let  $X$  be a topological space and  $d \geq 1$  an integer.

**Vector bundles.** A *vector bundle*  $\xi$  of dimension  $d$  over  $X$  consists of a topological space  $A = A(\xi)$ , the *total space*, a continuous map  $\pi = \pi(\xi): A \rightarrow X$ , the *projection map*, and for every  $x \in X$ , a structure of  $d$ -dimensional vector space on the *fiber*  $\pi^{-1}(\{x\})$ . Moreover,  $\xi$  must satisfy the local triviality condition: for every  $x \in X$ , there exists a neighborhood  $U \subseteq X$  of  $x$  and a homeomorphism  $h: U \times \mathbb{R}^d \rightarrow \pi^{-1}(U)$  such that for every  $y \in U$ , the map  $z \mapsto h(y, z)$  defines an isomorphism between the vector spaces  $\mathbb{R}^d$  and  $\pi^{-1}(\{y\})$ .

A *bundle map* between two vector bundles  $\xi$  and  $\eta$  with base spaces  $X$  and  $Y$  is a continuous map  $f: A(\xi) \rightarrow A(\eta)$  which sends each fiber  $\pi(\xi)^{-1}(\{x\})$  isomorphically into another fiber  $\pi(\eta)^{-1}(\{x'\})$ . If such a map exists, there exist a unique map  $\bar{f}$  which makes the following diagram commute:

$$\begin{array}{ccc} A(\xi) & \xrightarrow{f} & A(\eta) \\ \pi(\xi) \downarrow & & \downarrow \pi(\eta) \\ X & \xrightarrow{\bar{f}} & Y \end{array}$$

If  $X = Y$ , and if  $f$  is a homeomorphism, we say that  $f$  is an *isomorphism* of vector bundles, and that  $\xi$  and  $\eta$  are *isomorphic*. The *trivial bundle* of dimension  $d$  over  $X$ , denoted  $\epsilon$ , is defined with the total space  $A(\epsilon) = X \times \mathbb{R}^d$ , with the projection map  $\pi(\epsilon)$  being the projection on the first coordinate, and where each fiber is endowed with the usual vector space structure of  $\mathbb{R}^d$ . A vector bundle  $\xi$  over  $X$  is said *trivial* if it is isomorphic to  $\epsilon$ .

Let  $m \geq 0$ . The Grassmann manifold  $\mathcal{G}_d(\mathbb{R}^m)$  is a set which consists of all  $d$ -dimensional linear subspaces of  $\mathbb{R}^m$ . It can be given a smooth manifold structure. When  $d = 1$ ,  $\mathcal{G}_1(\mathbb{R}^m)$  corresponds to the real projective space  $\mathbb{P}_{m-1}(\mathbb{R})$ . In order to avoid mentioning  $m$ , it is convenient to consider the infinite Grassmannian. The infinite Grassmann manifold  $\mathcal{G}_d(\mathbb{R}^\infty)$  is the set of all  $d$ -dimensional linear subspaces of  $\mathbb{R}^\infty$ , where  $\mathbb{R}^\infty$  is the vector space of sequences with a finite number of nonzero terms.

Let  $X$  be a paracompact space. There exists a correspondence between the vector bundles over  $X$  (up to isomorphism) and the continuous maps  $X \rightarrow \mathcal{G}_d(\mathbb{R}^\infty)$  (up to homotopy). Such a map is called a *classifying map*. When  $X$  is compact, there exist an integer  $m \geq 1$  such that a classifying map factorizes through

$$X \longrightarrow \mathcal{G}_d(\mathbb{R}^m) \longleftarrow \mathcal{G}_d(\mathbb{R}^\infty).$$

Consequently, in the rest of this paper, we shall consider that vector bundles are given as a continuous maps  $X \rightarrow \mathcal{G}_d(\mathbb{R}^m)$  or  $X \rightarrow \mathcal{G}_d(\mathbb{R}^\infty)$ .

**Axioms for Stiefel-Whitney classes.** To each vector bundle  $\xi$  over a paracompact base space  $X$ , one associates a sequence of cohomology classes

$$w_i(\xi) \in H^i(X, \mathbb{Z}_2), \quad i \in \mathbb{N},$$

called the *Stiefel-Whitney classes* of  $\xi$ . These classes satisfy:

- **Axiom 1:**  $w_0$  is equal to  $1 \in H^0(X, \mathbb{Z}_2)$ , and if  $\xi$  is of dimension  $d$ , then  $w_i(\xi) = 0$  for  $i > d$ .
- **Axiom 2:** if  $f: \xi \rightarrow \eta$  is a bundle map, then  $w_i(\xi) = \bar{f}^* w_i(\eta)$ , where  $\bar{f}^*$  is the map in cohomology induced by the underlying map  $\bar{f}: X \rightarrow Y$  between base spaces.
- **Axiom 3:** if  $\xi, \eta$  are bundles over the same base space  $X$ , then for all  $k \in \mathbb{N}$ ,  $w_k(\xi \oplus \eta) = \sum_{i=0}^k w_i(\xi) \smile w_{k-i}(\eta)$ , where  $\oplus$  denotes the Whitney sum, and  $\smile$  denotes the cup product.
- **Axiom 4:** if  $\gamma_1^1$  denotes the tautological bundle of the projective line  $\mathcal{G}_1(\mathbb{R}^2)$ , then  $w_1(\gamma_1^1) \neq 0$ .

If such classes exist, then one proves that they are unique. A way to show that they actually exist relies on the cohomology of the Grassmannians.

**Construction of the Stiefel-Whitney classes.** The cohomology rings of the Grassmann manifolds admit a simple description:  $H^*(\mathcal{G}_d(\mathbb{R}^\infty), \mathbb{Z}_2)$  is the free abelian ring generated by  $d$  elements  $w_1, \dots, w_d$ . As a graded algebra, the degree of these elements are  $|w_1| = 1, \dots, |w_d| = d$ . Hence we can write

$$H^*(\mathcal{G}_d(\mathbb{R}^\infty), \mathbb{Z}_2) \cong \mathbb{Z}_2[w_1, \dots, w_d].$$

The generators  $w_1, \dots, w_d$  can be seen as the Stiefel-Whitney classes of a particular vector bundle on  $\mathcal{G}_d(\mathbb{R}^\infty)$ , called the *tautological bundle*. Now, for any vector bundle  $\xi$ , define

$$w_i(\xi) = f_\xi^*(w_i),$$

where  $f_\xi: X \rightarrow \mathcal{G}_d(\mathbb{R}^\infty)$  is a classifying map for  $\xi$ , and  $f_\xi^*: H^*(X) \leftarrow H^*(\mathcal{G}_d(\mathbb{R}^\infty))$  the induced map in cohomology. This construction yields the Stiefel-Whitney classes.

**Interpretation of the Stiefel-Whitney classes.** The Stiefel-Whitney classes are invariants of isomorphism classes of vector bundles, and carry topological information. Their main interpretation is the following: the Stiefel-Whitney classes are obstructions to the existence of nowhere vanishing sections of vector bundles. Let us explain this result. A *section* of a vector bundle  $\pi(\xi): A \rightarrow X$  is a continuous map  $s: X \rightarrow A$  such that  $s(x) \in \pi^{-1}(\{x\})$  for all  $x \in X$ . It is nowhere vanishing if  $s(x) \neq 0$  for all  $x \in X$ , where  $0$  denotes the origin of the vector space  $\pi^{-1}(\{x\})$ . Given  $k$  sections  $s_1, \dots, s_k$ , we say that they are *independent* if the family  $(s_1(x), \dots, s_k(x))$  is free for all  $x \in X$ . Then the following result holds: if a vector bundle  $\xi$  of dimension  $d$  admits  $k$  independent and nowhere vanishing sections, then the top  $k$  Stiefel-Whitney classes  $w_d(\xi), \dots, w_{d-k+1}(\xi)$  are zero.

Another property that we will use in this paper is the following: the first Stiefel-Whitney class detects orientability. More precisely, the first Stiefel-Whitney class  $w_1(\xi)$  is zero if and only if the vector bundle  $\xi$  is orientable. In the same vein, if  $X$  is a compact manifold and  $\tau$  its tangent bundle, then the manifold  $X$  is orientable if and only if  $w_1(\tau) = 0$ .

In this paper, we will particularly study *line bundles*, that is, vector bundles of dimension  $d = 1$ . As a consequence of being an obstruction to nowhere vanishing sections, a line bundle  $\xi$  on any topological space  $X$  is trivial if and only if  $w_1(\xi) = 0$ . More generally, the first Stiefel-Whitney class establishes a bijection between the isomorphism classes of line bundles over  $X$  and its first cohomology group  $H^1(X)$  over  $\mathbb{Z}_2$ . As an example, the circle  $\mathbb{S}_1$  has cohomology group  $H^1(\mathbb{S}_1) = \mathbb{Z}_2$ , hence admits only two isomorphism classes of line bundles. As another example, the sphere  $\mathbb{S}_2$  has trivial cohomology group  $H^1(\mathbb{S}_2) = 0$ , hence only admits trivial line bundles.

## 2.2 Simplicial approximation

We start by defining the simplicial complexes and their topology. We then describe the technique of simplicial approximation, based on the book of Hatcher (2002).

**Simplicial complexes.** A *simplicial complex* is a set  $K$  such that there exists a set  $V$ , the set of *vertices*, with  $K$  a collection of finite and non-empty subsets of  $V$ , and such that  $K$  satisfies the following condition: for every  $\sigma \in K$  and every non-empty subset  $\nu \subseteq \sigma$ ,  $\nu$  is in  $K$ . The elements of  $K$  are called *faces* or *simplices* of the simplicial complex  $K$ .

For every simplex  $\sigma \in K$ , we define its dimension  $\dim(\sigma) = \text{card}(\sigma) - 1$ . The *dimension* of  $K$ , denoted  $\dim(K)$ , is the maximal dimension of its simplices. For every  $i \geq 0$ , the  *$i$ -skeleton*  $K^i$  is defined as the subset of  $K$  consisting of simplices of dimension at most  $i$ . Note that  $K^0$  corresponds to the underlying vertex set  $V$ , and that  $K^1$  is a graph. Given a graph  $G$ , the corresponding *clique complex* is the simplicial complex whose simplices are the sets of vertices of the cliques of  $G$ . We say that a simplicial complex  $K$  is a *flag complex* if it is the clique complex of its 1-skeleton  $K^1$ .

Given a simplex  $\sigma \in K$ , its (*open*) *star*  $\text{St}(\sigma)$  is the set of all the simplices  $\nu \in K$  that contain  $\sigma$ . The open star is not a simplicial complex in general. We also define its *closed star*  $\overline{\text{St}}(\sigma)$  as the smallest simplicial subcomplex of  $K$  which contains  $\text{St}(\sigma)$ .

**Geometric realizations.** For every  $p \geq 0$ , the *standard  $p$ -simplex*  $\Delta^p$  is the topological space defined as the convex hull of the canonical basis vectors  $e_1, \dots, e_{p+1}$  of  $\mathbb{R}^{p+1}$ , endowed with the subspace topology. To each simplicial complex  $K$  is attached a *geometric realization*. It is a topological space, denoted  $|K|$ , obtained by gluing the simplices of  $K$  together. According to this construction, each simplex  $\sigma \in K$  admits a geometric realization  $|\sigma|$  which is a subset of  $|K|$ . The following set is a partition of  $|K|$ :

$$\{|\sigma|, \sigma \in K\}.$$

This allows to define the *face map* of  $K$ . It is the unique map  $\mathcal{F}_K: |K| \rightarrow K$  that satisfies  $x \in |\mathcal{F}_K(x)|$  for every  $x \in |K|$ .

If  $\sigma$  is a face of  $K$  of dimension at least 1, the subset  $|\sigma|$  is canonically homeomorphic to the interior of the standard  $p$ -simplex  $\Delta^p$ , where  $p = \dim(\sigma)$ . This allows to define on  $|K|$  the barycentric coordinates: for every face  $\sigma = [v_0, \dots, v_p] \in K$ , the points  $x \in |\sigma|$  can be written as

$$x = \sum_{i=0}^p \lambda_i v_i$$

with  $\lambda_0, \dots, \lambda_p > 0$  and  $\sum_{i=0}^p \lambda_i = 1$ .

If  $X$  is any a topological space, a *triangulation* of  $X$  consists of a simplicial complex  $K$  together with a homeomorphism  $h: X \rightarrow |K|$ .

**Simplicial approximation.** A *simplicial map* between simplicial complexes  $K$  and  $L$  is a map between geometric realizations  $g: |K| \rightarrow |L|$  which sends each simplex of  $K$  to a simplex of  $L$  by a linear maps that sends vertices to vertices. In other words, for every  $\sigma = [v_0, \dots, v_p] \in K$ , the subset  $[g(v_0), \dots, g(v_p)]$  is a simplex of  $L$ , and the map  $g$  restricted to  $|\sigma| \subset |K|$  can be written in barycentric coordinates as

$$\sum_{i=0}^p \lambda_i v_i \mapsto \sum_{i=0}^p \lambda_i g(v_i). \quad (2)$$

A simplicial map  $g: |K| \rightarrow |L|$  is uniquely determined by its restriction to the vertex sets  $g|_{K^0}: K^0 \rightarrow L^0$ . Reciprocally, let  $f: K^0 \rightarrow L^0$  be a map between vertex sets which satisfies the following condition:

$$\forall \sigma \in K, f(\sigma) \in L. \quad (3)$$

Then  $f$  induces a simplicial map via barycentric coordinates, denoted  $|f|: |K| \rightarrow |L|$ . In the rest of this paper, a simplicial map shall either refer to a map  $g: |K| \rightarrow |L|$  which satisfies Equation (2), to a map  $f: K^0 \rightarrow L^0$  which satisfies Equation (3), or to the induced map  $f: K \rightarrow L$ .

Let  $g: |K| \rightarrow |L|$  be any continuous map. The problem of *simplicial approximation* consists in finding a simplicial map  $f: K \rightarrow L$  with geometric realization  $|f|: |K| \rightarrow |L|$  homotopy equivalent to  $g$ . A way to solve this problem is to consider the following property (Hatcher, 2002, Proof of Theorem 2C.1): we say that the map  $g$  satisfies the *star condition* if for every vertex  $v$  of  $K$ , there exists a vertex  $w$  of  $L$  such that

$$g(|\overline{\text{St}}(v)|) \subseteq |\text{St}(w)|.$$

If this is the case, let  $f: K^0 \rightarrow L^0$  be any map between vertex sets such that for every vertex  $v$  of  $K$ , we have  $g(|\overline{\text{St}}(v)|) \subseteq |\text{St}(f(v))|$ . Equivalently,  $f$  satisfies

$$g(|\overline{\text{St}}(v)|) \subseteq \text{St}(f(v)).$$

Such a map is called a *simplicial approximation to  $g$* . One shows that it is a simplicial map, and that its geometric realization  $|f|$  is homotopic to  $g$ .

In general, a map  $g$  may not satisfy the star condition. However, there is always a way to subdivide the simplicial complex  $K$  in order to obtain an induced map which does. We describe this construction in the following paragraph.

We point out that, for some authors, such as Munkres (1984), the star condition is defined by the property  $g(|\text{St}(v)|) \subseteq |\text{St}(w)|$ . The definition we used above, although harder to satisfy than this one, will be enough for our purposes.

**Barycentric subdivisions.** Let  $\Delta^p$  denote the standard  $p$ -simplex, with vertices denoted  $v_0, \dots, v_p$ . The *barycentric subdivision* of  $\Delta^p$  consists in decomposing  $\Delta^p$  into  $(p+1)!$  simplices of dimension  $p$ . It is a simplicial complex, whose vertex set corresponds to the points  $\sum_{i=0}^p \lambda_i v_i$  for which some  $\lambda_i$  are zero and the other ones are equal. Equivalently, one can see this new set of vertices as a the power set of the set of vertices of  $\Delta^p$ .

More generally, if  $K$  is a simplicial complex, its barycentric subdivision  $\text{sub}(K)$  is the simplicial complex obtained by subdivising each of its faces. The set of vertices of  $\text{sub}(K)$  can be seen as a subset of the power set of the set of vertices of  $K$ .

If  $g: |K| \rightarrow |L|$  is any map, there exists a canonical extended map  $|\text{sub}(K)| \rightarrow |L|$ , still denoted  $g$ . The *simplicial approximation theorem* states that for any two simplicial complexes  $K, L$  with  $K$  finite, and  $g: |K| \rightarrow |L|$  any a continuous map, there exists  $n \geq 0$  such that  $g: |\text{sub}^n(K)| \rightarrow |L|$  satisfies the star condition. As a consequence, such a map  $g: |\text{sub}^n(K)| \rightarrow |L|$  admits a simplicial approximation.

### 2.3 Persistent cohomology

In this subsection, we write down the definitions of persistence modules, and their associated pseudo-distances, in the context of cohomology. Compared to the standard definitions of persistent homology, the arrows go backward. Let  $T \subseteq [0, +\infty)$  be an interval that contains 0, let  $E$  be a Euclidean space, and  $k$  a field.

**Persistence modules.** A *persistence module* over  $T$  is a pair  $(\mathbb{V}, \mathfrak{v})$  where  $\mathbb{V} = (V^t)_{t \in T}$  is a family of  $k$ -vector spaces, and  $\mathfrak{v} = (v_s^t)_{s \leq t \in T}$  is a family of linear maps  $v_s^t: V^s \leftarrow V^t$  such that:

- for every  $t \in T$ ,  $v_t^t: V^t \leftarrow V^t$  is the identity map,
- for every  $r, s, t \in T$  such that  $r \leq s \leq t$ ,  $v_r^s \circ v_s^t = v_r^t$ .

When there is no risk of confusion, we may denote a persistence module by  $\mathbb{V}$  instead of  $(\mathbb{V}, \mathfrak{v})$ . Given  $\epsilon \geq 0$ , an  $\epsilon$ -*morphism* between two persistence modules  $(\mathbb{V}, \mathfrak{v})$  and  $(\mathbb{W}, \mathfrak{w})$  is a family of linear maps  $(\phi_t: V^t \rightarrow W^{t-\epsilon})_{t \geq \epsilon}$  such that the following diagram commutes for every  $\epsilon \leq s \leq t$ :

$$\begin{array}{ccc} V^s & \xleftarrow{\quad} & V^t \\ \downarrow \phi_s & \begin{array}{c} v_s^t \\ \leftarrow \end{array} & \downarrow \phi_t \\ W^{s-\epsilon} & \xleftarrow{\quad} & W^{t-\epsilon} \\ & \begin{array}{c} w_{s-\epsilon}^{t-\epsilon} \\ \leftarrow \end{array} & \end{array}$$

If  $\epsilon = 0$  and each  $\phi_t$  is an isomorphism, the family  $(\phi_t)_{t \in T}$  is an *isomorphism* of persistence modules. An  $\epsilon$ -*interleaving* between two persistence modules  $(\mathbb{V}, \mathfrak{v})$  and  $(\mathbb{W}, \mathfrak{w})$  is a pair of  $\epsilon$ -morphisms  $(\phi_t: V^t \rightarrow W^{t-\epsilon})_{t \geq \epsilon}$  and  $(\psi_t: W^t \rightarrow V^{t-\epsilon})_{t \geq \epsilon}$  such that the following diagrams commute for every  $t \geq 2\epsilon$ :

$$\begin{array}{ccc} V^{t-2\epsilon} & \xleftarrow{\quad v_{t-2\epsilon}^t \quad} & V^t \\ \swarrow \psi_{t-\epsilon} & & \searrow \phi_t \\ & W^{t-\epsilon} & \end{array} \qquad \begin{array}{ccc} & V^{t-\epsilon} & \\ \swarrow \phi_{t-\epsilon} & & \searrow \psi_t \\ W^{t-2\epsilon} & \xleftarrow{\quad w_{t-2\epsilon}^t \quad} & W^t \end{array}$$

The interleaving pseudo-distance between  $(\mathbb{V}, \mathfrak{v})$  and  $(\mathbb{W}, \mathfrak{w})$  is defined as

$$d_i(\mathbb{V}, \mathbb{W}) = \inf\{\epsilon \geq 0, \mathbb{V} \text{ and } \mathbb{W} \text{ are } \epsilon\text{-interleaved}\}.$$



**Persistence barcodes.** A persistence module  $(\mathbb{V}, \mathfrak{v})$  is said to be *pointwise finite-dimensional* if for every  $t \in T$ ,  $V^t$  is finite-dimensional. This implies that we can define a notion of persistence barcode (Botnan and Crawley-Boevey, 2020). It comes from the algebraic decomposition of the persistence module into interval modules. Moreover, given two pointwise finite-dimensional persistence modules  $\mathbb{V}, \mathbb{W}$  with persistence barcodes  $\text{Barcode}(\mathbb{V})$  and  $\text{Barcode}(\mathbb{W})$ , the so-called isometry theorem states that

$$d_b(\text{Barcode}(\mathbb{V}), \text{Barcode}(\mathbb{W})) = d_i(\mathbb{V}, \mathbb{W}),$$

where  $d_i(\cdot, \cdot)$  denotes the interleaving distance between persistence modules, and  $d_b(\cdot, \cdot)$  denotes the bottleneck distance between barcodes.

More generally, the persistence module  $(\mathbb{V}, \mathfrak{v})$  is said to be *q-tame* if for every  $s, t \in T$  such that  $s < t$ , the map  $v_s^t$  is of finite rank. The *q-tameness* of a persistence module ensures that we can still define a notion of persistence barcode, even though the module may not be decomposable into interval modules. Moreover, the isometry theorem still holds (Chazal et al., 2016).

**Filtrations of sets and simplicial complexes.** A family of subsets  $\mathbb{X} = (X^t)_{t \in T}$  of  $E$  is a *filtration* if it is non-decreasing for the inclusion, i.e. for any  $s, t \in T$ , if  $s \leq t$  then  $X^s \subseteq X^t$ . Given  $\epsilon \geq 0$ , two filtrations  $\mathbb{X} = (X^t)_{t \in T}$  and  $\mathbb{Y} = (Y^t)_{t \in T}$  of  $E$  are  $\epsilon$ -*interleaved* if, for every  $t \in T$ ,  $X^t \subseteq Y^{t+\epsilon}$  and  $Y^t \subseteq X^{t+\epsilon}$ . The interleaving pseudo-distance between  $\mathbb{X}$  and  $\mathbb{Y}$  is defined as the infimum of such  $\epsilon$ :

$$d_i(\mathbb{X}, \mathbb{Y}) = \inf\{\epsilon, \mathbb{X} \text{ and } \mathbb{Y} \text{ are } \epsilon\text{-interleaved}\}.$$

Filtrations of simplicial complexes and their interleaving distance are similarly defined: given a simplicial complex  $S$ , a *filtration of  $S$*  is a non-decreasing family  $\mathbb{S} = (S^t)_{t \in T}$  of subcomplexes of  $S$ . The interleaving pseudo-distance between two filtrations  $(S_1^t)_{t \in T}$  and  $(S_2^t)_{t \in T}$  of  $S$  is the infimum of the  $\epsilon \geq 0$  such that they are  $\epsilon$ -interleaved, i.e., for any  $t \in T$ , we have  $S_1^t \subseteq S_2^{t+\epsilon}$  and  $S_2^t \subseteq S_1^{t+\epsilon}$ .

**Relation between filtrations and persistence modules.** Applying the singular cohomology functor to a set filtration gives rise to a persistence module whose linear maps between cohomology groups are induced by the inclusion maps between sets. As a consequence, if two filtrations are  $\epsilon$ -interleaved, then their associated persistence modules are also  $\epsilon$ -interleaved, the interleaving homomorphisms being induced by the interleaving inclusion maps. As a consequence of the isometry theorem, if the modules are *q-tame*, then the bottleneck distance between their persistence barcodes is upperbounded by  $\epsilon$ . The same remarks hold when applying the simplicial cohomology functor to simplicial filtrations.

## 2.4 Notations

We adopt the following notations:

- $I$  denotes a set,  $\text{card}(I)$  its cardinal and  $I^c$  its complement.
- if  $i$  and  $j$  are integers such that  $i \leq j$ ,  $\llbracket i, j \rrbracket$  denotes the set of integers between  $i$  and  $j$  included.
- $\mathbb{R}^n$  and  $\mathbb{R}^m$  denote the Euclidean spaces of dimension  $n$  and  $m$ ,  $E$  denotes a Euclidean space.
- $M(\mathbb{R}^m)$  denotes the vector space of  $m \times m$  matrices,  $\mathcal{G}_d(\mathbb{R}^m)$  the Grassmannian of  $d$ -subspaces of  $\mathbb{R}^m$ , and  $\mathbb{S}_k \subset \mathbb{R}^{k+1}$  the unit  $k$ -sphere.
- $\|\cdot\|$  denotes the usual Euclidean norm on  $\mathbb{R}^n$ ,  $\|\cdot\|_F$  the Frobenius norm on  $M(\mathbb{R}^m)$ ,  $\|\cdot\|_\gamma$  the norm on  $\mathbb{R}^n \times M(\mathbb{R}^m)$  defined as  $\|(x, A)\|_\gamma^2 = \|x\|^2 + \gamma^2 \|A\|_F^2$  where  $\gamma > 0$  is a parameter.

- $\mathbb{X} = (X^t)_{t \in T}$  denotes a set filtration.  $\mathbb{V}[\mathbb{X}]$  denotes the corresponding persistent cohomology module. If  $X$  is a subset of  $E$ , then  $\mathbb{X} = (X^t)_{t \in T}$  denotes the Čech set filtration of  $X$  (also called the offset filtration).
- $(\mathbb{V}, \mathbb{v})$  denotes a persistence module over  $T$ , with  $\mathbb{V} = (V^t)_{t \in T}$  a family of vector spaces, and  $\mathbb{v} = (v_s^t: X^s \leftarrow X^t)_{s \leq t \in T}$  a family of linear maps.
- $\mathcal{U}$  denotes a cover of a topological space, and  $\mathcal{N}(\mathcal{U})$  its nerve.  $\mathbb{S} = (S^t)_{t \in T}$  denotes a simplicial filtration.
- If  $X$  is a topological space,  $H^*(X)$  denotes its cohomology ring over  $\mathbb{Z}_2$  (the field with two elements), and  $H^i(X)$  its  $i^{\text{th}}$  cohomology group over  $\mathbb{Z}_2$ . If  $f: X \rightarrow Y$  is a continuous map,  $f^*: H^*(X) \leftarrow H^*(Y)$  is the map induced in cohomology.
- If  $\xi$  is a vector bundle,  $w_i(\xi)$  denotes its  $i^{\text{th}}$  Stiefel-Whitney class.
- If  $A$  is a subset of  $E$ , then  $\text{med}(A)$  denotes its medial axis,  $\text{reach}(A)$  its reach and  $\text{dist}(\cdot, A)$  the distance to  $A$ . The projection on  $A$  is denoted  $\text{proj}(\cdot, A)$  or  $\text{proj}_A(\cdot)$ .  $d_H(\cdot, \cdot)$  denotes the Hausdorff distance between two sets of  $E$ .
- If  $K$  is a simplicial complex,  $K^i$  denotes its  $i$ -skeleton. For every vertex  $v \in K^0$ ,  $\text{St}(v)$  and  $\overline{\text{St}}(v)$  denote its open and closed star. The geometric realization of  $K$  is denoted  $|K|$ , and the geometric realization of a simplex  $\sigma \in K$  is  $|\sigma|$ . The face map is denoted  $\mathcal{F}_K: |K| \rightarrow K$ .
- If  $f: K \rightarrow L$  is a simplicial map,  $|f|: |K| \rightarrow |L|$  denotes its geometric realization. The  $i^{\text{th}}$  barycentric subdivision of the simplicial complex  $K$  is denoted  $\text{sub}^i(K)$ .

### 3 Persistent Stiefel-Whitney classes

#### 3.1 Definition

Let  $E = \mathbb{R}^n$  be a Euclidean space, and  $\mathbb{X} = (X^t)_{t \in T}$  a set filtration of  $E$ . Let us denote by  $i_s^t$  the inclusion map from  $X^s$  to  $X^t$ . In order to define persistent Stiefel-Whitney classes, we have to give such a filtration a vector bundle structure. The infinite Grassmann manifold is denoted  $\mathcal{G}_d(\mathbb{R}^\infty)$ .

**Definition 3.1.** A *vector bundle filtration* of dimension  $d$  on  $E$  is a couple  $(\mathbb{X}, \mathbb{p})$  where  $\mathbb{X} = (X^t)_{t \in T}$  is a set filtration of  $E$  and  $\mathbb{p} = (p^t)_{t \in T}$  a family of continuous maps  $p^t: X^t \rightarrow \mathcal{G}_d(\mathbb{R}^\infty)$  such that, for every  $s, t \in T$  with  $s \leq t$ , we have  $p^t \circ i_s^t = p^s$ . In other words, the following diagram commutes:

$$\begin{array}{ccc} X^s & \xrightarrow{i_s^t} & X^t \\ & \searrow p^s & \swarrow p^t \\ & \mathcal{G}_d(\mathbb{R}^\infty) & \end{array}$$

Note that for any  $m \in \mathbb{N}$ , and by using the inclusion  $\mathcal{G}_d(\mathbb{R}^m) \hookrightarrow \mathcal{G}_d(\mathbb{R}^\infty)$ , one may define a vector bundle filtration by considering maps  $p^t: X^t \rightarrow \mathcal{G}_d(\mathbb{R}^m)$ .

Following Subsect. 2.1, the induced map in cohomology,  $(p^t)^*$ , allows to define the Stiefel-Whitney classes of this vector bundle. Let us introduce some notations. The Stiefel-Whitney classes of  $\mathcal{G}_d(\mathbb{R}^\infty)$  are denoted  $w_1, \dots, w_d$ . The Stiefel-Whitney classes of the vector bundle  $(X^t, p^t)$  are denoted  $w_1^t(\mathbb{p}), \dots, w_d^t(\mathbb{p})$ , and can be defined as  $w_i^t(\mathbb{p}) = (p^t)^*(w_i)$ .

$$\begin{array}{ccc} (p^t)^*: H^*(X^t) & \longleftarrow & H^*(\mathcal{G}_d(\mathbb{R}^\infty)) \\ & & \\ w_1^t(\mathbb{p}) & \longleftarrow & w_1 \\ & & \vdots \\ w_d^t(\mathbb{p}) & \longleftarrow & w_d \end{array}$$

Let  $(\mathbb{V}, \mathbb{v})$  denote the persistence module associated to the filtration  $\mathbb{X}$ , with  $\mathbb{V} = (V^t)_{t \in T}$  and  $\mathbb{v} = (v_s^t)_{s \leq t \in T}$ . Explicitly,  $V^t$  is the cohomology ring  $H^*(X^t)$ , and  $v_s^t$  is the induced map  $H^*(X^s) \leftarrow H^*(X^t)$ . For every  $t \in T$ , the classes  $w_1^t(\mathbb{p}), \dots, w_d^t(\mathbb{p})$  belong to the vector space  $V^t$ . The persistent Stiefel-Whitney classes are defined to be the collection of such classes over  $t$ .

**Definition 3.2.** Let  $(\mathbb{X}, \mathbb{p})$  be a vector bundle filtration. The *persistent Stiefel-Whitney classes* of  $(\mathbb{X}, \mathbb{p})$  are the families of classes

$$\begin{aligned} w_1(\mathbb{p}) &= (w_1^t(\mathbb{p}))_{t \in T} \\ &\vdots \\ w_d(\mathbb{p}) &= (w_d^t(\mathbb{p}))_{t \in T}. \end{aligned}$$

Let  $i \in \llbracket 1, d \rrbracket$ , and consider a persistent Stiefel-Whitney class  $w_i(\mathbb{p})$ . Note that it satisfies the following property: for all  $s, t \in T$  such that  $s \leq t$ ,  $w_i^s(\mathbb{p}) = v_s^t(w_i^t(\mathbb{p}))$ . As a consequence, if a class  $w_i^t(\mathbb{p})$  is given for a  $t \in T$ , one obtains all the others  $w_i^s(\mathbb{p})$ , with  $s \leq t$ , by applying the maps  $v_s^t$ . In particular, if  $w_i^t(\mathbb{p}) = 0$ , then  $w_i^s(\mathbb{p}) = 0$  for all  $s \in T$  such that  $s \leq t$ .

**Lifobar.** In order to visualize the evolution of a persistent Stiefel-Whitney class through the persistence module  $(\mathbb{V}, \mathbb{v})$ , we propose the following bar representation: the lifobar of  $w_i(\mathbb{p})$  is the set

$$\{t \in T, w_i^t(\mathbb{p}) \neq 0\}.$$

According to the last paragraph, the lifobar of a persistent class is an interval of  $T$ , of the form  $[t^\dagger, \sup(T))$  or  $(t^\dagger, \sup(T))$ , where

$$t^\dagger = \inf \{t \in T, w_i^t(\mathbb{p}) \neq 0\},$$

with the convention  $\inf(\emptyset) = \inf(T)$ . In order to distinguish the lifobar of a persistent Stiefel-Whitney class from the bars of the persistence barcodes, we draw the rest of the interval hatched (see Figure 1).



Figure 1: Example of a lifobar of a persistent Stiefel-Whitney class with  $t^\dagger = 0.2$  and  $\max(T) = 1$ .

### 3.2 Čech bundle filtrations

In this subsection, we propose a particular construction of vector bundle filtration, called the *Čech bundle filtration*. We shall work in the ambient space  $E = \mathbb{R}^n \times M(\mathbb{R}^m)$ . Let  $\|\cdot\|$  be the usual Euclidean norm on the space  $\mathbb{R}^n$ , and  $\|\cdot\|_F$  the Frobenius norm on  $M(\mathbb{R}^m)$ , the space of  $m \times m$  matrices. Let  $\gamma > 0$ . We endow the vector space  $E$  with the Euclidean norm  $\|\cdot\|_\gamma$  defined for every  $(x, A) \in E$  as

$$\|(x, A)\|_\gamma^2 = \|x\|^2 + \gamma^2 \|A\|_F^2. \quad (4)$$

See Subsection 5.4 for a discussion about the parameter  $\gamma$ .

In order to define the Čech bundle filtration, we shall first study the usual embedding of the Grassmann manifold  $\mathcal{G}_d(\mathbb{R}^m)$  into the matrix space  $M(\mathbb{R}^m)$ .

**Embedding of  $\mathcal{G}_d(\mathbb{R}^m)$ .** We embed the Grassmannian  $\mathcal{G}_d(\mathbb{R}^m)$  into  $M(\mathbb{R}^m)$  via the application which sends a  $d$ -dimensional subspace  $T \subset \mathbb{R}^m$  to its orthogonal projection matrix  $P_T$ . We can now see  $\mathcal{G}_d(\mathbb{R}^m)$  as a submanifold of  $M(\mathbb{R}^m)$ . Recall that  $M(\mathbb{R}^m)$  is endowed with the Frobenius norm. According to this metric,  $\mathcal{G}_d(\mathbb{R}^m)$  is included in the sphere of center 0 and radius  $\sqrt{d}$  of  $M(\mathbb{R}^m)$ .

In the metric space  $(M(\mathbb{R}^m), \|\cdot\|_F)$ , consider the distance function to  $\mathcal{G}_d(\mathbb{R}^m)$ , denoted  $\text{dist}(\cdot, \mathcal{G}_d(\mathbb{R}^m))$ . Let  $\text{med}(\mathcal{G}_d(\mathbb{R}^m))$  denote the medial axis of  $\mathcal{G}_d(\mathbb{R}^m)$ . It consists in the points  $A \in M(\mathbb{R}^m)$  which admit at least two projections on  $\mathcal{G}_d(\mathbb{R}^m)$ :

$$\begin{aligned} \text{med}(\mathcal{G}_d(\mathbb{R}^m)) &= \{A \in M(\mathbb{R}^m), \exists P, P' \in \mathcal{G}_d(\mathbb{R}^m), P \neq P', \\ &\quad \|A - P\|_F = \|A - P'\|_F = \text{dist}(A, \mathcal{G}_d(\mathbb{R}^m))\}. \end{aligned}$$

On the set  $M(\mathbb{R}^m) \setminus \text{med}(\mathcal{G}_d(\mathbb{R}^m))$ , the projection on  $\mathcal{G}_d(\mathbb{R}^m)$  is well-defined:

$$\begin{aligned} \text{proj}(\cdot, \mathcal{G}_d(\mathbb{R}^m)) : M(\mathbb{R}^m) \setminus \text{med}(\mathcal{G}_d(\mathbb{R}^m)) &\longrightarrow \mathcal{G}_d(\mathbb{R}^m) \subset M(\mathbb{R}^m) \\ A &\longmapsto P \text{ s.t. } \|P - A\|_F = \text{dist}(A, \mathcal{G}_d(\mathbb{R}^m)). \end{aligned}$$

The following lemma describes this projection explicitly.

**Lemma 3.1.** *For any  $A \in M(\mathbb{R}^m)$ , let  $A^s$  denote the matrix  $A^s = \frac{1}{2}(A + A')$ , where  $A'$  is the transpose of  $A$ , and let  $\lambda_1(A^s), \dots, \lambda_m(A^s)$  be the eigenvalues of  $A^s$  in decreasing order. The distance from  $A$  to  $\text{med}(\mathcal{G}_d(\mathbb{R}^m))$  is*

$$\text{dist}(A, \text{med}(\mathcal{G}_d(\mathbb{R}^m))) = \frac{\sqrt{2}}{2} |\lambda_d(A^s) - \lambda_{d+1}(A^s)|.$$

*If this distance is positive, the projection of  $A$  on  $\mathcal{G}_d(\mathbb{R}^m)$  can be described as follows: consider the symmetric matrix  $A^s$ , and let  $A^s = ODO'$ , with  $O$  an orthogonal matrix, and  $D$  the diagonal matrix containing the eigenvalues of  $A^s$  in decreasing order. Let  $J_d$  be the diagonal matrix whose first  $d$  terms are 1, and the other ones are zero. We have*

$$\text{proj}(A, \mathcal{G}_d(\mathbb{R}^m)) = OJ_dO'.$$

*Proof.* Note that  $\mathcal{G}_d(\mathbb{R}^m)$  is contained in the linear subspace  $\mathcal{S}$  of symmetric matrices. Therefore, to project a matrix  $A \in M(\mathbb{R}^m)$  onto  $\mathcal{G}_d(\mathbb{R}^m)$ , we may project on  $\mathcal{S}$  first. It is well known that the projection of  $A$  onto  $\mathcal{S}$  is the matrix  $A^s = \frac{1}{2}(A + A')$ .

Suppose now that we are given a symmetric matrix  $B$ . Let it be diagonalized as  $B = ODO'$  with  $O$  an orthogonal matrix. A projection of  $B$  onto  $\mathcal{G}_d(\mathbb{R}^m)$  is a matrix  $P$  which minimizes the following quantity:

$$\min_{P \in \mathcal{G}_d(E)} \|B - P\|_F. \quad (5)$$

By applying the transformation  $P \mapsto O'PO$ , we see that this problem is equivalent to  $\min_{P \in \mathcal{G}_d(E)} \|D - P\|_F$ . Now, let  $e_1, \dots, e_m$  denote the canonical basis of  $\mathbb{R}^m$ . We have

$$\begin{aligned} \|D - P\|_F^2 &= \|D\|_F^2 + \|P\|_F^2 - 2\langle D, P \rangle_F \\ &= \|D\|_F^2 + \|P\|_F^2 - 2 \sum \langle \lambda_i e_i, P(e_i) \rangle, \end{aligned}$$

where  $\langle \cdot, \cdot \rangle_F$  is the Frobenius inner product, and  $\langle \cdot, \cdot \rangle$  the usual inner product on  $\mathbb{R}^m$ . Therefore, Equation (5) is a problem equivalent to

$$\max_{P \in \mathcal{G}_d(E)} \sum \lambda_i \langle e_i, P(e_i) \rangle.$$

Since  $P$  is an orthogonal projection, we have  $\langle e_i, P(e_i) \rangle = \langle P(e_i), P(e_i) \rangle = \|P(e_i)\|^2$  for all  $i \in \llbracket 1, m \rrbracket$ . Moreover,  $d = \|P\|_F^2 = \sum \|P(e_i)\|^2$ . Denoting  $p_i = \|P(e_i)\|^2 \in [0, 1]$ , we finally obtain the following alternative formulation of Equation (5):

$$\max_{\substack{p_1, \dots, p_m \in [0, 1] \\ p_1 + \dots + p_m = d}} \sum \lambda_i p_i.$$

Using that  $\lambda_1 \geq \dots \geq \lambda_m$ , we see that this maximum is attained when  $p_1 = \dots = p_d = 1$  and  $p_{d+1} = \dots = p_m = 0$ . Consequently, a minimizer of Equation (5) is  $P = J_d$ , where  $J_d$  is the diagonal matrix whose first  $d$  terms are 1, and the other ones are zero. Moreover, it is unique if  $\lambda_d \neq \lambda_{d+1}$ . As a consequence of these considerations, we obtain the following characterization: for every  $B \in M(\mathbb{R}^m)$ ,

$$B \in \text{med}(\mathcal{G}_d(\mathbb{R}^m)) \iff \lambda_d(B^s) = \lambda_{d+1}(B^s). \quad (6)$$

Let us now show that for every matrix  $A \in M(\mathbb{R}^m)$ , we have

$$\text{dist}(A, \text{med}(\mathcal{G}_d(\mathbb{R}^m))) = \frac{\sqrt{2}}{2} |\lambda_d(A^s) - \lambda_{d+1}(A^s)|.$$

First, remark that

$$\text{dist}(A, \text{med}(\mathcal{G}_d(\mathbb{R}^m))) = \text{dist}(A^s, \text{med}(\mathcal{G}_d(\mathbb{R}^m))). \quad (7)$$

Indeed, if  $B$  is a projection of  $A$  on  $\text{med}(\mathcal{G}_d(\mathbb{R}^m))$ , then  $B^s$  is still in  $\text{med}(\mathcal{G}_d(\mathbb{R}^m))$  according to Equation (6), and

$$\text{dist}(A, \text{med}(\mathcal{G}_d(\mathbb{R}^m))) = \|A - B\|_{\mathbb{F}} \geq \|A^s - B^s\|_{\mathbb{F}} \geq \text{dist}(A^s, \text{med}(\mathcal{G}_d(\mathbb{R}^m))).$$

Conversely, if  $B$  is a projection of  $A^s$  on  $\text{med}(\mathcal{G}_d(\mathbb{R}^m))$ , then  $\hat{B} = B + A - A^s$  is still in  $\text{med}(\mathcal{G}_d(\mathbb{R}^m))$ , and

$$\text{dist}(A, \text{med}(\mathcal{G}_d(\mathbb{R}^m))) \leq \|A - \hat{B}\|_{\mathbb{F}} = \|A^s - B\|_{\mathbb{F}} = \text{dist}(A^s, \text{med}(\mathcal{G}_d(\mathbb{R}^m))).$$

We deduce Equation (7). Now, let  $A \in \mathcal{S}$  and  $B \in \text{med}(\mathcal{G}_d(\mathbb{R}^m))$ . Let  $e_1, \dots, e_m$  be a basis of  $\mathbb{R}^m$  that diagonalizes  $A$ . Writing  $\|A - B\|_{\mathbb{F}}^2 = \sum \|A(e_i) - B(e_i)\|^2 = \sum \|\lambda_i(A)e_i - B(e_i)\|^2$ , it is clear that the closest matrix  $B$  must satisfy  $B(e_i) = \lambda_i(B)e_i$ , with

- $\lambda_i(B) = \lambda_i(A)$  for  $i \notin \{d, d+1\}$ ,
- $\lambda_d(B) = \lambda_{d+1}(B) = \frac{1}{2}(\lambda_d(A) + \lambda_{d+1}(A))$ .

We finally compute

$$\begin{aligned} \|A - B\|_{\mathbb{F}}^2 &= \sum \|\lambda_i(A)e_i - \lambda_i(B)e_i\|^2 \\ &= |\lambda_d(A) - \lambda_d(B)|^2 + |\lambda_{d+1}(A) - \lambda_{d+1}(B)|^2 \\ &= \frac{1}{2} |\lambda_d(A) - \lambda_{d+1}(A)|^2 \end{aligned}$$

which yields the result.  $\square$

Observe that, as a consequence of this lemma, every point of  $\mathcal{G}_d(\mathbb{R}^m)$  is at equal distance from  $\text{med}(\mathcal{G}_d(\mathbb{R}^m))$ , and this distance is equal to  $\frac{\sqrt{2}}{2}$ . Therefore the reach of the subset  $\mathcal{G}_d(\mathbb{R}^m) \subset M(\mathbb{R}^m)$  is

$$\text{reach}(\mathcal{G}_d(\mathbb{R}^m)) = \frac{\sqrt{2}}{2}.$$

**Čech bundle filtration.** Let  $X$  be a subset of  $E = \mathbb{R}^n \times M(\mathbb{R}^m)$ . Consider the usual Čech filtration  $\mathbb{X} = (X^t)_{t \geq 0}$ , where  $X^t$  denotes the  $t$ -thickening of  $X$  in the metric space  $(E, \|\cdot\|_{\gamma})$ . It is also known as the offset filtration. In order to give this filtration a vector bundle structure, consider the map  $p^t$  defined as the composition

$$X^t \subset \mathbb{R}^n \times M(\mathbb{R}^m) \xrightarrow{\text{proj}_2} M(\mathbb{R}^m) \setminus \text{med}(\mathcal{G}_d(\mathbb{R}^m)) \xrightarrow{\text{proj}(\cdot, \mathcal{G}_d(\mathbb{R}^m))} \mathcal{G}_d(\mathbb{R}^m), \quad (8)$$

where  $\text{proj}_2$  represents the projection on the second coordinate of  $\mathbb{R}^n \times M(\mathbb{R}^m)$ , and  $\text{proj}(\cdot, \mathcal{G}_d(\mathbb{R}^m))$  the projection on  $\mathcal{G}_d(\mathbb{R}^m) \subset M(\mathbb{R}^m)$ . Note that  $p^t$  is well-defined only when  $X^t$  does not intersect  $\mathbb{R}^n \times \text{med}(\mathcal{G}_d(\mathbb{R}^m))$ . The supremum of such  $t$ 's is denoted  $t_\gamma^{\max}(X)$ . We have

$$t_\gamma^{\max}(X) = \inf \{ \text{dist}_\gamma(x, \mathbb{R}^n \times \text{med}(\mathcal{G}_d(\mathbb{R}^m))), x \in X \}, \quad (9)$$

where  $\text{dist}_\gamma(x, \mathbb{R}^n \times \text{med}(\mathcal{G}_d(\mathbb{R}^m)))$  is the distance between the point  $x \in \mathbb{R}^n \times M(\mathbb{R}^m)$  and the subspace  $\mathbb{R}^n \times \text{med}(\mathcal{G}_d(\mathbb{R}^m))$ , with respect to the norm  $\|\cdot\|_\gamma$ . By definition of  $\|\cdot\|_\gamma$ , Equation (9) rewrites as

$$t_\gamma^{\max}(X) = \gamma \cdot \inf \{ \text{dist}(A, \text{med}(\mathcal{G}_d(\mathbb{R}^m))), (y, A) \in X \},$$

where  $\text{dist}(A, \text{med}(\mathcal{G}_d(\mathbb{R}^m)))$  represents the distance between the matrix  $A$  and the subset  $\text{med}(\mathcal{G}_d(\mathbb{R}^m))$  with respect to the Frobenius norm  $\|\cdot\|_F$ . Denoting  $t^{\max}(X)$  the value  $t_\gamma^{\max}(X)$  for  $\gamma = 1$ , we obtain

$$t_\gamma^{\max}(X) = \gamma \cdot t^{\max}(X) \quad \text{and} \quad t^{\max}(X) = \inf \{ \text{dist}(A, \text{med}(\mathcal{G}_d(\mathbb{R}^m))), (y, A) \in X \}. \quad (10)$$

Note that the values  $t^{\max}(X)$  can be computed explicitly thanks to Lemma 3.1. In particular, if  $X$  is a subset of  $\mathbb{R}^n \times \mathcal{G}_d(\mathbb{R}^m)$ , then  $t^{\max}(X) = \frac{\sqrt{2}}{2}$ . Accordingly,

$$t_\gamma^{\max}(X) = \frac{\sqrt{2}}{2} \gamma. \quad (11)$$

**Definition 3.3.** Consider a subset  $X$  of  $E = \mathbb{R}^n \times M(\mathbb{R}^m)$ , and suppose that  $t^{\max}(X) > 0$ . The *Čech bundle filtration* associated to  $X$  in the ambient space  $(E, \|\cdot\|_\gamma)$  is the vector bundle filtration  $(\mathbb{X}, \mathbb{p})$  consisting of the Čech filtration  $\mathbb{X} = (X^t)_{t \in T}$ , and the maps  $\mathbb{p} = (p^t)_{t \in T}$  as defined in Equation (8). This vector bundle filtration is defined on the index set  $T = [0, t_\gamma^{\max}(X))$ , where  $t_\gamma^{\max}(X)$  is defined in Equation (10).

The  $i^{\text{th}}$  persistent Stiefel-Whitney class of the Čech bundle filtration  $(\mathbb{X}, \mathbb{p})$ , as in Definition 3.2, shall be denoted  $w_i(X)$  instead of  $w_i(\mathbb{p})$ .

**Example 3.2.** Let  $E = \mathbb{R}^2 \times M(\mathbb{R}^2)$ . Let  $X$  and  $Y$  be the subsets of  $E$  defined as:

$$X = \left\{ \left( \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \begin{pmatrix} \cos(\theta)^2 & \cos(\theta) \sin(\theta) \\ \cos(\theta) \sin(\theta) & \sin(\theta)^2 \end{pmatrix} \right), \theta \in [0, 2\pi) \right\}$$

$$Y = \left\{ \left( \begin{pmatrix} \cos(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) \end{pmatrix}, \begin{pmatrix} \cos(\frac{\theta}{2})^2 & \cos(\frac{\theta}{2}) \sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2}) \sin(\frac{\theta}{2}) & \sin(\frac{\theta}{2})^2 \end{pmatrix} \right), \theta \in [0, 2\pi) \right\}$$

The set  $X$  is to be seen as the normal bundle of the circle, and  $Y$  as the tautological bundle of the circle, also known as the Möbius strip. They are pictured in Figures 2 and 3. We have  $t^{\max}(X) = t^{\max}(Y) = \frac{\sqrt{2}}{2}$  as in Lemma 3.1. Let  $\gamma = 1$ .

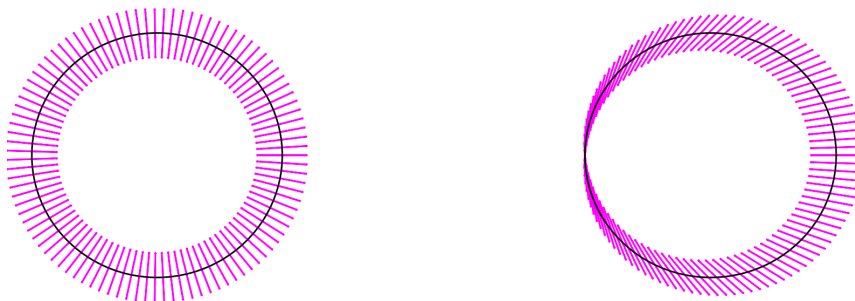


Figure 2: Representation of the sets  $X$  and  $Y \subset \mathbb{R}^2 \times M(\mathbb{R}^2)$ : the black points correspond to the  $\mathbb{R}^2$ -coordinate, and the pink segments over them correspond to the orientation of the  $M(\mathbb{R}^2)$ -coordinate.

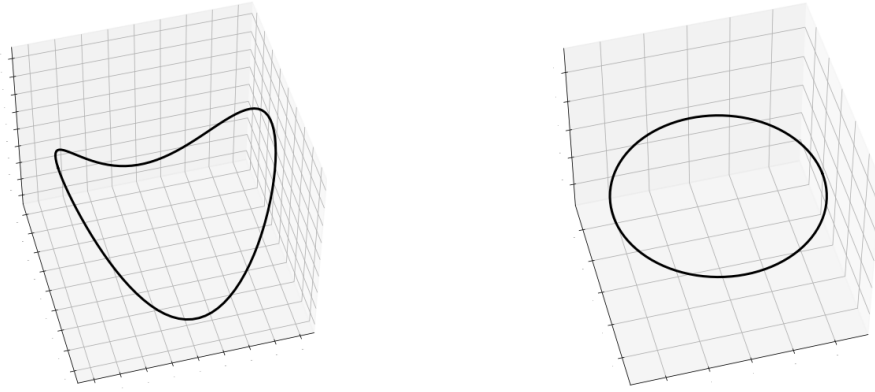


Figure 3: The sets  $X$  and  $Y \subset \mathbb{R}^2 \times M(\mathbb{R}^2)$ , projected in a 3-dimensional subspace of  $\mathbb{R}^3$  via Principal Component Analysis.

We now compute the persistence barcodes of the Čech filtrations of  $X$  and  $Y$  in the ambient space  $E$ , as represented in Figure 4.

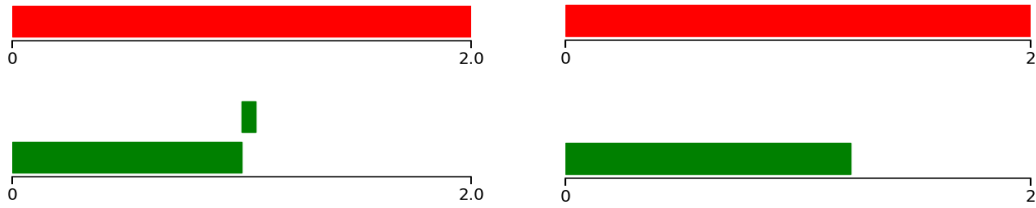


Figure 4:  $H^0$  and  $H^1$  persistence barcodes of the Čech filtration of  $X$  (left) and  $Y$  (right).

Consider the first persistent Stiefel-Whitney classes  $w_1(X)$  and  $w_1(Y)$  of the corresponding Čech bundle filtrations. We compute that their lifebars are  $\emptyset$  for  $w_1(X)$ , and  $[0, t^{\max}(Y)]$  for  $w_1(Y)$ . This is illustrated in Figure 5. One reads these bars as follows:  $w_1^t(X)$  is zero for every  $t \in [0, \frac{\sqrt{2}}{2})$ , while  $w_1^t(Y)$  is nonzero.



Figure 5: Lifebars of the first persistent Stiefel-Whitney classes  $w_1(X)$  and  $w_1(Y)$ .

### 3.3 Stability

In this subsection we derive a straightforward stability result for persistent Stiefel-Whitney classes. We start by defining a notion of interleavings for vector bundle filtrations, in the same vein as the usual interleavings of set filtrations.

**Definition 3.4.** Let  $\epsilon \geq 0$ , and consider two vector bundle filtrations  $(\mathbb{X}, \mathfrak{p})$ ,  $(\mathbb{Y}, \mathfrak{q})$  of dimension  $d$  on  $E$  with respective index sets  $T$  and  $U$ . They are  $\epsilon$ -interleaved if the underlying filtrations  $\mathbb{X} = (X^t)_{t \in T}$  and  $\mathbb{Y} = (Y^t)_{t \in U}$  are  $\epsilon$ -interleaved, and if the following diagrams commute for every  $t \in T \cap (U - \epsilon)$  and  $s \in U \cap (T - \epsilon)$ :

$$\begin{array}{ccc}
X^t & \xrightarrow{\quad} & Y^{t+\epsilon} \\
& \searrow p^t & \swarrow q^{t+\epsilon} \\
& & \mathcal{G}_d(\mathbb{R}^\infty)
\end{array}
\qquad
\begin{array}{ccc}
Y^s & \xrightarrow{\quad} & X^{s+\epsilon} \\
& \searrow q^s & \swarrow q^{s+\epsilon} \\
& & \mathcal{G}_d(\mathbb{R}^\infty)
\end{array}$$

The following theorem shows that interleavings of vector bundle filtrations give rise to interleavings of persistence modules which respect the persistent Stiefel-Whitney classes.

**Theorem 3.3.** *Consider two vector bundle filtrations  $(\mathbb{X}, \mathbb{p})$ ,  $(\mathbb{Y}, \mathbb{q})$  of dimension  $d$  with respective index sets  $T$  and  $U$ . Suppose that they are  $\epsilon$ -interleaved. Then there exists an  $\epsilon$ -interleaving  $(\phi, \psi)$  between their corresponding persistent cohomology modules which sends persistent Stiefel-Whitney classes on persistent Stiefel-Whitney classes. In other words, for every  $i \in \llbracket 1, d \rrbracket$ , and for every  $t \in (T + \epsilon) \cap U$  and  $s \in U \cap (T + \epsilon)$ , we have*

$$\begin{aligned}
\phi^t(w_i^t(\mathbb{p})) &= w_i^{t-\epsilon}(\mathbb{q}) \\
\text{and } \psi^s(w_i^s(\mathbb{p})) &= w_i^{s-\epsilon}(\mathbb{q}).
\end{aligned}$$

*Proof.* Define  $(\phi, \psi)$  to be the  $\epsilon$ -interleaving between the cohomology persistence modules  $\mathbb{V}(\mathbb{X})$  and  $\mathbb{V}(\mathbb{Y})$  given by the  $\epsilon$ -interleaving between the filtrations  $\mathbb{X}$  and  $\mathbb{Y}$ . Explicitly, if  $i_t^{t+\epsilon}$  denotes the inclusion  $X^t \hookrightarrow Y^{t+\epsilon}$  and  $j_s^{s+\epsilon}$  denotes the inclusion  $Y^s \hookrightarrow X^{s+\epsilon}$ , then  $\phi = (\phi^t)_{t \in (T+\epsilon) \cap U}$  is given by the induced maps in cohomology  $\phi^t = (i_t^{t+\epsilon})^*$ , and  $\psi = (\psi^s)_{s \in (U+\epsilon) \cap T}$  is given by  $\psi^s = (j_s^{s+\epsilon})^*$ .

Now, by functoriality, the diagrams of Definition 3.4 give rise to commutative diagrams in cohomology:

$$\begin{array}{ccc}
H^*(X^{t-\epsilon}) & \xleftarrow{\phi^t} & H^*(Y^t) \\
& \swarrow (p^{t-\epsilon})^* & \searrow (q^t)^* \\
& & H^*(\mathcal{G}_d(\mathbb{R}^\infty))
\end{array}
\qquad
\begin{array}{ccc}
H^*(Y^{s-\epsilon}) & \xleftarrow{\psi^s} & H^*(X^s) \\
& \swarrow (q^{s-\epsilon})^* & \searrow (p^s)^* \\
& & H^*(\mathcal{G}_d(\mathbb{R}^\infty))
\end{array}$$

Let  $i \in \llbracket 1, d \rrbracket$ . By definition, the persistent Stiefel-Whitney classes  $w_i(\mathbb{p}) = (w_i^t(\mathbb{p}))_{t \in T}$  and  $w_i(\mathbb{q}) = (w_i^s(\mathbb{q}))_{s \in U}$  are equal to  $w_i^t(\mathbb{p}) = (p^t)^*(w_i)$  and  $w_i^s(\mathbb{q}) = (q^s)^*(w_i)$ , where  $w_i$  is the  $i^{\text{th}}$  Stiefel-Whitney class of  $\mathcal{G}_d(\mathbb{R}^\infty)$ . The previous commutative diagrams then translates as  $\phi^t(w_i^t(\mathbb{p})) = w_i^{t-\epsilon}(\mathbb{q})$  and  $\psi^s(w_i^s(\mathbb{p})) = w_i^{s-\epsilon}(\mathbb{q})$ , as wanted.  $\square$

Consider two vector bundle filtrations  $(\mathbb{X}, \mathbb{p})$ ,  $(\mathbb{Y}, \mathbb{q})$  such that there exists an  $\epsilon$ -interleaving  $(\phi, \psi)$  between their persistent cohomology modules  $\mathbb{V}(\mathbb{X})$ ,  $\mathbb{V}(\mathbb{Y})$  which sends persistent Stiefel-Whitney classes on persistent Stiefel-Whitney classes. Let  $i \in \llbracket 1, d \rrbracket$ . Then the lifebars of their  $i^{\text{th}}$  persistent Stiefel-Whitney classes  $w_i(\mathbb{p})$  and  $w_i(\mathbb{q})$  are  $\epsilon$ -close in the following sense: if we denote  $t^\dagger(\mathbb{p}) = \inf\{t \in T, w_i^t(\mathbb{p}) \neq 0\}$  and  $t^\dagger(\mathbb{q}) = \inf\{t \in T, w_i^t(\mathbb{q}) \neq 0\}$ , then  $|t^\dagger(\mathbb{p}) - t^\dagger(\mathbb{q})| \leq \epsilon$ . This can be seen from their lifebars representations, as shown in Figure 6.

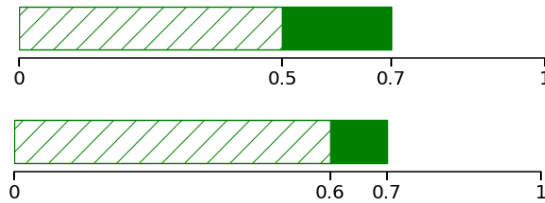


Figure 6: Two  $\epsilon$ -close lifebars, with  $\epsilon = 0.1$ .

Let us apply this result to the Čech bundle filtrations. Let  $X$  and  $Y$  be two subsets of  $E = \mathbb{R}^n \times M(\mathbb{R}^m)$ . Suppose that the Hausdorff distance  $d_H(X, Y)$ , with respect to the norm  $\|\cdot\|_\gamma$ , is lower than  $\epsilon$ , meaning that the  $\epsilon$ -thickenings  $X^\epsilon$  and  $Y^\epsilon$  satisfy  $Y \subseteq X^\epsilon$  and  $X \subseteq Y^\epsilon$ . It



is then clear that the vector bundle filtrations are  $\epsilon$ -interleaved, and we can apply Theorem 3.3 to obtain the following result.

**Corollary 3.4.** *If two subsets  $X, Y \subset E$  satisfy  $d_H(X, Y) \leq \epsilon$ , then there exists an  $\epsilon$ -interleaving between the persistent cohomology modules of their corresponding Čech bundle filtrations which sends persistent Stiefel-Whitney classes on persistent Stiefel-Whitney classes.*

**Example 3.5.** In order to illustrate Corollary 3.4, consider the sets  $X'$  and  $Y'$  represented in Figure 7. They are noisy samples of the sets  $X$  and  $Y$  defined in Example 3.2. They contain 50 points each.

Figure 8 represents the barcodes of the Čech filtrations of the sets  $X'$  and  $Y'$ , together with the lifebar of the first persistent Stiefel-Whitney class of their corresponding Čech bundle filtrations. We see that they are close to the original descriptors of  $X$  and  $Y$  (Figure 5). Experimentally, we computed that the Hausdorff distances between  $X, X'$  and  $Y, Y'$  are approximately  $d_H(X, X') \approx 0.5$  and  $d_H(Y, Y') \approx 0.4$ . We observe that this is coherent with the lifebar of  $w_1(Y')$ , which is  $\epsilon$ -close to the lifebar of  $w_1(Y)$  with  $\epsilon \approx 0.3 \leq 0.4$ .



Figure 7: Representation of the sets  $X', Y' \subset \mathbb{R}^2 \times M(\mathbb{R}^2)$ . The black points correspond to the  $\mathbb{R}^2$ -coordinate, and the pink segments over them correspond to the orientation of the  $M(\mathbb{R}^2)$ -coordinate.

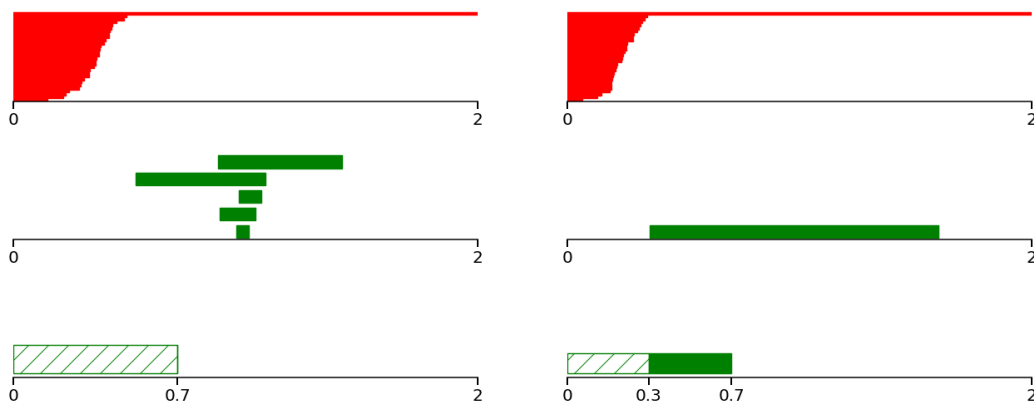


Figure 8: Left:  $H^0$  and  $H^1$  barcodes of  $X'$  and lifebar of  $w_1(X')$ . Right: same for  $Y'$ . Only bars of length larger than 0.05 are represented.

### 3.4 Consistency

In this subsection we describe a setting where the persistent Stiefel-Whitney classes  $w_i(X)$  of the Čech bundle filtration of a set  $X$  can be seen as consistent estimators of the Stiefel-Whitney classes of some underlying vector bundle.

Let  $\mathcal{M}_0$  be a compact  $\mathcal{C}^3$ -manifold, and  $u: \mathcal{M}_0 \rightarrow \mathcal{M} \subset \mathbb{R}^n$  an immersion. Suppose that  $\mathcal{M}_0$  is given a  $d$ -dimensional vector bundle structure  $p: \mathcal{M}_0 \rightarrow \mathcal{G}_d(\mathbb{R}^m)$ . Let  $E = \mathbb{R}^n \times M(\mathbb{R}^m)$ , and consider the set

$$\check{\mathcal{M}} = \{(u(x_0), P_{p(x_0)}), x_0 \in \mathcal{M}_0\} \subset E, \quad (12)$$

where  $P_{p(x_0)}$  denotes the orthogonal projection matrix onto the subspace  $p(x_0) \subset \mathbb{R}^m$ . The set  $\check{\mathcal{M}}$  is called the *lift* of  $\mathcal{M}_0$ . Consider the *lifting map* defined as

$$\begin{aligned} \check{u}: \mathcal{M}_0 &\longrightarrow \check{\mathcal{M}} \subset E \\ x_0 &\longmapsto (u(x_0), P_{p(x_0)}). \end{aligned} \quad (13)$$

We make the following assumption:  $\check{u}$  is an embedding. As a consequence,  $\check{\mathcal{M}}$  is a submanifold of  $E$ , and  $\mathcal{M}_0$  and  $\check{\mathcal{M}}$  are diffeomorphic.

The persistent cohomology of  $\check{\mathcal{M}}$  can be used to recover the cohomology of  $\mathcal{M}_0$ . To see this, select  $\gamma > 0$ , and denote by  $\text{reach}(\check{\mathcal{M}})$  the reach of  $\check{\mathcal{M}}$ , where  $E$  is endowed with the norm  $\|\cdot\|_\gamma$ . Since  $\check{\mathcal{M}}$  is a  $\mathcal{C}^2$ -submanifold,  $\text{reach}(\check{\mathcal{M}})$  is positive. Note that it depends on  $\gamma$ . Let  $V[\check{\mathcal{M}}] = (\check{\mathcal{M}}^t)_{t \geq 0}$  be the Čech set filtration of  $\check{\mathcal{M}}$  in the ambient space  $(E, \|\cdot\|_\gamma)$ , and let  $\mathbb{V}(\check{\mathcal{M}})$  be the corresponding persistent cohomology module. For every  $s, t \in [0, \text{reach}(\check{\mathcal{M}})]$  such that  $s \leq t$ , we know that the inclusion maps  $i_s^t: \check{\mathcal{M}}^s \hookrightarrow \check{\mathcal{M}}^t$  are homotopy equivalences. Hence the persistence module  $\mathbb{V}(\check{\mathcal{M}})$  is constant on the interval  $[0, \text{reach}(\check{\mathcal{M}})]$ , and is equal to the cohomology  $H^*(\check{\mathcal{M}}) = H^*(\mathcal{M}_0)$ .

Consider the Čech bundle filtration  $(V[\check{\mathcal{M}}], \mathbb{P})$  of  $\check{\mathcal{M}}$ . The following theorem shows that the persistent Stiefel-Whitney classes  $w_i^t(\check{\mathcal{M}})$  are also equal to the usual Stiefel-Whitney classes of the vector bundle  $(\mathcal{M}_0, p)$ .

**Theorem 3.6.** *Let  $\mathcal{M}_0$  be a compact  $\mathcal{C}^3$ -manifold,  $u: \mathcal{M}_0 \rightarrow \mathbb{R}^n$  an immersion and  $p: \mathcal{M}_0 \rightarrow \mathcal{G}_d(\mathbb{R}^m)$  a continuous map. Let  $\check{\mathcal{M}}$  be the lift of  $\mathcal{M}_0$  (Equation (12)) and  $\check{u}$  the lifting map (Equation (13)). Suppose that  $u$  is an embedding.*

*Let  $\gamma > 0$  and consider the Čech bundle filtration  $(V[\check{\mathcal{M}}], \mathbb{P})$  of  $\check{\mathcal{M}}$ . Its maximal filtration value is  $t_\gamma^{\max}(\check{\mathcal{M}}) = \frac{\sqrt{2}}{2}\gamma$ . Denote by  $w_i(\mathbb{P}) = (w_i^t(\mathbb{P}))_{t \in T}$  its persistent Stiefel-Whitney classes,  $i \in \llbracket 1, d \rrbracket$ . Denote also by  $i_0^t$  the inclusion  $\check{\mathcal{M}} \rightarrow \check{\mathcal{M}}^t$ , for  $t \in [0, \text{reach}(\check{\mathcal{M}})]$ .*

*Let  $t \geq 0$  be such that  $t < \min(\text{reach}(\check{\mathcal{M}}), t_\gamma^{\max}(\check{\mathcal{M}}))$ . Then the map  $i_0^t \circ \check{u}: \mathcal{M}_0 \rightarrow \check{\mathcal{M}}^t$  induces an isomorphism  $H^*(\mathcal{M}_0) \leftarrow H^*(\check{\mathcal{M}}^t)$  which maps the  $i^{\text{th}}$  persistent Stiefel-Whitney class  $w_i^t(\mathbb{P})$  of  $(V[\check{\mathcal{M}}], \mathbb{P})$  to the  $i^{\text{th}}$  Stiefel-Whitney class of  $(\mathcal{M}_0, p)$ .*

*Proof.* Consider the following commutative diagram, defined for every  $t < t_\gamma^{\max}(\check{\mathcal{M}})$ :

$$\begin{array}{ccccc} \mathcal{M}_0 & \xrightarrow{\check{u}} & \check{\mathcal{M}} & \xleftarrow{i_0^t} & \check{\mathcal{M}}^t \\ & \searrow p & & \swarrow p^t & \\ & & \mathcal{G}_d(\mathbb{R}^m) & & \end{array}$$

We obtain a commutative diagram in cohomology:

$$\begin{array}{ccccc} H^*(\mathcal{M}_0) & \xleftarrow{\check{u}^*} & H^*(\check{\mathcal{M}}) & \xleftarrow{(i_0^t)^*} & H^*(\check{\mathcal{M}}^t) \\ & \swarrow p^* & & \searrow (p^t)^* & \\ & & H^*(\mathcal{G}_d(\mathbb{R}^m)) & & \end{array}$$

Since  $t < \text{reach}(\check{\mathcal{M}})$ , the map  $(i_0^t)^*$  is an isomorphism. So is  $\check{u}^*$  since  $\check{u}$  is an embedding. As a consequence, the map  $i_0^t \circ \check{u}$  induces an isomorphism  $H^*(\mathcal{M}_0) \simeq H^*(\check{\mathcal{M}}^t)$ .

Let  $w_i$  denotes the  $i^{\text{th}}$  Stiefel-Whitney class of  $\mathcal{G}_d(\mathbb{R}^m)$ . By definition, the  $i^{\text{th}}$  Stiefel-Whitney class of  $(\mathcal{M}_0, p)$  is  $p^*(w_i)$ , and the  $i^{\text{th}}$  persistent Stiefel-Whitney class of  $(V[\check{\mathcal{M}}], \mathbb{P})$  is  $w_i^t(\mathbb{P}) = (p^t)^*(w_i)$ . By commutativity of the diagram, we obtain  $p^*(w_i) = (p^t)^*(w_i)$  under the identification  $H^*(\mathcal{M}_0) \simeq H^*(\check{\mathcal{M}}^t)$ .  $\square$

We deduce an estimation result.

**Corollary 3.7.** *Let  $X \subset E$  be any subset such that  $d_H(X, \check{\mathcal{M}}) \leq \frac{1}{17}\text{reach}(\check{\mathcal{M}})$ . Define  $\epsilon = d_H(X, \check{\mathcal{M}})$ . Then for every  $t \in [4\epsilon, \text{reach}(\check{\mathcal{M}}) - 3\epsilon)$ , the composition of inclusions  $\mathcal{M}_0 \hookrightarrow \check{\mathcal{M}} \hookrightarrow X^t$  induces an isomorphism  $H^*(\mathcal{M}_0) \leftarrow H^*(X^t)$  which sends the  $i^{\text{th}}$  persistent Stiefel-Whitney class  $w_i^t(X)$  of the Čech bundle filtration of  $X$  to the  $i^{\text{th}}$  Stiefel-Whitney class of  $(\mathcal{M}_0, p)$ .*

*Proof.* This is a consequence of Theorems 3.3 and 3.6 previously stated, and Proposition 3.1 of Chazal et al. (2009).  $\square$

As a consequence of this corollary, on the set  $[4\epsilon, \text{reach}(\check{\mathcal{M}}) - 3\epsilon)$ , the  $i^{\text{th}}$  persistent Stiefel-Whitney class of the Čech bundle filtration of  $X$  is zero if and only if the  $i^{\text{th}}$  Stiefel-Whitney class of  $(\mathcal{M}_0, p)$  is.

**Example 3.8.** In order to illustrate Corollary 3.7, consider the torus and the Klein bottle, immersed in  $\mathbb{R}^3$  as in Figure 9.



Figure 9: Immersion of the torus and the Klein bottle in  $\mathbb{R}^3$ .

Let them be endowed with their normal bundles. They can be seen as submanifolds  $\check{\mathcal{M}}, \check{\mathcal{M}}'$  of  $\mathbb{R}^3 \times M(\mathbb{R}^3)$ . We consider two samples  $X, X'$  of  $\check{\mathcal{M}}, \check{\mathcal{M}}'$ , represented in Figure 10. They contain respectively 346 and 1489 points. We computed experimentally the Hausdorff distances  $d_H(X, \check{\mathcal{M}}) \approx 0.6$  and  $d_H(X', \check{\mathcal{M}}') \approx 0.45$ , with respect to the norm  $\|\cdot\|_\gamma$  where  $\gamma = 1$ .

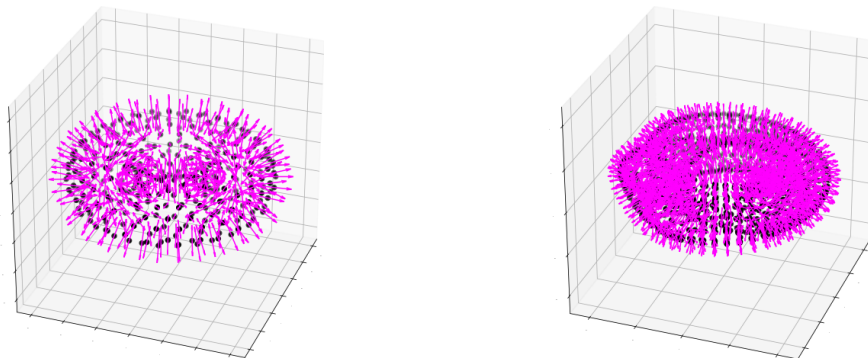


Figure 10: Samples  $X$  and  $X'$  of  $\check{\mathcal{M}}$  and  $\check{\mathcal{M}}'$ . The black points corresponds to the  $\mathbb{R}^3$ -coordinate, and the pink arrows over them correspond to the orientation of the  $M(\mathbb{R}^3)$ -coordinate.

Figure 11 represents the barcodes of the persistent cohomology of  $X$  and  $X'$ , and the lifebars of their first persistent Stiefel-Whitney classes  $w_1(X)$  and  $w_1(X')$ . Observe that  $w_1(X)$  is always zero, while  $w_1(X')$  is nonzero for  $t \geq 0.3$ . This is an indication that  $\check{\mathcal{M}}$ , the underlying manifold of  $X$ , is orientable, while  $\check{\mathcal{M}}'$  is not. Lemma 3.9, stated below, justifies this assertion. Therefore, one interprets these lifebars as follows:  $X$  is sampled on an orientable manifold, while  $X'$  is sampled on a non-orientable one.

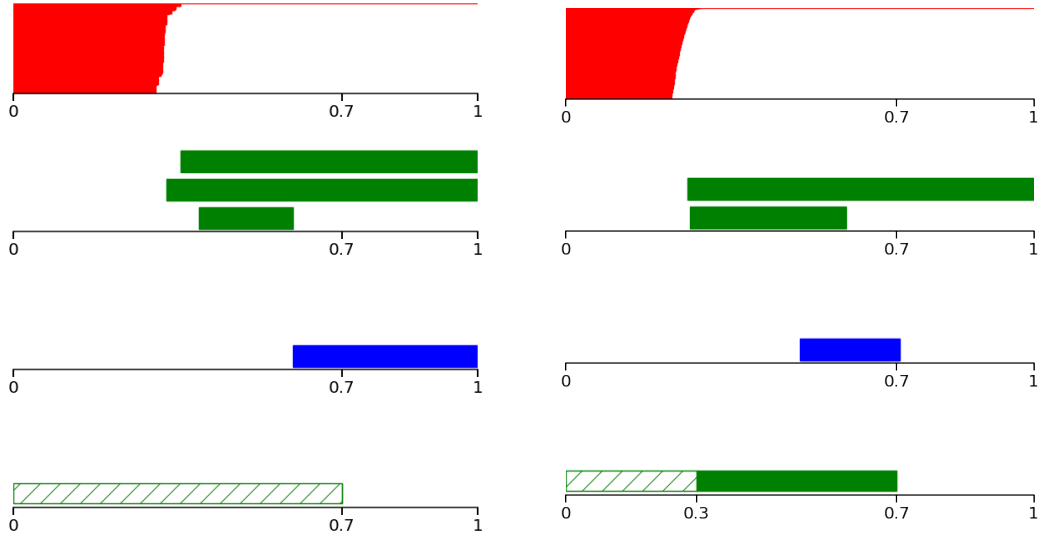


Figure 11: **Left:**  $H^0$ ,  $H^1$  and  $H^2$  barcodes of  $X$  and lifebar of  $w_1(X)$ . **Right:** same for  $X'$ . Only bars of length larger than 0.2 are represented.

**Lemma 3.9.** *Let  $\mathcal{M}_0 \rightarrow \mathcal{M}$  be an immersion of a manifold  $\mathcal{M}_0$  in a Euclidean space. Then  $\mathcal{M}_0$  is orientable if and only if the first Stiefel-Whitney class of its normal bundle is zero.*

*Proof.* Let  $\tau$  and  $\nu$  denote the tangent and normal bundles of  $\mathcal{M}_0$ . The Whitney sum  $\tau \oplus \nu$  is a trivial bundle, hence its first Stiefel-Whitney class is  $w_1(\tau \oplus \nu) = 0$ . Using Axioms 1 and 3 of the Stiefel-Whitney classes, we obtain

$$\begin{aligned} w_1(\tau \oplus \nu) &= w_1(\tau) \smile w_0(\nu) + w_0(\tau) \smile w_1(\nu) \\ &= w_1(\tau) \smile 1 + 1 \smile w_1(\nu) \\ &= w_1(\tau) + w_1(\nu). \end{aligned}$$

Therefore,  $w_1(\tau) = w_1(\nu)$ , hence  $w_1(\tau)$  is zero if and only if  $w_1(\nu)$  is zero. Besides, it is known that the first Stiefel-Whitney class of the tangent bundle of a manifold is zero if and only if the manifold is orientable. We deduce the result.  $\square$

## 4 Computation of persistent Stiefel-Whitney classes

In order to build an effective algorithm to compute the persistent Stiefel-Whitney classes, we have to find an equivalent formulation in terms of simplicial cohomology. We will make use of the well-known technique of simplicial approximation, as described in Subsect. 2.2.

### 4.1 Simplicial approximation to Čech bundle filtrations

Let  $X$  be a subset of  $E = \mathbb{R}^n \times M(\mathbb{R}^m)$ . Let us recall Definition 3.3: the Čech bundle filtration associated to  $X$  is the vector bundle filtration  $(\mathbb{X}, \mathbb{p})$  whose underlying filtration is the Čech filtration  $\mathbb{X} = (X^t)_{t \in T}$ , with  $T = [0, t_\gamma^{\max}(X))$ , and whose maps  $\mathbb{p} = (p^t)_{t \in T}$  are given by the following composition, as in Equation (8):

$$\begin{array}{ccc} X^t & \xrightarrow{\text{proj}_2} & M(\mathbb{R}^m) \setminus \text{med}(\mathcal{G}_d(\mathbb{R}^m)) \xrightarrow{\text{proj}(\cdot, \mathcal{G}_d(\mathbb{R}^m))} \mathcal{G}_d(\mathbb{R}^m). \\ & \searrow & \nearrow \\ & & p^t \end{array}$$

Let  $t \in T$ . The aim of this subsection is to describe a simplicial approximation to  $p^t: X^t \rightarrow \mathcal{G}_d(\mathbb{R}^m)$ . To do so, let us fix a triangulation  $L$  of  $\mathcal{G}_d(\mathbb{R}^m)$ . It comes with a homeomorphism  $h: \mathcal{G}_d(\mathbb{R}^m) \rightarrow |L|$ . We shall now triangulate the thickenings  $X^t$  of the Čech set filtration. The thickening  $X^t$  is a subset of the metric space  $(E, \|\cdot\|_\gamma)$  which consists in a union of closed balls centered around points of  $X$ :

$$X^t = \bigcup_{x \in X} \overline{\mathcal{B}}_\gamma(x, t),$$

where  $\overline{\mathcal{B}}_\gamma(x, t)$  denotes the closed ball of center  $x$  and radius  $t$  for the norm  $\|\cdot\|_\gamma$ . Let  $\mathcal{U}^t$  denote the cover  $\{\overline{\mathcal{B}}_\gamma(x, t), x \in X\}$  of  $X^t$ , and let  $\mathcal{N}(\mathcal{U}^t)$  be its nerve. By the nerve theorem for convex closed covers (Boissonnat et al., 2018, Theorem 2.9), the simplicial complex  $\mathcal{N}(\mathcal{U}^t)$  is homotopy equivalent to its underlying set  $X^t$ . That is to say, there exists a continuous map  $g^t: |\mathcal{N}(\mathcal{U}^t)| \rightarrow X^t$  which is a homotopy equivalence.

As a consequence, in cohomological terms, the map  $p^t: X^t \rightarrow \mathcal{G}_d(E)$  is equivalent to the map  $q^t$  defined as  $q^t = h \circ p^t \circ g^t$ .

$$\begin{array}{ccc} X^t & \xrightarrow{p^t} & \mathcal{G}_d(\mathbb{R}^m) \\ g^t \uparrow & & \downarrow h \\ |\mathcal{N}(\mathcal{U}^t)| & \xrightarrow{q^t} & |L| \end{array} \quad (14)$$

This gives a way to compute the induced map  $(p^t)^*: H^*(X^t) \leftarrow H^*(\mathcal{G}_d(\mathbb{R}^m))$  algorithmically:

- Subdivide  $\mathcal{N}(\mathcal{U}^t)$  until  $q^t$  satisfies the star condition.
- Choose a simplicial approximation  $f^t$  to  $q^t$ .
- Compute the induced map between simplicial cohomology groups  $(f^t)^*: H^*(\mathcal{N}(\mathcal{U}^t)) \leftarrow H^*(L)$ .

By correspondence between simplicial and singular cohomology, the map  $(f^t)^*$  corresponds to  $(p^t)^*$ . Hence the problem of computing  $(p^t)^*$  is solved, if it were not for the following issue: in practice, the map  $g^t: |\mathcal{N}(\mathcal{U}^t)| \rightarrow X^t$  given by the nerve theorem is not explicit. The rest of this subsection is devoted to showing that  $g^t$  can be chosen canonically as the *shadow map*.

**Shadow map.** We still consider  $X^t$ , the corresponding cover  $\mathcal{U}^t$  and its nerve  $\mathcal{N}(\mathcal{U}^t)$ . The underlying vertex set of the simplicial complex  $\mathcal{N}(\mathcal{U}^t)$  is the set  $X$  itself. The shadow map  $g^t: |\mathcal{N}(\mathcal{U}^t)| \rightarrow X^t$  is defined as follows: for every simplex  $\sigma = [x_0, \dots, x_p] \in \mathcal{N}(\mathcal{U}^t)$  and every point  $\sum_{i=0}^p \lambda_i x_i$  of  $|\sigma|$  written in barycentric coordinates, associate the point  $\sum_{i=0}^p \lambda_i x_i$  of  $E$ :

$$g^t: \sum_{i=0}^p \lambda_i x_i \in |\sigma| \mapsto \sum_{i=0}^p \lambda_i x_i \in E.$$

The following lemma states that this map is a homotopy equivalence. We are not aware whether the general position assumption can be removed.

**Lemma 4.1.** *Suppose that  $X$  is finite and in general position. Then the shadow map  $g^t: |\mathcal{N}(\mathcal{U}^t)| \rightarrow X^t$  is a homotopy equivalence. Consequently, it induces an isomorphism  $(g^t)^*: H^*(|\mathcal{N}(\mathcal{U}^t)|) \leftarrow H^*(X^t)$ .*

*Proof.* Recall that  $\mathcal{U}^t = \{\overline{\mathcal{B}}_\gamma(x, t), x \in X\}$ . Let us consider a smaller cover. For every  $x \in X$ , let  $\text{Vor}(x)$  denote the Voronoi cell of  $x$  in the ambient metric space  $(E, \|\cdot\|_\gamma)$ , and define

$$\mathcal{V}^t = \{\overline{\mathcal{B}}_\gamma(x, t) \cap \text{Vor}(x), x \in X\}.$$

The set  $\mathcal{V}^t$  is a cover of  $X^t$ , and its nerve  $\mathcal{N}(\mathcal{V}^t)$  is known as the Delaunay complex. Let  $h^t: |\mathcal{N}(\mathcal{V}^t)| \rightarrow X^t$  denote the shadow map of  $\mathcal{N}(\mathcal{V}^t)$ . The Delaunay complex is a subcomplex of the Čech complex, hence we can consider the following diagram:

$$\begin{array}{ccc}
& \xrightarrow{h^t} & \\
|\mathcal{N}(\mathcal{V}^t)| & \xleftrightarrow{\quad} & |\mathcal{N}(\mathcal{U}^t)| \xrightarrow{g^t} X^t.
\end{array}$$

Now, Edelsbrunner (1993, Theorem 3.2) has proven that the shadow map  $h^t: |\mathcal{N}(\mathcal{V}^t)| \rightarrow X^t$  is a homotopy equivalence (it is required here that  $X$  is in general position). Moreover, we know from Bauer and Edelsbrunner (2017, Theorem 5.10) that  $\mathcal{N}(\mathcal{U}^t)$  collapses to  $\mathcal{N}(\mathcal{V}^t)$ . Therefore the inclusion  $|\mathcal{N}(\mathcal{V}^t)| \hookrightarrow |\mathcal{N}(\mathcal{U}^t)|$  also is a homotopy equivalence. By the 2-out-of-3 property of homotopy equivalences, we conclude that  $g^t$  is a homotopy equivalence.  $\square$

## 4.2 A sketch of algorithm

Suppose that we are given a finite set  $X \subset E = \mathbb{R}^n \times M(\mathbb{R}^m)$ . Choose  $d \in \llbracket 1, n-1 \rrbracket$  and  $\gamma > 0$ . Consider the Čech bundle filtration of dimension  $d$  of  $X$ . Let  $T = [0, t_\gamma^{\max}(X))$ ,  $t \in T$  and  $i \in \llbracket 1, d \rrbracket$ . From the previous discussion we can infer an algorithm to solve the following problem:

Compute the persistent Stiefel-Whitney class  $w_i^t(X)$  of the Čech bundle filtration of  $X$ , using a cohomology computation software.

Denote:

- $\mathbb{X} = (X^t)_{t \geq 0}$  the Čech set filtration of  $X$ ,
- $\mathbb{S}$  the Čech simplicial filtration of  $X$ , and  $g^t: |S^t| \rightarrow X^t$  the shadow map,
- $L$  a triangulation of  $\mathcal{G}_d(\mathbb{R}^m)$  and  $h: \mathcal{G}_d(\mathbb{R}^m) \rightarrow |L|$  a homeomorphism,
- $(\mathbb{X}, \mathfrak{p})$  the Čech bundle filtration of  $X$ ,
- $(\mathbb{V}, \mathfrak{v})$  the persistent cohomology module of  $\mathbb{X}$ ,
- $w_i \in H^i(\mathcal{G}_d(\mathbb{R}^m))$  the  $i^{\text{th}}$  Stiefel-Whitney class of the Grassmannian.

Let  $t \in T$  and consider the map  $q^t$ , as defined in Equation (14):

$$\begin{array}{ccccccc}
& & & \xrightarrow{q^t} & & & \\
|S^t| & \xrightarrow{g^t} & X^t & \xrightarrow{p^t} & \mathcal{G}_d(\mathbb{R}^m) & \xrightarrow{h} & |L|.
\end{array}$$

We propose the following algorithm:

- Subdivide barycentrically  $S^t$  until  $q^t$  satisfies the star condition. Denote  $k$  the number of subdivisions needed.
- Consider a simplicial approximation  $f^t: \text{sub}^k(S^t) \rightarrow L$  to  $q^t$ .
- Compute the class  $(f^t)^*(w_i)$ .

The output  $(f^t)^*(w_i)$  is equal to the persistent Stiefel-Whitney class  $w_i^t(X)$  at time  $t$ , seen in the simplicial cohomology group  $H^i(S^t) = H^i(\text{sub}^k(S^t))$ . In the following section, we gather some technical details needed to implement this algorithm in practice.

Note that this also gives a way to compute the lifebar of  $w_i(X)$ . This bar is determined by the value  $t^\dagger = \inf\{t \in T, w_i(X) \neq 0\}$ . This quantity can be approximated by dichotomic search, by computing the classes  $w_i^t(X)$  for several values of  $t$ . We point out that, in order to compute the value  $t^\dagger$ , there may exist a better algorithm than evaluating the class  $w_i^t(X)$  several times.

Let us describe briefly such an algorithm when  $i = 1$ , that is, when the first persistent Stiefel-Whitney class  $w_1(X)$  is to be computed. First, we remind the reader that the first cohomology group  $H^1(\mathcal{G}_d(\mathbb{R}^m))$  of the Grassmannian is generated by one element, the first Stiefel-Whitney

class  $w_1$ . For any  $t \in T$ , consider the map  $p^t$  as above, and the map induced in cohomology,  $(p^t)^*: H^1(X^t) \leftarrow H^1(\mathcal{G}_d(\mathbb{R}^m))$ . Since  $w_1^t(X) = (p^t)^*(w_1)$ , and since  $H^1(\mathcal{G}_d(\mathbb{R}^m))$  has dimension 1, we have that  $w_1^t(X)$  is nonzero if and only if  $\text{rank}(p^t)^*$  is nonzero.

Next, let  $C(p^t)$  denotes the *mapping cone* of  $p^t: X^t \rightarrow \mathcal{G}_d(\mathbb{R}^m)$ . This is a usual construction in algebraic topology. In a few words, the mapping cone of a map is a topological space that contains information about the map. The mapping cone  $C(p^t)$  comes with a long exact sequence

$$\dots \longrightarrow H^k(X^t) \longrightarrow H^{k+1}(C(p^t)) \longrightarrow H^{k+1}(\mathcal{G}_d(\mathbb{R}^m)) \longrightarrow H^{k+1}(X^t) \longrightarrow \dots$$

from which we deduce the formula

$$\text{rank}(p^t)^* = \sum_{k=1}^{+\infty} (-1)^k \left( \dim H^k(X^t) - \dim H^{k+1}(C(p^t)) + \dim H^{k+1}(\mathcal{G}_d(\mathbb{R}^m)) \right)$$

On the simplicial side, is not complicated to build a triangulation  $C^t$  of the mapping cone  $C(p^t)$ , nor to build a simplicial filtration  $(C^t)_{t \in T}$  of  $(C(p^t))_{t \in T}$ . It relies on finding simplicial approximations to the maps  $q^t: S^t \rightarrow L$ , as in the proof of Hatcher (2002, Theorem 2C.5). We point out that, just as in the previous algorithm, we may have to apply barycentric subdivisions to  $S^t$  here. Now, the previous formula translates in simplicial cohomology as

$$\text{rank}(q^t)^* = \sum_{k=1}^{+\infty} (-1)^k \left( \dim H^k(S^t) - \dim H^{k+1}(C^t) + \dim H^{k+1}(L) \right).$$

All these terms can be computed efficiently by applying the persistent homology algorithm to the filtrations  $(S^t)_{t \in T}$ ,  $(C^t)_{t \in T}$  and  $(L)_{t \in T}$ . Finally, we identify the value  $t^\dagger$  as the first value of  $t \in T$  such that  $\text{rank}(q^t)^*$  is nonzero.

## 5 An algorithm when $d = 1$

Even though the last sections described a theoretical way to compute the persistent Stiefel-Whitney classes, some concrete issues are still to be discussed:

- verifying that the star condition is satisfied,
- the Grassmann manifold has to be triangulated,
- in practice, the Vietoris-Rips filtration is preferred to the Čech filtration,
- the parameter  $\gamma$  has to be tuned.

The following subsections will elucidate these points. Concerning the first one, we are not aware of a computational-explicit process to triangulate the Grassmann manifolds  $\mathcal{G}_d(\mathbb{R}^m)$ , except when  $d = 1$ , which corresponds to the projective spaces  $\mathcal{G}_1(\mathbb{R}^m)$ . We shall then restrict to the case  $d = 1$ . Note that, in this case, the only nonzero Stiefel-Whitney classes are the first two (by Axiom 1 of Stiefel-Whitney classes). Since  $w_0$  is always equal to 1, the only class to estimate is  $w_1$ .

### 5.1 The star condition in practice

Let us get back to the context of Subsect. 2.2:  $K, L$  are two simplicial complexes,  $K$  is finite, and  $g: |K| \rightarrow |L|$  is a continuous map. We have seen that finding a simplicial approximation to  $g$  reduces to finding a small enough barycentric subdivision  $\text{sub}^n(K)$  of  $K$  such that  $g: |\text{sub}^n(K)| \rightarrow |L|$  satisfies the star condition, that is, for every vertex  $v$  of  $\text{sub}^n(K)$ , there exists a vertex  $w$  of  $L$  such that

$$g(|\overline{\text{St}}(v)|) \subseteq |\text{St}(w)|.$$

In practice, one can compute the closed star  $\overline{\text{St}}(v)$  from the finite simplicial complex  $\text{sub}^n(K)$ . However, computing  $g(|\overline{\text{St}}(v)|)$  requires to evaluate  $g$  on the infinite set  $|\overline{\text{St}}(v)|$ . In order to reduce the problem to a finite number of evaluations of  $g$ , we shall consider a related property that we call the *weak star condition*.

**Definition 5.1.** A map  $g: |K| \rightarrow |L|$  between geometric realizations of simplicial complexes  $K$  and  $L$  satisfies the *weak star condition* if for every vertex  $v$  of  $\text{sub}^n(K)$ , there exists a vertex  $w$  of  $L$  such that

$$|g(\overline{\text{St}}(v)^0)| \subseteq |\text{St}(w)|,$$

where  $\overline{\text{St}}(v)^0$  denotes the 0-skeleton of  $\overline{\text{St}}(v)$ , i.e. its vertices.

Observe that the practical verification of the condition  $|g(\overline{\text{St}}(v)^0)| \subseteq |\text{St}(w)|$  requires only a finite number of computations. Indeed, one just has to check whether every neighbor  $v'$  of  $v$  in the graph  $K^1$ ,  $v$  included, satisfies  $g(v') \in |\text{St}(w)|$ . The following lemma rephrases this condition by using the face map  $\mathcal{F}_L: |L| \rightarrow L$ . We remind the reader that the face map is defined by the relation  $x \in \mathcal{F}_L(x)$  for all  $x \in |L|$ .

**Lemma 5.1.** *The map  $g$  satisfies the weak star condition if and only if for every vertex  $v$  of  $K$ , there exists a vertex  $w$  of  $L$  such that for every neighbor  $v'$  of  $v$  in  $K^1$ , we have*

$$w \in \mathcal{F}_L(g(v')).$$

*Proof.* Let us show that the assertion “ $w \in \mathcal{F}_L(g(v'))$ ” is equivalent to “ $g(v') \in |\text{St}(w)|$ ”. Recall that the open star  $\text{St}(w)$  consists of the simplices of  $L$  that contain  $w$ . Moreover, the geometric realization  $|\text{St}(w)|$  is the union of the  $|\sigma|$  for  $\sigma \in \text{St}(w)$ . As a consequence,  $g(v')$  belongs to  $|\text{St}(w)|$  if and only if it belongs to  $|\sigma|$  for some simplex  $\sigma \in L$  that contains  $w$ . Equivalently, the face map  $\mathcal{F}_L(g(v'))$  contains  $w$ .  $\square$

Suppose that  $g$  satisfies the weak star condition. Let  $f: K^0 \rightarrow L^0$  be a map, between vertex sets, such that for every  $v \in K^0$ ,

$$|g(\overline{\text{St}}(v)^0)| \subseteq |\text{St}(f(v))|.$$

According to the proof of Lemma 5.1, an equivalent formulation of this condition is: for all neighbor  $v'$  of  $v$  in  $K^1$ ,

$$f(v) \in \mathcal{F}_L(g(v')). \quad (15)$$

Such a map is called a *weak simplicial approximation to  $g$* . It plays a similar role as the simplicial approximations to  $g$ .

**Lemma 5.2.** *If  $f: K^0 \rightarrow L^0$  is a weak simplicial approximation to  $g: |K| \rightarrow |L|$ , then  $f$  is a simplicial map.*

*Proof.* Let  $\sigma = [v_0, \dots, v_n]$  be a simplex of  $K$ . We have to show that  $f(\sigma) = [f(v_0), \dots, f(v_n)]$  is a simplex of  $L$ . Note that each closed star  $\overline{\text{St}}(v_i)$  contains  $\sigma$ . Therefore each  $|g(\overline{\text{St}}(v_i)^0)|$  contains  $|g(\sigma^0)| = \{g(v_0), \dots, g(v_n)\}$ . Using the weak simplicial approximation property of  $f$ , we deduce that each  $|\text{St}(f(v_i))|$  contains  $\{g(v_0), \dots, g(v_n)\}$ . Using Lemma 5.3 stated below, we obtain that  $[f(v_0), \dots, f(v_n)]$  is a simplex of  $L$ .  $\square$

**Lemma 5.3** (Hatcher, 2002, Lemma 2C.2). *Let  $w_0, \dots, w_n$  be vertices of a simplicial complex  $L$ . Then  $\bigcap_{i=0}^n \text{St}(w_i) \neq \emptyset$  if and only if  $[w_0, \dots, w_n]$  is a simplex of  $L$ .*

As one can see from the definitions, the weak star condition is weaker than the star condition. Consequently, the simplicial approximation theorem admits the following corollary.



**Corollary 5.4.** *Consider two simplicial complexes  $K, L$  with  $K$  finite, and let  $g: |K| \rightarrow |L|$  be a continuous map. Then there exists  $n \geq 0$  such that  $g: |\text{sub}^n(K)| \rightarrow |L|$  satisfies the weak star condition.*

However, some weak simplicial approximations to  $g$  may not be simplicial approximations, and may not even be homotopic to  $g$ . Figure 12 gives such an example.

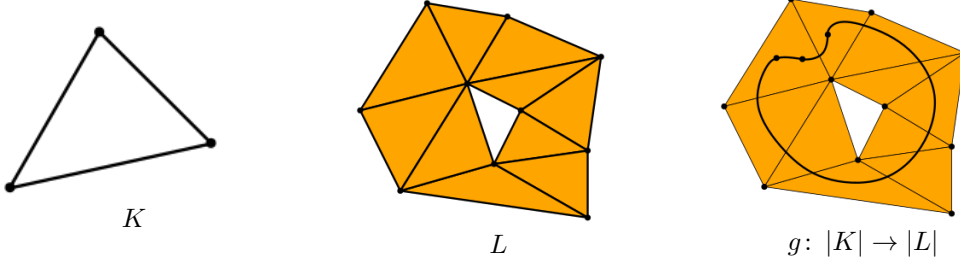


Figure 12: The map  $g$  is non-trivial but admits a weak simplicial approximation which is constant.

Fortunately, these two notions coincides under the star condition assumption:

**Proposition 5.5.** *Suppose that  $g$  satisfies the star condition. Then every weak simplicial approximation to  $g$  is a simplicial approximation.*

*Proof.* Let  $f$  be a weak simplicial approximation to  $g$ , and  $f'$  any simplicial approximation. Let us show that  $f$  and  $f'$  are contiguous simplicial maps. Let  $\sigma = [v_0, \dots, v_n]$  be a simplex of  $K$ . We have to show that  $[f(v_0), \dots, f(v_n), f'(v_0), \dots, f'(v_n)]$  is a simplex of  $L$ . As we have seen in the proof of Lemma 5.2, each  $|g(\overline{\text{St}}(v_i)^0)|$  contains  $\{g(v_0), \dots, g(v_n)\}$ . Therefore, by definition of weak simplicial approximations and simplicial approximations, each  $|\text{St}(f(v_i))|$  and  $|\text{St}(f'(v_i))|$  contains  $\{g(v_0), \dots, g(v_n)\}$ . We conclude by applying Lemma 5.3.  $\square$

Remark that the proof of this proposition can be adapted to obtain the following fact: any two weak simplicial approximations are equivalent—as well as any two simplicial approximations.

Let us comment Proposition 5.5. If  $K$  is subdivided enough, then every weak simplicial approximation to  $g$  is homotopic to  $g$ . We face the following problem in practice: the number of subdivisions needed by the star condition is not known. In order to work around this problem, we propose to subdivide the complex  $K$  until it satisfies the weak star condition, and then use a weak simplicial approximation to  $g$ . However, such a weak simplicial approximation may not be homotopic to  $g$ , and our algorithm would output a wrong result.

To close this subsection, we state a lemma that gives a quantitative idea of the number of subdivisions needed by the star condition. We say that a Lebesgue number for an open cover  $\mathcal{U}$  of a compact metric space  $X$  is a positive number  $\epsilon$  such that every subset of  $X$  with diameter less than  $\epsilon$  is included in some element of the cover  $\mathcal{U}$ .

**Lemma 5.6.** *Let  $|K|, |L|$  be endowed with metrics. Suppose that  $g: |K| \rightarrow |L|$  is  $l$ -Lipschitz with respect to these metrics. Let  $\epsilon$  be a Lebesgue number for the open cover  $\{|\text{St}(w)|, w \in L\}$  of  $|L|$ . Let  $p$  be the dimension of  $K$  and  $D$  an upper bound on the diameter of its faces. Then for  $n > \log(\frac{Dl}{\epsilon}) / \log(\frac{p+1}{p})$ , the map  $g: |\text{sub}^n(K)| \rightarrow |L|$  satisfies the star condition.*

*Proof.* The map  $g$  satisfies the star condition if for every vertex  $v$  of  $K$ , there exists a vertex  $w$  of  $L$  such that  $g(|\overline{\text{St}}(v)|) \subseteq |\text{St}(w)|$ . Since the cover  $\{|\text{St}(w)|, w \in L\}$  admits  $\epsilon$  as a Lebesgue number, it is enough for  $v$  to satisfy the following inequality:

$$\text{diam}(g(|\overline{\text{St}}(v)|)) < \epsilon. \quad (16)$$

Since  $g$  is  $l$ -Lipschitz, we have  $\text{diam}(g(|\overline{\text{St}}(v)|)) \leq l \cdot \text{diam}(|\overline{\text{St}}(v)|)$ . Using the hypothesis  $\text{diam}(|\overline{\text{St}}(v)|) \leq D$ , Equation (16) leads to the condition  $Dl < \epsilon$ . Now, we use the fact that a

barycentric subdivision reduces the diameter of each face by a factor  $\frac{p}{p+1}$ . After  $n$  barycentric subdivision, the last inequality rewrites  $\left(\frac{p}{p+1}\right)^n Dl < \epsilon$ . It admits  $n > \log\left(\frac{Dl}{\epsilon}\right) / \log\left(\frac{p+1}{p}\right)$  as a solution.  $\square$

## 5.2 Triangulating the projective spaces

As we described in Subsect. 5.1, the algorithm we propose rests on a triangulation  $L$  of the Grassmannian  $\mathcal{G}_1(\mathbb{R}^m)$ , together with the map  $\mathcal{F}_L \circ h: \mathcal{G}_1(\mathbb{R}^m) \rightarrow L$ , where  $h: \mathcal{G}_1(\mathbb{R}^m) \rightarrow |L|$  is a homeomorphism and  $\mathcal{F}_L: \mathcal{G}_1(\mathbb{R}^m) \rightarrow L$  is the face map. In the following, we also call  $\mathcal{F} := \mathcal{F}_L \circ h$  the face map.

We shall use the triangulation of the projective space  $\mathcal{G}_1(\mathbb{R}^m)$  by von Kühnel (1987). It uses the fact that the quotient of the sphere  $\mathbb{S}_{m-1}$  by the antipodal relation gives  $\mathcal{G}_1(\mathbb{R}^m)$ . Let  $\Delta^m$  denote the standard  $m$ -simplex,  $v_0, \dots, v_m$  its vertices, and  $\partial\Delta^m$  its boundary. The simplicial complex  $\partial\Delta^m$  is a triangulation of the sphere  $\mathbb{S}_{m-1}$ . Denote its first barycentric subdivision as  $\text{sub}^1(\partial\Delta^m)$ . The vertices of  $\text{sub}^1(\partial\Delta^m)$  are in bijection with the non-empty proper subsets of  $\{v_0, \dots, v_m\}$ . Consider the equivalence relation on these vertices which associates a vertex to its complement. The quotient simplicial complex under this relation,  $L$ , is a triangulation of  $\mathcal{G}_1(\mathbb{R}^m)$ . Figures 13 and 14 represent this construction.

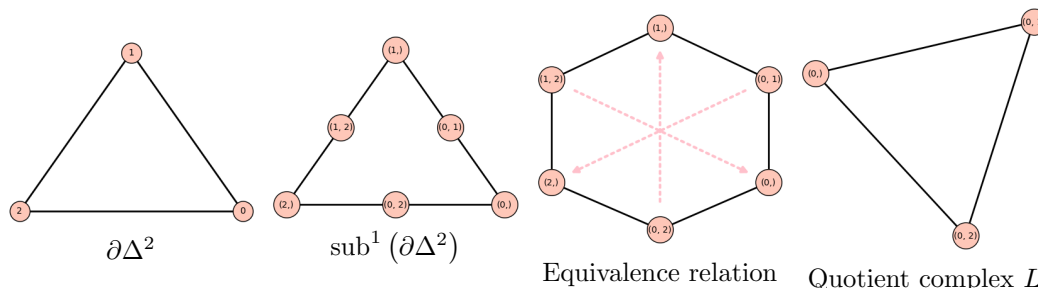


Figure 13: Triangulating  $\mathcal{G}_1(\mathbb{R}^2)$ .

Let us now describe how to define the homeomorphism  $h: \mathcal{G}_1(\mathbb{R}^m) \rightarrow |L|$ . First, embed  $\Delta^m$  in  $\mathbb{R}^{m+1}$  via  $v_i \mapsto (0, \dots, 0, 1, 0, \dots)$ , where 1 sits at the  $i^{\text{th}}$  coordinate. Its image lies on a  $m$ -dimensional affine subspace  $P$ , with origin being the barycenter of  $v_0, \dots, v_m$ . Seen in  $P$ , the vertices of  $\Delta^m$  now belong to the sphere centered at the origin and of radius  $\sqrt{\frac{m}{m+1}}$  (see Figure 14). Let us denote this sphere as  $\mathbb{S}_{m-1}$ . Next, subdivide barycentrically  $\partial\Delta^m$  once, and project each vertex of  $\text{sub}^1(\partial\Delta^m)$  on  $\mathbb{S}_{m-1}$ . By taking the convex hulls of its faces, we now see  $|\text{sub}^1(\partial\Delta^m)|$  as a subset of  $P$ . Define an application  $p: \mathbb{S}_{m-1} \rightarrow |\text{sub}^1(\partial\Delta^m)|$  as follows: for every  $x \in \mathbb{S}_{m-1}$ , the image  $p(x)$  is the unique intersection point between the segment  $[0, x]$  and the set  $|\text{sub}^1(\partial\Delta^m)|$ . The application  $p$  can also be seen as the inverse function of the projection on  $\mathbb{S}_{m-1}$ , written  $\text{proj}_{\mathbb{S}_{m-1}}: |\text{sub}^1(\partial\Delta^m)| \rightarrow \mathbb{S}_{m-1}$ . As a consequence, we can factorize  $p: \mathbb{S}_{m-1} \rightarrow |\text{sub}^1(\partial\Delta^m)|$  as

$$h: (\mathbb{S}_{m-1}/\sim) \rightarrow (|\text{sub}^1(\partial\Delta^m)|/\sim).$$

Using Lemma 5.7 stated below, we can identify these spaces with

$$h: \mathcal{G}_1(\mathbb{R}^m) \rightarrow |L|,$$

giving the desired triangulation.

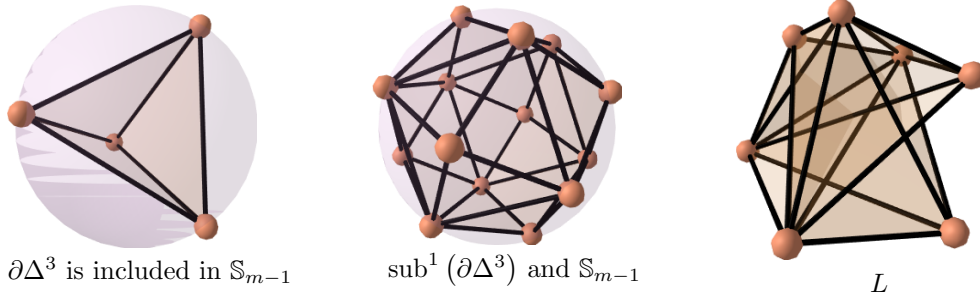


Figure 14: Triangulating  $\mathcal{G}_1(\mathbb{R}^3)$ .

**Lemma 5.7.** *For any vertex  $x \in \text{sub}^1(\partial\Delta^m)$ , denote by  $|x|$  its embedding in  $P$ . Let  $-|x|$  denote the image of  $|x|$  by the antipodal relation on  $\mathbb{S}_{m-1}$ . Denote by  $y$  the image of  $x$  by the relation on  $\text{sub}^1(\partial\Delta^m)$ . Then  $y = -|x|$ .*

*More generally, pulling back the antipodal relation onto  $|\text{sub}^1(\partial\Delta^m)|$  via  $p$  gives the relation we defined on  $\text{sub}^1(\partial\Delta^m)$ .*

*Proof.* Pick a vertex  $x$  of  $\text{sub}^1(\partial\Delta^m)$ . It can be described as a proper subset  $\{v_i, i \in I\}$  of the vertex set  $(\partial\Delta^m)^0 = \{v_0, \dots, v_m\}$ , where  $I \subset \llbracket 0, m \rrbracket$ . According to the relation on  $(\partial\Delta^m)$ , the vertex  $x$  is in relation with the vertex  $y$  described by the proper subset  $\{v_i, i \in I^c\}$ . The point  $x$  can be written in barycentric coordinates as  $\frac{1}{\text{card}(I)} \sum_{i \in I} |v_i|$ . Seen in  $P$ ,  $|x|$  can be written  $|x| = \text{proj}_{\mathbb{S}_{m-1}}(\sum_{i \in I} v_i)$ . Similarly,  $|y|$  can be written  $|y| = \text{proj}_{\mathbb{S}_{m-1}}(\sum_{i \in I^c} v_i)$ .

Now, denote by  $0$  the origin of the hyperplane  $P$ , and embed the vertices  $v_0, \dots, v_m$  in  $P$ . Observe that

$$0 = \sum_{i \leq 0} v_i = \sum_{i \in I} v_i + \sum_{i \in I^c} v_i.$$

Hence  $-\sum_{i \in I} v_i = \sum_{i \in I^c} v_i$ , and we deduce that

$$-|x| = \text{proj}_{\mathbb{S}_{m-1}}\left(-\sum_{i \in I} v_i\right) = \text{proj}_{\mathbb{S}_{m-1}}\left(\sum_{i \in I^c} v_i\right) = |y|.$$

Applying the same reasoning, one obtains the following result: for every simplex  $\sigma$  of  $\text{sub}^1(\partial\Delta^m)$ , if  $\nu$  denotes the image of  $\sigma$  by the relation on  $\text{sub}^1(\partial\Delta^m)$ , then the image of  $|\sigma|$  by the antipodal relation is also  $|\nu|$ . As a consequence, these two relations coincide.  $\square$

At a computational level, let us describe how to compute the face map  $\mathcal{F}: \mathcal{G}_1(\mathbb{R}^m) \rightarrow L$ . Since  $\mathcal{F}$  can be obtained as a quotient, it is enough to compute the face map of the sphere,  $\mathcal{F}': \mathbb{S}_{m-1} \rightarrow \text{sub}^1(\partial\Delta^m)$ , which corresponds to the homeomorphism  $p: \mathbb{S}_{m-1} \rightarrow |\text{sub}^1(\partial\Delta^m)|$ . It is given by the following lemma, which can be used in practice.

**Lemma 5.8.** *For every  $x \in \mathbb{S}_{m-1}$ , the image of  $x$  by  $\mathcal{F}'$  is equal to the intersection of all maximal faces  $\sigma = [w_0, \dots, w_m]$  of  $\text{sub}^1(\partial\Delta^m)$  that satisfies the following conditions: denoting by  $x_0$  any point of the affine hyperplane spanned by  $\{w_0, \dots, w_m\}$ , and by  $h$  a vector orthogonal to the corresponding linear hyperplane,*

- *the inner product  $\langle x, h \rangle$  has the same sign as  $\langle x_0, h \rangle$ ,*
- *the point  $\frac{\langle x_0, h \rangle}{\langle x, h \rangle} x$ , which is included in the affine hyperplane spanned by  $\{w_0, \dots, w_m\}$ , has nonnegative barycentric coordinates.*

*Proof.* Recall that for every  $x \in \mathbb{S}_{m-1}$ , the image  $p(x)$  is defined as the unique intersection point between the segment  $[0, x]$  and the set  $|\text{sub}^1(\partial\Delta^m)|$ . Besides, the face map  $\mathcal{F}'(x)$  is the unique

simplex  $\sigma \in \text{sub}^1(\partial\Delta^m)$  such that  $p(x) \in |\sigma|$ . Equivalently,  $\mathcal{F}'(x)$  is equal to the intersection of all maximal faces  $\sigma \in \text{sub}^1(\partial\Delta^m)$  such that  $p(x)$  belongs to the closure  $|\overline{\sigma}|$ .

Consider any maximal face  $\sigma = [w_0, \dots, w_m]$  of  $\text{sub}^1(\partial\Delta^m)$ . The first condition of the lemma ensures that the segment  $[0, x]$  intersects the affine hyperplane spanned by  $\{w_0, \dots, w_m\}$ . In this case, a computation shows that this intersection consists of the point  $\frac{\langle x_0, h \rangle}{\langle x, h \rangle} x$ . Then, the second condition of the lemma tests whether this point belongs to the convex hull of  $\{w_0, \dots, w_k\}$ . In conclusion, if  $\sigma$  satisfies these two conditions, then  $p(x) \in |\overline{\sigma}|$ .  $\square$

As a remark, let us point out that the verification of the conditions of this lemma is subject to numerical errors. In particular, the point  $\frac{\langle x_0, h \rangle}{\langle x, h \rangle} x$  may have nonnegative coordinates, yet mathematical softwares may return (small) negative values. Consequently, the algorithm may recognize less maximal faces that satisfy these conditions, hence return a simplex that strictly contains the wanted simplex  $\mathcal{F}'(x)$ . Nonetheless, such an error will not affect the output of the algorithm. Indeed, if we denote by  $\widetilde{\mathcal{F}'}$  the face map computed by the algorithm, we have that  $\mathcal{F}'(x) \subseteq \widetilde{\mathcal{F}'}(x)$  for all  $x \in \mathbb{S}_{m-1}$ . As a consequence of Lemma 5.1,  $\widetilde{\mathcal{F}'}$  satisfies the weak star condition if  $\mathcal{F}'$  does, and Equation (15) shows that every weak simplicial approximations for  $\mathcal{F}'$  are weak simplicial approximations for  $\widetilde{\mathcal{F}'}$ . Since every weak simplicial approximations are homotopic, we obtain that the induced maps in cohomology are equal, therefore the output of the algorithm is unchanged.

### 5.3 Vietoris-Rips version of the Čech bundle filtration

We still consider a subset  $X \subset \mathbb{R}^n \times M(\mathbb{R}^m)$ . Denote by  $\mathbb{X}$  the corresponding Čech set filtration, and by  $\mathbb{S} = (S^t)_{t \geq 0}$  the simplicial Čech filtration. For every  $t \geq 0$ , let  $R^t$  be the flag complex of  $S^t$ , i.e. the clique complex of the 1-skeleton  $(S^t)^1$  of  $S^t$ . It is known as the *Vietoris-Rips complex* of  $X$  at time  $t$ . The collection  $\mathbb{R} = (R^t)_{t \geq 0}$  is called the *Vietoris-Rips filtration* of  $X$ . The simplicial filtrations  $\mathbb{S}$  and  $\mathbb{R}$  are multiplicatively  $\sqrt{2}$ -interleaved (Bell et al., 2019, Theorem 3.1). In other words, for every  $t \geq 0$ , we have

$$S^t \subseteq R^t \subseteq S^{\sqrt{2}t}.$$

Let  $\gamma > 0$  and consider the Čech bundle filtration  $(\mathbb{X}, \mathbb{p})$  of  $X$ . Suppose that its maximal filtration value  $t_\gamma^{\max}(X)$  is positive. Let  $|\mathbb{R}| = (|R^t|)_{t \geq 0}$  denote the geometric realization of the Vietoris-Rips filtration. We can give  $|\mathbb{R}|$  a vector bundle filtration structure with  $(p')^t: |R^t| \rightarrow \mathcal{G}_d(\mathbb{R}^m)$  defined as

$$(p')^t = p^{\sqrt{2}t} \circ i^t,$$

where  $p^{\sqrt{2}t}$  denotes the maps of the Čech bundle filtration  $(\mathbb{X}, \mathbb{p})$ , and  $i^t$  denotes the inclusion  $|R^t| \hookrightarrow |S^{\sqrt{2}t}|$ . These maps fit in the following diagram:

$$\begin{array}{ccc} |R^t| & \xleftarrow{i^t} & |S^{\sqrt{2}t}| \\ & \searrow (p')^t & \downarrow p^{\sqrt{2}t} \\ & & \mathcal{G}_d(\mathbb{R}^m) \end{array} \quad = \quad \begin{array}{c} X^{\sqrt{2}t} \\ \downarrow p^{\sqrt{2}t} \\ \mathcal{G}_d(\mathbb{R}^m) \end{array}$$

This new vector bundle filtration is defined on the index set  $T' = \left[0, \frac{1}{\sqrt{2}} t_\gamma^{\max}(X)\right)$ .

It is clear from the construction that the vector bundle filtrations  $(\mathbb{X}, \mathbb{p})$  and  $(|\mathbb{R}|, \mathbb{p}')$  are multiplicatively  $\sqrt{2}$ -interleaved, with an interleaving that preserves the persistent Stiefel-Whitney classes. This property is a multiplicative equivalent of Theorem 3.3. Remember that if  $X$  is a subset of  $\mathbb{R}^n \times \mathcal{G}_d(\mathbb{R}^m)$ , then the maximal filtration value of the Čech bundle filtration on  $X$  is  $t_\gamma^{\max}(X) = \frac{\sqrt{2}}{2} \gamma$  (see Equation (11)). Consequently, the maximal filtration value of its Vietoris-Rips version is  $\frac{1}{2} \gamma$ .

From an application perspective, we choose to work with the Vietoris-Rips filtration since it is easier to compute. Indeed, its construction only relies on computing pairwise distances, and finding cliques in graphs.

## 5.4 Choice of the parameter $\gamma$

This subsection is devoted to discussing the influence of the parameter  $\gamma > 0$ . Recall that  $\gamma$  affects the norm  $\|\cdot\|_\gamma$  we chose on  $\mathbb{R}^n \times M(\mathbb{R}^m)$ :

$$\|(x, A)\|_\gamma^2 = \|x\|^2 + \gamma^2 \|A\|_F^2.$$

Let  $X \subset \mathbb{R}^n \times M(\mathbb{R}^m)$ . If  $\gamma_1 \leq \gamma_2$  are two positive real numbers, the corresponding Čech filtrations  $\mathbb{X}_1$  and  $\mathbb{X}_2$ , as well as the Čech bundle filtrations  $(\mathbb{X}_1, \mathbb{p}_1)$  and  $(\mathbb{X}_2, \mathbb{p}_2)$ , are  $\frac{\gamma_2}{\gamma_1}$ -interleaved multiplicatively. This comes from the straightforward inequality

$$\|\cdot\|_{\gamma_1} \leq \|\cdot\|_{\gamma_2} \leq \frac{\gamma_2}{\gamma_1} \|\cdot\|_{\gamma_1}.$$

Note that we also have the additive inequality

$$\|(x, A)\|_{\gamma_1} \leq \|(x, A)\|_{\gamma_2} \leq \|(x, A)\|_{\gamma_1} + \sqrt{\gamma_2^2 - \gamma_1^2} \|A\|_F.$$

One deduces that the Čech bundle filtrations  $(\mathbb{X}_1, \mathbb{p}_1)$  and  $(\mathbb{X}_2, \mathbb{p}_2)$  are  $\sqrt{\gamma_2^2 - \gamma_1^2} \cdot t^{\max}(X)$ -interleaved additively, where  $t^{\max}(X)$  is the maximal filtration value when  $\gamma = 1$ . As a consequence of these interleavings, when the values  $\gamma_1$  and  $\gamma_2$  are close, the persistence barcodes and the lifebars of the persistent Stiefel-Whitney classes are close (see Theorem 3.3).

As a general principle, one would choose the parameter  $\gamma$  to be large, since it would lead to large filtrations. More precisely, if  $t_{\gamma_1}^{\max}(X)$  and  $t_{\gamma_2}^{\max}(X)$  denote respectively the maximal filtration values of  $(\mathbb{X}_1, \mathbb{p}_1)$  and  $(\mathbb{X}_2, \mathbb{p}_2)$ , then  $t_{\gamma_1}^{\max}(X) = \gamma_1 \cdot t^{\max}(X)$  and  $t_{\gamma_2}^{\max}(X) = \gamma_2 \cdot t^{\max}(X)$ , as in Equation (10). Moreover, we have the following inclusion:

$$X_1^{t_{\gamma_1}^{\max}(X)} \subseteq X_2^{t_{\gamma_2}^{\max}(X)},$$

where  $X_1^{t_{\gamma_1}^{\max}(X)}$  denotes the thickening of  $X$  with respect to the norm  $\|\cdot\|_{\gamma_1}$ , and  $X_2^{t_{\gamma_2}^{\max}(X)}$  with respect to  $\|\cdot\|_{\gamma_2}$ . This inclusion can be proven from the following fact, valid for every  $x \in \mathbb{R}^n$  and  $A \in M(\mathbb{R}^m)$  such that  $\|A\|_F \leq t^{\max}(X)$ :

$$\|(x, A)\|_{\gamma_1} \leq t_{\gamma_1}^{\max}(X) \implies \|(x, A)\|_{\gamma_2} \leq t_{\gamma_2}^{\max}(X).$$

Hence larger parameters  $\gamma$  lead to larger maximal filtration values and larger filtrations.

However, as we show in the following examples, different values of  $\gamma$  may result in different behaviours of the persistent Stiefel-Whitney classes. In Example 5.10, large values of  $\gamma$  highlight properties of the dataset that are not consistent with the underlying vector bundle, which is orientable. Notice that, so far, we always picked the value  $\gamma = 1$ , for it seemed experimentally relevant with the datasets we chose.

**Example 5.9.** Consider the set  $Y \subset \mathbb{R}^2 \times M(\mathbb{R}^2)$  representing the Möbius strip, as in Example 3.2 of Subsect. 3.2:

$$Y = \left\{ \left( \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \begin{pmatrix} \cos(\frac{\theta}{2})^2 & \cos(\frac{\theta}{2}) \sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2}) \sin(\frac{\theta}{2}) & \sin(\frac{\theta}{2})^2 \end{pmatrix} \right), \theta \in [0, 2\pi) \right\}.$$

As we show in Appendix A.1,  $Y$  is a circle, included in a 2-dimensional affine subspace of  $\mathbb{R}^2 \times M(\mathbb{R}^2)$ . Its radius is  $\sqrt{1 + \frac{\gamma^2}{2}}$ . As a consequence, the persistence of the Čech filtration of  $Y$  consists of two bars: one  $H^0$ -feature, the bar  $[0, +\infty)$ , and one  $H^1$ -feature, the bar  $\left[0, \sqrt{1 + \frac{\gamma^2}{2}}\right)$ .

For any  $\gamma > 0$ , the maximal filtration value of the Čech bundle filtration of  $Y$  is  $t_\gamma^{\max}(Y) = \frac{\sqrt{2}}{2}\gamma$ . Moreover, the persistent Stiefel-Whitney class  $w_1^t(Y)$  is nonzero all along the filtration. In this example, we see that the parameter  $\gamma$  does not influence the qualitative interpretation of the persistent Stiefel-Whitney class. It is always nonzero where it is defined. The following example shows a case where  $\gamma$  does influence the persistent Stiefel-Whitney class.

**Example 5.10.** Consider the set  $X \subset \mathbb{R}^2 \times M(\mathbb{R}^2)$  representing the normal bundle of the circle  $\mathbb{S}_1$ , as in Example 3.2:

$$X = \left\{ \left( \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \begin{pmatrix} \cos(\theta)^2 & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin(\theta)^2 \end{pmatrix} \right), \theta \in [0, 2\pi) \right\}.$$

As we show in Appendix A.2,  $X$  is a subset of a 2-dimensional flat torus embedded in  $\mathbb{R}^2 \times M(\mathbb{R}^2)$ , hence can be seen as a torus knot.

Before studying the Čech bundle filtration of  $X$ , we discuss its Čech filtration  $\mathbb{X}$ . Its behaviour depends on  $\gamma$ :

- if  $\gamma \leq \frac{\sqrt{2}}{2}$ , then  $X^t$  retracts on a circle for  $t \in [0, 1)$ ,  $X^t$  retracts on a 3-sphere for  $t \in [1, \sqrt{1 + \frac{1}{2}\gamma^2})$ , and  $X^t$  retracts on a point for  $t \geq \sqrt{1 + \frac{1}{2}\gamma^2}$ .
- if  $\gamma \geq \frac{\sqrt{2}}{2}$ , then  $X^t$  retracts on a circle for  $t \in [0, 1)$ ,  $X^t$  retracts on another circle for  $t \in [1, \frac{\sqrt{2}}{2}\sqrt{1 + \gamma^2 + \frac{1}{4\gamma^2}})$ ,  $X^t$  retracts on a 3-sphere for  $t \in [\frac{\sqrt{2}}{2}\sqrt{1 + \gamma^2 + \frac{1}{4\gamma^2}}, \sqrt{1 + \frac{1}{2}\gamma^2})$ , and  $X^t$  has the homotopy type of a point for  $t \geq \sqrt{1 + \frac{1}{2}\gamma^2}$ .

Let us interpret these facts. If  $\gamma \leq \frac{\sqrt{2}}{2}$ , then the persistent cohomology of  $X$  looks similar to the persistent cohomology of the underlying set  $\left\{ \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \theta \in [0, 2\pi) \right\} \subset \mathbb{R}^2$ , but with a  $H^3$  cohomology feature added. Besides, if  $\gamma \geq \frac{\sqrt{2}}{2}$ , a new topological feature appears in the  $H^1$ -barcode: the bar  $[1, \sqrt{2}\sqrt{1 + \gamma^2 + \frac{1}{4\gamma^2}})$ . These barcodes are depicted in Figures 15 and 16.

Let us now discuss the corresponding Čech bundle filtrations. For any  $\gamma > 0$ , the maximal filtration value of the Čech bundle filtration of  $X$  is  $t_\gamma^{\max}(X) = \frac{\sqrt{2}}{2}\gamma$ . We observe two behaviours: if  $\gamma \leq \frac{\sqrt{2}}{2}$ , then  $w_1^t(X)$  is zero all along the filtration, and if  $\gamma > \frac{\sqrt{2}}{2}$ , then  $w_1^t(X)$  is nonzero from  $t^\dagger = 1$ . This is proven in Appendix A.2. To conclude, this persistent Stiefel-Whitney class is consistent with the underlying bundle—the normal bundle of the circle, which is trivial—only for  $t \leq 1$ .

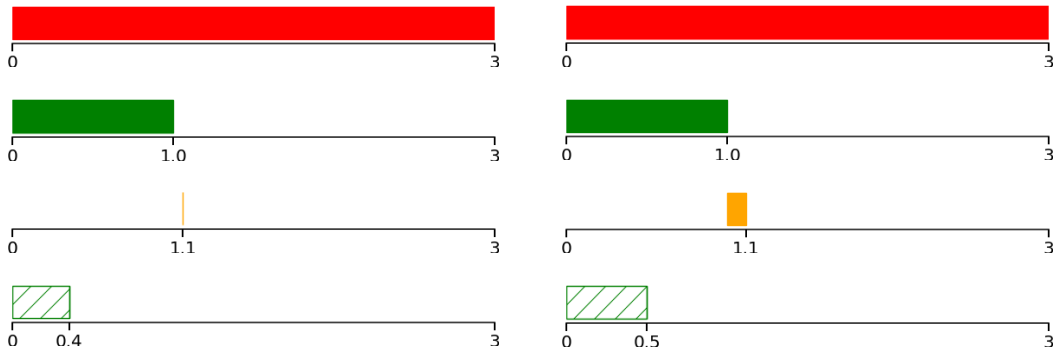


Figure 15:  $H^0$ -,  $H^1$ -,  $H^3$ -barcodes and lifebar of the first persistent Stiefel-Whitney class of  $X$  with  $\gamma = \frac{1}{2}$  (left) and  $\gamma = \frac{\sqrt{2}}{2}$  (right).

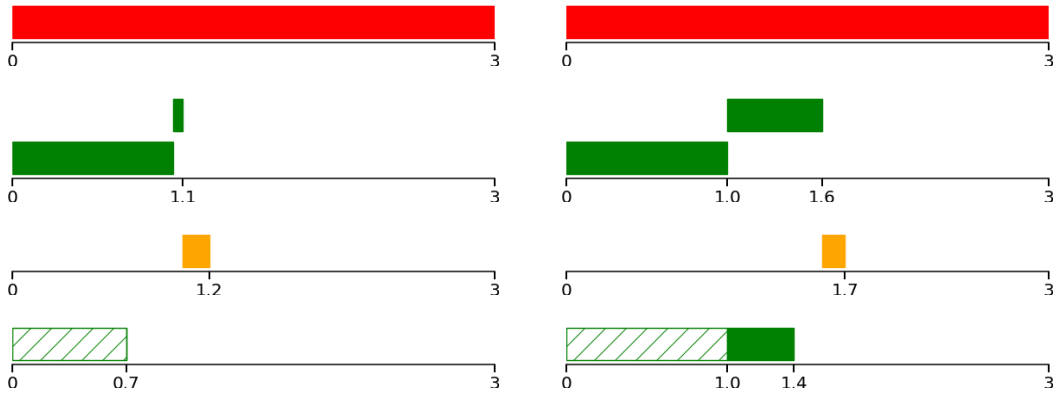


Figure 16:  $H^0$ -,  $H^1$ -,  $H^3$ -barcodes and lifebars of the first persistent Stiefel-Whitney class of  $X$  with  $\gamma = 1$  (left) and  $\gamma = 2$  (right).

## 6 Numerical experiments

In this last section, we propose to apply our algorithm on synthetic datasets, inspired by image analysis. We aim at illustrating how the first persistent Stiefel-Whitney class may reveal two properties: when the datasets contain certain symmetries, and when the datasets are close to non-orientables vector bundles. A Python notebook gathering these experiments can be found at <https://github.com/raphaeltinarrage/PersistentCharacteristicClasses/blob/master/Experiments.ipynb>.

### 6.1 Datasets with symmetries

**Giraffe moving forward.** Let us start with a simple dataset: it consists in a picture of a giraffe, that we translate to the right via cyclic permutations. The dataset contains of 150 images, each of size  $150 \times 300$  pixels, in RGB format. Since  $150 \times 300 \times 3 = 135\,000$ , the dataset can be embedded in  $\mathbb{R}^{135\,000}$ . Some of these images can be seen in Figure 17.



Figure 17: The dataset consists in a giraffe moving forward.

By performing a Principal Component Analysis (PCA), we project the dataset on the three principal axes. The result is a subset  $X = \{x_1, \dots, x_{150}\}$  of  $\mathbb{R}^3$ . As a last pre-processing step, we divide each point of  $X$  by the value  $\max\{\|x\|, x \in X\}$ , so that  $X$  becomes a subset of the unit ball  $\bar{B}(0, 1) \subset \mathbb{R}^3$ . The point cloud  $X$  is represented on Figure 18. Next to it, we give the persistence barcodes of its Rips filtration (we choose the coefficient field  $\mathbb{Z}_2$ , and represent  $H^0$  in red and  $H^1$  in green). Note that  $X$  actually lies close to the unit sphere  $\mathbb{S}_2$  of  $\mathbb{R}^3$ , and moreover that it describes a circle.

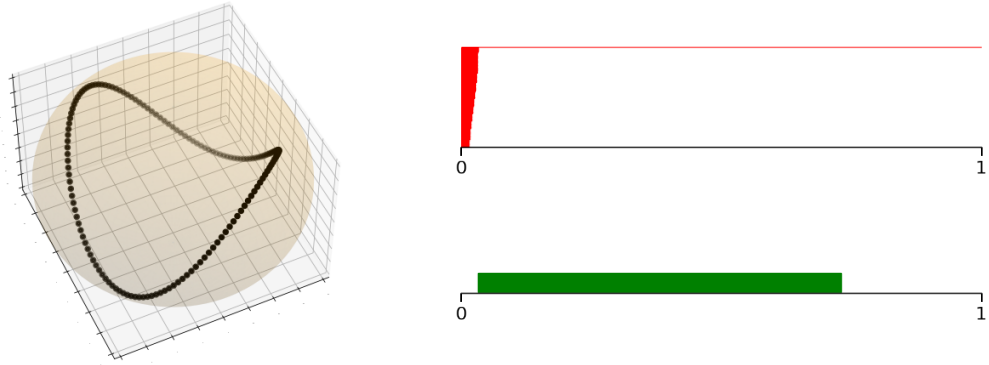


Figure 18: **Left:** The point cloud  $X \subset \mathbb{R}^3$ . **Right:** The barcode of its Rips filtration.

Let us consider on  $X$  two 1-dimensional bundles, defined via classifying maps  $X \rightarrow \mathcal{G}_1(\mathbb{R}^3) \subset M(\mathbb{R}^m)$ :

$$p: x_i \mapsto P(x_i) \quad \text{and} \quad p': x_i \mapsto P(x_{i+1} - x_{i-1})$$

where  $P(x)$  represents the orthogonal projection matrix on the 1-dimensional subspace of  $\mathbb{R}^3$  spanned by  $x$ . The vector bundle  $p$  is to be seen as the normal bundle of  $\mathbb{S}_2$  restricted to  $X$ , and  $p'$  is to be seen as the tangent bundle of the circle. These two theoretical bundles — restriction of the normal bundle of the sphere, and tangent bundle of a circle — are trivial. This follows from the general fact that 1-dimensional bundles on the sphere are trivial.

We represent on Figure 19 the vector bundles  $p$  and  $p'$ , seen in  $\mathbb{R}^3$ . One observes that, while going around the circle, the lines make a complete twist. This is the same behavior as the trivial bundle of the circle, that we studied earlier (see Figure 2).

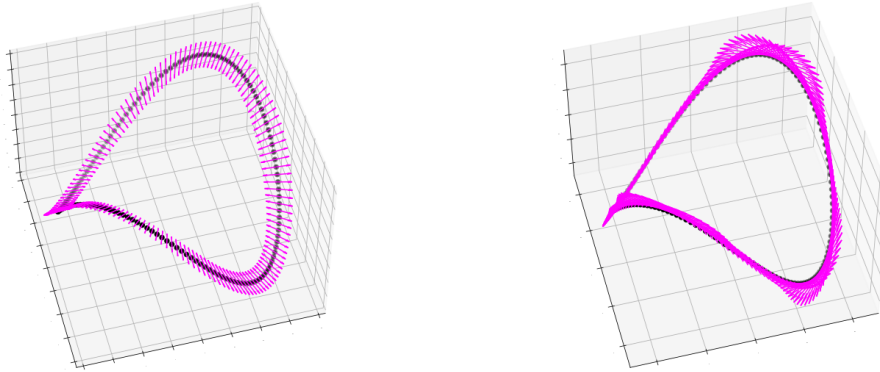


Figure 19: The vector bundles  $p$  and  $p'$ .

Next, let  $\gamma = 1$ , and consider the lifted sets  $\tilde{X} = \{(x, \gamma p(x)), x \in X\}$  and  $\tilde{X}' = \{(x, \gamma p'(x)), x \in X\}$ . We remind the reader that this construction has been studied in Subsection 3.4. We represent the point clouds  $\tilde{X}$  and  $\tilde{X}'$  on Figure 20, projected in  $\mathbb{R}^3$  via PCA, as well as the persistence barcodes of their Rips filtration. On these two barcodes, one observes one prominent  $H^0$ -feature, and one prominent  $H^1$ -feature. Below, we plot the lifebars of their first persistent Stiefel-Whitney classes, up to the maximal filtration value  $\frac{1}{2}\gamma$  (see Subsection 5.3). Both are empty, meaning that the persistent Stiefel-Whitney classes are zero all along the filtration. This is consistent with the fact that the underlying vector bundles are trivial.



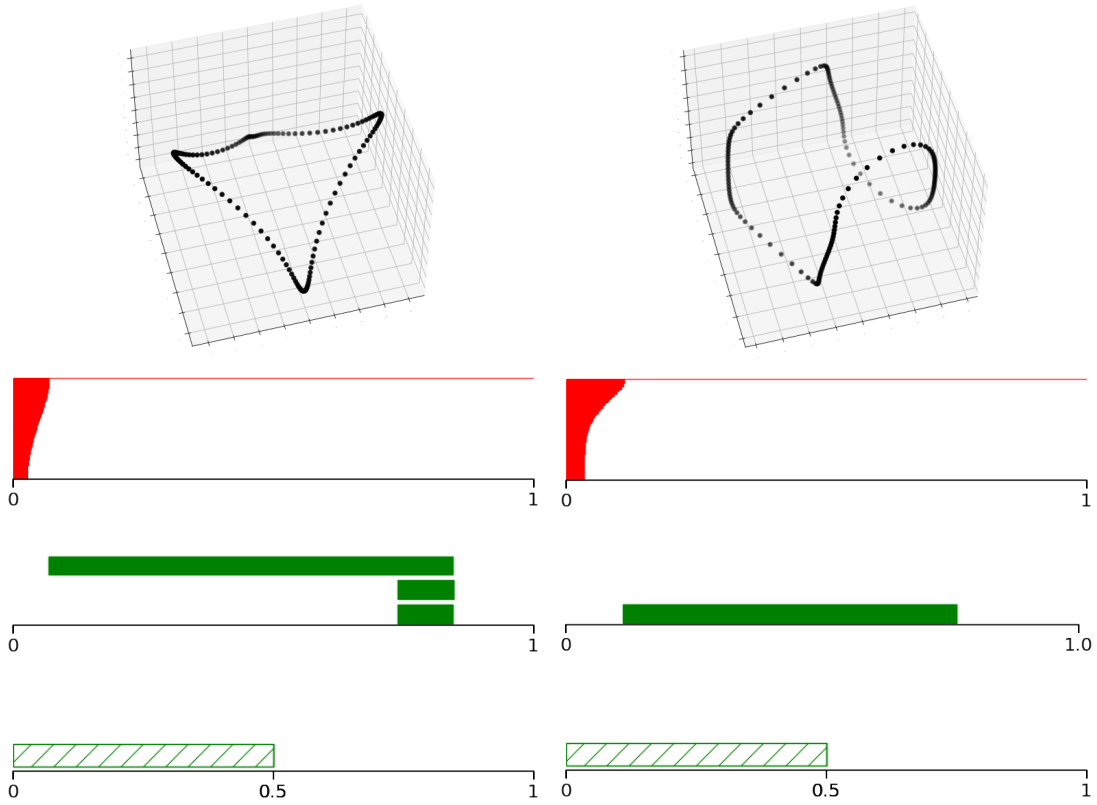


Figure 20: **Left:** The set  $\check{X}$  (projected in  $\mathbb{R}^3$  via PCA), the barcodes of its Rips filtration, and the lifebar of its first persistent Stiefel-Witney class. **Right:** Same for  $\check{X}'$ . Only green bars of length larger than 0.1 are represented.

**Giraffe moving behind two trees.** We will now study a variation of this dataset. As before, we consider a giraffe walking straight, but now with two trees on the foreground. The dataset consists of 150 images, each of size  $130 \times 300$  pixels, in RGB format. Since  $130 \times 300 \times 3 = 117\,000$ , the dataset can be seen as a subset of  $\mathbb{R}^{117\,000}$ . Some of the images can be seen in Figure 21.



Figure 21: The dataset consists in a giraffe moving forward, with two trees on the foreground.

Just as before, we project the point cloud in  $\mathbb{R}^3$  via PCA, and renormalize it. The corresponding point cloud, that we denote  $Y = \{y_1, \dots, y_{150}\}$ , and the barcode of its Rips filtration can be seen on Figure 22.

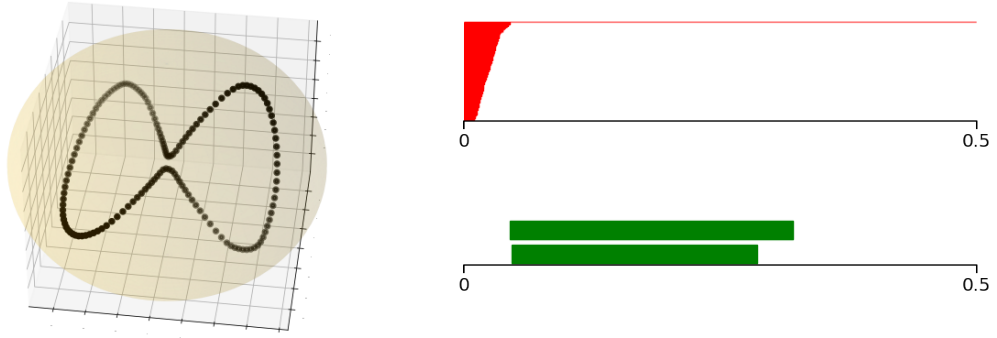


Figure 22: **Left:** The point cloud  $Y \subset \mathbb{R}^3$ . **Right:** The barcode of its Rips filtration. Only green bars of length larger than 0.05 are represented.

Note that, in this collection, the images where the giraffe goes behind the trees are close to each other. Indeed, only a few pixels differ between them. As a consequence, the point cloud  $Y$ , though lying close to a circle, seems to come closer to itself at some point. This behavior translates in its persistent cohomology as two prominent  $H^1$ -features.

We now consider the vector bundles  $q$  and  $q'$  defined as before:

$$q: y_i \mapsto P(y_i) \quad \text{and} \quad q': y_i \mapsto P(y_{i+1} - y_{i-1})$$

They are represented on Figure 23. Let  $\gamma = 0.5$ , and consider the lifted sets  $\check{Y} = \{(y, \gamma q(y)), y \in Y\}$  and  $\check{Y}' = \{(y, \gamma q'(y)), y \in Y\} \subset \mathbb{R}^3 \times \mathcal{G}_1(\mathbb{R}^3)$ . We represent them on Figure 24, as well as the persistence barcode of their Rips filtration, and the lifebars of their first Stiefel-Whitney classes, up to their maximal filtration value. We read that the second one is empty, while the first one is not.

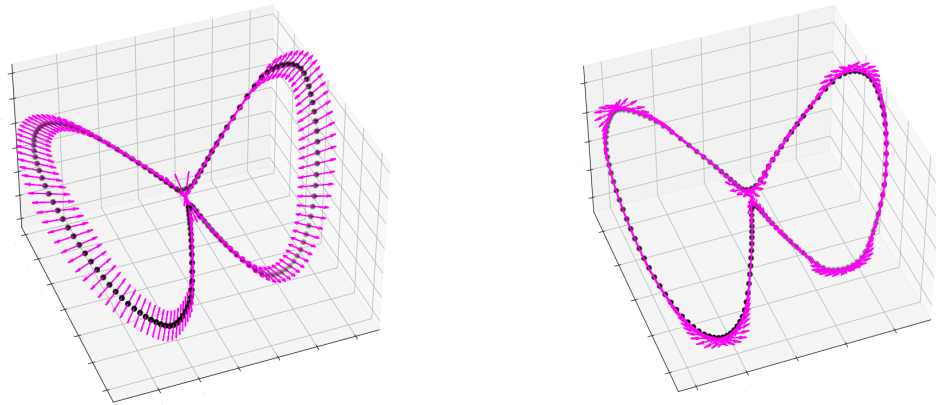


Figure 23: The vector bundles  $q$  and  $q'$ .

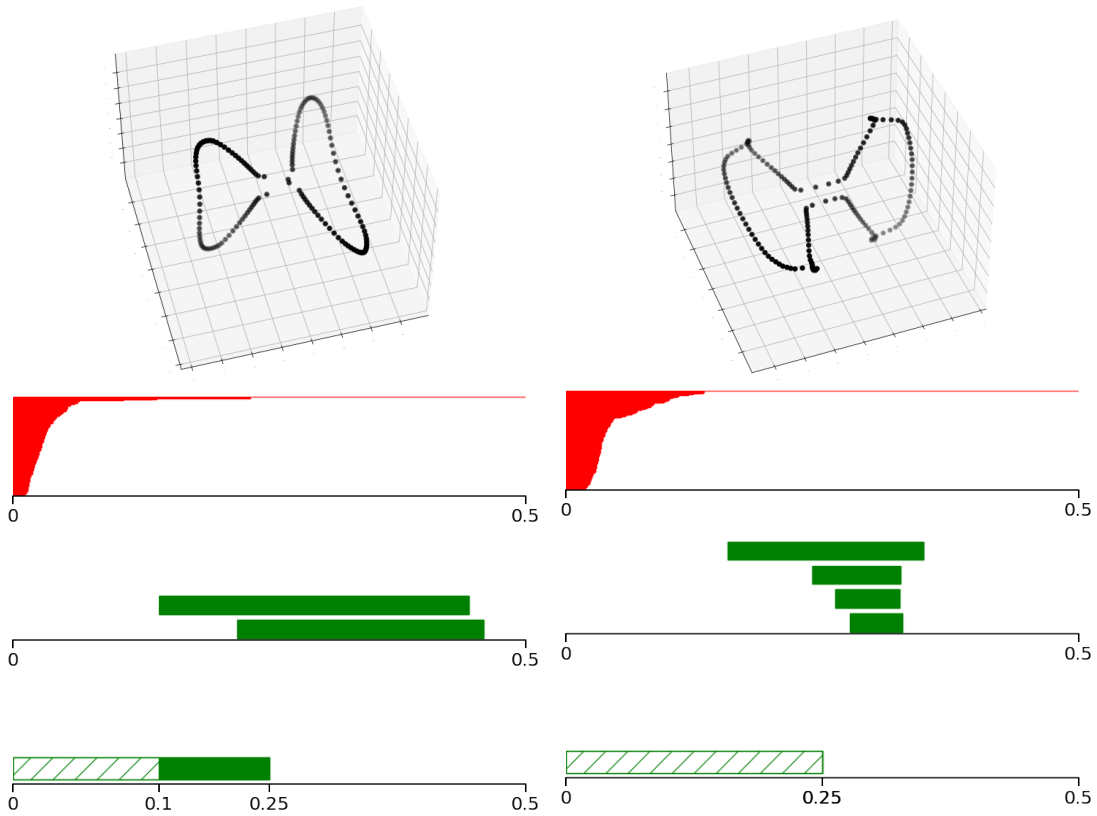


Figure 24: **Left:** The set  $\tilde{Y}$  (projected in  $\mathbb{R}^3$  via PCA), the barcodes of its Rips filtration, and the lifebar of its first persistent Stiefel-Whitney class. **Right:** Same for  $\tilde{Y}'$ . Only green bars of length larger than 0.05 are represented.

Let us explain this phenomenon. In both cases, as we can see on Figure 23, the lines make a full twist while turning around the circle. This corresponds to a trivial bundle. However, in the case of  $q$ , the points  $x$  close to the almost self-intersection of  $Y$  corresponds to lines  $q(x)$  close to each other. As a consequence, in the Rips filtration of the lifted set  $\tilde{Y}$ , these points will connect early. Hence the filtration behaves as if  $Y$  were composed of two loops. But, on these loops, the lines make only a half-twist. This correspond to the Mobius bundle on the circle (see Figure 2). Therefore we obtain a non-trivial Stiefel-Whitney class.

The bundle  $q'$  does not reflect this property. This is because the points  $x$  close to the self-intersection of  $Y$  does not correspond to lines  $q'(x)$  that are close to each other. In the Rips filtration of the lifted set  $\tilde{Y}'$ , the two loops connect late, hence the non-orientability does not appear on its persistent Stiefel-Whitney class.

**Rotating cylinders.** We now propose a dataset whose tangent bundle reflects non-orientability. Consider the union of two cylinders in  $\mathbb{R}^3$ , as represented on Figure 25. By applying rotations, we obtain a dataset of 100 pictures, in RGB format, of  $500 \times 500$  pixels.

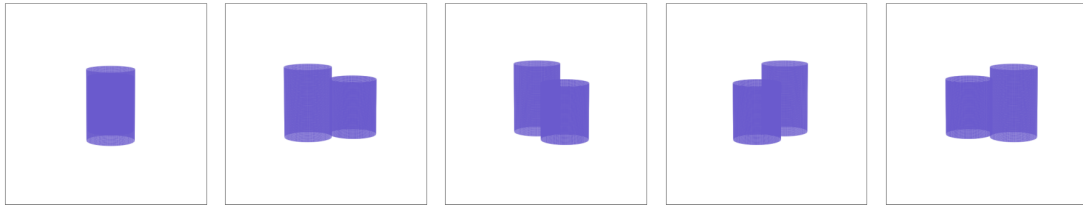


Figure 25: The dataset consists in different views of two cylinders in  $\mathbb{R}^3$ .

As before, we embed them into  $\mathbb{R}^{750000}$ , project them into  $\mathbb{R}^3$  via PCA, and renormalize them. The corresponding point cloud, denoted  $Z = \{z_1, \dots, z_{100}\}$ , and the barcodes of its Rips filtration are represented on Figure 26.

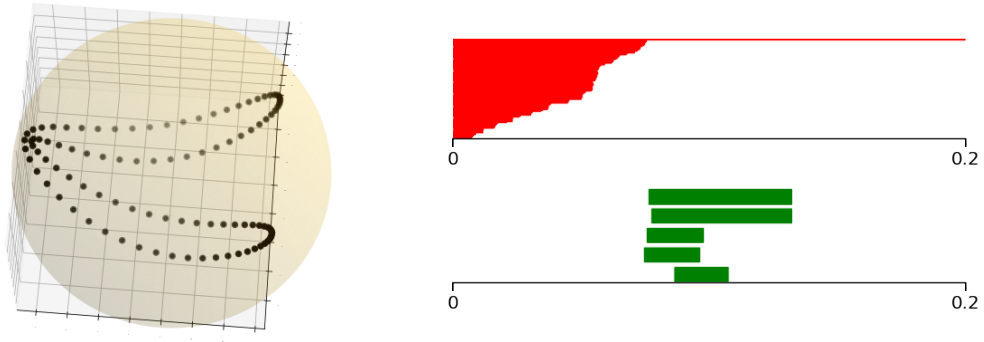


Figure 26: **Left:** The point cloud  $Z \subset \mathbb{R}^3$ . **Right:** The barcode of its Rips filtration. Only green bars of length larger than 0.02 are represented.

Define on  $Z$  the tangent bundle  $r: z_i \mapsto P(z_{i+1} - z_{i-1}) \in \mathcal{G}_1(\mathbb{R}^3)$ . Let  $\gamma = 0.2$  and consider the lifted set  $\tilde{Z} = \{(z, \gamma r(z)), z \in Z\} \subset \mathbb{R}^3 \times \mathcal{G}_1(\mathbb{R}^3)$ . The persistence barcodes of its Rips filtration, and the lifebar of its first persistent Stiefel-Whitney class are represented on Figure 27. We see that the lifebar is not trivial.

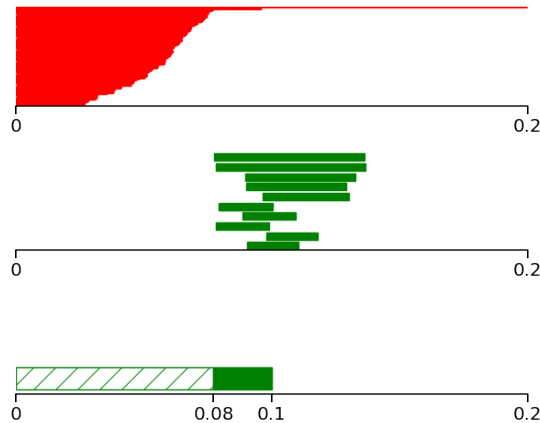


Figure 27: The barcodes of the Rips filtration of  $\tilde{Z}$ , and the lifebar of its first persistent Stiefel-Whitney class. Only green bars of length larger than 0.02 are represented.

Here, the same phenomenon as before occurs: there are two images of the collection which are almost equal (when the cylinders are one in front of the other), resulting in a point cloud

that almost intersect itself. Moreover, points  $z$  close to the area where  $Z$  almost intersect itself correspond to lines  $r(z)$  that are close to each other. Consequently, the persistent homology of  $\check{Z}$  shows two prominent loops, whose corresponding vector bundles are non-orientable.

In these two last experiments, we observed the following fact: trivial vector bundles, but whose underlying point clouds present self-similarity, or self-intersection, may result in a *shortcut* of the vector bundle, implying a non-trivial persistent Stiefel-Whitney class.

## 6.2 Datasets with intrinsic (non-)orientability

We will now present three datasets which reflect some underlying theoretical orientability or non-orientability.

**Gorilla on a torus.** Let us consider a picture of a gorilla, that we translate to the right and downwards via cyclic permutations (see Figure 28). Each image has size  $130 \times 120$  pixels, in RGB format. The dataset consists of 3900 images (65 vertical permutations and 60 horizontal permutations).

Note that the images behave as if the gorilla was on a torus. Indeed, the torus can be obtained from the square  $[0, 1]^2$  by gluing its opposite edges. Hence the various images can be seen as translations of the original image, whose gluing pattern follows the one of the torus.

Since  $130 \times 120 \times 3 = 46\,800$ , the dataset can be seen as a point cloud of  $\mathbb{R}^{46\,800}$ , that we project into  $\mathbb{R}^4$  via PCA. The resulting subset is denoted  $X = \{x_{i,j}, i \in [1, 65], j \in [1, 60]\}$ . The first index  $i$  corresponds to translations downwards, and the second index  $j$  corresponds to translations to the right.

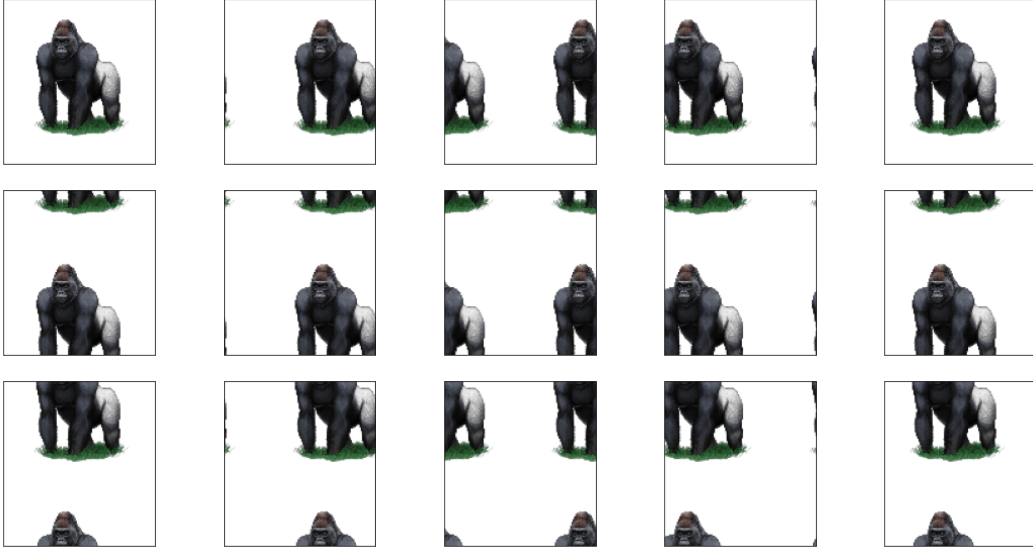


Figure 28: The dataset consists in a gorilla moving forward and downwards on a torus.

Consider on  $X$  the *vertical tangent bundle* and the *horizontal tangent bundle*, defined respectively as

$$p: x_{i,j} \mapsto P(x_{i-1,j} - x_{i+1,j}) \quad \text{and} \quad p': x_{i,j} \mapsto P(x_{i,j-1} - x_{i,j+1})$$

where we recall that  $P(x)$  is the orthogonal projection matrix on the line spanned by  $x$ . The vector bundle  $p$  is to be seen as the vertical component of the tangent bundle of a torus, and  $p'$  as the horizontal one. Both are trivial bundles.

Now, let  $\gamma = 0.3$ , and consider the corresponding lifted sets  $\check{X} = \{(x, \gamma p(x)), x \in X\}$  and  $\check{X}' = \{(x, \gamma p'(x)), x \in X\} \subset \mathbb{R}^3 \times \mathcal{G}_1(\mathbb{R}^3)$ . We represent on Figure 29 the barcodes of the Rips

filtrations of the sets  $\tilde{X}$  and  $\tilde{X}'$ , and the lifebars of their first persistent Stiefel-Whitney classes. We read that they are zero all along the filtration. This is consistent with the underlying line bundles on the torus being trivial.

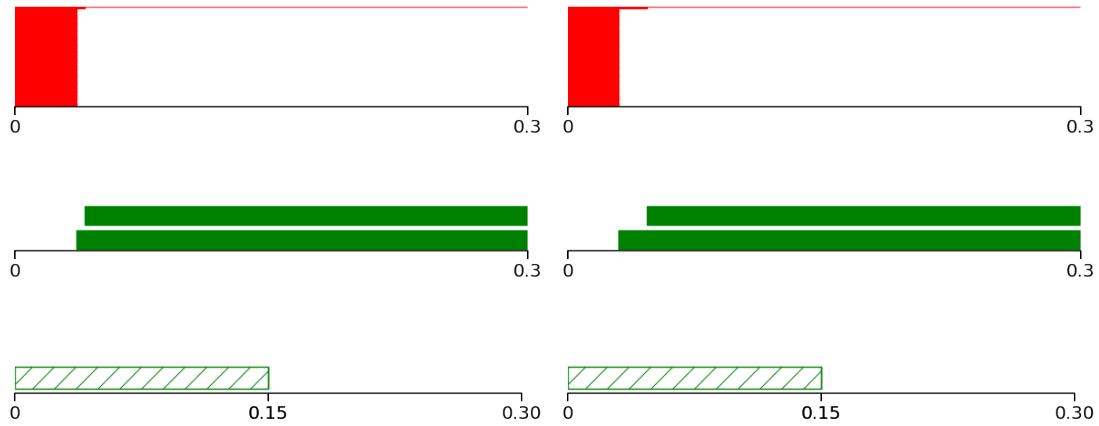


Figure 29: **Left:** The barcodes of the Rips filtration of  $\tilde{X}$ , and the lifebar of its first persistent Stiefel-Whitney class. **Right:** Same for  $\tilde{X}'$ . Only green bars of length larger than 0.03 are represented.

**Gorilla on a Klein bottle.** We will modify this dataset: we still translate the gorilla, but while translating it to the right, we inverse the part that arrives at the left (see Figure 30). It behaves as if the picture was glued to itself according to the gluing of a Klein bottle. Just as before, the dataset consists of 3900 images of size  $130 \times 120$  pixels. We embed the images into  $\mathbb{R}^{46800}$  and project them into  $\mathbb{R}^4$  via PCA. The resulting subset is denoted  $Y = \{y_{i,j}, i \in \llbracket 1, 65 \rrbracket, j \in \llbracket 1, 60 \rrbracket\}$ .

Again, consider the vector bundles  $q: y_{i,j} \mapsto P(y_{i-1,j} - y_{i+1,j})$  and  $q': y_{i,j} \mapsto P(y_{i,j-1} - y_{i,j+1})$ . They correspond to the horizontal and vertical components of the tangent bundle of a Klein bottle. Only one of them is non-trivial. Let  $\gamma = 0.3$ , and consider the corresponding lifted sets  $\tilde{Y}$  and  $\tilde{Y}'$ . We represent on Figure 31 the barcodes of their Rips filtrations, and the lifebars of their first persistent Stiefel-Whitney classes. The first lifebar is non-empty, reflecting the non-triviality of the underlying bundle.



Figure 30: The dataset consists in a gorilla moving forward and downwards on a Klein bottle.

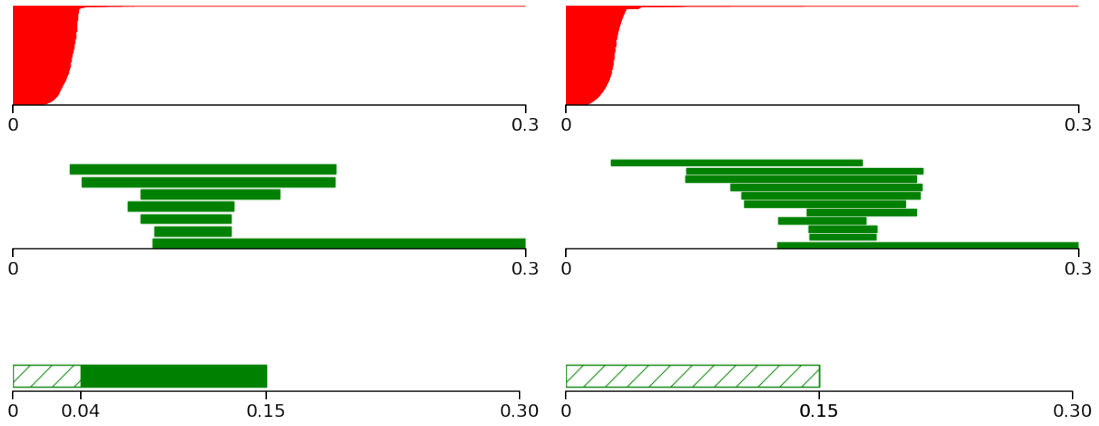


Figure 31: **Left:** The barcodes of the Rips filtration of  $\check{Y}$ , and the lifebar of its first persistent Stiefel-Whitney class. **Right:** Same for  $\check{Y}'$ . Only green bars of length larger than 0.03 are represented.

**Gorilla on the projective plane.** We close this subsection with a last variation of the dataset: the gorilla is translated to the right, with an inversion of the left part, and translated downwards, with also an inversion of the upper part (see Figure 32). It behaves as if the image was glued to itself according to the gluing of the projective plane  $\mathcal{G}_1(\mathbb{R}^3)$ . The dataset still consists of 3900 images of size  $130 \times 120$  pixels, that we embed into  $\mathbb{R}^{46800}$  and project into  $\mathbb{R}^4$  via PCA. The resulting subset is denoted  $Z = \{z_{i,j}, i \in \llbracket 1, 65 \rrbracket, j \in \llbracket 1, 60 \rrbracket\}$ .

We still consider the vector bundles  $r: z_{i,j} \mapsto P(z_{i-1,j} - z_{i+1,j})$  and  $r': z_{i,j} \mapsto P(z_{i,j-1} - z_{i,j+1})$ . Let  $\gamma = 0.3$ , and consider the corresponding lifted sets  $\check{Z}$  and  $\check{Z}'$ . We represent on Figure 33 the barcodes of their Rips filtrations, and the lifebars of their first persistent Stiefel-Whitney classes. Both lifebars are non-empty. We deduce that the underlying line bundles are non-trivial.

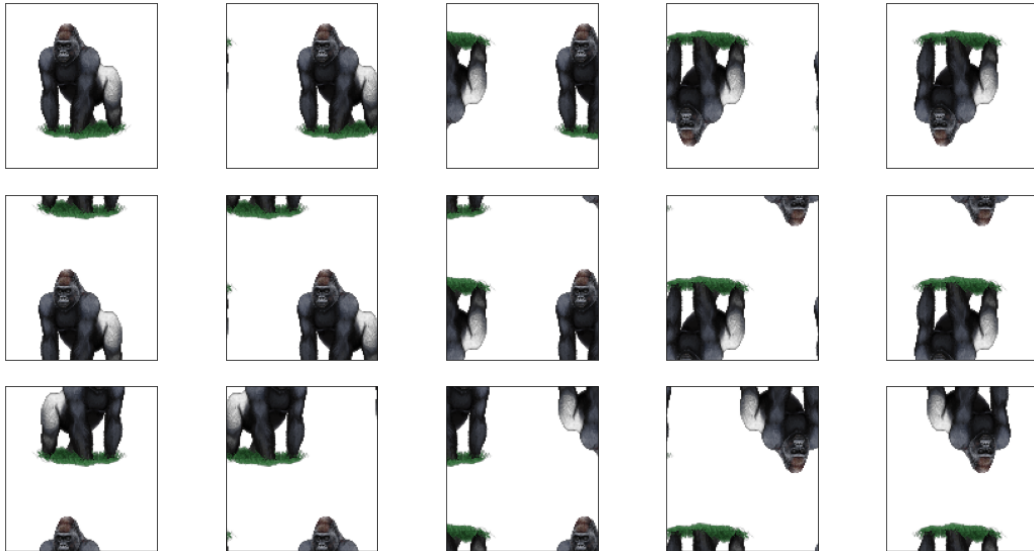


Figure 32: The dataset consists in a gorilla moving forward and downwards on the projective plane.

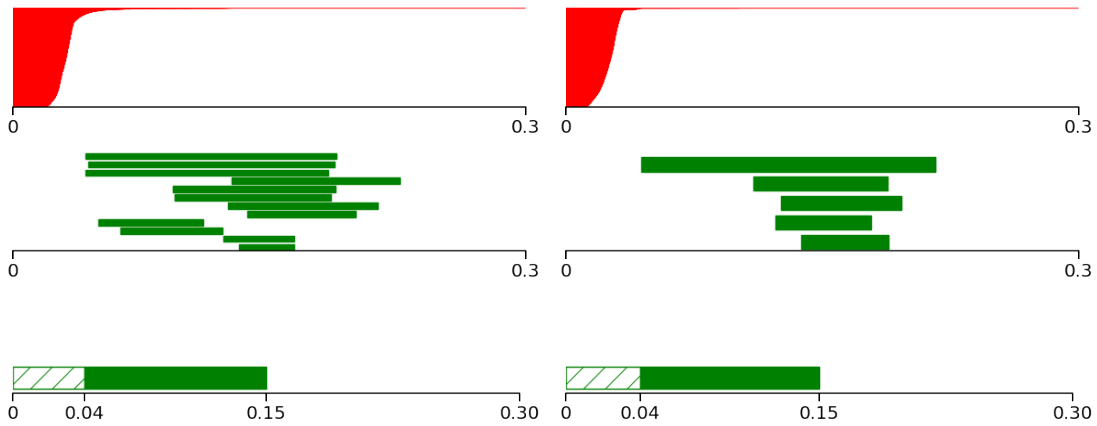


Figure 33: **Left:** The barcodes of the Rips filtration of  $\tilde{Z}$ , and the lifebar of its first persistent Stiefel-Whitney class. **Right:** Same for  $\tilde{Z}'$ . Only green bars of length larger than 0.03 are represented.

In these various experiments, we applied the same method: in ordered datasets, indexed by only one value  $i \mapsto x_i$ , two values  $(i, j) \mapsto x_{i,j}$  or more, one can consider directional line bundles by approximating partial derivatives, such as  $\frac{1}{2}(x_{i+1,j} - x_{i-1,j})$ . The first persistent Stiefel-Whitney classes of such bundles deliver information about the dataset in this particular direction.

## 7 Conclusion

In this paper we defined the persistent Stiefel-Whitney classes of vector bundle filtrations. We proved that they are stable with respect to the interleaving distance between vector bundle filtrations. We studied the particular case of Čech bundle filtrations of subsets of  $\mathbb{R}^n \times M(\mathbb{R}^m)$ , and showed that they yield consistent estimators of the usual Stiefel-Whitney classes of some underlying vector bundle.

Moreover, when the dimension of the bundle is 1 and  $X$  is finite, we proposed an algorithm to compute the first persistent Stiefel-Whitney class. We also described a way to compute their lifebars via mapping cones.

Our algorithm is limited to the bundles of dimension 1 since we only implemented triangulations of the Grassmannian  $\mathcal{G}_d(\mathbb{R}^m)$  when  $d = 1$ . However, any other triangulation of  $\mathcal{G}_d(\mathbb{R}^m)$ , with a computable face map, could be included in the algorithm without any modification. As far as we know, no triangulation of a Grassmannian  $\mathcal{G}_d(\mathbb{R}^m)$  with  $d > 1$  has never been given explicitly (i.e., as a list of simplices stored in a computer). This is a problem of interest, which will be addressed in a further work. A strategy could consist in using the usual CW-structures of the Grassmannians, and converting them into simplicial complexes, as done theoretically in Hatcher (2002, Theorem 2C.5). A recent result of Govc et al. (2020) gives an idea about the complexity of this problem: the number of simplices of minimal triangulations of  $\mathcal{G}_d(\mathbb{R}^m)$  must grow exponentially in both  $d$  and  $m$ .

**Acknowledgements.** I wish to thank Frédéric Chazal and Marc Glisse for fruitful discussions and corrections, as well as the anonymous reviewers for corrections and clarifications. I also thank Luis Scoccola for pointing out a strengthening of Lemma 4.1.



## A Supplementary material for Sect. 5

### A.1 Study of Example 5.9

We consider the set

$$X = \left\{ \left( \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \begin{pmatrix} \cos(\frac{\theta}{2})^2 & \cos(\frac{\theta}{2})\sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2})\sin(\frac{\theta}{2}) & \sin(\frac{\theta}{2})^2 \end{pmatrix} \right), \theta \in [0, 2\pi) \right\}.$$

In order to study the Čech filtration of  $X$ , we shall apply the following affine transformation: let  $Y$  be the subset of  $\mathbb{R}^2 \times M(\mathbb{R}^2)$  defined as

$$Y = \left\{ \left( \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \gamma \begin{pmatrix} \cos(\frac{\theta}{2})^2 & \cos(\frac{\theta}{2})\sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2})\sin(\frac{\theta}{2}) & \sin(\frac{\theta}{2})^2 \end{pmatrix} \right), \theta \in [0, 2\pi) \right\}.$$

and let  $\mathbb{Y} = (Y^t)_{t \geq 0}$  be the Čech filtration of  $Y$  in  $\mathbb{R}^2 \times M(\mathbb{R}^2)$  endowed with the norm  $\|(x, A)\|_1 = \sqrt{\|x\|^2 + \|A\|_F^2}$ . We recall that the Čech filtration of  $X$ , denoted  $\mathbb{X} = (X^t)_{t \geq 0}$ , is defined with respect to the norm  $\|\cdot\|_\gamma$ . It is clear that, for every  $t \geq 0$ , the thickenings  $X^t$  and  $Y^t$  are homeomorphic via the application

$$\begin{aligned} h: \mathbb{R}^2 \times M(\mathbb{R}^2) &\longrightarrow \mathbb{R}^2 \times M(\mathbb{R}^2) \\ (x, A) &\longmapsto (x, \gamma A). \end{aligned}$$

As a consequence, the persistence cohomology modules associated to  $\mathbb{X}$  and  $\mathbb{Y}$  are isomorphic.

Next, notice that  $Y$  is a subset of the affine subspace of dimension 2 of  $\mathbb{R}^2 \times M(\mathbb{R}^2)$  with origin  $O$  and spanned by the vectors  $e_1$  and  $e_2$ , where

$$O = \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \frac{\gamma}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right), \quad e_1 = \left( \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \frac{\gamma}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right), \quad e_2 = \left( \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \frac{\gamma}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right).$$

Indeed, using the equality

$$\begin{pmatrix} \cos(\frac{\theta}{2})^2 & \cos(\frac{\theta}{2})\sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2})\sin(\frac{\theta}{2}) & \sin(\frac{\theta}{2})^2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix},$$

we obtain

$$Y = O + \{\cos(\theta)e_1 + \sin(\theta)e_2, \theta \in [0, 2\pi)\}.$$

We see that  $Y$  is a circle, of radius  $\|e_1\| = \|e_2\| = \sqrt{1 + \frac{\gamma^2}{2}}$ .

Let  $E$  denotes the affine space with origin  $O$  and spanned by the vectors  $e_1$  and  $e_2$ . Lemma A.1, stated below, shows that the persistent cohomology of  $Y$ , seen in the ambient space  $\mathbb{R}^2 \times M(\mathbb{R}^2)$ , is the same as the persistent cohomology of  $Y$  restricted to the subspace  $E$ . As a consequence,  $Y$  has the same persistence as a circle of radius  $\sqrt{1 + \frac{\gamma^2}{2}}$  in the plane. Hence its barcode can be described as follows:

- one  $H^0$ -feature: the bar  $[0, +\infty)$ ,
- one  $H^1$ -feature: the bar  $\left[0, \sqrt{1 + \frac{\gamma^2}{2}}\right)$ .

**Lemma A.1.** *Let  $Y \subset \mathbb{R}^n$  be any subset, and define  $\check{Y} = Y \times \{(0, \dots, 0)\} \subset \mathbb{R}^n \times \mathbb{R}^m$ . Let these spaces be endowed with the usual Euclidean norms. Then the Čech filtrations of  $Y$  and  $\check{Y}$  yields isomorphic persistence modules.*

*Proof.* Let  $\text{proj}_n: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$  be the projection on the first  $n$  coordinates. One verifies that, for every  $t \geq 0$ , the map  $\text{proj}_n: \check{Y}^t \rightarrow Y^t$  is a homotopy equivalence. At cohomology level, these maps induce an isomorphism of persistence modules.  $\square$

Let us now study the Čech bundle filtration of  $Y$ , denoted  $(\mathbb{Y}, \mathbb{p})$ . According to Equation (11), its filtration maximal value is  $t^{\max}(Y) = t_{\gamma}^{\max}(X) = \frac{\gamma}{\sqrt{2}}$ . Note that  $\frac{\gamma}{\sqrt{2}}$  is lower than  $\sqrt{1 + \frac{\gamma^2}{2}}$ , which is the radius of the circle  $Y$ . Hence, for  $t < t^{\max}(Y)$ , the inclusion  $Y \hookrightarrow Y^t$  is a homotopy equivalence. Consider the following commutative diagram:

$$\begin{array}{ccc} Y & \xleftrightarrow{\quad} & Y^t \\ & \searrow p^0 & \swarrow p^t \\ & \mathcal{G}_1(\mathbb{R}^2) & \end{array}$$

It induces the following diagram in cohomology:

$$\begin{array}{ccc} H^*(Y) & \xleftarrow{\quad \sim \quad} & H^*(Y^t) \\ & \swarrow (p^0)^* & \searrow (p^t)^* \\ & H^*(\mathcal{G}_1(\mathbb{R}^2)) & \end{array}$$

The horizontal arrow is an isomorphism. Hence the map  $(p^t)^*: H^*(Y^t) \leftarrow H^*(\mathcal{G}_1(\mathbb{R}^2))$  is equal to  $(p^0)^*$ . We only have to understand  $(p^0)^*$ .

Remark that the map  $p^0: Y \rightarrow \mathcal{G}_1(\mathbb{R}^2)$  can be seen as the tautological bundle of the circle. It is then a standard result that  $(p^0)^*: H^*(Y) \leftarrow H^*(\mathcal{G}_1(\mathbb{R}^2))$  is nontrivial. Alternatively,  $p^0$  can be seen as a map between two circles. It is injective, hence its degree (modulo 2) is 1. We still deduce that  $(p^0)^*$  is nontrivial. As a consequence, the persistent Stiefel-Whitney class  $w_1^t(X)$  is nonzero for every  $t < t^{\max}(Y)$ .

## A.2 Study of Example 5.10

We consider the set

$$X = \left\{ \left( \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \begin{pmatrix} \cos(\theta)^2 & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin(\theta)^2 \end{pmatrix} \right), \theta \in [0, 2\pi) \right\}.$$

As we explained in the previous subsection, the Čech filtration of  $X$  with respect to the norm  $\|\cdot\|_{\gamma}$  yields the same persistence as the Čech filtration of  $Y$  with respect to the norm  $\|\cdot\|_1$ , where

$$Y = \left\{ \left( \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \gamma \begin{pmatrix} \cos(\theta)^2 & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin(\theta)^2 \end{pmatrix} \right), \theta \in [0, 2\pi) \right\}.$$

Notice that  $Y$  is a subset of the affine subspace of dimension 4 of  $\mathbb{R}^2 \times M(\mathbb{R}^2)$  with origin  $O = \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right)$  and spanned by the vectors  $e_1, e_2, e_3$  and  $e_4$ , where

$$\begin{aligned} e_1 &= \left( \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \right), & e_2 &= \left( \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \right), \\ e_3 &= \frac{1}{\sqrt{2}} \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right), & e_4 &= \frac{1}{\sqrt{2}} \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right). \end{aligned}$$

Indeed,  $Y$  can be written as

$$Y = O + \left\{ \cos(\theta)e_1 + \sin(\theta)e_2 + \frac{\gamma}{\sqrt{2}} \cos(2\theta)e_3 + \frac{\gamma}{\sqrt{2}} \sin(2\theta)e_4, \theta \in [0, 2\pi) \right\}.$$

This comes from the equality

$$\begin{pmatrix} \cos(\theta)^2 & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin(\theta)^2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix}.$$

Observe that  $Y$  is a torus knot, i.e. a simple closed curve included in the torus  $\mathbb{T}$ , defined as

$$\mathbb{T} = O + \left\{ \cos(\theta)e_1 + \sin(\theta)e_2 + \frac{\gamma}{\sqrt{2}} \cos(\nu)e_3 + \frac{\gamma}{\sqrt{2}} \sin(\nu)e_4, \quad \theta, \nu \in [0, 2\pi) \right\}.$$

The curve  $Y$  winds one time around the first circle of the torus, and two times around the second one, as represented in Figure 34. It is known as the torus knot  $(1, 2)$ .

Let  $E$  denotes the affine subspace with origin  $O$  and spanned by  $e_1, e_2, e_3, e_4$ . Since  $Y$  is a subset of  $E$ , it is equivalent to study the Čech filtration of  $Y$  restricted to  $E$  (as in Lemma A.1). We shall denote the coordinates of points  $x \in E$  with respect to the orthonormal basis  $(e_1, e_2, e_3, e_4)$ . That is, a tuple  $(x_1, x_2, x_3, x_4)$  shall refer to the point  $O + x_1e_1 + x_2e_2 + x_3e_3 + x_4e_4$  of  $E$ . Seen in  $E$ , the set  $Y$  can be written as

$$Y = \left\{ \left( \cos(\theta), \sin(\theta), \frac{\gamma}{\sqrt{2}} \cos(2\theta), \frac{\gamma}{\sqrt{2}} \sin(2\theta) \right), \theta \in [0, 2\pi) \right\}.$$

Moreover, for every  $\theta \in [0, 2\pi)$ , we shall denote  $y_\theta = \left( \cos(\theta), \sin(\theta), \frac{\gamma}{\sqrt{2}} \cos(2\theta), \frac{\gamma}{\sqrt{2}} \sin(2\theta) \right)$ .



Figure 34: Representations of the set  $Y$ , lying on a torus, for a small value of  $\gamma$  (left) and a large value of  $\gamma$  (right).

We now state two lemmas that will be useful in what follows.

**Lemma A.2.** *For every  $\theta \in [0, 2\pi)$ , the map  $\theta' \mapsto \|y_\theta - y_{\theta'}\|$  admits the following critical points:*

- $\theta' - \theta = 0$  and  $\theta' - \theta = \pi$  if  $\gamma \leq \frac{1}{\sqrt{2}}$ ,
- $\theta' - \theta = 0, \pi, \arccos(-\frac{1}{2\gamma^2})$  and  $-\arccos(-\frac{1}{2\gamma^2})$  if  $\gamma \geq \frac{1}{\sqrt{2}}$ .

They correspond to the values

- $\|y_\theta - y_{\theta'}\| = 0$  if  $\theta' - \theta = 0$ ,
- $\|y_\theta - y_{\theta'}\| = 2$  if  $\theta' - \theta = \pi$ ,
- $\|y_\theta - y_{\theta'}\| = \sqrt{2}\sqrt{1 + \gamma^2 + \frac{1}{4\gamma^2}}$  if  $\theta' - \theta = \pm \arccos(-\frac{1}{2\gamma^2})$ .

Moreover, we have  $\sqrt{2}\sqrt{1 + \gamma^2 + \frac{1}{4\gamma^2}} \geq 2$  when  $\gamma \geq \frac{1}{\sqrt{2}}$ .

*Proof.* Let  $\theta, \theta' \in [0, 2\pi)$ . One computes that

$$\|y_\theta - y_{\theta'}\|^2 = 4 \sin^2 \left( \frac{\theta - \theta'}{2} \right) + 2\gamma^2 \sin^2(\theta - \theta').$$

Consider the map  $f: x \in [0, 2\pi) \mapsto 4 \sin^2 \left( \frac{x}{2} \right) + 2\gamma^2 \sin^2(x)$ . Its derivative is

$$\begin{aligned} f'(x) &= 4 \cos \left( \frac{x}{2} \right) \sin \left( \frac{x}{2} \right) + 4\gamma^2 \cos(x) \sin(x) \\ &= 2 \sin(x) (1 + 2\gamma^2 \cos(x)). \end{aligned}$$

It vanishes when  $x = 0$ ,  $x = \pi$ , or  $x = \pm \arccos(-\frac{1}{2\gamma^2})$  if  $\gamma \geq \frac{1}{\sqrt{2}}$ . To conclude, a computation shows that  $f(0) = 0$ ,  $f(\pi) = 4$  and  $f \left( \pm \arccos \left( -\frac{1}{2\gamma^2} \right) \right) = 2 \left( 1 + \gamma^2 + \frac{1}{4\gamma^2} \right)$ .  $\square$

**Lemma A.3.** For every  $x \in E$  such that  $x \neq 0$ , the map  $\theta \mapsto \|x - y_\theta\|$  admits at most two local maxima and two local minima.

*Proof.* Consider the map  $g: \theta \in [0, 2\pi) \mapsto \|x - y_\theta\|^2$ . It can be written as

$$\begin{aligned} g(\theta) &= \|x\|^2 + \|y_\theta\|^2 - 2 \langle x, y_\theta \rangle \\ &= \|x\|^2 + 1 + \frac{\gamma^2}{2} - 2 \langle x, y_\theta \rangle. \end{aligned}$$

Let us show that its derivative  $g'$  vanishes at most four times on  $[0, 2\pi)$ , which will prove the result. Using the expression of  $y_\theta$ , we see that  $g'$  can be written as

$$g'(\theta) = a \cos(\theta) + b \sin(\theta) + c \cos(2\theta) + d \sin(2\theta),$$

where  $a, b, c, d \in \mathbb{R}$  are not all zero. Denoting  $\omega = \cos(\theta)$  and  $\xi = \sin(\theta)$ , we have  $\xi^2 = 1 - \omega^2$ ,  $\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta) = 2\omega^2 - 1$  and  $\sin(2\theta) = 2 \cos(\theta) \sin(\theta) = 2\omega\xi$ . Hence

$$g'(\theta) = a\omega + b\xi + 2c\omega^2 + 2d\omega\xi.$$

Now, if  $g'(\theta) = 0$ , we get

$$a\omega + 2c\omega^2 = -(b + 2d\omega)\xi \quad (17)$$

Squaring this equality yields  $(a\omega + 2c\omega^2)^2 = (b + 2d\omega)^2 (1 - \omega^2)$ . This degree four equation, with variable  $\omega$ , admits at most four roots. To each of these  $w$ , there exists a unique  $\xi = \pm\sqrt{1 - w^2}$  that satisfies Equation (17). In other words, the corresponding  $\theta \in [0, 2\pi)$  such that  $\omega = \cos(\theta)$  is unique. We deduce that  $g'$  vanishes at most four times on  $[0, 2\pi)$ .  $\square$

Before studying the Čech filtration of  $Y$ , let us describe some geometric quantities associated to it. Using a symbolic computation software, we see that the curvature of  $Y$  is constant and equal to

$$\rho = \frac{\sqrt{1 + 8\gamma^2}}{1 + 2\gamma^2}.$$

In particular, we have  $\rho \geq 1$  if  $\gamma \leq 1$ , and  $\rho < 1$  if  $\gamma > 1$ . We also have an expression for the diameter of  $Y$ :

$$\frac{1}{2} \text{diam}(Y) = \begin{cases} 1 & \text{if } \gamma \leq \frac{1}{\sqrt{2}}, \\ \frac{1}{\sqrt{2}} \sqrt{1 + \gamma^2 + \frac{1}{4\gamma^2}} & \text{if } \gamma \geq \frac{1}{\sqrt{2}}. \end{cases}$$

It is a consequence of Lemma A.2. We now describe the reach of  $Y$ :

$$\text{reach}(Y) = \begin{cases} \frac{1+2\gamma^2}{\sqrt{1+8\gamma^2}} & \text{if } \gamma \leq 1, \\ 1 & \text{if } \gamma \geq 1. \end{cases} \quad (18)$$

To prove this, we first define a *bottleneck* of  $Y$  as pair of distinct points  $(y, y') \in Y^2$  such that the open ball  $\mathcal{B}(\frac{1}{2}(y + y'), \frac{1}{2}\|y - y'\|)$  does not intersect  $Y$ . Its *length* is defined as  $\frac{1}{2}\|y - y'\|$ . According to the results of Aamari et al. (2019, Theorem 3.4), the reach of  $Y$  is equal to

$$\text{reach}(Y) = \min \left\{ \frac{1}{\rho}, \delta \right\},$$

where  $\frac{1}{\rho}$  is the inverse curvature of  $Y$ , and  $\delta$  is the minimal length of bottlenecks of  $Y$ . As we computed,  $\frac{1}{\rho}$  is equal to  $\frac{1+2\gamma^2}{\sqrt{1+8\gamma^2}}$ . Besides, according to Lemma A.2, a bottleneck  $(y_\theta, y_{\theta'})$  has to satisfy  $\theta' - \theta = \pi$  or  $\pm \arccos(-\frac{1}{2\gamma^2})$ . The smallest length is attained when  $\theta' - \theta = \pi$ , for which

$\frac{1}{2} \|y_\theta - y_{\theta'}\| = 1$ . It is straightforward to verify that the pair  $(y_\theta, y_{\theta'})$  is indeed a bottleneck. Therefore we have  $\delta = 1$ , and we deduce the expression of  $\text{reach}(Y)$ .

Last, the weak feature size of  $Y$  does not depend on  $\gamma$  and is equal to 1:

$$\text{wfs}(Y) = 1. \quad (19)$$

We shall prove it by using the characterization of Boissonnat et al. (2018):  $\text{wfs}(Y)$  is the infimum of distances  $\text{dist}(x, Y)$ , where  $x \in E$  is a critical point of the distance function  $d_Y$ . In this context,  $x$  is a critical point if it lies in the convex hull of its projections on  $Y$ . Remark that, if  $x \neq 0$ , then  $x$  admits at most two projections on  $Y$ . This follows from Lemma A.3. As a consequence, if  $x$  is a critical point, then there exists  $y, y' \in Y$  such that  $x$  lies in the middle of the segment  $[y, y']$ , and the open ball  $\mathcal{B}(x, \text{dist}(x, Y))$  does not intersect  $Y$ . Therefore  $y'$  is a critical point of  $y' \mapsto \|y - y'\|$ , hence Lemma A.2 gives that  $\|y - y'\| \geq 2$ . We deduce the result.

We now describe the thickenings  $Y^t$ . They present four different behaviours:

- $0 \leq t < 1$ :  $Y^t$  is homotopy equivalent to a circle,
- $1 \leq t < \frac{1}{2} \text{diam}(Y)$ :  $Y^t$  is homotopy equivalent to a circle,
- $\frac{1}{2} \text{diam}(Y) \leq t < \sqrt{1 + \frac{\gamma^2}{2}}$ :  $Y^t$  is homotopy equivalent to a 3-sphere,
- $t \geq \sqrt{1 + \frac{\gamma^2}{2}}$ :  $Y^t$  is homotopy equivalent to a point.

Recall that, in the case where  $\gamma \leq \frac{1}{\sqrt{2}}$ , we have  $\frac{1}{2} \text{diam}(Y) = 1$ . Consequently, the interval  $[1, \frac{1}{2} \text{diam}(Y))$  is empty, and the second point does not appear in this case.

*Study of the case  $0 \leq t < 1$ .* For  $t \in [0, 1)$ , let us show that  $Y^t$  deformation retracts on  $Y$ . According to Equation (19), we have  $\text{wfs}(Y) = 1$ . Moreover, Equation (18) gives that  $\text{reach}(Y) > 0$ . Using the results of Boissonnat et al. (2018), we deduce that  $Y^t$  is isotopic to  $Y$ .

*Study of the case  $1 \leq t < \frac{1}{2} \text{diam}(Y)$ .* Denote  $z_\theta = \left(0, 0, \frac{\gamma}{\sqrt{2}} \cos(2\theta), \frac{\gamma}{\sqrt{2}} \sin(2\theta)\right)$ , and define the circle  $Z = \{z_\theta, \theta \in [0, \pi)\}$ . It is represented in Figure 35.



Figure 35: Representation of the set  $Y$  (black) and the circle  $Z$  (red).

We claim that  $Y^t$  deformation retracts on  $Z$ . To prove so, we shall define a continuous application  $f: Y^t \rightarrow Z$  such that, for every  $y \in Y^t$ , the segment  $[y, f(y)]$  is included in  $Y^t$ . This would lead to a deformation retraction of  $Y^t$  onto  $Z$ , via

$$(s, y) \in [0, 1] \times Y^t \mapsto (1 - s)y + sf(y).$$

Equivalently, we shall define an application  $\Theta: Y^t \rightarrow [0, \pi)$  such that the segment  $[y, z_{\Theta(y)}]$  is included in  $Y^t$ .

Let  $y \in Y^t$ . According to Lemma A.3,  $y$  admits at most two projection on  $Y$ . We start with the case where  $y$  admits only one projection, namely  $y_\theta$  with  $\theta \in [0, 2\pi)$ . Let  $\bar{\theta} \in [0, \pi)$  be the reduction of  $\theta$  modulo  $\pi$ , and consider the point  $z_{\bar{\theta}}$  of  $Z$ . A computation shows that the distance  $\|y_\theta - z_{\bar{\theta}}\|$  is equal to 1. Besides, since  $y \in Y^t$ , the distance  $\|y_\theta - y\|$  is at most  $t$ . By

convexity, the segment  $[y, z_{\bar{\theta}}]$  is included in the ball  $\bar{B}(y_{\theta}, t)$ , which is a subset of  $Y^t$ . We then define  $\Theta(y) = \bar{\theta}$ .

Now suppose that  $y$  admits exactly two projection  $y_{\theta}$  and  $y_{\theta'}$ . According to Lemma A.2, these angles must satisfy  $\theta' - \theta = \pi$ . Indeed, the case  $\|y_{\theta} - y_{\theta'}\| = \sqrt{2}\sqrt{1 + \gamma^2 + \frac{1}{4\gamma^2}}$  does not occur since we chose  $t < \frac{1}{2}\text{diam}(Y) = \frac{\sqrt{2}}{2}\sqrt{1 + \gamma^2 + \frac{1}{4\gamma^2}}$ . The angles  $\theta$  and  $\theta'$  correspond to the same reduction modulo  $\pi$ , denoted  $\bar{\theta}$ , and we also define  $\Theta(y) = \bar{\theta}$ .

*Study of the case  $t \in \left[\frac{1}{2}\text{diam}(X), \sqrt{1 + \frac{\gamma^2}{2}}\right)$ .* Let  $\mathbb{S}_3$  denotes the unit sphere of  $E$ . For every  $v = (v_1, v_2, v_3, v_4) \in \mathbb{S}_3$ , we shall denote by  $\langle v \rangle$  the linear subspace spanned by  $v$ , and by  $\langle v \rangle_+$  the cone  $\{\lambda v, \lambda \geq 0\}$ . Moreover, we define the quantity

$$\delta(v) = \min_{y \in Y} \text{dist}(y, \langle v \rangle_+).$$

and the set

$$S = \{\delta(v)v, v \in \mathbb{S}_3\}.$$

The situation is depicted in Figure 36. We claim that  $S$  is a subset of  $Y^t$ , and that  $Y^t$  deform retracts on it. This follows from the two following facts: for every  $v \in \mathbb{S}_3$ ,

1.  $\delta(v)$  is not greater than  $\frac{1}{2}\text{diam}(Y)$ ,
2.  $\langle v \rangle_+ \cap Y^t$  consists of one connected component: an interval centered on  $\delta(v)v$ , that does not contain the point 0.

Suppose that these assertions are true. Then one defines a deformation retraction of  $Y^t$  on  $S$  by retracting each fiber  $\langle v \rangle_+ \cap Y^t$  linearly on the singleton  $\{\delta(v)v\}$ . We shall now prove the two items.

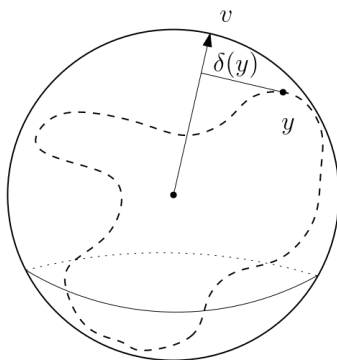


Figure 36: Representation of the set  $Y$  (dashed), lying on a 3-sphere of radius  $\sqrt{1 + \frac{\gamma^2}{2}}$ .

*Item 1.* Note that Item 1 can be reformulated as follows:

$$\max_{v \in \mathbb{S}_3} \min_{y \in Y} \text{dist}(y, \langle v \rangle_+) \leq \frac{1}{2}\text{diam}(Y). \quad (20)$$

Let us justify that the pairs  $(v, y)$  that attain this maximum-minimum are the same as in

$$\max_{v \in \mathbb{S}_3} \min_{y \in Y} \|y - v\|. \quad (21)$$

From the definition of  $Y = \{y_{\theta}, \theta \in [0, 2\pi)\}$ , we see that  $\min_{y \in Y} \text{dist}(y, \langle v \rangle_+) = \min_{y \in Y} \text{dist}(y, \langle v \rangle)$ . A vector  $v \in \mathbb{S}_3$  being fixed, let us show that  $y \mapsto \text{dist}(y, \langle v \rangle)$  is minimized

when  $y \mapsto \|v - y\|$  is. Let  $y \in Y$ . Since  $v$  is a unit vector, the projection of  $y$  on  $\langle v \rangle$  can be written as  $\langle y, v \rangle v$ . Hence  $\text{dist}(y, \langle v \rangle)^2 = \|\langle y, v \rangle v - y\|^2$ , and expanding this norm yields

$$\text{dist}(y, \langle v \rangle)^2 = \|y\|^2 - \langle y, v \rangle^2.$$

Expanding the norm  $\|y - v\|^2$  and using that  $\|y\|^2 = 1 + \frac{\gamma^2}{2}$ , we get  $\langle y, v \rangle = \frac{1}{2} \left( 2 + \frac{\gamma^2}{2} - \|y - v\|^2 \right)$ . We inject this relation in the preceding equation to obtain

$$\text{dist}(y, \langle v \rangle)^2 = - \left( \frac{\gamma}{2} \right)^4 + \gamma^2 + \frac{1}{4} \|y - v\|^2 \left( 4 + \gamma^2 - \|y - v\|^2 \right).$$

Now we can deduce that  $y \mapsto \text{dist}(y, \langle v \rangle)^2$  is minimized when  $y \mapsto \|y - v\|$  is minimized. Indeed, the map  $\|y - v\| \mapsto \frac{1}{4} \|y - v\|^2 \left( 4 + \gamma^2 - \|y - v\|^2 \right)$  is increasing on  $\left[ 0, \frac{1}{2}(4 + \gamma^2) \right]$ . But  $\|y - v\| \leq \|y\| + \|v\| = \frac{1}{2}(4 + \gamma^2)$ .

We deduce that studying the left hand term of Equation (20) is equivalent to studying Equation (21). We shall denote by  $g: \mathbb{S}_3 \rightarrow \mathbb{R}$  the map

$$g(v) = \min_{y \in Y} \|y - v\|. \quad (22)$$

Let  $v \in \mathbb{S}_3$  that attains the maximum of  $g$ , and let  $y$  be a corresponding point that attains the minimum of  $\|y - v\|$ . The points  $v$  and  $y$  attains the quantity in Equation (20). In order to prove that  $\text{dist}(y, \langle v \rangle) \leq \frac{1}{2} \text{diam}(Y)$ , let  $p(y)$  denotes the projection of  $y$  on  $\langle v \rangle$ . We shall show that there exists another point  $y' \in Y$  such that  $p(y)$  is equal to  $\frac{1}{2}(y + y')$ . Consequently, we would have  $\|y - p(y)\| = \frac{1}{2} \|y' - y\| \leq \frac{1}{2} \text{diam}(Y)$ , i.e.

$$\text{dist}(y, \langle v \rangle) \leq \frac{1}{2} \text{diam}(Y).$$

Remark the following fact: if  $w \in \mathbb{S}_3$  is a unit vector such that  $\langle p(y) - y, w \rangle > 0$ , then for  $\epsilon > 0$  small enough, we have

$$\text{dist}(y, \langle v + \epsilon w \rangle) > \text{dist}(y, \langle v \rangle).$$

Equivalently, this statement reformulates as  $0 \leq \left\langle y, \frac{1}{\|v + \epsilon w\|} (v + \epsilon w) \right\rangle < \langle y, v \rangle$ . Let us show that

$$\left\langle y, \frac{1}{\|v + \epsilon w\|} (v + \epsilon w) \right\rangle = \langle y, v \rangle - \epsilon \kappa + o(\epsilon), \quad (23)$$

where  $\kappa = \langle p(y) - y, w \rangle > 0$ , and where  $o(\epsilon)$  is the little-o notation. Note that  $\frac{1}{\|v + \epsilon w\|} = 1 - \epsilon \langle v, w \rangle + o(\epsilon)$ . We also have

$$\frac{1}{\|v + \epsilon w\|} (v + \epsilon w) = v + \epsilon (w - \langle v, w \rangle v) + o(\epsilon).$$

Expanding the inner product in Equation (23) gives

$$\begin{aligned} \left\langle y, \frac{1}{\|v + \epsilon w\|} (v + \epsilon w) \right\rangle &= \langle y, v \rangle + \epsilon \left( \langle y, w \rangle - \langle v, w \rangle \langle y, v \rangle \right) + o(\epsilon) \\ &= \langle y, v \rangle + \epsilon \left\langle y - \langle y, v \rangle v, w \right\rangle + o(\epsilon) \\ &= \langle y, v \rangle + \epsilon \langle y - p(y), w \rangle + o(\epsilon), \end{aligned}$$

and we obtain the result.

Next, let us prove that  $y$  is not the only point of  $Y$  that attains the minimum in Equation (22). Suppose that it is the case by contradiction. Let  $w \in \mathbb{S}_3$  be a unit vector such that

$\langle p(y) - y, w \rangle > 0$ . For  $\epsilon$  small enough, let us prove that the vector  $v' = \frac{1}{\|v + \epsilon w\|}(v + \epsilon w)$  of  $\mathbb{S}_3$  contradicts the maximality of  $v$ . That is, let us prove that  $g(v') > g(v)$ . Let  $y' \in Y$  be a minimizer  $\|y' - v'\|$ . We have to show that  $\|y' - v'\| > \|y - v\|$ . This would lead to  $g(v') > g(v)$ , hence the contradiction.

Expanding the norm yields

$$\|v' - y'\|^2 = \|v' - v + v - y'\|^2 \geq \|v' - v\|^2 + \|v - y'\|^2 - 2\langle v' - v, v - y' \rangle.$$

Using  $\|v' - v\|^2 \geq 0$  and  $\|v - y'\|^2 \geq \|v - y\|^2$  by definition of  $y$ , we obtain

$$\|v' - y'\|^2 \geq \|v - y\|^2 - 2\langle v' - v, v - y' \rangle.$$

We have to show that  $\langle v' - v, v - y' \rangle$  is positive for  $\epsilon$  small enough. By writing  $v - y' = v - y + (y - y')$  we get

$$\langle v' - v, v - y' \rangle = \langle v' - v, v \rangle - \langle v' - v, y \rangle + \langle v' - v, y - y' \rangle$$

According to Equation (23),  $-\langle v' - v, y \rangle = \epsilon\kappa + o(\epsilon)$ . Besides, using  $v' - v = \epsilon(w - \langle v, w \rangle v) + o(\epsilon)$ , we get  $\langle v' - v, v \rangle = o(\epsilon)$ . Last, Cauchy-Schwarz inequality gives  $|\langle v' - v, y - y' \rangle| \leq \|v' - v\| \|y - y'\|$ . Therefore,  $\langle v' - v, y - y' \rangle = O(\epsilon) \|y - y'\|$ , where  $O(\epsilon)$  is the big-o notation. Gathering these three equalities, we obtain

$$\langle v' - v, v - y' \rangle = o(\epsilon) + \epsilon\kappa + O(\epsilon) \|y - y'\|.$$

As we can read from this equation, if  $\|y - y'\|$  goes to zero as  $\epsilon$  does, then  $\langle v' - v, v - y' \rangle$  is positive for  $\epsilon$  small enough. Observe that  $v'$  goes to  $v$  when  $\epsilon$  goes to 0. By assumption  $y$  is the only minimizer in Equation (22). By continuity of  $g$ , we deduce that  $y'$  goes to  $y$ .

By contradiction, we deduce that there exists another point  $y'$  which attains the minimum in  $g(v)$ . Note that it is the only other one, according to Lemma A.3. Let us show that  $p(y)$  lies in the middle of the segment  $[y, y']$ . Suppose that it is not the case. Then  $p(y) - y$  is not equal to  $-(p(y') - y')$ , where  $p(y')$  denotes the projection of  $y'$  on  $\langle v \rangle$ . Consequently, the half-spaces  $\{w \in E, \langle p(y) - y, w \rangle > 0\}$  and  $\{w \in E, \langle p(y') - y', w \rangle > 0\}$  intersects. Let  $w$  be any vector in the intersection. For  $\epsilon > 0$ , denote  $v' = \frac{1}{\|v + \epsilon w\|}(v + \epsilon w)$ . If  $\epsilon$  is small enough, the same reasoning as before shows that  $v'$  contradicts the maximality of  $v$ . The situation is represented in Figure 37.

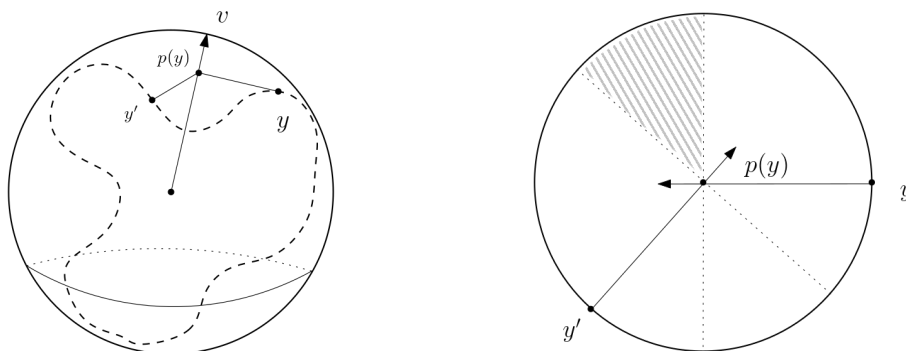


Figure 37: Left: Representation of the situation where  $y$  and  $y'$  are minimizers of Equation (22). Right: Representation in the plane passing through the points  $y$ ,  $y'$  and  $p(y)$ . The dashed area corresponds to the intersection of the half-spaces  $\{w \in E, \langle p(y) - y, w \rangle > 0\}$  and  $\{w \in E, \langle p(y') - y', w \rangle > 0\}$ .

Item 2. Let  $v \in \mathbb{S}_3$ . The set  $\langle v \rangle_+ \cap Y^t$  can be described as

$$\langle v \rangle_+ \cap \bigcup_{y \in Y} \overline{B}(y, t).$$



Let  $y \in Y$  such that  $\langle v \rangle_+ \cap \overline{\mathcal{B}}(y, t) \neq \emptyset$ . Denote by  $p(y)$  the projection of  $y$  on  $\langle v \rangle_+$ . It is equal to  $\langle y, v \rangle v$ . Using Pythagoras' theorem, we obtain that the set  $\langle v \rangle_+ \cap \overline{\mathcal{B}}(y, t)$  is equal to the interval

$$\left[ p(y) \pm \sqrt{t^2 - \text{dist}(y, \langle v \rangle)^2} v \right].$$

Using the identity  $\text{dist}(y, \langle v \rangle)^2 = \|y\|^2 - \langle y, v \rangle^2 = 1 + \frac{\gamma^2}{2} - \langle y, v \rangle^2$ , we can write this interval as

$$[I_1(y) \cdot v, I_2(y) \cdot v],$$

where  $I_1(y) = \langle y, v \rangle - \sqrt{\langle y, v \rangle^2 - (1 + \frac{\gamma^2}{2} - t^2)}$  and  $I_2(y) = \langle y, v \rangle + \sqrt{\langle y, v \rangle^2 - (1 + \frac{\gamma^2}{2} - t^2)}$ . Seen as functions of  $\langle y, v \rangle$ , the map  $I_1$  is decreasing, and the map  $I_2$  is increasing (see Figure 38). Let  $y^* \in Y$  that minimizes  $\text{dist}(y, \langle v \rangle)$ . Equivalently,  $y^*$  maximizes  $\langle y, v \rangle$ . It follows that the corresponding interval  $[I_1(y^*) \cdot v, I_2(y^*) \cdot v]$  contains all the others. We deduce that the set  $\langle v \rangle_+ \cap Y^t$  is equal to this interval.

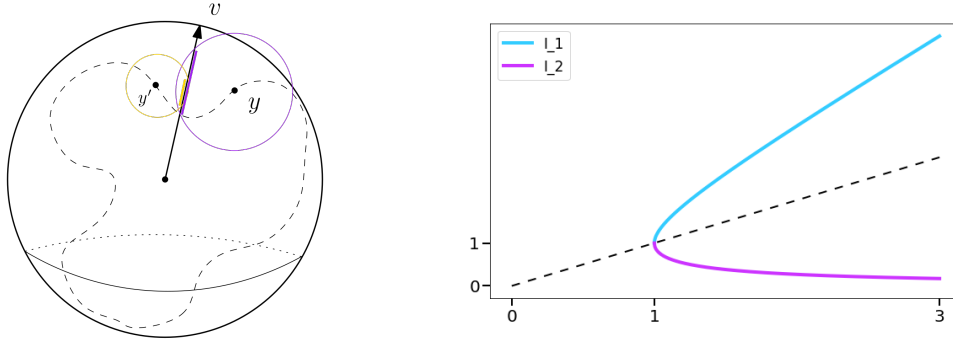


Figure 38: Left: Representation of two intervals  $\langle v \rangle_+ \cup \overline{\mathcal{B}}(y, t)$  and  $\langle v \rangle_+ \cup \overline{\mathcal{B}}(y', t)$ . Right: Representation of the maps  $x \mapsto x \pm \sqrt{x^2 - 1}$ .

*Study of the case  $t \geq \sqrt{1 + \frac{1}{2}\gamma^2}$ .* For every  $y \in Y$ , we have  $\|y\| = \sqrt{1 + \frac{1}{2}\gamma^2}$ . Therefore, if  $t \geq \sqrt{1 + \frac{1}{2}\gamma^2}$ , then  $Y^t$  is star shaped around the point 0, hence it deformation retracts on it.

To close this subsection, let us study the Čech bundle filtration  $(\mathbb{Y}, \mathbb{p})$  of  $Y$ . According to Equation (11), its filtration maximal value is  $t^{\max}(Y) = t_{\gamma}^{\max}(X) = \frac{\gamma}{\sqrt{2}}$ . Note that  $\frac{\gamma}{\sqrt{2}}$  is lower than  $\frac{1}{2}\text{diam}(Y)$ . Consequently, only two cases are to be studied:  $t \in [0, 1)$ , and  $t \in [1, \frac{1}{2}\text{diam}(Y))$ .

The same argument as in Subsect. A.2 yields that for every  $t \in [0, 1)$ , the persistent Stiefel-Whitney class  $w_1^t(Y)$  is equal to  $w_1^0(Y)$ . Accordingly, for every  $t \in [1, \frac{1}{2}\text{diam}(Y))$ , the class  $w_1^t(Y)$  is equal to  $w_1^1(Y)$ . Let us show that  $w_1^0(Y)$  is zero, and that  $w_1^1(Y)$  is not.

First, remark that the map  $p^0: Y \rightarrow \mathcal{G}_1(\mathbb{R}^2)$  can be seen as the normal bundle of the circle. Hence  $(p^0)^*: H^*(Y) \leftarrow H^*(\mathcal{G}_1(\mathbb{R}^2))$  is nontrivial, and we deduce that  $w_1^0(Y) = 0$ . As a consequence, the persistent Stiefel-Whitney class  $w_1^t(X)$  is nonzero for every  $t < 1$ .

Next, consider  $p^1: Y^1 \rightarrow \mathcal{G}_1(\mathbb{R}^2)$ . Recall that  $Y^1$  deformation retracts on the circle

$$Z = \left\{ \left( 0, 0, \frac{\gamma}{\sqrt{2}} \cos(2\theta), \frac{\gamma}{\sqrt{2}} \sin(2\theta) \right), \theta \in [0, \pi) \right\}.$$

Seen in  $\mathbb{R}^2 \times M(\mathbb{R}^2)$ , we have

$$Z = \left\{ \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \gamma \begin{pmatrix} \cos(\theta)^2 & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin(\theta)^2 \end{pmatrix} \right), \theta \in [0, \pi) \right\}.$$

Notice that the map  $q: Z \rightarrow \mathcal{G}_1(\mathbb{R}^2)$ , the projection on  $\mathcal{G}_1(\mathbb{R}^2)$ , is injective. Seen as a map between two circles, it has degree (modulo 2) equal to 1. We deduce that  $q^*: H^*(Z) \leftarrow H^*(\mathcal{G}_1(\mathbb{R}^2))$  is nontrivial. Now, remark that the map  $q$  factorizes through  $p^1$ :

$$\begin{array}{ccc} Z & \xleftarrow{\quad} & Y^1 \\ & \searrow q & \swarrow p^1 \\ & \mathcal{G}_1(\mathbb{R}^2) & \end{array}$$

It induces the following diagram in cohomology:

$$\begin{array}{ccc} H^*(Z) & \xleftarrow{\quad \sim \quad} & H^*(Y^1) \\ & \searrow q^* & \swarrow (p^1)^* \\ & H^*(\mathcal{G}_1(\mathbb{R}^2)) & \end{array}$$

Since  $q^*$  is nontrivial, this commutative diagram yields that the persistent Stiefel-Whitney class  $w_1^1(Y)$  is nonzero. As a consequence, the persistent Stiefel-Whitney class  $w_1^t(Y)$  is nonzero for every  $t \geq 1$ .

## References

- Aamari E, Kim J, Chazal F, Michel B, Rinaldo A, Wasserman L (2019) Estimating the Reach of a Manifold. *Electronic journal of statistics*
- Aubrey H (2011) Persistent cohomology operations. PhD thesis, Duke University
- Bauer U, Edelsbrunner H (2017) The Morse theory of Čech and Delaunay complexes. *Transactions of the American Mathematical Society* 369(5):3741–3762
- Bell G, Lawson A, Martin J, Rudzinski J, Smyth C (2019) Weighted persistent homology. *Involve, a Journal of Mathematics* 12(5):823–837
- Boissonnat JD, Chazal F, Yvinec M (2018) *Geometric and topological inference*, vol 57. Cambridge University Press
- Botnan M, Crawley-Boevey W (2020) Decomposition of persistence modules. *Proceedings of the American Mathematical Society* 148(11):4581–4596
- Chazal F, Cohen-Steiner D, Lieutier A (2009) A sampling theory for compact sets in Euclidean space. *Discrete & Computational Geometry* 41(3):461–479
- Chazal F, de Silva V, Glisse M, Oudot S (2016) *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics
- Edelsbrunner H (1993) The union of balls and its dual shape. In: *Proceedings of the ninth annual symposium on Computational geometry*, pp 218–231
- Govc D, Marzantowicz W, Pavešić P, et al. (2020) How many simplices are needed to triangulate a Grassmannian? *Topological Methods in Nonlinear Analysis*
- Hatcher A (2002) *Algebraic Topology*. Cambridge University Press
- von Kühnel W (1987) Minimal triangulations of kummer varieties. In: *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, Springer, vol 57, pp 7–20
- Milnor J, Stasheff JD (2016) *Characteristic Classes*.(AM-76), vol 76. Princeton university press
- Munkres JR (1984) *Elements of Algebraic Topology*. Addison-Wesley

Perea JA (2018) Multiscale projective coordinates via persistent cohomology of sparse filtrations. *Discrete & Computational Geometry* 59(1):175–225

Scoccola L, Perea JA (2021) Approximate and discrete euclidean vector bundles. arXiv preprint arXiv:210407563