



**HAL**  
open science

# Persistence Diagram Estimation: Beyond Plug-in Approaches

Hugo Henneuse

► **To cite this version:**

Hugo Henneuse. Persistence Diagram Estimation: Beyond Plug-in Approaches. 2024. hal-04593678v3

**HAL Id: hal-04593678**

**<https://hal.science/hal-04593678v3>**

Preprint submitted on 27 Aug 2024

**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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

# Persistence Diagram Estimation : Beyond Plug-in Approaches

Hugo Henneuse <sup>1,2</sup>

[hugo.henneuse@universite-paris-saclay.fr](mailto:hugo.henneuse@universite-paris-saclay.fr)

August 27, 2024

## Abstract

Persistent homology is a tool from Topological Data Analysis (TDA) used to summarize the topology underlying data. It can be conveniently represented through persistence diagrams. Observing a noisy signal, common strategies to infer its persistence diagram involve plug-in estimators, and convergence properties are then derived from sup-norm stability. This dependence on the sup-norm convergence of the preliminary estimator is restrictive, as it essentially imposes to consider regular classes of signals. Departing from these approaches, we design an estimator based on image persistence. In the context of the Gaussian white noise model, and for large classes of piecewise-constant signals, we prove that the proposed estimator is consistent and achieves parametric rates.

## Introduction

Topological Data Analysis (TDA) is a field that aims to provide representations to describe the “shape” of data. A central tool in TDA is persistent homology and its representation through persistence diagrams. Persistent homology permits to encode the evolution of topological features (in the homology sense) along a family of nested spaces, called filtration. Moving along indices, topological features (connected components, cycles, cavities, ...) can appear or die (existing connected components merge, cycles or cavities are filled, ...). Persistence diagrams then offer a practical, multiscale, summary of the topology underlying data. Although in its early years research on persistent homology has primarily focused on deterministic settings, it didn’t take long for statistical questions to emerge. A prominent topic is the estimation of persistence diagram and, more specifically, the investigation of convergence properties of considered estimators.

A first historical success in this direction is [Bubenik and Kim \(2006\)](#), that formalizes the problem of persistence diagram estimation from the (sub or super) level sets of a density, in several parametric settings. Notably, when observing points sampled from a Mises-Fischer distribution with fixed parameter  $\kappa$  on  $S^{p-1}$ , they show that a plug-in estimator from the maximum-likelihood estimation of  $\kappa$  permits to recover consistently the persistence diagram of the density.

[Fasy et al. \(2014\)](#) consider the problem of estimating the persistent homology of the support  $M$  of a distribution  $P$  (associated to its offset filtration  $(B_2(M, r))_{r \geq 0}$ ) while observing sampled points from  $P$ . Under geometric control over both  $M$  and  $P$  they show that the persistence diagram associated to the offsets of the sampled points provides a consistent estimator. They additionally derive confidence sets, showing that it suffices to remove points lying on a small band around the diagonal

---

<sup>1</sup>Laboratoire de Mathématiques d’Orsay, Université Paris-Saclay, Orsay, France

<sup>2</sup>DataShape, Inria Saclay, Palaiseau, France

of the persistence diagram to separate topological noise from true topological features. Their analysis exploits (a variant of) the sup-norm stability theorem (Barmannik, 1994; Cohen-Steiner et al., 2005; Chazal et al., 2009), showing that, the set of sampled points converges to  $M$  in Hausdorff distance. On the same problem, Chazal et al. (2014) provide consistent estimators and show that they achieve (nearly) minimax rates under more general assumptions. This work also exploits the sup-norm stability.

In the context of the non-parametric regression, closer to the setting considered in this work, Bubenik et al. (2009) study a plug-in estimator based on kernel estimation for the persistence diagram associated to sublevel sets of the regression function. Exploiting the sup-norm stability again, they prove the consistency of this estimator and that it achieves minimax rates on Hölder spaces. These rates coincide with the known minimax rates for the estimation of the regression function in sup-norm. It essentially means that the estimation of the persistence diagram from sublevel sets is as difficult as estimating the regression function in sup norm over Hölder spaces.

As highlighted, a common aspect of all the previously cited works is that they adopt plug-in approaches and convergence results are established through the sup-norm stability theorem (or variants of the sup-norm stability theorem). This dependence on the convergence of the preliminary estimators is limiting as it essentially imposes to consider regular objects (signal, regression function, density, ...). This limitation motivates investigations of persistent homology inference over wider classes, typically where consistent estimation of such objects in sup-norm is no longer possible. A difficulty is that it requires to depart from approaches that rely on sup-norm stability.

A first step in this direction is proposed by Bobrowski et al. (2017). For the non-parametric regression and the density model, they propose a non-plug-in approach to infer the persistence diagrams coming from the superlevel sets filtration of a function  $f$ , respectively, the regression function and the density. They consider, for all,  $L \in \mathbb{R}$ ,  $\widehat{\mathcal{D}}_L$  an estimator of the superlevel set  $f^{-1}([L, +\infty[)$  based on kernels. Instead of directly estimating the underlying persistent modules by taking the homology groups of the filtration  $(\widehat{\mathcal{D}}_L)_{L \in \mathbb{R}}$ , they consider, for all  $s \in \mathbb{N}$  and  $\lambda \in \mathbb{R}$ , the map :

$$i_{s,L} : H_s \left( \widehat{\mathcal{D}}_{L+\epsilon} \right) \rightarrow \widehat{\mathcal{D}}_{L-\epsilon}$$

induced by the inclusion  $\widehat{\mathcal{D}}_{L+\epsilon} \subset \widehat{\mathcal{D}}_{L-\epsilon}$ , with  $\epsilon > 0$  a chosen precision parameter. And then consider the modules  $(\text{Im}(i_{s,L}))_{L \in \mathbb{R}}$ . This additional step of constructing an image persistent module can be thought as a topological regularization, which allows considering wide classes of signal. In particular, supposing that  $f$  is bounded, and the module associated to its superlevel sets filtration is  $q$ -tame (see definition 3 in Section 1), they show that the bottleneck distance (see definition 4 in Section 1) between the diagram induced by the proposed image persistent module and the true persistence diagram is bounded by  $5\epsilon$  with high probability (depending also on  $\epsilon$ ). Their analysis relies on a weaker notion of stability : the algebraic stability (Chazal et al., 2009). Yet, in this framework the optimal calibration of  $\epsilon$  is not obvious, and thus this work does not permit to establish convergence rates or prove proper consistency. The difficulty stems from their framework being too wide, allowing for wildly irregular functions. Typically, it is easy to construct two functions satisfying the  $q$ -tameness and boundedness assumptions, that only differ on a null set (and thus statistically indistinguishable), that have (arbitrarily) far persistence diagrams in bottleneck distance.

Hence, it is interesting to identify narrower classes on which we can provide stronger and more precise results. Recently, in Henneuse (2024), we proposed a new approach that also departs from the sup norm stability. This work studies the inference of persistence diagram, from a minimax perspective, for the Gaussian white noise model and the non-parametric regression. Although we consider a plug-in estimator, our analysis of convergence properties did not rely on sup-norm stability but on algebraic stability. We introduced and studied classes of piecewise-Hölder continuous

functions with discontinuities set having a positive reach. The reach (Federer, 1959) is a popular curvature measure in geometric inference, that can be thought as a way to describe the geometric regularity of a set. Under this reach assumption, we manage to show that a histogram plug-in estimator permits to achieve the well known minimax rates for Hölder-continuous functions. Still, the positive reach assumption imposes some limitations on the shape of discontinuity sets, typically it does not allow corners or multiple points (i.e. self intersections of the discontinuities set). Relaxation of this assumption was already discussed in this earlier work. We highlighted that it would imply to consider other estimators, as, even in the noiseless setting, the histogram approximation falls short on simple examples (see Figure 5). This serves as a motivation to depart from plug-in approaches and as the starting point for this new work.

We consider the Gaussian white noise model given by the following stochastic equation,

$$dX_{t_1, \dots, t_d} = f(t_1, \dots, t_d) dt_1 \dots dt_d + \theta dW_{t_1, \dots, t_d}$$

with  $W$  a  $d$ -parameters Wiener field,  $f : [0, 1]^d \rightarrow \mathbb{R}$  a signal and  $\theta \geq 0$  the level of noise. In this context, our goal is to estimate  $\text{dgm}(f)$ , the persistence diagram coming from the sublevel sets of  $f$ .

For a set  $A \subset [0, 1]^d$ ,  $\bar{A}$  we denote its adherence and  $\partial A$  its boundary. We suppose that  $f$  verifies the following assumptions :

- **A1.**  $f$  is a piecewise constant function, i.e. there exist  $M_1, \dots, M_l$  open sets of  $[0, 1]^d$  and  $\lambda_1, \dots, \lambda_l$  in  $\mathbb{R}$  such that  $\bigcup_{i=1}^l \bar{M}_i = [0, 1]^d$  and  $f|_{M_i} = \lambda_i, \forall i \in \{1, \dots, l\}$ .
- **A2.**  $f$  verifies,  $\forall x_0 \in [0, 1]^d$ ,

$$\liminf_{x \in \bigcup_{i=1}^l M_i \rightarrow x_0} f(x) = f(x_0).$$

In this context, two signals, differing only on a null set, are statistically undistinguishable. Persistent homology is sensitive to point-wise irregularity, two signals differing only on a null set can have persistence diagrams that are arbitrarily far. Assumption **A2** prevents such scenario.

- **A3.** Let  $\mu \in ]0, 1]$  and  $R_\mu > 0$ , for all  $I \subset \{1, \dots, l\}$

$$\text{reach}_\mu \left( \bigcup_{i \in I} \partial M_i \right) \geq R_\mu$$

with  $\text{reach}_\mu$  denoting the  $\mu$ -reach.

The class of functions verifying **A1**, **A2** and **A3** is denoted  $S_d(\mu, R_\mu)$ . For our purpose, we show in Appendix A that persistence diagrams of signals in  $S_d(\mu, R_\mu)$  are well-defined. Compared to the classes, we introduced Henneuse (2024), our analytic assumption over  $f$  on the regular region are stronger, but we significantly relax the assumption on the geometry of the discontinuities allowing multiple points and corners in the discontinuities set.

## Contribution

We complete the main result of Henneuse (2024). Unlike the classes considered there, over the classes  $S_d(\mu, R_\mu)$ , a plug-in estimator from histogram estimation is no longer consistent (see Figure

5). To overcome this issue, we propose, a non-plug-in estimator of the persistence diagram of  $f$ , based on image persistence (Cohen-Steiner et al., 2009), denoted  $\widehat{\text{dgm}}(f)$ . Over  $S_d(\mu, R_\mu)$ , we show that this estimator is consistent and achieves parametric convergence rates. More precisely, our main result is the following theorem.

**Theorem 1.** *There exist  $\tilde{C}_0$  and  $\tilde{C}_1$  such that, for all  $t > 0$ ,*

$$\mathbb{P} \left( \sup_{f \in S_d(\mu, R_\mu)} d_b \left( \widehat{\text{dgm}}(f), \text{dgm}(f) \right) \geq t\theta \right) \leq \tilde{C}_0 \exp \left( -\tilde{C}_1 t^2 \right).$$

The paper is organized as follows. Section 1 recalls the necessary background on geometric measure theory and persistent homology, Section 2 describes our estimation procedure, and Section 3 is dedicated to the proof of Theorem 1. Proofs of technical lemmas can be found in appendix.

## 1 Background

This section provides the necessary background to follow this paper.

### 1.1 Persistent Homology

We here present briefly some notions related to persistence homology. For a broader overview and visual illustrations of persistent homology, we recommend Chazal and Michel (2021). For detailed and rigorous constructions, see Chazal et al. (2016). Additionally, since the construction discussed here involves (singular) homology, the reader can refer to Hatcher (2000).

**Definition 1.** Let  $\Lambda \subset \mathbb{R}$  be a set of indices. A **filtration** over  $\Lambda$  is a family  $(\mathcal{K}_\lambda)_{\lambda \in \Lambda}$  of topological spaces satisfying,  $\forall \lambda, \lambda' \in \Lambda, \lambda \leq \lambda'$ ,

$$\mathcal{K}_\lambda \subset \mathcal{K}_{\lambda'}.$$

A typical filtration that we will consider in this paper is, for a function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ , the family of sublevel sets  $(\mathcal{F}_\lambda)_{\lambda \in \mathbb{R}} = (f^{-1}([-\infty, \lambda]))_{\lambda \in \mathbb{R}}$ . The associated family of homology groups of degree  $s \in \mathbb{N}$ ,  $\mathbb{V}_{f,s} = (H_s(\mathcal{F}_\lambda))_{\lambda \in \mathbb{R}}$ , equipped with  $v_\lambda^s$  the linear application induced by the inclusion  $\mathcal{F}_\lambda \subset \mathcal{F}_{\lambda'}$ , for all  $\lambda \leq \lambda'$ , forms a persistence module. To be more precise, in this paper,  $H_s(\cdot)$  is the singular homology functor in degree  $s$  with coefficients in a field (typically  $\mathbb{Z}/2\mathbb{Z}$ ). Hence,  $H_s(\mathcal{F}_\lambda)$  is a vector space.

Still, the estimator we propose in Section 2, relies on image module, that does not come (directly) from sublevel sets filtration. Hence, we need the following, more general, definition.

**Definition 2.** Let  $\Lambda \subset \mathbb{R}$  be a set of indices. A **persistence module** over  $\lambda$  is a family  $\mathbb{V} = (\mathbb{V}_\lambda)_{\lambda \in \Lambda}$  of vector spaces equipped with linear application  $v_\lambda^{\lambda'} : \mathbb{V}_\lambda \rightarrow \mathbb{V}_{\lambda'}$  such that,  $\forall \lambda \leq \lambda' \leq \lambda'' \in \Lambda$ ,

$$v_\lambda^{\lambda'} = id$$

and

$$v_{\lambda'}^{\lambda''} \circ v_\lambda^{\lambda'} = v_\lambda^{\lambda''}.$$

Under  $q$ -tameness of the persistence module, it is possible to show that the algebraic structure of the persistence module encodes exactly the evolution of the topological features along the indices  $\Lambda$ .

**Definition 3.** A persistence module  $\mathbb{V}$  is said to be  $q$ -tame if  $\forall \lambda < \lambda' \in \Lambda, \text{rank}(v_\lambda^{\lambda'})$  is finite. By extension, when considering the persistence modules  $(\mathbb{V}_{f,s})_{s \in \mathbb{N}}$  coming from the sublevel sets filtration of a real function  $f$ , we say that  $f$  is  $q$ -tame if  $\mathbb{V}_{f,s}$  is for all  $s \in \mathbb{N}$ .

Furthermore, the algebraic structure of the persistence module can be summarized by a collection  $\{(b_i, d_i), i \in I\} \subset \overline{\mathbb{R}}^2$ , which defines the **persistence diagram**. Following previous remarks,  $b_i$  corresponds to the birth time of a topological feature,  $d_i$  to its death time and  $d_i - b_i$  to its lifetime. For a detailed construction of persistence diagrams, see [Chazal et al. \(2016\)](#).

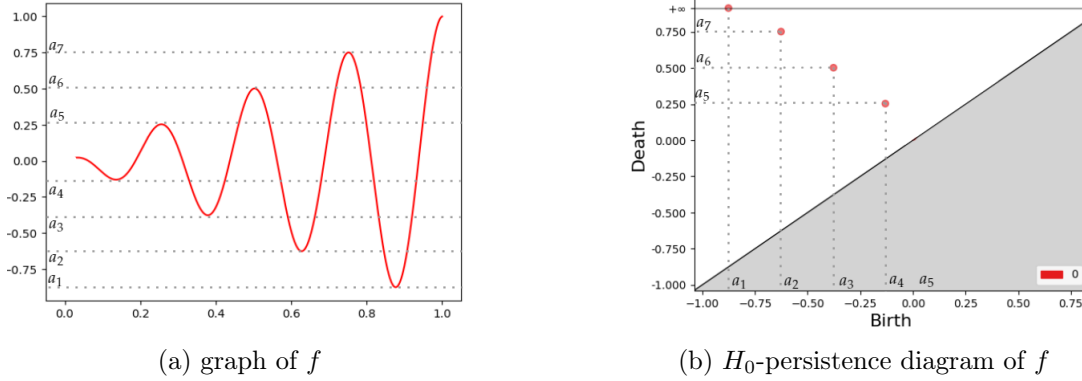


Figure 1: Graph of  $f(x) = x \cos(8\pi x)$  over  $[0, 1]$  and the persistence diagram associated to its sublevel sets filtration.  $a_1, \dots, a_4$  correspond to local minima of  $f$  and thus birth times in  $\text{dgm}(f)$ .  $a_5, \dots, a_7$  correspond to local maxima of  $f$  and thus death times in  $\text{dgm}(f)$ .

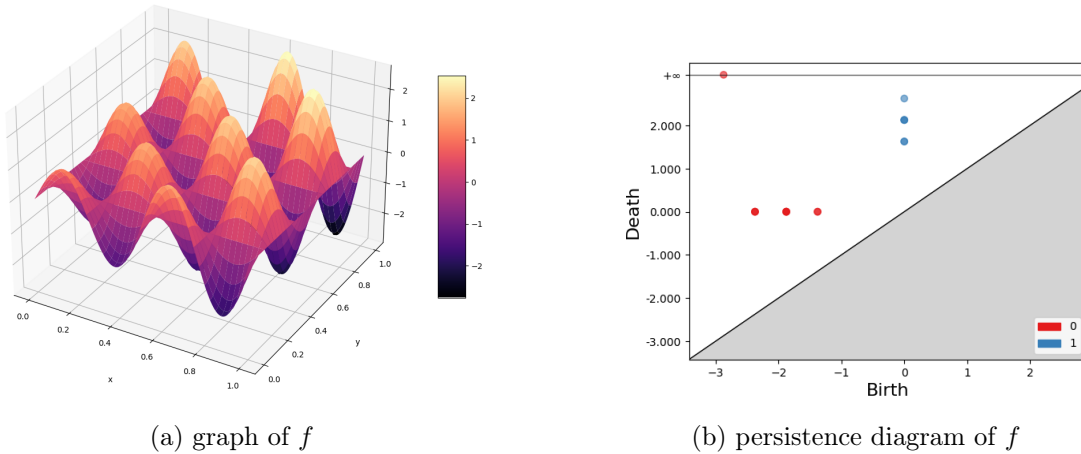


Figure 2: Graph of  $f(x, y) = \sin(4\pi x) \cos(4\pi y)(x + y + 1)$  over  $[0, 1]^2$  and the persistence diagram associated to its sublevel sets filtration, red points corresponds to the  $H_0$ -persistence diagram and blue points to the  $H_1$ -persistence diagram.

To compare persistence diagrams, a popular distance, especially in statistical contexts, is the bottleneck distance.

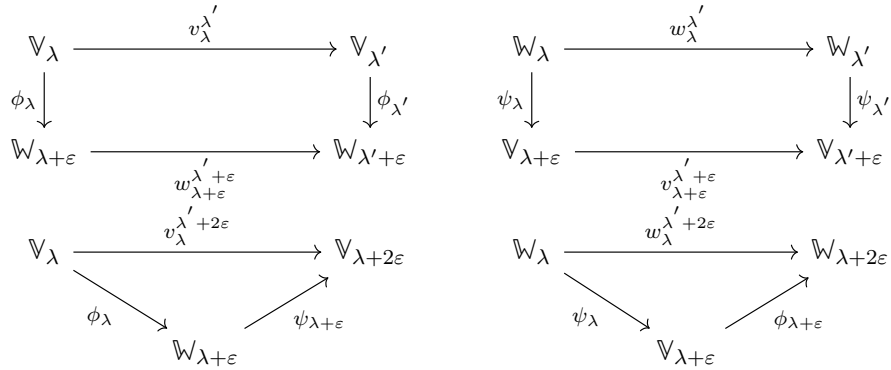
**Definition 4.** The **bottleneck distance** between two persistence diagrams  $D_1$  and  $D_2$  is,

$$d_b(D_1, D_2) = \inf_{\gamma \in \Gamma} \sup_{p \in D_1} \|\gamma(p) - p\|_\infty$$

with  $\Gamma$  the set of all bijections between  $D_1$  and  $D_2$  (both enriched with the diagonal).

We now introduce the algebraic stability theorem for the bottleneck distance. This theorem was the key for proving upper bounds in [Henneuse \(2024\)](#), it also plays a crucial role, here, in proving [Theorem 1](#). This theorem relies on interleaving between modules, a notion we extensively exploit.

**Definition 5.** Two persistence modules  $\mathbb{V} = (\mathbb{V}_\lambda)_{\lambda \in \Lambda \subset \mathbb{R}}$  and  $\mathbb{W} = (\mathbb{W}_\lambda)_{\lambda \in \Lambda \subset \mathbb{R}}$  are said to be  $\varepsilon$ -**interleaved** if there exist two families of applications  $\phi = (\phi_\lambda)_{\lambda \in \Lambda \subset \mathbb{R}}$  and  $\psi = (\psi_\lambda)_{\lambda \in \Lambda \subset \mathbb{R}}$  where  $\phi_\lambda : \mathbb{V}_\lambda \rightarrow \mathbb{W}_{\lambda+\varepsilon}$ ,  $\psi_\lambda : \mathbb{W}_\lambda \rightarrow \mathbb{V}_{\lambda+\varepsilon}$ , and for all  $\lambda < \lambda'$  the following diagrams commute,



**Theorem ([Chazal et al. 2009](#), "algebraic stability").** Let  $\mathbb{V}$  and  $\mathbb{W}$  two  $q$ -tame persistence modules. If  $\mathbb{V}$  and  $\mathbb{W}$  are  $\varepsilon$ -interleaved then,

$$d_b(\text{dgm}(\mathbb{V}), \text{dgm}(\mathbb{W})) \leq \varepsilon.$$

We now give a corollary of this result, proved earlier in special cases ([Barannikov, 1994](#); [Cohen-Steiner et al., 2005](#)). We insist on the fact that this is a strictly weaker result than algebraic stability.

**Theorem ("sup norm stability").** Let  $f$  and  $g$  two real-valued  $q$ -tame function, for all  $s \in \mathbb{N}$

$$d_b(\text{dgm}(\mathbb{V}_{f,s}), \text{dgm}(\mathbb{V}_{g,s})) \leq \|f - g\|_\infty.$$

This last property is often used to upper bound the errors (in bottleneck distance) of "plug-in" estimators of persistence diagrams. It enables the direct translation of convergence rates in sup-norm to convergence rates in bottleneck distance, which for regular classes of signals (e.g. Hölder spaces) provides minimax upper bounds. Still, for wider classes, this approach falls short. The alternative approach proposed in this work focuses instead on the estimation of the persistent modules  $\mathbb{V}_{f,s}$ , in the sense of interleaving, and then exploits algebraic stability.

## 1.2 Generalized gradient

We here present some concepts from geometric measure theory, extensively used in geometric inference and TDA. The first notion used in this paper is the distance function to a compact set.

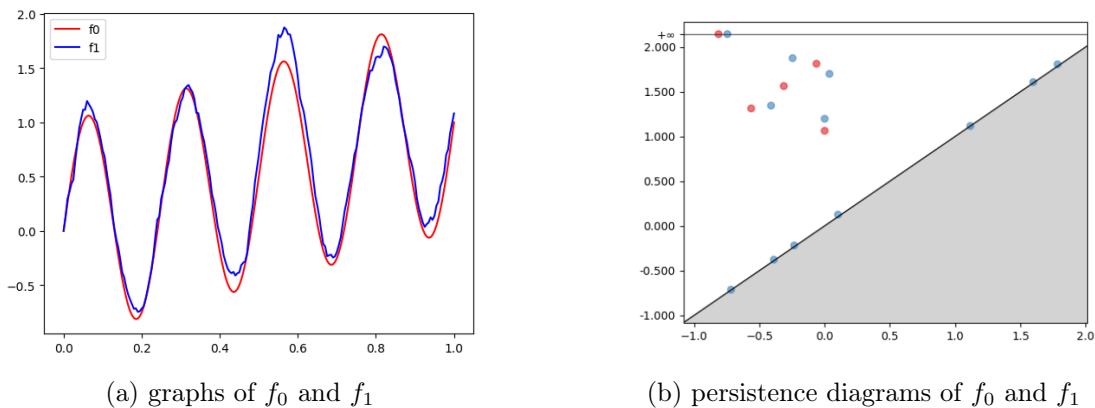


Figure 3: 1D Illustration of stability theorems.

**Definition 6.** Let  $K \subset [0, 1]^d$  a compact set, the **distance function**  $d_K$  is given by,

$$d_K : x \mapsto \min_{y \in K} \|x - y\|_2.$$

Generally, the distance function is not differentiable everywhere, still it is possible to define a generalized gradient function that matches the gradient at points where the distance function is differentiable.

**Definition 7.** Let,

$$\Gamma_K(x) = \{y \in K \mid \|x - y\|_2 = d_K(x)\}$$

the set of closest points to  $x$  in  $K$ . For  $x \in [0, 1]^d \setminus K$ , let denote  $\Theta_K(x)$  the center of the unique smallest ball enclosing  $\Gamma_K(x)$ , the **generalized gradient function**  $\nabla_K(x)$  is defined as,

$$\nabla_K(x) = \frac{x - \Theta_K(x)}{d_K(x)}.$$

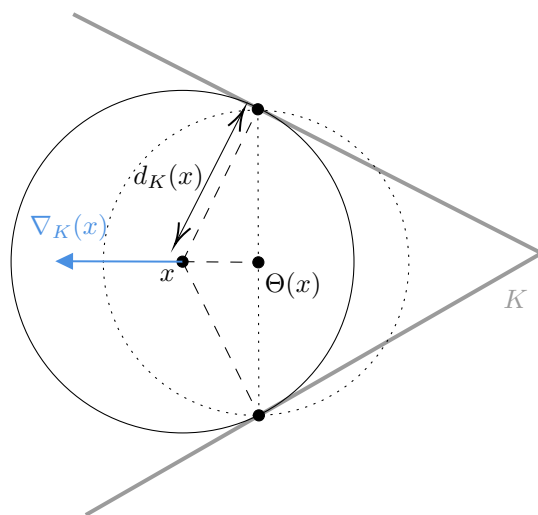


Figure 4: 2D example with 2 closest points

With these definitions we can now introduce the notion of  $\mu$ -reach (Chazal et al., 2006).



**Definition 8.** Let  $K \subset [0, 1]^d$  a compact set, its  $\mu$ -reach  $\text{reach}_\mu(K)$  is defined by,

$$\text{reach}_\mu(K) = \inf \left\{ r \mid \inf_{d_K^{-1}(r) \setminus K} \|\nabla_K\|_2 < \mu \right\}.$$

The 1-reach corresponds simply to the reach, the curvature measure introduced by Federer (1959), involved in our earlier work Henneuse (2024). The  $\mu$ -reach is positive for a large class of sets, indeed the union for all  $\mu > 0$  of compact sets with positive  $\mu$ -reach is dense in the space of compact sets (for the Hausdorff distance). In particular, for our purpose, Assumption **A3** allows considering discontinuity sets displaying corners and multiple points (i.e. self-intersections), cases that were excluded in Henneuse (2024).

We now state a particular property of sets with positive  $\mu$ -reach from Chazal et al. (2007). This property is used in the proof of Theorem 1 to quantify approximation errors. For a set  $K \subset [0, 1]^d$  and  $r \geq 0$  we denote  $B_2(K, r) = \{x \in [0, 1]^d \text{ s.t. } d_K(x) < r\}$  and  $\overline{B}_2(K, r) = \{x \in [0, 1]^d \text{ s.t. } d_K(x) \leq r\}$ .

**Lemma 1.** (Chazal et al., 2007, Lemma 3.1.) Let  $K \subset [0, 1]^d$  a compact set and let  $\mu > 0, r > 0$  be such that  $r < \text{reach}_\mu(K)$ . For any  $x \in \overline{B}_2(K, r) \setminus K$ , one has

$$d_{\partial \overline{B}_2(K, r)}(x) \leq \frac{r - d_K(x)}{\mu} \leq \frac{r}{\mu}.$$

## 2 Procedure description

We here propose a two-steps estimation procedure.

The first step is a "rough" estimation of sublevel sets via local averaging on a regular grid plus a thickening. We call these estimators rough, as they are not meant to perform well for the estimation of sublevel sets evaluated with standard metrics. Typically, over  $S_d(\mu, R_\mu)$ , for a fixed  $\lambda \in \mathbb{R}$ , our estimator  $\widehat{\mathcal{F}}_\lambda$  can be (arbitrarily) far in Hausdorff distance from  $\mathcal{F}_\lambda$ . But, it captures well the persistent topological features of the true signal.

Still, cubical approximation and noise may create cycles lying around the boundaries of sublevel sets estimators that do not correspond to any cycles of the true signal (as in Figure 5.b). These cycles are pretty benign when appearing in regular regions (they will have "short" lifetime), but may have arbitrarily long lifetime when appearing around the discontinuities of the signal. Hence, for the second step, in the same spirit that Bobrowski et al. (2017), instead of considering directly the modules coming from this estimated filtration, we construct an image persistence module. Again, this can be seen as a topological regularization step : it aims to eliminate such cycles (or at least make their lifetime short) without damaging (too much) the information inferred about true cycles. Note that the use of image modules to filter out topological noise is not a new idea, early works by Cohen-Steiner et al. (2009) and Chazal et al. (2011) already use image module to estimate persistence diagrams of signals while working on noisy domains.

In the following, for a set  $A \subset \mathbb{R}^d$  and  $b \geq 0$ , we denote,

$$A^b = \left\{ x \in \mathbb{R}^d \text{ s.t. } d_\infty(x, A) \leq b \right\}$$

with

$$d_\infty(x, A) = \inf_{y \in A} \|x - y\|_\infty.$$

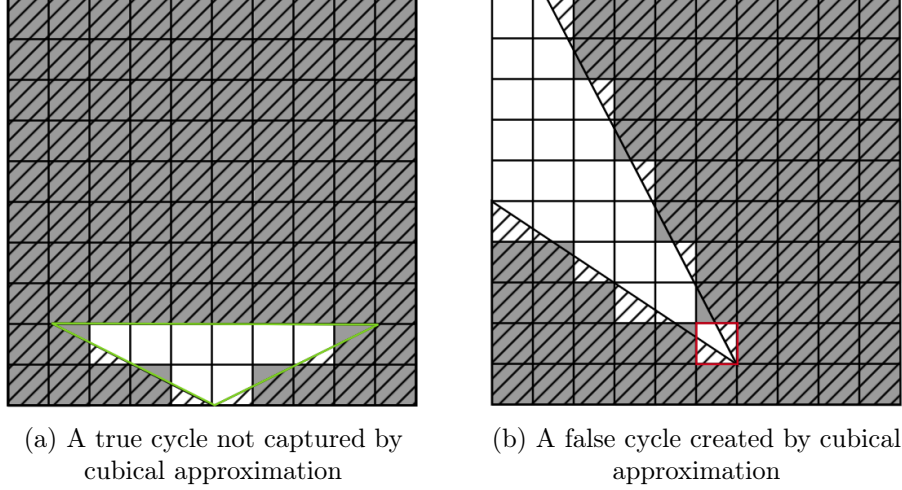


Figure 5:  $\lambda$ -sublevel cubical approximation for  $f$  the function defined as 0 on the hatched area and  $K$  outside (for arbitrarily large  $K$ ). (a) display a case where the histogram approximation fails to capture true cycle in green, for, at least, all  $0 < \lambda < K/2$ . (b) displays a case where the histogram approximation creates a cycle in red, not corresponding to any true cycle, with an arbitrarily long lifetime.

**Step 1 : rough sublevel sets estimation.** Let  $h > 0$  such that  $1/h$  is an integer, consider  $G_h$  the regular orthogonal grid over  $[0, 1]^d$  of step  $h$  and  $C_h$  the collection of all the hypercubes of side  $h$  composing  $G_h$ . Let  $r_1 > 0$ , we define, for all  $\lambda \in \mathbb{R}$ , the "rough"  $\lambda$ -sublevel estimator,

$$\widehat{\mathcal{F}}_\lambda = \left( \bigcup_{H \in C_{h,\lambda}} H \right)^{\lceil r_1 \rceil h} \quad \text{with } C_{h,\lambda} = \left\{ H \in C_h \text{ such that } \int_H dX - \int_H \lambda \leq 0 \right\}$$

and  $\lceil \cdot \rceil$  the ceiling function. This thickening of the sublevel sets obtained by histogram estimation of  $f$  aims to solve the problem exposed in Figure 5.a.

**Step 2 : construction of the image persistence module and associated diagram.** Let,  $r_2 > 0$  and,

$$\rho_\lambda : H_s \left( \widehat{\mathcal{F}}_\lambda \right) \rightarrow H_s \left( \widehat{\mathcal{F}}_\lambda^{\lceil r_2 \rceil h} \right)$$

the map induced by the inclusion  $\widehat{\mathcal{F}}_\lambda \subset \widehat{\mathcal{F}}_\lambda^{\lceil r_2 \rceil h}$ . Then denote,

$$\hat{v}_\lambda^{\lambda'} : H_s \left( \widehat{\mathcal{F}}_\lambda^{\lceil r_2 \rceil h} \right) \rightarrow H_s \left( \widehat{\mathcal{F}}_{\lambda'}^{\lceil r_2 \rceil h} \right)$$

the map induced by the inclusion.  $\widehat{\mathcal{F}}_\lambda^{\lceil r_2 \rceil h} \subset \widehat{\mathcal{F}}_{\lambda'}^{\lceil r_2 \rceil h}$ . Note that the following diagram commutes (all maps are induced by inclusions),

$$\begin{array}{ccc} H_s \left( \widehat{\mathcal{F}}_\lambda \right) & \longrightarrow & H_s \left( \widehat{\mathcal{F}}_{\lambda'} \right) \\ \downarrow \rho_\lambda & & \downarrow \rho_{\lambda'} \\ H_s \left( \widehat{\mathcal{F}}_\lambda^{\lceil r_2 \rceil h} \right) & \xrightarrow{\hat{v}_\lambda^{\lambda'}} & H_s \left( \widehat{\mathcal{F}}_{\lambda'}^{\lceil r_2 \rceil h} \right) \end{array}$$

Thus for all  $\lambda < \lambda'$ ,

$$\widehat{v}_{\lambda,h}^{\lambda'}(\text{Im}(\rho_\lambda)) \subset \text{Im}(\rho_{\lambda'}).$$

We now introduce  $\widehat{\mathbb{V}}_{f,s}$  the persistence module associated to  $(\text{Im}(\rho_\lambda))_{\lambda \in \mathbb{R}}$  equipped with the collection of maps  $(\widehat{v}_{\lambda,h}^{\lambda'})_{\lambda < \lambda'}$  for the  $s$ -th order homology,  $\widehat{\text{dgm}}_s(f)$  the associated persistence diagram, and  $\widehat{\text{dgm}}(f)$  the collection of such persistence diagrams for all  $s \in \mathbb{N}$ . This diagram is well-defined, as we prove in Appendix A that  $\widehat{\mathbb{V}}_{f,s}$  is  $q$ -tame. This second step aims to solve the problem exposed in Figure 5.b.

**Computation.** By construction, for all  $\lambda \in \mathbb{R}$ ,  $\widehat{\mathcal{F}}_\lambda$  and  $\widehat{\mathcal{F}}_\lambda^{\lceil r_2 \rceil h}$  are simply unions of cubes from the regular grid  $G_h$ , thus can be thought as (geometric realization of) cubical complexes (or even simplicial complexes). Hence, the computation of  $\widehat{\text{dgm}}(f)$  is made possible by the algorithm for image persistence gave in Cohen-Steiner et al. (2009).

Another strategy, would be to consider,

$$\widehat{\mathcal{F}}_\lambda = \overline{B}_2 \left( \bigcup_{H \in C_{h,\lambda}} c_H, r_1 h \right)$$

with  $c_H$  the center of the hypercube  $H$  and  $\widehat{\mathbb{V}}_{f,s}$ , the image module induced by,

$$\rho_\lambda : H_s(\widehat{\mathcal{F}}_\lambda) \rightarrow H_s(\overline{B}_2(\widehat{\mathcal{F}}_\lambda, r_2 h)).$$

One can show, adapting the proofs of Section 3, that the subsequent estimator achieve the same convergence rates (up to slight change in constant). As  $\widehat{\mathcal{F}}_\lambda$  and  $\overline{B}_2(\widehat{\mathcal{F}}_\lambda, r_2 h)$  are unions of Euclidean balls, they can be replaced by Čech complexes. To justify properly that it lead to the same estimated persistence diagram, one can adapt the "parameter nerve theorem" (Lemma 3.4) from Chazal and Oudot (2008). Again, the computation of this persistence diagram is then made possible by the algorithm provided in Cohen-Steiner et al. (2009).

### 3 Convergence results

This section is dedicated to the proof of Theorem 1.

#### 3.1 Main proof

The idea of proof follows similarly to the proof for upper bounds in Henneuse (2024). The general idea is to construct an interleaving between the module  $\mathbb{V}_{f,s}$  and  $\widehat{\mathbb{V}}_{f,s}$  and apply the algebraic stability theorem. This implies to construct two morphisms,  $\overline{\psi} : \mathbb{V}_{f,s} \rightarrow \widehat{\mathbb{V}}_{f,s}$  and  $\overline{\phi} : \widehat{\mathbb{V}}_{f,s} \rightarrow \mathbb{V}_{f,s}$  satisfying definition 5. Propositions 1 and 2 give the necessary ingredients.

Let,  $h > 0$  and,

$$\|W\|_h = \sup_{H \in C_h} \frac{|W(H)|}{h^d}$$

with  $W(H) = \int_H dW$ .

**Proposition 1.** Let  $f \in S_d(\mu, R_\mu)$ ,  $h < R_\mu \mu / \sqrt{d}$  and  $r_1 = \sqrt{d}/\mu$ . We have, for all  $\lambda \in \mathbb{R}$ ,

$$\mathcal{F}_{\lambda - \theta \|W\|_h} \subset \widehat{\mathcal{F}}_\lambda \subset \mathcal{F}_{\lambda + \theta \|W\|_h}^{(\sqrt{d} + \lceil \sqrt{d}/\mu \rceil)h}.$$

The proof of Proposition 1 can be found in Section 3.2. The lower inclusion given by this proposition induces directly a morphism from  $\mathbb{V}_{f,s}$  to  $\widehat{\mathbb{V}}_{f,s}$ , this will be our  $\overline{\psi}$ . But the upper inclusion is not sufficient to induce a converse morphism, the construction of  $\overline{\phi}$  requires additional work. To be able to use this upper inclusion, we need a map from the homology groups of  $(\mathcal{F}_\lambda^h)_{\lambda \in \mathbb{R}}$  (for all sufficiently small  $h > 0$ ) into the homology groups of  $(\mathcal{F}_\lambda)_{\lambda \in \mathbb{R}}$ . Proposition 2 provides such a map, exploiting the geometrical properties of the discontinuities set imposed by Assumption A3 to construct a deformation retraction from  $\overline{B}_2(\mathcal{F}_\lambda, h)$  to  $\mathcal{F}_\lambda$ .

**Definition 9.** A subspace  $A$  of  $X$  is called a **deformation retract** of  $X$  if there is a continuous  $F : X \times [0, 1] \rightarrow X$  (called a homotopy) such that for all  $x \in X$  and  $a \in A$ ,

- $F(x, 0) = x$
- $F(x, 1) \in A$
- $F(a, 1) = a$ .

The function  $F$  is called a **deformation retraction** from  $X$  to  $A$ .

Homotopy, and thus homology, is invariant under deformation retract. Thus, a deformation retraction from  $X$  to  $A$  induces isomorphism between homology groups. More precisely, for all  $t \in [0, 1]$ ,  $F(\cdot, t)$  induces a morphism  $F(\cdot, t)^\# : C_s(X) \rightarrow C_s(X)$  between  $s$ -chains of  $X$  defined by composing each singular  $s$ -simplex  $\sigma : \Delta^s \rightarrow X$  with  $F(\cdot, t)$  to get a singular  $s$ -simplex  $F^\#(\sigma, t) = F(\cdot, t) \circ \sigma : \Delta^s \rightarrow X$ , then extending  $F^\#(\cdot, t)$  linearly via  $F^\#(\sum_i n_i \sigma_i, t) = \sum_i n_i F^\#(\sigma_i, t) = \sum_i n_i F(\cdot, t) \circ \sigma_i$ . The morphism  $F^*(\cdot, t) : H_s(X) \rightarrow H_s(X)$  defined by  $[C] \mapsto [F^\#(C, t)]$  can be shown to be an isomorphism for all  $t \in [0, 1]$  (see Hatcher, 2000, pages 110-113). In particular,  $F^*(\cdot, 1) : H_s(X) \rightarrow H_s(A)$  is an isomorphism.

**Proposition 2.** For all  $0 < h < R_\mu$  and all  $\lambda \in \mathbb{R}$  there exists a map  $F_{\lambda,h} : \overline{B}_2(\mathcal{F}_\lambda, h) \times [0, 1] \rightarrow \overline{B}_2(\mathcal{F}_\lambda, h)$  such that,  $F_{\lambda,h}$  is a deformation retraction of  $\overline{B}_2(\mathcal{F}_\lambda, h)$  onto  $\mathcal{F}_\lambda$  and for all  $x \in \overline{B}_2(\mathcal{F}_\lambda, h)$  and all  $t \in [0, 1]$ ,

$$F_{\lambda,h}(x, t) \in \overline{B}_2(x, 2d_{\mathcal{F}_\lambda}(x)/\mu^2).$$

We now have all the necessary ingredients to construct  $\overline{\phi}$  and prove Theorem 1.

**Theorem 1.** Let  $r_1 = \sqrt{d}/\mu$  and  $r_2 = \sqrt{d}(1 + 2/\mu^2)(\sqrt{d} + \lceil r_1 \rceil)$  and  $h = R_\mu/(2\lceil r_2 \rceil)$ , there exist  $\tilde{C}_0$  and  $\tilde{C}_1$  such that, for all  $t > 0$ ,

$$\mathbb{P} \left( \sup_{f \in S_d(\mu, R_\mu)} d_b(\widehat{\text{dgm}}(f), \text{dgm}(f)) \geq \theta t \right) \leq \tilde{C}_0 \exp(-\tilde{C}_1 t^2).$$

*Proof.* Define the morphism,

$$\overline{\psi}_\lambda : H_s(\mathcal{F}_\lambda) \rightarrow H_s(\widehat{\mathcal{F}}_{\lambda + \theta \|W\|_h})$$

induced by the inclusion  $\mathcal{F}_\lambda \subset \widehat{\mathcal{F}}_{\lambda + \theta \|W\|_h}$  given by Proposition 1.

For the second morphism, denote  $k_2 = k_1 + \lceil r_2 \rceil$  and consider,

$$j_{1,\lambda} : \text{Im}(\rho_\lambda) \rightarrow H_s(\overline{B}_2(\mathcal{F}_{\lambda + \theta \|W\|_h}, k_2 h))$$

the map induced by the inclusion  $\widehat{\mathcal{F}}_\lambda^{\lceil r_2 \rceil h} \subset \overline{B}_2(\mathcal{F}_{\lambda+\theta\|W\|_h}, k_2h)$  given by Proposition 1, and,

$$F_{\lambda+\theta\|W\|_h, k_2h}^* : H_s(\overline{B}_2(\mathcal{F}_{\lambda+\theta\|W\|_h}, k_2h)) \rightarrow H_s(\mathcal{F}_{\lambda+\theta\|W\|_h})$$

induced by the deformation retract of Proposition 2. We then define,

$$\begin{cases} \overline{\phi}_\lambda : \text{Im}(\rho_\lambda) \rightarrow H_2(\mathcal{F}_{\lambda+\theta\|W\|_h}) \\ \overline{\phi}_\lambda = F_{\lambda+\theta\|W\|_h, k_2h}^* \circ j_{1,\lambda} \end{cases}$$

We now show that  $\overline{\psi}$  and  $\overline{\phi}$  induce an interleaving between  $\widehat{\mathbb{V}}_{f,s}$  and  $\mathbb{V}_{s,f}$ . More precisely, we show that the following diagrams commute, for all  $\lambda < \lambda'$ ,

$$\begin{array}{ccc} \text{Im}(\rho_\lambda) & \xrightarrow{\hat{v}_\lambda^{\lambda'}} & \text{Im}(\rho_{\lambda'}) \\ \downarrow \overline{\phi}_\lambda & & \downarrow \overline{\phi}_{\lambda'} \\ H_s(\mathcal{F}_{\lambda+\theta\|W\|_h}) & \xrightarrow{v_{\lambda+\theta\|W\|_h}^{\lambda'+\theta\|W\|_h}} & H_s(\mathcal{F}_{\lambda'+\theta\|W\|_h}) \end{array} \quad (1)$$

$$\begin{array}{ccc} H_s(\mathcal{F}_\lambda) & \xrightarrow{v_\lambda^{\lambda'}} & H_s(\mathcal{F}_{\lambda'}) \\ \downarrow \overline{\psi}_\lambda & & \downarrow \overline{\psi}_{\lambda'} \\ \text{Im}(\rho_{\lambda+\theta\|W\|_h}) & \xrightarrow{v_{\lambda+\theta\|W\|_h}^{\lambda'+\theta\|W\|_h}} & \text{Im}(\rho_{\lambda'+\theta\|W\|_h}) \end{array} \quad (2)$$

$$\begin{array}{ccc} \text{Im}(\rho_\lambda) & \xrightarrow{\hat{v}_\lambda^{\lambda+2\theta\|W\|_h}} & \text{Im}(\rho_{\lambda+2\theta\|W\|_h}) \\ \searrow \overline{\phi}_\lambda & & \nearrow \overline{\psi}_{\lambda+\theta\|W\|_h} \\ & H_s(\mathcal{F}_{\lambda+\theta\|W\|_h}) & \end{array} \quad (3)$$

$$\begin{array}{ccc} H_s(\mathcal{F}_\lambda) & \xrightarrow{v_\lambda^{\lambda+2\theta\|W\|_h}} & H_s(\mathcal{F}_{\lambda+2\theta\|W\|_h}) \\ \searrow \overline{\psi}_\lambda & & \nearrow \overline{\phi}_{\lambda+\theta\|W\|_h} \\ & \text{Im}(\rho_{\lambda+\theta\|W\|_h}) & \end{array} \quad (4)$$

- **Diagram 1** : We can rewrite the diagram as (unspecified maps are simply induced by set

inclusion),

$$\begin{array}{ccc}
\text{Im}(\rho_\lambda) & \longrightarrow & \text{Im}(\rho_{\lambda'}) \\
\downarrow & & \downarrow \\
H_s(\overline{B}_2(\mathcal{F}_{\lambda+\theta\|W\|_h}, k_2h)) & \longrightarrow & H_s(\overline{B}_2(\mathcal{F}_{\lambda'+\theta\|W\|_h}, k_2h)) \\
\uparrow F_{\lambda+\theta\|W\|_h, k_2h}^* & & \uparrow F_{\lambda'+\theta\|W\|_h, k_2h}^* \\
H_s(\mathcal{F}_{\lambda+\theta\|W\|_h}) & \longrightarrow & H_s(\mathcal{F}_{\lambda'+\theta\|W\|_h})
\end{array}$$

By inclusions the upper face commutes, and the lower face commutes, as for all  $\lambda \in \mathbb{R}$ ,  $F_{\lambda+\theta\|W\|_h, k_2h}^*$  comes from a deformation retract. Thus, the diagram is commutative.

- **Diagram 2:** One can check that it can be decomposed as diagram 1 and thus the same reasoning applies.
- **Diagram 3:** Let  $[C] \in \text{Im}(\rho_\lambda)$ . By construction, we can suppose  $C \in C_s(\widehat{\mathcal{F}}_\lambda)$ . The map  $\overline{\phi}_\lambda$  maps  $[C]$  to  $[C']$  with,

$$C' = F_{\lambda+\theta\|W\|_h, k_2h}^\#(C, 1).$$

Let denote  $\overline{F}_{\lambda+\theta\|W\|_h, k_2h}$  the restriction of  $F_{\lambda+\theta\|W\|_h, k_2h}$  to  $F_{\lambda+\theta\|W\|_h, k_2h}(\widehat{\mathcal{F}}_\lambda, [0, 1])$ . It is a deformation retraction from  $F_{\lambda+\theta\|W\|_h, k_2h}(\widehat{\mathcal{F}}_\lambda, [0, 1])$  onto  $F_{\lambda+\theta\|W\|_h, k_2h}(\widehat{\mathcal{F}}_\lambda, 1)$ . Thus, by homotopy invariance of singular homology,

$$\begin{aligned}
[C] &= [\overline{F}_{\lambda+\theta\|W\|_h, k_2h}^\#(C, 1)] \\
&= [F_{\lambda+\theta\|W\|_h, k_2h}^\#(C, 1)] \\
&= [C'] \in H_s(F_{\lambda+\theta\|W\|_h, k_2h}(\widehat{\mathcal{F}}_\lambda, [0, 1])).
\end{aligned}$$

From the last assertion of Proposition 2 and the upper inclusion of Proposition 1, we know that,

$$F_{\lambda+\theta\|W\|_h, k_2h}(\widehat{\mathcal{F}}_\lambda, [0, 1]) \subset \widehat{\mathcal{F}}_{\lambda+2\theta\|W\|_h}^{[r_2]h}$$

thus,

$$[C] = [C'] \in H_s[\widehat{\mathcal{F}}_{\lambda+2\theta\|W\|_h}^{[r_2]h}].$$

As  $\overline{\psi}_{\lambda+\theta\|W\|_h}$  is simply an inclusion map, it maps  $[C']$  to  $[C]$ , and hence,

$$[C] = [C'] = \overline{\psi}_{\lambda+\theta\|W\|_h}(\overline{\phi}_\lambda([C])) \in H_s[\widehat{\mathcal{F}}_{\lambda+2\theta\|W\|_h}^{[r_2]h}]$$

which proves the commutativity.

- **Diagram 4:** Let  $C \in C_s(\mathcal{F}_\lambda)$ ,  $\overline{\psi}_\lambda([C]) = [C]$  as it is simply an inclusion induced map and

$$\overline{\phi}_{\lambda+\theta\|W\|_h}([C]) = [F_{\lambda+\theta\|W\|_h, k_2h}^\#(C, 1)] = [C]$$

which establish the commutativity.

The commutativity of diagrams 1,2,3 and 4 means that  $\widehat{\mathbb{V}}_{f,s}$  and  $\mathbb{V}_{f,s}$  are  $2\theta\|W\|_h$  interleaved, and thus we get from the algebraic stability theorem (Chazal et al., 2009) that,

$$d_b\left(\text{dgm}\left(\widehat{\mathbb{V}}_{f,s}\right), \text{dgm}\left(\mathbb{V}_{f,s}\right)\right) \leq 2\theta\|W\|_h$$

and as it holds for all  $s \in \mathbb{N}$  and  $f \in S_d(\mu, R_\mu)$ ,

$$\sup_{f \in S_d(\mu, R_\mu)} d_b\left(\widehat{\text{dgm}}(f), \text{dgm}(f)\right) \leq 2\theta\|W\|_h.$$

We then conclude, using the concentration of  $\|W\|_h$ . As  $W(H)/h^{d/2}$  is a standard Gaussian, we have,

$$\begin{aligned} \mathbb{P}(\|W\|_h \leq t) &= \mathbb{P}\left(\bigcup_{H \in C_h} \{|W(H)/h^d| \leq t\}\right) \\ &\leq |C_h| \mathbb{P}\left(|W(H)/h^{d/2}| \leq th^{d/2}\right) \\ &\leq 2\left(\frac{1}{h}\right)^d \exp\left(-\frac{h^d t^2}{2}\right) \\ &:= C_0 \exp(-C_1 t^2). \end{aligned}$$

With  $C_1$  and  $C_2$  constants depending on  $\mu$ ,  $R_\mu$  and  $d$ . Then,

$$\begin{aligned} &\mathbb{P}\left(\sup_{f \in S_d(\mu, R_\mu)} d_b\left(\widehat{\text{dgm}}(f), \text{dgm}(f)\right) \geq \theta t\right) \\ &\leq \mathbb{P}(\|W\|_h \geq t/2) \\ &\leq C_0 \exp(-C_1 t^2/4) \end{aligned}$$

and the result follows. □

We can derive from this result bounds in expectation.

**Corollary 1.** *Let  $p \geq 1$ ,*

$$\sup_{f \in S_d(\mu, R_\mu)} \mathbb{E}\left(d_b\left(\widehat{\text{dgm}}(f), \text{dgm}(f)\right)^p\right) \lesssim \theta^p$$

*Proof.* The sub-Gaussian concentration provided by Theorem 1, gives that, for all  $t > 0$ ,

$$\mathbb{P}\left(\frac{d_b\left(\widehat{\text{dgm}}(f), \text{dgm}(f)\right)}{\theta} \geq t\right) \leq \tilde{C}_0 \exp\left(-\tilde{C}_1 t^2\right).$$

Now, we have,

$$\begin{aligned} &\mathbb{E}\left(\frac{d_b\left(\widehat{\text{dgm}}(f), \text{dgm}(f)\right)^p}{\theta^p}\right) \\ &= \int_0^{+\infty} \mathbb{P}\left(\frac{d_b\left(\widehat{\text{dgm}}(f), \text{dgm}(f)\right)^p}{\theta^p} \geq t\right) dt \\ &\leq \int_0^{+\infty} \tilde{C}_0 \exp\left(-\tilde{C}_1 t^{2/p}\right) dt < +\infty. \end{aligned}$$

□

### 3.2 Proof of Proposition 1

This section is dedicated to the proof of Proposition 1. It relies on Lemma 1 (stated in Section 1) and on the following lemma.

**Lemma 2.** *Let  $f : [0, 1]^d \rightarrow \mathbb{R}$  and  $h > 0$ . Let  $H \subset \mathcal{F}_{\lambda+\theta\|W\|_h}^c \cap C_h$  and  $H' \subset \mathcal{F}_{\lambda-\theta\|W\|_h} \cap C_h$ . We then have that,*

$$\int_H dX - \int_H \lambda > 0 \text{ and } \int_{H'} dX - \int_{H'} \lambda \leq 0.$$

*Proof.* Let first consider the case where in  $H' \subset \mathcal{F}_{\lambda-\theta\|W\|_h} \cap C_h$ . We have,

$$\begin{aligned} & \int_{H'} dX - \int_{H'} \lambda \\ &= \int_{H'} (f - \lambda) + \theta \int_{H'} dW \\ &\leq -\theta\|W\|_h h^d + \theta\|W\|_h h^d = 0. \end{aligned}$$

Now, let  $H \subset \mathcal{F}_{\lambda+\theta\|W\|_h}^c \cap C_h$ , we have,

$$\begin{aligned} & \int_H dX - \int_H \lambda \\ &= \int_H (f - \lambda) + \theta \int_H dW \\ &> \theta\|W\|_h h^d - \theta\|W\|_h h^d = 0. \end{aligned}$$

□

*Proof of Proposition 1.* We begin by proving the lower inclusion, let  $x \in \mathcal{F}_{\lambda-\theta\|W\|_h}$ . Without loss of generality, let suppose  $x \in M_i$ , or  $x \in \partial M_i$  and  $\liminf_{z \in M_i \rightarrow x} f(z) \leq f(x)$ . If,

$$\overline{B}_2(x, \sqrt{dh}) \subset \left( \bigcup_{i=1}^l \partial M_i \right)^c$$

then,  $H_{x,h}$ , the hypercube of  $C_{h,\lambda}$  containing  $x$  is included in  $\overline{M}_i$ . Assumption **A1** and **A2** then gives  $H_{x,h} \subset \mathcal{F}_{\lambda-\theta\|W\|_h}$ . Hence, it follows from Lemma 2 that  $H_{x,h} \in C_{h,\lambda}$  and consequently  $x \in \widehat{\mathcal{F}}_\lambda$ . Else, as  $\sqrt{dh}/\mu < R_\mu$ , by Lemma 1 (which we can apply thanks to Assumption **A3**), there exists,

$$y \in \left( B_2 \left( \bigcup_{i=1}^l \partial M_i, \sqrt{dh} \right) \right)^c \cap M_i \text{ such that } \|x - y\|_2 \leq \sqrt{dh}/\mu.$$

Let  $H_{y,h}$  the closed hypercube of  $C_h$  containing  $y$ . Hence,  $H_{y,h} \subset \overline{M}_i$ . Then, Assumption **A1** and **A2** ensure that,

$$H_{y,h} \subset \mathcal{F}_{\lambda-\theta\|W\|_h}.$$

Then, Lemma 2 gives  $H_{y,h} \in C_{h,\lambda}$  and thus, as  $x \in H_{y,h}^{\sqrt{dh}/\mu}$ ,  $x \in \widehat{\mathcal{F}}_\lambda$ , which proves the lower inclusion.

For the upper inclusion, let  $x \in \left( \mathcal{F}_{\lambda+\theta\|W\|_h}^c \right)^c$ , and  $H_{x,h}$  the hypercube of  $C_h$  containing  $x$ . We



then have,  $H_{x,h} \subset \mathcal{F}_{\lambda+\theta\|W\|_h}^c$ . Hence, Lemma 2 gives that,

$$H_{x,h} \subset \left( \bigcup_{H \in C_{h,\lambda}} H \right)^c$$

and thus,

$$\bigcup_{H \in C_{h,\lambda}} H \subset \mathcal{F}_{\lambda+\theta\|W\|_h}^{\sqrt{dh}}.$$

Consequently,

$$\left( \bigcup_{H \in C_{h,\lambda}} H \right)^{\lceil \sqrt{d}/\mu \rceil h} = \widehat{\mathcal{F}}_\lambda \subset \mathcal{F}_{\lambda+\theta\|W\|_h}^{(\sqrt{d} + \lceil \sqrt{d}/\mu \rceil)h}$$

and the proof is complete.  $\square$

### 3.3 Proof of Proposition 2

Proposition 2 is a direct corollary of Theorem 12 from Kim et al. (2020).

**Lemma 3.** (Kim et al., 2020, Theorem 12) Let  $K \subset [0, 1]^d$ , for all  $0 \leq r < \text{reach}_\mu(K)$ ,  $\overline{B}_2(K, r)$  retracts by deformation onto  $K$  and the associated deformation retraction  $F : \overline{B}_2(K, r) \times [0, 1] \rightarrow \overline{B}_2(K, r)$  verifies for all  $x \in \overline{B}_2(K, r)$  and all  $t \in [0, 1]$ ,  $F(x, t) \in \overline{B}_2(x, 2d_K(x)/\mu^2)$ .

The first part of the claim is Theorem 12 from Kim et al. (2020) and the second part follows easily from their construction. Elements of proof can be found in Appendix B.

*Proof of Lemma 2.* Let  $\lambda \in \mathbb{R}$  and denote  $I_\lambda = \{i \in \{1, \dots, l\} \text{ s.t. } M_i \cap \mathcal{F}_\lambda \neq \emptyset\}$ . By Assumption A1 and A2, we have  $\mathcal{F}_\lambda = \bigcup_{i \in I_\lambda} \overline{M}_i$ . Thus, by Assumption A3,

$$\text{reach}_\mu(\mathcal{F}_\lambda) \geq R_\mu.$$

Hence, by Lemma 3, for all  $h < R_\mu$ , there exists a deformation retraction  $F_{\lambda,h} : \overline{B}_2(\mathcal{F}_\lambda, h) \times [0, 1] \rightarrow \overline{B}_2(\mathcal{F}_\lambda, h)$  from  $\overline{B}_2(\mathcal{F}_\lambda, h)$  onto  $\mathcal{F}_\lambda$  that furthermore verifies, for all  $(x, t) \in \overline{B}_2(\mathcal{F}_\lambda, h) \times [0, 1]$ ,  $F_{\lambda,h}(x, t) \in \overline{B}_2(x, 2d_{\mathcal{F}_\lambda}(x)/\mu^2)$ .  $\square$

## 4 Discussion

This work extends the current scope on persistent homology inference from sublevel sets of a noisy signal. We motivate the need to depart from plug-in approaches, highlighting that cubical approximation fails to capture properly the persistent homology of piecewise-constant signals with discontinuities set having a positive  $\mu$ -reach. To overcome this issue, we propose a method based on image persistence. In the context of the Gaussian white noise model, we show that over the classes  $S_d(\mu, R_\mu)$  this approach is consistent and achieves parametric rates.

This study complements the results obtained in Henneuse (2024). Although we impose stronger conditions on the signals in the regular region, we significantly relax the conditions on the discontinuities set, allowing us to consider signals with more complex discontinuities, such as those with multiple points or arbitrarily narrow corners. This, once again, underscores the robustness of persistence diagram inference to signal irregularities and highlights the advantages of moving beyond

analyses that rely on sup-norm stability.

As demonstrated in Appendix C of [Henneuse \(2024\)](#), the method and results presented here can be easily extended to the setting of non-parametric regression. Motivated by applications to mode detection, we are also exploring an extension to the density model, which will be the focus of future work.

A practical limitation of the proposed method is that it requires some knowledge of the parameters  $\mu$  and  $R_\mu$ . This dependence is rather common in geometric inference and TDA. Developing adaptive approaches to handle such dependence on geometric measures is a research topic in its own right.

Also, from a computational perspective (to perform a numerical evaluation of this method), an important step forward would be to provide an efficient algorithm to compute the image persistence for general filtrations. This would permit to compare this method to plug-in approaches, specifically to the histogram estimator. While [Bauer and Schmahl \(2022\)](#) offers an efficient such algorithm, it is tailored for Rips complexes, complicating its direct application to our context. In particular, this would permit to investigate practical selection of the parameters  $h$ ,  $r_1$  and  $r_2$ .

## Acknowledgements

The author would like to thank Frédéric Chazal and Pascal Massart for our (many) helpful discussions. The author acknowledge the support of the ANR TopAI chair (ANR-19-CHIA-0001).

## References

- Serguei Barannikov. The framed morse complex and its invariants. *Adv. Soviet Math.*, 21:93–115, 1994.
- Ulrich Bauer and Maximilian Schmahl. Efficient computation of image persistence, 2022.
- Omer Bobrowski, Sayan Mukherjee, and Jonathan E. Taylor. Topological consistency via kernel estimation. *Bernoulli*, 23(1):288 – 328, 2017. doi: 10.3150/15-BEJ744.
- Peter Bubenik and Peter Kim. A statistical approach to persistent homology. *Homology, Homotopy and Applications*, 9, 2006.
- Peter Bubenik, Gunnar Carlsson, Peter Kim, and Zhiming Luo. Statistical topology via morse theory persistence and nonparametric estimation. *Contemporary Mathematics*, 516:75–92, 2009. doi: 10.1090/conm/516/10167.
- Frédéric Chazal and Steve Oudot. Towards persistence-based reconstruction in euclidean spaces. In *Proceedings of the Twenty-Fourth Annual Symposium on Computational Geometry*, SCG '08, page 232–241, New York, NY, USA, 2008. Association for Computing Machinery.
- Frédéric Chazal, David Cohen-Steiner, Marc Glisse, Leonidas J. Guibas, and Steve Y. Oudot. Proximity of persistence modules and their diagrams. In *Proceedings of the Twenty-Fifth Annual Symposium on Computational Geometry*, page 237–246. Association for Computing Machinery, 2009. doi: 10.1145/1542362.1542407.

- Frédéric Chazal, Steve Oudot, Marc Glisse, and Vin de Silva. *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics. Springer Verlag, 2016. doi: 10.1007/978-3-319-42545-0.
- Frédéric Chazal and Bertrand Michel. An introduction to topological data analysis: fundamental and practical aspects for data scientists, 2021.
- Frédéric Chazal, David Cohen-Steiner, and André Lieutier. A sampling theory for compact sets in euclidean space. *Discrete and Computational Geometry*, 41:461–479, 06 2006. doi: 10.1007/s00454-009-9144-8.
- Frédéric Chazal, David Cohen-Steiner, André Lieutier, and Boris Thibert. Shape smoothing using double offset. *Proceedings - SPM 2007: ACM Symposium on Solid and Physical Modeling*, 06 2007. doi: 10.1145/1236246.1236273.
- Frédéric Chazal, Leonidas Guibas, Steve Oudot, and Primoz Skraba. Scalar field analysis over point cloud data. *Discrete & Computational Geometry*, 46:743–775, 12 2011. doi: 10.1007/s00454-011-9360-x.
- Frédéric Chazal, Marc Glisse, Catherine Labruère Chazal, and Bertrand Michel. Convergence rates for persistence diagram estimation in topological data analysis. *31st International Conference on Machine Learning, ICML 2014*, 1, 2014.
- David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. *Discrete and Computational Geometry - DCG*, 37:263–271, 2005. doi: 10.1007/s00454-006-1276-5.
- David Cohen-Steiner, Herbert Edelsbrunner, John Harer, and Dmitriy Morozov. Persistent homology for kernels, images, and cokernels. *Proceedings of the 2009 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1011–1020, 2009. doi: 10.1137/1.9781611973068.110.
- William Crawley-Boevey. Decomposition of pointwise finite-dimensional persistence modules, 2012.
- Brittany Fasy, Fabrizio Lecci, Alessandro Rinaldo, Larry Wasserman, Sivaraman Balakrishnan, and Aarti Singh. Confidence sets for persistence diagrams. *The Annals of Statistics*, 42:2301–2339, 2014. doi: 10.1214/14-AOS1252.
- Herbert Federer. Curvature measures. *Trans. Amer. Math. Soc*, 1959.
- Allen Hatcher. *Algebraic topology*. Cambridge Univ. Press, Cambridge, 2000.
- Hugo Henneuse. Persistence diagram estimation of multivariate piecewise hölder-continuous signals. 2024.
- Jisu Kim, Jaehyeok Shin, Frédéric Chazal, Alessandro Rinaldo, and Larry Wasserman. Homotopy reconstruction via the cech complex and the vietoris-rips complex, 2020.

## A Proof for $q$ -tameness

This section is devoted to prove the claim that the persistence diagrams we consider and estimated persistence diagrams we propose are well-defined, by proving that the underlying persistence modules are  $q$ -tame.

**Proposition 3.** *Let  $f \in S_d(\mu, R_\mu)$  then  $f$  is  $q$ -tame.*

*Proof.* Let  $s \in \mathbb{N}$  and  $\mathbb{V}_{s,f}$  the persistence module (for the  $s$ -th homology) associated to the sublevel filtration,  $\mathcal{F}$  and for fixed levels  $\lambda < \lambda'$  let denote  $v_\lambda^{\lambda'}$  the associated map. Let  $h < R_\mu$  and denote  $i_{\lambda,h} : H_s(\mathcal{F}_\lambda) \rightarrow H_s(\overline{B}_2(\mathcal{F}_\lambda, h))$ . By Lemma 2, we have,

$$v_\lambda^{\lambda'} = v_\lambda^{\lambda'} \circ F_{\lambda,h}^* \circ i_{\lambda,h}.$$

By assumption **A1** and **A2**,  $\mathcal{F}_\lambda$  is compact. As  $[0, 1]^d$  is triangulable,  $\mathcal{F}_\lambda$  is covered by finitely many cells of the triangulation, and so there is a finite simplicial complex  $K$  such that  $\overline{\mathcal{F}_\lambda} \subset K \subset \mathcal{F}_\lambda^h$ . Consequently,  $i_{\lambda,h}$  factors through the finite dimensional space  $H_s(K)$  and is then of finite rank by Theorem 1.1 of [Crawley-Boevey \(2012\)](#). Thus,  $v_\lambda^{\lambda'}$  is of finite rank. Hence,  $f$  is  $q$ -tame.  $\square$

**Proposition 4.** *Let  $f : [0, 1]^d \rightarrow \mathbb{R}$  and  $h > 0$  then, for all  $s \in \mathbb{N}$ ,  $\widehat{\mathbb{V}}_{s,f}$  is  $q$ -tame.*

*Proof.* The results also follows from by Theorem 1.1 of [Crawley-Boevey \(2012\)](#). Let  $h > 0$  and  $\lambda \in \mathbb{R}$ .  $\widehat{\mathcal{F}}_\lambda$  and  $\widehat{\mathcal{F}}_\lambda^{\lceil r_2 \rceil h}$  are unions of cubes of  $C_h$  and thus finite dimensional. Hence,  $\text{Im}(\rho_\lambda)$  is finite dimensional. Thus  $\widehat{\mathbb{V}}_{s,f}$  is  $q$ -tame.  $\square$

## B Element of proof for Lemma 3

This section is dedicated to the proof of Lemma 3.

*Proof of Lemma 3.* The first part of the claim is Theorem 12 of [Kim et al. \(2020\)](#). A standard technique to construct deformation retraction in differential topology is to exploit the flow coming from a smooth underlying vector field. In [Kim et al. \(2020\)](#), their idea is to use the vector field defined on  $\overline{B}_2(K, r) \setminus K$ , by  $W(x) = -\nabla_K(x)$ . But this vector field is not continuous. To overcome this issue, they construct a locally finite covering  $(U_{x_i})_{i \in \mathbb{N}}$  of  $\overline{B}_2(K, r) \setminus K$  and an associated partition of the unity  $(\rho_i)_{i \in \mathbb{N}}$ , such that  $\overline{W}(x) = \sum_{i \in \mathbb{N}} \rho_i(x) w(x_i)$  shares the same dynamic as  $W$ . More precisely, they show that  $\overline{W}$  induces a smooth flow  $C$  that can be extended on  $\overline{B}_2(K, r) \times [0, +\infty[$  such that for all  $x \in \overline{B}_2(K, r)$ , for all  $t \geq 2d_K(x)/\mu^2$ ,  $C(x, t) = C(x, 2d_K(x)/\mu^2) \in K$ . We make an additional remark, denotes  $d_C$  the arc length distance along  $C$ , as  $\|W\| \leq 1$ , we have,

$$\begin{aligned} d_C(x, C(x, 2r/\mu^2)) &= \int_0^{2r/\mu^2} \left| \frac{\partial}{\partial t} C(x, t) \right| dt \\ &\leq \int_0^{2r/\mu^2} \|W(C(x, t))\|_2 dt \\ &\leq 2d_K(x)/\mu^2 \end{aligned} \tag{5}$$

Thus for all  $t \in [0, +\infty[$ ,  $\|x - C(x, t)\|_2 \leq 2r/\mu^2$ . Now taking,  $F(x, s) = C(x, 2rs/\mu^2)$ , provide a deformation retract of  $\overline{B}_2(K, r)$  onto  $K$  and the associated retraction  $R : x \mapsto F(x, 1)$  verifies,  $R(x) \in \overline{B}_2(x, 2d_K(x)/\mu^2)$  by (5).  $\square$