

Stable shape comparison by persistent homology

Massimo Ferri, Patrizio Frosini

Dipartimento di Matematica and ARCES
Università di Bologna, Italy

`massimo.ferri@unibo.it`, `patrizio.frosini@unibo.it`

Claudia Landi

Dipartimento di Scienze e Metodi dell'Ingegneria
Università di Modena e Reggio Emilia, Italy and

ARCES, Università di Bologna, Italy
`clandi@unimore.it`

December 15, 2011

Abstract

When shapes of objects are modeled as topological spaces endowed with functions, the shape comparison problem can be dealt with using persistent homology to provide shape descriptors, and the matching distance to measure dissimilarities. Motivated by the problem of dealing with incomplete or imprecise acquisition of data in computer vision and computer graphics, recent papers have studied stability properties of persistent Betti numbers with respect to perturbations both in the topological space and in the function. This paper reports on progress in this area of research.

2010 MSC: Primary: 55N35, Secondary: 68T10, 68U05, 55N05

Keywords: Persistent Betti numbers, persistence diagram, matching distance

1 Introduction

The problem of comparing shapes is well-studied in computer vision and computer graphics and many algorithms have been developed for this purpose. The aim of this paper is to present the methodological approach to shape comparison by persistent homology developed by the authors and their collaborators in recent years, shedding light on the unifying ideas underlying different papers while skipping technical details. This approach fits in the general scheme of associating a shape with a shape descriptor, or a signature, and comparing shapes by measuring dissimilarity between descriptors.

The starting observation is that, although there is no universally accepted definition for the notion of shape of an object, and most of the proposed techniques are

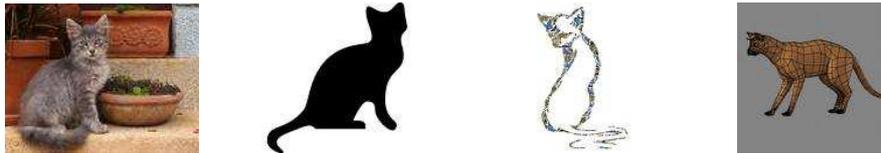


Figure 1: Different observations of a cat.

tailored for some particular interesting case (e.g., polyhedral rigid objects, planar curves, point cloud data, triangular meshes), tentative definitions are commonly based on observer's perceptions.

The dependence on observers implies large subjectivity. Depending on what the observer is observing, observations can be modeled as closed curves (outline), plane domains (silhouettes), triangular meshes (surfaces) and in many other ways (see Figure 1).

In this framework, observer's perceptions can be modeled as a function $\vec{\varphi} : X \rightarrow \mathbb{R}^n$. The function depends on the shape property the observer is perceiving: curvature, roundness, elongation, connectivity etc. For each observation $x \in X$, $\vec{\varphi}$ describes x as seen by the observer. Due to changes, e.g., in point of view, or distance from the object, or light conditions, also perceptions are subject to large subjectivity (see Figure 2).

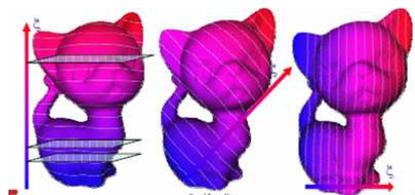


Figure 2: Different perceptions of the same space, depending on the view-point.

The second fundamental remark is that human judgments rely on persistent perceptions. Non-persistent properties can be considered as due to noise, whereas persistent properties concur to give a shape to objects. Therefore, a good shape descriptor for $(X, \vec{\varphi})$ should behave well also in the presence of small changes in the perceptions and in the observations. This yields a request for stability with respect to perturbations of the function $\vec{\varphi}$ and with respect to perturbations of the space X . The formalization of the first request motivates the assumption that X is a topological space (not only a set), and $\vec{\varphi}$ is a continuous function. The formalization of the second request calls for the availability of a metric between sets. This happens, e.g., if X is a compact subset of a metric space, so that we can use the Hausdorff metric to quantify perturbations of X .

As a result, we are led to study pairs $(X, \vec{\varphi})$ where X is a topological space and $\vec{\varphi} : X \rightarrow \mathbb{R}^n$ is a continuous function, usually called a *measuring* or *filtering*

function. These pairs are known in literature as *size pairs* (cf., e.g., [26, 31]). In Section 4 we shall also assume that X is a compact subset of \mathbb{R}^m .

The comparison between two size pairs $(X, \vec{\varphi})$, $(Y, \vec{\psi})$ can be done, in principle, by the *natural pseudodistance* (cf. [31, 18, 19, 20]). This pseudodistance equals the infimum of the value $\max_i \max_{x \in X} |\varphi_i(x) - \psi_i(h(x))|$, assuming that X and Y are homeomorphic and that h varies in the set of all homeomorphisms between these spaces. However, the computation of the natural pseudodistance is quite difficult to do in practice, due to high computational complexity.

Fortunately, lower bounds for this pseudodistance can be obtained by multi-dimensional persistent homology (cf. [12]), that is much easier to compute. This theory is often used for studying objects related to computer vision and computer graphics, and involves analyzing the qualitative and quantitative behavior of vector-valued functions $\vec{\varphi}$ defined over topological spaces X [4, 21]. This is achieved by considering the filtration obtained from the sequence of nested lower level sets of the function under study, and by encoding the scale at which a topological feature (e.g., a connected component, a tunnel, a void) is created, and when it is annihilated along this filtration. In this framework, multidimensional persistent homology groups capture the homology of a multi-parameter increasing family of spaces. For application purposes, these groups are further encoded by simply considering their rank, which yields a parameterized version of Betti numbers, called persistent Betti numbers [22] or rank invariants [8]. Varying the lower level sets, we get that persistent Betti numbers can be seen as functions taking pairs of vectors to the set of non-negative integers.

The use of persistent Betti numbers functions as shape descriptors when φ is scalar-valued (i.e. one-dimensional persistent Betti numbers) dates back to the beginning of the 1990s (see, e.g., [25, 34]) and has found a number of applications (see, e.g., [5, 10, 13, 16, 24, 32]).

The recent results obtained in [11, 12, 29] by the authors of this paper jointly with their collaborators when $\vec{\varphi}$ is vector-valued show that persistent Betti numbers are stable shape descriptors, behaving well both with respect to perturbations of the functions and with respect to perturbations of the space. This justifies the use of multidimensional persistent homology for shape comparison [1, 3, 14].

In this paper these recent results are surveyed. In Section 2 the definitions of persistent homology group and persistent Betti numbers function (briefly PBNs) are recalled. In Section 3 we illustrate the main results from [12], that are stability results with respect to functions perturbations. In particular, in Section 3.1, we report the definition of a distance, called *matching distance*, between one-dimensional PBNs useful to compare PBNs. Then the stability theorem for one-dimensional PBNs with respect to the matching distance is given. The novelty is that it does not require the tameness assumptions on the functions that were usually requested. Moreover, in Section 3.2, stability with respect to function perturbations is achieved also in the multidimensional case, by means of a so-called *foliation method* and an appropriate generalization of the matching distance. Section 4 shows how to use the multidimensional matching distance in order to cope also with changes in the studied topological space. In particular, Section 4.1 sur-

veys the results of [11], showing that the PBNs of a point sample of X are sufficient to recover the PBNs of X up to a controlled uncertainty. In section 4.2, the stability of PBNs with respect to domain perturbations measured by the Hausdorff distance as dealt with in [29] is reviewed.

We conclude this section by observing that the result about the stability of PBNs with respect to perturbation of multidimensional filtering functions has allowed a recent advance about the reconstruction of a size pair $(X, \vec{\varphi})$ up to vanishing natural pseudodistance, when the space X is a curve. We refer the interested reader to [30] for more details.

2 PBNs: Definitions and first properties

In this paper, the following relations \preceq and \prec are defined in \mathbb{R}^n : for $\vec{u} = (u_1, \dots, u_n)$ and $\vec{v} = (v_1, \dots, v_n)$, we say $\vec{u} \preceq \vec{v}$ (resp. $\vec{u} \prec \vec{v}$) if and only if $u_i \leq v_i$ (resp. $u_i < v_i$) for every index $i = 1, \dots, n$. Moreover, \mathbb{R}^n is endowed with the usual max-norm: $\|(u_1, u_2, \dots, u_n)\|_\infty = \max_{1 \leq i \leq n} |u_i|$.

We shall use the following notations: Δ^+ will be the open set $\{(\vec{u}, \vec{v}) \in \mathbb{R}^n \times \mathbb{R}^n : \vec{u} \prec \vec{v}\}$. For every n -tuple $\vec{u} = (u_1, \dots, u_n) \in \mathbb{R}^n$ and for every function $\vec{\varphi} : X \rightarrow \mathbb{R}^n$, we shall denote by $X\langle \vec{\varphi} \preceq \vec{u} \rangle$ the set $\{x \in X : \varphi_i(x) \leq u_i, i = 1, \dots, n\}$.

The definition below extends the concept of the persistent homology group to a multidimensional setting.

Definition 2.1 *Let $k \in \mathbb{Z}$. Let X be a topological space, and $\vec{\varphi} : X \rightarrow \mathbb{R}^n$ a continuous function. Let $\pi_k^{(\vec{u}, \vec{v})} : H_k(X\langle \vec{\varphi} \preceq \vec{u} \rangle) \rightarrow H_k(X\langle \vec{\varphi} \preceq \vec{v} \rangle)$ be the homomorphism induced in homology by the inclusion map $\pi^{(\vec{u}, \vec{v})} : X\langle \vec{\varphi} \preceq \vec{u} \rangle \hookrightarrow X\langle \vec{\varphi} \preceq \vec{v} \rangle$ with $\vec{u} \preceq \vec{v}$. If $\vec{u} \prec \vec{v}$, the image of $\pi_k^{(\vec{u}, \vec{v})}$ is called the multidimensional k th persistent homology group of $(X, \vec{\varphi})$ at (\vec{u}, \vec{v}) , and is denoted by $H_k^{(\vec{u}, \vec{v})}(X, \vec{\varphi})$.*

In other words, the group $H_k^{(\vec{u}, \vec{v})}(X, \vec{\varphi})$ contains all and only the homology classes of cycles born before or at \vec{u} and still alive at \vec{v} .

In what follows, we shall work with coefficients in a field \mathbb{K} , so that homology groups are vector spaces. Therefore, they can be completely described by their dimension, leading to the following definition (cf. [9, 22]).

Definition 2.2 *The function $\beta_{\vec{\varphi}} : \Delta^+ \rightarrow \mathbb{N} \cup \{\infty\}$ defined by*

$$\beta_{\vec{\varphi}}(\vec{u}, \vec{v}) = \dim \operatorname{im} \pi_k^{(\vec{u}, \vec{v})} = \dim \check{H}_k^{(\vec{u}, \vec{v})}(X, \vec{\varphi})$$

will be called the persistent Betti numbers function of $\vec{\varphi}$, briefly PBNs.

Obviously, for each $k \in \mathbb{Z}$, we have different PBNs $\beta_{\vec{\varphi}}$ of $\vec{\varphi}$ (which should be denoted $\beta_{\vec{\varphi}, k}$, say) but, for the sake of notational simplicity, we omit adding any reference to k . This will also apply to the notations used for other concepts in this paper, such as multiplicities and persistence diagrams.

The following two questions have recently been considered: Under which conditions do PBNs take only finite values? Which homology theory is better suited for the comparison of PBNs?

As for the first question, it has been proven in [7] that, if X is a compact and locally contractible space embeddable in some \mathbb{R}^m , then $\beta_{\tilde{\varphi}}$ never attains the value ∞ . For the sake of simplicity, in the rest of this paper we will assume X to be triangulable, i.e. the underlying space of a finite simplicial complex (up to a homeomorphism).

The second question has been thoroughly considered in [12], and has led us to the choice of working with Čech homology. The reason is that, having the continuity axiom, it allows us to completely represent one-dimensional PBNs by persistence diagrams. Even assuming tameness, this result would not hold for singular and simplicial theories, which guarantee a complete description of one-dimensional PBNs only outside a set of vanishing measure. For details about Čech homology, the reader can refer to [23, Ch. IX].

3 Stability with respect to functions perturbations

In this section we show that the persistent Betti numbers of nearby scalar or vector-valued filtering functions are close to each other in the sense expressed by a suitable matching distance. The proofs of the results presented in this section can be found in [12].

Our new stability results are not limited by the restrictions of tameness and max-tameness assumptions used in [15] to prove stability for scalar functions and in [6] to prove stability for vector-valued functions, respectively.

In what follows, we will refer to the case of scalar filtering functions as to the one-dimensional case, whereas the term multi-dimensional will refer to the case of vector-valued filtering functions.

3.1 Stability of one-dimensional PBNs

In this section we give a theorem stating the stability of PBNs for continuous scalar-valued filtering functions (Theorem 3.6). This result generalizes the main theorem in [15], which requires tame functions on triangulable spaces. The proof can be found in [12] and relies on a number of basic simple properties of PBNs that are completely analogous to those proved in [17, 27] and used to show the PBNs stability in the case of the 0th homology. We first recall the main ingredients.

Since now we confine ourselves to the case $n = 1$, for the sake of simplicity, the symbols $\tilde{\varphi}$, \tilde{u} , \tilde{v} will be replaced by φ , u , v , respectively. We remark that Δ^+ reduces to be the set $\{(u, v) \in \mathbb{R}^2 : u < v\}$. Moreover, we use the following notations: $\Delta = \partial\Delta^+$, $\Delta^* = \Delta^+ \cup \{(u, \infty) : u \in \mathbb{R}\}$, and $\tilde{\Delta}^* = \Delta^* \cup \Delta$.

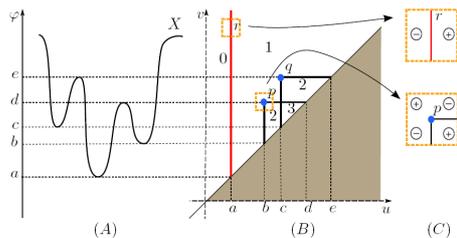


Figure 3: (A) A curve $X \subseteq \mathbb{R}^2$ represented by a continuous line, and the function $\varphi : X \rightarrow \mathbb{R}$ such that $\varphi(P) = y$ for every $P = (x, y) \in X$. (B) The points (proper and at infinity) of the persistence diagram D_φ . (C) Computation of multiplicities seen through lens.

Definition 3.1 For every point $p = (u, v) \in \Delta^+$, we define the number $\mu(p)$ as the minimum over all the positive real numbers ε , with $u + \varepsilon < v - \varepsilon$, of

$$\beta_\varphi(u + \varepsilon, v - \varepsilon) - \beta_\varphi(u - \varepsilon, v - \varepsilon) - \beta_\varphi(u + \varepsilon, v + \varepsilon) + \beta_\varphi(u - \varepsilon, v + \varepsilon).$$

The number $\mu(p)$ will be called the multiplicity of p for β_φ . Moreover, we shall call a proper cornerpoint for β_φ any point $p \in \Delta^+$ such that the number $\mu(p)$ is strictly positive.

Definition 3.2 For every vertical line r , with equation $u = \bar{u}$, $\bar{u} \in \mathbb{R}$, let us identify r with $(\bar{u}, \infty) \in \Delta^*$, and define the number $\mu(r)$ as the minimum over all the positive real numbers ε , with $\bar{u} + \varepsilon < 1/\varepsilon$, of

$$\beta_\varphi\left(\bar{u} + \varepsilon, \frac{1}{\varepsilon}\right) - \beta_\varphi\left(\bar{u} - \varepsilon, \frac{1}{\varepsilon}\right).$$

The number $\mu(r)$ will be called the multiplicity of r for β_φ . When this finite number is strictly positive, we call r a cornerpoint at infinity for β_φ .

The concept of cornerpoint allows us to introduce a representation of the PBNs, based on the following definition [15].

Definition 3.3 The persistence diagram D_φ is the multiset of all cornerpoints (both proper and at infinity) for β_φ , counted with their multiplicity, union the points of Δ , counted with infinite multiplicity.

An example of persistence diagram in zeroth homology degree is displayed in Figure 3. We recall that, in the case $k = 0$, $\beta_\varphi(u, v)$ counts the number of connected components born before or at the level u and still alive at level v . In this example we consider a curve X of \mathbb{R}^2 represented by the solid line in Figure 3(A), and the function $\varphi : X \rightarrow \mathbb{R}$ that associates with each point $P \in X$ its ordinate in the plane. The sole points (both proper and at infinity) of the associated persistence diagram D_φ are p , q , and r , and are shown in Figure 3(B). Here, solid lines

divide Δ^* into regions where the value taken by the zeroth PBNs of (X, φ) is constant. This value is displayed in each region. For instance, when $c \leq v < d$, $X\langle\varphi \leq v\rangle$ has three connected components. Only one of them contains at least one point of $X\langle\varphi \leq u\rangle$, when $a \leq u < b$; two of them contain at least one point of $X\langle\varphi \leq u\rangle$, when $b \leq u < c$; all of them contain at least one point of $X\langle\varphi \leq u\rangle$, when $c \leq u < v < d$. Therefore, when $c \leq v < d$, $\beta_\varphi(u, v) = 1$ for $a \leq u < b$; $\beta_\varphi(u, v) = 2$ for $b \leq u < c$; $\beta_\varphi(u, v) = 3$ for $c \leq u < v$.

Figure 3(C) zooms in on two points of the persistence diagram (one proper, p , and one at infinity, r) to explain how their multiplicity is computed. The alternating sum of the zeroth persistent Betti numbers at four points around p is $2 - 1 - 1 + 1$, giving $\mu(p) = 1$. The alternating sum of the zeroth persistent Betti numbers at two points next to r is $1 - 0$, giving $\mu(r) = 1$.

The PBNs of a scalar-valued filtering function can be completely described by a persistence diagram, as the following theorem states.

Theorem 3.4 *For every $(\bar{u}, \bar{v}) \in \Delta^+$, we have*

$$\beta_\varphi(\bar{u}, \bar{v}) = \sum_{\substack{(u,v) \in \Delta^* \\ u \leq \bar{u}, v > \bar{v}}} \mu((u, v)).$$

As an immediate consequence of Theorem 3.4, it follows that any distance between persistence diagrams induces a distance between one-dimensional PBNs. This justifies the introduction of the matching distance, recalled in the following definition.

Definition 3.5 *Let X, Y be triangulable spaces endowed with continuous functions $\varphi : X \rightarrow \mathbb{R}$, $\psi : Y \rightarrow \mathbb{R}$. The (extended) matching distance d_{match} between β_φ and β_ψ is defined by*

$$d_{match}(\beta_\varphi, \beta_\psi) = \inf_{\gamma} \sup_{p \in D_\varphi} \|p - \gamma(p)\|_\infty, \quad (1)$$

where γ ranges over all multi-bijections between D_φ and D_ψ , and, for every $p = (u, v), q = (u', v')$ in $\bar{\Delta}^*$,

$$\|p - q\|_\infty = \min \left\{ \max\{|u - u'|, |v - v'|\}, \max\left\{\frac{v - u}{2}, \frac{v' - u'}{2}\right\} \right\},$$

with the convention about points at infinity that $\infty - y = y - \infty = \infty$ when $y \neq \infty$, $\infty - \infty = 0$, $\frac{\infty}{2} = \infty$, $|\infty| = \infty$, $\min\{c, \infty\} = c$ and $\max\{c, \infty\} = \infty$.

In plain words, $\|\cdot\|_\infty$ measures the pseudo-distance between two points p and q as the minimum between the cost of moving one point onto the other and the cost of moving both points onto the diagonal, with respect to the max-norm and under the assumption that any two points of the diagonal have vanishing pseudo-distance.

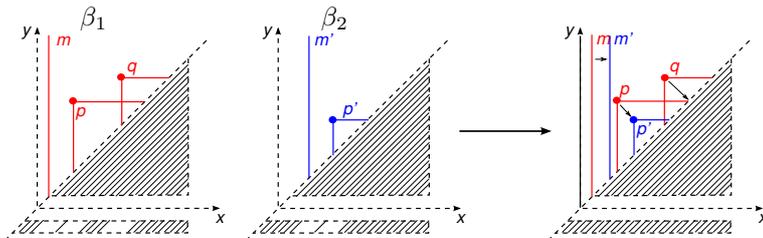


Figure 4: A matching between persistence diagrams used to compute the matching distance.

The term *extended* means that d_{match} can take the value $+\infty$. It will follow from our next Theorem 3.6 that d_{match} is finite when $X = Y$.

When the number of cornerpoints is finite, the matching of persistence diagrams is related to the bottleneck transportation problem, and the matching distance reduces to the bottleneck distance [15]. In our case, however, the number of cornerpoints may be countably infinite, because of our loose assumption on the filtering function, that is only required to be continuous.

Theorem 3.6 *Let X be a triangulable space, and $\varphi, \psi : X \rightarrow \mathbb{R}$ two continuous functions. Then $d_{match}(\beta_\varphi, \beta_\psi) \leq \max_{x \in X} |\varphi(x) - \psi(x)|$.*

For the 0th homology, the claim has been proved in [17, Thm. 25].

3.2 Stability of multidimensional PBNs

We now consider the stability of multidimensional PBNs, i.e. for vector-valued functions. It can be deduced following the same arguments given in [2] to prove the stability of multidimensional PBNs for the case of the 0th homology.

The key idea is that a foliation in half-planes of Δ^+ can be given, such that the restriction of the multidimensional PBNs function to these half-planes turns out to be a one-dimensional PBNs function in two scalar variables. This approach implies that the comparison of two multidimensional PBNs functions can be performed leaf by leaf by measuring the distance of appropriate one-dimensional PBNs functions. Therefore, the stability of multidimensional persistence is a consequence of the one-dimensional persistence stability.

We start by recalling that the following parameterized family of half-planes in $\mathbb{R}^n \times \mathbb{R}^n$ is a foliation of Δ^+ (cf. [2, Prop. 1]).

Definition 3.7 *For every vector $\vec{l} = (l_1, \dots, l_n)$ of \mathbb{R}^n such that $l_i > 0$ for $i = 1, \dots, n$, and $\sum_{i=1}^n l_i^2 = 1$, and for every vector $\vec{b} = (b_1, \dots, b_n)$ of \mathbb{R}^n such that $\sum_{i=1}^n b_i = 0$, we shall say that the pair (\vec{l}, \vec{b}) is admissible. We shall denote the set of all admissible pairs in $\mathbb{R}^n \times \mathbb{R}^n$ by Adm_n . Given an admissible pair (\vec{l}, \vec{b}) ,*

we define the half-plane $\pi_{(\vec{l}, \vec{b})}$ of $\mathbb{R}^n \times \mathbb{R}^n$ by the following parametric equations:

$$\begin{cases} \vec{u} = s\vec{l} + \vec{b} \\ \vec{v} = t\vec{l} + \vec{b} \end{cases}$$

for $s, t \in \mathbb{R}$, with $s < t$.

Since these half-planes $\pi_{(\vec{l}, \vec{b})}$ constitute a foliation of Δ^+ , for each $(\vec{u}, \vec{v}) \in \Delta^+$ there exists one and only one $(\vec{l}, \vec{b}) \in \text{Adm}_n$ such that $(\vec{u}, \vec{v}) \in \pi_{(\vec{l}, \vec{b})}$. Observe that \vec{l} and \vec{b} only depend on (\vec{u}, \vec{v}) .

A first property of this foliation is that the restriction of $\beta_{\vec{\varphi}}$ to each leaf can be seen as a particular one-dimensional PBNs function. Intuitively, on each half-plane $\pi_{(\vec{l}, \vec{b})}$ one can find the PBNs corresponding to the filtration of X obtained by sweeping the line through \vec{u} and \vec{v} parameterized by $\gamma_{(\vec{l}, \vec{b})} : \mathbb{R} \rightarrow \mathbb{R}^n$, with $\gamma_{(\vec{l}, \vec{b})}(\tau) = \tau\vec{l} + \vec{b}$.

A second property is that this filtration corresponds to the one given by the lower level sets of a certain scalar-valued continuous function. Both these properties are stated in the next theorem, analogous to [6, Thm. 2], and are intuitively shown in Figure 5.

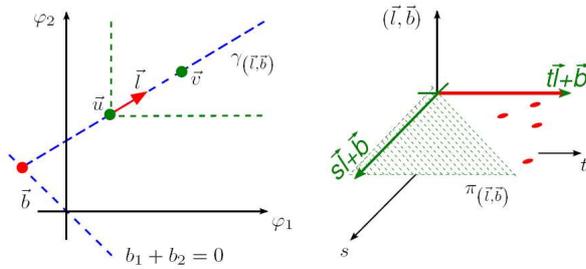


Figure 5: One-dimensional reduction of two-dimensional PBNs. Left: a one-dimensional filtration is constructed sweeping the line through \vec{u} and \vec{v} . A unit vector \vec{l} and a point \vec{b} are used to parameterize this line as $\gamma_{(\vec{l}, \vec{b})}(\tau) = \tau\vec{l} + \vec{b}$. Right: the persistence diagram of this filtration can be found on the leaf $\pi_{(\vec{l}, \vec{b})}$ of the foliation.

Theorem 3.8 For every $(\vec{u}, \vec{v}) \in \Delta^+$, let (\vec{l}, \vec{b}) be the only admissible pair such that $(\vec{u}, \vec{v}) = (s\vec{l} + \vec{b}, t\vec{l} + \vec{b}) \in \pi_{(\vec{l}, \vec{b})}$. Let moreover $\varphi_{(\vec{u}, \vec{v})} : X \rightarrow \mathbb{R}$ be the continuous filtering function defined by setting

$$\varphi_{(\vec{u}, \vec{v})}(x) = \min_i l_i \cdot \max_i \frac{\varphi_i(x) - b_i}{l_i}.$$

Then $X(\vec{\varphi} \preceq \vec{u}) = X((\min_i l_i)^{-1} \varphi_{(\vec{u}, \vec{v})} \leq s)$. Therefore

$$\beta_{\vec{\varphi}}(\vec{u}, \vec{v}) = \beta_{(\min_i l_i)^{-1} \varphi_{(\vec{u}, \vec{v})}}(s, t).$$

Finally, the most important property of our foliation is that it allows us to obtain an analogue of the distance d_{match} for the multidimensional case, denoted by D_{match} , having a particularly simple form, yet yielding the desired stability result.

D_{match} was introduced in [6] (see also [2]), although in the narrower setting of max-tame filtering functions, and can be rewritten as follows.

Definition 3.9 *Let X, Y be triangulable spaces endowed with continuous functions $\vec{\varphi} : X \rightarrow \mathbb{R}^n$, $\vec{\psi} : Y \rightarrow \mathbb{R}^n$. The (extended) multidimensional matching distance D_{match} between $\beta_{\vec{\varphi}}$ and $\beta_{\vec{\psi}}$ is defined as*

$$D_{match}(\beta_{\vec{\varphi}}, \beta_{\vec{\psi}}) = \sup_{(\vec{u}, \vec{v}) \in \Delta^+} d_{match}(\beta_{\varphi_{(\vec{u}, \vec{v})}}, \beta_{\psi_{(\vec{u}, \vec{v})}}). \quad (2)$$

We are now ready to state our result about the stability of multidimensional PBNs with respect to function perturbations.

Theorem 3.10 *If X is a triangulable space, then D_{match} is a distance on the set $\{\beta_{\vec{\varphi}} \mid \vec{\varphi} : X \rightarrow \mathbb{R}^n \text{ continuous}\}$. Moreover,*

$$D_{match}(\beta_{\vec{\varphi}}, \beta_{\vec{\psi}}) \leq \max_{x \in X} \|\vec{\varphi}(x) - \vec{\psi}(x)\|_{\infty}.$$

Roughly speaking, this theorem states that small changes in a vector-valued filtering function induce small changes in the associated multidimensional PBNs, with respect to the distance D_{match} .

4 Stability with respect to domain perturbations

In this section we describe how PBNs changes when the topological space is changed due to sampling or noise. This is as much important as the stability with respect to the change of measuring functions. First we will see that studying a submanifold of a Euclidean space through a finite sampling still conveys significant information about the PBNs of the submanifold itself. Then we show that by a change in perspective we can also achieve stability of PBNs with respect to domain perturbations.

4.1 Estimating multidimensional persistent homology through a finite sampling

An exact computation of the persistent Betti numbers of a submanifold X of a Euclidean space is possible only in a theoretical setting. In practical situations,

only a finite sample of X is available. We show that, under suitable density conditions, it is possible to estimate the multidimensional persistent Betti numbers of X from the ones of a union of balls centered on the sample points; this even yields the exact value in restricted areas of the domain. The proofs of the results presented in this section can be found in [11]. A similar study was performed in [15, Sect. 4].

Throughout this Section, X will be a compact Riemannian (triangulable) submanifold of \mathbb{R}^m . We want to get information on X out of a finite set of points. First, the points will be sampled on X itself, then even in a (narrow) neighborhood. In both cases, the idea is to consider a covering of X made of balls centered on the sampling points.

What we get, using a result of [33], is a double inequality which yields an estimate of the PBNs of X within a fixed distance from the discontinuity sets of the PBNs (meant as integer functions on Δ^+) of the union U of the balls of the covering, but even offers the exact value of it at points sufficiently far from the discontinuity sets.

Definition 4.1 *Let $\vec{\varphi} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ be a continuous function. Then, for $\varepsilon \in \mathbb{R}^+$, the modulus of continuity $\Omega(\varepsilon)$ of $\vec{\varphi}$ is:*

$$\Omega(\varepsilon) = \max_{j=1, \dots, n} \sup \left\{ |\varphi_j(\vec{p}) - \varphi_j(\vec{p}')| : \vec{p}, \vec{p}' \in \mathbb{R}^m, \|\vec{p} - \vec{p}'\| \leq \varepsilon \right\}.$$

In other words $\Omega(\varepsilon)$ is the maximum over all moduli of continuity of the single components of $\vec{\varphi}$.

A condition number $\frac{1}{\tau}$ is associated with a compact Riemannian submanifold X of \mathbb{R}^m .

Definition 4.2 *τ is the largest number such that every open normal bundle B about X of radius s is embedded in \mathbb{R}^m for $s < \tau$.*

Theorem 4.3 *Let $\delta < \sqrt{\frac{3}{5}}\tau$ and let $L = \{l_1, \dots, l_k\}$ be a set of points of X such that for every $p \in X$ there exists an $l_j \in L$ for which $\|p - l_j\| < \frac{\delta}{2}$. Let U be the union of the balls $B(l_j, \delta)$ of radius δ centered at l_j , $j = 1, \dots, k$.*

If (\vec{u}, \vec{v}) is a point of Δ^+ and if $\vec{u} + \vec{\omega}(\delta) \prec \vec{v} - \vec{\omega}(\delta)$, where $\vec{\omega}(\delta) = (\Omega(\delta), \dots, \Omega(\delta)) \in \mathbb{R}^n$, then

$$\beta_{\vec{\varphi}|_U}(\vec{u} - \vec{\omega}(\delta), \vec{v} + \vec{\omega}(\delta)) \leq \beta_{\vec{\varphi}|_X}(\vec{u}, \vec{v}) \leq \beta_{\vec{\varphi}|_U}(\vec{u} + \vec{\omega}(\delta), \vec{v} - \vec{\omega}(\delta)).$$

When we foliate the domain Δ^+ of the PBNs as in Section 3.2 — or simply when $n = 1$ — the discontinuity sets are (possibly infinite) line segments, and the regions of Δ^+ where only the inequality holds appear as strips around them (which we colloquially call “blind strips”). The width of such strips is a representation of the approximation error, in that it is directly related to $\Omega(\delta)$, where $1/\delta$ represents the density of the sampling.

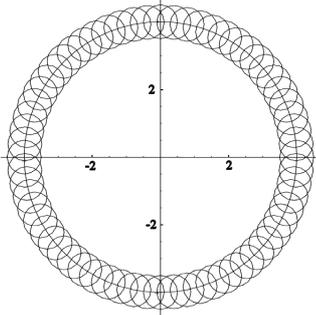


Figure 6: The circle of radius 4, X , covered by the ball union U .

The following examples show how Theorem 4.3 can be used for applications. Let X be a circle of radius 4 in \mathbb{R}^2 ; we observe that τ is exactly the radius of X , so $\tau = 4$. In order to create a well defined approximation we need that $\delta < \sqrt{\frac{3}{5}}\tau$.

In the first example we have taken $\delta = 0.5$. Now, to satisfy the assumptions of Theorem 4.3 (that for every $p \in X$ there exists an $l_j \in L$ such that $\|p - l_j\| < \frac{\delta}{2}$), we have chosen 64 points l_j on X . Moreover we have sampled X uniformly, so that there is a point every $\frac{\pi}{32}$ radians (Figure 6). We stick to the monodimensional case, choosing $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$, with $\varphi(x, y) = |y|$. U is the resulting ball union.

Figures 7 and 8 represent the PBN functions at degree zero of X and U respectively. In Figure 7 there is only a big triangle where the value 2 signals the two different connected components generated by $\varphi|_X$. The two connected components collapse to one at value 4. In Figure 8 there is also a big triangle representing the two connected components, but they collapse at value 3.53106. Moreover there are 4 other very small triangles near the diagonal, representing more connected components generated by the balls that are furthest from the x -axis. In the last figure (Figure 9) the blind strips around the discontinuity lines of $\beta_{\varphi|_U}$ are shown. The width of these strips, since $\Omega(\delta) = 0.5$, is equal to $2\Omega(\delta) = 1$. This figure illustrates the idea underlying Theorem 4.3. Taken a point (u, v) outside the strips, the values of the PBNs of U at $(u - \Omega(\delta), v + \Omega(\delta))$ and $(u + \Omega(\delta), v - \Omega(\delta))$ are the same. So also the value of the PBNs of X at (u, v) is determined.

So far we have approximated X by points picked up on X itself, but it is also possible to choose the points near X , by respecting some constraints. Once more, this is possible thanks to a result of [33].

Theorem 4.4 *Let $L = \{l_1, \dots, l_k\}$ be a set of points in the tubular neighborhood of radius s around X and $U = \bigcup_{j=1, \dots, k} B(l_j, \delta)$ be the union of the balls of \mathbb{R}^m centered at the points of L and with radius δ . If for every points $p \in X$, there exists a point $l_j \in L$ such that $\|p - l_j\| < s$, then U is a deformation retract of X , for all $s < (\sqrt{9} - \sqrt{8})\tau$ and $\delta \in \left(\frac{(s+\tau) - \sqrt{s^2 + \tau^2 - 6s\tau}}{2}, \frac{(s+\tau) + \sqrt{s^2 + \tau^2 - 6s\tau}}{2} \right)$.*

If (\vec{u}, \vec{v}) is a point of Δ^+ and if $\vec{u} + \vec{\omega}(\delta + s) \prec \vec{v} - \vec{\omega}(\delta + s)$, where $\vec{\omega}(\delta + s) =$

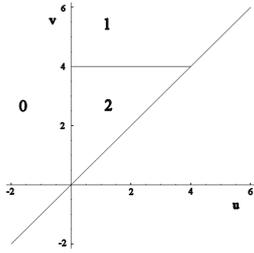


Figure 7: The representation of $\beta_{\varphi|_X}$, the 0-PBNs of X .

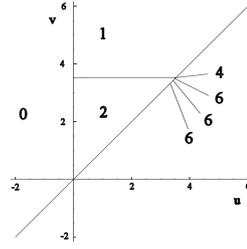


Figure 8: The representation of $\beta_{\varphi|_U}$, the 0-PBNs of the ball union U

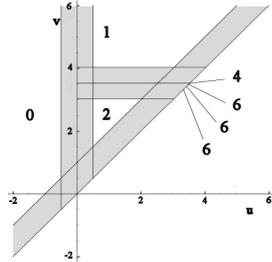


Figure 9: The blind strips of β_{φ} .

$(\Omega(\delta + s), \dots, \Omega(\delta + s)) \in \mathbb{R}^n$, then

$$\beta_{\bar{\varphi}|_U}(\vec{u} - \vec{\omega}(\delta + s), \vec{v} + \vec{\omega}(\delta + s)) \leq \beta_{\bar{\varphi}|_X}(\vec{u}, \vec{v}) \leq \beta_{\bar{\varphi}|_U}(\vec{u} + \vec{\omega}(\delta + s), \vec{v} - \vec{\omega}(\delta + s)).$$

In [11, Sect. 5] a similar double inequality, relating the PBNs of X with the ones of a simplicial complex derived from U , can also be found.

4.2 Stability of persistent Betti numbers with respect to noisy domains

In this section we propose a general approach to the problem of stability of persistent homology groups with respect to domain perturbations. The proofs of the results presented in this section can be found in [28].

Changes of the space under study can be measured in a number of different ways. Indeed, according to the kind of noise producing the perturbation, some distances are more suitable than other to compare sets. For example, the Hausdorff distance is useful to measure distortions of the domain, while the symmetric difference distance can cope with the presence of outliers.

Our main idea is to reduce the problem of stability with respect to changes of the topological space to that of stability with respect to changes of the measuring functions. This is achieved by substituting the domain K we are interested in with an appropriate function f_K defined on a fixed set X containing K , so that

the perturbation of the set K becomes a perturbation of the function f_K . As a consequence, the original measuring function $\vec{\varphi}|_K : K \rightarrow \mathbb{R}^n$ is replaced by a new measuring function $\vec{\Phi} : X \rightarrow \mathbb{R}^{n+1}$, $\vec{\Phi} = (f_K, \vec{\varphi})$. Persistent Betti numbers of $(X, \vec{\Phi})$ can be compared using the multidimensional matching distance. In this way we can obtain robustness of persistent homology groups under domain perturbations.

In particular, we use this strategy when sets are compared by the Hausdorff distance δ_H . In this case, taking f_K equal to the distance function from K , we show that the multidimensional matching distance between the PBNs associated with two compact sets K_1 and K_2 is always upperly bounded by the Hausdorff distance $\delta_H(K_1, K_2)$ between K_1 and K_2 (Theorem 4.5). At the same time, we show that, in our approach, the information about the original domain K and its original measuring function $\vec{\varphi}|_K$ is fully maintained in the persistent homology groups of $(X, \vec{\Phi})$ (Theorem 4.6).

Given two domains K_1 and K_2 , and two functions $\vec{\varphi}_1, \vec{\varphi}_2 : X \rightarrow \mathbb{R}^n$, our first result relates the distance D_{match} between the new pairs $(X, \vec{\Phi}_1), (X, \vec{\Phi}_2)$ to the change of the measuring functions $\vec{\varphi}_1$ and $\vec{\varphi}_2$, and to the Hausdorff distance between the original sets K_1, K_2 . More precisely, it states stability with respect to both set and function perturbations. Indeed, the change in the distance D_{match} is shown to be never greater than the maximum among the change in the Hausdorff distance between the domains K_1 and K_2 and the change in the sup-norm between the measuring functions $\vec{\varphi}_1$ and $\vec{\varphi}_2$. In particular, if $\vec{\varphi}_1$ and $\vec{\varphi}_2$ coincide then the change in the distance D_{match} is never greater than the Hausdorff distance $\delta_H(K_1, K_2)$ between K_1 and K_2 .

Theorem 4.5 *Let K_1, K_2 be non-empty closed subsets of a triangulable subspace X of \mathbb{R}^m . Let $d_{K_1}, d_{K_2} : X \rightarrow \mathbb{R}$ be their respective distance functions. Moreover, let $\vec{\varphi}_1, \vec{\varphi}_2 : X \rightarrow \mathbb{R}^n$ be vector-valued continuous functions. Then, defining $\vec{\Phi}_1, \vec{\Phi}_2 : X \rightarrow \mathbb{R}^{n+1}$ by $\vec{\Phi}_1 = (d_{K_1}, \vec{\varphi}_1)$ and $\vec{\Phi}_2 = (d_{K_2}, \vec{\varphi}_2)$, the following inequality holds:*

$$D_{match}(\beta_{\vec{\Phi}_1}, \beta_{\vec{\Phi}_2}) \leq \max\{\delta_H(K_1, K_2), \|\vec{\varphi}_1 - \vec{\varphi}_2\|_\infty\}.$$

An example illustrates this result. We work with the binary digital image represented in Figure 10 (left), and we corrupt this image by adding noise, as shown in Figure 10 (right).

Black pixels of left and right images represent the sets K_1, K_2 under study, respectively, whereas in both cases the rectangle of black and white pixels together constitute the set X . The so obtained noisy set K_2 is close to the original set K_1 with respect to the Hausdorff distance. A graph structure based on the local 4-neighbors adjacency relations of the digital points is used in order to topologize the images.

Fixed the point $c \in X$ corresponding to the center of mass of K_1 , the chosen measuring function for both instances is $\varphi : X \rightarrow \mathbb{R}$, $\varphi(p) = -\|p - c\|$.

Confining ourselves to zeroth homology, Figure 11 (left) shows the persistence diagram of the 1-dimensional 0th PBNs $\beta_{\varphi|_{K_1}}$. It displays eight relevant points in

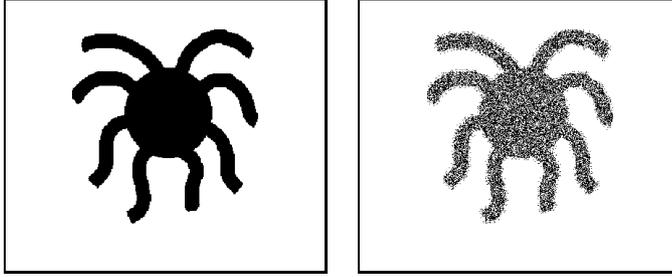


Figure 10: Two binary images of an octopus. The image on the right is a noisy version of that on the left.

the persistence diagram, corresponding to the eight tentacles of the octopus. Only one of these points is at infinity (and therefore depicted by a vertical line rather than by a circle) since K_1 has only one connected component. As for $\beta_{\varphi|_{K_2}}$, due to the presence of a great quantity of connected components in the noisy octopus, its persistence diagram has a very large number of points at infinity, and a figure showing them all would be hardly readable. For this reason Figure 11 (right) shows only a small subset of its persistence diagram. However it is sufficient to perceive how dissimilar it is from $\beta_{\varphi|_{K_1}}$.

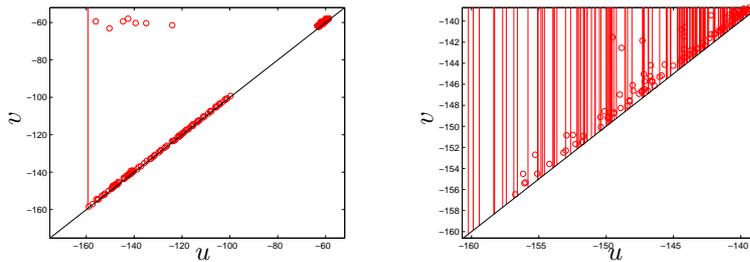


Figure 11: Left: The persistence diagram of $\beta_{\varphi|_{K_1}}$ corresponding to the original octopus image. Right: A detail of the persistence diagram of $\beta_{\varphi|_{K_2}}$ corresponding to the noisy octopus image.

As suggested by Theorem 4.5, if instead we compare K_1 and K_2 by means of the PBNs $\beta_{\vec{\Phi}_1}$ and $\beta_{\vec{\Phi}_2}$, where $\vec{\Phi}_1 : X \rightarrow \mathbb{R}^2$, $\vec{\Phi}_1 = (d_{K_1}, \varphi)$, and $\vec{\Phi}_2 : X \rightarrow \mathbb{R}^2$, $\vec{\Phi}_2 = (d_{K_2}, \varphi)$, we can see the similarity between K_1 and K_2 modulo the added noise.

Since the domain of $\beta_{\vec{\Phi}_1}$ and $\beta_{\vec{\Phi}_2}$ is in $\mathbb{R}^2 \times \mathbb{R}^2$, we explore these PBNs by means of the foliation method, allowing for a one-dimensional reduction. To this

end, we consider the functions $F_{(\vec{l}, \vec{b})}^{\vec{\Phi}_i} : X \rightarrow \mathbb{R}$ defined by setting, for every $x \in X$,

$$F_{(\vec{l}, \vec{b})}^{\vec{\Phi}_i}(x) = \max \left\{ \frac{d_{K_i}(x) - b_1}{l_1}, \frac{\varphi(x) - b_2}{l_2} \right\},$$

for $i = 1, 2$, and the number $\mu = \min\{l_1, l_2\}$.

The stability of our approach is illustrated in Figure 12. In Figure 12 (a)-(b), we show the PBNs $\beta_{\vec{\Phi}_1}$ and $\beta_{\vec{\Phi}_2}$ both restricted to the half-plane $\pi_{(\vec{l}, \vec{b})}$, with $\vec{l} = (0.1483, 0.9889)$ and $\vec{b} = (13.0434, -13.0434)$, that is the half-plane of the foliation containing the point $((0, -100), (3, -80))$. The considered half-plane has been chosen so that it contains points where PBNs take non-trivial values. Formally speaking, Figure 12 (a)-(b) shows the PBNs of $F_{(\vec{l}, \vec{b})}^{\vec{\Phi}_1}$ and $F_{(\vec{l}, \vec{b})}^{\vec{\Phi}_2}$, respectively. We can already appreciate their similarity, even if their distance is not necessarily smaller than the Hausdorff distance between K_1 and K_2 . In Figure 12 (c)-(d), we show the PBNs of $\mu \cdot F_{(\vec{l}, \vec{b})}^{\vec{\Phi}_1}$ and $\mu \cdot F_{(\vec{l}, \vec{b})}^{\vec{\Phi}_2}$, that are the functions appearing in the definition of D_{match} . This change of functions corresponds to “rescaling up” the domain of the PBNs. The stability of D_{match} guarantees that the distance between these two PBNs is not greater than the Hausdorff distance between K_1 and K_2 .

Until now, we have shown that we can obtain the wanted stability with respect to perturbation of the space by considering suitable vector-valued filtering functions. One could think that this is done at the price of forgetting information about the original problem. On the contrary, our method allows to retrieve the PBNs invariants of $(K, \vec{\varphi}|_K)$ from the PBNs of $(X, \vec{\Phi})$, with $\vec{\Phi} = (d_K, \vec{\varphi})$. This is stated in the next key result, showing that for any sufficiently small value of $\eta \in \mathbb{R}$ there exists a sufficiently small value $\varepsilon \in \mathbb{R}$ with $0 \leq \varepsilon < \eta$ such that $\beta_{\vec{\varphi}|_K}(\vec{u}, \vec{v}) = \beta_{\vec{\Phi}}((\varepsilon, \vec{u}), (\eta, \vec{v}))$.

Theorem 4.6 *Let K be a non-empty triangulable subset of a triangulable subspace X of \mathbb{R}^m . Moreover, let $\vec{\varphi} : X \rightarrow \mathbb{R}^n$ be a continuous function. Setting $\vec{\Phi} : X \rightarrow \mathbb{R}^{n+1}$, $\vec{\Phi} = (d_K, \vec{\varphi})$, for every $\vec{u}, \vec{v} \in \mathbb{R}^n$ with $\vec{u} \prec \vec{v}$, there exists a real number $\hat{\eta} > 0$ such that, for any $\eta \in \mathbb{R}$ with $0 < \eta \leq \hat{\eta}$, there exists a real number $\hat{\varepsilon} = \hat{\varepsilon}(\eta)$, with $0 < \hat{\varepsilon} < \eta$, for which*

$$\beta_{\vec{\varphi}|_K}(\vec{u}, \vec{v}) = \beta_{\vec{\Phi}}((\varepsilon, \vec{u}), (\eta, \vec{v})),$$

for every $\varepsilon \in \mathbb{R}$ with $0 \leq \varepsilon \leq \hat{\varepsilon}$. In particular,

$$\beta_{\vec{\varphi}|_K}(\vec{u}, \vec{v}) = \lim_{\eta \rightarrow 0^+} \beta_{\vec{\Phi}}((0, \vec{u}), (\eta, \vec{v})).$$

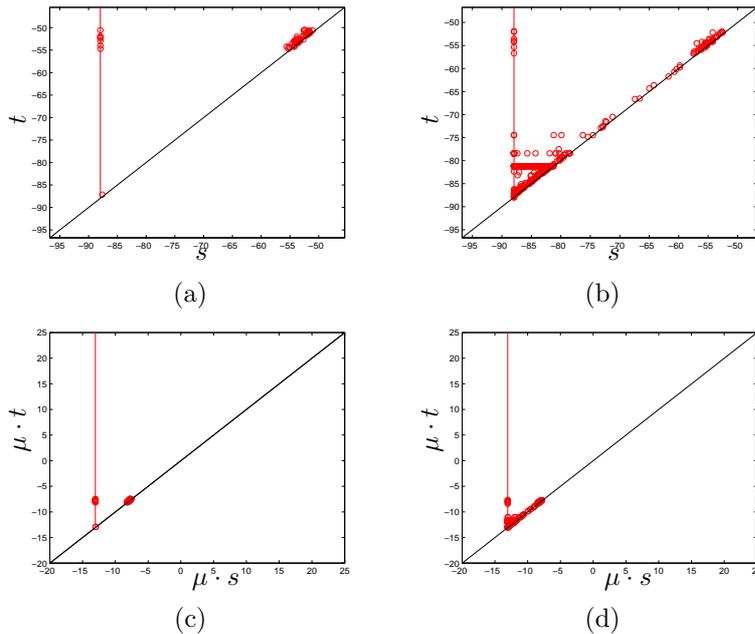


Figure 12: (a) The PBNs $\beta_{\vec{\Phi}_1}$ restricted to the half-plane $\pi_{(\vec{l}, \vec{b})}$, with $\vec{l} = (0.1483, 0.9889)$ and $\vec{b} = (13.0434, -13.0434)$, that is the half-plane of the foliation containing the point $((0, -100), (3, -80))$. (b) The PBNs $\beta_{\vec{\Phi}_2}$ restricted to the same half-plane. (c)-(d) The same restrictions as in (a)-(b), respectively, but rescaled by $\mu = \min\{l_1, l_2\}$.

Many applications require that the presence of single outliers does not affect the evaluation of similarity. In these cases, always assuming K triangulable, it is sufficient to study the closure of the interior of K instead of K itself. Indeed, applying Theorems 4.5 and 4.6 with the closure of the interior of K instead of K , we obtain a result of stability of persistent homology groups with respect to the perturbations of the studied set and a reconstruction result for the original persistent homology groups modulo perturbations of zero measure.

We underline once more that the results of this section are based on the idea of translating the problem of stability with respect to set perturbations into that of stability with respect to function perturbations. Therefore, the use of the distance function is only one among many ways to achieve this end and has the advantage of working well when sets are compared using the Hausdorff distance. One could conceive different ways, in connection with other methods to compare sets.

References

- [1] Silvia Biasotti, Andrea Cerri, Patrizio Frosini, and Daniela Giorgi. A new algorithm for computing the 2-dimensional matching distance between size functions. *Pattern Recognition Letters*, 32:1735–1746, 2011.
- [2] Silvia Biasotti, Andrea Cerri, Patrizio Frosini, Daniela Giorgi, and Claudia Landi. Multidimensional size functions for shape comparison. *J. Math. Imaging Vision*, 32(2):161–179, 2008.
- [3] Silvia Biasotti, Andrea Cerri, and Daniela Giorgi. k-dimensional size functions for shape description and comparison. In *Proceedings ICIAP 2007: 14th International Conference on Image Analysis and Processing*, pages 345–352. IEEE Computer Society, 2007.
- [4] Silvia Biasotti, Leila De Floriani, Bianca Falcidieno, Patrizio Frosini, Daniela Giorgi, Claudia Landi, Laura Papaleo, and Michela Spagnuolo. Describing shapes by geometrical-topological properties of real functions. *ACM Comput. Surv.*, 40(4):1–87, 2008.
- [5] Aldo Brucale, Michele d’Amico, Massimo Ferri, Luciano Gualandri, and Alberto Lovato. Size functions for image retrieval: A demonstrator on randomly generated curves. In M.S. Lew, N. Sebe, and J.P. Eakins, editors, *Proc. CIVR02, London*, volume 2383 of *LNCS*, pages 235–244. Springer-Verlag, 2002.
- [6] Francesca Cagliari, Barbara Di Fabio, and Massimo Ferri. One-dimensional reduction of multidimensional persistent homology. *Proc. Amer. Math. Soc.*, 138(8):3003–3017, 2010.
- [7] Francesca Cagliari and Claudia Landi. Finiteness of rank invariants of multidimensional persistent homology groups. *Applied Mathematics Letters*, 24(4):516 – 518, 2011.
- [8] Gunnar Carlsson and Afra J. Zomorodian. The theory of multidimensional persistence. In *SCG ’07: Proceedings of the twenty-third annual symposium on computational geometry*, pages 184–193, New York, NY, USA, 2007. ACM.
- [9] Gunnar Carlsson and Afra J. Zomorodian. The theory of multidimensional persistence. *Discrete & Computational Geometry*, 42(1):71–93, 2009.
- [10] Gunnar Carlsson, Afra J. Zomorodian, Anne Collins, and Leonidas J. Guibas. Persistence barcodes for shapes. *International Journal of Shape Modeling*, 11(2):149–187, 2005.
- [11] Niccolò Cavazza, Massimo Ferri, and Claudia Landi. Estimating multidimensional persistent homology through a finite sampling. Technical Report, Università di Bologna, 2010. <http://amsacta.cib.unibo.it/2861/>.

- [12] Andrea Cerri, Barbara Di Fabio, Massimo Ferri, Patrizio Frosini, and Claudia Landi. Betti numbers in multidimensional persistent homology are stable functions. Tech. Rep., Univ. di Bologna, November 2010. <http://amsacta.cib.unibo.it/2923/>.
- [13] Andrea Cerri, Massimo Ferri, and Daniela Giorgi. Retrieval of trademark images by means of size functions. *Graph. Models*, 68(5):451–471, 2006.
- [14] Andrea Cerri and Patrizio Frosini. Advances in multidimensional Size Theory. *Image Analysis and Stereology*, 29:19–26, 2010.
- [15] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. *Discrete Comput. Geom.*, 37(1):103–120, 2007.
- [16] Michele d’Amico, Massimo Ferri, and Ignazio Stanganelli. Qualitative asymmetry measure for melanoma detection. In *Proceedings of the 2004 IEEE International Symposium on Biomedical Imaging*, pages 1155–1158, 2004.
- [17] Michele d’Amico, Patrizio Frosini, and Claudia Landi. Natural pseudodistance and optimal matching between reduced size functions. *Acta. Appl. Math.*, 109:527–554, 2010.
- [18] Pietro Donatini and Patrizio Frosini. Natural pseudodistances between closed manifolds. *Forum Mathematicum*, 16(5):695–715, 2004.
- [19] Pietro Donatini and Patrizio Frosini. Natural pseudodistances between closed surfaces. *Journal of the European Mathematical Society*, 9(2):231253, 2007.
- [20] Pietro Donatini and Patrizio Frosini. Natural pseudodistances between closed curves. *Forum Mathematicum*, 21(6):981999, 2009.
- [21] Herbert Edelsbrunner and John Harer. Persistent homology—a survey. In *Surveys on discrete and computational geometry*, volume 453 of *Contemp. Math.*, pages 257–282. Amer. Math. Soc., Providence, RI, 2008.
- [22] Herbert Edelsbrunner, David Letscher, and Afra J. Zomorodian. Topological persistence and simplification. *Discrete & Computational Geometry*, 28(4):511–533, 2002.
- [23] Samuel Eilenberg and Norman Steenrod. *Foundations of algebraic topology*. Princeton University Press, Princeton, New Jersey, 1952.
- [24] Massimo Ferri, Sandra Lombardini, and Clemente Pallotti. Leukocyte classification by size functions. In *Proc. second IEEE Workshop on Applications of Computer Vision.*, IEEE Computer Society Press, pages 223–229, 1994.
- [25] Patrizio Frosini. Measuring shapes by size functions. In David P. Casasent, editor, *Proc. of SPIE, Intelligent Robots and Computer Vision X: Algorithms and Techniques*, Boston, MA, volume 1607, pages 122–133, 1991.

- [26] Patrizio Frosini and Claudia Landi. Size Theory as a topological tool for computer vision. *Pattern Recognition and Image Analysis*, 9:596–603, 1999.
- [27] Patrizio Frosini and Claudia Landi. Size functions and formal series. *Appl. Algebra Engrg. Comm. Comput.*, 12(4):327–349, 2001.
- [28] Patrizio Frosini and Claudia Landi. Stability of multidimensional persistent homology with respect to domain perturbations, 2010. Preprint: arXiv:1001.1078v2.
- [29] Patrizio Frosini and Claudia Landi. Persistent Betti numbers for a noise tolerant shape-based approach to image retrieval. In A. Berciano et al., editor, *Proc. CAIP 2011*, volume 6854 of *LNCS*, pages 294–301. Springer-Verlag, 2011.
- [30] Patrizio Frosini and Claudia Landi. Uniqueness of models in persistent homology: the case of curves. *Inverse Problems*, 27:124005, 2011.
- [31] Patrizio Frosini and Michele Mulazzani. Size homotopy groups for computation of natural size distances. *Bulletin of the Belgian Mathematical Society*, 6(3):455–464, 1999.
- [32] Mohamed Handouyaya, Djemel Ziou, and Shengrui Wang. Sign language recognition using moment-based size functions. In *Vision Interface 99, Trois-Rivières*, 1999.
- [33] Partha Niyogi, Stephen Smale, and Shmuel Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete Comput. Geom.*, 39(1):419–441, 2008.
- [34] Alessandro Verri, Claudio Uras, Patrizio Frosini, and Massimo Ferri. On the use of size functions for shape analysis. *Biol. Cybern.*, 70:99–107, 1993.