

ESTIMATING MULTIDIMENSIONAL PERSISTENT HOMOLOGY THROUGH A FINITE SAMPLING

NICCOLÒ CAVAZZA, MASSIMO FERRI, AND CLAUDIA LANDI

ABSTRACT. 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.

Using these inequalities we improve a previous lower bound for the natural pseudodistance to assess dissimilarity between the shapes of two objects from a sampling of them.

Similar inequalities are proved for the multidimensional persistent Betti numbers of the ball union and the one of a combinatorial description of it.

1. INTRODUCTION

Persistent Topology is an innovative way of matching topology and geometry, and it proves to be an effective mathematical tool in Pattern Recognition [18, 3], particularly in Shape Comparison [1]. This new research area is experiencing a period of intense theoretical progress, particularly in the form of the multidimensional *persistent Betti numbers* (PBNs; also called *rank invariant* in [4]). In order to express its full potential for applications, it has to interface with the typical environment of Computer Science: It must be possible to deal with a finite sampling of the object of interest, and with combinatorial representations of it.

A predecessor of the PBNs, the *size function* (i.e. PBNs at degree zero) already enjoys such a connection, in that it is possible to estimate it from a finite, sufficiently dense sampling [16], and it is possible to simplify the computation by processing a related graph [9]. Moreover, strict inequalities hold only in “blind strips”, i.e. in the ω -neighborhood of the discontinuity lines, where ω is the modulus of continuity of the filtering (also called *measuring*) function. Out of the blind strips, the values of the size function of the original object, of a ball covering of it, and of the related graph coincide. As a consequence it is possible to estimate dissimilarity of the shape of two objects by using size functions of two samplings.

The present paper extends this result to the PBNs of any degree, by using an article by P. Niyogi, S. Smale, and S. Weinberger [22], for a ball covering of the object X — a submanifold of a Euclidean space — with balls centered at dense enough points of X or near X : Theorems 4.1 and 4.2. A combinatorial representation of X , with the corresponding inequalities (Theorem 6.1) is based on a construction by H. Edelsbrunner [13]. In Section 5 we use Theorem 4.1 to get a lower bound for the natural pseudodistance between two objects of which only finite samplings are given.

All results are provided for multidimensional filtering functions, although most current applications just use monodimensional ones.

It should be noted that the same kind of problem has been addressed in [7] by using the notion of Weak Feature Size. An inequality rather similar to ours of Theorem 4.1, also extending the main result of [16], is exposed in the proof of the Homology Inference Theorem in Section 4 of [8]. The progress represented by the present paper consists in the stress on inequalities (and not just on the consequent equalities), in the use of multidimensional filtering functions, in the fact that any continuous function — not just distance — is considered, and in the estimate of the natural pseudodistance.

Some simple examples illustrate the results.

2. PRELIMINARY RESULTS

In this section, we report some results on the stability of the PBNs and some topological properties of compact Riemannian submanifolds of \mathbb{R}^m . For basic notions on homology and persistent homology we refer the reader to [15] and [14], for classical properties of submanifolds to [19].

2.1. Multidimensional persistent Betti numbers. In this paper we will always work with coefficients in a field \mathbb{K} , so that all homology modules are vector spaces. First we define the following relation \prec (resp. \preceq) in \mathbb{R}^n : if $u = (u_1, \dots, u_n)$ and $v = (v_1, \dots, v_n)$, we write $u \prec v$ (resp. $u \preceq v$) if and only if $u_j < v_j$ (resp. $u_j \leq v_j$) for $j = 1, \dots, n$. We also define Δ^+ as the open set $\{(u, v) \in \mathbb{R}^n \times \mathbb{R}^n \mid u \prec v\}$. As usual, a topological space X is *triangulable* if there is a finite simplicial complex whose underlying space is homeomorphic to X ; for a submanifold of a Euclidean space, it will mean that its triangulation can be extended to a domain containing it. We use Čech homology because it guarantees some useful continuity properties [5]. For terms and concepts concerning this homology, we refer to [15].

Persistent Betti numbers Let X be a triangulable space and $f = (f_1, \dots, f_n) : X \rightarrow \mathbb{R}^n$ be a continuous function. f is called a *filtering function*. We denote by $X\langle f \preceq u \rangle$ the lower level subset $\{p \in X \mid f_j(p) \leq u_j, j = 1, \dots, n\}$. Then, for each $i \in \mathbb{Z}$, the *i -th multidimensional persistent Betti numbers (briefly PBNs) function of (X, f)* is $\beta_{(X, f, i)} : \Delta^+ \rightarrow \mathbb{N}$ defined as $\beta_{(X, f, i)}(u, v) = \dim(\text{Im} \iota)$, with

$$\iota : \check{H}_i(X\langle f \preceq u \rangle) \rightarrow \check{H}_i(X\langle f \preceq v \rangle),$$

the homomorphism induced by the inclusion map of the sublevel set $X\langle f \preceq u \rangle \subseteq X\langle f \preceq v \rangle$. Here \check{H}_i denotes the i -th Čech homology module.

Remark When PBNs were considered in [8] the filtering functions were taken as to be tame, but in our case, since we need continuous maps, we refer to Section 2.2 of [5] for the relevant extension.

PBNs give us a way to analyze triangulable spaces through their homological properties. Then it is natural to introduce a distance for comparing them. This has been done, and through this distance it has been possible to prove stability of the PBNs under variations of the filtering function in the one-dimensional [8] and multidimensional case [5].

2.2. Topological properties of compact Riemannian submanifolds of \mathbb{R}^m .

As hinted in Section 1, we are interested in getting information on a submanifold of \mathbb{R}^m via a finite sampling of it and a related ball covering. To this goal, we now state some properties of compact Riemannian submanifolds of \mathbb{R}^m , especially referred to such an approximating covering. Definition 2.2 and Proposition 2.2 are due to P. Niyogi, S. Smale, and S. Weinberger [22]. The main idea is that, under suitable hypotheses, it is possible to get, from a sampling of a submanifold, a ball covering whose union retracts on it.

First, notice that the spaces \mathbb{R}^m and \mathbb{R}^n play two different rôles in our arguments: the ambient space of our submanifolds (which will always be \mathbb{R}^m) is endowed with the classical Euclidean norm and has no partial order relation on it. On the other hand, the codomain of the filtering functions (\mathbb{R}^n throughout) is endowed with the max norm and with the partial order relation \preceq , as defined at the beginning of Section 2.1.

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

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

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

For a given compact Riemannian submanifold X of \mathbb{R}^m , the normal space of X at a point $p \in X$ is the vector subspace $N_p X$ of the tangent space of \mathbb{R}^m at p , $T_p \mathbb{R}^m$, formed by the vectors orthogonal to $T_p X$, the tangent space of X at p . The *open normal bundle to X of radius s* is defined as the subset of $T\mathbb{R}^m$

$$\{(p, v) \in T\mathbb{R}^m \mid p \in X, v \in N_p X, \|v\| < s\}$$

By Theorem 10.19 of [20] and by compactness, there exists an embedding of the open normal bundle to X of radius s into \mathbb{R}^m for some s . Its image $Tub_s X$ is called a *tubular neighborhood* of X .

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

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

Proposition (Prop. 3.1 of [22]) Let X be a compact Riemannian submanifold of \mathbb{R}^m . Let $L = \{l_1, \dots, l_k\}$ be a collection of points of X , and let $U = \bigcup_{j=1,\dots,k} B(l_j, \delta)$ be the union of balls of \mathbb{R}^m with center at the points of L and radius δ . Now, if L is such that for every point $p \in X$ there exists an $l_j \in L$

such that $\|p - l_j\| < \frac{\delta}{2}$, then, for every $\delta < \sqrt{\frac{3}{5}}\tau$, X is a deformation retract of U . So they have the same homology.

Remark The proof of Proposition 2.2 gives us a way to construct a retraction $\pi : U \rightarrow X$ and a homotopy $F : U \times I \rightarrow U$ such that $F(q, 0) = q$ and $F(q, 1) = \pi(q)$. Let $\pi_0 : Tub_\tau \rightarrow X$ be the canonical projection from the tubular neighborhood of radius τ of X onto X . Then π is the restriction of π_0 to U for which it holds:

$$\pi(q) = \arg \min_{p \in X} \|q - p\|.$$

Then the homotopy is given by

$$F(q, t) = (1 - t)q + t\pi(q).$$

It is also important to observe that the retraction π moves the points of U less than δ ; this is because the trajectory of $\pi(q)$ always remains inside a ball of U that contains q (q can be contained in the intersection of different balls), for every $q \in U$. In fact $\pi^{-1}(q) = U \cap \text{Tan}_q^\perp \cap B_r(q)$ (for a complete argument we refer to Section 4 of [22]).

3. RETRACTS

Aim of this Section is to yield two rather general results, which will be specialized to Theorem 4.1 and Lemma 6.1. Throughout this Section, Y will be a compact Riemannian (hence triangulable) submanifold of \mathbb{R}^m and V will be a compact, triangulable subspace of \mathbb{R}^m such that Y is a deformation retract of V , with retraction r and homotopy $G : V \times I \rightarrow V$ from the identity of V , 1_V , to r . Moreover $\forall y \in Y$, $\forall v \in r^{-1}(y)$, $\forall t \in I$ we assume that $(r \circ G)(v, t) = y$.

Let also $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ be a continuous function, and f_Y and f_V be the restrictions of f to Y and V respectively.

Lemma $Y \langle f_Y \preceq x \rangle$ is a deformation retract of $V \langle f_Y \circ r \preceq x \rangle$.

Proof. Let $r_x : V \langle f_Y \circ r \preceq x \rangle \rightarrow Y \langle f_Y \preceq x \rangle$ be the restriction of r to $V \langle f_Y \circ r \preceq x \rangle$. It is well-defined since, by the definition of the two sets, $r_x(V \langle f_Y \circ r \preceq x \rangle) \subseteq Y \langle f_Y \preceq x \rangle$. We now set $G_x : V \langle f_Y \circ r \preceq x \rangle \times I \rightarrow V \langle f_Y \circ r \preceq x \rangle$ as the restriction of G to $V \langle f_Y \circ r \preceq x \rangle \times I$. This restriction is well-defined, because the path from v to $r_x(v)$ is all contained in $V \langle f_Y \circ r \preceq x \rangle$, thanks to the assumptions on G and r . Moreover, it is continuous and for every $v \in V \langle f_Y \circ r \preceq x \rangle$, $G_x(v, 0) = v$ and $G_x(v, 1) = r_x(v)$. So it is the searched for deformation retraction. \square

Remark Since the homotopy G is relative to Y (i.e. keeps the points of Y fixed throughout), this is what is called a *strong* deformation retract in [15].

Now let $\varepsilon = \max_{v \in V} \|r(v) - v\|$ and $\omega(\varepsilon) = (\Omega(\varepsilon), \dots, \Omega(\varepsilon)) \in \mathbb{R}^n$, where Ω is the modulus of continuity of f (Def. 2.2). For sake of simplicity, set now $\beta_V = \beta_{(V, f_V, i)}$, $\beta_Y = \beta_{(Y, f_Y, i)}$, $V_x = V \langle f_V \preceq x \rangle$, and $Y_x = Y \langle f_Y \preceq x \rangle$.

Lemma If (u, v) is a point of Δ^+ and if $u + \omega(\varepsilon) \prec v - \omega(\varepsilon)$, then

$$\beta_V(u - \omega(\varepsilon), v + \omega(\varepsilon)) \leq \beta_Y(u, v) \leq \beta_V(u + \omega(\varepsilon), v - \omega(\varepsilon))$$

Proof. For $\omega = \omega(\varepsilon)$ such that $u + \omega \prec v - \omega$ we can consider the following diagram

$$(1) \quad \begin{array}{ccccccc} V_{u-\omega} & \xrightarrow{\alpha} & V_{u+\omega} & \xrightarrow{\gamma} & V_{v-\omega} & \xrightarrow{\delta} & V_{v+\omega} \\ & \searrow \bar{r} & \nearrow \theta & & \searrow \hat{r} & \nearrow \iota & \\ & & Y_u & \xrightarrow{\eta} & Y_v & & \end{array}$$

where the maps $\alpha, \gamma, \delta, \eta, \theta, \iota$ are inclusions, \bar{r} is the composition of r_u , where r_u is as in the proof of Lemma 3 with $x = u$, with the inclusion i_u of $V_{u-\omega}$ into $V \langle f_Y \circ r \preceq u \rangle$, and \hat{r} is similarly defined as the composition of r_v with the inclusion i_v of $V_{v-\omega}$ into $V \langle f_Y \circ r \preceq v \rangle$. Let us observe that, passing to homology and using the symbol $*$ to denote the induced homomorphisms, the dimension of the image

of $\delta_* \circ \gamma_* \circ \alpha_*$ is precisely $\beta_V(u - \omega, v + \omega)$, that of η_* is $\beta_Y(u, v)$, and that of γ_* is $\beta_V(u + \omega, v - \omega)$.

The first claimed inequality will follow from the commutativity of the large trapezoid in diagram (1) up to homotopy. Indeed, if $\delta \circ \gamma \circ \alpha$ is homotopic to $\iota \circ \eta \circ \bar{r}$, then, passing to homology, the map $\delta_* \circ \gamma_* \circ \alpha_*$ is equal to the map $\iota_* \circ \eta_* \circ \bar{r}_*$. As a consequence the dimension of the image of $\delta_* \circ \gamma_* \circ \alpha_*$ is not greater than that of η_* . Moreover, the second claimed inequality will follow from the commutativity of the small trapezoid because, in this case, passing to homology we have $\eta_* = \hat{r}_* \circ \gamma_* \circ \theta_*$. Therefore, the dimension of the image of η_* is not greater than that of γ_* .

We begin proving that the small trapezoid commutes exactly. We observe that \hat{r} is the identity map on the points of Y . Since $Y_u \subseteq Y$, we have that $\hat{r} \circ \gamma \circ \theta$ is the canonical inclusion of Y_u in Y_v .

To prove the commutativity of the large trapezoid up to homotopy, using the exact commutativity of the small trapezoid, it is sufficient to prove that the two triangles in diagram (1) commute up to homotopy. As for the left triangle, consider $\bar{G} : V_{u-\omega} \times I \rightarrow V_{u+\omega}$ be the composition $\bar{G} = G_u \circ (i_u \times 1_I)$, where $G_{\bar{u}}$ is as in the proof of Lemma 3. Now, for every $v \in V_{u-\omega}$, we have $\bar{G}(v, 0) = G(v, 0) = v = \alpha(v)$ and $\bar{G}(v, 1) = G(v, 1) = r(v) = \bar{r}(v) = \theta \circ \bar{r}(v)$. Hence α is homotopic to $\theta \circ \bar{r}$. The proof that δ is homotopic to $\hat{r} \circ \iota$ is analogous. \square

Remark The diagram of the previous proof, setting $v = u + \omega$, shows that the persistence modules of Y and V are ω -interleaved [21] (strongly ω -interleaved, in the terminology of [6]), so that their interleaving distance is $\geq \omega$.

Lemma If (u, v) is a point of Δ^+ such that $u + \omega(\varepsilon) \prec v - \omega(\varepsilon)$ and

$$\beta_V(u - \omega(\varepsilon), v + \omega(\varepsilon)) = \beta_V(u + \omega(\varepsilon), v - \omega(\varepsilon))$$

then

$$\beta_Y(u, v) = \beta_V(u, v)$$

Proof. Straightforward from Lemma 3 and from the fact that β_V is non-decreasing in the first variable and non-increasing in the second one. \square

The next Lemma shows that there may be whole regions (hyperparallelepipeds) where the PBN's of Y and V are known to agree.

Lemma Let $u, \bar{u}, u', v, \bar{v}, v' \in \mathbb{R}^n$ be such that $u \preceq \bar{u} - \omega(\varepsilon)$, $\bar{u} + \omega(\varepsilon) \preceq u' \prec v \preceq \bar{v} - \omega(\varepsilon)$, $\bar{v} + \omega(\varepsilon) \preceq v'$. If

$$\beta_V(u, v') = \beta_V(u', v)$$

then

$$\beta_Y(\bar{u}, \bar{v}) = \beta_V(\bar{u}, \bar{v})$$

Proof. Since β_V is non-decreasing in the first variable and non-increasing in the second one, we have, with $\omega = \omega(\varepsilon)$,

$$\begin{aligned} \beta_V(u, v') &\leq \beta_V(u, \bar{v} + \omega) \leq \beta_V(\bar{u} - \omega, \bar{v} + \omega) \leq \\ &\leq \beta_V(\bar{u} + \omega, \bar{v} - \omega) \leq \beta_V(\bar{u} + \omega, v) \leq \beta_V(u', v) \end{aligned}$$

whence, by the hypothesis $\beta_V(u, v') = \beta_V(u', v)$, we get $\beta_V(\bar{u} - \omega, \bar{v} + \omega) = \beta_V(\bar{u} + \omega, \bar{v} - \omega)$. Lemma 3 then yields the thesis. \square

As a consequence of the preceding lemmas, the regions of Δ^+ where β_V and β_Y may possibly disagree can be precisely localized in a neighborhood of the discontinuity set of β_V (viewed as an integer function). In the case $n = 1$ where the discontinuity sets are proved to be (possibly infinite) line segments [17], these regions are called *blind strips*. We stress that the position of the blind strips is well-known, since it is determined by the position of the discontinuity lines of the PBNs of V .

Lemma If $\bar{u}, \bar{v} \in \mathbb{R}^n$ are such that

$$\beta_Y(\bar{u}, \bar{v}) \neq \beta_V(\bar{u}, \bar{v})$$

then there is at least a point (\tilde{u}, \tilde{v}) with max norm $\|(\bar{u}, \bar{v}) - (\tilde{u}, \tilde{v})\| \leq \Omega(\varepsilon)$, which is either an element of the boundary of Δ^+ or a discontinuity point of β_V .

Proof. We prove the contrapositive.

Let $(\bar{u}, \bar{v}) \in \mathbb{R}^n$ be a point of Δ^+ . Let D be the closed ball of radius Ω centered at (\bar{u}, \bar{v}) . Set $\omega = \omega(\varepsilon)$, $u = \bar{u} - \omega$, $u' = \bar{u} + \omega$, $v = \bar{v} - \omega$, $v' = \bar{v} + \omega$; D is the closed hypercube $\{(\hat{u}, \hat{v}) \in \mathbb{R}^n \mid (u, v) \preceq (\hat{u}, \hat{v}) \preceq (u', v')\}$ and its boundary contains (u, v') and (u', v) .

Now, assume that (\bar{u}, \bar{v}) is at distance greater than Ω from the boundary of Δ^+ (with the max norm distance). Then the point (u', v) , belonging to D , is contained in Δ^+ . This means that $u' \prec v$.

Moreover, assume that (\bar{u}, \bar{v}) is at distance greater than Ω from any discontinuity point of β_V . Then there is an open subset of Δ^+ , containing D , on whose points β_V is continuous. Since the range of β_V has the discrete topology, this implies that β_V is constant on this whole open set, hence also on D . Then $\beta_V(u, v') = \beta_V(u', v)$; therefore $\beta_Y(\bar{u}, \bar{v}) = \beta_V(\bar{u}, \bar{v})$ by Lemma 3. \square

4. BALL COVERINGS

Throughout this Section, X will be a compact Riemannian (triangulable) submanifold of \mathbb{R}^m . As hinted in Section 1, 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, 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 Lemma 3 even offers the exact value of it at points sufficiently far from the discontinuity sets.

4.1. Points on X . 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 , $j = 1, \dots, k$. So all conditions of Proposition 2.2 are satisfied. As before, set $\beta_U = \beta_{(U, f_U, i)}$ and $\beta_X = \beta_{(X, f_X, i)}$.

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

$$\beta_U(u - \omega(\delta), v + \omega(\delta)) \leq \beta_X(u, v) \leq \beta_U(u + \omega(\delta), v - \omega(\delta))$$

If $\bar{u}, \bar{v} \in \mathbb{R}^n$ are such that

$$\beta_X(\bar{u}, \bar{v}) \neq \beta_U(\bar{u}, \bar{v})$$

then there is at least a point (\tilde{u}, \tilde{v}) with max norm $\|(\bar{u}, \bar{v}) - (\tilde{u}, \tilde{v})\| \leq \Omega(\delta)$, which is either an element of the boundary of Δ^+ or a discontinuity point of β_U .

Proof. By Lemmas 3 and 3, with $Y = X$, $V = U$. \square

4.2. Points near X . 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 [22].

Proposition (Prop. 7.1 of [22]) 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 point $p \in X$, there exist 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 $\delta \in \left(\frac{(s+\tau) - \sqrt{s^2 + \tau^2 - 6s\tau}}{2}, \frac{(s+\tau) + \sqrt{s^2 + \tau^2 - 6s\tau}}{2} \right)$.

Then, as with Theorem 4.1, we have, with an analogous proof:

Theorem Under the hypotheses of Proposition 4.2, if (u, v) is a point of Δ^+ and if $u + \omega(\delta + s) \prec v - \omega(\delta + s)$, where $\omega(\delta + s) = (\Omega(\delta + s), \dots, \Omega(\delta + s)) \in \mathbb{R}^n$, then

$$\beta_U(u - \omega(\delta + s), v + \omega(\delta + s)) \leq \beta_X(u, v) \leq \beta_U(u + \omega(\delta + s), v - \omega(\delta + s))$$

If $\bar{u}, \bar{v} \in \mathbb{R}^n$ are such that

$$\beta_X(\bar{u}, \bar{v}) \neq \beta_U(\bar{u}, \bar{v})$$

then there is at least a point (\tilde{u}, \tilde{v}) with max norm $\|(\bar{u}, \bar{v}) - (\tilde{u}, \tilde{v})\| \leq \Omega(\delta + s)$, which is either an element of the boundary of Δ^+ or a discontinuity point of β_U .

4.3. The 1D case. We now show how Theorem 4.1 can be used for applications. In the case of filtering functions with one-dimensional range, and in the 1D reduction of Section 2.1 of [2] and Section 4 of [5], the subsets of \mathbb{R}^2 where the PBN functions are not continuous form (possibly infinite) line segments: [17]. We recall that by Lemma 3, the *blind strips*, i.e. the regions where the equality $\beta_Y(\bar{u}, \bar{v}) = \beta_V(\bar{u}, \bar{v})$ is not granted, are $2\omega(\varepsilon)$ wide strips around such segments.

In the case of ball coverings as in Section 4.1, the width of the blind 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.

Let X be a circle of radius 4 in \mathbb{R}^2 (Figure 1); 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 hypothesis of Theorem 4.1 (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 2). We stick to the monodimensional case, choosing $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, with $f(x, y) = \text{abs}(y)$. U is the resulting ball union.

Figures 3 and 4 represent the PBN functions at degree zero of X and U respectively. Δ^+ is the half-plane above the diagonal line, and the numbers are the values of the PBNs in the triangular regions they are written in. In Figure 3 there is only one big triangle where the value 2 signals the two different connected components generated by f_X . The two connected components collapse to one at value 4. In

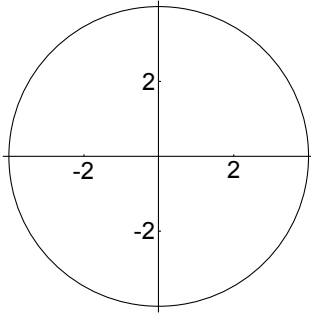


FIGURE 1. The circle of radius 4, X .

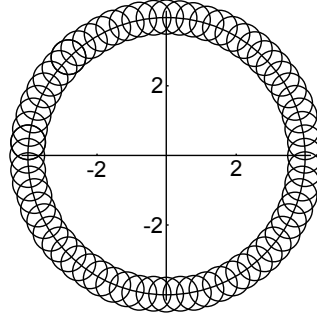


FIGURE 2. The ball union U .

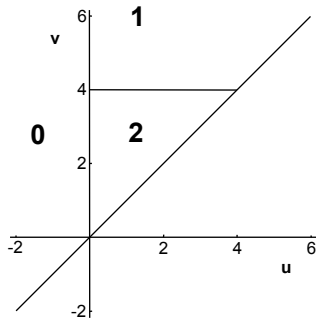


FIGURE 3. The representation of $\beta_{(X, f_X, 0)}$, the 0-PBNs of X .

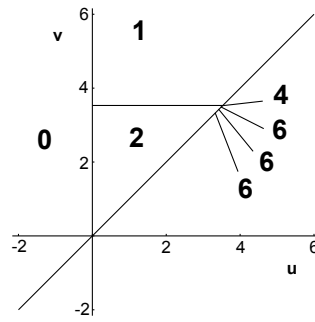


FIGURE 4. The representation of $\beta_{(U, f_U, 0)}$, the 0-PBNs of the ball union U .

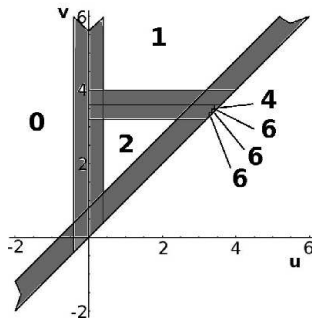


FIGURE 5. The blind strips of $\beta_{(U, f_U, 0)}$.

Figure 4 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 approximation error. In the last figure (Figure 5) the blind strips around the discontinuity lines of $\beta_{(U, f_U, 0)}$ 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 4.1. 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. Figures 6, 7, 8 depict, in analogous way, the (obviously much simpler) PBNs of degree 1.

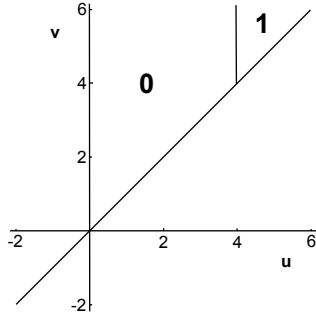


FIGURE 6. The representation of $\beta_{(X, f_X, 1)}$, the 1-PBNs of X .

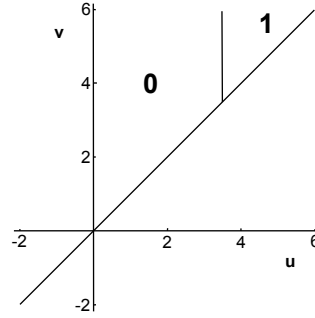


FIGURE 7. The representation of $\beta_{(U, f_U, 1)}$, the 1-PBNs of the ball union U .

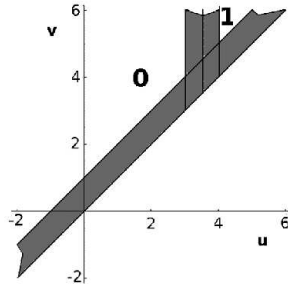


FIGURE 8. The blind strips of $\beta_{(U, f_U, 1)}$.

For a second example we have chosen the points l_j not necessarily on X . We have satisfied the hypothesis of Proposition 4.2, choosing $s = 0.25$ and $\delta = 0.55$. Then, in order to cover X well, we have chosen a point every $\frac{\pi}{48}$ radians, for a total of 96 points. But this time the points are either 0 or 0.1 or 0.2 away from X . Figure 9 shows the resulting ball union U' . As in the previous case, in the representation of $\beta_{(U', f_{U'}, 0)}$ (Figure 10) there is one big triangle showing two connected components and this time they collapse at value 3.40955. Compared to Figure 4, there are many more small triangles generated by the asymmetry of the sampling. The width of

the blind strips in Figure 11 is $2\Omega(\delta + s) = 1.6$, so there is still the central triangle. This means that, although the error in the approximation is much bigger, the blind strips do not cover the entire figure, leaving the topological information intact at least in some small areas of Δ^+ .

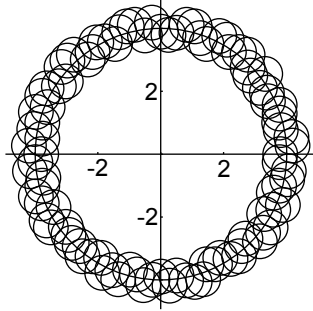


FIGURE 9. The ball union U' .

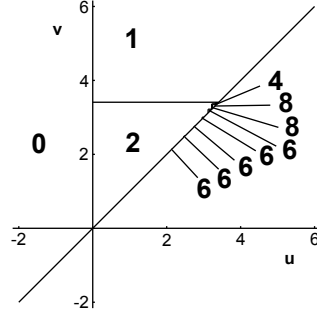


FIGURE 10. The representation of $\beta_{(U', f_{U'}, 0)}$, the 0-PBNs of the ball union U' .

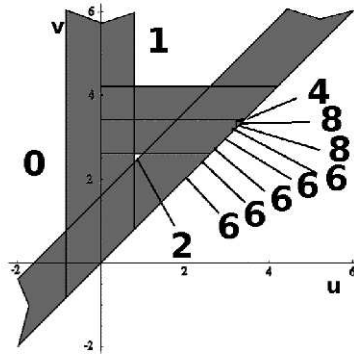


FIGURE 11. The blind strips of $\beta_{(U', f_{U'}, 0)}$.

5. SHAPE COMPARISON

So, outside the blind strips, we can get the values of PBNs of the sampled object X out of its ball covering U (as always, given a filtering function f defined on the ambient space \mathbb{R}^m). This fact gives us the possibility of assessing shape dissimilarity of two pairs (X, f) , (Y, g) although we only know ball coverings U , W of them, respectively. For this we need the notion of natural pseudodistance, introduced in [10] and further studied in [11, 12], and its relationship with PBNs.

We recall that, for any two topological spaces X, Y endowed with two continuous functions $f : X \rightarrow \mathbb{R}^n$, $g : Y \rightarrow \mathbb{R}^n$, we have the following definition.

Definition The *natural pseudodistance* between the pairs (X, f) and (Y, g) , denoted by $\delta((X, f), (Y, g))$, is

- (i) the number $\inf_h \max_{x \in X} \|f(x) - g(h(x))\|$ where h varies in the set $H(X, Y)$ of all the homeomorphisms between X and Y , if X and Y are homeomorphic;
- (ii) $+\infty$, if X and Y are not homeomorphic.

5.1. Lower bounds for the natural pseudodistance. The following theorem, which extends Theorem 1 of [10], provides us with a lower bound for the natural pseudodistance.

Theorem Let (X, f) and (Y, g) be two pairs. If, for some degree i , $\beta_{(X, f, i)}(u, v) > \beta_{(Y, g, i)}(u', v')$ then

$$\delta((X, f), (Y, g)) \geq \min \left\{ \min_r \{u'_r - u_r\}, \min_r \{v_r - v'_r\} \right\}$$

Proof. Set $\xi = \min \{ \min_r \{u'_r - u_r\}, \min_r \{v_r - v'_r\} \}$. If $\xi \leq 0$ then the thesis trivially holds; so assume $\xi > 0$. Since $\beta_{(X, f, i)}(u, v) > 0$ we have $X\langle f \preceq u \rangle \neq \emptyset$. Assume that the thesis does not hold, so $\delta((X, f), (Y, g)) < \xi$. Under this assumption there exists a homeomorphism $h : X \rightarrow Y$ such that

$$\max_{x \in X} \|f(x) - g(h(x))\| < \xi.$$

Set $\Theta(h) = \max_{x \in X} \|f(x) - g(h(x))\|$. So $\Theta(h) < \xi$, and, setting $\vec{\zeta} = (\xi, \dots, \xi)$ we have that h maps the lower level subset $X\langle f \preceq u \rangle$ into $Y\langle g \preceq u + \vec{\zeta} \rangle \subset Y\langle g \preceq u' \rangle$. Analogously, h^{-1} maps $Y\langle g \preceq v' \rangle$ into $X\langle f \preceq v' + \vec{\zeta} \rangle \subset X\langle f \preceq v \rangle$. Set $X_u = X\langle f \preceq u \rangle$, $X_{u'} = h^{-1}(Y\langle g \preceq u' \rangle)$, $X_{v'} = h^{-1}(Y\langle g \preceq v' \rangle)$, $X_v = X\langle f \preceq v \rangle$. Then there are inclusion maps

$$\iota_{u, u'} : X_u \hookrightarrow X_{u'}, \quad \iota_{u', v'} : X_{u'} \hookrightarrow X_{v'}, \quad \iota_{v', v} : X_{v'} \hookrightarrow X_v$$

and $\iota_{u, v} : X_u \hookrightarrow X_v$, with $\iota_{u, v} = \iota_{v', v} \circ \iota_{u', v'} \circ \iota_{u, u'}$.

The pairs $(Y\langle g \preceq v' \rangle, Y\langle g \preceq u' \rangle)$ and $(X_{v'}, X_{u'})$ are homeomorphic, so they can be interchanged in what follows. Our claim is that for each degree i the dimension of the image of the homomorphism induced in homology by $\iota_{u, v}$ is less than or equal to the one of the image of the homomorphism induced by $\iota_{u', v'}$. First, note that all dimensions considered here are finite (Theorem 2.3 of [5]). Then, each inclusion map ι induces a homology homomorphism ι_* , and $(\iota_{u, v})_* = (\iota_{v', v})_* \circ (\iota_{u', v'})_* \circ (\iota_{u, u'})_*$. But then $\dim \text{Im}(\iota_{u, v})_* \leq \dim \text{Im}(\iota_{u', v'})_*$, so $\beta_{(X, f, i)}(u, v) \leq \dim \text{Im}(\iota_{u', v'})_* = \beta_{(Y, g, i)}(u', v')$ against the hypothesis. \square

5.2. An example of comparison of sampled shapes. When only finite, dense enough samples - or, equivalently, ball coverings - of two objects are available, if there is a nonempty intersection of the complements of the blind strips, we can still assess the natural pseudodistance between them.

This is the case of the following example: We obtain a lower bound for $\delta(X, Y)$, where X is the circle and Y is the bean-shaped curve of Figure 12 and Figure 13.

For both spaces the filtering functions f and g are the restrictions of the absolute value of ordinate. The 0-PBNs for (X, f) and (Y, g) are depicted in Figures 14 and 15 respectively.

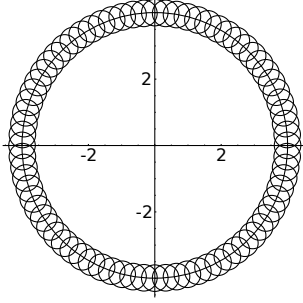


FIGURE 12. A circle X covered with balls of radius 0.4.

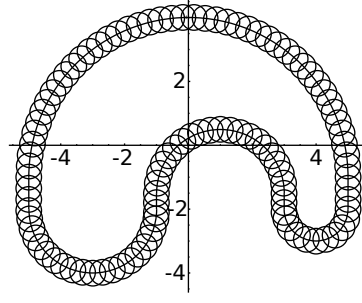


FIGURE 13. A different curve Y covered with balls of radius 0.4.

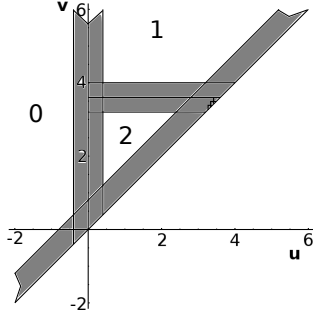


FIGURE 14. The 0-PBNs of the ball covering of X , with blind strips.

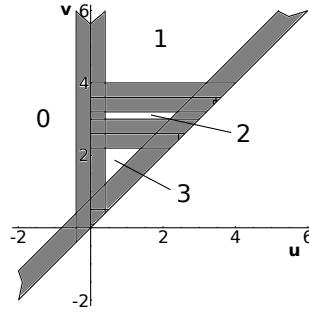


FIGURE 15. The 0-PBNs of the ball covering of Y , with blind strips.

There is a triangle, not covered by blind strips in both diagrams, where the PBN for X is 2 and for Y is 3 by Theorem 4.1. So we have

$$\beta_{(Y,g,0)}(0.4, 2.2) = 3 > 2 = \beta_{(X,f,0)}(1.1, 1.5)$$

and, by Theorem 5.1, $\delta((X, f), (Y, g)) \geq 0.7$.

If we made use of smaller - but denser - balls of radius 0.2, then we would have

$$\beta_{(Y,g,0)}(0.2, 2.6) = 3 > 2 = \beta_{(X,f,0)}(1.3, 1.5)$$

and $\delta((X, f), (Y, g)) \geq 1.1$. (The true natural pseudodistance of the two pairs can be computed directly from the definition, and equals 1.5.)

6. A COMBINATORIAL REPRESENTATION

The ball unions of Section 4, although generated by finite sets, are still continuous objects. It is desirable that the topological information on X , up to a certain approximation, be condensed in a combinatorial object. For size functions (i.e. for PBNs of degree 0) it was a graph; here, it has to be a simplicial complex. We shall build such a complex, by following [13], to which we refer for all definitions not reported here. Please note that [13] uses *weighted* Voronoi cells and diagrams, while we do not need to worry about that, since all of our balls have the same radius;

so the customary Euclidean distance can be used instead of the power distance employed in that paper.

Let X , $L = \{l_1, \dots, l_k\}$ and δ be as in Section 4.1 (the case of Section 4.2 is an immediate extension). Moreover, let the points of L be in general position. For each $l_j \in L$, let $B_j = B(l_j, \delta)$ be the ball of radius δ , centered at l_j . The set $B = \{B_1, \dots, B_k\}$ is a ball covering of X ; denote by U the corresponding ball union. Let now V_j be the *Voronoi cell* of B_j , i.e. the set of points of \mathbb{R}^m whose distance from l_j is not greater than the distance from any other $l_{j'}$.

The set $\mathcal{V} = \{V_1, \dots, V_k\}$ is the *Voronoi diagram* of B . From \mathcal{V} we get the collection of cells $\mathcal{Q} = \{V'_j = V_j \cap B_j \mid j = 1, \dots, k\}$, a decomposition of U .

The *nerve* $N(\mathcal{Q})$ of \mathcal{Q} is the abstract simplicial complex where vertices are the elements of \mathcal{Q} and, for a subset T of $\{1, \dots, k\}$, the set of vertices $\{V'_j \mid j \in T\}$ is a simplex if and only if $\bigcap_{j \in T} V'_j \neq \emptyset$.

For any $T \subseteq \{1, \dots, k\}, T \neq \emptyset$ we denote by σ_T the convex hull of $\{l_j \mid j \in T\}$

The *dual complex* of \mathcal{Q} is $\mathcal{K} = \{\sigma_T \mid \{V'_j \mid j \in T\} \in N(\mathcal{Q})\}$ and $\mathcal{S} = |\mathcal{K}|$, union of the simplices of \mathcal{K} , is the *dual shape* of U .

For a better understanding of the previous part we produce a toy example. Let X be a quarter of circle of radius 4 and U be the union of nine balls of radius 1, with centers near X (Figure 16). The Voronoi Diagram \mathcal{V} associated to this ball covering B is depicted in Figure 17.

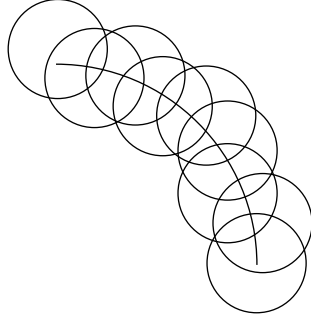


FIGURE 16. A quarter of circle of radius 4 covered by nine balls of radius 1.

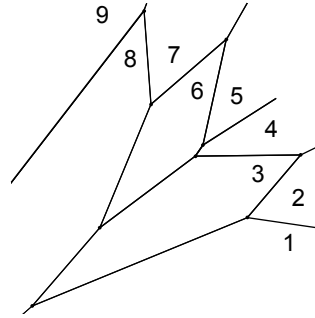
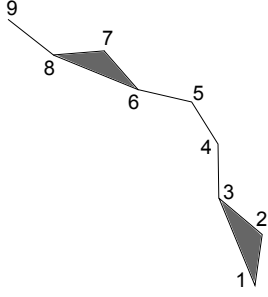
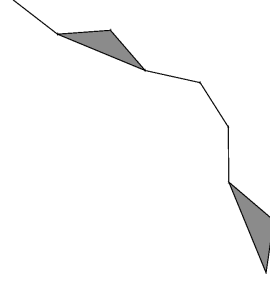


FIGURE 17. The Voronoi Diagram \mathcal{V} of B .

Now the main idea is that we can associate the dual complex \mathcal{K} with the submanifold X . In fact, by Theorem 3.2 of [13], its space \mathcal{S} is homotopically equivalent to U and, by transitivity, to X . Moreover, Section 3 of [13] explicitly builds a retraction r from U to \mathcal{S} and a homotopy H from the identity of U , to p , such that $\forall y \in \mathcal{S}, \forall v \in p^{-1}(y), \forall t \in I$ we have $(p \circ H)(v, t) = y$. For a complete description of the homotopy H and the retraction p we refer to the original article.

\mathcal{K} and \mathcal{S} are shown in Figures 18, 19 respectively.

6.1. Ball union and dual shape. Let $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ be a continuous function and let $f_{\mathcal{S}}$ and f_U be the restrictions of f to \mathcal{S} and U respectively. The notation for the PBN's β is simplified as before.

FIGURE 18. The dual complex \mathcal{K} .FIGURE 19. The dual shape \mathcal{S} .

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

$$\beta_U(u - \omega(\delta), v + \omega(\delta)) \leq \beta_S(u, v) \leq \beta_U(u + \omega(\delta), v - \omega(\delta)).$$

If $\bar{u}, \bar{v} \in \mathbb{R}^n$ are such that

$$\beta_U(\bar{u}, \bar{v}) \neq \beta_S(\bar{u}, \bar{v})$$

then there is at least a point (\tilde{u}, \tilde{v}) with max norm $\|(\bar{u}, \bar{v}) - (\tilde{u}, \tilde{v})\| \leq \Omega(\delta)$, which is either an element of the boundary of Δ^+ or a discontinuity point of β_U .

Proof. By Lemma 3, with $Y = \mathcal{S}$, $V = U$. □

Now we can get an estimate of the PBNs of X from the ones of \mathcal{S} . The blind strips of the 1D reduction will be doubly wide, with respect to the ones previously considered. Still, this can leave some regions of Δ^+ where the computation is exact. Also here the position of the blind strips is well determined by the position of the discontinuity lines of the PBNs of \mathcal{S} .

Theorem If (u, v) is a point of Δ^+ and if $u + 2\omega(\delta) \prec v - 2\omega(\delta)$, where $\omega(\delta) = (\Omega(\delta), \dots, \Omega(\delta)) \in \mathbb{R}^n$, then

$$\beta_S(u - 2\omega(\delta), v + 2\omega(\delta)) \leq \beta_X(u, v) \leq \beta_S(u + 2\omega(\delta), v - 2\omega(\delta)).$$

If $\bar{u}, \bar{v} \in \mathbb{R}^n$ are such that

$$\beta_X(\bar{u}, \bar{v}) \neq \beta_S(\bar{u}, \bar{v})$$

then there is at least a point (\tilde{u}, \tilde{v}) with max norm $\|(\bar{u}, \bar{v}) - (\tilde{u}, \tilde{v})\| \leq 2\Omega(\delta)$, which is either an element of the boundary of Δ^+ or a discontinuity point of β_S .

Proof. By Theorem 4.1, and with $\omega = \omega(\delta)$,

$$\beta_U(u - \omega, v + \omega) \leq \beta_X(u, v) \leq \beta_U(u + \omega, v - \omega)$$

Then we have

$$\beta_U(u + \omega, v - \omega) \leq \beta_S(u + 2\omega, v - 2\omega)$$

by Lemma 6.1 by substituting (u, v) with $(u + 2\omega, v - 2\omega)$, and

$$\beta_S(u - 2\omega, v + 2\omega) \leq \beta_U(u - \omega, v + \omega)$$

by Lemma 6.1 by substituting (u, v) with $(u - 2\omega, v + 2\omega)$.

The second part can be proved in an analogous way from the second part of Lemma 6.1. □

6.2. An example in 2D persistence. As a simple example, we now apply Theorems 5.1 and 6.1 to two pairs (X, f) and (Y, g) with 2-dimensional filtering functions. Let X and Y be two circles of radius r embedded in \mathbb{R}^2 and f, g be (unknown) continuous functions from \mathbb{R}^2 to \mathbb{R}^2 , whose restrictions to suitable neighbourhoods of X, Y have modulus of continuity $\Omega(\delta) = \frac{\delta}{2r}$. Here the range \mathbb{R}^2 parametrizes the plane Π of equation $R + G + B = \frac{4-\sqrt{2}}{2}$ (chosen as one which contains a fairly large square of points with nonnegative coordinates) of the color space RGB . A Cartesian reference frame x, y, z has been fixed in the RGB space, so that Π is the $z = 0$ plane. The change of reference is:

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \sqrt{\frac{2}{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} \frac{4-\sqrt{2}}{4} \\ \frac{4-\sqrt{2}}{4} \\ 0 \end{pmatrix}$$

Of the functions on two circles we know a sampling given by 80 regularly spaced points, represented in Figure 20.



FIGURE 20. The color samplings of (X, f) (left) and (Y, g) (right).

We can think of the 80 points on each circle as centers of disks of radius $\delta = 0.08r$. The distance between two consecutive points is $2r\pi \sin(\frac{\pi}{80}) \in]0, 078r, 0.079r[$, so the balls form a covering which respects the hypotheses of Proposition 2.2 since $\tau = r$ and $0.08 < \sqrt{\frac{3}{5}}$. The corresponding dual complex is a 1D cycle and the dual shape is a closed polygonal.

The point images in the color plane Π are along the edges of the square of vertices $(-0.4, 0), (-0.4, 0.8), (0.4, 0.8), (0.4, 0)$ for (X, f) and along the polygonal of vertices $(-0.4, 0.04), (0.36, 0.8), (0.4, 0.8), (0.4, 0.76), (-0.36, 0)$ (twice) for (Y, g) , two consecutive ones at distance $0.04 = \Omega(\delta)$ in the max norm.

A 1D reduction of the PBNs is possible through a foliation of the domain $\mathbb{R}^2 \times \mathbb{R}^2$. In the terminology of [1, 2], a suitable admissible pair \vec{l}, \vec{b} is given by $\vec{l} = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ and $\vec{b} = (-0.2, 0.2)$; in the corresponding leaf of the foliation, the PBN's of the dual shapes and the blind strips according to Theorem 6.1 are as shown in Figure 21.

In correspondence of points P, Q in the leaf, we get $(u, v), (u', v') \in \mathbb{R}^2 \times \mathbb{R}^2$ respectively, with $u = (-0.28, 0.12)$, $v = (0.32, 0.72)$, $u' = (-0.06, 0.34)$, $v' = (0.1, 0.5)$. Then we have

$$\beta_{(Y, g, 0)}(u, v) = 2 > 1 = \beta_{(X, f, 0)}(u', v')$$

and, by Theorem 5.1, $\delta((X, f), (Y, g)) \geq 0.22$.

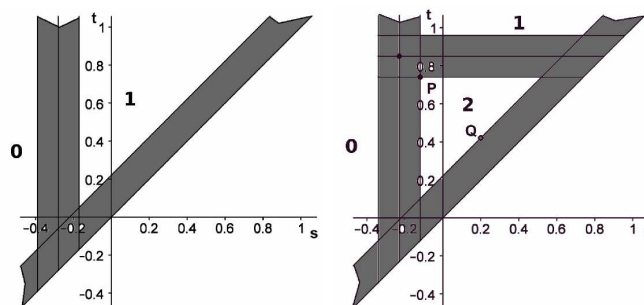


FIGURE 21. PBNs of the two colored circles

Acknowledgements. The authors wish to thank P. Frosini for the many helpful suggestions. This work was performed under the auspices of INdAM-GNSAGA and ARCES.

REFERENCES

- [1] S. Biasotti, A. Cerri, P. Frosini, D. Giorgi, and C. Landi. Multidimensional size functions for shape comparison. *J. Math. Imaging Vision*, 32(2):161–179, 2008.
- [2] F. Cagliari, B. Di Fabio, and M. Ferri. One-dimensional reduction of multidimensional persistent homology. *Proc. Amer. Math. Soc.*, 138(8):3003–3017, 2010.
- [3] G. Carlsson. Topology and data. *Bull. Amer. Math. Soc.*, 46(2):255–308, 2009.
- [4] G. Carlsson and A. Zomorodian. The theory of multidimensional persistence. *Discrete Comput. Geometry*, 42(1):71–93, 2009.
- [5] A. Cerri, B. Di Fabio, M. Ferri, P. Frosini, and C. Landi. Betti numbers in multidimensional persistent homology are stable functions. *Mathematical Methods in the Applied Sciences*, 36(12):1543–1557, 2013.
- [6] F. Chazal, D. Cohen-Steiner, M. Glisse, L. J. Guibas, and S. Y. Oudot. Proximity of persistence modules and their diagrams. In *SCG '09: Proceedings of the 25th annual symposium on Computational geometry*, pages 237–246, New York, NY, USA, 2009. ACM.
- [7] F. Chazal and A. Lieutier. Weak feature size and persistent homology: computing homology of solids in \mathbb{R}^n from noisy data samples. In *Proceedings of the twenty-first annual Symposium on Computational Geometry*, SCG '05, pages 255–262, New York, NY, USA, 2005. ACM.
- [8] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Stability of persistence diagrams. *Discrete Comput. Geom.*, 37(1):103–120, 2007.
- [9] M. d’Amico. A new optimal algorithm for computing size function of shapes. In *CVPRIP Algorithms III, Proceedings International Conference on Computer Vision, Pattern Recognition and Image Processing*, pages 107–110, 2000.
- [10] P. Donatini and P. Frosini. Lower bounds for natural pseudodistances via size functions. *Arch. Inequal. Appl.*, 2:1–12, 2004.
- [11] P. Donatini and P. Frosini. Natural pseudodistances between closed surfaces. *J. Europ. Math. Soc.*, 9(2):231–253, 2007.
- [12] P. Donatini and P. Frosini. Natural pseudodistances between closed curves. *Forum Math.*, 21(6):981–999, 2009.
- [13] H. Edelsbrunner. The union of balls and its dual shape. *Discrete Comput. Geom.*, 13:415–440, 1995.
- [14] H. Edelsbrunner and J. 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.
- [15] S. Eilenberg and N. Steenrod. *Foundations of algebraic topology*. Princeton University Press, Princeton, New Jersey, 1952.
- [16] P. Frosini. Discrete computation of size functions. *J. of Combin., Inf. System Sci.*, 17(3-4):232–250, 1992.

- [17] P. Frosini and C. Landi. Size functions and formal series. *Appl. Algebra Engrg. Comm. Comput.*, 12(4):327–349, 2001.
- [18] R. Ghrist. Barcodes: the persistent topology of data. *Bull. Amer. Math. Soc. (N.S.)*, 45(1):61–75 (electronic), 2008.
- [19] M. W. Hirsch. *Differential Topology*, volume 33 of *Graduate Text in Mathematics*. Springer, 1976.
- [20] J. M. Lee. *Introduction to Smooth Manifolds*, volume 218 of *Grad. Texts in Math*. Springer-Verlag, New York, 2003.
- [21] M. Lesnick. The theory of the interleaving distance on multidimensional persistence modules. *Foundations of Computational Mathematics*, pages 1–38, 2015.
- [22] P. Niyogi, S. Smale, and S. Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete Comput. Geom.*, 39(1):419–441, 2008.

NICCOLO.CAVAZZA@GMAIL.COM

DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DI BOLOGNA, ITALIA, MASSIMO.FERRI@UNIBO.IT

DIPARTIMENTO DI SCIENZE E METODI DELL'INGEGNERIA, UNIVERSITÀ DI MODENA E REGGIO EMILIA, ITALIA, CLANDI@UNIMORE.IT