STABLE COMPARISON OF MULTIDIMENSIONAL PERSISTENT HOMOLOGY GROUPS WITH TORSION

PATRIZIO FROSINI

This paper is dedicated to the memory of Padre Armando Felice Verde. G.g.

ABSTRACT. The present lack of a stable method to compare persistent homology groups with torsion is a relevant problem in current research about Persistent Homology and its applications in Pattern Recognition. In this paper we introduce a pseudo-distance d_T that represents a possible solution to this problem. Indeed, d_T is a pseudo-distance between multidimensional persistent homology groups with coefficients in an Abelian group, hence possibly having torsion. Our main theorem proves the stability of the new pseudo-distance with respect to the change of the filtering function, expressed both with respect to the max-norm and to the natural pseudo-distance between topological spaces endowed with $\mathbb{R}^n$-valued filtering functions. Furthermore, we prove a result showing the relationship between d_T and the matching distance in the 1-dimensional case, when the homology coefficients are taken in a field and hence the comparison can be made.

INTRODUCTION

The extension of the theory of Persistent Homology to the case of persistent homology groups with torsion is an important open problem. Indeed, the lack of a complete representation similar to the one given by persistent diagrams is a relevant obstacle to this development. However, this fact does not prevent us from using persistent homology groups with torsion for shape comparison. In order to show this, we introduce here a new pseudo-distance d_T between persistent homology groups with coefficients in an Abelian group (hence possibly having torsion), proving that d_T is stable with respect to changes of the filtering functions, measured by the max-norm. We also express this stability as a lower bound for the natural pseudo-distance between topological spaces endowed with $\mathbb{R}^n$-valued filtering functions. This approach opens the way to the use of persistent homology groups in concrete applications, extending some ideas developed in [22] for size homotopy groups to persistent homology.

Now, let us illustrate how persistent homology groups with torsion can appear in applications. As an example, let us imagine to have to study all the possible ways of grasping an object by a robotic hand endowed with m fingers. Obviously, not all grasps can be considered equivalent. For example, some of them can be difficult to realize in practice, due to the position of the object and the technical capability of the hand. Therefore, it is reasonable to associate each grasp with a "cost", represented by a real number. In one of the simplest cases, we can consider two fingers grasping an object whose surface $\mathcal{S}$ is diffeomorphic to a spherical surface, in presence of friction. Each grasp can be seen as an unordered pair $\{p, q\}$ of points on $\mathcal{S}$, and we can consider the set $\mathcal{G}$ of all possible grasps (endowed with the Hausdorff distance between unordered pairs). If the cost of the grasp is given by a value $\varphi_{\mathcal{G}}(\{p, q\})$, we are naturally led to the problem of comparing the similarity of the grasps possible for two different objects with surfaces $\mathcal{S}$ and $\mathcal{S}'$, with respect to the cost functions $\varphi_{\mathcal{G}} : \mathcal{G} \to \mathbb{R}$, $\varphi_{\mathcal{G}'} : \mathcal{G}' \to \mathbb{R}$. This

can be done by computing the (singular) persistent homology groups of each topological space $\mathcal{G}$ with respect to the filtering function $\varphi_{\mathcal{G}}$. In other words, we can consider the persistent homology groups $H_k^{(\mathcal{G},\varphi_{\mathcal{G}})}(u,v)$ given by the equivalence classes of cycles in $H_k(\mathcal{G}_v)$ that contain at least one cycle in $\mathcal{G}_u$, where $\mathcal{G}_t$ represents the set of points of $\mathcal{G}$ at which $\varphi_{\mathcal{G}}$ takes a value not greater than t. It is interesting to note that, in general, this procedure leads to groups with torsion, when the homology coefficients are the integer numbers.

To understand this, let us consider a particularly simple example. Let us assume that the cost of the grasp is given by the value $\varphi_{\mathcal{G}}(\{p,q\}) = -\|p-q\|$, so relating the cost of a grasp to the level of closure of the robotic hand. For the sake of simplicity, let us examine the case $\mathcal{S} = S^2 = \{(x,y,z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$. In this case, the group $H_1^{(\mathcal{G},\varphi_{\mathcal{G}})}(u,v)$ has torsion, when $u < v < 0$ and u is close enough to zero. Indeed, we can consider the continuous function f taking each unordered pair $\{p,q\} \in \mathcal{G}_v$ to the direction of the line l through p and q (note that $\{p,q\} \in \mathcal{G}_v$ implies $p \neq q$, so that l is well-defined). This direction can be seen as a point in the real projective plane $\mathbb{RP}^2$. Let us observe that, for v close enough to zero, the function $f : \mathcal{G}_v \to \mathbb{RP}^2$ is a homotopy equivalence (the homotopy inverse of f is the function $g : \mathbb{RP}^2 \to \mathcal{G}_v$ that takes each direction w to the unordered pair $\{p,q\}$, where p and q are the points at which the line through $(0,0,0)$, having direction w, meets S^2). Hence f induces an isomorphism between the singular homology groups of $\mathcal{G}_v$ and $\mathbb{RP}^2$. As a consequence, since $H_1(\mathbb{RP}^2)$ has torsion, the persistent homology group $H_1^{(\mathcal{G},\varphi_{\mathcal{G}})}(u,v)$ has torsion, too, when u is close enough to zero. For an introduction to Topological Robotics we refer the interested reader to [21].

Multidimensional persistence. Persistent homology has turned out to be a key mathematical method for studying the topology of data, with applications in an increasing number of fields, ranging from shape description (e.g., [5, 9, 24, 25]) to data simplification [20] and hole detection in sensor networks [14]. Recent surveys on the topic include [18, 19, 23, 26]. Persistent homology describes topological events occurring through the filtration of a topological space X (e.g., creation, merging, cancellation of connected components, tunnels, voids). Filtrations are usually expressed by real functions $\varphi : X \to \mathbb{R}$ called *filtering functions*. The main idea underlying this approach is that the most important piece of information enclosed in geometrical data is usually the one that is "*persistent*" with respect to the defining parameters.

Until recently, research on persistence has mainly focused on the use of scalar functions for describing filtrations. The extent to which this theory can be generalized to a situation in which two or more functions characterize the data is currently under investigation [1, 2, 3, 4, 7]. This generalization to vector-valued functions is usually known as the *Multidimensional Persistence Theory*, where the adjective multidimensional refers to the fact that filtering functions are vector-valued, and has no connections with the dimensionality of the space under study. The use of vector-valued filtering functions in this context enables the analysis of richer data structures.

An important topic in current research about multidimensional persistent homology is the *stability problem*. In plain words, we need to determine how the computation of invariants in this theory is affected by the unavoidable presence of noise and approximation errors. Indeed, it is clear that any data acquisition is subject to perturbations and, if persistent homology were not stable, then distinct computational investigations of the same object could produce completely different results. Obviously, this would make it impossible to use such a mathematical theory in real applications.

Up to our knowledge, no theoretical result is available for the metric comparison of persistent homology groups in presence of torsion, at the time we are writing.

Prior works. The problem of stability in persistent homology has been studied by Cohen-Steiner, Edelsbrunner and Harer in [11] for scalar filtering functions. By using a descriptor called a *persistence diagram*, they prove that persistent Betti numbers are stable under perturbations of filtering functions with respect to the max-norm, provided that the considered filtering functions are *tame*. The same

problem is studied in [12] for tame Lipschitz functions. In [10], Chazal *et al.* use the concept of *persistence module* and prove stability under the assumption that it is finite-dimensional. The problem of stability for scalar filtering functions is also approached in [13], where it is solved by assuming that the considered filtering functions are no more than continuous, but only for the 0th homology.

Multidimensional persistence was firstly investigated in [22] as regards homotopy groups, and by Carlsson and Zomorodian in [6] as regards homology modules. In this context, the first stability result has been obtained for the 0th homology in [1]: A distance between the 0th persistent Betti numbers, also called *size functions*, has been introduced and proven to be stable under perturbations of continuous vector-valued filtering functions. Such a result has been partially extended in [2] for all homology degrees, under the restrictive assumption that the vector-valued filtering functions are *max-tame*.

Contributions. In this paper we introduce a new pseudo-distance d_T between persistent homology groups, applicable also in the case that these groups have torsion (Theorem 2.3). Moreover, we prove that d_T is stable with respect to perturbation of the filtering function, measured by the max-norm (Corollary 2.11). This result is a consequence of a new lower bound for the natural pseudo-distance δ (Theorem 2.9), which also implies an easily computable lower bound for δ (Corollary 2.10). Furthermore, we investigate the link between d_T and the matching distance d_{match} between persistent diagrams, in the particular case that the filtering functions take values in $\mathbb{R}$ and the coefficients are taken in a field, so that also d_{match} can be applied (Theorem 2.5).

The structure of the paper is as follows. In Section 1 we recall the definition of persistent homology group and show some examples. In Section 2 the definition of d_T is given and our results are proven.

1. Some definitions and properties in multidimensional persistence

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

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 $\vec{u} = (u_1, \ldots, u_n) \in \mathbb{R}^n$ and for every function $\vec{\varphi} : X \to \mathbb{R}^n$, we shall denote by $X\langle \vec{\varphi} \preceq \vec{u} \rangle$ the set $\{x \in X : \varphi_i(x) \leq u_i, \ i = 1, \ldots, n\}$. If X is a compact topological space and $\vec{\varphi} : X \to \mathbb{R}^n$ is a continuous function, we shall set $\|\vec{\varphi}\|_\infty = \max_{x \in X} \|\vec{\varphi}(x)\|_\infty$.

Now we can recall the definition of multidimensional persistent homology group.

Definition 1.1 (Persistent homology group). Let $k \in \mathbb{Z}$. Let X be a topological space, and $\vec{\varphi} : X \to \mathbb{R}^n$ a continuous function. Let $i_* : H_k(X\langle \vec{\varphi} \preceq \vec{u} \rangle) \to H_k(X\langle \vec{\varphi} \preceq \vec{v} \rangle)$ be the homomorphism induced by the inclusion map $i : X\langle \vec{\varphi} \preceq \vec{u} \rangle \hookrightarrow X\langle \vec{\varphi} \preceq \vec{v} \rangle$ with $\vec{u} \preceq \vec{v}$, where H_k denotes the kth homology group with coefficients in an Abelian group G. If $\vec{u} \prec \vec{v}$, the image of i_* is called the *multidimensional kth persistent homology group of $(X, \vec{\varphi})$ at $(\vec{u}, \vec{v})$*. We shall denote it by the symbol $H_k^{(X,\vec{\varphi})}(\vec{u}, \vec{v})$. The function $H_k^{(X,\vec{\varphi})}$ is called the *multidimensional kth persistent homology group of $(X, \vec{\varphi})$*.

In other words, the group $H_k^{(X,\vec{\varphi})}(\vec{u}, \vec{v})$ contains all and only the homology classes of cycles born before or at $\vec{u}$ and still alive at $\vec{v}$. We observe that, up to Definition 1.1, each persistent homology group is a function $H_k^{(X,\vec{\varphi})}$, taking each pair $(\vec{u}, \vec{v}) \in \Delta^+$ to a group.

In the rest of this paper, the superscript arrow will be omitted when the filtering function takes values in $\mathbb{R}$.

Remark 1.2. The choice of the particular homology we use is not essential in this context, in the sense that the results proven in this paper do not depend on this choice. However, while we will not further discuss this point, we recall that some important properties of persistent homology groups depend on the homology that is considered. As an example, the use of Čech homology with real coefficients allows us to assume that the function $\beta_\varphi : \Delta^+ \to \mathbb{N} \cup \{\infty\}$ that takes each pair $(\vec{u}, \vec{v}) \in \Delta^+$ to the rank of

the multidimensional kth persistent homology group of $(X, \vec{\varphi})$ at $(\vec{u}, \vec{v})$, is right-continuous in both its variables (cf. [8]).

Example 1.3. Let $\mathbb{RP}^2$ be the real projective plane, represented by the unordered pairs $\{p, -p\}$ of opposite points of $S^2 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$. Let us consider the filtering functions $\varphi, \psi : \mathbb{RP}^2 \to \mathbb{R}$, defined by setting $\varphi(\{(x, y, z), -(x, y, z)\}) = |z|$ and $\psi(\{(x, y, z), -(x, y, z)\}) = 2|z|$. Let us also consider the filtering function $\chi : S^2 \to \mathbb{R}$, defined by setting $\chi(x, y, z) = |z|$. For each $(u, v) \in \Delta^+$, the persistent homology groups in degree 1 computed at (u, v) of φ, ψ and χ are displayed in Figure 1, from left to right. The coefficients of the homology are taken in $\mathbb{Z}$.
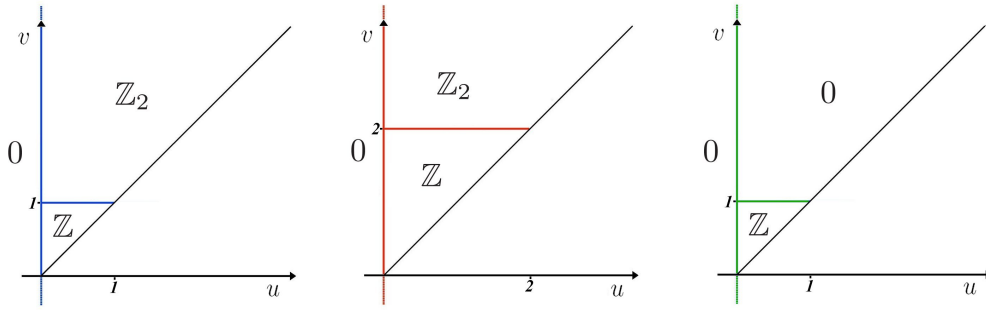

FIGURE 1. Three examples of persistent homology groups with integer coefficients, in degree 1. In the first and second case the topological space is the projective plane, represented by the unordered pairs $\{p, -p\}$ of opposite points of S^2, while in the third case the topological space is S^2. The filtering functions are $\varphi(\{(x, y, z), -(x, y, z)\}) = |z|$, $\psi(\{(x, y, z), -(x, y, z)\}) = 2|z|$ and $\chi(x, y, z) = |z|$, respectively. In each region, the persistent homology group at the points of that region is displayed. The symbol 0 denotes the trivial group.

Since we do not take coefficients in a field, the homology groups we consider have torsion, in the general case. As a consequence, they cannot be described by persistence diagrams (cf. [11]). It follows that the classical matching distance cannot be applied, if we are interested in the information given by the torsion part of our groups. However, we can define a pseudo-distance that does not suffer from this limitation. We shall do this in the next section.

2. The pseudo-distance d_T and the proof of its stability

In the following, Ab will represent the collection of all Abelian groups, considered up to isomorphisms.

Definition 2.1. Let X, Y be two topological spaces, and $\vec{\varphi} : X \to \mathbb{R}^n$, $\vec{\psi} : Y \to \mathbb{R}^n$ two continuous filtering functions. Let $H_k^{(X, \vec{\varphi})} : \Delta^+ \to Ab$ and $H_k^{(Y, \vec{\psi})} : \Delta^+ \to Ab$ be the multidimensional persistent k-homology groups associated with the pairs $(X, \vec{\varphi})$ and $(Y, \vec{\psi})$, respectively. We assume that the homology coefficients are taken in an Abelian group. Let us consider the set E of all $\epsilon \geq 0$ such that, setting $\vec{\epsilon} = (\epsilon, \dots, \epsilon) \in \mathbb{R}^n$, the following statements hold for each $(\vec{u}, \vec{v}) \in \Delta^+$:

(1) a surjective homomorphism from a subgroup of $H_k^{(X, \vec{\varphi})}(\vec{u}, \vec{v})$ onto $H_k^{(Y, \vec{\psi})}(\vec{u} - \vec{\epsilon}, \vec{v} + \vec{\epsilon})$ exists;

(2) a surjective homomorphism from a subgroup of $H_k^{(Y, \vec{\psi})}(\vec{u}, \vec{v})$ onto $H_k^{(X, \vec{\varphi})}(\vec{u} - \vec{\epsilon}, \vec{v} + \vec{\epsilon})$ exists.

We define $d_T\left(H_k^{(X, \vec{\varphi})}, H_k^{(Y, \vec{\psi})}\right)$ equal to $\inf E$ if E is not empty, and equal to ∞ otherwise.

Remark 2.2. We observe that if the set E in Definition 2.1 is not empty, then it is a half-line. In fact, if $\epsilon \in E$, we know that a surjective homomorphism f from a subgroup of $H_k^{(X,\vec{\varphi})}(\vec{u},\vec{v})$ onto $H_k^{(Y,\vec{\psi})}(\vec{u}-\vec{\epsilon},\vec{v}+\vec{\epsilon})$ exists. If $\epsilon' > \epsilon$, then it is easy to prove the existence of a surjective homomorphism g from the group $\bar{G} = H_k^{(Y,\vec{\psi})}(\vec{u}-\vec{\epsilon}',\vec{v}+\vec{\epsilon})$ (seen as a subgroup of $H_k^{(Y,\vec{\psi})}(\vec{u}-\vec{\epsilon},\vec{v}+\vec{\epsilon})$) onto $H_k^{(Y,\vec{\psi})}(\vec{u}-\vec{\epsilon}',\vec{v}+\vec{\epsilon}')$, because of the definition of persistent homology group. Therefore, if we consider the group $\bar{F} = f^{-1}(\bar{G})$, the map $g \circ f_{|\bar{F}}$ is a surjective homomorphism from the subgroup $\bar{F}$ of $H_k^{(X,\vec{\varphi})}(\vec{u},\vec{v})$ onto $H_k^{(Y,\vec{\psi})}(\vec{u}-\vec{\epsilon}',\vec{v}+\vec{\epsilon}')$.

Similarly, we can easily prove that a surjective homomorphism from a subgroup of $H_k^{(Y,\vec{\psi})}(\vec{u},\vec{v})$ onto $H_k^{(X,\vec{\varphi})}(\vec{u}-\vec{\epsilon}',\vec{v}+\vec{\epsilon}')$ exists. As a consequence, if $\epsilon \in E$ and $\epsilon' > \epsilon$, then $\epsilon' \in E$. This proves that E is a half-line.

Before proceeding, we recall that the term "pseudo-distance" means that the considered function verifies all the properties of a distance, with the possible exception of the one assuring that two elements having a vanishing distance must coincide. The term "extended" refers to the possibility that the function takes the value ∞.

Theorem 2.3. *The function d_T is an extended pseudo-distance.*

Proof. The following statements hold:

i)**:** d_T is non-negative by definition;

ii)**:** The equality $d_T\left(H_k^{(X,\vec{\varphi})}, H_k^{(X,\vec{\varphi})}\right) = 0$ is obtained by taking $\epsilon = 0$ and considering the identical homomorphism $id : H_k^{(X,\vec{\varphi})}(\vec{u},\vec{v}) \to H_k^{(X,\vec{\varphi})}(\vec{u}-\vec{\epsilon},\vec{v}+\vec{\epsilon})$;

iii)**:** d_T is symmetrical by definition;

iv)**:** Assume that $d_T\left(H_k^{(X,\vec{\varphi})}, H_k^{(Y,\vec{\psi})}\right) = \epsilon_1$ and $d_T\left(H_k^{(Y,\vec{\psi})}, H_k^{(Z,\vec{\chi})}\right) = \epsilon_2$. Let η be an arbitrarily small positive real number. The definition of d_T implies that for each $(\vec{u},\vec{v}) \in \Delta^+$ a surjective homomorphism f from a subgroup F of $H_k^{(X,\vec{\varphi})}(\vec{u},\vec{v})$ onto $H_k^{(Y,\vec{\psi})}(\vec{u}-(\vec{\epsilon}_1+\vec{\eta}),\vec{v}+(\vec{\epsilon}_1+\vec{\eta}))$, and a surjective homomorphism g from a subgroup G of $H_k^{(Y,\vec{\psi})}(\vec{u}-(\vec{\epsilon}_1+\vec{\eta}),\vec{v}+(\vec{\epsilon}_1+\vec{\eta}))$ onto $H_k^{(Z,\vec{\chi})}(\vec{u}-(\vec{\epsilon}_1+\vec{\eta})-(\vec{\epsilon}_2+\vec{\eta}),\vec{v}+(\vec{\epsilon}_1+\vec{\eta})+(\vec{\epsilon}_2+\vec{\eta}))$ exist. Now, we can consider the subgroup $\tilde{F} = f^{-1}(G) \subseteq H_k^{(X,\vec{\varphi})}(\vec{u},\vec{v})$. Obviously, the function $g \circ f_{|\tilde{F}}$ is a surjective homomorphism from the subgroup $\tilde{F}$ of $H_k^{(X,\vec{\varphi})}(\vec{u},\vec{v})$ onto $H_k^{(Z,\vec{\chi})}(\vec{u}-(\vec{\epsilon}_1+\vec{\eta})-(\vec{\epsilon}_2+\vec{\eta}),\vec{v}+(\vec{\epsilon}_1+\vec{\eta})+(\vec{\epsilon}_2+\vec{\eta}))$. Analogously, we can construct a surjective homomorphism from a subgroup of $H_k^{(Z,\vec{\chi})}(\vec{u},\vec{v})$ onto $H_k^{(X,\vec{\varphi})}(\vec{u}-(\vec{\epsilon}_1+\vec{\eta})-(\vec{\epsilon}_2+\vec{\eta}),\vec{v}+(\vec{\epsilon}_1+\vec{\eta})+(\vec{\epsilon}_2+\vec{\eta}))$. This implies that $d_T\left(H_k^{(X,\vec{\varphi})}, H_k^{(Z,\vec{\chi})}\right) \leq \epsilon_1+\epsilon_2+2\eta$. Since η can be taken arbitrarily small, it follows that $d_T\left(H_k^{(X,\vec{\varphi})}, H_k^{(Z,\vec{\chi})}\right) \leq \epsilon_1+\epsilon_2$, so proving the triangle inequality for d_T.

$\square$

As a simple example, we can easily check that d_T takes the value 1 between the first two persistent homology groups represented in Figure 1, and the value ∞ between the first and the third and between the second and the third persistent homology groups represented in the same figure.

Remark 2.4. We observe that the pseudo-distance d_T is not a distance. For example, we can consider the two Abelian groups $\bigoplus_{i=1}^{\infty} \mathbb{Z}$ (obtained by adding infinite copies of the integer numbers) and $\mathbb{Z}_2 \oplus \bigoplus_{i=1}^{\infty} \mathbb{Z}$. We can find two topological spaces X, Y such that their homology groups in degree 1 are these two groups, respectively. If we take the two filtering functions $\varphi : X \to \mathbb{R}$, $\psi : Y \to \mathbb{R}$ with $\varphi \equiv \psi \equiv 0$, we have that $H_k^{(X,\varphi)}(u,v) = \bigoplus_{i=1}^{\infty} \mathbb{Z}$ and $H_k^{(Y,\psi)}(u,v) = \mathbb{Z}_2 \oplus \bigoplus_{i=1}^{\infty} \mathbb{Z}$ for any $(u,v) \in \Delta^+$ with $u \geq 0$. Obviously, if $(u,v) \in \Delta^+$ with $u < 0$ we get that both $H_k^{(X,\varphi)}(u,v)$ and $H_k^{(Y,\psi)}(u,v)$ are the trivial

group. In any case, it is easy to find a surjective homomorphism from $H_k^{(X,\varphi)}(u,v)$ onto $H_k^{(Y,\psi)}(u,v)$ and a surjective homomorphism from $H_k^{(Y,\psi)}(u,v)$ onto $H_k^{(X,\varphi)}(u,v)$, so that $d_T\left(H_k^{(X,\varphi)}, H_k^{(Y,\psi)}\right) = 0$. However, the groups $H_k^{(X,\varphi)}(u,v), H_k^{(Y,\psi)}(u,v)$ are not isomorphic, for $(u,v) \in \Delta^+$ with $u \geq 0$.

2.1. Relationship between d_T and d_{match} in the 1-dimensional case, if there is no torsion.
Let us go back to the case of filtering functions taking values in $\mathbb{R}$. In this case, if the Abelian group G of homology coefficients is also a field, it is well known that the persistent homology groups can be described by persistence diagrams, under the assumption that the topological space is triangulable (cf. [11] for tame filtering functions, and [8] for continuous filtering functions). The distance that is usually used to compare persistence diagrams is the matching distance d_{match}.

In plain words, the persistence diagram is the collection of all points (u,v) where u and v represent the "birth" and "death" of a homology class, varying the sub-level sets of the topological space X with respect to the filtering function $\varphi : X \to \mathbb{R}$. Each one of these pairs (u,v) is endowed with a multiplicity. By definition, all the pairs (u,u) belong to the persistent diagram, each of them counted with infinite multiplicity. The matching distance d_{match} between two persistence diagrams is the infimum of the cost of a matching (i.e. a bijective correspondence) between two persistence diagrams. The cost of a matching f is defined to be the maximum displacement of the points of the persistence diagrams, induced by f. Each displacement is measured by the max-norm on the real plane.

For a formal and precise definition of persistence diagram and d_{match}, we refer the interested reader to [11, 8].

It is natural to wonder if there is any relationship between d_T and d_{match}, under the hypotheses that we have just made.

We can prove the following result.

Theorem 2.5. *Let X,Y be two triangulable spaces endowed with two continuous functions $\varphi : X \to \mathbb{R}$, $\psi : Y \to \mathbb{R}$. Let us consider the persistent homology groups $H_k^{(X,\varphi)}$, $H_k^{(Y,\varphi)}$ with homology coefficients in a field. Let D_φ and D_ψ be the persistent diagrams associated with the filtering functions φ and ψ, respectively. Then, for any integer k, $d_T\left(H_k^{(X,\varphi)}, H_k^{(Y,\psi)}\right) \leq d_{match}(D_\varphi, D_\psi)$.*

Proof. Since the coefficients are taken in a field, our persistent homology groups have no torsion, so that they can be described by the functions $\dim H_k^{(X,\varphi)}(u,v)$ and $\dim H_k^{(Y,\psi)}(u,v)$, for $(u,v) \in \Delta^+$.

Let us assume that the matching distance d_{match} between our persistence diagrams takes the value $\bar{\lambda}$. Then, from the Representation Theorem 2.17 in [8] (cf. also the k-triangle Lemma in [11]), it follows that for every $\lambda > \bar{\lambda}$ and every pair $(u,v) \in \Delta^+$ the inequalities $\dim H_k^{(X,\varphi)}(u-\lambda, v+\lambda) \leq \dim H_k^{(Y,\psi)}(u,v)$ and $\dim H_k^{(Y,\psi)}(u-\lambda, v+\lambda) \leq \dim H_k^{(X,\varphi)}(u,v)$ hold. This fact immediately implies that for each $(u,v) \in \Delta^+$

(1) a surjective homomorphism from a subgroup of $H_k^{(X,\varphi)}(u,v)$ onto $H_k^{(Y,\psi)}(u-\lambda, v+\lambda)$ exists;

(2) a surjective homomorphism from a subgroup of $H_k^{(Y,\psi)}(u,v)$ onto $H_k^{(X,\varphi)}(u-\lambda, v+\lambda)$ exists.

Hence $d_T\left(H_k^{(X,\varphi)}, H_k^{(Y,\psi)}\right) \leq \lambda$ for any $\lambda > \bar{\lambda}$. This implies our thesis. $\square$

Theorem 2.5 leaves open the possibility that d_T and d_{match} coincide when the persistent homology groups have no torsion and hence d_{match} is defined. This does not happen. Indeed, for any natural number $m \geq 2$ we can give an example for which $d_T\left(H_k^{(X,\varphi)}, H_k^{(Y,\psi)}\right) \leq \frac{2}{m} \cdot d_{match}(D_\varphi, D_\psi)$:

Example 2.6. Let us consider a natural number $m \geq 2$. It is easy to define two triangulable spaces endowed with filtering functions whose persistent diagrams consist of the points $(0,\infty), (0,m), (1, m+1), \ldots, (i, m+i), \ldots, (m, 2m)$ and $(0,\infty), (0,m), (1, m+1), \ldots, (i, m+i), \ldots, (m-1, 2m-1)$, respectively. Here all the points have multiplicity one. In Figure 2 we can see the case $m = 4$. The values of the

corresponding persistent Betti number functions are displayed in Figure 3. It is easy to check that in this case d_T and d_{match} take the value 1 and $\frac{m}{2}$, respectively.
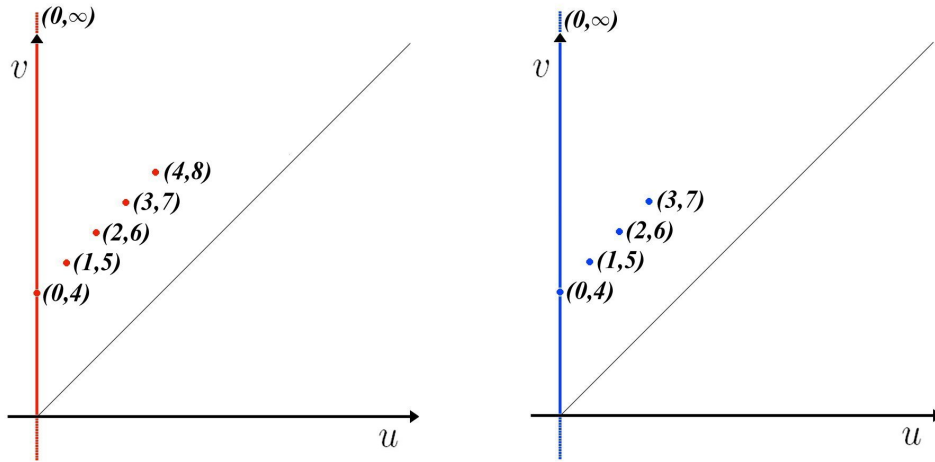

FIGURE 2. It is easy to check that the matching distance d_{match} between the two persistent diagrams displayed in this figure equals 2.
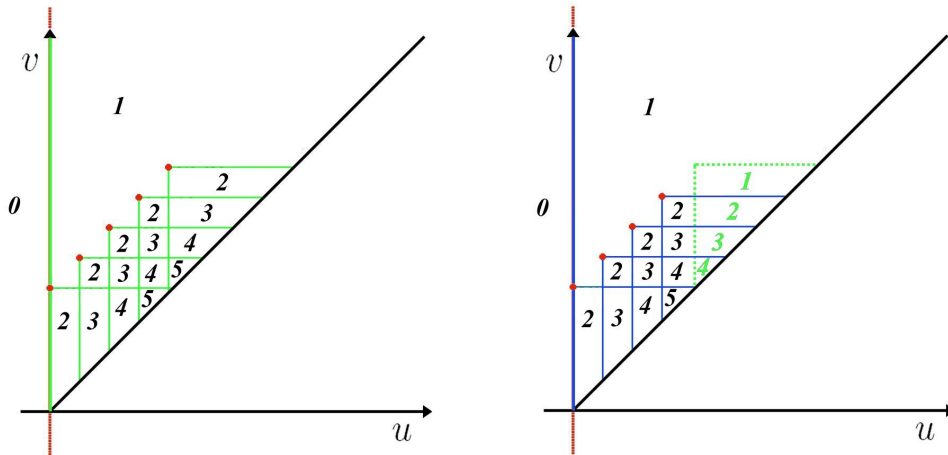

FIGURE 3. Values of the persistent Betti numbers functions represented by the persistent diagrams in Figure 2. In the picture on the right, the dotted triangle represents the part of domain where the persistent Betti numbers functions differ from each other. It is easy to check that the pseudo-distance d_T between the persistent homology groups described by the two persistent diagrams displayed in Figure 2 equals 1.

While Theorem 2.5 shows that the closeness with respect to d_{match} implies the closeness with respect to d_T, Example 2.6 proves that the opposite implication is false. In fact, the ratio between d_T and d_{match} can be arbitrarily small. This means that, in some sense, d_T is less informative than d_{match}. However, we recall that d_T can be applied also in presence of torsion, while d_{match} cannot.

2.2. **Stability of the pseudo-distance** d_T. Another relevant reason to study the pseudo-distance d_T is its stability. This stability can be expressed as a lower bound for the natural pseudo-distance δ. Before proceeding, let us recall the definition of δ.

Definition 2.7 (Natural pseudo-distance). Let X, Y be two compact topological spaces endowed with two continuous functions $\vec{\varphi} : X \to \mathbb{R}^n$, $\vec{\psi} : Y \to \mathbb{R}^n$. The *natural pseudo-distance* between the pairs $(X, \vec{\varphi})$ and $(Y, \vec{\psi})$, denoted by $\delta\left((X, \vec{\varphi}), (Y, \vec{\psi})\right)$, is

    (i) the number $\inf_h \|\vec{\varphi} - \vec{\psi} \circ h\|_\infty$ where h varies in the set of all the homeomorphisms between X and Y, if X and Y are homeomorphic;

    (ii) ∞, if X and Y are not homeomorphic.

We point out that the natural pseudo-distance is not a distance because it can vanish on two distinct pairs. However, it is symmetric, satisfies the triangular inequality, and vanishes on two equal pairs.

The natural pseudo-distance has been studied in [15, 16, 17] in the case of scalar-valued filtering functions on manifolds, and in [22] in the case of vector-valued filtering functions on manifolds.

We are now ready to show that the pseudo-distance d_T is stable. We begin by proving a useful lemma.

Lemma 2.8. *Let X, Y be two homeomorphic compact topological spaces endowed with two continuous functions $\vec{\varphi} : X \to \mathbb{R}^n$, $\vec{\psi} : Y \to \mathbb{R}^n$. Assume that $h : X \to Y$ is a homeomorphism such that $\|\vec{\varphi} - \vec{\psi} \circ h\|_\infty \le \epsilon$. Then for every pair $(\vec{u}, \vec{v}) \in \Delta^+$ a subgroup S of $H_k^{(X, \vec{\varphi})}(\vec{u}, \vec{v})$ and a surjective homomorphism $f : S \to H_k^{(Y, \vec{\psi})}(\vec{u} - \vec{\epsilon}, \vec{v} + \vec{\epsilon})$ exist.*

Proof. The homeomorphism h induces a chain isomorphism h_*. Obviously, h_* takes cycles to cycles and boundaries to boundaries. Let us consider the set S of the elements of $H_k^{(X, \vec{\varphi})}(\vec{u}, \vec{v})$ that contain at least one cycle whose image by h_* is a cycle in $Y\langle \vec{\psi} \preceq \vec{u} - \vec{\epsilon}\rangle$. We claim that S is a subgroup of $H_k^{(X, \vec{\varphi})}(\vec{u}, \vec{v})$. Indeed, if $a_1, a_2 \in S$ then two cycles $\gamma_1 \in a_1$, $\gamma_2 \in a_2$ exist, such that $h_*(\gamma_1), h_*(\gamma_2)$ are cycles in $Y\langle \vec{\psi} \preceq \vec{u} - \vec{\epsilon}\rangle$. Hence the cycle $h_*(\gamma_1) + h_*(\gamma_2) = h_*(\gamma_1 + \gamma_2)$ is a cycle in $Y\langle \vec{\psi} \preceq \vec{u} - \vec{\epsilon}\rangle$, too. Since $\gamma_1 + \gamma_2 \in a_1 + a_2$, it follows that the equivalence class $a_1 + a_2$ belongs to S. That means that S is closed with respect to the sum.

Furthermore, if $a \in S$ then a cycle $\gamma \in a$ exists, such that $h_*(\gamma)$ is a cycle in $Y\langle \vec{\psi} \preceq \vec{u} - \vec{\epsilon}\rangle$. Hence the cycle $h_*(-\gamma) = -h_*(\gamma)$ is a cycle in $Y\langle \vec{\psi} \preceq \vec{u} - \vec{\epsilon}\rangle$, too. Since $-\gamma \in -a$, it follows that the equivalence class $-a$ belongs to S. That means that S is closed with respect to the computation of the opposite.

In summary, S is a subgroup of $H_k^{(X, \vec{\varphi})}(\vec{u}, \vec{v})$.

Now, let us define the homomorphism $f : S \to H_k^{(Y, \vec{\psi})}(\vec{u} - \vec{\epsilon}, \vec{v} + \vec{\epsilon})$. In the following, for each $a \in S$, let us fix a cycle $\gamma_a \in a$ such that $h_*(\gamma_a)$ is a cycle in $Y\langle \vec{\psi} \preceq \vec{u} - \vec{\epsilon}\rangle$. For every $a \in S$ we define $f(a)$ as the element of $H_k^{(Y, \vec{\psi})}(\vec{u} - \vec{\epsilon}, \vec{v} + \vec{\epsilon})$ that contains $h_*(\gamma_a)$ (i.e., in symbols, $[h_*(\gamma_a)]$).

We claim that f is a homomorphism. In order to do that, let us consider three elements $a_1, a_2, a \in H_k^{(X, \vec{\varphi})}(u, v)$, such that $a_1 + a_2 = a$. Obviously,

$$f(a_1) + f(a_2) = [h_*(\gamma_{a_1})] + [h_*(\gamma_{a_2})] = [h_*(\gamma_{a_1}) + h_*(\gamma_{a_2})] = [h_*(\gamma_{a_1} + \gamma_{a_2})]$$

$$f(a) = [h_*(\gamma_a)].$$

Now, $[h_*(\gamma_{a_1} + \gamma_{a_2})] - [h_*(\gamma_a)] = [h_*(\gamma_{a_1} + \gamma_{a_2}) - h_*(\gamma_a)] = [h_*(\gamma_{a_1} + \gamma_{a_2} - \gamma_a)]$. Since $a_1 + a_2 = a$, a $(k+1)$-cycle τ in $X\langle \vec{\varphi} \preceq \vec{v} \rangle$ exists, such that $\partial \tau = \gamma_{a_1} + \gamma_{a_2} - \gamma_a$. Given that $\|\vec{\varphi} - \vec{\psi} \circ h\|_\infty \le \epsilon$, $h_*(\tau)$ is a $(k+1)$-cycle in $Y\langle \vec{\psi} \preceq \vec{v} + \vec{\epsilon} \rangle$. Moreover, recalling that h_* is a chain map, $\partial h_*(\tau) = h_*(\partial \tau) = h_*(\gamma_{a_1} + \gamma_{a_2} - \gamma_a)$, so that $h_*(\gamma_{a_1} + \gamma_{a_2} - \gamma_a)$ is a boundary in $Y\langle \vec{\psi} \preceq \vec{v} + \vec{\epsilon} \rangle$ and hence $[h_*(\gamma_{a_1} + \gamma_{a_2} - \gamma_a)]$ is the null element in $H_k^{(Y,\vec{\psi})}(\vec{u} - \vec{\epsilon}, \vec{v} + \vec{\epsilon})$. This proves that $f(a_1) + f(a_2) = f(a)$.

Finally, we claim that f is surjective. Indeed, let $b \in H_k^{(Y,\vec{\psi})}(\vec{u} - \vec{\epsilon}, \vec{v} + \vec{\epsilon})$ and $\beta \in b$, so that β is a cycle in $Y\langle \vec{\psi} \preceq \vec{u} - \vec{\epsilon} \rangle$. Let α be the cycle $\left((h_*)^{-1} \right)(\beta) = \left(h^{-1} \right)^*(\beta)$. Given that $\|\vec{\varphi} \circ h^{-1} - \vec{\psi}\|_\infty = \|\vec{\varphi} - \vec{\psi} \circ h\|_\infty \le \epsilon$, α is a cycle in $X\langle \vec{\varphi} \preceq \vec{u} \rangle$. Therefore the element $a \in H_k^{(X,\vec{\varphi})}(\vec{u}, \vec{v})$ the cycle α belongs to is an element of the set S. Then $f(a) = [h_*(\gamma_a)]$. Now, γ_a and α are homologous in $X\langle \vec{\varphi} \preceq \vec{v} \rangle$. Once again because of the inequality $\|\vec{\varphi} - \vec{\psi} \circ h\|_\infty \le \epsilon$, it follows that $h_*(\gamma_a)$ and $\beta = h_*(\alpha)$ are homologous in $Y\langle \vec{\psi} \preceq \vec{v} + \vec{\epsilon} \rangle$. Hence $f(a) = b$, and f is proven to be surjective. $\qquad \square$

Now we can prove that d_T gives a lower bound for the natural pseudo-distance δ.

Theorem 2.9. *Let X, Y be two compact topological spaces endowed with two continuous functions $\vec{\varphi} : X \to \mathbb{R}^n$, $\vec{\psi} : Y \to \mathbb{R}^n$. Then $d_T \left(H_k^{(X,\vec{\varphi})}, H_k^{(Y,\vec{\psi})} \right) \le \delta \left((X, \vec{\varphi}), (Y, \vec{\psi}) \right)$.*

Proof. Let us set $\bar{\delta} = \delta \left((X, \vec{\varphi}), (Y, \vec{\psi}) \right)$. Because of the definition of natural pseudo-distance, for each $\eta > 0$ we can find a homeomorphism $h_\eta : X \to Y$ such that $\|\vec{\varphi} - \vec{\psi} \circ h_\eta\|_\infty \le \bar{\delta} + \eta$. By applying Lemma 2.8 to both the homeomorphisms $h_\eta : X \to Y$ and $h_\eta^{-1} : Y \to X$, we get that $d_T \left(H_k^{(X,\vec{\varphi})}, H_k^{(Y,\vec{\psi})} \right) \le \bar{\delta} + \eta$. Since η can be chosen arbitrarily close to 0, our thesis follows. $\qquad \square$

Previous Theorem 2.9 gives an easy way to get lower bounds for the natural pseudo-distance δ:

Corollary 2.10. *Let X, Y be two compact topological spaces endowed with two continuous functions $\vec{\varphi} : X \to \mathbb{R}^n$, $\vec{\psi} : Y \to \mathbb{R}^n$. If no surjective homomorphism exists from a subgroup of $H_k^{(X,\vec{\varphi})}(\vec{u}, \vec{v})$ onto $H_k^{(Y,\vec{\psi})}(\vec{u}', \vec{v}')\}$, then $\delta \left((X, \vec{\varphi}), (Y, \vec{\psi}) \right) \ge \min_i \min\{u_i - u_i', v_i' - v_i\}$.*

Proof. We prove our statement by contradiction, setting $\epsilon = \min_i \min\{u_i - u_i', v_i' - v_i\}$ and assuming that $\delta \left((X, \vec{\varphi}), (Y, \vec{\psi}) \right) < \epsilon$. Then, it follows from Theorem 2.9 that $d_T \left(H_k^{(X,\vec{\varphi})}, H_k^{(Y,\vec{\psi})} \right) < \epsilon$. The definition of the pseudo-distance d_T (Definition 2.1) and Remark 2.2 imply that a surjective homomorphism f from a subgroup of $H_k^{(X,\vec{\varphi})}(\vec{u}, \vec{v})$ onto $H_k^{(Y,\vec{\psi})}(\vec{u} - \vec{\epsilon}, \vec{v} + \vec{\epsilon})$ exists, where $\vec{\epsilon} = (\epsilon, \dots, \epsilon) \in \mathbb{R}^n$.

Obviously, $u_i - \epsilon \ge u_i - (u_i - u_i') = u_i'$ and $v_i + \epsilon \le v_i + (v_i' - v_i) = v_i'$ for every index i, so that $\vec{u} - \vec{\epsilon} \succeq \vec{u}'$ and $\vec{v} + \vec{\epsilon} \preceq \vec{v}'$. Therefore, the definition of persistent homology group easily implies that a surjective homomorphism g exists from a subgroup G of $H_k^{(X,\vec{\varphi})}(\vec{u} - \vec{\epsilon}, \vec{v} + \vec{\epsilon})$ onto $H_k^{(Y,\vec{\psi})}(\vec{u}', \vec{v}')$ (cf. Definition 1.1). As a consequence, $g \circ f_{|F}$ is a surjective homomorphism from the subgroup $F = f^{-1}(G)$ of $H_k^{(X,\vec{\varphi})}(\vec{u}, \vec{v})$ onto $H_k^{(Y,\vec{\psi})}(\vec{u}', \vec{v}')$. This contradicts our assumptions. $\qquad \square$

Our last result illustrates the relationship between d_T and the change of the filtering function, measured by the max-norm. It shows that d_T is a stable distance.

Corollary 2.11. *Let X be a compact topological space endowed with two continuous functions $\vec{\varphi} : X \to \mathbb{R}^n$, $\vec{\psi} : X \to \mathbb{R}^n$. Then $d_T \left(H_k^{(X,\vec{\varphi})}, H_k^{(X,\vec{\psi})} \right) \le \|\vec{\varphi} - \vec{\psi}\|_\infty$.*

Proof. It follows from Theorem 2.9, observing that $\delta \left((X, \vec{\varphi}), (X, \vec{\psi}) \right) \le \|\vec{\varphi} - \vec{\psi}\|_\infty$. $\qquad \square$

Conclusions

The main contribution of this paper is the proof that it is possible to compare persistent homology groups in a stable way also in presence of torsion, although no complete representation by persistent diagrams or other compact descriptors is available, at the time we are writing. Our approach, based on the new pseudo-distance d_T, opens the way to experimentation in the case that the homology coefficients are taken in an Abelian group instead of a field. In this case, the classical matching distance between persistent diagrams cannot be applied without forgetting the information contained in the torsion of the persistent homology groups. We plan to extend these ideas to other concepts and problems in Multidimensional Persistence.

References

[1] S. Biasotti, A. Cerri, P. Frosini, D. Giorgi and C. Landi, *Multidimensional size functions for shape comparison*, Journal of Mathematical Imaging and Vision, vol. 32, no. 2, 161–179 (2008).

[2] F. Cagliari, B. Di Fabio, M. Ferri, *One-dimensional reduction of multidimensional persistent homology*, Proc. Amer. Math. Soc., vol. 138, no. 8, 3003–3017 (2010).

[3] F.Cagliari, C. Landi, *Finiteness of rank invariants of multidimensional persistent homology groups*, Applied Mathematics Letters, vol. 24, no. 4, 516–518 (2011).

[4] G. Carlsson, S. Gurjeet, A. Zomorodian, *Computing Multidimensional Persistence*, ISAAC '09: Proceedings of the 20th International Symposium on Algorithms and Computation, 730–739 (2009).

[5] G. Carlsson, A. Zomorodian, A. Collins, L. J. Guibas, *Persistence Barcodes for Shapes*, International Journal of Shape Modeling, vol. 11, no. 2, 149–187 (2005).

[6] G. Carlsson, A. Zomorodian, *The theory of multidimensional persistence*, SCG '07: Proceedings of the twenty-third annual symposium on Computational geometry, Gyeongju, South Korea, 184–193 (2007).

[7] G. Carlsson, A. Zomorodian, *The Theory of Multidimensional Persistence*, Discrete and Computational Geometry, vol. 42, no. 1, 71–93 (2009).

[8] A. Cerri, B. Di Fabio, M. Ferri, P. Frosini, and C. Landi. Multidimensional persistent homology is stable. Technical Report 2603, Università di Bologna, 2009. available at http://amsacta.cib.unibo.it/2603/.

[9] A. Cerri, M. Ferri, D. Giorgi, *Retrieval of trademark images by means of size functions*, Graph. Models, vol. 68, no. 5, 451–471 (2006).

[10] F. Chazal, D. Cohen-Steiner, M. Glisse, L.J. Guibas, S.Y. Oudot, *Proximity of persistence modules and their diagrams*, SCG '09: Proceedings of the 25th annual symposium on Computational geometry, Aarhus, Denmark, 237–246 (2009).

[11] D. Cohen-Steiner, H. Edelsbrunner, J. Harer, *Stability of Persistence Diagrams*, Discrete and Computational Geometry, vol. 37, no. 1, 103–120 (2007).

[12] D. Cohen-Steiner, H. Edelsbrunner, J. Harer, Y. Mileyko, *Lipschitz Functions Have L_p-Stable Persistence*, Foundations of Computational Mathematics, vol. 10, no. 2, 127–139 (2010).

[13] M. d'Amico, P. Frosini, C. Landi, *Natural pseudo-distance and optimal matching between reduced size functions*, Acta Applicandae Mathematicae, vol. 109, no. 2, 527–554 (2010).

[14] V. de Silva, R. Ghrist, *Coverage in sensor networks via persistent homology*, Algebr. Geom. Topol., vol. 7, 339–358 (2007).

[15] P. Donatini, P. Frosini, *Natural pseudodistances between closed manifolds*, Forum Mathematicum, vol. 16, no. 5, 695–715 (2004).

[16] P. Donatini, P. Frosini, *Natural pseudodistances between closed surfaces*, Journal of the European Mathematical Society, vol. 9, no. 2, 231–253 (2007).

[17] P. Donatini, P. Frosini, *Natural pseudodistances between closed curves*, Forum Mathematicum, vol. 21, no. 6, 981–999 (2009).

[18] H. Edelsbrunner, J. Harer, *Persistent homology—a survey*, Contemp. Math., vol. 453, 257–282 (2008).

[19] H. Edelsbrunner, J. Harer, Computational Topology: An Introduction, American Mathematical Society, 2009.

[20] H. Edelsbrunner, D. Letscher, A. Zomorodian, *Topological persistence and simplification*, Discrete & Computational Geometry, vol. 28, no. 4, 511–533 (2002).

[21] M. Farber, *Invitation to topological robotics*, European Mathematical Society, Zurich Lectures in Advanced Mathematics, 2008.

[22] P. Frosini, M. Mulazzani, *Size homotopy groups for computation of natural size distances*, Bulletin of the Belgian Mathematical Society, vol. 6, no. 3, 455–464 (1999).

[23] R. Ghrist, *Barcodes: the persistent topology of data*, Bull. Amer. Math. Soc. (N.S.), vol. 45, no. 1, 61–75 (2008).

[24] D. Moroni, M. Salvetti, O. Salvetti, *Multi-scale Representation and Persistency for Shape Description*, MDA '08: Proceedings of the 3rd international conference on Advances in Mass Data Analysis of Images and Signals in Medicine, Biotechnology, Chemistry and Food Industry, Leipzig, Germany, 123–138 (2008).

[25] A. Verri, C. Uras, P. Frosini, and M. Ferri, *On the use of size functions for shape analysis*, Biol. Cybern., vol. 70, 99–107 (1993).

[26] A. Zomorodian, Topology for computing, Cambridge Monographs on Applied and Computational Mathematics 16, Cambridge University Press, 2005.

Patrizio Frosini, ARCES, Università di Bologna, via Toffano 2/2, I-40135 Bologna, Italia

Dipartimento di Matematica, Università di Bologna, P.zza di Porta S. Donato 5, I-40126 Bologna, Italia, tel. +39-051-2094478, fax. +39-051-2094490

E-mail address: `patrizio.frosini@unibo.it`