ONE-DIMENSIONAL REDUCTION OF MULTIDIMENSIONAL PERSISTENT HOMOLOGY

FRANCESCA CAGLIARI, BARBARA DI FABIO, AND MASSIMO FERRI

(Communicated by Paul Goerss)

Abstract. A recent result on size functions is extended to higher homology modules: the persistent homology based on a multidimensional measuring function is reduced to a 1-dimensional one. This leads to a stable distance for multidimensional persistent homology. Some reflections on the \( i \)-essentiality of homological critical values conclude the paper.

1. Introduction

At the beginning of the 1990s, Patrizio Frosini introduced the size function \([14, 15, 19, \text{Sect. 8.4}]\), a topological-geometrical tool for describing, analyzing and comparing shapes. This was the origin of a rather large body of experimental research \([22, 23, 24]\). Subsequent papers generalized size functions in two directions, introducing the concepts of size homotopy groups \([17]\) and the size functor \([3]\).

At about the same time the size functor was developed, persistent homology was independently introduced \([13]\). All these theories turn out to be new and effective tools for shape recognition. They are constructed on some topological features of lower level sets of a continuous real-valued function defined on the object of interest. They also share an important advantage with respect to other methods of pattern recognition: they capture qualitative aspects of shape in a formal quantitative way; so, they turn out to be particularly suited to the analysis of “natural” shapes (blood cells, signatures, gestures, melanocytic lesions, ...). In retrospect, a size function can be identified with the rank of a 0-th persistent homology module, while the first persistent homology module is the Abelianization of the first size homotopy group, and the size functor is a functorial formalization of the direct sum of persistent homology modules. Extensive surveys on these items can be found in \([2]\) and \([12]\).

The results obtained recently involving the construction of size functions related to multidimensional measuring functions lead us to a generalization to persistent homology modules, which is the goal of this paper. As far as size theory is concerned, the main reason for such a generalization is that there are shape features that have a multidimensional nature (such as color) and whose description must be done necessarily by a multidimensional measuring function. Moreover, there are shapes which cannot be discriminated by \( n \) size functions related to \( n \) different real-valued measuring functions, but can be distinguished by the size function related...
to the $n$-dimensional measuring function of which those are the components (see Section 5).

A simple example intuitively shows this fact. Let us consider the ellipse $X = \{(u, v, w) \in \mathbb{R}^3 : u^2 + v^2 = 1, v = w\}$ and define $\varphi = (\varphi_1, \varphi_2), \psi = (\psi_1, \psi_2) : X \to \mathbb{R}^2$ such that $\varphi_1 = u, \psi_1 = v, \varphi_2 = \psi_2 = w$. As can be seen in Figure 1, each function among $\varphi_1, \varphi_2, \psi_1, \psi_2$ filters $X$ in such a way that the persistent homology modules of $(X, \varphi_1), (X, \psi_1), (X, \varphi_2) = (X, \psi_2)$ are identical in each homology degree. Instead, by considering the pairs $(X, \vec{\phi}), (X, \vec{\psi})$, at least in degree 0, their persistent homology modules differ from each other. Indeed, while the lower level sets of $\vec{\psi}$ are always either empty or connected, the lower level sets $X(\varphi \preceq (u, w)), \psi \leq a)$, with $0 < a < 1, \sqrt{1 - a^2} \leq w < 1$ consist of two connected components.

As mentioned in [18, Section 2.5], the study of multidimensional persistence is strongly motivated, but there are some objective obstacles. This paper tries to pave a way out of these difficulties.

After recalling some basic notions about multidimensional size functions and 1-dimensional persistent homology in Section 2, we adapt the arguments of [1] to multidimensional persistent homology in Section 3 for proving our main result (Theorem 2). This is a reduction theorem, which takes the detection of discontinuity points back to the case of 1-dimensional persistent homology.

In [4, Section 6] Carlsson and Zomorodian pointed out that there are very serious difficulties in coming to terms with the structure of the rank invariant $\rho_{X,i}$ (the function which assigns to a pair of lower level sets the rank of the image of the homology morphism induced by inclusion) in the multidimensional case. Our Theorem 2 shows that while the sets, on which these ranks are constant, are much more complicated than the triangles typical of the 1-dimensional case, they can be taken back to such triangles when properly “sliced” by a suitable foliation. Stable distances on the leaves of the foliation define (and approximate) a global distance for rank invariants. Examples and further remarks on a different kind of reduction conclude the paper.
2. Basic notions

In the first part of this section we will recall briefly the concept of the multidimensional size function and we will state the theorem that gives us the tools to calculate them (Theorem 1). It asserts, indeed, that a suitable foliation by half-planes of a 2n-dimensional real space makes an n-dimensional size function equal to a 1-dimensional one in correspondence with each plane [1]. In the second part we shall review the definitions of the persistent homology module and related concepts [7].

2.1. Multidimensional size functions and 1-dimensional reduction. In multidimensional size theory, any pair \((X, \tilde{f})\) where \(X\) is a non-empty compact and locally connected Hausdorff space and \(\tilde{f} = (f_1, \ldots, f_n) : X \to \mathbb{R}^n\) is a continuous function is called a size pair. The function \(\tilde{f}\) is called an n-dimensional measuring function. The following relations \(\preceq\) and \(\prec\) are defined in \(\mathbb{R}^n\): for \(\bar{u} = (u_1, \ldots, u_n)\) and \(\bar{v} = (v_1, \ldots, v_n)\), we say \(\bar{u} \preceq \bar{v}\) (resp. \(\bar{u} \prec \bar{v}\)) if and only if \(u_j \leq v_j\) (resp. \(u_j < v_j\)) for every index \(j = 1, \ldots, n\). For every \(n\)-tuple \(\bar{u} = (u_1, \ldots, u_n) \in \mathbb{R}^n\), let \(X(\bar{f} \preceq \bar{u})\) be the set \(\{P \in X : f_j(P) \leq u_j, j = 1, \ldots, n\}\), and let \(\Delta^*_n\) be the open set \(\{(\bar{u}, \bar{v}) \in \mathbb{R}^n \times \mathbb{R}^n : \bar{u} \prec \bar{v}\}\). Finally, let us consider \(\mathbb{R}^n\) endowed with the usual max-norm, \(\|u_1, \ldots, u_n\|_\infty = \max_{1 \leq j \leq n} |u_j|\).

**Definition 1.** For every \(n\)-tuple \(\bar{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n\), we say that two points \(P, Q \in X\) are \(\langle \bar{f} \preceq \bar{v} \rangle\)-connected if and only if a connected subset of \(X(\bar{f} \preceq \bar{u})\) exists containing both \(P\) and \(Q\).

**Proposition 1.** Let \((X, \tilde{f})\) be a size pair. Let also \((\bar{u}, \bar{v}) \in \Delta^*_n\). Then the set \(E\) of equivalence classes in the set \(X(\bar{f} \preceq \bar{u})\) is divided by the \(\langle \bar{f} \preceq \bar{v} \rangle\)-connectedness relation is finite.

**Proof.** First of all, we recall that \(X\) is assumed to be compact and locally connected. Set \(X_1 = X(\bar{f} \preceq \bar{u})\), \(X_2 = X(\bar{f} \prec \bar{v})\), \(X_3 = X(\bar{f} \preceq \bar{v})\), and let \(\iota_1 : X_1 \to X_2\), \(\iota_2 : X_2 \to X_3\) be the respective inclusion maps. \(X_1\) and \(X_3\) are closed, hence compact; \(X_2\) is open, hence locally connected. For \(r = 1, 2, 3\), let \(\mathcal{P}_r = \{C^r_i\}_{i \in S_r}\) be respectively the sets of connected components of \(X_r\), with index sets \(S_r\). For \(r = 1, 2\), let finally \(\iota^r_1 : \mathcal{P}_r \to \mathcal{P}_{r+1}\) be the applications induced by \(\iota_r\) respectively.

The set \(E\) of equivalence classes is in bijection with \(\iota^3_2 \circ \iota^1_1(\mathcal{P}_1)\). Its cardinality is not greater than that of \(\iota^1_1(\mathcal{P}_1)\); our claim is that the latter is finite.

In each connected component \(C^1_i\) of \(X_1\), choose one point \(q_i\) and form the set \(Q = \{q_i\}_{i \in S_1}\). Its closure \(\overline{Q}\) is a compact subset of \(X_1\). For each point \(p \in \overline{Q}\), let \(C_p\) be the connected component of \(X_2\) to which \(p\) belongs; by the local connectedness of \(X_2\), it is open. The set \(\{C_p\}_{p \in \overline{Q}}\) is then an open cover of \(\overline{Q}\); therefore, by the compactness of \(\overline{Q}\), there exists a finite subcover \(\{C_{p_1}, \ldots, C_{p_m}\}\) such that \(\bigcup_{k=1}^m C_{p_k}\) contains all points \(q_i\) \((i \in S_1)\), hence the connected components of \(X_1\). Therefore the connected components of \(X_1\) are contained in the union of a finite set of connected components of \(X_2\).

**Definition 2.** The \((n\text{-dimensional})\) size function associated with the size pair \((X, \tilde{f})\) is the function \(\ell_{(X, \tilde{f})} : \Delta^*_n \to \mathbb{N}\) defined by setting \(\ell_{(X, \tilde{f})}(\bar{u}, \bar{v})\) equal to the number of equivalence classes in which the set \(X(\bar{f} \preceq \bar{u})\) is divided by the \(\langle \bar{f} \preceq \bar{v} \rangle\)-connectedness relation.
An analogous definition for multidimensional persistent homology will be given in Definition 10.

The main goal of [1] for size functions, and of the present paper for persistent homology, is to reduce computation from the multidimensional to the 1-dimensional case. This is possible through particular foliations of \( \mathbb{R}^n \) by half-planes. They are determined by what are called “admissible” vector pairs.

**Definition 3.** For every vector \( \vec{l} = (l_1, \ldots, l_n) \) in \( \mathbb{R}^n \) such that \( \sqrt{\sum_{j=1}^n l_j^2} = 1 \), and \( l_j > 0 \) for \( j = 1, \ldots, n \), and for every vector \( \vec{b} = (b_1, \ldots, b_n) \) in \( \mathbb{R}^n \) such that \( \sum_{j=1}^n b_j = 0 \), we shall say that the pair \( (\vec{l}, \vec{b}) \) is admissible. We shall denote the set of all admissible pairs in \( \mathbb{R}^n \times \mathbb{R}^n \) by \( \text{Adm}_n \). Given an admissible pair \( (\vec{l}, \vec{b}) \), we define the half-plane \( \pi(\vec{l}, \vec{b}) \) in \( \mathbb{R}^n \times \mathbb{R}^n \) by the following parametric equations:

\[
\begin{align*}
\vec{u} &= sl + \vec{b} \\
\vec{v} &= tl + \vec{b}
\end{align*}
\]

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

The motivation for the previous definition is the fact that, for every \( (\vec{u}, \vec{v}) \in \Delta_n^+ \), there exists exactly one admissible pair \( (\vec{l}, \vec{b}) \) such that \( (\vec{u}, \vec{v}) \in \pi(\vec{l}, \vec{b}) \) [1, Prop. 1]. The following lemma is substantially contained in the proof of [1, Thm. 3].

**Lemma 1.** Let \( (\vec{l}, \vec{b}) \) be an admissible pair and let \( g : X \to \mathbb{R} \) be defined by setting

\[
g(P) = \max_{j=1,\ldots,n} \left\{ \frac{f_j(P) - b_j}{l_j} \right\}.
\]

Then, for every \( (\vec{u}, \vec{v}) = (sl + \vec{b}, tl + \vec{b}) \in \pi(\vec{l}, \vec{b}) \), the following equalities hold:

\[
X(\vec{f} \preceq \vec{u}) = X(g \leq s), \quad X(\vec{f} \preceq \vec{v}) = X(g \leq t).
\]

**Proof.** For every \( \vec{u} = (u_1, \ldots, u_n) \in \mathbb{R}^n \), with \( u_j = sl_j + b_j, j = 1, \ldots, n \), it holds that

\[
X(\vec{f} \preceq \vec{u}) = \{ P \in X : f_j(P) \leq u_j, j = 1, \ldots, n \} = \{ P \in X : f_j(P) \leq sl_j + b_j, j = 1, \ldots, n \} = \{ P \in X : \frac{f_j(P) - b_j}{l_j} \leq s, j = 1, \ldots, n \} = X(g \leq s).
\]

Analogously, for every \( \vec{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n \), with \( v_j = tl_j + b_j, j = 1, \ldots, n \), it can be proved that \( X(\vec{f} \preceq \vec{v}) = X(g \leq t) \). \( \square \)

From the above result, there follows the main theorem of [1].

**Theorem 1.** Let \( (\vec{l}, \vec{b}) \) and \( g \) be defined as in Lemma 1. Then the equality

\[
\ell(\pi, \vec{l}, \vec{b}) \equiv \ell(\pi, g)(s, t)
\]

holds for every \( (\vec{u}, \vec{v}) = (sl + \vec{b}, tl + \vec{b}) \in \pi(\vec{l}, \vec{b}) \).
This is indeed the theorem that we are going to extend, in Section 3, to persistent homology of all degrees. Its importance resides in the fact that essential discontinuity points (“cornerpoints” in the terminology of size theory) are the key to a stable distance between size functions. Unfortunately, cornerpoints do not form in general, discrete sets in the multidimensional case. This theorem makes it possible to find them “slice by slice” with the familiar technique of dimension one. A practical use is for sampling their sets, so getting bounds for a stable distance between size functions. Our extension will produce the same opportunity for persistent homology.

2.2. 1-dimensional persistent homology. Given a topological space $X$ and an integer $i$, we denote the $i$-th singular homology module of $X$ over a field $k$ by $H_i(X)$. Moreover, let $\Delta_+^i = \{(u, v) \in \mathbb{R}^2 : u < v\}$, and let $\Delta_1 = \{(u, v) \in \mathbb{R}^2 : u = v\}$.

Next we restate two definitions of [7].

Definition 4. Let $X$ be a topological space and let $f$ be a continuous real function on $X$. A homological critical value of $f$ is a real number $c$ for which there exists an integer $i$ such that, for all sufficiently small $\varepsilon > 0$, the map $H_i(X(f \leq c - \varepsilon)) \rightarrow H_i(X(f \geq c + \varepsilon))$ induced by inclusion is not an isomorphism.

This is called an $i$-essential critical value in the paper [3, Def. 2.6], dedicated to the size functor, a contemporary and not too different homological generalization of size functions.

Definition 5. A continuous function $f : X \rightarrow \mathbb{R}$ is tame if it has a finite number of homological critical values and the homology modules $H_i(X(f \leq c))$ are finite-dimensional for all $i \in \mathbb{Z}$ and $c \in \mathbb{R}$.

The reader should be warned that there exist other, different, meanings of “tame” in the current topological literature. Actually, “homologically tame” might be a better designation for such a type of function, but we adhere to this already current definition.

We write $F_i^u = H_i(X(f \leq u))$ for each $i \in \mathbb{Z}$, and, for $u < v$, we let $f_i^{u,v} : F_i^u \rightarrow F_i^v$ be the map induced by inclusion of the lower level set of $u$ in that of $v$, for a fixed integer $i$. Moreover, we indicate by $F_i^{u,v} = \text{Im} f_i^{u,v}$ the image of $F_i^u$ in $F_i^v$ that is called the $i$-th persistent homology module. For each homology degree $i \in \mathbb{Z}$ the rank invariant $\rho_{X,i} : \Delta_+^i \rightarrow \mathbb{N}$ is defined as $\rho_{X,i}(u, v) = \text{rank}(F_i^{u,v})$.

It is well-known for 1-dimensional tame measuring functions [7, 16, 21] that the rank invariant $\rho_{X,i}$ of a size pair $(X, f)$ can be recovered, for each degree $i$, from its multiset of cornerpoints, which is called the “persistence diagram”. We now recall the pertinent definitions.

Definition 6. For every point $p = (u, v) \in \Delta_+^i$, let $\mu_i(p)$ be the minimum over all the positive real numbers $\varepsilon$, with $u + \varepsilon < v - \varepsilon$, of

$$(2.1) \quad \rho_{X,i}(u + \varepsilon, v - \varepsilon) - \rho_{X,i}(u - \varepsilon, v - \varepsilon) - \rho_{X,i}(u + \varepsilon, v + \varepsilon) + \rho_{X,i}(u - \varepsilon, v + \varepsilon).$$

The number $\mu_i(p)$ will be called the multiplicity of $p$ for $\rho_{X,i}$. Moreover, we shall call a proper cornerpoint for $\rho_{X,i}$ any point $p \in \Delta_+^i$ such that the number $\mu_i(p)$ is strictly positive.
Definition 7. For every vertical line \( r \), with equation \( u = \bar{u} \), \( \bar{u} \in \mathbb{R} \), let us identify \( r \) with \((\bar{u}, \infty)\), and let \( \mu_i(r) \) be the minimum over all the positive real numbers \( \varepsilon \), with \( \bar{u} + \varepsilon < 1/\varepsilon \), of
\[
\rho_{X,i} \left( \bar{u} + \varepsilon, \frac{1}{\varepsilon} \right) - \rho_{X,i} \left( \bar{u} - \varepsilon, \frac{1}{\varepsilon} \right).
\]
The number \( \mu_i(r) \) will be called the multiplicity of \( r \) for \( \rho_{X,i} \). When this finite number is strictly positive, we call \( r \) a cornerpoint at infinity for \( \rho_{X,i} \).

Observe that the above numbers \( \mu_i(p) \) and \( \mu_i(r) \) exist because of the property of tameness of the measuring functions and since (2.1) and (2.2) are nonnegative and non-decreasingly monotone with respect to \( \varepsilon \).

The concept of cornerpoint allows us to introduce a representation of the rank invariant, based on the following definition, that substantially corresponds to that in [7, Sect. 2].

Definition 8. The persistence diagram \( D_i(X, f) \subset \Delta_1^+ \cup \Delta_1 \cup \{(k, \infty) : k \in \mathbb{R}\} \) is the multiset of all cornerpoints (both proper and at infinity) for \( \rho_{X,i} \), counted with their multiplicity, plus a countable infinity of points of \( \Delta_1 \).

3. Homological 1-dimensional reduction

In this section we define the \( i \)-th persistent homology module related to a continuous \( \mathbb{R}^n \)-valued function (substantially as in [4]). Then we show that the sets of points of \( \mathbb{R}^{2n} \), where the modules change, can be obtained by computing the cornerpoints of persistence diagrams of \( \mathbb{R} \)-valued functions defined on particular half-planes which foliate the \( 2n \)-space.

The first issue arises when one tries to compute the maximum between the components of a \( \mathbb{R}^n \)-valued function. In fact:

Remark 1. The maximum of two tame functions is not necessarily a tame function.

(We recall that “tame” has the meaning defined in Section 2.2)

As an example, let \( f_1, f_2 : \mathbb{R}^2 \to \mathbb{R} \) be two tame functions defined as
\[
f_1(u, v) = \begin{cases} 
v - u^2 \sin \left( \frac{1}{u} \right) & u \neq 0 \\
v & u = 0 \end{cases}, \quad f_2(u, v) = \begin{cases} 
-u - u^2 \sin \left( \frac{1}{u} \right) & u \neq 0 \\
-v & u = 0 \end{cases}
\]
and consider the function
\[
f = \max(f_1, f_2).
\]

Figures 2–4 show the sets where \( f_1, f_2 \) and \( f \), respectively, take a value smaller than or equal to 0. It is easily seen that, even if \( f_1 \) and \( f_2 \) are tame functions, with \( \mathbb{R}^2(f_1 \leq 0) \) and \( \mathbb{R}^2(f_2 \leq 0) \) being connected lower-level sets, \( f \) does not become tame, since \( H_0(\mathbb{R}^2(f \leq 0)) \) is not a finitely generated module.

Given this fault related to tame functions, the solution we propose is to introduce the following concept.

Definition 9. Let \( (X, \vec{f}) \) be a size pair. We shall say that \( \vec{f} \) is max-tame if, for every admissible pair \((\vec{l}, \vec{b})\), the function \( g(P) = \max_{j=1,\ldots,n} \left\{ \frac{f_j(P) - b_j}{l_j} \right\} \) is tame.
For conciseness, in the following, a size pair \((X, \tilde{f})\) with \(\tilde{f}\) a max-tame function will be called a max-tame size pair.

Choosing a size pair as above, let us define the multidimensional persistent homology modules.

**Definition 10.** Let \((X, \tilde{f})\) be a max-tame size pair. For each homology degree \(i \in \mathbb{Z}\), we put
\[
F_{\tilde{u}}^i = H_i(X \langle \tilde{f} \leq \tilde{u} \rangle)
\]
for each \(\tilde{u} \in \mathbb{R}^n\). For \(\tilde{u} \preceq \tilde{v}\), we let
\[
F_{\tilde{u}}^i : F_{\tilde{v}}^i \to F_{\tilde{u}}^i
\]
be the map induced by inclusion of the lower level set of \(\tilde{u}\) in that of \(\tilde{v}\) for a fixed integer \(i\), and we call
\[
F_{\tilde{u}}^i : = \text{Im} F_{\tilde{u}}^i
\]
the \(i\)-th multidimensional persistent homology module.

Note that the rank of \(F_{\tilde{u}}^i\) is what is called a rank invariant and is denoted by \(\rho_{X,i}(\tilde{u}, \tilde{v})\) in [4, Def. 12].

Let \(g(P) = \max_{j=1, \ldots, n} \{ l_j(P) - b_k \} \) for a fixed \((\tilde{L}, \tilde{b}) \in Adm_n\), let \(G_s^i = H_i(X \langle g \leq s \rangle)\) for all \(s \in \mathbb{R}\), and let \(i \in \mathbb{Z}\). For \(s < t\), we let \(G^s_t : G^s_i \to G^t_i\) be the map induced by inclusion of the lower level set of \(s\) in that of \(t\) for a fixed integer \(i\), and we denote by \(G^s_t = \text{Im} G^s_t\) the \(i\)-th persistent homology module.

Now we can state and prove the theorem which, in analogy with the main result of [1], enables us to reduce the computation of the multidimensional persistent homology to the 1-dimensional one. This is important, not so much for finding the
Theorem 2. Let \((\bar{l}, \bar{b})\) be an admissible pair and let \(\bar{f} = (f_1, \ldots, f_n) : X \to \mathbb{R}^n\) be a max-tame function. Then, for every \((\bar{u}, \bar{v}) = (s\bar{l} + \bar{b}, t\bar{l} + \bar{b}) \in \pi(\bar{l}, \bar{b})\), the equality
\[
F_i^{\bar{u}, \bar{v}} = C_i^{s,t}
\]
holds for all \(i \in \mathbb{Z}\) and \(s, t \in \mathbb{R}\) with \(s < t\).

Proof. By Lemma 1, we have \(X(\bar{f} \leq \bar{u}) = X(g \leq s)\) and \(X(\bar{f} \leq \bar{v}) = X(g \leq t)\) for every \((\bar{u}, \bar{v}) = (s\bar{l} + \bar{b}, t\bar{l} + \bar{b}) \in \pi(\bar{l}, \bar{b})\). Therefore, for every \(i \in \mathbb{Z}\), \(H_i(X(\bar{f} \leq \bar{u})) = H_i(X(g \leq s))\) and \(H_i(X(\bar{f} \leq \bar{v})) = H_i(X(g \leq t))\), respectively. Thus, since \(f_i^{\bar{u}, \bar{v}} : H_i(X(\bar{f} \leq \bar{u})) \to H_i(X(\bar{f} \leq \bar{v}))\) and \(g_i^{s,t} : H_i(X(g \leq s)) \to H_i(X(g \leq t))\) are homomorphisms induced by inclusion, having the same domain and codomain, \(f_i^{\bar{u}, \bar{v}} = g_i^{s,t}\) for every \(i\). Hence \(\text{Im } f_i^{\bar{u}, \bar{v}} = \text{Im } g_i^{s,t}\) and the claim is proved. \(\square\)

4. Multidimensional matching distance

In the sequel, to facilitate the reader, we will denote by \(\rho\) a multidimensional rank invariant and by \(\rho\) a 1-dimensional one.

According to [4, Def. 12], for a given measuring function \(\bar{f} : X \to \mathbb{R}^n\), for each homology degree \(i \in \mathbb{Z}\), the rank invariant \(\overline{\rho}_{X,i} : \Delta^+_{\bar{i}} \to \mathbb{N}\) is defined as \(\overline{\rho}_{X,i}(\bar{u}, \bar{v}) = \text{rank}(f_i^{\bar{u}, \bar{v}})\).

Let \((X, \bar{f}), (Y, \bar{f}''')\) be two max-tame size pairs, and let \(\overline{\rho}_{X,i}, \overline{\rho}_{Y,i}'\) be the respective rank invariants. Let an admissible pair \((\bar{l}, \bar{b})\) be fixed, and let \(g' : X \to \mathbb{R}\), \(g'' : Y \to \mathbb{R}\) be defined as in Lemma 1 by setting
\[
g'(P) = \max_{j=1, \ldots, n} \left\{ \frac{f'_j(P) - b_j}{l_j} \right\}, \quad g''(P) = \max_{j=1, \ldots, n} \left\{ \frac{f''_j(P) - b_j}{l_j} \right\}.
\]

As recalled in Section 2.2, the relevant information on the rank invariants \(\rho'_{X,i}, \rho''_{Y,i}\) of \(g'\) and \(g''\) respectively is contained, for each degree \(i\), in their multisets of cornerpoints. These are sets of points of the extended plane with multiplicities, augmented by adding a countable infinity of points of the diagonal \(y = x\); let them be called respectively \(C'\) and \(C''\). Each cornerpoint is determined by its coordinates \(x < y \leq \infty\). The distance of two cornerpoints is defined as
\[
\delta((a, b), (c, d)) = \min \left\{ \max\{|a-c|, |b-d|\}, \max\left\{ \frac{b-a}{2}, \frac{d-c}{2} \right\} \right\}.
\]
The adopted convention about \(\infty\) is that \(-\infty - a = a - \infty = \infty\) for \(a \neq \infty\), \(-\infty - \infty = 0\), \(\frac{\infty}{-\infty} = \infty\), \(\min\{\infty, a\} = a\), \(\max\{\infty, a\} = \infty\).

It has been proved in [7] that, using the matching (or bottleneck) distance, i.e.
\[
d(\rho'_{X,i}, \rho''_{Y,i}) = \inf_{\sigma} \sup_{P \in C'} \delta(P, \sigma(P)),
\]
where \(\sigma\) varies among all bijections from \(C'\) to \(C''\), persistence diagrams are robust against small perturbations of measuring functions. Mimicking [4] (and recalling that \(\rho'_{X,i}, \rho''_{Y,i}\) vary with \((\bar{l}, \bar{b})\)), we can use \(d\) to define distances \(D\) between the rank invariants of the original multidimensional persistent homologies.
Definition 11. Let \((X, \vec{f})\), \((Y, \vec{f}')\) be two max-tame size pairs and let \(\rho'_{X,i}, \rho''_{Y,i}\) be the respective rank invariants. Then the \(i\)-th multidimensional matching distance between rank invariants is defined as the extended distance

\[
D(\rho'_{X,i}, \rho''_{Y,i}) = \sup_{(l, b) \in \text{Adm}_n} \min_{j=1, \ldots, n} l_j \cdot d(\rho'_{X,i}, \rho''_{Y,i}).
\]

Note that \(D\) is by construction a global distance, i.e. not depending on \((l, b)\). However, this does not mean that every admissible pair should yield a smaller distance \(d\); this is due to the fact that \(D\) is not the supremum of the pair-depending distances \(d\), but of their products with coefficients \(l_j\) which are \(< 1\).

An easy corollary of our Theorem 2 is the following, which is the higher degree version of [1, Cor. 1].

Corollary 1. For each \(i \in \mathbb{Z}\), the identity \(\rho'_{X,i} \equiv \rho''_{Y,i}\) holds if and only if, for every admissible pair \((l, b)\), \(d(\rho'_{X,i}, \rho''_{Y,i}) = 0\).

With the same argument of the analogous Proposition 4 of [1], it is easy to prove the following inequality between the multidimensional matching distance and the 1-dimensional one obtained by considering the components of the measuring functions. That this inequality can be strict is shown in Section 5.

Proposition 2. Let \((X, \vec{f})\), \((Y, \vec{h})\) be two max-tame size pairs such that \(f_j, h_j\) are tame functions for every \(j = 1, \ldots, n\). For each \(i \in \mathbb{Z}\), let \(\rho^f_{X,i}, \rho^h_{Y,i}\) be the \(i\)-th rank invariants relative to the components \(f_j, h_j\) respectively; then let \(\rho'_{X,i}, \rho''_{Y,i}\) be the rank invariants relative to \(\vec{f}, \vec{h}\) respectively. Then it holds that

\[
d(\rho^f_{X,i}, \rho^h_{Y,i}) \leq D(\rho'_{X,i}, \rho''_{Y,i}).
\]

The persistence diagrams are known to be stable under possibly small perturbations of 1-dimensional measuring functions [7, Section 3.1] (see also [9, Thm. 25]). In the multidimensional setting, the stability of rank invariants under perturbations of 1-dimensional measuring functions on each leaf of the foliation is stated by the following proposition, whose proof is again a copy of that of [1, Prop. 2]. Here \(g', g'' : X \to \mathbb{R}\) are defined in correspondence to \((l, b)\) as at the beginning of this section.

Proposition 3. If \((X, \vec{f})\), \((X, \vec{f}')\) are max-tame size pairs, and \(\max_{P \in X} \|\vec{f}'(P) - \vec{f}(P)\|_{\infty} \leq \epsilon\), then, for each admissible pair \((l, b)\), with \(l = (l_1, \ldots, l_n)\), and for each \(i \in \mathbb{Z}\), it holds that

\[
d(\rho'_{X,i}, \rho''_{X,i}) \leq \frac{\epsilon}{\min_{j=1, \ldots, n} l_j},
\]

where \(\rho'_{X,i}, \rho''_{X,i}\) are the rank invariants at degree \(i\) of \((X, g')\), \((X, g'')\) respectively.

By the definition of \(D\), every 1-dimensional matching distance obtained in correspondence of an admissible pair yields a lower bound for the multidimensional matching distance \(D\). While \(D\) is difficult or even impossible to compute, we conjecture that a suitable, sufficiently fine sampling by admissible pairs produces approximations of arbitrary precision of it. This is likely to become the common practice in applications.
Of course, \( \max_{i \in \mathbb{Z}} D(\mathcal{P}_{X,i}^{\prime},\mathcal{P}_{Y,i}^{\prime}) \) is still a meaningful distance related to the size pairs \((X, \vec{f}^{\prime}), (Y, \vec{f}^{\prime\prime})\), and it would be interesting to study its relation with the \(n\)-dimensional natural pseudo-distance of [1, 10, 17].

5. Examples and remarks

This section provides two simple examples of shape comparison. Their aim is to demonstrate that the discriminatory power derivable from one comparison between \(n\)-dimensional persistent homology modules, i.e. computed with respect to an \(n\)-dimensional measuring function, is higher than the one achievable in \(n\) comparisons between the collection of the 1-dimensional persistent homology modules computed with respect to each component.

The first example we propose is the comparison of two shapes defined by two different topological spaces endowed with the restrictions of the same \(\mathbb{R}^2\)-valued measuring function. The dissimilarity of these shapes is computed in terms of the matching distance between the respective rank invariants.

In \(\mathbb{R}^3\) consider the set \(\Omega = [-1,1] \times [-1,1] \times [-1,1]\) and the sphere \(S\) of equation \(u^2 + v^2 + w^2 = 1\). Let also \(\vec{f} = (f_1, f_2) : \mathbb{R}^3 \to \mathbb{R}^2\) be a continuous function, defined as \(\vec{f}(u,v,w) = (|u|, |v|)\). In this setting, consider the size pairs \((\mathcal{C}, \varphi)\) and \((\mathcal{S}, \psi)\), where \(\mathcal{C} = \partial \Omega\) and \(\varphi\) and \(\psi\) are respectively the restrictions of \(\vec{f}\) to \(\mathcal{C}\) and \(S\).

In order to compare the persistent homology modules of \(\mathcal{C}\) and \(S\) defined by \(\vec{f}\), we are interested in studying the half-planes’ foliation of \(\mathbb{R}^4\), where \(\vec{l} = (\cos \theta, \sin \theta)\) with \(0 < \theta < \frac{\pi}{2}\), and \(\vec{b} = (a, -a)\) with \(a \in \mathbb{R}\). Any such half-plane is parameterized as

\[
\begin{align*}
&u_1 = s \cos \theta + a \\
&v_1 = s \sin \theta + a \\
&u_2 = t \cos \theta + a \\
&v_2 = t \sin \theta - a
\end{align*}
\]

with \(s, t \in \mathbb{R}, s < t\).

Let us choose, for example, the half-plane defined by \(\theta = \frac{\pi}{4}\) and \(a = 0\), i.e. \(\vec{l} = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})\) and \(\vec{b} = (0, 0)\). Then we obtain (Figure 5) that

\[
\begin{align*}
g' &= \sqrt{2} \max\{\varphi_1, \varphi_2\} = \sqrt{2} \max\{|u|, |v|\}, \\
g'' &= \sqrt{2} \max\{\psi_1, \psi_2\} = \sqrt{2} \max\{|u|, |v|\}.
\end{align*}
\]

Figure 5. Lower level sets \(\mathcal{C}(g' \leq 1)\) and \(S(g'' \leq 1)\)
For every \( i \in \mathbb{Z} \), let \( \rho_{C,i}', \rho_{S,i}' \) and \( \rho_{C,i}'' \) be respectively the 2-dimensional rank invariants associated with \((C, e')\) \( (S, \psi')\), and the 1-dimensional rank invariants associated with \((C, g'_i)\) \( (S, g''_i)\). So, writing \( G_{s,i}(C) = H_{i}(\mathcal{C}((g'_i \leq s))) \), \( G_{s,i}(S) = H_{i}(\mathcal{S}((g''_i \leq s))) \) and recalling that \( k \) is the coefficient field, we will show the persistent homology modules \( G_{s,t}^{s,t}(C) \) \( (S) \) and the respective rank invariants with bold typed values (right), and estimate the matching distance \( D(\rho_{C,i}', \rho_{S,i}'') \) by computing \( d(\rho_{C,i}', \rho_{S,i}'') \).

\[
G_{0,t}^{s,t}(C) = \begin{cases} 
0, & s < 0 \\
k^2, & 0 \leq s < \sqrt{2} \\
k, & \text{otherwise}
\end{cases}
\]

\[
G_{0,t}^{s,t}(S) = \begin{cases} 
0, & s < 0 \\
k^2, & 0 \leq s < 1 \\
k, & \text{otherwise}
\end{cases}
\]

\[
D(\rho_{C,0}'', \rho_{S,0}'') \geq \frac{\sqrt{2}}{2} d(\rho_{C,0}', \rho_{S,0}'') = \frac{\sqrt{2}}{2} (\sqrt{2} - 1);
\]

\[
G_{1,t}^{s,t}(C) = 0, \quad \text{for all } s, t \in \mathbb{R}
\]

\[
G_{1,t}^{s,t}(S) = \begin{cases} 
k^3, & 1 \leq s < \sqrt{2} \\
0, & \text{otherwise}
\end{cases}
\]

\[
D(\rho_{C,1}', \rho_{S,1}'') \geq \frac{\sqrt{2}}{2} d(\rho_{C,1}', \rho_{S,1}'') = \frac{\sqrt{2}}{2} (\sqrt{2} - \frac{1}{2});
\]

\[
G_{2,t}^{s,t}(C) = \begin{cases} 
k, & \sqrt{2} \leq s < t \\
0, & \text{otherwise}
\end{cases}
\]

\[
G_{2,t}^{s,t}(S) = \begin{cases} 
k, & \sqrt{2} \leq s < t \\
0, & \text{otherwise}
\end{cases}
\]

\[
D(\rho_{C,2}', \rho_{S,2}'') \geq \frac{\sqrt{2}}{2} d(\rho_{C,2}', \rho_{S,2}'') = 0.
\]
In other words, multidimensional persistent homology, with respect to \( \vec{\phi} \) and \( \vec{\psi} \), is able to discriminate the cube and the sphere, while the 1-dimensional one, with respect to \( \varphi_1, \varphi_2 \) and \( \psi_1, \psi_2 \), cannot do that. In fact, for either manifold the lower level sets of the single components (i.e., 1-dimensional measuring functions) are homeomorphic for all values: they are topologically either circles, or annuli, or spheres.

The second example concerns the comparison of two different size pairs having the same support.

Let us take into account the size pairs \((X, \vec{\phi})\), \((X, \vec{\psi})\) described in the Introduction (Figure 1), and for the sake of simplicity, consider again the half-plane used in the preceding example: \( \pi_{(l, b)} \), with \( l = (\sqrt{2}, 0) \) and \( b = (0, 0) \). Since 
\[
g'(\varphi_1, \varphi_2) = \sqrt{2} \max\{\varphi_1, \varphi_2\}, \quad g''(\psi_1, \psi_2) = \sqrt{2} \max\{\psi_1, \psi_2\}
\]

it holds that 
\[
G_{s,t}^0(X, g') = \begin{cases} 
0, & s < -1 \\
\frac{k^2}{2}, & 1 \leq s < t < \sqrt{2} \\
k, & \text{otherwise}
\end{cases}
\]
\[
G_{s,t}^0(X, g'') = \begin{cases} 
\frac{k}{2}, & -\sqrt{2} \leq s < t \\
0, & \text{otherwise}
\end{cases}
\]

\[
G_{s,t}^1(X, g') = \begin{cases} 
\frac{k}{2}, & \sqrt{2} \leq s < t \\
0, & \text{otherwise}
\end{cases}
\]
\[
G_{s,t}^1(X, g'') = \begin{cases} 
\frac{k}{2}, & \sqrt{2} \leq s < t \\
0, & \text{otherwise}
\end{cases}
\]

\[
\Rightarrow D(\rho(X, \vec{\phi}), 0, \rho(X, \vec{\psi}), 0) \geq \frac{\sqrt{2}}{2} d(\rho(X, g'), 0, \rho(X, g''), 0) = \frac{\sqrt{2}}{2} (\sqrt{2} - 1);
\]
\[
\Rightarrow D(\rho(X, \vec{\phi}), 1, \rho(X, \vec{\psi}), 1) \geq \frac{\sqrt{2}}{2} d(\rho(X, g'), 1, \rho(X, g''), 1) = 0.
\]

6. Reduction of \( i \)-essential critical values

The former example of the previous section suggests also some other considerations on the cooperation of measuring functions. We remind the reader that the adjective “essential” is used here with the meaning introduced in [3] and recalled in Section 2.2 after Definition 4, in a different way than, e.g., in [6].

A first remark regards the possibility of considering Jacobi sets [11] on \( C \) and \( S \). Loosely speaking, the Jacobi set of two Morse functions defined on a common
manifold is the set of critical points of the restrictions of one function to the level sets of the other function. In our case, the components of $\vec{f} : \mathbb{R}^3 \to \mathbb{R}^2$ (that we recall to be the function whose restrictions on $C$ and $S$ are $\vec{\varphi}$ and $\vec{\psi}$, respectively) are not Morse functions, nor is $C$ a manifold; nevertheless, considerations in this direction can be done. Indeed, note that although the persistent homology on single components of $\vec{f}$ cannot distinguish the two spaces, the persistent homology on $f_1$, for example, if restricted to lower-level sets of $f_2$, can, as shown in what follows. Consider again the sphere $S$ and $\vec{\psi} = f_\mid_S$. The value $1/\sqrt{2}$ (corresponding to the homological critical value 1 of $\vec{g}''$) is not homologically critical for the maps $\psi_1, \psi_2$ on $S$ itself, but it is indeed homologically critical for $\psi_2$ restricted to $S(\psi_1 \leq 1/\sqrt{2})$. We believe that, for every $i \in \mathbb{Z}$, homological critical values of the 1-dimensional reduction of multidimensional measuring functions are always clues to such phenomena (the case $i = 0$ has already been treated in [5]).

A further speculation on the use of cooperating measuring functions — from a completely different viewpoint than the one developed in the previous sections — is the following. A problem in 1-dimensional persistent homology, as well as for the size functor, is the computation of $i$-essential critical values for $i > 0$. A possibility is the use of several independent measuring functions for lowering $i$, i.e. the degree at which the passage through the critical value causes a homology change. Lowering $i$ is important, since 0-essential critical values are easily detected by graph-theoretical techniques [5]. The following example shows that a suitable choice of a second, auxiliary measuring function may actually take 1-essential critical values to 0-essential ones.

Let $T$ be a torus of revolution around the $x$ axis, with the innermost parallel circle of radius 2, the outermost of radius 3. On $T$ define $(f_1, f_2) = (z, -z)$. Suppose we are interested in the persistent homology of the size pair $(T, f_1)$. Then $(0, 0, 2)$ is a 1-essential critical point for $f_1$; i.e. it is a point at which 1-degree homology changes. Of course, there are computational methods (e.g. by the Euler–Poincaré characteristic) which enable us to detect it, but they will probably be tailored to the particular dimension of the manifold and to the particular homology degree.

The same point is 0-essential for its restriction to $T(f_2 \leq 1)$, so it can be recovered by the standard graph-theoretical techniques used in degree 0, i.e. for size functions. (The two functions need not be so strictly related: $f_2$ could be replaced by the Euclidean distance from $(0, 0, 3)$ with the same effect). We conjecture that — at least whenever torsion is not involved — one can recursively take the $i$-essential values of a measuring function to $(i-1)$-essential ones, down to (easily computable) 0-essential critical values by means of other (auxiliary) measuring functions, as in this example.

As a referee suggested, the behavior of the torus with respect to $f = (f_1, f_2)$ can be described by a filtered sequence of modules associated to a multi-filtration of the manifold induced by $f$. Knudson in paragraph 6.1 of [20] faces a situation of a similar kind, but the solution of this problem goes beyond the purpose of the present paper.

Acknowledgements

The authors wish to thank C. Landi and P. Frosini for the many helpful suggestions, their indispensable support and friendship.
This work was performed under the auspices of INdAM-GNSAGA, CIRAM, ARCES and the University of Bologna funds for selected research topics.

REFERENCES


Dipartimento di Matematica, Università di Bologna, Piazza di Porta S. Donato 5, I-40126 Bologna, Italia

E-mail address: cagliari@dm.unibo.it

Dipartimento di Matematica, Università di Bologna, Piazza di Porta S. Donato 5, I-40126 Bologna, Italia

E-mail address: difabio@dm.unibo.it

Dipartimento di Matematica, Università di Bologna, Piazza di Porta S. Donato 5, I-40126 Bologna, Italia

E-mail address: ferri@dm.unibo.it