Comparison of Persistent Homologies for Vector Functions: from continuous to discrete and back

Niccolò Cavazza, Marc Ethier, Patrizio Frosini,
Tomasz Kaczynski, and Claudia Landi

January 16, 2012

Abstract

The theory of multidimensional persistent homology was initially developed in the discrete setting, and involved the study of simplicial complexes filtered through an ordering of the simplices. Later, stability properties of multidimensional persistence have been proved to hold when topological spaces are filtered by continuous functions, i.e. for continuous data. This paper aims to provide a bridge between the continuous setting, where stability properties hold, and the discrete setting, where actual computations are carried out. More precisely, a stability preserving method is developed to compare rank invariants of vector functions obtained from discrete data. These advances confirm that multidimensional persistent homology is an appropriate tool for shape comparison in computer vision and computer graphics applications. The results are supported by numerical tests.

Keywords: Multidimensional persistent homology; axis-wise interpolation; filtration; matching distance; topological aliasing
Mathematics Subject Classification (2010): 55-04; 65D18

1 Introduction

In this paper we present a discrete counterpart of the theory of persistent homology of vector functions that still guarantees stability properties as the continuous framework. The theory of multidimensional persistence was developed in the discrete setting in [8], and involved the study of simplicial complexes filtered through an ordering of the simplices. On the other hand stability properties of multidimensional persistence are proved to hold when triangulable spaces are filtered by continuous functions, i.e. for continuous data [20, 9]. This paper aims to be a bridge between the continuous setting, where stability properties hold, and the discrete setting, where actual computations are carried out. More precisely, we

develop a method to compare persistent homologies of vector functions obtained from discrete data. We show that in the passage from the continuous to the discrete framework stability is preserved. These advances support the appropriateness of multidimensional persistent homology for shape comparison by functions.

The problem of comparing shapes is well-studied in computer vision and computer graphics and many algorithms have been developed for this purpose. A widely used scheme is to associate a shape with a shape descriptor, or a signature, and comparing shapes by measuring dissimilarity between descriptors. An important class of shape descriptors, which may be called *shape-from-functions* methods, is based on the common idea of performing a topological exploration of the shape according to some quantitative geometric properties provided by a (measuring) function defined on the shape and chosen to extract shape features [4].

The simplest topological attribute of a space is the number of its connected components. A well-known mathematical tool to count the number of connected components is the homology group H_0. More complex topological features are revealed by higher homology groups.

Persistent homology is a shape-from-functions method for shape description involving homology groups of any degree. The idea is to filter a space by the sublevel sets of the function and to analyze the homological changes of the sublevel sets across this filtration, due to the appearance or disappearance of topological attributes, such as connected components. Features with a short persistence along the filtration can be regarded as negligible information due to noise or very fine details. For application purposes, it is often sufficient to disregard the group structure of persistent homology and retain only the rank information. This gives rise to the notions of *rank invariant* [8], persistent Betti numbers [19], size functions [25].

The topic has been widely studied in the case of filtrations induced by scalar continuous functions (i.e. one-dimensional persistence), especially in connection with the stability problem [13, 12, 14, 15].

This theory has been generalized to a multidimensional situation in which a vector-valued function characterizes the data as suggested in [17, 18]. Results in this area are given in [3, 8, 5, 9]. This generalization is quite natural in view of the analogous generalization of Morse Theory [24]. Moreover, it is motivated by applications where data are more completely described by more than one function (e.g., curvature and torsion for space curves).

The passage from scalar to vector-valued functions presents new challenges. To begin with, critical points are no longer isolated even in non-degenerate situations [17]. Although the relevant points for persistent homology of vector functions are a subset of the critical points, precisely the Pareto critical points, these are still non-isolated [11]. For example, in the case of the sphere $x^2+y^2+z^2=1$ with the function $f=(y,z)$, the Pareto critical points are those in the set $x=0$, $y^2+z^2=1$, $yz\geq 0$.

Another delicate issue is passing from the comparison of continuous models to that of discrete models. This is an essential passage, and the core of this pa-

per. Indeed, for two given real-world objects X and Y, modeled as triangulable topological spaces (e.g., manifolds), we usually only know simplicial descriptions $\mathcal{K}$ and $\mathcal{L}$ of them, affected by approximation errors. For example, acquiring 3D models of real-world objects for computer graphics applications needs to account for errors due to sensor resolution, noise in the measurements, inaccuracy of sensor calibration [1]. Moreover, different techniques for reconstructing the geometry and topology of the scanned object yield different polyhedral approximations. Analogous considerations hold for any continuous measuring functions $f : X \to \mathbb{R}^k$ and $g : Y \to \mathbb{R}^k$, because we could only consider approximations $\tilde{\varphi} : K \to \mathbb{R}^k$, $\tilde{\psi} : L \to \mathbb{R}^k$ of f and g defined on finite polyhedra. Depending on the context of a specific application, these functions may or may not be given by explicit formulas. In either case, it is legitimate to assume that we are able to compute their values on vertices of $\mathcal{K}$ and $\mathcal{L}$. Hence, we only know the discrete maps $\varphi : \mathcal{V}(\mathcal{K}) \to \mathbb{R}^k$, $\psi : \mathcal{V}(\mathcal{L}) \to \mathbb{R}^k$ which are the restrictions of $\tilde{\varphi}$ and $\tilde{\psi}$ to vertices. Therefore a natural question is whether shape comparison by persistent homology of vector functions is numerically stable, i.e. whether the computation of a distance between rank invariants of discrete models gives a good approximation of the ideal distance between rank invariants of continuous models.

Our main result, Theorem 4.5, gives an affirmative answer to this question. It states that, in the passage from continuous to discrete data, the distance between rank invariants does not increase, provided that stability holds for the continuous model. We underline that at least one stable distance between rank invariants of continuous vector functions exists as proved in [9]. In order to profit from the stability theory in the continuous case, we give a new construction of axis-wise linear interpolation $\varphi^{\urcorner}$ which is generic in the sense that its persistent homology is exactly equal to that of the map defined on vertices φ.

The paper is organized as follows. In Section 2 the necessary background notions concerning persistence are reviewed and put in the context of our aims. Section 3 starts with the description of the simplicial framework and with Example 3.1 which is a simplicial analogue of the sphere example pointed above. The same example shows that, in the vector case, the linear extension of a map defined on vertices does not satisfy the genericity property described above. Topological artifacts of an interpolation method have been observed before. This phenomenon can be referred to as *topological aliasing*. Our example motivates the construction of our axis-wise linear interpolation. We next prove Theorem 3.3 on the deformation retraction of continuous sublevel sets of $\varphi^{\urcorner}$ onto the simplicial sublevel sets of φ. We introduce the notion of homological critical value for vector functions. As in the sphere example, the set of critical values need not be discrete, but we prove in Theorem 3.5 that in the case of $\varphi^{\urcorner}$ it has to be contained in a finite union of hyperplanes, thus it is a nowhere dense set and its k-dimensional Lebesgue measure is zero (Corollary 3.6).

Section 4 starts with Lemma 4.1 that provides an approximation of a distance between the rank invariants of continuous functions by that of the rank invariants of the corresponding axis-wise linear approximations. The genericity of $\varphi^{\urcorner}$ allows us to introduce the rank invariant for φ. Although this rank invariant is defined

for a discrete function φ and computed using only simplicial sublevel sets, it takes pairs of real vectors as variables, as it is in the case of the rank invariant for continuous functions. We show that this new rank invariant for φ is equal to that of $\varphi^{\urcorner}$. This allows us to derive the main result of the paper (Theorem 4.5).

Section 5 describes an algorithm which computes an approximate matching distance. Our algorithm is a modification of the algorithm described in [2], adapted to the rank invariants. The correctness of the algorithm is guaranteed by the results of Section 4. We next present tests of the algorithm performed on simplicial models in the case $k = 2$. Our tests revealed the same discrepancy as observed in Example 3.1, thus providing numerical confirmation of topological aliasing.

2 Basic notions and working assumptions

Let us consider a triangulable topological space X (i.e., a space homeomorphic to the carrier of a finite simplicial complex). A *filtration* of X is a family $\mathcal{F} = \{X_\alpha\}_{\alpha\in\mathbb{R}^k}$ of subsets of X that are nested with respect to inclusions, that is: $X_\alpha \subseteq X_\beta$, for every $\alpha \preceq \beta$, where $\alpha \preceq \beta$ if and only if $\alpha_j \leq \beta_j$ for all $j = 1, 2, \ldots, k$.

Persistence is based on analyzing the homological changes occurring along the filtration as α varies. This analysis is carried out by considering, for $\alpha \preceq \beta$, the homomorphism

$$H_*(i^{(\alpha,\beta)}) : H_*(X_\alpha) \to H_*(X_\beta).$$

induced by the inclusion map $i^{(\alpha,\beta)} : X_\alpha \hookrightarrow X_\beta$. We work with Čech homology with coefficients in a given field $\mathbb{F}$. When each X_α, $\alpha \in \mathbb{R}^k$, is triangulable, it reduces to simplicial homology. For simplicity of notation we write $H_*(X_\alpha)$ for the graded homology space $H_*(X_\alpha;\mathbb{F}) = \{H_q(X_\alpha;\mathbb{F})\}_{q\in\mathbb{Z}}$. The choice of a field is only made in experimentations, the most convenient in computations being $\mathbb{F} = \mathbb{Z}_p$, with p a prime number. Thus, for any $q \in \mathbb{Z}$, $H_q(X_\alpha)$ is a vector space of dimension equal to the q'th Betti number of X_α.

The image of the map $H_q(i^{(\alpha,\beta)})$ is a vector space known as the *q'th persistent homology group* of the filtration at (α, β). It contains the homology classes of order q born not later than α and still alive at β. The dimension of this vector space is called a *q'th persistent Betti number*.

A rank invariant is a function that encodes the changes in the persistent Betti numbers as α and β vary. Setting

$$\Delta^k_+ := \{(\alpha, \beta) \in \mathbb{R}^k \times \mathbb{R}^k \mid \alpha \prec \beta\},$$

where $\alpha \prec \beta$ if and only if $\alpha_j < \beta_j$ for all $j = 1, 2, \ldots, k$, the *q'th rank invariant* of the filtration $\mathcal{F}$ is the function $\rho^q_{\mathcal{F}} : \Delta^k_+ \to \mathbb{N} \cup \{\infty\}$ defined on each pair $(\alpha, \beta) \in \Delta^k_+$ as the rank of the map $H_q(i^{(\alpha,\beta)})$. In other words, $\rho^q_{\mathcal{F}}(\alpha, \beta) = \dim \mathrm{im} H_q(i^{(\alpha,\beta)})$.

In this paper, we will use the notation $\rho_{\mathcal{F}}$ to refer to rank invariants of arbitrary order. Ultimately, the shapes of two triangulable spaces X and Y, filtered by $\mathcal{F}$

and $\mathcal{G}$, respectively, can be compared by using an (extended) distance D between their rank invariants $\rho_{\mathcal{F}}$ and $\rho_{\mathcal{G}}$.

The framework described so far for general filtrations can be specialized in various directions. We now review the two most relevant ones for our paper.

2.1 Persistence of sublevel set filtrations

Given a continuous function $f : X \to \mathbb{R}^k$, it induces on X the so-called *sublevel set filtration*, defined as follows:

$$X_\alpha = \{x \in X \mid f(x) \preceq \alpha\}.$$

We will call the function f a *measuring function* and denote the rank invariant associated with this filtration by ρ_f.

Since X is assumed to be triangulable and f is continuous, $\rho_f(\alpha, \beta) < +\infty$ for every $\alpha \prec \beta \in \mathbb{R}^k$ (cf. [6]).

Among all the (extended) distances D between rank invariants of filtrations, we confine our study to those ones that, when applied to sublevel set filtrations, satisfy the following *stability property*:

(S) For every $f, f' : X \to \mathbb{R}^k$ continuous functions, $\mathrm{D}(\rho_f, \rho_{f'}) \leq \|f - f'\|_\infty$.

In [9] it has been shown that there is at least one distance between rank invariants, the *matching distance*, that has the stability property (S). An analogous stability property for a distance defined between modules is presented in [22].

The matching distance will be used for computations in the experiments described in Section 5. Until then, we will not need to specify which distance D we are using, provided it satisfies (S).

2.2 Persistence of simplicial complex filtrations

We consider a simplicial complex $\mathcal{K}$ consisting of closed geometric simplices and its *carrier* defined by

$$K = |\mathcal{K}| := \bigcup \mathcal{K}. \tag{1}$$

The set of all vertices of $\mathcal{K}$ is denoted by $\mathcal{V}(\mathcal{K})$ or by $\mathcal{V}$, if $\mathcal{K}$ is clear from the context. For $\sigma, \tau \in \mathcal{K}$, the relation τ *is a face of* σ is denoted by $\tau \leq \sigma$. For proper faces, we write $\tau < \sigma$.

In this discrete setting, we take a family $\{\mathcal{K}_\alpha\}_{\alpha \in \mathbb{R}^k}$ of simplicial subcomplexes of $\mathcal{K}$, such that $\mathcal{K}_\alpha$ is a subcomplex of $\mathcal{K}_\beta$, for every $\alpha \preceq \beta$. As a consequence, their carriers are nested with respect to inclusions, that is: $K_\alpha \subseteq K_\beta$, yielding a filtration of K.

In the next section we address the following problem: is any simplicial complex filtration induced by a suitable continuous function?

A positive answer to this question will allow us later to transfer the stability property of D from the continuous to the discrete setting.

3 From continuous to discrete vector functions

We let $\varphi : \mathcal{V}(\mathcal{K}) \to \mathbb{R}^k$ be a vector-valued function defined on vertices. We suppose that φ is a *discretization* of some continuous function $\tilde{\varphi} : K \to \mathbb{R}^k$. Reciprocally, $\tilde{\varphi}$ is an *interpolation* of φ. In this section we will simply assume that $\tilde{\varphi}$ is equal to φ on vertices of $\mathcal{K}$ but, of course, when it comes to computing, one has to set bounds for the rounding error. Although in some practical applications of persistent homology to the analysis of discrete multidimensional data $\tilde{\varphi} : K \to \mathbb{R}^k$ may be explicitly known, in some other cases we do not even have an explicit formula for $\tilde{\varphi}$: we assume that such a function exists, that we can estimate its modulus of uniform continuity (for the sake of simplicity, say, its Lipschitz constant), and that we can compute the values of $\tilde{\varphi}$ at grid points of arbitrary fine finite grids.

In a discrete model, we are interested in simplicial sublevel complexes

$$\mathcal{K}_\alpha := \{\sigma \in \mathcal{K} \mid \varphi(v) \preceq \alpha \text{ for all vertices } v \leq \sigma\}.$$

In Section 5, we compute the rank invariants for the discrete vector-valued function φ and we use this information for computing the distance between rank invariants for their continuous interpolations. In order to do this, we need to know that there exists a continuous function which is a *generic interpolation* of φ, in the sense that its rank invariant is exactly equal to that of φ.

In the case $k = 1$, that is, when φ has values in $\mathbb{R}$, it can be shown that such an interpolation can be obtained by extending φ to each simplex $\sigma \in \mathcal{K}$ by linearity. We shall denote this interpolation by $\overline{\varphi}$. In that case, one can show that K_α is a deformation retract of $K_{\overline{\varphi} \preceq \alpha}$, so the inclusion of one set into another induces an isomorphism in homology. This result belongs to "mathematical folklore": it is often implicitly used in computations without being proved. The arguments for that case are outlined in [23, Section 2.5] and Theorem 3.3 we prove in this section contains this result as a special case. Unfortunately, if $k > 1$, this result is no longer true as the following example shows:

Example 3.1 Let K be the boundary of the tetrahedron shown in Figure 1, homeomorphic to the 2D sphere. The corresponding simplicial complex $\mathcal{K}$ is made of all proper faces of the 3D simplex $[v_0, v_1, v_2, v_3]$ in $\mathbb{R}^3$, with vertices $v_0 = (0, 0, 0)$, $v_1 = (1, 0, 0)$, $v_2 = (0, 1, 0)$, $v_3 = (1/2, 0, 1)$. Hence $K = |\mathcal{K}|$ is homeomorphic to a 2D sphere. Let $\varphi : K \to \mathbb{R}^2$ be the restriction of the linear function $\overline{\varphi}$ given by $\overline{\varphi}(x, y, z) = (x, z)$ to the four vertices. Let $\alpha \in \mathbb{R}^2$ be any value chosen so that $1/2 < \alpha_1 < 1$ and $\alpha_2 = 2 - 2\alpha_1$. It is easy to see that $K_\alpha = [v_0, v_2]$. Its homology is trivial. Note that the set $K_{\overline{\varphi} \preceq \alpha}$ contains one point x on the edge $[v_1, v_3]$, namely $x = (\alpha_1, 0, \alpha_2)$, which closes a non-contractible path in $K_{\overline{\varphi} \preceq \alpha}$. We have $H_1(K_{\overline{\varphi} \preceq \alpha}) \cong \mathbb{F} \neq 0$.

The discrepancy between the discrete and linear interpolated models seen in Example 3.1 has been observed in applications to computer graphics and imaging, and has been recently referred to as topological aliasing.

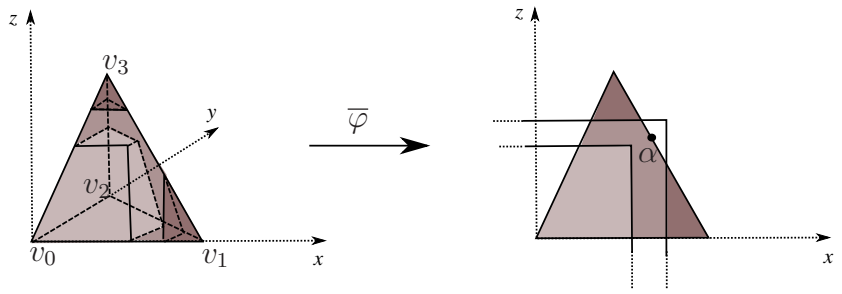

Figure 1: The tetrahedron boundary and sketches of two sublevel sets of the linear interpolation $\overline{\varphi}$ discussed in Example 3.1. The values taken at the displayed edge are critical.

Several interesting conclusions can be derived from this. First, K_α is not a deformation retract of $K_{\overline{\varphi} \preceq \alpha}$. This remains true if we slightly increase the value of α. Secondly, if we slightly decrease α, the set K_α does not change but the set $K_{\overline{\varphi} \preceq \alpha}$ becomes contractible. Hence, in the sense of Definition 3.4 presented further in this section, any value assumed at a point of the edge $[v_1, v_3]$ is a homological critical value. In particular, the set of such values may be uncountable. This is in contrast with the one-dimensional case, where a piecewise-linear function on a simplicial complex must have a discrete set of critical values.

We shall now construct a continuous function $\varphi^\urcorner : K \to \mathbb{R}^k$ called *axis-wise linear interpolation* of φ which will correct the problem encountered with the linear interpolation $\overline{\varphi}$ in the multidimensional case. First, given any $\sigma \in \mathcal{K}$, let $\mu(\sigma) \in \mathbb{R}^k$ be defined by

$$\mu_j(\sigma) = \max\{\varphi_j(v) \mid v \text{ is a vertex of } \sigma\}, j = 1, 2, \ldots, k. \tag{2}$$

Note that if $\tau \leq \sigma$, then $\mu(\tau) \preceq \mu(\sigma)$.

We will use induction on the dimension m of σ to define $\varphi^\urcorner : K \to \mathbb{R}^k$ on σ and a point $w_\sigma \in \sigma$ with the following properties:

(a) For all $x \in \sigma$, $\varphi^\urcorner(x) \preceq \varphi^\urcorner(w_\sigma) = \mu(\sigma)$;

(b) $\varphi^\urcorner$ is linear on any line segment $[w_\sigma, y]$ with y on the boundary of σ.

If $m = 0$, so that $\sigma = \{v\}$ is a vertex, $\varphi^\urcorner(v) = \varphi(v)$ and we put $w_{\{v\}} = v$. Let $m > 0$ and suppose $\varphi^\urcorner$ is constructed on simplices of lower dimensions. Let τ be a minimal face of σ such that $\mu(\tau) = \mu(\sigma)$. Consider two cases.

(i) If $\tau \neq \sigma$, then w_τ and $\varphi^\urcorner(w_\tau)$ are defined by the induction step. We put $w_\sigma = w_\tau$. Since σ is convex, any x in the interior of σ is on a line segment joining w_σ to a uniquely defined $y(x)$ on the boundary of σ. Since $\varphi^\urcorner(y(x))$ is defined by the induction step, we extend $\varphi^\urcorner$ to $[w_\sigma, y(x)]$ by linearity.

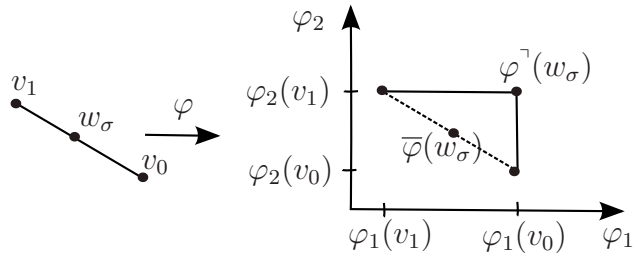

Figure 2: The linear (dashed line) and axis-wise linear (continuous line) interpolations of a function φ defined on the vertices of the simplex $\sigma = [v_0, v_1]$ with values in $\mathbb{R}^2$.

(ii) If $\tau = \sigma$, then let w_σ be the barycenter of σ and put $\varphi^\urcorner(w_\sigma) = \mu(\sigma)$. Again, any $x \neq w_\sigma$ in the interior of σ is on a line segment joining w_σ to a uniquely defined $y(x)$ on the boundary of σ and we proceed as before.

The property (a) follows from the fact that $\mu(\tau) \preceq \mu(\sigma)$ when $\tau \leq \sigma$, and from the linearity on joining segments. The property (b) is clear from the construction. By routine arguments from convex analysis, the point $y(x)$ on the boundary of σ is a continuous function of $x \in \sigma \setminus \{w_\sigma\}$, and the constructed function is continuous on σ. Since we proceeded by induction on the dimension of σ, the definitions on any two simplices coincide on their common face, so $\varphi^\urcorner$ extends continuously to K. The property (b) implies that if $k = 1$, and in certain cases of vector valued functions, $\varphi^\urcorner$ is equal to $\overline{\varphi}$, namely:

(c) $\varphi^\urcorner$ is piecewise linear on each simplex σ. In addition, if w_τ is a vertex of τ for each $\tau \leq \sigma$, then it is linear on σ.

The difference between the piecewise linear and the axis-wise linear interpolations $\overline{\varphi}$ and $\varphi^\urcorner$ is illustrated in Figure 2 for a 1-simplex $\sigma = [v_0, v_1]$ and a function φ defined on vertices.

Lemma 3.2 *The following statements hold:*

(i) For any $\alpha \in \mathbb{R}^k$, $K_\alpha \subset K_{\varphi^\urcorner \preceq \alpha}$.

(ii) Let $\sigma \in \mathcal{K}$ and $\alpha \in \mathbb{R}^k$. If $\sigma \cap K_{\varphi^\urcorner \preceq \alpha} \neq \emptyset$, then σ has at least one vertex in K_α.

PROOF: *(i)* Let $\sigma \in \mathcal{K}_\alpha$. It is clear from (2) that $\mu(\sigma) \preceq \alpha$. It follows from the property (a) in the definition of $\varphi^\urcorner$ that $\sigma \subset K_{\varphi^\urcorner \preceq \alpha}$.

(ii) We follow the induction steps in the construction of $\varphi^\urcorner$. If $\dim(\sigma) = 0$, σ is a vertex and there is nothing to prove. Let $\dim(\sigma) = m > 0$ and suppose the statement is proved for lower dimensions. Let $x \in \sigma \cap K_{\varphi^\urcorner \preceq \alpha}$. If $x = w_\sigma$, then $w_\sigma \in K_{\varphi^\urcorner \preceq \alpha}$. By the property (a) of $\varphi^\urcorner$, all σ is in K_α. If $x \neq w_\sigma$, then x is on a

line segment joining the point w_σ of σ with a point $y(x)$ of an $(m-1)$–simplex $\tau < \sigma$, where $\varphi^\urcorner(y(x))$ is defined by the induction hypothesis. We know that $\varphi^\urcorner$ is extended linearly to the line segment $[w_\sigma, y(x)]$. Also, $\varphi^\urcorner(y(x)) \preceq \varphi^\urcorner(w_\sigma)$ by the property (a). Hence $\varphi^\urcorner(y(x)) \preceq \varphi^\urcorner(x) \preceq \alpha$. It remains to use the induction hypothesis for $y(x)$ and τ to deduce that τ has a vertex in K_α. □

Theorem 3.3 *For any $\alpha \in \mathbb{R}^k$, K_α is a strong deformation retract of $K_{\varphi^\urcorner \preceq \alpha}$. Consequently, the inclusion $K_\alpha \hookrightarrow K_{\varphi^\urcorner \preceq \alpha}$ induces an isomorphism in homology.*

PROOF: Note that $K_{\varphi^\urcorner \preceq \alpha}$ is contained in a union of simplices $\sigma \in \mathcal{K}$ such that

$$\sigma_{\varphi^\urcorner \preceq \alpha} := \sigma \cap K_{\varphi^\urcorner \preceq \alpha} \neq \emptyset. \tag{3}$$

Given any such σ, consider the simplex σ_α defined as the convex hull of the set of vertices v of σ such that $\varphi(v) \preceq \alpha$. By the hypothesis on σ and by Lemma 3.2*(ii)*, $\sigma_\alpha \neq \emptyset$. Given any $\sigma \in \mathcal{K}$ for which $\sigma_{\varphi^\urcorner \preceq \alpha} \neq \emptyset$, we shall define a strong deformation retraction

$$H_\sigma : \sigma_{\varphi^\urcorner \preceq \alpha} \times [0,1] \to \sigma_{\varphi^\urcorner \preceq \alpha}$$

with $r = H(\cdot, 1)$ being a retraction of $\sigma_{\varphi^\urcorner \preceq \alpha}$ onto σ_α.

The construction goes by induction on the dimension m of σ following the induction steps in the construction of the function $\varphi^\urcorner$. If $\dim(\sigma) = 0$, σ is a vertex and there is nothing to prove. Now let $m > 0$. Suppose that the deformation retraction $H_\tau : \tau_{\varphi^\urcorner \preceq \alpha} \times [0,1] \to \tau_{\varphi^\urcorner \preceq \alpha}$ is defined for simplices τ of dimension $m' < m$ with $\tau_{\varphi^\urcorner \preceq \alpha} \neq \emptyset$ in such a way that $H_\tau(x,t) = x$ for any $(x,t) \in \tau_{\varphi^\urcorner \preceq \alpha} \times [0, 1/2^{m'}]$, and the values of $\varphi^\urcorner$ on $H_\tau(x,t)$ are decreasing with t. By "decreasing" we mean the weak inequality "$\preceq$". This hypothesis guarantees that the deformation has values in the set $\tau_{\varphi^\urcorner \preceq \alpha}$.

Let $x \in \sigma_{\varphi^\urcorner \preceq \alpha}$. If x is on a boundary of σ, we define $H_\sigma(x,t) = H_\tau(x,t)$, where τ is the smallest face of σ containing x and H_τ is defined by the induction hypothesis. Suppose x is in the interior of σ. Let w_σ and $y(x)$ be the points identified in the definition of $\varphi^\urcorner$. Note that, if $\varphi^\urcorner(w_\sigma) \preceq \alpha$, then $\sigma_{\varphi^\urcorner \preceq \alpha} = \sigma$, hence the deformation must be defined as the identity map for each t:

$$H_\sigma(x,t) := x \text{ for all } (x,t) \in \sigma \times [0,1].$$

Therefore, we may suppose that $w_\sigma \notin \sigma_{\varphi^\urcorner \preceq \alpha}$. Consider the smallest face τ of σ containing $y(x)$. Since $y(x)$ is on the boundary of σ, τ is a proper face of σ of dimension, say, $m' < m$. By the construction of $\varphi^\urcorner$,

$$\varphi^\urcorner(y(x)) \preceq \varphi^\urcorner(x) \preceq \varphi^\urcorner(w_\sigma). \tag{4}$$

Since $\varphi^\urcorner(x) \preceq \alpha$, we get $\varphi^\urcorner(y(x)) \preceq \alpha$ so $y(x) \in \tau_{\varphi^\urcorner \preceq \alpha} \neq \emptyset$. By the induction hypothesis, a deformation retraction

$$H_\tau : \tau_{\varphi^\urcorner \preceq \alpha} \times [0,1] \to \tau_{\varphi^\urcorner \preceq \alpha} \subset \sigma_{\varphi^\urcorner \preceq \alpha}$$

is defined so that the values of $\varphi^\urcorner$ on $H_\tau(x,t)$ decrease with t, and $H_\tau(x,t) = x$ for $t \in [0, 1/2^{m'}]$.

For any $t \in [0,1]$ and for x in the interior of σ we define

$$H_\sigma(x,t) := \begin{cases} x & \text{if } 0 \le t < 1/2^m \\ (2^m t - 1)y(x) - (2^m t - 2)x & \text{if } 1/2^m \le t < 1/2^{m-1} \\ H_\tau(y(x), t) & \text{if } 1/2^{m-1} \le t \le 1 \end{cases} .$$

It is easily checked that $H_\sigma(x, 1/2^m) = x$, $H_\sigma(x, 1/2^{m-1}) = y(x)$. Since $\varphi^\urcorner$ is linear on $[w_\sigma, y(x)]$, the inequality (4) implies that the values of $\varphi^\urcorner$ on $H_\sigma(x,t)$ decrease with t.

Thus we have defined H_σ both when x is on the boundary of σ and when it is in the interior of σ. By construction, for every $x \in \sigma_{\varphi^\urcorner \preceq \alpha}$, $H_\sigma(x,0) = x$, and $H_\sigma(x,1)$ belongs to σ_α, and moreover, for every $x \in \sigma_\alpha$, $H_\sigma(x,1) = x$. In order to conclude that H_σ is a deformation retraction of $\sigma_{\varphi^\urcorner \preceq \alpha}$ onto σ_α we must prove that H_σ is continuous. The continuity at a given point (x_0, t_0) with x_0 in the interior of σ follows from the continuity of $y(x)$ in x. The continuity at (x_0, t_0) with x_0 on the boundary of σ follows from the condition that $H_\tau(x,t) = x$ for any $t \in [0, 1/2^{m'}]$ and from the induction hypothesis.

In order to continuously extend H_σ to a deformation

$$H : K_{\varphi^\urcorner \preceq \alpha} \times [0,1] \to K_{\varphi^\urcorner \preceq \alpha},$$

it is enough to prove that, given two simplices σ_1 and σ_2 intersecting $K_{\varphi^\urcorner \preceq \alpha}$ and $\tau = \sigma_1 \cap \sigma_2$, the maps H_{σ_1} and H_{σ_2} agree at any $x \in \tau_{\varphi^\urcorner \preceq \alpha}$. It is clear from the definition that $H_{\sigma_1}(x,t) = H_{\sigma_2}(x,t) = H_\tau(x,t)$ for $x \in \tau$ and for all t, provided that H_τ is defined. But this is true, because $x \in \tau_{\varphi^\urcorner \preceq \alpha}$, so this is a nonempty set. □

In the next section, we use Theorem 3.3 to show that any distance between rank invariants of continuous functions that has property (S) can be approximated by the distance between rank invariants of discrete functions. We end this section with another application of Theorem 3.3 of interest in itself: a theorem on the structure of the set of critical values of the axis-wise interpolation $\varphi^\urcorner$. The following definition generalizes the notion of homological critical value given in [13] to vector functions. In plain words we call homological critical any value α for which any sufficiently small neighborhood contains two values whose sublevel sets are included one into the other but cannot be retracted one onto the other. Neighborhoods are taken with respect to the norm $\|\alpha\| = \max_{j=1,2,\ldots,k} |\alpha_j|$ in $\mathbb{R}^k$.

Definition 3.4 Let $\tilde{\varphi} : K \to \mathbb{R}^k$ be a continuous vector function. A value $\alpha \in \mathbb{R}^k$ is a *homological critical value* of $\tilde{\varphi}$ if there exists an integer q such that, for all sufficiently small real values $\epsilon > 0$, two values $\alpha', \alpha'' \in \mathbb{R}^k$ can be found with $\alpha' \preceq \alpha \preceq \alpha''$, $\|\alpha' - \alpha\| < \epsilon$, $\|\alpha'' - \alpha\| < \epsilon$, such that the map

$$H_q(K_{\tilde{\varphi} \preceq \alpha'}) \to H_q(K_{\tilde{\varphi} \preceq \alpha''})$$

induced by the inclusion $K_{\tilde{\varphi} \preceq \alpha'} \hookrightarrow K_{\tilde{\varphi} \preceq \alpha''}$ is not an isomorphism. If this condition fails, α is called a *homological regular value*.

Also note that, by the long exact sequence for the relative homology (see e.g. [21, Chapter 9]) the critical value definition is equivalent to the following condition on the graded relative homology

$$H_*(K_{\tilde{\varphi}\preceq\alpha'}, K_{\tilde{\varphi}\preceq\alpha''}) \neq 0.$$

For any $j = 1, 2, \ldots, k$ and a vertex $v \in \mathcal{V}(\mathcal{K})$, consider the hyperplane of $\mathbb{R}^k$ given by the equation $\alpha_j = \varphi^\urcorner{}_j(v)$ and a positive closed cone $C_j(v)$ contained in it, given by the formula

$$C_j(v) := \{\alpha \in \mathbb{R}^k \mid \alpha_j = \varphi^\urcorner{}_j(v) \text{ and } \alpha_i \geq \varphi^\urcorner{}_i(v) \text{ for all } i = 1, 2, \ldots, k\}.$$

Theorem 3.5 *The set of homological critical values of $\varphi^\urcorner$ is contained in the finite union of the described cones, namely, in the set*

$$C := \bigcup \{C_j(v) \mid v \in \mathcal{V}(\mathcal{K}) \textit{ and } j = 1, 2, \ldots, k\}.$$

PROOF: Consider any $\alpha \notin C$. We need to show that α is a homological regular value. Since C is a closed set, an $\epsilon > 0$ exists such that the set $\bar{Q}(\alpha, \epsilon) = \{\beta \in \mathbb{R}^k \mid \|\alpha - \beta\| \leq \epsilon\}$ does not meet C. If $\|\alpha - \beta\| \leq \epsilon$, then

$$K_\beta = K_\alpha. \tag{5}$$

Indeed, if this were not true, the segment joining α and β should contain a point of C, against the choice of ϵ.

Now, let us assume that $\alpha' \preceq \varphi(v) \preceq \alpha''$, $\|\alpha - \alpha'\| \leq \epsilon$ and $\|\alpha - \alpha''\| \leq \epsilon$. It follows from equation (5) and from Theorem 3.3 that the inclusions $i' : K_\alpha = K_{\alpha'} \hookrightarrow K_{\varphi^\urcorner \preceq \alpha'}$ and $i'' : K_\alpha = K_{\alpha''} \hookrightarrow K_{\varphi^\urcorner \preceq \alpha''}$ induce isomorphisms in homology. The inclusion $i^{(\alpha', \alpha'')} : K_{\varphi^\urcorner \preceq \alpha'} \hookrightarrow K_{\varphi^\urcorner \preceq \alpha''}$ can be written as $i^{(\alpha', \alpha'')} = i'' \circ r'$, where r' is the retraction homotopically inverse to i'. By the functoriality of homology, $H_*(i^{(\alpha', \alpha'')}) = H_*(i'') \circ H_*(r')$, hence it is also an isomorphism. □

For the sake of visualization, in Figure 3 the set C is shown in a simple case.

From the formula for C, we instantly get an analogy of a well-known result from differential geometry [24].

Corollary 3.6 *The set of homological critical values of $\varphi^\urcorner$ is a nowhere dense set in $\mathbb{R}^k$. Moreover its k-dimensional Lebesgue measure is zero.*

4 Approximation of distances between rank invariants

The goal of this section is to show that shape comparison by persistent homology of vector functions is numerically stable. In this passage from real (continuous) objects to their discretizations, the approximation error does not grow to be much larger when we compute the distance D between the rank invariants of φ and ψ

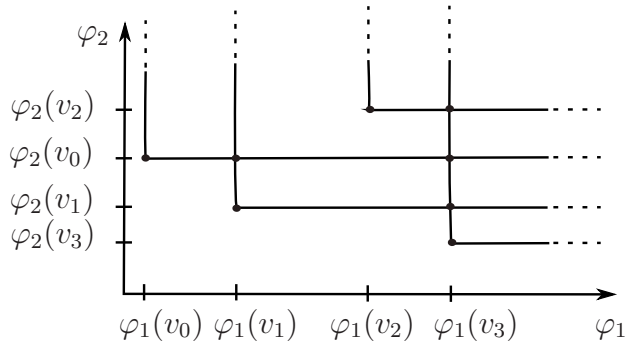

Figure 3: The set C defined in Theorem 3.5 is the union of closed cones with vertices at the values taken by φ. The set Λ introduced in Proposition 4.6 is a finite set whose elements are the highlighted points.

instead of the distance D between the rank invariants of f and g (Theorem 4.5). To this end, the stability property (S) of D defined in Section 2.1 in the continuous setting is crucial.

A description of this approximation procedure in concrete examples together with experiments exploiting the numerical stability of the comparison by persistent homology will be given in Section 5.

We end the section by showing that the set of homological critical values, although uncountable, admits a finite representative set.

We start from the following approximation lemma. It may happen that $\mathrm{D}\left(\rho_{\tilde{\varphi}}, \rho_{\tilde{\psi}}\right)$ and $\mathrm{D}\left(\rho_{\varphi^\urcorner}, \rho_{\psi^\urcorner}\right)$ are equal to $+\infty$. In the case of the matching distance, this occurs when $H_*(K) \neq H_*(L)$. In such a case, we adopt the convention $\infty - \infty = 0$.

Lemma 4.1 *Let $\tilde{\varphi} : K \to \mathbb{R}^k$, $\tilde{\psi} : L \to \mathbb{R}^k$ be two continuous measuring functions on the carrier of complexes $\mathcal{K}$ and $\mathcal{L}$. For any $\epsilon > 0$, there exists $\delta > 0$ such that if*

$$\max\{\mathrm{diam}\,\sigma \mid \sigma \in \mathcal{K} \text{ or } \sigma \in \mathcal{L}\} < \delta \tag{6}$$

then

$$\left|\mathrm{D}\left(\rho_{\tilde{\varphi}}, \rho_{\tilde{\psi}}\right) - \mathrm{D}\left(\rho_{\varphi^\urcorner}, \rho_{\psi^\urcorner}\right)\right| < \epsilon. \tag{7}$$

PROOF: Since K and L are compact, $\tilde{\varphi}$, $\tilde{\psi}$, $\varphi^\urcorner$, and $\psi^\urcorner$ are uniformly continuous. Hence for any $\epsilon > 0$, there exists $\delta > 0$ such that if (6) is satisfied then

$$\max\{\mathrm{diam}\,\tilde{\varphi}(\sigma) \mid \sigma \in \mathcal{K}\} < \epsilon/4 \tag{8}$$

and the same inequality holds for $\tilde{\psi}$, $\varphi^\urcorner$, and $\psi^\urcorner$. The diameters of $\tilde{\varphi}(\sigma)$ and σ are measured with respect to the maximum norm in the respective ambient spaces. Since φ is the restriction of $\tilde{\varphi}$ to the vertices, and $\varphi^\urcorner$ interpolates φ on the vertices, given any $x \in \sigma \in \mathcal{K}$, and any vertex v of σ, from (8) we get

$$\|\tilde{\varphi}(x) - \varphi^\urcorner(x)\| \leq \|\tilde{\varphi}(x) - \varphi(v)\| + \|\varphi(v) - \varphi^\urcorner(x)\| < \epsilon/2. \tag{9}$$

Hence, by the choice of the maximum norm in $\mathbb{R}^k$, $\|\tilde{\varphi} - \varphi^{\urcorner}\|_\infty < \epsilon/2$. By the same arguments, $\|\tilde{\psi} - \psi^{\urcorner}\|_\infty < \epsilon/2$. By the stability property (S) of D,

$$\begin{aligned} \mathrm{D}\left(\rho_{\tilde{\varphi}}, \rho_{\tilde{\psi}}\right) &\leq \mathrm{D}\left(\rho_{\tilde{\varphi}}, \rho_{\varphi^{\urcorner}}\right) + \mathrm{D}\left(\rho_{\varphi^{\urcorner}}, \rho_{\psi^{\urcorner}}\right) + \mathrm{D}\left(\rho_{\psi^{\urcorner}}, \rho_{\tilde{\psi}}\right) \\ &\leq \|\tilde{\varphi} - \varphi^{\urcorner}\|_\infty + \mathrm{D}\left(\rho_{\varphi^{\urcorner}}, \rho_{\psi^{\urcorner}}\right) + \|\psi^{\urcorner} - \tilde{\psi}\|_\infty \\ &< \mathrm{D}\left(\rho_{\varphi^{\urcorner}}, \rho_{\psi^{\urcorner}}\right) + \epsilon. \end{aligned}$$

Reversing the roles of $\tilde{\varphi}$, $\tilde{\psi}$ and $\varphi^{\urcorner}$, $\psi^{\urcorner}$, we get $\mathrm{D}\left(\rho_{\varphi^{\urcorner}}, \rho_{\psi^{\urcorner}}\right) < \mathrm{D}\left(\rho_{\tilde{\varphi}}, \rho_{\tilde{\psi}}\right) + \epsilon$ and the conclusion follows. □

Knowing Lemma 4.1, we now turn our attention to computing $\mathrm{D}\left(\rho_{\varphi^{\urcorner}}, \rho_{\psi^{\urcorner}}\right)$.

The following definition sets the notation for the rank invariant of the simplicial complex filtration obtained from a discrete map φ. Next, we show that this definition gives a rank invariant coinciding with the rank invariant of the continuous function $\varphi^{\urcorner}$. Thus it is a first step in the passage from the stability of rank invariants for continuous functions to that of discrete ones. Moreover, this definition is the one which we use to implement the reduction algorithm of [5] in our computations in Section 5.

Definition 4.2 Consider the discrete map $\varphi : \mathcal{V}(\mathcal{K}) \to \mathbb{R}^k$ defined on vertices of a simplicial complex $\mathcal{K}$. The *q'th real space variable rank invariant* or, shortly, *q'th real rank invariant* of φ is the function $\rho^q_\varphi : \Delta^k_+ \to \mathbb{N}$ defined on each pair $(\alpha, \beta) \in \Delta^k_+$ as the rank of the map

$$H_q(j^{(\alpha,\beta)}) : H_q(K_\alpha) \to H_q(K_\beta)$$

induced by the inclusion map $j^{(\alpha,\beta)} : K_\alpha \hookrightarrow K_\beta$ on simplicial sublevel sets.

Theorem 4.3 *Given any discrete function $\varphi : \mathcal{V}(\mathcal{K}) \to \mathbb{R}^k$ on the set of vertices of a simplicial complex $\mathcal{K}$ and its axis-wise interpolation $\varphi^{\urcorner}$, we have the equality of q'th real rank invariants*

$$\rho^q_\varphi = \rho^q_{\varphi^{\urcorner}}.$$

PROOF: Consider any $(\alpha, \beta) \in \Delta^k_+$, the inclusion maps $i^{(\alpha,\beta)} : K_{\varphi^{\urcorner} \leq \alpha} \hookrightarrow K_{\varphi^{\urcorner} \leq \beta}$, and $j^{(\alpha,\beta)} : K_\alpha \hookrightarrow K_\beta$. Theorem 3.3 implies that for every $q \in \mathbb{Z}$ we have the following commutative diagram

$$\begin{array}{ccc} H_q(K_{\varphi^{\urcorner} \preceq \alpha}) & \xrightarrow{H_q(i^{(\alpha,\beta)})} & H_q(K_{\varphi^{\urcorner} \preceq \beta}) \\ \downarrow \cong & & \downarrow \cong \\ H_q(K_\alpha) & \xrightarrow{H_q(j^{(\alpha,\beta)})} & H_q(K_\beta) \end{array}$$

where the vertical arrows are the isomorphisms induced by the corresponding retractions. Thus rank $H_q(i^{(\alpha,\beta)}) =$ rank $H_q(j^{(\alpha,\beta)})$.

□

In the sequel, we will once again use ρ_φ to refer to real rank invariants of arbitrary order. In conclusion we obtain that the distance between the rank invariants of two measuring (or interpolation) functions can be approximated using only the corresponding simplicial sublevel sets.

Corollary 4.4 *Let $\tilde{\varphi} : K \to \mathbb{R}^k$, $\tilde{\psi} : L \to \mathbb{R}^k$ be two continuous measuring functions on the carriers of complexes $\mathcal{K}$ and $\mathcal{L}$ and let $\varphi : \mathcal{V}(\mathcal{K}) \to \mathbb{R}^k$, $\psi : \mathcal{V}(\mathcal{L}) \to \mathbb{R}^k$ be the discretizations of $\tilde{\varphi}$ and $\tilde{\psi}$ on the sets of vertices of $\mathcal{K}$ and $\mathcal{L}$, respectively. For any $\epsilon > 0$ there exists $\delta > 0$ such that if*

$$\max\{\operatorname{diam}\sigma \mid \sigma \in \mathcal{K} \text{ or } \sigma \in \mathcal{L}\} < \delta$$

then

$$|\mathrm{D}\,(\rho_{\tilde{\varphi}}, \rho_{\tilde{\psi}}) - \mathrm{D}\,(\rho_\varphi, \rho_\psi)| < \epsilon. \tag{10}$$

PROOF: Immediate from Lemma 4.1 and Theorem 4.3. □

We are now ready to give the main result of this section.

Theorem 4.5 *Let X and Y be homeomorphic triangulable topological spaces, and let $f : X \to \mathbb{R}^k$, $g : Y \to \mathbb{R}^k$ be continuous functions. Let $(K, \tilde{\varphi})$ and $(L, \tilde{\psi})$, with K and L carriers of complexes $\mathcal{K}'$ and $\mathcal{L}'$, and $\tilde{\varphi} : K \to \mathbb{R}^k$, $\tilde{\psi} : L \to \mathbb{R}^k$ continuous measuring functions, approximate (X, f) and (Y, g), respectively, in the following sense: For a fixed $\epsilon > 0$, there exist a homeomorphism $\xi : K \to X$ with $\|\tilde{\varphi} - f \circ \xi\|_\infty \leq \epsilon/4$ and a homeomorphism $\zeta : L \to Y$ with $\|\tilde{\psi} - g \circ \zeta\|_\infty \leq \epsilon/4$. Then, for any sufficiently fine subdivision $\mathcal{K}$ of $\mathcal{K}'$ and $\mathcal{L}$ of $\mathcal{L}'$,*

$$|\mathrm{D}\,(\rho_f, \rho_g) - \mathrm{D}\,(\rho_\varphi, \rho_\psi)| \leq \epsilon,$$

$\varphi : \mathcal{V}(\mathcal{K}) \to \mathbb{R}^k$, $\psi : \mathcal{V}(\mathcal{L}) \to \mathbb{R}^k$ being restrictions of $\tilde{\varphi}$ and $\tilde{\psi}$ on the set of vertices of $\mathcal{K}$ and $\mathcal{L}$, respectively.

PROOF: By the triangle inequality

$$\mathrm{D}\,(\rho_f, \rho_g) \leq \mathrm{D}\,(\rho_f, \rho_{f\circ\xi}) + \mathrm{D}\,(\rho_{f\circ\xi}, \rho_{\tilde{\varphi}}) + \mathrm{D}\,(\rho_{\tilde{\varphi}}, \rho_{\tilde{\psi}}) + \mathrm{D}\,(\rho_{\tilde{\psi}}, \rho_{g\circ\zeta}) + \mathrm{D}\,(\rho_{g\circ\zeta}, \rho_g).$$

Since $\rho_f = \rho_{f\circ\xi}$ and $\rho_g = \rho_{g\circ\zeta}$, we have $\mathrm{D}\,(\rho_f, \rho_{f\circ\xi}) = 0$ and $\mathrm{D}\,(\rho_{g\circ\zeta}, \rho_g) = 0$. Moreover, by the stability property (S), since $\|\tilde{\varphi} - f \circ \xi\|_\infty \leq \epsilon/4$ and $\|\tilde{\psi} - g \circ \zeta\|_\infty \leq \epsilon/4$, we have $\mathrm{D}\,(\rho_{f\circ\xi}, \rho_{\tilde{\varphi}}) \leq \epsilon/4$ and $\mathrm{D}\,(\rho_{\tilde{\psi}}, \rho_{g\circ\zeta}) \leq \epsilon/4$. Therefore,

$$\mathrm{D}\,(\rho_f, \rho_g) \leq \mathrm{D}\,(\rho_{\tilde{\varphi}}, \rho_{\tilde{\psi}}) + \epsilon/2.$$

By Corollary 4.4, there exists $\delta > 0$ such that, if $\mathcal{K}$ and $\mathcal{L}$ are subdivisions of $\mathcal{K}'$ and $\mathcal{L}'$ with $\max\{\operatorname{diam}\sigma \mid \sigma \in \mathcal{K} \text{ or } \sigma \in \mathcal{L}\} < \delta$, then $\mathrm{D}\,(\rho_{\tilde{\varphi}}, \rho_{\tilde{\psi}}) \leq \mathrm{D}\,(\rho_\varphi, \rho_\psi) + \epsilon/2$. In conclusion we have proved that $\mathrm{D}\,(\rho_f, \rho_g) \leq \mathrm{D}\,(\rho_\varphi, \rho_\psi) + \epsilon$.

Reversing the roles of f, g and φ, ψ, we get $\mathrm{D}\,(\rho_\varphi, \rho_\psi) \leq \mathrm{D}\,(\rho_f, \rho_g) + \epsilon$, yielding the claim. □

We turn now to the question of the structure of the critical set of $\varphi^{\urcorner}$. Recall from the previous section that when $k > 1$, the set of homological critical values of a function on K with values in $\mathbb{R}^k$ may be an uncountable set, although contained in a nowhere dense set C by Theorem 3.5. However, the family $\{K_\alpha\}_{\alpha\in\mathbb{R}^k}$ of all subcomplexes of $\mathcal{K}$ is finite. Thus there exists a finite representative set $\Lambda \subset \mathbb{R}^k$ for C as the following proposition states.

Proposition 4.6 *For any $\alpha \in C$, there exists λ in*

$$\Lambda = \{\lambda \in C \mid \forall\, j = 1, 2, \ldots, k,\ \exists\, v \in \mathcal{V}(\mathcal{K}) \,:\, \lambda_j = \varphi_j(v)\}$$

such that $K_\alpha = K_\lambda$.

PROOF: Since $\alpha \in C$, $\mathcal{V}(\mathcal{K}_\alpha) \neq \emptyset$ and there exists j such that $\alpha_j = \varphi_j(\bar{v}_j)$ for some $\bar{v}_j \in \mathcal{V}(\mathcal{K})$, and $\alpha_i \geq \varphi_i(\bar{v}_j)$, for $1 \leq i \leq k$. For each $i \neq j$, let us take a vertex $\bar{v}_i \in \mathcal{V}(\mathcal{K}_\alpha)$ such that $\varphi_i(\bar{v}_i) \geq \varphi_i(v)$ for every $v \in \mathcal{V}(\alpha)$. Now we set $\lambda = (\lambda_1, \ldots, \lambda_k)$, with $\lambda_j = \varphi_j(\bar{v}_j)$. By construction λ belongs to Λ. Furthermore, it holds that $K_\alpha = K_\lambda$. Indeed, obviously, $K_\lambda \subseteq K_\alpha$. Moreover, for every $v \in K_\alpha$, by definition of $\bar{v}_i$ it holds that $\varphi_i(\bar{v}_i) \geq \varphi_i(v)$ for $1 \leq i \leq k$. Equivalently, $\lambda_i \geq \varphi_i(v)$ for $1 \leq i \leq k$, implying that $v \in K_\lambda$. □

The structure of set Λ is visualized in Figure 3. The previous proposition prompts for the following definition.

Definition 4.7 Consider the discrete map $\varphi : \mathcal{V}(\mathcal{K}) \to \mathbb{R}^k$ defined on vertices of $\mathcal{K}$. The *discrete rank invariant* of φ is the restriction of the real rank invariant ρ_φ to the finite domain $\Lambda^2_+ := \Delta^k_+ \cap (\Lambda \times \Lambda)$.

Definition 4.7 gives a discrete rank invariant which is similar to the one defined in [8, 7], except for the fact that we are using a different homological structure.

Defining a distance D directly on the basis of k-dimensional rank invariants would be a task impossible to accomplish. Even when $k = 2$, a pair of complexes $\mathcal{K}$ and $\mathcal{L}$ with an order of thousand vertices would result in computing ranks of millions of maps induced by inclusions. This motivates the one-dimensional reduction method described in the next section to compute the matching distance.

5 Algorithm and experimentation

For experimentation purposes, we now fix the distance between rank invariants that we will use to be the matching distance $\mathrm{D_m}$ defined in [5].

The one-dimensional reduction method presented in [5] to compute the matching distance consists of applying the one-dimensional rank invariant along the lines $t \mapsto \vec{b} + t\vec{l}$ parameterized by t and determined by pairs of vectors $(\vec{l}, \vec{b})$ in a chosen grid in $\mathbb{R}^k \times \mathbb{R}^k$, where $\vec{b}$ is an initial point and $\vec{l}$ directs the line. It is assumed that all components of $\vec{l}$ are positive, and that $\vec{l} \cdot \mathbf{1} = 1$, $\vec{b} \cdot \mathbf{1} = 0$, where $\mathbf{1} = (1, 1, \ldots, 1)$. For all $(\alpha, \beta) \in \Delta^k_+$, there exists a unique such pair, which

will be called *linearly admissible pair* or simply *admissible pair*, the set of which will be denoted $Ladm_k$. Also denote by $g_{(\vec{l},\vec{b})} : K \to \mathbb{R}$ and $h_{(\vec{l},\vec{b})} : L \to \mathbb{R}$ the one-dimensional functions given by $g_{(\vec{l},\vec{b})}(x) = \|(\varphi^{\urcorner}(x) - \vec{b})/\vec{l}\|$ and $h_{(\vec{l},\vec{b})}(x) = \|(\psi^{\urcorner}(x) - \vec{b})/\vec{l}\|$, where the division is coordinate-wise and the norm is the supremum norm. For ease of notation, the pair $(\vec{l},\vec{b})$ may be left out of g and h if it is unambiguous. By [5, Lemma 1], if $\alpha = \vec{b} + s\vec{l}$, then

$$K_{\varphi^{\urcorner} \preceq \alpha} = K_{g \leq s}.$$

This and Theorem 4.3 implies

Corollary 5.1 *Consider* $(\alpha,\beta) = (\vec{b} + s\vec{l}, \vec{b} + t\vec{l}) \in \Delta^k_+$, *for some* $(s,t) \in \Delta^1_+$. *Then*

$$\rho_{\varphi}(\alpha,\beta) = \rho_g(s,t).$$

The above theorem shows that it is legitimate to apply the reduction method of [5] to simplicial sublevel sets. Following [5, Definition 11], we define the *multi-dimensional matching distance* between the rank invariants ρ_φ and ρ_ψ to be

$$\mathrm{D_m}\left(\rho_\varphi, \rho_\psi\right) = \sup_{(\vec{l},\vec{b}) \in Ladm_k} \min_{i=1,\ldots,k} l_i \, \mathrm{D_m}\left(\rho_{g_{(\vec{l},\vec{b})}}, \rho_{h_{(\vec{l},\vec{b})}}\right).$$

In this section, the value $\min_{i=1,\ldots,k} l_i \, \mathrm{D_m}\left(\rho_{g_{(\vec{l},\vec{b})}}, \rho_{g_{(\vec{l},\vec{b})}}\right)$ will be denoted $\mathrm{d_m}(\rho_\varphi, \rho_\psi)$ or $\mathrm{d_{m(\vec{l},\vec{b})}}(\rho_\varphi, \rho_\psi)$ and referred to as the *rescaled one-dimensional matching distance*. The computational problem is, given a threshold value $\epsilon > 0$, computing an approximate matching distance $\widetilde{\mathrm{D_m}}\left(\rho_\varphi, \rho_\psi\right)$ on a suitable finite subset $A \subset Ladm_k$ such that

$$\widetilde{\mathrm{D_m}}\left(\rho_\varphi, \rho_\psi\right) \leq \mathrm{D_m}\left(\rho_\varphi, \rho_\psi\right) \leq \widetilde{\mathrm{D_m}}\left(\rho_\varphi, \rho_\psi\right) + \epsilon. \tag{11}$$

5.1 Algorithm

Our algorithm's inputs consist of lists of simplices of $\mathcal{K}$ and $\mathcal{L}$ of highest dimension together with their adjacency relations and vertices, and of the values of normalized measuring functions $\varphi : \mathcal{V}(\mathcal{K}) \to \mathbb{R}^2$ and $\psi : \mathcal{V}(\mathcal{L}) \to \mathbb{R}^2$, as well as a tolerance ϵ.[1] In our computations, we have confined ourselves to the case where $\mathcal{K}$ and $\mathcal{L}$ are triangular meshes. Its output is an approximate matching distance $\widetilde{\mathrm{D_m}}\left(\rho_\varphi, \rho_\psi\right)$. To compute the one-dimensional persistent homology on admissible pairs, we use the persistent homology software JPlex [16]. By default, JPlex computes the persistent Betti numbers over $\mathbb{Z}_{11}$ of a discretely indexed filtration of simplicial complexes. We build this filtration by adding simplices in the following recursive way. We first order the values attained by the one-dimensional measuring function g in increasing order, $\{g_1, \ldots, g_N\}$. A finite filtration $\{\mathcal{K}_1, \ldots, \mathcal{K}_N\}$

[1] Due to the finite precision of computer arithmetic, the codomain of the functions φ_i and ψ_i is in reality $10^{-p}\,\mathbb{Z}$ rather than $\mathbb{R}$. In our computations we tended to use $p = 6$, that is, a precision of up to six digits after the decimal point.

is then built by inserting simplices $\sigma \in \mathcal{K}$. If $\sigma = \{v\}$, where v is a vertex, we put $\{v\}$ into $\mathcal{K}_i$ if $g(v) \leq g_i$. Otherwise, $\sigma \in \mathcal{K}_i$ if all its vertices are in $\mathcal{K}_i$. Similarly, the function h is used to build a finite filtration $\{\mathcal{L}_1, \ldots, \mathcal{L}_M\}$ using simplices of $\mathcal{L}$.

The set of admissible pairs $Ladm_2$ is the set of quadruples $(a, 1-a, b, -b) \in \mathbb{R}^2 \times \mathbb{R}^2$ such that $0 < a < 1$. As described in [2, Remark 3.2], it is possible to avoid computation of the one-dimensional matching distance over a large portion of $Ladm_2$. Since the functions φ and ψ are normalized, $C = \max\{\|\varphi\|, \|\psi\|\} = 1$. Let $Ladm_2^*$ be the set of admissible pairs such that $|b| < 1$. Then, to compute the maximal value of $\mathrm{D_m}(\rho_g, \rho_h)$ over $Ladm_2 \backslash Ladm_2^*$, it is sufficient to consider the two admissible pairs $(a, 1-a, b, -b) = (1/2, 1/2, 2, -2)$ and $(1/2, 1/2, -2, 2)$. The details can be found in [2].

It follows from a generalization of the Error Bound Theorem ([2, Theorem 3.4] and [10]) to persistent homology of arbitrary order that if for $(\vec{l}, \vec{b})$ and $(\vec{l'}, \vec{b'}) \in Ladm_2$, $\|(\vec{l}, \vec{b}) - (\vec{l'}, \vec{b'})\| \leq \delta$, then for normalized functions φ and ψ

$$|\,\mathrm{d_m}_{(\vec{l},\vec{b})}(\rho_\varphi, \rho_\psi) - \mathrm{d_m}_{(\vec{l'},\vec{b'})}(\rho_\varphi, \rho_\psi)| \leq 18\delta.$$

This suggests that in order to satisfy Equation (11), it suffices to choose admissible pairs $(\vec{l}, \vec{b}) \in Ladm_2^*$ at a distance within $\epsilon/9$ of each other, guaranteeing that every member of $Ladm_2^*$ is within $\epsilon/18$ of a tested pair. In practice, our algorithm is reminiscent of the grid algorithm shown in Section 3 of [2], in the sense that we take pairs at a distance of $1/2^N$ of each other with N sufficiently large. We observe that the set $Ladm_2$ is in bijective correspondence with $\{(a, b) \in \mathbb{R}^2 \,|\, a \in (0, 1), b \in \mathbb{R}\}$, and so we will speak of computing $\mathrm{d_m}(\rho_\varphi, \rho_\psi)$ at a point $P = (a, b)$ of the preceding set. The lattice of points on which we compute this rescaled matching distance is chosen as follows: choose $N \in \mathbb{N}$ such that $1/2^N \leq \epsilon/18$, and choose $P_{ij} = (a_i, b_j), i = 0, \ldots, 2^N - 1, j = 0, \ldots, 2^{N+1} - 1$ such that $a_i = (2i+1)/2^{N+1}, b_j = 1 - (2j+1)/2^{N+1}$.

5.2 Examples of topological aliasing

Our experimentations have been made on triangular meshes of compact 2D surfaces. In doing so, the influence on experimental results of the concept of topological aliasing discussed in Section 3 became apparent. Namely, we used our algorithm to compare in a pairwise manner 10 cat models, a selection of which is found in Figure 4. We used for φ_i and ψ_i, $i = 1, 2$, the following functions. Assume that the model $\mathcal{K}$ is such that its vertex set $\mathcal{V}(\mathcal{K}) = \{v_1, \ldots, v_n\}$ and compute the following principal vector:

$$\vec{w} = \frac{\sum_{i=1}^n (v_i - c)\|v_i - c\|_2}{\sum_{i=1}^n \|v_i - c\|_2^2},$$

where c is the centre of mass of K defined by taking the weighted average of the centres of each triangle. Let d be the line passing through c having $\vec{w}$ as its

direction vector, and let π be the plane passing through c having $\vec{w}$ as its normal vector. We defined

$$\varphi_1(v_i) = 1 - \frac{\operatorname{dist}(v_i, d)}{\max_{j=1,\ldots,n} \operatorname{dist}(v_j, d)}$$

and

$$\varphi_2(v_i) = 1 - \frac{\operatorname{dist}(v_i, \pi)}{\max_{j=1,\ldots,n} \operatorname{dist}(v_j, \pi)},$$

where $\operatorname{dist}(v, d)$ and $\operatorname{dist}(v, \pi)$ are defined in the usual way, as the minimal Euclidean distance between v and the points on d or π. The functions ψ_1 and ψ_2 were defined similarly using the model $\mathcal{L}$. We then repeated the same procedure on the barycentric subdivisions of the models, with the value of the function at the new vertices defined using the linear interpolant. We found out that in this case the computed matching distance did not always yield the same result as when using the original unsubdivided models. However, replacing the linear interpolant by the axis-wise linear interpolant allowed us to retrieve the same results.

We show in Table 1 a selected subset of our results, and in Figure 4 images of the five models for which results are shown. The first two rows of numbers in each table represent the 1D matching distance computed using each of φ_i and ψ_i for $i = 1, 2$ respectively, while the last three rows represent the approximated 2D matching distance computed for three different tolerance levels ϵ. The column named "Nonsub" shows the distances computed on the original model, while "Linear" and "Axis-wise" show those computed on the subdivided models with respectively the linear and axis-wise linear interpolants. "Diff" and "% Diff" show the difference and relative difference between the matching distance results for the unsubdivided models and subdivided models with linear interpolation.

We can see that while the matching distance computed using the axis-wise linear interpolant is, for every tolerance level, equal to the matching distance between the original models, the matching distance computed using the linear interpolant can be quite different, and this both using 0th- and 1st-order persistent homology. However, this phenomenon is only seen when computing the 2D matching distance: the 1D matching distances (the numbers in the first two rows of each table) are always the same. Given that in our context topological aliasing can only be observed when using multi-measuring functions, this follows our expectations.

Acknowledgments

This work was partially supported by the following institutions: CRM-FQRNT (M.E.), University of Bologna under Marco Polo grant (N.C.), University of Modena and Reggio Emilia under Visiting Professor 2010 grant (T.K.), NSERC Canada Discovery Grant (T.K.), Fields Institute (P.F. and C.L.).

References

[1] F. Bernardini and H. E. Rushmeier. The 3D model acquisition pipeline. *Comput. Graph. Forum*, 21(2):149–172, 2002.

[2] S. Biasotti, A. Cerri, P. Frosini, and D. Giorgi. A new algorithm for computing the 2-dimensional matching distance between size functions. *Pattern Recognition Letters*, 32:1735–1746, 2011.

[3] 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.

[4] S. Biasotti, L. De Floriani, B. Falcidieno, P. Frosini, D. Giorgi, C. Landi, L. Papaleo, and M. Spagnuolo. Describing shapes by geometrical-topological properties of real functions. *ACM Comput. Surv.*, 40(4):1–87, 2008.

[5] F. Cagliari, B. Di Fabio, and M. Ferri. One-dimensional reduction of multidimensional persistent homology. *Proc. Amer. Math. Soc.*, 138(8):3003–3017, 2010.

[6] F. Cagliari and C. Landi. Finiteness of rank invariants of multidimensional persistent homology groups. *Applied Mathematics Letters*, 24(4):516 – 518, 2011.

[7] G. Carlsson, G. Singh, and A. Zomorodian. Computing multidimensional persistence. In *ISAAC '09: Proceedings of the 20th International Symposium on Algorithms and Computation*, pages 730–739, Berlin, Heidelberg, 2009. Springer-Verlag.

[8] G. Carlsson and A. Zomorodian. The theory of multidimensional persistence. In *SCG '07: Proceedings of the 23rd annual Symposium on Computational Geometry*, pages 184–193, New York, NY, USA, 2007. ACM.

[9] A. Cerri, B. Di Fabio, M. Ferri, P. Frosini, and C. Landi. Betti numbers in multidimensional persistent homology are stable functions. Tech. Rep., Univ. di Bologna, November 2010. http://amsacta.cib.unibo.it/2923/.

[10] A. Cerri and P. Frosini. A new approximation algorithm for the matching distance in multidimensional persistence. http://amsacta.cib.unibo.it/2971/.

[11] A. Cerri and P. Frosini. Advances in multidimensional size theory. *Image Analysis and Stereology*, 29:19–26, 2010.

[12] 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.

[13] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Stability of persistence diagrams. *Discrete Comput. Geom.*, 37(1):103–120, 2007.

[14] D. Cohen-Steiner, H. Edelsbrunner, J. Harer, and Y. Mileyko. Lipschitz functions have L_p-stable persistence. *Foundations of Computational Mathematics*, 10(2):127–139, 2010.

[15] M. d'Amico, P. Frosini, and C. Landi. Natural pseudo-distance and optimal matching between reduced size functions. *Acta. Appl. Math.*, 109:527–554, 2010.

[16] V. de Silva, G. Carlsson, and M. Vejdemo Johansson. Plex: A system for computational homology. http://comptop.stanford.edu/u/programs/jplex/.

[17] H. Edelsbrunner and J. Harer. Jacobi sets of multiple Morse functions. In *Foundations of Computational Mathematics*, pages 37–57, England, 2004. Cambridge Univ. Press.

[18] 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.

[19] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. *Discrete & Computational Geometry*, 28(4):511–533, 2002.

[20] P. Frosini and M. Mulazzani. Size homotopy groups for computation of natural size distances. *Bull. Belg. Math. Soc. Simon Stevin*, 6(3):455–464, 1999.

[21] T. Kaczynski, K. Mischaikow, and M. Mrozek. *Computational Homology*. Number 157 in Applied Mathematical Sciences. Springer-Verlag, 1st edition, 2004.

[22] M. Lesnick. The optimality of the interleaving distance on multidimensional persistence modules. Technical report, 2011. arXiv:1106.5305v2 [cs.CG].

[23] D. Morozov. *Homological Illusions of Persistence and Stability*. PhD thesis, Duke Univ., 2008.

[24] S. Smale. Global analysis and economics. I. Pareto optimum and a generalization of Morse theory. In *Dynamical systems (Proc. Sympos., Univ. Bahia, Salvador, 1971)*, pages 531–544. Academic Press, New York, 1973.

[25] A. Verri, C. Uras, P. Frosini, and M. Ferri. On the use of size functions for shape analysis. *Biological Cybernetics*, 70(2):99–107, 1993.

ARCES
Università di Bologna
via Toffano 2/2
40135 Bologna, Italia

{cavazza,frosini}@dm.unibo.it

Département de mathématiques
Université de Sherbrooke,
Sherbrooke (Québec), Canada J1K 2R1
{marc.ethier, t.kaczynski}@usherbrooke.ca

Dipartimento di Scienze e Metodi dell'Ingegneria
Università di Modena e Reggio Emilia
via Amendola 2, Pad. Morselli
42100 Reggio Emilia
clandi@unimore.it

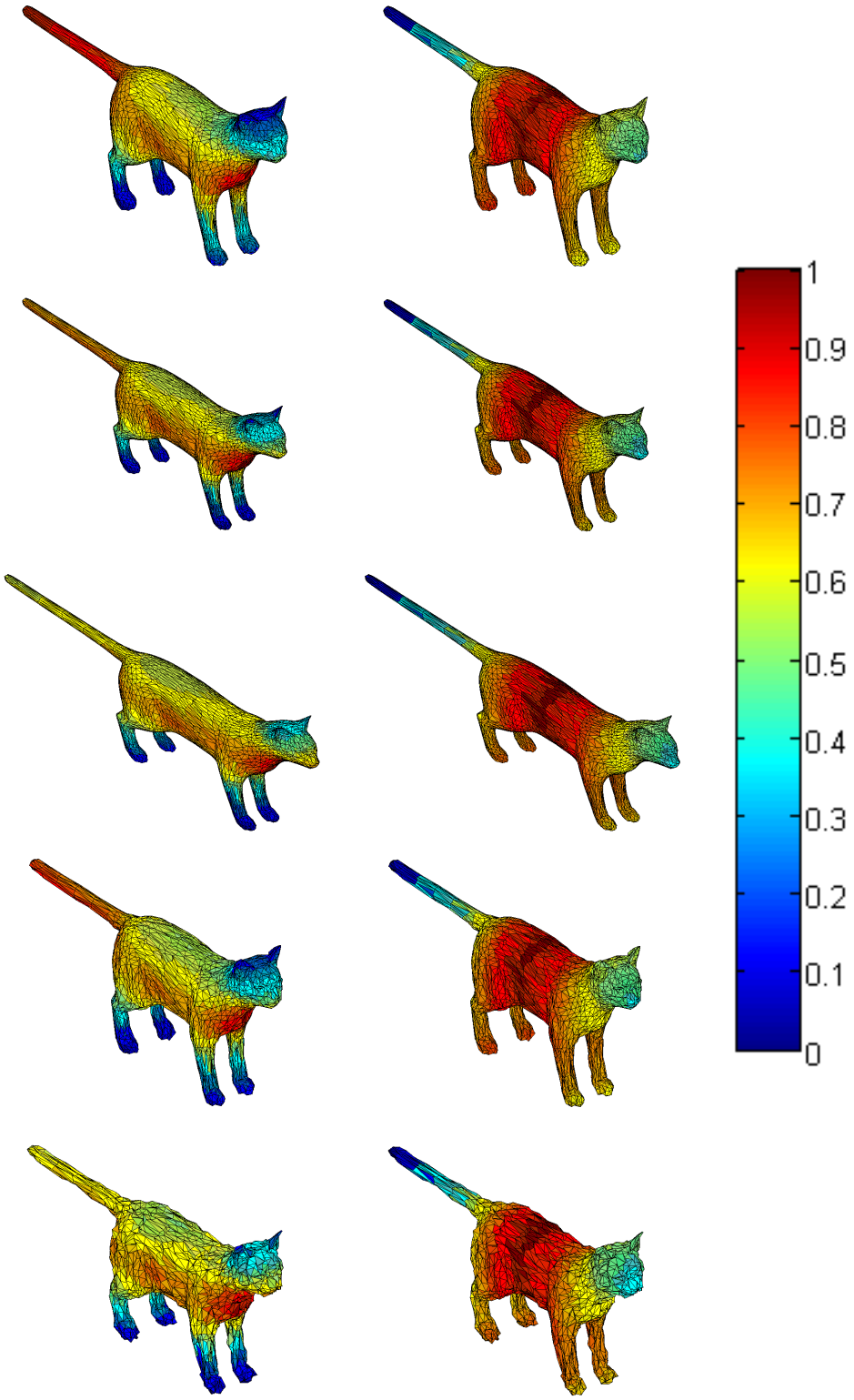

Figure 4: The models cat0, cat0-tran1-1, cat0-tran1-2, cat0-tran2-1 and cat0-tran2-2 are shown along with the values of φ_1 and φ_2. Models are courtesy of the authors of [2].

	cat0 vs. cat0-tran1-1				
	Nonsub	Linear	Axis-wise	Diff	% Diff
H_1	0.031129	0.031129	0.031129	*0.000000*	*0.000000*
	0.039497	0.039497	0.039497	*0.000000*	*0.000000*
	0.039497	0.039497	0.039497	*0.000000*	*0.000000*
	0.046150	0.039497	0.046150	*-0.006653*	*-16.844317*
	0.046150	**0.040576**	**0.046150**	***-0.005574***	***-13.737185***
H_0	0.118165	0.118165	0.118165	*0.000000*	*0.000000*
	0.032043	0.032043	0.032043	*0.000000*	*0.000000*
	0.194217	0.177001	0.194217	*-0.017216*	*-9.726499*
	0.224227	0.203102	0.224227	*-0.021125*	*-10.401178*
	0.225394	**0.207266**	**0.225394**	***-0.018128***	***-8.746249***

	cat0-tran1-2 vs. cat0-tran2-1				
	Nonsub	Linear	Axis-wise	Diff	% Diff
H_1	0.017272	0.017272	0.017272	*0.000000*	*0.000000*
	0.026101	0.026101	0.026101	*0.000000*	*0.000000*
	0.026101	0.028686	0.026101	*0.002585*	*9.903835*
	0.034314	0.028686	0.034314	*-0.005628*	*-19.619327*
	0.034314	**0.029188**	**0.034314**	***-0.005126***	***-17.562012***
H_0	0.182985	0.182985	0.182985	*0.000000*	*0.000000*
	0.018951	0.018951	0.018951	*0.000000*	*0.000000*
	0.192872	0.188365	0.192872	*-0.004507*	*-2.392695*
	0.207480	0.202844	0.207480	*-0.004636*	*-2.285500*
	0.208451	**0.204511**	**0.208451**	***-0.003940***	***-1.926547***

	cat0-tran2-1 vs. cat0-tran2-2				
	Nonsub	Linear	Axis-wise	Diff	% Diff
H_1	0.022001	0.022001	0.022001	*0.000000*	*0.000000*
	0.034288	0.034288	0.034288	*0.000000*	*0.000000*
	0.034288	0.034288	0.034288	*0.000000*	*0.000000*
	0.045545	0.035702	0.045545	*-0.009843*	*-27.569884*
	0.045545	**0.037061**	**0.045545**	***-0.008484***	***-22.891989***
H_0	0.095677	0.095677	0.095677	*0.000000*	*0.000000*
	0.032966	0.032966	0.032966	*0.000000*	*0.000000*
	0.178776	0.182322	0.178776	*0.003546*	*1.983488*
	0.202770	0.196977	0.202770	*-0.005793*	*-2.940952*
	0.212733	**0.208097**	**0.212733**	***-0.004636***	***-2.227807***

Table 1: The approximated matching distance computed by our algorithm for three decreasing tolerance values ($\epsilon = 9/8, 9/16$ and $9/32$) is shown for a few test cases (unsubdivided, subdivided with $\overline{\varphi}$ and subdivided with $\varphi^{\urcorner}$), with 0th- and 1st-order rank invariants.