Stability in multidimensional Size Theory

A. Cerri[1,2] P. Frosini[1,2] C. Landi[1,3,*]

[1] ARCES, Università di Bologna,
via Toffano 2/2, I-40135 Bologna, Italia

[2]Dipartimento di Matematica, Università di Bologna,
P.zza di Porta S. Donato 5, I-40126 Bologna, Italia

[3]DISMI, Università di Modena e Reggio Emilia,
via Amendola 2, Pad. Morselli, I-42100 Reggio Emilia, Italia

February 1, 2008

Abstract

This paper proves that in Size Theory the comparison of multidimensional size functions can be reduced to the 1-dimensional case by a suitable change of variables. Indeed, we show that a foliation in half-planes can be given, such that the restriction of a multidimensional size function to each of these half-planes turns out to be a classical size function in two scalar variables. This leads to the definition of a new distance between multidimensional size functions, and to the proof of their stability with respect to that distance.

Keywords: Multidimensional Size Function, Multidimensional Measuring Function, Natural Pseudo-distance.

Introduction

Shape comparison is probably one of the most challenging issues in Computer Vision and Pattern Recognition. In recent years many papers have been devoted to this subject and new

*Corresponding author. E-mail Address: `clandi@unimore.it`

mathematical techniques have been developed to deal with this problem. In the early 90's, Size Theory was proposed as a geometrical/topological approach to shape comparison. The main idea is to translate the comparison of two datasets (e.g. 3D-models, images or sounds) into the comparison of two suitable topological spaces $\mathcal{M}$ and $\mathcal{N}$, endowed with two continuous functions $\vec{\varphi} : \mathcal{M} \to \mathbb{R}^k$, $\vec{\psi} : \mathcal{N} \to \mathbb{R}^k$. These functions are called *k-dimensional measuring functions* and are chosen according to the application. In other words, they can be seen as descriptors of the properties considered relevant for the comparison. In [19] the definition of the natural pseudo-distance d between the pairs $(\mathcal{M}, \vec{\varphi})$, $(\mathcal{N}, \vec{\psi})$ was introduced, setting $d\left((\mathcal{M}, \vec{\varphi}), (\mathcal{N}, \vec{\psi})\right)$ equal to the infimum of the change of the measuring function, induced by composition with all the homeomorfisms from $\mathcal{M}$ to $\mathcal{N}$. Unfortunately, the study of d is quite difficult, even for $k = 1$, although strong properties can be proved in this case (cf. [9, 11]). Size Theory tackles this problem by introducing some mathematical tools that allow us to easily obtain lower bounds for d, such as *size homotopy groups* and *size functions* (cf. [19] and [10]). The idea is to study the pairs $(\mathcal{M}\langle\vec{\varphi} \preceq \vec{x}\rangle, \mathcal{M}\langle\vec{\varphi} \preceq \vec{y}\rangle)$, where $\mathcal{M}\langle\vec{\varphi} \preceq \vec{t}\,\rangle$ is defined by setting $\mathcal{M}\langle\vec{\varphi} \preceq \vec{t}\,\rangle = \{P \in \mathcal{M} : \varphi_i(P) \leq t_i, i = 1, \ldots, k\}$ for $\vec{t} = (t_1, \ldots, t_k) \in \mathbb{R}^k$. The k-th size homotopy group $\pi_k(\vec{x}, \vec{y})$ describes the non-trivial equivalence classes of k-dimensional loops in $\mathcal{M}\langle\vec{\varphi} \preceq \vec{x}\rangle$ that remain homotopically non-trivial also in $\mathcal{M}\langle\vec{\varphi} \preceq \vec{y}\rangle$. Size functions count the number of connected components in $\mathcal{M}\langle\vec{\varphi} \preceq \vec{y}\rangle$ that meet $\mathcal{M}\langle\vec{\varphi} \preceq \vec{x}\rangle$. It turns out that $\pi_0(\vec{x}, \vec{y})$ is a set whose cardinality is equal to the value taken by the size function at $(\vec{x}, \vec{y})$. From the homological point of view, an analogous approach, named *Size Functor*, has been developed in [2] for 1-dimensional measuring functions.

More recently, similar ideas have independently led Edelsbrunner et al. to the definition of *Persistent Homology* (cf. [12, 13]), and Allili et al. to the definition of the *Morse Homology Descriptor* (cf. [1]).

From the beginning of the 90's, size functions have been studied and applied in the case of 1-dimensional measuring functions (cf., e.g., [8, 14, 15, 16, 18, 20, 21, 22, 23, 24]). The multidimensional case presented more severe difficulties, since a concise, complete and stable description of multidimensional size functions was not available before this work.

In [3], Carlsson and Zomorodian examine the completeness problem by studying Multidimensional Persistent Homology. In that paper, it is claimed that multidimensional persistence has an essentially different character from its 1-dimensional version. Indeed, their approach does not seem to lead to a concise, complete and stable descriptor in the multidimensional case, whereas it does in classical Persistent Homology (see [4]).

The first result of this paper is the proof that in Size Theory the comparison of multidimensional size functions can be reduced to the 1-dimensional case by a suitable change of variables (Theorem 1). The key idea is to show that a foliation in half-planes can be given, such that the restriction of a multidimensional size function to these half-planes turns out to be a classical size function in two scalar variables. Our approach implies that each size function, with respect to a k-dimensional measuring function, can be completely and compactly described by a parameterized family of discrete descriptors (Remark 2). This follows from the results proved in [17] about the representation of classical size functions by means of formal series of points and lines, applied to each plane in our foliation. An important consequence is that we can easily prove the stability of this new descriptor (and hence of the corresponding k-dimensional size function) also in higher dimensions (Proposition 2), by using a recent result of stability proved for 1-dimensional size functions (cf. [6, 7]) and analogous to the result obtained in [4] for Persistence Homology. As a final contribution, we show that a matching distance between size functions, with respect to measuring functions taking values in $\mathbb{R}^k$, can easily be introduced (Definition 4). This matching distance provides a lower bound for the natural pseudo-distance, also in the multidimensional case (Theorem 2). All these results open the way to the use of Multidimensional Size Theory in real applications.

Outline. In Section 1 we give the definition of k-dimensional size function. In Section 2 the foliation we use is presented, and the reduction to the 1-dimensional case is proved. Section 3 shows the stability of our computational method, implying a lower bound for the natural pseudo-distance. Additionally, a new distance between multidimensional size functions is introduced. In Section 4 the effectiveness of the multidimensional approach is tested on an example. Section 5 examines some links between multidimensional size functions and the concept of *vineyard*, recently introduced in [5]. Section 6 concludes the paper, presenting some discussion and future work.

1 Definition of k-dimensional size function

For the present paper, $\mathcal{M}$, $\mathcal{N}$ denote two non-empty compact and locally connected Hausdorff spaces.

In Multidimensional Size Theory [19], any pair $(\mathcal{M},\vec{\varphi})$, where $\vec{\varphi}=(\varphi_1,\ldots,\varphi_k):\mathcal{M}\to\mathbb{R}^k$ is a continuous function, is called a *size pair*. The function $\vec{\varphi}$ is called a *k-dimensional measuring function*. The following relations $\preceq$ and $\prec$ are defined in $\mathbb{R}^k$: for $\vec{x} = (x_1, \ldots, x_k)$ and $\vec{y} = (y_1, \ldots, y_k)$, we say $\vec{x} \preceq \vec{y}$ (resp. $\vec{x} \prec \vec{y}$) if and only if $x_i \leq y_i$ (resp. $x_i < y_i$) for every index $i = 1, \ldots, k$. Moreover, $\mathbb{R}^k$ is endowed with the usual max-norm: $\|(x_1, x_2, \ldots, x_k)\|_\infty = \max_{1 \leq i \leq k} |x_i|$. In this framework, if $\mathcal{M}$ and $\mathcal{N}$ are homeomorphic, the size pairs $(\mathcal{M}, \vec{\varphi})$ and $(\mathcal{N}, \vec{\psi})$ can be compared by means of the *natural pseudo-distance* d, defined as

$$d((\mathcal{M}, \vec{\varphi}), (\mathcal{N}, \vec{\psi})) = \inf_f \max_{P \in \mathcal{M}} \|\vec{\varphi}(P) - \vec{\psi}(f(P))\|_\infty \ ,$$

where f varies among all the homeomorphisms between $\mathcal{M}$ and $\mathcal{N}$. The term pseudo-distance means that d can vanish even if the size pairs do not coincide. Here, and in what follows, $\mathbb{R}^k \times \mathbb{R}^k$ and $\mathbb{R}^{2k}$ are identified.

Now we introduce the k-dimensional analogue of size function for a size pair $(\mathcal{M}, \vec{\varphi})$. We shall use the following notations: Δ^+ will be the open set $\{(\vec{x}, \vec{y}) \in \mathbb{R}^k \times \mathbb{R}^k : \vec{x} \prec \vec{y}\}$, while $\Delta = \partial \Delta^+$. For every k-tuple $\vec{x} = (x_1, \ldots, x_k) \in \mathbb{R}^k$, let $\mathcal{M}\langle \vec{\varphi} \preceq \vec{x} \rangle$ be the set $\{P \in \mathcal{M} : \varphi_i(P) \leq x_i, \ i = 1, \ldots, k\}$.

Definition 1. *For every k-tuple $\vec{y} = (y_1, \ldots, y_k) \in \mathbb{R}^k$, we shall say that two points $P, Q \in \mathcal{M}$ are* $\langle \vec{\varphi} \preceq \vec{y} \rangle$-connected *if and only if a connected subset of $\mathcal{M}\langle \vec{\varphi} \preceq \vec{y} \rangle$ exists, containing P and Q.*

Definition 2. *We shall call* (k-dimensional) size function *associated with the size pair $(\mathcal{M}, \vec{\varphi})$ the function $\ell_{(\mathcal{M}, \vec{\varphi})} : \Delta^+ \to \mathbb{N}$, defined by setting $\ell_{(\mathcal{M}, \vec{\varphi})}(\vec{x}, \vec{y})$ equal to the number of equivalence classes in which the set $\mathcal{M}\langle \vec{\varphi} \preceq \vec{x} \rangle$ is divided by the $\langle \vec{\varphi} \preceq \vec{y} \rangle$-connectedness relation.*

Remark 1. *In other words, $\ell_{(\mathcal{M}, \vec{\varphi})}(\vec{x}, \vec{y})$ counts the connected components in $\mathcal{M}\langle \vec{\varphi} \preceq \vec{y} \rangle$ containing at least one point of $\mathcal{M}\langle \vec{\varphi} \preceq \vec{x} \rangle$.*

2 Reduction to the 1-dimensional case

In this section, we will show that there exists a parameterized family of half-planes in $\mathbb{R}^k \times \mathbb{R}^k$ such that the restriction of $\ell_{(\mathcal{M}, \vec{\varphi})}$ to each of these planes can be seen as a particular 1-dimensional size function.

Definition 3. *For every unit vector $\vec{l} = (l_1, \ldots, l_k)$ of $\mathbb{R}^k$ such that $l_i > 0$ for every $i = 1, \ldots, k$, and for every vector $\vec{b} = (b_1, \ldots, b_k)$ of $\mathbb{R}^k$ such that $\sum_{i=1}^k b_i = 0$, we shall say that the pair*

$(\vec{l},\vec{b})$ *is* admissible. *We shall denote the set of all admissible pairs in* $\mathbb{R}^k \times \mathbb{R}^k$ *by* Adm_k. *Given an admissible pair* $(\vec{l},\vec{b})$, *we define the half-plane* $\pi_{(\vec{l},\vec{b})}$ *of* $\mathbb{R}^k \times \mathbb{R}^k$ *by the following parametric equations:*

$$\begin{cases} \vec{x} = s\vec{l} + \vec{b} \\ \vec{y} = t\vec{l} + \vec{b} \end{cases}$$

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

Proposition 1. *For every* $(\vec{x},\vec{y}) \in \Delta^+$ *there exists one and only one admissible pair* $(\vec{l},\vec{b})$ *such that* $(\vec{x},\vec{y}) \in \pi_{(\vec{l},\vec{b})}$.

Proof. The claim immediately follows by taking, for $i = 1, \ldots, k$,

$$l_i = \frac{y_i - x_i}{\sqrt{\sum_{j=1}^k (y_j - x_j)^2}} \ , \quad b_i = \frac{x_i \sum_{j=1}^k y_j - y_i \sum_{j=1}^k x_j}{\sum_{j=1}^k (y_j - x_j)}.$$

Then, $\vec{x} = s\vec{l} + \vec{b}$, $\vec{y} = t\vec{l} + \vec{b}$, with

$$s = \frac{\sum_{j=1}^k x_j}{\sum_{j=1}^k l_j} = \frac{\sum_{j=1}^k x_j \sqrt{\sum_{j=1}^k (y_j - x_j)^2}}{\sum_{j=1}^k (y_j - x_j)}$$
$$t = \frac{\sum_{j=1}^k y_j}{\sum_{j=1}^k l_j} = \frac{\sum_{j=1}^k y_j \sqrt{\sum_{j=1}^k (y_j - x_j)^2}}{\sum_{j=1}^k (y_j - x_j)}.$$

□

Now we can prove the reduction to the 1-dimensional case.

Theorem 1. *Let* $(\vec{l},\vec{b})$ *be an admissible pair, and* $F^{\vec{\varphi}}_{(\vec{l},\vec{b})} : \mathcal{M} \to \mathbb{R}$ *be defined by setting*

$$F^{\vec{\varphi}}_{(\vec{l},\vec{b})}(P) = \max_{i=1,\ldots,k} \left\{ \frac{\varphi_i(P) - b_i}{l_i} \right\} \ .$$

Then, for every $(\vec{x},\vec{y}) = (s\vec{l} + \vec{b}, t\vec{l} + \vec{b}) \in \pi_{(\vec{l},\vec{b})}$ *the following equality holds:*

$$\ell_{(\mathcal{M},\vec{\varphi})}(\vec{x},\vec{y}) = \ell_{(\mathcal{M},F^{\vec{\varphi}}_{(\vec{l},\vec{b})})}(s,t) \ .$$

Proof. For every $\vec{x} = (x_1, \ldots, x_k) \in \mathbb{R}^k$, with $x_i = sl_i + b_i, \ i = 1, \ldots, k$, it holds that $\mathcal{M}\langle \vec{\varphi} \preceq \vec{x} \rangle = \mathcal{M}\langle F^{\vec{\varphi}}_{(\vec{l},\vec{b})} \le s \rangle$. This is true because

$$\begin{aligned} \mathcal{M}\langle \vec{\varphi} \preceq \vec{x} \rangle &= \{P \in \mathcal{M} : \varphi_i(P) \le x_i, \ i = 1, \ldots, k\} = \\ &= \{P \in \mathcal{M} : \varphi_i(P) \le sl_i + b_i, \ i = 1, \ldots, k\} = \\ &= \left\{P \in \mathcal{M} : \frac{\varphi_i(P) - b_i}{l_i} \le s, \ i = 1, \ldots, k\right\} = \mathcal{M}\langle F^{\vec{\varphi}}_{(\vec{l},\vec{b})} \le s \rangle. \end{aligned}$$

Analogously, for every $\vec{y} = (y_1, \ldots, y_k) \in \mathbb{R}^k$, with $y_i = tl_i + b_i$, $i = 1, \ldots, k$, it holds that $\mathcal{M}\langle \vec{\varphi} \preceq \vec{y} \rangle = \mathcal{M}\langle F^{\vec{\varphi}}_{(\vec{l},\vec{b})} \leq t \rangle$. Therefore Remark 1 implies the claim. $\square$

In the following, we shall use the symbol $F^{\vec{\varphi}}_{(\vec{l},\vec{b})}$ in the sense of Theorem 1.

Remark 2. *Theorem 1 allows us to represent each multidimensional size function as a parameterized family of formal series of points and lines, on the basis of the description introduced in [17] for the* 1*-dimensional case. Indeed, we can associate a formal series* $\sigma_{(\vec{l},\vec{b})}$ *with each admissible pair* $(\vec{l}, \vec{b})$*, with* $\sigma_{(\vec{l},\vec{b})}$ *describing the* 1*-dimensional size function* $\ell_{(\mathcal{M}, F^{\vec{\varphi}}_{(\vec{l},\vec{b})})}$*. The family* $\left\{ \sigma_{(\vec{l},\vec{b})} : (\vec{l}, \vec{b}) \in Adm_k \right\}$ *is a new complete descriptor for* $\ell_{(\mathcal{M},\vec{\varphi})}$*, in the sense that two multidimensional size functions coincide if and only if the corresponding parameterized families of formal series coincide.*

3 Lower bounds for the k-dimensional natural pseudo-distance

In [6, 7], it has been shown that 1-dimensional size functions can be compared by means of a distance, called *matching distance*. This distance is based on the observation that each 1-dimensional size function is the sum of characteristic functions of triangles. The matching distance is computed by finding an optimal matching between the sets of triangles describing two size functions. For a formal definition we refer to [6, 7] (see also [4] for the analogue of the matching distance in Persistent Homology). In the sequel, we shall denote by $d_{match}(\ell_{(\mathcal{M}, F^{\vec{\varphi}}_{(\vec{l},\vec{b})})}, \ell_{(\mathcal{N}, F^{\vec{\psi}}_{(\vec{l},\vec{b})})})$ the matching distance between the 1-dimensional size functions $\ell_{(\mathcal{M}, F^{\vec{\varphi}}_{(\vec{l},\vec{b})})}$ and $\ell_{(\mathcal{N}, F^{\vec{\psi}}_{(\vec{l},\vec{b})})}$.

The following result is an immediate consequence of Theorem 1 and Remark 2.

Corollary 1. *Let us consider the size pairs* $(\mathcal{M}, \vec{\varphi})$*,* $(\mathcal{N}, \vec{\psi})$*. Then, the identity* $\ell_{(\mathcal{M},\vec{\varphi})} \equiv \ell_{(\mathcal{N},\vec{\psi})}$ *holds if and only if* $d_{match}(\ell_{(\mathcal{M}, F^{\vec{\varphi}}_{(\vec{l},\vec{b})})}, \ell_{(\mathcal{N}, F^{\vec{\psi}}_{(\vec{l},\vec{b})})}) = 0$*, for every admissible pair* $(\vec{l}, \vec{b})$*.*

The next result proves that small enough changes in $\vec{\varphi}$ with respect to the max-norm induce small changes of $\ell_{(\mathcal{M}, F^{\vec{\varphi}}_{(\vec{l},\vec{b})})}$ with respect to the matching distance.

Proposition 2. *If* $(\mathcal{M}, \vec{\varphi})$*,* $(\mathcal{M}, \vec{\chi})$ *are size pairs and* $\max_{P \in \mathcal{M}} \|\vec{\varphi}(P) - \vec{\chi}(P)\|_\infty \leq \epsilon$*, then for each admissible pair* $(\vec{l}, \vec{b})$*, it holds that*

$$d_{match}(\ell_{(\mathcal{M}, F^{\vec{\varphi}}_{(\vec{l},\vec{b})})}, \ell_{(\mathcal{M}, F^{\vec{\chi}}_{(\vec{l},\vec{b})})}) \leq \frac{\epsilon}{\min_{i=1,\ldots,k} l_i}.$$

Proof. From the Matching Stability Theorem 25 in [6] (see also [7]) we obtain that

$$d_{match}(\ell_{(\mathcal{M},F^{\vec{\varphi}}_{(\vec{l},\vec{b})})},\ell_{(\mathcal{M},F^{\vec{\chi}}_{(\vec{l},\vec{b})})}) \quad\leq\quad \max_{P\in\mathcal{M}}|F^{\vec{\varphi}}_{(\vec{l},\vec{b})}(P)-F^{\vec{\chi}}_{(\vec{l},\vec{b})}(P)|.$$

Let us now fix $P\in\mathcal{M}$. Then, denoting by $\hat{\imath}$ the index for which $\max_i \frac{\varphi_i(P)-b_i}{l_i}$ is attained, by the definition of $F^{\vec{\varphi}}_{(\vec{l},\vec{b})}$ and $F^{\vec{\chi}}_{(\vec{l},\vec{b})}$ we have that

$$\begin{aligned} & F^{\vec{\varphi}}_{(\vec{l},\vec{b})}(P)-F^{\vec{\chi}}_{(\vec{l},\vec{b})}(P)=\max_i\frac{\varphi_i(P)-b_i}{l_i}-\max_i\frac{\chi_i(P)-b_i}{l_i}= \\ = & \frac{\varphi_{\hat{\imath}}(P)-b_{\hat{\imath}}}{l_{\hat{\imath}}}-\max_i\frac{\chi_i(P)-b_i}{l_i}\leq\frac{\varphi_{\hat{\imath}}(P)-b_{\hat{\imath}}}{l_{\hat{\imath}}}-\frac{\chi_{\hat{\imath}}(P)-b_{\hat{\imath}}}{l_{\hat{\imath}}}= \\ = & \frac{\varphi_{\hat{\imath}}(P)-\chi_{\hat{\imath}}(P)}{l_{\hat{\imath}}}\leq\frac{\|\vec{\varphi}(P)-\vec{\chi}(P)\|_\infty}{\min_{i=1,\ldots,k}l_i}. \end{aligned}$$

In the same way, we obtain $F^{\vec{\chi}}_{(\vec{l},\vec{b})}(P)-F^{\vec{\varphi}}_{(\vec{l},\vec{b})}(P)\leq\frac{\|\vec{\varphi}(P)-\vec{\chi}(P)\|_\infty}{\min_{i=1,\ldots,k}l_i}$. Therefore, if $\max_{P\in\mathcal{M}}\|\vec{\varphi}(P)-\vec{\chi}(P)\|_\infty\leq\epsilon$,

$$\max_{P\in\mathcal{M}}\left|F^{\vec{\varphi}}_{(\vec{l},\vec{b})}(P)-F^{\vec{\chi}}_{(\vec{l},\vec{b})}(P)\right|\leq\max_{P\in\mathcal{M}}\frac{\|\vec{\varphi}(P)-\vec{\chi}(P)\|_\infty}{\min_{i=1,\ldots,k}l_i}\leq\frac{\epsilon}{\min_{i=1,\ldots,k}l_i}.$$

□

Remark 3. *Analogously, it is easy to show that small enough changes in $(\vec{l},\vec{b})$ with respect to the* max-*norm induce small changes of $\ell_{(\mathcal{M},F^{\vec{\varphi}}_{(\vec{l},\vec{b})})}$ with respect to the matching distance.*

Proposition 2 and Remark 3 prove the stability of our computational approach.

Now we are able to prove our next result, showing that a lower bound exists for the multidimensional natural pseudo-distance.

Theorem 2. *Let $(\mathcal{M},\vec{\varphi})$ and $(\mathcal{N},\vec{\psi})$ be two size pairs, with $\mathcal{M}$, $\mathcal{N}$ homeomorphic. Setting $d((\mathcal{M},\vec{\varphi}),(\mathcal{N},\vec{\psi}))=\inf_f\max_{P\in\mathcal{M}}\|\vec{\varphi}(P)-\vec{\psi}(f(P))\|_\infty$, where f varies among all the homeomorphisms between $\mathcal{M}$ and $\mathcal{N}$, it holds that*

$$\sup_{(\vec{l},\vec{b})\in Adm_k}\min_{i=1,\ldots,k}l_i\cdot d_{match}(\ell_{(\mathcal{M},F^{\vec{\varphi}}_{(\vec{l},\vec{b})})},\ell_{(\mathcal{N},F^{\vec{\psi}}_{(\vec{l},\vec{b})})})\leq d((\mathcal{M},\vec{\varphi}),(\mathcal{N},\vec{\psi})).$$

Proof. For any homeomorphism f between $\mathcal{M}$ and $\mathcal{N}$, it holds that $\ell_{(\mathcal{N},F^{\vec{\psi}}_{(\vec{l},\vec{b})})}\equiv\ell_{(\mathcal{M},F^{\vec{\psi}}_{(\vec{l},\vec{b})}\circ f)}$. Moreover, by applying Proposition 2 with $\epsilon=\max_{P\in\mathcal{M}}\|\vec{\varphi}(P)-\vec{\psi}(f(P))\|_\infty$ and $\vec{\chi}=\vec{\psi}\circ f$, and observing that $F^{\vec{\psi}}_{(\vec{l},\vec{b})}\circ f\equiv F^{\vec{\psi}\circ f}_{(\vec{l},\vec{b})}$, we have

$$\min_{i=1,\ldots,k}l_i\cdot d_{match}(\ell_{(\mathcal{M},F^{\vec{\varphi}}_{(\vec{l},\vec{b})})},\ell_{(\mathcal{N},F^{\vec{\psi}}_{(\vec{l},\vec{b})})})\leq\max_{P\in\mathcal{M}}\|\vec{\varphi}(P)-\vec{\psi}(f(P))\|_\infty$$

for every admissible $(\vec{l},\vec{b})$. Since this is true for each homeomorphism f between $\mathcal{M}$ and $\mathcal{N}$, the claim immediately follows. □

Remark 4. *We observe that the left side of the inequality in Theorem 2 defines a distance between multidimensional size functions associated with homeomorphic spaces. When the spaces are not assumed to be homeomorphic, it still verifies all the properties of a distance, except for the fact that it may take the value $+\infty$. In other words, it defines an extended distance.*

Definition 4. *Let $(\mathcal{M}, \vec{\varphi})$ and $(\mathcal{N}, \vec{\psi})$ be two size pairs. We shall call* multidimensional matching distance *the extended distance defined by setting*

$$D_{match}(\ell_{(\mathcal{M},\vec{\varphi})}, \ell_{(\mathcal{N},\vec{\psi})}) = \sup_{(\vec{l},\vec{b})\in Adm_k} \min_{i=1,\dots,k} l_i \cdot d_{match}(\ell_{(\mathcal{M},F^{\vec{\varphi}}_{(\vec{l},\vec{b})})}, \ell_{(\mathcal{N},F^{\vec{\psi}}_{(\vec{l},\vec{b})})}).$$

Remark 5. *If we choose a non-empty subset $A \subseteq Adm_k$ and we substitute* $\sup_{(\vec{l},\vec{b})\in Adm_k}$ *with* $\sup_{(\vec{l},\vec{b})\in A}$ *in Definition 4, we obtain an (extended) pseudo-distance between multidimensional size functions. If A is finite, this pseudo-distance appears to be particularly suitable for applications, from a computational point of view.*

4 An example

In $\mathbb{R}^3$ consider the set $\mathcal{Q} = [-1,1]\times[-1,1]\times[-1,1]$ and the sphere $\mathcal{S}$ of equation $x^2+y^2+z^2 = 1$. Let also $\vec{\Phi} = (\Phi_1, \Phi_2) : \mathbb{R}^3 \to \mathbb{R}^2$ be the continuous function, defined as $\vec{\Phi}(x,y,z) = (|x|, |z|)$. In this setting, consider the size pairs $(\mathcal{M}, \vec{\varphi})$ and $(\mathcal{N}, \vec{\psi})$ where $\mathcal{M} = \partial\mathcal{Q}$, $\mathcal{N} = \mathcal{S}$, and $\vec{\varphi}$ and $\vec{\psi}$ are respectively the restrictions of $\vec{\Phi}$ to $\mathcal{M}$ and $\mathcal{N}$. In order to compare the size functions $\ell_{(\mathcal{M},\vec{\varphi})}$ and $\ell_{(\mathcal{N},\vec{\psi})}$, we are interested in studying the foliation in half-planes $\pi_{(\vec{l},\vec{b})}$, where $\vec{l} = (\cos\theta, \sin\theta)$ with $\theta \in (0, \frac{\pi}{2})$, and $\vec{b} = (a, -a)$ with $a \in \mathbb{R}$. Any such half-plane is represented by

$$\begin{cases} x_1 = s\cos\theta + a \\ x_2 = s\sin\theta - a \\ y_1 = t\cos\theta + a \\ y_2 = t\sin\theta - a \end{cases},$$

with $s,t \in \mathbb{R}$, $s < t$. Figure 1 shows the size functions $\ell_{(\mathcal{M},F^{\vec{\varphi}}_{(\vec{l},\vec{b})})}$ and $\ell_{(\mathcal{N},F^{\vec{\psi}}_{(\vec{l},\vec{b})})}$, for $\theta = \frac{\pi}{4}$ and $a = 0$, i.e. $\vec{l} = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ and $\vec{b} = (0,0)$. With this choice, we obtain that $F^{\vec{\varphi}}_{(\vec{l},\vec{b})} = \sqrt{2}\max\{\varphi_1, \varphi_2\} = \sqrt{2}\max\{|x|, |z|\}$ and $F^{\vec{\psi}}_{(\vec{l},\vec{b})} = \sqrt{2}\max\{\psi_1, \psi_2\} = \sqrt{2}\max\{|x|, |z|\}$. Therefore, Theorem 1 implies that, for every $(x_1, x_2, y_1, y_2) \in \pi_{(\vec{l},\vec{b})}$

$$\ell_{(\mathcal{M},\vec{\varphi})}(x_1, x_2, y_1, y_2) = \ell_{(\mathcal{M},\vec{\varphi})}\left(\frac{s}{\sqrt{2}}, \frac{s}{\sqrt{2}}, \frac{t}{\sqrt{2}}, \frac{t}{\sqrt{2}}\right) = \ell_{(\mathcal{M},F^{\vec{\varphi}}_{(\vec{l},\vec{b})})}(s,t)$$
$$\ell_{(\mathcal{N},\vec{\psi})}(x_1, x_2, y_1, y_2) = \ell_{(\mathcal{N},\vec{\psi})}\left(\frac{s}{\sqrt{2}}, \frac{s}{\sqrt{2}}, \frac{t}{\sqrt{2}}, \frac{t}{\sqrt{2}}\right) = \ell_{(\mathcal{N},F^{\vec{\psi}}_{(\vec{l},\vec{b})})}(s,t)\ .$$

In this case, by Theorem 2 and Remark 5 (applied for A containing just the admissible pair that we have chosen), a lower bound for the natural pseudo-distance $d((\mathcal{M},\vec{\varphi}),(\mathcal{N},\vec{\psi}))$ is given by

$$\frac{\sqrt{2}}{2} d_{match}(\ell_{(\mathcal{M},F^{\vec{\varphi}}_{(\vec{l},\vec{b})})}, \ell_{(\mathcal{N},F^{\vec{\psi}}_{(\vec{l},\vec{b})})}) = \frac{\sqrt{2}}{2}(\sqrt{2}-1) = 1 - \frac{\sqrt{2}}{2}.$$

Indeed, the matching distance $d_{match}(\ell_{(\mathcal{M},F^{\vec{\varphi}}_{(\vec{l},\vec{b})})}, \ell_{(\mathcal{N},F^{\vec{\psi}}_{(\vec{l},\vec{b})})})$ is equal to the cost of moving the point of coordinates $(0,\sqrt{2})$ onto the point of coordinates $(0,1)$, computed with respect to the max-norm. The points $(0,\sqrt{2})$ and $(0,1)$ are representative of the characteristic triangles of the size functions $\ell_{(\mathcal{M},F^{\vec{\varphi}}_{(\vec{l},\vec{b})})}$ and $\ell_{(\mathcal{N},F^{\vec{\psi}}_{(\vec{l},\vec{b})})}$, respectively.

We conclude by observing that $\ell_{(\mathcal{M},\varphi_1)} \equiv \ell_{(\mathcal{N},\psi_1)}$ and $\ell_{(\mathcal{M},\varphi_2)} \equiv \ell_{(\mathcal{N},\psi_2)}$. In other words, the multidimensional size functions, with respect to $\vec{\varphi}, \vec{\psi}$, are able to discriminate the cube and the sphere, while both the 1-dimensional size functions, with respect to φ_1, φ_2 and ψ_1, ψ_2, cannot do that. The higher discriminatory power of multidimensional size functions motivates their definition and use.

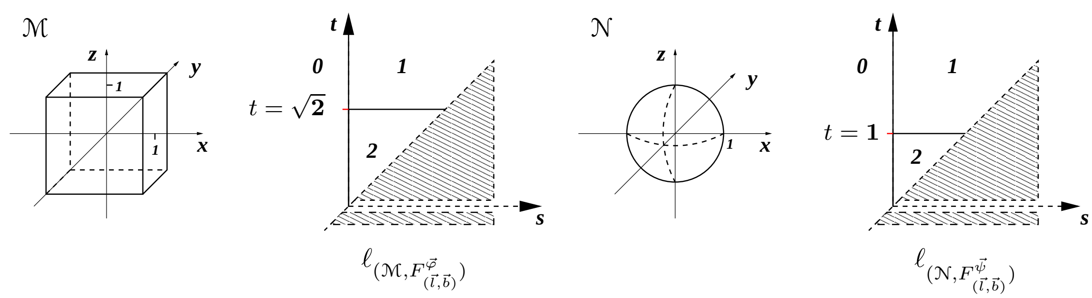

Figure 1: The topological spaces $\mathcal{M}$ and $\mathcal{N}$ and the size functions $\ell_{(\mathcal{M},F^{\vec{\varphi}}_{(\vec{l},\vec{b})})}, \ell_{(\mathcal{N},F^{\vec{\psi}}_{(\vec{l},\vec{b})})}$ associated with the half-plane $\pi_{(\vec{l},\vec{b})}$, for $\vec{l} = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$ and $\vec{b} = (0,0)$.

5 Links between dimension 0 vineyards and multidimensional size functions

In a recent paper [5], Cohen-Steiner et al. have introduced the concept of *vineyard*, that is a 1-parameter family of persistence diagrams associated with the homotopy f_t, interpolating between f_0 and f_1. These authors assume that the topological space is homeomorphic to the body of a simplicial complex, and that the measuring functions are *tame*. We shall do the same in this section. We recall that dimension p persistence diagrams are a concise representation of the function $rank\, H_p^{x,y}$, where $H_p^{x,y}$ denotes the dimension p persistent homology group computed at point (x, y) (cf. [5]). Therefore, the information described by vineyards is equivalent to the knowledge of the function $rank\, H_p^{x,y}$, computed with respect to the function f_t. We are interested in the case $p = 0$. Since, by definition, for $x < y$, $rank\, H_0^{x,y}$ coincides with the value taken by the size function $\ell_{(\mathcal{M},f_t)}(x, y)$, it follows that, for $x < y$, dimension 0 vineyards contain the same information as the 1-parameter family of size functions $\{\ell_{(\mathcal{M},f_t)}\}_{t\in[0,1]}$. Anyway, another interesting link exists between dimension 0 vineyards and multidimensional size functions. This link is expressed by the following theorem. In order to prove it, we need the next two lemmas. The former states that the relation of $\langle\vec{\varphi} \preceq \vec{y}\,\rangle$-connectedness passes to the limit.

Lemma 1. *Assume that $(\mathcal{M}, \vec{\varphi})$ is a size pair and $\vec{y} = (y_1, \ldots, y_k) \in \mathbb{R}^k$. If, for every $\varepsilon > 0$, P and Q are $\langle\vec{\varphi} \preceq (y_1 + \varepsilon, \ldots, y_k + \varepsilon)\,\rangle$-connected in $\mathcal{M}$, then they are also $\langle\vec{\varphi} \preceq \vec{y}\,\rangle$-connected.*

Proof. For every positive integer number n, let K_n be the connected component of $\mathcal{M}\langle\vec{\varphi} \preceq (y_1 + \frac{1}{n}, \ldots, y_k + \frac{1}{n})\rangle$ containing P and Q. Since connected components are closed sets and $\mathcal{M}$ is compact, each K_n is compact. The set $\bigcap_n K_n$ is the intersection of a family of connected compact Hausdorff subspaces with the property that $K_{n+1} \subseteq K_n$ for every n, and hence it is connected (cf. Theorem 28.2 in [25] p. 203). Moreover, $\bigcap_n K_n$ is a subset of $\mathcal{M}\langle\vec{\varphi} \preceq \vec{y}\rangle$ and contains both P and Q. Therefore, P and Q are $\langle\vec{\varphi} \preceq \vec{y}\,\rangle$-connected. $\square$

The following lemma allows us to study the behavior of multidimensional size functions near Δ (where they have not been defined because of instability problems when the measuring functions are not assumed to be tame).

Lemma 2. *Let $(\mathcal{M}, \vec{\varphi})$ be a size pair. If $\vec{x} \preceq \vec{y}$ then $\lim_{\varepsilon\to 0^+} \ell_{(\mathcal{M},\vec{\varphi})}\,(\vec{x}, (y_1 + \varepsilon, \ldots, y_k + \varepsilon))$ is equal to the number $L(\vec{x}, \vec{y})$ of equivalence classes of $\mathcal{M}\langle\vec{\varphi} \preceq \vec{x}\rangle$ quotiented with respect to the $\langle\vec{\varphi} \preceq \vec{y}\,\rangle$-connectedness relation.*

Note that, for $\vec{x} \prec \vec{y}$, $L(\vec{x}, \vec{y})$ simply coincides with $\ell_{(\mathcal{M},\vec{\varphi})}(\vec{x}, \vec{y})$.

Proof of Lemma 2. First of all we observe that the function $\ell_{(\mathcal{M},\vec{\varphi})}(\vec{x}, (y_1+\varepsilon, \ldots, y_k+\varepsilon))$ is nonincreasing in the variable ε, and hence the value $\lim_{\varepsilon\to 0^+} \ell_{(\mathcal{M},\vec{\varphi})}(\vec{x}, (y_1+\varepsilon, \ldots, y_k+\varepsilon))$ is defined. The statement of the theorem is trivial if $\lim_{\varepsilon\to 0^+} \ell_{(\mathcal{M},\vec{\varphi})}(\vec{x}, (y_1+\varepsilon, \ldots, y_k+\varepsilon)) = +\infty$, since, for every $\varepsilon > 0$, the inequality $\ell_{(\mathcal{M},\vec{\varphi})}(\vec{x}, (y_1+\varepsilon, \ldots, y_k+\varepsilon)) \leq L(\vec{x}, \vec{y})$ holds by definition, and hence the equality $L(\vec{x}, \vec{y}) = +\infty$ immediately follows. Let us now assume that $\lim_{\varepsilon\to 0^+} \ell_{(\mathcal{M},\vec{\varphi})}(\vec{x}, (y_1+\varepsilon, \ldots, y_k+\varepsilon)) = r < +\infty$. In this case a finite set $\mathcal{P} = \{P_1, \ldots, P_r\}$ of points in $\mathcal{M}\langle\vec{\varphi} \preceq \vec{x}\rangle$ exists such that, for every small enough $\varepsilon > 0$, every $P \in \mathcal{M}\langle\vec{\varphi} \preceq \vec{x}\rangle$ is $\langle\vec{\varphi} \preceq (y_1+\varepsilon, \ldots, y_k+\varepsilon)\rangle$-connected to a point $P_j \in \mathcal{P}$ in $\mathcal{M}$. Furthermore, for $i \neq j$ the points $P_i, P_j \in \mathcal{P}$ are not $\langle\vec{\varphi} \preceq (y_1+\varepsilon, \ldots, y_k+\varepsilon)\rangle$-connected and hence not $\langle\vec{\varphi} \preceq \vec{y}\rangle$-connected either. From Lemma 1 it follows that every $P \in \mathcal{M}\langle\vec{\varphi} \preceq \vec{x}\rangle$ is $\langle\vec{\varphi} \preceq \vec{y}\rangle$)-connected to a point $P_j \in \mathcal{P}$ in $\mathcal{M}$. Therefore $L(\vec{x}, \vec{y}) = r = \lim_{\varepsilon\to 0^+} \ell_{(\mathcal{M},\vec{\varphi})}(\vec{x}, (y_1+\varepsilon, \ldots, y_k+\varepsilon))$. □

Theorem 3. *For $t \in I = [0,1]$, consider the family of size pairs $(\mathcal{M}, f_t)$ where f_t is a homotopy between $f_0 : \mathcal{M} \to \mathbb{R}$ and $f_1 : \mathcal{M} \to \mathbb{R}$. Define $\vec{\chi} : \mathcal{M} \times I \to \mathbb{R}^3$ by $\vec{\chi}(P,t) = (f_t(P), t, -t)$. Then, for every $\bar{t} \in I$ and $\bar{x}, \bar{y} \in \mathbb{R}$ with $\bar{x} \leq \bar{y}$, it holds that*

$$rank\, H_0^{\bar{x},\bar{y}}(\bar{t}) = \lim_{\varepsilon\to 0^+} \ell_{(\mathcal{M}\times I,\vec{\chi})}(\bar{x}, \bar{t}, -\bar{t}, \bar{y}+\varepsilon, \bar{t}+\varepsilon, -\bar{t}+\varepsilon),$$

where $H_0^{\bar{x},\bar{y}}(\bar{t})$ denotes the dimension 0 persistent homology group computed at point $(\bar{x}, \bar{y})$ with respect to $f_{\bar{t}}$.

Proof. We know that $rank\, H_0^{\bar{x},\bar{y}}(\bar{t})$ is equal to the number of equivalence classes of $\mathcal{M}\langle f_{\bar{t}} \leq \bar{x}\rangle$ quotiented with respect to the $\langle f_{\bar{t}} \leq \bar{y}\rangle$-connectedness relation. On the other hand, Lemma 2 states that $\lim_{\varepsilon\to 0^+} \ell_{(\mathcal{M}\times I,\vec{\chi})}(\bar{x}, \bar{t}, -\bar{t}, \bar{y}+\varepsilon, \bar{t}+\varepsilon, -\bar{t}+\varepsilon)$ is equal to the number of equivalence classes of $\mathcal{M} \times I\langle\vec{\chi} \preceq (\bar{x}, \bar{t}, -\bar{t})\rangle$ quotiented, with respect to the $\langle\vec{\chi} \preceq (\bar{y}, \bar{t}, -\bar{t})\rangle$-connectedness relation. By definition of $\vec{\chi}$, this last number equals the number of equivalence classes of $\mathcal{M}\langle f_{\bar{t}} \leq \bar{x}\rangle$ quotiented, with respect to the $\langle f_{\bar{t}} \leq \bar{y}\rangle$-connectedness relation. This concludes our proof. □

However, although these two links exist, the concept of multidimensional size function has quite different purposes than that of vineyard. First of all, vineyards are based on a 1-parameter parallel foliation of $\mathbb{R}^3$, while the study of multidimensional size functions depends on a $(2k-2)$-parameter non-parallel foliation of $\Delta^+ \subseteq \mathbb{R}^k \times \mathbb{R}^k$. In fact, multidimensional size functions are associated with k-dimensional measuring functions, instead of with a homotopy between 1-dimensional measuring functions. Furthermore, [5] does not aim to identify distances for the comparison of vineyards, while we are interested in quantitative methods for comparing multidimensional size functions.

6 Conclusions and future work

In this paper we have proved that the theory of multidimensional size functions can be reduced to the 1-dimensional case by a suitable change of variables. This equivalence implies that multidimensional size functions are stable, with respect to the new distance D_{match}, and this allows us to use them in concrete applications by exploiting the existing computational techniques.

Many theoretical problems deserve further investigation, among them we list a few here.

- **Choice of the foliation.** Other foliations, different from the one we propose are possible. In general, we can choose a family Γ of continuous curves $\vec{\gamma}_{\vec{\alpha}} : \mathbb{R} \to \mathbb{R}^k$ such that (i) for $s < t$, $\vec{\gamma}_{\vec{\alpha}}(s) \prec \vec{\gamma}_{\vec{\alpha}}(t)$, (ii) for every $(\vec{x}, \vec{y}) \in \Delta^+$ there is one and only one $\vec{\gamma}_{\vec{\alpha}} \in \Gamma$ through $\vec{x}$, $\vec{y}$ and (iii) the curve $\gamma_{\vec{\alpha}}$ depends continuously on the parameter $\vec{\alpha}$ (this last hypothesis is important in computation for stability reasons). It would be interesting to study the dependence of our results on the choice of the foliation.

- **Extension to the algebraic context.** We think that the main results obtained in this paper for multidimensional size functions can be straightforwardly extended to the ranks of size homotopy groups and persistent homology groups.

Acknowledgements

Work performed within the activity of ARCES "E. De Castro", University of Bologna, under the auspices of INdAM-GNSAGA, of the University of Bologna, funds for selected research topics, and of the University of Modena and Reggio Emilia.

This paper is dedicated to the memory of Marco Gori.

References

[1] M. Allili and D. Corriveau and D. Ziou, *Morse Homology Descriptor for Shape Characterization*, ICPR'04 **4** (2004), 27–30.

[2] F. Cagliari, M. Ferri and P. Pozzi, *Size functions from a categorical viewpoint*, Acta Appl. Math. **67** (2001), 225-235.

[3] G. Carlsson and A. Zomorodian, *The Theory of Multidimensional Persistence* (preprint).

[4] D. Cohen-Steiner, H. Edelsbrunner, J. Harer, *Stability of persistence Diagrams*, Proc. 21st Sympos. Comput. Geom. 2005, 263–271.

[5] D. Cohen-Steiner, H. Edelsbrunner, D. Morozov, *Vines and Vineyards by Updating Persistence in Linear Time*, SCG'06, June 5-7, 2006, Sedona, Arizona, USA.

[6] M. d'Amico, P. Frosini and C. Landi, *Natural pseudo-distances and optimal matching between reduced size functions*, Tech. Rept. 66, DISMI, Univ. degli Studi di Modena e Reggio Emilia, Italy, 2005.

[7] M. d'Amico, P. Frosini and C. Landi, *Using matching distance in Size Theory: a survey*, International Journal of Imaging Systems and Technology (to appear).

[8] F. Dibos, P. Frosini and D. Pasquignon, *The use of Size Functions for Comparison of Shapes through Differential Invariants*, Journal of Mathematical Imaging and Vision **21** (2004), no. 2, 107–118.

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

[10] P. Donatini and P. Frosini, *Lower bounds for natural pseudodistances via size functions*, Archives of Inequalities and Applications **2** (2004), no. 1.

[11] P. Donatini and P. Frosini, *Natural pseudodistances between closed surfaces*, Journal of the European Mathematical Society (to appear).

[12] H. Edelsbrunner, D. Letscher and A. Zomorodian, *Topological Persistence and Simplification*, Proc. 41st Ann. IEEE Sympos. Found Comput. Sci. (2000), 454–463.

[13] H. Edelsbrunner, D. Letscher and A. Zomorodian, *Topological Persistence and Simplification*, Discrete Comput. Geom. **28** (2002), 511–533.

[14] P. Frosini, *A distance for similarity classes of submanifolds of a Euclidean space*, Bull. Austral. Math. Soc. **42** (1990), no. 3, 407–416.

[15] P. Frosini, *Measuring shapes by size functions*, Proc. of SPIE, Intelligent Robots and Computer Vision X: Algorithms and Techniques, Boston, MA 1607 (1991), 122–133.

[16] P. Frosini, *Connections between size functions and critical points*, Math. Meth. Appl. Sci. **19** (1996), 555–569.

[17] P. Frosini and C. Landi, *Size functions and formal series*, Appl. Algebra Engrg. Comm. Comput., **12** (2001), 327–349.

[18] P. Frosini and C. Landi, *Size functions and morphological transformations*, Acta Appl. Math. **49** (1997), no. 1, 85–104.

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

[20] T. Kaczynski, K. Mischaikow and M. Mrozek, Computational Homology, Applied Mathematical Sciences **157**, Springer-Verlag, New York (2004).

[21] C. Uras and A. Verri, *Computing size functions from edge maps*, Internat. J. Comput. Vision **23** (1997), no. 2, 169–183.

[22] A. Verri and C. Uras, *Invariant size functions*, Applications of invariance in computer vision. Lecture Notes in Comput. Sci. **825** (1993), 215–234.

[23] A. Verri and C. Uras, *Metric-topological approach to shape representation and recognition*, Image Vision Comput. **14** (1996), 189–207.

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

[25] S. Willard, *General Topology*, Addison-Wesley Publishing Company, 1970.