Journal of Symbolic Computation 78 (2017) 61–75

Contents lists available at ScienceDirect

Journal of Symbolic Computation

www.elsevier.com/locate/jsc

# Reducing complexes in multidimensional persistent homology theory

Madjid Allili [a,1], Tomasz Kaczynski [b,2], Claudia Landi [c,2]

[a] Bishop's University, Mathematics Department, Sherbrooke (Québec) J1M 1Z7, Canada
[b] Université de Sherbrooke, Département de Mathématiques, Sherbrooke (Québec) J1K 2R1, Canada
[c] Dipartimento di Scienze e Metodi dell'Ingegneria, Università di Modena e Reggio Emilia, Reggio Emilia, Italy

ARTICLE INFO

*Article history:*
Received 30 January 2015
Accepted 19 November 2015
Available online 25 March 2016

*Keywords:*
Multidimensional persistent homology
Discrete Morse theory
Acyclic partial matchings
Matching algorithm

ABSTRACT

Forman's discrete Morse theory appeared to be useful for providing filtration-preserving reductions of complexes in the study of persistent homology. So far, the algorithms computing discrete Morse matchings have only been used for one-dimensional filtrations. This paper is perhaps the first attempt in the direction of extending such algorithms to multidimensional filtrations. An initial framework related to Morse matchings for the multidimensional setting is proposed, and a matching algorithm given by King, Knudson, and Mramor is extended in this direction. The correctness of the algorithm is proved, and its complexity analyzed. The algorithm is used for establishing a reduction of a simplicial complex to a smaller but not necessarily optimal cellular complex. First experiments with filtrations of triangular meshes are presented.



## 1. Introduction

The *persistent homology* has been intensely developed in the last decade as a tool for studying problems of two kinds. One is the topological analysis of discrete data, e.g. *point-cloud data*, where the chosen framework is a discrete linear filtration of simplicial complexes. The first contributions in this

*E-mail addresses:* mallili@ubishops.ca (M. Allili), t.kaczynski@usherbrooke.ca (T. Kaczynski), claudia.landi@unimore.it (C. Landi).
[1] Tel. +1 (819) 822 9600 (2740); fax: +1 (819) 822 9661.
[2] This work was partially supported by the following institutions: INdAM-GNSAGA (C.L.), NSERC Canada Discovery Grant (T.K.).

http://dx.doi.org/10.1016/j.jsc.2015.11.020
0747-7171/

direction given by Edelsbrunner et al. (2002), and later by Zomorodian and Carlsson (2005) opened a new direction in research. The other one is the study of shape similarity by *shape-from-function* methods, where the framework is the filtration of a topological triangulable space by the values of a continuous function called *measuring function*. The 0-dimensional persistent homology case, where the topological invariants are based on the number of connected components, were known under the name of the *size function* theory since the paper by Frosini (1991). The applications of persistent homology to shape similarity are studied by Verri et al. (1993), Carlsson et al. (2005), Di Fabio and Landi (2012). The two frameworks, discrete and continuous, have been extended to the multiparameter filtration case called *multidimensional persistence*, where the filtration is set up with respect to a parameter space that is no longer ordered linearly (Carlsson and Zomorodian, 2007; Biasotti et al., 2008; Cagliari et al., 2010; Cerri et al., 2013). In the continuous setting this gives rise to *multidimensional measuring functions*, that is functions with values in $\mathbb{R}^k$. In Cavazza et al. (2013) the relation between the discrete and continuous settings is established.

In parallel, another mathematical theory which became increasingly popular in computational sciences is Forman's discrete Morse theory (Forman, 1998). We will not elaborate on all possible applications of this theory in visualization, imaging, computational geometry and other fields but just point out the one to computing persistence. The effective computation of the persistent homology is a challenge due to a huge size of complexes built from data, for instance, via meshing techniques. Discrete Morse theory enables algorithms reducing a given complex (simplicial, cubical, or cellular) to a much smaller cellular complex, homotopically equivalent to the initial one, by means of *Morse matching*, also called *Morse pairing*. An ultimate goal is often to reduce the complex to an optimal one, where all remaining cells are topologically significant. If a reduction by Morse pairings can be performed in a filtration-preserving way, that leads to a faster persistent homology computation. This goal motivated the contributions of King et al. (2005), Mischaikow and Nanda (2013), Robins et al. (2011), and Dłotko and Wagner (2014).

Given a complex and a partial pairing of its cells, the paired cells form a discrete vector field in the language of discrete Morse theory and can be reduced in pairs so to obtain at each step a new complex homotopically equivalent to the previous one. The final complex consists of unpaired cells that are also called critical cells. First, we give an algorithm that constructs a Morse matching for a given complex and we prove its correctness and analyze its complexity. Then, we go on proving that given a multifiltration on the initial complex, the reduction process yields a new multifiltration consisting of smaller complexes and which has the same persistent homology as the initial one. As pointed out in Mischaikow and Nanda (2013) for the one dimensional case, the complexity of computing multidimensional persistent homology of a filtration is essentially determined by the sizes of its complexes. This motivates this approach of reducing the initial complexes for achieving a low computational cost in the persistent homology computation. Our matching algorithm can be considered as an extension to the multidimensional setting of the algorithms given in King et al. (2005) and Cerri et al. (2011). The multidimensionality is symbolized by the function defined on the vertices of the complex. The algorithm is of iterative and recursive nature. It considers every vertex of the complex and builds a partial matching recursively on its lower link before extending it to the entire complex. When the dimension is fixed and the number of cofaces of every cell in the complex is bounded above by a fixed constant, we prove that the computational complexity of the algorithm is linear in the number of vertices of the initial complex.

So far, the algorithms for discrete Morse pairings have only been used for one-parameter filtrations. There does not yet exist a systematic extension of Forman's discrete Morse theory to the multiparameter case, and this goal offers challenges on both theoretical and computational levels. This paper is the first attempt in this direction.

The paper is organized as follows. In Section 2, we recall definitions and some known facts about the structure of S-complexes, multidimensional persistent homology, acyclic matchings, and an S-complex reduction. Although our main algorithm takes simplicial complexes as an input, in order to make use of it for the computation of persistent homology, one shall need a more general class of complexes introduced in Mrozek and Batko (2009) under the name of S-complexes. This structure is equivalent to the structure of a *complex* introduced by Lefschetz (1942) and currently called *Lefschetz complex*. Roughly speaking, an S-complex is a convenient combinatorial framework for

a cellular complex consisting of a canonical basis of elements called cells graded by dimension, and of adjacency relations which permit to define the boundary map in the associated chain complex. In Section 3, we propose initial definitions of Morse pairings for the multidimensional setting and we extend the algorithm given by King, Knudson, and Mramor. We next prove the correctness of the algorithm. Note that we do not claim to obtain an optimal cellular complex. The set of cells we call *critical* is simply the set of all unpaired cells and, typically, this is not an optimal complex. We next establish our filtration-preserving complex reduction method. In Section 4, we present our first experiments with multifiltrations of triangular meshes. These experiments show a fair rate of reduction but not an optimal one in the sense that the remaining cells are not all relevant in the computation of persistent homology. An improvement of our methods towards the optimality is a research in progress.

## 2. Preliminaries

### 2.1. $\mathbb{S}$-complexes

We shall use the combinatorial framework of $\mathbb{S}$-complexes introduced in Mrozek and Batko (2009). Let $R$ be a principal ideal domain (PID) whose invertible elements we call *units*. Given a finite set $X$, let $R(X)$ denote the free module over $R$ generated by $X$.

Let $\mathbb{S}$ be a finite set with a gradation $\mathbb{S}_q$ such that $\mathbb{S}_q = \emptyset$ for $q < 0$. Then $R(\mathbb{S}_q)$ is a gradation of $R(\mathbb{S})$ in the category of moduli over the ring $R$. For every element $\sigma \in \mathbb{S}$ there exists a unique number $q$ such that $\sigma \in \mathbb{S}_q$. This number will be referred to as the dimension of $\sigma$ and denoted $\dim\sigma$.

Let $\kappa : \mathbb{S} \times \mathbb{S} \to R$ be a function such that, if $\kappa(\sigma,\tau) \neq 0$, then $\dim\sigma = \dim\tau + 1$.

We say that $(\mathbb{S},\kappa)$ is an $\mathbb{S}$-*complex* if $(C_*(\mathbb{S}), \partial^\kappa_*)$ with $C_q(\mathbb{S}) := R(\mathbb{S}_q)$ and $\partial^\kappa_q : C_q(\mathbb{S}) \to C_{q-1}(\mathbb{S})$ defined on generators $\sigma \in \mathbb{S}$ by

$$\partial^\kappa(\sigma) := \sum_{\tau\in\mathbb{S}} \kappa(\sigma,\tau)\tau$$

is a free chain complex with base $\mathbb{S}$. The map $\kappa$ will be referred to as the *incidence number*. If $\kappa(\sigma,\tau) \neq 0$, then we say that $\tau$ is a *primary face* of $\sigma$ and $\sigma$ is a *primary coface* of $\tau$. We say that $\tau$ is a *face* of $\sigma$ and $\sigma$ is a *coface* of $\tau$ if there is a sequence of generators ordered by the primary face relation starting with $\tau$ and ending with $\sigma$.

By the homology of an $\mathbb{S}$-complex $(\mathbb{S},\kappa)$ we mean the homology of the chain complex $(C_*(\mathbb{S}), \partial^\kappa_*)$, and we denote it by $H_*(\mathbb{S},\kappa)$ or simply by $H_*(\mathbb{S})$.

The choice of $R$ a PID is made for the sake of homology computations, and also because in the proof of Proposition 2.1 we actually use the cancellation law.

A special case of an $\mathbb{S}$ complex is the simplicial complex. A $q$-simplex $\sigma = [v_0, v_1, \ldots, v_q]$ in $\mathbb{R}^d$ is the convex hull of $q+1$ affinely independent points $v_0, v_1, \ldots, v_q$ in $\mathbb{R}^d$, called the vertices of $\sigma$. The number $q$ is the dimension of the simplex. A face of $\sigma$ is a simplex whose vertices constitute a subset of $(v_0, v_1, \ldots, v_q)$. A simplicial complex consists of a collection $\mathbb{S}$ of simplices such that every face of a simplex in $\mathbb{S}$ is in $\mathbb{S}$, and the intersection of two simplices in $\mathbb{S}$ is their common face. The simplicial complex $\mathbb{S}$ has a natural gradation $(\mathbb{S}_q)$, where $\mathbb{S}_q$ consists of simplices of dimension $q$. Since a zero dimensional simplex is the singleton of its unique vertex, $\mathbb{S}_0$ may be identified with the collection of all vertices of all simplices in the simplicial complex $\mathbb{S}$.

Assume an ordering of $\mathbb{S}_0$ is given and every simplex $\sigma$ in $\mathbb{S}$ is coded as $[v_0, v_1, \ldots v_q]$, where the vertices $v_0, v_1, \ldots v_q$ are listed according to the prescribed ordering of $\mathbb{S}_0$. By putting

$$\kappa(\sigma,\tau) := \begin{cases} (-1)^i & \text{if } \sigma = [v_0, v_1, \ldots, v_q] \\ & \text{and } \tau = [v_0, v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_q] \\ 0 & \text{otherwise} \end{cases}$$

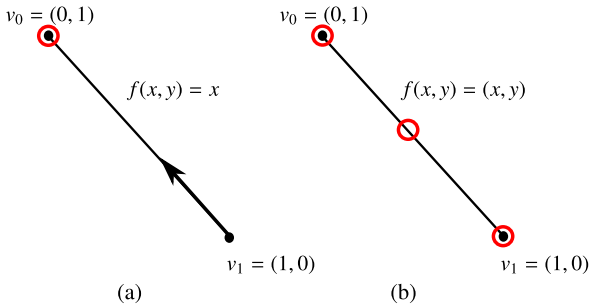



**Fig. 1.** The simplicial complex $\mathbb{S} = \{v_0, v_1, [v_0, v_1]\}$ is filtered by the 1D function $f(x, y) = x$ in (a), and by the 2D function $f(x, y) = (x, y)$ in (b). Simplexes whose insertion in the filtration changes the persistent homology and should be considered as critical are marked by red circles. The matchings of simplexes that can be removed in pairs without changing the persistent homology are represented by arrows. In (a), the distinct sublevel sets are $\emptyset \subset \mathbb{S}^0 = \{v_0\} \subset \mathbb{S}^1 = \mathbb{S}$. The only critical simplex of $\mathbb{S}$, that is a simplex which contributes to the change of homology when it enters a sublevel set, is $v_0$. The arrow shows that the edge $[v_0, v_1]$ is matched with $v_1$. In (b), the sublevel sets $\mathbb{S}^{(1,0)} = \{v_0\}$ and $\mathbb{S}^{(0,1)} = \{v_1\}$ are disjoint and they merge at $\alpha = (1, 1)$ when the edge $[v_0, v_1]$ is added to the filtration. Thus all three simplexes should be considered as critical. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

we obtain an $\mathbb{S}$-complex whose chain complex is the classical simplicial chain complex used in simplicial homology with the usual boundary map

$$\partial^{\kappa}([v_0, v_1, \dots v_q]) = \sum_{i=0}^{q} (-1)^i [v_0, v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_q].$$

Fig. 5 commented in Section 3.3 illustrates both simplicial and general $\mathbb{S}$ complexes as well as their interplay in the forthcoming reduction method.

### 2.2. Multidimensional persistent homology

Let $(\mathbb{S}, \kappa)$ be an $\mathbb{S}$-complex. A *multifiltration* of $\mathbb{S}$ is a family $\mathcal{F} = \{\mathbb{S}^{\alpha}\}_{\alpha \in \mathbb{R}^k}$ of subsets of $\mathbb{S}$ with the following properties:

(a) $\mathcal{F}$ is nested with respect to inclusions, that is $\mathbb{S}^{\alpha} \subseteq \mathbb{S}^{\beta}$, for every $\alpha \preceq \beta$, where $\alpha \preceq \beta$ if and only if $\alpha_i \leq \beta_i$ for all $i = 1, 2, \dots, k$;
(b) $\mathcal{F}$ is non-increasing on faces, that is, if $\sigma \in \mathbb{S}^{\alpha}$ and $\tau$ is a face of $\sigma$ then $\tau \in \mathbb{S}^{\alpha}$.

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

$$H_*(j^{(\alpha,\beta)}) : H_*(\mathbb{S}^{\alpha}) \to H_*(\mathbb{S}^{\beta})$$

induced by the inclusion map $j^{(\alpha,\beta)} : \mathbb{S}^{\alpha} \hookrightarrow \mathbb{S}^{\beta}$.

The image of the map $H_q(j^{(\alpha,\beta)})$ is known as the *q-th persistent homology group* of the filtration at $(\alpha, \beta)$ and we denote it by $H_q^{\alpha,\beta}(\mathbb{S})$. It contains the homology classes of order $q$ born not later than $\alpha$ and still alive at $\beta$.

The framework described so far for general filtrations can be specialized in various directions. A case relevant for a simplicial complex is when the filtration is induced by the values of a function defined at its vertices. Let $\mathbb{S}$ be a simplicial complex. Given a function $f : \mathbb{S}_0 \to \mathbb{R}^k$, it induces on $\mathbb{S}$ the so-called *sublevel set filtration*, defined as follows:

$$\mathbb{S}^{\alpha} = \{\sigma = [v_0, v_1, \dots, v_q] \in \mathbb{S} \mid f(v_i) \preceq \alpha,\ i = 0, \dots, q\}.$$

We will call the function $f$ a *measuring function*.

Fig. 1 illustrates the difference between sublevel set filtrations for 1D and 2D measuring functions.

### 2.3. Acyclic partial matchings

Let $(\mathtt{S}, \kappa)$ be an $\mathtt{S}$-complex. A *partial matching* $(\mathtt{A}, \mathtt{B}, \mathtt{C}, \mathtt{m})$ on $(\mathtt{S}, \kappa)$ is a partition of $\mathtt{S}$ into three sets $\mathtt{A}, \mathtt{B}, \mathtt{C}$ together with a bijective map $\mathtt{m} : \mathtt{A} \to \mathtt{B}$ such that, for each $\tau \in \mathtt{B}$, $\kappa(\mathtt{m}(\tau), \tau)$ is invertible. Observe that, in particular, $\mathtt{m}(\tau)$ is a primary coface of $\tau$.

A partial matching $(\mathtt{A}, \mathtt{B}, \mathtt{C}, \mathtt{m})$ on $(\mathtt{S}, \kappa)$ is called *acyclic* if there does not exist a sequence

$$\sigma_0, \tau_0, \sigma_1, \tau_1, \ldots, \sigma_p, \tau_p, \sigma_{p+1} \tag{1}$$

such that, $\sigma_{p+1} = \sigma_0$, and, for each $i = 0, \ldots, p$, $\sigma_{i+1} \neq \sigma_i$, $\tau_i = \mathtt{m}(\sigma_i)$, and $\tau_i$ is a primary coface of $\sigma_{i+1}$.

A convenient way to reformulate the definition of an acyclic partial matching is via Hasse diagrams. The *Hasse diagram* of $(\mathtt{S}, \kappa)$ is the directed graph whose vertices are elements of $\mathtt{S}$, and the edges are given by primary face relations and oriented from the larger element to the smaller one. Given a partial matching $(\mathtt{A}, \mathtt{B}, \mathtt{C}, \mathtt{m})$ on $(\mathtt{S}, \kappa)$, we change the orientation of the edge $(\tau, \sigma)$ whenever $\tau = \mathtt{m}(\sigma)$. The acyclicity condition says that the oriented graph obtained in this way, which is also called the *modified Hasse diagram* of $(\mathtt{S}, \kappa)$, has no nontrivial cycles. A directed graph with no directed cycles is called a directed acyclic graph (DAG). Thus, a partial matching $(\mathtt{A}, \mathtt{B}, \mathtt{C}, \mathtt{m})$ on $(\mathtt{S}, \kappa)$ is *acyclic* if its corresponding *modified Hasse diagram* is a DAG.

### 2.4. Reductions

We describe here a reduction construction which was introduced in Kaczynski et al. (1998) for finitely generated chain complexes, also presented in Kaczynski et al. (2004, Chapter 4). The construction was reused in Mrozek and Batko (2009) for the purposes of the coreduction method and, recently, in Mischaikow and Nanda (2013) for the one-dimensional filtration of $\mathtt{S}$-complexes, which is perhaps the closest reference for the purposes of this paper.

Let $(\mathtt{A}, \mathtt{B}, \mathtt{C}, \mathtt{m})$ be a partial matching (not necessarily acyclic) on an $\mathtt{S}$-complex $(\mathtt{S}, \kappa)$. Given $\sigma \in \mathtt{A}$, a new $\mathtt{S}$-complex $(\overline{\mathtt{S}}, \overline{\kappa})$ is constructed by setting $\overline{\mathtt{S}} = \mathtt{S} \setminus \{\mathtt{m}(\sigma), \sigma\}$, and $\overline{\kappa} : \overline{\mathtt{S}} \times \overline{\mathtt{S}} \to R$,

$$\overline{\kappa}(\eta, \xi) = \kappa(\eta, \xi) - \frac{\kappa(\eta, \sigma)\kappa(\mathtt{m}(\sigma), \xi)}{\kappa(\mathtt{m}(\sigma), \sigma)}. \tag{2}$$

Note that $\kappa(\mathtt{m}(\sigma), \sigma)$ is invertible by the definition of a partial matching. We say that $(\overline{\mathtt{S}}, \overline{\kappa})$ is obtained from $(\mathtt{S}, \kappa)$ by a *reduction* of the pair $(\mathtt{m}(\sigma), \sigma)$.

A pair of linear maps $\pi : C_*(\mathtt{S}) \to C_*(\overline{\mathtt{S}})$ and $\iota : C_*(\overline{\mathtt{S}}) \to C_*(\mathtt{S})$ is defined on generators by setting

$$\pi(\tau) = \begin{cases} 0 & \text{if } \tau = \mathtt{m}(\sigma) \\ -\sum_{\xi \in \overline{\mathtt{S}}} \frac{\kappa(\mathtt{m}(\sigma), \xi)}{\kappa(\mathtt{m}(\sigma), \sigma)} \xi & \text{if } \tau = \sigma \\ \tau & \text{otherwise} \end{cases} \tag{3}$$

and

$$\iota(\tau) = \tau - \frac{\kappa(\tau, \sigma)}{\kappa(\mathtt{m}(\sigma), \sigma)} \mathtt{m}(\sigma). \tag{4}$$

It is known (Kaczynski et al., 1998) that $C_*(\overline{\mathtt{S}})$ is a well-defined chain complex, and that $\pi$ and $\iota$ are chain equivalences with the chain homotopy $D_* : C_*(\mathtt{S}) \to C_{*+1}(\mathtt{S})$ given on generators $\tau \in \mathtt{S}_q$, $q \in \mathbb{Z}$, by

$$D_q(\tau) = \begin{cases} \frac{1}{\kappa(\mathtt{m}(\sigma), \sigma)} \mathtt{m}(\sigma) & \text{if } \tau = \sigma \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

As a consequence, $H_*(\mathtt{S}) = H_*(\overline{\mathtt{S}})$.

Let $(\mathtt{A}, \mathtt{B}, \mathtt{C}, \mathtt{m})$ be an acyclic partial matching on an $\mathtt{S}$-complex $(\mathtt{S}, \kappa)$. Let $(\overline{\mathtt{S}}, \overline{\kappa})$ be obtained from $(\mathtt{S}, \kappa)$ by reduction of the pair $(\mathtt{m}(\sigma), \sigma)$, $\sigma \in \mathtt{A}$.

**Proposition 2.1.** *If* $(\mathtt{A}, \mathtt{B}, \mathtt{C}, \mathtt{m})$ *is acyclic then, for any* $\tau \in \mathtt{A} \setminus \{\sigma\}$, $\overline{\kappa}(\mathtt{m}(\tau), \tau)$ *is invertible. Furthermore,* $\overline{\kappa}(\mathtt{m}(\tau), \tau) = \kappa(\mathtt{m}(\tau), \tau)$.

**Proof.** By definition,

$$\overline{\kappa}(\mathtt{m}(\tau), \tau) = \kappa(\mathtt{m}(\tau), \tau) - \frac{\kappa(\mathtt{m}(\tau), \sigma)\kappa(\mathtt{m}(\sigma), \tau)}{\kappa(\mathtt{m}(\sigma), \sigma)}.$$

If $\kappa(\mathtt{m}(\tau), \sigma)\kappa(\mathtt{m}(\sigma), \tau) = 0$, then $\overline{\kappa}(\mathtt{m}(\tau), \tau) = \kappa(\mathtt{m}(\tau), \tau)$ is invertible. Otherwise, $\kappa(\mathtt{m}(\tau), \sigma) \neq 0$ and $\kappa(\mathtt{m}(\sigma), \tau) \neq 0$. Hence $\sigma$ is a primary face of $\mathtt{m}(\tau)$ and $\tau$ is a primary face of $\mathtt{m}(\sigma)$. On the other hand, by definition of $\mathtt{m}$, $\sigma$ is a primary face of $\mathtt{m}(\sigma)$ and $\tau$ is a primary face of $\mathtt{m}(\tau)$. But this contradicts the assumption that the partial matching is acyclic. In conclusion, necessarily $\overline{\kappa}(\mathtt{m}(\tau), \tau) = \kappa(\mathtt{m}(\tau), \tau)$. □

**Corollary 2.2.** *Let* $(\mathtt{A}, \mathtt{B}, \mathtt{C}, \mathtt{m})$ *be an acyclic partial matching on* $(\mathtt{S}, \kappa)$. *Given a fixed* $\sigma \in \mathtt{A}$, *define* $\overline{\mathtt{A}} = \mathtt{A} \setminus \{\sigma\}$, $\overline{\mathtt{B}} = \mathtt{B} \setminus \{\mathtt{m}(\sigma)\}$, $\overline{\mathtt{m}} = \mathtt{m}_{|\overline{\mathtt{A}}}$, *and* $\overline{\mathtt{C}} = \mathtt{C}$. *Then* $(\overline{\mathtt{A}}, \overline{\mathtt{B}}, \overline{\mathtt{C}}, \overline{\mathtt{m}})$ *is an acyclic partial matching on* $(\overline{\mathtt{S}}, \overline{\kappa})$.

**Proof.** The bijectivity of $\overline{\mathtt{m}}$ is obvious by definition. The invertibility of $\overline{\kappa}(\mathtt{m}(\tau), \tau)$ has been already proved in Proposition 2.1. A cycle in the Hasse diagram of $(\overline{\mathtt{S}}, \overline{\kappa})$ is also a cycle in $(\mathtt{S}, \kappa)$, hence the acyclicity condition follows. □

Finally, we define the induced filtration on $\overline{\mathtt{S}}$.

**Definition 2.3.** *Let* $\mathcal{F} = \{\mathtt{S}^\alpha\}_{\alpha \in \mathbb{R}^k}$ *be a multifiltration on* $\mathtt{S}$. *Then* $\overline{\mathcal{F}} = \{\overline{\mathtt{S}}^\alpha\}_{\alpha \in \mathbb{R}^k}$ *is the multifiltration on* $\overline{\mathtt{S}}$ *defined by setting, for each* $\tau \in \overline{\mathtt{S}}$,

$$\tau \in \overline{\mathtt{S}}^\alpha \iff \tau \in \mathtt{S}^\alpha.$$

## 3. Main results

### 3.1. Matching algorithm

In this section we consider a finite simplicial complex $\mathtt{S}$ together with a function $f : \mathtt{S}_0 \to \mathbb{R}^k$ inducing the sublevel set filtration $\mathcal{F} = \{\mathtt{S}^\alpha\}_{\alpha \in \mathbb{R}^k}$. Given two values $\alpha = (\alpha_i), \beta = (\beta_i) \in \mathbb{R}^k$ we set $\alpha \prec \beta$ (resp. $\alpha \preceq \beta$) if and only if $\alpha_i < \beta_i$ (resp. $\alpha_i \leq \beta_i$) for every $i$ with $1 \leq i \leq k$. Moreover we write $\alpha \precneqq \beta$ whenever $\alpha \preceq \beta$ and $\alpha \neq \beta$.

#### 3.1.1. Indexing map for vertices

By definition, an indexing map on the vertices of the complex $\mathtt{S}$ is any one-to-one map $I : \mathtt{S}_0 \to \mathbb{N}$. Our objective is to build an indexing map $I$ such that, for each $v, w \in \mathtt{S}_0$ with $v \neq w$, $f(v) \precneqq f(w)$ implies $I(v) < I(w)$. For this purpose, we will use topological sorting of the vertices in $\mathtt{S}_0$.

We recall that a topological sorting of a directed graph is a linear ordering of its vertices such that for every directed edge $(u, v)$ from vertex $u$ to vertex $v$, $u$ precedes $v$ in the ordering. This ordering is possible if and only if the graph has no directed cycles, that is, if it is a DAG. A simple well known algorithm (see Wikipedia, 2014; Kahn, 1962) for this task consists in successively finding vertices of the DAG that have no incoming edges and placing them in a list for the final sorting. Note that at least one such vertex must exist in a DAG, otherwise, the graph must have at least one directed cycle. Let $\mathtt{L}$ denote the list that will contain the sorted vertices of $\mathtt{S}_0$ and $\mathtt{I}$ the list of vertices or nodes in the DAG with no incoming edges. The algorithm consists of two nested loops as summarized below.

**Algorithm 3.1** *(Topological sorting).*
**Input:** $\mathtt{I}$, the list of vertices.
**Output:** $\mathtt{L}$, the list of sorted vertices.
**while** there are vertices remaining in $\mathtt{I}$ **do**

**remove** a vertex u from I
**add** u to L
**for** each vertex v with an edge e from u to v **do**
**remove** edge e from the DAG
**if** v has no other incoming edges **then**
**insert** v into I
**End**

When the graph is a DAG, there exists at least one solution for the sorting problem, which is not necessarily unique. We can easily see that the algorithm visits at most once every node and every edge of the DAG, therefore its running time is linear in the number of nodes plus the number of edges in the DAG.

**Lemma 3.2.** *There exists an injective function* $I : \mathtt{S}_0 \to \mathbb{N}$ *such that, for each* $v, w \in \mathtt{S}_0$ *with* $v \neq w$, $f(v) \precneqq f(w)$ *implies* $I(v) < I(w)$.

**Proof.** Let us denote by $N$ the cardinality of $\mathtt{S}_0$. The set $\mathtt{S}_0$ can be represented in a directed graph where each vertex is a node, and a directed edge is drawn between two vertices $v, w \in \mathtt{S}_0$ if and only if $f(v) \precneqq f(w)$. It is easily seen that we actually obtain a directed acyclic graph (DAG), since a directed cycle in $\mathtt{S}_0$ leads to the relation $f(u) \precneqq f(u)$ for some vertex $u \in \mathtt{S}_0$, which is a contradiction. The topological sorting algorithm outlined above will allow to sort and store the vertices in $\mathtt{S}_0$ in an array L of size $N$, with indexes that can be chosen from 1 to $N$. It follows that the map $I : \mathtt{S}_0 \to \mathbb{N}$ that associates to every vertex its index in the array L is an injective map on $\mathtt{S}_0$. Moreover, and due to topological sorting, $I$ satisfies the constraint that for $v, w \in \mathtt{S}_0$ with $v \neq w$, $f(v) \precneqq f(w)$ implies $I(v) < I(w)$. □

Given a vertex $v$ of S and a simplex $\sigma \in \mathtt{S}$ with vertices affinely independent of $v$, we denote by $v * \sigma$ the *join* of $v$ and $\sigma$ which is, in our geometric setting, the convex hull of $\{v\} \cup \sigma$. We further denote by $\mathtt{S}'(v)$ the lower link of $v$ which is defined by the following formula

$$\mathtt{S}'(v) = \{\tau \in \mathtt{S} \mid v * \tau \in \mathtt{S} \wedge \forall \text{ vertex } w \in \tau,\ f(w) \precneqq f(v)\}. \tag{6}$$

**Algorithm 3.3** *(Matching).*
**Input:** A finite simplicial complex S with a function $f : \mathtt{S}_0 \to \mathbb{R}^k$ and an indexing $I : \mathtt{S}_0 \to \mathbb{N}$ on its vertices.
**Output:** Three lists A, B, C of simplices of S, and a function $\mathtt{m} : \mathtt{A} \to \mathtt{B}$.

**function** Partition (complex S, function $f$, indexing map $I$)
**Begin**

1. Initially, set $\mathtt{A}, \mathtt{B}, \mathtt{C} = \emptyset$.
2. For each $v \in \mathtt{S}_0$,
   (a) Compute $\mathtt{S}'(v)$, the lower link of $v$.
   (b) If $\mathtt{S}'(v)$ is empty, then add $v$ to C. Else
      i. add $v$ to A.
      ii. Let $f' : \mathtt{S}'_0(v) \to \mathbb{R}^k$ be the restriction of $f$ and $I' : \mathtt{S}'_0(v) \to \mathbb{N}$ be the restriction of $I$.
      iii. Call Partition (recursively) with input arguments $\mathtt{S}'(v)$, $f'$, and $I'$, and get the output $\mathtt{A}', \mathtt{B}', \mathtt{C}', \mathtt{m}'$.
      iv. Set $\mathtt{D}' = \{w \in \mathtt{C}'_0 \mid f(w)$ is minimal in $\mathtt{C}'_0$ w.r.t. $\precneqq\}$.
      v. Set $w_0$ as the vertex with smallest index $I'$ in $D'$.
      vi. Add $[v, w_0]$ to B and define $\mathtt{m}(v) = [v, w_0]$.
      vii. For each $\sigma \in \mathtt{C}' \setminus \{w_0\}$, add $v * \sigma$ to C.
      viii. For each $\sigma \in \mathtt{A}'$, add $v * \sigma$ to A, add $v * \mathtt{m}'(\sigma)$ to B, and define $\mathtt{m}(v * \sigma) = v * \mathtt{m}'(\sigma)$.

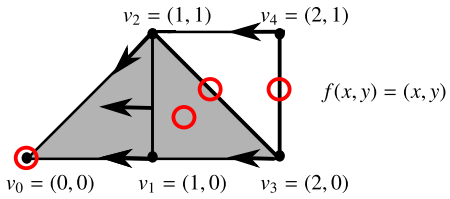



**Fig. 2.** Example of the output of Algorithm 3.3. Gray-shaded triangles are those which are present in the simplicial complex. The coordinates of the vertices are the values of the function $f$. The indexing is lexicographic on those coordinates and vertices are labeled by their indexes. Critical simplexes are marked by red circles and the matched simplexes are marked by arrows. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3. endfor.
4. For each $\sigma \in \mathtt{S} \setminus (\mathtt{A} \cup \mathtt{B} \cup \mathtt{C})$, add $\sigma$ to $\mathtt{C}$.
5. **return** $\mathtt{A}$, $\mathtt{B}$, $\mathtt{C}$, $\mathtt{m}$.

**End**

Fig. 2 illustrates the output of the algorithm on a simple example. Here are the main steps:

1. $\mathtt{S}'(v_0) = \emptyset$, add $v_0$ to $\mathtt{C}$.
2. $\mathtt{S}'(v_1) = \{v_0\}$, $\mathtt{m}(v_1) = [v_0, v_1]$.
3. $\mathtt{S}'(v_2) = \{v_0, v_1, [v_0, v_1]\}$, $v_0 \in \mathtt{C}'$, $v_1 \in \mathtt{A}'$, $[v_0, v_1] \in \mathtt{B}'$, by the recursive call (2)-(*iii*) of the Algorithm, $\mathtt{m}(v_2) = [v_0, v_2]$ and $\mathtt{m}([v_1, v_2]) = [v_0, v_1, v_2]$.
4. $\mathtt{S}'(v_3) = \{v_1\}$, $\mathtt{m}(v_3) = [v_1, v_3]$.
5. $\mathtt{S}'(v_4) = \{v_2, v_3\} = \mathtt{C}'$, by the recursive call (2)-(*iii*), $\mathtt{m}(v_4) = [v_2, v_4]$, add $[v_3, v_4]$ to $\mathtt{C}$.
6. Add remaining $[v_2, v_3]$ and $[v_1, v_2, v_3]$ to $\mathtt{C}$.

**Lemma 3.4.** $\mathtt{A}, \mathtt{B}, \mathtt{C}$ *is a partition of* $\mathtt{S}$ *and* $\mathtt{m}$ *is a bijective function from* $\mathtt{A}$ *to* $\mathtt{B}$.

**Proof.** $\mathtt{A} \cup \mathtt{B} \cup \mathtt{C} = \mathtt{S}$ by instruction 4. By construction (instruction viii), the map $\mathtt{m}$ is onto. We show, by induction on the dimension of simplices in $\mathtt{A}$ and $\mathtt{B}$, that $\mathtt{A} \cap \mathtt{B} = \emptyset$ and that $\mathtt{m}$ is injective. The proof of the equalities $\mathtt{A} \cap \mathtt{C} = \emptyset = \mathtt{B} \cap \mathtt{C}$ goes by similar arguments and we leave it to the reader. By instructions (b) and (i) vertices cannot belong to $\mathtt{B}$. Therefore the first claim is true for simplices of dimension 0. Moreover, the function $\mathtt{m}$ restricted to vertices of $\mathtt{A}$ is necessarily bijective. Indeed, if the edge $[v, w_0]$ is assigned to $v \in \mathtt{A}$ (instruction vi), that is $\mathtt{m}(v) = [v, w_0]$, it cannot be assigned again to $w_0$, because this would require that $v \in \mathtt{S}'(w_0)$. Then $f(v) \precneqq f(w_0)$ and $f(w_0) \precneqq f(v)$, a contradiction.

Let us now assume that the claim is true for simplices of dimension less than $n$. Let $\tau$ be a simplex of dimension $n$ in $\mathtt{A} \cap \mathtt{B}$. By instruction (viii) and since $\tau \in \mathtt{A}$, there exists $\sigma \in \mathtt{A}'$ (where $\mathtt{A}' = \mathtt{A}'(v)$) such that $\tau = v * \sigma$ and $\mathtt{m}(\tau) = v * \mathtt{m}'(\sigma)$. Since $\tau \in \mathtt{B}$, there exists $\beta \in \mathtt{A}$ such that $\tau = \mathtt{m}(\beta)$. Since $\dim \beta > 0$, there must exist a vertex $w$ and a simplex $\gamma \in \mathtt{A}'$ such that $\beta = w * \gamma$ and $\tau = w * \mathtt{m}'(\gamma)$, with $\mathtt{m}'(\gamma) \in \mathtt{B}'$. The vertices $v$ and $w$ must be equal, otherwise they must belong to the lower link of each other which would be a contradiction. It follows that $\sigma = \mathtt{m}'(\gamma) \in \mathtt{A}' \cap \mathtt{B}'$ (where $\mathtt{B}' = \mathtt{B}'(v)$) which violates the induction hypothesis.

We have proved that $\mathtt{A} \cap \mathtt{B} = \emptyset$ and we pass to the injectivity of $\mathtt{m}$. Let $\tau_1, \tau_2 \in \mathtt{A}$ be simplices of dimension $n$ such that $\mathtt{m}(\tau_1) = \mathtt{m}(\tau_2)$. There must exist vertices $v_1, v_2$ and simplices $\sigma_1, \sigma_2 \in \mathtt{A}'$ such that $\tau_1 = v_1 * \sigma_1$, $\tau_2 = v_2 * \sigma_2$ and

$$\mathtt{m}(\tau_1) = v_1 * \mathtt{m}'(\sigma_1) = \mathtt{m}(\tau_2) = v_2 * \mathtt{m}'(\sigma_2).$$

From what precedes, we can see that the vertices $v_1, v_2$ must be equal, otherwise they must belong to the lower link of each other. It follows that $\mathtt{m}'(\sigma_1) = \mathtt{m}'(\sigma_2)$ and, by the induction hypothesis, we must have $\sigma_1 = \sigma_2$ and therefore $\tau_1 = \tau_2$, which completes the proof. □

We define the map $\mathtt{maxI} : \mathtt{S} \to \mathbb{R}$ on simplices as follows

$$\mathtt{maxI}(\sigma) = \max_{v \text{ vertex of } \sigma} I(v).$$

**Lemma 3.5.** *The following statements hold.*

(a) *For every* $\sigma < \tau$, $\mathtt{maxI}(\sigma) \leq \mathtt{maxI}(\tau)$.
(b) *For every* $\sigma \in \mathtt{A}$, $\mathtt{maxI}(\sigma) = \mathtt{maxI}(\mathtt{m}(\sigma))$.

**Proof.** (a) is trivial from the definition of $\mathtt{maxI}$. (b) If $\sigma$ is a vertex $v$, then $\mathtt{m}(v) = [v, w]$ for some $w \in \mathtt{S}'(v)$. By Lemma 3.2, $I(w) < I(v)$ and hence $\mathtt{maxI}(v) = \mathtt{maxI}(\mathtt{m}(v))$. Let $\sigma \in \mathtt{A}$ be a simplex of dimension $n \geq 1$. There exist a vertex $v$ and a simplex $\sigma' \in \mathtt{A}' \subset \mathtt{S}'(v)$ such that $\sigma = v * \sigma'$ and $\mathtt{m}(\sigma) = v * \mathtt{m}'(\sigma')$. Since $\sigma'$ and $\mathtt{m}'(\sigma')$ are simplices of the lower link of $v$, it follows that both $\mathtt{maxI}(\sigma')$ and $\mathtt{maxI}(\mathtt{m}'(\sigma'))$ are smaller than $I(v)$. Thus

$$\mathtt{maxI}(\sigma) = \mathtt{maxI}(v * \sigma') = \mathtt{maxI}(v * \mathtt{m}(\sigma')) = \mathtt{maxI}(\mathtt{m}(\sigma)) = I(v). \qquad \square$$

**Theorem 3.6.** *Algorithm 3.3 produces a partial matching* $(\mathtt{A}, \mathtt{B}, \mathtt{C}, \mathtt{m})$ *that is acyclic. Moreover, if* $\sigma \in \mathtt{S}^\alpha$ *then* $\mathtt{m}(\sigma) \in \mathtt{S}^\alpha$.

**Proof.** The partial matching is acyclic if and only if it is a gradient vector field of a discrete Morse function. From Forman (2002, Theorem 6.2), this is equivalent to prove that there are no nontrivial closed directed paths in the modified Hasse diagram of the complex $\mathtt{S}$. Assume that

$$\ell : \sigma_0 \xrightarrow{\mathtt{m}} \tau_0 \xrightarrow{\succ} \sigma_1 \xrightarrow{\mathtt{m}} \tau_1 \ldots \xrightarrow{\mathtt{m}} \tau_n \xrightarrow{\succ} \sigma_0 \tag{7}$$

is a directed loop in the modified Hasse diagram, where $\mathtt{m}$ stands for the matching and the symbol $\succ$ for the face relation. From Lemma 3.5, we deduce that $\mathtt{maxI}$ is nondecreasing along any directed path in the modified Hasse diagram. It follows that $\mathtt{maxI}$ has to be constant along any directed loop. Thus, there must exist a unique vertex $v$ such that

$$\mathtt{maxI}(\sigma_0) = \mathtt{maxI}(\tau_0) = \mathtt{maxI}(\sigma_1) = \ldots = \mathtt{maxI}(\tau_n) = I(v),$$

and $v$ must belong to every $\sigma_i, \tau_i \in \ell$. We will prove by induction that this leads to a contradiction. First, observe that if $\dim \sigma_0 = \dim \sigma_1 = \ldots = 0$, then either these vertices are equal, in which case the loop is trivial, or they are distinct in which case $\mathtt{maxI} = I$ (on vertices) cannot be constant since it is injective. It follows that we cannot have a directed loop $\ell$ with cells of dimensions 0 and 1. Assume this claim is true up to dimensions $n-2$ and $n-1$, and suppose our directed loop $\ell$ in (7) is composed of cells $\sigma_i$ of dimension $n-1$ and cells $\tau_i$ of dimension $n$. We have proved that $v$ is a vertex of each $\sigma_i, \tau_i \in \ell$, so there exist simplices $\sigma_0', \sigma_1', \ldots, \sigma_n', \tau_0', \tau_1', \ldots, \tau_n'$ in $\mathtt{S}$ such that $\sigma_i = v * \sigma_i'$ and $\tau_i = v * \tau_i'$. It is easily seen that $\sigma_{i+1} \prec \tau_i$ implies that $\sigma_{i+1}' \prec \tau_i'$. On the other hand, $\tau_i = \mathtt{m}(\sigma_i)$ means that there must exist a vertex $w_i$ and a simplex $\gamma_i' \in A' \subset \mathtt{S}'(w_i)$ such that $\sigma_i = w_i * \gamma_i'$ and $\tau_i = w_i * \mathtt{m}'(\gamma_i')$. Using the same arguments as in the proof of Lemma 3.4, we conclude that we must have $v = w_i$ and therefore $\gamma_i' = \sigma_i'$ and $\mathtt{m}'(\gamma_i') = \tau_i'$. This shows also that $\sigma_i'$ and $\tau_i'$ have to be in $\mathtt{S}'(v)$. We can see now that we have a directed loop

$$\ell' : \sigma_0' \xrightarrow{\mathtt{m}'} \tau_0' \xrightarrow{\succ} \sigma_1' \xrightarrow{\mathtt{m}'} \ldots \xrightarrow{\mathtt{m}'} \tau_n' \xrightarrow{\succ} \sigma_0'$$

in the modified Hasse diagram of $\mathtt{S}'(v)$ with simplices of dimensions $n-2$ and $n-1$, which violates the induction hypothesis.

Let now $\sigma$ be a simplex of $(\mathtt{S}, \kappa)$ such that $\sigma \in \mathtt{S}^\alpha$. By definition of $\mathtt{m}$, there exist a vertex $v$ and simplices $\sigma', \tau' \in \mathtt{S}'(v)$ such that $\sigma = v * \sigma'$ and $\mathtt{m}(\sigma) = v * \tau'$. By definition of lower link and $\mathtt{S}^\alpha$, it follows that for every vertex $w$ in $\sigma'$ or $\tau'$, $f(w) \precneqq f(v) \preceq \alpha$. Hence, $\mathtt{m}(\sigma) \in \mathtt{S}^\alpha$. $\square$

**Remark 3.7.** A variation of partial matching may be obtained by replacing the lower link $\mathtt{S}'(v)$ in formula (6) with the *weak lower link* defined by

$$\mathtt{S}''(v) = \{\tau \in \mathtt{S} \mid v * \tau \in \mathtt{S} \wedge \forall \text{ vertex } w \in \tau,\ f(w) \preceq f(v)\},$$

and analogously replacing $\precneqq$ by $\preceq$ in the definition of $\mathtt{D}'$. The condition $v * \tau \in \mathtt{S}$ implies that $v$ is not in its weak lower link. The injectivity of $I$ and the instruction 2(b)-v of the algorithm permit carrying on the proofs. We considered this version of the algorithm with the hope of matching more cells, however our experiments did not show a significant improvement in terms of getting a more accurate set $\mathtt{C}$.

### 3.2. Complexity analysis

We first describe the computational complexity of Algorithm 3.3. Let $d$ be the dimension of the complex $\mathtt{S}$. For each $\sigma \in \mathtt{S}$, we define $\deg(\sigma)$ to be the cardinality of the set of all cofaces of $\sigma$.

We recall that $N$ is defined to be the cardinality of $\mathtt{S}_0$, i.e. the number of vertices in $\mathtt{S}$. For a vertex $v \in \mathtt{S}_0$, its lower link $\mathtt{S}'(v)$, which is a subcomplex of $\mathtt{S}$, consists of at most $\deg(v)$ simplices of dimensions smaller or equal to $d-1$. It follows that $\mathtt{S}'(v)$ has at most $\deg(v)d$ vertices. If we assume the worst case scenario where every vertex has a nonempty lower link, the first call to function Partition will result in $N$ subsequent calls for Partition, each for a fixed vertex $v \in \mathtt{S}_0$, with arguments $\mathtt{S}'(v)$ and the restrictions of $f$ and $I$ to $\mathtt{S}'(v)$. Since $\deg(v)$ varies for each vertex $v$, it is difficult to establish any complexity bounds for the algorithm without assuming some constraints on $\deg(v)$. We will assume hereafter that $\deg(v)$ is bounded above by a constant $\gamma$ for every $v \in \mathtt{S}_0$. This is a reasonable assumption when dealing with complexes of manifolds and approximating surface boundaries of objects. For each vertex $v \in \mathtt{S}_0$, we need to examine its set of cofaces (read directly from the structure storing the complex) to create its lower link which can be done first in at most $\gamma$ steps. The partition of the subcomplex $\mathtt{S}'(v)$ (resulting from the recursive call to Partition) will be visited once (in at most $\gamma$ steps) to execute the steps (b)-vi to (b)-viii of the algorithm. We assume that the vertices of $\mathtt{S}$ are already ordered with respect to the indexing function. It is easily seen that any subsequent call to Partition with a complex formed by a lower link of some vertex and of dimension $s < d$ is completed in a number of operations directly proportional to the number of simplices and vertices in the complex which are bounded by $\gamma$ and $\gamma(s+1)$ respectively. This number will be denoted by $\alpha(\gamma, s)$.

**Theorem 3.8.** *Algorithm 3.3 produces a partial matching* $(\mathtt{A}, \mathtt{B}, \mathtt{C}, \mathtt{m})$ *in less than* $2\gamma^d(d+1)!N$ *steps.*

**Proof.** From the discussion above, we deduce that the processing of each vertex of $\mathtt{S}$ is completed in less than $2\gamma + \alpha(\gamma, d-1)$ operations. Therefore, the number of operations for processing all the vertices (call it $\eta$) is bounded above by

$$N(2\gamma + \alpha(\gamma, d-1)).$$

Reasoning by induction and using the arguments above, each subsequent call to Partition on a complex of dimension $s$ of a lower link of a vertex costs less than $\alpha(\gamma, s)$ and we have

$$\alpha(\gamma, s) \leq \gamma(s+1)\left(2\gamma + \alpha(\gamma, s-1)\right).$$

Moreover, when the complex consists only of vertices, each of them will have an empty lower link in that complex. Thus, we can conclude that $\alpha(\gamma, 0) \leq \gamma$. Putting all together, we can conclude now that

$$\begin{aligned}\eta &\leq N\left[2\gamma + \alpha(\gamma, d-1)\right] \leq N\left[2\gamma + \gamma d\left[2\gamma + \alpha(\gamma, d-2)\right]\right] \\ &\leq N\left[2\gamma + 2\gamma^2 d + \gamma d\alpha(\gamma, d-2)\right].\end{aligned}$$

By induction, we can prove that

$$\begin{aligned}\eta &\leq N\left[2\gamma + 2\gamma^2 d + \ldots + 2\gamma^{(d-1)}d(d-1)\ldots 2 + \gamma^{(d-1)}d(d-1)\ldots 2\alpha(\gamma, 0)\right] \\ &\leq 2\gamma^d(d+1)!N. \qquad \square\end{aligned}$$

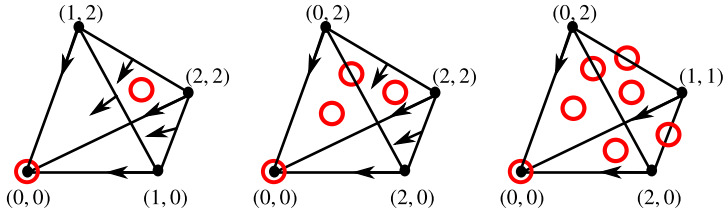



**Fig. 3.** Example of a sphere endowed with three different 2-dimensional maps on its vertices. Simplexes classified as critical by the algorithm are marked by red circles, whereas the matched simplexes are marked by arrows. As the number of vertices with non-comparable values increases, the number of matched simplexes decreases. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Let $n$ denote the total number of cells in the original complex $\mathtt{S}$. The computation of the rank invariant of a $d$-dimensional multifiltration of the complex $\mathtt{S}$ may be achieved with an algorithm that runs in $O(n^{2d+3})$ operations (see Carlsson et al. (2009) for more details). Our method which consists in using the acyclic matchings to perform homology preserving reductions on the original complex, allows to postpone the persistent homology computation until the complex is reduced to a smaller one which may yield a tremendous gain in the number of operations incurred. Let $m$ denote the number of cells in the final complex after all reductions yielded by the acyclic matching are performed. Thus, the computational cost of the multidimensional persistent homology of the complex $\mathtt{S}$ is reduced to $O(m^{2d+3})$. To illustrate the significance of our method, let us assume that our matching algorithm allows to reduce the complex by half its number of cells (a ratio that is comparable to the ones provided in our experimental results). In this case, the persistent homology computational cost is potentially reduced by a factor of $2^{2d+3}$ (slightly greater than 500 if $d = 3$) when the computation is performed on the reduced complex. This would be a major gain when comparing the computationally inexpensive reduction with the time consuming persistent homology computation.

Indeed, if we assume that we work under the constraint that $\deg(\sigma) \leq \gamma$ for every $\sigma \in \mathtt{S}$, we can easily prove that each elementary reduction is achieved in constant time. Hence, the time complexity of the total reduction process which runs through all the matching pairs $\{\mathtt{m}(\sigma), \sigma\}$ and performs the reductions is in the worst case linear in the number of cells of the complex.

Our aim is to make $m$ very small compared to $n$, or equivalently construct an optimal acyclic matching. However, this problem is known to be NP hard (see Joswig and Pfetsch, 2006) and there is no known procedure to minimize $m$ for arbitrary complexes. In our context where we are dealing with a multidimensional function and using an algorithm based on exploring lower links of vertices, it is possible to reach an outcome where no reduction is possible and every cell is critical. This point is illustrated in Fig. 4(a). In general, there will be cases in which the algorithm performs well and others in which it performs poorly in terms of the number of critical cells. The algorithm seems to work better when the number of non-comparable cells is lower. As an example, see Fig. 3.

Since our work is inspired by the work in King et al. (2005), it is natural to raise the question of whether it is possible to add some canceling step to reduce further the number of critical cells and allow a bigger number of reductions before proceeding with the persistent homology computation. Since our algorithm produces an acyclic matching of the complex, it is possible to build gradient paths and do cancellations when possible as defined in Forman (1998). However, the cancellation of critical cells is not necessarily desirable in this context because it works against providing a full account of the history of births and deaths of homology generators which is necessary for obtaining complete information about the persistent homology. This latter point can be easily illustrated with a simple example as shown in Fig. 4(b). A more extensive description on the consequences of cancellation of pairs of critical cells can be found in Delgado-Friedrichs et al. (2015), starting from an example very similar to ours.

### 3.3. Back to reductions

In this section, we prove that an acyclic matching on an $\mathtt{S}$-complex $(\mathtt{S}, \kappa)$ allows by means of reductions to replace the initial complex by a smaller one with the same persistent homology. The

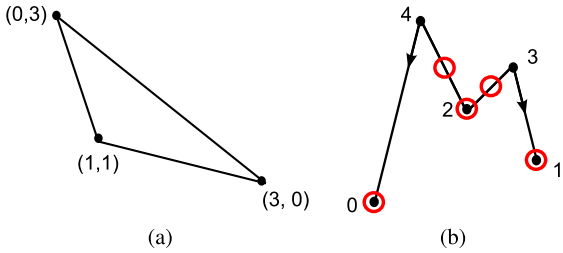



**Fig. 4.** (a) Example of a one dimensional complex with a 2-dimensional map on its vertices and in which every cell is critical according to Algorithm 3.3. (b) Example of a one dimensional complex with a one dimensional map on its vertices that can be extended to a Discrete Morse Function by assigning to each one dimensional cell the maximum of the values of its vertices. We can easily see that any cancellation of cells would lead to a change in the persistent homology of the complex.

motivation for this approach stems from the need to achieve a low computational cost in the persistent homology computation.

In the sequel, we assume that $(\mathtt{A}, \mathtt{B}, \mathtt{C}, \mathtt{m})$ is an acyclic matching on a filtered $\mathtt{S}$-complex $\mathtt{S}$ with the property:

$$\text{If } \sigma \in \mathtt{S}^\alpha \text{ then } \mathtt{m}(\sigma) \in \mathtt{S}^\alpha. \tag{8}$$

Theorem 3.6 asserts that the matching produced by Algorithm 3.3 on a filtered simplicial complex $\mathtt{S}$ has this property.

**Proposition 3.9.** *Let $\sigma \in \mathtt{A}$ and let $(\overline{\mathtt{S}}, \overline{\kappa})$ be obtained from $(\mathtt{S}, \kappa)$ by reduction of the pair $(\mathtt{m}(\sigma), \sigma)$. Let $\pi$, $\iota$, and $D$ be maps defined by formulas (3), (4), and (5) respectively. Then $\pi(C_*(\mathtt{S}^\alpha)) \subseteq C_*(\overline{\mathtt{S}}^\alpha)$, $\iota(C_*(\overline{\mathtt{S}}^\alpha)) \subseteq C_*(\mathtt{S}^\alpha)$, and $D_q(C_q(\mathtt{S}^\alpha)) \subseteq C_{q+1}(\overline{\mathtt{S}}^\alpha)$, for each $q \in \mathbb{Z}$.*

**Proof.** Let $\tau \in \mathtt{S}^\alpha$. We need to show that $\pi(\tau) \in C_*(\overline{\mathtt{S}}^\alpha)$. By definition of $\pi$, the only non-trivial case is when $\tau = \sigma$. In this case, $\sigma \in \mathtt{S}^\alpha$ and by (8), $\mathtt{m}(\sigma) \in \mathtt{S}^\alpha$. Note that the chain $\pi(\sigma)$ is supported in the union of cells $\xi \in \overline{\mathtt{S}}$ such that $\kappa(\mathtt{m}(\sigma), \xi) \neq 0$. Each such $\xi$ is a face of $\mathtt{m}(\sigma) \in \mathtt{S}^\alpha$, hence $\xi \in \overline{\mathtt{S}}^\alpha$.

Let now $\tau \in \overline{\mathtt{S}}^\alpha$. We need to show that $\iota(\tau) \in C_*(\mathtt{S}^\alpha)$. By definition of $\iota$, the only non-trivial case is when $\kappa(\tau, \sigma) \neq 0$. This implies that $\sigma$ is a face of $\tau$. Let $\tau \in \overline{\mathtt{S}}^\alpha$. By Definition 2.3, this means that $\tau \in \mathtt{S}^\alpha$. By definition of filtration, it follows that $\sigma \in \mathtt{S}^\alpha$. Again, by (8), $\mathtt{m}(\sigma) \in \mathtt{S}^\alpha$, proving the claim.

The statement on $D_*$ instantly follows by the same argument. □

**Lemma 3.10.** *The maps $\pi_{|C_*(\mathtt{S}^\alpha)} : C_*(\mathtt{S}^\alpha) \to C_*(\overline{\mathtt{S}}^\alpha)$ and $\iota_{|C_*(\overline{\mathtt{S}}^\alpha)} : C_*(\overline{\mathtt{S}}^\alpha) \to C_*(\mathtt{S}^\alpha)$ defined by restriction are chain homotopy equivalences. Moreover, the diagram*

$$\begin{array}{ccc} H_*(\mathtt{S}^\alpha) & \xrightarrow{H_*(j^{(\alpha,\beta)})} & H_*(\mathtt{S}^\beta) \\ \downarrow\cong & & \downarrow\cong \\ H_*(\overline{\mathtt{S}}^\alpha) & \xrightarrow{H_*(j^{(\alpha,\beta)})} & H_*(\overline{\mathtt{S}}^\beta) \end{array}$$

*commutes.*

**Proof.** By Proposition 3.9, we have the commutative diagram

$$\begin{array}{ccc} C_*(\mathtt{S}^\alpha) & \hookrightarrow & C_*(\mathtt{S}^\beta) \\ \downarrow{\scriptstyle \pi_{|C_*(\mathtt{S}^\alpha)}} & & \downarrow{\scriptstyle \pi_{|C_*(\mathtt{S}^\beta)}} \\ C_*(\overline{\mathtt{S}}^\alpha) & \hookrightarrow & C_*(\overline{\mathtt{S}}^\beta) \end{array}$$

where the vertical arrows are chain equivalences. The result follows by the functoriality of homology. □

This lemma immediately yields the following result.

**Theorem 3.11.** *For every $\alpha \preceq \beta \in \mathbb{R}^k$, $H_*^{\alpha,\beta}(\mathtt{S})$ is isomorphic to $H_*^{\alpha,\beta}(\overline{\mathtt{S}})$.*

Let us order $\mathtt{A}$ in a sequence

$$\mathtt{A} = \{\mathtt{A}(1), \mathtt{A}(2), \ldots, \mathtt{A}(n)\}$$

and set $\mathtt{B}(i) = \mathtt{m}(\mathtt{A}(i))$, $i = 1, 2, \ldots, n$. Put $\mathtt{S}(0) = \mathtt{S}$ and

$$\mathtt{S}(i) = \overline{\mathtt{S}(i-1)} = \mathtt{S}(i-1) \setminus \{\mathtt{B}(i), \mathtt{A}(i)\}, \;\; i = 1, 2, \ldots, n.$$

Since a partial matching defines a partition of $\mathtt{S}$, we have $\mathtt{S}(n) = \mathtt{C}$.

Note that, by Definition 2.3, the condition (8) carries through to the reduced complex. Consequently, Corollary 2.2, Lemma 3.10 and Theorem 3.11 extend by induction to any step of reduction. Hence, for any $\alpha \in \mathbb{R}^k$, we get a sequence of filtered $\mathtt{S}$-complexes

$$(\mathtt{S}^\alpha(0), \kappa^\alpha(0)), \; (\mathtt{S}^\alpha(1), \kappa^\alpha(1)), \; \ldots, \; (\mathtt{S}^\alpha(n), \kappa^\alpha(n)),$$

where $\kappa^\alpha(i) = \overline{\kappa^\alpha(i-1)}$, together with a sequence of chain equivalences

$$\pi^\alpha(i) : C_*(\mathtt{S}^\alpha(i-1)) \to C_*(\mathtt{S}^\alpha(i)), \;\; \iota^\alpha(i) : C_*(\mathtt{S}^\alpha(i)) \to C_*(\mathtt{S}^\alpha(i-1)).$$

Moreover, for any $\alpha \preceq \beta$, we get the sequence of inclusions

$$j^{(\alpha,\beta)}(i) : \mathtt{S}^\alpha(i) \hookrightarrow \mathtt{S}^\beta(i),$$

such that the commutative diagram of Lemma 3.10 applied to the $i$-th iterate gives the following.

$$\begin{array}{ccc} H_*(\mathtt{S}^\alpha(i-1)) & \xrightarrow{H_*(j^{(\alpha,\beta)}(i-1))} & H_*(\mathtt{S}^\beta(i-1)) \\ \downarrow\cong & & \downarrow\cong \\ H_*(\mathtt{S}^\alpha(i)) & \xrightarrow{H_*(j^{(\alpha,\beta)}(i))} & H_*(\mathtt{S}^\beta(i)) \end{array}$$

By induction, we get the following.

**Corollary 3.12.** *For every $\alpha \preceq \beta \in \mathbb{R}^k$, $H_*^{\alpha,\beta}(\mathtt{S})$ is isomorphic to $H_*^{\alpha,\beta}(\mathtt{C})$.*

The reduction procedure is illustrated by Fig. 5. The initial $\mathtt{S}$-complex $\mathtt{S} = \mathtt{S}(0)$ is a simplicial complex built of two edges $[v_0, v_1]$, $[v_0, v_2]$, one triangle $[v_1, v_2, v_3]$ and all their faces. The triangle $[v_0, v_1, v_2]$ is not a part of the complex. The arrows indicate an acyclic partial matching which is obtained by applying Algorithm 3.3 to the $1D$ function labeling the vertices. There are two critical simplices, the vertex $v_0$ and the edge $e_0 = [v_1, v_2]$. We remove the pairs $([v_1, v_2, v_3], [v_2, v_3])$, then $([v_1, v_3], v_3)$, next $([v_0, v_2], v_2)$ and finally $([v_0, v_1], v_1)$. In $\mathtt{S}(3)$, the boundary of the critical edge $e_0$ is replaced by $v_1 - v_0$ which is also the boundary of $[v_0, v_1]$. In $\mathtt{S}(4)$, the boundary of $e_0$ is set to zero.

## 4. Experimental results and conclusion

We considered four triangle meshes (available at The-GTS-Library, 2016.03.10). Each mesh was filtered by the 2-dimensional measuring function $f$ taking each vertex $v$ of coordinates $(x, y, z)$ to the pair $f(v) = (|x|, |y|)$.

In Table 1, the first row shows on the top line the number of vertices in each considered mesh, and in the middle line same quantities referred to the cell complex $\mathtt{C}$ obtained by using our matching algorithm to reduce $\mathtt{S}$. Finally, it also displays in the bottom line the ratio between the second and the first lines, expressing them in percentage points. The second and the third rows show similar information for the edges and the faces. Finally, the fourth row shows the same information for the total number of cells of each considered mesh $\mathtt{S}$.

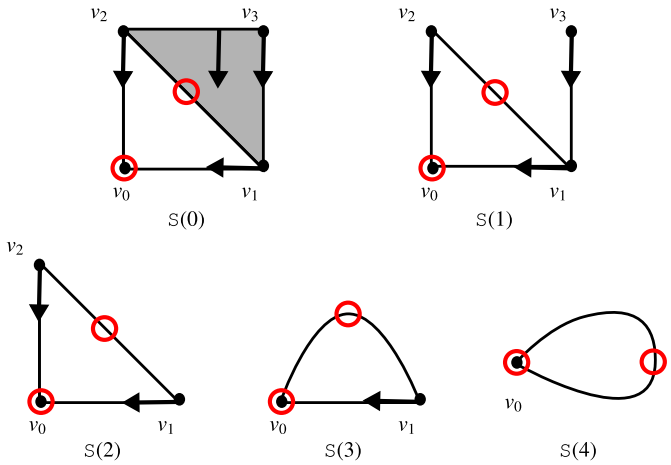
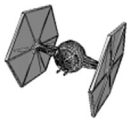



**Fig. 5.** Example of iterated S-complex reductions using the acyclic partial matching indicated by arrows. The initial complex S(0) as well as the first two iterates S(1) and S(2) are simplicial complexes. The last two iterates S(3) and S(4) cannot be geometrically realized as simplicial complexes: they require the framework of cellular complexes for that. Both simplicial and cellular complexes stay in the broad class of S-complexes.

**Table 1**
Reduction performance on some triangle meshes.

| |  tie |  space_shuttle |  x_wing |  space_station |
|---|---|---|---|---|
| $\#S_0$ | 2014 | 2376 | 3099 | 5749 |
| $\#C_0$ | 228 | 121 | 175 | 1879 |
| % | 11.3 | 5.1 | 5.6 | 32.7 |
| $\#S_1$ | 5944 | 6330 | 9190 | 15949 |
| $\#C_1$ | 3343 | 3699 | 3605 | 11158 |
| % | 56.2 | 58.4 | 39.2 | 70.0 |
| $\#S_2$ | 3827 | 3952 | 6076 | 10237 |
| $\#C_2$ | 3012 | 3576 | 3415 | 9316 |
| % | 78.7 | 90.5 | 56.2 | 91.0 |
| #S | 11785 | 12658 | 18365 | 31935 |
| #C | 6583 | 7396 | 7195 | 22353 |
| % | 55.9 | 58.4 | 39.2 | 70.0 |

Our experiments confirm that the current vertex-based matching algorithm does not produce optimal reduction of the complex so that every remaining cell is relevant in the computation of persistent homology. They show a fair rate of reduction for vertices, but the reduction rate for cells of dimensions 1 and 2 is not as significant as that for vertices. The discussion and the examples provided in subsection 3.2 illustrate some reasons for these findings.

## References

Biasotti, S., Cerri, A., Frosini, P., Giorgi, D., Landi, C., 2008. Multidimensional size functions for shape comparison. J. Math. Imaging Vis. 32 (2), 161–179.

Cagliari, F., Di Fabio, B., Ferri, M., 2010. One-dimensional reduction of multidimensional persistent homology. Proc. Am. Math. Soc. 138 (8), 3003–3017.

Carlsson, G., Singh, G., Zomorodian, A., 2009. Computing multidimensional persistence. In: ISAAC '09: Proceedings of the 20th International Symposium on Algorithms and Computation. Springer-Verlag, Berlin, Heidelberg, pp. 730–739.
Carlsson, G., Zomorodian, A., 2007. The theory of multidimensional persistence. In: SCG '07: Proceedings of the 23rd Annual Symposium on Computational Geometry. ACM, New York, NY, USA, pp. 184–193.
Carlsson, G., Zomorodian, A., Collins, A., Guibas, L.J., 2005. Persistence barcodes for shapes. Int. J. Shape Model. 11 (2), 149–187.
Cavazza, N., Ethier, M., Frosini, P., Kaczynski, T., Landi, C., 2013. Comparison of persistent homologies for vector functions: from continuous to discrete and back. Comput. Math. Appl. 66, 560–573.
Cerri, A., Di Fabio, B., Ferri, M., Frosini, P., Landi, C., 2013. Betti numbers in multidimensional persistent homology are stable functions. Math. Methods Appl. Sci. 36, 1543–1557.
Cerri, A., Frosini, P., Kropatsch, W.G., Landi, C., 2011. A global method for reducing multidimensional size graphs. In: Graph-Based Representations in Pattern Recognition. In: Lect. Notes Comput. Sci., vol. 6658, pp. 1–11.
Delgado-Friedrichs, O., Robins, V., Sheppard, A., 2015. Skeletonization and partitioning of digital images using discrete Morse theory. IEEE Trans. Pattern Anal. Mach. Intell. 37 (3), 654–666.
Di Fabio, B., Landi, C., 2012. Persistent homology and partial similarity of shapes. Pattern Recognit. Lett. 33, 1445–1450.
Dłotko, P., Wagner, H., 2014. Simplification of complexes for persistent homology computations. Homol. Homotopy Appl. 16 (1), 49–63.
Edelsbrunner, H., Letscher, D., Zomorodian, A., 2002. Topological persistence and simplification. Discrete Comput. Geom. 28 (4), 511–533.
Forman, N., 1998. Morse theory for cell complexes. Adv. Math. 134, 90–145.
Forman, R., 2002. A user's guide to discrete Morse theory. Sémin. Lothar. Comb. 48.
Frosini, P., 1991. Measuring shapes by size functions. In: Proc. of SPIE, Intelligent Robots and Computer Vision X: Algorithms and Techniques, vol. 1607. SPIE, Boston, MA, USA, pp. 122–133.
Joswig, M., Pfetsch, M., 2006. Computing optimal Morse matchings. SIAM J. Discrete Math. 20 (1), 11–25.
Kaczynski, T., Mischaikow, K., Mrozek, M., 2004. Computational Homology. Appl. Math. Sci., vol. 157. Springer-Verlag, New York.
Kaczynski, T., Mrozek, M., Slusarek, M., 1998. Homology computation by reduction of chain complexes. Comput. Math. Appl. 35 (4), 59–70.
Kahn, A.B., 1962. Topological sorting of large networks. Commun. ACM 5 (11), 558–562.
King, H., Knudson, K., Mramor, N., 2005. Generating discrete Morse functions from point data. Exp. Math. 14 (4), 435–444.
Lefschetz, S., 1942. Algebraic Topology. Colloq. Publ., vol. 27. Amer. Math. Soc., Providence.
Mischaikow, K., Nanda, V., 2013. Morse theory for filtrations and efficient computation of persistent homology. Discrete Comput. Geom. 50 (2), 330–353.
Mrozek, M., Batko, B., 2009. Coreduction homology algorithm. Discrete Comput. Geom. 41, 96–118.
Robins, V., Wood, P.J., Sheppard, A.P., 2011. Theory and algorithms for constructing discrete Morse complexes from grayscale digital images. IEEE Trans. Pattern Anal. Mach. Intell. 33 (8), 1646–1658.
The-GTS-Library, 2016.03.10. The GNU Triangulated Surface Library. http://gts.sourceforge.net/samples.html.
Verri, A., Uras, C., Frosini, P., Ferri, M., 1993. On the use of size functions for shape analysis. Biol. Cybern. 70 (2), 99–107.
Wikipedia, 2014. Algorithms and Data Structures, Part 7: Trees and Graphs. In: Creutzburg, Reiner, Knackmuß, Jenny (Eds.). Wikipedia Book. 160 pages.
Zomorodian, A., Carlsson, G., 2005. Computing persistent homology. Discrete Comput. Geom. 33 (2), 249–274.