Persistence-Sensitive Simplification of Functions on 2-Manifolds [*]

Herbert Edelsbrunner
Dept Computer Science,
Duke University, Durham
Geomagic, RTP
North Carolina, USA
edels@cs.duke.edu

Dmitriy Morozov
Dept Computer Science,
Duke University, Durham
North Carolina, USA
morozov@cs.duke.edu

Valerio Pascucci
Center for Appl. Sci. Comput.,
Lawrence Livermore Natl. Lab.
Livermore, California, USA
pascucci@llnl.gov

ABSTRACT

We continue the study of topological persistence [5] by investigating the problem of simplifying a function f in a way that removes topological noise as determined by its persistence diagram [2]. To state our results, we call a function g an ϵ-*simplification* of another function f if $\|f - g\|_\infty \leq \epsilon$, and the persistence diagrams of g are the same as those of f except all points within L_1-distance at most ϵ from the diagonal have been removed. We prove that for functions f on a 2-manifold such ϵ-simplification exists, and we give an algorithm to construct them in the piecewise linear case.

Categories and Subject Descriptors

F.2.2 [**Analysis of Algorithms and Problem Complexity**]: Non-numerical Algorithms and Problems—*Geometrical problems and computations, Computations on discrete structures*; G.2.1 [**Discrete Mathematics**]: Combinatorics—*Counting problems*

General Terms

Algorithms, Theory

1. INTRODUCTION

In this section, we briefly motivate the problem studied in this paper, review prior related work, and formally state our results.

Motivation. Scientists generate large quantities of continuous data, such as electron densities, temperature distributions. Topological analysis can be used to make sense of such data, to detect interesting features and to observe patterns that cannot be seen in the raw. Regardless of how the data is obtained, whether it is observed in experiments or computed in simulations, data is unfortunately always burdened with noise. While the source of the noise may range from purely physical such as imprecise measurements to purely computational such as the choice of a triangulation, the difficulties it creates always remain. In this paper we consider the problem of ridding the data of that noise by simplifying the function it defines.

It is important to note that whether something is noise or a feature is in the eyes of the beholder. We endorse the idea of Cohen-Steiner, Edelsbrunner, and Harer [2] that the importance of a feature can be quantified by the amount of change necessary to eliminate it. We therefore study the question of how one would eliminate a feature in order to both understand what parts of the domain it occupies, and what the function looks like without it.

Results and prior work. We build on the work of Edelsbrunner, Letscher, and Zomorodian who introduced the concept of topological persistence [5]. Applying this concept to continuous functions, we view the resulting sequence of persistence diagrams as a characterization in which each point represents a topological feature. The importance of a feature is quantified by the distance of this point from the diagonal. Points closer to the diagonal are deemed less important than others and may be interpreted as representing noise. This interpretation is in part justified by the stability of the representation [2]. To state our results, we first introduce the central concept of this paper. Let $\mathbb{X}$ be a topological space, $f : \mathbb{X} \to \mathbb{R}$ a continuous function, $\mathrm{D}_p(f)$ its dimension p persistence diagram, and ϵ a positive constant.

DEFINITION. A *dimension p ϵ-simplification* of f is a function $g : \mathbb{X} \to \mathbb{R}$ such that $\|f - g\|_\infty \leq \epsilon$ and all persistence diagrams of g are the same as those of f except for $\mathrm{D}_p(g)$ which is the same as $\mathrm{D}_p(f)$ but with all off-diagonal points at L_1-distances at most ϵ from the diagonal removed.

See Figure 1 for an illustration. Once we know that ϵ-simplifications exist for all dimensions, we can iterate the construction and erase the points close to the diagonal in all persistence diagrams. We refer to the resulting function as an ϵ-*simplification* of f. In this paper, we consider the problem of finding ϵ-simplifications of a function f, either restricted to a single dimension or iterated across all dimensions. Our main result is a constructive proof that for 2-manifolds such simplifications exist.

SIMPLIFICATION THEOREM FOR 2-MANIFOLDS.

A. Given a 2-manifold $\mathbb{M}$, a function $f : \mathbb{M} \to \mathbb{R}$, a constant $\epsilon > 0$, and a dimension $p = 0, 1$, there exists a dimension p ϵ-simplification $g : \mathbb{M} \to \mathbb{R}$.

B. For $p = 0, 1$ and all $\epsilon > \delta > 0$ there exists a 2-manifold $\mathbb{M}$ and a function $f : \mathbb{M} \to \mathbb{R}$ such that if $g : \mathbb{M} \to \mathbb{R}$ is a dimension p ϵ-simplification of f then $\|f - g\|_\infty > \epsilon - \delta$.
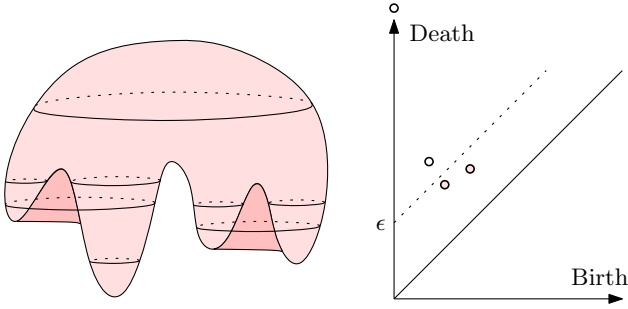
Figure 1: Left: Two embeddings of a 2-manifold $\mathbb{M}$ in $\mathbb{R}^3$. The functions $f, g : \mathbb{M} \to \mathbb{R}$ are the height functions of the light shaded and the combined light and dark shaded embeddings. Right: The dimension 0 persistence diagrams of f and g. The two points below the threshold distance ϵ from the diagonal are present in the persistence diagram of f but not in that of its ϵ-simplification g. The other two points appear in both diagrams.

The problem of simplifying continuous functions has been studied before, in many different areas and from many different angles. The work related most directly to ours is on the simplification of Morse-Smale complexes initiated in [4]. Such complexes capture information about the gradient vector field by partitioning the domain into regions of uniform flow. While the simplification algorithms given in [1, 4, 6] follow the persistence order, they only simplify the complex and not the function itself. The use of the simplified complex together with the original data may be tolerable for visualization purposes, but it is not satisfactory when the simplified data is used in the subsequent data analysis stage. It is worth noting that the example used for the proof of part B of the Simplification Theorem shows that the error bound of $\epsilon/2$ for simplification of a single pair of critical vertices claimed by Bremer et al. [1] is in general unachievable.

In their original paper [5], Edelsbrunner et al. also consider the question of topological simplification. However, there exist significant differences between their work and the results presented in this paper. The most obvious distinction comes from the problem statement itself. Edelsbrunner et al. propose to move all the points of the persistence diagram towards the diagonal regardless of their persistence; in this paper, we require points of persistence higher than ϵ to remain in place. In addition, we make explicit guarantees about the distance between the simplified and the original functions.

Outline. Section 2 reviews background material necessary for this paper. Section 3 gives an overview of our approach to constructing ϵ-simplifications. Section 4 presents the details of our main result, a procedure for simplifying a function in accordance with its persistence diagram. Section 5 exhibits functions for which the error bounds that we achieve are optimal. Section 6 concludes the paper.

2. BACKGROUND

We briefly review simplicial complexes and homology groups. We refer the reader to Hatcher [7] or Munkres [9] for a thorough study of these subjects. We also review the concept of topological persistence [2, 5, 11], restricting ourselves to modulo 2 arithmetic.

Complexes and homology. A *p-simplex* is the convex hull of $p+1$ affinely independent points. The convex hull of any subset of those points is again a simplex, and is called a *face* of the p-simplex. If τ is a face of σ, then σ is a *coface* of τ. A *simplicial complex* is the collection of faces of a finite number of simplices, any two of which are either disjoint or meet in a common face. If K is a simplicial complex in $\mathbb{R}^d$, then its *underlying space* is the union of its simplices together with the subspace topology inherited from $\mathbb{R}^d$. For a set of vertices U in K, we define its *star* as the set of simplices that have at least one vertex in U, and its *link* as the set of faces of simplices in the star that do not also belong to the star:

$$\mathrm{St}\, U = \{\sigma \in K \mid \exists u \in U, u \in \sigma\},$$
$$\mathrm{Lk}\, U = \{\tau \in K \mid \tau \subseteq \sigma \in \mathrm{St}\, U, \tau \notin \mathrm{St}\, U\}.$$

We consider a topological space $\mathbb{X}$ and a triangulation K of $\mathbb{X}$, i.e., a simplicial complex whose underlying space is homeomorphic to $\mathbb{X}$. In simplicial homology, a *p-chain* is a formal sum of p-simplices in K. We use modulo 2 arithmetic implying the coefficients in the formal sum are 0 or 1. We can therefore think of the p-chains as subsets of all p-simplices, namely the ones with coefficient 1. Adding chains modulo 2, we obtain the *group of p-chains*, denoted $\mathsf{C}_p(K)$. It is easy to see that $\mathsf{C}_p(K)$ is abelian. The *boundary* of a p-simplex is the set of its $(p-1)$-dimensional faces, and the boundary of a p-chain is the sum of the boundaries of its simplices. Denoting the boundary map by ∂_p, we observe that it is a homomorphism from $\mathsf{C}_p(K)$ to $\mathsf{C}_{p-1}(K)$. Noting that $\partial_p \partial_{p+1} = 0$, we take the sequence of groups together with the homomorphisms to obtain a *chain complex*,

$$\ldots \overset{\partial_{p+2}}{\to} \mathsf{C}_{p+1} \overset{\partial_{p+1}}{\to} \mathsf{C}_p \overset{\partial_p}{\to} \mathsf{C}_{p-1} \overset{\partial_{p-1}}{\to} \ldots$$

The *group of p-cycles* is the kernel of the p-th boundary homomorphism, $\mathsf{Z}_p(K) = \ker(\partial_p)$, and the *group of p-boundaries* is the image of the $(p+1)$-st boundary homomorphism, $\mathsf{B}_p(K) = \mathrm{im}(\partial_{p+1})$. Since $\partial_p \partial_{p+1} = 0$, $\mathsf{B}_p(K)$ is a subgroup of $\mathsf{Z}_p(K)$. The *p-th homology group* of K is the quotient of the two, $\mathsf{H}_p(K) = \mathsf{Z}_p(K)/\mathsf{B}_p(K)$. The *p-th Betti number* of K is the rank of its p-th homology group, $\beta_p(K) = \mathrm{rank}\, \mathsf{H}_p(K)$. Homology groups and therefore Betti numbers are invariants of the topological space $\mathbb{X}$, and do not depend on the choice of the triangulation K [7, 9].

A topological space is *contractible* if it is homotopy equivalent to a point. In this case, all Betti numbers vanish, except for β_0 which is 1.

Piecewise-linear framework. In this paper we consider real-valued, continuous functions $f : \mathbb{M} \to \mathbb{R}$ defined on a 2-manifold. More specifically, we restrict our attention to functions defined on the vertices of a triangulation K of $\mathbb{M}$ and interpolated linearly on all edges and triangles. Such functions are common in practice (when the underlying space is sampled at discrete points), and are of interest in scientific visualization. We assume that f is non-degenerate, i.e., the function values are different at all vertices. Using these function values, we refine the notions of star and link. Specifically, the *lower star* of a vertex u is the set of simplices in the star for which u has the maximum value of any vertex. The *lower link* of u is the set of faces of simplices in the lower star that do not also belong to the lower star:

$$\mathrm{St}_- u = \{\sigma \in \mathrm{St}\, u \mid v \in \sigma \Rightarrow f(v) \le f(u)\},$$
$$\mathrm{Lk}_- u = \{\tau \in \mathrm{Lk}\, u \mid v \in \tau \Rightarrow f(v) < f(u)\}.$$

Upper stars and *upper links* are defined symmetrically. Similar to the star, we extend the concept of lower star to a set of vertices, U, by taking the union of the individual lower stars: $\mathrm{St}_- U = \bigcup_{u \in U} \mathrm{St}_- u$. Observe that if f is non-degenerate, the lower and upper stars and links of a vertex do not depend on the function values but only on their ordering by function value.

We characterize all vertices by the Betti numbers of their lower links. Since K is a triangulation of a 2-manifold, the link of a vertex is a topological circle (a 1-sphere) and only β_0 and β_1 of the lower link can be non-zero. We call a vertex u a

$$\left.\begin{array}{r} \textit{regular point} \\ \textit{minimum} \\ \textit{saddle} \\ \textit{maximum} \end{array}\right\} \text{ if } \left\{\begin{array}{l} \beta_0 = 1 \text{ and } \beta_1 = 0, \\ \beta_0 = 0 \text{ and } \beta_1 = 0, \\ \beta_0 > 1 \text{ and } \beta_1 = 0, \\ \beta_0 = 1 \text{ and } \beta_1 = 1. \end{array}\right.$$

A saddle is *simple* if β_0 of its lower link is 2, otherwise, it is a *multi-saddle*. In the discussion below we assume that all the saddles in K are simple, since we can unfold all the multi-saddles following the procedure described in [4]. This assumption is not necessary but simplifies the exposition of the algorithm described in this paper. A vertex is a *critical point* unless it is a regular point, and we assign to it an *index* which is 0 for a minimum, 1 for a simple saddle, and 2 for a maximum.

Persistence. Let $\sigma_1, \sigma_2, \ldots, \sigma_N$ be a sequence of the simplices in K. Writing $K_i = \{\sigma_j \mid j \leq i\}$, we call the sequence $\emptyset = K_0 \subset K_1 \subset K_2 \subset \ldots \subset K_N = K$ a *filtration* of K if all K_i are complexes or, equivalently, the faces of every simplex precede the simplex in the given sequence. For a sequence $v_1, v_2, \ldots, v_n$ of the vertices in K, we can construct a sequence of the simplices by listing all lower stars in order and sorting the simplices within each lower star in the order of non-decreasing dimension. If the vertices are sorted in order of increasing function value, we call the resulting sequence of complexes the *lower-star filtration* of the function. From here on, all filtrations will be lower-star filtrations of functions defined at the vertices.

For $1 \leq i < j \leq N$, consider the homomorphisms α, β, and γ implied by the inclusions $K_{i-1} \subset K_i \subset K_{j-1} \subset K_j$:

$$\mathsf{H}_p(K_{i-1}) \xrightarrow{\alpha} \mathsf{H}_p(K_i) \xrightarrow{\beta} \mathsf{H}_p(K_{j-1}) \xrightarrow{\gamma} \mathsf{H}_p(K_j).$$

We say that a homology class $\lambda \in \mathsf{H}_p(K_i)$ is *born* in K_i if $0 \neq \lambda \notin \operatorname{im}(\alpha)$. If λ is born in K_i, we say that it *dies* entering K_j if $\beta(\lambda) \notin \operatorname{im}(\beta\alpha)$ and $\gamma\beta(\lambda) \in \operatorname{im}(\gamma\beta\alpha)$. Observe that since K_i and K_{i+1} differ by only one simplex, at most one homology class is born or dies at any step in the filtration. If there is a $\lambda \in \mathsf{H}_p(K_i)$ that is born in K_i, we call σ_i *positive*. Similarly, if a homology class dies entering K_j, we call σ_j *negative*. If there exists a homology class λ that is born in K_i and dies entering K_j, we pair simplices σ_i and σ_j and call (σ_i, σ_j) a *persistence pair*. It is easy to see that if σ_i is p-dimensional then σ_j is $(p+1)$-dimensional. Edelsbrunner, Letscher, and Zomorodian [5] give an algorithm for computing this persistence pairing in worst-case time cubic in the number of simplices in the given sequence.

There is a close relation between the pairing of simplices and the indices of critical points. To describe this, we consider a simplex σ in the lower star of a vertex s and a simplex τ in the lower star of a vertex t. Assuming (σ, τ) is a persistence pair we say σ and τ are *locally paired* if $s = t$ and they are *non-locally paired* if $s \neq t$. It is not difficult to prove the following observation.

DIMENSION-INDEX LEMMA. A vertex v is regular iff all simplices in its lower star are locally paired. Otherwise, it is critical with index equal to the dimension of the non-locally paired simplex in its lower star.

For every persistence pair of simplices, (σ, τ), we have the corresponding persistence pair of critical points, (s, t). We call the latter *improper* if $s = t$ and *proper* if $s \neq t$. We record information about all proper pairs by drawing the points $(f(s), f(t))$ in the plane. In addition, for each unpaired simplex we draw the point $(f(s), \infty)$,

and following [2] we draw all diagonal points, each infinitely often. By separating the points in which s has index 0 from those in which it has index 1 we get two multisets of points in the extended plane, $\bar{\mathbb{R}}^2$, which we refer to as the *dimension* 0 and the *dimension* 1 *persistence diagrams*, $\mathrm{D}_0(f)$ and $\mathrm{D}_1(f)$.

Stability and transpositions. Cohen-Steiner, Edelsbrunner, and Harer [2] proved a stability result for persistence diagrams. The following is its restriction to 2-manifolds. Given two functions $f, g : \mathbb{M} \to \mathbb{R}$, as above, we define the distance between them to be the L_∞-norm of their difference: $\|f - g\|_\infty = \sup_{x \in \mathbb{M}} |f(x) - g(x)|$. The *bottleneck distance* between the persistence diagrams of f and g is the infimum over all bijections $\gamma : \mathrm{D}_p(f) \to \mathrm{D}_p(g)$ of the supremum distance between corresponding points:

$$d_B(\mathrm{D}_p(f), \mathrm{D}_p(g)) = \inf_{\gamma} \sup_{u \in \mathrm{D}_p(f)} \|u - \gamma(u)\|_\infty.$$

For technical reasons the functions are required to be *tame*, by which we mean they have only finitely many critical values and any sublevel set has only finite Betti numbers.

STABILITY THEOREM. If $f, g : \mathbb{M} \to \mathbb{R}$ are two continuous, tame functions then for any $p \geq 0$, the bottleneck distance between their dimension p persistence diagrams is not greater than the distance between the functions: $d_B(\mathrm{D}_p(f), \mathrm{D}_p(g)) \leq \|f - g\|_\infty$.

Suppose that we continuously change the function values at the vertices. As a result the points in the persistence diagram move, but not more then the amount of change of the values. Even though the motion is therefore continuous, the pairs defining the points in the diagram can *switch* vertices, but only at moments in time when these vertices have the same value.

SWITCH LEMMA. Transposing two consecutive vertices v_i and v_{i+1} in the ordering defining the lower star-filtration can only affect the persistence pairs containing v_i and v_{i+1}.

Cohen-Steiner, Edelsbrunner, and Morozov [3] give an algorithm to maintain the pairing if two adjacent simplices are transposed and the new sequence of complexes remains a filtration. In the following sections, we will be transposing adjacent vertices v_i, v_{i+1} in the vertex ordering. The corresponding change in the lower-star filtration is obtained by transposing the lower stars of v_i and v_{i+1}, which reduces to a number of simplex transpositions. We get a first constraint on switches between persistence pairs by observing that the indices in each pair are contiguous and increasing.

SAME INDEX LEMMA. Transpositions between critical vertices with different indices preserve the persistence pairing.

A crucial second constraint on how switches between pairs can happen follows from the analysis in [3]. To describe it, we call two pairs of critical points, (v_i, v_j) and (v_k, v_l), *nested* if $i < k < l < j$ and *disjoint* if $i < j < k < l$. To use these notions for unpaired vertices, we consider them artificially paired with a dummy vertex with subscript equal to infinity and we permit equality when we compare subscripts that are infinite.

NESTED-DISJOINT LEMMA. During a transposition of two consecutive vertices, the pairs can switch these vertices iff the pairs are nested or disjoint both before and after the transposition.

This lemma in particular implies that if before the transposition there exist k and l with $k < i < i+1 < l$ such that v_k is paired with v_{i+1} and v_i is paired with v_l, then after v_i and v_{i+1} are transposed we still have the same two pairs.

3. OVERVIEW

In this section, we give a high-level view of our approach to finding an ϵ-simplification and present the necessary structural lemmas. We leave the details of the algorithm to the next section.

Basic strategy. Simplifications of a function are generated by cancelling critical points in pairs, minima with saddles and saddles with maxima. In order to cancel a pair, one's initial inclination may be to change the values of both critical points, i.e., lower the saddle and raise the minimum for a minimum-saddle pair, and raise the saddle and lower the maximum for a saddle-maximum pair. However, as the example in Section 5 shows, this may not always be possible because extrema can get stuck as they encounter other critical vertices. To avoid this difficulty, we leave the values of extrema unchanged, and move only the saddles. Below we describe the case of lowering a saddle to its matching minimum; the case of raising the saddle to its matching maximum is symmetric.

Let V be the ordering of the vertices by increasing function value, and let (s, t) be a minimum-saddle persistence pair of the lower-star filtration determined by V. To cancel (s, t), we lower a contiguous subsequence of vertices, T, which we imagine as a flat region the saddle drags along while being lowered. Initially, $T = \{t\}$. Since T is contiguous in V, it partitions V into three contiguous subsequences, W, T, U, as illustrated in Figure 2. Let w be the last vertex in W. Lowering T means either moving w past
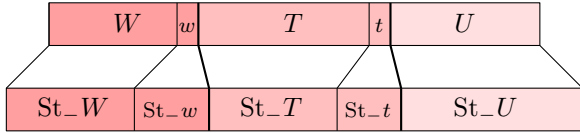

Figure 2: Top: The sequence of vertices is partitioned into W with the last vertex w, T with the last vertex t, and U. Bottom: The sequence of simplices defining the corresponding lower-star filtration.

T (by assigning all vertices in T a value slightly less than $f(w)$), or expanding T to include w (by setting the values of all vertices of T equal to $f(w)$). The former approach is preferable, and we use it when w is not in the link of T. A difficulty arises when w is in the link of T since moving T below w may change the type from regular to critical or vice versa or turn w into a multi-saddle if it is already a saddle. In this case, we expand T which preserves the type of the vertex. However, if w is a critical point with persistence higher than (s, t) then we cannot afford to move the corresponding point in the persistence diagram. We thus need to maintain a critical point with the same value and cannot immediately include w into T. This requirement dictates two properties we maintain as invariants, namely that t be the only critical vertex in T and that the star of T be contractible.

Encountering a minimum. If w is a minimum and belongs to the link of T then the following lemma tells us that w is paired with t, i.e., w is equal to s.

PAIRED MINIMUM LEMMA. If w is a minimum that immediately precedes T in V, the star of T is contractible, and the only critical vertex in T is a negative saddle t, then $w = s$ iff w belongs to the link of T.

PROOF. If $w = s$ then the definition of persistence pairs implies a path starting at w whose edges belong to the lower star of T. In particular the first edge connects w to a vertex in T implying that w is in the lower link of T, as required.

To prove the reverse direction, assume w belongs to the link of T. Starting with St_-W we proceed along the lower-star filtration by adding the simplices in St_-T until we arrive at the lower star of $W \cup T$. In St_-W, w forms its own component, and in $\mathrm{St}_-(W \cup T)$, w belongs to a component that contains all of St_-T. The latter component cannot just grow from w, by adding lower stars of regular vertices to it, because there is one negative saddle, t, and adding its lower star merges two components. One of these components contains w and the other was started by another, older vertex in W. Hence (w, t) is a persistence pair and $w = s$, as required. $\square$

Once we reach s, we add it to T and reorder the vertices in T so that all of them become regular, including s and t.

Encountering a saddle. If w is in the link of T, it cannot be a maximum, therefore the only remaining case is a saddle. To cope with this case, we subdivide some of the edges in the lower star of w in a way so that w is no longer a saddle, and the new saddle that replaces it is no longer in the link of T. To perform such a subdivision, we need vertices in the link of w that are in U. If there are no such vertices, we build a tunnel to U between the lower star of T and the lower star of W. To guarantee that this is possible, we use again the invariant that guarantees that the lower star of T is contractible. This ensures that the link of T is connected and we can travel to a vertex in U by following this link.

The following structural lemma will play a crucial role in the analysis presented in the next section. It assumes a partition of V into W, T, U and writes w for the last vertex of W, as usual. In a nutshell, the lemma says that unlike suggested by Figure 3, it is not possible to draw a path through w that enters the lower star of U as we move from w in both directions and which locally separates the vertices of T in the link of w.

NON-SEPARATION LEMMA. Suppose $\mathrm{St}_-T \cap \mathrm{Lk}\, w \neq \emptyset$, the star of T is contractible, and the only critical vertex in T is a negative saddle. Then Lk_-w merges all pieces of $\mathrm{St}_-T \cap \mathrm{Lk}\, w$ into a single component.

PROOF. Label the vertices in the link of w that belong to T in a counter-clockwise order around w as $t_1, t_2, \ldots, t_m$. To get a contradiction, we suppose there are two vertices, t_i and t_j, that are not in the same component of $(\mathrm{St}_-T \cap \mathrm{Lk}\, w) \cup \mathrm{Lk}_-w$. In other words, t_i and t_j are locally separated by a path that passes through w, connecting it on both sides to vertices in U, as in Figure 3.

Since the star of T is contractible, there exists a path that connects t_i and t_j entirely within the lower star of T. Adding $t_i w$ and $w t_j$ to the path forms a cycle. Since w is incident to triangles in the lower star of U on both sides, this cycle does not bound a 2-chain inside the lower star of $W \cup T$. Indeed, suppose that it does. Then that 2-chain must contain either $w t_i t_{i+1}$, or $w t_i t_{i-1}$ since they are the only two triangles in the lower star of $W \cup T$ that contain $w t_i$. Assuming it contains $w t_i t_{i+1}$ and noting that the cycle does not contain $w t_{i+1}$, the 2-chain also contains $w t_{i+1} t_{i+2}$. Continuing this way, the 2-chain must contain a triangle $w t_{i+k} u$, with $u \in U$, which is impossible since $w t_{i+k} u$ is not in the lower star of $W \cup T$.

Therefore, the cycle does not bound a 2-chain in the lower star of $W \cup T$. But this implies that adding the lower star of T to the lower star of W creates a non-zero class in the first homology group. It follows that there is a positive saddle in T, contradicting the supposition that the only critical vertex in T is a negative saddle. $\square$
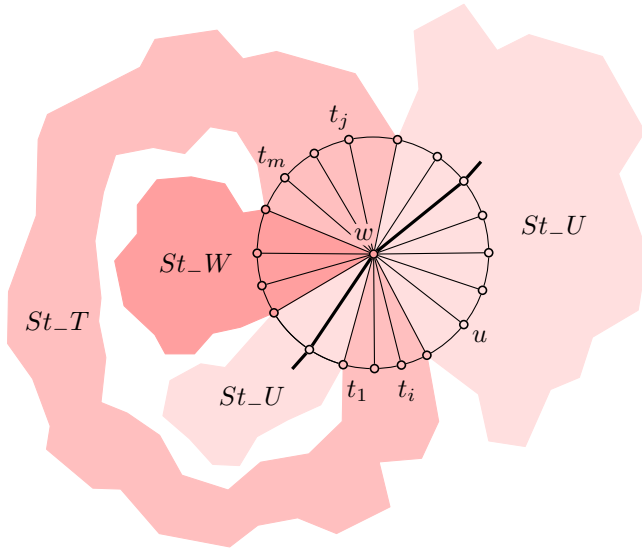
Figure 3: The addition of the lower star of T to the lower star of W creates a non-bounding cycle, which implies that T contains a positive saddle. The partially indicated path passing through w locally separates t_i and t_j.

Order of cancellations. The only remaining question is the order in which we consider the pairs of critical points. Most natural would be the order of increasing persistence. Unfortunately, with our technique such an order cannot guarantee that f is changed by at most ϵ. If two pairs overlap, canceling the one with higher persistence may drag a vertex that has already been lowered during the cancellation of a pair with lower persistence. Therefore, the change in the function values may compound. Instead, we consider critical point pairs (s, t) in the order of increasing values of t, i.e., we sweep the vertices from bottom to top and lower the saddles that belong to a pair of persistence less than ϵ. If a vertex was lowered during the cancellation of one pair and then again during the cancellation of another pair, then our technique guarantees that the first pair was nested in the second. Therefore, the total change in the function value of this vertex does not exceed ϵ. This implies $\|f - g\|_\infty \le \epsilon$, as required.

Naturally, if we cancel minimum-saddle and saddle-maximum pairs, we sweep the vertices twice, from bottom to top and from top to bottom. The change in function values is still bounded by ϵ since the former pass lowers vertices by at most ϵ and the latter pass raises them by at most ϵ.

4. SIMPLIFICATION DETAILS

As described in the previous section, we cancel pairs in the order of increasing values of the second vertex. To cancel a pair (s, t), we lower a collection of vertices, initializing it to $T = \{t\}$.

Case analysis. The algorithm proceeds by lowering or expanding T one vertex at a time. To guarantee progress at each step, we maintain three properties as invariants throughout the algorithm.

INVARIANT.

  I. T contains only one critical vertex, namely t;

  II. the star of T is contractible;

  III. T is a contiguous sequence in V.

Since we only consider cancellations of minimum-saddle pairs, the one critical point t in T can be assumed to be a negative saddle. Invariants I, II, III are trivially true when $T = \{t\}$. As in the previous section, since T is consecutive in V, it partitions V into three contiguous subsequences, W, T, U. Let w be the last vertex of W. We distinguish two cases, each with two subcases.

Case I. The vertex w does not belong to the link of T. Note that in this case exchanging w and T in V changes neither the lower star of w nor that of T.

Case I.1. w is regular. After exchanging w and T in V, w remains regular. The Stability Theorem thus implies that there are no changes in the pairing of the critical points.

Case I.2. w is critical. The Paired Minimum Lemma implies $w \ne s$. If w is a minimum or a maximum then the Same Index Lemma implies that exchanging w and T does not affect the pairing. If w is a positive saddle then it is either unpaired or paired with a maximum. In the first case, we consider it paired with a dummy vertex that succeeds all other vertices. In either case, the two pairs that contain w and t are neither nested nor disjoint. If w is a negative saddle, the fact it has not yet been cancelled implies its persistence exceeds ϵ. Since w precedes t, the pairs that contain w and t are neither nested nor disjoint. The Nested-Disjoint Lemma thus contradicts any switch in the pairing. In conclusion, exchanging w and T in V does not affect the pairing, as desired.

Case II. The vertex w belongs to the link of T.

Case II.1. w is regular. If $St_-T \cap Lk\, w$ is contractible, we add w to T by prepending it on the left. Then t is still the only critical vertex in T, the star of T is still contractible because $St\, (T - \{w\})$ is a deformation retract of $St\, T$, and T is still contiguous in V. In summary, Invariants I, II, III are preserved.

The situation is more complicated if $St_-T \cap Lk\, w$ is not contractible. By the Non-separation Lemma, the union of $St_-T \cap Lk\, w$ and Lk_-w is contractible, and by the regularity of w, $St_-T \cap Lk\, w$ consists of two components, as illustrated in Figure 4. Picking one of these components, we subdivide each edge connecting it to w with two new vertices. The value of the vertex closer to w is chosen above T but below U, and the value of the vertex further from w is chosen below T but above w. Within this range, we choose the values such that we get two monotonically increasing paths from Lk_-w to U, one passing through the new vertices above T and the other passing through the new vertices below T, as shown in Figure 4. Observe that all new vertices are regular. Indeed, each new vertex above T has a single vertex upper link and each new vertex below T has a single vertex lower link. The type of every other vertex remains unchanged as increasing edges in its star are replaced by increasing edges and decreasing edges are replaced by decreasing edges. After subdivision, we add the new vertices below T to T, observing that Invariants I, II, III are preserved. But now we are back in the case in which $St_-T \cap Lk\, w$ is contractible, so we can add w to T, as discussed earlier.

Case II.2. w is critical. It cannot be a maximum else its upper link would be empty and it could not be in the link of T.

Consider first the case in which w is a saddle. By assumption, all saddles are simple which implies that the lower link of w consists of two components and so does the upper link.
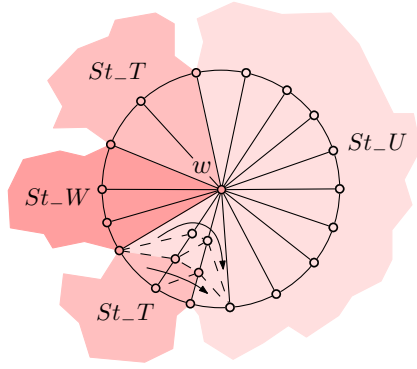
Figure 4: The vertex w is regular. Before the subdivision, $St_-T \cap Lk\, w$ is not contractible while $(St_-T \cap Lk\, w) \cup Lk_- w$ is. By subdividing, we reduce the problem to the case in which $St_-T \cap Lk\, w$ is contractible. The new edges are dashed and the lower stars after subdivision are indicated by the shading. The arrows indicate the direction in which the values of the new vertices increase.

Figure 5: The vertex w is a saddle. The lower star of U sandwiches a component of the lower star of w. By subdividing its edges, we turn w into a regular vertex, replacing it by the new saddle x. The shading shows the lower stars after the subdivision. The arrows indicate the direction in which the values of the new vertices increase.

The vertices in the upper link belong to T and to U, and here we consider the easy case in which there are vertices of U in both components of the upper link. We will show how to reduce the other case to this one shortly. By the Non-separation Lemma, the portions of the upper link that contain vertices of T sandwich one component of the lower link of w, and the portions that contain vertices of U sandwich the other component, as illustrated in Figure 5. We subdivide by placing a single vertex on each edge connecting w with the latter component of the lower link. The values of the new vertices are chosen above T and below U. Within this range we choose the values such that the path of new vertices first decreases, attains its minimum at a vertex x, and then increases, as in Figure 5. All new vertices are regular, except for x, which is a saddle. With these changes, w is regular and all other vertices retain their original type.

We argue that x replaces w in the pairing using a continuity argument. To start, we assign to each new vertex the value of the point where it is placed. At this time, all new vertices are regular and do not belong to any proper persistence pair. Next, we continuously change the values of the new vertices, updating the sequence through transpositions as we go. This is done so that all new vertices remain regular at all times, except for x which makes a crucial transposition with w in which x becomes a saddle and w a regular vertex. The Switch Lemma implies that x replaces w in its pair, as desired.

We finally continue the simplification process by moving x and the other vertices past T. Because of the subdivision, we are now in Case I implying that these transpositions do not affect the pairing. The vertex w is now regular, so we can add it to T as described in Case II.1.

Consider second the case in which w is a minimum. By the Paired Minimum Lemma, we have $w = s$. We add s to T and reorder T to make all its vertices regular, as described below. The reordering finally cancels the pair (s, t).

To measure progress, we count the vertices in W. Each step shrinks W, therefore the algorithm halts after a finite number of steps.
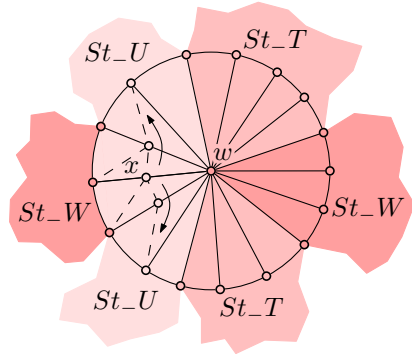
Tunneling. In Case II.2 when w is a saddle, we assumed that there are vertices of U in both components of the upper link of w. Now we describe additional actions that put such vertices in the upper link in case they are missing. A crucial property in this construction is Invariant II which implies that the link of T is connected. This link contains vertices both in W (for example, w) and in U (since the link of $t \in T$ contains vertices in U). This implies that we can walk on this link from w until we encounter a first vertex u in U. Let π be this path, as illustrated in Figure 6. By construction, all vertices in π other than u belong to W. To get a vertex of U into the upper link of w, we subdivide edges that connect interior vertices of π with U. More precisely, we construct a connected strip of triangles incident on π, starting with the triangle that connects the first edge of π with a vertex in T and ending with the triangle that connects the last edge of π with a vertex in T. We subdivide the interior edges in the strip, placing two new vertices on each. The value of the new vertex closer to π is chosen above T but below U and the value of the new vertex further from π is chosen below T but above w. Within these two ranges, we choose the values to get two monotonically increasing paths from w to u. Similar to Case II.1, all new vertices are regular and the types of the other vertices remain unchanged.

The new vertices in the path below T are now added to T. Note that this preserves Invariants I, II, III. The other new vertices belong to U so we succeeded in our goal of putting a vertex of U into the upper link of w. If necessary, we repeat this procedure for the second component of the upper link of w. Finally, we proceed as in Case II.2.

Reordering. We now discuss the last step of the algorithm in more detail, the reordering of the vertices in T. Recall that T is contiguous in V, it starts at s and ends at t, and all vertices in T are regular except for s and t. By removing t we decompose the lower star of T into components, and we let $S \subseteq T$ be the set of vertices in the same component as s. To reorder T, we

Step 1. remove S from V;

Step 2. reverse S;

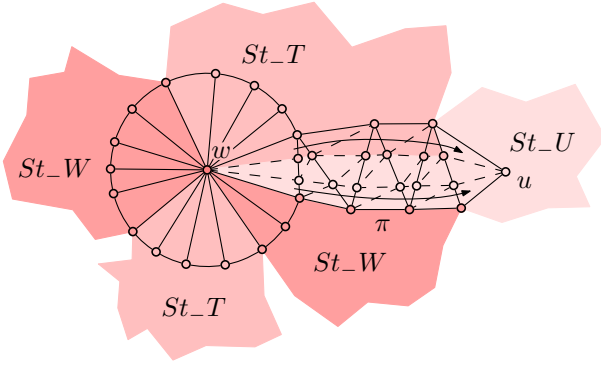**Step 3.** add the reversed sequence S right after t to V.

Figure 6: Connecting w to U in case the upper link of w has only vertices in T. The subdivision creates two monotonically increasing paths of new vertices parallel to the path π in the link of T. The shading shows the lower stars after the subdivision.
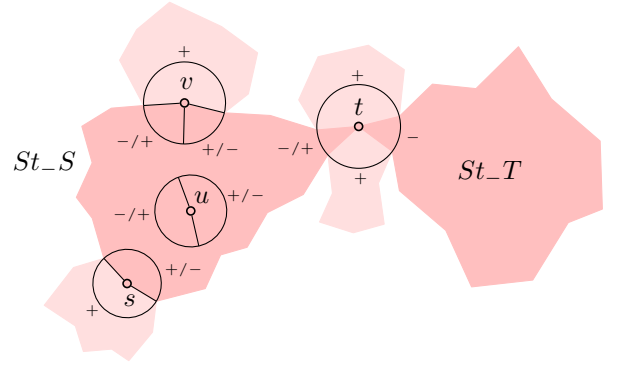

Figure 7: Reordering the vertices in T. Before the reordering, all vertices in T are regular except for s and t. After reordering, all vertices in T are regular. Plus and minus signs distinguish between upper and lower links. The cases in which the reordering swaps upper and lower links are marked by two signs.

The situations before and after the reordering are illustrated in Figure 7. The procedure is straightforward but it takes a bit of effort to show that it is correct. In particular, we prove that after reordering T all vertices in T are regular. This is clear for all vertices different from t that do not belong to S. We distinguish four cases.

Case i. The last vertex, $t \in T$. Before reordering, t is a saddle whose lower link consists of two components, one in the lower star of S. Step 3 effectively raises the vertices in S above t, making one of the components disappear and turning t into a regular vertex.

Case ii. The first vertex, $s \in T$. Since s is a minimum, its lower link is empty. Hence all neighbors of s belong either to T or to U, and the Non-separation Lemma implies that the portion of the upper link of s inside St_-T is connected. This is the same as $\text{Lk}_+ s \cap \text{St}_- S$, and because the star of T is contractible, at least some of the neighbors of s belong to U. This implies that the portion of the upper link defined by vertices in S is contractible. Step 2 turns this portion into the lower link and s into a regular vertex.

Case iii. A vertex $u \in S$ whose link is contained in the lower star of S. The upper link of u becomes its lower link and vice versa, implying that u remains regular.

Case iv. A vertex $v \in S$, different from s, whose lower link is not contained in the lower star of S. We first observe that v has no neighbors in W. To see this, we consider the lower-star filtration defined by V (before the reordering). Starting with the lower star of $W \cup \{s\}$ we add lower stars of vertices in T until we arrive at the lower star of $W \cup T$. For our argument only the vertices in the same component as s are relevant so we consider $\text{St}_-(W \cup S)$. But if v has neighbors in W then this process would have merged the component of s with another component, contradicting the fact that all vertices in $S - \{s\}$ are regular. This shows that all neighbors not in T belong to the upper link of v. The neighbors in T all belong to S and at least some of them are lower than v. These vertices form a contractible lower link of v, else v would not be regular. Similarly, the neighbors in S above v form a contractible portion of the upper link of v, else we would have gotten a contradiction to Invariant II at the time v was added to T. It thus follows that reversing S preserves v as a regular vertex.

Multi-saddles. We note that it is not necessary to unfold all multi-saddles for the algorithm to work. Generally, we distinguish saddles with persistence larger than ϵ, which do not have to be unfolded, and saddles with persistence at most ϵ, which have to be unfolded. However, this is oversimplifying the situation because a multi-saddle can be part of multiple pairs with persistence larger as well as smaller or equal to ϵ. The Nested-Disjoint Lemma implies that a multi-saddle can be unfolded such that the resulting positive simple saddles have higher function value than the resulting negative simple saddles. Similarly, pairs with smaller persistence can be nested within pairs of larger persistence. Finally, the resulting simple saddles with persistence larger than ϵ are assigned the function value of the multi-saddle so that the unfolding does not interfere with bounding the change of the function through simplification.

5. LOWER BOUND

In this section, we prove part B of the Simplification Theorem for 2-Manifolds stated in Section 1: for $p = 0, 1$ and all $\epsilon > \delta > 0$ there exists a 2-manifold $\mathbb{M}$ and a function $f : \mathbb{M} \to \mathbb{R}$ such that if $g : \mathbb{M} \to \mathbb{R}$ is a dimension p ϵ-simplification of f then $\|f - g\|_\infty > \epsilon - \delta$. The topology of the 2-manifold is less important for the proof than the details of the function. We thus let $\mathbb{M}$ be the 2-sphere and we choose f as the (vertical) height function of the embedding of $\mathbb{M}$ displayed in Figure 8. There are three critical points with similar heights, $f(P) = r - \epsilon$, $f(Q) = r - \delta$, $f(R) = r$, where $0 < \delta < \epsilon$. The two minima have function values $f(A) = a < f(B) = b$ that are both much smaller than r, and the maximum has a function value $f(Z) = z$ that is much larger than r. The critical points are paired as $(B, Q), (P, R)$, leaving A and Z unpaired. The off-diagonal points in the persistence diagrams are therefore

$$\begin{aligned}
\text{D}_0(f) &: \quad (a, \infty), (b, r - \delta); \\
\text{D}_1(f) &: \quad (r - \epsilon, r); \\
\text{D}_2(f) &: \quad (z, \infty).
\end{aligned}$$

All points have L_1-distance larger than ϵ from the diagonal, except for $(r - \epsilon, r)$ whose L_1-distance from the diagonal is ϵ. To get a dimension 1 ϵ-simplification, we thus need to cancel P with R and leave the other critical points in tact (or replace them by new critical points at the same height). It seems plausible that f does not have
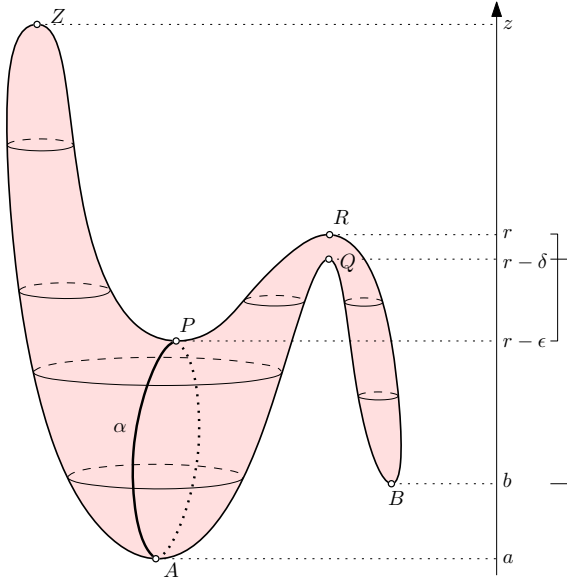
Figure 8: Embedding of the 2-sphere $\mathbb{M}$ in $\mathbb{R}^3$ such that $f : \mathbb{M} \to \mathbb{R}$ is its height function. There are two minima, A and B, two saddles, P and Q, and two maxima, R and Z. The two ascending paths from A to P decompose $\mathbb{M}$ into a left and a right hemisphere.

6. DISCUSSION

The main contribution of this paper is a constructive proof of the existence of ϵ-simplifications for continuous functions on 2-manifolds. The proof extends to 2-manifolds with boundary since we can convert those into 2-manifolds without boundary by gluing a disk to each boundary cycle. A curious aspect of our proof is that dimension 0 and dimension 1 homology can be simplified independently. Indeed, we can cancel all minimum-saddle pairs of persistence at most ϵ while leaving all saddle-maximum pairs intact, or vice versa. It is also worthwhile to mention that the algorithm is combinatorial and we are free to assign function values that are consistent with the computed ordering of the vertices. However, our algorithm is not incremental in the sense of continuously increasing the error threshold and this way generating a hierarchy of simplifications. The main reason for this shortcoming is that the sequence of pairs cancelled by our algorithm is generally not sorted by persistence. We leave the design of such an incremental simplification algorithm as an open question.

The authors consider the simplification of continuous functions as a central problem in visualization. It may be used to clean up Morse-Smale complexes [4] and Reeb graphs [8, 10], which are powerful tools in the study and visualization of continuous data in scientific computing. We therefore believe that the extension of our results to three- and higher-dimensional manifolds as well as to other topological spaces is important.

Acknowledgments

a dimension 1 $(\epsilon - \delta')$-simplification with $\delta < \delta' < \epsilon$. Indeed, we cannot lower R by more than δ since it gets stuck at Q. Hence we need to raise P by at least $\epsilon - \delta$. A more formal argument supporting this conclusion will be presented shortly. Since this works for arbitrarily small $\delta > 0$, this implies the claimed lower bound. To prove the same bound for $p = 0$ we use the construction upside-down, that is, we substitute $-f$ for f.

We now give the more formal argument for the claim that the difference between f and g is $\|f - g\|_\infty > \epsilon - \delta$. To get a contradiction, we assume there is a dimension 1 ϵ-simplification $g : \mathbb{M} \to \mathbb{R}$ of f with $\|f - g\|_\infty = \epsilon - \delta'$ for some $\delta' > \delta$. Let α be the cycle consisting of two monotonically increasing paths from A to P, as drawn in Figure 8. It decomposes the 2-sphere into a closed left hemisphere (containing Z) and a closed right hemisphere (containing B, Q, R). Consider the restrictions $\bar{f}$ and $\bar{g}$ of f and g to the right hemisphere. The diagram $\mathrm{D}_0(\bar{f})$ is the same as $\mathrm{D}_0(f)$. By the Stability Theorem, the diagram $\mathrm{D}_0(\bar{g})$ contains a point (b', q') at L_∞-distance at most ϵ from $(b, r - \delta)$ in $\mathrm{D}_0(\bar{f})$. The value q' is that of a saddle Q' of $\bar{g}$. By definition of ϵ-simplification, we have $g(Q') = q' = r - \delta$, which is larger than $g(x) \le f(x) + (\epsilon - \delta') < r - \delta$ for any point x on α. This implies that Q' lies in the interior of the right hemisphere and is therefore also a saddle of g. Furthermore, there are no other finite off-diagonal points in the persistence diagrams of g. It follows that g has only one saddle, namely Q'. A similar argument implies that g has only one maximum, Z', in the left hemisphere and that $g(Z') = z$. Since there is only one maximum and only one saddle, we can draw a path from Z' to Q' that monotonically decreases in g. This path crosses the cycle α. But the points x on α have $g(x) < r - \delta$ which is less than the values of Z' and Q' at the two ends. This contradicts the monotonicity of the path and implies $\|f - g\|_\infty > \epsilon - \delta$, as required.

7. REFERENCES

[1] PEER-TIMO BREMER, HERBERT EDELSBRUNNER, BERND HAMANN AND VALERIO PASCUCCI. Topological hierarchy for functions on triangulated surfaces. *IEEE Trans. Vis. Comput. Graphics*, 10:385–396, 2004.

[2] DAVID COHEN-STEINER, HERBERT EDELSBRUNNER AND JOHN HARER. Stability of persistence diagrams. In *Proc. 21st Ann. Sympos. Comput. Geom.*, pages 263–271, New York, NY, USA, 2005.

[3] DAVID COHEN-STEINER, HERBERT EDELSBRUNNER AND DMITRIY MOROZOV. Vines and vineyards by updating persistence in linear time. To appear in *Proc. 22st Ann. Sympos. Comput. Geom.*, Sedona, Arizona, USA, 2006.

[4] HERBERT EDELSBRUNNER, JOHN HARER AND AFRA ZOMORODIAN. Hierarchical Morse-Smale complexes for piecewise linear 2-manifolds. *Discrete Comput. Geom.*, 30: 87–107, 2003.

[5] HERBERT EDELSBRUNNER, DAVID LETSCHER AND AFRA ZOMORODIAN. Topological persistence and simplification. *Discrete Comput. Geom.*, 28:511–533, 2002.

[6] ATTILA GYULASSY, VIJAY NATARAJAN, VALERIO PASCUCCI, PEER-TIMO BREMER AND BERND HAMANN. Topology-based simplification for feature extraction from 3D scalar fields. In Proc. IEEE Conf. Visualization, pages 275–280, 2005.

[7] ALLEN HATCHER. *Algebraic Topology*. Cambridge University Press, Cambridge, England, 2001.

[8] MASAKI HILAGA, YOSHIHISA SHINAGAWA, TAKU KOHMURA AND TOSIYASU L. KUNII. Topology matching for fully automatic similarity estimation of 3D shapes. In *Proc. 28th Ann. Conf. on Comput. Graph. and Interact. Techn. (SIGGRAPH)*, pages 203–212, 2001.

[9] JAMES R. MUNKRES. *Elements of Algebraic Topology*. Addison Wesley, Redwood City, California, 1984.

[10] G. REEB. Sur les points singuliers d'une forme de Pfaff complèment intégrable ou d'une fonction numérique. *Comptes Rendus de L'Académie des Séances*, 222:847–849, 1946.

[11] AFRA ZOMORODIAN AND GUNNAR CARLSSON. Computing persistent homology. *Discrete Comput. Geom.*, 33:249–274, 2005.