Optimal Reconstruction Might be Hard*

[Extended Abstract]

Dominique Attali
Gipsa-lab, Grenoble, France
Dominique.Attali@grenoble-inp.fr

André Lieutier
Aix-en-Provence, France
Andre.Lieutier@wanadoo.fr

ABSTRACT

Sampling conditions for recovering the homology of a set using topological persistence are much weaker than sampling conditions required by any known polynomial time algorithm for producing a topologically correct reconstruction. Under the former sampling conditions which we call *weak sampling conditions*, we give an algorithm that outputs a topologically correct reconstruction. Unfortunately, even though the algorithm terminates, its time complexity is unbounded. Motivated by the question of knowing if a polynomial time algorithm for reconstruction exists under the weak sampling conditions, we identify at the heart of our algorithm a test which requires answering the following question: given two 2-dimensional simplicial complexes $L \subset K$, does there exist a simplicial complex containing L and contained in K which realizes the persistent homology of L into K? We call this problem the *homological simplification* of the pair (K, L) and prove that this problem is NP-complete, using a reduction from 3SAT.

Categories and Subject Descriptors

F.2.2 [**Analysis of Algorithms and Problem Complexity**]: Nonnumerical Algorithms and Problems—*Geometrical problems and computations, Computations on discrete structures*; I.3.5 [**Computer Graphics**]: Computational Geometry and Object Modeling

General Terms

Algorithms, Theory

Keywords

Shape reconstruction, topological persistence, sampling conditions, homological simplification, NP-completeness, 3SAT

*This work is partially supported by ANR Project *GIGA* ANR-09-BLAN-0331-01.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
SCG'10, June 13–16, 2010, Snowbird, Utah, USA.
Copyright 2010 ACM 978-1-4503-0016-2/10/06 ...$10.00.

1. INTRODUCTION

Previous works.

In the last decade, several authors have proposed algorithms for reconstructing shapes with topological guarantees. First results considered compact smooth surfaces embedded in the Euclidean three-dimensional space and assumed surfaces to be known through noise-free samples [5, 1, 3, 9, 22, 4]. Since then, an effort has been made to generalize such results to wider classes of shapes and sampling conditions [24, 21, 32, 15, 10, 25].

An important extension allows samples to be noisy: each point of the sample is required to lie within some distance of the sampled shape (the sample is *accurate*) and each point of the sampled shape must lie within some distance of a sample point (the sample is *dense*). When both distances are bounded by the same value ε, the sampling condition can be expressed by saying that the *Hausdorff distance* between the shape and the sample is upper bounded by ε.

In 2006, several algorithms for reconstructing non-smooth objects with topological guarantees have been designed. In [33, 11], Boissonnat and Oudot considered Lipschitz manifolds while Chazal, Cohen-Steiner and Lieutier in [12] considered a large class of non-smooth compact sets called *sets with positive μ-reach*. In particular, the latter gave a polynomial time algorithm that is able to output what we will call in the paper a *faithful reconstruction* (see Definition 1).

In 2002, a fruitful point of view in computational topology, called *topological persistence*, has been introduced [26]. Within this context, the theorem on the stability of persistence diagrams [18] allows to infer the homology of shapes known through a sample. The sampling condition is significantly milder than the one required by the previously mentioned algorithm for computing a faithful reconstruction. We refer to this mild sampling condition sufficient for recovering homology as the *weak sampling condition*. This condition is tight.

Optimal reconstruction.

We call any algorithm that would be able to produce a faithful reconstruction under the weak sampling condition an *optimal reconstruction algorithm*. We explain in Section 2.3 that, even though no realistic version of an optimal reconstruction algorithm is known today, the weak sampling condition ensures that the sample contains in principle enough information on the sampled shape to produce without ambiguity a faithful reconstruction of it. Starting from this observation, we give in Section 2.4 a "naive" algorithm which, at the expense of not being efficient, produces a faithful reconstruction under the weak sampling condition.

The main question we pursue is: can we do better? More precisely, does there exist a polynomial time optimal reconstruction algorithm? Such an algorithm would have applications in several fields such as shape reconstruction or machine learning. While [6] gives a partially positive answer for subsets of $\mathbb{R}^2$, the question remains open for higher dimensions. Indeed, this problem is closely related to the persistence-sensitive simplification of real-valued functions, whose goal is to filter out topological noise in sub-level sets. Indeed, reconstruction can be thought of as the simplification of distance functions to the samples. For functions defined on triangulated 2-manifolds, polynomial algorithms have been devised [27, 6, 30]. Still, persistence-sensitive simplification of functions in higher dimension remains elusive.

Homological simplification.

In Section 3, we focus on the test at the heart of our naive algorithm. This test requires to answer the following question: given two 2-dimensional simplicial complexes $L \subset K$, does there exist a simplicial complex X containing L and contained in K such that the maps induced by the inclusions $L \hookrightarrow X$ and $X \hookrightarrow K$ on all modulo 2 homology groups are respectively surjective and injective. We call this problem the *homological simplification* of the pair (K, L) and prove that it is NP-complete.

Although this result is negative, we believe that it casts new light on the problem of finding a faithful reconstruction under weak sampling conditions and opens further research tracks as mentioned in Section 4. In particular, it suggests that, in order to design efficient reconstruction algorithms under weak sampling conditions, one has to make additional hypotheses. For instance, one can assume shapes are embedded in $\mathbb{R}^3$ and/or require some specific properties on embedded simplicial complexes.

2. THE QUEST FOR AN OPTIMAL RECONSTRUCTION ALGORITHM

Section 2.1 presents the necessary background. Section 2.2 contains our definition of a (homological) faithful reconstruction. We then define weak sampling condition and optimal reconstruction algorithms in Section 2.3. Finally, Section 2.4 presents a naive algorithm that outputs a homological faithful reconstruction under a condition approaching the weak sampling condition.

2.1 Background

The goal of this section is to recall three closely related concepts useful for the expression of sampling conditions in shape reconstruction. Given a shape $\mathcal{A}$, we define the reach $r_1(\mathcal{A})$, the μ-reach $r_\mu(\mathcal{A})$ for any $\mu \in (0, 1]$ and the weak feature size $\mathrm{wfs}(\mathcal{A})$. As we shall see, these quantities are related by the following inequality: $r_1(\mathcal{A}) \leq r_\mu(\mathcal{A}) \leq \mathrm{wfs}(\mathcal{A})$. All three concepts can be derived from the critical function of the shape. This leads us to introduce the critical function, which requires first to define the norm of the gradient to the distance function.

The distance function to a compact set plays a central role in several recent works related to topologically guaranteed reconstruction [29, 23, 12]. For a compact set $\mathcal{A} \subset \mathbb{R}^N$, the distance function $d_\mathcal{A} : \mathbb{R}^N \to \mathbb{R}^+$ maps every point $q \in \mathbb{R}^N$ to

$$d_\mathcal{A}(q) \quad = \quad \min_{a \in \mathcal{A}} \|a - q\|.$$

Although not differentiable, $d_\mathcal{A}$ admits several extended notions of gradient [16, 29]. For our purpose, we will introduce a real valued function $\Psi_\mathcal{A} : \mathbb{R}^N \setminus \mathcal{A} \to [0, 1]$ which corresponds to the norm of the gradient defined in [29]. Let $\frac{\mathrm{d}}{\mathrm{d}t^+}(\cdot)_{|t=0}$ denote the right derivative with respect to the variable t at $t = 0$. For $q \in \mathbb{R}^N \setminus \mathcal{A}$ and $v \in \mathbb{S}^{N-1}$, one can check [29] that the quantity $\frac{\mathrm{d}}{\mathrm{d}t^+} d_\mathcal{A}(q+tv)_{|t=0}$ is well-defined and belongs to $[-1, 1]$. We define $\Psi_\mathcal{A}$ as:

$$\Psi_\mathcal{A}(q) \quad = \quad \max\left\{0, \sup_{v \in \mathbb{S}^{N-1}} \frac{\mathrm{d}}{\mathrm{d}t^+} d_\mathcal{A}(q + tv)_{|t=0}\right\}.$$

Roughly speaking, $\Psi_\mathcal{A}(q)$ quantifies at which maximal speed the distance function to $\mathcal{A}$ can increase in a neighborhood of q. We are now ready to recall the definition of the critical function $\chi_\mathcal{A}$ introduced in [12]. The critical function maps every positive real number $\rho > 0$ to the infimum of $\Psi_\mathcal{A}$ over the set of points at distance ρ from $\mathcal{A}$:

$$\chi_\mathcal{A}(\rho) \quad = \quad \inf_{d_\mathcal{A}(q) = \rho} \Psi_\mathcal{A}(q).$$

The critical function is lower semi-continuous [12] and allows to define two quantities, the *μ-reach* and the *weak feature size* of $\mathcal{A}$ denoted respectively $r_\mu(\mathcal{A})$ and $\mathrm{wfs}(\mathcal{A})$:

$$r_\mu(\mathcal{A}) \quad = \quad \inf\{\rho > 0,\, \chi_\mathcal{A}(\rho) < \mu\},$$
$$\mathrm{wfs}(\mathcal{A}) \quad = \quad \inf\{\rho > 0,\, \chi_\mathcal{A}(\rho) = 0\}.$$

The reach of $\mathcal{A}$ is equal to $r_1(\mathcal{A})$. From the definition, it is clear that $r_1(\mathcal{A}) \leq r_\mu(\mathcal{A}) \leq \mathrm{wfs}(\mathcal{A})$ for any $\mu \in (0, 1]$. Figure 1 shows the critical function $\chi_\mathcal{A}$ for a simple shape $\mathcal{A}$ in the Euclidean plane, which consists of the points at distance R from a full rectangle of width ℓ and length L.

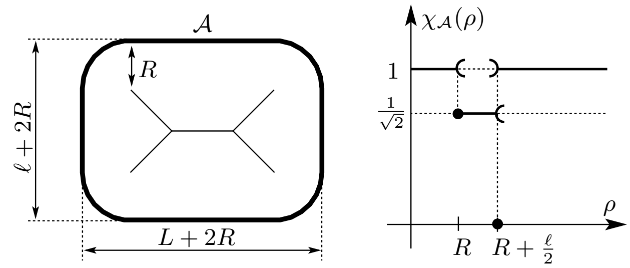

Figure 1: Left: the shape $\mathcal{A}$ is the outer closed thick curve and its medial axis consists of the five thin inner segments. Right: critical function $\chi_\mathcal{A}$. We have $r_\mu(\mathcal{A}) = R$ for $\mu > \frac{1}{\sqrt{2}}$ and $r_\mu(\mathcal{A}) = R + \frac{\ell}{2} = \mathrm{wfs}(A)$ for $\mu \leq \frac{1}{\sqrt{2}}$.

To shed light on these notions, it is useful to make some connections with the medial axis. The *medial axis* of $\mathcal{A}$ is the set of points $q \notin \mathcal{A}$ which have at least two closest points in $\mathcal{A}$. Alternatively, it is the locus of points q for which $\Psi_\mathcal{A}(q) < 1$. Any point q for which $\Psi_\mathcal{A}(q) = 0$ is called a *critical point* of the distance function and lies on the medial axis. The reach is the minimum of distances between points in $\mathcal{A}$ and in its medial axis. The weak feature size is the minimum of distances between points in $\mathcal{A}$ and critical points.

For instance, the function $\Psi_{\mathcal{A}}$ of the shape $\mathcal{A}$ depicted in Figure 1 evaluates to 0 on the horizontal line of the medial axis which constitutes the only critical points in this case, evaluates to $\frac{1}{\sqrt{2}}$ on the other points of the medial axis and evaluates to 0 on all points of the plane that neither belong to $\mathcal{A}$ nor to its medial axis.

For completeness, we also recall the related notion of local feature size, introduced by Amenta [2] for reconstructing smooth shapes. The *local feature size* is a real-valued function which maps every point of $\mathcal{A}$ to its distance to the medial axis. Notice that the local feature size and its infimum, the reach, vanish on non-smooth objects as soon as they contain a sharp concave corner or edge. For this reason, we will focus in Section 2.3 on sampling conditions based on the weak feature size and μ-reaches which apply to a large class of non-smooth shapes.

Given $\eta > 0$, the η-offset of $\mathcal{A}$ is the set of points at distance η or less from $\mathcal{A}$, $\mathcal{A}^\eta = d_{\mathcal{A}}^{-1}([0,\eta])$. As in Morse theory, topological changes in offsets occur only at critical values. More precisely, as stated in [28, 14]:

LEMMA 1 (TOPOLOGICAL STABILITY OF OFFSETS). *If $0 < x < y < \mathrm{wfs}(\mathcal{A})$, then the inclusion map $\mathcal{A}^x \hookrightarrow \mathcal{A}^y$ is a homotopy equivalence.*

2.2 Faithful reconstructions

Let us now give our definitions of a faithful reconstruction and a faithful homological reconstruction. For the second definition, we will consider a fixed field F and take coefficients in F for homology [31, Chapter 1]. Hence, the property of being a faithful homological reconstruction will depend on the choice of F.

DEFINITION 1. *We say that a subset $\mathcal{R} \subset \mathbb{R}^N$ is a* faithful reconstruction *of the compact set $\mathcal{A} \subset \mathbb{R}^N$ if there exist real numbers x, y such that $0 < x < y < \mathrm{wfs}(\mathcal{A})$ and the following two properties hold:*

- *$\mathcal{A}^x \subset \mathcal{R} \subset \mathcal{A}^y$*
- *the inclusion maps $\mathcal{A}^x \hookrightarrow \mathcal{R}$ and $\mathcal{R} \hookrightarrow \mathcal{A}^y$ are homotopy equivalences.*

We say that $\mathcal{R}$ is a faithful homological reconstruction *when the last condition is relaxed by:*

- *the inclusion maps $\mathcal{A}^x \hookrightarrow \mathcal{R}$ and $\mathcal{R} \hookrightarrow \mathcal{A}^y$ induces isomorphisms on all homology groups.*

A faithful reconstruction is always a faithful homological reconstruction. As expected, the converse is not true: a punctured Poincaré sphere nested between a point and a ball is an example where inclusions are not homotopy equivalences but yet induce isomorphisms on homology groups [17]. Interestingly, this example does not embed in $\mathbb{R}^3$.

Note that in the above definition, if one of the two inclusion maps $\mathcal{A}^x \hookrightarrow \mathcal{R}$ or $\mathcal{R} \hookrightarrow \mathcal{A}^y$ is a homotopy equivalence, so is the other one. Indeed, by Lemma 1, $\mathcal{A}^x \hookrightarrow \mathcal{A}^y$ is a homotopy equivalence and we can conclude by applying Lemma 2 below. A similar statement can be made for the second part of the definition.

LEMMA 2. *Consider three nested spaces $\mathcal{A} \subset \mathcal{B} \subset \mathcal{C}$. If two of the three inclusions $i : \mathcal{A} \hookrightarrow \mathcal{B}$, $j : \mathcal{B} \hookrightarrow \mathcal{C}$ and $k = j \circ i : \mathcal{A} \hookrightarrow \mathcal{C}$ are homotopy equivalences, then the third one is also a homotopy equivalence.*

PROOF. If i and j are homotopy equivalences with homotopy inverses i' and j' respectively, then $i' \circ j'$ is clearly a homotopy inverse of $k = j \circ i$.

If j and k are homotopy equivalences with homotopy inverses j' and k' respectively, then using $k = j \circ i$ we get that $j' \circ k = j' \circ j \circ i \simeq i$ and $k' \circ j$ is a homotopy inverse of $i \simeq j' \circ k$.

Similarly, if i and k are homotopy equivalences with homotopy inverses i' and k' respectively, then using $k = j \circ i$ we get that $k \circ i' = j \circ i \circ i' \simeq j$ and $i \circ k'$ is a homotopic inverse of $j \simeq k \circ i'$. □

2.3 Sampling conditions

In this section, we compare inputs, preconditions and outputs of two algorithms that infer information on a shape $\mathcal{A}$ known through a finite sample $\mathcal{S}$. Specifically, the first algorithm recovers Betti numbers of $\mathcal{A}$ and the second one constructs a faithful approximation of $\mathcal{A}$. Each algorithm relies on a key theorem that states sampling conditions ensuring correctness. Both algorithms are polynomial in the size of the sample. We then define an *optimal reconstruction algorithm* as one that would produce the output of the second algorithm with the input and precondition of the first algorithm.

We recall that the *Hausdorff distance* between two compact sets $\mathcal{A}$ and $\mathcal{A}'$ of $\mathbb{R}^N$ is defined by:

$$d_H(\mathcal{A}, \mathcal{A}') = \|d_{\mathcal{A}'} - d_{\mathcal{A}}\|_\infty = \sup_{q \in \mathbb{R}^N} |d_{\mathcal{A}'}(q) - d_{\mathcal{A}}(q)|.$$

Computing Betti numbers.

A powerful tool for inferring Betti numbers from geometric approximations is topological persistence [26]. Theorem 3 below is a corollary of the Persistence Stability Theorem [18] and can also be derived by flow based arguments [14]. Before stating it, we need the following definition.

DEFINITION 2. *Let $\mathcal{A} \subset \mathbb{R}^N$ be a compact set and let $0 \leq x \leq y$. The p-th (x,y)-*persistent Betti number *of $\mathcal{A}$ is the rank of the homomorphism induced by inclusion $\mathcal{A}^x \hookrightarrow \mathcal{A}^y$:*

$$\beta_p^{x,y}(\mathcal{A}) = \mathrm{rank}\left(\mathbf{H}_p(\mathcal{A}^x) \hookrightarrow \mathbf{H}_p(\mathcal{A}^y)\right)$$

It is worth noting that the (x,y)-persistent Betti numbers are finite whenever $x < y$ [20].

THEOREM 3 (HOMOLOGY INFERENCE [18, 14]). *Let $\mathcal{A}$ and $\mathcal{S}$ be two compact subsets of $\mathbb{R}^N$ and suppose there exists a real number $\alpha > 0$ such that*

$$d_H(\mathcal{S}, \mathcal{A}) < \alpha < \frac{1}{4}\,\mathrm{wfs}(\mathcal{A})$$

Then, $\beta_p(\mathcal{A}) = \beta_p^{\alpha,3\alpha}(\mathcal{S})$.

The above theorem leads immediately to a polynomial time algorithm for inferring Betti numbers of a shape $\mathcal{A}$ when the sample $\mathcal{S}$ of $\mathcal{A}$ is finite. Indeed, writing $K_\alpha(\mathcal{S})$ for the α-complex of $\mathcal{S}$, the persistent Betti numbers can be expressed as

$$\beta_p^{\alpha,3\alpha}(\mathcal{S}) = \mathrm{rank}\left(\mathbf{H}_p(K_\alpha(\mathcal{S})) \hookrightarrow \mathbf{H}_p(K_{3\alpha}(\mathcal{S}))\right).$$

In particular, they can be computed in time cubic the size of $K_{3\alpha}(\mathcal{S})$. Since for a fixed dimension, the size of α-complexes is polynomial in the number of vertices, it follows that $\beta_p(\mathcal{A})$ can also be computed in polynomial time the size of the sample.

Table 1: Input, precondition and output of two polynomial time algorithms derived from Theorem 3 and Theorem 4. The notation "$\exists\mathcal{A}$" stands for "there exists a compact set $\mathcal{A} \subset \mathbb{R}^N$". $\mathcal{S}$ designates a finite set of $\mathbb{R}^N$. α and μ designate two real numbers with $\alpha > 0$ and $\mu \in (0, 1]$.

Input	Precondition	Output
$\mathcal{S}$, α	$\exists\mathcal{A}, d_H(\mathcal{S},\mathcal{A}) < \alpha < \frac{1}{4}\,\mathrm{wfs}(\mathcal{A})$	Betti numbers of $\mathcal{A}$
$\mathcal{S}$, α, μ	$\exists\mathcal{A}, d_H(\mathcal{S},\mathcal{A}) < \alpha < \frac{\mu^2}{5\mu^2+12} r_\mu(\mathcal{A})$	a faithful reconstruction of $\mathcal{A}$

Computing a faithful reconstruction.

For a shape $\mathcal{A}$ with a positive μ-reach, authors in [12] describe a simple procedure for computing a faithful reconstruction of $\mathcal{A}$, given as input a sample $\mathcal{S}$ of $\mathcal{A}$. The procedure consists merely of outputting an r-offset of the sample, for a suitable value of the offset parameter r (see Theorem 4 below for the precise value of r). In practice, this computation can be replaced by the computation of $K_r(\mathcal{S})$, which shares the same homotopy type. Both computations can be done in polynomial time if the sample is finite. The sampling condition required by the procedure is that the Hausdorff distance between the sample $\mathcal{S}$ and the sampled shape $\mathcal{A}$ is less than a fraction the μ-reach of $\mathcal{A}$. More precisely:

Theorem 4 (Reconstruction Theorem [12]). *Let $\mathcal{A}$ and $\mathcal{S}$ be two compact subsets of $\mathbb{R}^N$ and suppose there exists two real numbers $\alpha > 0$ and $\mu \in (0, 1]$ such that*

$$d_H(\mathcal{S},\mathcal{A}) < \alpha < \frac{\mu^2}{5\mu^2+12} r_\mu(\mathcal{A})$$

Then, $\mathcal{S}^{\frac{4\alpha}{\mu^2}}$ is a faithful reconstruction of $\mathcal{A}$.

Comparing sampling conditions.

Table 1 summarizes inputs, preconditions and outputs of the two polynomial time algorithms described above and inspired by Theorems 3 and 4. Note that the precondition required by the first algorithm which we call the *weak precondition* is equivalent to saying that the following set is non-empty:

$$W(\mathcal{S},\alpha) = \left\{\mathcal{X} \mid d_H(\mathcal{S},\mathcal{X}) < \alpha < \frac{1}{4}\,\mathrm{wfs}(\mathcal{X})\right\} \neq \emptyset, \qquad (1)$$

where $\mathcal{X}$ ranges over all compact sets of $\mathbb{R}^N$. By Theorem 3, all shapes in $W(\mathcal{S},\alpha)$ share the same Betti numbers and the first algorithm returns the Betti numbers of any $\mathcal{A} \in W(\mathcal{S},\alpha)$. Since $r_\mu(\mathcal{A})$ is by definition upper bounded by $\mathrm{wfs}(\mathcal{A})$, the first precondition is significantly weaker than the second precondition, especially when μ is small. We now claim that, in some sense, the input of the first algorithm together with its weak precondition determines completely the output of the second. Indeed, let us recall from [14] the following theorem:

Theorem 5 ([14]). *Let $\mathcal{A}$ and $\mathcal{X}$ be two compact subsets of $\mathbb{R}^N$ and $\alpha > 0$ a real number such that*

$$d_H(\mathcal{A},\mathcal{X}) < 2\alpha < \frac{1}{2}\min\{\mathrm{wfs}(\mathcal{A}),\ \mathrm{wfs}(\mathcal{X})\}$$

Then, $\mathcal{X}^{2\alpha}$ is a faithful reconstruction of $\mathcal{A}$.

Suppose $\mathcal{A}$ and $\mathcal{X}$ both belong to $W(\mathcal{S},\alpha)$. Applying a triangular inequality, we get that $\mathcal{A}$, $\mathcal{X}$ and α fulfill conditions of Theorem 5 and therefore, $\mathcal{X}^{2\alpha}$ is a faithful reconstruction of $\mathcal{A}$. Hence, any 2α-offset of a shape $\mathcal{X} \in W(\mathcal{S},\alpha)$ is a faithful reconstruction of any shape $\mathcal{A} \in W(\mathcal{S},\alpha)$. In other words, all shapes in $W(\mathcal{S},\alpha)$ share a common set of faithful reconstructions which contains 2α-offsets of $W(\mathcal{S},\alpha)$. For this reason, we say that, under the weak precondition, inputs $\mathcal{S}$ and α carry in principle enough information about the unknown shape $\mathcal{A}$ to determine without ambiguity a faithful reconstruction of it.

Optimal reconstruction algorithms.

The weak precondition is tight, by which we mean that for any $\eta > 0$, the set $W_\eta(\mathcal{S},\alpha) = \{\mathcal{X}$ compact set of $\mathbb{R}^N \mid d_H(\mathcal{S},\mathcal{X}) < \alpha < \frac{1}{4}\,\mathrm{wfs}(\mathcal{X}) + \eta\}$ may contain objects that do not have the same homology. To construct such an example, consider the two shapes $\mathcal{O}$ and $\mathcal{U}$ described in [14] and the sample $\mathcal{S}$ pictured on Figure 2. By construction, we have $\mathrm{wfs}(\mathcal{O}) = 2$, $\mathrm{wfs}(\mathcal{U}) = +\infty$ and $d_H(S,\mathcal{O}) = d_H(S,\mathcal{U}) = \frac{1+\eta}{2}$. Hence, both $\mathcal{O}$ and $\mathcal{U}$ belong to $W_\eta(\mathcal{S}, \frac{1}{2} + \frac{3\eta}{4})$ but $\beta_1(\mathcal{O}) \neq \beta_1(\mathcal{U})$. Therefore, the weak precondition is the weakest amongst the preconditions expressed in terms of Hausdorff distance and critical functions that allows to retrieve a faithful reconstruction without ambiguity.

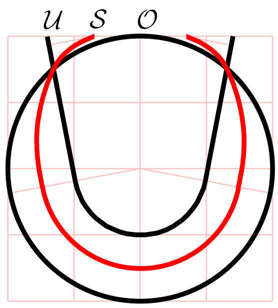

Figure 2: The angle between the two bars in shape $\mathcal{U}$ is adjusted such that $d_H(\mathcal{S},\mathcal{O}) = d_H(\mathcal{S},\mathcal{U}) = \frac{1+\eta}{2}$.

Because of that, we call any algorithm that would be able to output a faithful reconstruction under the weak precondition and associated inputs an *optimal reconstruction algorithm*. In Section 2.4, we describe a naive algorithm that outputs a faithful homological reconstruction under the weak precondition. We call it "naive" since it has an unbounded time complexity. The main question motivating our work is whether there exists a polynomial time optimal reconstruction algorithm. Section 3 suggests a negative answer if no additional condition is assumed.

2.4 Naive algorithms for reconstruction

Given as input a pair $(\mathcal{S}, \alpha)$ satisfying the weak precondition, the previous section suggests the following strategy for computing a faithful reconstruction: enumerate all compact sets in $\mathbb{R}^N$ and return the 2α-offset of the first compact set $\mathcal{X}$ that belongs to $W(\mathcal{S}, \alpha)$. Of course, this procedure is unrealistic and the goal of this section is to present an effective version of it. Specifically, given as input a sample $\mathcal{S}$ and two real numbers α and η satisfying the precondition:

$$\exists \mathcal{A},\ d_H(\mathcal{S}, \mathcal{A}) < \alpha < \frac{1}{4}(\mathrm{wfs}(\mathcal{A}) - \eta), \tag{2}$$

we describe an algorithm that computes a faithful homological reconstruction of $\mathcal{A}$ and whose pseudocode is given in Table 2. The idea is to replace the enumeration on compact sets by an enumeration on cubical sets and refine the size of the cubes until we find a solution. We also simplify the problem, replacing the search for a faithful reconstruction by the search of a faithful homological simplification. We proceed in two steps. First, we give an algorithm for shapes with a positive μ-reach (`NAIVE_1`) and then, derive an algorithm for shapes with a lower bounded weak feature size (`NAIVE_2`).

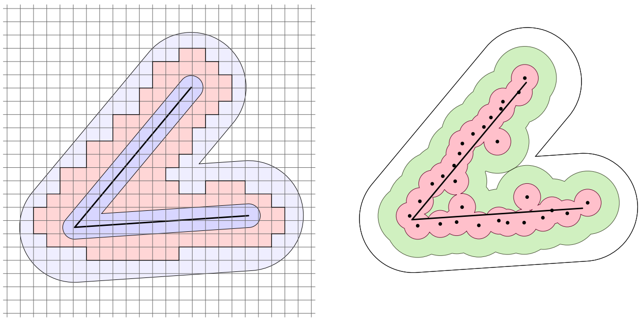

Figure 3: Left: the cubical set (in pink) is nested between two offsets (in light purple) of the V-shaped black curve and is a faithful reconstruction of it. Right: offsets $\mathcal{S}^l$ and $\mathcal{S}^k$ of the sample.

We start with some definitions. An *ε-voxel* is a closed cube with edge length ε and whose vertices belong to the lattice $\varepsilon\mathbb{Z}^N$. We call any finite union of ε-voxels an *ε-cubical set*. Assuming a shape $\mathcal{A}$ has a positive μ-reach, next lemma states the existence of a cubical set which is a faithful reconstruction of $\mathcal{A}$. This is a key ingredient in establishing the correctness of the naive algorithms.

LEMMA 6. *There exists a positive constant c_N depending only upon the ambient dimension N such that the following property holds: for all real numbers x, y and $\mu \in (0, 1]$ and all compact sets $\mathcal{A} \subset \mathbb{R}^N$ satisfying $r_\mu(\mathcal{A}) > y > x > 0$, there exists a $(c_N \mu(y - x))$-cubical set $\mathcal{X}$ such that $\mathcal{A}^x \subset \mathcal{X} \subset \mathcal{A}^y$ and the inclusion maps $\mathcal{A}^x \hookrightarrow \mathcal{X}$ and $\mathcal{X} \hookrightarrow \mathcal{A}^y$ are homotopy equivalences. In particular, the cubical set $\mathcal{X}$ is a faithful reconstruction of $\mathcal{A}$ (see Figure 3, left).*

The proof is in the appendix and relies on a result in [8].

First naive reconstruction algorithm.

Under precondition (2), the algorithm `NAIVE_1` outputs either the empty set or a faithful homological reconstruction

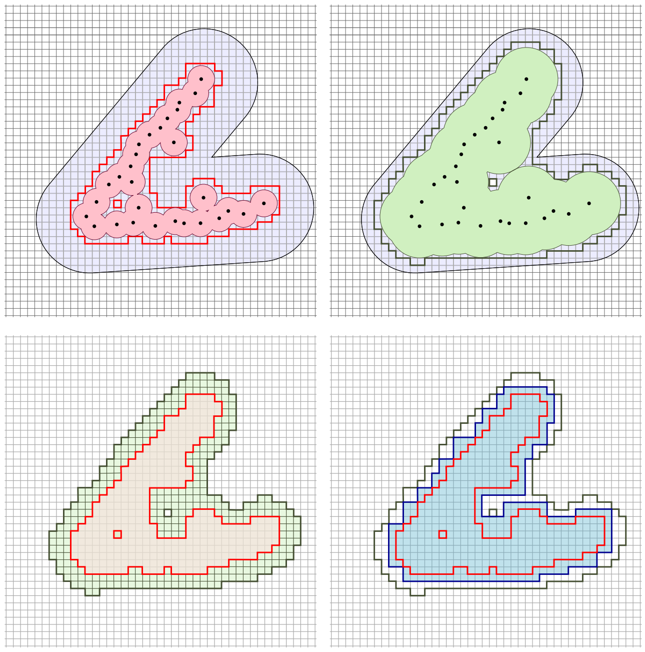

Figure 4: Illustration of algorithm `NAIVE_1`. Top: boundaries of $\mathcal{L} = V_\varepsilon(\mathcal{S}^l)$ and $\mathcal{K} = V_\varepsilon(\mathcal{S}^k)$ are depicted in red and dark green. Bottom: the cubical set $\mathcal{X}$ in blue is nested between $\mathcal{L}$ and $\mathcal{K}$ and is a faithful homological reconstruction of $\mathcal{A}$.

of $\mathcal{A}$. Its pseudocode is given in Table 2, left. The algorithm proceeds as follows. It chooses a voxel size ε, two offset parameters l and k and derives from the sample $\mathcal{S}$ two ε-cubical sets $\mathcal{L} = V_\varepsilon(\mathcal{S}^l)$ and $\mathcal{K} = V_\varepsilon(\mathcal{S}^k)$, obtained by collecting all ε-voxels intersecting respectively $\mathcal{S}^l$ and $\mathcal{S}^k$ (see Figures 3 and 4). For all cubical sets $\mathcal{X}$ containing $\mathcal{L}$ and contained in $\mathcal{K}$, the algorithm then considers three nested simplicial complexes $L \subset X \subset K$ triangulating the three cubical sets $\mathcal{L} \subset \mathcal{X} \subset \mathcal{K}$ in a way that is consistent with the grid. It then returns $\mathcal{X}$ if the simplicial complex X is a homological simplification of the pair (K, L) (see definition below). If no homological simplification X is found between L and K, the algorithm returns the empty set.

DEFINITION 3 (HOMOLOGICAL SIMPLIFICATION). *Let $L \subset K$ be two simplicial complexes. The simplicial complex X is said to be a* homological simplification *of the pair (K, L) if $L \subset X \subset K$ and the maps $j_* : \mathbf{H}_p(L) \to \mathbf{H}_p(X)$ and $i_* : \mathbf{H}_p(X) \to \mathbf{H}_p(K)$ induced by inclusions are respectively surjective and injective for all integers $p \geq 0$ (see Figure 5).*

$$\mathbf{H}_p(L) \xrightarrow{\ j_*\ } \mathbf{H}_p(X) \xrightarrow{\ i_*\ } \mathbf{H}_p(K)$$

Figure 5: Notations for Definition 3 and the proof of Lemma 7.

The correctness of the algorithm relies on several lemmas stated in the appendix. Under precondition (2), if X is a homological simplification of the pair (K, L), then Lemma 12 implies that $\mathcal{X} = |X|$ is a faithful homological reconstruction

Table 2: Naive reconstruction algorithms. $|X|$ denotes the underlying space of the simplicial complex X. $sc(\mathcal{X})$ denotes a triangulation of the cubical set $\mathcal{X}$ compatible with inclusion.

`NAIVE_1(` $\mathcal{S}, \alpha, \eta, \mu$ `)`
`PRECONDITION:` $\exists \mathcal{A},\ d_H(\mathcal{S}, \mathcal{A}) < \alpha < \frac{1}{4}(\mathrm{wfs}(\mathcal{A}) - \eta)$
`OUTPUT:` either $\emptyset$ or a faithful homological reconstruction of $\mathcal{A}$.

$\varepsilon \leftarrow \frac{\eta}{4\sqrt{N}+2/(c_N\mu)}$
$l \leftarrow \frac{\eta}{2} + \alpha$; $\quad k \leftarrow \eta + 3\alpha - \varepsilon\sqrt{N}$
$L \leftarrow \mathrm{sc}(V_\varepsilon(\mathcal{S}^l))$; $\quad K \leftarrow \mathrm{sc}(V_\varepsilon(\mathcal{S}^k))$

`for all (` X such that $L \subset X \subset K$ `)`
  `if (` X is a homological simplification of (K, L) `)`
  `return` $|X|$
`end for`
`return` $\emptyset$

`NAIVE_2(` $\mathcal{S}, \alpha, \eta$ `)`
`PRECONDITION:` $\exists \mathcal{A},\ d_H(\mathcal{S}, \mathcal{A}) < \alpha < \frac{1}{4}(\mathrm{wfs}(\mathcal{A}) - \eta)$
`OUTPUT:` a faithful homological reconstruction of $\mathcal{A}$

$\mu \leftarrow 1$
$\alpha' \leftarrow \alpha + \frac{\eta}{8}$
$\eta' \leftarrow \frac{\eta}{4}$

`while ( TRUE )`
  $\mathcal{X} \leftarrow$ `NAIVE_1(`$\mathcal{S}, \alpha', \eta', \mu$`)`
  `if (` $\mathcal{X}$ `!=` $\emptyset$ `) return` $\mathcal{X}$
  $\mu \leftarrow \frac{\mu}{2}$
`end while`

of $\mathcal{A}$. Furthermore, under the stronger precondition (3):

$$\exists \mathcal{A},\ d_H(\mathcal{S}, \mathcal{A}) < \alpha < \frac{1}{4}\left(r_\mu(\mathcal{A}) - \eta\right), \tag{3}$$

Lemma 11 guarantees that the algorithm always returns a faithful homological simplification (and not $\emptyset$). Let us bound the time complexity of a more efficient version of the algorithm in which voxels are not decomposed into simplices. Let D be the diameter of $\mathcal{S}$ and set $D' = D + 2(\nu + 3\alpha)$. It is not difficult to check that this simpler algorithm has time complexity $O\left(2^{|K|}|K|^3\right) = O\left(2^{\left(\frac{D'}{\varepsilon}\right)^N}\left(\frac{D'}{\varepsilon}\right)^{3N}\right)$. Indeed, the size of K is $O((D'/\varepsilon)^N)$. Checking if X is a homological simplification of (K, L) takes cubic time the size of K and the number of cubical sets $\mathcal{X}$ between $\mathcal{L}$ and $\mathcal{K}$ is $O(2^{|K|})$.

Second naive reconstruction algorithm.

Under precondition (2), the algorithm `NAIVE_2` outputs a faithful homological reconstruction of $\mathcal{A}$ after a finite number of iterations. Its pseudocode is given in Table 2, right. Starting with $\mu = 1$, it calls `NAIVE_1` with decreasing values of μ until `NAIVE_1` returns a faithful homological reconstruction. The algorithm terminates thanks to the lower semi-continuity of the critical function $\chi_\mathcal{A}$. Indeed, $\chi_\mathcal{A}$ attains its minimum $\mu' > 0$ over the interval $[\frac{\eta}{8}, 4\alpha + \frac{7\eta}{8}]$. Setting $\mathcal{A}' = \mathcal{A}^{\frac{\eta}{8}}$, $\alpha' = \alpha + \frac{\eta}{8}$ and $\eta' = \frac{\eta}{4}$, we have $r_{\mu'}(\mathcal{A}') > 4\alpha + \frac{3\eta}{4} = 4\alpha' + \eta'$ (see Figure 6 for an explanation) and therefore precondition (3) is satisfied for $\mathcal{A} = \mathcal{A}'$, $\alpha = \alpha'$, $\mu = \mu'$ and $\eta = \eta'$. Note that even though the algorithm terminates, its time complexity is unbounded.

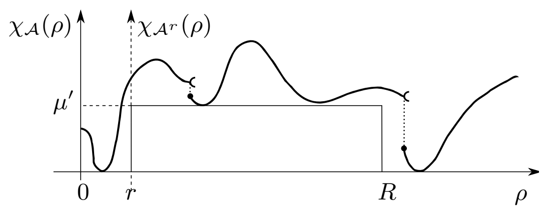

Figure 6: Performing an r-offset translates into translating the critical function to the left by r [12]. Thus, $\chi_\mathcal{A}(\rho) \geq \mu'$ on $[r, R]$ implies $r_{\mu'}(\mathcal{A}^r) > R - r$.

3. HOMOLOGICAL SIMPLIFICATION IS NP-COMPLETE

In this section, we focus on the problem of computing a homological simplification and prove that this problem is NP-complete. We denote the *p-th homology group* of K by $\mathbf{H}_p(K)$ and work with coefficients in the field $\mathbb{Z}_2$ of integers modulo 2. A *simplicial pair* (K, L) consists of a (finite) simplicial complex K and a subcomplex $L \subset K$. When clear from the context, we will simply speak of the pair (K, L) and omit "simplicial". We say that the pair (K, L) is *p-dimensional* if the simplicial complex K has dimension p.

DEFINITION 4. *The* homological simplification problem *takes as input a simplicial pair (K, L) and asks whether there exists a simplicial complex X which is a homological simplification of the pair (K, L).*

The size of the problem is the number of simplices in K. A useful observation is that since we are working with coefficients in $\mathbb{Z}_2$ and homology groups are finite-dimensional vector spaces, X is a homological simplification of the pair (K, L) if and only if X realizes the persistent homology of L into K. Formally, we have:

LEMMA 7. *Consider a sequence of simplicial complexes $L \subset X \subset K$. The simplicial complex X is a homological simplification of the pair (K, L) if and only if $\mathbf{H}_p(X)$ is isomorphic to the image of the homomorphism $\mathbf{H}_p(L) \to \mathbf{H}_p(K)$ induced by the inclusion $L \subset K$, for all integers $p \geq 0$.*

PROOF. See Figure 5. If X is a homological simplification of the pair (K, L), then the injectivity of the map $i_* : \mathbf{H}_p(X) \to \mathbf{H}_p(K)$ induced by the inclusion $X \subset K$ implies that $\mathbf{H}_p(X)$ is isomorphic to $i_*(\mathbf{H}_p(X))$ and the surjectivity of the map $j_* : \mathbf{H}_p(L) \to \mathbf{H}_p(X)$ induced by inclusion $L \subset X$ implies that $\mathbf{H}_p(X) = j_*(\mathbf{H}_p(L))$. It follows that $\mathbf{H}_p(X)$ is isomorphic to $i_* \circ j_*(\mathbf{H}_p(L))$ for all p.

Conversely, suppose $\mathbf{H}_p(X)$ is isomorphic to $i_* \circ j_*(\mathbf{H}_p(L))$. Then, j_* is surjective because $\dim j_*(\mathbf{H}_p(L)) \geq \dim i_* \circ j_*(\mathbf{H}_p(L)) = \dim \mathbf{H}_p(X)$. Furthermore, i_* is injective because the surjectivity of j_* implies $i_*(\mathbf{H}_p(X)) = i_* \circ j_*(\mathbf{H}_p(L))$ which is isomorphic to $\mathbf{H}_p(X)$. $\square$

We are now ready to state our main theorem:

Theorem 8. *The homological simplification problem of 2-dimensional simplicial pairs is NP-complete.*

Proof. To check that a candidate X is a homological simplification of the p-dimensional pair (K, L), it is enough to compute the dimension of the p-th homology group of X and compare it to the rank of the persistent p-th homology group of K into L, for all p. Since all computations can be done in time cubic in the number of simplices in K, we deduce that the homological simplification problem of p-dimensional simplicial pairs is in NP. In Section 3.1, we prove that this problem is NP-hard for $p = 2$ by reducing 3SAT to it in polynomial time. Figure 7 summarizes the reduction. $\square$

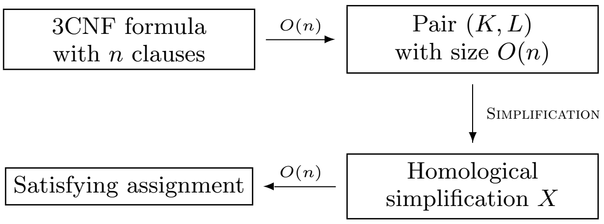

Figure 7: Diagram of the reduction.

3.1 Reduction from 3SAT

A Boolean formula E is in *3-conjunctive normal form*, or 3CNF, if it is a conjunction (AND) of n clauses $c_1, c_2, \ldots, c_n$, each of which is a disjunction (OR) of three literals, each literal being a Boolean variable or its negation [19]. Specifically, $E = \bigwedge_{1 \le i \le n} c_i$ and each clause c_i has the form

$$c_i = \left(e_i^1 v_{j_i^1}\right) \vee \left(e_i^2 v_{j_i^2}\right) \vee \left(e_i^3 v_{j_i^3}\right),$$

where $j_i^k \in \{1, \ldots, m\}$, $v_{j_i^k}$ is a Boolean variable and $e_i^k \in \{\mathbf{1}, \neg\}$ is either the identity symbol $\mathbf{1}$ or the negation symbol $\neg$, for $1 \le k \le 3$. The 3SAT problem takes as input a 3CNF formula E and determines whether one can assign a value TRUE or FALSE to each variable of E such that E evaluates to TRUE. An assignment of variables which makes E evaluates to TRUE is called a *satisfying assignment*. Since the number m of variables used in formula E is at most three times the number n of clauses, *i.e.* $m \le 3n$, we let n be the size of the 3SAT problem. 3SAT is known to be NP-complete.

Reduction algorithm.

We describe a reduction algorithm that transforms in linear time any instance E of the 3SAT problem into an instance (K, L) of the homological simplification problem in such a way that (K, L) has a homological simplification if and only if E has a satisfying assignment. Given a 3CNF formula E of n clauses $c_1, \ldots, c_n$ and m variables $v_1, \ldots, v_m$, we construct a 2-dimensional simplicial pair (K, L) as follows; see Figure 8. The simplicial complex L consists of

- a vertex A;
- two vertices B_i and C_i and three edges AB_i, B_iC_i and C_iA for each clause c_i;
- two vertices V_j and W_j and the edge V_jW_j for each variable v_j.

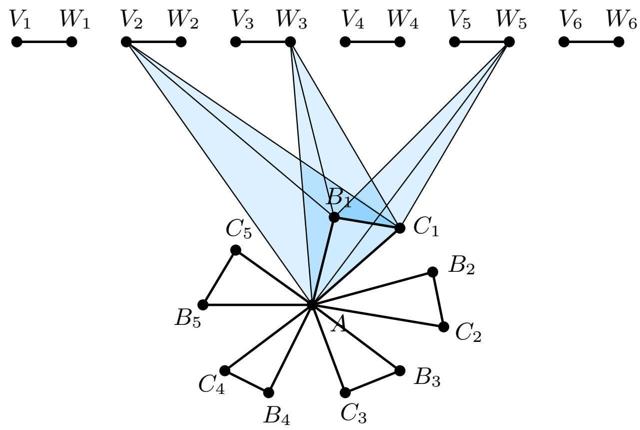

Figure 8: Simplicial complex L output by the reduction of a formula with five clauses and six variables and triangles in K created by clause $c_1 = v_2 \vee \neg v_3 \vee \neg v_5$.

Besides simplices in L, the simplicial complex K contains three triangles per literal and two edges per variable. Specifically, if $e_i^k = \mathbf{1}$, we add the three triangles $AB_iV_{j_i^k}$, $B_iC_iV_{j_i^k}$ and $C_iAV_{j_i^k}$ and their edges. If $e_i^k = \neg$, we add the three triangles $AB_iW_{j_i^k}$, $B_iC_iW_{j_i^k}$ and $C_iAW_{j_i^k}$ and their edges. Moreover, we add edges AV_j and AW_j for all $j \in \{1, \ldots, m\}$. Observe that the size of K is only a constant factor larger than the size of E and its construction requires linear time in n.

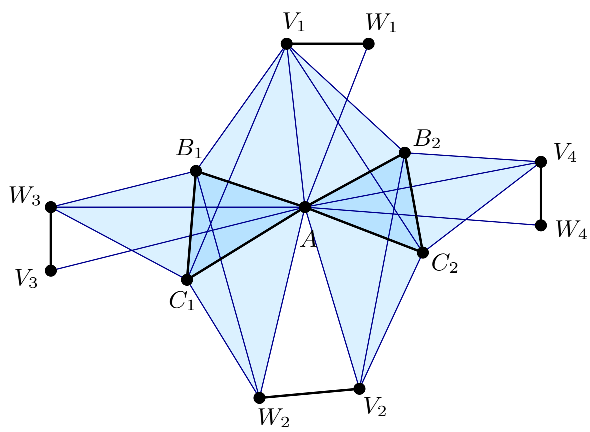

Figure 9: Pair (K, L) produced by the reduction of formula $(v_1 \vee \neg v_2 \vee \neg v_3) \wedge (v_1 \vee v_2 \vee v_4)$. L consists of the vertices and bold edges.

Let $f_* : \mathbf{H}_p(L) \to \mathbf{H}_p(K)$ be the homomorphism induced by the inclusion $L \subset K$. Since K is connected, we have $f_*(\mathbf{H}_0(L)) = \mathbb{Z}_2$. Furthermore, $f_*(\mathbf{H}_1(L)) = 0$ since a base of the 1-cycles in L is given by the n cycles $\sigma_i = AB_i + B_iC_i + C_iA$ and σ_i is homologous to 0 in K for each $i \in \{1, \ldots, n\}$. Last, $f_*(\mathbf{H}_2(L)) = 0$ since L contains no 2-simplices. By Lemma 7, we obtain that X is a homological simplification of the pair (K, L) if and only if $\mathbf{H}_0(X) = \mathbb{Z}_2$, $\mathbf{H}_1(X) = 0$ and $\mathbf{H}_2(X) = 0$. Keeping this in mind, we establish the following lemma, in which (K, L) designates the pair output by our reduction algorithm when applied to formula E.

Lemma 9. *The pair (K, L) has a homological simplification if and only if the formula E has a satisfying assignment. Furthermore, given a homological simplification of the pair (K, L), computing a satisfying assignment for E takes linear time.*

Proof. Suppose the pair (K, L) has a homological simplification X and let us prove that E has a satisfying assignment. First, we claim that X cannot contain both edges AV_j and AW_j, for $1 \leq j \leq m$. Indeed, if both edges AV_j and AW_j were in X, we could consider the cycle $\tau = AV_j + V_jW_j + W_jA$. Since the edge V_jW_j bounds no triangle in K, the cycle τ cannot be homologous to 0 in X, contradicting $\mathbf{H}_1(X) = 0$.

The claim allows us to assign to each variable v_j either the value `TRUE` if the edge AV_j belongs to X or the value `FALSE` if the edge AW_j belongs to X. If none of the edges AV_j and AW_j belong to X, then we assign to v_j an arbitrary value in $\{\texttt{TRUE}, \texttt{FALSE}\}$; see Figure 10. Note that the computation of this assignment can be done in linear time. We now check that this assignment of boolean values to the variables v_j is a satisfying assignment, in other words we show that all clauses c_i are satisfied for $1 \leq i \leq n$.

Since $\mathbf{H}_1(X) = 0$, the 1-cycle $AB_i + B_iC_i + C_iA$ is a boundary in X. This implies that at least one triangle of X contains AB_i on its boundary. By construction, AB_i belongs to exactly three triangles in K, namely the triangles $AB_iY_i^k$ for $1 \leq k \leq 3$ where Y_i^k designates $V_{j_i^k}$ if $e_i^k = \mathbf{1}$ and $W_{j_i^k}$ if $e_i^k = \neg$. It follows that one of the three triangles $AB_iY_i^k$ must belong to X and, in turn, at least one of the three edges AY_i^k for $1 \leq k \leq 3$ is in X. This implies that one of the three literals $e_i^k v_{j_i^k}$ in clause c_i evaluates to `TRUE` and hence c_i is satisfied.

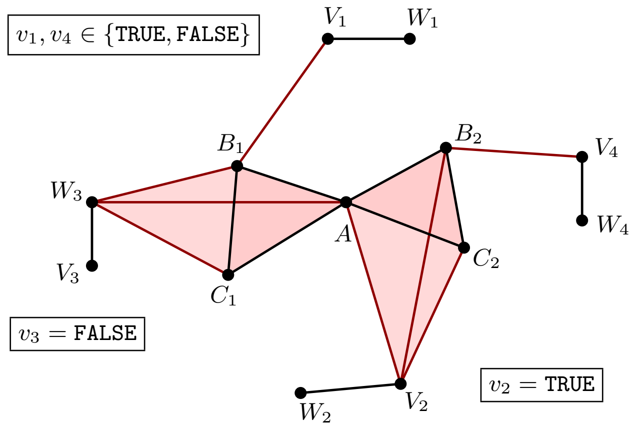

Figure 10: A homological simplification of the pair (K, L) drawn in Figure 9 and output by the reduction of formula $E = (v_1 \vee \neg v_2 \vee \neg v_3) \wedge (v_1 \vee v_2 \vee v_4)$. Corresponding satisfying assignments for E.

Conversely, suppose variables $v_1, \ldots, v_m$ have been assigned values that cause E to evaluate to `TRUE` and let us prove that the pair (K, L) has a homological simplification X. We construct X starting from L and adding some simplices of K as follows; see Figure 11. We begin by adding the edge AV_j if $v_j = \texttt{TRUE}$ and the edge AW_j if $v_j = \texttt{FALSE}$, for all $j \in \{1, \ldots, m\}$. Since values of $v_1, \ldots, v_m$ are a satisfying assignment, we can choose one literal $e_iv_{j_i}$ in each clause c_i that is true. Let $Y_i = V_{j_i}$ if $e_i = \mathbf{1}$ and $Y_i = W_{j_i}$ if $e_i = \neg$. Note that by construction, the edge AY_i is already in X. We then add the three triangles AB_iY_i, $B_iC_iY_i$ and C_iAY_i to X, for all $i \in \{1, \ldots, n\}$.

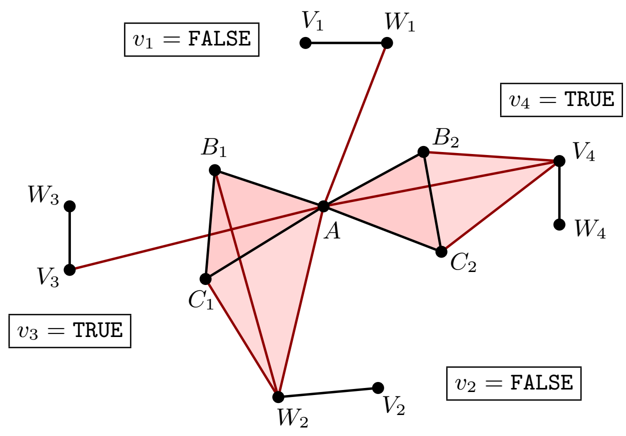

Figure 11: Satisfying assignment for formula $E = (v_1 \vee \neg v_2 \vee \neg v_3) \wedge (v_1 \vee v_2 \vee v_4)$ and corresponding homological simplification of (K, L).

Let us check that X is indeed a solution to the homological simplification problem, *i.e.* $\mathbf{H}_0(X) = \mathbb{Z}_2$, $\mathbf{H}_1(X) = 0$ and $\mathbf{H}_2(X) = 0$. For this, we check that X is contractible by collapsing X to A, using a sequence of elementary collapses. First, observe that exactly one of the two vertices V_j or W_j belongs to no other simplices than the edge V_jW_j. For instance, if $v_j = \texttt{TRUE}$, then by construction $AV_j \in X$ and $AW_j \notin X$. Thus, W_j belongs to no other simplices than V_jW_j and we can collapse the edge V_jW_j to the vertex V_j by removing the pair of simplices (W_j, V_jW_j). Similarly, if $v_j = \texttt{FALSE}$, we collapse the edge V_jW_j to the vertex W_j. For all $i \in \{1, \ldots, n\}$, we apply five elementary collapses, first removing the three triangles AB_iY_i, $B_iC_iY_i$ and C_iAY_i and their edges AB_i, B_iC_i and C_iA, then removing the edges B_iY_i and C_iY_i and their vertices B_i and C_i. In a last step, we collapse every edge AY_i for $1 \leq i \leq n$ to the vertex A. □

4. DISCUSSION

Our work raises several questions and research tracks.

Open questions.

Is there a version of Lemma 6 in which the voxel size does not depend on μ? Is the homological simplification problem in the same class of complexity if we constraint K to be a subcomplex of a triangulation of the sphere $\mathbb{S}^3$?

Optimistic research track.

If a polynomial time optimal reconstruction algorithm exists, it should take advantage of the embedding in Euclidean space or at least lead to a class of simplification problems sufficiently constrained to avoid constructions similar to ours.

Pessimistic research track.

Is it possible to encode 3-SAT as the homological simplification of a pair $(K_{3\alpha}(\mathcal{S}), K_{\alpha}(\mathcal{S}))$, where $(\mathcal{S}, \alpha)$ satisfies the weak sampling condition? Or, as the homological simplification of a pair of cubical complexes defined by offsets of the sample? If yes, in which minimal dimension?

5. REFERENCES

[1] N. Amenta and M. Bern. Surface reconstruction by Voronoi filtering. *Discrete and Computational Geometry*, 22(4):481–504, 1999.

[2] N. Amenta, M. Bern, and D. Eppstein. The crust and the β-skeleton: Combinatorial curve reconstruction. *Graphical Models and Image Processing*, 60(2):125–135, 1998.

[3] N. Amenta, S. Choi, T. Dey, and N. Leekha. A simple algorithm for homeomorphic surface reconstruction. In *Proceedings of the sixteenth annual symposium on Computational geometry*, pages 213–222. ACM New York, NY, USA, 2000.

[4] N. Amenta, S. Choi, and R. Kolluri. The power crust, unions of balls, and the medial axis transform. *Computational Geometry: Theory and Applications*, 19(2-3):127–153, 2001.

[5] D. Attali. r-regular shape reconstruction from unorganized points. *Computational Geometry: Theory and Applications*, 10:239–247, 1998.

[6] D. Attali, M. Glisse, S. Hornus, F. Lazarus, and D. Morozov. Persistence-sensitive simplification of functions on surfaces in linear time. *Manuscript, INRIA*, 2008.

[7] D. Attali and A. Lieutier. Reconstructing shapes with guarantees by unions of convex sets. http://hal.archives-ouvertes.fr/hal-00427035/en/.

[8] D. Attali and A. Lieutier. Reconstructing shapes with guarantees by unions of convex sets. In *Proc. ACM Symposium on Computational Geometry*, 2010. Submitted.

[9] J. Boissonnat and F. Cazals. Smooth surface reconstruction via natural neighbour interpolation of distance functions. *Computational Geometry: Theory and Applications*, 22(1-3):185–203, 2002.

[10] J. Boissonnat, L. Guibas, and S. Oudot. Manifold reconstruction in arbitrary dimensions using witness complexes. *Discrete and Computational Geometry*, 42(1):37–70, 2009.

[11] J. Boissonnat and S. Oudot. Provably good sampling and meshing of Lipschitz surfaces. In *Proceedings of the twenty-second annual symposium on Computational geometry*, page 346. ACM, 2006.

[12] F. Chazal, D. Cohen-Steiner, and A. Lieutier. A sampling theory for compact sets in Euclidean space. *Discrete and Computational Geometry*, 41(3):461–479, 2009.

[13] F. Chazal, D. Cohen-Steiner, A. Lieutier, and B. Thibert. Shape smoothing using double offsets. In *Proc. of the ACM symposium on Solid and physical modeling*, pages 183–192. ACM New York, NY, USA, 2007.

[14] F. Chazal and A. Lieutier. Stability and computation of topological invariants of solids in $\mathbb{R}^n$. *Discrete and Computational Geometry*, 37(4):601–617, 2007.

[15] F. Chazal and S. Oudot. Towards persistence-based reconstruction in Euclidean spaces. In *Proceedings of the twenty-fourth annual symposium on Computational geometry*, pages 232–241. ACM, 2008.

[16] F. Clarke. *Optimization and nonsmooth analysis.* Society for Industrial Mathematics, 1990.

[17] D. Cohen-Steiner. Personal communication. 2008.

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

[19] T. Cormen, C. Leiserson, R. Rivest, and C. Stein. Introduction to algorithms, 2001.

[20] V. de Silva. Personal communication. 2009.

[21] V. de Silva and G. Carlsson. Topological estimation using witness complexes. *Proc. Sympos. Point-Based Graphics*, pages 157–166, 2004.

[22] T. Dey, S. Funke, and E. Ramos. Surface reconstruction in almost linear time under locally uniform sampling. In *Abstracts 17th European Workshop Comput. Geom*, pages 129–132. Citeseer, 2001.

[23] T. Dey, J. Giesen, E. Ramos, and B. Sadri. Critical points of the distance to an epsilon-sampling of a surface and flow-complex-based surface reconstruction. In *Proc. of the twenty-first annual symposium on Computational geometry*, page 227. ACM, 2005.

[24] T. Dey and S. Goswami. Provable surface reconstruction from noisy samples. *Computational Geometry: Theory and Applications*, 35(1-2):124–141, 2006.

[25] T. Dey and K. Li. Cut locus and topology from surface point data. In *Proceedings of the 25th annual symposium on Computational geometry*, pages 125–134. ACM New York, NY, USA, 2009.

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

[27] H. Edelsbrunner, D. Morozov, and V. Pascucci. Persistence-sensitive simplification functions on 2-manifolds. In *Proceedings of the twenty-second annual symposium on Computational geometry*, page 134. ACM, 2006.

[28] K. Grove. Critical point theory for distance functions. In *Proc. of Symposia in Pure Mathematics*, volume 54, pages 357–386, 1993.

[29] A. Lieutier. Any open bounded subset of R n has the same homotopy type as its medial axis. *Computer-Aided Design*, 36(11):1029–1046, 2004.

[30] D. Morozov. Homological Illusions of Persistence and Stability. *Ph.D. Dissertation, Duke University*, 2008.

[31] J. Munkres. *Elements of algebraic topology.* Perseus Books, 1993.

[32] P. Niyogi, S. Smale, and S. Weinberger. Finding the Homology of Submanifolds with High Confidence from Random Samples. *Discrete Computational Geometry*, 39(1-3):419–441, 2008.

[33] S. Oudot. Échantillonnage et maillage de surfaces avec garanties. *Ph.D. Dissertation, Ecole Polytechnique*, 2005.

APPENDIX

This appendix establishes the correctness of our first naive reconstruction algorithm. First, we establish the existence of cubical sets that are faithful reconstruction of shapes with positive μ-reach, using Corollary 3 in [7]:

LEMMA 10 (COROLLARY 3 IN [7]). *For* $d_N = \frac{1}{40N^3\lceil\sqrt{N}\rceil}$ *and for all compact sets* $\mathcal{A} \subset \mathbb{R}^N$ *with reach greater than* $\rho > 0$*, there exists a* $(d_N\rho)$*-cubical set* $\mathcal{X}$ *such that* $\mathcal{A} \subset \mathcal{X} \subset \mathcal{A}^\rho$ *and the inclusion maps* $\mathcal{A} \hookrightarrow \mathcal{X}$ *and* $\mathcal{X} \hookrightarrow \mathcal{A}^\rho$ *are homotopy equivalences.*

PROOF OF LEMMA 6. The proof of Lemma 6 consists in extending the above lemma to the situation where compact sets have a positive μ-reach with the constant $c_N = \frac{d_N}{2}$. Given a set $\mathcal{Y} \subset \mathbb{R}^N$, we denote respectively by $\overline{\mathcal{Y}}$ and $\mathcal{Y}^{\mathbf{c}}$ the closure and the complement of $\mathcal{Y}$. For any compact set $\mathcal{X} \subset \mathbb{R}^N$ and any real number $\rho > 0$, let $\mathcal{X}^{-\rho} = \overline{((\mathcal{X}^{\mathbf{c}})^\rho)^{\mathbf{c}}}$ and consider the set $\mathcal{B} = (\mathcal{A}^y)^{-\mu(y-x)}$. We know from [13] that the reach of $\mathcal{B}$ is greater than or equal to $\mu(y-x)$ and the inclusion maps corresponding to the sequence

$$\mathcal{A}^x \subset \mathcal{B} \subset \mathcal{A}^y$$

are homotopy equivalences. We can now apply Corollary 3 in [7] to the set $\mathcal{B}$ whose reach is greater than $\rho = \frac{\mu(y-x)}{2}$. This gives the existence of a $(c_N\mu(y-x))$-cubical set $\mathcal{X}$ such that:

$$\mathcal{B} \subset \mathcal{X} \subset \mathcal{B}^\rho$$

and the maps corresponding to inclusions are homotopy equivalences. Using $\mathcal{B}^\rho = ((\mathcal{A}^y)^{-2\rho})^\rho \subset \mathcal{A}^y$, we get the sequence of inclusions

$$\mathcal{A}^x \subset \mathcal{X} \subset \mathcal{A}^y.$$

in which the inclusion map $\mathcal{A}^x \hookrightarrow \mathcal{X}$ is a homotopy equivalence. By Lemma 2, $\mathcal{X}$ is a faithful reconstruction of $\mathcal{A}$. $\square$

Under precondition 3, next lemma states the existence of a cubical set which is a faithful reconstruction nested between two cubical sets that can be deduced from the sample (see Figure 4). Given a compact subset $\mathcal{Y} \subset \mathbb{R}^N$, we recall that $V_\varepsilon(\mathcal{Y})$ designates the union of ε-voxels that intersect $\mathcal{Y}$. The underlying space of the simplicial complex X is denoted $|X|$.

LEMMA 11. *Let* $\alpha, \eta > 0$ *and* $\mu \in (0,1]$ *be real numbers and let* $\mathcal{A}$ *and* $\mathcal{S}$ *be compact subsets of* $\mathbb{R}^N$ *such that* $d_H(\mathcal{S},\mathcal{A}) < \alpha < \frac{1}{4}(r_\mu(\mathcal{A}) - \eta)$*. Then, for:*

$$\varepsilon = \frac{\eta}{4\sqrt{N} + \frac{2}{c_N\mu}}, \quad l = \frac{\eta}{2} + \alpha, \quad k = \eta + 3\alpha - \varepsilon\sqrt{N},$$

there exists an ε*-cubical set* $\mathcal{X}$ *such that:*

$$\mathcal{A}^{\frac{\eta}{2}} \subset V_\varepsilon(\mathcal{S}^l) \subset \mathcal{X} \subset V_\varepsilon(\mathcal{S}^k) \subset \mathcal{A}^{\eta+4\alpha}$$

and the inclusion maps $\mathcal{A}^{\frac{\eta}{2}} \hookrightarrow \mathcal{X}$ *and* $\mathcal{X} \hookrightarrow \mathcal{A}^{4\alpha+\eta}$ *are homotopy equivalences. In particular,* $\mathcal{X}$ *is a faithful reconstruction of* $\mathcal{A}$*. Furthermore, if we have three nested simplicial complexes* $L \subset X \subset K$ *such that* $V_\varepsilon(\mathcal{S}^l) = |L|$, $\mathcal{X} = |X|$ *and* $V_\varepsilon(\mathcal{S}^k) = |K|$*, then* X *is a homological simplification of the pair* (K, L)*.*

PROOF. Note that for all compact set $\mathcal{Y} \subset \mathbb{R}^N$, we have $\mathcal{Y} \subset V_\varepsilon(\mathcal{Y}) \subset \mathcal{Y}^{\varepsilon\sqrt{N}}$. It follows that for all $t \geq 0$, we have the following sequence of inclusions:

$$\mathcal{A}^t \subset \mathcal{S}^{t+\alpha} \subset V_\varepsilon(\mathcal{S}^{t+\alpha}) \subset \mathcal{S}^{t+\alpha+\varepsilon\sqrt{N}} \subset \mathcal{A}^{t+2\alpha+\varepsilon\sqrt{N}}.$$

Applying this sequence twice, once for $t = \frac{\eta}{2}$ and once for $t = \eta + 2\alpha - \varepsilon\sqrt{N}$, we get that

$$\begin{aligned} \mathcal{A}^{\frac{\eta}{2}} &\subset V_\varepsilon(\mathcal{S}^{\frac{\eta}{2}+\alpha}) \subset \mathcal{A}^{\frac{\eta}{2}+2\alpha+\varepsilon\sqrt{N}} \\ &\subset \mathcal{A}^{\eta+2\alpha-\varepsilon\sqrt{N}} \subset V_\varepsilon(\mathcal{S}^{\eta+3\alpha-\varepsilon\sqrt{N}}) \subset \mathcal{A}^{\eta+4\alpha}. \end{aligned}$$

The value ε has been chosen precisely such that the parameters of the two offsets of $\mathcal{A}$ in the middle differ by $\frac{\varepsilon}{c_N\mu}$. Specifically, writing $x = \frac{\eta}{2}+2\alpha+\varepsilon\sqrt{N}$ and $y = \eta+2\alpha-\varepsilon\sqrt{N}$, we have $y-x = \frac{\varepsilon}{c_N\mu}$. Hence, applying Lemma 6 to $\mathcal{A}$, we get the existence of an ε-cubical set $\mathcal{X}$ such that $\mathcal{A}^x \subset \mathcal{X} \subset \mathcal{A}^y$ and the maps corresponding to the inclusions are homotopy equivalences. The first part of the lemma follows. For the second part, notice that we have the following sequence of homomorphisms induced by inclusion maps:

$$\mathbf{H}_p(\mathcal{A}^{\frac{\eta}{2}}) \to \mathbf{H}_p(|L|) \to \mathbf{H}_p(|X|) \to \mathbf{H}_p(|K|) \to \mathbf{H}_p(\mathcal{A}^{\eta+4\alpha}).$$

We will use the following observation. Consider two functions $i : E \to F$ and $j : F \to G$ such that the composition $j \circ i$ is bijective. Then, i is injective and j is surjective. Since the map $\mathbf{H}_p(\mathcal{A}^{\frac{\eta}{2}}) \to \mathbf{H}_p(|X|)$ induced on homology groups by inclusion is bijective, the map $\mathbf{H}_p(|L|) \to \mathbf{H}_p(|X|)$ is surjective using the observation. Similarly, the map $\mathbf{H}_p(|X|) \to \mathbf{H}_p(|K|)$ is injective and therefore X is a homological simplification of the pair (K, L). $\square$

LEMMA 12. *Let* x_1, x_2, x_3 *be real numbers and* $\mathcal{A}$ *a compact subset of* $\mathbb{R}^N$ *such that* $0 < x_1 < x_2 < x_3 < \mathrm{wfs}(\mathcal{A})$*. Let* $L \subset K$ *be two simplicial complexes such that:*

$$\mathcal{A}^{x_1} \subset |L| \subset \mathcal{A}^{x_2} \subset |K| \subset \mathcal{A}^{x_3}$$

If X *is a homological simplification of the pair* (K, L)*, then* $|X|$ *is a faithful homological reconstruction of* $\mathcal{A}$*.*

PROOF. Inclusion maps induce the following commutative diagram between homology groups:

$$\begin{array}{ccccccc} & & & \mathbf{H}_p(\mathcal{A}^{x_2}) & & & \\ & & \nearrow & & \searrow & & \\ \mathbf{H}_p(\mathcal{A}^{x_1}) & \longrightarrow & \mathbf{H}_p(|L|) & & \mathbf{H}_p(|K|) & \longrightarrow & \mathbf{H}_p(\mathcal{A}^{x_3}) \\ & & \searrow & & \nearrow & & \\ & & & \mathbf{H}_p(|X|) & & & \end{array}$$

From Lemma 1, the maps $\mathbf{H}_p(\mathcal{A}^{x_1}) \to \mathbf{H}_p(\mathcal{A}^{x_2}) \to \mathbf{H}_p(\mathcal{A}^{x_3})$ are isomorphisms. Since we take coefficients in a field, homology groups are vector spaces. Since the map $\mathbf{H}_p(\mathcal{A}^{x_1}) \to \mathbf{H}_p(\mathcal{A}^{x_2})$ induced by inclusion is bijective, the map $\mathbf{H}_p(|L|) \to \mathbf{H}_p(\mathcal{A}^{x_2})$ is surjective. Since $\mathbf{H}_p(\mathcal{A}^{x_2}) \to \mathbf{H}_p(\mathcal{A}^{x_3})$ is bijective, $\mathbf{H}_p(\mathcal{A}^{x_2}) \to \mathbf{H}_p(|K|)$ is injective. By an argument similar to the one used in the proof of Lemma 7, we deduce that the dimensions of $\mathbf{H}_p(\mathcal{A}^{x_2})$ and $\mathbf{H}_p(|X|)$ are equal to the rank of $\mathbf{H}_p(|L|) \to \mathbf{H}_p(|K|)$ and therefore $\dim \mathbf{H}_p(|X|) = \dim \mathbf{H}_p(\mathcal{A}^{x_i})$ for all $1 \leq i \leq 3$. Since $\mathbf{H}_p(\mathcal{A}^{x_1}) \to \mathbf{H}_p(\mathcal{A}^{x_3})$ is bijective, $\mathbf{H}_p(\mathcal{A}^{x_1}) \to \mathbf{H}_p(|X|)$ is injective. Because its domain and image have same dimension, it follows that $\mathbf{H}_p(\mathcal{A}^{x_1}) \to \mathbf{H}_p(|X|)$ is an isomorphism. Similarly $\mathbf{H}_p(|X|) \to \mathbf{H}_p(\mathcal{A}^{x_3})$ is an isomorphism and $|X|$ is a faithful homological reconstruction of $\mathcal{A}$. $\square$