



HAL
open science

Homological Reconstruction and Simplification in R3

Dominique Attali, Ulrich Bauer, Olivier Devillers, Marc Glisse, André Lieutier

► **To cite this version:**

Dominique Attali, Ulrich Bauer, Olivier Devillers, Marc Glisse, André Lieutier. Homological Reconstruction and Simplification in R3. Computational Geometry, 2015, 48 (8), pp.606-621. 10.1016/j.comgeo.2014.08.010 . hal-01132440

HAL Id: hal-01132440

<https://hal.science/hal-01132440v1>

Submitted on 21 Apr 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Homological Reconstruction and Simplification in \mathbb{R}^3

Dominique Attali

Gipsa-lab, Saint Martin d'Hères, France

Ulrich Bauer*

IST Austria, Klosterneuburg, Austria

Olivier Devillers

INRIA Sophia Antipolis – Méditerranée, Sophia Antipolis, France

Marc Glisse

INRIA Saclay – Île-de-France, Orsay, France

André Lieutier

Dassault Système, Aix-en-Provence, France

Abstract

We consider the problem of deciding whether the persistent homology group of a simplicial pair (K, L) can be realized as the homology $H_*(X)$ of some complex X with $L \subset X \subset K$. We show that this problem is NP-complete even if K is embedded in \mathbb{R}^3 .

As a consequence, we show that it is NP-hard to simplify level and sublevel sets of scalar functions on \mathbb{S}^3 within a given tolerance constraint. This problem has relevance to the visualization of medical images by isosurfaces. We also show an implication to the theory of well groups of scalar functions: not every well group can be realized by some level set, and deciding whether a well group can be realized is NP-hard.

Keywords: NP-hard problems, homology, persistence

2010 MSC: 55U10

*Corresponding author.

Email addresses: mail@ulrich-bauer.org (Ulrich Bauer), Andre.LIEUTIER@3ds.com (André Lieutier)

URL: <http://www.gipsa-lab.grenoble-inp.fr/> (Dominique Attali),
<http://ulrich-bauer.org> (Ulrich Bauer),
<http://www-sop.inria.fr/members/Olivier.Devillers/> (Olivier Devillers),
<http://geometrica.saclay.inria.fr/team/Marc.Glisse/> (Marc Glisse)

1. Introduction

In this paper, we establish NP-completeness of a variety of related problems that ask for an object in \mathbb{R}^3 with certain prescribed topological constraints.

In the most basic setting, we have a point cloud in \mathbb{R}^d that samples a shape and want to retrieve information on the sampled shape. There exists a whole spectrum of possibilities regarding the type of sought information. At the coarsest level, we can content ourselves with the homology groups which record the “holes” of a given dimension, hereafter referred to as *homological features* (connected components, cycles, cavities and so on). At a finer level, we may be interested in building an approximation of the shape, reflecting as accurately as possible both its geometry and topology. The standard way is to construct a *simplicial complex* using the data points as vertices, such as for instance the α -complex, the Rips complex or the Čech complex [13, 12]. All three constructions have in common to depend upon a scale parameter α and to get bigger as α increases. In the ideal case, we expect the complex to have the right homology for some suitable value of α [21, 6, 7, 2]. Unfortunately, depending on the sampling, it may happen that such a value of α does not exist. Nonetheless, we might still be able to infer the true homology of the shape hidden in the noisy data using persistent homology [15, 10, 8]. Given two scale parameters α_1 and α_2 , the *persistent homology groups* record the homological features that persist from α_1 to α_2 . Under very weak hypotheses, we know that the persistent homology is precisely that of the sampled shape [10, 5]. The persistent homology can be computed efficiently (i.e., in polynomial time).

A natural question is then to ask for a complex that carries the persistent homology: given a complex K and a subcomplex L , can we find a subcomplex of K that contains L and whose homological features are precisely those common to L and K ? Our answer is that sometimes we cannot, and deciding whether we can is NP-complete. This answer was first given in the general case by Attali and Lieutier [1], who posed the restriction to complexes embedded in \mathbb{R}^3 as an open problem. We resolve this problem by proving NP-completeness even for complexes embedded in \mathbb{R}^3 .

The above problem concentrates on building a complex whose homology matches perfectly the persistent homology of L into K : all the homological noise has been removed. We call such an object a *homological reconstruction*. However, when it does not exist, it is still relevant to look for a complex nested between L and K and whose homology is as close as possible to the persistent homology of L into K : as much noise as possible has been removed. We call such a complex a *homological simplification* and prove that finding one is also an NP-hard problem.

In the field of visualization and image analysis, another common setting consists in describing a shape through a continuous function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ instead of a point cloud in \mathbb{R}^d . For instance, a medical image may be a collection of density measurements over a grid of 3D points and is best modeled as a continuous map over a certain domain of \mathbb{R}^3 . In the ideal case, the shape is a sublevel set of the function, $f^{-1}(-\infty, t]$. Unfortunately, noise can plague the data. As the parameter t increases, sublevel sets inflate and we can track the evolution of their homology. Features that appear and disappear quickly are considered topological noise, and we can consider the common features of two sublevel sets as those of a denoised sublevel set. The question now becomes: can we find another cleaner function, close enough to the original one, whose sublevel set has the

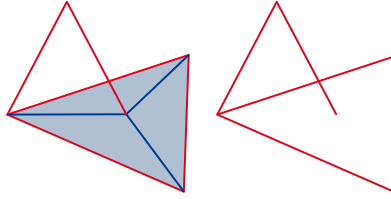


Figure 1: Example of a simplicial pair (K, L) embedded in \mathbb{R}^3 that has no homological reconstruction.

denoised homology, i.e., a *sublevel set reconstruction*? The corresponding optimization problem asks for a *sublevel set simplification*, i.e., a function close to the original one that minimizes the number of homological features of the sublevel set. We show that these two problems are the equivalents in the functional setting to the homological reconstruction and simplification of simplicial pairs described above.

Often, one is also interested in the homology of a level set, $f^{-1}(t)$. We show how it can be related to the (persistent) homology of sublevel sets, and consider the corresponding *level set reconstruction/simplification* problems.

Further in this direction, Edelsbrunner et al. introduced the well group [16, 3] as a denoised version of the homology group of a level set. Again, we can ask whether one can find a *realization* of the well group, i.e., a cleaner function whose level set has the same homology as the well group?

We shall see in this paper that all of these related problems are NP-hard, as a consequence of the NP-completeness of the homological reconstruction problem.

1.1. Background and notations

We are only concerned with topological spaces that are triangulable by a finite simplicial complex, so simplicial and singular homology are isomorphic and we make no distinction between the two. In particular, we use the simplicial versions of the Excision and Mayer-Vietoris sequence theorems, which have less restrictive assumptions than their singular counterparts. If K is an abstract simplicial complex, we denote by K its geometric realization. Throughout this article, we consider homology with coefficients in an arbitrary field \mathbb{F} , so the homology groups are finite-dimensional \mathbb{F} -vector spaces and there is no torsion. Note that for simplicial complexes K embedded in \mathbb{R}^3 , this is in fact not a restriction, since due to the absence of torsion in \mathbb{R}^3 the Betti numbers are independent of the choice of coefficients (see, e.g., [17, §3.3]).

Given a topological space \mathcal{K} , we write $H_*(\mathcal{K}) = \bigoplus_i H_i(\mathcal{K})$ for the direct sum of homology groups in all dimensions, and $\beta(\mathcal{K}) = \sum_{i \geq 0} \beta_i(\mathcal{K})$ for the *total Betti number*. If $(\mathcal{K}, \mathcal{L})$ is a pair of topological spaces $\mathcal{L} \subset \mathcal{K}$, the inclusion $\mathcal{L} \hookrightarrow \mathcal{K}$ induces a homomorphism $H_*(\mathcal{L}) \rightarrow H_*(\mathcal{K})$, which is denoted by $H_*(\mathcal{L} \hookrightarrow \mathcal{K})$. The rank of this map is the *persistent Betti number* of the inclusion $\mathcal{L} \hookrightarrow \mathcal{K}$ and is denoted by $\beta(\mathcal{L} \hookrightarrow \mathcal{K}) = \text{rank } H_*(\mathcal{L} \hookrightarrow \mathcal{K})$; the image $\text{im } H_*(\mathcal{L} \hookrightarrow \mathcal{K})$ is a *persistent homology group*. If (K, L) is a *simplicial pair*, that is, a pair of simplicial complexes such that $L \subset K$, then the persistent Betti number $\beta(L \hookrightarrow K)$ can be computed in time cubic in

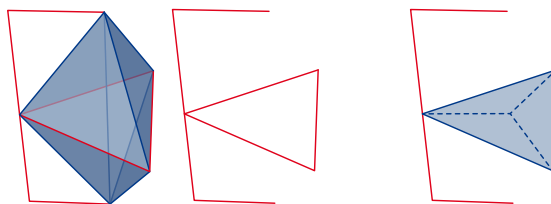


Figure 2: Example of a simplicial pair (left) having a homological reconstruction as a subspace (right), but not as a subcomplex. The simplicial complex on the left contains three tetrahedra sharing an edge and whose union forms a triangular bipyramid.

the number of simplices in K [15]. This cubic complexity can be improved to matrix multiplication time [19].

A *piecewise linear function* on a topological space \mathcal{K} is a continuous function $f : \mathcal{K} \rightarrow \mathbb{R}$ such that there exists a finite triangulation of \mathcal{K} on which f is simplexwise linear. Note that a *simplexwise linear function* must be linear on each simplex of the *given* triangulation, while a *piecewise linear function* is linear on each simplex of *some* arbitrary triangulation.

2. Reconstruction and simplification of simplicial pairs

In this section, we consider a simplicial pair and define the homological reconstruction problem and the homological simplification problem. We prove that both problems are NP-hard when the simplicial pair is embedded in \mathbb{R}^3 . We start with a simple lemma:

Lemma 1. *Consider a triple of topological spaces $\mathcal{L} \subset \mathcal{X} \subset \mathcal{K}$ with finite Betti numbers. Then*

$$\beta(\mathcal{X}) \geq \beta(\mathcal{L} \hookrightarrow \mathcal{K}).$$

Proof. This is a consequence of the fact that whenever we consider two linear maps $j : U \rightarrow V$ and $i : V \rightarrow W$ between finite dimensional vector spaces, then $\dim V \geq \text{rank } j \geq \text{rank } i \circ j$. \square

This property suggests the following definition:

Definition 1. *Consider a triple of topological spaces $\mathcal{L} \subset \mathcal{X} \subset \mathcal{K}$ with finite Betti numbers. Then \mathcal{X} is called a homological reconstruction of $(\mathcal{K}, \mathcal{L})$ if $\beta(\mathcal{X}) = \beta(\mathcal{L} \hookrightarrow \mathcal{K})$. Moreover, \mathcal{X} is called a homological p -reconstruction of $(\mathcal{K}, \mathcal{L})$ if $\beta_p(\mathcal{X}) = \beta_p(\mathcal{L} \hookrightarrow \mathcal{K})$.*

We will often omit “homological” since there is no ambiguity in this paper. An equivalent condition for \mathcal{X} being a reconstruction is that $H_*(\mathcal{L} \hookrightarrow \mathcal{X})$ is surjective and $H_*(\mathcal{X} \hookrightarrow \mathcal{K})$ is injective, as defined in [1]. Not every pair $(\mathcal{K}, \mathcal{L})$ admits a reconstruction; a simple counterexample is shown in Fig. 1. The use of topological spaces in the definition (as opposed to simplicial complexes) is motivated by the following observation. Let (K, L) be a simplicial pair. Then there might be a reconstruction of (K, L) , but not as a subcomplex of K . An example is shown in Fig. 2. A reconstruction that is a

subcomplex is called a *subcomplex reconstruction*. To emphasize the distinction to this case, we sometimes use the term *subspace reconstruction* to emphasize that the reconstruction is only required to be a subspace, not necessarily a subcomplex.

2.1. Homological reconstruction is NP-hard

We now focus our attention on spaces that are geometric realizations of finite simplicial complexes embedded in \mathbb{R}^3 .

Theorem 1. *The homological reconstruction problem is NP-hard: Given as input a simplicial pair (K, L) embedded in \mathbb{R}^3 , decide whether there exists a subspace reconstruction \mathcal{X} of (K, L) . The problem is NP-complete if \mathcal{X} is required to be a subcomplex.*

This section is devoted to the proof of Theorem 1 by reduction from 3-SAT. Recall that a Boolean formula Φ is in 3-CNF if it is a conjunction of several clauses, each of which is a disjunction of three literals, a literal being either a variable or its negation. Given a 3-CNF formula Φ , we construct a simplicial pair (K_Φ, L_Φ) embedded in \mathbb{R}^3 and prove that (K_Φ, L_Φ) has a reconstruction (as a subcomplex of K_Φ) if and only if Φ has a satisfying assignment (see Lemmas 2 and 3 below).

For this, we associate to the 3-CNF formula Φ a simplicial pair (K_Φ, L_Φ) with trivial persistent homology. Equivalently, any reconstruction \mathcal{X} of (K_Φ, L_Φ) has trivial homology, i.e.,

$$\beta_d(\mathcal{X}) = \beta_d(L_\Phi \hookrightarrow K_\Phi) = \begin{cases} 1 & \text{if } d = 0, \\ 0 & \text{otherwise.} \end{cases}$$

This means that \mathcal{X} has a single connected component, no loops, and no cavities. \mathcal{X} has to fill all loops or cavities in L_Φ and has to connect the different connected components of L_Φ by adding to L_Φ portions of K_Φ without creating any new loops or cavities.

The variable gadget. The variable gadget is a simplicial pair (V_i, W_i) as depicted in Fig. 3, top. The simplicial complex V_i contains 4 edges forming a cycle. The two bold edges do not belong to W_i . One of the bold edges will be called True_i and the other one will be called False_i . The key property of this construction is that any reconstruction of the pair (V_i, W_i) cannot contain both edges True_i and False_i , for otherwise they would form a 1-cycle with the remaining edges. This property will allow us to match the presence of the edge True_i to a *true* assignment of the variable v_i .

The clause gadget. The clause gadget is a simplicial pair (C_j, D_j) as depicted in Fig. 3, bottom. The simplicial complex D_j contains a cycle ABCDE. The cycle is closed with two surfaces in C_j (thereafter referred to as the lower hemisphere and the disk) thereby creating a cavity. Furthermore, the complex D_j contains an arc that ends inside the disk. Whenever we fill the cycle ABCDE with the disk, this connects the two endpoints of the arc, thus creating a new cycle, which we close twice in C_j by a left hemisphere and a right hemisphere. Consider one bold edge in the interior of each hemisphere, which is where the clause gadget will connect to the variable gadgets.

The key property of this clause gadget is that at least one of the 3 bold edges must be present in any reconstruction \mathcal{X} of the pair (C_j, D_j) . Indeed, the cycle ABCDE in \mathcal{X} must be filled up. If it is filled by the lower hemisphere, we are done. If it is filled by

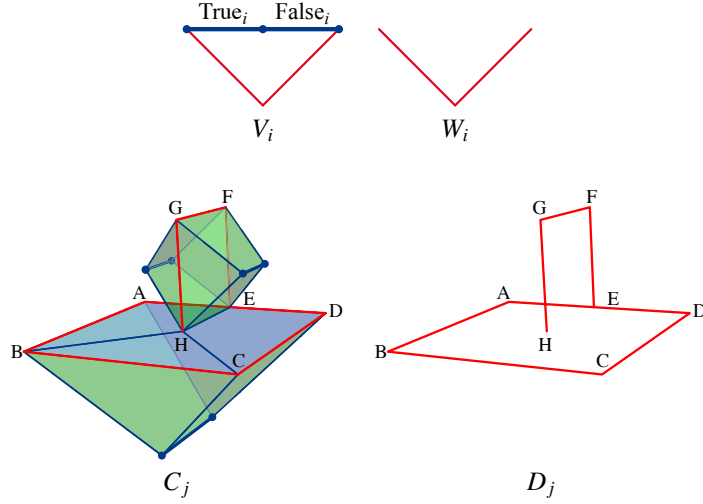


Figure 3: Variable (top) and clause (bottom) gadgets for the reduction of homological reconstruction to 3-SAT.

the disk, we have a new cycle EFGH in \mathcal{X} which in turn must be killed either by the left or by the right hemisphere. In any case, \mathcal{X} contains at least one of the hemispheres and thus one of the three bold edges.

Correspondence with a formula. Given a 3-CNF formula Φ with n clauses c_1, \dots, c_n and m variables $v_1 \dots, v_m$, we construct a 2-dimensional pair (K_Φ, L_Φ) as follows. For each variable v_i we take a copy (V_i, W_i) of the variable gadget. For each clause c_j , we take a copy (C_j, D_j) of the clause gadget; for each literal ev_i of c_j , we identify one of the bold edges of C_j to False_i if e is a negation and True_i otherwise. See Fig. 4 for an example.

First notice that $\beta_2(L_\Phi) = 0$ (i.e., L_Φ has no cavities). Second, we can assume that $\beta_0(K_\Phi) = 1$ (i.e., K_Φ is connected). Indeed, if K_Φ is disconnected, it means that the 3-SAT problem (and the reconstruction problem) can be decomposed into 2 independent subproblems with disjoint sets of variables, which can be solved separately. Last, $\beta_1(L_\Phi \hookrightarrow K_\Phi) = 0$ (i.e., the cycles in L_Φ are boundaries in K_Φ). Indeed, the only 1-cycles in L_Φ are the 1-cycles $ABCDE$ in each D_j , and they are filled in K_Φ . This means that we are looking for a reconstruction with trivial homology.

From a reconstruction to a satisfying assignment. Let \mathcal{X} be a homological reconstruction of the pair (K_Φ, L_Φ) . We do not assume that \mathcal{X} is the geometric realization of some subcomplex of K_Φ . Assign to each variable v_i the value *true* if the edge True_i is entirely contained in \mathcal{X} , and *false* otherwise. For each clause gadget (C_j, D_j) , at least one bold edge is contained in \mathcal{X} . If this edge corresponds to a positive literal v_i , this means that True_i is in \mathcal{X} , v_i is *true* and the clause is satisfied. If the edge corresponds to a negative literal $\neg v_i$, this implies that False_i is in \mathcal{X} . True_i is thus not in \mathcal{X} , so v_i was assigned *false* and the clause is satisfied. We have thus shown that the assignment of the variables makes the formula evaluate to *true*:

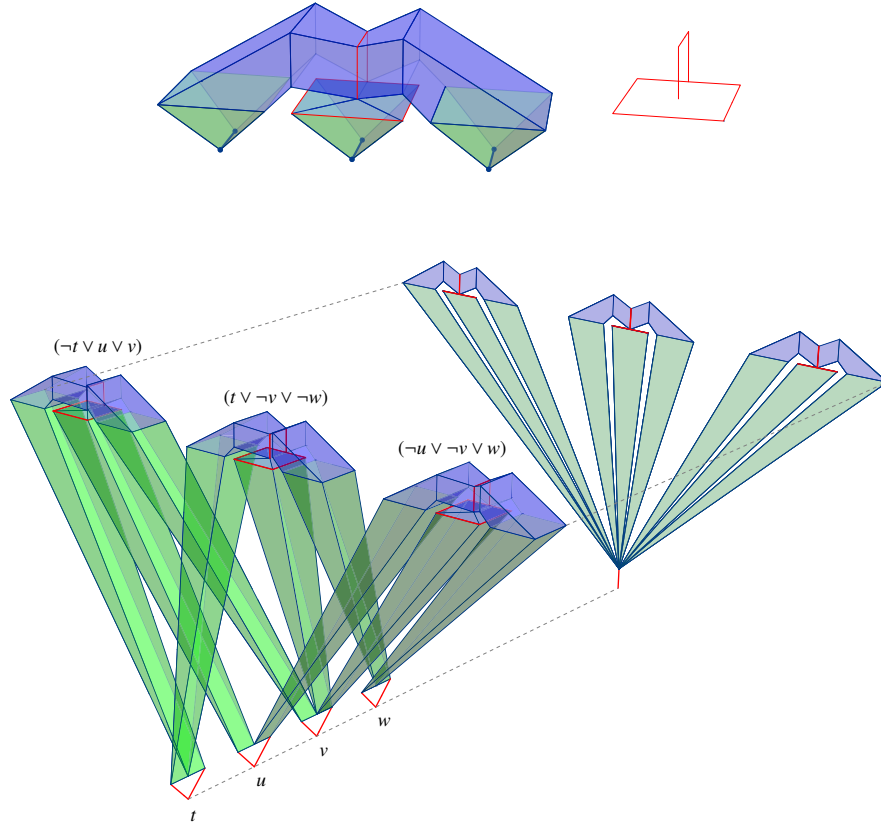


Figure 4: Embedding of the clause gadget with aligned hemispheres (top), and the simplicial complex K_Φ generated in the reduction from the 3-SAT instance $(\neg t \wedge u \wedge v) \vee (t \wedge \neg v \wedge \neg w) \vee (\neg u \wedge \neg v \wedge w)$ (bottom left), with parallel projection (bottom right) orthogonal to the alignment axis of the variable gadgets.

Lemma 2. *If (K_Φ, L_Φ) has a homological reconstruction, then Φ has a satisfying assignment.*

From a satisfying assignment to a reconstruction. Given a satisfying assignment for the formula Φ , we construct a subcomplex reconstruction X of (K_Φ, L_Φ) . We start with $X = L_\Phi$ and add to X a selected set of simplices from K_Φ . For each clause c_j , we pick one literal that evaluates to *true* and close the cycle in the clause gadget complex D_j correspondingly. If the literal corresponds to the bold edge of the lower hemisphere, we add this hemisphere. Otherwise, we add the disk and the hemisphere that contains the bold edge corresponding to the selected literal.

The only 2-cycles in K_Φ are in the clause gadgets. As we did not create any 2-cycle in X , it follows that $\beta_2(X) = 0$. By construction, filling the clause gadgets never introduced both True_i and False_i in X . Indeed, it could only introduce True_i if v_i was assigned the value *true* and False_i if it was assigned the value *false*. Therefore, the 1-cycle in

the variable gadgets do not appear in X . Also, for each clause gadget, we filled the ABCDE 1-cycle, and whenever we created an extra EFGH 1-cycle by adding the disk, we immediately filled it with the left or right hemisphere. Now we only need to check that the construction did not create any “non-local” 1-cycles. Since for each clause we have only used one of the literals which evaluate to *true*, the only contact a clause gadget in X has with the rest of X is through a single bold edge, and the clause gadget can be collapsed to that edge. After collapsing all clause gadgets, all that remains are disconnected variable gadgets with at most 3 edges each, and so $\beta_1(X) = 0$. We finally add to X just enough edges from K_Φ so that it becomes connected, without creating any extra cycles in the process. This is possible since we assumed that K_Φ is connected. Thus we have $\beta_0(X) = 1$. We conclude:

Lemma 3. *If Φ has a satisfying assignment, then (K_Φ, L_Φ) has a subcomplex reconstruction.*

We are now in a position to prove Theorem 1.

Proof of Theorem 1. First we show that the homological reconstruction problem is NP-hard. We proceed by reduction from 3-SAT. Let $(K, L) = (K_\Phi, L_\Phi)$ be a simplicial pair defined by a 3-SAT instance Φ . We show that the following propositions are equivalent:

- (a) (K, L) has a subspace reconstruction;
- (b) Φ has a satisfying assignment;
- (c) (K, L) has a subcomplex reconstruction.

The implication (a) \implies (b) is shown in Lemma 2; (b) \implies (c) is shown in Lemma 3. Finally, (c) \implies (a) is trivial.

The pair (K, L) can be constructed from Φ in time polynomial in the size of Φ . Together with the equivalence (a) \iff (b), this establishes NP-hardness of the homological reconstruction problem.

The equivalence (b) \iff (c) also yields NP-hardness of the subcomplex reconstruction problem. Moreover, given a subcomplex X as a polynomial size certificate, we can decide in polynomial time whether X is a reconstruction of (K, L) . Thus the problem is also in NP and hence NP-complete. \square

Embedding. Later, we have to consider not only an embedding of K_Φ , but also a triangulation of its complement. The following fact will be useful:

Lemma 4. *There is a triangulation of \mathbb{S}^3 with size polynomial in the size of K_Φ and having K_Φ as a subcomplex.*

Proof. First, referring to Fig. 4, it is clear that K_Φ can be embedded in \mathbb{R}^3 . Indeed, we can align the clause gadgets and the variable gadgets along two lines parallel to the coordinate axes and make each clause gadget look like a small body with three long tentacles that connect to the variable gadgets. Due to the way the variable and clause gadgets are aligned in the construction along skew axes, the tentacles do not intersect in their interior.

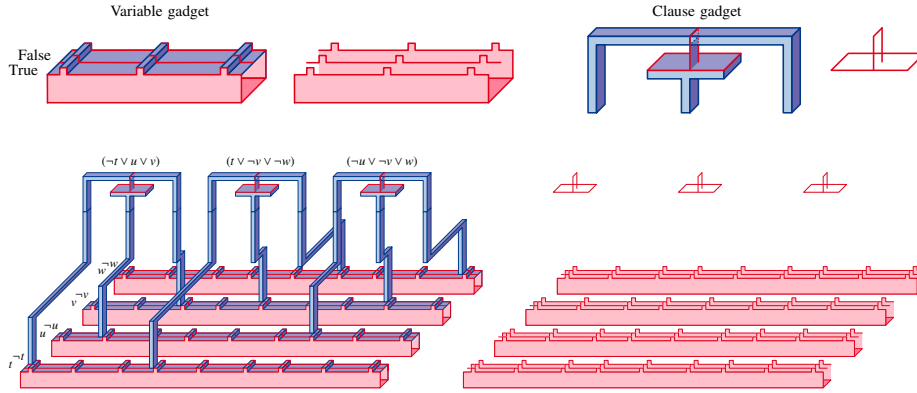


Figure 5: Example of 3-SAT reduction using a 3D grid embedding.

We can subdivide the space by first projecting K_Φ onto a plane orthogonal to the line carrying the variable gadgets. We get a polygonal region whose complement can easily be triangulated inside a bounding box without adding any new vertex and thus adding a linear number of edges. Extending each triangle in the direction of the projection, we get a collection of tubes, one for each triangle. The tubes can easily be triangulated while respecting K_Φ to obtain a polynomial size triangulation of a bounding box of the construction, which can trivially be extended to a polynomial size triangulation of \mathbb{S}^3 . \square

We want to remark that a similar construction can be realized even if we restrict edges and faces of L and K to be edges and faces of a 3D grid (see Fig. 5). This means that a variant of Theorem 1 can also be shown for cubical complexes arising from 3D image data.

Corollary 1. *The homological simplification problem is NP-hard: Given as input a simplicial pair (K, L) embedded in \mathbb{R}^3 , find a complex X minimizing $\beta(X)$ subject to $L \subset X \subset K$.*

Proof. We use a reduction from the subcomplex reconstruction problem. To determine if a subcomplex reconstruction exists, we can first find a complex X minimizing $\beta(X)$ subject to $L \subset X \subset K$. We then only need to check if its Betti number matches the lower bound $\beta(L \hookrightarrow K)$. \square

3. Reconstruction and simplification of level and sublevel sets

In this section, we consider a real-valued simplexwise linear function defined on a simplicial complex embedded in \mathbb{R}^3 and establish the NP-hardness of problems that ask for a nearby function with a simplified sublevel set (Section 3.1) and a simplified level set (Section 3.3).

Given a real-valued function f , we write F_t for the t -level set $f^{-1}(t)$, $F_{\leq t}$ for the (closed) t -sublevel set $f^{-1}((-\infty, t])$, and $F_{< t}$ for the open t -sublevel set $f^{-1}((-\infty, t))$. In

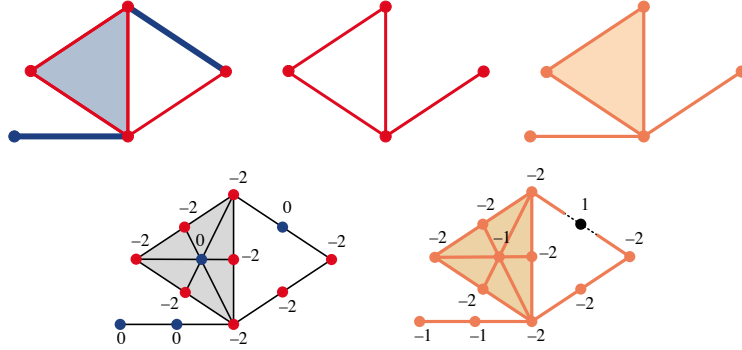


Figure 6: Top: A simplicial pair (K, L) and a homological reconstruction of (K, L) as a subcomplex. Bottom: values of f (left) and g (right) at the vertices of the barycentric subdivision $\text{sd } K$, as used in the proof of Theorem 2.

this paper we shall only consider real-valued piecewise linear functions. Note that level and sublevel sets of a simplexwise linear function on a simplicial complex K are not necessarily subcomplexes of K , but subcomplexes of an appropriate subdivision of K . Moreover, we have the following property:

Proposition 1 (Kühnel [18], Morozov [20]). *Let f be a simplexwise linear function on a simplicial complex K . Let $K(t)$ be the induced subcomplex of K on $\{v \in \text{vert } K : f(v) \leq t\}$. Then $K(t)$ is homotopy equivalent to the sublevel set $F_{\leq t}$. If $t \neq f(v)$ for all $v \in \text{vert } K$, then $K(t)$ is also homotopy equivalent to the open sublevel set $F_{< t}$.*

Definition 2. *Let f, g be piecewise linear functions and consider real parameters t and δ . The function g is called a sublevel set (t, δ) -reconstruction of f if $\|g - f\|_{\infty} \leq \delta$ and $G_{\leq t}$ is a reconstruction of the pair $(F_{\leq t+\delta}, F_{\leq t-\delta})$, i.e.,*

$$\beta(G_{\leq t}) = \beta(F_{\leq t-\delta} \leftrightarrow F_{\leq t+\delta}).$$

Note that

$$F_{\leq t-\delta} \subseteq G_{\leq t} \subseteq F_{\leq t+\delta},$$

so that

$$\beta(G_{\leq t}) \geq \beta(F_{\leq t-\delta} \leftrightarrow F_{\leq t+\delta}).$$

A sublevel set (t, δ) -reconstruction is thus also a minimizer of $\beta(G_{\leq t})$ subject to $\|g - f\|_{\infty} \leq \delta$.

3.1. Sublevel set reconstruction is NP-hard

Theorem 2. *The sublevel set reconstruction problem is NP-hard: Given as input a simplexwise linear function f on a simplicial complex embedded in \mathbb{R}^3 and parameters t and δ , decide whether there exists a sublevel set (t, δ) -reconstruction g of f .*

Proof. It suffices to establish the theorem for $t = 0$ and $\delta = 1$. We proceed by reduction from 3-SAT using the results of the previous section. Let $(K, L) = (K_\Phi, L_\Phi)$ be a simplicial pair defined by a 3-SAT instance Φ , as described in the proof of Theorem 1. We construct an instance of the level set simplification problem by defining a simplexwise linear function $f : \text{sd } K \rightarrow \mathbb{R}$ on the barycentric subdivision of K ; see Figure 6. Recall that the *barycentric subdivision* (or *derived subdivision*) of a simplicial complex K is the *order complex* of the face relation, i.e., the abstract simplicial complex $\text{sd } K$ whose vertices are the simplices of K and whose simplices are the totally ordered subsets of K with regard to the face relation. We define f via its values on the vertices of $\text{sd } K$. Using the fact that a vertex σ of $\text{sd } K$ is a simplex of K , we let

$$f : \sigma \mapsto \begin{cases} -2 & \text{if } \sigma \in L, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Note that for every function g with $\|g - f\|_\infty \leq 1$, the 0-sublevel set $G_{\leq 0}$ contains L and is contained in K . We show that the following propositions are equivalent to propositions (a)–(c) in the proof of Theorem 1:

- (d) f has a simplexwise linear sublevel set $(0, 1)$ -reconstruction g .
- (e) f has a sublevel set $(0, 1)$ -reconstruction g .

To show (c) \implies (d), we define a simplexwise linear function g on $\text{sd } K$ by its values on the vertices of $\text{sd } K$ (the simplices of K); see Figure 6:

$$g : \sigma \mapsto \begin{cases} -2 & \text{if } \sigma \in L, \\ -1 & \text{if } \sigma \in X \setminus L, \\ 1 & \text{if } \sigma \in K \setminus X. \end{cases} \quad (2)$$

We have $\|g - f\|_\infty = 1$. By Proposition 1, the sublevel set $G_{\leq 0}$ is homotopy equivalent to $|X|$ and hence is a reconstruction of the pair

$$(K, L) \simeq (F_{\leq 1}, F_{\leq -1}).$$

Finally, (d) \implies (e) is trivial and (e) \implies (a) follows directly with $G_{\leq 0}$ as a reconstruction of $(F_{\leq 1}, F_{\leq -1}) \simeq (K, L)$.

The function f can be constructed from the 3-SAT instance Φ in polynomial time. Together with the equivalence (b) \iff (e), this establishes NP-hardness of the sublevel set reconstruction problem.

The equivalence (b) \iff (d) also yields NP-hardness of the sublevel set reconstruction problem restricted to simplexwise linear functions $\text{sd } K \rightarrow \mathbb{R}$. \square

Theorem 3. *The sublevel set reconstruction problem is NP-complete if the reconstruction is required to be simplexwise linear on the same complex.*

Proof. By Theorem 2, it is sufficient to show that the problem is in NP, i.e. every “yes” instance $f : K \rightarrow \mathbb{R}$ has certificate with size polynomial in the size of K and f . Again, it suffices to establish the theorem for $t = 0$ and $\delta = 1$. We show that there is

a simplexwise linear $(0, 1)$ -reconstruction g iff there is a subset of vertices S whose induced subcomplex K_S is a reconstruction of $(K(1), K(-1))$, where $K(t)$ is the induced subcomplex of K on $\{v \in \text{vert } K : f(v) \leq t\}$ as in Proposition 1.

The subset of vertices v with $g(v) \leq 0$ induces a subcomplex that is homotopy equivalent to the sublevel set $G_{\leq 0}$, by Proposition 1. Vice versa, let S be a subset of vertices such that the induced subcomplex K_S is a reconstruction of $(K(1), K(-1))$. In particular, for all $v \in S$ we have $f(v) \leq 1$, and for all $v \notin S$ we have $f(v) > -1$. Define a simplexwise linear function by the vertex values

$$h : v \mapsto \begin{cases} f(v) - 1 & \text{if } v \in S, \\ f(v) + 1 & \text{if } v \notin S, \end{cases}$$

and note that for each vertex v , $h(v) \leq 0$ if and only if $v \in S$. By Proposition 1, the sublevel set $H_{\leq 0}$ is homotopy equivalent to the induced subcomplex K_S and hence is a reconstruction of the pair $(F_{\leq 1}, F_{\leq -1}) \simeq (K(1), K(-1))$. We conclude that h is a $(0, 1)$ -reconstruction of f .

Given a subset S as a polynomial size certificate, by computing and comparing $\beta(K_S)$ and $\beta(K(-1) \hookrightarrow K(1))$ we can verify in polynomial time the existence of a sublevel set $(0, 1)$ -reconstruction of f . Thus the problem is also in NP and hence NP-complete. \square

Corollary 2. *The sublevel set simplification problem is NP-hard: Given as input a simplexwise linear function f on a simplicial complex embedded in \mathbb{R}^3 and parameters t and δ , find a simplexwise linear function g minimizing $\beta(G_{\leq t})$ subject to $\|g - f\|_\infty \leq \delta$.*

Proof. We use a reduction from the sublevel set reconstruction problem. To determine if f has a sublevel set (t, δ) -reconstruction, we can first find a simplexwise linear minimizer of $\beta(G_{\leq t})$. We then only need to check if $\beta(G_{\leq t})$ matches the lower bound

$$\beta(F_{\leq t-\delta} \hookrightarrow F_{\leq t+\delta}),$$

which can be done in time polynomial in the size of K . \square

3.2. Betti numbers of level and sublevel sets

The Betti numbers of level and sublevel sets are related by the following formula:

Lemma 5. *Let f be a piecewise linear function on \mathbb{S}^n , $n > 1$, and let t be in the interior of the image of f , $t \in \text{int}(\text{im } f)$. Then*

$$\beta_d(F_t) = \beta_d(F_{\leq t}) + \beta_{n-d-1}(F_{< t}).$$

Proof. First recall that $F_{\leq t}$, $F_{\geq t}$, and $F_{\geq t}$ are subcomplexes of an appropriate triangulation of \mathbb{S}^n , so we can apply the simplicial version of the Mayer-Vietoris theorem [22, §4.6]. By exactness of the Mayer-Vietoris sequence for \mathbb{S}^n , $F_{\leq t}$, and $F_{\geq t}$, we have [14]

$$\beta_d(F_t) = \beta_d(F_{\leq t}) + \beta_d(F_{\geq t}) + \begin{cases} -1 & \text{if } d = 0, \\ 1 & \text{if } d = n - 1, \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

By Alexander duality [17, §3.3], the duality of homology and cohomology with field coefficients resulting from the universal coefficient theorem [17, §3.1], and isomorphism of dual finite-dimensional vector spaces, we have

$$\widetilde{H}_d(F_{\geq t}) \cong \widetilde{H}^{n-d-1}(F_{< t}) \cong \text{Hom}(\widetilde{H}_{n-d-1}(F_{< t}), \mathbb{F}) \cong \widetilde{H}_{n-d-1}(F_{< t}),$$

where \widetilde{H}_d denotes the d th reduced homology group and $\text{Hom}(\widetilde{H}_{n-d-1}(F_{< t}), \mathbb{F})$ is the dual vector space of $\widetilde{H}_{n-d-1}(F_{< t})$, i.e., the linear maps to \mathbb{F} . Recall that

$$\beta_d(X) = \text{rank}(\widetilde{H}_d(X)) + \begin{cases} 1 & \text{if } d = 0, \\ 0 & \text{otherwise.} \end{cases}$$

We thus have

$$\beta_d(F_{\geq t}) = \beta_{n-d-1}(F_{< t}) + \begin{cases} 1 & \text{if } d = 0, \\ -1 & \text{if } d = n - 1, \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

By combining Eqs. (3) and (4), we obtain the stated equality. \square

For all piecewise linear functions f, g on \mathbb{S}^n with $\|g - f\|_\infty \leq \delta$ and $t \pm \delta \in \text{int}(\text{im } f)$, we have $t \in \text{int}(\text{im } g)$ and thus by Lemmas 1 and 5,

$$\beta(G_t) \geq \beta(F_{\leq t-\delta} \hookrightarrow F_{\leq t+\delta}) + \beta(F_{< t-\delta} \hookrightarrow F_{< t+\delta}).$$

This motivates the following definition:

Definition 3. Let f, g be piecewise linear functions on \mathbb{S}^n and consider real parameters t and δ with $t \pm \delta \in \text{int}(\text{im } f)$. The function g is called a level set (t, δ) -reconstruction of f if $\|g - f\|_\infty \leq \delta$ and

$$\beta(G_t) = \beta(F_{\leq t-\delta} \hookrightarrow F_{\leq t+\delta}) + \beta(F_{< t-\delta} \hookrightarrow F_{< t+\delta}).$$

A level set (t, δ) -reconstruction is thus also a minimizer of $\beta(G_t)$ subject to $\|g - f\|_\infty \leq \delta$. Since the above equality can only be achieved if both inequalities

$$\begin{aligned} \beta(G_{\leq t}) &\geq \beta(F_{\leq t-\delta} \hookrightarrow F_{\leq t+\delta}) \quad \text{and} \\ \beta(G_{< t}) &\geq \beta(F_{< t-\delta} \hookrightarrow F_{< t+\delta}) \end{aligned}$$

derived from Lemma 1 hold with equality, we conclude:

Lemma 6. Let f, g be piecewise linear functions on \mathbb{S}^n . If g is a level set (t, δ) -reconstruction of f , then

$$\begin{aligned} \beta(G_{\leq t}) &= \beta(F_{\leq t-\delta} \hookrightarrow F_{\leq t+\delta}) \\ \beta(G_{< t}) &= \beta(F_{< t-\delta} \hookrightarrow F_{< t+\delta}) \end{aligned}$$

and in particular g is also a sublevel set (t, δ) -reconstruction of f .

We will show in the following that sublevel set reconstructions are also level set reconstructions, under some additional hypotheses.

3.3. Level set reconstruction is NP-hard

Definition 4. Let f be a piecewise linear function. A homological regular value of f is a number $t \in \mathbb{R}$ such that $H_*(F_{<t} \hookrightarrow F_{\leq t})$ is an isomorphism.

We remark that there exist several other notions of regularity in the literature, which do not match our definition when extended to general functions [10, 4]. For piecewise linear functions however, all these definitions are equivalent. Note also that regularity should be understood with respect to sublevel sets; t can be a regular value t even though $H_*(F_{>t} \hookrightarrow F_{\geq t})$ might not be an isomorphism.

Lemma 7. Let f be a piecewise linear function on \mathbb{S}^n , $n > 1$. If $t \pm \delta \in \text{int}(\text{im } f)$ are regular values of f and g is a level set (t, δ) -reconstruction of f , then t is a regular value of g .

Proof. By hypothesis $t \pm \delta$ are regular values of f , so

$$H_*(F_{<t-\delta} \hookrightarrow F_{\leq t-\delta}) \quad \text{and} \quad H_*(F_{<t+\delta} \hookrightarrow F_{\leq t+\delta})$$

are isomorphisms and

$$\beta(F_{<t-\delta} \hookrightarrow F_{<t+\delta}) = \beta(F_{<t-\delta} \hookrightarrow F_{\leq t+\delta}) = \beta(F_{\leq t-\delta} \hookrightarrow F_{\leq t+\delta}).$$

Since g is a level set (t, δ) -reconstruction of f , by Lemma 6, we have

$$\begin{aligned} \beta(G_{\leq t}) &= \beta(F_{\leq t-\delta} \hookrightarrow F_{\leq t+\delta}) \quad \text{and} \\ \beta(G_{<t}) &= \beta(F_{<t-\delta} \hookrightarrow F_{<t+\delta}) \end{aligned}$$

and hence

$$\beta(G_{\leq t}) = \beta(G_{<t}) = \beta(F_{<t-\delta} \hookrightarrow F_{\leq t+\delta}).$$

Observing that

$$F_{<t-\delta} \subset G_{<t} \subset G_{\leq t} \subset G_{\leq t+\delta}$$

and using the fact that whenever we have three linear maps

$$U \rightarrow V \rightarrow W \rightarrow X$$

between finite-dimensional vector spaces, then

$$\text{rank}(U \rightarrow X) \leq \text{rank}(V \rightarrow W),$$

we get

$$\beta(F_{<t-\delta} \hookrightarrow F_{\leq t+\delta}) \leq \beta(G_{<t} \hookrightarrow G_{\leq t}) \leq \beta(G_{\leq t}).$$

Combining all these relations, we deduce that

$$\beta(G_{\leq t}) = \beta(G_{<t}) = \beta(G_{<t} \hookrightarrow G_{\leq t}) = \beta(G_{\leq t}).$$

and conclude that $H_*(G_{<t} \hookrightarrow G_{\leq t})$ is an isomorphism. \square

Lemma 8. *Let f and g be piecewise linear functions on \mathbb{S}^n , $n > 1$. Assume that $t \pm \delta \in \text{int}(\text{im } f)$ are regular values of f and $t \in \text{int}(\text{im } g)$ is a regular value of g . Then g is a sublevel set (t, δ) -reconstruction of f if and only if g is a level set (t, δ) -reconstruction of f .*

Proof. By hypothesis, t is a regular value of g . Substituting into Lemma 5, we obtain the first equation below; the second equation comes from the fact that $t \pm \delta$ are regular values of f :

$$\begin{aligned} 2\beta(G_{\leq t}) &= \beta(G_t), \\ 2\beta(F_{\leq t-\delta} \hookrightarrow F_{\leq t+\delta}) &= \beta(F_{\leq t-\delta} \hookrightarrow F_{\leq t+\delta}) + \beta(F_{< t-\delta} \hookrightarrow F_{< t+\delta}). \end{aligned}$$

By definition, g is a sublevel set (t, δ) -reconstruction of f if and only if the left hand sides of the two equations above are equal. Similarly, g is a level set (t, δ) -reconstruction if and only if the right hand sides of the two equations above are equal. The result follows immediately. \square

Theorem 4. *The level set reconstruction problem is NP-hard: Given as input a simplexwise linear function on a triangulation of \mathbb{S}^3 and parameters t and δ , decide whether there exists a level set (t, δ) -reconstruction g of f . The problem is NP-complete if g is required to be simplexwise linear on this triangulation.*

Proof. We reuse the same reduction as in Theorem 2. Since we need functions defined on the sphere, we triangulate the complement of K to obtain a triangulation S of the sphere with size polynomial in the size of K and $K \subset S$ as in Lemma 4. We extend f from Eq. (1) to a simplexwise linear function \tilde{f} on $\text{sd } S$:

$$\tilde{f} : \sigma \mapsto \begin{cases} f(\sigma) & \text{if } \sigma \in K, \\ 2 & \text{otherwise.} \end{cases}$$

We then prove that propositions (a)–(e) in the proofs of Theorems 1 and 2 and (f), (g) below are equivalent.

- (f) \tilde{f} has a simplexwise linear level set $(0, 1)$ -reconstruction \tilde{g} .
- (g) \tilde{f} has a level set $(0, 1)$ -reconstruction \tilde{g} .

We trivially have (f) \implies (g). Now we prove that (g) \implies (d). Proposition 1 implies that the values ± 1 are regular values of \tilde{f} . By Lemma 7, the value 0 is a regular value of \tilde{g} . Lemma 8 then proves that \tilde{g} is a sublevel set reconstruction of \tilde{f} . Now let g be the restriction of \tilde{g} to K . Since the sublevel sets $F_{\leq t}$ and $\tilde{F}_{\leq t}$ are homotopy equivalent for $t \leq 1$, and the sublevel sets $G_{\leq t}$ and $\tilde{G}_{\leq t}$ are homotopy equivalent for $t \leq 0$, it follows that g is a sublevel set reconstruction of f .

Next, we prove that (c) \implies (f). Given a subcomplex reconstruction X of (K, L) , we define g using Eq. (2) and extend it to $\tilde{g} : \text{sd } S \rightarrow \mathbb{R}$ as above for \tilde{f} . Since 0 is a regular value of \tilde{g} , Lemma 8 implies that \tilde{g} is a level set reconstruction.

In analogy to the proof of Theorem 2, we obtain NP-hardness of the level set reconstruction problem and NP-completeness of the problem restricted to simplexwise linear functions. \square

Corollary 3. *The level set simplification problem is NP-hard: Given a piecewise linear function f on \mathbb{S}^3 and parameters t and δ , find a simplexwise linear function g minimizing $\beta(G_t)$ subject to $\|g - f\|_\infty \leq \delta$.*

Proof. To determine if f has a level set (t, δ) -reconstruction, we can first find a minimizer of $\beta(G_t)$. We then only need to check if $\beta(G_t)$ matches the lower bound

$$\beta(G_t) = \beta(G_{\leq t}) + \beta(G_{< t}) \leq \beta(F_{\leq t-\delta} \hookrightarrow F_{\leq t+\delta}) + \beta(F_{< t-\delta} \hookrightarrow F_{< t+\delta}),$$

which can be done in time polynomial in the size of the underlying triangulation. \square

4. Realizations of well groups

We now discuss how the previous results relate to the concept of well groups, which were introduced in [16] as a robust version of the homology group of a level set.

Let $f : \mathcal{K} \rightarrow \mathbb{R}$ be a piecewise linear function. For $\delta \geq 0$ and an interval $[a, b] \subset \mathbb{R}$, the $([a, b], \delta)$ -well group of f is defined as

$$W_*(f, [a, b], \delta) = \bigcap_{g: \|g-f\|_\infty \leq \delta} \text{im } H_*(G_{[a,b]} \hookrightarrow F_{[a-\delta, b+\delta]}),$$

where $F_{[a,b]} = f^{-1}([a, b])$. In fact, as shown in [3], the well group is already given by the intersection of just two persistent homology groups:

$$W_*(f, [a, b], \delta) = \text{im } H_*(F_{[a-\delta, b-\delta]} \hookrightarrow F_{[a-\delta, b+\delta]}) \cap \text{im } H_*(F_{[a+\delta, b+\delta]} \hookrightarrow F_{[a-\delta, b+\delta]}). \quad (5)$$

The following formula expresses the rank of the well group in terms of persistent Betti numbers using relative homology.

Theorem 5 (Bendich et al. [3]). *Let $f : \mathcal{K} \rightarrow \mathbb{R}$ be a piecewise linear function and let $a \leq b$ and $\delta \in \mathbb{R}$ be such that $a \pm \delta, b \pm \delta$ are regular values of f . Then*

$$\begin{aligned} \text{rank } W_*(f, [a, b], \delta) &= \beta(F_{\leq b-\delta} \hookrightarrow F_{\leq b+\delta}) \\ &\quad - \beta((F_{\leq b-\delta}, \emptyset) \hookrightarrow (\mathcal{K}, F_{\geq a+\delta})) \\ &\quad + \beta((\mathcal{K}, F_{\geq a+\delta}) \hookrightarrow (\mathcal{K}, F_{\geq a-\delta})) \\ &\quad - \beta((F_{\leq b+\delta}, \emptyset) \hookrightarrow (\mathcal{K}, F_{\geq a-\delta})). \end{aligned}$$

We are particularly interested in the case where the interval consists of a single point. We call $W_*(f, t, \delta) = W_*(f, [t, t], \delta)$ the (t, δ) -well group of f . Intuitively, it captures the homology common to all perturbed level sets.

Clearly, the rank of the well group provides a lower bound on the Betti number of the t -level set of any g with $\|g - f\|_\infty \leq \delta$:

$$\beta(G_t) \geq \beta(G_t \hookrightarrow F_{[t-\delta, t+\delta]}) \geq \text{rank } W_*(f, t, \delta).$$

We say that the well group is *realized* by such a function g if

$$\beta(G_t) = \text{rank } W_*(f, t, \delta),$$

or equivalently, if $H_*(G_t \hookrightarrow F_{[t-\delta, t+\delta]})$ maps $H_*(G_t)$ bijectively to $W_*(f, t, \delta)$. As we will show in Theorem 6, this lower bound cannot always be achieved, and hence not every well group is realizable.

4.1. Realizability of well groups is NP-hard

We now show that on \mathbb{S}^n , a realization of a well group is the same as a level set reconstruction:

Theorem 6. *Let f be a piecewise linear function on \mathbb{S}^n with $t \pm \delta \in \text{int}(\text{im } f)$. A piecewise linear function g realizes the well group $W_*(f, t, \delta)$ if and only if it is a level set (t, δ) -reconstruction of f .*

Proof. The number of critical values of f is finite, and so for every $s \in \mathbb{R}$, there is $\epsilon > 0$ such that all values in $[s - \epsilon, s)$ and in $(s, s + \epsilon]$ are regular, and hence

$$H_*(F_{\leq s-\epsilon} \hookrightarrow F_{< s}) \quad \text{and} \quad H_*(F_{\leq s} \hookrightarrow F_{\leq s+\epsilon})$$

are isomorphisms. Choose ϵ such that the above holds for $s = t \pm \delta$. Let $a = t - \epsilon$ and $b = t + \epsilon$. Now $a \pm \delta, b \pm \delta$ are regular values and we can apply Theorem 5.

The second and forth terms in the formula of Theorem 5 vanish. To see this, note that $t \pm \delta \in \text{int}(\text{im } f)$ implies

$$F_{\leq b \pm \delta} = F_{\leq t + \epsilon \pm \delta} \subsetneq \mathbb{S}^n$$

for ϵ small enough, and thus $\beta_n(F_{\leq b \pm \delta}) = 0$. Similarly,

$$F_{\geq a \pm \delta} = F_{\geq t - \epsilon \pm \delta} \neq \emptyset$$

and thus $\beta_0(\mathbb{S}^n, F_{\geq a \pm \delta}) = 0$. Moreover, $\beta_d(\mathbb{S}^n) = 0$ for $d \notin \{0, n\}$. Since the induced homomorphisms

$$H_*(F_{\leq b \pm \delta}, \emptyset) \hookrightarrow (\mathbb{S}^n, F_{\geq a \mp \delta})$$

factor as

$$H_*(F_{\leq b \pm \delta}) \rightarrow H_*(\mathbb{S}^n) \rightarrow H_*(\mathbb{S}^n, F_{\geq a \mp \delta}),$$

we have

$$\beta((F_{\leq b \pm \delta}, \emptyset) \hookrightarrow (\mathbb{S}^n, F_{\geq a \mp \delta})) = 0.$$

Moreover, by the duality theorem of extended persistence on manifolds [11], we can rewrite the third term in Theorem 5 as

$$\beta_d((\mathbb{S}^n, F_{\geq a + \delta}) \hookrightarrow (\mathbb{S}^n, F_{\geq a - \delta})) = \beta_{n-d}(F_{\leq a - \delta} \hookrightarrow F_{\leq a + \delta}).$$

Finally, by regularity of the values $[a \pm \delta, t \pm \delta]$ and $(t \pm \delta, b \pm \delta]$, we have isomorphisms

$$H_*(F_{t \pm \delta} \hookrightarrow F_{[a \pm \delta, b \pm \delta]}) \quad \text{and} \quad H_*(F_{[t - \delta, t + \delta]} \hookrightarrow F_{[a - \delta, b + \delta]})$$

and thus by Eq. (5)

$$W_*(f, t, \delta) \cong W_*(f, [a, b], \delta).$$

Altogether, this yields

$$\text{rank } W_*(f, t, \delta) = \beta(F_{\leq t - \delta} \hookrightarrow F_{\leq t + \delta}) + \beta(F_{< t - \delta} \hookrightarrow F_{< t + \delta}).$$

The statement now follows directly from the definitions. \square

Together with Theorem 4, we have:

Corollary 4. *The well group realization problem is NP-hard: Given a piecewise linear function $f : K \subseteq \mathbb{S}^3 \rightarrow \mathbb{R}$ and parameters t and δ , decide whether the well group $W_*(f, t, \delta)$ can be realized. The problem is NP-complete if the realization is required to be simplexwise linear on K .*

5. An easy case

In this section, we discuss an important special case in which the subcomplex reconstruction problem can in fact be solved in polynomial time.

5.1. Building a p -reconstruction of an easily p -reconstructible pair

We start by presenting a polynomial time algorithm which outputs a p -reconstruction of the pair (K, L) , assuming that (K, L) is p -reconstructible and enjoys an easiness property that we describe below.

Definition 5. *Given a simplicial complex K and a subcomplex L , we say that (K, L) is an easily p -reconstructible pair if*

- (a) (K, L) is p -reconstructible and
- (b) for all subcomplexes X such that $L \subseteq X \subseteq K$, the homomorphism $H_{p-1}(X \hookrightarrow K)$ induced by the inclusion $X \subseteq K$ is injective.

Condition (b) is equivalent to requiring that for every filtration \mathcal{F} containing the two simplicial complexes L and K , no $(p - 1)$ -cycle is destroyed in \mathcal{F} between L and K . In other words, each time a p -simplex is added in \mathcal{F} between L and K , it creates a p -cycle; see Figure 7. Using the terminology in [15], this means that the filtration \mathcal{F} has only *positive* p -simplices in $K \setminus L$. Note that for condition (b) to hold we only need the positivity of p -simplices in $K \setminus L$ for one filtration \mathcal{F} and not for every permutation. To see this, recall that given a filtration \mathcal{F} there exists a pairing between its positive p -simplices and negative $(p + 1)$ -simplices. The analysis in [9] shows that when we swap two consecutive simplices in the filtration, either they keep their pairings, or they swap them, but in no case can the number of negative p -simplices change. It follows that we can go from one filtration \mathcal{F} containing L and K to any other while preserving the positivity of p -simplices in $K \setminus L$. Hence, checking condition (b) boils down to computing the pairing of p -simplices in \mathcal{F} and thus takes polynomial time. In practice, checking the easiness property will not be necessary, as we shall see below.

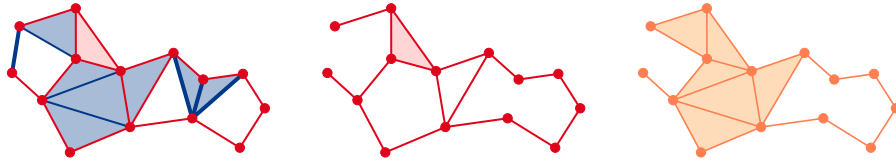


Figure 7: An easily 1-reconstructible pair (K, L) embedded in \mathbb{R}^2 and a 1-reconstruction obtained after removing from K the (K, L) -homology generating edges (bold edges) and their cofaces.

We now describe a polynomial time algorithm that constructs a solution to the dimension p reconstruction problem of the pair (K, L) , whenever (K, L) is easily p -reconstructible. The idea is to remove p -simplices from K in order to “break” p -cycles in K that do not correspond to cycles in L ; see Figure 7.

We say that a p -simplex $\sigma \in K \setminus L$ is (K, L) -homology generating if there is a chain $c \in C_p(L)$ such that $\partial\sigma = \partial c$ and $[\sigma + c]_K \notin \text{im } H_p(L \hookrightarrow K)$. Clearly, this implies that

σ cannot be contained in any homological reconstruction X of (K, L) , and hence the same is also true for every coface of σ . Writing $\text{st}_K \sigma$ for the set of cofaces of σ in K , we conclude:

Lemma 9. *Let X be a solution to the dimension p reconstruction problem of the pair (K, L) . For any (K, L) -homology generating p -simplex σ , we have $\text{st}_K \sigma \subseteq K \setminus X$.*

This lemma suggests the following algorithm for computing a p -reconstruction of an easily p -reconstructible pair (K, L) :

```

RECONSTRUCTION( $K, L, p$ )
   $K' \leftarrow K$ 
  while  $\exists$  a  $(K', L)$ -homology generating  $p$ -simplex  $\sigma$ 
    remove  $\text{st}_{K'} \sigma$  from  $K'$ 
  endwhile
  return  $K'$ 

```

We now show correctness of the algorithm.

Lemma 10. *Suppose (K, L) is an easily p -reconstructible pair. For any (K, L) -homology generating p -simplex σ , the pair (K', L) , where $K' = K \setminus \text{st}_K \sigma$, is an easily p -reconstructible pair. Moreover, every p -reconstruction of (K', L) is also a p -reconstruction of (K, L) .*

Proof. Let $K' = K \setminus \text{st}_K \sigma$. Let X be a solution to the dimension p reconstruction problem of the pair (K, L) . By Lemma 9, we have $L \subseteq X \subseteq K'$. Consider the commutative diagram of Figure 8 (left) where all maps are induced by inclusions. Since X is a p -

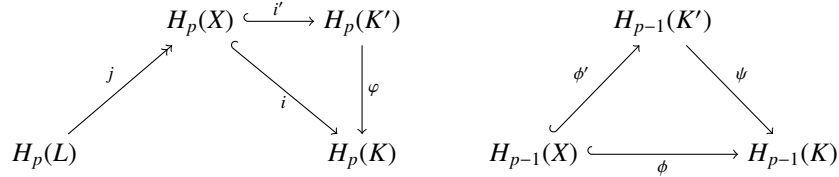


Figure 8: Commutative diagrams for the proof of Lemma 10. Left: injectivity of i implies injectivity of i' . Right: injectivity of ϕ implies injectivity of ϕ' .

reconstruction of the pair (K, L) , i is injective and j is surjective. Since $i = \phi \circ i'$, the map i' is also injective and thus $\text{im}(i' \circ j) \cong H_p(X)$, showing that X is also a p -reconstruction of the pair (K', L) .

We now use the easiness of the pair (K, L) to prove the easiness of the pair (K', L) . Consider an arbitrary simplicial complex X such that $L \subseteq X \subseteq K'$ and the commutative diagram of Figure 8 (right) where all maps are induced by inclusions. Since $\phi = \psi \circ \phi'$, the injectivity of ϕ implies the injectivity of ϕ' . \square

Suppose (K, L) is an easily p -reconstructible pair. From Lemma 10, it follows that at each step of the reconstruction algorithm, (K', L) is also an easily p -reconstructible pair.

If $H_p(K') \not\cong \text{im } H_p(L \hookrightarrow K')$, we claim that we can always find a (K', L) -homology generating p -simplex σ . Indeed, by assumption, every p -simplex σ in $K' \setminus L$ is positive for every filtration \mathcal{F} containing L and K' . This implies that $\partial\sigma$ is the boundary of some p -chain $c \in C_p(L)$ for every p -simplex σ in $K' \setminus L$. The classes $[\sigma + c]_{K'}$, where σ is a p -simplex in $K' \setminus L$, together with $\text{im } H_p(L \hookrightarrow K')$, generate $H_p(K')$. Since $H_p(K') \not\cong \text{im } H_p(L \hookrightarrow K')$, there must be a σ such that $[\sigma + c]_{K'} \notin \text{im } H_p(L \hookrightarrow K')$. Both finding a c for a given σ and deciding whether $[\sigma + c]_{K'} \in \text{im } H_p(L \hookrightarrow K')$ can be done in time polynomial in the size of K' .

The size of K' decreases strictly during the course of the algorithm. Since K is finite, the algorithm has to stop eventually, and when it stops, we have $H_p(K') \cong \text{im } H_p(L \hookrightarrow K') \cong \text{im } H_p(L \hookrightarrow K)$.

In practice, we need not test whether or not the pair (K, L) satisfies the easiness property. It suffices to run the algorithm and check if the resulting complex K' is a p -reconstruction. If the pair (K, L) is not easily p -reconstructible (as in Figure 1), the algorithm will output a simplicial complex X nested between L and K whose p -dimensional Betti number will differ from the persistent Betti number $\beta_p(L \hookrightarrow K)$. Nonetheless, if (K, L) has any p -reconstruction, it must be a subset of K' , and the algorithm may occasionally output a reconstruction even for pairs (K, L) that do not enjoy the easiness property.

5.2. Reconstruction in 3D

First, we review the use of persistent homology groups for homological inference, as proposed in [10, 5]. Second, we formulate the problem of reconstructing a 3D shape as one of finding a subcomplex reconstruction of a simplicial pair which enjoys the property to be easily 1-reconstructible. Assuming a solution exists, we then describe how to build it in polynomial time. We use the notation $\Omega^\alpha = \{x \in \mathbb{R}^n : d(x, \Omega) \leq \alpha\}$.

Definition 6. Let $\Omega \subset \mathbb{R}^n$ and let $S \subset \mathbb{R}^n$ be finite. We say that S is a homological (δ, ϵ) -sample of Ω if $\Omega \subseteq S^\delta$, $S \subseteq \Omega^\epsilon$, and both

$$H_*(\Omega \hookrightarrow \Omega^{\delta+\epsilon}) \quad \text{and} \quad H_*(\Omega^{\delta+\epsilon} \hookrightarrow \Omega^{2\delta+2\epsilon})$$

are isomorphisms.

Roughly, δ is a bound on the sampling density, and ϵ is a bound on the sampling error. If S is a homological (δ, ϵ) -sample of Ω , then the plain arrows in the following diagram commute:

$$\begin{array}{ccccc}
 H_*(\Omega) & \hookrightarrow & H_*(\Omega^{\delta+\epsilon}) & \hookrightarrow & H_*(\Omega^{2\delta+2\epsilon}) \\
 \searrow & & \nearrow & & \searrow \\
 & & H_*(S^\delta) & & H_*(S^{2\delta+\epsilon}) \\
 & & \searrow & & \nearrow \\
 & & & & \text{im } H_*(S^\delta \hookrightarrow S^{2\delta+\epsilon})
 \end{array}$$

(Note: A dashed vertical arrow points from $H_*(\Omega^{\delta+\epsilon})$ down to $\text{im } H_*(S^\delta \hookrightarrow S^{2\delta+\epsilon})$.)

Moreover, the morphism $\text{im } H_*(\Omega^{\delta+\epsilon} \hookrightarrow S^{2\delta+\epsilon})$ defines an isomorphism from $H_*(\Omega^{\delta+\epsilon})$ to $\text{im } H_*(S^\delta \hookrightarrow S^{2\delta+\epsilon})$. Hence, $\text{im } H_*(S^\delta \hookrightarrow S^{2\delta+\epsilon}) \cong H_*(\Omega)$.

Given as input a point set S that samples Ω , we are thus able to infer the homology groups of Ω from S by computing the persistent homology groups of the pair $(S^{2\delta+\epsilon}, S^\delta)$. Moreover we have the following lemma:

Lemma 11. *Let S be a homological (δ, ϵ) -sample of $\Omega \subset \mathbb{R}^n$ with $\delta \geq \epsilon$. Then $H_0(S^\delta \hookrightarrow S^{2\delta+\epsilon})$ is an isomorphism.*

Proof. First, note that $H_0(S^\delta \hookrightarrow S^{2\delta+\epsilon})$ is surjective, since every component of $S^{2\delta+\epsilon}$ contains a point of $S \subset S^\delta$. It remains to prove that $H_0(S^\delta \hookrightarrow S^{2\delta+\epsilon})$ is injective.

We first show that $H_0(\Omega \hookrightarrow S^\delta)$ is surjective. Let $x \in S^\delta$. There is $s \in S$ with $d(x, s) \leq \delta$. Moreover, there is $y \in \Omega$ with $d(s, y) \leq \epsilon$. Since $\epsilon \leq \delta$, the two points x and y are both contained in the ball of radius δ around s and hence in the same connected component of S^δ . In other words, every connected component of S^δ contains a point of Ω , so $H_0(\Omega \hookrightarrow S^\delta)$ is surjective.

Since $H_0(\Omega \hookrightarrow \Omega^{\delta+\epsilon})$ is an isomorphism, this implies that $H_0(S^\delta \hookrightarrow \Omega^{\delta+\epsilon})$ must be injective. By injectivity of $H_0(\Omega^{\delta+\epsilon} \hookrightarrow S^{2\delta+\epsilon})$, we obtain that $H_0(S^\delta \hookrightarrow S^{2\delta+\epsilon})$ is injective. \square

In practice, we replace each S^α in the pair by the corresponding α -complex of S which can be computed efficiently in \mathbb{R}^3 using the Delaunay triangulation. We recall that the *Delaunay triangulation* is the set of simplices $\sigma \subset S$ for which there exists a ball whose boundary contains the vertices of σ and which encloses no point of S in its interior. Such a ball is said to be *empty*. The α -complex, denoted $A_\alpha(S)$, is the subcomplex of the Delaunay triangulation obtained by keeping simplices that fit in an empty ball of radius α or less. It is a deformation retraction of the offset S^α .

We now focus our attention on the case $n = 3$ and $\delta \geq \epsilon$. It turns out that in this case we can find a subcomplex reconstruction of the pair of α -complexes $(K, L) = (A_{2\delta+\epsilon}(S), A_\delta(S))$ in polynomial time, if one exists. Note that for each $\alpha \geq 0$, the offset S^α deformation retracts to $A_\alpha(S)$, and the pair $(S^{2\delta+\epsilon}, S^\delta)$ has a subspace reconstruction $\Omega^{\delta+\epsilon}$. Note however that this does not imply that (K, L) has a (subcomplex or subspace) reconstruction. From now on, we assume that a subcomplex reconstruction of (K, L) exists. We next describe how to find one under this assumption.

Since $H_0(L \hookrightarrow K)$ is an isomorphism by Lemma 11 and since there are no new vertices in $K \setminus L$, this implies that every complex nested between L and K is a 0-reconstruction. Moreover, no edge in $K \setminus L$ joins two connected components of L ; in other words, (K, L) is easily 1-reconstructible. Construct a 1-reconstruction K' as described above. Recall that this takes time polynomial in the size of K .

By Alexander duality, the finite connected components of the complement $\mathbb{R}^3 \setminus K'$ correspond to classes in $H_2(K')$. Hence, K'' is a 2-reconstruction of (K', L) if and only if any two components of $\mathbb{R}^3 \setminus K'$ that lie in the same component of $\mathbb{R}^3 \setminus L$ are also contained in the same component of $\mathbb{R}^3 \setminus K''$ (see Figure 9 for an analogous illustration in \mathbb{R}^2). We note that such a 2-reconstruction of (K', L) is also a 2-reconstruction of (K, L) because $H_2(K' \hookrightarrow K)$ is injective as cavities in K' cannot be destroyed in K by construction of K' . In order to obtain a 2-reconstruction K'' , we now remove simplices

of dimension 2 and 3 from K' in order to connect all components in $\mathbb{R}^3 \setminus K'$ that are in the same component of $\mathbb{R}^3 \setminus L$.

Consider the dual graph G of K' , whose vertices are the tetrahedra of K' together with the connected components of $\mathbb{R}^3 \setminus K'$, and whose edges correspond to the triangles of K' ; see Figure 9. Let G' be the subgraph of G whose edges correspond to triangles in $K' \setminus L$. Now two components of $\mathbb{R}^3 \setminus K'$ lie in the same component of $\mathbb{R}^3 \setminus L$ if and only if they are connected by a path in G' . Removing the corresponding triangles and tetrahedra from K' merges the two components of the complement $\mathbb{R}^3 \setminus K'$. Repeating this procedure while there are mergeable components, we obtain a complex K'' with $H_2(K'') \cong \text{im } H_2(L \hookrightarrow K)$. The construction of K'' can also be done in polynomial time.

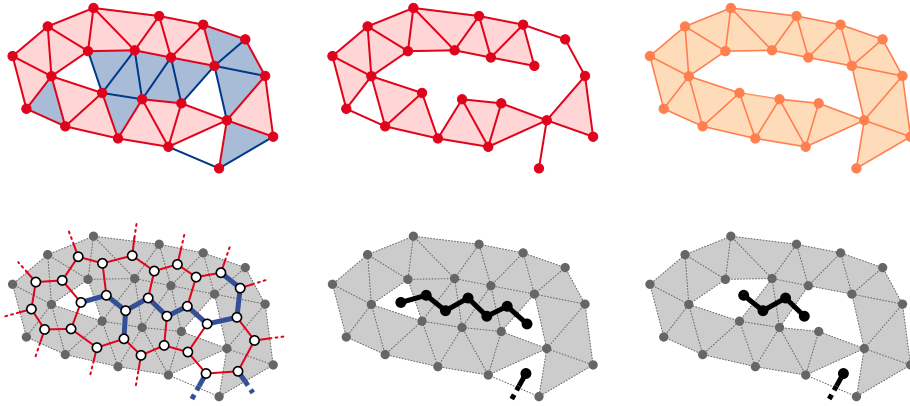


Figure 9: Top: A pair (K', L) embedded in \mathbb{R}^n and its $(n - 1)$ -reconstruction for $n = 2$. Bottom from left to right: graph G whose edges correspond to $(n - 1)$ -simplices of K' and subgraph G' whose (bold) edges correspond to $(n - 1)$ -simplices of $K' \setminus L$. Paths joining connected components of $\mathbb{R}^n \setminus K'$ in G' . Applying simplicial collapses along a path to merge two connected components.

Note that this procedure will not affect the property that the resulting complex K'' is a 1-reconstruction: two components in the complement of K' can be merged in the complement of K'' by a sequence of simplicial collapses on K' , followed by the removal of a 2-simplex of K' with empty coboundary; see Figure 9, bottom. This triangle must be positive, since its removal merges two components of the complement. In other words, its removal will not destroy any 1-cycle in the 1-reconstruction. The resulting complex K'' is thus a reconstruction of (K, L) . We conclude:

Theorem 7. *Let S be a homological (δ, ϵ) -sample of $\Omega \subset \mathbb{R}^3$ with $\delta \geq \epsilon$. Then the subcomplex reconstruction problem for the pair of α -complexes $(A_{2\delta+\epsilon}(S), A_\delta(S))$ can be solved in polynomial time. If a reconstruction exists, its homology is isomorphic to that of Ω .*

6. Conclusion

The homological reconstruction problem of simplicial pairs embedded in \mathbb{R}^3 is NP-hard. It is also NP-hard to decide whether a real-valued simplexwise linear function

in \mathbb{R}^3 has a level set or a sublevel set reconstruction. We deduce that simplifying the homology of a simplicial pair embedded in \mathbb{R}^3 is also NP-hard and so is the homological simplification of level and sublevel sets of real-valued simplexwise linear functions in \mathbb{R}^3 . On the other hand, such problems can be solved in polynomial time if we restrict ourselves to pairs of α -complexes in \mathbb{R}^3 that admit homological inference of a compact space, given an appropriate sample. Can we use this construction to devise a shape reconstruction algorithm with homological guarantees under the same sampling conditions?

Acknowledgements. This work was initiated during the 10th McGill–INRIA Workshop on Computational Geometry at the Bellairs Research Institute. The authors wish to thank all the participants for creating a pleasant and stimulating atmosphere, in particular Nina Amenta for discussions leading to a first version of Theorem 1. Some of the authors were partially supported by the GIGA ANR grant (contract ANR-09-BLAN-0331-01), the European project CG-Learning (contract 255827), and the TOPOSYS project FP7-ICT-318493-STREP.

- [1] D. Attali and A. Lieutier. [Optimal reconstruction might be hard](#). *Discrete & Computational Geometry*, 49(2):133–156, 2013.
- [2] D. Attali, A. Lieutier, and D. Salinas. [Vietoris–Rips complexes also provide topologically correct reconstructions of sampled shapes](#). *Computational Geometry*, 46(4):448–465, 2013.
- [3] P. Bendich, H. Edelsbrunner, D. Morozov, and A. Patel. [Homology and robustness of level and interlevel sets](#). *Homology, Homotopy and Applications*, 15(1):51–72, 2013.
- [4] P. Bubenik and J. Scott. [Categorification of persistent homology](#). *Discrete & Computational Geometry*, 2014. Available online.
- [5] 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.
- [6] F. Chazal and A. Lieutier. [Smooth manifold reconstruction from noisy and non-uniform approximation with guarantees](#). *Computational Geometry*, 40(2):156–170, 2008.
- [7] F. Chazal, D. Cohen-Steiner, and A. Lieutier. [A sampling theory for compact sets in Euclidean space](#). *Discrete & Computational Geometry*, 41(3):461–479, 2009.
- [8] F. Chazal, V. de Silva, M. Glisse, and S. Oudot. [The structure and stability of persistence modules](#). Preprint, 2012. [arXiv:1207.3674](#).
- [9] D. Cohen-Steiner, H. Edelsbrunner, and D. Morozov. [Vines and vineyards by updating persistence in linear time](#). In *Proceedings of the Twenty-second Annual Symposium on Computational Geometry, SCG '06*, pages 119–126. ACM, 2006.
- [10] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. [Stability of persistence diagrams](#). *Discrete & Computational Geometry*, 37(1):103–120, 2007.

- [11] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. [Extending persistence using Poincaré and Lefschetz duality](#). *Foundations of Computational Mathematics*, 9(1): 79–103, 2008.
- [12] V. de Silva and G. Carlsson. [Topological estimation using witness complexes](#). In *Eurographics Symposium on Point-Based Graphics*, pages 157–166, 2004.
- [13] H. Edelsbrunner. Alpha shapes — a survey. In R. van de Weygaert, G. Vegter, J. Ritzerveld, and V. Icke, editors, *Tessellations in the Sciences: Virtues, Techniques and Applications of Geometric Tilings*. Springer Verlag. To appear.
- [14] H. Edelsbrunner and M. Kerber. [Alexander duality for functions: the persistent behavior of land and water and shore](#). In *Proceedings of the 2012 symposium on Computational Geometry*, pages 249–258. ACM, 2012.
- [15] H. Edelsbrunner, D. Letscher, and A. Zomorodian. [Topological persistence and simplification](#). *Discrete & Computational Geometry*, 28(4):511–533, 2002.
- [16] H. Edelsbrunner, D. Morozov, and A. Patel. [Quantifying transversality by measuring the robustness of intersections](#). *Foundations of Computational Mathematics*, 11(3):345–361, 2011.
- [17] A. Hatcher. *Algebraic Topology*. Cambridge University Press, 2002.
- [18] W. Kühnel. Triangulations of manifolds with few vertices. In F. Triccerri, editor, *Advances in differential geometry and topology*, pages 59–114. World Scientific, Singapore, 1990.
- [19] N. Milosavljević, D. Morozov, and P. Skraba. [Zigzag persistent homology in matrix multiplication time](#). In *Proceedings of the twenty-seventh annual symposium on Computational geometry*, SoCG ’11, pages 216–225. ACM, 2011.
- [20] D. Morozov. *Homological Illusions of Persistence and Stability*. PhD thesis, Duke University, 2008.
- [21] 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.
- [22] E. H. Spanier. *Algebraic Topology*. Springer, 1994.