Minimum Bounded Chains and Minimum Homologous Chains in Embedded Simplicial Complexes

Glencora Borradaile
Oregon State University, Corvallis, OR, USA
glencora@eecs.oregonstate.edu

William Maxwell
Oregon State University, Corvallis, OR, USA
maxwellw@oregonstate.edu

Amir Nayyeri
Oregon State University, Corvallis, OR, USA
nayyeria@oregonstate.edu

Abstract

We study two optimization problems on simplicial complexes with homology over \( \mathbb{Z}_2 \), the minimum bounded chain problem: given a \( d \)-dimensional complex \( K \) embedded in \( \mathbb{R}^{d+1} \) and a null-homologous \((d-1)\)-cycle \( C \) in \( K \), find the minimum \( d \)-chain with boundary \( C \), and the minimum homologous chain problem: given a \((d+1)\)-manifold \( M \) and a \( d \)-chain \( D \) in \( M \), find the minimum \( d \)-chain homologous to \( D \). We show strong hardness results for both problems even for small values of \( d \); \( d = 2 \) for the former problem, and \( d = 1 \) for the latter problem. We show that both problems are \( \text{APX} \)-hard, and hard to approximate within any constant factor assuming the unique games conjecture. On the positive side, we show that both problems are fixed-parameter tractable with respect to the size of the optimal solution. Moreover, we provide an \( O(\sqrt{\log \beta_2}) \)-approximation algorithm for the minimum bounded chain problem where \( \beta_2 \) is the 2nd Betti number of \( K \). Finally, we provide an \( O(\sqrt{\log n_{d+1}}) \)-approximation algorithm for the minimum homologous chain problem where \( n_{d+1} \) is the number of \((d+1)\)-simplices in \( M \).

2012 ACM Subject Classification Theory of computation → Computational geometry

Keywords and phrases computational topology, algorithmic complexity, simplicial complexes

Digital Object Identifier 10.4230/LIPIcs.SoCG.2020.21

Related Version A full version of this paper is available at https://arxiv.org/abs/2003.02801.

Funding This material is based upon work supported by the National Science Foundation under Grant Nos. CCF-1617951 and CCF-1816442.

1 Introduction

Simplicial complexes are best known as a generalization of graphs, but have more structure than other generalizations such as hypergraphs. Despite the structure, simplicial complexes are sufficiently expressive to make many algorithmic questions computationally intractable. For example, the generalization of shortest path that we examine in this work is \( \text{NP} \)-hard in 2-dimensional simplicial complexes [16]. Since planar graphs (1-dimensional simplicial complexes embeddable in \( \mathbb{R}^2 \)) exhibit structure that is algorithmically useful, resulting in more efficient or more accurate algorithms than for general graphs, we ask whether 2-dimensional simplicial complexes that are embeddable in \( \mathbb{R}^3 \) (and more generally, \( d \)-complexes embeddable in \( \mathbb{R}^{d+1} \)) also have sufficient structure that can be exploited algorithmically. To this end, we examine the algebraic generalization of the shortest path problem in graphs to simplicial
complexes of higher dimension. This restriction via embedding in Euclidean space would still result in a useful algorithmic tool, given the connection of embedded simplicial complexes to meshes arising from physical systems.

Formally we study the **minimum bounded chain problem** which is the algebraic generalization of the shortest path problem in graphs [23]. The goal of the minimum bounded chain problem is to find a subcomplex whose boundary is a given input cycle \( C \). More precisely:

Given a \( d \)-dimensional simplicial complex \( K \) and a null-homologous \((d-1)\)-dimensional cycle \( C \subset K \), find a minimum-cost \( d \)-chain \( D \subset K \) whose boundary \( \partial D = C \). The requirement that the cycle be null-homologous is necessary and sufficient for the existence of a solution and we study the problem in the context of \( \mathbb{Z}_2 \)-homology.\(^1\) In \( \mathbb{Z}_2 \)-homology, a \( d \)-chain is a subset of \( d \)-simplices of the simplicial complex. We see this as a generalization of the shortest path problem in graphs as follows: Let \( K \) be a one dimensional simplicial complex (i.e. a graph). A pair of vertices in the same connected component, \( s \) and \( t \), is a null-homologous 0-chain and the minimum 1-chain whose boundary is \( \{s,t\} \) is the shortest \((s,t)\)-path. Grady has written on why this generalization is useful in the context of 3D graphics [18].

The minimum bounded chain problem is closely related to the minimum homologous chain problem which asks: given a \( d \)-chain \( D \), find a minimum-cost \( d \)-chain \( X \) such that the symmetric difference of \( D \) and \( X \) form the boundary of a \((d+1)\)-chain. Alternatively, \( X \) is the minimum-cost \( d \)-chain that is homologous to \( D \). Dunfield and Hirani [16] show the minimum bounded and homologous chain problems are equivalent under additional assumptions. We study the minimum homologous chain problem for \( d \)-chains in \((d+1)\)-manifolds.

### 1.1 Our results

We present approximation and fixed-parameter tractable algorithms for the minimum bounded chain and the minimum homologous chain problem. In this paper we consider both problems in the context of simplicial homology over \( \mathbb{Z}_2 \). We denote by \( n_d \) the number of \( d \)-simplices of the \( d \)-dimensional simplicial complex \( K \). Two of our results assume the unique games conjecture. For an overview of the unique games conjecture and its impact on computational topology we refer the reader to the work of Growchow and Tucker-Foltz [19].

- **Theorem 1.** There exists an \( O(\sqrt{\log \beta_d}) \)-approximation algorithm for the minimum bounded chain problem for a simplicial complex \( K \) embedded in \( \mathbb{R}^{d+1} \), with \( d \)th Betti number \( \beta_d \).

- **Theorem 2.** There exists an \( O(15^k \cdot k \cdot n_3^3) \) time exact algorithm for the minimum bounded chain problem for simplicial complexes embedded in \( \mathbb{R}^{d+1} \), where \( k \) is the number of \( d \)-simplices in the optimal solution.

- **Theorem 3.** There exists an \( O(\sqrt{\log n_{d+1}}) \)-approximation algorithm for the minimum homologous chain problem for \( d \)-chains in \((d+1)\)-manifolds.

- **Theorem 4.** There exists an \( O(15^k \cdot k \cdot n_3^3) \) time exact algorithm for the minimum homologous chain problem for \( d \)-chains in \((d+1)\)-manifolds, where \( k \) is the size of the optimal solution.

The running times for the first two theorems is computed assuming that the dual graph of the complex in \( \mathbb{R}^{d+1} \) is available. The last two theorems hold, more generally, for weak pseudomanifolds studied by Dey et al. in [14].

On the hardness side, we show that constant factor approximation algorithms for these problems (minimum bounded chain and minimum homologous chain) are unlikely.

---

\(^1\) Formal definitions are presented in Section 2.
Theorem 5. The minimum bounded chain problem is
(i) hard to approximate within a \((1 + \varepsilon)\) factor for some \(\varepsilon > 0\) assuming \(P \neq NP\), and
(ii) hard to approximate within any constant factor assuming the unique games conjecture, even if \(K\) is a 2-dimensional simplicial complex embedded in \(\mathbb{R}^3\) with input cycle \(C\) embedded on the boundary of the unbounded volume in \(\mathbb{R}^3 \setminus K\).

Theorem 6. The minimum homologous chain problem is
(i) hard to approximate within a \((1 + \varepsilon)\) factor for some \(\varepsilon > 0\) assuming \(P \neq NP\), and
(ii) hard to approximate within any constant factor assuming the unique games conjecture, even when the input chain is a 1-cycle on an orientable 2-manifold.

1.2 Related Work

1.2.1 Chain problems over \(\mathbb{Z}\) and \(\mathbb{R}\)

Research on the minimum bounded chain problem is limited to the case of \(\mathbb{Z}\)-homology, where linear programming techniques can be employed algorithmically. Sullivan described the problem as the discretization of the minimal spanning surface problem [28] with Kirsanov reducing the problem to an instance of minimum cut in the dual graph [23]. Sullivan’s work is on the closely related cellular complexes, but under the same restrictions we study (embedded in \(\mathbb{R}^d\)) and Kirsanov studies the problem in embedded simplicial complexes.

Likewise, research on minimum homologous chain has largely worked in \(\mathbb{Z}\)-homology. Dey, Hirani and Krishnamoorthy formulate the minimum homologous chain problem over \(\mathbb{Z}\) as an integer linear program and describe topological conditions for the linear program to be totally unimodular (and so, poly-time solvable) [13]. Of course, integer linear programming approaches do not extend to \(\mathbb{Z}_2\)-homology.

This linear programming approach was then applied to the minimum bounded chain problem (over \(\mathbb{Z}\)) by Dunfield and Hirani [16]. Moreover, they show the minimum bounded chain problem is \(NP\)-complete via a reduction from 1-in-3 SAT. The gadget they use was originally used by Agol, Hass and Thurston to show that the minimal spanning area problem is \(NP\)-complete [2].

Linear programming techniques have also been used by Chambers and Vejdemo-Johansson to solve the minimum bounded chain problem in the context of \(\mathbb{R}\)-homology [9]. In \(\mathbb{R}\)-homology Carvalho et al provide an algorithm finding a (not necessarily minimum) bounded chain in a manifold by searching the dual graph of the manifold [7].

1.2.2 Chain problems over \(\mathbb{Z}_2\)

Special cases of the minimum homologous chain problem have been studied in \(\mathbb{Z}_2\) homology. The homology localization problem is the case when the input chain is a cycle. The homology localization problem over \(\mathbb{Z}_2\) in surface-embedded graphs is known to be \(NP\)-hard via a reduction from maximum cut by Chambers et al. [8]; our reduction is from the complement problem minimum uncut. On the algorithmic side, Erickson and Nayyeri provide a \(2^{O(g)}n \log n\) time algorithm where \(g\) is the genus of the surface [17]. Using the idea of annotated simplices, Busaryev et al. generalize this algorithm for homology localization of 1-cycles in simplicial complexes; the algorithm runs in \(O(n^\omega) + 2^{O(g)}n^2 \log n\) time where \(\omega\) is the exponent of matrix multiplication, and \(g\) is the first homology rank of the complex [6].

Using a reduction from the nearest codeword problem Chen and Freedman showed that homology localization with coefficients over \(\mathbb{Z}_2\) is not only \(NP\)-hard, but it cannot be approximated within any constant factor in polynomial time [11]. These hardness results...
hold for a 2-dimensional simplicial complex, but not necessarily for 2-dimensional complexes embedded in $\mathbb{R}^3$. They also give a polynomial-time algorithm for the special case of $d$-dimensional simplicial complex that is embedded in $\mathbb{R}^d$. (This is different from our setting of a $d$-dimensional simplicial complex that is embedded in $\mathbb{R}^{d+1}$; however the algorithm also reduces to a minimum cut problem in a dual graph, much like that of Kirsanov and Gortler.)

1.2.3 Algebraic formulations

The minimum bounded chain problem over $\mathbb{Z}_2$ can be stated as a linear algebra problem, but this has little algorithmic use since the resulting problems are intractable. The algebraic formulation is to find a vector $x$ of minimum Hamming weight that solves an appropriately defined linear system $Ax = b$. (It is possible to reduce in the reverse direction, but the resulting complex is not embeddable in general, and so provides no new results.)

In coding theory this algebraic problem is a well studied decoding problem known as maximum likelihood decoding, and it was shown to be NP-hard by Berlekamp, McEliece and van Tilborg [4, 29]. Downey, Fellows, Vardy and Whittle show that maximum likelihood decoding is $W[1]$-hard [15]. Further, Austrin and Khot show that maximum likelihood decoding is hard to approximate within a factor of $2^{(\log n)^{O(1)}}$ under the assumption that $\mathbf{NP} \not\subset \mathbf{DTIME}(2^{(\log n)^{O(1)}})$ [3]. This work was continued by Bhattacharyya, Gadekar, Ghosal and Saket who showed that maximum likelihood decoding is still $W[1]$-hard when the problem is restricted to $O(k \log n) \times O(k \log n)$ sized matrices for some constant $k$ [5].

1.2.4 Paper organization

In Section 2, we give formal definitions for the paper. In Section 3, we present our approximation algorithms and fixed-parameter tractable algorithms. In Section 4, we present our hardness results.

2 Preliminaries

2.1 Simplicial complexes

Given a set of vertices $V$ we define an abstract simplicial complex $\mathcal{K}$ to be a subset of the power set of $V$ such that the following property holds: if $\sigma \in \mathcal{K}$ and $\tau \subset \sigma$ then $\tau \in \mathcal{K}$. We call any $\sigma \in \mathcal{K}$ a simplex and define the dimension of $\sigma$ to be $|\sigma| - 1$ if $|\sigma| - 1 = d$ we call $\sigma$ a $d$-simplex. Further, we call 0-simplices, 1-simplices, and 2-simplices vertices, edges, and triangles. We define the dimension of $\mathcal{K}$ to be equal to the largest dimension of any simplex in $\mathcal{K}$. If $\mathcal{K}$ has dimension $d$ we refer to $\mathcal{K}$ as a $d$-simplicial complex or $d$-complex. We refer to any subset of a $d$-simplex $\sigma$ as a face of $\sigma$.

2.2 Homology

In this paper we work in simplicial homology with coefficients over the finite field $\mathbb{Z}_2$. Here we briefly define the concepts from homology that will be used throughout this paper. We assume familiarity with the basics of algebraic topology, and refer the reader to standard references [20, 25] for the details.

Given a simplicial complex $\mathcal{K}$ we define the $d$th chain group of $\mathcal{K}$ to be the free abelian group, with coefficients over $\mathbb{Z}_2$, generated by the $d$-simplices in $\mathcal{K}$. We denote the chain group as $C_d(\mathcal{K})$ and note that its elements are expressed as formal sums $\bigoplus \alpha_i \sigma_i$, where
α_i ∈ Z_2 and σ_i ∈ K is a d-simplex. We call the elements of the chain group chains or more specifically d-chains. When working over Z_2 there is a one-to-one correspondence between d-chains and sets of d-simplices in K. It follows that adding two d-chains over Z_2 is the same thing as taking the symmetric difference of their corresponding sets. Hence, we use the notation σ ⊕ τ to denote the sum of two d-chains. By abuse of notation we will also use ⊕ to denote the symmetric difference of sets, but the context should always be clear.

For a d-simplex σ we define its boundary ∂σ to be the sum of the (d − 1)-simplices contained in σ. We extend this operation linearly to obtain the boundary operator on chain groups, ∂_d: C_d(K) → C_{d−1}(K). We will often drop the subscript when the context is clear. Note that the composition ∂_{d−1}∂_d is always equal to the zero map. If ∂σ = τ we say that σ is bounded by τ. We call a chain σ a cycle if ∂σ = 0.

By Z_d(K) we denote the dth cycle group of K. This is subgroup of C_p(K) generated by the d-simplices in Ker ∂_d. Similarly, by B_d(K) we denote the dth boundary group of K, which is the subgroup of C_p(K) generated by the d-simplices in img ∂_{d+1}. Since ∂_{d+1}∂_d = 0 we have that B_d(K) is a subgroup of Z_d(K). We define the dth homology group of K, denoted H_d(K), to be the quotient group Z_d(K)/B_d(K). The dth Betti number of K, denoted β_d, is defined to be the dimension of H_d(K). We call a d-chain σ null-homologous if it is a boundary, that is σ ∈ B_d(K). Further, we call two d-chains σ and τ homologous if their difference is a boundary, that is σ ⊕ τ ∈ B_d(K).

2.3 Embeddings and duality

Given a d-complex K an embedding of K is a function f: K → R^{d+1} such that f restricted to any simplex in K is an injection. Further, for any two simplices σ, τ ∈ K we require that f(σ) ∩ f(τ) = f(σ ∩ τ). That is, the images of two simplices only intersect at their common faces. The function f is an embedding of the abstract simplicial complex K. In this paper we make no distinction between K and an embedding of K. Hence, we use the notation K to refer to both and refer to a K as an embedded simplicial complex.

The Alexander duality theorem, a higher dimensional analog of the Jordan curve theorem, states that R^{d+1} \ K is partitioned into β_d + 1 connected components. Exactly one of these connected components is unbounded, and we refer to the unbounded component as V_∞. Using this partition we define the dual graph K* of K. K* has one vertex for each connected component of R^{d+1} \ K with the vertex corresponding to V_∞ denoted by v_∞. Further, K* has one edge for each d-simplex in K. There is an edge between two vertices representing connected components V_1 and V_2 in K if there is a d-simplex contained in the intersection of the topological closures of V_1 and V_2. Note that K* can have parallel edges and self-loops. Since each d-simplex can be in the closure of at most two connected components we have a one-to-one correspondence between d-simplices in K and edges in K*.

If S is a set of d-simplices in K we denote their corresponding edges in K* by S*. Similar to planar graphs, there is a duality between d-cycles in K and edge cuts in K*. There exists a one-to-one correspondence between d-cycles in K and minimal edge cuts in K*. We refer to this correspondence as cycle/cut duality, and it will play a central role in many of our proofs.

By shell(K) we denote the outer shell of K. This is defined to be the subcomplex of K consisting of all d-simplices whose corresponding edges in K* are incident to v_∞. Equivalently, it is also the subcomplex of K consisting of all d-simplices contained in the boundary of V_∞.

We endow the embedding of a simplicial complex K with the subspace topology inherited from R^{d+1}. We call K a d-dimensional manifold if every point in its embedding is contained in a neighborhood homeomorphic to R^d. If every point in the embedding of K is contained in a neighborhood homeomorphic to either R^d or the d-dimensional half-space we call K a manifold with boundary.
2.4 Graph cuts

Let $G = (V, E)$ be a graph. For any two subsets $V_1, V_2 \subset V$ a $(V_1, V_2)$-cut is a set of edges $E'$ such that the graph $G' = (V, E \setminus E')$ contains no path from $V_1$ to $V_2$. Often we will consider $(S, \overline{S})$-cuts for some $S \subset V$ where $\overline{S}$ denotes the complement of $S$ in $V$. By $E_S$ we refer to the edge set corresponding to all edges that have one endpoint in $S$ and the other in $\overline{S}$, which is the minimum $(S, \overline{S})$-cut. We extend this notation to vertices. For any two vertices $s, t \in V$ an $(s, t)$-cut refers to a set of edges whose removal disconnects $s$ from $t$.

2.5 The minimum bounded/homologous chain problems

Now we give the formal statement of the minimum bounded chain problem. Given a $d$-dimensional simplicial complex $K$ and a $(d-1)$-cycle $C$ contained in $K$ the minimum bounded chain problem $(K, C)$ asks to find a $d$-chain $X$ such that the cost of $X$ is minimized. The cost of $X$ is given by its $\ell_1$ norm $\|X\|_1$. Here we are treating $X$ as an $n$-dimensional indicator vector where $n$ is the number of $d$-simplices in $K$. The simplicial complex $K$ may be weighted by assigning a real number to each $d$-simplex in $K$. In this case the cost of $X$ is given by $\langle W, X \rangle$, where $W$ is a vector assigning weights to the $d$-simplices of $K$.

Now let $D$ be a $d$-chain, which may or may not be a cycle. The minimum homologous chain problem asks to find a minimum $d$-chain $X$ such that $X = D \oplus V$ for some $(d+1)$-chain $V$, equivalently, the minimum $d$-chain $X$ such that $D \oplus X$ is null-homologous. The cost of $X$ as well as the weighted problem are defined the same as in the previous paragraph.

In this paper, we study the minimum bounded chain problem for complexes embedded in $\mathbb{R}^{d+1}$, and the minimum homologous chain problem for $d$-chains in $(d+1)$-manifolds.

3 Approximation algorithm and fixed-parameter tractability

In this section, we describe approximation algorithms and parameterized algorithms for both minimum bounded chain and minimum homologous chain problems. Our algorithms work with the dual graph of the input space. In order to simplify our presentation we assume that the dual graph of the input complex contains no loops. The following lemma shows that we can make this assumption without any loss of generality. The proof can be found in the full version of the paper.

- **Lemma 7.** In polynomial time we can preprocess an instance of the minimum bounded chain problem $(K, C)$ into a new instance $(K', C')$ such that (i) $(K')^*$ contains no loops and (ii) an $\alpha$-approximation algorithm for $(K', C')$ implies an $\alpha$-approximation algorithm for $(K, C)$.

3.1 Reductions to the minimum cut completion problem

Given $G = (V, E)$ and $E' \subseteq E$, the minimum cut completion problem asks for a cut $(S, \overline{S})$ with edge set $E_S$ that minimizes $|E_S \oplus E'|$. First, we show that the minimum cut completion problem generalizes the minimum bounded chain problem.

- **Lemma 8.** For any $d$-dimensional instance of the minimum bounded chain problem, $(K, C)$, there exists an instance of the minimum cut completion problem $(G = (V, E), E')$ that can be computed in polynomial time, and a one-to-one correspondence between cuts in $G$ and $d$-chains with boundary $C$ in $K$. Moreover, if the cut $(S, \overline{S})$ with edge set $E_S$ in $G$ corresponds to the $d$-chain $Q$ in $K$ then $|E_S \oplus E'| = |Q|$.
Proof. Let $F$ be any $d$-chain such that $\partial F = C$, such an $F$ can be computed in polynomial time, by solving the linear system. In turn, let $G = K^*$, and $E' = F^*$.

Now, let $Q$ be any $d$-chain such that $\partial Q = C$. So, $\partial (Q \oplus F) = 0$. Thus, by cycle/cut duality $Q \oplus F$ partitions $\mathbb{R}^{d+1}$; let $(S, \overline{S})$ be the corresponding dual cut in $K^*$, and let $E_S$ be the edge set of this cut. We have $|E_S \oplus E'| = |E_S \oplus F^*| = |Q^*| = |Q|$.

On the other hand, let $(S, \overline{S})$ be a cut in $K^*$, with edge set $E_S$. By cycle/cut duality $\partial E^*_S = 0$. Now, let $Q = E^*_S \oplus F$. It follows that $\partial Q = C$. Moreover, we have $|Q| = |E_S \oplus F| = |E_S \oplus F^*| = |E^*_S \oplus E'|$. ▶

We show via a similar argument that the cut completion problem also generalizes the minimum homologous chain problem when the input complex is a weak pseudomanifold (see the full version of the paper for the proof). A weak pseudomanifold is a pure $d$-complex such that every $(d - 1)$-simplex is a face of at most two $d$-simplices. Weak pseudomanifolds generalize manifolds and the definition was first introduced by Dey et al. in [14]. Although recognizing $d$-manifolds is undecidable [12], weak pseudomanifolds can be recognized in polynomial time.

▶ Lemma 9. For any $d$-dimensional instance of the minimum homologous chain problem $(\mathcal{M}, D)$, where $\mathcal{M}$ is a weak pseudomanifold, there exists an instance of the minimum cut completion problem $(G = (V, E), E')$ that can be computed in polynomial time, and a one-to-one correspondence between cuts in $G$ and $d$-chains in $\mathcal{M}$ that are homologous to $D$. Moreover, if the cut $(S, \overline{S})$ with edge set $E_S$ in $G$ corresponds to the $d$-chain $Q$ in $\mathcal{M}$ then $|E_S \oplus E'| = |Q|$.

3.2 Algorithms for the minimum cut completion problem

We show an $O(\sqrt{\log |V|})$-approximation algorithm and a fixed-parameter tractable algorithm for the cut completion problem. We obtain both of these results via reduction to 2CNF Deletion: given an instance of 2SAT, find the minimum number of clauses to delete to make the instance satisfiable. Agarwal et al. [1] show an $O(\sqrt{\log n})$-approximation algorithm for 2CNF Deletion, where $n$ is the number of clauses, and Razgon and O’Sullivan show that the problem is fixed-parameter tractable.

▶ Lemma 10 (Agarwal et al.[1], Theorem 3.1). There is a randomized polynomial-time algorithm for finding an $O(\sqrt{\log n})$-approximation for the minimum disagreement 2CNF Deletion problem.

▶ Lemma 11 (Razgon and O’Sullivan [27], Theorem 7). Let $B$ be an instance of 2CNF Deletion problem with $m$ clauses that admits a solution of size $k$. There is an $O(15^k \cdot k \cdot m^3)$ time exact algorithm for solving $B$.

The next lemma shows similar results for the cut completion problem.

▶ Lemma 12. For the cut completion problem $(G = (V, E), E')$,

(i) there is a randomized polynomial-time $O(\sqrt{\log |V|})$-approximation algorithm, and

(ii) there is an $O(15^k \cdot k \cdot |E|^3)$ time exact algorithm, where $k$ is the size of the optimal solution.

Proof. Let $G = (V, E)$, and $E' \subseteq E$. We show a 2CNF Deletion instance $B_G$ such that for any cut $(S, \overline{S})$ with edge set $E_S$, the number of unsatisfied clauses in $B_G$ is exactly $|E_S \oplus E'|$. The statement of the lemma will follow from Lemma 10 and 11.
Minimum Bounded Chains

Let $B_G$ be the instance of the 2CNF Deletion problem defined on $G$ as follows:

- For each vertex $v \in V$, we have variable $b(v)$.
- For each edge $(u, v) \in E$:
  - if $(u, v) \in E'$, we add $b(u) \lor b(v)$ and $\neg b(u) \lor \neg b(v)$ to $B$, and
  - if $(u, v) \notin E'$, we add $b(u) \lor \neg b(v)$ and $\neg b(u) \lor b(v)$ to $B$.

(Note that in both cases, any assignment of $b(u)$ and $b(v)$ satisfies at least one of the clauses. Again in both cases, assignments exist that satisfy both clauses.)

Let $(S, \overline{S})$ be a cut with edge set $E_S$. Let $b_S$ be the natural boolean vector that corresponds to the cut: $b(v) = [v \in S]$ for all $v \in V$. We show that $|E_S \oplus E'|$ is equal to the number of clauses that are not satisfied in $B_G$. Specifically, we show (I) for each edge $(u, v) \in E_S \oplus E'$, exactly one of its corresponding clauses is satisfied, and (II) for each edge $(u, v) \notin E_S \oplus E'$ both of its corresponding clauses are satisfied.

If $(u, v) \in E_S \oplus F$ there are two cases to consider: (I.1) $(u, v) \in E_S$ and $(u, v) \notin E'$, that is $b(u) \neq b(v)$ and the corresponding clauses are $b(u) \lor \neg b(v)$ and $\neg b(u) \lor b(v)$. Exactly one of the clauses is satisfied. (I.2) $(u, v) \notin E_S$ and $(u, v) \in E'$, that is $b(u) = b(v)$, and the corresponding clauses are $b(u) \lor b(v)$ and $\neg b(u) \lor \neg b(v)$; exactly one of the clauses is satisfied.

If $(u, v) \notin E_S \oplus E'$ there are two cases to consider: (II.1) $(u, v) \in E_S$ and $(u, v) \in E'$, that is $b(u) \neq b(v)$ and the corresponding clauses are $b(u) \lor b(v)$ and $\neg b(u) \lor \neg b(v)$. Both of the clauses are satisfied. (II.2) $(u, v) \notin E_S$ and $(u, v) \notin E'$, that is $b(u) = b(v)$, and the corresponding clauses are $b(u) \lor \neg b(v)$ and $\neg b(u) \lor b(v)$. Both of the clauses are satisfied.

3.3 Wrap up (Proofs of Theorems 1, 2, 3, and 4)

Lemma 8 and Lemma 9 show that the bounded chain problem and the minimum homological chain problem are special cases of the cut completion problem, and Lemma 12 shows that we obtain $O(\sqrt{\log |V|})$-approximation algorithm and $O(15^k \cdot k \cdot |E|^3)$ time exact algorithm for the cut completion problem. The number of vertices $|V|$ translates to $\beta_d$ for simplicial complexes embedded in $\mathbb{R}^{d+1}$ (Theorem 1), and $n_{d+1}$, the number of $(d + 1)$-dimensional simplices for $(d + 1)$-manifolds (Theorem 3). The number of edges $|E|$ translates to $n_d$ in both simplicial complexes embedded in $\mathbb{R}^{d+1}$ (Theorem 2) $(d + 1)$-manifolds (Theorem 4).

4 Hardness of approximation

In this section, we show it is unlikely that either of the minimum bounded chain or minimum homologous chain problems admit constant factor approximation algorithms, even for their low dimensional instances. Our hardness results follow from reductions from the minimum cut completion problem, defined in the previous section.

4.1 Minimum bounded chain to minimum cut completion

We show that the minimum cut completion problem reduces to a 2-dimensional instance of the minimum bounded chain problem $(K, C)$, where $\text{shell}(K)$ is in fact a manifold and $C$ is a (possibly not connected) cycle on $\text{shell}(K)$. Our hardness of approximation result for the minimum bounded chain problem is based on this reduction.

Lemma 13. Let $(G = (V, E), E')$ be any instance of the minimum cut completion problem. There exists an instance of the 2-dimensional minimum bounded chain problem $(K, C)$ with $C$ on the outer shell of $K$ that can be computed in polynomial time, and a one-to-one correspondence between cuts in $G$ and 2-chains with boundary $C$ in $K$. Moreover, if the cut $(S, \overline{S})$ with edge set $E_S$ in $G$ corresponds to the 2-chain $Q$ in $K$ then...
\[
\frac{|Q|}{\tau} - 1 \leq |E_S \oplus E'| \leq \frac{|Q|}{\tau},
\]
where \( \tau = 58m + 2 \) and \( m \) is the number of edges in \( G \).

Proof. Our construction is simple in high-level. We start from any embedding of \( G \) in \( \mathbb{R}^3 \), and we thicken it to obtain a space, in which each edge corresponds to a tube. We insert a disk in the middle of each tube; we call these disks edge disks. Then we triangulate all of the 2-dimensional pieces. The dual of the complex that we build is almost \( G \), except for one extra vertex corresponding to its outer volume, and a set of extra edges, all incident to the extra vertex. We give our detailed construction below.

We consider the following piecewise linear embedding of \( G \) in \( \mathbb{R}^3 \); let \( n \) and \( m \) be the number of vertices and edges of \( G \), respectively. First, map the vertices of \( G \) into \( \{1, 2, \ldots, n\} \) on the \( x \)-axis. Now, consider \( m + 2 \) planes \( h_0, h_1, \ldots, h_{m+1} \) all containing the \( x \)-axis with normals being evenly spaced vectors ranging from \((0, 1, 1)\) to \((0, 1, -1)\). We use \( h_1, \ldots, h_m \) for drawing the edges \( G \). We arbitrarily assign edges of \( G \) to these plane, so each plane will contain exactly one edge. Each edge is drawn on its plane as a three-segment curve; the first and the last segment are orthogonal to \( x \)-axis and the middle one is parallel. All edges are drawn in the upper half-space of \( \mathbb{R}^3 \). See Figure 1, left.

Next, we place an axis parallel cube around each vertex. The size of the cubes must be so that they do not intersect, fix the width of each cube to be \( 1/10 \). We refer to these cubes as vertex cubes. Then, we replace the part of each edge outside the cubes with a cubical tube, called edge tube. We choose the thickness of these tubes sufficiently small so that they are disjoint. We also puncture the cubes so that the union of all vertex cubes and edges tubes form a surface; see Figure 1, left. (This surface will have genus \( m - n + 1 \) by Euler’s formula, which is the dimension of the cycle space of \( G \)).

Next, we subdivide each tube by placing a square in its middle; see Figure 1, right. We refer to these squares as edge squares. Edge squares partition the inside of the surface into \( n \) volumes. We observe that each of these volumes contains exactly one vertex of the drawing of \( G \), thus, we call them vertex volumes.

For our reduction to work, we need that the weight of each 2-cycle to be dominated by the weight of its edge squares. To achieve that we finely triangulate each edge square. For an edge tube, we first subdivide its surface to 16 quadrangles as shown in Figure 2, left. Then, we obtain a triangulation with 32 triangles by splitting each quadrangle into two triangles.
For a vertex cube, note that all the punctures are on the top face by our construction. We split all the other faces by dividing each of them into two triangles. For the top face, we can obtain a triangulation in polynomial time; this triangulation will have $4 \deg(v) + 8$ triangles by Euler’s formula, where $\deg(v)$ is the degree of the vertex corresponding to the cube. Therefore, the triangulation of each vertex cube will have $4 \deg(v) + 18$ triangles, see Figure 2, right. Therefore, there are $\left( \sum_{v \in V} 4 \deg(v) + 18 \right) + 32m \leq 58m$ triangles that are not part of edge squares. Finally, we triangulate each edge square into $58m + 2$ triangles so that the cost of one edge square is greater than the sum of all triangles not contained in edge squares. This triangulation can be done by starting with a square made up of two triangles and repeatedly subdividing triangles by inserting a new vertex in the interior and connecting it to the corners with edges. The subdivision is performed by inserting a vertex into the interior of the triangle and connecting it with an edge to each vertex on the boundary of the triangle. The result is a new complex, homeomorphic to the original, with two additional triangles. Overall, our complex $K$ has $O(m^2)$ triangles.

![Figure 2](image-url) Left: subdividing the surface of an edge-tube to quadrangles, right: triangulating the surface of a vertex cube.

We are now done with the construction of $K$. Let $B$ be the set of all triangles in edge squares that correspond to edges in $E'$. Then, let $C = \partial B$. We show an almost cost preserving one-to-one correspondence between cuts in the cut completion problem in $G$ and chains with boundary $C$ in $K$.

Let $(S, \overline{S})$ be a cut with edge set $E_S$, note that the cost of this cut is $|E_S \oplus E'|$ in the cut completion problem $(G, E')$. In $K$, let $V_S$ be the symmetric difference of the vertex volumes that correspond to vertices of $S$. The total weight of $V_S$ is between $|E_S|(58m + 2) + |E_S||(58m + 2) + 58m$. Similarly, the total weight of $V_S \oplus B$ is between $|E_S \oplus E'| (58m + 2)$ and $|E_S \oplus E'| (58m + 2) + 58m$. Since we cannot get an exact count on the number of edges in the subgraph induced by $S$ we have a range of values for the weight of $V_S$ instead of an exact weight. However, if $E_S$ and $E_S'$ are two cuts with $|E_S| < |E_S'|$ then the weight of $V_S$ is strictly less than the weight of $V_{S'}$ by the construction of the edge squares.

On the other hand, let $Q$ be a 2-chain with boundary $C$ in $K$. As $C$ does not intersect the interior of any edge square, for each edge square either $Q$ contains all of its triangles or none of them. Also, $Q \oplus B$ has no boundary, thus its complement $\mathbb{R}^3 \setminus (Q \oplus B)$ is disconnected. The interior of each vertex volume is completely inside one of the connected components of $\mathbb{R}^3 \setminus (Q \oplus B)$, as by the construction $Q \oplus B$ must either contain the entire vertex volume or none of it. Now, let $S$ be the set of all vertices whose corresponding vertex volumes are in the unbounded connected component of $\mathbb{R}^3 \setminus (Q \oplus B)$. The edges of the cut $(S, \overline{S})$ correspond to edge squares in $Q_s \oplus B$, where $Q_s$ is the set of edge square triangles of $Q$. As $B$ is in one-to-one correspondence to $E'$, it follows that the cut completion cost of $(S, \overline{S})$ is $\frac{|Q_s|}{58m + 2}$. 
We have $|Q| = |Q_s| + |Q_r|$ where $Q_r$ is the set of triangles in $Q$ not contained in edge squares. The size of $|Q_s|$ is $58m + 2$ per edge square, and $|Q_r| \leq 58m$ by construction. It follows that we have our desired inequality,

$$\frac{Q}{58m + 2} - 1 \leq |E_S \oplus E'| \leq \frac{Q}{58m + 2}.$$  

The next lemma shows that an approximation algorithm for the minimum bounded chain problem implies an approximation algorithm with almost the same quality for the minimum cut completion problem.

\begin{lemma}
Let $(G = (V, E), E')$ be any instance of the minimum cut completion problem. For any $\alpha \geq 1$ and any $\varepsilon > 0$, there exists an instance of the 2-dimensional minimum bounded chain problem $(K, C)$ that can be computed in polynomial time, such that an $\alpha$-approximation algorithm for $(K, C)$ implies a $(1 + \varepsilon)\alpha$-approximation algorithm for $(G, E')$, and $C$ is on the outer shell of $K$.
\end{lemma}

\begin{proof}
Let $\varepsilon > 0$. Given an $\alpha$-approximation algorithm for the minimum bounded chain problem, we describe an $((1 + \varepsilon)\alpha)$-approximation algorithm for the cut completion problem.

Let $G = (V, E)$, and $E' \subseteq E$ be any instance of the cut completion problem, and let $(S_{opt}, S_{opt}')$ be the corresponding 2-chain to $(S_{opt}, S_{opt}')$ in $K$. Thus, $\frac{Q_{\text{opt}}}{\tau} - 1 \leq |E_{S_{opt}} \oplus E'| \leq |Q_{\text{opt}}|$. In addition, let $Q$ be the surface with boundary $C$ that the $\alpha$-approximation algorithm finds, so $|Q| \leq \alpha \cdot |Q_{\text{opt}}|$. Finally, let $(S, S)$ be the cut corresponding to $Q$ in $G$ via the one-to-one correspondence of Lemma 13. Therefore, $\frac{Q}{\tau} - 1 \leq |E_S \oplus E'| \leq \frac{|Q|}{\tau}$. Putting everything together,

$$|E_S \oplus E'| \leq \frac{|Q|}{\tau} \leq \alpha \cdot \frac{|Q_{\text{opt}}|}{\tau} \leq \alpha \cdot (|E_{S_{opt}} \oplus E'| + 1).$$

(1)

Since $|E_{S_{opt}} \oplus E'| \geq 1/\varepsilon$, we have: $|E_{S_{opt}} \oplus E'| + 1 \leq (1 + \varepsilon) \cdot |E_{S_{opt}} \oplus E'|$. Therefore, together with (1), we have a $((1 + \varepsilon)\alpha)$-approximation algorithm, as desired.
\end{proof}

\section{Minimum homologous cycle to minimum cut completion}

We show a similar reduction from the cut completion problem to the minimum homologous cycle problem for 1-dimensional cycles on orientable 2-manifolds. The minimum homologous cycle problem is the special case of the minimum homologous chain problem when the input chain is required to be a cycle, so showing hardness of approximation for it implies hardness of approximation for the more general minimum homologous chain problem.

\begin{lemma}
Let $(G = (V, E), E')$ be any instance of the minimum cut completion problem. For any $\alpha \geq 1$, there exists an instance of the 1-dimensional minimum homologous cycle problem $(M, D)$ that can be computed in polynomial time such that an $\alpha$-approximation for $(M, D)$ implies an $\alpha$-approximation for $(G, E')$.
\end{lemma}
Proof. We construct a 2-manifold $M$ as in the proof of Lemma 13, but we omit the edge squares. Each edge of $G$ corresponds to a cycle with 4 edges in $M$; these cycles are the boundaries of the omitted edge squares. We call these cycles edge rings. The connected components of $M$ after removing the edge rings correspond to the vertices of $G$, we call these connected components vertex regions. We set $D$ to be equal to the set of edge rings corresponding to $E'$. Intuitively, if $X$ is the minimum cycle homologous to $D$ we do not want $X \oplus D$ to intersect the interior of any vertex region. That is, $X \oplus D$ is a collection of edge rings and corresponds to a cut in $G$. To achieve this, we subdivide each edge not contained in an edge ring into a long path. The result is an embedded graph with non-triangular faces, which is not a simplicial complex. To fix this, we triangulate the inside of each non-triangular face such that the shortest path between any two vertices on the face remains the shortest path after the triangulation. Given any $\alpha$-approximation of the new complex we can obtain a smaller solution using only the edge rings, which corresponds to a cut in $G$. Our formal construction follows.

Let $\tau = 4\lceil \alpha \rceil |E| + 1$; we subdivide each edge not contained in an edge ring $\tau$ times. For each face of length $\ell > 3$ we triangulate by adding $\ell + 1$ concentric cycles, each with $\ell$ vertices, labeled $\gamma_0, \ldots, \gamma_\ell$, where $\gamma_0$ is the original face from the subdivided version of $M$. By $v_{i,j}$ we denote the $j$th vertex in $\gamma_i$. We add the edges $(v_{i,j}, v_{i,j+1})$ and $(v_{i,j}, v_{i,j+1 \mod \ell})$. To complete the triangulation we add one additional vertex $v$ at the center of $\gamma_\ell$ and add an edge between it and each vertex on $\gamma_\ell$. We call the new simplicial complex $M'$. See Figure 3 for an example.

![Figure 3](image)

Figure 3 Subdividing a face of length five; the outer face with white vertices is the original face.

Let $(S_{\text{opt}}, S_{\text{opt}}')$ be an optimal solution to the minimum cut completion instance $(G, E')$. Suppose we can compute an $\alpha$-approximation $C$ of the minimum homologous cycle instance $(M', D)$, hence $|C| \leq \alpha |C_{\text{opt}}|$. By our construction an optimal solution to $(M', D)$ has the same size as an optimal solution to $(M, D)$. As $C$ is a cycle, if $C$ crosses a cycle $\gamma_0$ it must cross it an even number of times. For any two consecutive vertices $u, v \in \gamma_0$ in $C$ we replace the path between them with the shortest path contained in $\gamma_0$. We call the new cycle $C'$, since $C' \leq C$ we have that $C'$ is also an $\alpha$-approximation for $(M', D)$. Note that $C'$ is a union of edge rings, otherwise $|C'| > \alpha |C_{\text{opt}}|$. It follows that $C'$ corresponds to a cut $E_{S'}$ with $|C'| = 4|E_{S'} \oplus E'|$. Hence, we have $|E_{S'} \oplus E'| \leq \alpha |E_{S_{\text{opt}}'} \oplus E'|$. Thus, $E_{S'}$ is an $\alpha$-approximation for $(G, E')$. 

4.3 Wrap up

It remains to show that the cut completion problem is hard to approximate. We show this via a straightforward reduction from the minimum uncut problem: given a graph $G = (V, E)$, find a cut with minimum number of uncut edges. Note that the optimal cuts for the minimum uncut problem and the maximum cut problem coincide, yet, approximation algorithms for one problem do not necessarily imply approximation algorithm for the other one.
Lemma 16. The minimum uncut problem is a special case of the minimum cut completion problem.

Proof. Consider the cut completion problem for \( G = (V, E) \), and let \( E' = E \). Let \((S, \overline{S})\) be any cut with edge set \( E_S \). The cut completion cost of this cut is

\[
|E_S \oplus E'| = |E_S \oplus E| = |E \setminus E_S|,
\]

which is the number of uncut edges by \((S, \overline{S})\).

Now, we are ready to prove our hardness results.

Proof of Theorem 5 and 6. The minimum uncut problem is hard to approximate within \((1 + \varepsilon)\) for some \( \varepsilon > 0 \) [26]. In addition, it is hard to approximate within any constant factor assuming the unique games conjecture [24, 21, 10, 22]. By Lemma 16, the cut completion problem generalizes the minimum uncut problem. Finally, by Lemma 15 and 14, for any \( \alpha > 1 \) and \( \varepsilon > 0 \), an \( \alpha \)-approximation algorithm for the minimum bounded chain problem or the minimum homologous cycle problem implies a \((1 + \varepsilon)\alpha\)-approximation algorithm, or an \( \alpha \)-approximation for the cut completion problem, respectively.

References


21:14 Minimum Bounded Chains


