Fakultät für Mathematik
Technische Universität München

Technische Universität München

Fakultät für Mathematik

Algorithms in Discrete Morse Theory and Combinatorial Topology

Abhishek Rathod

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Prof. Dr. Stefan Weltge
Prüfer der Dissertation:	1.	Prof. Dr. Ulrich Bauer
	2.	Prof. Benjamin Burton
	3.	Prof. Herbert Edelsbrunner

Die Dissertation wurde am 25.10.2021 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 06.01.2022 angenommen.

ABSTRACT

This dissertation is a complexity theoretic study of well-known problems in combinatorial topology. In the first part, an open question concerning the (in)approximability of Morse matching is resolved, and existing results concerning the parameterized complexity of Morse matching are improved upon. The first part also provides a complexity-theoretic explanation for the effectiveness of the apparent pairs gradient. In the second part, certain natural problems in simple homotopy theory are shown to be hard from the point of view of parameterized complexity. The third part describes asymptotically fastest known algorithms for computing minimum 1-homology bases of simplicial complexes by exploiting the matroid structure of homology bases. In the fourth and the final part, the notion of cuts is generalized from graphs to simplicial complexes, and the parameterized complexity of these high-dimensional cuts is studied. The final part also provides a polynomial time algorithm for a high-dimensional cut problem in the special case of surfaces. The unifying theme across the four parts is the value of complexity theory in the study of problems in combinatorial topology.

ZUSAMMENFASSUNG

Diese Dissertation ist eine theoretische Studie der Komplexität bekannter Probleme der kombinatorischen Topologie. Im ersten Teil beantworten wir eine bisher offene Frage zur (mangelhaften) Approximierbarkeit von Morse Paarungen. Weiter verbessern wir bereits vorhandene Ergebnisse zur parametrisierten Komplexität von Morse Paarungen. Der erste Teil beinhaltet auch eine Erklärung der Wirksamkeit des Gradienten der sichtbaren Paare. Im zweiten Teil zeigen wir, dass natürliche Probleme aus der "einfachen Homotopietheorie" aus Sicht der parametrisierten Komplexitätstheorie schwierig sind. Im dritten Teil wird der Algorithmus mit der bisher besten asymptotischen Laufzeit zur Berechnung optimaler Basen der ersten Homologie simplizialer Komplexe beschrieben. Dazu nutzen wir die Struktur eines Matroiden auf der Menge der Basen der Homologie. Im vierten und letzten Teil verallgemeinern wir den Begriff eines Schnitts von Graphen auf simpliziale Komplexe und untersuchen die parametrisierte Komplexität dieser höher-dimensionalen Schnitte. In diesem letzten Teil beschreiben wir auch einen Algorithmus für ein höher-dimensionales Schnittproblem auf Flächen mit polynomieller Laufzeit. Das verbindende Thema der vier Teile ist der Wert der Komplexitätstheorie bei der Untersuchung von Problemen in der kombinatorischen Topologie.

In the depth of winter,
I learned that within me,
there lay an invincible summer.

Summer
Albert Camus

ACKNOWLEDGMENTS

It has been an absolute privilege to work with my PhD advisor, Uli Bauer. Working with you has always been inspiring and rewarding in so many ways!

- I am deeply grateful for the fact that you gave me full freedom to choose my research topics right from the very beginning of my PhD.
- I am really happy that this apprenticeship was fruitful, and we managed to collaborate on four finished (and two ongoing) projects.
- Throughout, you have been generous with your time and forthright with your advice. At every setback, you have been gracious, kind and supportive.
- Everything I know about mathematical writing, and scientific talk preparation, I have learnt from you. I have learnt that there's always a better way to say something no matter what else the clock or the calendar says.

I am deeply indebted to you for the exceptional guidance at every turn, for the unwavering support, and for the many opportunities I had at TUM to grow and evolve as a researcher. Thank you for showing me how it is done by being an exemplar of an academic and a mathematician. In sum, for the things that I can say, and for the things that I can't, "Thank you so much for *everything*!"

Next, I would like to thank Michael Lesnick, my academic mentor during the PhD years. I have the highest regard for your generosity and your friendship during this journey. Also, thank you for suggesting a great project for collaboration, and for initiating me into the rich world of multi-parameter persistence. It's a pity we couldn't finish it in time for my thesis.

Next, I would like to thank Tamal Dey for offering me a position in his research group. I have known you for many years, and I have always wanted to work with you. So I am really thrilled about this wonderful opportunity, and the exciting times ahead at Purdue in the year 2022. Thank you so much!

Next, I would like to thank Meirav Zehavi for offering me organizing a research visit, and in fact, collaborating with me on one of my PhD projects. I truly appreciate your kindness, warmth, and generosity. Working with you, and learning from you has been one of my most rewarding academic experiences – an experience I will cherish. Thank you so much! I look forward to being in Israel.

My deepest thanks to the members of the thesis committee: the chair of the committee Stefan Weltge, and the external examiners Herbert Edelsbrunner and Benjamin Burton for agreeing to read the thesis at short notice. I am *immensely grateful* for taking my timelines into account. I simply can't thank you enough.

I would like to thank Vijay Natarajan for advising me when I was working at the Indian Institute of Science. Thank you for collaborating with me on my first published paper. I will always be grateful for your help and support.

I would like to thank Jonathan Spreer for introducing me to parameterized complexity, and for suggesting one of the problems that is included in this thesis. Thank you for suggesting a reading list for 3-manifold topology, for being such an amazing collaborator, and for being supportive of my work.

I would like to thank Talha Bin Masood, who's my oldest collaborator and a very close friend.

I would like to thank Claudia Scheimbauer for introducing me to some interesting mathematical areas, and for sharing her perspective on recent advances in mathematics.

Many thanks to the members of the M10+ Lehrstuhl: Jürgen Richter-Gebert, Tim Hoffman, Hermann Vogel, Johann Hartl, Jutta Niebauer, Magnus, Florian, Tashi, Michael S., Jannik, Vanessa, and Alex P. for the warmth, the friendliness and the many shared meals. Thank you Carsten for the many German lessons punctuated with laughter. Thank you Bernhard for sharing the title page of your thesis. Thank you Fabian R. + Eilind for organizing the beer & barbecue meetups. Thank you Andrew N. + Erika for the discussions on random topology. Thank you Alex R. for providing some high-level feedback on my thesis. Thank you Ben and Fabians for the many math-related discussions, and thank you David for the several weeks of poring over Farber's "Topology of closed one-forms" together. Thank you Lena for covering for me multiple times, and thank you Lena + Aaron for planning the hikes and the board games. Thank you Eilind for the many fun conversations, and for providing feedback on my thesis. Thank you Ben, Lena and Zi for helping me beim Umzug, and for always being there for me – You guys are the best!

It has been a great pleasure to share the office space with Fabian L., a true-blue Renaissance man and a very dear friend. Thank you Fabian for inspiring me with your passion for mathematics, your zest for life, and your many pursuits outside of academia. Thank you for the many shared laughs, conversations and memories! And thank you so much for the careful perusal of my thesis and the detailed feedback! I really appreciate it.

Thank you Michael H. for inviting me to stay with your family to celebrate a very German Christmas. That was truly special. Thank you Harsha for staying true to your name by being my most cheerful friend – it is great to laugh with someone who really gets it! Finally, I would like to take this opportunity to thank *all* my non-TUM, non-academic friends: you know who you are, and you know how much I value you. *Dil Se*.

I would like to thank all the administrative staff at TUM. But first, a very, very special thanks to Isabella Wiegand and Diane Clayton-Winter for going out of the way to help me on too many occasions to count! And finally, many thanks to Lydia Weber for making the daunting process of thesis submission smooth and tractable, and for being so kind and supportive!

My deepest thanks to my parents for the several years of affection, compassion and toil they put in raising me, and educating me. I am really grateful for your care and support.

Finally, I would like to thank my wife Jital for being perfect to a fault. Thank you for the enduring love and the quiet strength. Thank you for making these years of my life rich, beautiful, and meaningful. Paraphrasing Blaise Pascal, "When one does not love too much, one does not love nearly enough."

LIST OF CONTRIBUTED ARTICLES FOR THE THESIS

The thesis is based on the following articles.

CORE ARTICLES AS PRINCIPAL AUTHOR

I Ulrich Bauer and Abhishek Rathod.
Hardness of Approximation for Morse Matching.
In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '19, pages 2663–2674.
(See article [13] in the bibliography.)

II Ulrich Bauer, Abhishek Rathod, and Jonathan Spreer.
Parametrized Complexity of Expansion Height.
In *27th Annual European Symposium on Algorithms* (ESA 2019), volume 144, pages 13:1–13:15.
(See article [15] in the bibliography.)

III Abhishek Rathod.
Fast Algorithms for Minimum Cycle Basis and Minimum Homology Basis.
In *36th International Symposium on Computational Geometry* (SoCG 2020), volume 164, pages 64:1–64:11.
(See article [136] in the bibliography.)

FURTHER ARTICLES AS PRINCIPAL AUTHOR

I Ulrich Bauer and Abhishek Rathod.
Parameterized Inapproximability of Morse Matching.
arXiv preprint arXiv:2109.04529, 2021.
Submitted to *CGTA Special Issue on Algorithmic Aspects of Computational and Applied Topology*
(See article [14] in the bibliography.)

II Ulrich Bauer, Abhishek Rathod, and Meirav Zehavi.
The Complexity of High-dimensional Cuts.
arXiv preprint arXiv:2108.10195, 2021.
To be submitted to a journal this year.
(See article [16] in the bibliography.)

I, Abhishek Rathod, am the principal author of all of the articles above.

ARTICLES THAT ARE NOT A PART OF THIS THESIS

Abhishek Rathod, Talha Bin Masood, and Vijay Natarajan.
Approximation algorithms for Max Morse Matching.
In *Computational Geometry, 61:1 – 23, 2017.*
(See article [137] in the bibliography.)

CONTENTS

1

INTRODUCTION

Computational topology is a relatively new area of mathematical research [52,67, 101,120]. In fact, most of the advances in this field have happened only in the last two decades. From the outset, computational topology has been multi-faceted and interdisciplinary. While the mathematical foundations of computational topology are rooted in algebraic topology, geometric topology, commutative algebra, representation theory and category theory [12,24,29,30,35,48,87,102,111, 128], it also has metric, statistical and machine learning aspects [3,28,36,75,110, 138,144,153]. The applications are wide-ranging from neuroscience and material science to imaging and dynamical systems [22, 34, 49, 56, 68, 85, 91, 101, 135]. However, arguably the complexity theoretic aspect of computational topology remains somewhat under-explored. Since the computational problems that arise in computational topology reduce to either combinatorics or (linear) algebra, a complexity theoretic exploration of topological problems constitutes a viable approach. In particular, it is often of interest to design algorithms that come with some kind of theoretical guarantees or to be able to rule out an entire class of algorithms for a problem under certain standard complexity theoretic assumptions. Many of the important computational problems that naturally arise in computational topology, for instance, the computation of interleaving distance, are hard for most categories [4, 20, 21]. In these cases, it is often reasonable to look for approximate solutions or to design algorithms with respect to suitable parameters. A refined approach to computational complexity can help bridge the gap between theory and practice. The guiding philosophy of this thesis is to bring the full machinery of modern theoretical computer science, i.e., computational complexity theory, approximation algorithms, parameterized algorithms and randomized algorithms into play. The hope is to obtain a fine-grained understanding of subtle complexity theoretic aspects in computational topology. In this section, we expound on our results and place them in the right context.

In our view, *computational topology* has two branches: *topological data analysis* and *algorithmic topology*. Topological data analysis is concerned with distilling topological information from data. The continued success of topological data analysis relies on developing efficient software tools that extract meaningful summaries from topological objects associated to data. So, as in the case of machine learning, topological data analysis is more concerned with practical performance than with complexity guarantees. Algorithmic topology, on the other hand, is closer in spirit to computational geometry, where algorithms for many common problems do not have implementations, and theoretical guarantees are ubiquitous. The raison d'être of this thesis is to make a case for

theory, even in the context of practical topological data analysis, and what is more, a case for the ability of theoretical computer science to bring to the fore latent structures in mathematical objects. It is *this* spirit of algorithms in the service of mathematics that is embodied throughout this work, and functions as a Leitmotiv in what is otherwise, by design, a *topics thesis*.

For the most part, in this thesis, we use parameterized algorithms and approximation algorithms. In what follows, we provide a rationale for our choice. By way of introduction, we recall the Cobham-Edmonds thesis [43, 69], which asserts that an optimization problem Π is tractable if there exists an algorithm for Π with the following properties:

ACCURACY The algorithm computes the optimal solution.

EFFICIENCY The runtime of the algorithm is polynomial in instance size.

Since the development of the theory of **NP**-completeness, it became increasingly apparent that there exist a large class of problems for which the two requirements can not be simultaneously satisfied. As pointed out in a recent survey article [76], there are two ways to relax the above requirements in order to tame the hardness associated to **NP**-hard optimization problems.

On the one hand, if one relaxes the accuracy condition and allows solutions within a factor α of the optimal, one enters the territory of approximation algorithms. Additionally, it is sometimes convenient to design approximation algorithms that are randomized. Such algorithms may be either Monte-Carlo (output an α-approximation with high probability), or Las Vegas (run in expected polynomial time).

On the other hand, if one relaxes the efficiency condition and allows algorithms with runtime of the form $f(k)n^{O(1)}$, where n denotes the size of the instance, k denotes an upper bound on a certain parameter of interest that is expected to be small, and f is any computable (possibly super-polynomial) function, one enters the domain of parameterized algorithms.

By relaxing both requirements simultaneously, one allows parameterized α-approximation algorithms, which computes a solution that is an α-approximation in $f(k)n^{O(1)}$ time. The factor α may depend on the size of the instance, or it may be a function of the parameter. Both variants are in use, and will be encountered in this thesis.

It barely needs to be mentioned that the "worst case polynomial time paradigm" or its relaxations mentioned above are far from perfect. Probably the most famous counterexample is the simplex method for linear programming, which does not terminate in polynomial time on all instance families, but is still the method of choice for most instances in practice [19, 108]. There is a recent book that addresses shortcomings of worst case analysis of algorithms called *Beyond Worst-Case Analysis of Algorithms* [145], which collects several essays on alternative paradigms that turn out to be more effective in specific scenarios. Even in context of this thesis, in Paper IV, we had to use an average case version

of approximability to explain the effectiveness of the apparent pairs gradient on random complexes.

In this thesis, for the most part, we study the parameterized complexity and approximability of some fundamental problems in computational topology. A brief introduction to parameterized and approximation algorithms can be found in Chapter 2.

In the subsequent sections, we provide a topic-wise summary of our main findings. The four topics of investigation in this thesis are: the computational complexity of discrete Morse theory, simple homotopy theory, minimum homology basis and high-dimensional cuts.

1.1 DISCRETE MORSE THEORY

To begin with, we would like to point out that the principal topological object of study in this thesis is a simplicial complex, whose definition we recall here, in order to fix notation.

A k-*simplex* τ is the convex hull of $k+1$ affinely independent points P in the Euclidean space of some dimension $d \geqslant k$. Here, k is called the *dimension* of τ. Any nonempty subset of P also spans a simplex that we call a *face* of τ. We say that τ is a *coface* of a simplex σ if and only if σ is a face of τ. If the dimensions of σ and τ differ by one, then we say that σ is a *facet* of τ, and τ is a *cofacet* of σ. We use the notation $\sigma \prec \tau$ to indicate that σ is a facet of τ. Now, we can define (geometric) simplicial complexes.

A *simplicial complex* K is a collection of simplices that satisfies the following conditions:

- any face of a simplex in K also belongs to K, and
- every non-empty intersection of two simplices $\sigma_1, \sigma_2 \in K$ is a face of both σ_1 and σ_2.

The set of p-simplices of K is denoted by $K^{(p)}$ and the simplicial complex K_p induced by the simplices in $K^{(p)}$ is called the p-*skeleton* of K. The underlying space of K, denoted by $|K|$, is the union of simplices of K.

An *abstract simplicial complex* on a set of vertices V is a collection $\mathcal{C}$ of subsets of V that is closed under inclusion. The elements of $\mathcal{C}$ are called its *simplices*. The collection of vertex sets of simplices in a geometric simplicial complex forms an abstract simplicial complex. On the other hand, an abstract simplicial complex $\mathcal{C}$ has a geometric realization $|\mathcal{C}|$ obtained by embedding the points in V in general position in a Euclidean space of sufficiently high dimension. Then, the geometric simplicial complex associated to $\mathcal{C}$ is defined as $\bigcup_{\sigma\in\mathcal{C}} |\sigma|$, where $|\sigma|$ denotes the span of points in σ. It is not very difficult to show that any two geometric realizations of an abstract simplicial complex are homeomorphic. Hence, going forward, we do not distinguish between abstract and geometric simplicial complexes.

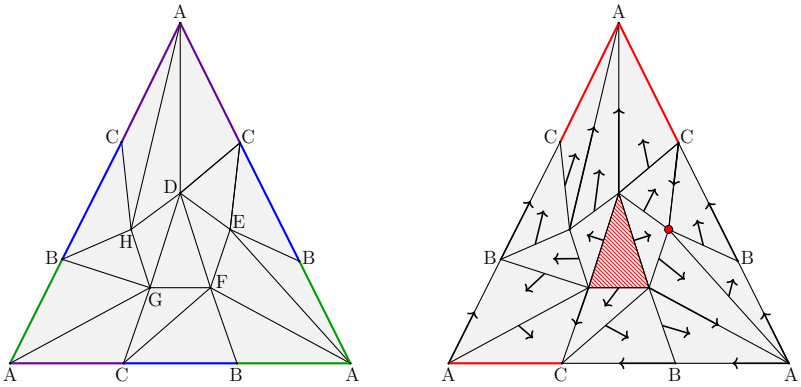


Figure 1: An optimal gradient vector field on a triangulated dunce hat. In the figure on the left, the edges along the boundary are glued as indicated by the labelling and the colors. In the figure on the right, the red colored simplices are the critical simplices.

Forman's discrete Morse theory [83,84] is an elegant combinatorial analogue of classical Morse theory, applicable to simplicial, and more generally CW complexes. In most cases, it provides a simple and effective method for substantially reducing the size of a complex while preserving its homotopy type. The most alluring quality of Forman's theory is that in spite of being extremely simple in its formulation, it affords immense theoretical and computational power.

Given a simplicial complex K, a simplex $\sigma \in K$ is said to be *free* if σ has a unique cofacet τ in K. An elementary collapse *pairs* a free face σ to its incident cofacet τ. Removing this pair of simplices gives rise to a smaller complex that is (simple) homotopy equivalent to the original one. However, what if one does not have any free faces in the complex? This happens to be the case with Zeeman's *dunce hat* [161], whose triangulation gives a simplicial complex that is contractible, but has no free faces, and is therefore not collapsible. In such cases, discrete Morse theory comes to the rescue by providing a theoretical framework for establishing homotopy equivalence through *internal collapses* that allow pairing of non-free faces. Informally, given a simplicial complex K, the collection of these internal collapses is called a discrete gradient vector field (or a *Morse matching*) on K. Please see Figure 1 for an illustration. The idea of viewing gradient pairs as internal collapses first appeared in Kozlov's exposition of discrete Morse theory [109]. For a brief introduction to discrete Morse theory, we refer the reader to Section 3.3.

In discrete Morse theory, it is possible that one starts with a simplicial complex and ends up with a (possibly non-regular) CW complex after executing a series of internal collapses. The fact that internal collapses induce simple homotopy equivalances in the cellular category was shown independently by Päkkilä [129] and Fernandez [79] in their respective thesis works. In particular, there is is an $(n+1)$-deformation from a CW complex K of dimension n

(equipped with vector field $\mathcal{V}$) to a Morse complex $\mathcal{M}_{\mathcal{V}}$ (with fewer cells). We briefly discuss their results in Section 3.4.

Now suppose that a simplicial complex L is contractible. Then, what are the obstructions to finding a gradient vector field that pairs all but one simplices of L? This question has a simple answer. The lack of free faces to begin with is a clear obstruction. Since the dunce hat is contractible, but not collapsible, any discrete gradient on a dunce hat has at least three critical simplices. We refer the reader to Figure 1 for an illustration.

Next, we ask a question that is a bit more refined. For instance, suppose that we begin with a 2-dimensional simplicial complex that does not have free faces. Then, upon the removal of a few critical simplices, new free faces (of lower dimension) may arise. The natural question here is: Can we always choose the simplices that are to be made critical wisely? In particular, is there a polynomial time algorithm for finding an optimal gradient vector field (that is, a vector field with fewest number of critical simplices) on a simplicial complex? Based on Lewiner's observations, and using a gadget by Eğecioğlu and Gonzalez [70], Joswig and Pfetsch [100] answered this question by showing that the problem is, indeed, **NP**-hard even for 2-complexes. Furthermore, they left the question of approximability of Morse matchings open. Since the problem is **NP**-hard, we study it from the point of view of parameterized complexity and approximability, and in fact, in Papers I and IV, resolve the open question posed by Joswig and Pfetsch.

Specifically, we study two variants of the Morse matching problem.

The Max-Morse Matching(MaxMM) can be described as follows: Given a simplicial complex K, compute a gradient vector field that maximizes the cardinality of matched (regular) simplices, over all possible gradient vectors fields on K. Equivalently, the goal is to maximize the number of gradient pairs. For the complementary problem Min-Morse Matching(MinMM), the goal is to compute a gradient vector field that minimizes the number of unmatched (critical) simplices, over all possible gradient vector fields on K. While the problem of finding an exact optimum are equivalent for MinMM and MaxMM, as expected, the approximation variants of these problems have vastly different flavors.

One of the early motivations for our investigations was the experiments related to the performance of the coreduction algorithm in [86, 137]. It has been independently corroborated by various research groups that the reduction and coreduction based discrete gradients drastically reduce the size of simplicial complexes [86, 99, 137]. This has been observed to be the case for a wide range of deterministic and random families of complexes. And yet, for Max-Morse Matching, the approximability bounds provided by coreduction (and other methods) are dimension dependent, and fairly modest [137]. The vast chasm between the modest theoretical bounds and the outstanding practical performance – in short, the unreasonable effectiveness of the coreduction method – is one of the enduring mysteries in our investigations. One explanation for

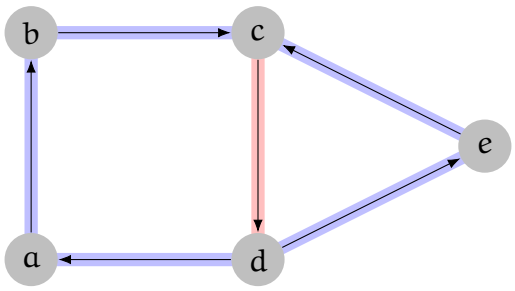


Figure 2: The figure shows a directed graph G. The single edge $\{c, d\}$ (shown in red) is a minimum feedback arc set for G, whereas, the edges complementary to the minimum feedback arc set induce a maximum acyclic subgraph, and are shown in blue.

this gap in understanding is that the analysis of the coreduction algorithm is not particularly tight. But it is far from clear how to improve upon the analysis in [137]. The coreduction algorithm is an iterative matching strategy, and deriving approximations bounds for these kind of iterative methods is extremely challenging. To unravel some of this mystery, in Papers I and IV of this thesis, we study the parameterized complexity and approximability of Morse matchings.

In order to prove an inapproximability bound for MaxMM, in Paper I, we carry out an L-reduction from a classical problem in graph theory that we describe next. Given directed graph $G = (V, E)$, the problem of finding a maximum subset $E_{max} \subseteq E$ for which the subgraph $G_{max} = (V, E_{max})$ has no directed cycles is called *maximum acyclic subgraph* (MAS). A *feedback arc set* is a set of edges whose removal leaves a directed acyclic graph. The problem of finding a feedback arc set of minimum cardinality is called *minimum feedback arc set* (minFAS). It is easy to check that minFAS is complementary to MAS. See Figure 2 for an example. A *directed degree-3 graph* is a directed graph with total degree (indegree plus outdegree) at most 3. The restriction of the problem MAS to directed degree-3 graphs is denoted by 3MAS.

The L-reduction from 3MAS to MaxMM, for establishing hardness of approximation for MaxMM uses a construction based on a modification of the dunce hat. While the dunce hat has no free faces, in contrast, the *modified dunce hat* is collapsible but only through a single free face. The specific triangulation of the modified dunce used by us in Paper I is as shown in Figure 3. Given a directed graph G, the reduction constructs a simplicial complex $K(G)$ built by systematically gluing together building blocks made of the modified dunce hat. Please refer to Section 4 of Paper I for further details on the construction. While we will not go into the details of the reduction, we will illustrate it here using the example in Figure 4. As a consequence of the L-reduction, we establish the following hardness result.

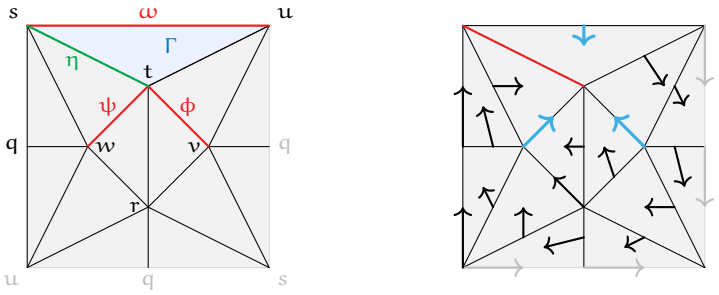


Figure 3: The modified dunce hat $\mathcal{D}$. Left: triangulation, with certain distinguished simplices highlighted. The complex is collapsible through the unique free face ω. Right: a discrete gradient $\mathcal{V}_{\mathcal{D}}$ on $\mathcal{D}$ that leaves only the vertices s, t and the edge η critical.

Theorem 1.1.1 (Theorem 4.1 in Paper I)**.** *It is* **NP***-hard to approximate* MaxMM *within a factor of* $\left(1 - \frac{1}{4914}\right) + \epsilon$ *and* **UGC***-hard to approximate it within a factor of* $\left(1 - \frac{1}{702}\right) + \epsilon$*, for any* $\epsilon > 0$*.*

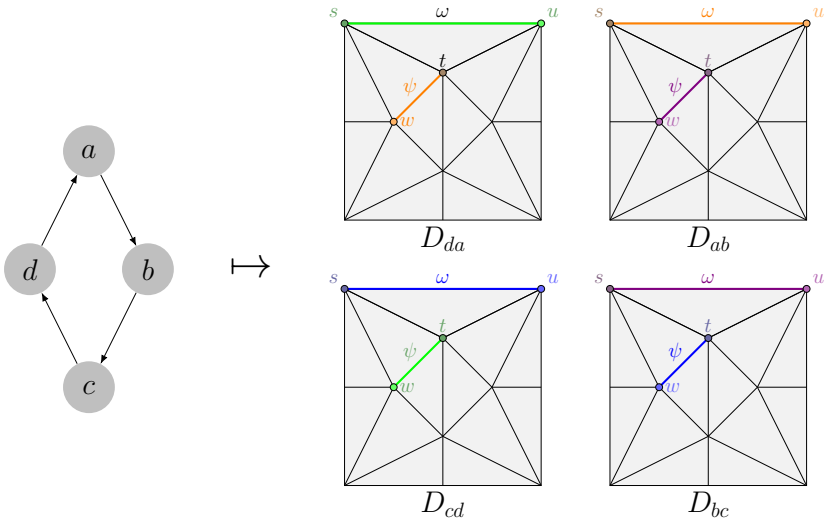
Remark 1. Here, we use the term **UGC**-hard to mean that assuming the unique games conjecture is true, the problem is **NP**-hard. Another way to state the second part of the theorem would be: If the Unique Games Conjecture is true, then MaxMM is **NP**-hard to approximate within a factor of $\left(1 - \frac{1}{702}\right) + \epsilon$, for any $\epsilon > 0$. The unique games conjecture is one of the most important open problems in theoretical computer science today. We refer the reader to the survey articles [107, 152] for further details on the conjecture.

Furthermore, building on Tancer's result [150] on **NP**-completeness of determining collapsibility of 3-dimensional simplicial complexes, in Paper I, we prove the following inapproximability result for MinMM for complexes of dimension $d \geqslant 3$.

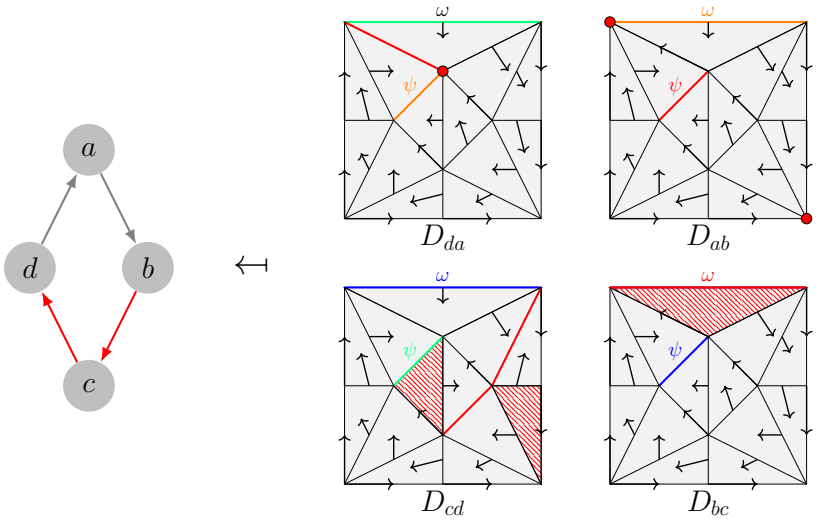
Theorem 1.1.2 (Corollary 3.1 in Paper I)**.** *For any* $\epsilon \in (0, 1]$*, there exists no* $O(n^{1-\epsilon})$*-factor approximation algorithm for* MinMM*, where* n *denotes the number of simplices of an input simplicial complex of dimension* $d \geqslant 3$*, unless* $\mathbf{P} = \mathbf{NP}$*.*

Prior to our work in Papers I and IV, Burton et al. [32] had shown that the related problem of Erasing Number is **W[P]**-hard with solution size as the parameter. The erasing number $\mathrm{er}(K)$ of a 2-dimensional simplicial complex K is the minimum number of 2-simplices that need to be removed so that the resulting complex collapses to a 1-complex. In other words, $\mathrm{er}(K)$ is the number of critical 2-simplices of an optimal gradient vector field on a 2-complex.

Note that one can not infer the **W[P]**-hardness of MinMM from Burton et al.'s reduction as the optimal number of critical simplices in their gadget can be of the order of the size of the complex because of the presence of 1-cycles (even when the number of critical 2-simplices is small). The questions of approximability of



(a) Graph G (b) Complex $\mathcal{K}(G)$



(a) Acyclic subgraph $\mathsf{A}(\mathcal{V})$ (b) Discrete gradient $\mathcal{V}_{\mathcal{K}(\mathsf{G})}$

Figure 4: Recall from Section 2.2 that an L-reduction from problem A to B constitutes of two maps. The first map that takes instances of A to instances of B, and the second map takes solutions of B to solutions of A. Accordingly, the above figure depicts an example of an L-reduction from 3MAS to MaxMM. The top figure shows a map that takes instances of 3MAS (directed graphs) to instances of MaxMM (simplicial complexes). The bottom figure shows a map that takes solutions of MaxMM (vector fields) to solutions of 3MAS (directed acyclic subgraphs). The red simplices in the complex on the bottom right are the critical simplices. The red edges in the graph on the bottom left are the feedback edges.

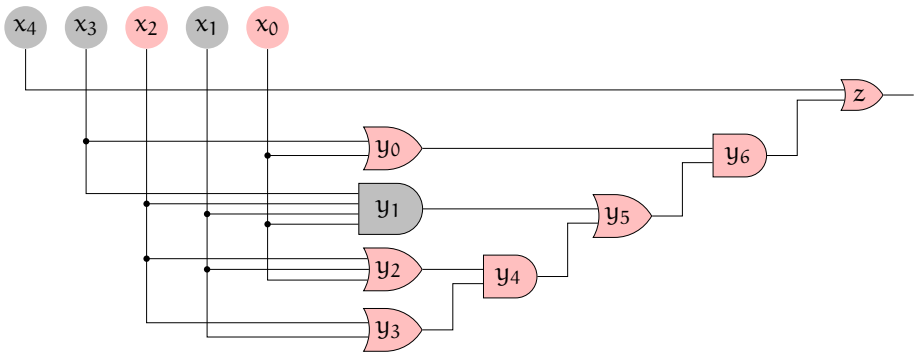


Figure 5: The figure shows a monotone circuit with a satisfying assignment. In this example the inputs x_0 and x_2 are set to 1. The gates that evaluate to 1 are called satisfied gates and those that evaluate to 0 are called unsatisfied gates. The satisfied gates are shown in red, and the unsatisfied gates are shown in gray. In this case, we say that the circuit is satisfied since the output gate is satisfied. Also, we say that $\{x_0, x_2\}$ is a satisfying assignment.

MinMM for 2-complexes, and the question of paramaterized approximability of MinMM for complexes of general dimension were also open at the time. This motivated our follow-up work on the complexity of computing optimal Morse matchings in Paper IV. In order to get a full picture of the complexity of Morse matchings, we needed to design the reduction gadget in a manner that allowed filling of 1-cycles that result from the gluing of building blocks of a reduction gadget. Also, we needed a hard problem to reduce from, a problem for which **W[P]**-hardness and (parameterized) inapproximability results were already known. We concluded that Min-Monotone Circuit Sat(MinMCS) was the right problem to reduce from. Informally, MinMCS can be described as follows: Given a monotone digital circuit (that is, a circuit consisting of and-gates and or-gates), how many inputs need to be activated in order to satisfy the circuit? For a formal definition, please refer to Section 3.3 of Paper IV. For an illustrative example, see Figure 5.

The reduction from MinMCS to MinMM in Paper IV is fairly complicated. So will refrain from discussing the reduction in detail. However, it would be remiss to not mention that the building block used for constructing the gadget is a subdivision of the modified dunce hat (Figure 3) from Paper I. Please refer to Figure 6 for an illustration of the subdivided modified dunce hat. The subdivision is necessary to ensure that the complex obtained after identifications is simplicial. Please refer to Section 4.2 of Paper IV for a full description of the construction of the gadget.

The same construction is used in establishing an L-reduction, a gap-preserving reduction and an **FPT**-reduction. As a consequence of this construction (followed by an additional technical step), we obtain the following results:

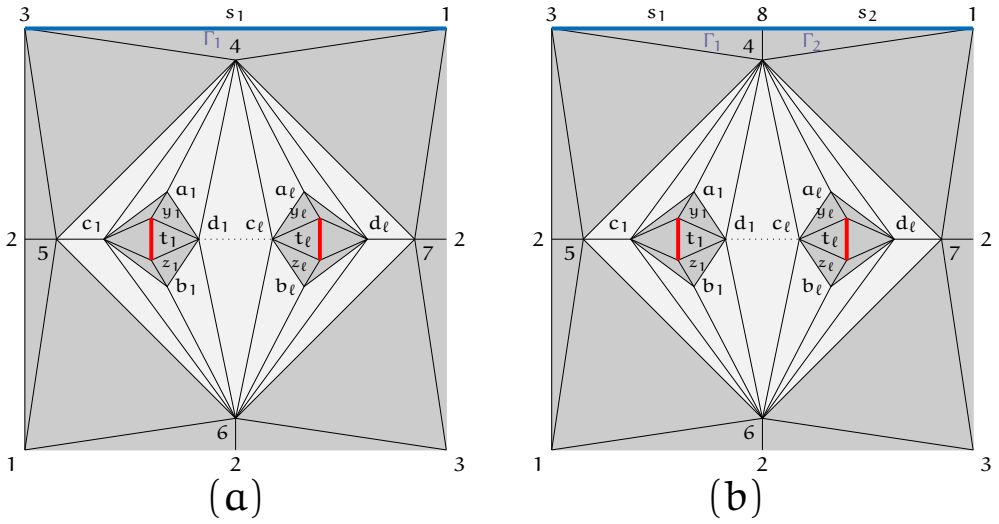


Figure 6: The figure (a) on the left depicts $\mathbf{D}_{1,\ell}$ that is collapsible through one free face, namely $s_1 = \{3, 1\}$. The figure (b) on the right depicts $\mathbf{D}_{2,\ell}$ that is collapsible through two free faces, namely $s_1 = \{3, 8\}$ and $s_2 = \{8, 1\}$. In both subfigures, the edges $\{1, 2\}$ and $\{2, 3\}$ on the right and at the bottom are shown in light grey to indicate that they are identified to $\{1, 2\}$ and $\{2, 3\}$ on the left. The bold edges highlighted in red and blue are the distinguished edges that are glued to distinguished edges from other building blocks to form a simplicial complex $K(C)$ associated to a circuit C.

- MinMM has no approximation within a factor of $2^{\log^{(1-\epsilon)} n}$, for any $\epsilon > 0$, unless **NP** $\subseteq$ **QP**,
- the standard parameterization of MinMM is **W[P]**-hard, and
- Assuming **FPT** $\neq$ **W**[**P**], MinMM with standard parameterization has no FPT approximation algorithm for *any* approximation ratio function ρ.

For further details about the reduction refer to Sections 5 and 6 in Paper IV. To complement the inapproximability results for MinMM on 2-complexes, in Section 7 of Paper IV, we provide a simple $\frac{n}{\log n}$-factor approximation algorithm for MinMM on 2-complexes. In Tables 1 to 3, the reader can find a full summary of parameterized complexity and approximability results concerning the computation of optimal discrete gradients on simplicial complexes.

Speaking of algorithms with good behavior, there is yet another discrete gradient algorithm of interest, namely, the *apparent pairs gradient*, which was first introduced in the context of stochastic topology, but has now been found to be very advantageous in speeding up persistence computations [11, 90]. The apparent pairs gradient can be stated as follows:

	Burton et al. [32]	Paper IV
W[P]-hardness of ERASING NUMBER (SP)	✓	✓
W[P]-hardness of MIN-MORSE MATCHING (SP)	✗	✓
FPT-inapproximability of MIN-MORSE MATCHING (SP)	✗	✓
FPT-algorithm for MIN-MORSE MATCHING (TW)	✓	✗

Table 1: Comparison of results: Parameterized complexity of Morse matching (dim. $\geqslant 2$). In the table above, SP denotes standard parameterization, whereas TW denotes treewidth parameterization.

	Paper I	Paper IV
IA for MIN-MORSE MATCHING (dim. $\geqslant 3$)	$O(n^{1-\epsilon})$	$2^{\log^{(1-\epsilon)} n}$ (weaker)
IA for MIN-MORSE MATCHING (dim. $\geqslant 2$)	✗	$2^{\log^{(1-\epsilon)} n}$
AA for MIN-MORSE MATCHING (dim. $= 2$)	✗	$O(\frac{n}{\log n})$

Table 2: Comparison of results: Approximability of Min-Morse matching. Here, IA denotes inapproximability, and AA denotes approximation algorithm.

	Paper I	Results from [137]
IA for MAX-MORSE MATCHING (dim. $\geqslant 2$)	$(1 - \frac{1}{4914}) + \epsilon$	✗
AA for MAX-MORSE MATCHING (dim. $\geqslant 2$)	✗	$\frac{d+1}{d^2+d+1}$

Table 3: Comparison of results: Approximability of Max-Morse matching, IA denotes inapproximability, AA denotes approximation algorithm, and d denotes the dimension of the complex.

Definition 1.1.1 (Apparent pairs gradient [11])**.** A pair of simplices (σ, τ) of K an *apparent pair* of K if both

- σ is the lexicographically highest facet of τ, and
- τ is the lexicographically lowest cofacet of σ.

The collection of all the apparent pairs in a complex K forms a discrete gradient that is referred to as the *apparent pairs gradient*.

In Section 8 of Paper IV, we extend Kahle's techniques for designing discrete gradients on random clique complexes to Costa–Farber complexes. While we will not define the various families of random complexes here (See Section 8 of Paper IV for definitions), we would like to mention that Costa–Farber complexes constitute a family of complexes that generalize the Linial–Meshulam complexes as well as the random clique complexes. In Paper IV, we show that for a wide range of parameter values for Costa–Farber complexes, there exist discrete gradients for which the ratio of expected number of critical r-simplices to the expected number of r-simplices (for any fixed dimension r) tends to zero.

Specializing the analysis to Linial–Meshulam complexes, followed by a few additional steps, we obtain a strong result about the "average case approxima-

bility" of Morse matchings (for Linial–Meshulam complexes), which we will state below.

Theorem 1.1.3. *For the regimes of Linial-Meshulam complexes* $Y_d(n,p)$ *that satisfy*

$$np \to \infty \quad \textit{or} \quad np \to 0$$

there exists a discrete gradient $\mathcal{V}$ *that satisfies* $\frac{\mathbb{E}(|\mathcal{V}|)}{\mathbb{E}(|\mathsf{opt}|)} \to 1$ *as* $n \to \infty$, *where* opt *denotes the optimal gradient on* $Y_d(n,p)$.

The discrete gradient $\mathcal{V}$, for the most part, is composed of apparent pairs. Here, we will refrain from providing further details about this result. Instead, we would like to mention that a statement analogous to Theorem 1.1.3 also holds for random clique complexes (but has not been explained in that form in Paper IV).

This concludes our discussion about the discrete Morse theory related results in this thesis.

1.1.1 *Outlook & Future directions*

STRENGTHENING OF ALGORITHMIC RESULTS In Section 8 of Paper IV, we study the behavior of the apparent pairs gradient (a deterministic algorithm) on random complexes. Complementing this idea, it might be interesting to study the behavior of random variants of the apparent pairs gradient on deterministic families of dense (or sparse) complexes. In this setting, there might even be some hope for obtaining worst-case approximability bounds.

STRENGTHENING OF HARDNESS RESULTS For the most part, we believe that the results in Papers I and IV, along with some of the earlier results that appeared in literature [32, 100, 137] provide a near-complete picture of the complexity of computing optimal Morse matchings. A minor strengthening is still possible. We conjecture that for simplicial complexes of dimension $d > 2$, Min-Morse Matching does not admit an $f(n)$-approximation algorithm for any $f = o(n)$. In particular, a result of this nature would show that while the problem is hard for complexes of all dimensions, it is, in fact, slightly harder for higher dimensional complexes when compared to 2-dimensional complexes, from an inapproximability standpoint. This would, for instance rule out an $O(\frac{n}{\log n})$-factor approximation algorithm for Min-Morse Matching for complexes of dimension $d > 2$, providing a contrast with the existence of such an algorithm for 2-complexes in Section 7 of Paper IV. That said, such a strengthening would have no bearing in practice, and is only a matter of complexity theory curiosity.

WIDER APPLICABILITY OF MODIFIED DUNCE HATS In Papers I and IV, we used modified dunce hats to obtain hardness results for Morse matching

problems, and in Paper II, we use them to obtain parameterized complexity results for problems in simple homotopy theory. Independent of our work, modified dunce hats were used by Santamaría-Galvis and Woodroofe to provide a simpler proof of NP-completeness of the shellability decision problem [147]. This raises the possibility of a wider applicability of these complexes for complexity questions in combinatorial topology.

GENERATING BENCHMARK DATASETS FOR MORSE MATCHINGS In the experiments from [137] it became clear that there are no known instance families of simplicial complexes that behave poorly with respect to the coreduction and reduction methods. To fix this anomaly, a natural question to ask is the following one: Is it possible to define a simple random model built out of gluing dunce hats geared specifically towards generating hard instances for MIN-MORSE MATCHING for a wide range of parameter values?

1.2 SIMPLE HOMOTOPY THEORY

The motivating problem and the starting point of discussion for Paper II was the topological Andrews-Curtis conjecture, which we state below.

TOPOLOGICAL ANDREWS-CURTIS CONJECTURE. If K is a finite contractible CW complex of dimension 2, then K can be 3-deformed to a point.

For a more detailed discussion on the importance of the Andrews-Curtis (and related) conjectures, we refer the reader to the discussion in Section 3.1. Most questions in (simple) homotopy theory revolve around computability. For instance, the algorithmic question related to the Andrews-Curtis conjecture, namely, the existence on an algorithm for determining triviality of balanced presentations is still open (See Section 3.1.3 for further details concerning this question). We wanted to bring algorithmic questions in simple homotopy theory from the computability realm to the complexity realm, if only, by studying the problem in restricted settings. In particular, we wanted to explore the possibility of parameterized algorithms for questions in simple homotopy theory.

In Paper II, motivated by the aforementioned problems, we investigate the parametrized complexity of a few of variants of the problem of deciding contractibility. More precisely, we focus on the problem of deciding whether a given 2-complex admits a simple homotopy to a 1-complex using at most p expansions, called ERASABILITY EXPANSION HEIGHT. In addition, we consider a variant, called ORDERED ERASABILITY EXPANSION HEIGHT, which requires that all expansions come at the very beginning of the sequence. It is worth noting that ERASABILITY EXPANSION HEIGHT and ORDERED ERASABILITY EXPANSION HEIGHT are equivalent for CW complexes for which one can readily swap the order of expansions and collapses [120, p. 34]. However, for simplicial complexes, the ordered and unordered expansion heights may differ.

No.	Assumptions ($\mathcal{A}$)		Conclusion (c)
Impl. 1.	$\{s_1\}$	$\Longrightarrow$	s_3
Impl. 2.	$\{s_2\}$	$\Longrightarrow$	s_3
Impl. 3.	$\{s_4\}$	$\Longrightarrow$	s_3
Impl. 4.	$\{s_3, s_1\}$	$\Longrightarrow$	s_4
Impl. 5.	$\{s_3, s_4\}$	$\Longrightarrow$	s_2
Impl. 6.	$\{s_1\}$	$\Longrightarrow$	s_2
Impl. 7.	$\{s_2, s_3, s_4\}$	$\Longrightarrow$	s_1

Table 4: Implication relation R

We begin the discussion about Paper II with an admission. Our efforts towards designing an FPT algorithm for Ordered Erasability Expansion Height and Erasability Expansion Height with respect to one of the most pliable parameters, that is, the treewidth of the Hasse graph of the complex, failed. This is because when one allows expansions, the treewidth of the complex increases in a manner that is difficult to control.

On the other hand, in Section 5 of Paper II, we were able to show that Erasability Expansion Height and Ordered Erasability Expansion Height with solution size as the parameter are both **W[P]**-complete. Moreover, in Section 6 of Paper II, we show that the problem of deciding whether a 2-complex is simple-homotopy equivalent to a 1-complex using only 3-dimensional expansions is **NP**-complete.

The membership in **W[P]** and **NP** of the respective problems crucially exploits the fact that erasability of 2-complexes can be deterministically checked in polynomial time, as shown by Tancer [150, Proposition 5].

To establish hardness results, we use an FPT reduction from Axiom Set, which is defined as follows: Given a finite set S of *sentences*, an *implication relation* R consisting of pairs $(\mathcal{A}, c)$, where $\mathcal{A} \subseteq S$ and $c \in S$, and a positive integer $p \leqslant |S|$, is there a set $S_0 \subseteq S$, called an *axiom set*, with $|S_0| \leqslant p$ and a positive integer n such that if we recursively define

$$S_i := S_{i-1} \cup \{s \in S \mid \exists \mathcal{A} \subseteq S_{i-1} : (\mathcal{A}, s) \in R\}$$

for $1 \leqslant i \leqslant n$, then $S_n = S$?

For example, an Axiom Set instance, with $S = \{s_1, s_2, s_3, s_4\}$ as the set of sentences, and with the implication relation as described in Table 4, has an axiom set of size 1. Specifically, $\{s_1\}$ is an axiom set.

Given an Axiom Set instance Π, the reduction constructs a complex $K(\Pi)$ whose building blocks consist of modified dunce hats encountered before in this introduction (in the context of Paper I and Paper IV). Interestingly, we were able to use the same construction to obtain **W[P]**-hardness of the unordered and the ordered variants of expansion height owing to the following equivalence established in Theorem 8 of Paper II:

(a) there exists a simple-homotopy equivalence from $K(\Pi)$ to a 1-dimensional complex using first at most p expansions, followed by a sequence of only collapses,

(b) there exists a simple-homotopy equivalence $K(\Pi) \nearrow\searrow L$ using at most p expansions, where L is a 1-dimensional complex, and

(c) there exists an axiom set $S_0 \subset S$ for $\Pi = (S, R, p)$ using at most p elements.

Note that we only considered Erasability Expansion Height and Ordered Erasability Expansion Height in paper II. The more natural variant (Expansion Height) would be to study the complexity of determining sequences of expansions and collapses (with at most p expansions) that take K to *a point*. At that time we did not consider the more natural variant because the gluings of modified dunce hats in $K(\Pi)$ give rise to 1-cycles. The technique for filling 1-cycles was introduced in Paper IV, and chronologically, Paper IV was written after Paper II. We expect the cycle filling techniques in Paper II to generalize in a straightforward way to expansion height problems. In particular, we expect Expansion Height to be **W[P]**-complete.

1.2.1 *Outlook and future directions.*

Loosely related to the theme in Paper II is the work by Fernandez [78], in which she uses discrete Morse theory based algorithms to find 3-deformations of 2-complexes to points in order to refute some potential counterexamples for the Andrews-Curtis conjecture.

Developing algorithmic tools in the context of simple homotopy theory is still a largely unexplored research area. The main open question, of course, is determining whether the triviality of balanced presentations is decidable.

Devising algorithms for problems in simple homotopy theory in limited settings, however, could still be fruitful. For instance, it is worth noting that even the word problem becomes polynomial time solvable in restricted settings [126]. So, are there interesting restricted settings for which the triviality of balanced presentations question becomes decidable? In a similar vein, one may ask if there exists an efficient algorithm to determine whether two lens spaces are simple homotopy equivalent.

1.3 MATROIDS, CYCLE BASES AND HOMOLOGY BASES

Notation 1. Throughout, In this section, we use n to denote the number of vertices of a graph (or a complex), m to denote the number of edges of a graph (or a complex), N to denote the number of simplices in a complex. and g to denote the rank of the first homology group of a complex.

MINIMUM CYCLE BASES FOR GRAPHS. Let graph $G = (V, E)$ be a connected graph, and let $w : E \to \mathbb{R}^+$ be a positive weight function on E. A subgraph of G which has even degree at every vertex is called a *cycle* of G. A cycle C is *elementary* if C is a connected subgraph in which every vertex has degree 2. To every cycle C, we associate an incidence vector $\mathbf{c}$ indexed by the edge set E as follows:

$$\mathbf{c}_e = \begin{cases} 1 & \text{for } e \in E \\ 0 & \text{for } e \notin E \end{cases}.$$

It is easy to check that the set of incidence vectors of cycles forms a vector space over $\mathbb{Z}_2$ called the *cycle space* of G of dimension ν, where $\nu = |E| - |V| + 1$. A cycle basis is a vector space basis of the cycle space.

The weight of a cycle is the sum of the weights of its edges, and the weight of a cycle basis is the sum of the weights of the basis elements. The problem of computing a cycle basis of minimum weight is called the minimum cycle basis problem. Since all weights are assumed to be positive, there always exists a minimum cycle basis of elementary cycles.

MINIMUM HOMOLOGY BASES FOR COMPLEXES. Given a simplicial complex K, a set of 1-cycles $\{\zeta_1, \ldots, \zeta_g\}$ is called a *homology cycle basis* if the set of classes given by $\{[\zeta_1], \ldots, [\zeta_g]\}$ forms a homology basis. For brevity, we abuse notation by using the term "homology basis" for $\{\zeta_1, \ldots, \zeta_g\}$. Assigning non-negative weights to the edges of K, the weight of a cycle is the sum of the weights of its edges, and the weight of a homology basis is the sum of the weights of the basis elements.

MATROIDS. A matroid $\mathscr{M}$ consists of a pair $(S, \mathscr{I})$, where S is a finite ground set and $\mathscr{I}$ is a family of subsets of S satisfying the following axioms:

1. $\emptyset \in \mathscr{I}$;
2. if $I \in \mathscr{I}$ and $J \subseteq I$, then $J \in \mathscr{I}$; and
3. if $I, K \in \mathscr{I}$ and $|I| < |K|$, then there is an element $e \in K \setminus I$ such that $I \bigcup \{e\} \in \mathscr{I}$.

If a set $I \subset S$ belongs to $\mathscr{I}$, then it is called an *independent set*, else it is called a *dependent set*. A *circuit* in a matroid $\mathscr{M}$ is a minimal dependent subset of S. All proper subsets of a circuits are independent sets. A maximal independent set is called a *basis* of the matroid.

Matroids admit a very useful property that goes by the name of *basis exchange property*: If A and B are distinct bases of a matroid and $a \in A \setminus B$, then there exists an element $b \in B \setminus A$ such that $A(\setminus \{a\}) \bigcup \{b\}$ is again a basis. A *weighted matroid* is a matroid $\mathscr{M}$ equipped with a weight function $w : S \to \mathbb{R}^+$ that additively extends to all subsets of S. From an algorithmic standpoint, the most

important property of weighted matroids is that there is a greedy algorithm for these matroids that computes the maximum (minimum) weight basis. It can be easily checked that B is a minimum weight basis of a matroid if and only if none of the cycles of B can be exchanged for a shorter cycle while still preserving linear independence. As an immediate consequence, if the cycles in two distinct minimum cycle (homology) bases are sorted by weight, then the ordered sets of sorted weights coincide.

Let the cycle space of G be the ground set S, and let $\mathscr{I}$ be defined as follows.

$$\mathscr{I} = \{I \mid I \text{ is a linearly independent set of cycles of } G\}$$

Then, the pair $\mathscr{M} = (S, \mathscr{I})$ forms a matroid. When combined with a weight function on edges, it becomes a weighted matroid. Cycle bases of G correspond to the bases of $\mathscr{M}$.

Analogously, let the nontrivial 1-cycles of a 2-complex K be the ground set S', and let $\mathscr{I}'$ be defined as follows.

$$\mathscr{I}' = \{I \mid I \text{ is a subset of some homology basis of } K\}$$

Then, the pair $\mathscr{M}' = (S', \mathscr{I}')$ together with a weight function on the edges forms a weighted matroid. The homology bases of K are the bases of matroid $\mathscr{M}'$.

The cycles in a cycle (homology) basis can be exchanged with other cycles while preserving independence. This property is key to Lemma 8 and Theorem 9 of Paper III.

1.3.1 *Background work on cycle bases*

The key structural results for minimum cycle basis can be found in the work of Horton [94] and de Pina [51]. We briefly recall them here. First we start with a definition, and then we recall the key theorems.

Definition 1.3.1 (Tight cycles)**.** A simple cycle C in a graph G is *tight* if it contains a shortest path between every pair of points in C. The set of all tight cycles in a graph is denoted by $\mathcal{T}(G)$.

Theorem 1.3.1 (Horton [94])**.** *A minimum cycle basis $\mathcal{M}(G)$ of a graph G consists only of tight cycles.*

Theorem 1.3.2 (de Pina [51])**.** *Cycles $C_1 \ldots, C_\nu$ form a minimum cycle basis if there are vectors $S_1, \ldots, S_\nu$ in the cycle space such that for all i, $1 \leqslant i \leqslant \nu$, the following conditions hold:*

PREFIX ORTHOGONALITY: $\langle C_j, S_i \rangle = 0$ *for all* $1 \leqslant j \leqslant i$.

NON-ORTHOGONALITY: $\langle C_i, S_i \rangle = 1$.

SHORTNESS: C_i *is a minimum weight cycle in* $\mathcal{T}$ *with* $\langle C_i, S_i \rangle = 1$.

The vectors $S_1, \ldots, S_\nu$ in Theorem 1.3.2 are called *support vectors*. The recent line of algorithmic work [5,51,105,106,121] on the minimum cycle basis problem strongly rely on Theorem 1.3.2. Building on this line of work, more recently, Amaldi et al. [5] proved another interesting structural result.

Theorem 1.3.3 (Amaldi et al. [5])**.** *The total length of the tight cycles in an undirected graph* G *is at most* $n\nu$*, and can be computed in* $O(nm)$ *time.*

For further details on background and history of this problem, we refer the reader to some survey articles [18, 104].

1.3.2 *Background work on homology bases*

For the special case when the input complex is a surface, Erickson and Whittlesey [73] gave a $O(N^2 \log N + gN^2 + g^3N)$-time algorithm. Recently, Borradaile et al. [23] gave an improved deterministic algorithm that runs in $O((h+c)^3 n \log n + m)$ where c denotes the number of boundary components, and h denotes the genus of the surface. For small values of c and h, the algorithm runs in nearly linear time.

For general complexes, Dey et al. [54] and Chen and Freedman [38] generalized the results by Erickson and Whittlesey [73] to arbitrary complexes. Subsequently, introducing the technique of annotations, Busaryev et al. [33] improved the complexity to $O(N^\omega + N^2 g^{\omega-1})$, where ω is the exponent of matrix multiplication. More recently, Dey et al. [53] designed an $O(N^\omega + N^2 g)$ time algorithm by adapting the divide and conquer algorithm for computing a minimum cycle basis of Kavitha et al. [106] for the purpose of computing a minimum homology basis.

Note that in Paper III, we focus solely on the bases of the first homology group since the problem of computing a shortest basis for higher homology groups with $\mathbb{Z}_2$ coefficients was shown to be **NP**-hard by Chen and Freedman [39].

For further details on background and history of minimum homology basis and related problems, we refer the reader to the survey article by Erickson [72].

1.3.3 *Background work on matrix algorithms*

The *column rank profile* (respectively *row rank profile*) of an $m \times n$ matrix A with rank r, is the lexicographically smallest sequence of r indices $[i_1, i_2, \ldots, i_r]$ (respectively $[j_1, j_2, \ldots, j_r]$) of linearly independent columns (respectively rows) of A.

It is well-known that classical Gaussian elimination can be used to compute the rank profile of an $m \times n$ matrix in $O(nmr)$ time. One of the keys to our design of fast algorithms for minimum cycle and homology basis is the use of state-of-the-art matrix rank profile algorithms. We recall the main results here.

Theorem 1.3.4 ([65, 98]). *There is a deterministic* $O(mnr^{\omega-2})$ *time algorithm to compute the column rank profile of an* $m \times n$ *matrix* A *with rank* r.

Theorem 1.3.5 (Storjohann and Yang [148, 149, 160]). *There exists a Monte Carlo algorithm for computing row (resp. column) rank profile of a (sparse) matrix* A *with* $\mathsf{nnz}(A)$ *non-zero entries and rank* r *that runs in* $\tilde{O}(\mathsf{nnz}(A) + r^{\omega})$ *time. The failure probability of this algorithm is* $1/2$.

1.3.4 *A structural result*

Previously, Erickson and Whittlesey [73] showed that the minimum homology basis of a complex K is contained is contained in the tight cycles of its 1-skeleton, considered as a graph.

Theorem 1.3.6 (Erickson and Whittlesey [73]). *With non-negative weights, every cycle in a shortest basis of* $H_1(K)$ *is tight. That is, if* $\mathcal{H}(K)$ *is any minimum homology basis of* K, *then* $\mathcal{H}(K) \subset \mathcal{T}(K_1)$.

In Paper III, we refine Theorems 1.3.1 and 1.3.6 to prove the following result.

Theorem 1.3.7 (Corollary 11 in Paper III). *Given a complex* K, *let* $\mathcal{T}(K_1)$ *denote the set of tight cycles of* K_1, *and let* $\mathcal{M}(K_1)$ *be a minimum cycle basis of* K_1. *Then, there exists a minimum homology basis* $\mathcal{H}(K)$ *of* K *such that*

$$\mathcal{H}(K) \subset \mathcal{M}(K_1) \subset \mathcal{T}(K_1).$$

1.3.5 *Algorithmic results*

The state-of-the-art algorithm for the minimum cycle basis computation is by Amaldi et al. [5], which runs in $O(m^{\omega})$ time. In Paper III, we devise a Monte Carlo algorithm for minimum cycle basis computation that computes the minimum cycle basis in $\tilde{O}(m^{\omega})$ time[1]. Please refer to Section 3 in Paper III for further details. While our algorithm for minimum cycle basis is slightly slower than the algorithm by Amaldi et al. [5], it enjoys a simpler high-level description, and is unique among the recent algorithms for minimum cycle basis, in the fact that it does not use support vectors (described in Section 1.3.1).

Further, in Paper III, we devise two Monte Carlo algorithms for minimum homology basis that run in $\tilde{O}(m^{\omega})$ and $O(m^{\omega} + Nm^{\omega-1})$ time, respectively, where as before, m denotes the number of edges in the complex, and N denotes the total number of simplices in the complex. These algorithms improve upon the state-of-the-art algorithm by Dey et al. [53], which runs in $O(N^{\omega} + N^2 g)$ time. Please refer to Section 5 in Paper III for further details.

Our algorithms for minimum cycle basis and minimum homology basis are fast and have a simple high-level description. These algorithms harness the following facts, observations and results.

1 Here, the $\tilde{O}()$ notation hides multiplicative polylogarithmic factors.

1. For minimum cycle basis and minimum homology basis, if one sorts the candidate cycles by weight and assembles them in a matrix, then the column rank profile algorithm is a greedy algorithm for minimum basis computation.

2. The minimum cycle bases and minimum homology bases have a weighted matroid structure, and hence column rank profile computation returns an optimal basis for both problems. The matroid structure of these problems is somewhat underemphasized in literature. The survey paper by Burger et al. [18] is a notable exception.

3. The matrix of tight cycles is a sparse matrix as proved by Amaldi et al. [5]. See Theorem 1.3.3.

4. The structural result proved in Theorem 1.3.7 was, in fact, the the main impetus behind Paper III.

5. Some of the state-of-the-art algorithms for column rank profile computation can be used as blackbox subroutines for computing minimum cycle basis and minimum homology basis.

1.3.6 *Outlook & Future directions*

In Paper III, we devised a randomized algorithm for minimum homology basis with complexity $\tilde{O}(m^\omega)$. This naturally leads to the question of how hard it would be to make significant improvements upon this result. The key to answering questions of this nature lies in the following result by Edelsbrunner and Parsa.

Theorem 1.3.8 (Edelsbrunner and Parsa [66])**.** *Let* $\mathbf{M}$ *be an* $n \times n$ *0-1 matrix with* p *non-zero entries. In time* $O(p)$*, we can build a 2-dimensional simplicial complex* K *with* $O(p)$ *simplices and a piecewise linear function* $g : |K| \to \mathbb{R}$ *such that the horizontal homology class* $H_1^{hor}(K)$ *is isomorphic to the null-space of* $\mathbf{M}$ *and* $H_2(K)$ *is isomorphic to the null-space of* $\mathbf{M}^T$*.*

The so-called horizontal homology classes $H_d^{hor}(K)$ and vertical homology classes $H_d^{ver}(K)$ were introduced in [45], and treated at length in [55]. In particular, we have the following relation: $\dim H_1^{hor}(K) + \dim H_1^{ver}(K) = \beta_1$. Also, as a consequence of [55, Claim 3.2], $H_1^{ver}(K)$ can be computed deterministically in time $O(p \log p)$ using the Reeb graph computation algorithm by Parsa [131]. Any algorithm for computing a minimum homology basis reveals β_1.

The algorithm for computing minimum homology basis in Paper III runs in $\tilde{O}(m^\omega)$ time, while the lower bound implied by Edelsbrunner and Parsa [66] is no stronger than $\Omega(m^2)$ (in the light of Wiedemann's algorithm). One reason for this discrepancy is that finding the column rank profile of a matrix is at least as hard as finding the rank. A second reason is that in order to compute a

minimum homology basis one needs to first identify a small (and sparse) list of candidate cycles, which incurs an additional computational overhead. This leaves open the question of whether a substantial improvement in asymptotic complexity bounds for computing the minimum homology basis (over the results obtained in Paper III) is possible.

On a slightly different note, while the complexity of computing *any* homology basis, and a minimum homology basis are both closely tied to rank computation, the situation for cycle bases is quite different. In particular, from any spanning tree of a graph, one can extract a (fundamental) cycles basis in linear time. Hence, computing a cycle basis of a graph is achievable in time linear in the number of edges. In comparison to finding any cycle basis of a graph, the problem of computing a minimum cycle basis of a graph seems to be a substantially harder problem, leading naturally to the interesting open question of establishing fine-grained complexity bounds for the problem of computing a minimum cycle basis of an undirected graph.

1.4 HIGH-DIMENSIONAL CUTS

1.4.1 *Motivation*

A *graph cut* is a partition of the vertices V of a graph $G = (V, E)$ into two disjoint subsets $(S, \bar{S})$. The set of edges $C \subset E$ that have one vertex lying in S and another one lying in $\bar{S}$ determines a *cut-set*. See Figure 7 for an illustration. Typically, the objective function to optimize involves the (weighted) size of the cut-set. Graph cuts have a ubiquitous presence in theoretical computer science. Additionally, cuts also have a rich mathematical aspect. For instance, cuts are also related to the spectra of the adjacency matrix of the graph leading to a beautiful mathematical theory [42]. Then, there is also a geometric aspect to cuts that comes from metric embeddings [57]. Cuts have found many real-world applications in clustering, shape matching,VLSI design, image segmentation and smoothing, and energy minimization problems in computer vision.

Cut problems are related to flow problems in graphs owing to the duality between cuts and flows. In fact, the max-flow min-cut theorem says that the maximum value of flow between a source vertex s and a target vertex t equals the value of the minimum cut that separates s and t. Figure 8 shows an example of an s-t cut on an undirected graph.

We begin with the observation that graphs are 1-dimensional simplicial complexes. As we will argue in the ensuing discussion, cuts have a natural homological interpretation. Then, it is natural to muse how would the notion of cuts have been defined had it first emerged in the context of simplicial complexes. Guided by a higher intuition that comes from topology, cuts as constructs need not be limited merely to achieve *separation*. But more generally,

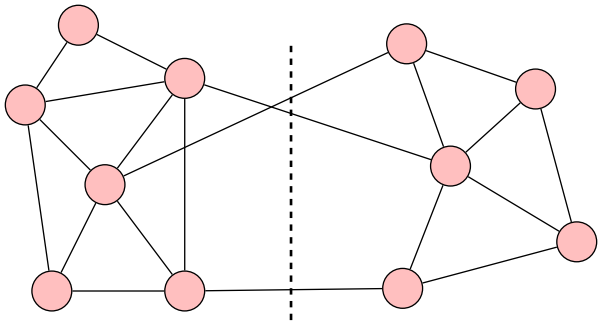


Figure 7: Partition $(S, \overline{S})$ of vertex set V.

one may view cuts as a mechanism to achieve *homology modification* via removal of simplices.

For instance, consider the following problem: What is the minimum number of edges you need to remove from a graph so that the vertices $\{s, t\}$ do not form a bounding 0-cycle of a 1-chain over $\mathbb{Z}_2$ in the resulting graph? Since we have an s-t cut if and only if there are no paths connecting s and t, it is easy to check that this problem is equivalent to finding the minimum s-t cut on graphs! Now, we ask the analogous question for complexes of higher dimension. In particular, the question we ask, namely BOUNDARY NONTRIVIALIZATION, is the following one: Given a bounding $\mathbb{Z}_2$ r-cycle ζ in a simplicial complex K, find a set S of $(r+1)$-dimensional simplices of minimum cardinality so that the removal of S from K makes ζ nontrivial.

For instance, consider the two complexes L_1 and L_2 shown in Figures 9 and 10, respectively. For complex L_1 shown in Figure 9, let the equator e of the sphere on the right be the bounding 1-cycle that we want to make nontrivial. Both hemispheres are bounded by the equator. So, the two highlighted triangles from the right sphere of the complex L_1 constitute the optimal solution for BOUNDARY NONTRIVIALIZATION. That is, removing these two triangles makes e a nontrivial 1-cycle. For complex L_2 shown in Figure 10, the circle of intersection of the two spheres, denoted by b, is the bounding 1-cycle of interest. Removing all the four highlighted triangles from complex L_2 makes b a nontrivial 1-cycle. This also happens to be the optimal solution for making b nontrivial. The global variant of BOUNDARY NONTRIVIALIZATION, that we refer to as GLOBAL BOUNDARY NONTRIVIALIZATION asks for the smallest number of $(r+1)$-simplices that need to be removed so that the dimension of the $(r+1)$-th boundary group reduces by one. The (weighted version of) GLOBAL BOUNDARY NONTRIVIALIZATION generalizes the classical minimum cut problem on undirected graphs.

Complementary to the question of removing the minimal number of $(r+1)$-simplices in order to make a bounding cycle nontrivial, is the problem of

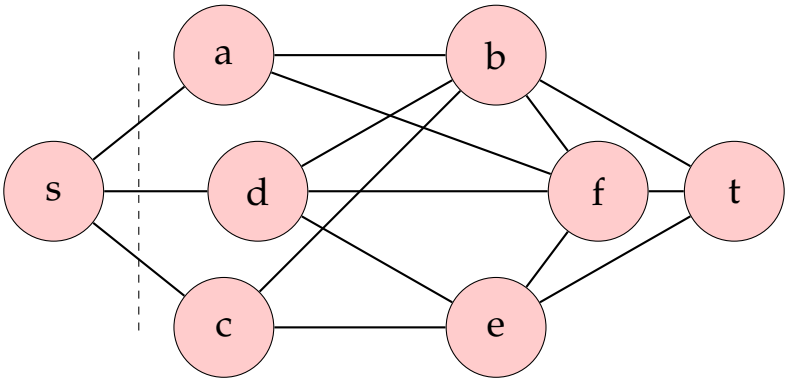


Figure 8: The dashed vertical line shows a minimum s-t cut in the graph.

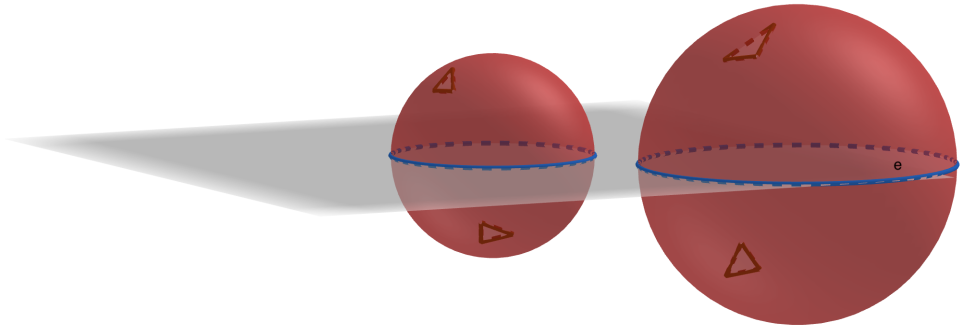


Figure 9: The complex L_1 consists of two disjoint triangulated spheres. We do not show the entire triangulation, only the four triangles of interest. The boundary of interest is the equator of the larger sphere on the right.

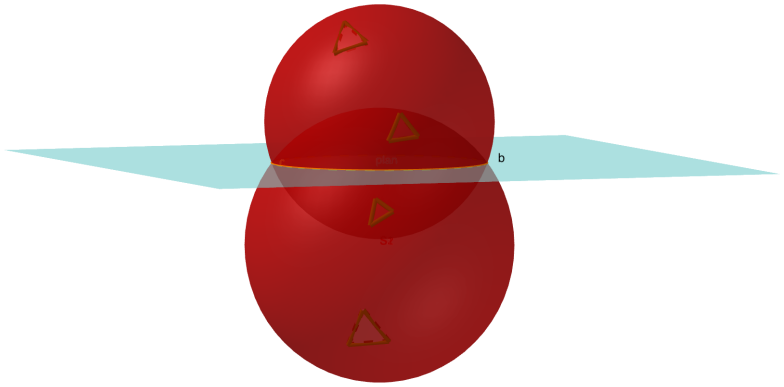


Figure 10: The complex L_2 consists of a triangulation of union of two spheres that intersect in a circle. As before, we do not show the entire triangulation, only the four triangles of interest. The boundary of interest is the circle of intersection of the two spheres.

removing the minimum number of r-simplices from a complex so that an entire homology class is destroyed. More formally, the problem Topological Hitting Set can be described as follows: given a nontrivial $\mathbb{Z}_2$ r-cycle ζ in a simplicial complex K, find a set S of r-dimensional simplices of minimum cardinality so that S meets every cycle homologous to ζ.

Topological Hitting Set on graphs can be described as follows: Suppose we are given a graph G with k components. Let C be one of the components of G. Then, $\beta_0(G) = |k|$, and each component determines a 0-cycle. So the question of Topological Hitting Set is to determine the minimum number of vertices you need to remove so that C is not a component anymore. The answer is trivial! One needs to remove all the vertices in C. For example in Figure 11, C_2 ceases to be a component if and only if all four vertices in C_2 are removed. It is worth noting that it is the unidimensionality of graphs that makes the problem trivial. What is more, the "cut aspect" of the problem is not immediately visible for graphs.

In contrast, for higher-dimensional complexes, the problem has a distinct cut flavor. For instance, consider the planar complex shown in Figure 12. The minimum number of edges that need to be removed so that every cycle homologous to ζ is destroyed is three. In Figure 12, an optimal set of edges is shown in red. Note that the edges happen to be in a "thin" portion of the complex, justifying our higher-dimensional homological viewpoint that (along with Boundary nontrivialization) this problem can also be seen as a high dimensional cut problem. The global variant of Topological Hitting Set that we call Global Topological Hitting Set asks for the minimum number of r-simplices that need to be removed so that *some* r-cycle is destroyed. For instance, for the complex shown in Figure 12, a set of two edges needs to be removed in order to destroy the class $[\xi]$, which corresponds to a solution to Global Topological Hitting Set for that complex.

In this work, we undertake an algorithmic study of the two high-dimensional cut problems: Boundary nontrivialization and Topological Hitting Set and their global variants.

TOPOLOGICAL HITTING SET ON SURFACES. The search for a tractable algorithm for Topological Hitting Set for the special case of surfaces was motivated by the intuitive example of 1-cycles on a torus, as shown in Figure 13. From the figure, it is clear that any cycle homologous to ζ must pass through cocycles η and ϑ. Additionally, ϑ is a cocycle with the smallest set of edges satisfying this property. Even more, every topological hitting set of ζ has at least as many edges as the support of ϑ.

Building on the intuition from the example in Figure 13 we provide an polynomial-time algorithm in Section 5 of Paper V for computing the optimal solution for Topological Hitting Set on closed triangulated surfaces. Here, we rely on an interesting characterization of the minimal solutions in terms of the cocycles of the surface. By a *connected cocycle* we mean a cocycle that induces

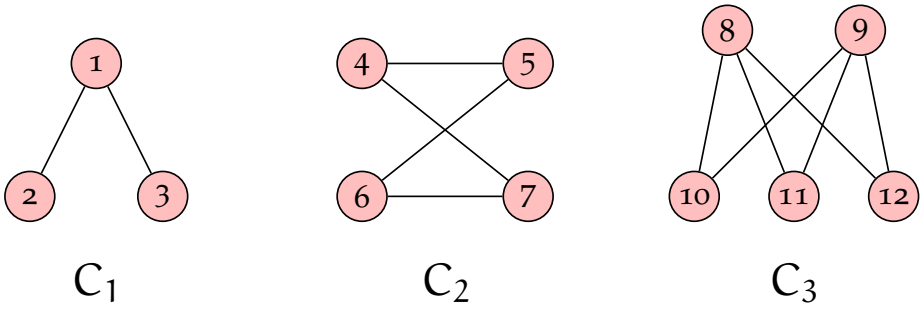


Figure 11: Graph G with three components

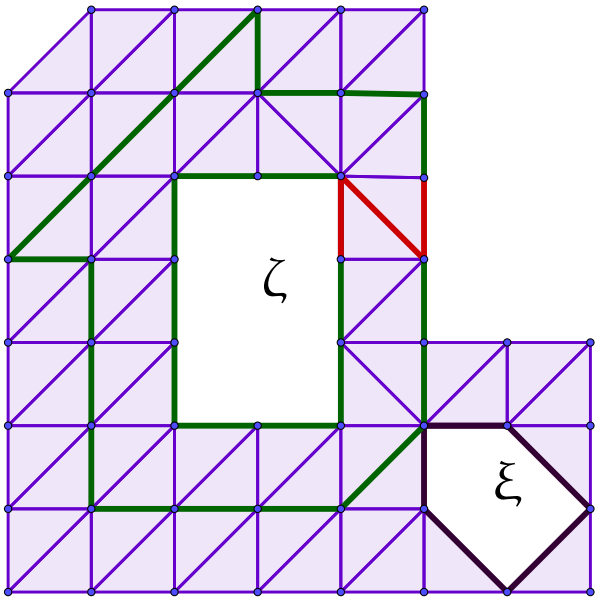


Figure 12: The figure shows two cycles that belong to $[\zeta]$ in green. Note that any cycle in $[\zeta]$ must pass through at least one of the three red edges. Thus, the set of red edges constitutes an optimal solution for Topological Hitting Set on this planar complex.

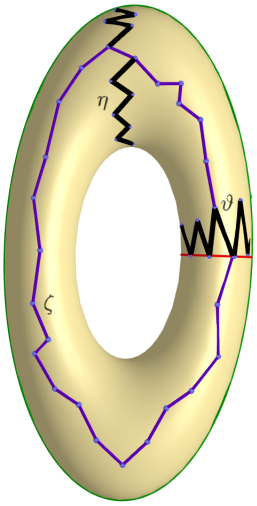


Figure 13: The figure shows a triangulated torus T. The cycle ζ is shown in purple. The cocycles η and ϑ "transversal" to ζ are shown in black. Any cycle homologous to ζ must pass through at least one of the edges of η (or ϑ). However, ϑ has a smaller support, and is, therefore, a more desirable "hitting set".

a connected component in the dual graph. By a *feasible set* for input ζ, we mean a set of edges that meets every cycle homologous to ζ. To begin with, we show that a minimal solution set is necessarily a nontrivial connected cocycle. Next, we show that for a connected cocycle η, the following are equivalent:

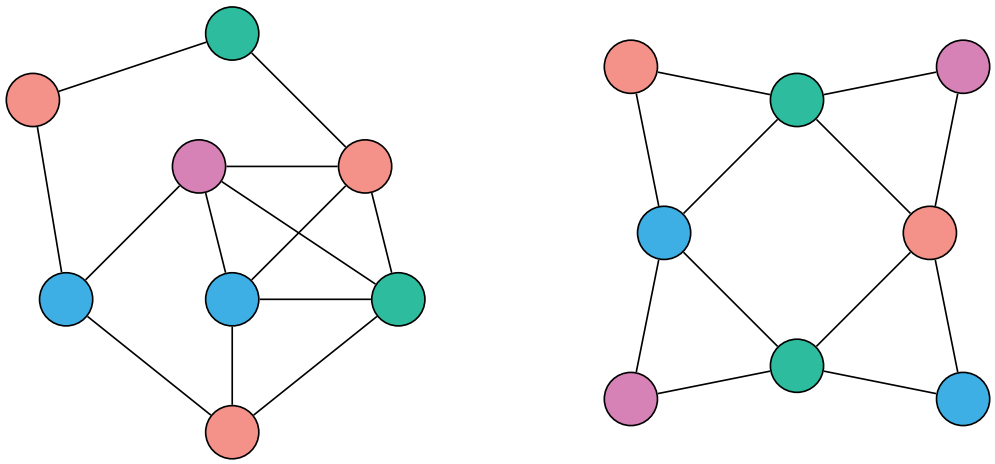
- η is a feasible set for the input cycle ζ.
- $\eta(\gamma) = 1$ for every cycle in $\gamma \in [\zeta]$.
- There exists a cycle $\gamma \in [\zeta]$, for which, $\eta[\gamma] = 1$.

Even more, cocycles ϑ for which $\vartheta(\zeta) = 0$, and any (cocycles cohomologous to) linear combinations of such cocycles are not solutions.

Specifically, this allows us to identify the nontrivial cocycles that are solutions based on a parity-based property. Having this characterization at hand, we eventually arrive at a very simple 3-step algorithm for Topological Hitting Set on surfaces that we describe below.

1. Find the optimal cohomology basis of the surface.
2. Sort the basis by ascending order of weights.
3. The lowest weight basis element η for which $\eta(\zeta) = 1$ is the desired solution.

For further details on the algorithm for Topological Hitting Set on surfaces, we refer the reader to Section 5 of Paper V. Finally, we remark that Boundary



G_1 G_2

Figure 14: In this figure, the graph G_1 on the left is a 4-colored graph with a multicolored 4-clique, whereas the graph G_2 on the right is a 4-colored graph that does not have a multicolored 4-clique.

NONTRIVIALIZATION is trivial for surfaces. In fact, it is easy to check that for some boundary b and a 2-chain ζ, if $\partial\zeta = b$, then removing any one of the triangles that appears in the chain ζ makes b nontrivial.

MULTICOLORED CLIQUES. A k-*clique* of a graph G is a complete subgraph of G with k vertices. A k-*coloring* of a graph G is an assignment of one of k possible colors to every vertex of G such that adjacent vertices do not receive the same color. A graph G equipped with a k-coloring is called a k-*colored graph*. A *multicolored* k-*clique* in a k-colored graph is a k-clique with a k-coloring. k-MULTICOLORED CLIQUE asks for the existence of a multicolored k-clique in a k-colored graph G. k-MULTICOLORED CLIQUE is known to be a **W[1]**-complete problem [77]. We establish **NP**-hardness and **W[1]**-hardness for TOPOLOGICAL HITTING SET and BOUNDARY NONTRIVIALIZATION by devising **FPT**-reductions from k-MULTICOLORED CLIQUE. Please see Figure 14 for an example and a non-example of a multicolored 4-clique.

W[1]-HARDNESS AND NP-HARDNESS. For general complexes, in Section 6.1 of Paper V, we show that TOPOLOGICAL HITTING SET is **W[1]**-hard with respect to the solution size k as the parameter, (and hence, it is also **NP**-hard). The proof is based on an **FPT**-reduction from the k-MULTICOLORED CLIQUE problem. In this reduction, given a k-colored graph G, we construct a complex $K(G)$. Owing to this reduction, if there exists an FPT algorithm for TOPOLOGICAL HITTING SET on $K(G)$, then there exists an FPT algorithm for k-MULTICOLORED CLIQUE on G. Since k-MULTICOLORED CLIQUE is known to be **W[1]**-hard [77], under standard

complexity theoretic assumptions, an FPT algorithm for TOPOLOGICAL HITTING SET with solution size as parameter is unlikely to exist.

In addition, in Section 6.2 of Paper V, we show that BOUNDARY NONTRIVIALIZATION is also **W[1]**-hard with respect to the solution size k as a parameter. The principles of this reduction follow the lines of the reduction for TOPOLOGICAL HITTING SET, though, here, both the description and the proof of the reduction are far more involved because of a special kind of subdivision that help avoid some unhelpful incidences.

FIXED-PARAMETER TRACTABILITY. On the positive side, in Section 7.1 of Paper V, we show that TOPOLOGICAL HITTING SET admits an FPT algorithm with respect to $k + \Delta$, where Δ is the maximum degree of the Hasse graph of the complex K. Here, the main insight is that a minimal solution induces a connected subgraph in the Hasse graph. Having this insight at hand, the algorithm follows: If we search across the geodesic ball of every r-simplex in the complex K, we will find a solution.

In contrast, we observe that BOUNDARY NONTRIVIALIZATION does not admit this property because minimal solutions can be disconnected. This motivates the search of another parameter that makes the problem tractable. Exploiting the set-cover like structure of the problem, in Section 7.2 of Paper V, we show that BOUNDARY NONTRIVIALIZATION with bounding r-cycles as input has an $O(\log n)$-approximation FPT algorithm with β_{r+1} (the Betti number) as the parameter, when the input complex K is $(r+1)$-dimensional. This provides a nice contrast with the hardness result: BOUNDARY NONTRIVIALIZATION is **W[1]**-hard even for $(r+1)$-dimensional complexes with solution size as the parameter since the hardness gadget used in Section 6.2 of Paper V is $(r+1)$-dimensional.

Finally, by exploiting the vector space structure of the homology groups and the boundary groups, in Sections 7.1 and 7.2, we also provide a randomized FPT algorithm for GLOBAL TOPOLOGICAL HITTING SET and a randomized FPT approximation algorithm for GLOBAL BOUNDARY NONTRIVIALIZATION respectively.

We believe that the FPT algorithms for GLOBAL TOPOLOGICAL HITTING SET and GLOBAL BOUNDARY NONTRIVIALIZATION can be derandomized. We also believe that the gadgets in Section 6 can be used to show **W[1]**-hardness of GLOBAL TOPOLOGICAL HITTING SET and GLOBAL BOUNDARY NONTRIVIALIZATION. However, these two results have not been written yet and are not a part of this thesis.

1.4.2 *Outlook & Future Directions: High-dimensional Cuts and Expansion*

Given a graph $G = (V, E)$, the *Cheeger constant* (or *edge expansion*) of a graph G, denoted by $h(G)$, is defined as

$$h(G) = \min\left\{ \frac{|E(S, V \setminus S)|}{|S|} \;\middle|\; S \subseteq V, |S| \in \left(0, \tfrac{|V|}{2}\right] \right\}.$$

A sequence of distinct graphs $\{G_i\}_{i=1}^{\infty}$ is called a *family of expander graphs* if there exists a constant $\delta > 0$ such that $h(G_i) > \delta$ for every $i \geqslant 1$. Expander graphs have many applications in computer science including design of robust computer networks, theory of error correcting codes and complexity theory [93]. For instance, Irit Dinur's simpler proof of the PCP theorem [59] was one of the landmark applications of expander graphs in complexity theory.

HIGH DIMENSIONAL EXPANSION. Generalizing the notion of expansion in graphs to simplicial complexes, a rich theory of high-dimensional expansion has emerged [116]. Already, in a short time since inception, high dimensional expanders have found applications in complexity theory [61], coding theory [58, 60], and counting and sampling algorithms [6, 7]. Here, we recall some of the basic definitions in order to motivate some connections to the ideas developed in Paper V.

First, note that when working with $\mathbb{Z}_2$ coefficients, the subsets S of $K^{(p)}$ are in one-to-one correspondence with the p-cochains of K. So notationally, we do not distinguish a set $S \subset K^{(p)}$ from the unique cochain associated to it. Further, let $\|\eta\|$ denote the size of the support of a cochain η. That is,

$$\operatorname{supp}(\eta) = \left\{\sigma \in K^{p+1} \mid \eta(\sigma) \neq 0\right\}, \quad \text{and} \quad \|\eta\| = |\operatorname{supp}(\eta)|.$$

The cosystolic norm of a p-cochain φ is defined as

$$\|[\varphi]\| = \min\{|\operatorname{supp}(\varphi + \delta\psi)| : \psi \in C^{p-1}(K)\}.$$

Note that if we let

$$\vartheta = \arg\min\{|\operatorname{supp}(\varphi + \delta\psi)| : \psi \in C^{p-1}(K)\},$$

then ϑ is the smallest cochain with $\delta\varphi$ as its coboundary.

Then, the p-*th coboundary expansion* of a complex K is defined as

$$h^p(K) = \min_{\zeta \in C^p(K) \setminus B^p(K)} \frac{\|\delta\zeta\|}{\|[\zeta]\|}.$$

Since a graph is a 1-dimensional complex, letting δS denote the coboundary of the 0-cochain S, the Cheeger constant can also be written as

$$\begin{aligned} h(G) &= \min\left\{ \frac{\|\delta S\|}{\|S\|} \;\middle|\; S \subseteq V, |S| \in \left(0, \tfrac{V}{2}\right] \right\} \\ &= \min_{S \in C^0(G)} \frac{\|\delta S\|}{\|[S]\|}. \end{aligned}$$

Just as the Cheeger constant of disconnected graphs is 0, the p-th coboundary expansion of complexes with nonzero p-th Betti number is 0. Hence, coboundary expansion is not a very useful notion to study the connectivity of complexes with nontrivial homology. In such cases, one of the alternative notions used is the so-called p-*th cocycle expansion* $z^p(K)$ of a simplicial complex K, defined as

$$z^p(K) = \min_{\zeta \in C^p(K) \setminus Z^p(K)} \frac{\|\delta\zeta\|}{\min_{\eta \in Z^p(K)} \|\eta + \zeta\|}.$$

Let ξ be a cochain for which $z^p(K) = \frac{\|\delta\xi\|}{\min_{\eta \in Z^p(K)} \|\eta+\xi\|}$.

Also, let $K_{\delta\xi}$ be the complex induced by the removal of simplices in the support of $\delta\xi$. Let ι be the inclusion map $\iota : K_{\delta\xi} \hookrightarrow K$. Then, it is easy to check that the induced map on cohomology $\hat{\iota} : H^p(K_{\delta\xi}) \to H^p(K)$ is not injective.

Also, we conclude that, $\xi \in Z^p(K_{\delta\xi}) \setminus B^p(K_{\delta\xi})$ because

- $\xi \in Z^p(K_{\delta\xi})$ since $\delta\xi = 0$ in $K_{\delta\xi}$,
- $\xi \in C^p(K) \setminus Z^p(K) \implies \xi \notin B^p(K)$,
- Finally, $B^p(K_{\delta\xi}) \cong B^p(K) \implies \xi \notin B^p(K_{\delta\xi})$.

In other words, ξ is a nontrivial cocycle in $K_{\delta\xi}$.

$$\text{Let } \vartheta = \arg\min_{\eta \in Z^p(K)} \|\eta + \xi\|.$$

If the cocycle ϑ created by the removal of $\delta\xi$ is large, then intuitively we can associate the removal of simplices in the support of $\delta\xi$ with the birth of a large feature. Then, if $z^p(K)$ is small, the coboundaries of K are *not robust* in the sense that removal of only a few $p+1$-simplices gives rise to large topological feature. In this sense, the quantity $z^p(K)$ captures the essence of *robustness of coboundaries*. If there exists a family of complexes $\mathscr{K}$ such that for every $K \in \mathscr{K}$, $z^p(K)$ is bounded away from 0, then the complexes in $\mathscr{K}$ are p-cocycle expanders. Alternatively, for our purposes, we say that the p-coboundaries of the complexes in $\mathscr{K}$ are robust. There is a further refinement of the notion of cocycle expansion that is called *cosystolic expansion*, which we will not define here. Please refer to [74] for a definition.

CYCLE EXPANSION. Motivated by the preceding discussion, we will define a notion analogous to cocycle expansion that we will call the p-*th (hitting) cycle expansion*. Let S be a p-chain of K, and let K_S be the complex induced by the removal of simplices in the support of S from K. Let $\iota : K_S \hookrightarrow K$ be the natural inclusion map. Then, we say that a p-chain S is a p-*topological hitting set* of complex K if the induced map on homology $\hat{\iota} : H_p(K_S) \to H_p(K)$ is not surjective. The collection of all p-topological hitting sets of K is denoted by $\mathscr{T}_p(K)$. We say that a nontrivial p-cycle ζ of K is *destroyed* by the removal of

simplices in the support of S if $[\zeta]$ is not in the image of $\hat{\imath}$. Let $\Upsilon_p^S(K)$ denote the set of all non-trivial p-cycles of K that are destroyed by the removal of S. We define the p-*th cycle expansion* of a simplicial complex K as

$$z_p(K) = \min_{S \in \mathscr{T}_p(K)} \frac{\|S\|}{\min\limits_{\zeta \in \Upsilon_p^S(K)} \|\zeta\|}.$$

Analogous to the earlier discussion,

$$\text{Let } \xi_S = \min_{\zeta \in \Upsilon_p^S(K)} \|\zeta\|.$$

If the smallest cycle ξ_S created by the removal of S is large, then intuitively we can associate the removal of S with the death of a large feature. Then, if $z_p(K)$ is small, the cycles of K are *not robust* in the sense that removal of only a few p-simplices kills a large topological feature. In this sense, the quantity $z_p(K)$ captures the essence of *robustness of cycles*. If there exists a family of complexes $\mathscr{K}$ such that for every $K \in \mathscr{K}$, the quantity $z_p(K)$ is bounded away from 0, then we say that the complexes in $\mathscr{K}$ have p-cycle expansion. Alternatively, for our purposes, we say that the p-cycles of the complexes in $\mathscr{K}$ are robust.

To the best of our knowledge, cycle expansion, as we define it here, has not been studied before. We believe that it captures "expansion-like" phenomena in higher dimensions, and studying the properties of this invariant would be of independent interest. We end this discussion, and also this section with a few open questions about p-cycle expansion.

1. We established in Paper V that optimal topological hitting sets induce a connected subgraph in the Hasse graph. What other topological properties can one establish for topological hitting sets, and what kind of statements hold true for cycle expanders? Dotterer et al. [62] showed that if K is a cosystolic expander, then it satisfies Gromov's topological overlap property, which can be described as follows. Given a d-dimensional complex K, for every continuous map $K \to \mathbb{R}^d$ there exists a point $p \in \mathbb{R}^d$ that is contained in the images of a positive fraction $\mu > 0$ of the d-simplices of K. Complexes that satisfy the topological overlap property are called *topological expanders*. The next question we ask is: how is cycle expansion related to topological expansion?

2. Like in the case of graphs, it is more interesting to look at *sparse* complexes that have good expansion properties. In this regard, there is a line of work in finding infinite families of bounded degree coboundary, cocycle and cosystolic expanders [74,103,117]. Here, by bounded degree we mean that each vertex in the complex is incident on a bounded number of simplices. This raises a natural question: For an integer $p > 1$, does there exist an infinite family of bounded degree complexes whose p-th cycle expansion is bounded below by a fixed constant? It is known that the so-called

> Ramanujan complexes have nontrivial homology, and hence they are not coboundary expanders [116]. And yet skeletons of Ramanujan complexes are cosystolic and topological expanders [74]. A natural question arises: Are Ramanujan (or related) complexes cycle expanders?

Just as graph expanders model the robustness of connectivity in networks, cocycle expanders and cycle expanders capture the robustness of (co)homology in complexes.

2

ALGORITHMIC PRELIMINARIES

2.1 PARAMETERIZED COMPLEXITY

Notation 2. In this section, we use $[m]$ to denote the set $\{1, 2, \ldots, m\}$ for any $m \in \mathbb{N}$, and $[a, b]$ to denote the set $\{a, a+1, \ldots, b\}$.

2.1.1 *Basic notions from graph theory*

An *(undirected) graph* is a pair $G = (V, E)$, where V denotes a set of elements called *vertices*, and E denotes a set of paired vertices, whose elements are called *edges*. A *simple graph* is a graph that does not have more than one edge between any two vertices and none of the edges start and end at the same vertex. In this section, we consider only simple graphs. Simple graphs are 1-dimensional simplicial complexes. A *subgraph* $H = (V_H, E_H)$ of a graph $G = (V, E)$ is a graph with $V_H \subseteq V$ and $E_H \subseteq E$. We say that H is a subgraph of G *induced* by the vertex set V_H if $E_H = \binom{V_H}{2} \bigcap E$. The deletion of a vertex v from graph G results in a graph denoted by $G - v$, which is the subgraph of G induced by the vertex set $V \setminus \{v\}$. The deletion of an edge e from a graph G results in a graph $G - e = (V, E \setminus \{e\})$. The contraction of an edge $f = \{u, v\}$ results in a graph $G \circ f = (V \circ f, E \circ f)$, where $V \circ f = V \setminus \{v\}$ and $E \circ f = E \setminus \{e \mid v \in e\} \bigcup \{\{u, w\} \mid \{v, w\} \in E, w \neq u\}$. An graph H is called a *minor* of a graph G if H can be obtained from G following a (possibly empty) sequence of edge contractions, edge deletions and vertex deletions.

A *complete graph* is a graph in which every pair of distinct vertices is connected by an edge. A complete graph on n vertices has $\binom{n}{2}$ edges and is denoted by K_n. A graph $G = (V, E)$ is called a *biparite graph* if V can be paritioned into two disjoint subsets V_1 and V_2 such that for any two vertices $u, v \in V$, if u and v belong to the same subset, then $\{u, v\} \notin E$. A bipartite graph $G = (V, E)$ with partition $V = V_1 \sqcup V_2$ is called a complete bipartite graph if for every $u \in V_1$ and $v \in V_2$, the edge $\{u, v\}$ is in E. A complete bipartite graph with $|V_1| = m$ and $|V_2| = n$ is denoted by $K_{m,n}$. Note that $K_{m,n}$ is isomorphic to $K_{n,m}$. A *planar graph* is a graph that can be drawn in the Euclidean plane with vertices represented by points, and edges represented by curves that connect neighboring vertices such that the edges intersect only at their endpoints. Equivalently, a planar graph is a graph that can be embedded in a plane. By Fáry's theorem [96], planar graphs can be drawn with edges as straight line segments without crossings. An *outerplanar graph* is a graph that can be embedded in the Euclidean plane such that all vertices lie on the outer face. A

path in a graph is a sequence of vertices joined by edges. A *tree* is a graph in which any two vertices are connected by exactly one path. A *forest* is a graph in which any two vertices are connected by at most one path. Given a graph $G = (V, E)$, the complement graph of G, denoted by $\overline{G} = (V, \overline{E})$, where $\overline{E}$ is given by $\overline{E} = \binom{V}{2} \setminus E$.

A *vertex cover* of a graph is a set of vertices that is incident on every edge of the graph. A *dominating set* for a graph $G = (V, E)$ is a set $D \subset V$ such that every vertex that is not in D is adjacent to at least one vertex in D. An *independent set* in a graph is a set of vertices, no two of which are connected by an edge. It is easy to check that a vertex set S is a vertex cover in a graph $G = (V, E)$, if and only if its complement $S' = V \setminus S$ is an independent set in G. Given a graph $G = (V, E)$, we say that a set $X \subseteq V$ is a *feedback vertex set* of G, if the subgraph of G induced by the vertex set $V \setminus X$ has no cycles. A k-*clique* of a graph G is a complete subgraph of G with k vertices. A k-*coloring* of a graph G is an assignment of one of k possible colors to every vertex of G (in other words, a vertex coloring) such that no two vertices that share an edge receive the same color. A graph G equipped with a k-coloring is called a k-*colored graph*. A *multicolored* k-*clique* in a k-colored graph is a k-clique with a k-coloring.

The notions of treewidth, pathwidth and branchwidth were introduced by Robertson and Seymour in their work on graph minors [140, 141, 142]. Since these constructs play an important role in parameterized complexity, we provide a brief description below.

First, we will define the notion of path decompositions. Given a graph $G = (V, E)$, a *bag* (of vertices in V) is simply a subset of vertices of V. Let $\mathcal{X}$ be a collection of bags indexed by W. That is, $\mathcal{X} = \{X_i \mid i \in W\}$, where each X_i is a bag of vertices in V. Let $\mathcal{P} = (W, F)$ be a path graph whose vertices are in one-to-one correspondence with the bags in $\mathcal{X}$. A *path decomposition* of G is a pair $(\mathcal{X}, \mathcal{P})$ that satisfies the following conditions.

- Every vertex of G is in at least one bag. That is, $\bigcup_{i \in W} X_i = V$.
- For every edge $e = \{u, v\} \in E$, there exists an element $q \in W$ such that the bag X_q contains both u and v.
- For all $i, j, k \in W$, if j is on the path from i to k in $\mathcal{P}$, then $X_i \bigcap X_k \subseteq X_j$. In other words, for every $v \in V$, the set of nodes of $\mathcal{P}$ whose corresponding bags contain v induces a sub-path in $\mathcal{P}$.

The width of a path decomposition $(\mathcal{X}, \mathcal{P})$ of a graph G is $\max_{\ell \in W} |X_\ell| - 1$. The pathwidth $\mathrm{pw}(G)$ of a graph G is the minimum possible width over all possible path decompositions of G. The pathwidth of a graph captures how close a graph is to a path. For instance, the pathwidth of a path graph with more than one edges is 1.

The notion of a tree decompositions generalizes the notion of a path decomposition in a natural way: A *tree decomposition* of a graph $G = (V, E)$ is a pair

$(\mathcal{X}, \mathcal{T})$, where $\mathcal{T} = (W, F)$ is a tree, and $\mathcal{X} = \{X_i \mid i \in W\}$ is a collection of bags of vertices in V, one for each node of $\mathcal{T}$ that satisfy the following conditions.

- Every vertex of G is in at least one bag. That is, $\bigcup_{i \in W} X_i = V$.
- For every edge $e = \{u, v\} \in E$, there exists an integer $q \in W$ such that the bag X_q contains both u and v.
- For every $v \in V$, the set of nodes of $\mathcal{T}$ whose corresponding bags contain v induces a connected subtree of $\mathcal{T}$.

The width of a tree decomposition $(\mathcal{X}, \mathcal{T})$ is $\max_{\ell \in W} |X_\ell| - 1$. The treewidth $tw(G)$ of a graph G is the minimum possible width over all possible tree decompositions of G. The treewidth of a graph is one of the most powerful tools for designing parameterized algorithms. It measures how tree-like a graph is. In particular, the treewidth of a tree with more than one edges is 1.

A branch decomposition [81] of a graph $G = (V, E)$ is a pair $(\mathcal{T}, \tau)$, where $\mathcal{T} = (V_T, E_T)$ is a tree whose vertices either have degree 1 (leaves) or degree 3 (internal nodes), and τ is a bijection from the set of leaves of $\mathcal{T}$ to E. Since $\mathcal{T}$ is a tree, the removal of an edge $e \in E_T$ from $\mathcal{T}$ results in two connected components. We say that a vertex $v \in V$ *lies in the cut* of $e \in E_T$, if there exist leaves $t_1, t_2 \in \mathcal{T}$ in different components with $v \in \tau(t_1)$ and $v \in \tau(t_2)$. The *cut-value* $\gamma(e)$ of an edge $e \in E_T$ is the number of vertices v of G that lie in the cut of e. The width of the branch decomposition $(\mathcal{T}, \tau)$ is defined as $\max_{e \in E_T} \gamma(e)$. The *branchwidth* $bw(G)$ of G is the minimum width over all the branch decompositions of G. The exposition in [81] provides a helpful illustration.

Theorem 2.1.1 ([142]). *If G is a graph, then*

$$bw(G) \leqslant tw(G) + 1 \leqslant \left\lfloor \frac{3}{2} bw(G) \right\rfloor .$$

As is evident from Theorem 2.1.1, the notions of branchwidth and treewidth are closely related. Although the notion of treewidth is the more commonly encountered one in the design of algorithms, two prominent exceptions are dynamic programming algorithms and algorithms for planar graphs. In these two cases, branchwidth is easier to work with [50].

2.1.2 *Mathematical origins of parameterized complexity*

Parameterized complexity was inspired by Robertson and Seymour's graph minor theory. Below, we provide a brief account of the mathematical underpinnings of parameterized complexity. For a short and informative account of graph minor theory we refer the reader to an article by Lovasz [115].

To begin with, it is easy to check that

1. Every graph is a minor of itself.

2. If K is a minor of H, and H is a minor of G, then K is a minor of G.

Hence, the minor ordering on graphs constitutes a preorder (that is, a reflexive, transitive relation). A *well-quasi-ordering* is a preorder that contains neither an infinite descending chain nor an infinite antichain.

A graph family $\mathcal{F}$ is a *minor-closed* family if any minor of any graph in $\mathcal{F}$ is also a member of $\mathcal{F}$. Examples of minor-closed families of graphs include forests, planar graphs, and outerplanar graphs.

We say that a graph parameter is *minor-closed* if it either decreases or remains the same upon edge contractions, edge deletions and vertex deletions. Examples of minor-closed graph parameters include vertex cover, feedback vertex set, treewidth, pathwidth and branchwidth.

The celebrated theorem by Wagner [154] characterizes planar graphs as graphs that do not have the complete graph K_5 or the complete bipartite graph $K_{3,3}$ as minors. A far-reaching generalization of Wagner's theorem is the Robertson-Seymour theorem. The graph minor theory of Robertson and Seymour took more than twenty years to develop and is regarded as one of the monumental feats of modern mathematics. The main result can be stated as follows.

Theorem 2.1.2 (Robertson-Seymour [143])**.** *Let $\mathcal{G}$ be a minor-closed family. Then, there exists a finite set* $\mathrm{ob}(\mathcal{G})$*, called the obstruction set of $\mathcal{G}$, such that for each graph* G*, we have* $G \in \mathcal{G}$ *if and only if none of its minors are in* $\mathrm{ob}(\mathcal{G})$*.*

It is easy to check that the Robertson-Seymour theorem can be formulated in the following equivalent way.

Corollary 1 (Robertson-Seymour)**.** *Undirected graphs partially ordered by the graph minor relation forms a well-quasi-ordering.*

Given a graph G of size n and a graph H of size c, one of the consequences of Robertson-Seymour theory is an $f(c)n^3$ time algorithm for checking if H is a minor of G, where f is a superpolynomial function [139]. This was later improved to $f(c)n^2$ time algorithm by Kawarabayashi et al. [95]. Combining these algorithms with Theorem 2.1.2, we obtain the following corollaries.

Corollary 2. *There exists polynomial-time algorithm for deciding the membership of a graph in a minor-closed family.*

Corollary 3. *Deciding if a minor-closed graph parameter is at most k can be done in $f(k)n^2$ time, where n denotes the size of the input graph, and f is a superpolynomial function.*

We make a few important remarks about the algorithmic consequences of the above corollaries.

Remark 2. The algorithm in Corollary 2 can be used only if the obstruction sets are known in advance for the specific minor-closed families under consideration. More often than not, this is not the case. So, without prior knowledge of obstruction sets, Corollary 2 should be seen as a non-constructive result about the existence of a polynomial time algorithm.

Remark 3. Distinct values of k in Corollary 3 give rise to distinct minor-closed graph families. Hence, Corollary 3 is a non-constructive result about the existence of a non-uniform polynomial time algorithm.

Remark 4. Although the algorithms in Corollaries 2 and 3 are polynomial in instance size, the hidden constants that depend on some fixed parameters may be huge.

It is this connection with algorithms and complexity that motivated Downey and Fellows to initiate the study of parameterized complexity. Please refer to Downey [64] for an insightful personal account about the origins of parameterized complexity. It is worth noting that there exist some fundamental problems arising from graph minor theory for which the best-known parameterized algorithms are "galactic" [114]. Galactic algorithms are algorithms with good-looking asymptotic bounds, but with concrete costs high enough to prevent their use on scales smaller than the known universe [112]. This can happen, for instance, when the parameter function grows at the rate of a tower of exponentials. Designing faster algorithms for these problems is an area of active research, and gets to the heart of some foundational questions [113, 114]. Having said that, it is also worth mentioning that the current practice of parameterized algorithms is only loosely connected to foundational algorithmic questions arising from graph minor theory. In fact, another trend in parameterized complexity is to design parameterized algorithms with practical use in mind [1]. Arguably, the parameterized algorithms in this thesis (specifically those in Paper V) have reasonable parameter dependence.

2.1.3 *Fixed parameter tractability*

Parameterized complexity was introduced by Downey and Fellows in [63] as a refinement of classical complexity theory. The theory hinges on the idea of developing complexity bounds based on instance size along with an additional *parameter*, which might be significantly smaller than the instance size. Table 5 provides a informal summary of key notions in parameterized complexity.

We now provide some formal definitions.

Definition 2.1.1 (Parameterized problem, FPT algorithm [80])**.** Let Σ be a finite alphabet, and Σ^* the set of strings over Σ.

1. A *parameter* of Σ^* is a function $\rho : \Sigma^* \to \mathbb{N}$ that associates a natural number $\rho(w)$ to every input $w \in \Sigma^*$.

	Classical complexity	Parameterized complexity
Instance	**Problem** A INSTANCE: Graph G or complex K of size n, $r \geqslant 0$. QUESTION: Is OPT $\leqslant r$?	**Problem** A' INSTANCE: G or K of size n, **param.** $\leqslant k$, $r, k \geqslant 0$. PARAMETER: k QUESTION: Is OPT $\leqslant r$?
Algorithms	Polynomial time algorithm: runs in $p(n)$ time	FPT algorithm: runs in $f(k) \cdot p(n)$ time
Complexity classes	**NP**	**W**-hierarchy
Reductions	Karp reductions to prove NP-hardness	FPT reductions to prove **W**-hardness

Table 5: The table provides a comparative summary of notions in classical and parameterized complexity. Pardigmatically, the desirable or the tractable algorithms in classical compexity are the polynomial time algorithms, whereas the tractable algorithms in the parameterized world are the fixed parameter tractable (FPT) algorithms. Under standard complexity theoretic assumptions, in the realm of classical complexity, polynomial time algorithms are ruled out by establishing a Karp reduction from a known **NP**-hard problem to the problem at hand. Likewise, FPT algorithms are ruled out by establishing FPT reductions.

2. A *parameterized problem* over Σ is a pair (P, ρ) that consists of a set $P \subseteq \Sigma^*$ and a parametrization $\rho : \Sigma^* \to \mathbb{N}$.

3. A parameterized problem (P, ρ) is said to be *fixed-parameter tractable* or *FPT in the parameter* ρ if the question

$$(x, p) \in \{(y, \rho(y)) \mid y \in P\}$$

can be decided in running time $O(g(p)) \cdot |x|^{O(1)}$, where g is an arbitrary computable function of the parameter p. The corresponding decision algorithm is called a *fixed parameter tractable (FPT) algorithm*.

SOLUTION SIZE AS THE PARAMETER. Let k-VERTEX COVER, k-CLIQUE, k-INDEPENDENT SET, and k-DOMINATING SET denote the vertex cover problem, the clique problem, the independent set problem, and the dominating set problem, respectively, each parameterized by the solution size. The formal definitions are as follows.

Problem 1 (k-VERTEX COVER).
INSTANCE: A graph $G = (V, E)$, and an integer k
PARAMETER: k
QUESTION: Does there exist a vertex cover in G of size k?

Problem 2 (k-CLIQUE).
INSTANCE: A graph $G = (V, E)$, and an integer k
PARAMETER: k
QUESTION: Does there exist a k-clique in G?

Problem 3 (k-INDEPENDENT SET).
INSTANCE: A graph $G = (V, E)$, and an integer k
PARAMETER: k
QUESTION: Does there exist an independent set in G of size k?

Problem 4 (k-DOMINATING SET).
INSTANCE: A graph $G = (V, E)$, and an integer k
PARAMETER: k
QUESTION: Does there exist a dominating set in G of size k?

Another problem of interest to us is called k−MULTICOLORED CLIQUE.

k−MULTICOLORED CLIQUE asks for the existence of a multicolored k-clique in a k-colored graph G. Formally, k−MULTICOLORED CLIQUE is defined as follows:

Problem 5 (k-MULTICOLORED CLIQUE).
INSTANCE: A graph $G = (V, E)$, and a vertex coloring $c : V \to [k]$.
PARAMETER: k.
QUESTION: Does there exist a multicolored k-clique H in G?

In parameterized complexity, one uses FPT reductions to establish hardness results. We begin with a definition.

Definition 2.1.2 (FPT reduction [80])**.** We say that there is an *FPT reduction* from a parameterized problem (P, k) to a parameterized problem (Q, k') if there exists a function ϕ that transforms parameterized instances of P to parameterized instances of Q while satisfying the following properties:

1. ϕ is computable by an FPT algorithm.
2. $\phi(x)$ is a yes-instance of (Q, k') if and only if x is a yes-instance of (P, k).
3. There exists a computable function $h\colon \mathbb{N} \to \mathbb{N}$ such that
$$k'(\phi(x)) \leqslant h(k(x)).$$

The obvious way of turning a minimization problem into a decision problem is to add a value k to the input instance, and seek a solution with cost at most k. If we take this value k appearing in the input as the parameter, it is called *the natural parameterization* of the minimization problem. Some researchers refer to such a parameterization as simply "parameterized by solution size".

For maximization problems, usually it is not clear when one can expect the solution size to be small. Hence, in most cases it would be unnatural to

parameterize by solution size. One alternative would be to parameterize the problem by $n - k$ instead of k, where k is the solution size and n is the instance size. For instance, in the case of the maximum independent set problem, parameterizing by $n - k$ where k is the solution size, is equivalent to parameterizing the minimum vertex cover by the solution size since the complement of an independent set is a vertex cover. This kind of reparameterization can substantially affect the complexity of the problem. In fact, the maximum independent set problem parameterized by solution size is **W[1]**-hard (a notion we encounter in Section 2.1.4), whereas the minimum vertex cover problem parameterized by solution size is fixed parameter tractable. In other words, the reduction $(G, k) \to (G, n - k)$ is not an FPT reduction from k-Independent Set to k-Vertex Cover.

From a theoretical standpoint, a parameter can be *any* function of the input instance, for example, the treewidth of the input graph, or the maximum degree of the input graph, or the solution size of some problem. In practice, it is desirable to establish tractability with respect to a parameter that is expected to be small for a fairly broad class of interesting instances.

We now provide simple examples of FPT reductions for some classical problems in complexity theory.

Example 2.1.1 (k-Clique to k-Independent Set). There exists a clique in G of size k if and only if there exists an independent set in $\overline{G}$ of size k . Hence, the reduction $(G, k) \to (\overline{G}, k)$ is an FPT reduction from k-Clique to k-Independent Set.

Example 2.1.2 (k-Clique to k-Multicolored Clique). Let $(G = (V, E), k)$ be a parameterized instance of k-Clique. Construct a graph $H = (V_H, E_H)$ as follows. To each vertex $v \in V$, associate k vertices $\{v_i \mid i \in [k]\}$ in V_H. For each $v \in V$, and $i \in [k]$, color v_i with color i. Thus, $V_H = k|V|$. In graph H, vertices of the same color are not connected by an edge. Furthermore, $\{u_i, v_j\} \in E_H$ for $i \neq j$ if and only if $\{u, v\} \in E$. It is easy to check that G has a k-clique if and only if H has a multicolored k-clique. This gives an FPT reduction from k-Clique to k-Multicolored Clique.

Next, we look at a few examples of FPT algorithms for some standard problems in graph theory.

The naïve algorithm for k-Vertex Cover checks if any of the $\binom{n}{k}$ vertex sets of G of size k forms a vertex cover. Clearly, this algorithm is not an FPT algorithm. However, there exists a simple branching based FPT algorithm for k-Vertex Cover based on the following observation.

Remark 5. If (G, k) is a yes-instance of k-Vertex Cover, then for every edge $(u, v) \in E$, either u or v belongs to the solution set. Hence, the graph G has a vertex cover of size k if and only if at least one of the graphs $G - u$ or $G - v$ has a vertex cover of size at most $k - 1$.

Now consider the following recursive algorithm.

Algorithm 1 FPT algorithm for k-VERTEX COVER

```
1: procedure VCover(G = (V, E), k)
2:     if E = ∅ then
3:         return true;
4:     else if k = 0 then
5:         return false;
6:     else
7:         Select an edge e = {u, v} ∈ E;
8:         return VCover(G − u, k − 1) ∨ VCover(G − v, k − 1);
9:     end if
10: end procedure
```

Note that Algorithm 1 terminates after most k levels of recursion, and computing $G - u$ from G takes at most $O(n)$ time. Hence, the running time of the algorithm is $O(n \cdot 2^k)$, and k-VERTEX COVER $\in$ **FPT**.

TREEWIDTH AS THE PARAMETER. We will briefly discuss one of the most commonly encountered parameter in parameterized complexity, namely, treewidth. By Courcelle's theorem, every graph property definable in the monadic second-order logic of graphs can be decided in linear time on graphs of bounded treewidth [50, Theorem 7.11]. Burton and Fellows [31] generalized Courcelle's theorem to d-manifolds by showing that if an algorithmic problem defined on simplicial manifolds can be expressed in monadic second-order logic, then it can be solved in linear time for triangulations whose dual graphs have bounded treewidth. However, the hidden constants in the complexity bounds for these theoretical algorithms (guaranteed by Courcelle's algorithm and its extensions) can be prohibitively large.

Having said that, if a graph property admits a monadic second-order logic formula, then this usually indicates the existence of algorithms with better parameter dependence. Practitioners of paramaterized complexity typically use the so-called nice tree decompositions to design dynamic programming algorithms that are linear in the instance size and have merely exponential dependence on treewidth. Please refer to [50, Section 7] for more details on FPT algorithms parameterized by treewidth and [82, Chapter 14.5].for a nice discussion on monadic second-order logic on graphs.

2.1.4 *Fixed parameter intractability*

First, we recall some elementary notions from Boolean circuits. In particular, by an *and-node*, we mean the digital logic gate that implements logical conjuction ($\wedge$), by an *or-node*, we mean the digital logic gate that implements logical dis-

junction ($\vee$), and by a *not-node*, we mean the digital logic gate that implements negation ($\neg$).

Definition 2.1.3 (Boolean circuit)**.** A *Boolean circuit* C is a directed acyclic graph, where each node is labeled in the following way:

1. every node with in-degree greater than 1 is either an *and-node* or an *or-node*,
2. each node of in-degree 1 is labeled as a *negation node*,
3. and each node of in-degree 0 is an *input node*.

Exactly one of the nodes with out-degree 0 is labeled as the *output node*.

Below, we recall some essential parameterized complexity results concerning circuits. We use the terms *gates* and *nodes* interchangeably. We say that a gate has *fan-in* k if its in-degree is at most k. A *big gate* in a circuit is a gate with more than two (or some other fixed constant) inputs. The *depth* of a circuit is the longest input-output path in the circuit. The *weft* of a circuit is the maximum number of big gates on an input-output path. Clearly, depth of a circuit is always greater than its weft. Alternatively, weft is the largest number of gates with unbounded fan-in on the input-output path. A gate is called an *ordinary gate* if it is neither an input gate nor an output gate. We denote the nodes and edges in C by $V(C)$ and $E(C)$ respectively. The size of a circuit C, denoted by $|C|$, is the total number of nodes and edges in C. That is, $|C| = |V(C)| + |E(C)|$. The *Hamming weight* of an assignment is the number of input gates receiving value 1. An assignment on the input nodes induces an assignment on all nodes according the logical operations on the labels of the vertices. So given an assignment from the input nodes of circuit C to $\{0, 1\}$, we say that the assignment satisfies C if the value of the output node is 1 for that assignment. Let $\mathscr{G}_I$ denote the set of input gates of C. Then, an assignment A can be viewed as a binary vector of size $|\mathscr{G}_I|$. In the Weighted Circuit Satisfiability (WCS) problem, we are given a circuit C and an integer k, and the task is to decide if C has a satisfying assignment of Hamming weight at most k. Accordingly, in the Min-Weighted Circuit Satisfiability (MinWCS) problem, we are given a circuit C, and the task is to find a satisfying assignment with minimum Hamming weight.

Definition 2.1.4 (**W[P]**)**.** A parameterized problem Π belongs to the class **W[P]** if it can be reduced to the standard parameterization of WCS.

The **W**-hierarchy is a nested collection of complexity classes in parameterized complexity theory. A parameterized problem Π is in the class $\mathbf{W}[\mathbf{t}]$, if every instance (x, k) of Π can be transformed in FPT-time to a constant depth circuit with weft at most t, such that (x, k) is a yes-instance if and only if the circuit has a satisfying assignment of Hamming weight k. In the definition of $\mathbf{W}[\mathbf{t}]$, t

is a fixed constant. It is known that **FPT** = **W**[**0**] [63, Chapter 29]. Hence, the classes in the **W**-hierarchy can be expressed as:

$$\mathbf{FPT} = \mathbf{W}[\mathbf{0}] \subseteq \mathbf{W}[\mathbf{1}] \subseteq \mathbf{W}[\mathbf{2}] \subseteq \cdots \subseteq \mathbf{W}[\mathbf{P}] \subseteq \mathbf{XP}.$$

Every class in the **W**-hierarchy is closed under FPT reductions. Fascinatingly, most naturally arising computational problems belong to either **W**[**1**] or **W**[**2**]. k-CLIQUE (and k-INDEPENDENT SET) are canonical **W[1]**-complete problems, whereas k-DOMINATING SET is the canonical **W[2]**-complete problem. For a a nice illustration of constant depth weft-1 circuit (weft-2 circuit) for k-INDEPENDENT SET (k-DOMINATING SET), please see Fig. 13.3 in [50].

Several refined results in parameterized complexity are obtained on the basis of certain hypotheses that have stood the test of time. The two most prominent hypotheses go by the name of EXPONENTIAL TIME HYPOTHESIS (ETH) and STRONG EXPONENTIAL TIME HYPOTHESIS (SETH). Informally, the two hypotheses can be stated as follows.

EXPONENTIAL TIME HYPOTHESIS There is no $2^{o(n)}$ time algorithm for 3-SAT.

STRONG EXPONENTIAL TIME HYPOTHESIS There is no $O^*((2-\epsilon)^n)$ time algorithm[1] for CNF-SAT.

For formal definitions, we refer the reader to [50, Chapter 14]. Several important results have been proved conditioned on the fact that ETH is true. For instance, assuming ETH, the **W**-heirarchy is strict. That is,

$$\mathbf{FPT} = \mathbf{W}[\mathbf{0}] \subsetneq \mathbf{W}[\mathbf{1}] \subsetneq \mathbf{W}[\mathbf{2}] \subsetneq \cdots \subsetneq \mathbf{W}[\mathbf{P}] \subsetneq \mathbf{XP}.$$

ETH is also an invaluable tool for proving strong quantitative results of the following kind.

Theorem 2.1.3 (Chen et al. [40])**.** *Assuming ETH, there is no* $f(k)n^{o(k)}$ *algorithm for* k*-*CLIQUE *for any computable function* f*.*

Finally, ETH and SETH are two of the most important hypotheses for proving fine-grained complexity bounds [157].

2.2 APPROXIMATION ALGORITHMS

The reader is referred to excellent resources [9, 46] for a rigorous and complete treatment of **NP**-completeness. Below, we informally recall some key ideas.

An optimization problem Π is called an **NP**-*optimization problem* if

- the size of a feasible solution is polynomially bounded in the size of a valid instance,

1 Here, the notation $O^*()$ is used to suppress multiplicative polynomial factors.

- the valid instances and feasible solutions can be recognized in polynomial time, and
- the cost (or the objective function) for a solution is polynomial time computable.

The natural decision problem associated to an **NP**-optimization problem lies in **NP**. Let x be an instance of an **NP**-optimization problem Π and $k \in \mathbb{N}$. For an instance x of Π, let $\mathsf{OPT}_\Pi(x)$ denote the optimal value of the objective function of problem Π on input x.

- Then, for minimization problems, the corresponding decision problem with input (x, k) can be stated as: Is $\mathsf{OPT}_\Pi(x) \leqslant k$?,
- And, for maximization problems, the corresponding decision problem with input (x, k) can be stated as: Is $\mathsf{OPT}_\Pi(x) \geqslant k$?.

The decision problems above are **NP**-complete. In this section, we restrict our attention to **NP**-optimization problems. An α-approximation algorithm $\mathbb{A}$ for an **NP**-optimization problem Π is a polynomial-time algorithm that computes a solution within a factor α of the value of an optimal solution. Let $c_{\mathbb{A}}(x)$ denote the value of the objective function of the solution obtained by Algorithm $\mathbb{A}$ on input x.

If Π is a minimization problem, then $\mathbb{A}$ is an α-approximation algorithm if for all instances x of Π,

$$\frac{c_{\mathbb{A}}(x)}{\mathsf{OPT}_\Pi(x)} \leqslant \alpha.$$

If Π is a maximization problem, then $\mathbb{A}$ is an α-approximation algorithm if for all instances x of Π,

$$\frac{c_{\mathbb{A}}(x)}{\mathsf{OPT}_\Pi(x)} \geqslant \alpha.$$

Clearly, for maximization problems, $\alpha \leqslant 1$ and for minimization problems $\alpha \geqslant 1$. The factor α is typically a function of the instance size n. If $\mathbb{A}$ is an α-approximation algorithm and $\mathbb{B}$ is a β-approximation algorithm for a minimization problem Π, and if $\alpha = o(\beta)$, then $\mathbb{A}$ is considered a better approximation algorithm than $\mathbb{B}$. Likewise, if $\mathbb{A}$ is an α-approximation algorithm and $\mathbb{B}$ is a β-approximation algorithm for a maximization problem Π, and if $\alpha = \omega(\beta)$, then $\mathbb{A}$ is considered a better approximation algorithm than $\mathbb{B}$.

COMPLEXITY CLASSES Constant factor approximation algorithms are common and are often desirable. The class **APX** is the set of **NP**-optimization problems that admit constant-factor approximation algorithms. More generally, $f(n)$-**APX** is the class of problems with an $O(f(n))$-approximation algorithm.

For instance, $O(\log n)$-**APX** or log-**APX** is the class of problems that have an $O(\log n)$-approximation algorithm.

In the heirarchy of approximation algorithms, one prefers the so-called *approximation schemes* to the $O(f(n))$-approximation algorithms. For an instance x of Π and a real number $\epsilon > 0$, an algorithm $\mathbb{A}$ is called an approximation scheme for Π, if on input (x, ϵ) it outputs a solution y that satisfies:

- $c_{\mathbb{A}}(x) \leqslant (1+\epsilon) \cdot \mathsf{OPT}_{\Pi}(x)$ if Π is a minimization problem, and
- $c_{\mathbb{A}}(x) \geqslant (1-\epsilon) \cdot \mathsf{OPT}_{\Pi}(x)$ if Π is a maximization problem.

We say that $\mathbb{A}$ is a *polynomial time approximation scheme* (or PTAS) for Π if for every fixed $\epsilon > 0$, the running time of $\mathbb{A}$ is bounded by a polynomial in the size of x for every fixed ϵ (but can be a different for different values of ϵ). For instance, an algorithm that runs in $O(|x|^{\exp(\frac{1}{\epsilon})})$ would be considered PTAS. An *efficient polynomial-time approximation scheme* (EPTAS) is a PTAS with running time of the form $f(\frac{1}{\epsilon}) \cdot |x|^{O(1)}$. Finally, $\mathbb{A}$ is said to be a *fully polynomial time approximation scheme* if the running time of $\mathbb{A}$ is bounded by a polynomial in the size of x and $\frac{1}{\epsilon}$. That is, $\mathbb{A}$ is an FPTAS if it runs in time $(\frac{1}{\epsilon})^{O(1)} \cdot |x|^{O(1)}$. The term **PTAS** (**FPTAS**) is used to refer to the class of optimization problems that admit a PTAS (FPTAS). Assuming $\mathbf{P} \neq \mathbf{NP}$, we have the following heirarchy [97, pg. 20].

$$\textbf{FPTAS} \subsetneq \textbf{PTAS} \subsetneq \textbf{APX} \subsetneq \log\text{-}\textbf{APX}.$$

REDUCTIONS. An *approximation preserving reduction* is a polynomial time procedure for transforming an optimization problem Π_1 to an optimization problem Π_2, such that an α-approximation algorithm for Π_2 implies an $f(\alpha)$-approximation algorithm for Π_1, for some function f. Then, if Π_1 is hard to approximate within factor $f(\alpha)$, the reduction implies that Π_2 is hard to approximate within factor α.

A particularly well-studied class of approximation preserving reductions is given by the *L-reductions*, which provide an effective tool in proving hardness of approximability results [130, 158]. Consider an optimization problem Π_1 with a non-negative integer valued objective function c_{Π_1}. Given an instance x of Π_1, the goal is to find a solution y optimizing the objective function $c_{\Pi_1}(x, y)$. Define $\mathsf{OPT}_{\Pi_1}(x)$ as the optimal value of the objective function on input x.

Definition 2.2.1 (L-reduction)**.** An *L-reduction* (with parameters μ and ν) from a optimization problem Π_1 to another optimization problem Π_2 is a pair of polynomial time computable functions f and g, and fixed constants $\mu, \nu > 0$, satisfying the following conditions:

1. The function f maps instances of Π_1 to instances of Π_2.
2. For any instance x of Π_1, we have
$$\mathsf{OPT}_{\Pi_2}(f(x)) \leqslant \mu\, \mathsf{OPT}_{\Pi_1}(x).$$

3. The function g maps an instance x of Π_1 and a solution y of the corresponding instance $f(x)$ of Π_2 to a solution of x.

4. For any instance x of Π_1, and any solution y of $f(x)$, we have

$$|c_{\Pi_1}(x, g(f(x), y)) - \mathsf{OPT}_{\Pi_1}(x)| \leqslant \nu \cdot |(c_{\Pi_2}(f(x), y) - \mathsf{OPT}_{\Pi_2}(f(x)))|.$$

If $\mu = \nu = 1$, the reduction is *strict*.

From the definition of L-reduction, it is fairly straightforward to prove the following lemmas [158].

Theorem 2.2.1. *If there is an L-reduction with parameters μ and ν from a minimization problem Π_1 to another minimizaton problem Π_2, and there is a $(1+\delta)$-approximation algorithm for Π_2, then there is a $(1+\mu\nu\delta)$-approximation algorithm for Π_1.*

Theorem 2.2.2. *If there is an L-reduction with parameters μ and ν from a maximization problem Π_1 to another maximization problem Π_2, and there is a $(1-\delta)$-approximation algorithm for Π_2, then there is a $(1-\mu\nu\delta)$-approximation algorithm for Π_1.*

Next, we will define another important class of reductions called the PTAS reductions.

Definition 2.2.2 (PTAS-reduction)**.** Let Π_1 and Π_2 be two optimization problems. A *PTAS reduction* from Π_1 to Π_2 uses three polynomial-time computable functions, f, g, and α, that satisfy the following properties:

1. f maps instances of Π_1 to instances of Π_2.

2. g maps solutions of Π_2 to solutions of Π_1.

3. α maps the error parameter for Π_1 to the error parameter for Π_2.

4. Let x be an instance of Π_1 and ϵ be the error parameter for Π_1. If a solution y to $f(x)$ is within a factor of $1+\alpha(\epsilon)$ of the optimal solution of Π_2, then g maps y to a solution of Π_1 that is within a factor of $1+\epsilon$ of the optimal solution of Π_1.

It can be shown that an L-reduction is a PTAS reduction. PTAS reductions play an important role in the theory of approximability since they are used to define **APX**-completeness.

Apart from L-reductions and PTAS-reductions, there are many other approximation preserving reductions in use, for instance, A-reductions, P-reductions, AP-reductions, and so on. We will eschew further discussion on this topic. Instead we refer the reader to a survey article [47], which treats various approximation preserving reductions in detail, while also showing how they are related to each other.

2.3 PARAMETERIZED APPROXIMABILITY

Parameterized approximability is an extension of the notion of classical approximability. There are several competing alternatives for a suitable notion for parameterized approximability.

- One of the most common approaches for combining the two paradigms of parameterized and approximation algorithms is to devise parameterized approximation algorithms with cost of the solution as the parameter, and with approximation factor ρ as a function of the parameter. In this thesis, specifically in Paper IV, we refer to algorithms of this kind as *FPT approximation algorithms with approximation factor* ρ.

- Sometimes it is convenient to devise an α-approximation algorithm that runs in **FPT** time (for some suitable parameter), where α is a function of the size of the instance. In Paper V of this thesis, we refer to algorithms of this kind as *α-approximation FPT algorithms*.

- An EPTAS (see Section 2.2 for a definition) is a parameterized approximation algorithm with the quality of approximation as the parameter.

In this section, we provide a brief account of FPT approximation algorithms. For a full discussion on (variants of) parameterized algorithms we refer the reader to the recent survey articles [76, 118].

Informally, an FPT approximation algorithm is an algorithm whose running time is fixed parameter tractable for the parameter *cost of the solution* and whose approximation factor ρ is a function of the parameter (and independent of the input size). For instance, every polynomial time approximation algorithm with constant approximation factor is automatically an FPT approximation algorithm, but an approximation algorithm with approximation factor $\Theta(\log n)$, where n denotes the input size, is not an FPT approximation algorithm. Next, following [119], for standard parameterization of minimization problems, we provide definitions for FPT approximation algorithms and FPT cost approximation algorithms. Analogous definitions for maximization problems are also considered in [119].

Definition 2.3.1 (FPT approximation algorithm [119])**.** Let P be an **NP** minimization problem, and let $\rho : \mathbb{N} \to \mathbb{R}_{\geqslant 1}$ be a computable function such that $k \mapsto k \cdot \rho(k)$ is nondecreasing. An *FPT approximation algorithm* for P (over some alphabet Σ) with approximation ratio ρ is an algorithm $\mathbb{A}$ with the following properties:

1. For every input (x, k) whose optimal solution has cost at most k, $\mathbb{A}$ computes a solution for x of cost at most $k \cdot \rho(k)$. For inputs (x, k) without a solution of cost at most k, the output can be arbitrary.

2. The runtime of $\mathbb{A}$ on input (x, k) is $O(g(k) \cdot |x|^{O(1)})$ for some computable function g.

It is often convenient to work with a weaker notion of approximability where an algorithm is only required to compute the cost of an optimal solution rather than an actual optimal solution, and to work with decision rather than optimization problems. With that in mind, the notion of *FPT cost approximability* was introduced in [41].

Definition 2.3.2 (FPT cost approximation algorithm [119])**.** Let P be an **NP** minimization problem (over the alphabet Σ), and $\rho : \mathbb{N} \to \mathbb{R}_{\geqslant 1}$ a computable function. For an instance x of P, let $\min(x)$ denote the value of an optimal solution of x. Then, a decision algorithm $\mathbb{A}$ is an *FPT cost approximation algorithm* for P with approximation ratio ρ if

1. For feasible instances x of P and parameterized instances (x, k), $\mathbb{A}$ satisfies:
   a) If $k \geqslant \min(x) \cdot \rho(\min(x))$, then $\mathbb{A}$ accepts (x, k).
   b) If $k < \min(x)$, then $\mathbb{A}$ rejects (x, k).
2. $\mathbb{A}$ is an FPT algorithm. That is, there exists a computable function f with the property that for an input (x, k), the running time of $\mathbb{A}$ is bounded by $f(k) \cdot |x|^{O(1)}$.

It can be readily checked that FPT-approximability implies FPT cost approximability with the same approximation factor. Please refer to Section 3.1 of [41] for more details.

Theorem 2.3.1 (Chen et al. [41])**.** *Let P be an* **NP** *minimization problem over the alphabet Σ, and let $\rho : \mathbb{N} \to \mathbb{R}_{\geqslant 1}$ be a computable function such that $k \cdot \rho(k)$ is nondecreasing and unbounded. Suppose that P is FPT approximable with approximation ratio ρ. Then P is FPT cost approximable with approximation ratio ρ.*

An immediate consequence of the theorem above is that if P is not FPT cost approximable with approximation ratio ρ (under certain complexity theory assumptions), then P is not FPT approximable with approximation ratio ρ (under the same assumptions).

Gap problems and gap-preserving reductions were originally introduced in the context of proving the PCP theorem [8] – a cornerstone in the theory of approximation algorithms. These notions have natural analogues in the parameterized approximability setting. Below, we follow the definitions as provided by Eickmeyer et al. [71].

Definition 2.3.3 (gap instance of a parameterized problem [71])**.** Let P be a minimization problem, and P' its standard parameterization. Let $\delta : \mathbb{N} \to \mathbb{R}_{\geqslant 1}$ be a function. An instance (x, k) is a δ-gap instance of P' if either $\min(x) \leqslant k$ or $\min(x) \geqslant k \cdot \delta(k)$.

Definition 2.3.4 (gap-preserving FPT reduction [71])**.** Let $\alpha, \beta \colon \mathbb{N} \to \mathbb{R}_{\geqslant 1}$ be two computable functions, and let P and Q be two minimization problems. Let P' and Q' be the natural parameterizations of P and Q, respectively. We say that a reduction R from P' to Q' is a (α, β)*-gap-preserving FPT reduction* if

1. R is an FPT reduction from P' to Q' and
2. for every α-gap instance (x, k) of P', the instance $R(x, k)$ is a β-gap instance of Q'.

Gap-preserving FPT reductions are used to establish FPT-inapproximability.

3

TOPOLOGICAL PRELIMINARIES

3.1 SIMPLE HOMOTOPY THEORY

Classical homotopy theory studies "continuous deformations" of topological spaces and maps between them. We recall some of the rudimentary notions in homotopy theory. But, we will not discuss the fundamental group or the higher homotopy groups. For an in-depth treatment, we refer the reader to standard texts in algebraic topology [26, 88, 125, 151].

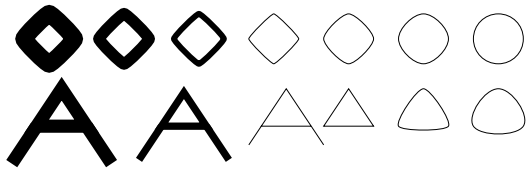
Definition 3.1.1 (Homotopy)**.** Let A and B be two topological spaces. We say that the continuous maps $f, g : A \to B$ are *homotopic* if there exists a continuous map $H : A \times [0,1] \to B$ with $H(x,0) = f(x)$ and $H(x,1) = g(x)$ for all $x \in X$. We call H a *homotopy* between f and g, and use the notation $f \simeq g$ to indicate that the maps f and g are homotopic.

Definition 3.1.2 (Homotopy equivalence, Contractible space)**.** Given a topological space X, let id_X denote the identity map on X. Two spaces A, B are said to be *homotopy equivalent*, if there exist continuous maps $f : A \to B$ and $g : B \to A$ such that $g \circ f \simeq \mathrm{id}_A$ and $f \circ g \simeq \mathrm{id}_B$. We use the notation $A \simeq B$ to indicate that A is homotopy equivalent to B. Each of the maps f and g is called a *homotopy equivalence*, and f and g are said to be *homotopy inverses* of each other. A topological space is said to be *contractible* if it is homotopy equivalent to a one point space.

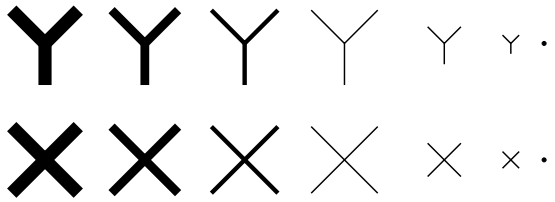
As the name suggests, homotopy equivalences indeed define an equivalence relation on topological spaces. See Figure 15 for some examples of homotopy equivalent spaces. Next, we look at a very specific kind of homotopy equivalence that is often encountered in practice.

Definition 3.1.3 ((Strong deformation) retraction)**.** Let B be a topological space, $A \subset B$ a subspace and $i : A \hookrightarrow B$ the inclusion map. Then, a continuous map $r : B \to A$ is called a *retraction* if r restricts to the identity map on A, that is, $r \circ i = r$. If $r : B \to A$ is a retraction, then A is called a *retract* of B. Additionally, if $i \circ r \simeq \mathrm{id}_B$, then r is called a *deformation retraction*. Let $F : B \times [0,1] \to B$ be a homotopy between $i \circ r$ and id_B. If $F(b,t) = b$ for all $t \in [0,1]$ and $b \in A$, then r is called a *strong deformation retracton*.

Motivated by the need to obtain a *combinatorial* description of homotopy theory, Whitehead introduced the notion of CW complexes, and developed simple homotopy theory [155, 156]. So, in short, simple homotopy theory is a refinement of homotopy theory, and hence can be formulated in terms of



(a) Homotopy equivalent to a circle



(b) Contractible space: homotopy equivalent to a one-point space

Figure 15: Examples of homotopy equivalent spaces

combinatorial moves on simplicial and CW complexes. Below, we introduce some basic notions in simple homotopy theory. For further details, we refer the reader to Cohen's book length treatment [44].

In this thesis, we limit ourselves in studying algorithmic problems on simplicial complexes. Yet we will provide a brief account of cellular expansions and collapses in Section 3.1.2, and discrete Morse theory on CW complexes in Section 3.3 so that we can connect the two arcs of the story in Section 3.4. But first we will look at simple homotopy equivalences in the simplicial world.

3.1.1 *Simplicial simple homotopy equivalence*

Given a simplicial complex K, a simplex $\sigma \in K$ is said to be *free* if σ has a unique cofacet τ in K.

Definition 3.1.4 (Elementary collapses and expansions)**.** Let K_0 be a simplicial complex, and let $\tau, \sigma \in K_0$ be an m-face and an $(m-1)$-face respectively such that $\sigma \subset \tau$, and σ is free in K_0.

We say that $K_1 = K_0 \setminus \{\tau, \sigma\}$ arises from K_0 by an *elementary collapse of dimension* m or *elementary* m*-collapse*, denoted by $K_0 \searrow K_1$. Its inverse, the operation $K_0 = K_1 \cup \{\tau, \sigma\}$ is called an *elementary expansion of dimension* m or *elementary* m*-expansion*, written $K_0 \nearrow K_1$. If the complex is implicit from the context, we denote elementary collapses by $\searrow^{\tau}_{\sigma}$ and elementary expansions by $\nearrow^{\tau}_{\sigma}$. An elementary collapse or an elementary expansion is sometimes referred to as an *elementary move*.

If there exists a sequence of elementary collapses turning a complex K_0 into K_1 we write $K_0 \searrow K_1$ and say that K_0 *collapses to* K_1. If K_1 is one-dimensional, we say that K_0 is *erasable*. If K_1 is a single point we call K_0 *collapsible*. Finally,

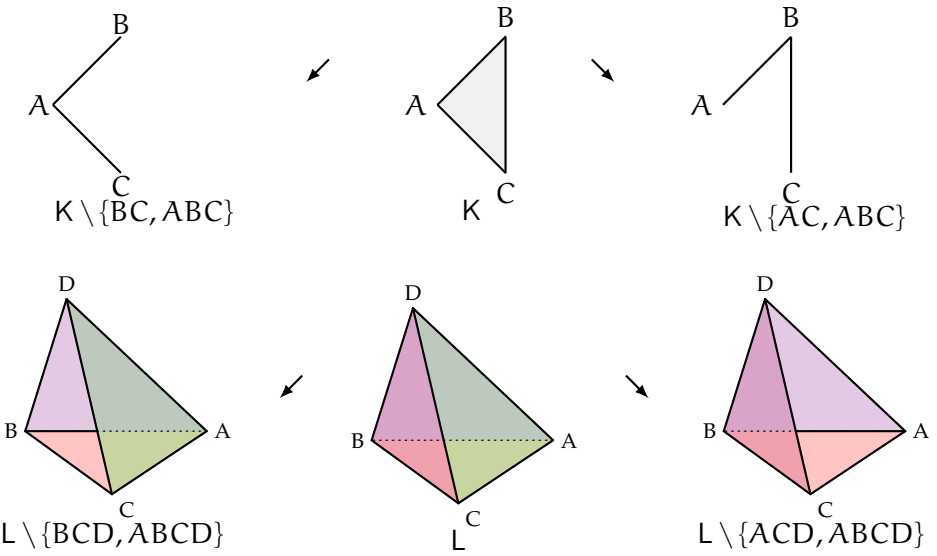


Figure 16: The top figure shows a 2-simplex collapsing to two different 2-horns. The bottom figure shows a 3-simplex collapsing to two different 3-horns.

we write $K_0 \nearrow K_1$ to indicate a sequence of expansions and say that K_0 *expands* to K_1.

It follows that an expansion $\nearrow^{\tau}_{\sigma}$ can only be performed in a simplicial complex K if all codimension 1 faces of τ except for σ are already in K. Hence, let τ be an m-face of a simplicial complex K, and let σ be one of its $(m-1)$-faces. An *(m-dimensional) horn* $\mathbf{H}(\tau, \sigma)$ associated to the pair (τ, σ) is the simplicial complex generated by the $(m-1)$-faces of τ apart from σ. Figure 16 shows examples of horns in dimensions 2 and 3.

Definition 3.1.5 (simple homotopy equivalence, simple homotopy graph)**.** Two simplicial complexes K and L are said to be *simple homotopy equivalent* or of coinciding *simple homotopy type*, written $K \nearrow\!\!\searrow L$, if there exists a sequence of elementary expansions and collapses turning one into the other.

The *dimension of a simple homotopy equivalence* is the maximum of the dimensions of K, L and of any elementary expansion or collapse in the sequence.

The *simple homotopy graph* is a graph whose nodes are simplicial complexes, and whose edges correspond to a pairs of complexes related by an elementary collapses. Naturally, its connected components are in one-to-one correspondence with simple homotopy types.

3.1.2 *Cellular simple homotopy equivalence*

Let I be a finite indexing set. *Let* X be a Hausdorff space, and let $\{e_i \mid i \in I\}$ be a finite family of open topological cells of varying dimensions. Sometimes we use the notation e_i^d to emphasize that the cell e_i is of dimension d. The collection of cells $\left\{e_i^d \mid i \in I, d \leqslant r\right\}$ is denoted by X_r. Then, we say that X is a finite CW complex if the following conditions are satisfied.

1. $X = \bigcup_{i \in I} e_i$ and $e_i \bigcap e_j = \emptyset$ when $i \neq j$.
2. For each cell e_i, there is a map $\varphi_i : \mathbb{B}^d \to X$ called the characteristic map of e_i, where $\mathbb{B}^d$ is a topological ball (that is homeomorphic to $\mathbb{I}^d = [0,1]^d$) of dimension d such that
   a) φ_i is a homemorphism from $\mathring{\mathbb{B}}^d$ onto e_i^d, where $\mathring{\mathbb{B}}^d$ denotes the interior of $\mathbb{B}^d$,
   b) $\varphi_i(\partial\mathbb{B}^d) \subset X_{d-1}$.

The map $\varphi|_{\partial\mathbb{B}^d}$ is called the *attaching map* for e_i^d. Let $J \subset I$. A *subcomplex* of a CW complex X is a subset Y along with a subfamily $\left\{e_j \mid j \in J\right\}$ of the cells of X such that $Y = \bigcup_{j \in J} e_j$ and the closure of each cell e_j, denoted by $\bar{e_j}$, is contained in Y. For example, for every r, the subset X_r, called the r-*dimensional skeleton* of X, is a subcomplex of X. A CW complex X is isomorphic to a CW complex Y if there exists a homeomorphism $h : X \to Y$ such that for every cell $e \in X$, $h(e)$ is a cell in Y.

Let X be a CW complex, and Y be a subcomplex of X. Then, we say that a CW complex X *collapses* to a subcomplex Y by an *elementary collapse* if the following conditions are satisfied.

1. There exists two cells e^p and f^{p-1} of X such that X can be expressed as follows:
$$X = Y \bigcup e^p \bigcup f^{p-1},$$
where e^p and f^{p-1} are not in Y.
2. There exists a pair of topological balls $(\mathbb{B}_1^p, \mathbb{B}_2^{p-1})$ along with a map $\varphi : \mathbb{B}_1^p \to X$ such that
   a) φ is a characteristic map for e^p.
   b) $\varphi|_{\mathbb{B}_2^{p-1}}$ is a characteristic map for f^{p-1}.
   c) $\varphi(\overline{\mathbb{B}}_3^{p-1}) \subset X^{p-1}$, where $\overline{\mathbb{B}}_3^{p-1}$ is the topological closure of the $(p-1)$-ball $\partial\mathbb{B}_1^p \setminus \mathbb{B}_2^{p-1}$.

It is easy to check that if $X \searrow Y$, then X is isomorphic to the adjunction space $Y \bigcup_{\psi} \mathbb{I}^p$, which is the mapping cylinder of the map $\psi : \mathbb{I}^{p-1} \to Y$. Hence, there is a strong deformation retraction from X to Y.

We say that X *collapses* to Y (denoted by, $X \searrow Y$) if there exists a finite sequence of elementary collapses

$$X = X_0 \searrow X_1 \searrow \cdots \searrow X_q = Y.$$

Equivalently, we say that Y *expands* to X (and denote it as $Y \nearrow X$).

We say that there is a *deformation* from X to Y (written as $X \nearrow\!\searrow Y$) if there exists a finite sequence of operations, each of which is either an elementary expansion or an elementary collapse. In this case, we say that X and Y have the same *simple homotopy type*.

Suppose that $X = X_0, X_1, \ldots, X_p = Y$ are the complexes that arise in the formal deformation from X to Y. In this sequence, for every $i \in [1, p]$, if $X_{i-1} \searrow X_i$, then the inclusion map $X_{i-1} \hookrightarrow X_i$ is a homotopy equivalence, and if $X_{i-1} \nearrow X_i$, then there is a cellular strong deformation retraction from X_{i-1} to X_i. Furthermore, we say that a complex X n-*deforms* to Y (written as $X \overset{n}{\nearrow\!\searrow} Y$) if every complex X_i, $i \in [0, p]$ is at most n-dimensional,

Clearly, $Y \nearrow\!\searrow X$ if and only if $X \nearrow\!\searrow Y$. In particular, simple homotopy equivalences define an equivalence relation on CW complexes.

Simplicial homotopy equivalences and cellular homotopy equivalences are related by the following two theorems.

Theorem 3.1.1 (mathoverflow [2], Rourke-Sanderson Appendix B5 [146])**.** *If* K *and* L *are simplicial complexes, then a map* f *is a simplicial simple homotopy equivalence between* $|K|$ *and* $|L|$ *if and only if it is a cellular simple homotopy equivalence between* K *and* L *with respect to the obvious CW structures on* $|K|$ *and* $|L|$*.*

Note that Theorem 3.1.1 justifies the omission of the adjective "simplicial" in Definition 3.1.5.

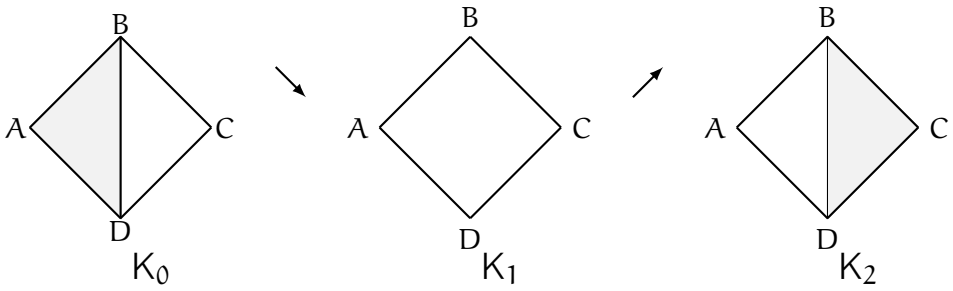
Theorem 3.1.2 (Cohen Theorem 7.2 [44])**.** *Every finite CW complex* X *has the simple homotopy type of a finite simplicial complex of the same dimension.*

Notwithstanding that all the important conjectures in simple homotopy theory are stated in the CW category, Theorems 3.1.1 and 3.1.2 justify our choice of limiting the algorithmic considerations in Paper II to simplicial collapses and expansions.

3.1.3 *Simple homotopy theory and combinatorial group theory*

We know that a simple homotopy equivalence is a homotopy equivalence, but is the converse true? In this regard, it is natural to make the following two conjectures in simple homotopy theory.

CONJECTURE 1 If there exists a homotopy equivalence from X to Y, then there exists a simple-homotopy equivalence from X to Y.



The figure shows a simple homotopy equivalence $K_0 \nearrow\!\searrow K_2$.

CONJECTURE II If $f : X \to Y$ is a homotopy equivalence, then f is a simple-homotopy equivalence.

The obstruction to a homotopy equivalence $f : X \to Y$ of finite CW-complexes being a simple homotopy equivalence is the Whitehead torsion $\tau(f)$, which is an element in the Whitehead group $\mathrm{Wh}(\pi_1(Y))$. In the special case when $\pi_1(Y)$ is trivial or $\mathbb{Z}$, both conjectures are true. In general, CONJECTURE I and CONJECTURE II are false. For a full discussion on the Whitehead torsion, we refer the reader to [44, 123].

The best known examples of homotopy equivalent complexes that are not simple homotopy equivalent come from lens spaces [44, 124], which we will briefly describe.

LENS SPACES. The three-dimensional lens spaces $L(p, q)$ were introduced by Tietze in in 1908 [124]. The three-dimensional lens spaces are formed by taking quotients of S^3 by $\mathbb{Z}_p$ actions. Formally,

Definition 3.1.6 (Lens spaces)**.** Let p and q be coprime integers. Then, the lens space $L(p, q)$ is defined as the quotient of S^3 by a certain free action of $\mathbb{Z}_p$ that depends on q:

$$L(p, q) = \frac{\left\{(z_1, z_2) \in \mathbb{C}^2 \mid |z_1|^2 + |z_2|^2 = 1\right\}}{(z_1, z_2) \sim (\xi \cdot z_1, \xi^q \cdot z_2)},$$

where $\xi = e^{\frac{2\pi i}{p}}$ is a root of unity.

It is easy to check that $L(2, 1) = \mathbb{R}P^3$ is the real projective 3-space. For $q_1 \neq q_2$, $L(p, q_1)$ and $L(p, q_2)$ have the same homology groups, and the same homotopy groups. However, there are finer topological aspects of lens spaces that may depend on the choice of q. In particular, $L(5, 1)$ and $L(5, 2)$ have the same homotopy groups, but they are not homotopy equivalent, whereas $L(7, 1)$ and $L(7, 2)$ are homotopy equivalent, but not simple homotopy equivalent.

For complexes that are simple homotopy equivalent , C.T.C. Wall proved the following result.

Theorem 3.1.3 (Wall [150], Matveev [120, Theorem 1.3.1]). *Let* K *and* L *be two simplicial (or CW) complexes of the same simple homotopy type and of dimension at most* $m > 2$*. Then, there exists a simple homotopy equivalence of dimension at most* $m+1$*, taking one to the other.*

For the case $m = 2$, Theorem 3.1.3 is still open and known as the (topological) Andrews–Curtis conjecture [10, 120, 127]. On the other hand, it is known that any contractible 2-complex is also simple homotopy equivalent to a point [156]. Hence, any pair of contractible 2-complexes can be connected by a simple homotopy equivalence of dimension at most four – but determining whether we can always decide if such a simple homotopy equivalence exists is an open question [17], equivalent to the triviality problem for balanced group presentations [92]. In what follows, we give a brief description of the Andrews–Curtis and related conjectures.

A *presentation of a group* is given by a pairs of sets $\langle X \mid R \rangle$, such that if we set $F = F(X)$ to be the free group on the set of *generators* $X = \{x_1, x_2 \dots x_n\}$, and if we let $N = N(R)$ be the normal closure of the set of *relators* $R = \{r_1, r_2 \dots r_m\}$, then there is an isomorphism from F/N to G. A presentation of a group is called *balanced* if it has the same number of generators and relators.

Given a balanced presentation of a group, the set of moves given by

(AC-1) replace r_i by $r_i r_j$ for some $j \neq i$

(AC-2) replace r_i by r_i^{-1}, and

(AC-3) replace r_i by $g r_i g^{-1}$ for some word $g \in F$,

is called *AC-moves* (or Andrews-Curtis moves). We say that a balanced presentation $\mathcal{P}$ of a group is *AC-trivializable* if $\mathcal{P}$ can be reduced to the trivial presentation (X, X) by applying a series of AC-moves.

ANDREWS-CURTIS CONJECTURE. If $\mathcal{P}$ is a balanced presentation of a trivial group, then $\mathcal{P}$ can be transformed into a trivial presentation by a sequence of AC-moves.

A *weaker* form of the *Andrews-Curtis conjecture* allows two additional moves for transforming presentations to trivial presentations:

(AC-4) Introduce a new generator x' and the corresponding relator x',

(AC-5) The converse of the operation above.

The moves AC-1 to AC-5 are called *stable AC moves*. We say that two presentations $\mathcal{P}$ and $\mathcal{P}'$ are *(stably) AC-equivalent* if $\mathcal{P}$ can be transformed to $\mathcal{P}'$ by a series of (stable) AC-moves. We use the notation $\mathcal{P} \sim \mathcal{P}'$ to indicate that $\mathcal{P}$ is stably AC-equivalent to $\mathcal{P}'$.

Given a connected CW-complex K, a presentation $\mathcal{P}_K$ for its fundamental group can be read off using a procedure described in [92, Section 1.3]:

1. Choose a vertex v of K as the basepoint.

2. Let T be a spanning tree in the 1-skeleton K^1, and let $e_1, e_2 \dots, e_n$ be the edges in $K^1 \setminus T$.

3. Orient every edge $e_i \in K^1 \setminus T$, $i \in [n]$. Every edge e_i determines a unique closed path x_i that starts and ends at v: from v to the source of e_i, followed by e_i itself, followed by target of e_i to v. Let $X = \{x_1, x_2 \dots x_n\}$. Then, the closed paths in X are the generators of $\pi_1(|K^1|, v)$, and $\pi_1(|K^1|, v) = F(X)$, that is, the free group over the set X.

4. Let $f_1, f_2 \dots f_m$ be the 2-cells of K. For every j, let ϕ_j be the attaching map of f_j.

5. Choose b_j in the image of ϕ_j, and let λ_j be the path in K^1 from b_j to v. Then, $r_j = \overline{\lambda_j} * \phi_j * \lambda_j$ defines as a word in $\pi_1(|K^1|, v)$.

6. Let $R = \{r_1, r_2 \dots r_m\}$.

7. The associated presentation $\mathcal{P}_K$ is given by $\mathcal{P}_K = \langle X \mid R \rangle$.

If $\mathcal{P}_K$ and $\mathcal{P}'_K$ are obtained from the procedure above with different choices of spanning trees, base points and attaching maps, then $\mathcal{P}_K \sim \mathcal{P}'_K$. We refer the reader to the PhD thesis by Fernandez [79, Proposition 1.4.4] for a proof.

On the other hand, an inverse of the procedure is also described in [92, Page 9]. This inverse procedure is easy to describe. Given a presentation $\mathcal{P} = \langle x_1, x_2 \dots x_n \mid r_1, r_2 \dots r_m \rangle$, one defines the complex $K^{\mathcal{P}}$ as follows.

1. The 0-skeleton $K_0^{\mathcal{P}}$ of $K^{\mathcal{P}}$ consists of a single vertex.

2. The 1-skeleton $K_1^{\mathcal{P}}$ of $K^{\mathcal{P}}$ consists of a bouquet of circles with one oriented edge e_i for each generator x_i.

3. The 2-cells f_j are in one-to-one correspondence with relators r_j. Subdividing the boundaries ∂D_j^2 of the m 2-dimensional disks D_j^2 into polygons, the attaching map $\phi_j : \partial D_j^2 \to K_1^{\mathcal{P}}$ takes each edge (of the polygon) to an edge of $K_1^{\mathcal{P}}$ or onto the vertex.

Example 3.1.1. It is easy to check that

- the complex $K_{\mathcal{P}}$ for the presentation $\mathcal{P} = \langle x \mid x^2 x^{-1} \rangle$ is a dunce hat,
- the complex $K_{\mathcal{P}'}$ for the presentation $\mathcal{P}' = \langle x_1, x_2 \mid x_1 x_2 x_1^{-1} x_2^{-1} \rangle$ is a torus, and
- the complex $K_{\hat{\mathcal{P}}}$ for the presentation $\hat{\mathcal{P}} = \langle x_1, x_2 \mid x_1 x_2 x_1^{-1} x_2 \rangle$ is a Klein bottle.

Strikingly, AC-equivalences between presentations give rise to simple homotopy equivalences between complexes [159]. What is even more interesting is that the expansions and collapses coming from AC-transformations strictly involve cells of dimension at most 3. Thus, the weak Andrews-Curtis conjecture is equivalent to the topological Andrews-Curtis conjecture stated below.

TOPOLOGICAL ANDREWS-CURTIS CONJECTURE. If K is a finite contractible CW complex of dimension 2, then K can be 3-deformed to a point.

The algorithmic challenge when it comes to the Andrews-Curtis conjecture is that it is not even clear if the triviality problem of balanced presentations is decidable. Some of the recent algorithmic work in context of the Andrews-Curtis concerns with refuting potential counterexamples to the Andrews-Curtis conjecture [25, 89, 122]. More recently, Fernandez used a discrete Morse theory based approach to refute some of the suggested potential counterexamples to the Andrews-Curtis conjecture [78].

One interesting complexity result in this context is by Bridson [27] who showed that there exists a family of balanced presentations for which the number of AC-moves required for trivialization grows faster than a tower of exponentials. As a counterpoint, Bridson notes that for this family of presentations despite the size of required AC-moves, there exists a cubic time algorithm [27, 7.3] that can determine triviality.

The (topological) Andrews-Curtis conjecture is related to a slew of other important conjectures in topology. We will look at at two of the most prominent ones, namely, the Zeeman conjecture and the Poincaré conjecture.

ZEEMAN CONJECTURE. Given a finite contractible 2-dimensional CW complex K, the complex $K \times I$ is collapsible, where I here denotes the interval $[0, 1]$.

Since $K \times I$ is a 3-complex, Zeeman conjecture (which is also still open) implies the Andrews-Curtis conjecture.

Next, we look at the Poincaré conjecture which says that if M is a compact 3-manifold without boundary that is homotopy equivalent to S^3, then M is homeomorphic to S^3. An equivalent formulation of the Poincaré conjecture is as follows.

POINCARÉ CONJECTURE. If L is a compact, contractible 3-manifold, then L is homeomorphic to the 3-ball.

In what was one of the most striking breakthroughs in mathematics, at the turn of the millennium, Perelman proved the Poincaré conjecture [132, 133, 134].

The Zeeman conjecture implies the Poincaré conjecture by the following (sketch of an) argument. Let L be a compact, contractible 3-manifold. Then, L collapses to a 2-complex K which is also contractible. By Zeeman's conjecture $K \times I$ is collapsible. Hence, $L \times I$ is collapsible. Since $L \times I$ is a 4-manifold that collapses to a point, $L \times I$ must be a 4-ball. The boundary of $L \times I$ is

homeomorphic to a 3-sphere. So, by construction, L is a compact, contractible 3-manifold that embeds in a 3-sphere. Therefore, L is homeomorphic to a 3-ball.

3.2 SIMPLICIAL HOMOLOGY AND COHOMOLOGY

In this section, we provide a brief introduction to simplicial homology with coefficients in $\mathbb{Z}_2$. For a general introduction to algebraic topology, we refer the reader to some standard texts [26, 88, 125, 151].

Let K be a simplicial complex. We consider the formal sums of simplices with coefficients in $\mathbb{Z}_2$, that is, sums of the form $\sum_{\sigma \in K^{(p)}} a_\sigma \sigma$, where each $a_\sigma \in \{0, 1\}$. The expression $\sum_{\sigma \in K^{(p)}} a_\sigma \sigma$ is called a p-*chain*. Since chains can be added to each other, they form an Abelian group, denoted by $C_p(K)$. Since we consider formal sums with coefficients coming from $\mathbb{Z}_2$, which is a field, $C_p(K)$, in this case, is a vector space of dimension n_p over $\mathbb{Z}_2$. The p-simplices in K form a (natural) basis for $C_p(K)$. This establishes a natural one-to-one correspondence between elements of $C_p(K)$ and subsets of $K^{(p)}$, and we will freely make use of this identification. The *boundary* of a p-simplex is a $(p-1)$-chain that corresponds to the set of its $(p-1)$-faces. This map can be linearly extended from p-simplices to p-chains, where the boundary of a chain is the $\mathbb{Z}_2$-sum of the boundaries of its elements. The resulting *boundary homomorphism* is denoted by $\partial_p : C_p(K) \to C_{p-1}(K)$. A chain $\zeta \in C_p(K)$ is called a p-*cycle* if $\partial_p \zeta = 0$, that is, $\zeta \in \ker \partial_p$. The group of p-dimensional cycles is denoted by $Z_p(K)$. As before, since we are working with $\mathbb{Z}_2$ coefficients, $Z_p(K)$ is a vector space over $\mathbb{Z}_2$. A chain $\eta \in C_p(K)$ is said to be a p-*boundary* if $\eta = \partial_{p+1} c$ for some chain $c \in C_{p+1}(K)$, that is, $\eta \in \mathrm{im}\partial_{p+1}$. The vector space of p-dimensional boundaries is denoted by $B_p(K)$.

In our case, $B_p(K)$ is also a vector space, and in fact a subspace of $C_p(K)$. Thus, we can consider the quotient space $H_p(K) = Z_p(K)/B_p(K)$. The elements of the vector space $H_p(K)$, known as the p-th *homology* of K, are equivalence classes of p-cycles, called *homology classes* where p-cycles are said to be *homologous* if their $\mathbb{Z}_2$-difference is a p-boundary. For a p-cycle ζ, its corresponding homology class is denoted by $[\zeta]$. Bases of $B_p(K)$, $Z_p(K)$ and $H_p(K)$ are called *boundary bases*, *cycle bases*, and *homology bases*, respectively. The dimension of the p-th homology of K is called the p-th *Betti number* of K, denoted by $\beta_p(K)$.

Using the natural bases for $C_p(K)$ and $C_{p-1}(K)$, the matrix $[\partial_p \sigma_1 \; \partial_p \sigma_2 \cdots \partial_p \sigma_{n_p}]$ whose column vectors are boundaries of p-simplices is called the p-*th boundary matrix*. Abusing notation, we also denote the p-th boundary matrix by ∂_p.

The dual vector space of $C_p(K)$ (the vector space of linear maps $C_p(K) \to \mathbb{Z}_2$) is called the space of *cochain*, which we denote by $C^p(K) = \mathrm{Hom}(C_p(K), \mathbb{Z}_2)$. Again, there is a natural basis corresponding to the p-simplices of K, with a p-simplex σ corresponding to the linear map η with values $\eta(\sigma) = 1$ and $\eta(\rho) = 0$ for every other p-simplex $\rho \neq \sigma$. The adjoint map to the boundary

map $\partial_{p+1}\colon C_{p+1}(K) \to C_p(K)$ is the *coboundary map* $\delta_p\colon C^p(K) \to C^{p+1}(K)$. Similarly to chains and boundary maps, we may define subspaces of *cocycles* $Z^p(K) = \ker \delta_p$ and *coboundaries* $B^p(K) = \mathrm{im}\delta_{p+1} \subseteq Z^p(K)$, and form their quotient $H^p(K) = Z^p(K)/B^p(K)$, which is the *cohomology* of K. Again, for a p-cocycle η, the corresponding cohomology class is denoted by $[\eta]$. The natural pairing of chains and cochains $C_p(K) \times C^p(K) \to \mathbb{Z}_2, (\zeta, \eta) \mapsto \eta(\zeta)$ induces a well-defined isomorphism $H^p(K) \times H_p(K) \to \mathbb{Z}_2, ([\zeta], [\eta]) \mapsto \eta(\zeta)$, identifying cohomology as the vector space dual to homology up to a natural isomorphism.

A set of p-cycles $\{\zeta_1, \ldots, \zeta_g\}$ is called a *homology cycle basis* if the set of classes $\{[\zeta_1], \ldots, [\zeta_g]\}$ forms a homology basis. For brevity, we abuse notation by using the term *(p-th) homology basis* for $\{\zeta_1, \ldots, \zeta_g\}$. Similarly, a set of p-cocycles $\{\eta_1, \ldots, \eta_g\}$ is called a *(p-th) cohomology cocycle basis* if the set of classes $\{[\eta_1], \ldots, [\eta_g]\}$ forms a cohomology basis.

Assigning non-negative weights to the edges of K, the weight of a cycle is the sum of the weights of its edges, and the weight of a homology basis is the sum of the weights of the basis elements. We call the problem of computing a minimum weight basis of $H_1(K)$ the *minimum homology basis* problem. Similarly, we call the problem of computing a minimum weight basis of $H^1(K)$, the *minimum cohomology basis* problem.

3.3 DISCRETE MORSE THEORY AND ERASABILITY

In this section, we provide a brief description of Forman's discrete Morse theory on CW complexes. For a comprehensive introduction, we refer the reader to [83].

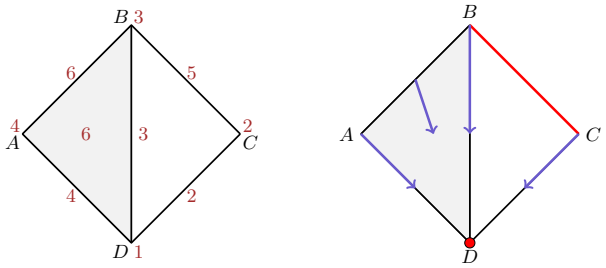
Notation 3. Let σ and τ be two cells that belongs to a CW complex K. We write $\tau \succ \sigma$ (or $\sigma \prec \tau$) if $\sigma \neq \tau$ and $\sigma \in \bar{\tau}$, where $\bar{\tau}$ is the closure of τ. In this case, we say that σ is a face of τ.

Definition 3.3.1 (Discrete Morse function, critical and regular cells)**.** A real-valued function f on a CW complex K is called a *discrete Morse function* if

- f is *monotonic*, i.e., $\sigma \prec \tau$ implies $f(\sigma) \leqslant f(\tau)$, and
- for all $t \in \mathrm{im}(f)$, the preimage $f^{-1}(t)$ is
  - either a singleton $\{\sigma\}$ (in which case σ is a *critical cell*)
  - or a pair $\{\sigma, \tau\}$, where σ is a facet of τ (in which case (σ, τ) form a *gradient pair* and σ and τ are *regular cells*).

Definition 3.3.2 (Discrete gradient vector field)**.** Given a discrete Morse function f on a CW complex K, the *discrete gradient vector field* $\mathcal{V}$ of f is the collection of pairs of cells (σ, τ), where (σ, τ) is in $\mathcal{V}$ if and only if σ is a facet of τ and $f(\sigma) = f(\tau)$.

Please see Figure 17 for an example.



(a) Morse function f (b) Vector field $\mathcal{V}$

Figure 17: The figure on the left shows a complex K with a discrete Morse function f on it. The figure on the right shows a discrete gradient $\mathcal{V}$ associated to the function f. Since K has the homotopy type of a circle, $\mathcal{V}$ is an optimal gradient.

Definition 3.3.3 (Sublevel complex)**.** Given a discrete Morse function f on a CW complex K, and a real number $c \in \mathbb{R}$, the *sublevel complex* $K_{\leqslant c}$ is defined as

$$K_{\leqslant c} = \{e \in K \mid e \in f^{-1}((-\infty, c])\}.$$

The next theorem captures the true spirit of Morse theory in the discrete realm.

Theorem 3.3.1 (Forman [83], Theorems 3.3 and 3.4)**.** *Let f be a discrete Morse function associated with a gradient vector field $\mathcal{V}$ on a regular CW complex K.*

1. *If the interval $[a, b]$ contains no critical cells, then $K_{\leqslant b} \searrow K_{\leqslant a}$. Moreover, $K_{\leqslant b} \setminus K_{\leqslant a}$ is a union of pairs in $\mathcal{V}$.*
2. *If σ^p is the unique critical cell in $f^{-1}([a, b])$, then $K_{\leqslant b}$ is homotopy equivalent to the space $K_{\leqslant a} \bigcup_\varphi \mathbb{B}^p$, where $\varphi : \partial\mathbb{B}^p \to K_{\leqslant a}$ is an attaching map for a cell of dimension p.*

From the first statement of Theorem 3.3.1, we learn that a cellular collapse can be encoded by a discrete gradient. In particular, if K is a CW complex with a vector field $\mathcal{V}$, and $L \subseteq K$ is a subcomplex such that $K \searrow L$, then $K \setminus L$ is a union of pairs in $\mathcal{V}$. In this case, we say that the collapse $K \searrow L$ is *induced by* the gradient $\mathcal{V}$.

The following theorem highlights the fact that discrete Morse theory leads to a condensed representation of topological datum, thereby motivating the algorithmic search for (near-)optimal Morse matchings.

Theorem 3.3.2 (Forman [83], Theorem 10.2)**.** *Let K be a CW complex with a discrete gradient vector field $\mathcal{V}$ and let m_d denote the number of critical cells of $\mathcal{V}$ of dimension d. Then K is homotopy equivalent to a CW complex $M_\mathcal{V}$ with exactly m_d cells of dimension d. We call $M_\mathcal{V}$ the* Morse complex *associated to gradient $\mathcal{V}$.*

One also obtains a (small) chain complex from Forman's theory. In particular, a discrete gradient vector field on a simplicial complex X with m_d critical cells of dimension d gives rise to a chain complex having dimension m_d in each degree d, whose homology is isomorphic to that of X. This is the content of Theorems 7.1 and 7.3 in the expository paper by Forman [84].

Discrete Morse theory on simplicial complexes

We now recall a few notions and results that are particularly useful when applying discrete Morse theory on complexes that are simplicial.

MORSE MATCHINGS. Discrete gradient vector fields have a useful interpretation in terms of acyclic graphs obtained from matchings on Hasse diagrams, due to Chari [37]. Let K be a simplicial complex, let $\mathcal{H}_K$ be its Hasse diagram, and let M be a matching in the underlying undirected graph $\mathcal{H}_K$. Let $\mathcal{H}_K(M)$ be the directed graph obtained from $\mathcal{H}_K$ by reversing the direction of each edge of the matching M. Then M is a *Morse matching* if and only if $\mathcal{H}_K(M)$ is a directed acyclic graph. Every Morse matching M on the Hasse diagram $\mathcal{H}_K$ corresponds to a unique gradient vector field $\mathcal{V}_M$ on complex K and vice versa. For a Morse matching M, the unmatched vertices correspond to critical simplices of $\mathcal{V}_M$, and the matched vertices correspond to the regular simplices of $\mathcal{V}_M$.

ERASABILITY. Following the terminology used in [13, 70], we make the following definitions: A maximal face τ in a simplicial complex K is called an *internal simplex* if it has no free face. If a 2-complex K collapses to a 1-complex, we say that K is *erasable*. Moreover, for a 2-complex K, the quantity $\mathrm{er}(\mathcal{K})$ is the minimum number of internal 2-simplices that need to be removed so that the resulting complex collapses to a 1-complex. Equivalently, it is the minimum number of critical 2-simplices of any discrete gradient on K. Furthermore, we say that a subcomplex $L \subseteq K$ is an *erasable subcomplex of* K *(through the gradient* $\mathcal{V}$*)* if there exists another subcomplex $C \subseteq K$ with $K \searrow C$ (induced by the gradient $\mathcal{V}$) such that the set of 2-dimensional simplices of these complexes satisfy the following relation: $L^{(2)} \subseteq K^{(2)} \setminus C^{(2)}$. We call such a gradient $\mathcal{V}$ an *erasing gradient*. Finally, we say that a simplex σ in a complex K is *eventually free (through the gradient* $\mathcal{V}$*)* if there exists a subcomplex L of K such that $K \searrow L$ (induced by $\mathcal{V}$) and σ is free in L. Equivalently, K collapses further to a subcomplex not containing σ.

3.4 DISCRETE MORSE THEORY AND SIMPLE HOMOTOPY THEORY

In Section 3.3, we learnt from Theorem 3.3.2 that for a simplicial complex K equipped with a gradient vector field $\mathcal{V}$, the Morse complex $M_{\mathcal{V}}$ constructed

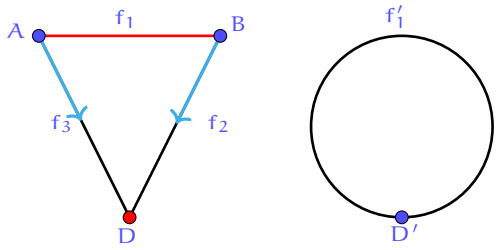


Figure 18: The figure on left shows a 1-dimensional simplicial complex K equipped with a gradient vector field $\mathcal{V} = \{(A, f_3), (B, f_2)\}$. The figure on the right shows the Morse complex $M_{\mathcal{V}}$ that is homotopy equivalent to K.

using Forman's theory is homotopy equivalent to K. Figure 18 shows an example. This raises a natural question: Is K also simple homotopy equivalent to $M_{\mathcal{V}}$? Päkkilä [129] and Fernandez [79] have independently answered this question in the affirmative.

Here, we follow the exposition by Fernandez [79]. But before we do that, we will try to understand the idea through the example in Figure 18. First, construct a 1-complex L_1 which is formed by taking the disjoint union of K and $M_{\mathcal{V}}$ and then identifying the point D to D'. Then, we use L_1 as a 1-skeleton for building a regular 2-complex L with a single 2-cell e^2. The boundary of e^2 is the whole of L_1. L looks like a cylinder (with the triangle K as one base and the circle $M_{\mathcal{V}}$ as the other base) squished in order to identify D to D'. We will not illustrate the L here, but we will instead rely on the reader's imagination. Then, $K \nearrow L$ via the pair (f_1', e^2), and $L \searrow M_{\mathcal{V}}$ via the pairs (f_1, e^2) followed by (A, f_3) and (B, f_2). Therefore, we have that $K \swarrow\!\!\!\searrow M_{\mathcal{V}}$. As to how the cell e^2 is obtained, is explained by Lemma 3.4.1 from Fernandez [79, Proposition 6.4.8].

Lemma 3.4.1 (Fernandez [79])**.** *Let X be a CW-complex of dimension at most p, and let $\varphi : \partial\mathbb{B}^p \to X$ be the attaching map of a p-cell e^p.*

If $X \searrow Y$, then $X \cup e^p \overset{p+1}{\swarrow\!\!\!\searrow} Y \cup \tilde{e}^p$, where the attaching map $\tilde{\varphi} : \partial\mathbb{B}^p \to X$ of $\tilde{e}^p$ is defined as $\tilde{\varphi} = r\varphi$ where $r : X \to Y$ is a strong deformation retract induced by the collapse $X \searrow Y$.

Proof. Let $\iota : Y \to X$ be the inclusion map and let $r : X \to Y$ be a strong deformation retract induced by the collapse $X \searrow Y$. Since $\iota r \cong \mathrm{id}_X$, composing by φ on both sides gives a homotopy $\iota r\varphi \cong_H \varphi$ with $H : \partial\mathbb{B}^p \times I \to X$ that allows us to perform the following sequence of elementary moves

$$X \cup_\varphi \mathbb{B}^p \overset{e}{\nearrow} (X \cup_\varphi \mathbb{B}^p) \cup_{\iota r\varphi} \mathbb{B}^p \cup_H \mathbb{B}^p \times I \overset{e}{\searrow} X \cup_{\iota r\varphi} \mathbb{B}^p$$

that comes from treating H as an attaching map. Finally, the collapse $X \searrow Y$ induces a collapse $X \cup_{\iota r \varphi} \mathbb{B}^p \searrow Y \cup_{r\varphi} \mathbb{B}^p$ since the image of the attaching map $\iota r \varphi$ is included in Y.

Since the dimension of the complex $(X \cup_\varphi \mathbb{B}^p) \cup_{\iota r \varphi} \mathbb{B}^p \cup_H \mathbb{B}^{p+1}$ is $p+1$, we conclude that

$$X \cup_\varphi \mathbb{B}^p \overset{p+1}{\nearrow\!\searrow} Y \cup_{\tilde{\varphi}} \mathbb{B}^p. \qquad \square$$

Through an inductive application of Lemma 3.4.1 along with some careful bookkeeping of attaching maps, Fernandez [79, Theorem 6.4.17] proved the following refinement of Theorem 3.3.2.

Theorem 3.4.2 (Fernandez [79]). *Let X be a regular CW complex of dimension n equipped with a gradient vector field $\mathcal{V}$. Then, $X \overset{n+1}{\nearrow\!\searrow} M_{\mathcal{V}}$, where $M_{\mathcal{V}}$ is the Morse complex associated to the gradient $\mathcal{V}$.*

BIBLIOGRAPHY

[1] Parameterized complexity and practical computing workshop. https://www.uib.no/en/rg/algo/127995/parameterized-complexity-and-practical-computing-workshop. Accessed: 2021-09-01.

[2] Simplicial simple homotopy vs. cellular simple homotopy. https://mathoverflow.net/questions/355891/simplicial-simple-homotopy-vs-cellular-simple-homotopy. Accessed: 2021-10-10.

[3] Henry Adams, Tegan Emerson, Michael Kirby, Rachel Neville, Chris Peterson, Patrick Shipman, Sofya Chepushtanova, Eric Hanson, Francis Motta, and Lori Ziegelmeier. Persistence images: A stable vector representation of persistent homology. *Journal of Machine Learning Research*, 18, 2017.

[4] Pankaj K Agarwal, Kyle Fox, Abhinandan Nath, Anastasios Sidiropoulos, and Yusu Wang. Computing the gromov-hausdorff distance for metric trees. *ACM Transactions on Algorithms (TALG)*, 14(2):1–20, 2018.

[5] Edoardo Amaldi, Claudio Iuliano, Tomasz Jurkiewicz, Kurt Mehlhorn, and Romeo Rizzi. Breaking the $O(m^2n)$ barrier for minimum cycle bases. In Amos Fiat and Peter Sanders, editors, *Algorithms - ESA 2009*, pages 301–312, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

[6] Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. *SIAM Journal on Computing*, (0):FOCS20–1, 2021.

[7] Nima Anari, Kuikui Liu, Shayan Oveis Gharan, and Cynthia Vinzant. Log-concave polynomials ii: high-dimensional walks and an fpras for counting bases of a matroid. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 1–12, 2019.

[8] Sanjeev Arora. *Probabilistic Checking of Proofs and Hardness of Approximation Problems*. PhD thesis, U C Berkeley, Berkeley, CA, USA, 1995. UMI Order No. GAX95-30468.

[9] Sanjeev Arora and Boaz Barak. Complexity theory: A modern approach, 2009.

[10] Jonathan A. Barmak. *Algebraic topology of finite topological spaces and applications*, volume 2032 of *Lecture Notes in Mathematics*. Springer,

Heidelberg, 2011. URL: https://doi.org/10.1007/978-3-642-22003-6, doi:10.1007/978-3-642-22003-6.

[11] Ulrich Bauer. Ripser: efficient computation of vietoris–rips persistence barcodes. *Journal of Applied and Computational Topology*, 2021. doi:10.1007/s41468-021-00071-5.

[12] Ulrich Bauer and Michael Lesnick. Induced matchings and the algebraic stability of persistence barcodes. *arXiv preprint arXiv:1311.3681*, 2013.

[13] Ulrich Bauer and Abhishek Rathod. Hardness of approximation for Morse Matching. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '19, pages 2663–2674, Philadelphia, PA, USA, 2019. Society for Industrial and Applied Mathematics.

[14] Ulrich Bauer and Abhishek Rathod. Parameterized inapproximability of morse matching. *arXiv preprint arXiv:2109.04529*, 2021.

[15] Ulrich Bauer, Abhishek Rathod, and Jonathan Spreer. Parametrized Complexity of Expansion Height. In Michael A. Bender, Ola Svensson, and Grzegorz Herman, editors, *27th Annual European Symposium on Algorithms (ESA 2019)*, volume 144 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 13:1–13:15, Dagstuhl, Germany, 2019. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. doi:10.4230/LIPIcs.ESA.2019.13.

[16] Ulrich Bauer, Abhishek Rathod, and Meirav Zehavi. The complexity of high-dimensional cuts. *arXiv preprint arXiv:2108.10195*, 2021.

[17] Gilbert Baumslag, Alexei G. Myasnikov, and Vladimir Shpilrain. Open problems in combinatorial group theory. Second edition. In *Combinatorial and geometric group theory (New York, 2000/Hoboken, NJ, 2001)*, volume 296 of *Contemp. Math.*, pages 1–38. Amer. Math. Soc., Providence, RI, 2002. URL: https://doi.org/10.1090/conm/296/05067, doi:10.1090/conm/296/05067.

[18] Franziska Berger, Peter Gritzmann, and Sven de Vries. Minimum cycle bases and their applications. In *Algorithmics of large and complex networks*, pages 34–49. Springer, 2009.

[19] Dimitris Bertsimas and John N Tsitsiklis. *Introduction to linear optimization*, volume 6. Athena Scientific Belmont, MA, 1997.

[20] Håvard Bakke Bjerkevik and Magnus Bakke Botnan. Computational Complexity of the Interleaving Distance. In Bettina Speckmann and Csaba D. Tóth, editors, *34th International Symposium on Computational Geometry (SoCG 2018)*, volume 99 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 13:1–13:15, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. URL: http://drops.dagstuhl.de/opus/volltexte/2018/8726, doi:10.4230/LIPIcs.SoCG.2018.13.

[21] Håvard Bakke Bjerkevik, Magnus Bakke Botnan, and Michael Kerber. Computing the interleaving distance is np-hard. *Foundations of Computational Mathematics*, 20(5):1237–1271, 2020.

[22] Jean-Daniel Boissonnat, Frédéric Chazal, and Mariette Yvinec. *Geometric and topological inference*, volume 57. Cambridge University Press, 2018.

[23] Glencora Borradaile, Erin Wolf Chambers, Kyle Fox, and Amir Nayyeri. Minimum cycle and homology bases of surface-embedded graphs. *JoCG*, 8(2):58–79, 2017.

[24] Magnus Botnan and William Crawley-Boevey. Decomposition of persistence modules. *Proceedings of the American Mathematical Society*, 148(11):4581–4596, 2020.

[25] R Sean Bowman and Stephen B McCaul. Fast searching for andrews–curtis trivializations. *Experimental Mathematics*, 15(2):193–197, 2006.

[26] Glen E Bredon. *Topology and geometry*, volume 139. Springer Science & Business Media, 2013.

[27] Martin R Bridson. The complexity of balanced presentations and the andrews-curtis conjecture. *arXiv preprint arXiv:1504.04187*, 2015.

[28] Peter Bubenik et al. Statistical topological data analysis using persistence landscapes. *J. Mach. Learn. Res.*, 16(1):77–102, 2015.

[29] Peter Bubenik and Jonathan A Scott. Categorification of persistent homology. *Discrete & Computational Geometry*, 51(3):600–627, 2014.

[30] Benjamin A Burton. Computational topology with regina: Algorithms, heuristics and implementations. *Geometry & Topology Down Under*, pages 195–224, 2013.

[31] Benjamin A. Burton and Rodney G. Downey. Courcelle's theorem for triangulations. *Journal of Combinatorial Theory, Series A*, 146:264–294, 2017. URL: https://www.sciencedirect.com/science/article/pii/S0097316516301005, doi:https://doi.org/10.1016/j.jcta.2016.10.001.

[32] Benjamin A. Burton, Thomas Lewiner, João Paixão, and Jonathan Spreer. Parameterized complexity of discrete Morse theory. *ACM Trans. Math. Software*, 42(1):Art. 6, 24, 2016. doi:10.1145/2738034.

[33] Oleksiy Busaryev, Sergio Cabello, Chao Chen, Tamal K. Dey, and Yusu Wang. Annotating simplices with a homology basis and its applications. In Fedor V. Fomin and Petteri Kaski, editors, *Algorithm Theory – SWAT 2012*, pages 189–200, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[34] Gunnar Carlsson. Topology and data. *Bulletin of the American Mathematical Society*, 46(2):255–308, 2009.

[35] Gunnar Carlsson and Vin De Silva. Zigzag persistence. *Foundations of computational mathematics*, 10(4):367–405, 2010.

[36] Mathieu Carriere, Marco Cuturi, and Steve Oudot. Sliced wasserstein kernel for persistence diagrams. In *International conference on machine learning*, pages 664–673. PMLR, 2017.

[37] Manoj K. Chari. On discrete Morse functions and combinatorial decompositions. *Discrete Mathematics*, 217(1-3):101–113, 2000. doi:10.1016/S0012-365X(99)00258-7.

[38] Chao Chen and Daniel Freedman. Measuring and computing natural generators for homology groups. *Comput. Geom. Theory Appl.*, 43(2):169–181, February 2010. doi:10.1016/j.comgeo.2009.06.004.

[39] Chao Chen and Daniel Freedman. Hardness results for homology localization. *Discrete & Computational Geometry*, 45(3):425–448, Apr 2011. doi:10.1007/s00454-010-9322-8.

[40] Jianer Chen, Benny Chor, Mike Fellows, Xiuzhen Huang, David Juedes, Iyad A Kanj, and Ge Xia. Tight lower bounds for certain parameterized np-hard problems. *Information and Computation*, 201(2):216–231, 2005.

[41] Yijia Chen, Martin Grohe, and Magdalena Grüber. On parameterized approximability. In Hans L. Bodlaender and Michael A. Langston, editors, *Parameterized and Exact Computation*, pages 109–120, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[42] F. R. K. Chung. *Spectral Graph Theory*. American Mathematical Society, 1997.

[43] Alan Cobham. The intrinsic computational difficulty of functions. In Yehoshua Bar-Hillel, editor, *Logic, Methodology and Philosophy of Science: Proceedings of the 1964 International Congress (Studies in Logic and the Foundations of Mathematics)*, pages 24–30. North-Holland Publishing, 1965.

[44] Marshall M Cohen. *A Course in Simple-Homotopy Theory*, volume 10. Springer Science & Business Media, 2012.

[45] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Extending persistence using Poincaré and Lefschetz duality. *Foundations of Computational Mathematics*, 9(1):79–103, Feb 2009. doi:10.1007/s10208-008-9027-z.

[46] Thomas H Cormen, Charles E Leiserson, Ronald L Rivest, and Clifford Stein. *Introduction to algorithms*. MIT press, 2009.

[47] P. Crescenzi. A short guide to approximation preserving reductions. In *Proceedings of the 12th Annual IEEE Conference on Computational Complexity*, CCC '97, page 262, USA, 1997. IEEE Computer Society.

[48] Justin Michael Curry. *Sheaves, cosheaves and applications*. University of Pennsylvania, 2014.

[49] Carina Curto. What can topology tell us about the neural code? *Bulletin of the American Mathematical Society*, 54(1):63–78, 2017.

[50] Marek Cygan, Fedor V. Fomin, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michal Pilipczuk, and Saket Saurabh. *Parameterized Algorithms*. Springer, 2015. URL: http://dx.doi.org/10.1007/978-3-319-21275-3, doi:10.1007/978-3-319-21275-3.

[51] José Coelho de Pina. *Applications of shortest path methods*. PhD thesis, Universiteit van Amsterdam, 1995.

[52] Tamal K Dey, Herbert Edelsbrunner, and Sumanta Guha. Computational topology. *Contemporary mathematics*, 223:109–144, 1999.

[53] Tamal K. Dey, Tianqi Li, and Yusu Wang. Efficient algorithms for computing a minimal homology basis. In Michael A. Bender, Martín Farach-Colton, and Miguel A. Mosteiro, editors, *LATIN 2018: Theoretical Informatics*, pages 376–398, Cham, 2018. Springer International Publishing.

[54] Tamal K. Dey, Jian Sun, and Yusu Wang. Approximating loops in a shortest homology basis from point data. In *Proceedings of the Twenty-sixth Annual Symposium on Computational Geometry*, SoCG '10, pages 166–175, New York, NY, USA, 2010. ACM. doi:10.1145/1810959.1810989.

[55] Tamal K. Dey and Yusu Wang. Reeb graphs: Approximation and persistence. *Discrete & Computational Geometry*, 49(1):46–73, Jan 2013. doi:10.1007/s00454-012-9463-z.

[56] Tamal Krishna Dey and Yusu Wang. *Computational Topology for Data Analysis*. Cambridge University Press, 2021.

[57] Michel Marie Deza and Monique Laurent. *Geometry of cuts and metrics*, volume 2. Springer, 1997.

[58] Yotam Dikstein, Irit Dinur, Prahladh Harsha, and Noga Ron-Zewi. Locally testable codes via high-dimensional expanders. *arXiv preprint arXiv:2005.01045*, 2020.

[59] Irit Dinur. The pcp theorem by gap amplification. *Journal of the ACM (JACM)*, 54(3):12–es, 2007.

[60] Irit Dinur, Prahladh Harsha, Tali Kaufman, Inbal Livni Navon, and Amnon Ta-Shma. List-decoding with double samplers. *SIAM Journal on Computing*, 50(2):301–349, 2021.

[61] Irit Dinur and Tali Kaufman. High dimensional expanders imply agreement expanders. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 974–985. IEEE, 2017.

[62] Dominic Dotterrer, Tali Kaufman, and Uli Wagner. On expansion and topological overlap. *Geometriae Dedicata*, 195(1):307–317, 2018. URL: https://doi.org/10.1007/s10711-017-0291-4, doi:10.1007/s10711-017-0291-4.

[63] R. G. Downey and M. R. Fellows. *Parameterized complexity*. Monographs in Computer Science. Springer-Verlag, New York, 1999. doi:10.1007/978-1-4612-0515-9.

[64] Rod Downey. *The Birth and Early Years of Parameterized Complexity*, pages 17–38. Springer-Verlag, Berlin, Heidelberg, 2012.

[65] Jean-Guillaume Dumas, Clément Pernet, and Ziad Sultan. Simultaneous computation of the row and column rank profiles. In *Proceedings of the 38th International Symposium on Symbolic and Algebraic Computation*, ISSAC '13, pages 181–188, New York, NY, USA, 2013. ACM. doi:10.1145/2465506.2465517.

[66] H. Edelsbrunner and S. Parsa. On the computational complexity of Betti numbers: Reductions from matrix rank. In *Proceedings of the Twenty-fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '14, pages 152–160, Philadelphia, PA, USA, 2014. Society for Industrial and Applied Mathematics.

[67] Herbert Edelsbrunner and John Harer. *Computational Topology: An Introduction*. 01 2010. doi:10.1007/978-3-540-33259-6_7.

[68] Herbert Edelsbrunner, John Harer, et al. Persistent homology-a survey. *Contemporary mathematics*, 453:257–282, 2008.

[69] Jack Edmonds. *Paths, Trees, and Flowers*, pages 361–379. Birkhäuser Boston, Boston, MA, 1987. URL: https://doi.org/10.1007/978-0-8176-4842-8_26, doi:10.1007/978-0-8176-4842-8_26.

[70] Ömer Eğecioğlu and Teofilo F. Gonzalez. A computationally intractable problem on simplicial complexes. *Computational Geometry*, 6(2):85 – 98, 1996. doi:https://doi.org/10.1016/0925-7721(95)00015-1.

[71] K. Eickmeyer, M. Grohe, and M. Grüber. Approximation of natural w[p]-complete minimisation problems is hard. In *2008 23rd Annual*

IEEE Conference on Computational Complexity, pages 8–18, June 2008. doi:10.1109/CCC.2008.24.

[72] Jeff Erickson. Combinatorial optimization of cycles and bases. *Advances in applied and computational topology*, 70:195–228, 2012.

[73] Jeff Erickson and Kim Whittlesey. Greedy optimal homotopy and homology generators. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '05, pages 1038–1046, Philadelphia, PA, USA, 2005. Society for Industrial and Applied Mathematics.

[74] Shai Evra and Tali Kaufman. Bounded degree cosystolic expanders of every dimension. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 36–48, 2016.

[75] Brittany Terese Fasy, Fabrizio Lecci, Alessandro Rinaldo, Larry Wasserman, Sivaraman Balakrishnan, and Aarti Singh. Confidence sets for persistence diagrams. *The Annals of Statistics*, 42(6):2301–2339, 2014.

[76] Andreas Emil Feldmann, Karthik C. S., Euiwoong Lee, and Pasin Manurangsi. A survey on approximation in parameterized complexity: Hardness and algorithms. *Algorithms*, 13(6), 2020. URL: https://www.mdpi.com/1999-4893/13/6/146, doi:10.3390/a13060146.

[77] Michael R. Fellows, Danny Hermelin, Frances Rosamond, and Stéphane Vialette. On the parameterized complexity of multiple-interval graph problems. *Theoretical Computer Science*, 410(1):53–61, 2009. URL: https://www.sciencedirect.com/science/article/pii/S0304397508007329, doi:https://doi.org/10.1016/j.tcs.2008.09.065.

[78] Ximena Fernández. 3-deformations of 2-complexes and morse theory. *arXiv preprint arXiv:1912.00115*, 2019.

[79] Ximena Laura Fernández. *Métodos combinatorios y algoritmos en topología de dimensiones bajas y la conjetura de Andrews-Curtis*. PhD thesis, Universidad de Buenos Aires. Facultad de Ciencias Exactas y Naturales, 2017.

[80] Jörg Flum and Martin Grohe. *Parameterized Complexity Theory*. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2006.

[81] Fedor Fomin and Dimitrios Thilikos. *Branchwidth of Graphs*, pages 101–105. Springer US, Boston, MA, 2008. URL: https://doi.org/10.1007/978-0-387-30162-4_55, doi:10.1007/978-0-387-30162-4_55.

[82] Fedor V Fomin, Daniel Lokshtanov, Saket Saurabh, and Meirav Zehavi. *Kernelization: Theory of Parameterized Preprocessing*. Cambridge University Press, 2019.

[83] Robin Forman. Morse theory for cell complexes. *Advances in Mathematics*, 134(1):90–145, 1998. doi:10.1006/aima.1997.1650.

[84] Robin Forman. A user's guide to discrete Morse theory. *Séminaire Lotharingien de Combinatoire*, B48c:1–35, 2002.

[85] Robert W Ghrist. *Elementary applied topology*, volume 1. Createspace Seattle, WA, 2014.

[86] Shawn Harker, Konstantin Mischaikow, Marian Mrozek, Vidit Nanda, Hubert Wagner, Mathias Juda, and Pavel Dlotko. The efficiency of a homology algorithm based on discrete Morse theory and coreductions. In *Proc. of 3rd Intl. Workshop on CTIC*, volume 1(1), 2010. doi:11441/26192.

[87] Joel Hass, Jeffrey C Lagarias, and Nicholas Pippenger. The computational complexity of knot and link problems. *Journal of the ACM (JACM)*, 46(2):185–211, 1999.

[88] Allen Hatcher. *Algebraic topology*. Cambridge University Press, Cambridge, 2002.

[89] George Havas and Colin Ramsay. Andrews-curtis and todd-coxeter proof words. *Groups St Andrews 2001 in Oxford*, pages 232–237, 2003.

[90] Gregory Henselman and Robert Ghrist. Matroid filtrations and computational persistent homology. *arXiv preprint arXiv:1606.00199*, 2016.

[91] Yasuaki Hiraoka, Takenobu Nakamura, Akihiko Hirata, Emerson G Escolar, Kaname Matsue, and Yasumasa Nishiura. Hierarchical structures of amorphous solids characterized by persistent homology. *Proceedings of the National Academy of Sciences*, 113(26):7035–7040, 2016.

[92] Cynthia Hog-Angeloni and Wolfgang Metzler, editors. *Two-dimensional homotopy and combinatorial group theory*, volume 197 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1993. URL: https://doi.org/10.1017/CBO9780511629358, doi:10.1017/CBO9780511629358.

[93] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006.

[94] J. D. Horton. A polynomial-time algorithm to find the shortest cycle basis of a graph. *SIAM J. Comput.*, 16(2):358–366, April 1987. doi:10.1137/0216026.

[95] Ken ichi Kawarabayashi, Yusuke Kobayashi, and Bruce Reed. The disjoint paths problem in quadratic time. *Journal of Combinatorial Theory, Series B*, 102(2):424–435, 2012. URL: https://www.sciencedirect.

com/science/article/pii/S0095895611000712, doi:https://doi.org/10.1016/j.jctb.2011.07.004.

[96] Fáry István. On straight-line representation of planar graphs. *Acta scientiarum mathematicarum*, 11(229-233):2, 1948.

[97] Thomas Jansen. Introduction to the theory of complexity and approximation algorithms. In *Lectures on Proof Verification and Approximation Algorithms*, pages 5–28. Springer, 1998.

[98] Claude-Pierre Jeannerod, Clément Pernet, and Arne Storjohann. Rank-profile revealing gaussian elimination and the cup matrix decomposition. *Journal of Symbolic Computation*, 56:46 – 68, 2013. doi:https://doi.org/10.1016/j.jsc.2013.04.004.

[99] Michael Joswig, Frank H. Lutz, and Mimi Tsuruga. Heuristics for sphere recognition. In Hoon Hong and Chee Yap, editors, *ICMS 2014: Mathematical Software*, pages 152–159, Berlin, Heidelberg, 2014. Springer Berlin Heidelberg. doi:10.1007/978-3-662-44199-2_26.

[100] Michael Joswig and Marc E. Pfetsch. Computing optimal Morse matchings. *SIAM Journal on Discrete Mathematics*, 20(1):11–25, 2006. doi:10.1137/S0895480104445885.

[101] Tomasz Kaczynski, Konstantin Michael Mischaikow, and Marian Mrozek. *Computational homology*, volume 3. Springer, 2004.

[102] Masaki Kashiwara and Pierre Schapira. Persistent homology and microlocal sheaf theory. *Journal of Applied and Computational Topology*, 2(1):83–113, 2018.

[103] Tali Kaufman, David Kazhdan, and Alexander Lubotzky. Ramanujan complexes and bounded degree topological expanders. In *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, pages 484–493. IEEE, 2014.

[104] Telikepalli Kavitha, Christian Liebchen, Kurt Mehlhorn, Dimitrios Michail, Romeo Rizzi, Torsten Ueckerdt, and Katharina A. Zweig. Cycle bases in graphs characterization, algorithms, complexity, and applications. *Computer Science Review*, 3(4):199 – 243, 2009. doi:https://doi.org/10.1016/j.cosrev.2009.08.001.

[105] Telikepalli Kavitha, Kurt Mehlhorn, and Dimitrios Michail. New approximation algorithms for minimum cycle bases of graphs. *Algorithmica*, 59(4):471–488, Apr 2011. doi:10.1007/s00453-009-9313-4.

[106] Telikepalli Kavitha, Kurt Mehlhorn, Dimitrios Michail, and Katarzyna Paluch. A faster algorithm for minimum cycle basis of graphs. In Josep

Díaz, Juhani Karhumäki, Arto Lepistö, and Donald Sannella, editors, *Automata, Languages and Programming*, pages 846–857, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.

[107] Subhash Khot. On the unique games conjecture (invited survey). In *2010 IEEE 25th Annual Conference on Computational Complexity*, pages 99–121, 2010. doi:10.1109/CCC.2010.19.

[108] Victor Klee and George J Minty. How good is the simplex algorithm. *Inequalities*, 3(3):159–175, 1972.

[109] Dmitry Kozlov. Discrete Morse Theory. In *Combinatorial Algebraic Topology*, volume 21 of *Algorithms and Computation in Mathematics*, chapter 11, pages 179–209. Springer Berlin Heidelberg, 2008. doi:10.1007/978-3-540-71962-5_11.

[110] Michael Lesnick. The theory of the interleaving distance on multidimensional persistence modules. *Foundations of Computational Mathematics*, 15(3):613–650, 2015.

[111] Michael Lesnick and Matthew Wright. Computing minimal presentations and bigraded betti numbers of 2-parameter persistent homology. *arXiv preprint arXiv:1902.05708*, 2019.

[112] Richard J. Lipton and Kenneth W. Regan. *David Johnson: Galactic Algorithms*, pages 109–112. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013. URL: https://doi.org/10.1007/978-3-642-41422-0_20, doi:10.1007/978-3-642-41422-0_20.

[113] Daniel Lokshtanov, Pranabendu Misra, Michał Pilipczuk, Saket Saurabh, and Meirav Zehavi. An exponential time parameterized algorithm for planar disjoint paths. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2020, pages 1307–1316, New York, NY, USA, 2020. Association for Computing Machinery. URL: https://doi.org/10.1145/3357713.3384250, doi:10.1145/3357713.3384250.

[114] Daniel Lokshtanov, Saket Saurabh, and Meirav Zehavi. *Efficient Graph Minors Theory and Parameterized Algorithms for (Planar) Disjoint Paths*, pages 112–128. Springer International Publishing, Cham, 2020. URL: https://doi.org/10.1007/978-3-030-42071-0_9, doi:10.1007/978-3-030-42071-0_9.

[115] László Lovász. Graph minor theory. *Bulletin of the American Mathematical Society*, 43(1):75–86, 2006.

[116] Alexander Lubotzky. High dimensional expanders. In *Proceedings of the International Congress of Mathematicians: Rio de Janeiro 2018*, pages 705–730. World Scientific, 2018.

[117] Alexander Lubotzky and Roy Meshulam. Random latin squares and 2-dimensional expanders. *Advances in Mathematics*, 272:743–760, 2015.

[118] Dániel Marx. Parameterized complexity and approximation algorithms. *The Computer Journal*, 51(1):60–78, 2008.

[119] Dániel Marx. Completely inapproximable monotone and antimonotone parameterized problems. *J. Comput. Syst. Sci.*, 79(1):144–151, February 2013. doi:10.1016/j.jcss.2012.09.001.

[120] Sergei Vladimirovich Matveev and SV Matveev. *Algorithmic topology and classification of 3-manifolds*, volume 9. Springer, 2007.

[121] Kurt Mehlhorn and Dimitrios Michail. Minimum cycle bases: Faster and simpler. *ACM Trans. Algorithms*, 6(1):8:1–8:13, December 2009. doi:10.1145/1644015.1644023.

[122] Alexei D Miasnikov. Genetic algorithms and the andrews–curtis conjecture. *International Journal of Algebra and Computation*, 9(06):671–686, 1999.

[123] John Milnor. Whitehead torsion. *Bulletin of the American Mathematical Society*, 72(3):358–426, 1966.

[124] Pavel Mnev. Lecture notes on torsions. *arXiv preprint arXiv:1406.3705*, 2014.

[125] James R. Munkres. *Elements of algebraic topology*. Addison-Wesley Publishing Company, Menlo Park, CA, 1984.

[126] Alexei Myasnikov, Alexander Ushakov, and Dong Wook Won. The word problem in the baumslag group with a non-elementary dehn function is polynomial time decidable. *Journal of Algebra*, 345(1):324–342, 2011.

[127] Alexei D. Myasnikov, Alexei G. Myasnikov, and Vladimir Shpilrain. On the Andrews-Curtis equivalence. In *Combinatorial and geometric group theory (New York, 2000/Hoboken, NJ, 2001)*, volume 296 of *Contemp. Math.*, pages 183–198. Amer. Math. Soc., Providence, RI, 2002. URL: https://doi.org/10.1090/conm/296/05074, doi:10.1090/conm/296/05074.

[128] Steve Y Oudot. *Persistence theory: from quiver representations to data analysis*, volume 209. American Mathematical Society Providence, 2015.

[129] Teemu Päkkilä et al. Discrete morse theory from a simple-homotopy point of view. 2014.

[130] Christos H. Papadimitriou and Mihalis Yannakakis. Optimization, approximation, and complexity classes. *Journal of Computer and System Sciences*, 43(3):425 – 440, 1991. doi:10.1016/0022-0000(91)90023-X.

[131] Salman Parsa. A deterministic O(m log m) time algorithm for the Reeb graph. In *Proceedings of the Twenty-eighth Annual Symposium on Computational Geometry*, SoCG '12, pages 269–276, New York, NY, USA, 2012. ACM. doi:10.1145/2261250.2261289.

[132] Grisha Perelman. The entropy formula for the ricci flow and its geometric applications. *arXiv preprint math/0211159*, 2002.

[133] Grisha Perelman. Finite extinction time for the solutions to the ricci flow on certain three-manifolds. *arXiv preprint math/0307245*, 2003.

[134] Grisha Perelman. Ricci flow with surgery on three-manifolds. *arXiv preprint math/0303109*, 2003.

[135] Raúl Rabadán and Andrew J Blumberg. *Topological data analysis for genomics and evolution: topology in biology*. Cambridge University Press, 2019.

[136] Abhishek Rathod. Fast Algorithms for Minimum Cycle Basis and Minimum Homology Basis. In *36th International Symposium on Computational Geometry (SoCG 2020)*, volume 164, pages 64:1–64:11, 2020.

[137] Abhishek Rathod, Talha Bin Masood, and Vijay Natarajan. Approximation algorithms for Max Morse Matching. *Computational Geometry*, 61:1 – 23, 2017. doi:10.1016/j.comgeo.2016.10.002.

[138] Jan Reininghaus, Stefan Huber, Ulrich Bauer, and Roland Kwitt. A stable multi-scale kernel for topological machine learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4741–4748, 2015.

[139] N. Robertson and P.D. Seymour. Graph minors .xiii. the disjoint paths problem. *Journal of Combinatorial Theory, Series B*, 63(1):65–110, 1995. URL: https://www.sciencedirect.com/science/article/pii/S0095895685710064, doi:https://doi.org/10.1006/jctb.1995.1006.

[140] Neil Robertson and P.D. Seymour. Graph minors. i. excluding a forest. *Journal of Combinatorial Theory, Series B*, 35(1):39–61, 1983. URL: https://www.sciencedirect.com/science/article/pii/0095895683900795, doi:https://doi.org/10.1016/0095-8956(83)90079-5.

[141] Neil Robertson and P.D Seymour. Graph minors. iii. planar tree-width. *Journal of Combinatorial Theory, Series B*, 36(1):49–64, 1984. URL: https://www.sciencedirect.com/science/article/pii/0095895684900133, doi:https://doi.org/10.1016/0095-8956(84)90013-3.

[142] Neil Robertson and P.D Seymour. Graph minors. x. obstructions to tree-decomposition. *Journal of Combinatorial Theory, Series B*, 52(2):153–190, 1991. URL: https://www.sciencedirect.com/science/article/

pii/009589569190061N, doi:https://doi.org/10.1016/0095-8956(91)90061-N.

[143] Neil Robertson and P.D. Seymour. Graph minors. xx. wagner's conjecture. *Journal of Combinatorial Theory, Series B*, 92(2):325–357, 2004. Special Issue Dedicated to Professor W.T. Tutte. URL: https://www.sciencedirect.com/science/article/pii/S0095895604000784, doi:https://doi.org/10.1016/j.jctb.2004.08.001.

[144] Alexander Rolle and Luis Scoccola. Stable and consistent density-based clustering. *arXiv preprint arXiv:2005.09048*, 2020.

[145] Tim Roughgarden, editor. *Beyond the Worst-Case Analysis of Algorithms*. Cambridge University Press, 2021. doi:10.1017/9781108637435.

[146] Colin P Rourke and Brian Joseph Sanderson. *Introduction to piecewise-linear topology*. Springer Science & Business Media, 2012.

[147] Andrés Santamaría-Galvis and Russ Woodroofe. Shellings from relative shellings, with an application to np-completeness. *Discrete & Computational Geometry*, 2021. doi:10.1007/s00454-020-00273-1.

[148] Arne Storjohann and Shiyun Yang. Linear independence oracles and applications to rectangular and low rank linear systems. In *Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation*, ISSAC '14, pages 381–388, New York, NY, USA, 2014. ACM. doi:10.1145/2608628.2608673.

[149] Arne Storjohann and Shiyun Yang. A relaxed algorithm for online matrix inversion. In *Proceedings of the 2015 ACM on International Symposium on Symbolic and Algebraic Computation*, ISSAC '15, pages 339–346, New York, NY, USA, 2015. ACM. doi:10.1145/2755996.2756672.

[150] Martin Tancer. Recognition of collapsible complexes is NP-complete. *Discrete Comput. Geom.*, 55(1):21–38, 2016. doi:10.1007/s00454-015-9747-1.

[151] Tammo tom Dieck. *Algebraic topology*, volume 8. European Mathematical Society, 2008.

[152] Luca Trevisan. On khot's unique games conjecture. *Bulletin of the American Mathematical Society*, 49(1):91–111, 2012.

[153] Katharine Turner, Sayan Mukherjee, and Doug M Boyer. Persistent homology transform for modeling shapes and surfaces. *Information and Inference: A Journal of the IMA*, 3(4):310–344, 2014.

[154] K. Wagner. Über eine eigenschaft der ebenen komplexe. *Mathematische Annalen*, 114(1):570–590, 1937. URL: https://doi.org/10.1007/BF01594196, doi:10.1007/BF01594196.

[155] J. H. C. Whitehead. Simplicial spaces, nuclei and m-groups. *Proceedings of the London Mathematical Society*, s2-45(1):243–327, 1939. URL: https://londmathsoc.onlinelibrary.wiley.com/doi/abs/10.1112/plms/s2-45.1.243, arXiv:https://londmathsoc.onlinelibrary.wiley.com/doi/pdf/10.1112/plms/s2-45.1.243, doi:https://doi.org/10.1112/plms/s2-45.1.243.

[156] J. H. C. Whitehead. Simple homotopy types. *American Journal of Mathematics*, 72(1):1–57, 1950. URL: http://www.jstor.org/stable/2372133.

[157] Virginia Vassilevska Williams. On some fine-grained questions in algorithms and complexity. In *Proceedings of the International Congress of Mathematicians: Rio de Janeiro 2018*, pages 3447–3487. World Scientific, 2018.

[158] David P. Williamson and David B. Shmoys. *The Design of Approximation Algorithms*. Cambridge University Press, 2011.

[159] Perrin Wright. Group presentations and formal deformations. *Transactions of the American Mathematical Society*, 208:161–169, 1975.

[160] Shiyun Yang. Algorithms for fast linear system solving and rank profile computation. Master's thesis, University of Waterloo, 2014.

[161] E. C. Zeeman. On the dunce hat. *Topology*, 2:341–358, 1964. doi:10.1016/0040-9383(63)90014-4.

APPENDICES

A

PAPER I: HARDNESS OF APPROXIMATION FOR MORSE MATCHING

Classical Morse theory provides a rich toolkit for analyzing and inferring the topology of a smooth manifolds by studying the critical points of smooth functions defined on it. Forman's discrete Morse theory is a combinatorial analogue of Morse theory that is applicable to regular CW complexes. It has become a popular tool in the visualization community, and is actively studied in algebraic combinatorics. Discrete Morse theory has also found applications as a preprocessing tool for speeding up computations in topological data analysis.

The principal construct in Forman's theory is the so-called discrete gradient vector field defined on a simplicial (or more generally a CW) complex. A discrete gradient vector field is a collection of facet-cofacet pairs satisfying a certain acyclicity condition. The simplices that do not belong to any of the pairs in the gradient are deemed critical. Forman's theory also has an equivalent graph theoretic formulation in which the acyclic matchings (also called *Morse matchings*) in the Hasse diagram of a simplicial complex correspond to the discrete gradient vector fields on the simplicial complex. There has been a lot of practical interest in computing gradient vector fields on simplicial complexes with (near-)optimal number of critical simplices. The problem of finding a gradient vector field with the minimum number of critical simplices in a simplicial complex was shown to be **NP**-hard by Joswig and Pfetsch, who also posed its approximability as an open question. In this article, we resolve the open question by establishing hardness of approximation results for the maximization and minimization variants of the Morse matching problem.

Following a brief introduction in Section 1, we cover the requisite background from approximation algorithms and discrete Morse theory in Section 2. In Section 3, we prove a strong hardness result for the minimization variant of Morse matching by utilizing Tancer's result on **NP**-hardness of determining collapsibility. In Section 4, we prove a hardness result for the maximization variant of Morse matching via an L-reduction from the Max-Acyclic Subgraph problem restricted to degree 3 graphs.

STATEMENT OF INDIVIDUAL CONTRIBUTION. I suggested this project to my advisor Uli Bauer. I found the main ideas behind the hardness results and presented them to him. Later, we had several discussions, where more details were fleshed out, and Uli Bauer made several key suggestions for improvements. In this collaboration, I was the main author responsible for finding the ideas and and for writing them down. Since this was the first article of my PhD, I got some invaluable writing lessons from my advisor.

COPYRIGHT TRANSFER AGREEMENT

In order for SIAM to include your paper in the SIAM Symposium on Discrete Algorithms (SODA19) proceedings, the following Copyright Transfer Agreement must be completed during the final paper upload process.

Title of Paper: Hardness of Approximation for Morse Matching

Paper ID Number: 622

Copyright to this paper is hereby irrevocably assigned to SIAM for publication in the ACM-SIAM Symposium on Discrete Algorithms (SODA19), January 6-9, 2019 at the Westin San Diego. SIAM has sole use for distribution in all forms and media, such as microfilm and anthologies, except that the author(s) or, in the case of a "work made for hire," the employer will retain:

> The right to use all or part of the content of the paper in future works of the author(s), including author's teaching, technical collaborations, conference presentations, lectures, or other scholarly works and professional activities as well as to the extent the fair use provisions of the U.S. Copyright Act permit. If the copyright is granted to SIAM, then the proper notice of the SIAM's copyright should be provided.
>
> The right to post the final draft of the paper on noncommercial pre-print serves like arXiv.org.
>
> The right to post the final version of the paper on the author's personal web site and on the web server of the author's institution, provided the proper notice of the SIAM's copyright is included and that no separate or additional fees are collected for access to or distribution of the paper.
>
> The right to refuse permission to third parties to republish all or part of the paper or translation thereof.

It is affirmed neither this paper nor portions of it have been published elsewhere and a copyright transfer agreement has not been signed permitting the publication of a similar paper in a journal or elsewhere. For multi-author works, the signing author agrees to notify all co-authors of his/her action.

Transfer of Copyright to the Publisher
SIAM strongly recommends this option. This transfer of copyright provides SIAM the legal basis not only to publish and to distribute the work, but also to pursue infringements of copyright (such as plagiarism and other forms of unauthorized use) and to grant permissions for the legitimate uses of the work by third parties. This option should not be selected if the work was prepared by a government office or employee as part of his or her official duties.

[✔] The Author hereby irrevocably assigns, conveys and transfers the copyright to the Work to SIAM. SIAM shall have sole rights of distribution and/or publication of the work in all forms and media, throughout the world, except for those rights given to the Author above.

[__] The Author DOES NOT assign, convey and transfer the copyright to the Work to SIAM. Please list in whose name the copyright should appear here:

_____

Alternate Copyright

If you do not select either of the options above, an alternate copyright is required. Check below to submit an alternative copyright. Send proposed text for alternate copyright, with a copy of this completed form, as a single document, to wilden@siam.org, Use the subject line "SIAM –SODA19 – Alternate Copyright – LAST NAME."

[__] I am submitting an alternative copyright.

Work Made for Hire

[__] Check here if signature is on behalf of employer in the event article is "work made for hire."

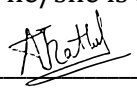
Previously Published

Have portions of this presentation been published elsewhere? If so, select "yes" and enclose appropriate credits and permissions to republish.

[__] Yes
[✔] No

Signatures

This form is to be signed by the Author or, in the case of "work-made-for-hire," by the employer. If each author's signature does not appear below, the signing author represents that he/she is authorized to execute this agreement for and on behalf of all the authors



_____ 31/10/2018

Signature Date

Abhishek Jayantilal Rathod_____

Print Name

Gmail

Abhishek Rathod <abhishek.rathod@gmail.com>

Permission to use material in doctoral thesis

Kelly Thomas <Thomas@siam.org> Tue, Sep 14, 2021 at 8:28 PM
To: Abhishek Rathod <abhishek.rathod@gmail.com>

Dear Mr. Rathod:

You do have permission to reuse material from the proceedings paper mentioned below in your thesis. Please acknowledge the original publication, using the complete bibliographic information.

Sincerely,

Kelly Thomas

Managing Editor

Society for Industrial and Applied Mathematics

3600 Market Street - 6th Floor
Philadelphia, PA 19104

thomas@siam.org / (267) 350-6387

From: Abhishek Rathod [mailto:abhishek.rathod@gmail.com]
Sent: Tuesday, September 14, 2021 5:14 AM
To: Kelly Thomas <Thomas@siam.org>
Subject: Permission to use material in doctoral thesis

[Quoted text hidden]

Hardness of Approximation for Morse Matching *

Ulrich Bauer † Abhishek Rathod ‡

Abstract

Discrete Morse theory has emerged as a powerful tool for a wide range of problems, including the computation of (persistent) homology. In this context, discrete Morse theory is used to reduce the problem of computing a topological invariant of an input simplicial complex to computing the same topological invariant of a (significantly smaller) collapsed cell or chain complex. Consequently, devising methods for obtaining gradient vector fields on complexes to reduce the size of the problem instance has become an emerging theme over the last decade. While computing the optimal gradient vector field on a simplicial complex is NP-hard, several heuristics have been observed to compute near-optimal gradient vector fields on a wide variety of datasets. Understanding the theoretical limits of these strategies is therefore a fundamental problem in computational topology. In this paper, we consider the approximability of maximization and minimization variants of the Morse matching problem. We establish hardness results for *Max-Morse matching* and *Min-Morse matching*, settling an open problem posed by Joswig and Pfetsch [20]. In particular, we show that, for a simplicial complex of dimension $d \geq 3$ with n simplices, it is NP-hard to approximate *Min-Morse matching* within a factor of $O(n^{1-\epsilon})$, for any $\epsilon > 0$. Moreover, we establish hardness of approximation results for *Max-Morse matching* for simplicial complexes of dimension $d \geq 2$, using an L-reduction from *Degree 3 Max-Acyclic Subgraph* to *Max-Morse matching*.

1 Introduction

Classical Morse theory [29] provides a method to analyze the topology of a smooth manifold by studying the critical points of smooth functions defined on it. Forman's discrete Morse theory is a combinatorial analogue of Morse theory that is applicable to regular cell complexes [10]. It has become a popular tool in computational topology and visualization [5, 8, 36], and is actively studied in algebraic, geometric, and topological combinatorics [18, 23, 28]. In Forman's theory, discrete Morse functions play the role of smooth Morse functions, whereas discrete gradient vector fields are the analogues of gradient-like vector fields. Forman's theory also has an equivalent graph theoretic formulation [6], in which the acyclic matchings (or *Morse matchings*) in the Hasse diagram of a simplicial complex correspond to the discrete gradient vector fields on the simplicial complex. We shall use the terms *gradient vector fields* and *Morse matchings* interchangeably. In the next subsection, we will elaborate on the practical interest [1, 3, 4, 14, 15, 26] in computing gradient vector fields on simplicial complexes with (near-)optimal number of critical simplices (unmatched nodes in the Hasse diagram).

1.1 Motivation The idea of using discrete Morse theory to speedup the computation of homology [15], persistent homology [30], zigzag persistence [27], multidimensional persistence [1], and cellular sheaf cohomology [7] hinges on the promise that discrete Morse theory may reduce the problem of computing homology groups of an input simplicial complex to computing homology groups of a smaller collapsed cell or chain complex. In fact, certain state-of-the-art methods for computing homology groups of complexes [15] like RedHom and CHomP crucially depend on discrete Morse theory. In a follow-up work, Harker et al. [14] devised a discrete Morse theory based framework to efficiently compute the induced map on homology, a problem that arises in Conley index computations. More recently, Brendel et al. [3] designed a discrete Morse theory based algorithm to compute (typically small) presentations of the fundamental group of finite regular CW-complexes, and certain knot invariants. Despite the wide applicability of discrete Morse theory (and associated claims of success), the reader should be aware that discrete Morse theory is not a silver bullet for speeding up topological computations. We advocate a degree of circumspection in evaluating some of the claims about speedup made in the literature. In particular, the reductions based on discrete Morse theory should be considered as heuristics that are often observed to lead to improvements in practice, but usually do not come with any guarantees on the size of the reduced chain complex (both in terms of its rank and the number of nonzero entries of the boundary matrix) or on the complexity of the result-

*Research supported by the DFG Collaborative Research Center TRR 109 *Discretization in Geometry and Dynamics.*

†mail@ulrich-bauer.org, Technical University of Munich.

‡rathod@ma.tum.de, Technical University of Munich.

ing homology computation. In other words, there is no theoretical guarantee that reducing the original chain complex and computing homology of the reduced complex together is any more efficient than the computation for the original complex, and observed improvements in running time might be specific to the implementation.

In terms of computational complexity, it is known that finding an optimal gradient vector field for a simplicial complex is an NP-hard problem, as shown by Joswig et al. [20] via a reduction from the *erasability problem* introduced by Eğecioğlu and Gonzalez [9]. On the other hand, certain heuristics for Morse matching have been reported to be highly effective, often achieving optimality in practice [15, 19, 25]. This naturally raises the question of approximability: to what extent is it feasible to obtain near-optimal solutions for Morse matching in polynomial time? By establishing bounds on the hardness of approximation, we make it evident that for certain instances, the otherwise effective heuristics would fail to compute near-optimal Morse matchings.

1.2 The Morse Matching Problems The Max-Morse Matching problem (MaxMM) can be described as follows: Given a simplicial complex $\mathcal{K}$, compute a gradient vector field that maximizes the cardinality of matched (regular) simplices, over all possible gradient vectors fields on $\mathcal{K}$. Equivalently, the goal is to maximize the number of gradient pairs. For the complementary problem Min-Morse Matching (MinMM), the goal is to compute a gradient vector field that minimizes the number of unmatched (critical) simplices, over all possible gradient vector fields on $\mathcal{K}$. While the problem of finding an exact optimum are equivalent for MinMM and MaxMM, the approximation variants of these problems have vastly different flavors, as we shall note in §3 and §4.

1.3 Related work Based on the relationship between erasability and Morse Matching observed by Lewiner [24, 25], Joswig et al. [20] established NP-completeness of the Morse Matching Problem, using a reduction that is not approximation preserving, and posing the approximability of Morse matching as an open problem [20, Sec. 4]. The algorithmic question of finding optimal discrete Morse functions on simplicial complexes is a well studied problem. Most methods so far have relied on effective heuristics [2, 15, 17, 20, 25]. The first theoretical result in context of Morse matchings was established by Burton et al. [4], who developed a fixed parameter tractable algorithm for computing optimal Morse functions on 3-manifolds. More recently, Rathod et al. [35] proposed the first approximation algorithms for MaxMM on simplicial complexes that provide constant factor approximation bounds for fixed dimension.

1.4 Our contributions In §3, using Tancer's result [37] about NP-completeness of *collapsibility*, we provide a straightforward proof of inapproximability of MinMM on simplicial complexes with dimension $d \geq 3$. In particular, we prove that, assuming $P \neq NP$, there is no $O(n^{1-\epsilon})$-factor approximation algorithm for MinMM for any $\epsilon > 0$, where n denotes the total number of simplices in a given complex $\mathcal{K}$. Then, in §4, we prove that, for any $\epsilon > 0$, approximating MaxMM for simplicial complexes of dimension $d \geq 2$ within a factor of $\left(1 - \frac{1}{4914}\right) + \epsilon$ is NP-hard and approximating it within a factor of $\left(1 - \frac{1}{702}\right) + \epsilon$ is UGC-hard. In particular, this shows that MaxMM has no PTAS unless $P = NP$.

2 Background and Preliminaries

2.1 Simplicial complexes A *k-simplex* $\sigma = \operatorname{conv} V$ is the convex hull of a set V of $(k+1)$ affinely independent points in $\mathbb{R}^d$. We call k the dimension of σ. We say that σ is *spanned* by the points V. Any nonempty subset of V also spans a simplex, a *face* of σ. A simplex σ is said to be a *coface* of a simplex τ if and only if τ is face of σ. We say that σ is a *facet* of τ if σ is a face of τ with $\dim \sigma = \dim \tau - 1$. A *simplicial complex* $\mathcal{K}$ is a collection of simplices that satisfies the following conditions:

- any face of a simplex in $\mathcal{K}$ also belongs to $\mathcal{K}$, and
- the intersection of two simplices $\sigma_1, \sigma_2 \in \mathcal{K}$ is either empty or a face of both σ_1 and σ_2.

For a complex $\mathcal{K}$, we denote the set of d-simplices of $\mathcal{K}$ by $\mathcal{K}^{(d)}$. The *n-skeleton* of a simplicial complex $\mathcal{K}$ is the simplicial complex $\bigcup_{m=0}^{n} \mathcal{K}^{(m)}$. A simplex σ is called a *maximal face* of a simplicial complex $\mathcal{K}$ if it is not a strict subset of any other simplex $\tau \in \mathcal{K}$.

The *underlying space* of $\mathcal{K}$ is the union of its simplices, denoted by $|\mathcal{K}|$. The underlying space is implicitly used whenever we refer to $\mathcal{K}$ as a topological space.

An *abstract simplicial complex* $\mathcal{S}$ is a collection of finite nonempty sets $A \in \mathcal{S}$ such that every nonempty subset of A is also contained in $\mathcal{S}$. The sets in $\mathcal{S}$ are called its *simplices*. For example, the vertex sets of the simplices in a geometric complex form an abstract simplicial complex, called its *vertex scheme*. If $\mathcal{K}$ is a geometric simplicial complex whose vertex scheme is isomorphic to an abstract simplicial complex $\mathcal{S}$, then $\mathcal{K}$ is a *geometric realization* of $\mathcal{S}$. It is unique up to simplicial isomorphism.

We will use the construction of a *pasting map* [31] to perform vertex and edge identifications on simplicial

complexes. Given a finite abstract simplicial complex $\mathcal{L}$, a *labelling* of the vertices of $\mathcal{L}$ is a surjective map $f : \mathcal{L}^{(0)} \to C$ where the set C is called the set of *vertex labels*. Then, the set $\{f(\sigma) \mid \sigma \in \mathcal{L}\}$ is an abstract simplicial complex. Let $\mathcal{K}$ be a geometric realization. Then f induces a simplicial quotient map $g : |\mathcal{L}| \to |\mathcal{K}|$, called the *pasting map* associated with f.

In particular, given an equivalence relation $\sim$ on the vertices of $\mathcal{L}$, the surjection sending each vertex to its equivalence class induces a pasting map. We will use this construction to perform identifications of simplices. For example, the wedge sum of a collection of *pointed simplicial complexes* (complexes with a distinguished vertex, called the *basepoint*) can be constructed this way using the equivalence relation identifying all basepoints.

2.2 Discrete Morse theory and Erasability Our focus in this paper is limited to simplicial complexes, and hence we restrict the discussion of Forman's discrete Morse theory to simplicial complexes. We refer to [11] for a compelling expository introduction.

A function f on a simplicial complex $\mathcal{K}$ is called a *discrete Morse function* if

- f is *monotonic*, i.e., $\sigma \subseteq \tau$ implies $f(\sigma) \leq f(\tau)$, and
- for all $t \in \mathrm{im}(f)$, $f^{-1}(t)$ is either a singleton $\{\sigma\}$ (in which case σ is a *critical simplex*) or a pair $\{\sigma, \tau\}$, where σ is a facet of τ (in which case (σ, τ) form a *gradient pair* and σ and τ are *regular simplices*).

Given a discrete Morse function f defined on complex $\mathcal{K}$, the *discrete gradient vector field* $\mathcal{V}$ of f is the collection of pairs of simplices (σ, τ), where (σ, τ) is in $\mathcal{V}$ if and only if σ is a facet of τ and $f(\sigma) = f(\tau)$.

Discrete gradient vector fields have a useful interpretation in terms of acyclic graphs obtained from matchings on Hasse diagrams, due to Chari [6]. Let $\mathcal{K}$ be a simplicial complex, let $H_{\mathcal{K}}$ be its Hasse diagram, and let M be a matching in the underlying undirected graph $H_{\mathcal{K}}$. Let $H_{\mathcal{K}}(M)$ be the directed graph obtained from $H_{\mathcal{K}}$ by reversing the direction of each edge of the matching M. Then M is a *Morse matching* if and only if $H_{\mathcal{K}}(M)$ is a directed acyclic graph. Every Morse matching M on the Hasse diagram $H_{\mathcal{K}}$ corresponds to a unique gradient vector field V_M on complex $\mathcal{K}$ and vice versa. For a Morse matching M, the unmatched vertices correspond to critical simplices of V_M, and the matched vertices correspond to the regular simplices of $\mathcal{V}_M$.

A non-maximal face $\sigma \in \mathcal{K}$ is said to be a *free face* if it is contained in a unique maximal face $\tau \in \mathcal{K}$. If $\dim \tau = \dim \sigma + 1$, we say that $\mathcal{K}' = \mathcal{K} \setminus \{\sigma, \tau\}$ arises from $\mathcal{K}$ by an *elementary collapse*, denoted by $\mathcal{K} \searrow^e \mathcal{K}'$. Furthermore, we say that $\mathcal{K}$ *collapses* to $\mathcal{L}$, denoted by $\mathcal{K} \searrow \mathcal{L}$, if there exists a sequence $\mathcal{K} = \mathcal{K}_1, \mathcal{K}_2, \ldots \mathcal{K}_n = \mathcal{L}$ such that $\mathcal{K}_i \searrow^e \mathcal{K}_{i+1}$ for all i. If $\mathcal{K} \searrow \mathcal{L}$, or more generally, if $\mathcal{K}$ and $\mathcal{L}$ are related through a sequence collapses and *expansions* (inverses of collapses), then the two complexes are *simple-homotopy equivalent type*. In particular, $\mathcal{K}$ and $\mathcal{L}$ are homotopy equivalent. Furthermore, if $\mathcal{K}$ collapses to a point, one says that $\mathcal{K}$ is collapsible and writes $\mathcal{K} \searrow 0$.

A simplicial collapse can be encoded by a discrete gradient.

THEOREM 2.1. (FORMAN [10], THEOREM 3.3) *Let $\mathcal{K}$ be a simplicial complex with a discrete gradient vector field $\mathcal{V}$, and let $\mathcal{L} \subseteq \mathcal{K}$ be a subcomplex. If $\mathcal{K} \setminus \mathcal{L}$ is a union of pairs in $\mathcal{V}$, then $\mathcal{K} \searrow \mathcal{L}$.*

In this case, we say that the collapse $\mathcal{K} \searrow \mathcal{L}$ is *induced by* the gradient $\mathcal{V}$. As a consequence of this theorem, we obtain:

THEOREM 2.2. (FORMAN [10], COROLLARY 3.5) *Let $\mathcal{K}$ be a simplicial complex with a discrete gradient vector field $\mathcal{V}$ and let m_d denote the number of critical simplices of $\mathcal{V}$ of dimension d. Then $\mathcal{K}$ is homotopy equivalent to a CW complex with exactly m_d cells of dimension d.*

In particular, a discrete gradient vector field on $\mathcal{K}$ with m_d critical simplices of dimension d gives rise to a chain complex having dimension m_d in each degree d, whose homology is isomorphic to that of $\mathcal{K}$. This condensed representation motivates the algorithmic search for (near-)optimal Morse matchings.

We will later use the following elementary lemma about gradient vector fields.

LEMMA 2.1. *Let $\mathcal{K}$ be a connected simplicial complex, let p be a vertex of $\mathcal{K}$, and let $\mathcal{V}_1$ be a discrete gradient on $\mathcal{K}$ with $m_0 > 1$ critical vertices and m critical simplices in total. Then there exists another gradient vector field $\widetilde{\mathcal{V}}$ on $\mathcal{K}$ with p as the only critical simplex of dimension 0 and $m - 2(m_0 - 1)$ critical simplices in total.*

Proof. Let $\mathcal{L}$ be the set of all the 1-simplices paired with 2-simplices in $\mathcal{V}$. Let $\mathcal{Q}$ be the 1-skeleton of $\mathcal{K}$. Then, by [20, Lemma 4.2], the 1-complex $\mathcal{Q} \setminus \mathcal{L}$ is connected, and one can compute a gradient vector field $\mathcal{V}_1$ on $\mathcal{Q} \setminus \mathcal{L}$ with p as the single critical vertex using depth first search starting from p (see, e.g., [35]). Let $\mathcal{W} \subset \mathcal{V}$ consist of all gradient pairs of $\mathcal{V}$ contained in $\mathcal{K} \setminus (\mathcal{Q} \setminus \mathcal{L}) = (\mathcal{K} \setminus \mathcal{Q}) \cup \mathcal{L}$. Note that, by construction, $\mathcal{W}$ does not contain any pairs of dimensions $(0, 1)$, while $\mathcal{V}_1$ has only such pairs. Since the gradient vector fields $\mathcal{V}_1$ and $\mathcal{W}$ are defined on disjoint sets of simplices, it follows that $\widetilde{\mathcal{V}} = \mathcal{V}_1 \cup \mathcal{W}$ is a gradient vector field with the desired property.

COROLLARY 2.1. *Given a collapsible simplicial complex $\mathcal{K}$ and an arbitrary vertex $p \in \mathcal{K}$, there exists a gradient vector field $\mathcal{V}$ on $\mathcal{K}$ with p as the unique critical simplex of $\mathcal{V}$.*

Borrowing and extending the terminology used in [9], we make the following definitions: A maximal face τ in a simplicial complex $\mathcal{K}$ is called an *internal simplex* if it has no free face. If a 2-complex $\mathcal{K}$ collapses to a 1-complex, we say that $\mathcal{K}$ is *erasable*. Moreover, for a 2-complex $\mathcal{K}$, the quantity $\mathrm{er}(\mathcal{K})$ is the minimum number of internal 2-simplices that need to be removed so that the resulting complex collapses to a 1-complex. Equivalently, it is the minimum number of critical 2-simplices of any discrete gradient on $\mathcal{K}$.

DEFINITION 2.1. (ERASABLE SUBCOMPLEX) *Given a 2-complex $\mathcal{K}$, we say that a subcomplex $\mathcal{L} \subseteq \mathcal{K}$ is an* erasable subcomplex of $\mathcal{K}$ (through the gradient $\mathcal{V}$) *if there exists another subcomplex $\mathcal{M} \subseteq \mathcal{K}$ with $\mathcal{K} \searrow \mathcal{M}$ (induced by the gradient $\mathcal{V}$) such that $\mathcal{K} \setminus \mathcal{M} \subseteq \mathcal{L}$ and $\mathcal{K}^{(2)} \setminus \mathcal{M}^{(2)} = \mathcal{L}^{(2)}$.*

DEFINITION 2.2. (EVENTUALLY FREE) *We say that a simplex σ is* eventually free (through the gradient $\mathcal{V}$) *in a complex $\mathcal{K}$ if there exists a subcomplex $\mathcal{L}$ of $\mathcal{K}$ such that $\mathcal{K} \searrow \mathcal{L}$ (induced by $\mathcal{V}$) and σ is free in $\mathcal{L}$. Equivalently, $\mathcal{K}$ collapses further to a subcomplex not containing σ.*

LEMMA 2.2. *If $\mathcal{L}_1, \mathcal{L}_2$ are erasable subcomplexes of a 2-complex $\mathcal{K}$, then so is their union.*

Proof. Let $\mathcal{V}_1$ be a discrete gradient erasing $\mathcal{L}_1$, and $\mathcal{V}_2$ a discrete gradient erasing $\mathcal{L}_2$. Without loss of generality, we may assume that both gradients have only pairs (σ, τ) of dimension $(1, 2)$, and that all such pairs are in $\mathcal{L}_1$ or $\mathcal{L}_2$, respectively, and σ is eventually free; removing all other pairs still yields an erasing gradient. Now consider the collapse $\mathcal{K} \searrow \mathcal{M}_1$ induced by $\mathcal{V}_1$ and the collapse $\mathcal{K} \searrow \mathcal{M}_2$ induced by $\mathcal{V}_2$. Restricting the gradient $\mathcal{V}_2$ to the subcomplex $\mathcal{M}_1$, we obtain a gradient $\mathcal{V}_{12} = \{(\sigma, \tau) \in \mathcal{V}_2 \mid \sigma, \tau \in \mathcal{M}_1\}$. By induction, each σ appearing in such a pair (σ, τ) is eventually free in $\mathcal{M}_1$, since any 2-simplex $\psi \in \mathcal{M}_1$ that is a coface of σ other than τ must appear in a pair $(\phi, \psi) \in \mathcal{V}_{12}$ by the definition of $\mathcal{V}_{12}$ and the assumption that ϕ is eventually free through the gradient $\mathcal{V}_2$. Thus, $\mathcal{K}$ collapses to a complex that contains no 2-simplices of either $\mathcal{L}_1$ or $\mathcal{L}_2$, as claimed.

LEMMA 2.3. *If $\mathcal{K}$ is an erasable complex, then any subcomplex $\mathcal{L} \subset \mathcal{K}$ is also erasable.*

Proof. Let $\mathcal{V}$ be a discrete gradient erasing $\mathcal{K}$. Restricting the gradient $\mathcal{V}$ to the subcomplex $\mathcal{L}$, we obtain a gradient $\mathcal{V}' = \{(\sigma, \tau) \in \mathcal{V} \mid \sigma, \tau \in \mathcal{L}\}$. Moreover, if τ is a 2-simplex in $\mathcal{L} \subset \mathcal{K}$, then τ occurs in a gradient pair $(\sigma, \tau) \in \mathcal{V}$, where some $\sigma \subset \tau$ is a 1-simplex in $\mathcal{K}$. Since $\mathcal{L}$ is a simplicial complex, σ is also a simplex in $\mathcal{L}$. Hence, $(\sigma, \tau) \in \mathcal{V}'$. Thus, every 2-simplex in $\mathcal{L}$ occurs in a gradient pair in $\mathcal{V}'$, meaning that $\mathcal{V}'$ is a discrete gradient erasing $\mathcal{L}$.

2.3 Approximation algorithms An *α-approximation algorithm* for an optimization problem is a polynomial-time algorithm that, for all instances of the problem, produces a solution whose value is within a factor α of the value of an optimal solution. The factor α is called the *approximation ratio* of the algorithm. An *approximation preserving reduction* is a procedure for transforming an optimization problem A to an optimization problem B, such that an α-approximation algorithm for B implies an $h(\alpha)$-approximation algorithm for A, for some function h. Then, if A is hard to approximate within factor $h(\alpha)$, the reduction implies that B is hard to approximate within factor α.

We will use a particular important and well-studied class of approximation preserving reductions, called *L-reductions*, which provide a simple and effective tool in proving hardness of approximability results [33, 39]. To give the definition, consider a maximization problem A with a non-negative integer valued objective function m_A. Given an instance x of A, the goal is to find a solution y (among a finite set of feasible solutions) maximizing the objective function $m_A(x, y)$. Define $\mathrm{OPT_A}(x)$ as the maximum value of the objective function on input x.

An *L-reduction* from one optimization problem A to another optimization problem B is a pair of functions f and g that are computable in polynomial time and satisfy the following conditions:

1. The function f maps instances of A to instances of B.
2. There is a positive constant μ such that, for all instances x of A,
$$\mathrm{OPT_B}(f(x)) \leq \mu \, \mathrm{OPT_A}(x).$$
3. The function g maps instances of A and solutions of B to solutions of A.
4. There is a positive constant ν such that, for any instance x of A and any solution y of $f(x)$, we have $\mathrm{OPT_A}(x) - m_A(x, g(x, y)) \leq \nu \left(\mathrm{OPT_B}(f(x)) - m_B(f(x), y)\right)$.

If $\mu = \nu = 1$, the reduction is *strict*.

We will use the following straightforward fact about L-reductions which appears as Theorem 16.5 in a book by Williamson and Shmoys [39].

THEOREM 2.3. *If there is an L-reduction with parameters μ and ν from a maximization problem A to another maximization problem B, and there is a $(1-\delta)$-approximation algorithm for B, then there is a $(1-\mu\nu\delta)$-approximation algorithm for A.*

2.4 Acyclic subgraphs We recall some concepts and problems from graph theory that will be used in our reductions. A directed graph G with vertex set V and edge set E is written as $G = (V, E)$. A directed graph is called an *oriented graph* if no pair of vertices is connected by an anti-parallel pair of edges. In other words, an oriented graph is a directed graph without 2-cycles or loops. Note that in contrast to a general directed graph, an oriented graph always has a simple *underlying undirected graph*, which is therefore a simplicial complex. We will be making use of this fact in §4.

The problem of finding the *maximum acyclic subgraph* (MAS) of a given directed graph $G = (V, E)$ consists of determining a maximum subset $E_{max} \subseteq E$ for which the subgraph $G_{max} = (V, E_{max})$ has no directed cycles. A *feedback arc set* is a set of edges whose removal leaves a directed acyclic graph. A *minimum feedback arc set* is a feedback arc set of minimum cardinality. The problem minFAS of finding such a set is thus complementary to MAS.

A *directed degree-3 graph* is a directed graph with total degree (indegree plus outdegree) at most 3. The restriction of the problem MAS to directed degree-3 graphs is denoted by 3MAS. Moreover, the problem MAS restricted to oriented graphs is denoted by OMAS, and the restriction to oriented degree-3 graphs is denoted by 3OMAS.

We will show that there is a L-reduction from MAS to OMAS, allowing us to consider only oriented graphs later.

THEOREM 2.4. *There is a strict reduction from MAS to OMAS, and from 3MAS to 3OMAS.*

Proof. The map f transforming an instance of MAS (a directed graph G) to an instance of OMAS (an oriented graph $f(G)$) is given by removing from G all loops and all pairs of anti-parallel edges. Furthermore, the map g transforming a solution of OMAS for the instance $f(G)$ (an acyclic subgraph A of $f(G)$) to a solution of MAS for the instance G (an acyclic subgraph $B = g(G, A)$ of G) is given as follows: Extend the acyclic graph A to a subgraph B of G by adding for each anti-parallel pair of edges in G one edge whose orientation is consistent with the partial order induced by A. By construction, the subgraph B is still acyclic.

By definition, $m_{\mathsf{OMAS}}(f(G), A)$ is the number of edges in A, and $m_{\mathsf{MAS}}(G, B)$ is the number of edges in B. Let k be the number of pairs of anti-parallel edges in G. Then,

$$m_{\mathsf{MAS}}(G, B) = m_{\mathsf{OMAS}}(f(G), A) + k.$$

In particular, if A is optimal, this yields

$$\begin{aligned}\mathrm{OPT}_{\mathsf{MAS}}(G) &\geq m_{\mathsf{MAS}}(G, B)\\ &= m_{\mathsf{OMAS}}(f(G), A) + k\\ &= \mathrm{OPT}_{\mathsf{OMAS}}(f(G)) + k.\end{aligned}$$

On the other hand, any acyclic subgraph C of G restricts to an acyclic subgraph C' of $f(G)$ by removing at most k edges. In particular, for an optimal C, we get

$$\begin{aligned}\mathrm{OPT}_{\mathsf{MAS}}(G) &= m_{\mathsf{MAS}}(G, C)\\ &\leq m_{\mathsf{OMAS}}(f(G), C') + k\\ &\leq \mathrm{OPT}_{\mathsf{OMAS}}(f(G)) + k.\end{aligned}$$

Thus, we have

$$\begin{aligned}m_{\mathsf{MAS}}(G, B) &= m_{\mathsf{OMAS}}(f(G), A) + k,\\ \mathrm{OPT}_{\mathsf{MAS}}(G) &= \mathrm{OPT}_{\mathsf{OMAS}}(f(G)) + k\end{aligned}$$

and thus

$$\mathrm{OPT}_{\mathsf{OMAS}}(f(G)) \leq \mathrm{OPT}_{\mathsf{MAS}}(G)$$

and

$$\begin{aligned}&\mathrm{OPT}_{\mathsf{MAS}}(G) - m_{\mathsf{MAS}}(G, B)\\ &= \mathrm{OPT}_{\mathsf{OMAS}}(f(G)) - m_{\mathsf{OMAS}}(f(G), A),\end{aligned}$$

establishing a strict reduction from MAS to OMAS. The same construction restricts to a strict reduction from 3MAS to 3OMAS.

We state a few known hardness of approximation results for MAS and related problems.

THEOREM 2.5. (NEWMAN [32], THEOREM 3) *It is NP-hard to approximate 3MAS to within $\left(1 - \frac{1}{126}\right) + \epsilon$ for any $\epsilon > 0$.*

Moreover, the following result establishes hardness with respect to the *unique games conjecture* (UGC) [21]. A problem is said to be *UGC-hard* (or *UG-hard*) if the unique games conjecture implies that the problem is NP-hard. We refer to [22] for a detailed account on this conjecture.

Furthermore, Guruswami et al. proved the following result, which appears as Theorem 1.1 of [12].

THEOREM 2.6. *Let $\delta \in (0, \frac{1}{2})$. If for any directed graph G with an acyclic subgraph consisting of a fraction $(1-\delta)$ of its edges, one can efficiently find an acyclic subgraph of G with more than $(\frac{1}{2}+\delta)$ of its edges, then the UGC is false. In particular, it is UGC-hard to approximate MAS within a factor of $\frac{1}{2}+\delta$ for any $\delta > 0$.*

By Theorem 2.4, the same is true for OMAS. Moreover, Newman [32] established an approximation preserving reduction from MAS to 3MAS, with the following consequence:

THEOREM 2.7. (NEWMAN [32], THEOREM 5) *For any constant $\epsilon > 0$, if there exists a $\left(\left(1-\frac{1}{18}\right)+\epsilon\right)$-approximation algorithm for 3MAS, then there exists a $\left(\frac{1}{2}+\delta\right)$-approximation algorithm for MAS for some constant $\delta > 0$.*

From Theorems 2.4, 2.5, 2.6, and 2.7, we conclude:

COROLLARY 2.2. *It is UGC-hard to approximate 3MAS and 3OMAS within a factor of $\left(1-\frac{1}{18}\right)+\epsilon$, and NP-hard to approximate 3MAS and 3OMAS to within $\left(1-\frac{1}{126}\right)+\epsilon$, for any $\epsilon > 0$.*

3 Hardness of Approximation of Min-Morse Matching

In this section, we work with abstract connected simplicial complexes. Recall that an abstract simplicial complex is connected if its 1-skeleton is connected as a graph.

DEFINITION 3.1. (AMPLIFIED COMPLEX) *Given a pointed simplicial complex $\mathcal{K}$ with n simplices and some integer $c > 0$, the amplified complex $\widehat{\mathcal{K}}_c$ is defined as the wedge sum of m copies of $\mathcal{K}$, with $m = n^{c-1}$.*

LEMMA 3.1. *Given a complex $\mathcal{K}$ of size n and integer c, consider the amplified complex $\widehat{\mathcal{K}}_c$ of $\mathcal{K}$. Let $\widehat{\mathcal{V}}$ be a gradient vector field on $\widehat{\mathcal{K}}_c$. Then*

(i) $\widehat{\mathcal{K}}_c$ is collapsible if and only if $\mathcal{K}$ is collapsible.

(ii) If $\mathcal{K}$ is not collapsible, then $\widehat{\mathcal{V}}$ has more than n^{c-1} critical simplices.

Proof. Suppose that the complex $\mathcal{K}$ is collapsible. Then there exists a gradient vector field $\mathcal{V}_{\mathcal{K}}$ with a unique critical simplex $q \in \mathcal{K}$. Let p be an arbitrarily chosen distinguished vertex of $\mathcal{K}$ that will be used to construct amplified complex $\widehat{\mathcal{K}}_c$. Using Corollary 2.1, without loss of generality the vector field $\mathcal{V}_{\mathcal{K}}$ has p as its unique critical simplex. Now the gradient vector field on $\widehat{\mathcal{K}}_c$, say $\widehat{\mathcal{V}}$, is simply the gradient vector field $\mathcal{V}_{\mathcal{K}}$ repeated on each identical copy of $\mathcal{K}$. Since p is the unique critical simplex of $\widehat{\mathcal{V}}$ on $\widehat{\mathcal{K}}_c$, we conclude that $\widehat{\mathcal{K}}_c$ is collapsible.

Conversely, suppose that the complex $\widehat{\mathcal{K}}_c$ is collapsible. Then, by Corollary 2.1, we can obtain a gradient vector field $\widehat{\mathcal{V}}$ on $\widehat{\mathcal{K}}_c$ with the distinguished vertex p as its unique critical simplex. If we consider the gradient vector field $\widehat{\mathcal{V}}$ restricted to any one of the copies of $\widehat{\mathcal{K}}_c$, it follows immediately that $\mathcal{K}$ is collapsible.

Now suppose that $\mathcal{K}$ is not collapsible and $\widehat{\mathcal{V}}$ has less than or equal to n^{c-1} critical simplices. By Lemma 2.1, without loss of generality we may assume that $\widehat{\mathcal{V}}$ has the distinguished vertex p as the unique critical vertex. Now consider $\widehat{\mathcal{V}}$ restricted to each of the individual copies of $\mathcal{K}$. Then clearly at least one of the copies has p as its unique critical simplex (else we would have more than n^{c-1} critical simplices in total). But this immediately implies that $\mathcal{K}$ is collapsible, a contradiction. Hence, if $\mathcal{K}$ is not collapsible, then $\widehat{\mathcal{V}}$ has more than n^{c-1} critical simplices.

PROPOSITION 3.1. *For any $\epsilon \in (0, 1]$, if there exists an $O(n^{1-\epsilon})$-factor approximation algorithm for MinMM, where n denotes the number of simplices of an input simplicial complex, then there exists a polynomial time algorithm for deciding collapsibility of simplicial complexes.*

Proof. Given any $\epsilon \in (0, 1]$, suppose there exists an $O(n^{1-\epsilon})$-factor approximation algorithm for MinMM. Specifically, there exist $p, M > 0$ such that for all $n \geq M$, the approximation ratio is bounded above by $pn^{1-\epsilon}$. Now choose c to be the smallest positive integer with the property $\frac{1}{c} < \epsilon$, i.e., $c = \left\lfloor \frac{1}{\epsilon} + 1 \right\rfloor$. Consider an arbitrary connected complex $\mathcal{K}$ with n simplices and construct the amplified complex $\widehat{\mathcal{K}}_c$. Note that the total number of simplices in $\widehat{\mathcal{K}}_c$ is $\hat{n} = (n-1)\, n^{c-1} + 1 = n^c - n^{c-1} + 1$. Also, if $n < \max\{M, p^{\frac{1}{1+c\epsilon}}\}$, i.e., if n is bounded by a constant, the collapsibility of $\mathcal{K}$ can easily be checked in constant time. So, without loss of generality, we assume that $n \geq \max\{M, p^{\frac{1}{1+c\epsilon}}\}$.

We now use the following Algorithm B to decide collapsibility of the complex $\mathcal{K}$. We execute the $O(n^{1-\epsilon})$-factor approximation Algorithm A for MinMM on the amplified complex $\widehat{\mathcal{K}}_c$. If the number C_A of critical simplices returned by Algorithm A is less than n^{c-1}, we report that the complex $\mathcal{K}$ is collapsible, else we declare that $\mathcal{K}$ is not collapsible.

When the complex is collapsible, the number C_A of critical simplices returned by Algorithm A can be bounded as follows:

$$C_A \leq p\hat{n}^{1-\epsilon} < p\,(n^c)^{1-\epsilon} \leq n^{c-1}.$$

The bound $C_A < n^{c-1}$ for a collapsible complex $\mathcal{K}$ with n simplices, along with part (ii) of Lemma 3.1, estab-

lishes the correctness of Algorithm B for determining collapsibility of the complex $\mathcal{K}$. Also, since $\hat{n} < n^c$, Algorithm B runs in time polynomial in n.

Recently, Tancer [37] proved the following theorem about collapsibility of 3-complexes.

THEOREM 3.1. (TANCER [37], THEOREM 1) *It is NP-complete to decide whether a given 3-complex is collapsible.*

COROLLARY 3.1. *For any $\epsilon \in (0, 1]$, there exists no $O(n^{1-\epsilon})$-factor approximation algorithm for MinMM, where n denotes the number of simplices of an input simplicial complex, unless $P = NP$.*

4 Hardness of Approximation for Max-Morse Matching

In this section, we describe an L-reduction from 3OMAS to MaxMM, establishing hardness of approximation for MaxMM. Our construction is based on a modification of Zeeman's *dunce hat* [40]. The dunce hat is a simplicial complex which is contractible but has no free faces and is therefore not collapsible. In contrast, the *modified dunce hat* is collapsible but only through a single free face. The triangulation is given in Fig. 1. An equivalent triangulation has been described by Hachimori [13, p. 108]. Its number of simplices is minimal among all complexes that are collapsible through a single free face, as can be verified by an exhaustive search [34]. The 1-simplex ω is the unique free face of the modified dunce hat $\mathcal{D}$.

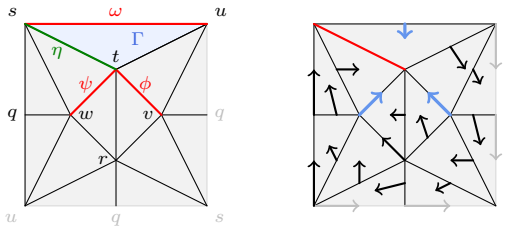


Figure 1: The modified dunce hat $\mathcal{D}$. Left: triangulation, with certain distinguished simplices highlighted. Note that the vertices s, q, u appearing multiple times are identified accordingly. The complex is collapsible through the unique free face ω. Right: a discrete gradient $\mathcal{V}_\mathcal{D}$ on $\mathcal{D}$ that leaves only the vertices s, t and the edge η critical. The three highlighted arrows (blue, thick) correspond to pairs in $\mathcal{V}_\mathcal{D}$ that will in some cases be discarded when assembling a gradient on the entire complex $\mathcal{K}(G)$.

For the remainder of the section, we will use the following notation: For a graph $G = (V, E)$, the indegree of a vertex v, in G, is denoted $\deg^-_G(v)$, and its outdegree is denoted $\deg^+_G(v)$. We now construct a complex $\mathcal{K}(G)$ from G, and more generally, a complex $\mathcal{K}(G, H)$ for any subgraph $H = (V_H, E_H)$ of G. Throughout this section, the graph G denotes an oriented connected degree-3 graph. Note that the connectedness assumption is not a restriction for the 3OMAS problem.

In order to aid the reader's intuition, we first outline the motivation behind some of the design choices for the gadget, before giving the formal description. Our aim is to construct a complex $\mathcal{K}(G)$ so that there is a one-to-one correspondence between the cardinality of a (minimum) feedback arc set (Proposition 4.1) and the (minimum) number of simplices that need to be removed from $\mathcal{K}(G)$ to make it erasable. To this end, we start with a disjoint union of copies of $\mathcal{D}$, one for each edge $e \in E_H$, and make identifications of vertices to obtain a complex $\mathcal{K}(G)$ which is homotopy equivalent to the undirected graph underlying G. However, we make additional identifications of edges to ensure that any subcomplex $\mathcal{K}(G, H)$, corresponding to a subgraph H, collapses to a 1-dimensional complex isomorphic to the undirected graph underlying H if and only if H is acyclic. In order to obtain an L-reduction, we ensure that the size of the complex $\mathcal{K}(G)$ is linear in size of graph G.

We now describe the construction of the complexes $\mathcal{K}(G, H)$ and $\mathcal{K}(G)$ for an oriented degree-3 graph G, by classifying the vertices of H into five types.

1. Consider an arbitrary total order $\prec$ on the edge set E of G.

2. Start with a disjoint union of copies of $\mathcal{D}$, one for each edge $e \in E_H$, denoted by $\mathcal{D}_e$.

3. Using the pasting map construction as defined in §2.1, construct the complex $\mathcal{K}(G, H)$ by identifying some of the distinguished simplices of each gadget $\mathcal{D}_e$ based on the following rules, as applied to each vertex $v \in V_H$ according to its indegree and outdegree. Note that vertex and edge labels are as in §1 with a subscript added identifying the respective copy of $\mathcal{D}$.

   (a) For every vertex $v \in V_H$ with $\deg^-_H(v) = \deg^+_H(v) = 0$, no identifications are made.

   (b) For every vertex $v \in V_H$ with $\deg^-_H(v) = 0$ and $\deg^+_H(v) > 0$, identify all vertices s_e for every outgoing edge $e \in E_H$.

   (c) For every vertex $v \in V_H$ with $\deg^+_H(v) = 0$ and $\deg^-_H(v) > 0$, identify all vertices t_e for every incoming edge $e \in E_H$.

(d) For every vertex $v \in V_H$ with $\deg_G^-(v) = 1$, $\deg_G^+(v) = 2$, $\deg_H^-(v) > 0$ and $\deg_H^+(v) > 0$, let k and l denote the outgoing edges of v in G, with $k \prec l$, and let j denote the incoming edge of v in G.
   i. If $j, k \in E_H$, identify the 1-simplices $\phi_j \sim \omega_k$ such that the incident vertices are identified as $u_k \sim v_j$ and $s_k \sim t_j$.
   ii. Similarly, if $j, l \in E_H$, identify $\psi_j \sim \omega_l$ such that $u_l \sim w_j$ and $s_l \sim t_j$.

(e) For every vertex $v \in V_H$ with $\deg_G^-(v) \in \{1, 2\}$, $\deg_G^+(v) = 1$, $\deg_H^-(v) > 0$ and $\deg_H^+(v) > 0$, let k (and possibly l, with $k \prec l$) denote the incoming edges of v in G, and let j denote the outgoing edge of v in G.
   i. If $j, k \in E_H$, identify the 1-simplices $\omega_j \sim \phi_k$ such that $u_j \sim v_k$ and $s_j \sim t_k$.
   ii. (Similarly, if $\deg_H^-(v) = 2$ and $j, l \in E_H$, identify $\omega_j \sim \phi_l$ such that $u_j \sim v_l$ and $s_j \sim t_l$.)

4. Furthermore, we define $\mathcal{K}(G) = \mathcal{K}(G, G)$.

REMARK 4.1. *We choose an arbitrary linear order since there is no natural choice to determine some of the attachments in the construction. Using an (arbitrarily chosen) linear order on the edge set E of G allows us to make the construction of complexes $\mathcal{K}(G, H)$ explicit and concrete. While it is clear that different linear orders on the edge set E may result in different complexes, the hardness results in this section do not depend on the choice of the linear order $\prec$.*

REMARK 4.2. *Since G is an oriented graph, it is easy to verify that $\mathcal{K}(G, H)$ is a simplicial complex. Also, by construction, $\mathcal{K}(G, H)$ is a subcomplex of $\mathcal{K}(G)$ whenever H is a subgraph of G.*

REMARK 4.3. *From Fig. 1, we observe that for a modified dunce hat $\mathcal{D}_e$, ψ_e is incident on exactly two 2-simplices of $\mathcal{D}_e$. The same holds true for ϕ_e. Also, note that ω_e is incident on a unique 2-simplex of $\mathcal{D}_e$, namely Γ_e.*

4.1 Structural properties of the reduction

LEMMA 4.1. *For a subgraph $H = (V_H, E_H)$ of a directed degree-3 graph G and an edge $e \in E_H$:*

(i) If ω_e is eventually free in $\mathcal{K}(G, H)$, then $\mathcal{D}_e$ is erasable in $\mathcal{K}(G, H)$.

(ii) If $\mathcal{D}_e$ is erasable in $\mathcal{K}(G, H)$ through a gradient $\mathcal{V}$, then (ω_e, Γ_e) is a gradient pair in $\mathcal{V}$. If f is a discrete Morse function with gradient $\mathcal{V}$, then for any simplex $\sigma \in \mathcal{D}_e$ such that $\sigma \notin \{\omega_e, \Gamma_e\}$ we have $f(\omega_e) > f(\sigma)$.

Proof. Suppose ω_e is eventually free in $\mathcal{K}(G, H)$. Then there exists a subcomplex $\mathcal{L}$ of $\mathcal{K}(G, H)$ such that $\mathcal{K}(G, H) \searrow \mathcal{L}$ and ω_e is free in $\mathcal{L}$. Note that, by construction of $\mathcal{D}$, this implies that $\mathcal{D}_e$ is a subcomplex of $\mathcal{L}$. Now using the gradient specified in Fig. 1 all the 2-simplices of $\mathcal{D}_e$ can be collapsed, making $\mathcal{D}_e$ erasable in $\mathcal{K}(G, H)$. This proves the first statement of the lemma. The second statement of the lemma immediately follows from observing that ω_e is the unique free 1-simplex in complex $\mathcal{D}_e$, Γ_e is the unique coface incident on ω_e, and $\mathcal{D}_e$ is erasable in $\mathcal{K}(G, H)$ through the gradient $\mathcal{V}$ of f.

LEMMA 4.2. *For a subgraph H of a directed degree-3 graph G and a vertex $v \in V_H$ with $\deg_H^+(v) = l > 0$ and outgoing edges $\{d_1, \ldots, d_l\} \in E_H$, we have:*

(i) If $\deg_H^-(v) = 0$, then each ω_{d_j} is free in $\mathcal{K}(G, H)$.

(ii) If $\deg_H^-(v) = k > 0$, let $\{e_1, \ldots, e_k\} \in E_H$ be the set of incoming edges of v. If there is a gradient $\mathcal{V}$ such that each ω_{e_i} is eventually free in $\mathcal{K}(G, H)$ through $\mathcal{V}$, then each ω_{d_j} is eventually free in $\mathcal{K}(G, H)$ through $\mathcal{V}$ as well.

Proof. If $\deg_H^-(v) = 0$, then each ω_{d_j} is free by construction of $\mathcal{K}(G, H)$. Now suppose that $\deg_H^-(v) = k > 0$ and each ω_{e_i} is eventually free in $\mathcal{K}(G, H)$. From Remark 4.3 and from the construction of the complex $\mathcal{K}(G, H)$, it can be deduced that for any edge d_j, the only 2-simplices incident on ω_{d_j} in the complex $\mathcal{K}(G, H)$ are Γ_{d_j} and one pair of 2-simplices of $\mathcal{D}_{e_i}$ for each e_i. By assumption, each ω_{e_i} is eventually free, and so by part (i) of Lemma 4.1, each $\mathcal{D}_{e_i}$ is erasable. Hence, by Lemma 2.2, their union is erasable too. This means that $\mathcal{K}(G, H)$ collapses to a complex $\mathcal{M}$ in which each ω_{d_j} is free, proving the claim.

LEMMA 4.3. *A subgraph H of a directed degree-3 graph G is acyclic if and only if the corresponding complex $\mathcal{K}(G, H)$ is erasable.*

Proof. Suppose that the given subgraph H of G, with n vertices , is acyclic. Consider an arbitrary total order on the vertices of H consistent with the partial order induced by the edges in H, and index the vertices $\{v_1, v_2, \ldots, v_n\}$ according to this total order. We can now apply Lemma 4.2 and part (i) of Lemma 4.1 inductively for all v_i from v_1 to v_{n-1} to establish the erasability of $\mathcal{D}_{d_j}$ for each of the outgoing edges d_j of v_i in H. Hence, the entire complex $\mathcal{K}(G, H)$ is erasable.

To show the reverse implication, we prove that if H has directed cycles, then $\mathcal{K}(G,H)$ is not erasable. Assume for a contradiction that $\mathcal{K}(G,H)$ is erasable through a gradient $\mathcal{V}$, and let f be a discrete Morse function with that gradient. Let a, b be two consecutive edges in a directed cycle of $\mathcal{K}(G,H)$. Then, by construction of $\mathcal{K}(G,H)$, either $\phi_a \sim \omega_b$ or $\psi_a \sim \omega_b$, and so by part (ii) of Lemma 4.1 we have $f(\omega_a) > f(\omega_b)$. Applying this argument to each pair of consecutive edges in the cycle yields a contradiction. Hence, if H has directed cycles, then $\mathcal{K}(G,H)$ is not erasable.

LEMMA 4.4. *For any edge $e \in E(G)$, the subcomplex $\mathcal{D}_e \setminus \{\Gamma_e\}$ is erasable in $\mathcal{K}(G)$.*

Proof. Consider the discrete gradient specified in Fig. 1 as a gradient $\mathcal{V}_e$ on $\mathcal{D}_e \subseteq \mathcal{K}(G)$. First note that $\mathcal{D}_e \setminus \{\Gamma_e\}$ is erasable in $\mathcal{D}_e$ through the gradient $\mathcal{V}_e \setminus \{(\omega_e, \Gamma_e)\}$. Moreover, all 1-simplices of $\mathcal{D}_e$ that are paired in $\mathcal{V}_e$ with a 2-simplex do not appear in $\mathcal{D}_c$ for any edge $c \neq e$. It follows that $\mathcal{D}_e \setminus \{\Gamma_e\}$ is erasable in $\mathcal{K}(G)$.

LEMMA 4.5. *Let C be a set of 2-simplices such that $\mathcal{K}(G) \setminus C$ is erasable, and let $F = \{f \in E \mid C \cap \mathcal{D}_f \neq \emptyset\}$. Then F is a feedback arc set of G.*

Proof. Each 2-simplex $\sigma \in C$ lies in $\mathcal{D}_f$ for a unique $f \in E$, which implies $f \in F$. In particular, $\sigma \in C$ implies $\sigma \notin \mathcal{D}_e$ for any $e \in E \setminus F$. Now consider the subgraph $H = (V, E \setminus F)$ of G. Then, using Lemma 2.3, $\mathcal{K}(G,H) \subseteq \mathcal{K}(G) \setminus C$ is also erasable. Hence, by Lemma 4.3, H is acyclic, i.e., F is a feedback arc set of G.

PROPOSITION 4.1. *Given an oriented degree-3 graph G and the corresponding complex $\mathcal{K}(G)$, $\mathrm{er}(\mathcal{K}(G)) = \mathrm{OPT}_{\mathit{minFAS}}(G)$.*

Proof. Given a graph $G = (V, E)$, let $F \subseteq E$ be a minimum feedback arc set of G, and let $H = (V, E \setminus F)$ be the corresponding maximum acyclic subgraph. We construct a new complex $\mathcal{K}'$ from $\mathcal{K}(G)$ as follows: For every $f \in F$, we remove $\{\Gamma_f\}$ from $\mathcal{D}_f \subseteq \mathcal{K}(G)$ and erase $\mathcal{D}_f \setminus \{\Gamma_f\}$ in $\mathcal{K}(G)$ using Lemma 4.4. Note that $(\mathcal{K}(G) \setminus \{\Gamma_f \mid f \in F\}) \searrow \mathcal{K}'$. In order to show that $\mathcal{K}'$ is erasable, it suffices to show erasability of the pure subcomplex of $\mathcal{K}'$ induced by its 2-simplices. It is easy to check that the subcomplex of $\mathcal{K}'$ induced by the 2-simplices in $\mathcal{K}'$ is precisely $\mathcal{K}(G,H)$. However, from Lemma 4.3, we can deduce that $\mathcal{K}(G,H)$ is erasable. This implies that $\mathcal{K}'$ is erasable, and hence $(\mathcal{K}(G) \setminus \{\Gamma_f \mid f \in F\})$ is erasable. Since the total number of 2-simplices that were removed to erase $\mathcal{K}(G)$ is equal to $|F| = \mathrm{OPT}_{\mathsf{minFAS}}(G)$, we have established that $\mathrm{er}(\mathcal{K}(G)) \leq \mathrm{OPT}_{\mathsf{minFAS}}(G)$.

Now assume for a contradiction that $\mathrm{er}(\mathcal{K}(G)) < \mathrm{OPT}_{\mathsf{minFAS}}(G)$. Let C be a minimal set of 2-simplices that need to be removed to erase $\mathcal{K}(G)$, i.e., $|C| = \mathrm{er}(\mathcal{K}(G))$. Let $F' = \{f \in E \mid C \cap \mathcal{D}_f \neq \emptyset\}$. By Lemma 4.5, the graph $(V, E \setminus F')$ is acyclic and F' is a feedback arc set. Since each 2-simplex lies in $\mathcal{D}_e$ for some unique $e \in E$, it follows that $|F'| \leq |C|$. We conclude that $|F'| < \mathrm{OPT}_{\mathsf{minFAS}}(G)$, which contradicts the minimality of $\mathrm{OPT}_{\mathsf{minFAS}}(G)$. Hence, the claim follows.

In order to relate the homotopy type of G with that of $\mathcal{K}(G)$, we construct a new complex $\widetilde{\mathcal{K}}(G)$ as follows:

1. Start with a disjoint union of copies of $\mathcal{D}$, one for each edge in G, denoted by $\mathcal{D}_e$.
2. Similar to the construction of $\mathcal{K}(G)$, the complex $\widetilde{\mathcal{K}}(G)$ is constructed by identifying some of the distinguished vertices of each gadget $\mathcal{D}_e$ based on the following rules, as applied to each vertex of G based on its indegree and outdegree:
   (a) For every vertex of G that has incoming as well as outgoing edges, identify s_j with t_i for every incoming edge i and outgoing edge j of G.
   (b) For every vertex of G that has only incoming edges, identify all vertices t_e for every incoming edge e.
   (c) For ever vertex of G that has only outgoing edges, identify all vertices s_e for every outgoing edge e.

LEMMA 4.6. *$\widetilde{\mathcal{K}}(G)$ is homotopy equivalent to $\mathcal{K}(G)$.*

Proof. Comparing the two constructions, first note that $\mathcal{K}(G)$ can be obtained from $\widetilde{\mathcal{K}}(G)$ by further identifying certain 1-simplices $\phi_e \sim \omega_f$ or $\psi_e \sim \omega_f$ (together with vertices $v_e \sim u_f$ or $w_e \sim u_f$) in subcomplexes $\mathcal{D}_e, \mathcal{D}_f \subseteq \widetilde{\mathcal{K}}(G)$, where these two 1-simplices already have a common vertex by the identification $t_e \sim s_f$ in the construction of $\widetilde{\mathcal{K}}(G)$, and are otherwise not connected by another 1-simplex. In both $\widetilde{\mathcal{K}}(G)$ and $\mathcal{K}(G)$, the union of the two 1-simplices is contractible, and so each complex is homotopy equivalent to the space that further identifies each such pair of 1-simplices to a single point [16, Proposition 0.17]. The claim follows.

LEMMA 4.7. *$\widetilde{\mathcal{K}}(G)$ collapses to the undirected graph underlying G.*

Proof. First note that ω_e is free in $\widetilde{\mathcal{K}}(G)$ for each e in G. Moreover, the only simplices that are possibly shared by

gadgets $\mathcal{D}_i$ and $\mathcal{D}_j$ for $i \neq j$ are the vertices s and t of $\mathcal{D}_i$ and $\mathcal{D}_j$. Therefore, we can use the gradient vector field depicted in Figure 1 to collapse each gadget $\mathcal{D}_e$ to the 1-simplex η_e together with the vertices s_e and t_e. Thus, $\widetilde{\mathcal{K}}(G)$ collapses to the subcomplex $\mathcal{Q}$ of $\widetilde{\mathcal{K}}(G)$ induced by these 1-simplices. By construction of $\widetilde{\mathcal{K}}(G)$, this complex $\mathcal{Q}$ is isomorphic to the undirected graph underlying G.

COROLLARY 4.1. *$\mathcal{K}(G)$ is homotopy equivalent to the undirected graph underlying G.*

4.2 Inapproximability results Given an oriented connected degree-3 graph $G = (V, E)$ and the corresponding complex $\mathcal{K}(G)$, let $\mathrm{OPT}_{\mathsf{3OMAS}}(G)$ denote the optimal value of the 3OMAS problem on G, and let $\mathrm{OPT}_{\mathsf{MaxMM}}(\mathcal{K}(G))$ denote the optimal value of the MaxMM problem on $\mathcal{K}(G)$.

We now describe an L-reduction from 3OMAS to MaxMM. The map $\mathcal{K} : G \mapsto \mathcal{K}(G)$ transforms instances of 3OMAS (directed graphs) to instances of MaxMM (simplicial complexes). The map A that transforms solutions of MaxMM (discrete gradients $\mathcal{V}$ on $\mathcal{K}(G)$) to solutions of 3OMAS (acyclic subgraphs $A(G, \mathcal{V})$) is defined as follows: Let

$$F = \{f \in E \mid \exists \sigma \in \mathcal{D}_f : \sigma \text{ is a critical 2-simplex in } \mathcal{V}\}.$$

By Lemma 4.5, F is a feedback arc set. The corresponding solution $A(G, \mathcal{V})$ for 3OMAS is then simply the subgraph of G with edges $E \setminus F$. The value of the objective function $m_{\mathsf{MaxMM}}(\mathcal{K}(G), \mathcal{V})$ is the number of regular simplices in $\mathcal{V}$; the value of the objective function $m_{\mathsf{3OMAS}}(G, A(G, \mathcal{V}))$ is the number of edges of the acyclic subgraph, $|E \setminus F|$.

For a discrete gradient $\mathcal{V}$ on $\mathcal{K}(G)$, let n denote the number of simplices in $\mathcal{K}(G)$, let m denote the total number of critical simplices in $\mathcal{V}$, and let m_d denote the number of critical simplices in dimension d. Also, let β_d denote the Betti number of $\mathcal{K}(G)$ in dimension d, and let β be the sum of all Betti numbers.

LEMMA 4.8. *With the above notation,*

$$m_{\mathit{MaxMM}}(\mathcal{K}(G), \mathcal{V}) \leq n - 2\,m_2 - \beta$$

and

$$m_{\mathit{3OMAS}}(G, A(G, \mathcal{V})) \geq |E| - m_2.$$

Proof. By the Morse inequalities [10, Theorem 3.7], we have $m_0 \geq \beta_0$ and

$$m_2 - m_1 + m_0 = \beta_2 - \beta_1 + \beta_0.$$

From Corollary 4.1, we have $\beta_0 = 1$ and $\beta_2 = 0$. This gives us:

$$m_1 = \beta_1 + m_2 + (m_0 - \beta_0). \tag{4.1}$$

Moreover,

$$\begin{aligned} m &= m_2 + m_1 + m_0 && \text{(by definition)} \\ &= 2\,m_2 + \beta_1 + 2\,m_0 - \beta_0 && \text{(from Equation (4.1))} \\ &= 2\,m_2 + \beta + 2\,(m_0 - \beta_0) && \text{(since } \beta_2 = 0\text{)} \\ &\geq 2\,m_2 + \beta && \text{(since } m_0 \geq \beta_0\text{)}. \end{aligned}$$

Hence, $m_{\mathsf{MaxMM}}(\mathcal{K}(G), \mathcal{V}) = n - m \leq n - 2\,m_2 - \beta$.

In the construction of the acyclic subgraph $A(G, \mathcal{V})$, for every critical 2-simplex in $\mathcal{V}$, we remove at most one edge in G. Hence, we conclude that $m_{\mathsf{3OMAS}}(G, A(G, \mathcal{V})) \geq |E| - m_2$.

LEMMA 4.9. *Given a graph G and the corresponding complex $\mathcal{K}(G)$, we have $\mathrm{OPT}_{\mathit{MaxMM}}(\mathcal{K}(G)) = n - 2\,\mathrm{er}(\mathcal{K}(G)) - \beta = n - 2\,\mathrm{OPT}_{\mathit{minFAS}}(G) - \beta$.*

Proof. First note that for an optimal gradient vector field on $\mathcal{K}(G)$, we have $m_0 = 1$ and $m_2 = \mathrm{er}(\mathcal{K}(G))$. The first equality now follows by observing that in the proof of Lemma 4.8, equality $m_{\mathsf{MaxMM}}(\mathcal{K}(G), \mathcal{V}) = n - 2\,m_2 - \beta$ is obtained for $\beta_0 = m_0$. The second equality immediately follows from Proposition 4.1.

LEMMA 4.10. $\mathrm{OPT}_{\mathit{MaxMM}}(\mathcal{K}(G)) \leq 78\,\mathrm{OPT}_{\mathit{3OMAS}}(G)$.

Proof. For the Max-Acyclic Subgraph problem, from the trivial $\frac{1}{2}$-factor approximation algorithm mentioned in [38, Ch. 1], one knows that it is always possible to find an acyclic subgraph A_H of a directed graph H that contains at least half the number of edges in H. Clearly, this bound continues to hold when the class of graphs is restricted to degree-3 oriented graphs. This gives the following inequality:

$$\mathrm{OPT}_{\mathsf{3OMAS}}(G) \geq \frac{|E|}{2}. \tag{4.2}$$

First note that the number of simplices in the modified dunce hat $\mathcal{D}$ is $7 + 19 + 13 = 39$. The complex $\mathcal{K}(G)$ described in §4 is constructed from a disjoint union of $|E|$ copies of $\mathcal{D}$ with several simplices identified, giving us

$$n \leq 39\ |E|\,. \tag{4.3}$$

From Lemma 4.9 and Eqs. (4.2) and (4.3) we obtain the bound

$$\mathrm{OPT}_{\mathsf{MaxMM}}(\mathcal{K}(G)) \leq n \leq 39\ |E| \leq 78\,\mathrm{OPT}_{\mathsf{3OMAS}}(G).$$

LEMMA 4.11. $\mathrm{OPT}_{\mathit{3OMAS}}(G) - m_{\mathit{3OMAS}}(G, A(G, \mathcal{V})) \leq \frac{1}{2}\left(\mathrm{OPT}_{\mathit{MaxMM}}(\mathcal{K}(G)) - m_{\mathit{MaxMM}}(\mathcal{K}(G), \mathcal{V})\right)$.

Proof. By definition, $\mathrm{OPT}_{\mathsf{3OMAS}}(G) = E - \mathrm{OPT}_{\mathsf{minFAS}}(G)$. By Lemma 4.8,

$$m_{\mathsf{3OMAS}}(G, A(G, \mathcal{V})) \geq |E| - m_2.$$

Hence,

$$(\mathrm{OPT}_{\mathsf{3OMAS}}(G) - m_{\mathsf{3OMAS}}(G, A(G,\mathcal{V}))) \leq m_2 - \mathrm{OPT}_{\mathsf{minFAS}}(G). \tag{4.4}$$

Using Lemma 4.8 and Lemma 4.9, we obtain

$$(\mathrm{OPT}_{\mathsf{MaxMM}}(\mathcal{K}(G)) - m_{\mathsf{MaxMM}}(\mathcal{K}(G),\mathcal{V})) \geq 2\,(m_2 - \mathrm{OPT}_{\mathsf{minFAS}}(G)). \tag{4.5}$$

Substituting Equation 4.4 in Equation 4.5, we obtain the lemma.

THEOREM 4.1. *It is NP-hard to approximate* MaxMM *within a factor of* $\left(1 - \frac{1}{4914}\right) + \epsilon$ *and UGC-hard to approximate it within a factor of* $\left(1 - \frac{1}{702}\right) + \epsilon$*, for any* $\epsilon > 0$.

Proof. From Lemma 4.10 and Lemma 4.11 we conclude that the reduction from 3OMAS to MaxMM is an L-reduction with parameters $\mu = 78$ and $\nu = \frac{1}{2}$. By Theorem 2.3, if there exists a $\left(1 - \frac{1}{\mu\nu}\delta + \epsilon\right)$-approximation algorithm for MaxMM, then there exists a $(1 - \delta + \mu\nu\epsilon)$-algorithm for 3OMAS. Using Corollary 2.2, we choose $\delta = \frac{1}{126}$ to deduce that it is NP-hard to approximate MaxMM within a factor of $\left(1 - \frac{1}{4914}\right) + \epsilon$, and choose $\delta = \frac{1}{18}$ to deduce that it is UGC-hard to approximate MaxMM within a factor of $\left(1 - \frac{1}{702}\right) + \epsilon$.

5 Conclusion & Discussion

In this paper, we provide the first hardness of approximation results for the maximization and the minimization variants of the Morse matching problems.

While we established a hardness result for Min-Morse Matching on simplicial complexes of dimension $d \geq 3$, the question of hardness of approximation for Min-Morse matching for 2-dimensional simplicial complexes remains open. We will address this question in future work.

For the Max-Morse Matching problem on d-dimensional simplicial complexes, although our work clears a major hurdle of going beyond NP-hardness, a gap remains between the best approximability and inapproximability bounds. The best known approximation algorithm for Max-morse matching on simplicial complexes yields an approximation ratio of $\frac{d+1}{d^2+d+1}$ [35]. We believe that our result and techniques will pave way for further work in improving the gap, and in placing Max-Morse Matching in the right kind of approximation-algorithms related complexity class. In particular, devising an approximation algorithm for Max-Morse Matching with an approximation factor independent of the dimension of the complex, or establishing a hardness of approximation result for Max-Morse Matching that is dependent on the dimension is a challenging open problem.

We close the discussion with some additional open problems. Note that the complex $\mathcal{K}(G)$ employed in the hardness result for Max-Morse Matching described in §4 is not a manifold. Hence, the question of hardness of approximation for Max-Morse Matching on simplicial manifolds is open. The best known approximation algorithm for Max-Morse Matching on d-dimensional simplicial manifolds has approximation factor $\frac{2}{d}$ [35].

Finally, we also leave it as an open question to investigate sharper inapproximability bounds for Max-Morse Matching on regular cell complexes.

References

[1] M. Allili, T. Kaczynski, C. Landi, and F. Masoni. Algorithmic construction of acyclic partial matchings for multidimensional persistence. In W. G. Kropatsch, N. M. Artner, and I. Janusch, editors, *DGCI 2017: Discrete Geometry for Computer Imagery*, pages 375–387. Springer, 2017.

[2] U. Bauer, C. Lange, and M. Wardetzky. Optimal topological simplification of discrete functions on surfaces. *Discrete & Computational Geometry*, 47(2):347–377, 2012.

[3] P. Brendel, P. Dlotko, G. Ellis, M. Juda, and M. Mrozek. Computing fundamental groups from point clouds. *Appl. Algebra Eng., Commun. Comput.*, 26(1-2):27–48, Mar. 2015.

[4] B. A. Burton, T. Lewiner, J. a. Paixão, and J. Spreer. Parameterized complexity of discrete Morse theory. *ACM Trans. Math. Softw.*, 42(1):6:1–6:24, Mar. 2016.

[5] F. Cazals, F. Chazal, and T. Lewiner. Molecular shape analysis based upon the Morse–Smale complex and the Connolly function. In *Proceedings of the Nineteenth Annual Symposium on Computational Geometry*, SCG '03, pages 351–360. ACM, 2003.

[6] M. K. Chari. On discrete Morse functions and combinatorial decompositions. *Discrete Mathematics*, 217(1-3):101–113, 2000.

[7] J. Curry, R. Ghrist, and V. Nanda. Discrete morse theory for computing cellular sheaf cohomology. *Found. Comput. Math.*, 16(4):875–897, Aug. 2016.

[8] T. K. Dey, J. Wang, and Y. Wang. Graph Reconstruction by Discrete Morse Theory. In *34th International Symposium on Computational Geometry (SoCG 2018)*, 2018.

[9] Ö. Eğecioğlu and T. F. Gonzalez. A computationally intractable problem on simplicial complexes. *Computational Geometry*, 6(2):85 – 98, 1996.

[10] R. Forman. Morse theory for cell complexes. *Advances in Mathematics*, 134(1):90–145, 1998.

[11] R. Forman. A user's guide to discrete Morse theory. *Séminaire Lotharingien de Combinatoire*, B48c:1–35, 2002.

[12] V. Guruswami, J. Håstad, R. Manokaran, P. Raghavendra, and M. Charikar. Beating the random ordering is hard: Every ordering CSP is approximation resistant. *SIAM Journal on Computing*, 40(3):878–914, 2011.

[13] M. Hachimori. *Combinatorics of constructible complexes*. PhD thesis, Tokyo University, 2000.

[14] S. Harker, K. Mischaikow, M. Mrozek, and V. Nanda. Discrete Morse theoretic algorithms for computing homology of complexes and maps. *Foundations of Computational Mathematics*, 14(1):151–184, Feb 2014.

[15] S. Harker, K. Mischaikow, M. Mrozek, V. Nanda, H. Wagner, M. Juda, and P. Dlotko. The efficiency of a homology algorithm based on discrete Morse theory and coreductions. In *Proc. of 3rd Intl. Workshop on CTIC*, volume 1(1), 2010.

[16] A. Hatcher. *Algebraic Topology*. Cambridge University Press, 2002.

[17] P. Hersh. On optimizing discrete Morse functions. *Advances in Applied Mathematics*, 35(3):294–322, Sept. 2005.

[18] M. Jöllenbeck and V. Welker. *Minimal Resolutions via Algebraic Discrete Morse Theory*, volume 923 of *Memoirs of the American Mathematical Society*. American Mathematical Society, 2009.

[19] M. Joswig, F. H. Lutz, and M. Tsuruga. Heuristics for sphere recognition. In H. Hong and C. Yap, editors, *ICMS 2014: Mathematical Software*, pages 152–159, Berlin, Heidelberg, 2014. Springer Berlin Heidelberg.

[20] M. Joswig and M. E. Pfetsch. Computing optimal Morse matchings. *SIAM Journal on Discrete Mathematics*, 20(1):11–25, 2006.

[21] S. Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the Thiry-fourth Annual ACM Symposium on Theory of Computing*, STOC '02, pages 767–775, New York, NY, USA, 2002. ACM.

[22] S. Khot. On the unique games conjecture (invited survey). In *Proceedings of the 2010 IEEE 25th Annual Conference on Computational Complexity*, CCC '10, pages 99–121, Washington, DC, USA, 2010. IEEE Computer Society.

[23] D. Kozlov. Discrete Morse Theory. In *Combinatorial Algebraic Topology*, volume 21 of *Algorithms and Computation in Mathematics*, chapter 11, pages 179–209. Springer Berlin Heidelberg, 2008.

[24] T. Lewiner. Constructing discrete Morse functions. Master's thesis, Department of Mathematics, PUC-Rio, 2002.

[25] T. Lewiner, H. Lopes, and G. Tavares. Toward optimality in discrete Morse theory. *Experimental Mathematics*, 12(3):271–285, 2003.

[26] T. Lewiner, H. Lopes, and G. Tavares. Applications of forman's discrete morse theory to topology visualization and mesh compression. *IEEE Transactions on Visualization and Computer Graphics*, 10(5):499–508, Sept 2004.

[27] C. Maria and H. Schreiber. Discrete Morse theory for computing zigzag persistence, July 2018. Preprint.

[28] E. Miller, V. Reiner, and B. Sturmfels. *Geometric combinatorics*, volume 13 of *IAS/Park City Mathematics Series*. American Mathematical Society, 2007.

[29] J. Milnor. *Morse theory*. Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, No. 51. Princeton University Press, Princeton, N.J., 1963.

[30] K. Mischaikow and V. Nanda. Morse theory for filtrations and efficient computation of persistent homology. *Discrete & Computational Geometry*, 50(2):330–353, 2013.

[31] J. R. Munkres. *Elements of algebraic topology*. Addison-Wesley, 1984.

[32] A. Newman. The maximum acyclic subgraph problem and degree-3 graphs. In *Approximation, Randomization, and Combinatorial Optimization: Algorithms and Techniques*, volume 2129 of *Lecture Notes in Computer Science*, pages 147–158. Springer, 2001.

[33] C. H. Papadimitriou and M. Yannakakis. Optimization, approximation, and complexity classes. *Journal of Computer and System Sciences*, 43(3):425 – 440, 1991.

[34] P. Patak, 2018. Personal communication.

[35] A. Rathod, T. B. Masood, and V. Natarajan. Approximation algorithms for Max Morse Matching. *Computational Geometry*, 61:1 – 23, 2017.

[36] N. Shivashankar, S. M, and V. Natarajan. Parallel computation of 2D Morse–Smale complexes. *IEEE Transactions on Visualization and Computer Graphics*, 18(10):1757–1770, Oct 2012.

[37] M. Tancer. Recognition of collapsible complexes is NP-complete. 55(1):21–38, 2016.

[38] V. V. Vazirani. *Approximation algorithms*. Springer, 2003.

[39] D. P. Williamson and D. B. Shmoys. *The Design of Approximation Algorithms*. Cambridge University Press, 2011.

[40] E. Zeeman. On the dunce hat. *Topology*, 2(4):341 – 358, 1963.

B

PAPER II: PARAMETRIZED COMPLEXITY OF EXPANSION HEIGHT.

In order to obtain a combinatorial description of homotopy theory, Whitehead developed a refinement of homotopy theory, called *simple homotopy theory*. Simple homotopy theory considers sequences of elementary homotopy equivalences defined on simplicial complexes (or, more generally, CW complexes): an *elementary collapse*, which takes a face of a complex contained only in a single proper coface and removes both faces, and its inverse operation, called an *elementary expansion*. Two simplicial complexes are then said to be of the same simple homotopy type if one can be transformed into the other by a sequence of elementary collapses and expansions. Complexes of the same simple homotopy type are homotopy equivalent, but the converse is not always true.

The question of whether a finite contractible 2-dimensional CW complex K is simple homotopy equivalent to a point through expansions only up to dimension 3 is a famous open problem known as the Andrews-Curtis conjecture. The Andrews-Curtis conjecture is related to other important problems in mathematics like the Zeeman conjecture and the Poincaré conjecture. In this article, motivated by some of the important open problems in simple homotopy theory, we investigate the parametrized complexity of a few variants of the problem of deciding contractibility.

In particular, we focus on two problems:

- Decide whether a given 2-dimensional simplicial complex admits a simple-homotopy to a 1-complex using at most p expansions, called Erasability Expansion Height.
- Decide whether a given 2-dimensional simplicial complex admits a simple-homotopy to a 1-complex using at most p expansions, where all expansions come at the beginning of the sequence, called Erasability Expansion Height.

In general, the two heights may differ.

We begin with an introduction in Section 1, followed by preliminaries in Section 2. In Section 3, we describe the variants of the problems for which complexity bounds are established in later sections. In Section 4, we describe the main building block used in the reduction, which is a subdivision of the so-called modified dunce hat introduced in Paper I. In Section 5.1, we prove that Erasability Expansion Height is **W[P]**-hard. The proof uses a reduction from Axiom Set. The same reduction also establishes **W[P]**-hardness of Ordered Erasability Expansion Height. In Section 5.2, we show that Erasability

Expansion Height and Ordered Erasability Expansion Height are both in **W[P]**, and hence also **W[P]**-complete. Both results rest on the key observation that a 2-complex is erasable if and only if greedily collapsing triangles yields a 1-dimensional complex. In Section 6, we show that as a consequence of the above reduction, the problem of deciding whether a 2-dimensional simplicial complex can be shown to be simple-homotopy equivalent to a 1-complex using only 3-dimensional expansions is **NP**-complete.

STATEMENT OF INDIVIDUAL CONTRIBUTION. Jonathan Spreer suggested the idea for this project to Uli Bauer during a visit to Munich. Uli Bauer later involved me in the discussions. In this collaboration, I was the principal author and I was significantly involved in finding the ideas and carrying out the scientific work of all parts of this article. Even though I was the main person responsible for writing up this article, I received substantial writing help from my collaborators, especially from Jonathan Spreer, in Sections 4 and 5.

27th Annual European Symposium on Algorithms

ESA 2019, September 9–11, 2019, Munich/Garching, Germany

Edited by

Michael A. Bender

Ola Svensson

Grzegorz Herman

Editors

Michael A. Bender
Stony Brook University, NY, USA
bender@cs.stonybrook.edu

Ola Svensson
EPFL, Lausanne, Switzerland
ola.svensson@epfl.ch

Grzegorz Herman
Jagiellonian University, Kraków, Poland
gherman@tcs.uj.edu.pl

ACM Classification 2012
Applied computing → Transportation; Computing methodologies → Algebraic algorithms; Hardware → External storage; Human-centered computing → Graph drawings; Information systems → Data structures; Mathematics of computing → Algebraic topology; Mathematics of computing → Approximation algorithms; Mathematics of computing → Combinatorial algorithms; Mathematics of computing → Combinatorial optimization; Mathematics of computing → Graph algorithms ; Mathematics of computing → Graph algorithms; Mathematics of computing → Graphs and surfaces; Mathematics of computing → Graph theory; Mathematics of computing → Graph Theory; Mathematics of computing → Integer programming; Mathematics of computing → Linear programming; Mathematics of computing → Network flows; Mathematics of computing → Network optimization; Mathematics of computing → Paths and connectivity problems; Mathematics of computing → Probabilistic algorithms; Mathematics of computing → Probability and statistics; Mathematics of computing → Semidefinite programming; Mathematics of computing → Spectra of graphs; Networks → Network algorithms; Networks → Routing protocols; Theory of computation → Algorithm design techniques; Theory of computation → Algorithmic game theory and mechanism design; Theory of computation → Algorithmic mechanism design; Theory of computation → Approximation algorithms analysis; Theory of computation → Computational complexity and cryptography; Theory of computation → Computational geometry; Theory of computation → Convex optimization; Theory of computation → Data compression; Theory of computation → Data structures design and analysis; Theory of computation → Design and analysis of algorithms; Theory of computation → Dynamic graph algorithms; Theory of computation → Exact and approximate computation of equilibria; Theory of computation → Facility location and clustering; Theory of computation → Fixed parameter tractability; Theory of computation → Graph algorithms analysis; Theory of computation → Interactive proof systems; Theory of computation → Linear programming; Theory of computation → Lower bounds and information complexity; Theory of computation → Massively parallel algorithms; Theory of computation → Mathematical optimization; Theory of computation → Nearest neighbor algorithms; Theory of computation → Network flows; Theory of computation → Numeric approximation algorithms; Theory of computation → Online algorithms; Theory of computation → Packing and covering problems; Theory of computation → Parallel algorithms; Theory of computation → Parameterized complexity and exact algorithms; Theory of computation → Pattern matching; Theory of computation → Problems, reductions and completeness; Theory of computation → Quantum computation theory; Theory of computation → Quantum query complexity; Theory of computation → Routing and network design problems; Theory of computation → Scheduling algorithms; Theory of computation → Shortest paths; Theory of computation → Sketching and sampling; Theory of computation → Sparsification and spanners; Theory of computation → Streaming, sublinear and near linear time algorithms; Theory of computation → Submodular optimization and polymatroids; Theory of computation → Unsupervised learning and clustering; Theory of computation → W hierarchy

ISBN 978-3-95977-124-5

Published online and open access by
Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing, Saarbrücken/Wadern, Germany. Online available at https://www.dagstuhl.de/dagpub/978-3-95977-124-5.

Publication date
September, 2019

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at https://portal.dnb.de.

Digital Object Identifier: 10.4230/LIPIcs.ESA.2019.0
ISBN 978-3-95977-124-5 **ISSN 1868-8969** **https://www.dagstuhl.de/lipics**

Parametrized Complexity of Expansion Height

Ulrich Bauer
Department of Mathematics, Technical University of Munich (TUM),
Boltzmannstr. 3, 85748 Garching b. München, Germany
ulrich.bauer@tum.de

Abhishek Rathod
Department of Mathematics, Technical University of Munich (TUM),
Boltzmannstr. 3, 85748 Garching b. München, Germany
abhishek.rathod@tum.de

Jonathan Spreer
School of Mathematics and Statistics, The University of Sydney, NSW 2006 Australia
jonathan.spreer@sydney.edu.au

Abstract

Deciding whether two simplicial complexes are homotopy equivalent is a fundamental problem in topology, which is famously undecidable. There exists a combinatorial refinement of this concept, called simple-homotopy equivalence: two simplicial complexes are of the same simple-homotopy type if they can be transformed into each other by a sequence of two basic homotopy equivalences, an elementary collapse and its inverse, an elementary expansion. In this article we consider the following related problem: given a 2-dimensional simplicial complex, is there a simple-homotopy equivalence to a 1-dimensional simplicial complex using at most p expansions? We show that the problem, which we call the *erasability expansion height*, is $\mathbf{W}[\mathbf{P}]$-complete in the natural parameter p.

2012 ACM Subject Classification Mathematics of computing → Algebraic topology; Theory of computation → Problems, reductions and completeness; Theory of computation → W hierarchy

Keywords and phrases Simple-homotopy theory, simple-homotopy type, parametrized complexity theory, simplicial complexes, (modified) dunce hat

Digital Object Identifier 10.4230/LIPIcs.ESA.2019.13

Funding This research has been supported by the DFG Collaborative Research Center SFB/TRR 109 "Discretization in Geometry and Dynamics".
Jonathan Spreer: The author is partially supported by grant EVF-2015-230 of the Einstein Foundation Berlin. The authors acknowledges support by The University of Sydney, where parts of this work were finished.

Acknowledgements We want to thank João Paixão for very helpful discussions.

1 Introduction

Homotopy theory lies at the heart of algebraic topology. In an attempt to make the concept of homotopy equivalence more amenable to combinatorial methods, Whitehead developed what turned out to be a combinatorial refinement of the theory, called *simple-homotopy theory*. Simple-homotopy theory considers sequences of elementary homotopy equivalences defined on simplicial complexes (or, more generally, CW complexes): an *elementary collapse*, which takes a face of a complex contained only in a single proper coface and removes both faces, and its inverse operation, called an *elementary expansion*. Two simplicial complexes are then said to be of the same simple-homotopy type if one can be transformed into the other by a sequence of elementary collapses and expansions. Complexes of the same simple-homotopy type are homotopy equivalent, but the converse is not always true [18], the obstruction being an element of the *Whitehead group* of the fundamental group. However, Whitehead

proved that all homotopy-equivalent complexes with a trivial fundamental group are in fact of the same simple-homotopy type [19], and thus in this particular case the notions of simple-homotopy and homotopy coincide.

A presentation of the fundamental group can be read off from a two-dimensional complex such that the presentation is balanced and describes the trivial group if and only if this complex is contractible [12]. Since the decidability of the triviality problem for balanced presentations is open [4], the same is also true for the decidability of contractibility of 2-complexes. Hence, the decidability of the existence of a simple-homotopy equivalence from a 2-complex to a point is also open. In contrast, the problem of deciding whether a given complex has trivial fundamental group is famously undecidable already for 2-complexes through its connection to the word problem, see, for instance, [5]. It follows that sequences of elementary collapses and expansions proving simple-homotopy equivalence between a 2-complex and a point can be expected to be long, if not unbounded. Nonetheless, understanding these sequences offers a great reward: the statement that any contractible 2-complex contracts to a point using only expansions up to dimension three is equivalent to a weaker variant [13, p. 34–35] of the Andrews–Curtis conjecture [1, 20].

In this article, motivated by the aforementioned problems, we investigate the computational (parametrized) complexity of a number of variants of the problem of deciding contractibility. More precisely, we focus on the problem of deciding whether a given 2-complex admits a simple-homotopy to a 1-complex using at most p expansions, called Erasibility Expansion Height. In addition, we consider a variant, called Ordered Erasibility Expansion Height, which requires that all expansions come at the very beginning of the sequence. It is worth noting that Erasibility Expansion Height and Ordered Erasibility Expansion Height are equivalent for CW complexes for which one can readily swap the order of expansions and collapses [13, p. 34]. However, for simplicial complexes, the ordered and unordered expansion heights may differ.

In Section 5.1, we prove that Erasibility Expansion Height is **W[P]**-hard, see Theorem 8. The proof uses a reduction from Axiom Set. The same reduction also establishes **W[P]**-hardness of Ordered Erasibility Expansion Height. Note that a reduction from Axiom Set is also used by the third author and others in [6] to establish **W[P]**-hardness of a parametrized version of Optimal Morse Matching. However, unlike in [6], the use of combinatorial and topological properties of the dunce hat is a key ingredient of the reduction used in this paper. In particular, there is only one gadget in the reduction – a subdivision of the so-called *modified dunce hat* [3], see Figure 2. In this sense the techniques used in this paper are also related to recent work by the first and second author in [3], where they show hardness of approximation for some Morse matching problems.

In Section 5.2, we show that Erasibility Expansion Height and Ordered Erasibility Expansion Height are both in **W[P]**, and hence also **W[P]**-complete, see Theorems 8, 13 and 14. Both results rest on the key observation that a 2-complex is erasable if and only if greedily collapsing triangles yields a 1-dimensional complex (Proposition 11), as shown by Tancer [17, Proposition 5].

In Section 6 we show that, as a consequence of the above reduction, the problem of deciding whether a 2-complex can be shown to be simple-homotopy equivalent to a 1-complex using only 3-dimensional expansions, called Erasibility 3-Expansion Height, is **NP**-complete, see Theorem 17.

2 Definitions and Preliminaries

2.1 Simplicial complexes

A (finite) *abstract simplicial complex* is a collection K of subsets of a finite ground set V such that if τ is an element of K, and σ is a nonempty subset of τ, then σ is an element of K. The ground set V is called the *set of vertices* of K. Since simplicial complexes are determined by their facets, we sometimes present simplicial complexes by listing their facets. A *subcomplex* of K is a subset $L \subseteq K$ which is itself a simplicial complex. Given a subset $W \subseteq V$ of the vertices of K, the *induced subcomplex on* W consists of all simplices of K that are subsets of W.

The elements of K are referred to as its *faces*. The *dimension of a face* is defined to be its cardinality minus one, and the *dimension of* K equals the largest dimension of its faces. For brevity, we sometimes refer to a d-dimensional simplicial complex as a *d-complex* and to a d-dimensional face as a *d-face*. The 0-, 1-, and 2-faces of a d-complex K are called its *vertices*, *edges*, and *triangles* respectively. Faces of K which are not properly contained in any other face are called *facets*. An $(m-1)$-face $\sigma \in K$ which is contained in exactly one m-face $\tau \in K$ is called *free*.

The *star* of a vertex v of complex K, written $\mathrm{star}_K(v)$, is the subcomplex consisting of all faces of K containing v, together with their faces. If a map $\phi : V \to W$ between the vertex sets of two simplicial complexes K and L, respectively, sends every simplex $\sigma \in K$ to a simplex $\phi(\sigma) \in L$, then the induced map $f : K \to L, \sigma \mapsto \phi(\sigma)$, is said to be *simplicial*.

2.2 Simple-homotopy

We introduce the basic notions of simple-homotopy used in the present paper. The general concept of simple-homotopy can be understood independently from the notion of homotopy. In this sense this article aims to be self-contained. For further reading on homotopy theory we refer to [11].

In short, a simple-homotopy equivalence is a refinement of a homotopy equivalence. It can be described purely combinatorially with the help of the following definition.

▶ **Definition 1** (Elementary collapses and expansions). *Let K_0 be a simplicial complex, and let $\tau, \sigma \in K_0$ be an m-face and an $(m-1)$-face respectively such that $\sigma \subset \tau$, and σ is free in K_0.*

We say that $K_1 = K_0 \setminus \{\tau, \sigma\}$ arises from K_0 by an elementary collapse of dimension m *or* elementary m-collapse*, denoted by $K_0 \searrow K_1$. Its inverse, the operation $K_0 = K_1 \cup \{\tau, \sigma\}$ is called an* elementary expansion of dimension m *or* elementary m-expansion*, written $K_0 \nearrow K_1$. If the complex is implicit from context, we denote elementary collapses by $\searrow^{\tau}_{\sigma}$ and elementary expansions by $\nearrow^{\tau}_{\sigma}$. An elementary collapse or an elementary expansion is sometimes referred to as an* elementary move*, or simply a* move*.*

If there exists a sequence of elementary collapses turning a complex K_0 into K_1 we write $K_0 \searrow K_1$ and say that K_0 collapses to *K_1. If K_1 is one-dimensional, we say that K_0 is* erasable*. If K_1 is merely a point we call K_0* collapsible*.*

Finally, we write $K_0 \nearrow K_1$ to indicate a sequence of expansions and say that K_0 expands *to K_1.*

It follows that an expansion $\nearrow^{\tau}_{\sigma}$ can only be performed in a simplicial complex K if all codimension 1 faces of τ except for σ are already in K. Hence, let τ be an m-face of a simplicial complex K, and let σ be one of its $(m-1)$-faces. An *(m-dimensional) horn* $\mathbf{H}(\tau, \sigma)$ associated to the pair (τ, σ) is the simplicial complex generated by the $(m-1)$-faces of τ apart from σ.

All m-expansions and m-collapses with $m > 1$ leave the vertex set unchanged.

▶ **Definition 2** (Simple-homotopy equivalence, simple-homotopy graph)**.** *Two simplicial complexes K and L are said to be* simple homotopy equivalent *or of coinciding* simple-homotopy type*, written $K \nearrow\!\searrow L$, if there exists a sequence $\mathcal{S}$ of elementary moves turning one into the other. In this case, we write $\mathcal{S} : K \nearrow\!\searrow L$.*

The dimension of a simple-homotopy equivalence *is the maximum of the dimensions of K, L and of any elementary expansion or collapse in the sequence.*

The graph whose nodes are simplicial complexes, and two nodes are adjacent if their corresponding complexes are related by an elementary collapse is called simple-homotopy graph*. Naturally, its connected components are in one-to-one correspondence with simple-homotopy types.*

Two simplicial complexes of the same simple-homotopy type are homotopy equivalent, but the converse is not true, see, for instance, [18]. For simple-homotopy equivalent simplicial complexes we know the following.

▶ **Theorem 3** (Wall [17], Matveev [13, Theorem 1.3.5])**.** *Let K and L be two simplicial complexes of the same simple-homotopy type and of dimension at most $m > 2$. Then there exists a simple-homotopy equivalence of dimension at most $m + 1$, taking one to the other.*

For the case $m = 2$, Theorem 3 is still open and known as the (topological or geometric) Andrews–Curtis conjecture [2, 13, 15]. On the other hand, it is known that any contractible 2-complex is also simple-homotopy equivalent to a point [19]. Hence, any pair of contractible 2-complexes can be connected by a simple-homotopy equivalence of dimension at most four – but determining whether we can always decide if such a simple-homotopy equivalence exists is an open question [4], equivalent to the triviality problem for balanced group presentations [12].

2.3 Parametrized complexity

Parametrized complexity, as introduced by Downey and Fellows in [7], is a refinement of classical complexity theory. The theory revolves around the general idea of developing complexity bounds for instances of a problem not just based on their size, but also involving an additional *parameter*, which might be significantly smaller than the size. Specifically, we have the following definition.

▶ **Definition 4** (Parameter, parametrized problem)**.** *Let Σ be a finite alphabet.*

1. *A* parameter *of Σ^*, the set of strings over Σ, is a function $\rho : \Sigma^* \to \mathbb{N}$, attaching to every input $w \in \Sigma^*$ a natural number $\rho(w)$.*
2. *A* parametrized problem *over Σ is a pair (P, ρ) consisting of a set $P \subseteq \Sigma^*$ and a parametrization $\rho : \Sigma^* \to \mathbb{N}$.*

In this article we consider the complexity class $\mathbf{W[P]}$ for parametrized problems, following the definition by Flum and Grohe [8].

▶ **Definition 5** (Complexity Class **W[P]**)**.** *Let Σ be an alphabet and $\rho : \Sigma^* \to \mathbb{N}$ a parametrization. A nondeterministic Turing machine $\mathbb{M}$ with input alphabet Σ is called* ρ-restricted *if there are computable functions $f, h : \mathbb{N} \to \mathbb{N}$ and a polynomial p (with coefficients in the set of natural numbers) such that on every run with input $x \in \Sigma^*$ the machine $\mathbb{M}$ performs at most $f(k) \cdot p(|x|)$ steps, at most $h(k) \cdot \log |x|$ of them being nondeterministic, where $k := \rho(x)$. **W[P]** is the class of all parametrized problems (P, ρ) that can be decided by a ρ-restricted nondeterministic Turing machine.*

3 Problems

In this article we consider the following parametrized problems.

▶ **Problem 1** (Erasibility Expansion Height).

Instance: A 2-dimensional simplicial complex K and a natural number p.
Parameter: p.
Question: Is there a path in the simple-homotopy graph connecting K to a 1-complex using at most p expansions?

▶ **Problem 2** (Ordered Erasibility Expansion Height).

Instance: A 2-dimensional simplicial complex K and a natural number p.
Parameter: p.
Question: Is there a path in the simple-homotopy graph connecting K to a 1-complex using first at most p expansions, followed by a sequence of only collapses?

In Section 5, we establish **W[P]**-completenes for Erasibility Expansion Height and Ordered Erasibility Expansion Height.

The hardness proof works via a parametrized reduction using the Axiom Set problem, which is a classical **NP**-complete problem [9, p. 263] that is well-known to be **W[P]**-complete with respect to the appropriate parameter [7, p. 473].

▶ **Problem 3** (Axiom Set).

Instance: A finite set S of *sentences*, an *implication relation* R consisting of pairs (U, s) where $U \subseteq S$ and $s \in S$, and a positive integer $p \leq |S|$.
Parameter: p.
Question: Is there a set $S_0 \subseteq S$, called an *axiom set*, with $|S_0| \leq p$ and a positive integer n such that if we recursively define

$$S_i := S_{i-1} \cup \{s \in S \mid \exists U \subseteq S_{i-1} : (U, s) \in R\}$$

for $1 \leq i \leq n$, then $S_n = S$?

▶ Remark 6. Note that every instance of Axiom Set can be reduced in polynomial time to an instance for which every sentence must occur in at least one implication relation: First iteratively remove all sentences from the instance which do not feature in at least one implication relation. Then, for each of them, reduce p by one (note that each of them must necessarily be an axiom). It follows that solving the reduced instance is equivalent to solving the original instance.

Similarly, note that if there exists an implication $(U, s) \in R$, $s \in U$, we can simply omit it and, if this deletes s from the instance altogether, decrease p by one.

In Section 6, we show that the following variants of the expansion height problem are **NP**-complete.

▶ **Problem 4** (Erasibility 3-Expansion Height).

Instance: A finite 2-dimensional simplicial complex K and a natural number p.
Question: Is there a path in the simple-homotopy graph connecting K to a 1-complex using at most p expansions, all of which are 3-expansions?

▶ **Problem 5** (Ordered Erasibilty 3-Expansion Height).

Instance: A finite 2-dimensional simplicial complex K and a natural number p.
Question: Is there a path in the simple-homotopy graph connecting K to a 1-complex using first at most p expansions, all of which are 3-expansions, followed by a sequence of only collapses?

4 Contractibility and collapsibility for 2-complexes

The main gadget used in the proof of our main result, Theorem 8, is based on the simplest 2-dimensional contractible complex which is not collapsible to a point – the *dunce hat*. Hence, before we describe our main gadget in detail, we start this section by briefly discussing minimal triangulations of the dunce hat, and a variant that is collapsible through a unique free edge, the *modified dunce hat*.

4.1 The dunce hat

In the category of CW complexes, the dunce hat can be obtained by identifying two boundary edges of a triangle to build a cone and then gluing the third edge along the seam of the first gluing. The resulting complex does not have a collapsible triangulation. On the other hand, the dunce hat is known to be contractible [21].

The smallest simplicial complexes realizing this construction have 8 vertices, 24 edges and 17 triangles. There are seven such minimal triangulations of the dunce hat [16]. One such triangulation, denoted by $\mathbf{D}$, is shown in Figure 1. The dunce hat $\mathbf{D}$ has two horns, namely $\mathbf{H}(\{2,7,8\},\{1,2,7,8\})$ and $\mathbf{H}(\{3,5,6\},\{1,3,5,6\})$, and hence admits two 3-expansions, namely $\nearrow^{\{1,2,7,8\}}_{\{2,7,8\}}$ and $\nearrow^{\{1,3,5,6\}}_{\{3,5,6\}}$ respectively. They are shown by the shaded areas in Figure 1.

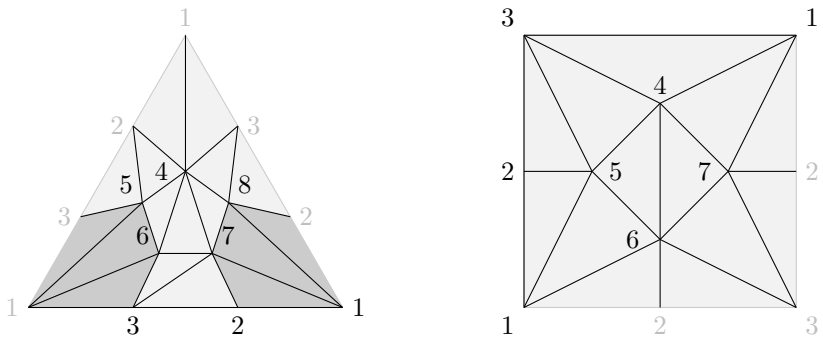


Figure 1 Left: The 8-vertex triangulation $\mathbf{D}$ of the dunce hat. The two expansions turning it collapsible are highlighted. Right: The 7-vertex triangulation $\mathbf{P}$ of the modified dunce hat.

Note that after any of these two expansions we obtain a collapsible complex: After the expansion $\nearrow^{\{1,2,7,8\}}_{\{2,7,8\}}$ and the collapses $\searrow^{\{1,2,7,8\}}_{\{1,7,8\}}$, $\searrow^{\{1,2,7\}}_{\{1,7\}}$, and $\searrow^{\{1,2,8\}}_{\{1,8\}}$, the edge $\{1,2\}$ becomes free and thus $\mathbf{D}$ becomes collapsible. Similarly, starting with $\nearrow^{\{1,3,5,6\}}_{\{3,5,6\}}$, one may perform the collapse $\searrow^{\{1,3,5,6\}}_{\{1,5,6\}}$ and proceed in an analogous way. In particular, this shows that the dunce hat has the simple-homotopy type of a point, and in fact can be made collapsible by using a single expansion.

4.2 The modified dunce hat

Rather than working with the dunce hat directly, we base the construction of our gadget for the proof of Theorem 8 on the *modified dunce hat* [10]. More precisely, we "insert" a free edge into the dunce hat. For instance, Figure 1 depicts a triangulation of the modified dunce hat, which we denote by $\mathbf{P}$, with $\{1,3\}$ as the unique free edge. This particular triangulation of the modified dunce hat uses only 7 vertices, 19 edges, and 13 triangles. The modified dunce hat has previously been used as a gadget to show hardness of approximation for Morse matchings [3].

If we assume that $\mathbf{P}$ is part of a larger complex K, in which edge $\{1,3\}$ is glued to triangles not lying in $\mathbf{P}$, then $\{1,3\}$ is not free. In this case, the triangles of $\mathbf{P}$ can be collapsed away in essentially two distinct ways. Either, at some point in a simple-homotopy on K, the edge $\{1,3\}$ becomes free and thus the triangles of $\mathbf{P}$ collapse, or the triangles of $\mathbf{P}$ become collapsible by performing one of two possible 3-expansions on $\mathbf{P}$. Looking at the latter case in more detail, we have the following sequences of expansions and collapses:

$$\nearrow^{\{1,2,5,6\}}_{\{2,5,6\}}, \quad \searrow^{\{1,2,5,6\}}_{\{1,5,6\}}, \quad \searrow^{\{1,2,5\}}_{\{1,5\}}, \quad \searrow^{\{1,2,6\}}_{\{1,6\}}$$

and

$$\nearrow^{\{2,3,6,7\}}_{\{2,6,7\}}, \quad \searrow^{\{2,3,6,7\}}_{\{3,6,7\}}, \quad \searrow^{\{2,3,6\}}_{\{3,6\}}, \quad \searrow^{\{2,3,7\}}_{\{3,7\}}.$$

In the first case, the edge $\{1,2\}$ is freed, in the second case, the edge $\{2,3\}$ is freed. Both sequences can be extended to a collapsing sequence of the entire complex $\mathbf{P}$.

4.3 The main gadget

Our gadget for the proof of Theorem 8 is a subdivided version of the modified dunce hat $\mathbf{P}$ from Section 4.2. More precisely, it is determined by two positive integers m and ℓ, denoted by $\mathbf{P}_{m,\ell}$, and can be constructed from the complex $\mathbf{P}$ in essentially two steps.

1. Subdivide the edge $\{1,3\}$ of $\mathbf{P}$ $(m-1)$ times, thereby introducing vertices $x_1, \dots x_{m-1}$. Relabel $1 \to x_0$ and $3 \to x_m$ to obtain m free edges $f_i = \{x_{i-1}, x_i\}$, $1 \le i \le m$.
2. Remove the edge $\{4,6\}$ and place ℓ vertex-disjoint copies of the disk

$$\{\{c_j, a_j, y_j\}, \{c_j, y_j, z_j\}, \{c_j, z_j, b_j\}, \{d_j, a_j, y_j\}, \{d_j, y_j, z_j\}, \{d_j, z_j, b_j\}\},$$

$1 \le j \le \ell$, inside the 4-gon in the center of $\mathbf{P}$ bounded by 4, 5, 6, and 7. Triangulate the remaining space in the interior of the 4-gon. This creates edges $e_j = \{y_j, z_j\}$, $1 \le j \le \ell$, with pairwise vertex disjoint stars disjoint to 4, 5, 6, and 7 (now $\{4,6\}$ reappears as a path from 4 to 6, and thus $\mathbf{P}_{m,\ell}$ is in fact a proper subdivision of $\mathbf{P}$). See Figure 2 for an illustration.

One key property of $\mathbf{P}_{m,\ell}$ is that we do not subdivide any faces of $\mathbf{P}$ near to the two available 3-expansions. As a result, again, assuming that $\mathbf{P}_{m,\ell}$ is part of a larger complex K where all free edges of $\mathbf{P}_{m,\ell}$ are glued to other triangles of K outside of $\mathbf{P}_{m,\ell}$ and thus are not free, the triangles of $\mathbf{P}_{m,\ell}$ can be collapsed according to the following observation:

▶ Remark 7. Let K be a two-dimensional simplicial complex such that $\mathbf{P}_{m,\ell}$ is a subcomplex whose vertices do not span any other faces of K (i.e., $\mathbf{P}_{m,\ell}$ is an induced subcomplex of K), and $K \nearrow\!\!\searrow L$ where L is a 1-complex. Then, at least one of the following three statements holds true at some point in $K \nearrow\!\!\searrow L$, enabling us to eventually collapse away all the triangles of $\mathbf{P}_{m,\ell}$.

1. one of the edges $f_i \in \mathbf{P}_{m,\ell}$ becomes free;
2. one of two 3-expansions on $\mathbf{P}_{m,\ell}$: $\nearrow^{\{1,2,5,6\}}_{\{2,5,6\}}$ or $\nearrow^{\{2,3,6,7\}}_{\{2,6,7\}}$ is performed;
3. multiple expansions result in a complex in which all the triangles of $\mathbf{P}_{m,\ell}$ can be collapsed.

In other words, if one of the edges $f_i \in \mathbf{P}_{m,\ell}$ does not become free at some point in $K \nearrow\!\!\searrow L$, then one is forced to use 3-expansions (either directly on $\mathbf{P}_{m,\ell}$, or after performing additional expansions) to collapse away the triangles of $\mathbf{P}_{m,\ell}$.

In Section 5.1 we use this gadget to reduce an instance $A = (S, R, p)$ of Axiom Set to Erasibility Expansion Height: Every sentence $s \in S$ is associated with one copy of $\mathbf{P}_{m,\ell}$, the edges f_i correspond to implications $(U, s) \in R$, and the edges e_j correspond to whenever $s \in U$ for some implication $(U, u) \in R$.

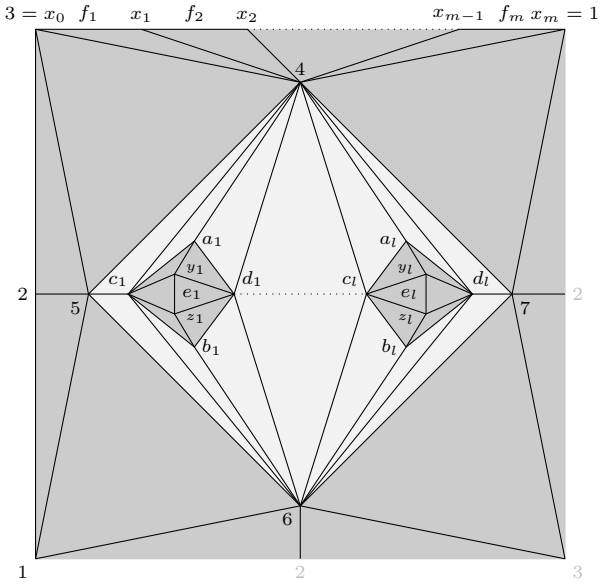


Figure 2 The main gadget of the construction $\mathbf{P}_{m,\ell}$.

5 Parametrized complexity of Erasibility Expansion Height

In this section, we first prove that Erasibility Expansion Height and Ordered Erasibility Expansion Height are **W**[**P**]-hard by a reduction from Axiom Set, a problem known to be **W**[**P**]-complete. We then show that the two problems are also contained in **W**[**P**].

5.1 W[P]-hardness of expansion height problems

▶ **Theorem 8.** Erasibility Expansion Height *and* Ordered Erasibility Expansion Height *are* ***W[P]-hard problems.***

The following lemma is used to assemble the gadgets in our reduction into a simplicial complex K.

▶ **Lemma 9** (Munkres, [14, Lemma 3.2])**.** *Let C be a finite set, let K be a simplicial complex with set of vertices V, and let $f : V \to C$ be a surjective map associating to each vertex of K a color from C. The coloring f extends to a simplicial map $g : K \to K_f$ where K_f has vertex set C and is obtained from K by identifying vertices with equal color.*

If for all pairs $v, w \in V$, $f(v) = f(w)$ implies that their stars $\mathrm{star}_K(v)$ and $\mathrm{star}_K(w)$ are vertex disjoint, then, for all faces $\tau, \sigma \in K$ we have that

- *τ and $g(\tau)$ have the same dimension, and*
- *$g(\tau) = g(\sigma)$ implies that either $\tau = \sigma$ or τ and σ are vertex disjoint in K.*

Lemma 9 provides a way of gluing faces of a simplicial complex by a simplicial quotient map obtained from vertex identifications, and tells us when this gluing does not create unwanted identifications.

Proof of Theorem 8. We want to reduce Axiom Set to Erasibility Expansion Height.

Fix an instance $A = (S, R, p)$ of Axiom Set such that every sentence $s \in S$ is subject to at least one implication $(U, s) \in R$ and such that $(U, s) \in R$ implies $s \notin U$. By Remark 6, this is not a restriction, since every instance of Axiom Set can be reduced to such an

instance in polynomial time. For every sentence $s \in S$, take a copy $\mathbf{P}_s$ of the gadget $\mathbf{P}_{m,\ell}$ to model s, where $\ell \geq 0$ is the number of implications $(U, u) \in R$ with $s \in U$ and $m \geq 1$ is the number of implications $(U, s) \in R$. Thus, for all values $\ell \geq 0$ and $m \geq 1$ the gadget $\mathbf{P}_{m,\ell}$ is a simplicial complex without any unintended identifications. Denote the free edges of $\mathbf{P}_s$ by $f_i^s = \{x_{i-1}^s, x_i^s\}$, $1 \leq i \leq m$, and its edges of type e_j by $e_j^s = \{y_j^s, z_j^s\}$, $1 \leq j \leq \ell$.

For a fixed $s \in S$, endow the set of implications $(U, s) \in R$ of s with an arbitrary order $(U_1, s), \ldots, (U_m, s)$. Similarly, for a fixed $u \in S$, order the set of implications in R containing u arbitrarily as $(U^1, s^1), \ldots, (U^\ell, s^\ell)$. Now for every $(U_i, s) \in R$ and every $u \in U_i = U^j$ (i.e., $s^j = s$), glue the edge $f_i^s = \{x_{i-1}^s, x_i^s\}$ of the gadget $\mathbf{P}_s$ to the edge $e_j^u = \{y_j^u, z_j^u\}$ of $\mathbf{P}_u$ by identifying x_{i-1}^s with y_j^u and x_i^s with z_j^u.

Performing these identifications for all implications in R yields a complex, which we denote by K. Note that, fixing $s \in S$, and $0 \leq i \leq m$, the only vertices to which x_i^s can possibly be identified to in K are y_j^u, z_j^u $(1 \leq j \leq \ell)$.

More precisely, using the orderings from above for vertex x_i^s, let (U_i, s) and (U_{i+1}, s) be the i-th and $(i+1)$-st implication of s in R (if $i \in \{0, m\}$ there is only one implication to consider) and denote their sentences by $u_1, \ldots, u_r$ and $u^1, \ldots, u^t$, where $r = |U_i|$ and $t = |U_{i+1}|$. Moreover, let U_i (resp. U_{i+1}) be the j_f-th (resp. j^g-th) implication where the sentence u_f (resp. u^g) occurs, for $1 \leq f \leq r$ (resp. $1 \leq g \leq t$). Then x_i^s is identified with $z_{j_f}^{u_f}$ $(1 \leq f \leq r)$ and $y_{j^g}^{u^g}$ $(1 \leq g \leq t)$.

Now since every fixed edge of type e_j^u is only identified with one edge of type f_i^s, those vertices are not identified with any other vertices of K. Since, by construction, the set of the vertex stars of $z_{j_f}^{u_f}$ $(1 \leq f \leq r)$, $y_{j^g}^{u^g}$ $(1 \leq g \leq t)$, and x_i^s are pairwise vertex disjoint, we can apply Lemma 9 to ensure that no unwanted identifications occur in building up K. In particular, every gadget $\mathbf{P}_s$ is a subcomplex of K via the canonical isomorphism given by the gluing map.

We now show that the following three statements are equivalent for our complex K:

(a) there exists a simple-homotopy equivalence turning K into a 1-dimensional complex using first at most p expansions, followed by a sequence of only collapses,

(b) there exists a simple-homotopy equivalence $K \nearrow\!\searrow L$ turning K into a 1-dimensional complex L using at most p expansions, and

(c) there exists an axiom set $S_0 \subset S$ for $A = (S, R, p)$ using at most p elements.

We trivially have that (a) $\Longrightarrow$ (b).

In order to show that (c) $\Longrightarrow$ (a), assume that there exists an axiom set $S_0 \subset S$ of size p, and perform one 3-expansion on each gadget $\mathbf{P}_u$ with $u \in S_0$. As described in Section 4.3, these expansions admit all triangles of these gadgets to collapse. This, in turn, frees all edges f_i^s where $s \in S$ has an implication $(U, s) \in R$ with $U \subset S_0$. Consequently, all triangles of such gadgets $\mathbf{P}_s$ can be collapsed. Since S_0 is an axiom set, repeating this process eventually collapses away all tetrahedra and triangles, leaving a 1-complex.

In order to show that (b) $\Longrightarrow$ (c), we start with a few definitions. For a complex K' with $K \nearrow\!\searrow K'$, we say that a gadget $\mathbf{P}_s \subseteq K$ is *touched* with respect to a simple-homotopy sequence $\mathcal{S} : K \nearrow\!\searrow K'$ if at some point in the simple-homotopy sequence one of the triangles of $\mathbf{P}_s$ is removed. Otherwise $\mathbf{P}_s$ is said to be *untouched*. Note that even if all triangles of $\mathbf{P}_s$ are present in K', $\mathbf{P}_s$ might still be touched. Being touched or untouched is a property of the sequence $\mathcal{S} : K \nearrow\!\searrow K'$, not of the complex K'.

We build the axiom set $S_0 \subset S$ for $A = (S, R, p)$ in the following way: A sentence $s \in S$ is in S_0 if and only if a triangle in $\mathbf{P}_s$ is removed by a 3-collapse of the given simple-homotopy sequence $\mathcal{S} : K \nearrow\!\searrow L$. We first inductively prove a claim about

$$S^k = \{s \in S \mid \mathbf{P}_s \text{ is touched by a 3-collapse in the first } k \text{ moves of } K \nearrow\!\searrow L\}.$$

We first inductively prove a claim about

$$S^k = \{s \in S_0 \mid \mathbf{P}_s \text{ is by a 3-collapse in the first } k \text{ moves of } K \searrow L\}.$$

▷ Claim 10. For $s \in S$, if the gadget $\mathbf{P}_s$ is touched by the first k elementary moves in $\mathcal{S} : K \searrow L$, then s is implied by sentences in S^k.

Proof. First note that all gadgets are untouched in K and $S^0 = \emptyset$.

By induction hypothesis, if a gadget $\mathbf{P}_s$ is touched by one of the first $k-1$ moves in $\mathcal{S} : K \searrow L$, then s is implied by sentences in S^{k-1}.

The induction claim is trivially true if $\mathbf{P}_s$ is touched in the first $k-1$ moves, or if $\mathbf{P}_s$ is touched by a 3-collapse in the k-th move ($s \in S^k \setminus S^{k-1}$), causing the sentence s to be included in S_0.

So, suppose that this is not the case. That is, suppose that $\mathbf{P}_s$ is untouched in the length $k-1$ prefix $\mathcal{S}' : K \searrow K'$ of $\mathcal{S} : K \searrow L$ and touched by a 2-collapse in the k-th move. This implies that $S^k = S^{k-1}$ and that $\mathbf{P}_s$ is a subcomplex of K' and one of the edges f_i^s must be free in K'. Now let $\mathbf{P}_{u_1}, \mathbf{P}_{u_2}, \ldots, \mathbf{P}_{u_q}$ be the set of other gadgets containing triangles glued to f_i^s in the original complex K (that is, there is an implication $(\{u_1, u_2, \ldots, u_q\}, s) \in R$). Since none of these triangles are present in K', all gadgets $\mathbf{P}_{u_1}, \mathbf{P}_{u_2}, \ldots, \mathbf{P}_{u_q}$ must be touched in $\mathcal{S}' : K \searrow K'$. Thus, either they were touched by a 3-collapse and their corresponding sentences are part of S^{k-1}, or they were touched by a 2-collapse and, by the induction hypothesis, their corresponding sentences are implied by sentences in S^k. It follows that s is implied by sentences in $S^k = S^{k-1}$, proving the claim. ◁

By assumption, K is simple homotopy equivalent to a 1-complex L. That is, $\mathcal{S}' : K \searrow L$ eventually removes all triangles from K. Hence, every sentence $s \in S$ is touched as a result of a 2-collapse or a 3-collapse. Let m be the number of elementary moves needed to reach L starting from K. Then, by Claim 10, $S^m = S_0$ is the desired axiom set. Also, since a sentence s is included in S_0 only if a triangle belonging to gadget $\mathbf{P}_s$ is removed as part of a 3-collapse, and since a triangle belonging to gadget $\mathbf{P}_s$ does not belong to any other gadget $\mathbf{P}_u$ for $u \neq s$, S_0 cannot contain more elements than the number of 3-collapses (and hence 3-expansions).

Finally, we infer **W[P]**-hardness of Erasibility Expansion Height and Ordered Erasibility Expansion Height from the above equivalence and the **W[P]**-hardness of Axiom Set [7, p. 473]. ◀

5.2 W[P]-membership of Erasibility Expansion Height

We now show that Ordered Erasibility Expansion Height and Erasibility Expansion Height are in **W[P]** by describing suitable nondeterministic algorithms for deciding both problems. We begin with a well-known fact about checking collapsibility of 2-complexes.

▶ **Proposition 11** (Tancer [17], Proposition 5)**.** *Let K be a 2-complex that collapses to a 1-complex L and to another 2-complex M. Then M also collapses to a 1-complex.*

▶ Remark 12. The proposition above implies that we can collapse an input 2-complex K greedily until no more 2-collapses are possible, and if K collapses to a 1-complex L, the algorithm is guaranteed to terminate with a 1-complex as well.

▶ **Theorem 13.** Ordered Erasibility Expansion Height *is in* ***W[P].***

Proof. Let K be a simplicial complex with n simplices. First, note that if there exists a simple homotopy sequence $\mathcal{S}$ taking K to a 1-complex with p expansions that all come at the beginning of the sequence, then there also exists a simple homotopy sequence $\mathcal{S}_{\mathbb{M}}$ taking K to a 1-complex where p expansions are followed by collapses such that, for each d, all collapses of dimension $d+1$ are executed before collapses of dimension d. This follows from observing that, for any two d-collapses $\searrow_{\sigma}^{\tau}$ and $\searrow_{\alpha}^{\beta}$, if the d-collapse $\searrow_{\sigma}^{\tau}$ is executed before the d-collapse $\searrow_{\alpha}^{\beta}$ in $\mathcal{S}$, then the same can be carried out in $\mathcal{S}_{\mathbb{M}}$. Also, in any simple homotopy sequence $\mathcal{S}$ that takes K to a 1-complex, for every $d > 2$, the number of d-expansions equals the number of d-collapses in $\mathcal{S}$. This follows from a simple inductive argument starting with highest dimensional moves.

Denoting the total number of d-collapses, $d > 2$, in $\mathcal{S}$ by $q \leq p$, it follows that, if there exists a simple homotopy sequence $\mathcal{S}$ with p expansions that come at the beginning, then there exists a simple homotopy sequence $\mathcal{S}_{\mathbb{M}}$ with p expansions in the beginning followed by q collapses that gives rise to a 2-complex K' with $O(n^3)$ faces. The faces can be as many as $O(n^3)$ since $\mathbb{M}$ does not guess any 2-collapses. Furthermore, if $\mathcal{K}$ is erasable through the simple homotopy sequence $\mathcal{S}$, then K' is also erasable, once again, because the 2-collapses of $\mathcal{S}$ can be carried out in the same order in $\mathcal{S}_{\mathbb{M}}$. Hence, the non-deterministic Turing machine $\mathbb{M}$ can now be described as follows:

1. Guess p expansions and q collapses non-deterministically to obtain a complex K'.
2. Deterministically check if K' is erasable.

By Remark 12, erasability of K' can be deterministically checked in time polynomial in n.

Since any simplex in the desired simple homotopy sequence has at most $n+p$ vertices, the number of bits required to encode a single vertex is $O(\log(n+p))$. Also, because the dimension of the faces involved in expansions and collapses is certainly in $O(p)$, and since an expansion or a collapse can be fully described by a pair of simplices, the number of bits required to encode an expansion or a collapse is $O(p\log(n+p))$. Hence, in order to guess $p+q$ moves, it suffices for $\mathbb{M}$ to guess $O(p^2 \cdot \log(n+p))$ bits in total since $q \leq p$. Now, assuming $n, p \geq 2$, we have

$$p^2\log(n+p) \leq p^2\log(np) = p^2\log(n) + p^2\log(p) \leq p^2(1+\log(p))\log(n).$$

Hence, for sufficiently large n and p, the number of bits guessed by $\mathbb{M}$ is bounded by a function of the form $f(p)\log n$. Thus, $\mathbb{M}$ is a p-restricted Turing machine, and Ordered Erasibility Expansion Height is in $\mathbf{W}[\mathbf{P}]$. ◀

▶ **Theorem 14.** Erasibility Expansion Height *is in* ***W[P].***

Proof. Assume that there exists a simple homotopy that takes K to a 1-complex using no more than p expansions. The Turing machine $\mathbb{M}$ needs to generate one such sequence. Below, we show that, in order to achieve this, $\mathbb{M}$ does not have to guess an entire simple homotopy sequence, but only a subsequence, and the remaining part of the sequence can be found deterministically by $\mathbb{M}$.

Given a 2-dimensional complex K with n faces, $\mathbb{M}$ first nondeterministically guesses p expansions and p collapses, and the order in which they are to be executed. These moves are referred to in the following as *prescribed moves*. While these moves are meant to appear in a specified order, they need not appear consecutively. The moves that are not prescribed are computed deterministically by $\mathbb{M}$. A simple homotopy sequence of K, that takes K to a 1-complex, in which all the prescribed moves occur as a subsequence, is called a sequence *compatible* with the prescribed moves. By assumption, there exists a set of prescribed moves for which a compatible sequence exists.

In order to give a description of $\mathbb{M}$, we introduce some additional terminology. Let $\mathcal{S}_q^j$ be an intermediate simple homotopy sequence computed by $\mathbb{M}$, such that the first j prescribed moves guessed by $\mathbb{M}$ form a subsequence of $\mathcal{S}_q^j$, q is the total number of moves in $\mathcal{S}_q^j$, and $\mathcal{S}_q^j$ is a prefix of a set of sequences $\mathcal{S}$ compatible with the prescribed set of moves. Let $K \nearrow\!\!\searrow K_q^j$ be the complex obtained by executing the moves in $\mathcal{S}_q^j$. Then, a collapse $\searrow_\sigma^\tau$ in K_q^j is *valid* for this prefix if appending the collapse still leaves a compatible prefix. That is, $\mathcal{S}_q^j$ appended with the collapse $\searrow_\sigma^\tau$ (giving $\mathcal{S}_{q+1}^j$) continues to be a prefix of at least one compatible sequence $\mathcal{S} \in \mathcal{S}$. A collapse that is not valid is said to be *forbidden*.

Note that labelling vertices of a complex C by natural numbers determines a lexicographic order on the simplices of C. The lexicographic order $<_C$ on simplices of C can be extended to a lexicographic order $\prec$ on collapses as follows: If $(\searrow_\sigma^\tau), (\searrow_\alpha^\beta)$ are two collapses in C, then $(\searrow_\sigma^\tau) \prec (\searrow_\alpha^\beta)$ if $\sigma <_C \alpha$.

The Turing machine $\mathbb{M}$ for deciding Erasibility Expansion Height can be described as follows:

1. Guess $2p$ prescribed moves non-deterministically.
2. Execute 2-collapses in lexicographic order until no more 2-collapses are valid.
3. Repeat until all prescribed moves have been executed:
   a. Execute the next prescribed move.
   b. Execute 2-collapses in lexicographic order until no more collapses are valid.

Let $\mathcal{S}$ be a sequence compatible with an ordered set of prescribed moves X (of cardinality $2p$), and let $\mathcal{S}_\mathbb{M}$ be a simple homotopy sequence computed by $\mathbb{M}$ as above such that X is a subsequence of $\mathcal{S}_\mathbb{M}$. Now, let σ be a free edge associated with a 2-collapse $\searrow_\sigma^\tau$ in $\mathcal{S}_q^j$ for some j and q, where $\mathcal{S}_q^j$ is a subsequence of $\mathcal{S}_\mathbb{M}$. If there exist future prescribed moves including cofaces of σ, then the next prescribed move including cofaces of σ that is not an expansion involving σ is denoted by m_1. Similarly, if there exist future prescribed moves including cofaces of τ, then the next prescribed move including cofaces of τ that is not an expansion involving τ is denoted by m_2. Note that m_2 cannot come before m_1 but we may have $m_1 = m_2$. Then, the 2-collapse $\searrow_\sigma^\tau$ is forbidden if and only if m_1 exists and is not preceded by a future prescribed expansion involving σ or m_2 exists, and is not preceded by a future prescribed expansion involving τ. It follows that, for each free edge, checking if a collapse is forbidden (or valid) can be done deterministically in time polynomial in p. To see this note that the most expensive atomic operation is to check if a simplex (of dimension 1 or 2) is a face of a simplex that is at most p dimensional, and the number of prescribed moves is at most $2p$. Altogether, the set of valid collapses can be computed in time polynomial in n and p, which can also be lexicographically ordered in polynomial time.

Finally, let K' denote the complex obtained from K by the sequence $\mathcal{S}_\mathbb{M}$. Then, the following claim establishes the effectiveness of the greedy strategy employed by $\mathbb{M}$.

▷ Claim 15. If there exists a simple homotopy sequence with at most p expansions that takes K to a 1-complex, then there exists an execution branch of the Turing machine that terminates successfully, i.e., the complex K' obtained by $\mathbb{M}$ is a 1-complex.

Proof. Let $\mathcal{S}$ be a simple homotopy sequence with p expansions that takes K to a 1-complex. Let X_e be the ordered set of expansions in $\mathcal{S}$. Thus, $|X_e| = p$. Moreover, let X_e^+ (X_c^+) denote the d-expansions (d-collapses) in $\mathcal{S}$ with $d > 2$. As in Theorem 13, by a simple inductive argument starting from the highest dimension it can be shown that $|X_c^+| = |X_e^+|$. To the p expansions of $\mathcal{S}$, we associate a set X_c of collapses of $\mathcal{S}$ as follows: If $|X_c^+| < p$, then let X_c^- be an arbitrary set of d-collapses in $\mathcal{S}$ with $d \leq 2$, and $|X_c^-| = p - |X_c^+|$. Now, let $X_c = X_c^+ \cup X_c^-$, so that $|X_c| = p$. Finally, let the ordered set X of prescribed moves be the set containing all elements of $X_c \cup X_e$ seen as a subsequence of $\mathcal{S}$.

We assume that the non-deterministic Turing machine $\mathbb{M}$ correctly guesses the specified sequence of prescribed moves X. It now suffices to show the following claim about the sequence $\mathcal{S}_{\mathbb{M}}$. ◁

▷ Claim 16. $\mathcal{S}_{\mathbb{M}}$ is compatible with the prescribed moves X, and $\mathcal{S}_{\mathbb{M}}$ takes K to a 1-complex if $\mathcal{S}$ takes K to a 1-complex.

Proof. Let $K \nearrow\!\!\searrow K^j$ denote the complex obtained from $\mathcal{S}$ after executing the j-th prescribed move in $\mathcal{S}$. We show that there exists a complex $K^j_{\mathbb{M}}$ obtained from $\mathcal{S}_{\mathbb{M}}$ after executing the j-th prescribed move in $\mathcal{S}_{\mathbb{M}}$. Also, let T^j ($T^j_{\mathbb{M}}$) denote the set of 2-simplices of K^j ($K^j_{\mathbb{M}}$).

First observe that, $K^0_{\mathbb{M}} = K$ exists and that $T^0_{\mathbb{M}} = T^0$. We now show that the following claim is inductively true: $T^j_{\mathbb{M}} \subset T^j$ for all $j \in [1, 2p]$. Suppose we make the induction hypothesis that $T^{j-1}_{\mathbb{M}} \subset T^{j-1}$ for some $j \in [1, 2p]$. Then, the set of forbidden collapses for $\mathcal{S}$ and $\mathcal{S}_{\mathbb{M}}$ are the same until the j-th move in X can be reached. Let τ_1 be the first 2-face of K^{j-1} that is removed as part of a 2-collapse after $j-1$ prescribed moves have been executed in $\mathcal{S}$. Without loss of generality, assume that the 2-collapse that removes τ_1 is non-prescribed. Then, there exists an edge $\sigma_1 \subset \tau_1$ such that τ_1 is the unique coface of σ_1 in K^{j-1}. By induction hypothesis, since $T^{j-1}_{\mathbb{M}} \subset T^{j-1}$ the same is also true for $K^{j-1}_{\mathbb{M}}$. Since $\mathbb{M}$ greedily removes every valid collapse it can (in lexicographic order), at some appropriate lexicographic index, τ_1 is also removed from $K^{j-1}_{\mathbb{M}}$ (possibly along with σ_1). Now, let $\tau_1, \tau_2, \ldots, \tau_{q-1}$ be the first $q-1$ 2-faces removed from K^{j-1} (as part of non-prescribed collapses). Assume that $\tau_1, \tau_2, \ldots, \tau_{q-1}$ have also been removed from $K^{j-1}_{\mathbb{M}}$. By the same reasoning as before, if τ_q is the q-th face to be removed from $K^{j-1}_{\mathbb{M}}$ (as part of non-prescribed collapses), then τ_q may also be removed from $\mathcal{S}_{\mathbb{M}}$ as part of a valid collapse. Hence, by induction, $T^j_{\mathbb{M}} \subset T^j$ for all $j \in [1, 2p]$.

Finally, since by assumption, K^{2p} collapses to a 1-complex, by applying arguments analogous to the induction above, the same is true for $K^{2p}_{\mathbb{M}}$ since $T^{2p}_{\mathbb{M}} \subset T^{2p}$. ◁

Since given a 2-complex with n faces, $\mathbb{M}$ non-deterministically guesses $2p$ moves, as in Theorem 13, the number of bits guessed by $\mathbb{M}$ is bounded by $f(p)\log(n)$, where $f(p) = O(p^2(1+\log(p)))$. Hence, $\mathbb{M}$ is a p-restricted Turing machine, and Erasibility Expansion Height is in $\mathbf{W}[\mathbf{P}]$. ◀

6 NP-completeness of Erasibility 3-Expansion Height

Note that the parametrized reduction from Axiom Set to Erasibility Expansion Height (and Ordered Erasibility Expansion Height) is also a polynomial-time reduction (or Karp reduction) from Axiom Set to Erasibility 3-Expansion Height (and Ordered Erasibilty 3-Expansion Height), since the complexity of reduction is independent of the parameter p and depends only on the size of the input complex. This observation leads us to the following result.

▶ **Theorem 17.** *The decision problems* Erasibility 3-Expansion Height *and* Ordered Erasibilty 3-Expansion Height *are* ***NP****-hard.*

Proof. Since the Axiom Set problem is known to be **NP**-hard [9], it follows that Erasibility 3-Expansion Height and Ordered Erasibilty 3-Expansion Height are also **NP**-hard. ◀

For the rest of the section, we assume that K is a 2-complex K with n faces and m vertices. The total number of simplices that one can encounter in any simple homotopy

sequence of K using only 3-expansions is at most $M = O(m^4)$. (Note that the ground set of K is fixed since we do not allow 1-expansions). Hence, the total number of *elementary moves* that may be available at any given point in the sequence is bounded by $O(M)$. That is, p itself is bounded by $O(M)$.

▶ **Theorem 18.** Erasibility 3-Expansion Height *is in* ***NP****.*

Proof. The non-deterministic algorithm $\mathbb{M}$ for deciding Erasibility 3-Expansion Height first guesses at each point in the simple homotopy sequence starting with K, one elementary move (out of at most $O(M)$ available moves), and constructs a new complex from the move. The total number of moves made by $\mathbb{M}$ is bounded by $(\frac{n+2p-1}{2})$. Finally, $\mathbb{M}$ checks if the final complex is a 1-complex. ◀

▶ **Theorem 19.** Ordered Erasibilty 3-Expansion Height *is in* ***NP****.*

Proof. The non-deterministic algorithm $\mathbb{M}$ for deciding Ordered Erasibilty 3-Expansion Height first guesses at most p 3-expansions followed by an equal number of 3-collapses, resulting in a 2-complex K' with n faces. From Remark 12, the erasability of K' can be deterministically checked in time polynomial in n, proving the claim. ◀

References

1 J. J. Andrews and M. L. Curtis. Free groups and handlebodies. *Proc. Amer. Math. Soc.*, 16:192–195, 1965. doi:10.2307/2033843.

2 Jonathan A. Barmak. *Algebraic topology of finite topological spaces and applications*, volume 2032 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2011. doi:10.1007/978-3-642-22003-6.

3 Ulrich Bauer and Abhishek Rathod. Hardness of approximation for Morse matching. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2663–2674. SIAM, Philadelphia, PA, 2019. doi:10.1137/1.9781611975482.165.

4 Gilbert Baumslag, Alexei G. Myasnikov, and Vladimir Shpilrain. Open problems in combinatorial group theory. Second edition. In *Combinatorial and geometric group theory (New York, 2000/Hoboken, NJ, 2001)*, volume 296 of *Contemp. Math.*, pages 1–38. Amer. Math. Soc., Providence, RI, 2002. doi:10.1090/conm/296/05067.

5 V. V. Borisov. Simple examples of groups with unsolvable word problem. *Mat. Zametki*, 6:521–532, 1969.

6 Benjamin A. Burton, Thomas Lewiner, João Paixão, and Jonathan Spreer. Parameterized complexity of discrete Morse theory. *ACM Trans. Math. Software*, 42(1):Art. 6, 24, 2016. doi:10.1145/2738034.

7 R. G. Downey and M. R. Fellows. *Parameterized complexity*. Monographs in Computer Science. Springer-Verlag, New York, 1999. doi:10.1007/978-1-4612-0515-9.

8 J. Flum and M. Grohe. *Parameterized complexity theory*. Texts in Theoretical Computer Science. An EATCS Series. Springer-Verlag, Berlin, 2006.

9 Michael R. Garey and David S. Johnson. *Computers and intractability*. W. H. Freeman and Co., San Francisco, Calif., 1979. A guide to the theory of NP-completeness, A Series of Books in the Mathematical Sciences.

10 Masahiro Hachimori. *Combinatorics of constructible complexes*. PhD thesis, Tokyo University, 2000. URL: https://pdfs.semanticscholar.org/81ee/09ed08ca6c7487e8e18639a7a05c110781c3.pdf.

11 Allen Hatcher. *Algebraic topology*. Cambridge University Press, Cambridge, 2002.

12 Cynthia Hog-Angeloni and Wolfgang Metzler, editors. *Two-dimensional homotopy and combinatorial group theory*, volume 197 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1993. doi:10.1017/CBO9780511629358.

13 Sergei Matveev. *Algorithmic topology and classification of 3-manifolds*, volume 9 of *Algorithms and Computation in Mathematics*. Springer, Berlin, second edition, 2007.

14 James R. Munkres. *Elements of algebraic topology*. Addison-Wesley Publishing Company, Menlo Park, CA, 1984.

15 Alexei D. Myasnikov, Alexei G. Myasnikov, and Vladimir Shpilrain. On the Andrews-Curtis equivalence. In *Combinatorial and geometric group theory (New York, 2000/Hoboken, NJ, 2001)*, volume 296 of *Contemp. Math.*, pages 183–198. Amer. Math. Soc., Providence, RI, 2002. doi:10.1090/conm/296/05074.

16 J. Paixão and J. Spreer. Random collapsibility and 3-sphere recognition, 2015. Preprint, 18 pages, 6 figures. arXiv:1509.07607.

17 Martin Tancer. Recognition of collapsible complexes is NP-complete. *Discrete Comput. Geom.*, 55(1):21–38, 2016. doi:10.1007/s00454-015-9747-1.

18 J. H. C. Whitehead. On incidence matrices, nuclei and homotopy types. *Ann. of Math. (2)*, 42:1197–1239, 1941. doi:10.2307/1970465.

19 J. H. C. Whitehead. Simple homotopy types. *Amer. J. Math.*, 72:1–57, 1950. doi:10.2307/2372133.

20 Perrin Wright. Group presentations and formal deformations. *Trans. Amer. Math. Soc.*, 208:161–169, 1975. doi:10.2307/1997282.

21 E. C. Zeeman. On the dunce hat. *Topology*, 2:341–358, 1964. doi:10.1016/0040-9383(63)90014-4.

C

PAPER III: FAST ALGORITHMS FOR MINIMUM CYCLE BASIS AND MINIMUM HOMOLOGY BASIS.

Minimum cycle bases in graphs have several applications from analysis of electrical networks and biological pathways to surface reconstruction and graph drawing. Cycle bases of small size offer a compact description that can be particularly advantageous from an application viewpoint. For this reason, the problem of computing a minimum cycle basis is a well-studied problem.

In topological data analysis, "holes" of different dimensions in a geometric dataset are regarded as "features" of the data. Algebraic topology offers a rigorous language to formalize our intuitive picture of holes, tunnels and cavities in these geometric objects. Specifically, a basis for the first homology group H_1 can be taken as a representative for the one-dimensional holes in the geometric object. The advantages of using minimum homology bases are twofold: firstly, one can bring geometry in picture by assigning appropriate weights to edges if these weights also correspond to distances between vertices, and secondly, cycles with smaller support are easier to understand and analyze, especially visually. In this article, we focus solely on the bases of the first homology group since the problem of computing a shortest basis for higher homology groups with $\mathbb{Z}_2$ coefficients was shown to be **NP**-hard by Chen and Freedman. Hence, going forward, by homology basis, we mean homology basis for H_1.

The article is structured as follows. We begin with an introduction in Section 1. In Section 2.1, we recall some structural results regarding minimum cycle bases from the works of Horton, de Pina, and Amaldi et al. In Section 2.2, we recall some of the state-of-the-art results from literature concerning column rank profile computation. In Section 3, we devise an algorithm for computing a minimum cycle basis of a graph. This algorithm has a particularly simple high-level description. In Section 4, we prove a structural result concerning minimum homology bases for simplicial complexes. In Section 5, we devise two algorithms for computing minimum homology bases for simplicial complexes, which improve upon the state-of-the-art algorithms for this problem by Dey et al. We end the article with a discussion section.

STATEMENT OF INDIVIDUAL CONTRIBUTION. As soon as I had the idea for this paper, I discussed it with Uli Bauer, who was supportive of the project, and asked me to write it up. Since I am the sole author of this article, I am fully responsible for all parts.

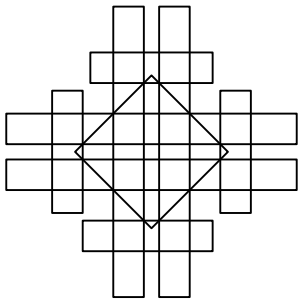
36th International Symposium on Computational Geometry

SoCG 2020, June 23–26, 2020, Zürich, Switzerland (Virtual Conference)

Edited by

Sergio Cabello

Danny Z. Chen



Editors

Sergio Cabello
University of Ljubljana, Ljubljana, Slovenia
sergio.cabello@fmf.uni-lj.si

Danny Z. Chen
University of Notre Dame, Indiana, USA
dchen@nd.edu

ACM Classification 2012
Theory of computation → Computational geometry; Theory of computation → Design and analysis of algorithms; Mathematics of computing → Combinatorics; Mathematics of computing → Graph algorithms

ISBN 978-3-95977-143-6

Published online and open access by
Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing, Saarbrücken/Wadern, Germany. Online available at https://www.dagstuhl.de/dagpub/978-3-95977-143-6.

Publication date
June, 2020

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at https://portal.dnb.de.

Digital Object Identifier: 10.4230/LIPIcs.SoCG.2020.0

Fast Algorithms for Minimum Cycle Basis and Minimum Homology Basis

Abhishek Rathod
Department of Mathematics, Technical University of Munich (TUM),
Boltzmannstr. 3, 85748 Garching b. München, Germany
abhishek.rathod@tum.de

Abstract

We study the problem of finding a minimum homology basis, that is, a shortest set of cycles that generates the 1-dimensional homology classes with $\mathbb{Z}_2$ coefficients in a given simplicial complex K. This problem has been extensively studied in the last few years. For general complexes, the current best deterministic algorithm, by Dey et al. [8], runs in $O(N^\omega + N^2 g)$ time, where N denotes the number of simplices in K, g denotes the rank of the 1-homology group of K, and ω denotes the exponent of matrix multiplication. In this paper, we present two conceptually simple randomized algorithms that compute a minimum homology basis of a general simplicial complex K. The first algorithm runs in $\tilde{O}(m^\omega)$ time, where m denotes the number of edges in K, whereas the second algorithm runs in $O(m^\omega + Nm^{\omega-1})$ time.

We also study the problem of finding a minimum cycle basis in an undirected graph G with n vertices and m edges. The best known algorithm for this problem runs in $O(m^\omega)$ time. Our algorithm, which has a simpler high-level description, but is slightly more expensive, runs in $\tilde{O}(m^\omega)$ time.

2012 ACM Subject Classification Theory of computation → Computational geometry; Mathematics of computing → Algebraic topology

Keywords and phrases Computational topology, Minimum homology basis, Minimum cycle basis, Simplicial complexes, Matrix computations

Digital Object Identifier 10.4230/LIPIcs.SoCG.2020.64

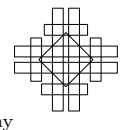
Funding This research has been supported by the DFG Collaborative Research Center SFB/TRR 109 "Discretization in Geometry and Dynamics".

Acknowledgements The author would like to thank Ulrich Bauer and Michael Lesnick for valuable discussions, and anonymous reviewers for their useful comments.

1 Introduction

Minimum cycle bases in graphs have several applications, for instance, in analysis of electrical networks, analysis of chemical and biological pathways, periodic scheduling, surface reconstruction and graph drawing. Also, algorithms from diverse application domains like electrical circuit theory and structural engineering require cycle basis computation as a preprocessing step. Cycle bases of small size offer a compact description that is advantageous from a mathematical as well as from an application viewpoint. For this reason, the problem of computing a minimum cycle basis has received a lot of attention, both in its general setting as well as in special classes of graphs such as planar graphs, sparse graphs, dense graphs, network graphs, and so on. We refer the reader to [15] for a comprehensive survey.

In topological data analysis, "holes" of different dimensions in a geometric dataset constitute "features" of the data. Algebraic topology offers a rigorous language to formalize our intuitive picture of holes in these geometric objects. More precisely, a basis for the first homology group H_1 can be taken as a representative of the one-dimensional holes in the geometric object. The advantages of using minimum homology bases are twofold: firstly, one can bring geometry in picture by assigning appropriate weights to edges, and secondly,




smaller cycles are easier to understand and analyze, especially visually. We focus solely on the bases of the first homology group since the problem of computing a shortest basis for higher homology groups with $\mathbb{Z}_2$ coefficients was shown to be NP-hard by Chen and Freedman [5].

2 Background and Preliminaries

2.1 Cycle Basis

Let $G = (V, E)$ be a connected graph. A subgraph of G which has even degree for each vertex is called a *cycle* of G. A cycle is called *elementary* if the set of edges form a connected subgraph in which each vertex has degree 2. We associate an incidence vector C, indexed on E, to each cycle, so that $C_e = 1$ if e is an edge of the cycle, and $C_e = 0$ otherwise. The set of incidence vectors of cycles forms a vector space over $\mathbb{Z}_2$, called the *cycle space* of G. It is a well-known fact that for a connected graph G, the cycle space is of dimension $|E| - |V| + 1$. Throughout, we use ν to denote the dimension of the cycle space of a graph. A basis of the cycle space, that is, a maximal linearly independent set of cycles is called a *cycle basis*.

Suppose that the edges of G have non-negative weights. Then, the weight of a cycle is the sum of the weights of its edges, and the weight of a cycle basis is the sum of the weights of the basis elements. The problem of computing a cycle basis of minimum weight is called the *minimum cycle basis* problem. Since we assume all edge weights to be non-negative, there always exists a minimum cycle basis of elementary cycles, allowing us to focus on minimum cycle basis comprising entirely of elementary cycles.

A simple cycle C is *tight* if it contains a shortest path between every pair of points in C. We denote the set of all tight cycles in the graph by $\mathcal{T}$. Tight cycles are sometimes also referred to as *isometric* cycles [1, 15]. Tight cycles play an important role in designing algorithms for minimum cycle basis, owing to the following theorem by Horton.

▶ **Theorem 1** (Horton [13]). *A minimum cycle basis $\mathcal{M}$ consists only of tight cycles.*

A key structural property about minimum cycle bases was proved by de Pina.

▶ **Theorem 2** (de Pina [7]). *Cycles $C_1 \ldots, C_\nu$ form a minimum cycle basis if there are vectors $S_1, \ldots, S_\nu$ such that for all i, $1 \leq i \leq \nu$, the following hold:*
Prefix Orthogonality: *$\langle C_j, S_i \rangle = 0$ for all $1 \leq j \leq i$.*
Non-Orthogonality: *$\langle C_i, S_i \rangle = 1$.*
Shortness: *C_i is a minimum weight cycle in $\mathcal{T}$ with $\langle C_i, S_i \rangle = 1$.*

The vectors $S_1, \ldots, S_\nu$ in Theorem 2 are called *support vectors*. The recent line of algorithmic work [1, 7, 16, 17, 18] on the minimum cycle basis problem rely on Theorem 2. In fact, these algorithms may all be seen as refinements of the algorithm by de Pina, see Algorithm 1.

Algorithm 1 De Pina's Algorithm for computing a minimum cycle basis.

```
Initialize S_i to the i-th unit vector e_i for 1 ≤ i ≤ ν
for i ← 1, ..., ν do
    Compute a minimum weight cycle C_i with ⟨C_i, S_i⟩ = 1.
    for j ← i + 1, ..., ν do
        S_j = S_j + ⟨C_i, S_j⟩ S_i
    end for
end for
Return {C_1, ..., C_ν}.
```

Algorithm 1 works by inductively maintaining a set of support vectors $\{S_i\}$ so that the conditions of Theorem 2 are satisfied when the algorithm terminates. In particular, Lines 4 and 5 of the algorithm ensure that the set of vectors S_j for $j > i$ are orthogonal to vectors $C_1, \ldots, C_i$. Updating the vectors S_j as outlined in Lines 4 and 5 of Algorithm 1 takes time $O(m^3)$ time in total. Using a divide and conquer procedure for maintaining S_j, Kavitha et al. [17] improved the cost of maintaining the support vectors to $O(m^\omega)$. See Algorithm 2.

Algorithm 2 Divide and conquer procedure for fast computation of support vectors by Kavitha et al. [17].

1: Initialize S_i to the i-th unit vector e_i for $1 \leq i \leq \nu$.
2: MinCycleBasis$(1, \nu)$.

3: **procedure** MinCycleBasis(ℓ, u)
4: **if** $\ell = u$ **then**
5: Compute a minimum weight cycle C_ℓ with $\langle C_\ell, S_\ell \rangle = 1$.
6: **else**
7: $q \leftarrow \lfloor (\ell+u)/2 \rfloor$.
8: MinCycleBasis(ℓ, q).
9: $\mathbf{C} \leftarrow [C_\ell, \ldots, C_q]$.
10: $\mathbf{W} \leftarrow (\mathbf{C}^T[S_\ell, \ldots, S_q])^{-1}\mathbf{C}^T[S_{q+1}, \ldots, S_u]$.
11: $[S_{q+1}, \ldots, S_u] \leftarrow [S_{q+1}, \ldots, S_u] + [S_\ell, \ldots, S_q]\mathbf{W}$.
12: MinCycleBasis$(q+1, u)$.
13: **end if**
14: **end procedure**
15: Return $\{C_1, \ldots, C_\nu\}$.

▶ **Lemma 3** (Lemma 5.6 in [15]). *The total number of arithmetic operations performed in lines 9 to 11 of Algorithm 2 is $O(m^\omega)$. That is, the support vectors satisfying conditions of Theorem 2 can be maintained in $O(m^\omega)$ time.*

Finally, in [1], Amaldi et al. designed an $O(m^\omega)$ time algorithm for computing a minimum cycle basis by improving the complexity of Line 5 of Algorithm 2 to $o(m^\omega)$ (from $O(m^2 n)$ in [17]), while using the $O(m^\omega)$ time divide-and-conquer template for maintaining the support vectors as presented in Algorithm 2. The $o(m^\omega)$ procedure for Line 3 is achieved by performing a Monte Carlo binary search on the set of tight cycles (sorted by weight) to find a minimum weight cycle C_i that satisfies $\langle C_i, S_i \rangle = 1$. An efficient binary search is made possible on account of the following key structural property about tight cycles.

▶ **Theorem 4** (Amaldi et al. [1]). *The total length of the tight cycles is at most $n\nu$.*

Amaldi et al. [1] also show that there exists an $O(nm)$ algorithm to compute the set of all the tight cycles of an undirected graph G. See Sections 2 and 3 of [1] for details about Amaldi et al.'s algorithm.

2.2 Matrix operations

The *column rank profile* (respectively *row rank profile*) of an $m \times n$ matrix A with rank r, is the lexicographically smallest sequence of r indices $[i_1, i_2, \ldots, i_r]$ (respectively $[j_1, j_2, \ldots, j_r]$) of linearly independent columns (respectively rows) of A. Suppose that $\{a_1, a_2, \ldots, a_n\}$ represent the columns of A. Then, following Busaryev et al. [3], we define the *earliest basis* of A as the set of columns $\mathcal{E}(A) = \{a_{i_1}, a_{i_2}, \ldots, a_{i_r}\}$.

It is well-known that classical Gaussian elimination can be used to compute rank profile in $O(nmr)$ time. The current state-of-the-art deterministic matrix rank profile algorithms run in $O(mnr^{\omega-2})$ time.

▶ **Theorem 5** ([10, 14]). *There is a deterministic $O(mnr^{\omega-2})$ time algorithm to compute the column rank profile of an $m \times n$ matrix A.*

In case of randomized algorithms, Cheung, Kwok and Lau [6] presented a breakthrough Monte Carlo algorithm for rank computation that runs in $(\mathsf{nnz}(A) + r^\omega)^{1+o(1)}$ time, where $o(1)$ in the exponent captures some missing multiplicative $\log n$ and $\log m$ factors, and $\mathsf{nnz}(A)$ denotes the number of nonzero entries in A. Equivalently, the complexity for Cheung et al.'s algorithm can also be written as $\tilde{O}(\mathsf{nnz}(A) + r^\omega)$. The notation $\tilde{O}(\cdot)$ is often used in literature to hide small polylogarithmic factors in time bounds. While the algorithm by Cheung et al. also computes r linearly independent columns of A, the columns may not correspond to the column rank profile. However, building upon Cheung et al.'s work, Storjohann and Yang established the following result.

▶ **Theorem 6** (Storjohann and Yang [19, 20, 21]). *There exists a Monte Carlo algorithm for computing row (resp. column) rank profile of a matrix A that runs in $(\mathsf{nnz}(A) + r^\omega)^{1+o(1)}$ time. The failure probability of this algorithm is $1/2$.*

Once again, the $o(1)$ in the exponent captures some missing multiplicative $\log n$ and $\log m$ factors, see [19], and hence the complexity can also be written as $\tilde{O}(\mathsf{nnz}(A) + r^\omega)$.

2.3 Homology

In this work, we restrict our attention to simplicial homology with $\mathbb{Z}_2$ coefficients. For a general introduction to algebraic topology, we refer the reader to [12]. Below we give a brief description of homology over $\mathbb{Z}_2$.

Let K be a connected simplicial complex. We will denote by $K^{(p)}$ the set of p-dimensional simplices in K, and n_p the number of p-dimensional simplices in K. Also, the p-dimensional skeleton of K will be denoted by K_p. In particular, the 1-skeleton of K (which is an undirected graph) will be denoted by K_1.

We consider formal sums of simplices with $\mathbb{Z}_2$ coefficients, that is, sums of the form $\sum_{\sigma \in K^{(p)}} a_\sigma \sigma$, where each $a_\sigma \in \{0, 1\}$. The expression $\sum_{\sigma \in K^{(p)}} a_\sigma \sigma$ is called a *p-chain*. Since chains can be added to each other, they form an Abelian group, denoted by $\mathsf{C}_p(K)$. Since we consider formal sums with coefficients coming from $\mathbb{Z}_2$, which is a field, $\mathsf{C}_p(K)$, in this case, is a vector space of dimension n_p over $\mathbb{Z}_2$. The p-simplices in K form a (natural) basis for $\mathsf{C}_p(K)$. This establishes a natural one-to-one correspondence between elements of $\mathsf{C}_p(K)$ and subsets of $K^{(p)}$. Thus, associated with each chain is an incidence vector v, indexed on $K^{(p)}$, where $v_\sigma = 1$ if σ is a simplex of v, and $v_\sigma = 0$ otherwise. The *boundary* of a p-simplex is a $(p-1)$-chain that corresponds to the set of its $(p-1)$-faces. This map can be linearly extended from p-simplices to p-chains, where the boundary of a chain is the $\mathbb{Z}_2$-sum of the boundaries of its elements. Such an extension is known as the *boundary homomorphism*, and denoted by $\partial_p : \mathsf{C}_p(K) \to \mathsf{C}_{p-1}(K)$. A chain $\zeta \in \mathsf{C}_p(K)$ is called a *p-cycle* if $\partial_p \zeta = 0$, that is, $\zeta \in \ker \partial_p$. The group of p-dimensional cycles is denoted by $\mathsf{Z}_p(K)$. As before, since we are working with $\mathbb{Z}_2$ coefficients, $\mathsf{Z}_p(K)$ is a vector space over $\mathbb{Z}_2$. A chain $\eta \in \mathsf{C}_p(K)$ is said to be a p-boundary if $\eta = \partial_{p+1} c$ for some chain $c \in \mathsf{C}_{p+1}(K)$, that is, $\eta \in \operatorname{im} \partial_{p+1}$. The group of p-dimensional boundaries is denoted by $\mathsf{B}_p(K)$. In our case, $\mathsf{B}_p(K)$ is also a vector space, and in fact a subspace of $\mathsf{C}_p(K)$.

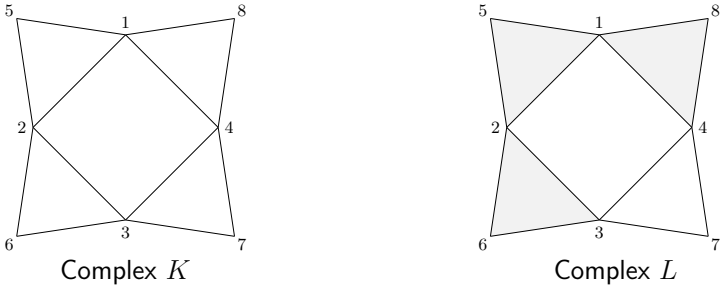


Figure 1 Consider complexes K and L in the figure above with unit weights on the edges. Since K has no 2-simplices, its 1-skeleton K_1 is identical to K itself. The set of cycles $\mathcal{C} = \{\{1,2,5\},\{1,4,8\},\{3,4,7\},\{2,3,6\},\{1,2,3,4\}\}$ constitutes a minimum cycle basis for the respective 1-skeletons K_1 and L_1 (viewed as graphs). The set $\mathcal{C}$ also constitutes a minimum homology basis for K. The set $\mathcal{C}' = \{\{1,2,3,4\},\{3,4,7\}\}$ constitutes a minimum homology basis for L.

Thus, we can consider the quotient space $\mathsf{H}_p(K) = \mathsf{Z}_p(K)/\mathsf{B}_p(K)$. The elements of the vector space $\mathsf{H}_p(K)$, known as the p-th *homology group* of K, are equivalence classes of p-cycles, where p-cycles are equivalent if their $\mathbb{Z}_2$-difference is a p-boundary. Equivalent cycles are said to be *homologous*. For a p-cycle ζ, its corresponding homology class is denoted by $[\zeta]$. Bases of $\mathsf{B}_p(K)$, $\mathsf{Z}_p(K)$ and $\mathsf{H}_p(K)$ are called *boundary bases*, *cycle bases*, and *homology bases* respectively. The dimension of the p-th homology group of K is called the p-th *Betti number* of K, denoted by $\beta_p(K)$. We are primarily interested in the first Betti number $\beta_1(K)$. For notational convenience, let $g = \beta_1(K)$, and denote the dimension of $\mathsf{B}_1(K)$ by b.

Using the natural bases for $\mathsf{C}_p(K)$ and $\mathsf{C}_{p-1}(K)$, the matrix $[\partial_p\sigma_1\ \partial_p\sigma_2 \cdots \partial_p\sigma_{n_p}]$ whose column vectors are boundaries of p-simplices is called the p-th boundary matrix. Abusing notation, we denote the p-th boundary matrix by ∂_p. For the rest of the paper, we use n, m and N to denote the number of vertices, edges and simplices in the complex respectively.

A set of p-cycles $\{\zeta_1, \ldots, \zeta_g\}$ is called a *homology cycle basis* if the set of classes $\{[\zeta_1], \ldots, [\zeta_g]\}$ forms a homology basis. For brevity, we abuse notation by using the term "homology basis" for $\{\zeta_1, \ldots, \zeta_g\}$. Assigning non-negative weights to the edges of K, the weight of a cycle is the sum of the weights of its edges, and the weight of a homology basis is the sum of the weights of the basis elements. The problem of computing a minimum weight basis of $\mathsf{H}_1(K)$ is called the *minimum homology basis* problem. Note that, when the input simplicial complex is a graph, the notions of homology basis and cycle basis coincide. Please refer to Figure 1 for an example.

For the special case when the input complex is a surface, Erickson and Whittlesey [11] gave a $O(N^2 \log N + gN^2 + g^3N)$-time algorithm. Recently, Borradaile et al. [2] gave an improved deterministic algorithm that runs in $O((h+c)^3 n \log n + m)$ where c denotes the number of boundary components, and h denotes the genus of the surface. For small values of c and h, the algorithm runs in nearly linear time.

For general complexes, Dey et al. [9] and Chen and Freedman [4] generalized the results by Erickson and Whittlesey [11] to arbitrary complexes. Subsequently, introducing the technique of annotations, Busaryev et al. [3] improved the complexity to $O(N^\omega + N^2 g^{\omega-1})$. More recently, Dey et al. [8] designed an $O(N^\omega + N^2 g)$ time algorithm by adapting the divide and conquer algorithm for computing a minimum cycle basis of Kavitha et al. [17] for the purpose of computing a minimum homology basis. Dey et al. also designed a randomized 2-approximation algorithm for the same problem that runs in $O(N^\omega \sqrt{N \log N})$ expected time.

3 An algorithm for computing minimum cycle basis

Given a graph $G = (V, E)$, let $\{C_1, \ldots, C_{|\mathcal{T}|}\}$ be the list of tight cycles in G sorted by weight, and let $\mathbf{M}_{\mathcal{T}}(G) = [C_1\, C_2 \ldots\, C_{|\mathcal{T}|}]$ be the matrix formed with cycles C_i as its columns. Using Theorem 4, since the total length of tight cycles is at most $n\nu$, and since each tight cycle consists of at least three edges, we have that $|\mathcal{T}| \leq \frac{n\nu}{3}$. Also, the rank of $\mathbf{M}_{\mathcal{T}}(G)$ is ν and $\mathbf{M}_{\mathcal{T}}(G)$ is a sparse matrix with $\mathsf{nnz}(\mathbf{M}_{\mathcal{T}}(G))$ bounded by $n\nu$. This sparsity is implicitly used in the design of the Monte Carlo binary search algorithm for computing minimum cycle basis, as described in [1]. We now present a simple and fast algorithm for minimum cycle basis that exploits the sparsity and the low rank of $\mathbf{M}_{\mathcal{T}}(G)$ more directly.

Algorithm 3 Algorithm for minimum cycle basis.

1: Compute the sorted list of tight cycles in G, and assemble the matrix $\mathbf{M}_{\mathcal{T}}(G)$.
2: Compute the column rank profile $[i_1, i_2, \ldots, i_\nu]$ of $\mathbf{M}_{\mathcal{T}}(G)$ using Storjohann and Yang's algorithm described in [20].
3: RETURN $\mathcal{E}(\mathbf{M}_{\mathcal{T}}(G))$.

▶ **Theorem 7.** *There is a Monte Carlo algorithm that computes the minimum cycle basis in $\tilde{O}(m^\omega)$ time, with failure probability at most $1/2$.*

Proof. The correctness of the algorithm follows immediately from Theorem 1. For instance, if $\mathcal{E}(\mathbf{M}_{\mathcal{T}}(G))$ is not a minimum cycle basis, then let k be the smallest integer such that the k-th smallest cycle in a minimum cycle basis contained in $\mathbf{M}_{\mathcal{T}}(G)$ is smaller than the k-th smallest cycle in $\mathcal{E}(\mathbf{M}_{\mathcal{T}}(G))$. Since the columns in $\mathbf{M}_{\mathcal{T}}(G)$ are sorted by weight, the existence of such a k contradicts the fact that $\mathcal{E}(\mathbf{M}_{\mathcal{T}}(G))$ is the earliest basis of $\mathbf{M}_{\mathcal{T}}(G)$.

The list of tight cycles in G can be computed in $O(nm)$ time using the algorithm described in Section 2 of [1]. Hence, Step 1 of Algorithm 3 takes $O(nm \log(nm))$ time (which in turn is same as $O(nm \log n)$ time). Moreover, using Theorem 6, the complexity of Step 2 is bounded by $\tilde{O}(n\nu + \nu^\omega)$. Since $n, \nu < m$, the complexity of Algorithm 3 is bounded by $\tilde{O}(m^\omega)$. Using Theorem 6, the failure probability of the algorithm is at most $1/2$. ◀

4 Minimum homology basis, minimum cycle basis and tight cycles

To begin with, note that since every graph is a 1-dimensional simplicial complex, the minimum cycle basis problem is a restriction of the minimum homology basis problem to instances (simplicial complexes) that have no 2-simplices. In this section, we refine this observation by deriving a closer relation between the two problems.

We assume that we are provided a complex K in which all edges are assigned non-negative weights. Given a non-negative weight $w(\sigma)$ for each edge σ, we define the weight of a cycle z as the sum of the weights of the edges, $w(z) = \sum_{\sigma \in z} w(\sigma)$. Let $\mathcal{B} = \{\eta_1, \ldots, \eta_b\}$ be a basis for the boundary vector space $\mathsf{B}_1(K)$ indexed so that $w(\eta_i) \leq w(\eta_{i+1})$, $1 \leq i < b$ (with ties broken arbitrarily). Also, let $\mathcal{H} = \{\zeta_1, \ldots, \zeta_g\}$ be a minimum homology basis of K indexed so that $w(\zeta_i) \leq w(\zeta_{i+1})$, $1 \leq i < g$ (with ties broken arbitrarily). Then, the set $\mathcal{C} = \{\eta_1, \ldots, \eta_b, \zeta_1, \ldots, \zeta_g\}$ is a cycle basis for K_1. Let $\mathcal{M}$ be a minimum cycle basis of K_1. Every element $C \in \mathcal{M}$ is homologous to a cycle $\sum_{i=1}^{g} a_i \zeta_i$ where $a_i \in \{0, 1\}$ for each i. Then, for some fixed integers p and q, $\mathcal{M} = \{B_1, \ldots, B_q, C_1, \ldots, C_p\}$ is indexed so that the elements $B_1, \ldots, B_q$ are null-homologous and the elements $C_1, \ldots, C_p$ are non-bounding cycles. Also, we have $w(B_j) \leq w(B_{j+1})$ for $1 \leq j < q$ (with ties broken arbitrarily), and $w(C_j) \leq w(C_{j+1})$ for $1 \leq j < p$ (with ties broken arbitrarily).

▶ **Lemma 8.**
1. *For every minimum homology basis,* $w(\zeta_1) = w(C_1)$.
2. *There exists a minimum homology basis* $\overline{\mathcal{H}}$ *with* ζ_1 *homologous to* C_1.

Proof. Suppose there exists a minimum homology basis with $w(\zeta_1) < w(C_1)$. Let $\zeta_1 = \sum_{i=1}^{p} a_i C_i + \sum_{j=1}^{q} b_j B_j$, where $a_i \in \{0,1\}$ for each i and $b_j \in \{0,1\}$ for each j. Since ζ_1 is a non-bounding cycle, there exists at least one i with $a_i = 1$. Let $\ell \in [1,p]$ be the largest index in the above equation with $a_\ell = 1$. Rewriting the equation, we obtain $C_\ell = \sum_{i=1}^{\ell-1} a_i C_i + \sum_{j=1}^{q} b_j B_j + \zeta_1$. Since $w(\zeta_1) < w(C_1)$ by assumption, we have $w(\zeta_1) < w(C_\ell)$ because $w(C_\ell) \geq w(C_1)$ by indexing of $\mathcal{M}$. It follows that the basis obtained by exchanging C_ℓ for ζ_1, that is, $\{B_1, \ldots, B_q, \zeta_1, C_1, \ldots, C_{\ell-1}, C_{\ell+1}, \ldots, C_p\}$ gives a smaller cycle basis than the minimum one, a contradiction.

Now, suppose there exists a minimum homology basis with $w(\zeta_1) > w(C_1)$. Let $C_1 = \sum_{i=1}^{g} a_i \zeta_i + \sum_{j=1}^{b} b_j \eta_j$. As before, since C_1 is not null-homologous, there exists at least one i with $a_i = 1$. Let $\ell \in [1,g]$ be the largest index in the above equation with $a_\ell = 1$. Then, $\zeta_\ell = \sum_{i=1}^{\ell-1} a_i \zeta_i + \sum_{j=1}^{b} b_j \eta_j + C_1$. Note that $w(\zeta_\ell) \geq w(\zeta_1)$ because of the indexing, and $w(\zeta_1) > w(C_1)$ by assumption. Therefore, the set $\{C_1, \zeta_1, \ldots, \zeta_{\ell-1}, \zeta_{\ell+1}, \ldots, \zeta_p\}$ obtained by exchanging ζ_ℓ for C_1 gives a smaller homology basis than the minimum one, a contradiction. This proves the first part of the lemma.

From the first part of the lemma, we have $w(\zeta_1) = w(C_1)$ for every minimum homology basis. Let $\mathcal{H}$ be an arbitrary minimum homology basis. Then, if C_1 is not homologous to $\zeta_1 \in \mathcal{H}$, by using basis exchange we can obtain $\overline{\mathcal{H}} = \{C_1, \zeta_1, \ldots, \zeta_{\ell-1}, \zeta_{\ell+1}, \ldots, \zeta_p\}$, which is the minimum homology basis with its first element homologous to C_1, and having the same weight as $w(C_1)$, proving the claim. ◀

We now prove a theorem which allows us to harness fast algorithms for minimum cycle basis in service of improving time complexity of algorithms for minimum homology basis.

▶ **Theorem 9.** *Given a simplicial complex* K*, and a minimum cycle basis* $\mathcal{M} = \{B_1, \ldots, B_q, C_1, \ldots, C_p\}$ *of* K_1*, there exists a minimum homology basis* $\overline{\mathcal{H}}$ *of* K*, and a set* $\{C_{i_1}, \ldots, C_{i_g}\} \subset \{C_1, \ldots, C_p\} \subset \mathcal{M}$ *such that, for every* $k \in [1,g]$*, we have* C_{i_k} *homologous to a cycle spanned by* $\zeta_1, \ldots, \zeta_k$*, and* $w(C_{i_k}) = w(\zeta_k)$*. Moreover,* $i_1 = 1$*, and* i_k *for* $k > 1$ *is the smallest index for which* C_{i_k} *is not homologous to any cycle spanned by* $\{C_{i_1}, \ldots, C_{i_{k-1}}\}$*. In particular, the set* $\{C_{i_1}, \ldots, C_{i_g}\} \subset \mathcal{M}$ *constitutes a minimum homology basis of* K*.*

Proof. The key argument is essentially the same as for the proof of Lemma 8. Nonetheless, we present it here for the sake of completeness. We shall prove the claim by induction. Lemma 8 covers the base case. By induction hypothesis, there is an integer k, and a minimum homology basis $\mathcal{H} = \{\zeta_1, \ldots, \zeta_g\}$, for which, vectors $\{C_{i_1}, \ldots, C_{i_k}\} \subseteq \{C_1, \ldots, C_p\}$ are such that, for every $j \in [1,k]$, we have C_{i_j} homologous to a cycle spanned by $\zeta_1, \ldots, \zeta_j$, and $w(C_{i_j}) = w(\zeta_j)$. Let i_{k+1} be the smallest index for which $C_{i_{k+1}} \in \{C_1, \ldots, C_p\}$ is not homologous to any cycle spanned by $\{C_{i_1}, \ldots, C_{i_k}\}$.

Suppose that $w(\zeta_{k+1}) < w(C_{i_{k+1}})$. Let $\zeta_{k+1} = \sum_{i=1}^{p} a_i C_i + \sum_{j=1}^{q} b_j B_j$. Let $\ell \in [1,p]$ be the largest index in the above equation with $a_\ell = 1$. Then, $C_\ell = \sum_{i=1}^{\ell-1} a_i C_i + \sum_{j=1}^{q} b_j B_j + \zeta_{k+1}$. From the induction hypothesis, we can infer that $\ell \geq i_{k+1}$, and hence $w(C_\ell) \geq w(C_{i_{k+1}})$ by indexing of $\mathcal{M}$. Thus, if $w(\zeta_{k+1}) < w(C_{i_{k+1}})$, then we have $w(\zeta_{k+1}) < w(C_\ell)$. It follows that, $\{B_1, \ldots, B_q, \zeta_{k+1}, C_1, \ldots, C_{\ell-1}, C_{\ell+1}, \ldots, C_p\}$ obtained by exchanging C_ℓ for ζ_{k+1} gives a smaller cycle basis than the minimum one, contradicting the minimality of $\mathcal{H}$.

Now, suppose that $w(\zeta_{k+1}) > w(C_{i_{k+1}})$. Let $C_{i_{k+1}} = \sum_{i=1}^{g} a_i \zeta_i + \sum_{j=1}^{b} b_j \eta_j$. Let $\ell \in [1,g]$ be the largest index in the above equation with $a_\ell = 1$. Rewriting the equation, we obtain $\zeta_\ell = \sum_{i=1}^{\ell-1} a_i \zeta_i + \sum_{j=1}^{b} b_j \eta_j + C_{i_{k+1}}$. Again, using the induction hypothesis, $\ell \geq k+1$, and

hence, $w(\zeta_\ell) \geq w(\zeta_{k+1})$ because of the indexing. Since we have assumed $w(\zeta_{k+1}) > w(C_{i_{k+1}})$, this gives us $w(\zeta_\ell) > w(C_{i_{k+1}})$. Hence, the set $\{C_{i_{k+1}}, \zeta_1, \ldots, \zeta_{\ell-1}, \zeta_{\ell+1}, \ldots, \zeta_p\}$ obtained by exchanging ζ_ℓ for $C_{i_{k+1}}$ gives a smaller homology basis than the minimum one, contradicting the minimality of $\mathcal{H}$.

From the first part of the proof, we have established that $w(C_{i_{k+1}}) = w(\zeta_{k+1})$. So, if $C_{i_{k+1}}$ is not homologous to $\zeta_{k+1} \in \mathcal{H}$ and $w(\zeta_{k+1}) = w(C_{i_{k+1}})$, then $\overline{\mathcal{H}} = \{C_{i_{k+1}}, \zeta_1, \ldots, \zeta_{\ell-1}, \zeta_{\ell+1}, \ldots, \zeta_p\}$ obtained by exchanging ζ_ℓ for $C_{i_{k+1}}$ is the desired minimum homology basis, proving the induction claim. ◀

Previously, it was known from Erickson and Whittlesey [11] that $\mathcal{H}$ is contained in $\mathcal{T}$.

▶ **Theorem 10** (Erickson and Whittlesey [11]). *With non-negative weights, every cycle in a shortest basis of $\mathsf{H}_1(K)$ is tight. That is, if $\mathcal{H}$ is any minimum homology basis of K, then $\mathcal{H} \subset \mathcal{T}$.*

Using Theorems 1 and 9, we can refine the above observation.

▶ **Corollary 11.** *Let $\mathcal{T}$ denote the set of tight cycles of K_1, and let $\mathcal{M}$ be a minimum cycle basis of K_1. Then, there exists a minimum homology basis $\mathcal{H}$ of K such that $\mathcal{H} \subset \mathcal{M} \subset \mathcal{T}$.*

5 Algorithms for minimum homology basis

To begin with, note that since $\mathsf{C}_p(K), \mathsf{Z}_p(K), \mathsf{B}_p(K)$ and $\mathsf{H}_p(K)$ are vector spaces, the problem of computing a minimum homology basis can be couched in terms of matrix operations.

Given a complex K, let $\{C_1, \ldots, C_{|\mathcal{T}|}\}$ be the list of tight cycles in K_1 sorted by weight, and let $\mathbf{M}_{\mathcal{T}}(K_1) = [C_1\, C_2 \ldots\, C_{|\mathcal{T}|}]$ be the matrix formed with cycles C_i as its columns. Then, the matrix $\hat{\mathbf{Z}} = [\partial_2 \mid \mathbf{M}_{\mathcal{T}}(K_1)]$ has $O(N + n\nu)$ columns and $O(N + n\nu)$ non-zero entries since $\mathbf{M}_{\mathcal{T}}(K_1)$ has $O(n\nu)$ columns and $O(n\nu)$ non-zero entries by Theorem 4, and ∂_2 has $O(N)$ columns and $O(N)$ non-zero entries. Since $\hat{\mathbf{Z}}$ has m rows, the rank of $\hat{\mathbf{Z}}$ is bounded by m. This immediately suggests an algorithm for computing minimum homology basis analogous to Algorithm 3.

Algorithm 4 Algorithm for minimum homology basis.

1: Compute the sorted list of tight cycles in $\mathbf{M}_{\mathcal{T}}(K_1)$, and assemble matrix $\hat{\mathbf{Z}}$.
2: Compute the column rank profile $[j_1, j_2, \ldots, j_b, i_1, i_2, \ldots, i_g]$ of $\hat{\mathbf{Z}}$ using Storjohann and Yang's algorithm [20], where columns $\{\hat{\mathbf{Z}}_{j_k}\}$ and $\{\hat{\mathbf{Z}}_{i_\ell}\}$ are linearly independent columns of ∂_2 and $\mathbf{M}_{\mathcal{T}}(K_1)$ respectively.
3: Return Columns $\{\hat{\mathbf{Z}}_{i_1}, \hat{\mathbf{Z}}_{i_2}, \ldots, \hat{\mathbf{Z}}_{i_g}\}$.

▶ **Theorem 12.** *Algorithm 4 is a Monte Carlo algorithm for computing a minimum homology basis that runs in $\tilde{O}(m^\omega)$ time with failure probability at most $\frac{1}{2}$.*

Proof. The correctness of the algorithm is an immediate consequence of Theorem 9 since, by definition, i_k is the smallest index for which $\hat{\mathbf{Z}}_{i_k}$ is not homologous to any cycle spanned by $\{\hat{\mathbf{Z}}_{i_1}, \ldots, \hat{\mathbf{Z}}_{i_{k-1}}\}$.

The list of tight cycles in G can be computed in $O(nm)$ time using the algorithm described in Section 2 of [1]. Hence, Step 1 of Algorithm 4 takes $O(nm \log n)$ time. Moreover, using Theorem 6, the complexity of Step 2 is bounded by $\tilde{O}(N + n\nu + m^\omega)$, which is the same as $\tilde{O}(m^\omega)$ since N and $n\nu$ are both in $\tilde{O}(m^\omega)$, and the failure probability is at most 1/2. ◀

When the number of 2-simplices in complex K is significantly smaller than the number of edges, the complexity for minimum homology can be slightly improved by decoupling the minimum homology basis computation from the minimum cycle basis computation, as illustrated in Algorithm 5.

Algorithm 5 Algorithm for minimum homology basis.

1: Compute a minimum cycle basis $\mathcal{M}$ of K_1 using the Monte Carlo algorithm by Amaldi et al. [1]. Let $\mathbf{B}_{\mathcal{M}}$ be the matrix whose columns are cycle vectors in $\mathcal{M}$ sorted by weight.
2: Assemble the matrix $\tilde{\mathbf{Z}} = [\partial_2 \mid \mathbf{B}_{\mathcal{M}}]$.
3: Compute the column rank profile $[j_1, j_2, \ldots, j_b, i_1, i_2, \ldots, i_g]$ of $\tilde{\mathbf{Z}}$ using the deterministic algorithm by Jeannerod et al. [14], where columns $\{\tilde{\mathbf{Z}}_{j_k}\}$ and $\{\tilde{\mathbf{Z}}_{i_\ell}\}$ are linearly independent columns of ∂_2 and $\mathbf{B}_{\mathcal{M}}$ respectively.
4: RETURN Columns $\{\tilde{\mathbf{Z}}_{i_1}, \tilde{\mathbf{Z}}_{i_2}, \ldots, \tilde{\mathbf{Z}}_{i_g}\}$.

▶ **Theorem 13.** *Minimum homology basis can be computed in $O(m^\omega + Nm^{\omega-1})$ time using the Monte Carlo algorithm described in Algorithm 5. The algorithm fails with probability at most $\nu \log(nm)\, 2^{-k}$, where $k = m^{0.1}$.*

Proof. As in Theorem 12, the correctness of the algorithm is an immediate consequence of Theorem 9. The algorithm fails only when Step 1 returns an incorrect answer, the probability of which is as low as $\nu \log(nm)\, 2^{-k}$, where $k = m^{0.1}$, see Theorem 3.2 of [1].

The minimum cycle basis algorithm by Amaldi et al. [1] runs in $O(m^\omega)$ time (assuming the current exponent of matrix multiplication $\omega > 2$). Furthermore, using Theorem 5, the complexity of Line 3 is bounded by $O(Nm^{\omega-1})$. So, the overall complexity of the algorithm is $O(m^\omega + Nm^{\omega-1})$. ◀

Note that in Line 3 of Algorithm 5, it is possible to replace the deterministic algorithm by Jeannerod et al. [14] with the Monte Carlo algorithm by Storjohann and Yang's algorithm [20]. In that case, the complexity of the algorithm will once again be $\tilde{O}(m^\omega)$, and the failure probability will be at most $1 - \frac{1}{2}(1 - \nu \log(nm) 2^{-k})$.

We would like to point out that the complexities of Algorithm 4 and Algorithm 5 are, *in general*, not comparable. For instance, for families of complexes with $N^{1-\epsilon} = \omega(m)$, for some $\epsilon > 0$, Algorithm 4 is faster than Algorithm 5. However, for families of complexes with $N = o(m)$, Algorithm 5 is faster than Algorithm 4. Moreover, for families of complexes with $g = \Theta(N)$, where, as before, g denotes the rank of $\mathsf{H}_1(K)$, Algorithms 4 and 5 are both faster than Dey et al.'s algorithm [8] (which runs in $O(N^\omega + N^2 g)$ time).

6 Discussion

In this paper, we show that questions about minimum cycle basis and minimum homology basis can be naturally recast into the problem of computing rank profiles of matrices, leading to fast algorithms with simple and elegant high-level descriptions. The column rank profile (or the earliest basis) of a matrix has previously been used to compute the minimum homology basis of a simplicial complex [3, 8]. Such a greedy approach that picks, at each step, an independent cycle of the smallest index, works because of the matroid structure of homology bases and cycle bases. What's novel about our approach is that we point out that, for both problems, independence can be efficiently checked owing to the sparsity of the matrices comprising of candidate cycles.

It is also worth noting that for the algorithms presented in this paper, the simplicity of high-level description doesn't translate to simple algorithms that can be easily implemented because the black-box subroutines employed by these algorithms are fairly complex.

Maintenance of support vectors has served as a key ingredient in designing algorithms for minimum cycle basis since de Pina. Our algorithm, however, does not explicitly maintain support vectors, and in that sense, is somewhat conceptually different from the recent algorithms for computing minimum cycle bases.

References

1 Edoardo Amaldi, Claudio Iuliano, Tomasz Jurkiewicz, Kurt Mehlhorn, and Romeo Rizzi. Breaking the $O(m^2n)$ barrier for minimum cycle bases. In Amos Fiat and Peter Sanders, editors, *Algorithms - ESA 2009*, pages 301–312, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

2 Glencora Borradaile, Erin Wolf Chambers, Kyle Fox, and Amir Nayyeri. Minimum cycle and homology bases of surface-embedded graphs. *JoCG*, 8(2):58–79, 2017. doi:10.20382/jocg.v8i2a4.

3 Oleksiy Busaryev, Sergio Cabello, Chao Chen, Tamal K. Dey, and Yusu Wang. Annotating simplices with a homology basis and its applications. In Fedor V. Fomin and Petteri Kaski, editors, *Algorithm Theory – SWAT 2012*, pages 189–200, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

4 Chao Chen and Daniel Freedman. Measuring and computing natural generators for homology groups. *Comput. Geom. Theory Appl.*, 43(2):169–181, February 2010. doi:10.1016/j.comgeo.2009.06.004.

5 Chao Chen and Daniel Freedman. Hardness results for homology localization. *Discrete & Computational Geometry*, 45(3):425–448, April 2011. doi:10.1007/s00454-010-9322-8.

6 Ho Yee Cheung, Tsz Chiu Kwok, and Lap Chi Lau. Fast matrix rank algorithms and applications. *J. ACM*, 60(5):31:1–31:25, October 2013. doi:10.1145/2528404.

7 José Coelho de Pina. *Applications of shortest path methods.* PhD thesis, Universiteit van Amsterdam, 1995.

8 Tamal K. Dey, Tianqi Li, and Yusu Wang. Efficient algorithms for computing a minimal homology basis. In Michael A. Bender, Martín Farach-Colton, and Miguel A. Mosteiro, editors, *LATIN 2018: Theoretical Informatics*, pages 376–398, Cham, 2018. Springer International Publishing.

9 Tamal K. Dey, Jian Sun, and Yusu Wang. Approximating loops in a shortest homology basis from point data. In *Proceedings of the Twenty-sixth Annual Symposium on Computational Geometry*, SoCG '10, pages 166–175, New York, NY, USA, 2010. ACM. doi:10.1145/1810959.1810989.

10 Jean-Guillaume Dumas, Clément Pernet, and Ziad Sultan. Simultaneous computation of the row and column rank profiles. In *Proceedings of the 38th International Symposium on Symbolic and Algebraic Computation*, ISSAC '13, pages 181–188, New York, NY, USA, 2013. ACM. doi:10.1145/2465506.2465517.

11 Jeff Erickson and Kim Whittlesey. Greedy optimal homotopy and homology generators. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '05, pages 1038–1046, Philadelphia, PA, USA, 2005. Society for Industrial and Applied Mathematics.

12 Allen Hatcher. *Algebraic topology.* Cambridge University Press, Cambridge, 2002.

13 J. D. Horton. A polynomial-time algorithm to find the shortest cycle basis of a graph. *SIAM J. Comput.*, 16(2):358–366, April 1987. doi:10.1137/0216026.

14 Claude-Pierre Jeannerod, Clément Pernet, and Arne Storjohann. Rank-profile revealing gaussian elimination and the cup matrix decomposition. *Journal of Symbolic Computation*, 56:46–68, 2013. doi:10.1016/j.jsc.2013.04.004.

15 Telikepalli Kavitha, Christian Liebchen, Kurt Mehlhorn, Dimitrios Michail, Romeo Rizzi, Torsten Ueckerdt, and Katharina A. Zweig. Cycle bases in graphs characterization, algorithms, complexity, and applications. *Computer Science Review*, 3(4):199–243, 2009. doi:10.1016/j.cosrev.2009.08.001.

16 Telikepalli Kavitha, Kurt Mehlhorn, and Dimitrios Michail. New approximation algorithms for minimum cycle bases of graphs. *Algorithmica*, 59(4):471–488, April 2011. doi:10.1007/s00453-009-9313-4.

17 Telikepalli Kavitha, Kurt Mehlhorn, Dimitrios Michail, and Katarzyna Paluch. A faster algorithm for minimum cycle basis of graphs. In Josep Díaz, Juhani Karhumäki, Arto Lepistö, and Donald Sannella, editors, *Automata, Languages and Programming*, pages 846–857, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.

18 Kurt Mehlhorn and Dimitrios Michail. Minimum cycle bases: Faster and simpler. *ACM Trans. Algorithms*, 6(1):8:1–8:13, December 2009. doi:10.1145/1644015.1644023.

19 Arne Storjohann and Shiyun Yang. Linear independence oracles and applications to rectangular and low rank linear systems. In *Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation*, ISSAC '14, pages 381–388, New York, NY, USA, 2014. ACM. doi:10.1145/2608628.2608673.

20 Arne Storjohann and Shiyun Yang. A relaxed algorithm for online matrix inversion. In *Proceedings of the 2015 ACM on International Symposium on Symbolic and Algebraic Computation*, ISSAC '15, pages 339–346, New York, NY, USA, 2015. ACM. doi:10.1145/2755996.2756672.

21 Shiyun Yang. Algorithms for fast linear system solving and rank profile computation. Master's thesis, University of Waterloo, 2014.

D

PAPER IV: PARAMETERIZED INAPPROXIMABILITY OF MORSE MATCHING.

Classical Morse theory provides a rich toolkit for analyzing and inferring the topology of a smooth manifolds by studying the critical points of smooth functions defined on it. Forman's discrete Morse theory is a combinatorial analogue of Morse theory that is applicable to regular CW complexes. It has become a popular tool in the visualization community, and is actively studied in algebraic combinatorics. Discrete Morse theory has also found applications as a preprocessing tool for speeding up computations in topological data analysis.

The principal construct in Forman's theory is the so-called discrete gradient vector field defined on a simplicial (or a CW) complex. A discrete gradient vector field is a collection of facet-cofacet pairs satisfying a certain acyclicity condition. The simplices that do not belong to any of the pairs in the gradient are deemed critical. Forman's theory also has an equivalent graph theoretic formulation in which the acyclic matchings (also called *Morse matchings*) in the Hasse diagram of a simplicial complex correspond to the discrete gradient vector fields on the simplicial complex. There has been a lot of practical interest in computing gradient vector fields on simplicial complexes with (near-)optimal number of critical simplices. The problem of finding a gradient vector field with the minimum number of critical simplices in a simplicial complex was shown to be **NP**-hard by Joswig and Pfetsch, who also posed its approximability as an open question. In Paper I, we resolved the open question by establishing hardness of approximation results for the maximization and minimization variants of the Morse matching problem. This paper is a follow-up paper to Paper I.

Despite what was already known at the time about the complexity of computing optimal Morse matchings, there were some glaring gaps. For instance, although Burton et al. proved that the related problem of computing the Erasing number of a 2-complex is **W[P]**-hard, it was not clear if computing optimal Morse matchings parameterized by solution size is also **W[P]**-hard. The approximability status of the minimization variant of Morse matchings restricted to 2-complexes was also open. The motivation for this paper was to close all the loose ends and to provide a complete picture for the complexity of computing optimal Morse matchings.

In Sections 1-3, we start with an the introduction and then follow it up with by topological and algorithmic preliminaries. In Section 4, we outline the reduction from Min-Monotone Circuit Sat to Min-Morse Matching (MinMM) on 2-complexes. As a consequence of this reduction, in Sections 5 and 6, we prove the following statements.

- MinMM has no approximation within a factor of $2^{\log^{(1-\epsilon)} n}$, for any $\epsilon > 0$, unless **NP** $\subseteq$ **QP** (throughout this paper, log denotes the logarithm by base 2),
- the standard parameterization of MinMM is **W[P]**-hard, and
- MinMM with standard parameterization has no FPT approximation algorithm for *any* approximation ratio function ρ, unless **FPT** $=$ **W**[**P**].

In Section 7, we design an $O(\frac{n}{\log n})$-factor algorithm for MinMM on 2-complexes to complement the inapproximability bound from Sections 6.

In Section 8, we observe that Kahle's techniques for designing discrete gradients on random clique complexes generalize to Costa–Farber complexes. Specifically, we show that for a wide range of parameter values, there exist discrete gradients for which the ratio of expected number of critical r-simplices to the expected number of r-simplices (for any fixed dimension r) tends to zero. When (re-)specialized to Linial–Meshulam complexes, we obtain a very appealing statement about the average case approximability of Morse matchings. (See Theorem 53). Finally, we conclude the paper in section 9 with a discussion about future directions.

STATEMENT OF INDIVIDUAL CONTRIBUTION. I am the principal author of this article. The idea for this work came up during a discussion about possible related projects after the publication of Article I. After doing some literature search, I concluded that for the kind of hardness results we were looking for, it would be best to reduce from MIN-MONOTONE CIRCUIT SAT. After I sketched the high-level ideas, Uli Bauer and I had a discussion, where he suggested some important simplifications to the ideas I presented. After a few more discussions, I proved the results and wrote the article. Once the article was written up, he suggested some expository changes.

Parameterized inapproximability of Morse matching

Ulrich Bauer
Department of Mathematics, Technical University of Munich (TUM)
Boltzmannstr. 3, 85748 Garching b. München, Germany

Abhishek Rathod
Department of Mathematics, Technical University of Munich (TUM)
Boltzmannstr. 3, 85748 Garching b. München, Germany

Abstract

We study the problem of minimizing the number of critical simplices from the point of view of inapproximability and parameterized complexity. We first show inapproximability of Min-Morse Matching within a factor of $2^{\log^{(1-\epsilon)} n}$. Our second result shows that Min-Morse Matching is **W**[**P**]-hard with respect to the standard parameter. Next, we show that Min-Morse Matching with standard parameterization has no FPT approximation algorithm for *any* approximation factor ρ. The above hardness results are applicable to complexes of dimension ≥ 2.

On the positive side, we provide a factor $O(\frac{n}{\log n})$ approximation algorithm for Min-Morse Matching on 2-complexes, noting that no such algorithm is known for higher dimensional complexes. Finally, we devise discrete gradients with very few critical simplices for typical instances drawn from a fairly wide range of parameter values of the Costa–Farber model of random complexes.

2012 ACM Subject Classification Mathematics of computing → Algebraic topology

Keywords and phrases Discrete Morse theory, Hardness of approximation, W[P]-hardness, L-reduction, FPT reduction, Gap preserving reduction, Parameterized complexity, FPT-approximability, Stochastic topology

Funding This research has been supported by the DFG Collaborative Research Center SFB/TRR 109 "Discretization in Geometry and Dynamics"

1 Introduction

Classical Morse theory [56] is an analytical tool for studying topology of smooth manifolds. Forman's discrete Morse theory is a combinatorial analogue of Morse theory that is applicable to simplicial complexes, and more generally regular cell complexes [29]. In Forman's theory, discrete Morse functions play the role of smooth Morse functions, whereas discrete gradient vector fields are the analogues of gradient-like vector fields. The principal objects of study are, therefore, the so-called discrete gradient vector fields (or discrete gradients) on simplicial complexes. Discrete gradients are partial face-coface matchings that satisfy certain acyclicity conditions. Forman's theory also has an elegant graph theoretic formulation [14], in which the acyclic matchings (or *Morse matchings*) in the Hasse diagram of a simplicial complex are in one-to-one correspondence with the discrete gradients on the simplicial complex. For this reason, we use the terms *gradient vector fields* and *Morse matchings* interchangeably.

Discrete Morse theory has become a popular tool in computational topology, image processing and visualization [6, 8, 13, 36, 43, 58, 62, 68], and is actively studied in algebraic and toplogical combinatorics [37, 44, 45, 55, 66]. Over the period of last decade, it has emerged a powerful computational tool for several problems in topological data analysis [21, 35, 42, 61]. Because of the wide array of applications there is a lot of practical interest in computing gradient vector fields on simplicial complexes with a (near-)optimal number of critical simplices [2, 7, 11, 12, 32, 33, 47]. The idea of using discrete Morse theory to speed up the computation of (co)homology [20,33,46], persistent homology [7,57], zigzag persistence [26,52],

and multiparameter persistent homology [65] relies on the fact that discrete Morse theory can be employed to reduce the problem of computing the homology of an input simplicial complex to that of a much smaller chain complex.

The effectiveness of certain heuristics for Morse matching raises an important question: to what extent is it feasible to obtain near-optimal solutions for Morse matching in polynomial time? To this end, inapproximability results for Min-Morse Matching for simplicial complexes of dimension $d \geq 3$ were established in [9]. To this date, however, we are unaware of any hardness results for Min-Morse Matching on 2-complexes from the perspective of inapproximability or parameterized complexity (although the related Erasability problem was shown to be **W[P]**-hard by Burton et al. [12]). With this paper, we seek to close the knowledge gap. By establishing various hardness results, we demonstrate the limitations of polynomial time methods for computing near-optimal Morse matchings. On the other hand, by devising an approximation algorithm, we make it evident that Min-Morse Matching on 2-complexes is not entirely inapproximable. We also observe that the typical Morse matching instances drawn from a wide range of parameter values of the Costa–Farber complexes are a lot easier in contrast to the discouraging worst case inapproximability bounds.

1.1 Related work

Joswig and Pfetsch [39] showed that finding an optimal gradient vector field is an **NP**-hard problem based on the relationship between erasability and Morse Matching observed by Lewiner [48, 49]. The *erasability problem* was first studied by Eğecioğlu and Gonzalez [24]. Joswig and Pfetsch also posed the question of approximability of optimal Morse matching as an open problem. On the positive side, Rathod et al. [60] devised the first approximation algorithms for Max-Morse Matching on simplicial complexes that provide constant factor approximation bounds for fixed dimension. Complementing these results, Bauer and Rathod [9] showed that for simplicial complexes of dimension $d \geq 3$ with n simplices, it is **NP**-hard to approximate Min-Morse Matching within a factor of $O(n^{1-\epsilon})$, for any $\epsilon > 0$. However, the question of approximability of Min-Morse Matching for 2-complexes is left unanswered in [9].

Next, Burton et al. [12] showed that the Erasability problem (that is, finding the number of 2-simplices that need to be removed to make a 2-complex erasable) is **W[P]**-complete. We note the **W[P]**-hardness of erasability can be inferred from our methods as well, and therefore our result can be seen as a strengthening of the hardness result from [12]. Moreover, our parameterized inapproximability results rely on the machinery developed by Eickmeyer et al. [25] and Marx [53].

Our reduction techniques have a flavor that is similar to the techniques used by Malgouryes and Francés [51] and Tancer [69] for proving **NP**-hardness of certain collapsibility problems. In particular, Tancer [69] also describes a procedure for 'filling 1-cycles with disks' to make the complex contractible (and even collapsible for satisfiable inputs). Our technique of filling 1-cycles is however entirely different (and arguably simpler) than Tancer's procedure. Our work is also related to [9, 10] in that we use the so-called modified dunce hats for constructing the gadget used in the reduction. Recently, modified dunce hats were used to provide a simpler proof of **NP**-completeness of the shellability decision problem [64].

1.2 The Morse Matching Problems

The Max-Morse Matching problem (MaxMM) can be described as follows: Given a simplicial complex $\mathcal{K}$, compute a gradient vector field that maximizes the cardinality of

matched (regular) simplices, over all possible gradient vectors fields on $\mathcal{K}$. Equivalently, the goal is to maximize the number of gradient pairs. For the complementary problem Min-Morse Matching (MinMM), the goal is to compute a gradient vector field that minimizes the number of unmatched (critical) simplices, over all possible gradient vector fields on K. While the problem of finding the exact optimum are equivalent for MinMM and MaxMM, the approximation variants behave quite differently.

Additionally, we define another variant of the minimization problem for 2-dimensional complexes, namely Min-Reduced Morse Matching (MinrMM). For this problem, we seek to minimize the total number of critical simplices minus one. This variant is natural, since any discrete gradient necessarily has at least one critical 0-simplex. It corresponds to a variant definition of simplicial complexes commonly used in combinatorics, which also consider the empty set as a simplex of dimension -1.

1.3 Our contributions

	Burton et al. [12]	This paper
W[**P**]-hardness of Erasability (SP)	✓	✓
W[**P**]-hardness of Min-Morse Matching (SP)	✗	✓
FPT-inapproximability of Min-Morse Matching (SP)	✗	✓
FPT-algorithm for Min-Morse Matching (TW)	✓	✗

Table 1 Comparison of results: Parameterized complexity of Morse matching (dim. ≥ 2). In the above table, SP denotes standard parameterization, whereas TW denotes treewidth parameterization.

	Results from [9]	This paper
IA for Max-Morse Matching (dim. ≥ 2)	$\left(1-\frac{1}{4914}\right)+\epsilon$	✗
IA for Min-Morse Matching (dim. ≥ 3)	$O(n^{1-\epsilon})$	$2^{\log^{(1-\epsilon)} n}$ (weaker)
IA for Min-Morse Matching (dim. ≥ 2)	✗	$2^{\log^{(1-\epsilon)} n}$
AA for Min-Morse Matching (dim. ≥ 2)	✗	$O(\frac{n}{\log n})$

Table 2 Comparison of results: Approximability of Morse matching. IA denotes inapproximability. AA denotes approximation algorithm.

In Section 5, we establish several hardness results for MinrMM, using a reduction from Min-Monotone Circuit Sat. In particular, we show the following:

- MinrMM has no approximation within a factor of $2^{\log^{(1-\epsilon)} n}$, for any $\epsilon > 0$, unless **NP** $\subseteq$ **QP** (throughout this paper, log denotes the logarithm by base 2),
- the standard parameterization of MinrMM is **W**[**P**]-hard, and
- MinrMM with standard parameterization has no FPT approximation algorithm for *any* approximation ratio function ρ, unless **FPT** = **W**[**P**].

In Section 6, we first show that the **W**[**P**]-hardness result and FPT-inapproximability results easily carry over from MinrMM to MinMM. To the best of our knowledge, this constitutes the first FPT-inapproximability result in computational topology. Using the amplified complex construction introduced in [9], we observe that the inapproximability result also carries over from MinrMM to MinMM. In particular, we show that even for 2-complexes MinMM cannot be approximated within a factor of $2^{\log^{(1-\epsilon)} n}$, for any $\epsilon > 0$, unless **NP** $\subseteq$ **QP**, where n denotes the number of simplices in the complex.

Sections 7 and 8 are concerned with some positive results. First, in Section 7, we design an $O(\frac{n}{\log n})$-factor algorithm for MinMM on 2-complexes. Then, in Section 8, we make the observation that Kahle's techniques [40] for designing discrete gradients on random clique complexes generalize to Costa–Farber random complexes. Specifically, we show that for a wide range of parameter values, there exist discrete gradients for which the ratio of expected number of critical r-simplices to the expected number of r-simplices (for any fixed dimension r) tends to zero. Although these methods do not lead to approximation algorithms, they fall under the general paradigm of beyond worst-case analysis [63].

Note that we do not distinguish between abstract and geometric simplicial complexes since every abstract simplicial complex can be embedded in a Euclidean space of appropriate dimension. As a final remark, we believe that with this paper we tie all the loose ends regarding complexity questions in discrete Morse theory.

2 Topological preliminaries

2.1 Simplicial complexes

A *k-simplex* $\sigma = \operatorname{conv} V$ is the convex hull of a set V of $(k+1)$ affinely independent points in $\mathbb{R}^d$. We call k the dimension of σ. We say that σ is *spanned* by the points V. Any nonempty subset of V also spans a simplex, a *face* of σ. A simplex σ is said to be a *coface* of a simplex τ if and only if τ is face of σ. We say that σ is a *facet* of τ if σ is a face of τ with $\dim \sigma = \dim \tau - 1$. A *simplicial complex* K is a collection of simplices that satisfies the following conditions:

- any face of a simplex in K also belongs to K, and
- the intersection of two simplices $\sigma_1, \sigma_2 \in K$ is either empty or a face of both σ_1 and σ_2.

For a complex K, we denote the set of d-simplices of K by $K^{(d)}$. The *n-skeleton* of a simplicial complex K is the simplicial complex $\bigcup_{m=0}^{n} K^{(m)}$. A simplex σ is called a *maximal face* of a simplicial complex K if it is not a strict subset of any other simplex $\tau \in K$. The *underlying space* of K is the union of its simplices, denoted by $|K|$. The underlying space is implicitly used whenever we refer to K as a topological space.

An *abstract simplicial complex* $\mathcal{S}$ is a collection of finite nonempty sets $A \in \mathcal{S}$ such that every nonempty subset of A is also contained in $\mathcal{S}$. The sets in $\mathcal{S}$ are called its *simplices*. A subcomplex of K is an abstract simplicial complex L such that every face of L belongs to K; denoted as $L \subset K$. For example, the vertex sets of the simplices in a geometric complex form an abstract simplicial complex, called its *vertex scheme*. Given an abstract simplicial complex K with n simplices, we can associate a *pointed simplicial complex* to it by choosing an arbitrary vertex and regarding it as the distinguished basepoint of K. The *m^{th} wedge sum of K* is then the quotient space of a disjoint union of m copies of K with the distinguished basepoints of each of the copies of K identified.

2.2 Discrete Morse theory and Erasability

We assume that the reader is familiar with simplicial complexes. Section 2.1 summarizes the key definitions. In this section, we provide a brief description of Forman's discrete Morse theory on simplicial complexes. For a comprehensive expository introduction, we refer the reader to [30].

A real-valued function f on a simplicial complex K is called a *discrete Morse function* if

- f is *monotonic*, i.e., $\sigma \subseteq \tau$ implies $f(\sigma) \leq f(\tau)$, and

- for all $t \in \mathrm{im}(f)$, the preimage $f^{-1}(t)$ is either a singleton $\{\sigma\}$ (in which case σ is a *critical simplex*) or a pair $\{\sigma, \tau\}$, where σ is a facet of τ (in which case (σ, τ) form a *gradient pair* and σ and τ are *regular simplices*).

Given a discrete Morse function f defined on complex K, the *discrete gradient vector field* $\mathcal{V}$ of f is the collection of pairs of simplices (σ, τ), where (σ, τ) is in $\mathcal{V}$ if and only if σ is a facet of τ and $f(\sigma) = f(\tau)$.

Discrete gradient vector fields have a useful interpretation in terms of acyclic graphs obtained from matchings on Hasse diagrams, due to Chari [14]. Let K be a simplicial complex, let H_K be its Hasse diagram, and let M be a matching in the underlying undirected graph H_K. Let $H_K(M)$ be the directed graph obtained from H_K by reversing the direction of each edge of the matching M. Then M is a *Morse matching* if and only if $H_K(M)$ is a directed acyclic graph. Every Morse matching M on the Hasse diagram H_K corresponds to a unique gradient vector field $\mathcal{V}_M$ on complex K and vice versa. For a Morse matching M, the unmatched vertices correspond to critical simplices of $\mathcal{V}_M$, and the matched vertices correspond to the regular simplices of $\mathcal{V}_M$.

A non-maximal face $\sigma \in K$ is said to be a *free face* if it is contained in a unique maximal simplex $\tau \in K$. If $d = \dim \tau = \dim \sigma + 1$, we say that $K' = K \setminus \{\sigma, \tau\}$ arises from K by an *elementary collapse*, or an *elementary d-collapse* denoted by $K \searrow^e K'$. Furthermore, we say that K *collapses* to L, denoted by $K \searrow L$, if there exists a sequence $K = K_1, K_2, \dots K_n = L$ such that $K_i \searrow^e K_{i+1}$ for all i. If K collapses to a point, one says that K is collapsible.

A simplicial collapse can be encoded by a discrete gradient.

▶ **Theorem 1** (Forman [29], Theorem 3.3)**.** *Let K be a simplicial complex with a vector field $\mathcal{V}$, and let $L \subseteq K$ be a subcomplex. If $K \setminus L$ is a union of pairs in $\mathcal{V}$, then $K \searrow L$.*

In this case, we say that the collapse $K \searrow L$ is *induced by* the gradient $\mathcal{V}$. As a consequence of this theorem, we obtain:

▶ **Theorem 2** (Forman [29], Corollary 3.5)**.** *Let K be a simplicial complex with a discrete gradient vector field $\mathcal{V}$ and let m_d denote the number of critical simplices of $\mathcal{V}$ of dimension d. Then K is homotopy equivalent to a CW complex with exactly m_d cells of dimension d.*

In particular, a discrete gradient vector field on K with m_d critical simplices of dimension d gives rise to a chain complex having dimension m_d in each degree d, whose homology is isomorphic to that of K. This condensed representation motivates the algorithmic search for (near-)optimal Morse matchings.

Following the terminology used in [9, 24], we make the following definitions: A maximal face τ in a simplicial complex K is called an *internal simplex* if it has no free face. If a 2-complex K collapses to a 1-complex, we say that K is *erasable*. Moreover, for a 2-complex K, the quantity $\mathrm{er}(K)$ is the minimum number of internal 2-simplices that need to be removed so that the resulting complex collapses to a 1-complex. Equivalently, it is the minimum number of critical 2-simplices of any discrete gradient on K. Furthermore, we say that a subcomplex $L \subseteq K$ is an *erasable subcomplex of K (through the gradient $\mathcal{V}$)* if there exists another subcomplex $\mathcal{M} \subseteq K$ with $K \searrow \mathcal{M}$ (induced by the gradient $\mathcal{V}$) such that the set of 2-dimensional simplices of these complexes satisfy the following relation: $L^{(2)} \subseteq K^{(2)} \setminus \mathcal{M}^{(2)}$. We call such a gradient $\mathcal{V}$ an *erasing gradient*. Finally, we say that a simplex σ in a complex K is *eventually free (through the gradient $\mathcal{V}$)* if there exists a subcomplex L of K such that $K \searrow L$ (induced by $\mathcal{V}$) and σ is free in L. Equivalently, K collapses further to a subcomplex not containing σ.

We recall the following results from [9].

▶ **Lemma 3** ([9], Lemma 2.1). *Let K be a connected simplicial complex, let p be a vertex of K, and let $\mathcal{V}_1$ be a discrete gradient on K with $m_0 > 1$ critical simplices of dimension 0 and m critical simplices in total. Then there exists a polynomial time algorithm to compute another gradient vector field $\widetilde{\mathcal{V}}$ on K with p as the only critical simplex of dimension 0 and $m - 2(m_0 - 1)$ critical simplices in total.*

▶ **Lemma 4** ([9], Lemma 2.3). *If K is an erasable complex, then any subcomplex $L \subset K$ is also erasable.*

▶ **Lemma 5.** *Suppose that we are given a complex K and a set M of simplices in K with the property that simplices in M have no cofaces in K. Then $L = K \setminus M$ is a subcomplex of K with the property that the gradient vector fields on K with all simplices in M critical are in one-to-one correspondence with gradient vector fields on L.*

Proof. Given a gradient vector field on L, we extend it to a gradient vector field on K by making all simplices in M critical. Given a vector field $\mathcal{V}$ on K, the restriction $\mathcal{V}|_L$ is a gradient vector field on L. ◀

▶ Notation 1. For the remainder of the paper, we use $[m]$ to denote the set $\{1, 2, \ldots, m\}$ for any $m \in \mathbb{N}$, and $[i, j]$ to denote the set $\{i, i+1, \ldots, j\}$ for any $i, j \in \mathbb{N}$.

3 Algorithmic preliminaries

3.1 Approximation algorithms

An α-*approximation algorithm* for an optimization problem is a polynomial-time algorithm that, for all instances of the problem, produces a solution whose objective value is within a factor α of the objective value of an optimal solution. The factor α is called the *approximation ratio* (or *approximation factor*) of the algorithm.

An *approximation preserving reduction* is a polynomial time procedure for transforming an optimization problem A to an optimization problem B, such that an α-approximation algorithm for B implies an $f(\alpha)$-approximation algorithm for A, for some function f. Then, if A is hard to approximate within factor $f(\alpha)$, the reduction implies that B is hard to approximate within factor α. A particularly well-studied class of approximation preserving reductions is given by the *L-reductions*, which provide an effective tool in proving hardness of approximability results [59, 71].

Now, consider a minimization problem A with a non-negative integer valued objective function m_A. Given an instance x of A, the goal is to find a solution y minimizing the objective function $m_A(x, y)$. Define $\mathrm{OPT_A}(x)$ as the minimum value of the objective function on input x. An *L-reduction* (with parameters μ and ν) from a minimization problem A to another minimization problem B is a pair of polynomial time computable functions f and g, and fixed constants $\mu, \nu > 0$, satisfying the following conditions:

1. The function f maps instances of A to instances of B.
2. For any instance x of A, we have
$$\mathrm{OPT_B}(f(x)) \leq \mu \, \mathrm{OPT_A}(x).$$
3. The function g maps an instance x of A and a solution of the corresponding instance $f(x)$ of B to a solution of x.
4. For any instance x of A, and any solution y of $f(x)$, we have
$$m_A(x, g(x, y)) - \mathrm{OPT_A}(x) \leq \nu \left(m_B(f(x), y) - \mathrm{OPT_B}(f(x)) \right).$$

If $\mu = \nu = 1$, the reduction is *strict*.

We will use the following straightforward fact about L-reductions, which appears as Theorem 16.6 in a book by Williamson and Shmoys [71].

▶ **Theorem 6.** *If there is an L-reduction with parameters μ and ν from a minimization problem A to another minimization problem B, and there is a $(1+\delta)$-approximation algorithm for B, then there is a $(1+\mu\nu\delta)$-approximation algorithm for A.*

3.2 Parameterized complexity

Parameterized complexity, as introduced by Downey and Fellows in [22], is a refinement of classical complexity theory. The theory revolves around the general idea of developing complexity bounds for instances of a problem not just based on their size, but also involving an additional *parameter*, which might be significantly smaller than the size. Specifically, we have the following definition.

▶ **Definition 7** (Parameter, parameterized problem [28])**.** *Let Σ be a finite alphabet.*

1. *A* parameter *of Σ^*, the set of strings over Σ, is a function $\rho : \Sigma^* \to \mathbb{N}$, attaching to every input $w \in \Sigma^*$ a natural number $\rho(w)$.*
2. *A* parameterized problem *over Σ is a pair (P, ρ) consisting of a set $P \subseteq \Sigma^*$ and a (polynomial time computable) parametrization $\rho : \Sigma^* \to \mathbb{N}$.*
3. *A parameterized problem (P, ρ) is said to be* fixed-parameter tractable *or* FPT in the parameter ρ *if the question*

$$(x, p) \in \{(y, \rho(y)) \mid y \in P\}$$

can be decided in running time $O(g(p))\cdot|x|^{O(1)}$, where $g\colon \mathbb{N} \to \mathbb{N}$ is an arbitrary computable function depending only on the parameter p.

FPT reductions provide a principal tool to establish hardness results in the parameterized complexity landscape.

▶ **Definition 8** (FPT reduction [28])**.** *Given two parameterized problems (P, k) and (Q, k'), we say that there is an* FPT reduction *from (P, k) and (Q, k'), if there exists a functions φ that transforms parameterized instances of P to parameterized instances of Q while satisfying the following properties:*

1. *φ is computable by an FPT algorithm,*
2. *$\varphi(x)$ is a yes-instance of (Q, k') if and only if x is a yes-instance of (P, k).*
3. *There exists a computable function $g\colon \mathbb{N} \to \mathbb{N}$ such that $k'(\varphi(x)) \leq g(k(x))$.*

The natural way of turning a minimization problem into a decision problem is to add a value k to the input instance, and seek a solution with cost at most k. Taking this value k appearing in the input as the parameter is called *the standard parameterization* of the minimization problem (sometimes also referred to as *the natural parameterization*). In general, the parameter can be any function of the input instance, for example, the treewidth of the input graph, or the maximum degree of the input graph.

Parameterized approximability is an extension of the notion of classical approximability. Informally, an FPT approximation algorithm is an algorithm whose running time is fixed parameter tractable for the parameter *cost of the solution* and whose approximation factor ρ is a function of the parameter (and independent of the input size). For instance, every

polynomial time approximation algorithm with constant approximation factor is automatically an FPT approximation algorithm, but an approximation algorithm with approximation factor $\Theta(\sqrt{n})$, where n denotes the input size, is not an FPT approximation algorithm. Next, following [53], for standard parameterization of minimization problems, we provide definitions for FPT approximation algorithms and FPT cost approximation algorithms. Analogous definitions for maximization problems are also considered in [53].

▶ **Definition 9** (FPT approximation algorithm [53]). *Let P be an* **NP** *minimization problem, and let $\rho : \mathbb{N} \to \mathbb{R}_{\geq 1}$ be a computable function such that $k \mapsto k \cdot \rho(k)$ is nondecreasing. An* FPT approximation algorithm *for P (over some alphabet Σ) with approximation ratio ρ is an algorithm $\mathbb{A}$ with the following properties:*

1. *For every input (x, k) whose optimal solution has cost at most k, $\mathbb{A}$ computes a solution for x of cost at most $k \cdot \rho(k)$. For inputs (x, k) without a solution of cost at most k, the output can be arbitrary.*
2. *The runtime of $\mathbb{A}$ on input (x, k) is $O(g(k) \cdot |x|^{O(1)})$ for some computable function g.*

It is often convenient to work with a weaker notion of approximability where an algorithm is only required to compute the cost of an optimal solution rather than an actual optimal solution, and to work with decision rather than optimization problems. With that in mind, the notion of *FPT cost approximability* was introduced in [15].

▶ **Definition 10** (FPT cost approximation algorithm [53]). *Let P be an* **NP** *minimization problem (over the alphabet Σ), and $\rho : \mathbb{N} \to \mathbb{R}_{\geq 1}$ a computable function. For an instance x of P, let $\min(x)$ denote its optimal value. Then, a decision algorithm $\mathbb{A}$ is an* FPT cost approximation algorithm *for P with approximation ratio ρ if*

1. *For feasible instances x of P and parameterized instances (x, k), $\mathbb{A}$ satisfies:*
   a. *If $k \geq \min(x) \cdot \rho(\min(x))$, then $\mathbb{A}$ accepts (x, k).*
   b. *If $k < \min(x)$, then $\mathbb{A}$ rejects (x, k).*
2. *$\mathbb{A}$ is an FPT algorithm. That is, there exists a computable function f with the property that for an input (x, k), the running time of $\mathbb{A}$ is bounded by $f(k) \cdot |x|^{O(1)}$.*

It can be readily checked that FPT-approximability implies FPT cost approximability with the same approximation factor. Please refer to Section 3.1 of [15] for more details.

▶ **Theorem 11** (Chen et al. [15]). *Let P be an* **NP** *minimization problem over the alphabet Σ, and let $\rho : \mathbb{N} \to \mathbb{R}_{\geq 1}$ be a computable function such that $k \cdot \rho(k)$ is nondecreasing and unbounded. Suppose that P is FPT approximable with approximation ratio ρ. Then P is FPT cost approximable with approximation ratio ρ.*

An immediate consequence of the theorem above is that if P is not FPT cost approximable with approximation ratio ρ (under certain complexity theory assumptions), then P is not FPT approximable with approximation ratio ρ (under the same assumptions).

Gap problems and gap-preserving reductions were originally introduced in the context of proving the PCP theorem [5] – a cornerstone in the theory of approximation algorithms. These notions have natural analogues in the parameterized approximability setting. Below, we follow the definitions as provided by Eickmeyer et al. [25].

▶ **Definition 12** (gap instance of a parameterized problem [25]). *Let $\delta : \mathbb{N} \to \mathbb{R}_{\geq 1}$ be a function, P a minimization problem, and P' its standard parameterization. An instance (x, k) is a δ-gap instance of P' if either $\min(x) \leq k$ or $\min(x) \geq k \cdot \delta(k)$.*

▶ **Definition 13** (gap-preserving FPT reduction [25]). *Let $\alpha, \beta \colon \mathbb{N} \to \mathbb{R}_{\geq 1}$ be two computable functions, and let P and Q be two minimization problems. Let P' and Q' be the natural parameterizations of P and Q, respectively. We say that a reduction R from P' to Q' is a* (α, β)-gap-preserving FPT reduction *if*

1. *R is an FPT reduction from P' to Q',*
2. *for every α-gap instance (x, k) of P', the instance $R(x, k)$ is a β-gap instance of Q'.*

We use gap-preserving FPT reductions to establish FPT-inapproximability.

3.3 Circuits

First, we recall some elementary notions from Boolean circuits. In particular, by an *and-node*, we mean the digital logic gate that implements logical conjuction ($\wedge$), by an *or-node*, we mean the digital logic gate that implements logical disjunction ($\vee$), and by a *not-node*, we mean the digital logic gate that implements negation ($\neg$).

▶ **Definition 14** (Boolean circuit). *A* Boolean circuit *C is a directed acyclic graph, where each node is labeled in the following way:*

1. *every node with in-degree greater than* 1 *is either an* and-node *or an* or-node*,*
2. *each node of in-degree* 1 *is labeled as a* negation node*,*
3. *and each node of in-degree* 0 *is an* input node*.*

Moreover, exactly one of the nodes with out-degree 0 is labeled as the output node*.*

Below, we recall some essential parameterized complexity results concerning circuits.

We use the terms *gates* and *nodes* interchangeably. We say that a gate has *fan-in k* if its in-degree is at most k. We say that a gate is an *ordinary gate* if it is neither an input gate nor an output gate. We denote the nodes and edges in C by $V(C)$ and $E(C)$ respectively. The size of a circuit C, denoted by $|C|$, is the total number of nodes and edges in C. That is, $|C| = |V(C)| + |E(C)|$. The *Hamming weight* of an assignment is the number of input gates receiving value 1. An assignment on the input nodes induces an assignment on all nodes. So given an assignment from the input nodes of circuit C to $\{0, 1\}$, we say that the assignment satisfies C if the value of the output node is 1 for that assignment. Let $\mathcal{G}_I$ denote the set of input gates of C. Then, an assignment A can be viewed as a binary vector of size $|\mathcal{G}_I|$. In the Weighted Circuit Satisfiability (WCS) problem, we are given a circuit C and an integer k, and the task is to decide if C has a satisfying assignment of Hamming weight at most k. Accordingly, in the Min-Weighted Circuit Satisfiability (MinWCS) problem, we are given a circuit C, and the task is to find a satisfying assignment with minimum Hamming weight.

▶ **Definition 15** ($\mathbf{W}[\mathbf{P}]$). *A parameterized problem W belongs to the class $\mathbf{W}[\mathbf{P}]$ if it can be reduced to the standard parameterization of* WCS*.*

A Boolean circuit is monotone if it does not contain any negation nodes. Let $\mathcal{C}^+$ be the class of all monotone Boolean circuits. Then, Min-Monotone Circuit Sat (MinMCS) is the restriction of the problem MinWCS to input circuits belonging to $\mathcal{C}^+$.

The following result seems to be folklore and appears in the standard literature [23, 27].

▶ **Theorem 16** (Theorem 3.14 [28]). *The standard parameterization of* MinMCS *is $\mathbf{W}[\mathbf{P}]$-complete.*

Furthermore, Eickmeyer et al. [25] showed that unless $\mathbf{W}[\mathbf{P}] = \mathbf{FPT}$, MinMCS does not have an FPT approximation algorithm with polylogarithmic approximation factor ρ. The FPT-inapproximability result was subsequently improved by Marx [53] as follows.

▶ **Theorem 17** (Marx [53]). MinMCS *is not FPT cost approximable, unless* $\mathbf{FPT} = \mathbf{W}[\mathbf{P}]$.

Combined with Theorem 11, the above theorem implies that MinMCS is not FPT-approximable for any function ρ, unless $\mathbf{FPT} = \mathbf{W}[\mathbf{P}]$.

▶ Remark 18 (Fan-in 2 circuits). We note that it is possible to transform a monotone circuit C to another monotone circuit C' such that both circuits are satisfied on the same inputs, and every gate of C' has fan-in 2. This is achieved as follows: Each or-gate of in-degree k in C is replaced by a *tree* of or-gates with in-degree-2 in C', and each and-gate of in-degree k in C is replaced by a *tree* of and-gates with in-degree-2 in C'. In each case, we transform a single gate having fan-in k to a sub-circuit of $\Theta(k)$ gates having depth $\Theta(\log k)$ and fan-in 2. In fact, it is easy to check that $|C'|$ is a polynomial function of $|C|$, and C' can be computed from C in time polynomial in C. Since the number of input gates for C and C' is the same, for the rest of the paper we will assume without loss of generality that an input circuit instance has fan-in 2.

4 Reducing MinMCS to MinrMM

In this section, we describe how to construct a 2-complex $K(C)$ that corresponds to a monotone circuit $C(V, E)$. By Remark 18, we assume without loss of generality that C has fan-in 2. For the rest of the paper, we denote the number of gates in C by n. Also, throughout, we use the notation $j \in [a, b]$ to mean that j takes integer values in the interval $[a, b]$.

Following the notation from Section 3.1, given a monotone circuit $C = (\mathcal{V}, E)$ and the associated complex $K(C)$, let $\mathrm{OPT}_{\mathsf{MinMCS}}(C)$ denote the optimal value of the MinMCS problem on C, and let $\mathrm{OPT}_{\mathsf{MinrMM}}(K(C))$ denote the optimal value of the MinrMM problem on $K(C)$. The value of the objective function $m_{\mathsf{MinrMM}}(K(C), \mathcal{V})$ is the number of critical simplices in $\mathcal{V}$ minus one; the value of the objective function $m_{\mathsf{MinMCS}}(C, \mathcal{I}(C, \mathcal{V}))$ is the Hamming weight of the input assignment. In Section 4.2, we describe the map K that transforms instances of MinMCS (monotone circuits C) to instances of MinrMM (simplicial complexes $K(C)$), and the map $\mathcal{I}$ that transforms solutions of MinrMM (discrete gradients $\mathcal{V}$ on $K(C)$) to solutions of MinMCS (satisfying input assignments $\mathcal{I}(C, \mathcal{V})$ of circuit C).

4.1 The building block for the gadget

We shall first describe a complex that serves as the principal building block for the gadget in our reduction. The building block is based on a modification of Zeeman's *dunce hat* [73]. The dunce hat is a simplicial complex that is contractible (i.e. has the homotopy type of a point) but has no free faces and is therefore not collapsible. In contrast, we work with *modified dunce hats* [31] that are collapsible through either one or two free edges. The modified dunce hat has been previously used to show hardness of approximation of Max-Morse Matching [9] and $\mathbf{W}[\mathbf{P}]$-hardness of Erasability Expansion Height [10], and is discussed extensively in these papers.

Figure 1 depicts two triangulations of modified dunce hats, which we denote by $\mathbf{D}_{m,\ell}$. We use the subscript m, ℓ to designate the numbers of distinguished edges of $\mathbf{D}_{m,\ell}$, which come in two types: the free edges of $\mathbf{D}_{m,\ell}$ denoted by s_i, $1 \leq i \leq m$, and the edges of type

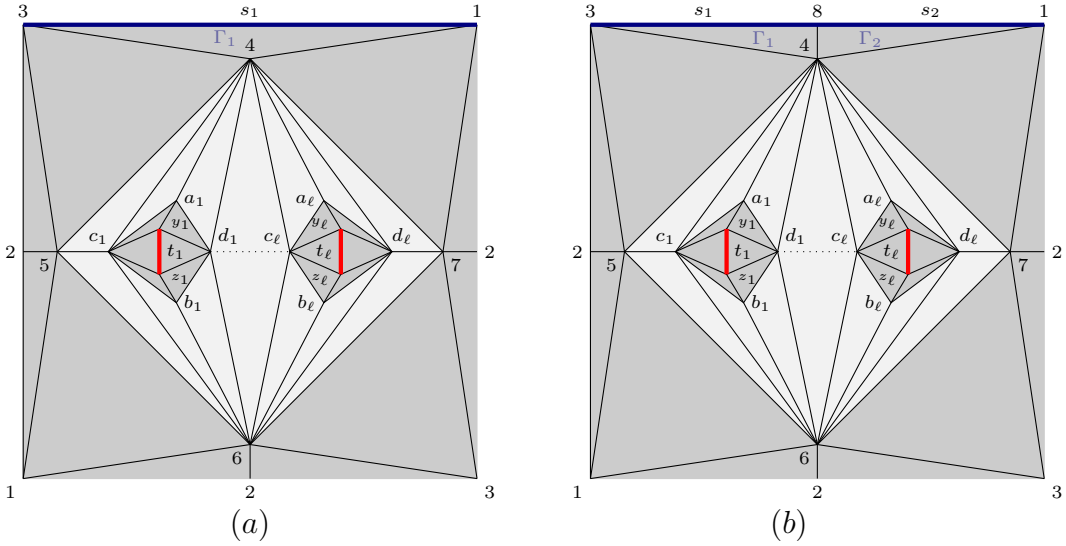


Figure 1 The figure (a) on the left depicts $\mathbf{D}_{1,\ell}$ that is collapsible through one free face, namely $s_1 = \{3, 1\}$. The figure (b) on the right depicts $\mathbf{D}_{2,\ell}$ that is collapsible through two free faces, namely $s_1 = \{3, 8\}$ and $s_2 = \{8, 1\}$. The edges $\{1, 2\}$ and $\{2, 3\}$ on the right and at the bottom of both subfigures are shown in light grey to indicate that they are identified to $\{1, 2\}$ and $\{2, 3\}$ on the left.

$t_j = \{y_j, z_j\}$, $1 \leq j \leq \ell$. These distinguished edges are precisely the edges that are identified to edges from other building blocks. In Figure 1, we depict $\mathbf{D}_{1,\ell}$, and $\mathbf{D}_{2,\ell}$, with distinguished edges s_i, and t_j highlighted.

Note that, in this paper, we only consider modified dunce hats with either one or two free edges. That is, for the purpose of this paper, $m \in \{1, 2\}$. Also, abusing terminology, we often refer to 'modified dunce hats' as simply 'dunce hats'.

▶ Remark 19. It is easy to check that after executing a series of elementary 2-collapses, $\mathbf{D}_{m,\ell}$ collapses to a complex induced by edges

$$\{\{1, 2\}, \{2, 3\}, \{2, 6\}, \{6, 5\}, \{6, 7\}, \{6, b_i\}, \{6, c_i\}, \{6, d_i\}, \{b_i, z_i\}, \{z_i, y_i\}, \{y_i, a_i\}\} \cup F$$

for $i \in [1, \ell]$, and,

- $F = \{\{v, 4\}\}$ if $m = 1$, where $v = 6$, if ℓ is even, and $v = {}^{(\ell-1)}/_2 + 1$ if ℓ is odd,
- $F = \{s_1, \{4, 8\}\}$ if $m = 2$ and the collapse starts with a gradient pair involving s_2,
- $F = \{s_2, \{4, 8\}\}$ if $m = 2$ and the collapse starts with a gradient pair involving s_1.

The edges that are left behind after executing all the 2-collapses are highlighted using examples in Figure 2. (a) depicts the case where ℓ is odd and $m = 1$. (b) depicts the case where ℓ is even, $m = 2$ and the collapse starts with a gradient pair involving s_1.

4.2 Construction of the complex $K(C)$ and the map $\mathcal{I}(C, \mathcal{V})$

Given a circuit C, we first explain the construction of an intermediate complex $K'(C)$. We use the notation $\mathbf{D}^{(i,j)}_{m,\ell}$ to refer to the j-th copy of the the dunce hat associated to gate i, having m s-edges and ℓ t-edges. Sometimes, we suppress the subscript, and use the notation $\mathbf{D}^{(i,j)}$ to talk about $\mathbf{D}^{(i,j)}_{m,\ell}$. As illustrated in Figure 3, to each input gate G_i we associate a dunce hat $\mathbf{D}^{(i,1)}$. To the output gate G_o, we associate n copies of dunce hats $\{\mathbf{D}^{(o,j)}\}_{j=1}^{n}$. Moreover, Figure 4 depicts how we associate to each ordinary gate G_i n blocks

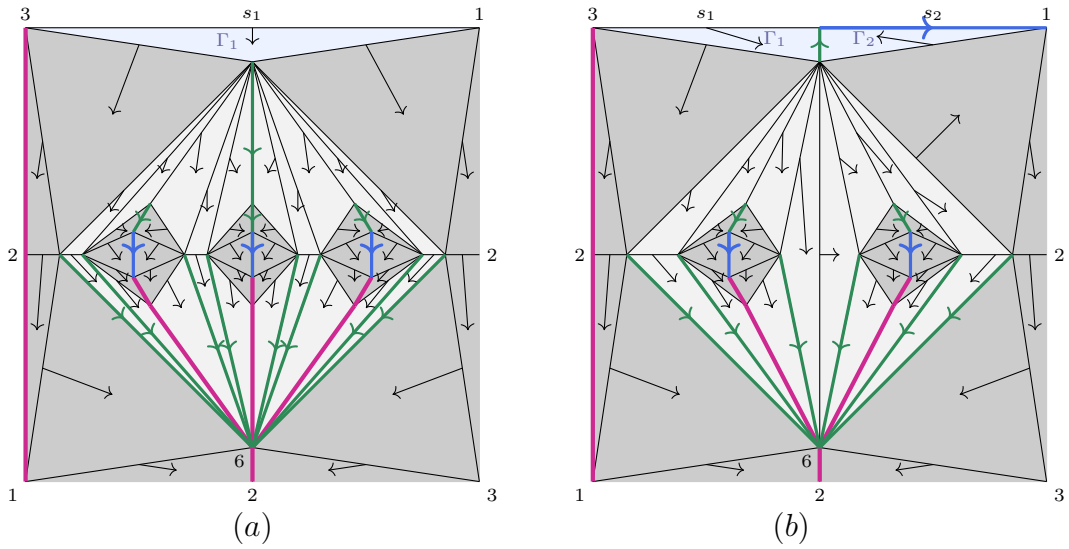


Figure 2 The figure (a) on the left depicts $\mathbf{D}_{1,3}$ that is collapsible through a unique free face, namely $s_1 = \{3, 1\}$. The complex $\mathbf{D}_{1,3}$ collapses to the subcomplex induced by the highlighted edges. When s_1 cannot be made free (because of edge identifications from other dunce hats), then s_1 and Γ_1 are made critical. In any case, the remaining 2-collapses are executed as shown in figure (a). The figure (b) on the right depicts $\mathbf{D}_{2,2}$ that is collapsible through two free faces, namely $s_1 = \{3, 8\}$ and $s_2 = \{8, 1\}$. There exists a collapsing sequence for $\mathbf{D}_{2,2}$ starting from the gradient pair (s_1, Γ_1) such that $\mathbf{D}_{2,2}$ collapses to the subcomplex induced by the highlighted edges. A symmetric statement can be made for a collapse starting from (s_2, Γ_2).

$\{{}^1\mathbf{D}^{(i,j)}, {}^2\mathbf{D}^{(i,j)}, {}^3\mathbf{D}^{(i,j)}\}_{j=1}^{n}$. The superscript to the left indexes dunce hats internal to the block. We call ${}^3\mathbf{D}^{(i,j)}$ the *output component* of block j associated to G_i. Likewise, we call ${}^1\mathbf{D}^{(i,j)}$ and ${}^2\mathbf{D}^{(i,j)}$ the *input components* of block j associated to G_i. If G_p serves as one of the two inputs to G_q, we say that G_p is a *predecessor* of G_q, and G_q is the *successor* of G_p.

A simplex labeled σ in $\mathbf{D}_{m,\ell}$ is correspondingly labeled as $\sigma^{(p,j)}$ in $\mathbf{D}_{m,\ell}^{(p,j)}$, (respectively as ${}^k\sigma^{(p,j)}$ in ${}^k\mathbf{D}_{m,\ell}^{(p,j)}$). We call the unique s-edge of the dunce hat associated to an input gate G_i, namely $s_1^{(i,1)}$, its *feedback edge*. As depicted in Figure 4, for an ordinary gate G_p, for each $j \in [1, n]$, the s_2 edges of ${}^1\mathbf{D}^{(p,j)}$ and ${}^2\mathbf{D}^{(p,j)}$, namely ${}^1s_2^{(p,j)}$ and ${}^2s_2^{(p,j)}$ respectively, are called the *feedback edges* of the j-th block associated to G_p. For an ordinary gate G_p, for each $j \in [1, n]$, the s_1 edges of ${}^1\mathbf{D}^{(p,j)}$ and ${}^2\mathbf{D}^{(p,j)}$, namely ${}^1s_1^{(p,j)}$ and ${}^2s_1^{(p,j)}$ respectively, are called the *input edges* of the j-th block associated to G_p. For the output gate G_o, the s_1 and s_2 edges of $\mathbf{D}^{(o,j)}$, namely $s_1^{(o,j)}$ and $s_2^{(o,j)}$ respectively, are called the *input edges* of the j-th copy associated to G_o.

To bring the notation of the edges closer to their function in the gadget, for the rest of the paper, we use the following alternative notation for s-edges. We denote the feedback edges $s_1^{(i,1)}$, ${}^1s_2^{(p,j)}$ and ${}^2s_2^{(p,j)}$ described above as $s_f^{(i,1)}$, $s_{f_1}^{(p,j)}$ and $s_{f_2}^{(p,j)}$ respectively. Also, we denote the input edges ${}^1s_1^{(p,j)}$, ${}^2s_1^{(p,j)}$, $s_1^{(o,j)}$ and $s_2^{(o,j)}$ described above by $s_{\iota_1}^{(p,j)}$, $s_{\iota_2}^{(p,j)}$, $s_{\iota_1}^{(o,j)}$ and $s_{\iota_2}^{(o,j)}$ respectively. Please see Figure 5 for an example.

We start with a disjoint union of dunce hats (or blocks) associated to each gate. Then, for an ordinary gate G_p that is a predecessor of G_q, for all $j, k \in [1, n]$ two distinct t-edges from the j-th copy (output component of the block) associated to G_q are identified to the two feedback edges of the k-th block associated to G_p. Also, for all $j, k \in [1, n]$ a t-edge from

the output component of the k-th block associated to G_p is identified to an input edge of the j-th copy (block) associated to G_q.

For an input gate G_p that is a predecessor of G_q, for all $j \in [1, n]$, a t-edge from the j-th copy (output component of the block) associated to G_q is identified to the feedback edge of the unique dunce hat associated to G_p. Also, for all $j \in [1, n]$, a t-edge from the dunce hat associated to G_p is attached to the input edge of the j-th copy (block) associated to G_q.

Moreover, these identifications are done to ensure that: a feedback edge of a copy (block) associated to G_p is free only if all the copies (output components of all the blocks) associated to *all* the successors of G_p have been erased, and an input edge of a copy (block) associated to G_q is free only if the unique copy (all the output components of all the block) associated to the predecessors of G_p have been erased. Please refer to Figure 5 for an example illustrating the identifications.

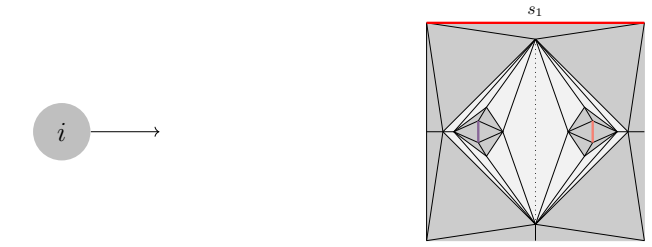
It is important to note that the gluing is done so that the s-edges from two different copies (blocks) associated to the same gate are *never* identified as an outcome of gluing, nor do they intersect in a vertex. In particular if G_p is a gate with G_{p_1} and G_{p_2} as inputs, where, for instance, if G_{p_1} is an input gate and G_{p_2} is an ordinary gate, then for every $k \in [1, n]$, $s_{\iota_1}^{(p,k)}$ is identified to a unique t-edge from the dunce hat associated to G_{p_1}, and $s_{\iota_2}^{(p,k)}$ is identified to n t-edges each from a block associated to G_{p_2}. These are the only identifications for edges $s_{\iota_1}^{(p,k)}$ and $s_{\iota_2}^{(p,k)}$. For every non-output gate G_p, let θ_p denote the number of successors of G_p. Then, for all $k \in [1, n]$, $s_{f_1}^{(p,k)}$ and $s_{f_2}^{(p,k)}$ each have $\theta_p n$ identifications from t-edges coming from each of the blocks associated to each of the successors of G_p. These are the only identifications for $s_{f_1}^{(p,k)}$ and $s_{f_2}^{(p,k)}$. If G_p is an input gate, then $s_f^{(p,1)}$ is identifies to $\theta_p n$ t-edges from blocks associated successors of G_p. Finally, the input s-edges of the k-th copy associated to the output gate G_o is identified to either one or n t-edges coming from dunce hats associated to predecessor gates, depending on whether the predecessor is an input gate or an ordinary gate. We refrain from providing indices for the identified t-edges as this would needlessly complicate the exposition.

For every non-input gate G_i, set $\phi_i = 1$ if the first input to G_i is from an input gate, and set $\phi_i = 2n$ otherwise. Similarly, for every non-input gate G_i, set $\psi_i = 1$, if the second input to G_i is from an input gate, and set $\psi_i = 2n$ otherwise.

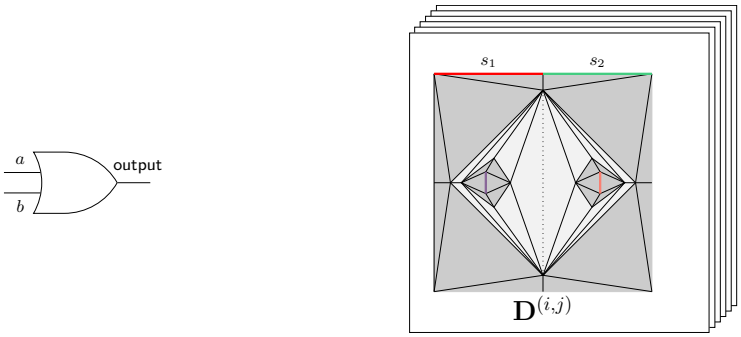
Now we can readily check the following: In our construction, for a dunce hat ${}^3\mathbf{D}_{m,\ell}^{(i,j)}$ associated to an ordinary gate G_i, we have $m = 1$ or $m = 2$ (depending on whether it is an and-gate or an or-gate), and $\ell = \theta_i n + \phi_i + \psi_i$. For dunce hats ${}^2\mathbf{D}_{m,\ell}^{(i,j)}$ and ${}^1\mathbf{D}_{m,\ell}^{(i,j)}$ associated to an ordinary gate G_i, we have $m = 2$, and $\ell = 1$. For a dunce hat $\mathbf{D}_{m,\ell}^{(i,j)}$ associated to an output gate G_i, we have $m = 1$ or $m = 2$, and $\ell = \phi_i + \psi_i$. Finally, for the dunce hat $\mathbf{D}_{m,\ell}^{(i,1)}$ associated to an input gate G_i, we have $m = 1$, and $\ell = \theta_i n$.

▶ Remark 20. We reindex the dunce hats described above using the indexing set Ξ. That is, for every dunce hat in $K'(C)$ there exists a unique $\zeta \in [1, |\Xi|]$ such that $\mathbf{D}_{m,\ell}^{\zeta}$ identifies the dunce hat of interest. Sometimes in our exposition it is more convenient to refer to dunce hats with a single index as opposed to using two or three indices in the superscript.

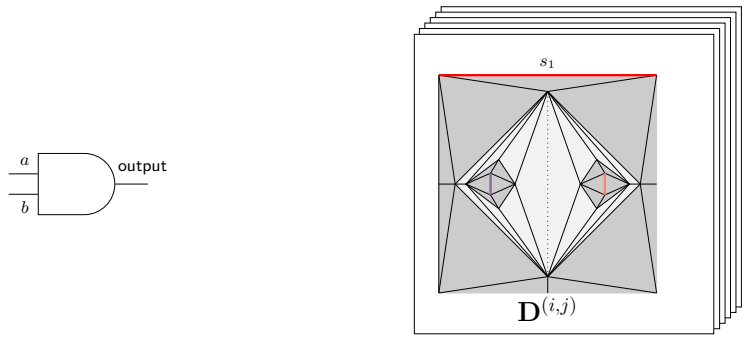
Let ζ be the indexing variable, and Ξ the indexing set as described in Remark 20. For a dunce hat $\mathbf{D}_{m,\ell}^{\zeta}$, we call the complex induced by the edges $\{\{1^\zeta, 2^\zeta\}, \{2^\zeta, 3^\zeta\}, \{2^\zeta, 6^\zeta\}, \{6^\zeta, b_k^\zeta\}, \{b_k^\zeta, z_k^\zeta\}\}$ (i.e., the pink edges of $\mathbf{D}_{m,\ell}$ in Figure 2) the *stem* of $\mathbf{D}_{m,\ell}^{\zeta}$. Then, in complex $K'(C)$, let H be the 1-dimensional subcomplex formed by the union of stems of $\mathbf{D}_{m,\ell}^{\zeta}$, for all $\zeta \in [1, |\Xi|]$. We call H the *stem of the complex* $K'(C)$. It can be shown that a basis for the first homology group of the complex $K'(C)$ is supported by the edges in the stem of the complex. The complex $K(C)$ is formed as follows: We first assemble the minimal



Input gate G_i | Associated dunce hat $\mathbf{D}^{(i,1)}$

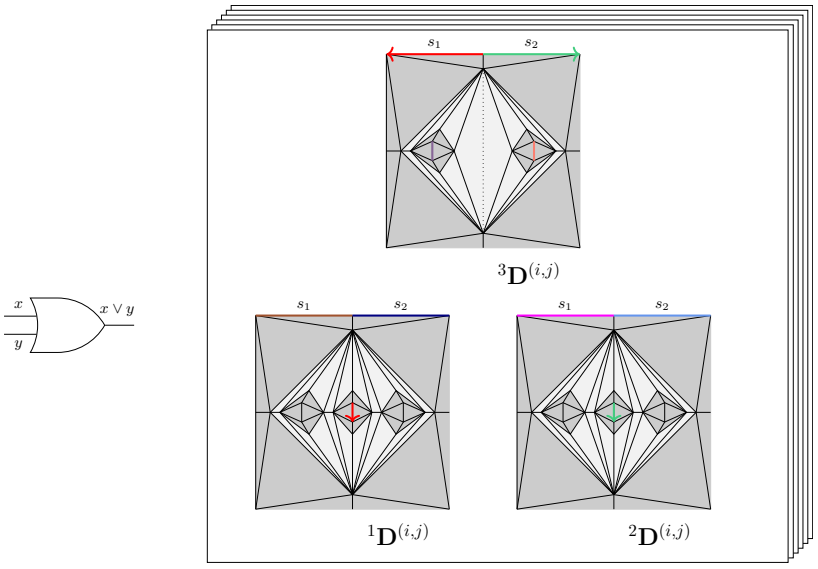


Output OR gate G_i | n associated copies $\{\mathbf{D}^{(i,j)}\}_{j=1}^{n}$

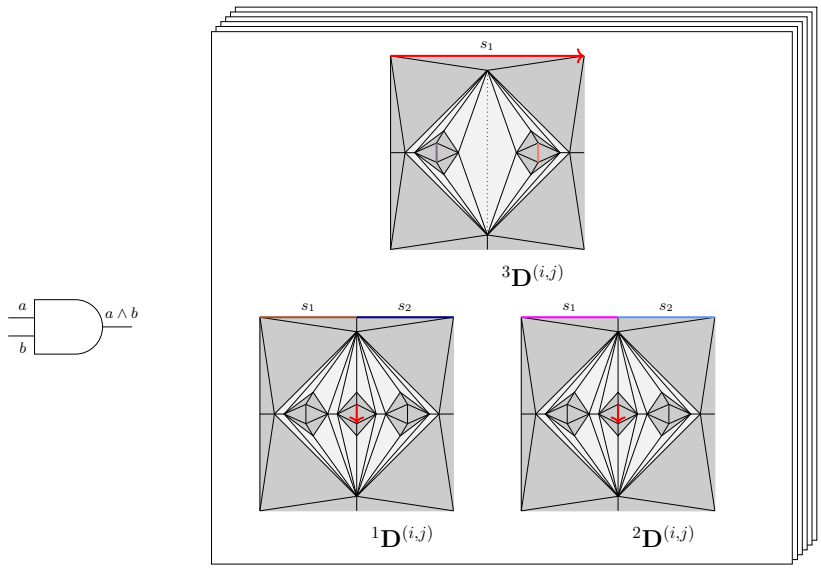


Output AND gate G_i | n associated copies $\{\mathbf{D}^{(i,j)}\}_{j=1}^{n}$

Figure 3 In all three figures, the distinguished edges are highlighted. The top figure shows an input gate and the dunce hat associated to it. We conceive the input gate as activated when the associated dunce hat has a critical 2-simplex in it. If the dunce hat doesn't have critical 2-simplices, then s_1 must be paired to its coface for the dunce hat to be erased. The edge s_1 supports a feedback mechanism. In particular, if all the dunce hats associated to the output gate are erased without activating G_i, then we need an *alternative* means to erase G_i, which is provided by s_1. The figure in the middle (resp. bottom) shows an output or-gate (resp. an output and-gate) and the associated n copies of dunce hats. In both cases, the j-th copy consists of a single dunce hat $\mathbf{D}^{(i,j)}$, where $j \in [1, n]$. The idea behind the dunce hat associated to the or-gate is that if either s_1 or s_2 is free, then $\mathbf{D}^{(i,j)}$ can be erased. The idea behind the dunce hat associated to the and-gate is that if s_1 is free, then $\mathbf{D}^{(i,j)}$ can be erased. Finally, we have n copies instead of a single copy per gate to ensure that optimum values of MinMCS and MinrMM are the same.



Non-output OR gate G_i n associated blocks $\{{}^1\mathbf{D}^{(i,j)}, {}^2\mathbf{D}^{(i,j)}, {}^3\mathbf{D}^{(i,j)}\}_{j=1}^n$



Non-output AND gate G_i n associated blocks $\{{}^1\mathbf{D}^{(i,j)}, {}^2\mathbf{D}^{(i,j)}, {}^3\mathbf{D}^{(i,j)}\}_{j=1}^n$

Figure 4 The figure on the top (resp. bottom) shows a non-output or-gate (resp. a non-output and-gate) and the associated n blocks of dunce hats. In both cases j-th block consists of 3 dunce hats $\{{}^1\mathbf{D}^{(i,j)}, {}^2\mathbf{D}^{(i,j)}, {}^3\mathbf{D}^{(i,j)}\}$, where $j \in [1, n]$. All distinguished edges are highlighted, and identical color coding indicates identifications. That is, red edges are glued to red edges and green to green. The arrows on the highlighted edges show the orientations of identifications. The idea behind the blocks associated to the or-gate is that if either the s_1 edge of ${}^1\mathbf{D}^{(i,j)}$ or the s_1 edge of ${}^2\mathbf{D}^{(i,j)}$ is free, then all three dunce hats in the j-th block can be erased. The idea behind the blocks associated to the and-gate is that if the s_1 edge of ${}^1\mathbf{D}^{(i,j)}$ and the s_1 edge of ${}^2\mathbf{D}^{(i,j)}$ are free, then all three dunce hats in the j-th block can be erased. For each block, the dark and the light blue s_2 edges of ${}^1\mathbf{D}^{(i,j)}$ and ${}^2\mathbf{D}^{(i,j)}$ respectively support a feedback mechanism. In particular, if the dunce hats associated to the output gate are erased then, we need an *alternative* means to erase all the dunce hats, since the satisfaction of the output gate is all we really care about. Finally, we have n blocks instead of a single block per gate to ensure that optimum values of MinMCS and MinrMM are the same.

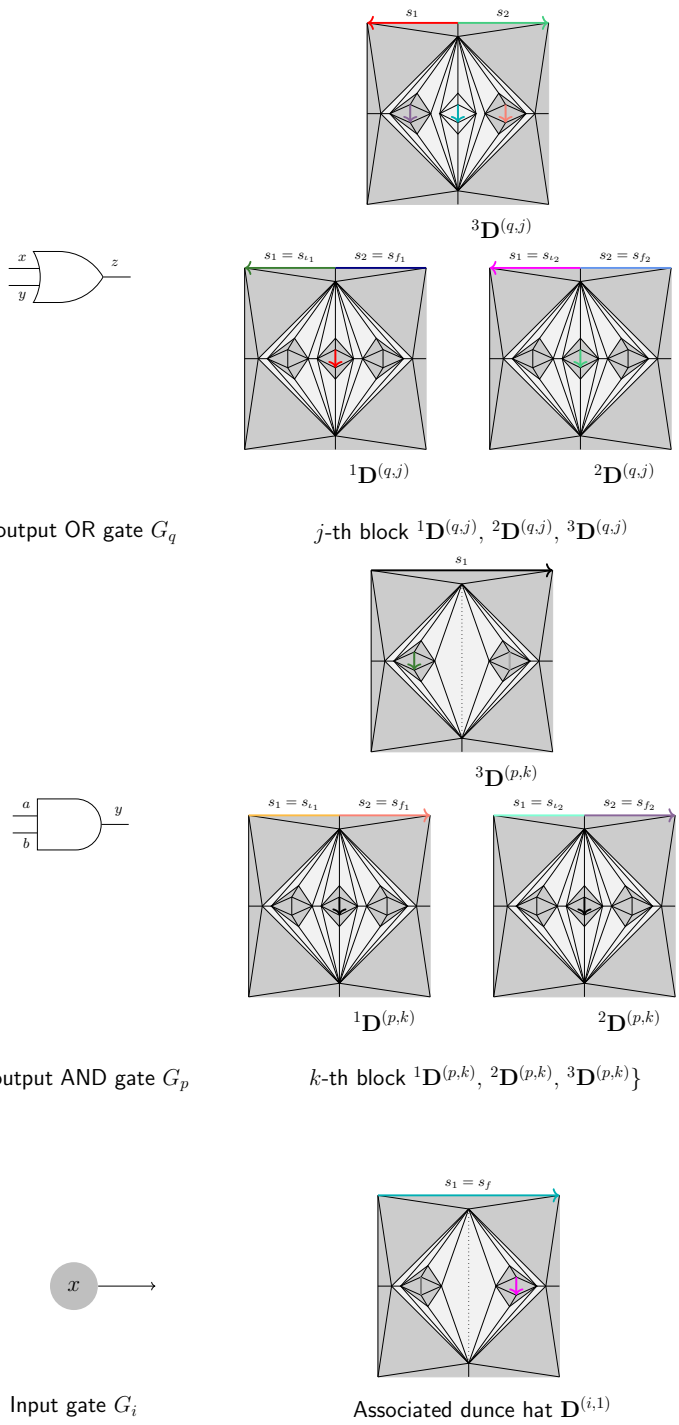


Figure 5 In this figure we depict the part of the complex associated to the (partial) circuit that implements $z = (a \wedge b) \vee x$, where x is an input to the circuit. Identical color coding indicates identifications, and the arrows indicate orientations of idenitifications. Here we only show identifications for j-th block of G_q and k-th block of G_p for arbitrary $j, k \in [1, n]$. Similar identifications occur across all respective associated blocks.

cycle basis of the stem of the complex $K'(C)$ in a matrix $\mathbf{M}$ and then make this matrix upper-triangular. Then, each cycle (column in $\mathbf{M}$) is filled with a triangulated disk, giving us the desired complex $K(C)$. Please refer to Section 4.3 for further details.

▶ Remark 21 (Design choices for ordinary gates). At this point we would like to remark that, in principle, one could construct a complex for circuits with arbitrary fan-ins wherein the or-gate and and-gate like behaviour can easily be implemented with a single (suitably sub-divided) dunce hat having two or more free edges. The problem with this approach is that it is much harder to control where the 1-cycles in the complex appear, and this makes the cycle filling procedure far more technical. This motivates our approach to first pass to circuits with fan-in two and then implement or-gates and and-gates with blocks of three instead of single dunce hats. As we shall see later, this leads to a straightforward instance-independent description of the 1-homology basis of $K'(C)$, which in turn simplifies cycle filling.

Given a gradient vector field $\tilde{\mathcal{V}}$ on K, we construct the map $\mathcal{I}(C, \tilde{\mathcal{V}})$ as follows: For every input gate G_i whose associated dunce hat has a critical 2-simplex in $\tilde{\mathcal{V}}$, we set $\mathcal{I}(C, \tilde{\mathcal{V}})(G_i) = 1$. Please refer to Appendix A.2 for further details.

4.3 Construction of the complex $K(C)$

From the identifications described in Section 4.2, it is easy to check that H is, in fact, a connected graph. The procedure for constructing $K(C)$ is described in Algorithm 1.

Algorithm 1 Procedure for constructing $K(C)$ from $K'(C)$

$K(C) \leftarrow K'(C)$
▷ Initially, $K(C)$ consists only of simplices from $K'(C)$.
Compute a minimal cycle basis $\mathcal{B}$ of H with $\mathbb{Z}_2$ coefficients
Assemble the basis vectors in a matrix $\mathbf{M}$
Make $\mathbf{M}$ upper-triangular using column operations
▷ The columns of $\mathbf{M}$ now represent a new basis $\mathcal{B}'$ of the cycle space of H.
▷ Let $\{e_i^j\}$ denote the edges of a cycle $z_i \in \mathcal{B}'$. Let n_i denote the number of edges in z_i.
for $i \leftarrow 1, |\mathcal{B}'|$ **do**
    Add a new vertex v_i to $K(C)$
    for $j \leftarrow 1, n_i$ **do**
        Add to $K(C)$ a 2-simplex $\sigma_i^j = e_i^j * v_i$ for each edge e_i^j of z_i
        Add to $K(C)$ all of the faces of simplices σ_i^j
    end for
end for
$\mathcal{D} \leftarrow K(C) \setminus K'(C)$
RETURN $K(C), \mathcal{D}$

In the construction described in Algorithm 1, the star of the vertex v_i may be viewed as a 'disk' that fills the cycle z_i. See Figure 6 for an illustration.

▶ Remark 22. It can be shown that

- the second homology groups $H_2(K'(C))$ and $H_2(K(C))$ are trivial,
- the classes $\{[z_i]\}$ are nontrivial and form a basis for $H_1(K'(C))$, where $\mathcal{B}' = \{z_i\}$ as in Algorithm 1, and
- $K(C)$ is contractible.

However, our hardness results can be established without proving any of the statements above. Having said that, it is important to bear in mind that the procedure of going from $K'(C)$ to $K(C)$ is, in fact, a 1-cycle filling procedure. As will be clear in the following section, the reason for using the basis $\mathcal{B}'$ instead of $\mathcal{B}$ is to make edges associated to successive pivot entries in matrix $\mathbf{M}$ free after collapsing all 2-simplices that fill the preceding cycle.

To establish hardness results, we introduce some additional notation. Given a monotone circuit $C = (\mathcal{V}, E)$ let $K(C)$ be its associated complex. Now let $\mathrm{OPT}_{\mathsf{MinMCS}}(C)$ denote the optimal value of the MinMCS problem on C, and let $\mathrm{OPT}_{\mathsf{MinrMM}}(K(C))$ denote the optimal value of the MinrMM problem on $K(C)$. The value of the objective function $m_{\mathsf{MinrMM}}(K(C), \mathcal{V})$ is the number of critical simplices in $\mathcal{V}$ minus one; the value of the objective function $m_{\mathsf{MinMCS}}(C, \mathcal{I}(C, \mathcal{V}))$ is the Hamming weight of the input assignment.

4.4 Reducing MinMCS to MinrMM: Forward direction

Given a circuit C, suppose that we are given an input assignment A that satisfies the circuit $C = (V(C), E(C))$. Let $\mathcal{S}$ be the set of gates that are satisfied by the assignment, and let $I(\mathcal{S})$ be the set of input gates that are assigned 1. Clearly, $I(\mathcal{S}) \subset \mathcal{S}$, and also the output gate $G_o \in \mathcal{S}$. Let $\overline{\mathcal{S}} = V(C) \setminus \mathcal{S}$ denote the set of gates that are not satisfied by the input A. Clearly, the subgraph $C_{\mathcal{S}}$ of C induced by the gates in $\mathcal{S}$ is a connected graph. Also, since C is a directed acyclic graph, the induced subgraph $C_{\mathcal{S}}$ is also directed acyclic. Let $\prec_{\mathcal{S}}$ be some total order on $\mathcal{S}$ consistent with the partial order imposed by $C_{\mathcal{S}}$, and let $\prec_C$ be some total order on $V(C)$ consistent with the partial order imposed by C.

Next, given an assignment A on C, we describe how to obtain a gradient vector field $\mathcal{V}$ on $K(C)$. We denote the complex obtained after i-th step by $K^i(C)$.

Step 1: Erase satisfied input gates

First, for every input gate $G_i \in \mathcal{S}$, we make $\Gamma_1^{(i,1)}$ critical. By Lemma 5, this is akin to removing $\Gamma_1^{(i,1)}$ from $\mathbf{D}^{(i,1)}$. Next, we make all $s_f^{(i,1)}$ for all $G_i \in \mathcal{S}$ critical. We then use Lemma 56 from Appendix A.1 to erase all the dunce hats $\mathbf{D}^{(i,1)}$ associated to satisfied input gates G_i, giving $K^1(C)$.

Step 2: Forward collapsing

Assume throughout Step 2 that the gates in $\mathcal{S}$ are indexed from 1 to $|\mathcal{S}|$ so that

$$\text{for all } G_i, G_j \in \mathcal{S}, \quad i < j \Leftrightarrow G_i \prec_{\mathcal{S}} G_j.$$

▶ **Lemma 23.** *Let $G_p \in \mathcal{S} \setminus I(\mathcal{S})$. Suppose that all the gates in $I(\mathcal{S})$ have been erased, and for all gates $G_k \in \mathcal{S} \setminus I(\mathcal{S})$ with $k < p$ the associated dunce hats ${}^3\mathbf{D}^{(k,r)}$ for all $r \in [1, n]$ have been erased. Then, the dunce hats ${}^3\mathbf{D}^{(p,j)}$, for all $j \in [1, n]$ associated to G_p can be erased.*

Proof. Let G_{p_1} and G_{p_2} be inputs to G_p. Assume without loss of generality that G_{p_1}, G_{p_2} are non-input gates. By our assumption on indexing, $p_1 < p$ and $p_2 < p$. By construction, the only identifications to $s_{\iota_1}^{(p,j)} \in {}^1\mathbf{D}^{(p,j)}$ are from t-edges that belong to ${}^3\mathbf{D}^{(p_1,r)}$ for all $r \in [1, n]$, and the only identifications to $s_{\iota_2}^{(p,j)} \in {}^2\mathbf{D}^{(p,j)}$ are from t-edges that belong to ${}^3\mathbf{D}^{(p_2,r)}$ for all $r \in [1, n]$. We have two cases:

▶ Case 1. Assume that G_p is a satisfied or-gate. Then, either $G_{p_1} \in \mathcal{S}$ or $G_{p_2} \in \mathcal{S}$. Without loss of generality, we assume that $G_{p_1} \in \mathcal{S}$. Then, for all j, $s_{\iota_1}^{(p,j)}$ become free since, by assumption, the dunce hats ${}^3\mathbf{D}^{(p_1,r)}$ associated to G_{p_1} have been erased. So using Lemma 55

from Appendix A.1, for all j, ${}^1\mathbf{D}^{(p,j)}$ can be erased. For each j, the unique identification to ${}^3s_1^{(p,j)}$ is from a t-edge in ${}^1\mathbf{D}^{(p,j)}$. Hence, for all j, ${}^3s_1^{(p,j)}$ becomes free, making it possible to erase ${}^3\mathbf{D}^{(p,j)}$ for all j.

▶ Case 2. Now, assume that G_p is a satisfied and-gate. Then, both $G_{p_1} \in \mathcal{S}$ and $G_{p_2} \in \mathcal{S}$. Thus, for all j, $s_{\iota_1}^{(p,j)}$ and $s_{\iota_2}^{(p,j)}$ become free since, by assumption, all the dunce hats ${}^3\mathbf{D}^{(p_1,p)}$ for all $p \in [1,n]$ associated to G_{p_1} and all dunce hats ${}^3\mathbf{D}^{(p_2,q)}$, for all $q \in [1,n]$ associated to G_{p_2} have been erased. So, using Lemma 55 from Appendix A.1, for all $j \in [1,n]$, ${}^1\mathbf{D}^{(p,j)}$ and ${}^2\mathbf{D}^{(p,j)}$ can be erased. For all $j \in [1,n]$, the only two edges identified to ${}^3s_1^{(p,j)}$ belong to ${}^1\mathbf{D}^{(p,j)}$ and ${}^2\mathbf{D}^{(p,j)}$ respectively. Hence, for all $j \in [1,n]$, ${}^3s_1^{(p,j)}$ becomes free, making it possible to erase ${}^3\mathbf{D}^{(p,j)}$ for all $j \in [1,n]$. Thus, the dunce hats ${}^3\mathbf{D}^{(p,j)}$ for $j \in [1,n]$ associated to G_p can be erased.

The argument is identical for the case when G_{p_1} or G_{p_2} is an input gate. ◀

▶ **Lemma 24.** *All dunce hats associated to the output gate are erased.*

Proof. Note that a satisfying assignment A that satisfies the circuit, in particular, also satisfies the output gate. A simple inductive argument using Lemma 23 proves the lemma. ◀

After applying Step 1, we apply Step 2, which comprises of executing the collapses described by Lemmas 23 and 24. This immediately gives us the following claim.

▷ Claim 25. If there exists an assignment satisfying a circuit C with Hamming weight m, then there exists a gradient vector field on $K(C)$ such that after making m 2-cells critical, all the dunce hats ${}^3\mathbf{D}^{(p,j)}$ associated to the satisfied non-output gates G_p and all dunce hats associated to the output gate can be erased.

The complex obtained after erasing executing Step 2 is denoted by $K^2(C)$. We have, $K^1(C) \searrow K^2(C)$.

▶ Remark 26. Note that the forward collapses do not erase all the dunce hats associated to satisfied gates. For instance, for a satisfied or-gate G_p, if one of the input gates, G_{p_1}, is satisfied and the other, G_{p_2}, is not, then ${}^1\mathbf{D}^{(p,j)}$ and ${}^3\mathbf{D}^{(p,j)}$ will be erased, but ${}^2\mathbf{D}^{(p,j)}$ will not be erased. The dunce hats associated to the unsatisfied gates and the unerased dunce hats associated to the satisfied gates are erased in the next step while executing the backward collapses.

Step 3: Backward collapsing

Assume throughout Step 3 that the gates in $V(C)$ are indexed from 1 to n so that

$$\text{for all } G_i, G_j \in V(C), \quad i < j \Leftrightarrow G_i \prec_C G_j.$$

The idea behind backward collapsing is that the feedback edges become successively free when one starts the collapse from dunce hats associated to highest indexed gate and proceeds in descending order of index.

▶ **Lemma 27.** *If all the dunce hats associated to gates G_k, where $k > i$, have been erased, then the dunce hats associated to G_i can be erased.*

Proof. We have three cases to verify:

▶ Case 1. First, assume that G_i is an ordinary gate. The only identifications to edges $s_{f_1}^{(i,j)} \in {}^1\mathbf{D}^{(i,j)}$ and $s_{f_2}^{(i,j)} \in {}^1\mathbf{D}^{(i,j)}$ respectively are from the t-edges in dunce hats associated to successors of G_i. By assumption, all dunce hats ${}^3\mathbf{D}^{(k,p)}$ associated to ordinary gates G_k where $k > i$ have been erased, and all dunce hats $\mathbf{D}^{(o,q)}$ associated to the output gate G_o have been erased. Hence, $s_{f_1}^{(i,j)}$ and $s_{f_2}^{(i,j)}$ are free, for every j. Therefore, for every j, dunce hats ${}^1\mathbf{D}^{(i,j)}$ and ${}^2\mathbf{D}^{(i,j)}$ can be erased.

▶ Case 2. If G_i is an unsatisfied gate, then for all j, the only identifications to s-edge(s) of ${}^3\mathbf{D}^{(i,j)}$ are from t-edges of ${}^1\mathbf{D}^{(i,j)}$ and ${}^2\mathbf{D}^{(i,j)}$. So the s-edge(s) of ${}^3\mathbf{D}^{(i,j)}$ become free for all j, allowing us to erase ${}^3\mathbf{D}^{(i,j)}$, for all j. Thus, all dunce hats associated to G_i can be erased.

▶ Case 3. Now assume that G_i is an input gate. Then, the unique s-edge of the unique copy associated to G_i is identified to t-edges of dunce hats associated to successors of G_i. Since, by assumption, all dunce hats associated to gates G_k, where $k > i$, have been erased, $s_1^{(i,1)}$ becomes free, allowing us to erase $\mathbf{D}^{(i,1)}$.

Note that in the proof of this lemma, for ordinary satisfied gates only Case 1 may be relevant, whereas for ordinary unsatisfied gates both Case 1 and Case 2 apply. ◀

▷ Claim 28. If there exists an assignment satisfying a circuit C with Hamming weight m, then there exists a gradient vector field on $K(C)$ with exactly m critical 2-cells.

Proof. We prove the claim by induction. The base step of the induction is provided by Claim 25. Then, we repeatedly apply the steps below until all gates in $K(C)$ are erased:

1. Choose the highest indexed gate whose associated dunce hats haven't been erased.
2. Apply the collapses described in Lemma 27 to erase dunce hats associated to G_k.

◀

The complex obtained after erasing all dunce hats in $K(C)$ is denoted by $K^3(C)$. We have, $K^1(C) \searrow K^2(C) \searrow K^3(C)$.

Step 4: Deleting critical 1-simplices

Note that in complex $K^3(C)$, the s-edges $s_f^{(i,1)}$ from $\mathbf{D}^{(i,1)}$, for all $G_i \in \mathcal{S}$ have no cofaces. Since they were already made critical in Step 1, by Lemma 5, we can delete $s_1^{(i,1)}$ from $K^3(C)$ for all $G_i \in \mathcal{S}$, and continue designing the gradient vector field on the subcomplex $K^4(C)$ obtained after the deletion.

Step 5: Removing dangling edges

Since the 2-collapses executed in Steps 1-3 are as described in Remark 19 and Figure 2, it is easy to check that for each $\mathbf{D}_{m,\ell}^{\zeta_1} \subset K(C)$, the edges that remain are of the form: $\{\{1,2\}^{\zeta_1}, \{2,3\}^{\zeta_1}, \{2,6\}^{\zeta_1}, \{5,6\}^{\zeta_1}, \{7,6\}^{\zeta_1}, \{b_k,6\}^{\zeta_1}, \{c_k,6\}^{\zeta_1}, \{d_k,6\}^{\zeta_1}, \{b_k,z_k\}^{\zeta_1}, \{z_k,y_k\}^{\zeta_1}, \{y_k,a_k\}^{\zeta_1}\} \cup F$ for $k \in [1,\ell]$, and,

- $F = \{\{v,4\}\}$ if $m = 1$, where $v = 6$, if ℓ is even, and $v = {}^{(\ell-1)}/_2 + 1$ if ℓ is odd,
- $F = \{s_1^{\zeta_1}, \{4,8\}^{\zeta_1}\}$ if $m = 2$ and $s_2^{\zeta_1}$ is removed as part of a 2-collapse,
- $F = \{s_2^{\zeta_1}, \{4,8\}^{\zeta_1}\}$ if $m = 2$ and $s_1^{\zeta_1}$ is removed as part of a 2-collapse.

We now execute the following 1-collapses (1-3 highlighted in green, and 4 highlighted in blue as illustrated in Figure 2).

1. Since ${5^{\zeta_1}}$, ${7^{\zeta_1}}$, ${c_k}^{\zeta_1}$, and ${d_k}^{\zeta_1}$ are free for all $k \in [1, \ell]$, for all $\mathbf{D}^{\zeta_1}_{m,\ell} \subset K(C)$, we execute the following collapses for all $k \in [1, \ell]$, for all $\mathbf{D}^{\zeta_1}_{m,\ell} \subset K(C)$:

$$(5^{\zeta_1}, \{5,6\}^{\zeta_1}), \quad (7^{\zeta_1}, \{7,6\}^{\zeta_1}), \quad (c_k^{\zeta_1}, \{c_k,6\}^{\zeta_1}), \quad \text{and} \quad (d_k^{\zeta_1}, \{d_k,6\}^{\zeta_1}).$$

2. Since the vertices ${4^{\zeta_1}}$ are free for all $\mathbf{D}^{\zeta_1}_{m,\ell} \subset K(C)$, for all $\mathbf{D}^{\zeta_1}_{m,\ell} \subset K(C)$:
   - if $m = 1$, we execute the collapse $(4^{\zeta_1}, \{4, v\}^{\zeta_1})$, where $v = 6$, if ℓ is even, and $v = {}^{(\ell-1)}\!/_2 + 1$ if ℓ is odd,
   - if $m = 2$, we execute the collapse $(4^{\zeta_1}, \{4,8\}^{\zeta_1})$.
3. Now, ${a_k}^{\zeta_1}$ become free for all $k \in [1, \ell]$, for all $\mathbf{D}^{\zeta_1}_{m,\ell} \subset K(C)$. So, we execute the collapses $({a_k}^{\zeta_2}, \{a_k, y_k\}^{\zeta_2})$ for all $k \in [1, \ell]$, for all $\mathbf{D}^{\zeta_1}_{m,\ell} \subset K(C)$.
4. Now, ${8^{\zeta_1}}$ become free for all $\mathbf{D}^{\zeta_1}_{2,\ell} \subset K(C)$. So, for all $\mathbf{D}^{\zeta_1}_{2,\ell} \subset K(C)$:
   - If $s_2^{\zeta_1}$ was removed as part of a 2-collapse, we execute the collapse $(8^{\zeta_1}, s_1^{\zeta_1})$,
   - else if $s_1^{\zeta_1}$ was removed as part of a 2-collapse, we execute the collapse $(8^{\zeta_1}, s_2^{\zeta_1})$.

   Note that because of the identifications, there may exist several $\mathbf{D}^{\zeta_2}_{m,\ell} \subset K(C)$ with points ${y_k}^{\zeta_2} \in \mathbf{D}^{\zeta_2}_{m,\ell}$ that are identical to ${8^{\zeta_1}}$. So, the above collapses $(8^{\zeta_1}, s_r^{\zeta_1}), r \in [1,2]$ may appear as $({y_k}^{\zeta_2}, \{y_k, z_k\}^{\zeta_2})$ in other dunce hats $\mathbf{D}^{\zeta_2}_{m,\ell} \subset K(C)$.

   The complex obtained after collapsing all the dangling edges is denoted by $K^5(C)$. So far, we have, $K^1(C) \searrow K^2(C) \searrow K^3(C)$ and $K^4(C) \searrow K^5(C)$.

Step 6: Collapsing the cycle-filling disks

The 1-complex H formed by the union of stems of $\mathbf{D}^{\zeta}_{m,\ell}$, for all $\zeta \in [1, |\Xi|]$ described in Section 4.3 is clearly a subcomplex of $K^5(C)$. Let $\mathcal{D} = K(C) \setminus K'(C)$ be the set described in Algorithm 1 obtained while building $K(C)$ from $K'(C)$. It is, in fact, easy to check that $K^5(C) = H \sqcup \mathcal{D}$. Next, we show that H is a connected graph.

▶ **Lemma 29.** *H is connected.*

Proof. First note that for every $\zeta \in [1, |\Xi|]$, the stem of $\mathbf{D}^{\zeta}_{m,\ell}$ is connected. In particular, the stem of $\mathbf{D}^{\zeta}_{m,\ell}$ connects 1^{ζ} and 3^{ζ} to z_k^{ζ} all $k \in \ell$.

Suppose G_i and G_j are two gates in C such that G_i is the predecessor of G_j. Then, in every dunce hat associated to G_i, there exists a t-edge that is connected to an s-edge to every dunce hat associated to G_j. That is, for all $p, q \in [1, n]$ there exists a $z_k^{(i,p)}$ that is identified to either $1^{(j,q)}$ or $3^{(j,q)}$. Thus, the stems of G_i are connected to the stems of G_j. Now, since C itself is a connected directed acyclic graph, it follows that the complex H which is the the union of stems of $\mathbf{D}^{\zeta}_{m,\ell}$, for all $\zeta \in [1, |\Xi|]$ is also connected. ◀

Now, as in Algorithm 1, let $\mathbf{M}$ be the upper-triangular matrix whose columns represent a basis $\mathcal{B}'$ of the cycle space of H. The cycle z_i is represented by column M^i. Let n_i denote the number of edges in z_i. Let the vertices $v_i^j \in z_i, j \in [1, n_i]$ and the edges $e_i^j \in z_i, j \in [1, n_i]$ be indexed so that e_i^1 represents the lowest entry (that is the pivot) for column M^i, and v_i^j and v_i^{j+1} form the endpoints of e_i^j. Simplices σ_i^j are indexed so that the vertices incident on σ_i^j are v_i^j and v_i^{j+1} and v_i. Please refer to Figure 6 for an example of a cycle z_i with six edges. The procedure to collapse all the disks in $K^5(C) \subset K(C)$ corresponding to cycles $z_i \in \mathcal{B}'$ is described in Algorithm 2.

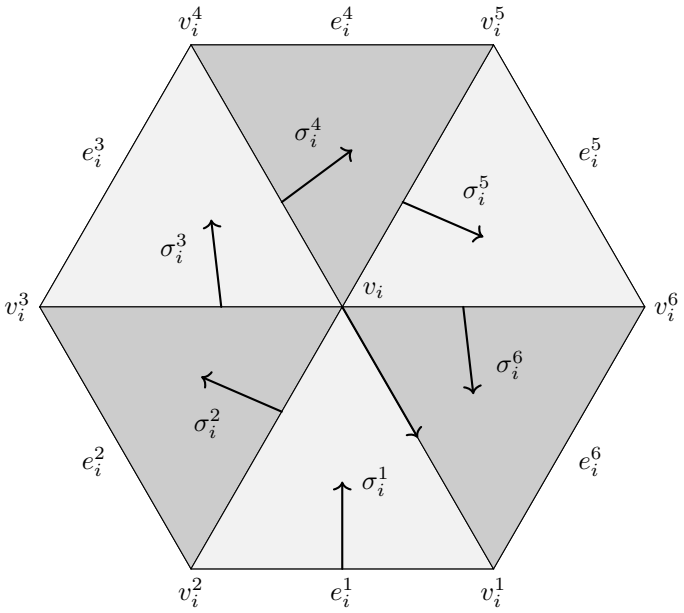


Figure 6 The above figure shows a triangulated disk that fills the cycle z_i. Here, e_i^1 is the pivot edge of z_i. The gradient field starts with a gradient pair that includes the pivot edge.

Algorithm 2 Procedure for collapsing cycle-filling disks

1: **for** $i \leftarrow 1, |\mathcal{B}'|$ **do**
2: Execute the collapse (e_i^1, σ_i^1).
3: **for** $j \leftarrow 2, n_i$ **do**
4: Execute the collapse $(\{v_i, v_i^j\}, \sigma_i^j)$.
5: **end for**
6: Execute the collapse $(v_i, \{v_i, v_i^1\})$.
7: **end for**
8: RETURN T.

Note that in Algorithm 2, it is possible to execute the collapse (e_i^1, σ_i^1) for each i because the matrix M of basis $\mathcal{B}'$ is upper-triangular. This guarantees that after collapsing all the disks corresponding to cycles $z_k, k \in [1, i-1]$, e_i^1 is free.

Denote the complex obtained at the end of Algorithm 2 as T.

Step 7: Collapsing the tree

Now observe that Algorithm 2 removes all simplices in $\mathcal{D}$ from $K^5(C)$. So, in particular, $T \subset H \subset K^5(C)$. Moreover, the pivot edges e_i^1 from cycles z_i are also removed as part of 2-collapses in Line 2 of Algorithm 2. In other words, $T = H \setminus \bigcup_{i=1}^{|\mathcal{B}'|} \{e_i^1\}$, where $\mathcal{B}'$ forms a basis for cycle space of H.

▷ Claim 30. T is a tree.

Proof. By Lemma 29, H is connected. Removal of each edge e_i^1 from H, decreases the β_1 of H by 1, whereas β_0 of H is unaffected. Hence, T is connected. Moreover, since we destroy all $|\mathcal{B}'|$ cycles of H, T has no cycles, proving the claim. ◀

Next, we greedily collapse the tree T to a vertex $v_0 \in K(C)$, which can be done in time linear in the size of T. Finally, we make v_0 critical.

Let $\mathcal{V}$ be the collection of gradient pairs arising out of all the collapses from Steps 1-7. Also, note that $K^1(C)$ is obtained from $K(C)$ by deletion of m critical 2-simplices. Then, $K^1(C) \searrow K^3(C)$. Then, $K^4(C)$ is obtained from $K^3(C)$ by deleting m critical 1-simplices. Then, $K^5(C)$ is obtained from $K^4(C)$ by executing some 1-collapses. Finally, $K^5(C) \searrow T \searrow v_0$. So, using Lemma 5, we conclude that given a circuit C with a satisfying assignment A of Hamming weight m, we can obtain a vector field $\mathcal{V}$ on $K(C)$ with m critical 2-simplices, m critical 1-simplices and a single critical vertex. Now, for a circuit C if the assignment A is, in fact, optimal, that is, assuming $m = \mathrm{OPT}_{\mathsf{MinMCS}}(C)$, then MinrMM for complex $K(C)$ has a solution of size $(2m+1)-1$ giving us the following proposition.

▶ **Proposition 31.** $\mathrm{OPT}_{\mathsf{MinrMM}}(K(C)) \leq 2 \cdot \mathrm{OPT}_{\mathsf{MinMCS}}(C)$.

We highlight the entire collapsing sequence in Figure 2 (a) and (b). First we perform the 2-collapses as described in Steps 2-3. Then, the 1-collapses for highlighted edges (in green) are executed. This is followed by 1-collapses for highlighted edges (in blue), whenever these edges are available. These edges may not be available if they are involved in 2-collapses in other dunce hats, or if they are made critical. After executing the above collapses, for edges in green and blue, we first execute the 2-collapses to erase all the cycle-filling disks, which leaves behind a tree supported by the edges in pink. The tree is then collapsed to a point.

5 Hardness results for Min-Reduced Morse Matching

For maps K and $\mathcal{I}$ described in Section 4.2, we can establish the following relations.

▶ **Proposition 32.** $\mathrm{OPT}_{\mathsf{MinrMM}}(K(C)) \leq 2 \cdot \mathrm{OPT}_{\mathsf{MinMCS}}(C)$.

Proof. For proof, please refer to Proposition 31 in Section 4.4. ◀

▶ **Proposition 33.** $m_{\mathsf{MinrMM}}(K(C), \tilde{\mathcal{V}}) \geq 2 \cdot m_{\mathsf{MinMCS}}(C, \mathcal{I}(C, \tilde{\mathcal{V}}))$

Proof. For proof, please refer to Proposition 62 in Appendix A.2. ◀

▶ **Proposition 34.** $\mathrm{OPT}_{\mathsf{MinrMM}}(K(C)) = 2\mathrm{OPT}_{\mathsf{MinMCS}}(C)$.

Proof. For proof, please refer to Proposition 64 in Appendix A.2. ◀

▶ **Proposition 35.**

$$m_{\mathsf{MinMCS}}(C, \mathcal{I}(C, \mathcal{V})) - \mathrm{OPT}_{\mathsf{MinMCS}}(C) \leq \frac{1}{2} m_{\mathsf{MinrMM}}(K(C), \mathcal{V}) - \mathrm{OPT}_{\mathsf{MinrMM}}(K(C)))$$

Proof. Combining Propositions 32 and 33 proves the claim. ◀

We will use the following straightforward fact about L-reductions.

▶ **Theorem 36** (Williamson, Shmoys [71])**.** *If there is an L-reduction with parameters μ and ν from a minimization problem A to a minimization problem B, and there is a $(1+\delta)$-approximation algorithm for B, then there is a $(1+\mu\nu\delta)$-approximation algorithm for A.*

Next, we shall use the following result by Alekhnovich et al. [1].

▶ **Theorem 37** (Theorem 3, [1])**.** *Unless* **NP** $\subseteq$ **QP**, *there is no polynomial time algorithm which can approximate* MinMCS *within a factor of $2^{\log^{(1-\epsilon)} n}$, for any $\epsilon > 0$.*

▶ **Theorem 38.** MinrMM *cannot be approximated within a factor of* $2^{\log^{(1-\epsilon)} n}$*, for any* $\epsilon > 0$ *unless* $\mathbf{NP} \subseteq \mathbf{QP}$*.*

Proof. From Proposition 33 and Proposition 35, we conclude that the reduction from MinMCS to MinrMM is a strict reduction with parameters $\mu = 2$ and $\nu = \frac{1}{2}$. By Theorem 36, if there exists a $(1+\delta)$-approximation algorithm for MinrMM, then there exists a $(1+\mu\nu\delta)$-algorithm for MinMCS. Using Theorem 37, the claim follows. ◀

Denote the standard parameterizations for MinMCS and MinrMM by MinMCS′ and MinrMM′ respectively. Using the map $K : C \to K(C)$ that transforms instances of MinMCS to instances of MinrMM, we define a new map $\hat{K} : (C, k) \to (K(C), k')$ that transforms instances of MinMCS′ to instances of MinrMM′, where we set $k' = 2k$.

▶ **Proposition 39.** *The map* $\hat{K}$ *from* MinMCS′ *to* MinrMM′ *is*

1. *an FPT reduction,*
2. *a* (δ, δ')*-gap preserving reduction for every function* δ*, where* $\delta'(k) = \delta(\lfloor \frac{k}{2} \rfloor)$*.*

Proof. 1. First note that, using Proposition 34, for any value of parameter k,

$$\mathrm{OPT}_{\mathsf{MinMCS}}(C) \leq k \Leftrightarrow \mathrm{OPT}_{\mathsf{MinrMM}}(K(C)) \leq 2k.$$

Then, the conclusion follows immediately from observing that complex $K(C)$ can be constructed in time polynomial in the size of C.

2. Suppose an instance (C, k) is a δ-gap instance of MinMCS′. That is, either $\mathrm{OPT}_{\mathsf{MinMCS}}(C) \leq k$ or $\mathrm{OPT}_{\mathsf{MinMCS}}(C) \geq k\delta(k)$. So, we have two cases to check:
Uusing Proposition 34,
   - $\mathrm{OPT}_{\mathsf{MinMCS}}(C) \leq k \Rightarrow \mathrm{OPT}_{\mathsf{MinrMM}}(K(C)) \leq 2k = k'$.
   - If $\mathrm{OPT}_{\mathsf{MinMCS}}(C) \geq k\delta(k) \Rightarrow \mathrm{OPT}_{\mathsf{MinrMM}}(K(C)) \geq 2k\delta(k) = k'\delta(\frac{k'}{2}) = k'\delta'(k')$. ◀

▶ **Theorem 40.** 1. MinrMM *is* $\mathbf{W}[\mathbf{P}]$*-hard.*
2. MinrMM *has no fixed-parameter tractable approximation algorithm with any approximation ratio function* ρ*, unless* $\mathbf{FPT} = \mathbf{W}[\mathbf{P}]$*.*

Proof. The first statement follows immediately from Proposition 39 and Theorem 16.

Eickmeyer et al. [25] provides a standard template to carry over FPT inapproximability results using gap preserving FPT reductions. Accordingly to prove the second statement, we closely follow the line of argument from [25, Corollary 12]. In this case, the strong FPT inapproximability result for MinMCS from Theorem 17 is carried over to MinrMM. We first reduce MinMCS′ to the approximation variant of MinrMM′. Assume there exists an FPT cost approximation algorithm for MinrMM with approximation ratio ρ, where ρ is any computable function.

Given an input (C, k) for MinMCS′, we first use the construction described in the proof of [25, Theorem 6]. Using this construction, we obtain a circuit C of size $|C| = f(k) \cdot |C|^{O(1)}$ for some computable function f in FPT time (with parameter k), such that

- $(C', \alpha(k))$ is a δ-gap instance for some $\alpha : \mathbb{N} \to \mathbb{N}$ and $\delta : \mathbb{N} \to \mathbb{R}_{>1}$,
- and $\rho(2\alpha(k)) < \delta(\alpha(k))$.

Note that satisfying the second condition becomes possible since we have no restriction on the function δ.

Using the FPT gap-preserving reduction described in Proposition 39 from MinMCS to MinrMM on the δ-gap instance $(C, \alpha(k))$, we get a δ'-gap instance $(K(C), 2\alpha(k))$ of MinrMM′ with $\delta'(2\alpha(k)) = \delta(\alpha(k))$. We run $\mathbb{A}$ on $(K(C), \rho(2\alpha(k)) \cdot 2\alpha(k))$.

If $\mathrm{OPT}_{\mathsf{MinrMM}}(K(C)) \leq 2\alpha(k)$, then

$$\rho(2\alpha(k)) \cdot 2\alpha(k) \geq \rho(\mathrm{OPT}_{\mathsf{MinrMM}}(K(C))) \cdot \mathrm{OPT}_{\mathsf{MinrMM}}(K(C))$$

and $\mathbb{A}$ accepts. If, on the other hand, $\mathrm{OPT}_{\mathsf{MinrMM}}(K(C)) \geq \delta'(2\alpha(k))2\alpha(k)$ then

$$\rho(2\alpha(k)) \cdot 2\alpha(k) < \delta(\alpha(k)) \cdot 2\alpha(k) = \delta'(2\alpha(k)) \cdot 2\alpha(k) \leq \mathrm{OPT}_{\mathsf{MinrMM}}(K(C)),$$

and $\mathbb{A}$ rejects.

Hence, using such an algorithm $\mathbb{A}$ we could devise an FPT cost approximable algorithm for MinMCS some computable function ρ, which in turn would imply $\mathbf{W}[\mathbf{P}] = \mathbf{FPT}$ using Theorem 17. ◀

6 Hardness results for Min-Morse Matching

Denoting the standard parameterizations for MinMM by MinMM′, we now consider the map $\tilde{K} : (K, p) \mapsto (K, p+1)$ that transforms instances of MinrMM′ (simplicial complexes) to instances of MinMM′ (identical simplicial complexes).

▶ **Proposition 41.** *The map $\tilde{K}$ from* MinrMM′ *to* MinMM′ *is*

1. *an FPT reduction,*
2. *a (δ, δ')-gap preserving reduction for every function δ, where $\delta'(p) = \frac{(p-1)\delta(p-1)+1}{p}$.*

Proof. 1. By definition, $\mathrm{OPT}_{\mathsf{MinMM}}(K) = \mathrm{OPT}_{\mathsf{MinrMM}}(K) + 1$. So, for any value of p,

$$\mathrm{OPT}_{\mathsf{MinrMM}}(K) \leq p \Leftrightarrow \mathrm{OPT}_{\mathsf{MinMM}}(K) \leq p + 1.$$

So, the conclusion follows immediately.

2. Suppose an instance (K, p) is a δ-gap instance of MinrMM′. That is, either $\mathrm{OPT}_{\mathsf{MinrMM}}(K) \leq p$ or $\mathrm{OPT}_{\mathsf{MinrMM}}(K) \geq p\delta(p)$. So, we have two cases to check:
   - If $\mathrm{OPT}_{\mathsf{MinrMM}}(K) \leq p$, then
   $$\mathrm{OPT}_{\mathsf{MinMM}}(K) \leq p + 1 = p'.$$
   - If $\mathrm{OPT}_{\mathsf{MinrMM}}(K) \geq p\delta(p)$, then
   $$\mathrm{OPT}_{\mathsf{MinMM}}(K) \geq p\delta(p) + 1 = p'\delta(p').$$ ◀

Combining Theorem 40 and Proposition 41, we obtain the following result:

▶ **Theorem 42.** MinMM *is* $\mathbf{W}[\mathbf{P}]$*-hard. Furthermore, it has no fixed-parameter tractable approximation algorithm within any approximation ratio function ρ, unless* $\mathbf{FPT} = \mathbf{W}[\mathbf{P}]$.

▶ **Definition 43** (Amplified complex)**.** *Given a pointed simplicial complex K with n simplices, the amplified complex $\widehat{K}$ is defined as the wedge sum of n copies of K.*

▶ **Lemma 44.** *For any 2-complex K, $\mathrm{OPT}_{\mathsf{MinMM}}(\widehat{K}) = n \cdot \mathrm{OPT}_{\mathsf{MinrMM}}(K) + 1$.*

Proof. It is easy to check that the optimal vector field on $\widehat{K}$ is obtained by repeating the optimal vector field on K on each of the n copies of K in $\widehat{K}$, while making the distinguished vertex of $\widehat{K}$ the unique critical vertex in $\widehat{K}$. ◀

▶ **Lemma 45.** *Using a vector field $\widehat{\mathcal{V}}$ on $\widehat{K}$ with $m+1$ critical simplices, one can compute a vector field $\mathcal{V}$ on K with at most $\lfloor \frac{m}{n} \rfloor + 1$ critical simplices in polynomial time.*

Proof. Using Lemma 3, we can assume without loss of generality that $\widehat{\mathcal{V}}$ has the distinguished vertex as its unique critical simplex. Restricting $\widehat{\mathcal{V}}$ to each of the n copies of K, the claim follows. ◀

▶ **Proposition 46.** *For a fixed $\epsilon > 0$, let $\rho = f(n)$, where $f(n) = o(n)$. Then, for any $\delta \in (0,1)$ and $\varrho = f(n) - \delta$, if there exists a ϱ-factor approximation algorithm for* MinMM*, then there exists a ρ-factor approximation algorithm for* MinrMM*.*

Proof. For a complex K, the optimal value of MinrMM on K is denoted by $\mathrm{OPT}_{\mathsf{MinrMM}}(K)$. Suppose that there exists a ϱ-factor approximation algorithm $\mathbb{A}$ for MinMM. If we apply $\mathbb{A}$ on $\widehat{K}$, then using Lemma 44, we obtain a vector field with at most $\varrho\,(n \cdot \mathrm{OPT}_{\mathsf{MinrMM}}(K) + 1)$ critical simplices. Then, using Lemma 45, we can compute a vector field $\mathcal{V}$ on K with at most $m(\mathcal{V})$ critical simplices, where

$$\begin{aligned}
m(\mathcal{V}) &\leq \left\lfloor \frac{\varrho \cdot n \cdot \mathrm{OPT}_{\mathsf{MinrMM}}(K) + \varrho - 1}{n} \right\rfloor + 1 \\
&\leq \left\lfloor \varrho \cdot \mathrm{OPT}_{\mathsf{MinrMM}}(K) + \frac{\varrho - 1}{n} \right\rfloor + 1 \\
&\leq \lfloor \varrho \cdot \mathrm{OPT}_{\mathsf{MinrMM}}(K) \rfloor + \left\lfloor \frac{\varrho - 1}{n} \right\rfloor + 2 && \text{using } \lfloor x + y \rfloor \leq \lfloor x \rfloor + \lfloor y \rfloor + 1 \\
&= \lfloor \varrho \cdot \mathrm{OPT}_{\mathsf{MinrMM}}(K) \rfloor + 2 && \text{using } \left\lfloor \frac{\varrho - 1}{n} \right\rfloor = 0 \text{ for large } n,
\end{aligned}$$

which gives us

$$\begin{aligned}
m(\mathcal{V}) - 1 &\leq \varrho \cdot \mathrm{OPT}_{\mathsf{MinrMM}}(K) + 1 \\
m(\mathcal{V}) - 1 &\leq \rho \cdot \mathrm{OPT}_{\mathsf{MinrMM}}(K) - \delta \cdot \mathrm{OPT}_{\mathsf{MinrMM}}(K) + 1 \\
&\leq \rho \cdot \mathrm{OPT}_{\mathsf{MinrMM}}(K) && \text{assuming } \mathrm{OPT}_{\mathsf{MinrMM}}(K) > \frac{1}{\delta}.
\end{aligned}$$

The above analysis shows that one can obtain a ρ-factor approximation algorithm for MinrMM assuming a ϱ factor approximation algorithm for MinMM. Note that $n^{\frac{1}{\delta}}$ is bounded by a polynomial in n given the fact that $\frac{1}{\delta}$ is a constant. So, we can assume without loss of generality that $\mathrm{OPT}_{\mathsf{MinrMM}}(K) > \frac{1}{\delta}$ based on the observation by Joswig and Pfetsch [39] that if $\mathrm{OPT}_{\mathsf{MinrMM}}(K) \leq c$, for some constant c, then one can find the optimum in $O(n^c)$ time. ◀

Combining Theorem 38 and Proposition 46, we conclude that for a fixed $\epsilon > 0$, MinMM cannot be approximated within a factor of $2^{\log^{(1-\epsilon)} n} - \delta$, for any $\delta > 0$, unless $\mathbf{NP} \subseteq \mathbf{QP}$. But, in order to get rid of the δ-term in the inapproximability bound for MinMM, we can do slightly better by allowing ϵ to vary. To make this precise, suppose there exists an $\iota > 0$ such that MinMM can be approximated within a factor of $2^{\log^{(1-\iota)} n}$, and let $\delta \in (0,1)$. Then, using Proposition 46, this would give a $2^{\log^{(1-\iota)} n} + \delta$ approximation algorithm for MinrMM. However, one can always find an $\epsilon > 0$ such that $2^{\log^{(1-\iota)} n} + \delta = O(2^{\log^{(1-\epsilon)} n})$. Then, for sufficiently large n, $2^{\log^{(1-\iota)} n} + \delta < 2^{\log^{(1-\epsilon)} n}$.

Hence, the assumption of a $2^{\log^{(1-\iota)} n}$-factor approximation algorithm for MinMM contradicts Theorem 38. We can thus make the following claim.

▶ **Theorem 47.** *For any $\epsilon > 0$,* MinMM *cannot be approximated within a factor of $2^{\log^{(1-\epsilon)} n}$, unless* $\mathbf{NP} \subseteq \mathbf{QP}$.

7 An approximation algorithm for Min-Morse Matching

In this section, we assume without loss of generality that the input complex K is connected. The algorithm can be described as follows. Given a 2-complex K, let n be the number of 2-simplices. Assume without loss of generality that $\log n$ is an integer that divides n. Partition the set of 2-simplices of K arbitrarily into $\log n$ parts each of size $\frac{n}{\log n}$. Writing S for the partition, we note that the power set $P(S)$ of the parts has n elements. The 2-simplices that belong to a part $s \in S$ is denoted by $K_s^{(2)}$. Each element of $P(S)$ gives us a subset $\hat{S}$ of S. To each $\hat{S}$ we can associate a binary incidence vector $\mathbf{j}(\hat{S})$ of length $\log n$ in the natural way. Let $\hat{K}$ be a complex induced by the 2-simplices belonging to the parts that belong to some $\hat{S} \subset S$. In this case, we may also write $\hat{K}$ as $\hat{K} = K(\mathbf{j}(\hat{S}))$ to emphasize the data from which $\hat{K}$ can be constructed. Compute such a complex $\hat{K}$ for each subset $\hat{S}$, and let $\hat{S}_{\max}$ be the subset of largest cardinality whose induced complex $\hat{K}_{\max}$ is erasable. In particular, $\hat{K}_{\max} \searrow L$ where L is a 1-complex. Make all the 2-simplices in $K \setminus \hat{K}_{\max}$ critical. The gradient on $\hat{K}_{\max}$ is comprised of the erasing gradient of $\hat{K}_{\max}$, namely $\mathcal{V}^2$, combined with the optimal gradient for L, namely $\mathcal{V}^1$. In what follows, we will show that this simple algorithm provides a $O(\frac{n}{\log n})$-factor approximation for Min-Morse Matching on 2-complexes.

▶ **Lemma 48.** *Let $\hat{K}_{\max} = K(\mathbf{j}(\hat{S}))$ for some $\hat{S}$. Let $w_{\mathbf{j}}$ be the Hamming weight of $\mathbf{j}(\hat{S})$, and let $\gamma = \log n - w_{\mathbf{j}}$. Then, every Morse matching on K has at least γ critical 2-simplices.*

Proof. Suppose that there exists a gradient vector field $\mathcal{V}^2$ with μ critical 2-simplices where $\mu < \gamma$. Let Ψ denote the critical 2-simplices of $\mathcal{V}^2$. Define $\hat{K}^{(2)}_{new}$ as follows:

$$\hat{K}^{(2)}_{new} = \bigcup_{\substack{s \in S, \\ \Psi \cap K_s^{(2)} = \emptyset}} K_s^{(2)}.$$

As before, let $\hat{K}_{new}$ be the complex induced by simplices in $\hat{K}^{(2)}_{new}$. Then, $\hat{K}_{new} \subset K \setminus \Psi$. However, $K \setminus \Psi$ is erasable via gradient $\mathcal{V}^2$. So, by Lemma 4, $\hat{K}_{new}$ is erasable. But this contradicts the maximality of $\hat{K}_{\max}$, proving the claim. ◀

We denote the critical k-simplices of $\mathcal{V}^2 \cup \mathcal{V}^1$ by c_k.

▶ **Lemma 49.** *The gradient vector field $\mathcal{V}^2 \cup \mathcal{V}^1$ over K has at most $\beta_1 - \beta_2 + 1 + 2\gamma \cdot \left(\frac{n}{\log n}\right)$ critical simplices.*

Proof. From Lemma 48, we have

$$c_2 = \gamma \cdot \left(\frac{n}{\log n}\right). \tag{1}$$

By [39, Lemma 4.2], K^1 is connected, and one can compute a gradient vector field $\mathcal{V}^1$ on K^1 with a single critical vertex in linear time using depth first search starting from an arbitrary vertex in K^1 (see, e.g., [60]). We have by [29, Theorem 1.7],

$$c_0 - c_1 + c_2 = \beta_0 - \beta_1 + \beta_2.$$

Since $\beta_0 = c_0 = 1$, we have,

$$c_2 - \beta_2 = c_1 - \beta_1. \tag{2}$$

Thus, combining Equation (1) and Equation (2), we have

$$c_1 = \beta_1 - \beta_2 + \gamma \cdot \left(\frac{n}{\log n}\right).$$

The claim follows. ◀

▶ **Theorem 50.** *The exists a $O(\frac{n}{\log n})$-factor approximation algorithm for* Min-Morse Matching *on 2-complexes.*

Proof. To begin with, we know from Tancer [69, Proposition 5] that a 2-complex is erasable if and only if greedily collapsing triangles yields a 1-dimensional complex. That is, erasability of a complex can checked in polynomial time. Since we check the erasability of $O(n)$ complexes each of size $O(n)$, the algorithm terminates in polynomial time.

Now, let $\mathcal{V}_{\min}$ be an optimal gradient vector field. By Lemma 48, $\mathcal{V}_{\min}$ has at least γ critical 2-simplices. By weak Morse inequalities [29, Theorem 1.7], the number of critical 1-simplices of $\mathcal{V}_{\min}$ is at least β_1, and the number of critical 0-simplices of $\mathcal{V}_{\min}$ is $\beta_0 = 1$. Thus, an optimal gradient vector field has at least $O(\gamma + \beta_1)$ critical simplices. Combining this observation with Lemma 49, it follows that the algorithm described in this section provides an $O(\frac{n}{\log n})$-factor approximation for Min-Morse Matching on 2-complexes. ◀

8 Morse matchings for Costa–Farber complexes

The strong hardness results established in Section 6 belie what is observed in computer experiments for both structured as well as random instances [33, 60]. In particular, the structured instances generated by Lutz [38, 50] and by the RedHom and CHomP groups [33] and the random instances that come from Linial–Meshulam model [54] and the Costa–Farber model (referred to as type-2 random complexes in [60]) turn out to be 'easy' for Morse matching [60]. We use the terms 'easy' and 'hard' in an informal sense. Here, by *easy* instances, we mean those instances for which simple heuristics give near-optimal matchings, and by *hard* instances we mean instances for which known heuristics produce matchings that are far from optimal. Unfortunately, the approximation bounds in [60], and in Section 7 of this paper do not explain the superior performance of simple heuristics in obtaining near-optimal matchings. Below, we provide some justification for this phenomena from the lens of random complexes. We start with the definition of the *apparent pairs gradient* [7]. Apparent pairs gradient is a powerful optimization tool in softwares like Ripser [7] and Eirene [34] that compute persistent homology of Rips filtrations.

Let K be a d-dimensional simplicial complex, and let V denote the set of vertices in K. Suppose that the vertices in V are equipped with an indexing. For two simplices $\sigma, \tau \in K$, we write $\sigma \prec_{\mathsf{K}} \tau$ if σ comes before τ in the lexicographic ordering.

Following [7], we call a pair of simplices (σ, τ) of K an *apparent pair* of K if both

- σ is the lexicographically highest facet of τ, and
- τ is the lexicographically lowest cofacet of σ.

As observed in [7, Lemma 3.5], the collection of all the apparent pairs in $\mathsf{X}(n, \mathbf{p})$ forms a discrete gradient on $\mathsf{X}(n, \mathbf{p})$. We denote this gradient by $\mathcal{V}_1$.

In [40, 41], (a variation of) the apparent pairs gradient was used to construct matchings with provably few critical simplices on Vietoris-Rips complexes built on a random set of points in space (and on random clique complexes). The Morse numbers from this matching are then used to obtain upper bounds on the Betti numbers of respective degrees. In what

follows, we observe that the apparent pairs gradient also provides very efficient matchings for an even larger class of random combinatorial complexes, namely, the multiparameteric complexes of Costa and Farber [16, 17, 18, 19]. Our analysis closely follows Kahle's work on random clique complexes [40, Section 7].

The Costa–Farber complex $\mathsf{X}(n, \mathbf{p})$ on a vertex set V of size n and a probability vector $\mathbf{p} = \{p_1, \dots, p_{n-1}\}$ can be described as follows. First, add all the vertices in V to the complex $\mathsf{X}(n, \mathbf{p})$. Next, include every possible edge independently with probability p_1. So far, this is the same as the Erdős–Rényi graph $G(n, p_1)$. That is, the 1-skeleton $\mathsf{X}_1 = G(n, p_1)$. Subsequently, for every 3-clique in $G(n, p_1)$, include a 2-simplex independently with probability p_2 to obtain the 2-skeleton X_2. More generally, consider an r-simplex σ defined on an $r+1$-element vertex set $V_r \subset V$. If all the simplices of the set $\partial\sigma$ are present in X_{r-1}, then include σ in X_r with probability p_r. Do this for every for every $r+1$-element subset of V to obtain the r-skeleton X_r. Following this process for every $r \in [n-1]$ gives the complex $\mathsf{X}_{n-1} = \mathsf{X}(n, \mathbf{p})$.

Note that the Costa–Farber model generalizes both the d-dimensional Linial–Meshulam model $\mathsf{Y}_d(n, p)$ as well as the random clique complex model $\mathsf{X}(n, p)$. For instance, when $p_1 = p$ and $p_i = 1$ for $i \in [2, n-1]$, we obtain the random clique complex model with parameter p. When $p_i = 1$ for $i \in [d-1]$, $p_d = p$, and $p_i = 0$ for $i \in [d+1, n-1]$, we recover the d-dimensional Linial–Meshulam model with parameter p.

Let $\sigma = \{v_0, v_1 \dots v_r\}$ be an r-dimensional simplex of $\mathsf{X}(n, \mathbf{p})$, where the vertices are labelled so that $v_i \in [n]$. Then, σ is a critical simplex in $\mathcal{V}_1$ if and only if:

1. $\sigma \in \mathsf{X}(n, \mathbf{p})$, and
2. Every lexicographically lower cofacet τ of σ is not in $\mathsf{X}(n, \mathbf{p})$. That is, if $\tau = \sigma \bigcup \{v'\}$ and $v' \prec_{\mathsf{K}} v_r$, then $\tau \notin \mathsf{X}(n, \mathbf{p})$.

Using independence, the probability that σ is critical is given by

$$\prod_{i=1}^{r} p_i^{\binom{r+1}{i+1}} \left(1 - \prod_{\ell=1}^{r+1} p_\ell^{\binom{r+1}{\ell}}\right)^{v_0 - 1}$$

Let m_r denote the total number of critical r-simplices. Since there are $\binom{n-j}{r}$ possible choices for σ with $v_0 = j$, and since $v_0 \in [1, n-r]$, we obtain the following expression for $\mathbb{E}(m_r)$.

$$\begin{aligned}
\mathbb{E}(m_r) &= \sum_{j=1}^{n-r} \binom{n-j}{r} \prod_{i=1}^{r} p_i^{\binom{r+1}{i+1}} \left(1 - \prod_{\ell=1}^{r+1} p_\ell^{\binom{r+1}{\ell}}\right)^{j-1} \\
&\leq \binom{n}{r} \prod_{i=1}^{r} p_i^{\binom{r+1}{i+1}} \sum_{j=1}^{n-r} \left(1 - \prod_{\ell=1}^{r+1} p_\ell^{\binom{r+1}{\ell}}\right)^{j-1} \\
&\leq \binom{n}{r} \prod_{i=1}^{r} p_i^{\binom{r+1}{i+1}} \sum_{j=1}^{\infty} \left(1 - \prod_{\ell=1}^{r+1} p_\ell^{\binom{r+1}{\ell}}\right)^{j-1} \\
&= \left(\binom{n}{r} \prod_{i=1}^{r} p_i^{\binom{r+1}{i+1}}\right) \cdot \left(\prod_{\ell=1}^{r+1} p_\ell^{-\binom{r+1}{\ell}}\right)
\end{aligned}$$

Let c_r denote the total number of r-simplices in $\mathsf{X}(n, \mathbf{p})$. Then, the expected number of r-simplices in $\mathsf{X}(n, \mathbf{p})$ is given by

$$\mathbb{E}(c_r) = \binom{n}{r+1} \prod_{i=1}^{r} p_i^{\binom{r+1}{i+1}}.$$

Therefore,

$$\frac{\mathbb{E}(m_r)}{\mathbb{E}(c_r)} \leq \frac{\left(\binom{n}{r}\prod_{i=1}^{r} p_i^{\binom{r+1}{i+1}}\right)\cdot\left(\prod_{\ell=1}^{r+1} p_\ell^{-\binom{r+1}{\ell}}\right)}{\binom{n}{r+1}\prod_{i=1}^{r} p_i^{\binom{r+1}{i+1}}} = \frac{(r+1)}{(n-r)\prod_{\ell=1}^{r+1} p_\ell^{\binom{r+1}{\ell}}}$$

When r is a fixed constant,

$$\frac{\mathbb{E}(m_r)}{\mathbb{E}(c_r)} = O\left(\frac{1}{n}\prod_{\ell=1}^{r+1} p_\ell^{-\binom{r+1}{\ell}}\right). \tag{3}$$

Assuming the denominator $n\prod_{\ell=1}^{r+1} p_\ell^{\binom{r+1}{\ell}} \to \infty$, we obtain $\frac{\mathbb{E}(m_r)}{\mathbb{E}(c_r)} = o(1)$.

The second gradient.

Kahle [40, Section 7] describes an alternative method for designing gradients on random clique complexes with parameter p_1 for which the following holds true.

$$\frac{\mathbb{E}(m_r)}{\mathbb{E}(c_r)} = \frac{\binom{r+2}{2}\binom{n}{r+2}}{\binom{n}{r+1}}p_1^r. \tag{4}$$

To establish Equation (4), Kahle considers the following strategy: Randomly match every r-simplex to one of its facets. This strategy doesn't give you a discrete gradient on the nose as there will be $r-1$-simplices that are matched to more than one cofacets, and there might also be some cycles. These events are termed as bad events. It suffices to make one pair of simplices critical per bad event. Once the corresponding simplices associated to all bad events are made critical, one indeed obtains a discrete gradient $\mathcal{V}_2$. Bounding the expected number of bad events $\mathcal{B}_r$ therefore gives a bound on the expected number of critical simplices. It is then shown that the total number of bad events for dimension r is given by

$$\mathbb{E}(\mathcal{B}_r) = \binom{r+2}{2}\binom{n}{r+2}p_1^{\binom{r+2}{2}-1}.$$

This is because each bad event contains at least one pair of r-simplices meeting in an $(r-1)$-simplex. The total number of vertices involved are, therefore, $r+2$. So there are $\binom{n}{r+2}$ choices of $(r+2)$-vertex sets, and for every choice of an $(r+2)$-vertex set, there are $\binom{r+2}{2}$ choices of pairs of adjacent r-simplices. Finally, for such a bad pair to be actually present all but one edge must be present among all the $(r+2)$ vertices. The expected number of simplices of dimension r is given by

$$\mathbb{E}(c_r) = \binom{n}{r+1}p_1^{\frac{r(r+1)}{2}}.$$

Dividing the two, we obtain

$$\frac{\mathbb{E}(\mathcal{B}_r)}{\mathbb{E}(c_r)} = \frac{\binom{r+2}{2}\binom{n}{r+2}}{\binom{n}{r+1}}p_1^r. \tag{5}$$

Note that if r is a fixed constant, and $np_1^r \to 0$, then

$$\frac{\mathbb{E}(\mathcal{B}_r)}{\mathbb{E}(c_r)} = o(1).$$

Kahle's method [40] for constructing gradients on (typically) sparse random clique complexes easily extends to the Costa–Farber model, and Equation (5) generalizes as follows:

$$\frac{\mathbb{E}(\mathcal{B}_r)}{\mathbb{E}(c_r)} = \frac{\binom{r+2}{2}\binom{n}{r+2}}{\binom{n}{r+1}} \prod_{j=1}^{r} p_j^{\binom{r}{j}}.$$

Let r be a fixed constant, and m_r denote the critical simplices of $\mathcal{V}_2$. Then, we obtain

$$\frac{\mathbb{E}(\mathcal{B}_r)}{\mathbb{E}(c_r)} = \frac{\mathbb{E}(m_r)}{\mathbb{E}(c_r)} = O\left(n \prod_{j=1}^{r} p_j^{\binom{r}{j}}\right). \tag{6}$$

When the parameters $\mathbf{p}$ in $\mathsf{X}(n, \mathbf{p})$ are such that

$$n \prod_{j=1}^{r} p_j^{\binom{r}{j}} \to 0, \quad \text{we obtain} \quad \frac{\mathbb{E}(m_r)}{\mathbb{E}(c_r)} = o(1).$$

Note that $n \prod_{\ell=1}^{r+1} p_\ell^{\binom{r+1}{\ell}} \to \infty$ typically leads to dense complexes, while $n \prod_{j=1}^{r} p_j^{\binom{r}{j}} \to 0$ typically leads to sparse complexes. From Equations (3) and (6), in both cases, we obtain very good discrete gradients for typical instances. In particular, we obtain the following theorem.

▶ **Theorem 51.** *Let r be a fixed dimension. Then, for the regimes of Costa–Farber complexes $\mathsf{X}(n, \boldsymbol{p})$ that satisfy*

$$n \prod_{\ell=1}^{r+1} p_\ell^{\binom{r+1}{\ell}} \to \infty \quad \text{or} \quad n \prod_{j=1}^{r} p_j^{\binom{r}{j}} \to 0$$

there exist respective discrete gradients that satisfy $\frac{\mathbb{E}(m_r)}{\mathbb{E}(c_r)} = o(1)$.

The following statement is obtained by specializing the above analysis to Linial–Meshulam complexes.

▶ **Corollary 52.** *For the regimes of Linial-Meshulam complexes $\mathsf{Y}_d(n, p)$ that satisfy*

$$(np \to \infty \text{ and } r + 1 = d) \quad \text{or} \quad (np \to 0 \text{ and } r = d)$$

there exist respective discrete gradients that satisfy $\frac{\mathbb{E}(m_r)}{\mathbb{E}(c_r)} = o(1)$.

In other words, the above corollary says that if ($p = o(\frac{1}{n})$ and $r + 1 = d$), or if ($p = \omega(\frac{1}{n})$ and $r = d$) for Linial–Meshulam complexes $\mathsf{Y}_d(n, p)$, then there exist respective discrete gradients that satisfy $\frac{\mathbb{E}(m_r)}{\mathbb{E}(c_r)} = o(1)$.

To further refine our analysis, we now define a gradient $\mathcal{U}$ on the entire Linial–Meshulam complex $\mathsf{Y}_d(n, p)$ as follows:

- when $np \to \infty$, let $\mathcal{U}$ be the apparent pairs gradient on $\mathsf{Y}_d(n, p)$.
- when $np \to 0$, $\mathcal{U}$ is comprised of
  - the apparent pairs gradient for matching $k-1$-simplices to k-simplices for $k \in [d-1]$
  - $\mathcal{V}_2$ for matching the remaining $(d-1)$-simplices to d-simplices.

Also, let opt be the optimal discrete gradient on $\mathsf{Y}_d(n, p)$. Then, the following holds true.

▶ **Theorem 53.** *For the regimes of Linial-Meshulam complexes $\mathsf{Y}_d(n, p)$ that satisfy*

$$np \to \infty \quad \text{or} \quad np \to 0$$

the discrete gradient $\mathcal{U}$ satisfies $\frac{\mathbb{E}(|\mathcal{U}|)}{\mathbb{E}(|\mathsf{opt}|)} \to 1$ as $n \to \infty$.

For proof, we refer the reader to Appendix B.

We would like to contrast the above observations with a known result from literature. We start with a definition. If a d-dimensional simplicial complex collapses to a $d-1$-dimensional complex, then we say that it is *d-collapsible*. The following result concerning the d-collapsibility threshold was established by Aronshtam et al. [3, 4]. See also Theorem 23.3.17 in Chapter 23 of [70].

▶ **Theorem 54** ([3,4]). *There exists a dimension dependent constant c_d for Linial–Meshulam complexes $\mathsf{Y}_d(n,p)$ such that*

- *If $p \geq \frac{c}{n}$ where $c > c_d$ then with high probability $\mathsf{Y}_d(n,p)$ is not d-collapsible,*
- *and if $p \leq \frac{c}{n}$ where $c < c_d$ then $\mathsf{Y}_d(n,p)$ is d-collapsible with probability bounded away from zero.*

Therefore, from Theorems 53 and 54, we conclude that away from the d-collapsibility threshold, we expect to have very good gradients. It is natural to ask what happens at the threshold? In relation to what is known for hard satisfiability instances [67, 72], are complexes of dimension larger than 2 sampled at the collapsibility thresholds of $\mathsf{Y}_d(n,p)$ and more generally $\mathsf{X}(n,\mathbf{p})$ hard? The experiments in [60] do not address this question. Secondly, for 2-complexes is it possible to define a simple random model built out of gluing dunce hats geared specifically towards generating hard instances for Min-Morse Matching for a wide range of parameter values? We are optimistic about affirmative answers to both questions, but leave this topic for future investigation.

9 Conclusion and Discussion

In this paper, we establish several hardness results for Min-Morse Matching. In particular, we show that for complexes of all dimensions, Min-Morse Matching with standard parameterization is $\mathbf{W}[\mathbf{P}]$-hard and has no FPT approximation algorithm for any approximation factor. We also establish novel (in)approximability bounds for Min-Morse Matching on 2-complexes. While we believe that this paper provides a nearly complete picture of complexity of Morse matchings, we conclude the paper with two remarks.

Strengthening of hardness results

We conjecture that for complexes of dimension $d > 2$, Min-Morse Matching does not admit an $f(n)$-approximation algorithm for any $f = o(n)$. In particular, a result of this nature would show that while the problem is hard for complexes of all dimensions, it is, in fact, slightly harder for higher dimensional complexes when compared to 2-dimensional complexes, from an inapproximability standpoint.

Hardness of other related combinatorial problems

In [10], the complexity of the following problem (Erasability Expansion Height) was studied: Given a 2-dimensional simplicial complex K and a natural number p, does there exist a sequence of expansions and collapses that take K to a 1-complex such that this sequence has at most p expansions? A more natural variant (Expansion Height) would be to study the complexity of determining sequences of expansions and collapses (with at most p expansions) that take K to a point. From what we understand, the only obstruction in [10] towards considering the complexity of determining whether K is simple homotopy equivalent to a point with bounded number of expansions is that the gadgets used in [10] have 1-cycles.

We believe that an immediate application of the cycle filling method introduced in this paper would be towards establishing $\mathbf{W}[\mathbf{P}]$-completeness for Expansion Height.

References

1 Michael Alekhnovich, Sam Buss, Shlomo Moran, and Toniann Pitassi. Minimum propositional proof length is np-hard to linearly approximate. In Luboš Brim, Jozef Gruska, and Jiří Zlatuška, editors, *Mathematical Foundations of Computer Science 1998*, pages 176–184. Springer Berlin Heidelberg, 1998.

2 Madjid Allili, Tomasz Kaczynski, Claudia Landi, and Filippo Masoni. Algorithmic construction of acyclic partial matchings for multidimensional persistence. In Walter G. Kropatsch, Nicole M. Artner, and Ines Janusch, editors, *DGCI 2017: Discrete Geometry for Computer Imagery*, pages 375–387. Springer, 2017. doi:10.1007/978-3-319-66272-5_30.

3 Lior Aronshtam and Nathan Linial. When does the top homology of a random simplicial complex vanish? *Random Structures & Algorithms*, 46(1):26–35, 2015.

4 Lior Aronshtam, Nathan Linial, Tomasz Łuczak, and Roy Meshulam. Collapsibility and vanishing of top homology in random simplicial complexes. *Discrete & Computational Geometry*, 49(2):317–334, 2013.

5 Sanjeev Arora. *Probabilistic Checking of Proofs and Hardness of Approximation Problems*. PhD thesis, U C Berkeley, Berkeley, CA, USA, 1995. UMI Order No. GAX95-30468.

6 U. Bauer, H. Edelsbrunner, G. Jabłoński, and M. Mrozek. Čech–delaunay gradient flow and homology inference for self-maps. *Journal of Applied and Computational Topology*, 4(4):455–480, 2020. URL: https://doi.org/10.1007/s41468-020-00058-8, doi:10.1007/s41468-020-00058-8.

7 Ulrich Bauer. Ripser: efficient computation of vietoris–rips persistence barcodes. *Journal of Applied and Computational Topology*, 2021. doi:10.1007/s41468-021-00071-5.

8 Ulrich Bauer and Herbert Edelsbrunner. The morse theory of Čech and delaunay complexes. *Transactions of the American Mathematical Society*, 369(5):3741–3762, 2017. URL: http://dx.doi.org/10.1090/tran/6991, doi:10.1090/tran/6991.

9 Ulrich Bauer and Abhishek Rathod. Hardness of approximation for Morse Matching. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '19, pages 2663–2674, Philadelphia, PA, USA, 2019. Society for Industrial and Applied Mathematics.

10 Ulrich Bauer, Abhishek Rathod, and Jonathan Spreer. Parametrized Complexity of Expansion Height. In Michael A. Bender, Ola Svensson, and Grzegorz Herman, editors, *27th Annual European Symposium on Algorithms (ESA 2019)*, volume 144 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 13:1–13:15, Dagstuhl, Germany, 2019. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. doi:10.4230/LIPIcs.ESA.2019.13.

11 Piotr Brendel, Pawel Dlotko, Graham Ellis, Mateusz Juda, and Marian Mrozek. Computing fundamental groups from point clouds. *Appl. Algebra Eng., Commun. Comput.*, 26(1-2):27–48, March 2015. doi:10.1007/s00200-014-0244-1.

12 Benjamin A. Burton, Thomas Lewiner, João Paixão, and Jonathan Spreer. Parameterized complexity of discrete Morse theory. *ACM Trans. Math. Software*, 42(1):Art. 6, 24, 2016. doi:10.1145/2738034.

13 F. Cazals, F. Chazal, and T. Lewiner. Molecular shape analysis based upon the Morse–Smale complex and the Connolly function. In *Proceedings of the Nineteenth Annual Symposium on Computational Geometry*, SCG '03, pages 351–360. ACM, 2003. doi:10.1145/777792.777845.

14 Manoj K. Chari. On discrete Morse functions and combinatorial decompositions. *Discrete Mathematics*, 217(1-3):101–113, 2000. doi:10.1016/S0012-365X(99)00258-7.

15 Yijia Chen, Martin Grohe, and Magdalena Grüber. On parameterized approximability. In Hans L. Bodlaender and Michael A. Langston, editors, *Parameterized and Exact Computation*, pages 109–120, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

16 A. Costa and M. Farber. Large random simplicial complexes, i. *Journal of Topology and Analysis*, 08(03):399–429, 2016. doi:10.1142/S179352531650014X.

17 A. Costa and M. Farber. Large random simplicial complexes, ii; the fundamental group. *Journal of Topology and Analysis*, 09(03):441–483, 2017. doi:10.1142/S1793525317500170.

18 A. Costa and M. Farber. Large random simplicial complexes, iii the critical dimension. *Journal of Knot Theory and Its Ramifications*, 26(02):1740010, 2017. doi:10.1142/S0218216517400107.

19 Armindo Costa and Michael Farber. *Random Simplicial Complexes*, pages 129–153. Springer International Publishing, Cham, 2016.

20 Justin Curry, Robert Ghrist, and Vidit Nanda. Discrete morse theory for computing cellular sheaf cohomology. *Foundations of Computational Mathematics*, 16(4):875–897, 2016. URL: https://doi.org/10.1007/s10208-015-9266-8, doi:10.1007/s10208-015-9266-8.

21 Tamal K. Dey, Jiayuan Wang, and Yusu Wang. Graph Reconstruction by Discrete Morse Theory. In *34th International Symposium on Computational Geometry (SoCG 2018)*, 2018. doi:10.4230/LIPIcs.SoCG.2018.31.

22 R. G. Downey and M. R. Fellows. *Parameterized complexity*. Monographs in Computer Science. Springer-Verlag, New York, 1999. doi:10.1007/978-1-4612-0515-9.

23 Rodney G. Downey and Michael R. Fellows. *Fundamentals of Parameterized Complexity*. Springer Publishing Company, Incorporated, 2013.

24 Ömer Eğecioğlu and Teofilo F. Gonzalez. A computationally intractable problem on simplicial complexes. *Computational Geometry*, 6(2):85 – 98, 1996. doi:https://doi.org/10.1016/0925-7721(95)00015-1.

25 K. Eickmeyer, M. Grohe, and M. Grüber. Approximation of natural w[p]-complete minimisation problems is hard. In *2008 23rd Annual IEEE Conference on Computational Complexity*, pages 8–18, June 2008. doi:10.1109/CCC.2008.24.

26 Emerson G. Escolar and Y. Hiraoka. Morse reduction for zigzag complexes. *Journal of the Indonesian Mathematical Society*, 20:47–75, 2014.

27 J. Flum and M. Grohe. *Parameterized complexity theory*. Texts in Theoretical Computer Science. An EATCS Series. Springer-Verlag, Berlin, 2006.

28 Jörg Flum and Martin Grohe. *Parameterized Complexity Theory*. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2006.

29 Robin Forman. Morse theory for cell complexes. *Advances in Mathematics*, 134(1):90–145, 1998. doi:10.1006/aima.1997.1650.

30 Robin Forman. A user's guide to discrete Morse theory. *Séminaire Lotharingien de Combinatoire*, B48c:1–35, 2002.

31 Masahiro Hachimori. *Combinatorics of constructible complexes*. PhD thesis, Tokyo University, 2000.

32 Shaun Harker, Konstantin Mischaikow, Marian Mrozek, and Vidit Nanda. Discrete Morse theoretic algorithms for computing homology of complexes and maps. *Foundations of Computational Mathematics*, 14(1):151–184, Feb 2014. doi:10.1007/s10208-013-9145-0.

33 Shawn Harker, Konstantin Mischaikow, Marian Mrozek, Vidit Nanda, Hubert Wagner, Mathias Juda, and Pavel Dlotko. The efficiency of a homology algorithm based on discrete Morse theory and coreductions. In *Proc. of 3rd Intl. Workshop on CTIC*, volume 1(1), 2010. doi:11441/26192.

34 Gregory Henselman and Robert Ghrist. Matroid filtrations and computational persistent homology. *arXiv preprint arXiv:1606.00199*, 2016.

35 Benjamin Holmgren, Bradley McCoy, Brittany Fasy, and David Millman. If you must choose among your children, pick the right one. *arXiv preprint arXiv:2103.13882*, 2021.

36 Xiaoling Hu, Yusu Wang, Li Fuxin, Dimitris Samaras, and Chao Chen. Topology-aware segmentation using discrete morse theory. In *International Conference on Learning Representations*, 2021. URL: https://openreview.net/forum?id=LGgdb4TS4Z.

37 Michael Jöllenbeck and Volkmar Welker. *Minimal Resolutions via Algebraic Discrete Morse Theory*, volume 923 of *Memoirs of the American Mathematical Society*. American Mathematical Society, 2009.

38 Michael Joswig, Frank H. Lutz, and Mimi Tsuruga. Heuristics for sphere recognition. In Hoon Hong and Chee Yap, editors, *ICMS 2014: Mathematical Software*, pages 152–159, Berlin, Heidelberg, 2014. Springer Berlin Heidelberg. doi:10.1007/978-3-662-44199-2_26.

39 Michael Joswig and Marc E. Pfetsch. Computing optimal Morse matchings. *SIAM Journal on Discrete Mathematics*, 20(1):11–25, 2006. doi:10.1137/S0895480104445885.

40 Matthew Kahle. Topology of random clique complexes. *Discrete Mathematics*, 309(6):1658–1671, 2009. doi:https://doi.org/10.1016/j.disc.2008.02.037.

41 Matthew Kahle. Random geometric complexes. *Discrete & Computational Geometry*, 45(3):553–573, 2011. doi:10.1007/s00454-010-9319-3.

42 Henry King, Kevin Knudson, and Neža Mramor. Generating discrete morse functions from point data. *Experimental Mathematics*, 14(4):435–444, 2005.

43 Kevin Knudson and Bei Wang. Discrete Stratified Morse Theory: A User's Guide. In *34th International Symposium on Computational Geometry (SoCG 2018)*, volume 99 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 54:1–54:14, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. doi:10.4230/LIPIcs.SoCG.2018.54.

44 Dmitry Kozlov. Discrete Morse Theory. In *Combinatorial Algebraic Topology*, volume 21 of *Algorithms and Computation in Mathematics*, chapter 11, pages 179–209. Springer Berlin Heidelberg, 2008. doi:10.1007/978-3-540-71962-5_11.

45 Dmitry N Kozlov. *Organized collapse: An introduction to discrete Morse theory*, volume 207. American Mathematical Society, 2021.

46 Leon Lampret and Aleš Vavpetič. (co)homology of lie algebras via algebraic morse theory. *Journal of Algebra*, 463:254–277, 2016. URL: https://www.sciencedirect.com/science/article/pii/S0021869316301235, doi:https://doi.org/10.1016/j.jalgebra.2016.04.036.

47 T. Lewiner, H. Lopes, and G. Tavares. Applications of forman's discrete morse theory to topology visualization and mesh compression. *IEEE Transactions on Visualization and Computer Graphics*, 10(5):499–508, Sept 2004. doi:10.1109/TVCG.2004.18.

48 Thomas Lewiner. Constructing discrete Morse functions. Master's thesis, Department of Mathematics, PUC-Rio, 2002. URL: http://thomas.lewiner.org/pdfs/tomlew_msc.pdf.

49 Thomas Lewiner, Hélio Lopes, and Geovan Tavares. Toward optimality in discrete Morse theory. *Experimental Mathematics*, 12(3):271–285, 2003. doi:10.1080/10586458.2003.10504498.

50 F. Lutz. Vertex-transitive triangulations. http://page.math.tu-berlin.de/~lutz/stellar/vertex-transitive-triangulations.html, 2011.

51 Rémy Malgouyres and Angel R. Francés. Determining whether a simplicial 3-complex collapses to a 1-complex is NP-complete. In *DGCI 2008: Discrete Geometry for Computer Imagery*, pages 177–188, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg. doi:10.1007/978-3-540-79126-3_17.

52 Clément Maria and Hannah Schreiber. Discrete morse theory for computing zigzag persistence. In Zachary Friggstad, Jörg-Rüdiger Sack, and Mohammad R Salavatipour, editors, *Algorithms and Data Structures*, pages 538–552, Cham, 2019. Springer International Publishing.

53 Dániel Marx. Completely inapproximable monotone and antimonotone parameterized problems. *J. Comput. Syst. Sci.*, 79(1):144–151, February 2013. doi:10.1016/j.jcss.2012.09.001.

54 R. Meshulam and N. Wallach. Homological connectivity of random k-dimensional complexes. *Random Struct. Algorithms*, 34(3):408–417, May 2009. doi:10.1002/rsa.v34:3.

55 Ezra Miller, Victor Reiner, and Bernd Sturmfels. *Geometric combinatorics*, volume 13 of *IAS/Park City Mathematics Series*. American Mathematical Society, 2007. doi:10.1090/pcms/013.

56 J. Milnor. *Morse theory*. Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, No. 51. Princeton University Press, Princeton, N.J., 1963.

57 Konstantin Mischaikow and Vidit Nanda. Morse theory for filtrations and efficient computation of persistent homology. *Discrete & Computational Geometry*, 50(2):330–353, 2013. doi:10.1007/s00454-013-9529-6.

58 Soham Mukherjee. Denoising with discrete morse theory. *The Visual Computer*, 2021. URL: https://doi.org/10.1007/s00371-021-02255-7, doi:10.1007/s00371-021-02255-7.

59 Christos H. Papadimitriou and Mihalis Yannakakis. Optimization, approximation, and complexity classes. *Journal of Computer and System Sciences*, 43(3):425 – 440, 1991. doi:10.1016/0022-0000(91)90023-X.

60 Abhishek Rathod, Talha Bin Masood, and Vijay Natarajan. Approximation algorithms for Max Morse Matching. *Computational Geometry*, 61:1 – 23, 2017. doi:10.1016/j.comgeo.2016.10.002.

61 Jan Reininghaus, David Günther, Ingrid Hotz, Steffen Prohaska, and Hans-Christian Hege. Tadd: A computational framework for data analysis using discrete morse theory. In *International Congress on Mathematical Software*, pages 198–208. Springer, 2010.

62 Vanessa Robins, Peter John Wood, and Adrian P. Sheppard. Theory and algorithms for constructing discrete morse complexes from grayscale digital images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(8):1646–1658, 2011. doi:10.1109/TPAMI.2011.95.

63 Tim Roughgarden, editor. *Beyond the Worst-Case Analysis of Algorithms*. Cambridge University Press, 2021. doi:10.1017/9781108637435.

64 Andrés Santamaría-Galvis and Russ Woodroofe. Shellings from relative shellings, with an application to np-completeness. *Discrete & Computational Geometry*, 2021. doi:10.1007/s00454-020-00273-1.

65 Sara Scaramuccia, Federico Iuricich, Leila De Floriani, and Claudia Landi. Computing multiparameter persistent homology through a discrete morse-based approach. *Computational Geometry*, 89:101623, 2020. URL: https://www.sciencedirect.com/science/article/pii/S0925772120300171, doi:https://doi.org/10.1016/j.comgeo.2020.101623.

66 Nicholas A Scoville. *Discrete Morse Theory*, volume 90. American Mathematical Soc., 2019.

67 Bart Selman, David G. Mitchell, and Hector J. Levesque. Generating hard satisfiability problems. *Artificial Intelligence*, 81(1):17–29, 1996. Frontiers in Problem Solving: Phase Transitions and Complexity.

68 N. Shivashankar, S. M, and V. Natarajan. Parallel computation of 2D Morse–Smale complexes. *IEEE Transactions on Visualization and Computer Graphics*, 18(10):1757–1770, Oct 2012. doi:10.1109/TVCG.2011.284.

69 Martin Tancer. Recognition of collapsible complexes is NP-complete. *Discrete Comput. Geom.*, 55(1):21–38, 2016. doi:10.1007/s00454-015-9747-1.

70 Csaba D Toth, Joseph O'Rourke, and Jacob E Goodman, editors. *Handbook of discrete and computational geometry*. CRC press, 2017.

71 David P. Williamson and David B. Shmoys. *The Design of Approximation Algorithms*. Cambridge University Press, 2011.

72 Ke Xu, Frédéric Boussemart, Fred Hemery, and Christophe Lecoutre. A simple model to generate hard satisfiable instances. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence*, IJCAI'05, pages 337–342, San Francisco, CA, USA, 2005. Morgan Kaufmann Publishers Inc.

73 E. C. Zeeman. On the dunce hat. *Topology*, 2:341–358, 1964. doi:10.1016/0040-9383(63)90014-4.

A Reducing MinMCS to MinrMM

A.1 Structural properties of the reduction

Note that Lemmas 55 and 56 appear as Lemmas 4.1 and 4.4 in [9], but with slightly different notation. For the sake of completeness, we restate the lemma with the notation introduced in this paper.

▶ **Lemma 55.** *For a circuit $C = (V(C), E(C))$, let F is a discrete Morse function on $K(C)$ with gradient $\mathcal{V}$:*

(i) If $s_i^\zeta \in \mathbf{D}_{m,\ell}^\zeta$ is eventually free in $K(C)$, then $\mathbf{D}_{m,\ell}^\zeta$ is erasable in $K(C)$.

(ii) Suppose that $\mathbf{D}_{m,\ell}^\zeta$ is erasable in $K(C)$ through a gradient $\mathcal{V}$,

- *If $m = 1$, then $\left(s_1^\zeta, \Gamma_1^\zeta\right)$ is a gradient pair in $\mathcal{V}$, and for any simplex $\sigma^\zeta \in \mathbf{D}_{1,\ell}^\zeta$ such that $\sigma^\zeta \notin \{s_1^\zeta, \Gamma_1^\zeta\}$ we have $F(s_1^\zeta) > F(\sigma^\zeta)$.*
- *If $m = 2$, then $\left(s_1^\zeta, \Gamma_1^\zeta\right) \in \mathcal{V}$ or $\left(s_2^\zeta, \Gamma_2^\zeta\right) \in \mathcal{V}$, and for any simplex $\sigma^\zeta \in \mathbf{D}_{m,\ell}^\zeta$ such that $\sigma^\zeta \notin \{s_i^\zeta, \Gamma_i^\zeta\}$ for $i = \{1, 2\}$, then we have, $\max(F(s_1^\zeta), F(s_2^\zeta)) > F(\sigma^\zeta)$.*

Proof. Suppose s_i^ζ is eventually free in $K(C)$. Then there exists a subcomplex L of $K(C)$ such that $K(C) \searrow L$ and s_i^ζ is free in L. Note that, by construction of $\mathbf{D}_{m,\ell}$, this implies that $\mathbf{D}_{m,\ell}^\zeta$ is a subcomplex of L. Now using the gradient specified in Figure 2 all the 2-simplices of $\mathbf{D}_{m,\ell}^\zeta$ can be collapsed, making $\mathbf{D}_{m,\ell}^\zeta$ erasable in $K(C)$. This proves the first statement of the lemma. The last two statements of the lemma immediately follows from observing that the s-edges are the only free edges in complex $\mathbf{D}_{m,\ell}^\zeta$, the simplices $\{\Gamma_i^\zeta\}$ are the unique cofaces incident on edges $\{s_i^\zeta\}$ respectively, and $\mathbf{D}_{m,\ell}^\zeta$ is erasable in $K(C)$ through the gradient $\mathcal{V}$ of F. ◀

▶ **Lemma 56.** *For any input gate G_i, the subcomplex $\mathbf{D}^{(i,1)} \setminus \{\Gamma_1^{(i,1)}\}$ is erasable in $K(C)$.*

Proof. Under the reindexing scheme described in Remark 20, let ζ_1 be such that $\mathbf{D}^{(i,1)} = \mathbf{D}^{\zeta_1}$. Consider the discrete gradient specified in Figure 2 (a) as a gradient $\mathcal{V}^{(i,1)}$ on $\mathbf{D}^{(i,1)} \subseteq K(C)$. First note that $\mathbf{D}^{(i,1)} \setminus \{\Gamma^{(i,1)}\}$ is erasable in $\mathbf{D}^{(i,1)}$ through the gradient $\mathcal{V}^{(i,1)} \setminus \{(s_1^{(i,1)}, \Gamma^{(i,1)})\}$. Moreover, all 1-simplices of $\mathbf{D}^{(i,1)}$ that are paired in $\mathcal{V}^{(i,1)}$ with a 2-simplex do not appear in $\mathbf{D}^{\zeta_2}$ for any edge $\zeta_1 \neq \zeta_2$. It follows that $\mathbf{D}^{(i,1)} \setminus \{\Gamma_1^{(i,1)}\}$ is erasable in $K(C)$. ◀

A.2 Reducing MinMCS to MinrMM: Backward direction

We intend to establish an L-reduction from MinMCS to MinrMM. To this end, in Section 4.2 and Section 4.3 we described the map $K : C \mapsto K(C)$ that transforms instances of MinMCS (monotone circuits) to instances of MinrMM (simplicial complexes). In this section, we seek to construct a map $\mathcal{I}$ that transforms solutions of MinrMM (discrete gradients $\mathcal{V}$ on $K(C)$) to solutions of MinMCS (satisfying input assignments $\mathcal{I}(C, \mathcal{V})$ of circuit C). Recall that $m_{\mathsf{MinrMM}}(K(C), \mathcal{V})$ denotes the objective value of some solution $\mathcal{V}$ on $K(C)$ for MinrMM, whereas $m_{\mathsf{MinMCS}}(C, \mathcal{I}(C, \mathcal{V}))$ denotes the objective value of a solution $\mathcal{I}(C, \mathcal{V})$ on C for MinMCS.

Suppose that we are given a circuit $C = (V(C), E(C))$ with $n = |V(C)|$ number of nodes. Also, for a vector field $\tilde{\mathcal{V}}$ on $K(C)$, we denote the critical simplices of dimension $2, 1$ and 0 by $m_2(\tilde{\mathcal{V}}), m_1(\tilde{\mathcal{V}})$ and $m_0(\tilde{\mathcal{V}})$ respectively. Then, by definition,

$$m_{\mathsf{MinrMM}}(K(C), \tilde{\mathcal{V}}) = m_2(\tilde{\mathcal{V}}) + m_1(\tilde{\mathcal{V}}) + m_0(\tilde{\mathcal{V}}) - 1. \tag{7}$$

In Section 4.4, we designed a gradient vector field $\mathcal{V}$ on $K(C)$ with $m_2(\mathcal{V}) = m, m_1(\mathcal{V}) = m$ and $m_0(\mathcal{V}) = 1$, for some $m \leq n$. We have from [29, Theorem 1.7],

$$m_0(\tilde{\mathcal{V}}) - m_1(\tilde{\mathcal{V}}) + m_2(\tilde{\mathcal{V}}) = m_0(\mathcal{V}) - m_1(\mathcal{V}) + m_2(\mathcal{V}).$$

which gives $m_0(\tilde{\mathcal{V}}) - m_1(\tilde{\mathcal{V}}) + m_2(\tilde{\mathcal{V}}) = 1$. Since $m_0(\tilde{\mathcal{V}}) \geq 1$, this gives, for any vector field $\tilde{\mathcal{V}}$ on $K(C)$, the following inequality

$$m_2(\tilde{\mathcal{V}}) \leq m_1(\tilde{\mathcal{V}}). \tag{8}$$

In particular, from Equation (7) and Equation (8), we obtain

$$m_{\mathsf{MinrMM}}(K(C), \tilde{\mathcal{V}}) \geq 2m_2(\tilde{\mathcal{V}}). \tag{9}$$

Now, if $m_2(\tilde{\mathcal{V}}) \geq n$, we set $\mathcal{I}(C, \tilde{\mathcal{V}})$ to be the set of all input gates of C. Clearly, this gives a satisfying assignment and using Equation (9) also satisfies

$$m_{\mathsf{MinrMM}}(K(C), \tilde{\mathcal{V}}) \geq 2 \cdot m_{\mathsf{MinMCS}}(C, \mathcal{I}(C, \tilde{\mathcal{V}})).$$

So, for the remainder of this section, we assume that $m_2(\tilde{\mathcal{V}}) < n$. In particular, for any non-output gate G_i with n blocks, at most $n-1$ of them may have critical 2-simplices.

▶ **Definition 57** (2-paired edges). *Given a vector field $\mathcal{V}$ on a 2-complex K, we say that an edge $e \in K$ is* 2-paired *in $\mathcal{V}$ if it is paired to a 2-simplex in $\mathcal{V}$.*

▶ **Definition 58** (properly satisfied gates). *Suppose that we are given a circuit C, and a vector field $\tilde{\mathcal{V}}$ on the associated complex $K(C)$. Then,*

1. *an ordinary gate G_q is said to be* properly satisfied *if there exists a $j \in [1, n]$ such that*
   - *for an or-gate G_q at least of the two edges $s_{\iota_1}^{(q,j)}$, $s_{\iota_2}^{(q,j)}$ is 2-paired (in $\tilde{\mathcal{V}}$), or*
   - *for an and-gate G_q both the edges $s_{\iota_1}^{(q,j)}$, $s_{\iota_2}^{(q,j)}$ are 2-paired (in $\tilde{\mathcal{V}}$),*
   - *in both cases, the j-th block has no critical 2-simplices;*
2. *an input gate G_i is said to be* properly satisfied *if the dunce hat associated to it contains at least one critical 2-simplex,*
3. *the output gate G_o is said to be* properly satisfied *if G_o is an or-gate and at least one of the two inputs gates of G_o is properly satisfied, or if G_o is an and-gate and both input gates of G_o are properly satisfied.*

▶ **Lemma 59.** *Suppose that G_k is a non-output gate that is properly satisfied. Then,*

1. *if G_k is an or-gate, then at least one of the two gates that serve as inputs to G_k is also properly satisfied.*
2. *if G_k is an and-gate, then both gates that serve as inputs to G_k are also properly satisfied.*

Proof. Assume without loss of generality that G_k is an and-gate, and the two inputs that go into G_k, namely G_ℓ and G_j are both input gates. Since G_k is properly satisfied, there exists $p \in [1, n]$ such that $s_{\iota_1}^{(k,p)}$, $s_{\iota_2}^{(k,p)}$ are 2-paired and the p-th block of G_k has no critical 2-simplices. Now suppose that either G_ℓ or G_j is not properly satisfied. For the sake of argument, suppose that G_ℓ is not properly satisfied. That is, $\mathbf{D}^{(\ell,1)}$ has no critical 2-simplices and $s_f^{(\ell,1)}$ is 2-paired. Note that $s_{\iota_1}^{(k,p)}$ is identified to a t-edge in $\mathbf{D}^{(\ell,1)}$. Using Lemma 55, we obtain $\tilde{F}(s_f^{(\ell,1)}) > \tilde{F}(s_{\iota_1}^{(k,p)})$. Since ${}^3s_1^{(k,p)}$ occurs as a t-edge in ${}^1\mathbf{D}^{(k,p)}$, using Lemma 55, we obtain $\tilde{F}(s_{\iota_1}^{(k,p)}) > \tilde{F}({}^3s_1^{(k,p)})$. Combining the two inequalities we obtain

$$\tilde{F}(s_f^{(\ell,1)}) > \tilde{F}({}^3s_1^{(k,p)}) \tag{10}$$

Moreover, $s_f^{(\ell,1)}$ is identified to a t-edge in ${}^3\mathbf{D}^{(k,p)}$, and by assumption ${}^3\mathbf{D}^{(k,p)}$ has no critical 2-simplices and, hence ${}^3s_1^{(k,p)}$ is 2-paired. Therefore, once again, using Lemma 55, we obtain

$$\tilde{F}({}^3s_1^{(k,p)}) > \tilde{F}(s_f^{(\ell,1)}) \tag{11}$$

Since Equation (10) and Equation (11) combine to give a contradiction, we conclude that G_ℓ is properly satisfied. All combinations of G_k as an {and-gate, or-gate}, and G_ℓ and G_j as {or-gates, input gates, and-gates} give similar contradictions, proving the claim. ◀

▶ **Lemma 60.** *Given a Morse function $\tilde{F}$ on $K(C)$ with vector field $\tilde{\mathcal{V}}$, the output gate G_o is properly satisfied.*

Proof. Assume without loss of generality that G_o is an or-gate, and the two inputs to G_o, namely G_ℓ and G_j are non-input and-gates. Let $k \in [1, n]$ be such that the k-th copy of G_o has no critical 2-simplices. Such a k exists because by assumption we have less than n critical simplices. Now, suppose that neither G_ℓ nor G_j is properly satisfied.

Since, G_ℓ is not properly satisfied there exists a $p \in [1, n]$ such that either $s_{\iota_1}^{(\ell,p)}$ or $s_{\iota_2}^{(\ell,p)}$ is not 2-paired and p-th block has no critical 2-simplices (because by assumption we have less than n critical simplces). Assume without loss of generality that $s_{\iota_1}^{(\ell,p)}$ is not 2-paired. Then, $s_{f_1}^{(\ell,p)}$ is 2-paired. Using Lemma 55, we obtain $\tilde{F}(s_{f_1}^{(\ell,p)}) > \tilde{F}({}^3s_1^{(\ell,p)})$. Now, $s_{\iota_1}^{(o,k)}$ is identified to a t-edge in ${}^3\mathbf{D}^{(\ell,p)}$. So, using Lemma 55, we obtain $\tilde{F}({}^3s_1^{(\ell,p)}) > \tilde{F}(s_{\iota_1}^{(o,k)})$. Combining the two inequalities, we obtain,

$$\tilde{F}(s_{f_1}^{(\ell,p)}) > \tilde{F}(s_{\iota_1}^{(o,k)}). \tag{12}$$

Similarly, there exists a $q \in [1, n]$ such that either $s_{\iota_1}^{(j,q)}$ or $s_{\iota_2}^{(j,q)}$ is not 2-paired. Assume without loss of generality that $s_{\iota_1}^{(j,q)}$ is not 2-paired. Hence, we can show that

$$\tilde{F}(s_{f_1}^{(j,q)}) > \tilde{F}(s_{\iota_2}^{(o,k)}). \tag{13}$$

Combining Equation (12) and Equation (13), we obtain:

$$\max(\tilde{F}(s_{f_1}^{(\ell,p)}), \tilde{F}(s_{f_1}^{(j,q)})) > \max(\tilde{F}(s_{\iota_1}^{(o,k)}), \tilde{F}(s_{\iota_2}^{(o,k)})). \tag{14}$$

But, $s_{f_1}^{(\ell,p)}$ and $s_{f_1}^{(j,q)}$ appear as t-edges in $\mathbf{D}^{(o,k)}$. So, once again, using Lemma 55,

$$\max(\tilde{F}(s_{\iota_1}^{(o,k)}), \tilde{F}(s_{\iota_2}^{(o,k)})) > \tilde{F}(s_{f_1}^{(\ell,p)}) \text{ and } \max(\tilde{F}(s_{\iota_1}^{(o,k)}), \tilde{F}(s_{\iota_2}^{(o,k)})) > \tilde{F}(s_{f_1}^{(j,q)}),$$

which combine to give:

$$\max(\tilde{F}(s_{\iota_1}^{(o,k)}), \tilde{F}(s_{\iota_2}^{(o,k)})) > \max(\tilde{F}(s_{f_1}^{(\ell,p)}), \tilde{F}(s_{f_1}^{(j,q)})) \tag{15}$$

Since Equation (14) and Equation (15) combine to give a contradiction, we conclude that G_o is properly satisfied, and at least one of the two gates G_ℓ and G_j are also properly satisfied. All combinations of G_o as an {and-gate, or-gate}, and G_ℓ and G_j as {or-gates, input gates, and-gates} give similar contradictions, proving the claim. Moreover, if G_o is a properly satisfied and-gate, then both G_ℓ and G_j are also be properly satisfied. ◀

Now, we construct the map $\mathcal{I}(C, \tilde{\mathcal{V}})$ as follows: For every input gate G_ℓ whose associated dunce hat is properly satisfied, we set $\mathcal{I}(C, \tilde{\mathcal{V}})(G_\ell) = 1$. That is, our assignment $\mathcal{I}(C, \tilde{\mathcal{V}})(\cdot)$ ensures that an input gate is satisfied if and only if it is properly satisfied.

▷ Claim 61. With input assignment $\mathcal{I}(C, \tilde{\mathcal{V}})(\cdot)$, the circuit C is satisfied.

Proof. We prove the following claim inductively: Every gate of C that is properly satisfied is also satisfied.

To begin with, let $\prec_C$ be some total order on $V(C)$ consistent with the partial order imposed by C. Assume that the gates in C are indexed from 1 to $|C|$ so that

$$\text{for all } G_i, G_j \in C, \quad i < j \Leftrightarrow G_i \prec_C G_j.$$

Let $\mathcal{P}$ denotes the set of properly satisfied gates. Let $i_1, i_2, \ldots i_{|\mathcal{P}|}$ denote the indices of the properly satisfied gates, where $i_k > i_{k-1}$ for all k. By repeated application of Lemma 59, it follows that G_{i_1} is an input gate. Then, from our construction of $\mathcal{I}(C, \tilde{\mathcal{V}})(\cdot)$, we can conclude that G_{i_1} is also satisfied, giving us the base case.

Now, we make the inductive hypothesis that the gates $G_{i_1} \ldots G_{i_{k-1}}$ are satisfied. Suppose that G_{i_k} is an or-gate. Then, by Lemma 59, one of the inputs to G_{i_k}, say G_{i_j} is properly satisfied. As a consequence of our indexing we have $j \in [1, k-1]$, and owing to the inductive hypothesis, G_{i_j} is satisfied. But, since G_{i_k} is an or-gate, this implies that G_{i_k} is also satisfied. Suppose that G_{i_k} be an and-gate. Then, by Lemma 59, both the inputs to G_{i_k}, say G_{i_j}, G_{i_p} are properly satisfied. As a consequence of our indexing we have $j, p \in [1, k-1]$, and owing to the inductive hypothesis, G_{i_j}, G_{i_p} are satisfied. But, since G_{i_k} is an and-gate, this implies that G_{i_k} is also satisfied, completing the induction.

Finally, using Lemma 60, the output gate is properly satisfied, and by the argument above it is also satisified. ◀

An immediate consequence of Claim 61 is the following:

$$m_2(\tilde{\mathcal{V}}) \geq m_{\mathsf{MinMCS}}(C, \mathcal{I}(C, \tilde{\mathcal{V}})) \tag{16}$$

▶ **Proposition 62.** $m_{\mathsf{MinrMM}}(K(C), \tilde{\mathcal{V}}) \geq 2 \cdot m_{\mathsf{MinMCS}}(C, \mathcal{I}(C, \tilde{\mathcal{V}}))$

Proof. This follows immmediately by combining Equation (9) and Equation (16). ◀

Now, if the gradient vector field $\tilde{\mathcal{V}}$ is, in fact optimal for $K(C)$, then $\tilde{\mathcal{V}}$ has a single critical 0-simplex. That is, $m_0(\tilde{\mathcal{V}}) = 1$ Recall that in Section 4.4, we designed a gradient vector field $\mathcal{V}$ on $K(C)$ with $m_2(\mathcal{V}) = m, m_1(\mathcal{V}) = m$ and $m_0(\mathcal{V}) = 1$, for some $m \leq n$. From [29, Theorem 1.7], we have

$$m_0(\tilde{\mathcal{V}}) - m_1(\tilde{\mathcal{V}}) + m_2(\tilde{\mathcal{V}}) = m_0(\mathcal{V}) - m_1(\mathcal{V}) + m_2(\mathcal{V}).$$

which gives us

$$-m_1(\tilde{\mathcal{V}}) + m_2(\tilde{\mathcal{V}}) = 0 \tag{17}$$

From Equation (17), we conclude that $\mathrm{OPT}_{\mathsf{MinrMM}}(K(C)) = 2m_2(\tilde{\mathcal{V}})$.

By Equation (16), we have

$$m_2(\tilde{\mathcal{V}}) \geq m_{\mathsf{MinMCS}}(C, \mathcal{I}(C, \tilde{\mathcal{V}})).$$

Since by definition

$$m_{\mathsf{MinMCS}}(C, \mathcal{I}(C, \tilde{\mathcal{V}})) \geq \mathrm{OPT}_{\mathsf{MinMCS}}(C),$$

we have the following proposition

▶ **Proposition 63.** $\mathrm{OPT}_{\mathsf{MinrMM}}(K(C)) \geq 2\mathrm{OPT}_{\mathsf{MinMCS}}(C)$.

Combining Proposition 31 and Proposition 63, we obtain the following proposition.

▶ **Proposition 64.** $\mathrm{OPT}_{\mathsf{MinrMM}}(K(C)) = 2\mathrm{OPT}_{\mathsf{MinMCS}}(C)$.

B Morse matchings for Linial–Meshulam complexes

For Linial–Meshulam complexes $\mathsf{Y}_d(n,p)$,

- when $np \to \infty$, let $\mathcal{U}$ be the apparent pairs gradient on $\mathsf{Y}_d(n,p)$.
- when $np \to 0$, $\mathcal{U}$ is comprised of
  - the apparent pairs gradient for matching $k-1$-simplices to k-simplices for $k \in [d-1]$
  - $\mathcal{V}_2$ for matching the remaining $(d-1)$-simplices to d-simplices.

The apparent pairs gradient, and the gradient $\mathcal{V}_2$ are as described in Section 8.

Let V be the vertex set of $\mathsf{Y}_d(n,p)$, and v' be the lexicographically lowest vertex of V.

For each $r \in [0,d]$, let c_r denote the total number of r-dimensional simplices in $\mathsf{Y}_d(n,p)$. Let m_r denote the total number of critical r-simplices of $\mathcal{U}$ and $\overline{m_r}$ be the total number of regular simplices of $\mathcal{U}$. Also, let n_r and $\overline{n_r}$ denote the total number of critical r-simplices and regular r-simplices respectively of the optimal discrete gradient on $\mathsf{Y}_d(n,p)$.

▶ **Lemma 65.** *All the k-simplices of $\mathsf{Y}_d(n,p)$ for $k \in [0,d-2]$ are matched by $\mathcal{U}$. In particular, $\overline{n_k} = \overline{m_k} = c_k$ for $k \in [d-2]$, and $\overline{n_0} = \overline{m_0} = V - 1$.*

Proof. Let σ be a k-simplex, where $k \in [d-2]$. If $v' \in \sigma$, then $(\sigma \setminus \{v\}, \sigma) \in \mathcal{U}$, whereas if $v' \notin \sigma$, then $(\sigma, \sigma \cup \{v'\}) \in \mathcal{U}$. That is for $k \in [0,d-2]$, $n_k = m_k = 0$, and $\overline{n_k} = \overline{m_k} = c_k$. On the other hand, any vertex $v \neq v'$ is matched to the edge $\{v, v'\}$. ◀

Note that

$$\frac{\mathbb{E}(|\mathcal{U}|)}{\mathbb{E}(|\mathsf{opt}|)} = \frac{\mathbb{E}(2|\mathcal{U}|)}{\mathbb{E}(2|\mathsf{opt}|)} = \frac{\mathbb{E}(\sum_{k=0}^{d} \overline{m_k})}{\mathbb{E}(\sum_{k=0}^{d} \overline{n_k})}$$

Case 1: $np \to \infty$

By definition,

$$\overline{m_d} \geq c_{d-1} - m_{d-1} - c_{d-2} \tag{18}$$

Also since the complex is d-dimensional, we get

$$\overline{n_d} \leq c_{d-1}. \tag{19}$$

Using Equations (18) and (19) and Lemma 65, we obtain

$$\mathbb{E}(\sum_{k=0}^{d} \overline{n_k}) \leq \mathbb{E}(\sum_{k=0}^{d-1} c_k + c_{d-1}) = \sum_{k=0}^{d-1} c_k + c_{d-1}.$$

$$\begin{aligned}\mathbb{E}(\sum_{k=0}^{d} \overline{m_k}) &\geq \mathbb{E}(\sum_{k=0}^{d-2} c_k + c_{d-1} - m_{d-1} + c_{d-1} - m_{d-1} - c_{d-2}) \\ &= \sum_{k=0}^{d-1} c_k + c_{d-1} - c_{d-2} - 2\mathbb{E}(m_{d-1}).\end{aligned}$$

Therefore,

$$\begin{aligned}\frac{\mathbb{E}(|\mathcal{U}|)}{\mathbb{E}(|\mathsf{opt}|)} &= \frac{\mathbb{E}(\sum_{k=0}^{d}\overline{m_k})}{\mathbb{E}(\sum_{k=0}^{d}\overline{n_k})}\\ &\geq \frac{\sum_{k=0}^{d-1}c_k + c_{d-1} - c_{d-2} - 2\mathbb{E}(m_{d-1})}{\sum_{k=0}^{d-1}c_k + c_{d-1}}\\ &= 1 + \frac{-c_{d-2} - 2\mathbb{E}(m_{d-1})}{\sum_{k=0}^{d-1}c_k + c_{d-1}}.\end{aligned}$$

Using Corollary 52, and the fact that in $\mathsf{Y}_d(n,p)$, $\frac{c_j}{c_k} \to 0$ for $j < k$ and $j,k \in [0,d-1]$, we conclude that

$$\frac{\mathbb{E}(|\mathcal{U}|)}{\mathbb{E}(|\mathsf{opt}|)} \to 1.$$

Case 2: $np \to 0$

Since a regular $(d-1)$-simplex is paired to either a d-simplex or a $(d-2)$-simplex, we obtain

$$\overline{n_{d-1}} \leq c_d + c_{d-2}, \tag{20}$$

$$\overline{m_{d-1}} \geq c_d - m_d + c_{d-2} - c_{d-3}. \tag{21}$$

Therefore, using Equations (20) and (21) and Lemma 65,

$$\begin{aligned}\frac{\mathbb{E}(|\mathcal{U}|)}{\mathbb{E}(|\mathsf{opt}|)} &= \frac{\mathbb{E}(\sum_{k=0}^{d}\overline{m_k})}{\mathbb{E}(\sum_{k=0}^{d}\overline{n_k})}\\ &\geq \frac{\sum_{k=0}^{d-2}c_k + (\mathbb{E}(c_d) - \mathbb{E}(m_d) + c_{d-2} - c_{d-3}) + (\mathbb{E}(c_d) - \mathbb{E}(m_d))}{\sum_{k=0}^{d-2}c_k + \mathbb{E}(c_d) + c_{d-2} + \mathbb{E}(c_d)}\\ &= 1 + \frac{-c_{d-3} - 2\mathbb{E}(m_d)}{\sum_{k=0}^{d-2}c_k + c_{d-2} + 2\mathbb{E}(c_d)}.\end{aligned}$$

Using Corollary 52, and the fact that in $\mathsf{Y}_d(n,p)$, $\frac{c_j}{c_k} \to 0$ for $j < k$ and $j,k \in [0,d-1]$, we conclude that

$$\frac{\mathbb{E}(|\mathcal{U}|)}{\mathbb{E}(|\mathsf{opt}|)} \to 1.$$

E

PAPER V: THE COMPLEXITY OF HIGH-DIMENSIONAL CUTS

Generalizing the notion of cuts from graphs to simplicial complexes, we study two problems.

- Boundary nontrivialization which can be described as follows: Given a bounding $\mathbb{Z}_2$ r-cycle ζ in a simplicial complex K, find a set S of $(r+1)$-dimensional simplices of minimum cardinality so that the removal of S from K makes ζ nontrivial.

- Topological Hitting Set which can be described as follows: Given a nontrivial $\mathbb{Z}_2$ r-cycle ζ in a simplicial complex K, find a set S of r-dimensional simplices of minimum cardinality so that S meets every cycle homologous to ζ.

We also study the global variants of the problem that we call Global Boundary nontrivialization and Global Topological Hitting Set, respectively.

The first four sections of the paper consist of a well-motivated introduction, followed by requisite background on homology with $\mathbb{Z}_2$ coefficients and parameterized complexity.

Surfaces. Our first result, explained in Section 5, is the following: Topological Hitting Set admits a polynomial-time algorithm on triangulations of closed surfaces. At the heart of our proof lies an appealing characterization of the optimal solutions in terms of the cocycles of the surface, which is of independent interest. Specifically, we show that a minimal solution set is necessarily a nontrivial cocycle. Further, we show that the following are equivalent: 1. A connected cocycle η is a feasible set for the input cycle ζ. 2. Every cycle in $[\zeta]$ intersects a connected cocycle η in an odd number of edges. 3. One of the cycles in $[\zeta]$ intersects a connected cocycle η in an odd number of edges. In particular, this allows us to identify the nontrivial cocycles that are solutions based on a parity-based property. Having this characterization at hand, we proceed to characterize cohomology classes that are solutions. Eventually, we arrive at a very simple 3-step algorithm for Topological Hitting Set on surfaces.

W[1]-hardness and NP-hardness. For general complexes, in 6.1, we show that Topological Hitting Set is **W[1]**-hard with respect to the solution size k as the parameter, (and hence, it is also **NP**-hard). The proof is based on a reduction from the k-Multicolored Clique problem. Here, the reduction shows the essence of hardness: its description is short, but its proof exposes various "behaviors" that we find interesting. In particular, the forward direction requires

a nontrivial parity based argument, while the reverse direction shows how to "trace" a solution through the complex.

In addition, in 6.2, we show that BOUNDARY NONTRIVIALIZATION is also **W[1]**-hard with respect to the solution size k as a parameter. The principles of this reduction follow the lines of the reduction for TOPOLOGICAL HITTING SET, though, here, both the description and the proof of the reduction are more involved because of subdivisions that help avoid some unhelpful incidences.

Fixed-parameter tractability. On the positive side, in Section 7.1, we show that TOPOLOGICAL HITTING SET admits an FPT algorithm with respect to $k + \Delta$, where Δ is the maximum degree of the Hasse graph of the complex K. Here, the main insight is that a minimal solution must be connected. Having this insight at hand, the algorithm follows: If we search across the geodesic ball of every r-simplex in the complex K, we will find a solution.

In contrast, we observe that BOUNDARY NONTRIVIALIZATION does not admit this property because minimal solutions can be disconnected. This motivates the search of another parameter that makes the problem tractable. Exploiting the set-cover like structure of the problem, in Section 7.2, we show that BOUNDARY NONTRIVIALIZATION with bounding r-cycles as input has an $O(\log n)$-approximation FPT algorithm with β_{r+1} (the Betti number) as the parameter, when the input complex K is $(r+1)$-dimensional. It is worth noting that BOUNDARY NONTRIVIALIZATION is **W[1]**-hard even for $(r+1)$-dimensional complexes with solution size as the parameter since the hardness gadget used in 7.2 is $(r+1)$-dimensional.

Finally, by exploiting the vector space structure of the homology groups and the boundary groups, in Sections 7.1 and 7.2, we also provide a randomized FPT algorithm for GLOBAL TOPOLOGICAL HITTING SET and a randomized FPT approximation algorithm for GLOBAL BOUNDARY NONTRIVIALIZATION respectively.

STATEMENT OF INDIVIDUAL CONTRIBUTION. I had been reading some literature of high-dimensional expansion, and I wanted to understand these questions from a complexity standpoint. So I suggested the questions addressed in this paper as a research project to my two collaborators. In this collaboration, I was the principal author and I was significantly involved in finding the ideas and carrying out the scientific work of all parts of this article. I was in charge of writing down this article. After the writeup was ready, I made several expository and corrective changes based on the feedback I received from my collaborators.

The complexity of high-dimensional cuts

Ulrich Bauer
Department of Mathematics, Technical University of Munich (TUM)
Boltzmannstr. 3, 85748 Garching b. München, Germany

Abhishek Rathod
Department of Mathematics, Technical University of Munich (TUM)
Boltzmannstr. 3, 85748 Garching b. München, Germany

Meirav Zehavi
Ben-Gurion University of the Negev, Beer-Sheva, Israel

Abstract

Cut problems form one of the most fundamental classes of problems in algorithmic graph theory. For instance, the minimum cut, the minimum s-t cut, the minimum multiway cut, and the minimum k-way cut are some of the commonly encountered cut problems. Many of these problems have been extensively studied over several decades. In this paper, we initiate the algorithmic study of some cut problems in high dimensions.

The first problem we study, namely, Topological Hitting Set (THS), is defined as follows: Given a nontrivial r-cycle ζ in a simplicial complex K, find a set $\mathcal{S}$ of r-dimensional simplices of minimum cardinality so that $\mathcal{S}$ meets every cycle homologous to ζ. Our main result is that this problem admits a polynomial time solution on triangulations of closed surfaces. Interestingly, the optimal solution is given in terms of the cocycles of the surface. For general complexes, we show that THS is **W[1]**-hard with respect to the solution size k. On the positive side, we show that THS admits an **FPT** algorithm with respect to $k + d$, where d is the maximum degree of the Hasse graph of the complex K.

We also define a problem called Boundary Nontrivialization (BNT): Given a bounding r-cycle ζ in a simplicial complex K, find a set $\mathcal{S}$ of $(r+1)$-dimensional simplices of minimum cardinality so that the removal of $\mathcal{S}$ from K makes ζ non-bounding. We show that BNT is **W[1]**-hard with respect to the solution size as the parameter, and has an $O(\log n)$-approximation **FPT** algorithm for $(r + 1)$-dimensional complexes with the $(r + 1)$-th Betti number β_{r+1} as the parameter. Finally, we provide randomized (approximation) **FPT** algorithms for the global variants of THS and BNT.

2012 ACM Subject Classification Theory of computation → Computational geometry; Mathematics of computing → Algebraic topology

Keywords and phrases Algorithmic topology, Cut problems, Surfaces, Parameterized complexity, FPT algorithms

1 Introduction

A *graph cut* is a partition of the vertices of a graph into two disjoint subsets. The set of edges that have one vertex lying in each of the two subsets determines a so-called *cut-set*. Typically, the objective function to optimize involves the size of the cut-set. Graph cuts have a ubiquitous presence in theoretical computer science. Cuts are also related to the spectra of the adjacency matrix of the graph leading to a beautiful mathematical theory [14]. Cuts have also found many real-world applications in clustering, shape matching, image segmentation and smoothing, and energy minimization problems in computer vision.

Cut problems are related to flow problems in graphs due to the duality between cuts and flows. In fact, the max-flow min-cut theorem tells us that the maximum value of flow between a vertex s and and vertex t equals the value of the minimum cut that separates s and t. Figure 1 shows an example of an s-t cut on an undirected graph.

Incidentally, graphs happen to be 1-dimensional simplicial complexes. And some of the

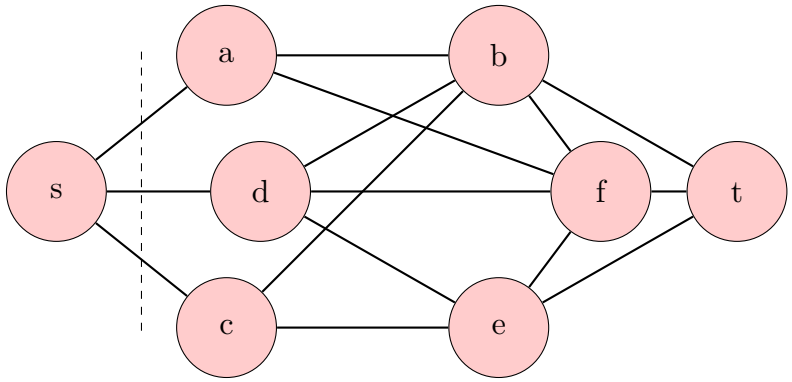


Figure 1 The dashed vertical line shows a minimum s-t cut in the graph.

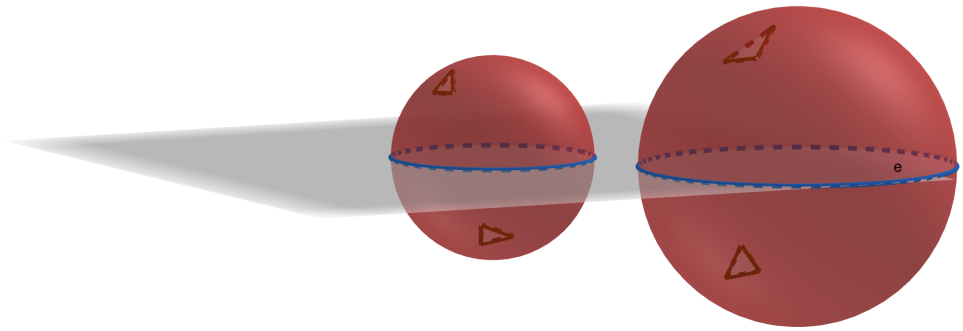


Figure 2 The complex L_1 consists of two disjoint triangulated spheres. We do not show the entire triangulation, only the four triangles of interest. The boundary of interest is the equator of the larger sphere on the right.

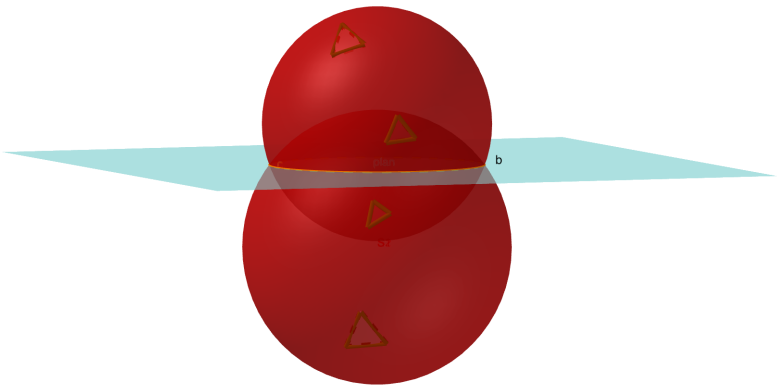


Figure 3 The complex L_2 consists of a triangulation of union of two spheres that intersect in a circle. As before, we do not show the entire triangulation, only the four triangles of interest. The boundary of interest is the circle of intersection of the two spheres.

cut problems have a natural homological interpretation. For instance, consider the following problem: What is the minimum number of edges you need to remove from a graph so that the vertices $\{s, t\}$ do not form a bounding 0-cycle of a 1-chain over $\mathbb{Z}_2$ in the resulting graph? Since we have an s-t cut if and only if there are no paths connecting s and t, it is easy to check that this problem is equivalent to finding the minimum s-t cut on graphs! It is natural to ask the analogous question for complexes of higher dimension. In particular, the question we ask, namely Boundary Nontrivialization, is the following one: Given a bounding $\mathbb{Z}_2$ r-cycle ζ in a simplicial complex K, find a set $\mathcal{S}$ of $(r+1)$-dimensional simplices of minimum cardinality so that the removal of $\mathcal{S}$ from K makes ζ nontrivial.

For instance, consider the two complexes L_1 and L_2 shown in Figures 2 and 3, respectively. For complex L_1 shown in Figure 2, let the equator e of the sphere on the right be the bounding 1-cycle that we want to make nontrivial. Both hemispheres are bounded by the equator. So, the two highlighted triangles from the right sphere of the complex L_1 constitute the optimal solution for Boundary Nontrivialization. That is, removing these two triangles makes e a nontrivial 1-cycle. For complex L_2 shown in Figure 3, the circle of intersection of the two spheres is the bounding 1-cycle of interest denoted by b. Removing all the four highlighted triangles from complex L_2 makes b a nontrivial 1-cycle. This also happens to be the optimal solution for making b nontrivial.

Complementary to the question of removing the minimal number of $r+1$-simplices in order to make a bounding cycle nontrivial, is the problem of removing the minimum number of r-simplices from a complex so that an entire homology class is destroyed. More formally, the problem Topological Hitting Set can be described as follows: given a nontrivial $\mathbb{Z}_2$ r-cycle ζ in a simplicial complex K, find a set $\mathcal{S}$ of r-dimensional simplices of minimum cardinality so that $\mathcal{S}$ meets every cycle homologous to ζ.

Topological Hitting Set on graphs can be described as follows: Suppose we are given a graph G with k components. Let C be one of the components of G. Then, $\beta_0(G) = |k|$, and each component determines a 0-cycle. So the question of Topological Hitting Set is to determine the minimum number of vertices you need to remove so that C is not a component anymore. The answer is trivial! One needs to remove all the vertices in C. For example in Figure 4, C_2 ceases to be a component if and only if all four vertices in C_2 are removed. It is worth noting that it is the unidimensionality of graphs that makes the problem trivial. What is more, even the 'cut' aspect of the problem is not immediately visible for graphs.

In contrast, for higher-dimensional complexes, the problem has a distinct cut flavor. For instance, consider the planar complex shown in Figure 5. The minimum number of edges that need to be removed so that every cycle homologous to ζ is destroyed is three. In Figure 5, an optimal set of edges is shown in red. Note that the edges happen to be in a 'thin' portion of the complex, justifying our standpoint that (along with Boundary Nontrivialization) this problem can also be seen as a high dimensional cut problem.

In this work, we undertake an algorithmic study of the two high-dimensional cut problems: Boundary Nontrivialization and Topological Hitting Set.

1.1 Related work

Duval et al. [21] study the vector spaces and integer lattices of cuts and flows associated to CW complexes and their relationships to group invariants. Ghrist and Krishnan [26] prove a topological version of the max-flow min-cut theorem for directed networks using methods from sheaf theory. Then, there is also a long line of work on cuts in surface embedded graphs [2, 3, 8, 9, 10, 11], which is algorithmic in spirit and is loosely related to our work.

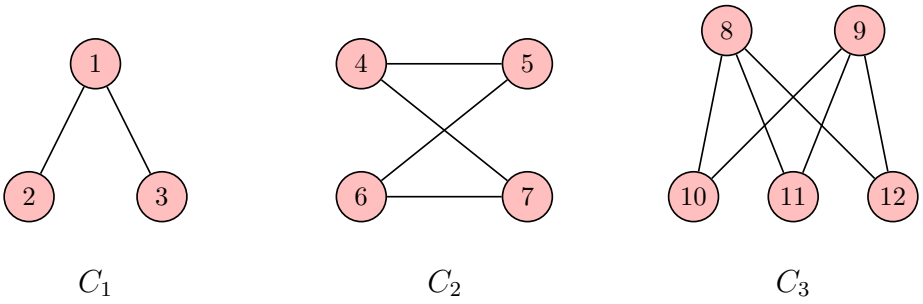


Figure 4 Graph G with three components

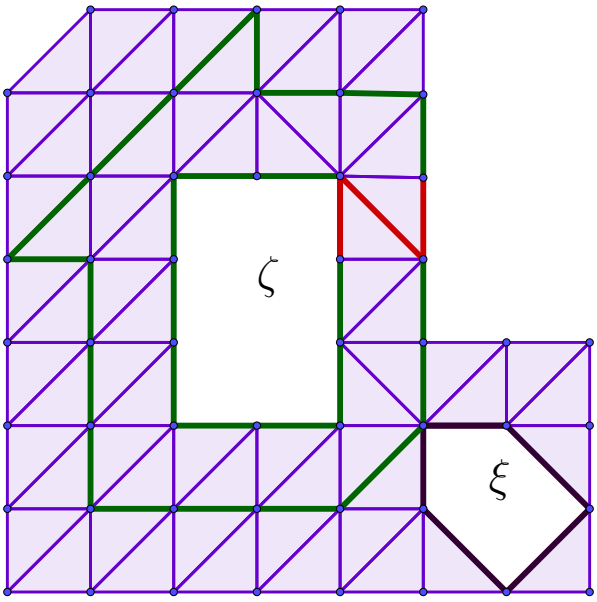


Figure 5 The figure shows two cycles that belong to $[\zeta]$ in green. Note that any cycle in $[\zeta]$ must pass through at least one of the three red edges. Thus, the set of red edges constitutes an optimal solution for Topological Hitting Set on this planar complex.

There is a growing body of work on parameterized complexity in topology [1, 4, 5, 6, 28, 30, 31, 35], and much of this paper can be characterized as such.

During the preparation of this article, we became aware of a recent paper by Maxwell and Nayyeri [32] that studies problems similar to the ones we define but from a completely different point of view. While our focus was on surfaces and parameterized complexity, the main focus of their work was to find out the extent to which the conceptual and the algorithmic framework of max-flow min-cut duality generalizes to the case of simplicial complexes. While we focus only on cuts, they study both cuts and flows.

We summarize the main results of Maxwell and Nayyeri [32] as we understand them: They define a topological max-flow and a topological min-cut problem, and also a combinatorial min-cut problem. They show that unlike in the case of graphs, computing maximum integral flows and combinatorial cuts on simplicial complexes is **NP**-hard. Moreover, they describe conditions under which the linear program gives the optimal value of a combinatorial cut, and also provide a generalization of the Ford-Fulkerson algorithm to the case of simplicial complexes. Their definition of combinatorial cut coincides with our definition of Boundary Nontrivialization, except for some important differences: they are interested in real coefficients and co-dimension one cycles, whereas we work with $\mathbb{Z}_2$ coefficients and cycles of all dimensions. We implore the reader to look up their interesting results [32].

We note that while their paper is in the same spirit as ours, their focus is quite different from ours, and there is very little overlap in terms of hardness or algorithmic results. In particular, they show **NP**-hardness for combinatorial cuts with real coefficients, and we show **NP**-hardness and **W[1]**-hardness for the same problem with $\mathbb{Z}_2$ coefficients.

2 Summary of results

Surfaces. Our first result, expounded in Section 5, is the following: Topological Hitting Set admits a polynomial-time algorithm on triangulations of closed surfaces. At the heart of our proof lies an appealing characterization of the optimal solutions in terms of the cocycles of the surface, which is of independent interest. Specifically, we show that a minimal solution set is necessarily a nontrivial cocycle. Further, we show that the following are equivalent: **1.** A connected cocycle η is a feasible set for the input cycle ζ. **2.** Every cycle in $[\zeta]$ intersects a connected cocycle η in an odd number of edges. **3.** One of the cycles in $[\zeta]$ intersects a connected cocycle η in an odd number of edges.

In particular, this allows us to identify the nontrivial cocycles that are solutions based on a parity-based property. Having this characterization at hand, we proceed to characterize cohomology classes that are solutions. Eventually, we arrive at a very simple 3-step algorithm for Topological Hitting Set on surfaces.

We remark that Boundary Nontrivialization is trivial for surfaces. In fact, it is easy to check that for some boundary b and a 2-chain ζ, if $\partial\zeta = b$, then removing any one of the triangles that appears in the chain ζ makes b nontrivial.

W[1]-hardness and NP-hardness. For general complexes, in Section 6.1, we show that Topological Hitting Set is **W[1]**-hard with respect to the solution size k as the parameter, (and hence, it is also **NP**-hard). The proof is based on a reduction from the k-Multicolored Clique problem. Here, the reduction shows the essence of hardness: its description is short, but its proof exposes various "behaviors" that we find interesting. In particular, the forward direction requires a nontrivial parity based argument, while the reverse direction shows how to "trace" a solution through the complex.

In addition, in Section 6.2, we show that Boundary Nontrivialization is also **W[1]**-hard with respect to the solution size k as a parameter. The principles of this reduction follow the lines of the reduction for Topological Hitting Set, though, here, both the description and the proof of the reduction are more involved because of subdivisions that help avoid some unhelpful incidences.

Fixed-parameter tractability. On the positive side, in Section 7.1, we show that Topological Hitting Set admits an **FPT** algorithm with respect to $k + \Delta$, where Δ is the maximum degree of the Hasse graph of the complex K. Here, the main insight is that a minimal solution must be connected. Having this insight at hand, the algorithm follows: If we search across the geodesic ball of every r-simplex in the complex K, we will find a solution.

In contrast, we observe that Boundary Nontrivialization does not admit this property because minimal solutions can be disconnected. This motivates the search of another parameter that makes the problem tractable. Exploiting the set-cover like structure of the problem, in Section 7.2, we show that Boundary Nontrivialization with bounding r-cycles as input has an $O(\log n)$-approximation **FPT** algorithm with β_{r+1} (the Betti number) as the parameter, when the input complex K is $(r+1)$-dimensional. It is worth noting that Boundary Nontrivialization is **W[1]**-hard even for $(r+1)$-dimensional complexes with solution size as the parameter since the hardness gadget used in Section 6.2 is $(r+1)$-dimensional.

By exploiting the vector space structure of the homology groups and the boundary groups, in Sections 7.1.1 and 7.2.1, we provide a randomized **FPT** algorithm for Global Topological Hitting Set and a randomized **FPT** approximation algorithm for Global Boundary Nontrivialization respectively.

3 Preliminaries

3.1 Simplicial complexes

A *k-simplex* σ is the convex hull of a set V of $(k+1)$ affinely independent points in the Euclidean space of dimension $d \geq k$. We call k the dimension of σ. Any nonempty subset of V also spans a simplex, which we call a *face* of σ. A simplex σ is said to be a *coface* of a simplex τ if and only if τ is face of σ. We say that σ is a *facet* of τ, and τ a *cofacet* of σ, if σ is a face of τ with $\dim \sigma = \dim \tau - 1$. We denote a facet-cofacet pair by $\sigma \prec \tau$. A *simplicial complex* K is a collection of simplices that satisfies the following conditions:

- any face of a simplex in K also belongs to K, and
- the intersection of two simplices $\sigma_1, \sigma_2 \in \mathsf{K}$ is either empty or a face of both σ_1 and σ_2.

An *abstract simplicial complex* K on a set of vertices V is a collection of subsets of V that is closed under inclusion. The elements of K are called its *simplices*. An abstract simplicial complex L is said to be a *subcomplex* of K if every simplex of L belongs to K.

The collection of vertex sets of simplices in a geometric simplicial complex forms an abstract simplicial complex. On the other hand, an abstract simplicial complex K has a geometric realization $|\mathsf{K}|$ obtained by embedding the points in V in general position in a high-dimensional Euclidean space. Then, the complex $|\mathsf{K}|$ is defined as $\bigcup_{\sigma \in \mathsf{K}} |\sigma|$, where $|\sigma|$ denotes the span of points in σ. It is not very difficult to show that any two geometric realizations of an abstract simplicial complex are homeomorphic. Hence, going forward, we do not distinguish between abstract and geometric simplicial complexes.

The *star* of a vertex v of complex K, written $\mathrm{star}_{\mathsf{K}}(v)$, is the subcomplex consisting of all faces of K containing v, together with their faces.

Let V be the vertex set of K, W be the vertex set of L and ϕ be a map from V to W. If for every simplex $\{v_0, v_1, \ldots, v_r\} \in \mathsf{K}$, the vertices $\{\phi(v_0), \phi(v_1), \ldots, \phi(v_r)\}$ span a simplex in L, then the ϕ induces a map, say f, from K to L. The induced map $f : \mathsf{K} \to \mathsf{L}$, is said to be *simplicial*.

We will denote by $\mathsf{K}^{(p)}$ the set of p-dimensional simplices in K, and n_p the number of p-dimensional simplices in K. The complex induced by $\mathsf{K}^{(p)}$ is called the *p-dimensional skeleton* of K, and is denoted by K_p. Given a simplicial complex K, we denote by the H_{K}, the *Hasse graph* of K, which is simply the graph that has a node for every simplex of the complex, and an edge for every facet-cofacet pair. Given a triangulated closed surface K, we denote by D_{K}, the *dual graph* of K, which is simply the graph that has a node for every 2-simplex and an edge connecting two nodes if the corresponding 2-simplices are incident on a common edge in the complex. The *stellar subdivision* of a simplex (or a polytope) is the complex formed by taking a cone over its boundary.

▶ Notation 1. We use $[m]$ to denote the set $\{1, 2, ..., m\}$ for any $m \in \mathbb{N}$.

3.2 Homology and cohomology

In this work, we restrict our attention to simplicial homology with $\mathbb{Z}_2$ coefficients. For a general introduction to algebraic topology, we refer the reader to [27]. Below we give a brief description of homology over $\mathbb{Z}_2$.

Let K be a connected simplicial complex. We consider formal sums of simplices with $\mathbb{Z}_2$ coefficients, that is, sums of the form $\sum_{\sigma \in \mathsf{K}^{(p)}} a_\sigma \sigma$, where each $a_\sigma \in \{0, 1\}$. The expression $\sum_{\sigma \in \mathsf{K}^{(p)}} a_\sigma \sigma$ is called a *p-chain*. Since chains can be added to each other, they form an Abelian group, denoted by $\mathsf{C}_p(\mathsf{K})$. Since we consider formal sums with coefficients coming from $\mathbb{Z}_2$, which is a field, $\mathsf{C}_p(\mathsf{K})$, in this case, is a vector space of dimension n_p over $\mathbb{Z}_2$. The p-simplices in K form a (natural) basis for $\mathsf{C}_p(\mathsf{K})$. This establishes a natural one-to-one correspondence between elements of $\mathsf{C}_p(\mathsf{K})$ and subsets of $\mathsf{K}^{(p)}$, and we will freely make use of this identification. The *boundary* of a p-simplex is a $(p-1)$-chain that corresponds to the set of its $(p-1)$-faces. This map can be linearly extended from p-simplices to p-chains, where the boundary of a chain is the $\mathbb{Z}_2$-sum of the boundaries of its elements. The resulting *boundary homomorphism* is denoted by $\partial_p : \mathsf{C}_p(\mathsf{K}) \to \mathsf{C}_{p-1}(\mathsf{K})$. A chain $\zeta \in \mathsf{C}_p(\mathsf{K})$ is called a *p-cycle* if $\partial_p \zeta = 0$, that is, $\zeta \in \ker \partial_p$. The group of p-dimensional cycles is denoted by $\mathsf{Z}_p(\mathsf{K})$. As before, since we are working with $\mathbb{Z}_2$ coefficients, $\mathsf{Z}_p(\mathsf{K})$ is a vector space over $\mathbb{Z}_2$. A chain $\eta \in \mathsf{C}_p(\mathsf{K})$ is said to be a *p-boundary* if $\eta = \partial_{p+1} c$ for some chain $c \in \mathsf{C}_{p+1}(\mathsf{K})$, that is, $\eta \in \operatorname{im} \partial_{p+1}$. The vector space of p-dimensional boundaries is denoted by $\mathsf{B}_p(\mathsf{K})$.

In our case, $\mathsf{B}_p(\mathsf{K})$ is also a vector space, and in fact a subspace of $\mathsf{C}_p(\mathsf{K})$. Thus, we can consider the quotient space $\mathsf{H}_p(\mathsf{K}) = \mathsf{Z}_p(\mathsf{K})/\mathsf{B}_p(\mathsf{K})$. The elements of the vector space $\mathsf{H}_p(\mathsf{K})$, known as the p-th *homology* of K, are equivalence classes of p-cycles, called *homology classes* where p-cycles are said to be *homologous* if their $\mathbb{Z}_2$-difference is a p-boundary. For a p-cycle ζ, its corresponding homology class is denoted by $[\zeta]$. Bases of $\mathsf{B}_p(\mathsf{K})$, $\mathsf{Z}_p(\mathsf{K})$ and $\mathsf{H}_p(\mathsf{K})$ are called *boundary bases*, *cycle bases*, and *homology bases*, respectively. The dimension of the p-th homology of K is called the p-th *Betti number* of K, denoted by $\beta_p(\mathsf{K})$.

Using the natural bases for $\mathsf{C}_p(\mathsf{K})$ and $\mathsf{C}_{p-1}(\mathsf{K})$, the matrix $[\partial_p \sigma_1 \; \partial_p \sigma_2 \cdots \partial_p \sigma_{n_p}]$ whose column vectors are boundaries of p-simplices is called the *p-th boundary matrix*. Abusing notation, we also denote the p-th boundary matrix by ∂_p.

The dual vector space of $\mathsf{C}_p(\mathsf{K})$ (the vector space of linear maps $\mathsf{C}_p(\mathsf{K}) \to \mathbb{Z}_2$) is called the space of *cochain*, denoted by $\mathsf{C}^p(\mathsf{K}) = \operatorname{Hom}(\mathsf{C}_p(\mathsf{K}), \mathbb{Z}_2)$. Again, there is a natural basis corresponding to the p-simplices of K, with a p-simplex σ corresponding to the linear map η with values $\eta(\sigma) = 1$ and $\eta(\rho) = 0$ for every other p-simplex $\rho \neq \sigma$. The adjoint map to

the boundary map $\partial_{p+1}\colon \mathsf{C}_{p+1}(\mathsf{K}) \to \mathsf{C}_p(\mathsf{K})$ is the *coboundary map* $\delta_p\colon \mathsf{C}^p(\mathsf{K}) \to \mathsf{C}^{p+1}(\mathsf{K})$. Similarly to chains and boundary maps, we may define subspaces of *cocycles* $\mathsf{Z}^p(\mathsf{K}) = \ker \delta_p$ and *coboundaries* $\mathsf{B}^p(\mathsf{K}) = \operatorname{im} \delta_{p+1} \subseteq \mathsf{Z}^p(\mathsf{K})$, and form their quotient $\mathsf{H}^p(\mathsf{K}) = \mathsf{Z}^p(\mathsf{K})/\mathsf{B}^p(\mathsf{K})$, which is the *cohomology* of K. Again, for a p-cocycle η, the corresponding cohomology class is denoted by $[\eta]$. The natural pairing of chains and cochains $\mathsf{C}_p(\mathsf{K}) \times \mathsf{C}^p(\mathsf{K}) \to \mathbb{Z}_2, (\zeta, \eta) \mapsto \eta(\zeta)$ induces a well-defined isomorphism $\mathsf{H}^p(\mathsf{K}) \times \mathsf{H}_p(\mathsf{K}) \to \mathbb{Z}_2, ([\zeta], [\eta]) \mapsto \eta(\zeta)$, identifying cohomology as the vector space dual to homology up to a natural isomorphism.

A set of p-cycles $\{\zeta_1, \ldots, \zeta_g\}$ is called a *homology cycle basis* if the set of classes $\{[\zeta_1], \ldots, [\zeta_g]\}$ forms a homology basis. For brevity, we abuse notation by using the term *(p-th) homology basis* for $\{\zeta_1, \ldots, \zeta_g\}$. Similarly, a set of p-cocycles $\{\eta_1, \ldots, \eta_g\}$ is called a *(p-th) cohomology cocycle basis* if the set of classes $\{[\eta_1], \ldots, [\eta_g]\}$ forms a cohomology basis.

Assigning non-negative weights to the edges of K, the weight of a cycle is the sum of the weights of its edges, and the weight of a homology basis is the sum of the weights of the basis elements. We call the problem of computing a minimum weight basis of $\mathsf{H}_1(\mathsf{K})$ the *minimum homology basis* problem. Similarly, we call the problem of computing a minimum weight basis of $\mathsf{H}^1(\mathsf{K})$, the *minimum cohomology basis* problem.

▶ Notation 2. Since there is a 1-to-1 correspondence between the p-chains of a complex K and the subsets of $\mathsf{K}^{(p)}$, we abuse notation by writing $\partial\mathcal{C}$ in place of $\partial(\sum_{\sigma\in\mathcal{C}} \sigma)$, for $\mathcal{C} \subset \mathsf{K}^{(p)}$. Likewise, for p-cochains $\delta(\sum_{\tau\in\mathcal{C}'} \tau)$, we often write $\delta\mathcal{C}'$.

We also abuse notation in the other direction. That is, we treat chains and cochains as sets. For instance, sometimes we say that a (co)chain γ intersects a (co)chain ζ, when we actually mean that the corresponding sets of simplices of the respective (co)chains intersect. Also, we say that a simplex $\sigma \in \zeta$, when indeed the simplex σ belongs to the set associated to ζ.

3.3 Parameterized complexity

Let Π be an **NP**-hard problem. In the framework of Parameterized Complexity, each instance of Π is associated with a *parameter* k. Here, the goal is to confine the combinatorial explosion in the running time of an algorithm for Π to depend only on k. Formally, we say that Π is *fixed-parameter tractable (***FPT***)* if any instance (I, k) of Π is solvable in time $f(k) \cdot |I|^{O(1)}$, where f is an arbitrary computable function of k.

A weaker request is that for every fixed k, the problem Π would be solvable in polynomial time. Formally, we say that Π is *slice-wise polynomial (***XP***)* if any instance (I, k) of Π is solvable in time $f(k) \cdot |I|^{g(k)}$, where f and g are arbitrary computable functions of k. In other words, for a fixed k, Π has a polynomial time algorithm, and we refer to such an algorithm as an **XP** algorithm for Π. Nowadays, Parameterized Complexity supplies a rich toolkit to design **FPT** and **XP** algorithms [16, 20, 24].

Parameterized Complexity also provides methods to show that a problem is unlikely to be **FPT**. The main technique is the one of parameterized reductions analogous to those employed in classical complexity. Here, the concept of **W**-hardness replaces the one of NP-hardness, and for reductions we need not only construct an equivalent instance in **FPT** time, but also ensure that the size of the parameter in the new instance depends only on the size of the parameter in the original one.

▶ **Definition 1 (Parameterized Reduction).** *Let Π and Π' be two parameterized problems. A* parameterized reduction *from Π to Π' is an algorithm that, given an instance (I, k) of Π, outputs an instance (I', k') of Π' such that:*

- *(I, k) is a yes-instance of Π if and only if (I', k') is a yes-instance of Π'.*

- $k' \leq g(k)$ *for some computable function* g.
- *The running time is* $f(k) \cdot |\Pi|^{O(1)}$ *for some computable function* f.

If there exists such a reduction transforming a problem known to be **W[1]**-hard to another problem Π, then the problem Π is **W[1]**-hard as well. Central **W[1]**-hard problems include, for example, deciding whether a nondeterministic single-tape Turing machine accepts within k steps, Clique parameterized by solution size, and Independent Set parameterized by solution size. To show that a problem Π is not **XP** unless $\mathbf{P} = \mathbf{NP}$, it is *sufficient* to show that there exists a fixed k such that Π is **NP**-hard. If the problem Π is in **NP** for a fixed k then it is said to be in para-**NP**, and if is **NP**-hard for a fixed k then it is said to be para-**NP**-hard.

Now, suppose that the parameter k does not depend on the sought solution size, but it is a structural parameter. Then, we say that a minimization (maximization) problem Π admits a c-approximation **FPT** (with respect to k) if it admits a $f(k) \cdot |I|^{O(1)}$-time algorithm that, given an instance (I, k) of Π, outputs a solution for (I, k) that is larger (smaller) than the optimal solution for (I, k) by a factor of at most c. When the parameter k does depend on the sought solution size, the notion of a c-approximation **FPT** algorithm is defined as well, but this definition is slightly more complicated and is not required in this paper. For more information on Parameterized Complexity, we refer the reader to recent books such as [16, 20, 24].

4 Problem definitions

In this section, we define the two key problems of interest, namely, Topological Hitting Set and Boundary Nontrivialization along with their global variants, namely, Global Topological Hitting Set and Global Boundary Nontrivialization, respectively. Also, we observe that all four problems lie in **NP** and in **XP** with respect to the solution size as the parameter.

4.1 Topological Hitting Set

Problem 1 (Topological Hitting Set).

Instance:	Given a d-dimensional simplicial complex K, a natural number k, a natural number $r < d$ and a non-bounding cycle $\zeta \in \mathsf{Z}_r(\mathsf{K})$.
Parameter:	k.
Question:	Does there exists a set $\mathcal{S}$ of r-dimensional simplices with $\lvert\mathcal{S}\rvert \leq k$ such that $\mathcal{S}$ meets every cycle homologous to ζ?

Let $\mathsf{K}_{\mathcal{C}}$ denote the complex obtained from K upon removal of the set of r-simplices $\mathcal{C}$ along with all the cofaces of the simplices in $\mathcal{C}$. In particular, the homology class $[\zeta]$ does not survive in $\mathsf{K}_{\mathcal{C}}$.

Let $\{\alpha_i\}$ for $i \in [\beta_r(\mathsf{K}_{\mathcal{C}})]$ be a homology basis for $\mathsf{K}_{\mathcal{C}}$. The inclusion map $\iota : \mathsf{K}_{\mathcal{C}} \hookrightarrow \mathsf{K}$ induces a map $\hat{\iota} : \mathsf{Z}_r(\mathsf{K}_{\mathcal{C}}) \to \mathsf{Z}_r(\mathsf{K})$ and also a map $\tilde{\iota} : \mathsf{H}_r(\mathsf{K}_{\mathcal{C}}) \to \mathsf{H}_r(\mathsf{K})$. Let $\hat{\alpha}_i = \hat{\iota}(\alpha_i)$. Let $\mathbf{A}$ denote the matrix with nontrivial r-cycles $\hat{\alpha}_i$ as its columns. Let $\mathbf{M}$ denote the matrix $[\mathbf{A} \mid \partial_{r+1}(\mathsf{K})]$ and $C(\mathbf{M})$ the column space of $\mathbf{M}$. The following lemma ensures polynomial time verification for the decision variant of Topological Hitting Set.

▶ **Lemma 2.** $\zeta \notin$ *column space of* $\boldsymbol{M}$ *if and only if* $\mathcal{S}$ *meets every cycle homologous to* ζ.

Proof. ($\Longrightarrow$) Let ρ be a cycle homologous to ζ such that $\mathcal{S}$ does not meet ρ. Then, $\rho \in C(\mathbf{M})$ since it survives in $\mathsf{K}_{\mathcal{C}}$. The claim follows from observing that ζ is homologous to ρ.

($\Longleftarrow$) Suppose that $\mathcal{S}$ meets every cycle that is homologous to ζ. Thus, at least one simplex is removed from every cycle homologous to ζ. Then, a cycle homologous to ζ (in K) is not present in $\mathsf{K}_{\mathcal{C}}$. The claim follows. ◀

Lemma 2 provides an easy way to check if a set constitutes a feasible solution.

▶ **Theorem 3.** *Checking if a set $\mathcal{S}$ is a feasible solution to* Topological Hitting Set *amounts to solving a linear system of equations, and can be done in $O(n^{\omega})$ time, where ω is the exponent of matrix multiplication, and n is the size of the complex.*

▶ **Corollary 4.** Topological Hitting Set *is in* **NP**, *and is in* **XP** *with respect to the solution size k as the parameter.*

We now define the global variant of Topological Hitting Set.

Problem 2 (Global Topological Hitting Set).

Instance:	Given a d-dimensional simplicial complex K, a natural number k, a natural number $r < d$.
Parameter:	k.
Question:	Does there exists a set $\mathcal{S}$ of r-dimensional simplices with $\lvert\mathcal{S}\rvert \leq k$ such that the induced map on homology $\tilde{\iota} : \mathsf{H}_r(\mathsf{K}_{\mathcal{S}}) \to \mathsf{H}_r(\mathsf{K})$ is non-surjective?

For a complex L, let $\mathcal{H}_r(\mathsf{L})$ denote an r-th homology basis of L. It is well-known that such a basis can always be computed in polynomial time.

▶ **Theorem 5.** Global Topological Hitting Set *is in* **NP**, *and in* **XP** *with respect to the solution size k as the parameter.*

Proof. $\mathcal{S}$ is a solution for Global Topological Hitting Set if and only if one of the two conditions is satisfied:

- $\beta_r(\mathsf{K}_{\mathcal{S}}) < \beta_r(\mathsf{K})$ or
- $\beta_r(\mathsf{K}_{\mathcal{S}}) \geq \beta_r(\mathsf{K})$ and in the column rank profile of the matrix $[\partial_{r+1}(\mathsf{K}) \mid \mathcal{H}_r(\mathsf{K}_{\mathcal{S}})]$, of the last $\beta_r(\mathsf{K}_{\mathcal{S}})$ columns, exactly $\beta_r(\mathsf{K})$ columns are nonzero.

The two conditions can be verified in polynomial time, proving the claim. ◀

4.2 Boundary Nontrivialization

Problem 3 (Boundary Nontrivialization).

Instance:	Given a d-dimensional simplicial complex K, a natural number k, a natural number $r < d$ and a bounding cycle $\zeta \in \mathsf{B}_r(\mathsf{K})$.
Parameter:	k.
Question:	Does there exists a set $\mathcal{S}$ of $r+1$-dimensional simplices with $\lvert\mathcal{S}\rvert \leq k$ such that removal of $\mathcal{S}$ from the K makes ζ non-bounding?

▶ **Theorem 6.** Boundary Nontrivialization *is in* **NP**, *and is in* **XP** *with respect to the solution size k as the parameter.*

Proof. A set $\mathcal{S}$ is a solution if and only if the system of equations $\partial_{r+1}(\mathsf{K}_{\mathcal{S}}) \cdot \mathbf{x} = \zeta$ has no solution, which can be checked in polynomial time. ◀

The global variant of Boundary Nontrivialization can be described as follows.

Problem 4 (Global Boundary Nontrivialization).

Instance:	Given a d-dimensional simplicial complex K, and a natural number k.
Parameter:	k.
Question:	Does there exists a set $\mathcal{S}$ of $r+1$-dimensional simplices with $\lvert\mathcal{S}\rvert \leq k$ such that the column space of $\partial_{r+1}(\mathsf{K}_{\mathcal{S}})$ is a strictly smaller subspace of the column space of $\partial_{r+1}(\mathsf{K})$?

▶ **Theorem 7.** Global Boundary Nontrivialization *is in* **NP**, *and is in* **XP** *with respect to the solution size k as the parameter.*

Proof. It is easy to check in polynomial time if the column space of $\partial_{r+1}(\mathsf{K}_{\mathcal{S}})$ is a strictly smaller subspace of $\partial_{r+1}(\mathsf{K})$. ◀

5 Topological Hitting Set on surfaces

In this section we describe a polynomial time algorithm for Topological Hitting Set on surfaces. Let ζ be a nontrivial 1-cycle in a triangulated closed surface K. The algorithm for surfaces has a very simple high-level description as detailed in Algorithm 1.

▶ Notation 3. Note that if we evaluate the r-cocycle η at an r-cycle ζ, then by linearity,

$$\eta(\zeta) = \eta(\sum_{\sigma_i \in \zeta} \sigma_i) = \sum_{\sigma_i \in \zeta} \eta(\sigma_i).$$

Because of $\mathbb{Z}_2$ addition, $\eta(\zeta)$ is either 0 or 1.

Algorithm 1 The algorithm for Topological Hitting Set on surfaces with input cycle ζ

1: Find the optimal cohomology basis of K with unit weights on edges.
2: Arrange the cocycles in the basis in ascending order of weight.
3: Pick the smallest weight cocycle η with $\eta(\zeta) = 1$.

In what follows, we will establish a series of structural results about the solution set for Topological Hitting Set on surfaces in order to prove the correctness of Algorithm 1. We begin with a few definitions.

▶ **Definition 8** (Connected cocycles). *A cocycle η is said to be* connected *if it induces a connected component in the dual graph, else we say that it is* disconnected.

In Lemma 11, we show that a minimal solution is, in fact, a connected cocycle. Since cocycles can be potential solutions for Topological Hitting Set, we make the following definitions.

▶ **Definition 9.** *We say that a cocycle η is said to be a* feasible set *if every cycle $\zeta' \in [\zeta]$ meets η in an edge. A cocycle that is not a feasible set, is said to be an* infeasible set.

Next, we provide a useful characterization of cocycles that constitute feasible sets.

▶ **Lemma 10.** *If there exists a cycle in $[\zeta]$ that intersects a connected cocycle η in an odd number of edges, then every cycle in $[\zeta]$ intersects η in an odd number of edges.*

Proof. Suppose that a cycle $\gamma \in [\zeta]$ intersects η in an odd number of edges. Let $\gamma' = \gamma + \partial\sigma$. We claim that γ' intersects η in an odd number of edges. We have four cases to consider.

1. The simplex boundary $\partial\sigma$ is not incident on η. Then, the homologous cycle obtained by addition of $\partial\sigma$ maintains odd incidence.
2. The simplex boundary $\partial\sigma$ intersects η in two edges and both edges also belong to γ. Then, addition of $\partial\sigma$ to γ reduces the number of incident edges on η by two, and the number stays odd.
3. The simplex boundary $\partial\sigma$ intersects η in two edges and none of the edges belong to γ. Then, addition of $\partial\sigma$ to γ increases the number of incident edges on η by two, and the number stays odd.
4. The simplex boundary $\partial\sigma$ intersects η in two edges one of which belongs to γ. Then, upon addition of $\partial\sigma$ to γ, the incident edge is exchanged with the non-incident one, and the incidence number stays the same.

Figure 6 illustrates the four cases. Any cycle in $[\zeta]$ can be obtained by adding simplex boundaries $\sum_i \partial\sigma_i$ to γ. So, applying the four cases inductively, we see that every cycle in $[\zeta]$ has odd incidence on η. ◀

Note that any connected cocycle η induces a cycle graph which is a subgraph of the dual graph D_{K} of the surface K. We denote the cycle graph by C_η.

▶ **Lemma 11.** *A minimal solution set is a cocycle η that induces a circle subgraph C_η in the dual graph D_{K}.*

Proof. Let e_1 be an edge in the minimal solution set $\mathcal{S}$. Let σ be a 2-simplex incident on e_1, and let e_2 and e_3 be the other two edges incident on σ. Let $\gamma \in [\zeta]$ be a cycle with e_1 as the *unique* edge incident on $\mathcal{S}$. We know that such a cycle exists because of minimality of $\mathcal{S}$. Then, there exists a cycle $\gamma' = \gamma + \partial\sigma$ with e_2 and e_3 incident on it. Since γ and γ' differ only by a boundary $\partial\sigma$, using the fact that e_1 is the unique edge incident on $\mathcal{S}$, either e_2 or e_3 must be incident on $\mathcal{S}$. Without loss of generality, assume that e_2 is incident on $\mathcal{S}$. Now, consider the 2-simplex $\tau \neq \sigma$ incident on e_2. Using the same argument as before, and proceeding by induction, we obtain a sequence of edges in $\mathcal{S}$ starting from e_1, each connected by a 2-simplex. Then, there must exist a sequence starting at e_1 and ending at an edge e' such that both e' and e_1 are incident on a common cofacet $\rho \neq \sigma$, for if this is not the case, then we can find a a cycle $\gamma' \in [\zeta]$ which is not incident on $\mathcal{S}$. The sequence of edges from e_1 to e' forms a cocycle, say η, where $\eta \subset \mathcal{S}$.

Targeting a contradiction, assume that $\eta \neq \mathcal{S}$. By Lemma 63, $\mathcal{S}$ induces a connected subgraph in the Hasse graph, which implies that there exists an edge $e' \in \mathcal{S} \setminus \eta$ and a 2-simplex τ such that $e' \prec \tau$ and the other two edges $e_1, e_2 \prec \tau$ belong to η. Let γ be a cycle with e_1 as the unique edge from $\mathcal{S}$ incident on it, and let γ' be a cycle with e' as the unique edge from $\mathcal{S}$ incident on it. Let $\mathcal{B}$ be the set of boundaries of 2-simplices added to γ in order to obtain γ' from γ. Since γ intersects η in an odd number of edges, using Lemma 10, any cycle homologous to γ will also intersect η in an odd number of edges. Hence, γ' is incident on at least one of the edges of η. But this contradicts the existence of γ' since by assumption, γ' is not incident on η. Therefore, an edge $e' \in \mathcal{S} \setminus \eta$ that shares a cofacet incident on two of the edges of η does not exist.

Finally, using Lemma 63 from Section 7.1, for any edge $f' \in \mathcal{S} \setminus \eta$, and a cofacet ρ of f', there must be a path in the dual graph from ρ to a 2-simplex ϱ incident on two of the edges of η. But for such a path to exist, there must exist an edge $e' \in \mathcal{S} \setminus \eta$ and a 2-simplex τ such that $e' \prec \tau$ and the other two edges $e_1, e_2 \prec \tau$ belong to η. But we showed that such an edge

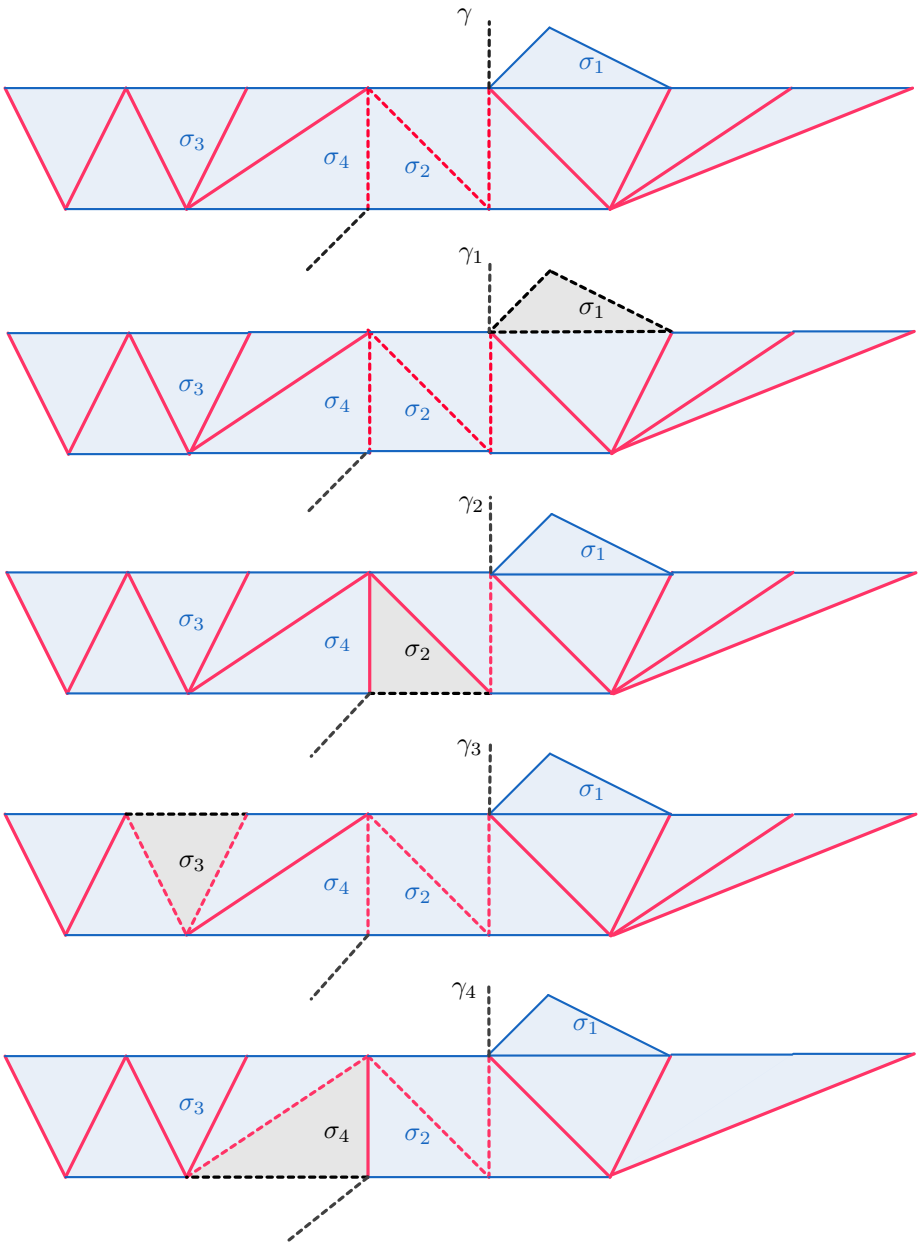


Figure 6 In this figure, we illustrate the four cases discussed in Lemma 10. Let ζ be the input cycle and $\gamma \in [\zeta]$. The dotted edges in the topmost figure belong to γ. Let η be a cocycle that meets γ in an odd number of edges. The edges of η in each of the 5 figures are shown in pink. The part of γ that does not intersect η is shown in black and the part of γ that intersects η is shown in pink. As shown in the topmost figure, γ intersects η in three (pink-dotted) edges. In the four cases labelled (I-IV), $\gamma_i = \gamma + \partial\sigma_i$, where $i \in [4]$. As before, for every $i \in [4]$, the edges of γ_i that intersect the edges of η are shown as pink-dotted edges and the edges of γ_i that do not intersect η are shown as black-dotted edges. Note that in each of the four cases, the number of pink-dotted edges is odd.

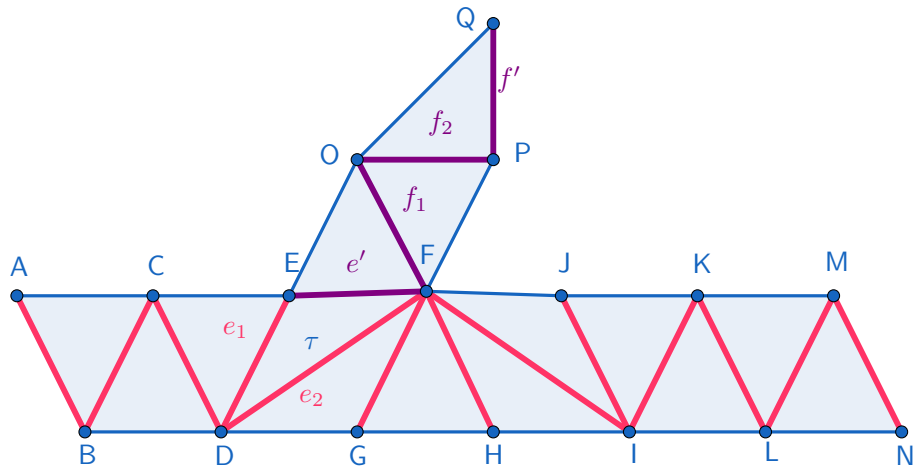


Figure 7 The cocycle η is shown in red. e_1, e_2, e', f' and τ are as in Lemma 11. Note that there exists a path from e' to f' in the dual graph.

e' does not exist. So, the set $\mathcal{S} \setminus \eta$ is empty. Since $\eta = \mathcal{S}$, the claim follows. Please refer to Figure 7 for an example. Note that the final part of the argument is specific to surfaces. ◀

▶ **Lemma 12.** *A trivial cocycle is not a minimal solution set.*

Proof. Let η be a trivial cocycle. Then, $\eta = \delta(\mathcal{S})$, where $\mathcal{S}$ is a collection of points. Let e be an edge of η. By the assumption on minimality of the solution set, there exists a cycle $\zeta' \in [\zeta]$ such that e is the only edge of η incident on ζ'. One of the vertices of e, say v_1, belongs to $\mathcal{S}$. We write the cycle ζ' as a sequence of vertices $v_1, v_2, \ldots, v_q = v_1$, for some q, such that an edge connects subsequent vertices, the sequence starts and ends at v_1, and the edge $\{v_{q-1}, v_1\} = e$. Then, the path from v_1 to v_{q-1} must pass through a vertex v' such that $v' \in \mathcal{S}$. But this is only possible if ζ' also contains an edge of η other than e, which in turn, contradicts the minimality of the solution set. Hence, a trivial cocycle is not a minimal solution set. See Figure 8 for an example. ◀

Lemmas 11 and 12 combine to give the following theorem.

▶ **Theorem 13.** *A minimal solution set is a nontrivial cocycle.*

▶ **Lemma 14.** *If a connected cocycle η intersects a cycle $\zeta_0 \in [\zeta]$ in m edges, then there exists another cycle $\gamma \in [\zeta]$ such that γ also intersects η in m edges, and the intersection of γ and η induces a connected component in the dual graph.*

Proof. To begin with, note that the intersection of $\zeta_0 \in [\zeta]$ with η induces a (possibly disconnected) subgraph of the cycle graph C_η, which we denote by $C_\eta^{\zeta_0}$. Both C_η and $C_\eta^{\zeta_0}$ are subgraphs of the dual graph D_K. Let $\mathcal{C}_1, \ldots, \mathcal{C}_k$ be the k connected components of $C_\eta^{\zeta_0}$. If $k = 1$, then the lemma is already satisfied. So without loss of generality, assume $k > 1$. We say that a component $\mathcal{C}_i$ is a *neighbor* of a component $\mathcal{C}_j$ if there exists a vertex in $\mathcal{C}_i$ that has a path to a vertex in $\mathcal{C}_j$ that does not intersect the edges of $C_\eta^{\zeta_0}$. It is easy to check that every component of $C_\eta^{\zeta_0}$ has exactly two (possibly non-distinct) neighbors. Choose any two neighboring components $\mathcal{C}_i$ and $\mathcal{C}_j$ of $C_\eta^{\zeta_0}$. Let $v \in \mathcal{C}_i$ and $v' \in \mathcal{C}_j$ be two vertices that have a simple path $\mathcal{P}$ with vertices $v = u_1, u_2, \ldots, u_\ell = v'$ such that $\mathcal{P}$ does not intersect the

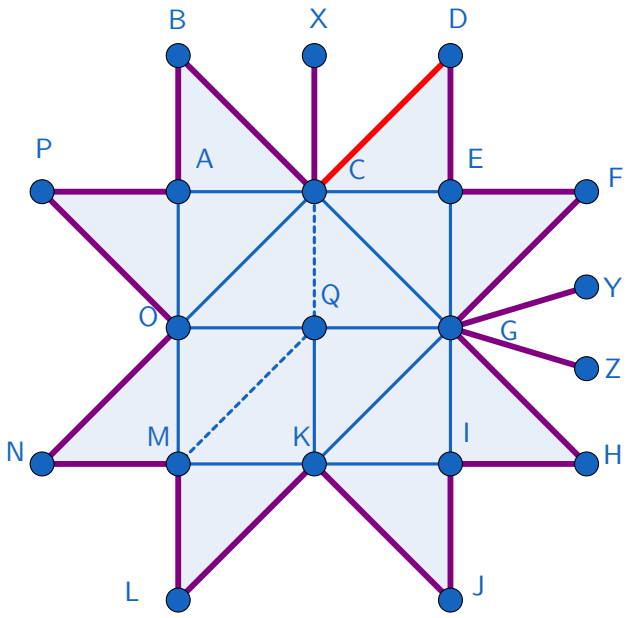


Figure 8 The edges in purple and red together form a trivial cocycle η given by $\delta(A+C+E+O+Q+G+M+K+I)$. If η forms a minimal solution set, then it intersects a cycle $\zeta' \in [\zeta]$ in a unique edge, say, edge DC. But any path that passes through DC must pass through another edge of η. In this example, DC-CQ-QM-MN is one such path.

edges of C_η. Every vertex u_t corresponds to a simplex σ_t in K. Now, adding the simplex boundaries $\sum_{t=2}^{\ell} \partial\sigma_i$ to ζ_0 gives rise to a cycle homologous to ζ_0 such that $\mathcal{C}_i$ has one vertex (and one edge) more and $\mathcal{C}_j$ has one vertex (and one edge) less. We repeat this process inductively until all edges are 'transported' from $\mathcal{C}_j$ to $\mathcal{C}_i$ and $\mathcal{C}_j$ becomes empty. That is, the new cycle ζ_1 we obtain is homologous to ζ_0 and has $k-1$ components. We denote the subgraph induced by intersection of ζ_1 and η by $C_\eta^{\zeta_1}$.

We apply the same procedure to $C_\eta^{\zeta_1}$ as above and get a a cycle $\zeta_2 \in [\zeta]$ whose induced subgraph $C_\eta^{\zeta_2}$ has $k-2$ components. Proceeding inductively, we finally obtain a cycle $\gamma = \zeta_{k-1} \in [\zeta]$ whose induced subgraph $C_\eta^{\zeta_{k-1}}$ is a connected subgraph of the dual graph. Moreover, by design, the total number of edges in every induced graph $C_\eta^{\zeta_i}$ for $i \in [0, k-1]$ is m. ◀

▶ **Lemma 15.** *The following are equivalent.*
(a.) A connected cocycle η is a feasible set for the input cycle ζ.
(b.) Every cycle in $[\zeta]$ intersects a connected cocycle η in an odd number of edges.
(c.) There exists a cycle in $[\zeta]$ that intersects a connected cocycle η in an odd number of edges.

Proof. (a.) $\implies$ (b.) Assume that there exists a cycle $\xi \in [\zeta]$ that intersects η in an even number of edges. The intersection of ξ with η induces a (possibly disconnected) subgraph of C_η, which we denote by C_η^ξ. Using Lemma 14, there exists another cycle γ homologous to ξ such that γ intersects η in the same number of edges, and the intersection of γ and η induces a connected graph in the dual graph, which we denote by C_η^γ. There are two cases:

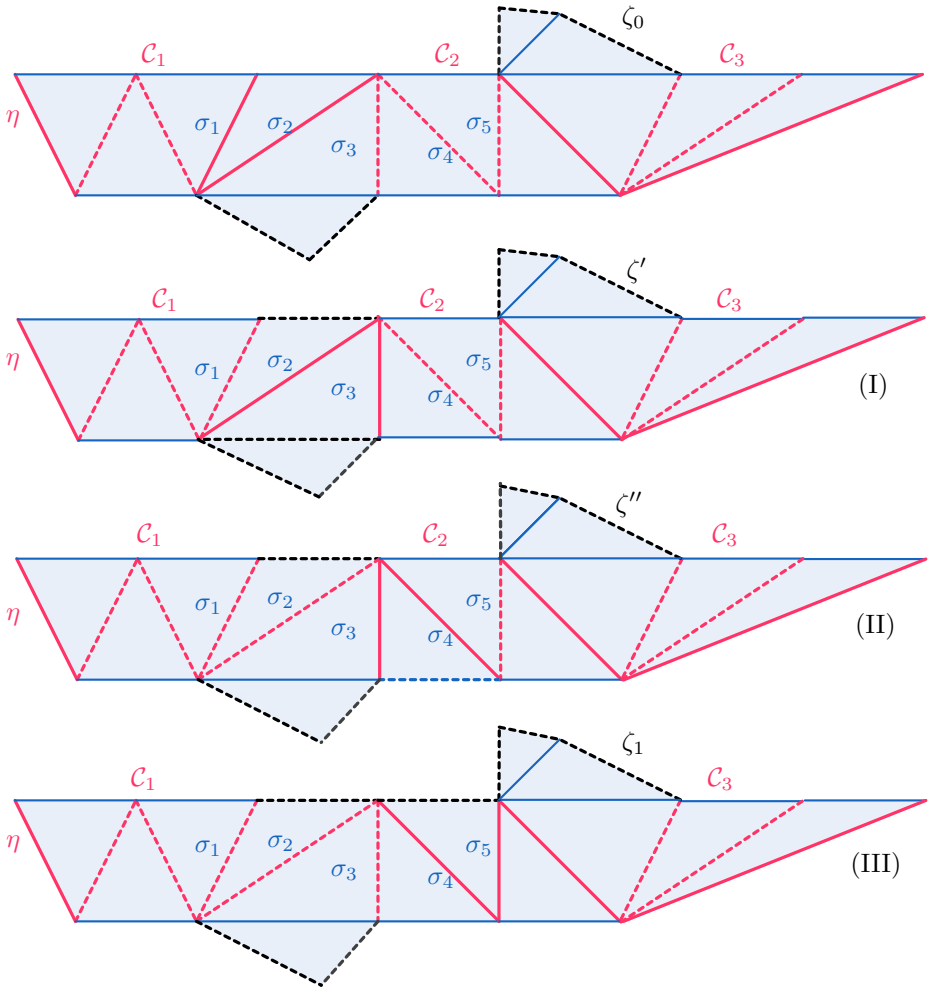


Figure 9 In this figure, we provide an illustrative example of the cycle modification technique from Lemma 14. As in Figure 6, the edges of a cycle that intersect the edges of the cocycle η are shown as pink-dotted edges and the edges of a cycle that do not intersect η are shown as black-dotted edges. The cocycle η is shown in pink. The intersection of ζ_0 and η induces a disconnected subgraph $\mathcal{C}_\eta^{\zeta_0}$ of the dual graph with several components, only three of which are shown, namely, $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{C}_3$. Here, $\mathcal{C}_1$ and $\mathcal{C}_2$ are neighbors, and $\mathcal{C}_2$ and $\mathcal{C}_3$ are neighbors. The path from $\mathcal{C}_1$ to $\mathcal{C}_2$ in D_K consists of simplices $\sigma_1, \sigma_2, \sigma_3$ (which are vertices in the dual graph). So we add $\partial\sigma_2 + \partial\sigma_3$ to ζ_0 to obtain ζ'. Next, we add $\partial\sigma_3 + \partial\sigma_4$ to ζ' to obtain ζ''. Finally, we add $\partial\sigma_4 + \partial\sigma_5$ to ζ'' to obtain ζ_1. The number of connected components of $\mathcal{C}_\eta^{\zeta_1}$ is one less than the number of connected components of $\mathcal{C}_\eta^{\zeta_0}$. This was achieved by 'transporting' edges in $\mathcal{C}_2$ to $\mathcal{C}_1$.

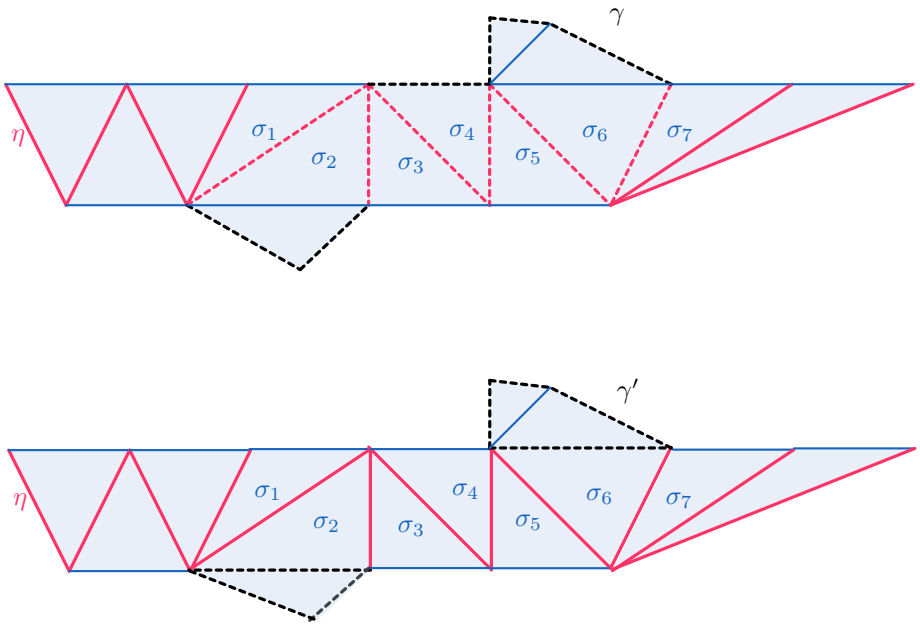


Figure 10 In this figure, we provide an illustrative example of the cycle modification scheme in ((**a.**) $\implies$ (**b.**)) from Lemma 15. As in Figures 6 and 9, the edges of a cycle that intersect the edges of the cocycle η are shown as pink-dotted edges and the edges of a cycle that do not intersect η are shown as black-dotted edges. The cocycle η is shown in pink. The intersection of γ and η induces a connected graph in the dual graph, namely C_η^γ. The number of edges in C_η^γ is even. Let $\gamma' = \gamma + \partial\sigma_2 + \partial\sigma_4 + \partial\sigma_6$. The cycle γ' does not intersect η.

Case 1 C_η^γ is not identical to C_η (the cycle graph induced by the entire cocycle η), and C_η^γ is a path graph $\mathcal{P}$ with vertices $u_1, u_2, \ldots, u_{m+1}$, where each vertex u_i is distinct and corresponds to a simplex σ_i in K.

Case 2 C_η^γ is the same as the entire cycle graph C_η with vertices $u_1, u_2, \ldots, u_m, u_{m+1}$, where $u_{m+1} = u_1$.

In either case, upon adding the simplex boundaries $\sum_{i=1}^{m/2} \partial\sigma_{2i}$ to γ, we obtain a cycle that has an empty intersection with η. That is, there exists a cycle in $[\zeta]$ which does not meet η in any of its edges. In other words, η is an infeasible set.

(b.) $\implies$ **(a.)** This is true by the definition of a feasible set.

(c.) $\implies$ **(b.)** This is the content of Lemma 10.

(b.) $\implies$ **(c.)** This is trivially true. ◀

This completes the proof.

Using Notation 3, Lemma 15 can be written as follows.

▶ **Lemma 16.** *The following are equivalent.*

(a.) A connected cocycle η is a feasible set for the input cycle ζ.

(b.) For a connected cocycle η, and any cycle $\zeta' \in [\zeta]$, $\eta(\zeta') = 1$.

(c.) For a connected cocycle η, there exists a cycle $\zeta' \in [\zeta]$ such that $\eta(\zeta') = 1$.

▶ **Lemma 17.** *A connected cocycle η is a feasible set if and only if a connected cocycle cohomologous to it is a feasible set.*

Proof. A cocycle η' cohomologous to η can be written as $\eta' = \eta + \delta(S)$, where S is a collection of vertices. Then, by linearity,

$$\eta'(\zeta) = \eta(\zeta) + \delta(S)(\zeta) = \eta(\zeta) + \sum_{v \in S} \delta(v)(\zeta).$$

$\delta(v)$ is a connected trivial cocycle. So, using Lemma 12 and $\neg(\textbf{(c.)} \implies \textbf{(a.)})$ in Lemma 16, $\delta(v)(\zeta)$ is 0 for every $v \in S$. Hence, $\eta'(\zeta) = 1$ if and only if $\eta(\zeta) = 1$. So, the claim follows from $(\textbf{(c.)} \implies \textbf{(a.)})$ in Lemma 16. ◀

Next, we prove an important generalization of Lemma 17.

▶ **Lemma 18.** *Let $k > 1$ be an integer. Let η_i for $i \in [k]$ be connected cocycles. On the one hand, if η_i for $i \in [k]$ are infeasible sets for the input cycle ζ, then any cocycle ϑ cohomologous to $\sum_{i=1}^{k} \eta_i$ is an infeasible set. On the other hand, if η_k is a feasible set, and η_i for $i \in [k-1]$ are infeasible sets for the input cycle ζ, then any cocycle ϑ cohomologous to $\sum_{i=1}^{k} \eta_i$ is a feasible set.*

Proof. A cocycle ϑ cohomologous to $\sum_{i=1}^{k} \eta_i$ can be written as $\sum_{i=1}^{k} \eta_i + \delta(S)$ where S is a collection of vertices. Then, by linearity,

$$\vartheta(\zeta) = \sum_{i=1}^{k} \eta_i(\zeta) + \delta(S)(\zeta) = \sum_{i=1}^{k} \eta_i(\zeta) + \sum_{v \in S} \delta(v)(\zeta).$$

Using Lemmas 12 and 16, $\delta(v)(\zeta) = 0$ for every $v \in S$, and $\eta_i(\zeta) = 0$ for $i \in [k-1]$. Hence, $\vartheta(\zeta) = 1$ if and only if $\eta_k(\zeta) = 1$. So, the claim follows from Lemma 16. ◀

▶ Remark 19 (Computing optimal (co)homology basis for surfaces). For simplicial complexes with n vertices, m edges and N simplices in total, we recall some of the known results from literature. For the special case when the input complex is an surface, Erickson and Whittlesey [22] devised a $O(N^2 \log N + gN^2 + g^3N)$-time algorithm for computing an optimal homology basis. Borradaile et al. [2] improved on this result by providing a $O((h+c)^3 n \log n + m)$-time algorithm for the same problem. Here c denotes the number of boundary components, and h denotes the genus of the surface. Dlotko [2] generalized the algorithm from [22] for computing an optimal cohomology basis for surfaces. For general complexes, Dey et al. [17, 18], Chen and Freedman [12], Busaryev et al. [7], and Rathod [34] provided progressively faster algorithms for computing an optimal homology basis.

Although we expect this to be fairly well known, for the sake of completeness, we describe an algorithm for computing minimum cohomology basis of a triangulated surface that uses the minimum homology basis algorithm as a subroutine.

▶ **Lemma 20.** *The minimum cohomology basis problem on surfaces can be solved in the same time as the minimum homology basis problem on surfaces.*

Proof. Let K be a surface with a weight function w on its edges. Let $\hat{\mathsf{K}}$ be the dual cell complex of K. Then, to every edge e of K there is a unique corresponding edge $\hat{e}$ in $\hat{\mathsf{K}}$. We now define a weight function on the edges of $\hat{\mathsf{K}}$ in the obvious way: $w(\hat{e}) = w(e)$. Let $\hat{\mathsf{K}}'$ be

the simplicial complex obtained from the stellar subdivision of each of the 2-cells of $\hat{\mathsf{K}}$. The weight function on edges of $\hat{\mathsf{K}}$ is extended to a weight function on edges of $\hat{\mathsf{K}}'$ by assigning weight ∞ to every newly added edge during the stellar subdivision. Such a complex $\hat{\mathsf{K}}'$ can be computed in linear time. It is easy to check that the cocycles of K are in one-to-one correspondence with the cycles of $\hat{\mathsf{K}}$, and the cycles of $\hat{\mathsf{K}}$ are in one-to-one correspondence with finite weight cycles of $\hat{\mathsf{K}}'$. Moreover, if η is a cocycle of K, and if $\hat{\eta}$ and $\hat{\eta}'$ are the corresponding cycles in $\hat{\mathsf{K}}$ and $\hat{\mathsf{K}}'$, respectively, then $w(\eta) = w(\hat{\eta}) = w(\hat{\eta}')$. Hence, computing a minimum homology basis for $\hat{\mathsf{K}}'$ gives a minimum cohomology basis for K. ◀

▶ **Theorem 21.** *Algorithm 1 provides a polynomial time algorithm for computing an optimal solution for* Topological Hitting Set *on surfaces.*

Proof. Let $\{\nu_i \mid i \in [m]\}$ be an optimal cohomology basis for K. Then, by Theorem 13, any optimal solution set is a cocycle. So, we can let k be the smallest integer for which a cocycle cohomologous to some cocycle in the span of $\{\nu_i \mid i \in [k]\}$ is a feasible solution set. Because the algorithm confirms that each ν_i, $i \in [k-1]$ is an infeasible set, by Lemma 18, any *connected* cocycle cohomologous to $\sum_{i=1}^{k-1} \nu_i$ is an infeasible set. On the other hand, since there exists a feasible set $\theta = \sum_{j_i} \nu_{j_i} + \nu_k + \beta$ where $j_i \in [k-1]$, and β is a coboundary, by Lemma 18, $\nu_k = \sum_{j_i} \nu_{j_i} + \beta + \theta$ is also a feasible set. Because $\{\nu_i \mid i \in [m]\}$ is an optimal cohomology basis, ν_k is, in fact, a minimal solution set.

From Lemma 20, we know that Step-1 of Algorithm 1 can be computed in polynomial time. Step-2 can be implemented by a simple sorting algorithm. Finally, Step-3 can be executed in linear time. ◀

▶ Remark 22. The algorithmic results in this section motivate severals questions: To what extent can this machinery be extended from surfaces to general complexes?

1. Are the optimal solutions sets for Topological Hitting Set nontrivial cocycles for general complexes? To the best of our knowledge, this question is open.
2. Can the optimal solution sets for Topological Hitting Set be computed efficiently for general complexes? We answer this question in the negative in Section 6.1 by showing that for general complexes Topological Hitting Set is **NP**-hard and **W[1]**-hard. Intriguingly, for the gadgets used in the reduction the optimal solution sets for Topological Hitting Set are cocycles! So they do not provide (a family of) counterexamples for the first question.
3. We believe that it should be possible to dualize the hardness results of Chen and Freedman [13] to show that computing an optimal cohomology basis for general complexes is **NP**-hard. So, in general, knowing that the optimal solutions sets are cocycles is not enough to guarantee tractability. One also needs an efficient algorithm for computing an optimal cohomology basis.

6 W[1]-hardness results

In this section, we obtain **W[1]**-hardness results for Topological Hitting Set and Boundary Nontrivialization with respect to the solution size k as the parameter via parameterized reductions from k-Multicolored Clique. We begin this section by recalling some common notions from graph theory.

A *k-clique* in a graph G is a complete subgraph of G with k vertices. Next, a *k-coloring* of a graph G is an assignment of one of k possible colors to every vertex of G (that is, a

vertex coloring) such that no two vertices that share an edge receive the same color. A graph G equipped with a k-coloring is called a *k-colored graph*. Then, a *multicolored k-clique* in a colored graph is a k-clique with a k-coloring. k-Multicolored Clique asks for the existence of a multicolored k-clique in a k-colored graph G. We remark that reducing from k-Multicolored Clique is a highly effective tool for showing $\mathbf{W[1]}$-hardness [23]. Formally, k-Multicolored Clique is defined as follows:

Problem 5 (k-Multicolored Clique).

Instance:	Given a graph $G = (V, E)$, and a vertex coloring $c : V \to [k]$.
Parameter:	k.
Question:	Does there exist a multicolored k-clique H in G?

▶ **Theorem 23** (Fellows et al. [23]). *k-Multicolored Clique is $\mathbf{W[1]}$-complete.*

6.1 W[1]-hardness for Topological Hitting Set

For $i \in [k]$, the subset of vertices of color i is denoted by V_i. Clearly, the vertex coloring c induces a partition on V:

$$V = \bigcup_{i=1}^{k} V_i, \quad \text{and} \quad V_i \bigcap V_j = \emptyset \text{ for all } i, j \in [k].$$

We now provide a parameterized reduction from k-Multicolored Clique to Topological Hitting Set. For $r = |V| - 1$, we define an $(r+1)$-dimensional complex $\mathsf{K}(G)$ associated to the given colored graph G as follows.

Vertices.

The set of vertices of $\mathsf{K}(G)$ contains the disjoint union of the vertices V in the graph G, the set of colors $[k]$, and an additional dummy vertex d. Altogether, we have $r + k + 2$ vertices in $\mathsf{K}(G)$ so far. In what follows, further vertices are added to $\mathsf{K}(G)$.

Simplices.

Below, we describe the simplices that constitute the complex $\mathsf{K}(G)$.

The cycle ζ. First, add the r-simplex V corresponding to vertex set V of the graph G. Next, add the r-simplices $(V \setminus \{u\}) \bigcup \{d\}$ for every $u \in V$. The collection of these $r+2$ simplices of dimension r forms a nontrivial r-cycle ζ.

The simplices in $\mathscr{X}_1$. $\mathscr{X}_1 = \{\sigma_i \mid i \in [k]\}$.

- For every color $i \in [k]$,
  - add an $(r+1)$-simplex $\sigma_i = V \bigcup \{i\}$.

▶ **Definition 24** (Admissible and undesirable facets of σ_i). *A facet $(V \setminus \{v\}) \bigcup \{i\}$ of σ_i is said to be* undesirable *if and only if $v \notin V_i$. All other facets of σ_i are deemed* admissible. *In particular, V is admissible.*

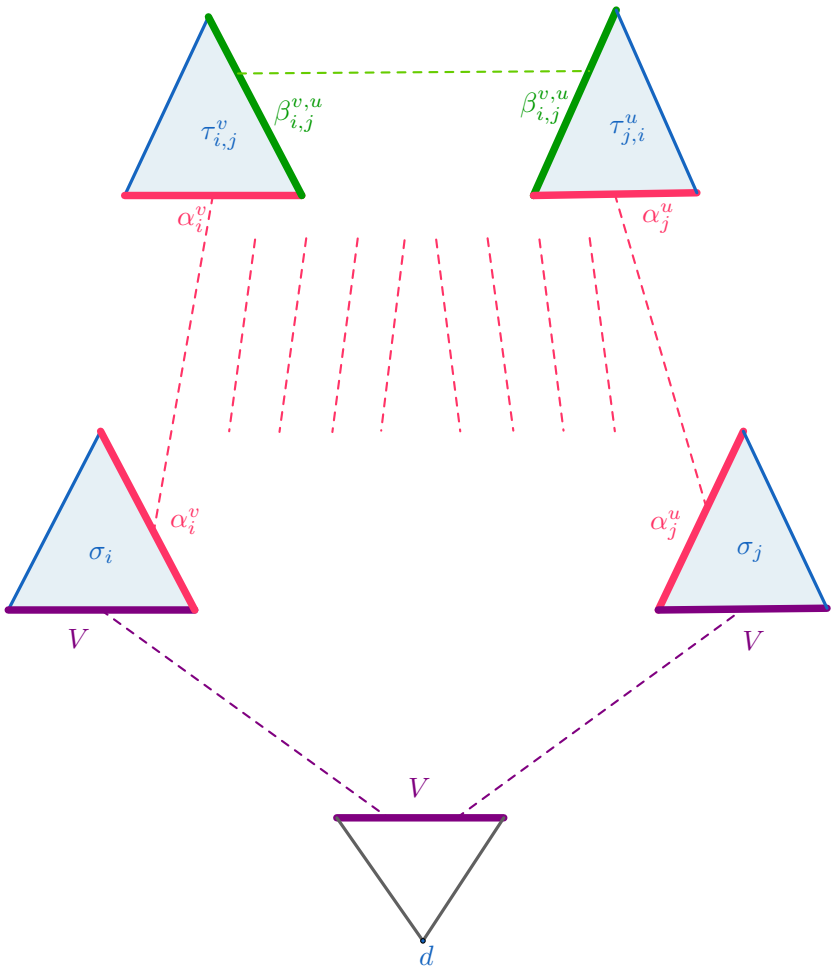


Figure 11 The figure shows some of the attachments in complex $\mathsf{K}(G)$. In particular, α^v_i is the common face of $\tau^v_{i,j}$ and σ_i, α^u_j is the common face of $\tau^u_{j,i}$ and σ_j, and $\beta^{v,u}_{i,j}$ is the common face of $\tau^v_{i,j}$ and $\tau^u_{j,i}$. The dashed lines indicate identifications along facets. The set of r-simplices supported by the vertices $V \bigcup \{d\}$ forms a nontrivial r-cycle in $\mathsf{K}(G)$.

The idea here is that including an admissible simplex of the form $(V \setminus \{v\}) \bigcup \{i\}$ in $\mathcal{S}$ is akin to picking the vertex v of color i for constructing the colorful clique. Including undesirable simplices in the solution will be made prohibitively expensive as the coloring specified by undesirable simplices is incompatible with the coloring c that the graph G comes equipped with.

The simplices in $\mathcal{X}_2$. $\mathcal{X}_2 = \left\{\tau^v_{i,j} \mid i \in [k],\ v \in V_i,\ j \in [k] \setminus \{i\}\right\}$.

- For every color $i \in [k]$,
  - for every vertex v in V_i and every color $j \in [k] \setminus \{i\}$,
    - add an $(r+1)$-simplex $\tau^v_{i,j} = (V \setminus \{v\}) \bigcup \{i, j\}$.

▶ **Definition 25** (Admissible and undesirable facets of $\tau^v_{i,j}$)**.** *The* admissible *facets of* $\tau^v_{i,j}$ *are:*
- $(V \setminus \{v, u\}) \bigcup \{i, j\}$ *with* $u \in V_j$ *and* $\{u, v\} \in E$*, and*
- $(V \setminus \{v\}) \bigcup \{i\}$*,*

A facet of $\tau^v_{i,j}$ *that is not admissible is* undesirable*.*

The intuition here is that picking an admissible facet of the form $(V \setminus \{v, u\}) \bigcup \{i, j\}$ is akin to picking the edge $\{u, v\}$ of color $\{i, j\}$ for constructing the colorful clique, whereas the admissible facet $(V \setminus \{v\}) \bigcup \{i\}$ is common with σ_i. Including undesirable simplices in the solution will be made prohibitively expensive (as explained later). Undesirable simplices of $\tau^v_{i,j}$ correspond either to coloring that is incompatible with c or with edges that are not even present in E.

Undesirable and inadmissible simplices. The undesirability of certain r-simplices is implemented in the gadget as follows: Let $m = n^3$. Then, to every undesirable r-simplex $\omega = \{v_1, v_2, \ldots, v_{r+1}\}$, associate m new vertices $\mathcal{U}^\omega = \{u^\omega_1, u^\omega_2, \ldots, u^\omega_m\}$. Now introduce m new $r+1$-simplices

$$\Upsilon^\omega = \{\mu_i(\omega) = \{v_1, v_2, \ldots, v_{r+1}, u^\omega_i\} \mid i \in [m]\}$$

that are cofacets of ω. See Figure 12 for an illustrative example.

▶ **Definition 26** (Set of inadmissible simplices associated to an undesirable simplex ω)**.** *The set of* r*-simplices in* $\{\{$*facets of* $\mu_i(\omega)\} \mid i \in [m]\}$ *is denoted by* $[\omega]$*. The simplices in the set* $[\omega]$ *are said to be* inadmissible*. In particular,* ω *itself is inadmissible.*

Further, note that the set of vertices in $\mathcal{U}^\omega$ and r-simplices in Υ^ω are unique to ω. As we observe later, introducing these new simplices makes inclusion of ω in the solution set prohibitively expensive. Denote by $\mathcal{Y}$ the set of all $r+1$-simplices added in this step.

This completes the construction of complex $\mathsf{K}(G)$. It is easy to check that the inadmissible and admissible simplices of $\mathsf{K}(G)$ partition the set of r-simplices of $\mathsf{K}(G)$.

▶ Notation 4. The admissible facets $(V \setminus \{v\}) \bigcup \{i\}$ and $(V \setminus \{v, u\}) \bigcup \{i, j\}$ are denoted by α^v_i and $\beta^{v,u}_{i,j}$, respectively. For every vertex $v \in V$ of color i, there is a facet α^v_i. For every edge $\{u, v\} \in E$, there is a facet $\beta^{v,u}_{i,j}$, where i is the color of v and j is the color of u.

▶ Remark 27 (Meaning of superscripts and subscripts of simplices). A simple mnemonic for remembering the meaning of the notation for simplices is as follows: the indices in the subscript are the included colors, and the vertices in the superscript indicate the vertices

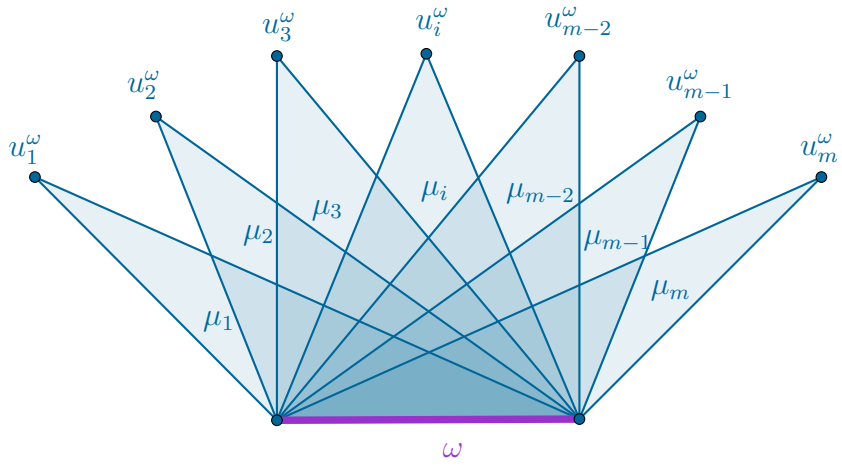


Figure 12 For every undesirable simplex $\omega = \{v_1, v_2, \ldots, v_{r+1}\}$ m new vertices $\mathcal{U}^\omega = \{u_1^\omega, u_2^\omega, \ldots, u_m^\omega\}$ are added to $\mathsf{K}(G)$. Moreover, m new $r+1$-simplices $\Upsilon^\omega = \{\mu_i(\omega) = \{v_1, v_2, \ldots, v_{r+1}, u_i^\omega\} \mid i \in [m]\}$, where $\omega \prec \mu_i(\omega)$ for every $i \in [m]$ are also added to $\mathsf{K}(G)$. The facets of $\mu_i(\omega)$ for every $i \in [m]$ are the inadmissible simplices associated to ω and denoted by $[\omega]$.

excluded from V. For instance, $\beta_{i,j}^{v,u}$ is the full simplex on the vertex set $(V \setminus \{v, u\}) \bigcup \{i, j\}$. In this case, colors i and j are included and vertices u and v are excluded. The same notational rule applies for α_i^v , σ_i and $\tau_{i,j}^v$.

▶ Remark 28 (Correspondence between colors and vertices in α_i^v $\beta_{i,j}^{v,u}$ and $\tau_{i,j}^v$). In our notation, the first color corresponds to the first vertex, the second color to the second vertex, and so on. For instance,

- In α_i^v, vertex v is of color i.
- In $\beta_{i,j}^{v,u}$, v is of color i and u is of color j.
- In $\tau_{i,j}^v$, v is of color i and the vertex associated to color j is not specified. It is, in fact, chosen through a facet $\beta_{i,j}^{v,u} \prec \tau_{i,j}^v$.

Choice of parameter.

Let $(k + \binom{k}{2} + 1 = \binom{k+1}{2} + 1)$ be the parameter for Topological Hitting Set on the complex $\mathsf{K}(G)$.

▶ Remark 29 (Size of $\mathsf{K}(G)$). We note that every subset of vertices of G is a simplex in $\mathsf{K}(G)$. However, $\mathsf{K}(G)$ is represented implicitly, and the simplices of dimensions other than r and $r+1$ are not used in the reduction. Thus, although $\mathsf{K}(G)$ as a simplicial complex is exponential in the size of G, the reduction itself is polynomial in the size of G because the number of r and $r+1$ dimensional simplices of $\mathsf{K}(G)$ are polynomial in size of G, even after inadmissible simplices.

▶ **Lemma 30.** *If there exists a multicolored k-clique $H = (V_H, E_H)$ of G, then there exists a topological hitting set $\mathcal{S}$ for ζ consisting of $\binom{k+1}{2} + 1$ r-simplices.*

Proof. We construct a set $\mathcal{S}$ of r-simplices that mimics the graphical structure of H as follows:

$$\mathcal{S}_\alpha = \left\{\alpha_i^v \mid v \in V_i \bigcap V_H\right\}$$

$$\mathcal{S}_\beta = \left\{\beta^{v,u}_{i,j} \mid v \in V_i,\, u \in V_j,\, \{i,j\} \in E_H\right\}$$

First, set $\mathcal{S} = \{V\} \bigcup \mathcal{S}_\alpha \bigcup \mathcal{S}_\beta$. Next, note that every cycle $\zeta' \in [\zeta]$ can be expressed as

$$\zeta' = \zeta + \sum_{\nu_i \in \mathcal{X}'} \partial \nu_i + \sum_{\mu_j \in \mathcal{Y}'} \partial \mu_j$$

for some $\mathcal{X}' \subset \mathcal{X}_1 \bigcup \mathcal{X}_2$ and $\mathcal{Y}' \subset \mathcal{Y}$. Let $\mathcal{X}'_1 = \mathcal{X}' \bigcap \mathcal{X}_1$, and $\mathcal{X}'_2 = \mathcal{X}' \bigcap \mathcal{X}_2$. Now, we claim that removing $\mathcal{S}$ from $\mathsf{K}(G)$ destroys every cycle $\zeta' \in [\zeta]$. We show this by establishing that the coefficient in every $\zeta' \in [\zeta]$ of at least one of the simplicies of $\mathcal{S}$ is 1. In other words, $\mathcal{S} \bigcap \zeta \neq \emptyset$ for every $\zeta' \in [\zeta]$.

Case 1: $\mathcal{X}' = \emptyset$. Then, $V \in \mathcal{S}$ has coefficient 1 in cycle ζ'. This is because simplices in $\mathcal{Y}$ are not incident on V, and $V \in \zeta$.

Case 2: $\mathcal{X}'_1 \neq \emptyset, \mathcal{X}'_2 = \emptyset$. Then, the cycle ζ' can be written as

$$\zeta' = \zeta + \sum_{\sigma_j \in \mathcal{X}'_1} \partial \sigma_j + \sum_{\mu_\ell \in \mathcal{Y}'} \partial \mu_\ell$$

Then, every $\alpha^v_j \in \mathcal{S}$ for $\sigma_j \in \mathcal{X}'_1$ and $v \in V_H \bigcap V_j$ has coefficient 1 in cycle ζ'. This is because $\alpha^v_j \in \partial \sigma_j$ for every $\sigma_j \in \mathcal{X}'_1$, but $\alpha^v_j \notin \zeta$ and $\alpha^v_j \notin \partial \mu_\ell$ for any $\mu_\ell \in \mathcal{Y}'$.

Case 3: $\mathcal{X}'_1 = \emptyset, \mathcal{X}'_2 \neq \emptyset$. This case is identical to Case-1, because $V \in \mathcal{S}$ has coefficient 1 in cycle ζ'.

Case 4: $\mathcal{X}'_1 \neq \emptyset,\ \mathcal{X}'_2 \neq \emptyset$. If every simplex $\tau^v_{p,q} \in \mathcal{X}'_2$ is such that $v \in V_p \setminus V_H$, then this case becomes identical to Case 2. So we will assume without loss of generality that the set $\mathcal{X}''_2 = \left\{\tau^v_{p,q} \mid p,q \in [k], v \in V_p \bigcap V_H, \tau^v_{p,q} \in \mathcal{X}'_2\right\}$ is non-empty. For some $\{u,v\} \in E_H$ and $u \in V_q$, $v \in V_p$, if $\tau^v_{p,q} \in \mathcal{X}''_2$ and $\tau^u_{q,p} \notin \mathcal{X}''_2$, then the coefficient of $\beta^{v,u}_{p,q} \in \mathcal{S}$ in ζ' is 1 because the only two $(r+1)$-simplices incident on $\beta^{v,u}_{p,q}$ are $\tau^v_{p,q}$ and $\tau^u_{q,p}$. So, without loss of generality assume that the symmetric simplex $\tau^u_{q,p}$ is also in $\mathcal{X}''_2$. In other words, $|\mathcal{X}''_2|$ is even. Note that for every $\tau^v_{p,q} \in \mathcal{X}''_2$, exactly one facet of $\tau^v_{p,q}$ lies in $\mathcal{S}_\alpha$, namely α^v_p. Hence the cardinality of the multiset $\mathcal{T} = \left\{\partial \tau^v_{i,j} \bigcap \mathcal{S}_\alpha \mid \tau^v_{i,j} \in \mathcal{X}''_2\right\}$ is even. Let $\sigma_i \in \mathcal{E}$ if and only if the cardinality of the set $\left\{\tau^v_{i,j} \in \mathcal{X}''_2 \mid \alpha^v_i \in \partial \tau^v_{i,j} \bigcap \mathcal{S}_\alpha\right\}$ is even, and $\sigma_i \in \mathcal{O}$ if and only if the cardinality of the set $\left\{\tau^v_{i,j} \in \mathcal{X}''_2 \mid \alpha^v_i \in \partial \tau^v_{i,j} \bigcap \mathcal{S}_\alpha\right\}$ is odd. It is easy to check that $\mathcal{X}'_1 \subseteq \mathcal{O} \bigcup \mathcal{E}$. Note that since $|\mathcal{T}| = |\mathcal{O}| + |\mathcal{E}|$, $|\mathcal{O}|$ must be even.

Now, if $\sigma_i \in \mathcal{E} \bigcap \mathcal{X}'_1$, then the coefficient of $\alpha^v_i \in \mathcal{S}$ in ζ' is 1 because the only $(r+1)$-simplices incident on α^v_i are $\left\{\tau^v_{i,j} \in \mathcal{X}''_2 \mid \alpha^v_i \in \partial \tau^v_{i,j}\right\} \bigcup \{\sigma_i\}$, and $|\left\{\tau^v_{i,j} \in \mathcal{X}''_2 \mid \alpha^v_i \in \partial \tau^v_{i,j}\right\}|$ is even when $\sigma_i \in \mathcal{E}$. So, without loss of generality assume that $\mathcal{E} \bigcap \mathcal{X}'_1$ is empty. That is, we assume that $\mathcal{X}'_1 \subseteq \mathcal{O}$. But if, $\sigma \in \mathcal{O} \setminus \mathcal{X}'_1$, then the coefficient of $\alpha^v_i \in \mathcal{S}_\alpha$ in ζ' is 1 because in that case the only $(r+1)$-simplices incident on α^v_i will be $\left\{\tau^v_{i,j} \in \mathcal{X}''_2 \mid \alpha^v_i \in \partial \tau^v_{i,j}\right\}$ which has odd cardinality. So, we assume that $\mathcal{O} = \mathcal{X}'_1$. But if $\mathcal{O} = \mathcal{X}'_1$, then $V \in \mathcal{S}$ has coefficient 1 in ζ' because $\mathcal{O}$ is even and $V \in \zeta$. This completes the proof. Please see Figure 13 for the final part of the argument. ◀

The next few lemmas provide a method to extract a multi-colored k-clique from G given a solution set $\mathcal{R}$ for Topological Hitting Set on $\mathsf{K}(G)$.

▶ **Lemma 31.** *If there exists a cycle $\zeta' \in [\zeta]$ such that only the inadmissible simplices of $\mathcal{R}$ have coefficient 1 in ζ', then the size of $\mathcal{R}$ is at least m.*

Proof. We consider two cases.

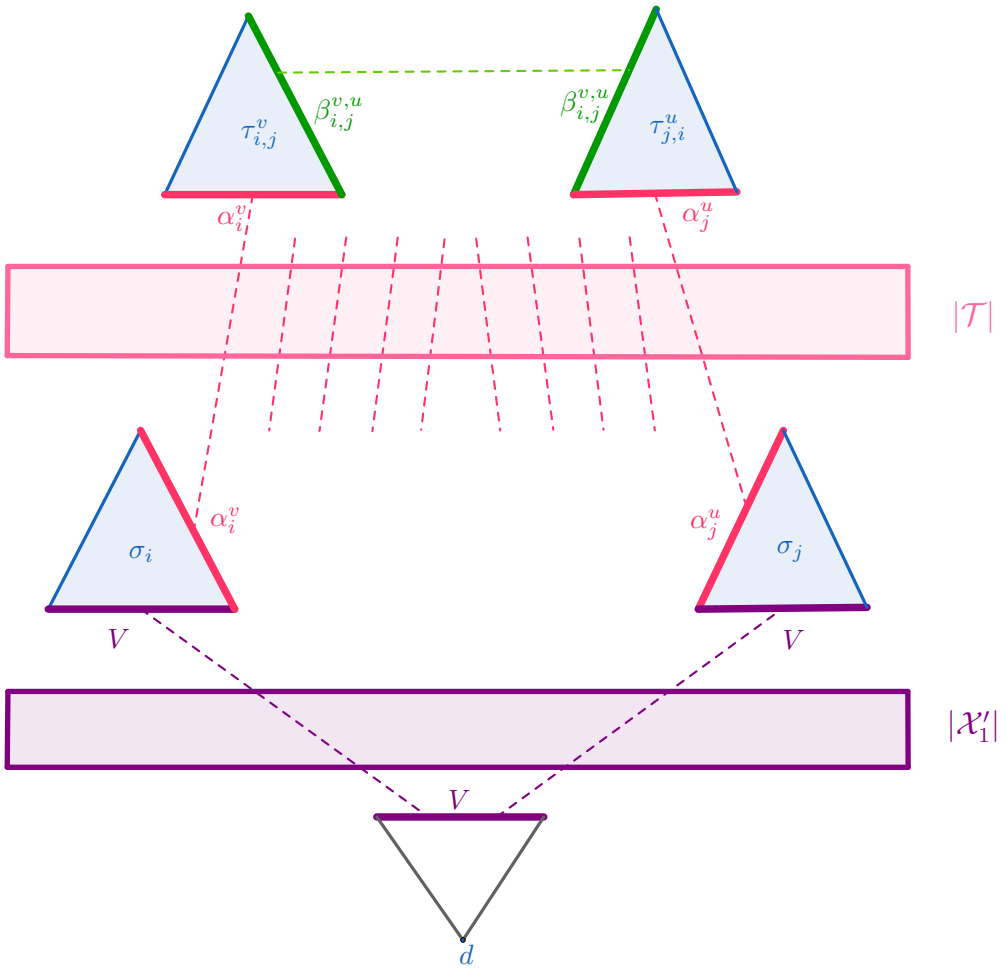


Figure 13 Like in Figure 13, the dashed lines indicate identifications along facets. Additionally, in this figure, $\tau^v_{i,j}$ and $\tau^u_{j,i}$ belong to $\mathcal{X}''_2$, σ_i and σ_j belong to $\mathcal{X}'_1$, and α^v_i and α^u_j belong to $\mathcal{S}_\alpha$. $\mathcal{T}$ accounts for all the incidences of the boundaries of simplices in $\mathcal{X}''_2$ on simplices in $\mathcal{S}_\alpha$. The final part of the argument in Case 4 of Lemma 30 is depicted here. In Case 4 of Lemma 30, we use the fact that $|\mathcal{T}|$ is even. Also, if $|\mathcal{T}|$ is even and if the coefficient of all the simplices of $\mathcal{S} \setminus \{V\}$ have coefficient 0 in some cycle $\zeta' \in [\zeta]$, then $\mathcal{X}' = \mathcal{O}$, and $\mathcal{O}$ has even cardinality. But if this is so, then the coefficient of V in ζ' is $(1 + |\mathcal{O}|) \bmod 2 = 1$.

Case 1: A cycle ζ' has a unique inadmissible simplex ν with coefficient 1 in $\mathcal{R}$.

Suppose that $\nu = \{v_1, v_2, \ldots, v_{r+1}\}$ is, in fact, an undesirable simplex. Assume that ν is the unique simplex in $\mathcal{R}$ with coefficient 1 in ζ'. Let $\mu_i = \{v_1, v_2, \ldots, v_{r+1}, u_i^\nu\}$, $i \in [m]$ be the inadmissible simplices in $[\nu]$. Then, a simplex in $\mathcal{R}$ will have coefficient 1 in the cycle $\zeta_i' = \zeta' + \partial\mu_i$ only if one of the simplices in $\partial\mu_i \setminus \{\nu\}$ for every $i \in [m]$ belongs to $\mathcal{R}$. Since $\partial\mu_i \setminus \{\nu\}$ for $i \in [m]$ are disjoint sets, the size of $\mathcal{R}$ is at least m.

Next, suppose $\nu = \{v_1, v_2, \ldots, v_{r+1}\} \in [\omega]$, where ω is a undesirable simplex, and $\nu \neq \omega$. Then, ω is a facet of $\mu_j(\omega)$ for some $\mu_j(\omega) = \{v_1, v_2, \ldots, v_{r+1}, u_j^\nu\}$, $j \in [m]$. Since $\zeta' \in [\zeta]$ is a *cycle*, all simplices in $\partial\mu_j(\omega)$ must have coefficient 1 in ζ'. Then, a simplex in $\mathcal{R}$ will have coefficient 1 in each of the cycles $\zeta_i' = \zeta' + \partial\mu_i(\omega) + \partial\mu_j(\omega)$, $i \in [m] \setminus \{j\}$ only if one of the simplices in $\partial\mu_i(\omega) \setminus \{\nu\}$ for each i belongs to $\mathcal{R}$, where $\mu_i(\omega) = \{v_1, v_2, \ldots, v_{r+1}, u_i^\nu\}$. Also, the cycle $\zeta'' = \zeta + \partial\mu_j(\omega)$ has a simplex in $\mathcal{R}$ with coefficient 1 only if $\omega \in \mathcal{R}$. Hence, in both cases, the size of $\mathcal{R}$ is at least m.

Case 2: A cycle ζ' has multiple inadmissible simplices with coefficient 1 in $\mathcal{R}$.

More generally, suppose there exist more than one inadmissible simplices in $\mathcal{R}$ with coefficient 1 in ζ', for some cycle $\zeta' \in [\zeta]$. For an undesirable simplex ω, we say that $[\omega]$ *belongs to* ζ' if it there exists a simplex in $[\omega]$ that has coefficient 1 in ζ'. Let J be an indexing set for the undesirable simplices of $\mathsf{K}(G)$ whose classes belong to ζ'. That is, for all $j \in J$, $[\omega^j]$ belongs to ζ'. Define the sets P and Q as follows.

$$P = \{\mu_i(\omega^j) \mid j \in J,\ i \in [m],\ \text{ a simplex in } \partial\mu_i(\omega^j) \setminus \{\omega^j\} \text{ belongs to } \zeta' \text{ and } \mathcal{R}\}.$$

and

$$Q = \{\mu_k(\omega^j) \mid j \in J,\ k \in [m],\ \mu_k(\omega^j) \notin P\}.$$

Define $\zeta_{\overline{Q}}$ as follows.

$$\zeta_{\overline{Q}} = \zeta' + \sum_{\substack{\mu_i(\omega^j) \in P, \\ j \in J}} \partial\mu_i(\omega^j) + \sum_{\substack{\mu_k(\omega^j) \in \overline{Q}, \\ j \in J}} \partial\mu_k(\omega^j).$$

for all $\overline{Q} \subset Q$.

Then, a simplex in $\mathcal{R}$ will have coefficient 1 in each of the cycles $\zeta_{\overline{Q}}$ if and only if one of the r-simplices in $\partial\mu_k(\omega^j)$ for every $\mu_k(\omega^j) \in Q$ belongs to $\mathcal{R}$. Clearly, $|P + Q| \geq m$, proving the claim. ◀

▶ **Lemma 32.** *Let $\mathcal{R}$ be a solution set for* Topological Hitting Set *on complex $\mathsf{K}(G)$ such that $|\mathcal{R}| \leq \binom{k+1}{2} + 1$. Then,*

(1.) $V \in \mathcal{R}$.
(2.) For every σ_i, there is at least one simplex α_i^v (with $v \in V_i$) that is included in $\mathcal{R}$.
(3.) For every unordered pair (i, j), where $i, j \in [k]$, there exists a simplex $\beta_{i,j}^{v,u}$ for some v, u that is included in $\mathcal{R}$.
(4.) $|\mathcal{R}| = |A_\mathcal{R}| = \binom{k+1}{2} + 1$, where $A_\mathcal{R}$ denotes the set of admissible simplices of $\mathcal{R}$.

Proof. If $\zeta' \in [\zeta]$ is such that $\zeta' \bigcap A_\mathcal{R} = \emptyset$, then we are forced to include some simplex $\omega \in \zeta'$ in $\mathcal{R}$ such that ω is inadmissible. In that case, Lemma 31 applies and we are forced to include at least m simplices. But, if we include more than n^3 facets in $\mathcal{R}$, we exceed the budget of $\binom{k+1}{2} + 1$. So, going forward, we assume that at least one simplex in $A_\mathcal{R}$ has coefficient 1 for every $\zeta' \in [\zeta]$.

Note that if at least one simplex in $A_\mathcal{R}$ has coefficient 1 for every cycle in $[\zeta]$, then we do not need inadmissible simplices in $\mathcal{R}$. We now prove the four statements of the lemma.

(1.) Since V is the only admissible facet of ζ, it must be included.
(2.) Since $\zeta' = \zeta + \partial\sigma_i$ is a cycle homologous to ζ, and the coefficient of V in ζ' is zero, the admissible simplices in ζ' are given by the set $\{\alpha_i^v \mid v \in V_i\}$. One of the simplices in this set must be included in $\mathcal{R}$ for each i, for $\mathcal{R}$ to be a solution set.
(3.) Note that for a fixed i and $j \in [k] \setminus \{i\}$ unless some admissible facet $\beta_{i,j}^{v,u}$ for some v, u is included in $\mathcal{R}$, the coefficient of all admissible simplices in $\zeta' = \zeta + \partial\sigma_i + \sum_{v \in V_i} \partial\tau_{i,j}^v$ will be zero. The claim follows.
(4.) This follows from the first three parts of the lemma. By (1.) we must include V in $\mathcal{R}$, by (2.) we must include at least k α faces in $\mathcal{R}$, and by (3.), we must include at least $\binom{k}{2}$ faces n $\mathcal{R}$. Since $\binom{k}{2} + k + 1 = \binom{k+1}{2} + 1$, the claim follows. ◀

▶ **Lemma 33.** *If $|\mathcal{R}| = \binom{k+1}{2} + 1$, then one can obtain a k-clique H of G from $\mathcal{R}$.*

Proof. If $|\mathcal{R}| = \binom{k+1}{2} + 1$, then using (4.), $|\mathcal{R}| = |A_{\mathcal{R}}|$. Therefore, $\mathcal{R}$ consists entirely of admissible simplices. We now provide four conditions that characterize a solution of size $\binom{k+1}{2} + 1$.

As noted in Section 6.1 (1.), V is part of any solution set. Using Section 6.1 (2.), for $\mathcal{R}$ to be a solution set, at least one facet (other than V) of σ_i for every i must belong to $\mathcal{R}$.

Condition 1. For every $i \in [k]$, the *only* facet of σ_i (other than V) that belongs $\mathcal{R}$ is an admissible simplex α_i^v, for some $v \in V_i$.

Now, α_i^v is incident on $k-1$ simplices, namely, $\tau_{i,j}^v$ for all $j \neq i$. Using Section 6.1 (3.), for $\mathcal{R}$ to be a solution, we must include in $\mathcal{R}$ at least one admissible facet $\beta_{i,j}^{v,u}$ of $\tau_{i,j}^v$ (for all $j \neq i$).

Condition 2. For every $v \in V_i$ such that $\alpha_i^v \in \mathcal{R}$, and every $j \in [k] \setminus \{i\}$, the *only* facet of $\tau_{i,j}^v$ (other than α_i^v) that belongs $\mathcal{R}$ is an admissible simplex $\beta_{i,j}^{v,u}$, for some $u \in V_j$.

Note that $\beta_{i,j}^{v,u}$ is also incident on both $\tau_{j,i}^u$ and $\tau_{j,i}^u$. Then, since there must exist an admissible simplex in $\mathcal{R}$ with coefficient 1 in $\zeta' = \zeta + \partial\sigma_i + \partial\tau_{i,j}^v + \partial\tau_{j,i}^u$, it is necessary that at least one facet of $\tau_{i,j}^u$ other than $\beta_{i,j}^{v,u}$ is included in $\mathcal{R}$. That is, α_j^u must be included in $\mathcal{R}$.

Repeating the same argument for every σ_i and every $\tau_{i,j}^v$, it is easy to check that the only way to construct $\mathcal{R}$ without exceeding the budget of $\binom{k+1}{2} + 1$ is by making these choices consistent. Thus, we obtain two additional conditions.

Condition 3. for every i and v such that $\alpha_i^v \in \mathcal{R}$, and every $j \in [k] \setminus \{i\}$, if α_i^v is in $\mathcal{R}$, and $\beta_{i,j}^{v,u}$ is in $\mathcal{R}$, then α_j^u is in $\mathcal{R}$.

Condition 4. for every i and v if $\alpha_i^v \notin \mathcal{R}$ (from choices made for σ_i's in Condition 1.), then for every $j \in [k] \setminus \{i\}$, the facet $\beta_{i,j}^{u,v}$ of $\tau_{i,j}^v$ in not included in $\mathcal{R}$.

The fact that such a set $\mathcal{R}$ is indeed a solution set follows the same argument as in Claim 30. It is clear that such a solution set $\mathcal{R}$ satisfies Conditions 1-4 if and only if $|\mathcal{R}| = \binom{k+1}{2} + 1$. If any of the conditions are not satisfied, then either we are forced to choose more than one vertices per color, or we have that the choice of vertices u, v in $\beta_{i,j}^{v,u}$ that is included in $\mathcal{R}$ as per Section 6.1 (3.) for pairs (i, j) and (j, i) is inconsistent. In both cases, $|\mathcal{R}| \geq \binom{k+1}{2} + 1$.

Finally, the graph H is constructed from $\mathcal{R}$ by first including one vertex v per color i for every $\alpha_i^v \in \mathcal{R}$, and the edges $\{u, v\}$ for every simplex $\beta_{i,j}^{v,u} \in \mathcal{R}$. ◀

Lemma 30 and Lemma 33 together provide a parameterized reduction from k-Multicolored Clique to Topological Hitting Set. Using Theorem 23, we obtain the following result.

▶ **Theorem 34.** Topological Hitting Set *is* $\mathbf{W}[\mathbf{1}]$*-hard.*

6.2 W[1]-hardness for Boundary Nontrivialization

Next, we provide a parameterized reduction from k-Multicolored Clique to Boundary Nontrivialization. This reduction bears some similarities with reduction from k-Multicolored Clique to Topological Hitting Set. So towards the end, we skip some of the details that are common to both the reductions.

Recalling some notation from Section 6.1

In Section 6.1, given a k-colored graph $G = (V, E)$, we defined two sets of $(r+1)$-simplices, namely, $\mathcal{X}_1$ and $\mathcal{X}_2$.

$$\mathcal{X}_1 = \{\sigma_i \mid i \in [k]\},$$

where $\sigma_i = V \bigcup \{i\}$, and

$$\mathcal{X}_2 = \left\{\tau^v_{i,j} \mid i \in [k],\ v \in V_i,\ j \in [k] \setminus \{i\}\right\},$$

where $\tau^v_{i,j} = (V \setminus \{v\}) \bigcup \{i, j\}$.
Furthermore, recall from Section 6.1, that for every $i \in [k]$, and $v \in V_i$, then

$$\alpha^v_i = (V \setminus \{v\}) \bigcup \{i\}$$

is an *admissible* facet of σ_i and $\tau^v_{i,j}$ respectively.
Also, for every $i \in [k]$, $j \in [k] \setminus \{i\}$, $v \in V_i$, $u \in V_j$, and $\{u, v\} \in E$,

$$\beta^{v,u}_{i,j} = (V \setminus \{v, u\}) \bigcup \{i, j\}$$

is an *admissible* facet of $\tau^v_{i,j}$ and $\tau^u_{j,i}$ respectively.

Overview of the reduction for Boundary Nontrivialization

In this section, given a k-colored graph $G = (V, E)$, an r-dimensional complex $\mathsf{L}(G)$ is constructed, where $r = |V| - 1$. Here, we provide an overview of the construction.

Let $\sigma_i = V \bigcup \{i\}$ as in Section 6.1, and let $\hat{\partial}\sigma_i$ be the r-complex $\partial\sigma_i \setminus \{V\}$. The complex $\hat{\Delta}\sigma_i$ is formed from $\hat{\partial}\sigma_i$ by the so-called S-subdivision of some of the faces of $\hat{\partial}\sigma_i$. The S-subdivision of a simplex is described in Section 6.2.1. The construction of $\hat{\Delta}\sigma_i$ from $\hat{\partial}\sigma_i$ is described in Algorithm 3. The lexicographically highest simplex of an S-subdivided face of $\hat{\partial}\sigma_i$ is a *distinguished simplex* in $\hat{\Delta}\sigma_i$. The simplices in $\hat{\partial}\sigma_i$ that are not distinguished are called *undesirable*. We wish to exclude undesirable simplices from solutions of small size. To implement the undesirability of simplices, we add further simplices to $\hat{\Delta}\sigma_i$. The newly added simplices and the undesirable simplices are together called *inadmissible simplices* of $\hat{\Delta}\sigma_i$. That completes the high-level description of $\hat{\Delta}\sigma_i$. Next, a subcomplex $\mathcal{Z}_1$ is built out of the union of subcomplexes $\hat{\Delta}\sigma_i$. That is,

$$\mathcal{Z}_1 = \bigcup_{i \in [k]} \hat{\Delta}\sigma_i.$$

Let $T = |V| \cdot (k-1)$. It is easy to check that $|\mathcal{X}_2| = T$. Let $t \in [T]$ be an indexing variable such that there is a unique t that corresponds to a triple (i, j, v), where $i \in [k]$, $j \in [k] \setminus \{i\}$ and $v \in V_i$. Now, for every $t \in [T]$, add a *new* set of vertices V_t. Here the vertex set V_t is in one-to-one correspondence with the vertex set $(V \setminus \{v\}) \bigcup \{i, j\}$. Let $\tau^v_{i,j}$ be the full simplex on the vertex set V_t. The complex $\Delta\tau^v_{i,j}$ is formed from $\partial\tau^v_{i,j}$ following an S-subdivision of

some of the faces of $\partial\tau^v_{i,j}$. The construction of $\Delta\tau^v_{i,j}$ from $\partial\tau^v_{i,j}$ is described in Algorithm 4. The *distinguished, undesirable,* and *inadmissible* simplices of $\Delta\tau^v_{i,j}$ are built in a manner analogous to the distinguished, undesirable and inadmissible simplices of $\hat{\Delta}\sigma_i$. For further details, please refer to Section 6.2.2. Next, a subcomplex $\mathcal{Z}_2$ is built out of the union of subcomplexes $\Delta\tau^v_{i,j}$. That is,

$$\mathcal{Z}_2 = \bigcup_{\substack{i\in[k]\\ v\in V_i}} \bigcup_{j\in[k]\setminus\{i\}} \Delta\tau^v_{i,j}.$$

The complex $\mathsf{L}(G)$ is obtained from G by identifying the distinguished faces of $\mathcal{Z}_1 \bigcup \mathcal{Z}_2$ as per the procedure described in Algorithm 5. The distinguished faces upon identifications are called the *admissible simplices* of $\mathsf{L}(G)$. In what should remind the reader of the notation used in Section 6.1, for every $i \in [k]$, and $v \in V_i$, there is an admissible simplex denoted by α^v_i that belongs to $\mathsf{L}(G)$. And for every $i \in [k]$, $j \in [k] \setminus \{i\}$ with $v \in V_i$, $u \in V_j$ and $\{u, v\} \in E$, there is an admissible simplex $\beta^{v,u}_{i,j}$ that belongs to $\mathsf{L}(G)$. The admissible simplices of $\mathsf{L}(G)$ encode the connectivity and coloring information of G. Analogous to the construction of the complex $\mathsf{K}(G)$ described in Section 6.1, in complex $\mathsf{L}(G)$, α^v_i is an admissible r-simplex belonging to the simplicial manifolds $\hat{\Delta}\sigma_i$ and $\Delta\tau^v_{i,j}$, where $\hat{\Delta}\sigma_i$ and $\Delta\tau^v_{i,j}$ are subcomplexes of $\mathsf{L}(G)$. Also, $\beta^{v,u}_{i,j}$ of is an admissible r-simplex belonging to the simplicial manifolds $\Delta\tau^v_{i,j}$ and $\Delta\tau^u_{j,i}$. We ask the reader to compare Figures 11 and 16.

Finally, Proposition 50 and Lemma 60 combine to show that the k-multicolored cliques of G are in one-to-one correspondence with $\binom{k+1}{2}$-sized solutions for Boundary Nontrivialization with $\mathsf{L}(G)$ as the instance. In fact, the reduction is a parameterized reduction that establishes the **W[1]**-hardness of Boundary Nontrivialization as a consequence.

▶ Remark 35. Note that we use the notation $\hat{\partial}\sigma_i$ and $\hat{\Delta}\sigma_i$ for the complexes associated to σ_i, and $\partial\tau^v_{i,j}$ and $\Delta\tau^v_{i,j}$ for the complexes associated to $\tau^v_{i,j}$. This disparity in notation (that is the use of $\hat{}$ for σ_i) is to remind the reader that in the case of $\hat{\partial}\sigma_i$, a face is deleted from simplex boundary of σ_i, whereas in the case of $\partial\tau^v_{i,j}$, the full simplex boundary is used.

6.2.1 S-subdivisions of simplices

Next, we recall a lemma from Munkres [33, Lemma 3.2] that will be used to provide a guarantee that the complex described in Section 6.2.2 is, in fact, a simplicial complex.

▶ **Lemma 36** (Munkres, [33, Lemma 3.2]). *Let $\mathcal{L}$ be a finite set of labels. Let K be a simplicial complex defined on a set of vertices V. Also, let $f : V \to \mathcal{L}$ be a surjective map associating to each vertex of K a label from $\mathcal{L}$. The labeling f extends to a simplicial map $g : \mathsf{K} \to \mathsf{K}_f$ where K_f has vertex set V and is obtained from K by identifying vertices with the same label.*

If for all pairs $v, w \in V$, $f(v) = f(w)$ implies that their stars $\mathrm{star}_{\mathsf{K}}(v)$ and $\mathrm{star}_{\mathsf{K}}(w)$ are vertex disjoint, then, for all faces $\eta, \omega \in \mathsf{K}$ we have that

- *η and $g(\eta)$ have the same dimension, and*
- *$g(\eta) = g(\omega)$ implies that either $\eta = \omega$ or η and ω are vertex disjoint in K.*

Lemma 36 provides a way of gluing faces of a simplicial complex by a simplicial quotient map obtained from vertex identifications. In particular, Lemma 36 provides conditions under which the gluing does not create unwanted identifications, and the resulting complex thus obtained is also a simplicial complex. Now, we describe a special kind of subdivision, which we call an *S-subdivision* of a d-simplex, with a later application of Lemma 36 in mind.

Let ν be an d-simplex, and let U be the vertex set of ν equipped with an ordering $\succ_\nu$. We construct a complex C_ν obtained from a subdivision of ν such that an r-simplex $\Omega \in \mathsf{C}_\nu$ has the following property: for every vertex $v \in \Omega$, $(\mathrm{star}_{\mathsf{C}_\nu} v) \bigcap U = \emptyset$. The construction of the complex C_ν is described in Algorithm 2.

Algorithm 2 S-subdivision of simplex ν

procedure S-SUBDIVIDE($\nu, \succ_\nu$)
  Let U denote the vertices of ν;
  Let $\mathsf{C}_0 \leftarrow \{\nu\}$; $\Omega_1 \leftarrow \nu$; $U_0 \leftarrow U$;
  for $i = 1$ to $2(d+1)$ **do**
    Perform a stellar subdivision of Ω_i to obtain C_i from C_{i-1};
    Let v_i be the new vertex introduced during the stellar subdivision;
    $U_i \leftarrow U_{i-1} \bigcup \{v_i\}$;
    Extend $\succ_\nu$ as follows: Set $v_i \succ_\nu v$ for all $v \in U_{i-1}$;
    Let Ω_{i+1} be the lexicographically highest d-simplex of C_i;
  end for
  $\mathsf{C}_\nu \leftarrow \mathsf{C}_i$;
  return C_ν, $\Omega_{2(d+1)+1}$;
end procedure

Please refer to Figure 14 for an illustrative example. In Figure 14, $U = \{A, B, C\}$ and $\Omega = \{G, H, I\}$, and the stars of G, H and I do not intersect U.

▶ Remark 37. The total number of d-simplices in C_i for $i \in [0, d+1]$ are $2i \cdot d + 1$. So C_ν has $2d(d+1)+1$ d-simplices. Also, by construction, $\Omega_{2(d+1)+1}$ is the lexicographically highest d-simplex of C_ν.

▶ **Lemma 38.** *For every $i \in [2(d+1)+1]$, Ω_i is the full simplex on the $d+1$ lexicographically highest vertices of C_{i-1}.*

Proof. This is trivially true for $i = 1$ as C_0 has only $d+1$ vertices. Suppose that the statement of the lemma holds true for all $i \in [j]$ for some $j > 1$. Let $\{u_0, u_1, \ldots, u_d\}$ be the vertices of Ω_j where $u_k \succ_\nu u_{k-1}$ for $k \in [d]$. Then, by construction, $\Omega_{j+1} = \{u_1, \ldots, u_d, v_j\}$, which coincides with the set of lexicographically highest vertices of C_j. ◀

▶ **Lemma 39.** *Let v_i be the vertex introduced during the i-th iteration of the algorithm. If $\{u, v_i\}$ is an edge in C_ν, then it has two types.*

- *(type-1) $v_i \succ_\nu u$, or*
- *(type-2) $u \succ_\nu v_i$ and there are at most d vertices w_j, $j \in [d]$ such that $u \succ_\nu w_j \succ_\nu v_i$.*

Proof. In complex C_i, v_i has degree $d+1$. In particular, denoting the vertices of Ω_i by $\{u_0, u_1, \ldots, u_d\}$, the edges $\{v_i, u_k\}$ for $k \in [0, d]$ belong to C_i. So all edges of C_i incident on v_i are of type-1.

Moreover, for every $i' > i$, every vertex $v' \neq v_i$ of $\Omega_{i'}$ satisifes $v' \succ_\nu v_i$ by Lemma 38. Hence, the newly added edges in $\mathsf{C}_{i'}$ for $i' > i$ that are incident on v_i are of type-2. Again, using Lemma 38 inductively, there can be at most $(d+1)$ such vertices v' in the final complex C_ν. ◀

▶ **Proposition 40.** *Let v be a vertex of $\Omega_{2(d+1)+1}$. Then, $(\mathrm{star}_{\mathsf{C}_\nu} v) \bigcap U = \emptyset$.*

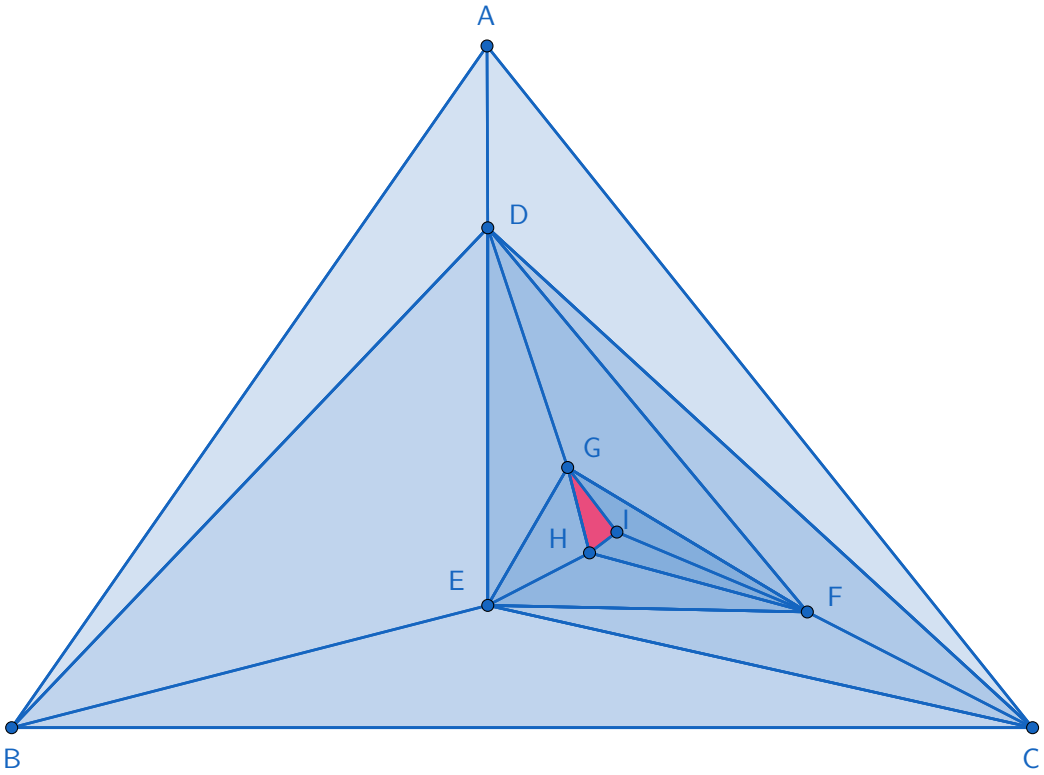


Figure 14 The figure shows a specific subdivision of the 2-simplex ABC defined on $d+1=3$ vertices. The vertices that are higher in the alphabetical order are also higher with respect to the ordering $\succ$. For instance, $G \succ D \succ A$. The above triangulation is obtained as follows: First, the simplex ABC is stellar subdivided by the introduction of the vertex D. Since BCD is the lexicographically highest simplex, it is the only one that is stellar subdivided by the introduction of the vertex E. Then, CDE which is the lexicographically highest simplex is subdivided by the introduction of the vertex F, and so on. Note that at each step the lexicographically highest vertices always span a simplex, and that simplex is the one that is subdivided. For instance, after the first subdivision, BCD is a simplex, after the second subdivision CDE is a simplex, after the third subdivision DEF is a simplex, and so on. The process stops after $2(d+1)$ subdivisions. In this case, we perform six subdivisions. The total number of d-simplices introduced is $2d(d+1)+1$, which in this case is 13. Note that the vertices A, B and C do not lie in the respective stars of the vertices of the highlighted triangle GHI.

Proof. The complex C_ν has $3(d+1)$ vertices totally ordered by $\succ_\nu$. By Lemma 38, the vertices of $\Omega_{2(d+1)+1}$ are the highest $d+1$ vertices ordered by $\succ_\nu$. By construction, the vertices in U are are the lowest $d+1$ vertices ordered by $\succ_\nu$.

By Lemma 39, the vertices of $\Omega_{2(d+1)+1}$ do not have any edges in common with vertices in U. The claim follows. ◀

6.2.2 Description of the reduction

We now give a detailed description of the reduction. As before, associated to a k-colored graph $G = (V, E)$, we define an r-dimensional complex $\mathsf{L}(G)$ as follows.

Vertices.

Let $V' = V \bigcup [k]$, and $r = |V| - 1$. Then, $|V'| = r + k + 1$. Include the vertex set V' in $\mathsf{L}(G)$. In what follows, we add further vertices to $\mathsf{L}(G)$.

Ordering relation $\succ_{V'}$ on vertices of $\mathsf{L}(G)$.

We now impose the following ordering relation on V'. Enumerate the vertices of G according to a fixed total order $V = \{v_1, v_2, \ldots, v_{r+1}\}$. For every color $i \in [k]$ and $j \in [r+1]$, we have $i \succ_{V'} v_j$. For $i_2 \geq i_1$, we have $i_2 \succ_{V'} i_1$, and for $j_2 \geq j_1$, we have $v_{j_2} \succ_{V'} v_{j_1}$.

▶ Remark 41 (Implementing undesirability). The undesirability of certain r-simplices is implemented in the gadget as follows: Let $m = n^3$. Then, to every undesirable r-simplex $\omega = \{v_1, v_2, \ldots, v_{r+1}\}$, associate m new vertices $\mathcal{U}^\omega = \{u_1^\omega, u_2^\omega, \ldots, u_m^\omega\}$. For every $\ell \in [m]$, let $\mu_\ell = \{v_1, v_2, \ldots, v_{r+1}, u_\ell^\omega\}$. Now introduce $m(r+1)$ *new* r-simplices

$$\Upsilon^\omega = \{\{\text{facets of } \mu_\ell\} \setminus \{\omega\} \mid \ell \in [m]\}.$$

Note that for any two undesirable simplices ω_1 and ω_2 we have, $\mathcal{U}^{\omega_1} \bigcap \mathcal{U}^{\omega_2} = \emptyset$ and $\Upsilon^{\omega_1} \bigcap \Upsilon^{\omega_2} = \emptyset$. As observed later, introducing these new simplices makes inclusion of ω in the solution set prohibitively expensive. Please refer to Figure 15 for an illustrative example. For undesirable simplices, we denote the set of r-simplices in $\Upsilon^\omega \bigcup \omega$ by $[\omega]$. For admissible simplices, $[\omega] = \omega$.

Gadgets.

The complex $\mathsf{L}(G)$ is constructed by gluing the distinguished faces of two types of gadgets. Next, we describe these two types of gadgets.

Gadgets of type-1.

The construction of gadgets of type-1 is explained in detail in the pseudocode of Algorithm 3. Below, we provide a high-level sketch.

First, we describe the subroutine SubdivideDelta1. In this subroutine, given an index $i \in [k]$, we begin our construction with the complex $\hat{\partial}\sigma_i = \{\text{facets of } \sigma_i\} \setminus \{V\}$, where as in Section 6.1, $\sigma_i = V \bigcup \{i\}$. The vertices of $\hat{\partial}\sigma_i$ inherit an order from $\succ_{V'}$.

▶ **Definition 42** (Pre-admissible and non-pre-admissible simplices of $\hat{\partial}\sigma_i$). *For every $v \in V_i$, the simplices $a_i^v = (V \setminus \{v\}) \bigcup \{i\}$ are called the* pre-admissible simplices *of $\hat{\partial}\sigma_i$, and all other simplices of $\hat{\partial}\sigma_i$ are called* non-pre-admissible.

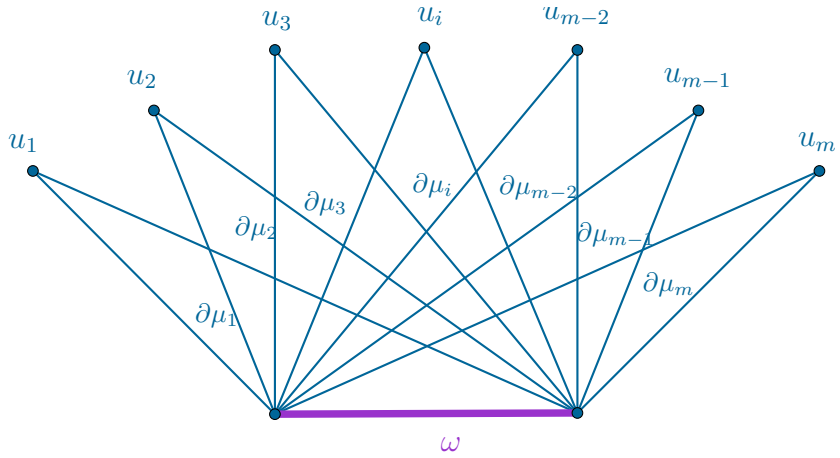


Figure 15 For every undesirable simplex $\omega = \{v_1, v_2, \ldots, v_{r+1}\}$ m new vertices $\mathcal{U}^\omega = \{u_1^\omega, u_2^\omega, \ldots, u_m^\omega\}$ are added to $\mathsf{L}(G)$. Furthermore, $m(r+1)$ new r-simplices $\Upsilon^\omega = \{\{\text{facets of } \mu_\ell\} \setminus \{\omega\} \mid \ell \in [m]\}$ are added to $\mathsf{L}(G)$. The r-simplices in $\Upsilon^\omega \bigcup \omega$ are denoted by $[\omega]$.

The procedure S-SUBDIVIDE described in Section 6.2.1 takes an r-simplex ν as input and returns a subdivision of ν along with the (lexicographically highest) distinguished simplex from within the subdivided simplex. For every pre-admissible simplex a_i^v, its subdivision is denoted by C_i^v, and the distinguished simplex of C_i^v is denoted by $\alpha_i^v(\sigma_i)$. The complex C is formed by taking the union of the subdivided pre-admissible simplices.

Let A denote the collection of *distinguished simplices* in C, and let W denote the set of non-pre-admissible r-simplices of $\hat{\partial}\sigma_i$. Since there are V_i pre-admissible simplices for color i, $|A| = |V_i|$. Finally, the complex $\hat{\Delta}\sigma_i$ is formed by taking the union of the non-pre-admissible simplices, namely W, with the collection of subdivisions of the pre-admissible simplices, namely C. We end the description of SUBDIVIDEDELTA1 with one last definition.

▶ **Definition 43** (Undesirable simplices of $\hat{\Delta}\sigma_i$)**.** *At the end of the procedure* SUBDIVIDEDELTA1*, the simplices in* $\hat{\Delta}\sigma_i \setminus A$ *are called the* undesirable simplices *of* $\hat{\Delta}\sigma_i$*.*

In procedure TYPEZ1, the complex $\mathcal{Z}_1$ is constructed. To begin with, the subroutine SUBDIVIDEDELTA1 is invoked for every $i \in [k]$, which returns the complex $\hat{\Delta}\sigma_i$ along with its *distinguished simplices* A_{σ_i}. Next, we add further simplices to $\hat{\Delta}\sigma_i$ in order to implement undesirability of simplices as per Remark 41. As per the notation used in TYPEZ1, $\hat{\Delta}\sigma_i \setminus A_{\sigma_i}$ are the undesirable simplices of $\hat{\Delta}\sigma_i$. Then, to every undesirable simplex $\omega \in \hat{\Delta}\sigma_i \setminus A_{\sigma_i}$, we add $(r+1)m$ simplices Υ^ω to $\hat{\Delta}\sigma_i$, completing the construction of $\hat{\Delta}\sigma_i$. The complex $\mathcal{Z}_1$ is then given by the union of all simplices in $\hat{\Delta}\sigma_i$ for every i. We end the description of TYPEZ1 with a definition.

▶ **Definition 44** (Inadmissible simplices of $\hat{\Delta}\sigma_i$)**.** *At the end of the procedure* TYPEZ1*, the simplices in* $\hat{\Delta}\sigma_i \setminus A_{\sigma_i}$ *are the* inadmissible simplices *of* $\hat{\Delta}\sigma_i$*.*

Gadgets of type-2.

We now provide a high-level description of gadgets of type-2, the pseudocode of which is provided in Algorithm 4. The type-2 gadgets are indexed by t.

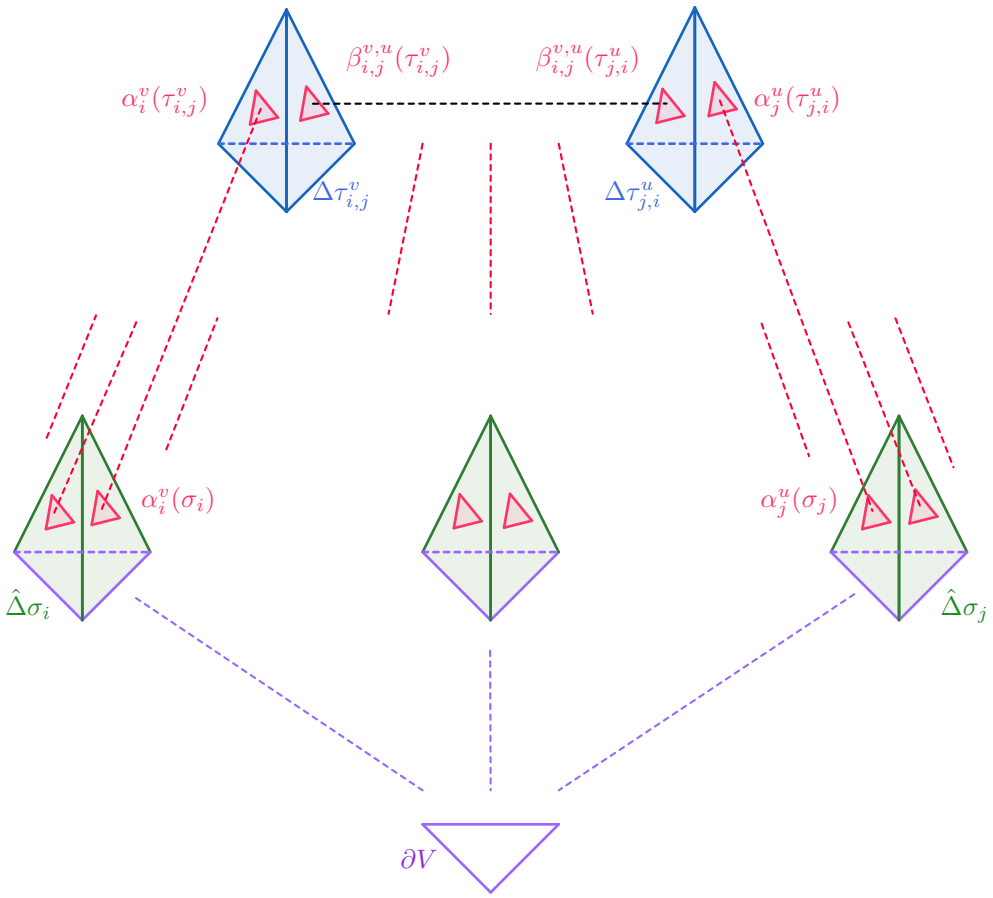


Figure 16 The figure depicts the construction of $\mathsf{L}(G)$ via identifications of various gadgets as described in Algorithm 5. In particular, the dashed red lines show identifications of the (red) congruent faces of type-1 gadgets (shown in green) and type-2 gadgets (shown in blue). The dashed black line shows identifications of the (red) congruent faces of two distinct type-2 gadgets. Note that some of the red dashed lines are only partially drawn. The red faces are the lexicographically highest distinguished faces obtained by S-subdivisions described in Section 6.2.1. The construction of the type-1 green gadgets is described in Algorithm 3, and the construction of the type-2 blue gadgets is described in Algorithm 4. Note that the full simplex with vertex set V (or its subdivision) does not appear as a simplex in any of the type-1 gadgets. In fact the type-1 gadget are r-manifolds with ∂V as their common boundary, and the dashed purple lines depict precisely that. The type-2 gadgets are topological r-spheres.

Algorithm 3 Construction of complex $\hat{\Delta}\sigma_i$

procedure SUBDIVIDEDELTA1(i)
  $\sigma_i \leftarrow V \bigcup \{i\}$;
  Let $\hat{\partial}\sigma_i$ be the r-complex $\partial\sigma_i \setminus \{V\}$;
  The vertex set U of $\hat{\partial}\sigma_i$ is ordered by $\succ$, obtained by restricting $\succ_{V'}$ to U;
  Every r-simplex $a_i^v = (V \setminus \{v\}) \bigcup \{i\}$ of $\hat{\partial}\sigma_i$ with $v \in V_i$ is deemed *pre-admissible*;
  Let W denote the set of r-simplices of $\hat{\partial}\sigma_i$ that are not pre-admissible;
  $A \leftarrow \emptyset$; $\mathsf{C} \leftarrow \emptyset$;
  for each pre-admissible simplex a_i^v of $\hat{\partial}\sigma_i$ **do**
    C_i^v, $\alpha_i^v(\sigma_i) \leftarrow$ S-SUBDIVIDE$(a_i^v, \succ)$;
    $A \leftarrow A \bigcup \{\alpha_i^v(\sigma_i)\}$;
    $\mathsf{C} \leftarrow \mathsf{C} \bigcup \mathsf{C}_i^v$;
  end for
  $\hat{\Delta}\sigma_i \leftarrow \mathsf{C} \bigcup W$;
  return $\hat{\Delta}\sigma_i$, A; ▷ The simplices in $\hat{\Delta}\sigma_i \setminus A$ are undesirable.
end procedure

procedure TYPEZ1
  $\mathcal{Z}_1 \leftarrow \emptyset$;
  for $i = 1$ to k **do**
    $\hat{\Delta}\sigma_i$, $A_{\sigma_i} \leftarrow$ SUBDIVIDEDELTA1(i) ;
    The simplices in A_{σ_i} are the distinguished simplices of $\hat{\Delta}\sigma_i$;
    The simplices in $\hat{\Delta}\sigma_i \setminus A_{\sigma_i}$ are deemed the undesirable simplices of $\hat{\Delta}\sigma_i$;
    for every undesirable simplex ω in $\hat{\Delta}\sigma_i$ **do** ▷ as described in Remark 41
      Add $(r+1)m$ simplices Υ^ω to $\hat{\Delta}\sigma_i$;
    end for ▷ The simplices in $\hat{\Delta}\sigma_i \setminus A_{\sigma_i}$ are inadmissible.
    $\mathcal{Z}_1 \leftarrow \mathcal{Z}_1 \bigcup \hat{\Delta}\sigma_i$;
  end for
end procedure

First, we describe the subroutine SubdivideDelta2. In this subroutine, given a vertex $v \in V_i$ a color $j \neq i$, and an index t, we introduce a vertex set V_t whose vertices are in one-to-one correspondence with the vertices $(V \setminus \{v\}) \bigcup \{i, j\}$. Let $\tau^v_{i,j}$ be the full $(r+1)$-simplex on V_t, and $\partial\tau^v_{i,j}$ be the complex induced by the facets of $\tau^v_{i,j}$. The vertices of $\partial\tau^v_{i,j}$ are ordered according to the same rules as $\succ_{V'}$.

▶ **Definition 45** (Pre-admissible and non-pre-admissible simplices of $\partial\tau^v_{i,j}$)**.** *The simplices $a^v_i = V_t \setminus \{j_t\}$ for every $v \in V_i$, and the simplices $b^{v,u}_{i,j} = V_t \setminus \{u_t\}$, where $u \in V_j$ and $\{u, v\} \in E$ are also said to be the* pre-admissible simplices *of $\partial\tau^v_{i,j}$. All other r-simplices of $\partial\tau^v_{i,j}$ are deemed* non-pre-admissible.

We invoke the procedure S-Subdivide described in Section 6.2.1 to subdivide the pre-admissible simplices of $\partial\tau^v_{i,j}$. The subdivision of a pre-admissible simplex $b^{v,u}_{i,j}$ is denoted by $\mathsf{C}^{v,u}_{i,j}$ and the distinguished simplex of $\mathsf{C}^{v,u}_{i,j}$ is denoted by $\beta^{v,u}_{i,j}(\tau^u_{j,i})$. The subdivision of a pre-admissible simplex $\alpha^v_i(\tau^v_{i,j})$ is denoted by C^v_i and the distinguished simplex of C^v_i is denoted by a^v_i. The complex C is formed by taking the union of the subdivided pre-admissible simplices. Furthermore, the collection of all the distinguished simplices of the subdivided pre-admissible simplices is denoted by A. It is easy to check that, $|A| = k$. Finally, the complex $\Delta\tau^v_{i,j}$ is formed by taking the union of the non-pre-admissible simplices, namely W, with the collection of subdivisions of the pre-admissible simplices, namely C. We conclude the description of SubdivideDelta2 with a definition.

▶ **Definition 46** (Undesirable simplices of $\Delta\tau^v_{i,j}$)**.** *At the end of procedure* SubdivideDelta2*, the simplices in $\Delta\tau^v_{i,j} \setminus A$ are said to be the* undesirable simplices *of $\Delta\tau^v_{i,j}$.*

In the procedure TypeZ2, the complex $\mathcal{Z}_2$ is constructed. To do this, the subroutine SubdivideDelta2 is invoked for every color i, every vertex v in V_i, and every color j where $j \neq i$, which returns the complex $\Delta\tau^v_{i,j}$, along with its set of distinguished simplices $A_{\tau^v_{i,j}}$. Next, we add further simplices to $\Delta\tau^v_{i,j}$, for every $v \in V_i$ and $j \in [k] \setminus \{i\}$, in order to implement undesirability of simplices as per Remark 41. We start the construction with the undesirable simplices of $\Delta\tau^v_{i,j}$, namely $\Delta\tau^v_{i,j} \setminus A_{\tau^v_{i,j}}$. To every undesirable simplex $\omega \in \Delta\tau^v_{i,j} \setminus A_{\tau^v_{i,j}}$, we add $(r+1)m$ simplices Υ^ω to $\Delta\tau^v_{i,j}$, completing the construction of $\Delta\tau^v_{i,j}$. The complex $\mathcal{Z}_2$ is then given by the union of all simplices in $\Delta\tau^v_{i,j}$ for every $i \in [k]$, every vertex $v \in V_i$ and every $j \in [k] \setminus \{i\}$. We conclude the description of TypeZ2 with a definition.

▶ **Definition 47** (Inadmissible simplices of $\Delta\tau^v_{i,j}$)**.** *At the end of procedure* TypeZ2*, the simplices in $\Delta\tau^v_{i,j} \setminus A_{\tau^v_{i,j}}$ are the* inadmissible simplices *of $\Delta\tau^v_{i,j}$.*

Attachments.

Let $\mathsf{K}' = \mathcal{Z}_1 \bigcup \mathcal{Z}_2$. Then, complex $\mathsf{L}(G)$ is formed from K' after making the attachments described in Algorithm 5.

Algorithm 4 Construction of complex $\Delta\tau^v_{i,j}$

procedure SUBDIVIDEDELTA2(v, i, j, t)
  Let $V_t \leftarrow \emptyset$;
  for every $u \in V \setminus v$ **do**
    Add a vertex u_t to V_t;
  end for
  $V_t \leftarrow V_t \bigcup \{i_t, j_t\}$;
  Let $\tau^v_{i,j}$ be the full $(r+1)$-simplex on V_t;
  Let $\partial\tau^v_{i,j}$ be the r-complex induced by the facets of $\tau^v_{i,j}$;
  The vertices V_t of $\partial\tau^v_{i,j}$ are in a natural one-to-one correspondence to a subset of vertices in V'. The ordering $\succ_t$ on V_t is defined using the same rules as for $\succ_{V'}$;
  The r-simplex $a^v_i = V_t \setminus \{j_t\}$ is deemed *pre-admissible*;
  The r-simplices $\left\{b^{v,u}_{i,j} \mid b^{v,u}_{i,j} = V_t \setminus \{u_t\},\ \text{where } u \in V_j \text{ and } \{u, v\} \in E\right\}$ are also deemed *pre-admissible*;
  Let W denote the set of r-simplices of $\partial\tau^v_{i,j}$ that are not pre-admissible;
  $\mathsf{C}^v_i,\ \alpha^v_i(\tau^v_{i,j}) \leftarrow$ S-SUBDIVIDE$(a^v_i, \succ_t)$;
  $A \leftarrow \{\alpha^v_i(\tau^v_{i,j})\}$; $\mathsf{C} \leftarrow \mathsf{C}^v_i$;
  for each pre-admissible simplex $b^{v,u}_{i,j}$ of $\partial\tau^v_{i,j}$ **do**
    $\mathsf{C}^{v,u}_{i,j},\ \beta^{v,u}_{i,j}(\tau^u_{j,i}) \leftarrow$ S-SUBDIVIDE$(b^{v,u}_{i,j}, \succ_t)$;
    $A \leftarrow A \bigcup \{\beta^{v,u}_{i,j}(\tau^u_{j,i})\}$;
    $\mathsf{C} \leftarrow \mathsf{C} \bigcup \mathsf{C}^{v,u}_{i,j}$;
  end for
  $\Delta\tau^v_{i,j} \leftarrow \mathsf{C} \bigcup W$;
  return $\Delta\tau^v_{i,j}$, A; ▷ The simplices in $\Delta\tau^v_{i,j} \setminus A$ are undesirable.
end procedure

procedure TYPEZ2
  $\mathcal{Z}_2 \leftarrow \emptyset$; $t = 0$;
  for $i = 1$ to k **do**
    for every vertex v in V_i, and a color $j \in [k] \setminus \{i\}$ **do**
      $t = t + 1$;
      $\Delta\tau^v_{i,j},\ A_{\tau^v_{i,j}} \leftarrow$ SUBDIVIDEDELTA2(v, i, j, t) ;
      The simplices in $A_{\tau^v_{i,j}}$ are the distinguished simplices of $\Delta\tau^v_{i,j}$;
      The simplices in $\Delta\tau^v_{i,j} \setminus A_{\tau^v_{i,j}}$ are deemed the undesirable simplices of $\Delta\tau^v_{i,j}$;
      for every undesirable simplex ω in $\Delta\tau^v_{i,j}$ **do** ▷ as described in Remark 41
        Add another $(r+1)m$ simplices Υ^ω to $\Delta\tau^v_{i,j}$;
      end for ▷ The simplices in $\Delta\tau^v_{i,j} \setminus A_{\tau^v_{i,j}}$ are inadmissible.
      $\mathcal{Z}_2 \leftarrow \mathcal{Z}_2 \bigcup \Delta\tau^v_{i,j}$;
    end for
  end for
end procedure

Algorithm 5 Construction of complex $\mathsf{L}(G)$

1: **for every** $v \in V_i$ and $j \neq i$ **do**
2: Identify the r-simplices: $\alpha_i^v(\tau_{i,j}^v) \sim \alpha_i^v(\sigma_i)$, where the identifications of vertices are consistent with respective lexicographic orderings. Denote the identified simplex by α_i^v;
3: **end for**
4: **for every** edge $\{u, v\} \in E$ with $v \in V_i$ and $u \in V_j$ **do**
5: Identify the r-simplices: $\beta_{i,j}^{v,u}(\tau_{j,i}^u) \sim \beta_{i,j}^{v,u}(\tau_{i,j}^v)$, by respecting the respective lexicographic orderings. Denote the identified simplex by $\beta_{i,j}^{v,u}$;
6: **end for**

▶ **Definition 48** (Admissible simplices of $\mathsf{L}(G)$)**.** *The simplices α_i^v for every $i \in [j]$ and $v \in V_i$, and the simplices $\beta_{i,j}^{v,u}$ for every edge $\{u, v\} \in E$ with $v \in V_i$ and $u \in V_j$ are said to be the* admissible simplices *of* $\mathsf{L}(G)$.

▶ **Proposition 49.** *The complex $\mathsf{L}(G)$ formed from identifying vertices in K' is a simplicial complex.*

Proof. This follows immediately from Lemma 36 and Proposition 40. ◀

This completes the construction of complex $\mathsf{L}(G)$. Please refer to Figure 19 for a schematic illustration.

Choice of input boundary.

For the abstract simplex V, let ∂V denote the set of facets of V. Note that although $V \notin \mathsf{L}(G)$, every facet of V is in $\mathsf{L}(G)$. In fact, the complex $\hat{\Delta}\sigma_i$ for every i, is a simplical r-manifold with ∂V as its boundary. We choose ∂V as our input boundary that we want to make nontrivial.

Choice of parameter.

Let $\left(k + \binom{k}{2} = \binom{k+1}{2}\right)$ be the parameter for Boundary Nontrivialization on the complex $\mathsf{L}(G)$.

▶ **Proposition 50.** *If there exists a k-clique H of G such that every vertex of H has a different color, then a set of $\binom{k+1}{2}$ r-simplices in $\mathsf{L}(G)$ meets every chain ξ with $\partial\xi = \partial V$.*

Proof. As in Section 6.1, we construct a set $\mathcal{S}$ of r-simplices that mimics the graphical structure of H as follows:

$$\mathcal{S}_\alpha = \left\{\alpha_i^v \mid v \in V_i \bigcap V_H\right\}$$

$$\mathcal{S}_\beta = \left\{\beta_{i,j}^{v,u} \mid v \in V_i,\, u \in V_j,\, \{i, j\} \in E_H\right\}$$

Set $\mathcal{S} = \mathcal{S}_\alpha \bigcup \mathcal{S}_\beta$.

Now we want to show that at least one element from the solution set $\mathcal{S}$ has coefficient 1 in every chain ξ that satisfies $\partial\xi = \partial V$. Thus, we aim to show that removing $\mathcal{S}$ from $\mathsf{L}(G)$ makes ∂V nontrivial. Before we proceed, we introduce some notations and definitions. To begin with let $\mathcal{A}$ denote the set of all admissible simplices in $\mathsf{L}(G)$ (described in Definition 48).

▶ Notation 5. For an r-simplex ω, let $[\omega]_\xi$ denote the simplices of $[\omega]$ in ξ.

▶ **Definition 51** (Type-1 gadget belonging to chain ξ)**.** *If there exists an r-simplex $\omega \in \hat{\Delta}\sigma_i \setminus \mathcal{A}$ such that $\partial([\omega]_\xi) = 1$, then we say that σ_i* belongs to *ξ.*

▶ **Definition 52** (Type-2 gadget belonging to chain ξ)**.** *If there exists an r-simplex $\omega \in \Delta\tau^v_{i,j} \setminus \mathcal{A}$ such that $\partial([\omega]_\xi) = 1$, then we say that $\tau^v_{i,j}$* belongs to *ξ.*

Before we can finish the proof of Proposition 50, we need a few auxillary lemmas. For the lemmas that follow, we let ξ be a chain that satisfies $\partial\xi = \partial V$.

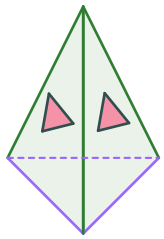


Figure 17 The figure depicts the gadget $\hat{\Delta}\sigma_i$ in a simplistic manner, that is, without the full triangulation and without the inadmissible simplices. Also, in this figure, σ_i *belongs to* ξ. That is, there exists an r-simplex $\omega \in \hat{\Delta}\sigma_i \setminus \mathcal{A}$ such that $\partial([\omega]_\xi) = 1$. The simplices of $\hat{\Delta}\sigma_i$ that lie in $\mathcal{A}$ are shown in red. Then, according to Lemma 53, the boundary of the part of the complex in green equals the boundary of triangles in red (i.e., the black edges) + ∂V (shown in purple).

▶ **Lemma 53.** *If σ_i belongs to ξ, then $\partial\left(\left(\hat{\Delta}\sigma_i \setminus \mathcal{A}\right) \bigcap \xi\right) = \partial V + \partial\left(\hat{\Delta}\sigma_i \bigcap \mathcal{A}\right)$.*

Proof. Please refer to Figure 17 for an illustration of the statement of the lemma. Since σ_i belongs to ξ, there exists an r-simplex $\omega_1 \in \hat{\Delta}\sigma_i \setminus \mathcal{A}$ such that $\partial([\omega_1]_\xi) \neq 0$. This implies that there exists a facet ς of ω_1 such that $\varsigma \in \partial([\omega_1]_\xi)$ and $\varsigma \notin \partial\xi$. Hence, there must be an r-simplex ω_2 with ς as a facet such that $\omega_2 \in \hat{\Delta}\sigma_i \setminus \mathcal{A}$, $\partial([\omega_2]_\xi) \neq 0$ and ς vanishes in $\partial([\omega_1]_\xi + [\omega_2]_\xi)$. Repeating the argument above, we inductively add classes $[\omega_j]_\xi$, where $\omega_j \in \hat{\Delta}\sigma_i \setminus \mathcal{A}$ such that $\partial([\omega_j]_\xi) \neq 0$. Note that by construction, $\bigcup_j [\omega_j]_\xi = \left(\hat{\Delta}\sigma_i \setminus \mathcal{A}\right) \bigcap \xi$, where j indexes the simplices in $\hat{\Delta}\sigma_i \setminus \mathcal{A}$. Clearly, the induction stops when $\partial(\bigcup_j [\omega_j]_\xi) = \partial V + \partial\left(\hat{\Delta}\sigma_i \bigcap \mathcal{A}\right)$. ◀

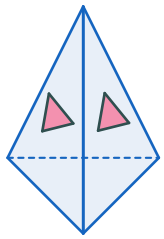


Figure 18 The figure is a simplistic depiction of gadget $\Delta\tau^v_{i,j}$. In particular, the full triangulation and the the inadmissible simplices of $\Delta\tau^v_{i,j}$ are not shown. In this figure, $\tau^v_{i,j}$ *belongs to* ξ. That is, there exists an r-simplex $\omega \in \Delta\tau^v_{i,j} \setminus \mathcal{A}$ such that $\partial([\omega]_\xi) = 1$. The simplices of $\Delta\tau^v_{i,j}$ that lie in $\mathcal{A}$ are shown in red. Then, according to Lemma 54, the boundary of the part of the complex in blue equals the boundary of triangles in red (i.e., the black edges).

▶ **Lemma 54.** *If $\tau^v_{i,j}$ belongs to ξ, then $\partial\left(\left(\Delta\tau^v_{i,j} \setminus \mathcal{A}\right) \bigcap \xi\right) = \partial\left(\Delta\tau^v_{i,j} \bigcap \mathcal{A}\right)$.*

Proof. The argument is identical to the proof of Lemma 53. Please refer to Figure 18 for an illustration of the statement of the lemma. ◀

▶ **Lemma 55.** *The cardinality of the set $\{i \in [k] \mid \sigma_i$ belongs to $\xi\}$ is odd.*

Proof. Please see (the bottom portion of) Figure 19 for an illustration of the statement of the lemma. First, note that by construction, $\partial V \bigcap \partial(\mathcal{S}) = \emptyset$. Then, using Lemma 53, we have $\partial V \subset \partial\left((\hat{\Delta}\sigma_i \setminus \mathcal{A}) \bigcap \xi\right)$ for every σ_i that belongs to ξ. Since ∂V only occurs in the boundaries of type-1 gadgets and $\partial\xi = \partial V$, the cardinality of the set $\{i \in [k] \mid \sigma_i$ belongs to $\xi\}$ must be odd. ◀

▶ **Lemma 56.** *If $\partial\xi \bigcap \mathcal{S} = \emptyset$, and if σ_i belongs to ξ for some $i \in [k]$, then the cardinality of*

$$I = \left\{\tau^v_{i,j} \;\middle|\; v \in V_i \bigcap V_H,\ j \in [k] \setminus \{i\} \text{ and } \tau^v_{i,j} \text{ belongs to } \xi\right\}$$

is odd. On the other hand, if σ_i does not belong to ξ, then I is even.

Proof. Please see (the middle portion of) Figure 19 for an illustration of the statement of the lemma.

Case 1: σ_i belongs to ξ.

Since H is a multicolored clique, for color i, there exists a vertex $v \in V_i \bigcap V_H$. Hence, by construction, $\alpha^v_i \in \mathcal{S}$. Moreover, α^v_i is the only r-simplex that is common to $\hat{\Delta}\sigma_i$ and $\Delta\tau^v_{i,j}$ for every $\tau^v_{i,j} \in I$. Note that $\mathcal{S} \subseteq \mathcal{A}$, and simplices in $\mathcal{A}$ have disjoint boundaries. Since σ_i belongs to ξ, we obtain

$$\partial\alpha^v_i \subset \partial\left(\hat{\Delta}\sigma_i \bigcap \mathcal{S}\right) \subset \partial\left(\hat{\Delta}\sigma_i \bigcap \mathcal{A}\right) = \partial\left(\left(\hat{\Delta}\sigma_i \setminus \mathcal{A}\right) \bigcap \xi\right). \tag{1}$$

where the last equality uses Lemma 53.

Moreover for every $\tau^v_{i,j} \in I$, we obtain

$$\partial\alpha^v_i \subset \partial\left(\Delta\tau^v_{i,j} \bigcap \mathcal{S}\right) \subset \partial\left(\Delta\tau^v_{i,j} \bigcap \mathcal{A}\right) = \partial\left(\left(\Delta\tau^v_{i,j} \setminus \mathcal{A}\right) \bigcap \xi\right), \tag{2}$$

where the last equality uses Lemma 54.

For $j \in [k] \setminus \{i\}$ such that $\tau^v_{i,j} \notin I$,

$$\partial\left(\left(\Delta\tau^v_{i,j} \setminus \mathcal{A}\right) \bigcap \xi\right) = 0, \tag{3}$$

which is a simple consequence of Definition 52.

Using the assumption $\partial\xi \bigcap \mathcal{S} = \emptyset$, and Equations (1)–(3) we get

$$((I+1) \bmod 2) \cdot \partial\alpha^v_i \subset \partial\xi.$$

Since $\partial\xi = \partial V$, and $\alpha^v_i \bigcap \partial V = \emptyset$, $I+1$ should be even, proving the first claim.

Case 2: σ_i does not belong to ξ.

In this case,

$$\partial\alpha^v_i \subset \partial\left(\hat{\Delta}\sigma_i \bigcap \mathcal{S}\right) \subset \partial\left(\hat{\Delta}\sigma_i \bigcap \mathcal{A}\right) \not\subset \partial\left(\left(\hat{\Delta}\sigma_i \setminus \mathcal{A}\right) \bigcap \xi\right),$$

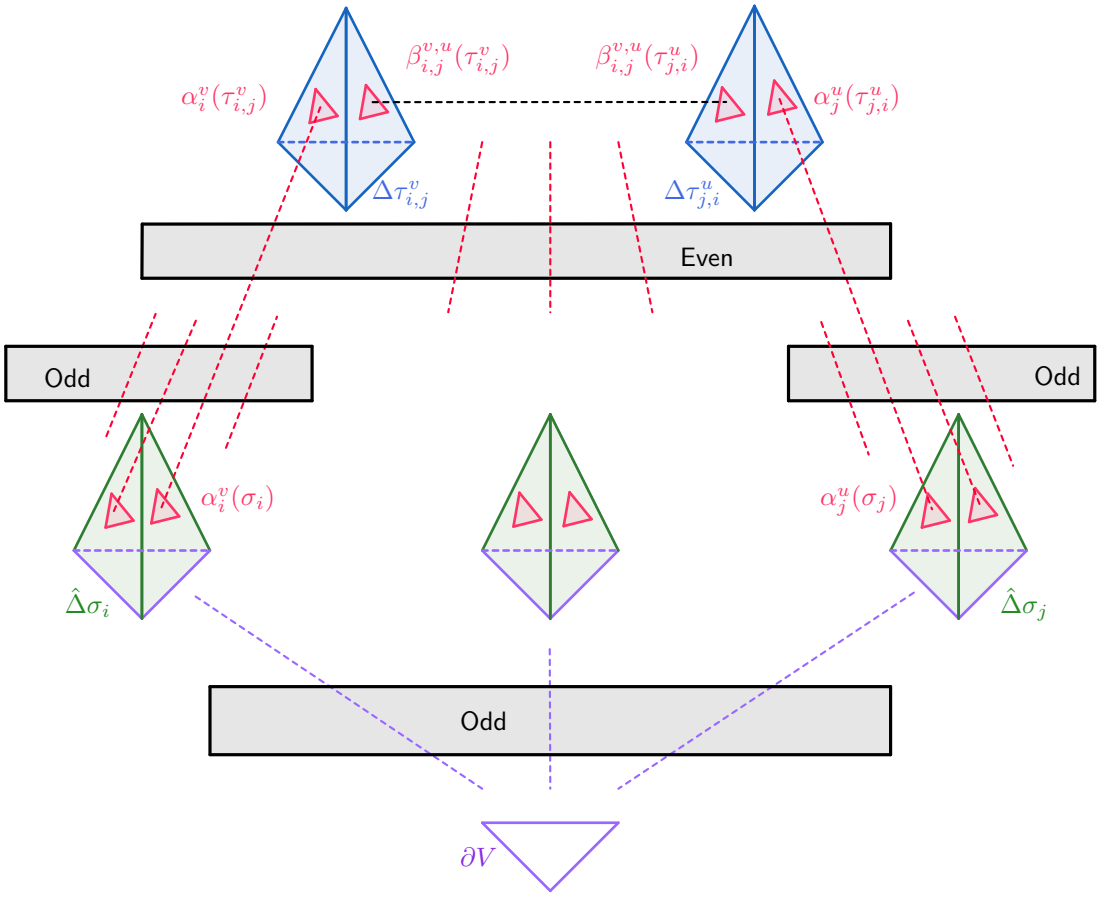


Figure 19 As in the case of Figure 16, this figure also depicts some of the gadgets of $\mathsf{L}(G)$. However, we depict only those gadgets that *belong* to some chain ξ. See Definitions 51 and 52 for what it means for a gadget to belong to a chain. The dashed red lines show identifications of the (red) congruent faces of type-1 gadgets that belong to ξ to the (red) congruent faces of type-2 gadgets that belong to ξ. The dashed black line shows identifications along the (red) congruent faces of two distinct type-2 gadgets that belong to ξ.
The odd count of purple dashed lines is the content of the Lemma 55. The odd count of each group of red dashed lines in the middle is the content of Lemma 56. Hence, for every purple dashed line, there is a group of red dashed lines of odd cardinality. On the one hand, since an odd sum of odd numbers is odd, by Lemmas 55 and 56, the total number of red dashed lines should be odd. On the other hand, Lemma 57 says that the cardinality of the red dashed lines (counted from above) is even. The main idea of Proposition 50 which uses Lemmas 55–57, and a proof by contradiction is that an odd sum of odd numbers cannot be even.

where the last non-inclusion follows from $\partial\left(\left(\hat{\Delta}\sigma_i \setminus \mathcal{A}\right)\bigcap \xi\right) = 0$ (as a simple consequence of Definition 51). But for every $\tau^v_{i,j} \in I$, we still have

$$\partial\alpha^v_i \subset \partial\left(\Delta\tau^v_{i,j}\bigcap \mathcal{S}\right) \subset \partial\left(\Delta\tau^v_{i,j}\bigcap \mathcal{A}\right) = \partial\left(\left(\Delta\tau^v_{i,j}\setminus \mathcal{A}\right)\bigcap \xi\right).$$

which gives

$$(I \bmod 2)\cdot \partial\alpha^v_i \subset \partial\xi.$$

Since $\partial\xi = \partial V$, and $\alpha^v_i \bigcap \partial V = \emptyset$, I should be even, proving the second claim. ◀

▶ **Lemma 57.** *Assuming $\partial\xi \bigcap \mathcal{S} = \emptyset$, we define the set $\mathcal{P}$ as*

$$\mathcal{P} = \left\{(i,j)\;\middle|\; \sigma_i \text{ and } \tau^v_{i,j} \text{ for some } v \in V_i \bigcap V_H \text{ belong to } \xi\right\}.$$

Then, $|\mathcal{P}|$ is even.

Proof. Please see (the top portion of) Figure 19 for an illustration of the statement of the lemma. First, we define $\mathcal{P}'$ as follows.

$$\mathcal{P}' = \left\{(i,j)\;\middle|\; \tau^v_{i,j} \text{ for some } v \in V_i \bigcap V_H \text{ belongs to } \xi\right\}.$$

Suppose $(i,j) \in \mathcal{P}'$ for some $i \in [k]$, and $j \neq i$. Since H is a multicolored clique, for color i, there exists a vertex $v \in V_i \bigcap V_H$. Also, there exists a vertex $u \in V_j \bigcap V_H$ and an edge $\{u,v\} \in E_H$. By construction of $\mathcal{S}$, $\beta^{v,u}_{i,j} \in \mathcal{S}$. Once again, we will use the facts: **1.** $\mathcal{S} \subseteq \mathcal{A}$, and **2.** the simplices in $\mathcal{A}$ have disjoint boundaries.
For every $\tau^v_{i,j} \in I$, we obtain

$$\partial\beta^{v,u}_{i,j} \subset \partial\left(\Delta\tau^v_{i,j}\bigcap \mathcal{S}\right) \subset \partial\left(\Delta\tau^v_{i,j}\bigcap \mathcal{A}\right) = \partial\left(\left(\Delta\tau^v_{i,j}\setminus \mathcal{A}\right)\bigcap \xi\right), \tag{4}$$

where the last equality uses Lemma 54.
Moreover, by construction, $\beta^{v,u}_{i,j}$ belongs to only two gadgets of $\mathsf{L}(G)$: $\Delta\tau^u_{j,i}$ and $\Delta\tau^v_{i,j}$.
Using $\partial\xi = \partial V$ and $\partial V \bigcap \partial\beta^{v,u}_{i,j} = \emptyset$, we deduce that $\partial\beta^{v,u}_{i,j} \not\subset \partial\xi$. Then, using $\partial\xi \bigcap \mathcal{S} = \emptyset$, we have

$$\partial\beta^{v,u}_{i,j} \subset \partial\left(\left(\Delta\tau^u_{j,i}\setminus \mathcal{A}\right)\bigcap \xi\right) \tag{5}$$

But this forces $\tau^u_{j,i}$ to belong to ξ, and hence the pair (j,i) belongs to $\mathcal{P}'$. Therefore, using Equations (4) and (5), $\mathcal{P}'$ is of even cardinality. Now, define $\mathcal{P}''$ as follows.

$$\mathcal{P}'' = \left\{(i,j)\;\middle|\; \tau^v_{i,j} \text{ for some } v \in V_i \bigcap V_H \text{ belongs to } \xi, \text{ and } \sigma_i \text{ does not belong to } \xi\right\}.$$

By inductively applying Case 2 of Lemma 56, we deduce that $\mathcal{P}''$ is of even cardinality. Finally, $\mathcal{P} = \mathcal{P}' - \mathcal{P}''$. Hence, $\mathcal{P}$ is of even cardinality. ◀

Now, observe that if the conditions of Lemmas 55–57 are simultaneously satisfied, then we reach a contradiction. This is because using Lemmas 55 and 56, $|\mathcal{P}|$ is an odd set of odd numbers and hence odd, whereas according to Lemma 57, $|\mathcal{P}|$ is even. So if the chain ξ has ∂V as its boundary, then the assumption $\partial\xi \bigcap \mathcal{S} = \emptyset$ cannot be satisfied. This concludes the proof of Proposition 50. ◀

▶ **Lemma 58.** *If there exists a chain ξ' with $\partial\xi' = \partial V$ such that only the inadmissible simplices of $\mathcal{R}$ have coefficient 1 in ξ', then the size of $\mathcal{R}$ is at least m.*

Proof. We skip the proof since it is identical to the proof of Lemma 31. ◀

▶ **Lemma 59.** *Let $\mathcal{R}$ be a solution set for* Boundary Nontrivialization *on complex $\mathsf{L}(G)$. Then,*

1. *For every $\hat{\Delta}\sigma_i$, there is at least one facet α_i^v with $v \in V_i$ that is included in $\mathcal{R}$.*
2. *For every unordered pair (i, j), where $i, j \in [k]$, there exists a simplex $\beta_{i,j}^{v,u}$ for some v, u that is included in $\mathcal{R}$.*
3. *If $|\mathcal{R}| \leq \binom{k+1}{2}$, then $|\mathcal{R}| = |A_{\mathcal{R}}| = \binom{k+1}{2}$, where $A_{\mathcal{R}}$ denotes the set of admissible simplices of $\mathcal{R}$.*

Proof. The proof is analogous to the proof of Section 6.1. We repeat it here for the sake of clarity and completeness.

Let $A_{\mathcal{R}}$ denote the set of admissible simplices of $\mathcal{R}$. If ξ is such that $\partial\xi = \partial V$ and $\xi \bigcap A_{\mathcal{R}} = \emptyset$, then we are forced to include inadmissible simplices. In that case, Lemma 58 applies, and $\mathcal{R}$ is of cardinality at least $m = n^3$. But, if we include a total of (more than) n^3 facets in $\mathcal{R}$, we exceed the budget of $\binom{k+1}{2}$. So, going forward, we assume that at least one simplex in $A_{\mathcal{R}}$ has coefficient 1 in every chain ξ, where $\xi = \partial V$.

Note that if at least one simplex from $A_{\mathcal{R}}$ has coefficient 1 in every chain ξ with $\xi = \partial V$, then we do not need simplices that are inadmissible in $\mathcal{R}$. Next, we prove the three claims in the lemma.

1. Let $\xi = \hat{\Delta}\sigma_i$ for some $i \in [k]$. Then, $\partial(\hat{\Delta}\sigma_i) = \partial V$. So if we do not include an admissible simplex α_i^v for some $v \in V_i$ in $\mathcal{R}$, then we would be forced to include some inadmissible simplices of $\hat{\Delta}\sigma_i$.
2. Next, for some fixed i and $j \in [k] \setminus \{i\}$, let $\xi = \hat{\Delta}\sigma_i + \sum_{v \in V_i} (\Delta\tau_{i,j}^v)$. Then, $\partial(\hat{\Delta}\sigma_i) + \sum_{v \in V_i} \partial(\Delta\tau_{i,j}^v) = \partial V$. So unless some admissible facet $\beta_{i,j}^{v,u}$ for some v, u is included in $\mathcal{R}$, the coefficient of all admissible simplices in ξ will be zero, and we would be forced to include inadmissible simplices, which according to Lemma 58 is prohibitively expensive.
3. The third claim follows immediately from the first two. ◀

▶ **Lemma 60.** *If $|\mathcal{R}| = |A_{\mathcal{R}}| = \binom{k+1}{2}$, then one can obtain a k-clique H of G from $\mathcal{R}$.*

Proof. Structurally the proof is identical to Lemma 33. The roles of σ_i and $\tau_{i,j}^u$ are played by $\hat{\Delta}\sigma_i$ and $\Delta\tau_{i,j}^v$, respectively. Moreover, there is a difference of 1 in the cardinality of solution set $\mathcal{R}$, because for Topological Hitting Set, we need to remove V whereas the simplex V is not a part of the complex $\mathsf{L}(G)$ in Boundary Nontrivialization. ◀

Proposition 50 and Lemma 60 together provide a parameterized reduction from k-Multicolored Clique to Boundary Nontrivialization. Using Theorem 23, we obtain the following result.

▶ **Theorem 61.** Boundary Nontrivialization *is* $\mathbf{W}[\mathbf{1}]$*-hard.*

7 FPT algorithms

7.1 FPT algorithm for Topological Hitting Set

In Section 6.1 we showed that Topological Hitting Set is $\mathbf{W}[\mathbf{1}]$-hard with the solution size k as the parameter. This motivates the search of other meaningful parameters that make the problem tractable. With that in mind, in this section, we prove an important structural

property about the connectivity of the minimal solution sets for Topological Hitting Set. First, we start with a definition.

▶ **Definition 62** (Induced subgraphs in Hasse graphs)**.** *Given a d-dimensional complex K with Hasse graph H_{K}, and a set $\mathcal{S}$ of r-simplices for some $r < d$, the subgraph of H_{K}* induced by *$\mathcal{S}$ is the union of $\mathcal{S}$ with the set of $(r+1)$-dimensional simplices incident on $\mathcal{S}$.*

▶ **Lemma 63.** *Given a d-dimensional complex K, a minimal solution of* Topological Hitting Set *for a non-bounding cycle $\zeta \in \mathsf{Z}_r(\mathsf{K})$ for some $r < d$ induces a connected subgraph of H_{K}.*

Proof. Let $H_{\mathcal{S}}$ be the subgraph of the Hasse graph H_{K} induced by a minimum topological hitting set $\mathcal{S}$ of a non-bounding cycle $\zeta \in \mathsf{Z}_r(\mathsf{K})$. Targeting a contradiction, assume there exist two components C_1 and C_2 such that C_1 and C_2 have no edges in common. Note that we do not assume that C_1 and C_2 are connected components, merely that they are components that do not share an edge. Since $\mathcal{S}$ is minimal, there exists a cycle $\phi \in [\zeta]$ that is incident on an r-simplex in C_1 but not on any r-simplices in C_2, and a cycle $\psi \in [\zeta]$ that is incident on an r-simplex in C_2 but not on any r-simplices in C_1. Then, $\phi = \psi + \partial b$, for some $(r+1)$ chain b. Let b' be an $(r+1)$-chain obtained from b by removing exactly those $(r+1)$-simplices that are incident on C_1. Now, let $\phi' = \psi + \partial b'$. By construction, ϕ' is not incident on C_1. Also, because C_1 and C_2 are disconnected, the simplices removed from b to obtain b' are not incident on C_2. Hence, ϕ' is not incident on C_2. In other words, $\phi' \in [\zeta]$ does not meet $\mathcal{S}$, and $\mathcal{S}$ is not a hitting set, a contradiction. Therefore, the induced subgraph of $\mathcal{S}$ is connected. ◀

Note that the path from any r-simplex to a neighboring r-simplex in the Hasse graph is of size 2. So it follows from Lemma 63 that any minimal solution of size at most k lies in some geodesic ball of radius $2k$ of some r-simplex in the Hasse graph. In particular, if we search across the geodesic ball of every r-simplex in the complex K, we will find a solution if one exists. So, if we choose $k + \Delta$, where Δ is the maximum degree of the Hasse graph, the search becomes tractable. In fact, we can even count the number of minimal solutions. We remark that the degree Δ of the Hasse graph H_{K} is bounded when the dimension of the complex is bounded and the number of incident cofacets on every simplex is bounded.

Algorithm 6 FPT Algorithm for Topological Hitting Set with $k + \Delta$ as the parameter

```
min ← |K|;  SOL = K;
for each r-simplex τ of K do
    Consider the set S_τ of all simplices within the graph distance 2k (in H_K) of τ.
    if a connected subset S ⊆ S_τ with |S| ≤ k is a hitting set of ζ and |S| < min then
        min = |S|;  SOL = S;
    end if
end for
if min < k then return SOL;
end if
```

Correctness.

The correctness of the algorithm immediately follows from Lemma 63.

Complexity.

Note that in Line 4 of Algorithm 6, we need to enumerate only the connected subsets S of cardinality less than or equal to k. We use Lemmas 64 and 65 by Fomin and Villanger [25] that provide very good bounds for enumerating connected subgraphs of graphs. First, we introduce some notation.

▶ Notation 6. The neighborhood of a vertex v is denoted by $\mathrm{nbd}(v) = \{u \in V : u, v \in E\}$, whereas the neighborhood of a vertex set $S \subseteq V$ is set to be $\mathrm{nbd}(S) = \bigcup_{v \in S} N(v) \setminus S$.

▶ **Lemma 64** ([25, Lemma 3.1]). *Let $G = (V, E)$ be a graph. For every $v \in V$, and $b, d \geq 0$, the number of connected vertex subsets $\mathcal{C} \subseteq V$ such that*

1. *$v \in B$,*
2. *$|B| = b + 1$, and*
3. *$|\mathrm{nbd}(B)| = d$*

is at most $\binom{b+d}{b}$.

▶ **Lemma 65** ([25, Lemma 3.2]). *All connected vertex sets of size $b + 1$ with f neighbors of an n-vertex graph G can be enumerated in time $O(n^2 \cdot b \cdot (b + d) \cdot \binom{b+d}{b})$ by making use of polynomial space.*

In Algorithm 6, $b = O(k)$ and $d = O(k\Delta)$. Therefore,

$$\binom{b+d}{b} = \binom{O(k\Delta)}{O(k)} \leq (k\Delta)^{O(k)} = 2^{O(k \log(k\Delta))}. \tag{6}$$

Hence, by Lemma 65, for a single r-simplex, the number of connected sets enumerated in Line 4 is $O(n^2 \cdot O(k) \cdot O(k\Delta) \cdot 2^{O(k \log(k\Delta))}) = O(n^5 \cdot 2^{O(k \log(k\Delta))})$ time. Since we do this for every r-simplex τ in K, the total time in enumerating all candidate sets in Lines 4-6 is at most $O(n^6 \cdot 2^{O(k \log(k\Delta))})$. Using Theorem 3, one can check if the set is a feasible solution in time $O(n^\omega)$, where ω is the exponent of matrix multiplication. Hence, the algorithm runs in $O(n^{6+\omega} \cdot 2^{O(k \log(k\Delta))})$ time, which is fixed parameter tractable in $k + \Delta$.

▶ **Theorem 66.** Topological Hitting Set *admits an FPT algorithm with respect to the parameter $k + \Delta$, where Δ is the maximum degree of the Hasse graph and k is the solution size. The algorithm runs in $O(n^{6+\omega} \cdot 2^{O(k \log(k\Delta))})$ time.*

7.1.1 Randomized FPT algorithm for Global Topological Hitting Set

Ostensibly, Global Topological Hitting Set looks a lot harder than Topological Hitting Set. However, this is not really the case. Fortunately, we can exploit the vector space structure of homology to design a randomized algorithm for Global Topological Hitting Set that uses the deterministic FPT algorithm for Topological Hitting Set as a subroutine.

Algorithm 7 Randomized FPT Algorithm for Global Topological Hitting Set with $k + \Delta$ as the parameter

1: Find the r-th homology basis of K. Denote the basis by $\mathcal{B}$. Here, $|\mathcal{B}| = \beta_r(\mathsf{K})$.
2: Arrange the cycles in $\mathcal{B}$ in a matrix. Denote the matrix by $\mathbf{B}$.
3: Let $\mathbf{x}$ be a uniformly distributed random binary vector of dimension $\beta_r(\mathsf{K})$.
4: With $\mathbf{B} \cdot \mathbf{x}$ as the input cycle, and $k + \Delta$ as the parameter, invoke Algorithm 6.

▶ **Proposition 67.** *The probability that a minimal topological hitting set of the cycle $\boldsymbol{B} \cdot \boldsymbol{x}$ is the optimal solution to* Global Topological Hitting Set *is at least* 1/2.

Proof. Note that the total number of nontrivial r-th homology classes of K is $2^{\beta_r(\mathsf{K})}$. Let $\mathcal{S}$ be a optimal solution to Global Topological Hitting Set. Then, because of the vector space structure of homology groups, the total number of nontrivial homology classes of $\mathsf{K}_{\mathcal{S}}$ is at most $2^{\beta_r(\mathsf{K})-1}$. In other words, $\mathcal{S}$ is a topological hitting set of at least $2^{\beta_r(\mathsf{K})-1}$ nontrivial classes. Let $\mathcal{C}$ be the set of r-th homology classes for which $\mathcal{S}$ is a topological hitting set. Then, the probability that a uniformly random homology class chosen by $\mathbf{B} \cdot \mathbf{x}$ belongs to $\mathcal{C}$ is at least $\frac{2^{\beta_r(\mathsf{K})-1}}{2^{\beta_r(\mathsf{K})}} = \frac{1}{2}$. ◀

From the proposition above, the following corollary follows immediately.

▶ **Corollary 68.** *Algorithm 7 is a randomized FPT algorithm for* Global Topological Hitting Set *with $k + \Delta$ as the parameter.*

7.2 FPT approximation algorithm for Boundary Nontrivialization

It turns out that we do not have a connectivity lemma analogous to Lemma 63 for Boundary Nontrivialization. For instance, consider the triangulation of a sphere as the input complex K, and let the boundary that needs to be made nontrivial be the equator of the sphere. Then, the two triangles at the north pole and the south pole constitute an optimal solution for Boundary Nontrivialization, as the removal of these triangles makes the boundary nontrivial. Clearly, the solution set consisting of these two triangles is not connected. Please refer to Figure 2 from Section 1. So it is not clear if there is an FPT algorithm for Boundary Nontrivialization with $k + \Delta$ as the parameter.

This motivates the search of another parameter that makes the problem tractable. To this end, we first make a few elementary observations.

▶ **Lemma 69.** *If there are two $(r+1)$-chains ξ and ξ' with b as a boundary, then their sum is an $(r+1)$-cycle. Also, if an $(r+1)$-chain ξ has b as a boundary, and ζ is an $r+1$-cycle, then $\xi + \zeta$ has b as a boundary.*

Proof. If $\partial\xi = b$ and $\partial\xi' = b$, then we have $\partial(\xi + \xi') = 0$.

Next, if $\partial\xi = b$ and $\partial\zeta = 0$, then we have $\partial(\xi + \zeta) = b$. ◀

In other words, the number of chains that have b as a boundary is precisely $2^{\dim \mathsf{Z}_{r+1}(\mathsf{K})}$. When the complex is $(r+1)$-dimensional, $\mathsf{Z}_{r+1}(\mathsf{K}) = \mathsf{H}_{r+1}(\mathsf{K})$. So, for $(r+1)$-dimensional complexes we provide an FPT approximation algorithm with $\beta_{r+1}(\mathsf{K})$ as a parameter. Let $\mathbf{B}$ be the $(r+1)$-th boundary matrix of complex K. The algorithm can be described as follows.

Algorithm 8 FPT approximation algorithm for Boundary Nontrivialization

1: $\mathbf{B}' \leftarrow \mathbf{B}$; $X = \{\}$;
2: **while** $\mathbf{B}' \cdot \mathbf{x} = \zeta$ has a solution **do**
3: $X \leftarrow X \cup \{\mathbf{x}\}$.
4: $Y \leftarrow$ the set of all chains generated from odd linear combinations of elements in X.
5: Let $\mathbf{Y}$ denote the matrix with the chains of Y as its columns.
6: Let $\mathcal{R}$ be the collection of row indices of $\mathbf{Y}$ and $\mathcal{C}$ be the collection of column indices.
7: In the natural way, interpret $\mathcal{R}$ as a collection of sets, $\mathcal{C}$ as a collection of elements, and $\mathbf{Y}$ as the incidence matrix between sets and elements.
8: Solve the Set Cover problem approximately for the instance described above using the greedy method [36, Chapter 2.1].
9: Let $\mathcal{S} \subseteq \mathcal{R}$ be the approximate solution for the setcover problem.
10: Let $\mathbf{B}'$ be the matrix formed by deleting from $\mathbf{B}$ the columns specified by the $r+1$-simplices in $\mathcal{S}$.
11: **end while**
12: **Return** $\mathcal{S}$.

▶ **Lemma 70.** *The algorithm terminates in $O(2^\beta \beta n \cdot \min(n, 2^\beta))$ time, where $\beta = \beta_{r+1}(\mathsf{K})+1$ and n is the number of simplices in K.*

Proof. First, we note that the algorithm terminates. This is because in each iteration of the while loop we add a vector $\mathbf{x}$ that is linearly independent to the vectors in set X. By Lemma 69, the number of iterations is bounded by $\beta = \beta_{r+1}(\mathsf{K}) + 1$. With an appropriate choice of data structures, Lines 4 and 5 can be executed in $O(2^\beta n)$ time. Note that the resulting matrix $\mathbf{Y}$ has $O(2^\beta)$ columns and $O(n)$ rows.

The most expensive step in the while loop is Line 8. A simple implementation of the greedy approximation algorithm for Set Cover runs in $O(2^\beta n \cdot \min(n, 2^\beta))$ time [15, Chapter 35.3]. ◀

▶ **Lemma 71.** *When the while loop terminates $\mathcal{S}$ covers every chain whose boundary is ζ. The set $\mathcal{S}$ returned at Line 11 provides an $O(\log n)$-factor approximation to* Boundary Nontrivialization.

Proof. This follows from the fact that after deleting some columns of $\mathbf{B}$ specified by set $\mathcal{S}$, if the loop terminates, then the new matrix $\mathbf{B}'$ has no solution to the equation $\mathbf{B}' \cdot \mathbf{x} = \zeta$. The algorithm provides an $O(\log n)$ factor approximation because we use the approximation algorithm for Set Cover as a subroutine in Line 8. ◀

Lemma 70 and Lemma 71 combine to give the following theorem.

▶ **Theorem 72.** Boundary Nontrivialization *has an $O(\log n)$-factor FPT approximation algorithm that takes bounding r-cycles as input on $(r+1)$-dimensional complexes, and runs in $O(2^\beta \beta n \cdot \min(n, 2^\beta))$ time, where $\beta = \beta_{r+1}(\mathsf{K}) + 1$ and n is the number of simplices in K.*

7.2.1 Randomized FPT approximation algorithm for Global Boundary Nontrivialization

As in the case of Global Topological Hitting Set in Section 7.1.1, we now exploit the vector space structure of the boundary group to design a randomized algorithm for Global

Boundary Nontrivialization that uses the deterministic FPT approximation algorithm for Boundary Nontrivialization as a subroutine.

Algorithm 9 Randomized FPT approximation algorithm for Global Topological Hitting Set with β as the parameter

1: Find a basis $\mathcal{B}$ for the column space of $\partial_{r+1}(\mathsf{K})$.
2: Arrange the bounding cycles from $\mathcal{B}$ in a matrix $\mathbf{B}$.
3: Let $\mathbf{x}$ be a uniformly distributed random binary vector of dimension $|\mathcal{B}|$.
4: With $\mathbf{B} \cdot \mathbf{x}$ at the input boundary, and β as the parameter, invoke Algorithm 8.

▶ **Proposition 73.** *Let $\mathcal{R}$ be a minimal set of simplices whose removal from K makes $\boldsymbol{B} \cdot \boldsymbol{x}$ nontrivial. The probability that $\mathcal{R}$ is the optimal solution to* Global Boundary Nontrivialization *is at least $1/2$.*

Proof. The total number of elements in the range of $\mathbf{B}$ is $2^{|\mathcal{B}|}$. Let $\mathcal{S}$ be a optimal solution to Global Boundary Nontrivialization. Suppose that $\mathbf{c}$ is a bounding cycle that is made nontrivial by removal of $\mathcal{S}$ from K. Then, $\mathbf{c} \notin \mathrm{im}(\partial_{r+1}(\mathsf{K}_{\mathcal{S}}))$. Suppose that $\mathbf{c}'$ is a bounding cycle that continues to be trivial following removal of $\mathcal{S}$ from K. That is, $\mathbf{c}' = \partial_{r+1}(\mathsf{K}_{\mathcal{S}}) \cdot \mathbf{x}$ for some $\mathbf{x}$. Then, $\mathbf{c} + \mathbf{c}' \notin \mathrm{im}(\partial_{r+1}(\mathsf{K}_{\mathcal{S}}))$, for otherwise, there would exist a vector $\mathbf{y}$ such that $\mathbf{c} + \mathbf{c}' = \partial_{r+1}(\mathsf{K}_{\mathcal{S}}) \cdot \mathbf{y}$, which gives $\mathbf{c} = \partial_{r+1}(\mathsf{K}_{\mathcal{S}}) \cdot (\mathbf{x} + \mathbf{y})$, a contradiction.

Now, assume that none of the bounding cycles in the basis $\mathcal{B}$ are made nontrivial by the removal of $\mathcal{S}$ from K. But that implies that any linear combination of cycles in $\mathcal{B}$ also belongs to $\mathrm{im}(\partial_{r+1}(\mathsf{K}_{\mathcal{S}}))$. This contradicts the existence of $\mathbf{c}$. So there exists at least one bounding cycle $\mathbf{b} \in \mathcal{B}$ which is made nontrivial by the removal of $\mathcal{S}$. Let $\mathcal{B}_{\mathcal{S}}$ be the subset of cycles in $\mathcal{B}$ that are made nontrivial by the removal of $\mathcal{S}$. Then, we have two cases:

Case 1: Suppose $\mathbf{b}$ is the only cycle in $\mathcal{B}_{\mathcal{S}}$.

Now, let $\mathbf{z}$ be any cycle that lies in the span of $\mathcal{B} \setminus \{\mathbf{b}\}$. By the argument above, $\mathbf{b} + \mathbf{z}$ is also made nontrivial by the removal of $\mathcal{S}$. So, the total number of bounding cycles that are made nontrivial by the removal of $\mathsf{K}_{\mathcal{S}}$ is $2^{|\mathcal{B}|-1}$.

Case 2: Suppose $\mathcal{B}_{\mathcal{S}} \setminus \{\mathbf{b}\}$ is nonempty.

Then, one obtains a new set of vectors $\mathcal{B}'$ from $\mathcal{B}$ as follows: For every $\mathbf{a} \neq \mathbf{b} \in \mathcal{B}$ such that $\mathbf{a}$ is made nontrivial by the removal of $\mathcal{S}$ and $\mathbf{a} + \mathbf{b}$ is in $\mathrm{im}(\partial_{r+1}(\mathsf{K}_{\mathcal{S}}))$, replace $\mathbf{a}$ by $\mathbf{a} + \mathbf{b}$. It is easy to check that $\mathcal{B}'$ is also a basis for the column space of $\partial_{r+1}(\mathsf{K})$. Moreover, if $\mathbf{z}'$ is any cycle that lies in the span of $\mathcal{B}' \setminus \{\mathbf{b}\}$, then $\mathbf{b} + \mathbf{z}'$ is also made nontrivial by the removal of $\mathcal{S}$. So, the total number of bounding cycles that are made nontrivial by the removal of $\mathsf{K}_{\mathcal{S}}$ is at least $2^{|\mathcal{B}|-1}$.

From the above analysis, we conclude that there are at least $2^{|\mathcal{B}|-1}$ bounding r-cycles that are made nontrivial by the removal of $\mathcal{S}$. Let $\mathcal{C}$ be the set of bounding r-cycles for which $\mathcal{S}$ is a Boundary Nontrivialization solution. Then, the probability that a uniformly random bounding cycle chosen by $\mathbf{B} \cdot \mathbf{x}$ belongs to $\mathcal{C}$ is at least $\frac{2^{|\mathcal{B}|-1}}{2^{|\mathcal{B}|}} = \frac{1}{2}$. ◀

From the proposition above, we obtain the following corollary immediately.

▶ **Corollary 74.** Global Boundary Nontrivialization *has an $O(\log n)$-factor randomized FPT approximation algorithm for r-th homology on $(r+1)$-dimensional complexes, with β as the parameter. The algorithm runs in $O(2^{\beta}\beta n \cdot \min(n, 2^{\beta}))$ time.*

8 Conclusion and Discussion

In this paper, we devise a polynomial time algorithm for Topological Hitting Set on closed surfaces. We believe that our algorithm should also easily generalize to surfaces with boundary.

Moreover, we show how certain cut problems generalize naturally from graphs to simplicial complexes, motivating a complexity theoretic study of these problems. For future work, it remains to be shown that Global Topological Hitting Set and Global Boundary Nontrivialization are also **W[1]**-hard. We believe that the **W[1]**-hardness reductions for Topological Hitting Set and Boundary Nontrivialization can be extended to establish hardness results for the global variants. Finally, a theoretical future direction of our work is to investigate how (the global variants of) Topological Hitting Set and Boundary Nontrivialization may be used to study high dimensional expansion in simplicial complexes [19, 29].

References

1 Ulrich Bauer, Abhishek Rathod, and Jonathan Spreer. Parametrized complexity of expansion height. In Michael A. Bender, Ola Svensson, and Grzegorz Herman, editors, *27th Annual European Symposium on Algorithms, ESA 2019, September 9-11, 2019, Munich/Garching, Germany*, volume 144 of *LIPIcs*, pages 13:1–13:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPIcs.ESA.2019.13.

2 Glencora Borradaile, Erin Wolf Chambers, Kyle Fox, and Amir Nayyeri. Minimum cycle and homology bases of surface-embedded graphs. *JoCG*, 8(2):58–79, 2017.

3 Glencora Borradaile, David Eppstein, Amir Nayyeri, and Christian Wulff-Nilsen. All-Pairs Minimum Cuts in Near-Linear Time for Surface-Embedded Graphs. In Sándor Fekete and Anna Lubiw, editors, *32nd International Symposium on Computational Geometry (SoCG 2016)*, volume 51 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 22:1–22:16, Dagstuhl, Germany, 2016. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. URL: http://drops.dagstuhl.de/opus/volltexte/2016/5914, doi:10.4230/LIPIcs.SoCG.2016.22.

4 Benjamin A. Burton, Sergio Cabello, Stefan Kratsch, and William Pettersson. The parameterized complexity of finding a 2-sphere in a simplicial complex. *SIAM J. Discret. Math.*, 33(4):2092–2110, 2019. doi:10.1137/18M1168704.

5 Benjamin A. Burton, Thomas Lewiner, João Paixão, and Jonathan Spreer. Parameterized complexity of discrete morse theory. *ACM Trans. Math. Softw.*, 42(1):6:1–6:24, 2016. doi:10.1145/2738034.

6 Benjamin A. Burton and William Pettersson. Fixed parameter tractable algorithms in combinatorial topology. In Zhipeng Cai, Alex Zelikovsky, and Anu G. Bourgeois, editors, *Computing and Combinatorics - 20th International Conference, COCOON 2014, Atlanta, GA, USA, August 4-6, 2014. Proceedings*, volume 8591 of *Lecture Notes in Computer Science*, pages 300–311. Springer, 2014. doi:10.1007/978-3-319-08783-2_26.

7 Oleksiy Busaryev, Sergio Cabello, Chao Chen, Tamal K. Dey, and Yusu Wang. Annotating simplices with a homology basis and its applications. In Fedor V. Fomin and Petteri Kaski, editors, *Algorithm Theory – SWAT 2012*, pages 189–200, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

8 Erin W. Chambers, Jeff Erickson, K. Fox, and A. Nayyeri. Minimum cuts in surface graphs. *ArXiv*, abs/1910.04278, 2019.

9 Erin W. Chambers, Jeff Erickson, Kyle Fox, and Amir Nayyeri. *Global Minimum Cuts in Surface-Embedded Graphs*, pages 852–856. Springer New York, New York, NY, 2016. doi:10.1007/978-1-4939-2864-4_683.

10 Erin W. Chambers, Jeff Erickson, and Amir Nayyeri. Homology flows, cohomology cuts. In *Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing*, STOC '09,

pages 273–282, New York, NY, USA, 2009. Association for Computing Machinery. URL: https://doi.org/10.1145/1536414.1536453, doi:10.1145/1536414.1536453.

11 Erin W. Chambers, Jeff Erickson, and Amir Nayyeri. Minimum cuts and shortest homologous cycles. In *Proceedings of the Twenty-Fifth Annual Symposium on Computational Geometry*, SCG '09, pages 377–385, New York, NY, USA, 2009. Association for Computing Machinery. URL: https://doi.org/10.1145/1542362.1542426, doi:10.1145/1542362.1542426.

12 Chao Chen and Daniel Freedman. Measuring and computing natural generators for homology groups. *Comput. Geom. Theory Appl.*, 43(2):169–181, February 2010. doi:10.1016/j.comgeo.2009.06.004.

13 Chao Chen and Daniel Freedman. Hardness results for homology localization. *Discrete & Computational Geometry*, 45(3):425–448, Apr 2011. doi:10.1007/s00454-010-9322-8.

14 F. R. K. Chung. *Spectral Graph Theory*. American Mathematical Society, 1997.

15 Thomas H Cormen, Charles E Leiserson, Ronald L Rivest, and Clifford Stein. *Introduction to algorithms*. MIT press, 2009.

16 Marek Cygan, Fedor V. Fomin, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michal Pilipczuk, and Saket Saurabh. *Parameterized Algorithms*. Springer, 2015. URL: http://dx.doi.org/10.1007/978-3-319-21275-3, doi:10.1007/978-3-319-21275-3.

17 Tamal K. Dey, Tianqi Li, and Yusu Wang. Efficient algorithms for computing a minimal homology basis. In Michael A. Bender, Martín Farach-Colton, and Miguel A. Mosteiro, editors, *LATIN 2018: Theoretical Informatics*, pages 376–398, Cham, 2018. Springer International Publishing.

18 Tamal K. Dey, Jian Sun, and Yusu Wang. Approximating loops in a shortest homology basis from point data. In *Proceedings of the Twenty-sixth Annual Symposium on Computational Geometry*, SoCG '10, pages 166–175, New York, NY, USA, 2010. ACM. doi:10.1145/1810959.1810989.

19 Dominic Dotterrer, Tali Kaufman, and Uli Wagner. On expansion and topological overlap. *Geometriae Dedicata*, 195(1):307–317, 2018. URL: https://doi.org/10.1007/s10711-017-0291-4, doi:10.1007/s10711-017-0291-4.

20 Rodney G. Downey and Michael R. Fellows. *Fundamentals of Parameterized Complexity*. Texts in Computer Science. Springer, 2013. URL: http://dx.doi.org/10.1007/978-1-4471-5559-1, doi:10.1007/978-1-4471-5559-1.

21 Art M. Duval, Caroline J. Klivans, and Jeremy L. Martin. Cuts and flows of cell complexes. *Journal of Algebraic Combinatorics*, 41(4):969–999, 2015. URL: https://doi.org/10.1007/s10801-014-0561-2, doi:10.1007/s10801-014-0561-2.

22 Jeff Erickson and Kim Whittlesey. Greedy optimal homotopy and homology generators. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '05, pages 1038–1046, Philadelphia, PA, USA, 2005. Society for Industrial and Applied Mathematics.

23 Michael R. Fellows, Danny Hermelin, Frances Rosamond, and Stéphane Vialette. On the parameterized complexity of multiple-interval graph problems. *Theoretical Computer Science*, 410(1):53–61, 2009. URL: https://www.sciencedirect.com/science/article/pii/S0304397508007329, doi:https://doi.org/10.1016/j.tcs.2008.09.065.

24 Fedor V Fomin, Daniel Lokshtanov, Saket Saurabh, and Meirav Zehavi. *Kernelization: Theory of Parameterized Preprocessing*. Cambridge University Press, 2019.

25 Fedor V. Fomin and Yngve Villanger. Treewidth computation and extremal combinatorics. *Comb.*, 32(3):289–308, 2012. URL: https://doi.org/10.1007/s00493-012-2536-z, doi:10.1007/s00493-012-2536-z.

26 Robert Ghrist and Sanjeevi Krishnan. A topological max-flow-min-cut theorem. In *2013 IEEE Global Conference on Signal and Information Processing*, pages 815–818, 2013. doi:10.1109/GlobalSIP.2013.6737016.

27 Allen Hatcher. *Algebraic topology*. Cambridge University Press, Cambridge, 2002.

28 Kristóf Huszár, Jonathan Spreer, and Uli Wagner. On the treewidth of triangulated 3-manifolds. *J. Comput. Geom.*, 10(2):70–98, 2019. doi:10.20382/jogc.v10i2a5.

29 Alexander Lubotzky. High dimensional expanders. In *Proceedings of the International Congress of Mathematicians: Rio de Janeiro 2018*, pages 705–730. World Scientific, 2018.

30 Clément Maria. Parameterized complexity of quantum knot invariants. In Kevin Buchin and Éric Colin de Verdière, editors, *37th International Symposium on Computational Geometry, SoCG 2021, June 7-11, 2021, Buffalo, NY, USA (Virtual Conference)*, volume 189 of *LIPIcs*, pages 53:1–53:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPIcs.SoCG.2021.53.

31 Clément Maria and Jonathan Spreer. A polynomial-time algorithm to compute turaev-viro invariants $TV_{4,q}$ of 3-manifolds with bounded first betti number. *Found. Comput. Math.*, 20(5):1013–1034, 2020. doi:10.1007/s10208-019-09438-8.

32 William Maxwell and Amir Nayyeri. Generalized max-flows and min-cuts in simplicial complexes. *arXiv preprint arXiv:2106.14116*, 2021.

33 James R. Munkres. *Elements of algebraic topology*. Addison-Wesley Publishing Company, Menlo Park, CA, 1984.

34 Abhishek Rathod. Fast Algorithms for Minimum Cycle Basis and Minimum Homology Basis. In *36th International Symposium on Computational Geometry (SoCG 2020)*, volume 164, pages 64:1–64:11, 2020.

35 Elena Farahbakhsh Touli and Yusu Wang. FPT-Algorithms for Computing Gromov-Hausdorff and Interleaving Distances Between Trees. In Michael A. Bender, Ola Svensson, and Grzegorz Herman, editors, *27th Annual European Symposium on Algorithms (ESA 2019)*, volume 144 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 83:1–83:14, Dagstuhl, Germany, 2019. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. doi:10.4230/LIPIcs.ESA.2019.83.

36 Vijay V Vazirani. *Approximation algorithms*. Springer Science & Business Media, 2013.