TAL TECH

**DOCTORAL THESIS**

# Computational Aspects of Rewriting in Higher-Dimensional Diagrams

Diana-Maria Kessler

TALLINNA TEHNIKAÜLIKOOL
TALLINN UNIVERSITY OF TECHNOLOGY
TALLINN 2025

TALLINN UNIVERSITY OF TECHNOLOGY
DOCTORAL THESIS
47/2025

# Computational Aspects of Rewriting in Higher-Dimensional Diagrams

DIANA-MARIA KESSLER

TAL TECH PRESS

TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies
Department of Software Science
**The dissertation was accepted for the defence of the degree of Doctor of Philosophy on 16 June 2025**

**Supervisor:** Doctor Amar Hadzihasanovic
Department of Software Science
School of Information Technologies
Tallinn University of Technology
Tallinn, Estonia

**Supervisor:** Professor Paweł Sobociński,
Department of Software Science
School of Information Technologies
Tallinn University of Technology
Tallinn, Estonia

**Opponents:** Professor Samuel Mimram
Laboratoire d'informatique
École Polytechnique
Paris, France

Professor Fabio Zanasi
Department of Computer Science
University College London
London, United Kingdom

**Defence of the thesis:** 27 June 2025, Tallinn

**Declaration:**
*Hereby I declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology, has not been submitted for any academic degree elsewhere.*

Diana-Maria Kessler _____

signature


ISSN 2585-6898 (publication)
ISBN 978-9916-80-330-1 (publication)
ISSN 2585-6901 (PDF)
ISBN 978-9916-80-331-8 (PDF)
DOI https://doi.org/10.23658/taltech.47/2025
Printed by EVG Print

Kessler, D.-M. (2025). *Computational Aspects of Rewriting in Higher-Dimensional Diagrams* [TalTech Press]. https://doi.org/10.23658/taltech.47/2025

TALLINNA TEHNIKAÜLIKOOL
DOKTORITÖÖ
47/2025

# Kõrgemamõõtmeliste diagrammide ümberkirjutamise arvutuslikud aspektid

DIANA-MARIA KESSLER

TAL TECH
KIRJASTUS

# Contents

# Abstract

Higher-dimensional rewriting is founded on the observation that certain rewrite systems correspond to directed cell complexes in different dimensions. This gives us a geometric grip on rewrite systems and rewrite derivations, connecting computational mathematics to higher categories and homotopy theory: rewrite systems are directed cell complexes and rewrite rules are directed homotopies.
In this thesis we adopt a computational perspective and view higher-dimensional rewriting as a model of computation in which an $n$-dimensional rewrite is a computation on $(n-1)$-dimensional data. Our motivation is the question of whether such a model of computation would be a feasible one.
With this in mind, we start by studying the computational complexity of building diagrams in the setting of diagrammatic sets. One of the main results is a *traversal algorithm* for solving the isomorphism problem in time $O(n^2 \log(n))$. We continue by studying the higher-dimensional subdiagram matching problem for rewritable subdiagrams. We provide an algorithm for solving this problem in arbitrary dimension as well as an upper bound on its running time. In the general case, the problem turns out to be in NP. We further show that under certain acyclicity conditions (that are satisfied by all diagrams of dimension less than or equal to 3), the running time of the algorithm is polynomial in the size of the diagram.
The diagrams mentioned so far correspond to pasting diagrams. We end the thesis with the study of acyclicity conditions in a more general framework of diagrams. We show that under one of the acyclicity conditions that we study for the subdiagram matching problem, the $\omega$-category presented by a diagram shape is freely generated in the sense of polygraphs. We further show that under stronger acyclicity conditions this $\omega$-category is equivalent to the one obtained from an augmented directed chain complex in the sense of Steiner, or consists only of subsets of cells in the diagram.



# Kokkuvõte

Kõrgemamõõtmeline ümberkirjutamine põhineb tähelepanekul, et teatud ümberkirjutussüsteemid vastavad suunatud rakukompleksidele erinevates mõõtmetes. See annab meile võimaluse ümberkirjutussüsteeme ja ümberkirjutustuletusi käsitleda geomeetriliselt, seostades arvutusmatemaatika kõrgemate kategooriate ja homotoopiateooriaga: ümberkirjutussüsteemid on suunatud rakukompleksid ja ümberkirjutusreeglid on suunatud homotoopiad.
Käesolevas töös võtame arvutusliku vaatenurga ja käsitleme kõrgemamõõtmelist ümberkirjutamist arvutusmudelina, kus $n$-mõõtmeline ümberkirjutamine on arvutamine $(n-1)$-mõõtmelistel andmetel. Meid ajendab küsimus, kas selline arvutusmudel võiks olla praktiline.
Kõigepealt uurime me diagrammide konstrueerimise arvutuslikku keerukust diagrammiliste hulkade kontekstis. Üks peamisi tulemusi on läbimisalgoritm isomorfismiprobleemi lahendamiseks $O(n^2 \log(n))$ ajas. Seejärel uurime kõrgemamõõtmelist alamdiagrammide sobitamise probleemi ümberkirjutatavate alamdiagrammide jaoks. Esitame algoritmi selle ülesande lahendamiseks suvalises mõõtmes koos ülemise tõkkega tema täitmiseks kuluvale ajale—üldjuhul kuulub see ülesanne

klassi NP. Lisaks näitame, et teatud mittetsüklilisuse tingimuste korral (mis on täidetud kõikide kuni 3-mõõtmeliste diagrammide puhul) on algoritmi täitmiseks kuluv aeg diagrammi suuruse suhtes polünomiaalne.
Diagrammid eeltoodus on kleepimisdiagrammid. Doktoritöö lõpuosas uurime tsüklivabaduse tingimusi üldisemate diagrammide jaoks. Näitame, et ühe tsüklivabaduse tingimuse korral, mida alamdiagrammide sobitamise probleemi jaoks uurime, on diagrammikujuga esitatud $\omega$-kategooria polügraafide mõttes vabalt genereeritud. Samuti näitame, et tugevamate tsüklivabaduse tingimuste korral on see $\omega$-kategooria kas ekvivalentne Steineri laiendatud suunatud ahelakompleksist saadud $\omega$-kategooriaga või koosneb ainult diagrammi rakkude alamhulkadest.

# List of Publications

This thesis includes material from articles I, II and III. As customary in mathematics, for articles I, II, III, and V the authors are listed in alphabetical order, and papers are assumed to be equal collaborations between all of the listed authors.
Article V started during the Adjoint School 2024 (where the author was a teaching assistant) and hence the students are listed in alphabetical order at the beginning and the mentor (Koko Muroya) and TA's (Diana Kessler and Juan F. Meleiro) are listed in alphabetical order at the end.
For article IV, Fabian Wiesner is the first author, Ziad Chaoui, Diana Kessler and Anna Pappa are in alphabetical order as second authors while Martti Karvonen is the senior author. We all contributed to the final manuscript.

I A. Hadzihasanovic and D. Kessler. "Data Structures for Topologically Sound Higher-Dimensional Diagram Rewriting". In: *Electronic Proceedings in Theoretical Computer Science* 380 (2023), pp. 111–127. DOI: 10.4204/eptcs.380.7

II A. Hadzihasanovic and D. Kessler. "Higher-dimensional subdiagram matching". In: *2023 38th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. IEEE. 2023, pp. 1–13. DOI: 10.1109/LICS56636.2023.10175726

III A. Hadzihasanovic and D. Kessler. "Acyclicity Conditions on Pasting Diagrams". In: *Applied Categorical Structures* 32 (Oct. 2024). DOI: 10.1007/s10485-024-09784-x

IV F. Wiesner, Z. Chaoui, D. Kessler, A. Pappa, and M. Karvonen. *Why quantum state verification cannot be both efficient and secure: a categorical approach.* Online preprint arXiv:2411.04767. 2024

V A. Matsui, I. Obi, G. Sabbagh, L. Torres, D. Kessler, J. F. Meleiro, and K. Muroya. "A Critical Pair Enumeration Algorithm for String Diagram Rewriting". In: *Applied Category Theory (ACT)* (2025). to appear

# Author's Contribution to the Publications

We refer to the list of publications on the previous page.

I The traversal algorithm and its runtime analysis are the joint work of the two authors. The author was the lead contributor on the complexity/runtime analysis aspects.

II The algorithms and runtime analysis of the algorithms presented in this paper are the joint work of the two authors. The author was the lead contributor on the complexity/runtime analysis aspects.

III The statements, proofs, and examples in this paper are the joint work of the two authors. The author was the lead contributor in the comparison between the regular directed complexes presented in this thesis and Steiner's augmented directed complexes.

IV The author contributed on the category theory part of the paper.

V The author contributed on the initial development of the algorithm as well as to the final structure and presentation of the paper.

# Acknowledgements

First and foremost, my thanks go to my supervisors.
To Amar Hadzihasanovic, to whom I am deeply grateful—for always having my back, for believing in me, and for being there when I was struggling. It was a pleasure to work with you and I learned a lot from our collaboration—not only how to think more deeply and clearly, but also how to turn creative ideas into concise mathematical ideas and research. Thank you for encouraging me to explore, to question, and to write and express my thoughts and ideas.
To Paweł Sobociński, for giving me the freedom to pursue this research project and for welcoming me into the beautiful group he created in Tallinn where I not only learned a lot but also formed lifelong friendships. It was a privilege to be part of this group and to see it grow.
Moreover, I want to thank both Amar and Paweł for supporting my travels to conferences and for giving me the chance to be involved in more than just the PhD work—learning how grant applications are written and how conferences are organised.
I am deeply grateful to Michael Johnson, whom I had the pleasure to meet at ACT 2023. Our lovely and lively discussions and your wonderful intuitions have shaped the way I view higher rewriting and its computational aspects which is especially reflected in the introduction and future work of this thesis. Thank you for the influence, the great stories I will always cherish and encouragement.
Special thanks to Jamie Vicary, for inviting me to visit his group in Cambridge and for his valuable advice and deep insights on the problem of obstructions.
To my PhD sister and dear friend Clémence Chanavat—thank you for the insightful discussions over the years. You have a beautiful mind, and I am looking forward to reading your thesis.
My PhD journey would not have been the same without our sister group, the Logic and Semantics Group. Many thanks to:
Tarmo Uustalu, for helping me settle into Estonia, for the career and research advice, and for the careful translation of the title and abstract of this thesis. Also, thank you for agreeing to be part of the defence committee.
Hellis Tamm, for the nice lunch breaks, your cheerful and caring spirit that always lifted me up, and the beautiful trips around Estonia with Toomas. Thank you for welcoming me into Estonia – it meant more than I can say. Also, many thanks for your help with the Estonian translation.
Monika Perkmann for your cheerful spirit, our fun chats and your advice.
To friends from both our group and the sister group:
Philipp Joram, for your kind and caring nature. You were always there to check in on me, remind me to rest, and gently pull me back when I was pushing too hard. I will always be grateful. I will always cherish our endless discussions and debates, the fun times we had in Tallinn and the excursions you organised.
Kat Zhuchko, for being an incredible friend on this journey—whether we were working in cafés, sharing brunches, or just chatting or debating. Thank you for always truly listening, for showing up when I needed someone, and for making even the most stressful moments feel lighter. Your steady presence and constant support meant more than I can say.

Michele De Pascalis, for noticing when I needed support (or snacks) and encouraging me to keep space for my hobbies. Thank you for sharing your knowledge—and passion—for food, coffee, and ale. I learnt so much from you and the way you brought our group together with weekend lunches made Tallinn feel more like home.
To Andrea Laretto, for the piano breaks, for bringing a much-needed dose of realism and for his dry humour that was often exactly what I needed.
To Elena Di Lavore and Mario Román, for sharing the ups and downs of our PhD journeys and for the unforgettable times at conferences. Thank you for the laughter and for being there in the tougher moments, too – I really appreciate it.
Elina Shakhnovich - your bubbly spirit, courage and perseverance are inspiring.
To Tanguy Massacrier, for your patience and quiet kindness, and the good times which made things feel more like home.
Matt Earnshaw and Eigil Fjeldgren Rischel, for sharing the (sometimes chaotic) process of writing our theses and for being there through the highs, lows, and everything in between. Matt, for starting the PhD together and for your support throughout the journey. Eigil, for the insightful math discussions, your matter-of-fact jokes, and gentle reminders not to overwork.
To Fosco Loregian for movie nights, engaging conversations and sharing his math knowledge.
To Ed Morehouse, for the beautiful conversations, for making difficult concepts feel intuitive and fun, and for being there with support and kindness during the difficult moments.
To Chad Nester - for always going the extra mile to welcome new members and make sure they feel part of the group.

To the rest of the two groups, for advice, coffee break laughs, and conversations about anything: Hendrik Maarand (thank you also for help with the Estonian translation), Nathanael Arkor (thank you for advice on the introduction), Niels Voorneveld (for the board game nights and agreeing to be part of the committee), Niccolò Veltri, Nathan Haydon, Vahur Kotkas, Maksym Bortin, Silvio Capobianco. And to the new members who enriched our community: Alkis Ioannidis, Callum Reader (thank you for agreeing to be part of the defence committee), Eva Graversen, Sofiya Taskova, Ioannis Andreou, Alessandro Di Giorgio (thank you for agreeing to be part of the defence committee) and Priyaa Srinivasan.

Many thanks to Bryce Clarke, Samuel Mimram and Fabio Zanasi for being part of the committee and agreeing to review this thesis - I appreciated your feedback.

I also want to thank TalTech for supporting some of my conference trips, and to the wonderful people working behind the scenes who made everything run smoothly (including, but not limited to): Ruth Laos, Ülle Ainsoo, Kaisa Saarman, Liisi Ilu, Ülle Lõhmus, Tiina Hagen, Elena Vaarmets, and especially Kristi Ainen, who always went the extra mile to make sure everything goes well for the group.

Additionally, I want to acknowledge the research collaborations that enriched my PhD experience: I am grateful to Martti Karvonen, Ziad Chaoui, and Fabian Wiesner, for the fun and fruitful collaboration on categorical cryptography. You made it a safe and inspiring space to ask questions and express ideas. Many thanks to the ACT Adjoint School, for the opportunity to participate, and to Koko Muroya, Juan Meleiro, Anna Matsui, Innocent Obi, Guillaume Sabbagh, and Léo Torres — I'm grateful we were all part of the same group and for our collaboration.

I am thankful to Chris Heunen who introduced me to this field and to Rik Sarkar who encouraged me to pursue this direction.
Furthermore, I would like to acknowledge the ARIA grant, which provided me with the opportunity to explore my research from a different perspective and to collaborate with researchers from this community.
Beyond research, I am grateful to my friends, old and new, especially Andreea Gavrilă, Alina Ion, Manali Sharma, Adelina Badea, Daniela Groza, Diana Tănase, Vlad Iordan, Raluca Nae, Justyna Vădan and Zhe Deng for supporting me during my PhD journey and not only.
I also want to acknowledge the teachers who shaped me over the years: my math and physics teachers — Doina Enache, George Cazacu, Mioara Ioniță, Florina Bărbulescu, and Viorica Stănescu — as well as others who left a deep and lasting imprint: Simona Kalamar, Angela Dogaru, Mihaela Stoenescu, Simona Mirea, Magdalena Bartoș, Gerard Enache, Iulia Pop, Liana Lungu, Denisa Berindei, Manuela Iana-Mihăilescu and Dragoș Mihăilescu.
I would not have been here without the support of my family. I want to thank my parents (Nicoleta and Mihai) for all the love, support and encouragement. Mulțumesc! Sunt norocoasă să vă am ca părinți. I also want to thank both my maternal and paternal grandparents for their immense love and kindness, for always being there for me and my parents and for the wonderful holidays and hikes at Oravița that turned my childhood from beautiful to magical.
I will always be grateful to my sister, Raluca, for her always cheerful and optimistic spirit and for her kind and quiet support - you have always managed to lift my spirits up.
I am also grateful to my extended family who contributed to where I am now, among whom I mention: Mihai Corciova (for his continuous encouragements and kind advice), Nela and Constantin Kessler, Raluca Manole, Cornelia and Lorenț Cocard.
And last but not least, to the person who had to put up with me and shared all the ups and downs of this journey – my husband, Ioan Stanciu. Thank you for always believing in me, not shying away from problems, but on the contrary brainstorming for solutions and for being there in every aspect of life. I'm more grateful than words can express.

# Chapter 1

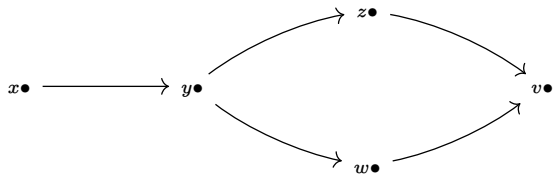
# Introduction

## 1.1 Higher rewriting

In the theory of computing, rewriting has been used as a mechanism of computation which relies on the idea of replacing a "subterm" in a formula with another "subterm" to obtain a new formula. This substitution is done according to some derivation rules called rewrites. Examples of models of computation that exhibit rewriting as a computation mechanism include Petri nets via transitions - see [43]- and $\lambda$-calculus via the $\beta$-reduction (see [21] for some examples). Rewriting theory has been applied in different branches of computer science, including term rewriting [3], graph transformation techniques [49, Chapters 3 and 4] [19, 46, 44, 45], rewriting logic [39, 13] as well as in the development of automated theorem provers [14, 35]. Rewriting theory also has applications in abstract algebra as a tool for studying the word problem in a monoid (or a group) or more generally as equational logic in universal algebras.

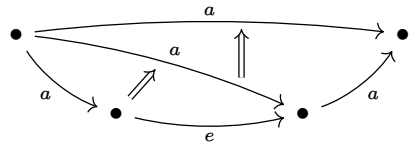
Higher-dimensional rewriting was founded on the observation that particular rewrite systems correspond to directed cell complexes in different dimensions. This gives us a geometric grip on rewrite systems and rewrite derivations: rewrite systems are directed cell complexes and rewrite rules are directed homotopies. Examples of higher-dimensional rewrite systems include presentations of higher categories, presentations of monoidal categories and presentations of (higher) operads. This notion was independently introduced under the names of "polygraphs" and "computads" by Burroni [11] and Street [52], respectively.

An example of a 1-dimensional rewrite system is an *abstract rewrite system* which can be represented by a directed graph corresponding to a 1-dimensional cell complex. Consider the following set of generators: $\{x, y, z, w, v\}$ with the following derivation rules: $\{x \to y, y \to z, y \to w, z \to v, w \to v\}$.

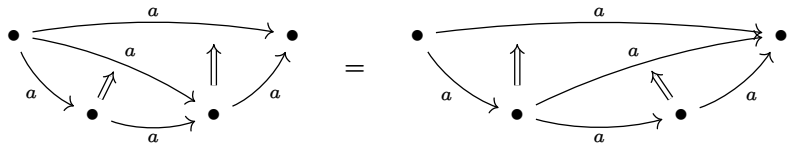



(1.1)

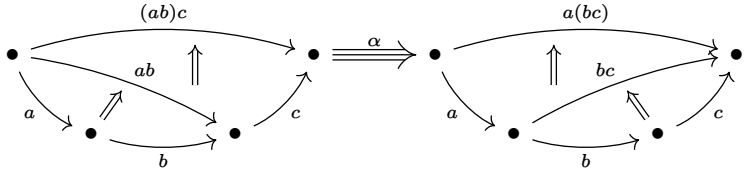
Figure 1.1 shows two different possible derivations from $x$ to $v$: according to the relations above, $x$ can be rewritten into $y$, $y$ into $z$ and so on. We are rewriting 0-dimensional elements and the rewrites are 1-dimensional. In other words, one 0-element can be turned into another if there is a path between them.
From here on, we may refer to individual elements of such rewrite systems as *cells*. The 1-dimensional example already hints at the idea of two elements being related if there is a path between them. We can make this intuition even more clear in the 2-dimensional case. By analogy with the first example, the elements being rewritten are composable configurations of 1-dimensional cells - which can be viewed as paths in a graph - and the rewrites are now 2-dimensional surfaces (see [32]). An example of a 2-dimensional rewrite system is a *string rewrite system* used for presentations of monoids or groups. Consider the example of a monoid generated by $\{a\}$ with relation $aa = a$. Now equality is a reversible relation that does not give us too many details about the computations going on. We can weaken the notion of equality and replace it by a (2-dimensional) rewrite rule: $aa \Rightarrow a$. This is a non-symmetric, irreversible operation that captures the idea of computation: turn an expression into another expression according to some rules given beforehand. The following picture represents an example of the unit law followed by an application of the rewrite rule (the orientation of the surface is given by the 2-dimensional arrow):




By replacing equality with (non-symmetric) directed transformations, one may study the word problem on monoids by using methods tailored to rewriting theory such as confluence and termination [22].
As examples for the 3-dimensional case we have *term-rewriting* [3] and rewriting on string diagrams. The terms are composable configurations of 2-dimensional elements which are themselves rewrites on 1-dimensional paths. Note how above we secretly replaced another equality relation with a rewrite rule: the unit law. But elements in a monoid have to obey one more identity law: associativity. In our 2-dimensional example, associativity can be represented by the following picture:



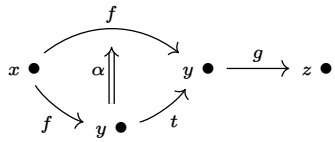

This is a relation between 2-dimensional surfaces. By the same argument presented in the 2-dimensional case, we can replace this equality with a 3-dimensional rewrite rule obtaining an example of an associator:



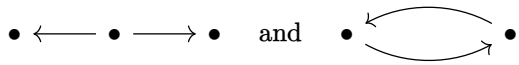
In this example, the set of generators contains one 0-dimensional element, $\bullet$, one 1-dimensional element, $a$, one 2-dimensional element that stands for composition and one 3-dimensional element which is the associator above.
Continuing in this way, one obtains an $n$-dimensional rewrite system in which the $(n-1)$-dimensional surfaces being rewritten are themselves rewrites on surfaces one dimension lower and so on.

## 1.2 Pasting diagrams

Pasting diagrams are a central tool for studying the composition of cells in higher-dimensional categories. The notion of 2-categorical pasting was introduced by Bénabou in his treatment of bicategories [6]; in the 1980s and 1990s, a number of frameworks for $n$-categorical pasting emerged, with corresponding *pasting theorems* guaranteeing that a pasting diagram admits a suitably unique composite [33, 47, 53, 51]; see [20] for a recent survey.
A pasting diagram is, informally, a *composable* configuration of cells in an $n$-category, such as the following:

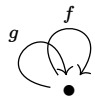



Examples of non-pasting diagrams include:

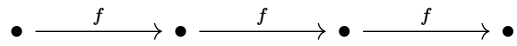



(1.2)

In the first example, the shape is not composable, while in the second one it is ambiguous what the composite should be.
We show on an example how the higher-dimensional pasting diagrams arise from the study of composition in a category. Take the following category (also a monoid in this case) presented by the diagram below:

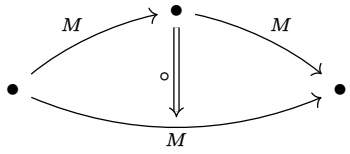



For a category theorist it is perfectly fine to represent $f \circ f \circ f$ in the loopy way pictured above. However, if we want to reason about composition in a category, the picture above is ambiguous. This becomes intractable as we move to higher-dimensional categories that may have higher-dimensional endo-arrows. To counter this problem, we need to have an "unloopy" way to refer to this composition. In the graphical representation of pasting diagrams, this becomes:

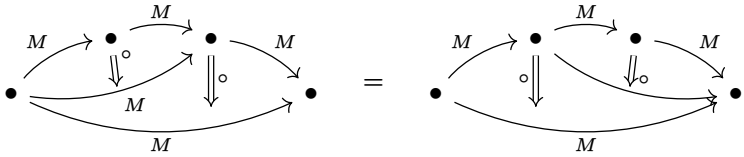


We can think of it as a parametrisation of the composition; the four dots and three arrows are the "parametrising" objects, while the labelling is the "parametrisation". Throughout this thesis, the parametrising object will be called the *shape of a diagram*, while the parametrisation will be called a *diagram*.

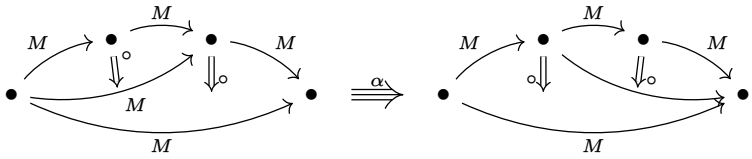
From another point of view, consider our monoid as a 1-object category. We can enrich this category over **Set** seen as a 1-object bicategory (the delooping of **Set** as a monoidal category) by sending $(\bullet, \bullet)$ to a 1-arrow, $M$, representing our monoid. Then the monoid operation $\circ : M \times M \to M$ in the bicategory can be represented by a 2-cell in the following fashion:




Now, to satisfy the equations of a monoid, for every quadruple of objects, (three endoarrows), the two ways in which we can compose have to be equal:

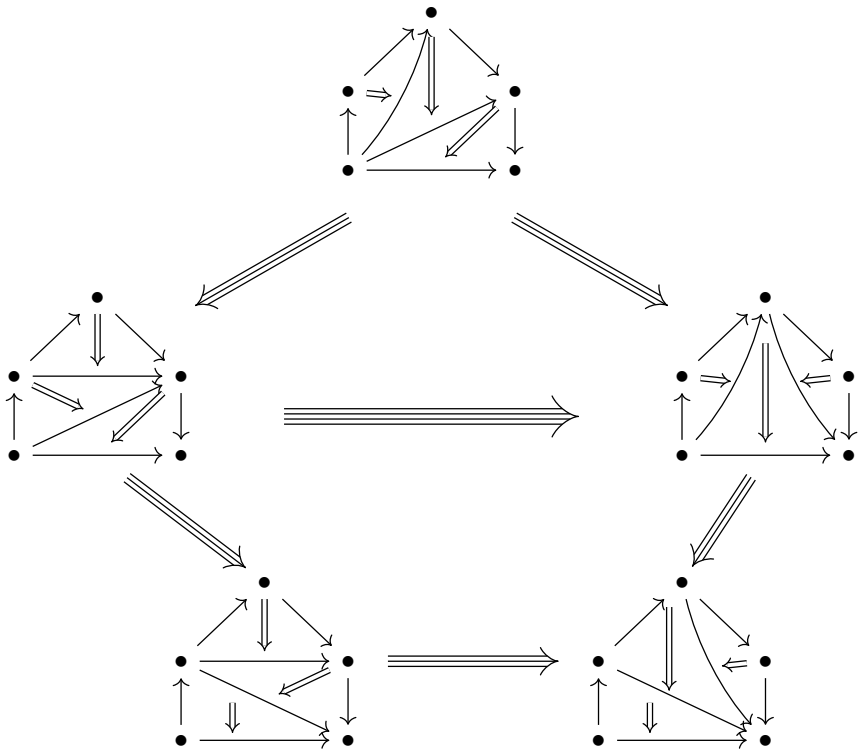



However, we are in a bicategory in which morphism composition is associative up to a natural isomorphism, $\alpha : (M \circ M) \circ M \to M \circ (M \circ M)$. This adds another dimension to our diagrams:

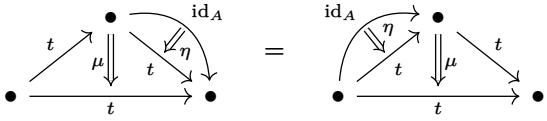



Since we do not have a strict equality between the two sides, this associator has to satisfy a further coherence condition, MacLane's pentagon, which is the 4-simplex [1]:

[1] On a slightly different note (and to make the connection with the previous paragraph), see [5] where the author shows that the equations for monoidal categories can be obtained by considering a convergent term rewrite system on the equational theory of monoids.



We have seen how the higher-dimensional pasting diagrams for associativity and MacLane's pentagon arised from the study of composition in a category enriched over a bicategory. The use of pasting diagrams has been leveraged for the study of equations in $n$-categories. For example, the following diagrams appear as the left hand side and right-hand side of the unit laws of a monad in a 2-category:

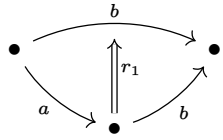



Though it does not constitute the topic of this thesis, pasting diagrams are also used in the study of weak higher categories, where the weak equalities are represented by higher-dimensional cells - see [12] for an overview.

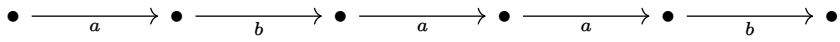
## 1.3 Higher-dimensional rewriting as a model of computation

In this thesis, we turn our attention to higher-dimensional rewriting seen as a model of computation. Namely, an $n$-dimensional rewrite is an $n$-cell in which the left hand side and right-hand side are composable configurations (pasting diagrams) of

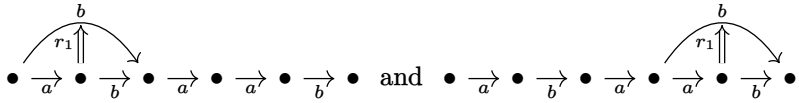
$(n-1)$-dimensional cells. As mentioned previously, these rewrites are not reversible and hence, an $n$-dimensional rewrite can be seen as a computation step on $(n-1)$-dimensional data.
One feature of this approach is the encoding of computations as data. More specifically, an $n$-dimensional computation whose boundary is made of $(n-1)$-dimensional data, can be part of the data of a computation one dimension higher. This allows us to encode programs as data and compare the simulations of various machines within the model.
A second advantage is that cells in a higher-dimensional diagram rewrite system represent the data, the computations and the space in which the data lives (and the computations take place). As a consequence, the operations that are part of a program can be "seen" explicitly; for example, certain steps that are usually hidden become now explicit, such as the operations of duplicating data or pointers management (allocation and deallocation) [10]. Moreover, the way in which data is processed or manipulated (e.g., the possibility of parallelism or accessing it as a stack or in free-order) becomes internal topological constraints.
The mechanism of computation or a basic computation step in our model is as follows: given a list of $(n+1)$-dimensional rewrite rules and a wider $n$-dimensional diagram, $d$, we want to find matches of the left-hand side of the rewrite rules into $d$. Once such a match is found, we apply the rewrite by substituting the input of the rewrite rule in $d$ with its output.
For example, in a string rewrite system with rule $r_1 \coloneqq ab \to b$, represented by the 2-dimensional diagram



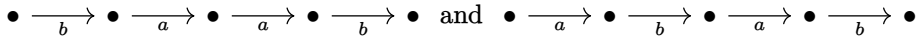

and the string *abaab* corresponding to the following 1-dimensional diagram, $d$:




there are two matches of the input of $r_1$ into $d$. The diagrams embodying these two substitutions are the following (respectively):

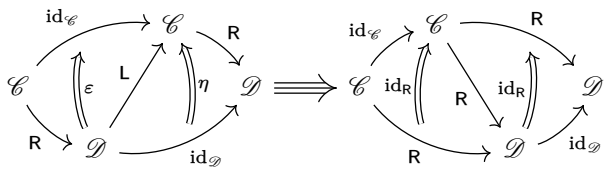



The strings resulted after the rewrite was applied are *baab* and *abab* corresponding to:




Now, the question that we are asking is: would a model of computation working in the way explained above be a feasible model of computation? More explicitly,

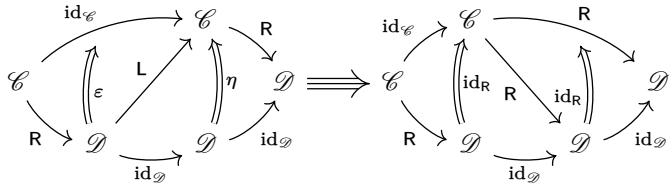
“ Is the obvious cost model that attributes constant cost to each rewrite step a “reasonable” cost model?”
To answer if such a machine would be a feasible model of computation, we start by studying the computational complexity of building the shapes of diagrams in the framework of diagrammatic sets introduced in [26]. This is a model of higher-dimensional rewriting that is topologically sound - i.e., diagrams admit a functorial interpretation as homotopies in CW complexes. The main contribution in this part is the algorithm for deciding whether two shapes of diagrams are isomorphic - Procedure 2.4.8. We show that this decision algorithm is correct (Theorem 2.4.19) and admits a low-degree polynomial time solution in the size of the diagram (Theorem 2.4.21). This is not a trivial result, since the more general problem of graph isomorphism has a quasipolynomial solution.
However, an inexpensive solution for building shapes of diagrams is not enough to establish if our model would be a feasible model of computation. This question is answered in the affirmative if and only if the *subdiagram matching problem* admits a low-degree polynomial time algorithm with respect to a reasonable size measure for diagrams. Since the problem of subgraph matching is NP-complete, the answer to this question is not obvious. Our approach to the higher-dimensional subdiagram matching problem is split in two main parts. Procedure 3.4.1, returns all inclusions of a shape of diagram into a larger shape of diagram; the algorithm has a low-degree polynomial time solution (Theorem 3.4.2). The second part is more difficult and deals with verifying whether such an inclusion is indeed a subdiagram inclusion. The solution to this problem (Procedure 4.2.1) involved developing more theory about pasting decompositions for our shapes of diagrams. The algorithm for the general case has a factorial running time and is in NP (Theorem 4.2.4). However, under certain acyclicity conditions (that all shapes of diagram of dimension less than or equal to 3 satisfy) the algorithm runs in linear time in the size of the data structure encoding the diagram - Proposition 4.3.11.
The acyclicity conditions give rise to nice results in the problem of subdiagram matching. In the last chapter of the thesis, we turn our attention to the study of acyclicity conditions of various strength in a more general setting of diagrams.
In the previous formalisms, the acyclicity conditions were used to exclude examples of non-pasting diagrams such as the “loopy” one in Figure 1.2. However, imposing such conditions resulted in cutting out commonly occurring shapes that appear in category theory already from dimension 3. One such example is the “weakened” form of one of the *triangle equations* in the theory of pseudoadjunctions of 2-categories pictured below:




(1.3)

Moreover, in the study of strict higher categories generated by diagram shapes, one might also want to consider non-pasting diagrams, such as the ones in Figure 1.2 in which case, imposing such acyclicity conditions would exclude the shape on the right.

In this thesis, the definitions for pasting diagram and the more general shapes of non-pasting diagrams that we are using were first introduced in [26, 24] inspired by Steiner's directed complexes [51]. They allow examples of pasting diagrams containing cycles, as the one exemplified above. One important property that all our diagrams satisfy is regularity which imposes that the input and output boundaries of a cell shape to be topologically closed balls.
At this point it is worth noting that (at least in the work of Johnson [33]) even though they were excluding the shape above as a "parametrising" object, their approach was to recover the behaviour we exemplified by a "parametrisation" of a different (acyclic) shape:




Apart from excluding "bad" examples, the acyclicity conditions served at least two more purposes:

- to guarantee that an $n$-category can be formed out of *subdiagrams*, or "composable subsets" of cells in the diagram;
- to ensure that the presented $n$-category is *freely generated* in the sense of computads or polygraphs [52, 11, 1].

We continue the study of the acyclicity conditions of various strength to answer the two questions above in the more general framework of *regular directed complexes* which are diagrams that are locally pasting diagrams with one greatest element (and which include both shapes from Figure 1.2). We investigate the connection between our framework and Steiner's work on augmented directed complexes [50] and study the stability of the acyclicity conditions under the operations of joins, Gray products and suspensions.

## 1.4 Related work

Similar work studying the computational complexity aspects of higher rewriting is present in [15] and [10]. In [15], the authors present the data structures and algorithms for rewriting in the homotopy.io proof assistant [48] which is based on the theory of associative $n$-categories developed by Dorn, Douglas and Vicary [18]. [2] In [10], the underlying framework is based on polygraphs but only up to dimension 2.
The works of Bonchi, Gadducci, Kissinger, Sobocinski, and Zanasi [7, 8, 9] investigate string diagram rewriting, a form of three-dimensional rewrite system. Their

[2] Both diagrammatic sets and associative n-categories are frameworks for semi-strict higher categories. However, the point of divergence is that associative $n$-categories have strict units and associators and weak interchangers while diagrammatic sets have strict associators and interchangers and weak units.

approach involves encoding string diagrams as hypergraphs, with rewriting subsequently carried out using the double pushout (DPO) method [19], a well-established technique in graph transformations. These studies establish a framework for certain types of rewrite systems that include additional equations, such as those with a Frobenius structure or a symmetric monoidal structure, and explore rewriting-theoretic questions, such as confluence and termination, in relation to these systems. A more computational complexity approach to string diagram rewriting is taken by Vicary and Delpeuch [16, 54].
The research in [36, 23, 42] focuses on the computational properties of rewrite systems, such as confluence, termination or computing critical pairs in polygraphs . Another notable contribution is by Plump [44, 45], who investigates a distinct computational model known as term graphs. Term graphs, which are defined on hypergraphs in a manner similar to the approach in [7, 8, 9], offer a framework for rewriting on two-dimensional data. This work provides an alternative model of computation, expanding on the concept of hypergraphs [46].

## 1.5 Structure of the thesis

In the second chapter, we introduce the basic structure that we use to encode our shapes of diagrams, *oriented graded posets*. We give the inductive definition of *molecules* which are our notion of pasting diagrams and prove some fundamental results about them. We continue by defining two notions of graphs -*flow graphs* and *graphs with open edges* - that will be used as tools in some of the proofs and can also aid in visualizing the behavior of molecules in the higher dimensions. The chapter concludes with a presentation of the basic data structures and algorithms for constructing molecules, with the primary contribution being an *isomorphism algorithm* for molecules (Procedure 2.4.8) which admits a low-degree polynomial time solution in the size of the diagram (Theorem 2.4.21).
In the third chapter, we define the problem of higher-dimensional subdiagram matching, which can be broken down into three subproblems, with the last one having a trivial solution in our framework. We introduce the definitions of *submolecules* and *layerings*, where a *layering* is a special kind of pasting decomposition of molecules. We end the chapter with the molecule matching algorithm (Procedure 3.4.1), the first step in the problem of higher-dimensional subdiagram matching.
The fourth chapter focuses on the second step of the subdiagram matching problem: the rewritable submolecule problem. This problem is more complex and requires additional results related to layerings. We establish a connection between layerings and specific orderings of the top-dimensional elements of a molecule, and use this link to determine when the inclusion of molecules corresponds to a submolecule inclusion - Theorem 4.1.20. We prove that the solution to the subdiagram matching problem in the general case has a factorial running time (Theorem 4.2.4). We present the acyclicity conditions on molecules under which the runtime of this algorithm is improved. Moreover, we show that all molecules of dimension less than or equal to 3 satisfy a certain acyclicity condition for which the runtime of this algorithm is linear - Proposition 4.3.11.
The last chapter presents the $\omega$-categorical framework of our diagrams. We give the definition of regular directed complexes (a more general framework for diagrams).

We show that the set of isomorphism classes in the slice category of molecules over an oriented graded poset, $P$, has the structure of an $\omega$-category (Proposition 5.1.17). In a more constrained version of an oriented graded poset with frame-acyclic molecules, we show that this category admits the structure of a polygraph (Theorem 5.1.30). We show that for an even stronger acyclicity condition for the molecules of a regular directed complex, $P$, this category consists only of subsets of $P$ (Corollary 5.4.27). We study the connection between our framework and Steiner's theory on augmented directed complexes. Finally, we end this chapter by studying the stability of these acyclicity conditions under the operations of suspensions, joins and Gray products.

## 1.6 Main contributions

This thesis is based on work published in [28, 29, 27].

- The traversal algorithm (Procedure 2.4.8) together with its runtime analysis (Theorem 2.4.21).
- The molecule matching algorithm (Procedure 3.4.1) and its runtime analysis (Theorem 3.4.2).
- The rewritable submolecule algorithm (Procedure 4.2.1), its runtime analysis (Theorem 4.2.4) and the improvement of the running time when the molecules are stably-frame acyclic (Proposition 4.3.11).
- The construction of the functor from the category of oriented graded posets to $\omega$-cat (Proposition 5.1.17); showing that the image of an oriented graded poset with frame-acyclic molecules has the structure of a polygraph (Theorem 5.1.30). Studying the relation between our setting and the one of Steiner's augmented directed complexes (Proposition 5.4.10 and Theorem 5.4.15).

## 1.7 Author's note

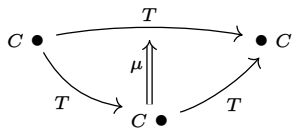
This thesis makes extensive use of pasting diagrams. This is on purpose. Apart from the fact that the author believes that visualising is a great way to gain the desired intuition, the author finds that the literature lacks examples in the form of pasting diagrams. Usually, most examples of pasting diagrams include commonly occurring shapes - i.e., simplices, cubes etc - that rarely go as high as dimension 3. However, these shapes are not general enough to model what we want. For this reason, this thesis will include various examples of diagrams that do not fall into these categories of commonly used examples.
To let the reader get accustomed to this graphical representation, the shapes of diagrams that we will be using will be introduced along with the theory used to encode them. Moreover, most of the times, a combinatorial or algebraic intuition or description will be provided along with the diagram.

# Chapter 2

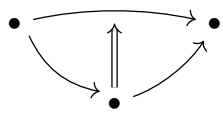
# Oriented graded posets and molecules

We split the information representing a diagram in two. A diagram is formed out of its shape and a labelling of its elements into a set of variables. Shapes of diagrams represent our notion of pasting diagrams - i.e., composable configurations of cells in an $n$-category. We will give the precise definition for a shape of diagram in the section about molecules and a categorical definition for diagrams in a strict $\omega$-category in 5.1.23.

For the moment we concern ourselves with how we can encode the information of a diagram. For example, take the following diagram which may correspond to monad multiplication in the 2-category of categories, functors and natural transformations.




As already mentioned, we can split the information in the picture above into two; The first is the shape of the diagram :



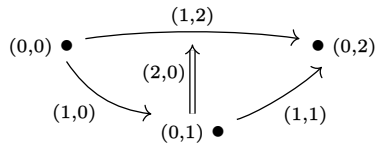
The second is the labeling of its elements:

$$\begin{array}{lcl} 0-\text{dimensional} \\ \text{elements} & \longmapsto & C \\ \\ 1-\text{dimensional} \\ \text{elements} & \longmapsto & T \\ \\ 2-\text{dimensional} \\ \text{elements} & \longmapsto & \mu \end{array}$$

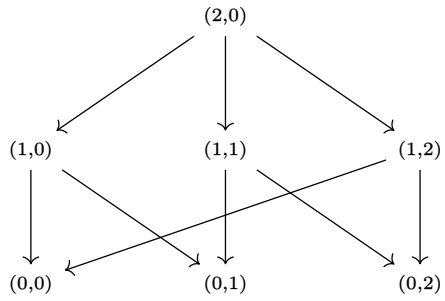
A shape of a diagram together with a labelling is a diagram. We label the elements of a shape of diagram from a set of generating cells. When dealing with shapes of diagram and their computational aspects, we want to refer to each element individually. Each such element will be written in the form $(n, k)$, where:

- $n$ is the dimension of the element.
- $k$ is the position of the elements in the *traversal order* - see Procedure 2.4.2.

So, when reasoning about shapes of diagrams, we think of them as:

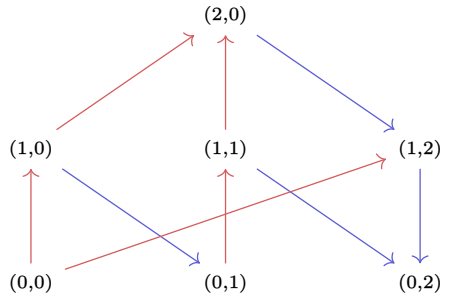



The unoriented version of our shapes of diagrams are regular cell complexes. We take the approach from combinatorial topology and encode the shape of a diagram using its face poset with extra structure to account for the directedness of our cells. This structure is called an *oriented graded poset* and is represented with a Hasse diagram with oriented edges. The poset encoding the shape above, without taking account of the direction of the elements is:




How does this work? First of all, note that the elements of each dimension are on a level. This diagram has dimension 2, and so the Hasse diagram representation has

3 levels: on the bottom level, are the 0-dimensional elements, on the middle level are the 1-dimensional elements and on the last level are the 2-dimensional elements. We call a poset with this property *graded*. The poset structure records whether an $n$-dimensional element is in the boundary of an $(n+1)$-dimensional element. For example, the 0-dimensional elements $(0,0)$ and $(0,1)$ are in the boundary of the 1-dimensional element $(1,0)$ so there is an edge between $(0,0)$ and $(1,0)$ and between $(0,1)$ and $(1,0)$. We mentioned previously that the boundary of an $n$-dimensional rewrite splits into an input $(n-1)$-dimensional half and an output $(n-1)$-dimensional half. We can encode the membership of an element to one of the faces by assigning an orientation(or colour) to the edges to indicate whether an $n$-dimensional element is in the input boundary of an $(n+1)$-dimensional element (arrows pointing upwards) or in the output boundary of an $(n+1)$-dimensional element (arrows pointing downwards).




For example, (0, 0) is in the input boundary of (1, 0) – arrow pointing upwards or red – and (0, 1) is in the output boundary of (1, 0) – arrow pointing downwards or blue – while for (2, 0), it has (1, 0) and (1, 1) in its input and (1, 3) in its output. Oriented graded posets are the basic data structure that we use in describing our shapes of diagrams. Since we use these structures to encode shapes of diagrams, for the purpose of this thesis we are only considering finite posets. However, their definition does not exclude the non-composable shapes. Our notion of pasting diagrams is captured by *molecules*, which are an inductively defined subclass of oriented graded posets. By defining molecules this way, we eliminate the "bad" examples like those mentioned in the introduction. The combinatorial nature of the definition makes it both straightforward to implement and easy to manipulate by a computer. In the framework of rewriting as computations we are interested in a model of computation on molecules and hence we are interested in the computational cost of building such molecules which was the main focus of [28]. It turns out that most of the operations that are necessary to build these molecules have straightforward low-degree polynomial time complexity, with one exception: the isomorphism problem for molecules. We show that the isomorphism problem extended to all oriented graded posets is equivalent to the graph isomorphism problem. However, due to the rigid nature of molecules, the isomorphism problem restricted only to shapes of diagrams can be solved in time $O(n^2 \log n)$, where $n$ is the number of vertices and edges in the Hasse diagram representation of a molecule (Theorem 2.4.21).

We begin by introducing the fundamental definitions for the structure used to encode the shapes of diagrams. Next, we define molecules and conclude the chap-

ter by presenting the algorithms and data structures used in constructing these molecules. Specifically, the final (sub)section focuses on illustrating on an example the algorithm used to solve the molecule isomorphism problem.

## 2.1 Basic definitions

***Definition* 2.1.1** (Covering relation) — Let $P$ be a poset. Given elements $x, y \in P$, we say that $y$ *covers* $x$ if $x < y$ and, for all $y' \in P$, if $x < y' \leq y$ then $y' = y$.

Many operations described from here on are operations on closed subsets of posets. Also, many times when thinking about elements of a diagram, we think about the element together with its "boundary" or the elements that are beneath it.

***Definition* 2.1.2** (Closure of a subset) — Let $P$ be a poset and $U \subseteq P$. The *closure of* $U$ is the subset of $P$

$$\operatorname{cl} U \coloneqq \{x \in P \mid \text{there exists } y \in U \text{ such that } x \leq y\}.$$

We say that $U$ is *closed* if $U = \operatorname{cl} U$.

We encode the covering relation (whether an $n$-dimensional element is related to an $(n+1)$-dimensional element) using faces. The dual notion is the one of cofaces; we define both of them below:

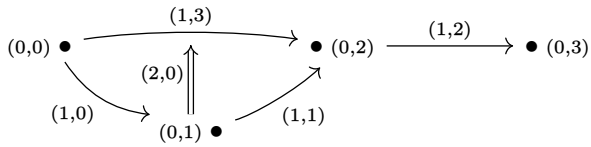
***Definition* 2.1.3** (Faces and cofaces) — Let $P$ be a poset and $x \in P$. The set of *faces* of $x$ is

$$\Delta x \coloneqq \{y \in P \mid x \text{ covers } y\}$$

and the set of *cofaces* of $x$ is

$$\nabla x \coloneqq \{y \in P \mid y \text{ covers } x\}.$$

***Example* 2.1.4** — Take the following shape as an example.




Then, $\Delta(2,0) = \{(1,0),(1,1),(1,3)\}$, $\Delta(0,2) = \emptyset$, $\nabla(0,2) = \{(1,1),(1,2),(1,3)\}$ and $\nabla(2,0) = \emptyset$.

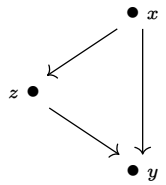
***Definition* 2.1.5** (Hasse diagram) — Let $P$ be a poset. The *Hasse diagram* of $P$ is the directed acyclic graph $\mathscr{H}P$ whose

- set of vertices is the underlying set of $P$, and
- set of edges is
$$\{(y, x) \mid x \in \Delta y\},$$
with $s\colon (y,x) \mapsto y$ called the source and $t\colon (y,x) \mapsto x$ called the target.

***Definition* 2.1.6** (Maximal path) — Let $P$ be a poset. A maximal path starting at $x \in P$ is a sequence $x_1 < x_2 < \ldots < x_n = x$ of elements of $P$ such that for all $i \in \{1, \ldots, n-1\}$ if $x_i \leq y \leq x_{i+1}$ then $y = x_i$ or $y = x_{i+1}$ and if $y \leq x_1$, then $y = x_1$. We call $n$ the length of this path.

***Definition* 2.1.7** (Graded poset) — A poset $P$ is *graded* if, for all $x \in P$, there exists $n$ such that all maximal paths in $\mathrm{cl}\{x\}$ have the same finite size $n$.

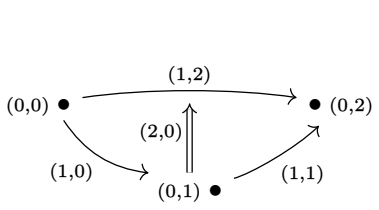
We already saw what being graded means and how to intuitively calculate the dimension of an element. The poset below is not a graded poset. Element $x$ has height 2 when taking the path going through $z$ and height 1 when taking the other path.

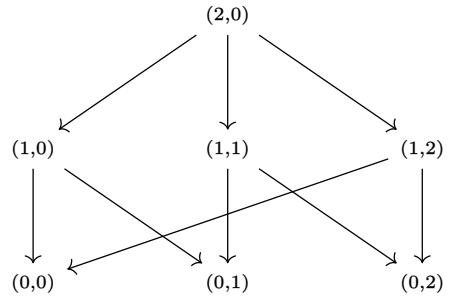



Now that we defined the basic structure we will be working with, we can proceed to further define the properties that were used at the beginning of this chapter but not articulated properly.

***Definition* 2.1.8** (Dimension of an element) — Let $P$ be a graded poset and $x \in P$. The *dimension* of $x$ is the length $\dim x$ of a maximal path starting from $x$ in $\mathscr{H}P$.

***Example* 2.1.9** — Let $U$ be the following shape of diagram, together with its Hasse diagram representation:




As discussed in the introduction, the dimension of $(1,0)$ is 1 and there are two maximal paths from $(1,0)$ - one to $(0,\ 0)$, another one to $(0,\ 1)$ - both having length 1. Similarly, for $(2,0)$: there are six maximal paths starting at $(2,0)$, all having length 2.

***Definition* 2.1.10** (Dimension of a subset) — Let $U$ be a closed subset of a graded poset. The *dimension* of $U$ is the integer

$$\dim U \coloneqq \begin{cases} \max\{\dim x \mid x \in U\} & \text{if } U \text{ is inhabited,} \\ -1 & \text{if } U \text{ is empty.} \end{cases}$$

***Example* 2.1.11** — Consider the graded poset from example 2.1.9 and let $U = \mathrm{cl}\{(1,0),(1,1)\} = \{(0,0),(1,0),(0,1),(1,1),(0,2)\}$. Then $\dim U = 1$ , since $(1,0)$ and $(1,1)$ are the elements with the greatest dimension.

***Definition* 2.1.12** (Maximal element) — Let $P$ be a poset, $x \in P$. We say that $x$ is *maximal in* $P$ if, for all $y \in P$, if $x \leq y$ then $x = y$. We write $\mathscr{M}ax\, P$ for the set of maximal elements in $P$.

We make a distinction between maximal elements and *top-dimensional* elements. The shape from Example 2.1.4 has two maximal elements - $(2,0)$ and $(1,2)$- but only one top-dimensional element: $(2,0)$. That is, if $x$ is an element in a closed subset $U$ of a graded poset $P$, then $x$ is top-dimensional if $\dim x = \dim U$.
Throughout the thesis we will want to talk about the $n$-dimensional elements of a subset.

***Definition* 2.1.13** (Grading of a subset) — Let $U$ be a subset of a graded poset. For each $n \in \mathbb{N}$, we write $U_n \coloneqq \{x \in U \mid \dim x = n\}$. We have $U = \sum_{n\in\mathbb{N}} U_n$.

***Example* 2.1.14** — If we let $U$ be the shape of diagram from Example 2.1.9 seen as a graded poset, then $U_0 = \{(0,0),(0,1),(0,2)\}$, $U_1 = \{(1,0),(1,1),(1,2)\}$ and $U_2 = \{(2,0)\}$.

We call a graded poset whose maximal elements are top-dimensional elements *pure*:

***Definition* 2.1.15** (Pure subset) — Let $U$ be a closed subset of a graded poset, $n \coloneqq \dim U$. We say that $U$ is *pure* if all the maximal elements of $U$ have dimension $n$, that is, $\mathscr{M}ax\, U = U_n$.

Finally, Definition 2.1.16 gives us the structure that we need to encode our shapes of diagrams.

***Definition* 2.1.16** (Orientation on a graded poset) — Let $P$ be a graded poset. An *orientation* on $P$ is an edge-labelling of $\mathscr{H}P$ with values in $\{+,-\}$.

***Definition* 2.1.17** (Oriented graded poset) — An *oriented graded poset* is a graded poset $P$ together with an orientation on $P$.

The orientation of the poset allows to partition the sets of faces and cofaces into sets of input and output faces and cofaces.

***Definition* 2.1.18** (Input and output faces and cofaces) — Let $P$ be an oriented graded poset and $x \in P$. The set of *input faces* of $x$ is

$$\Delta^- x \coloneqq \{y \in P \mid x \text{ covers } y \text{ with orientation } -\}$$

and the set of *output faces* of $x$ is

$$\Delta^+ x \coloneqq \{y \in P \mid x \text{ covers } y \text{ with orientation } +\}\,.$$

Dually, the set of *input cofaces* of $x$ is

$$\nabla^- x \coloneqq \{y \in P \mid y \text{ covers } x \text{ with orientation } -\}$$

and the set of *output cofaces* of $x$ is

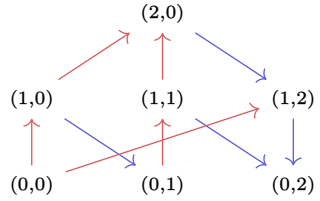
$$\nabla^+ x \coloneqq \{y \in P \mid y \text{ covers } x \text{ with orientation } +\}\,.$$

We have $\Delta x = \Delta^+ x \cup \Delta^- x$ and $\nabla x = \nabla^+ x \cup \nabla^- x$.

***Definition* 2.1.19** (Oriented Hasse diagram) — Let $P$ be an oriented graded poset. The *oriented Hasse diagram of* $P$ is the directed graph $\vec{\mathscr{H}}P$ whose

- set of vertices is the underlying set of $P$, and
- set of edges is
$$\{(y,x) \mid y \in \Delta^- x \text{ or } x \in \Delta^+ y\},$$
with $s\colon (y,x) \mapsto y$ called the source and $t\colon (y,x) \mapsto x$ called the target.

***Example* 2.1.20** — The oriented Hasse diagram of the shape of diagram from Example 2.1.9 is:



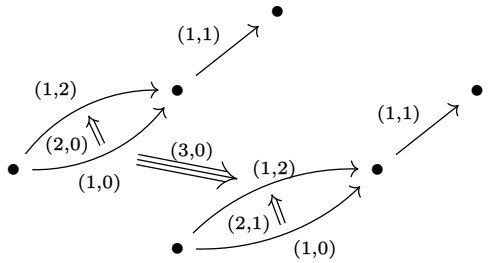

***Definition* 2.1.21** (Input and output $n$-boundaries) — Let $U$ be a closed subset of an oriented graded poset. For all $\alpha \in \{+,-\}$ and $n \in \mathbb{N}$, let
$$\Delta^\alpha_n U \coloneqq \{x \in U_n \mid \nabla^{-\alpha} x \cap U = \varnothing\}.$$
For each $n \in \mathbb{N}$, the *input $n$-boundary* of $U$ is the closed subset
$$\partial^-_n U \coloneqq \mathrm{cl}\,(\Delta^-_n U) \cup \bigcup_{k<n} \mathrm{cl}\,(\mathscr{M}ax\, U)_k$$
and the *output $n$-boundary* of $U$ is the closed subset
$$\partial^+_n U \coloneqq \mathrm{cl}\,(\Delta^+_n U) \cup \bigcup_{k<n} \mathrm{cl}\,(\mathscr{M}ax\, U)_k.$$
For $n < 0$, we let $\Delta^\alpha_n U = \partial^\alpha_n U \coloneqq \varnothing$.

***Definition* 2.1.22** (Notation for boundaries) — We will use the following notations, for $x$ an element in an oriented graded poset, $U$ a closed subset, $n \in \mathbb{N}$, and $\alpha \in \{+,-\}$:
$$\partial^\alpha_n x \coloneqq \partial^\alpha_n \mathrm{cl}\,\{x\}, \qquad \partial_n U \coloneqq \partial^-_n U \cup \partial^+_n U, \qquad \Delta_n U \coloneqq \Delta^-_n U \cup \Delta^+_n U.$$
Note that $\Delta^\alpha_k U$ are not closed subsets of oriented graded posets as they only contain $k$-dimensional elements. However, the sets $\partial^\alpha_n U$ are closed. We will later see that when $U$ is a molecule, $\partial^\alpha_k U$ are themselves molecules for any $k \geq 0$.

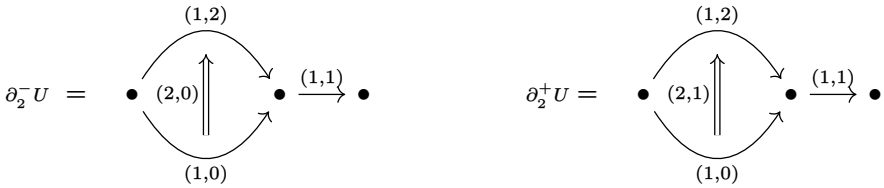
***Example* 2.1.23** — Consider the shape from Example 2.1.4. Then, $\Delta^-_0 = \{(0,0)\}$, $\Delta^+_0 = \{(0,3)\}$, $\Delta^-_1 = \{(1,0),(1,1),(1,2)\}$, $\Delta^+_1 = \{(1,3),(1,2)\}$. For the boundaries: $\partial^-_1 = \{(1,0),(1,1),(1,2),(0,0),(0,1),(0,2),(0,3)\}$, $\partial^+_1 = \{(1,3),(1,2),(0,0),(0,2),(0,3)\}$, $\partial^-_0 = \{(0,0)\}$, $\partial^+_0 = \{(0,3)\}$. We represent graphically the input and output 1-dimensional boundaries of $U$:
$$\partial^-_1 U = \ (0,0) \bullet \xrightarrow[(1,0)]{} (0,1) \bullet \xrightarrow[(1,1)]{} \bullet\ (0,2) \xrightarrow[(1,2)]{} \bullet\ (0,3)$$
$$\partial^+_1 U = \ (0,0) \bullet \xrightarrow[(1,3)]{} \bullet\ (0,2) \xrightarrow[(1,2)]{} \bullet\ (0,3)$$

The example above does not illustrate why the definition of boundaries includes the union of the closure of the maximal elements. Below, we present a 3-dimensional example in which this condition is used.

***Example* 2.1.24 —** Let $U$ be the following shape of diagram:




We omitted to add the labels for the 0-cells to not overcrowd the image, but on both faces they are - from left to right- $(0,0),(0,1),(0,2)$.
Note that the 1-cell, $(1,1)$, is not in the closure of $(3,0)$. We drew the diagram slightly tilted to help visualise that. More formally, $\Delta_1^- \operatorname{cl}(3,0) = \{(1,0)\}$, $\Delta_1^+ \operatorname{cl}(3,0) = \{(1,2)\}$ and $\mathscr{M}ax\, U = \{(1,1),(3,0)\}$. Geometrically, $U$ is a wedge of a 3-dimensional ball and an interval. Now, $\Delta_2^- U = \{(2,0)\}$ and $\Delta_2^+ U = \{(2,1)\}$. However, $\partial_2^- U = \{(2,0),(1,0),(1,2),(0,0),(0,1)\} \cup \{(1,1),(0,2),(0,1)\}$ and $\partial_2^+ U = \{(2,1),(1,0),(1,2),(0,0),(0,1)\} \cup \{(1,1),(0,2),(0,1)\}$. Graphically,




From the definitions previously presented, the following results hold:

**Lemma 2.1.25.** *Let $U$ be a closed subset of an oriented graded poset, $n \in \mathbb{N}$, and $\alpha \in \{+,-\}$. Then $\dim \partial_n^\alpha U \leq n$.*

*Proof.* Let $x \in \partial_n^\alpha U$. By definition there exists $y$ such that $x \leq y$ and either $y \in \Delta_n^\alpha U$, so $\dim y = n$, or $y \in (\mathscr{M}ax\, U)_k$, and $\dim y = k < n$. In either case, by the monotonicity of dimension ([25, Lemma 1.2.33]), $\dim x \leq \dim y \leq n$. ∎

**Lemma 2.1.26.** *Let $U$ be a closed subset of an oriented graded poset, $n \in \mathbb{N}$, and $\alpha \in \{+,-\}$. Then*

*I* $(\partial_n^\alpha U)_n = \Delta_n^\alpha U$,

*II* $(\mathscr{M}ax\,(\partial_n^\alpha U))_k = (\mathscr{M}ax\, U)_k$ *for all* $k < n$.

*Proof.* Let $x \in \partial_n^\alpha U$. Then by definition there exists $y$ such that $x \leq y$ and either $y \in \Delta_n^\alpha U$ or $y \in (\mathscr{M}ax\, U)_k$ for some $k < n$. If $x$ is maximal, necessarily $x = y$, and we obtain one inclusion. The converse inclusions are evident. ∎

**Lemma 2.1.27.** *Let $U$ be a closed subset of an oriented graded poset, $n \in \mathbb{N}$, and $\alpha \in \{+,-\}$. Then*

*I* $(\mathscr{M}ax\, U)_n = \Delta^+_n U \cap \Delta^-_n U$,

*II if* $n = \dim U$, *then* $(\mathscr{M}ax\, U)_n = \Delta^\alpha_n U = U_n$.

*Proof.* Let $x \in U$, $\dim x = n$. Then $x$ is maximal if and only if it has no cofaces in $U$, if and only if $\nabla^{-\alpha} x \cap U = \nabla^\alpha x \cap U = \varnothing$, if and only if $x \in \Delta^+_n U \cap \Delta^-_n U$. If $n = \dim U$, then every element of $U_n$ is maximal in $U$, so

$$U_n = (\mathscr{M}ax\, U)_n \subseteq \Delta^\alpha_n U \subseteq U_n$$

using the first part of the proof, and we conclude that they are all equal. ∎

Now we can state the result that one might have guessed from our previous examples: if $U$ is a closed subset of an oriented graded poset and $n < \dim U$, then the $n$-dimensional boundaries are subsets of $U$.

**Lemma 2.1.28.** *Let $U$ be a closed subset of an oriented graded poset, $n \in \mathbb{N}$, and $\alpha \in \{+,-\}$. Then*

*I* $\partial^\alpha_n U \subseteq U$,

*II* $\partial^\alpha_n U = U$ *if and only if* $n \geq \dim U$.

*Proof.* See [25, Lemma 2.1.20]. ∎

Now that we presented the data used to describe the shapes of diagrams, we define morphisms of oriented graded posets and describe the category **ogPos**. We will use the properties of this category in the next section where we'll talk about *gluing* and define *molecules*.

***Definition* 2.1.29** (Morphism of oriented graded posets) — Let $P, Q$ be oriented graded posets. A *morphism* $f\colon P \to Q$ is a function of their underlying sets which, for all $x \in P$ and $\alpha \in \{+,-\}$, induces a bijection between $\Delta^\alpha x$ and $\Delta^\alpha f(x)$.

Note that the definition above is different than the one used in [28, 29] as the condition of preserving the faces is more general while still capturing the behaviour that we want in order to define molecules.

***Definition* 2.1.30** (The category **ogPos**) — We let **ogPos** denote the category whose objects are oriented graded posets and morphisms are morphisms of oriented graded posets.

We state some properties of oriented graded posets without proving them. One can consult [25, Section 2.2] for the full proofs.

**Lemma 2.1.31.** *Let $f\colon P \to Q$ be a morphism of oriented graded posets. Then*

*I $f$ is order-preserving,*

*II $f$ is closed,*

*III* $f$ *is dimension-preserving, that is, for all* $x \in P$, $\dim f(x) = \dim x$.

*Proof.* See [25, Lemma 2.2.3]. ∎

***Definition* 2.1.32** (Inclusion of oriented graded posets) — An *inclusion* is an injective morphism of oriented graded posets.

**Lemma 2.1.33.** *Let* $f : P \to Q$ *be a morphism of oriented graded posets. The following are equivalent:*

*(a)* $f$ *is a surjective inclusion;*

*(b)* $f$ *is an isomorphism of oriented graded posets.*

*Proof.* See [25, Lemma 2.2.10]. ∎

**Lemma 2.1.34.** *Let* $\imath : P \hookrightarrow Q$ *be an inclusion of oriented graded posets and* $U \subseteq P$ *a closed subset. For all* $n \in \mathbb{N}$ *and* $\alpha \in \{+, -\}$,

$$\imath(\partial_n^\alpha U) = \partial_n^\alpha \imath(U).$$

*Proof.* See [25, Corollary 2.2.13]. ∎

**Proposition 2.1.35.** *The category* **ogPos** *has*

*I a strict initial object* $\varnothing$,

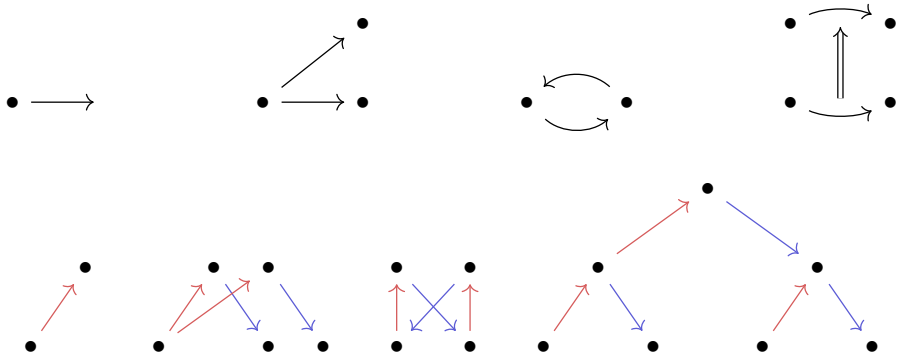
*II pushouts of inclusions along inclusions.*

*These colimits are computed as in the category of posets and order preserving maps.*

*Proof.* See [25, Proposition 2.2.21]. ∎

## 2.2 Molecules

Until now we have seen some examples of shapes of diagrams together with their oriented graded posets. But not all oriented graded posets describe well formed shapes of diagrams. Below we provide some examples of diagrams that are not pasting diagrams together with their Hasse diagram representations:



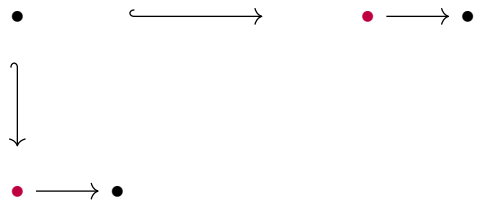
This section introduces molecules, our notion of pasting diagrams and builds up on their theory and properties. We build molecules via gluing which we describe in the following subsection.

### 2.2.1 Gluing

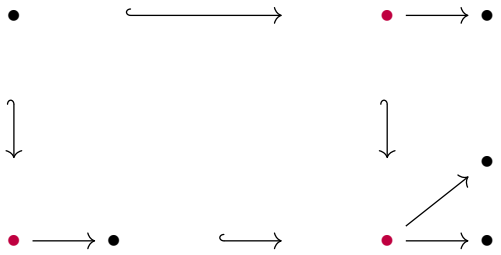
We take the approach from algebraic topology and construct what we will call molecules by gluing. This is a technique that works by identifying common boundaries of spaces and gluing them together along these boundaries to form more complicated objects. A good exposition of this technique can be found in [30].

We glue two oriented graded posets by taking the pushout of inclusions along the identified elements. Before giving the formal definition, let us provide some intuition about how the gluing operation works. To further highlight the definition of molecules, we include examples of non-pasting diagrams in this subsection.

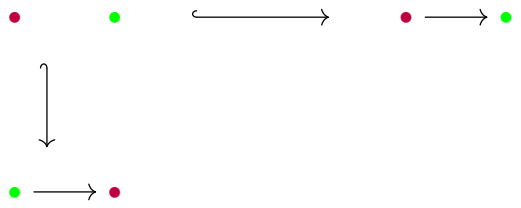
Let $U$ and $V$ be intervals (1-dimensional cells) and let's say we want to glue them along a 0-dimensional element. Then, we need to find an inclusion of the said element in both diagrams. Suppose we find the following inclusion (identified with purple):



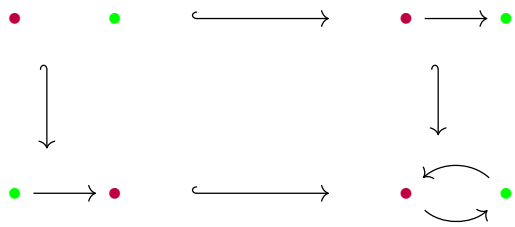
The pushout is the union of $U$ and $V$ where we identify the common elements of the inclusions (this is equivalent to quotienting the union by the common elements in the image):



Let's look at another example. Again, suppose that $U$ and $V$ are 1-dimensional cells and we have the following inclusion of 0-dimensional cells:



By taking the pushout of the diagram above, we obtain:



### 2.2.2 The inductive definition of molecules

We want to avoid bad situations like those exemplified at the beginning of the section. Our setting still allows for cycles to appear in higher-dimensional diagrams; this is not an issue, but on the contrary a desired result. While we exclude the bad examples mentioned above, our framework is lax enough to allow shapes of diagrams representing "phenomena" that are commonly occurring in category theory like the one in Figure 1.3.
In our work the concept of a pasting diagram is captured by molecules. The term is taken from [51]. Intuitively and similarly to pasting schemes, molecules are configurations of composable cells. We define the class of molecules as an inductive subclass of oriented graded posets using the *paste* and *rewrite* constructions defined below:

***Definition*** **2.2.1** (Pasting construction) — Let $U$, $V$ be oriented graded posets, $k \in \mathbb{N}$, and let $\varphi \colon \partial_k^+ U \xhookrightarrow{\sim} \partial_k^- V$ be an isomorphism. The *pasting of $U$ and $V$ at the $k$-boundary along $\varphi$* is the oriented graded poset $U \#_k^\varphi V$ obtained as the pushout

$$\begin{array}{ccccc} \partial_k^+ U & \xhookrightarrow{\varphi} & \partial_k^- V & \hookrightarrow & V \\ \downarrow & & & & \downarrow{\scriptstyle \imath_V} \\ U & & \xhookrightarrow{\imath_U} & & U \#_k^\varphi V \end{array}$$

in **ogPos**.

***Definition*** **2.2.2** (Rewrite construction) — Let $U$, $V$ be oriented graded posets of the same finite dimension $n$, and suppose $\varphi \colon \partial U \xhookrightarrow{\sim} \partial V$ is an isomorphism restricting to isomorphisms $\varphi^\alpha \colon \partial^\alpha U \xhookrightarrow{\sim} \partial^\alpha V$ for each $\alpha \in \{+, -\}$. Construct the pushout

$$\begin{array}{ccccc} \partial U & \xhookrightarrow{\varphi} & \partial V & \hookrightarrow & V \\ \downarrow & & & & \downarrow \\ U & & \hookrightarrow & & \partial(U \Rightarrow^\varphi V) \end{array}$$

in **ogPos**. The *rewrite of $U$ into $V$ along $\varphi$* is the oriented graded poset $U \Rightarrow^\varphi V$ obtained by adjoining a single $(n+1)$-dimensional element $\top$ to $\partial(U \Rightarrow^\varphi V)$, with

$$\Delta^- \top \coloneqq U_n, \qquad \Delta^+ \top \coloneqq V_n.$$

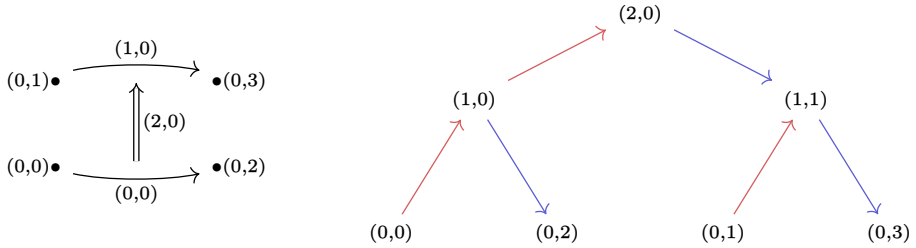
Pasting diagrams also possess the property of *globularity*, which ensures that repeated operation of taking lower-dimensional input or output boundaries can be reduced to a single instance.

***Definition* 2.2.3** (Globularity) — Let $U$ be an oriented graded poset. We say that $U$ is *globular* if, for all $k, n \in \mathbb{N}$ and $\alpha, \beta \in \{+, -\}$, if $k < n$ then

$$\partial_k^\alpha(\partial_n^\beta U) = \partial_k^\alpha U.$$

The last example of a non-pasting diagram from the beginning of this section is not globular. But let us see why:

***Example* 2.2.4** — Let $U$ be the following oriented graded poset:




Note that $\partial_0^- U = \{(0,0),(0,1)\}$, while $\partial_0^-(\partial_1^- U) = \{(0,0)\}$ and $\partial_0^-(\partial_1^+ U) = \{(0,1)\}$.

*Roundness*, a stronger condition than globularity, is the last ingredient that we need to define before introducing molecules. Intuitively, the property of roundness ensures that the shape of an $n$-dimensional diagram is a closed topological $n$-ball. As we will see later on, being round also has the consequence that the shape of diagrams is "connected" at its $(n-1)$-boundaries.

***Definition* 2.2.5** (Roundness) — Let $U$ be an oriented graded poset. We say that $U$ is *round* if it is globular and, for all $n < \dim U$,

$$\partial_n^- U \cap \partial_n^+ U = \partial_{n-1} U.$$

***Example* 2.2.6** — The molecule in example 2.1.4 seen as an oriented graded poset is not a round molecule as $\partial_1^- U \cap \partial_1^+ U = \{(0,0),(0,2),(0,3)\}$ while $\partial_0 U = \{(0,0),(0,3)\}$.

***Definition* 2.2.7** (Point) — The *point* is the oriented graded poset 1 with a single element and trivial orientation.

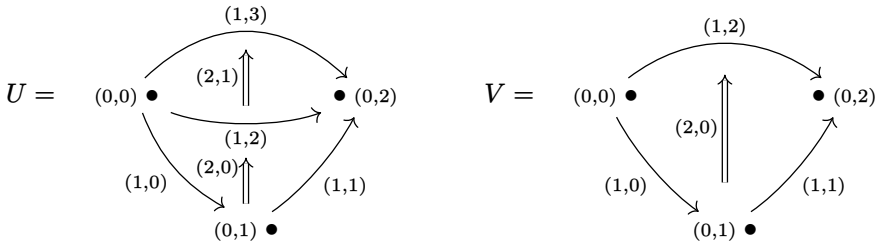
***Definition* 2.2.8** (Molecule) — The class of *molecules* is the inductive subclass of oriented graded posets closed under isomorphisms and generated by the following clauses.

I (*Point*). The point is a molecule.

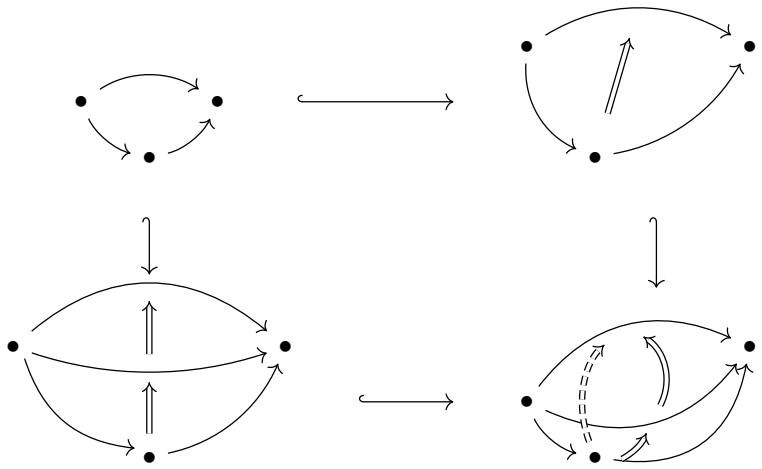
II (*Paste*). Let $U$, $V$ be molecules, let $k < \min\{\dim U, \dim V\}$, and consider $\varphi\colon \partial_k^+ U \xhookrightarrow{\sim} \partial_k^- V$ an isomorphism. Then $U \#_k^\varphi V$ is a molecule.

III (*Atom*). Let $U$, $V$ be *round* molecules of the same finite dimension and let $\varphi\colon \partial U \xrightarrow{\sim} \partial V$ be an isomorphism restricting to $\varphi^{\alpha}\colon \partial^{\alpha} U \xrightarrow{\sim} \partial^{\alpha} V$ for each $\alpha \in \{+,-\}$. Then $U \Rightarrow^{\varphi} V$ is a molecule.

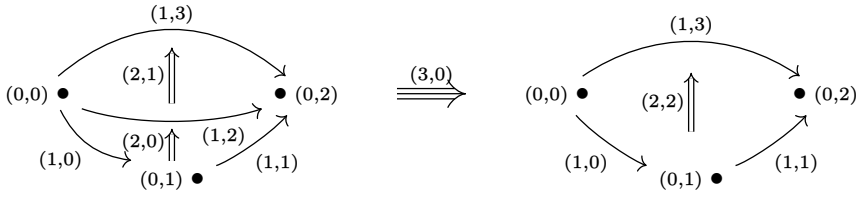
Let's see how these molecules are constructed in practice. We start with the *Atom* rule and build a 3-dimensional shape out of two 2-dimensional shapes of diagrams. We chose to exemplify in dimension three, because the 0-dimensional elements (the points) are trivially isomorphic. These points will be the input or output of the 1-dimensional shapes that will form the input or output boundary of our 2-dimensional shape, so building a 2-dimensional atom is always possible as long as we have 1-dimensional well-formed pasting diagrams.
Consider the following two 2-dimensional shapes of diagrams:



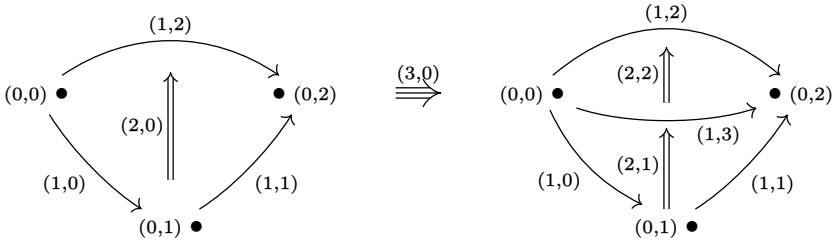

We identify that the input and output boundaries of $U$ and $V$ are isomorphic and glue the two shapes along this isomorphism. The gluing is done by taking the pushout in the category **ogPos**. In our example, this corresponds to the following picture:



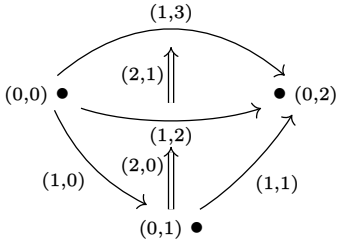
What we obtain by taking the pushout is a sphere to which we want to add the volume - the 3-dimensional cell. One way is to imagine some kind of a shell or deflated balloon in which $U$ is at the front and $V$ at the bottom. To finish building the atom, we now add a new 3-dimensional cell, $(3,0)$ going from $U$ to $V$. The new molecule $U \Rightarrow V$ will look as follows:

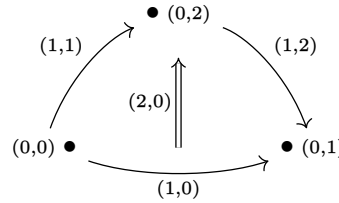



Note that it is perfectly valid to do $V \Rightarrow U$ in which case our shape looks like:

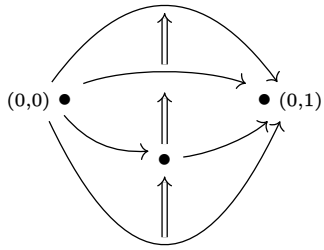


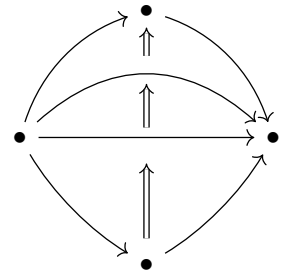

The Paste construction works similarly to the atom construction. Consider the following two shapes of diagrams:

$U =$ 

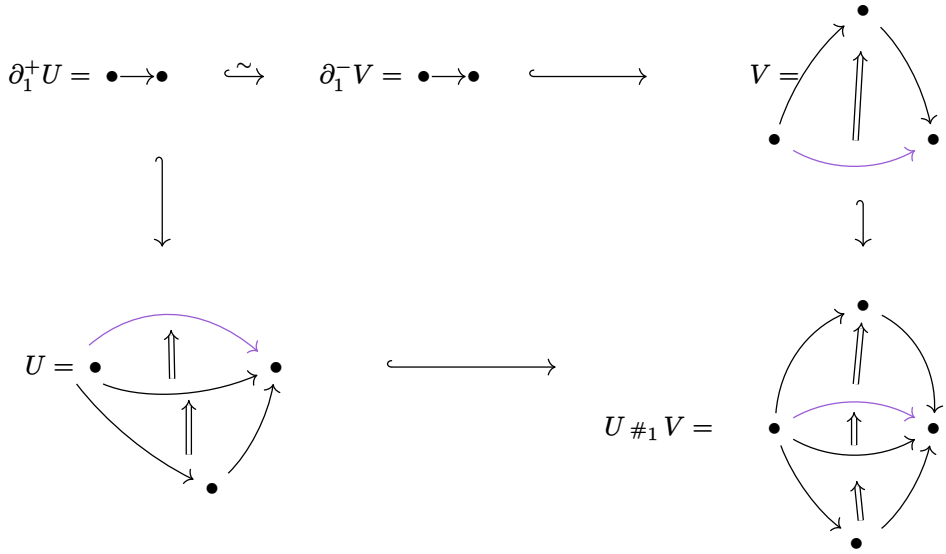

and $V =$ 


First note that $\dim U = \dim V = 2$, so $k \in \{0, 1\}$. We can paste at dimensions 0 and 1 as long as the boundaries are isomorphic. Let's choose $k = 1$. Note that both $\partial_1^+ U \cong \partial_1^- V$ and $\partial_1^+ V \cong \partial_1^- U$ and we get the following shapes:

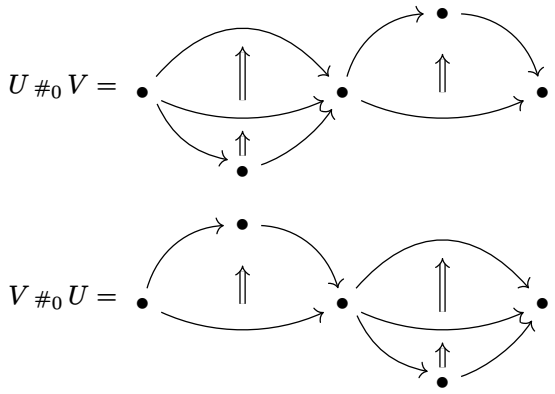
$V \#_1 U =$ 


and $U \#_1 V =$ 

The pasting operation is done via taking the pushout along the isomorphic boundaries:




For $k = 0$, we always have an isomorphism, hence the shapes $U \#_0 V$ and $V \#_0 U$ are:




### 2.2.3 Basic properties of molecules

We prove or state some basic properties of molecules. Namely, that all molecules are globular, that the boundaries of a molecule are themselves molecules and that for any element in a molecule, its closure is a molecule. Some of the properties presented below also hold for oriented graded posets. But since the thesis is dedicated to molecules, we choose to present them at this stage.
We start by showing some properties about the operations of pasting and atom on molecules. We introduce the concept of an atom - a molecule with a greatest element. In Lemma 2.2.10, we show that the pasting of two molecules is always

a molecule, whenever an isomorphism between their respective boundaries exists while in Lemma 2.2.14 we give a concise description of what it means for a molecule to be an *atom*.

**Lemma 2.2.9.** *Let $U$ be a molecule. Then $|U|$ is finite.*

*Proof.* By induction on the construction of $U$. If $U$ was produced by (*Point*), then $|U| = 1$. If $U$ was produced by (*Paste*), then it is equal to $V \#_k^\varphi W$ for some molecules $V$, $W$ and $k < \min\{\dim V, \dim W\}$. Then $|U| \leq |V| + |W|$, which is finite by the inductive hypothesis. If $U$ was produced by (*Atom*), then it is of the form $V \Rightarrow^\varphi W$ for some molecules $V$, $W$. Then $|U| \leq |V| + |W| + 1$, which is finite by the inductive hypothesis. ∎

**Lemma 2.2.10.** *Let $U$, $V$ be molecules, $k \in \mathbb{N}$, and $\varphi \colon \partial_k^+ U \xhookrightarrow{\sim} \partial_k^- V$ an isomorphism. Then $U \#_k^\varphi V$ is a molecule.*

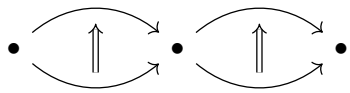
*Proof.* If $k < \min\{\dim U, \dim V\}$, then this is an application of the (*Paste*) constructor. If $k \geq \dim U$, then $\partial_k^+ U = U$ and $U \#_k V$ is isomorphic to $V$, which is a molecule by assumption. Similarly, if $k \geq \dim V$, then $U \#_k V$ is isomorphic to $U$. ∎

**Lemma 2.2.11.** *Let $U$ be a round oriented graded poset. If $U$ is finite-dimensional, then $U$ is pure.*

*Proof.* Suppose that $U$ is not pure. Then there exists a maximal element $x$ in $U$ with $k \coloneqq \dim x < \dim U$. By Lemma 2.1.27, $x \in \partial_k^- U \cap \partial_k^+ U$. Then $\partial_k^- U \cap \partial_k^+ U$ is $k$-dimensional and cannot be equal to $\partial_{k-1} U$, which is $(k-1)$-dimensional. It follows that $U$ is not round. ∎

The converse of the lemma above is not true:

***Example* 2.2.12** (Molecule that is pure but not round) — Before presenting the example, let's argue why this would not be the case. We saw from the definition that one can glue two atoms along their $k$-dimensional boundary with $k < n - 1$ which will result in the two maximal elements not sharing $(n-1)$-dimensional elements.



***Definition* 2.2.13** (Atom) — An *atom* is a molecule with one maximal element.

**Lemma 2.2.14.** *Let $U$ be a molecule. The following are equivalent:*

*(a) $U$ is an atom;*

*(b) the final constructor producing $U$ is (Point) or (Atom).*

*Proof.* If $U$ was produced by (*Point*), then $U$ is the point, which trivially has a greatest element.
If $U$ was produced by (*Paste*), then $U$ splits into a union $V \cup W$, where $V \cap W = \partial_k^+ V = \partial_k^- W$ and $k < \max\{\dim V, \dim W\}$. Then there exist elements $x_1 \in V$ and $x_2 \in W$ such that

I $x_1$ is maximal in $V$ and $x_2$ is maximal in $W$,

II $\dim x_1 > k$ and $\dim x_2 > k$.

By Lemma 2.1.25, $\dim(V \cap W) \leq k$, so neither $x_1$ nor $x_2$ are contained in $V \cap W$. It follows that $x_1$ and $x_2$ are distinct maximal elements of $U$, so $U$ does not have a greatest element.
If $U$ was produced by $(Atom)$, then $U$ splits into $(U_- \cup U_+) + \{\top\}$, where $U_-$ and $U_+$ are round molecules of dimension $n$, and $\Delta^\alpha \top = (U_\alpha)_n$ for each $\alpha \in \{+, -\}$. By Lemma 2.2.11, we have $U_\alpha = \mathrm{cl}\,(U_\alpha)_n$, so $U_\alpha = \partial^\alpha \top \subseteq \mathrm{cl}\,\{\top\}$. It follows that all elements of $U$ are in the closure of $x$, that is, $x$ is the greatest element of $U$. ∎

We have the following results about the boundaries of molecules, which we prove more generally for oriented graded posets.

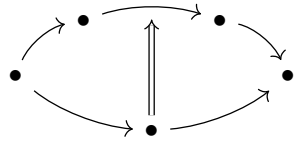
**Lemma 2.2.15.** *Let $U$, $V$ be oriented graded posets and suppose $U \Rightarrow^\varphi V$ is defined. Then*

*I* $\partial^-(U \Rightarrow^\varphi V) = U$,

*II* $\partial^+(U \Rightarrow^\varphi V) = V$.

*Proof.* See [25, Lemma 3.2.3]. ∎

***Example* 2.2.16 —** The lemma above follows from the definition.

If we let $U \Rightarrow V =$  , then

$$\partial^-(U \Rightarrow V) = \bullet \longrightarrow \bullet \longrightarrow \bullet = U$$

and

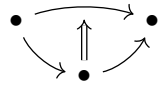
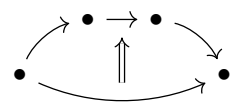
$$\partial^+(U \Rightarrow V) = \bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet = V.$$

**Lemma 2.2.17.** *Let $U$, $V$ be oriented graded posets, $k \in \mathbb{N}$, and suppose $U \#_k^\varphi V$ is defined. Then*

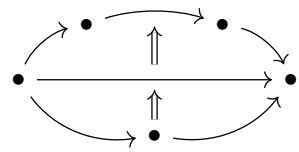
*I* $\partial_k^-(U \#_k^\varphi V) = \partial_k^- U$,

*II* $\partial_k^+(U \#_k^\varphi V) = \partial_k^+ V$.

*Proof.* See [25, Lemma 3.1.5]. ∎

***Example* 2.2.18 —** Let $U =$  and $V =$  .

Then

$U \#_1 V =$ 

and

$$\partial_1^-(U \#_1 V) = \quad \bullet \longrightarrow \bullet \longrightarrow \bullet \quad = \partial_1^- U$$

and

$$\partial_1^+(U \#_1 V) = \quad \bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet \quad = \partial_1^+ V.$$

**Lemma 2.2.19.** *Let $U$, $V$ be globular oriented graded posets, $k \in \mathbb{N}$, and suppose $U \#_k^\varphi V$ is defined. For all $j < k$ and $\alpha \in \{+,-\}$,*

$$\partial_j^\alpha U = \partial_j^\alpha V = \partial_j^\alpha (U \#_k^\varphi V).$$

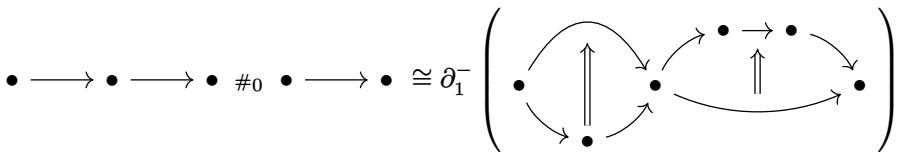
*Proof.* See [25, Lemma 3.1.14]. ∎

Lemma 2.2.19 holds for any molecule that is formed as a pasting of two molecules. Since $U$ and $V$ are pasted at dimension $k$, by the property of globularity, all the elements of dimension less than $k$ are identified with their isomorphic image. Symetrically to Lemma 2.2.19, we get:

**Lemma 2.2.20.** *Let $U$, $V$ be globular oriented graded posets, $k \in \mathbb{N}$, and suppose $U \#_k^\varphi V$ is defined. For all $n > k$ and $\alpha \in \{+,-\}$, the pasting $\partial_n^\alpha U \#_k^\varphi \partial_n^\alpha V$ is defined and maps isomorphically onto $\partial_n^\alpha (U \#_k^\varphi V)$.*

*Proof.* See [25, Lemma 3.1.15]. ∎

***Example* 2.2.21 —** Consider the molecules defined in Example 2.2.18; we then have $\partial_1^- U \#_0^\varphi \partial_1^- V \cong \partial_1^-(U \#_0^\varphi V)$. In pasting diagrams this looks like:



We show that globularity is preserved under all the operations used to construct molecules. In particular, we show that every molecule is globular; we moreover show that molecules are preserved under taking boundaries. For this, we need a few additional lemmas that hold for all oriented graded posets.

**Lemma 2.2.22.** *Let $U$ be a globular oriented graded poset, $n \in \mathbb{N}$, and $\beta \in \{+,-\}$. Then $\partial_n^\alpha U$ is globular.*

*Proof.* Let $k < m$ be natural numbers and $\alpha, \gamma \in \{+,-\}$. If $m < n$, using globularity of $U$ twice,

$$\partial_k^\alpha(\partial_m^\gamma(\partial_n^\beta U)) = \partial_k^\alpha(\partial_m^\gamma U) = \partial_k^\alpha U = \partial_k^\alpha(\partial_n^\beta U).$$

If $m \geq n$, by Lemma 2.1.28 we have $\partial_m^\gamma(\partial_n^\beta U) = \partial_n^\beta U$, so

$$\partial_k^\alpha(\partial_m^\gamma(\partial_n^\beta U)) = \partial_k^\alpha(\partial_n^\beta U).$$ ∎

**Lemma 2.2.23.** *Let $U$, $V$ be globular oriented graded posets, $k \in \mathbb{N}$, and suppose $U \#_k^\varphi V$ is defined. Then $U \#_k^\varphi V$ is globular.*

*Proof.* Let $m, n \in \mathbb{N}$ such that $m < n$, and $\alpha, \beta \in \{+, -\}$. If $n < k$, by Lemma 2.2.19

$$\partial_m^\alpha(\partial_n^\beta(U \#_k^\varphi V)) = \partial_m^\alpha(\partial_n^\beta U) = \partial_m^\alpha(U) = \partial_m^\alpha(U \#_k^\varphi V).$$

If $n = k$, by Lemma 2.2.17 and Lemma 2.2.19,

$$\partial_m^\alpha(\partial_n^-(U \#_k^\varphi V)) = \partial_m^\alpha(\partial_n^- U) = \partial_m^\alpha(U) = \partial_m^\alpha(U \#_k^\varphi V)$$

and

$$\partial_m^\alpha(\partial_n^+(U \#_k^\varphi V)) = \partial_m^\alpha(\partial_n^+ V) = \partial_m^\alpha(V) = \partial_m^\alpha(U \#_k^\varphi V).$$

Finally, if $n > k$, by Lemma 2.2.20 we have

$$\partial_m^\alpha(\partial_n^\beta(U \#_k^\varphi V)) = \partial_m^\alpha(\partial_n^\beta U \#_k^\varphi \partial_n^\beta V),$$

and by Lemma 2.2.22 $\partial_n^\beta U$ and $\partial_n^\beta V$ are globular. If $m < k$ we use Lemma 2.2.19 to obtain

$$\partial_m^\alpha(\partial_n^\beta U \#_k^\varphi \partial_n^\beta V) = \partial_m^\alpha(\partial_n^\beta U) = \partial_m^\alpha U = \partial_m^\alpha(U \#_k^\varphi V).$$

If $m = k$ we use Lemma 2.2.17 instead to obtain

$$\partial_m^-(\partial_n^\beta U \#_k^\varphi \partial_n^\beta V) = \partial_m^-(\partial_n^\beta U) = \partial_m^- U = \partial_m^-(U \#_k^\varphi V)$$

and similarly

$$\partial_m^+(\partial_n^\beta U \#_k^\varphi \partial_n^\beta V) = \partial_m^+(\partial_n^\beta V) = \partial_m^+ V = \partial_m^+(U \#_k^\varphi V).$$

Finally, if $m > k$ we use Lemma 2.2.20 once more to obtain

$$\partial_m^\alpha(\partial_n^\beta U \#_k^\varphi \partial_n^\beta V) = \partial_m^\alpha(\partial_n^\beta U) \#_k^\varphi \partial_m^\alpha(\partial_n^\beta V) = \partial_m^\alpha U \#_k^\varphi \partial_m^\alpha V$$

and once more to obtain

$$\partial_m^\alpha U \#_k^\varphi \partial_m^\alpha V = \partial_m^\alpha(U \#_k^\varphi V). \qquad \blacksquare$$

**Lemma 2.2.24.** *Let $U$, $V$ be oriented graded posets and suppose $U \Rightarrow^\varphi V$ is defined. Then $U \Rightarrow^\varphi V$ is globular.*

*Proof.* For all $k < \dim U = \dim V$ and $\alpha \in \{+, -\}$, we have

$$\partial_k^\alpha U = \partial_k^\alpha(\partial^\beta U) = \partial_k^\alpha(\partial^\beta V) = \partial_k^\alpha V$$

since $\partial^\beta U = \partial^\beta V$ and $U$, $V$ are globular. It then suffices to show that, for all $k < \dim U$ and $\alpha \in \{+, -\}$,

$$\partial_k^\alpha(U \Rightarrow^\varphi V) = \partial_k^\alpha U.$$

Indeed, suppose this holds, and let $k < n < \dim(U \Rightarrow^\varphi V)$ and $\alpha, \beta \in \{+, -\}$. If $n = \dim U$, then by Lemma 2.2.15

$$\partial_k^\alpha(\partial_n^-(U \Rightarrow^\varphi V)) = \partial_k^\alpha U = \partial_k^\alpha(U \Rightarrow^\varphi V)$$

and similarly

$$\partial_k^\alpha(\partial_n^+(U \Rightarrow^\varphi V)) = \partial_k^\alpha V = \partial_k^\alpha U = \partial_k^\alpha(U \Rightarrow^\varphi V).$$

If $n < \dim U$, then

$$\partial_k^\alpha(\partial_n^\beta(U \Rightarrow^\varphi V)) = \partial_k^\alpha(\partial_n^\beta U) = \partial_k^\alpha U = \partial_k^\alpha(U \Rightarrow^\varphi V)$$

using the globularity of $U$.
Let then $k < \dim U$ and $\alpha \in \{+,-\}$. We have $\Delta_k^\alpha(U \Rightarrow^\varphi V) = \Delta_k^\alpha(U \cup V)$. Since $\Delta_k^\alpha U = \Delta_k^\alpha V$, by [25, Lemma 2.1.22] we have $\Delta_k^\alpha(U \cup V) = \Delta_k^\alpha U$. Similarly, we prove that for all $j < k$ we have $(\mathscr{M}ax\,(U \cup V))_j = (\mathscr{M}ax\,U)_j$. It follows that $\partial_k^\alpha(U \Rightarrow^\varphi V) = \partial_k^\alpha U$. ∎

**Proposition 2.2.25.** *Let $U$ be a molecule, $n \in \mathbb{N}$, $\alpha \in \{+,-\}$. Then*

*I $U$ is globular,*

*II $\partial_n^\alpha U$ is a molecule,*

*III if $n \leq \dim U$, then $\dim \partial_n^\alpha U = n$.*

*Proof.* By induction on the construction of $U$. Suppose $U$ was produced by (*Point*). Then $U$ is the point, it has no non-trivial boundaries, and is trivially globular.
Suppose $U$ was produced by (*Paste*). Then $U = V \#_k^\varphi W$ for some molecules $V$, $W$. By the inductive hypothesis, $V$ and $W$ are globular, and by Lemma 2.2.23 so is $U$. We have $k < \min\{\dim V, \dim W\}$. If $n = k$, then by Lemma 2.2.17 $\partial_n^- U$ is equal to $\partial_n^- V$ and $\partial_n^+ U$ to $\partial_n^+ W$. By the inductive hypothesis, both of these are $n$-dimensional molecules. If $n < k$, then by Lemma 2.2.19 $\partial_n^\alpha U$ is equal to $\partial_n^\alpha V$, and again the inductive hypothesis applies. If $n > k$, then by Lemma 2.2.20 $\partial_n^\alpha U$ is equal to $\partial_n^\alpha V \#_k^\varphi \partial_n^\alpha W$. By the inductive hypothesis, $\partial_n^\alpha V$ and $\partial_n^\alpha W$ are molecules, and if $n < \dim U = \max\{\dim V, \dim W\}$, at least one of them is $n$-dimensional.
Finally, suppose $U$ was produced by (*Atom*). Then $U = V \Rightarrow^\varphi W$ for some round molecules $V$, $W$ of the same dimension. By the inductive hypothesis, $V$ and $W$ are globular, and by Lemma 2.2.24 so is $U$. If $n \geq \dim U$, then $\partial_n^\alpha U = U$ is by assumption a molecule. If $n = \dim U - 1$, then by Lemma 2.2.15 $\partial^- U$ is equal to $V$ and $\partial^+ U$ to $W$, both molecules of dimension $n$. If $n < \dim U - 1$, then $\partial_n^\alpha U = \partial_n^\alpha V = \partial_n^\alpha W$ by globularity, and the inductive hypothesis applies. ∎

Roundness, however is preserved only under some of these operations:

**Lemma 2.2.26.** *Let $U$ be round, $n \in \mathbb{N}$, and $\alpha \in \{+,-\}$. Then $\partial_n^\alpha U$ is round.*

*Proof.* If $n \geq \dim U$ there is nothing to prove, so suppose $n < \dim U$. By Lemma 2.2.22, $\partial_n^\alpha U$ is globular. Let $k < \dim(\partial_n^\alpha U) \leq n$. Then

$$\partial_k^-(\partial_n^\alpha U) \cap \partial_k^+(\partial_n^\alpha U) = \partial_k^- U \cap \partial_k^+ U = \partial_{k-1} U = \partial_{k-1}(\partial_n^\alpha U)$$

using roundness of $U$. ∎

**Lemma 2.2.27.** *Let $U$, $V$ be round and suppose $U \Rightarrow^\varphi V$ is defined. Then $U \Rightarrow^\varphi V$ is round.*

*Proof.* Globularity follows from Lemma 2.2.24, so we only need to prove roundness. Let $n \coloneqq \dim U = \dim V$. By Lemma 2.2.15

$$\partial^-(U \Rightarrow^\varphi V) \cap \partial^+(U \Rightarrow^\varphi V) = U \cap V = \partial U = \partial V,$$

and by globularity $\partial U = \partial(\partial^-(U \Rightarrow^\varphi V)) = \partial_{n-1}(U \Rightarrow^\varphi V)$. Finally, for $k < n$

$$\partial_k^-(U \Rightarrow^\varphi V) \cap \partial_k^+(U \Rightarrow^\varphi V) = \partial_k^- U \cap \partial_k^+ U = \partial_{k-1} U = \partial_{k-1}(U \Rightarrow^\varphi V)$$

by globularity of $U \Rightarrow^\varphi V$ and roundness of $U$. ∎

**Corollary 2.2.28.** *All atoms are round.*

*Proof.* Let $U$ be an atom. If it was produced by (*Point*), it is trivially round. If it was produced by (*Atom*), it is round by Lemma 2.2.27. ∎

**Remark 2.2.29.** *Note that roundness is not preserved under pasting. Example 2.2.12 motivates why: the molecule can be written as a 0-pasting of two atoms.*

**Lemma 2.2.30.** *Let $U$ be a molecule, $x \in U$. Then $\operatorname{cl}\{x\}$ is an atom.*

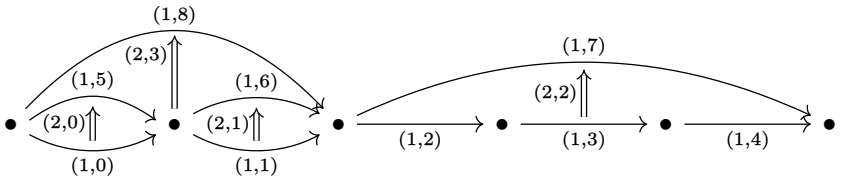
*Proof.* By induction on the construction of $U$. If $U$ was produced by (*Point*), then $x$ must be the unique element of $U$ whose closure is $U$ itself. If $U$ was produced by (*Paste*), it splits into $V \cup W$, and $x \in V$ or $x \in W$; the inductive hypothesis applies. If $U$ was produced by (*Atom*), it is equal to $(V \cup W) + \{\top\}$, and either $x \in V$ or $x \in W$, in which case the inductive hypothesis applies, or $x = \top$, and $\operatorname{cl}\{x\} = U$ is an atom by definition. ∎

## 2.3 Graphs associated with diagrams

As one might have noticed, the shapes of diagrams are already quite complicated and beyond dimension 3, graphical representation of molecules becomes hard to manipulate. Moreover, the Hasse diagram representation does not give us a good idea or too much intuition on how our shapes look like: they get complicated as the shapes become more and more complex and deriving any intuition from them becomes cumbersome. Many times, we want to get an idea on how the top few dimensions - i.e., from $n$ to $(n-2)$ - look like and derive results for them or we want to know how the elements in the diagram relate to each other. For these purposes, we introduce two notions of graphs that we use to study the behaviour of molecules at the top two dimensions and to study how their elements are connected. Apart from their usefulness in helping us better visualise the diagrams we are working with, the graphs presented will also prove to be great tools in proving properties about molecules. The directed graphs with open edges are a valuable tool for proving some important properties about the molecules, such as the the fact that there exists at most one isomorphism between two molecules. This allows us to treat two isomorphic molecules as being equal. Moreover, flow graphs are used at a later time in the thesis to study certain (a)cyclicity conditions of the molecules.

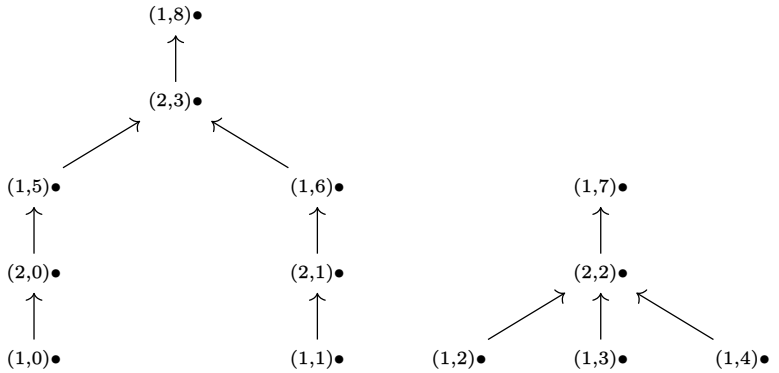
### 2.3.1 Directed graph with open edges

We want a way to reason about the shapes that are part of the diagram and the way the elements of a certain dimension connect. Take for example the following 2-dimensional diagram:




(2.1)

We are interested in the number of faces and cofaces of the elements in the top two dimensions. To represent the diagram above we use the following graph:




(2.2)

Note that the graphs above have two types of nodes: node vertices, which represent the top-dimensional elements and wire vertices representing the $(n-1)$-dimensional elements. A similar representation can be found in [17] where it is called an *open graph* or in [34] where it is called a *graph*. We call it *directed graph with open edges*:

***Definition* 2.3.1** (Directed graph with open edges) — A *directed graph with open edges* is a directed graph

$$\mathscr{G} \coloneqq E_{\mathscr{G}} \underset{t}{\overset{s}{\rightrightarrows}} N_{\mathscr{G}} + W_{\mathscr{G}}$$

with set of vertices bipartite into a set $N_{\mathscr{G}}$ of *node vertices* and a set $W_{\mathscr{G}}$ of *wire vertices*, satisfying the following properties:

I the bipartition $N_{\mathscr{G}} + W_{\mathscr{G}}$ exhibits $\mathscr{G}$ as a bipartite graph, that is, every edge connects a node vertex to a wire vertex or vice versa;

II each wire vertex is the source of at most one edge and the target of at most one edge.

***Definition* 2.3.2** (Boundary of a directed graph with open edges) — Let $\mathscr{G}$ be a directed graph with open edges. The *input boundary* of $\mathscr{G}$ is the set

$$\Delta^{-}\mathscr{G} \coloneqq \left\{x \in W_{\mathscr{G}} \mid t^{-1}(x) = \varnothing\right\}$$

and the *output boundary* of $\mathscr{G}$ is the set

$$\Delta^+\mathscr{G} \coloneqq \left\{x \in W_{\mathscr{G}} \mid s^{-1}(x) = \varnothing\right\}.$$

***Example* 2.3.3** — In the shape from 2.1 whose graph is 2.2, $\Delta^+\mathscr{G} = \{(1,7),(1,8)\}$ and $\Delta^-\mathscr{G} = \{(1,0),(1,1),(1,2),(1,3),(1,4)\}$.

When reasoning about molecules, we are usually interested about what happens at dimensions $n$ and $(n-1)$:

***Definition* 2.3.4** (Graph of a molecule) — Let $U$ be a molecule, $n \coloneqq \dim U$. The *graph of* $U$ is the directed graph

$$\mathscr{G}U \coloneqq \quad E_{\mathscr{G}U} \underset{t}{\overset{s}{\rightrightarrows}} N_{\mathscr{G}U} + W_{\mathscr{G}U},$$

where

- $E_{\mathscr{G}U} \coloneqq \{(x,y) \mid x \in U_n, y \in \Delta^+x\} + \{(x,y) \mid y \in U_n, x \in \Delta^-y\}$,
- $N_{\mathscr{G}U} \coloneqq U_n$,
- $W_{\mathscr{G}U} \coloneqq U_{n-1}$,
- $s\colon (x,y) \mapsto x$,
- $t\colon (x,y) \mapsto y$.

***Definition* 2.3.5** (Induced subgraph) — Let $\mathscr{G}$ be a directed graph and let $W \subseteq V_{\mathscr{G}}$. The *induced subgraph* of $\mathscr{G}$ on $W$ is the directed graph

$$\mathscr{G}|_W \coloneqq \quad E' \underset{t|_{E'}}{\overset{s|_{E'}}{\rightrightarrows}} W$$

where $E' \coloneqq \{e \in E_{\mathscr{G}} \mid s(e), t(e) \in W\}$.

**Remark 2.3.6.** *Note that, if we forget the separation of the vertex set into $N_{\mathscr{G}U}$ and $W_{\mathscr{G}U}$, then $\mathscr{G}U$ is the induced subgraph of $\vec{\mathscr{H}}U$ on vertices of dimension $n$ and $n-1$.*

**Proposition 2.3.7.** *Let $U$ be a molecule. Then*

*I $\mathscr{G}U$ is a directed graph with open edges,*

*II $\mathscr{G}U$ is acyclic,*

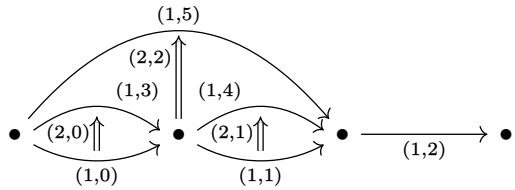
*III $\Delta^\alpha\mathscr{G}U = \Delta^\alpha U$ for all $\alpha \in \{+,-\}$.*

*Proof.* The fact that $\Delta^\alpha\mathscr{G}U = \Delta^\alpha U$ for all $\alpha \in \{+,-\}$ is immediate from the definitions. Moreover, $\mathscr{G}U$ is bipartite by construction, so it suffices to check the other conditions.
We proceed by induction on the construction of $U$. If $U$ was produced by (*Point*) or by (*Atom*), then by Lemma 2.2.14 it has a greatest element $\top$. In this case,

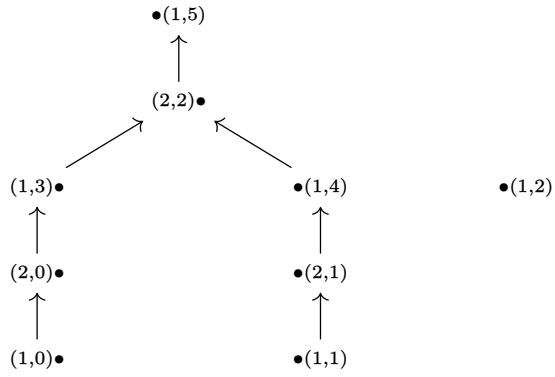
$\mathscr{G}U$ has a single edge $(x, \top)$ for each $x \in \Delta^- \top$ and a single edge $(\top, x)$ for each $x \in \Delta^+ \top$. Since $\Delta^- \top \cap \Delta^+ \top = \varnothing$, the graph is acyclic.
If $U$ was produced by (*Paste*), it is of the form $V \#_k^\varphi W$. Let $n \coloneqq \dim U$. If $k < n - 1$, then $\mathscr{G}U$ is the disjoint union of the induced subgraphs on the vertices in the image of $V$ and $W$, respectively. If $n = \dim V = \dim W$ we can conclude by the inductive hypothesis. Otherwise, the inductive hypothesis applies to one of the components, while the other is a discrete graph with no node vertices, trivially satisfying the conditions of an acyclic directed graph with open edges.
If $k = n - 1$, observe first that necessarily $\dim V = \dim W = n$. Then $\mathscr{G}U$ is the union of $\mathscr{G}V$ and $\mathscr{G}W$, and their intersection consists of the wire vertices in $\Delta^+_{n-1} V = \Delta^-_{n-1} W$. Let $x$ be a wire vertex. If $x \in V \setminus W$ or $x \in V \setminus W$, it is the source of at most one edge and the target of at most one edge by the inductive hypothesis applied to $\mathscr{G}V$ and $\mathscr{G}W$. If $x \in V \cap W$, then $x \in \Delta^+ \mathscr{G}V$, so it is the source of no edge of $\mathscr{G}V$ and at most one edge of $\mathscr{G}W$, and $x \in \Delta^- \mathscr{G}W$, so it is the target of no edge of $\mathscr{G}W$ and the source of at most one edge of $\mathscr{G}V$.
Finally, suppose there is a cycle in $\mathscr{G}U$. Because $\mathscr{G}V$ and $\mathscr{G}W$ are separately acyclic, such a cycle needs to cross from $V$ to $W \setminus V$ and back. However, a path entering $V$ from $W \setminus V$ must enter a wire vertex $y$ from a node vertex $x \in W$ such that $y \in \Delta^+ x$. But $(V \cap W)_{n-1} = \Delta^- W$, so this is impossible. We conclude that $\mathscr{G}U$ is acyclic. ■

Before continuing, let's look at another example.

***Example* 2.3.8 —** Take the following molecule:




Then, $\mathscr{G}U$ is:




Note that we have the following three situations for the wire vertices in Example 2.3.8:

- they have no incoming or outgoing edges,
- they have either one incoming edge or one outgoing edge,
- they have exactly one incoming edge and exactly one outgoing edge.

This observation is true for all the elements of codimension 1 in a molecule and is captured by the following corollary:

**Corollary 2.3.9.** *Let $U$ be a molecule, $x \in U$ with $\mathrm{codim}_U(x) = 1$, and $\alpha \in \{+,-\}$. Then*

*I $x \in \mathscr{M}ax\, U$ if and only if $|\nabla x| = 0$,*

*II $x \in \Delta^\alpha U \setminus \Delta^{-\alpha} U$ if and only if $|\nabla^\alpha x| = 1$ and $|\nabla^{-\alpha} x| = 0$,*

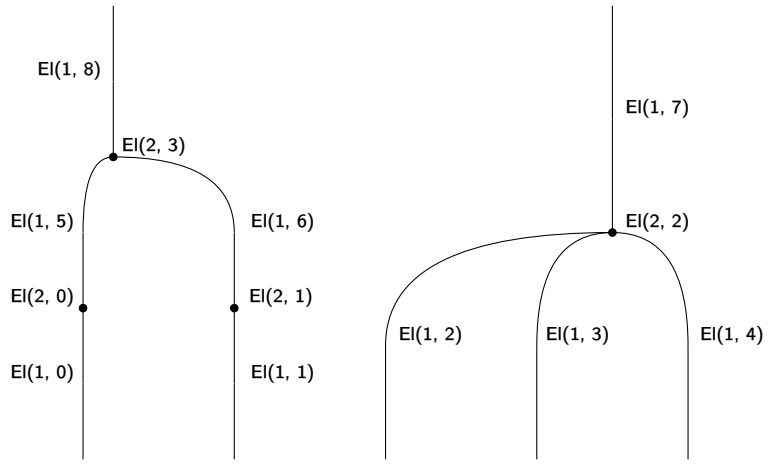
*III $x \notin \Delta U$ if and only if $|\nabla^+ x| = |\nabla^- x| = 1$.*

*Proof.* By Proposition 2.3.7, $\mathscr{G}U$ is a directed graph with open edges, and by construction we can identify $\nabla^- x$ with $s^{-1}(x)$ and $\nabla^+ x$ with $t^{-1}(x)$. It follows that $|\nabla^\alpha x| \leq 1$. The statement then follows from the isomorphism between $\Delta^\alpha U$ and $\Delta^\alpha \mathscr{G}U$, combined with Lemma 2.1.27. ∎

Note that we can employ these graphs with open edges to obtain a formalisation of string diagrammatic representation of the top two dimensions of our shapes of diagrams. This is done according to the following technique:

- Each node vertex stays a vertex.
- If a wire vertex has only one incoming or outgoing arrow, then we replace the arrow and the wire vertex with an edge and connect it to the node vertex that was acting as the arrow's source or target.
- If a wire vertex has one incoming and one outgoing arrow, then we connect the two edges that replaced the arrows.
- If we have a wire vertex with no incoming or outgoing edges, we replace it with an edge, thinking of it as the midpoint of that edge.

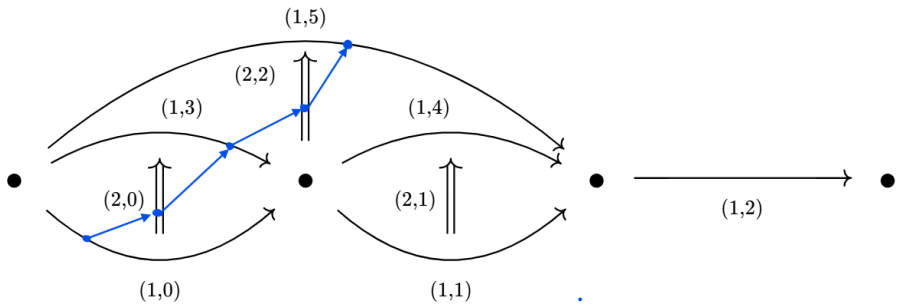
For example, the string diagram representation of the molecule from figure 2.1 obtained from 2.2 (where we extended the input and output edges to reach the same height) is:




**Lemma 2.3.10.** *Let $U$ be a molecule, $n \coloneqq \dim U > 0$, and $x \in U_n$. Then there exist $y_- \in \Delta^- U$ and $y_+ \in \Delta^+ U$ such that there is a path from $y_-$ to $y_+$ passing through $x$ in $\mathscr{G}U$.*

*Proof.* We construct a path $x = x_0 \to y_0 \to \ldots \to x_m \to y_+$ by successive extensions; the construction of a path from $y_-$ to $x$ is dual. Suppose we have reached $x_i$. By Lemma 2.2.30 $\operatorname{cl}\{x_i\}$ is an atom, so $\partial^+ x_i$ is $(n-1)$-dimensional and $\Delta^+ x_i$ is non-empty. Pick $y_i$ in $\Delta^+ x_i$. If $y_i$ has no input cofaces, then $y_i \in \Delta^+ U$, so we can let $m \coloneqq i$ and $y_+ \coloneqq y_i$. Otherwise, pick $x_{i+1} \in \nabla^- y_i$. Since $\mathscr{G}U$ is finite and acyclic by Proposition 2.3.7, this procedure must terminate after a finite number of steps. ∎

***Example* 2.3.11 —** By Lemma 2.3.10, given an $n$-dimensional element, $x$, we can "walk" through the molecule by alternating between its $n$-dimensional and $(n-1)$-dimensional elements, starting with a $y_-$ in the input face and ending with a $y_+$ in the output face.
Suppose we take the element $(2,2)$ from the molecule from Example 2.3.8. Then, we can construct a path $y_- = (1,0) \to (2,0) \to (1,3) \to x = (2,2) \to y_+ = (1,5)$ which is the following path in $U$:




The directed graphs with open edges allow us to prove that whenever an isomorphism exists between two molecules, then that isomorphism is unique.

**Proposition 2.3.12.** *Let $U$, $V$ be molecules. If $U$ and $V$ are isomorphic, there exists a unique isomorphism $\varphi\colon U \xrightarrow{\sim} V$.*

*Proof.* It suffices to prove that if $\imath : U \xrightarrow{\sim} U$ is an automorphism, then it is the identity.
We proceed by induction on $n \coloneqq \dim U$. If $n = 0$, then $U = 1$ by Lemma 4.4.1 (note that the proof of the lemma does not produce any cross dependencies), and the only endomorphism of 1 is the identity.
Suppose $n > 0$ and let $\alpha \in \{+, -\}$. By Proposition 2.2.25, $\partial^\alpha U$ is a molecule of dimension $n-1$, and $\imath(\partial^\alpha U) = \partial^\alpha U$. By the inductive hypothesis, the restriction of $\imath$ to $\partial^\alpha U$ is the identity.
Let $x \in \mathscr{M}ax\, U$, and suppose $\imath(x) = x$. Then $\imath(\partial^\alpha x) = \partial^\alpha x$. By Lemma 2.2.30, $\operatorname{cl}\{x\}$ is an atom, so $\partial^\alpha x$ is a molecule of dimension strictly lower than $n$. By the inductive hypothesis the restriction of $\imath$ to $\partial^\alpha x$ is the identity. Since $\operatorname{cl}\{x\} = (\partial^- x \cup \partial^+ x) + \{x\}$, it follows that $\imath$ restricts to the identity on $\operatorname{cl}\{x\}$. Therefore, it suffices to prove that $\imath$ fixes all $x \in \mathscr{M}ax\, U$.
If $\dim x < n$, then $x \in \partial^\alpha U$, and we have already proved $\imath(x) = x$. Suppose then $\dim x = n$, and construct a path $y_- = y_0 \to x_0 \to \ldots \to y_m \to x_m = x$ in $\mathscr{G}U$ as in Lemma 2.3.10. Since $\imath$ preserves the covering relation and orientations, it maps this path to another path in $\mathscr{G}U$. We have $y_0 \in \partial^- U$, so $\imath(y_0) = y_0$. Suppose $\imath(y_i) = y_i$. Since $y_i$ is a wire vertex in a directed graph with open edges, $x_i$ is the only node vertex with an edge from $y_i$, so necessarily $\imath(x_i) = x_i$. If $i < m$, then $\imath$ is the identity on $\operatorname{cl}\{x_i\}$, so $\imath(y_{i+1}) = y_{i+1}$. Iterating until we reach $m$, we conclude. ∎

So from now, on we will refer to $U \#_k V \coloneqq U \#_k^\varphi V$ simply as the pasting of $U$ and $V$ along their $k$-dimensional boundary. Similarly, for the atom rule, we will write $U \Rightarrow V \coloneqq U \Rightarrow^\varphi V$ and say that $U$ rewrites into $V$.
Moreover, pasting of molecules has the following properties:

**Proposition 2.3.13.** *The following hold for all molecules:*

*I Let $U$, $V$, $W$ be molecules and $k \in \mathbb{N}$ such that $U \#_k V$ and $V \#_k W$ are both defined. Then $(U \#_k V) \#_k W$ and $U \#_k (V \#_k W)$ are both defined and uniquely isomorphic.*

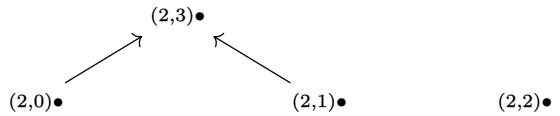
*II Let $U$ be a molecule and $k \in \mathbb{N}$. Then $U \#_k \partial_k^+ U$ and $\partial_k^- U \#_k U$ are both defined and uniquely isomorphic to $U$.*

*III Let $U, U', V, V'$ be molecules and $k < n \in \mathbb{N}$ such that $(U \#_n U') \#_k (V \#_n V')$ is defined. Then $(U \#_k V) \#_n (U' \#_k V')$ is defined and uniquely isomorphic to $(U \#_n U') \#_k (V \#_n V')$.*

The reader familiar with higher category theory, may recognise the first property as the associativity of pasting, the second property as unitality and the third as the interchange law. In fact, in Chapter 5 we will show that the molecules together with the operations $\#_k$ and $\partial_k^\alpha$ with $\alpha \in \{+, -\}$ and $n \in \mathbb{N}$ form a *strict $\omega$-category*.

### 2.3.2 Flow graphs

The connectivity of a diagram at dimension $k$ is represented with a graph as the one below. Note that in this case we are not concerned with the number of $k$-dimensional elements that are shared between elements of dimension greater than $k$; we are interested whether such an element exists. What we care about is whether we can "travel" from the $n$-dimensional element $x$ to the $m$-dimensional element $y$ through $k$-dimensional elements. For the diagram from figure 2.1 and $k = 1$, such a graph would look like:




We call this the *flow graph* and it has the following definition:

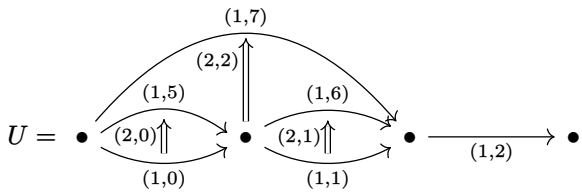
***Definition* 2.3.14** (Flow graph) **—** Let $P$ be an oriented graded poset, $k \geq -1$. The *$k$-flow graph of $P$* is the directed graph $\mathscr{F}_k P$ whose

- set of vertices is $\bigcup_{i>k} P_i$, and
- set of edges is
$$\left\{(x, y) \mid \Delta_k^+ x \cap \Delta_k^- y \neq \varnothing\right\},$$
with $s\colon (x, y) \mapsto x$ and $t\colon (x, y) \mapsto y$.

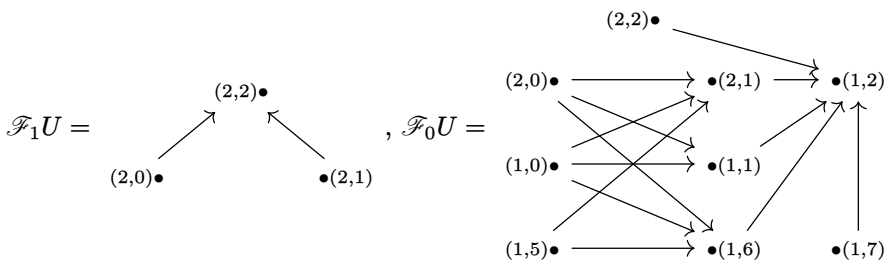
***Definition* 2.3.15** (Maximal flow graph) **—** Let $P$ be a finite-dimensional oriented graded poset, $k \geq -1$. The *maximal $k$-flow graph of $P$* is the induced subgraph $\mathscr{M}_k P$ of $\mathscr{F}_k P$ on the vertex set
$$\bigcup_{i>k} (\mathscr{M}ax\, P)_i \subseteq \bigcup_{i>k} P_i.$$

**Remark 2.3.16.** *For $k \coloneqq \dim P - 1$, $\mathscr{F}_k P$ and $\mathscr{M}_k P$ coincide.*

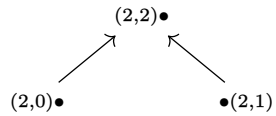
***Example* 2.3.17 —** Take the following molecule:

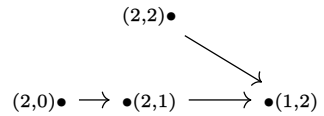



The flow and maximal flow graphs are:



Note how in $\mathscr{F}_0 U$ there are no edges between $(2,0)$ and $(2,2)$ and between $(2,1)$ and $(2,2)$ as $\Delta_0^+(2,0) \cap \Delta_0^-(2,2) = \emptyset$ and $\Delta_0^+(2,1) \cap \Delta_0^-(2,2) = \emptyset$.

$\mathscr{M}_1 U = \mathscr{F}_1 U =$ 

, $\mathscr{M}_0 U =$ 


Also note that the graphs above are all acyclic. This will become relevant when we talk about the acyclicity conditions of molecules and will lead to nice results in the problem of subdiagram matching.

***Definition* 2.3.18** (Topological sort) — Let $\mathscr{G}$ be a directed acyclic graph with finite set of vertices, $m \coloneqq |V_{\mathscr{G}}|$. A *topological sort of* $\mathscr{G}$ is a linear ordering $(x^{(i)})_{i=1}^m$ of $V_{\mathscr{G}}$ such that, for all edges $e \in E_{\mathscr{G}}$, if $s(e) = x^{(i)}$ and $t(e) = x^{(j)}$, then $i < j$.

***Definition* 2.3.19** (Ordering of a molecule) — Let $U$ be a molecule, $k \geq -1$, and suppose $\mathscr{M}_k U$ is acyclic. A *$k$-ordering of $U$* is a topological sort of $\mathscr{M}_k U$.

***Example* 2.3.20** — Consider the molecule from Example 2.3.17. It has two 1-orderings: $[(2,0),(2,1),(2,2)]$, $[(2,1),(2,0),(2,2)]$ and three 0-orderings given by: $[(2,0),(2,1),(2,2),(1,2)]$, $[(2,0),(2,2),(2,1),(1,2)]$ and $[(2,2),(2,0),(2,1),(1,2)]$.

**Lemma 2.3.21.** *Let $\imath \colon V \hookrightarrow U$ be an inclusion of oriented graded posets, $k \geq -1$. Then $\mathscr{F}_k V$ is isomorphic to the induced subgraph of $\mathscr{F}_k U$ on the vertices in the image of $\imath$.*

*Proof.* See [25, Lemma 8.3.9]. ∎

The following lemma shows that the acyclicity of the maximal flow graph is preserved under pasting. We will use it in Chapter 4 when we will make the connection between orderings and a certain type of pasting decompositions.

**Lemma 2.3.22.** *Let $U$, $V$ be molecules and $k < \min\{\dim U, \dim V\}$ such that $U \#_k V$ is defined. If $\mathscr{M}_k U$ and $\mathscr{M}_k V$ are acyclic, then $\mathscr{M}_k(U \#_k V)$ is acyclic.*

## 2.4 Computational aspects of building molecules

To study our framework from the point of view of a model of computation, we need to make sure that working in this model of computation is indeed feasible. We want to ensure that building the shapes of diagrams in our framework is computationally inexpensive. Out of the operations that we use to construct our molecules, all but one admit straightforward low-degree polynomial time solutions. The problem of molecule isomorphism generalised to all oriented graded posets is equivalent to the graph isomorphism problem (Proposition 2.4.5) whose best known running time to day is quasipolynomial [4]. However, due to their rigid structure, we show that the isomorphism problem restricted to the class of molecules admits a low-degree polynomial time solution. Our solution to the molecule isomorphism problem is a deterministic algorithm, the *traversal algorithm*, which outputs for each molecule

a unique ordering of its elements. Moreover, our solution justifies why we can represent the elements of a molecule in the (dimension, position) format.
Most of the content of this chapter is part of Sections 1 and 2 from [28]. The algorithms and the associated data structures presented in this chapter are part of a tool called `rewalt`. Even though we do not present the tool in this thesis, we will sometimes make a reference to the implementation details that are part of it.

### 2.4.1 Data structures for oriented graded posets

We mentioned in the introduction of this thesis the distinction between a diagram and its shape. Moreover, in our treatment of molecules we named each element individually. Note that this naming convention, which will be a consequence of the traversal algorithm is different than the labelling that is part of the definition of a diagram.

***Definition* 2.4.1 —** Let $\mathbb{V}$ be a set of variables. A diagram is a pair $(U, t)$, where $U$ is a molecule and $t : U \to \mathbb{V}$ is a labelling. We call $t$ a diagram of shape $U$.

***Definition* 2.4.2 —** A *cell* is a diagram whose shape is an atom.

**Remark 2.4.3.** *We can extend the definition of atom, $\#_k$ and $\partial^\alpha_n$ operators, for all $n, k \in \mathbb{N}$ and $\alpha \in \{+, -\}$ to the setting of diagrams.*
*The input/output $k-$boundary of $t$ is the diagram: $\partial^\alpha_k t = \imath^\alpha_k ; t$, where $\imath^\alpha_k : \partial^\alpha_k U \hookrightarrow U$. If $t \colon U \to \mathbb{V}$ is a diagram such that $U$ decomposes as $U = U_1 \#_k U_2$ then for $\imath_i \colon U_i \hookrightarrow U$, we can write $t = t_1 \#_k t_2$, where $t_i = \imath_i ; t$ for $i \in \{1, 2\}$. If $U$ is an atom such that $U = \operatorname{cl} x$, then $t = u^- \overset{t(x)}{\Rightarrow} u^+$ where $u^\alpha = \partial^\alpha t$.*

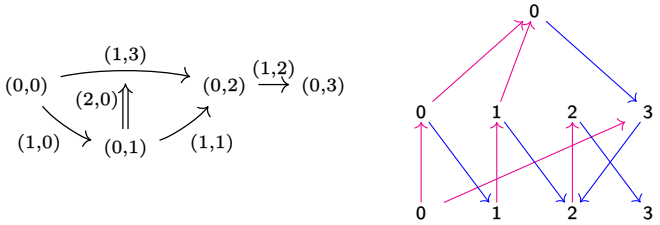
We continue with the treatment of data structures and algorithms for molecules. By Remark 2.4.3, these results can be extended to diagrams. Before presenting our solution to the isomorphism problem, we introduce the data structures that we use to represent oriented graded posets.
If we linearly order the elements of an oriented graded poset in each dimension, each element $x$ is uniquely identified by a pair of integers $(n, k)$, where $n$ is the *dimension* of $x$, and $k$ is the *position* of $x$ in the linear ordering of $n$-dimensional elements. In the implementation, there are methods that return, for each element, the first and second projection - i.e., its dimension and position.
We then represent an oriented graded poset as a pair $(\mathsf{face\_data}, \mathsf{coface\_data})$ of *arrays of arrays of pairs of sets of integers*, where

I $j \in \mathsf{face\_data}[n][k][i]$ if and only if $(n-1, j)$ is covered by $(n, k)$, and
II $j \in \mathsf{coface\_data}[n][k][i]$ if and only if $(n+1, j)$ covers $(n, k)$

with orientation $-$ $(i = 0)$ or $+$ $(i = 1)$. We may implement the sets of integers as sorted arrays, or another data type which supports binary search in logarithmic time. This defines a data type OgPoset.
This representation is essentially an adjacency list representation of the poset's Hasse diagram, with vertices separated according to their dimension, and incoming and outgoing edges separated according to their label. If $P$ is an oriented graded poset and $E_P$ is the set of edges of the Hasse diagram of $P$, the OgPoset representation of $P$ takes space $O(|P| + |E_P|)$.

Below we present an example of a shape of a diagram together with its oriented Hasse diagram representation and its face and coface data:

***Example* 2.4.4 —**




face_data:

| | | | |
|---|---|---|---|
| $([],[])$ | $([],[])$ | $([],[])$ | $([],[])$ |
| $([0],[1])$ | $([1],[2])$ | $([2],[3])$ | $([0],[2])$ |
| $([0,1],[3])$ | | | |

coface_data:

| | | | |
|---|---|---|---|
| $([0,3],[])$ | $([1],[0])$ | $([2],[1,3])$ | $([],[2])$ |
| $([0],[])$ | $([0],[])$ | $([],[])$ | $([],[0])$ |
| $([],[])$ | | | |

Each row in the face or coface data represents the dimension of the element - the first index of the array. Each column represents the position of the element in the linear order - the second index in the array. For example, the tuple containing $([0],[1])$ from face_data corresponds to element $(1,0)$ in the pasting diagram, while the tuple containing $([0,3],[])$ from the coface_data corresponds to $(0,0)$.
Now, let us look at the face_data. The first set from each pair represents the input faces while the second set represents the output faces. For example, element $(2,0)$ - the pair on the third row, first column - has elements $(1,0)$ and $(1,1)$ in its input face and element $(1,3)$ in its output face. So we store $0,1$ in the first component of the pair and 3 in the second component.
In case of the coface_data, the first component from each pair represents the input cofaces while the second component represents the output cofaces. For example, element $(0,2)$ is covered with $+$ by elements $(1,1)$ and $(1,3)$ and with $-$ by element $(1,2)$, hence the pair on the first row, third column in the coface_data has $[2]$ in its first component and $[1,3]$ in its second component.

Storing both face_data and coface_data is redundant since these are uniquely determined by each other. However, most of the computations we need to perform on oriented graded posets require regular access both to faces (covered elements) and cofaces (covering elements) of a given element, so it is advantageous to be able to access them in constant time. If we want to get the coface data from the face data, we do the following for each element $(n,k)$ in the face data: for each $j$ in the input or output set of $(n,k)$, add $k$ to the input or output set of $(n-1,j)$ in the coface data. Let's look at Example 2.4.4. Take element $(2,0)$ in the face data; the information tells us that $(1,0)$ and $(1,1)$ are covered with $-$ by $(2,0)$ and $(1,3)$

is covered with $+$ by $(2,0)$. In coface_data this translates as: in the second row we add 0 to the first component of the first two pairs and we add 0 to the second component of the last tuple.
We represent a set of elements of an OgPoset as an array of sets of positions, indexed by dimensions. This allows us to access the subset of elements of a given dimension in constant time. The size of arrays can be fixed to be equal to the dimension of a specific OgPoset, or dynamically adjusted to the dimension of each set of elements. Sets of positions can again be implemented as sorted arrays. This defines a data type for graded sets.
Moreover, we represent a map $f : P \to Q$ as an array of arrays of pairs of integers called `mapping`, together with pointers source, target to the OgPoset representations of $P$ and $Q$. This defines a data type OgMap. As an array of arrays, `mapping` has the same size of $P$'s face_data, and is defined by:

$$\texttt{mapping}[n][k] = (m, j) \text{ if and only if } f((n,k)) = (m,j).$$

This representation takes space $O(|P|)$.

### 2.4.2 The traversal algorithm

There are four steps involved in constructing molecules:

I compute the input and output boundaries of a molecule,

II check if a closed subset is round,

III determine if two molecules are isomorphic,

IV compute the pushout of a span of inclusions.

The first, second and fourth operations admit straightforward low-degree polynomial time solutions which we provide in subsection 2.4.5. However, as mentioned in the chapter's introduction, the third problem generalised to all oriented graded posets is equivalent to the graph isomorphism problem (GI) whose best running time to date is quasipolynomial [4].

**Proposition 2.4.5.** *The isomorphism problem for oriented graded posets is* GI *complete.*

*Proof.* Deciding isomorphism of oriented graded posets is equivalent to deciding isomorphism of their Hasse diagrams with $\{+,-\}$-labelled edges. The isomorphism problem for edge-labelled finite graphs is an instance of the isomorphism problem for finite relational structures, which is GI-complete [40].
Conversely, one can represent a directed graph by its "oriented incidence poset": the 0-dimensional elements are the vertices, the 1-dimensional elements are the edges, the only input face of an edge is its source, and the only output face of an edge is its target. Two directed graphs are isomorphic if and only if their oriented incidence posets are isomorphic. Since GI reduces to the isomorphism problem for directed graphs, it reduces to the isomorphism problem for 1-dimensional oriented graded posets. ∎

In the rest of the section we show that restricting to the subclass of oriented graded posets given by molecules, we can do much better.
The isomorphism problem for diagrams is the following decision problem:

***Definition* 2.4.6** (Diagram isomorphism problem) — Given diagrams $t\colon U \to V$ and $t'\colon U' \to V$, does there exist an isomorphism $\phi\colon U \to U'$ of their shapes such that $t = \phi; t'$?

Then, by Remark 2.4.3, once we found an isomorphism between $U$ and $U'$, verifying that the labellings of the diagrams match can be checked in linear time.
A high-level description of the molecule isomorphism problem for two molecules is as follows: check that the face data matches - this takes quadratic time in the size of the molecule. If this fails, the molecules are not isomorphic; otherwise, we run the traversal algorithm for both of them to obtain a linear order. We finally verify if two molecules are isomorphic by linearly ordering their elements in each dimension according to the traversal order and checking that these orders match.
Checking that the face data of two molecules matches is equivalent to verifying that:

- the molecules have the same dimension,
- in each dimension both molecules have the same number of elements,
- the molecules are made of the same shapes (given by the number of elements in the input and output face).

The procedure for checking these is:

```
    procedure PREVIOUSCHECKS(U, V : molecule)
        if face_data(U).size() == face_data(V).size() then
3:          for u ∈ face_data(U) and v ∈ face_data(V) do
                if u.size() == v.size() then
                    for x ∈ u do
6:                      Search for a y ∈ v such that x.fst.size() == y.fst.size()
    and x.snd.size() == y.snd.size().
                        Delete y from v.
                    if v ≠ ∅ then return False
9:              else return False
            return True
        return False
```

Line 2 of the algorithm verifies that the dimension of $U$ is equal to the dimension of $V$. In Line 4 we check that $U$ and $V$ have the same number of elements in each dimension. Inside the for loop from line 5 we check whether $U$ and $V$ have the same number of input and output faces. The runtime of this algorithm is bounded above by $O(|U|)^2$, where $|U|$ is the cardinality of $U$, i.e. the number of vertices in the Hasse diagram representation.

**Remark 2.4.7.** *Note that since the face and coface data can be determined from each other, it is sufficient to only run these checks for one of them. The exact same procedure as above can be applied for checking that the coface data of $U$ and $V$ match.*

We solve the isomorphism problem between two molecules with the help of a deterministic *traversal algorithm*. This algorithm takes a molecule of type OgPoset and returns a unique ordering of its elements relying only on the intrinsic structure of the oriented graded poset.

***Procedure* 2.4.8** (Traversal algorithm) — The procedure takes as input a molecule $U$ and returns a list of its elements in the order in which they are *marked*. It uses an auxiliary stack of molecules $V \subseteq U$.
At the beginning, only $U$ is on the stack and all elements are unmarked. We iterate the *main loop* until the stack is empty, at which point the procedure terminates.
At each iteration, suppose $V$ is on top of the stack. If all elements of $V$ are marked, then we pop $V$ from the stack and iterate. Else, if any elements of $\partial^- V$ are unmarked, we push $\partial^- V$ to the top of the stack and iterate. Else, if $V = \mathrm{cl}\,\{x\}$ for some $x \in U$, we

I mark $x$ and pop $V$ from the stack,

II if any elements of $\partial^+ V$ are unmarked, we push $\partial^+ V$ to the top of the stack, and

III we iterate.

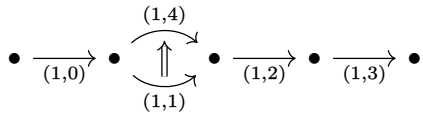
If $V \neq \mathrm{cl}\,\{x\}$, we let $y$ be the earliest marked element such that $\dim y = \dim V - 1$ and there is an unmarked $x \in \nabla^- y \cap V$. Such a $y$ always exists, and then $\nabla^- y \cap V = \{x\}$. We push $\mathrm{cl}\,\{x\}$ to the top of the stack and iterate.

Below we give the pseudocode of the traversal algorithm:

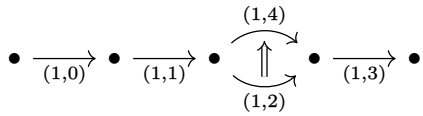
```
    procedure TRAVERSE(U : molecule)
        marked ← []
        stack ← [U]
        while stack is not empty do
 5:         focus ← top of stack
            dim ← dim(focus)
            if focus ⊆ marked then
                pop focus from top of stack
            else
10:             if ∂⁻focus ⊈ marked then
                    push ∂⁻focus to top of stack
                else
                    if focus = cl{x} for some x then
                        append x to marked
15:                     pop focus from top of stack
                        if ∂⁺focus ⊈ marked then
                            push ∂⁺focus on top of stack
                    else
                        y ← first item of dimension dim − 1 in marked such
20:                         that y has an unmarked input coface in focus
                        x ← unique input coface of y in focus
                        push cl{x} on top of stack
        return marked
```

Let's take a look at some examples.

***Example* 2.4.9 —** Consider the following two molecules in which the 0-dimensional elements are labelled in both molecules from left to right as follows: $(0,0), (0,1), (0,2), (0,3), (0,4)$. The 2-dimensional element is $(2,0)$.




face_data:
([], [])  ([], [])  ([], [])  ([], [])  ([], [])
([0], [1])  ([1], [2])  ([2], [3])  ([3], [4])  ([1], [2])
([1], [4])
coface_data:
([0], [])  ([1, 4], [0])  ([2], [1, 4])  ([3], [2])  ([], [3])
([], [])  ([0], [])  ([], [])  ([], [])  ([], [0])
([], [])

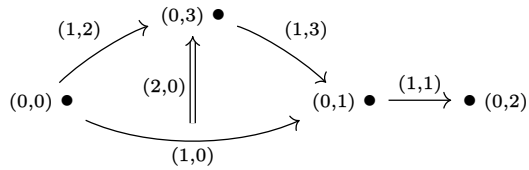



face_data:
([], [])  ([], [])  ([], [])  ([], [])  ([], [])
([0], [1])  ([1], [2])  ([2], [3])  ([3], [4])  ([1], [2])
([2], [4])
coface_data:
([0], [])  ([1], [0])  ([2, 4], [1])  ([3], [2, 4])  ([], [3])
([], [])  ([], [])  ([0], [])  ([], [])  ([], [0])
([], [])

With the elements ordered according to the traversal order, one can see that we can distinguish between the two lists.

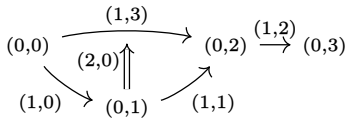
***Example* 2.4.10 —** The traversal order for the molecule in Example 2.1.4 is:
{ (0, 0), (1, 0), (0, 1), (1, 1), (0, 2), (1, 2), (0, 3), (2, 0), (1, 3) }.

***Example* 2.4.11 —** Consider the molecule below. Its traversal order is: {(0, 0), (1, 0), (0, 1), (1, 1), (0, 2), (2, 0), (1, 2), (0, 3), (1, 3)} which is different from the one in Example 2.1.4 that we used to exemplify the algorithm.




However, the traversal algorithm alone is not sufficient to solve the isomorphism problem as the following example shows:

***Example* 2.4.12 —** Consider the following two shapes together with their face and coface data in which we sorted the elements inside the face and coface data according to traversal order.

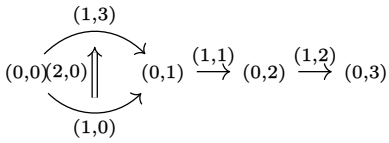



face_data:

| | | | |
|---|---|---|---|
| ([], []) | ([], []) | ([], []) | ([], []) |
| ([0], [1]) | ([1], [2]) | ([2], [3]) | ([0], [2]) |
| ([0, 1], [3]) | | | |

coface_data:

| | | | |
|---|---|---|---|
| ([0, 3], []) | ([1], [0]) | ([2], [1, 3]) | ([], [2]) |
| ([0], []) | ([0], []) | ([], []) | ([], [0]) |
| ([], []) | | | |




face_data:

| | | | |
|---|---|---|---|
| ([], []) | ([], []) | ([], []) | ([], []) |
| ([0], [1]) | ([1], [2]) | ([2], [3]) | ([0], [1]) |
| ([0], [3]) | | | |

coface_data:

| | | | |
|---|---|---|---|
| ([0, 3], []) | ([1], [0, 3]) | ([2], [1]) | ([], [2]) |
| ([0], []) | ([], []) | ([], []) | ([], [0]) |
| ([], []) | | | |

The traversal order for both molecules is: $\{(0,0),(1,0),(0,1),(1,1),(0,2),(1,2),(0,3),(2,0),(1,3)\}$, but clearly the two shapes are different. There are two points to be drawn from this example. Firstly, it is the dimension-wise order that needs to be checked for isomorphism (which can be seen by looking at the face data of the two molecules). Secondly, the traversal algorithm is an expensive procedure. Verifying the face data for these molecules would have made the isomorphism algorithm to return false at an earlier stage.

Combining the two methods, we obtained the desired results, i.e., two molecules are isomorphic if and only if they have the same traversal order dimension-wise and their face data matches. The reverse implication can be proved by defining a bijective map between the corresponding traversal order which is compatible with the faces operator. These checks are independent of each other so they may be performed in either order.

### 2.4.3 Correctness and runtime analysis

The algorithm that we use to check for isomorphism of two molecules, $U$ and $V$ is the following:

- first run `PreviousChecks` and then check that the traversal algorithm returns the same list for both $U$ and $V$.

```
    procedure IsIso(U, V : molecule)
        if PreviousChecks(U, V) then
3:          U' = Traverse(U)
            V' = Traverse(V)
            return U' == V'
6:      else return False
```

The worst case running time of this algorithm is $O(n^2 log(n))$ where $n$ is the number of vertices and edges of the Hasse diagram representation of the molecule. Verifying whether the face_data matches takes $O(m^2)$, where $m$ is the number of elements in the molecule (vertices in the Hasse diagram).
In the remainder of the subsection, we prove the correctness of the isomorphism algorithm. More specifically, we show that for two molecules, $U$ and $V$, if $U' = \text{traverse}(U)$ and $V' = \text{traverse}(V)$ (where $U'$ and $V'$ are the versions of $U$ and $V$ with the elements reordered in each dimension according to their traversal order), then:

$$U \simeq V \text{ if and only if } U' \equiv V'.$$

We will show that, with this strategy, we can solve the isomorphism problem for molecules in time $O(n^2 \log n)$. A more precise upper bound is given in Theorem 2.4.21 below. In addition to solving the isomorphism problem, the traversal order gives us a canonical form for molecules. If we linearly order the elements in each dimension according to the traversal order, we obtain a unique representation for shapes of diagram. Note that the second component in our naming convention for elements in a molecule - (dimension, position) - comes from the traversal order.

**Remark 2.4.13.** *We can implement the constructors $\partial_n^\alpha$, atom and $\#_k$ in such a way that the elements are arranged in traversal order after each step. With this implementation, we can check molecules for equality instead of for isomorphism.*

In order to prove the correctness of the traversal algorithm, we need some preliminary results.

**Lemma 2.4.14.** *Let $V$ be an item on the stack. Then $V$ is a molecule. If $W$ is below $V$ on the stack, then $V$ is a proper subset of $W$.*

*Proof.* Initially, the stack only contains $U$, which is a molecule by assumption. Assume, inductively, that the statement is true at the beginning of the current iteration with focus $V$, and that a set $V'$ is pushed onto the stack at the end. Then either

I $V' = \partial^\alpha V$ for some $\alpha \in \{+, -\}$, or

II $V' = \text{cl}\{x\}$ for some $x \in V$.

In both cases, $V'$ is a molecule and a proper subset of $V$ (hence also of each item below $V$), under the assumption that $V$ is a molecule. ∎

**Remark 2.4.15.** *In fact, any $V$ that appears on the stack is either $\partial_k^- U$, which we call "$U$-linked", or it is $\text{cl}\{x\}$ or $\partial_k^\alpha x$, which we call "$x$-linked", for some $x \in U$. In the latter case, $V$ is* round, *which implies that it is also* pure *2.2.11: its maximal elements all have the same dimension.*

**Lemma 2.4.16.** *Suppose $V$ is on the stack. Then all elements of $V$ must be marked before any item below $V$ is accessed, or before any proper superset of $V$ becomes the focus.*

*Proof.* By Lemma 2.4.14, as long as $V$ is on the stack, only $V$ and its proper subsets can be on top. It follows that, for a proper superset of $V$ to be the focus, $V$ must be popped from the stack at the end of an iteration where $V$ is the focus. There are only two ways this can happen:

- $V$ was already fully marked before the current loop iteration, or
- $\partial^- V$ was fully marked and $V = \mathrm{cl}\{x\}$ for some $x$ which is marked at the current loop iteration.

In both cases, $\partial^- V$ was already fully marked before the current loop iteration. In the latter case, if $\partial^+ V$ is already fully marked, then $V = \{x\} \cup \partial^- V \cup \partial^+ V$ is also fully marked. Otherwise, $\partial^+ V \subsetneq V$ gets pushed onto the stack to replace $V$, and must be popped before any superset of $V$ becomes the focus. By the same case distinction, whenever $\partial^+ V$ is popped, either

- it was fully marked, in which case $V$ was fully marked, or
- it is of the form $\mathrm{cl}\{y\}$ for some $y$ which is marked at the current loop iteration.

In both situations, since all molecules satisfy the *globularity* property $\partial^\alpha(\partial^+ V) = \partial^\alpha(\partial^- V) \subseteq \partial^- V$, we know that $\partial^+ V$, hence $V$, is fully marked at the end of the iteration, and nothing is added to the stack. ∎

**Lemma 2.4.17.** *Any subset $V$ of $U$ can be pushed onto the stack at most once.*

*Proof.* Suppose $V$ is pushed onto the stack. As long as $V$ is on the stack, any subsequent addition to the stack must be a proper subset of $V$, so it cannot be equal to $V$.
If $V$ is popped from the stack, by Lemma 2.4.16, it must be fully marked before any item below it is accessed. Since the algorithm checks if a set is fully marked before pushing it onto the stack, $V$ can never appear again. ∎

**Lemma 2.4.18.** *Let $V$ be the focus, $n \coloneqq \dim V$. Then either $V$ is fully marked, or there exists an $n$-dimensional element of $V$ which is unmarked.*

*Proof.* First, we prove a weaker result: either $V$ is fully marked, or there exists a *maximal* element of $V$ which is unmarked.
Let $x \in V$ be marked. At some prior iteration, $\mathrm{cl}\{x\}$ must have been the focus, and by Lemma 2.4.16, in order for $V$ to become the focus, $\mathrm{cl}\{x\}$ must have been fully marked as well. Because

$$V = \bigcup_{k \le n} \mathrm{cl}\,(\mathit{Max}\, V)_k = \bigcup_{k \le n} \bigcup_{x \in (\mathit{Max}\, V)_k} \mathrm{cl}\,\{x\},$$

it follows that $V$ is fully marked if and only if its maximal elements are all marked.

Now, $V$ has one of the two forms in Remark 2.4.15. If $V$ is of the second form, its maximal elements all have the top dimension, so we only need to consider the case $V = \partial^-_k U$.
At the start of the algorithm, $U, \dots, \partial^-_0 U$ are all consecutively added to the stack. So $\partial^-_k U$ becomes the focus either at this stage, in which case *all* its elements are unmarked, or after $\partial^-_{k-1} U$ is fully marked. In the latter case, any maximal element of $\partial^-_k U$ of dimension strictly smaller than $k$ also belongs to $\partial^-_{k-1} U$. ∎

Finally, we show that the traversal algorithm always terminates; given a molecule, $U$, it returns a unique ordering of its elements.

**Theorem 2.4.19.** *The traversal algorithm is correct: given a molecule $U$, it terminates returning a unique linear ordering of the elements of $U$.*

*Proof.* As a particular case of Lemma 2.4.16, $U$ must be fully marked before the stack is emptied. Therefore, the algorithm either terminates after all elements have been traversed, or it does not terminate.
To prove that the algorithm does always terminate, it suffices to show that, unless all elements are already marked, it always finds an element to mark. First of all, observe that, from any state, the algorithm first goes through the following sequence of steps:

I popping all fully marked subsets from the top of the stack;

II once it reaches a subset which is not fully marked, successively pushing its lower-dimensional input boundaries that are not fully marked onto the stack.

At the end of this sequence, we always reach a state in which the focus $V$ is not fully marked, but $\partial^- V$ is fully marked. Let us call such a $V$ a *proper* focus.
We proceed by induction on dimension and proper subsets of a proper focus. If $\dim V = 0$, since a 0-molecule always consists of a single element, $V = \{x\}$, and $x$ gets marked at the current iteration.
Let $n \coloneqq \dim V$. By Lemma 2.4.18, there is an unmarked $x \in V_n$. If $V = \mathrm{cl}\{x\}$, then $x$ is marked at the current iteration, and we are done. Otherwise, we prove that there always exists a pair $(y, x)$ where $x \in V_n$ is unmarked, and $y$ is a marked input face of $x$. By Corollary 2.3.9 applied to $V$, the coface $x$ is unique given $y$, so among such pairs we can pick the one where $y$ comes *earliest* in the list of marked elements, and this selects a unique $x$.
Let $x \in V_n$ be unmarked. By Lemma 2.3.10, there exists a sequence

$$y_0 \to x_0 \to \dots \to y_m \to x_m = x$$

where $y_0 \in \Delta^-_{n-1} V$, $x_i \in V_n$, $y_i$ is an input face of $x_i$, and $y_{i+1}$ is an output face of $x_i$. Since $V$ is a proper focus, $y_0$ is marked. Let $k$ be the smallest index such that $x_k$ is unmarked; because $x_m$ is unmarked, such a $k$ exists. Then $x_i$ is marked for all $i < k$, hence $\mathrm{cl}\{x_i\}$ is also marked. It follows that $y_k \in \partial^+ x_{k-1}$ is marked, and the pair $(y_k, x_k)$ satisfies our requirement.
Thus, the algorithm will find a unique $x \in V_n$ and push $\mathrm{cl}\{x\}$ onto the stack. The next proper focus will necessarily be a proper subset of $V$, and we conclude by the inductive hypothesis. ∎

***Definition* 2.4.20 —** For a molecule $U$, and all $k \in \mathbb{N}$, we let

$$\begin{aligned} \mathscr{E}_k U &\coloneqq \sum_{x \in U_k} \Delta x = \sum_{y \in U_{k-1}} \nabla y, \\ |U_\vee| &\coloneqq \max \{|U_i|\}_{i \in \mathbb{N}}, \\ |\mathscr{E}_\vee U| &\coloneqq \max(\{|\mathscr{E}_i U|\}_{i \in \mathbb{N}} \cup \{1\}). \end{aligned}$$

Note that $\mathscr{E}_k U$ is the set of edges between $k$ and $(k-1)$-dimensional elements in $\mathscr{H}U$. We have $|U_k| \leq |\mathscr{E}_k U|$ for all $k > 0$, while $|\mathscr{E}_0 U| = 0$. Since the maximum of the $|\mathscr{E}_k U|$ is 0 only when $U$ is 0-dimensional, in which case $|U_k| = 1$, with our definition we always have $|U_\vee| \leq |\mathscr{E}_\vee U|$.

**Theorem 2.4.21.** *The traversal algorithm admits an implementation running in time* $O(|U| \; |\mathscr{E}_\vee U| \; \log |\mathscr{E}_\vee U|)$.

*Proof.* First of all, we represent any closed set on the stack with its graded set of maximal elements. To initialise the algorithm, we only need to compute the maximal elements of $U$. This can be done in time $O(|U|)$ by going through the elements of $U$ and checking if their set of cofaces is empty.
We start by showing that there are $O(|U|)$ iterations of the main loop.
We let $k \leq \dim U$ and we count the number of loop iterations where a $k$-dimensional subset $V$ is on top of the stack. This can happen in two ways:

- $V$ is either $U$ or $\partial^\alpha W$ for some $W$ with $\dim W > k$, where $W$ was earlier (and may still be) on the stack,
- $V$ is $\operatorname{cl}\{x\}$ for some $x \in W$, where $\dim W = k$ and $W$ is below $V$ on the stack.

Let $(V^{(i)})_{i=1}^m$ be the sequence of all $k$-dimensional subsets appearing on the stack *in the first way* during the run, in the order in which they appear. For all $j < i \in \{1, \dots, m\}$, by Lemma 2.4.16 $V^{(j)}$ must be fully marked before $V^{(i)}$ can appear on the stack. Moreover, $V^{(i)}$ can be on top at most

I once to push $\partial^- V^{(i)}$ to the top,

II once every time we push $\operatorname{cl}\{x\}$ to the top for an unmarked $x \in (V^{(i)})_k$,

III once to pop $V^{(i)}$ from the stack.

Any $k$-dimensional $\operatorname{cl}\{x\}$ appearing *in the second way* appears while a unique $V^{(i)}$ is on the stack, and at most

I once to push $\partial^- x$ to the top,

II once to mark $x$ and pop $\operatorname{cl}\{x\}$ from the stack.

Let $U_k^{(i)} \coloneqq (V^{(i)})_k \setminus \bigcup_{j<i} (V^{(j)})_k$. Then $U_k^{(i)}$ is precisely the set of unmarked $k$-dimensional elements of $V^{(i)}$ when $V^{(i)}$ first appears on the stack. It follows that the number of loop iterations with a $k$-dimensional subset on top of the stack *while* $V^{(i)}$ *is on the stack* at most $2 + 3\left|U_k^{(i)}\right|$ times.

Since at the end of the procedure all $k$-dimensional elements of $U$ are marked, the $(U_k^{(i)})_{i=1}^m$ form a partition of $U_k$. Thus, the total number of loop iterations where a $k$-dimensional subset is on top of the stack is bounded above by

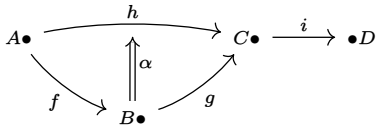
$$\sum_{i=1}^{m} \left(2 + 3\left|U_k^{(i)}\right|\right) = 2m + 3\sum_{i=1}^{m}\left|U_k^{(i)}\right| = 2m + 3\,|U_k| \leq 5\,|U_k|\,.$$

Summing over all dimensions, we get an upper bound of $5\,|U|$ iterations.
What is now left to show is the cost of one loop iteration with focus $V$, where $\dim V = n$. In our implementation, we split the list of marked elements into three objects: a list order (for the total traversal order), an array of lists grorder (for the traversal order split by dimension), and a graded set marked (for the set of marked elements).
**(Line 7).** By Lemma 2.4.18, to check if $V$ is fully marked, it suffices to check whether $V_n \subseteq \mathsf{marked}_n$. Since both are sorted arrays of integers, they can be compared in time linear in $|V_n| + |\mathsf{marked}_n|$, which is $O(|U_n|)$. At this stage, we may also record the unmarked $n$-dimensional elements of $V$ in a sorted array unmarked without affecting the complexity.
**(Lines 10, 16).** To compute the maximal elements of $\partial^- V$ and $\partial^+ V$, we may use different strategies depending on whether $V$ is "$U$-linked" or not.
If $V = \partial_n^- U$, we compute the $(n-1)$-dimensional elements of $\partial^- V = \partial_{n-1}^- U$ simply by going through the elements of $U_{n-1}$ and checking which ones have empty sets of output cofaces, in time $O(|U_{n-1}|)$. Lower-dimensional maximal elements are shared between $V$ and $\partial^- V$, so we may then point from the latter to the former, at no extra cost.
If $V$ is not $U$-linked, $V$ and its boundaries are pure, so the set of maximal elements of $\partial^\alpha V$ is equal to $\Delta^\alpha V$, and each of its elements is covered by an element of $V_n$. To compute it, we add all the input and output faces of all $x \in V_n$ to sets in_faces and out_faces, respectively, then use the relations $\Delta^- V = \mathsf{in\_faces} \setminus \mathsf{out\_faces}$ and $\Delta^+ V = \mathsf{out\_faces} \setminus \mathsf{in\_faces}$.
There are $O(|\mathscr{E}_n U|)$ faces of elements of $V_n$, and we can sort in_faces and out_faces, remove duplicates, and compute their difference in $O(|\mathscr{E}_n U| \cdot \log |\mathscr{E}_n U|)$ which is bounded above by $O(|\mathscr{E}_\vee U| \log |\mathscr{E}_\vee U|)$.
At this stage, we also create an associative array candidates as follows: whenever $x \in V_n$ is in unmarked, and $y$ is an input face of $x$, we add the position of $x$ as a value to candidates, indexed by the position of $y$. We then sort the indices of candidates. This also takes time $O(|\mathscr{E}_n U| \cdot \log |\mathscr{E}_n U|)$ which is bounded above by $O(|\mathscr{E}_\vee U| \log |\mathscr{E}_\vee U|)$ and it does not affect the overall complexity.
By the same reasoning applied to line 7, checking if $\partial^- V$ and $\partial^+ V$ are fully marked takes time $O(|U_{n-1}|)$.
**(Line 14).** If $V_n$ has a single element that we mark, adding it to order and grorder takes constant time with an appropriate implementation of lists. Adding it to marked takes $O(|U_n|)$. That is because one needs to traverse the array at dimension $n$ until an empty position is found to add $x$.
**(Lines 19—21).** To select the next focus we traverse $\mathsf{grorder}_{n-1}$ starting from the first item and search for each item in the indices of candidates until we find a hit $y$. This takes time $O(|U_{n-1}| \cdot \log |U_{n-1}|)$ in the worst case. The next focus will be $\mathrm{cl}\{x\}$, where $x$ is the value corresponding to index $y$.

Overall, we obtain that the worst-case complexity is $O(|U_n| + |\mathscr{E}_\vee U| \cdot \log|\mathscr{E}_\vee U| + |U_{n-1}| \cdot \log|U_{n-1}|)$. Using the bounds $|U_n|, |U_{n-1}| \leq |U_\vee|$ and the fact that $|U_\vee| \leq |\mathscr{E}_\vee U|$ from Definition 2.4.20, we obtain an upper bound of $O(|\mathscr{E}_\vee U| \log|\mathscr{E}_\vee U|)$ for one iteration of the while loop. Multiplying by our bound on the number of iterations, we conclude. ∎

### 2.4.4 Traversal algorithm on an example

Consider the following shape of diagram, in which we name the elements to ease the reading:




First, we start with the whole molecule being in the focus. Since *marked* is empty, we push $\partial^- U$ and $\partial^-(\partial^- U)$ into the stack. After two iterations of the while loop, *marked* is empty and *stack* contains the following molecules:

$$\partial_0^- U = \ A\bullet \ , \partial^- U = \quad \begin{array}{ccccc} A\bullet & & & & C\bullet \xrightarrow{i} \bullet D \\ & {\scriptstyle f}\searrow & & \nearrow{\scriptstyle g} & \\ & & B\bullet & & \end{array} \quad , U = \quad \begin{array}{ccccc} A\bullet & & \xrightarrow{h} & & C\bullet \xrightarrow{i} \bullet D \\ & {\scriptstyle f}\searrow & \Uparrow{\scriptstyle \alpha} & \nearrow{\scriptstyle g} & \\ & & B\bullet & & \end{array}$$

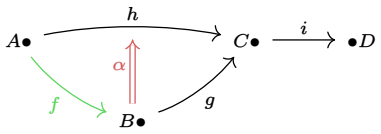
After the third iteration, we get $\partial_0^- U$ into *marked* (**Line 14** of the algorithm) and since $\partial^+(\partial_0^- U) = \emptyset$, *stack*$= \{\partial^- U, U\}$. On the fourth iteration, *focus*$= \partial^- U$, the procedure enters **else** on **Line 18** and $x$ and $y$ become $f$ and $A$, respectively.

$$\partial^- U = \quad \begin{array}{ccccc} A\bullet & & & & C\bullet \xrightarrow{i} \bullet D \\ & {\scriptstyle f}\searrow & & \nearrow{\scriptstyle g} & \\ & & B\bullet & & \end{array}$$

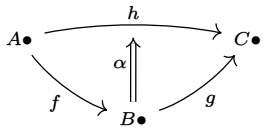
At the end of this iteration, *stack* contains the three sets below:

$$\mathrm{cl}\, f = \ A\bullet \xrightarrow[f]{} B\bullet \ , \partial^- U, U.$$

On the fifth iteration, *focus* $= \mathrm{cl}\, f$; the element $f$ is appended into *marked* and $\partial^+ f$ added to the stack. After the fifth iteration we get *marked* $= \{A\bullet, \xrightarrow{f}\}$ and *stack* $= \{\partial^+ f = B\bullet, \partial^- U, U\}$. In the sixth iteration, $\partial^+ x = \{B\}$ gets into *marked* and from iteration 7 the algorithm adds the elements of $\partial^- U$ in order following the same steps as in iterations 4-6.
Once the whole of $\partial^- U$ was marked (iteration 12), we pop $\partial^- U$ from the stack (iteration 13) and *focus* $= U$ (in iteration 14). Since $U$ is not an atom, the algorithm does not enter the **if** clause on line 13, instead it enters **else** on Line 18. Again, $x$ and $y$ are marked with red and green, respectively (where $x = \alpha$ and $y = f$):



At the end of iteration 14, $marked = \{A\bullet, \xrightarrow{f}, B\bullet, \xrightarrow{g}, C\bullet, \xrightarrow{i}, \bullet D\}$ and *stack* is:

$\mathrm{cl}\,\alpha =$ 

, $U$.

On iteration 15, the 2-dimensional element is appended to *marked* and $\partial^+\alpha$ gets into the stack. In iteration 16, $\partial^+\alpha$ gets into *marked* and on iteration 17, since the whole of $U$ is marked, it is popped from the stack and the procedure ends returning *marked*. The returned list looks as follows: $\{A\bullet, \xrightarrow{f}, B\bullet, \xrightarrow{g}, C\bullet, \xrightarrow{i}, \bullet D \overset{\alpha}{\Rightarrow}, \xrightarrow{h}\}$. By ordering the elements in each dimension according to the list above and using the naming convention described in Chapter 2 we obtain the labelling from Example 2.1.4; the outputted *marked* list becomes: { (0, 0), (1, 0), (0, 1), (1, 1), (0, 2), (1, 2), (0, 3), (2, 0), (1, 3) }.

### 2.4.5 Algorithms for building molecules

Recall that we can represent a closed subset of an oriented graded poset with a graded set of its maximal elements. (These graded sets can be implemented using sorted arrays - where the elements are sorted by their dimension and position.) We provide the algorithms for a molecule $U$, but they also work when $U$ is a closed subset of an oriented graded poset, in which case in the `maximal()` and `boundary()` algorithms, we check if the coface of an element, $x$, is disjoint from the subset of elements of the poset.

```
    procedure MAXIMAL(U : molecule)
2:      maximal ← GrSet()
        for x ∈ U do
4:          if cofaces(x).isEmpty() then
                maximal.add(x)
        return maximal
```

The runtime of `maximal()` is bounded above by the number of elements of $U$, $O(|U|)$ (which is the number of vertices in the Hasse diagram representation).

```
    procedure BOUNDARY(U : molecule, dim, sign)
        maximal ← maximal(U)
3:      boundary_max ← maximal[0, ..., dim]
        for x ∈ U_dim do
            if cofaces(x, (− sign)).isEmpty() then
6:              boundary_max.add(x)
        return clos{boundary_max}
```

For `boundary()` algorithm: Line 3 in the algorithm is again bounded above by $|U|$ (if we do not want to point to the array of maximal elements) or it takes constant time if we want to point to this array. The for loop on line 4 is bounded above by $|U_{\dim}|$, while returning the closure is bounded above by $|U|$ since we need to go through the whole of $U$ and add all its elements in the worst case. So the worst-case running time for `boundary()` is $O(|U|)$.

```
    procedure isRound(U : molecule)
        if U is not pure then return False
3:      bound_in ← boundary(U, (n − 1), −)
        bound_out ← boundary(U, (n − 1), +)
        intersection ← bound_in ∩ bound_out
6:      for (k = dim U − 2, k = −1, k − −) do
            bound_in ← boundary(bound_in, (n − 1), −)
            bound_out ← boundary(bound_out, (n − 1), +)
9:          if intersection ≠ bound_in ∪ bound_out then return False
            intersection ← bound_in ∩ bound_out
        return True
```

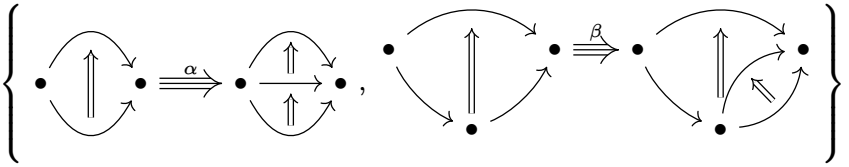
By computing the number of maximal elements of a molecule, $U$, the if statement in Line 2 can be checked in constant time with a sorted array implementation of graded sets. So, Line 2 is bounded above by $O(|U|)$. By linearly ordering the elements of a molecule according to its traversal order, the set theoretic operations can again be computed in linear time in the size of the sets considered, which is bounded above by $O(|U|)$. So the worst-case running time of `isRound()` is $O(\dim(U)\,|U|)$.

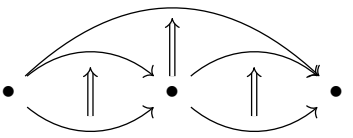
The pushout of inclusions is computed as in the category of sets; that is, given two inclusions, $\imath_1 : U \hookrightarrow V$ and $\imath_2 : U \hookrightarrow W$, the pushout is the disjoint union of $V$ and $W$ quotiented by the equivalence relation $\imath_1(u, 0) \simeq \imath_2(u, 1)$ for all $u \in U$. These set operations are supported by the implementation of our data structures.

# Chapter 3

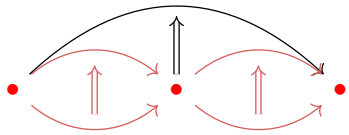
# The higher-dimensional subdiagram matching problem

In the paradigm of higher dimensional rewriting we are working in, a rewrite system on diagrams looks as follows: we have a set $\mathcal{S}$ of $(n+1)$-dimensional cells - $r_i \colon V_i \to \mathbb{V}$ - which are the rewrite rules and a target diagram $\mathcal{D} \colon U \to \mathbb{V}$ we want to rewrite on. Since diagrams are labellings of molecules into a set of variables, verifying whether two labellings match can be trivially done in linear time in the size of the underlying molecule. So we can turn our attention to rewriting on shapes of diagrams - i.e., $\mathcal{S}$ is a list of atoms, $V_i$ and $U$ is our target molecule. To rewrite on $U$, we look, for each rewrite $V_i \in \mathcal{S}$, for matches of $\partial^- V_i$ in $U$. Take the following set of rewrite rules (where we named the top-dimensional elements in each atom to make the text easier to read):

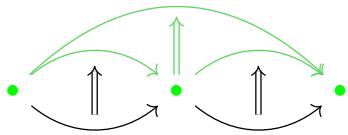
$\mathcal{S} =$ 


and let $U =$  . The first step for rewriting in this setting is searching in the target diagram for matches of the inputs of the rewrites

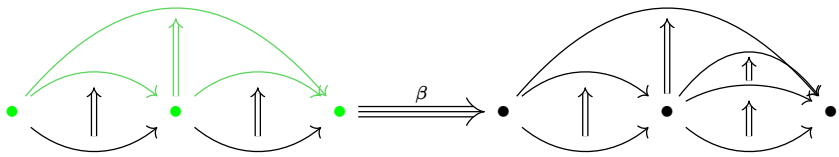
in the list, $\mathcal{S}$. There are two matches for the input of $\alpha$ (highlighted in red):



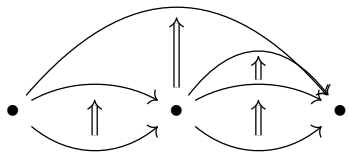
And one match for the input of $\beta$ (highlighted in green):



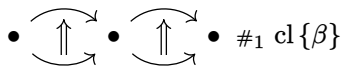
Applying one of the rewrites (say, $\beta$) means rewriting the part of the diagram where the input of $\beta$ was identified as a submolecule:




More specifically, it corresponds to the operation of substituting the input of $\beta$ with the output of $\beta$ in $U$, producing a new diagram that we will denote by $U[\partial^+\beta/\partial^-\beta]$:



Note that unlike we do in pasting, we rewrite only on portions of a diagram. These portions are identified as *submolecules* of the diagram. A rewrite is then represented by *substituting* the input of the atom that was identified with its output. A rewrite on an $n$-dimensional diagram can also be seen as an $(n+1)$-dimensional diagram - the whiskered rewrite rule was pasted on the target diagram (which by Proposition 2.3.13 is isomorphic to the whiskered rewrite rule).
For example, applying $\beta$ to $U$ can be seen as the following whiskering:




Note that for the time being, we do not concern ourselves with the properties of termination and confluence of such a rewrite system which would hopefully be part of future work. The focus is on the algorithmic part of the problem: finding a

subdiagram inclusion algorithm and studying its running time. Ideally, we would like to have a low polynomial time algorithm that solves this problem.

Even though in our example from this introduction it seems quite intuitive to identify the portions we want to rewrite on, this is only the case for diagrams of dimensions lower or equal than two and it does not generalise to arbitrary dimensions. Indeed, from dimension 3 onwards not all inclusions of molecules represent *rewritable* portions of a diagram.

We make a distinction between identifying inclusions of molecules (which is the *molecule matching problem*) and verifying whether such an inclusion represents a rewritable portion of a diagram (which is the *rewritable submolecule problem*). As we will see later on in the next chapter, deciding if $\partial^- V_i$ is a rewritable portion of $U$ is equivalent to $\partial^- V_i$ appearing as a factor in a special kind of pasting decomposition of $U$. These decompositions are called *layerings* and we will present them in this chapter.

We begin by introducing the concepts of submolecules and substitution. We then move on to the theory of layerings, where we present the upper and lower bounds for the dimensions at which a molecule admits a layering. We end the chapter with the *molecule matching algorithm*, an algorithm that identifies all the molecule inclusions between two molecules. This algorithm represents the initial step in solving the subdiagram matching problem. Most of the content of this chapter is part of [29] Section 2 and the beginning of Section 3. For a more detailed exposition on the theory of layerings, check [25].

## 3.1 Submolecules and substitutions

***Definition* 3.1.1** (Submolecule inclusion) — The class of *submolecule inclusions* is the smallest subclass of inclusions of molecules such that

I all isomorphisms are submolecule inclusions,

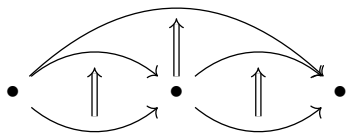
II for all molecules $U, V$ and all $k \in \mathbb{N}$ such that $U \#_k V$ is defined, $U \hookrightarrow (U \#_k V)$ and $V \hookrightarrow (U \#_k V)$ are submolecule inclusions,

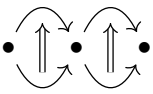
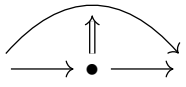
III the composite of two submolecule inclusions is a submolecule inclusion.

A closed subset $V \subseteq U$ is a *submolecule* if its inclusion in $U$ is a submolecule inclusion. In that case we write $V \sqsubseteq U$.

***Definition* 3.1.2** — We also let $\varnothing \sqsubseteq \varnothing$ to take care of some corner cases.

***Example* 3.1.3** — Let $U$ be the following molecule



Then, by definition, $U \sqsubseteq U$. Also, $V =$  is a submolecule of $U$ since $U = V \#_1$  . Moreover, $W = \mathrm{cl}\{(2,0)\} \sqsubseteq U$ since $V = W \#_0$  and $W \hookrightarrow V \hookrightarrow U$.

As one may expect, for any molecule, $U$, and $n \in \mathbb{N}$ its input or output $n$-dimensional boundaries are always submolecules of $U$ and for any element $x \in U$, $\mathrm{cl}\{x\}$ is a submolecule of $U$. The following lemma proves this intuition:

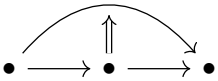
**Lemma 3.1.4.** *Let $U$ be a molecule. Then*

*I for all $n \in \mathbb{N}$ and $\alpha \in \{+,-\}$, $\partial_n^\alpha U \sqsubseteq U$;*

*II for all $x \in U$, $\mathrm{cl}\{x\} \sqsubseteq U$.*

*Proof.* We have by proposition 2.2.25, $\partial_n^\alpha U$ is a molecule. Further, by Proposition 2.3.13, the pastings $U \#_n \partial_n^+ U$ and $\partial_n^- U \#_n U$ are both defined and uniquely isomorphic to $U$. The inclusion of $\partial_n^- U$ into $U$ factors as the inclusion $\partial_n^- U \hookrightarrow (\partial_n^- U \#_n U)$ followed by an isomorphism, and the inclusion of $\partial_n^+ U$ factors as the inclusion $\partial_n^+ U \hookrightarrow (U \#_n \partial_n^+ U)$ followed by an isomorphism.
By Lemma 2.2.30, $\mathrm{cl}\{x\}$ is a molecule. We proceed by induction on the construction of $U$. If $U$ was produced by (*Point*), then $x$ must be the unique element of $U$, so $\mathrm{cl}\{x\} = U$. If $U$ was produced by (*Paste*), it splits into $V \cup W$ with $V$, $W \sqsubseteq U$, and $x \in V$ or $x \in W$. By the inductive hypothesis, $\mathrm{cl}\{x\} \sqsubseteq V$ or $\mathrm{cl}\{x\} \sqsubseteq W$. If $U$ was produced by (*Atom*), it is equal to $(V \cup W) + \{\top\}$ with $V$, $W \sqsubseteq U$ by the first point from this same Lemma, and either $x \in V$ or $x \in W$, in which case the inductive hypothesis applies, or $x = \top$, and $\mathrm{cl}\{x\} = U$. ∎

***Definition* 3.1.5** (Rewritable submolecule) — A submolecule $V \sqsubseteq U$ is *rewritable* if $\dim V = \dim U$ and $V$ is round.

***Example* 3.1.6** — For the molecule in example 3.1.3,  is a rewritable submolecule of $U$, while $V$ and any $\partial_k^\alpha U$, for $k < 2$ are not since $V$ is not round and the dimension of the boundaries is strictly lower than the dimension of $U$.

***Definition* 3.1.7** (Rewritable subdiagram) — Let $t : U \to \mathbb{V}$ be a diagram. A *rewritable subdiagram* of $t$ is the restriction of $t$ to a rewritable submolecule $V \sqsubseteq U$.

***Definition* 3.1.8** — We extend boundary operations to diagrams $t : U \to \mathbb{V}$ by $\partial_n^\alpha t \coloneqq t|_{\partial_n^\alpha U}$ for all $n \in \mathbb{N}$ and $\alpha \in \{+,-\}$.

***Definition* 3.1.9** (Subdiagram matching problem) — The *subdiagram matching problem* is the following search problem: given diagrams $t : U \to \mathbb{V}$ and $s : V \to \mathbb{V}$ such that $\dim U = \dim V$ and $V$ is round, find, if any, the submolecule inclusions $\imath : V \hookrightarrow U$ such that $s = \imath ; t$. This can be split into three subproblems.

I (*Molecule matching problem*). Find, if any, the inclusions $\imath \colon V \hookrightarrow U$.

II (*Rewritable submolecule problem*). Decide if $\imath(V) \sqsubseteq U$.

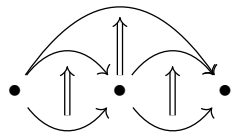
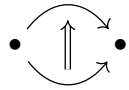
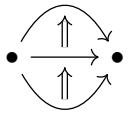
III Decide if $s = \imath; t$.

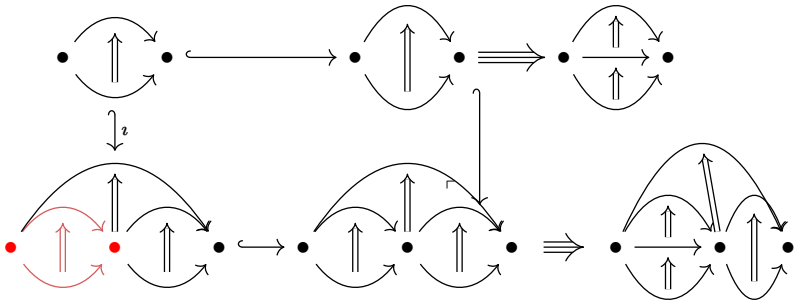
In this chapter, we focus on the molecule matching problem, and in the next on the rewritable submolecule problem. The third problem is trivial.

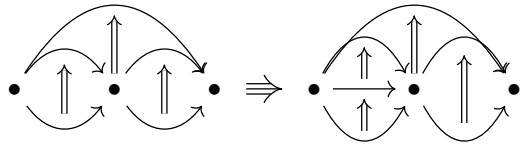
***Definition* 3.1.10** (Substitution) — Let $U, V, W$ be molecules of equal dimension, $\imath \colon V \hookrightarrow U$ an inclusion, and suppose that $V \Rightarrow W$ is defined. Consider the pushout

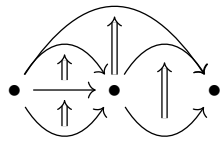
$$\begin{array}{ccc} V & \hookrightarrow & V \Rightarrow W \\ \downarrow{\scriptstyle \imath} & \ulcorner & \downarrow \\ U & \hookrightarrow & U \cup (V \Rightarrow W) \end{array} \tag{3.1}$$

in **ogPos**. The *substitution of $W$ for $\imath \colon V \hookrightarrow U$* is the oriented graded poset $U[W/\imath(V)] \coloneqq \partial^+(U \cup (V \Rightarrow W))$. When $\imath$ is the inclusion of a closed subset we write simply $U[W/V]$.

***Example* 3.1.11** — Let $U =$  , $V =$  and $W =$  . Moreover, let $\imath$ be the inclusion highlighted in red in the following pushout diagram in **ogPos**:



Then the output boundary of  is:



which we write as $U[W/V]$.

So, we can say that in our model performing a rewrite $\alpha$ inside a diagram $\mathcal{D}$ corresponds to the following substitution: $\mathcal{D}[\partial^+\alpha/\partial^-\alpha]$.
Also, note that the result of the pushout in Figure 3.1 is a molecule, which we prove below:

**Lemma 3.1.12.** *Let $V$ be a molecule, $n < \dim V$, $\alpha \in \{+,-\}$. Consider a pushout diagram of the form*

$$\begin{array}{ccc} \partial^\alpha_n V & \hookrightarrow & V \\ \downarrow{\scriptstyle \imath} & & \downarrow{\scriptstyle j_V} \\ U & \xhookrightarrow{j_U} & V \cup U \end{array}$$

*in* **ogPos**. *If $\dim U = n$ and $\imath$ is a submolecule inclusion, then*

*I $V \cup U$ is a molecule,*

*II $j_U$ maps $U$ onto $\partial^\alpha_n(V \cup U)$,*

*III $j_V(V) \sqsubseteq V \cup U$ and $j_V(\partial_n^{-\alpha} V) \sqsubseteq \partial_n^{-\alpha}(V \cup U)$.*

*Proof.* By induction on the construction of $\imath$. If $\imath$ is an isomorphism, then $j_V$ is also an isomorphism, and all the statements are trivially satisfied.
Suppose $U$ is of the form $\partial^\alpha_n V \#_k W$ for some $k \in \mathbb{N}$, and $\imath$ is the inclusion of $\partial^\alpha_n V$ into the pasting. Since $\dim U = n$, necessarily $\dim W \leq n$, so $\partial^\alpha_n W = W$ by Lemma 2.1.25. If $k \geq n$, then also $k \geq \dim W$, and in this case $\imath$ and $j_V$ are again isomorphisms. Suppose that $k < n$. Identifying $V$ with its image through $j_V$, $V \cup U$ splits into $V \cup W$ with

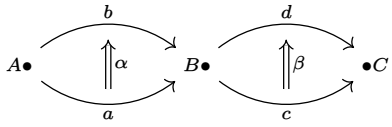
$$V \cap W = \partial^\alpha_n V \cap W = \partial^-_k W = \partial^+_k(\partial^\alpha_n V) = \partial^+_k V$$

where the final equation uses globularity of $V$. This exhibits $V \cup U$ as $V \#_k W$, with $j_V$ the inclusion of $V$ into the pasting, and $j_U$ maps $\partial^\alpha_n V \#_k W$ onto $\partial^\alpha_n(V \#_k W)$ by Lemma 2.2.20. By the same result, $\partial_n^{-\alpha} V \sqsubseteq \partial_n^{-\alpha}(V \#_k W)$. The case where $U$ is of the form $W \#_k \partial^\alpha_n V$ is dual.
By the pasting law for pushout squares, if the statement is true of two submolecule inclusions, it is also true of their composite. ∎

## 3.2 Layerings

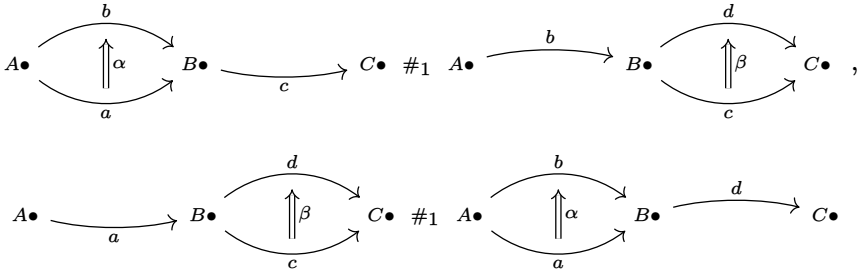
With our current definition of submolecule inclusion, determining whether an inclusion of closed subsets is a submolecule inclusion is cumbersome. Our results on

the subdiagram matching problem rely on pasting decompositions of molecules. A molecule can have several pasting decompositions - i.e., we can "deconstruct" it into molecules such that the pasting of those molecules is isomorphic to the starting molecule. In this section we introduce layerings - a special kind of pasting decomposition - and present the condition for the maximal $k$ such that a molecule admits a $k$-layering. The results of this section were discussed in [29], but for a better exposition, please consult [25, Section 4.2].
Consider the following molecule:

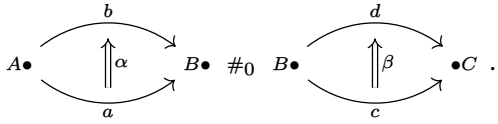



(3.2)

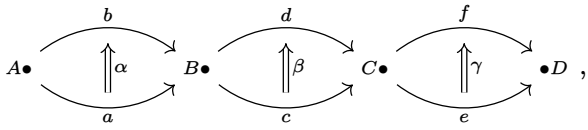
It has three non-trivial pasting decompositions - two at dimension 1:




and one at dimension 0:

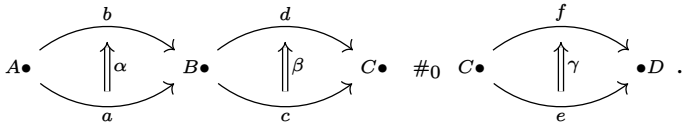



A pasting decomposition does not necessarily need to contain exactly one maximal element in each component. For example, in the molecule below,




(3.3)

the following is a valid pasting decomposition:



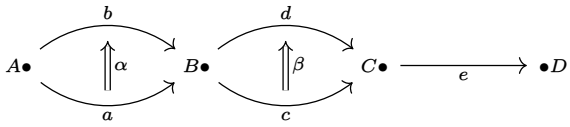

We continue by presenting the theory of layerings which is a way of decomposing a molecule into $k$-layers, where $k$ is less than the dimension of the molecule such that each layer contains at most one maximal element of dimension greater than $k$.

***Definition* 3.2.1 —** Let $U$ be a molecule, $-1 \leq k < \dim U$, and

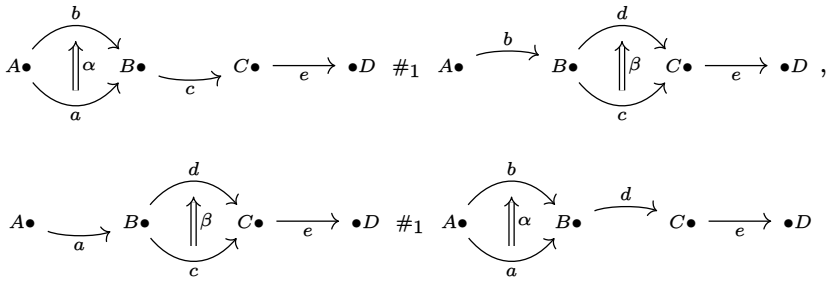
$$m \coloneqq \left| \bigcup_{i>k} (\mathscr{M}ax\, U)_i \right|.$$

A *$k$-layering of $U$* is a sequence $(U^{(i)})_{i=1}^{m}$ of molecules such that $U$ is isomorphic to $U^{(1)} \#_k \dots \#_k U^{(m)}$ and $\dim U^{(i)} > k$ for all $i \in \{1, \dots, m\}$.
For $k = -1$, it is implied that $m = 1$, and $U$ is an atom. We will regularly identify the molecules in a layering of $U$ with their isomorphic images in $U$, which are submolecules.

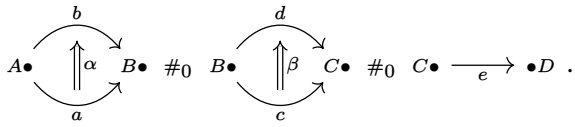
***Example* 3.2.2 —** The molecule from Figure 3.2 has two 1-layerings and one 0-layering. The molecule from Figure 3.3 has six 1-layerings and one 0-layering.
The following molecule




has two 1-layerings:




and one 0-layering:




As expected, the following result holds:

**Lemma 3.2.3.** *Let $U$ be a molecule, $k < \dim U$, and let $(U^{(i)})_{i=1}^{m}$ be a $k$-layering of $U$. Then, for all $i \in \{1, \dots, m\}$, $U^{(i)}$ contains a single maximal element of dimension $> k$.*

*Proof.* See [25, Lemma 4.2.5]. ∎

**Lemma 3.2.4.** *Let $U$ be a molecule, $k \leq \ell < \dim U$. If $U$ admits a $k$-layering, then $U$ admits an $\ell$-layering.*

*Proof.* Let $(U^{(i)})_{i=1}^m$ be a $k$-layering of $U$. For all $i \in \{1, \dots, m\}$, let

$$V^{(i)} \coloneqq \partial_\ell^+ U^{(1)} \#_k \dots \#_k \partial_\ell^+ U^{(i-1)} \#_k U^{(i)} \#_k \partial_\ell^- U^{(i+1)} \#_k \dots \#_k \partial_\ell^- U^{(m)}.$$

By repeated applications of point 3 from Proposition 2.3.13 followed by point 2 from Proposition 2.3.13, $U$ is isomorphic to

$$V^{(1)} \#_\ell \dots \#_\ell V^{(m)}.$$

Restricting to the subsequence of $(V^{(i)})_{i=1}^m$ on those $i \in \{1, \dots, m\}$ such that $\dim V^{(i)} > \ell$, which does not change the result by point 2 from Proposition 2.3.13, we obtain an $\ell$-layering of $U$. ∎

***Example* 3.2.5** — Consider the molecule from Example 3.2.2 together with its 0-layering $(U^{(i)})_{i=1}^3$ and let

$$V^{(1)} = A\bullet \overset{b}{\underset{a}{\Uparrow \alpha}} B\bullet \;\#_0\; B\bullet \xrightarrow[c]{} \bullet C \;\#_0\; C\bullet \xrightarrow[e]{} \bullet D$$

$$V^{(2)} = A\bullet \xrightarrow[b]{} \bullet B \;\#_0\; B\bullet \overset{d}{\underset{c}{\Uparrow \beta}} C\bullet \;\#_0\; C\bullet \xrightarrow[e]{} \bullet D$$

$$V^{(3)} = A\bullet \xrightarrow[b]{} \bullet B \;\#_0\; B\bullet \xrightarrow[d]{} \bullet C \;\#_0\; C\bullet \xrightarrow[e]{} \bullet D$$

Then we obtain a 1-layering:

$$V^{(1)} \#_1 V^{(2)} = A\bullet \overset{b}{\underset{a}{\Uparrow \alpha}} B\bullet \xrightarrow{c} C\bullet \xrightarrow{e} \bullet D \;\#_1\; A\bullet \xrightarrow{b} B\bullet \overset{d}{\underset{c}{\Uparrow \beta}} C\bullet \xrightarrow{e} \bullet D$$

where $V^{(3)} = \partial^+ V^{(2)}$, and by unitality of pasting (Proposition 2.3.13) we get that $V^{(2)} \#_1 V^{(3)}$ is uniquely isomorphic to $V^{(2)}$.

***Definition* 3.2.6** (Layering dimension) — Let $U$ be a molecule. The *layering dimension* of $U$ is the integer

$$\operatorname{lydim} U \coloneqq \min \left\{ k \geq -1 \mid \left| \bigcup_{i > k+1} (\mathscr{M}ax\, U)_i \right| \leq 1 \right\}.$$

**Lemma 3.2.7.** *Let $U$ be a molecule, $n \coloneqq \dim U$. Then*

*I* $\operatorname{lydim} U \leq n - 1$,

*II* $\operatorname{lydim} U = n - 1$ *if and only if* $|U_n| > 1$.

*Proof.* We have

$$\left|\bigcup_{i>n}(\mathscr{M}ax\,U)_i\right| = |\varnothing| = 0,$$

so $\operatorname{lydim} U \leq n-1$, with equality if and only if

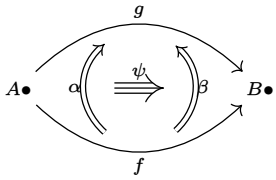
$$\left|\bigcup_{i>n-1}(\mathscr{M}ax\,U)_i\right| = |(\mathscr{M}ax\,U)_n| = |U_n| > 1,$$

where we used Lemma 2.1.27. ∎

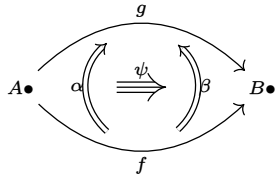
The following results justify why we can prove results for molecules using induction on the layering dimension. For this we use two results about lower dimensional molecules that will be proved in Section 4.4. The proofs of these results are self-contained and do not rely on any content from the current chapter or subsequent ones.

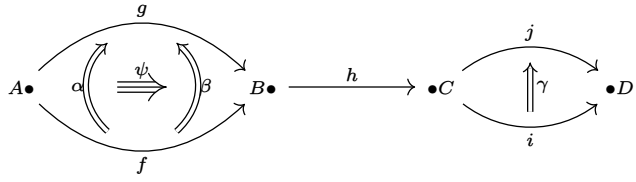
**Lemma 3.2.8.** *Let $U$ be a molecule. Then* $\operatorname{lydim} U = -1$ *if and only if $U$ is an atom.*

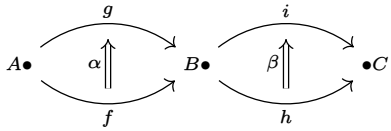
*Proof.* Suppose $\operatorname{lydim} U = -1$. Either $\left|\bigcup_{i>0}(\mathscr{M}ax\,U)_i\right| = 0$, so $\dim U = 0$ and we conclude by Lemma 4.4.1, or $1 = \left|\bigcup_{i>0}(\mathscr{M}ax\,U)_i\right| = |\mathscr{M}ax\,U|$ by Lemma 4.4.2. In either case, $U$ has a greatest element. Conversely, if $U$ has a greatest element, $\left|\bigcup_{i>0}(\mathscr{M}ax\,U)_i\right| \leq |\mathscr{M}ax\,U| = 1$. ∎

***Example* 3.2.9 —** The layering dimension of 

is $-1$ since it contains only one maximal element, $\psi$.

The layering dimension of 

is 0 while the layering dimension of 

is 1 and the layering dimension of 

is also 1.

**Lemma 3.2.10.** *Let $U$ be a molecule, $k \coloneqq \operatorname{lydim} U$. Suppose $k \geq 0$, and let $(U^{(i)})_{i=1}^m$ be a $k$-layering of $U$. Then*

*I* $m > 1$,

*II for each* $i \in \{1, \dots, m\}$, $\operatorname{lydim} U^{(i)} < k$,

*III at most one of the* $U^{(i)}$ *contains an element of dimension* $> k+1$.

*Proof.* By definition of $\operatorname{lydim} U$, if $k \geq 0$ and a $k$-layering exists, then $m > 1$, for otherwise $k - 1 \leq \operatorname{lydim} U$, a contradiction. Moreover, $U$ contains at most one element of dimension $> k+1$, which can be contained at most in one of the $U^{(i)}$. Finally, by Lemma 3.2.3, we have $\left|\bigcup_{j>k}(\mathit{Max}\, U^{(i)})_j\right| = 1$, so $\operatorname{lydim} U^{(i)} \leq k - 1 < k$. ∎

**Theorem 3.2.11.** *Let* $U$ *be a molecule,* $\operatorname{lydim} U \leq k < \dim U$. *Then* $U$ *admits a* $k$*-layering.*

*Proof.* Let $k \coloneqq \operatorname{lydim} U$. If $k = -1$, then $U$ is an atom and admits the trivial layering $U = U^{(1)}$. If $k \geq 0$, by Lemma 3.2.8 $U$ is not an atom, so we can assume that $U$ was produced by (*Paste*). Then $U$ is equal to $V \#_\ell W$ for some molecules $V$, $W$ and $\ell < \min\{\dim V, \dim W\}$. By the inductive hypothesis, we have layerings

$$V^{(1)} \#_{k_V} \cdots \#_{k_V} V^{(m_V)}, \qquad W^{(1)} \#_{k_W} \cdots \#_{k_W} W^{(m_W)}$$

of $V$ and $W$, respectively, for $k_V \coloneqq \operatorname{lydim} V$ and $k_W \coloneqq \operatorname{lydim} W$. Furthermore, by [25, Lemma 4.2.10], we know that $k \geq \max\{k_V, k_W, \ell\}$. Let

$$n_V \coloneqq \begin{cases} m_V & \text{if } k_V = k, \\ 1 & \text{if } k_V < k \text{ and } \dim V > k, \\ 0 & \text{if } k_V < \dim V < k, \end{cases}$$

$$n_W \coloneqq \begin{cases} m_W & \text{if } k_W = k, \\ 1 & \text{if } k_W < k \text{ and } \dim W > k, \\ 0 & \text{if } k_W < \dim W < k. \end{cases}$$

Notice that it can never be the case that $n_V = n_W = 0$. We claim that we can decompose $V$ as

$$\tilde{V}^{(1)} \#_k \cdots \#_k \tilde{V}^{(n_V)} \#_k \underbrace{\partial_k^+ V \#_k \cdots \#_k \partial_k^+ V}_{n_W \text{ times}}, \tag{3.4}$$

where each $\tilde{V}^{(i)}$ is a molecule containing exactly one maximal element of dimension $> k$. If $k_V = k$, we let $\tilde{V}^{(i)} \coloneqq V^{(i)}$ for all $i \in \{1, \dots, m_V\}$. If $k_V < k$, then $V$ contains at most one maximal element of dimension $> k_V + 1$, hence at most one maximal element of dimension $> k$. If $\dim V > k$, it contains exactly one, and we let $\tilde{V}^{(1)} \coloneqq V$. If $\dim V < k$, then $V = \partial_k^+ V$. By Proposition 2.3.13, pasting copies of $\partial_k^+ V$ does not change the result up to unique isomorphism. Similarly, we can decompose $W$ as

$$\underbrace{\partial_k^- W \#_k \cdots \#_k \partial_k^- W}_{n_V \text{ times}} \#_k \tilde{W}^{(1)} \#_k \cdots \#_k \tilde{W}^{(n_W)} \tag{3.5}$$

where each $\tilde{W}^{(i)}$ contains exactly one maximal element of dimension $> k$.
If $\ell = k$, since $\ell < \min\{\dim V, \dim W\}$, we have $0 < \min\{n_V, n_W\}$. Then

$$\tilde{V}^{(1)} \#_k \ldots \#_k \tilde{V}^{(n_V)} \#_k \tilde{W}^{(1)} \#_k \ldots \#_k \tilde{W}^{(n_W)}$$

is a $k$-layering of $U$. If $\ell < k$, let

$$U^{(i)} \coloneqq \begin{cases} \tilde{V}^{(i)} \#_\ell \partial_k^- W & \text{if } i \leq n_V, \\ \partial_k^+ V \#_\ell \tilde{W}^{(i-n_V)} & \text{if } n_V < i \leq n_V + n_W. \end{cases}$$

Since $\dim \partial_k^- V = \dim \partial_k^+ W = k$, each $U^{(i)}$ still contains exactly one maximal element of dimension $> k$. Plugging (3.4) and (3.5) in $V \#_\ell W$ and using the third point from Proposition 2.3.13 repeatedly, we deduce that $V \#_\ell W$ is isomorphic to

$$U^{(1)} \#_k \ldots \#_k U^{(n_V+n_W)},$$

which has the desired properties. Necessarily, $n_V + n_W = m$.
For $\operatorname{lydim} U < k < \dim U$, the statement follows from Lemma 3.2.4. ∎

Using Lemma 3.2.10, Lemma 3.2.8 and Theorem 3.2.11, we can prove that a property holds for all molecules $U$ using induction on the layering dimension. To do this it suffices to:

- prove that it holds when $\operatorname{lydim} U = -1$, that is, when $U$ is an atom,
- prove that it holds when $k \coloneqq \operatorname{lydim} U \geq 0$, assuming that it holds of all the $(U^{(i)})_{i=1}^m$ in a $k$-layering of $U$.

We prove the following lemma that will be used in the next section using this technique:

**Lemma 3.2.12.** *Let $U$ be a molecule, $k \in \mathbb{N}$, and suppose*

$$\bigcup_{i>k} (\mathscr{M}ax\, U)_i = \{x\}.$$

*Then, for all $\alpha \in \{+,-\}$,*

*I $\partial_k^\alpha x \sqsubseteq \partial_k^\alpha U$,*

*II $\partial_k^\alpha U$ is isomorphic to $\partial_k^{-\alpha} U[\partial_k^\alpha x / \partial_k^{-\alpha} x]$.*

*Proof.* We proceed by induction on $\operatorname{lydim} U$. If $\operatorname{lydim} U = -1$, then $U$ is an atom and equal to $\operatorname{cl}\{x\}$. It follows that $\partial_k^\alpha x = \partial_k^\alpha U$, which is trivially a submolecule, and is isomorphic to $\partial_k^{-\alpha} U[\partial_k^\alpha x / \partial_k^{-\alpha} x]$ by [25, Lemma 4.1.12].
Suppose $\ell \coloneqq \operatorname{lydim} U \geq 0$, and let $(U^{(i)})_{i=1}^m$ be an $\ell$-layering of $U$. Then $\ell \leq k-1 < k$ because $\left|\bigcup_{i>k} (\mathscr{M}ax\, U)_i\right| = 1$. By Lemma 2.2.20, $\partial_k^\alpha U$ is isomorphic to

$$\partial_k^\alpha U^{(1)} \#_\ell \ldots \#_\ell \partial_k^\alpha U^{(m)}.$$

Now $x$ is contained in a single $U^{(i)}$. By the inductive hypothesis, $\partial_k^\alpha x \sqsubseteq \partial_k^\alpha U^{(i)}$, and the latter is isomorphic to $\partial_k^{-\alpha} U^{(i)}[\partial_k^\alpha x / \partial_k^{-\alpha} x]$. We conclude by [25, Lemma 4.1.16]. ∎

## 3.3 Frame-dimension

The frame-dimension of a molecule is the number that gives the maximal dimension at which two maximal elements intersect (are glued together). For example, a round molecule always has frame dimension $n-1$ (as we will prove in Proposition 3.3.10). We show that the frame-dimension provides a lower bound for which a molecule, $U$, accepts a layering. While it is true that all molecules of dimension less than or equal to 3 accept a layering in their frame dimension (see Lemma 4.4.15 and Corollary 4.3.6), this does not hold in general. The first example of such behaviour is 4-dimensional (see Example 3.3.8).

***Definition* 3.3.1** (Frame dimension) — Let $U$ be a molecule. The *frame dimension of* $U$ is the integer

$$\operatorname{frdim} U \coloneqq \dim \bigcup \{\operatorname{cl}\{x\} \cap \operatorname{cl}\{y\} \mid x, y \in \mathscr{M}ax\, U, x \neq y\}.$$

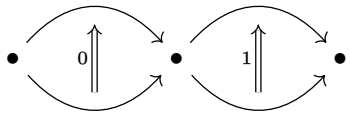
**Lemma 3.3.2.** *Let $U$ be a molecule. Then $\operatorname{frdim} U = -1$ if and only if $U$ is an atom.*

*Proof.* If $U$ is an atom, then there does not exist a pair of distinct elements of $\mathscr{M}ax\, U$, so $\operatorname{frdim} U = \dim \varnothing = -1$. Otherwise, suppose that $\operatorname{frdim} U = -1$, and let $x \in \mathscr{M}ax\, U$. Then, letting $V \coloneqq \operatorname{cl}((\mathscr{M}ax\, U) \setminus \{x\})$, we have $U = \operatorname{cl}\{x\} \cup V$ and $\operatorname{cl}\{x\} \cap V = \varnothing$. By [25, Lemma 3.3.13], $V = \varnothing$, so $x$ is the greatest element of $U$. ∎

**Lemma 3.3.3.** *Let $U$ be a molecule. Then $\operatorname{frdim} U \leq \operatorname{lydim} U$.*

*Proof.* Let $r \coloneqq \operatorname{frdim} U$. If $r = -1$, by Lemma 3.3.2 $U$ is an atom, and by Lemma 3.2.8 $\operatorname{lydim} U$ is also $-1$. Suppose $r \geq 0$. Then there exist distinct maximal elements $x, y \in U$ such that $\dim(\operatorname{cl}\{x\} \cap \operatorname{cl}\{y\}) = r$. Necessarily $r < \min\{\dim x, \dim y\}$, so $x, y \in \bigcup_{i>r}(\mathscr{M}ax\, U)_i$ and $\left|\bigcup_{i>r}(\mathscr{M}ax\, U)_i\right| \geq 2$. It follows that $r - 1 < \operatorname{lydim} U$, that is, $r \leq \operatorname{lydim} U$. ∎

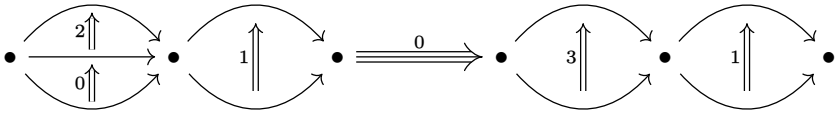
***Example* 3.3.4** — Let $U$ be the following molecule:



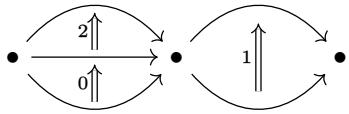

Then $U$ has $\operatorname{frdim} U = 0$ and $\operatorname{lydim} U = 1$.

Unlike layering dimension, frame dimension is not stable under submolecule inclusions. That means that we will not prove properties about molecules using induction on the frame dimension of the molecule.

***Example* 3.3.5** — Let $U$ be the following molecule, where we omit the dimension of the elements from the naming convention:



The maximal elements are $(3,0)$ and $(2,1)$ and $\operatorname{frdim} U = 0$. The input boundary, $\partial^- U$ is:




The maximal elements in $\partial^- U$ are $(2,0)$, $(2,1)$ and $(2,2)$ and $\operatorname{frdim} \partial^- U = 1$ since $\operatorname{frdim} \{(2,0),(2,2)\} = 1$.

**Lemma 3.3.6.** *Let $U$ be a molecule, $k \geq -1$. If $U$ admits a $k$-layering, then $k \geq \operatorname{frdim} U$.*

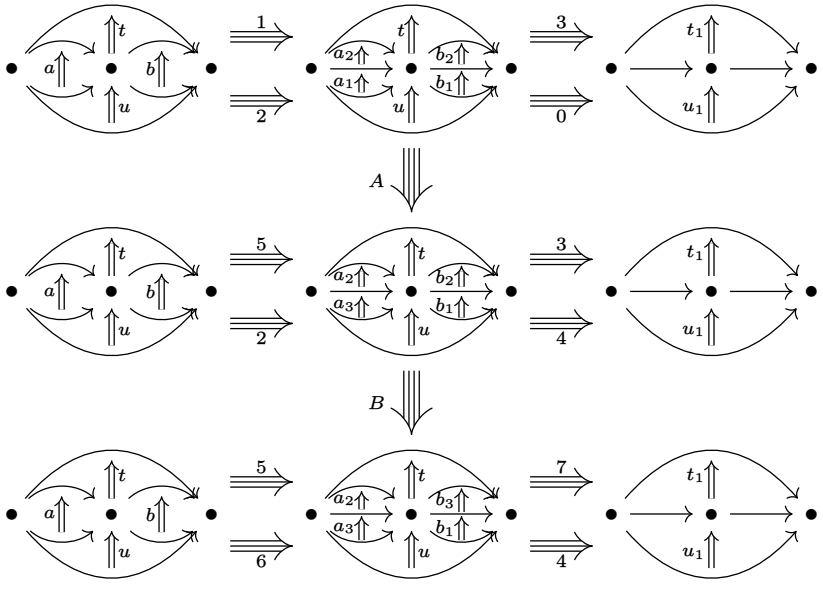
*Proof.* See [25, Lemma 4.4.5]. ∎

**Corollary 3.3.7.** *Let $U$ be a molecule. Then*

$$\operatorname{frdim} U \leq \min \{k \geq -1 \mid U \text{ admits a } k\text{-layering}\} \leq \operatorname{lydim} U.$$

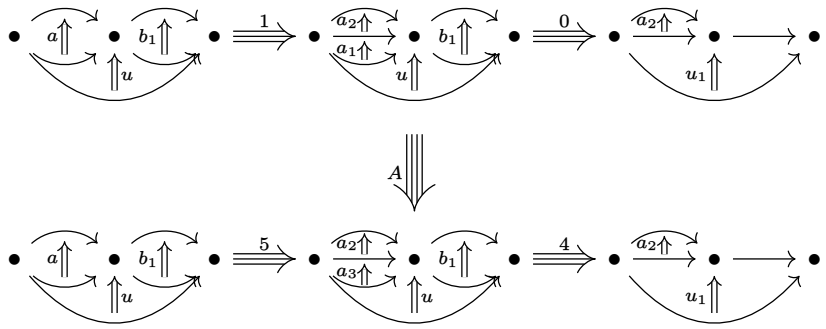
*Proof.* Follows from Lemma 3.3.6 and Theorem 3.2.11. ∎

***Example* 3.3.8 —** Below we show an example of a molecule with frame dimension 2 that does not admit a 2-layering. Let $U$ be the following 4-dimensional molecule:

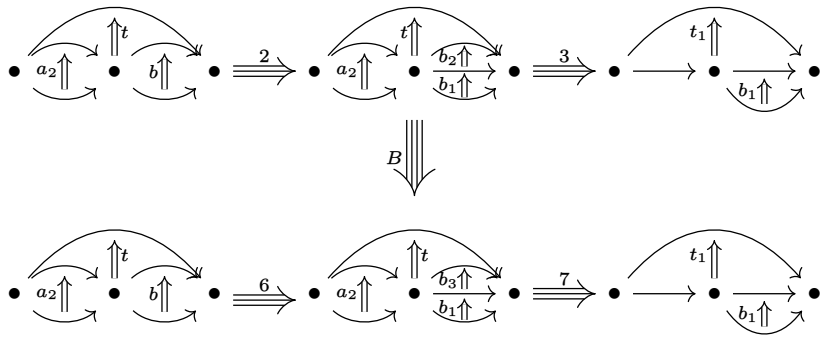



The 4-dimensional element, $A$, rewrites the composable configuration of volumes $\operatorname{cl}\{1,0\}$ to $\operatorname{cl}\{5,4\}$, while $B$ rewrites $\operatorname{cl}\{2,3\}$ to $\operatorname{cl}\{6,7\}$. Note that $\operatorname{frdim}\{A,B\} = 2$ since $\operatorname{cl}\{A\}$ and $\operatorname{cl}\{B\}$ do not share any 3-dimensional elements.

Then, $\mathrm{cl}\{A\}$ is:




and $\mathrm{cl}\{B\}$ is:




If we want to write $U$ as a layering $U_A \#_2 U_B$, where $U_A$ is a layer containing $A$ and $U_B$ is a layer containing $B$, then we would need to whisker $\partial_2^- A$ with $(\partial^+ a_2 \#_0 b) \#_1 t$ which is impossible since $\partial^- b \neq \partial^+ b_1$.
Similarly, if we want to write $U$ as a layering $U_B \#_2 U_A$, then we would need to whisker $\partial_2^- B$ with $(a \#_0 \partial^- b_1) \#_1 u$ which is impossible since $\partial^+ a \neq \partial^- a_2$.
We will understand this example better once we make the connection between layerings and orderings more clear.

Finally, we obtain the following result, for any molecule, $U$.

**Corollary 3.3.9.** *Let $U$ be a molecule, $n \coloneqq \dim U$. Then $U$ admits an $(n-1)$-layering.*

*Proof.* Follows from Theorem 3.2.11 together with Lemma 3.2.7. ∎

Finally, we can prove:

**Proposition 3.3.10.** *Let $U$ be a molecule, $n \coloneqq \dim U$. If $U$ is round, then $\mathscr{F}_{n-1}U$ is connected.*

*Proof.* First of all, if $U$ is round, then it is pure, so the vertices of $\mathscr{F}_{n-1}U$ are the elements of $U_n$. If $U$ is an atom, then $\mathscr{F}_{n-1}U$ consists of a single vertex and no

edges, so it is trivially connected. In particular this is true when $n = 0$ by Lemma 4.4.1, so we can proceed by induction on $n$.

Suppose $n > 0$ and $|U_n| > 1$, which by Lemma 3.2.7 implies $\operatorname{lydim} U = n - 1$. Assume by way of contradiction that $\mathscr{F}_{n-1}U$ is not connected. Then there is a bipartition $U_n = A + B$ such that there are no edges in $\mathscr{F}_{n-1}U$ between vertices in $A$ and vertices in $B$. By Corollary 2.3.9, no element of codimension 1 in $U$ can be covered by two elements with the same orientation, so this implies that $\dim(\operatorname{cl} A \cap \operatorname{cl} B) < n - 1$. Let

$$A' \coloneqq \left\{x \in \Delta^- U \mid \nabla^- x \subseteq A\right\}, \qquad B' \coloneqq \left\{x \in \Delta^- U \mid \nabla^- x \subseteq B\right\}.$$

Then $A' + B'$ is a bipartition of $\Delta^- U$. By Lemma 2.2.26, $\partial^- U$ is round, so by the inductive hypothesis $\mathscr{F}_{n-2}(\partial^- U)$ is connected. It follows that there exist $\alpha \in \{+, -\}$, $x \in A'$, $y \in B'$, and $z \in U_{n-2}$ such that $z \in \Delta^\alpha x \cap \Delta^{-\alpha} y$. Then $z$ has two distinct cofaces in $\partial^- U$, so by Corollary 2.3.9 $z \notin \partial(\partial^- U) = \partial_{n-2} U$. We claim that $z \in \partial^+ U$, contradicting the roundness of $U$.

By Theorem 3.2.11, there exists an $(n-1)$-layering $(U^{(i)})_{i=1}^m$ of $U$; we will identify the $U^{(i)}$ with their isomorphic images in $U$. Let $V_0 \coloneqq \partial^- U$ and $V_i \coloneqq \partial^+ U^{(i)}$ for each $i \in \{1, \dots, m\}$. We will prove that, for all $i \in \{0, \dots, m\}$,

I $z \in V_i$,

II there exist $x_i \in \operatorname{cl} A$ and $y_i \in \operatorname{cl} B$ such that $\nabla^\alpha z \cap V_i = \{x_i\}$ and $\nabla^{-\alpha} z \cap V_i = \{y_i\}$.

For $i = 0$, we have already established this with $x_0 \coloneqq x$, $y_0 \coloneqq y$. Let $i \geq 0$, and assume this holds for $i - 1$. By Lemma 3.2.3, there is a single $n$-dimensional element $x^{(i)}$ in $U^{(i)}$, and by Lemma 3.2.12

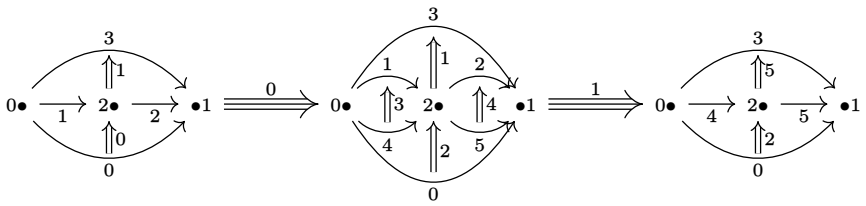
$$V_i = \partial^- U^{(i)}[\partial^+ x^{(i)} / \partial^- x^{(i)}] = V_{i-1}[\partial^+ x^{(i)} / \partial^- x^{(i)}].$$

Suppose $x^{(i)} \in A$. Then $y_{i-1} \notin \operatorname{cl}\left\{x^{(i)}\right\}$, so $y_{i-1} \in V_i$, and we let $y_i \coloneqq y_{i-1}$. If $x_{i-1} \notin \operatorname{cl}\left\{x^{(i)}\right\}$ then also $x_{i-1} \in V_i$, and we let $x_i \coloneqq x_{i-1}$. Otherwise, $x_{i-1}$ is the only coface of $z$ in $\partial^- x^{(i)}$, so by Corollary 2.3.9 we have $z \in \partial^\alpha(\partial^- x^{(i)}) = \partial^\alpha(\partial^+ x^{(i)})$. It follows that $z \in V_i$ and there exists a unique $x_i$ such that $\nabla^\alpha z \cap \partial^+ x^{(i)} = \{x_i\}$. The case $x^{(i)} \in B$ is analogous.

Since $V_m = \partial^+ U$, we have proved that $z \in \partial^+ U$, a contradiction. ∎

The converse of this proposition is not true, not even when the molecule is pure.

***Example* 3.3.11 —** [25, Example 4.5.8] Let $U$ be the following molecule presented in its pasting diagram representation in which the elements are labelled according to their traversal order, but we omitted the dimension from our usual notation:




Element $(0, 2)$ is in $\partial_2^- U \cap \partial_2^+ U$, but $(0, 2) \notin \partial_1 U$. Geometrically, one can note that the intersection of $(3, 0)$ and $(3, 1)$ is not round.

## 3.4 Molecule matching algorithm

We are now ready to present the solution to the first step of the subdiagram matching problem. The algorithm takes two $n$-dimensional molecules, $U$ and $V$, where $V$ is round and returns a list of all the inclusions $V \hookrightarrow U$. The algorithm first constructs the $(n-1)$-dimensional flow graphs for both $U$ and $V$. It then chooses an $n$-element in $V$ (according to some order, which can be the traversal order of the $n$-dimensional elements of $V$) and using the traversal algorithm tries to find a match (i.e., an isomorphism) of this element with an n-dimensional element in $U$. Once such a match is found, there is exactly one way to extend this match to all top-dimensional elements in $V$. By Proposition 3.3.10 we can find an $n$-dimensional element that is connected to an already matched element. That is, the two elements cover an $(n-1)$-dimensional element, $z$, that was already matched at a previous round. By Corollary 2.3.9, $z$ has exactly two cofaces in $U$, out of which one was already matched. Then, it must be that the next match is the other coface of $z$.

We provide both the pseudo code for the algorithm as well as a description of its behaviour. We then continue with its runtime analysis and illustrate how it runs on an example.

***Procedure* 3.4.1** (Molecule matching algorithm) — The procedure takes as input two molecules $U, V$ such that $\dim U = \dim V$ and $V$ is round, and it returns all inclusions $V \hookrightarrow U$.

Let $n \coloneqq \dim U$. To begin, we pick an arbitrary ordering $(x^{(i)})_{i=1}^{m}$, for example the traversal order, of the elements of $U_n$. Moreover, we pick an ordering $(y^{(j)})_{j=1}^{p}$ of the elements of $V_n$ with the property that, for all $k \in \{1, \dots, p\}$, the induced subgraph of $\mathscr{F}_{n-1}V$ on $(y^{(j)})_{j=1}^{k}$ is connected. This is possible because $\mathscr{F}_{n-1}V$ is connected by Proposition 3.3.10. For each $k \in \{1, \dots, p\}$, we let $V^{(k)} \coloneqq \bigcup_{j \le k} \operatorname{cl}\{y^{(j)}\}$. We have $V^{(i)} \subseteq V^{(j)}$ whenever $i \le j$, and $V^{(p)} = V$ since $V$ is pure by Lemma 2.2.11. For each $i \in \{1, \dots, m\}$, we try to construct a sequence of inclusions $(\imath^{(i,j)} \colon V^{(j)} \hookrightarrow U)_{j=1}^{p}$ such that the restriction of $\imath^{(i,j')}$ to $V^{(j)}$ is equal to $\imath^{(i,j)}$ when $j \le j'$, iterating on $k \in \{1, \dots, p\}$. When $k = 1$, if $V^{(1)} = \operatorname{cl}\{y^{(1)}\}$ is isomorphic to $\operatorname{cl}\{x^{(i)}\}$, we let $\imath^{(i,1)}$ be the unique isomorphism $V^{(1)} \xhookrightarrow{\sim} \operatorname{cl}\{x^{(i)}\}$ followed by the inclusion $\operatorname{cl}\{x^{(i)}\} \subseteq U$, and iterate on $k$. Else, we iterate on $i$.

When $k > 1$, we let $j$ be the least value such that there exists an edge between $y^{(j)}$ and $y^{(k)}$ in $\mathscr{F}_{n-1}V$. Then $j < k$ because of our connectedness assumption, and there exists $z \in \Delta^{\alpha} y^{(j)} \cap \Delta^{-\alpha} y^{(k)}$ for some $\alpha \in \{+, -\}$. We pick the least such $z$ with respect to some ordering of $V_{n-1}$, for example the traversal order. Since $\imath$ is a morphism of oriented graded posets, $\imath^{(i,k-1)}(y^{(j)})$ is one coface of $\imath^{(i,k-1)}(z)$ in $U$. If $\imath^{(i,k-1)}(z)$ has no other cofaces, then we iterate on $i$. Else, by Corollary 2.3.9, $\imath^{(i,k-1)}(z)$ has exactly one other coface, call it $x$; note that $x$ cannot be in the image of $\imath^{(i,k-1)}$, since $y^{(j)}$ and $y^{(k)}$ are the only cofaces of $z$ in $V$. If $\operatorname{cl}\{y^{(k)}\}$ is isomorphic to $\operatorname{cl}\{x\}$, and the unique isomorphism $\operatorname{cl}\{y^{(k)}\} \xhookrightarrow{\sim} \operatorname{cl}\{x\}$ followed by the inclusion $\operatorname{cl}\{x\} \subseteq U$ matches $\imath^{(i,k-1)}$ on $\operatorname{cl}\{y^{(k)}\} \cap V^{(k-1)}$, then we let $\imath^{(i,k)}$ be the unique extension of $\imath^{(i,k-1)}$ that restricts to $\operatorname{cl}\{y^{(k)}\} \xhookrightarrow{\sim} \operatorname{cl}\{x\} \subseteq U$. Else, we iterate on $i$.

If we succeed to construct $\imath^{(i,p)}$, we add it to the list of inclusions $V \hookrightarrow U$, then iterate on $i$.

The pseudocode is described in a simplified version below:

```
procedure MoleculeMatching(U, V : molecule, V is round)
    n ← dim U (= dim V)
    ı = []
    Pick (x^(i))_{i=1}^m the traversal order of U_n.
    Pick (y^(i))_{i=1}^p an ordering of V_n with the property: for all
k ∈ {1,...,p}, the induced subgraph of 𝒻_{n-1}V on (y^(i))_{i=1}^k is connected.
                                  ▷ notation: V^(k) := ⋃_{j≤k} cl{y^(j)}
    for i ∈ {1,...,m} do
        if V^(1) ↪̃ cl{x^(i)} then
            ı^(i,1) ← V^(1) ↪̃ cl{x^(i)}
            for k ← {2,...,p} do
                j ← least value such that there exists an edge between
                      y^(k) and y^(j) in 𝒻_{n-1}V.
                z ← first element in Δ^α y^(j) ∩ Δ^{-α} y^(k) according to the
                      traversal order of V_{n-1}.
                x' ← ı^(i,k-1)(y^(j)) where x' ∈ Δ^α ı^(i,k-1)(z)
                if |∇^{-α} ı^(i,k-1)(z)| = 0 then
                    break;
                x ← ∇^{-α} ı^(i,k-1)(z)[0]                  ▷ By Corollary 2.3.9.
                if cl{y^(k)} ↪̃ cl{x} then
                    iso ← cl{y^(k)} ↪̃ cl{x}
                    if iso matches ı^(i,k-1) on cl{y^(k)} ∩ V^(k) then
                        ı^(i,k) ← unique extension of ı^(i,k-1) that restricts to
                            cl{y^(k)} ↪̃ cl{x}
                    else break;
                else break
            ı.add(ı^(i,p))
    return ı
```

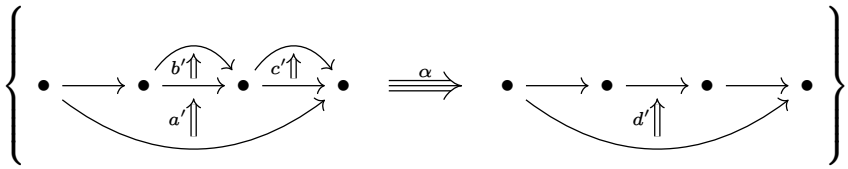
**Theorem 3.4.2.** *The molecule matching problem in dimension $n$ can be solved in time*

$$O(|U_n|\ |V_n|\ |V|\ |\mathscr{E}_\vee V|\ \log|\mathscr{E}_\vee V|).$$

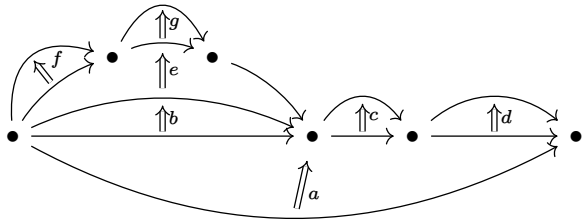
*Proof.* We suppose $n > 0$ since the case $n = 0$ is trivial. First of all, with our choice of data structures both $U_n$ and $V_{n-1}$ already come with a linear order when one is needed. Moreover, we can both construct $\mathscr{F}_{n-1}V$ and order its vertices in the desired way by traversing the "slice" of $\mathscr{H}V$ on the elements of dimension $n$ and $(n-1)$. Since $\max\{|V_n|, |V_{n-1}|\} \leq |\mathscr{E}_n V|$, this can be done in time $O(|\mathscr{E}_n V|)$ with a standard traversal algorithm.
In the main part of the algorithm, we have exactly $|U_n|$ iterations. At each iteration, we need to solve at most $|V_n|$ isomorphism problems for submolecules of $V$. The time complexity of each can be bounded above by the time complexity of the isomorphism problem for $V$, which is $O(|V|\ |\mathscr{E}_\vee V|\ \log|\mathscr{E}_\vee V|)$ by Theorem

2.4.21. It is straightforward to verify that all other operations, such as checking that the isomorphisms match on intersections or finding the next match, have lower complexity. Since $|\mathscr{E}_n V| \leq |\mathscr{E}_\vee V|$, we can ignore the $O(|\mathscr{E}_n V|)$ summand, and conclude. ■

***Example* 3.4.3 —** Let's show how the algorithm above works on an example. Let $\mathcal{S}$ be the list containing one 3-dimensional atom:

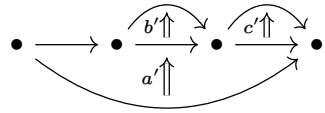



and let $U$ be the following diagram:



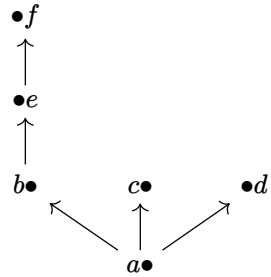

We let

$$V \coloneqq \partial^- \alpha =$$

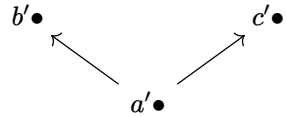



The $(n-1)$-flow graphs of $U$ and $V$ are:

$\mathscr{F}_{n-1}U =$



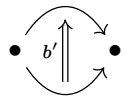

$\mathscr{F}_{n-1}V =$




The algorithm takes the "first" element in $V$ and tries to match it with each element in $U$. Once a match is found, we try to extend it by picking the "next" element in $V$. Since the shapes of diagram are rigid, there is at most one choice for a match. If this fails, we return to the main loop and try to find another match for the "first" element in $V$.

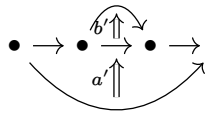
We start with an empty list of inclusions, call it $\imath = []$ and pick the traversal order for the elements in $U$ - $[a, b, c, d, e, f]$ - and an order respecting the condition for all $k \in \{1, \dots, p\}$, the induced subgraph of $\mathscr{F}_{n-1}V$ on $(y^{(j)})_{j=1}^{k}$ is connected for the elements in $V$ - $[b', a', c']$. Note that this order is different than the traversal order of $V$ which is $[a', b', c']$ - this would have also been a valid order. For each $k \in \{1, 2, 3\}$, we let $V^{(k)} \coloneqq \bigcup_{j \leq k} \operatorname{cl}\{y^{(j)}\}$. In this case, since $V$ only has three maximal elements,

$$V^{(1)} = \operatorname{cl} b' =$$




,

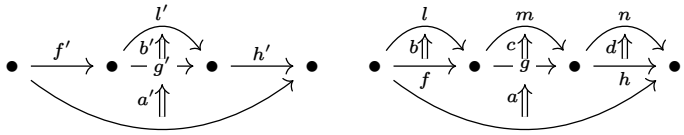
$$V^{(2)} = \operatorname{cl}\{a', b'\} =$$




,

$$V^{(3)} = \operatorname{cl}\{a', b', c'\} = V.$$

The algorithm first tries to match $\operatorname{cl} b'$ to $\operatorname{cl} a$; they are not isomorphic so it iterates to the next element in $U$. For $i = 2$, it tries to match $\operatorname{cl} b'$ to $\operatorname{cl} b$; $V^{(1)}$ is isomorphic to $\operatorname{cl} b$ so we let $\imath^{(2,1)}$ be the inclusion mapping $\operatorname{cl} b'$ to $\operatorname{cl} b$. Next, we try to extend the match by finding an inclusion from $V^{(2)}$ to $U$. In this case $k = 2$ and $j = 1$. Let's name the 1-dimensional elements in $U$ and $V$ that are involved at this step in the algorithm:

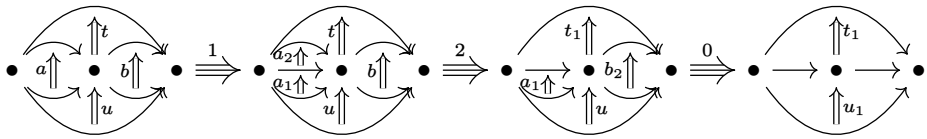



Because of our connectedness assumption (there is an edge in the flow graph between $b'$ and $a'$), there exists an element $z \in \Delta^{\alpha} b' \cap \Delta^{-\alpha} a'$, for $\alpha \in \{+, -\}$. In this case, we have only one such element, $g'$, so we do not have to choose one. Now, $\imath^{(2,1)}(b') = b$ is one coface of $\imath^{(2,1)}(g') = f$ in $U$. Element $f$ has exactly one other coface in $U$, $a$, so we check if $\operatorname{cl} a'$ and $\operatorname{cl} a$ are isomorphic. There is indeed an isomophism $\operatorname{cl}\{a'\} \xhookrightarrow{\sim} \operatorname{cl}\{a\}$, but this isomorphism does not match the assignment that we had from the previous inclusion: $g'$ is mapped to $g$ by this (latest) morphism, but $\operatorname{cl}\{a'\} \cap V^{(1)} = g'$ which is mapped to $f$ by $\imath^{(2,1)}$. So, we cannot extend the inclusion $\imath^{(2,1)}$. We break the inner loop and continue with $i = 3$. The algorithm now finds an isomorphism $\imath^{(3,1)} : \operatorname{cl}\{b'\} \xhookrightarrow{\sim} \operatorname{cl}\{c\}$. Again, for $k = 2$, $V^{(2)} = \operatorname{cl}\{a', b'\}$ (because $b'$ is the only previous element in the order to which $a'$ is connected) and $g'$ is the only element in $\Delta^{\alpha} b' \cap \Delta^{-\alpha} a'$. The only other element (apart form $c$) covering $g$ in $U$ is $a$. Again, there is an isomorphism $\operatorname{cl}\{a'\} \xhookrightarrow{\sim} \operatorname{cl}\{a\}$ which matches $\imath^{(3,1)}$ on $\operatorname{cl}\{a'\} \cap V^{(1)}$. So we let $\imath^{(3,2)}$ be the unique extension of $V^{(3,1)}$ restricted to $\operatorname{cl} a' \xhookrightarrow{\sim} \operatorname{cl} a$. We continue by iterating on $k = 3$; $V^{(3)} = V$ and $y^{(3)} = c'$. The element $y^{(j)}$ is now $a'$ (the only element $c'$ is connected to) so $j = 2$ and $h' \in \Delta^{+} a' \cap \Delta^{-} c'$. We repeat the same argument as above: $\operatorname{cl} c' \xhookrightarrow{\sim} \operatorname{cl} d$

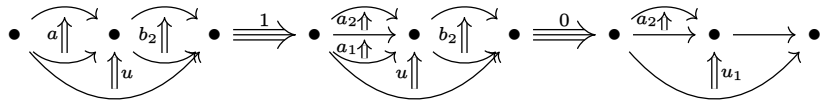
which matches $\imath^{(3,2)}$ on $\operatorname{cl}\{c'\} \cap V^{(2)}$ and we extend $\imath^{(3,2)}$ to $\imath^{(3,3)}$ by restricting to $\operatorname{cl} c' \xhookrightarrow{\sim} \operatorname{cl} d$. Since $k = 3$, and the whole of $V$ has been traversed, we let $\imath = [\imath^{(3,3)}]$. The algorithm continues by setting $i = 4$ and trying to match $b'$ with $d$. In the end, the algorithm returns $\imath = [\imath^{(3,3)}]$.

The molecule matching algorithm is a stepping stone for identifying submolecule inclusions, but not all inclusions that it returns are submolecule inclusions. As we will see later on, the molecule matching algorithm identifies submolecule inclusions in dimensions less than or equal to 2. However, from dimension 3 onwards, more conditions are necessary. Below we provide a counterexample of a 3-dimensional molecule inclusion found by the algorithm that is not a submolecule inclusion and in the next chapter we discuss the submolecule inclusion algorithm on the general case of $n$-dimensional molecules.

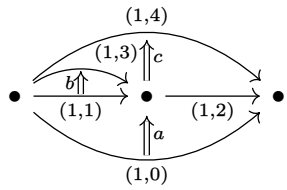
***Example* 3.4.4 —** Let $U$ be:

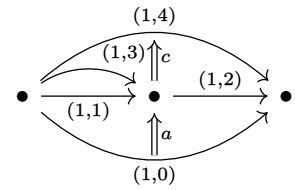



and let $V$ be:




There is an inclusion of $V$ into $U$ that maps the elements labelwise. But this inclusion is not a submolecule inclusion because $\operatorname{cl}\{(3,1),(3,0)\}$ does not appear as a factor in a pasting decomposition - i.e., $U$ cannot be written as a pasting in which 0 and 1 are consecutive.
Intuitively what stops it from being a submolecule inclusion is the presence of the 3-dimensional rewrite 2 in $U$ which creates a dependency between the 3-elements 0 and 1: the input of 0 is dependent on the output of 2 and the input of 2 is dependent on the output of 1. Thus, the images of the elements 0 and 1 from $V$ cannot appear in consecutive order in a pasting decomposition of $U$.
There is an extra condition that inclusions in dimension 3 should satisfy in order to be submolecule inclusions. However, we are not aware of such conditions for molecules of dimension 4 or greater. We believe there is a gap in complexity between deciding if an inclusion is a submolecule inclusion for molecules of dimension less than or equal to 3 versus in the general case. It is exactly these dependencies that gives rise to the phenomena causing an exponential running time in the rewritable submolecule problem.
Situations like the one above do not occur in dimension 2. If it were to occur, then $V$ would not be a molecule. We leave below an example of a *dependency* in dimension 2 which is similar to the situation above:

Let $U =$ 

and let $V =$ 


Firstly, $V$ is not a molecule because of the lack of a 2-dimensional element between the two arrows, $(1,1)$ and $(1,3)$. Also note that the molecule matching algorithm would not return $V$ as a molecule inclusion into $U$; $V$ is the closure of the atoms $a$ and $c$, however, $V$ cannot be a submolecule of $U$ because of the dependency between $a, b, c$ in $U$. We will come back to this later when discussing obstructions.

# Chapter 4

# The rewritable submolecule problem

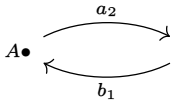
In the last chapter we presented an algorithm that returns all the inclusions of a molecule $V$ into another molecule of the same dimension, $U$, whenever $V$ is round. However, by Example 3.4.4 this is not enough to decide whether $V$ is a submolecule of $U$. The inductive definition of submolecule inclusion makes it challenging to directly translate into an algorithm. To counteract this we start studying the relation between the layerings and orderings of a molecule. We present the conditions for which an ordering is induced by a layering. Since by Definition 2.3.19, orderings are topological sorts of a graph, we obtain a deterministic and easy to manipulate criterion to find valid layerings for a molecule. We continue by tying the submolecule inclusion to the existence of a certain layering that corresponds to the intuition we gave at the beginning of the previous chapter: a submolecule appears as a factor in a pasting decomposition. Finally, we link the submolecule inclusion to the existence of a certain ordering. The rewritable submolecule algorithm is then based on this result.
In the general case, the rewritable submolecule decision algorithm has a factorial running time. However, under certain acyclicity conditions, the running time is improved. We present these acyclicity conditions and show that the subdiagram matching problem for molecules of dimension less than or equal to 3 takes linear time in the size of the Hasse diagram representation of the molecule. This makes rewriting on 3-dimensional diagrams computationally feasible.
Most of the theory presented in this chapter is part of [29, Section 3 and Section 4]. For a detailed discussion on the connection between layerings and orderings and the acyclicity conditions, check [27].

## 4.1 Relation between layerings and orderings

In this section we present the connection between layerings and orderings via path-induced subgraphs. The connection is motivated by the following proposition.

**Proposition 4.1.1.** *Let $U$ be a molecule, $k \geq -1$. If $U$ admits a $k$-layering, then $\mathscr{M}_k U$ is acyclic, and $U$ admits a $k$-ordering.*

*Proof.* Let $(U^{(i)})_{i=1}^m$ be a $k$-layering of $U$. For each $i \in \{1, \dots, m\}$, the graph $\mathscr{M}_k U^{(i)}$ is trivially acyclic by Lemma 3.2.3. We conclude by applying Lemma 2.3.22 repeatedly. ∎

***Example* 4.1.2 —** Recall the molecule from Example 3.3.8 whose frame dimension is 2, but it did not have a 2-layering. Its maximal flow graph, $\mathscr{M}_2 U$ contains a cycle: 

.

**Corollary 4.1.3.** *Let $U$ be a molecule, $n \coloneqq \dim U$. Then $\mathscr{F}_{n-1} U$ is acyclic.*

*Proof.* By Theorem 3.2.11, $U$ always admits an $(n-1)$-layering. We conclude by Proposition 4.1.1 combined with the fact that $\mathscr{F}_{n-1} U = \mathscr{M}_{n-1} U$. ∎

***Definition* 4.1.4 —** Let $U$ be a molecule, $k \geq -1$. We let

$$\mathscr{L}ay_k U \coloneqq \left\{ k\text{-layerings } (U^{(i)})_{i=1}^m \text{ of } U \text{ up to layer-wise isomorphism} \right\},$$
$$\mathscr{O}rd_k U \coloneqq \left\{ k\text{-orderings } (x^{(i)})_{i=1}^m \text{ of } U \right\},$$

**Proposition 4.1.5.** *Let $U$ be a molecule, $k \geq -1$. For each $k$-layering $(U^{(i)})_{i=1}^m$ of $U$ and each $i \in \{1, \dots, m\}$, let $x^{(i)}$ be the only element of $\bigcup_{j>k} (\mathscr{M}ax\, U)_j$ in the image of $U^{(i)}$. Then the assignment*

$$\mathsf{o}_{k,U} \colon (U^{(i)})_{i=1}^m \mapsto (x^{(i)})_{i=1}^m \tag{4.1}$$

*determines an injective function $\mathscr{L}ay_k U \hookrightarrow \mathscr{O}rd_k U$.*

*Proof.* By Lemma 3.2.3, the assignment $(U^{(i)})_{i=1}^m \mapsto (x^{(i)})_{i=1}^m$ is well-defined. Let $i, j \in \{1, \dots, m\}$, and suppose that there is an edge from $x^{(i)}$ to $x^{(j)}$ in $\mathscr{M}_k U$, that is, there exists $z \in \Delta_k^+ x^{(i)} \cap \Delta_k^- x^{(j)}$. By Proposition 4.1.1, $\mathscr{M}_k U$ is acyclic, so necessarily $i \neq j$. If $j < i$, then $U^{(j)} \cap U^{(i)} \subseteq \partial_k^+ U^{(j)} \cap \partial_k^- U^{(i)}$ by [25, Lemma 4.2.4], contradicting the existence of $z$. It follows that $i < j$, so $(x^{(i)})_{i=1}^m$ is a $k$-ordering of $U$.
Let $(V^{(i)})_{i=1}^m$ be another $k$-layering, and suppose it determines the same $k$-ordering as $(U^{(i)})_{i=1}^m$. Then the image of both $U^{(1)}$ and $V^{(1)}$ in $U$ is

$$\mathrm{cl}\left\{x^{(1)}\right\} \cup \partial^- U,$$

so $U^{(1)}$ is isomorphic to $V^{(1)}$. If $m = 1$ we are done. Otherwise, $(U^{(i)})_{i=2}^m$ and $(V^{(i)})_{i=2}^m$ are $k$-layerings inducing the same $k$-ordering on their image. By recursion, we conclude that they are layer-wise isomorphic. ∎

We can now formally define a rewrite step in our framework. As discussed in the introduction, rewriting on an $n$-dimensional molecule is equivalent to gluing the

$(n+1)$-dimensional diagram representing the rewrite to the target diagram. This process can be understood as pasting the whiskered $(n+1)$-dimensional diagram onto the target diagram, which, by Proposition 2.3.13 (unitality), is isomorphic to the already whiskered diagram. Consequently, applying an $(n+1)$-dimensional rewrite to an $n$-dimensional diagram is equivalent to replacing the input of the rewrite rule in the target diagram with its output. Moreover, the result of applying a sequence of $(n+1)$-dimensional rewrites on an $n$-dimensional diagram is the output boundary of the $(n+1)$-dimensional diagram obtained by sequentially gluing the rewrite rules. We formalize this intuition below:

***Definition* 4.1.6** (Rewrite steps) **—** Let $U$ be a molecule, $k \geq -1$, and let $(U^{(i)})_{i=1}^{m}$ be a $k$-layering of $U$. The sequence $(U^{\mathsf{st}(i)})_{i=0}^{m}$ of *rewrite steps* associated with $(U^{(i)})_{i=1}^{m}$ is defined recursively by

- $U^{\mathsf{st}(0)} \coloneqq \partial_k^- U$,
- $U^{\mathsf{st}(i)} \coloneqq \partial_k^+ U^{(i)}$ for $i \in \{1, \dots, m\}$.

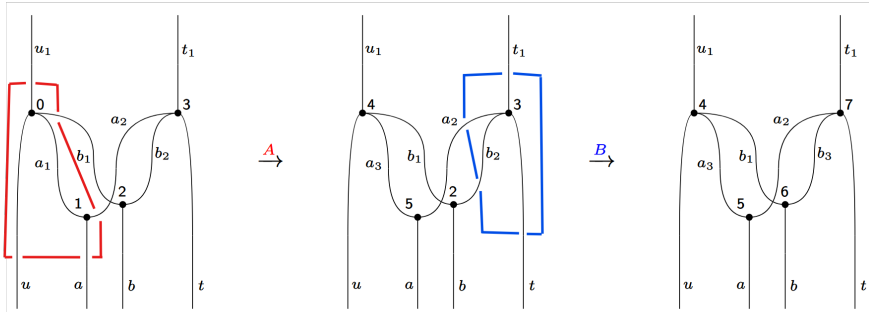
We can now finally prove the intuition we introduced about rewrite steps and layerings: i.e., if we have a sequence of rewrite steps associated to a $k$-layering induced by a $k$-ordering, then the result of applying the $i$-th rewrite is equivalent to substituting the input of $x^{(i)}$ with the output of $x^{(i)}$ in $U^{(i-1)}$.

**Corollary 4.1.7.** *Let $U$ be a molecule, $k \geq -1$, $(U^{(i)})_{i=1}^{m}$ a $k$-layering of $U$, and $(x^{(i)})_{i=1}^{m}$ the associated $k$-ordering. For all $i \in \{1, \dots, m\}$, the $i$-th rewrite step $U^{\mathsf{st}(i)}$ is isomorphic to $U^{\mathsf{st}(i-1)}[\partial_k^+ x^{(i)} / \partial_k^- x^{(i)}]$.*

*Proof.* By Proposition 3.2.3, $x^{(i)}$ is the only element of dimension $> k$ in $U^{(i)}$. The result then follows from repeated application of Lemma 3.2.12. ∎

***Example* 4.1.8 —** We have already secretly used rewrite steps when representing the shapes of diagrams of dimension greater than 3 as a sequence of rewrites. More specifically, in the molecule, $U$, from Example 3.3.8, its input boundary is a sequence of 3-dimensional rewrites on 2-dimensional diagrams. This representation is possible due to the existence of layerings and each such sequence is a rewrite step. In the end, we depict the whole of the 4-dimensional molecule as a sequence of sequences of rewrites. One advantage of this representation is that it lets one show which $(n-1)$-dimensional elements are selected to be in the input of the rewrite one wants to apply. The ability of writing a diagram as a sequence of rewrite steps is of great importance in helping us visualise a higher dimensional shape of diagram. We illustrate this below by looking at the string diagram representation of $U$.

Note that in Example 3.3.8 we had to state which elements are being rewritten by both $A$ and $B$. Now, this will be clear from the representation. To help illustrate, we circled (and colour coded) the areas containing the elements being rewritten. The red area corresponds to the input of $A$, while the blue one represents the input of $B$.




Note how $A$ rewrites $\mathrm{cl}\{0,1\}$ into $\mathrm{cl}\{4,5\}$, while preserving the input and output boundaries.

The following proposition gives the criteria for which a $k$-ordering is a $k$-layering.

**Proposition 4.1.9.** *Let $U$ be a molecule, $k \geq -1$, and let $(x^{(i)})_{i=1}^m$ be a $k$-ordering of $U$. Let*

$$\begin{aligned} U^{(0)} &\coloneqq \partial_k^- U, \\ U^{(i)} &\coloneqq \partial_k^+ U^{(i-1)} \cup \mathrm{cl}\left\{x^{(i)}\right\} \quad \text{for } i \in \{1,\dots,m\}. \end{aligned}$$

*The following are equivalent:*

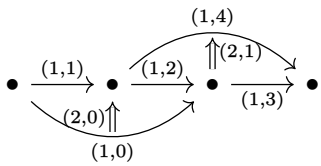
*(a) $(U^{(i)})_{i=1}^m$ is a $k$-layering of $U$;*

*(b) for all $i \in \{1,\dots,m\}$, $\partial_k^- x^{(i)} \sqsubseteq \partial_k^- U^{(i)}$.*

*Moreover, for all $i \in \{1,\dots,m-1\}$, if $\partial_k^- x^{(i)} \sqsubseteq \partial_k^- U^{(i)}$, then $U^{(i)}$ and $\partial_k^+ U^{(i)} = \partial_k^- U^{(i+1)}$ are molecules.*

*Proof.* Suppose $(U^{(i)})_{i=1}^m$ is a $k$-layering. Then, for all $i \in \{1,\dots,m\}$, $U^{(i)}$ is a molecule, and by Proposition 3.2.3 $x^{(i)}$ is the only element of dimension $> k$ in $U^{(i)}$. By Lemma 3.2.12, $\partial_k^- x^{(i)} \sqsubseteq \partial_k^- U^{(i)}$.
Conversely, it follows from Lemma 3.1.12 that for all $i \in \{1,\dots,m\}$, if $\partial_k^- U^{(i)}$ is a molecule and $\partial_k^- x^{(i)} \sqsubseteq \partial_k^- U^{(i)}$, then $U^{(i)}$ is a molecule, hence $\partial_k^+ U^{(i)}$ is a molecule. Moreover, since $(x^{(i)})_{i=1}^m$ is a $k$-ordering, it is straightforward to prove that $U^{(i)} \cap U^{(i+1)} = \partial_k^+ U^{(i)} = \partial_k^- U^{(i+1)}$ for all $i \in \{1,\dots,m-1\}$. Since $\partial^- U^{(1)} = \partial^- U$ is a molecule, it follows by induction, assuming condition (b), that $U^{(i)}$ is a molecule for all $i \in \{1,\dots,m\}$. This proves that $(U^{(i)})_{i=1}^m$ is a $k$-layering of $U$. ∎

Consider the following molecule, $U$:



We pick the 1-ordering $\{(2,0),(2,1)\}$, where

$$U^{(0)} = \bullet \xrightarrow[(1,0)]{} \bullet \xrightarrow[(1,3)]{} \bullet ,$$

$$U^{(1)} = \bullet \xrightarrow{(1,1)} \bullet \xrightarrow{(1,2)} \bullet \xrightarrow[(1,3)]{} \bullet , \quad (2,0) \Uparrow \ \text{over} \ (1,0)$$

$$U^{(2)} = \bullet \xrightarrow{(1,1)} \bullet \xrightarrow{(1,2)} \bullet \xrightarrow[(1,3)]{} \bullet . \quad \Uparrow (2,1)$$

By Proposition 4.1.9 if $(U^{(i)})_{i=1}^{2}$ is a 1-layering of $U$, then $(2,0) \sqsubseteq \partial_1^- U^{(1)}$ and $(2,1) \sqsubseteq \partial_1^- U^{(2)}$.
This proposition applies in the general case, but it is still far from helping us with the subdiagram matching problem. Our approach is to relate the submolecule inclusion relation to $(n-1)$-layerings and later on to $(n-1)$-orderings.
As mentioned in the chapter introduction, the criteria for $V$ being a submolecule of $U$ relies on $V$ appearing as a factor in a pasting decomposition of $U$. In our solution, we will consider layerings of $U$. Since $V$ is rewritable, its maximal elements will appear amongst the layers of $U$. By Proposition 4.1.1 we are looking for a layering inducing an ordering in which the maximal elements of $V$ are consecutive. Since any ordering of the elements of $V$ suffices and $V$ is round, we can contract the graph $\mathscr{F}_{n-1}U$ by *merging* the maximal elements of $V$. By merging the maximal elements of $V$, we obtain a molecule which is an atom, $\langle V \rangle$, whose input boundary is $\partial^- V$ and whose output boundary is $\partial^+ V$. Our problem now simplifies to searching through the topological sorts of a contracted graph, $\mathscr{F}_{n-1}U / \mathscr{F}_{n-1}V$.
Recall the definition of an induced subgraph from 2.3.5.

***Definition* 4.1.10** (Path-induced subgraph) — Let $\mathscr{G}$ be a directed graph and $W$ a subset of its vertex set. We say that induced subgraph $\mathscr{G}|_W$ is *path-induced* if, for all $x, y \in W$, every path from $x$ to $y$ in $\mathscr{G}$ is included in $\mathscr{G}|_W$.

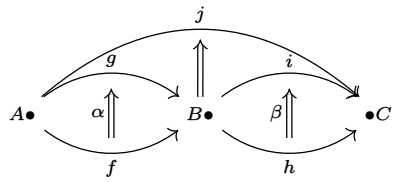
Path-induced subgraphs are also called *convex subgraphs*, for example in [8].

***Definition* 4.1.11** (Contraction of a connected subgraph) — Let $\mathscr{G}$ be a directed graph and $W$ a subset of its vertex set such that $\mathscr{G}|_W$ is connected. The *contraction of* $\mathscr{G}|_W$ *in* $\mathscr{G}$ is the graph minor $\mathscr{G}/(\mathscr{G}|_W)$ obtained by contracting every edge in $\mathscr{G}|_W$.

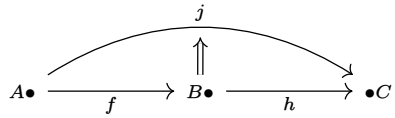
***Definition* 4.1.12** (Merger of a round molecule) — Let $U$ be a round molecule. The *merger* of $U$ is $\langle U \rangle \coloneqq \partial^- U \Rightarrow \partial^+ U$.

Another way to look at the merger is as the "unbiased" composition of the top-dimensional elements of a round molecule.

***Example* 4.1.13 —** Let $U$ be the following 2-dimensional molecule:




Its merger, $\langle U \rangle$, is:




Note that if $U$ is an atom, $\dim U > 0$, then it is easy to check that it is isomorphic to $\langle U \rangle$.

**Lemma 4.1.14.** *Let $\mathscr{G}$ be a directed acyclic graph and $W \subseteq V_{\mathscr{G}}$ such that $\mathscr{G}|_W$ is connected. The following are equivalent:*

*(a) $\mathscr{G}|_W$ is path-induced;*

*(b) $\mathscr{G}/(\mathscr{G}|_W)$ is acyclic;*

*(c) there is a topological sort of $\mathscr{G}$ in which vertices of $W$ are consecutive.*

*Moreover, under any of the equivalent conditions, there is a bijection between*

- *topological sorts of $\mathscr{G}$ in which vertices of $W$ are consecutive,*
- *pairs of a topological sort of $\mathscr{G}|_W$ and a topological sort of $\mathscr{G}/(\mathscr{G}|_W)$.*

*Proof.* We prove the contrapositive of the implication from (a) to (b). Suppose $\mathscr{G}/(\mathscr{G}|_W)$ has a cycle. If the cycle does not pass through $x_W$, then it lifts to a cycle in $\mathscr{G}$, contradicting the assumption that $\mathscr{G}$ is acyclic. It follows that the cycle contains a segment of the form $x_W \to x_1 \to \ldots \to x_m \to x_W$, where $m > 0$ and $x_i \neq x_W$ for all $i \in \{1, \ldots, m\}$. Then there exist $y, z \in W$ and a path $y \to x_1 \to \ldots \to x_m \to z$ in $\mathscr{G}$, so $\mathscr{G}|_W$ is not path-induced.
Next, suppose that $\mathscr{G}/(\mathscr{G}|_W)$ is acyclic. Then both $\mathscr{G}/(\mathscr{G}|_W)$ and $\mathscr{G}|_W$ are acyclic, so they admit topological sorts $(x^{(i)})_{i=1}^m$ and $(y^{(j)})_{j=1}^p$, respectively. For exactly one $q \in \{1, \ldots, m\}$, $x^{(i)} = x_W$. We claim that

$$((x^{(i)})_{i=1}^{q-1}, (y^{(j)})_{j=1}^p, (x^{(i)})_{i=q+1}^m)$$

is a topological sort of $\mathscr{G}$. Indeed, for all edges from $x$ to $x'$ in $\mathscr{G}$,

- if $x, x' \notin W$, then $x = x^{(i)}$, $x' = x^{(i')}$ for some $i, i' \in \{1, \ldots, m\} \setminus \{q\}$, and there is an edge from $x$ to $x'$ in $\mathscr{G}/(\mathscr{G}|_W)$, so $i < i'$;
- if $x, x' \in W$, then $x = y^{(j)}$, $x' = y^{(j')}$ for some $j, j' \in \{1, \ldots, p\}$, and there is an edge from $x$ to $x'$ in $\mathscr{G}|_W$, so $j < j'$;

- if $x \in W$, $x' \notin W$, then $x = y^{(j)}$, $x' = x^{(i)}$ for some $i \in \{1, \dots, m\} \setminus \{q\}$, $j \in \{1, \dots, p\}$, and there is an edge from $x_W$ to $x'$ in $\mathscr{G}/(\mathscr{G}|_W)$, so $q < i$;

- if $x \notin W$, $x' \in W$, then $x = x^{(i)}$, $x' = y^{(j)}$ for some $i \in \{1, \dots, m\} \setminus \{q\}$, $j \in \{1, \dots, p\}$, and there is an edge from $x$ to $x_W$ in $\mathscr{G}/(\mathscr{G}|_W)$, so $i < q$.

This proves the implication from (b) to (c). Moreover, it defines an injection from pairs of a topological sort of $\mathscr{G}|_W$ and a topological sort of $\mathscr{G}/(\mathscr{G}|_W)$ to topological sorts of $\mathscr{G}$ in which the vertices of $W$ are consecutive. This will prove to be a bijection as soon as we have proven the converse implication.
Finally, we prove the contrapositive of the implication from (c) to (a). Suppose $\mathscr{G}|_W$ is not path-induced, that is, there is a path $x \to x_1 \to \dots \to x_m \to y$ in $\mathscr{G}$ such that $m > 0$, $x, y \in W$, and $x_i \notin W$ for all $i \in \{1, \dots, m\}$. It follows that the $x_i$ must come between $x$ and $y$ in every topological sort of $\mathscr{G}$, so the vertices of $W$ can never be consecutive. ∎

**Proposition 4.1.15.** *Let $\imath \colon V \hookrightarrow U$ be an inclusion of molecules such that $\dim V = \dim U$ and $V$ is round. The following are equivalent:*

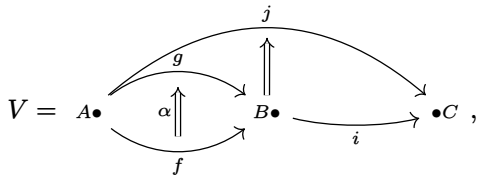
*(a) $\imath$ is a submolecule inclusion;*

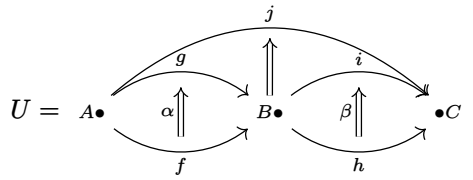
*(b) for all molecules $W$ such that $V \Rightarrow W$ is defined, $U[W/\imath(V)]$ is a molecule and $j \colon W \hookrightarrow U[W/\imath(V)]$ is a submolecule inclusion;*

*(c) $U[\langle V \rangle / \imath(V)]$ is a molecule.*

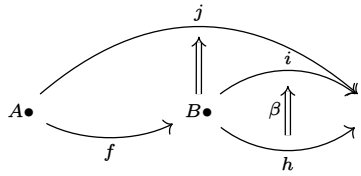
*Proof.* If $\imath$ is a submolecule inclusion, by Lemma 3.1.12 $U \cup (V \Rightarrow W)$ and its output boundary $U[W/\imath(V)]$ are molecules, and the inclusion of $W$ into $U[W/\imath(V)]$ is a submolecule inclusion.
If $V$ is a round molecule, then $\langle V \rangle$ is an atom, which is round by Corollary 2.2.28, and has boundaries isomorphic to those of $V$ by Lemma 2.2.15. By [25, Corollary 3.4.14], $V \Rightarrow \langle V \rangle$ is defined, so the third condition is a special case of the second one.
Finally, suppose $U[\langle V \rangle / \imath(V)]$ is a molecule. By Lemma 3.1.4, since $\langle V \rangle$ is an atom, its inclusion $j$ into $U[\langle V \rangle / \imath(V)]$ is a submolecule inclusion. Using Lemma 3.1.12 as in the first part, we deduce that $(U[\langle V \rangle / \imath(V)])[V/j(\langle V \rangle)]$ is a molecule, and the inclusion of $V$ into it is a submolecule inclusion. By Lemma [25, Lemma 4.1.14], $(U[\langle V \rangle / \imath(V)])[V/j(\langle V \rangle)]$ is isomorphic to $U$, and $\imath$ factors as this submolecule inclusion followed by an isomorphism. ∎

***Example* 4.1.16 —** By Proposition 4.1.15 if we have an inclusion $V \hookrightarrow U$, where



$U =$ 

,

then $U[\langle V\rangle / \imath(V)] =$ 

is a molecule.

**Lemma 4.1.17.** *Let $\imath\colon V \hookrightarrow U$ be an inclusion of molecules such that $n \coloneqq \dim U = \dim V$ and $V$ is round. Then $\mathscr{F}_{n-1}U[\langle V\rangle/\imath(V)]$ is isomorphic to $\mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$.*

*Proof.* By Lemma 2.3.21 and Proposition 3.3.10, $\mathscr{F}_{n-1}V$ is a connected induced subgraph of $\mathscr{F}_{n-1}U$, so its contraction is well-defined. Now, the vertices of the graph $\mathscr{F}_{n-1}U[\langle V\rangle/\imath(V)]$ are either

- $x \in U_n \setminus V_n$, or
- $x_V$ such that the image of $\langle V\rangle$ in $U[\langle V\rangle/\imath(V)]$ is $\operatorname{cl}\{x_V\}$.

Let $x, y$ be two vertices of $\mathscr{F}_{n-1}U[\langle V\rangle/\imath(V)]$.

- If $x, y \in U_n \setminus V_n$, then $\Delta^+ x \cap \Delta^- y$ is the same in $U[\langle V\rangle/\imath(V)]$ as in $U$, so there is an edge from $x$ to $y$ in $\mathscr{F}_{n-1}U$ if and only if there is an edge in $\mathscr{F}_{n-1}U[\langle V\rangle/\imath(V)]$.
- If $x = x_V$ then $\Delta^+ x_V \cap \Delta^- y$ is in bijection with $\Delta^+ V \cap \Delta^- y$ in $U$. For all $z \in \Delta^+ V$, since $V$ is pure and $n$-dimensional, there exists $w \in \nabla^+ z$. If $\Delta^+ x_V \cap \Delta^- y$ is non-empty, it follows that $\Delta^+ z \cap \Delta^- y$ is non-empty in $U$ for some $z \in \imath(V)_n$. Thus there exist $z \in \imath(V)_n$ and an edge from $z$ to $y$ in $\mathscr{F}_{n-1}U$.
- Dually, if $y = x_V$, there is an edge from $x$ to $y$ in $\mathscr{F}_{n-1}U[\langle V\rangle/\imath(V)]$ if and only if there exist $z \in \imath(V)_n$ and an edge from $x$ to $z$ in $\mathscr{F}_{n-1}U$.
- Finally, $\Delta^+ V \cap \Delta^- V = \varnothing$ because $V$ is pure, so $\Delta^+ x_V \cap \Delta^- x_V$ and there is no edge from $x_V$ to $x_V$.

It is then straightforward to establish an isomorphism with the explicit description of $\mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$. ∎

**Proposition 4.1.18.** *Let $\imath\colon V \hookrightarrow U$ be an inclusion of molecules such that $n \coloneqq \dim U = \dim V$ and $V$ is round. If $\imath$ is a submolecule inclusion, then $\mathscr{F}_{n-1}V$ is a path-induced subgraph of $\mathscr{F}_{n-1}U$.*

*Proof.* By Proposition 4.1.15, if $\imath$ is a submolecule inclusion then $U[\langle V\rangle/\imath(V)]$ is a molecule. By Corollary 4.1.3 $\mathscr{F}_{n-1}U[\langle V\rangle/\imath(V)]$ is acyclic, and by Lemma 4.1.17 it is isomorphic to $\mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$. It follows from Lemma 4.1.14 that $\mathscr{F}_{n-1}V$ is a path-induced subgraph of $\mathscr{F}_{n-1}U$. ∎

Finally, the following results are the ones on which the rewritable submolecule algorithm relies on.

**Lemma 4.1.19.** *Let $\imath\colon V \hookrightarrow U$ be an inclusion of molecules such that $n \coloneqq \dim U = \dim V$ and $V$ is round, and let $(y^{(i)})_{i=1}^{p}$ be an $(n-1)$-ordering induced by an $(n-1)$-layering of $V$. The following are equivalent:*

*(a) $\imath$ is a submolecule inclusion;*

*(b) there exist an $(n-1)$-ordering $(x^{(i)})_{i=1}^{m}$ induced by an $(n-1)$-layering $(U^{(i)})_{i=1}^{m}$ of $U$, and $q \in \{1,\dots,m\}$ such that*

*1. $(x^{(i)})_{i=q}^{p+q-1} = (\imath(y^{(i)}))_{i=1}^{p}$,*

*2. $\imath(\partial^- V) \sqsubseteq \partial^- U^{(q)}$.*

*Proof.* Identify $V$ with its isomorphic image through $\imath$, and suppose that $\imath$ is a submolecule inclusion. Then $\tilde{U} \coloneqq U[\langle V\rangle/V]$ is a molecule by Proposition 4.1.15, and admits an $(n-1)$-layering $(\tilde{U}^{(i)})_{i=1}^{m-p+1}$ by Theorem 3.2.11. Let $\operatorname{cl}\{x\}$ be the image of $\langle V\rangle$ in $\tilde{U}$; then $x \in \tilde{U}^{(q)}$ for exactly one $q \in \{1,\dots,m-p+1\}$. Then $W \coloneqq \tilde{U}^{(q)}[V/\operatorname{cl}\{x\}]$ is defined, and by [25, Lemma 4.1.16] combined with Lemma [25, Lemma 4.1.14], $U$ is isomorphic to

$$\tilde{U}^{(1)} \#_{n-1} \cdots \#_{n-1} \tilde{U}^{(q-1)} \#_{n-1} W \#_{n-1} \tilde{U}^{(q+1)} \#_{n-1} \cdots \#_{n-1} \tilde{U}^{(m-p+1)}.$$

By Lemma 3.2.12, $\partial^- x \sqsubseteq \partial^- \tilde{U}^{(q)}$, so by [25, Lemma 4.1.15] $\partial^- V \sqsubseteq \partial^- W$. We can apply the criterion of Proposition 4.1.9 to deduce that $(y^{(i)})_{i=1}^{p}$ is an $(n-1)$-ordering of $W$ induced by an $(n-1)$-layering $(W^{(i)})_{i=1}^{p}$. Letting

$$(U^{(i)})_{i=1}^{m} \coloneqq ((\tilde{U}^{(i)})_{i=1}^{q-1}, (W^{(i)})_{i=1}^{p}, (\tilde{U}^{(i)})_{i=q+1}^{m-p+1}),$$

produces an $(n-1)$-layering of $U$, hence also an $(n-1)$-ordering $(x^{(i)})_{i=1}^{m}$ of $U$, with the property that $(x^{(i)})_{i=q}^{p+q-1} = (y^{(i)})_{i=1}^{p}$.
Conversely, let $(U^{(i)})_{i=1}^{m}$ be an $(n-1)$-layering of $U$ satisfying the properties in the statement, and let $W \sqsubseteq U$ be the image of $U^{(q)} \#_{n-1} \cdots \#_{n-1} U^{(p+q-1)}$ in $U$. Then $W_n = V_n$, so

$$W = V \cup \partial^- W.$$

Because $\partial^- V \sqsubseteq \partial^- U^{(q)} = \partial^- W$, by Lemma 3.1.12 $V \sqsubseteq W \sqsubseteq U$. ∎

**Theorem 4.1.20.** *Let $\imath\colon V \hookrightarrow U$ be an inclusion of molecules such that $n \coloneqq \dim U = \dim V$ and $V$ is round, $m \coloneqq |U_n|$, $p \coloneqq |V_n|$. The following are equivalent:*

*(a) $\imath$ is a submolecule inclusion;*

*(b) there is a topological sort $((x^{(i)})_{i=1}^{q-1}, x_V, (x^{(i)})_{i=q+1}^{m-p+1})$ of $\mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$ such that, letting*

$$\begin{aligned} U^{(0)} &\coloneqq \partial^- U, \\ U^{(q)} &\coloneqq \partial^+_{n-1} U^{(q-1)} \cup \imath(V), \\ U^{(i)} &\coloneqq \partial^+_{n-1} U^{(i-1)} \cup \operatorname{cl}\left\{x^{(i)}\right\} \quad \text{for } i \neq q, \end{aligned}$$

*we have $\imath(\partial^- V) \sqsubseteq \partial^- U^{(q)}$ and $\partial^- x^{(i)} \sqsubseteq \partial^- U^{(i)}$ for all $i \neq q$.*

*Proof.* Identify $V$ with its isomorphic image through $\imath$, and suppose that $\imath$ is a submolecule inclusion. Then $\tilde{U} \coloneqq U[\langle V \rangle / V]$ is a molecule by Proposition 4.1.15, so it admits an $(n-1)$-layering $(\tilde{U}^{(i)})_{i=1}^{m-p+1}$, which induces an $(n-1)$-ordering. By Lemma 4.1.17, this $(n-1)$-ordering can be identified with a topological sort $((x^{(i)})_{i=1}^{q-1}, x_V, (x^{(i)})_{i=q+1}^{m-p+1})$ of $\mathscr{F}_{n-1}U / \mathscr{F}_{n-1}V$. By Lemma 3.2.12, we have $\partial^- x_V \sqsubseteq \partial^- \tilde{U}^{(q)}$ and $\partial^- x^{(i)} \sqsubseteq \partial^- \tilde{U}^{(i)}$ for $i \neq q$. By [25, Lemma 4.1.16] combined with Lemma [25, Lemma 4.1.14], letting $W \coloneqq \tilde{U}^{(q)}[V / \mathrm{cl}\,\{x_V\}]$, $U$ is isomorphic to

$$\tilde{U}^{(1)} \#_{n-1} \cdots \#_{n-1} \tilde{U}^{(q-1)} \#_{n-1} W \#_{n-1} \tilde{U}^{(q+1)} \#_{n-1} \cdots \#_{n-1} \tilde{U}^{(m-p+1)},$$

and $W$ is isomorphic to $U^{(q)}$, while $\tilde{U}^{(i)}$ is isomorphic to $U^{(i)}$ for all $i \neq q$. We conclude by [25, Lemma 4.1.15].
Conversely, let $(y^{(i)})_{i=1}^{p}$ be an $(n-1)$-ordering induced by an $(n-1)$-layering of $V$. Then $((x^{(i)})_{i=1}^{q-1}, (y^{(i)})_{i=1}^{p}, (x^{(i)})_{i=q+1}^{m-p+1})$ is an $(n-1)$-ordering of $U$, and by the criterion of Proposition 4.1.9 it is induced by an $(n-1)$-layering. We conclude by Lemma 4.1.19. ∎

Lemma 4.1.19 describes the relation between layerings and orderings as a criteria for rewritable submolecule inclusion. Theorem 4.1.20 however shifts the layering criteria to finding a certain topological sort and recursively checking that the boundary of the top-dimensional element is a submolecule of the layer it is part of.

## 4.2 The rewritable submolecule decision algorithm

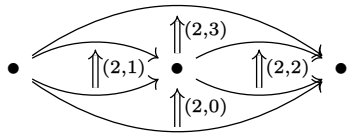
We saw in Example 3.4.4 a matching that is not a submolecule inclusion. As it turns out, the rewritable submolecule problem is more difficult. In this chapter we present an algorithm for deciding when a molecule inclusion is a submolecule inclusion. This algorithm has a factorial running time, but under certain acyclicity conditions (satisfied by all molecules of dimension less than or equal to 3), the runtime is linear.

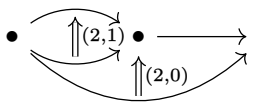
***Procedure* 4.2.1** (Rewritable submolecule decision algorithm) — The procedure takes as input an inclusion $V \subseteq U$ of molecules such that $n \coloneqq \dim U = \dim V$ and $V$ is round, and it returns whether $V \sqsubseteq U$. We let $m \coloneqq |U_n|$ and $p \coloneqq |V_n|$.
We construct the graph $\mathscr{G} \coloneqq \mathscr{F}_{n-1}U / \mathscr{F}_{n-1}V$. Then we start a loop. At each iteration, we search for a new topological sort of $\mathscr{G}$. If we cannot find one, we return *false*. Else, let $((x^{(i)})_{i=1}^{q-1}, x_V, (x^{(i)})_{i=q+1}^{m-p+1})$ be the new topological sort, and let $(U^{(i)})_{i=1}^{m-p+1}$ be as in Theorem 4.1.20.
For each $i \in \{1, \dots, m-p+1\}$, we start a recursive call to the algorithm to decide whether $\partial^- x^{(i)} \sqsubseteq \partial^- U^{(i)}$ if $i \neq q$, and $\partial^- V \sqsubseteq \partial^- U^{(q)}$ if $i = q$. If this returns *false*, we break the iteration on $i$ and iterate the main loop. If this returns *true*, we iterate on $i$. At the end of the iteration on $i$, we return *true*.

Below we present the algorithm in pseudocode.

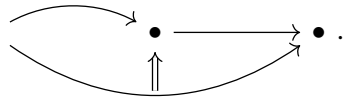
```
procedure RewritableSubmol(V ⊆ U:inclusion of molecules, V is round, n := dim U = dim V)
    n := dim U
    m := |U_n|
    p := |V_n|
    Construct 𝒢 := 𝓕_{n-1}U/𝓕_{n-1}V.
    for t ∈ tsorts of 𝒢 do
                                  ▷ t is of the form ((x^(i))_{i=1}^{q-1}, x_V, (x^(i))_{i=q+1}^{m-p+1})
        Let (U^(i))_{i=1}^{m-p+1} be as in Theorem 4.1.20.
        for i ← {1, ... m - p + 1} do
            if i ≠ q then
                if RewritableSubmol (∂⁻x^(i) ⊆ ∂⁻U^(i)) == False then
                    break
                else if RewritableSubmol (∂⁻V ⊆ ∂⁻U^(q)) == False then
                    break
        return True
    return False
```

To begin with, consider an example on 2-dimensional diagrams. Even though, in the 2-dimensional case, an inclusion is already a submolecule inclusion, it is instructive to see how the algorithm works on the lower dimensional cases.

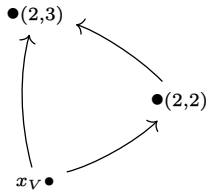
Let $U =$ 

and let $V =$ 

.

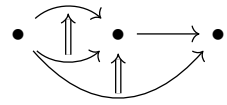
The merger of $V$ is $\langle V \rangle =$ 
.

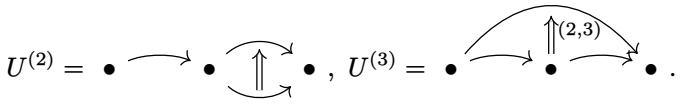
The inclusion maps $(2,0)$ in $V$ to $(2,0)$ in $U$ and $(2,1)$ in $V$ to $(2,1)$ in $U$. Then $|U_n| = 4$, $|V_n| = 2$ and the graph $\mathcal{G}$ is:

$\mathcal{G} := \mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V =$ 

.

The only topological sort of $\mathcal{G}$ is $[x_v, (2,2), (2,3)]$ and $(U^{(i)})_{i=1}^{m-p+1}$ is for each $i$, where $q = 1$:

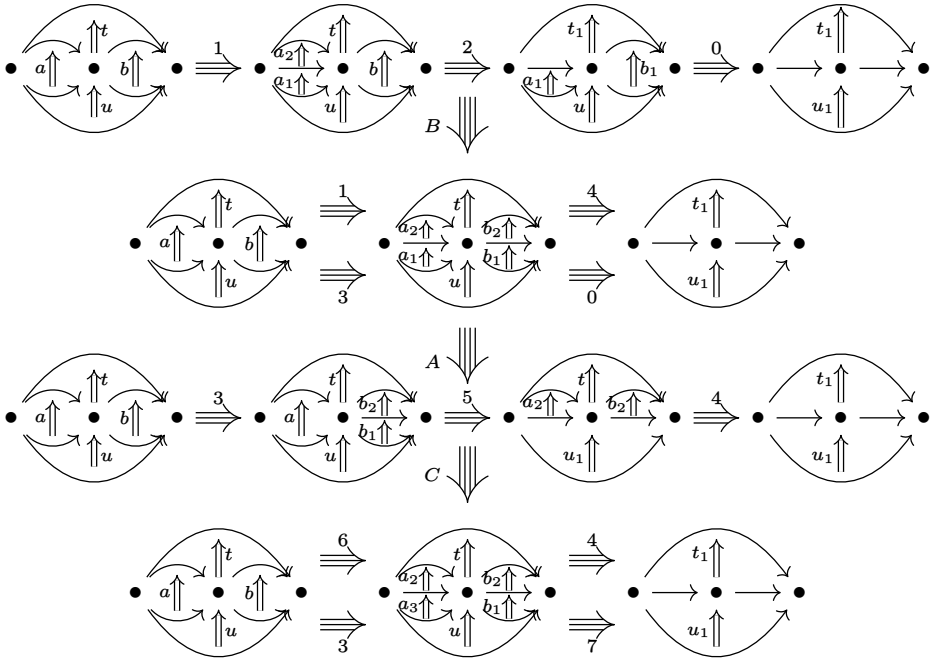
$U^{(0)} =$ 
, $U^{(1)} = \imath(V) =$ 
,




Then for each $i$, we make a recursive call to check that $\partial^- x^{(i)} \sqsubseteq \partial^- U^{(i)}$ if $i \neq q$, and $\partial^- V \sqsubseteq \partial^- U^{(q)}$ if $i = q$. In this example, we have by Lemma 3.1.4 that $\partial^- V \sqsubseteq \partial^- U^{(1)}$ since $\partial^- V$ is an atom and for each $x^{(i)}$, $\partial^- x^{(i)} \sqsubseteq \partial^- U^{(i)}$ since the boundaries of $(2,2)$ and $(2,3)$ are atoms so we can conclude by Lemma 3.1.4. However, this situation rarely occurs and is not covered by the algorithm as a special case. The algorithm will make a second recursive call to check the input boundaries of each $\partial^- x^{(i)}$ and $\partial^- V$ are submolecule inclusions (i.e., $\partial^-(\partial^- x^{(i)}) \sqsubseteq \partial^- \partial^- U^{(i)}$ and $\partial^- \partial^- V \sqsubseteq \partial^- \partial^- U^{(1)}$).

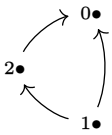
At the end of this example, we have that $V \sqsubseteq U$.

Before going into the analysis of this algorithm, let's simulate it on the boundary of a 4-dimensional molecule. This will help when we introduce the acyclicity conditions and obstructions. Consider the following 4-dimensional shape of diagram, call it $U$. What we want to check is: "Is $\partial^- A$ a submolecule of $\partial^- U$?" (Note that these are the molecules from Example 3.4.4).

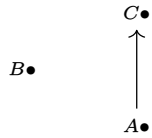



(4.2)

The answer to the question is no, because there is a dependency in the input of $A$ on element $2 \in \Delta^- B$. But let's see how the algorithm determines this. The inclusion map returned by the molecule matching algorithm is as expected. Note that $\partial^- A = \mathrm{cl}\{0,1\}$, so $|(\partial^- U)_3| = 3$ and $|(\partial^- A)_3| = 2$. The flowgraphs are:

$\mathscr{F}_2\partial^- U =$  and $\mathcal{G} \coloneqq \mathscr{F}_2\partial^- U / \mathscr{F}_2\partial^- A =$  .

Since $\mathcal{G}$ contains a cycle, there is no topological sort of it. So, $\partial^- A \not\sqsubseteq \partial^- U$. Moreover, note that the flow graph of $U$ is:



This flow graph has three orderings: $[A, B, C], [A, C, B], [B, A, C]$, however because $\partial^- A \not\sqsubseteq \partial^- U$ only $[B, A, C]$ is induced by a valid layering. For a more detailed discussion on why this is the case, see Section 4.5.

**Remark 4.2.2.** *The rewritable submolecule algorithm returns whether an inclusion of molecules is a submolecule inclusion and not whether a layering is a valid layering. For example, in the molecule above, by letting $U_A$ be a layer that contains the element $A$ and $U_B$ being a layer that contains the element $B$, the layering $U_A \#_3 U_B$ is not a valid layering, but the algorithm will return that $V = \operatorname{cl}\{A, B\}$ is a submolecule of $U$ since $U_B \#_3 U_A$ appears as a factor in a pasting decomposition of $U$.*

**Theorem 4.2.3.** *The rewritable submolecule decision algorithm is correct: it always terminates, and returns* true *if and only if $V \sqsubseteq U$.*

*Proof.* We proceed by induction on the dimension $n$ of $U$ and $V$. If $n = 0$, this is straightforward, so let $n > 0$.
The number of iterations of the main loop is bounded by the number of topological sorts of $\mathscr{F}_{n-1}U / \mathscr{F}_{n-1}V$, which is finite. Consider one such iteration, producing a topological sort $((x^{(i)})_{i=1}^{q-1}, x_V, (x^{(i)})_{i=q+1}^{m-p+1})$. Let us write $V^{(q)} \coloneqq V$ and $V^{(i)} \coloneqq \operatorname{cl}\{x^{(i)}\}$ for $i \neq q$. For all $i \in \{1, \dots, m-p+1\}$, we have a call to the decision algorithm with input $\partial^- V^{(i)} \subseteq \partial^- U^{(i)}$, assuming that the calls for $j < i$ all returned *true*.
Now, $\partial^- V^{(i)}$ is round by Proposition 2.2.26 and Lemma 2.2.30 and Corollary 2.2.28. Moreover, $\partial^- U^{(1)} = \partial^- U$, which is a molecule. For $i > 1$, assuming that $\partial^- V^{(i-1)} \sqsubseteq \partial^- U^{(i-1)}$, we may apply Proposition 4.1.15 to derive that $U^{(i-1)}$ and $\partial^+ U^{(i-1)} = \partial^- U^{(i)}$ are molecules. Thus

I the input of the first call is well-formed,

II for $i > 1$, assuming that the $(i-1)$-th call correctly returned *true*, the input of the $i$-th call is well-formed.

Since all of these are in dimension $(n-1)$, by the inductive hypothesis, each call terminates returning the correct answer. By Theorem 4.1.20, this proves both correctness and termination in dimension $n$. ∎

**Theorem 4.2.4.** *The rewritable submolecule problem in dimension $n$ can be solved in time*

$$O\left(\prod_{k\leq n} |U_k|!\,|U_k|\right).$$

*Proof.* For $n = 0$, this is obvious, so let $n > 0$. The number of iterations of the main loop is bounded above by the number of topological sorts of $\mathscr{G} \coloneqq \mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$. This reaches its maximum when $\mathscr{G}$ is a discrete graph, in which case the number is $(|U_n| - |V_n| + 1)!$, tightly bounded above by $|U_n|!$.
At each iteration of the main loop, we have at most $|U_n| - |V_n| + 1$ calls to the algorithm on molecules of dimension $n-1$ contained in $U$. By the inductive hypothesis, these take time $O(\prod_{k\leq n-1} |U_k|!\,|U_k|)$.
All other operations have lower complexity: both finding topological sorts and computing the boundaries of the $U^{(i)}$ take linear time in $|\mathscr{E}_n U|$, but this can be bounded above by $|U_n|\,|U_{n-1}|$, and we conclude. ∎

It still remains an open question whether the subdiagram matching problem admits a polynomial-time algorithm in arbitrary dimension. On the other hand, we show that the problem is in NP.

**Proposition 4.2.5.** *For all $n \in \mathbb{N}$, the $n$-dimensional subdiagram matching problem is in* NP*.*

*Proof.* It suffices to prove by induction on $n$ that the rewritable submolecule problem in dimension $n$ is in NP. When $n = 0$, the problem is trivial. In dimension $n > 0$, a polynomial-size certificate that $V \sqsubseteq U$ is given by

I a topological sort $((x^{(i)})_{i=1}^{q-1}, x_V, (x^{(i)})_{i=q+1}^{m-p+1})$ of the graph $\mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$, and

II polynomial-size certificates that $\partial^- V \sqsubseteq \partial^- U^{(q)}$ and $\partial^- x^{(i)} \sqsubseteq \partial^- U^{(i)}$ for all $i \neq q$,

with the notations of Theorem 4.1.20. By the inductive hypothesis this exists and is verifiable in polynomial time. ∎

## 4.3 Runtime improvements under acyclicity conditions

The running time of the algorithm above is factorial in the worst case due to the iteration of the topological sorts. However, under certain acyclcicity conditions, this runtime can be improved.

***Definition* 4.3.1** (Frame-acyclic molecule) — Let $U$ be a molecule. We say that $U$ is *frame-acyclic* if for all submolecules $V \sqsubseteq U$, if $r \coloneqq \operatorname{frdim} V$, then $\mathscr{M}_r V$ is acyclic.

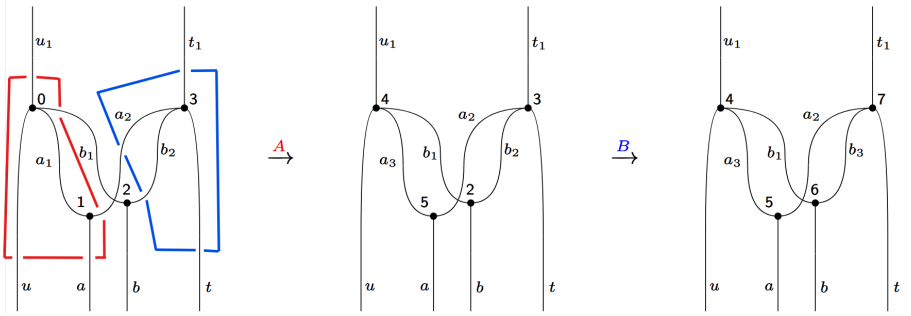
**Lemma 4.3.2.** *Let $U$ be a molecule, $V \sqsubseteq U$. If $U$ is frame-acyclic, then $V$ is frame-acyclic.*

*Proof.* Straightforward. ∎

As we will see later in this section, the first example of a non-frame acyclic molecule first appears in dimension 4. The molecule in figure 4.2 is not frame acyclic and we have already seen that its flow graph admits topological sorts that are not induced by layerings. This is the main reason for the factorial increase in complexity in our algorithm. However, in the example from figure 4.2 there is one more phenomenon going on which we discuss at the end of this section. For this reason, we present below a "nicer" shape of diagram in which the phenomenon mentioned previously does not occur. Moreover, we exemplify how to spot frame-acyclicity in a string diagram representation of a molecule.

***Example* 4.3.3 —** Let $U$ be the 4-dimensional molecule from Example 3.3.8.

Note that $\mathscr{M}_2 U$ contains a cycle: $A \bullet \overset{a_2}{\underset{b_1}{\rightleftarrows}} \bullet B$ . However, the dependency is more apparent in the string diagram representation.




The portion enclosed by the red line is the input of $A$, namely $\partial^- A$, while the one inside the blue line is the input of $B$, i.e., $\partial^- B$. The line is interrupted when it comes across the elements that are part of the input or output 2-boundary of $A$ or $B$. We identify the elements that are part of the input boundary as the lines that cross the coloured line and go into 3-dimensional elements. Similarly, the output boundaries are identified as the lines that cross the coloured enclosure and go out of the 3-dimensional elements. For example, the elements in the 2-input of $A$ are $u, a, a_1$, while the elements in its 2-output are $a_2, u_1$. Similarly for $B$: $b, t, a_2$ are in its 2-input while $b_1, t_1$ are in its 2-output.

To see that the molecule is not frame-acyclic, note how one of the 2-dimensional elements in the output of $A$ - $a_2$ which is going out of the highlighted red area - is required in the 2-dimensional input of $B$ (it is going into the highlighted blue area), while one of the elements in the 2-dimensional output of $B$ - $b_1$ which is going out of the highlighted blue area - is required in the input of $A$.

The reason why $U$ is a molecule and we can do either of the rewrites is because these dependencies do not cause any "deadlock"; no matter which of them we want to apply first, the 3-dimensional input of the respective layer can be written such that the elements of the rewrite we want to apply appear in consecutive order.

**Lemma 4.3.4.** *Let $U$ be a molecule. Suppose that for all submolecules $V \sqsubseteq U$, if $r \coloneqq \operatorname{frdim} V$, then $V$ admits an $r$-layering. Then for all $k \geq \operatorname{frdim} U$ the function*

$\mathsf{o}_{k,U} : \mathscr{L}ay_k U \hookrightarrow \mathscr{O}rd_k U$ *is a bijection.*

*Proof.* See [25, Lemma 8.1.3]. ∎

**Theorem 4.3.5.** *Let $U$ be a molecule, $r \coloneqq \operatorname{frdim} U$. If $U$ is frame-acyclic, then $U$ admits an $r$-layering.*

*Proof.* By Lemma 4.3.2, we can proceed by induction on submolecules. For all $x \in U_0$, we have $\operatorname{frdim}\{x\} = -1$, and $\{x\}$ admits the trivial $(-1)$-layering, which proves the base case.
We construct a finite plane tree of submolecules $U^{(j_1,\dots,j_p)} \sqsubseteq U$, as follows:

- the root is $U^{()} \coloneqq U$;
- if $\operatorname{lydim} U^{(j_1,\dots,j_p)} \leq r$, then we let $\operatorname{lydim} U^{(j_1,\dots,j_p)}$ be a leaf;
- if $k \coloneqq \operatorname{lydim} U^{(j_1,\dots,j_p)} > r$, then we pick a $k$-layering $(V^{(i)})_{i=1}^{q}$ of $U^{(j_1,\dots,j_p)}$, which is possible by Theorem 3.2.11, and for each $i \in \{1,\dots,q\}$, we let the image of $V^{(i)}$ be a child $U^{(j_1,\dots,j_p,i)}$ of $U^{(j_1,\dots,j_p)}$.

By Lemma 3.2.10, the layering dimension of the children of a node is strictly smaller than that of the node, so the procedure terminates.
Fix an $r$-ordering $(x^{(i)})_{i=1}^{m}$ of $U$; this is possible because $\mathscr{M}_r U$ is acyclic. Let $V \coloneqq U^{(j_1,\dots,j_p)}$ be a node of the tree. We have

$$\bigcup_{j>r} (\mathscr{M}ax\, V)_j = \sum_{i=1}^{m} \bigcup_{j>r} \left( (\mathscr{M}ax\, V)_j \cap \operatorname{cl}\left\{x^{(i)}\right\} \right) =: \sum_{i=1}^{m} M^{(i)};$$

the $M^{(i)}$ form a partition because $\operatorname{frdim} U = r$, so every element of dimension $> r$ is in the closure of $x^{(i)}$ for a unique $i \in \{1,\dots,m\}$. We claim that $V$ is isomorphic to

$$V^{(1)} \#_r \dots \#_r V^{(m)}$$

for some molecules $(V^{(i)})_{i=1}^{m}$ such that, for each $i \in \{1,\dots,m\}$, identifying $V^{(i)}$ with its image in $V$, we have

$$\bigcup_{j>r} (\mathscr{M}ax\, V^{(i)})_j = M^{(i)}.$$

We will prove this by backward induction on the tree $U^{(j_1,\dots,j_p)}$.
Suppose $V$ is a leaf, so $\operatorname{lydim} V \leq r$. Then $V$ admits an $r$-layering. For each $i \in \{1,\dots,m\}$, fix a topological sort $(y^{(i,j)})_{j=1}^{p_i}$ of the induced subgraph $\mathscr{M}_r V|_{M^{(i)}}$. We claim that $((y^{(i,j)})_{j=1}^{p_i})_{i=1}^{m}$ is an $r$-ordering of $V$.
Suppose there is an edge from $x$ to $x'$ in $\mathscr{M}_r V$. Then $x \in M^{(i)}$, $x' \in M^{(i')}$ for a unique pair $i, i' \in \{1,\dots,m\}$. If $i = i'$, then $x = y^{(i,j)}$ and $x' = y^{(i,j')}$ for some $j, j' \in \{1,\dots,p_i\}$, and $j < j'$ because $(y^{(i,j)})_{j=1}^{p_i}$ is a topological sort of $\mathscr{M}_r V|_{M^{(i)}}$. If $i \neq i'$, then there exists

$$z \in \Delta_r^+ x \cap \Delta_r^- x' \subseteq \operatorname{cl}\left\{x^{(i)}\right\} \cap \operatorname{cl}\left\{x^{(i')}\right\}.$$

Since $\partial_r^\alpha x^{(i)}$ and $\partial_r^\alpha x^{(i')}$ is pure and $r$-dimensional for all $\alpha \in \{+,-\}$, by [25, Proposition 4.4.8]

$$z \in (\Delta_r^+ x^{(i)} \cap \Delta_r^- x^{(i')}) \cup (\Delta_r^- x^{(i)} \cap \Delta_r^+ x^{(i')}),$$

and by [25, Lemma 2.1.24] $\Delta_r^- x^{(i)} \cap \operatorname{cl}\{x\} \subseteq \Delta_r^- x$ which is disjoint from $\Delta_r^+ x$, so $z \in \Delta_r^+ x^{(i)} \cap \Delta_r^- x^{(i')}$. It follows that there is an edge from $x^{(i)}$ to $x^{(i')}$ in $\mathscr{M}_r U$, so $i < i'$ because $(x^{(i)})_{i=1}^m$ is a topological sort of $\mathscr{M}_r U$. This proves that $((y^{(i,j)})_{j=1}^{p_i})_{i=1}^m$ is an $r$-ordering of $V$.
Let $W \sqsubseteq V$, $\ell \coloneqq \operatorname{frdim} W$. If $V \neq U$ or $W \neq U$, then $W$ admits an $\ell$-layering by the inductive hypothesis on proper submolecules of $U$. If $W = V = U$ then $\ell = r$ and $W$ admits an $\ell$-layering by Theorem 3.2.11. In either case, $V$ satisfies the conditions of Lemma 4.3.4, and since $r \geq \operatorname{lydim} V \geq \operatorname{frdim} V$, every $r$-ordering of $V$ comes from an $r$-layering of $V$.
It follows that $((y^{(i,j)})_{j=1}^{p_i})_{i=1}^m$ comes from an $r$-layering $((W^{(i,j)})_{j=1}^{p_i})_{i=1}^m$, and we can define

$$V^{(i)} \coloneqq W^{(i,1)} \#_r \dots \#_r W^{(i,p_i)}$$

for each $i \in \{1,\dots,m\}$, satisfying the desired condition.
Now, suppose that $V$ is not a leaf, so $k \coloneqq \operatorname{lydim} V > r$, and $V$ has children $(W^{(j)})_{j=1}^q$ forming a $k$-layering of $V$. By the inductive hypothesis, each of the $W^{(j)}$ has a decomposition

$$W^{(j,1)} \#_r \dots \#_r W^{(j,m)}$$

such that the maximal elements of dimension $> r$ in the image of $W^{(j,i)}$ are contained in $\operatorname{cl}\{x^{(i)}\}$. Then, for each $i \in \{1,\dots,m\}$ and $j, j' \in \{1,\dots,q\}$,

$$W^{(j,i)} \cap W^{(j')} \subseteq W^{(j',i)},$$

so $V^{(i)} \coloneqq W^{(1,i)} \#_k \dots \#_k W^{(q,i)}$ is defined. Using point 3 from Proposition 2.3.13 repeatedly, we conclude that $V$ is isomorphic to $V^{(1)} \#_r \dots \#_r V^{(m)}$.
This concludes the induction on the tree $U^{(j_1,\dots,j_p)}$. In particular, for the root $U^{()} = U$, the decomposition $U^{(1)} \#_r \dots \#_r U^{(m)}$ satisfies

$$\bigcup_{j>r} (\mathscr{M}ax\, U^{(i)})_j = \left\{x^{(i)}\right\},$$

that is, $(U^{(i)})_{i=1}^m$ is an $r$-layering of $U$. ∎

Frame-acyclicity has further implications in the relation between layerings and orderings, the most important (or used in the treatment of computational complexity aspects for diagram rewriting) being the one in which the $k$-orderings and $k$-layerings are in bijection:

**Corollary 4.3.6.** *Let $U$ be a molecule. The following are equivalent:*

*(a) $U$ is frame-acyclic;*

*(b) for all $V \sqsubseteq U$ and all $\operatorname{frdim} V \leq k < \dim V$, $V$ admits a $k$-layering;*

*(c) for all $V \sqsubseteq U$ and all $\operatorname{frdim} V \leq k < \dim V$, the sets $\mathscr{L}ay_k V$ and $\mathscr{O}rd_k V$ are non-empty and equinumerous.*

*Proof.* The implication from (a) to (b) is a consequence of Theorem 4.3.5 together with 4.3.2 and Lemma 3.2.4. The implication from (b) to (c) is Lemma 4.3.4. Finally, the implication from (c) to (a) follows from Proposition 4.1.1. ∎

***Example* 4.3.7 —** The molecule in Example 3.3.8 is not frame-acyclic and it does not admit a 2-layering even if its frame dimension is 2.

**Proposition 4.3.8.** *If $U$ is guaranteed to be frame-acyclic, the rewritable submolecule problem in dimension $n$ can be solved in time*

$$O\left(\prod_{k \leq n} |U_k|!\right).$$

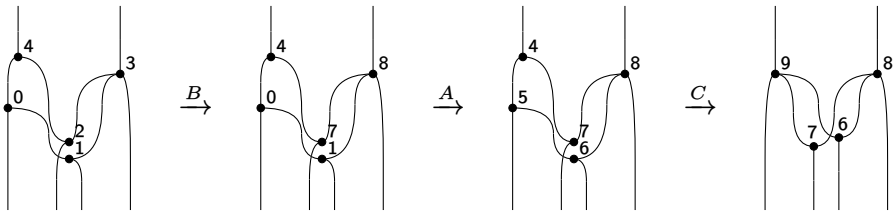
*Proof.* Given a topological sort $((x^{(i)})_{i=1}^{q-1}, x_V, (x^{(i)})_{i=q+1}^{m-p+1})$ of $\mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$, by Lemma 4.1.14, substituting any $(n-1)$-ordering of $V$ for $x_V$ produces an $(n-1)$-ordering of $U$ in which the elements of $V$ are consecutive. By Theorem 4.3.6, this is induced by an $(n-1)$-layering of $U$. By Lemma 4.1.19, it suffices to check that $\partial^- V \sqsubseteq \partial^- U^{(q)}$ to conclude that $V \sqsubseteq U$, so we have a single recursive call instead of $O(|U_n|)$ many. Since $\partial^- U^{(q)} \sqsubseteq U^{(q)} \sqsubseteq U$, it is frame-acyclic, and we can proceed inductively. ∎

Note that in this case, the running time is still factorial but this suggests a stronger acyclicity condition might reduce the running time to polynomial. Once a topological sort is found, we know it is induced by a layering in $U$, so to verify if $V \sqsubseteq U$ by Lemma 4.1.19 we only need to check that $\partial^- V \sqsubseteq \partial^- U^{(q)}$. The trick is that the ordering we find is induced by a layering of $U$, but not necessarily by a layering of $U[\langle V \rangle / V]$. So, we still need to iterate searching for a topological sort of the contracted flow graph since $U[\langle V \rangle / V]$ is not necessarily frame-acyclic. The next condition has the consequence that if a topological sort of $U[\langle V \rangle / V]$ exists, then it must be induced by a layering.

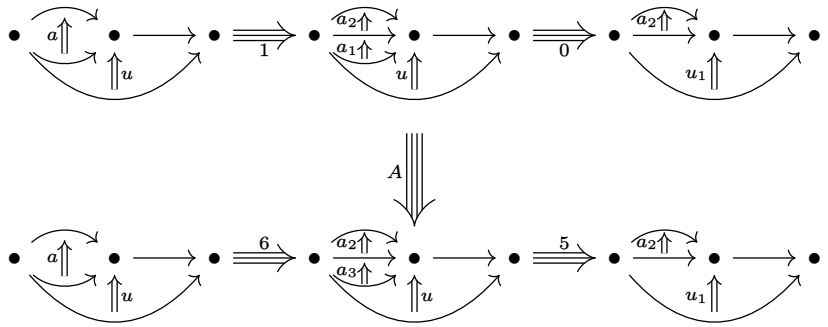
***Definition* 4.3.9** (Stably frame-acyclic molecule) **—** Let $U$ be a molecule. We say that $U$ is *stably frame-acyclic* if for all submolecules $V \sqsubseteq U$ and all rewritable submolecules $W \sqsubseteq V$, the molecule $V[\langle W \rangle / W]$ is frame-acyclic.

Every stably frame-acyclic molecule is frame-acyclic: if we take $V \sqsubseteq U$ to be an atom, the substitution $U[\langle V \rangle / V]$ is trivial. Moreover, every submolecule of a (stably) frame-acyclic molecule is (stably) frame-acyclic.
However, the converse of the statement above is not true; that is, not every frame-acyclic molecule is stably frame-acyclic. We show in the example below that the classes of frame-acyclic and stably frame-acyclic molecules do not coincide:

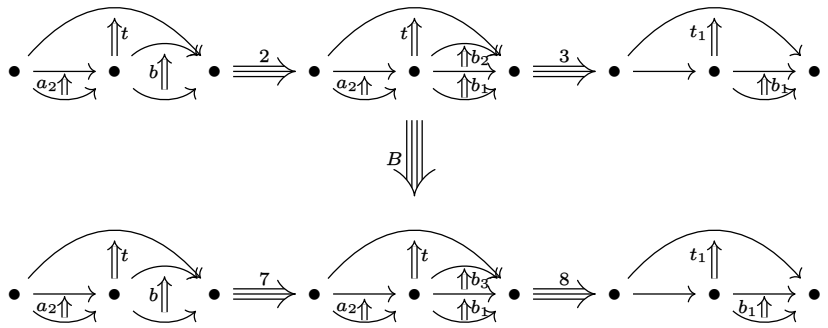
***Example* 4.3.10 —** There exists a 4-dimensional molecule $U$ whose representation as a sequence of string diagram rewrites (in the style of Definition 4.1.6 and Example 4.1.8) is the following:




Since $U$ is frame-acyclic and $W = \operatorname{cl}\{A, B\}$ is a submolecule of $U$ with $\operatorname{frdim}(W) = 2$, then by Theorem 4.3.5 it admits a 2-layering. To see that, let us spell out $A$ and $B$ in their pasting diagram representation. As a pasting diagram, $\operatorname{cl} A$ is:




while $\operatorname{cl} B$ is:

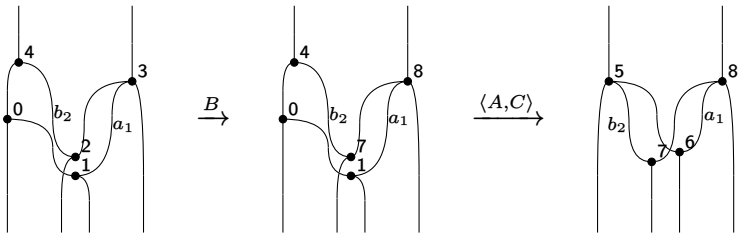



The maximal 2-flow graph of $W$ is:

$$\mathscr{M}_2 W = \quad A\bullet \xrightarrow{\quad a_2 \quad} \bullet B$$

There is only one 2-ordering of $W$ which can be induced by the 2-layering $U_A \#_2 U_B$, where $U_A$ is a layer containing $A$ and $U_B$ is a layer containing $B$. Then $U_A \#_2 U_B$ is indeed a valid 2-layering where $U_A$ is obtained by whiskering $A$ with the 2-elements $b$ and $t$ and $U_b$ is the layer obtained by whiskering $B$ with $u_1$.
Let $V = \operatorname{cl}\{A, C\}$ be a submolecule of $U$. Then $U[\langle V \rangle / V]$ is the molecule which

corresponds to the following sequence of rewrites:




The maximal flow graph of $U[\langle V\rangle/V]$ has the following cycle:

$$\mathscr{M}_2 U[\langle V\rangle/V]: \quad \langle A, C\rangle \underset{b_2}{\overset{a_1}{\rightleftarrows}} B$$

showing that $U$ is not stably frame-acyclic.

**Proposition 4.3.11.** *If $U$ is guaranteed to be stably frame-acyclic, the rewritable submolecule problem can be solved in linear time in the size of $\mathscr{H}U$.*

*Proof.* If $V \sqsubseteq U$, by assumption $U[\langle V\rangle/V]$ is frame-acyclic. By Theorem 4.3.6, *all* its $(n-1)$-orderings are induced by $(n-1)$-layerings, and by Lemma 4.1.17 they are in bijection with topological sorts of $\mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$. From Proposition 4.1.9 it follows that, if any topological sort of $\mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$ fails to satisfy the conditions of Theorem 4.1.20, then $V$ is not a submolecule of $U$, so in the decision algorithm we can stop after the first iteration of the main loop.
This involves finding a single topological sort and computing $\partial^- U^{(q)}$, both of which take time $O(|\mathscr{E}_n U|)$; then, as in Proposition 4.3.8, we make a single call to the decision algorithm for $\partial^- V \sqsubseteq \partial^- U^{(q)}$. Since $\partial^- U^{(q)} \sqsubseteq U^{(q)} \sqsubseteq U$, it is stably frame-acyclic, and we can proceed inductively. ∎

The difference between the two conditions is that if $U$ is only frame-acyclic, the molecule $U[\langle V\rangle/V]$ is not necessarily frame-acyclic so its orderings are not in bijection with its layerings. The improvement in the running time is because $U[\langle V\rangle/V]$ is frame-acyclic, so any $(n-1)$-ordering is induced by an $(n-1)$-layering; by Proposition 4.1.9, $\partial^- x_V \sqsubseteq \partial^- U^{(q)}$ so any ordering should satisfy the conditions of Theorem 4.1.20.

**Corollary 4.3.12.** *The subdiagram matching problem restricted to diagrams with stably frame-acyclic shape is in* P.

The two conditions of frame-acyclicity and stably frame-acyclicity seem difficult to verify in practice. However, they are implied by stronger acyclicity conditions.

***Definition* 4.3.13** (Acyclic molecule) — Let $U$ be a molecule. We say that $U$ is *acyclic* if $\vec{\mathscr{H}}U$ is acyclic.

***Definition* 4.3.14** (Dimension-wise acyclic molecule) — Let $U$ be a molecule. We say that $U$ is *dimension-wise acyclic* if, for all $k \in \mathbb{N}$, $\mathscr{F}_k U$ is acyclic.

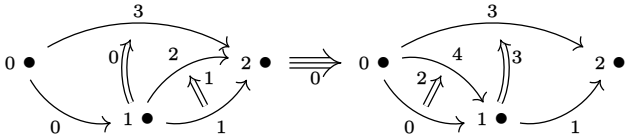
**Proposition 4.3.15.** *Let $U$ be a molecule. Then*

*I if $U$ is acyclic, it is dimension-wise acyclic;*

*II if $U$ is dimension-wise acyclic, it is frame-acyclic.*

*Proof.* Suppose $U$ is acyclic. Let $k \in \mathbb{N}$ and suppose there is a cycle $x_0 \to x_1 \to \ldots \to x_m = x_0$ in $\mathscr{F}_k U$. By definition, for all $i \in \{1, \ldots, m\}$ there exists $y_i \in \Delta_k^+ x_{i-1} \cap \Delta_k^- x_i$. By Lemma 2.3.9, there exist paths from $x_{i-1}$ to $x_i$ passing through $y_i$ in $\vec{\mathscr{H}} U$. Concatenating all these paths, we obtain a cycle in $\vec{\mathscr{H}} U$.
Suppose $U$ is dimension-wise acyclic, and let $V \sqsubseteq U$ be a submolecule inclusion, $r \coloneqq \operatorname{frdim} V$. Then $\mathscr{F}_r U$ is acyclic, hence so are its induced subgraphs $\mathscr{F}_r V$ and $\mathscr{M}_r V$. ∎

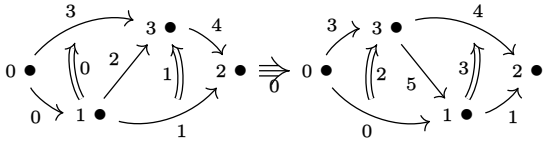
***Example* 4.3.16 —** [29, Example 109] Both implications are strict. The 3-dimensional atom




(based on [51, Fig. 2]) is not acyclic, since its oriented Hasse diagram contains the cycle

$$(0,1) \to (1,1) \to (2,1) \to (3,0) \to (2,2) \to (1,4) \to (0,1),$$

but it is dimension-wise acyclic. The 3-dimensional atom




(based on [51, Fig. 4]) is not dimension-wise acyclic, since its 0-flow graph contains the cycle $(1,2) \to (1,5) \to (1,2)$, but as we will soon see it is stably frame-acyclic by Theorem 4.4.16.

## 4.4 Results in lower dimensions

We continue with the results in lower dimensions that were previously mentioned and end the section with a 4-dimensional diagram exemplifying when the factorial running time is arising.

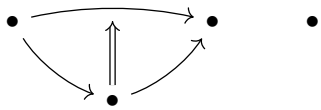
**Lemma 4.4.1.** *Let $U$ be a molecule. If $\dim U = 0$, then $U$ is isomorphic to the point.*

*Proof.* By induction on the construction of $U$. If $U$ was produced by (*Point*), then $U = 1$ and $\dim U = 0$. If $U$ was produced by (*Paste*), then it is equal to $V \#_k W$ where $V$, $W$ are molecules with $k < \min\{\dim V, \dim W\}$. Then $\dim U = \max\{\dim V, \dim W\} > k \geq 0$. If $U$ was produced by (*Atom*), then it is of the form $V \Rightarrow W$, and $\dim U = \dim V + 1 = \dim W + 1 > 0$. ∎

**Lemma 4.4.2.** *Let $U$ be a molecule. Then $U$ has a maximal 0-dimensional element if and only if* $\dim U = 0$.

*Proof.* If $\dim U = 0$, then $U$ is the point by Lemma 4.4.1, hence has a greatest 0-dimensional element. Conversely, let $x$ be maximal and 0-dimensional, and let $V \coloneqq \mathrm{cl}\left((\mathscr{M}ax\, U) \setminus \{x\}\right)$. Then $\{x\}$ is closed, $U = V \cup \{x\}$, and $V \cap \{x\} = \varnothing$. By [25, Lemma 3.3.13], $V = \varnothing$, so $U = \{x\}$. ∎

***Example* 4.4.3 —** By Lemma 4.4.2, we cannot have a molecule as the oriented graded poset below, as it is not a composable configuration of cells:



**Lemma 4.4.4.** *Let $U$ be a 1-dimensional molecule, $m \coloneqq |U_1|$. Then $U$ is isomorphic to $m\vec{I} \coloneqq \underbrace{\vec{I} \#_0 \ldots \#_0 \vec{I}}_{m \text{ times}}$.*

*Proof.* By Lemma 3.2.7, either $\mathrm{lydim}\, U = -1$ or $\mathrm{lydim}\, U = 0$. In the first case, $U$ is an atom by Lemma 3.2.8. Because by Lemma 4.4.1 the point is the only 0-dimensional molecule up to isomorphism, the arrow is the only 1-dimensional atom, so $U$ is isomorphic to $\vec{I}$. In the second case, $U$ admits a 0-layering $(U^{(i)})_{i=1}^{m}$ by Theorem 3.2.11, and by Lemma 3.2.10, for each $i \in \{1, \ldots, m\}$, necessarily $\mathrm{lydim}\, U^{(i)} = -1$. By the first part, $U^{(i)}$ is isomorphic to $\vec{I}$. ∎

**Lemma 4.4.5.** *Let $U$ be a molecule,* $\dim U \leq 1$. *Then*

*I $U$ is round,*

*II $U$ is acyclic,*

*III $\vec{\mathscr{H}}U$ is a linear graph with $|U|$ vertices,*

*IV $\mathscr{F}_0 U$ is a linear graph with $|U_1|$ vertices,*

*V $U$ admits a unique 0-ordering.*

*Proof.* All straightforward checks using Lemma 4.4.1 and Lemma 4.4.4. ∎

**Proposition 4.4.6.** *Let $\imath \colon V \hookrightarrow U$ be an inclusion of* 1*-dimensional molecules. Then $\imath$ is a submolecule inclusion.*

*Proof.* By Lemma 2.3.21 $\mathscr{F}_0 V$ is an induced subgraph of $\mathscr{F}_0 U$. By Lemma 4.4.5 both of them are linear graphs, and an induced subgraph of a linear graph is a linear graph if and only if its vertices are consecutive in the ambient graph. All other conditions of Lemma 4.1.19 are trivially satisfied. ∎

**Lemma 4.4.7.** *Let $U, V$ be molecules and $k < \min\{\dim U, \dim V\}$ such that $U \#_k V$ is defined. If $U$ and $V$ are acyclic, then $U \#_k V$ is acyclic.*

*Proof.* See [25, Lemma 8.3.26]. ∎

**Lemma 4.4.8.** *Let $U$ be a 2-dimensional atom, $n \coloneqq |\Delta^- U|$, $m \coloneqq |\Delta^+ U|$. Then $U$ is isomorphic to $D_{n,m} \coloneqq n\vec{I} \Rightarrow m\vec{I}$.*

*Proof.* Immediate from Lemma 4.4.4. ∎

**Proposition 4.4.9.** *Let $U$ be a molecule, $\dim U \leq 2$. Then $U$ is acyclic.*

*Proof.* If $\dim U < 2$ this is part of Lemma 4.4.5, and if $U$ is a 2-dimensional atom it can be checked directly using Lemma 4.4.8. The statement then follows by an easy induction from Lemma 4.4.7. ∎

In order to prove the main result about molecules of dimension 2, we need to introduce *horizontal* and *vertical* orders.

***Definition* 4.4.10** (Horizontal and vertical order) — Let $U$ be a molecule, $\dim U \leq 2$. The *horizontal order* $\preceq_h$ and the *vertical order* $\preceq_v$ on the set $U_1$ of 1-dimensional elements of $U$ are defined by

- $x \preceq_h y$ if and only if there is a path from $x$ to $y$ in $\vec{\mathscr{H}}U$ only passing through elements of dimension 0 and 1,
- $x \preceq_v y$ if and only if there is a path from $x$ to $y$ in $\vec{\mathscr{H}}U$ only passing through elements of dimension 1 and 2.

**Lemma 4.4.11.** *Let $U$ be a molecule, $\dim U \leq 2$. Then*

*I the union of $\preceq_h$ and $\preceq_v$ is a linear order on $U_1$,*

*II the intersection of $\preceq_h$ and $\preceq_v$ is the identity relation on $U_1$.*

*Proof.* If $\dim U < 2$, then $\preceq_v$ is trivially the identity relation, and $\preceq_h$ is a linear order by Lemma 4.4.5. If $U$ is a 2-dimensional atom, then $U_1 = \Delta^- U + \Delta^+ U$, and that $\preceq_h$ is a linear order on $\Delta^\alpha U$ for each $\alpha \in \{+, -\}$ separately, so we have $x \preceq_v y$ for all $x \in \Delta^- U$ and $y \in \Delta^+ U$, and no other relations exist.
Otherwise, we proceed exactly as in the proof of Proposition 4.4.9, defining a 1-ordering $(x^{(i)})_{i=1}^m$ and a sequence $(V^{(i)})_{i=1}^m$ of increasing submolecules of $U$. We let $\preceq_h^{(i)}$ and $\preceq_v^{(i)}$ be the orders determined by paths in $\vec{\mathscr{H}}V^{(i)}$, which are increasing in $i$, and proceed by induction. Since $\dim V^{(0)} = 1$, we have already proved the base case.
Let $i > 0$, assume that the statement holds of the orders $\preceq_h^{(i-1)}$ and $\preceq_v^{(i-1)}$ on $(V^{(i-1)})_1$. Let $x, y \in (V^{(i)})_1$; we will show that $x$ and $y$ are comparable via $\preceq_h^{(i)}$ or $\preceq_v^{(i)}$. If $x, y \in V^{(i-1)}$ or $x, y \in \operatorname{cl}\left\{x^{(i)}\right\}$, we can apply the inductive hypothesis or the atom case, so it suffices to consider the case

$$x \in V^{(i-1)} \setminus \Delta^- x^{(i)}, \quad y \in \Delta^+ x^{(i)}.$$

Let $(z^{(j)})_{j=1}^p$ be the unique linear 0-ordering on $\partial^- x^{(i)}$, so $z^{(j)} \preceq_h z^{(j')}$ if $j \leq j'$. For all $j \in \{1, \dots, p\}$, we have $x \neq z^{(j)}$, and by the inductive hypothesis $x$ and $z^{(j)}$ are comparable via $\preceq_h^{(i-1)}$ or they are comparable via $\preceq_v^{(i-1)}$.

- Suppose there exists $j$ such that $x$ and $z^{(j)}$ are comparable via $\preceq_v^{(i-1)}$. Then necessarily $x \preceq_v^{(i-1)} z^{(j)}$, because $\Delta^- x^{(i)} \subseteq \Delta^+ V^{(i-1)}$. Since $z^{(j)} \preceq_v^{(i)} y$, we have $x \preceq_v^{(i)} y$.

- Otherwise, $x$ and $z^{(j)}$ are comparable via $\preceq_h^{(i-1)}$ for all $j \in \{1, \dots, p\}$. Suppose that $x \preceq_h^{(i-1)} z^{(1)}$, in which case $x \preceq_h^{(i-1)} z^{(j)}$ for all $j \in \{1, \dots, p\}$. Then the path from $x$ to $z^{(1)}$ through elements of dimension 0 and 1 must enter $z^{(1)}$ from $\partial^- z^{(1)} = \partial_0^- x^{(i)}$. Since there is a path in $\vec{\mathscr{H}} V_i$ from $\partial_0^- x^{(i)}$ to $y$, we have $x \preceq_h^{(i)} y$.

- Otherwise, there is a greatest $j$ such that $z^{(j)} \preceq_h^{(i-1)} x$. If $j < p$, then $z^{(j)} \preceq_h^{(i-1)} x \preceq_h^{(i-1)} z^{(j+1)}$. Because all three are distinct, letting $\partial^+ z^{(j)} = \partial^- z^{(j+1)} = \{w\}$, there is a non-trivial cycle in $\vec{\mathscr{H}} V^{(i-1)}$ from $w$ to $x$ and back to $w$, a contradiction to Proposition 4.4.9. It follows that $z^{(p)} \preceq_h^{(i-1)} x$, and the path between the two must leave $z^{(p)}$ through $\partial_0^+ x^{(i)}$, so $y \preceq_h^{(i)} x$.

This proves that the union of $\preceq_h^{(i)}$ and $\preceq_v^{(i)}$ is a linear order on $(V^{(i)})_1$.
Suppose that $x \preceq_h^{(i)} y$ and $x \preceq_v^{(i)} y$; we will prove that $x = y$. If $x \in \operatorname{cl}\left\{x^{(i)}\right\}$, then $x \preceq_v^{(i)} y$ implies that $y \in \operatorname{cl}\left\{x^{(i)}\right\}$, and any path from $x$ to $y$ in $\vec{\mathscr{H}} V^{(i)}$ is entirely contained in $\operatorname{cl}\left\{x^{(i)}\right\}$, so we can apply the atom case. Suppose that $x \notin \operatorname{cl}\left\{x^{(i)}\right\}$. If $y \in \Delta^+ x^{(i)}$, then any path from $x$ to $y$ through elements of dimension 1 and 2 consists of a path contained in $V^{(i-1)}$ to some $z \in \Delta^- x^{(i)}$ followed by the path $z \to x^{(i)} \to y$; while any path through elements of dimension 0 and 1 consists of a path contained in $V^{(i-1)}$ to $\partial_0^- x^{(i)}$ followed by a path contained in $\partial^+ x^{(i)}$. Since there is a path from $\partial_0^- x^{(i)}$ to any $z \in \Delta^- x^{(i)}$ through elements of dimension 0 and 1 in $V^{(i-1)}$, we have that $x \preceq_h^{(i-1)} z$ and $x \preceq_v^{(i-1)} z$ for some $z \in \Delta^- x^{(i)}$. By the inductive hypothesis, $x = z$, a contradiction since $z \in \operatorname{cl}\left\{x^{(i)}\right\}$.
It follows that $y \in V^{(i-1)}$. Then any path from $x$ to $y$ through elements of dimension 1 and 2 is entirely contained in $V^{(i-1)}$, so $x \preceq_v^{(i-1)} y$; while a path through elements of dimension 0 and 1 is either contained in $V^{(i-1)}$, or it enters $\partial^+ x^{(i)}$ through $\partial_0^- x^{(i)}$, traverses it in its entirety, and leaves from $\partial_0^+ x^{(i)}$. Such a path segment can be replaced with the one that traverses $\partial^- x^{(i)}$ in its entirety, so in either case $x \preceq_h^{(i-1)} y$, and we conclude that $x = y$ by the inductive hypothesis. This concludes the proof of the statement for $V^{(i)}$. Since $V^{(m)} = U$, we conclude. ∎

**Lemma 4.4.12.** *Let $\imath \colon V \hookrightarrow U$ be an inclusion of molecules, $\dim U \leq 2$. Then $\mathscr{F}_1 V$ is a path-induced subgraph of $\mathscr{F}_1 U$.*

*Proof.* Both $U$ and $\imath(V)$ are molecules of dimension $\leq 2$. Let $\preceq_h$, $\preceq_v$ be the horizontal and vertical order on $U$, and $\preceq_h^V$, $\preceq_v^V$ those on $\imath(V)$, which are subsets of those on $U$.
Suppose by way of contradiction that $\mathscr{F}_1 V$ is not path-induced. Then there exists a path $x_0 \to \ldots \to x_m$ in $\mathscr{F}_1 U$ such that $m > 1$, $x_0, x_m \in \imath(V)$, and $x_i \notin \imath(V)$ for all $i \in \{1, \dots, m-1\}$. By definition, there exist 1-dimensional elements $y_i \in \Delta^+ x_{i-1} \cap \Delta^- x_i$ for all $i \in \{1, \dots, m\}$. Then $y_1 \preceq_v y_m$ and $y_1 \neq y_m$. By Corollary

2.3.9, $x_{i-1}$ and $x_i$ are the only cofaces of $y_i$ for all $i \in \{1, \dots, m\}$. Necessarily, then, $y_1 \in \Delta^+ \imath(V)$ and $y_m \in \Delta^- \imath(V)$, so it is not possible that $y_1 \preceq_v^V y_m$. Then by Lemma 4.4.11 applied to $\imath(V)$, one of $y_1 \preceq_h^V y_m$, $y_m \preceq_h^V y_1$, or $y_m \preceq_v^V y_1$ must hold. Combined with $y_1 \preceq_v y_m$, each one of these implies $y_1 = y_m$ by Lemma 4.4.11 applied to $U$, a contradiction. ∎

Finally, we can state and prove the result for which we presented the intuition when discussing the molecule matching, namely that any inclusion of molecules $V \hookrightarrow U$, where the dimension of $U$ and $V$ is 2 and $V$ is round, is a submolecule inclusion.

**Theorem 4.4.13.** *Let $\imath \colon V \hookrightarrow U$ be an inclusion of molecules such that $\dim U = \dim V = 2$ and $V$ is round. Then $\imath$ is a submolecule inclusion.*

*Proof.* By Lemma 4.4.12 combined with Lemma 4.1.14, there exists a 1-ordering $(x^{(i)})_{i=1}^m$ of $U$ in which the elements of $\imath(V)$ are consecutive, that is, $x^{(i)} \in \imath(V)$ if and only if $p \le i \le q$ for some $p, q \in \{1, \dots, m\}$.
By Proposition 4.4.9 $U$ is acyclic, so by Lemma 4.3.15 it is frame-acyclic, and by Corollary 4.3.6 the 1-ordering comes from a 1-layering $(U^{(i)})_{i=1}^m$ such that $\imath(\partial^- V) \subseteq \partial^- U^{(p)}$. Since both are 1-dimensional molecules, by Proposition 4.4.6 $\imath(\partial^- V) \sqsubseteq \partial^- U^{(p)}$, and we conclude by Lemma 4.1.19. ∎

**Corollary 4.4.14.** *The rewritable submolecule problem in dimension $\le 2$ has a trivial constant-time solution.*

**Lemma 4.4.15.** *Let $U$ be a molecule, $\dim U \le 3$. Then $U$ is frame-acyclic.*

*Proof.* See [25, Theorem 8.4.11]. ∎

**Theorem 4.4.16.** *Let $U$ be a molecule, $\dim U \le 3$. Then $U$ is stably frame-acyclic.*

*Proof.* By Lemma 4.4.15 $U$ is frame-acyclic. Since for all $V \sqsubseteq U$ and all rewritable $W \sqsubseteq V$, $V[\langle W \rangle / W]$ is still a molecule of dimension $\le 3$, it is frame-acyclic. Hence $U$ is stably frame-acyclic. ∎

**Theorem 4.4.17.** *Let $\imath \colon V \hookrightarrow U$ be an inclusion of molecules such that $\dim U = \dim V = 3$ and $V$ is round. The following are equivalent:*

*(a) $\imath$ is a submolecule inclusion;*

*(b) $\mathscr{F}_2 V$ is a path-induced subgraph of $\mathscr{F}_2 U$.*

*Proof.* One implication is Proposition 4.1.18, so we only need to prove the converse. By Lemma 4.1.14, if $\mathscr{F}_2 V$ is path-induced, then there exists a 2-ordering $(x^{(i)})_{i=1}^m$ of $U$ in which the elements of $\imath(V)$ are consecutive, that is, $x^{(i)} \in \imath(V)$ if and only if $p \le i \le q$ for some $p, q \in \{1, \dots, m\}$.
By Lemma 4.4.15, $U$ is frame-acyclic, so by Corollary 4.3.6 the 2-ordering comes from a 2-layering $(U^{(i)})_{i=1}^m$ such that $\imath(\partial^- V) \subseteq \partial^- U^{(p)}$. Since both are 2-dimensional molecules and $\partial^- V$ is round, we have $\imath(\partial^- V) \sqsubseteq \partial^- U^{(p)}$ by Theorem 4.4.13, and we conclude by Lemma 4.1.19. ∎

**Theorem 4.4.18.** *The rewritable submolecule problem in dimension 3 can be solved in time $O(|\mathscr{E}_3 U|)$.*

*Proof.* By Theorem 4.4.17 combined with Lemma 4.1.14, it suffices to construct $\mathscr{F}_2 U / \mathscr{F}_2 V$ and check if it is acyclic. The first can be done while traversing the induced subgraph of $\mathscr{H} U$ on $U_3 \cup U_2$, which takes time $O(|\mathscr{E}_3 U|)$. Both the number of vertices and the number of edges of $\mathscr{F}_2 U / \mathscr{F}_2 V$ is also $O(|\mathscr{E}_3 U|)$, and we conclude. ∎

To match a diagram $s \colon V \to \mathbb{V}$ in $t \colon U \to \mathbb{V}$ in dimension 3, according to our results we need

- $O(|U_3|\ |V_3|\ |V|\ |\mathscr{E}_\vee V|\ \log |\mathscr{E}_\vee V|)$ time to find all inclusions $V \hookrightarrow U$, of which there are $O(|U_3|)$,
- $O(|\mathscr{E}_3 U|)$ time to check whether each of them is a submolecule inclusion,
- $O(|V|)$ time to compare labellings on each, assuming labels can be compared in constant time,

leading to an overall

$$O(|U_3|\,(|\mathscr{E}_3 U| + |V_3|\ |V|\ |\mathscr{E}_\vee V|\ \log |\mathscr{E}_\vee V|))$$

upper bound.

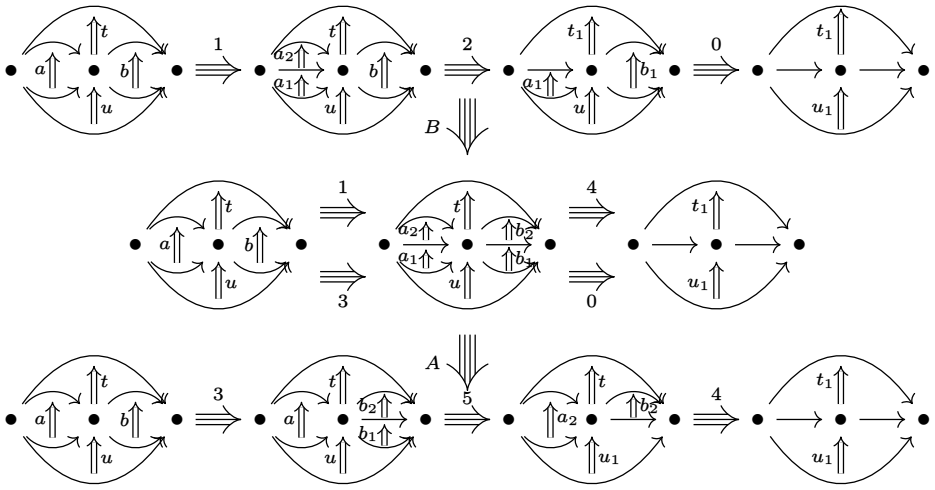
Here we used the bound on molecule matching in generic dimension; it is possible that this can be improved by using strategies tailored to dimension 3, as it is certainly the case in dimension $\leq 2$.

If we consider a machine operating by rewriting 3-dimensional diagrams, which has a fixed finite list of rewrite rules, the variables linked to $V$ can be considered as constant parameters of the machine. Our results then imply that such a machine can be simulated with $O(|U_3|\ |\mathscr{E}_3 U|)$ overhead in a standard model of computation.

It is still an open question whether the rewritable submolecule problem in dimension greater than 3 is NP-complete.

Note that even though the molecule in Example 4.3.3 is not frame-acyclic, its 3-orderings and 3-layerings are in bijection. Below, we provide an example of a 4-dimensional molecule for which, even though $\mathscr{F}_3 U$ is a discrete graph with two 3-orderings, only one ordering is induced by a layering. This behaviour gives rise

to the notion of *obstruction* which will be the topic of study for the next section.



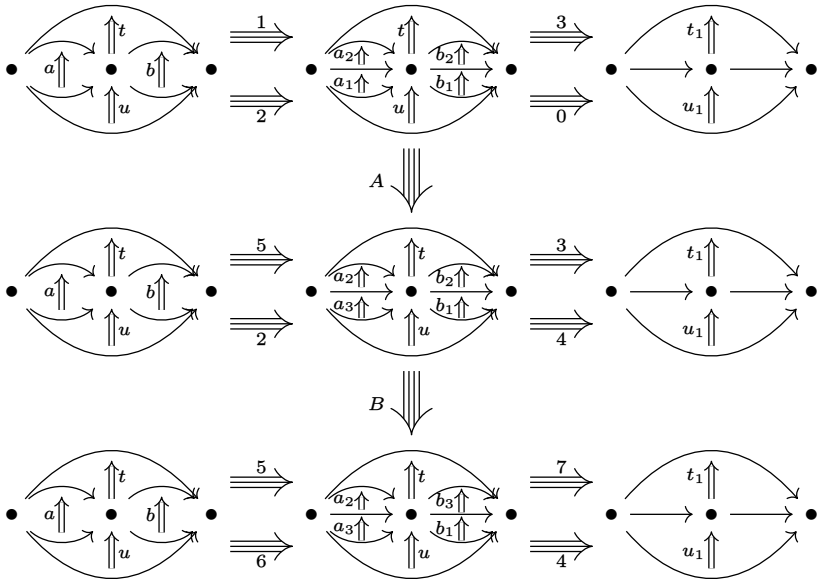

Even though the 3-flow graph of $U$ is $\mathscr{F}_3 U =$ $A\bullet$ $\bullet B$, only the ordering $[B, A]$ comes from a valid layering because the input of $A$, $\partial^- A = \mathrm{cl}\{1, 0\}$ is not a submolecule of $\partial^- U$. We say that $b \in \Delta^- B$ obstructs $A$ as it acts as an obstacle to $A$.

## 4.5 Obstructions

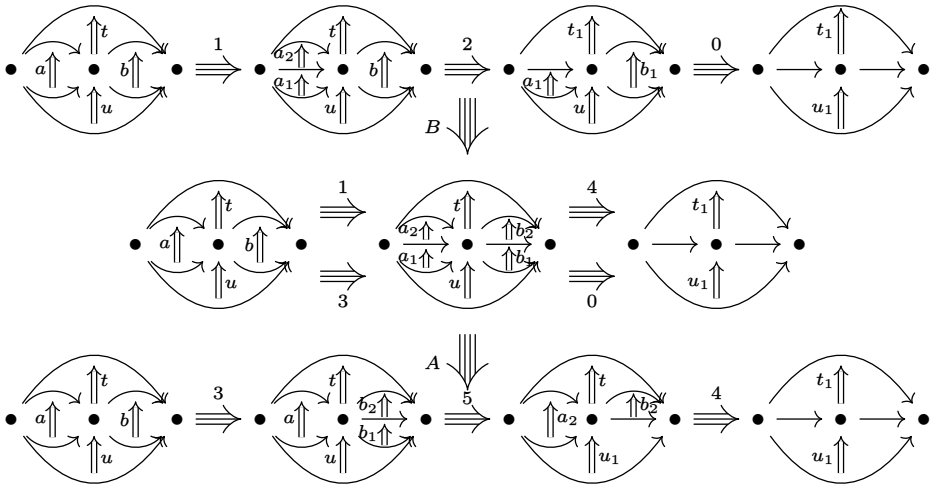
The study of obstructions in higher-dimensional diagrams stemmed from the study of the computational complexity of the subdiagram matching problem. In Chapter 4 we noticed that under the condition of stably frame-acyclicity the runtime of the algorithm above is greatly improved - i.e., linear instead of factorial.

The work in this chapter starts from the observation made in the following example:

***Example* 4.5.1 —** Let $U_1$ be:




and let $U_2$ be




The first diagram is not frame-acyclic, however any ordering of its 3-flow graphs is induced by a layering, so finding a valid layering is "easy" in this case. The second diagram contains what we will call an *obstruction*. By looking at the string diagram representation of $\partial^- U_2$ in Figure 4.1 one can see that in order to produce the 3-dimensional input of $A$, we need an element that is part of the input of $B$; in turn, this element is dependent on the 2-dimensional output from element 1 which is in the input of $A$. We obtain some kind of "circularity": the 2-input of 0 depends on the 2-output of 2 whose 2-input depends on the 2-output of 1. We call this a

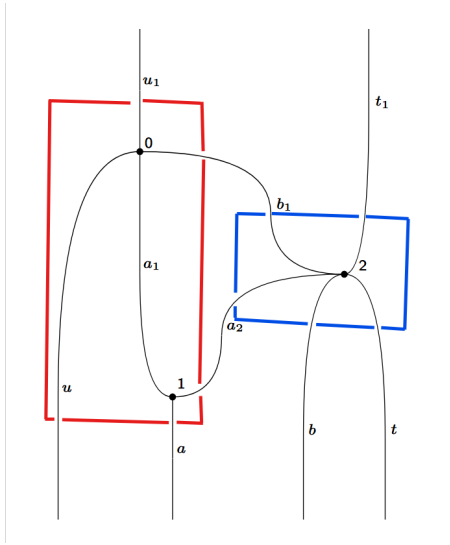



*Figure 4.1: Input boundary of $U$.*

dependency of 0 on 1 that goes through 2 or a dependency in the input of $A$ going through $2 \in \Delta^{\alpha} B$.
The difference between an obstruction and a molecule not being frame-acyclic is that the dependency in convex but not frame-acyclic molecules is disjoint which can be seen in the string diagram representation from Example 4.3.3.
When situations like the one described above arise we say that $B$ obstructs $A$. As argued in the previous sections we would like to find an ordering that is induced by a layering. It is clear that rewrite $A$ cannot appear first in a layering since its input is not a submolecule of $U$. We hope that by applying $B$, we can "parallelise" the 2-dimensional faces causing an obstruction and get rid of the dependency of 0 and 1 on 2. We will call this technique "solving a non-convexity or obstruction". Note that in the example above, even if $\mathscr{F}_3 U$ is a discrete graph, with two orderings, we only have one valid layering which induces the ordering: $B, A$.

One of our goals was to come up with a definition that captures the intuition presented above, while containing minimal constraints. In the remainder of the section, we provide the definition for the basic case of a molecule with an obstruction: for an $n$-dimensional molecule $U$, there are only two $n$-dimensional elements involved in the obstruction called $A$ and $B$ and the obstruction passes only through one element $b \in \Delta^{\alpha} B$. We prove that the molecules of dimension lower or equal than three have no obstructions and present a result in dimension 4 that covers the intuition presented in the example above.
We discuss the potential generalisations in the Future work chapter.

***Definition* 4.5.2 —** (Restricted flow graph) Let $U$ be a molecule, $k \geq -1$, $k \leq n$. The $n$-restricted $k$-flow graph of $U$ is the induced subgraph $\mathscr{F}_k^n U$ of $\mathscr{F}_k U$ on the

vertex set $U_n \subseteq \bigcup_{i>k} U_i$.

At the moment, we will work with a simple definition, that for example only allow obstructions that pass through exactly one element. We will think of possible generalisations later on.

***Definition* 4.5.3 —** Let $U$ be an $n$-dimensional molecule and let $A, B \in U_n$. $U$ has an obstruction $(\Delta_{n-1}A, B)$, if there exists a path $a_1 \to b \to a_2$ in $\mathscr{F}_k^{n-1}U$ , where $a_1, a_2 \in \Delta^- A$, $b \in \Delta^\alpha B$ and $k = \mathrm{frdim}(A, B)$.

***Definition* 4.5.4 —** A molecule is *convex* if it has no obstructions.

The following proofs are meant to capture the intuition that we already have about non-convexities: it is impossible for a molecule of dimension less than 3 to have an obstruction. The intuition behind this lies in the following fact: assume we have such an obstruction. The obstruction cannot pass through the $(n-1)$-dimensional elements, otherwise we would get a contradiction with the flow graph being acyclic. So the obstructions can only pass through elements of dimension 0 or 1. These elements are not parallelisable and hence we would obtain a shape that is not a molecule (again a contradiction).

**Lemma 4.5.5.** *Let $U$ be a molecule,* $\dim(U) \leq 1$. *Then $U$ is convex.*

*Proof.* If $\dim(U) = 0$, then $U$ is a point, $\vec{\mathscr{H}}U$ is also a point and there is no path in $\vec{\mathscr{H}}U$. So $U$ is convex.
Now let $\dim(U) = 1$ and suppose $U$ has an obstruction $(\Delta^- A, B)$, with $A, B \in U_1$. Then, there should be a path that is not a cycle in $\vec{\mathscr{H}}U$ starting and ending in $\Delta_0^- A$. This is impossible since $\left|\Delta_0^- A\right| = 1$ and $\vec{\mathscr{H}}U$ is linear (and acyclic). ∎

To illustrate the technique used in the proofs, when $U$-is 2 and 3-dimensional, we will prove an easier case before proving the general one.

**Lemma 4.5.6.** *Let $U$ be a molecule,* $\dim(U) = 2$, $U_2 = \{A, B\}$. *Then $U$ is convex.*

*Proof.* Case $\mathrm{frdim}(U) = 0$: then $\mathscr{F}_0^1 U$ is: $1 \in \Delta^- A \to 2 \in \Delta^\alpha B \to 0 \in \Delta^- A$ which leads to a cycle $A \to B \to A$ in $\mathscr{F}_0 U$. This would mean that $U$ is not dimension-wise acyclic, which is a contradiction with Proposition 4.4.9 (every 2-dim molecule is acyclic). We conclude by Proposition 4.3.15.
Case $\mathrm{frdim}(U) = 1$: This is impossible from the definition since $\mathscr{F}_1 U$ is acyclic. ∎

More generally, one proves:

**Lemma 4.5.7.** *Let $U$ be a molecule,* $\dim(U) = 2$. *Then $U$ is convex.*

*Proof.* Suppose $U$ has an obstruction $(\Delta^- A, B)$.
If $\mathrm{frdim}(\{A, B\}) = 1$, then we obtain a contradiction with $\mathscr{F}_1 U$ being acyclic. If $\mathrm{frdim}(\{A, B\}) = 0$, then $\mathscr{F}_0 U$ has a cycle $A \to B \to A$ which is a contradiction with $U$ being dimension-wise acyclic. ∎

**Lemma 4.5.8.** *Let $U$ be a molecule,* $\dim(U) = 3$, $U_3 = \{A, B\}$ *and $U$ is pure. Then $U$ is convex.*

*Proof.* Suppose that $U$ has an obstruction $(\Delta^- A, B)$.
Case $\operatorname{frdim}(U) = 2$ : then we would get a cycle $A \to B \to A$ in $\mathscr{F}_2 U$. Contradiction.
Case $\operatorname{frdim}(U) \leq 1$ : then we would get a cycle $A \to B \to A$ in $\mathscr{M}_{\operatorname{frdim}(U)} U$. Contradiction with $U$ being stably frame-acyclic by Theorem 4.4.16. ∎

More generally, one proves:

**Lemma 4.5.9.** *Let $U$ be a molecule, $\dim(U) = 3$ and $U$ is pure. Then $U$ is convex.*

*Proof.* Suppose $U$ has an obstruction $(\Delta^- A, B)$, where $A, B \in U_3$.
Case $\operatorname{frdim}(\{A, B\}) = 2$ : then we would get a cycle $A \to B \to A$ in $\mathscr{F}_2 U$. Contradiction.
Case $\operatorname{frdim}(\{A, B\}) \leq 1$. The trick we'll employ is the following: $A, B$ are not connected in $\mathscr{F}_2 U$, so there exists a 2-layering of $U$ which is induced by an ordering in which $A$ and $B$ are consecutive. Then, $\operatorname{cl}\{A, B\}$ is a submolecule and we will show $\mathscr{M}_1(\operatorname{cl}\{A, B\})$ has a cycle, contradicting the fact that $U$ is frame-acyclic.
Since $U$ is 3-dimensional and frame-acyclic, the 2-orderings and layerings are in bijection by Corollary 4.3.6. Also $\operatorname{frdim}(A, B) = 1$ which means they are not connected in $\mathscr{F}_2 U$. Hence, there exists a layering that induces an ordering in which $A$ and $B$ are consecutive. Without loss of generality, consider the layering $U = U_1 \#_2 \ldots \#_2 U_A \#_2 U_B \#_2 \ldots \#_2 U_m$ and let $U_{AB} = U_A \#_2 U_B$. This is well defined since it is part of a valid layering of $U$ and by definition, it is a submolecule of $U$ with $\operatorname{frdim}(U_{AB}) = 1$. Since $\operatorname{frdim}(A, B) = 1$ and there is an obstruction between $A$ and $B$, we obtain a cycle in $\mathscr{M}_1(U_{A,B})$ which by Lemma 4.4.15 is a contradiction with $U$ being frame-acyclic. ∎

As it looks right now, it seems that a molecule that has an obstruction is not frame-acyclic. There is a choice of whether to make that part of the definition or not. In the first case, the results up to this point would be trivial, since any molecule of dimension $\leq 3$ is frame-acyclic. However, we chose a more general definition that has the potential to be generalised even further to oriented graded posets.
The following lemma is meant to prove the intuition that was explained at the beginning of the section. We show that given a molecule $U$ with an obstruction, $(\Delta^- A, B)$, if $b$ is in the input boundary of $B$, then in any layering of $U$, $B$ comes before $A$ and if $b \in \Delta^+ B$, then $A$ comes before $B$ in any layering of $U$. We can rephrase the statement above as follows:

- if $b \in \Delta^- B$, then $b$ causes an obstruction in the input of $A$. In other words, $A$ cannot happen before $B$ because its input is not submolecule. By doing the rewrite $B$ first, the obstruction is "solved" and we can then apply rewrite $A$.
- if $b \in \Delta^+ B$: the output of $B$ causes an obstruction in the input of $A$. To avoid this, rewrite $A$ must happen before the rewrite $B$.

At the moment, we prove this for a pure 4-dimensional molecule with only two elements.

**Lemma 4.5.10.** *Let $U$ be a 4-dimensional molecule, $U_4 = \{A, B\}$, $U$ is pure, $\operatorname{frdim} U \geq 2$ and suppose $U$ has an obstruction $(\Delta^- A, B)$. Then:*

- *if* $b \in \Delta^- B$, *then* $\partial^- A \not\sqsubseteq \partial^- U$ *and* $U$ *admits the layering* $U_B \#_{n-1} U_A$.
- *if* $b \in \Delta^+ B$, *then* $\partial^- A \not\sqsubseteq \partial^+ U_B$ *and* $U$ *admits the layering* $U_A \#_{n-1} U_B$.

*Proof.* We start with the proofs for $b \in \Delta^- B$.
Case $\mathrm{frdim}(U) = 3$. Then $U$ admits the layering induced by the 3-ordering.
Case $\mathrm{frdim}(U) = 2$. Then there is a path $a_1 \to x \to b \to y \to a_2$ in $\vec{\mathscr{H}} U$, with $a_1, a_2 \in \Delta^- A$. Then $\mathscr{M}_2 U$ has a cycle $A \xrightarrow{x} B \xrightarrow{y} A$. Let $b \in \Delta^- B$. Since $\mathrm{frdim}\, U = 2$, $b \notin Int(U)$, so $b \in \partial^- U$. Similarly for the elements in $\Delta^- A$. Consider the molecule $\partial^- U$. By Remark 2.3.16 we know that $\mathscr{M}_2 \partial^- U = \mathscr{F}_2 \partial^- U$. Again, since $\mathrm{frdim}\, U = 2$, these graphs contain $b$ and all the elements in $\Delta^- A$. In particular, from the obstruction, we get the following subgraph in $\mathscr{M}_2 \partial^- U$: $a_1 \to b \to a_2$ from which we deduce there is no valid ordering in which the elements of $\Delta^- A$ are consecutive. Thus, by Theorem 4.4.18, $\partial^- A \not\sqsubseteq \partial^- U$ and the only layering available for $U$ is $U_B \#_3 U_A$. The case $b \in \Delta^+ B$ and showing that $\partial^- A \not\sqsubseteq U[\partial^- B / \partial^+ B]$ is symmetrical.

■

# Chapter 5

# Categorical framework

The theory presented in this thesis so far is of very combinatorial flavour. However, given the topological underlying structure, it is very natural to construct a categorical framework. As alluded in the introduction of the thesis, our notion of molecules is intimately related to the theory of pasting diagrams in strict $\omega$-categories.

In this chapter we turn our attention to regular directed complexes, a more general notion of diagrams which are oriented graded posets where the closure of each element is an atom. We continue our study to answer the remaining two questions from the introduction: subdiagrams form an $n$-category and that the presented $n$-category is free in the sense of a polygraph. We show that the first of these two can be achieved by considering suitable subdiagrams of regular directed complexes. For the second property, we lift the notions of acyclity developed in the previous chapter to the framework of regular directed complexes and show that when a regular directed complex is frame-acyclic, the $\omega$-category obtained from it is freely-generated.

As already shown in the previous chapter for molecules, it remains true for regular directed complexes that frame-acyclicity is difficult to check in practice. However, as it was the case before, this condition is implied by stronger acyclicity conditions. We study these conditions in relation to Steiner's theory on augmented directed complexes [50]. We show that our notion of dimension-wise acyclic regular directed complexes translates to augmented directed complexes with *loop-free* basis, called a *Steiner complex*. We prove that when a regular directed complex, $P$, is frame-acyclic, the $\omega$- category of "subdiagrams" of $P$ $Mol/P$ is isomorphic to the $\omega$-category obtained by applying Steiner's $\nu$ functor to the Steiner complex obtained from $P$. Moreover, when $P$ is strongly dimension-wise acyclic - an even stronger acyclicity condition that implies dimension-wise acyclicity - then the $\omega$-category $Mol/P$ consists only of "subsets" of $P$. Finally, when $P$ is acyclic, the corresponding augmented directed complex has a strongly loop-free basis. We end the chapter by studying the stability of the acyclicity conditions mentioned above under the operations of pasting, suspensions, Gray products and joins.

The results of this chapter are part of the article [27] and appear with slightly different proofs in [25, Chapter 8 and Chapter 11].

## 5.1 The category $Mol/P$

***Definition* 5.1.1** (Regular directed complex) — A *regular directed complex* is an oriented graded poset $P$ with the property that, for all $x \in P$, the closed subset $\mathrm{cl}\{x\}$ is an atom. We write **RDC** for the full subcategory of **ogPos** whose objects are regular directed complexes.

**Remark 5.1.2.** *While superficially different, this is equivalent to the definition of regular directed complex given in [26]. On the other hand, our definition of* molecule *here corresponds to a* molecule in a regular directed complex*, or* molecule*, and is more restrictive than the definition in [51]. Note that by Lemma 2.2.30, every molecule is a regular directed complex. /*

Before defining the functor from oriented graded posets to $\omega$-categories, let us describe an important property that all regular directed complexes satisfy.

***Definition* 5.1.3** (Positive least element) — Let $P$ be an oriented graded poset, $\bot \in P$. We say that $\bot$ is a *positive least element of* $P$ if $\bot$ is the least element of $P$ and $\nabla\bot = \nabla^+\bot$. We let $\mathbf{ogPos}^+$ denote the full subcategory of **ogPos** on oriented graded posets with a positive least element.

Freely adjoining a positive least element and, respectively, deleting the least element exhibit an equivalence between **ogPos** and $\mathbf{ogPos}^+$.

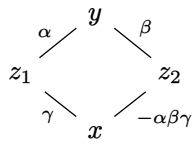
**Proposition 5.1.4.** *There exists a pair of functors*

$$(-)_\bot : \mathbf{ogPos} \to \mathbf{ogPos}^+, \qquad (-)_{\not\bot} : \mathbf{ogPos}^+ \to \mathbf{ogPos}$$

*inverse to each other up to natural isomorphism.*

The following is a useful property of regular directed complexes.

***Definition* 5.1.5** (Oriented thin graded poset) — Let $P$ be an oriented graded poset with a positive least element. We say that $P$ is *oriented thin* if, for all $x, y \in P$ such that $x \leq y$ and $\mathrm{codim}_y(x) = 2$, the interval $[x, y]$ is of the form




for exactly two elements $z_1, z_2$, and for some $\alpha, \beta, \gamma \in \{+, -\}$.

**Proposition 5.1.6.** *Let $P$ be a regular directed complex. Then $P_\bot$ is an oriented thin graded poset.*

*Proof.* See [25, Proposition 5.3.4] ∎

The connection between oriented graded posets and strict $\omega$-categories is given by the fact that (isomorphism classes of) molecules form a strict $\omega$-category with pasting at the $k$-boundary as $k$-composition. The fibred version of this result

implies that (isomorphism classes of) molecules *over an oriented graded poset* $P$ form a strict $\omega$-category.
In what follows, we recall the *single-set* definition of strict $\omega$-category, which is most natural in this context, and state these results more precisely.

***Definition* 5.1.7** (Reflexive $\omega$-graph) — A *reflexive $\omega$-graph* is a set $X$, whose elements are called *cells*, together with, for all $n \in \mathbb{N}$, operators $\partial_n^-, \partial_n^+ \colon X \to X$ called *input* and *output $n$-boundary*, satisfying the following axioms.

I (*Finite dimension*). For all $t \in X$, there exists $n \in \mathbb{N}$ such that

$$\partial_n^- t = \partial_n^+ t = t.$$

II (*Globularity*). For all $t \in X$, $k, n \in \mathbb{N}$, and $\alpha, \beta \in \{+, -\}$,

$$\partial_k^\alpha(\partial_n^\beta t) = \begin{cases} \partial_k^\alpha t & \text{if } k < n, \\ \partial_n^\beta t & \text{if } k \geq n. \end{cases}$$

If $t$ is a cell in a reflexive $\omega$-graph, the *dimension of* $t$ is the natural number $\dim t \coloneqq \min\{n \in \mathbb{N} \mid \partial_n^- t = \partial_n^+ t = t\}$.

***Definition* 5.1.8** (Composable pair of cells) — Let $t, u$ be a pair of cells in a reflexive $\omega$-graph, $k \in \mathbb{N}$. We say that $t$ and $u$ are *$k$-composable* if $\partial_k^+ t = \partial_k^- u$. We write

$$X \times_k X \coloneqq \left\{(t, u) \in X \times X \mid \partial_k^+ t = \partial_k^- u\right\}.$$

for the set of $k$-composable pairs of cells in $X$.

***Definition* 5.1.9** (Strict $\omega$-category) — A *strict $\omega$-category* is a reflexive $\omega$-graph $X$ together with, for all $k \in \mathbb{N}$, an operation $- \#_k - \colon X \times_k X \to X$ called *$k$-composition*, satisfying the following axioms.

I (*Compatibility with boundaries*). For all $k$-composable pairs of cells $t, u$, all $n \in \mathbb{N}$, and $\alpha \in \{+, -\}$,

$$\partial_n^\alpha(t \#_k u) = \begin{cases} \partial_n^\alpha t = \partial_n^\alpha u & \text{if } n < k, \\ \partial_k^- t & \text{if } n = k,\ \alpha = -, \\ \partial_k^+ u & \text{if } n = k,\ \alpha = +, \\ \partial_n^\alpha t \#_k \partial_n^\alpha u & \text{if } n > k. \end{cases}$$

II (*Associativity*). For all cells $t, u, v$ such that either side of the equation is defined, $(t \#_k u) \#_k v = t \#_k (u \#_k v)$.

III (*Unitality*). For all cells $t$, $t \#_k \partial_k^+ t = \partial_k^- t \#_k t = t$.

IV (*Interchange*). For all cells $t, t', u, u'$ and $n > k$ such that the left-hand side is defined, $(t \#_n t') \#_k (u \#_n u') = (t \#_k u) \#_n (t' \#_k u')$.

Given a strict $\omega$-category $X$ and $n \in \mathbb{N}$, we let $\sigma_{\leq n} X$ denote its *$n$-skeleton*, that is, its restriction to cells of dimension $\leq n$. A strict $\omega$-category is a *strict $n$-category* if it is equal to its $n$-skeleton.

***Definition* 5.1.10** (Strict functor of strict $\omega$-categories) — Let $X, Y$ be strict $\omega$-categories. A *strict functor* $f\colon X \to Y$ is a function such that, for all $k, n \in \mathbb{N}$, $\alpha \in \{+,-\}$, and $k$-composable cells $t, u$ in $X$,

$$f(\partial_n^\alpha t) = \partial_n^\alpha f(t), \qquad f(t \#_k u) = f(t) \#_k f(u).$$

Strict $\omega$-categories and strict functors form a category $\omega\mathbf{Cat}$.

***Definition* 5.1.11** (Generating sets and bases) — Let $X$ be a strict $\omega$-category and $\mathscr{S}$ a set of cells in $X$. The set $\operatorname{span}\mathscr{S}$ is the smallest set such that

I if $t \in \mathscr{S}$, then $t \in \operatorname{span}\mathscr{S}$,

II for all $k \in \mathbb{N}$, if $t, u \in \operatorname{span}\mathscr{S}$ are $k$-composable, then $t \#_k u \in \operatorname{span}\mathscr{S}$.

A *generating set for* $X$ is a set $\mathscr{S}$ of cells such that $\operatorname{span}\mathscr{S}$ contains every cell in $X$. A *basis for* $X$ is a minimal generating set.

**Lemma 5.1.12.** *Let $f, g\colon X \to Y$ be strict functors and let $\mathscr{S}$ be a generating set for $X$. If $f(t) = g(t)$ for all $t \in \mathscr{S}$, then $f = g$.*

***Definition* 5.1.13** (Isomorphism classes of molecules) — For each oriented graded poset $P$, let $[P]$ denote its isomorphism class in $\mathbf{ogPos}$. We let

$$\begin{aligned} \mathit{Mol} &\coloneqq \{[U] \mid U \text{ is a molecule}\}, \\ \mathit{Atom} &\coloneqq \{[U] \mid U \text{ is an atom}\} \subset \mathit{Mol}. \end{aligned}$$

Note that we can use the traversal algorithm (Procedure 2.4.2) to choose a representative for each isomorphism class of molecules.

**Proposition 5.1.14.** *For all $n, k \in \mathbb{N}$ and $\alpha \in \{+,-\}$, let*

$$\begin{aligned} \partial_n^\alpha &\colon \mathit{Mol} \to \mathit{Mol}, & [U] &\mapsto [\partial_n^\alpha U], \\ - \#_k - &\colon \mathit{Mol} \times_k \mathit{Mol} \to \mathit{Mol}, & [U], [V] &\mapsto [U \#_k V]. \end{aligned}$$

*Then Mol together with these operations is a strict $\omega$-category.*

*Proof.* We have by Lemma 2.2.10 and Proposition 2.2.25 that the class of molecules is preserved under taking boundaries and pushouts of inclusions. Furthermore taking boundaries and pushouts of inclusions preserves isomorphisms, so $\partial_n^+, \partial_n^-$ and $\#_n$ are well defined in *Mol*.
Finite dimension follows from Lemma 2.2.9 and applying Proposition 2.2.25, we conclude that *Mol* is a reflexive $\omega$-graph. We are left to check the axioms of strict $\omega$-category.
Compatibility with boundaries follows by combining Lemma 2.2.17, Lemma 2.2.19 and Lemma 2.2.20. Associativity, interchange and unitality follow from Proposition 2.3.13. ∎

**Proposition 5.1.15.** *The set Atom is a basis for the strict $\omega$-category Mol.*

*Proof.* First we show that *Atom* is a spanning set. We will induct on $\operatorname{lydim} U$. If $\operatorname{lydim} U = -1$, then by Lemma 3.2.8, $U$ is an atom and we are done. Otherwise, if $\operatorname{lydim} U = k$, we have by Theorem 3.2.11 that $U$ admits a $k$-layering $U^{(1)} \#_k U^{(2)} \dots \#_k U^{(m)}$. By Lemma 3.2.10, $\operatorname{lydim} U^{(i)} < k$, so we may apply the inductive hypothesis on $[U^{(i)}]$ to conclude that $[U] \in \operatorname{span} \mathit{Atom}$.
For the other direction, let $[U] \in \mathit{Atom}$ and suppose $[U] = [U_1] \#_k [U_2] = [U_1 \#_k U_2]$ for some $[U_1], [U_2] \in \operatorname{span} \mathit{Atom}$. If $k < \min\{\dim U_1, \dim U_2\}$, then either $[U] = [U_1]$ or $[U] = [U_2]$ and we are done. Otherwise, we obtain a contradiction with $U$ being an atom by Lemma 3.2.8. ∎

***Definition* 5.1.16** (Molecules over an oriented graded poset) — For each morphism $f\colon U \to P$ of oriented graded posets, let $[f]$ denote its isomorphism class in the slice category $\mathbf{ogPos}/P$. Given an oriented graded poset $P$, we let

$$\begin{aligned} \mathit{Mol}/P &\coloneqq \{[f\colon U \to P] \mid U \text{ is a molecule}\}, \\ \mathit{Atom}/P &\coloneqq \{[f\colon U \to P] \mid U \text{ is an atom}\} \subseteq \mathit{Mol}/P, \end{aligned}$$

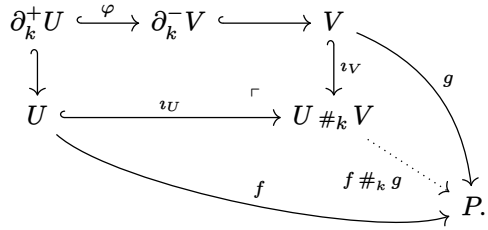
which we call *molecules* and *atoms over* $P$. For all $k \in \mathbb{N}$ and $\alpha \in \{+,-\}$,

$$\partial^\alpha_k\colon \mathit{Mol}/P \to \mathit{Mol}/P, \qquad [f\colon U \to P] \mapsto [f|_{\partial^\alpha_k U}\colon \partial^\alpha_k U \to P]$$

make $\mathit{Mol}/P$ a reflexive $\omega$-graph. If $[f\colon U \to P]$, $[g\colon V \to P]$ are $k$-composable molecules over $P$, then there exists a unique isomorphism $\varphi$ such that

$$\begin{array}{ccccc} \partial^+_k U & \xrightarrow{\varphi} & \partial^-_k V & \hookrightarrow & V \\ \downarrow & & & & \downarrow{\scriptstyle g} \\ U & & \xrightarrow{\quad f \quad} & & P \end{array}$$

commutes, which induces, by the universal property of $U \#_k V$, a unique morphism $f \#_k g$ such that the following diagram commutes:




**Proposition 5.1.17.** *Let $P$ be an oriented graded poset and, for each $k \in \mathbb{N}$,*

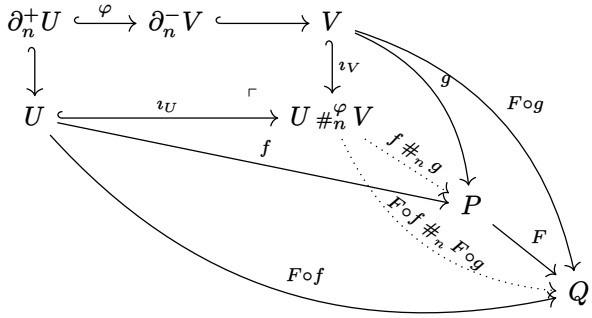
$$\begin{aligned} -\#_k -\colon \mathit{Mol}/P \times_k \mathit{Mol}/P &\to \mathit{Mol}/P, \\ [f\colon U \to P], [g\colon V \to P] &\mapsto [f \#_k g\colon U \#_k V \to P]. \end{aligned}$$

*If $F : P \to Q$ is a morphism of oriented graded posets, we let $\mathit{Mol}/F : \mathit{Mol}/P \to \mathit{Mol}/Q$ be defined by:*

$$F([f : U \to P]) = [F \circ f : U \to Q].$$

*This assignment extends to a functor $\mathit{Mol}/\_\colon \mathbf{ogPos} \to \omega\mathbf{Cat}$.*

*Proof.* The fact that $Mol/_-$ is a functor on objects is just a fibrant variant of Proposition 5.1.14. Note that $Mol/_F$ is well-defined since precomposing with $F$ preserves isomorphisms in the slice category. Compatibility of $F$ with boundaries follows by checking the equality element-wise. The compatibility with the $\#_n$ operator comes from the universal property of the pushout, as follows:




We have that $F \circ (f \#_n g) \circ \imath_U = F \circ f$ (by definition); and by the uniqueness property of the pushout, we obtain that $F \circ (f \#_n g) = F \circ f \#_n F \circ g$. Therefore $Mol/_F$ is a strict $\omega$-functor. Functoriality of the assigment $F \to Mol/_F$ is routine to check.

■

**Remark 5.1.18.** *The set $Atom/_P$ is a basis for $Mol/_P$.*

In particular $Mol/_-$ restricts to a functor from **RDC** to $\omega$**Cat**. When $P$ is a regular directed complex, $Mol/_P$ admits a basis whose elements are in bijection with the elements of $P$. This comes from very strong rigidity properties of atoms, which do not generalise to other oriented graded posets. To prove this we need the following result:

**Theorem 5.1.19.** *Let $f : U \to V$ be a morphism of atoms of the same dimension. Then $f$ is an isomorphism.*

*Proof.* See [25, Theorem 5.3.7]. ■

**Corollary 5.1.20.** *If $P$ is a regular directed complex, $\{[\mathrm{cl}\,\{x\} \hookrightarrow P] \mid x \in P\}$ is a basis for the $\omega$-category $Mol/_P$.*

*Proof.* We have by Lemma 2.1.31 that $f(\mathrm{cl}\, x) = \mathrm{cl}\, f(x)$ is an atom, so by Theorem 5.1.19, $f$ restricts to an isomorphism with its image. Therefore every morphism $\mathrm{cl}\, x \to P$ is an inclusion. The claim follows from Remark 5.1.18. ■

***Definition* 5.1.21** (Local embedding of oriented graded posets) — A morphism $f\colon P \to Q$ of oriented graded posets is a *local embedding* if, for all $x \in P$, the restriction $f|_{\mathrm{cl}\{x\}}$ is an inclusion, hence determines an isomorphism between $\mathrm{cl}\,\{x\}$ and its image $\mathrm{cl}\,\{f(x)\}$.

**Proposition 5.1.22.** *Let $f\colon P \to Q$ be a morphism of regular directed complexes. Then $f$ is a local embedding.*

*Proof.* Follows immediately from Theorem 5.1.19. ∎

***Definition* 5.1.23** (Diagram in a strict $\omega$-category) — Let $X$ be a strict $\omega$-category and $P$ a regular directed complex. A *diagram of shape $P$ in $X$* is a strict functor $d\colon Mol/P \to X$. A diagram is a *pasting diagram* if its shape is a molecule.

Note that by [25, Proposition 5.3.15] a diagram in a strict $\omega$-category induces a diagram in the sense used in the previous chapters by picking an injective encoding of non-degenerate cells in the set of variables.

### 5.1.1 Polygraphs

***Definition* 5.1.24** (Cellular extension of a strict $\omega$-category) — Let $X$ be a strict $\omega$-category. A *cellular extension of $X$* is a strict $\omega$-category $X_{\mathscr{S}}$ together with a pushout diagram

$$\begin{array}{ccc} \coprod_{e\in\mathscr{S}} Mol/\partial U_e & \xhookrightarrow{\coprod_{e\in\mathscr{S}} Mol/\imath_e} & \coprod_{e\in\mathscr{S}} Mol/U_e \\ \downarrow{\scriptstyle (\partial e)_{e\in\mathscr{S}}} & & \downarrow{\scriptstyle (e)_{e\in\mathscr{S}}} \\ X & \xhookrightarrow{\qquad\qquad} & X_{\mathscr{S}} \end{array}$$

in $\omega\mathbf{Cat}$, where, for each $e \in \mathscr{S}$, $U_e$ is an atom and $\imath_e\colon \partial U_e \hookrightarrow U_e$ is the inclusion of its boundary.

**Remark 5.1.25.** *The functor $X \hookrightarrow X_{\mathscr{S}}$ in a cellular extension is always injective, as shown in [37, Section 4].*
*This is a non-standard definition of cellular extension, allowing any atom as a potential cell shape; the usual definition only uses* globes*. However, the two are equivalent in the sense that a cellular extension in our sense can always be turned into a cellular extension in the more restrictive sense.*

***Definition* 5.1.26** (Polygraph) — A *polygraph*, also known as *computad*, is a strict $\omega$-category $X$ together with, for each $n \in \mathbb{N}$, a pushout diagram

$$\begin{array}{ccc} \coprod_{e\in\mathscr{S}_n} Mol/\partial U_e & \xhookrightarrow{\coprod_{e\in\mathscr{S}_n} Mol/\imath_e} & \coprod_{e\in\mathscr{S}_n} Mol/U_e \\ \downarrow{\scriptstyle (\partial e)_{e\in\mathscr{S}_n}} & & \downarrow{\scriptstyle (e)_{e\in\mathscr{S}_n}} \\ \sigma_{\leq n-1}X & \xhookrightarrow{\qquad\qquad} & \sigma_{\leq n}X \end{array}$$

in $\omega\mathbf{Cat}$, exhibiting $\sigma_{\leq n}X$ as a cellular extension of $\sigma_{\leq n-1}X$, such that $U_e$ is an $n$-dimensional atom for all $e \in \mathscr{S}_n$. The set

$$\mathscr{S} \coloneqq \sum_{n\in\mathbb{N}} \{e[\mathrm{id}_{U_e}] \mid e \in \mathscr{S}_n\}$$

is called the set of *generating cells* of the polygraph. We write $(X, \mathscr{S})$ for a polygraph $X$ with set $\mathscr{S}$ of generating cells.

**Lemma 5.1.27.** *Let $(X, \mathscr{S})$ be a polygraph. Then $\mathscr{S}$ is a basis for $X$.*

*Proof.* The fact that $\mathscr{S}$ is a generating set and its minimality are consequences of [1, Proposition 15.1.8 and Lemma 16.6.2], respectively. ∎

***Definition* 5.1.28** (Oriented graded poset with frame-acyclic molecules) — Let $P$ be an oriented graded poset. We say that $P$ *has frame-acyclic molecules* if, for all molecules $U$, if there exists a morphism $f : U \to P$, then $U$ is frame-acyclic.

**Lemma 5.1.29.** *Let $P$ be an oriented graded poset, $n \in \mathbb{N}$, and let $\mathscr{S}_n$ be a set containing one pasting diagram*

$$e \equiv Mol\!/_e : Mol\!/_{U_e} \to \sigma_{\leq n} Mol\!/_P$$

*for each $[e : U_e \to P]$ in $Atom\!/_P$ such that $\dim U_e = n$. If $\sigma_{\leq n} P$ has frame-acyclic molecules, then*

$$\begin{array}{ccc} \coprod_{e \in \mathscr{S}_n} Mol\!/_{\partial U_e} & \xhookrightarrow{\coprod_{e \in \mathscr{S}_n} Mol\!/_{\iota_e}} & \coprod_{e \in \mathscr{S}_n} Mol\!/_{U_e} \\ \downarrow{\scriptstyle (\partial e)_{e \in \mathscr{S}_n}} & & \downarrow{\scriptstyle (e)_{e \in \mathscr{S}_n}} \\ \sigma_{\leq n-1} Mol\!/_P & \xhookrightarrow{\qquad\qquad} & \sigma_{\leq n} Mol\!/_P \end{array}$$

*is a pushout diagram in $\omega\mathbf{Cat}$, exhibiting $\sigma_{\leq n} Mol\!/_P$ as a cellular extension of $\sigma_{\leq n-1} Mol\!/_P$.*

*Proof.* Let $X$ be a strict $\omega$-category and let

$$\begin{array}{ccc} \coprod_{e \in \mathscr{S}_n} Mol\!/_{\partial U_e} & \xhookrightarrow{\coprod_{e \in \mathscr{S}_n} Mol\!/_{\iota_e}} & \coprod_{e \in \mathscr{S}_n} Mol\!/_{U_e} \\ \downarrow{\scriptstyle (\partial e)_{e \in \mathscr{S}_n}} & & \downarrow{\scriptstyle k} \\ \sigma_{\leq n-1} Mol\!/_P & \xrightarrow{\qquad h \qquad} & X \end{array}$$

be a commutative diagram of strict functors. We define $\bar{h} : \sigma_{\leq n} Mol\!/_P \to X$ as follows on cells $[f : U \to P]$ in $\sigma_{\leq n} Mol\!/_P$. If $\dim U < n$, then we let

$$\bar{h}[f] \coloneqq h[f].$$

Suppose $\dim U = n$; we proceed by induction on $\operatorname{lydim} U$. If $\operatorname{lydim} U = -1$, then by Lemma 3.2.8 $U$ is an atom, so there exists a unique $Mol\!/_e \in \mathscr{S}_n$ such that $[f] = [e]$, and we let $\bar{h}[f] \coloneqq k[\mathrm{id}_{U_e}]$. If $\operatorname{lydim} U = k \geq 0$, then $U$ admits a $k$-layering $(U^{(i)})_{i=1}^m$, and each layer $U^{(i)}$ has strictly lower layering dimension. Then we let

$$\bar{h}[f] \coloneqq \bar{h}[f|_{U^{(1)}}] \#_k \ldots \#_k \bar{h}[f|_{U^{(m)}}].$$

By construction, if $\bar{h}$ is well-defined, then it is a strict functor satisfying $\bar{h} \circ (e)_{e \in \mathscr{S}_n} = k$ and restricting to $h$ on $\sigma_{\leq n-1} Mol\!/_P$. Moreover, let $h'$ be another strict functor with the same property. Then $h'$ agrees with $\bar{h}$ on all atoms of dimension $\leq n$, which form a basis of $\sigma_{\leq n} Mol\!/_P$. It follows from Lemma 5.1.12 that $h' = \bar{h}$. It only remains to show that $\bar{h}$ is well-defined, that is, it is independent of the choice of a $k$-layering of $U$ when $\dim U = n$ and $k \coloneqq \operatorname{lydim} U \geq 0$.

We may assume, inductively, that $\bar{h}$ is well-defined on all cells $[g\colon V \to P]$ such that $\dim V < n$ or $\operatorname{lydim} V < k$. Let $(U^{(i)})_{i=1}^m$ and $(V^{(i)})_{i=1}^m$ be two $k$-layerings of $U$ and let $(x^{(i)})_{i=1}^m$, $(y^{(i)})_{i=1}^m$ be the induced $k$-orderings. Further, let $\sigma$ be the unique permutation such that $x^{(i)} = y^{(\sigma(i))}$ for all $i \in \{1, \dots, m\}$ and

$$d \coloneqq \mathsf{d}((x^{(i)})_{i=1}^m, (y^{(i)})_{i=1}^m)$$

the number of pairs $(j, j')$ such that $j < j'$ but $\sigma(j') < \sigma(j)$, and proceeding by induction on $d$. If $d = 0$, then the two layerings are equal up to layer-wise isomorphism. If $d > 0$, then there exists $j < m$ such that $\sigma(j+1) < \sigma(j)$, and we let $W \sqsubseteq U$ be the image of $U^{(j)} \#_k U^{(j+1)}$ in $U$. Then $W$ contains exactly two elements $z_1 \coloneqq x^{(j)} = x^{(\sigma(j))}$ and $z_2 \coloneqq x^{(j+1)} = y^{(\sigma(j+1))}$ of dimension $> k$, yet there can be no edge between them in $\mathscr{M}_k U$, from which we deduce that $r \coloneqq \operatorname{frdim} W < k$. By assumption, $W$ is frame-acyclic, so by Theorem 4.3.5 there exists an $r$-layering of $W$, hence also a pair of molecules $W^{(1)}$, $W^{(2)}$, each containing a single element of dimension $> k$, such that $W$ is isomorphic to $W^{(1)} \#_r W^{(2)}$. We may assume, without loss of generality, that $z_1$ is in the image of $W^{(1)}$ and $z_2$ in the image of $W^{(2)}$. We then have

$$\begin{aligned}
&\bar{h}[f|_{U^{(j)}}] \#_k \bar{h}[f|_{U^{(j+1)}}] = \\
&\quad = \left(\bar{h}[f|_{W^{(1)}}] \#_r \bar{h}[f|_{\partial_k^- W^{(2)}}]\right) \#_k \left(\bar{h}[f|_{\partial_k^+ W^{(1)}}] \#_r \bar{h}[f|_{W^{(2)}}]\right),
\end{aligned}$$

which by interchange and unitality in $X$ is equal to

$$\begin{aligned}
&\bar{h}[f|_{W^{(1)}}] \#_r \bar{h}[f|_{W^{(2)}}] = \\
&\quad = \left(\bar{h}[f|_{\partial_k^- W^{(1)}}] \#_r \bar{h}[f|_{W^{(2)}}]\right) \#_k \left(\bar{h}[f|_{W^{(1)}}] \#_r \bar{h}[f|_{\partial_k^+ W^{(2)}}]\right) = \\
&\quad = \bar{h}[f|_{\tilde{U}^{(j)}}] \#_k \bar{h}[f|_{\tilde{U}^{(j+1)}}],
\end{aligned}$$

where we let $\tilde{U}^{(j)} \coloneqq \partial_k^- W^{(1)} \#_r W^{(2)}$ and $\tilde{U}^{(j+1)} \coloneqq W^{(1)} \#_r \partial_k^+ W^{(2)}$. Notice that all the $n$-dimensional cells in this calculation involve molecules whose layering dimension is $< k$, so $\bar{h}$ is well-defined on each of them. Letting $\tilde{U}^{(i)} \coloneqq U^{(i)}$ for all $i \notin \{j, j+1\}$, we have that

I $(\tilde{U}^{(i)})_{i=1}^m$ is a $k$-layering of $U$,

II the definition of $\bar{h}[f]$ using $(U^{(i)})_{i=1}^m$ is equal to the one using $(\tilde{U}^{(i)})$, and

III the induced $k$-ordering $(\tilde{x}^{(i)})_{i=1}^m \coloneqq (x^{(1)}, \dots, x^{(j+1)}, x^{(j)}, \dots, x^{(m)})$ satisfies $\mathsf{d}((\tilde{x}^{(i)})_{i=1}^m, (y^{(i)})_{i=1}^m) < d$,

so, by the inductive hypothesis on $d$, the definition of $\bar{h}[f]$ using $(\tilde{U}^{(i)})_{i=1}^m$ is equal to the definition using $(V^{(i)})_{i=1}^m$. We conclude that $\bar{h}[f]$ is well-defined, which completes the proof. ∎

**Theorem 5.1.30.** *Let $P$ be an oriented graded poset with frame-acyclic molecules. Then $\mathit{Mol}/P$ is a polygraph whose set of generating cells is $\mathit{Atom}/P$.*

*Proof.* If $P$ has frame-acyclic molecules, then $\sigma_{\le n} P$ has frame-acyclic molecules for all $n \in \mathbb{N}$. The statement then follows from Lemma 5.1.29. ∎

**Remark 5.1.31.** *In fact, by the roundness property of atoms, if $\mathit{Mol}/P$ is a polygraph, then it is a* regular polygraph *in the sense of [31].*

## 5.2 Augmented directed complexes

In this section we introduce *augmented directed complexes* which play a crucial part in Steiner theory [50]. We define a functor from the category of regular directed complexes to the category of augmented directed complexes.

***Definition* 5.2.1** (Chain complex) — A chain complex $(A, d)$ is a sequence of abelian groups,

$$\ldots \xleftarrow{d_0} A_0 \xleftarrow{d_1} A_1 \xleftarrow{d_2} \ldots \xleftarrow{d_n} A_n \xleftarrow{d_{n+1}} \ldots ,$$

such that the composite of two consecutive maps is the zero morphism, i.e. $d_n \circ d_{n+1} = 0$ or $d^2 = 0$.

***Definition* 5.2.2** (Augmented directed complex) — An *augmented chain complex* $C$ is a chain complex of abelian groups in non-negative degree

$$\ldots \xrightarrow{\mathrm{d}} C_n \xrightarrow{\mathrm{d}} C_{n-1} \xrightarrow{\mathrm{d}} \ldots \xrightarrow{\mathrm{d}} C_1 \xrightarrow{\mathrm{d}} C_0$$

together with a homomorphism $\mathrm{e} \colon C_0 \to \mathbb{Z}$ satisfying $\mathrm{e} \circ \mathrm{d} = 0$. A *direction* on $C$ is a choice of a commutative submonoid $C_n^{\to}$ of $C_n$ for each $n \in \mathbb{N}$. An *augmented directed complex* is an augmented chain complex $C$ together with a direction on its underlying chain complex.

***Definition* 5.2.3** (Homomorphism of augmented directed complexes) — Let $C, D$ be augmented directed complexes. A *homomorphism* $f \colon C \to D$ is a homomorphism of the underlying augmented chain complexes, that is, a sequence $(f_n \colon C_n \to D_n)_{n \in \mathbb{N}}$ of homomorphisms of abelian groups satisfying

$$\mathrm{d} \circ f_{n+1} = f_n \circ \mathrm{d}, \qquad e \circ f_0 = e,$$

which is compatible with directions in the sense that

$$f_n(C_n^{\to}) \subseteq D_n^{\to}$$

for all $n \in \mathbb{N}$. Augmented directed complexes with their homomorphisms form a category $\mathbf{ADC}$.

We are now ready to construct augmented directed complexes from regular directed complexes.

***Definition* 5.2.4** (Augmented directed complex of a regular directed complex) — Let $P$ be a regular directed complex. The *augmented directed complex of* $P$, denoted by $\vec{\mathbb{Z}}P$, is the augmented chain complex

$$\ldots \xrightarrow{\mathrm{d}} \mathbb{Z}P_n \xrightarrow{\mathrm{d}} \mathbb{Z}P_{n-1} \xrightarrow{\mathrm{d}} \ldots \xrightarrow{\mathrm{d}} \mathbb{Z}P_1 \xrightarrow{\mathrm{d}} \mathbb{Z}P_0$$

where $\mathbb{Z}P_n$ is the free abelian group on the set $P_n$ and, for each $n > 0$, the homomorphism $\mathrm{d} \colon \mathbb{Z}P_n \to \mathbb{Z}P_{n-1}$ is defined on the generators $x \in P_n$ by

$$x \mapsto \sum_{y \in \Delta^+ x} y - \sum_{y \in \Delta^- x} y, \tag{5.1}$$

together with the homomorphism $\mathrm{e} \colon \mathbb{Z}P_0 \to \mathbb{Z}$ defined on the generators $x \in P_0$ by $x \mapsto 1$ and the direction given by $\mathbb{Z}P_n^{\to} := \mathbb{N}P_n$ for each $n \in \mathbb{N}$.

To prove that $\vec{\mathbb{Z}}P$ is indeed an augmented directed complex, we will use the following lemma:

**Lemma 5.2.5.** *Let $U$ be a molecule. Then, $\Delta^\alpha(\partial^+ U) = \Delta^\alpha(\partial^- U)$.*

*Proof.* Let $x \in \Delta^\alpha(\partial^+ U)$. Then, by definition, $x \in \partial^\alpha(\partial^+ U) = \partial^\alpha(\partial^- U)$ where the latter equality is by globularity. If $x$ is not maximal in $U$, then by Proposition 5.1.6, $x \in \Delta^\alpha(\partial^- U)$. If $x$ is maximal, then $x$ is not covered by anything and we can conclude that $x \in \Delta^\alpha(\partial^- U)$. ∎

**Proposition 5.2.6.** *Let $P$ be a regular directed complex. Then $\vec{\mathbb{Z}}P$ is an augmented directed complex.*

*Proof.* To show $\vec{\mathbb{Z}}P$ defined is an augmented directed complex, we need to show that $d^2 = 0$ and $ed = 0$. First, let $d^+ \coloneqq \sum_{y \in \Delta^+ x} y$ and $d^- \coloneqq \sum_{y \in \Delta^- x} y$. Then,

$$dd^+ x = \sum_{z \in \Delta^+(\partial^+ x)} z - \sum_{z \in \Delta^-(\partial^+ x)} z, \text{ and}$$
$$dd^- x = \sum_{z \in \Delta^+(\partial^- x)} z - \sum_{z \in \Delta^-(\partial^- x)} z.$$

Since $\partial^\alpha_n x$ is a molecule and $z$ is covered by $y \in \partial^\alpha_n x$ with orientation $\beta$, then by Corollary 2.3.9 it is either only covered by $y$ and we have $z \in \Delta^\beta \partial^\alpha_n x$ or is covered by another element $y'$ with orientation $-\beta$ and in that case it cancels.
Because $\Delta^+(\partial^\alpha x)$ and $\Delta^-(\partial^\alpha x)$ are disjoint, we want to show that $dd^+ x = dd^- x$, that is, $\Delta^+(\partial^+ x) = \Delta^+(\partial^- x)$ and $\Delta^-(\partial^+ x) = \Delta^-(\partial^- x)$. By Lemma 5.2.5 the two equalities hold and we get $d^2 = 0$.
Now, we need to show that $ed = 0$. In this case, $x$ is 1-dimensional and $|\Delta^+ x| = |\Delta^- x| = 1$. So $ed(x) = 1 - 1 = 0$. ∎

If $f : P \to Q$ is a morphism of regular directed complexes, we define $\vec{\mathbb{Z}}f : \vec{\mathbb{Z}}P \to \vec{\mathbb{Z}}Q$ by freely extending $\vec{\mathbb{Z}}f(x) = f(x)$ and letting $\vec{\mathbb{Z}}f(n) = n$ for $n \in \mathbb{Z}$.

**Lemma 5.2.7.** *Let $f : P \to Q$ be as above. Then $\vec{\mathbb{Z}}f : \vec{\mathbb{Z}}P \to \vec{\mathbb{Z}}Q$ is a morphism of augmented directed complexes.*

*Proof.* It follows from Lemma 2.1.31 that $f$ preserves the dimension so it maps $P_n$ to $Q_n$. For $x \in P_n$, we have

$$d(\vec{\mathbb{Z}}f(x)) = d(f(x)) = \sum_{y \in \Delta^+ f(x)} y - \sum_{y' \in \Delta^- f(x)} y',$$
$$\vec{\mathbb{Z}}f(d(x)) = \vec{\mathbb{Z}}f\Big(\sum_{y \in \Delta^+ x} y - \sum_{y' \in \Delta^- x} y'\Big) = \sum_{y \in \Delta^+ x} f(y) - \sum_{y' \in \Delta^- x} f(y').$$

We have by the defintion of morphisms that $f(\Delta^\alpha(x)) = \Delta^\alpha(f(x))$, so we obtain $d(\vec{\mathbb{Z}}f(x)) = \vec{\mathbb{Z}}f(d(x)))$. The compatibility of $\vec{\mathbb{Z}}f$ with $e$ follows easily, thus $\vec{\mathbb{Z}}f$ is indeed a morphism of augmented directed complexes. ∎

**Corollary 5.2.8.** *The assignment $\vec{\mathbb{Z}}$: **RDC** $\to$ **ADC** is a functor.*

*Proof.* Follows from Proposition 5.2.6 and Lemma 5.2.7 together with the easy compatibility checks. ∎

## 5.3 Image of the functor $\vec{\mathbb{Z}}$

The functor $\vec{\mathbb{Z}}$ is not full; to understand which augmented directed complexes may be constructed from regular directed complexes, we need to recall more terminology due to Steiner.

***Definition* 5.3.1 —** (Basis for an augmented directed complex). Let $C$ be an augmented directed complex. A *basis* for $C$ is a set $B \subset \cup_{n\in\mathbb{N}} C_n$ such that each $C_n$ is a free abelian group with basis $B \cap C_n$ and each $C_n^{\to}$ is the submoid of $C$ generated by $B \cap C_n$.

We can impose a partial order on $C_n$ via $x \leq y$ if $y - x \in C_n^{\to}$. The basis elements in $C_n$ can be characterised as the minimal non-zero elements in $C_n^{\to}$, thus if $C$ has a basis, it is unique. Furthermore, $C_n$ is a lattice so we may make the following definition:

***Definition* 5.3.2 —** Let $x, y \in C_n$. We denote $x \vee y$ the least upper bound of $x$ and $y$ and $x \wedge y$ the greatest lower bound.
If $x \in C_{n+1}$, there are elements denoted $d^- x, d^+ x \in C_n$ such that

$$dx = d^+ x - d^- x, \quad d^+ x \wedge d^- x = 0_{C_n}.$$

In fact if $dx$ is expressed as a linear combination of distinct basis elements, then $d^+ x$ is the sum of terms with positive coefficients and $-d^- x$ is the sum of elements with negative coefficients.

***Definition* 5.3.3 —** Let $b$ be a basis element and suppose that $b \in C_n$. We define elements $\langle b \rangle_i^+$ and $\langle b \rangle_i^-$ via downward recursion:

$$\langle b \rangle_i^\alpha = \begin{cases} 0 \text{ for } i > n, \\ b \text{ for } i = n, \\ d^\alpha \langle b \rangle_{i+1}^\alpha \text{ for } i < n. \end{cases} \tag{5.2}$$

***Definition* 5.3.4 —** (Unital augmented directed complex). A basis $B$ for an augmented directed complex is unital if for all $b \in B$ such that $\dim b = n$, we have

$$e(\langle b \rangle_0^+) = e(\langle b \rangle_0^-) = 0.$$

We shall prove that under the functor $\vec{\mathbb{Z}}$, regular directed complexes map to unital augmented directed complexes. To do this, we need the following lemma:

**Lemma 5.3.5.** *Let $P$ be a regular directed complex and let $b \in \mathbb{Z}P_n$. Then,*

$$\langle b \rangle_k^\alpha = \sum_{y \in \Delta_k^\alpha b} y$$

*Proof.* We prove this by downward induction.
Case $k = n$. Then $\langle b \rangle_n^\alpha = b = \sum_{y\in\Delta_n^\alpha b} y$.
Case $k = n-1$. Then $db = \sum_{y\in\Delta_n^+ b} y - \sum_{y\in\Delta_n^- b} y$. By definition 5.3.2 we get that $db^+ = \sum_{y\in\Delta_n^+ b} y$ and $db^- = \sum_{y\in\Delta_n^- b} y$. Note that this definition coincides with

the notation we use in the proof of Proposition 5.2.6.
Suppose the statement holds for the case $k = i$ - i.e., $\langle b\rangle_i^\alpha = \sum_{y\in\Delta_i^\alpha b} y$; we prove that it holds for $k = i-1$. we have

$$d\langle b\rangle_i^\alpha = d(\sum_{y\in\Delta_i^\alpha b} y) = \sum_{y\in\Delta_i^\alpha b} dy = \sum_{y\in\Delta_i^\alpha b}(\sum_{z\in\Delta^+ y} z - \sum_{z\in\Delta^- y} z). \tag{5.3}$$

Since $\mathrm{cl}\, b$ is an atom and the only maximal element in $\mathrm{cl}\, b$ is $b$, $\mathrm{cl}\,\Delta_i^\alpha b = \partial_i^\alpha b$. Since $z$ is a codimension 1 element in $\partial_i^\alpha b$, and and there are no maximal elements other than $b$, we have by Corollary 2.3.9, that $z$ is either in internal in $\partial_i^\alpha b$ which means it is covered by exactly one element with orientation $+$ and exactly one element with orientation $-$ and it cancels in (5.3) or $z$ is in $\Delta^\beta\partial_i^\alpha b$.
By Corollary 2.3.9 and globularity, we can write the sum above as:

$$d\langle b\rangle_i^+ = \sum_{z\in\Delta_{i-1}^+ b} z - \sum_{z\in\Delta_{i-1}^- b} z.$$

By definition 5.3.2 we can write $d^+\langle b\rangle_i^\alpha = \sum_{z\in\Delta_{i-1}^+ b} z$ and $d^-\langle b\rangle_i^\alpha = \sum_{z\in\Delta_{i-1}^- b} z$. And we can conclude that

$$\langle b\rangle_{i-1}^\alpha = d^\alpha\langle b\rangle_i^\beta = \sum_{z\in\Delta_{i-1}^\alpha b} z.$$

■

**Proposition 5.3.6.** *Let $P$ be a regular directed complex. Then $\vec{\mathbb{Z}}P$ is an augmented directed complex with unital basis.*

*Proof.* Let $X_n$ be the set of elements of dimension $n$ in $P$ and $X = \cup_{n\in\mathbb{N}} X_n$. By construction, it is clear that $X$ is a basis for $\vec{\mathbb{Z}}P$, so we are left to prove it is unital. Let $x \in X_n$.
If $n = 0$, then $ex = 1$ by definition, so we may assume $n \geq 1$ and let $b \in \mathbb{Z}P_n$ be a basis element. We have by Lemma 5.3.5 that $\langle b\rangle_0^\alpha = \sum_{y\in\Delta_0^\alpha b} y$. Therefore, we get $e(\langle b\rangle_0^\alpha) = |\Delta_0^\alpha b| = 1$. ■

## 5.4 Acyclicity conditions

Now that we made the connection between regular directed complexes and augmented directed complexes, let us recall Steiner's functors between augmented directed complexes and $\omega$-categories.

***Definition* 5.4.1 —** [50, Definition 2.6] Let $C$ be a chain complex concentrated in positive degree. One constructs $\mu C$ to be the abelian group of double sequences $(x_0^-, x_0^+, x_1^-, x_1^+, \ldots)$ such that

- $x_n^-, x_n^+ \in C_n$,
- $x_n^- = x_n^+$ for all but finitely many $n$,

- $x_n^+ - x_n^- = dx_{n+1}^- = dx_{n+1}^+$ for $n \geq 0$.

We impose a structure of a $\omega$-category on $\mu C$ by letting

$$\partial_n^\alpha(x) := (x_0^-, x_0^+, x_1^-, x_1^+, \dots, x_n^\alpha, x_n^\alpha, 0, 0 \dots)$$

and if $\partial_n^+ x = \partial_n^- y$

$$x \#_n y = (x_0^-, y_0^+, \dots, x_n^-, y_n^+, x_{n+1}^- + y_{n+1}^-, x_{n+1}^+ + y_{n+1}^+, \dots).$$

***Definition* 5.4.2 —** [50, Definition 2.8] Let $\nu : \mathbf{ADC} \to \omega\text{-cat}$ be the functor defined as follows; for an augmented directed complex $C$, we let $\nu C$ be the sub-$\omega$-category of $\mu C$ consisting of elements

$$(x_0^-, x_0^+, x_1^-, x_1^+, \dots)$$

such that $x_i^\alpha \in C_n^\to$ and $e(x_0^-) = e(x_0^+) = 1$.
If $f : C \to D$ is a morphism of augmented directed complex, we define $\mu(f) : \mu(C) \to \mu(D)$ via

$$\mu(f)(x_0^-, x_0^+, x_1^-, x_1^+, \dots) := (f(x_0^-), f(x_0^+), f(x_1^-), f(x_1^+), \dots)$$

***Definition* 5.4.3 —** Let $C$ be an augmented directed complex with a unital basis and let $b$ be a basis element. We let $\langle b \rangle = (\langle b \rangle_0^+, \langle b \rangle_0^-, \langle b \rangle_1^+, \langle b \rangle_1^-, \dots ...) \in \nu C$.

***Definition* 5.4.4 —** [50, Definition 2.4] The functor $\lambda : \omega\text{-cat} \to \mathbf{ADC}$ is defined as follows. Let $C$ be an $\omega$-category. Then the chain group $(\lambda C)_n$ for $n \geq 0$ is the group generated by the elements $[x]_n$ for $x \in C_n$ (where $C_n$ is an $n$-category) subject to relations

$$[x \#_m y]_n = [x]_n + [y]_n \text{ for } m < n,$$

the boundary homomorphism $d : (\lambda C)_{n+1} \to (\lambda C)_n$ for $n \geq 0$ is given by

$$d[x]_{n+1} = [\partial_n^+ x]_n - [\partial_n^- x]_n,$$

the augmentation $e : (\lambda C)_0 \to \mathbf{Z}$ is given by

$$e[x]_0 = 1,$$

the orientation $\lambda C_n^\to$ is the submonoid generated by the elements $[x]_n$.

**Theorem 5.4.5.** *[50, Theorem 2.11] The functors $\lambda : \omega\mathbf{Cat} \to \mathbf{ADC}$ and $\nu : \mathbf{ADC} \to \omega\mathbf{Cat}$ form an adjoint pair with unit $\eta : id \to \nu\lambda$ and counit $\pi : \lambda\nu \to id$ defined by:*

$$\begin{aligned} \eta x &= ([d_0^- x]_0, [d_0^+ x]_0, [d_1^- x]_1, [d_1^+ x]_1, \dots) \\ \pi[x]_n &= x_n^- = x_n^+ \end{aligned} \tag{5.4}$$

There are now two ways to get from regular directed complexes to augmented directed complexes: either directly via the $\vec{\mathbb{Z}}$ functor, or by first passing to the $\omega$-category $Mol/\_$, then applying Steiner's linearisation functor $\lambda$. Dually, there are two ways to get from a regular directed complex to an $\omega$-category: either via the $\omega$-category $Mol/\_$, or by applying the functor $\nu$ after $\vec{\mathbb{Z}}$. In the following, we will see how the combinatorial acyclicity conditions give rise to natural isomorphisms of functors.

### 5.4.1 Frame acyclicity

***Definition* 5.4.6** — Let $P$ be a regular directed complex that has frame-acyclic molecules. We define $G : \lambda(Mol/P) \to \vec{\mathbb{Z}}P$ as follows: $G([\operatorname{cl} b])_n = b \in \mathbb{Z}P_n$, if $\operatorname{cl} b$ is an atom and $\dim b = n$ and $G([\operatorname{cl} b])_n = 0$ if $\dim \operatorname{cl} b < n$.
We freely extend $G$ as follows: if $m < n$ and $x$ and $y$ are atoms then $G([x \#_m y]) = G([x]_n + [y]_n) = G([x]_n) + G([y]_n)$. If $U$ was obtained via the paste rule, with $\dim U = n$, let $x_1, x_2 \dots x_m$ be the maximal elements of dimension $n$. Then we let $G([U])_n = x_1 + \dots + x_m \in \mathbb{Z}P_n$.
Finally, to make $G$ compatible with the augmentation, we let $G(n) = n$ for $n \in \mathbb{Z}$.

We will now show that $G$ is an isomorphism of ADCs. Before proceeding, we need to check that $G$ is well-defined. For this, we will need two additional lemmas.

**Lemma 5.4.7.** *Let $U$ be a molecule, $r = \operatorname{frdim} U$, $r \geq 0$. Then, for any decomposition $U = U' \#_r U''$, $\operatorname{frdim} U' \leq \operatorname{frdim} U$ and $\operatorname{frdim} U'' \leq \operatorname{frdim} U$.*

*Proof.* Let $x_1$ and $x_2$ be two maximal elements in $U$. First assume that $x_1$, $x_2$ are in $U'$. If $\min\{\dim x_1,\ \dim x_2\} \leq r$, then $\dim(\operatorname{cl} x_1 \cap \operatorname{cl} x_2) \leq r \leq \operatorname{frdim} U$.
If $\dim x_1 > r$ and $\dim x_2 > r$, then $x_1$ and $x_2$ are maximal in $U' \#_r U''$, because $\partial_r^+ U' = \partial_r^- U''$ and by Lemma 2.1.25 $\dim(U' \cap U'') = r \leq \operatorname{frdim}(U' \#_r U'')$. Hence, by definition, $\dim(\operatorname{cl} x_1 \cap \operatorname{cl} x_2) \leq \operatorname{frdim}(U' \#_r U'')$.
The case where $x_1$ and $x_2$ are in $U''$ is similar. If $x_1 \in U'$ and $x_2 \in U''$, then $\dim \operatorname{cl} x_1 \cap \operatorname{cl} x_2 \leq \operatorname{frdim} U$, from the definition of frame dimension.

■

**Lemma 5.4.8.** *Let $U$ be a frame-acyclic molecule, with $\operatorname{frdim} U = r$. Then $U$ admits a decomposition into atoms using $\#_m$, where $m \leq r$.*

*Proof.* We will prove this by induction on the submolecules of $U$. Suppose that all submolecules properly contained in $U$ admit a an atom decomposition. If $U$ is an atom ($r = -1$), then we get the trivial decomposition. Alternatively, if $U$ is not an atom then it was generated using the Paste rule. By Theorem 4.3.5 and Corollary 4.3.6, $U$ admits an $r$-layering induced by an $r$-ordering of $\mathscr{M}_r U$, say $U_1 \#_r \dots \#_r U_m$. By grouping the factors appropriately, we can write $U = U' \#_r U''$, where $U'$ and $U''$ are submolecules of $U$. From the definition, $U'$ and $U''$ are frame-acyclic and by 5.4.7, we get that $\operatorname{frdim} U' \leq r$ and $\operatorname{frdim} U'' \leq r$. It follows by the inductive hypothesis that $U'$ and $U''$ have an atom decomposition for some $m \leq \operatorname{frdim} U$. So $U$ has a decomposition into atoms using $\#_m$. ■

**Lemma 5.4.9.** *Let $P$ be a regular directed complex with frame-acyclic molecules and let $G$ be a morphism defined as above. Let $U$ be a molecule in $P$ with $\dim U = n$ and $\{y_i\}_{i=1}^m$ the $n$-dimensional elements of $U$. Then $G([U]_n) = y_1 + \dots + y_m \in \vec{\mathbb{Z}}P_n$.*

*Proof.* To show that $G$ is well defined, we need to show that for any two decompositions of $U$, $d_1$ and $d_2$, $G(d_1) = G(d_2)$. We have by Proposition 5.4.8 that $U$ has a decomposition into atoms using $\#_l$, with $l < \operatorname{frdim} U$. Let $U = x_1 \#_{i_1} x_2 \#_{i_2} \dots \#_{i_{k-1}} x_k$ be an atom decomposition for $U$, with $k > m$. Since $\{x_i\}_{i=1}^k$ is an atom decomposition, each $y_i$ appears in at least one of $x_1, \dots, x_k$.

What we need to show is that $y_i$ appears in exactly one of the terms. Suppose for contradiction that $y_i = x_a$ and $y_i = x_b$, with $a < b$. Define $V, W, Z$ as follows:

$$\begin{aligned} V &= x_1 \#_{i_1} \dots \#_{i_{a-1}} x_a, \\ W &= \partial^+_{i_a} x_a \#_{i_a} \dots \#_{i_{b-2}} x_{b-1}, \\ Z &= \partial^+_{i_{b-1}} x_{b-1} \#_{i_{b-1}} x_b \#_{i_b} \dots \#_{i_{k-1}} x_k. \end{aligned}$$

Then $U$ splits into $V \#_{i_a} (W \#_{i_{b-1}} Z)$ with $\partial^+_{i_a} x_a = \partial^+_{i_a} V = \partial^-_{i_a}(W \#_{i_{b-1}} Z) = = V \cap (W \#_{i_{b-1}} Z)$. Now $\mathrm{cl}\{x_a\} \subseteq V$ and $\mathrm{cl}\{x_b\} \subseteq (W \#_{i_{b-1}} Z)$. Then $\mathrm{cl}\{x_a\} \cap \mathrm{cl}\{x_b\} \subseteq V \cap (W \#_{i_{b-1}} Z) = \partial^+_{i_a} x_a$. Since $\dim x_b = n$ by assumption and $\partial^+_{i_a} x_a$ contains elements of dimension less than or equal to $i_a < n$, we obtain a contradiction. ∎

**Proposition 5.4.10.** *Let $G : \lambda(Mol/P) \to \vec{\mathbb{Z}}P$ be defined as in 5.4.6. Then $G$ is an isomorphism of augmented directed complexes.*

*Proof.* We need to show that if $\dim U = n+1$, then $G(d[U]_{n+1}) = dG([U]_{n+1})$ and if $\dim U = 0$, then $G(e[U]_0) = eG([U]_0)$.
We start by first showing that $G$ preserves the boundary homomorphism. We know that $d[U]_{n+1} = [\partial^+_n U]_n - [\partial^-_n U]_n$. So,

$$\begin{aligned} G([\partial^+_n U]_n - [\partial^-_n U]_n) &\overset{G \text{ is free}}{=} G([\partial^+_n U]_n) - G([\partial^-_n U]_n) \\ &\overset{5.4.9}{=} \sum_{\substack{y \in \partial^+_n U \\ \dim(y)=n}} y - \sum_{\substack{z \in \partial^-_n U \\ \dim(z)=n}} z \overset{2.1.26}{=} \\ &= \sum_{y \in \Delta^+_n U} y - \sum_{z \in \Delta^-_n U} z := s_1. \end{aligned}$$

Now,

$$\begin{aligned} dG([U]_{n+1}) &\overset{5.4.9}{=} d\Big(\sum_{x \in U_{n+1}} x\Big) = d(x_1 + \dots + x_m) \\ &= dx_1 + \dots + dx_m = \\ &= \Big(\sum_{y_1 \in \Delta^+ x_1} y_1 - \sum_{y'_1 \in \Delta^- x_1} y'_1\Big) + \dots + \Big(\sum_{y_m \in \Delta^+ x_m} y_m - \sum_{y'_m \in \Delta^- x_m} y'_m\Big) \\ &:= s_2. \end{aligned}$$

Take $y \in \Delta^\alpha x_1$. Then, by Corollary 2.3.9 $y$ is either covered only with orientation $\alpha$, so $y \in \Delta^\alpha U$, or $y$ is covered by exactly two elements with opposite orientations. In the first scenario $y$ appears exactly once in $s_1$, while in the second situation there is a cancellation in $s_2$.
On the other hand, if $y \in \Delta^\alpha U$, there exists $x_i$ such that $y \in \Delta^\alpha x_i$. Therefore

$$dG([U]_{n+1}) = \sum_{y \in \Delta^+_n U} y - \sum_{z \in \Delta^-_n U} z = G(d[U_{n+1}]).$$

Since the only 0-dimensional elements are the points, $G[U]_0 = U_0$ and $e(U_0) = 1$ by definition. Alternatively, $G(e[U]_0) = G(1) = 1$ as defined in 5.4.6. So $G$ is a morphism of augmented directed complexes.
By Theorem 5.1.30, we have that $(\lambda(Mol/_P))_n$ is freely generated by its atoms of dimension $n$. We get that $G$ is an isomorphism since (by contruction) $G$ is a bijective map between the generators of the free groups of dimension $n$.

■

### 5.4.2 Dimension-wise acyclicity

Recall by Proposition 4.3.15 that dimension-wise acyclicity is a strictly stronger condition that implies frame-acyclicity. We show that under the functor $\vec{\mathbb{Z}}$ a dimension-wise acyclic regular directed complex gives rise to a Steiner complex.

***Definition* 5.4.11 —** Let $P$ be a regular directed complex. We say that $P$ is *dimension-wise acyclic* if, for all $n \in \mathbb{N}$, $\mathscr{F}_n P$ is acyclic.

***Definition* 5.4.12 —** A basis $B$ for an augmented directed complex is loop-free if there are partial orderings $\leq_0, \leq_1, \ldots$ such that $a <_n b$ whenever $\langle a \rangle_n^+ \wedge \langle b \rangle_n^- > 0, |a|, |b| > n$. An augmented directed complex is loop-free if it has a loop-free basis. A *Steiner* complex is an augmented directed complex with a loop-free basis. We let $\mathbf{ADC}_{St}$ be the full subcategory of $\mathbf{ADC}$ on Steiner complexes.

The following is the fundamental theorem of Steiner theory.

**Theorem 5.4.13.** *The restriction of $\nu\colon \mathbf{ADC} \to \omega\mathbf{Cat}$ to $\mathbf{ADC}_{St}$ is full and faithful. Moreover, if $C$ is a Steiner complex with basis $(\mathscr{B}_n)_{n\in\mathbb{N}}$, then $\nu C$ is a polygraph whose set of generating cells is*

$$\left\{ \langle b \rangle \mid b \in \bigcup_{n\in\mathbb{N}} \mathscr{B}_n \right\}.$$

*Proof.* See [50, Theorem 5.6 and Theorem 6.1]. ■

**Proposition 5.4.14.** *Let $P$ be a dimension-wise acyclic regular directed complex. Then $\vec{\mathbb{Z}}P$ is a Steiner complex.*

*Proof.* By Proposition 5.3.6, $\vec{\mathbb{Z}}P$ has a basis given by the elements of $P$. We define partial orderings $\leq_i$ in $\vec{\mathbb{Z}}P$ via $a \leq_i b$ if there is a path from $a$ to $b$ in $\mathscr{F}_i P$ with equality if and only if $a = b$. Since $P$ is dimension-wise acyclic, $\leq_i$ are well-defined as partial orderings - i.e., we cannot have $a \leq_i b$ and $b \leq_i a$ for $a \neq b$.
Let $n \in \mathbb{N}$ and $a, b \in \vec{\mathbb{Z}}P_n$ with $\dim(a), \dim(b) > n$ such that $\langle a \rangle_n^+ \wedge \langle b \rangle_n^- > 0$. We have by Lemma 5.3.5 that

$$\langle a \rangle_n^+ = \sum_{y \in \Delta_n^+ a} y \quad \text{and}$$
$$\langle b \rangle_n^- = \sum_{y' \in \Delta_n^- b} y'$$

Since elements of $P$ form a basis for $\vec{\mathbb{Z}}P_n$ and $\langle a\rangle_n^+ \wedge \langle b\rangle_n^- > 0$, we get that there exists $z \in \Delta_n^+ a \cap \Delta_n^- b$, thus an edge from $a$ to $b$ in $\mathscr{F}_n P$, so $a <_n b$. ∎

**Theorem 5.4.15.** *Let $P$ be a dimension-wise acyclic regular directed complex. Then $\nu\vec{\mathbb{Z}}P \cong Mol/P$ as $\omega$-categories.*

*Proof.* Let $f : Mol/P \to \nu(\vec{\mathbb{Z}}P)$, $f = \eta_{Mol/P}; \nu(G)$ be a morphism of $\omega$-categories, where $\eta_{Mol/P}$ is the unit of the adjunction between $\lambda$ and $\nu$. We will show that $f$ defined as such is an isomorphism sending basis elements into basis elements.
Let $\operatorname{cl} x$ be an atom in $Mol/P$. By Theorem 5.4.5,

$$\eta_{Mol/P}(x) = ([\partial_0^- x]_0, [\partial_0^+ x]_0, [\partial_1^- x]_1, [\partial_1^+ x]_1, \ldots),$$

so by definition 5.4.2

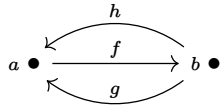
$$f(\operatorname{cl} x) = \nu(G)(\eta_{Mol/P}(x)) = (G[\partial_0^- x]_0, G[\partial_0^+ x]_0, G[\partial_1^- x]_1, G[\partial_1^+ x]_1, \ldots).$$

We aim to prove that $f(\operatorname{cl} x) = \langle x\rangle$; we will do this componentwise. Let $n = \dim x$. If $k > n$, then $G([\partial_k^\alpha x]_k) = G([x]_k) = 0$ by [50, Proposition 2.5]. If $k = n$, then $G([\partial_k^\alpha x]_k) = G([x]_k) = x \in \vec{\mathbb{Z}}P_n$. If $k = n - p$, then

$$G([\partial_{n-p}^\alpha x]_{n-p}) = \sum_{y \in \Delta^\alpha(\partial^\alpha \ldots \partial^\alpha x)} y = \sum_{y \in \Delta^\alpha_{n-p} x} y \overset{5.3.5}{=} \langle x\rangle^\alpha_{n-p}.$$

By Proposition 5.3.6, the basis of $\vec{\mathbb{Z}}P$ consists of the elements of $P$ and by Proposition 5.4.14 $\vec{\mathbb{Z}}P$ is a Steiner complex. By Theorem 5.4.13, $\langle b\rangle$ is a basis for $\nu\vec{\mathbb{Z}}P$. Therefore $f$ determines a bijection between the generating cells of $Mol/P$ and of $\nu\vec{\mathbb{Z}}P$. By [1, Proposition 16.2.12], we conclude that $f$ is an isomorphism of polygraphs. ∎

***Example* 5.4.16 —** Theorem 5.4.15 does not extend beyond dimension-wise acyclic regular directed complexes. Let $P$ be the regular directed complex encoding the 1-dimensional diagram




(5.5)

which is evidently not dimension-wise acyclic. Then $Mol/P$ is isomorphic to the free category on the directed graph (5.5). However, in $\nu\vec{\mathbb{Z}}P$, let

$$x \coloneqq \langle f\rangle \#_0 \langle g\rangle, \qquad y \coloneqq \langle f\rangle \#_0 \langle h\rangle,$$

We then have $\langle f\rangle = (a, b, f, f, 0, 0, \ldots)$, $\langle g\rangle = (b, a, g, g, 0, 0, \ldots)$ and $\langle h\rangle = (b, a, g, g,$ $0,0,\ldots)$, so we get

$$x = (a, a, f+g, f+g, 0, 0, \ldots), \quad y = (a, a, f+h, f+h, 0, 0\ldots).$$

Then

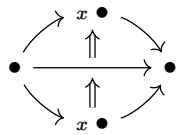
$$x \#_0 y = y \#_0 x = (a, a, 2f+g+h, 2f+g+h, 0, 0, \ldots).$$

We conclude that $\nu\vec{\mathbb{Z}}P$ is not free, so it is not isomorphic to $Mol/P$ which is free by Theorem 5.1.30.

### 5.4.3 Strongly dimension-wise acyclity

While dimension-wise acyclicity is a more manageable sufficient condition for frame-acyclicity, it does not guarantee that the $\omega$-category of molecules over $P$ consists only of *subsets* of $P$.

***Example* 5.4.17** — Let $U$ be the 2-dimensional molecule encoding the shape of the pasting diagram




and let $P$ be the result of identifying the two 0-dimensional cells marked with $x$. Then $P$ is a dimension-wise acyclic regular directed complex, and the canonical quotient map $q\colon U \to P$ is a molecule over $P$. However, $q$ is evidently not injective.

In this section, following [51], we consider a strengthening of dimension-wise acyclicity which does guarantee this property at least for regular directed complexes.

***Definition* 5.4.18** (Extended flow graph) — Let $P$ be an oriented graded poset, $k \geq -1$. The *extended $k$-flow graph of $P$* is the bipartite directed graph $\overline{\mathscr{F}}_k P$ whose

- set of vertices is
$$P = \bigcup_{i \leq k} P_i + \bigcup_{i > k} P_i,$$
- set of edges is $E_- + E_+$, where
$$E_- \coloneqq \left\{ (y,x) \mid y \in \bigcup_{i \leq k} P_i, x \in \bigcup_{i > k} P_i, y \in \operatorname{int} \partial_k^- x \right\},$$
$$E_+ \coloneqq \left\{ (y,x) \mid y \in \bigcup_{i > k} P_i, x \in \bigcup_{i \leq k} P_i, x \in \operatorname{int} \partial_k^+ y \right\},$$
where the source of $(y,x)$ is $y$ and the target is $x$.

***Definition* 5.4.19** (Strongly dimension-wise acyclic oriented graded poset) — Let $P$ be an oriented graded poset. We say that $P$ is *strongly dimension-wise acyclic* if, for all $k \geq -1$, $\overline{\mathscr{F}}_k P$ is acyclic.

**Remark 5.4.20.** *Strong dimension-wise acyclicity is essentially the same as* loop-freeness *in the sense of [51].*

**Lemma 5.4.21.** *Let $P$ be an oriented graded poset, $k \geq -1$, and suppose $x, y \in \bigcup_{i>k} P_i$. If there exists a path from $x$ to $y$ in $\mathscr{F}_k P$, then there exists a path from $x$ to $y$ in $\overline{\mathscr{F}}_k P$.*

*Proof.* Consider a path $x = x_0 \to x_1 \to \ldots \to x_m \to y$ from $x$ to $y$ in $\mathscr{F}_k P$. By definition of the $k$-flow graph, for all $i \in \{1, \ldots, m\}$, there exists $z_i \in \Delta_k^+ x_{i-1} \cap \Delta_k^- x_i$. By definition of the extended $k$-flow graph, there exist edges $x_{i-1} \to z_i$ and $z_i \to x_i$ in $\overline{\mathscr{F}}_k P$. Concatenating all the two-step paths $x_{i-1} \to z_i \to x_i$, we obtain a path from $x$ to $y$ in $\overline{\mathscr{F}}_k P$. ∎

**Proposition 5.4.22.** *Let $P$ be a strongly dimension-wise acyclic oriented graded poset. Then $P$ is dimension-wise acyclic.*

*Proof.* By Lemma 5.4.21 a cycle in $\mathscr{F}_k U$ induces a cycle in $\overline{\mathscr{F}}_k U$. ∎

**Proposition 5.4.23.** *Let $f : P \to Q$ be a local embedding of oriented graded posets. For all $k \geq -1$, $f$ induces a homomorphism $\overline{\mathscr{F}}_k f : \overline{\mathscr{F}}_k P \to \overline{\mathscr{F}}_k Q$.*

*Proof.* See [27, Proposition 5.7]. ∎

As a corollary, we obtain immediately:

**Corollary 5.4.24.** *Let $f : P \to Q$ be a local embedding of oriented graded posets. If $Q$ is strongly dimension-wise acyclic, then so is $P$.*

**Lemma 5.4.25.** *Let $U$ be a frame-acyclic molecule, $x, y \in U$. Then there exists $k \geq -1$ such that there is a path from $x$ to $y$ or a path from $y$ to $x$ in $\overline{\mathscr{F}}_k U$.*

*Proof.* Follows by the same arguments as in [51, Theorem 2.16]. ∎

**Proposition 5.4.26.** *Let $U$ be a molecule, $P$ a strongly dimension-wise acyclic oriented graded poset, and $f : U \to P$ a local embedding. Then $f$ is an inclusion.*

*Proof.* Let $x, y \in U$ with $x \neq y$ and suppose that $f(x) = f(y)$. By Corollary 5.4.24, $U$ is strongly dimension-wise acyclic. By Lemma 5.4.25 there exists $k \geq -1$ such that there is a path from $x$ to $y$ or a path from $y$ to $x$ in $\overline{\mathscr{F}}_k U$. Then by Proposition 5.4.23 $\overline{\mathscr{F}}_k f$ maps this onto a cycle in $\overline{\mathscr{F}}_k P$, a contradiction. We conclude that $f$ is injective. ∎

**Corollary 5.4.27.** *Let $P$ be a strongly dimension-wise acyclic regular directed complex. Then*

$$Mol/P = \{[U \hookrightarrow P] \mid U \subseteq P, U \text{ is a molecule}\}.$$

*Proof.* Follows from Proposition 5.4.26 together with Proposition 5.1.22. ∎

In particular, if $P$ is finite and strongly dimension-wise acyclic, it follows that $Mol/P$ has finitely many cells.

### 5.4.4 Acyclic regular directed complexes

We now consider an even stronger acyclicity condition that implies all the other conditions described in the previous subsections. Our acyclicity is essentially the same as *total loop-freeness* in [51]. As we will see, it is also related to *strong loop-freeness* in [50].

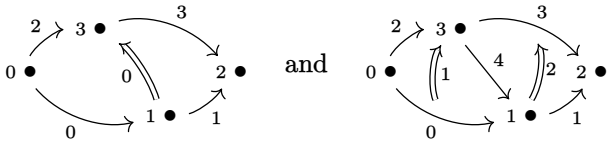
***Definition* 5.4.28 —** Let $P$ be a regular directed complex. We say that $P$ is acyclic if $\vec{\mathscr{H}}P$ is acyclic as a graph.

**Proposition 5.4.29.** *Let $P$ be an acyclic regular directed complex, $x, y \in P$, and $k \geq -1$. If there is a path from $x$ to $y$ in $\overline{\mathscr{F}}_k P$, then there is a path from $x$ to $y$ in $\vec{\mathscr{H}}P$. Consequently, $P$ is strongly dimension-wise acyclic.*

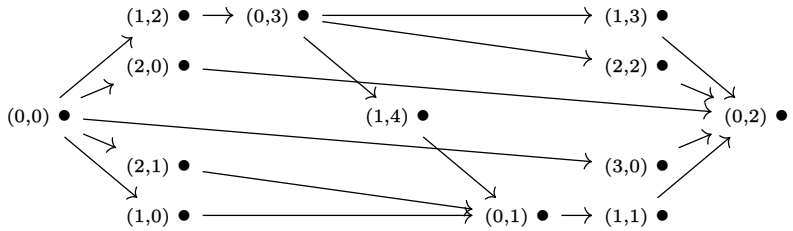
*Proof.* See [51, Proposition 2.15 and Proposition 5.2]. ∎

The following example shows that acylicity is a tighter condition that strongly dimension-wise acyclicity.

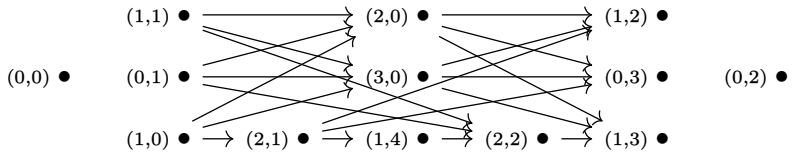
***Example* 5.4.30 —** Let $U$ be a 3-dimensional atom whose input and output boundaries encode the pasting diagrams



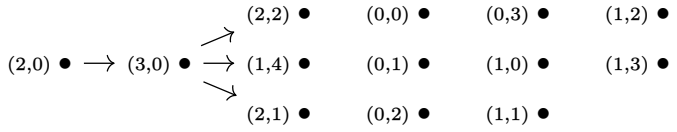

respectively, and let $(n, k)$ denote the $n$-dimensional cell labelled with $k$. Then the extended 0-flow graph $\overline{\mathscr{F}}_0 U$ is




while the extended 1-flow graph $\overline{\mathscr{F}}_1 U$ is




and the extended 2-flow graph $\overline{\mathscr{F}}_2 U$ is



all of which are acyclic. All other extended flow graphs are discrete, so $U$ is strongly dimension-wise acyclic. However, $\vec{\mathscr{H}}U$ contains the cycle

$$(0,1) \to (1,1) \to (2,0) \to (3,0) \to (2,1) \to (1,4) \to (0,1)$$

so $U$ is not acyclic.

***Definition* 5.4.31** — A basis $B$ for an augmented directed complex is *strongly loop-free* if there is a partial ordering $\leq_N$ on $B$ such that $a <_N b$ whenever $a \leq d^- b$ or $d^+ a \geq b$. An augmented complex is called a *strong Steiner complex* if it has a strongly loop-free basis.

**Proposition 5.4.32.** *Let $P$ be an acyclic regular directed complex. Then $\vec{\mathbb{Z}}P$ is a strong Steiner complex.*

*Proof.* Recall that $\vec{\mathbb{Z}}P$ has a basis given by the elements of $P$. We define a partial order via $a <_N b$ if there is a path from $a$ to $b$ in $\vec{\mathscr{H}}P$ and $a \neq b$. Since $\vec{\mathscr{H}}P$ is acyclic, this is a well-defined partial order.
Suppose first that $a \leq d^- b$. We have by Lemma 5.3.5 that $d^- b = \sum_{y \in \Delta^- b} y$ and since the elements $y \in \Delta^- b$ are part of the basis, we obtain $a \in \Delta^- b$. Thus by [25, Lemma 8.1.11] there is a path from $a$ to $b$ in $\vec{\mathscr{H}}P$, so $a <_N b$.
Similarly, if $d^+ a \leq b$, we have $d^+ a = \sum_{y \in \Delta^+ a} y$, so $b \in \Delta^+ a$. Again from Lemma [25, Lemma 8.1.11] there is a path from $a$ to $b$ in $\vec{\mathscr{H}}P$, so $a <_N b$. ∎

## 5.5 Stability under constructions and operations

In this section, we consider some operations under which the classes of molecules and regular directed complexes are closed — pastings, suspensions, Gray products, joins, and duals — and study the stability of acyclicity conditions under these operations.

***Definition* 5.5.1** (Suspension of an oriented graded poset) — Let $P$ be an oriented graded poset. The *suspension of $P$* is the oriented graded poset $\mathsf{S}P$ whose

- underlying set is $\{\mathsf{S}x \mid x \in P\} + \{\bot^+, \bot^-\}$,
- order and orientation are defined, for all $x \in \mathsf{S}P$ and $\alpha \in \{+,-\}$, by

$$\nabla^\alpha x \coloneqq \begin{cases} \{\mathsf{S}y \mid y \in \nabla^\alpha x'\} & \text{if } x = \mathsf{S}x',\ x' \in P, \\ \{\mathsf{S}y \mid y \in P_0\} & \text{if } x = \bot^\alpha, \\ \varnothing & \text{if } x = \bot^{-\alpha}. \end{cases}$$

***Definition* 5.5.2** (Gray product of oriented graded posets) — Let $P$, $Q$ be oriented graded posets. The *Gray product* of $P$ and $Q$ is the oriented graded poset $P \otimes Q$ whose

- underlying graded poset is the product $P \times Q$ of the underlying posets,
- orientation is defined, for all $(x,y) \in P \times Q$ and all $\alpha \in \{+,-\}$, by $\Delta^\alpha(x,y) \coloneqq \Delta^\alpha x \times \{y\} + \{x\} \times \Delta^{(-)^{\dim x}\alpha} y$.

Gray products determine a monoidal structure $(\mathbf{ogPos}, \otimes, 1)$ on $\mathbf{ogPos}$.

***Definition* 5.5.3** (Join of oriented graded posets) — Let $P$, $Q$ be oriented graded posets. The *join of $P$ and $Q$* is the oriented graded poset $P \star Q \coloneqq (P_\perp \otimes Q_\perp)_{/\!\!/}$. Joins determine a monoidal structure $(\mathbf{ogPos}, \star, \varnothing)$ on $\mathbf{ogPos}$.

The following collects a number of non-trivial results of [25, Chapter 7].

**Proposition 5.5.4.** *Both the classes of molecules and of regular directed complexes are closed under suspensions, Gray products and joins.*

We now move on to considering the stability of our acyclicity conditions.

**Proposition 5.5.5.** *Let $U$ be an acyclic regular directed complex. Then $\mathsf{S}U$ is acyclic.*

*Proof.* Let $U$ be an acyclic molecule. Suppose $\mathsf{S}U$ has a cycle in $\vec{\mathscr{H}}\mathsf{S}U$: $x_1 \to x_2 \to \ldots \to x_n$. Then, there are two cases.
If $x_i \neq \perp^\alpha$ for all $i$, then the cycle $x_1 \to x_2 \to \ldots \to x_n$ is a cycle in $\vec{\mathscr{H}}U$, which is a contradiction with the assumption.
Otherwise, there exists an element in the cycle, $x_i$ such that $x_i = \perp^\alpha$. Again, there are two cases. First, suppose $x_i = \perp^-$. Since $\Delta_0^- \mathsf{S}U = \{\perp^-\}$, there is no arrow pointing towards $\perp^-$. So, if $\perp^-$ appears in the cycle it can only be for $i = 1$; now let $x_1 = \perp^-$ in the cycle. Then, since $\perp^- \in \Delta_0^- \mathsf{S}U$ there is no edge $x_n \to \perp^-$ in $\vec{\mathscr{H}}\mathsf{S}U$. Now, suppose $x_i = \perp^+$. We argue similarly to the case above. Since $\Delta_0^+ \mathsf{S}U = \{\perp^+\}$, a cycle cannot start with $x_i$ and once $x_i$ is reached, the sequence ends. So we cannot have a cycle in $\vec{\mathscr{H}}\mathsf{S}U$ if $x_i = \perp^+$. Hence, $\mathsf{S}U$ is acyclic. ∎

To prove that strongly-dimension wise and dimension wise acyclicity is preserved we will need the following Lemma:

**Lemma 5.5.6.** *[25, Lemma 7.3.13] Let $U$ be an $n$-dimensional molecule. Then $\mathsf{S}U$ is an $(n+1)$-dimensional molecule with $\Delta_0^\alpha \mathsf{S}U = \{\perp^\alpha\}$ and $\Delta_k^\alpha \mathsf{S}U = \mathsf{S}\Delta_{k-1}^\alpha U$, for all $k > 0$.*

As a corollary, we obtain:

**Corollary 5.5.7.** *Let $P$ be an oriented graded poset, $k \in \mathbb{N}$. Then*

*I $x \mapsto \mathsf{S}x$ induces an isomorphism of directed graphs $\mathscr{F}_k P \xrightarrow{\sim} \mathscr{F}_{k+1}\mathsf{S}P$, restricting to an isomorphism $\mathscr{M}_k P \xrightarrow{\sim} \mathscr{M}_{k+1}\mathsf{S}P$;*

*II $x \mapsto \mathsf{S}x$ induces an embedding of directed graphs $\overline{\mathscr{F}}_k P \hookrightarrow \overline{\mathscr{F}}_{k+1}\mathsf{S}P$, whose complement is the discrete graph on $\{\perp^-, \perp^+\}$.*

**Corollary 5.5.8.** *Let $P$ be a (strongly) dimension-wise acyclic regular directed complex. Then $\mathsf{S}P$ is a (strongly) dimension-wise acyclic regular directed*

*Proof.* Immediate from Corollary 5.5.7. ∎

To prove that frame-acyclicity is also preserved, we need two additional results.

**Lemma 5.5.9.** *Let $U$ be a molecule and $\mathsf{S}U$ its suspension. Then, for any $x \in U$, $\mathrm{cl}\,\mathsf{S}x = \mathsf{S}\mathrm{cl}\,x$.*

*Proof.* By definition, $\mathrm{cl}\,\mathsf{S}x = \{\mathsf{S}y | \mathsf{S}x \geq \mathsf{S}y\} + \{\bot^-, \bot^+\}$. On the other hand, $\mathsf{S}\mathrm{cl}\,x = \mathsf{S}\{y | x \geq y\} = \{\mathsf{S}y | \mathsf{S}x \geq \mathsf{S}y\} + \{\bot^-, \bot^+\}$. ∎

**Lemma 5.5.10.** *Let $U$ be a molecule. Then $V \sqsubseteq \mathsf{S}U$ if and only if $V = \bot^-$, $V = \bot^+$ or $V = \mathsf{S}V'$, where $V' \sqsubseteq U$.*

*Proof.* We prove this by induction on the construction of $V$. If $V$ was produced by (Point), then either $V = \bot^-$ or $V = \bot^+$. If $V$ was produced by the (Atom) rule, then $V = \mathrm{cl}\,\mathsf{S}x$, for some $\mathsf{S}x \in \mathsf{S}U$. By Lemma 5.5.9, $V = \mathsf{S}\mathrm{cl}\,x$, where $x \in U$ and by Lemma 3.1.4 $\mathrm{cl}\,x \sqsubseteq U$. Now suppose $V$ was produced using the (Paste) rule. Then, $V = W \#_k Z$. By the induction hypothesis, $W = \mathsf{S}W'$ and $Z = \mathsf{S}Z'$, for $W', Z' \sqsubseteq U$. We want to prove that $V = \mathsf{S}(W' \#_{k-1} Z')$, where $W' \#_{k-1} Z' \sqsubseteq U$. Since $V = W \#_k Z$, we know that $\partial_k^+ \mathsf{S}W' = \partial_k^- \mathsf{S}Z'$. By Lemma 5.5.6, $\partial_k^+ \mathsf{S}W' = \mathsf{S}\partial_{k-1}^+ W'$ and $\partial_k^- \mathsf{S}Z' = \mathsf{S}\partial_{k-1}^- Z'$ and hence $\partial_{k-1}^+ W' = \partial_{k-1}^- Z'$. We thus know that $W' \#_k Z' \sqsubseteq U$. We claim $\mathsf{S}W' \#_k \mathsf{S}Z' = \mathsf{S}(W' \#_{k-1} Z')$.
Now, $\mathsf{S}W' \#_k \mathsf{S}Z' = \{\mathsf{S}x | \mathsf{S}x \in \mathsf{S}W' - \mathsf{S}\partial_{k-1}^+ W'\} + \{\mathsf{S}x | \mathsf{S}x \in \mathsf{S}Z' - \mathsf{S}\partial_{k-1}^- Z'\} + \{\mathsf{S}x | \mathsf{S}x \in \mathsf{S}\partial_{k-1}^+ W'\} + \{\bot^-, \bot^+\}$.
On the other hand, $\mathsf{S}(W' \#_{k-1} Z') = \{\mathsf{S}x | x \in W' \#_{k-1} Z'\} + \{\bot^-, \bot^+\} = \{\mathsf{S}x | x \in W' - \partial_{k-1}^+ W'\} + \{\mathsf{S}x | x \in Z' - \partial_{k-1}^- Z'\} + \{\mathsf{S}x | x \in \partial_{k-1}^+ W'\} + \{\bot^-, \bot^+\} = \mathsf{S}W' \#_k \mathsf{S}Z'$.
So, $V = \mathsf{S}(W' \#_{k-1} Z')$, where $W' \#_{k-1} Z' \sqsubseteq U$. ∎

**Proposition 5.5.11.** *Let $P$ be a regular directed complex that has frame-acyclic molecules. Then $\mathsf{S}P$ has frame-acyclic molecules.*

*Proof.* Follows from Lemma 5.5.10 and Corollary 5.5.7. ∎

**Proposition 5.5.12.** *Let $P$, $Q$ be acyclic oriented graded posets. Then $P \otimes Q$ is acyclic.*

*Proof.* We will prove this by contradiction. Suppose I have a cycle $(x_1, y_1) \to \ldots \to (x_n, y_n) \to (x_1, y_1)$ in $\vec{\mathscr{H}}(P \otimes Q)$. Also note that there exists an edge $(x_i, y_i) \to (x_j, y_j)$ in $\vec{\mathscr{H}}(P \otimes Q)$ when either $x_i = x_j$ or $y_i = y_j$. Assume $y_1 = y_2 = \ldots = y_n$. Then the cycle above will be a cycle in $x_1 \to \ldots \to x_n \to x_1$ in $\vec{\mathscr{H}}P$ which is a contradiction with the assumption.
Now, suppose not all $y_i$ with $0 \leq i \leq n$ are equal. We have three cases.
Case $x_1 = x_2 = \ldots = x_n$ all equal. Then, we will get a cycle $y_1 \to y_2 \to \ldots \to y_n \to y_1$ in $\vec{\mathscr{H}}Q$ if $x_1$ is even or a cycle $y_1 \to y_n \to \ldots \to y_2 \to y_1$ in $\vec{\mathscr{H}}Q$ if $x_1$ is odd which contradicts the assumption.
Case $x_n \neq x_1$. Then, necessarily $y_n = y_1$ and we have an edge from $x_n$ to $x_1$ in $\vec{\mathscr{H}}P$. We construct the following sequence. Start with $x_1$ and go through each element, $(x_i, y_i)$ of the cycle in $\vec{\mathscr{H}}(P \otimes Q)$. Every time $x_i = x_{i+1}$ in the cycle in $\vec{\mathscr{H}}(P \otimes Q)$, do nothing. When $x_i \neq x_{i+1}$, add $x_{i+1}$ to the sequence. Since $x_n \neq x_1$, this will happen at most once. Since an edge $(x_i, y_i) \to (x_{i+1}, y_{i+1})$ in $\vec{\mathscr{H}}(P \otimes Q)$ corresponds to an edge $x_i \to x_{i+1}$ in $\vec{\mathscr{H}}P$ whenever $y_i = y_{i+1}$, the sequence we constructed above is a cycle in $\vec{\mathscr{H}}P$ - contradiction with the assumption.

Case $x_n = x_1$. We treated the case when all $x_i$ are equal above. Suppose now that not all $x_i$ are equal. We proceed as in the previous case and build the following sequence. We start with $x_1$ and go through each element, $(x_i, y_i)$, of the cycle in $\vec{\mathscr{H}}(P \otimes Q)$; as before, every time $x_i = x_{i+1}$ in the cycle in $\vec{\mathscr{H}}(P \otimes Q)$, do nothing. When $x_i \neq x_{i+1}$, add $x_{i+1}$ to the sequence. From the assumption that not all $x_i$ are equal, the last step will happen at least once. Also, since $x_n = x_1$, we know that the sequence ends with $x_1$. By a similar argument with the one from the previous case, this sequence is a cycle in $\vec{\mathscr{H}}P$ - contradiction.
Therefore, $P \otimes Q$ is acyclic. ∎

**Proposition 5.5.13.** *Let $P$, $Q$ be acyclic oriented graded posets. Then, $P \star Q$ is acyclic.*

*Proof.* We prove this by contradiction. Suppose $P \star Q$ is not acyclic. Then there exists a cycle $(x_1, y_1) \to \ldots \to (x_n, y_n) \to (x_1, y_1)$ in $\vec{\mathscr{H}}(P \star Q)$. If none of the elements contains $\perp$ in either component, then the same argument as in Lemma 5.5.12 applies. Otherwise, we have the following two cases.
In the first case, $\perp$ appears in the first component - i.e., $(x_i, y_i) = (\perp, y_i)$. We make the following claim: once we reach an element of the form $(\perp, y_i)$, then all the elements following it in the sequence have $\perp$ in their first component. The argument goes as follows: once $(\perp, y_i)$ is reached, we have two ways to continue the cycle:

I $(x_{i+1}, y_{i+1}) = (x_{i+1}, y_i)$ covers $(\perp, y_i)$ with "-",

II $(x_{i+1}, y_{i+1}) = (\perp, y_{i+1})$ covers $(\perp, y_i)$ with "-".

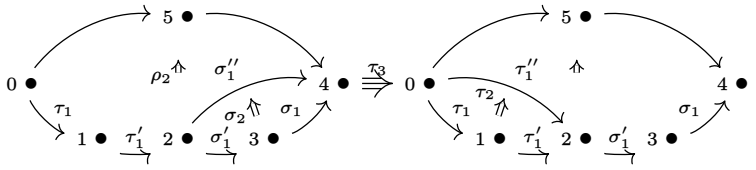
The first case is impossible to happen. Since $y_i$ is fixed, $x_{i+1}$ must cover $\perp$ with "-" in $\vec{\mathscr{H}}P_\perp$ (which is not possible by the definition of the join). In the second case, since $\dim \perp = 0$ in $\vec{\mathscr{H}}P_\perp$, having an edge $(\perp, y_i) \to (\perp, y_{i+1}$ in $\vec{\mathscr{H}}P \star Q$ means $y_{i+1}$ covers $y_i$ with "-" in $\vec{\mathscr{H}}Q_\perp$. By the argument above it follows that if $(\perp, y_i)$ appears anywhere in the cycle, then $x_1 = \perp$. Then, having a cycle $(\perp, y_1) \to \ldots \to (\perp, y_n) \to (\perp, y_1)$ in $\vec{\mathscr{H}}(P \star Q)$ means having a cycle in $\vec{\mathscr{H}}(Q)$ (because $\dim(\perp) = 0$) which is a contradiction.
In the second case, $\perp$ appears in the second component - i.e., $(x_i, y_i = (x_i, \perp)$. We have two further cases. If $x_i \neq x_{i+1}$ for some $i$, then since $\dim \perp = 0$ in $\vec{\mathscr{H}}Q_\perp$, having a cycle $(x_1, \perp) \to \ldots \to (x_n, \perp) \to (x_1, \perp)$ in $\vec{\mathscr{H}}(P \star Q)$ gives a cycle $x_1 \to \ldots \to x_n \to x_1$ in $\vec{\mathscr{H}}P$. Otherwise, let $x_1 = x_2 = \ldots = x_n$ and suppose $(x_1, y_i) = (x_1, \perp)$. We claim that if such an element appears in a sequence $(x_1, y_1) \to \ldots \to (x_1, y_n)$ then that sequence cannot be a cycle (which contradicts the assumption we started with). We argue on the parity of $\dim x_1$. If $\dim x_1$ is odd, then $(x_1, \perp)$ is covered only with "+" and the sequence ends once $(x_1, \perp)$ is reached. If $\dim x_1$ is even, then $(x_1, \perp)$ is covered only with "-" in which case a sequence can start with $(x_1, \perp)$, but it cannot be ended with $(x_1, \perp)$. Hence, $\vec{\mathscr{H}}(P \star Q)$ is acyclic if $\perp$ appears in the second component and $x_i = x_j$, for all $i, j$.
Therefore, $P \star Q$ is acyclic. ∎

**Remark 5.5.14.** *Propositions 5.5.12 and 5.5.13, in conjunction with the results of [2] and Theorem 5.4.15, can be used to show that $Mol/\_$ is compatible with Gray*

*products and joins of strict $\omega$-categories when restricted to acyclic regular directed complexes.*

Strongly dimension-wise and dimension-wise acyclicity are not preserved by Gray products and joins as the following examples show:

***Example* 5.5.15 —** [51] Let $U$ be the following 3-dimensional molecule:




Then we have the following cycle in $\mathscr{F}_2(U \otimes U)$: $(2,\tau_3) \to (2,\tau_2) \to (\sigma_1',\tau_2) \to (\sigma_1',\tau_1') \to (\sigma_2,\tau_1') \to (\sigma_2,2) \to (\tau_3,2) \to (\tau_2,2) \to (\tau_2,\sigma_1'') \to (\tau_1'',\sigma_1'') \to (\tau_1'',\sigma_2) \to (2,\sigma_2) \to (2,\tau_3)$.

***Example* 5.5.16 —** Let $U$ be the 3-dimensional molecule from Example 5.5.15. By taking $\mathsf{S}U \star U$, we obtain the following cycle in $\mathscr{F}_5(\mathsf{S}U \star U)$: $(2,\tau_3) \to (2,\tau_2) \to (\sigma_1',\tau_2) \to (\sigma_1',\tau_1') \to (\sigma_2,\tau_1') \to (\sigma_2,2) \to (\tau_3,2) \to (\tau_2,2) \to (\tau_2,\sigma_1'') \to (\tau_1'',\sigma_1'') \to (\tau_1'',\sigma_2) \to (2,\sigma_2) \to (2,\tau_3)$.

In what follows we present some results that move us closer to proving whether frame-acyclicity is preserved under taking Gray products. Namely, with the results below, we believe that if there exists a counter example to the statement above, then the example must be a submolecule that is not the Gray product of two submolecules of $U$ and $V$, respectively.

**Lemma 5.5.17.** *Let $U$ and $V$ be molecules. Then*

$$\operatorname{frdim} U \otimes V = \max\{\operatorname{frdim} U + \dim V, \dim U + \operatorname{frdim} V\}.$$

*Proof.* The atoms of $U \otimes V$ are of the form $\operatorname{cl} a \times \operatorname{cl} b$, where $\operatorname{cl} a$ is an atom in $U$ and $\operatorname{cl} b$ is an atom in $V$.
Let $\operatorname{cl} a \times \operatorname{cl} b$ and $\operatorname{cl} c \times \operatorname{cl} d$ be two distinct atom in $U \otimes V$.
We distinguish three cases.

I $a \neq c$ and $b \neq d$. Then

$$\begin{aligned} \dim(\operatorname{cl} a \times \operatorname{cl} b \cap \operatorname{cl} c \times \operatorname{cl} d) &= \dim(\operatorname{cl} a \cap \operatorname{cl} c \times \operatorname{cl} b \cap \operatorname{cl} d) \\ &= \dim(\operatorname{cl} a \cap \operatorname{cl} c) + \dim(\operatorname{cl} b \cap \operatorname{cl} d) \\ &\leq \operatorname{frdim} U + \operatorname{frdim} V \end{aligned} \tag{5.6}$$

II $a = c$ and $b \neq d$. Then:

$$\begin{aligned} \dim(\operatorname{cl} a \times \operatorname{cl} b \cap \operatorname{cl} a \times \operatorname{cl} d) &= a \times (\operatorname{cl} b \cap \operatorname{cl} d) \\ &= \dim a + \dim(\operatorname{cl} b \cap \operatorname{cl} d) \\ &\leq \dim U + \operatorname{frdim} V \end{aligned} \tag{5.7}$$

III $a \neq c$ and $b = d$. Then:

$$\begin{aligned} \dim(\operatorname{cl} a \times \operatorname{cl} b \cap \operatorname{cl} c \times \operatorname{cl} b) &= \dim(\operatorname{cl} a \cap \operatorname{cl} c \times \operatorname{cl} b) \\ &= \dim(\operatorname{cl} a \cap \operatorname{cl} c) + \dim b \\ &\leq \operatorname{frdim} U + \dim V \end{aligned} \tag{5.8}$$

Hence, we get that $\operatorname{frdim}(U \otimes V) \leq \max\{\operatorname{frdim} U + \dim V, \dim U \operatorname{frdim} V\}$ using the fact that $\operatorname{frdim} X < \dim X$ for any molecule $X$ by combining Lemma 3.2.7 and Corollary 3.3.7.

To see that the equality can be achieved suppose without loss of generality that $\max\{\operatorname{frdim} U + \dim V, \dim U + \operatorname{frdim} V\} = \operatorname{frdim} U + \dim V$. Let $\operatorname{cl} a$ and $\operatorname{cl} c$ be atoms such that $\dim(\operatorname{cl} a \cap \operatorname{cl} c) = \operatorname{frdim} U$ and $\operatorname{cl} b$ be such that $\dim(\operatorname{cl} b) = \dim V$. Then $\dim(\operatorname{cl} a \times \operatorname{cl} b \cap \operatorname{cl} c \times \operatorname{cl} b) = \operatorname{frdim} U + \dim V$.

■

**Lemma 5.5.18.** *Let $U$ and $V$ be molecules and let $n := \operatorname{frdim}(U \otimes V)$. If there is an edge from $x \otimes y \to x' \otimes y'$ in $\mathscr{M}_n(U \otimes V)$ then either $x = x'$ or $y = y'$.*

*Proof.* Suppose for a contradiction that $x \neq x', y \neq y'$. Since there is an edge from $x \otimes y \to x' \otimes y'$ in $\mathscr{M}_n(U \otimes V)$, then there exists

$$a \otimes b \in \Delta_n^+(x \otimes y) \cap \Delta_n^-(x' \otimes y').$$

Then we have by [25, Lemma 7.2.9] that there exists $i$ such that $a \in \Delta_i^+ x \cap \Delta_i^- x'$ and $b \in \Delta_{n-i}^{(-)^i} y \cap \Delta_i^{(-)^{i+1}} y'$. Then by assumption, we get $a \in \operatorname{cl} x \cap \operatorname{cl} x'$ and $b \in \operatorname{cl} y \cap \operatorname{cl} y'$. Therefore, we get $\dim a \leq \operatorname{frdim} U$ and $\dim b \leq \operatorname{frdim} V$, so $\dim(a \otimes b) \leq \operatorname{frdim} U + \operatorname{frdim} V$ and by Lemma 5.5.17, we obtain that $\operatorname{frdim} U + \operatorname{frdim} V < \operatorname{frdim}(U \otimes V) = n$, which contradicts $a \otimes b \in \Delta_n^+(x \otimes y)$. ■

**Lemma 5.5.19.** *Let $U$ and $V$ be molecules and let $n := \operatorname{frdim}(U \otimes V)$. If there is an edge from $x \otimes y \to x' \otimes y$ in $\mathscr{M}_n(U \otimes V)$, then there is an edge from $x$ to $x'$ in $\mathscr{M}_{\operatorname{frdim} U} U$.*

*Proof.* Let let $a \otimes b \in \Delta_n^+(x \otimes y) \cap \Delta_n^-(x' \otimes y)$. We have again by [25, Lemma 7.2.9] that there exists $i$ such that $a \in \Delta_i^+ x \cap \Delta_i^- x'$. If $i < \operatorname{frdim} U$, then $\dim(a \otimes b) = \dim(a) + \dim(b) < \operatorname{frdim} U + \dim V = \operatorname{frdim}(U \otimes V)$ (by Lemma 5.5.17), contradicting the fact that $a \otimes b \in \Delta_n^+(x \otimes y)$. Therefore, $i = \operatorname{frdim} U$, so there is an edge from $x$ to $x'$ in $\mathscr{M}_{\operatorname{frdim} U} U$. ■

Similarly, we may prove:

**Lemma 5.5.20.** *Let $U$ and $V$ be molecules and let $n := \operatorname{frdim}(U \otimes V)$. If there is an edge from $x \otimes y \to x \otimes y'$ in $\mathscr{M}_n(U \otimes V)$, then:*

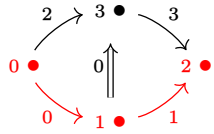
- *If $\dim x$ is even there is an edge from $y$ to $y'$ in $\mathscr{M}_{\operatorname{frdim} V} V$.*
- *If $\dim x$ is odd there is an edge from $y'$ to $y$ in $\mathscr{M}_{\operatorname{frdim} V} V$.*

**Proposition 5.5.21.** *Let $U$ and $V$ be molecules and let $n := \operatorname{frdim}(U \otimes V)$. If $\mathscr{M}_{\operatorname{frdim} U} U$ and $\mathscr{M}_{\operatorname{frdim} V} V$ are acyclic then so is $\mathscr{M}_n(U \otimes V)$.*

*Proof.* Assume for a contradiction that there exits $x_1 \otimes y_1, \ldots, x_k \otimes y_k, x_1 \otimes y_1$ a cycle in $\mathscr{M}_n(U \otimes V)$. We have by Lemma 5.5.18 and Lemma 5.5.19 that either $x_i = x_{i+1}$ or there is an edge from $x_i$ to $x_{i+1}$ in $\mathscr{M}_{\mathrm{frdim}\, U} U$ for $1 \leq i \leq k-1$. Thus if there exists $i$ such that $x_i \neq x_1$, by considering the set $x_i$, we would obtain a cycle in $\mathscr{M}_{\mathrm{frdim}\, U} U$. Otherwise, if $x_i = x_1$ for all $i$ by applying Lemma 5.5.18 and Lemma 5.5.20, we obtain a cycle $\mathscr{M}_{\mathrm{frdim}\, V} V$ with the direction of the cycle controlled by the parity of $\dim x_1$. ∎

Note that the results above are not enough to show that frame-acyclicity is preserved under taking Gray products since not all submolecules of $U \otimes V$ are of the form $u \otimes v$ with $u \sqsubseteq U$ and $v \sqsubseteq V$. Below we present an example of a submolecule which is not the Gray product of two submolecules.

***Example* 5.5.22 —** Let $U$ be the 1-dimensional atom. Then, $U \otimes U$ is the following molecule:




The input boundary of $U \otimes U$ (highlighted in red) is an example of a submolecule which is not the product of two submolecules.

# Chapter 6

# Future work

One of the questions that remains to be answered is:

**Question 6.0.1.** *Is the rewritable submolecule problem NP-hard?*

Motivated by the examples that we saw, we believe the answer to this question is affirmative for the general case. This points toward a possible complexity gap between the 3-dimensional and 4-dimensional rewrite systems. However, proving such results — even in the 4-dimensional case — remains challenging. Molecule $U_1$ from Example 4.5.1 hinted at the fact that there may be a better running time for molecules that are not frame-acyclic, but are convex - thus generalising our results. At the moment, it seems that progress on this problem depends on the following:

- Generalise the definition of obstructions to include more than two top-dimensional elements.
- Generalise the definition of obstructions to account for obstructions that may happen at lower dimensions.
- Give a precise definition to what it means for a rewrite to be "parallelisable" and for the concept of "solving an obstruction".

We believe that by developing the theory of obstructions and investigating the running time of the subdiagram matching algorithm for convex molecules, would bring us closer to answering Question 6.0.1.
Moreover, in the paradigm of rewrites as computations, studying how the lower-dimensional elements of two (or more) rewrites interact, will provide a basis for studying situations such as concurrency or deadlock in the framework of higher-dimensional rewriting. For example, let's take a look at molecule $U_1$ from Example 4.5.1. Even though $U_1$ is not frame-acyclic and we have the case of the 2-dimensional rewrites in the closures of $A$ and $B$ interacting in such a way that hints towards a dependency of $A$ on $B$ or vice-versa, we know that these rewrites can be parallelised and do not impact the order in which $A$ or $B$ can be applied - i.e., the order in which the computations are executed does not affect the outcome. However, molecule $U_2$ exhibits what we called a dependency: rewrite $A$ has to wait for $B$ to happen.

When studying the nature of higher-dimensional computations, it is important to know the nature of the computations that are going on at a lower dimension and know how they interact to avoid situations such as deadlocks or reaching a stage when a computation that ought to happen cannot occur since its "requirements" are not met.
With this in mind, our long-term goal is to study the rewrite-theoretic aspects - such as critical pair analysis, confluence and termination - of higher-dimensional rewrite systems within the framework described in this thesis. Work in this vein has been done for 2-dimensional or 3-dimensional rewrite systems in [41, 42] in the framework of computads or polygraphs [1, 11, 52] or in [9] where the authors study the confluence for string diagram rewrite systems in symmetric monoidal categories, which are a special case of the rewrite systems that can be expressed in our framework. We believe that the results about obstructions will provide valuable insights into the study of these rewrite theoretic aspects for rewrite systems of dimension greater or equal than 4.

# Bibliography


[1] D. Ara, A. Burroni, Y. Guiraud, P. Malbos, F. Métayer, and S. Mimram. "Polygraphs: From Rewriting to Higher Categories". In: London Mathematical Society Lecture Note Series. Cambridge University Press, 2025, pp. 572–583. DOI: 10.1017/9781009498968.

[2] D. Ara and G. Maltsiniotis. *Joint et tranches pour les $\infty$-catégories strictes*. Société Mathématique de France, 2020. DOI: 10.24033/msmf.473.

[3] F. Baader and T. Nipkow. *Term Rewriting and All That*. Cambridge University Press, 1998. DOI: 10.1017/CBO9781139172752.

[4] L. Babai. "Graph isomorphism in quasipolynomial time [extended abstract]". In: *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*. ACM, 2016, pp. 684–697. DOI: 10.1145/2897518.2897542.

[5] T. Beke. "Categorification, term rewriting and the Knuth–Bendix procedure". In: *Journal of Pure and Applied Algebra* 215.5 (2011), pp. 728–740. ISSN: 0022-4049. DOI: 10.1016/j.jpaa.2010.06.019.

[6] J. Bénabou. "Introduction to bicategories". In: *Reports of the Midwest Category Seminar*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1967, pp. 1–77. DOI: 10.1007/BFb0074299.

[7] F. Bonchi, F. Gadducci, A. Kissinger, P. Sobocinski, and F. Zanasi. "String diagram rewrite theory I: Rewriting with Frobenius structure". In: *Journal of the ACM (JACM)* 69.2 (2022), pp. 1–58. DOI: 10.1145/3502719.

[8] F. Bonchi, F. Gadducci, A. Kissinger, P. Sobocinski, and F. Zanasi. "String diagram rewrite theory II: rewriting with symmetric monoidal structure". In: *Mathematical Structures in Computer Science* 32.4 (2022), pp. 511–541. DOI: 10.1017/S0960129522000317.

[9] F. Bonchi, F. Gadducci, A. Kissinger, P. Sobocinski, and F. Zanasi. "String Diagram Rewrite Theory III: Confluence with and without Frobenius". In: *Mathematical Structures in Computer Science* 32.7 (2022), pp. 829–869. DOI: 10.1017/S0960129522000123.

[10] G. Bonfante and Y. Guiraud. "Polygraphic programs and polynomial-time functions". In: *Logical Methods in Computer Science* 5.2: 14 (2009), pp. 1–37. DOI: 10.2168/LMCS-5(2:14)2009.

[11] A. Burroni. "Higher-dimensional word problems with applications to equational logic". In: *Theoretical Computer Science* 115.1 (1993), pp. 43–62. DOI: 10.1016/0304-3975(93)90054-W.

[12] E. Cheng and A. Lauda. "Higher-dimensional categories: an illustrated guide book". In: (2004). Online preprint. URL: https://eugeniacheng.com/wp-content/uploads/2017/02/cheng-lauda-guidebook.pdf.

[13] M. Clavel, F. Durán, S. Eker, S. Escobar, P. Lincoln, N. Martí-Oliet, and C. Talcott. "Two decades of maude". In: *Logic, Rewriting, and Concurrency: Essays Dedicated to José Meseguer on the Occasion of His 65th Birthday* (2015), pp. 232–254. DOI: 10.1007/978-3-319-23165-5_11.

[14] J. Cockx, N. Tabareau, and T. Winterhalter. "The taming of the rew: a type theory with computational assumptions". In: *Proc. ACM Program. Lang.* 5.POPL (Jan. 2021). DOI: 10.1145/3434341.

[15] N. Corbyn, L. Heidemann, N. Hu, C. Sarti, C. Tataru, and J. Vicary. "homotopy.io: a proof assistant for finitely-presented globular $n$-categories". Online preprint arXiv:2402.13179. 2024.

[16] A. Delpeuch and J. Vicary. "The Word Problem for Braided Monoidal Categories is Unknot-Hard". In: *Electronic Proceedings in Theoretical Computer Science* 372 (2022), pp. 72–87. DOI: 10.4204/EPTCS.372.6.

[17] L. Dixon and A. Kissinger. "Open-graphs and monoidal theories". In: *Mathematical Structures in Computer Science* 23.02 (2013), pp. 308–359. DOI: 10.1017/S0960129512000138.

[18] C. Dorn. "Associative $n$-categories". Online preprint arxiv:1812.10586. PhD thesis. 2018.

[19] H. Ehrig, M. Pfender, and H. J. Schneider. "Graph-grammars: An algebraic approach". In: *14th Annual Symposium on Switching and Automata Theory (swat 1973)*. 1973, pp. 167–180. DOI: 10.1109/SWAT.1973.11.

[20] S. Forest. "Unifying notions of pasting diagrams". In: *Higher Structures* 6.1 (2022), pp. 1–79. DOI: 10.21136/HS.2022.01.

[21] Y. Guiraud. "Présentations d'opérades et systèmes de réécriture". Theses. Université Montpellier II - Sciences et Techniques du Languedoc, June 2004. URL: https://theses.hal.science/tel-00006863.

[22] Y. Guiraud. Rewriting methods in higher algebra. Thèse d'habilitation à diriger des recherches, Université Paris 7. 2019. URL: https://hal.science/tel-02161197v1.

[23] Y. Guiraud. "Termination orders for three-dimensional rewriting". In: *Journal of Pure and Applied Algebra* 207.2 (Oct. 2006), pp. 341–371. ISSN: 0022-4049. DOI: 10.1016/j.jpaa.2005.10.011.

[24] A. Hadzihasanovic. "A combinatorial-topological shape category for polygraphs". In: *Applied Categorical Structures* 28.3 (2020), pp. 419–476. DOI: 10.1007/s10485-019-09586-6.

[25] A. Hadzihasanovic. *Combinatorics of higher-categorical diagrams*. Online preprint arXiv:2404.07273, 2024.

[26] A. Hadzihasanovic. “Diagrammatic sets and rewriting in weak higher categories”. Online preprint arXiv:2007.14505. 2020.

[27] A. Hadzihasanovic and D. Kessler. “Acyclicity Conditions on Pasting Diagrams”. In: *Applied Categorical Structures* 32 (Oct. 2024). DOI: 10.1007/s10485-024-09784-x.

[28] A. Hadzihasanovic and D. Kessler. “Data Structures for Topologically Sound Higher-Dimensional Diagram Rewriting”. In: *Electronic Proceedings in Theoretical Computer Science* 380 (2023), pp. 111–127. DOI: 10.4204/eptcs.380.7.

[29] A. Hadzihasanovic and D. Kessler. “Higher-dimensional subdiagram matching”. In: *2023 38th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. IEEE. 2023, pp. 1–13. DOI: 10.1109/LICS56636.2023.10175726.

[30] A. Hatcher. *Algebraic topology*. Cambridge University Press, 2002. ISBN: 0-521-79540-0. URL: https://pi.math.cornell.edu/~hatcher/AT/AT.pdf.

[31] S. Henry. “Regular polygraphs and the Simpson conjecture”. Online preprint arXiv:1807.02627. 2018.

[32] M. Johnson. “Linear term rewriting systems are higher dimensional string rewriting systems”. In: *The Unified Computation Laboratory: Modelling, Specifications, and Tools*. Oxford University Press, Inc., 1992, pp. 101–110. ISBN: 0198536844. URL: https://web.science.mq.edu.au/~mike/pubclas.html.

[33] M. Johnson. “The combinatorics of $n$-categorical pasting”. In: *Journal of Pure and Applied Algebra* 62.3 (1989), pp. 211–225. DOI: 10.1016/0022-4049(89)90136-9.

[34] J. Kock. “Graphs, hypergraphs, and properads”. In: *Collectanea Mathematica* 67.2 (2016), pp. 155–190. DOI: 10.1007/s13348-015-0160-0.

[35] A. Krauss, C. Sternagel, R. Thiemann, C. Fuhs, and J. Giesl. “Termination of Isabelle Functions via Termination of Rewriting”. In: *Interactive Theorem Proving*. Ed. by Marko van Eekelen, Herman Geuvers, Julien Schmaltz, and Freek Wiedijk. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 152–167. ISBN: 978-3-642-22863-6. DOI: /10.1007/978-3-642-22863-6_13.

[36] Y. Lafont. “Towards an algebraic theory of Boolean circuits”. In: *Journal of Pure and Applied Algebra* 184.2 (2003), pp. 257–310. ISSN: 0022-4049. DOI: 10.1016/S0022-4049(03)00069-0.

[37] M. Makkai. “The word problem for computads”. Available at http://www.math.mcgill.ca/makkai. 2005.

[38] A. Matsui, I. Obi, G. Sabbagh, L. Torres, D. Kessler, J. F. Meleiro, and K. Muroya. “A Critical Pair Enumeration Algorithm for String Diagram Rewriting”. In: *Applied Category Theory (ACT)* (2025). to appear.

[39] J. Meseguer. “Twenty years of rewriting logic”. In: *The Journal of Logic and Algebraic Programming* 81.7 (2012). Rewriting Logic and its Applications, pp. 721–781. ISSN: 1567-8326. DOI: 10.1016/j.jlap.2012.06.003.

[40] G. L. Miller. “Graph isomorphism, general remarks”. In: *Journal of Computer and System Sciences* 18.2 (1979), pp. 128–142. ISSN: 0022-0000. DOI: 10.1016/0022-0000(79)90043-6.

[41] S. Mimram. “Computing Critical Pairs in 2-Dimensional Rewriting Systems”. In: *Leibniz International Proceedings in Informatics, LIPIcs* 6 (Apr. 2010). DOI: 10.4230/LIPIcs.RTA.2010.227.

[42] S. Mimram. “Towards 3-Dimensional Rewriting Theory”. In: *Logical Methods in Computer Science* Volume 10, Issue 2 (2014). ISSN: 1860-5974. DOI: 10.2168/lmcs-10(2:1)2014.

[43] T. Murata. “Petri nets: Properties, analysis and applications”. In: *Proceedings of the IEEE* 77.4 (1989), pp. 541–580. DOI: 10.1109/5.24143.

[44] D. Plump. “Critical pairs in term graph rewriting”. In: *Mathematical Foundations of Computer Science 1994*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1994, pp. 556–566. DOI: 10.1007/3-540-58338-6_102.

[45] D. Plump. “Essentials of Term Graph Rewriting”. In: *Electronic Notes in Theoretical Computer Science* 51 (2002). GETGRATS Closing Workshop, pp. 277–289. ISSN: 1571-0661. DOI: 10.1016/S1571-0661(04)80210-X.

[46] D. Plump. “Hypergraph rewriting: critical pairs and undecidability of confluence”. In: *Term graph rewriting: theory and practice*. 1993, pp. 201–213. URL: https://www-users.york.ac.uk/~djp10/Papers/wiley.93.pdf.

[47] J. Power. “An $n$-categorical pasting theorem”. In: *Lecture Notes in Mathematics*. Springer Nature, 1991, pp. 326–358. DOI: 10.1007/BFb0084230.

[48] D. Reutter and J. Vicary. “High-level methods for homotopy construction in associative $n$-categories”. In: *2019 34th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. IEEE. 2019, pp. 1–13. URL: https://dl.acm.org/doi/10.5555/3470152.3470214.

[49] G. Rozenberg, ed. *Handbook of graph grammars and computing by graph transformation: volume I. foundations*. USA: World Scientific Publishing Co., Inc., 1997. ISBN: 9810228848. URL: https://dl.acm.org/doi/10.5555/278918.

[50] R. Steiner. “Omega-categories and chain complexes”. In: *Homology, Homotopy and Applications* 6.1 (2004), pp. 175–200. DOI: 10.4310/HHA.2004.v6.n1.a12.

[51] R. Steiner. “The algebra of directed complexes”. In: *Applied Categorical Structures* 1.3 (1993), pp. 247–284. DOI: 10.1007/BF00873990.

[52] R. Street. “Limits indexed by category-valued 2-functors”. In: *Journal of Pure and Applied Algebra* 8.2 (1976), pp. 149–181. DOI: 10.1016/0022-4049(76)90013-X.

[53] R. Street. “Parity complexes”. In: *Cahiers de topologie et géométrie différentielle catégoriques* 32.4 (1991), pp. 315–343. URL: https://www.numdam.org/item/CTGDC_1991__32_4_315_0/.

[54] J. Vicary and A. Delpeuch. “Normalization for planar string diagrams and a quadratic equivalence algorithm”. In: *Logical Methods in Computer Science* 18 (2022). DOI: 10.46298/LMCS-18(1:10)2022.

[55] F. Wiesner, Z. Chaoui, D. Kessler, A. Pappa, and M. Karvonen. *Why quantum state verification cannot be both efficient and secure: a categorical approach.* Online preprint arXiv:2411.04767. 2024.

# Appendix A

# Paper I

A. Hadzihasanovic and D. Kessler. "Data Structures for Topologically Sound Higher-Dimensional Diagram Rewriting". In: *Electronic Proceedings in Theoretical Computer Science* 380 (2023), pp. 111–127. DOI: 10.4204/eptcs.380.7

# Data Structures for Topologically Sound Higher-Dimensional Diagram Rewriting

Amar Hadzihasanovic
[1] Tallinn University of Technology
[2] Quantinuum, 17 Beaumont Street, Oxford, UK
amar@ioc.ee

Diana Kessler
Tallinn University of Technology
diana-maria.kessler@taltech.ee

We present a computational implementation of diagrammatic sets, a model of higher-dimensional diagram rewriting that is "topologically sound": diagrams admit a functorial interpretation as homotopies in cell complexes. This has potential applications both in the formalisation of higher algebra and category theory and in computational algebraic topology. We describe data structures for well-formed shapes of diagrams of arbitrary dimensions and provide a solution to their isomorphism problem in time $O(n^3 \log n)$. On top of this, we define a type theory for rewriting in diagrammatic sets and provide a semantic characterisation of its syntactic category. All data structures and algorithms are implemented in the Python library `rewalt`, which also supports various visualisations of diagrams.

## Introduction

This article concerns the computational implementation of higher-dimensional diagrams in the sense of higher category theory, and contains some first steps in the computational complexity theory of diagrammatic rewriting in arbitrary dimensions.

Higher-dimensional rewriting, as emergent from the theory of polygraphs [5] – see [12] for a survey – is founded on an interpretation of *rewrites as directed homotopies*. A particular aim of our work is provable *topological soundness*, namely, the existence of a functorial interpretation of rewrite systems as cell complexes, and of rewrites as homotopies. This ensures that our implementation of higher-dimensional rewriting can act as a formal system for homotopical algebra and higher category theory in all generality.

With this aim, we turn to the *diagrammatic set* model [13] developed by the first author as a combinatorial alternative to polygraphs. Diagrammatic sets have a dual nature as higher-dimensional rewrite systems and "combinatorial directed cell complexes". They support a model of weak higher categories and, unlike polygraphs, are topologically sound.

Beside the formalisation of higher algebra and category theory, potential applications are manifold. *String diagram rewriting*, which is a form of 3-dimensional rewriting, is arguably the characteristic computational mechanism of applied category theory. It has been suggested [4] that even "classical" forms of rewriting are more faithfully represented as diagram rewriting: for example, term rewriting implemented as rewriting in monoidal categories with cartesian structure explicitates the "hidden costs" of copying and deleting terms. In these contexts, it is important to have a grasp on the computational complexity of the basic operations of diagram rewriting, to ensure that one's cost model for a machine operating by diagram rewriting is reasonable.

*Via* topological soundness, we also envisage applications to computational algebraic topology. Directedness of cells gives an *algebraic* grip on their pasting, which lends itself better to computation. Directed cell complexes are also equipped with an orientation on their cells, which makes them naturally suited to the computation of cellular homology.

J. Master & M. Lewis (Eds.): Fifth International
Conference on Applied Category Theory (ACT 2022).
EPTCS 380, 2023, pp. 111–127, doi:10.4204/EPTCS.380.7

### Structure of the paper

In Section 1, we present some basic data structures from the theory of diagrammatic sets, together with their formal encoding: in particular, *oriented graded posets* which are used to encode shapes of diagrams.

In Section 2, we focus on the implementation of *regular molecules*, the inductive subclass of oriented graded posets corresponding to well-formed shapes of diagrams. To construct regular molecules, we need to decide their isomorphism problem; for general oriented graded posets, this is equivalent to the graph isomorphism problem (Proposition 2.11), not known to be in P. Our main result is a solution to the isomorphism problem for regular molecules in time $O(n^3 \log n)$ (Theorem 2.19), which also gives us a canonical form, hence a unique representation of shapes of diagrams.

In Section 3, we move on to the formalisation of diagrams and diagrammatic sets. We present this in the form of a type theory DiagSet living "on top" of our implementation of shapes of diagrams: the terms, corresponding to diagrams, are "filtered by regular molecules". This allows us to define formal semantics and give a semantic characterisation of our formal system (Theorem 3.10).

### Related work

A number of type theories for higher-categorical structures of arbitrary dimension have been defined in recent years: most notably, Finster and Mimram's CaTT [8], implementing the Maltsiniotis model of weak higher categories [3], together with its "strictly associative" [10] and "strictly unital" [9] variants; and the *opetopic* type theories by Ho Thanh, Curien, and Mimram [15, 6].

The former are not particularly concerned with diagram rewriting, and focus instead on the implementation of coherent globular composition; the link to our work is tenuous. The latter have some commonality, albeit with a focus on a more restrictive class of shapes. In fact, DiagSet takes some inspiration not from one of the published opetopic type theories, but from a privately communicated variant due to Curien, which similarly rests on a "black-boxed" implementation of opetopic shapes.

Most closely related is the work by Vicary, Bar, Dorn, and others on quasistrict [2] and later associative [7, 17] $n$-categories, serving as the foundation of the homotopy.io proof assistant. While the aim is nearly the same, we believe that our framework has a number of advantages over associative $n$-categories.

From a theoretical perspective, it is only conjectural that associative $n$-categories, in general, are topologically sound or satisfy the homotopy hypothesis. They also currently lack connections with other models of higher categories and a clear functorial viewpoint. On the other hand, diagrammatic sets are topologically sound, satisfy a version of the homotopy hypothesis, and support a model of weak higher categories with concrete functorial ties to well-established models.

From a user perspective, the main point of divergence is that diagrams in associative $n$-categories have "strict units" but "weak interchange", while our diagrams have "strict interchange" but need weak units to model "nullary" inputs or outputs. For rewrite systems with many "nullary" generators, associative $n$-categories may have a practical advantage, while diagrammatic sets are otherwise favoured.

Finally, in associative $n$-categories, diagram shapes are essentially descriptions of cubical tilings, and by lack of strict interchange, each rewrite gets by default its own "layer" in the tiling. This makes it so a "local" rewrite on a portion of a diagram leads to an inefficient "global" duplication of information. Our "face poset" representation of diagrams, on the other hand, allows local rewrites to stay local, which is more efficient and will be beneficial to the parallelisability of diagram rewriting.

### Implementation

All data structures, algorithms, and systems discussed in this article were implemented by the authors as part of a Python library for higher-dimensional rewriting and algebra, called `rewalt`.[1] An example of `rewalt` code is included in Example 3.13. The library also supports various kinds of visualisation for diagrams, optionally in the form of TikZ output. All the Hasse and string diagrams in this article were generated by `rewalt` and included here with no subsequent retouching.

### Acknowledgements

This work was supported by the ESF funded Estonian IT Academy research measure (project 2014-2020.4.05.19-0001) and by the Estonian Research Council grant PSG764.

## 1 Basic data structures

***1.1.*** In the theory of diagrammatic sets, the shape of a pasting diagram is encoded by its *face poset*, recording whether a cell is located in the boundary of another cell, together with *orientation* data which specifies whether an $(n-1)$-dimensional cell is in the *input* or *output* half of the boundary of an $n$-dimensional cell. We call the mathematical structure containing these data an *oriented graded poset*. This is essentially the same as what Steiner calls a *directed precomplex* [18] and Forest an *ω-hypergraph* [11].

***1.2*** (Graded poset)**.** Let $P$ be a finite poset with order relation $\leq$ and let $P_\perp$ be $P$ extended with a least element $\perp$. We say that $P$ is *graded* if, for all $x \in P$, all directed paths from $x$ to $\perp$ in the Hasse diagram $\mathscr{H}P_\perp$, with edges going from covering to covered elements, have the same length. If this length is $n+1$, we let $\dim(x) := n$ be the *dimension* of $x$. We write $P_n$ for the subset of $n$-dimensional elements of $P$.

***1.3*** (Oriented graded poset)**.** An *orientation* on a finite poset $P$ is an edge-labelling of its Hasse diagram with values in $\{+,-\}$. An *oriented graded poset* is a finite graded poset with an orientation.

*Implementation 1.4.* If we linearly order the elements of an oriented graded poset in each dimension, each element $x$ is uniquely identified by a pair of integers $(n,k)$, where $n$ is the *dimension* of $x$, and $k$ is the *position* of $x$ in the linear ordering of $n$-dimensional elements.
We then represent an oriented graded poset as a pair $(\mathsf{face\_data}, \mathsf{coface\_data})$ of *arrays of arrays of pairs of sets of integers*, where

1. $j \in \mathsf{face\_data}[n][k][i]$ if and only if $(n-1, j)$ is covered by $(n,k)$, and
2. $j \in \mathsf{coface\_data}[n][k][i]$ if and only if $(n+1, j)$ covers $(n,k)$

with orientation $-$ $(i=0)$ or $+$ $(i=1)$. We may implement the sets of integers as sorted arrays, or another data type which supports binary search in logarithmic time. This defines a data type OgPoset.

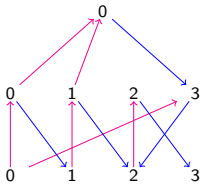
This representation is essentially an adjacency list representation of the poset's Hasse diagram, with vertices separated according to their dimension, and incoming and outgoing edges separated according to their label. If $E_P$ is the set of edges of the Hasse diagram of $P$, the OgPoset representation of $P$ takes space $O(|P| + |E_P|)$.

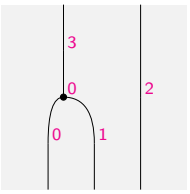
Storing both face_data and coface_data is redundant since these are uniquely determined by each other. However, most of the computations we need to perform on oriented graded posets require regular access both to faces (covered elements) and cofaces (covering elements) of a given element, so it is advantageous to be able to access them in constant time.

[1] Code: `https://github.com/ahadziha/rewalt`. Documentation: `https://rewalt.readthedocs.io`.

*Example 1.5.* Consider a diagram formed of one 2-cell with two input 1-cells and a single output 1-cell, whiskered to the right with a single 1-cell. The following are representations of its shape as

- an oriented face poset, pictured as a Hasse diagram with input faces pointing upwards (in magenta) and output faces downwards (in blue);
- a string diagram (0-cells are unlabelled, but correspond to bounded regions of the plane);
- the pair of face_data and coface_data (rows are outer array indices and columns inner array indices).




face_data:

| | | | |
|---|---|---|---|
| ([], []) | ([], []) | ([], []) | ([], []) |
| ([0], [1]) | ([1], [2]) | ([2], [3]) | ([0], [2]) |
| ([0, 1], [3]) | | | |

coface_data:

| | | | |
|---|---|---|---|
| ([0, 3], []) | ([1], [0]) | ([2], [1, 3]) | ([], [2]) |
| ([0], []) | ([0], []) | ([], []) | ([], [0]) |
| ([], []) | | | |

*Remark 1.6.* The representation of an oriented graded poset (up to isomorphism) is not unique: any permutation of the linear order on elements in each dimension leads to an equivalent representation.

***1.7.*** Many important computations are performed on *(downwards) closed subsets*, rather than the whole of an oriented graded poset. In particular, the structure of an oriented graded poset supports a purely combinatorial definition of the input and output boundary of a closed subset.

***1.8*** (Closed subsets)**.** Let $P$ be an oriented graded poset and $U \subseteq P$. We say that $U$ is *closed* if, for all $y \in U$ and $x \in P$, if $x \leq y$ then $x \in U$. The *closure* of $U$ is the subset $\mathrm{cl}U \coloneqq \{x \in P \mid \exists y \in U\ x \leq y\}$.

We let $\dim(U)$ be the maximum of $\dim(x)$ for $x \in U$, or $-1$ if $U$ is empty.

***1.9*** (Input and output boundaries)**.** Let $P$ be an oriented graded poset and $U \subseteq P$ a closed subset. For all $\alpha \in \{+,-\}$ and $n \in \mathbb{N}$, let

- $\Delta_n^\alpha U \subseteq U$ be the subset of elements $x$ such that $\dim(x) = n$ and, if $y \in U$ covers $x$, then it covers it with orientation $\alpha$;
- $\mathscr{M}_n U \subseteq U$ be the subset of elements $x$ such that $\dim(x) = n$ and $x$ is maximal in $U$ (not covered by any other element of $U$).

The *input* ($\alpha \coloneqq -$) or *output* ($\alpha \coloneqq +$) *$n$-boundary* of $U$ is the closed subset

$$\partial_n^\alpha U \coloneqq \mathrm{cl}\Big(\Delta_n^\alpha U \cup \bigcup_{k<n} \mathscr{M}_k U\Big).$$

We let $\partial_n U \coloneqq \partial_n^+ U \cup \partial_n^- U$ and omit $n$ when $n = \dim(U) - 1$. For all $x \in P$, we let $\partial_n^\alpha x \coloneqq \partial_n^\alpha \mathrm{cl}\{x\}$.

*Remark 1.10.* It is convenient to also let $\partial_{-1}^\alpha U = \partial_{-2}^\alpha U \coloneqq \emptyset$, so that $\partial^\alpha U$ is defined for all $U \subseteq P$.

*Example 1.11.* Let $U$ be the oriented face poset of Example 1.5. Then

$$\begin{aligned}
\partial_1^- U &= \{(0,0),(0,1),(0,2),(0,3),(1,0),(1,1),(1,2)\},\\
\partial_1^+ U &= \{(0,0),(0,2),(0,3),(1,2),(1,3)\},\\
\partial_0^- U &= \{(0,0)\}, \qquad \partial_0^+ U = \{(0,3)\}.
\end{aligned}$$

*Implementation 1.12.* We represent a set of elements of an OgPoset as an *array of sets of positions, indexed by dimensions*. This allows us to access the subset of elements of a given dimension in constant time. The size of arrays can be fixed to be equal to the dimension of a specific OgPoset, or dynamically adjusted to the dimension of each set of elements. Sets of positions can again be implemented as sorted arrays. This defines a data type GrSet (for *graded set*).

***1.13*** (Map of oriented graded posets)**.** A *map* $f\colon P \to Q$ of oriented graded posets is a function of their underlying sets that satisfies $\partial^\alpha_n f(x) = f(\partial^\alpha_n x)$ for all $x \in P$, $n \in \mathbb{N}$, and $\alpha \in \{+,-\}$. We call an injective map an *inclusion*. Oriented graded posets and their maps form a category **ogPos**.

*Example 1.14.* A closed subset of an oriented graded poset inherits the structure of an oriented graded poset by restriction. Its subset inclusion is an inclusion of oriented graded posets.

*Implementation 1.15.* We represent a map $f\colon P \to Q$ as an *array of arrays of pairs of integers* mapping, together with pointers source, target to OgPoset representations of $P$ and $Q$. This defines a data type OgMap. As an array of arrays, mapping has the same size of $P$'s face_data, and is defined by

$$\mathsf{mapping}[n][k] = (m, j) \text{ if and only if } f((n,k)) = (m, j).$$

This representation takes space $O(|P|)$.

## 2 Unique representation of shapes of diagrams

***2.1.*** In the theory of diagrammatic sets, shapes of diagrams form an inductively generated class of oriented graded posets, called regular *molecules* after Steiner [18].

***2.2*** (Round subset)**.** Let $U$ be a closed subset of an oriented graded poset, $n \coloneqq \dim(U)$. We say that $U$ is *round* if, for all $k < n$,

$$\partial^+_k U \cap \partial^-_k U = \partial_{k-1} U.$$

*Remark 2.3.* Roundness is called "spherical boundary" in [13].

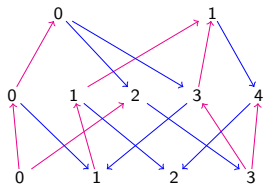
*Example 2.4.* Shapes of 2-dimensional diagrams, as oriented face posets, are round precisely when

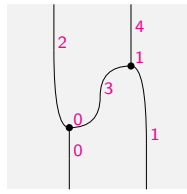
1. their string diagram representation is *connected*, and
2. all nodes of the string diagram have at least one input and one output wire.

For example, the oriented graded poset of Example 1.5 is not round: we have

$$\partial_0 U = \{(0,0),(0,3)\} \subsetneq \partial^+_1 U \cap \partial^-_1 U = \{(0,0),(0,2),(0,3)\}.$$

On the other hand, the following oriented graded poset is round:




***2.5*** (Regular molecules)**.** The class of regular molecules is generated by the following clauses.

- (Point). The terminal oriented graded poset $\bullet$ is a regular molecule.
- (Atom). Let $U, V$ be *round* regular molecules such that $\dim(U) = \dim(V)$ and, for all $\alpha \in \{+,-\}$, $\partial^\alpha U$ is isomorphic to $\partial^\alpha V$. Then $U \Rightarrow V$ is a regular molecule, where $U \Rightarrow V$ is the essentially unique oriented graded poset $U \Rightarrow V$ with the property that
  1. $U \Rightarrow V$ has a greatest element, and
  2. $\partial^-(U \Rightarrow V)$ is isomorphic to $U$, while $\partial^+(U \Rightarrow V)$ is isomorphic to $V$.
- (Paste). Let $U, V$ be regular molecules and $k < \min(\dim(U), \dim(V))$, such that $\partial^+_k U$ is isomorphic to $\partial^-_k V$. Then the pushout $U \#_k V$ of the span $\partial^+_k U \hookrightarrow U$, $\partial^+_k U \xhookrightarrow{\sim} \partial^-_k V \hookrightarrow V$ is a regular molecule.

A regular molecule is an *atom* if it has a greatest element; these are precisely the molecules whose final generating clause is (Point) or (Atom).

The *submolecule* relation $U \sqsubseteq V$ is the preorder generated by $U, V \sqsubseteq U \Rightarrow V$ and $U, V \sqsubseteq U \#_k V$.

*Comment 2.6.* The properties of regular molecules are explored in [13, Sections 1, 2]. Importantly, the following results ensure that §2.5 is a valid definition:

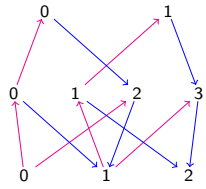
1. the category **ogPos** has pushouts of inclusions;
2. if $U$ and $V$ are isomorphic regular molecules, they are isomorphic *in a unique way*;
3. input and output boundaries of regular molecules are regular molecules;
4. if $U$ and $V$ are round, then a pair of isomorphisms between $\partial^\alpha U$ and $\partial^\alpha V$ for $\alpha \in \{+,-\}$ extends uniquely to an isomorphism between $\partial U$ and $\partial V$.

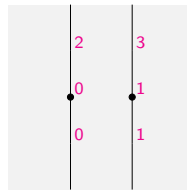
The first three imply that $U \#_k V$ is well-defined and does not depend on a choice of isomorphism between $\partial^+_k U$ and $\partial^-_k V$. The fourth implies that $U \Rightarrow V$ can be uniquely constructed by extending the isomorphisms $\partial^\alpha U \xhookrightarrow{\sim} \partial^\alpha V$ to an isomorphism $\partial U \xhookrightarrow{\sim} \partial V$, then gluing $U$ and $V$ along this isomorphism, and finally adding a greatest element with the appropriate orientation.

*Example 2.7.* Let $\mathsf{arrow} \coloneqq (\bullet \Rightarrow \bullet)$ and $\mathsf{binary} \coloneqq ((\mathsf{arrow} \#_0 \mathsf{arrow}) \Rightarrow \mathsf{arrow})$. The shape of the diagram of Example 1.5 is generated as $\mathsf{binary} \#_0 \mathsf{arrow}$, while the oriented graded poset of Example 2.4 is generated as $(\mathsf{cobinary} \#_0 \mathsf{arrow}) \#_1 (\mathsf{arrow} \#_0 \mathsf{binary})$, where $\mathsf{cobinary} \coloneqq (\mathsf{arrow} \Rightarrow (\mathsf{arrow} \#_0 \mathsf{arrow}))$.

*Remark 2.8.* As discussed in [13, §2.1], the pasting constructions $- \#_k -$ satisfy the equations of composition in strict $\omega$-categories *up to unique isomorphism*. It follows that the "same" regular molecule may be constructed in different ways. For example, letting $\mathsf{globe} \coloneqq (\mathsf{arrow} \Rightarrow \mathsf{arrow})$, we have

$$(\mathsf{globe} \#_0 \mathsf{arrow}) \#_1 (\mathsf{arrow} \#_0 \mathsf{globe}) \simeq \mathsf{globe} \#_0 \mathsf{globe} \simeq (\mathsf{arrow} \#_0 \mathsf{globe}) \#_1 (\mathsf{globe} \#_0 \mathsf{arrow}).$$




*Implementation 2.9.* We want to implement regular molecules as a subtype $\mathsf{Shape}$ of $\mathsf{OgPoset}$ with a nullary constructor $\mathsf{point}$ and partial binary constructors $\mathsf{atom}(-,-)$ and $\mathsf{paste}_k(-,-)$ for $k \in \mathbb{N}$. In order to implement the constructors, we need to be able to perform the following operations:

1. compute input and output $k$-boundaries;

2. check if a closed subset is round;
3. determine if two regular molecules are isomorphic;
4. compute the pushout of a span of inclusions.

The first, second, and fourth of these admit straightforward algorithms of low-degree polynomial time complexity, that do not rely on any special properties of regular molecules. The third problem, however, is non-trivial. Indeed, the isomorphism problem generalised to all oriented graded posets is equivalent to the *graph isomorphism* (GI) problem, which is not known to be in P; the best known algorithm, due to Babai, runs in quasipolynomial time [1].

*Remark 2.10.* As customary in this context, a graph is a *simple* graph (no loops or multiple edges).

***Proposition 2.11 —*** *The isomorphism problem for oriented graded posets is* GI*-complete.*

*Proof.* Deciding isomorphism of oriented graded posets is equivalent to deciding isomorphism of their Hasse diagrams with $\{+,-\}$-labelled edges. The isomorphism problem for edge-labelled finite graphs is an instance of the isomorphism problem for finite relational structures, which is GI-complete [16].

Conversely, a directed graph can be represented by its "oriented incidence poset": the 0-dimensional elements are the vertices, the 1-dimensional elements are the edges, the only input face of an edge is its source, and the only output face of an edge is its target. Two directed graphs are isomorphic if and only if their oriented incidence posets are isomorphic. Since GI reduces to the isomorphism problem for directed graphs, it reduces to the isomorphism problem for 1-dimensional oriented graded posets. ∎

Nevertheless, in the special case of regular molecules, we can do much better. Our strategy is to describe a deterministic *traversal algorithm*, where the traversal order depends only on the intrinsic structure of a regular molecule as an oriented graded poset and not on its representation.

Given $U,V$ : OgPoset representing regular molecules, we traverse both $U$ and $V$, and then *reorder* their elements in each dimension according to their traversal order. If $U',V'$ : OgPoset are the reordered versions of $U,V$, we then have

$$U \simeq V \text{ if and only if } U' \equiv V'.$$

We will show that, with this strategy, we can solve the isomorphism problem for regular molecules in time $O(n^3 \log n)$. A more precise upper bound is given in Theorem 2.19 below.

In addition to solving the isomorphism problem for regular molecules, the traversal order gives us a canonical form for regular molecules in OgPoset form. If we implement the constructors of Shape in such a way that they always produce an OgPoset in traversal order, we obtain that

$$\text{for all } U,V : \mathsf{Shape},\ U \simeq V \text{ if and only if } U \equiv V,$$

that is, we have a *unique representation* for shapes of diagrams.

The algorithm is described in Figure 1. At each iteration of the main loop (line 4), the current state is fully described by the *stack* – including its top element, the *focus* – and by the list of *marked* elements.

***Lemma 2.12 —*** *Let $V$ be an item on the stack. Then $V$ is a regular molecule. If $W$ is below $V$ on the stack, then $V$ is a proper subset of $W$.*

*Proof.* Initially, the stack only contains $U$, which is a regular molecule by assumption. Assume, inductively, that the statement is true at the beginning of the current iteration with focus $V$, and that a set $V'$ is pushed onto the stack at the end. Then either

1. $V' = \partial^\alpha V$ for some $\alpha \in \{+,-\}$, or

```
    procedure TRAVERSE(U : regular molecule)
        marked ← []
        stack ← [U]
        while stack is not empty do
 5:         focus ← top of stack
            dim ← dim(focus)
            if focus ⊆ marked then
                pop focus from top of stack
            else
10:             if ∂⁻focus ⊈ marked then
                    push ∂⁻focus to top of stack
                else
                    if focus = cl{x} for some x then
                        append x to marked
15:                     pop focus from top of stack
                        if ∂⁺focus ⊈ marked then
                            push ∂⁺focus on top of stack
                    else
                        y ← first item of dimension dim − 1 in marked such that
20:                         y has an unmarked input coface in focus
                        x ← unique input coface of y in focus
                        push cl{x} on top of stack
        return marked
```

Figure 1: The traversal algorithm.

2. $V' = \mathrm{cl}\{x\}$ for some $x \in V$.

In both cases, $V'$ is a regular molecule and a proper subset of $V$ (hence also of each item below $V$), under the assumption that $V$ is a regular molecule. ∎

*Remark 2.13.* In fact, any $V$ that appears on the stack is either $\partial_k^- U$, which we call "$U$-linked", or it is $\mathrm{cl}\{x\}$ or $\partial_k^\alpha x$, which we call "$x$-linked", for some $x \in U$. In the latter case, $V$ is *round*, which implies that it is also *pure* [13, Lemma 1.35]: its maximal elements all have the same dimension.

***Lemma 2.14 —*** *Suppose $V$ is on the stack. Then all elements of $V$ must be marked before any item below $V$ is accessed, or before any proper superset of $V$ becomes the focus.*

*Proof.* By Lemma 2.12, as long as $V$ is on the stack, only $V$ and its proper subsets can be on top. It follows that, for a proper superset of $V$ to be the focus, $V$ must be popped from the stack at the end of an iteration where $V$ is the focus. There are only two ways this can happen:

- $V$ was already fully marked before the current loop iteration, or
- $\partial^- V$ was fully marked and $V = \mathrm{cl}\{x\}$ for some $x$ which is marked at the current loop iteration.

In both cases, $\partial^- V$ was already fully marked before the current loop iteration. In the latter case, if $\partial^+ V$ is already fully marked, then $V = \{x\} \cup \partial^- V \cup \partial^+ V$ is also fully marked. Otherwise, $\partial^+ V \subsetneq V$ gets pushed onto the stack to replace $V$, and must be popped before any superset of $V$ becomes the focus. By the same case distinction, whenever $\partial^+ V$ is popped, either

- it was fully marked, in which case $V$ was fully marked, or

- it is of the form $\mathrm{cl}\{y\}$ for some $y$ which is marked at the current loop iteration.

Either way, since all regular molecules satisfy the *globularity* property $\partial^\alpha(\partial^+ V) = \partial^\alpha(\partial^- V) \subseteq \partial^- V$, we know that $\partial^+ V$, hence $V$, is fully marked at the end of the iteration, and nothing is added to the stack. ∎

***Lemma 2.15 —*** *Any subset $V$ of $U$ can be pushed onto the stack at most once.*

*Proof.* Suppose $V$ is pushed onto the stack. As long as $V$ is on the stack, any subsequent addition to the stack must be a proper subset of $V$, so it cannot be equal to $V$.

If $V$ is popped from the stack, by Lemma 2.14, it must be fully marked before any item below it is accessed. Since the algorithm checks if a set is fully marked before pushing it onto the stack, $V$ can never appear again. ∎

***Lemma 2.16 —*** *Let $V$ be the focus, $n \coloneqq \dim(V)$. Then either $V$ is fully marked, or there exists an $n$-dimensional element of $V$ which is unmarked.*

*Proof.* First, we prove a weaker result: either $V$ is fully marked, or there exists a *maximal* element of $V$ which is unmarked.

Let $x \in V$ be marked. At some prior iteration, $\mathrm{cl}\{x\}$ must have been the focus, and by Lemma 2.14, in order for $V$ to become the focus, $\mathrm{cl}\{x\}$ must have been fully marked as well. Because

$$V = \bigcup_{k \leq n} \mathrm{cl}\mathscr{M}_k V = \bigcup_{k \leq n} \bigcup_{x \in \mathscr{M}_k V} \mathrm{cl}\{x\},$$

it follows that $V$ is fully marked if and only if its maximal elements are all marked.

Now, $V$ has one of the two forms in Remark 2.13. If $V$ is of the second form, its maximal elements all have the top dimension, so we only need to consider the case $V = \partial^-_k U$.

At the start of the algorithm, $U, \dots, \partial^-_0 U$ are all consecutively added to the stack. So $\partial^-_k U$ becomes the focus either at this stage, in which case *all* its elements are unmarked, or after $\partial^-_{k-1} U$ is fully marked. In the latter case, any maximal element of $\partial^-_k U$ of dimension strictly smaller than $k$ also belongs to $\partial^-_{k-1} U$. ∎

***Theorem 2.17 —*** *The traversal algorithm is correct: given a regular molecule $U$, it terminates returning a unique linear ordering of the elements of $U$.*

*Proof.* As a particular case of Lemma 2.14, $U$ must be fully marked before the stack is emptied. Therefore, the algorithm either terminates after all elements have been traversed, or it does not terminate.

To prove that the algorithm does always terminate, it suffices to show that, unless all elements are already marked, it always finds an element to mark. First of all, observe that, from any state, the algorithm first goes through the following sequence of steps:

1. popping all fully marked subsets from the top of the stack;
2. once it reaches a subset which is not fully marked, successively pushing its lower-dimensional input boundaries that are not fully marked onto the stack.

At the end of this sequence, we always reach a state in which the focus $V$ is not fully marked, but $\partial^- V$ is fully marked. Let us call such a $V$ a *proper* focus.

We proceed by induction on dimension and proper subsets of a proper focus. If $\dim(V) = 0$, since a 0-molecule always consists of a single element, $V = \{x\}$, and $x$ gets marked at the current iteration.

Let $n \coloneqq \dim(V)$. By Lemma 2.16, there is an unmarked $x \in V_n$. If $V = \mathrm{cl}\{x\}$, then $x$ is marked at the current iteration, and we are done. Otherwise, we prove that there always exists a pair $(y,x)$ where $x \in V_n$ is unmarked, and $y$ is a marked input face of $x$. By [13, Lemma 1.16] applied to $V$, the coface $x$ is unique given $y$, so among such pairs we can pick the one where $y$ comes *earliest* in the list of marked elements, and this selects a unique $x$.

Let $x \in V_n$ be unmarked. By a dual version of [*ibid.*, Lemma 1.37], there exists a sequence

$$y_0 \to x_0 \to \ldots \to y_m \to x_m = x$$

where $y_0 \in \Delta^-_{n-1}V$, $x_i \in V_n$, $y_i$ is an input face of $x_i$, and $y_{i+1}$ is an output face of $x_i$. Since $V$ is a proper focus, $y_0$ is marked. Let $k$ be the smallest index such that $x_k$ is unmarked; because $x_m$ is unmarked, such a $k$ exists. Then $x_i$ is marked for all $i < k$, hence $\mathrm{cl}\{x_i\}$ is also marked. It follows that $y_k \in \partial^+ x_{k-1}$ is marked, and the pair $(y_k, x_k)$ satisfies our requirement.

Thus, the algorithm will find a unique $x \in V_n$ and push $\mathrm{cl}\{x\}$ onto the stack. The next proper focus will necessarily be a proper subset of $V$, and we conclude by the inductive hypothesis. ∎

***2.18.*** In what follows, for a fixed regular molecule $U$, we let $|E_n|$ be the number of edges between $n$ and $(n-1)$-dimensional elements in the Hasse diagram of $U$, and we let

$$|U_{\max}| \coloneqq \max_n |U_n|, \qquad |E_{\max}| \coloneqq \max_n |E_n|.$$

***Theorem 2.19 —*** *The traversal algorithm admits an implementation running in time*

$$O\big(|U|^2(|E_{\max}| \cdot \log|E_{\max}| + |U_{\max}| \cdot \log|U_{\max}|)\big).$$

*Proof.* First of all, we represent any closed set on the stack with its graded set of maximal elements. To initialise the algorithm, we only need to compute the maximal elements of $U$. This can be done in time $O(|U|)$ by going through the elements of $U$ and checking if their set of cofaces is empty.

Next, let us find an upper bound for the number of iterations of the main loop (line 4). Let $V$ be a set on the stack, $n \coloneqq \dim(V)$. Then $V$ can become the focus

- at most once before pushing $\partial^- V$ onto the stack (line 11),
- at most once before pushing $\mathrm{cl}\{y\}$ onto the stack for each $y \in V_n$ (line 22), and
- at most once to be popped from the stack (line 8),

after which, by Lemma 2.15, it can never appear again. Thus, the number of loop iterations with $V$ as focus is bounded by $|V_n| + 2$.

By Remark 2.13, every set $V$ on the stack is either "$U$-linked" or "$x$-linked" for some $x \in U$. There are $(\dim(U)+1)$ many $U$-linked focusses and $(2\dim(x)+1)$ many $x$-linked focusses. Then

- the number of loop iterations with $U$-linked focusses is bounded by $|U| + 2\dim(U) + 2$, and
- for each $x$, the number of iterations with $x$-linked focusses is bounded by $|\mathrm{cl}\{x\}| + 4\dim(x) + 2$.

Since there are $|U|$ elements, $|\mathrm{cl}\{x\}| \le |U|$, and $\dim(x) \le \dim(U)$, we have a coarse upper bound of $(|U|+1)(|U|+4\dim(U)+2)$ on the total number of iterations, which is $O(|U|^2)$.

Next, in our implementation, we split the list of marked elements into three objects: a list `order` (for the total traversal order), an array of lists `grorder` (for the traversal order split by dimension), and a graded set `marked` (for the set of marked elements).

Consider a single loop iteration with focus $V$, $n \coloneqq \dim(V)$.

**(Line 7).** By Lemma 2.16, to check if $V$ is fully marked, it suffices to check whether $V_n \subseteq \mathsf{marked}_n$. Since both are sorted arrays of integers, they can be compared in time linear in $|V_n| + |\mathsf{marked}_n|$, which is $O(|U_n|)$. At this stage, we may also record the unmarked $n$-dimensional elements of $V$ in a sorted array unmarked without affecting the complexity.

**(Line 10).** To compute the maximal elements of $\partial^- V$ and $\partial^+ V$, we may use different strategies depending on whether $V$ is "$U$-linked" or not.

If $V = \partial^-_n U$, we compute the $(n-1)$-dimensional elements of $\partial^- V = \partial^-_{n-1} U$ simply by going through the elements of $U_{n-1}$ and checking which ones have empty sets of output cofaces, in time $O(|U_{n-1}|)$. Lower-dimensional maximal elements are shared between $V$ and $\partial^- V$, so we may then point from the latter to the former, at no extra cost.

If $V$ is not $U$-linked, $V$ and its boundaries are pure, so the set of maximal elements of $\partial^\alpha V$ is equal to $\Delta^\alpha V$, and each of its elements is covered by an element of $V_n$. To compute it, we add all the input and output faces of all $x \in V_n$ to sets in_faces and out_faces, respectively, then use the relations $\Delta^- V = \mathsf{in\_faces} \setminus \mathsf{out\_faces}$ and $\Delta^+ V = \mathsf{out\_faces} \setminus \mathsf{in\_faces}$.

There are $O(|E_n|)$ faces of elements of $V_n$, and we can sort in_faces and out_faces, remove duplicates, and compute their difference in time $O(|E_n| \cdot \log |E_n|)$.

At this stage, we also create an associative array candidates as follows: whenever $x \in V_n$ is in unmarked, and $y$ is an input face of $x$, we add the position of $x$ as a value to candidates, indexed by the position of $y$. We then sort the indices of candidates. This also takes time $O(|E_n| \cdot \log |E_n|)$ so it does not affect the overall complexity.

**(Lines 10, 16).** By the same reasoning applied to line 7, checking if $\partial^- V$ and $\partial^+ V$ are fully marked takes time $O(|U_{n-1}|)$.

**(Line 14).** If $V_n$ has a single element that we mark, adding it to order and grorder takes constant time with an appropriate implementation of lists. Adding it to marked takes $O(|U_n|)$.

**(Lines 19—21).** To select the next focus we traverse $\mathsf{grorder}_{n-1}$ starting from the first item and search for each item in the indices of candidates until we find a hit $y$. This takes time $O(|U_{n-1}| \cdot \log |U_{n-1}|)$ in the worst case. The next focus will be $\mathrm{cl}\{x\}$, where $x$ is the value corresponding to index $y$.

Overall, the worst-case complexity is $O(|U_n| + |E_n| \cdot \log |E_n| + |U_{n-1}| \cdot \log |U_{n-1}|)$. Using the bounds $|U_n|, |U_{n-1}| \leq |U_{\max}|$ and $|E_n| \leq |E_{\max}|$, and multiplying by our bound on the number of iterations, we conclude. ∎

## 3 A type theory for higher-dimensional rewriting

***3.1.*** We rapidly go through the definitions of diagrammatic sets and some related notions. For a thorough treatment, we refer to [13, Section 4 and onwards], and to [14, Section V] for diagrammatic complexes as presentations of higher-dimensional theories.

***3.2*** (Diagrammatic set)**.** Let $\dot{\odot}$ (to be read *atom*) be a skeleton of the full subcategory of **ogPos** on the atoms of every dimension. A *diagrammatic set* is a presheaf on $\dot{\odot}$. Diagrammatic sets and their morphisms of presheaves form a category $\dot{\odot}$**Set**.

***3.3.*** We identify $\dot{\odot}$ with a full subcategory $\dot{\odot} \hookrightarrow \dot{\odot}$**Set** via the Yoneda embedding. With this identification, we use morphisms in $\dot{\odot}$**Set** as our notation for both elements and structural operations of a diagrammatic set $X$:

- $x \in X(U)$ becomes $x \colon U \to X$, and
- for each map $f \colon V \to U$ in $\dot{\odot}$, $X(f)(x) \in X(V)$ becomes $f;x \colon V \to X$.

The embedding $\odot \hookrightarrow \odot\mathbf{Set}$ extends along pushouts of inclusions to the full subcategory of **ogPos** on the regular molecules.

***3.4*** (Diagrams and cells)**.** Let $X$ be a diagrammatic set and $U$ a regular molecule. A *diagram of shape $U$ in $X$* is a morphism $x\colon U \to X$. A diagram is a *cell* if $U$ is an atom. For all $n \in \mathbb{N}$, we say that $x$ is an *$n$-diagram* or an *$n$-cell* when $\dim(U) = n$.

If $U$ decomposes as $U_1 \#_k U_2$, we write $x = x_1 \#_k x_2$ for $x_i := \iota_i; x$, where $\iota_i$ is the inclusion $U_i \hookrightarrow U$ for $i \in \{1,2\}$. Let $\iota_k^\alpha\colon \partial_k^\alpha U \hookrightarrow U$ be the inclusions of the $k$-boundaries of $U$. The *input $k$-boundary* of $x$ is the diagram $\partial_k^- x := \iota_k^-; x$ and the *output $k$-boundary* of $x$ is the diagram $\partial_k^+ x := \iota_k^+; x$. We write $x\colon y^- \Rightarrow y^+$ to express that $\partial_k^\alpha x = y^\alpha$ for each $\alpha \in \{+,-\}$.

***3.5*** (Diagrammatic complex)**.** For each $n \in \mathbb{N}$, let $\odot_n$ be the full subcategory of $\odot$ on the atoms of dimension $\leq n$, and let $\odot_{-1}$ be the empty subcategory. The restriction functor $\odot\mathbf{Set} \to \mathrm{PSh}(\odot_n)$ has a left adjoint; let $\sigma_{\leq n}$ be the comonad induced by this adjunction. The *$n$-skeleton* of a diagrammatic set $X$ is the counit $\sigma_{\leq n}X \to X$. For all $k \leq n$, the $k$-skeleton factors uniquely through the $n$-skeleton of $X$.

A *diagrammatic complex* is a diagrammatic set $X$ together with a set $\mathscr{X} = \sum_{n\in\mathbb{N}} \mathscr{X}_n$ of *generating* cells such that, for all $n \in \mathbb{N}$,

$$\begin{array}{ccc} \bigsqcup_{x\in\mathscr{X}_n} \partial U(x) & \hookrightarrow & \bigsqcup_{x\in\mathscr{X}_n} U(x) \\ \downarrow{\scriptstyle (\partial x)_{x\in\mathscr{X}_n}} & & \downarrow{\scriptstyle (x)_{x\in\mathscr{X}_n}} \\ \sigma_{\leq n-1}X & \hookrightarrow & \sigma_{\leq n}X \end{array}$$

is a pushout in $\odot\mathbf{Set}$, where $U(x)$ denotes the shape of $x$. A diagrammatic complex is *finite* if $\mathscr{X}$ is finite.

***3.6*** (Support-based diagrammatic complex)**.** Each cell in a diagrammatic complex $(X, \mathscr{X})$ is uniquely of the form $(p\colon U \twoheadrightarrow V, x\colon V \to X)$, where $p$ is a surjective map of atoms and $x \in \mathscr{X}$. We let $\mathsf{supp}(p,x) := x$, the *support* of $(p,x)$.

A *support-based diagrammatic complex* is the quotient of a diagrammatic complex by the relations

$$x \sim y \text{ if and only if } \mathsf{supp}(\iota; x) = \mathsf{supp}(\iota; y) \text{ for all inclusions of atoms } \iota\colon V \hookrightarrow U, \tag{1}$$

for all atoms $U$ and cells $x, y\colon U \to X$. We let $\odot\mathbf{Cpx}_{fsb}$ denote the category of finite, support-based diagrammatic complexes with morphisms of their underlying diagrammatic sets.

***3.7.*** We define a dependent type theory for diagrammatic sets – more precisely, for finite, support-based diagrammatic complexes – that relies on an underlying unique representation of regular molecules and their maps, treated as a "black box". Of course, in the previous section we have provided such an implementation and proved that it is computationally feasible. Nevertheless, it is useful to separate its abstract properties from the implementation details.

***3.8*** (DiagSet)**.** Let $\mathbb{V}$ be an infinite set of variables. We define a type theory DiagSet as follows.

**Terms.** A term $t$ is a pair of a regular molecule $U$, the *shape* of $t$, and a function $t\colon U \to \mathbb{V}$. We write $t/_U$ to express that $t$ is a term of shape $U$. Maps $p\colon U \to V$ act on terms by precomposition: if $t/_V$ is a term, then $p^*t := (p;t)/_U$. In particular, we let $\partial_k^\alpha t := (\iota_k^\alpha; t)/_{\partial_k^\alpha V}$ for all $k \in \mathbb{N}$ and $\alpha \in \{+,-\}$.

**Types.** A type $A$ is either $\varnothing$ or an expression $t \Rightarrow s$ where $t, s$ are terms. We may annotate a term $t$ of shape $U$ with the type $A := \varnothing$ if $U \equiv \bullet$, and $A := \partial^- t \Rightarrow \partial^+ t$ otherwise.

**Contexts.** A context $\Gamma$ is a list $x_1 : A_1, \ldots, x_n : A_n$ of typed variables. We consider two contexts to be equal if they are equal up to a permutation. If $x : A$ is a typed variable, we say that $x$ has *shape* $\bullet$ if $A \equiv \varnothing$, and $U \Rightarrow V$ if $A \equiv t/_U \Rightarrow s/_V$. We write $x/_U : A$ to express that $x : A$ has shape $U$.

**Substitutions.** A substitution $\sigma$ is a list $x_1 \mapsto t_1, \dots, x_n \mapsto t_n$ of assignments of terms to variables. We consider two substitutions to be equal if they are equal up to a permutation.
**Judgments.** We consider three kinds of judgments:

- $\Gamma \vdash$ meaning that $\Gamma$ is a well-formed context,
- $\Gamma \vdash t$ meaning that $t$ is a well-formed term in context $\Gamma$, and
- $\Delta \vdash \sigma : \Gamma$ meaning that $\sigma$ is a well-formed substitution from context $\Delta$ to context $\Gamma$.

The inference rules of DiagSet are the following. We use $\langle\rangle$ to indicate the empty list.

**Rules for contexts.**

$$\frac{}{\langle\rangle \vdash}\ \textsf{init} \qquad \frac{\Gamma \vdash}{\Gamma, x : \varnothing \vdash}\ \textsf{pt} \qquad \frac{\Gamma \vdash t/_U : r^- \Rightarrow r^+ \quad \Gamma \vdash s/_V : r^- \Rightarrow r^+ \quad U, V \ \textsf{round}}{\Gamma, x : t \Rightarrow s \vdash}\ \textsf{gen}$$

(where $x \in \mathbb{V}$ is fresh)

**Rules for terms.**

$$\frac{\Gamma \vdash \qquad (x/_V : A) \in \Gamma \qquad U \ \textsf{atom} \qquad p \colon U \twoheadrightarrow V \ \textsf{surjective}}{\Gamma \vdash p^* \widehat{x}/_U}\ \textsf{cell}$$

$$\frac{\Gamma \vdash t/_U \qquad \Gamma \vdash s/_V \qquad \partial^+_k t \equiv \partial^-_k s}{\Gamma \vdash (t \#_k s)/(U \#_k V)}\ \textsf{paste}_k, \quad k < \min(\dim(U), \dim(V))$$

**Rules for substitutions.**

$$\frac{\Gamma \vdash}{\Gamma \vdash \langle\rangle : \Gamma}\ \textsf{id} \qquad \frac{\Delta \vdash \sigma : \Gamma \qquad \Gamma, x : s/_U \Rightarrow r/_V \vdash \qquad \Delta \vdash t/_U \Rightarrow V : s[\sigma] \Rightarrow r[\sigma]}{\Delta \vdash \langle \sigma, x \mapsto t \rangle : (\Gamma, x : s \Rightarrow r)}\ \textsf{ext}$$

In the rules cell and paste, the terms $\widehat{x}$ and $t \#_k s$ are defined as follows:

- $\widehat{x}$ is the unique term of shape $V$ which sends the greatest element of $V$ to $x$, and, if $A \equiv t \Rightarrow s$, is equal to $t$ on $\partial^- V$ and to $s$ on $\partial^+ V$;
- $t \#_k s$ is the unique term of shape $U \#_k V$ that is equal to $t$ on $U \hookrightarrow (U \#_k V)$ and to $s$ on $V \hookrightarrow (U \#_k V)$.

The side conditions for gen and paste ensure that this is well-defined.

To define the action $t[\sigma]$ of a well-formed substitution $\sigma$ on a term $t$, we extend $\sigma$ to a function $\mathbb{V} \to \mathbb{V}$ as follows: for all $x \in \mathbb{V}$, if $(x \mapsto t/_U) \in \sigma$, we let $\sigma(x) \coloneqq t(\top)$, where $\top$ is the greatest element of $U$; otherwise, $\sigma(x) \coloneqq x$. Then $t[\sigma]$ is the composite of $t \colon U \to \mathbb{V}$ and $\sigma \colon \mathbb{V} \to \mathbb{V}$. Note that this is well-defined because a well-formed substitution assigns to each variable a term whose shape is an atom.

***3.9*** (Syntactic category)**.** The syntactic category $\mathbf{Ctx}[\mathsf{DiagSet}]$ has

- well-formed contexts $\Gamma$ as objects, and
- well-formed substitutions as morphisms from $\Delta$ to $\Gamma$,

with the obvious composition of substitutions, and empty substitutions as identities.

***Theorem 3.10 —*** *The category* $\mathbf{Ctx}[\mathsf{DiagSet}]^{\mathrm{op}}$ *is equivalent to* $\odot\mathbf{Cpx}_{fsb}$.

*Sketch of proof.* We define an encoding $\mathsf{enc}$ of finite support-based diagrammatic complexes, diagrams, and morphisms as contexts, terms, and substitutions. Given $(X, \mathscr{X})$, we pick an injective function $\mathsf{name}\colon \mathscr{X} \to \mathbb{V}$, assigning unique variable names to the generating cells of $X$.

For all diagrams $d\colon U \to X$, we define a term $\mathsf{enc}(d)$ as follows: for all $x \in U$, we let $\mathsf{enc}(d)(x)$ be equal to $\mathsf{name}(\mathsf{supp}(d|_{\mathrm{cl}\{x\}}))$. Since $(X, \mathscr{X})$ is support-based, $\mathsf{enc}(d) \equiv \mathsf{enc}(d')$ implies $d = d'$.

Let $n$ be the greatest dimension in which $\mathscr{X}_n$ is non-empty, and pick a linear ordering $x_1, \dots, x_{m_k}$ of $\mathscr{X}_k$ for all $k \leq n$. We let $\mathsf{enc}(X, \mathscr{X}) \coloneqq \Gamma_0, \dots, \Gamma_n$, where

$$\Gamma_k \coloneqq \mathsf{name}(x_1) : \mathsf{enc}(\partial^- x_1) \Rightarrow \mathsf{enc}(\partial^+ x_1), \dots, \mathsf{name}(x_{m_k}) : \mathsf{enc}(\partial^- x_{m_k}) \Rightarrow \mathsf{enc}(\partial^+ x_{m_k}).$$

By the construction of $X$ as a colimit of its generating cells, any map $X \to Y$ is uniquely determined by what it does on $\mathscr{X}$. Given a map $f\colon (X, \mathscr{X}) \to (Y, \mathscr{Y})$ in $\odot\mathbf{Cpx}_{fsb}$, we let $\mathsf{enc}(f)$ be the substitution

$$\langle \mathsf{name}_X(x) \mapsto \mathsf{enc}_Y(f(x)) \rangle_{x \in \mathscr{X}}.$$

Conversely, we define an interpretation $[\![-]\!]$ of well-formed contexts, terms, and substitutions by induction on inference rules of $\mathsf{DiagSet}$. At each step the interpretation $[\![\Gamma]\!]$ of a well-formed context is a support-based diagrammatic complex with one generator $[\![\widehat{x}]\!]$ of shape $U$ for each variable $x/U$ in $\Gamma$.

- ($\mathsf{init}$) The interpretation of the empty context is the initial diagrammatic set.
- ($\mathsf{pt}$) Suppose $[\![\Gamma]\!]$ is defined. The interpretation of $\Gamma, x : \varnothing$ is the coproduct $[\![\Gamma]\!] + \bullet$. The interpretation of $\widehat{x}$ is the inclusion $\bullet \hookrightarrow [\![\Gamma]\!] + \bullet$.
- ($\mathsf{gen}$) Suppose $[\![\Gamma]\!]$ and $[\![t/U]\!], [\![s/V]\!]$ are defined. The interpretation of $\Gamma, x : t \Rightarrow s$ is the pushout of $\partial[\![\widehat{x}]\!]\colon \partial(U \Rightarrow V) \to [\![\Gamma]\!]$ and $\partial(U \Rightarrow V) \hookrightarrow (U \Rightarrow V)$, quotiented by the equations (1), where $\partial[\![\widehat{x}]\!]$ is equal to $[\![t]\!]$ on $\partial^-(U \Rightarrow V)$ and to $[\![s]\!]$ on $\partial^+(U \Rightarrow V)$.
- ($\mathsf{cell}$) Suppose $[\![\Gamma]\!]$ is defined and has a generating cell $[\![\widehat{x}]\!]$. The interpretation of $p^*\widehat{x}$ is $p; [\![\widehat{x}]\!]$.
- ($\mathsf{paste}_k$) Suppose $[\![\Gamma]\!]$ and $[\![t]\!], [\![s]\!]$ are defined with $\partial_k^+[\![t]\!] = [\![\partial_k^+ t]\!] = [\![\partial_k^- s]\!] = \partial_k^-[\![s]\!]$. The interpretation of $t \#_k s$ is the diagram $[\![t]\!] \#_k [\![s]\!]$.
- ($\mathsf{id}$) The interpretation of the empty substitution in context $\Gamma$ is the identity of $[\![\Gamma]\!]$.
- ($\mathsf{ext}$) Suppose $[\![\sigma]\!]$ and $[\![t]\!]$ are defined, where $[\![\widehat{x}]\!]$ and $[\![t]\!]$ both have the same shape $U$. By the construction of $[\![\Gamma, x]\!]$ as a colimit of $[\![\Gamma]\!]$ and $U$, the pair of $[\![\sigma]\!]\colon [\![\Gamma]\!] \to [\![\Delta]\!]$ and $[\![t]\!]\colon U \to [\![\Gamma]\!]$ induces a unique morphism $[\![\sigma, x \mapsto t]\!]\colon [\![\Gamma, x]\!] \to [\![\Delta]\!]$.

It is routine to check that $\mathsf{enc}$ and $[\![-]\!]$ define contravariant functors between $\odot\mathbf{Cpx}_{fsb}$ and $\mathbf{Ctx}[\mathsf{DiagSet}]$, and that they are each other's inverse up to natural isomorphism. ∎

*Remark 3.11.* The proof of Theorem 3.10 gives a semantic characterisation of well-formed terms as diagrams in a diagrammatic set. An immediate consequence is that the following rule is admissible:

$$\frac{\Gamma \vdash t/V \qquad p\colon U \to V \ \mathsf{map}}{\Gamma \vdash p^* t/U}\ \mathsf{pb}$$

where $p$ is an arbitrary map of regular molecules.

*Comment 3.12.* A sticking point in our type theory is the fact that $\mathsf{cell}$ is parametrised by an arbitrary surjective map of atoms $p$. This is necessary to access the "weak units" and degenerate cells which in our framework are needed, among other things, to model nullary operations in an algebraic theory.

In practice, however, this is the one point in which the underlying implementation of regular molecules and their maps has to be explicitly accessed in order to define $p$ and its domain. To avoid this, in a practical implementation, we want to include explicitly some extra admissible rules, corresponding to the application of useful maps that are parametric in their codomain.

In particular, we want to explicitly include

- the trivial case $p \equiv \mathrm{id}_U$:

$$\frac{\Gamma \vdash \qquad (x/_U : A) \in \Gamma}{\Gamma \vdash \widehat{x}/_U}\ \mathsf{cell}',$$

- *unit* rules, modelling [13, §4.16]:

$$\frac{\Gamma \vdash t/_U}{\Gamma \vdash \mathsf{unit}(t) \coloneqq \tau^*(t) : t \Rightarrow t}\ \mathsf{unit},$$

- *left and right unitor rules*, modelling [*ibid.*, §4.17]:

$$\frac{\Gamma \vdash t/_U \qquad V \sqsubseteq \partial^- U\ \mathsf{round}}{\Gamma \vdash \mathsf{lunitor}_V(t) \coloneqq (\ell^-_{V \hookrightarrow U})^* t}\ \mathsf{lunitor} \qquad \frac{\Gamma \vdash t/_U \qquad V \sqsubseteq \partial^+ U\ \mathsf{round}}{\Gamma \vdash \mathsf{runitor}_V(t) \coloneqq (r^-_{V \hookrightarrow U})^* t}\ \mathsf{runitor}$$

  where $V$ can be specified, for example, by the set of positions of its maximal elements.

We may also have extra rules for *simplex and cube degeneracy* maps and for *cube connection* maps, in the case where $U$ is an oriented simplex or cube as in [*ibid.*, §3.33]. All of these are implemented as diagram methods in `rewalt`.

*Example 3.13.* As an example, we give a presentation in DiagSet of the theory of a left-unital binary operation, together with its implementation in `rewalt`. In the framework of diagrammatic sets, a many-sorted "monoidal theory" is presented by a diagrammatic complex with a single 0-cell; this is analogous to the way a monoidal category is a bicategory with a single 0-cell. The sorts are generating 1-cells, the basic operations are generating 2-cells, and "oriented equations" are generating 3-cells.

First, we add a single 0-cell $x$ and a single sort $a$.

$$\dfrac{\dfrac{\dfrac{}{\langle\rangle \vdash}\,\mathsf{init}}{x : \varnothing}\,\mathsf{pt}}{x : \varnothing \vdash \widehat{x}}\,\mathsf{cell}' \qquad \dfrac{x : \varnothing \vdash \widehat{x} \qquad x : \varnothing \vdash \widehat{x}}{\dfrac{x : \varnothing, a : \widehat{x} \Rightarrow \widehat{x} \vdash}{x : \varnothing, a : \widehat{x} \Rightarrow \widehat{x} \vdash \widehat{a}}\,\mathsf{cell}'}\,\mathsf{gen}$$

```
import rewalt
Lun = rewalt.DiagSet()
x = Lun.add('x')
a = Lun.add('a', x, x)
```

Let $\Gamma \coloneqq x : \varnothing, a : \widehat{x} \Rightarrow \widehat{x}$. We add a binary operation $m$.

$$\dfrac{\dfrac{\dfrac{\Gamma \vdash \widehat{a} \qquad \Gamma \vdash \widehat{a}}{\Gamma \vdash \widehat{a} \#_0 \widehat{a}}\,\mathsf{paste}_0 \qquad \Gamma \vdash \widehat{a}}{\Gamma, m : \widehat{a} \#_0 \widehat{a} \Rightarrow \widehat{a} \vdash}\,\mathsf{gen}}{\Gamma, m : \widehat{a} \#_0 \widehat{a} \Rightarrow \widehat{a} \vdash \widehat{m}}\,\mathsf{cell}'$$

```
m = Lun.add('m', a.paste(a), a)
```

Let $\Gamma' \coloneqq \Gamma, m : \widehat{a} \#_0 \widehat{a} \Rightarrow \widehat{a}$. We produce a weak unit on $x$ and add a nullary operation $u$.

$$\dfrac{\dfrac{\dfrac{\Gamma' \vdash \widehat{x}}{\Gamma' \vdash \mathsf{unit}(\widehat{x})}\,\mathsf{unit} \qquad \Gamma' \vdash \widehat{a}}{\Gamma', u : \mathsf{unit}(\widehat{x}) \Rightarrow \widehat{a} \vdash}\,\mathsf{gen}}{\Gamma', u : \mathsf{unit}(\widehat{x}) \Rightarrow \widehat{a} \vdash \widehat{u}}\,\mathsf{cell}'$$

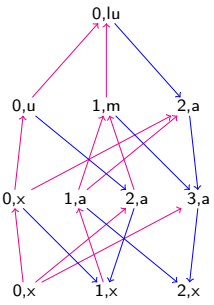
```
u = Lun.add('u', x.unit(), a)
```

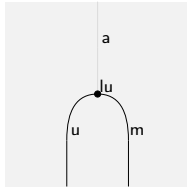
Let $\Gamma'' \coloneqq \Gamma', u : \mathsf{unit}(\widehat{x}) \Rightarrow \widehat{a}$. We produce a left unitor 2-cell on $a$, and add an "oriented equation" exhibiting the fact that $u$ is a left unit for $m$.

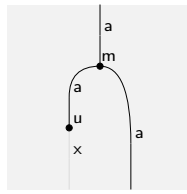
$$\dfrac{\dfrac{\dfrac{\Gamma'' \vdash \widehat{u} \qquad \Gamma'' \vdash \widehat{a}}{\Gamma'' \vdash \widehat{u} \#_0 \widehat{a}}\,\mathsf{paste}_0 \qquad \Gamma'' \vdash \widehat{m}}{\Gamma'' \vdash (\widehat{u} \#_0 \widehat{a}) \#_1 \widehat{m}}\,\mathsf{paste}_1 \qquad \dfrac{\Gamma'' \vdash \widehat{a}/_{\mathsf{arrow}}}{\Gamma'' \vdash \mathsf{lunitor}_{\partial^- \mathsf{arrow}}(\widehat{a})}\,\mathsf{lunitor}}{\Gamma'', lu : ((\widehat{u} \#_0 \widehat{a}) \#_1 \widehat{m}) \Rightarrow \mathsf{lunitor}_{\partial^- \mathsf{arrow}}(\widehat{a}) \vdash}\,\mathsf{gen}$$

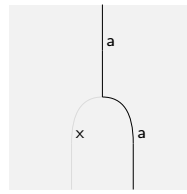
```
7 lu = Lun.add('lu', u.paste(a).paste(m), a.lunitor())
```

The following is a representation of *lu* as a term of DiagSet, that is, an oriented graded poset labelled with names, together with string diagram representations of *lu*, its input boundary, and its output boundary, and the `rewalt` code that generated them.




```
8  lu.hasse(tikz=True)
9  lu.draw(bgcolor='gray!10', tikz=True)
10 lu.input.draw(bgcolor='gray!10', tikz=True)
11 lu.output.draw(bgcolor='gray!10', tikz=True)
```

*Comment 3.14.* Provided we have a unique underlying representation of shapes, as described in Section 2, every term of DiagSet also has a unique representation. In this sense, terms of DiagSet are "noncomputational": all the computation, which consists exclusively of computing and matching shapes, happens under the hood before a term is even created, so the equality theory of terms is trivial.

This is intended. Rather than a computational theory in itself, DiagSet is intended as a *substrate for computational theories* according to the paradigm of higher-dimensional rewriting. A term $t : r^- \Rightarrow r^+$ can be seen as a rewrite of the "lower-dimensional" term $r^-$ to the term $r^+$, and the extension of $t$ via the $\mathsf{paste}_k$ rules establishes how the rewrite can happen in a wider context. In this sense, every well-formed context in DiagSet contains its own internal computational theory on terms of each dimension.

*Remark 3.15.* While "rewrites in context" can be built with the $\mathsf{paste}_k$ rules, this is quite impractical. In practice, one wants to start from a diagram and apply a generating rewrite directly to a subdiagram. This is modelled by *pasting along a subdiagram* [13, §4.12] in the theory of diagrammatic sets.

Pasting along a subdiagram is implemented in `rewalt` with methods `to_inputs` and `to_outputs`. These invoke a procedure for recognising subdiagrams, which currently uses a quite naive algorithm. The issue of recognising subdiagrams deserves further study, so we leave it to future work.

## Conclusions and outlook

We have provided a formal implementation of "plain" diagrammatic sets. An obvious next step is the formalisation of *weakly invertible* cells, and then of diagrammatic sets with weak composites, a model of

weak higher categories [13, Sections 5, 6]. This is in fact part of `rewalt`, but still lacks a formal analysis.

In addition, we still have a limited range of high-level methods for handling weak units. We may want, for example, flexible higher-dimensional versions of "Mac Lane triangle" rules for shuffling weak units around. Development of these methods, and others tailored to specific applications, will likely go hand in hand with practical experience in the use of `rewalt` as a proof assistant.

To conclude, we have only scratched the surface of the algorithm and complexity theory of diagram rewriting in higher dimensions. In particular, we have not yet studied the problem of searching for a subdiagram within another diagram, whose solution is essential to any form of fully automated or assisted diagram rewriting. We plan to tackle this problem in future work.

## References


[1] L. Babai (2016): *Graph isomorphism in quasipolynomial time [extended abstract]*. In: *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, ACM, pp. 684–697, doi:10.1145/2897518.2897542.

[2] K. Bar & J. Vicary (2017): *Data structures for quasistrict higher categories*. In: *2017 32nd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, IEEE, doi:10.1109/lics.2017.8005147.

[3] T. Benjamin, E. Finster & S. Mimram (2021): *Globular weak $\omega$-categories as models of a type theory*. arXiv:2106.04475.

[4] G. Bonfante & Y. Guiraud (2009): *Polygraphic programs and polynomial-time functions*. *Logical Methods in Computer Science* 5(2), doi:10.2168/lmcs-5(2:14)2009.

[5] A. Burroni (1993): *Higher-dimensional word problems with applications to equational logic*. *Theoretical Computer Science* 115(1), pp. 43–62, doi:10.1016/0304-3975(93)90054-w.

[6] P.L. Curien, C. Ho Thanh & S. Mimram (2019): *Syntactic approaches to opetopes*.

[7] C. Dorn (2018): *Associative n-categories*. Ph.D. thesis, University of Oxford.

[8] E. Finster & S. Mimram (2017): *A type-theoretical definition of weak $\omega$-categories*. In: *2017 32nd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, IEEE, doi:10.1109/lics.2017.8005124.

[9] E. Finster, D. Reutter & J. Vicary (2020): *A type theory for strictly unital $\infty$-categories*. arXiv:2007.08307.

[10] E. Finster, A. Rice & J. Vicary (2021): *A type theory for strictly associative infinity categories*. arXiv:2109.01513.

[11] S. Forest (2019): *Unifying notions of pasting diagrams*.

[12] Y. Guiraud (2019): Rewriting methods in higher algebra. Thèse d'habilitation à diriger des recherches, Université Paris 7.

[13] A. Hadzihasanovic (2020): *Diagrammatic sets and rewriting in weak higher categories*. arXiv:2007.14505.

[14] A. Hadzihasanovic (2021): *The smash product of monoidal theories*. In: *2021 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, IEEE, doi:10.1109/lics52264.2021.9470575.

[15] C. Ho Thanh, P.L. Curien & S. Mimram (2019): *A Sequent Calculus for Opetopes*. In: *2019 34th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, IEEE, doi:10.1109/lics.2019.8785667.

[16] G.L. Miller (1979): *Graph isomorphism, general remarks*. *Journal of Computer and System Sciences* 18(2), pp. 128–142, doi:10.1016/0022-0000(79)90043-6.

[17] D. Reutter & J. Vicary (2019): *High-level methods for homotopy construction in associative n-categories*. In: *2019 34th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, IEEE, doi:10.1109/lics.2019.8785895.

[18] R. Steiner (1993): *The algebra of directed complexes*. *Applied Categorical Structures* 1(3), pp. 247–284, doi:10.1007/bf00873990.

# Appendix B

# Paper II

A. Hadzihasanovic and D. Kessler. "Higher-dimensional subdiagram matching". In: *2023 38th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. IEEE. 2023, pp. 1–13. DOI: 10.1109/LICS56636.2023.10175726

# Higher-Dimensional Subdiagram Matching

Amar Hadzihasanovic
[1] Department of Software Science
Tallinn University of Technology
[2] Quantinuum, 17 Beaumont Street
Oxford OX1 2NA, United Kingdom

Diana Kessler
Department of Software Science
Tallinn University of Technology



***Abstract*—Higher-dimensional rewriting is founded on a duality of rewrite systems and cell complexes, connecting computational mathematics to higher categories and homotopy theory: the two sides of a rewrite rule are two halves of the boundary of an $(n+1)$-cell, which are diagrams of $n$-cells. We study higher-dimensional diagram rewriting as a mechanism of computation, focussing on the matching problem for rewritable subdiagrams within the combinatorial framework of diagrammatic sets. We provide an algorithm for subdiagram matching in arbitrary dimensions, based on new results on layerings of diagrams, and derive upper bounds on its time complexity. We show that these superpolynomial bounds can be improved to polynomial bounds under certain acyclicity conditions, and that these conditions hold in general for diagrams up to dimension 3. We discuss the challenges that arise in dimension 4.**

## Introduction

Higher-dimensional rewriting [1], [2] is founded, as a field, on the observation that different varieties of rewrite systems are instances of *directed cell complexes* in different dimensions. These, in turn, can be seen as presentations of higher or monoidal categories and groupoids, situating objects traditionally associated with syntactic and quantitative aspects of computation in the same universe as objects traditionally associated with semantic and logical aspects.

Informally, a directed cell complex is an object assembled from "directed $n$-cells", models of topological $n$-balls whose boundary, an $(n-1)$-sphere, is subdivided into an *input* and an *output* half, modelling the left and right-hand sides of a rewrite rule. A 1-dimensional directed cell complex is a directed graph, which is the same as an abstract rewrite system up to interpretive nuances. A 2-dimensional directed cell complex is the presentation with "oriented equations" of a category; when the category has a single object, it is the presentation of a *monoid*, also known as a string rewrite system. One dimension higher, we find presentations with oriented equations of monoidal categories, which subsume, *via* functorial semantics [3], term rewrite systems such as presentations of algebraic theories.

The duality of rewrite systems and higher structures has, so far, been leveraged mostly on the side of higher algebra, for example in the study of homotopically coherent presentations of algebras [4] or the development of proof assistants for homotopy theory and higher category theory [5]. We would like to argue that, also *as a mechanism of computation*, higher-dimensional rewriting has some unique characteristics which may be significant in fundamental computer science, yet have remained underexplored.

One of these is a kind of *uniformity of data and computations*. In most models, computations or executions are objects of different nature from the data that they manipulate: for example, sequences of configurations as opposed to terms or strings of characters. There is no *internal* way to manipulate computations as data, that does not require encoding in some external "reference" machine, to which the same consideration applies. In higher-dimensional rewriting, on the other hand, data is given in the form of a *diagram* (sometimes called a *pasting diagram*) of $n$-cells, and a computation is then embodied by *a diagram of* $(n+1)$*-cells*, which is naturally the data of a computation one dimension above. This suggests that higher-dimensional rewriting may, for example, be a natural framework for the formal study of simulations of machines by other machines.

Another point of interest is that a higher-dimensional rewrite system presents not only the admissible data and computations, but also *the space in which computations happen*. The possibility of parallelism, or whether the data is accessed as a stack or in free order, become internal topological features rather than externally imposed constraints. This is particularly interesting in relation to quantum models where the topological features themselves drive the computation [6].

This article is an effort to establish some fundamental facts about higher-dimensional rewriting as a mechanism of computation. In particular, we try to answer the following question: *is a machine that operates by higher-dimensional rewriting a "reasonable" machine, according to the standards of computational complexity theory?* More precisely: is the obvious cost model that attributes constant cost to each rewrite step a "reasonable" cost model?

The basic computational step of any such machine may be described as follows. The machine has a list $(r_i)_{i=1}^m$ of rewrite rules, which are $(n+1)$-cells, and whose input boundaries $(\partial^- r_i)_{i=1}^m$ and output boundaries $(\partial^+ r_i)_{i=1}^m$ are $n$-dimensional diagrams. Given an $n$-dimensional diagram $t$ as input, the machine tries to match one of the input boundaries with a rewritable *subdiagram* of $t$. If it finds a match with $\partial^- r_i$ for some $i \in \{1, \dots, m\}$, it substitutes $\partial^+ r_i$ for the match in $t$; otherwise it stops.

Accepted for LICS 2023

Evidently, our question is answered in the affirmative for such a machine if and only if the *subdiagram matching problem* is feasible in dimension $n$, which we read as: *admits a (preferably low-degree) polynomial-time algorithm with respect to a reasonable size measure for diagrams*. Since a cognate problem such as subgraph matching is notoriously NP-complete, it is not at all obvious that this should be true.

With this in mind, in this article we study the higher-dimensional subdiagram matching problem. As a first step, we must fix a particular model of higher-dimensional diagrams. We adopt the *diagrammatic set* model defined by the first author [7] after a combinatorial approach to diagrams started by Steiner [8]. This model combines expressiveness with the property of *topological soundness*: diagrams admit a functorial interpretation as homotopies in CW complexes. Furthermore, it uses inductive data structures that are suitable for computation; we began a study of their algorithms and complexity in [9].

Our main contribution is an algorithm for subdiagram matching in arbitrary dimension, relying on new combinatorial results on *layerings* of diagrams (decompositions into layers containing each a single cell of highest dimension). Only one stage in this algorithm takes superpolynomial time in the worst case. We show that this can be avoided under certain acyclicity conditions on diagrams, which hold in general up to dimension 3. We derive that subdiagram matching is feasible up to dimension 3, which should cover most current applications of higher-dimensional rewriting. We then discuss the case of dimension 4 and higher, showing through a counterexample why there is no obvious patch to the algorithm. We leave the question of feasibility in higher dimensions open.

*Related work*

Complexity-theoretic aspects of higher-dimensional rewriting in the proper sense were considered by Bonfante and Guiraud in [10], but this work focussed only on a particular class of 2-dimensional rewrite systems.

Other works focus on *string diagram rewriting*, which is connected to higher-dimensional rewriting in that both of them have semantics in higher and monoidal categories, but the flavour is altogether different: models of string diagram rewrite systems are typically 2-dimensional but have extra structure which makes the diagrams "graph-like", and it can be convenient to reflect that in the data structures. Works in this vein include the series by Bonchi, Gadducci, Kissinger, Sobocinski, and Zanasi [11], [12], [13] with a focus on rewriting-theoretic questions, and articles by Delpeuch and Vicary [14], [15] with a more complexity-theoretic focus.

*Structure of the article*

Section I presents the combinatorial framework together with the data structures used to represent diagrams. It also provides an improved complexity upper bound for the diagram isomorphism problem. Section II defines rewritable subdiagrams and their matching problem, which can be split into subproblems. It then deals with the first subproblem, which is to match the shape of a diagram in another diagram, irrespective of whether it is a subdiagram. Section III deals with the second subproblem, which is to recognise which matches are, in fact, rewritable subdiagrams. It presents a theory of *layerings* and *orderings* of diagrams as a way to an algorithm solving this problem in any dimension. This algorithm has only a superpolynomial upper bound, but it is shown that it can be improved to a polynomial bound under certain acyclicity conditions. Section IV shows that these acyclicity conditions hold automatically up to dimension 3, and the algorithms can be further simplified in low dimensions. It then discusses an example in dimension 4 which highlights why the strategies that work in low dimensions do not have obvious extensions to higher dimensions.

Proofs of combinatorial results are attached in the appendix. A full development will be presented in a forthcoming technical monograph [16].

## I. The data structures

***1.*** This section is for the largest part a recap of [9]. We refer the reader there for more details.

***2.*** In the framework of diagrammatic sets, a diagram $t$ is specified by the data of

1) its *shape* $U$,
2) a labelling $t\colon U \to \mathbb{V}$ in a set of variables.

The shape of a diagram records its cells, together with the information of which $(n-1)$-dimensional cells are located in the input or output half of the boundary of an $n$-dimensional cell. This is similar to the data of an abstract polytope or polytopal complex, but comes with additional orientation data. We present these data in the form of an *oriented graded poset*.

***3*** (Covering relation)**.** Let $P$ be a finite poset with order relation $\leq$. Given elements $x, y \in P$, we say that $y$ *covers* $x$ if $x < y$ and, for all $y' \in P$, if $x < y' \leq y$ then $y' = y$.

***4*** (Hasse diagram)**.** Let $P$ be a finite poset. The *Hasse diagram* of $P$ is the directed acyclic graph $\mathscr{H}P$ whose

- set of vertices is the underlying set of $P$, and
- for all vertices $x, y$, there is an edge from $y$ to $x$ if and only if $y$ covers $x$ in $P$.

***5*** (Graded poset)**.** Let $P$ be a finite poset. We say that $P$ is *graded* if, for all $x \in P$, all maximal paths starting from $x$ in $\mathscr{H}P$ have the same length.

***6*** (Dimension of an element)**.** Let $P$ be a graded poset and $x \in P$. The *dimension* of $x$ is the length $\dim x$ of a maximal path starting from $x$ in $\mathscr{H}P$. For each $U \subseteq P$ and $n \in \mathbb{N}$, we write $U_n \coloneqq \{x \in U \mid \dim x = n\}$.

***7*** (Oriented graded poset)**.** Let $P$ be a finite poset. An *orientation* on $P$ is an edge-labelling of $\mathscr{H}P$ with values in $\{+, -\}$. An *oriented graded poset* is a graded poset $P$ together with an orientation on $P$.

***8*** (Faces and cofaces)**.** Let $P$ be an oriented graded poset, $x \in P$, $\alpha \in \{+,-\}$. The set of *input* ($\alpha = -$) or *output* ($\alpha = +$) *faces* of $x$ is

$$\Delta^\alpha x \coloneqq \{y \in P \mid x \text{ covers } y \text{ with orientation } \alpha\}.$$

The set of *input* ($\alpha = -$) or *output* ($\alpha = +$) *cofaces* of $x$ is

$$\nabla^\alpha x \coloneqq \{y \in P \mid y \text{ covers } x \text{ with orientation } \alpha\}.$$

We let $\Delta x \coloneqq \Delta^- x \cup \Delta^+ x$ and $\nabla x \coloneqq \nabla^- x \cup \nabla^+ x$.

***9*** (Oriented Hasse diagram)**.** Let $P$ be an oriented graded poset. The *oriented Hasse diagram of* $P$ is the directed graph $\vec{\mathscr{H}} P$ whose

- set of vertices is the underlying set of $P$, and
- for all vertices $x, y$, there is an edge from $y$ to $x$ if and only if $y \in \Delta^- x$ or $x \in \Delta^+ y$.

***10*.** To represent an oriented graded poset $P$, we linearly order $P_n$ in each dimension $n$, so that each element $x \in P$ is uniquely identified by a pair of integers $(n, k)$, where $n \coloneqq \dim x$ and $k$ is the position of $x$ in the linear order on $P_n$. Then we represent $P$ as a pair $(\mathsf{face\_data}, \mathsf{coface\_data})$ of arrays of arrays of pairs of sets of integers, where

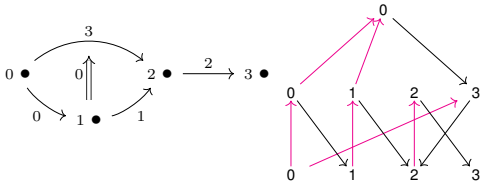
1) $j \in \mathsf{face\_data}[n][k][i]$ iff $(n-1, j) \in \Delta^{\alpha(i)}(n, k)$,
2) $j \in \mathsf{coface\_data}[n][k][i]$ iff $(n+1, j) \in \nabla^{\alpha(i)}(n, k)$;

here the index $i \in \{0, 1\}$ is used to encode pairs, $\alpha(0) \coloneqq -$ and $\alpha(1) \coloneqq +$. Sets of integers may be implemented as any data type supporting binary search in logarithmic time. Note that this representation is redundant: face_data and coface_data can be reconstructed from each other.

This is essentially an adjacency list representation of $\mathscr{H} P$, with vertices separated according to their dimension, and incoming and outgoing edges separated according to their label. It is not unique: any permutation of the dimension-wise linear orders produces an equivalent representation.

*Example 11.* The following are representations of the same 2-dimensional diagram shape as

- a typical drawing of a pasting diagram;
- an oriented Hasse diagram, with input edges in pink, and dimension increasing from bottom to top;
- the pair of face_data and coface_data (rows are outer array indices, increasing from top to bottom, and columns inner array indices, increasing from left to right).




face_data:

| | | | |
|---|---|---|---|
| $(\{\},\{\})$ | $(\{\},\{\})$ | $(\{\},\{\})$ | $(\{\},\{\})$ |
| $(\{0\},\{1\})$ | $(\{1\},\{2\})$ | $(\{2\},\{3\})$ | $(\{0\},\{2\})$ |
| $(\{0,1\},\{3\})$ | | | |

coface_data:

| | | | |
|---|---|---|---|
| $(\{0,3\},\{\})$ | $(\{1\},\{0\})$ | $(\{2\},\{1,3\})$ | $(\{\},\{2\})$ |
| $(\{0\},\{\})$ | $(\{0\},\{\})$ | $(\{\},\{\})$ | $(\{\},\{0\})$ |
| $(\{\},\{\})$ | | | |

***12*** (Closure of a subset)**.** Let $P$ be a poset, $U \subseteq P$. The *closure of* $U$ is the subset $\operatorname{cl} U \coloneqq \{x \in P \mid \exists y \in U\, x \le y\}$. We say that $U$ is *closed* if $U = \operatorname{cl} U$.

***13*** (Dimension of a subset)**.** Let $U$ be a closed subset of a graded poset. The *dimension* of $U$ is the integer

$$\dim U \coloneqq \begin{cases} \max\{\dim x \mid x \in U\} & \text{if } U \text{ is inhabited,} \\ -1 & \text{if } U \text{ is empty.} \end{cases}$$

***14*** (Input and output boundaries)**.** Let $U$ be a closed subset of an oriented graded poset. For all $\alpha \in \{+,-\}$ and $n \in \mathbb{N}$, let

$$\Delta^\alpha_n U \coloneqq \left\{x \in U_n \mid \nabla^{-\alpha} x \cap U = \varnothing\right\}.$$

Note that, if $\mathscr{M}\!ax\, U$ is the set of maximal elements of $U$, then

$$\Delta^-_n U \cap \Delta^+_n U = (\mathscr{M}\!ax\, U)_n.$$

The *input* ($\alpha = -$) and *output* ($\alpha = +$) *$n$-boundary* of $U$ is the closed subset

$$\partial^\alpha_n U \coloneqq \operatorname{cl}(\Delta^\alpha_n U) \cup \bigcup_{k<n} \operatorname{cl}(\mathscr{M}\!ax\, U)_k.$$

We omit the subscript when $n = \dim U - 1$, and for $n \in \{-1, -2\}$, we let $\Delta^\alpha_n U = \partial^\alpha_n U \coloneqq \varnothing$.

***15*.** We use the following notations, for $x$ an element in an oriented graded poset, $U$ a closed subset, $n \in \mathbb{N}$, and $\alpha \in \{+,-\}$: $\partial^\alpha_n x \coloneqq \partial^\alpha_n \operatorname{cl}\{x\}$, $\partial_n U \coloneqq \partial^-_n U \cup \partial^+_n U$, $\Delta_n U \coloneqq \Delta^-_n U \cup \Delta^+_n U$.

***Lemma 16.*** *Let $U$ be a closed subset of an oriented graded poset, $n \in \mathbb{N}$, and $\alpha \in \{+,-\}$. Then*

*1) $\partial^\alpha_n U \subseteq U$,*
*2) $\partial^\alpha_n U = U$ if and only if $n \ge \dim U$.*

*Example 17.* Let $U$ be the oriented graded poset of Example 11. Then $\partial^-_2 U = \partial^+_2 U = U$, and

$$\partial^-_1 U = \{(0,0),(0,1),(0,2),(0,3),(1,0),(1,1),(1,2)\},$$
$$\partial^+_1 U = \{(0,0),(0,2),(0,3),(1,2),(1,3)\},$$
$$\partial^-_0 U = \{(0,0)\}, \qquad \partial^+_0 U = \{(0,3)\}.$$

***18*** (Map of oriented graded posets)**.** Let $P, Q$ be oriented graded posets. A *map* $f\colon P \to Q$ is a function of their underlying sets that satisfies

$$f(\partial^\alpha_n x) = \partial^\alpha_n f(x)$$

for all $x \in P$, $n \in \mathbb{N}$, and $\alpha \in \{+,-\}$. There is a category $\mathbf{ogPos}$ whose objects are oriented graded posets and morphisms are maps.

***19*** (Inclusion of oriented graded posets)**.** An *inclusion* is an injective map of oriented graded posets.

***Lemma 20.*** *Let $\imath\colon P \hookrightarrow Q$ be an inclusion of oriented graded posets. Then*

*1) $\imath$ is both order-preserving and order-reflecting, that is, $\imath(x) \le \imath(y)$ if and only if $x \le y$;*

2) $\imath$ *preserves dimensions, that is,* $\dim \imath(x) = \dim x$ *for all* $x \in P$*;*
3) $\imath$ *preserves the covering relation and orientations, that is, if* $y$ *covers* $x$ *in* $P$ *with orientation* $\alpha$*, then* $\imath(y)$ *covers* $\imath(x)$ *in* $Q$ *with orientation* $\alpha$*;*
4) *for all closed* $U \subseteq P$*,* $n \in \mathbb{N}$ *and* $\alpha \in \{+,-\}$*,* $\imath$ *maps* $\partial^\alpha_n U$ *isomorphically onto* $\partial^\alpha_n \imath(U)$*.*

*Remark 21.* Every closed subset of an oriented graded poset inherits an orientation by restriction, and its inclusion as a subset is an inclusion of oriented graded posets. The fact that inclusions preserve all boundaries justifies us in identifying an oriented graded poset with its image through an inclusion, which we will often do implicitly.

***Proposition 22.*** *The category* **ogPos** *has*

1) *a terminal object* 1*,*
2) *an initial object* $\varnothing$*,*
3) *pushouts of inclusions.*

***23.*** Not all oriented graded posets describe well-formed diagram shapes. In our framework, the well-formed shapes form an inductive subclass, the *regular molecules*, generated by the following constructions.

***24*** (Pasting construction)**.** Let $U, V$ be oriented graded posets, $k \in \mathbb{N}$, and let $\varphi \colon \partial^+_k U \xhookrightarrow{\sim} \partial^-_k V$ be an isomorphism. The *pasting of* $U$ *and* $V$ *at the* $k$*-boundary along* $\varphi$ is the oriented graded poset $U \#^\varphi_k V$ obtained in **ogPos** as the pushout

$$\begin{array}{ccccc} \partial^+_k U & \xhookrightarrow{\varphi} & \partial^-_k V & \hookrightarrow & V \\ \downarrow & & & & \downarrow \\ U & & \hookrightarrow & & U \#^\varphi_k V. \end{array}$$

***25*** (Rewrite construction)**.** Let $U, V$ be oriented graded posets of the same dimension $n$, and suppose $\varphi \colon \partial U \xhookrightarrow{\sim} \partial V$ is an isomorphism restricting to isomorphisms $\varphi^\alpha \colon \partial^\alpha U \xhookrightarrow{\sim} \partial^\alpha V$ for each $\alpha \in \{+,-\}$. Construct the pushout

$$\begin{array}{ccccc} \partial U & \xhookrightarrow{\varphi} & \partial V & \hookrightarrow & V \\ \downarrow & & & & \downarrow \\ U & & \hookrightarrow & & \partial(U \Rightarrow^\varphi V) \end{array}$$

in **ogPos**. The *rewrite of* $U$ *into* $V$ *along* $\varphi$ is the oriented graded poset $U \Rightarrow^\varphi V$ obtained by adjoining a single $(n+1)$-dimensional element $\top$ to $\partial(U \Rightarrow^\varphi V)$ such that $\Delta^- \top \coloneqq U_n$ and $\Delta^+ \top \coloneqq V_n$.

***Lemma 26.*** *Let* $U, V$ *be oriented graded posets and suppose* $U \Rightarrow^\varphi V$ *is defined. Then*

1) $\partial^-(U \Rightarrow^\varphi V)$ *is isomorphic to* $U$*,*
2) $\partial^+(U \Rightarrow^\varphi V)$ *is isomorphic to* $V$*.*

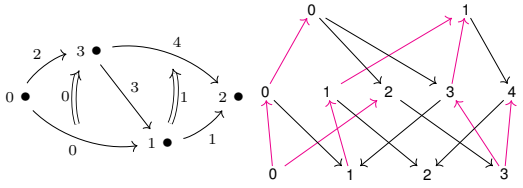
*Remark 27.* The notation $\partial(U \Rightarrow^\varphi V)$ for the pushout in the rewrite construction is *a posteriori* justified, that is, the pushout indeed constructs the boundary of $U \Rightarrow^\varphi V$.

***28*** (Roundness)**.** Let $U$ be an oriented graded poset. We say that $U$ is *round* if, for all $n < \dim U$, $\partial^-_n U \cap \partial^+_n U = \partial_{n-1} U$.

*Example 29.* The shape of a pasting diagram is round when, intuitively, the diagram is shaped as a topological ball of the appropriate dimension. For example, the oriented graded poset of Example 11 is not round, since

$$\begin{aligned} \partial_0 U &= \{(0,0),(0,3)\} \\ \subsetneq \partial^+_1 U \cap \partial^-_1 U &= \{(0,0),(0,2),(1,2),(0,3)\}, \end{aligned}$$

and, indeed, the pasting diagram is shaped as the wedge of a 2-ball (disc) with a 1-ball (interval). However, the following pasting diagram is round:




***30*** (Pure subset)**.** Let $U$ be a closed subset of a graded poset, $n \coloneqq \dim U$. We say that $U$ is *pure* if all the maximal elements of $U$ have dimension $n$, that is, $\mathscr{M}\!ax\, U = U_n$.

***Lemma 31.*** *If* $U$ *is round, then it is pure.*

***32*** (Regular molecule)**.** The class of *regular molecules* is the inductive subclass of oriented graded posets closed under isomorphisms and generated by the following clauses.

1) (*Point*). The terminal oriented graded poset 1 is a regular molecule.
2) (*Paste*). If $U, V$ are regular molecules, $\varphi \colon \partial^+_k U \xhookrightarrow{\sim} \partial^-_k V$ is an isomorphism with $k < \min\{\dim U, \dim V\}$, then $U \#^\varphi_k V$ is a regular molecule.
3) (*Atom*). If $U$, $V$ are *round* regular molecules of the same dimension and $\varphi \colon \partial U \xhookrightarrow{\sim} \partial V$ is an isomorphism restricting to $\varphi^\alpha \colon \partial^\alpha U \xhookrightarrow{\sim} \partial^\alpha V$ for each $\alpha \in \{+,-\}$, then $U \Rightarrow^\varphi V$ is a regular molecule.

***33.*** We summarise the essential properties of regular molecules. The second point of Proposition 34 allows us to write $U \#_k V$ and $U \Rightarrow V$ instead of $U \#^\varphi_k V$ and $U \Rightarrow^\varphi V$ when the latter are defined and $U, V$ are regular molecules.

***Proposition 34.*** *Let* $U, V$ *be regular molecules,* $k \in \mathbb{N}$*. Then*

1) *if* $U$ *and* $V$ *are isomorphic, they are isomorphic in a unique way;*
2) *if* $U \#^\varphi_k V$ *or* $U \Rightarrow^\varphi V$ *is defined, it is defined for a unique* $\varphi$*;*
3) *for all* $n \in \mathbb{N}$*,* $\alpha \in \{+,-\}$*,* $\partial^\alpha_n U$ *is a regular molecule;*
4) $U$ *is globular, that is, for all* $k, n \in \mathbb{N}$ *and* $\alpha, \beta \in \{+,-\}$*, if* $k < n$ *then* $\partial^\alpha_k(\partial^\beta_n U) = \partial^\alpha_k U$*;*
5) *if* $U$ *is round, for all* $n \in \mathbb{N}$*,* $\alpha \in \{+,-\}$*,* $\partial^\alpha_n U$ *is round.*

*Example 35.* Let

$$\begin{aligned}\mathsf{arrow} &\coloneqq 1 \Rightarrow 1,\\ \mathsf{binary} &\coloneqq (\mathsf{arrow} \#_0 \mathsf{arrow}) \Rightarrow \mathsf{arrow},\\ \mathsf{cobinary} &\coloneqq \mathsf{arrow} \Rightarrow (\mathsf{arrow} \#_0 \mathsf{arrow}).\end{aligned}$$

Then the shape of Example 11 is a regular molecule constructed as $\mathsf{binary} \#_0 \mathsf{arrow}$, while the shape of Example 29 is a regular molecule constructed as

$$(\mathsf{cobinary} \#_0 \mathsf{arrow}) \#_1 (\mathsf{arrow} \#_0 \mathsf{binary}).$$

***36.*** The following results imply, together, that pasting of regular molecules satisfies the equations of strict $\omega$-categories up to unique isomorphism. In particular, the associativity result allows us to write multiple pastings in the same dimension without bracketing.

***Proposition 37.*** *Let $U, V, W$ be regular molecules and let $k \in \mathbb{N}$ such that $U \#_k V$ and $V \#_k W$ are both defined. Then $(U \#_k V) \#_k W$ and $U \#_k (V \#_k W)$ are both defined and uniquely isomorphic.*

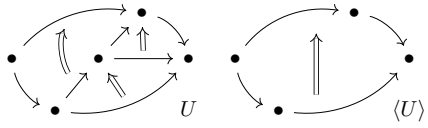
***Proposition 38.*** *Let $U$ be a regular molecule and $k \in \mathbb{N}$. Then $U \#_k \partial_k^+ U$ and $\partial_k^- U \#_k U$ are both defined and uniquely isomorphic to $U$.*

***Proposition 39.*** *Let $U, U', V, V'$ be regular molecules and $k < n \in \mathbb{N}$ such that $(U \#_n U') \#_k (V \#_n V')$ is defined. Then $(U \#_k V) \#_n (U' \#_k V')$ is defined and uniquely isomorphic to $(U \#_n U') \#_k (V \#_n V')$.*

***40*** (Atom)**.** An *atom* is a regular molecule with a greatest element.

***41*** (Merger of a round regular molecule)**.** Let $U$ be a round regular molecule of dimension $> 0$. The *merger* of $U$ is the atom $\langle U \rangle \coloneqq \partial^- U \Rightarrow \partial^+ U$.

*Example 42.* The following pair of pasting diagrams depicts a round regular molecule $U$ and its merger $\langle U \rangle$, an atom.




***Proposition 43.*** *Let $U$ be a regular molecule. Then*

1) *$U$ is an atom if and only if it was produced by (Point) or (Atom);*
2) *for all $x \in U$, $\mathrm{cl}\{x\}$ is an atom;*
3) *if $U$ is an atom, then $U$ is round, and if $\dim U > 0$ then $U$ is isomorphic to $\langle U \rangle$.*

***44.*** For us, a *cell* is a diagram whose shape is an atom.

***45*** (Diagram isomorphism problem)**.** The *diagram isomorphism problem* is the following decision problem: given diagrams $t \colon U \to \mathbb{V}$ and $t' \colon U' \to \mathbb{V}$, does there exist an isomorphism $\varphi \colon U \xrightarrow{\sim} U'$ of their shapes such that $t = \varphi; t'$?

By Proposition 34, if $\varphi$ exists it is unique, and if $\varphi$ is found the labellings $t$ and $\varphi; t'$ can be compared in linear time, so this problem reduces to the isomorphism problem for regular molecules. One of the main results of [9] is a polynomial-time solution to this problem, relying on a deterministic *traversal algorithm* for regular molecules:

1) all elements of a regular molecule can be traversed in polynomial time in such a way that the traversal order is invariant under isomorphism;
2) consequently, $U$ and $U'$ are isomorphic if and only if their representations using the traversal order dimension-wise are identical.

The traversal order also gives a *canonical representation* of regular molecules. The constructors for regular molecules can be implemented in such a way that the elements are rearranged in traversal order after each step, so that regular molecules can be checked for equality rather than isomorphism.

The complexity upper bound given in [9, Theorem 2.19] is in fact the result of overcounting. We describe the traversal algorithm, whose correctness is proved in [9, Theorem 2.17], and give an improved upper bound.

***46*** (Traversal algorithm)**.** The procedure takes as input a regular molecule $U$ and returns a list of its elements in the order in which they are *marked*. It uses an auxiliary stack of regular molecules $V \subseteq U$.

At the beginning, only $U$ is on the stack and all elements are unmarked. We iterate the *main loop* until the stack is empty, at which point the procedure terminates.

At each iteration, suppose $V$ is on top of the stack. If all elements of $V$ are marked, then we pop $V$ from the stack and iterate. Else, if any elements of $\partial^- V$ are unmarked, we push $\partial^- V$ to the top of the stack and iterate. Else, if $V = \mathrm{cl}\{x\}$ for some $x \in U$, we

1) mark $x$ and pop $V$ from the stack,
2) if any elements of $\partial^+ V$ are unmarked, we push $\partial^+ V$ to the top of the stack, and
3) we iterate.

Else, we let $y$ be the earliest marked element such that $\dim y = \dim V - 1$ and there is an unmarked $x \in \nabla^- y \cap V$. Such a $y$ always exists, and then $\nabla^- y \cap V = \{x\}$. We push $\mathrm{cl}\{x\}$ to the top of the stack and iterate.

***47.*** For a regular molecule $U$, and all $k \in \mathbb{N}$, we let

$$\begin{aligned}\mathscr{E}_k U &\coloneqq \coprod_{x \in U_k} \Delta x = \coprod_{y \in U_{k-1}} \nabla y,\\ |U_\vee| &\coloneqq \max\{|U_i|\}_{i \in \mathbb{N}},\\ |\mathscr{E}_\vee U| &\coloneqq \max(\{|\mathscr{E}_i U|\}_{i \in \mathbb{N}} \cup \{1\}).\end{aligned}$$

Note that $\mathscr{E}_k U$ is the set of edges between $k$ and $(k-1)$-dimensional elements in $\mathscr{H}U$. We have $|U_k| \leq |\mathscr{E}_k U|$ for all $k > 0$, while $|\mathscr{E}_0 U| = 0$. Since the maximum of the $|\mathscr{E}_k U|$ is 0 only when $U$ is 0-dimensional, in which case $|U_k| = 1$, with our definition we always have $|U_\vee| \leq |\mathscr{E}_\vee U|$.

***Theorem 48.*** *The traversal algorithm admits an implementation running in time* $O(|U|\ |\mathscr{E}_\vee U|\ \log|\mathscr{E}_\vee U|)$.

*Proof.* The upper bound of $O(|\mathscr{E}_\vee U|\ \log|\mathscr{E}_\vee U|)$ on each loop iteration is derived as in [9, Theorem 2.19], simplified by our modified definition of $|\mathscr{E}_\vee U|$, so we only need to show that there are $O(|U|)$ loop iterations.

We let $k \le \dim U$ and we count the number of loop iterations where a $k$-dimensional subset $V$ is on top of the stack. This can happen in two ways:

- $V$ is either $U$ or $\partial^\alpha W$ for some $W$ with $\dim W > k$, where $W$ was earlier (and may still be) on the stack,
- $V$ is $\operatorname{cl}\{x\}$ for some $x \in W$, where $\dim W = k$ and $W$ is below $V$ on the stack.

Let $(V^{(i)})_{i=1}^m$ be the sequence of all $k$-dimensional subsets appearing on the stack *in the first way* during the run, in the order in which they appear. For all $j < i \in \{1,\dots,m\}$, by [9, Lemma 2.14] $V^{(j)}$ must be fully marked before $V^{(i)}$ can appear on the stack. Moreover, $V^{(i)}$ can be on top at most

1) once to push $\partial^- V^{(i)}$ to the top,
2) once every time we push $\operatorname{cl}\{x\}$ to the top for an unmarked $x \in (V^{(i)})_k$,
3) once to pop $V^{(i)}$ from the stack.

Any $k$-dimensional $\operatorname{cl}\{x\}$ appearing *in the second way* appears while a unique $V^{(i)}$ is on the stack, and at most

1) once to push $\partial^- x$ to the top,
2) once to mark $x$ and pop $\operatorname{cl}\{x\}$ from the stack.

Let $U_k^{(i)} \coloneqq (V^{(i)})_k \setminus \bigcup_{j<i}(V^{(j)})_k$. Then $U_k^{(i)}$ is precisely the set of unmarked $k$-dimensional elements of $V^{(i)}$ when $V^{(i)}$ first appears on the stack. It follows that the number of loop iterations with a $k$-dimensional subset on top of the stack *while* $V^{(i)}$ *is on the stack* is at most $2 + 3\left|U_k^{(i)}\right|$.

Since at the end of the procedure all $k$-dimensional elements of $U$ are marked, the $(U_k^{(i)})_{i=1}^m$ form a partition of $U_k$. Thus, the total number of loop iterations where a $k$-dimensional subset is on top of the stack is bounded above by

$$\sum_{i=1}^{m}\left(2+3\left|U_k^{(i)}\right|\right) = 2m + 3\sum_{i=1}^{m}\left|U_k^{(i)}\right| = 2m + 3\,|U_k|\,,$$

which is bounded above by $5\,|U_k|$. Summing over all dimensions, we get an upper bound of $5\,|U|$ iterations. ∎

## II. The subdiagram matching problem

***49*** (Submolecule inclusion)**.** The class of *submolecule inclusions* is the smallest subclass of inclusions of regular molecules such that

1) all isomorphisms are submolecule inclusions,
2) for all regular molecules $U, V$ and all $k \in \mathbb{N}$ such that $U \#_k V$ is defined, $U \hookrightarrow (U \#_k V)$ and $V \hookrightarrow (U \#_k V)$ are submolecule inclusions,
3) the composite of two submolecule inclusions is a submolecule inclusion.

A closed subset $V \subseteq U$ is a *submolecule* if its inclusion in $U$ is a submolecule inclusion. In that case we write $V \sqsubseteq U$.

***50.*** We also let $\varnothing \sqsubseteq \varnothing$ to take care of some corner cases.

***Lemma 51.*** *Let $U$ be a regular molecule. Then*

*1) for all $n \in \mathbb{N}$ and $\alpha \in \{+,-\}$, $\partial^\alpha_n U \sqsubseteq U$;*
*2) for all $x \in U$, $\operatorname{cl}\{x\} \sqsubseteq U$.*

***52*** (Substitution)**.** Let $U, V, W$ be regular molecules with $\dim U = \dim V = \dim W$, $\imath\colon V \hookrightarrow U$ an inclusion, and suppose that $V \Rightarrow W$ is defined. Consider the pushout

$$\begin{array}{ccc} V & \hookrightarrow & V \Rightarrow W \\ \downarrow{\scriptstyle \imath} & \ulcorner & \downarrow \\ U & \hookrightarrow & U \cup (V \Rightarrow W) \end{array} \tag{1}$$

in $\mathbf{ogPos}$. The *substitution of $W$ for $\imath$* $U[W/\imath(V)] \coloneqq \partial^+(U \cup (V \Rightarrow W))$. When $\imath$ is the inclusion of a closed subset we write simply $U[W/V]$.

***Proposition 53.*** *Let $\imath\colon V \hookrightarrow U$ be an inclusion of regular molecules such that $\dim V = \dim U$ and $V$ is round. The following are equivalent:*

*(a) $\imath$ is a submolecule inclusion;*
*(b) for all regular molecules $W$ such that $V \Rightarrow W$ is defined, $U \cup (V \Rightarrow W)$ in (1) is a regular molecule;*
*(c) for all regular molecules $W$ such that $V \Rightarrow W$ is defined, $U[W/\imath(V)]$ is a regular molecule;*
*(d) $U[\langle V\rangle/\imath(V)]$ is a regular molecule.*

***54*** (Rewritable submolecule)**.** A submolecule $V \sqsubseteq U$ is *rewritable* if $\dim V = \dim U$ and $V$ is round.

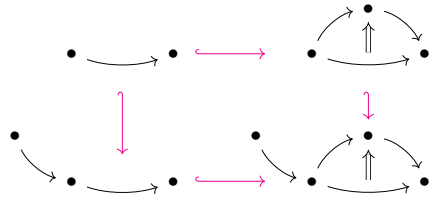
***55*** (Rewritable subdiagram)**.** Let $t\colon U \to \mathbb{V}$ be a diagram. A *rewritable subdiagram* of $t$ is the restriction of $t$ to a rewritable submolecule $V \sqsubseteq U$.

***56.*** We extend boundary operations to diagrams $t\colon U \to \mathbb{V}$ by $\partial^\alpha_n t \coloneqq t|_{\partial^\alpha_n U}$ for all $n \in \mathbb{N}$ and $\alpha \in \{+,-\}$.

***57.*** Suppose $t$ is an $n$-dimensional diagram of shape $U$ and $r$ a rewrite rule, in the form of an $(n+1)$-dimensional cell of shape $V \Rightarrow W$. If there is an inclusion $\imath\colon V \hookrightarrow U$ such that $\partial^- r = \imath; t$, then the application of the rewrite $r$ to $t$ is modelled by an $(n+1)$-dimensional diagram $t \cup r$ of shape $U \cup (V \Rightarrow W)$ as in (1).

By Proposition 53, $\partial^+(t \cup r)$, which models the substitution of $\partial^+ r$ for $\partial^- r$ in $t$, is guaranteed to be an $n$-dimensional diagram precisely when $\imath$ is a submolecule inclusion, that is, $\partial^- r$ is isomorphic to a rewritable subdiagram of $t$.

*Example 58.* The following is a depiction of diagram (1) when $V \Rightarrow W \coloneqq \mathsf{cobinary}$, $U \coloneqq \mathsf{arrow} \#_0 \mathsf{arrow}$, and $\imath$ is the inclusion of the second arrow into the pasting.



The result of the substitution is the output boundary of the bottom right diagram, isomorphic to arrow $\#_0$ arrow $\#_0$ arrow.

***59*** (Subdiagram matching problem)**.** The *subdiagram matching problem* is the following search problem: given diagrams $t\colon U \to \mathbb{V}$ and $s\colon V \to \mathbb{V}$ such that $\dim U = \dim V$ and $V$ is round, find, if any, the submolecule inclusions $\imath\colon V \hookrightarrow U$ such that $s = \imath; t$. This can be split into three subproblems.

1) (*Molecule matching problem*). Find, if any, the inclusions $\imath\colon V \hookrightarrow U$.
2) (*Rewritable submolecule problem*). Decide if $\imath(V) \sqsubseteq U$.
3) Decide if $s = \imath; t$.

In this section, we will focus on the molecule matching problem, and in the next on the rewritable submolecule problem. The third problem is trivial.

***Lemma 60.*** *Let $U$ be a regular molecule, $n \coloneqq \dim U$, $x \in U_{n-1}$, and $\alpha \in \{+,-\}$. Then*

*1) $x \in \mathscr{M}\!ax\, U$ if and only if $|\nabla x| = 0$,*
*2) $x \in \Delta^{\alpha} U \setminus \Delta^{-\alpha} U$ if and only if $|\nabla^{\alpha} x| = 1$ and $|\nabla^{-\alpha} x| = 0$,*
*3) $x \notin \Delta U$ if and only if $|\nabla^{+} x| = |\nabla^{-} x| = 1$.*

***61*** (Flow graph)**.** Let $U$ be a regular molecule, $k \geq -1$. The *$k$-flow graph of $U$* is the directed graph $\mathscr{F}_k U$ whose

- set of vertices is $\bigcup_{i>k} U_i$, and
- for all vertices $x, y$, there is an edge from $x$ to $y$ if and only if $\Delta_k^{+} x \cap \Delta_k^{-} y$ is non-empty.

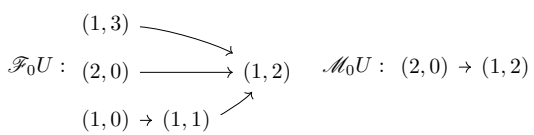
***62*** (Induced subgraph)**.** Let $\mathscr{G}$ be a directed graph and let $W$ be a subset of its vertex set. The *induced subgraph* of $\mathscr{G}$ on $W$ is the directed graph $\mathscr{G}|_W$ whose vertex set is $W$, and there is an edge from $x$ to $y$ for every edge from $x$ to $y$ in $\mathscr{G}$.

***63*** (Maximal flow graph)**.** Let $U$ be a regular molecule, $k \geq -1$. The *maximal $k$-flow graph of $U$* is the induced subgraph $\mathscr{M}_k U$ of $\mathscr{F}_k U$ on the vertex set

$$\bigcup_{i>k} (\mathscr{M}\!ax\, U)_i \subseteq \bigcup_{i>k} U_i.$$

Note that, if $k = \dim U - 1$, then $\mathscr{F}_k U = \mathscr{M}_k U$.

*Example 64.* If $U$ is the regular molecule of Example 11,




$$\mathscr{F}_1 U = \mathscr{M}_1 U : \ (2,0)$$

If $U$ is the round regular molecule of Example 29,

$$\mathscr{F}_1 U = \mathscr{M}_1 U : \ (2,0) \to (2,1)$$

***Lemma 65.*** *Let $\imath\colon V \hookrightarrow U$ be an inclusion of regular molecules, $k \geq -1$. Then $\mathscr{F}_k V$ is isomorphic to the induced subgraph of $\mathscr{F}_k U$ on the vertices in the image of $\imath$.*

***Proposition 66.*** *Let $U$ be a regular molecule, $n \coloneqq \dim U$. If $U$ is round, then $\mathscr{F}_{n-1} U$ is connected.*

***67.*** We describe an algorithm for the molecule matching problem. The main idea is the following: if we succeed in matching only *one* top-dimensional atom in $V$ with one top-dimensional atom in $U$, then there is only *one possible matching* of all other top-dimensional atoms of $V$ to atoms in $U$. This is because, by Proposition 66, we can try to match top-dimensional elements of $V$ in such an order that the next element to match is connected by an edge in $\mathscr{F}_{n-1} V$ to a previously matched element; that is, it shares a face $z$ with a previously matched element. In particular, $z$ has already been matched. By Lemma 60, in order for us to continue, the match of $z$ must have exactly two cofaces in $U$, one of which is the previously matched top-dimensional element. Necessarily, the other coface is the next match.

***68*** (Molecule matching algorithm)**.** The procedure takes as input two regular molecules $U, V$ such that $\dim U = \dim V$ and $V$ is round, and it returns all inclusions $V \hookrightarrow U$.

Let $n \coloneqq \dim U$. To begin, we pick an arbitrary ordering $(x^{(i)})_{i=1}^{m}$, for example the traversal order, of the elements of $U_n$. Moreover, we pick an ordering $(y^{(j)})_{j=1}^{p}$ of the elements of $V_n$ with the property that, for all $k \in \{1, \dots, p\}$, the induced subgraph of $\mathscr{F}_{n-1} V$ on $(y^{(j)})_{j=1}^{k}$ is connected. This is possible because $\mathscr{F}_{n-1} V$ is connected by Proposition 66. For each $k \in \{1, \dots, p\}$, we let $V^{(k)} \coloneqq \bigcup_{j \leq k} \operatorname{cl}\{y^{(j)}\}$. We have $V^{(i)} \subseteq V^{(j)}$ whenever $i \leq j$, and $V^{(p)} = V$ since $V$ is pure by Lemma 31.

For each $i \in \{1, \dots, m\}$, we attempt to construct a sequence of inclusions $(\imath^{(i,j)}\colon V^{(j)} \hookrightarrow U)_{j=1}^{p}$ such that the restriction of $\imath^{(i,j')}$ to $V^{(j)}$ is equal to $\imath^{(i,j)}$ when $j \leq j'$, iterating on $k \in \{1, \dots, p\}$. When $k = 1$, if $V^{(1)} = \operatorname{cl}\{y^{(1)}\}$ is isomorphic to $\operatorname{cl}\{x^{(i)}\}$, we let $\imath^{(i,1)}$ be the unique isomorphism $V^{(1)} \xrightarrow{\sim} \operatorname{cl}\{x^{(i)}\}$ followed by the inclusion $\operatorname{cl}\{x^{(i)}\} \subseteq U$, and iterate on $k$. Else, we iterate on $i$.

When $k > 1$, we let $j$ be the least value such that there exists an edge between $y^{(j)}$ and $y^{(k)}$ in $\mathscr{F}_{n-1} V$. Then $j < k$ because of our connectedness assumption, and there exists $z \in \Delta^{\alpha} y^{(j)} \cap \Delta^{-\alpha} y^{(k)}$ for some $\alpha \in \{+,-\}$. We pick the least such $z$ with respect to some ordering of $V_{n-1}$, for example the traversal order. By Lemma 20, $\imath^{(i,k-1)}(y^{(j)})$ is one coface of $\imath^{(i,k-1)}(z)$ in $U$. If $\imath^{(i,k-1)}(z)$ has no other cofaces, then we iterate on $i$. Else, by Lemma 60, $\imath^{(i,k-1)}(z)$ has exactly one other coface, call it $x$; note that $x$ cannot be in the image of $\imath^{(i,k-1)}$, since $y^{(j)}$ and $y^{(k)}$ are the only cofaces of $z$ in $V$. If $\operatorname{cl}\{y^{(k)}\}$ is isomorphic to $\operatorname{cl}\{x\}$, and the unique isomorphism $\operatorname{cl}\{y^{(k)}\} \xrightarrow{\sim} \operatorname{cl}\{x\}$ followed by the inclusion $\operatorname{cl}\{x\} \subseteq U$ matches $\imath^{(i,k-1)}$ on $\operatorname{cl}\{y^{(k)}\} \cap V^{(k-1)}$, then we let $\imath^{(i,k)}$ be the unique extension of $\imath^{(i,k-1)}$ that restricts to $\operatorname{cl}\{y^{(k)}\} \xrightarrow{\sim} \operatorname{cl}\{x\} \subseteq U$. Else, we iterate on $i$.

If we succeed to construct $\imath^{(i,p)}$, we add it to the list of inclusions $V \hookrightarrow U$, then iterate on $i$.

***Theorem 69.*** *The molecule matching problem in dimension $n$*

*can be solved in time*

$$O(|U_n|\ |V_n|\ |V|\ |\mathscr{E}_\vee V|\ \log|\mathscr{E}_\vee V|).$$

*Proof.* We suppose $n > 0$ since the case $n = 0$ is trivial. First of all, with our choice of data structures both $U_n$ and $V_{n-1}$ already come with a linear order when one is needed. Moreover, we can both construct $\mathscr{F}_{n-1}V$ and order its vertices in the desired way by traversing the "slice" of $\mathscr{H}V$ on the elements of dimension $n$ and $(n-1)$. Since $\max\{|V_n|, |V_{n-1}|\} \leq |\mathscr{E}_n V|$, this can be done in time $O(|\mathscr{E}_n V|)$ with a standard traversal algorithm.

In the main part of the algorithm, we have exactly $|U_n|$ iterations. At each iteration, we need to solve at most $|V_n|$ isomorphism problems for submolecules of $V$. The time complexity of each can be bounded above by the time complexity of the isomorphism problem for $V$, which is $O(|V|\ |\mathscr{E}_\vee V|\ \log|\mathscr{E}_\vee V|)$ by Theorem 48. It is straightforward to verify that all other operations, such as checking that the isomorphisms match on intersections or finding the next match, have lower complexity. Since $|\mathscr{E}_n V| \leq |\mathscr{E}_\vee V|$, we can ignore the $O(|\mathscr{E}_n V|)$ summand, and conclude. ∎

## III. The rewritable submolecule problem

***70.*** Our solution to the rewritable submolecule problem requires us to develop new results about *layerings* of diagrams, and their associated *orderings*.

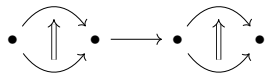
***71*** (Layering of a regular molecule)**.** Let $U$ be a regular molecule, $-1 \leq k < \dim U$, and

$$m \coloneqq \left| \bigcup_{i>k} (\mathscr{M}ax\, U)_i \right|.$$

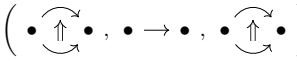
A *$k$-layering of $U$* is a sequence $(U^{(i)})_{i=1}^m$ of regular molecules such that $U$ is isomorphic to $U^{(1)} \#_k \dots \#_k U^{(m)}$ and $\dim U^{(i)} > k$ for all $i \in \{1, \dots, m\}$.

For $k = -1$, it is implied that $m = 1$, and $U$ is an atom. We will regularly identify the regular molecules in a layering of $U$ with their isomorphic images in $U$, which are submolecules.

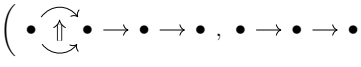
*Example 72.* The shape of the pasting diagram



admits no $(-1)$-layerings, a single 0-layering



and two 1-layerings:



***Lemma 73.*** *Let $U$ be a regular molecule, $k < \dim U$, and suppose $U$ admits a $k$-layering $(U^{(i)})_{i=1}^m$. Then*

1) *for all $i \in \{1, \dots, m\}$, $U^{(i)}$ contains a single maximal element of dimension $> k$,*
2) *for all $k \leq \ell < \dim U$, $U$ admits an $\ell$-layering.*

***74.*** A regular molecule does not, in general, admit a $k$-layering for each $k < \dim U$; however, by Lemma 73, when it does admit a $k$-layering, it also admits a layering in dimensions higher than $k$. The next result shows that every regular molecule does admit a $k$-layering for some $k$, and that the smallest such $k$ falls into a particular range.

***75*** (Layering dimension)**.** Let $U$ be a regular molecule. The *layering dimension* of $U$ is the integer

$$\operatorname{lydim} U \coloneqq \min\left\{ k \geq -1 \mid \left| \bigcup_{i>k+1} (\mathscr{M}ax\, U)_i \right| \leq 1 \right\}.$$

***76*** (Frame dimension)**.** Let $U$ be a regular molecule. The *frame dimension of $U$* is the integer

$$\operatorname{frdim} U \coloneqq \dim \bigcup \{\operatorname{cl}\{x\} \cap \operatorname{cl}\{y\} \mid x, y \in \mathscr{M}ax\, U, x \neq y\}.$$

***Theorem 77.*** *Let $U$ be a regular molecule. Then there exists $k < \dim U$ such that $U$ admits a $k$-layering. Moreover,*

$$\operatorname{frdim} U \leq \min\{k \mid U \text{ admits a } k\text{-layering}\} \leq \operatorname{lydim} U.$$

***Corollary 78.*** *Let $U$ be a regular molecule, $n \coloneqq \dim U$. Then $U$ admits an $(n-1)$-layering.*

***79*** (Ordering of a regular molecule)**.** Let $U$ be a regular molecule, $k \geq -1$, and suppose $\mathscr{M}_k U$ is acyclic. A *$k$-ordering of $U$* is a topological sort of $\mathscr{M}_k U$.

***Proposition 80.*** *Let $U$ be a regular molecule, $k \geq -1$. If $U$ admits a $k$-layering, then $\mathscr{M}_k U$ is acyclic, hence $U$ admits a $k$-ordering.*

***Corollary 81.*** *Let $U$ be a regular molecule, $n \coloneqq \dim U$. Then $\mathscr{F}_{n-1}U$ is acyclic.*

*Proof.* Follows from Corollary 78 and Proposition 80 combined with the fact that $\mathscr{F}_{n-1}U = \mathscr{M}_{n-1}U$. ∎

***82.*** Let $U$ be a regular molecule, $k \geq -1$. We let

$$\mathscr{L}ay_k U \coloneqq \left\{ k\text{-layerings } (U^{(i)})_{i=1}^m \text{ of } U \right\},$$
$$\mathscr{O}rd_k U \coloneqq \left\{ k\text{-orderings } (x^{(i)})_{i=1}^m \text{ of } U \right\},$$

where layerings are considered up to layer-wise isomorphism.

***Proposition 83.*** *Let $U$ be a regular molecule, $k \geq -1$. For each $k$-layering $(U^{(i)})_{i=1}^m$ of $U$ and each $i \in \{1, \dots, m\}$, let $x^{(i)}$ be the only element of $\bigcup_{j>k}(\mathscr{M}ax\, U)_j$ in the image of $U^{(i)}$. Then the assignment*

$$\mathsf{m}_{k,U} \colon (U^{(i)})_{i=1}^m \mapsto (x^{(i)})_{i=1}^m \tag{2}$$

*determines an injective function $\mathscr{L}ay_k U \hookrightarrow \mathscr{O}rd_k U$.*

***84.*** In general, the function $\mathsf{m}_{k,U}$ is not surjective, that is, not every $k$-ordering is induced by a $k$-layering. The following

is a criterion for deciding when a $k$-ordering comes from a $k$-layering.

***Proposition 85.*** *Let $U$ be a regular molecule, $k \geq -1$, and let $(x^{(i)})_{i=1}^m$ be a $k$-ordering of $U$. Let*

$$U^{(0)} \coloneqq \partial_k^- U,$$
$$U^{(i)} \coloneqq \partial_k^+ U^{(i-1)} \cup \operatorname{cl}\left\{x^{(i)}\right\} \quad \text{for } i \in \{1, \dots, m\}.$$

*The following are equivalent:*

*(a) $(U^{(i)})_{i=1}^m$ is a $k$-layering of $U$;*
*(b) for all $i \in \{1, \dots, m\}$, $\partial_k^- x^{(i)} \sqsubseteq \partial_k^- U^{(i)}$.*

*Moreover, for all $i \in \{1, \dots, m-1\}$, if $\partial_k^- x^{(i)} \sqsubseteq \partial_k^- U^{(i)}$, then $U^{(i)}$ and $\partial_k^+ U^{(i)} = \partial_k^- U^{(i+1)}$ are regular molecules.*

***86*** (Path-induced subgraph)**.** Let $\mathscr{G}$ be a directed graph and $W$ a subset of its vertex set. We say that $\mathscr{G}|_W$ is *path-induced* if, for all $x, y \in W$, every path from $x$ to $y$ in $\mathscr{G}$ is included in $\mathscr{G}|_W$.

***87.*** Path-induced subgraphs are also called *convex subgraphs*, for example in [12].

***88*** (Contraction of a connected subgraph)**.** Let $\mathscr{G}$ be a directed graph and $W$ a subset of its vertex set such that $\mathscr{G}|_W$ is connected. The *contraction of $\mathscr{G}|_W$ in $\mathscr{G}$* is the graph minor $\mathscr{G}/(\mathscr{G}|_W)$ obtained by contracting every edge in $\mathscr{G}|_W$.

***Lemma 89.*** *Let $\mathscr{G}$ be a directed acyclic graph and $W$ a subset of its vertex set such that $\mathscr{G}|_W$ is connected. The following are equivalent:*

*(a) $\mathscr{G}|_W$ is path-induced;*
*(b) $\mathscr{G}/(\mathscr{G}|_W)$ is acyclic;*
*(c) there is a topological sort of $\mathscr{G}$ in which vertices of $W$ are consecutive.*

*Moreover, under any of the equivalent conditions, there is a bijection between*

- *topological sorts of $\mathscr{G}$ in which vertices of $W$ are consecutive,*
- *pairs of a topological sort of $\mathscr{G}|_W$ and a topological sort of $\mathscr{G}/(\mathscr{G}|_W)$.*

***Lemma 90.*** *Let $\imath \colon V \hookrightarrow U$ be an inclusion of regular molecules such that $n \coloneqq \dim U = \dim V$ and $V$ is round. Then $\mathscr{F}_{n-1}U[\langle V\rangle/\imath(V)]$ is isomorphic to $\mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$.*

***Proposition 91.*** *Let $\imath \colon V \hookrightarrow U$ be an inclusion of regular molecules such that $n \coloneqq \dim U = \dim V$ and $V$ is round. If $\imath$ is a submolecule inclusion, then $\mathscr{F}_{n-1}V$ is a path-induced subgraph of $\mathscr{F}_{n-1}U$.*

*Proof.* By Proposition 53, if $\imath$ is a submolecule inclusion then $U[\langle V\rangle/\imath(V)]$ is a regular molecule. By Corollary 81 $\mathscr{F}_{n-1}U[\langle V\rangle/\imath(V)]$ is acyclic, and by Lemma 90 it is isomorphic to $\mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$. It follows from Lemma 89 that $\mathscr{F}_{n-1}V$ is a path-induced subgraph of $\mathscr{F}_{n-1}U$. ∎

***92.*** Given an inclusion $\imath \colon V \hookrightarrow U$ of regular molecules such that $n \coloneqq \dim U = \dim V$ and $V$ is round, the vertices of $\mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$ are either

- $x \in U_n \setminus \imath(V_n)$, or
- $x_V$, obtained from contracting all vertices in $\imath(V_n)$.

The following results will justify our algorithm for the rewritable submolecule problem.

***Lemma 93.*** *Let $\imath \colon V \hookrightarrow U$ be an inclusion of regular molecules such that $n \coloneqq \dim U = \dim V$ and $V$ is round, and let $(y^{(i)})_{i=1}^p$ be an $(n-1)$-ordering induced by an $(n-1)$-layering of $V$. The following are equivalent:*

*(a) $\imath$ is a submolecule inclusion;*
*(b) there exist an $(n-1)$-ordering $(x^{(i)})_{i=1}^m$ induced by an $(n-1)$-layering $(U^{(i)})_{i=1}^m$ of $U$, and $q \in \{1, \dots, m\}$ such that*
  *1. $(x^{(i)})_{i=q}^{p+q-1} = (\imath(y^{(i)}))_{i=1}^p$,*
  *2. $\imath(\partial^- V) \sqsubseteq \partial^- U^{(q)}$.*

***Theorem 94.*** *Let $\imath \colon V \hookrightarrow U$ be an inclusion of regular molecules such that $n \coloneqq \dim U = \dim V$ and $V$ is round, $m \coloneqq |U_n|$, $p \coloneqq |V_n|$. The following are equivalent:*

*(a) $\imath$ is a submolecule inclusion;*
*(b) there is a topological sort $((x^{(i)})_{i=1}^{q-1}, x_V, (x^{(i)})_{i=q+1}^{m-p+1})$ of $\mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$ such that, letting*

$$U^{(0)} \coloneqq \partial^- U,$$
$$U^{(q)} \coloneqq \partial_{n-1}^+ U^{(q-1)} \cup \imath(V),$$
$$U^{(i)} \coloneqq \partial_{n-1}^+ U^{(i-1)} \cup \operatorname{cl}\left\{x^{(i)}\right\} \quad \text{for } i \neq q,$$

*we have $\imath(\partial^- V) \sqsubseteq \partial^- U^{(q)}$ and $\partial^- x^{(i)} \sqsubseteq \partial^- U^{(i)}$ for all $i \neq q$.*

*Sketch of proof.* By Proposition 53, if $\imath$ is a submolecule inclusion then $U[\langle V\rangle/\imath(V)]$ is a regular molecule, so by Corollary 78 it admits an $(n-1)$-layering inducing an $(n-1)$-ordering. By Lemma 90 this can be identified with a topological sort of $\mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$. The properties of the $U^{(i)}$ follow from the properties of the layering of $U[\langle V\rangle/\imath(V)]$, as established by Proposition 85, after we "reverse" the substitution of $\langle V\rangle$ for $\imath(V)$, producing a regular molecule isomorphic to $U$.

Conversely, if $(y^{(i)})_{i=1}^p$ is an $(n-1)$-ordering induced by an $(n-1)$-layering of $V$, then $((x^{(i)})_{i=1}^{q-1}, (y^{(i)})_{i=1}^p, (x^{(i)})_{i=q+1}^{m-p+1})$ is an $(n-1)$-ordering of $U$, which by the criterion of Proposition 85 is induced by an $(n-1)$-layering. We conclude by Lemma 93. ∎

***95*** (Rewritable submolecule decision algorithm)**.** The procedure takes as input an inclusion $V \subseteq U$ of regular molecules such that $n \coloneqq \dim U = \dim V$ and $V$ is round, and it returns whether $V \sqsubseteq U$. We let $m \coloneqq |U_n|$ and $p \coloneqq |V_n|$.

We construct the graph $\mathscr{G} \coloneqq \mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$. Then we start a loop. At each iteration, we search for a new topological sort of $\mathscr{G}$. If we cannot find one, we return *false*. Else, let $((x^{(i)})_{i=1}^{q-1}, x_V, (x^{(i)})_{i=q+1}^{m-p+1})$ be the new topological sort, and let $(U^{(i)})_{i=1}^{m-p+1}$ be as in Theorem 94.

For each $i \in \{1, \dots, m-p+1\}$, we start a recursive call to the algorithm to decide whether $\partial^- x^{(i)} \sqsubseteq \partial^- U^{(i)}$ if $i \neq q$, and $\partial^- V \sqsubseteq \partial^- U^{(q)}$ if $i = q$. If this returns *false*, we break

the iteration on $i$ and iterate the main loop. If this returns *true*, we iterate on $i$. At the end of the iteration on $i$, we return *true*.

***Theorem 96.*** *The rewritable submolecule decision algorithm is correct: it always terminates, and returns* true *if and only if* $V \sqsubseteq U$.

*Proof.* We proceed by induction on the dimension $n$ of $U$ and $V$. If $n = 0$, this is straightforward, so let $n > 0$.

The number of iterations of the main loop is bounded by the number of topological sorts of $\mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$, which is finite. Consider one such iteration, producing a topological sort $((x^{(i)})_{i=1}^{q-1}, x_V, (x^{(i)})_{i=q+1}^{m-p+1})$. Let us write $V^{(q)} \coloneqq V$ and $V^{(i)} \coloneqq \operatorname{cl}\{x^{(i)}\}$ for $i \neq q$. For all $i \in \{1, \dots, m-p+1\}$, we have a call to the decision algorithm with input $\partial^- V^{(i)} \subseteq \partial^- U^{(i)}$, assuming that the calls for $j < i$ all returned *true*.

Now, $\partial^- V^{(i)}$ is round by Proposition 34 and Proposition 43. Moreover, $\partial^- U^{(1)} = \partial^- U$, which is a regular molecule. For $i > 1$, assuming that $\partial^- V^{(i-1)} \sqsubseteq \partial^- U^{(i-1)}$, we may apply Proposition 53 to derive that $U^{(i-1)}$ and $\partial^+ U^{(i-1)} = \partial^- U^{(i)}$ are regular molecules. Thus

1) the input of the first call is well-formed,
2) for $i > 1$, assuming that the $(i-1)$-th call correctly returned *true*, the input of the $i$-th call is well-formed.

Since all of these are in dimension $(n-1)$, by the inductive hypothesis, each call terminates returning the correct answer. By Theorem 94, this proves both correctness and termination in dimension $n$. ∎

***Theorem 97.*** *The rewritable submolecule problem in dimension $n$ can be solved in time*

$$O\left(\prod_{k \leq n} |U_k|! \, |U_k|\right).$$

*Proof.* For $n = 0$, this is obvious, so let $n > 0$. The number of iterations of the main loop is bounded above by the number of topological sorts of $\mathscr{G} \coloneqq \mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$. This reaches its maximum when $\mathscr{G}$ is a discrete graph, in which case the number is $(|U_n| - |V_n| + 1)!$, tightly bounded above by $|U_n|!$.

At each iteration of the main loop, we have at most $|U_n| - |V_n| + 1$ calls to the algorithm on regular molecules of dimension $n-1$ contained in $U$. By the inductive hypothesis, these take time $O(\prod_{k \leq n-1} |U_k|! \, |U_k|)$.

All other operations have lower complexity: both finding topological sorts and computing the boundaries of the $U^{(i)}$ take linear time in $|\mathscr{E}_n U|$, but this can be bounded above by $|U_n| \, |U_{n-1}|$, and we conclude. ∎

***98.*** The superpolynomial upper bound on the rewritable submolecule problem leaves it inconclusive whether subdiagram matching admits a polynomial-time algorithm in arbitrary dimension. Nevertheless, we are at least able to prove that the problem is in NP.

***Proposition 99.*** *For all $n \in \mathbb{N}$, the $n$-dimensional subdiagram matching problem is in* NP.

*Proof.* It suffices to prove by induction on $n$ that the rewritable submolecule problem in dimension $n$ is in NP. When $n = 0$, the problem is trivial. In dimension $n > 0$, a polynomial-size certificate that $V \sqsubseteq U$ is given by

1) a topological sort $((x^{(i)})_{i=1}^{q-1}, x_V, (x^{(i)})_{i=q+1}^{m-p+1})$ of the graph $\mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$, and
2) polynomial-size certificates that $\partial^- V \sqsubseteq \partial^- U^{(q)}$ and $\partial^- x^{(i)} \sqsubseteq \partial^- U^{(i)}$ for all $i \neq q$,

with the notations of Theorem 94. By the inductive hypothesis this exists and is verifiable in polynomial time. ∎

***100.*** We conclude this section by considering some improvements on the subdiagram matching algorithm conditional on *acyclicity* properties.

***101*** (Frame-acyclic molecule)**.** Let $U$ be a regular molecule. We say that $U$ is *frame-acyclic* if for all submolecules $V \sqsubseteq U$, if $r \coloneqq \operatorname{frdim} V$, then $\mathscr{M}_r V$ is acyclic.

***Theorem 102.*** *Let $U$ be a regular molecule. The following are equivalent:*

*(a) $U$ is frame-acyclic;*
*(b) for all $V \sqsubseteq U$ and all $\operatorname{frdim} V \leq k < \dim V$, $V$ admits a $k$-layering;*
*(c) for all $V \sqsubseteq U$ and all $\operatorname{frdim} V \leq k < \dim V$, the sets $\mathscr{L}\mathit{ay}_k V$ and $\mathscr{O}\mathit{rd}_k V$ are non-empty and equinumerous.*

***103*** (Stably frame-acyclic molecule)**.** Let $U$ be a regular molecule. We say that $U$ is *stably frame-acyclic* if for all submolecules $V \sqsubseteq U$ and all rewritable submolecules $W \sqsubseteq V$, the regular molecule $V[\langle W \rangle / W]$ is frame-acyclic.

***104.*** Every stably frame-acyclic regular molecule is frame-acyclic: if we take $V \sqsubseteq U$ to be an atom, the substitution $U[\langle V \rangle / V]$ is trivial. Moreover, every submolecule of a (stably) frame-acyclic regular molecule is (stably) frame-acyclic.

We are not aware of examples of regular molecules that are stably frame-acyclic but not frame-acyclic (as we will see in the next section, any such example is at least 4-dimensional), so we cannot exclude that the two classes coincide, but neither it seems clear that they do.

***105.*** In general, frame-acyclicity seems difficult to check. However, it is implied by stronger acyclicity conditions that are easier to check. We do not know any easily verifiable sufficient conditions for stable frame-acyclicity.

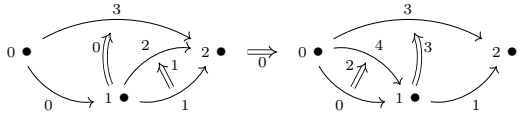
***106*** (Acyclic molecule)**.** Let $U$ be a regular molecule. We say that $U$ is *acyclic* if $\mathring{\mathscr{H}}U$ is acyclic.

***107*** (Dimension-wise acyclic molecule)**.** Let $U$ be a regular molecule. We say that $U$ is *dimension-wise acyclic* if, for all $k \in \mathbb{N}$, $\mathscr{F}_k U$ is acyclic.

***Lemma 108.*** *Let $U$ be a regular molecule. Then*

*1) if $U$ is acyclic, it is dimension-wise acyclic;*
*2) if $U$ is dimension-wise acyclic, it is frame-acyclic.*

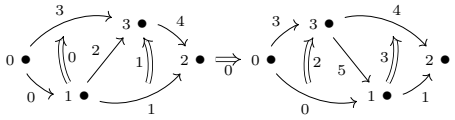
*Example 109.* Both implications are strict. The 3-dimensional atom



(based on [8, Fig. 2]) is not acyclic, since its oriented Hasse diagram contains the cycle

$$(0,1) \to (1,1) \to (2,1) \to (3,0) \to (2,2) \to (1,4) \to (0,1),$$

but it is dimension-wise acyclic. The 3-dimensional atom



(based on [8, Fig. 4]) is not dimension-wise acyclic, since its 0-flow graph contains the cycle $(1,2) \to (1,5) \to (1,2)$, but it is frame-acyclic by Theorem 121.

***Proposition 110.*** *If $U$ is guaranteed to be frame-acyclic, the rewritable submolecule problem in dimension $n$ can be solved in time*

$$O\left(\prod_{k \le n} |U_k|!\right).$$

*Proof.* Given a topological sort $((x^{(i)})_{i=1}^{q-1}, x_V, (x^{(i)})_{i=q+1}^{m-p+1})$ of $\mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$, by Lemma 89, substituting any $(n-1)$-ordering of $V$ for $x_V$ produces an $(n-1)$-ordering of $U$ in which the elements of $V$ are consecutive. By Theorem 102, this is induced by an $(n-1)$-layering of $U$. By Lemma 93, it suffices to check that $\partial^- V \sqsubseteq \partial^- U^{(q)}$ to conclude that $V \sqsubseteq U$, so we have a single recursive call instead of $O(|U_n|)$ many. Since $\partial^- U^{(q)} \sqsubseteq U^{(q)} \sqsubseteq U$, it is frame-acyclic, and we can proceed inductively. ∎

***Proposition 111.*** *If $U$ is guaranteed to be stably frame-acyclic, the rewritable submolecule problem can be solved in linear time in the size of $\mathscr{H}U$.*

*Proof.* If $V \sqsubseteq U$, by assumption $U[\langle V \rangle / V]$ is frame-acyclic. By Theorem 102, *all* its $(n-1)$-orderings are induced by $(n-1)$-layerings, and by Lemma 90 they are in bijection with topological sorts of $\mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$. It follows that, if any topological sort of $\mathscr{F}_{n-1}U/\mathscr{F}_{n-1}V$ fails to satisfy the conditions of Theorem 94, then $V$ is not a submolecule of $U$, so in the decision algorithm we can stop after the first iteration of the main loop.

This involves finding a single topological sort and computing $\partial^- U^{(q)}$, both of which take time $O(|\mathscr{E}_n U|)$; then, as in Proposition 110, we make a single call to the decision algorithm for $\partial^- V \sqsubseteq \partial^- U^{(q)}$. Since $\partial^- U^{(q)} \sqsubseteq U^{(q)} \sqsubseteq U$, it is stably frame-acyclic, and we can proceed inductively. ∎

***Corollary 112.*** *The subdiagram matching problem restricted to diagrams with stably frame-acyclic shape is in* P.

***113.*** Proposition 111 begs the question: why not just develop a higher-dimensional rewriting theory around stably frame-acyclic shapes of diagrams? The reason is that, in general, acyclicity properties are *global* properties of diagrams, that are not stable under local substitutions, essential to rewriting theory. Indeed, the inductive definition of regular molecules makes them "minimal" for a class of shapes closed under pasting and rewrites of round diagrams, with roundness seemingly the natural condition ensuring both topological soundness and a good combinatorial account of substitution. Any further restriction would almost certainly be impractical from a rewriting-theoretic perspective.

Nevertheless, the following section will show that *up to dimension 3* there is no restriction at all: all regular molecules are stably frame-acyclic, and in fact we can further simplify our algorithms.

## IV. In low dimensions

***114.*** Some results in this section come from [17, Section 3].

***Lemma 115.*** *Let $U$ be a 1-dimensional regular molecule, $m := |U_1|$. Then $\mathring{\mathscr{H}}U$ is a linear graph with $(2m+1)$ vertices, and $\mathscr{F}_0 U$ is a linear graph with $m$ vertices.*

***Proposition 116.*** *Let $\imath \colon V \hookrightarrow U$ be an inclusion of regular molecules, $\dim V = \dim U = 1$. Then $\imath$ is a submolecule inclusion.*

*Proof.* By Lemma 65 $\mathscr{F}_0 V$ is an induced subgraph of $\mathscr{F}_0 U$. By Lemma 115 both of them are linear graphs, and an induced subgraph of a linear graph is a linear graph if and only if its vertices are consecutive in the ambient graph. All other conditions of Lemma 93 are trivially satisfied. ∎

***Proposition 117.*** *Let $U$ be a regular molecule, $\dim U \le 2$. Then $U$ is acyclic.*

***Lemma 118.*** *Let $\imath \colon V \hookrightarrow U$ be an inclusion of regular molecules, $\dim U \le 2$. Then $\mathscr{F}_1 V$ is a path-induced subgraph of $\mathscr{F}_1 U$.*

***Theorem 119.*** *Let $\imath \colon V \hookrightarrow U$ be an inclusion of regular molecules such that $\dim U = \dim V = 2$ and $V$ is round. Then $\imath$ is a submolecule inclusion.*

*Proof.* By Lemma 118 combined with Lemma 89, there exists a 1-ordering $(x^{(i)})_{i=1}^m$ of $U$ in which the elements of $\imath(V)$ are consecutive, that is, $x^{(i)} \in \imath(V)$ if and only if $p \le i \le q$ for some $p, q \in \{1, \dots, m\}$.

By Proposition 117 $U$ is acyclic, so by Lemma 108 it is frame-acyclic, and by Theorem 102 the 1-ordering comes from a 1-layering $(U^{(i)})_{i=1}^m$ such that $\imath(\partial^- V) \subseteq \partial^- U^{(p)}$. Since both are 1-dimensional regular molecules, by Proposition 116 $\imath(\partial^- V) \sqsubseteq \partial^- U^{(p)}$, and Lemma 93 allows us to conclude. ∎

***Corollary 120.*** *The rewritable submolecule problem in dimension $\le 2$ has a trivial constant-time solution.*

***Theorem 121.*** *Let $U$ be a regular molecule, $\dim U \le 3$. Then $U$ is stably frame-acyclic.*

***Theorem 122.*** *Let $\imath\colon V \hookrightarrow U$ be an inclusion of regular molecules such that $\dim U = \dim V = 3$ and $V$ is round. The following are equivalent:*

*(a) $\imath$ is a submolecule inclusion;*
*(b) $\mathscr{F}_2 V$ is a path-induced subgraph of $\mathscr{F}_2 U$.*

*Proof.* One implication is Proposition 91, so we only need to prove the converse. By Lemma 89, if $\mathscr{F}_2 V$ is path-induced, then there exists a 2-ordering $(x^{(i)})_{i=1}^m$ of $U$ in which the elements of $\imath(V)$ are consecutive, that is, $x^{(i)} \in \imath(V)$ if and only if $p \le i \le q$ for some $p, q \in \{1, \dots, m\}$. By Theorem 121, $U$ is frame-acyclic, so by Theorem 102 the 2-ordering comes from a 2-layering $(U^{(i)})_{i=1}^m$ such that $\imath(\partial^- V) \subseteq \partial^- U^{(p)}$. Since both are 2-dimensional regular molecules and $\partial^- V$ is round, by Theorem 119 $\imath(\partial^- V) \sqsubseteq \partial^- U^{(p)}$, and we conclude by Lemma 93. ∎

***Theorem 123.*** *The rewritable submolecule problem in dimension 3 can be solved in time $O(|\mathscr{E}_3 U|)$.*

*Proof.* By Theorem 122 combined with Lemma 89, it suffices to construct $\mathscr{F}_2 U / \mathscr{F}_2 V$ and check if it is acyclic. The first can be done while traversing the induced subgraph of $\mathscr{H}U$ on $U_3 \cup U_2$, which takes time $O(|\mathscr{E}_3 U|)$. Both the number of vertices and the number of edges of $\mathscr{F}_2 U / \mathscr{F}_2 V$ is also $O(|\mathscr{E}_3 U|)$, and we conclude. ∎

***124.*** To match a diagram $s\colon V \to \mathbb{V}$ in $t\colon U \to \mathbb{V}$ in dimension 3, according to our results we need

- $O(|U_3|\ |V_3|\ |V|\ |\mathscr{E}_\vee V|\ \log |\mathscr{E}_\vee V|)$ time to find all inclusions $V \hookrightarrow U$, of which there are $O(|U_3|)$,
- $O(|\mathscr{E}_3 U|)$ time to check whether each of them is a submolecule inclusion,
- $O(|V|)$ time to compare labellings on each, assuming labels can be compared in constant time,

leading to an overall

$$O(|U_3|\,(|\mathscr{E}_3 U| + |V_3|\ |V|\ |\mathscr{E}_\vee V|\ \log |\mathscr{E}_\vee V|))$$

upper bound. Here we used the bound on molecule matching in generic dimension; it is possible that this can be improved by using strategies tailored to dimension 3, as it is certainly the case in dimension $\le 2$.

If we consider a machine operating by rewriting 3-dimensional diagrams, which has a fixed finite list of rewrite rules, the variables linked to $V$ can be considered as constant parameters of the machine. Our results then imply that such a machine can be simulated with $O(|U_3|\ |\mathscr{E}_3 U|)$ overhead in a standard model of computation.

***125.*** We leave the existence of a polynomial algorithm for subdiagram matching in dimension 4 or higher as an open problem. The main obstacle to overcome is the expensive iteration on topological sorts, motivated by the fact that, from dimension 4 onwards, not all of them arise from layerings. One may hope, perhaps, that this is due to flow graphs "missing" some relations, and that it should be possible to supplement them with extra information, in such a way as to restore the bijective correspondence between layerings and topological sorts. Unfortunately, this cannot be the case in general, as the following counterexample shows.

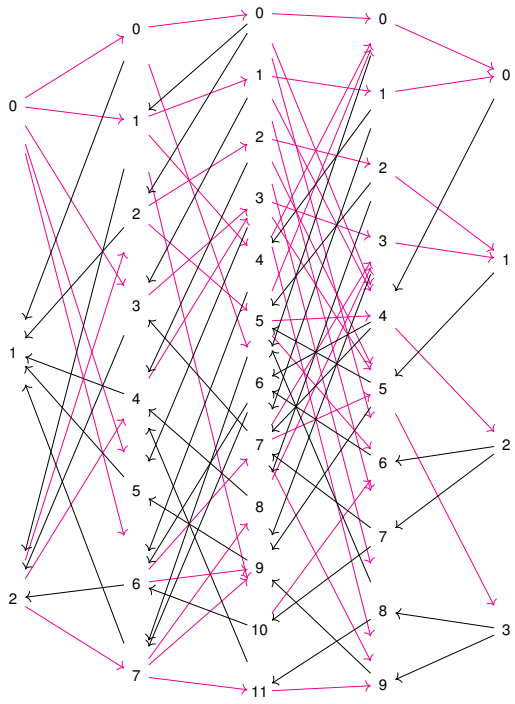


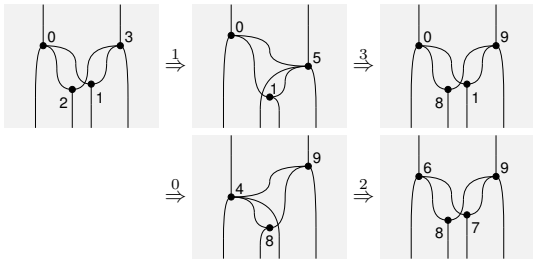
Figure 1. Oriented Hasse diagram of Example 126.

*Example 126.* This example is a 4-dimensional regular molecule $U$ which is not frame-acyclic. Its Hasse diagram is given in Figure 1, with dimensions increasing from left to right.

To understand this example, we can picture the 4-dimensional elements of $U$ as rewrites of planar projections of 3-dimensional diagrams, portrayed as string diagrams. Then $U$ has one 3-layering inducing the 3-ordering

$$((4,1),(4,3),(4,0),(4,2))$$

corresponding to the sequence of rewrite steps



where nodes represent 3-dimensional elements and incoming and outgoing wires represent their input and output faces. It has one other 3-layering inducing the 3-ordering

$$((4,0),(4,2),(4,1),(4,3)).$$

These are the only two 3-layerings of $U$. Indeed, the application of the rewrite $(4,1)$ creates a "non-convexity" in the input boundary of $(4,0)$, in the form of a path $(3,1) \to (3,5) \to (3,0)$ in $\mathscr{F}_2 U$, which can also be spotted as the upward path $1 \to 5 \to 0$ in the second string diagram. Before $(4,0)$ can be applied, this needs to be resolved by the application of $(4,3)$, which removes the non-convexity. Similarly, if $(4,0)$ is applied first, it creates a non-convexity in the input boundary of $(4,1)$, a path $(3,2) \to (3,4) \to (3,3)$ in $\mathscr{F}_2 U$, which needs to be resolved by the application of $(4,2)$.

Now we have examples of all the following.

1) *A 3-ordering that is not induced by a 3-layering.* The graph $\mathscr{F}_3 U$ is simply

$$(4,0) \longrightarrow (4,2)$$

$$(4,1) \longrightarrow (4,3)$$

so $U$ admits four other 3-orderings which do not determine 3-layerings; for example, $((4,0),(4,1),(4,2),(4,3))$. What is more, there is no extension of $\mathscr{F}_3 U$, and more in general no graph whose vertex set is $U_4$, whose topological sorts correspond to the 3-layerings of $U$.

2) *A regular molecule that is not frame-acyclic.* We can deduce that $U$ is not frame-acyclic using Theorem 102. More directly, $V \coloneqq \operatorname{cl}\{(4,0),(4,3)\}$ is a submolecule of $U$ such that $\operatorname{frdim} V = 2$, but $\mathscr{M}_2 V$ contains a cycle $(4,3) \to (4,0) \to (4,3)$, since

$$(2,5) \in \Delta_2^+(4,3) \cap \Delta_2^-(4,0),$$
$$(2,7) \in \Delta_2^+(4,0) \cap \Delta_2^-(4,3).$$

3) *An inclusion of a round regular molecule that is not a submolecule inclusion.* We have that $\partial^-(4,0) = \operatorname{cl}\{(3,0),(3,1)\}$ is a round 3-dimensional regular molecule, included in the 3-dimensional regular molecule $W \coloneqq \operatorname{cl}\{(3,0),(3,1),(3,5)\}$. However, it is *not* a submolecule of $W$, due to the presence of the path $(3,1) \to (3,5) \to (3,0)$ in $\mathscr{M}_2 W$.

## Conclusions and outlook

We have taken the first steps into the study of machines based on higher-dimensional rewriting in all dimensions. We have presented algorithms by which they could be simulated in standard models of computation, and shown that this requires only low-degree polynomial time overhead in dimension $\leq 3$.

Feasibility in dimension $> 3$ is the most obvious open question. We hope that a deeper understanding of cases like Example 126 will lead either to an improved algorithm, or to a proof of NP-completeness. The way in which 4-dimensional rewrites can introduce obstructions to "disjoint" rewrites, in a non-local way, may be a hint that the latter is more likely. In either case, we are actively working on the problem.

Beyond the more immediate questions, we hope to have laid the groundwork for an approach to complexity theory based on higher-dimensional rewriting, that leverages its unique characteristics as described in the introduction. For instance, we believe that the coexistence of higher algebraic structures and rewrite systems within the same category, as made possible by the theory of diagrammatic sets or their variants, may lead to a unified and compositional understanding of interpretations of rewrite systems, such as polynomial and matrix interpretations, which are one of the key techniques in implicit computational complexity. We plan to develop various aspects of this programme in future work.

*Acknowledgements*

This work was supported by the ESF funded Estonian IT Academy research measure (project 2014-2020.4.05.19-0001) and by the Estonian Research Council grant PSG764. We thank the anonymous referees for their helpful comments on an earlier draft.

## References

[1] A. Burroni, "Higher-dimensional word problems with applications to equational logic," *Theoretical Computer Science*, vol. 115, no. 1, pp. 43–62, 1993.

[2] Y. Guiraud, "*Rewriting methods in higher algebra*," Thèse d'habilitation à diriger des recherches, Université Paris 7, 2019.

[3] F. Lawvere, "Functorial semantics of algebraic theories," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 50, no. 5, p. 869, 1963.

[4] S. Gaussent, Y. Guiraud, and P. Malbos, "Coherent presentations of Artin monoids," *Compositio Mathematica*, vol. 151, no. 5, pp. 957–998, 2015.

[5] D. Reutter and J. Vicary, "High-level methods for homotopy construction in associative $n$-categories," in *2019 34th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. IEEE, 2019, pp. 1–13.

[6] M. Freedman, A. Kitaev, M. Larsen, and Z. Wang, "Topological quantum computation," *Bulletin of the American Mathematical Society*, vol. 40, no. 1, pp. 31–38, 2003.

[7] A. Hadzihasanovic, "Diagrammatic sets and rewriting in weak higher categories," *arXiv preprint arXiv:2007.14505*, 2020.

[8] R. Steiner, "The algebra of directed complexes," *Applied Categorical Structures*, vol. 1, no. 3, pp. 247–284, 1993.

[9] A. Hadzihasanovic and D. Kessler, "Data structures for topologically sound higher-dimensional diagram rewriting," in *Proceedings Fifth International Conference on Applied Category Theory (ACT2022)*, 2022.

[10] G. Bonfante and Y. Guiraud, "Polygraphic programs and polynomial-time functions," *Logical Methods in Computer Science*, vol. 5, no. 2: 14, pp. 1–37, 2009.

[11] F. Bonchi, F. Gadducci, A. Kissinger, P. Sobocinski, and F. Zanasi, "String diagram rewrite theory I: Rewriting with Frobenius structure," *Journal of the ACM (JACM)*, vol. 69, no. 2, pp. 1–58, 2022.

[12] ——, "String diagram rewrite theory II: Rewriting with symmetric monoidal structure," *Mathematical Structures in Computer Science*, vol. 32, no. 4, pp. 511–541, 2022.

[13] ——, "String diagram rewrite theory III: Confluence with and without Frobenius," *arXiv preprint arXiv:2109.06049*, 2021.

[14] A. Delpeuch and J. Vicary, "The word problem for braided monoidal categories is unknot-hard," *arXiv preprint arXiv:2105.04237*, 2021.

[15] J. Vicary and A. Delpeuch, "Normalization for planar string diagrams and a quadratic equivalence algorithm," *Logical Methods in Computer Science*, vol. 18, 2022.

[16] A. Hadzihasanovic, "Higher-categorical pasting diagrams," to appear.

[17] ——, "The smash product of monoidal theories," in *2021 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. IEEE, 2021.

## APPENDIX
## ADDITIONAL PROOFS

### A. Proofs for Section I

***127.*** Most of the proofs are taken from a forthcoming monograph on the combinatorics of pasting diagrams [16]. To avoid making this appendix longer than it already is, we cite results of [8], [7], [17] whenever possible, even though they use slightly different definitions; all the results used have been independently reproved.

*Proof of Lemma 20.* This is a combination of [7, Lemma 1.9 and Lemma 1.11]. ∎

*Proof of Proposition 22.* By [7, Lemma 1.9], there is a forgetful functor $\mathsf{U}$ from $\mathbf{ogPos}$ to the category $\mathbf{Pos}$ of posets and order-preserving maps. All these limits and colimits exist in $\mathbf{Pos}$, so it suffices to prove that they can be lifted to $\mathbf{ogPos}$.

The terminal poset is the poset with a single element, and the initial poset is the empty poset. Both of them admit a unique orientation. Let $P$ be an oriented graded poset. Both the unique map from $\mathsf{U}P$ to the terminal poset and the unique map from the initial poset trivially preserve boundaries, so they lift to maps of oriented graded posets.

Let $\imath_1 \colon Q \hookrightarrow P_1$ and $\imath_2 \colon Q \hookrightarrow P_2$ be inclusions of oriented graded posets. Computing their pushout in $\mathbf{Pos}$ determines two order-preserving maps

$$j_1 \colon \mathsf{U}P_1 \to \mathsf{U}P_1 \cup \mathsf{U}P_2, \qquad j_2 \colon \mathsf{U}P_2 \to \mathsf{U}P_1 \cup \mathsf{U}P_2.$$

Since $\mathsf{U}\imath_1$ and $\mathsf{U}\imath_2$ are closed embeddings, it is an exercise to show that $j_1$ and $j_2$ are also closed embeddings, and deduce that $\mathsf{U}P_1 \cup \mathsf{U}P_2$ is a graded poset. Since $j_1$ and $j_2$ preserve the covering relation and are jointly surjective, we can put a unique orientation on $\mathsf{U}P_1 \cup \mathsf{U}P_2$ in such a way that $j_1$ and $j_2$ both preserve orientations; overlaps are resolved by the fact that $(\mathsf{U}\imath_1); j_1 = (\mathsf{U}\imath_2); j_2$ and $\imath_1$ and $\imath_2$ preserve orientations. This choice of orientation determines a unique lift of the pushout to $\mathbf{ogPos}$. ∎

***Lemma 128.*** *Let $U$ be a closed subset of an oriented graded poset, $n \in \mathbb{N}$, and $\alpha \in \{+,-\}$. Then*

1) $(\partial_n^\alpha U)_n = \Delta_n^\alpha U$,
2) $(\mathit{Max}\,(\partial_n^\alpha U))_k = (\mathit{Max}\, U)_k$ *for all* $k < n$.

*Proof.* Let $x \in \partial_n^\alpha U$. Then by definition there exists $y$ such that $x \leq y$ and either $y \in \Delta_n^\alpha U$ or $y \in (\mathit{Max}\, U)_k$ for some $k < n$. If $x$ is maximal, necessarily $x = y$, and we obtain one inclusion. The converse inclusions are evident. ∎

***Lemma 129.*** *Let $U$ be a closed subset in an oriented graded poset, $n \in \mathbb{N}$, and $\alpha \in \{+,-\}$. Then*

1) $(\mathit{Max}\, U)_n = \Delta_n^+ U \cap \Delta_n^- U$,
2) *if* $n = \dim U$ *then* $(\mathit{Max}\, U)_n = \Delta_n^\alpha U = U_n$.

*Proof.* Let $x \in U$, $\dim x = n$. Then $x$ is maximal if and only if it has no cofaces in $U$, if and only if $\nabla^{-\alpha} x \cap U = \nabla^\alpha x \cap U = \varnothing$, if and only if $x \in \Delta_n^+ U \cap \Delta_n^- U$. If $n = \dim U$, then every element of $U_n$ is maximal in $U$, so

$$U_n = (\mathit{Max}\, U)_n \subseteq \Delta_n^\alpha U \subseteq U_n$$

using the first part of the proof, and we conclude that they are all equal. ∎

***Lemma 130.*** *Let $U, V$ be closed subsets of an oriented graded poset, $n \in \mathbb{N}$, and $\alpha \in \{+,-\}$. Then*

1) $\mathit{Max}\,(U \cup V) = (\mathit{Max}\, U \cap \mathit{Max}\, V) + (\mathit{Max}\, U \setminus V) + (\mathit{Max}\, V \setminus U)$,
2) $\Delta_n^\alpha(U \cup V) = (\Delta_n^\alpha U \cap \Delta_n^\alpha V) + (\Delta_n^\alpha U \setminus V) + (\Delta_n^\alpha V \setminus U)$.

*Proof.* Follows straightforwardly from the definitions using the decomposition $U \cup V = (U \cap V) + (U \setminus V) + (V \setminus U)$. ∎

*Proof of Lemma 26.* Identifying $U$ and $V$ with their isomorphic images, we will prove that $\partial^-(U \Rightarrow^\varphi V) = U$ and $\partial^+(U \Rightarrow^\varphi V) = V$. Let $n \coloneqq \dim U = \dim V$. By construction, we have $\Delta_n^-(U \Rightarrow^\varphi V) = U_n$ and $\Delta_n^+(U \Rightarrow^\varphi V) = V_n$.

For all $k < n$, the set $(\mathit{Max}\,(U \Rightarrow^\varphi V))_k$ is equal to $(\mathit{Max}\,(U \cup V))_k$. We claim that this is equal to both $(\mathit{Max}\, U)_k$ and $(\mathit{Max}\, V)_k$. For $k < n-1$,

$$(\mathit{Max}\, U)_k = (\mathit{Max}\, \partial^\alpha U)_k = (\mathit{Max}\, \partial^\alpha V)_k = (\mathit{Max}\, V)_k$$

by Lemma 128. For $k = n-1$, by Lemma 129

$$(\mathit{Max}\, U)_{n-1} = \Delta^- U \cap \Delta^+ U = \Delta^- V \cap \Delta^+ V = (\mathit{Max}\, V)_{n-1}.$$

We then conclude by Lemma 130. ∎

*Proof of Lemma 31.* We will prove the contrapositive. Suppose that $U$ is not pure. Then there exists a maximal element $x$ in $U$ with $k \coloneqq \dim x < \dim U$.

Since $(\mathit{Max}\, U)_k = \Delta_k^- U \cap \Delta_k^+ U$, we have $x \in \partial_k^- U \cap \partial_k^+ U$. Then $\partial_k^- U \cap \partial_k^+ U$ is $k$-dimensional and cannot be equal to $\partial_{k-1} U$, which is $(k-1)$-dimensional. It follows that $U$ is not round. ∎

*Proof of Proposition 34.* The first point follows from [7, Proposition 1.38], the third and fourth point from [7, Proposition 1.23]. The second point follows from the first and the third for the pasting construction, and [7, Lemma 2.2] for the rewrite construction. The fifth point is an immediate consequence of the definition of roundness combined with globularity. ∎

*Proof of Propositions 37, 38, 39.* These all follow from [7, Proposition 1.23] in conjunction with uniqueness of isomorphisms of regular molecules. ∎

*Proof of Proposition 43.* We prove the first point by induction. If $U$ was produced by (*Point*), then $U$ is the terminal oriented graded poset, which trivially has a greatest element. If $U$ was produced by (*Paste*), then $U$ splits into a union $V \cup W$, where $V \cap W = \partial_k^+ V = \partial_k^- W$ and $k < \max\{\dim V, \dim W\}$. Then there exist elements $x_1 \in V$ and $x_2 \in W$ such that

1) $x_1$ is maximal in $V$ and $x_2$ is maximal in $W$,
2) $\dim x_1 > k$ and $\dim x_2 > k$.

We have $\dim(V \cap W) \leq k$, so neither $x_1$ nor $x_2$ are contained in $V \cap W$. It follows that $x_1$ and $x_2$ are distinct maximal elements of $U$, so $U$ does not have a greatest element. If $U$ was produced by (*Atom*), then $U$ splits into $(U_- \cup U_+) + \{\top\}$, where $U_-$ and $U_+$ are round regular molecules of dimension $n$, and $\Delta^\alpha \top = (U_\alpha)_n$ for each $\alpha \in \{+,-\}$. By Lemma 31,

we have $U_\alpha = \mathrm{cl}\,(U_\alpha)_n$, so $U_\alpha = \partial^\alpha \top \subseteq \mathrm{cl}\,\{\top\}$. It follows that all elements of $U$ are in the closure of $x$, that is, $x$ is the greatest element of $U$.

We prove the second point also by induction. If $U$ was produced by (*Point*), then $x$ must be the unique element of $U$ whose closure is $U$ itself. If $U$ was produced by (*Paste*), it splits into $V \cup W$, and $x \in V$ or $x \in W$; the inductive hypothesis applies. If $U$ was produced by (*Atom*), it is equal to $(V \cup W) + \{\top\}$, and either $x \in V$ or $x \in W$, in which case the inductive hypothesis applies, or $x = \top$, and $\mathrm{cl}\,\{x\} = U$ is an atom by definition.

For the third point, if $U$ was produced by (*Point*), it is trivially round. If it was produced by (*Atom*), it is of the form $V \Rightarrow W$ where $\partial^- U$ is isomorphic to $V$ and $\partial^+ U$ to $W$, and by definition of the rewrite construction their intersection is uniquely isomorphic to $\partial V$ and $\partial W$; we conclude by globularity. If $\dim U > 0$, then $U$ was produced by (*Atom*), so it is of the form $V \Rightarrow W$, with $\partial^- U$ isomorphic to $V$ and $\partial^+ U$ to $W$. By uniqueness of these isomorphisms, $U$ is isomorphic to $\partial^- U \Rightarrow \partial^+ U$. ∎

### B. Proofs for Section II

*Proof of Lemma 51.* For the first point, $\partial^\alpha_n U$ is a regular molecule by Proposition 34. By Proposition 38, the pastings $U \#_n \partial^+_n U$ and $\partial^-_n U \#_n U$ are both defined and uniquely isomorphic to $U$. The inclusion of $\partial^-_n U$ into $U$ factors as the inclusion $\partial^-_n U \hookrightarrow (\partial^-_n U \#_n U)$ followed by an isomorphism, and the inclusion of $\partial^+_n U$ factors as the inclusion $\partial^+_n U \hookrightarrow (U \#_n \partial^+_n U)$ followed by an isomorphism.

For the second point, $\mathrm{cl}\,\{x\}$ is an atom by Proposition 43. We proceed by induction on the construction of $U$. If $U$ was produced by (*Point*), then $x$ must be the unique element of $U$, so $\mathrm{cl}\,\{x\} = U$. If $U$ was produced by (*Paste*), it splits into $V \cup W$ with $V, W \sqsubseteq U$, and $x \in V$ or $x \in W$. By the inductive hypothesis, $\mathrm{cl}\,\{x\} \sqsubseteq V$ or $\mathrm{cl}\,\{x\} \sqsubseteq W$. If $U$ was produced by (*Atom*), it splits into $(V \cup W) + \{\top\}$ with $V, W \sqsubseteq U$, and either $x \in V$ or $x \in W$, in which case the inductive hypothesis applies since $V, W \sqsubseteq U$ by Lemma 51, or $x = \top$, and $\mathrm{cl}\,\{x\} = U$. ∎

***Lemma 131.*** *Let $V$ be a regular molecule, $n < \dim V$, $\alpha \in \{+, -\}$. Consider a pushout diagram of the form*

$$\begin{array}{ccc} \partial^\alpha_n V & \longrightarrow & V \\ \downarrow{\scriptstyle \imath} & & \downarrow{\scriptstyle j_V} \\ U & \xrightarrow{\ j_U\ } & V \cup U \end{array}$$

*in* **ogPos**. *If $\dim U = n$ and $\imath$ is a submolecule inclusion, then*

1) *$V \cup U$ is a regular molecule,*
2) *$j_U$ maps $U$ onto $\partial^\alpha_n(V \cup U)$,*
3) *$j_V(V) \sqsubseteq V \cup U$ and $j_V(\partial_n^{-\alpha} V) \sqsubseteq \partial_n^{-\alpha}(V \cup U)$.*

*Proof.* By induction on the construction of $\imath$. If $\imath$ is an isomorphism, then $j_V$ is also an isomorphism, and all the statements are trivially satisfied.

Suppose $U$ is of the form $\partial^\alpha_n V \#_k W$ for some $k \in \mathbb{N}$, and $\imath$ is the inclusion of $\partial^\alpha_n V$ into the pasting. Since $\dim U = n$, necessarily $\dim W \leq n$, so $\partial^\alpha_n W = W$. If $k \geq n$, then also $k \geq \dim W$, and in this case $\imath$ and $j_V$ are again isomorphisms. Suppose that $k < n$. Identifying $V$ with its image through $j_V$, $V \cup U$ splits into $V \cup W$ with

$$V \cap W = \partial^\alpha_n V \cap W = \partial^-_k W = \partial^+_k(\partial^\alpha_n V) = \partial^+_k V$$

where the final equation uses globularity of $V$. This exhibits $V \cup U$ as $V \#_k W$, with $j_V$ the inclusion of $V$ into the pasting, and $j_U$ maps $\partial^\alpha_n V \#_k W$ onto $\partial^\alpha_n(V \#_k W)$ by [7, Proposition 1.23] and the axioms of strict $\omega$-categories. Similarly, $\partial_n^{-\alpha} V \sqsubseteq \partial_n^{-\alpha}(V \#_k W)$. The case where $U$ is of the form $W \#_k \partial^\alpha_n V$ is dual.

By the pasting law for pushout squares, if the statement is true of two submolecule inclusions, it is also true of their composite. ∎

***Lemma 132.*** *Let $U$ be an oriented graded poset, $V, W$ round regular molecules, and $\imath \colon V \hookrightarrow U$ an inclusion such that $U[W/\imath(V)]$ is defined. Let $\varphi \colon \partial V \xrightarrow{\sim} \partial W$ be the isomorphism used in the construction of $V \Rightarrow W$. Then $U[W/\imath(V)]$ can be constructed as the pushout*

$$\begin{array}{ccccc} \partial V & \xhookrightarrow{\ \varphi\ } & \partial W & \hookrightarrow & W \\ \downarrow & & & & \downarrow{\scriptstyle j} \\ U \setminus (\imath(V) \setminus \imath(\partial V)) & & \xhookrightarrow{\qquad} & & U[W/\imath(V)]. \end{array} \tag{3}$$

*Proof.* We can safely identify $V$ with its image through $\imath$, and treat it as a closed subset of $U$. First of all, observe that $U \setminus (V \setminus \partial V)$ is the complement of the complement of a closed subset in a closed subset, so it is closed in $U$, and well-defined as an oriented graded poset.

Let $n \coloneqq \dim U$, so $\dim(U \cup (V \Rightarrow W))$ and $\dim(V \Rightarrow W)$ are both equal to $n + 1$. Then

$$\Delta^+_n(V \Rightarrow W) = W_n, \qquad \Delta^+_n U = U_n$$

and since $U \cap (V \Rightarrow W) = V$, by Lemma 130

$$\Delta^+_n(U \cup (V \Rightarrow W)) = W_n + (U_n \setminus V_n) = W_n + (U \setminus (V \setminus \partial V))_n,$$

while for all $k < n$

$$\begin{aligned}(\mathit{Max}\,(U \cup (V \Rightarrow W)))_k &= (\mathit{Max}\,U)_k = \\ &= (\mathit{Max}\,(U \setminus (V \setminus \partial V)))_k\end{aligned}$$

because both $V$ and $V \Rightarrow W$ are round, hence pure, and do not contain any maximal elements of dimension $k$.

It follows that $\partial^+(U \cup (V \Rightarrow W))$ is the union of $W$ and $(U \setminus (V \setminus \partial V))$, with intersection $\partial W = \partial V$. ∎

***Lemma 133.*** *Let $U$ be an oriented graded poset, $V, W$ round regular molecules, and $\imath \colon V \hookrightarrow U$ an inclusion such that $U[W/\imath(V)]$ is defined. Let $j \colon W \hookrightarrow U[W/\imath(V)]$ be the right side of (3). Then $(U[W/\imath(V)])[V/j(W)]$ is defined and isomorphic to $U$.*

*Proof.* Since $W \Rightarrow V$ is defined whenever $V \Rightarrow W$ is defined, it follows that $(U[W/\imath(V)])[V/j(W)]$ is defined.

The isomorphism with $U$ is straightforward algebra of closed subsets using Lemma 132 twice. ∎

*Proof of Proposition 53.* If $\imath$ is a submolecule inclusion, by Lemma 131 $U \cup (V \Rightarrow W)$ and its output boundary $U[W/\imath(V)]$ are regular molecules, and the inclusion of $W$ into $U[W/\imath(V)]$ is a submolecule inclusion.

If $V$ is a round regular molecule, then $\langle V \rangle$ is an atom, which is round by Proposition 43, and has boundaries isomorphic to those of $V$ by Lemma 26. Then $V \Rightarrow \langle V \rangle$ is defined, so the fourth condition is a special case of the third one.

Finally, suppose $U[\langle V \rangle / \imath(V)]$ is a regular molecule. By Lemma 51, since $\langle V \rangle$ is an atom, its inclusion $j$ into $U[\langle V \rangle / \imath(V)]$ is a submolecule inclusion.

Using Lemma 131 as in the first part, we deduce that $(U[\langle V \rangle / \imath(V)])[V / j(\langle V \rangle)]$ is a regular molecule, and the inclusion of $V$ into it is a submolecule inclusion. By Lemma 133, $(U[\langle V \rangle / \imath(V)])[V / j(\langle V \rangle)]$ is isomorphic to $U$, and $\imath$ factors as this submolecule inclusion followed by an isomorphism. ∎

*Proof of Lemma 60.* This is [7, Lemma 1.16]. ∎

*Proof of Lemma 65.* It follows from Lemma 20 that, for all $x \in V$ and $\alpha \in \{+,-\}$, the set $\Delta^\alpha_k x$ is isomorphic to $\Delta^\alpha_k \imath(x)$. It follows that, for all $x, y \in \bigcup_{i>k} V_i$, there is an edge between $x$ and $y$ in $\mathscr{F}_k V$ if and only if there is an edge between $\imath(x)$ and $\imath(y)$ in $\mathscr{F}_k U$. ∎

***134.*** The following proof uses results proved in the following sections; none of their proofs use it, so there is no circularity.

*Proof of Proposition 66.* First of all, if $U$ is round, then it is pure, so the vertices of $\mathscr{F}_{n-1}U$ are the elements of $U_n$. If $U$ is an atom, then $\mathscr{F}_{n-1}U$ consists of a single vertex and no edges, so it is trivially connected. In particular this is true when $n = 0$ since $U$ is then the point, so we can proceed by induction on $n$.

Suppose $n > 0$ and $|U_n| > 1$, which by Lemma 136 implies $\operatorname{lydim} U = n - 1$. Assume by way of contradiction that $\mathscr{F}_{n-1}U$ is not connected. Then there is a bipartition $U_n = A + B$ such that there are no edges in $\mathscr{F}_{n-1}U$ between vertices in $A$ and vertices in $B$. By Lemma 60, no element of $U_{n-1}$ can be covered by two elements with the same orientation, so this implies that $\dim(\operatorname{cl} A \cap \operatorname{cl} B) < n - 1$. Let

$$A' \coloneqq \left\{x \in \Delta^- U \mid \nabla^- x \subseteq A\right\},$$
$$B' \coloneqq \left\{x \in \Delta^- U \mid \nabla^- x \subseteq B\right\}.$$

Then $A' + B'$ is a bipartition of $\Delta^- U$. By Proposition 34, $\partial^- U$ is round, so by the inductive hypothesis $\mathscr{F}_{n-2}(\partial^- U)$ is connected. It follows that there exist $\alpha \in \{+,-\}$, $x \in A'$, $y \in B'$, and $z \in U_{n-2}$ such that $z \in \Delta^\alpha x \cap \Delta^{-\alpha} y$. Then $z$ has two distinct cofaces in $\partial^- U$, so by Lemma 60 $z \notin \partial(\partial^- U) = \partial_{n-2} U$. We claim that $z \in \partial^+ U$, contradicting the roundness of $U$.

By Theorem 77, there exists an $(n-1)$-layering $(U^{(i)})_{i=1}^m$ of $U$; we will identify the $U^{(i)}$ with their isomorphic images in $U$. Let $V_0 \coloneqq \partial^- U$ and $V_i \coloneqq \partial^+ U^{(i)}$ for each $i \in \{1, \dots, m\}$. We will prove that, for all $i \in \{0, \dots, m\}$,

1) $z \in V_i$,
2) there exist elements $x_i \in \operatorname{cl} A$ and $y_i \in \operatorname{cl} B$ such that $\nabla^\alpha z \cap V_i = \{x_i\}$ and $\nabla^{-\alpha} z \cap V_i = \{y_i\}$.

For $i = 0$, we have already established this with $x_0 \coloneqq x$, $y_0 \coloneqq y$. Let $i \geq 0$, and assume this holds for $i - 1$. By Lemma 73, there is a single $n$-dimensional element $x^{(i)}$ in $U^{(i)}$, and by Lemma 144

$$V_i = \partial^- U^{(i)}[\partial^+ x^{(i)} / \partial^- x^{(i)}] = V_{i-1}[\partial^+ x^{(i)} / \partial^- x^{(i)}].$$

Suppose $x^{(i)} \in A$. Then $y_{i-1} \notin \operatorname{cl}\left\{x^{(i)}\right\}$, so $y_{i-1} \in V_i$, and we let $y_i \coloneqq y_{i-1}$. If $x_{i-1} \notin \operatorname{cl}\left\{x^{(i)}\right\}$ then also $x_{i-1} \in V_i$, and we let $x_i \coloneqq x_{i-1}$. Otherwise, $x_{i-1}$ is the only coface of $z$ in $\partial^- x^{(i)}$, so by Lemma 60 we have $z \in \partial^\alpha(\partial^- x^{(i)}) = \partial^\alpha(\partial^+ x^{(i)})$. It follows that $z \in V_i$ and there exists a unique $x_i$ such that $\nabla^\alpha z \cap \partial^+ x^{(i)} = \{x_i\}$. The case $x^{(i)} \in B$ is analogous.

Since $V_m = \partial^+ U$, we have proved that $z \in \partial^+ U$, a contradiction. ∎

### C. Proofs for Section III

***Lemma 135.*** *Let $U$ be a regular molecule, $-1 \leq k < \dim U$, and $(U^{(i)})_{i=1}^m$ a $k$-layering of $U$. For all $i < j \in \{1, \dots, m\}$,*

$$U^{(i)} \cap U^{(j)} = \partial^+_k U^{(i)} \cap \partial^-_k U^{(j)}.$$

*Proof.* Let $i < j \in \{1, \dots, m\}$, and

$$V \coloneqq U^{(1)} \#_k \dots \#_k U^{(i)},$$
$$W \coloneqq \partial^+_k U^{(i)} \#_k U^{(i+1)} \#_k \dots \#_k U^{(j-1)},$$
$$Z \coloneqq U^{(j)} \#_k \dots \#_k U^{(m)}.$$

Then $U$ splits into $V \cup (W \#_k Z)$ along the $k$-boundary, so

$$\partial^+_k U^{(i)} = \partial^+_k V = \partial^-_k (W \#_k Z) = V \cap (W \#_k Z).$$

Since $U^{(i)} \subseteq V$ and $U^{(j)} \subseteq (W \#_k Z)$, it follows that $U^{(i)} \cap U^{(j)} \subseteq \partial^+_k U^{(i)}$. Dually, from the fact that $U$ splits into $(V \#_k W) \cup Z$ along the $k$-boundary, we derive $U^{(i)} \cap U^{(j)} \subseteq \partial^-_k U^{(j)}$. ∎

*Proof of Lemma 73.* For the first point, since $\dim U^{(i)} > k$ each $U^{(i)}$ contains at least one maximal element of dimension $> k$, and because

$$\dim(U^{(i)} \cap U^{(j)}) = \dim(\partial^+_k U^{(i)} \cap \partial^-_k U^{(j)}) \leq k$$

by Lemma 135, no such maximal element is contained in two of them. Since there are exactly $m$ maximal elements of dimension $> k$, it follows that each $U^{(i)}$ contains exactly one of them.

For the second point, for all $i \in \{1, \dots, m\}$, let

$$V^{(i)} \coloneqq \partial^+_\ell U^{(1)} \#_k \dots \#_k \partial^+_\ell U^{(i-1)} \#_k \\ \#_k U^{(i)} \#_k \partial^-_\ell U^{(i+1)} \#_k \dots \#_k \partial^-_\ell U^{(m)}.$$

By repeated applications of Proposition 39 followed by Proposition 38, $U$ is isomorphic to

$$V^{(1)} \#_\ell \dots \#_\ell V^{(m)}.$$

Restricting to the subsequence of $(V^{(i)})_{i=1}^m$ on those $i$ such that $\dim V^{(i)} > \ell$, which does not change the result by Proposition 38, we obtain an $\ell$-layering of $U$. ∎

***Lemma 136.*** *Let $U$ be a regular molecule, $n \coloneqq \dim U$. Then*

*1)* $\operatorname{lydim} U \leq n-1$,
*2)* $\operatorname{lydim} U = n-1$ *if and only if* $|U_n| > 1$.

*Proof.* We have

$$\left|\bigcup_{i>n} (\mathscr{M}\mathit{ax}\, U)_i\right| = |\varnothing| = 0,$$

so $\operatorname{lydim} U \leq n-1$, with equality if and only if

$$\left|\bigcup_{i>n-1} (\mathscr{M}\mathit{ax}\, U)_i\right| = |(\mathscr{M}\mathit{ax}\, U)_n| = |U_n| > 1,$$

where we used Lemma 129. ∎

***Lemma 137.*** *Let $U$ be a regular molecule. Then* $\operatorname{lydim} U$ *is* $-1$ *if and only if $U$ is an atom.*

*Proof.* Suppose $\operatorname{lydim} U = -1$. If $\left|\bigcup_{i>0}(\mathscr{M}\mathit{ax}\, U)_i\right| = 0$, then $\dim U = 0$ and $U$ is the point.

Otherwise, $1 = \left|\bigcup_{i>0}(\mathscr{M}\mathit{ax}\, U)_i\right| = |\mathscr{M}\mathit{ax}\, U|$ because a regular molecule which is not 0-dimensional cannot have a 0-dimensional maximal element. In either case, $U$ has a greatest element. Conversely, if $U$ has a greatest element, $\left|\bigcup_{i>0}(\mathscr{M}\mathit{ax}\, U)_i\right| \leq |\mathscr{M}\mathit{ax}\, U| = 1$. ∎

***Lemma 138.*** *Let $U, V$ be regular molecules, and suppose $U \#_k V$ is defined for some $k < \min\{\dim U, \dim V\}$. Then*

$$\operatorname{lydim}(U \#_k V) \geq \max\{\operatorname{lydim} U, \operatorname{lydim} V, k\}.$$

*Proof.* Identifying $U$ and $V$ with their isomorphic images, $U \#_k V$ splits into $U \cup V$ with $\dim(U \cap V) = \dim \partial_k^+ U = k$. By Lemma 130, for all $i > k$,

$$(\mathscr{M}\mathit{ax}\,(U \#_k V))_i = (\mathscr{M}\mathit{ax}\, U)_i + (\mathscr{M}\mathit{ax}\, V)_i,$$

and since $k < \min\{\dim U, \dim V\}$, both $U$ and $V$ have at least one maximal element of dimension strictly larger than $k$. It follows that

$$\left|\bigcup_{i>k}(\mathscr{M}\mathit{ax}\,(U \#_k V))_i\right| = \\ = \left|\bigcup_{i>k}(\mathscr{M}\mathit{ax}\, U)_i\right| + \left|\bigcup_{i>k}(\mathscr{M}\mathit{ax}\, V)_i\right| \geq 2,$$

so $k - 1 < \operatorname{lydim}(U \#_k V)$, that is, $k \leq \operatorname{lydim}(U \#_k V)$. Furthermore, letting $n \coloneqq \operatorname{lydim}(U \#_k V)$, since $n + 1 > k$,

$$\left|\bigcup_{i>n+1} (\mathscr{M}\mathit{ax}\, U)_i\right| + \left|\bigcup_{i>n+1} (\mathscr{M}\mathit{ax}\, V)_i\right| = \\ = \left|\bigcup_{i>n+1} (\mathscr{M}\mathit{ax}\,(U \#_k V))_i\right| \leq 1,$$

which implies that

$$\left|\bigcup_{i>n+1} (\mathscr{M}\mathit{ax}\, U)_i\right| \leq 1, \left|\bigcup_{i>n+1} (\mathscr{M}\mathit{ax}\, V)_i\right| \leq 1.$$

Then $\max\{\operatorname{lydim} U, \operatorname{lydim} V\} \leq \operatorname{lydim}(U \#_k V)$. ∎

***Lemma 139.*** *Let $U$ be a regular molecule, $k \coloneqq \operatorname{lydim} U$. Suppose $k \geq 0$, and let $(U^{(i)})_{i=1}^m$ be a $k$-layering of $U$. Then*

*1)* $m > 1$,
*2) for each* $i \in \{1, \ldots, m\}$, $\operatorname{lydim} U^{(i)} < k$,
*3) at most one of the* $U^{(i)}$ *contains an element of dimension* $> k+1$.

*Proof.* By definition of $\operatorname{lydim} U$, if $k \geq 0$ and a $k$-layering exists, then $m > 1$, for otherwise $k - 1 \leq \operatorname{lydim} U$, a contradiction. Moreover, $U$ contains at most one element of dimension $> k+1$, which can be contained at most in one of the $U^{(i)}$. Finally, by Lemma 73, we have $\left|\bigcup_{j>k}(\mathscr{M}\mathit{ax}\, U^{(i)})_j\right| = 1$, so $\operatorname{lydim} U^{(i)} \leq k - 1 < k$. ∎

*Proof of Theorem 77.* Let $k \coloneqq \operatorname{lydim} U$. If $k = -1$, then $U$ is an atom and admits the trivial layering $U = U^{(1)}$. If $k \geq 0$, by Lemma 137 $U$ is not an atom, so we can assume that $U$ was produced by (*Paste*). Then $U$ is equal to $V \#_\ell W$ for some regular molecules $V, W$ and $\ell < \min\{\dim V, \dim W\}$. By the inductive hypothesis, we have layerings

$$V^{(1)} \#_{k_V} \cdots \#_{k_V} V^{(m_V)}, \quad W^{(1)} \#_{k_W} \cdots \#_{k_W} W^{(m_W)}$$

of $V$ and $W$, where $k_V \coloneqq \operatorname{lydim} V$ and $k_W \coloneqq \operatorname{lydim} W$. Furthermore, by Lemma 138, $k \geq \max\{k_V, k_W, \ell\}$. Let

$$n_V \coloneqq \begin{cases} m_V & \text{if } k_V = k, \\ 1 & \text{if } k_V < k \text{ and } \dim V > k, \\ 0 & \text{if } k_V < \dim V < k, \end{cases}$$

$$n_W \coloneqq \begin{cases} m_W & \text{if } k_W = k, \\ 1 & \text{if } k_W < k \text{ and } \dim W > k, \\ 0 & \text{if } k_W < \dim W < k. \end{cases}$$

Notice that it can never be the case that $n_V = n_W = 0$. We claim that we can decompose $V$ as

$$\tilde{V}^{(1)} \#_k \cdots \#_k \tilde{V}^{(n_V)} \#_k \underbrace{\partial_k^+ V \#_k \cdots \#_k \partial_k^+ V}_{n_W \text{ times}}, \tag{4}$$

where each $\tilde{V}^{(i)}$ is a regular molecule containing exactly one maximal element of dimension $> k$. If $k_V = k$, we let $\tilde{V}^{(i)} \coloneqq V^{(i)}$ for all $i \in \{1, \ldots, m_V\}$. If $k_V < k$, then $V$ contains at most one maximal element of dimension $> k_V + 1$, hence at most one maximal element of dimension $> k$. If $\dim V > k$, it contains exactly one, and we let $\tilde{V}^{(1)} \coloneqq V$. If $\dim V < k$, then $V = \partial_k^+ V$. By Proposition 38, pasting copies of $\partial_k^+ V$ does not change the result up to unique isomorphism. Similarly, we can decompose $W$ as

$$\underbrace{\partial_k^- W \#_k \cdots \#_k \partial_k^- W}_{n_V \text{ times}} \#_k \tilde{W}^{(1)} \#_k \cdots \#_k \tilde{W}^{(n_W)} \tag{5}$$

where each $\tilde{W}^{(i)}$ contains exactly one maximal element of dimension $> k$.

If $\ell = k$, since $\ell < \min\{\dim V, \dim W\}$, we have $0 < \min\{n_V, n_W\}$. Then

$$\tilde{V}^{(1)} \#_k \ldots \#_k \tilde{V}^{(n_V)} \#_k \tilde{W}^{(1)} \#_k \ldots \#_k \tilde{W}^{(n_W)}$$

is a $k$-layering of $U$. If $\ell < k$, let

$$U^{(i)} \coloneqq \begin{cases} \tilde{V}^{(i)} \#_\ell \partial_k^- W & \text{if } i \le n_V, \\ \partial_k^+ V \#_\ell \tilde{W}^{(i-n_V)} & \text{if } n_V < i \le n_V + n_W. \end{cases}$$

Since $\dim \partial_k^- V = \dim \partial_k^+ W = k$, each $U^{(i)}$ still contains exactly one maximal element of dimension $> k$. Plugging (4) and (5) in $V \#_\ell W$ and using Proposition 39 repeatedly, we deduce that $V \#_\ell W$ is isomorphic to

$$U^{(1)} \#_k \ldots \#_k U^{(n_V+n_W)},$$

which has the desired properties. Necessarily, $n_V + n_W = m$.

This proves that $U$ has a $k$-layering. It remains to show that $\operatorname{frdim} U \le k$. Let $x, y$ be distinct maximal elements of $U$. If $\min\{\dim x, \dim y\} \le k$, then $\dim(\operatorname{cl}\{x\} \cap \operatorname{cl}\{y\}) < k$.

Suppose that $k < \min\{\dim x, \dim y\}$, and let $(U^{(i)})_{i=1}^m$ be a $k$-layering of $U$. By Lemma 73 there exist $i \ne j$ such that $x \in U^{(i)}$ and $y \in U^{(j)}$. By Lemma 135, there exists $\alpha \in \{+,-\}$ such that $U^{(i)} \cap U^{(j)} = \partial_k^\alpha U^{(i)} \cap \partial_k^{-\alpha} U^{(j)}$. Then $\operatorname{cl}\{x\} \cap \operatorname{cl}\{y\} \subseteq \partial_k^\alpha U^{(i)} \cap \partial_k^{-\alpha} U^{(j)}$, which is at most $k$-dimensional. ∎

*Proof of Corollary 78.* Follows from Theorem 77 together with Lemma 136. ∎

***Lemma 140.*** *Let $U, V$ be regular molecules and suppose $U \#_k V$ is defined for some $k < \min\{\dim U, \dim V\}$. If $\mathscr{M}_k U$ and $\mathscr{M}_k V$ are acyclic, then $\mathscr{M}_k(U \#_k V)$ is acyclic.*

*Proof.* Suppose that $\mathscr{M}_k U$ and $\mathscr{M}_k V$ are acyclic. We may identify $U$ and $V$ with their images in $U \#_k V$. By Lemma 130, since $\dim(U \cap V) = k$,

$$\bigcup_{i>k} (\mathscr{M}ax\,(U \#_k V))_i = \bigcup_{i>k} (\mathscr{M}ax\, U)_i + \bigcup_{i>k} (\mathscr{M}ax\, V)_i,$$

so $\mathscr{M}_k U$ and $\mathscr{M}_k V$ are isomorphic to the induced subgraphs of $\mathscr{M}_k(U \#_k V)$ on the vertices in $U$ and $V$, respectively. It follows that a cycle in $\mathscr{M}_k(U \#_k V)$ cannot remain in $U$ or $V$, but has to visit vertices in both. In particular, such a cycle has to go through an edge from $x \in V$ to $y \in U$, induced by the existence of $z \in \Delta_k^+ x \cap \Delta_k^- y$. But then $z \notin \partial_k^- V$ and $z \notin \partial_k^+ U$, yet $z \in U \cap V$, a contradiction. ∎

*Proof of Proposition 80.* Let $(U^{(i)})_{i=1}^m$ be a $k$-layering of $U$. For each $i \in \{1, \ldots, m\}$, the graph $\mathscr{M}_k U^{(i)}$ is trivially acyclic by Lemma 73. We conclude by applying Lemma 140 repeatedly. ∎

*Proof of Proposition 83.* The function $(U^{(i)})_{i=1}^m \mapsto (x^{(i)})_{i=1}^m$ is well-defined by Lemma 73. Let $i, j \in \{1, \ldots, m\}$, and suppose that there is an edge from $x^{(i)}$ to $x^{(j)}$ in $\mathscr{M}_k U$, that is, there exists $z \in \Delta_k^+ x^{(i)} \cap \Delta_k^- x^{(j)}$. By Proposition 80, $\mathscr{M}_k U$ is acyclic, so necessarily $i \ne j$. If $j < i$, then $U^{(j)} \cap U^{(i)} \subseteq \partial_k^+ U^{(j)} \cap \partial_k^- U^{(i)}$ by Lemma 135, contradicting the existence of $z$. It follows that $i < j$, so $(x^{(i)})_{i=1}^m$ is a $k$-ordering of $U$.

Let $(V^{(i)})_{i=1}^m$ be another $k$-layering, and suppose it determines the same $k$-ordering as $(U^{(i)})_{i=1}^m$. Then the image of both $U^{(1)}$ and $V^{(1)}$ in $U$ is

$$\operatorname{cl}\left\{x^{(1)}\right\} \cup \partial^- U,$$

so $U^{(1)}$ is isomorphic to $V^{(1)}$. If $m = 1$ we are done. Otherwise, $(U^{(i)})_{i=2}^m$ and $(V^{(i)})_{i=2}^m$ are $k$-layerings inducing the same $k$-ordering on their image. By recursion, we conclude that they are layer-wise isomorphic. ∎

***Lemma 141.*** *Let $U, V, W$ be regular molecules, $k < \dim U$, $\alpha \in \{+,-\}$, and let $\imath\colon V \hookrightarrow U$ be a submolecule inclusion such that $U[W/\imath(V)]$ is defined. Then $\partial_k^\alpha U$ is isomorphic to $\partial_k^\alpha(U[W/\imath(V)])$.*

*Proof.* By Lemma 131, $U \cup (V \Rightarrow W)$ is a regular molecule and $U$ is isomorphic to its input boundary. By globularity, $\partial_k^\alpha U$ is isomorphic to

$$\partial_k^\alpha(\partial^+(U \cup (V \Rightarrow W))) = \partial_k^\alpha(U[W/\imath(V)]).$$ ∎

***Lemma 142.*** *Let $U, V, W, U', U''$ be regular molecules, let $k < \dim U$, and let $\imath\colon V \hookrightarrow U$ be a submolecule inclusion such that*

$$U \#_k U', U'' \#_k U, U[W/\imath(V)]$$

*are defined. Then*

1) *$U[W/\imath(V)] \#_k U'$ and $U'' \#_k U[W/\imath(V)]$ are defined,*
2) *if $\dim U' \le \dim U$, then $(U \#_k U')[W/\imath_U(\imath(V))]$ is defined and isomorphic to $U[W/\imath(V)] \#_k U'$,*
3) *if $\dim U'' \le \dim U$, then $(U'' \#_k U)[W/\imath_U(\imath(V))]$ is defined and isomorphic to $U'' \#_k U[W/\imath(V)]$.*

*Proof.* It follows from Lemma 141 that $U[W/\imath(V)] \#_k U'$ and $U'' \#_k U[W/\imath(V)]$ are defined.

The substitution $(U \#_k U')[W/\imath_U(\imath(V))]$ is then defined if and only if $\dim(U \#_k U') = \dim U$, equivalently, if and only if $\dim U' \le \dim U$. Similarly, $(U'' \#_k U)[W/\imath_U(\imath(V))]$ is defined if and only if $\dim U'' \le \dim U$. The isomorphisms follow straightforwardly from the definitions using the pasting law for pushout squares. ∎

***143.*** Theorem 77 in conjunction with Lemma 139 and Lemma 137 allows us to prove properties of regular molecules *by induction on their layering dimension*. That is, to prove that a property holds of all regular molecules $U$, it suffices to

- prove that it holds when $\operatorname{lydim} U = -1$, that is, when $U$ is an atom,
- prove that it holds when $k \coloneqq \operatorname{lydim} U \ge 0$, assuming that it holds of all the $(U^{(i)})_{i=1}^m$ in a $k$-layering of $U$.

***Lemma 144.*** *Let $U$ be a regular molecule, $k \in \mathbb{N}$, and suppose*

$$\bigcup_{i>k} (\mathscr{M}ax\, U)_i = \{x\}.$$

*Then, for all $\alpha \in \{+,-\}$,*

*1)* $\partial^\alpha_k x \sqsubseteq \partial^\alpha_k U$,
*2)* $\partial^\alpha_k U$ *is isomorphic to* $\partial^{-\alpha}_k U[\partial^\alpha_k x / \partial^{-\alpha}_k x]$.

*Proof.* We proceed by induction on $\operatorname{lydim} U$.

If $\operatorname{lydim} U = -1$, then $U$ is an atom and equal to $\operatorname{cl}\{x\}$. It follows that $\partial^\alpha_k x = \partial^\alpha_k U$, which is trivially a submolecule, and is isomorphic to $\partial^{-\alpha}_k U[\partial^\alpha_k x / \partial^{-\alpha}_k x]$.

Suppose $\ell \coloneqq \operatorname{lydim} U \geq 0$, and let $(U^{(i)})_{i=1}^m$ be an $\ell$-layering of $U$. Then $\ell \leq k-1 < k$ because $\left|\bigcup_{i>k} (\mathit{Max}\, U)_i\right| = 1$. By [7, Proposition 1.23] and the axioms of strict $\omega$-categories, $\partial^\alpha_k U$ is isomorphic to

$$\partial^\alpha_k U^{(1)} \#_\ell \ldots \#_\ell \partial^\alpha_k U^{(m)}.$$

Now $x$ is contained in a single $U^{(i)}$. By the inductive hypothesis, $\partial^\alpha_k x \sqsubseteq \partial^\alpha_k U^{(i)}$, and the latter is isomorphic to $\partial^{-\alpha}_k U^{(i)}[\partial^\alpha_k x / \partial^{-\alpha}_k x]$. We conclude by Lemma 142. ∎

*Proof of Proposition 85.* Suppose $(U^{(i)})_{i=1}^m$ is a $k$-layering. Then, for all $i \in \{1,\ldots,m\}$, $U^{(i)}$ is a regular molecule, and by Proposition 73 $x^{(i)}$ is the only element of dimension $> k$ in $U^{(i)}$. By Lemma 144, $\partial^-_k x^{(i)} \sqsubseteq \partial^-_k U^{(i)}$.

Conversely, it follows from Lemma 131 that for all $i$, if $\partial^-_k U^{(i)}$ is a regular molecule and $\partial^-_k x^{(i)}$ is its submolecule, then $U^{(i)}$ is a regular molecule, hence $\partial^+_k U^{(i)}$ is a regular molecule. Moreover, since $(x^{(i)})_{i=1}^m$ is a $k$-ordering, it is straightforward to prove that $U^{(i)} \cap U^{(i+1)} = \partial^+_k U^{(i)} = \partial^-_k U^{(i+1)}$ for all $i \in \{1,\ldots,m-1\}$. Since $\partial^- U^{(1)} = \partial^- U$ is a regular molecule, it follows by induction, assuming condition (b), that $U^{(i)}$ is a regular molecule for all $i \in \{1,\ldots,m\}$. This proves that $(U^{(i)})_{i=1}^m$ is a $k$-layering of $U$. ∎

***145.*** In the following, we use the following explicit construction of $\mathscr{G}/(\mathscr{G}|_W)$. Its set of vertices is $(V_\mathscr{G} \setminus W) + \{x_W\}$, and for all pair of vertices $x, y$,

- if $x, y \neq x_W$, there is an edge between $x$ and $y$ for each edge between $x$ and $y$ in $\mathscr{G}$,
- if $x = x_W$ and $y \neq x_W$, there is an edge from $x$ to $y$ for each pair of a vertex $z \in W$ and an edge from $x$ to $y$ in $\mathscr{G}$,
- if $x \neq x_W$ and $y = x_W$, there is an edge from $x$ to $y$ for each pair of a vertex $z \in W$ and an edge from $x$ to $z$ in $\mathscr{G}$,
- there are no edges from $x_W$ to $x_W$.

*Proof of Lemma 89.* We prove the contrapositive of the implication from (a) to (b). Suppose $\mathscr{G}/(\mathscr{G}|_W)$ has a cycle. If the cycle does not pass through $x_W$, then it lifts to a cycle in $\mathscr{G}$, contradicting the assumption that $\mathscr{G}$ is acyclic. It follows that the cycle contains a segment of the form $x_W \to x_1 \to \ldots \to x_m \to x_W$, where $m > 0$ and $x_i \neq x_W$ for all $i \in \{1,\ldots,m\}$. Then there exist $y, z \in W$ and a path $y \to x_1 \to \ldots \to x_m \to z$ in $\mathscr{G}$, so $\mathscr{G}|_W$ is not path-induced.

Next, suppose that $\mathscr{G}/(\mathscr{G}|_W)$ is acyclic. Then both the graphs $\mathscr{G}/(\mathscr{G}|_W)$ and $\mathscr{G}|_W$ are acyclic, so they admit topological sorts $(x^{(i)})_{i=1}^m$ and $(y^{(j)})_{j=1}^p$, respectively. For exactly one $q \in \{1,\ldots,m\}$, $x^{(i)} = x_W$. We claim that

$$((x^{(i)})_{i=1}^{q-1}, (y^{(j)})_{j=1}^p, (x^{(i)})_{i=q+1}^m)$$

is a topological sort of $\mathscr{G}$. Indeed, for all edges from $x$ to $x'$ in $\mathscr{G}$,

- if $x, x' \notin W$, then $x = x^{(i)}$, $x' = x^{(i')}$ for some $i, i' \in \{1,\ldots,m\} \setminus \{q\}$, and there is an edge from $x$ to $x'$ in $\mathscr{G}/(\mathscr{G}|_W)$, so $i < i'$;
- if $x, x' \in W$, then $x = y^{(j)}$, $x' = y^{(j')}$ for some $j, j' \in \{1,\ldots,p\}$, and there is an edge from $x$ to $x'$ in $\mathscr{G}|_W$, so $j < j'$;
- if $x \in W$, $x' \notin W$, then $x = y^{(j)}$, $x' = x^{(i)}$ for some $i \in \{1,\ldots,m\} \setminus \{q\}$, $j \in \{1,\ldots,p\}$, and there is an edge from $x_W$ to $x'$ in $\mathscr{G}/(\mathscr{G}|_W)$, so $q < i$;
- if $x \notin W$, $x' \in W$, then $x = x^{(i)}$, $x' = y^{(j)}$ for some $i \in \{1,\ldots,m\} \setminus \{q\}$, $j \in \{1,\ldots,p\}$, and there is an edge from $x$ to $x_W$ in $\mathscr{G}/(\mathscr{G}|_W)$, so $i < q$.

This proves the implication from (b) to (c). Moreover, it defines an injection from pairs of a topological sort of $\mathscr{G}|_W$ and a topological sort of $\mathscr{G}/(\mathscr{G}|_W)$ to topological sorts of $\mathscr{G}$ in which the vertices of $W$ are consecutive. This will prove to be a bijection as soon as we have proven the converse implication.

Finally, we prove the contrapositive of the implication from (c) to (a). Suppose $\mathscr{G}|_W$ is not path-induced, that is, there is a path $x \to x_1 \to \ldots \to x_m \to y$ in $\mathscr{G}$ such that $m > 0$, $x, y \in W$, and $x_i \notin W$ for all $i \in \{1,\ldots,m\}$. It follows that the $x_i$ must come between $x$ and $y$ in every topological sort of $\mathscr{G}$, so the vertices of $W$ can never be consecutive. ∎

*Proof of Lemma 90.* By Lemma 65 combined with Proposition 66, $\mathscr{F}_{n-1}V$ is a connected induced subgraph of $\mathscr{F}_{n-1}U$, so its contraction is well-defined. Now, the vertices of the graph $\mathscr{F}_{n-1}U[\langle V\rangle / \imath(V)]$ are either

- $x \in U_n \setminus V_n$, or
- $x_V$ such that the image of $\langle V\rangle$ in $U[\langle V\rangle / \imath(V)]$ is $\operatorname{cl}\{x_V\}$.

Let $x, y$ be two vertices of $\mathscr{F}_{n-1}U[\langle V\rangle / \imath(V)]$.

- If $x, y \in U_n \setminus V_n$, then $\Delta^+ x \cap \Delta^- y$ is the same in $U[\langle V\rangle / \imath(V)]$ as in $U$, so there is an edge from $x$ to $y$ in $\mathscr{F}_{n-1}U$ if and only if there is an edge in $\mathscr{F}_{n-1}U[\langle V\rangle / \imath(V)]$.
- If $x = x_V$ then $\Delta^+ x_V \cap \Delta^- y$ is in bijection with $\Delta^+ V \cap \Delta^- y$ in $U$. For all $z \in \Delta^+ V$, since $V$ is pure and $n$-dimensional, there exists $w \in \nabla^+ z$. If $\Delta^+ x_V \cap \Delta^- y$ is non-empty, it follows that $\Delta^+ z \cap \Delta^- y$ is non-empty in $U$ for some $z \in \imath(V)_n$. Thus there exist $z \in \imath(V)_n$ and an edge from $z$ to $y$ in $\mathscr{F}_{n-1}U$.
- Dually, if $y = x_V$, there is an edge from $x$ to $y$ in $\mathscr{F}_{n-1}U[\langle V\rangle / \imath(V)]$ if and only if there exist $z \in \imath(V)_n$ and an edge from $x$ to $z$ in $\mathscr{F}_{n-1}U$.
- Finally, $\Delta^+ V \cap \Delta^- V = \varnothing$ because $V$ is pure, so $\Delta^+ x_V \cap \Delta^- x_V$ and there is no edge from $x_V$ to $x_V$.

It is then straightforward to establish an isomorphism with the explicit description of $\mathscr{F}_{n-1}U / \mathscr{F}_{n-1}V$. ∎

*Proof of Lemma 93.* Identify $V$ with its isomorphic image through $\imath$, and suppose that $\imath$ is a submolecule inclusion. Then $\tilde{U} \coloneqq U[\langle V\rangle / V]$ is a regular molecule by Proposition 53, and admits an $(n-1)$-layering $(\tilde{U}^{(i)})_{i=1}^{m-p+1}$ by Theorem 77. Let $\operatorname{cl}\{x\}$ be the image of $\langle V\rangle$ in $\tilde{U}$; then $x \in \tilde{U}^{(q)}$ for exactly

one $q \in \{1, \dots, m-p+1\}$. Then $W \coloneqq \tilde{U}^{(q)}[V/\mathrm{cl}\,\{x\}]$ is defined, and by Lemma 142 combined with Lemma 133, $U$ is isomorphic to

$$\tilde{U}^{(1)} \#_{n-1} \cdots \#_{n-1} \tilde{U}^{(q-1)} \#_{n-1} \\ \#_{n-1} W \#_{n-1} \tilde{U}^{(q+1)} \#_{n-1} \cdots \#_{n-1} \tilde{U}^{(m-p+1)}.$$

By Lemma 144, $\partial^- x \sqsubseteq \partial^- \tilde{U}^{(q)}$, so by Lemma 141 $\partial^- V \sqsubseteq \partial^- W$. We can apply the criterion of Proposition 85 to deduce that $(y^{(i)})_{i=1}^p$ is an $(n-1)$-ordering of $W$ induced by an $(n-1)$-layering $(W^{(i)})_{i=1}^p$. Letting

$$(U^{(i)})_{i=1}^m \coloneqq ((\tilde{U}^{(i)})_{i=1}^{q-1}, (W^{(i)})_{i=1}^p, (\tilde{U}^{(i)})_{i=q+1}^{m-p+1}),$$

produces an $(n-1)$-layering, hence an $(n-1)$-ordering $(x^{(i)})_{i=1}^m$ of $U$, with the property that $(x^{(i)})_{i=q}^{p+q-1} = (y^{(i)})_{i=1}^p$.

Conversely, let $(U^{(i)})_{i=1}^m$ be an $(n-1)$-layering of $U$ satisfying the properties in the statement, and let $W \sqsubseteq U$ be the image of $U^{(q)} \#_{n-1} \cdots \#_{n-1} U^{(p+q-1)}$ in $U$. Then $W_n = V_n$, so

$$W = V \cup \partial^- W.$$

Because $\partial^- V \sqsubseteq \partial^- U^{(q)}$ which is equal to $\partial^- W$, by Lemma 131 $V \sqsubseteq W \sqsubseteq U$. ∎

*Full proof of Theorem 94.* Identify $V$ with its isomorphic image through $\imath$, and suppose that $\imath$ is a submolecule inclusion. Then $\tilde{U} \coloneqq U[\langle V \rangle / V]$ is a regular molecule by Proposition 53, so it admits an $(n-1)$-layering $(\tilde{U}^{(i)})_{i=1}^{m-p+1}$, which induces an $(n-1)$-ordering. By Lemma 90, this $(n-1)$-ordering can be identified with a topological sort $((x^{(i)})_{i=1}^{q-1}, x_V, (x^{(i)})_{i=q+1}^{m-p+1})$ of $\mathscr{F}_{n-1}U / \mathscr{F}_{n-1}V$. By Lemma 144, we have $\partial^- x_V \sqsubseteq \partial^- \tilde{U}^{(q)}$ and $\partial^- x^{(i)} \sqsubseteq \partial^- \tilde{U}^{(i)}$ for $i \neq q$. By Lemma 142 combined with Lemma 133, letting $W \coloneqq \tilde{U}^{(q)}[V/\mathrm{cl}\,\{x_V\}]$, $U$ is isomorphic to

$$\tilde{U}^{(1)} \#_{n-1} \cdots \#_{n-1} \tilde{U}^{(q-1)} \#_{n-1} \\ \#_{n-1} W \#_{n-1} \tilde{U}^{(q+1)} \#_{n-1} \cdots \#_{n-1} \tilde{U}^{(m-p+1)}.$$

and $W$ is isomorphic to $U^{(q)}$, while $\tilde{U}^{(i)}$ is isomorphic to $U^{(i)}$ for all $i \neq q$. We conclude by Lemma 141.

The converse implication has already been fully proved. ∎

***Lemma 146.*** *Let $U$ be a regular molecule, $\ell \geq -1$. If $U$ has an $\ell$-layering, then for all $k > \ell$ the function $\mathsf{m}_{k,U} \colon \mathscr{L}\mathrm{a}\mathrm{y}_k U \hookrightarrow \mathscr{O}\mathrm{r}\mathrm{d}_k U$ is a bijection.*

*Proof.* Let $(U^{(i)})_{i=1}^m$ be an $\ell$-layering of $U$, and let $(x^{(i)})_{i=1}^m$ be its image through $\mathsf{m}_{\ell,U}$. For $k > \ell$, let $(y^{(i)})_{i=1}^p$ be a $k$-ordering of $U$. Then there exists a unique injection $\mathsf{j} \colon \{1, \dots, p\} \hookrightarrow \{1, \dots, m\}$ such that $y^{(i)} = x^{(\mathsf{j}(i))}$ for all $i \in \{1, \dots, p\}$. Let

$$V^{(i)} \coloneqq \partial_k^{\alpha(i,1)} U^{(1)} \#_\ell \cdots \#_\ell U^{(\mathsf{j}(i))} \#_\ell \cdots \#_\ell \partial_k^{\alpha(i,m)} U^{(m)},$$

$$\alpha(i,j) \coloneqq \begin{cases} + & \text{if } j = \mathsf{j}(i') \text{ for some } i' < i, \\ - & \text{otherwise.} \end{cases}$$

Applying Proposition 38 and Proposition 39 repeatedly, we find that $(V^{(i)})_{i=1}^p$ is a $k$-layering of $U$ and $(y^{(i)})_{i=1}^p$ is its image through $\mathsf{m}_{k,U}$. This proves that $\mathsf{m}_{k,U}$ is surjective, and we conclude by Proposition 83. ∎

***Lemma 147.*** *Let $U$ be a regular molecule. Suppose that for all submolecules $V \sqsubseteq U$, if $r \coloneqq \mathrm{frdim}\, V$, then $V$ admits an $r$-layering. Then for all $k \geq \mathrm{frdim}\, U$ the function $\mathsf{m}_{k,U} \colon \mathscr{L}\mathrm{a}\mathrm{y}_k U \hookrightarrow \mathscr{O}\mathrm{r}\mathrm{d}_k U$ is a bijection.*

*Proof.* Let $r \coloneqq \mathrm{frdim}\, U$. By assumption, there exists an $r$-layering of $U$, so by Lemma 146 it suffices to show that $\mathsf{m}_{r,U}$ is a bijection.

Given two $r$-orderings $(x^{(i)})_{i=1}^m$ and $(y^{(i)})_{i=1}^m$, there exists a unique permutation $\sigma$ such that $x^{(i)} = y^{(\sigma(i))}$ for all $i \in \{1, \dots, m\}$. Let $d((x^{(i)})_{i=1}^m, (y^{(i)})_{i=1}^m)$ be the number of pairs $(j, j')$ such that $j < j'$ but $\sigma(j') < \sigma(j)$. Under the assumption that $(x^{(i)})_{i=1}^m$ is in the image of $\mathsf{m}_{r,U}$, we will prove that $(y^{(i)})_{i=1}^m$ is also in the image of $\mathsf{m}_{r,U}$ by induction on $d((x^{(i)})_{i=1}^m, (y^{(i)})_{i=1}^m)$. Since the image of $\mathsf{m}_{r,U}$ is not empty, this will suffice to prove that $\mathsf{m}_{r,U}$ is surjective, hence bijective by Proposition 83.

If $d((x^{(i)})_{i=1}^m, (y^{(i)})_{i=1}^m) = 0$, then $x^{(i)} = y^{(i)}$ for all $i \in \{1, \dots, m\}$, and there is nothing left to prove.

Suppose $d((x^{(i)})_{i=1}^m, (y^{(i)})_{i=1}^m) > 0$. Then there exists $j < m$ such that $\sigma(j+1) < \sigma(j)$. Suppose $(x^{(i)})_{i=1}^m$ is the image of the $r$-layering $(U^{(i)})_{i=1}^m$. Let $V \sqsubseteq U$ be the image of $U^{(j)} \#_r U^{(j+1)}$ in $U$, and let

$$z_1 \coloneqq x^{(j)} = y^{(\sigma(j))}, \quad z_2 \coloneqq x^{(j+1)} = y^{(\sigma(j+1))}.$$

Because $z_1$ comes before $z_2$ in one $r$-ordering, but after in another, there can be no edge between them in $\mathscr{M}_r U$, so

$$\dim(\mathrm{cl}\,\{z_1\} \cap \mathrm{cl}\,\{z_2\}) < r.$$

Since $z_1, z_2$ are the only maximal elements of dimension $> r$ in $V$, we deduce that $\ell \coloneqq \mathrm{frdim}\, V < r$. By assumption, there exists an $\ell$-layering of $V$. In particular, there exist regular molecules $V^{(1)}, V^{(2)}$ such that

1) $z_i$ is in the image of $V^{(i)}$ for all $i \in \{1, 2\}$, and
2) $V$ is isomorphic to $V^{(1)} \#_\ell V^{(2)}$ or to $V^{(2)} \#_\ell V^{(1)}$.

Without loss of generality suppose that $V$ is isomorphic to $V^{(1)} \#_\ell V^{(2)}$. By Proposition 38 and Proposition 39, letting

$$\tilde{U}^{(j)} \coloneqq \partial_r^- V^{(1)} \#_\ell V^{(2)},$$
$$\tilde{U}^{(j+1)} \coloneqq V^{(1)} \#_\ell \partial_r^+ V^{(2)},$$

we have that $V$ is isomorphic to $\tilde{U}^{(j)} \#_r \tilde{U}^{(j+1)}$. Letting $\tilde{U}^{(i)} \coloneqq U^{(i)}$ for $i \notin \{j, j+1\}$, we have that $(\tilde{U}^{(i)})_{i=1}^m$ is an $r$-layering of $U$, and

$$\mathsf{m}_{r,U} \colon (\tilde{U}^{(i)})_{i=1}^m \mapsto (\tilde{x}^{(i)})_{i=1}^m = \\ = (x^{(1)}, \dots, x^{(j+1)}, x^{(j)}, \dots, x^{(m)}).$$

Then $d((\tilde{x}^{(i)})_{i=1}^m, (y^{(i)})_{i=1}^m) < d((x^{(i)})_{i=1}^m, (y^{(i)})_{i=1}^m)$ and $(\tilde{x}^{(i)})_{i=1}^m$ is in the image of $\mathsf{m}_{r,U}$. We conclude by the inductive hypothesis. ∎

***Lemma 148.*** *Let $U$ be a regular molecule, $r \coloneqq \mathrm{frdim}\, U$. If $U$ is frame-acyclic, then $U$ admits an $r$-layering.*

*Proof.* Since submolecules of frame-acyclic regular molecules are frame-acyclic, we can proceed by induction on submolecules. For all $x \in U_0$, we have $\operatorname{frdim}\{x\} = -1$, and $\{x\}$ admits a trivial $(-1)$-layering, proving the base case.

We construct an ordered tree of submolecules $U^{(j_1,\dots,j_p)}$ of $U$, as follows:

- the root is $U^{()} \coloneqq U$;
- if $\operatorname{lydim} U^{(j_1,\dots,j_p)} \leq r$, then we let $\operatorname{lydim} U^{(j_1,\dots,j_p)}$ be a leaf;
- if $k \coloneqq \operatorname{lydim} U^{(j_1,\dots,j_p)} > r$, then we pick a $k$-layering $(V^{(i)})_{i=1}^q$ of $U^{(j_1,\dots,j_p)}$, which is possible by Theorem 77, and for each $i \in \{1,\dots,q\}$, we let the image of $V^{(i)}$ be a child $U^{(j_1,\dots,j_p,i)}$ of $U^{(j_1,\dots,j_p)}$.

By Lemma 139, the layering dimension of the children of a node is strictly smaller than that of the node, so the procedure terminates.

Fix an $r$-ordering $(x^{(i)})_{i=1}^m$ of $U$; this is possible because $\mathscr{M}_r U$ is acyclic. Let $V \coloneqq U^{(j_1,\dots,j_p)}$ be a node of the tree. We have

$$\begin{aligned}\bigcup_{j>r} (\mathscr{M}\!ax\, V)_j &= \sum_{i=1}^m \bigcup_{j>r} \left((\mathscr{M}\!ax\, V)_j \cap \operatorname{cl}\left\{x^{(i)}\right\}\right) \\ &=: \sum_{i=1}^m M^{(i)};\end{aligned}$$

the $M^{(i)}$ form a partition because $\operatorname{frdim} U = r$, so every element of dimension $> r$ is in the closure of $x^{(i)}$ for a unique $i \in \{1,\dots,m\}$. We claim that $V$ is isomorphic to

$$V^{(1)} \#_r \dots \#_r V^{(m)}$$

for some regular molecules $(V^{(i)})_{i=1}^m$ with the following property: for each $i \in \{1,\dots,m\}$, identifying $V^{(i)}$ with its image in $V$, we have

$$\bigcup_{j>r} (\mathscr{M}\!ax\, V^{(i)})_j = M^{(i)}.$$

We will prove this by backward induction on the tree $U^{(j_1,\dots,j_p)}$.

Suppose $V$ is a leaf, so $\operatorname{lydim} V \leq r$. Then $V$ admits an $r$-layering. For each $i \in \{1,\dots,m\}$, fix a topological sort $(y^{(i,j)})_{j=1}^{p_i}$ of the induced subgraph $\mathscr{M}_r V|_{M^{(i)}}$. We claim that $((y^{(i,j)})_{j=1}^{p_i})_{i=1}^m$ is an $r$-ordering of $V$.

Suppose there is an edge from $x$ to $x'$ in $\mathscr{M}_r V$. Then $x \in M^{(i)}$, $x' \in M^{(i')}$ for a unique pair $i, i' \in \{1,\dots,m\}$. If $i = i'$, then $x = y^{(i,j)}$ and $x' = y^{(i,j')}$ for some $j, j' \in \{1,\dots,p_i\}$, and $j < j'$ because $(y^{(i,j)})_{j=1}^{p_i}$ is a topological sort of $\mathscr{M}_r V|_{M^{(i)}}$. If $i \neq i'$, then there exists

$$z \in \Delta_r^+ x \cap \Delta_r^- x' \subseteq \operatorname{cl}\left\{x^{(i)}\right\} \cap \operatorname{cl}\left\{x^{(i')}\right\}.$$

Since $\partial_r^\alpha x^{(i)}$ and $\partial_r^\alpha x^{(i')}$ is pure and $r$-dimensional for all $\alpha \in \{+,-\}$, by [8, Proposition 6.4]

$$z \in (\Delta_r^+ x^{(i)} \cap \Delta_r^- x^{(i')}) \cup (\Delta_r^- x^{(i)} \cap \Delta_r^+ x^{(i')}),$$

and $\Delta_r^- x^{(i)} \cap \operatorname{cl}\{x\} \subseteq \Delta_r^- x$ which is disjoint from $\Delta_r^+ x$, so $z \in \Delta_r^+ x^{(i)} \cap \Delta_r^- x^{(i')}$. It follows that there is an edge from $x^{(i)}$ to $x^{(i')}$ in $\mathscr{M}_r U$, so $i < i'$ because $(x^{(i)})_{i=1}^m$ is a topological sort of $\mathscr{M}_r U$. This proves that $((y^{(i,j)})_{j=1}^{p_i})_{i=1}^m$ is an $r$-ordering of $V$.

Let $W \sqsubseteq V$, $\ell \coloneqq \operatorname{frdim} W$. If $V \neq U$ or $W \neq U$, then $W$ admits an $\ell$-layering by the inductive hypothesis on proper submolecules of $U$. If $W = V = U$ then $\ell = r$ and $W$ admits an $\ell$-layering by Theorem 77. In either case, $V$ satisfies the conditions of Lemma 147, and since $r \geq \operatorname{lydim} V \geq \operatorname{frdim} V$, every $r$-ordering of $V$ comes from an $r$-layering of $V$.

It follows that $((y^{(i,j)})_{j=1}^{p_i})_{i=1}^m$ comes from an $r$-layering $((W^{(i,j)})_{j=1}^{p_i})_{i=1}^m$, and we can define

$$V^{(i)} \coloneqq W^{(i,1)} \#_r \dots \#_r W^{(i,p_i)}$$

for each $i \in \{1,\dots,m\}$, satisfying the desired condition.

Now, suppose that $V$ is not a leaf, so $k \coloneqq \operatorname{lydim} V > r$, and $V$ has children $(W^{(j)})_{j=1}^q$ forming a $k$-layering of $V$. By the inductive hypothesis, each of the $W^{(j)}$ has a decomposition

$$W^{(j,1)} \#_r \dots \#_r W^{(j,m)}$$

such that the maximal elements of dimension $> r$ in the image of $W^{(j,i)}$ are contained in $\operatorname{cl}\{x^{(i)}\}$. Then, for each $i \in \{1,\dots,m\}$ and $j, j' \in \{1,\dots,q\}$,

$$W^{(j,i)} \cap W^{(j')} \subseteq W^{(j',i)},$$

so $V^{(i)} \coloneqq W^{(1,i)} \#_k \dots \#_k W^{(q,i)}$ is defined. Using Proposition 39 repeatedly, we conclude that $V$ is isomorphic to $V^{(1)} \#_r \dots \#_r V^{(m)}$.

This concludes the induction on $U^{(j_1,\dots,j_p)}$. In particular, for the root $U^{()} = U$, the decomposition $U^{(1)} \#_r \dots \#_r U^{(m)}$ satisfies

$$\bigcup_{j>r} (\mathscr{M}\!ax\, U^{(i)})_j = \left\{x^{(i)}\right\},$$

that is, $(U^{(i)})_{i=1}^m$ is an $r$-layering of $U$. ∎

*Proof of Theorem 102.* The implication from (a) to (b) is a consequence of Lemma 148 together Lemma 73. The implication from (b) to (c) is Lemma 147. Finally, the implication from (c) to (a) follows from Proposition 80. ∎

*Proof of Lemma 108.* Suppose $U$ is acyclic. Let $k \in \mathbb{N}$ and suppose there is a cycle $x_0 \to x_1 \to \dots \to x_m = x_0$ in $\mathscr{F}_k U$. By definition, for all $i \in \{1,\dots,m\}$ there exists $y_i \in \Delta_k^+ x_{i-1} \cap \Delta_k^- x_i$. By [7, Lemma 1.37], there exist paths from $x_{i-1}$ to $y_i$ and then from $y_i$ to $x_i$ in $\mathscr{H}U$. Concatenating all these paths, we obtain a cycle in $\mathscr{H}U$.

Suppose $U$ is dimension-wise acyclic, and let $V \sqsubseteq U$ be a submolecule inclusion, $r \coloneqq \operatorname{frdim} V$. Then $\mathscr{F}_r U$ is acyclic, hence so are its induced subgraphs $\mathscr{F}_r V$ and $\mathscr{M}_r V$. ∎

### D. Proofs for Section IV

*Proof of Lemma 115.* We will show that $U$ is isomorphic to

$$\underbrace{\mathsf{arrow} \#_0 \dots \#_0 \mathsf{arrow}}_{m \text{ times}}.$$

By Lemma 136, either $\operatorname{lydim} U = -1$ or $\operatorname{lydim} U = 0$. In the first case, $U$ is an atom by Lemma 137. Because the point is

the only 0-dimensional molecule up to isomorphism, the arrow is the only 1-dimensional atom, so $U$ is isomorphic to arrow. In the second case, $U$ admits a 0-layering $(U^{(i)})_{i=1}^m$ by Theorem 77, and by Lemma 139, for each $i \in \{1,\dots,m\}$, necessarily $\operatorname{lydim} U^{(i)} = -1$. By the first part, $U^{(i)}$ is isomorphic to arrow. ∎

*Proof of Proposition 117.* This is [17, Proposition 5]. ∎

***149*** (Horizontal and vertical order)**.** Let $U$ be a regular molecule, $\dim U \leq 2$. The *horizontal order* $\preceq_0$ and the *vertical order* $\preceq_2$ on the set $U_1$ of 1-dimensional elements of $U$ are defined by

- $x \preceq_0 y$ if and only if there is a path from $x$ to $y$ in $\mathring{\mathscr{H}}U$ only passing through elements of dimension 0 and 1,
- $x \preceq_2 y$ if and only if there is a path from $x$ to $y$ in $\mathring{\mathscr{H}}U$ only passing through elements of dimension 1 and 2.

***Lemma 150.*** *Let $U$ be a regular molecule, $\dim U \leq 2$. Then*

1) *the union of $\preceq_0$ and $\preceq_2$ is a linear order on $U_1$,*
2) *the intersection of $\preceq_0$ and $\preceq_2$ is the identity relation on $U_1$.*

*Proof.* If $\dim U < 2$, then $\preceq_2$ is trivially the identity relation, and $\preceq_0$ is a linear order by 115. If $U$ is a 2-dimensional atom, then $U_1 = \Delta^- U + \Delta^+ U$, and that $\preceq_0$ is a linear order on $\Delta^\alpha U$ for each $\alpha \in \{+,-\}$ separately, so we have $x \preceq_2 y$ for all $x \in \Delta^- U$ and $y \in \Delta^+ U$, and no other relations exist.

Otherwise, we proceed exactly as in the proof of Proposition 117, defining a 1-ordering $(x^{(i)})_{i=1}^m$ and a sequence $(V^{(i)})_{i=1}^m$ of increasing submolecules of $U$. We let $\preceq_0^{(i)}$ and $\preceq_2^{(i)}$ be the orders determined by paths in $\mathring{\mathscr{H}}V^{(i)}$, which are increasing in $i$, and proceed by induction. Since $\dim V^{(0)} = 1$, we have already proved the base case.

Let $i > 0$, assume that the statement holds of the orders $\preceq_0^{(i-1)}$ and $\preceq_2^{(i-1)}$ on $(V^{(i-1)})_1$. Let $x, y \in (V^{(i)})_1$; we will show that $x$ and $y$ are comparable via $\preceq_0^{(i)}$ or $\preceq_2^{(i)}$. If $x, y \in V^{(i-1)}$ or $x, y \in \operatorname{cl}\left\{x^{(i)}\right\}$, we can apply the inductive hypothesis or the atom case, so it suffices to consider the case

$$x \in V^{(i-1)} \setminus \Delta^- x^{(i)}, \quad y \in \Delta^+ x^{(i)}.$$

Let $(z^{(j)})_{j=1}^p$ be the unique linear 0-ordering on $\partial^- x^{(i)}$, so $z^{(j)} \preceq_0 z^{(j')}$ if $j \leq j'$. For all $j \in \{1,\dots,p\}$, we have $x \neq z^{(j)}$, and by the inductive hypothesis $x$ and $z^{(j)}$ are comparable via $\preceq_0^{(i-1)}$ or they are comparable via $\preceq_2^{(i-1)}$.

- Suppose there exists $j$ such that $x$ and $z^{(j)}$ are comparable via $\preceq_2^{(i-1)}$. Then necessarily $x \preceq_2^{(i-1)} z^{(j)}$, because $\Delta^- x^{(i)} \subseteq \Delta^+ V^{(i-1)}$. Since $z^{(j)} \preceq_2^{(i)} y$, we have $x \preceq_2^{(i)} y$.
- Otherwise, $x$ and $z^{(j)}$ are comparable via $\preceq_0^{(i-1)}$ for all $j \in \{1,\dots,p\}$. Suppose that $x \preceq_0^{(i-1)} z^{(1)}$, in which case $x \preceq_0^{(i-1)} z^{(j)}$ for all $j \in \{1,\dots,p\}$. Then the path from $x$ to $z^{(1)}$ through elements of dimension 0 and 1 must enter $z^{(1)}$ from $\partial^- z^{(1)} = \partial_0^- x^{(i)}$. Since there is a path in $\mathring{\mathscr{H}}V_i$ from $\partial_0^- x^{(i)}$ to $y$, we have $x \preceq_0^{(i)} y$.
- Otherwise, there is a greatest $j$ such that $z^{(j)} \preceq_0^{(i-1)} x$. If $j < p$, then $z^{(j)} \preceq_0^{(i-1)} x \preceq_0^{(i-1)} z^{(j+1)}$. Because all three are distinct, letting $\partial^+ z^{(j)} = \partial^- z^{(j+1)} = \{w\}$, there is a non-trivial cycle in $\mathring{\mathscr{H}}V^{(i-1)}$ from $w$ to $x$ and back to $w$, a contradiction to Proposition 117. It follows that $z^{(p)} \preceq_0^{(i-1)} x$, and the path between the two must leave $z^{(p)}$ through $\partial_0^+ x^{(i)}$, so $y \preceq_0^{(i)} x$.

This proves that the union of $\preceq_0^{(i)}$ and $\preceq_2^{(i)}$ is a linear order on $(V^{(i)})_1$.

Suppose that $x \preceq_0^{(i)} y$ and $x \preceq_2^{(i)} y$; we will prove that $x = y$. If $x \in \operatorname{cl}\left\{x^{(i)}\right\}$, then $x \preceq_2^{(i)} y$ implies that $y \in \operatorname{cl}\left\{x^{(i)}\right\}$, and any path from $x$ to $y$ in $\mathring{\mathscr{H}}V^{(i)}$ is entirely contained in $\operatorname{cl}\left\{x^{(i)}\right\}$, so we can apply the atom case. Suppose that $x \notin \operatorname{cl}\left\{x^{(i)}\right\}$.

If $y \in \Delta^+ x^{(i)}$, then any path from $x$ to $y$ through elements of dimension 1 and 2 consists of a path contained in $V^{(i-1)}$ to some $z \in \Delta^- x^{(i)}$ followed by the path $z \to x^{(i)} \to y$; while any path through elements of dimension 0 and 1 consists of a path contained in $V^{(i-1)}$ to $\partial_0^- x^{(i)}$ followed by a path contained in $\partial^+ x^{(i)}$. Since there is a path from $\partial_0^- x^{(i)}$ to any $z \in \Delta^- x^{(i)}$ through elements of dimension 0 and 1 in $V^{(i-1)}$, we have that $x \preceq_0^{(i-1)} z$ and $x \preceq_2^{(i-1)} z$ for some $z \in \Delta^- x^{(i)}$. By the inductive hypothesis, $x = z$, a contradiction since $z \in \operatorname{cl}\left\{x^{(i)}\right\}$.

It follows that $y \in V^{(i-1)}$. Then any path from $x$ to $y$ through elements of dimension 1 and 2 is entirely contained in $V^{(i-1)}$, so $x \preceq_2^{(i-1)} y$; while a path through elements of dimension 0 and 1 is either contained in $V^{(i-1)}$, or it enters $\partial^+ x^{(i)}$ through $\partial_0^- x^{(i)}$, traverses it in its entirety, and leaves from $\partial_0^+ x^{(i)}$. Such a path segment can be replaced with the one that traverses $\partial^- x^{(i)}$ in its entirety, so in either case $x \preceq_0^{(i-1)} y$, and we conclude that $x = y$ by the inductive hypothesis. This concludes the proof of the statement for $V^{(i)}$. Since $V^{(m)} = U$, we conclude. ∎

*Proof of Lemma 118.* Both $U$ and $\imath(V)$ are regular molecules of dimension $\leq 2$. Let $\preceq_0$, $\preceq_2$ be the horizontal and vertical order on $U$, and $\preceq_0^V$, $\preceq_2^V$ those on $\imath(V)$, which are subsets of those on $U$.

Suppose by way of contradiction that $\mathscr{F}_1 V$ is not path-induced. Then there exists a path $x_0 \to \dots \to x_m$ in $\mathscr{F}_1 U$ such that $m > 1$, $x_0, x_m \in \imath(V)$, and $x_i \notin \imath(V)$ for all $i \in \{1,\dots,m-1\}$. By definition, there exist 1-dimensional elements $y_i \in \Delta^+ x_{i-1} \cap \Delta^- x_i$ for all $i \in \{1,\dots,m\}$. Then $y_1 \preceq_2 y_m$ and $y_1 \neq y_m$. By Lemma 60, $x_{i-1}$ and $x_i$ are the only cofaces of $y_i$ for all $i \in \{1,\dots,m\}$. Necessarily, then, $y_1 \in \Delta^+ \imath(V)$ and $y_m \in \Delta^- \imath(V)$, so it is not possible that $y_1 \preceq_2^V y_m$. Then by Lemma 150 applied to $\imath(V)$, one of $y_1 \preceq_0^V y_m$, $y_m \preceq_0^V y_1$, or $y_m \preceq_2^V y_1$ must hold. Combined with $y_1 \preceq_2 y_m$, each one of these implies $y_1 = y_m$ by Lemma 150 applied to $U$, a contradiction. ∎

*Proof of Theorem 121.* By [17, Theorem 1] $U$ is frame-acyclic. Since for all $V \sqsubseteq U$ and all rewritable $W \sqsubseteq V$, $V[\langle W \rangle / W]$ is still a regular molecule of dimension $\leq 3$, it is frame-acyclic. Hence $U$ is stably frame-acyclic. ∎

# Appendix C

# Paper III

A. Hadzihasanovic and D. Kessler. "Acyclicity Conditions on Pasting Diagrams". In: *Applied Categorical Structures* 32 (Oct. 2024). DOI: 10.1007/s10485-024-09784-x

# Acyclicity conditions on pasting diagrams

Amar Hadzihasanovic and Diana Kessler

*Tallinn University of Technology*

*Abstract.* We study various acyclicity conditions on higher-categorical pasting diagrams in the combinatorial framework of regular directed complexes. We present an apparently weakest acyclicity condition under which the $\omega$-category presented by a diagram shape is freely generated in the sense of polygraphs. We then consider stronger conditions under which this $\omega$-category is equivalent to one obtained from an augmented directed chain complex in the sense of Steiner, or consists only of subsets of cells in the diagram. Finally, we study the stability of these conditions under the operations of pasting, suspensions, Gray products, joins and duals.

*Current version:* 2nd September 2024

Contents

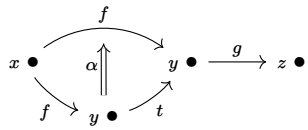


## Introduction

Pasting diagrams are a central tool for studying the composition of cells in higher-dimensional categories. The notion of 2-categorical pasting was introduced by Bénabou [Bén67]; in the 1980s and 1990s, a number of frameworks for $n$-categorical pasting emerged, with corresponding *pasting theorems* guaranteeing that a pasting diagram admits a suitably unique composite [Joh89, Pow91, Str91, Ste93], see [For22] for a recent survey.

A pasting diagram is, informally, a *composable* configuration of cells in an

$n$-category, such as the following:



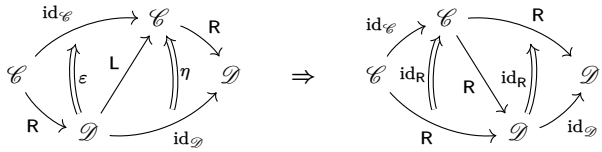

More in general, one considers "non-pasting" diagram shapes, that do not admit a composite, such as the following:

Various formalisms for diagrams have tried to encode, either *combinatorially* or *topologically* the information of such a diagram, in a way that reflects (and generalises in higher dimensions) the content of these pictures, and has a univocal interpretation, in the form of a higher category *presented* by the diagram shape, together with a functor out of it.

Most of these formalisms include, as part of the definition of a pasting diagram shape, some *acyclicity* conditions of varying strength, barring — at the very least — the existence of "direct loops" where a cell may appear more than one time in a composite. These conditions serve at least three purposes:

1. to restrict the class of admissible structures so that some undesirable examples are not part of it;
2. to guarantee that an $n$-category can be formed out of *subdiagrams*, or "composable subsets" of cells in the diagram;
3. to ensure that the presented $n$-category is *freely generated* in the sense of polygraphs or computads [Str76, Bur93, ABG$^+$23].

On the other hand, imposing such conditions comes with a cost. Firstly, they exclude commonly occuring shapes of pasting diagrams that appear in dimension 3: for example, a 3-cell of the form




appears as a "weakened" form of one of the *triangle equations* in the theory of pseudoadjunctions of 2-categories, but its shape is *not* acyclic, not even in a weak sense, due to the 1-cells in the interiors of the two sides forming a direct loop. If we look at non-pasting diagrams, simple counterexamples appear already in dimension 1, where looping diagrams of shape

are perfectly well-defined, yet any result proved with an assumption of acyclicity will not extend to them. Secondly, acyclicity properties are *global* properties that tend to be unstable, not preserved under common operations. For example, stronger acyclicity conditions are not preserved under direction-reversing duality operations, and weaker acyclicity conditions are not preserved under pasting and various forms of products.

In [Had20a, Had20b], the first-named author started exploring a framework for diagrams inspired by Steiner's approach in [Ste93], but based on the "local" property of *regularity*, which requires the input and output boundaries of cells in a diagram to be closed balls in a topological sense, and is remarkably stable under all sorts of constructions. This has resulted, recently, in the book-length exposition [Had24]. The structures encoding "diagram shapes" in this framework are called *regular directed complexes*.

The inductive definition of pasting diagram shapes, called *molecules* in this framework, ensures that "bad" examples are left out, but is general enough in the sense that it allows further shapes that one might wish to have. However, in general, the two other properties listed above — subdiagrams form an $n$-category, the $n$-category is a polygraph — are no longer guaranteed. In this article we use regular directed complexes as a backdrop for a more refined study of acyclicity conditions, and their role in achieving these properties.

For the first, we argue that, to a certain extent, it is a non-problem in that one can replace *subsets* with more general *morphisms* whose domains are molecules in order to obtain an $\omega$-category even without acyclicity. This is akin to the situation with directed graphs, where *linear subgraphs* only form a category if the graph is acyclic, but *paths* always form a category.

For the "free generation" property, we show that this is achieved by a weaker notion of acyclicity, called *frame-acyclicity*, which is shared by all regular directed complexes of dimension lower or equal then 3 — including the non-acyclic examples above. Frame-acyclicity for molecules is equivalent to *splitness* in the sense of [Ste93], and also has interesting algorithmic consequences as studied by the authors in [HK23]. However, frame-acyclicity is very technical, difficult to check, and its stability properties are unclear. so it is useful to consider other conditions which are easier to check in practice, but more restrictive.

The first such notion is *dimension-wise acyclicity* which is tied to Steiner's loop-freeness property in the theory of *augmented directed chain complexes* [Ste04], and allows us to make a precise connection between this and our framework. In particular, we prove the existence of an isomorphism between the $\omega$-category presented by a dimension-wise acyclic regular directed complex, and the $\omega$-category obtained by first passing to an augmented directed chain complex and then applying Steiner's functor $\nu$. A slightly stronger notion, that we call *strong dimension-wise acyclicity*, is what guarantees that every morphism from a molecule is injective, hence "subsets suffice". These two conditions have the nice property of being closed under direction-reversing

duals, but not under other operations such as pasting (for molecules), Gray products, or joins. For this reason, we finally consider an even stronger notion, *acyclicity*, corresponding to total loop-freeness in [Ste93], and which is stable under the latter operations.

*Structure of the article*

In Section 1, we introduce oriented graded posets, our basic data structure on which we define regular directed complexes and the inductive class of molecules, together with their category **ogPos**. We then present the construction of a strict $\omega$-category $Mol/P$ from an oriented graded poset $P$, whose cells are morphisms from a molecule to $P$, taken up to isomorphism in the slice category $\mathbf{ogPos}/P$. Section 2 is dedicated to the theory of layerings of molecules, which are ways of writing a molecule as a pasting decomposition in which each term or layer contains exactly one maximal element of dimension greater then the pasting dimension. In Section 3, we discuss the frame-acyclicity condition, giving full proofs of some results outlined in [HK23]. We end the section with the proof that, if an oriented graded poset $P$ has frame-acyclic molecules, then $Mol/P$ is a polygraph. In Section 4 we make the connection between our framework and Steiner's theory of augmented directed chain complexes. We prove that for a dimension-wise acyclic regular directed complex $P$, the $\omega$-category $Mol/P$ is isomorphic to the $\omega$-category obtained by applying Steiner's $\nu$ functor to the Steiner complex obtained from $P$. In Section 5 we study the stronger acyclicity conditions implying that the $\omega$-category $Mol/P$ consists only of subsets of $P$. We also show that the strongest condition translates to the associated Steiner complex being a "strong Steiner complex". Finally, in Section 6 we study the stability of the acyclicity conditions presented above under the operations of pasting, suspensions, Gray products, joins and duals.

*Note*

The content of this article was recently exposed, with more detail and all results reproved from scratch, in [Had24, Chapter 8 and Chapter 11], as part of a reference book written by the first-named author. The results, however, have been developed in cooperation by the two authors, and many of them have not appeared in print before. The purpose of this article is both to give a clearer picture of the original research developments — in contrast to the book, we do not reprove results when a proof with roughly the same content has appeared before, even when the definitions are slightly different — and to offer a concise, self-contained treatment of a topic which seems particularly subtle and somewhat misunderstood in the theory of higher-categorical diagrams.

*Acknowledgements*

The first-named author was supported by Estonian Research Council grant PSG764. We thank Guillaume Laplante-Anfossi for discussions which helped

shape the article and Clémence Chanavat for discussions about [Ste04].

## 1. Regular directed complexes and $\omega$-categories

In this section, we give an overview of our combinatorial framework for higher-categorical diagrams, and state without proof some of the foundational results. This is exposed in much more detail in the first chapters of [Had24]. The basic structure that we use to represent shapes of diagrams is called an *oriented graded poset*. Before introducing it, we first recall some notions about posets with order relation $\leq$.

***1.1*** (Faces and cofaces). Let $P$ be a poset. Given elements $x, y \in P$, we say that $y$ *covers* $x$ if $x < y$ and, for all $y' \in P$, if $x < y' \leq y$ then $y' = y$. For each $x \in P$, the sets of *faces* and *cofaces* of $x$ are, respectively,

$$\Delta x \coloneqq \{y \in P \mid x \text{ covers } y\} \quad \text{and} \quad \nabla x \coloneqq \{y \in P \mid y \text{ covers } x\}.$$

***1.2*** (Closed subsets). Let $P$ be a poset and $U \subseteq P$. The *closure of* $U$ is the subset $\operatorname{cl} U \coloneqq \{x \in P \mid \text{there exists } y \in U \text{ such that } x \leq y\}$. We say that $U$ is *closed* if $U = \operatorname{cl} U$.

***1.3*** (Graded poset). A poset $P$ is *graded* if, for all $x \in P$, all maximal chains in $\operatorname{cl}\{x\}$ have the same finite size $n$. In this case, we let $\dim x$, the *dimension* of $x$, be equal to $n-1$. If $P$ is a graded poset, the *dimension* of $P$ is

$$\dim P \coloneqq \begin{cases} \max\left(\{-1\} \cup \{\dim x \mid x \in P\}\right) & \text{if defined,} \\ \infty & \text{otherwise.} \end{cases}$$

For each $n \in \mathbb{N}$, we write $P_n \coloneqq \{x \in P \mid \dim x = n\}$.

*Remark 1.4* — In a graded poset, if $y \in \Delta x$, then $\dim y = \dim x - 1$.

***1.5*** (Oriented graded poset). An *oriented graded poset* is a graded poset $P$ together with, for all $x \in P$, a bipartition $\Delta x = \Delta^- x + \Delta^+ x$ of the set of faces of $x$ into a set $\Delta^- x$ of *input faces* and a set $\Delta^+ x$ of *output faces*.

*Remark 1.6* — By duality, this induces a bipartition $\nabla x = \nabla^+ x + \nabla^- x$ of the set of cofaces of each element $x$.

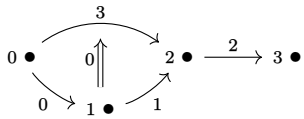
We will use $\alpha, \beta, \ldots$ for variables ranging over $\{+, -\}$. We let $-\alpha$ be $-$ if $\alpha = +$ and $+$ if $\alpha = -$.

***1.7*** (Oriented Hasse diagram). Let $P$ be an oriented graded poset. The *oriented Hasse diagram of* $P$ is the directed graph $\vec{\mathscr{H}} P$ whose

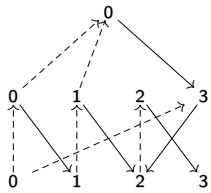
- set of vertices is the underlying set of $P$, and
- set of edges is $\{(x, y) \mid x \in \Delta^- y \text{ or } y \in \Delta^+ x\}$, where the source of $(x, y)$ is $x$ and the target is $y$.

The oriented Hasse diagram is the usual Hasse diagram of a poset, with edges representing input and output faces given opposite orientations (input from lower to higher dimension, and output from higher to lower dimension). An oriented graded poset is uniquely specified by its oriented Hasse diagram, together with the dim function, which graphically can be encoded by height, as in the following example.

*Example 1.8* — Consider the 2-dimensional pasting diagram shape




where we used progressive natural numbers for cells of each dimension. This is encoded by the oriented graded poset whose oriented Hasse diagram is




where we also (redundantly) represented the "input" edges as densely dashed lines for extra emphasis.

***1.9*** (Morphism of oriented graded posets). Let $P, Q$ be oriented graded posets. A *morphism* $f\colon P \to Q$ is a function of their underlying sets which, for all $x \in P$ and $\alpha \in \{+,-\}$, induces a bijection between $\Delta^\alpha x$ and $\Delta^\alpha f(x)$. An *inclusion* of oriented graded posets is an injective morphism.

We let **ogPos** denote the category whose objects are oriented graded posets and morphisms are morphisms of oriented graded posets.

We list some basic properties of morphisms of oriented graded posets. By *closed map*, we mean a map that sends closed subsets to closed subsets.

***Lemma 1.10*** — *Let $f\colon P \to Q$ be a morphism of oriented graded posets. Then $f$ is an order-preserving, closed, dimension-preserving map of the underlying graded posets.*

*Moreover, if $f$ is an inclusion, then $f$ is order-reflecting, and reflects input and output faces, that is, if $f(x) \in \Delta^\alpha f(y)$, then $x \in \Delta^\alpha y$.*

*In particular, $f$ is an isomorphism if and only if it is a surjective inclusion.*

*Remark 1.11* — Consequently, there is a forgetful functor $\mathsf{U}\colon \mathbf{ogPos} \to \mathbf{Pos}$, where **Pos** is the category of posets and order-preserving maps.

*Remark 1.12* — The inclusion of a closed subset $U \subseteq P$ of an oriented graded poset with the induced order and orientation is always an inclusion of oriented graded posets.

***Proposition 1.13*** — *The category* **ogPos** *has a strict initial object* $\varnothing$ *and pushouts of inclusions along inclusions, which are both preserved and reflected by* $\mathsf{U} \colon \mathbf{ogPos} \to \mathbf{Pos}$*. Moreover,*

1. *the pushout of an inclusion along an inclusion is an inclusion,*
2. *a pushout square of inclusions is also a pullback square.*

***1.14*** (Input and output $n$-boundaries). Let $U$ be a closed subset of an oriented graded poset, and let $\mathscr{M}ax\, U \subseteq U$ be its subset of maximal elements. For all $\alpha \in \{+,-\}$ and $n \in \mathbb{N}$, let $\Delta^\alpha_n U \coloneqq \{x \in U_n \mid \nabla^{-\alpha} x \cap U = \varnothing\}$. The *input* and *output* $n$*-boundary* of $U$ are, respectively, the closed subsets

$$\partial^-_n U \coloneqq \mathrm{cl}\,(\Delta^-_n U) \cup \bigcup_{k<n} \mathrm{cl}\,(\mathscr{M}ax\, U)_k, \quad \partial^+_n U \coloneqq \mathrm{cl}\,(\Delta^+_n U) \cup \bigcup_{k<n} \mathrm{cl}\,(\mathscr{M}ax\, U)_k.$$

We let $\partial_n U \coloneqq \partial^+_n U \cup \partial^-_n U$. For $n < 0$, we let $\Delta^\alpha_n U = \partial^\alpha_n U \coloneqq \varnothing$.

***Lemma 1.15*** — *Let* $V \subseteq U$ *be closed subsets of an oriented graded poset,* $n \in \mathbb{N}$*, and* $\alpha \in \{+,-\}$*. Then* $V \cap \Delta^\alpha_n U \subseteq \Delta^\alpha_n V$*.*

***Proposition 1.16*** — *Let* $\imath \colon P \hookrightarrow Q$ *be an inclusion of oriented graded posets and* $U \subseteq P$ *a closed subset. For all* $n \in \mathbb{N}$ *and* $\alpha \in \{+,-\}$*,* $\imath(\Delta^\alpha_n U) = \Delta^\alpha_n \imath(U)$ *and* $\imath((\mathscr{M}ax\, U)_n) = (\mathscr{M}ax\, \imath(U))_n$*. Consequently,* $\imath(\partial^\alpha_n U) = \partial^\alpha_n \imath(U)$*.*

Not all oriented graded posets represent shapes of diagrams. In what follows, we introduce the two constructions that we use to build shapes of diagrams, then give the inductive construction of molecules, the subclass of oriented graded posets that represents well-formed shapes of diagrams.

***1.17*** (Pasting construction). Let $U$, $V$ be oriented graded posets, $k \in \mathbb{N}$, and let $\varphi \colon \partial^+_k U \xhookrightarrow{\sim} \partial^-_k V$ be an isomorphism. The *pasting of* $U$ *and* $V$ *at the* $k$*-boundary along* $\varphi$ is the oriented graded poset $U \#^\varphi_k V$ obtained in **ogPos** as the pushout

$$\begin{array}{ccc} \partial^+_k U & \xhookrightarrow{\varphi} & \partial^-_k V \hookrightarrow V \\ \downarrow & & \downarrow{\scriptstyle \imath_V} \\ U & \xhookrightarrow{\imath_U} & U \#^\varphi_k V. \end{array}$$

***1.18*** (Globularity). Let $U$ be an oriented graded poset. We say that $U$ is *globular* if, for all $k, n \in \mathbb{N}$ and $\alpha, \beta \in \{+,-\}$, if $k < n$ then

$$\partial^\alpha_k(\partial^\beta_n U) = \partial^\alpha_k U.$$

Let $U$ be a globular oriented graded poset such that $n \coloneqq \dim U < \infty$. For all $\alpha \in \{+,-\}$, we write $\partial^\alpha U \coloneqq \partial^\alpha_{n-1} U$, and $\partial U \coloneqq \partial^+ U \cup \partial^- U$. We also write $\mathrm{int}\, U \coloneqq U \setminus \partial U$.

***1.19*** (Roundness). Let $U$ be an oriented graded poset. We say that $U$ is *round* if it is globular and, for all $n < \dim U$,

$$\partial_n^- U \cap \partial_n^+ U = \partial_{n-1} U.$$

***1.20*** (Rewrite construction). Let $U$, $V$ be round oriented graded posets of the same finite dimension $n$, and suppose $\varphi \colon \partial U \xhookrightarrow{\sim} \partial V$ is an isomorphism restricting to isomorphisms $\varphi^\alpha \colon \partial^\alpha U \xhookrightarrow{\sim} \partial^\alpha V$ for each $\alpha \in \{+, -\}$. Construct the pushout in **ogPos**

$$\begin{array}{ccc} \partial U & \xhookrightarrow{\varphi} & \partial V & \hookrightarrow & V \\ \downarrow & & & & \downarrow \\ U & & \hookrightarrow & & \partial(U \Rightarrow^\varphi V). \end{array}$$

The *rewrite of $U$ into $V$ along $\varphi$* is the oriented graded poset $U \Rightarrow^\varphi V$ obtained by adjoining a single $(n+1)$-dimensional element $\top$ to $\partial(U \Rightarrow^\varphi V)$, with

$$\Delta^- \top \coloneqq U_n, \qquad \Delta^+ \top \coloneqq V_n.$$

***1.21*** (Point). The *point* is the oriented graded poset 1 with a single element and trivial orientation.

We are now ready to give the definition of molecules.

***1.22*** (Molecules and atoms). The class of *molecules* is the inductive subclass of oriented graded posets closed under isomorphisms and generated by the following clauses.

1. (*Point*). The point is a molecule.
2. (*Paste*). Let $U$, $V$ be molecules, let $k < \min\{\dim U, \dim V\}$, and let $\varphi \colon \partial_k^+ U \xhookrightarrow{\sim} \partial_k^- V$ be an isomorphism. Then $U \#_k^\varphi V$ is a molecule.
3. (*Atom*). Let $U$, $V$ be round molecules of the same finite dimension and let $\varphi \colon \partial U \xhookrightarrow{\sim} \partial V$ be an isomorphism restricting to $\varphi^\alpha \colon \partial^\alpha U \xhookrightarrow{\sim} \partial^\alpha V$ for each $\alpha \in \{+, -\}$. Then $U \Rightarrow^\varphi V$ is a molecule.

An *atom* is a molecule with a greatest element. Equivalently, it is a molecule whose final generating clause is either (*Point*) or (*Atom*).

The following summarises some basic properties of molecules.

***Proposition 1.23*** — *Let $U$ be a molecule, $n \in \mathbb{N}$, $\alpha \in \{+, -\}$, $x \in U$. Then*

1. *$U$ is globular,*
2. *$\partial_n^\alpha U$ is a molecule,*
3. *$\mathrm{cl}\{x\}$ is an atom and is round.*

***Proposition 1.24*** — *Let $U$, $V$ be molecules, $k \in \mathbb{N}$. Then*

1. *if $U$ and $V$ are isomorphic, there exists a unique isomorphism $\varphi \colon U \xrightarrow{\sim} V$,*
2. *if $U \#_k^{\varphi} V$ or $U \Rightarrow^{\varphi} V$ are defined, they are defined for a unique $\varphi$.*

*Remark 1.25* — This allows us to write $U \#_k V$ and $U \Rightarrow V$, omitting the specific isomorphism, when $U$ and $V$ are molecules and the constructions are defined. It will also allow us to be relaxed about the distinction between isomorphism and equality of molecules.

***1.26*** (Regular directed complex). A *regular directed complex* is an oriented graded poset $P$ with the property that, for all $x \in P$, the closed subset $\mathrm{cl}\{x\}$ is an atom. We write $\mathbf{RDCpx}$ for the full subcategory of $\mathbf{ogPos}$ whose objects are regular directed complexes.

*Comment 1.27* — While superficially different, this is equivalent to the definition of regular directed complex given in [Had20b]. On the other hand, our definition of *molecule* here corresponds to a *molecule in a regular directed complex*, or *regular molecule*, and is more restrictive than the definition in [Ste93]. Note that by Proposition 1.23, every molecule is a regular directed complex.

***1.28*** (Positive least element). Let $P$ be an oriented graded poset, $\bot \in P$. We say that $\bot$ is a *positive least element of $P$* if $\bot$ is the least element of $P$ and $\nabla \bot = \nabla^+ \bot$. We let $\mathbf{ogPos}^+$ denote the full subcategory of $\mathbf{ogPos}$ on oriented graded posets with a positive least element.

Freely adjoining a positive least element and, respectively, deleting the least element exhibit an equivalence between $\mathbf{ogPos}$ and $\mathbf{ogPos}^+$.

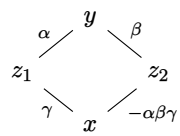
***Proposition 1.29*** — *There exists a pair of functors*

$$(-)_\bot \colon \mathbf{ogPos} \to \mathbf{ogPos}^+, \qquad (-)_{\not\bot} \colon \mathbf{ogPos}^+ \to \mathbf{ogPos}$$

*inverse to each other up to natural isomorphism.*

The following is a useful property of regular directed complexes.

***1.30*** (Oriented thin graded poset). Let $P$ be an oriented graded poset with a positive least element. We say that $P$ is *oriented thin* if, for all $x, y \in P$ such that $x \leq y$ and $\mathrm{codim}_y(x) = 2$, the interval $[x, y]$ is of the form




for exactly two elements $z_1, z_2$, and for some $\alpha, \beta, \gamma \in \{+, -\}$.

***Proposition 1.31*** — *Let $P$ be a regular directed complex. Then $P_\perp$ is an oriented thin graded poset.*

The connection between oriented graded posets and strict $\omega$-categories is given by the fact that (isomorphism classes of) molecules form a strict $\omega$-category with pasting at the $k$-boundary as $k$-composition. The fibred version of this result implies that (isomorphism classes of) molecules *over an oriented graded poset $P$* form a strict $\omega$-category.

In what follows, we recall the *single-set* definition of strict $\omega$-category, which is most natural in this context, and state these results more precisely.

***1.32*** (Reflexive $\omega$-graph). A *reflexive $\omega$-graph* is a set $X$, whose elements are called *cells*, together with, for all $n \in \mathbb{N}$, operators $\partial_n^-, \partial_n^+ \colon X \to X$ called *input* and *output $n$-boundary*, satisfying the following axioms.

1. (*Finite dimension*). For all $t \in X$, there exists $n \in \mathbb{N}$ such that
$$\partial_n^- t = \partial_n^+ t = t.$$
2. (*Globularity*). For all $t \in X$, $k, n \in \mathbb{N}$, and $\alpha, \beta \in \{+, -\}$,
$$\partial_k^\alpha(\partial_n^\beta t) = \begin{cases} \partial_k^\alpha t & \text{if } k < n, \\ \partial_n^\beta t & \text{if } k \geq n. \end{cases}$$

If $t$ is a cell in a reflexive $\omega$-graph, the *dimension of $t$* is the natural number $\dim t \coloneqq \min\{n \in \mathbb{N} \mid \partial_n^- t = \partial_n^+ t = t\}$.

***1.33*** (Composable pair of cells). Let $t, u$ be a pair of cells in a reflexive $\omega$-graph, $k \in \mathbb{N}$. We say that $t$ and $u$ are *$k$-composable* if $\partial_k^+ t = \partial_k^- u$. We write
$$X \times_k X \coloneqq \left\{(t, u) \in X \times X \mid \partial_k^+ t = \partial_k^- u\right\}.$$
for the set of $k$-composable pairs of cells in $X$.

***1.34*** (Strict $\omega$-category). A *strict $\omega$-category* is a reflexive $\omega$-graph $X$ together with, for all $k \in \mathbb{N}$, an operation $- \#_k - \colon X \times_k X \to X$ called *$k$-composition*, satisfying the following axioms.

1. (*Compatibility with boundaries*). For all $k$-composable pairs of cells $t, u$, all $n \in \mathbb{N}$, and $\alpha \in \{+, -\}$,
$$\partial_n^\alpha(t \#_k u) = \begin{cases} \partial_n^\alpha t = \partial_n^\alpha u & \text{if } n < k, \\ \partial_k^- t & \text{if } n = k,\ \alpha = -, \\ \partial_k^+ u & \text{if } n = k,\ \alpha = +, \\ \partial_n^\alpha t \#_k \partial_n^\alpha u & \text{if } n > k. \end{cases}$$

2. (*Associativity*). For all cells $t, u, v$ such that either side of the equation is defined, $(t \#_k u) \#_k v = t \#_k (u \#_k v)$.
3. (*Unitality*). For all cells $t$, $t \#_k \partial_k^+ t = \partial_k^- t \#_k t = t$.
4. (*Interchange*). For all cells $t, t', u, u'$ and $n > k$ such that the left-hand side is defined, $(t \#_n t') \#_k (u \#_n u') = (t \#_k u) \#_n (t' \#_k u')$.

Given a strict $\omega$-category $X$ and $n \in \mathbb{N}$, we let $\sigma_{\leq n} X$ denote its *$n$-skeleton*, that is, its restriction to cells of dimension $\leq n$. A strict $\omega$-category is a *strict $n$-category* if it is equal to its $n$-skeleton.

***1.35*** (Strict functor of strict $\omega$-categories). Let $X, Y$ be strict $\omega$-categories. A *strict functor* $f \colon X \to Y$ is a function such that, for all $k, n \in \mathbb{N}$, $\alpha \in \{+, -\}$, and $k$-composable cells $t, u$ in $X$,

$$f(\partial_n^\alpha t) = \partial_n^\alpha f(t), \qquad f(t \#_k u) = f(t) \#_k f(u).$$

Strict $\omega$-categories and strict functors form a category $\omega\mathbf{Cat}$.

***1.36*** (Generating sets and bases). Let $X$ be a strict $\omega$-category and $\mathscr{S}$ a set of cells in $X$. The set $\operatorname{span} \mathscr{S}$ is the smallest set such that

1. if $t \in \mathscr{S}$, then $t \in \operatorname{span} \mathscr{S}$,
2. for all $k \in \mathbb{N}$, if $t, u \in \operatorname{span} \mathscr{S}$ are $k$-composable, then $t \#_k u \in \operatorname{span} \mathscr{S}$.

A *generating set for $X$* is a set $\mathscr{S}$ of cells such that $\operatorname{span} \mathscr{S}$ contains every cell in $X$. A *basis for $X$* is a minimal generating set.

***Lemma 1.37*** — *Let $f, g \colon X \to Y$ be strict functors and let $\mathscr{S}$ be a generating set for $X$. If $f(t) = g(t)$ for all $t \in \mathscr{S}$, then $f = g$.*

***1.38*** (Isomorphism classes of molecules). For each oriented graded poset $P$, let $[P]$ denote its isomorphism class in $\mathbf{ogPos}$. We let

$$\begin{aligned} \mathit{Mol} &\coloneqq \{[U] \mid U \text{ is a molecule}\}, \\ \mathit{Atom} &\coloneqq \{[U] \mid U \text{ is an atom}\} \subset \mathit{Mol}. \end{aligned}$$

***Proposition 1.39*** — *For all $n, k \in \mathbb{N}$ and $\alpha \in \{+, -\}$, let*

$$\begin{aligned} \partial_n^\alpha &\colon \mathit{Mol} \to \mathit{Mol}, & [U] &\mapsto [\partial_n^\alpha U], \\ - \#_k - &\colon \mathit{Mol} \times_k \mathit{Mol} \to \mathit{Mol}, & [U], [V] &\mapsto [U \#_k V]. \end{aligned}$$

*Then Mol together with these operations is a strict $\omega$-category. Moreover,*

1. *for all molecules $U$, $\dim [U] = \dim U$,*
2. *Atom is a basis for Mol.*

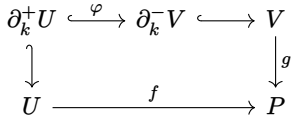
***1.40*** (Molecules over an oriented graded poset). For each morphism $f : U \to P$ of oriented graded posets, let $[f]$ denote its isomorphism class in the slice category $\mathbf{ogPos}/P$. Given an oriented graded poset $P$, we let

$$\begin{aligned} \mathit{Mol}/P &\coloneqq \{[f : U \to P] \mid U \text{ is a molecule}\}, \\ \mathit{Atom}/P &\coloneqq \{[f : U \to P] \mid U \text{ is an atom}\} \subseteq \mathit{Mol}/P, \end{aligned}$$

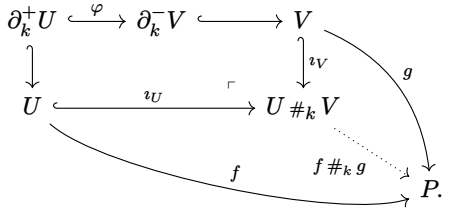
which we call *molecules* and *atoms over* $P$. For all $k \in \mathbb{N}$ and $\alpha \in \{+, -\}$,

$$\partial^\alpha_k : \mathit{Mol}/P \to \mathit{Mol}/P, \qquad [f : U \to P] \mapsto [f|_{\partial^\alpha_k U} : \partial^\alpha_k U \to P]$$

make $\mathit{Mol}/P$ a reflexive $\omega$-graph. If $[f : U \to P]$, $[g : V \to P]$ are $k$-composable molecules over $P$, then there exists a unique isomorphism $\varphi$ such that




commutes, which induces, by the universal property of $U \#_k V$, a unique morphism $f \#_k g$ such that the following diagram commutes:




***Proposition 1.41*** — *Let $P$ be an oriented graded poset and, for each $k \in \mathbb{N}$,*

$$\begin{aligned} - \#_k - &: \mathit{Mol}/P \times_k \mathit{Mol}/P \to \mathit{Mol}/P, \\ &[f : U \to P], [g : V \to P] \mapsto [f \#_k g : U \#_k V \to P]. \end{aligned}$$

*Then $\mathit{Mol}/P$ together with these composition operations is a strict $\omega$-category, which has the set $\mathit{Atom}/P$ as a basis. In particular, if $\dim P \leq n$, then $\mathit{Mol}/P$ is a strict $n$-category. This assignment extends to a functor $\mathit{Mol}/\_ : \mathbf{ogPos} \to \omega\mathbf{Cat}$.*

When $P$ is a regular directed complex, $\mathit{Mol}/P$ admits a basis whose elements are in bijection with the elements of $P$, as a consequence of the following result. This comes from very strong rigidity properties of atoms, which do not generalise to other oriented graded posets.

***1.42*** (Local embedding of oriented graded posets). A morphism $f\colon P \to Q$ of oriented graded posets is a *local embedding* if, for all $x \in P$, the restriction $f|_{\mathrm{cl}\{x\}}$ is an inclusion, hence determines an isomorphism between $\mathrm{cl}\,\{x\}$ and its image $\mathrm{cl}\,\{f(x)\}$.

***Proposition 1.43*** — *Let $f\colon P \to Q$ be a morphism of regular directed complexes. Then $f$ is a local embedding.*

***Corollary 1.44*** — *If $P$ is a regular directed complex, $\{[\mathrm{cl}\,\{x\} \hookrightarrow P] \mid x \in P\}$ is a basis for the $\omega$-category $Mol/P$.*

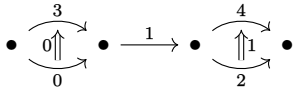
***1.45*** (Diagram in a strict $\omega$-category). Let $X$ be a strict $\omega$-category and $P$ a regular directed complex. A *diagram of shape $P$ in $X$* is a strict functor $d\colon Mol/P \to X$. A diagram is a *pasting diagram* if its shape is a molecule.

## 2. Layerings, flow graphs, and orderings

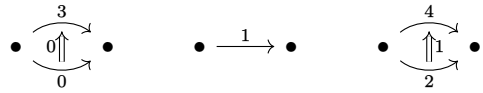
Since pasting of molecules satisfies the axioms of strict $\omega$-categories, it is clear that every molecule admits multiple pasting decompositions. However, the space of possible decompositions can at least in part be constrained by considering decompositions of a special type, called *layerings*, where each factor (layer) contains exactly one maximal cell of dimension higher than the pasting dimension. Such decompositions have played a role in most past approaches to higher-categorical diagrams — see for example [For22] — but we first studied them systematically in [HK23]. We give an overview of the main notions and results, and refer to [Had24, Chapter 4] for proofs.

***2.1*** (Layering). Let $U$ be a molecule, $-1 \leq k < \dim U$. A *$k$-layering of $U$* is a sequence $(U^{(i)})_{i=1}^{m}$ of molecules such that $U = U^{(1)} \#_k \dots \#_k U^{(m)}$ and $\left|\bigcup_{i>k}(\mathscr{M}ax\, U^{(j)})_i\right| = 1$ for all $j \in \{1,\dots,m\}$, that is, each "layer" $U^{(j)}$ contains a single maximal element of dimension $> k$.

*Example 2.2* — If $U$ is the molecule encoding the 2-dimensional pasting diagram shape




then $U$ admits a single 0-layering with layers




and two 1-layerings with layers


respectively.

***2.3*** (Frame and layering dimension). Let $U$ be a molecule. The *frame dimension* of $U$ is the integer

$$\operatorname{frdim} U \coloneqq \dim \bigcup \{(\operatorname{cl}\{x\} \cap \operatorname{cl}\{y\}) \mid x, y \in \mathscr{M}ax\, U, x \neq y\}.$$

The *layering dimension* of $U$ is the integer

$$\operatorname{lydim} U \coloneqq \min \left\{ k \geq -1 \mid \left| \bigcup_{i>k+1} (\mathscr{M}ax\, U)_i \right| \leq 1 \right\}.$$

***Proposition 2.4*** — *Let $U$ be a molecule. Then*

1. *there exists $k \geq -1$ such that $U$ admits a $k$-layering,*
2. *if $U$ admits a $k$-layering, it admits an $\ell$-layering for all $k \leq \ell < \dim U$,*
3. $\operatorname{frdim} U \leq \min\{k \geq -1 \mid U \text{ admits a } k\text{-layering}\} \leq \operatorname{lydim} U$.

***Lemma 2.5*** — *Let $U$ be a molecule. Then*

1. $\operatorname{lydim} U = -1$ *if and only if* $\operatorname{frdim} U = -1$ *if and only if $U$ is an atom,*
2. *if $k \geq 0$ and $(U^{(i)})_{i=1}^m$ is a $k$-layering of $U$, then for each $i \in \{1, \dots, m\}$,* $\operatorname{lydim} U^{(i)} < k$.

*Comment 2.6* — Proposition 2.4 in conjunction with Lemma 2.5 allows us to prove properties of molecules *by induction on their layering dimension.* That is, to prove that a property holds of all molecules $U$, it suffices to

- prove that it holds when $\operatorname{lydim} U = -1$, that is, when $U$ is an atom,
- prove that it holds when $k \coloneqq \operatorname{lydim} U \geq 0$, assuming that it holds of all the $(U^{(i)})_{i=1}^m$ in a $k$-layering of $U$.

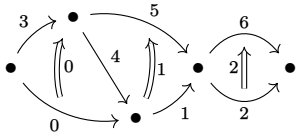
***2.7*** (Flow graph). Let $P$ be an oriented graded poset, $k \geq -1$. The *$k$-flow graph of $P$* is the directed graph $\mathscr{F}_k P$ whose

- set of vertices is $\bigcup_{i>k} P_i$, and
- set of edges is
$$\left\{(x, y) \mid \Delta_k^+ x \cap \Delta_k^- y \neq \varnothing\right\},$$
where the source of $(x, y)$ is $x$ and the target is $y$.

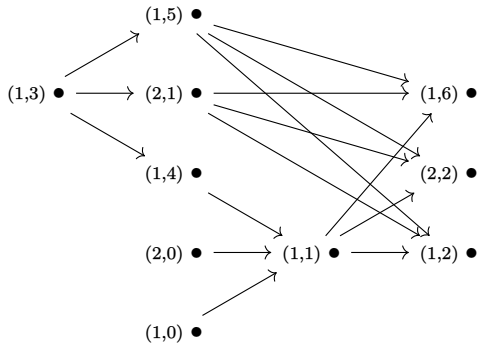
***2.8*** (Maximal flow graph). Let $P$ be a finite-dimensional oriented graded poset, $k \geq -1$. The *maximal $k$-flow graph of $P$* is the induced subgraph $\mathscr{M}_k P$ of $\mathscr{F}_k P$ on the vertex set

$$\bigcup_{i>k} (\mathscr{M}ax\, P)_i \subseteq \bigcup_{i>k} P_i.$$

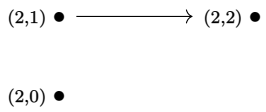
*Example 2.9* — If $U$ is the molecule encoding the 2-dimensional pasting diagram shape




the 0-flow graph $\mathscr{F}_0 U$ is




and the maximal 0-flow graph $\mathscr{M}_0 U$ is its induced subgraph




while the 1-flow graph $\mathscr{F}_1 U$ is


and it is equal to $\mathscr{M}_1 U$, since every 2-dimensional element of $U$ is maximal.

***2.10*** (Ordering of a molecule). Let $U$ be a molecule, $k \geq -1$. A *$k$-ordering of $U$* is a topological sort of $\mathscr{M}_k U$.

*Remark 2.11* — A $k$-ordering of $U$ exists if and only if $\mathscr{M}_k U$ is acyclic.

***Proposition 2.12*** — *Let $U$ be a molecule, $k \geq -1$, and let*

$$\mathscr{L}ay_k U \coloneqq \left\{ k\text{-layerings } (U^{(i)})_{i=1}^m \text{ of } U \text{ up to layer-wise isomorphism} \right\},$$
$$\mathscr{O}rd_k U \coloneqq \left\{ k\text{-orderings } (x^{(i)})_{i=1}^m \text{ of } U \right\}.$$

*For each $k$-layering $(U^{(i)})_{i=1}^m$ of $U$ and each $i \in \{1, \dots, m\}$, let $x^{(i)}$ be the only element of $\bigcup_{j>k}(\mathscr{M}ax\, U)_j$ in the layer $U^{(i)}$. Then the assignment*

$$\mathsf{o}_{k,U} \colon (U^{(i)})_{i=1}^m \mapsto (x^{(i)})_{i=1}^m$$

*determines an injective function $\mathscr{L}ay_k U \hookrightarrow \mathscr{O}rd_k U$. Moreover, if $U$ admits a $k$-layering, then for all $\ell > k$, the function $\mathsf{o}_{k,U}$ is a bijection.*

## 3. Frame-acyclicity and polygraphs

In this section, we study what seems to be the mildest acyclicity condition on an oriented graded poset $P$ guaranteeing that $Mol/P$ is freely generated in the sense of polygraphs. On molecules, this condition, which we call *frame-acyclicity*, is (non-trivially) equivalent to what is called being *split* in [Ste93]; our treatment elucidates its status as an acyclicity condition, which was not originally recognised.

Frame-acyclicity was first defined in [Had21], and in [HK23] we studied its role in algorithmic properties of higher-dimensional rewriting. However, self-contained proofs of the main results related to frame-acyclicity have not appeared in print; this section is meant to fix this gap.

***3.1*** (Submolecules). Let $V \subseteq U$ be molecules. We say that $V$ is a *submolecule* of $U$, and write $V \sqsubseteq U$, if $V$ is a factor in a pasting decomposition of $U$.

*Remark 3.2* — Equivalently, $\sqsubseteq$ can be characterised as the smallest partial order on molecules such that $U, V \sqsubseteq U \#_k V$ whenever the latter is defined, once $U$ and $V$ are identified with their images in the pasting.

***3.3*** (Frame-acyclic molecule). Let $U$ be a molecule. We say that $U$ is *frame-acyclic* if for all submolecules $V \sqsubseteq U$, if $r \coloneqq \operatorname{frdim} V$, then $\mathscr{M}_r V$ is acyclic.

***3.4*** (Oriented graded poset with frame-acyclic molecules). Let $P$ be an oriented graded poset. We say that $P$ *has frame-acyclic molecules* if, for all molecules $U$, if there exists a morphism $f \colon U \to P$, then $U$ is frame-acyclic.

We recall without proof the following result [HK23, Theorem 121], which implies that this condition is only non-trivial starting from dimension 4.

***Theorem 3.5*** — *Let $U$ be a molecule, $\dim U \leq 3$. Then $U$ is frame-acyclic.*

***Corollary 3.6*** — *Let $P$ be an oriented graded poset, $\dim P \leq 3$. Then $P$ has frame-acyclic molecules.*

*Example 3.7* — The dimensional bound in Theorem 3.5 is strict: [HK23, Example 126], exhibits a 4-dimensional molecule which is not frame-acyclic.

***Lemma 3.8 —*** *Let $U$ be a molecule. Suppose that for all submolecules $V \sqsubseteq U$, if $r \coloneqq \operatorname{frdim} V$, then $V$ admits an $r$-layering. Then for all $k \geq \operatorname{frdim} U$ the function $\mathsf{o}_{k,U} \colon \mathscr{L}ay_k U \hookrightarrow \mathscr{O}rd_k U$ is a bijection.*

*Proof.* Let $r \coloneqq \operatorname{frdim} U$. By assumption, there exists an $r$-layering of $U$, so by Proposition 2.12 it suffices to show that $\mathsf{o}_{r,U}$ is a bijection.

Given two $r$-orderings $(x^{(i)})_{i=1}^m$ and $(y^{(i)})_{i=1}^m$, there exists a unique permutation $\sigma$ such that $x^{(i)} = y^{(\sigma(i))}$ for all $i \in \{1, \dots, m\}$. Let $\mathsf{d}((x^{(i)})_{i=1}^m, (y^{(i)})_{i=1}^m)$ be the number of pairs $(j, j')$ such that $j < j'$ but $\sigma(j') < \sigma(j)$. Under the assumption that $(x^{(i)})_{i=1}^m$ is in the image of $\mathsf{o}_{r,U}$, we will prove that $(y^{(i)})_{i=1}^m$ is also in the image of $\mathsf{o}_{r,U}$ by induction on $\mathsf{d}((x^{(i)})_{i=1}^m, (y^{(i)})_{i=1}^m)$. Since the image of $\mathsf{o}_{r,U}$ is not empty, this will suffice to prove that $\mathsf{o}_{r,U}$ is surjective, hence bijective by Proposition 2.12.

If $\mathsf{d}((x^{(i)})_{i=1}^m, (y^{(i)})_{i=1}^m) = 0$, then $x^{(i)} = y^{(i)}$ for all $i \in \{1, \dots, m\}$, and there is nothing left to prove.

Suppose $\mathsf{d}((x^{(i)})_{i=1}^m, (y^{(i)})_{i=1}^m) > 0$. Then there exists $j < m$ such that $\sigma(j+1) < \sigma(j)$. Suppose $(x^{(i)})_{i=1}^m$ is the image of the $r$-layering $(U^{(i)})_{i=1}^m$. Let $V \sqsubseteq U$ be the image of $U^{(j)} \#_r U^{(j+1)}$ in $U$, and let

$$z_1 \coloneqq x^{(j)} = y^{(\sigma(j))}, \quad z_2 \coloneqq x^{(j+1)} = y^{(\sigma(j+1))}.$$

Because $z_1$ comes before $z_2$ in one $r$-ordering, but after in another, there can be no edge between them in $\mathscr{M}_r U$, so

$$\dim \left( \operatorname{cl}\{z_1\} \cap \operatorname{cl}\{z_2\} \right) < r.$$

Since $z_1, z_2$ are the only maximal elements of dimension $> r$ in $V$, we deduce that $\ell \coloneqq \operatorname{frdim} V < r$. By assumption, there exists an $\ell$-layering of $V$. In particular, there exist molecules $V^{(1)}, V^{(2)}$ such that

1. $z_i$ is in the image of $V^{(i)}$ for all $i \in \{1, 2\}$, and
2. $V$ is isomorphic to $V^{(1)} \#_\ell V^{(2)}$ or to $V^{(2)} \#_\ell V^{(1)}$.

Without loss of generality suppose that $V$ is isomorphic to $V^{(1)} \#_\ell V^{(2)}$. By the unitality and interchange properties of pasting, letting

$$\begin{aligned} \tilde{U}^{(j)} &\coloneqq \partial_r^- V^{(1)} \#_\ell V^{(2)}, \\ \tilde{U}^{(j+1)} &\coloneqq V^{(1)} \#_\ell \partial_r^+ V^{(2)}, \end{aligned}$$

we have that $V = \tilde{U}^{(j)} \#_r \tilde{U}^{(j+1)}$. Letting $\tilde{U}^{(i)} \coloneqq U^{(i)}$ for $i \notin \{j, j+1\}$, we have that $(\tilde{U}^{(i)})_{i=1}^m$ is an $r$-layering of $U$, and

$$\mathsf{o}_{r,U} \colon (\tilde{U}^{(i)})_{i=1}^m \mapsto (\tilde{x}^{(i)})_{i=1}^m = (x^{(1)}, \dots, x^{(j+1)}, x^{(j)}, \dots, x^{(m)}).$$

Then $\mathsf{d}((\tilde{x}^{(i)})_{i=1}^m, (y^{(i)})_{i=1}^m) < \mathsf{d}((x^{(i)})_{i=1}^m, (y^{(i)})_{i=1}^m)$ and $(\tilde{x}^{(i)})_{i=1}^m$ is in the image of $\mathsf{o}_{r,U}$. We conclude by the inductive hypothesis. ∎

*Comment 3.9* — Let P be a property of molecules such that, whenever P holds of a molecule $U$, then P holds of every submolecule $V \sqsubseteq U$; the property of frame-acyclicity is of this sort. Because every proper submolecule of $U$ has strictly fewer elements than $U$, the submolecule relation on $U$ is well-founded, and its minimal elements are the 0-dimensional one-element subsets $\{x\} \sqsubseteq U$ for each $x \in U_0$.

If we want to prove that P implies Q for all molecules, we can then proceed by *induction on submolecules*: assume that a molecule $U$ satisfies P, then

- prove that $\{x\}$ satisfies Q for all $x \in U_0$,
- prove that $U$ satisfies Q under the assumption that every proper submolecule $V \sqsubseteq U$ satisfies Q.

***Lemma 3.10*** — *Let $U$ be a molecule and let $x, y \in \mathscr{M}ax\, U$ such that $x \neq y$. For all $k \geq \operatorname{frdim} U$, $\operatorname{cl}\{x\} \cap \operatorname{cl}\{y\} = (\partial_k^- x \cap \partial_k^+ y) \cup (\partial_k^+ x \cap \partial_k^- y)$.*

*Proof.* See [Ste93, Proposition 6.4]. ∎

***Theorem 3.11*** — *Let $U$ be a molecule, $r \coloneqq \operatorname{frdim} U$. If $U$ is frame-acyclic, then $U$ admits an $r$-layering.*

*Proof.* We proceed by induction on submolecules. For all $x \in U_0$, we have $\operatorname{frdim}\{x\} = -1$, and $\{x\}$ admits the trivial $(-1)$-layering, which proves the base case.

We construct a finite plane tree of submolecules $U^{(j_1,\dots,j_p)} \sqsubseteq U$, as follows:

- the root is $U^{()} \coloneqq U$;
- if $\operatorname{lydim} U^{(j_1,\dots,j_p)} \leq r$, then we let $\operatorname{lydim} U^{(j_1,\dots,j_p)}$ be a leaf;
- if $k \coloneqq \operatorname{lydim} U^{(j_1,\dots,j_p)} > r$, then we pick a $k$-layering $(V^{(i)})_{i=1}^q$ of $U^{(j_1,\dots,j_p)}$, and for each $i \in \{1,\dots,q\}$, we let the image of $V^{(i)}$ be a child $U^{(j_1,\dots,j_p,i)}$ of $U^{(j_1,\dots,j_p)}$.

By Lemma 2.5, the layering dimension of the children of a node is strictly smaller than that of the node, so the procedure terminates.

Fix an $r$-ordering $(x^{(i)})_{i=1}^m$ of $U$; this is possible because $\mathscr{M}_r U$ is acyclic. Let $V \coloneqq U^{(j_1,\dots,j_p)}$ be a node of the tree. We have

$$\bigcup_{j>r} (\mathscr{M}ax\, V)_j = \sum_{i=1}^m \bigcup_{j>r} \left( (\mathscr{M}ax\, V)_j \cap \operatorname{cl}\left\{x^{(i)}\right\} \right) =: \sum_{i=1}^m M^{(i)};$$

the $M^{(i)}$ form a partition because $\operatorname{frdim} U = r$, so every element of dimension $> r$ is in the closure of $x^{(i)}$ for a unique $i \in \{1,\dots,m\}$. We claim that $V$ is isomorphic to $V^{(1)} \#_r \dots \#_r V^{(m)}$ for some molecules $(V^{(i)})_{i=1}^m$ such that, for each $i \in \{1,\dots,m\}$, identifying $V^{(i)}$ with its image in $V$, we have

$$\bigcup_{j>r} (\mathscr{M}ax\, V^{(i)})_j = M^{(i)}.$$

We will prove this by backward induction on the tree $U^{(j_1,\dots,j_p)}$.

Suppose $V$ is a leaf, so $\operatorname{lydim} V \le r$. Then $V$ admits an $r$-layering. For each $i \in \{1,\dots,m\}$, fix a topological sort $(y^{(i,j)})_{j=1}^{p_i}$ of the induced subgraph $\mathscr{M}_r V|_{M^{(i)}}$. We claim that $((y^{(i,j)})_{j=1}^{p_i})_{i=1}^{m}$ is an $r$-ordering of $V$.

Suppose there is an edge from $x$ to $x'$ in $\mathscr{M}_r V$. Then $x \in M^{(i)}$, $x' \in M^{(i')}$ for a unique pair $i, i' \in \{1,\dots,m\}$. If $i = i'$, then $x = y^{(i,j)}$ and $x' = y^{(i,j')}$ for some $j, j' \in \{1,\dots,p_i\}$, and $j < j'$ because $(y^{(i,j)})_{j=1}^{p_i}$ is a topological sort of $\mathscr{M}_r V|_{M^{(i)}}$. If $i \ne i'$, then there exists

$$z \in \Delta_r^+ x \cap \Delta_r^- x' \subseteq \operatorname{cl}\left\{x^{(i)}\right\} \cap \operatorname{cl}\left\{x^{(i')}\right\}.$$

Since $\partial_r^\alpha x^{(i)}$ and $\partial_r^\alpha x^{(i')}$ is pure and $r$-dimensional for all $\alpha \in \{+,-\}$, by Lemma 3.10

$$z \in (\Delta_r^+ x^{(i)} \cap \Delta_r^- x^{(i')}) \cup (\Delta_r^- x^{(i)} \cap \Delta_r^+ x^{(i')}),$$

and by Lemma 1.15 $\Delta_r^- x^{(i)} \cap \operatorname{cl}\{x\} \subseteq \Delta_r^- x$ which is disjoint from $\Delta_r^+ x$, so $z \in \Delta_r^+ x^{(i)} \cap \Delta_r^- x^{(i')}$. It follows that there is an edge from $x^{(i)}$ to $x^{(i')}$ in $\mathscr{M}_r U$, so $i < i'$ because $(x^{(i)})_{i=1}^m$ is a topological sort of $\mathscr{M}_r U$. This proves that $((y^{(i,j)})_{j=1}^{p_i})_{i=1}^m$ is an $r$-ordering of $V$.

Let $W \sqsubseteq V$, $\ell \coloneqq \operatorname{frdim} W$. If $V \ne U$ or $W \ne U$, then $W$ admits an $\ell$-layering by the inductive hypothesis on proper submolecules of $U$. If $W = V = U$ then $\ell = r$ and $W$ admits an $\ell$-layering by Proposition 2.4. In either case, $V$ satisfies the conditions of Lemma 3.8, and since $r \ge \operatorname{lydim} V \ge \operatorname{frdim} V$, every $r$-ordering of $V$ comes from an $r$-layering of $V$.

It follows that $((y^{(i,j)})_{j=1}^{p_i})_{i=1}^m$ comes from an $r$-layering $((W^{(i,j)})_{j=1}^{p_i})_{i=1}^m$, and we can define $V^{(i)} \coloneqq W^{(i,1)} \#_r \dots \#_r W^{(i,p_i)}$ for each $i \in \{1,\dots,m\}$, satisfying the desired condition.

Now, suppose that $V$ is not a leaf, so $k \coloneqq \operatorname{lydim} V > r$, and $V$ has children $(W^{(j)})_{j=1}^q$ forming a $k$-layering of $V$. By the inductive hypothesis, each of the $W^{(j)}$ has a decomposition $W^{(j,1)} \#_r \dots \#_r W^{(j,m)}$ such that the maximal elements of dimension $> r$ in the image of $W^{(j,i)}$ are contained in $\operatorname{cl}\left\{x^{(i)}\right\}$. Then, for each $i \in \{1,\dots,m\}$ and $j, j' \in \{1,\dots,q\}$, $W^{(j,i)} \cap W^{(j')} \subseteq W^{(j',i)}$, so $V^{(i)} \coloneqq W^{(1,i)} \#_k \dots \#_k W^{(q,i)}$ is defined. Using interchange repeatedly, we conclude that $V$ is isomorphic to $V^{(1)} \#_r \dots \#_r V^{(m)}$.

This concludes the induction on the tree $U^{(j_1,\dots,j_p)}$. In particular, for the root $U^{()} = U$, the decomposition $U^{(1)} \#_r \dots \#_r U^{(m)}$ satisfies

$$\bigcup_{j>r} (\mathscr{M}ax\, U^{(i)})_j = \left\{x^{(i)}\right\},$$

that is, $(U^{(i)})_{i=1}^m$ is an $r$-layering of $U$. ∎

***Corollary 3.12*** — *Let $U$ be a molecule. The following are equivalent:*

*(a) $U$ is frame-acyclic;*

*(b) for all $V \sqsubseteq U$ and all* $\operatorname{frdim} V \leq k < \dim V$*, $V$ admits a $k$-layering;*

*(c) for all $V \sqsubseteq U$ and all* $\operatorname{frdim} V \leq k < \dim V$*, the sets $\mathscr{L}ay_k V$ and $\mathscr{O}rd_k V$ are non-empty and equinumerous.*

*Proof.* The implication from (a) to (b) is a consequence of Theorem 3.11 together with Proposition 2.4. The implication from (b) to (c) is Lemma 3.8. Finally, the implication from (c) to (a) follows from Proposition 2.12. ∎

*Remark 3.13* — Corollary 3.12 implies that a molecule is frame-acyclic if and only if it is *split* in the sense of [Ste93]. Thus this condition rephrases the splitness condition as an acyclicity condition, which will help us elucidate its connection to other, stronger acyclicity conditions.

***3.14*** (Cellular extension of a strict $\omega$-category). Let $X$ be a strict $\omega$-category. A *cellular extension of $X$* is a strict $\omega$-category $X_{\mathscr{S}}$ together with a pushout diagram

$$\begin{array}{ccc} \coprod_{e \in \mathscr{S}} Mol/\partial U_e & \xhookrightarrow{\coprod_{e \in \mathscr{S}} Mol/\imath_e} & \coprod_{e \in \mathscr{S}} Mol/U_e \\ \downarrow{\scriptstyle (\partial e)_{e \in \mathscr{S}}} & & \downarrow{\scriptstyle (e)_{e \in \mathscr{S}}} \\ X & \xhookrightarrow{\qquad\qquad} & X_{\mathscr{S}} \end{array}$$

in $\omega\mathbf{Cat}$, where, for each $e \in \mathscr{S}$, $U_e$ is an atom and $\imath_e \colon \partial U_e \hookrightarrow U_e$ is the inclusion of its boundary.

*Comment 3.15* — The functor $X \hookrightarrow X_{\mathscr{S}}$ in a cellular extension is always injective, as shown in [Mak05, Section 4].

This is a non-standard definition of cellular extension, allowing any atom as a potential cell shape; the usual definition only uses *globes*. However, the two are equivalent in the sense that a cellular extension in our sense can always be turned into a cellular extension in the more restrictive sense.

***3.16*** (Polygraph). A *polygraph*, also known as *computad*, is a strict $\omega$-category $X$ together with, for each $n \in \mathbb{N}$, a pushout diagram

$$\begin{array}{ccc} \coprod_{e \in \mathscr{S}_n} Mol/\partial U_e & \xhookrightarrow{\coprod_{e \in \mathscr{S}_n} Mol/\imath_e} & \coprod_{e \in \mathscr{S}_n} Mol/U_e \\ \downarrow{\scriptstyle (\partial e)_{e \in \mathscr{S}_n}} & & \downarrow{\scriptstyle (e)_{e \in \mathscr{S}_n}} \\ \sigma_{\leq n-1} X & \xhookrightarrow{\qquad\qquad} & \sigma_{\leq n} X \end{array}$$

in $\omega\mathbf{Cat}$, exhibiting $\sigma_{\leq n} X$ as a cellular extension of $\sigma_{\leq n-1} X$, such that $U_e$ is an $n$-dimensional atom for all $e \in \mathscr{S}_n$. The set

$$\mathscr{S} \coloneqq \sum_{n \in \mathbb{N}} \{e[\mathrm{id}_{U_e}] \mid e \in \mathscr{S}_n\}$$

is called the set of *generating cells* of the polygraph. We write $(X, \mathscr{S})$ for a polygraph $X$ with set $\mathscr{S}$ of generating cells.

***Lemma 3.17*** — *Let $(X, \mathscr{S})$ be a polygraph. Then $\mathscr{S}$ is a basis for $X$.*

*Proof.* The fact that $\mathscr{S}$ is a generating set and its minimality are consequences of [ABG+23, Proposition 15.1.8 and Lemma 16.6.2], respectively. ∎

***Lemma 3.18*** — *Let $P$ be an oriented graded poset, $n \in \mathbb{N}$, and let $\mathscr{S}_n$ be a set containing one pasting diagram*

$$e \equiv Mol\!/_e \colon Mol\!/_{U_e} \to \sigma_{\leq n} Mol\!/_P$$

*for each $[e \colon U_e \to P]$ in $Atom\!/_P$ such that $\dim U_e = n$. If $\sigma_{\leq n} P$ has frame-acyclic molecules, then*

$$\begin{array}{ccc} \coprod_{e \in \mathscr{S}_n} Mol\!/_{\partial U_e} & \xhookrightarrow{\coprod_{e \in \mathscr{S}_n} Mol\!/_{\imath_e}} & \coprod_{e \in \mathscr{S}_n} Mol\!/_{U_e} \\ \downarrow{\scriptstyle (\partial e)_{e \in \mathscr{S}_n}} & & \downarrow{\scriptstyle (e)_{e \in \mathscr{S}_n}} \\ \sigma_{\leq n-1} Mol\!/_P & \xhookrightarrow{\qquad\qquad} & \sigma_{\leq n} Mol\!/_P \end{array}$$

*is a pushout diagram in $\omega\mathbf{Cat}$, exhibiting $\sigma_{\leq n} Mol\!/_P$ as a cellular extension of $\sigma_{\leq n-1} Mol\!/_P$.*

*Proof.* Let $X$ be a strict $\omega$-category and let

$$\begin{array}{ccc} \coprod_{e \in \mathscr{S}_n} Mol\!/_{\partial U_e} & \xhookrightarrow{\coprod_{e \in \mathscr{S}_n} Mol\!/_{\imath_e}} & \coprod_{e \in \mathscr{S}_n} Mol\!/_{U_e} \\ \downarrow{\scriptstyle (\partial e)_{e \in \mathscr{S}_n}} & & \downarrow{\scriptstyle \ell} \\ \sigma_{\leq n-1} Mol\!/_P & \xrightarrow{\qquad h \qquad} & X \end{array}$$

be a commutative diagram of strict functors. We define $\bar{h} \colon \sigma_{\leq n} Mol\!/_P \to X$ as follows on cells $[f \colon U \to P]$ in $\sigma_{\leq n} Mol\!/_P$. If $\dim U < n$, then we let

$$\bar{h}[f] \coloneqq h[f].$$

Suppose $\dim U = n$; we proceed by induction on $\operatorname{lydim} U$. If $\operatorname{lydim} U = -1$, then by Lemma 2.5 $U$ is an atom, so there exists a unique $Mol\!/_e \in \mathscr{S}_n$ such that $[f] = [e]$, and we let $\bar{h}[f] \coloneqq \ell[\mathrm{id}_{U_e}]$. If $\operatorname{lydim} U = k \geq 0$, then $U$ admits a $k$-layering $(U^{(i)})_{i=1}^m$, and each layer $U^{(i)}$ has strictly lower layering dimension. Then we let

$$\bar{h}[f] \coloneqq \bar{h}[f|_{U^{(1)}}] \#_k \cdots \#_k \bar{h}[f|_{U^{(m)}}].$$

By construction, if $\bar{h}$ is well-defined, then it is a strict functor satisfying $\bar{h} \circ (e)_{e \in \mathscr{S}_n} = \ell$ and restricting to $h$ on $\sigma_{\leq n-1} Mol\!/_P$. Moreover, let $h'$ be

another strict functor with the same property. Then $h'$ agrees with $\bar{h}$ on all atoms of dimension $\leq n$, which form a basis of $\sigma_{\leq n}Mol/P$. It follows from Lemma 1.37 that $h' = \bar{h}$. It only remains to show that $\bar{h}$ is well-defined, that is, it is independent of the choice of a $k$-layering of $U$ when $\dim U = n$ and $k \coloneqq \operatorname{lydim} U \geq 0$.

We may assume, inductively, that $\bar{h}$ is well-defined on all cells $[g\colon V \to P]$ such that $\dim V < n$ or $\operatorname{lydim} V < k$. Let $(U^{(i)})_{i=1}^m$ and $(V^{(i)})_{i=1}^m$ be two $k$-layerings of $U$ and let $(x^{(i)})_{i=1}^m$, $(y^{(i)})_{i=1}^m$ be the induced $k$-orderings. We now proceed as in the proof of Lemma 3.8, letting $\sigma$ be the unique permutation such that $x^{(i)} = y^{(\sigma(i))}$ for all $i \in \{1, \dots, m\}$, letting

$$d \coloneqq \mathsf{d}((x^{(i)})_{i=1}^m, (y^{(i)})_{i=1}^m)$$

be the number of pairs $(j, j')$ such that $j < j'$ but $\sigma(j') < \sigma(j)$, and proceeding by induction on $d$. If $d = 0$, then the two layerings are equal up to layer-wise isomorphism. If $d > 0$, then there exists $j < m$ such that $\sigma(j+1) < \sigma(j)$, and we let $W \sqsubseteq U$ be the image of $U^{(j)} \#_k U^{(j+1)}$ in $U$. Then $W$ contains exactly two elements $z_1 \coloneqq x^{(j)} = y^{(\sigma(j))}$ and $z_2 \coloneqq x^{(j+1)} = y^{(\sigma(j+1))}$ of dimension $> k$, yet there can be no edge between them in $\mathscr{M}_k U$, from which we deduce that $r \coloneqq \operatorname{frdim} W < k$. By assumption, $W$ is frame-acyclic, so by Theorem 3.11 there exists an $r$-layering of $W$, hence also a pair of molecules $W^{(1)}$, $W^{(2)}$, each containing a single element of dimension $> k$, such that $W$ is isomorphic to $W^{(1)} \#_r W^{(2)}$. We may assume, without loss of generality, that $z_1$ is in the image of $W^{(1)}$ and $z_2$ in the image of $W^{(2)}$. We then have

$$\begin{aligned}
&\bar{h}[f|_{U^{(j)}}] \#_k \bar{h}[f|_{U^{(j+1)}}] = \\
&\quad = \left(\bar{h}[f|_{W^{(1)}}] \#_r \bar{h}[f|_{\partial_k^- W^{(2)}}]\right) \#_k \left(\bar{h}[f|_{\partial_k^+ W^{(1)}}] \#_r \bar{h}[f|_{W^{(2)}}]\right),
\end{aligned}$$

which by interchange and unitality in $X$ is equal to

$$\begin{aligned}
&\bar{h}[f|_{W^{(1)}}] \#_r \bar{h}[f|_{W^{(2)}}] = \\
&\quad = \left(\bar{h}[f|_{\partial_k^- W^{(1)}}] \#_r \bar{h}[f|_{W^{(2)}}]\right) \#_k \left(\bar{h}[f|_{W^{(1)}}] \#_r \bar{h}[f|_{\partial_k^+ W^{(2)}}]\right) = \\
&\quad = \bar{h}[f|_{\tilde{U}^{(j)}}] \#_k \bar{h}[f|_{\tilde{U}^{(j+1)}}],
\end{aligned}$$

where we let $\tilde{U}^{(j)} \coloneqq \partial_k^- W^{(1)} \#_r W^{(2)}$ and $\tilde{U}^{(j+1)} \coloneqq W^{(1)} \#_r \partial_k^+ W^{(2)}$. Notice that all the $n$-dimensional cells in this calculation involve molecules whose layering dimension is $< k$, so $\bar{h}$ is well-defined on each of them. Letting $\tilde{U}^{(i)} \coloneqq U^{(i)}$ for all $i \notin \{j, j+1\}$, we have that

1. $(\tilde{U}^{(i)})_{i=1}^m$ is a $k$-layering of $U$,
2. the definition of $\bar{h}[f]$ using $(U^{(i)})_{i=1}^m$ is equal to the one using $(\tilde{U}^{(i)})$, and
3. the induced $k$-ordering $(\tilde{x}^{(i)})_{i=1}^m \coloneqq (x^{(1)}, \dots, x^{(j+1)}, x^{(j)}, \dots, x^{(m)})$ satisfies $\mathsf{d}((\tilde{x}^{(i)})_{i=1}^m, (y^{(i)})_{i=1}^m) < d$,

so, by the inductive hypothesis on $d$, the definition of $\bar{h}[f]$ using $(\tilde{U}^{(i)})_{i=1}^{m}$ is equal to the definition using $(V^{(i)})_{i=1}^{m}$. We conclude that $\bar{h}[f]$ is well-defined, which completes the proof. ∎

***Theorem 3.19*** — *Let $P$ be an oriented graded poset with frame-acyclic molecules. Then $Mol/P$ is a polygraph whose set of generating cells is $Atom/P$.*

*Proof.* If $P$ has frame-acyclic molecules, then $\sigma_{\leq n}P$ has frame-acyclic molecules for all $n \in \mathbb{N}$. The statement then follows from Lemma 3.18. ∎

*Comment 3.20* — In fact, by the roundness property of atoms, if $Mol/P$ is a polygraph, then it is a *regular polygraph* in the sense of [Hen18].

*Comment 3.21* — Frame-acyclic molecules seem to be the tightest combinatorial condition ensuring that $Mol/P$ is a polygraph: the 4-dimensional molecule of [HK23, Example 126] does not present a polygraph, since its two possible 3-layerings cannot be related by applications of the interchange equation. This does not mean that having frame-acyclic molecules is *equivalent* to $Mol/P$ being a polygraph: one can engineer a variant of this example where a 4-dimensional molecule only has one valid 3-layering (even though it has multiple 3-orderings), so there are no "extra equations". However the algebraic freeness of such examples can be seen as accidental.

## 4. Dimension-wise acyclicity and Steiner complexes

In the previouse section we saw that the property of frame-acyclic molecules is sufficient to obtain freeness of the $\omega$-category presented by an oriented graded poset. Unfortunately, beyond low dimensions where it holds automatically, this property is difficult to check, since one needs to verify acyclicity of one graph *for each molecule* over an oriented graded poset. Its stability properties are also unclear. However, it is implied by stronger acyclicity conditions which are more restrictive, but easier both to check and to work with in practice.

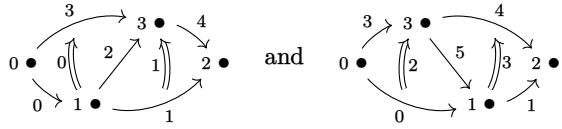
In this section, we start by considering a property we call *dimension-wise acyclicity*. This is of interest because it corresponds in a precise sense to the loop-freeness property of *augmented directed chain complexes* considered in [Ste04]; see also [AM20]. This will allow us to make a precise connection between our framework and what has come to be known as *Steiner theory*.

***4.1*** (Dimension-wise acyclic oriented graded poset). Let $P$ be an oriented graded poset. We say that $P$ is *dimension-wise acyclic* if, for all $k \geq -1$, $\mathscr{F}_k P$ is acyclic.

***Proposition 4.2*** — *Let $U$ be a molecule. If $U$ is dimension-wise acyclic, then $U$ is frame-acyclic.*

*Proof.* Let $V \sqsubseteq U$ be a submolecule inclusion, $r \coloneqq \operatorname{frdim} V$. If $U$ is dimension-wise acyclic, then $\mathscr{F}_r U$ is acyclic. Then $\mathscr{F}_r V$ is the induced subgraph of $\mathscr{F}_r U$ on the vertices contained in $V$, while $\mathscr{M}_r V$ is an induced subgraph of $\mathscr{F}_r V$, and an induced subgraph of an acyclic graph is always acyclic. ∎

*Example 4.3* — This example, which is essentially [Ste93, Fig. 4], shows that the implication of Proposition 4.2 is strict. Let $U$ be a 3-dimensional atom whose input and output boundaries correspond to the pasting diagrams




respectively. Let $(n, k)$ denote the $n$-dimensional cell labelled $k$. Then $\mathscr{F}_0 U$ contains the cycle

$$(1,2) \to (1,5) \to (1,2),$$

so $U$ is not dimension-wise acyclic. However, since $U$ is 3-dimensional, it is frame-acyclic as a consequence of Theorem 3.5.

***Proposition 4.4*** — *Let $f \colon P \to Q$ be a local embedding of oriented graded posets. Then, for all $k \geq -1$, $f$ induces a homomorphism $\mathscr{F}_k f \colon \mathscr{F}_k P \to \mathscr{F}_k Q$.*

*Proof.* For all $x \in P$, the restriction $f|_{\mathrm{cl}\{x\}}$ is an inclusion. By Proposition 1.16, for all $\alpha \in \{+, -\}$, if $y \in \Delta^\alpha_k x$, then $f(y) \in \Delta^\alpha_k f(x)$, so if there is an edge between $x$ and $y$ in $\mathscr{F}_k P$, then there is an edge between $f(x)$ and $f(y)$ in $\mathscr{F}_k Q$. ∎

***Corollary 4.5*** — *Let $f \colon P \to Q$ be a local embedding of oriented graded posets. If $Q$ is dimension-wise acyclic, then so is $P$.*

***Proposition 4.6*** — *Let $P$ be a dimension-wise acyclic regular directed complex. Then $P$ has frame-acyclic molecules.*

*Proof.* Let $U$ be a molecule and $f \colon U \to P$ a morphism. By Proposition 1.43, $f$ is a local embedding, so by Corollary 4.5 $U$ is dimension-wise acyclic. It follows from Proposition 4.2 that $U$ is frame-acyclic. ∎

***Corollary 4.7*** — *Let $P$ be a dimension-wise acyclic regular directed complex. Then $Mol/P$ is a polygraph.*

***4.8*** (Augmented directed chain complex). An *augmented chain complex* $C$ is a chain complex of abelian groups in non-negative degree

$$\dots \xrightarrow{\mathrm{d}} C_n \xrightarrow{\mathrm{d}} C_{n-1} \xrightarrow{\mathrm{d}} \dots \xrightarrow{\mathrm{d}} C_1 \xrightarrow{\mathrm{d}} C_0$$

together with a homomorphism $\mathrm{e}\colon C_0 \to \mathbb{Z}$ satisfying $\mathrm{e} \circ \mathrm{d} = 0$. A *direction* on $C$ is a choice of a commutative submonoid $C_n^{\to}$ of $C_n$ for each $n \in \mathbb{N}$. An *augmented directed chain complex* is an augmented chain complex $C$ together with a direction on its underlying chain complex.

***4.9*** (Homomorphism of augmented directed chain complexes). Let $C, D$ be augmented directed chain complexes. A *homomorphism* $f\colon C \to D$ is a homomorphism of the underlying augmented chain complexes, that is, a sequence $(f_n\colon C_n \to D_n)_{n\in\mathbb{N}}$ of homomorphisms of abelian groups satisfying

$$\mathrm{d} \circ f_{n+1} = f_n \circ \mathrm{d}, \qquad e \circ f_0 = e,$$

which is compatible with directions in the sense that

$$f_n(C_n^{\to}) \subseteq D_n^{\to}$$

for all $n \in \mathbb{N}$. Augmented directed chain complexes with their homomorphisms form a category $\mathbf{DCh}^{+}$.

***4.10*** (Linearisation of a strict $\omega$-category). Let $X$ be a strict $\omega$-category. The *linearisation of* $X$ is the augmented directed chain complex $\lambda X$ whose underlying augmented chain complex is defined by

$$\lambda X_n \coloneqq \frac{\mathbb{Z}(\sigma_{\le n}X)}{\langle t \#_k u - t - u \mid t, u \in \sigma_{\le n}X,\ k < n\rangle}$$

for all $n \in \mathbb{N}$, where $\mathbb{Z}(\sigma_{\le n}X)$ denotes the free abelian group on the set of cells of the $n$-skeleton of $X$, with the homomorphisms determined by

$$\mathrm{d}\colon t \in \sigma_{\le n}X \mapsto (\partial^{+}_{n-1}t - \partial^{-}_{n-1}t), \qquad \mathrm{e}\colon t \in \sigma_{\le 0}X \mapsto 1$$

for each $n > 0$, and the direction defined by

$$\lambda X_n^{\to} \coloneqq \mathrm{Im}\,(\mathbb{N}(\sigma_{\le n}X) \hookrightarrow \mathbb{Z}(\sigma_{\le n}X) \to \lambda X_n)$$

for each $n \in \mathbb{N}$, where $\mathbb{Z}(\sigma_{\le n}X) \to \lambda X_n$ is the canonical quotient homomorphism. This extends to a functor $\omega\mathbf{Cat} \to \mathbf{DCh}^{+}$, see [Ste04, Definition 2.4].

***4.11*** (Globular table in an augmented directed chain complex). Let $C$ be an augmented directed chain complex. A *globular table in* $C$ is a double sequence

$$x \equiv (x_n^{\alpha})_{n\in\mathbb{N},\,\alpha\in\{+,-\}}$$

such that

1. $x_n^{\alpha} \in C_n^{\to}$ for all $n \in \mathbb{N}$ and $\alpha \in \{+,-\}$,
2. $\mathrm{d}x_n^{\alpha} = x_{n-1}^{+} - x_{n-1}^{-}$ for all $n > 0$ and $\alpha \in \{+,-\}$,
3. $\mathrm{e}x_0^{\alpha} = 1$ for all $\alpha \in \{+,-\}$,

4. there exists $m \in \mathbb{N}$ such that $x^\alpha_n = 0$ for all $n > m$ and $\alpha \in \{+, -\}$.

***4.12*** (Strict $\omega$-category of globular tables). Let $C$ be an augmented directed chain complex. The *strict $\omega$-category of globular tables in $C$* is the strict $\omega$-category $\nu C$ whose set of cells is $\{x \mid x \text{ is a globular table in } C\}$, with the boundary operators defined, for all $n \in \mathbb{N}$ and $\alpha \in \{+, -\}$, by

$$(\partial^\alpha_n x)^\beta_m \coloneqq \begin{cases} x^\beta_m & \text{if } m < n, \\ x^\alpha_n & \text{if } m = n, \\ 0 & \text{if } m > n, \end{cases}$$

and the $k$-composition operations defined, for all $k \in \mathbb{N}$ and $k$-composable pairs $x, y$ of globular tables, by

$$(x \#_k y)^\alpha_n \coloneqq x^\alpha_n - (\partial^+_k x)^\alpha_n + y^\alpha_n.$$

This extends to a functor $\nu \colon \mathbf{DCh}^+ \to \omega\mathbf{Cat}$, see [Ste04, Definition 2.8].

***Proposition 4.13*** — *The functor $\lambda$ is left adjoint to $\nu$.*

*Proof.* See [Ste04, Theorem 2.11]. ∎

***4.14*** (Augmented directed chain complex of an oriented thin graded poset). Let $P$ be an oriented graded poset such that $P_\perp$ is oriented thin. The *augmented directed chain complex of $P$*, denoted by $\vec{\mathbb{Z}}P$, is the augmented chain complex

$$\dots \xrightarrow{\mathrm{d}} \mathbb{Z}P_n \xrightarrow{\mathrm{d}} \mathbb{Z}P_{n-1} \xrightarrow{\mathrm{d}} \dots \xrightarrow{\mathrm{d}} \mathbb{Z}P_1 \xrightarrow{\mathrm{d}} \mathbb{Z}P_0$$

where $\mathbb{Z}P_n$ is the free abelian group on the set $P_n$ and, for each $n > 0$, the homomorphism $\mathrm{d} \colon \mathbb{Z}P_n \to \mathbb{Z}P_{n-1}$ is defined on the generators $x \in P_n$ by

$$x \mapsto \sum_{y \in \Delta^+ x} y - \sum_{y \in \Delta^- x} y, \tag{1}$$

together with the homomorphism $\mathrm{e} \colon \mathbb{Z}P_0 \to \mathbb{Z}$ defined on the generators $x \in P_0$ by $x \mapsto 1$ and the direction given by $\mathbb{Z}P^{\to}_n \coloneqq \mathbb{N}P_n$ for each $n \in \mathbb{N}$.

***Proposition 4.15*** — *Let $P$ be an oriented graded poset such that $P_\perp$ is oriented thin. Then $\vec{\mathbb{Z}}P$ is well-defined as an augmented directed chain complex. Moreover, if $f \colon P \to Q$ is a morphism of oriented graded posets such that $P_\perp$, $Q_\perp$ are oriented thin, then the sequence of homomorphisms*

$$\begin{aligned} \mathbb{Z}f_n \colon \mathbb{Z}P_n &\to \mathbb{Z}Q_n, \\ x \in P_n &\mapsto f(x) \in Q_n \end{aligned}$$

*is a homomorphism $\vec{\mathbb{Z}}f \colon \vec{\mathbb{Z}}P \to \vec{\mathbb{Z}}Q$ of augmented directed chain complexes.*

By Proposition 1.31, this gives us an assignment $\vec{\mathbb{Z}}\colon \mathbf{RDCpx} \to \mathbf{DCh}^+$ which is easily determined to be functorial. There are now two ways to get from regular directed complexes to augmented directed chain complexes: either directly via $\vec{\mathbb{Z}}$, or by first passing to the $\omega$-category of molecules, then applying Steiner's linearisation functor. Fortunately, the two coincide up to natural isomorphism.

***Proposition 4.16*** — *Let $P$ be a regular directed complex. Then the assignment, for each $n \in \mathbb{N}$,*

$$\begin{aligned}\varphi_n &\colon \mathbb{Z}P_n \to (\lambda Mol/_P)_n\,,\\ &\quad x \in P_n \mapsto [\mathrm{cl}\,\{x\} \hookrightarrow P] \in \sigma_{\leq n} Mol/_P\end{aligned}$$

*is a natural isomorphism between $\vec{\mathbb{Z}}P$ and $\lambda Mol/_P$.*

Dually, there are two ways to get from a regular directed complex to an $\omega$-category: either via the $\omega$-category of molecules, or applying the right adjoint $\nu$ after $\vec{\mathbb{Z}}$. These two do *not*, in general, coincide. The rest of this section is dedicated to showing that they do coincide when $P$ is dimension-wise acyclic, in which case $\vec{\mathbb{Z}}P$ is a *Steiner complex* in the sense of [AM20].

***4.17*** (Basis of an augmented directed chain complex). Let $C$ be an augmented directed chain complex. A *basis for $C$* is a sequence of subsets $(\mathscr{B}_n \subseteq C_n)_{n\in\mathbb{N}}$ such that, for all $n \in \mathbb{N}$, $C_n$ is isomorphic to $\mathbb{Z}\mathscr{B}_n$ and $C_n^{\to}$ to $\mathbb{N}\mathscr{B}_n$.

***4.18*** (Support of a chain). Let $C$ be an augmented directed chain complex with basis $(\mathscr{B}_n)_{n\in\mathbb{N}}$, $n \in \mathbb{N}$, and $x \equiv \sum_{b\in\mathscr{B}_n} x_b b \in C_n$. The *support of $x$* is the subset

$$\operatorname{supp} x \coloneqq \{b \in \mathscr{B}_n \mid x_b \neq 0\} \subseteq \mathscr{B}_n.$$

If $C$ has a basis, then for all $x$ there exist unique $x^+, x^- \in C_n^{\to}$ such that $x = x^+ - x^-$ and $\operatorname{supp} x^+ \cap \operatorname{supp} x^- = \varnothing$.

***4.19*** (Unital basis). Let $C$ be an augmented directed chain complex with basis $(\mathscr{B}_n)_{n\in\mathbb{N}}$. For all $n \in \mathbb{N}$ and $b \in \mathscr{B}_n$, let

$$\langle b\rangle_m^\alpha \coloneqq \begin{cases} 0 & \text{if } m > n,\\ b & \text{if } m = n,\\ (\mathrm{d}\langle b\rangle_{m+1}^\alpha)^\alpha & \text{if } m < n\end{cases}$$

for each $m \in \mathbb{N}$ and $\alpha \in \{+,-\}$, where the definition is obtained by downward recursion when $m \leq n$. We say that the basis $(\mathscr{B}_n)_{n\in\mathbb{N}}$ is *unital* if, for all $n \in \mathbb{N}$ and $b \in \mathscr{B}_n$,

$$\langle b\rangle \equiv (\langle b\rangle_m^\alpha)_{m\in\mathbb{N},\,\alpha\in\{+,-\}}$$

is a globular table, or, equivalently, if $\mathrm{e}\langle b\rangle_0^+ = \mathrm{e}\langle b\rangle_0^- = 1$.

***4.20*** (Flow graph of an augmented directed chain complex with basis). Let $C$ be an augmented directed chain complex with basis $(\mathscr{B}_n)_{n\in\mathbb{N}}$, $k \in \mathbb{N}$. The *$k$-flow graph of $C$* is the directed graph $\mathscr{F}_k C$ whose

- set of vertices is $\bigcup_{i>k} \mathscr{B}_i$, and
- set of edges is $\left\{(b,c) \mid \operatorname{supp} \langle b\rangle_k^+ \cap \operatorname{supp} \langle c\rangle_k^- \neq \varnothing\right\}$, where the source of $(b,c)$ is $b$ and the target is $c$.

***4.21*** (Steiner complex). A *Steiner complex* is an augmented directed chain complex $C$ with a unital basis such that, for all $k \in \mathbb{N}$, $\mathscr{F}_k C$ is acyclic. We let $\mathbf{DCh}^+_{St}$ denote the full subcategory of $\mathbf{DCh}^+$ on the Steiner complexes.

The following is the fundamental theorem of Steiner theory.

***Theorem 4.22*** — *The restriction of $\nu\colon \mathbf{DCh}^+ \to \omega\mathbf{Cat}$ to $\mathbf{DCh}^+_{St}$ is full and faithful. Moreover, if $C$ is a Steiner complex with basis $(\mathscr{B}_n)_{n\in\mathbb{N}}$, then $\nu C$ is a polygraph whose set of generating cells is*

$$\left\{\langle b\rangle \mid b \in \bigcup_{n\in\mathbb{N}} \mathscr{B}_n\right\}.$$

*Proof.* See [Ste04, Theorem 5.6 and Theorem 6.1]. ∎

We will need the following result, which we state without proof.

***Lemma 4.23*** — *Let $P$ be a regular directed complex and let $U \subseteq P$ be a molecule, $n \coloneqq \dim U > 0$. Then, in $\vec{\mathbb{Z}}P$,*

$$\mathrm{d}\left(\sum_{x\in U_n} x\right) = \sum_{y\in\Delta^+ U} y - \sum_{y\in\Delta^- U} y.$$

***Lemma 4.24*** — *Let $P$ be a regular directed complex, $x \in P$, $m \in \mathbb{N}$, and $\alpha \in \{+,-\}$. Then, in $\vec{\mathbb{Z}}P$,*

$$\langle x\rangle_m^\alpha = \sum_{y\in\Delta_m^\alpha x} y.$$

*Proof.* Let $n \coloneqq \dim x$, so $x \in P_n$. By definition, for $m > n$, $\langle x\rangle_m^\alpha = 0$, while $\Delta_m^\alpha x = \varnothing$, and the equality holds. For $m \leq n$, we proceed by downward recursion. If $m = n$, we have $\langle x\rangle_m^\alpha = x$, while $\Delta_m^\alpha x = \{x\}$, and the equality holds. Let $m < n$. Then

$$\mathrm{d}\langle x\rangle_{m+1}^\alpha = \mathrm{d}\left(\sum_{y\in\Delta_{m+1}^\alpha x} y\right)$$

by the inductive hypothesis, and $\Delta^\alpha_{m+1}x = (\partial^\alpha_{m+1}x)_{m+1}$. By Lemma 4.23 and globularity of $\mathrm{cl}\,\{x\}$, this is equal to

$$\sum_{y\in\Delta^+(\partial^\alpha_{m+1}x)} y \;-\; \sum_{y\in\Delta^-(\partial^\alpha_{m+1}x)} y \;=\; \sum_{y\in\Delta^+_m x} y \;-\; \sum_{y\in\Delta^-_m x} y\,,$$

hence by definition

$$\langle x\rangle^+_m = \sum_{y\in\Delta^+_m x} y\,, \qquad \langle x\rangle^-_m = \sum_{y\in\Delta^-_m x} y\,.$$

This completes the proof. ∎

***Corollary 4.25*** — *Let $P$ be a regular directed complex, $x \in P$, $n \in \mathbb{N}$, $\alpha \in \{+,-\}$. Then* $\operatorname{supp}\langle x\rangle^\alpha_n = \Delta^\alpha_n x$.

***Proposition 4.26*** — *Let $P$ be a regular directed complex. Then $(P_n)_{n\in\mathbb{N}}$ is a unital basis of $\vec{\mathbb{Z}}P$.*

*Proof.* Let $x \in P$. For all $\alpha \in \{+,-\}$, $\Delta^\alpha_0 x = \partial^\alpha_0 x = \{x^\alpha\}$ for a unique $x^\alpha \in P_0$, since the point is the only 0-dimensional molecule. Then, by Lemma 4.24, $\mathrm{e}\langle x\rangle^\alpha_0 = \mathrm{e}x^\alpha = 1$, so the basis $(P_n)_{n\in\mathbb{N}}$ is unital. ∎

***Lemma 4.27*** — *Let $P$ be a regular directed complex. Then $\mathscr{F}_k\vec{\mathbb{Z}}P$ is isomorphic to $\mathscr{F}_k P$.*

*Proof.* Immediate from Corollary 4.25. ∎

***Proposition 4.28*** — *Let $P$ be a dimension-wise acyclic regular directed complex. Then $\vec{\mathbb{Z}}P$ is a Steiner complex.*

*Proof.* Follows from Proposition 4.26 and Lemma 4.27. ∎

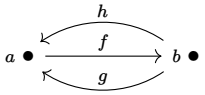
***Theorem 4.29*** — *Let $P$ be a dimension-wise acyclic regular directed complex. Then $\nu\vec{\mathbb{Z}}P$ is naturally isomorphic to $Mol/P$.*

*Proof.* Composing the component $\eta\colon Mol/P \to \nu\lambda Mol/P$ of the unit of the adjunction between $\lambda$ and $\nu$ with the natural isomorphism between $\lambda Mol/P$ and $\vec{\mathbb{Z}}P$ from Theorem 4.16, we obtain a strict functor

$$\varphi\colon Mol/P \to \nu\vec{\mathbb{Z}}P.$$

By Corollary 4.7, $Mol/P$ is a polygraph generated by $\{[\mathrm{cl}\,\{x\} \hookrightarrow P] \mid x \in P\}$, while by Theorem 4.22 combined with Proposition 4.28, $\nu\vec{\mathbb{Z}}P$ is a polygraph whose set of generating cells is $\{\langle x\rangle \mid x \in P\}$. By sending $[\mathrm{cl}\,\{x\} \hookrightarrow P]$ to $\langle x\rangle$, $\varphi$ determines a bijection between the generating cells of $Mol/P$ and of $\nu\vec{\mathbb{Z}}P$. By [ABG+23, Proposition 16.2.12], we conclude that $\varphi$ is an isomorphism of polygraphs. ∎

*Example 4.30* — Theorem 4.29 does not extend beyond dimension-wise acyclic regular directed complexes. Let $P$ be the regular directed complex encoding the 1-dimensional diagram




(2)

which is evidently not dimension-wise acyclic. Then $Mol/_{\vec{\mathbb{Z}}P}$ is isomorphic to the free category on the directed graph (2). However, in $\nu\vec{\mathbb{Z}}P$, let

$$x \coloneqq \langle f \rangle \#_0 \langle g \rangle, \qquad y \coloneqq \langle f \rangle \#_0 \langle h \rangle,$$

which as globular tables are defined, for all $\alpha \in \{+, -\}$, by

$$x^\alpha_n \coloneqq \begin{cases} a & \text{if } n = 0, \\ f + g & \text{if } n = 1, \\ 0 & \text{if } n > 1, \end{cases} \quad y^\alpha_n \coloneqq \begin{cases} a & \text{if } n = 0, \\ f + h & \text{if } n = 1, \\ 0 & \text{if } n > 1. \end{cases}$$

Then $x \#_0 y$ and $y \#_0 x$ are both equal to the globular table $z$ defined, for all $\alpha \in \{+, -\}$, by

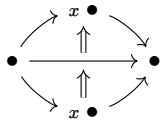
$$z^\alpha_n \coloneqq \begin{cases} a & \text{if } n = 0, \\ 2f + g + h & \text{if } n = 1, \\ 0 & \text{if } n > 1. \end{cases}$$

We conclude that $\nu\vec{\mathbb{Z}}P$ is not free, so it is not isomorphic to $Mol/_P$.

## 5. Stronger acyclicity conditions

While dimension-wise acyclicity is a more manageable sufficient condition for frame-acyclicity, it does not guarantee a second property that we considered in the introduction, that is, that the $\omega$-category of molecules over $P$ consists only of *subsets* of $P$.

*Example 5.1* — Let $U$ be the 2-dimensional molecule encoding the shape of the pasting diagram




and let $P$ be the result of identifying the two 0-dimensional cells marked with $x$. Then $P$ is a dimension-wise acyclic regular directed complex, and the canonical quotient map $q \colon U \to P$ is a molecule over $P$. However, $q$ is evidently not injective.

In this section, following [Ste93], we consider a strengthening of dimension-wise acyclicity which does guarantee this property at least for regular directed complexes. Then, we consider an even stronger acyclicity property, relying on acyclicity of a *single* directed graph, which, as we will see in Section 6, has better stability properties with respect to a number of constructions.

***5.2*** (Extended flow graph). Let $P$ be an oriented graded poset, $k \geq -1$. The *extended $k$-flow graph of $P$* is the bipartite directed graph $\overline{\mathscr{F}}_k P$ whose

- set of vertices is
$$P = \bigcup_{i \leq k} P_i + \bigcup_{i > k} P_i,$$
- set of edges is $E_- + E_+$, where
$$E_- \coloneqq \left\{ (y, x) \mid y \in \bigcup_{i \leq k} P_i, x \in \bigcup_{i > k} P_i, y \in \operatorname{int} \partial_k^- x \right\},$$
$$E_+ \coloneqq \left\{ (y, x) \mid y \in \bigcup_{i > k} P_i, x \in \bigcup_{i \leq k} P_i, x \in \operatorname{int} \partial_k^+ y \right\},$$
where the source of $(y, x)$ is $y$ and the target is $x$.

***5.3*** (Strongly dimension-wise acyclic oriented graded poset). Let $P$ be an oriented graded poset. We say that $P$ is *strongly dimension-wise acyclic* if, for all $k \geq -1$, $\overline{\mathscr{F}}_k P$ is acyclic.

*Comment 5.4* — Strong dimension-wise acyclicity is essentially the same as *loop-freeness* in the sense of [Ste93].

***Lemma 5.5*** — *Let $P$ be an oriented graded poset, $k \geq -1$, and suppose $x, y \in \bigcup_{i>k} P_i$. If there exists a path from $x$ to $y$ in $\mathscr{F}_k P$, then there exists a path from $x$ to $y$ in $\overline{\mathscr{F}}_k P$.*

*Proof.* Consider a path $x = x_0 \to x_1 \to \ldots \to x_m \to y$ from $x$ to $y$ in $\mathscr{F}_k P$. By definition of the $k$-flow graph, for all $i \in \{1, \ldots, m\}$, there exists $z_i \in \Delta_k^+ x_{i-1} \cap \Delta_k^- x_i$. By definition of the extended $k$-flow graph, there exist edges $x_{i-1} \to z_i$ and $z_i \to x_i$ in $\overline{\mathscr{F}}_k P$. Concatenating all the two-step paths $x_{i-1} \to z_i \to x_i$, we obtain a path from $x$ to $y$ in $\overline{\mathscr{F}}_k P$. ∎

***Proposition 5.6*** — *Let $P$ be a strongly dimension-wise acyclic oriented graded poset. Then $P$ is dimension-wise acyclic.*

*Proof.* By Lemma 5.5 a cycle in $\mathscr{F}_k U$ induces a cycle in $\overline{\mathscr{F}}_k U$. ∎

***Proposition 5.7*** — *Let $f \colon P \to Q$ be a local embedding of oriented graded posets. For all $k \geq -1$, $f$ induces a homomorphism $\overline{\mathscr{F}}_k f \colon \overline{\mathscr{F}}_k P \to \overline{\mathscr{F}}_k Q$.*

*Proof.* Similar to the proof of Proposition 4.4, using the fact that, by Proposition 1.16, inclusions preserve boundaries, hence they also preserve interiors. ∎

***Corollary 5.8*** — *Let $f\colon P \to Q$ be a local embedding of oriented graded posets. If $Q$ is strongly dimension-wise acyclic, then so is $P$.*

***Lemma 5.9*** — *Let $U$ be a frame-acyclic molecule, $x, y \in U$. Then there exists $k \geq -1$ such that there is a path from $x$ to $y$ or a path from $y$ to $x$ in $\overline{\mathscr{F}}_k U$.*

*Proof.* See [Ste93, Theorem 2.16], which applies by Remark 3.13. ∎

***Proposition 5.10*** — *Let $U$ be a molecule, $P$ a strongly dimension-wise acyclic oriented graded poset, and $f\colon U \to P$ a local embedding. Then $f$ is an inclusion.*

*Proof.* Let $x, y \in U$ and suppose that $f(x) = f(y)$. By Corollary 4.5, $U$ is strongly dimension-wise acyclic. It follows from Proposition 5.6 and 4.2 that $U$ is frame acyclic, so by Lemma 5.9 there exists $k \geq -1$ such that there is a path from $x$ to $y$ or a path from $y$ to $x$ in $\overline{\mathscr{F}}_k U$. Then by Proposition 5.7 $\overline{\mathscr{F}}_k f$ maps this onto a cycle in $\overline{\mathscr{F}}_k P$, a contradiction, unless $x = y$ and the path is constant. We conclude that $f$ is injective. ∎

***Corollary 5.11*** — *Let $P$ be a strongly dimension-wise acyclic regular directed complex. Then*

$$Mol/_P = \{[U \hookrightarrow P] \mid U \subseteq P, U \text{ is a molecule}\}.$$

*Proof.* Follows from Proposition 5.10 together with Proposition 1.43. ∎

*Remark 5.12* — In particular, if $P$ is finite and strongly dimension-wise acyclic, it follows that $Mol/_P$ has finitely many cells.

***5.13*** (Acyclic oriented graded poset). Let $P$ be an oriented graded poset. We say that $P$ is *acyclic* if $\vec{\mathscr{H}} P$ is acyclic.

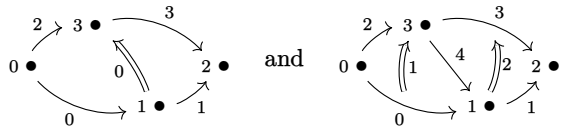
*Comment 5.14* — Acyclicity is essentially the same as *total loop-freeness* in [Ste93]. As we will see, it is also related to *strong loop-freeness* in [Ste04].

***Proposition 5.15*** — *Let $P$ be an acyclic regular directed complex, $x, y \in P$, and $k \geq -1$. If there is a path from $x$ to $y$ in $\overline{\mathscr{F}}_k P$, then there is a path from $x$ to $y$ in $\vec{\mathscr{H}} P$. Consequently, $P$ is strongly dimension-wise acyclic.*

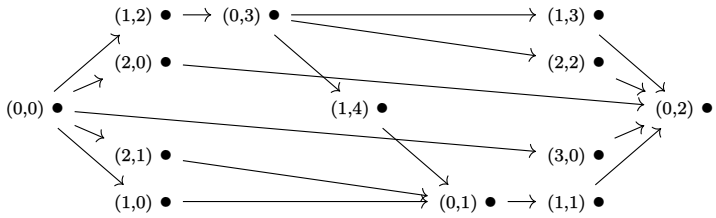
*Proof.* See [Ste93, Proposition 2.15 and Proposition 5.2]. ∎

*Remark 5.16* — In fact, the following, stronger fact holds: in *every* regular directed complex $P$, if there is a path from $x$ to $y$ in $\overline{\mathscr{F}}_k P$, then there is a path from $x$ to $y$ in $\vec{\mathscr{H}} P$. However, the proof is quite lengthy and technical and does not add much for our purposes.

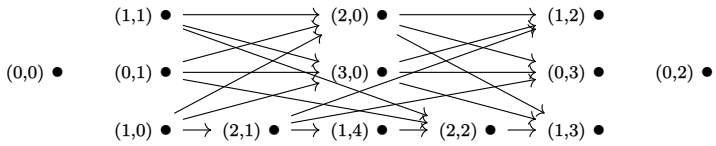
*Example 5.17* — Let $U$ be a 3-dimensional atom whose input and output boundaries encode the pasting diagrams



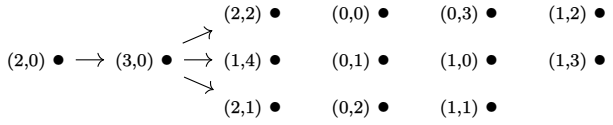

respectively, and let $(n, k)$ denote the $n$-dimensional cell labelled with $k$. Then the extended 0-flow graph $\overline{\mathscr{F}}_0 U$ is




while the extended 1-flow graph $\overline{\mathscr{F}}_1 U$ is




and the extended 2-flow graph $\overline{\mathscr{F}}_2 U$ is




all of which are acyclic. All other extended flow graphs are discrete, so $U$ is strongly dimension-wise acyclic. However, $\vec{\mathscr{H}} U$ contains the cycle

$$(0,1) \to (1,1) \to (2,0) \to (3,0) \to (2,1) \to (1,4) \to (0,1)$$

so $U$ is not acyclic.

***Proposition 5.18*** — *Let $f \colon P \to Q$ be a morphism of oriented graded posets. Then $f$ induces a homomorphism of directed graphs $\vec{\mathscr{H}} f \colon \vec{\mathscr{H}} P \to \vec{\mathscr{H}} Q$.*

*Proof.* Let $x, y \in P$ and suppose there is an edge from $x$ to $y$ in $\vec{\mathscr{H}} P$. Then either $x \in \Delta^- y$, hence $f(x) \in \Delta^- f(y)$, or $y \in \Delta^+ x$, hence $f(y) \in \Delta^+ f(x)$. In either case there is an edge from $f(x)$ to $f(y)$ in $\vec{\mathscr{H}} Q$. ∎

***Lemma 5.19*** — *Let $f\colon P \to Q$ be a morphism of oriented graded posets. If $Q$ is acyclic, then $P$ is acyclic.*

*Proof.* Suppose that there is a cycle in $\vec{\mathscr{H}}P$. By Proposition 5.18, $\vec{\mathscr{H}}f$ maps it onto a cycle in $\vec{\mathscr{H}}Q$. ∎

***Proposition 5.20*** — *Let $P$ be an acyclic oriented graded poset. Then $P$ has frame-acyclic molecules.*

*Proof.* Let $U$ be a molecule and $f\colon U \to P$ be a morphism. By Lemma 5.19, $U$ is an acyclic regular directed complex, so by Proposition 5.15, Proposition 5.6, and Proposition 4.6, it is frame-acyclic. ∎

***Corollary 5.21*** — *Let $U$ be an acyclic molecule, $x, y \in U$. Then there exists a path from $x$ to $y$ or from $y$ to $x$ in $\vec{\mathscr{H}}U$.*

*Proof.* Follows from Proposition 5.20 in combination with Lemma 5.9 and Proposition 5.15. ∎

***Corollary 5.22*** — *Let $P$ be an acyclic oriented graded poset. Then $Mol/P$ is a polygraph.*

***Proposition 5.23*** — *Let $U$ be a molecule, $P$ an acyclic oriented graded poset, and $f\colon U \to P$ a morphism. Then $f$ is an inclusion.*

*Proof.* Let $x, y \in U$ and suppose that $f(x) = f(y)$. By Corollary 5.21, there is a path from $x$ to $y$ or a path from $y$ to $x$ in $\vec{\mathscr{H}}U$. Then $\vec{\mathscr{H}}f$ maps this onto a cycle in $\vec{\mathscr{H}}P$, a contradiction, unless $x = y$ and the path is constant. We conclude that $f$ is injective. ∎

***Corollary 5.24*** — *Let $P$ be an acyclic oriented graded poset. Then*

$$\begin{aligned} Mol/P &= \{[U \hookrightarrow P] \mid U \subseteq P, U \text{ is a molecule}\}, \\ Atom/P &= \{[\mathrm{cl}\,\{x\} \hookrightarrow P] \mid x \in P, \mathrm{cl}\,\{x\} \text{ is an atom}\}. \end{aligned}$$

*Proof.* By Proposition 5.23, every morphism from a molecule to $P$ is an inclusion, equivalent to a subset inclusion $U \hookrightarrow P$ for some closed subset $U \subseteq P$. In particular, every morphism from an atom to $P$ is equivalent to the inclusion $\mathrm{cl}\,\{x\} \hookrightarrow P$ for some $x \in P$. ∎

*Remark 5.25* — Corollary 5.24 also implies that $Mol/P$ has finitely many cells as soon as $P$ is finite.

*Remark 5.26* — Observe that Corollary 5.24 does not require $P$ to be a regular directed complex, unlike Corollary 5.11.

We conclude by showing that acyclic regular directed complexes determine *strong Steiner complexes* in the sense of [AM20].

***5.27*** (Oriented Hasse diagram of an augmented directed chain complex with basis). Let $C$ be an augmented directed chain complex with basis $(\mathscr{B}_n)_{n\in\mathbb{N}}$. The *oriented Hasse diagram of* $C$ is the directed graph $\vec{\mathscr{H}}C$ whose

- set of vertices is $\bigcup_{n\in\mathbb{N}} \mathscr{B}_n$,
- set of edges is $\{(b,c) \mid b \in \operatorname{supp}(\mathrm{d}c)^- \text{ or } c \in \operatorname{supp}(\mathrm{d}b)^+\}$, where the source of $(b,c)$ is $b$ and the target is $c$.

***5.28*** (Strong Steiner complex). A *strong Steiner complex* is an augmented directed chain complex $C$ with a unital basis such that $\vec{\mathscr{H}}C$ is acyclic. We let $\mathbf{DCh}^+_{sSt}$ denote the full subcategory of $\mathbf{DCh}^+$ on strong Steiner complexes.

***Lemma 5.29*** — *Let $P$ be an oriented graded poset such that $P_\perp$ is oriented thin. Then $\vec{\mathscr{H}}\vec{\mathbb{Z}}P$ is isomorphic to $\vec{\mathscr{H}}P$.*

*Proof.* By construction, for all $x \in P$,

$$\operatorname{supp}(\mathrm{d}x)^+ = \Delta^+ x, \qquad \operatorname{supp}(\mathrm{d}x)^- = \Delta^- x,$$

so the definitions of $\vec{\mathscr{H}}\vec{\mathbb{Z}}P$ and of $\vec{\mathscr{H}}P$ coincide. ∎

***Proposition 5.30*** — *Let $P$ be an acyclic regular directed complex. Then $\vec{\mathbb{Z}}P$ is a strong Steiner complex.*

*Proof.* Follows from Proposition 4.26 and Lemma 5.29. ∎

## 6. Stability under constructions and operations

In this section, we consider some operations under which the classes of molecules and regular directed complexes are closed — pastings, suspensions, Gray products, joins, and duals — and study the stability of acyclicity conditions under these operations.

***6.1*** (Suspension of an oriented graded poset). Let $P$ be an oriented graded poset. The *suspension of* $P$ is the oriented graded poset $\mathsf{S}P$ whose

- underlying set is $\{\mathsf{S}x \mid x \in P\} + \{\perp^+, \perp^-\}$,
- order and orientation are defined, for all $x \in \mathsf{S}P$ and $\alpha \in \{+,-\}$, by

$$\nabla^\alpha x \coloneqq \begin{cases} \{\mathsf{S}y \mid y \in \nabla^\alpha x'\} & \text{if } x = \mathsf{S}x',\ x' \in P, \\ \{\mathsf{S}y \mid y \in P_0\} & \text{if } x = \perp^\alpha, \\ \varnothing & \text{if } x = \perp^{-\alpha}. \end{cases}$$

***6.2*** (Gray product of oriented graded posets). Let $P$, $Q$ be oriented graded posets. The *Gray product* of $P$ and $Q$ is the oriented graded poset $P \otimes Q$ whose

- underlying graded poset is the product $P \times Q$ of the underlying posets,
- orientation is defined, for all $(x, y) \in P \times Q$ and all $\alpha \in \{+, -\}$, by $\Delta^\alpha(x, y) \coloneqq \Delta^\alpha x \times \{y\} + \{x\} \times \Delta^{(-)^{\dim x}\alpha} y$.

Gray products determine a monoidal structure $(\mathbf{ogPos}, \otimes, 1)$ on $\mathbf{ogPos}$.

The monoidal structure $(\mathbf{ogPos}, \otimes, 1)$ restricts to $\mathbf{ogPos}^+$, and through the equivalence $(-)_{\not\perp}$ induces a different monoidal structure on $\mathbf{ogPos}$.

***6.3*** (Join of oriented graded posets). Let $P$, $Q$ be oriented graded posets. The *join of $P$ and $Q$* is the oriented graded poset $P \star Q \coloneqq (P_\perp \otimes Q_\perp)_{\not\perp}$. Joins determine a monoidal structure $(\mathbf{ogPos}, \star, \varnothing)$ on $\mathbf{ogPos}$.

***6.4*** (Duals of an oriented graded poset). Let $P$ be an oriented graded poset, $J \subseteq \mathbb{N} \setminus \{0\}$. The *$J$-dual of $P$* is the oriented graded poset $\mathsf{D}_J P$ whose

- underlying set is $\{\mathsf{D}_J x \mid x \in P\}$,
- partial order and orientation are defined by

$$\Delta^\alpha \mathsf{D}_J x \coloneqq \begin{cases} \{\mathsf{D}_J y \mid y \in \Delta^{-\alpha} x\} & \text{if } \dim x \in J, \\ \{\mathsf{D}_J y \mid y \in \Delta^\alpha x\} & \text{if } \dim x \notin J \end{cases}$$

  for all $x \in P$ and $\alpha \in \{+, -\}$.

When $J = \mathbb{N} \setminus \{0\}$, we write $P^\circ$ for $\mathsf{D}_J P$, and call it the *total dual of $P$*.

The following collects a number of non-trivial results of [Had24, Chapter 7].

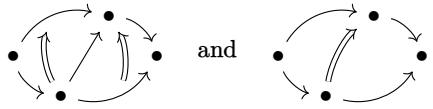
***Proposition 6.5*** — *Both the classes of molecules and of regular directed complexes are closed under suspensions, Gray products, joins, and all duals.*

We now move on to considering the stability of our acyclicity conditions.

***Proposition 6.6*** — *Let $U, V$ be molecules and $k \in \mathbb{N}$ such that $U \#_k V$ is defined. If $U$ and $V$ are acyclic, then $U \#_k V$ is acyclic.*

*Proof.* See the proof of [Ste93, Theorem 2.18]. ∎

*Example 6.7* — We show that Proposition 6.6 does not extend to weaker acyclicity conditions. Let $V$ be a 3-dimensional atom whose input and output boundaries correspond to the pasting diagrams




respectively, and let $U$ be the 3-dimensional atom from Example 5.17. Then both $V$ and $U$ are strongly dimension-wise acyclic. However, the boundary of the pasting $V \#_2 U$ is isomorphic to the boundary of the 3-dimensional atom from Example 4.3, which contains a cycle in its 0-flow graph. We conclude that $V \#_2 U$ is not dimension-wise acyclic.

The following lemma has a straightforward proof.

***Lemma 6.8*** — *Let $P$ be an oriented graded poset, $k \in \mathbb{N}$. Then*

1. *$x \mapsto \mathsf{S}x$ induces an isomorphism of directed graphs $\mathscr{F}_k P \xrightarrow{\sim} \mathscr{F}_{k+1}\mathsf{S}P$, restricting to an isomorphism $\mathscr{M}_k P \xrightarrow{\sim} \mathscr{M}_{k+1}\mathsf{S}P$;*
2. *$x \mapsto \mathsf{S}x$ induces an embedding of directed graphs $\overline{\mathscr{F}}_k P \hookrightarrow \overline{\mathscr{F}}_{k+1}\mathsf{S}P$, whose complement is the discrete graph on $\{\bot^-, \bot^+\}$.*

***Proposition 6.9*** — *Let $P$ be an oriented graded poset. If $P$ is acyclic (strongly dimension-wise acyclic, dimension-wise acyclic), then so is $\mathsf{S}P$.*

*Proof.* The strongly dimension-wise acyclic and dimension-wise acyclic cases follow from Lemma 6.8, together with the observation that $\mathscr{F}_0\mathsf{S}P$ is always a discrete graph, and that $\overline{\mathscr{F}}_0\mathsf{S}P$ has

- an edge from $\bot^-$ to every element of the form $\mathsf{S}x$,
- an edge from every element of the form $\mathsf{S}x$ to $\bot^+$,

and no other edges. The acyclic case is part of [Ste93, Theorem 2.19]. ∎

Next, we prove that frame-acyclicity is also preserved under suspension.

***Lemma 6.10*** — *Let $U$ be a frame-acyclic molecule. Then $\mathsf{S}U$ is frame-acyclic.*

*Proof.* The following facts are straightforward: the submolecules of $\mathsf{S}U$ are either $\{\bot^+\}$, $\{\bot^-\}$, or $\mathsf{S}V$ for $V \sqsubseteq U$, and

$$\operatorname{frdim} \mathsf{S}V = \begin{cases} -1 & \text{if } V \text{ is an atom,} \\ \operatorname{frdim} V + 1 & \text{otherwise.} \end{cases}$$

Given a submolecule $V' \sqsubseteq \mathsf{S}U$, then, either $V'$ is an atom, in which case $\mathscr{M}_{-1}V'$ is trivially acyclic, or $V' = \mathsf{S}V$ for $V \sqsubseteq U$ with $\operatorname{frdim} V = r \geq 0$, in which case $\operatorname{frdim} V' = r + 1$ and, by Lemma 6.8, $\mathscr{M}_{r+1}V'$ is isomorphic to $\mathscr{M}_r V$, which is acyclic by assumption. ∎

***Proposition 6.11*** — *Let $P$ be a regular directed complex with frame-acyclic molecules. Then $\mathsf{S}P$ has frame-acyclic molecules.*

*Proof.* By Lemma 6.10, it suffices to show that every molecule over $\mathsf{S}P$ is, up to isomorphism, either $\{\bot^\alpha\} \hookrightarrow P$ or of the form $\mathsf{S}f \colon \mathsf{S}U \to \mathsf{S}P$ for some molecule $U$ and morphism $f \colon U \to P$. Let $f' \colon U' \to \mathsf{S}P$ be a molecule over $\mathsf{S}P$; we can proceed by induction on submolecules. If $U'$ is an atom, by Proposition 1.43 $f'$ is isomorphic to the inclusion $\operatorname{cl}\{x\} \hookrightarrow \mathsf{S}P$ for some $x \in \mathsf{S}P$, which is either $\{\bot^\alpha\} \hookrightarrow \mathsf{S}P$ or $\operatorname{cl}\{\mathsf{S}x'\} \hookrightarrow \mathsf{S}P$ for some $x' \in P$, and the latter is isomorphic to $\mathsf{S}\operatorname{cl}\{x'\} \hookrightarrow \mathsf{S}P$. Otherwise, $f'$ is isomorphic to $g' \#_k h' \colon V' \#_k W' \to \mathsf{S}P$ with $k < \min\{\dim V', \dim W'\}$. Then $V'$ and $W'$ are not 0-dimensional, so by the inductive hypothesis $g' = \mathsf{S}g$ and $h' = \mathsf{S}h$ for

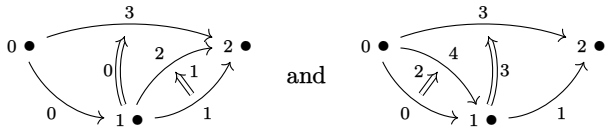
some $g\colon V \to P$ and $h\colon W \to P$. Moreover, for all $\alpha \in \{+,-\}$, necessarily $\partial^\alpha_0 g' = \partial^\alpha_0 h' = (\{\bot^\alpha\} \hookrightarrow \mathsf{S}P)$, so $g'$ and $h'$ cannot be 0-composable, and $k > 0$. Then $g$ and $h$ are $(k-1)$-composable and $g' \#_k h'$ is equal to $\mathsf{S}(g \#_{k-1} h)$ up to isomorphism. ∎

***Proposition 6.12*** — *Let $P$, $Q$ be acyclic oriented graded posets. Then $P \otimes Q$ and $P \star Q$ are acyclic.*

*Proof.* This is a part of [Ste93, Theorem 2.19]. ∎

*Comment 6.13* — This, in conjunction with the results of [AM20] and Theorem 4.29, can be used to show that $Mol/_{\_}$ is compatible with Gray products and joins of strict $\omega$-categories when restricted to acyclic regular directed complexes.

*Example 6.14* — We show that strongly dimension-wise acyclic and dimension-wise acyclic molecules are not closed under Gray products. Let $U$ be a 3-dimensional atom whose input and output boundary correspond to the pasting diagrams




respectively. Then $U$ is strongly dimension-wise acyclic. However, in $U \otimes U$, writing $x \otimes y$ instead of $(x, y)$ for better readability, we have

$$
\begin{aligned}
&(0,1) \otimes (2,2) \in \Delta^+((0,1) \otimes (3,0)) \cap \Delta^-((1,1) \otimes (2,2)),\\
&(1,1) \otimes (1,0) \in \Delta^+((1,1) \otimes (2,2)) \cap \Delta^-((2,1) \otimes (1,0)),\\
&(2,1) \otimes (0,1) \in \Delta^+((2,1) \otimes (1,0)) \cap \Delta^-((3,0) \otimes (0,1)),\\
&(2,2) \otimes (0,1) \in \Delta^+((3,0) \otimes (0,1)) \cap \Delta^-((2,2) \otimes (1,2)),\\
&(1,4) \otimes (1,2) \in \Delta^+((2,2) \otimes (1,2)) \cap \Delta^-((1,4) \otimes (2,1)),\\
&(0,1) \otimes (2,1) \in \Delta^+((1,4) \otimes (2,1)) \cap \Delta^-((0,1) \otimes (3,0)).
\end{aligned}
$$

These relations determine a cycle in $\mathscr{F}_2(U \otimes U)$. This proves that $U \otimes U$ is not dimension-wise acyclic.

***6.15*** (Converse of a directed graph). Let $\mathscr{G}$ be a directed graph. The *converse of $\mathscr{G}$* is the directed graph $\mathscr{G}^\circ$ with

- the same sets of vertices and edges as $\mathscr{G}$,
- source and target functions swapped with respect to $\mathscr{G}$.

***Lemma 6.16*** — *Let $P$ be an oriented graded poset, $J \subseteq \mathbb{N} \setminus \{0\}$, $k \geq -1$, and consider the bijection $\mathsf{D}_J\colon x \mapsto \mathsf{D}_J x$ between the underlying sets of $P$ and $\mathsf{D}_J P$. Then*

1. *if $k+1 \in J$, then $\mathsf{D}_J$ induces isomorphisms of directed graphs*

$$(\mathscr{M}_k P)^\circ \xrightarrow{\sim} \mathscr{M}_k \mathsf{D}_J P, \quad (\mathscr{F}_k P)^\circ \xrightarrow{\sim} \mathscr{F}_k \mathsf{D}_J P, \quad (\overline{\mathscr{F}}_k P)^\circ \xrightarrow{\sim} \overline{\mathscr{F}}_k \mathsf{D}_J P,$$

2. *if $k+1 \notin J$, then $\mathsf{D}_J$ induces isomorphisms of directed graphs*

$$\mathscr{M}_k P \xrightarrow{\sim} \mathscr{M}_k \mathsf{D}_J P, \quad \mathscr{F}_k P \xrightarrow{\sim} \mathscr{F}_k \mathsf{D}_J P, \qquad \overline{\mathscr{F}}_k P \xrightarrow{\sim} \overline{\mathscr{F}}_k \mathsf{D}_J P.$$

***Proposition 6.17*** — *Let $P$ be an oriented graded poset, $J \subseteq \mathbb{N} \setminus \{0\}$. Then*

1. *if $P$ is frame-acyclic, then so is $\mathsf{D}_J P$,*
2. *if $P$ is dimension-wise acyclic, then so is $\mathsf{D}_J P$,*
3. *if $P$ is strongly dimension-wise acyclic, then so is $\mathsf{D}_J P$.*

*Proof.* Follows from Lemma 6.16, combined with the fact that a directed graph is acyclic if and only if its converse is acyclic. ∎

*Example 6.18* — We show that strongly dimension-wise acyclic and dimension-wise acyclic molecules are not closed under joins. Let $U$ be the same 3-dimensional atom as in Example 6.14. Since $U$ is strongly dimension-wise acyclic, by Proposition 6.17 so is its total dual $U^\circ$. Using the isomorphism between $(U \star U^\circ)_\perp$ and $U_\perp \otimes (U^\circ)_\perp$, since the total dual counteracts the orientation reversal on faces of the second factor due to dimensions being raised by 1, we see that the cycle in $\mathscr{F}_2(U \otimes U)$ maps to a cycle

$$\begin{aligned}(0,1) \star (3,0)^\circ \to (1,1) \star (2,2)^\circ \to (2,1) \star (1,0)^\circ \to (3,0) \star (0,1)^\circ \to \\ \to (2,2) \star (1,2)^\circ \to (1,4) \star (2,1)^\circ \to (0,1) \star (3,0)^\circ\end{aligned}$$

in $\mathscr{F}_3(U \star U^\circ)$. This proves that $U \star U^\circ$ is not dimension-wise acyclic.

***Lemma 6.19*** — *Let $P$ be an oriented graded poset. Then the bijection $x \mapsto x^\circ$ induces an isomorphism $(\vec{\mathscr{H}} P)^\circ \xrightarrow{\sim} \vec{\mathscr{H}}(P^\circ)$ of directed graphs.*

***Proposition 6.20*** — *Let $P$ be an acyclic oriented graded poset. Then $P^\circ$ is acyclic.*

*Proof.* Immediate from Lemma 6.19. ∎

*Example 6.21* — Let $U$ be the 3-dimensional atom of Example 5.17. Then $U$ is not acyclic, but $\mathsf{D}_{\{1\}} U$ is acyclic. Since every dual is involutive, we conclude that acyclicity is not in general stable under duals.

## References


[ABG+23] D. Ara, A. Burroni, Y. Guiraud, P. Malbos, F. Métayer, and S. Mimram. *Polygraphs: from rewriting to higher categories.* Online preprint arXiv:2312.00429, 2023.

[AM20] D. Ara and G. Maltsiniotis. *Joint et tranches pour les $\infty$-catégories strictes.* Société Mathématique de France, 2020.

[Bén67] J. Bénabou. Introduction to bicategories. In *Reports of the Midwest Category Seminar*, pages 1–77, Berlin, Heidelberg, 1967. Springer Berlin Heidelberg.

[Bur93] A. Burroni. Higher-dimensional word problems with applications to equational logic. *Theoretical Computer Science*, 115(1):43–62, 1993.

[For22] S. Forest. Unifying notions of pasting diagrams. *Higher Structures*, 6(1):1–79, 2022.

[Had20a] A. Hadzihasanovic. A combinatorial-topological shape category for polygraphs. *Applied Categorical Structures*, 28(3):419–476, 2020.

[Had20b] A. Hadzihasanovic. Diagrammatic sets and rewriting in weak higher categories. Online preprint arXiv:2007.14505, 2020.

[Had21] A. Hadzihasanovic. The smash product of monoidal theories. In *2021 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 1–13. IEEE, 2021.

[Had24] A. Hadzihasanovic. *Combinatorics of higher-categorical diagrams.* Online preprint arXiv:2404.07273, 2024.

[Hen18] S. Henry. Regular polygraphs and the Simpson conjecture. Online preprint arXiv:1807.02627, 2018.

[HK23] A. Hadzihasanovic and D. Kessler. Higher-dimensional subdiagram matching. In *2023 38th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 1–13. IEEE, 2023.

[Joh89] M. Johnson. The combinatorics of $n$-categorical pasting. *Journal of Pure and Applied Algebra*, 62(3):211–225, 1989.

[Mak05] M. Makkai. The word problem for computads. Available at `http://www.math.mcgill.ca/makkai`, 2005.

[Pow91] J. Power. An $n$-categorical pasting theorem. In *Lecture Notes in Mathematics*, pages 326–358. Springer Nature, 1991.

[Ste93] R. Steiner. The algebra of directed complexes. *Applied Categorical Structures*, 1(3):247–284, 1993.

[Ste04] R. Steiner. Omega-categories and chain complexes. *Homology, Homotopy and Applications*, 6(1):175–200, 2004.

[Str76] R. Street. Limits indexed by category-valued 2-functors. *Journal of Pure and Applied Algebra*, 8(2):149–181, 1976.

[Str91] R. Street. Parity complexes. *Cahiers de topologie et géométrie différentielle catégoriques*, 32(4):315–343, 1991.

# Appendix D

# Paper IV

F. Wiesner, Z. Chaoui, D. Kessler, A. Pappa, and M. Karvonen. *Why quantum state verification cannot be both efficient and secure: a categorical approach.* Online preprint arXiv:2411.04767. 2024

# Why quantum state verification cannot be both efficient and secure: a categorical approach

Fabian Wiesner[1], Ziad Chaoui[1], Diana Kessler[2], Anna Pappa[1], and Martti Karvonen[3]

[1] Technische Universität Berlin {f.wiesner,ziad.chaoui,anna.pappa}@tu-berlin.de
[2] Tallinn University of Technology diana-maria.kessler@taltech.ee
[3] University College London martti.karvonen@ucl.ac.uk

**Abstract.** The advantage of quantum protocols lies in the inherent properties of the shared quantum states. These states are sometimes provided by sources that are not trusted, and therefore need to be verified. Finding secure and efficient quantum state verification protocols remains a big challenge, and recent works illustrate trade-offs between efficiency and security for different groups of states in restricted settings. However, whether a universal trade-off exists for all quantum states and all verification strategies remains unknown. In this work, we instantiate the categorical composable cryptography framework to show a fundamental limit for quantum state verification for all cut-and-choose approaches used to verify arbitrary quantum states. Our findings show that the prevailing cut-and-choose techniques cannot lead to quantum state verification protocols that are both efficient and secure.

**Keywords:** Quantum state verification · Categorical cryptography · Security limitations.

## 1 Introduction

For much of cryptography's history, security has been assumed but not proven. Even today, we rely on protocols without proven security, which are rather based on intuitive arguments [12]. In the comparably young field of quantum cryptography, many protocols claim provable security under the assumption that the devices used in these protocols are trustworthy. While the protocols offer a real advantage in tackling modern cryptographic challenges [7,21], they often come with two caveats:

1. Quantum hardware is expensive and difficult to operate and maintain. This is particularly true for quantum computers or their main building blocks, such as implementations of entangling gates [22].
2. The devices might not be trustworthy. To assume otherwise might in fact be a very strong assumption; someone untrusted could be operating the device, or there could be a hardware-based attack that leaks important information, as was done in the past for quantum key distribution systems [15].

Interestingly, these two issues are connected. Indeed, one way to address the first issue is to delegate some complex tasks to other parties while making sure that these perform the tasks as requested. Quantum correlations provide a way to check that the operations and tasks at hand are executed correctly. In the most general case, this is done through a framework called 'Device-Independence' [1], where the parties involved in a protocol can verify that the operations performed are correct, without putting any trust on the hardware.
In this paper, we focus on one specific task: quantum state verification. In quantum state verification protocols, an untrusted source prepares quantum states and distributes them among the clients who are sometimes considered honest. If the source is honest, it always prepares the target state, i.e. the state the clients desire to hold, and the clients accept the result. However, if the source is dishonest, it might not always send the target state, and the clients should ideally reject it. By virtue of the no-cloning theorem, the clients cannot simply measure and then use the quantum states the source sent. Hence the need to verify that the quantum states are indeed correct. A typical way to verify quantum states is for the source to send several copies of the state, some of which are then measured by the clients. If enough measurements correspond to the expected the states,the clients are convinced that the source is honest. They can then use the states they did not measure for further tasks. Indeed, some protocols use quantum state verification as a subroutine, for example when the clients don't have the local resources to create the state or the network resources to distribute it [8]. This modular use of quantum state verification begs for composable security but until recently, it wasn't clear if a quantum state verification protocol could be composably secure, especially if the clients are not trusted. Then, in [27], the authors showed composable security for the protocol in [20] but only against a dishonest source. Following a different approach for the clients, the authors of [5] demonstrated that stand-alone security implies composable security for many target states.
In this work we provide a no-go result showing that a quantum state verification protocol cannot be composably secure and efficient at the same time. We use the novel framework of categorical composable cryptography [2, 3] to prove this result. The motivation for using this framework lies in its combination of rigor and flexibility. On the one hand, modeling quantum processes and protocols as morphisms in a category provides a precise, albeit abstract, machine model which, by design, prevents mistakes and hidden assumptions. On the other hand, the ability to define attack models in the framework formally but still freely allows for the flexibility to investigate more complex adversarial settings such as 'honest-but-curious' or notions of i.i.d.-restrictions. Although we use a rather general attack model for our result, we will see that by the nature of the actual attacks, one could use more restrictive attack models as well. However, we stress that we believe that our results could be proven with essentially the same proofs in other frameworks for composably secure quantum cryptography [16, 17, 25].

### 1.1 Our contribution and related work

Many protocols implement quantum state verification for different types of states, e.g. [18, 20, 23, 24]. However, all protocols suffer from the same efficiency vs. security

trade-off: a quantum state verification protocol cannot be secure and efficient. We investigate this trade-off in a general setting and find fundamental limitations for quantum state verification.

**Theorem 1.1 (Main result (informal)).** *Let $\pi$ be a protocol for quantum state verification with the following properties:*

- *the clients cannot prepare the target state, and*
- *if the clients output a state received from the source, they perform no map on it.*

*At least one of the following statements about $\pi$ with security parameter $\lambda$ is false:*

1. *$\pi$ rejects the target state with a probability negligible in $\lambda$.*
2. *If the source is dishonest, either the probability to accept or the distance to the target state is negligible in $\lambda$.*
3. *The number of rounds $N$ is polynomial in $\lambda$.*

*Moreover, we find with $\varepsilon_H$ being the distinguishability to the idealized process if the source is honest and $\varepsilon_D$ if it is not*

$$\varepsilon_H + \varepsilon_D \geq \begin{cases} 1/8\sqrt{N} \text{ if the target state is pure.} \\ 1/27N \text{ if the target state is mixed.} \end{cases}$$

This trade-off has been proven in other works before, albeit in more restrictive settings. In both [19] and [29], the authors showed that for a quantum verification protocol for pure target states with a fixed number of rounds, the worst-case infidelity ($1-Fid.$) scales with the inverse of number of rounds. Although, in [19] the authors argue that this is not a restriction, both assume that the clients perform single round tests, i.e. do not use collective measurements. However, our work differs from [19, 29] in many aspects. First, the assumptions differ: we do not consider a fixed number of rounds, we allow for collective measurements, and, very importantly, we derive a bound for mixed states as well. To illustrate the importance of the latter, consider quantum state verification for a pure target state $|\phi\rangle\langle\phi|$. One can then circumvent the results from [19, 29] by using a protocol with a target state $(1-f(N))|\phi\rangle\langle\phi| + f(N)|\phi^\perp\rangle\langle\phi^\perp|$, where $f$ can be even negligible in the number of rounds $N$. Our result closes such loopholes. Further, the perspectives on the topic are different. In [19, 29], the authors utilize the hypothesis testing framework[4], which is useful for quantum state verification but is not common in other areas of quantum cryptography. We argue that quantum state verification should be viewed as a building block of larger protocols and hence use categorical composable cryptography. Because of this difference, we developed a novel proof technique which we expect to also be adaptable to other settings.

Our results provide bounds for self-testing as well [30]. Self-testing is slightly different from quantum state verification, since there is a single client that does not trust any of their devices, including preparation and measurement apparatus. Self-testing can therefore be seen as a stricter case of quantum state verification. Hence, any attack on quantum state verification implies an attack on self-testing.

[4] See [28] for a review on quantum state verification focused on the hypothesis testing approach.

### 1.2 Structure

Our work is structured as follows. In section 2, we first present results from quantum information theory that we need for our security analysis. We then provide a gentle introduction to category theory and our category **CPTP** alongside morphisms used to express the algorithms in later sections. We conclude section 2 with the n-comb construction necessary to define the resource theories for the subsequent work. We outline the resource theory that we work with in section 3, guided by [2,3]. We then give a formal security definition and a description of our ideal resource before presenting our main results in section 4. In section 4, we present our no-go result first for a simple type of protocols and then for general quantum state verification protocols. In both cases, we prove the no-go result in the single- and multi-client cases. Finally, we discuss open questions and possible implications of our work in section 5.

## 2 Preliminaries

### 2.1 Quantum Information Theory

Before instantiating the categorical framework we need for our security analysis, we present some preliminaries on quantum information theory. All results in this section are taken from [26].
In the following, we write our definitions with respect to density operators and quantum channels, although they hold for general linear operators and linear maps. A density operator on a space $\mathcal{X}$ is a positive semidefinite operator with trace equal to one. $D(\mathcal{X})$ is the space of density operators. A quantum channel from $\mathcal{X}$ to $\mathcal{Y}$ is a completely positive trace preserving map from the space of linear operators on $\mathcal{X}$, $\mathcal{L}(\mathcal{X})$ to $\mathcal{L}(\mathcal{Y})$. $\mathcal{C}(\mathcal{X},\mathcal{Y})$ denotes the space of quantum channels from $\mathcal{X}$ to $\mathcal{Y}$.

**Definition 2.1 (Trace norm, diamond norm, diamond distance).** *For a density operator $\rho\in D(\mathcal{X})$, we define the trace norm to be $\|\rho\|_1=\mathrm{Tr}\left(\sqrt{\rho\rho^\dagger}\right)$. The induced trace norm of a quantum channel $\Phi\in\mathcal{C}(\mathcal{X},\mathcal{Y})$ is then $\|\Phi\|_1=\max\{|\Phi(\rho)|_1 : \rho\in D(\mathcal{X})\}$. The diamond norm of $\Phi$ is then defined as*

$$\|\Phi\|_\diamond=\|\Phi\otimes\mathrm{id}_{L(\mathcal{X})}\|_1.$$

*Finally we define the diamond distance between $\Phi,\Psi\in\mathcal{C}(\mathcal{X},\mathcal{Y})$ as*

$$d_\diamond(\Phi,\Psi)=\|\Phi-\Psi\|_\diamond. \tag{2.1}$$

Two properties of the diamond distance that hold for CPTP maps $\Phi_0,\Phi_1,\Psi_0,\Psi_1\in\mathcal{C}(\mathcal{X},\mathcal{Y})$ and any space $\mathcal{Z}$ are

$$d_\diamond(\Psi_1\Psi_0,\Phi_1\Phi_0)\leq d_\diamond(\Psi_1,\Phi_1)+d_\diamond(\Psi_0,\Phi_0), \tag{2.2}$$

$$d_\diamond\big(\Phi\otimes\mathrm{id}_{L(\mathcal{Z})},\Psi\otimes\mathrm{id}_{L(\mathcal{Z})}\big)\leq d_\diamond(\Phi,\Psi). \tag{2.3}$$

The distances above also have an operational interpretation. Indeed the trace distance yields a bound on the achievable distinguishing advantage between two density operators given by the Holevo-Helstrom Theorem.

**Theorem 2.2 (Holevo-Helstrom Theorem).** *Let $\rho_0, \rho_1 \in D(\mathcal{X})$ be density operators, and let $\lambda \in [0,1]$. For any measurement $\mu : \{0,1\} \to Pos(\mathcal{X})$ (where $Pos(\mathcal{X})$ denotes all positive definite operators on $\mathcal{X}$) it then holds*

$$\lambda\langle\mu(0)|\rho_0\rangle + (1-\lambda)\langle\mu(1)|\rho_1\rangle \leq \frac{1}{2} + \frac{1}{2}\|\lambda\rho_0 - (1-\lambda)\rho_1\|_1. \tag{2.4}$$

*Moreover there exists a projective measurement $\mu : \{0,1\} \to Pos(\mathcal{X})$ for which equality is achieved in (2.4).*

To see that this actually gives a bound on the distinguishing advantage we set $\lambda = \frac{1}{2}$ in (2.4) and we obtain

$$\frac{1}{2}\langle\mu(0)|\rho_0\rangle + \frac{1}{2}\langle\mu(1)|\rho_1\rangle \leq \frac{1}{2} + \frac{1}{4}\|\rho_0 - \rho_1\|_1 \tag{2.5}$$

$$\Leftrightarrow \langle\mu(1)|\rho_1\rangle + (\langle\mu(0)|\rho_0\rangle - 1) = \langle\mu(1)|\rho_1\rangle - \langle\mu(1)|\rho_0\rangle \leq \frac{1}{2}\|\rho_0 - \rho_1\|_1. \tag{2.6}$$

Using (2.6) with an adequate choice of measurement $\mu$ that satisfies $\langle\mu(0)|\rho_0\rangle \geq \frac{1}{2}$ and $\langle\mu(1)|\rho_1\rangle \geq \frac{1}{2}$, we find the distinguishing advantage.
Another important quantity is the fidelity. The fidelity between two density operators $\rho_0$, $\rho_1$ is given by

$$F(\rho_0, \rho_1) = \mathrm{Tr}\left(\sqrt{\sqrt{\rho_0}\rho_1\sqrt{\rho_0}}\right)^2.$$

The Fuchs-van de Graaf inequalities link the trace distance to the fidelity [26].

**Theorem 2.3 (Fuchs-van de Graaf Inequalities).** *Let $\rho_0, \rho_1 \in D(\mathcal{X})$ be density operators, it holds that*

$$1 - \sqrt{F(\rho_0,\rho_1)} \leq \frac{1}{2}\|\rho_0 - \rho_1\|_1 \leq \sqrt{1 - F(\rho_0,\rho_1)} \tag{2.7}$$

$$\left(1 - \frac{1}{2}\|\rho_0 - \rho_1\|_1\right)^2 \leq F(\rho_0,\rho_1) \leq 1 - \frac{1}{2}\|\rho_0 - \rho_1\|_1^2 \tag{2.8}$$

**Lemma 2.4 (Bounds for multi copy distinction).** *For $\rho_0, \rho_1 \in D(\mathcal{X})$ we have*

$$1 - \left(\sqrt{1 - \frac{1}{2}\|\rho_0 - \rho_1\|_1^2}\right)^n \leq \frac{1}{2}\|\rho_0^{\otimes n} - \rho_1^{\otimes n}\|_1 \leq \sqrt{1 - (1 - \|\rho_0 - \rho_1\|_1)^n}. \tag{2.9}$$

*Proof.* An important property of the fidelity function is

$$F(\rho_0^{\otimes n}, \rho_1^{\otimes n}) = F(\rho_0, \rho_1)^n. \tag{2.10}$$

Using this and the Fuchs-van de Graaf inequalities we first derive the left hand side of (2.9)

$$1 - \left(\sqrt{1 - \frac{1}{2}\|\rho_0 - \rho_1\|_1^2}\right)^n \overset{(2.7)}{\leq} 1 - \sqrt{F(\rho_0,\rho_1)}^n$$
$$\overset{(2.10)}{=} 1 - \sqrt{F(\rho_0^{\otimes n}, \rho_1^{\otimes n})} \overset{(2.7)}{\leq} \frac{1}{2}\|\rho_0^{\otimes n} - \rho_1^{\otimes n}\|_1.$$

For the right hand side we have

$$\begin{aligned}\frac{1}{2}\|\rho_0^{\otimes n}-\rho_1^{\otimes n}\|_1 &\overset{(2.7)}{\leq} \sqrt{1-F(\rho_0^{\otimes n},\rho_1^{\otimes n})}\\ \overset{(2.10)}{=} \sqrt{1-F(\rho_0,\rho_1)^n} &\overset{(2.8)}{\leq} \sqrt{1-(1-\frac{1}{2}\|\rho_0-\rho_1\|_1)^{2n}}\\ &\leq \sqrt{1-(1-\|\rho_0-\rho_1\|_1)^n}.\end{aligned}$$

□

For pure states however, it holds that

$$\||\psi\rangle\langle\psi|-|\phi\rangle\langle\phi|\|_1 = 2\sqrt{1-|\langle\psi|\phi\rangle|^2},$$

which implies

$$\frac{1}{2}\|\rho_0^{\otimes n}-\rho_1^{\otimes n}\|_1 = \sqrt{1-|\langle\psi|\phi\rangle|^{2n}}. \tag{2.11}$$

### 2.2 Category Theory

To introduce the framework of categorical composable cryptography, we need to introduce some notions about category theory. Informally, a category is a collection of objects - usually denoted $A,B,C,\ldots$ - and morphisms - $f,g,h,\ldots$ - between objects. Whenever we have two morphisms $f: A\to B$, $g: B\to C$ such that the domain of $g$ and the codomain of $f$ coincide, we can compose them to obtain the morphism $g\circ f: A\to C$. This composition operation is required to be associative and for every object $A$, there should exist a morphism $\mathrm{id}_A$ which acts as identity on morphism composition.

*Example 2.5.*

1. Sets and functions between them form a category, **Set**, in which the objects are sets and the morphisms are functions. Morphism composition is function composition and the identity morphism is the identity function, $f(x)=x$.
2. The category **FHilb**, is the category in which the objects are finite dimensional Hilbert spaces and the morphisms are linear transformations between them.
3. The category **Met** of extended pseudometric spaces has extended pseudometric spaces as its objects: these are pairs $(X,d)$ where $X$ is a set and $d\colon X\times X\to[0,\infty]$ satisfies the axioms of a pseudometric[5], except that we allow for points with infinite distance[6]. The morphisms in **Met** are given by the *short* (or distance non-increasing) maps, so that maps $(X,d)\to(Y,e)$ are given by functions $f\colon X\to Y$ satisfying $e(f(x),f(y))\leq d(x,y)$ for all $x,y\in X$.
4. Recall that a monoid is basically a group without inverses, i.e., a set $M$ equipped with a binary operation $\cdot\colon M\times M\to M$ that is associative and has a unit element. Any monoid $(M,\cdot)$ can be viewed as a category with one object $\bullet$, with the morphisms $\bullet\to\bullet$ given by elements of $M$ and composition given by $\cdot$.

[5] These are almost the axioms of a metric, except distinct points can have distance zero.

[6] This corresponds to the adjective "extended", and is mostly for mathematical convenience. This can be ignored as in the sequel as all metrics we use take finite values.

5. Any partially ordered set $(P,\leq)$ induces a category whose objects are given by the elements of $P$, and there exists a a unique morphism $x\to y$ iff $x\leq y$.

For any two objects in a category $\mathbf{C}$, we denote the set of all morphisms $A\to B$ by $\mathbf{C}(A,B)$.
In many categories of interest one can not only compose morphisms sequentially, but also in parallel. For instance in **Set**, given two morphisms $f\colon A\to B$ and $g\colon C\to D$, we can form the morphism $f\times g\colon A\times C\to B\times D$. This parallel composition is almost associative, commutative and has a unit. For instance, there is an obvious bijection relating the sets $A\times(B\times C)$ and $(A\times B)\times C$ that merely re-brackets the data. This idea is made precise by the notion of a *symmetric monoidal category* (SMC). We begin by introducing a stricter notion which is easier to define precisely although it fails to capture many examples of interest.

**Definition 2.6 (Symmetric strict monoidal category).**
*A strict monoidal category $(\mathbf{C},\otimes,I)$ is a category $\mathbf{C}$ equipped with an object $I$ called the monoidal unit and a monoidal product $\otimes$ sending a pair $(A,B)$ of objects to an object $A\otimes B$, and two morphisms $f\colon A\to B$ and $g\colon C\to D$ to a morphism $f\otimes g\colon A\otimes C\to B\otimes D$. The operation $\otimes$ must respect identity morphisms in that* $\mathrm{id}_A\otimes\mathrm{id}_B=\mathrm{id}_{A\otimes B}$. *Moreover, the operation $\otimes$ should satisfy the* interchange law, *which states that whenever $g\circ f$ and $i\circ h$ are defined, then*

$$(g\circ f)\otimes(i\circ h)=(g\otimes i)\circ(f\otimes h) \tag{2.12}$$

*Finally, the operation $\otimes$ should be associative and unital in that for all objects $A,B,C$ and morphisms $f,g,h$, we have:*

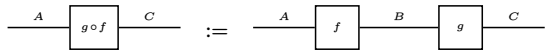
$$\begin{aligned}&(A\otimes B)\otimes C=A\otimes(B\otimes C),\\&I\otimes A=A\otimes I=A,\\&(f\otimes g)\otimes h=f\otimes(g\otimes h),\\&\mathrm{id}_I\otimes f=f\otimes\mathrm{id}_I=f.\end{aligned}$$

*A symmetric strict monoidal category is a strict monoidal category with chosen isomorphisms $\sigma_{A,B}\colon A\otimes B\to B\otimes A$ for all $A,B$ such that (i) $\sigma_{B,A}\circ\sigma_{A,B}=\mathrm{id}_{A\otimes B}$ (ii) the isomorphisms $\{\sigma_{A,B}\}_{A,B}$ are natural in the sense that $(g\otimes f)\circ\sigma=\sigma\circ(f\otimes g)$.*

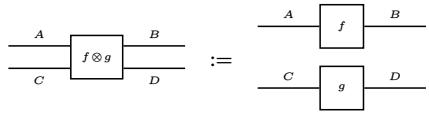
In the general (not necessarily strict) case, a monoidal category has an operation $\otimes$ as above, but the associativity and unit equations of it are replaced by isomorphisms (satisfying some further conditions), see [9] for more details.

*Example 2.7.* 1. The monoidal structure of **Set** is described as follows: the monoidal product is the cartesian product, $\times$, and the unit object is a chosen one element set: $\{\bullet\}$.
2. The monoidal structure of **FHilb** has the tensor product of Hilbert spaces as the monoidal product and the one-dimensional Hilbert space, $\mathbb{C}$, as the unit object.
3. We equip **Met** with a monoidal structure as follows: we define $(X,d)\otimes(Y,e)$ to be the set $X\times Y$ equipped with the $\ell^1$ distance, so that the distance between $(x,y)$ and $(x',y')$ in $X\otimes Y$ is given by the sum $d(x,x')+e(y,y')$.

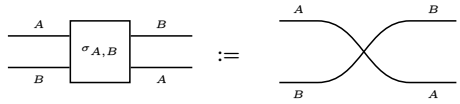
A nice feature of (symmetric) monoidal categories is that there is an intuitive yet precise graphical syntax for describing morphisms in them. We next introduce these *string diagrams* and our conventions for them. We draw string diagrams from left to right, just like quantum circuits, although the reader should be warned that different papers might have their string diagrams drawn from top to bottom or from bottom to top instead. The sequential composition looks like:



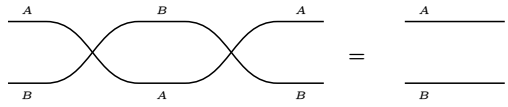

While the tensor composition is simply drawing two morphisms in parallel.



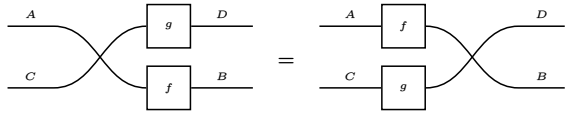

The symmetry isomorphisms are drawn as wire crossings :



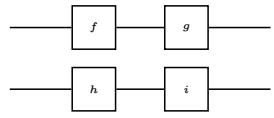

The string diagrams make the axioms intuitive: for instance, the condition $\sigma_{B,A} \circ \sigma_{A,B} = \mathrm{id}_{A \otimes B}$ becomes




so that two crossings undo each other, and the naturality condition for the symmetry can be pictured as us being allowed to slide the boxes corresponding to the morphisms $f$ and $g$ through the crossing:




Other axioms are used in the pictures implicitly: for instance, when drawing three parallel lines we don't add in any brackets, and the interchange law (2.12) guarantees that the following picture is unambiguous:



We will also need the definition of a bimonoidal (rig) category, which we state formally below. This is a "categorified" version of a rig/semiring ($\approx$ ring without negatives, like the natural numbers with addition and multiplication), just like a monoidal category is a categorified version of a monoid.

**Definition 2.8.** *A bimonoidal category is a category equipped with two monoidal products - a symmetric monoidal structure* $(\mathbf{C},\oplus,0)$ *and a monoidal structure* $(\mathbf{C},\otimes,I)$*- such that there exist distributivity isomorphisms*

$$d_l : A\otimes(B\oplus C)\to(A\otimes B)\oplus(A\otimes C),$$
$$d_r : (A\oplus B)\otimes C\to(A\otimes C)\oplus(B\otimes C),$$

*and absorption isomorphisms*

$$a_l : A\otimes 0\to 0,$$
$$a_r : 0\otimes A\to 0$$

*that satisfy some coherence laws [14].*

**Base Categorical modeling** We will model our quantum systems of interest as finite-dimensional $C^*$-algebras and our quantum processes as quantum channels i.e., CPTP maps between them. We will only sketch these informally and refer the reader to [13] for full details. A paradigmatic example of a $C^*$-algebra is given by the space of bounded operators on a Hilbert space. First of all this is a vector space over the complex numbers and composition of operators makes it into an algebra. Moreover, the operation of taking the adjoint equips this algebra with an involution, and the general notion of a $C^*$-algebra abstracts away from this by axiomatizing important interactions between these structures and the operator norm. It is standard that any $C^*$-algebra embeds into one of this form, much like any group embeds into a permutation group.
In the finite-dimensional case, a paradigmatic example of a $C^*$-algebra is given by $M_n(\mathbb{C})$, the $n\times n$ complex matrices. Any finite-dimensional $C^*$-algebra is isomorphic to a finite direct sum of such matrix algebras (see e.g. [6, Theorem III.I.1]), and hence can be captured by a list $[n_1,...n_k]$ of non-negative natural numbers specifying the dimension of each matrix algebra.
The main reason we work with general (but finite-dimensional) $C^*$ algebras is that they allow us to treat quantum and classical systems. For example, the state-space of a qubit is modeled by the $C^*$-algebra $M_2(\mathbb{C})$, whereas the state-space of a classical bit is modeled by $M_1(\mathbb{C})\oplus M_1(\mathbb{C})\cong\mathbb{C}\oplus\mathbb{C}$. The act of destructively measuring a qubit in the standard basis is then represented by the CPTP-map $M_2(\mathbb{C})\to\mathbb{C}\oplus\mathbb{C}$ acting by $\begin{pmatrix} a & b \\ c & d \end{pmatrix}\mapsto(a\ d)$, and a non-destructive measurement of a qubit could be modeled as a map $M_2(\mathbb{C})\to M_2(\mathbb{C})\otimes(\mathbb{C}\oplus\mathbb{C})$.

**Definition 2.9 (CPTP).** *The category* **CPTP** *of quantum channels is defined as follows: the objects are finite-dimensional* $C^*$*-algebras and the maps are completely positive trace preserving maps.*

The category of $C^*$-algebras admits two monoidal structures, $\oplus$ and $\otimes$ given by the natural direct sum and direct product of the underlying vector space. Thus, if $A \cong M_{n_1}(\mathbb{C}) \oplus ... \oplus M_{n_k}(\mathbb{C})$ and $B \cong M_{m_1}(\mathbb{C}) \oplus ... \oplus M_{m_p}(\mathbb{C})$, then

$$A \oplus B = M_{n_1}(\mathbb{C}) \oplus ... \oplus M_{n_k}(\mathbb{C}) \oplus M_{m_1}(\mathbb{C}) \oplus ... \oplus M_{m_p}(\mathbb{C})$$
$$A \otimes B = M_{n_1 m_1}(\mathbb{C}) \oplus M_{n_1 m_2}(\mathbb{C}) \oplus ... M_{n_1 m_p}(\mathbb{C}) \oplus M_{n_2 m_1}(\mathbb{C}) \oplus ... M_{n_k m_p}(\mathbb{C})$$

In shorthand, this can be represented as

$$[n_1,...,n_k] \oplus [m_1,...,m_p] = [n_1,...n_k,m_1,...,m_p]$$
$$[n_1,...,n_k] \otimes [m_1,...,m_p] = [n_1 m_1, n_1 m_2, ..., n_k m_p]$$

With respect to the $\oplus$ product, the 0-dimensional $C^*$-algebra is the unit object, while with respect to the $\otimes$ product, $\mathbb{C} \cong M_1(\mathbb{C})$ is the unit object.

**Lemma 2.10 (CPTP is bimonoidal).** *The category* **CPTP** *is bimonoidal with product operations $\oplus$ and $\otimes$.*

*Proof.* See [10, Definition 2.10].

**Pseudo-Code** To express algorithms in this category in a simple fashion, we introduce translations from pseudo-code to morphisms in **CPTP**. This translation relies on the bimonoidal structure of **CPTP**. We start with Branch-up and Branch-down – isomorphisms that essentially state that there are two ways of expressing classical distributions of quantum states.

**Definition 2.11 (Branch-up/ Branch-down).** *Let $A$ be an object in* **CPTP**. *Then, we define the isomorphism $B^{up}_{n,A}$ using the unitors and distributors as follows:*

$$B^{up}_{n,A} : I^{\oplus n} \otimes A \xrightarrow{\simeq} (I^{\oplus n-1} \otimes A) \oplus (I \otimes A) \xrightarrow{\simeq} (I^{\oplus n-1} \otimes A) \oplus A \xrightarrow{\simeq} ... \xrightarrow{\simeq} A^{\oplus n}.$$

*So, $B^{up}_{n,A} : I^{\oplus n} \otimes A \xrightarrow{\simeq} A^{\oplus n}$.*
*Symmetrically, we define $B^{down}_{n,A} : A^{\oplus n} \xrightarrow{\simeq} I^{\oplus n} \otimes A$.*

In particular, for algorithms in the context of verification, explicit branching is essential. We allow for branching using the following definition.

**Definition 2.12 (If-Else).** *Let $A,B,C$ be objects in* **CPTP** *and let $f : A \to B$, $g : A \to C$ be morphisms in* **CPTP**. *Then, $\mathsf{ife}(f,g) : (I \oplus I) \otimes A \to B \oplus C$ is a morphism (channel) defined as:*

$$\mathsf{ife}(f,g) = (f \oplus g) \circ B^{up}_{2,A}$$

*More generally, an if-else channel applied to $n$ arguments $f_1,...,f_n$ (and corresponding to an if-then-else structure with $(n-2)$ else-if structures) is a morphism $\mathsf{elif}((f_i)_{i=1}^n) : (I^{\oplus n} \otimes A \to \bigoplus_{i=1}^n B_i)$ defined as*

$$\mathsf{elif}((f_i)_{i=1}^n) = \left( \bigoplus_{i=1}^n f_i \right) \circ B^{up}_{n,A},$$

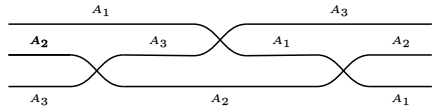
*where $f_i : A \to B_i$.*

Using the symmetry of the category allows for swapping registers. To denote this formally in an algorithm, we introduce the corresponding pseudo-code.

**Definition 2.13 (Swap).** *Given $A_1,...,A_n$, objects in* **CPTP**, *to define a map* $\mathsf{swap_{k,l}}: A_1 \otimes ... \otimes A_k \otimes ... \otimes A_l \otimes ... \otimes A_n \to A_1 \otimes ... \otimes A_l \otimes ... \otimes A_k \otimes ... \otimes A_n$ *we first let* $\Sigma_{k,k+1} = \bigotimes_{i=1}^{k-1} I \otimes \sigma_{k,k+1} \otimes \bigotimes_{i=k+2}^{n} I$. *Then,*

$$\mathsf{swap_{k,l}} = \Sigma_{k,l-1} \circ ... \circ \Sigma_{k,k+1} \circ \Sigma_{k,l} \circ ... \circ \Sigma_{l-2,l} \circ \Sigma_{l-1,l}.$$

*Graphically, for the case $n=3$, if we want to swap $A_1$ and $A_3$ the equation looks like (note that we do not draw the tensor units):*




**Definition 2.14 (Move-back).** *Given $A_1, \dots, A_n$ objects in* **CPTP**, *the map* $\mathsf{move-back}_{k,n}: A_1 \otimes ... \otimes A_k \otimes ... \otimes A_n \to A_1 \otimes ... \otimes A_{k-1} \otimes A_{k+1} \otimes ... \otimes A_n \otimes A_k$ *is defined as:*

$$\mathsf{move-back}_{k,n} = \circ_{i=k}^{n} \mathsf{swap_{i,i+1}}.$$

At last, we need to implement a different kind of branching. So far, we are only representing explicit branching. Hence, one can learn from the outside which branch the program chose. However, if the meta-data of the state, especially the dimensionality, does not give away which branch was chosen, one can choose not to output this information. In this case, we shall delete this information after branching with $\mathsf{elif}$.

**Definition 2.15 (Forget-branch).** *Given an object $A$ in* **CPTP**, *the morphism* $\mathsf{forget-branch}_{n,A}: A^{\oplus n} \to A$ *is defined as follows:*

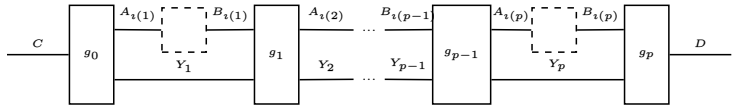
$$\mathsf{forget-branch}_{n,A} = (Tr_{I^{\oplus n}} \otimes id_A) \circ B_{n,A}^{down},$$

*where $Tr_{(-)}$ is the map to the monoidal unit.*

**n-combs** In the following, we wish to formalize settings where we can model useful cryptographic resources based on quantum channels shared between $n$ parties. In cryptographic protocols, each party acts locally on their system, and the parties interact with one another over multiple rounds. Our category therefore needs to model both local actions on systems as well as multiple rounds of interactions between parties. To this end, we present a slightly modified version of the n-comb category defined in [3, Definition 3.2]. In n-comb morphisms represent a given agent's part of a protocol. In contrast to the one in [3], our definition allows for settings, where the protocols don't necessarily use all shared resources.

**Definition 2.16.** *Given an SMC* $\mathbf{C}$, *the category* n-comb($\mathbf{C}$) *is defined as follows: objects of* n-comb($\mathbf{C}$) *are finite lists* $(A_i, B_i)_{i=1}^{m}$ *of pairs of objects of* $\mathbf{C}$. *Morphisms*

*are defined in two stages: For $p \leq m$, a morphism $(A_i,B_i)_{i=1}^m \to (C,D)$ is given by an injection $\imath\colon \{1,...,p\} \to \{1,...,m\}$ and a p-comb*




*in $\mathbf{C}$. Formally, a p-comb is an equivalence class of tuples $(g_0,...,g_p)$ of maps in $\mathbf{C}$, where $g_0\colon C \to A_{\imath(1)} \otimes Y_1$, $g_i\colon B_{\imath(i)} \otimes Y_i \to A_{\imath(i+1)} \otimes Y_{i+1}$ for $i = 1...p-1$ and $g_p\colon B_{\imath(p)} \otimes Y_p \to D$ for some objects $Y_i$. Two such tuples are identified if, whenever one "plugs the holes" with maps of the form $Z_i \otimes A_{\imath(i)} \to Z_i \otimes B_{\imath(i)}$, the resulting maps in $\mathbf{C}$ are equal.*
*A morphism $(A_i,B_i)_{i=1}^m \to (C_j,D_j)_{j=1}^k$ is given by a function $f\colon \{1,...,m\} \to \{1,...,k\}$ and a morphism $(A_i,B_i)_{i \in f^{-1}(j)} \to (C_j,D_j)$ for each $j$. Composition is defined by nesting circuits into circuits, and the monoidal product is given by concatenation of lists.*

Note that the monoidal product in the underlying category $\mathbf{C}$ is different from that of in n-comb($\mathbf{C}$).

## 3 Categorical Composable Cryptography

One of the main contributions of [2, 3] are highly general composition theorems. These can be viewed as giving a blueprint for numerous models of composable cryptography: one gets a specific model by fixing each degree of freedom in the formalism. To fix these, one first needs to choose two SMCs $\mathbf{D}$ and $\mathbf{C}$, where $\mathbf{D}$ models the protocols, and $\mathbf{C}$ models the relevant kind of (computational) processes, which may or may not be more general than those given in $\mathbf{D}$. One also needs to fix a map $\mathbf{D} \to \mathbf{C}$ of SMCs which interprets protocols into processes. One also needs to give a map out of $\mathbf{C}$, which gives for each object (thought of as a system type) the resources of that type, and specifies how processes in $\mathbf{C}$ act on these resources. If one requires perfect security, this operation $R$ can be modeled as a suitable kind of map of SMCs $\mathbf{C} \to \mathbf{Set}$, so that in particular for each object $A$ we have a set $R(A)$ of resources of that type. If we want to model security up to (computational) indistinguishability, $R(A)$ should be equipped with an equivalence relation, and if we want to do security up to some notion of distance, then $R(A)$ should be a (pseudo)metric space. The chain of maps $\mathbf{D} \to \mathbf{C} \to \mathbf{Set}$ (or $\mathbf{D} \to \mathbf{C} \to \mathbf{Met}$) then induces a *resource theory* of correct conversions between resources. To add in a notion of security, one needs a further structure called an *attack model* on $\mathbf{C}$, which in a nutshell specifies the way that adversaries can force a protocol to deviate from its intended behavior. One can then form the SMC of (suitably correct) resource conversions that are secure against this attack model, and the fact that this results in an SMC is the heart of the composition theorem—secure conversions are closed under sequential and parallel composition.

For a detailed exposition on how these resource theories arise we refer the reader to [2,3]. For a more general study of resource theories one can consult [4].
In this section we adapt this framework for our analysis of quantum state verification protocols. We present the resource theories we work in for single- and multi-client verification protocols. Based on this we can present our security definition and finally we give a formal definition of our ideal resources.

### 3.1 The relevant resource theories

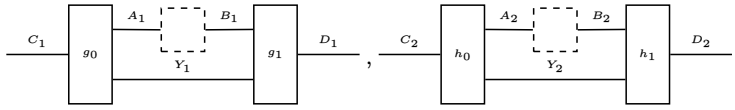
We wish to consider "security up to $\varepsilon$", so our mapping $R$ specifying the resources of a given type should be a (pseudo)metric space. In other words, we wish for $R$ to land in the SMC **Met** of extended (pseudo)metric spaces and short maps from examples 2.5 and 2.7.
In fact, we will define two different, albeit similar, resource theories for multi- and single-party quantum state verification. In quantum state verification, we consider a source that can perform any arbitrary quantum operation. This corresponds to the category **CPTP**, which we denote from now on with $\mathbf{C}$ to simplify notation. In the single-party case we consider the receiving party, that wishes to verify the quantum state, to be able to measure the state, but not to create the state. We define $\mathbf{D}$ to be the sub-SMCs of $\mathbf{C}$ generated by morphisms that are destructive quantum measurements and by arbitrary maps between classical systems (which correspond to stochastic maps).
For multi-party verification, we again consider a source corresponding to $\mathbf{C}$ along with $k$ clients each also acting in $\mathbf{C}$. The clients can only act locally and cannot create entanglement with one another. This restriction is represented by the Cartesian product $\mathbf{C}^k$. The resource theories of single- and multi-party quantum state verification respectively are induced by the maps

$$\text{n-comb}(\mathbf{D}\times\mathbf{C})\to\text{n-comb}(\mathbf{C})\to\mathbf{Met} \tag{3.1}$$

$$\text{n-comb}(\mathbf{C}^k\times\mathbf{C})\to\text{n-comb}(\mathbf{C})\to\mathbf{Met} \tag{3.2}$$

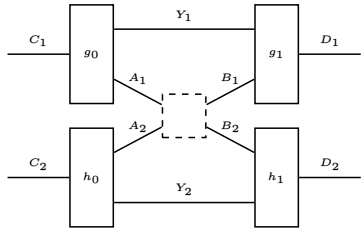
The morphisms on the left are the monoidal functors induced by the ($k$-)fold tensor product $\mathbf{D}\times\mathbf{C}\hookrightarrow\mathbf{C}$ and $\mathbf{C}^k\times\mathbf{C}\hookrightarrow\mathbf{C}$. The second map is given by $\text{n-comb}(\mathbf{C})(I,-)$, where $I$ is the tensor unit in $\text{n-comb}(\mathbf{C})$.
Let us now explain what these abstract definitions amount to concretely, starting from the simpler case of (3.1). We first unwind the definitions. An object of $\text{n-comb}(\mathbf{D}\times\mathbf{C})$ is given by a finite list $(A_i,B_i)_{i=1}^n$ of objects of $\mathbf{D}\times\mathbf{C}$, but we'll first focus on lists $(A,B)$ of length one. In turn, an object of $\mathbf{D}\times\mathbf{C}$ is a pair of objects: one of $\mathbf{D}$ and one of $\mathbf{C}$. Thus each $(A,B)$ is of the form $((A_1,A_2),(B_1,B_2))$, and one can then show that the map (3.1) sends $((A_1,A_2),(B_1,B_2))$ to $\mathbf{CPTP}(A_1\otimes A_2,B_1\otimes B_2)$. It follows that a resource of type $((A_1,A_2),(B_1,B_2))$ is given by a bipartite quantum channel $A_1\otimes A_2\to B_1\otimes B_2$, where we think of the first input and output belonging to the first party (the verifier) and the second input and output belong to the second party (the source). More generally, a resource of type $(A_i,B_i)_{i=1}^n$ is a list of $n$ such bipartite channels.

Given a starting resource $f\colon A_1\otimes A_2\to B_1\otimes B_2$ (of type $((A_1,A_2),(B_1,B_2))$ and a target resource $g\colon C_1\otimes C_2\to D_1\otimes D_2$ (of type $((C_1,C_2),(D_1,D_2))$, a resource conversion $f\to g$ can be depicted by two 1-combs, one for each party, as in




(3.3)

where we require that the first part, belonging to the verifier, lives in the category $\mathbf{D}$ (i.e., that $g_0$ and $g_1$ are morphisms in $\mathbf{D}$). This resource conversion is correct, exactly if, when applied to $f$, it produces $g$, i.e., filling the hole in




(3.4)

with $f$ results in $g$. Typically, but not necessarily, the resources used enable communication between the parties, so that one could then think of these pictures as depicting a 2-party 1-round protocol. A more general resource conversion $(f_1,\dots f_n)\to g$ is similar, except that (i) there's more holes in the pictures (corresponding to more rounds in the protocol) and (ii) the parties have to agree on the order they call the shared resources $f_i$ (which, in the case of communication, amounts to agreeing what kind of information is sent at each round). The parties can also agree to not use some of the shared resources. We note that since n-combs are finite and holes represent rounds, we are de facto setting an upper limit on the number of rounds. However, this does not pose a problem since the size of the n-comb can be arbitrarily chosen. In fact, any sensible model would not allow for an infinite number of rounds and would abort after a preset number of rounds.
We consider security of such protocols in the next subsection, and conclude this subsection by verifying carefully that n-comb$(\mathbf{C})(I,-)$ is indeed a map to $\mathbf{Met}$. Objects in n-comb$(\mathbf{C})$ are finite lists, and the tensor unit $I$ of n-comb$(\mathbf{C})$ is the empty list. For $C,D\in\mathbf{C}$ we have $(C,D)\in$n-comb$(C)$ and

$$\text{n-comb}(\mathbf{C})(I,(C,D))=\mathbf{C}(C,D).$$

By endowing $\mathbf{C}(C,D)$ with the diamond distance (2.1), we obtain a metric space $(\mathbf{C}(C,D),d_\diamond)$. For an object $(A_i,B_i)_{i=1}^n$, one can show that n-comb$(\mathbf{C})(I,-)$ maps it to a product space

$$\text{n-comb}(\mathbf{C})(I,(A_i,B_i)_{i=1}^n)=\mathbf{C}(A_1,B_1)\times\cdots\times\mathbf{C}(A_n,B_n). \tag{3.5}$$

For this product space we use the monoidal structure of **Met**, so it is equipped with the $\ell^1$-distance given by the sum of diamond distance on each of the hom spaces. Let $d := \sum_{i=1}^{n} d_\diamond$ denote said distance. The product space (3.5) is then also a metric space $\left(\bigtimes_{i=1}^{n} \mathbf{C}(A_i,B_i), d\right)$.
Next, we show that n-combs induce morphisms in **Met**, which are short maps. A n-comb maps a list $(A_i,B_i)_{i=1}^{l}$ with $l \geq n$ to $(C,D)$ This induces a morphism

$$\gamma : \text{n-comb}(\mathbf{C})(I,(A_i,B_i)_{i=1}^{l}) \to \text{n-comb}(\mathbf{C})(I,(C,D)).$$

that is, a morphism

$$\mathbf{C}(A_1,B_1) \times \cdots \times \mathbf{C}(A_n,B_n) \to \mathbf{C}(C,D)$$

that acts on $(a_1,...a_n)$ by filling the holes in the comb with $a_i$. We then need to check that $\gamma$ itself induces a short map, i.e., a morphism in **Met**. Let $x = (\imath, (h_i)_{i=0}^{n})$ with $h_0 : C \to A_{\imath(1)} \otimes Y_1$, $h_1 : B_{\imath(i)} \otimes Y_i \to A_{\imath(i+1)} \otimes Y_{i+1}$ for $i = 1,\ldots,n-1$ and $h_n : B_{\imath(n)} \otimes Y_n \to D$ specify an n-comb, and let $x(\bar{a})$ be the n-comb filled with the tuple of CPTP maps $\bar{a} = (a_1,...,a_l)$, where $a_i : A_i \to B_i$ and $l \geq n$. And $Y_k$ denotes an auxiliary register. To show that $\gamma$ induces a short map we need to show that $d(\bar{a},\bar{b}) \geq d_\diamond(x(\bar{a}),x(\bar{b}))$. Indeed using the notation as in Definition 2.16, with $\imath$ being an injection from $\{1,...,n\}$ to $\{1,...,l\}$, we can write $x(\bar{a})$ as

$$x(\bar{a}) = h_n \circ_{i=1}^{n} [(a_{\imath(i)} \otimes \mathrm{id}_{Y_i}) \circ h_{i-1}].$$

By the properties of the diamond distance (2.2),(2.3) and the fact that $d_\diamond(f,f) = 0$ by virtue of it being a metric, it then follows

$$\begin{aligned}
&d_\diamond(x(\bar{a}),x(\bar{b})) \\
=&d_\diamond(h_n \circ_{i=1}^{n} [(a_{\imath(i)} \otimes \mathrm{id}_{Y_i}) \circ h_{i-1}], h_n \circ_{i=1}^{n} [(b_{\imath(i)} \otimes \mathrm{id}_{Y_i}) \circ h_{i-1}]) \\
\leq&d_\diamond(h_n,h_n) + d_\diamond(\circ_{i=1}^{n} [(a_{\imath(i)} \otimes \mathrm{id}_{Y_i}) \circ h_{i-1}], \circ_{i=1}^{n} [(b_{\imath(i)} \otimes \mathrm{id}_{Y_i}) \circ h_{i-1}]) \\
\leq&\sum_{i=1}^{n} d_\diamond((a_{\imath(i)} \otimes \mathrm{id}_{Y_i}) \circ h_{i-1}, (b_{\imath(i)} \otimes \mathrm{id}_{Y_i}) \circ h_{i-1}) \\
\leq&\sum_{i=1}^{n} [d_\diamond((a_{\imath(i)} \otimes \mathrm{id}_{Y_i}), (b_{\imath(i)} \otimes \mathrm{id}_{Y_i})) + d_\diamond(h_{i-1},h_{i-1})] \\
\leq&\sum_{i=1}^{n} d_\diamond(a_{\imath(i)},b_{\imath(i)}) \leq \sum_{i=1}^{l} d_\diamond(a_i,b_i) = d(\bar{a},\bar{b}).
\end{aligned}$$

### 3.2 Security Definition

Using the resource theories (3.1) and (3.2), we can present our security definition based on [2, 3]. The security definition relies on an attack model $\mathcal{A}$. An attack model on a category $\mathbf{C}$ gives for every morphism $f$ in $\mathbf{C}$ a class of morphisms $\mathcal{A}(f)$ that fulfills certain properties which are stated in the original work.
To capture both of the situations above in a single situation, we model the situation

with $K+1$ parties, where the last party acts maliciously in $\mathbf{C}$, and let $\mathbf{E}$ be a sub-SMC of $\mathbf{C}$ (where in the above, we either have $\mathbf{E}=\mathbf{C}$ or $\mathbf{E}=\mathbf{D}$). We define an attack model on n-comb$(\mathbf{E}^K \times \mathbf{C})$ derived from a general attack model in [3]. The attack model $\mathcal{A}$ consists of allowing the last party to change their part of any $m$-comb arbitrarily while leaving everything else in the morphisms of n-comb$(\mathbf{E}^K \times \mathbf{C})$ unchanged. For example, in the 2-party 1-round case depicted in (3.3) and (3.4), this amounts to allowing the second party to change their comb (and hence their resulting input $C_2$ and output $D_2$) arbitrarily, provided they do send and receive something of type $A_2$ and $B_2$ respectively into the shared resource.
We now give a formal definition of the attack model, but note that the intuitive definition above is sufficient for many purposes.

**Definition 3.1 (Attack model $\mathcal{A}$ on** n-comb$(\mathbf{E}^K \times \mathbf{C})$ **).** *We define an attack model $\mathcal{A}$ on* n-comb$(\mathbf{E}^K \times \mathbf{C})$ *corresponding to $K$ honest parties and one malicious party. Consider a morphism in* n-comb$(\mathbf{E}^K \times \mathbf{C})$ *given by an injection $\iota : \{1,...,m\} \to \{1,...,l\}$ and a $m$-tuple of morphisms in* n-comb$(\mathbf{C}^k)$*, $(g_0,...,g_m)$, that is*

$$(\iota,(g_0,...,g_m)) : (A_i,B_i)_{i=1}^{l} \to (C,D)$$

*Each $g_j$ is a morphism in $\mathbf{E}^K \times \mathbf{C}$, and as such itself a tuple of morphisms in $\mathbf{E}$ and one in $\mathbf{C}$. We write $\pi_j : \mathbf{E}^K \times \mathbf{C} \to \mathbf{E}$ for the $j$-the projection. The attack model is then defined as*

$$\mathcal{A}((\iota,(g_0,...,g_m))) := \{(\iota,(h_0,...,h_m)) | \pi_j(h_\ell) = \pi_j(g_\ell) \text{ for all } \ell \text{ and } 1 \leq j \leq K\}.$$

In the following we study quantum state verification protocols with $K \geq 1$ honest clients and a dishonest source. The source is then modeled in $\mathbf{C}$ and the clients correspond to $\mathbf{E}^K$ in Definition 3.1.
Intuitively, security against a dishonest source means that for every attack $a$ in the attack model in $\mathcal{A}(\pi)$ applied on the protocol, there is an attack $b$ in the attack model $\mathcal{A}(\mathrm{id}_s)$ on the identity for the ideal resource such that $a$ applied on the real resources and $b$ on the ideal resource are indistinguishable up to $\varepsilon$. With this intuition and the attack model we defined before, we now define security formally.

**Definition 3.2 (Security against the source).** *Let $\mathbf{E}$ be any sub-SMC in $\mathbf{C}$. We further consider $F :$* n-comb$(\mathbf{E}^K \times \mathbf{C}) \to$ n-comb$(\mathbf{C})$ *being the injection $\mathbf{E} \hookrightarrow \mathbf{C}$ followed by the $(K+1)$-fold tensor product and $R :$* n-comb$(\mathbf{C}) \to \mathbf{Met}$ *given by* n-comb$(\mathbf{C})(I,-)$*. A protocol (morphism in* n-comb$(\mathbf{E}^K \times \mathbf{C})$*) $\pi : (A,\bar{r}) \to (B,\bar{q})$ $\varepsilon$-securely implements $(B,\bar{s})$ with an untrusted source if*

$$\forall a \in \mathcal{A}(\pi)\ \exists b \in \mathcal{A}(\mathrm{id}_B) : \frac{1}{2} d_\diamond(RF(a)\bar{r}, RF(b)\bar{s}) \leq \varepsilon,$$

*where $\mathcal{A}$ is the attack model as defined in Def. 3.1.*

Apart from security, we need a definition of correctness, i.e. that the implementation is close to the ideal resource if all parties act honestly.

**Definition 3.3 (Correctness).** *Let* $\mathbf{E}$ *be any sub-SMC in* $\mathbf{C}$*. We further consider* $F : \text{n-comb}(\mathbf{E}^K \times \mathbf{C}) \to \text{n-comb}(\mathbf{C})$ *being the injection* $\mathbf{E} \hookrightarrow \mathbf{C}$ *followed by the* $(K+1)$*-fold tensor product and* $R : \text{n-comb}(\mathbf{C}) \to \mathbf{Met}$ *given by* $\text{n-comb}(\mathbf{C})(I,-)$*. A protocol (morphism in* $\text{n-comb}(\mathbf{E}^K \times \mathbf{C})$*)* $\pi : (A,\vec{r}) \to (B,\vec{q})$ *ε-correctly implements* $(B,\vec{s})$ *if*

$$\frac{1}{2} d_\diamond(RF(\pi)\vec{r}, RF(\mathrm{id}_B)\vec{s}) \leq \varepsilon.$$

### 3.3 Ideal resource

To prove that there is no efficient and secure quantum state verification, we also need to define the ideal resource. The ideal resource we consider is the same as in [5]. To ease the reading of the ideal resource as morphisms, we will write out tensor units explicitly when they represent the input or output of parties. Further, we use the tensor unit as a constant signal, which can either be the start signal or the abort/end signal, depending if it is the input or output of the morphism.

**Ideal resource 1:** $\mathcal{S}_{\phi,K}^{QSV}$, the ideal quantum state verification resource for $K$ clients and one source.

**Input:** The clients input $I$.
**Input:** The source inputs $c \in \{0,1\}$.
  **if** $c = 0$ **then**
    $\xi \leftarrow \phi$
  **else**
    $\xi \leftarrow I^{\otimes n}$
  **end if**
**Output:** The client receives $\xi$.

*Remark 3.4.* In the category **CPTP**, morphisms are quantum channels defined on finite-dimensional $C^*$-algebras, or concretely on direct sums of matrix algebras $M_n(\mathbb{C})$. In quantum cryptography, however, we work with density matrices. Density matrices form a subset of all complex matrices, that is $D\left(\bigotimes_{i=1}^k \mathbb{C}^{n_i}\right) \subseteq \bigotimes_{i=1}^k M_{n_i}(\mathbb{C})$ for any tuple $(n_1,\ldots,n_k) \in \mathbb{N}^k$. Moreover, a quantum channel always maps density operators to density operators. Therefore, all the preliminaries in Section 2.1 also hold for the morphisms in the category **CPTP**. In the subsequent work we can therefore restrict our analysis to density operators while still working with the morphisms in **CPTP**.

**Definition 3.5 (Quantum state verification).** *Let* $\mathcal{S}_{\phi,K}^{QSV}$ *be the* quantum state verification *resource for* $K$ *clients* $C = \{i\}_{i=1}^K$*, a source* $S$ *and a target state* $\phi \in D\left(\bigotimes_{i=1}^K \mathbb{C}^{n_i}\right)$*. The source decides with their input* $c \in \{0,1\}$ *if the clients*

*receive the target state $\phi$ or the tensor unit $I$. As a morphism, we can type $S_{\phi,K}^{QSV}$ as follows*

$$\mathcal{S}_{\phi,K}^{QSV} : \left(\bigotimes_{i=1}^{K} I\right) \otimes (I \oplus I) \to \left(\left(\bigotimes_{i=1}^{K} M_{n_i}(\mathbb{C})\right) \oplus \left(\bigotimes_{i=1}^{K} I\right)\right) \otimes I.$$

*We show the ideal resource in* ***Ideal resource 1****.*

We further introduce the 1-comb $t_{\sharp_S}$ that takes the role of a filter, i.e. an operation applied to the ideal resource that shields access that should be only available to dishonest parties. We apply a filter in the honest case to ensure that the source cannot force the ideal resource to abort. Again, we write tensor units explicitly to represent the inputs or outputs of the different parties:

$$t_{\sharp_S} = \left(I^{\otimes K+1} \to I^{\otimes K} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \mathrm{id}_{\left(\left(\bigotimes_{i=1}^{K} M_{n_i}(\mathbb{C})\right) \oplus \left(\bigotimes_{i=1}^{K} I\right)\right) \otimes I}\right).$$

We define $\sharp_S = (1 \mapsto 1, t_{\sharp_S})$.
Now we have all definitions we need to define $\varepsilon$-implementations of quantum state verification.

**Definition 3.6 (Implementation).** *Let $\bar{r}$ be any sequence of resources, and $\pi$ be a protocol in form of a morphism in* n-comb$(\mathbf{E}^K \times \mathbf{C})$ *applicable to $\bar{a}$. We say $\pi(\bar{a})$ is an $\varepsilon$-implementation for quantum state verification if*

- *$\pi$ $\varepsilon$-correctly implements $\sharp_S\left(\mathcal{S}_{\phi,K}^{QSV}\right)$ and*
- *$\pi$ $\varepsilon$-securely implements $\mathcal{S}_{\phi,K}^{QSV}$*

*from $\bar{r}$.*

We still need to list the resources we are considering for the implementation. While we shouldn't be overly restrictive, we must ensure that the clients cannot use these resources alone to fully prepare the target state. Nevertheless, the resource should provide the necessary communication structure. Motivated by this contrast, we describe the abilities provided by the resources and their restrictions below.

- $\mathcal{W}$ is a resource allowing the clients to coordinate their verification. It is assumed, that $\mathcal{W}$ either doesn't allow to or is not used to distribute the output state.
- $\mathcal{Q}$ is a quantum communication channel from the source to the client in single-client quantum verification.
- $\mathcal{T}$ is a quantum communication channel from the source to all clients and allows some quantum communication among the clients. Nevertheless, the graph representing the connectivity of the clients is not connected. We assume $\mathcal{T}$ for quantum state verification.
- $\mathcal{V}$ is a resource allowing the clients to sample whether they query another state and reveals the decision to the source. We assume once, $\mathcal{V}$ outputted that no further state should be queried, the parties don't use it or ignore its outputs.

# 4 No-Go result

## 4.1 Simple protocols

We first consider a simple type of quantum state verification protocols. In this simple setting, an honest source sends $N+1$ copies of the ideal state to the client(s). The client(s) perform a measurement on a random subset of size $N$. If the measurement outcome is 0 they accept the verification and output the remaining state to the environment. If the measurement outcome is 1, they output the abort signal $I$ to the environment.

**Definition 4.1 (Simple protocol type).** *Let $N \geq 0$ be an integer, $\eta \in I^{\oplus N+1}$ a probability distribution, $\phi \in D\left(\bigotimes_{i=1}^{K} \mathbb{C}^{n_i}\right)$ the target state for $K \geq 1$ client(s) and $\mu : \left(\bigotimes_{i=1}^{K} M_{n_i}(\mathbb{C})\right)^{\otimes N} \to I \oplus I$ a measurement. $\pi^{SP}$ is defined by the two algorithms $\pi_S^{SP}$ and $\pi_C^{SP}$, where $\pi_S^{SP}$ describes the protocol followed by the source preparing the states and $\pi_C^{SP}$ the protocol followed by the client(s) to verify the states.*

---

**Protocol 2:** The protocol $\pi^{SP}$ of the source and the (joint) protocol of the client(s). $N$, $K$ and $\phi$ are publicly known and fixed per protocol instance.

---

**Source's protocol $\pi_S^{SP}$:**

1: Prepare $N+1$ copies of the target state, i.e. $I \to \phi^{\otimes(N+1)}$.
2: Send these copies to the client(s).

**Client's protocol $\pi_C^{SP}$:**

1: The client(s) receive their respective share of each of the $N+1$ states in $D\left(\bigotimes_{i=1}^{K} \mathbb{C}^{n_i}\right)$, i.e. $I \to \bigotimes_{i=1}^{N+1} \rho_i$.
2: The client(s) sample the output register: $r \leftarrow \eta$
3: **if** $r = 1$ **then**
4: $\quad \bigotimes_{i=1}^{N+1} \rho_i' \leftarrow \texttt{MOVE-BACK}_{N+1,1}(\bigotimes_{i=1}^{N+1} \rho_i)$.
5: **else if** $r = \ldots$ **then**
6: $\quad \vdots$
7: **else**
8: $\quad \bigotimes_{i=1}^{N+1} \rho_i' \leftarrow \texttt{MOVE-BACK}_{N+1,N+1}(\bigotimes_{i=1}^{N+1} \rho_i)$.
9: **end if**
10: Perform $\texttt{FORGET-BRANCH}_{N+1,\mathcal{H}_c}$ and get $\bigotimes_{i=1}^{N+1} \rho_i''$.
11: Perform the measurement $\mu$ on the first $N$ registers, the result is $s$. The remaining register is now called $\rho'''$.
12: **if** $s = 0$ **then**
13: $\quad$ Output $\rho'''$, distributed to the clients.
14: **else**
15: $\quad$ Output $\mathrm{Tr}(\rho''')$ to each client.
16: **end if**

---

Where the else-if structure, `MOVE-BACK`, and `FORGET-BRANCH` are defined in Definitions 2.12, 2.14, and 2.15. With the definition of the simple protocol type, we can derive our first result. We show that there is no composably secure single- or multi-client quantum state verification protocol that is efficient. Intuitively, both players, the client(s) and the distinguisher, are limited by the Holevo-Helstrom theorem. The crucial observation is, that the distinguisher can choose a state that is likely to pass the test of the client(s), but is still non-negligibly far away from the target state. We start with single-client state verification, i.e. we consider a single client, who is not able to prepare a state at all. For any simple protocol $\pi$ with $N+1$ rounds we show that $\pi$ cannot be an $negl(N)$-implementation as defined in 3.6, i.e., $\pi$ cannot be negligible close to the ideal functionality and efficient.

**Theorem 4.2 (No efficient single-client state verification with fixed number of rounds).** *Let $\pi=\{\pi_S,\pi_C\}$ be a simple protocol (see Def.4.1). Then there exists a morphism $\mathbf{A}_\rho\in\mathcal{A}(\pi)$ such that for all $\mathbf{B}_\rho\in\mathcal{A}\left(\mathrm{id}_{\mathcal{S}_{\phi,1}^{QSV}}\right)$ it holds*

$$\frac{1}{2}d_\diamond\left(RF(\pi)(\bar{a}),RF(\sharp_S)\left(\mathcal{S}_{\phi,1}^{QSV}\right)\right)+\frac{1}{2}d_\diamond\left(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\left(\mathcal{S}_{\phi,1}^{QSV}\right)\right)\geq\varepsilon,$$

*where $\bar{a}=(\mathcal{Q})^{\times N+1}$ consists of $N+1$ copies on quantum communication from the source to the client, $\varepsilon=1/27N$, if $\phi$ is mixed and $\varepsilon=1/8\sqrt{N}$, if $\phi$ is pure.*

*Proof.* First, we consider the setting, where the source is honest, i.e. $\frac{1}{2}d_\diamond\left(RF(\pi)(\bar{a}),RF(\sharp_S)\left(\mathcal{S}_{\phi,1}^{QSV}\right)\right)$. Half the diamond distance being the distinguishing advantage, we know that this quantity is lower bounded by the difference of the probabilities that the real and ideal resource output 0, that is

$$\frac{1}{2}d_\diamond\left(RF(\pi)(\bar{a}),RF(\sharp_S)\left(\mathcal{S}_{\phi,1}^{QSV}\right)\right)\geq\left|Pr[(\pi(\bar{a}))(I)=0]-Pr\left[\left(\sharp_S\left(\mathcal{S}_{\phi,1}^{QSV}\right)\right)(I)=0\right]\right|. \tag{4.1}$$

We consider the channel $\mathcal{M}=\mathrm{Tr}_{M_n(\mathbb{C})}\oplus\mathrm{id}_I$. If we apply $\mathcal{M}$ to the resource's output, we find

$$\begin{aligned}Pr[(\mathcal{M}\circ\pi(\bar{a}))(I)=0]&=\left\langle\mu(0)\middle|\phi^{\otimes N}\right\rangle\\ Pr\left[\left(\mathcal{M}\circ\sharp_S\left(\mathcal{S}_{\phi,1}^{QSV}\right)\right)(I)=0\right]&=1.\end{aligned}$$

However the diamond distance is contractive with respect to CPTP maps, i.e. $\mathcal{M}$ cannot increase the diamond distance, which implies

$$\begin{aligned}&\frac{1}{2}d_\diamond\left(RF(\pi)(\bar{a}),RF(\sharp_S)\left(\mathcal{S}_{\phi,1}^{QSV}\right)\right)\\&\geq\frac{1}{2}d_\diamond\left(RF(\mathcal{M}\circ\pi)(\bar{a}),RF(\mathcal{M}\circ\sharp_S)\left(\mathcal{S}_{\phi,1}^{QSV}\right)\right)\geq1-\left\langle\mu(0)\middle|\phi^{\otimes N}\right\rangle.\end{aligned} \tag{4.2}$$

If the source is dishonest, we use the family of attacks $\{\mathbf{A}_\rho\}_{\rho\in D(\mathbb{C}^n)}$. We apply these attacks to $\mathcal{Q}^{\times N+1}$ and obtain a channel from $I$ to $M_n(\mathbb{C})\oplus I$. $\mathbf{A}_\rho$ inputs

$N+1$ copies of as state $\rho\in D(\mathbb{C}^n)$ and implements $\pi_C$ on the client side. To bound $1/2 d_\diamond\left(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\left(\mathcal{S}_{\phi,1}^{QSV}\right)\right)$ we need to model any attack on $\mathrm{id}_{\mathcal{S}_{\phi,1}^{QSV}}$. For the input part of the comb we consider any channel from $I$ to $I\oplus I$. The output part can only be the identity, since there is no output on the side of the source and the client is honest. Hence, $\mathbf{B}_\rho$ can only input a single classical bit $c$ to the ideal resource $\mathcal{S}_{\phi,1}^{QSV}$. Therefore, we can assume without loss of generality, that there is a probability $p_N(\rho)$ for the simulator to input $c=1$, i.e. $p_N(\rho)=Pr\left[c=0\,|\,c\leftarrow\mathbf{B}_\rho\left(\mathcal{S}_{\phi,1}^{QSV}\right)\right]$. As in the honest case, we want to apply a measurement channel to the outputs of the channels obtained by $\mathbf{A}_\rho(\bar{a})$ and $\mathbf{B}_\rho\left(\mathcal{S}_{\phi,1}^{QSV}\right)$. Let $\mathcal{M}_D=\{\gamma(0)\oplus 1,\gamma(1)\oplus 0\}$ be a measurement channel, where $\gamma$ is an arbitrary binary measurement. It follows[7]

$$\begin{aligned}Pr[\mathcal{M}_D\circ\mathbf{A}_\rho(\bar{a})(I)=1]&=\langle\gamma(1)|\rho\rangle\langle\mu(0)|\rho^{\otimes N}\rangle\\ Pr\left[\mathcal{M}_D\circ\mathbf{B}_\rho\left(\mathcal{S}_{\phi,1}^{QSV}\right)(I)=1\right]&=\langle\gamma(1)|\phi\rangle p_N(\rho).\end{aligned}$$

And the Holevo-Helstrom Theorem implies

$$\frac{1}{2}d_\diamond\left(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\left(\mathcal{S}_{\phi,1}^{QSV}\right)\right)\geq\left|\langle\gamma(1)|\rho\rangle\langle\mu(0)|\rho^{\otimes N}\rangle-\langle\gamma(1)|\phi\rangle p_N(\rho)\right|. \tag{4.3}$$

When considering $\gamma(1)=\mathbb{1}_{\mathbb{C}^n}$, we find

$$\frac{1}{2}d_\diamond\left(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\left(\mathcal{S}_{\phi,1}^{QSV}\right)\right)\geq\left|\langle\mu(0)|\rho^{\otimes N}\rangle-p_N(\rho)\right|.$$

We denote $p_N(\rho)=\langle\mu(0)|\rho^{\otimes N}\rangle+\delta(N,\rho)$ and find

$$\frac{1}{2}d_\diamond\left(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\left(\mathcal{S}_{\phi,1}^{QSV}\right)\right)\geq|\delta(N,\rho)|.$$

For an arbitrary $\gamma(1)$, we then have

$$\begin{aligned}&\frac{1}{2}d_\diamond\left(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\left(\mathcal{S}_{\phi,1}^{QSV}\right)\right)\\ \geq&\left|\langle\mu(0)|\rho^{\otimes N}\rangle\langle\gamma(1)|\rho\rangle-p_N(\rho)\langle\gamma(1)|\phi\rangle\right|\\ =&\left|\langle\mu(0)|\rho^{\otimes N}\rangle(\langle\gamma(1)|\rho\rangle-\langle\gamma(1)|\phi\rangle)-\delta(N,\rho)\langle\gamma(1)|\phi\rangle\right|\\ \geq&\left|\langle\mu(0)|\rho^{\otimes N}\rangle(\langle\gamma(1)|\rho\rangle-\langle\gamma(1)|\phi\rangle)\right|-|\delta(N,\rho)|\\ \geq&\langle\mu(0)|\rho^{\otimes N}\rangle|\langle\gamma(1)|\rho\rangle-\langle\gamma(1)|\phi\rangle|-\frac{1}{2}d_\diamond\left(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\left(\mathcal{S}_{\phi,1}^{QSV}\right)\right).\end{aligned}$$

If we consider the optimal measurement $\{\gamma(0),\gamma(1)\}$ to distinguish $\rho$ and $\phi$, i.e., the measurement that saturates the Holevo-Helstrom bound, then the above inequality is yields

$$\frac{1}{2}d_\diamond\left(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\left(\mathcal{S}_{\phi,1}^{QSV}\right)\right)\geq\frac{1}{2}\langle\mu(0)|\rho^{\otimes N}\rangle\frac{1}{2}\|\rho-\phi\|_1. \tag{4.4}$$

[7] Note, that the attack is i.i.d., hence the probability distribution $\eta$ used by the client(s), does not occur in the analysis.

Adding both honesty-setting (4.2) and (4.4), the Holevo-Helstrom bound yields

$$\begin{aligned}
&\frac{1}{2}d_\diamond\Big(RF(\pi)(\bar{a}),RF(\sharp_S)\Big(\mathcal{S}_{\phi,1}^{QSV}\Big)\Big)+\frac{1}{2}d_\diamond\Big(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\Big(\mathcal{S}_{\phi,1}^{QSV}\Big)\Big)\\
&\geq\big(1-\langle\mu(0)|\phi^{\otimes N}\rangle\big)+\frac{1}{2}\langle\mu(0)|\rho^{\otimes N}\rangle\frac{1}{2}\|\rho-\phi\|_1\\
&\geq\frac{1}{4}\|\rho-\phi\|_1\big(1-\big(\langle\mu(0)|\rho^{\otimes N}\rangle-\langle\mu(0)|\phi^{\otimes N}\rangle\big)\big)\\
&\geq\frac{1}{4}\|\rho-\phi\|_1\left(1-\frac{1}{2}\|\rho^{\otimes N}-\phi^{\otimes N}\|_1\right).
\end{aligned}\tag{4.5}$$

Depending on whether target state $\phi$ is pure or mixed we obtain a different bound. We first consider the mixed case. Using Lemma 2.4, we find

$$\begin{aligned}
&\frac{1}{2}d_\diamond\Big(RF(\pi)(\bar{a}),RF(\sharp_S)\Big(\mathcal{S}_{\phi,1}^{QSV}\Big)\Big)+\frac{1}{2}d_\diamond\Big(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\Big(\mathcal{S}_{\phi,1}^{QSV}\Big)\Big)\\
&\geq\frac{1}{4}\|\rho-\phi\|_1\left(1-\sqrt{1-(1-\|\rho-\phi\|_1)^N}\right)
\end{aligned}$$

Now we fix $\mathbf{A}_\rho$ such that $\frac{1}{2}\|\rho-\phi\|_1=\alpha/N$, which give us

$$\begin{aligned}
&\frac{1}{2}d_\diamond\Big(RF(\pi)(\bar{a}),RF(\sharp_S)\Big(\mathcal{S}_{\phi,1}^{QSV}\Big)\Big)+\frac{1}{2}d_\diamond\Big(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\Big(\mathcal{S}_{\phi,1}^{QSV}\Big)\Big)\\
&\geq\frac{\alpha}{2N}\left(1-\sqrt{1-\left(1-\frac{2\alpha}{N}\right)^N}\right)\\
&\geq\frac{\alpha}{2N}\Big(1-\sqrt{2\alpha}\Big),
\end{aligned}$$

where we used $(1-\frac{\beta}{k})^k\geq 1-\beta$ for $k\in\mathbb{N}$ and $|\beta|\leq k$. This is maximized for $\alpha=2/9$ which gives

$$\frac{1}{2}d_\diamond\Big(RF(\pi)(\bar{a}),RF(\sharp_S)\Big(\mathcal{S}_{\phi,1}^{QSV}\Big)\Big)+\frac{1}{2}d_\diamond\Big(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\Big(\mathcal{S}_{\phi,1}^{QSV}\Big)\Big)\geq\frac{1}{27N}.\tag{4.6}$$

Next, we consider $\phi=|\phi\rangle\langle\phi|$ to be a pure state, and we can choose a pure state $\rho=|\psi\rangle\langle\psi|$ as well. Using (2.11) plugged into (4.5) we obtain

$$\begin{aligned}
&\frac{1}{2}d_\diamond\Big(RF(\pi)(\bar{a}),RF(\sharp_S)\Big(\mathcal{S}_{\phi,1}^{QSV}\Big)\Big)+\frac{1}{2}d_\diamond\Big(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\Big(\mathcal{S}_{\phi,1}^{QSV}\Big)\Big)\\
&\geq\frac{1}{2}\sqrt{1-|\langle\psi|\phi\rangle|^2}\left(1-\sqrt{1-|\langle\psi|\phi\rangle|^{2N}}\right).
\end{aligned}$$

Replacing $\sqrt{1-|\langle\psi|\phi\rangle|^2}$ with $\tau$ yields

$$\begin{aligned}
&\frac{1}{2}d_\diamond\Big(RF(\pi)(\bar{a}),RF(\sharp_S)\Big(\mathcal{S}_{\phi,1}^{QSV}\Big)\Big)+\frac{1}{2}d_\diamond\Big(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\Big(\mathcal{S}_{\phi,1}^{QSV}\Big)\Big)\\
&\geq\frac{\tau}{2}\left(1-\sqrt{1-(1-\tau^2)^N}\right).
\end{aligned}$$

Now we choose $|\psi\rangle$ such that $\tau = 1/2\sqrt{N}$ and we find again using $(1-\frac{\beta}{k})^k \geq 1-\beta$ for $k \in \mathbb{N}$ and $|\beta| \leq k$

$$\frac{1}{2} d_\diamond \left( RF(\pi)(\bar{a}), RF(\sharp_S)\left(\mathcal{S}_{\phi,1}^{QSV}\right)\right) + \frac{1}{2} d_\diamond \left( RF(\mathbf{A}_\rho)(\bar{a}), RF(\mathbf{B}_\rho)\left(\mathcal{S}_{\phi,1}^{QSV}\right)\right) \geq \frac{1}{8\sqrt{N}}. \tag{4.7}$$

□

We can also extend the result for simple protocols for multi-client quantum state verification.

**Theorem 4.3 (No efficient secure state verification with fixed number of rounds for entangled states).** *Let $\pi = \{\pi_S\} \cup \{\pi_i\}_{i \in C}$ be a simple protocol as defined in Def. 4.1. Then there exists a morphism $\mathbf{A}_\rho \in \mathcal{A}(\pi)$ such that for all $\mathbf{B}_\rho \in \mathcal{A}\left(\mathrm{id}_{\mathcal{S}_{\phi,K}^{QSV}}\right)$ it holds*

$$\frac{1}{2} d_\diamond \left( \pi(\bar{a}), \sharp_S \left(\mathcal{S}_{\phi,K}^{QSV}\right)\right) + \frac{1}{2} d_\diamond \left( \mathbf{A}_\rho(\bar{a}), \mathbf{B}_\rho \left(\mathcal{S}_{\phi,K}^{QSV}\right)\right) \geq \varepsilon,$$

*where $\bar{a} = (\mathcal{T}^{\times N+1}, \mathcal{W})$ consits of $N+1$ copies of $\mathcal{T}$ used to distribute the states and one copy of $\mathcal{W}$ used to coordinate the verification task. Further, we find $\varepsilon = 1/27N$, if $\phi$ is mixed and $\varepsilon = 1/8\sqrt{N}$, if $\phi$ is pure.*

The proof of Theorem 4.3 is just a repetition of the proof of Theorem 4.2 – having multiple clients does not change the fundamental inequalities we used.

### 4.2 General protocols

Next, we consider general protocols of quantum state verification. In this setting, the client(s) sample $r$ and $i$, where $r$ is the number of states to be sent by the source and $i$, the number of those states that are then measured to verify the state. To this end the clients use a joint probability distribution $p(r,i)$. Again, if the measurement outcome is 0, they accept the verification and output one of the remaining $r-i$ states to the environment. If the measurement outcome is 1, they output the abort signal $I$ to the environment.

For the general protocol, we need to adapt our categorical model for the multi-client case. For the single-client case, the client is still not able to create a state and needs to output a state created by the distinguisher. This translates to $p(l,l) = 0$ for all $l$. In the multi-client case, the clients can prepare a state. However, the state they can prepare is separable with respect to a partition with respect to which to the target state is entangled. Otherwise, the clients would not require an external source in the first place. In the proof of Theorems 4.4 and 4.5 we allow for an arbitrary number of verification rounds, i.e. $r$ and $i$ are arbitrary positive integers. However, in our categorical modeling, we set an upper bound on the number of rounds. The bound can be arbitrarily chosen and the two settings are then equivalent. Indeed let $D$ be the upper bound, then we can set $p(r,i) = 0$ for all $r \leq D$. On the other hand, the upper bound $D$ can be

chosen arbitrarily large and is therefore no significant restriction. The probabilities for $r > D$ can be chosen arbitrarily small, but non-zero, for every distribution.

With this general formulation of protocols, we find that any implementation of quantum state verification either has a non-negligible distance to the ideal functionality according to Def. 3.6 or is inefficient in the number of rounds.

**Theorem 4.4 (No efficient secure state verification for entangled states).** *Let $\pi = \{\pi_S\} \cup \{\pi_i\}_{i \in C}$ be any protocol applied on resources $\bar{a} = (\mathcal{T}^{\times M}, \mathcal{V}^{\times M}, \mathcal{W})$, and $D$ is an upper bound for the number of rounds. $\pi(\bar{a})$ is an implementation with the following properties*

- *$K$ clients sample the number of rounds $r$ and the number of verification rounds $0 \leq i \leq r$ from any joint distribution $p(r,i)$,*
- *the clients perform a measurement $\mu_{r,i}$ and accept the outcome if the result is $0$,*
- *if $r = i$, the clients prepare a state $\chi \in D\left(\bigotimes_{j=1}^{K} \mathbb{C}^{n_j}\right)$, where $\chi$ and the target state $\phi$ are separable respectively entangled with respect to a particular partition,*
- *if $r \neq i$, the clients output one of the $r - i$ states drawn to any unspecified distribution.*

*It then follows that there exists $\mathbf{A}_\rho \in \mathcal{A}(\pi)$ such that for all $\mathbf{B}_\rho \in \mathcal{A}\left(\mathrm{id}_{\mathcal{S}_{\phi,K}^{QSV}}\right)$ it holds that*

$$\frac{1}{2} d_\diamond\left(\pi(\bar{a}), \sharp_S\left(\mathcal{S}_{\phi,K}^{QSV}\right)\right) + \frac{1}{2} d_\diamond\left(\mathbf{A}_\rho(\bar{a}), \mathbf{B}_\rho\left(\mathcal{S}_{\phi,K}^{QSV}\right)\right) \geq \varepsilon,$$

*where $\varepsilon = 1/27N$, if $\phi$ is mixed and $\varepsilon = 1/8\sqrt{N}$, if $\phi$ is pure and $N$ is the expected number of rounds.*

*Proof (sketch).* The proof is very similar to the proof of Thm. 4.2. Hence, we only sketch the proof idea and provide the formal proof in the appendix A.1.

1. If the source is honest, we use the same measurement $\mathcal{M}$ as in (4.1) to find a lower bound for $1/2 d_\diamond\left(\pi(\bar{a}), \sharp_S\left(\mathcal{S}_{\phi,K}^{QSV}\right)\right)$.

2. If the source is dishonest, we use a similar i.i.d. attack and the measurement $\mathcal{M}_D$ used for (4.3), which yields a lower bound $1/2 d_\diamond\left(\mathbf{A}_\rho(\bar{a}), \mathbf{B}_\rho\left(\mathcal{S}_{\phi,K}^{QSV}\right)\right)$.

3. Next, we consider $\gamma(1) = \mathbb{1}_{\bigotimes_{i=1}^{K} \mathbb{C}^{n_i}}$, find a $\delta$ which we use to eliminate the acceptance probability of $\mathbf{B}_\rho$.

4. An argument based on the direction of $\rho$ and $\gamma$ shows that $\min(\langle\gamma(1)|\chi\rangle - \langle\gamma(1)|\phi\rangle, \langle\gamma(1)|\rho\rangle - \langle\gamma(1)|\phi\rangle) \geq 0$. We then bring both honesty configurations together.

5. At last, we use Jensen's inequality, and the fact that $\lambda \geq 1/2\sqrt{N} \geq 2/9N$ in the asymptotic limit, as $1/2\|\chi - \phi\|_1 > 0$ is constant, to deduce the same lower bounds as in (4.6) and (4.7).

In fact, the same proof works for single-client quantum state verification as well. The difference is now that $p(l,l) = 0$ for every $l \geq 0$ as the client is not able to prepare states at all. Hence, we provide a bound for single-client quantum state verification without restating the proof.

**Theorem 4.5 (No efficient secure single-client quantum state verification).** *Let $\pi = \{\pi_C, \pi_S\}$ be any protocol applied to resources $\bar{a} = (\mathcal{Q}^{\times M}, \mathcal{V}^{\times M}, \mathcal{W})$ and $D$ be an upper bound on the number of rounds. $\pi(\bar{a})$ is an implementation with the following properties*

- *the client samples the number of rounds $r$ and the number of verification rounds $0 \leq i < r$ from any joint distribution $p(r,i)$,*
- *the client performs a measurement $\mu_{r,i}$ and accepts the outcome if the result is $0$,*
- *the client outputs one of the $r-i$ states drawn to any unspecified distribution.*

*It then follows that there exists $\mathbf{A}_\rho \in \mathcal{A}(\pi)$ such that for $\mathbf{B}_\rho \in \mathcal{A}\left(\mathrm{id}_{\mathcal{S}_{\phi,1}^{QSV}}\right)$ it holds*

$$\frac{1}{2} d_\diamond \left( \pi(\bar{a}), \sharp_S \left( \mathcal{S}_{\phi,K}^{QSV} \right) \right) + \frac{1}{2} d_\diamond \left( \mathbf{A}_\rho, \mathbf{B}_\rho \left( \mathcal{S}_{\phi,1}^{QSV} \right) \right) \geq \varepsilon,$$

*where $\varepsilon = 1/27N$, if $\phi$ is mixed and $\varepsilon = 1/8\sqrt{N}$, if $\phi$ is pure and $N$ is the expected number of rounds.*

*Remark 4.6 (Categorical modeling).* Our categorical modeling is more restrictive than necessary. Indeed, in the single-client case, the no-state-preparation assumption is overly restrictive. It is enough to assume that the client cannot prepare states that are too close to the target state, otherwise a source would be superfluous. A reasonable assumption could be that the client can only prepare states that are outside a finite-size ball around the target state. Nevertheless, the categorical model remains simpler and more intuitive with stronger restrictions.
In the multi-client case, we only need to ensure that the target state is entangled with respect to a partition for which the clients cannot generate entanglement. This holds as long as there is a subset of clients which is not connected via quantum channels. If the clients are not connected via quantum channels they can only perform separable operations with respect to that particular partition. Since separable operations cannot create entanglement, we can be certain that any state the clients prepare will be far enough from any target state that is entangled with respect to the same partition. This restriction is sensible, as otherwise, an external source is superfluous.

## 5 Discussion

### 5.1 Summary

In our work, we first present how to use the categorical composable cryptography framework for quantum cryptography. For that, we introduce a resource theory based on n-combs on **CPTP**. The instantiation of the framework we presented can contribute to a deeper understanding of composable quantum cryptography

as it defines protocols, resources and attacks rigorously while still being applicable without additional effort.
Using this instantiation of the framework, we prove that quantum state verification can not be efficient and secure if one relies on the usual cut-and-choose technique, i.e. uses one of the rounds directly as output. Indeed we show that in the usual cut-and-chose regime a quantum state verification protocol is either to far from the ideal quantum state verification resource and therefore insecure or it is inefficient in the number of rounds. Our result is agnostic about the target state, the number of clients, and used resources, except for a few reasonable restrictions. These restrictions should only prevent the clients from preparing the target state themselves and are the motivation to use such a protocol in the first place. Although we only consider quantum state verification for our results, one finds direct implications for other primitives. One example is self-testing, in which a party prepares states with an untrusted device and measures them with a different untrusted device to verify the preparation. It is easy to see that this is even harder than quantum state verification as the measurement device is not trustworthy in self-testing, i.e. our result extends naturally to self-testing.

### 5.2 Discussion of the assumptions

The strength and generality of our results stems from the fact that we only use very few and simple assumptions. A fundamental assumption is the inability of the clients to prepare the state themselves. While we argue that this assumption comes naturally in the setting of quantum state verification, we modeled the categorical representation of this assumption stricter than necessary. For the proof to work, we only need that the client can not prepare states that are far enough from the target state, especially since we consider the asymptotic behavior. However, this restriction is complicated to model in a category for the client, which motivates the stricter modeling. We leave it open to future work to find less restrictive categories which implement the assumption.

Another assumption is that the clients output the state as received. While this seems to be the natural approach for verification, our work shows that it fails. In fact, questioning this assumption might lead to a workaround, which we discuss in more detail in the next section.

At last, one might be tempted to see the framework we used as an assumption. Because of that, we emphasize that one can find the same lower bounds for implementations of quantum state verification in other composability frameworks [16,17,25]. This fact is already reflected by how we present the proof. Indeed, measuring the output and input choice would be part of a distinguisher in other frameworks, such as abstract cryptography. The simulator would implement the attack $\mathbf{B}_\rho$ on the ideal resource but it would have to obey the same restrictions regarding the input of the ideal resource as the attack. In the end, the inequalities are the same.

However, the explicit and strict typing of the categorical composable cryptography framework allows for rigorous proofs without significant overhead once the user understands the framework. Further, the flexibility of choosing the appropriate attack model enables the user to analyse more restricted or complicated adversarial situations such as honest-but-curios or non-colluding adversaries. Our proof provides an example of this flexibility: One could restrict the attack model to i.i.d. attacks and still find the same result.

### 5.3 Possible workarounds and open questions

The lower bounds we presented are an inherent property of quantum state verification in a cut-and-choose fashion. They raise the question of how to circumvent this lower bound and what consequences follow. First, one should recall the implication of the result: One can not use quantum state verification in a modular manner for cryptography as one can not have efficiency and security. This no-go result holds not only for composable cryptography; in the recent work [5], the authors show that stand-alone secure protocols for quantum state verification are composable secure, where the $\varepsilon$ for composable security is a polynomial of the one for the stand-alone security definition. This lifting implies that our result extends to stand-alone security as well. Either way, the implication is about the modular use or as a protocol for its own sake. However, most times verification is used in the context of a larger protocol, which raises the question of what happens in a non-modular setting. We investigated this question to some extent by post-composing the ideal resource with different kinds of channels on the client side. We found that the lower bounds similarly extend to post-composition with unital channels and measurements in a basis. We present this result in more detail in the Appendix A.2. The idea of post-composition also leads to other approaches. One of these is error-detection: If the server has to prepare the target state in an error-detection code and the clients run the verification on the encoding, they can decode the output and eliminate or detect errors introduced by a dishonest server. Similar techniques are already used in verifiable delegated quantum computing [11], which already indicates that other primitives using verification in quantum cryptography could be affected by similar lower bounds. So, while we show that the naive approach to quantum state verification is doomed to fail, many open questions remain, and possible workarounds may exist.

## References

1. Acín, A., Brunner, N., Gisin, N., Massar, S., Pironio, S., Scarani, V.: Device-independent security of quantum cryptography against collective attacks. Physical Review Letters **98**(23) (Jun 2007). https://doi.org/10.1103/physrevlett.98.230501
2. Broadbent, A., Karvonen, M.: Categorical composable cryptography. In: Foundations of Software Science and Computation Structures (FoSSaCS). Lecture Notes in Computer Science, vol. 13242, pp. 161–183. Springer (2022). https://doi.org/10.1007/978-3-030-99253-8_9
3. Broadbent, A., Karvonen, M.: Categorical composable cryptography: extended version. Logical Methods in Computer Science **19**, 30:1–30:46 (2023). https://doi.org/10.46298/LMCS-19(4:30)2023

4. Coecke, B., Fritz, T., Spekkens, R.W.: A mathematical theory of resources. Information and Computation **250**, 59–86 (Oct 2016). https://doi.org/10.1016/j.ic.2016.02.008
5. Colisson, L., Markham, D., Yehia, R.: All graph state verification protocols are composably secure (2024)
6. Davidson, K.: $c^*$-Algebras by Example. American Mathematical Society (1996)
7. Gisin, N., Ribordy, G., Tittel, W., Zbinden, H.: Quantum cryptography. Rev. Mod. Phys. **74**, 145–195 (Mar 2002). https://doi.org/10.1103/RevModPhys.74.145
8. Hahn, F., de Jong, J., Pappa, A.: Anonymous quantum conference key agreement. PRX Quantum **1**, 020325 (Dec 2020). https://doi.org/10.1103/PRXQuantum.1.020325
9. Heunen, C., Vicary, J.: Categories for quantum theory: an introduction. Oxford University Press (Nov 2019)
10. Huot, M., Staton, S.: Quantum channels as a categorical completion. In: 2019 34th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS). IEEE (Jun 2019). https://doi.org/10.1109/lics.2019.8785700
11. Kashefi, E., Wallden, P.: Optimised resource construction for verifiable quantum computation. Journal of Physics A: Mathematical and Theoretical **50**(14), 145306 (mar 2017). https://doi.org/10.1088/1751-8121/aa5dac
12. Katz, J., Lindell, Y.: Introduction to Modern Cryptography, Second Edition. Chapman & Hall/CRC, 2nd edn. (2014)
13. Keyl, M.: Fundamentals of quantum information theory. Physics Reports **369**(5), 431–548 (Oct 2002). https://doi.org/10.1016/s0370-1573(02)00266-1
14. Laplaza, M.L.: A new result of coherence for distributivity. In: Coherence in Categories. pp. 214–235. Springer Berlin Heidelberg, Berlin, Heidelberg (1972)
15. Lydersen, L., Wiechers, C., Wittmann, C., Elser, D., Skaar, J., Makarov, V.: Hacking commercial quantum cryptography systems by tailored bright illumination. Nature Photonics **4**, 686–689 (2010). https://doi.org/https://doi.org/10.1038/nphoton.2010.214
16. Maurer, U.: Constructive cryptography – a new paradigm for security definitions and proofs. In: Mödersheim, S., Palamidessi, C. (eds.) Theory of Security and Applications. pp. 33–56. Springer Berlin Heidelberg, Berlin, Heidelberg (2012)
17. Maurer, U., Renner, R.: Abstract cryptography. In: Chazelle, B. (ed.) The Second Symposium on Innovations in Computer Science, ICS 2011. pp. 1–21. Tsinghua University Press (1 2011)
18. Morimae, T., Takeuchi, Y., Hayashi, M.: Verification of hypergraph states. Phys. Rev. A **96**, 062321 (Dec 2017). https://doi.org/10.1103/PhysRevA.96.062321
19. Pallister, S., Linden, N., Montanaro, A.: Optimal verification of entangled states with local measurements. Physical Review Letters **120**(17) (Apr 2018). https://doi.org/10.1103/physrevlett.120.170502
20. Pappa, A., Chailloux, A., Wehner, S., Diamanti, E., Kerenidis, I.: Multipartite entanglement verification resistant against dishonest parties. Physical Review Letters **108**(26) (Jun 2012). https://doi.org/10.1103/physrevlett.108.260502
21. Pirandola, S., Andersen, U.L., Banchi, L., Berta, M., Bunandar, D., Colbeck, R., Englund, D., Gehring, T., Lupo, C., Ottaviani, C., Pereira, J.L., Razavi, M., Shaari, J.S., Tomamichel, M., Usenko, V.C., Vallone, G., Villoresi, P., Wallden, P.: Advances in quantum cryptography. Adv. Opt. Photon. **12**(4), 1012–1236 (Dec 2020). https://doi.org/10.1364/AOP.361502
22. Preskill, J.: Quantum computing in the nisq era and beyond. Quantum **2**, 79 (Aug 2018). https://doi.org/10.22331/q-2018-08-06-79
23. Takeuchi, Y., Morimae, T.: Verification of many-qubit states. Physical Review X **8**(2) (Jun 2018). https://doi.org/10.1103/physrevx.8.021060

24. Unnikrishnan, A., Markham, D.: Verification of graph states in an untrusted network. Physical Review A **105**(5) (May 2022). https://doi.org/10.1103/physreva.105.052420
25. Unruh, D.: Universally composable quantum multi-party computation. In: Gilbert, H. (ed.) Advances in Cryptology – EUROCRYPT 2010. pp. 486–505. Springer Berlin Heidelberg, Berlin, Heidelberg (2010)
26. Watrous, J.: The Theory of Quantum Information. Cambridge University Press, 1 edn. (Apr 2018). https://doi.org/10.1017/9781316848142
27. Yehia, R., Diamanti, E., Kerenidis, I.: Composable security for multipartite entanglement verification. Physical Review A **103**(5) (May 2021). https://doi.org/10.1103/physreva.103.052609
28. Yu, X., Shang, J., Gühne, O.: Statistical methods for quantum state verification and fidelity estimation. Advanced Quantum Technologies **5**(5) (Mar 2022). https://doi.org/10.1002/qute.202100126
29. Zhu, H., Hayashi, M.: Efficient verification of pure quantum states in the adversarial scenario. Phys. Rev. Lett. **123**, 260504 (Dec 2019). https://doi.org/10.1103/PhysRevLett.123.260504
30. Šupić, I., Bowles, J.: Self-testing of quantum systems: a review. Quantum **4**, 337 (Sep 2020). https://doi.org/10.22331/q-2020-09-30-337

# A Appendix

## A.1 Detailed proof for general protocols

Before we restate the theorem, we note that we'll encounter expectation values of function of the following type in the proof, for $0<a<1$

$$f_a(X)=\sqrt{1-a^X}. \tag{A.1}$$

The functions $f_a$ are concave, since their second derivative is negative, indeed

$$f_a''(Z)=\frac{d^2f}{dX^2}(Z)=-\frac{a^Z\ln(a)^2\left(2-a^Z\right)}{4\left(\sqrt{1-a^Z}\right)^3}<0. \tag{A.2}$$

Since $f_a$ is concave, we can use Jensen's inequality for the expectation value of $f_a$ and we find

$$E(f_a(X))\leq f_a(E(X)). \tag{A.3}$$

**Theorem A.1 (No efficient secure state verification for entangled states).** *Let $\pi=\{\pi_S\}\cup\{\pi_i\}_{i\in C}$ be any protocol applied on resources $\bar{a}=(\mathcal{T}^{\times M},\mathcal{V}^{\times M},\mathcal{W})$, and $D$ is an upper bound for the number of rounds. $\pi(\bar{a})$ is an implementation with the following properties*

- *$K$ clients sample the number of rounds $r$ and the number of verification rounds $0\leq i\leq r$ from any joint distribution $p(r,i)$,*
- *the clients perform a measurement $\mu_{r,i}$ and accept the outcome if the result is $0$,*

- *if $r=i$, the clients prepare a state $\chi\in D\left(\bigotimes_{j=1}^{K}\mathbb{C}^{n_i}\right)$, where $\chi$ and the target state $\phi$ are separable respectively entangled with respect to a particular partition,*
- *if $r\neq i$, the clients output one of the $r-i$ states drawn to any unspecified distribution.*

*It then follows that there exists $\mathbf{A}_\rho\in\mathcal{A}(\pi)$ such that for all $\mathbf{B}_\rho\in\mathcal{A}\left(\mathrm{id}_{\mathcal{S}_{\phi,K}^{QSV}}\right)$ it holds that*

$$\frac{1}{2}d_\diamond\left(RF(\pi)(\bar{a}),RF(\sharp_S)\left(\mathcal{S}_{\phi,K}^{QSV}\right)\right)+\frac{1}{2}d_\diamond\left(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\left(\mathcal{S}_{\phi,K}^{QSV}\right)\right)\geq\varepsilon, \tag{A.4}$$

*where $\varepsilon=1/27N$, if $\phi$ is mixed and $\varepsilon=1/8\sqrt{N}$, if $\phi$ is pure and $N$ is the expected number of rounds.*

*Proof.* First we consider correctness, i.e. assume the source is honest. Again, we can bound the diamond-distance by composing with a $\mathcal{M}=Tr_{\bigotimes_{i\in C}M_{n_i}(\mathbb{C})}\otimes\mathrm{id}_I$ and find

$$\frac{1}{2}d_\diamond\left(RF(\pi)(\bar{a}),RF(\sharp_S)\left(\mathcal{S}_{\phi,K}^{QSV}\right)\right)\geq 1-\sum_{r=0}^{\infty}\sum_{i=0}^{r}p(r,i)\left\langle\mu_{r,i}(0)\middle|\phi^{\otimes i}\right\rangle. \tag{A.5}$$

If the source is dishonest, we consider a family of attacks $\{\mathbf{A}_\rho\}_{\rho\in D\left(\bigotimes_{i\in C}\mathbb{C}^{n_i}\right)}$ which prepares and inputs for every query the state $\rho$ on the source's side and implements $\pi_C$ as the clients are considered to be honest. As $\mathrm{dom}(\mathbf{A}_\rho(\bar{a}))=I$, the domain of any suitable attack $\mathbf{B}_\rho\in\mathcal{A}\left(\mathrm{id}_{\mathcal{S}_{\phi,K}^{QSV}}\right)$ must have the same domain, i.e. prepares and inputs a binary distribution $\{q(\rho),1-q(\rho)\}$, inputs this at the source's interface and acts as the identity on the client's side. Again, with $\mathcal{M}_D=\{\gamma(0)\oplus 1,\gamma(1)\oplus 0\}$ we find

$$\begin{aligned}&Pr[\mathcal{M}_D\circ\mathbf{A}_\rho(\bar{a})(I)=1]\\&=\langle\gamma(1)|\rho\rangle\sum_{r=1}^{\infty}\sum_{i=0}^{r-1}p(r,i)\left\langle\mu_{r,i}(0)\middle|\rho^{\otimes i}\right\rangle+\langle\gamma(1)|\chi\rangle\sum_{r=0}^{\infty}p(r,r)\left\langle\mu_{r,r}(0)\middle|\rho^{\otimes r}\right\rangle\end{aligned} \tag{A.6}$$

$$Pr[\mathcal{M}_D\circ\mathbf{B}_\rho\left(\mathcal{S}_{\phi,K}^{QSV}\right)(I)=1]=\langle\gamma(1)|\phi\rangle q(\rho) \tag{A.7}$$

With $\langle\gamma(1)|\phi\rangle=\langle\gamma(1)|\rho\rangle=\langle\gamma(1)|\chi\rangle=1$ we find:

$$\frac{1}{2}d_\diamond\left(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\left(\mathcal{S}_{\phi,K}^{QSV}\right)\right)\geq\left|\left(\sum_{r=0}^{\infty}\sum_{i=0}^{r}p(r,i)\left\langle\mu_{r,i}(0)\middle|\rho^{\otimes i}\right\rangle\right)-q(\rho)\right|=|\delta|, \tag{A.8}$$

with

$$\delta=\left(\sum_{r=0}^{\infty}\sum_{i=0}^{r}p(r,i)\left\langle\mu_{r,i}(0)\middle|\rho^{\otimes i}\right\rangle\right)-q(\rho). \tag{A.9}$$

With that, we find for any measurement $\gamma$

$$\begin{aligned}\frac{1}{2}d_\diamond\Big(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\Big(\mathcal{S}_{\phi,K}^{QSV}\Big)\Big) \geq &\Bigg|\langle\gamma(1)|\rho\rangle\sum_{r=1}^{\infty}\sum_{i=0}^{r-1}p(r,i)\langle\mu_{r,i}(0)\big|\rho^{\otimes i}\rangle\\ &+\langle\gamma(1)|\chi\rangle\sum_{r=0}^{\infty}p(r,r)\langle\mu_{r,r}(0)\big|\rho^{\otimes r}\rangle-\langle\gamma(1)|\phi\rangle q(\rho)\Bigg|\\ =&\Bigg|(\langle\gamma(1)|\rho\rangle-\langle\gamma(1)|\phi\rangle)\sum_{r=1}^{\infty}\sum_{i=0}^{r-1}p(r,i)\langle\mu_{r,i}(0)\big|\rho^{\otimes i}\rangle\\ &+(\langle\gamma(1)|\chi\rangle-\langle\gamma(1)|\phi\rangle)\sum_{r=0}^{\infty}p(r,r)\langle\mu_{r,r}(0)\big|\rho^{\otimes r}\rangle+\langle\gamma(1)|\phi\rangle(\delta)\Bigg|\\ \geq&\Bigg|(\langle\gamma(1)|\rho\rangle-\langle\gamma(1)|\phi\rangle)\sum_{r=1}^{\infty}\sum_{i=0}^{r-1}p(r,i)\langle\mu_{r,i}(0)\big|\rho^{\otimes i}\rangle\\ &+(\langle\gamma(1)|\chi\rangle-\langle\gamma(1)|\phi\rangle)\sum_{r=0}^{\infty}p(r,r)\langle\mu_{r,r}(0)\big|\rho^{\otimes r}\rangle\Bigg|-|\delta|,\end{aligned} \tag{A.10}$$

which implies

$$\begin{aligned}\frac{1}{2}d_\diamond\Big(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\Big(\mathcal{S}_{\phi,K}^{QSV}\Big)\Big) \geq \frac{1}{2}\Bigg|&(\langle\gamma(1)|\rho\rangle-\langle\gamma(1)|\phi\rangle)\sum_{r=1}^{\infty}\sum_{i=0}^{r-1}p(r,i)\langle\mu_{r,i}(0)\big|\rho^{\otimes i}\rangle\\ &+(\langle\gamma(1)|\chi\rangle-\langle\gamma(1)|\phi\rangle)\sum_{r=0}^{\infty}p(r,r)\langle\mu_{r,r}(0)\big|\rho^{\otimes r}\rangle\Bigg|\end{aligned} \tag{A.11}$$

We can choose the direction of $\rho$ and $\gamma(1)$ such that $\langle\gamma(1)|\chi\rangle\geq\langle\gamma(1)|\phi\rangle\leq\langle\gamma(1)|\rho\rangle$ and define $\lambda=\min(\langle\gamma(1)|\chi\rangle-\langle\gamma(1)|\phi\rangle,\langle\gamma(1)|\rho\rangle-\langle\gamma(1)|\phi\rangle)$:

$$\begin{aligned}\frac{1}{2}d_\diamond\Big(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\Big(\mathcal{S}_{\phi,K}^{QSV}\Big)\Big) \geq &\frac{\lambda}{2}\Bigg|\sum_{r=1}^{\infty}\sum_{i=0}^{r-1}p(r,i)\langle\mu_{r,i}(0)\big|\rho^{\otimes i}\rangle+\sum_{r=0}^{\infty}p(r,r)\langle\mu_{r,r}(0)\big|\rho^{\otimes r}\rangle\Bigg|\\ =&\frac{\lambda}{2}\Bigg|\sum_{r=0}^{\infty}\sum_{i=0}^{r}p(r,i)\langle\mu_{r,i}(0)\big|\rho^{\otimes i}\rangle\Bigg|.\end{aligned} \tag{A.12}$$

Now, we again consider both honesty configurations together

$$\begin{aligned}
&\frac{1}{2}d_\diamond\left(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\left(\mathcal{S}^{QSV}_{\phi,K}\right)\right)+\frac{1}{2}d_\diamond\left(RF(\pi)(\bar{a}),RF(\sharp_S)\left(\mathcal{S}^{QSV}_{\phi,K}\right)\right)\\
&\geq\frac{\lambda}{2}\left|\sum_{r=0}^{\infty}\sum_{i=0}^{r}p(r,i)\langle\mu_{r,i}(0)|\rho^{\otimes i}\rangle\right|+\left|1-\sum_{r=0}^{\infty}\sum_{i=0}^{r}p(r,i)\langle\mu_{r,i}(0)|\phi^{\otimes i}\rangle\right|\\
&\geq\frac{\lambda}{2}\left|1+\sum_{r=0}^{\infty}\sum_{i=0}^{r}p(r,i)\left(\langle\mu_{r,i}(0)|\rho^{\otimes i}\rangle-\langle\mu_{r,i}(0)|\phi^{\otimes i}\rangle\right)\right|\\
&\geq\frac{\lambda}{2}\left(1-\sum_{r=0}^{\infty}\sum_{i=0}^{r}p(r,i)\left|\langle\mu_{r,i}(0)|\rho^{\otimes i}\rangle-\langle\mu_{r,i}(0)|\phi^{\otimes i}\rangle\right|\right)\\
&\geq\frac{\lambda}{2}\left(1-\sum_{r=0}^{\infty}\sum_{i=0}^{r}p(r,i)\|\rho^{\otimes i}-\phi^{\otimes i}\|_1\right).
\end{aligned}\tag{A.13}$$

For mixed states we find with $1/2\|\rho-\phi\|_1=2/9N$

$$\begin{aligned}
&\frac{1}{2}d_\diamond\left(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\left(\mathcal{S}^{QSV}_{\phi,K}\right)\right)+\frac{1}{2}d_\diamond\left(RF(\pi)(\bar{a}),RF(\sharp_S)\left(\mathcal{S}^{QSV}_{\phi,K}\right)\right)\\
&\geq\frac{\lambda}{2}\left(1-\sum_{r=0}^{\infty}\sum_{i=0}^{r}p(r,i)\sqrt{1-(1-4/9N)^i}\right)\\
&\geq\frac{\lambda}{2}\left(1-\sum_{r=0}^{\infty}\sqrt{1-(1-4/9N)^r}\left(\sum_{i=0}^{r}p(r,i)\right)\right)\geq\frac{\lambda}{2}\left(1-\sqrt{1-(1-4/9N)^N}\right),
\end{aligned}\tag{A.14}$$

where we used Jensen's inequality with $N$ being the average number of rounds. If $N$ is large enough and the clients are not able to prepare $\phi$, we find the same expression as in the proof of Thm. 4.2 and find $1/2d_\diamond\left(\mathbf{A}_\rho(\bar{a}),\mathbf{B}_\rho\left(\mathcal{S}^{QSV}_{\phi,K}\right)\right)+1/2d_\diamond\left(\pi(\bar{a}),\sharp_S\left(\mathcal{S}^{QSV}_{\phi,K}\right)\right)\geq 1/27N$.

For pure states, we use again a pure state for $\rho$, $\rho=|\psi\rangle\langle\psi|$ and find if $N$ is large enough:

$$\begin{aligned}
&\frac{1}{2}d_\diamond\left(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\left(\mathcal{S}^{QSV}_{\phi,K}\right)\right)+\frac{1}{2}d_\diamond\left(RF(\pi)(\bar{a}),RF(\sharp_S)\left(\mathcal{S}^{QSV}_{\phi,K}\right)\right)\\
&\geq\frac{\sqrt{1-|\langle\psi|\phi\rangle|^2}}{2}\left(1-\sum_{r=0}^{\infty}\sum_{i=0}^{r}p(r,i)\sqrt{1-\sqrt{1-|\langle\psi|\phi\rangle|^{2i}}}\right)\\
&\geq\frac{\sqrt{1-|\langle\psi|\phi\rangle|^2}}{2}\left(1-\sum_{r=0}^{\infty}\sqrt{1-\sqrt{1-|\langle\psi|\phi\rangle|^{2r}}}\sum_{i=0}^{r}p(r,i)\right)\\
&\geq\frac{\sqrt{1-|\langle\psi|\phi\rangle|^2}}{2}\left(1-\sqrt{1-|\langle\psi|\phi\rangle|^{2N}}\right),
\end{aligned}\tag{A.15}$$

i.e. we find again $\frac{1}{2}d_\diamond\left(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)\left(\mathcal{S}^{QSV}_{\phi,K}\right)\right)+\frac{1}{2}d_\diamond\left(RF(\pi)(\bar{a}),RF(\sharp_S)\left(\mathcal{S}^{QSV}_{\phi,K}\right)\right)\geq 1/8\sqrt{N}$ □

### A.2 Extending to post-composition with channels

We also need to consider the situation where the clients try to overcome or at least improve the no-go result. The clients could do so by applying a channel either before or after the verification. Since for any channel $\phi$ it holds that

$$\frac{1}{2}\|\rho-\phi\|_1 \geq \frac{1}{2}\|\Lambda(\rho)-\Lambda(\phi)\|_1, \tag{A.16}$$

precomposing with a channel is of no help, as it could decrease the clients' chance of catching the source cheating. However, applying a channel after accepting the state could yield a good implementation as it cannot increase the distinguishing advantage. For a channel $\Lambda$ performed on the output of $\mathcal{V}_\phi^f$ in the case of no abort, we denote the resulting (ideal) resource $\Lambda \circ \mathcal{V}_\phi^f$. The question is then whether there are channels that can either improve the lower bound or avoid the no-go result all together.

Different channels lead to vastly different results. In fact, the analysis in the previous section breaks down completely for some channels. For example, consider a replacement channel:

$$\Lambda_{repl}^{\chi}(\rho) = Tr(\rho)\chi. \tag{A.17}$$

We find that $\frac{1}{2}\|\Lambda_{repl}^{\chi}(\rho) - \Lambda_{repl}^{\chi}(\phi)\|_1 = \frac{1}{2}\|\chi - \chi\|_1 = 0$, i.e. there is no chance a distinguisher could distinguish the implementation and the ideal resource. However, replacement channels are not interesting because they imply that the clients were able to prepare the desired state in the first place, making the source and therefore the verification obsolete.

Since the trace distance is unitarily invariant, the no-go result is upheld under post-composition with unitary channels. In the following we take a closer look at measurement and unital channels.

**Measurement channels** We consider the scenario where the clients, after accepting the verification, measure the state and output the outcome. Let the same happen in the ideal setting. Can this be composably secure with negligible distinguishing advantage? The verification works as in the general case. We just need to specify how the distinguisher could distinguish the outputs. We use the general setting described in section 4.2, but with $p(i,i) = 0$ for all $i$. Let $d > 2$ be the dimension of the output space, then we denote the measurement channel as

$$\mathcal{M}(\rho) = \sum_{j=1}^{d} |j\rangle\langle\xi_j|\rho|\xi_j\rangle\langle j|, \tag{A.18}$$

The distinguisher now fixes one $\tilde{j}$ and only outputs 1 if the measurement outcome is $\tilde{j}$. We restrict our analysis to pure states. For a large enough $N$, we have that there

is an attack $\mathbf{A}_\psi \in \mathcal{A}(\pi')$ such that for every attack $\mathbf{B}_\psi \in \mathcal{A}(\mathrm{id}_{\mathcal{S}_{\phi,K}^{QSV}})$ holds (A.15)

$$\begin{aligned}&\frac{1}{2}d_\diamond\left(RF(\mathbf{A}_\psi)(\bar{a}),RF(\mathbf{B}_\psi)\left(\mathcal{S}_{\phi,K}^{QSV}\right)\right)+\frac{1}{2}d_\diamond\left(\pi'(\bar{a}),\sharp_S\left(\mathcal{S}_{\phi,K}^{QSV}\right)\right)\\&\geq\frac{\sqrt{1-|\langle\psi|\phi\rangle|^2}}{2}\left(1-\sqrt{1-|\langle\psi|\phi\rangle|^{2N}}\right).=\end{aligned}\tag{A.19}$$

For our current setting the protocol $\pi$ then includes the measurement and if we set $|\xi\rangle=\left|\xi_{\vec{j}}\right\rangle$, we find that there is an attack $\mathbf{A}_\psi \in \mathcal{A}(\pi)$ such that for every attack $\mathbf{B}_\psi \in \mathcal{A}(\mathrm{id}_{\mathcal{M}\circ\mathcal{S}_{\phi,K}^{QSV}})$

$$\begin{aligned}&\frac{1}{2}d_\diamond\left(RF(\mathbf{A}_\psi)(\bar{a}),RF(\mathbf{B}_\psi)\left(\mathcal{M}\circ\mathcal{S}_{\phi,K}^{QSV}\right)\right)+\frac{1}{2}d_\diamond\left(RF(\pi)(\bar{a}),RF(\sharp_S)\left(\mathcal{M}\circ\mathcal{S}_{\phi,K}^{QSV}\right)\right)\\&\geq\frac{\left||\langle\xi|\psi\rangle|^2-|\langle\xi|\phi\rangle|^2\right|}{2}\left(1-\sqrt{1-|\langle\psi|\phi\rangle|^{2N}}\right).\end{aligned}\tag{A.20}$$

We choose $|\xi\rangle$ such that $1/\sqrt{2}>1/\sqrt{d}\geq|\langle\xi|\phi\rangle|\geq 0$. This is possible since we assume $d\geq 3$. We now choose $\pi/2>\eta\geq 0$ and $\pi/2>\theta\geq 0$ such that $\cos(\theta)=|\langle\xi|\phi\rangle|$ and $\cos(\eta)=|\langle\psi|\phi\rangle|$. Next, we fix $|\psi\rangle$ to be in the plane spanned by $|\phi\rangle$ and $|\xi\rangle$. We define the following basis

$$|b_0\rangle=|\phi\rangle\tag{A.21}$$

$$|b_1\rangle=\frac{|\xi\rangle-\langle\phi|\xi\rangle|\phi\rangle}{\sqrt{1-|\langle\phi|\xi\rangle|^2}}=\frac{|\xi\rangle-\langle\phi|\xi\rangle|\phi\rangle}{\sqrt{1-\cos(\theta)^2}}.\tag{A.22}$$

With $\langle\xi|\phi\rangle=\cos(\theta)e^{i\alpha}$ we can express $|\psi\rangle$ in the basis as

$$|\psi\rangle=\cos(\eta)e^{-i\alpha}|b_0\rangle+\sin(\eta)|b_1\rangle,\tag{A.23}$$

With that we find

$$\langle\xi|\psi\rangle=\cos(\eta)\cos(\theta)+\sin(\eta)\frac{1-\cos(\theta)^2}{\sqrt{1-\cos(\theta)^2}}\tag{A.24}$$

$$=\cos(\eta)\cos(\theta)+\sin(\eta)\sin(\theta)=\cos(\theta-\eta).\tag{A.25}$$

We can now obtain a lower bound on $\left||\langle\xi|\psi\rangle|^2-|\langle\xi|\phi\rangle|^2\right|$ as follows

$$\left||\langle\xi|\psi\rangle|^2-|\langle\xi|\phi\rangle|^2\right|=\left|\cos(\theta-\eta)^2-\cos(\theta)^2\right| \tag{A.26}$$

$$=\left|\frac{e^{2i\theta-2i\eta}+2+e^{2i\eta-2i\theta}}{4}-\frac{e^{2i\theta}+2+e^{-2i\theta}}{4}\right| \tag{A.27}$$

$$=\left|\frac{e^{2i\theta-2i\eta}+e^{2i\eta-2i\theta}-e^{2i\theta}-e^{-2i\theta}}{4}\right| \tag{A.28}$$

$$=\left|\frac{\left(e^{-2i\eta}-1\right)e^{2i\theta}+\left(e^{2i\eta}-1\right)e^{-2i\theta}}{4}\right| \tag{A.29}$$

$$=\left|\frac{\left(e^{-i\eta}-e^{i\eta}\right)e^{2i\theta-i\eta}+\left(e^{i\eta}-e^{-i\eta}\right)e^{-2i\theta+i\eta}}{4}\right| \tag{A.30}$$

$$=\left|\frac{\left(e^{-i\eta}-e^{i\eta}\right)\left(e^{i(2\theta-\eta)}-e^{-i(2\theta+\eta)}\right)}{4}\right| \tag{A.31}$$

$$=|\sin(\eta)\sin(2\theta-\eta)|=|\sin(\eta)||\sin(2\theta-\eta)| \tag{A.32}$$

From $0\leq\cos(\theta)<1/\sqrt{2}$ it follows that $\pi/2>\theta>\pi/4$ and we choose $\eta$ such that $\sin(\eta)=\frac{1}{2\sqrt{N}}$ and $\sin(2\theta-\eta)\geq\sin(\eta)$. Using A.20

$$\begin{aligned}&\frac{1}{2}d_\diamond\left(RF(\mathbf{A}_\psi)(\bar{a}),RF(\mathbf{B}_\psi)\left(\mathcal{M}\circ\mathcal{S}_{\phi,K}^{QSV}\right)\right)+\frac{1}{2}d_\diamond\left(RF(\pi)(\bar{a}),RF(\sharp_S)\left(\mathcal{M}\circ\mathcal{S}_{\phi,K}^{QSV}\right)\right)\\&\geq\frac{\sin^2(\eta)}{2}\left(1-\sqrt{1-\cos(\eta)^{2N}}\right)\\&=\frac{\sin^2(\eta)}{2}\left(1-\sqrt{1-(1-\sin(\eta)^2)^N}\right)\geq\frac{1}{16N}.\end{aligned} \tag{A.33}$$

**Unital channels** A channel is unital when it preserves the identity. That is, for a space $M_a(\mathbb{C})$ a unital channel $\Lambda$ maps $M_a(\mathbb{C})$ to $M_a(\mathbb{C})$ and it holds that

$$\Lambda(\mathbb{1}_{M_a(\mathbb{C})})=\mathbb{1}_{M_a(\mathbb{C})}. \tag{A.34}$$

Again, we consider a setting as described in section 4.2 with $p(i,i)=0$ where we post-compose the protocol with a unital channel $\Lambda$. We need to find a lower bound on the distinguishing advantage after applying the unital channel. Using the bound in Theorem 4.4 for mixed states, because $\Lambda(\phi)$ might be mixed even when $\phi$ is pure, we find for arbitrary $\rho$

$$\begin{aligned}&\frac{1}{2}d_\diamond\left(RF(\mathbf{A}_\rho)(\bar{a}),RF(\mathbf{B}_\rho)(\Lambda\circ\mathcal{S}_{\phi,K}^{QSV})\right)+\frac{1}{2}d_\diamond\left(RF(\pi)(\bar{a}),RF(\sharp_S)\left(\Lambda\circ\mathcal{S}_{\phi,K}^{QSV}\right)\right)\geq\\&\qquad\qquad\frac{\|\Lambda(\rho)-\Lambda(\phi)\|_1}{4}\left(1-\sqrt{1-(1-\|\rho-\phi\|_1)^N}\right).\end{aligned} \tag{A.35}$$

We can freely choose $\rho$, and set the following:

$$\rho = \alpha\phi + (1-\alpha)\frac{\mathbb{1}-\phi}{d-1}. \tag{A.36}$$

And we find

$$\|\Lambda(\phi)-\Lambda(\rho)\|_1 = \left\|\phi' - \left(\alpha\phi' + (1-\alpha)\frac{\mathbb{1}-\phi'}{d-1}\right)\right\|_1 \tag{A.37}$$

$$= \left\|(1-\alpha)\left(\phi' - \frac{\mathbb{1}-\phi'}{d-1}\right)\right\|_1 = (1-\alpha)\left\|\frac{(d-1)\phi' - \mathbb{1} + \phi'}{d-1}\right\|_1 \tag{A.38}$$

$$= (1-\alpha)\frac{d}{d-1}\left\|\phi' - \frac{\mathbb{1}}{d}\right\|_1. \tag{A.39}$$

With $\beta = 1-\alpha$, we obtain

$$\frac{1}{2}\|\Lambda(\phi)-\Lambda(\rho)\|_1 = \frac{\beta d}{d-1}\frac{1}{2}\|\Lambda(\phi) - {}^{\mathbb{1}}/_{d}\|_1. \tag{A.40}$$

Similarly, the trace distance of the inputs is then

$$\frac{1}{2}\|\rho-\phi\|_1 = \frac{1}{2}\left\|\phi - \left(\alpha\phi + (1-\alpha)\frac{\mathbb{1}-\phi}{d-1}\right)\right\|_1 = \frac{\beta d}{(d-1)}\frac{1}{2}\|\phi - {}^{\mathbb{1}}/_{d}\|_1. \tag{A.41}$$

We define

$$\omega = \frac{d}{d-1}\frac{1}{2}\|\phi - \mathbb{1}/d\|_1, \tag{A.42}$$

$$\omega' = \frac{d}{d-1}\frac{1}{2}\|\Lambda(\phi) - \mathbb{1}/d\|_1. \tag{A.43}$$

We can then rewrite A.35 as follows

$$\frac{1}{2}d_\diamond\left(RF(\mathbf{A}_\rho)(\bar{a}), RF(\mathbf{B}_\rho)(\Lambda\circ\mathcal{S}^{QSV}_{\phi,K})\right) + \frac{1}{2}d_\diamond\left(RF(\pi)(\bar{a}), RF(\sharp_S)\left(\Lambda\circ\mathcal{S}^{QSV}_{\phi,K}\right)\right)$$
$$\geq \frac{\omega'\beta}{2}\left(1 - \sqrt{1-(1-2\omega\beta)^N}\right). \tag{A.44}$$

We set $\beta = \frac{1}{2\omega N}$, and we can assume that $N \geq \frac{1}{2\omega N}$, since $\phi$ does not depend on the number of rounds. Using $(1-\frac{\beta}{k})^k \geq 1-\beta$ for $k\in\mathbb{N}$ and $|\beta| \leq k$ we find:

$$\frac{1}{2}d_\diamond\left(RF(\mathbf{A}_\rho)(\bar{a}), RF(\mathbf{B}_\rho)(\Lambda\circ\mathcal{S}^{QSV}_{\phi,K})\right) + \frac{1}{2}d_\diamond\left(RF(\pi)(\bar{a}), RF(\sharp_S)\left(\Lambda\circ\mathcal{S}^{QSV}_{\phi,K}\right)\right)$$
$$\geq \frac{\omega'}{2(2\omega N)}\left(1 - \sqrt{1-\left(1-\frac{1}{N}\right)^N}\right) \geq \frac{\omega'}{4\omega N}. \tag{A.45}$$

# Appendix E

# Paper V

A. Matsui, I. Obi, G. Sabbagh, L. Torres, D. Kessler, J. F. Meleiro, and K. Muroya. "A Critical Pair Enumeration Algorithm for String Diagram Rewriting". In: *Applied Category Theory (ACT)* (2025). to appear

# A Critical Pair Enumeration Algorithm for String Diagram Rewriting*

Anna Matsui
Johns Hopkins University, USA
amatsui1@jhu.edu

Innocent Obi
University of Washington, USA
innoobi@cs.washington.edu

Guillaume Sabbagh
University of Technology of Compiègne, France†
guillaume.sabbagh@utc.fr

Leo Torres
Universidad Nacional de Córdoba, Argentina
leo.torres@mi.unc.edu.ar

Diana Kessler
Tallinn University of Technology, Estonia‡
diana-maria.kessler@taltech.ee

Juan F. Meleiro
University of São Paulo, Brazil
juan.meleiro@usp.br

Koko Muroya
National Institute of Informatics, Japan§
kmuroya@nii.ac.jp

Critical pair analysis provides a convenient and computable criterion of confluence, which is a fundamental property in rewriting theory, for a wide variety of rewriting systems. Bonchi et al. showed validity of critical pair analysis for rewriting on string diagrams in symmetric monoidal categories. This work aims at automation of critical pair analysis for string diagram rewriting, and develops an algorithm that implements the core part of critical pair analysis. The algorithm enumerates all critical pairs of a given left-connected string diagram rewriting system, and it can be realised by concrete manipulation of hypergraphs. We prove correctness and exhaustiveness of the algorithm, for string diagrams in symmetric monoidal categories without a Frobenius structure.

## 1 Introduction

### 1.1 Rewriting Theory and Critical Pair Analysis

Mathematical reasoning often involves derivation of a (complex) equation from known (typically simpler) equations, which is sometimes called **equational reasoning**. Equations can be between various mathematical objects, e.g. terms, programs, graphs, processes, and objects/morphisms in a category.
Rewriting theory has been established (see e.g. [1, 14]), with equational reasoning as one application. The starting point is to turn known equations $a = a'$ into directed[1] **rewrite rules** $a \multimap a'$. Each step $b \to b'$ of rewrite modifies a part of $b$ by applying one rewrite rule. Derivation of an equation $c \stackrel{?}{=} c'$ then boils down to finding some $d$ with two chains $c \to \cdots \to d$ and $c' \to \cdots \to d$ of rewrites. These chains altogether imply a chain of equations $c = \cdots = d = \cdots = c'$, which concludes the desired equation $c \stackrel{?}{=} c'$.

*We are grateful to organisers of Adjoint School 2024 in which our collaboration started.
†G.S. was funded by BNP Paribas CIB EMEA and the French Ministry of Research under CIFRE project No. 2021/1502.
‡D.K. was funded by an Advanced Research + Invention Agency (ARIA) Safeguarded AI: TA1.1 Theory grant.
§K.M. was funded by JSPS, KAKENHI Project No. 22K17850, Japan.
[1]The direction is typically chosen so that $a'$ is "simpler" than $a$.

Submitted to:
ACT 2025

**Confluence** is a fundamental property in rewriting theory, intuitively meaning that ordering of rewrites does not matter. Rewrites $\to$ are said to be confluent if any two diverging chains $b \leftarrow \cdots \leftarrow a \to \cdots \to b'$ of rewrites are **joinable**, that is, there exists $c$ with converging chains $b \to \cdots \to c \leftarrow \cdots \leftarrow b'$. **Local confluence** is a variant of confluence in which the two diverging chains are in fact given by two single rewrites, i.e. $b \leftarrow a \to b'$.

For some pair of two diverging rewrites $b \leftarrow a \to b'$, joinability is obvious. An example is the so-called **parallel** case, namely when the two rewrites change different, independent, parts of $a$. Consequently, checking local confluence boils down to analysing joinability of non-parallel pairs.

**Critical pair analysis** is a well-established technique for automatically checking local confluence, providing a convenient and computable criterion. It reduces local confluence to joinability of **critical pairs** that are finitely many representatives of non-parallel pairs. Critical pairs can be enumerated from a given set of rewrite rules $a \multimap a'$. This enumeration plays a central role in automating local-confluence check.

## 1.2 String Diagram Rewriting

**String diagrams** [11, 15] provide a graphical syntax of category theory. They are particularly useful in equational reasoning on morphisms of a category, because they trivialise certain equations as graph isomorphism.

Rewriting theory for string diagrams has been developed by Bonchi et al. [2, 3, 4], targeting at string diagrams for symmetric monoidal categories (with and without a Frobenius structure). String diagrams are combinatorially represented using hypergraphs, and rewrites on string diagrams are categorically modelled using **double pushout rewriting** (DPO rewriting in short) [6]. A key concept in string diagram rewriting theory is that of **interface**. An interface of a hypergraph specifies how other hypergraphs can be connected to the hypergraph.

Bonchi et al. showed validity of critical pair analysis for string diagram rewriting [4]. They defined critical pairs for an adaptation of DPO rewriting (dubbed **convex**[2] **DPOI rewriting**) that takes interface into account, and proved that joinability of critical pairs implies local confluence. Their development focuses on a theoretical side, and automation, which is an important aspect of critical pair analysis, has not been investigated.

## 1.3 Contributions

We aim at automation of critical pair analysis for string diagram rewriting, and develop an algorithm that implements the core part of the automation. The algorithm enumerates of all critical pairs for a given set of DPOI rewrite rules. We focus on the so-called **left-connected** DPOI rewrite rules [3, 4]. Left-connectivity allows us to reducing enumeration of critical pairs to enumeration of certain cospans in the category of hypergraphs. While it is an arguably powerful restriction, it still accommodates various concrete string diagram rewriting systems from the literature [8, 9, 12].

Each critical pair is associated with two DPOI rewrite rules, which are given by spans $L_1 \leftarrow K_1 \to R_1$ and $L_2 \leftarrow K_2 \to R_2$ in the category of hypergraphs. Thanks to left-connectivity, the critical pair is uniquely determined by a certain cospan of the form $L_1 + L_2 \twoheadrightarrow S \leftarrow J$. Its left leg is, in particular, an epimorphism given by the coupling of monomorphisms.

Our key idea is that the cospan, in particular the hypergraph $S$, can be generated by suitably **gluing** hyperedges and nodes of $L_1 + L_2$ (i.e. the hypergraph that puts $L_1$ and $L_2$ in parallel). We observe that the

[2] Convexity is for dealing with the absence of a Frobenius structure.

gluing process can be realised in two steps: (1) repeatedly merge a hyperedge from $L_1$ with a hyperedge from $L_2$, and (2) repeatedly merge a node from $L_1$ with a node from $L_2$ without merging any hyperedges. Our contributions can be summarised as follows.

- We develop an algorithm (Algo. 3) that enumerates all critical pairs of a given set of left-connected DPOI rewrite rules by implementing the two-fold gluing process.
- We prove that the algorithm generates all critical pairs and nothing else (correctness and exhaustiveness; Thm. 3.9).
- We provide a proof-of-concept Haskell implementation[3].
- We present an optimised algorithm (Algo. 4) that enumerates less but sufficient critical pairs to decide local confluence by only performing the first step of the two-fold gluing process.

In the rest of this paper, Sec. 2 recalls relevant concepts (e.g. hypergraph, interface, DPOI rewriting, critical pair) from string diagram rewriting theory [2, 3, 4]. Sec. 3 presents our main contribution, the critical pair enumeration algorithm with a proof of its correctness and exhaustiveness. Sec. 4 provides the optimised algorithm. Examples and some proofs can be found in Appendix.

**Related work.** For term rewriting, rewrite rules typically use variables as placeholders to succinctly represent a family of rewrite rules, e.g. $x+y \multimap y+x$. To deal with variables, enumeration of critical pairs employ a technique called **unification**. In contrast, for string diagram rewriting, rewrite rules are always concrete without placeholders. We can therefore take a direct approach and generate a critical pair by suitably gluing hyperedges and nodes of left-hand sides of rewrite rules. There are some attempts to enumerating critical pairs for variations of graph rewriting (graph transformation), e.g. [13, 5, 10].

## 2 Critical Pairs for String Diagram Rewriting

We denote the composition of morphisms $f\colon A \to B$ and $g\colon B \to C$ by $f;g$. We denote coprojections of a coproduct by $\iota_1, \iota_2$. Given a set $A$, the free monoid on $A$ is denoted by $A^*$. For a function $f\colon A \to B$, we denote as $f^*\colon A^* \to B^*$ the pointwise application of $f$ over a list of elements. Let $\mathbb{N}$ be the set of natural numbers.

### 2.1 Hypergraphs with Interface

When a symmetric monoidal category is equipped with a Frobenius structure, string diagrams in the category can be combinatorially represented as (edge-labelled) hypergraphs with interface [2].

**Definition 2.1** (Hypergraphs)**.** A (directed) **hypergraph** is a tuple $G = (V, E, s : E \to V^*, t : E \to V^*)$ where $V$ and $E$ are finite sets of nodes and hyperedges, $s$ maps each hyperedge to a list of source nodes and $t$ maps each hyperedge to a list of target nodes. The **arity** of a hyperedge is the number of its sources, the **coarity** of a hyperedge is the number of its targets.
We refer to $V$ as $Nodes(G)$ and $E$ as $HEdges(G)$.
Let $\sigma$ be an alphabet, a **signature** $\Sigma$ on $\sigma$ is a subset of $\sigma \times \mathbb{N} \times \mathbb{N}$. A triplet $(x, n, m)$ represents a label $x$ for morphisms with arity $n$ and coarity $m$.
A **$\Sigma$-labelled hypergraph** ($\Sigma$-hypergraph in short) is a hypergraph equipped with a labelling function $l : E \to \Sigma$ such that $l_E$ maps a hyperedge with arity $n$ and coarity $m$ to a triplet $(x, n, m)$.

[3] available online at `https://github.com/GuiSab/hypergraphrewriting`

**Definition 2.2** (Σ-hypergraph morphisms)**.** A **Σ-hypergraph morphism** between Σ-hypergraphs $(V_0, E_0, s_0, t_0, l_0)$ and $(V_1, E_1, s_1, t_1, l_1)$ is a pair of functions $f_V : V_0 \to V_1$ and $f_E : E_0 \to E_1$ that respects sources, targets and labels; that is, that satisfies $f_V^* \circ s_0 = s_1 \circ f_E$, $f_V^* \circ t_0 = t_1 \circ f_E$ and $l_0 = l_1 \circ f_E$.

Given a signature Σ, Σ-hypergraphs and Σ-hypergraph morphisms form a category $\mathbf{Hyp}_\Sigma$. It has all small limits and colimits (since it is a presheaf category); in particular, it has pushouts, coproducts and coequalizers. We can spell them out in set-theoretic terms, which makes them suitable for an algorithmic implementation; see Appendix B for details.
Interface specifies nodes of a hypergraph to which other hypergraphs can be connected.

**Definition 2.3** (Σ-hypergraph with interfaces)**.** A discrete Σ-hypergraph is a Σ-hypergraph with no hyperedges. A **Σ-hypergraph with interface** is a cospan $n \longrightarrow G \longleftarrow m$ in $\mathbf{Hyp}_\Sigma$ where $n$, $m$ are finite discrete Σ-hypergraphs.

The discrete hypergraphs $n$ and $m$ specify input interface and output interface, respectively. We sometimes identify a Σ-hypergraph with interface $n \longrightarrow G \longleftarrow m$ by a single morphism $G \leftarrow n + m$.
For a general symmetric monoidal category without a Frobenius structure, the combinatorial representation of string diagrams requires extra conditions on hypergraphs [3]: monogamy and acyclicity.

**Definition 2.4** (Paths)**.** A **path** $P$ in a Σ-hypergraph is a list of hyperedges $[e_1, e_2, \cdots, e_n]$ such that for every consecutive pair of hyperedges $(e_k, e_{k+1})$, there is at least one target of $e_k$ equal to a source of $e_{k+1}$. A **cycle** $C$ in a Σ-hypergraph is a path such that at least one source of $e_1$ is a target of $e_n$.

**Definition 2.5** (Monogamous acyclicity)**.** A Σ-hypergraph is **monogamous acyclic** (ma-hypergraph) if

1. it contains no cycles (acyclicity) ;
2. every node has at most in- and out- degree 1 (monogamy).

Here in- (out-) degree of a node $v$ in a Σ-hypergraph $H$ is the number of pairs $(e, i)$ where $e$ is a hyperedge of $H$ with $v$ as its $i$-th target (source). We call input nodes those with in-degree 0, denoted by $in(H)$. Similarly, output nodes have out-degree 0 and are denoted by $out(H)$.
A Σ-hypergraph with interface $n \xrightarrow{f} H \xleftarrow{g} m$ is **monogamous acyclic** (ma-cospan) if $H$ is an ma-hypergraph, $f$ is mono and its image is $in(H)$, and $g$ is mono and its image is $out(H)$.

## 2.2 Convex DPOI Rewriting

Rewriting on string diagrams can be modelled categorically [2, 3], by adapting DPO rewriting [6]. We first recall DPO rewriting in $\mathbf{Hyp}_\Sigma$.
A rewrite rule is a span $L \longleftarrow K \longrightarrow R$ in $\mathbf{Hyp}_\Sigma$. A rewrite system $\mathscr{R}$ is a finite set of rewrite rules. We say that a Σ-hypergraph $G$ rewrites into a Σ-hypergraph $H$ if there exists a rule $L \leftarrow K \rightarrow R$, a morphism $L \xrightarrow{m} G$ (called **match**) and an object $C \in \mathbf{Hyp}_\Sigma$ such that the following two squares are pushouts:

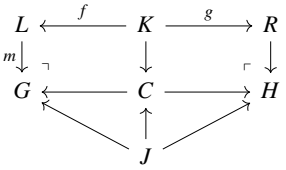
$$\begin{array}{ccccc} L & \xleftarrow{f} & K & \xrightarrow{g} & R \\ {\scriptstyle m}\downarrow & \lrcorner & \downarrow & \lrcorner & \downarrow \\ G & \longleftarrow & C & \longrightarrow & H \end{array}$$

The above rewrite works as follows. Computing the pushout complement removes the image of $L$ (the left-hand side of the rewrite rule) in $G$ while keeping the image of $K$ intact. By computing the pushout of $C \longleftarrow K \xrightarrow{g} R$, we glue $R$ and $C$ along the image of $K$, thus replacing the image of $L$ in $G$ with $R$. More intuitively, what this procedure does is to take away the part that corresponds to the matching of the left-hand side, $L$, of a rewrite rule and replace it by its right-hand side, $R$. See Example C.1 in Appendix.
In this work we focus on **left-connected** rewrite rules.

**Definition 2.6** (Strong connectivity)**.** An ma-hypergraph $G$ is **strongly connected** if for every input $x \in in(G)$ and output $y \in out(G)$ there exists a path from $x$ to $y$ in $G$.

A **left-connected rewrite rule** is a span $L \xleftarrow{[i_L,o_L]} I+O \xrightarrow{[i_R,o_R]} R$ such that $I \xrightarrow{i_L} L \xleftarrow{o_L} O$ and $I \xrightarrow{i_R} R \xleftarrow{o_R} O$ are ma-cospans, $[i_L,o_L]$ is mono (we say that the rule is left-linear) and $L$ is strongly connected.
A **left-connected rewriting system** is a set of left-connected rewrite rules.

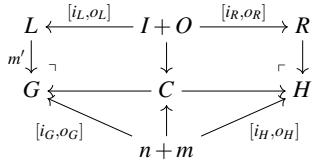
The first adaptation of DPO rewriting for string diagrams is to accommodate interface. This yields DPOI rewriting [2]. Given two hypergraphs with interfaces, $G \leftarrow J$ and $H \leftarrow J$, we say that $G$ rewrites into $H$ if there exists a rewrite rule $L \xleftarrow{f} K \xrightarrow{g} R$, a match $L \xrightarrow{m} G$ and a hypergraph with interface $C \leftarrow J$, such that the squares below are pushouts and the whole diagram commutes:




The second adaptation of DPO rewriting is to impose convexity on matches. This is necessary to deal with the absence of a Frobenius structure [3].

**Definition 2.7** (Convex matches)**.** A $\Sigma$-hypergraph morphism $m : L \to G$ is a **convex match** if it is mono and its image $m(L)$ is convex, i.e. for any nodes $v, v'$ in $m(L)$ and any path $p$ from $v$ to $v'$ in $G$, every hyperedge in $p$ is also in $m(L)$.

**Definition 2.8** (Convex rewriting)**.** Given a left-connected rewrite system $\mathscr{R}$, we say that an ma-cospan $n \xrightarrow{i_G} G \xleftarrow{o_G} m$ **rewrites convexly** into $n \xrightarrow{i_H} H \xleftarrow{o_H} m$ if there is a convex match $m' : L \to G$, a rewrite rule $L \xleftarrow{[i_L,o_L]} I+O \xrightarrow{[i_R,o_R]} R$ in $\mathscr{R}$ and a $\Sigma$-hypergraph $C$ such that the following diagram commutes, the left square is a boundary complement [3, Definition 30] and the right square is a pushout:




(1)

We will write $n \xrightarrow{i_G} G \xleftarrow{o_G} m \Rrightarrow_{\mathscr{R}} n \xrightarrow{i_H} H \xleftarrow{o_H} m$ and call it a **derivation**.

Thanks to left-connectedness, a derivation can be uniquely determined by an ma-cospan $n \to G \leftarrow m$, a mono match $L \xrightarrow{m'} G$ and a rewrite rule $L \leftarrow I+O \to R$.

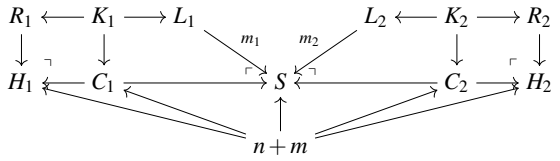
**Proposition 2.9.** *In left-connected rewrite systems, the boundary complement condition is always met. In left-connected rewrite systems, a mono match is always convex.*

**Proposition 2.10.** *In left-connected rewrite systems, given a rewrite rule and a mono matching the pushout complement C always uniquely exists.*

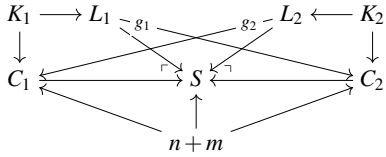
## 2.3 Critical Pairs

We finally recall the definition of critical pairs [4].

**Definition 2.11** (Critical pairs)**.** Let $\mathscr{R}$ be a left-connected rewrite system, and $L_1 \longleftarrow K_1 \longrightarrow R_1$ and $L_2 \longleftarrow K_2 \longrightarrow R_2$ be its two rewrite rules. Consider two derivations with common source $n \to S \leftarrow m$:




1. We say that $(n \to H_1 \leftarrow m) \Leftarrow_{\mathscr{R}} (n \to S \leftarrow m) \Rightarrow_{\mathscr{R}} (n \to H_2 \leftarrow m)$ is a **pre-critical pair** if $[m_1, m_2] : L_1 + L_2 \to S$ is epi.
2. The pre-critical pair is **joinable** if there exists $W$ such that $(n \to H_1 \leftarrow m) \overset{*}{\Rightarrow}_{\mathscr{R}} (n \to W \leftarrow m) \overset{*}{\Leftarrow}_{\mathscr{R}} (n \to H_2 \leftarrow m)$.
3. The pre-critical pair is a **parallel pair** if there exist $g_1 : L_1 \to C_2$ and $g_2 : L_2 \to C_1$ making the diagram below commute:




4. The pre-critical pair is a **critical pair** if it is not parallel.

Thanks to Prop. 2.9 and Prop. 2.10, for left-connected rewrite systems, the pre-critical pair $(n \to H_1 \leftarrow m) \Leftarrow_{\mathscr{R}} (n \to S \leftarrow m) \Rightarrow_{\mathscr{R}} (n \to H_2 \leftarrow m)$ can uniquely be determined by a cospan (dubbed **cp-cospan**) $L_1 + L_2 \twoheadrightarrow S \leftarrow I + O$ such that: (i) $L_1 + L_2 \twoheadrightarrow S$ is an epimorphism given by the coupling of two mono matches from $L_1$ and $L_2$ to $S$; and (ii) $I \to S \leftarrow O$ is a ma-cospan.

## 3 A Critical Pair Enumeration Algorithm

Our goal is to enumerate automatically all critical pairs for a given left-connected rewrite system $\mathscr{R}$. To do so, we have to enumerate all relevant epimorphisms $L_i + L_j \twoheadrightarrow S$ where $L_i$ and $L_j$ are left hand sides in rewrite rules of $\mathscr{R}$. We will glue different nodes and hyperedges of $L_i + L_j$ to enumerate these epimorphisms.

**Definition 3.1.** Let $L_1$ and $L_2$ be two $\Sigma$-hypergraphs. A **gluing scheme** is given by a $\Sigma$-hypergraph $G$ and two $\Sigma$-hypergraph morphisms $g_1 : G \to L_1 + L_2$ and $g_2 : G \to L_1 + L_2$. The **gluing** is the coequalizer of $g_1$ and $g_2$. For two nodes (or hyperedges) $x$ and $x'$ of $L_1 + L_2$, they are **glued** if there exists a node (or a hyperedge) $y$ of $G$ such that $g_1(y) = x$ and $g_2(y) = x'$.

$$G \underset{g_2}{\overset{g_1}{\rightrightarrows}} L_1 + L_2 \xrightarrow[\varepsilon]{} \mathtt{coeq}_G(g_1, g_2)$$

Each gluing scheme $(g_1 : G \to L_1 + L_2, g_2 : G \to L_1 + L_2)$ induces a cospan $L_1 + L_2 \twoheadrightarrow \mathtt{coeq}_G(g_1, g_2) \xleftarrow{[\subseteq,\subseteq]} \mathit{in}(\mathtt{coeq}_G(g_1, g_2)) + \mathit{out}(\mathtt{coeq}_G(g_1, g_2))$. We call the coequaliser $\mathtt{coeq}_G(g_1, g_2)$ **candidate source**. As observed in Sec. 2.3, this cospan uniquely determines a pre-critical pair if it is a cp-cospan. Now the question is: what are necessary conditions on the gluing scheme so that the cospan becomes a cp-cospan and hence induces a (pre-)critical pair?

We first observe that a gluing scheme should not glue nodes nor hyperedges within $L_1$ and $L_2$.

**Proposition 3.2.** *If two nodes from $L_1$ (resp. $L_2$) are glued, then the gluing scheme does not yield a critical pair. If two hyperedges from $L_1$ (resp. $L_2$) are glued, then the gluing scheme does not yield a critical pair.*

*Proof.* If two nodes from $L_1$ are glued in $\mathtt{coeq}_G(g_1, g_2)$, then $\iota_1; \varepsilon$ is not mono and is thus not a valid convex matching. The same proof works for hyperedges. □

Secondly we observe that nodes, separately from $L_1$ and $L_2$, should be glued in a specific way.

**Proposition 3.3.** *If a node $A$ from $L_1$ and a node $B$ from $L_2$ are glued and the gluing scheme yields a valid critical pair, then either*

- *$A$ and $B$ are the $k$-th source of a glued hyperedge ;*
- *$A$ and $B$ are the $i$-th target of a glued hyperedge ;*
- *$A$ is an output of $L_1$, $B$ is an input of $L_2$.*
- *$A$ is an input of $L_1$, $B$ is an output of $L_2$;*

**Proposition 3.4.** *If a node $A$ from $L_1$ and a node $B$ from $L_2$ are glued and the gluing scheme yields a valid critical pair, then:*

- *if $A$ is an output of $L_1$, $B$ is an input of $L_2$, then no input of $L_1$ is glued to an output of $L_2$;*
- *if $A$ is an input of $L_1$ and $B$ is an output of $L_2$, then no output of $L_1$ is glued to an input of $L_2$.*

*Proof.* Suppose that a node $A$ from $L_1$ and a node $B$ from $L_2$ are glued, the gluing scheme yields a valid critical pair, $A$ is an output of $L_1$ and $B$ is an input of $L_2$.
Suppose $X$ an input of $L_1$ is glued to $Y$ an output of $L_2$. There are paths $X \rightsquigarrow A$ and $B \rightsquigarrow Y$ by strong connectedness of $L_1$ and $L_2$. Then, we obtain a cycle $[X] \rightsquigarrow [A] = [B] \rightsquigarrow [Y] = [X]$ which contradicts the acyclicity property.
The second point follows a similar argument. □

Finally, we can obtain a sufficient and necessary condition for a gluing scheme to induce a critical pair.

**Proposition 3.5.** *A pre-critical pair $L_1 + L_2 \twoheadrightarrow \mathtt{coeq}_G(g_1, g_2) \xleftarrow{[\subseteq,\subseteq]} in(\mathtt{coeq}_G(g_1, g_2)) + out(\mathtt{coeq}_G(g_1, g_2))$ yielded by a gluing scheme is parallel iff the following holds:*

1. *no hyperedges from $L_1$ and $L_2$ are glued, and*
2. *if two nodes from $L_1$ and $L_2$ are glued, they are in interfaces of $L_1$ and $L_2$.*

**Proposition 3.6.** *A pre-critical pair $L_1 + L_2 \twoheadrightarrow \mathtt{coeq}_G(g_1, g_2) \xleftarrow{[\subseteq,\subseteq]} in(\mathtt{coeq}_G(g_1, g_2)) + out(\mathtt{coeq}_G(g_1, g_2))$ yielded by a gluing scheme is a critical pair iff there are hyperedges separately from $L_1$ and $L_2$ that are glued.*

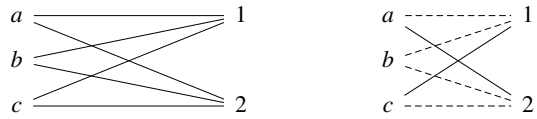
*Proof.* This is a consequence of Prop. 3.5 and $\mathtt{coeq}_G(g_1, g_2)$ being an ma-hypergraph. □

These observations suggest the following two-fold gluing process to yield a suitable gluing scheme $(g_1 \colon G \to L_1 + L_2, g_2 \colon G \to L_1 + L_2)$ that induces a critical pair: (1) glue (at least one pair of) hyperedges that are separately from $L_1$ and $L_2$, and (2) glue inputs/outputs that are separately from $L_1$ and $L_2$.
To compute such a gluing scheme, we use **independent edge sets** on complete bipartite graphs.

**Definition 3.7.** Given two sets $A$ and $B$, the **complete bipartite graph** $K_{A,B}$ is defined as follows: its vertices are $A + B$ and there is an edge between every element of $A$ and every element of $B$.
A **independent edge set**[4] on $K_{A,B}$ is a set of edges such that no two edges share common vertices.

[4] Independent edge sets are also called "matching" in graph theory.

The following shows an example of an independent edge set.




$K_{\{a,b,c\},\{1,2\}}$

For each independent edge set on hyperedges, i.e. on $K_{HEdges(L_1),HEdges(L_2)}$ (or nodes, i.e. on $K_{Nodes(L_1),Nodes(L_2)}$), we can construct the **induced hypergraph** $\gamma$ as follows: for each edge connecting two vertices in the independent edge set, we add the pair of hyperedges (or pair of nodes) associated with those vertices together with their induced pairs of sources and targets. We let $p_1^\gamma : \gamma \to L_1$ be the first projection of $\gamma$ into $L_1$ and $p_2^\gamma : \gamma \to L_2$ be the second projection of $\gamma$ into $L_2$. Consequently, we obtain a gluing scheme $(p_1^\gamma; \iota_1 : \gamma \to L_1 + L_2, p_2^\gamma; \iota_2 : \gamma \to L_1 + L_2)$.

By suitably generating independent edge sets firstly on hyperedges (i.e. $K_{HEdges(L_1),HEdges(L_2)}$), and secondly on nodes (i.e. $K_{Nodes(L_1),Nodes(L_2)}$), in particular its restriction on inputs and outputs, we can compute gluing schemes that glue hyperedges and nodes as specified by the independent edge sets and hence induces a critical pair; if there is an edge between two vertices of a bipartite graph, the two endpoints of the edge gets merged (glued).

Our algorithm uses subroutines to enumerate independent edge sets on a complete bipartite graph $K_{a,b}$. We assume that the sets $a$ and $b$ are totally ordered. We believe this is a reasonable assumption, because the set of hyperedges and the set of nodes, of a hypergraph, are typically implemented using a totally ordered data structure. The zip function turns two lists of the same length into a list of pairs.

**Algorithm 1:** An algorithm for enumerating independent edge sets

**Input:** a set $a$ and a set $b$
**Output:** independent edge sets on $K_{a,b}$
**return** $\bigcup_{k \in [\![0,\min(|a|,|b|)]\!]}$ enumerateKIndependentEdgeSets$(a,b,k)$ ;

**Algorithm 2:** An algorithm for enumerating $k$ independent edge sets

**Input :** a set $a$, a set $b$, and a number $k$
**Output:** independent edge sets on $K_{a,b}$ with $k$ edges
**for** $x \subseteq a$ **such that** $|x| = k$ **do**
  **for** $y$ **being a partial permutation of** $y' \subseteq b$ **such that** $|y| = k$ **do**
    **yield** zip(x,y)
  **end**
**end**

**Proposition 3.8.** *There are* $\sum_{0 \le k \le min(|a|,|b|)} k! \binom{|a|}{k} \binom{|b|}{k}$ *independent edge sets on* $K_{a,b}$.

*Proof.* There are $\binom{|a|}{k}$ choices of $k$ elements from a set containing $|a|$ elements and $k!\binom{|b|}{k}$ partial permutations of $k$ elements from a set containing $|b|$ elements. □

We can now present our critical pair enumeration algorithm (Algo. 3). It uses a subroutine *InducedHypergraphs* that computes induced hypergraphs of a given set of independent edge sets on hyperedges or nodes.

**Algorithm 3:** An algorithm for enumerating all critical pairs

**Input :** rewrite rules $\rho = \{L_i \xleftarrow{f_i} K_i \xrightarrow{g_i} R_i\}_{i\in I}$

**Output:** epimorphisms with interface $\{\{L_i + L_j \overset{\varepsilon}{\twoheadrightarrow} S_{ij\gamma} \leftarrow I + O\}_{\gamma\in I_{ij}}\}_{(i,j)\in I^2}$

1 **for** $(i,j) \in I^2$ **do**
2  **for** $\gamma \in InducedHypergraphs(\bigcup_{l\in\Sigma} enumerateIndependentEdgeSets($
3  $\{e \mid e \in Hyperedges(L_i), label(e) = l\}, \{e \mid e \in Hyperedges(L_j), label(e) = l\}))$ **do**
4  **if** $\gamma$ **has at least a hyperedge then**
5  $(S_{ij\gamma}, \varepsilon_{ij\gamma}) = \mathtt{coeq}_\gamma(p_1^\gamma; \iota_1, p_2^\gamma; \iota_2)$;
  /* the coequalizer ($\gamma \overset{p_1^\gamma;\iota_1}{\underset{p_2^\gamma;\iota_2}{\rightrightarrows}} L_i + L_j \overset{\varepsilon_{ij\gamma}}{\cdots\!\!\rightarrow} S_{ij\gamma}$ in $\mathbf{Hyp}_\Sigma$) */
6  $I_1 = in(S_{ij\gamma}) \cap \varepsilon_{ij\gamma}(\iota_1(in(L_i)))$;
7  $I_2 = in(S_{ij\gamma}) \cap \varepsilon_{ij\gamma}(\iota_2(in(L_j)))$;
8  $O_1 = out(S_{ij\gamma}) \cap \varepsilon_{ij\gamma}(\iota_1(out(L_i)))$;
9  $O_2 = out(S_{ij\gamma}) \cap \varepsilon_{ij\gamma}(\iota_2(out(L_j)))$;
10  **for** $\gamma' \in InducedHypergraphs(enumerateIndependentEdgeSets(I_1, O_2) +$
11  $enumerateIndependentEdgeSets(I_2, O_1))$ **do**
12  $(S_{ij\gamma'}, \varepsilon_{ij\gamma'}) = \mathtt{coeq}_{\gamma'}(p_1^{\gamma'}, p_2^{\gamma'})$;
  /* the coequalizer ($\gamma' \overset{p_1^{\gamma'}}{\underset{p_2^{\gamma'}}{\rightrightarrows}} S_{ij\gamma} \overset{\varepsilon_{ij\gamma}}{\cdots\!\!\rightarrow} S_{ij\gamma'}$) */
13  $I' = in(S_{ij\gamma'})$;
14  $O' = out(S_{ij\gamma'})$;
15  **if** $I' \overset{\subseteq}{\rightarrow} S_{ij\gamma'} \overset{\subseteq}{\leftarrow} O'$ **is a ma-cospan then**
16  **yield** $L_i + L_j \overset{\varepsilon_{ij\gamma};\varepsilon_{ij\gamma'}}{\twoheadrightarrow} S_{ij\gamma'} \overset{[\subseteq,\subseteq]}{\longleftarrow} I' + O'$;
17  **end**
18  **end**
19  **end**
20  **end**
21 **end**

**Theorem 3.9.** *Algorithm 3 is correct and exhaustive. That is,*

**Correctness** *Each result* $L_i + L_j \overset{\varepsilon_{ij\gamma'}}{\twoheadrightarrow} S_{ij\gamma'} \overset{[\subseteq,\subseteq]}{\longleftarrow} I' + O'$ *of Algorithm 3 is a critical pair.*

**Exhaustiveness** *Any critical pair of the form* $L_i + L_j \overset{\varepsilon_{ij\gamma'}}{\twoheadrightarrow} S_{ij\gamma'} \overset{[\subseteq,\subseteq]}{\longleftarrow} I' + O'$ *can be yielded by Algorithm 3.*

*Proof. Proof of correctness.* A coequalizer is an epimorphism, thus $\varepsilon_{ij\gamma}$ and $\varepsilon_{ij\gamma'}$ is epi. Their composition is therefore epi. $I' \overset{\subseteq}{\rightarrow} S_{ij\gamma'} \overset{\subseteq}{\leftarrow} O'$ is a ma-cospan as required by the pre-critical pair with interface definition because of the if statement in line 15. Moreover, the matchings $\iota_1; \varepsilon_{ij\gamma}; \varepsilon_{ij\gamma'}$ and $\iota_2; \varepsilon_{ij\gamma}; \varepsilon_{ij\gamma'}$ are mono, because the gluing schemes $\gamma$ and $\gamma'$ do not glue nodes and hyperedges of the same hypergraph. The gluing scheme $\gamma$ also glues at least a pair of hyperedges (line 4). Therefore, by Prop. 3.6, each result yielded is a critical pair.

*Proof of exhaustiveness.* Let $L_i + L_j \overset{\varepsilon}{\twoheadrightarrow} X \overset{F}{\leftarrow} J$ be a precritical-pair with interface.

An epimorphism of hypergraphs is surjective on nodes and on hyperedges because the category of hypergraphs is a presheaf category.
Thus, each node and each hyperedge of $X$ has a non-empty preimage set by $\varepsilon$. Moreover, each preimage set by $\varepsilon$ contains: (i) at most two elements, and (ii) is there are two elements, they come separately from $L_1$ and $L_2$. This is because $\iota_i;\varepsilon$ and $\iota_j;\varepsilon$ are mono.
We construct a hypergraph $\gamma$ whose nodes are given by preimage sets $\varepsilon^{-1}(v)$ with size 2 for $v \in Nodes(X)$, and hyperedges are given by preimage sets $\varepsilon^{-1}(e)$ with size 2 for $e \in HEdges(X)$. It comes with two hypergraph homomorphisms $F_1 : \gamma \to L_1 + L_2$ and $F_2 : \gamma \to L_1 + L_2$, such that $F_1$ maps a preimage set to its element from $L_1$ and $F_2$ maps a preimage set to its element from $L_2$. We obtain a gluing scheme $(F_1, F_2)$.
Because this gluing scheme induces a critical pair, it satisfies the necessary conditions of the propositions in Sec. 3. Namely:

1. If it glues edges, they are separetely from $L_1$ and $L_2$.
2. It glues at least a pair of hyperedges separately from $L_1$ and $L_2$.
3. If it glues nodes, they are either a source/target of glued hyperedges, or input/output separately from $L_1$ and $L_2$.

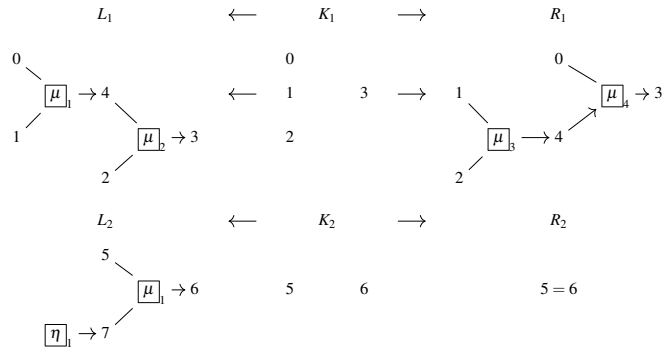
These conditions are realised by Algo. 3, respectively by lines 2 & 10, line 4, and line 10.
We can therefore conclude that the merging of hyperedges and nodes specified by the gluing scheme $(F_1, F_2)$ is impelemented by Algo. 3.

□

We implement Algo. 3 in Haskell[5], and test it using the example of non-commutative bimonoids [4, Sec. 6.1]. While there are 22 critical pairs, the implementation outputs 58 critical pairs. This is due to duplication caused by isomorphic gluing schemes $\gamma, \gamma'$. Our implementation currently does not check for isomorphisms of hypergraphs.

### 3.1 An Example Run

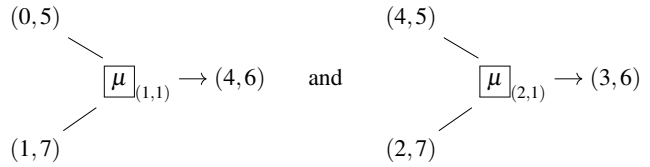
We compute critical pairs associated to the following pair of rules that is taken from the example of non-commutative bimonoids [4, Sec. 6.1].



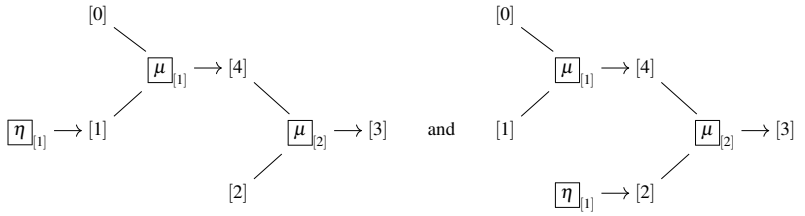

We first enumerate the independent edge sets associated to the labels $\mu$: there are three independent edge sets, namely $\{\}, \{(\boxed{\mu}_1, \boxed{\mu}_1)\}$ and $\{(\boxed{\mu}_2, \boxed{\mu}_1)\}$. There is only one independent edge set for the label

[5] Available online at `https://github.com/GuiSab/hypergraphrewriting`

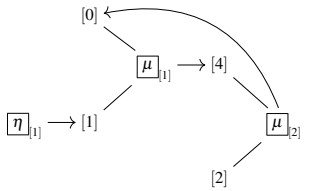
$\eta$, namely the empty one. We thus have 2 gluing schemes for the hyperedges which are not empty:



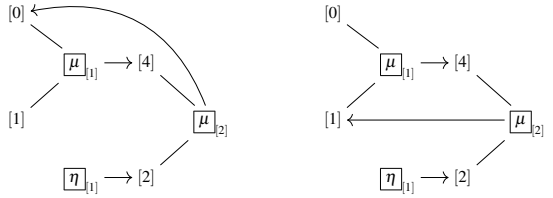

The gluings associated to the gluing schemes are the following:



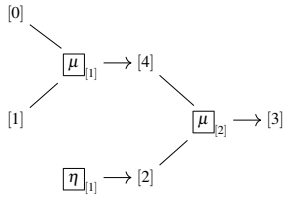

For the first gluing, we compute $I_1 = \{[0],[2]\}$, $I_2 = \{[0]\}$, $O_1 = \{[3]\}$, $O_2 = \{\}$. The only independent edge sets on the nodes are $\{\}$ and $\{([0],[3])\}$. The gluing associated to $\{([0],[3]\}$ is not acyclic:




Therefore, we only yield the critical pair given by the first gluing.
For the second gluing, we compute $I_1 = \{[0],[1]\}$, $I_2 = \{\}$, $O_1 = \{[3]\}$, $O_2 = \{3\}$. The independent edge sets on the nodes are $\{\}$, $\{([0],[3])\}$ and $\{([1],[3])\}$. The gluing associated to $\{([0],[3]\}$ and $\{([1],[3])\}$ are not acyclic:




Therefore, we only yield the critical pair given by the second gluing:



# 4 An Optimisation

Algo. 3 implements the two-fold gluing process, firstly gluing hyperedges and secondly gluing inputs/outputs. We can in fact prove that the second step is redundant, for the purpose of critical pair analysis (and local-confluence check).
Let $S_{ij\gamma}$ be a gluing of hyperedges of algorithm 3 and $S_{ij'_\gamma}$ a gluing of nodes on $S_{ij\gamma}$.

**Proposition 4.1.** *If $S_{ij'_\gamma}$ yields a valid critical pair, then $S_{ij_\gamma}$ yields a valid critical pair as well.*

*Proof.* If $S_{ij'_\gamma}$ is monogamous acyclic, then the hypergraph $S_{ij_\gamma}$ in which no nodes were glued cannot be cyclic; moreover, it will respect the monogamy condition. □

We now suppose that $S_{ij'_\gamma}$ yields a valid critical pair (and thus $S_{ij_\gamma}$ yields a valid critical pair as well).

**Proposition 4.2.** *Any convex match in $I \to S_{ij\gamma} \leftarrow O$ is a convex match in $I' \to S_{ij'_\gamma} \leftarrow O'$.*

**Corollary 4.3.** *Any rewriting sequence on $I \to S_{ij\gamma} \leftarrow O$ is also a rewriting sequence on $I' \to S_{ij'_\gamma} \leftarrow O'$.*

**Corollary 4.4.** *If $I' \to S_{ij'_\gamma} \leftarrow O'$ yields a critical pair that is not joinable, then so does $I \to S_{ij\gamma} \leftarrow O$.*

By Corollary 4.4, it suffices to enumerate the critical pairs where only hyperedges are glued, to determine if a left-connected rewrite system is locally confluent or not. Algo. 4 enumerates a sufficient subset of critical pairs necessary to determine local confluence.

**Algorithm 4:** An algorithm for enumerating sufficient critical pairs

**Input :** rewrite rules $\rho = \{L_i \xleftarrow{f_i} K_i \xrightarrow{g_i} R_i\}_{i\in I}$
**Output:** epimorphisms with interface $\{\{L_i + L_j \xrightarrow{\varepsilon} S_{ij\gamma} \leftarrow I + O\}_{\gamma\in I_{ij}}\}_{(i,j)\in I^2}$

**for** $(i,j) \in I^2$ **do**
  **for** $\gamma \in InducedHypergraphs(\prod_{l\in\Sigma} enumerateIndependentEdgeSets($
  $\{e \mid e \in Hyperedges(L_i), label(e) = l\}, \{e \mid e \in Hyperedges(L_j), label(e) = l\})$ **do**
    **if** $\gamma$ **has at least a hyperedge then**
      $(S_{ij\gamma}, \varepsilon_{ij\gamma}) = \mathtt{coeq}_\gamma(p_1^\gamma; q_1, p_2^\gamma; q_2)$;
      /* the coequaliser of $(\gamma \rightrightarrows L_1 + L_2 \dashrightarrow S_{ij\gamma}$ in $\mathbf{Hyp}_\Sigma)$ */
      $I = in(S_{ij\gamma})$;
      $O = out(S_{ij\gamma})$;
      **if** $I \xrightarrow{\subseteq} S_{ij\gamma} \xleftarrow{\subseteq} O$ **is a ma-cospan then**
        **yield** $L_i + L_j \xrightarrow{\varepsilon_{ij\gamma}} S_{ij\gamma} \xleftarrow{[\subseteq,\subseteq]} I + O$;
      **end**
    **end**
  **end**
**end**

## References


[1] Baader, F., Nipkow, T.: Term rewriting and all that. Cambridge University Press (1998)

[2] Bonchi, F., Gadducci, F., Kissinger, A., Sobocinski, P., Zanasi, F.: String diagram rewrite theory I: rewriting with frobenius structure. J. ACM **69**(2), 14:1–14:58 (2022). https://doi.org/10.1145/3502719, https://doi.org/10.1145/3502719

[3] Bonchi, F., Gadducci, F., Kissinger, A., Sobocinski, P., Zanasi, F.: String diagram rewrite theory II: rewriting with symmetric monoidal structure. Math. Struct. Comput. Sci. **32**(4), 511–541 (2022). https://doi.org/10.1017/S0960129522000317, https://doi.org/10.1017/S0960129522000317

[4] Bonchi, F., Gadducci, F., Kissinger, A., Sobocinski, P., Zanasi, F.: String diagram rewrite theory III: confluence with and without frobenius. Math. Struct. Comput. Sci. **32**(7), 829–869 (2022). https://doi.org/10.1017/S0960129522000123, https://doi.org/10.1017/S0960129522000123

[5] Born, K., Taentzer, G.: An algorithm for the critical pair analysis of amalgamated graph transformations. In: Echahed, R., Minas, M. (eds.) Graph Transformation. pp. 118–134. Springer International Publishing, Cham (2016)

[6] Ehrig, H., Pfender, M., Schneider, H.J.: Graph-grammars: An algebraic approach. In: 14th Annual Symposium on Switching and Automata Theory (swat 1973). pp. 167–180 (1973). https://doi.org/10.1109/SWAT.1973.11

[7] Ehrig, H.: Fundamentals of algebraic graph transformation. Springer, Berlin (2006), oCLC: 79651676

[8] Fiore, M.P., Campos, M.D.: The algebra of directed acyclic graphs. In: Coecke, B., Ong, L., Panangaden, P. (eds.) Computation, Logic, Games, and Quantum Foundations. The Many Facets of Samson Abramsky - Essays Dedicated to Samson Abramsky on the Occasion of His 60th Birthday. Lecture Notes in Computer Science, vol. 7860, pp. 37–51. Springer (2013). https://doi.org/10.1007/978-3-642-38164-5_4, https://doi.org/10.1007/978-3-642-38164-5_4

[9] Ghica, D.R.: Diagrammatic reasoning for delay-insensitive asynchronous circuits. In: Coecke, B., Ong, L., Panangaden, P. (eds.) Computation, Logic, Games, and Quantum Foundations. The Many Facets of Samson Abramsky - Essays Dedicated to Samson Abramsky on the Occasion of His 60th Birthday. Lecture Notes in Computer Science, vol. 7860, pp. 52–68. Springer (2013). https://doi.org/10.1007/978-3-642-38164-5_5, https://doi.org/10.1007/978-3-642-38164-5_5

[10] Hristakiev, I.: Confluence analysis for a graph programming language. Ph.D. thesis, University of York, UK (2018), https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.745783

[11] Joyal, A., Street, R.: The geometry of tensor calculus, I. Advances in Mathematics **88**(1), 55–112 (1991). https://doi.org/https://doi.org/10.1016/0001-8708(91)90003-P, https://www.sciencedirect.com/science/article/pii/000187089190003P

[12] Lafont, Y.: Towards an algebraic theory of boolean circuits. Journal of Pure and Applied Algebra **184**(2), 257–310 (2003). https://doi.org/https://doi.org/10.1016/S0022-4049(03)00069-0, https://www.sciencedirect.com/science/article/pii/S0022404903000690

[13] Mimram, S.: Computing critical pairs in 2-dimensional rewriting systems. In: Lynch, C. (ed.) Proceedings of the 21st International Conference on Rewriting Techniques and Applications, RTA 2010, July 11-13, 2010, Edinburgh, Scottland, UK. LIPIcs, vol. 6, pp. 227–242. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2010). https://doi.org/10.4230/LIPICS.RTA.2010.227, https://doi.org/10.4230/LIPIcs.RTA.2010.227

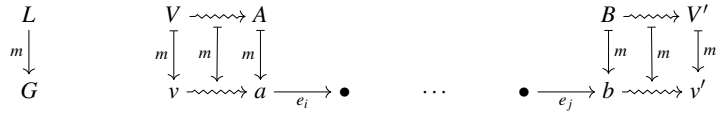
[14] Rozenberg, G.: Handbook of Graph Grammars and Computing by Graph Transformation. WORLD SCIENTIFIC (1997). https://doi.org/10.1142/3303, https://www.worldscientific.com/doi/abs/10.1142/3303

[15] Selinger, P.: A Survey of Graphical Languages for Monoidal Categories, pp. 289–355. Springer Berlin Heidelberg, Berlin, Heidelberg (2010). https://doi.org/10.1007/978-3-642-12821-9_4, https://doi.org/10.1007/978-3-642-12821-9_4

# A Proofs

## A.1 Proofs for Section 2

*Proof of Proposition 2.9.* Let's prove that in left-connected rewrite systems, the boundary complement condition is always met. All rules are left-linear in a left-connected rewrite system, the morphism $[i_L, o_L]$ is thus mono and the pushout complement defined by the left square of the diagram is unique up to isomorphism because $\mathbf{Hyp}_\Sigma$ is an adhesive category (it is a presheaf category).
Now let's prove that the convexity condition of a convex match is also always met in left-connected rewrite systems.
Suppose that a match $m : L \to G$ is not convex, meaning that there are nodes $v$ and $v'$ in $m(L)$, a path $p$ from $v$ to $v'$ and a hyperedge $e$ of $p$ not in $m(L)$.
Suppose $e$ is the last hyperedge of the path $p$. Then $v'$ which is a target of $e$ could not be in $m(L)$ because $e$ is not in $m(L)$. Therefore, $e$ cannot be the last hyperedge of the path $p$. Similarly, $e$ cannot be the first hyperedge of the path. We consider the largest subpath $[e_i, \cdots, e \cdots, e_j]$ of $p$ not in the image $m(L)$. Let $a$ be a target of the predecessor of $e_i$ in $p$ which is a source of $e_i$ and $b$ a source of the successor of $e_j$ in $p$ which is a target of $e_j$.




Let's prove that there is no hyperedge in $L$ with source $A$ the preimage of $a$ by $m$.
Suppose that there is a hyperedge $x$ with source $A$ in $L$. Then $m(x)$ is a hyperedge in $G$ with source $a$ but it cannot be $e_i$ because $e_i$ is not in $m(L)$. However, $e_i$ and $m(x)$ share a common source $a$ which is not possible because $G$ is monogamous. Therefore, $x$ cannot exist.
By the same reasoning there is no hyperedge in $L$ with target $B$ the preimage of $b$ by $m$.
Therefore, $A$ is an output node of $L$ and $B$ is an input node of $L$. By strong connectedness of $L$, there is a path from $B$ to $A$. This path has an image by $m$ and forms a cycle when composed with $[e_i, \cdots, e_j]$ which contradicts the fact that $G$ is acyclic. This refutes our first hypothesis and $m$ is convex.
Thus, a match in a left-connected rewrite system is always convex. □

*Proof of Proposition 2.10.* We recall a result on pushout complements in the category of hypergraphs which can be found in [7, pp.44,45].
Let $RR = L \xleftarrow{[i_L,o_L]} K \xrightarrow{[i_R,o_R]} R$ be a rewrite rule and $m : L \to X$ be a matching of the left-hand side of the rule.

- The gluing points $GP$ are the nodes and hyperedges in $L$ that are not deleted by the rewrite rule, namely the nodes and hyperedges in the image of $[i_L, o_L]$, *i.e.*, $GP = [i_L, o_L](V_K) \cup [i_L, o_L](E_K)$;
- the identification points $IP$ are the nodes and hyperedges in $L$ that are identified by $m$, namely the nodes and hyperedges with at least two different preimages by $m$, *i.e.*,

$$IP = \{v \in V_L \mid (\exists w \in V_L) w \neq v \wedge m(v) = m(w)\}$$
$$\cup \{e \in E_L \mid (\exists f \in E_L) f \neq e \wedge m(e) = m(f)\}$$

- the dangling points $DP$ are the nodes in $L$ whose images under $m$ are the source or target of an edge in $X$ with no preimage by $m$, *i.e.*,

$$DP = \{v \in V_L \mid (\exists e \in E_X \setminus m(E_L)), s(e) = m(v) \vee t(e) = m(v)\}$$

The pushout complement $\begin{array}{ccc} L & \xleftarrow{[i_L,o_L]} & K \\ {\scriptstyle m}\downarrow & & \downarrow \\ X & \longleftarrow & C \end{array}$ exists if and only if $IP \cup DP \subseteq GP$.

In our case, $GP$ are the interface nodes (nodes images of $[i_L, o_L]$). $IP$ is empty because a matching is mono (by the definition of a convex match). $DP$ are necessarily input and output nodes by monogamy of $X$, therefore they must be included in the interface defined by $[i_L, o_L]$.

Finally, uniqueness of the pushout complement follows from the mono condition on the match [2, Prop. 3.18]. □

## A.2 Proofs for Section 3

*Proof of Proposition 3.3.* Suppose that a node $A$ from $L_1$ and a node $B$ from $L_2$ are glued and the gluing scheme yields a valid critical pair.

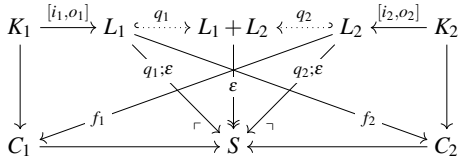
Suppose that $A$ is not a source or a target of any hyperedge in $L_1$. Then $A$ is both an input node and an output node. Therefore, $[i_{L_1}, o_{L_1}]$ is not mono and the first rule is not left-linear which contradicts our hypothesis that $\mathscr{R}$ is a left-connected rewrite system. Thus, $A$ is at least a source or target node of a hyperedge $e_1$ in $L_1$.

By a similar reasoning we get that $B$ is at least a source or target node of a hyperedge $e_2$ in $L_2$.

1. Suppose $A$ is the $i$-th source of $e_1$ and $B$ is the $j$-th source of $e_2$.
   - Suppose $e_1$ is glued to $e_2$. Suppose $i \neq j$, we then have $[e_1]$ having $[A]$ as its $i$-th source and $j$-th source which contradicts the monogamy of the candidate source. Thus $i = j$ and $A$ and $B$ are both the $i$-th source of a glued hyperedge.
   - Suppose on the contrary that $e_1$ and $e_2$ are not glued, then $[A]$ is the source of $[e_1]$ and $[e_2]$ which contradicts the monogamy of the candidate source. $\bot$
2. By a similar reasoning, if $A$ and $B$ are both targets then $A$ and $B$ are both the $i$-th target of a glued hyperedge.
3. Suppose that $A$ is the $i$-th target of $e_1$ and $B$ is the $j$-th source of $e_2$.
   - Suppose that $A$ is not an output of $L_1$. Then $A$ is also a source of a hyperedge $e_1'$ and we get the case 1. A similar argument applies to the case: $B$ is not an input of $L_2$.
   - The other two cases are: $A$ is an output of $L_1$ and $B$ is an input of $L_2$ or $A$ is an input of $L_1$ and $B$ is an output of $L_2$.

□

*Proof of Proposition 3.5 Proof of* $\Rightarrow$. Suppose we have a parallel pair

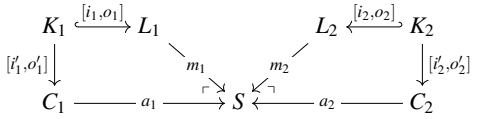



- Let $v_1 \in L_1, v_2 \in L_2$ be two glued nodes (they must come from different hypergraphs by proposition 3.2). Since this is a parallel pair, there are mappings $f_1 : L_2 \to C_1$, $f_2 : L_1 \to C_2$ such that the diagram above commutes. Because the triangle commutes, $f_1(v_2)$ must be sent to $[v_2] = [v_1] \in S$. Moreover, $(q_1;\varepsilon)(v_1) = [v_1] = [v_2] \in S$. But $S$ is a pushout, so the identified elements of $C_1$ and $L_1$

must be present in $K_1$, so $v_1$ and $f_1(v_2)$ have a preimage in $K_1$, and $v_1 \in [i_1,o_1](K_1)$. By symmetry of the argument we have $v_2 \in [i_2,o_2](K_2)$.

- Suppose there are glued hyperedges. Let $e_1 \in L_1$ and $e_2 \in L_2$ be mapped to the same $e \in S$. As before, for the identified nodes we can conclude that they are in the interfaces $[i_1,o_1],[i_2,o_2]$ by diagram chasing. But $K_1,K_2$ are discrete, so such hyperedges cannot exist - there are no elements $e'_1 \in C_2$, $e'_2 \in C_1$ such that $f_2^{-1}(e'_1) = e_1$ and $f_1^{-1}(e'_2) = e_2$ and the diagram commutes.

*Proof of* $\Leftarrow$. Suppose the assumptions 1 and 2 hold, let's prove that the pre-critical pair is parallel. Consider




The proof is by diagram chasing.

If no nodes are glued, then no hyperedges are glued and $S \cong L_1 + L_2$ which obviously is a parallel pair.

Otherwise, consider a pair of nodes ($v_1 \in L_1$, $v_2 \in L_2$) with the same image $v = m_1(v_1) = m_2(v_2)$ in $S$. By assumption, $v_2$ is in the image of $[i_2,o_2] : K_2 \to L_2$. Let $f_2 : L_1 \to C_2$ be the morphism sending $v_1$ to $[i'_2,o'_2](v_2)$ and acting as an identity on the rest of the nodes and edges (since by assumption no edges from $L_1$ and $L_2$ glued).

By construction this will make the triangle commute.

Similarly for $f_1 : L_2 \to C_1$. □

### A.3 Proofs for Section 4

*Proof of Proposition 4.2.* Let $m : L \to S_{ij\gamma}$ be a convex match. We will prove that $m;\varepsilon_{ij\gamma} : L \to S_{ij\gamma}$ is a convex match as well.

Let $e_1$ and $e_2$ be two hyperedges of $L$ such that $(m;\varepsilon_{ij\gamma})(e_1) = (m;\varepsilon_{ij\gamma})(e_2)$, meaning that $\varepsilon_{ij\gamma}(m(e_1)) = \varepsilon_{ij\gamma}(m(e_2))$. We must have $m(e_1) = m(e_2)$ because $\varepsilon_{ij\gamma}$ only glues nodes, it is thus mono on hyperedges. We then deduce $e_1 = e_2$ because $m$ is mono. $m;\varepsilon_{ij\gamma}$ is therefore mono on hyperedges.

Let $v_1$ and $v_2$ be two nodes of $L$ such that $(m;\varepsilon_{ij\gamma})(v_1) = (m;\varepsilon_{ij\gamma})(v_2)$, meaning that $\varepsilon_{ij\gamma}(m(v_1)) = \varepsilon_{ij\gamma}(m(v_2))$.

Suppose $m(v_1) \neq m(v_2)$. $\varepsilon_{ij\gamma}$ would glue $m(v_1)$ with $m(v_2)$. By construction of $\varepsilon_{ij\gamma}$ we either have $m(v_1)$ an input of $S_{ij\gamma}$ and $m(v_2)$ an output of $S_{ij\gamma}$ or $m(v_1)$ an output of $S_{ij\gamma}$ and $m(v_2)$ an input of $S_{ij\gamma}$. Let's suppose WLOG the first case. As $m$ is a convex match, there is a path $[m(e_1),\cdots,m(e_n)]$ from $m(v_1)$ to $m(v_2)$ in the image of $m$. By applying $\varepsilon_{ij\gamma}$ to the path, we get $[\varepsilon_{ij\gamma}(m(e_1)),\cdots,\varepsilon_{ij\gamma}(m(e_2))]$ a path from $\varepsilon_{ij\gamma}(m(v_1))$ to $\varepsilon_{ij\gamma}(m(v_2)) = \varepsilon_{ij\gamma}(m(v_1))$. $S_{ij\gamma}$ would not be acyclic which contradicts the hypothesis that $S_{ij'_\gamma}$ yields a valid critical pair. Therefore we proved $m(v_1) = m(v_2)$. We then deduce $v_1 = v_2$ because $m$ is mono. $m;\varepsilon_{ij\gamma}$ is therefore mono on nodes.

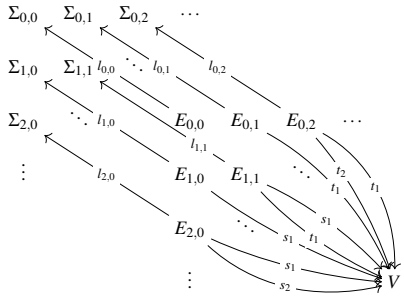
$m;\varepsilon_{ij\gamma}$ is mono, we can therefore deduce that $m;\varepsilon_{ij\gamma} : L \to S_{ij\gamma}$ is a convex match. □

## B Categorical constructions in the category of hypergraphs

The category of Σ-hypergraphs considered in this paper is a functor category $[\mathbb{E}_{\Sigma-hypergraph}, \mathbf{FinSet}]$ where $\mathbb{E}_{\Sigma-hypergraph}$ is the category generated by a quiver having:

- an object $V$ corresponding to the nodes of the hyperedge;

- for each couple of natural numbers $(i,j)$ an object $E_{i,j}$ corresponding to the hyperedges of arity $i$ and coarity $j$;
- for each couple of natural numbers $(i,j)$ an object $\Sigma_{i,j}$ corresponding to the signature of hyperedges of arity $i$ and coarity $j$;
- for each couple of natural numbers $(i,j)$ and for each natural number $k$ in $[\![1,i]\!]$ an arrow $s_k : E_{i,j} \to V$ corresponding to the $k^{\text{th}}$ source function on hyperedges of arity $i$ and coarity $j$;
- for each couple of natural numbers $(i,j)$ and for each natural number $k$ in $[\![1,j]\!]$ an arrow $t_k : E_{i,j} \to V$ corresponding to the $k^{\text{th}}$ target function on hyperedges of arity $i$ and coarity $j$;
- for each couple of natural numbers $(i,j)$ an arrow $l_{i,j} : \Sigma_{i,j} \to V$ corresponding to the labelling function on hyperedges of arity $i$ and coarity $j$.




Therefore, limits and colimits are computed pointwise: there are isomorphisms $[\mathscr{I},[\mathbb{E}_{\Sigma-hypergraph},\mathbf{FinSet}]] \cong [\mathscr{I}\times\mathbb{E}_{\Sigma-hypergraph},\mathbf{FinSet}] \cong [\mathbb{E}_{\Sigma-hypergraph}\times\mathscr{I},\mathbf{FinSet}] \cong [\mathbb{E}_{\Sigma-hypergraph},[\mathscr{I}\to\mathbf{FinSet}]]$ by curryfication and symmetry of the product. Thus a diagram $D$ element of $[\mathscr{I},[\mathbb{E}_{\Sigma-hypergraph},\mathbf{FinSet}]]$ yields a hypergraph $H$ of diagrams in **FinSet** element of $[\mathbb{E}_{\Sigma-hypergraph},[\mathscr{I},\mathbf{FinSet}]]$, the (co)limit of $D$ is the hypergraph given by the (co)limits of the various images $H(-)$.
For example, we spell out what a product is in $[\mathbb{E}_{\Sigma-hypergraph},\mathbf{FinSet}]$. Let $H_1$ and $H_2$ be two $\Sigma-$hypergraphs, we could also say that $H_1$ and $H_2$ are functors $H_1 : \mathbb{E}_{\Sigma-hypergraph} \to \mathbf{FinSet}$ and $H_2 : \mathbb{E}_{\Sigma-hypergraph} \to \mathbf{FinSet}$. The product $H_1 \times H_2 : \mathbb{E}_{\Sigma-hypergraph} \to \mathbf{FinSet}$ is then defined as follows:

- $(H_1\times H_2)(V) \stackrel{\text{def}}{=} H_1(V)\times H_2(V)$;
- $(\forall (i,j)\in\mathbb{N}) \qquad (H_1\times H_2)(E_{i,j}) \stackrel{\text{def}}{=} H_1(E_{i,j})\times H_2(E_{i,j})$;
- $(\forall (i,j)\in\mathbb{N}) \qquad (H_1\times H_2)(\Sigma_{i,j}) \stackrel{\text{def}}{=} H_1(\Sigma_{i,j})\times H_2(\Sigma_{i,j})$;
- $(\forall (i,j)\in\mathbb{N})(\forall k\in[\![1,i]\!]) \qquad (H_1\times H_2)(s_k)((e_1,e_2)) \stackrel{\text{def}}{=} (H_1(s_k)(e_1),H_2(s_k)(e_2))$;
- $(\forall (i,j)\in\mathbb{N})(\forall k\in[\![1,j]\!]) \qquad (H_1\times H_2)(t_k)((e_1,e_2)) \stackrel{\text{def}}{=} (H_1(t_k)(e_1),H_2(t_k)(e_2))$;
- $(\forall (i,j)\in\mathbb{N}) \qquad (H_1\times H_2)(l_{i,j})((e_1,e_2)) \stackrel{\text{def}}{=} (H_1(l_{i,j})(e_1),H_2(l_{i,j})(e_2))$.

The projections are the obvious ones. The coproduct is defined dually.
We spell out what a coequalizer is in $[\mathbb{E}_{\Sigma-hypergraph},\mathbf{FinSet}]$. Let $f : H_1 \to H_2$ and $g : H_1 \to H_2$ be two $\Sigma-$hypergraph morphisms, we could also say that $f$ and $g$ are natural transformations $f : H_1 \Rightarrow H_2$ and $g : H_1 \Rightarrow H_2$. The coequalizer $\mathtt{coeq}_{H_1}(f,g)$ is then defined as follows:

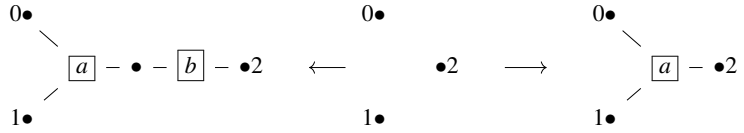
- $\mathtt{coeq}_{H_1}(f,g)(V) \stackrel{\text{def}}{=} \mathtt{coeq}_{H_1}(f_V,g_V) = H_2(V)/\mathscr{R}(V)$ where $\mathscr{R}(V)$ is the smallest equivalence relation generated by the relation $\{f_V(x)\sim g_V(x) \quad | \quad x\in H_1(V)\}$;
- $(\forall (i,j)\in\mathbb{N}) \qquad \mathtt{coeq}_{H_1}(f,g)(E_{i,j}) \stackrel{\text{def}}{=} \mathtt{coeq}_{H_1}(f_{E_{i,j}},g_{E_{i,j}}) = H_2(E_{i,j})/\mathscr{R}(E_{i,j})$ where $\mathscr{R}(E_{i,j})$ is the smallest equivalence relation generated by the relation $\{f_{E_{i,j}}(x)\sim g_{E_{i,j}}(x) \quad | \quad x\in H_1(E_{i,j})\}$;

- $(\forall (i,j) \in \mathbb{N}) \qquad \mathtt{coeq}_{H_1}(f,g)(\Sigma_{i,j}) \stackrel{\text{def}}{=} \mathtt{coeq}_{H_1}(f_{\Sigma_{i,j}}, g_{\Sigma_{i,j}}) = H_2(\Sigma_{i,j})/\mathscr{R}(\Sigma_{i,j})$ where $\mathscr{R}(\Sigma_{i,j})$ is the smallest equivalence relation generated by the relation $\{f_{\Sigma_{i,j}}(x) \sim g_{\Sigma_{i,j}}(x) \quad | \quad x \in H_1(\Sigma_{i,j})\}$;
- $(\forall (i,j) \in \mathbb{N})(\forall k \in [\![1,i]\!]) \qquad \mathtt{coeq}_{H_1}(f,g)(s_k)([e]_{H_2(E_{i,j})/\mathscr{R}(E_{i,j})}) \stackrel{\text{def}}{=} [s_k(e)]_{H_2(V)/\mathscr{R}(V)}$ where $[x]_{E/\mathscr{R}}$ is the notation for the equivalence class associated to an element $x$ of $E$ in the quotient set $E/\mathscr{R}$;
- $(\forall (i,j) \in \mathbb{N})(\forall k \in [\![1,j]\!]) \qquad \mathtt{coeq}_{H_1}(f,g)(t_k)([e]_{H_2(E_{i,j})/\mathscr{R}(E_{i,j})}) \stackrel{\text{def}}{=} [t_k(e)]_{H_2(V)/\mathscr{R}(V)}$;
- $(\forall (i,j) \in \mathbb{N}) \qquad \mathtt{coeq}_{H_1}(f,g)(l_{i,j})([e]_{H_2(E_{i,j})/\mathscr{R}(E_{i,j})}) \stackrel{\text{def}}{=} [l_{i,j}(e)]_{H_2(\Sigma_{i,j})/\mathscr{R}(\Sigma_{i,j})}$.

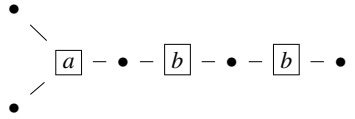
The coprojection is the obvious one (the mapping sending each element to its equivalence class). The pushout computation is deduced from the coproduct and the coequalizer.

# C Examples

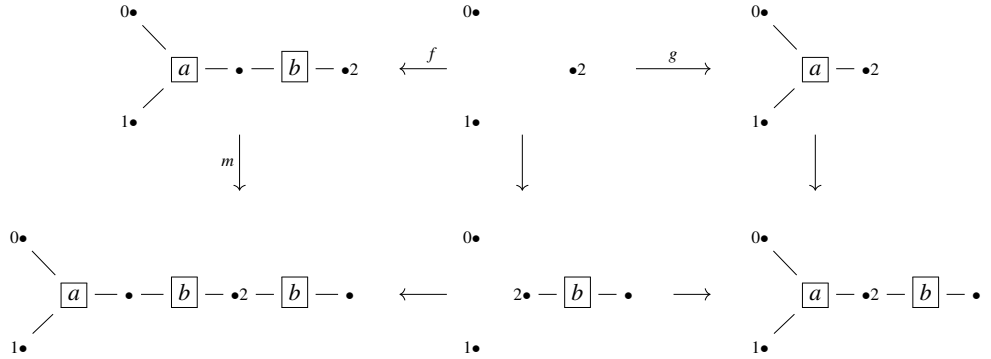
**Example C.1** (An example of DPO rewriting in $\mathbf{Hyp}_\Sigma$)**.** *Take the following rewrite rule that removes one b from a graph (we identified the mapping by numbering the dots):*




*and let G be the following hypergraph:*




*Then, by using DPO rewriting, we obtain H by the following procedure:*

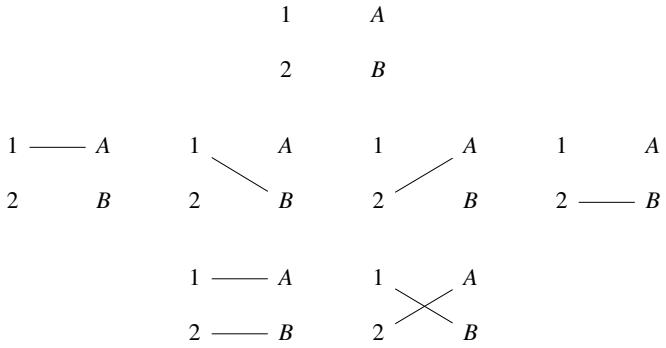



**Example C.2** (Examples of independent edge sets)**.** *Given two sets* $\{1,2\}$ *and* $\{A,B\}$*, independent edge sets on* $K_{\{1,2\},\{A,B\}}$ *are:*

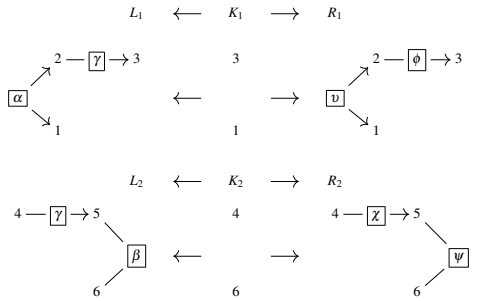
- *for size* 0*:* $\emptyset$*,*
- *for size* 1*:* $\{(1,A)\}$*,* $\{(1,B)\}$*,* $\{(2,A)\}$*,* $\{(2,B)\}$*,*

- *for size* 2*:* $\{(1,A),(2,B)\}$, $\{(1,B),(2,A)\}$.

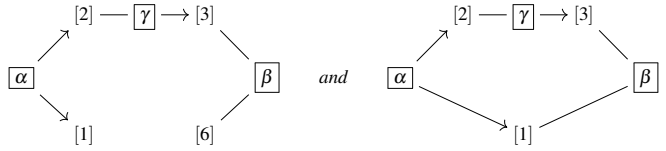
*These independent edge sets correspond to the following bipartite graphs:*




**Example C.3** (Example yielding a critical pair in which two nodes are glued)**.** *Let's give an example of rewrite rules where the gluing of nodes matter: we will study the critical pairs associated to the rules below.*




*The two critical pairs yielded by the algorithm are given by the following gluings:*



## Curriculum vitae

**Diana-Maria Kessler**

| | |
|---|---|
| ORCID | 0009-0006-6355-1581 |
| Homepage | https://sites.google.com/view/dianakessler/home |


**Institutions and positions**

| | |
|---|---|
| 19.11.2020–... | Tallinn University of Technology, School of Information Technologies, Department of Software Science, Junior Research Fellow (1,00) |
| 01.04.2019–12.07.2019 | Adobe Systems, Software Engineer Intern |
| 01.06.2017–18.08.2017 | Metaswitch, Software Engineering Intern |
| 01.06.2016–30.08.2016 | University of Edinburgh, Summer Intern |

**Academic degrees**

Diana-Maria Kessler, Master's Degree, 2020, (sup) Chris Heunen, Master of Informatics, University of Edinbuirgh

Diana - Maria Kessler, PhD student, (sup) Paweł Maria Sobociński; Amar Hadzihasanovic, Models of computation, compositionally, Tallinn University of Technology School of Information Technologies, Department of Software Science

**Honours & awards**

2022 Distinguished presentation, Applied Category Theory (ACT)

**Additional information**

During my PhD (September 2023): Academic visit to Jamie Vicary at the University of Cambridge.

During my time at Edinburgh University (2019-2020), I tutored for the first year courses ``Computation and Logic'' and ``Introduction to Object-Oriented Programming'' and the third year course ``Software testing''.

## Education

2014–2020 MSc in Informatics, University of Edinburgh

## Teaching

### Functional Programming (ITI0212)

I was a teaching assistant (TA) in 2022 and an instructor/TA in 2023.

## Publications

**2024**

Hadzihasanovic, Amar; Kessler, Diana (2024). Acyclicity conditions on pasting diagrams. Applied Categorical Structures, 32 (6), 1−34. DOI: 10.1007/s10485-024-09784-x.

**2023**

Hadziahasanovic, Amar; Kessler, Diana (2023). Data structures for topologically sound higher-dimensional diagram rewriting. Proceedings Fifth International Conference on Applied Category Theory (ACT 2022), Glasgow, United Kingdom, 18-22 July 2022. Ed. Master, Jade; Lewis, Martha. Sydney: Open Publishing Association, 111−127. (Electronic Proceedings in Theoretical Computer Science; 380). DOI: 10.4204/EPTCS.380.7.

Hadzihasanovic, Amar; Kessler, Diana (2023). Higher-dimensional subdiagram matching. 2023 38th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), 26 - 29 June, 2023, Boston, USA. ACM, 13 pp. DOI: 10.1109/LICS56636.2023.10175726.

# Elulookirjeldus

**Diana-Maria Kessler**

**Töökohad ja ametid**

| | |
|---|---|
| 19.11.2020–... | Tallinna Tehnikaülikool, Infotehnoloogia teaduskond, Tarkvarateaduse instituut, doktorant-nooremteadur (1,00) |
| 01.04.2019–12.07.2019 | Adobe Systems, Tarkvaraarenduse praktikant |
| 01.06.2017–18.08.2017 | Metaswitch, Tarkvaraarenduse praktikant |
| 01.06.2016–30.08.2016 | Edinburghi Ülikool, Suvepraktikant |

**Teaduskraadid**

Diana - Maria Kessler, doktorant, (juh) Paweł Maria Sobociński; Amar Hadzihasanovic, Models of computation, compositionally (Arvutamise mudelid, kompositsiooniliselt), Tallinna Tehnikaülikool, Infotehnoloogia teaduskond, Tarkvarateaduse instituut

**Teaduspreemiad ja tunnustused**

2022 Silmapaistev ettekanne, Applied Category Theory (ACT)

**Lisainfo**

Doktorantuuri ajal (september 2023): akadeemiline visiit Cambridge'i Ülikooli Jamie Vicary juurde.

Edinburghi Ülikoolis (2019–2020) juhendasin esimese aasta kursusi "Arvutus ja loogika" ning "Sissejuhatus objektorienteeritud programmeerimisse" ja kolmanda aasta kursust "Tarkvara testimine".

**Haridustee**

| | |
|---|---|
| 2014–2020 | Infotehnoloogia magistrikraad, Edinburghi Ülikool |

## Õppetöö

### Funktsionaalne programmeerimine (ITI0212)

Olin 2022 aastal õppeassistent (TA) ja 2023 aastal juhendaja/õppeassistent (TA).

## Publikatsioonid

**2024**

Hadzihasanovic, Amar; Kessler, Diana (2024). Acyclicity conditions on pasting diagrams. Applied Categorical Structures, 32 (6), 1−34. DOI: 10.1007/s10485-024-09784-x.

**2023**

Hadziahasanovic, Amar; Kessler, Diana (2023). Data structures for topologically sound higher-dimensional diagram rewriting. Proceedings Fifth International Conference on Applied Category Theory (ACT 2022), Glasgow, United Kingdom, 18-22 July 2022. Ed. Master, Jade; Lewis, Martha. Sydney: Open Publishing Association, 111−127. (Electronic Proceedings in Theoretical Computer Science; 380). DOI: 10.4204/EPTCS.380.7.

Hadzihasanovic, Amar; Kessler, Diana (2023). Higher-dimensional subdiagram matching. 2023 38th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), 26 - 29 June, 2023, Boston, USA. ACM, 13 pp. DOI: 10.1109/LICS56636.2023.10175726.

ISSN 2585-6901 (PDF)
ISBN 978-9916-80-331-8 (PDF)