2020

Algebraic and Geometric Properties of Hierarchical Models

Aida Maraj
University of Kentucky, aida.maraj@uky.edu
Digital Object Identifier: https://doi.org/10.13023/etd.2020.232

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Aida Maraj, Student

Dr. Uwe Nagel, Major Professor

Dr. Peter Hislop, Director of Graduate Studies

Algebraic and Geometric Properties of Hierarchical Models

DISSERTATION

A dissertation submitted in partial
fulfillment of the requirements for
the degree of Doctor of Philosophy
in the College of Arts and Sciences
at the University of Kentucky

By
Aida Maraj
Lexington, Kentucky

Director: Dr. Uwe Nagel, Professor of Mathematics
Lexington, Kentucky
2020

ABSTRACT OF DISSERTATION

Algebraic and Geometric Properties of Hierarchical Models

In this dissertation filtrations of ideals arising from hierarchical models in statistics related by a group action are are studied. These filtrations lead to ideals in polynomial rings in infinitely many variables, which require innovative tools. Regular languages and finite automata are used to prove and explicitly compute the rationality of some multivariate power series that record important quantitative information about the ideals. Some work regarding Markov bases for non-reducible models is shown, together with advances in the polyhedral geometry of binary hierarchical models.

KEYWORDS: hierarchical models, toric ideals, Markov bases, stabilization, equivariant Hilbert series, polyhedral geometry

Aida Maraj

May 11, 2020

Algebraic and Geometric Properties of Hierarchical Models

By
Aida Maraj

Dr. Uwe Nagel
Director of Dissertation

Dr. Peter Hislop
Director of Graduate Studies

May 11, 2020
Date

Dedicated to my mother.
Dedikuar nënës sime.

ACKNOWLEDGMENTS

. . . dhe drita e diturisë përpara do të na shpjerë!/. . . and the light of knowledge will lead us forward!–Naim Frashëri.

While I write this part of my thesis, I realize how lucky I am to have found mentors, professors, collaborators, colleagues, and friends that have played significant positive roles in my graduate school experience.

I wish, first of all, to express my deep gratitude to my advisor Uwe Nagel. I am ending this experience with knowledge and math curiosity led or triggered by him, which doesn't end with this work! A special thanks to him for introducing me to algebraic statistics. Through my advisor, I have found myself with a supportive academic family – Rachelle Bouchat, Nathan Fieldsteel, Sema Güntürkün, Jennifer Kenkel, Patricia Klein, Sonja Petrović, and William Trok.

I would like to thank my committee members Alberto Corso, Heide Luerssen-Gluesing, Katherine Thompson, Xiangrong Yin, and the outside examiner, Michael Samers, for the time and cooperation spent during this approval process.

I am very grateful to my letter-writers, Ben Braun, Daniel Erman, Chris Manon, and Seth Sullivant, for their time and their support. They have been role models for me and inspire me how to be a good member in the mathematics community. I am greatly indebted to Seth for his patience with me, and for inviting me to visit his institution, which led to my collaboration with two of his graduate students.

I would also like to express deep gratitude to Bernd Sturmfels and to Andrew Snowden for their support on my career. I am very looking forward to talk math with them in their institutions!

It has been a joy to discover new math with my collaborators and my good friends Jane Ivy Coons, Joseph Cummings and Ben Hollering! I also thank Brian Davis and

Robert Krone for the all the productive math conversations!

While traveling to conferences and summer schools, I have found very welcoming environments in both algebraic statistics and commutative algebra to which I am very excited to continue to be part of. Here I am mentioning Ayah Almousa, Juliette Bruce, Kaie Kubjas, Kuei-Nuan Lin, Abraham Martin Del Campo, Lily Silverstein, Ola Sobjeska, among a lot.

I am grateful to the University of Kentucky Department of Mathematics for providing summer support for my travel and research. Here I have had a very welcoming and supportive environment of professors and graduate students. The department has an amazing administrative group of Rejeana Cassady, Christine Levitt, and Sheri Rhine, that positively affects the lives of each of us.

In particular, I want to thank erica Whitaker for all the teaching related help and support, and for showing care about my health. During conversations with erica, she left me with this inspiring quote *"None of us in academia are perfect, and if you wait until you feel perfect to communicate with colleagues it will be a lonely place indeed."*

A would like to extend my deepest gratitude to my friend Darleen and to her family. While sharing winter and spring breaks with the Perez-Lavin family, I have received love, support, peace, and a second family. Thank you mama Elizabeth for all the warm hugs and messages!

A special thanks to my friend Jared Antrobus taking care of me while I was sick. To this, I thank my doctors, Lori and Tina, for their valiant efforts on improving my health. I also thank my amazing non-math friend Titay Ayano, who was my bridge for life outside Patterson Office Tower.

Në përfundim, falenderoj me të rëndësishmit e jetës sime, prindërit e mi Agim dhe Hajdi, dhe vëllain tim Albi. Falenderoj gjithashtu tezet Zajde dhe Mira, dhe kushërirën time të vogël Eliza për dashurinë e dhënë. E di që këto pesë vite të jetuar të ndarë nga një oqean nuk kanë qënë të lehtë për asnjë nga ne. I jam mirënjohese në veçanti mamit për gjithë sakrificat dhe për mbështetetjen e zgjedhjeve të mija. Falenderoj mësuesit ndër vite që kanë ndikuar pozitivisht në formimin tim. Në veçanti, falenderoj Elton Paskun që më drejtoi në Universitetin e Kentucky-it.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1 Introduction

This dissertation presents results inspired by problems in algebraic statistics. Algebraic statistics is the research area that uses sources including algebraic geometry, commutative algebra, combinatorics, and symbolic computation to solve problems in probability theory and statistics. This relatively new area, with the very first paper in 1998 [16], provides new techniques for statistical problems, and has led to many interesting developments in mathematics. The book [42] contains a good sampling of algebraic statistics topics, and their development.

Since 1998, with the work of Diaconis and Sturmfels [16], it has become clear that algebra can be useful in analyzing statistical models. Some good examples are hierarchical log-linear models, which describe the dependency relationships among random variables in observed data. This class of statistical models can be described parametrically in terms of algebraic conditions on a natural parameter space. Hierarchical models are defined by the number of states each random variable can take and the collection of subsets of dependent random variables. Each random variable has some fixed finite number of states. Intuitively, if one increases the number of states in a model, one expects essential properties to be preserved. In order to make this precise, one takes advantage of symmetry and studies asymptotic behaviours of related models. Algebraically, increasing the number of states leads to ideals with more generators in polynomial rings in increasingly many variables. There are two approaches we can take: we can work with sequences of related ideals in larger and larger polynomial rings, or we can work with a single ideal in a polynomial ring with infinitely many variables. This is a new point of view, less than a decade old, which opens up vast new territories to explore in algebra. In representation theory and topology, similar ideas have led to many breakthroughs using the theory of FI-modules.

This dissertation studies problems related to quantitative and qualitative asymptotic behaviours of hierarchical models, such as the generating sets and the Hilbert series of hierarchical models with increasing number of states, and presents some advances in the half-space description of the underlying polytope.

1.1 Stabilization of Hierarchical Models

Chapter 2 concerns the Markov bases for the hierarchical model. Markov bases are an essential tool for performing statistical tests with hierarchical models. Diaconis and Sturmfels proved in [16] that Markov bases of a hierarchical model can be obtained by computing generating sets of a toric ideal arising from the parametric description. One difficulty is that these ideals have very large minimal generating sets for large number of states, even if the dependency relations are simple. The number of generators grows rapidly when we increase the number of states incrementally. However, one can control this growth via symmetry. The question of finding a finite set of "master generators", that via an action of a product of symmetric groups produces generating sets for any model in a family of models that share the same dependency relations among random variables is widely studied by mathematicians and statisticians [2, 10, 15, 22, 24, 36, 37]. The challenging hierarchical models for which very little is known are the non-reducible ones. Chapter 2 provides a conjecture when these non-reducible hierarchical models have the underlying dependency structure of a cycle, together with some progress toward the conjecture. We work in details a class of $four$-cycle models, where we prove that a set of Markov moves induced by some *three*-cycle models is Markov basis for it. The work of Hoşten and Sullivant in [24] is used to guess a Markov bases for the $four$-cycle model, and the approach of Aoki and Takemura in [2] is used to prove this is indeed a Markov basis. Additionally, some interpretation on the stabilization of filtrations of hierarchical models with a decomposable, reducible, and non-reducible dependency structure is given.

1.2 Equivariant Hilbert Series for Hierarchical Models

Chapter 3 concerns finding quantitative information for the hierarchical models via their toric ideals. In work of Hoşten and Sullivant [24], a formula on the dimension of these ideals is given. Detailed quantitative information is recorded in the Hilbert series of the ideal. Hilbert series are formal power series in one variable, which are proven to always have rational presentation. The rational presentation of the Hilbert series captures in a compact form many other quantitative facts about the model. Unfortunately, obtaining a rational presentation of Hilbert series is a very difficult problem, even for a simple model or ideal.

Nagel–Römer [32] take a new approach on the Hilbert series of related problems. They discovered that studying many ideals simultaneously is a powerful technique. They introduce a new formal power series in two variables that records all the information the Hilbert series has, but for infinitely many such ideals that are related by the symmetric group action on the indices. They show these power series always have a rational form. Krone, Leykin, and Snowden [28] give an algorithm to explicitly compute such power series via regular languages and finite automata. Inspired by this, in this chapter we define a multivariate equivariant Hilbert series for ideals of related hierarchical models together with a hypothesis which ensures that these are rational functions with rational coefficients, and allow us to derive quantitative information about each individual ideal. The proof uses regular languages and finite automata theory. We implemented an algorithm for explicitly computing such rational presentations in the symbolic computation software Macaulay2. The code can be found in the Appendix.

The concept of multivariate Hilbert series using regular languages and finite automata to compute them are new approaches that offer a lot of promise for extending the above results, as well as for analyzing the size of other algebraic objects. More generally, these techniques should be useful in other contexts, where one attempts a quantitative asymptotic analysis.

1.3 Polyhedral Geometry of Hierarchical Models

Chapter 4 is based on joint work with Jane Ivy Coons, Joseph Cummings, and Ben Hollering, and has a combinatorial flavour. It involves some famous polytopes: correlation polytopes (see [11, Chapter 5],[35]), cut polytopes (see [11]), and marginal polytopes (see [42, Chapter 9]). Each of them has a multitude of uses beyond combinatorics. Correlation polytopes arise in the theory of probability and propositional logic. The theory of cut polytopes is developed by their use in combinatorial optimization. The polytopes associated to toric ideals of hierarchical models are marginal polytopes.

We extend work of Sturmfels and Sullivant [41] on the use of the cut polytopes for graphical binary hierarchical models, and introduce generalized cut polytopes defined over any simplicial simplex. As hoped, the new polytope, the correlation polytope, and the binary marginal polytope over the same simplicial complex are isomorphic. Moreover, these generalized cut polytopes are full dimensional, which promises general results on the half-space description of binary marginal polytopes, which are very far from being full dimensional. We provide results on half-space descriptions of these polytopes when the dependency relations are representable by the boundary of the simplex. In the proof we use Gale transformations and switch operators.

Chapter 2 Preliminaries

Chapter 2 provides background on a number of concepts that play a prominent role in this dissertation. Readers familiar with any of the concepts therein should feel comfortable skipping the corresponding section in it. Section 2.1 reviews toric ideals over polynomial rings, their generating sets, and a Hilbert series of its quotient ring. The content of Section 2.1 is required for Sections 2.3 and 2.4, and all the chapters. As the title indicates, hierarchical models, which are introduced in Section 2.2, play an important role in this thesis, which will be used in all chapters. In this dissertation, toric ideals are used as analogous objects to hierarchical models. Section 2.3 presents how one constructs such ideals, and their close connection to hierarchical models. These ideals are used in all chapters. In studying asymptotic behaviours of hierarchical models in Chapters 3 and 4, one needs filtrations, which are introduced in Section 2.4. Section 2.5 introduced regular languages, finite automata, and their formal power series used in Chapter 4. Lastly, Gale transformations are introduced in Section 2.6, and the switch operators are used in Section 2.7. Both concepts will be used in Chapter 5.

2.1 Toric Ideals

Given a commutative ring $(R, +, \cdot)$, a nonempty subset I of R is an ideal in R if I contains the zero element of R, is closed under addition i.e. $I + I \subseteq I$, and is closed under multiplication by R, i.e. $R \cdot I = I \cdot R \subseteq I$. In this dissertation we consider certain ideals in polynomial rings over fields, called toric ideals. Throughout this work we use $\mathbb{N}$ and $\mathbb{N}_0$ to denote the set of positive integers and the set of non-negative integers, respectively.

Let $\mathbb{K}[x] = \mathbb{K}[x_1, \ldots, x_n]$ be a polynomial ring over a field $\mathbb{K}$, vectors $\mathscr{A} = \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\} \subseteq \mathbb{Z}^d$, and $\mathbb{K}[y^{\pm 1}] = \mathbb{K}[y_1^{\pm 1}, \ldots, y_d^{\pm 1}]$ a Laurent polynomial ring. The matrix $A = [\mathbf{a}_1\ \mathbf{a}_2\ \ldots\ \mathbf{a}_n]$ induces a semigroup homomorphism:

$$\varphi : \mathbb{N}^n \xrightarrow{A} \mathbb{Z}^d, \quad \mathbf{u} \to \mathbf{a}_1 u_1 + \cdots + \mathbf{a}_n u_n = A \cdot \mathbf{u}, \tag{2.1.1}$$

which lifts to a homomorphism of semigroup algebras

$$\Phi : \mathbb{K}[x] \xrightarrow{A} \mathbb{K}[y^{\pm 1}], \quad x_i \to y^{\mathbf{a}_i}. \tag{2.1.2}$$

Denote by $I_{\mathscr{A}}$ the kernel of such a homomorphism.

Lemma 2.1.1. *[40, Lemma 4.2] The toric ideal $I_{\mathscr{A}}$ is spanned as a $\mathbb{K}$-vector space by the set of binomials*

$$\{\mathbf{x}^{\mathbf{u}} - \mathbf{x}^{\mathbf{v}} \mid \mathbf{u}, \mathbf{v} \in \mathbb{N}^n \text{ with } \varphi(\mathbf{u}) = \varphi(\mathbf{v})\}$$

Every vector $\mathbf{u} \in \mathbb{Z}^n$ can be written uniquely as $\mathbf{u} = \mathbf{u}^+ - \mathbf{u}^-$, where $\mathbf{u}^+$ and $\mathbf{u}^-$ are non-negative and have disjoint support. More precisely, the i-th coordinate of $\mathbf{u}^+$ equals u_i if $u_i > 0$, and it equals 0 otherwise. Similarly, the i-th coordinate of $\mathbf{u}^+$ equals $-u_i$ if $u_i < 0$, and 0 otherwise. We write $\ker(\varphi)$ for the sublattice of $\mathbb{Z}^n$ consisting of all vectors $\mathbf{u}$ such that $\varphi(\mathbf{u}^+) = \varphi(\mathbf{u}^-)$. With this, Lemma 2.1.1 can be rewritten as follows:

$$I_{\mathscr{A}} =< \{\mathbf{x}^{\mathbf{u}^+} - \mathbf{x}^{\mathbf{u}^-} \mid \mathbf{u} = ker(\varphi)\} >$$

Lemma 2.1.2. *[40, Lemma 4.2] The Krull dimension of the residue ring $\mathbb{K}[x]/I_{\mathscr{A}}$ is equal to $\dim \mathscr{A}$. The latter is one more than the dimension of the polytope in $\mathbb{R}^d$ with the columns of $\mathscr{A}$ as its vertex set.*

One records detailed quantitative data of an ideal in the coefficients of a formal power series called the Hilbert series. Throughout this work, we will assume that polynomial rings have standard grading, i.e. each variable has degree one, and the ideals will be homogeneous with respect to this grading. The quotient ring R/I of an ideal I can be written as a direct sum of its graded components, i.e.

$$R/I = [R/I]_0 \oplus [R/I]_1 \oplus \cdots \oplus [R/I]_d \oplus \dots,$$

where $[R/I]_d$ denotes the collection of all homogeneous polynomials of degree d in R/I. Since the graded components are vector spaces over $\mathbb{K}$, one can talk about their vector dimension, denoted $\dim_{\mathbb{K}}[R/I]_d$.

Definition 2.1.3. *Let I be a homogeneous ideal in a polynomial ring R in finitely many variables over some field $\mathbb{K}$. Thus, $R/I = \oplus_{d\geq 0}[R/I]_d$ is a standard graded $\mathbb{K}$-algebra. Its Hilbert series is the formal power series*

$$H_{R/I}(t) = \sum_{d\geq 0} \dim_{\mathbb{K}}[R/I]_j t^d.$$

By Hilbert's theorem (see, e.g., [4, Corollary 4.1.8]), it is rational and can be uniquely written as

$$H_{R/I}(t) = \frac{g(t)}{(1-t)^{\dim R/I}},$$

with $g(t) \in \mathbb{Z}[t]$ and $g(1) > 0$, unless $I = R$. The number $g(1)$ is called the *degree* of I.

2.2 Hierarchical Models

When one uses random variables to analyze data, one is interested in determining dependencies among the random variables. Hierarchical models are log-linear statistical models used to achieve this goal.

Example 2.2.1. *Reye's syndrome is a condition that causes swelling in the liver and brain. It can affect children and teenagers after they have an infection. It is suspected that using Aspirin to treat an infection increases the chances of getting Reye's syndrome. The question is if among patients with Reye's syndrome, there is a dependency relation between the type of infection and the use of Aspirin to treat that infection. The contingency table below cross-classifies data from 1070 patients with this syndrome in US from 1980 to 1997 according to the criteria of interest [3].*

		Type of Infection		
		Varicella	Influenza	Gastroenteritis
Use Aspirin regularly	Yes	29	21	2
	No	704	188	125

Data on the use of Aspirin from 1070 patients with Reye's Syndrome in US from 1980 to 1997.

Figure 2.1: Data

Assign the use of Aspirin and the type of infection the discrete random variables Z_1 and Z_2,respectively. We code their possible outcomes with $[2] = \{1, 2\}$ and $[3] = \{1, 2, 3\}$. Denote $\mathbf{p} = (p_{ij})$ the unknown probability distribution, where

$$p_{ij} = P(Z_1 = i, Z_2 = j),$$

and $\mathscr{P}_{2\times 3}$ the probability simplex i.e.

$$\mathscr{P}_{2\times 3} = \{\mathbf{p} = (p_{ij}) \in \mathbb{R}^{2\times 3} \ : \ p_{ij} \geq 0, \ and \ \sum_{i,j} p_{ij} = 1\}.$$

The collection of all possible probability distributions that record certain dependency relations among random variables is a hierarchical model. Denote

$$p_{i\bullet} = \sum_j p_{ij}, \ \ and \ p_{\bullet j} = \sum_i p_{ij}.$$

Two random variables are independent iff $p_{ij} = p_{i\bullet}p_{\bullet j}$ for every (i, j) in the state space $\Omega_1 \times \Omega_2$. Hence, the collection

$$\mathcal{M} = \{\mathbf{p} = (p_{ij}) \in \mathscr{P}_{2\times 3} \,|\, p_{ij} = p_{i\bullet}p_{\bullet j}, for\ all\ i \in \Omega_1, j \in \Omega_2\}$$

is the hierarchical model for Z_1 and Z_2 being independent.

In general, given discrete random variables $Z_1, \dots, Z_m$ with $r_1, \dots, r_m$ number of states, respectively, we will refer to $\mathbf{r} = (r_1, \dots, r_m)$ as the vector of states. The dependency relations among random variables are describable via a collection $\Delta = \{F_1, \cdots, F_q\}$ of non-empty subsets of $[m]$, with $\bigcup_{j \in [q]} F_j = [m]$, where every set F_j indicates a dependency among the variables indicated in the set. We refer to elements of Δ as *faces*. All subsets of a face are faces, and Δ itself is a simplicial complex. If a face is maximal, i.e. no other face in Δ contains it, we will refer to it as *facet*, and we denote the collection of facets with $\text{facet}(\Delta)$. Note that facets of Δ fully determines Δ. The notations $\mathcal{M}(\mathbf{r}, \Delta)$ and $\mathcal{M}(\mathbf{r}, \text{facet}(\Delta))$ will be used interchangeably to call the same hierarchical model. For any subset $F = \{i_1, i_2, \dots, i_s\} \subseteq [m]$, we write

$$\mathbf{r}_F = (r_{i_1}, r_{i_2}, \dots, r_{i_s}) \in \mathbb{N}^s \text{ and } [\mathbf{r}_F] = [r_{i_1}] \times [r_{i_2}] \times \cdots \times [r_{i_s}] \subseteq \mathbb{N}^s.$$

In particular, $[\mathbf{r}_{[m]}] = [\mathbf{r}] \subseteq \mathbb{N}^m$. For any F in Δ and $i_F \in [\mathbf{r}_F]$, denote the marginal sum

$$p_{\mathbf{i}_F} = \sum_{i_{[m] \setminus F}} p_{i_1 \dots i_m}.$$

With this notation one has the following definition.

Definition 2.2.2. *The hierarchical model $\mathcal{M}(\Delta, \mathbf{r})$ in m parameters with vector of states $\mathbf{r} = (r_1, \dots, r_m) \in \mathbb{N}^m$ and dependency relations $\Delta \subseteq 2^{[m]}$ is the collection of probability distributions $\mathbf{p} = (p_{i_1 \dots i_m}) \in \mathscr{P}_{r_1 \times \cdots \times r_m}$ that satisfy the independence equations*

$$p_{\mathbf{i}} = \frac{1}{Z(\mathbf{p})} \prod_{F \in facet(\Delta)} p_{\mathbf{i}_F} \text{ for all } \mathbf{i} = (i_1 \dots i_m) \in [r_1] \times \cdots \times [r_m],$$

where $Z(\mathbf{p})$ is the normalizing constant

$$Z(\mathbf{p}) = \sum_{\mathbf{i} \in [\mathbf{r}]} \prod_{F \in facet(\Delta)} p_{\mathbf{i}_F}.$$

In modeling the unknown probability distribution, one needs some critical information about the hierarchical models. Markov bases of a hierarchical models are very useful, since one can use them to run algorithms, which detect if the collected data has enough information to refuse that model. Such an example is the Metropolis-Hasting Algorithm (see[18] 1.1.13). The basic idea is to compare the given contingency table with all tables of the same size and with some extra conditions, called the same Δ-marginal sums, which will be described later in this section.

In an $r_1 \times \cdots \times r_m$ dimensional contingency table (table of counts) N, the entry $N_{\mathbf{i}}$ in position $i \in [\mathbf{r}]$ is a non-negative integer that denotes the number of units or individuals sharing the same attributes $\mathbf{i}$. Figure 2.2 visualizes a $2 \times 3 \times 2$ contingency table, with n_{ijk} being the count of items with criteria i, j and k, respectively.

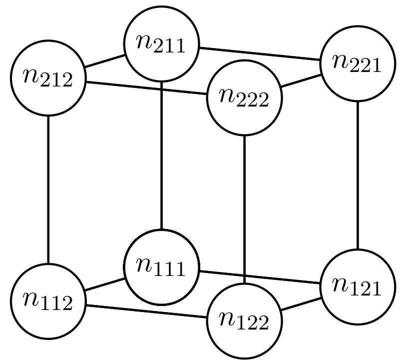


Figure 2.2: $2 \times 2 \times 2$ contingency tables

Given $F = \{i_1, \ldots, i_s\} \in \Delta$, the *F-marginal* of N is the $r_{i_1} \times \cdots \times r_{i_s}$ dimensional table N^F, where for every $\mathbf{i}_F \in [\mathbf{r}_F]$, the entry $N^F_{\mathbf{i}_F}$ is the sum over all entries of N in position $\mathbf{i}_F$, i.e.

$$N^F_{\mathbf{i}_F} = \sum_{\mathbf{i}_{[m]\setminus F} \in [\mathbf{r}_{[m]\setminus F}]} N_{\mathbf{i}}.$$

In Example 2.2.1, $N^{\{1\}} = \begin{bmatrix} 52 & 1017 \end{bmatrix}$, and $N^{\{2\}} = \begin{bmatrix} 733 & 209 & 127 \end{bmatrix}$.

One adds a table to another of the same size via summing their respective cell entries. The difference $B = N - N'$ of two tables N and N' can be interpreted as a move between tables N and $N' = N + B$. In particular, a primitive move has two entries equal to 1, and two entries equal to -1, while the remaining entries are 0. We are interested in moves B for $\mathcal{M}$ that leave the marginals in Δ unchanged, i.e the tables N^F are the same for all of them and $B^F = \mathbf{0}$ for all $F \in \Delta$.

Definition 2.2.3. *Given a table of counts $N \in \mathbb{N}^{\mathbf{r}}$, and a hierarchical model $\mathcal{M}(\mathbf{r}, \Delta)$, denote by $T^N(\mathcal{M})$ the set of tables with non-negative elements that have all Δ-marginals, $F \in \Delta$, equal to the corresponding marginals of N. A move B is allowable for $\mathcal{M}$ if $N + B$ belongs to $T^N(\mathcal{M})$.*

Definition 2.2.4. *A Markov basis $\mathcal{B}$ for a hierarchical model $\mathcal{M}(\mathbf{r}, \Delta)$ is a finite collection of moves that preserve the Δ-marginals and connect any two $r_1 \times \cdots \times r_m$-dimensional tables with the same Δ-marginals. In other words, for any table N' that belongs to $T^N(\mathcal{M})$, there exists a sequence of moves $B_1, \ldots, B_k$ in $\mathcal{B} \in \mathcal{B}$, and $\epsilon_1, \ldots, \epsilon_k \in \{\pm 1\}$, such that $N' - N = \sum_{i=1}^{k} \epsilon_i B_i$, and $N + \sum_{i=1}^{l} \epsilon_i B_i \in T^N(\mathcal{M})$, for $1 \leq l \leq k$.*

The equations say that the table N can be transformed into any table N' with the same Δ-*marginals* in $T^N(\mathcal{M})$ by employing moves in $\mathcal{B}$.

Example 2.2.5. *Let $\mathcal{M} = \mathcal{M}(\mathbf{r}, \{F_1, F_2\})$, where F_1 and F_2 are disjoint. Possibly permuting the positions of the entries of a vector $\mathbf{i} \in [\mathbf{r}] = [\mathbf{r}_{F_1 \cup F_2}]$, we write $x_{\mathbf{i}_{F_1}, \mathbf{i}_{F_2}}$ instead of $x_{\mathbf{i}}$. This corresponds to a bijection $[\mathbf{r}_{F_1 \cup F_2}] \to [\mathbf{r}_{F_1}] \times [\mathbf{r}_{F_2}]$. Using this* *notation, a Markov basis of $\mathcal{M}$ is the set of $r_1 \times \cdots \times r_m$-dimensional primitive moves (see, e.g.,[45, 10])*

$$B^{\mathbf{i}_{F_1},\mathbf{i'}_{F_1},\mathbf{i}_{F_2},\mathbf{i'}_{F_2}} = \begin{cases} 1 & \text{at positions } \{(\mathbf{i}_{F_1},\mathbf{i}_{F_2}),(\mathbf{i'}_{F_1},\mathbf{i'}_{F_2})\} \\ -1 & \text{at positions } \{(\mathbf{i'}_{F_1},\mathbf{i}_{F_2}),(\mathbf{i}_{F_1},\mathbf{i'}_{F_2})\} \\ 0 & \text{else} \end{cases}$$

for any $\mathbf{i}_{F_1} < \mathbf{i'}_{F_1} \in [\mathbf{r}_{F_2}]$, $\mathbf{i}_{F_2} < \mathbf{i'}_{F_2} \in [\mathbf{r}_{F_2}]$, compared component-wise.

The challenges of working with high dimensional tables, and the difficulty of finding a Markov basis for hierarchical models led to construction of certain toric ideals.

2.3 Toric Ideals of Hierarchical Models

Given a field $\mathbb{K}$, usually the real numbers, and a hierarchical model $\mathcal{M} = \mathcal{M}(\mathbf{r},\Delta)$, consider the following ring homomorphism:

$$\begin{aligned} \Phi_{\mathcal{M}}\colon R_{\mathbf{r}} = \mathbb{K}[x_{\mathbf{i}} \mid \mathbf{i} \in [\mathbf{r}]] &\xrightarrow{A} S_{\mathcal{M}} = \mathbb{K}[y_{j,\mathbf{i}_{F_j}} \mid F_j \in \text{facet}(\Delta), \mathbf{i}_{F_j} \in [\mathbf{r}_{F_j}]], \\ x_{\mathbf{i}} &\longmapsto \prod_{F_j \in \Delta} y_{j,\mathbf{i}_{F_j}}. \end{aligned} \tag{2.3.1}$$

The kernel of this homomorphism, denoted $I_{\mathcal{M}}$, is called the *toric ideal* to the hierarchical model $\mathcal{M}$. We also refer to $R_{\mathbf{r}}/I_{\mathcal{M}}$ as the *coordinate ring* of the model $\mathcal{M}$. Given $\mathbf{a} \in \mathbb{Z}_{\geq 0}^{r_1\times\cdots\times r_m}$, denote $\mathbf{x}^{\mathbf{a}} = \prod_{\mathbf{i}\in[\mathbf{r}]} x_{\mathbf{i}}^{a_{\mathbf{i}}}$.

Theorem 2.3.1. *[16, Theorem 3.1] A subset $\mathcal{B} \subseteq ker_{\mathbb{Z}}A$ is a Markov Basis for $\mathcal{M}$ if and only if the corresponding set of binomials $\{\mathbf{x}^{B^+} - \mathbf{x}^{B^-} \mid B = B^+ - B^- \in \mathcal{B}\}$ generates the ideal $I_{\mathcal{M}}$.*

Example 2.3.2. *(i) Let facet$(\Delta) = \{F_1, F_2\}$ with F_1 and F_2 disjoint as in example 2.2.5. A generating set of $I_{\mathcal{M}}$ is*

$$G(\mathcal{M}(\mathbf{r},\{F_1,F_2\})) = \{x_{\mathbf{i}_{F_1},\mathbf{i}_{F_2}} x_{\mathbf{i'}_{F_1},\mathbf{i'}_{F_2}} - x_{\mathbf{i}_{F_1},\mathbf{i'}_{F_2}} x_{\mathbf{i'}_{F_1},\mathbf{i}_{F_2}} \mid \mathbf{i}_{F_1} < \mathbf{i'}_{F_1} \in [\mathbf{r}_{F_1}],\ \mathbf{i}_{F_2} < \mathbf{i'}_{F_2} \in [\mathbf{r}_{F_2}]\}.$$

In the special case, where $m = 2$ and, say, $F_1 = \{1\}, F_2 = \{2\}$, this set becomes

$$\{x_{i_1,i_2} x_{i'_1,i'_2} - x_{i_1,i'_2} x_{i'_1,i_2} \mid 1 \leq i_1 \leq i'_1 \leq r_1,\ 1 \leq i_2 \leq i'_2 \leq r_2\},$$

which is the set of 2×2 minors of a generic $r_1 \times r_2$ matrix with entries x_{i_1,i_2}. The image of the map $\Phi_{\mathcal{M}}$ in this case is known in algebraic geometry as the coordinate ring of the Segre product $\mathbb{P}^{r_1-1} \times \mathbb{P}^{r_2-1}$ whose homogeneous ideal is $I_{\mathcal{M}}$.

(ii) Consider now the general case, where F_1 and F_2 are not necessarily disjoint. Note that $[m]$ is the disjoint union of $F_1\backslash F_2, F_2\backslash F_1$ and $F_1 \cap F_2$. Thus, possibly permuting the positions of the entries of $\mathbf{i} \in [r]$ as above, we write $x_{\mathbf{i}_{F_1\backslash F_2},\mathbf{i}_{F_1\cap F_2},\mathbf{i}_{F_2\backslash F_1}}$

for $x_{\mathbf{i}}$*. Fixing a vector* $\mathbf{c} \in [\mathbf{r}_{F_1 \cap F_2}]$*, we define a set* $G^{\mathbf{c}}(\mathcal{M}(\mathbf{r}_{[m]\setminus F_1\cap F_2}, \{F_1\setminus F_2, F_2\setminus F_1\})$ *whose elements are*

$$x_{\mathbf{i}_{F_1\setminus F_2},\mathbf{c},\mathbf{i}_{F_2\setminus F_1}} x_{\mathbf{i'}_{F_1\setminus F_2},\mathbf{c},\mathbf{i'}_{F_2\setminus F_1}} - x_{\mathbf{i'}_{F_1\setminus F_2},\mathbf{c},\mathbf{i}_{F_2\setminus F_1}} x_{\mathbf{i}_{F_1\setminus F_2},\mathbf{c},\mathbf{i'}_{F_2\setminus F_1}},$$

where

$$\mathbf{i}_{F_1\setminus F_2} < \mathbf{i'}_{F_1\setminus F_2} \in [\mathbf{r}_{F_1\setminus F_2}] \text{ and } \mathbf{i}_{F_2\setminus F_1} < \mathbf{i'}_{F_2\setminus F_1} \in [\mathbf{r}_{F_2\setminus F_1}].$$

The collection

$$G(\mathcal{M}(\mathbf{r}, \{F_1, F_2\})) = \bigcup_{\mathbf{c}\in[\mathbf{r}_{F_1\cap F_2}]} G^{\mathbf{c}}(\mathcal{M}(\mathbf{r}_{[m]\setminus F_1\cap F_2}, \{F_1\setminus F_2, F_2\setminus F_1\}))$$

is a generating set for the ideal $I_{\mathcal{M}(\mathbf{r},\{F_1,F_2\})}$ *(see [10]).*

Example 2.3.3. *Using Macaulay2, one computes a generating set for the ideal for* $\mathcal{M}(\{12, 23, 13\}, (2, 2, 2))$ *as the kernel of the map*

$$\begin{aligned} \Phi_{\mathcal{M}}\colon R_{222} = \mathbb{K}[x_{111}\dots x_{222}] &\longrightarrow S_{\mathcal{M}} = \mathbb{K}[y_{11}\dots y_{22}, z_{11}\dots z_{22}, w_{11}\dots w_{22}], \\ x_{ijk} &\longmapsto y_{ij}z_{jk}w_{ik} \end{aligned} \quad (2.3.2)$$

and obtains the only generator

$$x_{111}x_{122}x_{212}x_{221} - x_{112}x_{121}x_{211}x_{222},$$

which, using Theorem 2.3.1, induces the Markov basis with the following move.

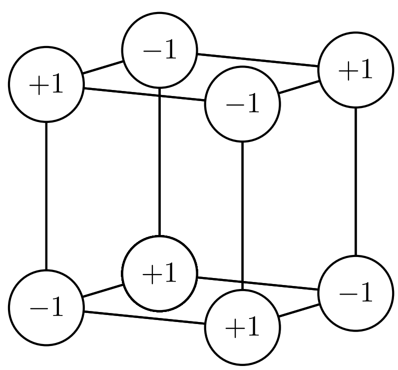


Figure 2.3: A Markov basis for $\mathcal{M}(\{12, 23, 13\}, (2, 2, 2))$

Hoşten and Sullivant in give a formula on the dimension of such ideals.

Proposition 2.3.4. *[24, Corollary 2.7] The dimension of the toric ideal* $I_{\mathcal{M}}$ *of a hierarchical model* $\mathcal{M}(\Delta, \mathbf{r})$ *is*

$$dim(R/I_{\mathcal{M}}) = \sum_{F\in\Delta} \prod_{i\in F} (r_i - 1). \quad (2.3.3)$$

2.4 On Filtrations and their Stabilization

Even in the simple cases of Example 2.3.2, the number of minimal generators of a toric ideal $I_{\mathcal{M}}$ is large if the entries of $\mathbf{r}$ are large. However, many of these generators have similar shape. This can be made precise using symmetry. Indeed, denote by S_n the symmetric group in n letters. Set $S_{[\mathbf{r}]} = S_{r_1} \times \cdots \times S_{r_m}$. This group acts on the polynomial ring $R_{\mathbf{r}}$ by permuting the indices of its variables, that is,

$$(\sigma_1, \ldots, \sigma_m) \cdot x_{\mathbf{i}} = x_{(\sigma_1(i_1), \ldots, \sigma_m(i_m))}.$$

Theorem 2.3.1 says that toric ideals have minimal generating sets consisting of binomials. Thus, the definition of the homomorphism $\Phi_{\mathcal{M}}$ in (2.3.1) implies that the ideal $I_{\mathcal{M}}$ is $S_{[\mathbf{r}]}$-invariant, that is, $\sigma \cdot f \in I_{\mathcal{M}}$ whenever $\sigma \in S_{[\mathbf{r}]}$ and $f \in I_{\mathcal{M}}$. In some cases, this invariance can be used to obtain all minimal generators of $I_{\mathcal{M}}$ from a subset by using symmetry. For example, in the special case $m = q = 2$, $F_1 = \{1\}, F_2 = \{2\}$ with $r_1, r_2 \geq 2$, the set $G(\mathcal{M}(\mathbf{r}, \{F_1, F_2\}))$ can be obtained from

$$x_{1,1}x_{2,2} - x_{1,2}x_{2,1}$$

using the action of $S_{r_1} \times S_{r_2}$. Note that this is true for every vector $\mathbf{r} = (r_1, r_2)$. There is a vast generalization of this observation using the concept of an invariant filtration.

The symmetric group S_n is naturally embedded into S_{n+1} as the stabilizer of $\{n+1\}$. Using this construction component-wise, we get an embedding of $S_{[\mathbf{r}]}$ into $S_{[\mathbf{r}']}$ if $\mathbf{r} \leq \mathbf{r}'$. Set

$$S_{\infty}^{m} = \bigcup_{\mathbf{r} \in \mathbb{N}^m} S_{[\mathbf{r}]}.$$

Definition 2.4.1. *An* S_{∞}^{m}-invariant filtration *is a family* $(I_{\mathbf{r}})_{\mathbf{r} \in \mathbb{N}^m}$ *of ideals* $I_{\mathbf{r}} \subseteq R_{\mathbf{r}}$ *such that every ideal* $I_{\mathbf{r}}$ *is* $S_{[\mathbf{r}]}$*-invariant and, as subsets of* $R_{\mathbf{r}'}$*,*

$$S_{[\mathbf{r}']} \cdot I_{\mathbf{r}} \subseteq I_{\mathbf{r}'} \quad \textit{whenever } \mathbf{r} \leq \mathbf{r}'.$$

Note that fixing Δ, the ideals $(I_{\mathcal{M}(\Delta,\mathbf{r})})_{\mathbf{r} \in \mathbb{N}^m}$ form an S_{∞}^{m}-invariant filtration. It is useful to extend these ideas.

Remark 2.4.2. *Let* T *be any non-empty subset of* $[m]$*. For vectors* $\mathbf{r} \in \mathbb{N}^m$*, we want to fix the entries in positions supported at* $[m]\backslash T$*, but vary the other entries. To this end write* $(\mathbf{r}_{[m]\backslash T}, \mathbf{r}_T)$ *instead of* $\mathbf{r}$*. Fix a vector* $\mathbf{c} \in \mathbb{N}^{m-\#T}$*. Let* $(I_{\mathbf{r}})_{\mathbf{r} \in \mathbb{N}^m}$ *be an* S_{∞}^{m}*-invariant filtration. Restricting* $S_{[\mathbf{r}]}$ *and its action to components supported at* T*, we get an* $S_{\infty}^{\#T}$*-invariant filtration of ideals* $I_{\mathbf{c},\mathbf{n}} \subseteq R_{\mathbf{c},\mathbf{n}}$ *with* $\mathbf{n} \in \mathbb{N}^{\#T}$*.*

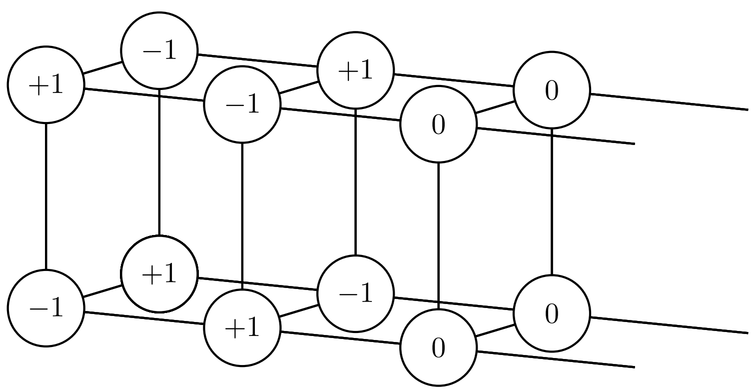
Note that this idea applies to the ideals $I_{\mathcal{M}(\Delta,\mathbf{r})}$ with fixed Δ and fixed entries in the vector of states $\mathbf{c} \in \mathbb{N}^{m-\#T}$. We can now state the mentioned extension of the Example 2.3.2 part (i). It is called the Independent Set Theorem and has been established by Hillar and Sullivant in [22, Theorem 4.7] (see also [12]).

Theorem 2.4.3. *Fix Δ and consider a subset $T \subseteq [m]$ such that $\#(F \cap T) \leq 1$ for every $F \in \Delta$. Assume the number of states of every parameter $j \in [m] \backslash T$ is fixed, and consider the hierarchical models $\mathcal{M}_{\mathbf{n}} = \mathcal{M}(\Delta, (\mathbf{c}, \mathbf{n}))$, where $\mathbf{c} \in \mathbb{N}^{m-\#T}$ and $\mathbf{n} \in \mathbb{N}^{\#T}$. Then the ideals $\mathscr{I}_{\Delta,\mathbf{c}} = (I_{\mathcal{M}_{\mathbf{n}}})$, form an $S_{\infty}^{\#T}$-invariant filtration, that is, there is some $\mathbf{d} \in \mathbb{N}^{\#T}$ such that $S_{[\mathbf{n}]} \cdot I_{\mathcal{M}_{\mathbf{d}}}$ generates in $R_{\mathbf{n}}$ the ideal $I_{\mathcal{M}_{\mathbf{n}}}$ whenever $\mathbf{n} \geq \mathbf{d}$.*

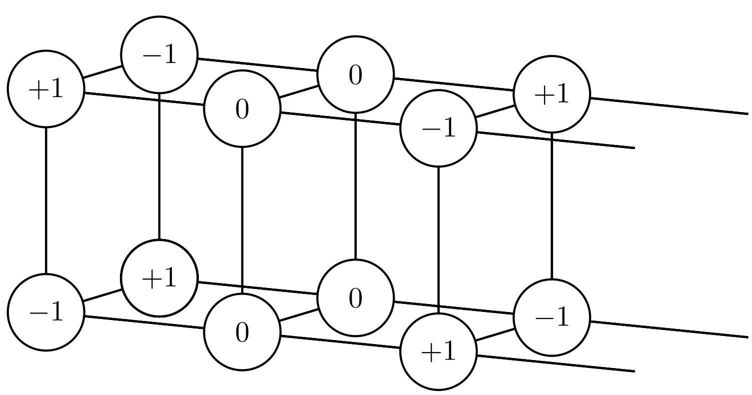
In other words, this result says that a generating set of the ideal $I_{\mathcal{M}(\Delta,\mathbf{r})}$ can be obtained from a fixed finite minimal generating set of $I_{\mathcal{M}(\Delta,(\mathbf{c},\mathbf{d}))}$ by applying suitable permutations whenever the number of states of every parameter in $[m] \backslash T$ is large enough.

Example 2.4.4. *Let $\mathscr{I} = (I_r)_{r \geq 2}$ be the filtration of ideals from hierarchical models $\mathcal{M}_r = \mathcal{M}(\{12, 23, 13\}, (2, r, 2))$, associated with the action by S_{∞} on the second component in the index vector of each variable, i.e. $\sigma \cdot x_{ijk} = x_{i\sigma(j)k}$. For any $r > 2$, the set $G_r = S_r(x_{111}x_{122}x_{212}x_{221} - x_{112}x_{121}x_{211}x_{222})$ is a generating set for I_r.*

To visualize what happens to the Markov moves when one applies such a group action, first observe that a move of a hierarchical model can be embedded as a move of another hierarchical model with the same dependency relation and with larger number of states by padding zeros in the added entries. The move in Figure 2.3 of $\mathcal{M}_2$ can be treated as a move of $\mathcal{M}_r, r > 2$ as below:



For $\sigma = (23)$, $\sigma \cdot f$ induces the Markov move:



Theorem 2.4.3 is not true without an assumption on the set T (see [22, Example 4.3]).

In order to study asymptotic properties of ideals in an S_∞-invariant filtration, an equivariant Hilbert series was introduced in [32]. Here we study an extension of this concept for S_∞^m-invariant filtrations.

Definition 2.4.5. *The* equivariant Hilbert series *of an S_∞^m-invariant filtration $\mathscr{I} = (I_\mathbf{r})_{\mathbf{r}\in\mathbb{N}^m}$ of ideals $I_\mathbf{r} \subseteq R_\mathbf{r}$ is the formal power series in variables $s_1, \dots, s_m, t$*

$$\begin{aligned} equivH_{\mathscr{I}}(s_1, \dots, s_m, t) &= \sum_{\mathbf{r}\in\mathbb{N}^m} H_{R_\mathbf{r}/I_\mathbf{r}}(t) \cdot s_1^{r_1} \cdots s_m^{r_m} \\ &= \sum_{\mathbf{r}\in\mathbb{N}^m} \sum_{j\geq 0} \dim_\mathbb{K}[R_\mathbf{r}/I_\mathbf{r}]_j \cdot s_1^{r_1} \cdots s_m^{r_m} t^j. \end{aligned}$$

If $m = 1$, that is, $\mathscr{I}$ is an S_∞-invariant filtration, the Hilbert series of $\mathscr{I}$ is always rational by [32, Theorem 7.8] or [28, Theorem 4.3]. For $m \geq 1$, one can also consider another formal power series by focusing on components whose degree is on the diagonal of $\mathbb{N}^m$. This gives

$$\sum_{r\geq 1} H_{R_{(r,\dots,r)}/I_{(r,\dots,r)}}(t) \cdot s^r.$$

It is open whether this formal power series is rational when $m \geq 2$, even if the ideals are trivial (see Example 4.5.1).

Remark 2.4.6. *An S_∞^m-invariant filtration can also be described using a categorical framework. Indeed, in the case $m = 1$, this approach has been used in [33] to study also sequences of modules by using the category* FI*, whose objects are finite sets and whose morphisms are injections. This approach can be extended to any $m \geq 1$ using the category* FI^m *(see, e.g., [29] in the case of modules over a fixed ring). For conceptual simplicity we prefer to use invariant filtrations.*

2.5 Regular Languages, Finite Automata, and their Power Series

Let Σ be a collection of symbols. We refer to Σ as an alphabet, and to the elements of Σ as letters. Let Σ^* be the free monoid on Σ. We refer to its elements as words. The empty word is denoted by ϵ. A *formal language* with words in the alphabet Σ is a subset of Σ^*. The class of regular languages on Σ is the smallest class of languages containing the singleton languages for each letter in the alphabet, the empty word language ϵ, and is closed under union, concatenation, and Kleene star. Kleene star $\mathcal{L}^*$ of a language $\mathcal{L} \subseteq \Sigma^*$ is the collection of words of the form $w_1 \dots w_n$, where $w_i \in \mathcal{L}$, i.e. is the submonoid of Σ^* generated by $\mathcal{L}$.

Let $T = \mathbb{K}[s_1, \dots, s_k]$ be a polynomial ring in k variables and denote by $\mathrm{Mon}(T)$ the set of monomials in T. A *weight function* on Σ^* is a monoid homomorphism

$\rho\colon \Sigma^* \to \mathrm{Mon}(T)$ such that $\rho(w) = 1$ only if w is the empty word. The corresponding generating function is a formal power series in variables $s_1, \ldots, s_k$:

$$P_{\mathcal{L},\rho}(s_1, .., s_k) = \sum_{w \in \mathcal{L}} \rho(w).$$

Theorem 2.5.1 ([25])**.** *If ρ is any weight function on a regular language $\mathcal{L}$ then $P_{\mathcal{L},\rho}$ is a rational function in $\mathbb{Q}(s_1, .., s_k)$.*

Example 2.5.2. *The language $\mathcal{L} = \{ab^n \mid n \in \mathbb{N}\} \subset \{a, b\}^*$ is a regular language as concatenation of $\{a\}^*$ and $\{b\}$, which are both regular. The weight function with $\rho(a) = t$ and $\rho(b) = s$ induces*

$$P_{\mathcal{L},\rho}(t, s) = \sum_{n \geq 1} \rho(ab^n) = \sum_{n \geq 1} ts^n = \frac{ts}{1 - s}.$$

One can show that a language $\mathcal{L}$ is a regular language by proving that $\mathcal{L}$ is recognizable by a finite automaton ([27, Theorems 3.4 and 3.7]). A *finite automaton* on an alphabet Σ is a 5-tuple $\mathcal{A} = (P, \Sigma, \delta, p_0, F)$ that consist of a finite set P of *states*, an *initial state* $p_0 \in P$, a set $F \subseteq P$ of *accepting states* and a *transition map* $\delta\colon D \to P$, where D is some subset of $P \times \Sigma$. The automaton $\mathcal{A}$ *recognizes* or *accepts* a word $w = a_1 a_2 \ldots a_s \in \Sigma^*$ if there is a sequence of states $r_0, r_1, \ldots, r_s$ satisfying $r_0 = p_0$, $r_s \in F$ and

$$r_{j+1} = \delta(r_j, a_{j+1}) \quad \text{whenever } 0 \leq j < s.$$

In words, the automaton starts in state p_0 and transitions from state r_j to a state r_{j+1} based on the input a_{j+1}. The word w is accepted if r_s is an accepting state. If $\delta(p, a)$ is not defined the machine halts. The automaton $\mathcal{A}$ *recognizes a formal language* $\mathcal{L} \subseteq \Sigma^*$ if $\mathcal{L}$ is precisely the set of words in Σ^* that are accepted by $\mathcal{A}$.

Remark 2.5.3. *Any finite automaton $\mathcal{A} = (P, \Sigma, \delta, p_0, F)$ can be represented by a labeled directed graph whose vertex set is the set of states P. Accepting states are indicated by double circles. There is an edge from vertex p to vertex p' if there is a transition $\delta(p, a) = p'$. In that case, the edge is labeled by all $a \in \Sigma$ such that $\delta(p, a) = p'$.*

Example 2.5.4. *The following finite automaton recognizes $\mathcal{L}$ from Example 2.5.2.*

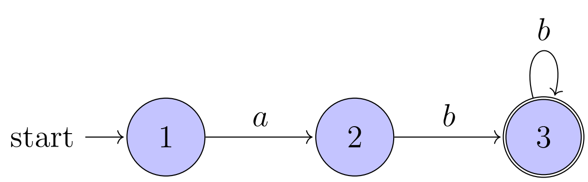


Figure 2.4: An example of a finite automaton

2.6 Gale Transformations

Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be vectors in $\mathbb{R}^{d-1}$ whose affine hull is full dimensional and let P be their convex hull. Then, the matrix below has rank d.

$$M := \begin{bmatrix} 1 & 1 & \dots & 1 \\ \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \end{bmatrix} \in \mathbb{R}^{d \times n}.$$

By the Rank-Nullity Theorem, the dimension of the kernel of M is $n - d$. Let $B_1, \dots, B_{n-d} \in \mathbb{R}^n$ be a basis for the vector space $\ker(A)$. If we organize these vectors as the columns of an $n \times (n-d)$ matrix,

$$B := [\, B_1 \ \dots \ B_{n-d} \,],$$

we see that $MB = \mathbf{0}$.

Definition 2.6.1. *Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_{n-d}\} \in \mathbb{R}^{n-d}$ be the $n - d$ ordered rows of B. Then $\mathcal{B}$ is called a* Gale transformation *of $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$.*

Example 2.6.2. *Consider the polytope defined as the convex hull of the columns $\mathbf{v}_1, \dots, \mathbf{v}_8$ of the following matrix A:*

$$A = \begin{array}{c} \begin{matrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 & \mathbf{v}_5 & \mathbf{v}_6 & \mathbf{v}_7 & \mathbf{v}_8 \end{matrix} \\ \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \end{bmatrix} \end{array}$$

The kernel for $M = \begin{bmatrix} \mathbf{1} & A \end{bmatrix}^T$ is generated by the vector

$$\begin{matrix} \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 & \mathbf{b}_4 & \mathbf{b}_5 & \mathbf{b}_6 & \mathbf{b}_7 & \mathbf{b}_8 \\ [\,1 & -1 & -1 & 1 & -1 & 1 & 1 & -1\,] \end{matrix},$$

which is the only row in the matrix B. From here one obtains the Gale transformation:

$$\mathbf{b}_1 = \mathbf{b}_4 = \mathbf{b}_6 = \mathbf{b}_7 = 1, \ \mathbf{b}_2 = \mathbf{b}_3 = \mathbf{b}_5 = \mathbf{b}_8 = -1.$$

The interior of a polytope in $\mathbb{R}^d$ is the set of all points in the polytope such that we can fit a d- dimensional ball centered at this point, of infinitesimal radius, entirely inside the polytope. A polytope has an interior if and only if it is full-dimensional. However, given a non-full-dimensional polytope in $\mathbb{R}^d$, the polytope has an interior if considered as a polytope in its affine hull, where it is a full-dimensional polytope. This is known as the *relative interior* of the polytope, denoted *relint*(P).

Theorem 2.6.3. *Let $P = conv(\{\mathbf{v}_1, \ldots, \mathbf{v}_n\})$, $\mathbf{v}_i \in \mathbb{R}^{d-1}$, and let $\mathcal{B}$ be a Gale transformation of $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Then, for $J \subseteq [n]$, $P_J = \{\mathbf{v}_j \mid j \notin J\}$ is a face of P if and only if either $J = \emptyset$ or $\mathbf{0} \in relint(conv(\{\mathbf{b}_k \mid k \in J\}))$.*

Example 2.6.4. *Given $J = \{i, j\}$, for $i \in \{1, 4, 6, 7\}$ and $j \in \{2, 3, 5, 8\}$, the set P_J is a face for the polytope in Example 2.6.2.*

2.7 Switch Operators for Cut Polytopes

Definition 2.7.1. *Let $G([m], E)$ be a graph. The cut polytope of G, denoted $cutP_\Delta$, is the convex hull of vectors $\{\mathbf{v}^{(A|B)} \mid A \cup B \subset [m], A \cap B = \emptyset\}$ in $\mathbb{R}^E$, and*

$$\mathbf{v}_e^{(A|B)} = \begin{cases} 1 & \#(e \cap A) = 1, \\ 0 & \text{else.} \end{cases}$$

Since a partition $(A|B)$ is fully defined by one of its parts, in the future $\mathbf{v}_e^{(A|B)}$ will be denoted $\mathbf{v}_e^A$.

Given the partition $(A|B)$ of $[m]$, the set of edges $\mathcal{C}_A = \{e \in E \mid \#(e \cap A) = 1\}$ is called a cut of the graph, since it cuts the graph in two pieces and every edge in $\mathcal{C}_A$ connects two components of $G([m], E \backslash \mathcal{C}_A)$.

Given $T \subseteq E$, define the collection of symmetric differences with other sets in E:

$$\mathcal{S}_T = \{S \triangle T \mid S \subseteq E\}.$$

The set of cuts is closed under taking symmetric differences. For $\mathbf{a} \in \mathbb{R}^E$, and $S \in [m]$, let $\mathbf{a}(S) \in \mathbb{R}^E$ be defined by

$$\mathbf{a}(S)_e := \begin{cases} -\mathbf{a}_e^S & e \in S, \\ \mathbf{a}_e^S & \text{else.} \end{cases}$$

If $\chi^S \in \mathbb{R}^E$ is the indicator vector, consider the mapping $r_S : \mathbb{R}^E \to \mathbb{R}^E$ defined by $r_S(\mathbf{x}) := \mathbf{x}(S) + \chi^B$, for $\mathbf{x} \in \mathbb{R}^E$, i.e.,

$$(r_S(\mathbf{x})))_e = \begin{cases} 1 - \mathbf{x}_e & e \in S, \\ \mathbf{x}_e & \text{else.} \end{cases}$$

The mapping r_S is an affine bijection of the space $\mathbb{R}^E$, called *switching mapping*. One has the following result ([11]).

Proposition 2.7.2. *Given $\mathbf{a} \in \mathbb{R}^E$, $\mathbf{a}^t$ its transpose, $c \in R$, and $T \subseteq [m]$, the following assertions are equivalent.*

i. The inequality $\mathbf{ax} \leq c$ is valid or facet inducing for the cut polytope, respectively.

ii. The inequality $(\mathbf{a}(S))\mathbf{x} \leq c - \mathbf{a}^t\mathbf{v}^S$ is valid or facet inducing for the cut polytope, respectively.

Chapter 3 Stabilizations of Markov Bases for Hierarchical Models

This chapter concerns the Markov bases for non-reducible models, with an emphasis on the cycle models. We work in details a class of $four$-cycle models in Theorem 3.3.3. The proof combines both the algebraic and the statistical approach. In particular, toric ideals are used to guess a generating set up to symmetry for the ideals of hierarcical models. We prove that the corresponding set of moves satisfies the conditions of a Markov basis. Some analysis on the stabilization of filtrations of hierarchical models with a decomposable, reducible, and non-reducible dependency structure is given.

3.1 Preliminaries

It is a basic problem to describe a Markov basis of a hierarchical model since Markov bases are essential tools in statistical sampling. For example, the algorithm in [18, Chapter 1] takes as input a Markov basis and a contingency table, and produces an aperiodic, reversible and irreducible Markov chain that has stationary distribution equal to the conditional distribution in the fiber $\mathcal{T}^N(\mathcal{M})$ of the chi-square statistics $\chi^2(N)$. Unfortunately, due to the conditions in their definition, and their large size, Markov bases are very difficult to compute. We seek for alternative ways of computing them. Diaconis and Sturmfels [16, Theorem 3.1] showed that Markov bases can be found in the exponent vectors of generating sets of certain toric ideals that are induced by the independence equations in Definition 2.2.2. From here, one can use the theory of toric ideals, software such as `Macaulay2` and `FourTiTwo`, and combinations of different methods, to compute these generating sets. One difficulty is that these ideals have large minimal generating sets for large number of states, even if the simplicial complex Δ is simple. The number of generators grows rapidly when we increase r_i's incrementally. However, one can control this growth via symmetry.

Example 3.1.1. *Any ideal I_{r_1,r_2} arising from facet$(\Delta) = \{\{1\},\{2\}\}$ and $r_1, r_2 \geq 2$ has a minimal generating set of $\binom{r_1}{2}\binom{r_2}{2}$ elements, which can be obtained via applying $S_{r_1} \times S_{r_2}$ to the indices of the binomial $x_{12}x_{21} - x_{11}x_{22}$. The latter one serves as a generating set for the ideal with vector of states $(2,2)$. We say that the ideals arising from Δ stabilize up to S_∞^2–symmetry at the vector of states $(2,2)$.*

Unfortunately stabilization does not happen in every case. The smallest such examples are ideals with a 3-cycle as their simplicial complex; these ideals do not stabilize up to S_∞^3-symmetry. Hoşten and Sullivant in [24] prove that the stabilization results for models depend on the stabilization results for the non-reducible parts of the simplicial complex. From here, the plan is to discover stabilization results for the non-reducible models. The 3-cycle model is of interest for many mathematicians (for example [2, 15, 22]), and the work done until now confirms how surprisingly complicated they are. Sullivant introduced a new approach with toric fiber products in [40] and, with Rauh, they used the toric fiber products to produce Markov bases for some non-reducible models [36]. This approach demands the semigroup algebra for the toric ideal to be normal, which is not the case for the majority of non-reducible models.

Throughout this chapter $V(\Delta)$ will denote the vertex set used by a simplicial complex Δ. Given $\Gamma \subseteq \Delta$ a simplicial complex, denote $\mathbf{r}_{V(\Gamma)}$ to be the restriction of vector $\mathbf{r}$ to coordinates used by Δ. The hierarchical model $\mathcal{M}(\Gamma, \mathbf{r}_{V(\Gamma)})$ is the hierarchical model induced by Γ. If Γ uses all the ground set of Δ, i.e. $\mathbf{r}_{V(\Gamma)} = \mathbf{r}$, the corresponding ideals are contained in the same ring. Moreover, since Γ has less conditions to be satisfied, one must have that $I_{\mathcal{M}(\Delta,\mathbf{r})} \subseteq \mathcal{M}(\Gamma, \mathbf{r})$.

To make the exposition easier, we involve the tableau notation for monomials presented in [22]. To each monomial $x_{\mathbf{k}_1} \dots x_{\mathbf{k}_d}$, where $\mathbf{k}_j = (k_{j1}, \dots, k_{jm})$ for $j = 1, \dots, d$, we associate the $d \times n$ tableau

$$\begin{bmatrix} k_{11} & \dots & k_{1m} \\ \vdots & \ddots & \vdots \\ k_{d1} & \dots & k_{dm} \end{bmatrix}.$$

If a variable occurs to its pth power in a monomial, its corresponding index set occurs p times in the tableau. Also, two tableaus are equal when they are equal up to a permutation of the rows, since the rows of the tableau are the indices for commuting variables. The following example describes the transitions between the monomial, its tableau notation, and the Markov move for it from the Fundamental Theorem of Markov Bases.

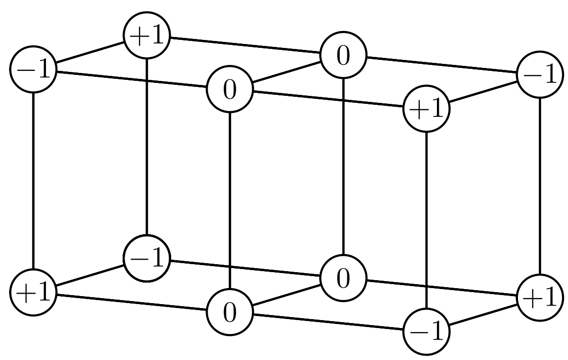
Example 3.1.2. *The binomial*

$$x_{111}x_{132}x_{212}x_{231} - x_{112}x_{131}x_{211}x_{232}$$

in Example 2.3.3 has tableau notation

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 2 & 1 & 2 \\ 2 & 3 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 1 & 2 \\ 1 & 3 & 1 \\ 2 & 1 & 1 \\ 2 & 3 & 2 \end{bmatrix}.$$

and corresponds to the Markov move below (see Example 2.4.4).



3.2 Reducible Models

Definition 3.2.1. *A simplicial complex* Δ *is a reducible simplicial complex if it is a simplex or there exists a proper decomposition* (Δ_1, S, Δ_2) *of* Δ*, i.e.* $\Delta = \Delta_1 \cup \Delta_2$ *and* $\Delta_1 \cap \Delta_2 = 2^S$ *is a simplex. A simplicial complex* Δ *is decomposable simplicial complex if it is a simplex or there exists a proper decomposition* (Δ_1, S, Δ_2) *of* Δ *with* Δ_1 *and* Δ_2 *both decomposable simplicial complexes.*

The set S is often called a separator. As noticed by their definitions, the decomposable simplicial complexes are reducible. Hierachical models with a decomposable simplicial complex as Δ are called *decomposable hierarchical models*, and the ones with a reducible simplicial complex are called *reducible hierarchical models*. Given a reducible model $\mathcal{M}$, denote $\mathcal{M}_1$ and $\mathcal{M}_2$ the models induced by its components Δ_1 and Δ_2.

Example 3.2.2. *The hierarchical model* $\mathcal{M} = \mathcal{M}(\{12, 23, 34, 14, 13\}, (r_1, r_2, r_3, r_4))$ *is reducible. The hierarchical models*

$$\mathcal{M}_1(\{\{1,2\},\{1,3\},\{2,3\}\},(r_1,r_2,r_3)) \text{ and } \mathcal{M}_2(\{\{1,3\},\{3,4\},\{1,4\}\},(r_1,r_3,r_4)),$$

are its reduced models. On the other side, $\mathcal{M}_1$ *and* $\mathcal{M}_2$ *are not reducible.*

Given a reducible hierarchical model $\mathcal{M}(\Delta, \mathbf{r})$, with G_i sets of binomials in $I_{\mathcal{M}_i}$, define the sets of binomials $Ext(G_1 \to I_{\mathcal{M}})$, $Ext(G_2 \to I_{\mathcal{M}})$ and $G(\{V(\Delta_1), V(\Delta_2)\}, \mathbf{r})$ as follows.

Definition 3.2.3. *Given* $\Delta = \Delta_1 \cup \Delta_2$*, with* $\Delta_1 \cap \Delta_2 = 2^S$*, define* $G(\{V(\Delta_1), V(\Delta_2)\}, \mathbf{r})$ *to be the collection of quadratics*

$$f = \begin{bmatrix} p_1 & s & q_1 \\ p_2 & s & q_2 \end{bmatrix} - \begin{bmatrix} p_1 & s & q_2 \\ p_2 & s & q_1 \end{bmatrix},$$

for any $p_1 \neq p_2 \in [\mathbf{r}_{V(\Delta_1 \setminus S)}]$*, and* $s \in [\mathbf{r}_S]$*, and* $q_1, q_2 \in [\mathbf{r}_{V(\Delta_2 \setminus S)}]$*.*

Definition 3.2.4. *Let* H *in either* $I_{\mathcal{M}_1}$ *or* $I_{\mathcal{M}_2}$ *be a set of binomials*

$$f = \begin{bmatrix} p_1 & q_1 \\ \vdots & \vdots \\ p_d & q_d \end{bmatrix} - \begin{bmatrix} p'_1 & q_1 \\ \vdots & \vdots \\ p'_d & q_d \end{bmatrix},$$

with $p_i, p'_i \in [\mathbf{r}_{V(\Delta_1 \setminus S)}]$*, and* $q_i \in [\mathbf{r}_S]$*. For any* $r_1, \dots, r_d \in [\mathbf{r}_{V(\Delta_2 \setminus S)}]$ *Define the binomials* f' *from* f *by appending the* r_i *as*

$$f' = \begin{bmatrix} p_1 & q_1 & r_1 \\ \vdots & \vdots & \\ p_d & q_d & r_d \end{bmatrix} - \begin{bmatrix} p'_1 & q_1 & r_1 \\ \vdots & \vdots & \\ p'_d & q_d & r_d \end{bmatrix},$$

and define $Ext(H \to I_{\mathcal{M}})$ *to be the set of all* f' *as* f *ranges over* H *and* $r_1, \dots, r_d \in [\mathbf{r}_{V(\Delta_2 \setminus S)}]$*.*

Theorem 3.2.5. *[24, Theorem 4.17] Let $\mathcal{M}$ be a reducible model with induced submodels $\mathcal{M}_1$ and $\mathcal{M}_2$, and Grobner bases G_1 and G_2 with respect to some orderings. Then the set*

$$G = Ext(G_1 \to I_{\mathcal{M}}) \cup Ext(G_2 \to I_{\mathcal{M}}) \cup G(\{V(\Delta_1), V(\Delta_2)\}, \mathbf{r})$$

is a Gröbner basis for $I_{\mathcal{M}}$ with respect to a term order described in Lemma 4.16 of [22].

Remark 3.2.6. *In [40, Section 3.2], Sullivant introduces the toric fiber product as an alternative way of producing generating sets for reducible hierarchical models out of generating sets for the reduced parts. More specifically, if $S \subseteq F_i \cap F_j$ is a separator for the reducible simplicial complex Δ of the hierarchical model $\mathcal{M}$, then*

$$I_{\mathcal{M}} = I_{\mathcal{M}_1} \times_{\mathcal{A}} I_{\mathcal{M}_2},$$

where elements in $\mathcal{A}$ are the linearly independent multi-degree vectors for the variables in $S_{\mathcal{M}}$:

$$a_{F,\mathbf{j}_F} = \deg(y_{F,\mathbf{j}_F}) = \begin{cases} e_{(\mathbf{j}_F)_S} & if F \in \{F_1, F_2\} \\ 0 & otherwise. \end{cases}$$

The vector $e_{\mathbf{i}_S}$ for $\mathbf{i}_S \in [\mathbf{r}_S]$ is the standard unit vector in $\mathbb{Z}^{[\mathbf{r}_S]}$ with a 1 in the $\mathbf{i}_S$ position and a zero elsewhere.

The generating set in Theorem 3.2.5 induces the following result for the filtration of ideals from reducible hierarchical models.

Corollary 3.2.7. *Let $\mathscr{I}_{\Delta,\mathbf{c}}$ be a filtration of ideals with respect to $S_{\infty}^{\#T}$ induced by hierarchical models with a reducible simplicial complex Δ. The filtration stabilizes iff the induced filtrations from Δ_1 and Δ_2 stabilize. More particularly, let S be a separator of Δ, and $\mathbf{a} \in \mathbb{N}^{V(\Delta_1)}$ $\mathbf{b} \in \mathbb{N}^{V(\Delta_1)}$ be vectors of stabilization for $\mathcal{M}_1$ and $\mathcal{M}_2$ respectively. A vector of stabilization for $\mathscr{I}_{\Delta,\mathbf{c}}$ has entries*

$$r_i^0 = \begin{cases} a_i & if\ i \in V(\Delta_1)\backslash S \\ max\{a_i, b_i\} & if\ i \in S \\ b_i & if\ i \in V(\Delta_2)\backslash S. \end{cases}$$

Decomposable hierarchical models are a nice class of reducible models that have square-free quadratic Gröbner bases (see [10]). The vector with all entries equal to two works as a vector of stabilization for any filtration of ideals from hierarchical models with a decomposable simplicial complex.

Corollary 3.2.8. *Let $\mathscr{I}_{\Delta,\mathbf{c}}$ be an invariant filtration ideals with respect to $S_{\infty}^{\#T}$, induced by hierarchical models with a decomposable simplicial complex as Δ. The filtration stabilizes after $r_t \geq 2$ for all $t \in T$. This bound is sharp only when $\#\mathrm{facet}(\Delta) \leq 2$.*

Proof. The corollary will be proved via induction on the number of facets in Δ. If $\#\text{facet}(\Delta) = 1$, then $I_{\mathcal{M}(\mathbf{r},\Delta)} = \{0\}$, and the filtration stabilizes at $\mathbf{r}^0 = \mathbf{1}$. When $\text{facet}(\Delta) = \{F_1, F_2\}$, the vector $\mathbf{r}^0$ with $\mathbf{r}^0_{F_1 \triangle F2} = \mathbf{2}$ and $\mathbf{r}^0_{F_1 \cap F2} = \mathbf{1}$ serves as a vector of stabilization. Example 2.3.2 describes a generating set for these ideals. Assume the Corollary 3.2.8 is true for hierarchical models with less than n facets. Let $\mathcal{M}$ be a decomposable hierarchical model with n facets. Its reduced models Δ_1 and Δ_2 must have strictly less than n facets. Corollary 3.2.7 ends the proof. □

3.3 Non-reducible Hierarchical Models

We will call a non reducible hierarchical model any hierarchical model that is not reducible. The smallest such examples are hierarchical models that arise from the three-cycle simplicial complex, i.e. when $\text{facet}(\Delta) = \{12, 23, 13\}$. In general, any m-cycle induces a non-reducible hierarchical models. There is literature about them, (see for example [2, 1, 36]), but still little is known due to their complicated bases. De Loera and Onn in [15] prove that even the *three-cycle model* is enough complicated, by proving that for any positive integer, there exists a vector of states $\mathbf{r} = (r_1, r_2, r_3)$ that requires generators of that degree. Here we describe a Markov basis for the hierarchical model with $C_4 = \{12, 23, 34, 14\}$ as the facets of its simplicial complex and vector of states $\mathbf{r} = (r_1, 2, r_3, 2)$, where r_1, r_2 are any two natural numbers. The proof uses some binomials in the ideal for the three-cycle model, the result for the reducible case Theorem 3.2.5, and the technique used by Aoki and Takemura in [2] in proving that a set of moves is a Markov basis. Rauh and Sullivant in [36] give a proof using the toric fiber products. We hope that the approach presented in this work provides some intuition on easier ways to produce Markov bases for non-reducible hierarchical models.

Lemma 3.3.1. *Given facet*$(\Delta) = \{12, 23, 13\}$*, and* $\mathbf{r} = (r_1, r_2, r_3) \in \mathbb{N}^3$*, the ideal* $I_{\mathcal{M}(\Delta,\mathbf{r})}$ *contains the set* Q *of quartics:*

$$\begin{bmatrix} i_1 & j_1 & k_1 \\ i_1 & j_2 & k_2 \\ i_2 & j_1 & k_2 \\ i_2 & j_2 & k_1 \end{bmatrix} - \begin{bmatrix} i_2 & j_1 & k_1 \\ i_2 & j_2 & k_2 \\ i_1 & j_1 & k_2 \\ i_1 & j_2 & k_1 \end{bmatrix},$$

where $i_1 \neq i_2 \in [r_1]$*,* $j_1 \neq j_2 \in [r_2]$*,* $k_1 \neq k_2 \in [r_3]$*.*

Proof. From Example 2.3.3 the quartic binomial

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 2 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 1 & 1 \\ 2 & 2 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

is in the ideal $I_{2,2,2}$. The invariant property of filtrations Definition 2.4.1 ends the proof. □

Consider the *three*-cycle hierarchical models

$$\mathcal{M}_1 = \mathcal{M}(\{23, 34, 24\}, (2, r_3, 2)),\ \mathcal{M}_2 = \mathcal{M}(\{13, 34, 14\}, (r_1, r_3, 2)),$$

$$\mathcal{M}_3 = \mathcal{M}(\{12, 24, 14\}, (r_1, 2, 2)),\ \mathcal{M}_4 = \mathcal{M}(\{12, 23, 13\}, (r_1, 2, r_3)),$$

with their respective sets of quartics Q_1, Q_2, Q_3, Q_4 from Lemma 3.3.1, and the reducible hierarchical models

$$\mathcal{M}_{13} = \mathcal{M}(\{12, 23, 34, 14, 13\}, (r_1, 2, r_3, 2)),\ \mathcal{M}_{24} = \mathcal{M}(\{12, 23, 34, 14, 24\}, (r_1, 2, r_3, 2).$$

Extend the quartics Q_i, using Definition 3.2.4 to obtain the sets of binomials

$$Ext(Q_3 \to I_{\mathcal{M}_{24}}) \cup Ext(Q_1 \to I_{\mathcal{M}_{24}}) \subseteq I_{\mathcal{M}_{24}}$$

and

$$Ext(Q_2 \to I_{\mathcal{M}_{13}}) \cup Ext(Q_4 \to I_{\mathcal{M}_{13}}) \subseteq I_{\mathcal{M}_{13}}.$$

Example 3.3.2. *Given any $k_1, k_2, k_3, k_4 \in [r_3]$, the following quartic in Q_3:*

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 1 & 1 \\ 2 & 2 & 2 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

extends to the quartic in M_{24}:

$$\begin{bmatrix} 1 & 1 & k_1 & 1 \\ 1 & 2 & k_2 & 2 \\ 2 & 1 & k_3 & 2 \\ 2 & 2 & k_4 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 1 & k_1 & 1 \\ 2 & 2 & k_2 & 2 \\ 1 & 1 & k_3 & 2 \\ 1 & 2 & k_4 & 1 \end{bmatrix}.$$

Denote G_{13} and G_{24} the collections of binomials in $\mathcal{M}_{13}$ and $\mathcal{M}_{24}$, respectively:

$$G_{13} = Ext(Q_2 \to I_{\mathcal{M}_{13}}) \cup Ext(Q_4 \to I_{\mathcal{M}_{13}}) \cup G(\{123, 134\}, (r_1, 2, r_3, 2)),$$

$$G_{24} = Ext(Q_1 \to I_{\mathcal{M}_{24}}) \cup Ext(G_3 \to I_{\mathcal{M}_{24}}) \cup G(\{124, 234\}, (r_1, 2, r_3, 2)).$$

Now we are ready for the main theorem.

Theorem 3.3.3. *The collection of moves*

$$\mathcal{B} = \{B = B^+ - B^- \mid x^{B^+} - x^{B^-} \in G_{13} \cup G_{24}\}$$

is a Markov basis for the hierarchical model $\mathcal{M}(\{12, 23, 34, 14\}, (r_1, 2, r_3, 2))$.

Let N and N' be two 4-way contingency tables of the same size, with the same given marginal totals, i.e.

$$n_{ij\bullet\bullet} = \sum_{k,l} n_{ijkl} = \sum_{k,l} n'_{ijkl} = n'_{ij\bullet\bullet}, \quad n_{i\bullet\bullet l} = \sum_{j,k} n_{ijkl} = \sum_{j,k} n'_{ijkl} = n'_{i\bullet\bullet l},$$

$$n_{\bullet\bullet kl} = \sum_{ij} n_{ijkl} = \sum_{ij} n'_{ijkl} = n'_{\bullet\bullet kl}, \quad n_{\bullet jk\bullet} = \sum_{i,l} n_{ijkl} = \sum_{i,l} n'_{ijkl} = n'_{\bullet jk\bullet},$$

Note that these marginal totals for the table $A = N - N'$ are all zero.

$$a_{ij\bullet\bullet} = \sum_{k,l} a_{ijkl} = 0, \qquad (3.3.1) \qquad\qquad a_{i\bullet\bullet l} = \sum_{j,k} a_{ijkl} = 0, \qquad (3.3.2)$$

$$a_{\bullet\bullet kl} = \sum_{ij} a_{ijkl} = 0, \qquad (3.3.3) \qquad\qquad a_{\bullet jk\bullet} = \sum_{i,l} a_{ijkl} = 0. \qquad (3.3.4)$$

Figure 3.1 visualizes the marginal conditions (entries adjacent with the same color must sum to zero).

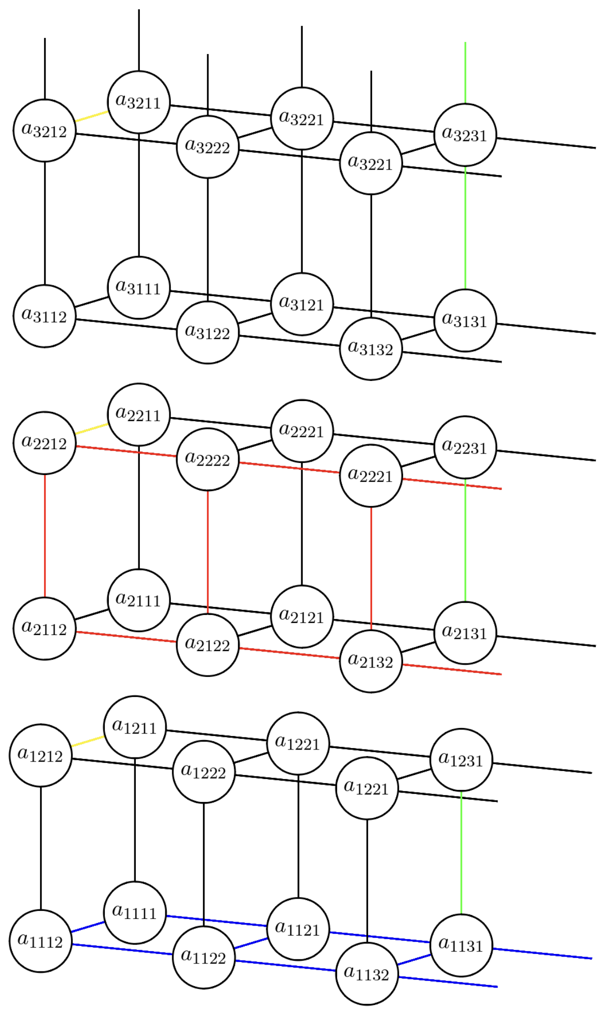


Figure 3.1: Marginal sums

The following is an example of such a move and how one obtains other moves via the action of symmetry group.

Example 3.3.4. *For $k_1 = k_2 = k_3 = k_4 = 1$ in example 3.3.2 one has the following Markov move*

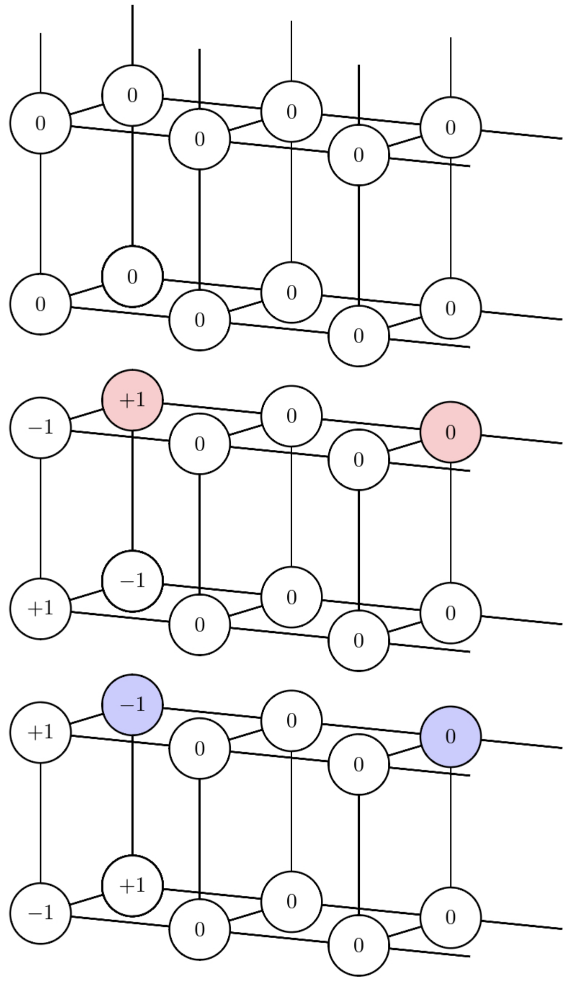


Figure 3.2: Moves of degree four

Changing the value of $k_4 = 1$ say to $k_4 = 3$ can be interpreted as permuting the entry $a_{2211} = 1$ to the entry $a_{2231} = 0$ (red), and $a_{1211} = -1$ to $a_{1231} = 0$ (blue).

The following observation will be useful in reducing the number of cases in the proof of Theorem 3.3.3.

Proposition 3.3.5. *Given $A = (a_{ijkl}) \in \mathbb{Z}^{r_1 \times 2 \times r_3 \times 2}$ with marginal conditions*

$$a_{\bullet\bullet kl} = a_{i\bullet\bullet l} = a_{ij\bullet\bullet} = a_{\bullet jk\bullet} = 0,$$

the matrices $A^{(13)} = (a_{kjil}) \in \mathbb{Z}^{r_3 \times 2 \times r_1 \times 2}$ and $A^{(24)} = (a_{ilkj}) \in \mathbb{Z}^{r_1 \times 2 \times r_3 \times 2}$ preserve the same marginal conditions.

The idea of the proof for Theorem 3.3.3 is based on the following observation. Suppose that a set of moves $\mathcal{B}$ is given. If we make N and N' as close as possible, i.e.

make $A = N - N' = (a_{ijkl})$ as small as possible by applying moves B in $\mathcal{B}$ without causing negative entries on the way, then it follows that

$$|A| = \sum_{i,j,k,l} |a_{ijkl}| \text{ can be decreased to } 0 \Longleftrightarrow \mathcal{B} \text{ is a Markov basis.}$$

This shows that we only have to consider the patterns of A, after making $|A|$ as small as possible by applying moves from $\mathcal{B}$. Suppose this is not true, i.e. one cannot apply any more moves in $\mathcal{B}$, but there are still nonzero entries in A. This means that A has at least one positive entry. Without loos of generality, we will assume that $a_{1111} > 0$. Before continuing with the argument, we establish the following restrictions on the positive entries of A.

Proposition 3.3.6. *Given* $i, i', i_1, \ldots, i_4 \in [r_1]$, $j, j', j_1, \ldots, j_4 \in [2]$, $k, k', k_1, \ldots, k_4 \in [r_3]$, *and* $l, l', l_1, \ldots, l_4 \in [2]$, *one has the following:*

(i) There are no $a_{ijkl}, a_{i'jk'l} > 0$, *with* $i \neq i', k \neq k'$.

(ii) There are no $a_{ijkl}, a_{ij'kl'} > 0$, *with* $j \neq j', l \neq l'$.

(iii) There are no $a_{i_1j_1k_1l_1}, a_{i_2j_1k_2l_2}, a_{i_3j_2k_1l_2}, a_{i_4j_2k_2l_1} > 0$, *with* $j_1 \neq j_2, k_1 \neq k_2, l_1 \neq l_2$.

(iv) There are no $a_{i_1j_1k_1l_1}, a_{i_1j_2k_2l_2}, a_{i_2j_3k_1l_2}, a_{i_2j_4k_2l_1} > 0$, *with* $i_1 \neq i_2, k_1 \neq k_2, l_1 \neq l_2$.

(v) There are no $a_{i_1j_1k_1l_1}, a_{i_1j_2k_2l_2}, a_{i_2j_1k_3l_2}, a_{i_2j_2k_4l_1} > 0$, *with* $i_1 \neq i_2, j_1 \neq j_2, l_1 \neq l_2$.

(vi) There are no $a_{i_1j_1k_1l_1}, a_{i_1j_2k_2l_2}, a_{i_2j_1k_2l_3}, a_{i_2j_2k_1l_4} > 0$, *for* $i_1 \neq i_2, j_1 \neq j_2, k_1 \neq k_2$.

Proof. We will prove that the above scenarios cannot happen by showing that there are moves in $\mathcal{B}$ that do not increase $|A|$ and do not cause negative entries.

In (i) consider the move B from the following binomial in $G(\{124, 234\}, \mathbf{r})$:

$$\begin{bmatrix} i & j & k & l \\ i' & j & k' & l \end{bmatrix} - \begin{bmatrix} i & j & k' & l \\ i' & j & k & l \end{bmatrix}.$$

Applying $-B$ to A subtracts one from the entries a_{ijkl} and $a_{i'jk'l}$, adds one to entries $a_{i'jkl}$ and $a_{ijk'l}$, while every other entry in A doesn't change. Since $a_{ijkl} - 1, a_{i'jk'l} - 1$ remain non-negative, one concludes that applying $-B$ to A doesn't cause negative entries. More, using that $|a+1| - |a| \leq 1$ for any $a \in \mathbb{Z}$, one proves thar $|A - B| \leq |A|$.

$$\begin{aligned} |A - B| = |A| &- |a_{ijkl}| - |a_{i'jk'l}| - |a_{i'jkl}| - |a_{ijk'l}| \\ &+ |a_{ijkl} - 1| + |a_{i'jk'l} - 1| + |a_{i'jkl} + 1| + |a_{ijk'l} + 1| \\ &\leq |A| - 2 + 2 = |A|. \end{aligned}$$

In (ii) one uses the same argument using the following binomial in $G(\{1243, 134\})$:

$$\begin{bmatrix} i & j & k & l \\ i & j' & k & l' \end{bmatrix} - \begin{bmatrix} i & j & k & l' \\ i & j' & k & l \end{bmatrix}.$$

In (iii) one uses the same argument using the following binomial in $Ext(Q_1 \to I_{\mathcal{M}_{24}})$:

$$\begin{bmatrix} i_1 & j_1 & k_1 & l_1 \\ i_2 & j_1 & k_2 & l_2 \\ i_3 & j_2 & k_1 & l_2 \\ i_4 & j_2 & k_2 & l_1 \end{bmatrix} - \begin{bmatrix} i_1 & j_1 & k_2 & l_1 \\ i_2 & j_1 & k_1 & l_2 \\ i_3 & j_2 & k_2 & l_2 \\ i_4 & j_2 & k_1 & l_1 \end{bmatrix}.$$

$$\begin{aligned} |A-B| = |A| &- |a_{i_1j_1k_1l_1}| - |a_{i_2j_1k_2l_2}| - |a_{i_3j_2k_1l_2}| - |a_{i_4j_2k_2l_1}| \\ &- |a_{i_1j_1k_2l_1}| - |a_{i_2j_1k_1l_2}| - |a_{i_3j_2k_2l_2}| - |a_{i_4j_2k_1l_1}| \\ &+ (a_{i_1j_1k_1l_1} - 1) + |a_{i_2j_1k_2l_2} - 1| + |a_{i_3j_2k_1l_2} - 1| + |a_{i_4j_2k_2l_1} - 1| \\ &+ |a_{i_1j_1k_2l_1} - 1| + |a_{i_2j_1k_1l_2} - 1| + |a_{i_3j_2k_2l_2} - 1| + |a_{i_4j_2k_1l_1} - 1| \\ &\leq |A| - 4 + 4 = |A| \end{aligned}$$

In (iv) one uses the same argument using the following binomial in $Ext(Q_2 \to I_{\mathcal{M}_{13}})$:

$$\begin{bmatrix} i_1 & j_1 & k_1 & l_1 \\ i_1 & j_2 & k_2 & l_2 \\ i_2 & j_3 & k_1 & l_2 \\ i_2 & j_4 & k_2 & l_1 \end{bmatrix} - \begin{bmatrix} i_1 & j_1 & k_1 & l_2 \\ i_1 & j_2 & k_2 & l_1 \\ i_2 & j_3 & k_1 & l_1 \\ i_2 & j_4 & k_2 & l_2 \end{bmatrix}.$$

In (v) one uses the same argument using the following binomial in $Ext(Q_3 \to I_{\mathcal{M}_{24}})$:

$$\begin{bmatrix} i_1 & j_1 & k_1 & l_1 \\ i_1 & j_2 & k_2 & l_2 \\ i_2 & j_1 & k_3 & l_2 \\ i_2 & j_2 & k_4 & l_1 \end{bmatrix} - \begin{bmatrix} i_2 & j_1 & k_1 & l_1 \\ i_2 & j_2 & k_2 & l_2 \\ i_1 & j_1 & k_3 & l_2 \\ i_1 & j_2 & k_4 & l_1 \end{bmatrix}.$$

In (vi) one uses the same argument using the following binomial in $Ext(Q_4 \to I_{\mathcal{M}_{13}})$:

$$\begin{bmatrix} i_1 & j_1 & k_1 & l_1 \\ i_1 & j_2 & k_2 & l_2 \\ i_2 & j_1 & k_2 & l_3 \\ i_2 & j_2 & k_1 & l_4 \end{bmatrix} - \begin{bmatrix} i_1 & j_2 & k_1 & l_1 \\ i_1 & j_1 & k_2 & l_2 \\ i_2 & j_2 & k_2 & l_3 \\ i_2 & j_1 & k_1 & l_4 \end{bmatrix}.$$

□

For the rest of the work, we will associate the steps of the proof with visualizations, where a light blue filled in circle will imply a non-positive entry, a dark blue will indicate a strictly negative entry, a light red will mean a non-negative entry, and a dark red will mean a strictly positive entry.

Lemma 3.3.7. *Given $a_{1111} > 0$, then any a_{1211} and a_{1112} must be non-positive.*

Proof. Assume that booth a_{1111} and a_{1211} are positive. Points (i) and (ii) in Proposition 3.3.6 imply that $a_{ijk1} \leq 0$ for all $1 < i \in [r_1]$, $j \in [2]$, and $1 < k \in [r_3]$. Figure 3.3 (a) is a visualization of this information.

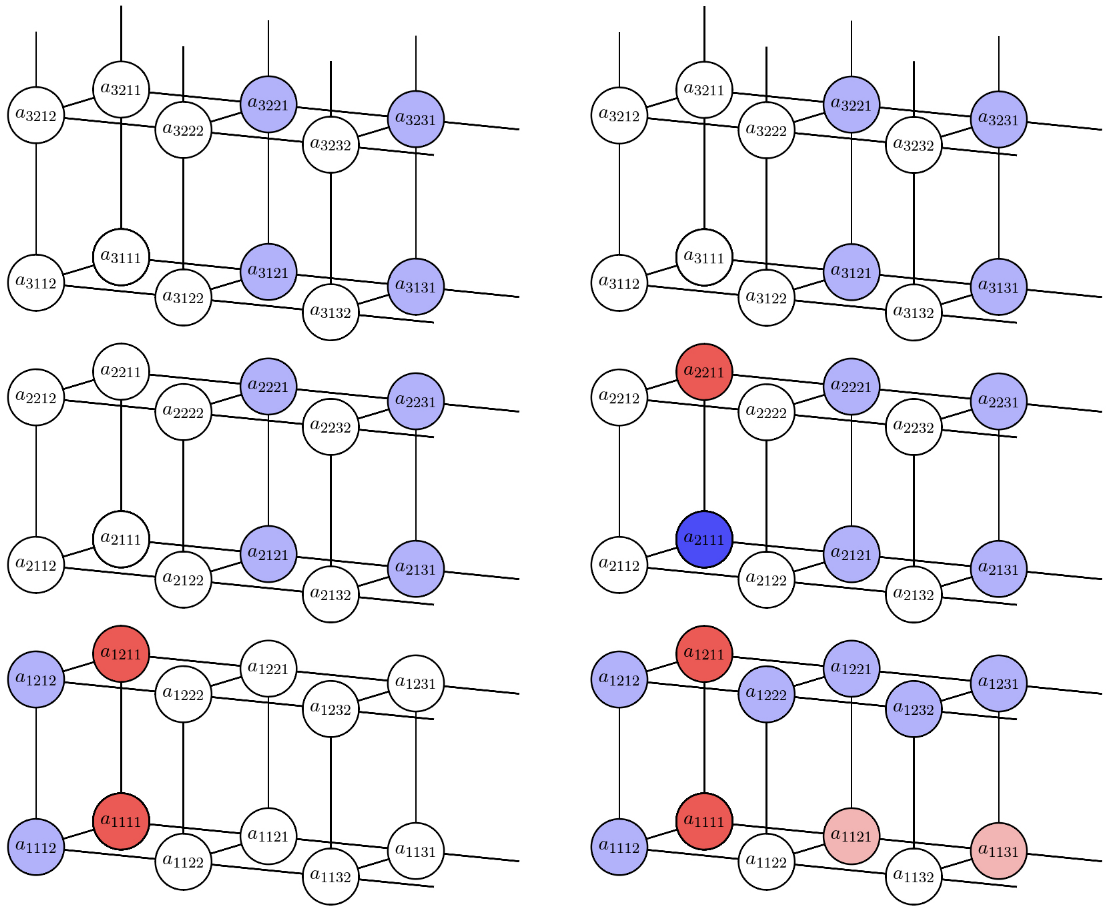


Figure 3.3: Tables (a) and (b) of Lemma 3.3.7

The marginal condition $a_{\bullet\bullet 11} = \sum_{ij} a_{ij11} = 0$ (Equation (3.3.3)), together with $a_{1111}, a_{1211} > 0$, imply that $a_{ij11} < 0$ for some $i > 1$. Since for any given $i > 1$ the slices are identical, assume this happens when $i = 2$. Assume further that $a_{2111} < 0$ since the other case $a_{2211} < 0$ frames the same situation. The marginal condition $a_{2\bullet\bullet 1} = \sum_{j,k} a_{2jk1} = 0$ (Equation (3.3.2)), $a_{2jk1} \leq 0$ for all $k > 1$, and $a_{2111} < 0$, give $a_{2211} > 0$. Proposition 3.3.6 part (i), provides $a_{12k1} \leq 0$, for all $k > 1$. Assume for a moment that under these conditions, for some $k > 1$, $a_{12k2} > 0$. Then any a_{21k2} must be non-positive from Proposition 3.3.6 part (v). The entries $a_{21k2} \leq 0$ for all $k \in [r_3]$, together with $a_{21k1} \leq 0$ for all $1 < k \in [r_3]$, and $a_{2111} < 0$, imply that $a_{21\bullet\bullet} < 0$, which is a contradiction to Equation (3.3.1). Hence, $a_{12k2} \leq 0$ for all $k \in [r_3]$. Each a_{11k1} must be non-negative since the rest of terms in the marginal sum $a_{\bullet\bullet k1} = 0$ are non-positive. This implies that $a_{11\bullet 1} > 0$. Since $a_{1\bullet\bullet 1} = a_{11\bullet 1} + a_{12\bullet 1} = 0$, one has that $a_{12\bullet 1} < 0$ and that $a_{12\bullet\bullet} = a_{12\bullet 1} + a_{12\bullet 2} < 0$, which is in contradiction to Equation (3.3.1). See Figure 3.3 (b) for a vizualization of the situation. The entry a_{1112} must be non-positive by Proposition 3.3.5. □

Next we will prove that given $a_{1111} > 0$, the entries $a_{1jk1}, a_{11kl}, a_{i11l}$,and a_{ij11}, for $i, j, k, l \neq 1$ must be non-negative, but first we need to prove the following two results.

Proposition 3.3.8. *Given $a_{1111}, a_{2211} > 0$, then a_{12k2}, a_{21k2}, a_{i212}, and a_{i112} are non-positive for all $i \in [r_1]$ and $[k \in [r_3]$.*

Proof. Cases $k = 1, i = 1$ are true by Proposition 3.3.6(i). Assume there is some $1 < k \in [r_3]$ with $a_{12k2} > 0$. Proposition 3.3.6 (i,ii) and Lemma 3.3.7 produce information on the other indices. Additionally, given $a_{1111}, a_{2211}, a_{11k2} > 0$, proposition 3.3.6 (iii) implies that $a_{21k2} \leq 0$.For $k = 2$ one has the table in Figure 3.4 (a).

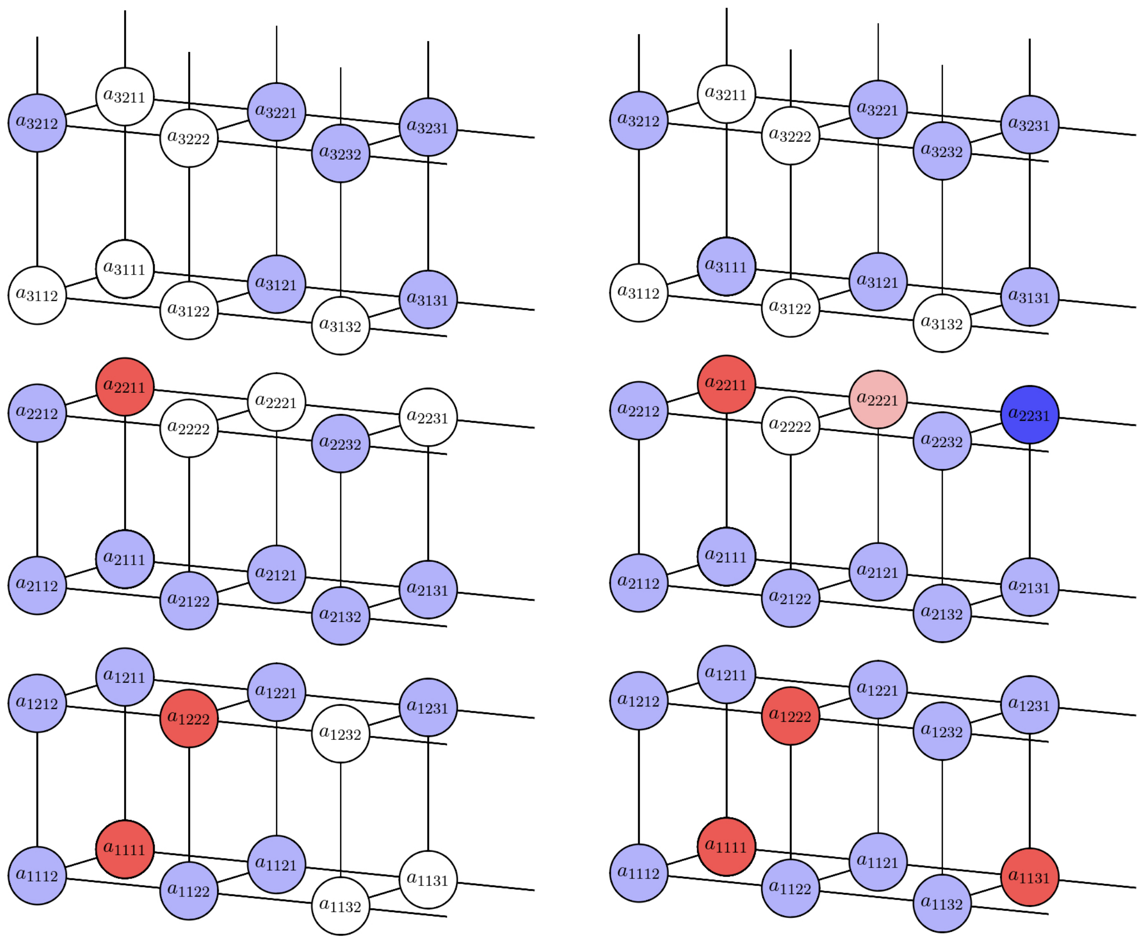


Figure 3.4: Tables (a) and (b) in Proposition 3.3.8

Each $a_{21kl} \leq 0$ and $a_{21\bullet\bullet} = 0$ imply that $a_{21kl} = 0$, for any possible k and l. The latter one and $0 = a_{2\bullet\bullet 1} = a_{21\bullet 1} + a_{22\bullet 1} = 0$ imply that $a_{22\bullet 1} = 0$. Since $a_{2211} > 0$, there must be some $k' > 1$ in $[r_1]$ with $a_{22k'1} < 0$. The index k' should be different from k used in $a_{11k2} > 0$, since otherwise for this k one obtains $a_{\bullet\bullet k1} < 0$, which is not possible. Hence $k' \neq k$. The marginal sum $a_{\bullet\bullet k'1} = 0$ implies that $a_{11k'1}$ must be positive. Figure 3.4 (b) visualizes such a table for $k = 2$ and $k' = 3$. Under these conditions, $a_{\bullet 2k'\bullet} < 0$, which concludes that the assumption $a_{12k2} > 0$ is incorrect. The proof is analogous for a_{21k2}. $\square$

Proposition 3.3.9. *Given $a_{1111}, a_{2211} > 0$, then all $a_{11k1}, a_{22k1}, 1 < k \in [r_3]$ are negative or zero.*

Proof. Assume $a_{1111}, a_{2211}, a_{11k1} > 0$, for some $k > 1$. When $k = 2$, using Proposition 3.3.6 (i,ii), Lemma 3.3.7, one has the situation in Figure 3.5 (a).

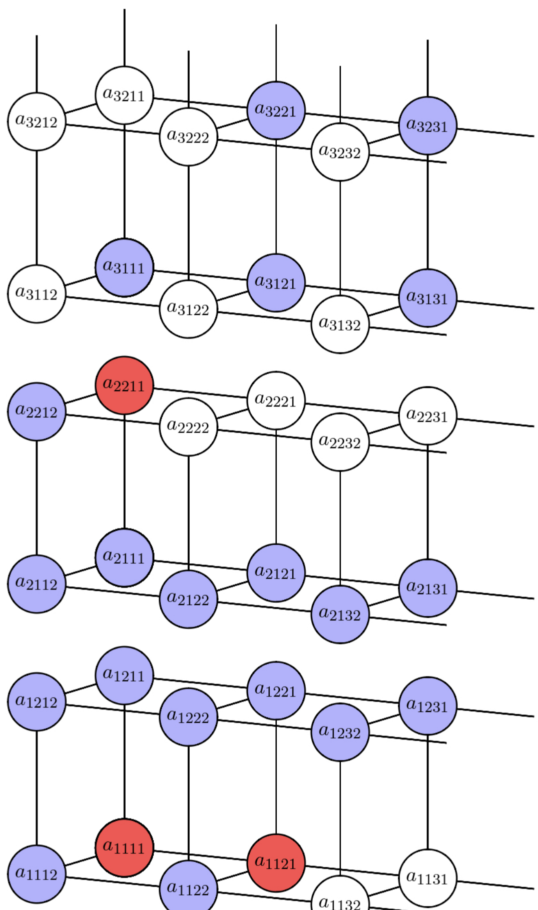


Figure 3.5: Tables (a) and (b) in Proposition 3.3.9

Immediate consequences are $a_{12kl} = a_{21kl} = 0$ for all k, l. The marginal condition $a_{ij\bullet\bullet} = 0$, and $a_{i111} \leq 0$, $a_{1111}, a_{2211} > 0, a_{1211}, a_{2111} = 0$, imply that there is some $i > 2$ with $a_{i211} < 0$. For this particular i, one has that the marginal sum $a_{i\bullet\bullet 1}$ is negative, which proves that the assumption $a_{11k1} > 0$ is incorrect. Figure below visualizes when $i = 3$. Proposition 3.3.5 and symmetric group actions S_{r_1} applied to the first index, and S_{r_3} applied to the third index, solve the other cases of this proposition.

□

Lemma 3.3.10. *If $a_{1111} > 0$, then any $a_{12k1}, a_{11kl}, a_{i112}$, and a_{i211}, for $1 \neq i \in [r_1]$ and $1 \neq k \in [r_2]$ must be non-positive.*

Proof. We will prove that a_{2211} must be non-positive. The rest follows by the symmetric group action and Proposition 3.3.5. Assume that both a_{1111} and a_{2211} are positive. Using Lemmas and Propositions above, one has the situation as in Figure 3.6.

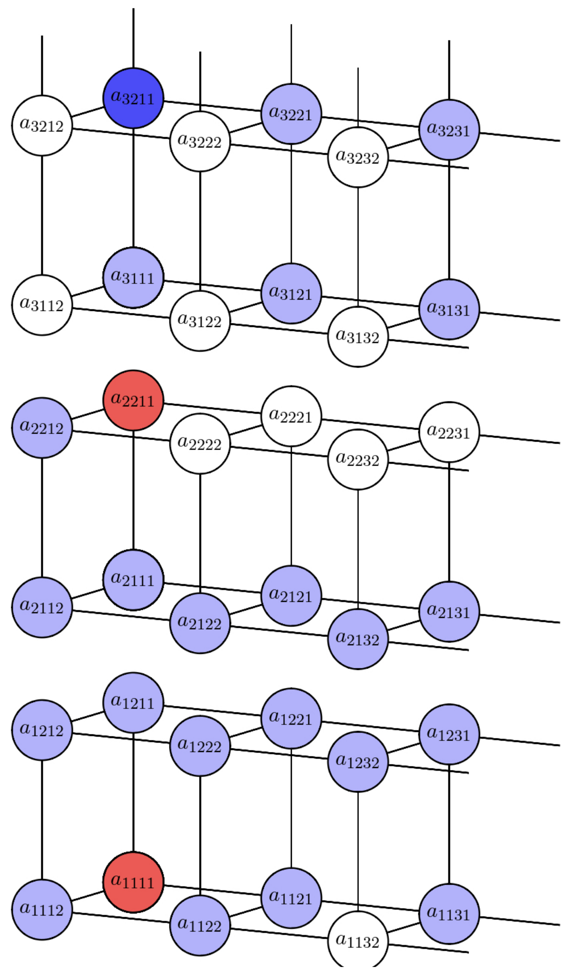


Figure 3.6: The table for Lemma 3.3.10

We will first prove that under these circumstances no a_{12k2} can be positive. Indeed, notice that $a_{12kl} \leq 0$ for all k, l and $a_{12\bullet\bullet} = 0$, which implies that all a_{12kl} are zero. Since $0 = a_{1\bullet\bullet 1} = a_{11\bullet 1} + a_{12\bullet 1}$, one has $a_{12\bullet 1} = 0$. Given $a_{1111} > 0$, for some $k > 1$, one of a_{11k1} must be negative. This means that the marginal sum $a_{ij\bullet\bullet}$ is negative, which concludes that the assumption $a_{2211} > 0$ is incorrect. □

Lemma 3.3.11. *If $a_{1111} > 0$, then a_{12k2}, a_{i212} must be non-negative for all $i \in [r_1]$ and $k \in [r_3]$.*

Proof. From Proposition 3.3.5, it is enough to prove one of the cases. Assume for contradiction that $a_{1222} > 0$, and consider the consequences of the two positive entries obtained by Proposition 3.3.6 and lemmas 3.3.7 and 3.3.10 visualized in Figure 3.7(a).

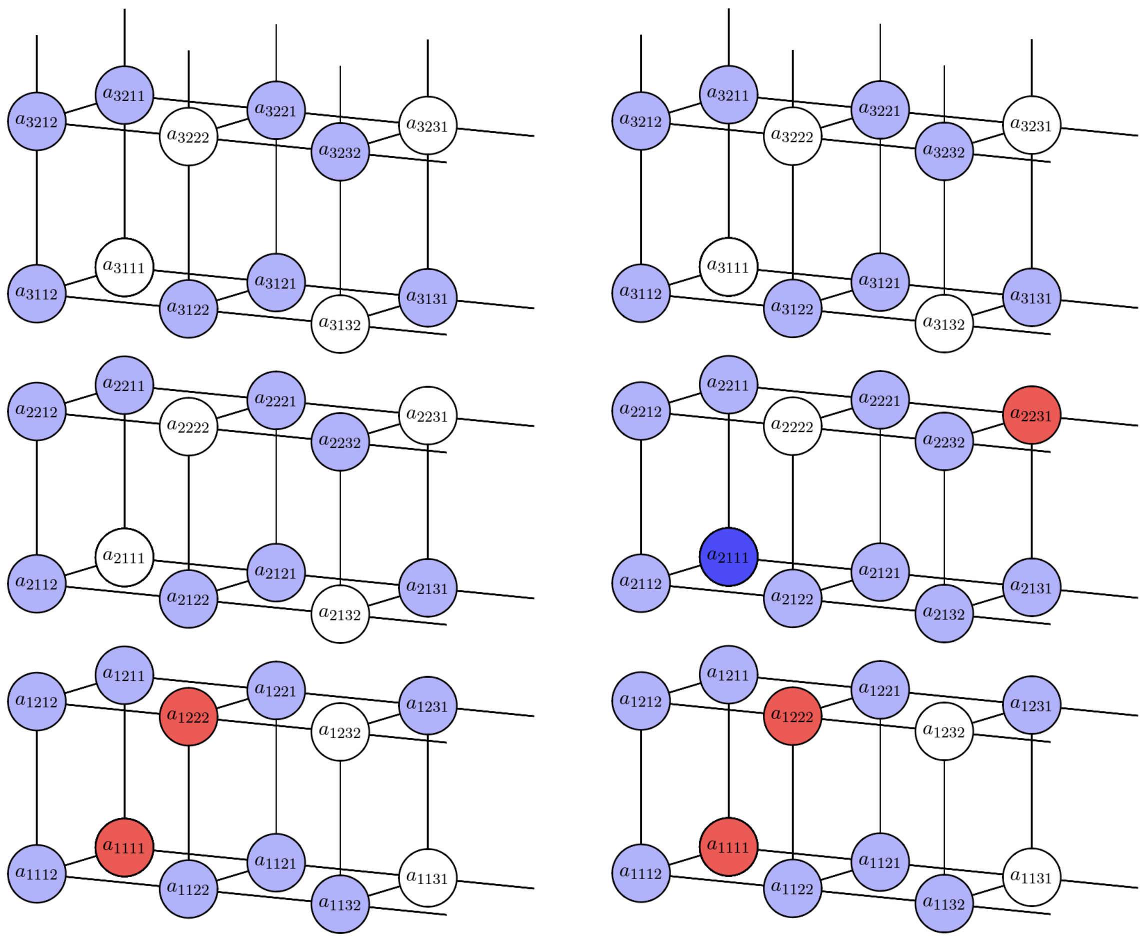


Figure 3.7: Tables (a) and (b) for Lemma 3.3.11

The entries $a_{i21l} \leq 0$ for all i, l, and $a_{\bullet 21 \bullet} = 0$, and so each $a_{i21l} = 0$, and $a_{\bullet 211} = 0$. The later one, together with and $a_{\bullet\bullet 11} = a_{\bullet 111} + a_{\bullet 211} = 0$ imply that $a_{\bullet 111} = 0$. Since $a_{1111} > 0$, A must have a negative entry a_{i111}, for some $i > 1$. Case $i = 2$ is visualized below. This negative entry and $a_{2\bullet\bullet 1} = 0$ imply that there exists some $k > 2$ with $a_{i2k1} > 0$. The new positive entry induces conditions described in Proposition 3.3.6(i,ii,iii) and Lemma 3.3.7. When $i = 2$ and $k = 3$ one has the table in Figure 3.7 (b). Observe that in this situation $a_{21\bullet\bullet}$ is positive with is in contradiction to *eq.* (3.3.1). Hence, the assumption that a_{1222} is positive is incorrect. □

Proof of Theorem 3.3.3. The entry $a_{1111} > 0$ induces all conditions in proposition 3.3.6, lemmas 3.3.7, 3.3.10 and 3.3.11. Figure 3.8 (a) provides a visualization of these conditions.

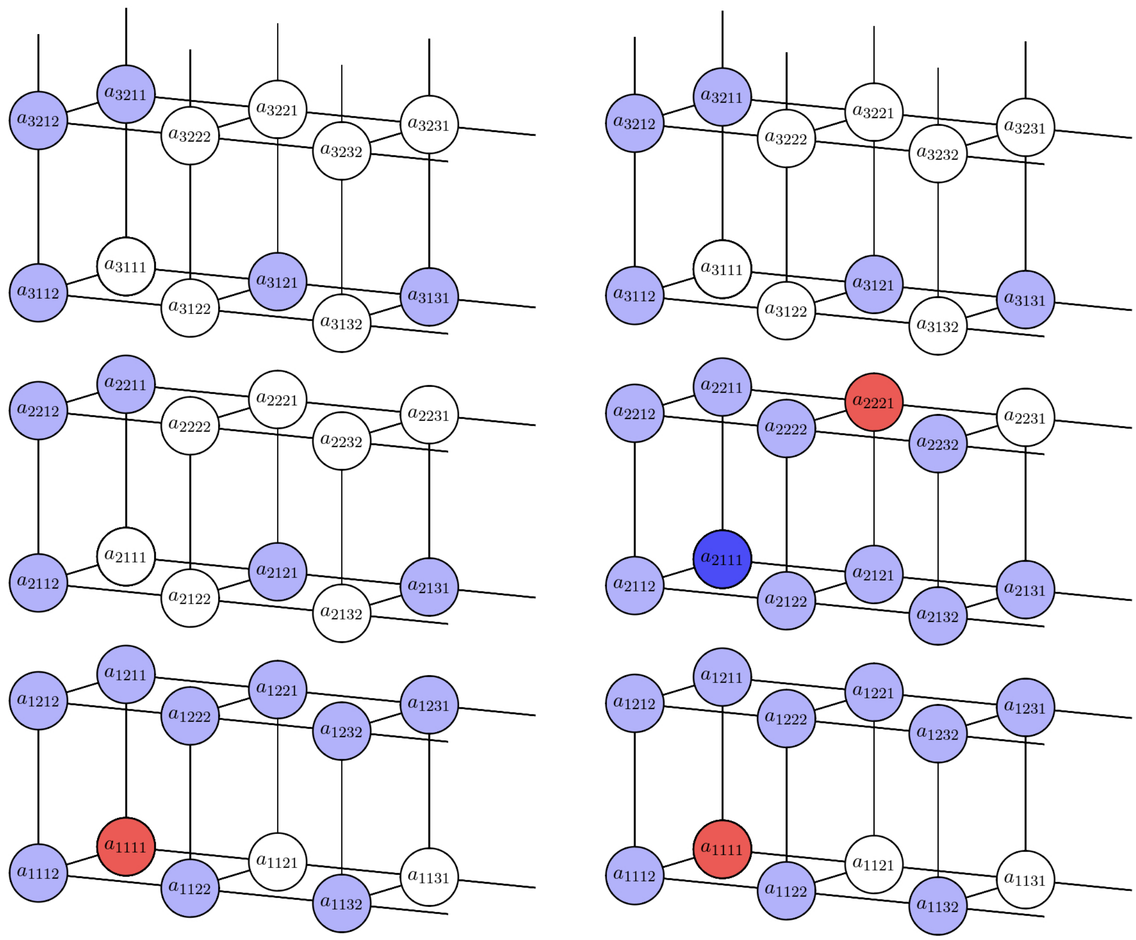


Figure 3.8: Tables (a) and (b) for Theorem 3.3.3

Since all $a_{i21l} = 0$, $a_{1111} > 0$, and $a_{\bullet\bullet 11} = 0$, one must have that $a_{i111} < 0$ for some $i > 1$. As result, for this particular i, there exist a $k > 1$ such that $a_{i2k1} > 0$. Case $i = 2$, $k = 2$ is visualized in Figure 3.8. Under these conditions, one has $a_{21\bullet\bullet} < 0$ which contradicts eq. (3.3.1). Hence, the assumption that a_{1111} is positive is incorrect, and A must be the zero table. □

The filtration of ideals $(I_{r_1,r_3})_{r_1,r_3 \in \mathbb{N}}$, where $I_{r_1,r_3} = I_{\mathcal{M}(C_4,(r_1,2,r_3,2))}$ with respect to the group action

$$S_{r_1} \times S_{r_3} \curvearrowright R_{r_1,r_3} = \mathbb{K}[x_{ijkl} \mid i \in [r_1], j \in [2], k \in [r_3], l \in [2]],$$
$$(\sigma, \tau) \cdot x_{ijkl} \to x_{\sigma(i)j\tau(k)l}.$$

Observe that binomials in Theorem 3.3.3 are all of the form $(\sigma, \tau) \cdot f$, with $\sigma \in S_{r_1}, \tau \in S_{r_3}$, and $f \in I_{4,4}$, which induce the next result.

Corollary 3.3.12. *The invariant filtration* $(I_{r_1,r_3})_{r_1,r_3 \in \mathbb{N}}$ *with* $I_{r_1,r_3} = I_{\mathcal{M}(C_4,(r_1,2,r_3,2))}$ *stabilizes for* $r_1 = r_3 = 4$.

In general, given the hierarchical model $\mathcal{M}(C_m, \mathbf{r})$, denote $\mathcal{M}_{1,m-1} = \mathcal{M}(C_m \cup \{1, m-1\}, \mathbf{r})$ and $\mathcal{M}_{2,m} = \mathcal{M}(C_m \cup \{2, m\}, \mathbf{r})$. Let $G_{1,m-1}$ and $G_{2,m}$ be generating sets for $\mathcal{M}_{1,m-1}$ and $\mathcal{M}_{2,m}$, respectively. Let $G^{\leq d}$ be the set of binomials of degree at most d in G. With the support of examples computed in software $4ti2$, we suspect the following.

Conjecture 3.3.13. *Given the filtration* $\mathscr{I}_{C_m,c} = (I_{\mathbf{r}_T})_{\mathbf{r}_T \in \mathbb{N}^{\#T}}$, *with* $\#(T \cap F) \leq 1$ *for all* $F \in C_m$, *there exists* $\mathbf{r}_T^0 \in \mathbb{N}^{\#T}$ *and a positive integer* d *such that the set of binomials* $G_{1,m-1}^d$ *and* $G_{2,m}^d$ *arising from the model* $\mathcal{M}(C_m, (\mathbf{c}, \mathbf{r}_T^0))$, *is a generating set for the filtration* $\mathscr{I}_{C_m,c}$.

If the conjecture is true, the problem of generating sets for any hierarchical model with the structure of a cycle reduces to finding generating sets of the three-cycle model. The more we know about three-cycle models, the closer we are in proving the conjecture. In particular, using Markov bases for three-cycle models with one binary state in [16, Section 4], we expect that that the filtration for the four-cycle and $r = (r_1, c, r_3, 2)$, where $r_1, r_3 \in \mathbb{N}$ and c is some fixed positive integer, has $r_1^0 = r_3^0 = c$ and $d = 2c$.

Chapter 4 Equivariant Hilbert Series of Hierarchical Models

This chapter concerns the asymptotic quantitative behaviours of ideals of hierarchical models. Toric ideals to hierarchical models are invariant under the action of a product of symmetric groups. Taking the number of factors, say m, into account, we introduce and study invariant filtrations and their equivariant Hilbert series. In Theorem 4.2.1 we present a condition that guarantees the equivariant Hilbert series is a rational function in $m+1$ variables with rational coefficients. Furthermore, in Section 4.4, we give explicit formulas for the rational functions with coefficients in a number field and an algorithm for determining the rational functions with rational coefficients. The key is to construct finite automata that recognize languages corresponding to invariant filtrations.

4.1 Preliminaries

In this chapter we restrict ourselves to considering ideals of hierarchical models $\mathcal{M}(\Delta, \mathbf{r})$. As pointed out in Remark 2.4.2, for any subset $T \neq \emptyset$ of $[m]$, these ideals give rise to $S_{\infty}^{\#T}$-invariant filtrations. To study their equivariant Hilbert series, it is convenient to simplify notation. We may assume that $T = \{m - \#T + 1, \dots, m\}$ and fix the entries of $\mathbf{r}$ in positions supported on $[m] \setminus T$, that is, we fix $\mathbf{c} \in \mathbb{N}^{m-\#T}$ and set $\mathbf{n} = (n_1, \dots, n_{\#T})$ to obtain $\mathbf{r} = (\mathbf{c}, \mathbf{n})$. We write $\mathcal{M}_{\mathbf{n}}$ instead of $\mathcal{M}(\Delta, (\mathbf{c}, \mathbf{n}))$ and denote the resulting $S_{\infty}^{\#T}$-invariant filtration $(I_{\mathcal{M}_{\mathbf{n}}})_{\mathbf{n} \in \mathbb{N}^{\#T}}$ by $\mathscr{I}_{\Delta,\mathbf{c}}$, as in the Independent Set theorem. Its equivariant Hilbert series is

$$equivH_{\mathscr{I}_{\Delta,\mathbf{c}}}(s_1, s_2, \dots, s_{\#T}, t) = \sum_{\mathbf{n} \in \mathbb{N}^{\#T}} H_{R_{\mathbf{c}}/I_{\mathcal{M}_{\mathbf{n}}}}(t) \cdot s_1^{n_1} \cdots s_{\#T}^{n_{\#T}}.$$

The Independent Set theorem (Theorem 2.4.3) guarantees stabilization of the filtration. This suggests the following problem.

Question 4.1.1. *If $T \subseteq [m]$ satisfies $\#(F \cap T) \leq 1$ for every facet F of Δ, is then the equivariant Hilbert series of $\mathscr{I}_{\Delta,\mathbf{c}}$ rational?*

The answer is affirmative if T consists of exactly one element.

Proposition 4.1.2. *If $\#T = 1$, then the equivariant Hilbert series of $\mathscr{I}_{\Delta,\mathbf{c}}$ is rational.*

Proof. The assumption means $T = \{m\}$ and $\mathbf{r} = (\mathbf{c}, n)$ with $\mathbf{c} \in \mathbb{N}^{m-1}$ and $n \in \mathbb{N}$. Set $c = c_1 \cdots c_{m-1}$ and fix a bijection

$$\psi \colon [\mathbf{c}] = [c_1] \times \cdots \times [c_{m-1}] \to [c].$$

For every $n \in \mathbb{N}$, it induces a ring isomorphism

$$\begin{aligned} R_n = \mathbb{K}[x_{\mathbf{i},j} \mid (\mathbf{i}, j) \in [\mathbf{c}] \times [n]] &\longrightarrow \mathbb{K}[x_{i,j} \mid (i,j) \in [c] \times [n]] = R'_n \\ x_{\mathbf{i},j} &\mapsto x_{\psi(\mathbf{i}),j}. \end{aligned}$$

This isomorphism maps every ideal $I_{\mathcal{M}_n}$ corresponding to the model $\mathcal{M}(\Delta, (\mathbf{c}, n))$ onto an S_n-invariant ideal I_n. In particular, the rings $R_n/I_{\mathcal{M}_n}$ and R'_n/I_n have the same Hilbert series and the family $(I_n)_{n\in\mathbb{N}}$ is an S_∞-invariant filtration. Thus, its equivariant Hilbert series is rational by [32, Theorem 7.8] or [28, Theorem 4.3]. $\square$

4.2 The Generalized Independence Hierarchical Models

Our main result in this section describes further cases in which the answer to Question 4.1.1 is affirmative.

Theorem 4.2.1. *The equivariant Hilbert series of $\mathscr{I}_{\Delta,\mathbf{c}}$ is a rational function with rational coefficients if*

1. $F_i \cap F_j = \emptyset$ *for any distinct* $F_i, F_j \in$ *facet*(Δ).
2. $|F \cap T| \leq 1$ *for any* $F \in$ *facet*(Δ).

This results applies in particular to the independence model, where it takes an attractive form.

Example 4.2.2. *A hierarchical model describing m independent parameters is called* independence model*. Its collection of facets is facet$(\Delta) = \{\{1\}, \{2\}, \ldots, \{m\}\}$. Thus, we may apply Theorem 4.2.1 with any subset T of $[m]$. Using $T = [m]$, we show in Example 4.4.5 below that*

$$\begin{aligned} equivH_{\mathscr{I}_\Delta}(s_1, s_2, \ldots, s_m, t) &= \sum_{\mathbf{n}\in\mathbb{N}^m} H_{R_\mathbf{n}/I_{\mathcal{M}_\mathbf{n}}}(t) \cdot s_1^{n_1} \cdots s_m^{n_m} \\ &= \frac{s_1 \cdots s_m}{(1-s_1)\cdots(1-s_m) - t}. \end{aligned}$$

The proof of Theorem 4.2.1 will be given in two steps. First we show that it is enough to prove the result in a special case where every facet consists of two elements. Second, we use regular languages to show the desired rationality in the following section. In the remainder of this section we establish the reduction step.

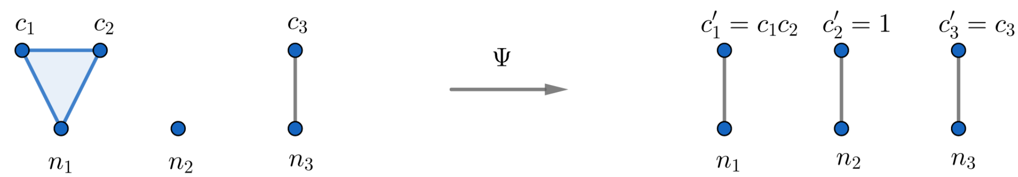
Lemma 4.2.3. *Consider a collection $\Delta = \{F_1, \ldots, F_q\}$ on vertex set $[m]$ and a subset T of $[m]$ satisfying*

1. $F_i \cap F_j = \emptyset$ *for any* $F_i, F_j \in$ *facet*(Δ).
2. $|F \cap T| = 1$ *for any* $F \in$ *facet*(Δ).

Then there is a collection facet$(\Delta') = \{F'_1, \ldots, F'_q\}$ on vertex set $[2q]$ consisting of two element facets and also satisfying conditions (1) and (2) with the property that, for every $\mathbf{c} \in \mathbb{N}^{m-\#T}$ there is some $\mathbf{c}' \in \mathbb{N}^{2q-\#T}$ such that the filtrations corresponding to the models $\mathcal{M}(\Delta, (\mathbf{c}, \mathbf{n}))$ and $\mathcal{M}(\Delta', (\mathbf{c}', \mathbf{n}))$ with $\mathbf{n} \in \mathbb{N}^{\#T}$ have the same equivariant Hilbert series.

Proof. The assumptions imply that T must have q elements. We may assume that every facet in Δ has at least two elements. Indeed, if $F \in \text{facet}(\Delta)$ has only one element then we may replace F by the union F' of F and a new vertex. Assigning to the parameter corresponding to the new vertex exactly one possible state gives a new model whose coordinate ring has the same Hilbert series as the original model. Given such a hierarchical model $\mathcal{M}_{\mathbf{n}} = \mathcal{M}(\Delta, (\mathbf{c}, \mathbf{n}))$ on vertex set $[m]$, we will construct a new hierarchical model $\mathcal{M}'_{\mathbf{n}} = \mathcal{M}(\Delta', (\mathbf{c}', \mathbf{n}))$ on $2q$ vertices that has the same Hilbert series. The new vertex set is the disjoint union of the q vertices in $F_j \cap T$ with $j \in [q]$ and a set V of q other vertices, say $V = [q]$. For $j \in [q]$, set $F'_j = \{j\} \cup (F_j \cap T)$. Thus, the sets F'_j are pairwise disjoint because $F_1, \ldots, F_q$ have this property, and each F'_j has exactly two elements. In particular, $\text{facet}(\Delta') = \{F'_1, \ldots, F'_q\}$ and T satisfy conditions (1) and (2).

Now let $c'_j = \prod_{e \in F_j \setminus T} c_e = \#[\mathbf{c}_{F_j \setminus T}]$ be the number of states of the parameter corresponding to vertex $j \in F'_j$. Furthermore, for every $j \in [q]$, let the parameter corresponding to vertex $F'_j \cap T$ have the same number of states as $F_j \cap T$ has in $\mathcal{M}_{\mathbf{n}}$. This completes the definition of a new hierarchical model $\mathcal{M}'_{\mathbf{n}} = \mathcal{M}(\Delta', (\mathbf{c}', \mathbf{n}))$. The passage form $\mathcal{M}_{\mathbf{n}}$ to $\mathcal{M}'_{\mathbf{n}}$ is illustrated in an example below.



$$\mathcal{M}(\{124, 5, 36\}, (c_1, c_2, c_3, n_1, n_2, n_3)) \longrightarrow \mathcal{M}(\{14, 25, 36\}, (c'_1, 1, c'_3, n_1, n_2, n_3))$$

Figure 4.1: Reduced facets

Varying $\mathbf{n} \in \mathbb{N}^q$, the ideals $I_{\mathcal{M}'_{\mathbf{n}}}$ form an S^q_∞-invariant filtration. Thus, to establish the assertion it is enough to prove that for every $\mathbf{n} \in \mathbb{N}^q$, the quotient rings $R_{\mathbf{n}}/I_{\mathcal{M}_{\mathbf{n}}}$ and $R'_{\mathbf{n}}/I_{\mathcal{M}'_{\mathbf{n}}}$ are isomorphic.

For every $F_j \in \text{facet}(\Delta)$, the sets $[\mathbf{c}_{F_j \setminus T}]$ and $[c'_j]$ have the same finite cardinality. Choose a bijection

$$\psi_j : [\mathbf{c}_{F_j \setminus T}] \longrightarrow [c'_j].$$

These choices determine two further bijections:

$$(\psi_1, \ldots, \psi_q, \text{id}_{[\mathbf{n}]}) \colon [\mathbf{c}_{F_1 \setminus T}] \times \cdots \times [\mathbf{c}_{F_q \setminus T}] \times [\mathbf{n}] \longrightarrow [c'_1] \times \cdots \times [c'_q] \times [\mathbf{n}] \tag{4.2.1}$$

and

$$(\psi_j, \text{id}_{[n_j]}) \colon [\mathbf{c}_{F_j \setminus T}] \times [n_j] \longrightarrow [c'_j] \times [n_j]. \tag{4.2.2}$$

Bijection (4.2.1) induces the following isomorphism of polynomial rings

$$\begin{aligned}\Psi\colon R_{\mathbf{n}} = \mathbb{K}[x_{\mathbf{i}_{F_1\setminus T},\ldots,\mathbf{i}_{F_q\setminus T},\mathbf{k}} \mid \mathbf{i}_{F_q\setminus T} \in [\mathbf{c}_{F_q\setminus T}], \mathbf{k} \in [\mathbf{n}]] \\ \longrightarrow \mathbb{K}[x_{i_1,\ldots,i_q,\mathbf{k}} \mid i_j \in [c'_j], \mathbf{k} \in [\mathbf{n}]] = R'_{\mathbf{n}} \\ x_{\mathbf{i}_{F_1\setminus T},\ldots,\mathbf{i}_{F_q\setminus T},\mathbf{k}} \mapsto x_{\psi_1(\mathbf{i}_{F_1\setminus T}),\ldots,\psi_q(\mathbf{i}_{F_q\setminus T}),\mathbf{k}}.\end{aligned}$$

Similarly, Bijection (4.2.2) induces an isomorphism of polynomial rings

$$\begin{aligned}\Psi'\colon S_{\mathbf{n}} = \mathbb{K}[y_{j,\mathbf{i}_{F_j\setminus T},k_j} \mid 1 \le j \le q, \mathbf{i}_{F_j\setminus T} \in [\mathbf{c}_{F_j\setminus T}], k_j \in [n_j]] \\ \longrightarrow \mathbb{K}[y_{j,i_j,k_j} \mid 1 \le j \le q, i_j \in [c_j], k_j \in [n_j]] = S'_{\mathbf{n}} \\ y_{j,\mathbf{i}_{F_j\setminus T},k_j} \mapsto y_{j,\psi_j(\mathbf{i}_{F_j\setminus T}),k_j}.\end{aligned}$$

We claim that the following diagram is commutative:

$$\begin{array}{ccc} R_{\mathbf{n}} & \xrightarrow{\Phi_{\mathcal{M}}} & S_{\mathbf{n}} \\ \Psi\downarrow & & \downarrow\Psi' \\ R'_{\mathbf{n}} & \xrightarrow{\Phi_{\mathcal{M}'}} & S'_{\mathbf{n}} \end{array} \tag{4.2.3}$$

Indeed, it suffices to check this for variables variables. In this case commutative is shown by the diagram:

$$\begin{array}{ccc} x_{\mathbf{i}_{F_1\setminus T},\ldots,\mathbf{i}_{F_q\setminus T},\mathbf{k}} & \xmapsto{\Phi_{\mathcal{M}}} & \prod_{j=1}^{q} y_{j,\mathbf{i}_{F_j\setminus T},k_j} \\ \Psi\downarrow & & \downarrow\Psi' \\ x_{\psi_1(\mathbf{i}_{F_1\setminus T}),\ldots,\psi_q(\mathbf{i}_{F_q\setminus T}),\mathbf{k}} & \xmapsto{\Phi_{\mathcal{M}'}} & \prod_{j=1}^{q} y_{j,\psi_1(\mathbf{i}_{F_q\setminus T}),k_j} \end{array}$$

Since Ψ and Ψ' are isomorphisms commutativity of Diagram (4.2.3) implies that $\operatorname{im}(\Phi) \cong \operatorname{im}(\Phi')$, which concludes the proof. □

We also need the following result.

Proposition 4.2.4. *Let $\mathscr{I} = \{I_{\mathbf{n}}\}_{\mathbf{n}\in\mathbb{N}^q}$ be the S^q_∞-invariant filtration corresponding to hierarchical models $\mathcal{M}(\Delta, (\mathbf{c}, \mathbf{n}))$ with facet(Δ) consisting of q 2-element disjoint facets $F_1, \ldots, F_q$, each meeting T in exactly one vertex. Then the equivariant Hilbert series of $\mathscr{I}$ is a rational function in $s_1, \ldots, s_q, t$ with rational coefficients.*

This will be shown in the following section. Assuming the result, we complete the argument for establishing Theorem 4.2.1.

Proof of Theorem 4.2.1. Let ν be the number of facets in Δ whose intersection with T is empty. We use induction on $\nu \ge 0$. If $\nu = 0$ the claimed rationality follows by combining Lemma 4.2.3 and Proposition 4.2.4.

Let $\nu \geq 1$. We may assume that $F_1 \cap T = \emptyset$ and that vertex 1 is in F_1. By assumption, it has c_1 states. Set $\tilde{\mathbf{n}} = (n_1, \mathbf{n})$, $\tilde{\mathbf{c}} = (c_2, \ldots, c_{\#T})$ and $\tilde{T} = T \cup \{1\}$. Then the hierarchical models $\tilde{\mathcal{M}}(\Delta, (\tilde{\mathbf{c}}, \tilde{\mathbf{n}}))$ give rise to a filtration $\tilde{\mathscr{I}} = \mathscr{I}_{\Delta, \tilde{\mathbf{c}}}$. By induction on ν, it has a rational equivariant Hilbert series. Since $equivH_{\mathscr{I}}$ is obtained by evaluating $\dfrac{1}{c_1!} \dfrac{\partial^{c_1} equivH_{\tilde{\mathscr{I}}}}{\partial s_1^{c_1}}$ at $s_1 = 0$, it follows that also $equivH_{\mathscr{I}}$ is rational. $\square$

4.3 Regular Languages

The goal of this section is to establish Proposition 4.2.4. We adopt its notation.

Fix $\mathbf{c} \in \mathbb{N}^q$. As above, we write $x_{\mathbf{i},\mathbf{k}}$, where $(\mathbf{i}, \mathbf{k}) = (i_1, \ldots, i_q, k_1, \ldots, k_q) \in [\mathbf{c}] \times [\mathbf{n}] \subseteq \mathbb{N}^{2q}$. Thus, $y_{j, \mathbf{i}_{F_j}, \mathbf{k}_{F_j}}$ is simply y_{j,i_j,k_j}. For any $\mathbf{n} \in \mathbb{N}^q$, the homomorphism associated to the model $\mathcal{M}_{\mathbf{n}} = \mathcal{M}(\Delta, (\mathbf{c}, \mathbf{n}))$ is

$$\begin{aligned} \Phi_{\mathbf{n}} \colon R_{\mathbf{n}} = \mathbb{K}[x_{\mathbf{i},\mathbf{k}} \mid (\mathbf{i}, \mathbf{k}) \in [\mathbf{c}] \times [\mathbf{n}]] &\to \mathbb{K}[y_{j,i_j,k_j} \mid j \in [q], i_j \in c_j, k_j \in [n_j]] = S_{\mathbf{n}} \\ x_{\mathbf{i},\mathbf{k}} &\longmapsto \prod_{j=1}^{q} y_{j,i_j,k_j}. \end{aligned}$$

Set

$$A_{\mathbf{n}} = \operatorname{im} \Phi_{\mathbf{n}} = \mathbb{K}\left[\prod_{j=1}^{q} y_{j,i_j,k_j} \mid i_j \in c_j, k_j \in [n_j]\right].$$

We denote the set of monomials in $A_{\mathbf{n}}$ and $S_{\mathbf{n}}$ by $\operatorname{Mon}(A_{\mathbf{n}})$ and $\operatorname{Mon}(S_{\mathbf{n}})$, respectively. Define $\operatorname{Mon}(A)$ as the disjoint union of the sets $\operatorname{Mon}(A_{\mathbf{n}})$ with $\mathbf{n} \in \mathbb{N}^q$ and similarly $\operatorname{Mon}(S)$. Our next goal is to show that the elements of $\operatorname{Mon}(A)$ are in bijection to the words of a suitable formal language.

Consider a set

$$\Sigma = \{\zeta_{\mathbf{i}}, \tau_j \mid \mathbf{i} \in [\mathbf{c}], j \in [q]\}$$

with $q + \prod_{j=1}^{q} c_j$ elements. Let Σ^* be the free monoid on Σ. In order to compare subsets of Σ^* with $\operatorname{Mon}(A)$ we need suitable maps. For $j \in [q]$, define a shift operator $T_j \colon \operatorname{Mon}(S) \to \operatorname{Mon}(S)$ by

$$T_j(y_{l,i,k}) = \begin{cases} y_{l,i,k+1} & \text{if } l = j; \\ y_{l,i,k} & \text{if } l \neq j, \end{cases}$$

extended multiplicatively to $\operatorname{Mon}(S)$. Define a map $\mathbf{m} \colon \Sigma^* \to \operatorname{Mon}(S)$ inductively using the three rules

$$\text{(a) } \mathbf{m}(\epsilon) = 1, \quad \text{(b) } \mathbf{m}(\zeta_{\mathbf{i}} w) = \prod_{j=1}^{q} y_{j,i_j,1} \mathbf{m}(w), \quad \text{(c) } \mathbf{m}(\tau_j w) = T_j(\mathbf{m}(w)),$$

where $w \in \Sigma^*$. In particular, this gives $\mathbf{m}(\zeta_{\mathbf{i}}) = \Phi_{\mathbf{n}}(x_{\mathbf{i},\mathbf{1}})$ for any $\mathbf{n} \in \mathbb{N}^q$, where $\mathbf{1}$ is the q-tuple whose entries are all equal to 1.

Example 4.3.1. *If $c_1 = c_2 = q = 2$, one has $\Sigma = \{\zeta_{1,1}, \zeta_{1,2}, \zeta_{2,1}, \zeta_{2,2}, \tau_1, \tau_2\}$, and, for any $\mathbf{n} \geq (2,3)$,*

$$\begin{aligned}\mathbf{m}(\tau_1\tau_2\zeta_{1,2}\tau_2\zeta_{1,1}\tau_1) &= T_1(T_2(y_{1,1,1}y_{2,2,1}T_2(y_{1,1,1}y_{2,1,1}T_1(1)))) \\ &= T_1(T_2(y_{1,1,1}y_{2,2,1}y_{1,1,1}y_{2,1,2})) \\ &= y_{1,1,2}y_{2,2,2}y_{1,1,2}y_{2,1,3} \\ &= \Phi_{\mathbf{n}}(x_{(1,2),(2,2)})\Phi_{\mathbf{n}}(x_{(1,1),(2,3)}).\end{aligned}$$

The map $\mathbf{m}$ is certainly not injective because the variables $y_{j,i,k}$ commute. For example, if $q = 2$ one has $\mathbf{m}(\tau_1\tau_2) = \mathbf{m}(\tau_2\tau_1)$ and $\mathbf{m}(\zeta_{2,1}\zeta_{1,2}) = \mathbf{m}(\zeta_{1,2}\zeta_{2,1}) = \mathbf{m}(\zeta_{1,1}\zeta_{2,2})$ and $\mathbf{m}(\tau_1\zeta_{1,2}\tau_2\zeta_{2,1}) = \mathbf{m}(\tau_1\zeta_{2,2}\tau_2\zeta_{1,1})$. Thus, we introduce a suitable subset of Σ^*.

Definition 4.3.2. *Let $\mathcal{L}$ be the set of words in Σ^* that satisfy the following conditions:*

1. *Every substring $\tau_i\tau_j$ has $i \leq j$.*
2. *In every substring with no τ_j, if $\zeta_{\mathbf{i}}$ occurs to the left of some $\zeta_{\mathbf{i}'}$, then the j-th entry of $\mathbf{i}$ is less than or equal to the j-th entry of $\mathbf{i}'$.*

To avoid triple subscripts below, we denote the j-th entry of a q-tuple $\mathbf{k}_l$ by $k_{(l,j)}$, that is, we write

$$\mathbf{k}_l = (k_{(l,1)}, k_{(l,2)}, \ldots, k_{(l,q)}) \in \mathbb{N}^q.$$

Using multi-indices, we write τ^a for $\tau_1^{a_1}\tau_2^{a_2}\ldots\tau_q^{a_q}$ with $a = (a_1, a_2, \ldots, a_q)$. A string consisting only of τ-letters can be written as $\tau^{\mathbf{k}}$ if and only if it satisfies Condition (1) in Definition 4.3.2. With this notation, one gets immediately the following explicit description of the words in $\mathcal{L}$.

Lemma 4.3.3. *The elements of the formal language $\mathcal{L}$ are precisely the words of the form*

$$\tau^{\mathbf{k}_1}\zeta_{\mathbf{i}_1}\tau^{\mathbf{k}_2}\zeta_{\mathbf{i}_2}\ldots\tau^{\mathbf{k}_d}\zeta_{\mathbf{i}_d}\tau^{\mathbf{k}_{d+1}},$$

where $\mathbf{i}_1, \ldots, \mathbf{i}_d \in [\mathbf{c}]$, $\mathbf{k}_1, \ldots, \mathbf{k}_{d+1} \in \mathbb{N}_0^q$, and $i_{(l-1,j)} \leq i_{(l,j)}$ whenever $k_{(l,j)} = 0$ for some (l,j) with $2 \leq l \leq d$ and $j \in [q]$.

The following elementary observation is useful.

Lemma 4.3.4. *Every monomial in $\operatorname{Mon}(A)$ can be uniquely written as a string of variables such that one has, the variable in any position l is of the form y_{j,i_j,k_j} with $j = l \bmod q$ and, for each $j \in [q]$, if a variable y_{j,i_j,k_j} appears to the left of y_{j,i'_j,k'_j}, then either $k_j < k'_j$ or $k_j = k'_j$ and $i_j \leq i'_j$.*

Proof. If for some j, two variables y_{j,i_j,k_j} and y_{j,i'_j,k'_j} appearing in a monomial do not satisfy the stated condition, then swap their positions. Repeating this step as long as needed results in a string meeting the requirement. It is unique, because the given condition induces an order on the variables $y_{j,i,k}$ with fixed j. In the desired string, for each fixed j, the variables $y_{j,i,k}$ occur in this order when one reads the string from left to right. □

We illustrate the above argument.

Example 4.3.5. *Let $q = 2$. To simplify notation write y_{jk} instead of $y_{1,j,k}$ and z_{jk} instead of $y_{2,j,k}$. Then one gets, for example,*

$$y_{22}z_{21}y_{14}z_{11}y_{31}z_{21} \quad \begin{matrix} y_{22}y_{14}y_{31} \;\text{———}\; y_{31}y_{22}y_{14} \\ z_{21}z_{11}z_{21} \;\text{———}\; z_{11}z_{21}z_{21} \end{matrix} \quad y_{31}z_{11}y_{22}z_{21}y_{14}z_{21}$$

We observed above that the map $\mathbf{m}$ sends each letter $\zeta_{\mathbf{i}}$ to the monomial $\Phi_{\mathbf{n}}(x_{\mathbf{i},\mathbf{1}})$. It follows that $\mathbf{m}(\Sigma^*)$ is a subset of $\mathrm{Mon}(A)$. In fact, one has the following result.

Proposition 4.3.6. *For any $\mathbf{n} \in \mathbb{N}_0^q$, denote by $\mathcal{L}_{\mathbf{n}}$ the set of words in $\mathcal{L}$ in which, for each $j \in [q]$, the letter τ_j occurs precisely n_j times. Then $\mathbf{m}$ induces for every $\mathbf{n} \in \mathbb{N}_0^q$ a bijection*

$$\mathbf{m}_{\mathbf{n}} \colon \mathcal{L}_{\mathbf{n}} \to \mathrm{Mon}(A_{\mathbf{n}+\mathbf{1}}),\; w \mapsto \mathbf{m}(w).$$

Proof. The definition of $\mathbf{m}$ readily implies $\mathbf{m}(w) \in \mathrm{Mon}(A_{\mathbf{n}+\mathbf{1}})$ if $w \in \mathcal{L}_{\mathbf{n}}$. First we show that $\mathbf{m}_{\mathbf{n}}$ is surjective. Let $m \in \mathrm{Mon}(A_{\mathbf{n}+\mathbf{1}})$ be any monomial. Its degree is dq for some $d \in \mathbb{N}_0$. By Lemma 4.3.4, m can be written as

$$m = \prod_{l=1}^{d} \left(\prod_{j=1}^{q} y_{j,i_{(l,j)},k_{(l,j)}} \right) = \prod_{l=1}^{d} \Phi_{\mathbf{n}}(x_{\mathbf{i}_l,\mathbf{k}_l})$$

such that, for each $j \in [q]$, one has

$$1 \le k_{(1,j)} \le \cdots \le k_{(d,j)} \le n_j + 1$$

and

$$i_{(l-1,j)} \le i_{(l,j)} \quad \text{ if } k_{(l,j)} = 0 \text{ for some } l.$$

The first condition implies that all the q-tuples $\mathbf{k}_1 - \mathbf{1}, \mathbf{k}_2 - \mathbf{k}_1, \ldots, \mathbf{k}_d - \mathbf{k}_{d-1}$ and $\mathbf{n} + \mathbf{1} - \mathbf{k}_d$ are in $\mathbb{N}_0^q$. Hence the string

$$w = \tau^{\mathbf{k}_1 - \mathbf{1}} \zeta_{\mathbf{i}_1} \tau^{\mathbf{k}_2 - \mathbf{k}_1} \zeta_{\mathbf{i}_2} \ldots \tau^{\mathbf{k}_d - \mathbf{k}_{d-1}} \zeta_{\mathbf{i}_d} \tau^{\mathbf{n}+\mathbf{1}-\mathbf{k}_d}$$

is defined. The two conditions together combined with Lemma 4.3.3 show that in fact m is in $\mathcal{L}_{\mathbf{n}}$. Hence $\mathbf{m}(w) = m$ proves the claimed surjectivity.

Second, we establish that $\mathbf{m}_{\mathbf{n}}$ is injective. Consider any two words $w, w' \in \mathcal{L}_{\mathbf{n}}$ with $\mathbf{m}(w) = \mathbf{m}(w')$. We will show $w = w'$.

Write w and w' as in Lemma 4.3.3:

$$w = \tau^{\mathbf{k}_1} \zeta_{\mathbf{i}_1} \tau^{\mathbf{k}_2} \zeta_{\mathbf{i}_2} \ldots \tau^{\mathbf{k}_d} \zeta_{\mathbf{i}_d} \tau^{\mathbf{k}_{d+1}}, \quad w' = \tau^{\mathbf{k'}_1} \zeta_{\mathbf{i'}_1} \tau^{\mathbf{k'}_2} \zeta_{\mathbf{i'}_2} \ldots \tau^{\mathbf{k'}_{d'}} \zeta_{\mathbf{i'}_{d'}} \tau^{\mathbf{k'}_{d'+1}}$$

Since $\mathbf{m}(w)$ has degree dq and $\mathbf{m}(w')$ has degree $d'q$, we conclude $d = d'$. Evaluating $\mathbf{m}$ we obtain

$$\prod_{l=1}^{d} (\prod_{j=1}^{q} y_{j,i_{(l,j)},f_{(l,j)}}) = \prod_{e=1}^{d} (\prod_{j=1}^{q} y_{j,i'_{(l,j)},f'_{(l,j)}}), \tag{4.3.1}$$

where $f_{(l,j)} = k_{(1,j)} + \cdots + k_{(l,j)} + 1$ and $f'_{(l,j)} = k'_{(1,j)} + \cdots + k'_{(l,j)} + 1$. Fix any $j \in [q]$. Comparing the third indices of the variables whose first index equals j and using that every index is non-negative, we get for each $l \in [d]$,

$$k_{(1,j)} + \cdots + k_{(l,j)} = k'_{(1,j)} + \cdots + k'_{(l,j)}.$$

It follows that $\mathbf{k}_l = \mathbf{k}'_l$ for each $l \in [d]$. Since w and w' are in $\mathcal{L}_{\mathbf{n}}$, we have $\mathbf{k}_{d+1} = \mathbf{n} - (\mathbf{k}_1 + \mathbf{k}_2 + \cdots + \mathbf{k}_d)$ and an analogous equation for $\mathbf{k}'_{d+1}$, which gives $\mathbf{k}_{d+1} = \mathbf{k}'_{d+1}$. It remains to show $\mathbf{i}_l = \mathbf{i}'_l$ for every $l \in [d]$. Fix any $j \in [q]$. If for some l there is only one variable of the form $y_{j,\mu,f_{(l,j)}}$ with $\mu \in [c_j]$ that divides $\mathbf{m}(w)$, this implies $i_{(l,j)} = i'_{(l,j)} = \mu$, as desired. Otherwise, there is a maximal interval of consecutive indices $k_{(l,j)}$ that are equal to zero, that is, there any integers a, b such that $1 \le a \le b \le d$ and

- $k_{(l,j)} = 0$ if $a \le l \le b$,
- $k_{(a-1,j)} > 0$, unless $a = 1$, and
- $k_{(b+1,j)} > 0$, unless $b = d$.

Thus, the number of variables of the form $y_{j,\mu,f_{(l,j)}}$ that divide $\mathbf{m}(w)$ is $b - a + 2$ if $a \ge 2$ and $b - a + 1$ if $a = 1$. Considering these variables, Lemma 4.3.3 gives

$$i_{(a-1,j)} \le i_{(a,j)} \le \cdots \le i_{(b,j)} \quad \text{and} \quad i'_{(a-1,j)} \le i'_{(a,j)} \le \cdots \le i'_{(b,j)},$$

where $i_{(a-1,j)}$ and $i'_{(a-1,j)}$ are omitted if $a = 1$. Using (4.3.1), it now follows that $i_{(l,j)} = i'_{(l,j)}$ whenever $a - 1 \le l \le b$, unless $a = 1$. If $a = 1$, the latter equality is true whenever $a \le l \le b$. Applying the latter argument to any interval of consecutive zero indices $k_{(l,j)}$, we conclude $i_{(l,j)} = i'_{(l,j)}$ for every $l \in [d]$. This completes the argument. □

Our next goal is to show that $\mathcal{L}$ is a regular language. By [27, Theorems 3.4 and 3.7], this is equivalent to proving that $\mathcal{L}$ is recognizable by a finite automaton. Recall that a finite automaton on an alphabet Σ is a 5-tuple $\mathcal{A} = (P, \Sigma, \delta, p_0, F)$ consisting of a finite set P of states, an initial state $p_0 \in P$, a set $F \subseteq P$ of accepting states and a transition map $\delta \colon D \to P$, where D is some subset of $P \times \Sigma$. The automaton $\mathcal{A}$ recognizes or accepts a word $w = a_1 a_2 \ldots a_s \in \Sigma^*$ if there is a sequence of states $r_0, r_1, \ldots, r_s$ satisfying $r_0 = p_0$, $r_s \in F$ and $r_{j+1} = \delta(r_j, a_{j+1})$ whenever $0 \le j < s$. The automaton $\mathcal{A}$ recognizes a formal language $\mathcal{L} \subseteq \Sigma^*$ if $\mathcal{L}$ is precisely the set of words in Σ^* that are accepted by $\mathcal{A}$.

Returning to the formal language $\mathcal{L}$ specified in Definition 4.3.2, we are ready to show:

Proposition 4.3.7. *The language $\mathcal{L}$ is recognized by a finite automaton.*

Proof. We need some further notation. We say that a sequence C of $l \ge 0$ integers $j_1, j_2, \ldots, j_l$ is an *increasing chain in* $[q]$ if $1 \le j_1 < j_2 < \cdots < j_l \le q$. Define $\max(C)$

as the largest element j_l of C. We put $\max(\emptyset) = 0$. We denote the set of increasing chains in $[q]$ by $\mathcal{C}$. Thus, the cardinality of $\mathcal{C}$ is 2^q. We write $j \in C$ if j occurs in the chain C. For any $\mathbf{k} \in \mathbb{N}_0^q$, we define the sequence of indices j with $k_j > 0$ as its support $\mathrm{Supp}(\mathbf{k})$. It is an element of $\mathcal{C}$. For example, one has $\mathrm{Supp}(7,0,1,5,0) = (1,3,4)$.

Now we define an automaton $\mathcal{A}$ as follows: Let

$$P = \{p_j, p_{\mathbf{i}}, p_{\mathbf{i},C,k} \mid 0 \leq j \leq q,\ \mathbf{i} \in [\mathbf{c}],\ C \in \mathcal{C},\ k \in C\}$$

be the set of states, where p_0 is the initial state of $\mathcal{A}$. Let

$$F = \{p_j, p_{\mathbf{i}}, p_{\mathbf{i},C,k} \mid 0 \leq j \leq q, \mathbf{i} \in [\mathbf{c}],\ C \in \mathcal{C},\ k = \max(C)\}$$

be the set of accepting states. Furthermore, define transitions

$$\delta(p_j, \tau_{j'}) = p_{j'} \text{ if } j = 0 < j' \leq q \text{ or } 1 \leq j \leq j' \leq q, \tag{4.3.2}$$
$$\delta(p_j, \zeta_{\mathbf{i}}) = p_{\mathbf{i}} \text{ if } 0 \leq j \leq q,\ \mathbf{i} \in [\mathbf{c}], \tag{4.3.3}$$
$$\delta(p_{\mathbf{i}}, \tau_j) = p_{\mathbf{i},C,j} \text{ if } \mathbf{i} \in [\mathbf{c}],\ C \in \mathcal{C},\ j \in C, \tag{4.3.4}$$
$$\delta(p_{\mathbf{i}}, \zeta_{\mathbf{i}'}) = p_{\mathbf{i}'} \text{ if } \mathbf{i}, \mathbf{i}' \in [\mathbf{c}],\ \mathbf{i} \leq \mathbf{i}', \tag{4.3.5}$$
$$\delta(p_{\mathbf{i},C,j}, \tau_k) = p_{\mathbf{i},C,k}, \text{ if } \mathbf{i} \in [\mathbf{c}],\ C \in \mathcal{C},\ j \in C,\ k \text{ directly follows } j \text{ in } C \text{ or } k = j, \tag{4.3.6}$$
$$\delta(p_{\mathbf{i},C,j}, \zeta_{\mathbf{i}'}) = p_{\mathbf{i}'} \text{ if } \mathbf{i}, \mathbf{i}' \in [\mathbf{c}],\ j = \max(C),\ i_k \leq i'_k \text{ whenever } k \notin C. \tag{4.3.7}$$

If an element of $P \times \Sigma$ does not satisfy any of the above six conditions then it is not in the domain of δ.

We claim that $\mathcal{A}$ recognizes $\mathcal{L}$. Indeed, let $w \in \Sigma^*$ be a word with exactly $d \geq 0$ ζ-letters. We show by induction on d that w is recognized by $\mathcal{A}$ if $w \in \mathcal{L}$, but any word in $\Sigma^* \backslash \mathcal{L}$ is not accepted by $\mathcal{A}$. It turns out that $w \in \mathcal{L}$ is accepted

- at a state p_j for some $0 \leq j \leq q$ if $d = 0$,
- at a state $p_{\mathbf{i}}$ for some $\mathbf{i} \in [c]$ if $d \geq 1$ and w ends with a ζ-letter, and
- at a state $p_{\mathbf{i},C,j}$ for some $\mathbf{i} \in [\mathbf{c}]$, $C \in \mathcal{C}$, $j = \max(C)$ if $d \geq 1$ and w ends with a τ-letter.

In particular, this explains the set of accepting states.

Consider any word $w \in \Sigma^*$ with exactly $d \geq 0$ ζ-letters. Assume $d = 0$, that is, $w = \tau_{l_1}\tau_{l_2}\ldots\tau_{l_t}$. By transition rule (4.3.2), $\mathcal{A}$ transitions from state p_0 to any state p_j with $j \in [q]$ using input τ_j. From any p_j with $j \in [q]$ the automaton can transition to any state $p_{j'}$ with $j \leq j' \leq q$ by using input $\tau_{j'}$. Thus, w is accepted by $\mathcal{A}$ if and only if $l_1 \leq l_2 \leq \cdots \leq l_t$, that is, $w \in \mathcal{L}$ (see Lemma 4.3.3).

Assume now that $d \geq 1$. We proceed in several steps.

(I) Assume $d = 1$ and w ends with a ζ-letter, that is,

$$w = \tau_{l_1}\tau_{l_2}\dots\tau_{l_t}\zeta_{\mathbf{i}}$$

for some $t \geq 0$. The argument for $d = 0$ shows that $\tau_{l_1}\tau_{l_2}\dots\tau_{l_t}$ is accepted if and only if it can be written as some $\tau^{\mathbf{k}}$. Processing input $\tau^{\mathbf{k}}$, the automaton arrives at state p_j with $j = \max(\mathrm{Supp}(\mathbf{k}))$. Using input $\zeta_{\mathbf{i}}$, it then transitions to $p_{\mathbf{i}} \in F$ by Rule (4.3.3). Hence w is accepted if and only of $w \in \mathcal{L}$.

(II) Let $d \geq 1$ and assume w ends with a τ-letter, that is, w can be written as

$$w = w'\zeta_{\mathbf{i}}\tau_{l_1}\tau_{l_2}\dots\tau_{l_t}$$

with $t \geq 1$. Furthermore assume that $w'\zeta_{\mathbf{i}}$ is accepted by $\mathcal{A}$ in state $p_{\mathbf{i}}$. We show that w is accepted by $\mathcal{A}$ if and only if $w = w'\zeta_{\mathbf{i}}\tau^{\mathbf{k}}$ for some $\mathbf{k} \in \mathbb{N}_0^q$. If w is recognized, it is accepted in state $p_{\mathbf{i},C,\max(C)}$, where $C = \mathrm{Supp}(\mathbf{k})$.

Indeed, let $C \in \mathcal{C}$ be the chain corresponding to the set $\{l_1, \dots, l_t\}$. Processing input τ_{l_1}, Rule (4.3.3) yields that $\mathcal{A}$ transitions to state $p_{\mathbf{i},C,l_1}$. If $t = 1$, then $l_1 = \max(C)$ and w is accepted in $p_{\mathbf{i},C,l_1} \in F$, as claimed. If $t \geq 2$, Rule (4.3.6) shows that $\mathcal{A}$ can transition from $p_{\mathbf{i},C,l_1}$ using input τ_{l_2} precisely if $l_2 \geq l_1$. If transition is possible $\mathcal{A}$ gets to state $p_{\mathbf{i},C,l_2}$. Hence Rule (4.3.6) guarantees that $\tau_{l_1}\tau_{l_2}\dots\tau_{l_t}$ can be processed by $\mathcal{A}$ if and only if $\tau_{l_1}\tau_{l_2}\dots\tau_{l_t} = \tau^{\mathbf{k}}$ for some non-zero $\mathbf{k} \in \mathbb{N}_0^q$. In this case $w = w'\zeta_{\mathbf{i}}\tau^{\mathbf{k}}$ is accepted by $\mathcal{A}$ in state $p_{\mathbf{i},C,\max(C)}$, where $C = \mathrm{Supp}(\mathbf{k})$.

(III) Assume now $w \in \Sigma^*$ ends with a ζ-letter, that is, w is of the form

$$w = w'\tau_{l_1}\tau_{l_2}\dots\tau_{l_t}\zeta_{\mathbf{i}},$$

where $w' \in \mathcal{L}$ is either empty or ends with a ζ-letter and $t \geq 0$. We show by induction on $d \geq 1$ that w is recognized by $\mathcal{A}$ if and only if $w \in \mathcal{L}$. In this case, w is accepted in state $p_{\mathbf{i}}$.

Indeed, if $d = 1$, i.e., w' is the empty word, this has been shown in Step (I). If $d \geq 2$ write $w' = w''\zeta_{\mathbf{i}'}$. If w' is not accepted by $\mathcal{A}$, then so is w. Furthermore, the induction hypothesis gives $w' \notin \mathcal{L}$, which implies $w \notin \mathcal{L}$.

If $w' = w''\zeta_{\mathbf{i}'}$ is recognized by $\mathcal{A}$ the induction hypothesis yields $w' \in \mathcal{L}$ and w' is accepted in state $p_{\mathbf{i}'}$. Step (II) shows that $w''\zeta_{\mathbf{i}'}\tau_{l_1}\tau_{l_2}\dots\tau_{l_t}$ is accepted by $\mathcal{A}$ if and only if it can be written as $w''\zeta_{\mathbf{i}'}\tau^{\mathbf{k}}$ for some $\mathbf{k} \in \mathbb{N}_0^q$, and so

$$w = w''\zeta_{\mathbf{i}'}\tau^{\mathbf{k}}\zeta_{\mathbf{i}}.$$

We consider two cases.

Case 1. Suppose $\mathbf{k}$ is zero, i.e., $\mathrm{Supp}(\mathbf{k}) = \emptyset$. Thus, $\mathcal{A}$ accepted $w''\zeta_{\mathbf{i}'} \in \mathcal{L}$ in state $p_{\mathbf{i}'}$. Using input $\zeta_{\mathbf{i}}$, Rule (4.3.5) shows that $\mathcal{A}$ does not halt in $p_{\mathbf{i}'}$ if and only if $\mathbf{i}' \leq \mathbf{i}$. By Lemma 4.3.3, this is equivalent to $w = w''\zeta_{\mathbf{i}'}\zeta_{\mathbf{i}} \in \mathcal{L}$. Furthermore, if w is in $\mathcal{L}$ it is accepted in state $p_{\mathbf{i}}$, as claimed.

Case 2. Suppose $\mathrm{Supp}(\mathbf{k}) \neq \emptyset$. Set $C = \mathrm{Supp}(\mathbf{k})$. By Step (II), $w''\zeta_{\mathbf{i}'}\tau^{\mathbf{k}}$ is accepted in state $p_{\mathbf{i}',C,j}$, where $j = \max(C)$. Hence Rule (4.3.7) gives that input $\zeta_{\mathbf{i}}$ can be processed if and only if $i'_l \leq i_l$ whenever $l \notin C$. By Lemma 4.3.3, this is equivalent to $w = w''\zeta_{\mathbf{i}'}\tau^{\mathbf{k}}\zeta_{\mathbf{i}} \in \mathcal{L}$. Moreover, if w is recognized it is accepted in state $p_{\mathbf{i}}$, as claimed.

(IV) By Steps (I) and (III) it remains to consider the case, where w ends with a τ-letter, i.e., $w = w'\zeta_{\mathbf{i}}\tau_{l_1}\tau_{l_2}\dots\tau_{l_t}$ with $t \geq 1$. By Step (III), $w'\zeta_{\mathbf{i}}$ is recognized by $\mathcal{A}$ if and only of $w'\zeta_{\mathbf{i}} \in \mathcal{L}$. Furthermore, if $w'\zeta_{\mathbf{i}} \in \mathcal{L}$ then it is accepted in state $p_{\mathbf{i}}$. Hence, the assumption in Step (II) is satisfied and we conclude that w is accepted if and only if $w = w'\zeta_{\mathbf{i}}\tau^{\mathbf{k}}$. The latter is equivalent to $w'\zeta_{\mathbf{i}}\tau^{\mathbf{k}} \in \mathcal{L}$ because $w'\zeta_{\mathbf{i}}$ is in $\mathcal{L}$. This completes the argument. □

We illustrate the automata constructed in Proposition 4.3.7 using the graphical representation in Remark 2.5.3.

Example 4.3.8. *Let $\mathcal{A}$ be the automaton constructed in Proposition 4.3.7 if $q = 3$ and $\mathbf{c} = (1, 1, 1)$. Note the only element in $[\mathbf{c}]$ is $\mathbf{1} = (1, 1, 1)$. To simplify notation, we write ζ for $\zeta_{1,1,1}$ and $p_{\mathbf{1}}$ for $p_{(1,1,1)}$. We denote the non-empty increasing chains in the interval $[3]$ as follows: $C_1 = \{1\}$, $C_2 = \{2\}$, $C_3 = \{3\}$, $C_4 = \{1, 2\}$, $C_5 = \{1, 3\}$, $C_6 = \{2, 3\}$, $C_7 = \{1, 2, 3\}$ and write $p_{i,j}$ instead of $p_{\mathbf{1},C_i,j}$. Using this notation, the constructed automaton $\mathcal{A}$ is represented by the following graph:*

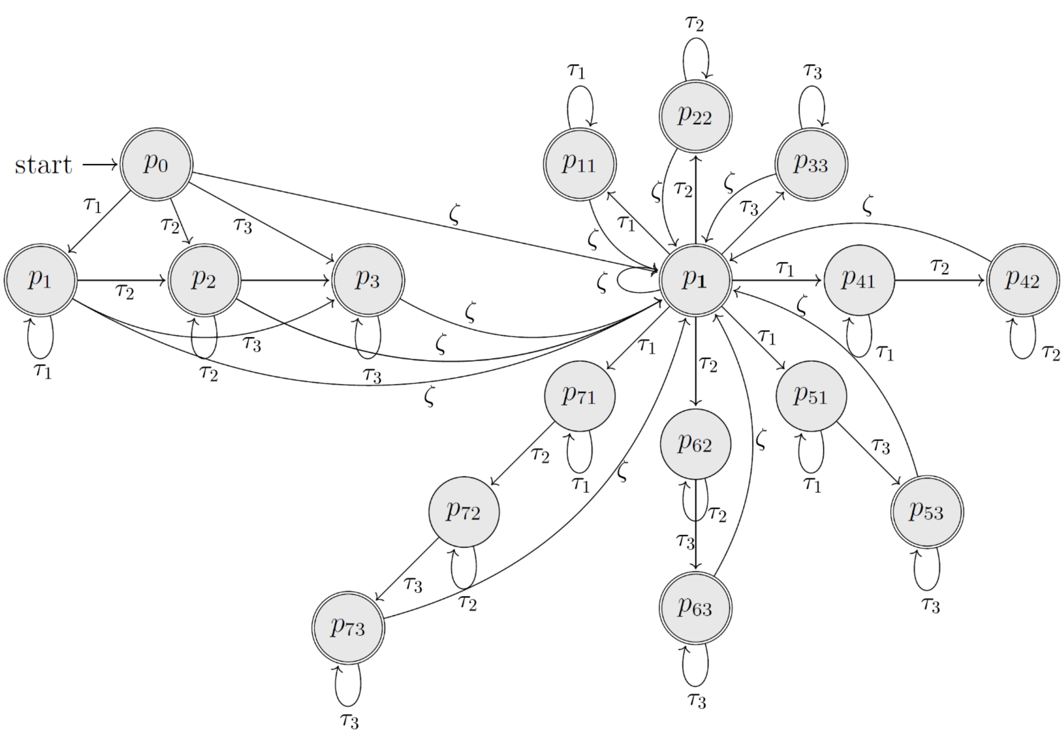


Figure 4.2: The automaton for $\mathbf{c} = (1, 1, 1)$ and $T = [3]$.

Remark 4.3.9. *The automaton constructed in Proposition 4.3.7 is often not the smallest automaton that recognizes the language $\mathcal{L}$. Using the minimization technique described in [27, Theorem 4.26], one can obtain an automaton with fewer states that also recognizes $\mathcal{L}$. For example, if $\mathbf{c} = (1, 1, 1)$, this produces an automaton with only four states in Figure 4.3.*

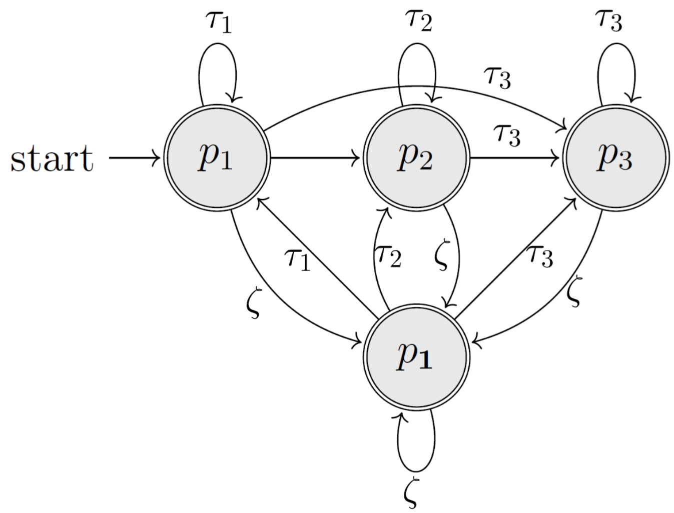


Figure 4.3: The reduced automaton for $\mathbf{c} = (1, 1, 1)$ and $T = [3]$.

In order to relate a language $\mathcal{L}$ on an alphabet Σ to a Hilbert series we need a suitable weight function. We will use theorem 2.5.1 to end the proof of Theorem 4.2.1 whose proof we had postponed.

Proof of Proposition 4.2.4. Since $I_{\mathbf{n}} = \ker \Phi_n$ and Φ_n is a homomorphism of degree q, we get $R_{\mathbf{n}}/I_n \cong A_{\mathbf{n}}$ and, for each $d \in \mathbb{Z}$,

$$\dim_{\mathbb{K}}[R_{\mathbf{n}}/I_{\mathbf{n}}]_d = \dim_{\mathbb{K}}[A_{\mathbf{n}}]_{dq}.$$

Recall that the algebra $A_{\mathbf{n}}$ is generated by monomials. Hence, every graded component has a $\mathbb{K}$-basis consisting of monomials. It follows that $\dim_{\mathbb{K}}[A_{\mathbf{n}}]_{dq} = \#\operatorname{Mon}([A_{\mathbf{n}}]_{dq})$. Therefore we get for the equivariant Hilbert series of the filtration $\mathscr{I}$:

$$equivH_{\mathscr{I}}(s_1, .., s_q, t) = \sum_{\mathbf{n} \in \mathbb{N}^q} \sum_{d \geq 0} \#\operatorname{Mon}([A_{\mathbf{n}}]_{dq}) \cdot s^{\mathbf{n}} t^d,$$

where $s^{\mathbf{n}} = s_1^{n_1} \cdots s_q^{n_q}$ if $\mathbf{n} = (n_1, \ldots, n_q)$.

Consider now the language $\mathcal{L}$ described in Definition 4.3.2. Define a weight function $\rho \colon \Sigma^* \to \operatorname{Mon}(\mathbb{K}[s_1, \ldots, s_q, t])$ by $\rho(\tau_j) = s_j$ and $\rho(\zeta_{\mathbf{i}}) = t$ for $\mathbf{i} \in [\mathbf{c}]$. Thus, for $w \in \mathcal{L}$, one obtains $\rho(w) = s^{\mathbf{n}} t^d$ if d is the number of ζ-letters occurring in w and n_j is the number of appearances of τ_j in w. Hence Proposition 4.3.6 gives that the number of words $w \in \mathcal{L}_{\mathbf{n}}$ with $\rho(w) = s^{\mathbf{n}} t^d$ is precisely $\#\operatorname{Mon}([A_{\mathbf{n+1}}]_{dq})$. Since $\mathcal{L}$ is the disjoint union of all $\mathcal{L}_{\mathbf{n}}$, it follows

$$s_1 \cdots s_q \cdot equivH_{\mathscr{I}}(s_1, .., s_q, t) = \sum_{\mathbf{n} \in \mathbb{N}_0^q} \sum_{w \in \mathcal{L}_n} \rho(w) = P_{\mathcal{L},\rho}(s_1, .., s_q, t). \tag{4.3.8}$$

As the right-hand side is rational by Theorem 2.5.1, the claim follows. □

Remark 4.3.10. *The method of proof for Theorem 4.2.1 is rather general and can also be used in other situations. An easy generalization is obtained as follows. Fix* $(a_1, \ldots, a_q) \in \mathbb{N}^q$*. For* $\mathbf{n} \in \mathbb{N}^q$*, consider the homomorphism*

$$\widetilde{\Phi}_{\mathbf{n}} \colon R_{\mathbf{n}} = \mathbb{K}[x_{\mathbf{i},\mathbf{k}} \mid (\mathbf{i}, \mathbf{k}) \in [\mathbf{c}] \times [\mathbf{n}]] \to \mathbb{K}[y_{j,i_j,k_j} \mid j \in [q], i_j \in c_j, k_j \in [n_j]] = S_{\mathbf{n}}$$
$$x_{\mathbf{i},\mathbf{k}} \longmapsto \prod_{j=1}^{q} y_{j,i_j,k_j}^{a_j},$$

and set $\widetilde{A}_{\mathbf{n}} = \operatorname{im} \Phi_{\mathbf{n}} = \mathbb{K}\left[\prod_{j=1}^{q} y_{j,i_j,k_j}^{a_j} \mid i_j \in c_j, k_j \in [n_j]\right]$*,* $\widetilde{I}_{\mathbf{n}} = \ker \widetilde{\Phi}_{\mathbf{n}}$*. Then* $\widetilde{\mathscr{I}} = \{\widetilde{I}_{\mathbf{n}}\}_{\mathbf{n} \in \mathbb{N}^q}$ *also is an* S_∞^q*-invariant filtration whose equivariant Hilbert series is rational. Indeed, this follows using the language* $\mathcal{L}$ *as above with the following modifications. In the definition of the map* $\boldsymbol{m}$ *change Rule (b) to* $\widetilde{\boldsymbol{m}}(\zeta_{\mathbf{i}} w) = \prod_{j=1}^{q} y_{j,i_j,1}^{a_j} \widetilde{\boldsymbol{m}}(w)$*, but keep Rules (a), (c) to obtain a map* $\widetilde{\boldsymbol{m}} \colon \Sigma^* \to \operatorname{Mon}(S)$*. It induces bijections* $\mathcal{L}_n \to \operatorname{Mon}(\widetilde{A}_{\mathbf{n+1}})$ *as in Proposition 4.3.6. Observe that* $[R_{\mathbf{n}}/\widetilde{I}_{\mathbf{n}}]_d \cong [\widetilde{A}_{\mathbf{n}}]_{da}$*, where* $a = a_1 + \cdots + a_q$*. Thus, using the same weight function* ρ *as above, we obtain* $s_1 \cdots s_q \cdot equivH_{\mathscr{I}}(s_1, .., s_q, t) = P_{\mathcal{L},\rho}(s_1, .., s_q, t)$*.*

A systematic study of substantial generalizations will be presented in [30].

4.4 Explicit formulas

We provide explicit formulas for the Hilbert series of hierarchical models considered in Theorem 4.2.1.

It is useful to begin by discussing Segre products more generally. To this end we temporarily use some new notation.

Lemma 4.4.1. *Let* $A = \mathbb{K}[a_1, \ldots, a_s] \subseteq R$ *and* $B = \mathbb{K}[b_1, \ldots, b_t] \subseteq S$ *be subalgebras of polynomial rings* $R = \mathbb{K}[x_1, \ldots, x_m]$ *and* $S = \mathbb{K}[y_1, \ldots, y_n]$ *that are generated by monomials* $a_1, \ldots, a_s$ *of degree* d_1 *and monomials* $b_1, \ldots, b_t$ *of degree* d_2*, respectively. Let* C *be the subalgebra of* $\mathbb{K}[x_1, \ldots, x_m, y_1, \ldots, y_n]$ *that is generated by all monomials* $a_i b_j$ *with* $i \in [s]$ *and* $j \in [t]$*. Using the gradings induced from the corresponding polynomials rings one has, for all* $k \in \mathbb{Z}$*,*

$$\dim_{\mathbb{K}}[C]_{k(d_1+d_2)} = \dim_{\mathbb{K}}[A]_{kd_1} \cdot \dim_{\mathbb{K}}[B]_{kd_2}.$$

Proof. This follows from the fact that the non-trivial degree components of the algebras A, B, C have $\mathbb{K}$-bases generated by monomials in the respective algebra generators of suitable degrees. □

It is customary to consider the algebras occurring in Lemma 4.4.1 as standard graded algebras that are generated in degree one by redefining their grading. In the new gradings, the degree k elements of A are elements that have degree kd_1, considered as polynomials in R, and similarly the degree k elements of C have degree $k(d_1 + d_2)$ when considered as elements of $\mathbb{K}[x_1, \ldots, x_m, y_1, \ldots, y_n]$. Using this new grading, the statement in the above lemma reads

$$\dim_{\mathbb{K}}[C]_d = \dim_{\mathbb{K}}[A]_d \cdot \dim_{\mathbb{K}}[B]_d. \tag{4.4.1}$$

This justifies to call C the *Segre product* of the algebras A and B. We denote it by $A \boxtimes B$.

Iterating the above construction we get the following consequence.

Corollary 4.4.2. *Let $A_1, \ldots, A_k$ be subalgebras of polynomial rings and assume every A_i generated by finitely many monomials of degrees d_i. Regrade such that every A_i is an algebra that is generated in degree one. Then one has*

$$\dim_{\mathbb{K}}[A_1 \boxtimes \cdots \boxtimes A_k]_d = \prod_{i=1}^{k} \dim_{\mathbb{K}}[A_i]_d.$$

We need an elementary observation.

Lemma 4.4.3. *Let $\omega \in \mathbb{C}$ be a primitive k-th root of unity. If*

$$f(t) = \sum_{n=0}^{\infty} c_n t^n$$

is a formal power series in t with complex coefficients, then

$$\sum_{n=0}^{\infty} c_{kn} x^{kn} = \frac{1}{k}\left[f(t) + f(\omega t) + \cdots + f(\omega^{k-1} t)\right].$$

Proof. Using geometric sums one gets, for every $n \in \mathbb{N}_0$,

$$\sum_{j=0}^{k-1} (\omega^j)^n = \begin{cases} k & \text{if } k \text{ divides } n \\ 0 & \text{else} \end{cases}$$

The claim follows. □

Proposition 4.4.4. *Fix any $q \in \mathbb{N}$ and let $\mathscr{I}$ be the S_∞^q-invariant filtration considered in Proposition 4.2.4. For $j \in [q]$, let ω_j be a c_j-th primitive root of unity. Then the equivariant Hilbert series of $\mathscr{I}$ is*

$$equivH_{\mathscr{I}}(s_1, \ldots, s_q, t) = \frac{1}{c_1 \cdots c_q} \sum_{m_1 \in [c_1], \ldots, m_q \in [c_q]} \frac{\omega_1^{m_1} s_1^{\frac{1}{c_1}} \cdots \omega_q^{m_q} s_q^{\frac{1}{c_q}}}{(1 - \omega_1^{m_1} s_1^{\frac{1}{c_1}}) \cdots (1 - \omega_q^{m_q} s_q^{\frac{1}{c_q}}) - t}.$$

Proof. By definition of the map $\Phi_{\mathcal{M}_{\mathbf{n}}}$, its image is isomorphic to the Segre product of polynomial rings of dimension $c_j n_j$ with $j = 1, \ldots, q$. Hence Corollary 4.4.2 gives for the equivariant Hilbert series

$$\begin{aligned} equivH_{\mathscr{I}}(s_1, \ldots, s_q, t) &= \sum_{d \geq 0, \mathbf{n} \in \mathbb{N}^q} \binom{c_1 n_1 + d - 1}{d} \cdots \binom{c_q n_q + d - 1}{d} s_1^{n_1} \ldots s_q^{n_q} t^d \\ &= \sum_{d \geq 0} \left\{ \prod_{j=1}^{q} \left[\sum_{n_j \in \mathbb{N}} \binom{c_j n_j + d - 1}{d} s_j^{n_j} \right] \right\} t^d \qquad (4.4.2) \end{aligned}$$

For any integer $d \geq 0$, one computes

$$\sum_{n\in\mathbb{N}} \binom{n+d-1}{d} s^n = s \sum_{n\in\mathbb{N}_0} \binom{d+n}{n} s^n = \frac{s}{(1-s)^{d+1}}.$$

Combined with Lemma 4.4.3 and using a c-th primitive root of unity $\omega \in \mathbb{C}$, we obtain, for any integer $c > 0$,

$$\sum_{n\in\mathbb{N}} \binom{cn+d-1}{d} s^n = \frac{1}{c} \sum_{m\in[c]} \frac{\omega^m s^{\frac{1}{c}}}{(1-\omega^m s^{\frac{1}{c}})^{d+1}}.$$

Applying the last formula to the inner sums in Equation (4.4.2) we get

$$\begin{aligned}
&equivH_{\mathscr{I}}(s_1, \dots, s_q, t) \\
&= \sum_{d\geq 0} \left\{ \prod_{j=1}^{q} \left[\frac{1}{c_j} \frac{\omega_j^m s_j^{\frac{1}{c_j}}}{(1-\omega_j^m s_j^{\frac{1}{c_j}})^{d+1}} \right] \right\} t^d \\
&= \sum_{d\geq 0} \frac{1}{c_1 \cdots c_q} \left\{ \sum_{m_1\in[c_1],\dots,m_q\in[c_q]} \frac{\omega_1^{m_1} s_1^{\frac{1}{c_1}}}{(1-\omega_1^{m_1} s_1^{\frac{1}{c_1}})^{d+1}} \cdots \frac{\omega_q^{m_q} s_q^{\frac{1}{c_q}}}{(1-\omega_q^{m_q} s_q^{\frac{1}{c_q}})^{d+1}} \right\} t^d \\
&= \frac{1}{c_1 \cdots c_q} \sum_{m_1\in[c_1],\dots,m_q\in[c_q]} \frac{\omega_1^{m_1} s_1^{\frac{1}{c_1}} \cdots \omega_q^{m_q} s_q^{\frac{1}{c_q}}}{(1-\omega_1^{m_1} s_1^{\frac{1}{c_1}}) \cdots (1-\omega_q^{m_q} s_q^{\frac{1}{c_q}}) - t},
\end{aligned}$$

as claimed. □

By Theorem 4.2.1, the above formula for the equivariant Hilbert series can be re-written as a rational function with rational coefficients.

Example 4.4.5. *(i) Let $c_1 = \cdots = c_q = 1$. Then Proposition 4.4.4 gives*

$$equivH_{\mathscr{I}}(s_1, s_2, \dots, s_q, t) = \frac{s_1 \dots s_q}{(1-s_1)\dots(1-s_q) - t}.$$

By the argument at the beginning of the proof of Lemma 4.2.3, this model has the same equivariant Hilbert series as the corresponding independence model (see Example 4.2.2).

(ii) Let $q = c_1 = c_2 = 2$. Then Proposition 4.4.4 yields

$$\begin{aligned}
4 \cdot equivH_{\mathscr{I}}(s_1, s_2, t) &= \frac{\sqrt{s_1 s_2}}{(1-\sqrt{s_1})(1-\sqrt{s_2}) - t} - \frac{\sqrt{s_1 s_2}}{(1-\sqrt{s_1})(1+\sqrt{s_2}) - t} \\
&\quad - \frac{\sqrt{s_1 s_2}}{(1+\sqrt{s_1})(1-\sqrt{s_2}) - t} + \frac{\sqrt{s_1 s_2}}{(1+\sqrt{s_1})(1+\sqrt{s_2}) - t}.
\end{aligned}$$

Now a straightforward computation gives

$$equivH_{\mathscr{I}}(s_1, s_2, t) = \frac{s_1 s_2 (s_1 s_2 - s_1 - s_2 - t^2)}{f},$$

where

$$f = s_1s_2(s_1-2)(s_2-2) + s_1(s_1-2) + s_2(s_2-2) \\ - 2t^2(s_1s_2+s_1+s_2) - 4t(s_1s_2-s_1-s_2) + (1-t)^4.$$

There is an alternative method to determine the equivariant Hilbert series whose rationality is guaranteed by Proposition 4.2.4. It directly produces a rational function with rational coefficients. This approach applies to any equivariant Hilbert series that is equal to the generating function $P_{\mathcal{L},\rho}$ determined by a weight function ρ on a regular language $\mathcal{L}$. Indeed, let $\mathcal{A} = (P, \Sigma, \delta, p_0, F)$ be a finite automaton that recognizes $\mathcal{L}$. Suppose P has N elements $p_0, \dots, p_{N-1}$. For every letter $a \in \Sigma$ define a $0-1$ matrix $M_{\mathcal{A},a}$ of size $N \times N$. Its entry at position (i,j) is 1 precisely if there is a transition $\delta(p_j, a) = p_i$. Let $\mathbf{e}_i \in \mathbb{K}^N$ be the canonical basis vector corresponding to state p_{i-1}. Let $\mathbf{u} = \sum_{p_{i-1} \in F} \mathbf{e}_i \in \mathbb{K}^N$ be the sum of the basis vectors corresponding to the accepting states. Then, for any word $w = w_1 \dots w_d$ with $w_i \in \Sigma$, one has

$$\mathbf{u}^T M_{\mathcal{A},w_d} \dots A_{\mathcal{A},w_1}\mathbf{e}_1 = \begin{cases} 1 & \text{if } \mathcal{A} \text{ accepts } w \\ 0 & \text{if } \mathcal{A} \text{ rejects } w. \end{cases}$$

Let $\rho : \Sigma^* \to \mathrm{Mon}(\mathbb{K}[s_1, \dots, s_k])$ be a weight function. Thus, $\rho(w_1w_2) = \rho(w_1) \cdot \rho(w_2)$ for any $w_1, w_2 \in \Sigma^*$. It follows (see, e.g, [39, Section 4.7]):

$$\begin{aligned} P_{\mathcal{L},\rho}(s_1, \dots, s_k) &= \sum_{w \in \mathcal{L}} \rho(w) = \sum_{d \geq 0} \sum_{w_1, \dots, w_d \in \Sigma} \mathbf{u}^T \left(\rho(w_1 \dots w_d) M_{\mathcal{A},w_d} \dots A_{\mathcal{A},w_1}\right)\mathbf{e}_1 \\ &= \sum_{d \geq 0} \mathbf{u}^T \left(\sum_{a \in \Sigma} \rho(a) M_{\mathcal{A},a}\right)^d \mathbf{e}_1 = \mathbf{u}^T \left(\mathrm{id}_N - \sum_{a \in \Sigma} \rho(a) M_{\mathcal{A},a}\right)^{-1} \mathbf{e}_1. \end{aligned}$$

Thus, the generating function $P_{\mathcal{L},\rho}(s_1, \dots, s_k)$ is rational with rational coefficients and can be explicitly computed from the automaton $\mathcal{A}$ using linear algebra.

In the proof of Proposition 4.2.4, we showed (see Equation (4.3.8)) that the equivariant Hilbert series of a considered filtration is, up to a degree shift, equal to a generating function. Hence, the above approach can be used to compute directly this Hilbert series as a rational function with rational coefficients. We implemented the resulting algorithm in Macaulay2 [19].

Example 4.4.6. *In Proposition 4.2.4, consider the case where* $\mathbf{c} = (1, 1 \dots, 1) \in \mathbb{N}^q$. *The automaton constructed in Proposition 4.3.7 can be reduced to one with only* $q+1$ *states (see Remark 4.3.9 if* $q = 3$*):*

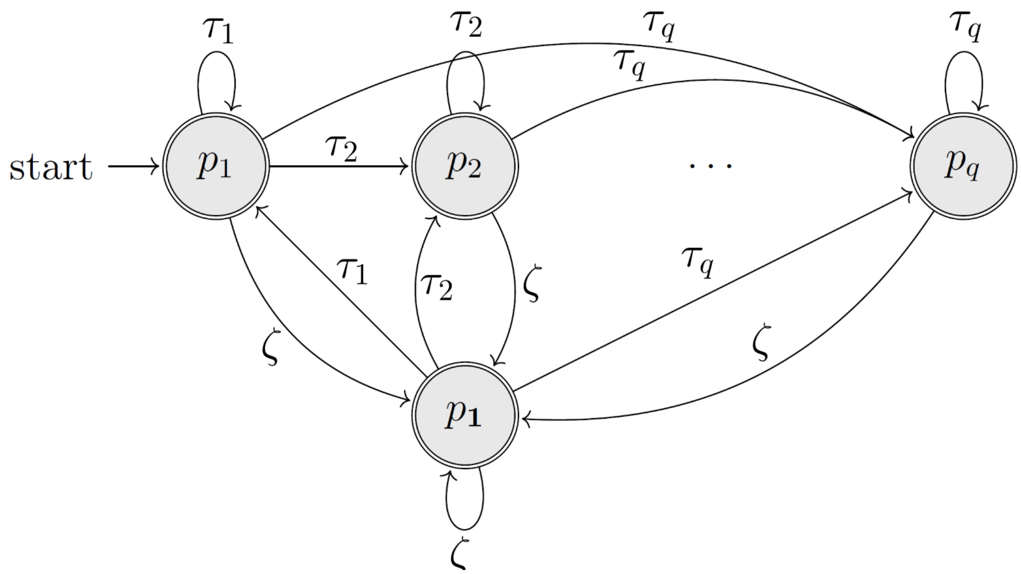


Figure 4.4: The reduced automaton for $\mathbf{c} = (1, \dots, 1)$ and $T = [q]$.

Hence, listing $p_{\mathbf{1}}$ as the last state we obtain for the equivariant Hilbert series of the filtration $\mathscr{I}$:

$$
\begin{aligned}
&\mathit{equivH}_{\mathscr{I}}(s_1, \dots, s_q, t) \\
&\quad = s_1 s_2 \cdots s_q \cdot \mathbf{u}^T \left(\mathrm{id}_{q+1} - \sum_{a \in \Sigma} \rho(a) M_{\mathcal{A},w} \right)^{-1} \mathbf{e}_1 \\
&\quad = s_1 s_2 \cdots s_q \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}^T \begin{bmatrix} 1-s_1 & 0 & 0 & \dots & 0 & 0 & -s_1 \\ -s_2 & 1-s_2 & 0 & \dots & 0 & 0 & -s_2 \\ -s_3 & -s_3 & 1-s_3 & \dots & 0 & 0 & -s_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ -s_{q-1} & -s_{q-1} & -s_{q-1} & \dots & 1-s_{q-1} & 0 & -s_{q-1} \\ -s_q & -s_q & -s_q & \dots & -s_q & 1-s_q & -s_q \\ -t & -t & -t & \dots & -t & -t & 1-t \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\
&\quad = \frac{s_1 \cdots s_q}{(1-s_1) \cdots (1-s_q) - t},
\end{aligned}
$$

where the first column of the inverse matrix can be determined using suitable minors. Of course, the result is the same as in Example 4.4.5.

4.5 Open Questions

One immediate question is if one can further generalize the conditions on Theorem 4.2.1. In the case of filtrations of ideals on polynomial rings of the form $\mathbb{K}[X_{[c] \times \mathbb{N}}]$ all filtrations that stabilize up to symmetric group action have a rational equivariant Hilbert series. We hope this is true for any invariant filtration of ideals that stabilizes.

Unfortunately, this seems to be less clear if that is true when the set T doesn't satisfy the Independence Set Theorem, i.e. $\#(T \cap F) \leq 1$ for all $F \in \Delta$. The following example shows how difficult this question is even for the simplest case that does not satisfy this condition.

Example 4.5.1. *Let $m = 2$ and consider the filtration $\mathscr{I} = (I_{\mathbf{r}})$, where every ideal $I_{\mathbf{r}}$ is zero. Since the ring $R_{(r_1,r_2)}$ has dimension $r_1 r_2$, one obtains*

$$
\begin{aligned}
equivH_{\mathscr{I}}(s_1, s_2, t) &= \sum_{(r_1,r_2)\in\mathbb{N}^2} H_{R_{(r_1,r_2)}}(t) \cdot s_1^{r_1} s_2^{r_2} = \sum_{(r_1,r_2)\in\mathbb{N}^2} \frac{1}{(1-t)^{r_1 r_2}} \cdot s_1^{r_1} s_2^{r_2} \\
&= \sum_{r_1 \geq 1} \left[-1 + \frac{(1-t)^{r_1}}{(1-t)^{r_1} - s_2} s_1^{r_1} \right].
\end{aligned}
$$

We do not know if this is a rational function in s_1, s_2 and t. However, if one considers the more standard Hilbert series with $r = r_1 = r_2$ one gets

$$
\sum_{r \geq 0} H_{R_{(r,r)}}(t) \cdot s^r = \sum_{n \geq 1} \frac{1}{(1-t)^{r^2}} \cdot s^r.
$$

This is not a rational function because the sequence $\left(\frac{1}{(1-t)^{r^2}}\right)_{r\in\mathbb{N}}$ does not satisfy a finite linear recurrence relation with coefficients in $\mathbb{Q}(t)$.

An indicator on the difficulty of obtaining a rational presentation is the growth of the Krull dimension of the ideal in such a filtration. Using Equation (2.3.3), one has the following corollaries.

Corollary 4.5.2. *(i) Let $\mathscr{I}_{\Delta,\mathbf{c}}$ be a filtration of ideals arising from the family of hierarchical models with the same structure Δ and fixed entries $\mathbf{c}$ in their vector of states. Treat Equation (2.3.3) as in equation in variables $\{r_t, t \in T\}$. Then Equation (2.3.3) is a polynomial in $\mathbb{Z}[r_t, t \in T]$ of degree $max\{|T \cap F|, for\ all\ F \in \text{facet}(\Delta)\}$.*

(ii) When $|F_j \cap T| \leq 1$ for any $F_j \in \text{facet}(\Delta)$, the Krull dimension formula has a linear form. The case when $\#T = 1$ corresponds to [32, Theorem 7.9].

Combining the position of the dimension R/I in the rational presentation of the Hilbert series, Corollary 4.5.2, and Definition 2.1.3, one has that the equivariant Hilbert series sums over rational functions with denominators that grow linearly in the case of the Independence Set Theorem, and that grow very fast (not linearly) in the other cases.

Another important topic is analyzing this rational form of the equivariant Hilbert series, if it exists. Can one predict the coefficients, the maximum degrees of the numerator and the denominator? Most importantly, what other information about the ideals in the filtration does this rational presentation provide? Nagel and Romer in [32, Section 6,7] have answered similar questions when $\#T = 1$, but in general these are all open questions.

We end this chapter with the promise that the rationality of an equivariant Hilbert series are useful in studying other filtrations of algebraic objects stabilized by some group action. In an ongoing joint project with Uwe Nagel, we study the rationality of equivariant Hilbert series of some filtrations of algebras and their Segre product, and we use regular languages for the proofs similarly to the work shown in this chapter.

Chapter 5 Polyhedral Geometry of Hierarchical Models

This chapter is based on joint work with Jane Ivy Coons, Joseph Cummings, and Ben Hollering. We study the polyhedral geometry of convex polytopes associated to the hierarchical models with binary states, known commonly as binary marginal polytopes. It is challenging to study marginal polytopes since they typically have a very high dimension (see [24, Proposition 9.3.10] for a formula on the dimension of the marginal polytope), and they are not full dimensional. The latter one causes non-unique half-space descriptions of the marginal polytope. One can avoid this difficulty by studying the polyhedral geometry of full dimensional polytopes that are affinely isomorphic to the marginal polytope. We introduce the generalized cut polytopes. These polytopes are full dimensional, affinely isomorphic to the binary marginal polytope, and agrees with the definition of the cut polytope of the suspension of a graph defined in [41]. We also involve the correlation polytopes, since they are affinely isomorphic to the binary marginal polytopes [42, Theorem 19.1.20]. All maps among polytopes are described explicitly. These maps allow us to easily transfer results about the polyhedral description among polytopes. We end the chapter by giving an explicit half-space description of the generalized cut polytope when Δ is the boundary of a simplex. The proof uses Gale tranformations and switch operators.

5.1 Preliminaries

This work uses [47] and [42, Chapter 8] as references to an introduction to polytopes and their polyhedral geometry. Given points $V = \{\mathbf{v}_1, \ldots, , \mathbf{v}_n\}$ in the affine space R^d, the convex polytope for V is its convex hull, i.e.

$$P(V) = \text{conv}(V) = \{\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n \mid \alpha_1, \ldots, \alpha_n \geq 0, \ \sum_i \alpha_i = 1\}.$$

The *dimension* of a polytope is defined to be the dimension of the smallest affine space containing it, and it is denoted $dim(P)$. Lemma 2.1.2 describes the connection between the dimension of the polytope and its toric ideal.

Polytopes can also be defined as intersection of half-spaces of the form $H = \{\mathbf{x} \in \mathbb{R}^d \mid a_1x_1 + \cdots + a_dx_d \leq b\}$, for some $\mathbf{a} \in R^d$ and $b \in \mathbb{R}$. Both descriptions are useful in analysing properties of the toric ideal of the matrix $A = \begin{bmatrix}\mathbf{v}_1 & \ldots & \mathbf{v}_n\end{bmatrix}$ (see Equation (2.1.2)). Often in applications, the second definition is valuable. Work of Rauh and Sullivant in [36] is a good example how one can make use of the half-space description to compute Markov bases for the hierarchical models. Books [26, 9] are good references on the connections of toric ideals with their corresponding polytopes.

Recall that a hierarchical model on m random variables is defined by a vector $\mathbf{r} = (r_1, \dots, r_m) \in \mathbb{N}^m$, and a simplicial complex Δ on $[m]$ that indicates the dependency relations among random variables.

Given facet(Δ) is the collection of facets of Δ, the toric ideal for the hierarchical model $\mathcal{M}(\Delta, \mathbf{r})$ can be written as the kernel of the following monomial map:

$$\Phi_{\mathcal{M}} \colon R_{\mathbf{r}} = \mathbb{K}[x_{\mathbf{i}} \mid \mathbf{i} \in [\mathbf{r}]] \xrightarrow{A} S_{\mathcal{M}} = \mathbb{K}[y_{T,\mathbf{i}_T} \mid T \in \mathcal{T}(\Delta), \mathbf{i}_T \in [\mathbf{r}_T]],$$
$$x_{\mathbf{i}} \longmapsto \prod_{T \in \Delta} y_{T,\mathbf{i}_T}.$$

The polytope $P_{\mathcal{M}}$ for the hierarchical model $\mathcal{M}$ is the convex hull of the column vectors of the matrix A. Marginal polytopes with all binary number of states, i.e. $r_i = 2$ for all $i \in [m]$, are denoted *binary marginal polytopes.* This polytope is called *marginal polytope.* Formally, the binary marginal polytope margP_Δ is the convex hull of vectors $\{\mathbf{v}^{\mathbf{i}} \mid \mathbf{i} \in \{1,2\}^m\}$ in $\mathbb{R}^C$, where $C = \{(T, \mathbf{j}_T) \mid T \in \text{facet}(\Delta), \mathbf{j}_T \in \{1,2\}^{\#T}\}$, has entries

$$\mathbf{v}^{\mathbf{i}}_{(T,\mathbf{j}_T)} = \begin{cases} 1 & \mathbf{i}_T = \mathbf{j}_T, \\ 0 & \text{else.} \end{cases}$$

Example 5.1.1. *The columns of A are vertices of the 6 dimensional binary marginal polytope for facet(Δ) $= C_3$, i.e. $\Delta = 2^{[3]} - [3]$.*

$$A = \begin{array}{c} \\ (\{1,2\},11) \\ (\{1,2\},21) \\ (\{1,2\},12) \\ (\{1,2\},22) \\ (\{1,3\},11) \\ (\{1,3\},21) \\ (\{1,3\},12) \\ (\{1,3\},22) \\ (\{2,3\},11) \\ (\{2,3\},21) \\ (\{2,3\},12) \\ (\{2,3\},22) \end{array} \begin{array}{c} \begin{array}{cccccccc} \mathbf{v}^{\{111\}} & \mathbf{v}^{\{211\}} & \mathbf{v}^{\{121\}} & \mathbf{v}^{\{221\}} & \mathbf{v}^{\{112\}} & \mathbf{v}^{\{212\}} & \mathbf{v}^{\{122\}} & \mathbf{v}^{\{222\}} \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \end{array}.$$

If the binary marginal polytope has a graph as its underlying structure, Sturmfels and Sullivant in [41] show that the toric ideal of a binary hierarchical model with a graph structure is isomorphic to the cut ideal for the suspension of that graph. Cut ideals and their polytopes are defined only on graphs, and they are very useful in optimization problems (see [11, 14]). In the following section we define a generalized version of the cut polytopes over any simplicial complex.

5.2 The Correlation Polytope and the Generalized Cut Polytope

Let Δ be a simplicial complex on a ground set $[m]$.

Definition 5.2.1. *The correlation polytope associated to Δ, denoted corrP_Δ is the convex hull of vectors $\{\mathbf{v}^S \mid S \subseteq [m]\}$ in $\mathbb{R}^{\Delta-\{\emptyset\}}$, with entries on $\mathbf{v}^S$ defined by*

$$\mathbf{v}^S_F = \begin{cases} 1 & F \subseteq S, \\ 0 & \text{else.} \end{cases}$$

Example 5.2.2. *The columns of B are vertices of the 6 dimensional correlation polytope for facet$(\Delta) = C_3$.*

$$B = \begin{array}{c} \\ \{1\} \\ \{2\} \\ \{1,2\} \\ \{3\} \\ \{1,3\} \\ \{2,3\} \end{array} \begin{array}{c} \begin{array}{cccccccc} \mathbf{d}^{\{\emptyset\}} & \mathbf{d}^{\{1\}} & \mathbf{d}^{\{2\}} & \mathbf{d}^{\{1,2\}} & \mathbf{d}^{\{3\}} & \mathbf{d}^{\{1,3\}} & \mathbf{d}^{\{2,3\}} & \mathbf{d}^{\{1,2,3\}} \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \end{array}.$$

We introduce the correlation polytope because it is isomorphic to the polytope of the hierarchical models associated to Δ where random variables have two states [41, Theorem 19.1.20]. In the case where Δ is a graph, the correlation polytope corrP$_\Delta$ is isomorphic to the cut polytope of the suspension of Δ. Here is the definition of the suspension of a graph G.

Definition 5.2.3. *Let $G([m], E)$ be a graph on the vertex set $[m]$. The suspension of G, denoted $\hat{G}$, is a graph on the vertex set $[m+1]$, and edges $\hat{E} = E \cup \{\{i, m+1\} \mid i \in [m]\}$. In other words, $\hat{G}$ is obtained by introducing a new vertex to G that is adjacent to all other vertices.*

Theorem 5.2.4. *[42, Proposition 19.1.21] Let $\hat{G}$ be the suspension of graph G. Then the polytope corrP_Δ is linearly isomorphic to cut$P_{\hat{G}}$.*

The proof is rooted in the covariance map on $\mathbb{R}^E$ defined by

$$\begin{aligned} \phi(\mathbf{v})_{i,m+1} &= \mathbf{v}_i && \text{for } i \in [m], \\ \phi(\mathbf{v})_{i,j} &= \mathbf{v}_i + \mathbf{v}_j - 2\mathbf{v}_{ij} && \text{for } \{i,j\} \in G. \end{aligned}$$

Definition 5.2.5. *The generalized cut polytope, gcutP_Δ, is the convex hull of vectors $\{\mathbf{d}^S \mid S \subseteq [m]\}$ in $\mathbb{R}^{\Delta-\{\emptyset\}}$, with*

$$\mathbf{d}^S_F = \begin{cases} 1 & \text{if } \#(F \cap S) \text{ odd,} \\ 0 & \text{if } \#(F \cap S) \text{ even.} \end{cases}$$

Definition 5.2.6. *The* generalized covariance mapping ϕ *is the map from* $\mathbb{R}^{\Delta-\{\emptyset\}}$ *to itself defined by*

$$\phi(\mathbf{v})_F = \sum_{\emptyset \neq H \subseteq F} (-2)^{\#H-1} \mathbf{v}_H.$$

Example 5.2.7. *Let* $\Delta = 2^{[3]} - [3]$ *be the the boundary of the* 3*-simplex. The columns of* C *are vertices of the 6 dimensional generalized cut polytope for* $\Delta = 2^{[3]} - [3]$.

$$A = \begin{array}{c} \{1\} \\ \{2\} \\ \{1,2\} \\ \{3\} \\ \{1,3\} \\ \{2,3\} \end{array} \overset{\begin{array}{cccccccc} \mathbf{d}^{\{\emptyset\}} & \mathbf{d}^{\{1\}} & \mathbf{d}^{\{2\}} & \mathbf{d}^{\{1,2\}} & \mathbf{d}^{\{3\}} & \mathbf{d}^{\{1,3\}} & \mathbf{d}^{\{2,3\}} & \mathbf{d}^{\{1,2,3\}} \end{array}}{\begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \end{bmatrix}}.$$

Proposition 5.2.8. *Given a simplicial complex* Δ*, the image of corr*P_Δ *under the generalized covariance map is gcut*P_Δ*.*

Proof. Let $S \subseteq [m]$. We will show that $\phi(\mathbf{v}^S)_F = \mathbf{d}^S_F$ by induction on the size of F. As a base case, we let $\#F = 1$. Then $\phi(\mathbf{v}^S)_F = (-2)^0 \mathbf{v}^S_F$ since F is the only nonempty subset of F. Therefore by definition of $\mathbf{v}^S$,

$$\begin{aligned} \phi(\mathbf{v}^S)_F &= \begin{cases} 1 & \text{if } F \subseteq S \\ 0 & \text{else} \end{cases} \\ &= \begin{cases} 1 & \text{if } \#(F \cap S) \text{ is odd} \\ 0 & \text{if } \#(F \cap S) \text{ is even} \end{cases} \\ &= \mathbf{d}^S_F. \end{aligned}$$

Let $\#F = k \geq 1$. Without loss of generality, assume $F = [k]$. There are two cases.
Case 1: $F \not\subseteq S$. Without loss of generality, suppose $k \notin S$. This implies that $[k] \cap S = [k-1] \cap S$. For any nonempty face $F' \subseteq F$, if $k \in H$ then $\mathbf{v}^S_H = 0$. So we have that

$$\phi(\mathbf{v}^S)_{[k]} = \sum_{\substack{\emptyset \neq H \subseteq [k], \\ k \notin H}} (-2)^{\#H-1} \mathbf{v}^S_H = \sum_{\emptyset \neq H \subseteq [k-1]} (-2)^{\#H-1} \mathbf{v}^S_H \qquad = \phi(\mathbf{v}^S)_{[k-1]}.$$

By the induction hypothesis, and that $[k] \cap S = [k-1] \cap S$, one has

$$\phi(\mathbf{v}^S)_{[k-1]} = \begin{cases} 1 & \text{if } \#([k-1] \cap S) \text{ is odd} \\ 0 & \text{if } \#([k-1] \cap S) \text{ is even} \end{cases} = \begin{cases} 1 & \text{if } \#(F \cap S) \text{ is odd} \\ 0 & \text{if } \#(F \cap S) \text{ is even,} \end{cases}$$

as needed.
Case 2: Let $F \subseteq S$. Then $\#(F \cap S) = k$ and for all $H \subseteq F$, $\mathbf{v}^S_H = 1$. So we have

$$\phi(\mathbf{v}^S)_F = \sum_{\substack{H \subseteq F \\ H \neq \emptyset}} (-2)^{\#H-1} = \sum_{i=1}^{k} \binom{k}{i} (-2)^{i-1} = \frac{(1+(-2))^k - 1}{-2}$$
$$= \begin{cases} 1 & \text{if } k \text{ is odd} \\ 0 & \text{if } k \text{ is even,} \end{cases}$$

as needed.

□

Proposition 5.2.9. *The generalized covariance map is invertible.*

Proof. The generalized covariance map in Proposition 5.2.8 defines the matrix A of the linear transformation from $\mathbb{R}^{\Delta - \{\emptyset\}}$ to itself that sends corrP_Δ to gcutP_Δ as

$$A^H_F = \begin{cases} (-2)^{\#H-1} & if\ H \subseteq F \\ 0 & if\ H \not\subseteq F. \end{cases}$$

Consider the linear transformation B from $\mathbb{R}^{\Delta - \{\emptyset\}}$ to itself as

$$B^G_H = \begin{cases} \dfrac{(-1)^{\#G-1}}{2^{\#H-1}} & if\ G \subseteq H \\ 0 & if\ G \not\subseteq H. \end{cases}$$

This linear transformation is the inverse of the generalized convariance map, i.e. $AB = \mathbf{I}_{\Delta - \{\emptyset\}}$. Indeed,

$$(AB)^G_F = A_F \cdot B^G = \sum_{H \in \Delta - \{\emptyset\}} A^H_F \cdot B^G_H.$$

Given H, the entry $A^H_F \neq 0$ only when $H \subseteq F$, and the entry $B^G_H \neq 0$ only when $G \subseteq H$. Hence, these entries are both nonzero only when $G \subseteq H \subseteq F$. In the diagonal entries of AB, i.e. when $G = F$, one has that these entries are nonzero only when $H = F = G$. In this situation one has

$$(AB)^G_F = (-2)^{\#F} \cdot \frac{(-1)^{\#F-1}}{2^{\#F-1}} = 1.$$

When $F \neq G$, one has

$$\begin{aligned}
(AB)_F^G &= \sum_{G \subseteq H \subseteq F} (-2)^{\#H} \cdot \frac{(-1)^{\#G-1}}{2^{\#H-1}} \\
&= (-1)^{\#G-1} \sum_{G \subseteq H \subseteq F} (-1)^{\#H-1} \\
&= (-1)^{\#G-1} \sum_{G \subseteq H \subseteq F} (-1)^{\#G+\#(H \setminus G)-1} \\
&= (-1)^{2(\#G-1)} \sum_{H' \subseteq F \setminus G} (-1)^{\#G-1}(-1)^{\#(H')} \\
&= 1 \cdot 0 = 0.
\end{aligned}$$

□

Denote the facets of a simplicial complex Δ with $\mathcal{T}(\Delta)$. The bijection between sequences of $\{1,2\}^m$ and subsets of $[m]$ induces a bijection between sets $D = \{(T,F) \mid F \subseteq T \in \text{facet}(\Delta)\}$ and $\{(T, \mathbf{i}_T) \mid T \in \text{facet}(\Delta), \mathbf{i}_T \in \{1,2\}^{\#T}\}$. One can rewrite the definition of the binary marginal polytope as the convex hull of vectors $\{\mathbf{v}^S \mid S \in [m]\}$ in $\mathbb{R}^D$, where

$$\mathbf{v}^S_{(T,F)} = \begin{cases} 1 & if\ S \cap T = F \\ 0 & \text{else.} \end{cases} \tag{5.2.1}$$

Define the affine transformation $A\mathbf{x} + \mathbf{b}$ from $\mathbb{R}^D$ to $\mathbb{R}^{\Delta - \{\emptyset\}}$ with $\mathbf{b} = \mathbf{v}^{\{\emptyset\}}$, and

$$A^H_{(T,F)} = \begin{cases} (-1)^{\#F+\#H} & if\ F \subseteq H \subseteq T \\ 0 & \text{else.} \end{cases}$$

Proposition 5.2.10. *[42, Proposition 19.1.20] Given a simplicial complex Δ, corrP$_\Delta$ is affinely isomorphic to margP$_\Delta$.*

Proof. It will be proven that the affine transformation $A\mathbf{x} + \mathbf{v}^{\{\emptyset\}}$ sends corrP$_\Delta$ to margP$_\Delta$. The "translation" vector $\mathbf{b}$ must be $\mathbf{v}^{\{\emptyset\}}$ since the vector in the correlation polytope $\mathbf{u}^{\{\emptyset\}}$, which coincides to the zero vector, must be mapped to $\mathbf{v}^{\{\emptyset\}}$, i.e.

$$A\mathbf{u}^{\{\emptyset\}} + \mathbf{b} = \mathbf{v}^{\{\emptyset\}} \rightarrow \mathbf{b} = \mathbf{v}^{\{\emptyset\}}.$$

The new polytope margP'$_\Delta$ that arises from translating margP$_\Delta$ by $\mathbf{b}$ has vertex description

$$\mathbf{w}^S_{(T,F)} = \begin{cases} 1 & if\ S \cap T = F, F \neq \{\emptyset\} \\ -1 & if\ S \cap T \neq F, F = \{\emptyset\}, \\ 0 & \text{else.} \end{cases}$$

To prove the proposition it is enough to prove that the linear transformation with A as its matrix sends corrP_Δ to margP'$_\Delta$.

$$A_{(T,F)} \cdot \mathbf{w}^S = \sum_{H \in \Delta \setminus \{\emptyset\}} A^H_{(F,T)} \mathbf{w}^S_H = \sum_{\substack{H \in \Delta - \{\emptyset\} \\ F \subseteq H \subseteq T \cap S}} A^H_{(F,T)} \mathbf{w}^S_H.$$

The latter equality is true since $A^H_{(T,F)}$ and $\mathbf{w}^H_S$ are not zero only when $F \subseteq H \subseteq T$, and $H \subseteq S$, respectively.

Case 1: $S \cap T = F$, $F = \{\emptyset\}$. There is no nonzero face in Δ with $F \subseteq H \subseteq T \cap S$. Hence, $A_{(T,F)} \cdot \mathbf{w}^S = 0$.

Case 2: When $S \cap T = F$ and $F \neq \{\emptyset\}$, the only the nonzero face in Δ that satisfies $F \subseteq H \subseteq T \cap S$ is $H = F$. Hence, $A_{(T,F)} \cdot \mathbf{w}^S = (-1)^{\#F + \#F} = 1$.

Case 3: When $S \cap T \neq F$ and $F = \{\emptyset\}$, one has

$$A_{(T,F)} \cdot \mathbf{w}^S = \sum_{\substack{H \in \Delta - \{\emptyset\} \\ H \subseteq T \cap S}} (-1)^{\#H} = \sum_{\{\emptyset\} \neq H \subseteq T \cap S} (-1)^{\#H} = \sum_{H \subseteq T \cap S} (-1)^{\#H} - 1 = 0 - 1 = -1.$$

Case 4: Lastly, when $S \cap T \neq F$, $F \neq \{\emptyset\}$ induces

$$\begin{aligned} A_{(T,F)} \cdot \mathbf{w}^S &= \sum_{\substack{H \in \Delta - \{\emptyset\} \\ F \subseteq H \subseteq T \cap S}} (-1)^{\#H + \#F} \\ &= (-1)^{\#F} \sum_{\substack{H \in \Delta - \{\emptyset\} \\ F \subseteq H \subseteq T \cap S}} (-1)^{\#H} \\ &= (-1)^{2\#F} \sum_{H' \subseteq (T \cap S) \setminus F} (-1)^{\#H'} \\ &= 0. \end{aligned}$$

□

Proposition 5.2.11. *The generalized cut polytope and the correlation polytope of a simplicial complex Δ are full dimensional; that, is the number of nonempty faces of Δ.*

Proof. The correlation polytope for Δ is isomorphic to the polytope of the hierarchical model associated to Δ where each random variable has a binary choice. The formula on the dimension of such polytopes in [42, Proposition 9.3.10] induces

$$\dim(\mathrm{gcutP}_\Delta) = \left[\sum_{F \in \Delta} \prod_{f \in F} (2-1) \right] - 1 = \left[\sum_{F \in \Delta} 1 \right] - 1 = \sum_{\emptyset \neq F \in \Delta} 1 = \#\Delta - 1.$$

□

Note that for a graph G, we do *not* have that $\text{gcutP}_\Delta \cong \text{cutP}_\Delta$. However, we do have that $\text{gcutP}_G \cong \text{cutP}_{\hat{G}}$. We can see this by simply renaming $\mathbf{d}^{A|B} \in \text{cutP}_{\hat{G}}$ by $\mathbf{d}^A \in \text{gcutP}_\Delta$ and relabeling the $\{i, m+1\}$ coordinates of $\text{cutP}_{\hat{G}}$ by $\{i\}$.

Lastly, let $I_{gcut(\Delta)}$ be the toric ideal induced by the vectors that define the generalized cut polytope, and $I_{Marg(\Delta)}$ be the toric ideal induced by vectors in Equation (5.2.1). By the way the Equation (5.2.1) got defined, one has $I_{\mathcal{M}(\Delta,\mathbf{2})} \cong I_{Marg(\Delta)}$. The affine transformations in Proposition 5.2.8 and Proposition 5.2.10 together Equation (2.1.1) and Equation (2.1.2) induce the following result on the toric ideals.

Corollary 5.2.12. *Given Δ a simplicial complex, $I_{gcut(\Delta)} = I_{Marg(\Delta)} \cong I_{\mathcal{M}(\Delta,\mathbf{2})}$.*

5.3 The Switching Operation for the Generalized Cut Polytopes

Definition 5.3.1. *Let $\mathbf{ax} \leq c$ be a valid inequality for $gcutP_\Delta$, and let $S \subseteq [m]$. We define the map $\mathbf{a}(S)$ on $R^{\Delta - \{\emptyset\}}$ by*

$$(\mathbf{a}(S))_F = (-1)^{\#(S \cap F)} \mathbf{a}_F.$$

The switching *of the inequality $\mathbf{ax} \leq c$ with respect to the set S is the inequality,*

$$\mathbf{a}(S)\mathbf{x} \leq c - \mathbf{ad}^S.$$

Given the inequality $\mathbf{ax} \leq c$ and $\mathcal{S} \subseteq 2^{[m]}$, denote

$$switch_{\mathcal{S}}(\mathbf{ax} \leq c) = \{\mathbf{a}(S)\mathbf{x} \leq c - \mathbf{ad}^S \mid S \in \mathcal{S}\}.$$

For each $S \subseteq [m]$, we denote by $supp(\mathbf{d}^S)$ the support of $\mathbf{d}^S$. That is,

$$supp(\mathbf{d}^S) = \{F \in \Delta \mid \#(F \cap S) \text{ is odd}\}.$$

Proposition 5.3.2. *The set of all supports of vertices of $gcutP_\Delta$ is closed under taking symmetric differences.*

Proof. Let $S, T \subseteq [m]$. We will show that $supp(\mathbf{d}^S) \triangle supp(\mathbf{d}^T)$ is the support of $\mathbf{d}^{S \triangle T}$. Let $S' = S \backslash T$ and let $T' = T \backslash S$. Let $F \in \Delta$. Let $c = \#(F \cap S \cap T)$ so that

$$\#(F \cap S) = \#(F \cap S') + c,$$

and similarly for T. There are three cases.

Case 1: Let $\mathbf{d}^S_F = \mathbf{d}^T_F = 1$ so that $F \in supp(\mathbf{d}^S)$ and $F \in supp(\mathbf{d}^T)$. So $F \notin supp(\mathbf{d}^{S \triangle T})$. Then we can write $\#(F \cap S) = 2a+1$ and $\#(F \cap T) = 2b+1$ for some non-negative integers a and b. So $\#(F \cap S') = 2a + 1 - c$ and $\#(F \cap T') = 2b + 1 - c$. So

$$\#(F \cap (S \triangle T)) = \#(F \cap S') + \#(F \cap T') = 2(a + b - c + 1)$$

is an even number. Therefore $\mathbf{d}^{S \triangle T}_F = 0$, and $F \notin supp(\mathbf{d}^{S \triangle T})$, as needed.

Case 2: Without loss of generality, let $\mathbf{d}_F^S = 1$ and $\mathbf{d}_F^T = 0$ so that $F \in supp(\mathbf{d}^S)$ and $F \notin supp(\mathbf{d}^T)$. So $F \in supp(\mathbf{d}^S) \triangle supp(\mathbf{d}^T)$. We have that $\#(F \cap S') = 2a+1-c$ and $\#(F \cap T') = 2b - c$ for some nonnegative integers a and b. So

$$\#(F \cap (S \triangle T)) = \#(F \cap S') + \#(F \cap T') = 2(a + b - c) + 1$$

is an odd number. Therefore $\mathbf{d}_F^{S \triangle T} = 1$, and $F \in supp(\mathbf{d}^{S \triangle T})$, as needed.

Case 3: Let $\mathbf{d}_F^S = \mathbf{d}_F^T = 0$ so that $F \notin supp(\mathbf{d}^S), supp(\mathbf{d}^T)$. Then $F \notin supp(\mathbf{d}^S) \triangle supp(\mathbf{d}^T)$. We have that $\#(F \cap S') = 2a - c$ and $\#(F \cap T') = 2b - c$ for some non-negative integers a and b. So

$$\#(F \cap (S \triangle T)) = \#(F \cap S') + \#(F \cap T') = 2(a + b - c)$$

is an even number. Therefore $\mathbf{d}_F^{S \triangle T} = 0$, and $F \notin supp(\mathbf{d}^{S \triangle T})$, as needed. □

Corollary 5.3.3. *Let $S \subseteq [m]$. The inequality $\mathbf{ax} \leq c$ is valid (resp. facet-defining) for $gcutP_\Delta$ if and only of its switching $\mathbf{a}(S)\mathbf{x} \leq c - \mathbf{ad}^S$ is valid (resp. facet defining) for $gcutP_\Delta$.*

Proof. This corollary holds for any polytope for which the set of supports of the vertices is closed under taking symmetric differences [11, Chapter 26.3]. □

5.4 The Boundary of a Simplex

Let Δ be the boundary of an m-dimensional simplex. Since Δ is a non-reducible simplicial complex, its induced hierarchical models are non-reducible, and little is known about them. The authors of [24] showed that the marginal polytope associated to Δ, and therefore gcutP_Δ has 4^{m-1} facets. We will identify these facets using the Gale transform of gcutP_Δ.

Given $S \subseteq [m]$ denote,

$$\mathcal{E}_S = \{A \subseteq [m] \mid \#(A \cap S) \text{ } even\}.$$

Theorem 5.4.1. *Let $\Delta = 2^{[m]} - [m]$ be the boundary of a simplex. Its generalized cut polytope is defined by the collection of half-spaces*

$$\mathscr{H}(gcutP_\Delta) = \{switch_E(\mathbf{a}^S \cdot \mathbf{x} \leq 2^{m-2}) \mid S \subseteq [m], \#S \text{ } odd, \ E \in \mathcal{E}_S\},$$

where

$$\mathbf{a}_F^S = \begin{cases} 0 & if \ \#(S \cap F) \text{ } is \text{ } odd \\ 1 & if \ \#(S \cap F) \text{ } is \text{ } even. \end{cases}$$

The following lemmas and propositions will be used to prove this theorem.

Lemma 5.4.2. *Fix Δ to be the boundary of the m-dimensional simplex. Then a Gale transform of $gcutP_\Delta$ is $\mathcal{B} = \{(-1)^{\#S} \mid S \subseteq [m]\} \subseteq \mathbb{R}$.*

Proof. Let M be the $2^m - 1 \times 2^m$ matrix whose columns are $[1\ \mathbf{d}^A]^T$. Since gcutP_Δ has dimension $\#\Delta - 1 = 2^m - 2$, the dimension of the kernel of M is 1. If $\{e_S \mid S \subseteq [m]\}$ is the standard basis for $\mathbb{R}^{2^m}$, it will be shown that this kernel is generated by the vector $\alpha \in \mathbb{R}^{2^m}$

$$\alpha = \sum_{S \subseteq [m]} (-1)^{\#S} e_S.$$

It is enough to prove that

$$M\alpha = \sum_{S \subseteq [m]} (-1)^{\#S} \begin{pmatrix} 1 \\ \mathbf{d}^A \end{pmatrix} = \mathbf{0}.$$

The first entry of $M\alpha$ is equal to $\sum_{S \subseteq [m]} (-1)^{\#S}$, which is zero by the binomial theorem. The rest of the entries are indexed by the faces in Δ. For any $F \in \Delta$, the F-th coordinate of $M\alpha$ is

$$\begin{aligned} [M\alpha]_F &= \sum_{S \subseteq [m]} (-1)^{\#S} \mathbf{d}^S_F \\ &= \sum_{\substack{S \subseteq [m] \\ \#(S \cap F) \text{ odd}}} (-1)^{\#S}. \end{aligned}$$

One can prove that the last sum is zero via the following sign-reversing involution. F cannot be $[m]$, since $[m]$ is not in Δ. Hence, there exists a $k \in [m]$, such that $k \notin F$. Take the involution

$$\begin{aligned} f\colon \{S \subseteq [m] \mid \#(S \cap F)\ \mathit{odd}\} &\longrightarrow \{S \subseteq [m] \mid \#(S \cap F)\ \mathit{odd}\}, \\ S &\longmapsto S \triangle \{k\}. \end{aligned}$$

Take

$$\begin{aligned} \epsilon\colon \{S \subseteq [m] \mid \#(S \cap F)\ \mathit{odd}\} &\longrightarrow \{-1, 1\}, \\ S &\longmapsto (-1)^{\#S}. \end{aligned}$$

The equality $\epsilon(f(S)) = -\epsilon(S)$ makes f a sign-reversing involution. Since for every such S, $S\triangle\{k\}$ is either $S\backslash\{k\}$ or $S \cup \{k\}$, f doesn't have fixed points, and the sum in $*$ must be zero. Finally, the Gale transformation of gcutP_Δ is given by the rows of α which are as claimed. □

Lemma 5.4.3. *The facets of $gcutP_\Delta$ are given by $P_{E,T} := conv(\mathbf{d}^S \mid S \notin \{E, T\})$ where $\#E$ and $\#T$ have different parity.*

Proof. Let $\mathcal{B}$ be as in Lemma 5.4.3. By Theorem 2.6.3, given $\mathcal{J} \subseteq 2^{[m]}$, a set of vectors $\{(-1)^{\#S} \mid S \in \mathcal{J}\}$ in $\mathcal{B}$ lie on a common face if and only if $0 \in \mathrm{relint}(\mathrm{conv}(b^S \mid S \notin \mathcal{J}))$. The latter one happens when $\mathrm{relint}(\mathrm{conv}(b^S \mid S \notin \mathcal{J})$ contains -1 and 1, i.e. there are $E, T \subseteq [m]$ not in $\mathcal{J}$ with $\#E$ even and $\#T$ odd. Such maximal sets $\mathcal{J}$ are the ones that do not contain exactly a set E of even size and a set T of odd size. □

Given E of even cardinality and T of odd cardinality in $[m]$, denote $\mathcal{V}^{E,T} = \{\mathbf{d}^S \mid S \notin \{E,T\}\}$. In the proof the following lemmas, we will frequently make use of the fact that for any set S, there are $2^{\#S-1}$ subsets of S of even cardinality, and $2^{\#S-1}$ subsets of S of odd cardinality.

Lemma 5.4.4. *Let $T \subseteq [m]$ have odd cardinality. The facet of $gcutP_\Delta$ with vertex set $V^{T,\{\emptyset\}}$ is defined by the equation*

$$\mathbf{a}^T\mathbf{x} = 2^{m-2}. \tag{5.4.1}$$

Proof. First note that $\mathbf{a}^T\mathbf{d}^{\{\emptyset\}} = 0$ and $\mathbf{a}^T\mathbf{d}^T = 0$. So $\mathbf{d}^{\{\emptyset\}}$ and $\mathbf{d}^T$ do not lie on the hyperplane defined by Equation (5.4.1). We claim that for any $S \neq \{\emptyset\}, T$, one has $\mathbf{a}^T\mathbf{d}^S = 2^{m-2}$. Recalling the definitions of $\mathbf{a}^T$ and $\mathbf{d}^S$, it is enough to prove that $\mathbf{a}^T_F = \mathbf{d}^S_F = 1$ for exactly 2^{m-2} faces F in Δ. i.e. there are exactly 2^{m-2} elements $F \in \Delta$ with $\#(F \cap T)$ even and $\#(F \cap S)$ odd.

Case 1: $S \subsetneq T$ ($T \subsetneq S$ is analogous). Given $F \subseteq [m]$, one has

$$F = (F \cap S) \sqcup (F \cap T\backslash S) \sqcup (F \cap [m]\backslash T),$$

and

$$F \cap T = (F \cap S) \sqcup (F \cap T\backslash S).$$

The first equation says that one can describe F as the union of its disjoint intersections with S, $T\backslash S$, and $[m]\backslash T$, respectively. The second equation gives information about the parity of each of the intersections; $\#(F \cap T)$ has even parity and $F \cap S$ has odd parity, $\#(F \cap T\backslash S)$ must have odd parity, and $\#(F \cap [m]\backslash T)$ can have any parity since it doesn't contribute on the parity of $\#(F \cap T)$ and $\#(F \cap S)$. There are $2^{\#S-1}$ ways to choose a set F_1 of odd cardinality in S. There are $2^{(\#T-\#S)-1}$ ways to choose a set F_2 of odd cardinality in $T\backslash S$, and there are $2^{m-\#T}$ ways to choose a set in $[m]\backslash T$. Hence, there are

$$2^{\#S-1}2^{(\#T-\#S)-1}2^{m-\#T} = 2^{m-2}$$

ways to pick $F = F_1 \sqcup F_2 \sqcup F_3$ in Δ.

Case 2: $S \nsubseteq T$ and $S \nsubseteq T$. Here one has

$$F = (F \cap (S \cap T)) \sqcup (F \cap T\backslash(S \cap T)) \sqcup (F \cap S\backslash(S \cap T)) \sqcup (F \cap [m]\backslash(T \cup S)),$$

$$F \cap T = (F \cap (S \cap T)) \sqcup (F \cap T\backslash(S \cap T)),$$

and

$$F \cap S = (F \cap (S \cap T)) \sqcup (F \cap S\backslash(S \cap T)).$$

The first equation says that F is union of disjoints sets in $S\cap T$, $T\backslash(S\cap T)$, $S\backslash(S\cap T)$, and $[m]\backslash(T \cup S)$. The second and third equations show that $\#(F \cap (S \cap T))$, $\#(F \cap T\backslash(S \cap T))$ have the same parity, $\#(F \cap (S \cap T))$, $\#(F \cap S\backslash(S \cap T))$ have different parity, and the parity of $\#([m]\backslash(S \cup T))$ doesn't matter. There are $2^{\#S\cap T}$ ways to choose a set F_1 in $S \cap T$. There are $2^{(\#T-\#S\cap T)-1}$ to choose a set F_2 of the same parity as F_1 in $T\backslash(S \cap T)$. There are $2^{(\#S-\#S\cap T)-1}$ to choose a set F_3 of different parity to F_1 in $S\backslash(S\cap T)$. There are $2^{m-\#S\cup T}$ ways to pick a set F_4 in $[m]\backslash(S\cup T)$. Therefore $F = F_1 \sqcup F_2 \sqcup F_3 \sqcup F_4$ can be chosen in

$$2^{\#S\cap T}2^{(\#T-\#S\cap T)-1}2^{(\#S-\#S\cap T)-1}2^{m-\#S-\#T-\#(S\cap T)} = 2^{m-2}$$

ways, as needed.

Lastly, such an $F \in \Delta - \{\emptyset\}$, where $\Delta = 2^{[m]} - [m]$. If $F = [m]$, one has $T = T \cap [m]$, and $\#T$ must be even, in contradiction to the assumption that T has odd cardinality. If $F = \{\emptyset\}$, then $\{\emptyset\} = \{\emptyset\} \cap S$ is odd, in contradiction to the parity of $\{\emptyset\}$ being even. $\square$

Lemma 5.4.5. *Let $T \subseteq [m]$ have odd cardinality. Let $S_1, S_2 \in \Delta$ such that $\#(S_1 \cap T)$ and $\#(S_2 \cap T)$ are even and $S_1 \neq S_2$. Then the switched linear functionals $\mathbf{a}^T(S_1)$ and $\mathbf{a}^T(S_2)$ are not equal.*

Proof of Theorem 2.8. . By [24, Theorem 2.8], the given binary marginal polytope has 4^{m-1} facets. Lemma 5.4.4 provides 2^{m-1} distinct half-spaces with positive coefficients, one for each set in $[m]$ of odd size. By Lemma 5.4.5, each of these half-spaces induces 2^{m-1} other distinct half-spaces via the switch operators with respect to sets in $[m]$ that intersect evenly with the chosen set of odd size. The total number of these half-spaces is $2^{m-1}2^{m-2} = 4^{m-1}$, which ends the proof. $\square$

Example 5.4.6. *The generalized cut polytope for $\Delta = 2^{[3]} - [3]$ has as vertices the columns of matrix A in Example 5.2.7, which coincides with matrix A in Example 2.6.2. A Gale transformation for it is $\mathcal{B} = \{\mathbf{b}_S \mid S \subseteq [3]\} \subseteq \mathbb{R}$, with*

$$\mathbf{b}_{\{\emptyset\}} = \mathbf{b}_{\{1,2\}} = \mathbf{b}_{\{1,3\}} = \mathbf{b}_{\{2,3\}} = 1,\ \mathbf{b}_{\{1\}} = \mathbf{b}_{\{2\}} = \mathbf{b}_{\{3\}} = \mathbf{b}_{\{1,2,3\}} = -1.$$

The facets of the $gcutP_\Delta$ are $P_{E,T} = conv(\mathbf{d}^S \mid S \notin \{E, T\})$, with $E \subseteq [3]$ of even size, and with $T \subseteq [3]$ of odd size. The valid facet inequalities from Lemma 5.4.4 are

$$\{1\} \to x_{\{1\}} + x_{\{1,2\}} + x_{\{1,3\}} \leq 2,$$

$$\{2\} \to x_{\{2\}} + x_{\{1,2\}} + x_{\{2,3\}} \leq 2,$$

$$\{3\} \to x_{\{3\}} + x_{\{1,3\}} + x_{\{2,3\}} \leq 2,$$

$$\{1,2,3\} \to x_{\{1\}} + x_{\{2\}} + x_{\{3\}} \leq 2.$$

Applying the switch operators with respect to $\mathcal{E}_{\{1\}} = \{\{\emptyset\}, \{2\}, \{3\}, \{2,3\}\}$ to the inequality corresponding to $\{1\}$, one obtains the new valid facet inequalities

$$\{2\} \to x_{\{1\}} - x_{\{1,2\}} + x_{\{1,3\}} \leq 0,$$

$$\{3\} \to x_{\{1\}} + x_{\{1,2\}} - x_{\{1,3\}} \leq 0,$$

$$\{2,3\} \to x_{\{1\}} - x_{\{1,2\}} - x_{\{1,3\}} \leq 0.$$

Similarly, one obtains the rest of inequalities

$$x_{\{2\}} - x_{\{1,2\}} + x_{\{2,3\}} \leq 0,\ x_{\{2\}} + x_{\{1,2\}} - x_{\{2,3\}} \leq 0,\ x_{\{2\}} - x_{\{1,2\}} - x_{\{2,3\}} \leq 0,$$

$$x_{\{3\}} - x_{\{1,3\}} + x_{\{2,3\}} \leq 0,\ x_{\{3\}} + x_{\{1,3\}} - x_{\{2,3\}} \leq 0,\ x_{\{3\}} - x_{\{1,3\}} - x_{\{2,3\}} \leq 0,$$
$$-x_{\{1\}} - x_{\{2\}} + x_{\{3\}} \leq 0,\ -x_{\{1\}} + x_{\{2\}} - x_{\{3\}} \leq 0,\ x_{\{1\}} - x_{\{2\}} - x_{\{3\}} \leq 0.$$

One uses transformations in the proofs of Proposition 5.2.8 and Proposition 5.2.10 to conclude half-space descriptions for the corresponding correlation and marginal polytopes.

We end this chapter with the hopeful note that the generalized cut polytopes can provide more results regarding the polyhedral geometry of the binary marginal polytopes. For example, combining results in papers [5, 6] with the polyhedral geometry of the generalized cut polytopes, one obtains advances in the half-space description of binary marginal polytopes for the unimodular hierarchical models. On the other hand, since the generalized cut polytopes are generalizations of the cut polytope, it is worth exploring their place in optimization problems, and other possible applications.

Appendices

Appendix A: The Hilbert Series of Hierarchical Models

This appendix describes the rational form of the Hilbert for ideals of hierarchical models.

Given $\mathcal{M}(\Delta, \mathbf{r})$ a hierarchical model, denote f_i the product $\prod_{i \in F_i} r_i$. If $\text{facet}(\Delta) = \{F_1\}$, then $I_{\mathcal{M}(\Delta,\mathbf{r})}$ is the zero ideal. Hence, the Hilbert series for $R_{\mathbf{r}}/I_{\mathcal{M}(\Delta,\mathbf{r})}$ is

$$H_{R_{\mathbf{r}}/I_{\mathcal{M}(\Delta,\mathbf{r})}}(t) = \frac{1}{(1-t)^{f_1}}.$$

If $\text{facet}(\Delta) = \{F_1, F_2\}$, and $F_1 \cap F_2 = \emptyset$, the image of $\Phi_{\mathcal{M}}$ is isomorphic to the coordinate ring of the Segre product $\mathbb{P}^{f_1-1} \times \mathbb{P}^{f_2-1}$, whose homogeneous ideal is $I_{\mathcal{M}}$ (see Example 2.3.2). Using [7], and assuming that $f_1 \leq f_2$, one has the following closed formula for the Hilbert series of $I_{\mathcal{M}(\Delta,\mathbf{r})}$

$$H_{R_{\mathbf{r}}/I_{\mathcal{M}(\Delta,\mathbf{r})}}(t) = \frac{\sum_{i=0}^{f_1-1} \binom{f_1-1}{i}\binom{f_2-1}{i} t^i}{(1-t)^{f_1+f_2-1}}. \tag{.0.2}$$

If $\text{facet}(\Delta) = \{F_1, F_2\}$, and $F_1 \cap F_2 \neq \emptyset$, let $f = \prod_{i \in F_1 \cap F_2} r_i$, and denote I the ideal for the hierarchical model $\mathcal{M}(\{F_1 \backslash F_2, F_2 \backslash F_1\}, \mathbf{r}_{F_1 \triangle F_2})$ in $R = R_{\mathbf{r}_{F_1 \triangle F_2}}$. Given any $\mathbf{c} \in [\mathbf{r}_{F_1 \cap F_2}]$, let $I_{\mathbf{c}}$ be the ideal I embedded naturally in $R_{\mathbf{c}} = \mathbb{K}[x_{\mathbf{i}} \mid \mathbf{i} \in [\mathbf{r}], \mathbf{i}_{F_1 \cap F_2} = \mathbf{c}]$. One has that all ideals $I_{\mathbf{c}} \subseteq R_{\mathbf{c}} \subseteq R_{\mathbf{r}}$ are disjoint, i.e. $I_{\mathbf{c}} \cap I_{\mathbf{c}'} = \{0\}$ and

$$R_{\mathbf{r}}/I_{\mathcal{M}} = R_{\mathbf{r}}/(\sum_{\mathbf{c}} I_{\mathbf{c}}) \cong \bigotimes_{\mathbf{c}} R_{\mathbf{c}}/I_{\mathbf{c}}.$$

This, and Equation (.0.2) on the disjoint union of two facets, induce the following result.

Proposition .0.7. *Let* $\text{facet}(\Delta) = \{F_1, F_2\}$. *The Hilbert series of ideal for the hierarchical model* $\mathcal{M}(\Delta, \mathbf{r})$ *has the following reduced rational form:*

$$H_{R_{\mathbf{r}}/I_{\mathcal{M}(\Delta,\mathbf{r})}}(t) = \left[H_{R/I}(t)\right]^f = \frac{\left[\sum_{i=0}^{f_1-f-1} \binom{f_1-f-1}{i}\binom{f_2-f-1}{i} t^i\right]^f}{(1-t)^{f_1+f_2-f}}.$$

Observe that $f = 1$ corresponds to the case $F_1 \cap F_2 = \emptyset$. This makes sense since the quotient ring $R_{\mathbf{r}}/I_{\mathcal{M}(\Delta,\mathbf{r})}$ is isomorphic to the quotient ring obtained by the model $\mathcal{M}(\Delta, \mathbf{r})$ restricted to the vertex set $V = \{i \in [m] \mid r_1 > 1\}$.

Next we propose a way to obtain a recursive formula with respect to $\mathbf{r}$ on the Hilbert series of decomposable hierarchical models. In cases, this method provides explicit rational presentations. First, we need an introduction to edge ideals.

Given a simple undirected graph $G([n], E)$ on the set of vertices $[n]$ and edge set E, the edge ideal I_G for it is a square-free monomial ideal in $R = \mathbb{K}[x_1, \ldots, x_n]$, generated by a set of quadratics induced by the edges of the graph G. More specifically, $x_i x_j \in I_G$ if and only if $\{i, j\} \in E$ (see [46]). For a variable $k \in [n]$, denote the link of k in G by $\text{lk}_G(k) = \{\{i, k\} \in E \mid i \in [n]\}$. The colon ideal $(I : x_k) = \{a \in R \mid ax_k \in I\}$ is an edge ideal. The graph associated to it is a graph on vertex set $[n]$ and edges $E - \bigcup\limits_{\{i,k\} \in E} \text{lk}(i)$. The ideal $(I, x_k) = \{a + bx_k \mid a \in I, b \in R\}$ is an edge ideal. The graph associated to it is a graph on vertex set $[n]$ and edges $= E - \text{lk}_G(k)$. The variable x_k induces the following short exact sequence

$$0 \; \to \; R/(I : x_k)(-1) \; \xrightarrow{\cdot x_k} \; R/I \; \to \; R/(I, x_k) \; \to \; 0, \tag{.0.3}$$

which produces the equation

$$H_{R/I}(t) = tH_{R/(I:x_k)}(t) + H_{R/(I,x_k)}(t). \tag{.0.4}$$

If the quotient rings $R/I, R/(I : x_k), R/(I, x_k)$ are all Cohen-Macaulay, have the same dimension, and $g(t)$, $g_1(t)$, and $g_2(t)$ are the numerators in the reduced rational form of their Hilbert series, one has the following relation

$$g(t) = tg_1(t) + g_2(t). \tag{.0.5}$$

See [34, Remark 2.4(iii)] as a reference. Back to hierarchical models, given Δ a simplicial complex, denote $g_{\mathbf{r}}(t)$ the numerator in the reduced form of the Hilbert series of $\mathcal{M}(\Delta, \mathbf{r})$. Decomposable hierarchical models have square free quadratic Gröbner bases (see Theorem 3.2.5). This means that their initial ideals are edge ideals. Lastly, since the Hilbert series of an ideal equals to the Hilbert series of its initial ideal, we can study the Hilbert series of the edge ideals of decomposable hierarchical models.

Proposition .0.8. *(i) Assume Δ has k pairwise disjoint facets. Assuming $r_m > 1$, the Hilbert series for $I_{\mathcal{M}(\Delta,\mathbf{r})}$ is*

$$H_{R_{\mathbf{r}}/I_{\mathcal{M}(\Delta,\mathbf{r})}}(t) = \frac{g_{\mathbf{r}}(t)}{(1-t)^{f_1 + \cdots f_k - k + 1}},$$

where

$$g_{\mathbf{r}}(t) = t \sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{r}-\mathbf{1}} g_{\mathbf{i}_{[m-1]}-\mathbf{1},1}(t) g_{\mathbf{r}_{[m-1]}-\mathbf{i}_{[m-1]},r_m-1}(t) + (1-t) g_{\mathbf{r}_{[m-1]},r_m-1}(t).$$

(ii) *The ideal for the hierarchical model with facets facet*$(\Delta) = \{\{1,2\},\{2,4\},\{3,4\}\}$ *and* $\mathbf{r} = (2,2,2,n+1)$ *has*

$$H_{R_{\mathbf{r}}/I_{\mathcal{M}(\Delta,\mathbf{r})}}(t) = \frac{g_{n+1}(t)}{(1-t)^{3n-5}},$$

where

$$g_{n+1}(t) = (1+t)g_n(t)+t^2(1+t)^n(3+t)+(2t^2+4t)(1+t)^{n-1}((n-1)t^2+2nt+1).$$

Proof. Let $I^0_{\mathcal{M}}$ be the initial ideal for $I_{\mathcal{M}(\Delta,\mathbf{r})}$. Both parts of the proposition use the same technique. Start with the Equation (.0.3) for $I^0_{\mathcal{M}}$ and the variable $x_{\mathbf{r}}$. Equation (.0.4) transfers the problem to the quotient rings $R_{\mathbf{r}}/(I^0_{\mathcal{M}} : x_{\mathbf{r}})$ and $R_{\mathbf{r}}/(I^0_{\mathcal{M}}, x_{\mathbf{r}})$. Apply Equation (.0.3) to both these new rings, with respect to $x_{\mathbf{r}_{[m-1],r_m-1}}$. Apply again this procedure to the new quotient rings, using $x_{\mathbf{r}_{[m-1],r_m-2}}$, and so on. We stop when we have used all $x_{\mathbf{r}_{[m-1],i}}$, $i \in [r_m]$. The resulting equations are the recursive formulas in this proposition. □

Corollary .0.9. *Proposition .0.8(i) provides another way to prove Equation* (.0.2). *One needs to continue the same technique further. In an analogous way, Proposition .0.8(i) produces the rational form*

$$H_{R_{\mathbf{r}}/I_{\mathcal{M}(\Delta,\mathbf{r})}}(t) = \frac{1+(3n-2)t+(n-1)^2t^2}{(1-t)^{n+2}},$$

for the models with facet$(\Delta) = \{\{1\},\{2\},\{3\}\}$ *and* $r = (2,2,n)$.

Remark .0.10. *Given a hierarchical model* $\mathcal{M}(\Delta,\mathbf{r})$*, let* $h_{\mathcal{M}}(\mathbf{d}) = \dim_{\mathbb{K}}[R/I_{\mathcal{M}}]_{\mathbf{d}}$ *for some multi-degree vector* $\mathbf{d}$*, be the multi-graded Hilbert function evaluated at* $\mathbf{d}$*. For a reducible hierarchical model* $\mathcal{M}$ *with reduced models* $\mathcal{M}_1$ *and* $\mathcal{M}_2$ *as in Remark 3.2.6, [40, Corollary 2.12] produces that*

$$h_{\mathcal{M}}(d) = h_{\mathcal{M}_1}(d) \cdot h_{\mathcal{M}_2}(d).$$

Its multi-graded Hilbert series is the Hallamard product of the multigraded Hilbert series of the reduced parts. In the case of decomposable hierarchical models, the formula above turns in product of multigraded Hilbert series for polynomial rings.

Appendix B: Macaulay2 Code

Here we describe implementations of algorithms developed in this dissertation, using the computer algebra software `Macaulay2`.

The following function serves to construct the toric ideal of a hierarchical model. The function takes as inputs the vector of states `r` and the collection facets of Δ denoted `Facets`, The output is the toric ideal $I_{\mathcal{M}(\Delta,r)}$ and the transformation matrix A.

```
makeModel = (r,Facets)->(
R=QQ[x_(splice{#r:0})..x_r]; -- the source ring
S=QQ; --initiate S to be the ring QQ and use iteration to construct
   ↪ the target ring
for i from 0 to (#Facets-1) do (S= tensor (S, QQ[y_(splice{#Facets_i
   ↪ :0,i})..y_(append (r_(Facets_i),i))]));
listOfImages={};
apply(flatten entries vars R, j->(c=toString(j);
    c=substring(2,#c-2,c);
        alpha = value c;
    accumulateMonomial = 1;
    for i from 0 to (#Facets-1)do (accumulateMonomial=
       ↪ accumulateMonomial*y_(append(alpha_(Facets_i),i)));
    listOfImages=append (listOfImages,accumulateMonomial););
PSI=map(S,R,listOfImages); -- maps from R to S where the i-th variable
   ↪  goes to the i-th monomial in the list of images.
B=mutableMatrix (ZZ, numgens S, numgens R);
for j from 0 to ( #listOfImages-1) do(
for i from 0 to (numgens S-1) do(if gcd(S_i,listOfImages_j)!=1 then B_
   ↪ (i,j)=1));
A= matrix entries B;
I=ker PSI;
return (I,A))

TEST
r={1,1,1} -- the vector of states
F1={0,1} --give the facet 1
F2={1,2} --give the facet 2
Facets = {F1,F2} --collection of facets
makeModel(r,Facets)
```

The function `equivariantHS(I,R)` computes the rational form of the equivariant Hilbert series for fitrations of ideals from hierarchical models $\mathscr{I}_{\Delta,\mathbf{r}}$ with $\#T = 1$ (see Proposition 4.1.2). This function takes as input a finitely generated ideal `I` and a ring $\mathtt{R} = \mathtt{QQ}[\mathtt{x}_{1..1} \mathinner{..} \mathtt{x}_{\mathtt{r}}]$ that contains the ideal `I`. Any ideal I_{r_0} with $S_r(I_{r_0}) = I_r$, for any $r \geq r_0$ can be inputed as the ideal I. One can use the previous function `makeModel` to compute I_{r_0}. The function `equivariantHS(I,R)` adapts I to an isomorphic ideal in a ring of the form $K[X_{[c]\times\mathbb{N}}]$. From here, we use functions in `http://rckr.one/eHilbert.m2` written by Krone, Leykin, and Snowden to compute the equivariant Hilbert series of the filtration.

```
loadPackage"EquivariantGB"
```

```
equivariantHS=(I,R)->( --EMBED THE IDEAL INTO THE INFINITE POLY. RING
Indices={};
apply(flatten entries vars R, j->(c=toString(j);
        c=substring(2,#c-2,c);
    alpha=value c;
    Indices=append(Indices, alpha)));
ListOfVariables={};
apply(Indices,i->(
    c=concatenate{toString X, toList apply(#r-1, j-> toString i_j)};
alpha= value c;
ListOfVariables=append(ListOfVariables,alpha)));
L=unique ListOfVariables; -- create the variables
ER = buildERing(L,toList(#L:1),QQ, r_(#r-1)+1); -- create the infinite
   ↪  ring with the variables above
eR= gens ER;
g=map (ER,R,toList apply (#gens R ,i-> eR_i)); -- inject the finite
   ↪ ring to the infinite ring
J=g(I);--inject the ideal to the infinite ring
A = idealAutomaton flatten{J_*}; --create the automaton from the ideal
   ↪  J
--CALCULATE THE EQUIVARIANT HILBERT SERIES
T=frac(QQ[s,t]);--the world of equivariant Hilbert Series
W={};
apply(#gens R, i->(t;
   W=prepend(t,W)));
W=flatten{s,W}; --the vector of weights for the letters of alphabet
h = automatonHS(A,W); -- equivariant Hilbert series for I
H=1/((1-t)^(#L)-s)-h; -- equivariant Hilbert series for R/I
return (H))

--
TEST
r={2,1} -- the vector of states
Facets = {{0},{1}} --collection of facets
makeModel(r,Facets)
equivariantHS(I,R)
--
r={1,1,1} -- the vector of states
F1={0,1} --give the facet 1
F2={1,2} --give the facet 2
Facets = {{0,1},[1,2]} --collection of facets
makeModel(r,Facets)
equivariantHS(I,R)
```

Lastly, the function equivH(c) computes the rational form of the equivariant Hilbert series for the generalized independence models with fixed number of states c (see Chapter 4).The function takes as input the vector c. Its output is the rational form of the equivariant Hilbert series for the filtration of ideals arising from hierarchical models with Δ as in Proposition 4.2.4. The function encodes the formula presented at the end of Chapter 4, with some simplifications done beforehand to the automaton. More specifically, we reduce the automata construceted in Proposition 4.3.7 by minimizing the number states with the technique in [27, Theorem 4.26].

```
equivH=(c)->(
m=1;
for i from 0 to #c-1 do m=m*c_i;--m is the number of ZETA letters
L=splice{#c:1}..c; --(#l=M) the vectors associated to each letter ZETA
n=0;
d={};
for i from 0 to #c-1 do(d=drop(c,{i,i});
    p=1;
    for j from 0 to #d-1 do(p=p*d_j);
    n=n+p); --n is the number of acceptable states represented by TAU
M=mutableMatrix(ZZ,#c,#c);
for j from 0 to #c-1 do( for i from 0 to #c-1 do( (M_(i,j)=c_j)));
for i from 0 to #c-1 do(M_(i,i)=1);
 M=entries matrix M;
I={};
for i from 0 to #c-1 do(Q={splice{#M_i:1}..M_i};
    I= append(I,Q));
I=flatten I; -- records the vectors associated to TAU letters/states
J=apply(#I,i->#I_i); --records how many times one has states with the
    ↪ TAU_i loops
b=0;
h=mutableMatrix(ZZ,#J,1);
for i from 0 to #J-1 do( b=b+J_i;
    h_(i,0)=b);
h=matrix entries h;
h=0||h;
H=apply(#J,i->h_(i,0)); -- records the sums of the first i-th terms of
    ↪  J
R=QQ[t,s_1..s_(#c)]; --#c is the a number of words TAU for shifts, t
    ↪ is the weight for ZETA-s
          -- and s_i is the weight for TAU_i-s
S=frac R; -- the fraction field where the equivariant Hilbert series
--Below is the matrix that records the relationships in the automaton
--FIRST PART-- relations in automaton from ZETA to ZETA
L=splice{#c:1}..c;
```

```
A=mutableMatrix (S, #L,#L);
N=mutableMatrix (ZZ,#c,1);
for k from 0 to #c-1 do N_(k,0)=1;
N; -- the vector with all 1 and length #c.
for j from 0 to #L-1 do( for i from 0 to #L-1 do( a=mutableMatrix(ZZ,#
    ↪ c,1);
for k from 0 to #c-1 do(if (L_j)_k<=(L_i)_k then a_(k,0)=1);
if a==N then A_(i,j)=t));
A= matrix entries A;
--SECOND PART--relations in automata from TAU to ZETA
B=mutableMatrix(S,#L,n);
for i from 0 to #I-1 do( for e from 0 to #I_i-1 do( for j from 0 to #L
    ↪ -1 do( if drop(I_i_e,{i,i})<=drop(L_j,{i,i}) then B_(j,H_i+e)=t
    ↪ )));
B=matrix entries B; --
--THIRD PART--relations in automata from ZETA to TAU
C=mutableMatrix(S,n,#L);
for i from 0 to #I-1 do( for e from 0 to #I_i-1 do( for j from 0 to #L
    ↪ -1 do( if drop(I_i_e,{i,i})==drop(L_j,{i,i}) then C_(H_i+e,j)=
    ↪ s_(i+1))));
C=matrix entries C;
--FOURTH PART-- relations in automata from TAU to TAU
D=mutableMatrix(S,n,n);
for i from 0 to #I-1 do( for j from 0 to #I-1 do(for e from 0 to #(I_i
    ↪ )-1 do( for l from 0 to #(I_j)-1 do( if i<=j and I_j_l_i==1 and
    ↪  drop(I_i_e,{j,j})==drop(I_j_l,{j,j}) then D_(H_j+l,H_i+e)=s_(j
    ↪ +1)))));
D=matrix entries D;
M=(A|B)||(C|D); -- the square matrix describing all the relations
k=m+n; -- the size of the matrix
M=id_(S^k)-M; -- the matrix that we want to use the inverse of
D=determinant M;
N = inverse M;
v=mutableMatrix(S,1,k);
for i from 0 to k-1 do(v_(0,i)=1);
v=matrix entries v; -- the vector that records the accepting states
e=mutableMatrix(S,k,1);
for i from 0 to k-1 do(e_(0,0)=1);
e= matrix entries e; -- the unit vector
g=v*N*e;
m=1;
for i from 0 to #c-1 do m=m*s_(i+1);
m;
h=m*determinant g; -- the equivariant Hilbert series
return (h))
```

```
TEST
c={1,1,1}
equivH(c)

c={2,2}
equivH(c)
```

Bibliography

[1] S. Aoki, H. Hara, and A. Takemura, *Markov bases in algebraic statistics*, Springer Series in Statistics, Springer, New York, 2012.

[2] S. Aoki and A. Takemura, *Minimal basis for connected Markov chain over* $3 \times 3 \times K$ *contingency tables with two-dimensional marginals*, Australian and New Zealand Journal of Statistics, 229–249, 2003.

[3] E. Belay, J. Bresee and +4, *Reye's Syndrome in the United States from 1981 through 1997*, The New Eng. J. of Med. **340**, 1377–1382, 1999.

[4] W. Bruns and J. Herzog, *Cohen-Macaulay rings*, Revised edition, Cambridge Studies in Advanced Mathematics **39**, Cambridge University Press, 1998.

[5] D. Bernstein, and C. O'Neill. *Unimodular hierarchical models and their Graver bases*, Journal of Algebraic Statistics **8.2**, 2017.

[6] D. Bernstein, and S. Sullivant, *Unimodular binary hierarchical models*, Journal of Combinatorial Theory, Series B, **123**, 97-125, 2015.

[7] A. Conca, J. Herzog, *On the Hilbert function of determinantal rings and their canonical module*, Proc. Amer. Math. Soc. **112**, 677–681, 1994.

[8] J. Coons, J. Cummings, B. Hollering, and A. Maraj, *Polyhedral geometry of generalized cut polytopes*, in progress.

[9] D. Cox,, J Little, and H Schenck. *Toric varieties*, American Mathematical Soc., 2011.

[10] A. Dobra, *Markov bases for decomposable graphical models*, Bernoulli **9**, 1093–1108, 2003.

[11] M. Deza and M. Laurent, *Geometry of cuts and metrics*, Algorithms and Combinatorics, **15**, Springer-Verlag, Berlin, 1997.

[12] J. Draisma, R. H. Eggermont, R. Krone, and A. Leykin, *Noetherianity for infinite-dimensional toric varieties*, Algebra Number Theory **9**, 1857–1880, 2015.

[13] P. Diaconis, A. Gangolli, *Rectangular arrays with fixed margins*, In: Discrete probability and algorithms (Minneapolis, MN, 1993), IMA Vol. Math. Appl. **72**, Springer, New York, 15–41, 1995.

[14] M. Deza and M. Laurent, *Applications of cut polyhedra—II*, Journal of Computational and Applied Mathematics **55**, 217-247, 1994.

[15] J. De Loera, S. Onn, *Markov bases of three-way tables are arbitrarily complicated*, Journal of Symbolic Computation **41**, 173–181, 2006.

[16] P. Diaconis and B. Sturmfels, *Algebraic Algorithms for sampling from conditional distributions*, Ann. Statist. **26**, 1998.

[17] A. Dobra and S. Sullivant, *A divide-and-conquer algorithm for generating Markov bases of multi-way tables*, Comput. Statist. 19, no. 3, 347–366, 2004.

[18] M. Drton, and B. Sturmfels, and S. Sullivant, *Lectures on algebraic statistics*, Algebra Number Theory **39**, Springer Science & Business Media, 2008.

[19] D. Grayson, M. Stillman, *Macaulay2*, a software system for research in algebraic geometry; available at http://www.math.uiuc.edu/Macaulay2/.

[20] B. Grunbaum, *Convex Polytopes*, Interscience, London, (1967); revised edition, Graduate Texts in Math., Springer-Verlag, 2003.

[21] R. Hemmecke and R. Hemmecke, *4ti2 - Software for computation of Hilbert bases, Graver bases, toric Gröbner bases, and more,* Available at http://www.4ti2.de, 2003.

[22] C. Hillar and S. Sullivant, *Finite Gröbner bases in infinite dimensional polynomial rings and applications*, Adv. Math. **229**, 1–25, 2012.

[23] S. Hoşten and S. Sullivant, *A finiteness theorem for Markov bases of hierarchical models*, J. Combin. Theory Ser. A **114**, 311–321, 2007.

[24] S.Hoşten and S. Sullivant, *Gröbner bases and polyhedral geometry of reducible and cyclic models*, J. Combin. Theory, **100**, 277–301, 2002.

[25] J. Honkala, *A necessary condition for the rationality of the zeta function of a regular language*, Theor. Comput. Sci. **66**, 341–347, 1989.

[26] J Herzog, H Takayuki, and O Hidefumi. *Binomial ideals.* **279**, Springer, 2018.

[27] J. E. Hopcroft and J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation*, Addison-Wesley Series in Computer Science, Addison-Wesley Publishing Co., Reading, Mass., 1979.

[28] R. Krone, A. Leykin and A. Snowden, *Hilbert series of symmetric ideals in infinite polynomial rings via formal languages*, J. Algebra **485**, 353–362, 2017.

[29] L. Li and N. Yu, FI^m*-modules over Noetherian rings*, J. Pure Appl. Algebra **223**, 3436–3460, 2019.

[30] A. Maraj and U. Nagel, *Equivariant Hilbert series for hierarchical modes*, arxiv: 1909.13026, 2019.

[31] E. Miller and B. Sturmfels, *Combinatorial Commutative Algebra*, Graduate Texts in Mathematics, **227**, Springer-Verlag, New York, 2004.

[32] U. Nagel and T. Römer, *Equivariant Hilbert series in non-Noetherian Polynomial Rings*, J. Algebra **486**, 204–245, 2017.

[33] U. Nagel and T. Römer, *FI- and OI-modules with varying coefficients*, J. Algebra **535**, 286–322, 2019.

[34] U. Nagel, and T. Römer, *Glicci simplicial complexes.* Journal of Pure and Applied Algebra **212(10)**, 2250-2258, 2008.

[35] I. Pitowsky, *Correlation polytopes: their geometry and complexity*, Mathematical Programming **50**, 395-414, 1991.

[36] J. Rauh and S. Sullivant, *Lifting Markov bases and higher codimension toric fiber products*, J. Symbolic Comput. **74**, 276–307, 2016.

[37] F. Santos and B. Sturmfels, *Higher Lawrence configurations*, J.Combin. Theory Ser. A **103**, no. 1, 151–164, 2003.

[38] A. Snowden, *Syzygies of Segre embeddings and* Δ*-modules*, Duke Math. J. **162**, 225-277, 2013.

[39] R. Stanley, *Enumerative Combinatorics, Volume 1*, second edition, Cambridge Studies in Advanced Mathematics **49**, Cambridge University Press, Cambridge, Cambridge, 2012.

[40] B. Sturmfels, *Gröbner Bases and Convex Polytopes*, University Lecture Series **8** American Mathematical Society, Providence, RI, 1996.

[41] B. Sturmfels and S. Sullivant, *Toric geometry of cuts and splits*, Michigan Math. J. **57**, 689–709, Special volume in honor of Melvin Hochster, 2008.

[42] S. Sullivant, *Algebraic Statistics*, Graduate Studies in Mathematics **194** American Mathematical Society, Providence, RI, 2018.

[43] S. Sullivant, *Toric fiber products*, J. Algebra 316, no. 2, 560–577, 2007.

[44] S. Sullivant, *Normal binary graph models*, Ann. Inst. Statist. Math. **62**, no. 4, 717–726, 2010.

[45] A. Takken, *Monte Carlo goodness-of-fit tests for discrete data*, Ph.D. thesis, Stanford University, 1999.

[46] R. Villarreal, *Cohen-macaulay graphs.* manuscripta mathematica 66, **1** 277-293, 1990.

[47] G. M. Ziegler, *Lectures on Polytopes*, Graduate Texts in Mathematics, **152**, Springer-Verlag, New York, 1995.

Vita

Aida Maraj

Education:

- Ph.D in Mathematics, University of Kentucky, Fall 2015–Spring 2020.
- M.Sc in Mathematics, University of Tirana, Albania, Fall 2012–Spring 2014.
- B.S. in Mathematics, University of Tirana, Albania, Fall 2009–Spring 2012.

Professional Positions:

- Postdoctoral Assistant Professor, University of Michigan–Ann Arbor, Fall 2021–
- Postdoctoral Position, Max Plank Institute for Mathematics in the Sciences–Leipzig, Fall 2020–Spring 2021.
- Teaching Assistant, University of Kentucky, Fall 2015–Spring 2020.
- Full Time Faculty Position, University of Vlora, Albania, Fall 2014–Spring 2015.

Honors

- Clifford J. Swauger, Jr. Graduate Fellowship, University of Kentucky, Summer 2019.
- Arts and Sciences Outstanding Teaching Award, University of Kentucky, Spring 2019.
- Edgar Enochs Departmental Award in Algebra, University of Kentucky, Spring 2018.
- Summer Research Support Departmental Award, University of Kentucky, Summer 2018.
- Outstanding Undergraduate Student Scholarship, University of Tirana, Fall 2010–Spring 2012.

Publications & Preprints:

- J. Coons, J. Cummings, B. Hollering, and A. Maraj, *Polyhedral geometry of generalized cut polytopes, in preparation.*
- A. Maraj and U. Nagel, *Equivariant Hilbert Series for Hierarchical Models.* arXiv:1909.13026, *submitted to the Algebraic Statistics journal.*
- A. Maraj, U. Nagel, *Equivariant Hilbert series for some families of related algebras, and their Segre product. in preparation.*