

ABSTRACT

COONS, JANE IVY. Applications of Toric Geometry in Algebraic Statistics. (Under the direction of Seth Sullivant.)

In the field of algebraic statistics, we use tools from algebra, geometry and combinatorics to answer questions about statistical models viewed as algebraic varieties. In the present thesis, we focus on two specific statistical models both of whose associated varieties are toric.

The first models that we discuss are two-way quasi-independence models, or independence models with structural zeros. We classify the two-way quasi-independence models that have rational maximum likelihood estimators, or MLEs. We give a necessary and sufficient condition on the bipartite graph associated to the model for the MLE to be rational. In this case, we give an explicit formula for the MLE in terms of combinatorial features of this graph. We also use the Horn uniformization to show that for general log-linear models \mathcal{M} with rational MLE, any model obtained by restricting to a face of the cone of sufficient statistics of \mathcal{M} also has rational MLE.

Next, we discuss the toric geometry of the Cavender-Farris-Neyman model with a molecular clock, or CFN-MC model. We give a combinatorial description of the toric ideal of invariants of the CFN-MC model on a rooted binary phylogenetic tree and prove results about the polytope associated to this toric ideal. Key results about the polyhedral structure include that the number of vertices of this polytope is a Fibonacci number, the facets of the polytope can be described using the combinatorial "cluster" structure of the underlying rooted tree, and the volume is equal to an Euler zig-zag number. The toric ideal of invariants of the CFN-MC model has a quadratic Gröbner basis with squarefree initial terms. We show that the Ehrhart polynomial of these polytopes, and therefore the Hilbert series of the ideals, depends only on the number of leaves of the underlying binary tree, and not on the topology of the tree itself. We give a formula for the numerator of the Ehrhart series of these polytopes using the combinatorics of alternating permutations. These results are analogous to classic results for the Cavender-Farris-Neyman model without a molecular clock. However, new techniques are required because the molecular clock assumption destroys the toric fiber product structure that governs group-based models without the molecular clock.

© Copyright 2021 by Jane Ivy Coons

All Rights Reserved

Applications of Toric Geometry in Algebraic Statistics

by
Jane Ivy Coons

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Mathematics

Raleigh, North Carolina

2021

APPROVED BY:

Nathan Reading

Agnes Szanto

Cynthia Vinzant

Seth Sullivant
Chair of Advisory Committee

DEDICATION

To my parents, Sally and David, and my sister, Rachel.

BIOGRAPHY

Jane Ivy Coons grew up in Great Neck, New York, where she attended Great Neck South High School. She graduated Summa Cum Laude from the State University of New York at Geneseo in May 2015 with majors in Vocal Performance and Mathematics. Jane began her graduate work in mathematics at North Carolina State University in August 2015. In 2017, she was awarded the National Science Foundation Graduate Research Fellowship. Jane will continue her academic career as a Supernumerary Teaching Fellow at St. John's College at the University of Oxford in the fall of 2021. In her spare time, Jane loves to sing in choir, read books, drink coffee, cook new recipes, and participate in social justice activism.

ACKNOWLEDGEMENTS

My success in graduate school was made possible by my fantastic support network of friends, family, professors and colleagues.

I am so grateful to my advisor, Seth Sullivant, who has been a fixture in my academic life since my very first year at NC State. His feedback, encouragement and insight have been fundamental to my success as a graduate student. As my advisor, Seth went above and beyond guiding me to successfully complete my dissertation research. He helped me establish myself as a member of our academic community by encouraging me to attend conferences and workshops. Moreover, he's given me crucial feedback on my academic writing, job applications and conference presentations to ensure that I represent myself and my research in the best possible way. Perhaps most importantly, Seth believed in my abilities as a mathematician even at the times when I did not. Thank you, Seth!

The professors here at NC State have been great resources to me as I completed my coursework and my thesis research. I would especially like to thank my committee members, Gavin Conant, Nathan Reading, Seth Sullivant, Agnes Szanto and Cynthia Vinzant. Thank you for asking great questions and for offering your unique perspectives on my work.

Most people who know me know how much I value the time I spent as an undergraduate at SUNY Geneseo. I would especially like to thank Stacey Edgar, Patrick Rault, Chris Leary and Jeff Johannes, each of whom encouraged me to pursue my Ph.D. in math. I may never have discovered how much I love research if I hadn't participated in the Research Experience for Undergraduates at Winthrop University. I am especially grateful to my REU mentor, Joseph Rusinko.

I feel very fortunate to have brilliant collaborators who I can also call friends. I am especially grateful to Aida Maraj, Ben Hollering, Joe Cummings, Miruna-Stefana Sorea, Pratik Misra, Orlando Marigliano and Michael Ruddy. I am constantly learning from you. When I talk with other graduate students who are trying to decide on a research area, I tell them that the applied algebra community is, hands-down, the best. We have a friendly, welcoming and diverse research community, and I am thankful to all of the mathematicians who have worked hard to make it this way. Thanks especially to my letter-writers, Marta Casanellas, Andrew Cooper, Elizabeth Gross, Nathan Reading, Seth Sullivant and Cynthia Vinzant. I will soon be gainfully employed because of you!

The friendships we make in graduate school are special because of how much we learn and grow together. I've met too many wonderful mathematicians during my time at NC State to name here, but please know I'm grateful to each and every one of you. There are, however, some "math friends" who must be thanked individually. Ben Hollering, Ashley Tharp, Joe Johnson, Pratik Misra, Cash Bortner, Aida Maraj, Stefana Sorea and Katherine Harris all watched me practice for my defense, some more than once. Thank you for your generosity with your time and your invaluable feedback.

To Katie Ahrens and Katherine Harris: I just couldn't have done grad school without you. I'm so proud of how far we've come since our days in the Combinatorics Five; it only took uncountably-many cups of coffee and episodes of Doctor Who. Thank you for being by my side every step of the way. To Aida Maraj: we quickly decided to become collaborators a few years ago at the Triangle

Lectures in Combinatorics, but I didn't realize at the time that I'd also made a dear friend for life. Thank you for your unwavering support. To Justin Harden (an honorary mathematician): thank you for being my number one cheerleader, especially through the trials and tribulations of job applications. Thanks especially for fully organizing and being the sole attendee of my COVID-times defense party.

No acknowledgements section would be complete without a partial list of the best places in Raleigh to get coffee and other study fuel. Thanks to Cup-A-Joe, Global Village, Iris Coffee Lab, Liquid State and, of course, Mitch's. I *really* couldn't have done it without you.

Finally, I am eternally grateful for the love and support of my family. I feel so lucky to belong to a family that loves learning and that never pressured me to be anyone other than who I am. My sister Rachel is my dearest friend, and is always full of insight and encouragement exactly when I need it most. My father is my fan club president, and follows my academic career with remarkable diligence. At this point, he knows the names and research interests of the mathematics faculty at NC State better than I do. My mother has never failed to throw one hundred percent of her support behind my dreams, and has never been anything but proud, even if my plans didn't exactly pan out. When I called my mom after a week in graduate school to tell her how hard my first combinatorics problem set was, she immediately told me I could come home if I wanted to. I didn't end up taking her up on that, and it all seems to have worked out, but I have always been grateful for that little piece of unconditional support. From the bottom of my heart, thank you.

TABLE OF CONTENTS

LIST OF FIGURES	vii
Chapter 1 INTRODUCTION	1
1.1 Ideals and Varieties	1
1.1.1 Parametrized Varieties	5
1.1.2 Toric Varieties	6
1.2 Polytopes	8
1.2.1 Ehrhart Theory	10
1.2.2 Toric Ideals and Polytopes	12
1.3 Algebraic Statistical Models	12
1.3.1 Log-Linear Models	13
1.4 Maximum Likelihood Estimation in Log-Linear Models	14
1.5 Group-Based Phylogenetic Models	17
1.5.1 Preliminaries on Trees	18
1.5.2 The Cavender-Farris-Neyman Model	19
1.6 Summary of Results	21
Chapter 2 FACIAL SUBMODELS AND RATIONAL MAXIMUM LIKELIHOOD ESTIMATES .	23
2.1 Preliminaries	23
2.2 Main Result	24
Chapter 3 QUASI-INDEPENDENCE MODELS WITH RATIONAL MAXIMUM LIKELIHOOD ESTIMATOR	27
3.1 Preliminaries	27
3.2 Quasi-independence Models with Non-Rational MLE	28
3.3 The Clique Formula for the MLE	32
3.4 Intersections of Cliques with a Fixed Column	34
3.5 Checking the Conditions of Birch's Theorem	46
Chapter 4 THE CAVENDER-FARRIS-NEYMAN MODEL WITH A MOLECULAR CLOCK . . .	49
4.1 The Discrete Fourier Transform	49
4.2 The CFN-MC Polytope	57
4.2.1 Toric Fiber Products	68
4.2.2 Cluster Trees	72
Chapter 5 EHRHART THEORY OF THE CFN-MC MODEL	79
5.1 Alternating Permutations and Order Polytopes	79
5.1.1 Alternating Permutations	79
5.1.2 Order Polytopes	80
5.2 Ehrhart Function of the CFN-MC Polytope	82
5.2.1 Caterpillar Trees	82
5.2.2 The Ehrhart Function and Rotations	83
5.3 The h^* -Polynomial of the Order Polytope of the Zig-Zag Poset	96
5.4 The Swap Statistic Via Rank Selection	101
BIBLIOGRAPHY	106

LIST OF FIGURES

Figure 1.1	The polytope described in Example 1.2.5	9
Figure 1.2	A rooted binary tree.	18
Figure 1.3	A labeling of all nodes of tree T with elements of \mathbb{Z}_2	20
Figure 3.1	The double-square graph associated to the matrix in Example 3.2.1	29
Figure 3.2	The bipartite graph associated the matrix S in Example 3.3.3	33
Figure 3.3	The poset $P(j_0)$ for Example 3.4.7	38
Figure 4.1	The tree T referenced in Example 4.1.6	52
Figure 4.2	The four-leaf tree in Example 4.2.7	60
Figure 4.3	The edges and nodes surrounding node r	60
Figure 4.4	An example of a cluster b, c , which are marked with triangles, and the elements of their neighbor set, which are marked with squares.	62
Figure 4.5	The tree T in Example 4.2.12	63
Figure 4.6	The tree construction for T_3 described in Example 4.2.14. "Spine" nodes are marked with a circle.	68
Figure 4.7	Decomposition of T into T' and T'' via a node adjacent to exactly two internal nodes	69
Figure 5.1	The zig-zag poset \mathcal{Z}_4	82
Figure 5.2	A rotation performed by pruning c and its right subtree and reattaching in the right subtree of b	84
Figure 5.3	Nodes a, c, d, e and g are the nodes that are blocked in top-set $\{a, d, e, g\}$	85
Figure 5.4	Proposition 5.2.8 (iii): In this case, \overline{P}_1 contains all dashed edges descended from d . \hat{P}_1 contains all dotted edges descended from e . \overline{P} contains all thick solid edges descended from f . \hat{P}_2 contains all thick solid edges above b	86
Figure 5.5	Proof of Lemma 5.2.12, The first row of trees contain path systems – one whose top-set is b -nonmaintaining, and one without b or c in its top-set, but with d not blocked in its top-set. The second row of trees are the path systems obtained by performing the operation in the proof of Lemma 5.2.12. Note that the representation given by the path systems in the second row is df -compressed.	89
Figure 5.6	Proof of Lemma 5.2.12, The first row of trees contain path systems – one whose top-set is c -nonmaintaining, and one without b or c in its top-set, but with f not blocked in its top-set. The second row of trees are the path systems obtained by performing the operation in the proof of Lemma 5.2.12. Note that the representation given by the path systems in the second row is df -compressed.	90
Figure 5.7	Proof of Lemma 5.2.13. This figure illustrates the case where $\{[\Omega_1], \dots, [\Omega_m]\}$ contains fewer b -nonmaintaining vertices than $\{[\mathfrak{P}_1], \dots, [\mathfrak{P}_m]\}$. The first row of trees are path systems that realize $\mathbf{v} = [\mathfrak{P}_1] + [\mathfrak{P}_2] + [\mathfrak{P}_3] \in 3R_T$. The second row of trees are path systems that realize the vector $\mathbf{u} = [\Omega_1] + [\Omega_2] + [\Omega_3] \in 3R_T$ that satisfies the assumptions of the lemma. The third row of trees are a new set of path systems that realize \mathbf{v} using fewer b -nonmaintaining vertices, which we obtained by applying the procedure discussed in the proof of the lemma.	93

Figure 5.8	Proof of Lemma 5.2.13. $\{[\mathcal{Q}_1], \dots, [\mathcal{Q}_m]\}$ contains fewer c -nonmaintaining vertices than $\{[\mathcal{P}_1], \dots, [\mathcal{P}_m]\}$. The first row of trees are path systems that realize $\mathbf{v} = [\mathcal{P}_1] + [\mathcal{P}_2] + [\mathcal{P}_3] \in 3R_T$. The second row of trees are path systems that realize $\mathbf{u} = [\mathcal{Q}_1] + [\mathcal{Q}_2] + [\mathcal{Q}_3] \in 3R_T$, which satisfies the assumptions of the lemma. The third row of trees are a new set of path systems that realize \mathbf{v} using fewer c -nonmaintaining vertices, which we obtained by applying the procedure discussed in the proof of the lemma. 94
Figure 5.9	The zig-zag poset \mathcal{Z}_7 104

CHAPTER

1

INTRODUCTION

In this chapter, we outline some background material that will be used in the present thesis. We give a brief introduction to ideals, varieties and polytopes. We also introduce the notion of algebraic statistical models and their maximum likelihood estimates. Finally, we discuss group-based phylogenetic models and, in particular, the Cavender-Farris-Neyman model.

1.1 Ideals and Varieties

Let \mathbb{C} denote the field of complex numbers and let

$$\mathbb{C}[\mathbf{x}] := \mathbb{C}[x_1, \dots, x_n]$$

be a polynomial ring in n variables. For each $\mathbf{a} \in \mathbb{N}^n$, we write $\mathbf{x}^{\mathbf{a}} := \prod_{i=1}^n x_i^{a_i}$. All definitions and theorems in this section can be found in [22] unless otherwise cited.

Definition 1.1.1. An *ideal* in $\mathbb{C}[\mathbf{x}]$ is a set of polynomials, $I \subset \mathbb{C}[\mathbf{x}]$ such that

- if $f, g \in I$, then $f + g \in I$, and
- if $f \in I$ and $h \in \mathbb{C}[\mathbf{x}]$, then $fh \in I$.

A set of polynomials $\{f_1, \dots, f_r\} \subset I$ is a *generating set* for I if every element of I has the form

$$\sum_{i=1}^r h_i f_i,$$

for some $h_i \in \mathbb{C}[\mathbf{x}]$. If I has generating set $\{f_1, \dots, f_r\}$, then we write $I = \langle f_1, \dots, f_r \rangle$. The *Hilbert basis theorem* guarantees that every ideal in $\mathbb{C}[\mathbf{x}]$ has a finite generating set. A generating set is *minimal* if no proper subset of it is also a generating set. Suppose that for some $\mathbf{a} \in \mathbb{C}^n$, $f_1(\mathbf{a}) = \dots = f_r(\mathbf{a}) = 0$. Then by definition of an ideal, $f(\mathbf{a}) = 0$ for all $f \in \langle f_1, \dots, f_r \rangle$. Thus, we may associate to each ideal the following geometric object.

Definition 1.1.2. Let $I \subset \mathbb{C}[\mathbf{x}]$. The *affine variety* of I is

$$\{\mathbf{a} \in \mathbb{C}^n \mid f(\mathbf{a}) = 0 \text{ for all } f \in I\}.$$

A set of points in \mathbb{C}^n is *algebraic* if it is an affine variety. Given any $M \subset \mathbb{C}^n$, the *Zariski closure* of M , denoted \overline{M} is the inclusion-smallest algebraic set containing M . The *vanishing ideal* of M is

$$I(M) := \{f \in \mathbb{C}[\mathbf{x}] \mid f(\mathbf{a}) = 0 \text{ for all } \mathbf{a} \in M\}.$$

A polynomial is *homogeneous* if each of its terms has the same degree. An ideal is *homogeneous* if it has a generating set consisting of homogeneous polynomials. In this case, one may also define the projective variety associated to I . Define the equivalence relation \sim on $\mathbb{C}^{n+1} - \{0\}$ by $\mathbf{a} \sim \mathbf{a}'$ if and only if $\mathbf{a} = \lambda \mathbf{a}'$ for some $\lambda \neq 0$. The *n-dimensional complex projective space* \mathbb{P}^n is

$$\mathbb{P}^n := (\mathbb{C}^{n+1} - \{0\}) / \sim.$$

For each $a \in \mathbb{P}^n$ and homogeneous polynomial f , the value $f(a)$ is defined up to a nonzero constant; however, the zero locus of f is well-defined.

Definition 1.1.3. Let $I \subset \mathbb{C}[\mathbf{x}]$ be a homogeneous ideal. The *projective variety* associated to I is

$$V(I) := \{\mathbf{a} \in \mathbb{P}^{n-1} \mid f(\mathbf{a}) = 0 \text{ for all } f \in I\}.$$

All varieties discussed in the following chapters are projective.

Example 1.1.4. Let $R = \mathbb{C}[x_{11}, x_{12}, x_{13}, x_{21}, x_{22}, x_{23}]$. Consider the ideal generated by the 2×2 minors of the generic 2×3 matrix,

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{bmatrix}.$$

That is, $I = \langle x_{11}x_{22} - x_{12}x_{21}, x_{11}x_{23} - x_{21}x_{13}, x_{12}x_{23} - x_{13}x_{22} \rangle$. The affine variety of I is the set of all 2×3 complex matrices all of whose 2×2 minors vanish. This is the set of all 2×3 complex matrices of rank less than or equal to 1. The projective variety, $V(I)$, lies in the projective space \mathbb{P}^5 . The elements of this projective space are equivalence classes of nonzero matrices, where $A \sim B$ if and only if $A = \lambda B$ for some $\lambda \neq 0$. Scaling a matrix by a nonzero constant does not change its rank, so $V(I)$ is the set of all \sim -equivalence classes of matrices whose rank is equal to 1. Note that $V(I)$ does not contain the matrix of all 0s since this matrix is not in \mathbb{P}^5 .

In order to determine if a polynomial f lies in an ideal I , we wish to apply the division algorithm to f using generators of I ; if the remainder of f upon division is zero, then f is in the ideal. However, it is not necessarily the case that for any generating set G of I and any $f \in I$, the remainder of f upon division by G is 0. In fact, one must use a special type of generating set called a *Gröbner basis*; the following definitions prepare us to introduce these.

Definition 1.1.5. A *term order* on $\mathbb{C}[\mathbf{x}]$ is a relation $<$ on the monomials $\mathbf{x}^{\mathbf{a}}$ for $\mathbf{a} \in \mathbb{N}^n$ such that

- $<$ is a total order,
- if $\mathbf{x}^{\mathbf{a}} < \mathbf{x}^{\mathbf{b}}$, then $\mathbf{x}^{\mathbf{a}}\mathbf{x}^{\mathbf{c}} < \mathbf{x}^{\mathbf{b}}\mathbf{x}^{\mathbf{c}}$ for all $\mathbf{c} \in \mathbb{N}^n$, and
- every non-empty set of monomials has a $<$ -smallest element.

Many of the term orders that appear in the present thesis are known as weight orders. Fix $\omega \in \mathbb{N}^n$. The *weight order* defined by ω is denoted $<_{\omega}$. We say that $\mathbf{x}^{\mathbf{a}} <_{\omega} \mathbf{x}^{\mathbf{b}}$ if $\omega \cdot \mathbf{a} < \omega \cdot \mathbf{b}$. In order to make $<_{\omega}$ a total order, one fixes another term order to break ties among monomials with the same weight.

Definition 1.1.6. Let $f = \sum_{i=1}^r c_i \mathbf{x}^{\mathbf{a}_i}$ be a polynomial in $\mathbb{C}[\mathbf{x}]$ and let $<$ be a term order on $\mathbb{C}[\mathbf{x}]$. The *leading monomial*, or *initial monomial*, of f under the term order $<$, denoted $\text{in}_{<}(f)$ is the monomial $\mathbf{x}^{\mathbf{a}_j}$ that is $<$ -maximal over all terms of f . If $\text{in}_{<}(f) = \mathbf{x}^{\mathbf{a}_j}$, then the *leading term* of f is $\text{LT}_{<}(f) := c_j \mathbf{x}^{\mathbf{a}_j}$. The *initial ideal* of an ideal $I \subset \mathbb{C}[\mathbf{x}]$ is the ideal

$$\text{in}_{<}(I) := \langle \text{in}_{<}(f) \mid f \in I \rangle.$$

Example (Example 1.1.4, continued.). Consider the weight order on R defined by $\omega = (1, 2, 4, 2, 1, 3)$ corresponding to the variables $(x_{11}, x_{12}, x_{13}, x_{21}, x_{22}, x_{23})$ of R . Let $f_1 = x_{11}x_{22} - x_{12}x_{21}$. The weight of the terms of f_1 are $\omega \cdot (1, 0, 0, 1, 0, 0) = 2$ for $x_{11}x_{22}$ and $\omega \cdot (0, 1, 0, 0, 1, 0) = 4$ for $x_{12}x_{21}$. So $\text{in}_{<_{\omega}}(f_1) = x_{12}x_{21}$. We often denote this by $f_1 = \underline{x_{12}x_{21}} - x_{11}x_{22}$ when the term order has already been defined.

We can now define a Gröbner basis of an ideal. As we shall see, Gröbner bases are important when one wishes to apply the division algorithm to test ideal membership. Furthermore, in Section 1.1.1, we will see that Gröbner bases are essential for computing the vanishing ideal of the image of a polynomial map.

Definition 1.1.7. Let $I \subset \mathbb{C}[\mathbf{x}]$ be an ideal. Fix a term order $<$ on $\mathbb{C}[\mathbf{x}]$. The set $\{g_1, \dots, g_s\} \subset I$ is a *Gröbner basis* if $\text{in}_{<}(I) = \langle \text{in}_{<}(g_1), \dots, \text{in}_{<}(g_s) \rangle$.

In order to check whether a set of polynomials \mathcal{G} is a Gröbner basis, one can use the notion of the *S-polynomial* of a pair of polynomials. Let $f_1, f_2 \in \mathbb{C}[\mathbf{x}]$ and let $<$ be a term order on $\mathbb{C}[\mathbf{x}]$. For two monomials $\mathbf{x}^{\mathbf{a}}, \mathbf{x}^{\mathbf{b}} \in \mathbb{C}[\mathbf{x}]$, their *least common multiple* is the monomial

$$\text{LCM}(\mathbf{x}^{\mathbf{a}}, \mathbf{x}^{\mathbf{b}}) := \prod_{i=1}^n x_i^{\max\{a_i, b_i\}}.$$

Let $\mathbf{m} = \text{LCM}(\text{in}_<(f_1), \text{in}_<(f_2))$. The S -polynomial of f_1 and f_2 is

$$S(f_1, f_2) := \frac{\mathbf{m}}{\text{LT}_<(f_1)} f_1 - \frac{\mathbf{m}}{\text{LT}_<(f_2)} f_2.$$

Theorem 1.1.8 (Buchberger's Criterion). *Fix a term order $<$ and a set of polynomials $\mathcal{G} = \{g_1, \dots, g_s\}$. Let $I = \langle g_1, \dots, g_s \rangle$. The following are equivalent:*

1. \mathcal{G} is a Gröbner basis for I .
2. The remainder of each S -polynomial $S(g_i, g_j)$ upon division by \mathcal{G} is zero.

More over, this criterion provides an algorithm for computing Gröbner bases. Corollary 2.29 of [22] describes this algorithm. In fact, Buchberger's algorithm naturally outputs a Gröbner basis that satisfies a special property called being *reduced*.

Definition 1.1.9. A Gröbner basis \mathcal{G} with respect to term order $<$ is *reduced* if for each $g, g' \in \mathcal{G}$, no term of g' is divisible by $\text{in}_<(g)$.

Example (Example 1.1.4, continued). Consider the weight order on R with $\omega = (1, 2, 4, 2, 1, 3)$. Under this weight order, the generators of I along with their leading terms are $f_1 = \underline{x_{12}x_{21}} - x_{11}x_{22}$, $f_2 = \underline{x_{21}x_{13}} - x_{11}x_{23}$ and $f_3 = \underline{x_{12}x_{23}} - x_{13}x_{22}$. We compute the S -polynomial of f_1 and f_2 . We have that $\text{LCM}(\text{LT}_{<\omega}(f_1), \text{LT}_{<\omega}(f_2)) = x_{12}x_{13}x_{21}$. Therefore

$$\begin{aligned} S(f_1, f_2) &= \frac{x_{12}x_{13}x_{21}}{x_{12}x_{21}}(x_{12}x_{21} - x_{11}x_{22}) - \frac{x_{12}x_{13}x_{21}}{x_{21}x_{13}}(x_{21}x_{13} - x_{11}x_{23}) \\ &= x_{11}x_{12}x_{23} - x_{11}x_{13}x_{22} \\ &= x_{11}f_3. \end{aligned}$$

Since $S(f_1, f_2)$ is a multiple of f_3 , its remainder upon division by $\{f_1, f_2, f_3\}$ is zero. One can check that the same is true for both of the other generators of I . So by Buchberger's criterion, $\{f_1, f_2, f_3\}$ is a Gröbner basis for I with respect to the term order $<\omega$.

A Gröbner basis of an ideal is always a generating set. Moreover, one can use a Gröbner basis to test ideal membership using the division algorithm.

Theorem 1.1.10. *Let $I \subset \mathbb{C}[\mathbf{x}]$ be an ideal with Gröbner basis \mathcal{G} . Let $f \in \mathbb{C}[\mathbf{x}]$. Then $f \in I$ if and only if the remainder of f upon division by \mathcal{G} is zero.*

A key feature of algebraic varieties is whether or not they can be written as a union of proper subvarieties. A variety V is *irreducible* if whenever $V = V_1 \cup V_2$ for two varieties V_1 and V_2 , either $V_1 \subset V_2 = V$ or vice versa. An ideal I is *prime* if whenever $fg \in I$, either $f \in I$ or $g \in I$. These two notions are related by the following theorem.

Proposition 1.1.11. *A variety V is irreducible if and only if its vanishing ideal $I(V)$ is prime.*

Let V be a projective variety. The *coordinate ring* of V is the $\mathbb{C}[\mathbf{x}]$ -module, $\mathbb{C}[V] := \mathbb{C}[\mathbf{x}]/I(V)$. Since $I(V)$ is homogeneous, $\mathbb{C}[V]$ is finitely generated and graded. Therefore, when V is a projective variety, $\mathbb{C}[V]$ is a finitely generated, graded $\mathbb{C}[\mathbf{x}]$ -module. As such, it can be written as a direct sum,

$$\mathbb{C}[V] = \bigoplus_{k \geq 0} \mathbb{C}[V]_k,$$

where $\mathbb{C}[V]_k$ is the degree k graded piece of $\mathbb{C}[V]$.

Definition 1.1.12. The *Hilbert function* of V is

$$\text{HF}_V(k) = \dim_{\mathbb{C}} \mathbb{C}[V]_k.$$

The *Hilbert series* of V is the formal power series,

$$\text{Hilb}_V(t) = \sum_{k \geq 0} \text{HF}_V(k) t^k.$$

The Hilbert series of V is always a rational function in t .

The Hilbert series encodes quantitative geometric information about a variety. The dimension of V is one less than the degree of the denominator of its Hilbert series. Moreover, the *degree* of the variety is the numerator of the Hilbert series evaluated at zero; this is equal to the number of intersection points of V with a generic linear space of complementary dimension. To compute the Hilbert series in practice, we use the fact that the Hilbert series of I is equal to that of $\text{in}_{< \omega}(I)$. So, knowledge of a Gröbner basis of I is a key ingredient for the computation of a Hilbert series.

Example (Example 1.1.4 continued). The initial ideal of I using the weight order defined by ω is $\langle x_{12}x_{21}, x_{13}x_{21}, x_{12}x_{23} \rangle$. The k th graded piece of the vector space $R/\text{in}_{< \omega}(I)$ has as its basis the set of all monomials of degree k not divisible by a generator of $\text{in}_{< \omega}(I)$. So the first few values of the Hilbert function of $\text{in}_{< \omega}(I)$, and therefore of $V(I)$, are as follows. First, $\text{HF}_{V(I)}(0) = 1$, since 1 is the only degree 0 monomial. Then $\text{HF}_{V(I)}(1) = 6$, since each variable has degree 1 and none of them are in $\text{in}_{< \omega}(I)$. The number of quadratic monomials in R is $\binom{7}{2} = 21$. All of these monomials are in $(R/\text{in}_{< \omega}(I))_2$ except for the three generators of $\text{in}_{< \omega}(I)$. So $\text{HF}_{V(I)}(2) = 18$. Using Macaulay2 or another computer algebra software, one can compute that

$$\text{Hilb}_{V(I)}(t) = 1 + 6t + 18t^2 + \cdots = \frac{1 + 2t}{(1 - t)^4}.$$

So the dimension of $V(I)$ is 3 and the degree of $V(I)$ is 3.

1.1.1 Parametrized Varieties

A common way to describe a variety is to parametrize it by a polynomial map. We will do this by defining a map of polynomial rings and describing the map it induces between varieties. Let

$\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{y} = (y_1, \dots, y_m)$. Define the function,

$$\begin{aligned}\phi : \mathbb{C}[\mathbf{y}] &\rightarrow \mathbb{C}[\mathbf{x}] \\ y_i &\mapsto \phi_i(\mathbf{x})\end{aligned}$$

where each ϕ_i is a polynomial. The polynomial map ϕ induces a morphism $\phi_* : \mathbb{C}^n \rightarrow \mathbb{C}^m$ defined by $\phi_*(\mathbf{a}) = (\phi_1(\mathbf{a}), \dots, \phi_m(\mathbf{a}))$. Then the *variety parametrized by ϕ_** is the Zariski closure of the image of ϕ_* , denoted $\overline{\text{im}(\phi_*)}$. The varieties discussed in the present work are all parametrized by polynomial (in fact, monomial) maps; however, one can also parametrize a variety via a rational map. The vanishing ideal of a parametrized variety is the kernel of ϕ ; we use the *Elimination Theorem* to compute the generators of this vanishing ideal.

Definition 1.1.13. Let $R = \mathbb{C}[x_1, \dots, x_n, y_1, \dots, y_m]$ be a polynomial ring. An *elimination order*, or *block order*, $<$ for x_1, \dots, x_n is a monomial order on R such that for all $f \in R$, if some y_i divides $\text{in}_<(f)$, then $f \in \mathbb{C}[y_1, \dots, y_m]$.

Theorem 1.1.14. Let $\phi_* : \mathbb{C}^n \rightarrow \mathbb{C}^m$ be a morphism of varieties with coordinate functions (ϕ_1, \dots, ϕ_m) . Let $J \subset \mathbb{C}[x_1, \dots, x_n, y_1, \dots, y_m]$ be the ideal generated by $y_i - \phi_i(x_1, \dots, x_n)$ for $1 \leq i \leq m$. Let $<$ be an elimination order for x_1, \dots, x_n on R and let \mathcal{G} be a Gröbner basis for J with respect to $<$. Then the vanishing ideal of the parametrized variety $\overline{\text{im}(\phi_*)}$ is generated by $\mathcal{G} \cap \mathbb{C}[y_1, \dots, y_m]$.

In other words, in order to find the vanishing ideal of $\text{im}(\phi_*)$, one can compute a Gröbner basis for J with respect to an elimination order and take all elements of the Gröbner basis that only use the variables y_1, \dots, y_m . In the next section, we will see a parametrization of the variety from Example 1.1.4. One nice property of parametrized varieties is that they are irreducible; equivalently, their vanishing ideals are prime.

Theorem 1.1.15. Let $\phi : \mathbb{C}[\mathbf{y}] \rightarrow \mathbb{C}[\mathbf{x}]$ be a polynomial map and let $\phi_* : \mathbb{C}^n \rightarrow \mathbb{C}^m$ be its associated morphism of varieties. Then $\overline{\text{im}(\phi_*)}$ is an irreducible variety. Its vanishing ideal, $\ker(\phi)$ is prime.

1.1.2 Toric Varieties

This section introduces toric varieties and ideals and some of their key properties. Much of the content of this section can be found in Chapter 4 of [38] For a very thorough reference on toric varieties, we refer the reader to [14]. There are many equivalent definitions of toric ideals and varieties. The one that is the most relevant to the current work is that a toric ideal can be realized as the kernel of a monomial map. We will now describe this construction.

Let $A \in \mathbb{Z}^{d \times n}$ be an integer matrix. Let $\mathbb{C}[\mathbf{t}]$ and $\mathbb{C}[\mathbf{x}]$ be polynomial rings in d and n variables respectively. The matrix A has a monomial map naturally associated to it. We define $\phi_A : \mathbb{C}[\mathbf{x}] \rightarrow \mathbb{C}[\mathbf{t}]$ by

$$\phi_A(x_i) = \prod_{j=1}^d t_j^{a_{ij}}$$

for each i . In words, the i th column of A is the exponent vector $\phi_A(x_i)$.

Definition 1.1.16. Let $A \in \mathbb{Z}^{d \times n}$ with associated monomial map $\phi_A : \mathbb{C}[\mathbf{x}] \rightarrow \mathbb{C}[\mathbf{t}]$. The *toric ideal* associated to A , denoted $I_A \subset \mathbb{C}[\mathbf{x}]$ is the kernel of ϕ_A . The toric variety associated to A is the variety of I_A .

Toric ideals are always generated by binomials. In fact, these binomials can be understood in terms of the integer kernel of A . These results are summarized in the following theorem.

Theorem 1.1.17. *The ideal I_A is generated by all binomials of the form $\mathbf{x}^{\mathbf{u}} - \mathbf{x}^{\mathbf{v}}$ where $A\mathbf{u} = A\mathbf{v}$, or equivalently, where $\mathbf{u} - \mathbf{v} \in \ker(A)$. If $\mathbf{1} \in \text{rowspan}(A)$, then I_A is homogeneous. Every reduced Gröbner basis of I_A consists solely of binomials.*

Of course, the generating set described in Theorem 1.1.17 is infinite. However, one may compute several finite generating sets by choosing different term orders and computing a Gröbner basis. A generating set of a toric ideal is often known as a *Markov basis*.

Example 1.1.18 (Example 1.1.4 continued). Let R and I be as in Example 1.1.4. The matrix $A \in \mathbb{Z}^{5 \times 6}$ be the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

where the rows of A are indexed by s_1, s_2, t_1, t_2, t_3 and the columns of A are indexed by $x_{11}, x_{12}, x_{13}, x_{21}, x_{22}, x_{23}$. Then the monomial map

$$\phi_A : R \rightarrow \mathbb{C}[s_1, s_2, t_1, t_2, t_3].$$

is defined by $\phi_A(x_{ij}) = s_i t_j$. Let $S = \mathbb{C}[\mathbf{x}, s_1, s_2, t_1, t_2, t_3]$ and let $J = \langle x_{ij} - s_i t_j \rangle \subset S$. Let $<$ be an elimination order for $\{s_1, s_2, t_1, t_2, t_3\}$. We use Macaulay2 to compute a Gröbner basis for J under one such elimination order and obtain the Gröbner basis,

$$\begin{aligned} \mathcal{G} = \{ & x_{13}x_{22} - x_{12}x_{23}, x_{13}x_{21} - x_{11}x_{23}, x_{12}x_{21} - x_{11}x_{22}, t_3x_{22} - t_2x_{23}, t_3x_{21} - t_1x_{23}, t_2x_{21} - t_1x_{22}, \\ & s_2x_{13} - s_1x_{23}, t_3x_{12}t_2x_{13}, s_2x_{12} - s_1x_{22}, t_3x_{11} - t_1x_{13}, t_2x_{11}t_1x_{12}, s_2x_{11} - s_1x_{21}, \\ & s_2t_3 - x_{23}, s_1t_3 - x_{13}, s_2t_2 - x_{22}, s_1t_2 - x_{12}, s_2t_1 - x_{21}, s_1t_1 - x_{11} \}. \end{aligned}$$

Only the first three polynomials in \mathcal{G} lie in R . So

$$\ker(\phi_A) = \langle x_{13}x_{22} - x_{12}x_{23}, x_{13}x_{21} - x_{11}x_{23}, x_{12}x_{21} - x_{11}x_{22} \rangle = I,$$

the ideal described in Example 1.1.4. Consider the binomial $f_1 = x_{11}x_{22} - x_{12}x_{21}$. The support vectors of the two monomials of f_1 are $\mathbf{u} = (1, 0, 0, 0, 1, 0)^T$ and $\mathbf{v} = (0, 1, 0, 1, 0, 0)^T$ respectively. Note that $\mathbf{u} - \mathbf{v} \in \ker(A)$.

Toric varieties find applications in algebraic statistics as log-linear models, which we shall explore in Section 1.3. Moreover, they arise in phylogenetics as group-based phylogenetic models after a linear change of coordinates. Toric ideals are especially nice to study as they have an associated convex polytope that encodes some of their geometric data. In the next section, we introduce polytopes and some key definitions and facts pertaining to them. Then we describe their relationship with toric ideals.

1.2 Polytopes

In this section, we review polytopes and their Ehrhart theory. For a standard reference on polytopes, we refer the reader to [45].

Definition 1.2.1. A *V-polytope* is the convex hull of finitely many points in \mathbb{R}^d . For a matrix $A \in \mathbb{R}^{d \times n}$, we denote the convex hull of its columns by $\text{conv}(A)$.

Definition 1.2.2. An *H-polytope* is the bounded intersection of finitely many half-spaces in \mathbb{R}^d . For a matrix $B \in \mathbb{R}^{m \times d}$ and vector $\mathbf{z} \in \mathbb{R}^m$, we write

$$P(B, \mathbf{z}) = \{\mathbf{x} \in \mathbb{R}^d \mid B\mathbf{x} \leq \mathbf{z}\}.$$

In fact, these two definitions are equivalent. This is the content of the following theorem.

Theorem 1.2.3. A set $P \subset \mathbb{R}^d$ is a *V-polytope* if and only if P is an *H-polytope*.

Thus, we can simply call these objects *polytopes*. When a polytope P is presented as the convex hull of finitely many points, this is known as a *V-description*. When it is presented as a bounded intersection of half-spaces, this is known as an *H-description*. In order to compute a V-description from an H-description or vice versa, one implements Fourier-Motzkin elimination. This process is described in Chapter 1.2 of [45]. More generally, if we project a polytope P to obtain a polytope Q , one can use Fourier-Motzkin elimination to use the H-description for P to compute an H-description of Q . We take this approach for polytopes arising in mathematical phylogenetics in Chapter 4.

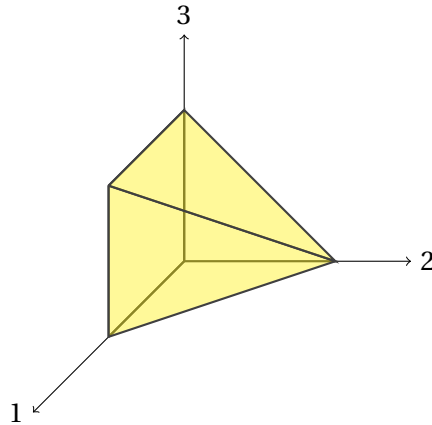
Definition 1.2.4. The *dimension* of a polytope P is the dimension of the affine hull of P . A *face* of a polytope P is the set of all points of P on which some linear functional is maximized; that is, it is a set of the form

$$F = \{\mathbf{x} \in P \mid \mathbf{a} \cdot \mathbf{x} \geq \mathbf{a} \cdot \mathbf{y} \text{ for all } \mathbf{y} \in P\}.$$

for some \mathbf{a} . We also include the empty set in the set of faces of P . A *vertex* of P is a 0-dimensional face of P . If P has dimension d , then a *facet* of P is a $(d-1)$ -dimensional face of P .

Every polytope can be uniquely described as the convex hull of its vertices. When P is full-dimensional in its ambient space, it also has a unique H-description in terms of its facets.

Figure 1.1 The polytope described in Example 1.2.5



Example 1.2.5. Let $P = \text{conv}(V)$ where

$$V = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

In Chapter 3, we will show that $P = P(B, \mathbf{z})$ where

$$B = \begin{bmatrix} -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{z} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

In other words, the facets of P are defined by

$$\begin{aligned} 0 \leq x_1 \leq 1 & & x_1 + x_2 \leq 1 \\ 0 \leq x_2 \leq 1 & & x_2 + x_3 \leq 1 \\ 0 \leq x_3 \leq 1. & & \end{aligned}$$

The polytope P has dimension 3 as it contains the standard simplex in \mathbb{R}^3 . It has 5 vertices and 8 2-dimensional facets. It is a square pyramid with a shifted “point”, as pictured in Figure 1.1.

1.2.1 Ehrhart Theory

We turn our attention to the study of Ehrhart functions and series of lattice polytopes. Let $P \subset \mathbb{R}^d$ be a d -dimensional polytope with integer vertices. Since P is an integer polytope, its volume is a rational number of the form $\text{Vol}(P) = V/d!$ where V is an integer; in this case, the *normalized volume* of P is $V = d!\text{Vol}(P)$. One can compute the normalized volume by finding unimodular triangulations of P ; we will introduce these after the following preliminary definitions.

The *Ehrhart function*, $i_P(m)$, counts the integer points in dilates of P ; that is,

$$i_P(m) = \#(\mathbb{Z}^d \cap mP),$$

where $mP = \{m\mathbf{v} \mid \mathbf{v} \in P\}$ denotes the m th dilate of P . The Ehrhart function is, in fact, a polynomial in m [4, Chapter 3]. We further define the *Ehrhart series* of P to be the generating function

$$\text{Ehr}_P(t) = \sum_{m \geq 0} i_P(m)t^m.$$

The Ehrhart series is of the form

$$\text{Ehr}_P(t) = \frac{h_P^*(t)}{(1-t)^{d+1}},$$

where d is the dimension of P and $h_P^*(t)$ is a polynomial in t of degree at most d . Often we just write $h^*(t)$ when the particular polytope is clear. The coefficients of $h^*(t)$ have an interpretation in terms of a *shelling* of a unimodular triangulation of P , if such a shellable unimodular triangulation exists.

Definition 1.2.6. A *triangulation* of P is a set of d -dimensional simplices $\Delta_1, \dots, \Delta_s$ such that

1. the vertices of each Δ_i are also vertices of P
2. $P = \bigcup_{i=1}^s \Delta_i$, and
3. for all $i \neq j$, $\Delta_i \cap \Delta_j$ is a proper face of each.

A triangulation of P is *unimodular* if each Δ_i has normalized volume 1. If P has a unimodular triangulation with s simplices, then the normalized volume of P is s .

The set of simplices in a triangulation of P forms the set of facets of a simplicial complex. To each simplicial complex Δ , we may associate a monomial ideal called its *Stanley-Reisner ideal* as follows. Let Δ have ν vertices. For each $A \subset [\nu]$, let $\mathbf{i}_A \in \{0, 1\}^\nu$ be the indicator vector for A . Then the Stanley-Reisner ideal, denoted I_Δ is an ideal $\mathbb{C}[x_1, \dots, x_\nu]$ defined by

$$I_\Delta := \langle \mathbf{x}^{\mathbf{i}_A} \mid A \notin \Delta \rangle.$$

The ideal I_Δ is minimally generated by all inclusion-minimal non-faces of Δ . Note that the generators of I_Δ are squarefree.

Example (Example 1.2.5 continued). Let P be as in Example 1.2.5. A unimodular triangulation of P is given by

$$\Delta_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \Delta_2 = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

So the normalized volume of P is 2. Let $\mathbb{C}[\mathbf{x}] = \mathbb{C}[x_1, \dots, x_5]$ where each indeterminate corresponds naturally to a vertex of P in the order given in Example 1.2.5. Let Δ be the simplicial complex with facets Δ_1 and Δ_2 . The unique minimal non-face of this triangulation is the edge between $(0, 0, 0)^T$ and $(1, 0, 1)^T$. So the Stanley-Reisner ideal is

$$I_\Delta = \langle x_1 x_5 \rangle.$$

We can now define the notion of a shelling of a triangulation.

Definition 1.2.7. Let Δ be a pure d -dimensional simplicial complex with facets $\Delta_1, \dots, \Delta_s$. An ordering $\Delta_1, \Delta_2, \dots, \Delta_s$ on these facets is a *shelling order* if for all $1 < r \leq s$,

$$\bigcup_{i=1}^{r-1} (\Delta_i \cap \Delta_r)$$

is a union of facets of Δ_r .

Equivalently, the order $\Delta_1, \Delta_2, \dots, \Delta_s$ is a shelling order if and only if for all $r \leq s$ and $k < r$, there exists an $i < r$ such that $\Delta_k \cap \Delta_r \subset \Delta_i \cap \Delta_r$ and $\Delta_i \cap \Delta_r$ is a facet of Δ_r . This means that when we build our simplicial complex by adding facets in the order prescribed by the shelling order, we add each simplex along its highest dimensional faces. Keeping track of the number of facets that each simplex is added along gives the following relationship between shellings of a triangulation of an integer polytope and the Ehrhart series of the polytope, which is proved in [4, Chapter 3].

Theorem 1.2.8. Let P be a polytope with integer vertices. Let $\{\Delta_1, \dots, \Delta_s\}$ be a regular unimodular triangulation of P using no new vertices. Denote by h_j^* the coefficient of t^j in the h^* polynomial of P . If $\Delta_1, \dots, \Delta_s$ is a shelling order, then h_j^* is the number of Δ_i that are added along j of their facets in this shelling. Equivalently,

$$h^*(t) = \sum_{i=1}^s t^{a_i},$$

where $a_i = \#\{k < i \mid \Delta_k \cap \Delta_i \text{ is a facet of } \Delta_i\}$.

Example (Example 1.2.5 continued). The order Δ_1, Δ_2 is a shelling order in which Δ_1 is added along none of its facets and Δ_2 is added along one of its facets. So since P is 3-dimensional, its Ehrhart series is

$$\text{Ehr}_P(t) = \frac{1+t}{(1-t)^4}.$$

One can also see this by computing the Ehrhart polynomial of P . We can see from Figure 1.1 that we add a square of lattice points for each integer dilate of P ; indeed, from the $(k-1)$ st to

k th dilate of P , we add a square with edges of length k ; this contains $(k + 1)^2$ lattice points. So $i_P(k) = \sum_{i=0}^k (k + 1)^2 = \frac{n(n+1)(2n+1)}{6}$. Multiplying by t^k and summing over all k yields the Ehrhart series.

1.2.2 Toric Ideals and Polytopes

We now describe the connections between toric ideals and polytopes. The following material can be found in [38].

Let $A \in \mathbb{Z}^{d \times n}$ be the matrix defining the monomial map ϕ_A and let I_A be the toric ideal obtained as the kernel of ϕ_A . Then the *polytope associated to I_A* , denoted P_A , is the convex hull of the columns of A . We are especially interested in the case where P_A is *normal*.

Definition 1.2.9. A polytope $P \subset \mathbb{R}^d$ is *normal* if every lattice point in $\mathbb{Z}^d \cap kP$ can be written as a sum of exactly k lattice points in $\mathbb{Z}^d \cap P$.

One can study the polytope P_A to learn combinatorial information about I_A as summarized in the following theorem.

Theorem 1.2.10. Let I_A be a homogeneous toric ideal and let P_A be its associated polytope.

1. The dimension of the projective variety $V(I_A)$ is the dimension of P_A .
2. Suppose that A is such that the integer span of the columns of A is equal to \mathbb{Z}^d . Then the degree of I_A is the normalized volume of P_A .
3. The polytope P_A is normal if and only if the Hilbert series of I_A is equal to the Ehrhart series of P_A .
4. The square-free initial ideals of I_A are the Stanley-Reisner ideals of the regular unimodular triangulations of P_A .

Example (Example 1.2.5). Let A' be obtained from the matrix A in Example 1.2.5 by adding a row of all ones. This embeds the polytope in an affine slice of \mathbb{R}^4 . On the level of the toric ideal, this ensures that the parametrization and therefore the resulting ideal is homogeneous as described in Theorem 1.1.17. The matrix A' has a one-dimensional kernel that is generated by $(1, -1, 0, -1, 1)$. So $I_{A'} = \langle x_1 x_5 - x_2 x_4 \rangle$. Since $I_{A'}$ is generated by a single quadratic binomial, its degree is 2; note that this is also equal to the normalized volume of $P_{A'}$. One can also check that the Hilbert series of $I_{A'}$ is also $(1 + t)/(1 - t)^4$. Let $<$ be a term order which picks $x_1 x_5$ as the initial term of the generator of $I_{A'}$; then the initial ideal is $\text{in}_{<}(I_{A'}) = \langle x_1 x_5 \rangle$. This is the Stanley-Reisner ideal of the triangulation given in the previous example.

1.3 Algebraic Statistical Models

A *statistical model* is a set of probability distributions or density functions. Typically, a statistical model is specified as a parametrized family of distributions or as the set of all distributions that satisfy

certain properties. All statistical models considered in the present thesis are defined parametrically. A typical objective in the field of algebraic statistics is to use algebra, geometry and combinatorics to learn information about some statistical model. We refer the reader to [42] for a reference on algebraic statistics.

1.3.1 Log-Linear Models

Let $A \in \mathbb{Z}^{d \times r}$ with entries a_{ij} . Denote by $\mathbf{1}$ the vector of all ones in \mathbb{Z}^r . We assume throughout that $\mathbf{1} \in \text{rowspan}(A)$. Let Δ_{r-1} denote the $(r-1)$ -dimensional probability simplex in \mathbb{R}^r .

Definition 1.3.1. The *log-linear model* associated to A is the set of probability distributions,

$$\mathcal{M}_A := \{\mathbf{p} \in \Delta_{r-1} \mid \log \mathbf{p} \in \text{rowspan}(A)\}.$$

These are also known as *toric models* because, as we will soon see, their Zariski closures are toric varieties. The matrix A is called the *design matrix* of the log-linear model.

Log-linear models are discrete exponential families. Algebraic and combinatorial tools are well-suited for the study of log-linear models since these models have monomial parametrizations. Let ϕ_A denote the monomial map specified by A . Then we have that $\mathcal{M}_A = \phi_A(\mathbb{R}^d) \cap \Delta_{r-1}$. Indeed, when each coordinate of $\phi_A(\mathbf{t})$ is positive, the logarithm of $\phi_A(\mathbf{t})$ is a linear combination of the rows of A with coefficients t_1, \dots, t_d . Background on log-linear models can be found in [42, Chapter 6.2]. Denote by $\mathbb{C}[\mathbf{p}] := \mathbb{C}[p_1, \dots, p_r]$ the polynomial ring in r indeterminates. Let $I_A \subset \mathbb{C}[\mathbf{p}]$ denote the vanishing ideal of $\phi_A(\mathbb{R}^d)$ over the algebraically closed field \mathbb{C} . Since ϕ_A is a monomial map and the row of all ones is in the rowspan of A , I_A is a homogeneous toric ideal. So, as described in Theorem 1.1.17, it is generated by binomials whose exponent vectors correspond to elements of the integer kernel of A .

One of the most common examples of a log-linear model is the *independence model*. To understand this model, we must first describe some notation and define the notion of independence for discrete random variables.

Let X and Y be discrete random variables on state spaces $[m]$ and $[n]$ respectively. Denote by $P(X = i)$ the probability that X takes the value $i \in [m]$, and similarly for Y . Assume that $P(X = i), P(Y = j) > 0$ for all $i \in [m], j \in [n]$. We write $P(X = i, Y = j)$ to mean the probability that X takes the value i and Y takes the value j . The *probability of $X = i$ given $Y = j$* is denoted $P(X = i \mid Y = j)$ and defined by

$$P(X = i \mid Y = j) = \frac{P(X = i, Y = j)}{P(Y = j)};$$

this is known as a *conditional probability*. The random variables X and Y are *independent* if

$$P(X = i, Y = j) = P(X = i)P(Y = j)$$

for all $i \in [m]$ and $j \in [n]$. Equivalently, this means that $P(X = i \mid Y = j) = P(X = i)$; so knowledge of

Y does not give us any more information about X .

Definition 1.3.2. The $m \times n$ independence model $\mathcal{M}_{m,n}$ consists of all joint probability distributions on discrete random variables X and Y on state spaces $[m]$ and $[n]$ respectively such that $P(X = i, Y = j) = P(X = i)P(Y = j)$ for all $i \in [m], j \in [n]$.

We claim that $\mathcal{M}_{m,n}$ is a log-linear model; in fact, its Zariski closure is the projective variety of all $m \times n$ matrices of rank 1. It is parametrized by the monomial map

$$\phi_{m,n} : \mathbb{C}[p_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n] \rightarrow \mathbb{C}[s_1, \dots, s_m, t_1, \dots, t_n]$$

which sends p_{ij} to $s_i t_j$. In the positive orthant, one may think of s_i as $P(X = i)$, t_j as $P(Y = j)$, and p_{ij} as $P(X = i, Y = j)$. So this parametrization reflects that X and Y are independent. If we view (p_{ij}) as an $m \times n$ matrix, we see that it has rank at most one; indeed, it is written as the outer product $(s_1, \dots, s_m)^T (t_1, \dots, t_n)$.

Example 1.3.3. The toric variety from Example 1.1.4 is the Zariski closure of the 2×3 independence model. Its vanishing ideal is the ideal of 2×2 minors of a generic 2×3 matrix.

1.4 Maximum Likelihood Estimation in Log-Linear Models

Let $A \in \mathbb{Z}^{d \times r}$ be the design matrix for the log-linear statistical model \mathcal{M}_A . Given independent, identically distributed (iid) data $\mathbf{u} \in \mathbb{N}^r$, we wish to infer the distribution $\mathbf{p} \in \mathcal{M}_A$ that is “most likely” to have generated it. This is the central problem of maximum likelihood estimation.

Definition 1.4.1. Let \mathcal{M} be a discrete statistical model in \mathbb{R}^r and let $\mathbf{u} \in \mathbb{N}^r$ be an iid vector of counts recording the number of occurrences of each outcome in an experiment. The *likelihood function* is

$$L(\mathbf{p} \mid \mathbf{u}) = \prod_{i=1}^r p_i^{u_i}.$$

The *maximum likelihood estimate*, or MLE, for \mathbf{u} is the distribution in \mathcal{M} that maximizes the likelihood function; that is, it is the distribution

$$\hat{\mathbf{p}} = \operatorname{argmax}_{\mathbf{p} \in \mathcal{M}} L(\mathbf{p} \mid \mathbf{u}).$$

Note that for a fixed $\mathbf{p} \in \mathcal{M}$, $L(\mathbf{p} \mid \mathbf{u})$ is exactly the probability of observing \mathbf{u} from the distribution \mathbf{p} . Hence, the MLE for \mathbf{u} is the distribution $\hat{\mathbf{p}} \in \mathcal{M}$ that maximizes the probability of observing \mathbf{u} . The map $\mathbf{u} \mapsto \hat{\mathbf{p}}$ is a function of the data known as the *maximum likelihood estimator*. We are particularly interested in the case when the coordinate functions of the maximum likelihood estimator are rational functions of the data. In this case, we say that \mathcal{M} has *rational MLE*.

The *log-likelihood function* $\ell(\mathbf{p} \mid \mathbf{u})$ is the natural logarithm of $L(\mathbf{p} \mid \mathbf{u})$. Note that since the natural log is a concave function, $\ell(\mathbf{p} \mid \mathbf{u})$ and $L(\mathbf{p} \mid \mathbf{u})$ have the same maximizers. We define the *maximum*

likelihood degree of \mathcal{M} to be the number of complex critical points of $\ell(\mathbf{p} \mid \mathbf{u})$ for generic \mathbf{u} . Huh and Sturmfels [25] show that the maximum likelihood degree is well-defined. In particular, \mathcal{M} has maximum likelihood degree 1 if and only if it has rational maximum likelihood estimator [24]. The following result of Huh gives a characterization of the form of this maximum likelihood estimator, when it exists.

Theorem 1.4.2 ([24]). *A discrete statistical model \mathcal{M} has maximum likelihood degree 1 if and only if there exists $\mathbf{h} = (h_1, \dots, h_r) \in (\mathbb{C}^*)^r$, a positive integer d , and a matrix $B \in \mathbb{Z}^{d \times r}$ with entries b_{ij} whose column sums are zero such that the map*

$$\Psi : \mathbb{P}^{r-1} \dashrightarrow (\mathbb{C}^*)^r$$

with coordinate function

$$\Psi_k(u_1, \dots, u_r) = h_k \prod_{i=1}^d \left(\sum_{j=1}^r b_{ij} u_j \right)^{b_{ik}}$$

maps dominantly onto $\overline{\mathcal{M}}$. In this case, the function Ψ is the maximum likelihood estimator for \mathcal{M} .

In this context, the pair (B, \mathbf{h}) is called the *Horn pair* that defines Ψ , and Ψ is called the *Horn map*. For more details about the Horn map and its connection to the theory of A -discriminants, we refer the reader to [16] and [24].

Example 1.4.3. Consider the matrix,

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

and the log-linear model \mathcal{M}_A that it defines. Let

$$S = \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2)\}.$$

We index the columns of A by the elements of S in the given order. In Chapter 3, we shall see that \mathcal{M}_A is the quasi-independence model defined by S , and that \mathcal{M}_A has rational MLE. In particular,

we shall show that the Horn pair associated to \mathcal{M}_A is

$$B = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 \\ -1 & -1 & 0 & -1 & -1 & 0 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \end{bmatrix}$$

with $\mathbf{h} = (-1, -1, 1, -1, -1, 1, 1, 1)$. The columns of B and \mathbf{h} are also indexed by the elements of S . We can use this Horn pair to write the MLE as a rational function of the data.

Let $\mathbf{u} \in \mathbb{N}^S$ be a vector of counts of iid data for the model \mathcal{M}_A . Denote by u_{++} the sum of all entries of \mathbf{u} , and abbreviate each ordered pair $(i, j) \in S$ by ij . Then for example, the $(1, 3)$ coordinate of the MLE is

$$\begin{aligned} \hat{p}_{13} &= h_{13}(u_{11} + u_{12} + u_{13})^1 (u_{13} + u_{23})^1 (u_{11} + u_{12} + u_{13} + u_{21} + u_{22} + u_{23})^{-1} u_{++}^{-1} \\ &= \frac{(u_{11} + u_{12} + u_{13})(u_{13} + u_{23})}{u_{++}(u_{11} + u_{12} + u_{13} + u_{21} + u_{22} + u_{23})}. \end{aligned}$$

Similarly, the $(2, 3)$ coordinate is

$$\hat{p}_{23} = \frac{(u_{21} + u_{22} + u_{23})(u_{13} + u_{23})}{u_{++}(u_{11} + u_{12} + u_{13} + u_{21} + u_{22} + u_{23})}.$$

The following theorem, known as Birch's Theorem, says that the maximum likelihood estimate for \mathbf{u} in a log-linear model \mathcal{M}_A , if it exists, is the unique distribution $\hat{\mathbf{p}}$ in \mathcal{M}_A with the same sufficient statistics as the normalized data. A proof of this result can be found in [42, Chapter 7].

Theorem 1.4.4 (Birch's Theorem). *Let $A \in \mathbb{Z}^{n \times r}$ such that $\mathbf{1} \in \text{rowspan}(A)$. Let $\mathbf{u} \in \mathbb{R}_{\geq 0}^r$ and let $u_+ = u_1 + \dots + u_r$. Then the maximum likelihood estimate in the log-linear model \mathcal{M}_A given data \mathbf{u} is the unique solution, if it exists, to the equations $A\mathbf{u} = u_+ \mathbf{A}\mathbf{p}$ subject to $\mathbf{p} \in \mathcal{M}_A$.*

Example (Example 1.4.3, continued). Consider the last row a_6 of the matrix A . One sufficient statistic

of \mathcal{M}_A is $a_6 \cdot \mathbf{u} = u_{13} + u_{23}$. We must check that $a_6 \cdot \mathbf{u} = u_{++} a_6 \cdot \hat{\mathbf{p}}$. Indeed, we compute that

$$\begin{aligned} a_6 \cdot \hat{\mathbf{p}} &= \frac{(u_{11} + u_{12} + u_{13})(u_{13} + u_{23})}{u_{++}(u_{11} + u_{12} + u_{13} + u_{21} + u_{22} + u_{23})} + \frac{(u_{21} + u_{22} + u_{23})(u_{13} + u_{23})}{u_{++}(u_{11} + u_{12} + u_{13} + u_{21} + u_{22} + u_{23})} \\ &= (u_{13} + u_{23}) \frac{(u_{11} + u_{12} + u_{13} + u_{21} + u_{22} + u_{23})}{u_{++}(u_{11} + u_{12} + u_{13} + u_{21} + u_{22} + u_{23})} \\ &= \frac{u_{13} + u_{23}}{u_{++}}, \end{aligned}$$

as needed.

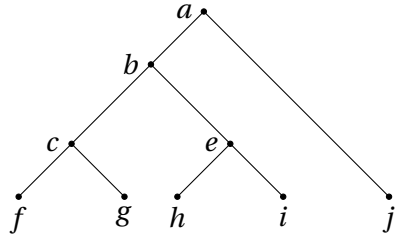
In the language of toric geometry, Birch's theorem says that the maximum likelihood estimate is the unique intersection point of the positive part of the toric variety $\overline{\mathcal{M}_A}$ with the affine linear space defined by $A\mathbf{u} = u_+ A\mathbf{p}$. In practice, to check that a distribution $\hat{\mathbf{p}}$ is the MLE for data \mathbf{u} , one computes the vanishing ideal of \mathcal{M}_A and checks that $\hat{\mathbf{p}}$ satisfies the generators of \mathcal{M}_A and lies in this affine linear space. We shall employ this technique to compute MLEs in quasi-independence models in Chapter 3.

1.5 Group-Based Phylogenetic Models

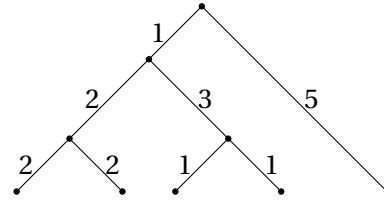
The field of phylogenetics is concerned with reconstructing evolutionary histories of different species or other taxonomic units (*taxa* for short), such as genes or bacterial strains. See [18, 31] for general background on mathematical phylogenetics. In phylogenetics, we use trees to model these evolutionary histories. The leaves of these trees represent the extant taxa of interest, while the internal nodes represent their extinct common ancestors. Branching within the tree represents speciation events, wherein two species diverged from a single common ancestor. One may use combinatorial trees to depict only the evolutionary relationships between the taxa, or include branch lengths to represent time or amount of genetic mutation.

An organism's DNA is made up of chemical compounds called *nucleotides*, or *bases*. There are four different types of bases: adenine, thymine, guanine and cytosine, abbreviated A,T,G and C. These bases are split into two different types based upon their chemical structure. Adenine and guanine are *purines*, and thymine and cytosine are *pyrimidines*. Evolution occurs via a series of *substitutions* within the DNA of a taxon, wherein one nucleotide gets swapped out for another. In the phylogenetic models discussed in the present thesis, we assume that base substitutions occur as a continuous-time Markov process along the edges of a tree, where the edge lengths are the time parameters in the Markov process [26, 27]. The entries of the transition matrices in this Markov process are the probabilities of observing a substitution from one base to another at a site in the genome at the end of the given time interval.

In this section, we discuss a standard phylogenetic model called the Cavender-Farris-Neyman, or CFN, model. The two possible states of the CFN, model are purine and pyrimidine. This is based on the observed fact that within group substitutions (purine-purine or pyrimidine-pyrimidine) are much more common, so a two state model like the CFN model only focuses on cross group



(a) An example of a rooted binary tree



(b) Branch lengths that make this tree equidistant.

Figure 1.2 A rooted binary tree.

substitutions (which are known as *transversions*). In the CFN model we further assume that the rate of substitution from purine to pyrimidine is equal to the rate of substitution from pyrimidine to purine. The algebraic and combinatorial structure of the CFN-model has been studied by a number of authors and key results include descriptions of the generating set of the vanishing ideal [39], Gröbner bases [39], polyhedral geometry [8], Hilbert series [8], and connections to the Hilbert scheme and toric degenerations [40].

1.5.1 Preliminaries on Trees

This section provides a brief background on combinatorial trees and metric trees. A more detailed description can be found in [18, 31].

Definition 1.5.1. A *tree* is a connected graph with no cycles. A *leaf* of the tree T is a node of T of degree 1. A tree is *rooted* if it has a distinguished node of degree 2, called the *root*. A *rooted binary tree* is a rooted tree in which all non-leaf, non-root nodes have degree 3. An internal node of T is a *cherry node* if it is adjacent to two leaves.

Example 1.5.2. Consider the rooted binary tree in Figure 1.2a. It is rooted with root a . The leaves of this tree are f, g, h, i and j . This tree is binary, since the three nodes b, c and e that are not the root or the leaves have degree three. The nodes c and e are both cherry nodes.

Typically, we orient trees with the root at the top of the page and the leaves toward the bottom. This allows us to think of the tree as being directed, so that time starts at the root and progresses in the direction of leaves. Labeling the leaves of the tree with the taxa $\{1, \dots, n\}$ gives a proposed evolutionary history of these taxa. For any tree T , there exists a unique path between any two nodes in the tree. This allows us to say that if a and b are nodes of the rooted tree T , then a is an *ancestor* of b and b is a *descendant* of a if a lies along the path from the root of T to b . Furthermore, a rooted binary tree on n leaves has $n - 1$ internal nodes and $2n - 2$ edges. Proofs of these facts can be found in Chapter 2.1 of [44].

Trees may also come equipped with *branch lengths*, which can represent time, amount of substitution, etc. The branch lengths are assignments of positive real numbers to each edge in the

tree. The assignment of branch lengths to the edges of a tree induces a metric on the nodes in the tree, where the distance between a pair of nodes is the sum of the branch lengths on the unique path in the tree connecting those nodes. Not every metric on a finite set arises in this way; for example, the resulting *tree metrics* must satisfy the four-point condition [9]. In this paper we are interested in the following restricted class of tree metrics.

Definition 1.5.3. An *equidistant* tree T is a rooted tree with positive branch lengths such that the distance between the root and any leaf is the same.

The tree pictured in Figure 1.2b is an example of an equidistant tree. In the phylogenetic modeling literature, this is known as imposing the *molecular clock* condition on the model. In other contexts, an equidistant tree metric is also known as an *ultrametric*.

1.5.2 The Cavender-Farris-Neyman Model

In this section, we introduce the Cavender-Farris-Neyman model, or CFN model. Note that the CFN model is also referred to as the binary Jukes-Cantor model and the binary symmetric model throughout the literature. We will use these results to provide a combinatorial description of the toric ideal of phylogenetic invariants of the CFN model with the molecular clock, which is the main object of study in Chapters 4 and 5.

The CFN model describes substitutions at a single site in the gene sequences of the taxa in question. It is a two-state model, where the states are purine (adenine and guanine) and pyrimidine (thymine and cytosine). We denote purines with U and pyrimidines with Y . The CFN model assumes a continuous-time Markov process along a fixed rooted binary tree with positive branch lengths. The rate matrix for the Markov process in the CFN model is

$$Q = \begin{array}{cc} & \begin{array}{c} U \quad Y \end{array} \\ \begin{array}{c} U \\ Y \end{array} & \begin{bmatrix} -\alpha & \alpha \\ \alpha & -\alpha \end{bmatrix} \end{array},$$

for some parameter $\alpha > 0$ that describes the rate of change of purines to pyrimidines or vice versa. Note that in the CFN model, we assume that the rate of substitution from purine to pyrimidine is equal to the rate of substitution from pyrimidine to purine. We also assume that the distribution of states at the root of the tree, or *root distribution*, is uniform.

Let $t_e > 0$ be the branch length of an edge e in the rooted binary tree T . The transition matrix M^e associated to the edge e is the matrix exponential,

$$\begin{aligned} M^e &= \exp(Q t_e) \\ &= \begin{bmatrix} (1 + e^{-2\alpha t_e})/2 & (1 - e^{-2\alpha t_e})/2 \\ (1 - e^{-2\alpha t_e})/2 & (1 + e^{-2\alpha t_e})/2 \end{bmatrix}. \end{aligned}$$

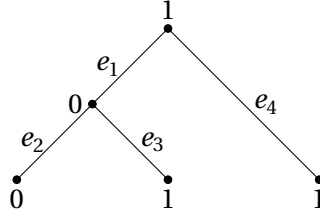


Figure 1.3 A labeling of all nodes of tree T with elements of \mathbb{Z}_2 .

Denote by $a(e)$ and $d(e)$ the two nodes adjacent to e so that $d(e)$ is a descendant of $a(e)$. Then the (i, j) th entry of M^e , $M^e(i, j)$, is the probability that $d(e)$ has state j given that $a(e)$ has state i for all $i, j \in \{U, Y\}$.

Let T be a rooted binary tree with edge set E and nodes $1, \dots, 2n - 1$. For the following section, we label these nodes so that $1, \dots, n$ are leaf labels. We can identify the set of states $\{U, Y\}$ with elements of the two element group \mathbb{Z}_2 . (Note that it does not matter which identification is chosen; either $U = 0, Y = 1$, or $Y = 0, U = 1$ produce the same results.) Let $\mathbf{u} \in \mathbb{Z}_2^{2n-1}$ be a labeling of all of the nodes of T by states in the state space, and let u_i denote the i th coordinate of \mathbf{u} , which is the labeling of node i . Then the probability of observing the set of states \mathbf{u} is

$$\frac{1}{2} \prod_{e \in E} M^e(u_{a(e)}, u_{d(e)}). \quad (1.1)$$

We note that the factor of $\frac{1}{2}$ appears in this formula because the distribution of states at the root of the tree is uniform.

Example 1.5.4. For the tree in Figure 1.3, the probability of observing the states $(0, 1, 1, 0, 1)$ (left to right and bottom to top) is

$$\frac{1}{2} M^{e_1}(1, 0) M^{e_2}(0, 0) M^{e_3}(0, 1) M^{e_4}(1, 1).$$

We have described the CFN model thus far with all variables observed. However, in typical phylogenetic analysis we do not have access to the DNA of the unknown ancestral species, and hence we need to consider a hidden variable model where all internal nodes correspond to hidden states. In this case, to determine the probability of observing a certain set of states at the leaves, we sum the probabilities given by Equation (1.1) over all possible labelings of the internal nodes of the tree. Let $\mathbf{v} \in \mathbb{Z}_2^n$ be a labeling of the leaves of T . Since the CFN-MC model assumes a uniform distribution of states at the root, the probability of observing the set of states \mathbf{v} at the leaves is

$$p(v_1, \dots, v_n) = \frac{1}{2} \sum_{(v_{n+1}, \dots, v_{2n-1}) \in \mathbb{Z}_2^{n-1}} \prod_{e \in E} M^e(v_{a(e)}, v_{d(e)}). \quad (1.2)$$

Note that $d(e)$ might be a leaf, in which case $v_{d(e)} = v_i$ for the appropriate value of i .

Example 1.5.5. Consider the tree from Figure 1.3. We use Equation (1.2) to compute the probability $p(0, 1, 1)$ of observing states $(0, 1, 1)$ at the leaves of T . Summing over all possible labelings of the internal nodes of T yields

$$\begin{aligned} p(0, 1, 1) = & \frac{1}{2} \left(M^{e_1}(0, 0)M^{e_2}(0, 0)M^{e_3}(0, 1)M^{e_4}(0, 1) \right. \\ & + M^{e_1}(1, 0)M^{e_2}(0, 0)M^{e_3}(0, 1)M^{e_4}(1, 1) \\ & + M^{e_1}(0, 1)M^{e_2}(1, 0)M^{e_3}(1, 1)M^{e_4}(0, 1) \\ & \left. + M^{e_1}(1, 1)M^{e_2}(1, 0)M^{e_3}(1, 1)M^{e_4}(1, 1) \right). \end{aligned}$$

1.6 Summary of Results

The contents of Chapters 2 and 3 of the present work were published in the Journal of Symbolic Computation [12]. In Chapter 2, we introduce the notion of a facial submodel of a log-linear model. We prove the following result regarding MLEs of facial submodels of log-linear models with rational MLE.

Theorem (Theorem 2.2.1). *Let \mathcal{M} be a log-linear model with rational MLE and let \mathcal{M}' be a facial submodel of \mathcal{M} . Then \mathcal{M}' also has rational MLE.*

The full version of this theorem describes how to obtain the Horn pair for \mathcal{M}' from that of \mathcal{M} .

In Chapter 3, we introduce two-way quasi-independence models and the bipartite graphs that they encode. We characterize when a quasi-independence model has rational MLE and compute the MLE in these cases. In particular, we prove the following theorem.

Theorem (Theorem 3.3.4). *Let \mathcal{M} be a quasi-independence model with associated bipartite graph G . Then \mathcal{M} has rational MLE if and only if G is doubly chordal bipartite. In these cases, the MLE of \mathcal{M} has an explicit formula using complete bipartite subgraphs of G .*

In Chapter 4, we turn our attention to the Cavendar-Farris-Neyman model with a molecular clock, or CFN-MC model. The contents of this chapter were published in Advances in Applied Mathematics [13]. Let T be a rooted binary phylogenetic tree on n leaves. We apply a linear change of coordinates called the *discrete Fourier transform* to the CFN-MC model on T to obtain a toric variety. Let I_T denote the toric vanishing ideal of this variety. Let R_T denote the polytope associated to I_T . We give an explicit vertex description of this polytope and prove the following results regarding the V- and H-descriptions of R_T .

Theorem (Proposition 4.2.3 and Corollary 4.2.13). *The polytope R_T has F_n vertices, where F_n denotes the n th Fibonacci number. The facets of R_T can be described using combinatorial features of T called clusters.*

The full version of this theorem gives an explicit facet description of R_T .

We then consider the problem of finding a Gröbner basis for I_T . We define a special term order on the polynomial ring of I_T and use it to prove the following theorem.

Theorem (Theorem 4.2.15). *For any rooted binary phylogenetic tree T , the CFN-MC ideal I_T has a Gröbner basis consisting of quadratic binomials with squarefree initial terms.*

The full version of this theorem describes how to compute these Gröbner bases.

In Chapter 5, we study the Ehrhart theory of the CFN-MC polytopes. In particular, we show that for a specific tree called a *caterpillar tree*, the CFN-MC polytope is affinely isomorphic to the order polytope of the zig-zag poset. The Ehrhart theory of this order polytope is governed by the combinatorics of alternating permutations; in particular, its volume is equal to the number of alternating permutations on $(n - 1)$ letters. This is the $(n - 1)$ st *Euler zig-zag number*. Then we show that the Ehrhart polynomial of R_T is equal to that of the caterpillar tree for all T . This leads to the following result, which was also published in [13].

Theorem (Theorem 5.2.1). *The CFN-MC polytope R_T has normalized volume equal to the $(n - 1)$ st Euler zig-zag number.*

Finally, we study the Ehrhart series of these polytopes by analyzing that of the order polytope of the zig-zag poset. This is the content of our preprint, [11]. We compute the h^* -polynomial of R_T by introducing the *swap statistic* on alternating permutations. We introduce a family of shellings of the order polytope of the zig-zag poset and show that the swap statistic counts the number of facets that each simplex is added along. Let A_n denote the set of alternating permutations on n letters. We prove the following theorem.

Theorem (Theorem 5.3.1). *The h^* -polynomial of R_T is $\sum_{\sigma \in A_{n-1}} t^{\text{swap}(\sigma)}$.*

CHAPTER

2

FACIAL SUBMODELS AND RATIONAL MAXIMUM LIKELIHOOD ESTIMATES

In this chapter, we introduce the notion of facial submodels of log-linear models, and show that a facial submodel of a model with rational MLE also has rational MLE. The proof of this result utilizes Birch's theorem and the Horn uniformization of a discrete statistical model with rational MLE.

2.1 Preliminaries

Let $A \in \mathbb{Z}^{n \times r}$ be the matrix defining the monomial map for the log-linear model \mathcal{M}_A . Let I_A denote the vanishing ideal of the Zariski closure of \mathcal{M}_A . We assume throughout that $\mathbf{1} \in \text{rowspan}(A)$. Let $P_A = \text{conv}(A)$, where $\text{conv}(A)$ denotes the convex hull of the columns $\mathbf{a}_1, \dots, \mathbf{a}_r$ of A .

We assume throughout that P_A has n facets, F_1, \dots, F_n , and that the ij entry of A , denoted a_{ij} is equal to the lattice distance between the j th column of A and facet F_i . This is not a restriction, since one can always reparametrize a log-linear model in this way [30, Theorem 27]. Indeed, given a polytope Q , a matrix A that satisfies the above condition is a *slack matrix* of Q , and the convex hull of the columns of A is affinely isomorphic to Q [21].

Let \bar{A} be a matrix whose columns are a subset of A . Without loss of generality, assume that the columns of \bar{A} are $\mathbf{a}_1, \dots, \mathbf{a}_s$.

Definition 2.1.1. The submatrix \bar{A} is called a *facial submatrix* of A if $P_{\bar{A}}$ is a face of P_A . The corresponding statistical model $\mathcal{M}_{\bar{A}}$ is called a *facial submodel* of \mathcal{M}_A .¹

¹Note that the term “facial submodel” is a slight abuse of terminology because $\mathcal{M}_{\bar{A}}$ is not a submodel of \mathcal{M}_A . This is

Let \mathbf{e}_i denote the i th standard basis vector in \mathbb{R}^n . Then $\mathbf{e}_i \cdot \mathbf{a}_j = 0$ if \mathbf{a}_j lies on F_i and $\mathbf{e}_i \cdot \mathbf{a}_j \geq 1$ otherwise. So under our assumptions on A , this definition of a facial submatrix of A aligns with the one given in [19] and [30].

2.2 Main Result

We prove the following result concerning the maximum likelihood estimator for $\mathcal{M}_{\bar{A}}$ when \bar{A} is a facial submatrix of A . This result was used implicitly in the proof of Theorem 4.4 of [19].

Theorem 2.2.1. *Let $A \in \mathbb{Z}^{n \times r}$ and let $\bar{A} \in \mathbb{Z}^{n \times s}$ consist of the first s columns of A . Suppose that \bar{A} is a facial submatrix of A . Let \mathcal{M}_A have rational maximum likelihood estimator Ψ given by the Horn pair (B, \mathbf{h}) where $B \in \mathbb{Z}^{d \times r}$ and $\mathbf{h} \in (\mathbb{C}^*)^r$. Let \bar{B} denote the submatrix consisting of the first s columns of B and let $\bar{\mathbf{h}} = (h_1, \dots, h_s)$. Then $\mathcal{M}_{\bar{A}}$ has rational maximum likelihood estimator $\bar{\Psi}$ given by the Horn pair $(\bar{B}, \bar{\mathbf{h}})$.*

In order to prove Theorem 2.2.1, we check the conditions of Birch's theorem. We do this using the following Lemmas.

Lemma 2.2.2. *Let $\bar{\Psi}$ be as in Theorem 2.2.1. Then for generic $\bar{\mathbf{u}} \in \mathbb{R}_{\geq 0}^s$, $\bar{\Psi}(\bar{\mathbf{u}})$ is defined. In this case, $\bar{\Psi}(\bar{\mathbf{u}})$ is in the Zariski closure of $\mathcal{M}_{\bar{A}}$.*

Proof. Let $\mathbf{u} \in \mathbb{R}_{\geq 0}^r$ be given by $u_i = \bar{u}_i$ if $i \leq s$ and $u_i = 0$ if $i > s$. We claim that when $\Psi(\mathbf{u})$ is defined, $\bar{\Psi}_k(\bar{\mathbf{u}}) = \Psi_k(\mathbf{u})$ for $k \leq s$. Indeed, each factor of $\Psi_k(\mathbf{u})$ is of the form

$$\left(\sum_{j=1}^r b_{ij} u_j \right)^{b_{ik}}$$

for each $i = 1, \dots, d$. If the i th factor of Ψ_k is not identically equal to one, then $b_{ik} \neq 0$. So the i th factor has the nonzero summand $b_{ik} u_k$ and is generically nonzero when evaluated at a point u of the given form. In particular, this implies that $\Psi_k(\mathbf{u})$ is defined for a generic u of the given form since having $u_j = 0$ for $j > s$ does not make any factor of Ψ_k identically equal to zero. Setting each $b_{ij} = 0$ when $j > s$ gives that $\bar{\Psi}_k(\bar{\mathbf{u}}) = \Psi_k(\mathbf{u})$ when $k \leq s$.

The elements of $I_{\bar{A}}$ are those elements of I_A that belong to the polynomial ring $k[p_1, \dots, p_s]$. Let $f \in I_{\bar{A}}$. Since $f \in I_A$ as well, $f(\bar{\Psi}(\bar{\mathbf{u}})) = f(\Psi(\mathbf{u})) = 0$, as needed. \square

Next we check that the sufficient statistics $\bar{A}\bar{\mathbf{u}}/\bar{u}_+$ are equal to those of $\bar{\Psi}(\bar{\mathbf{u}})$.

Lemma 2.2.3. *Let $\bar{\mathbf{c}}$ be a row of \bar{A} . Then*

$$\frac{\bar{\mathbf{c}} \cdot \bar{\mathbf{u}}}{\bar{u}_+} = \bar{\mathbf{c}} \cdot \bar{\Psi}(\bar{\mathbf{u}}).$$

because the log-linear model \mathcal{M}_A does not include distributions on the boundary of the probability simplex. Technically, $\mathcal{M}_{\bar{A}}$ is a submodel of the closure of \mathcal{M}_A .

Proof. Let \mathbf{c} be the row of A corresponding to $\bar{\mathbf{c}}$. Define a sequence $\mathbf{u}^{(i)} \in \mathbb{R}_{\geq 0}^r$ by

$$\mathbf{u}_j^{(i)} = \begin{cases} \bar{u}_j & \text{if } j \leq s \\ \epsilon_j^{(i)} & \text{if } j > s, \end{cases}$$

where $\lim_{i \rightarrow \infty} \epsilon_j^{(i)} = 0$ for each j . We choose each $\epsilon_j^{(i)} > 0$ generically so that $\Psi(\mathbf{u}^{(i)})$ is defined for all i .

Since $\bar{\mathbf{u}}$ is generic, we have that $\lim_{i \rightarrow \infty} u_+^{(i)} = \bar{u}_+ \neq 0$. Similarly, we have that $\lim_{i \rightarrow \infty} \mathbf{c} \cdot \mathbf{u}^{(i)} = \bar{\mathbf{c}} \cdot \bar{\mathbf{u}}$. So

$$\lim_{i \rightarrow \infty} \frac{\mathbf{c} \cdot \mathbf{u}^{(i)}}{u_+^{(i)}} = \frac{\bar{\mathbf{c}} \cdot \bar{\mathbf{u}}}{\bar{u}_+}.$$

Since $\Psi(\mathbf{u}^{(i)})$ is the maximum likelihood estimate in \mathcal{M}_A for each $\mathbf{u}^{(i)}$, by Birch's theorem we have that

$$\begin{aligned} \frac{\mathbf{c} \cdot \mathbf{u}^{(i)}}{u_+^{(i)}} &= \mathbf{c} \cdot \Psi(\mathbf{u}^{(i)}) \\ &= \sum_{j=1}^s c_j \Psi_j(\mathbf{u}^{(i)}) + \sum_{j=s+1}^r c_j \Psi_j(\mathbf{u}^{(i)}). \end{aligned}$$

By the arguments in the proof of Lemma 2.2.2, when $k \leq s$, no factor of $\Psi_k(\mathbf{u}^{(i)})$ involves only summands $u_j^{(i)}$ for $j > s$. So $\lim_{i \rightarrow \infty} \Psi_k(\mathbf{u}^{(i)}) = \bar{\Psi}_k(\bar{\mathbf{u}})$.

Finally, we claim that for $k > s$, $\lim_{i \rightarrow \infty} \Psi_k(\mathbf{u}^{(i)}) = 0$. Without loss of generality, we may assume that $P_{\bar{A}}$ is a facet of P_A . Indeed, if it were not, we could simply iterate these arguments over a saturated chain of faces between $P_{\bar{A}}$ and P_A in the face lattice of P_A . Let $\boldsymbol{\alpha} = (a_1, \dots, a_r)$ be the row of A corresponding to the facet $P_{\bar{A}}$ of P_A . Then $a_j = 0$ if $j \leq s$ and $a_j \geq 1$ if $j > s$. Since $\Psi(\mathbf{u}^{(i)})$ is the maximum likelihood estimate in \mathcal{M}_A for $\mathbf{u}^{(i)}$, by Birch's theorem we have that

$$\begin{aligned} \boldsymbol{\alpha} \cdot \Psi(\mathbf{u}^{(i)}) &= \frac{1}{u_+^{(i)}} (a_{s+1} u_{s+1}^{(i)} + \dots + a_r u_r^{(i)}) \\ &= \frac{1}{u_+^{(i)}} (a_{s+1} \epsilon_{s+1}^{(i)} + \dots + a_r \epsilon_r^{(i)}). \end{aligned}$$

Since $\bar{u}_+ \neq 0$, we also have that

$$\begin{aligned} \lim_{i \rightarrow \infty} \boldsymbol{\alpha} \cdot \Psi(\mathbf{u}^{(i)}) &= \lim_{i \rightarrow \infty} \frac{1}{u_+^{(i)}} (a_{s+1} \epsilon_{s+1}^{(i)} + \dots + a_r \epsilon_r^{(i)}) \\ &= \frac{1}{\bar{u}_+} \lim_{i \rightarrow \infty} (a_{s+1} \epsilon_{s+1}^{(i)} + \dots + a_r \epsilon_r^{(i)}) \\ &= 0. \end{aligned}$$

Furthermore, for all i and k , $\Psi_k(\mathbf{u}^{(i)}) > 0$. So $\lim_{i \rightarrow \infty} \Psi_k(\mathbf{u}^{(i)}) \geq 0$. Since each $a_i > 0$ for $i > s$, this

implies that $\lim_{i \rightarrow \infty} \Psi_k(\mathbf{u}^{(i)}) = 0$ for all $k > s$.

So we have that

$$\begin{aligned} \frac{\bar{\mathbf{c}} \cdot \bar{\mathbf{u}}}{\bar{u}_+} &= \lim_{i \rightarrow \infty} \frac{\mathbf{c} \cdot \mathbf{u}^{(i)}}{u_+^{(i)}} \\ &= \lim_{i \rightarrow \infty} \mathbf{c} \cdot \Psi(\mathbf{u}^{(i)}) \\ &= \bar{\mathbf{c}} \cdot \bar{\Psi}(\bar{\mathbf{u}}) + \sum_{j=s+1}^r c_j \left(\lim_{i \rightarrow \infty} \Psi_j(\mathbf{u}^{(i)}) \right) \\ &= \bar{\mathbf{c}} \cdot \bar{\Psi}(\bar{\mathbf{u}}), \end{aligned}$$

as needed. □

Proof of Theorem 2.2.1. First, note that $\bar{\Psi}$ is still a rational function of degree zero since deleting columns of B does not affect the remaining column sums. So $(\bar{B}, \bar{\mathbf{h}})$ is a Horn pair.

By Lemma 2.2.2, we have that $\bar{\Psi}(\bar{\mathbf{u}}) \in \overline{\mathcal{M}_A}$. Since $\mathbf{1} \in \text{rowspan}(\bar{A})$, it follows from Lemma 2.2.3 that $\sum_{k=1}^s \bar{\Psi}_k(\bar{\mathbf{u}}) = 1$. Defining a sequence $\{\mathbf{u}^{(i)}\}_{i=1}^{\infty}$ as in the proof of Lemma 2.2.3, we have that $\bar{\Psi}_k(\bar{\mathbf{u}}) = \lim_{i \rightarrow \infty} \Psi_k(\mathbf{u}^{(i)})$. So $\bar{\Psi}_k(\bar{\mathbf{u}}) \geq 0$ since each $\Psi_k(\mathbf{u}^{(i)}) > 0$. Furthermore, for generic choices of $\bar{\mathbf{u}}$, we cannot have $\bar{\Psi}_k(\bar{\mathbf{u}}) = 0$. Indeed, for $k \leq s$, the i th factor of $\bar{\Psi}_k(\bar{\mathbf{u}})$ has nonzero summand $b_{ik} u_k$. So none of these factors is zero for generic choices of \mathbf{u} of the given form. Therefore $\bar{\Psi}(\bar{\mathbf{u}}) \in \overline{\mathcal{M}_A} = \overline{\mathcal{M}_A} \cap \Delta_{s-1}$.

By Lemma 2.2.3,

$$\frac{\bar{A} \cdot \bar{\mathbf{u}}}{\bar{u}_+} = \bar{A} \cdot \bar{\Psi}(\bar{\mathbf{u}}).$$

So by Birch's theorem, $\bar{\Psi}$ is the maximum likelihood estimator for $\overline{\mathcal{M}_A}$. □

Note that $\bar{\Psi}$ is a dominant map. Indeed, for generic $\mathbf{p} \in \overline{\mathcal{M}_A}$, $\bar{\Psi}(\mathbf{p})$ is defined. Since \mathbf{p} is a probability distribution, $p_+ = 1$. By Birch's Theorem, \mathbf{p} is the MLE for data vector \mathbf{p} . So $\bar{\Psi}(\mathbf{p}) = \mathbf{p}$.

We close this chapter by noting that we believe that a natural generalization of Theorem 2.2.1 is also true.

Conjecture 2.2.4. *Let $A \in \mathbb{Z}^{n \times r}$ and $\bar{A} \in \mathbb{Z}^{n \times s}$ a facial submatrix of A . Then the maximum likelihood degree of $\overline{\mathcal{M}_A}$ is greater than or equal to the maximum likelihood degree of $\overline{\mathcal{M}_{\bar{A}}}$.*

CHAPTER

3

QUASI-INDEPENDENCE MODELS WITH RATIONAL MAXIMUM LIKELIHOOD ESTIMATOR

Let X and Y be two discrete random variables with m and n states, respectively. Quasi-independence models describe the situation in which some combinations of states of X and Y cannot occur together, but X and Y are otherwise independent of one another. This condition is known as quasi-independence in the statistics literature [5]. Quasi-independence models are basic models that arise in data analysis with log-linear models. For example, quasi-independence models arise in the biomedical field as rater agreement models [1, 29] and in engineering to model system failures at nuclear plants [10]. There is a great deal of literature regarding hypothesis testing under the assumption of quasi-independence, see, for example, [7, 20, 34]. Results about existence and uniqueness of the maximum likelihood estimate in quasi-independence models as well as explicit computations in some cases can be found in [5, Chapter 5]. The main result of the present chapter is 3.3.4, which gives a complete classification of quasi-independence models with rational MLE and a formula for the MLE in these cases.

3.1 Preliminaries

In order to define quasi-independence models, let $S \subset [m] \times [n]$ be a set of indices, where $[m] = \{1, 2, \dots, m\}$. These correspond to a matrix with structural zeros whose observed entries are given

by the indices in S . We often use S to refer to both the set of indices and the matrix representation of this set and abbreviate the ordered pairs (i, j) in S by ij . For all r , we denote by Δ_{r-1} the open $(r-1)$ -dimensional probability simplex in \mathbb{R}^r ,

$$\Delta_{r-1} := \{\mathbf{x} \in \mathbb{R}^r \mid x_i > 0 \text{ for all } i \text{ and } \sum_{i=1}^r x_i = 1\}.$$

Definition 3.1.1. Let $S \subset [m] \times [n]$. Index the coordinates of \mathbb{R}^{m+n} by $(s_1, \dots, s_m, t_1, \dots, t_n) = (\mathbf{s}, \mathbf{t})$. Let \mathbb{R}^S denote the real vector space of dimension $\#S$ whose coordinates are indexed by S . Define the monomial map $\phi^S : \mathbb{R}^{m+n} \rightarrow \mathbb{R}^S$ by

$$\phi_{ij}^S(\mathbf{s}, \mathbf{t}) = s_i t_j.$$

Let $A(S)$ be the 0/1-matrix that defines this monomial map. The *quasi-independence model* associated to S is the model,

$$\mathcal{M}_S := \phi^S(\mathbb{R}^{m+n}) \cap \Delta_{\#S-1}.$$

We note that the Zariski closure of \mathcal{M}_S is a toric variety since it is parametrized by monomials. To any quasi-independence model, we can associate a bipartite graph in the following way.

Definition 3.1.2. The *bipartite graph associated to S* , denoted G_S , is the bipartite graph with independent sets $[m]$ and $[n]$ with an edge between i and j if and only if $(i, j) \in S$. The graph G_S is *chordal bipartite* if every cycle of length greater than or equal to 6 has a chord. The graph G_S is *doubly chordal bipartite* if every cycle of length greater than or equal 6 has at least two chords. We say that S is doubly chordal bipartite if G_S is doubly chordal bipartite.

We can now state the key result of this chapter.

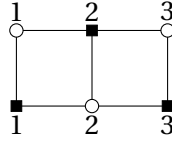
Theorem 3.1.3. *Let $S \subset [m] \times [n]$ and let \mathcal{M}_S be the associated quasi-independence model. Let G_S be the bipartite graph associated to S . Then \mathcal{M}_S has rational maximum likelihood estimate if and only if G_S is doubly chordal bipartite.*

Theorem 3.3.4 is a strengthened version of Theorem 3.1.3 in which we give an explicit formula for the MLE when G_S is doubly chordal bipartite. The outline of the rest of the chapter is as follows. In Section 3.2, we apply the results of Chapter 2 to show that if G_S is not doubly chordal bipartite, then \mathcal{M}_S does not have rational MLE. The main bulk of the chapter is in Sections 3.3, 3.4 and 3.5, where we show that if G_S is doubly chordal bipartite, then the MLE is rational and we give an explicit formula for it. Section 3.3 covers combinatorial features of doubly chordal bipartite graphs and gives the statement of the main Theorem 3.3.4. Sections 3.4 and 3.5 are concerned with the verification that the formula for the MLE is correct.

3.2 Quasi-independence Models with Non-Rational MLE

In this section, we show that when S is not doubly chordal bipartite, the ML-degree of \mathcal{M}_S is strictly greater than one. We can apply Theorem 2.2.1 to quasi-independence models whose associated

Figure 3.1 The double-square graph associated to the matrix in Example 3.2.1



bipartite graphs are not doubly chordal bipartite using cycles and the following “double square” structure.

Example 3.2.1. The minimal example of a chordal bipartite graph that is not doubly chordal bipartite is the *double-square graph*. The matrix of the double-square graph has the form

$$\begin{bmatrix} \star & \star & 0 \\ \star & \star & \star \\ 0 & \star & \star \end{bmatrix},$$

or any permutation of the rows and columns of this matrix. The resulting graph, pictured in Figure 3.1 is two squares joined along an edge. This is a 6-cycle with exactly one chord and as such, is not doubly chordal bipartite.

Remark 3.2.2. A bipartite graph is doubly chordal bipartite if and only if it is chordal bipartite and does not have the double-square graph as an induced subgraph.

We now compute the maximum likelihood degree of models associated to the double square and to cycles of length greater than or equal to 6.

Proposition 3.2.3. *The maximum likelihood degree of the quasi-independence model whose associated graph is the double square is 2.*

Proof. Without loss of generality, let

$$S = \{11, 12, 21, 22, 23, 32, 33\},$$

so that G_S is a double-square graph. Then the vanishing ideal of \mathcal{M}_S is the ideal $I(\mathcal{M}_S) \subset \mathbb{C}[p_{ij} \mid i, j \in S]$ given by

$$I(\mathcal{M}_S) = \langle p_{11}p_{22} - p_{12}p_{21}, p_{22}p_{33} - p_{23}p_{32} \rangle.$$

Define the hyperplane arrangement

$$\mathcal{H} := \{\mathbf{p} \in \mathbb{C}^S \mid p_{++} \prod_{i,j \in S} p_{ij} = 0\},$$

where p_{++} denotes the sum of all the coordinates of \mathbf{p} . Then Proposition 7 of [2] implies that the

ML-degree of \mathcal{M}_S is the number of solutions to the system

$$I(\mathcal{M}_S) + \langle A(S)\mathbf{u} + u_+A(S)\mathbf{p} \rangle$$

that lie outside of \mathcal{H} for generic \mathbf{u} . Since $A(S)$ encodes the row and column marginals of \mathbf{u} , the MLE for \mathbf{u} can be written in matrix form as

$$\begin{bmatrix} u_{11} + \alpha & u_{12} - \alpha & 0 \\ u_{21} - \alpha & u_{22} + \alpha + \beta & u_{23} - \beta \\ 0 & u_{32} - \beta & u_{33} + \beta \end{bmatrix}$$

for some α and β . So computing the MLE is equivalent to solving for α and β in the system

$$\begin{aligned} (u_{11} + \alpha)(u_{22} + \alpha + \beta) - (u_{12} - \alpha)(u_{21} - \alpha) &= 0 \\ (u_{22} + \alpha + \beta)(u_{33} + \beta) - (u_{23} - \beta)(u_{32} - \beta) &= 0. \end{aligned}$$

Expanding gives two equations of the form

$$\begin{aligned} \alpha\beta + c_1\alpha + c_2\beta + c_3 &= 0 \\ \alpha\beta + d_1\alpha + d_2\beta + d_3 &= 0, \end{aligned} \tag{3.1}$$

where each c_i, d_i are polynomials in the entries of \mathbf{u} .

Solving for $\alpha = -(c_2\beta + c_3)/(\beta + c_1)$ in the first equation of (3.1) and substituting into the second gives a degree 2 function of β , which can have at most two solutions. Indeed, for generic choices of u , this equation has exactly two solutions, neither of which lie on \mathcal{H} . For example, take $u_{11} = u_{12} = u_{21} = u_{22} = 1$ and $u_{23} = u_{32} = u_{33} = 2$. By performing this substitution in (3.1) with these values for \mathbf{u} , we obtain the degree 2 equation

$$\frac{-\beta^2}{\beta + 4} + \frac{7\beta}{\beta + 4} + 2\beta - 2 = 0. \tag{3.2}$$

After clearing denominators, we obtain that $\beta^2 + 13\beta - 8 = 0$. This polynomial has two distinct roots neither of which lie on \mathcal{H} , and (3.2) is defined at both of these roots. These are generic conditions on the data; so since there exists a \mathbf{u} for which (3.1) has exactly two solutions, the ML-degree of \mathcal{M}_S is 2. \square

Proposition 3.2.4. *Let $S_k \subset [k] \times [k]$ be a collection of indices such that G_{S_k} is a cycle of length $2k$. Then the ML-degree of \mathcal{M}_{S_k} is k if k is odd and $(k-1)$ if k is even.*

Proof. Without loss of generality, we may assume that $S_k = \{(i, i) \mid i \in [k]\} \cup \{(i, i+1) \mid i \in [k-1]\} \cup$

$\{(k, 1)\}$. Since G_{S_k} consists of a single cycle, the ideal $I(\mathcal{M}_{S_k})$ is principal. Indeed, it is given by

$$I(\mathcal{M}_{S_k}) = \left\langle \prod_{i=1}^k p_{i,i} - \prod_{i=1}^k p_{i,i+1} \right\rangle, \quad (3.3)$$

where we set $p_{k,k+1} = p_{k,1}$. Let \mathcal{H} be the hyperplane arrangement,

$$\mathcal{H} = \left\{ \mathbf{p} \mid p_{++} \prod_{i,j \in S} p_{i,j} = 0 \right\}.$$

By Proposition 7 of [2], ML-degree of \mathcal{M}_{S_k} is the number of solutions to

$$I(\mathcal{M}_{S_k}) + \langle A(S_k)\mathbf{u} - u_+ A(S_k)\mathbf{p} \rangle. \quad (3.4)$$

that lie outside of \mathcal{H} .

The sufficient statistics of \mathbf{u} are of the form $u_{i,i} + u_{i,i+1}$ and $u_{i-1,i} + u_{i,i}$ where we set $u_{0,1} = u_{k,1}$. So computing solutions to Equation (3.4) is equivalent to solving for $\alpha \in \mathbb{C}$ in the equation

$$\prod_{i=1}^k (u_{i,i} + \alpha) - \prod_{i=1}^k (u_{i,i+1} - \alpha) = 0. \quad (3.5)$$

The MLE is then of the form $p_{i,i} = (u_{i,i} + \alpha)/u_{++}$ and $p_{i,i+1} = (u_{i,i+1} - \alpha)/u_{++}$. The degree of this polynomial is k when k is odd and $k - 1$ when k is even.

Furthermore, we claim that for generic \mathbf{u} , none of these solutions lie in \mathcal{H} . Indeed, without loss of generality, suppose that $\bar{\mathbf{p}}$ is a solution to (3.4) with $\bar{p}_{1,1} = 0$. Then we have that $\alpha = -u_{1,1}$. So the first term of (3.5) is 0. But then there exists an i such that

$$u_{i,i+1} - \alpha = u_{i,i+1} + u_{1,1} = 0,$$

which is a non-generic condition on \mathbf{u} . Similarly, since u is generic, we may assume that $u_{++} \neq 0$. But if $\bar{p}_{++} = 0$, then since each $\bar{p}_{i,i} = (u_{i,i} + \alpha)/u_{++}$ and $\bar{p}_{i,i+1} = (u_{i,i+1} - \alpha)/u_{++}$, this implies that $u_{++} = 0$, which is a contradiction. So for generic values of u , the roots of (3.5) give rise to exactly k , resp. $k - 1$, solutions to (3.4) that lie outside of \mathcal{H} . So the ML-degree of \mathcal{M}_{S_k} is k if k is odd and $k - 1$ if k is even. \square

Theorem 3.2.5. *Let S be such that G_S is not doubly chordal bipartite. Then \mathcal{M}_S does not have rational MLE.*

Proof. Suppose that G_S is not doubly chordal bipartite. Then it has an induced subgraph H that is either a double square or a cycle of length greater than or equal to 6. Without loss of generality, let the edge set $E(H)$ be a subset of $[k] \times [k]$. Let $A = A(S)$ and let \bar{A} be the submatrix of A consisting of the columns indexed by elements of $E(H)$.

Let the coordinates of P_A and $P_{\bar{A}}$ be indexed by $(x_1, \dots, x_m, y_1, \dots, y_n)$. We claim that \bar{A} is a facial

submatrix of A . Indeed, \bar{A} consists of exactly the vertices of P_A that satisfy $x_i = 0$ for $k < i \leq m$ and $y_j = 0$ for $k < j \leq n$. Since P_A is a 0/1 polytope, the inequalities $x_i \geq 0$ and $y_j \geq 0$ are valid. So this constitutes a face of P_A .

Therefore, by Propositions 3.2.3 and 3.2.4, A has a facial submatrix \bar{A} such that $\mathcal{M}_{\bar{A}}$ has ML-degree strictly greater than 1. So by Theorem 2.2.1, the ML-degree of $\mathcal{M}_A = \mathcal{M}_S$ is also strictly greater than 1, as needed. \square

3.3 The Clique Formula for the MLE

In this section we state the main result of the chapter, which gives the specific form of the rational maximum likelihood estimates for quasi-independence models when they exist. These are described in terms of the complete bipartite subgraphs of the associated graph G_S . A complete bipartite subgraph of G_S corresponds to an entirely nonzero submatrix of S . This motivates our use of the word ‘‘clique’’ in the following definition.

Definition 3.3.1. A set of indices $C = \{i_1, \dots, i_r\} \times \{j_1, \dots, j_s\}$ is a *clique* in S if $(i_\alpha, j_\beta) \in S$ for all $1 \leq \alpha \leq r$ and $1 \leq \beta \leq s$. A clique C is maximal if it is not contained in any other clique in S .

We now describe some important sets of cliques in S .

Notation 3.3.2. For every pair of indices $(i, j) \in S$, we let $\text{Max}(i, j)$ be the set of all maximal cliques in S that contain (i, j) . We let $\text{Int}(i, j)$ be the set of all containment-maximal pairwise intersections of elements of $\text{Max}(i, j)$. Similarly, we let $\text{Max}(S)$ denote the set of all maximal cliques in S and $\text{Int}(S)$ denote the set of all maximal intersections of maximal cliques in S .

Example 3.3.3. Let $m = 8$ and $n = 9$. Consider the set of indices

$$S = \{11, 12, 21, 22, 23, 28, 31, 32, 33, 34, 41, 45, 51, 56, 57, 65, 76, 86, 87, 89\},$$

where we replace (i, j) with ij for the sake of brevity. The corresponding matrix with structural zeros is

$$\begin{bmatrix} * & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 & 0 & * & 0 \\ * & * & * & * & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & * & * & 0 & * \end{bmatrix}.$$

We will use this as a running example. The bipartite graph G_S associated to S is pictured in Figure 3.2. In this figure, we use white circles to denote vertices corresponding to rows in S and black squares

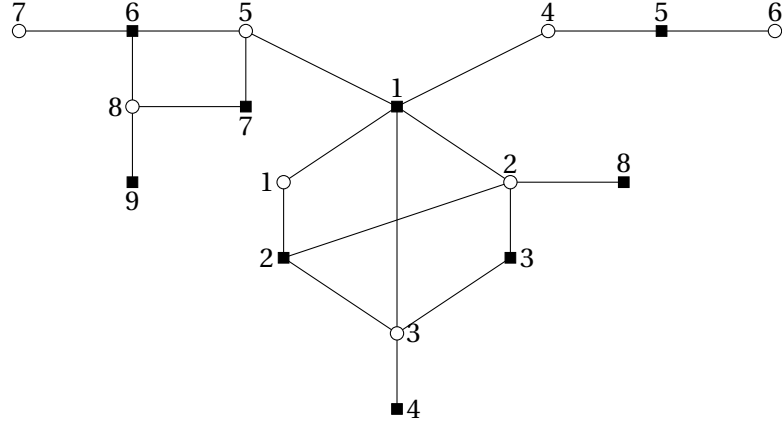


Figure 3.2 The bipartite graph associated the matrix S in Example 3.3.3

to denote vertices corresponding to columns in S . Note that G_S is doubly chordal bipartite since its only cycle of length 6 has two chords.

In this case, the set of maximal cliques in S is

$$\text{Max}(S) = \{ \{11, 21, 31, 41, 51\}, \{11, 12, 21, 22, 31, 32\}, \{21, 22, 23, 31, 32, 33\}, \{21, 22, 23, 28\}, \\ \{31, 32, 33, 34\}, \{41, 45\}, \{51, 56, 57\}, \{45, 65\}, \{56, 76, 86\}, \{56, 57, 86, 87\}, \{86, 87, 89\} \}.$$

The set of maximal intersections of maximal cliques in S is

$$\text{Int}(S) = \{ \{11, 21, 31\}, \{21, 22, 31, 32\}, \{21, 22, 23\}, \{31, 32, 33\}, \{41\}, \{51\}, \{45\}, \\ \{56, 57\}, \{56, 86\}, \{86, 87\} \}.$$

Note, for example, that $\{31, 32\}$ is the intersection of the two maximal cliques $\{11, 12, 21, 22, 31, 32\}$ and $\{31, 32, 33, 34\}$. However it is not in $\text{Int}(S)$ because it is properly contained in the intersection of maximal cliques,

$$\{11, 12, 21, 22, 31, 32\} \cap \{21, 22, 23, 31, 32, 33\} = \{21, 22, 31, 32\}.$$

Let $\mathbf{u} = (u_{ij} \mid (i, j) \in S)$ be a matrix of counts. For any $C \subset S$, we let C^+ denote the sum of all the entries of \mathbf{u} whose indices are in C . That is,

$$C^+ = \sum_{(i,j) \in C} u_{ij}.$$

Similarly, we denote the row and column marginals $u_{i+} = \sum_{j:(i,j) \in S} u_{ij}$ and $u_{+j} = \sum_{i:(i,j) \in S} u_{ij}$. The sum of all entries of \mathbf{u} is $u_{++} = \sum_{(i,j) \in S} u_{ij}$.

Theorem 3.3.4. *Let $S \subset [m] \times [n]$ be a set of indices with associated bipartite graph G_S and quasi-independence model \mathcal{M}_S . Then \mathcal{M}_S has rational maximum likelihood estimate if and only if G_S is doubly chordal bipartite. In particular, if $\mathbf{u} = (u_{ij} \mid (i, j) \in S)$ is a matrix of counts, the maximum likelihood estimate for \mathbf{u} has ij th entry*

$$\hat{p}_{ij} = \frac{u_{i+} u_{+j} \prod_{C \in \text{Int}(ij)} C^+}{u_{++} \prod_{D \in \text{Max}(ij)} D^+}$$

where the sets $\text{Max}(ij)$ and $\text{Int}(ij)$ are as in Notation 3.3.2.

Over the course of the next two sections, we prove various lemmas that ultimately allow us to prove Theorem 3.3.4.

Example 3.3.5. Consider the set of indices S from Example 3.3.3. Let \mathbf{u} be a matrix of counts. Consider the maximum likelihood estimate for the $(2, 1)$ entry, \hat{p}_{21} . The maximal cliques that contain 21 are $\{11, 21, 31, 41, 51\}$, $\{11, 12, 21, 22, 31, 32\}$, $\{21, 22, 23, 28\}$ and $\{21, 22, 23, 31, 32, 33\}$. The maximal intersections of maximal cliques that contain 21 are $\{11, 21, 31\}$, $\{21, 22, 23\}$ and $\{21, 22, 31, 32\}$. Since S is doubly chordal bipartite, we apply Theorem 3.3.4 to obtain that the numerator of \hat{p}_{21} is

$$(u_{21} + u_{22} + u_{23} + u_{28})(u_{11} + u_{21} + u_{31} + u_{41} + u_{51})(u_{11} + u_{21} + u_{31})(u_{21} + u_{22} + u_{23})(u_{21} + u_{22} + u_{31} + u_{32}).$$

The denominator of \hat{p}_{21} is

$$u_{++}(u_{11} + u_{21} + u_{31} + u_{41} + u_{51})(u_{11} + u_{12} + u_{21} + u_{22} + u_{31} + u_{32})(u_{21} + u_{22} + u_{23} + u_{28})(u_{21} + u_{22} + u_{23} + u_{31} + u_{32} + u_{33}).$$

We note that when a maximal clique is a single row or column, as is the case with $\{21, 22, 23, 28\}$ and $\{11, 21, 31, 41, 51\}$, we have cancellation between the numerator and denominator.

In order to prove Theorem 3.3.4, we show that $\hat{\mathbf{p}}$ satisfies the conditions of Birch's theorem. First, we investigate the intersections of a fixed column of the matrix with structural zeros with maximal cliques and their intersections. We prove useful lemmas about the form that these maximal cliques have that allow us to show that the conditions of Birch's Theorem are satisfied. In particular, we use them to prove Corollary 3.5.1, which states that the column marginal of the formula in Theorem 3.3.4 given by the fixed column is equal to that of the normalized data.

3.4 Intersections of Cliques with a Fixed Column

In this section we prove some results that will set the stage for the proof of Theorem 3.3.4 that appears in Section 3.5. To prove that our formulas satisfy Birch's theorem, we need to understand what happens to sums of these formulas over certain sets of indices.

Let $S \subset [m] \times [n]$ and let $j_0 \in [n]$. Without loss of generality, we assume that $(1, j_0), \dots, (r, j_0) \in S$,

and that the last $(i, j_0) \notin S$ for all $i > r$. Let

$$N_{j_0} := \{(1, j_0), \dots, (r, j_0)\}.$$

We consider j_0 to be the index of a column in the matrix representation of S , and $1, \dots, r$ to be the indices of its nonzero rows. Now let $T_0 | \dots | T_h$ be the coarsest partition of $[n]$ with the property that whenever $j, k \in T_\ell$,

$$\{i \in [r] \mid (i, j) \in S\} = \{i \in [r] \mid (i, k) \in S\}.$$

In the matrix representation of S , each T_ℓ corresponds to a set of columns whose first r rows are identical. The fact that we take $T_0 | \dots | T_h$ to be the coarsest such partition ensures that the supports of the columns in distinct parts of the partition are distinct.

Define the partition $B_0 | \dots | B_h$ of $S \cap ([r] \times [n])$ by $B_\ell = \{(i, j) \mid j \in T_\ell\}$. Note that one of the B_ℓ may be empty, in which case we exclude it from the partition. We call these B_ℓ the *blocks* of S corresponding to column j_0 . We fix j_0 and B_0, \dots, B_h for the entirety of this section, and we assume without loss of generality that $j_0 \in T_0$.

Denote by $\text{rows}^{j_0}(B_\alpha)$ the set of all $i \in [r]$ such that $(i, j) \in B_\alpha$ for some column index j . Note that this is a subset of the first $1, \dots, r$ rows of S , and that in the matrix representation of S , the columns whose indices are in B_α may not have the same zero patterns in rows $r+1, \dots, m$. Similarly, for each $j \in [n]$, define $\text{rows}^{j_0}(j)$ to be the set of all $i \in [r]$ such that $(i, j) \in S$; that is, the elements of $\text{rows}^{j_0}(j)$ are the row indices of the nonzero entries of column j in the first r rows of S . Note that the dependence on j_0 in this notation stems from the fact that the column j_0 is used to obtain the partition $B_0 | \dots | B_h$.

Example 3.4.1. Consider the running example S from Example 3.3.3, and let $j_0 = 1$ be the first column of S . In this case, $r = 5$ since only the first 5 rows entries of column j_0 are nonzero. Then the blocks associated to j_0 consist of the following columns.

$$\begin{array}{ll} T_0 = \{j_0\} = \{1\} & T_1 = \{2\} \\ T_2 = \{3\} & T_3 = \{4\} \\ T_4 = \{5\} & T_5 = \{6, 7\} \\ T_6 = \{8\} & T_7 = \{9\}. \end{array}$$

We note that although columns 6 and 7 are not the same over the whole matrix, their first five rows are the same. Since these are the nonzero rows of column j_0 , columns 6 and 7 belong to the same block.

The B_i associated to each of these sets of column indices are

$$\begin{array}{ll}
B_0 = \{11, 21, 31, 41, 51\} & B_1 = \{12, 22, 32\} \\
B_2 = \{23, 33\} & B_3 = \{34\} \\
B_4 = \{45\} & B_5 = \{56, 57\} \\
B_6 = \{28\} & B_7 = \emptyset
\end{array}$$

For instance, $\text{rows}^{j_0}(B_1) = \{1, 2, 3\}$ and $\text{rows}^{j_0}(B_5) = \{5\}$.

The following proposition characterizes what configurations of the rows of the B_α 's are allowable in order to avoid a cycle with exactly one chord. We call the condition outlined in Proposition 3.4.2 the double-squarefree, or *DS-free* condition.

Proposition 3.4.2 (DS-free condition). *Let S be doubly chordal bipartite. Let $\alpha, \beta \in [h]$. If $\text{rows}^{j_0}(B_\alpha) \cap \text{rows}^{j_0}(B_\beta)$ is nonempty, then $\text{rows}^{j_0}(B_\alpha) \subset \text{rows}^{j_0}(B_\beta)$ or $\text{rows}^{j_0}(B_\beta) \subset \text{rows}^{j_0}(B_\alpha)$.*

Proof. For the sake of contradiction, suppose without loss of generality that $\text{rows}^{j_0}(B_1) \cap \text{rows}^{j_0}(B_2)$ is nonempty but neither is contained in the other. Then let $i_0, i_1 \in \text{rows}^{j_0}(B_1)$ and $i_1, i_2 \in \text{rows}^{j_0}(B_2)$ so that $i_0 \notin \text{rows}^{j_0}(B_2)$ and $i_2 \notin \text{rows}^{j_0}(B_1)$. We have $i_0, i_1, i_2 \in \text{rows}^{j_0}(j_0)$ by definition. Let $j_1 \in T_1$ and $j_2 \in T_2$. Then the $\{i_0, i_1, i_2\} \times \{j_0, j_1, j_2\}$ submatrix of S is the matrix of a double-square, which contradicts that S is doubly chordal bipartite. \square

Proposition 3.4.2 implies that the sets $\text{rows}^{j_0}(B_\alpha)$ over all α have a tree structure ordered by containment. In fact, we will see that this gives a tree structure on the maximal cliques in S that intersect N_{j_0} . (Recall that $N_{j_0} = \{(i, j_0) \in S\} = [r] \times \{j_0\}$).

Example 3.4.3. The matrix S from Example 3.3.3 is doubly chordal bipartite, and as such, satisfies the DS-free condition. If we append a tenth column, $(0, *, *, *, 0, 0, 0, 0)^T$ to obtain a matrix S' , this introduces a new block B_8 which just contains column 10. This matrix violates the DS-free condition since $\text{rows}^{j_0}(B_1) = \{1, 2, 3\}$ and $\text{rows}^{j_0}(B_8) = \{2, 3, 4\}$. Their intersection is nonempty, but neither is contained in another. Indeed, the $\{1, 2, 4\} \times \{1, 2, 10\}$ submatrix of S' is the matrix of a double-square.

For each pair of indices i, j such that $(i, j) \in S$, let x_{ij} be the polynomial obtained from \hat{p}_{ij} by simultaneously clearing the denominators of all \hat{p}_{kl} . That is, to obtain x_{ij} , we multiply \hat{p}_{ij} by $u_{++} \prod_{D \in \text{Max}(S)} D^+$ so that

$$x_{ij} = u_{i+} u_{+j} \prod_{C \in \text{Int}(ij)} C^+ \prod_{D \in \text{Max}(S) \setminus \text{Max}(ij)} D^+.$$

Our main goal in this section is to derive a formula for the sum,

$$\sum_{i \in \text{rows}^{j_0}(B_\alpha)} x_{ij_0}.$$

This is the content of Lemma 3.4.9. This formula allows us to verify that the j_0 column marginal of $\hat{\mathbf{p}}$ matches that of the normalized data. In order to simplify this sum, we must first understand how maximal cliques and their intersections intersect N_{j_0} . For each B_α with $\alpha \in [h]$ and $\text{rows}^{j_0}(B_\alpha) \neq \emptyset$, we let D_α be the clique,

$$D_\alpha = \{(i, j) \mid i \in \text{rows}^{j_0}(B_\alpha) \text{ and } \text{rows}^{j_0}(B_\alpha) \subset \text{rows}^{j_0}(j)\}.$$

In other words, D_α is the largest clique that contains B_α and intersects N_{j_0} . We call D_α the *clique induced by B_α* .

Example 3.4.4. Consider our running example S with $j_0 = 1$ and blocks B_0, \dots, B_7 as described in Example 3.4.1. Then $N_{j_0} = N_1 = \{11, 21, 31, 41, 51\}$. The cliques induced by B_0, \dots, B_7 are

$$D_0 = \{11, 21, 31, 41, 51\},$$

$$D_1 = \{11, 12, 21, 22, 31, 32\},$$

$$D_2 = \{21, 22, 23, 31, 32, 33\},$$

$$D_3 = \{31, 32, 33, 34\},$$

$$D_4 = \{41, 45\},$$

$$D_5 = \{51, 56, 57\}, \text{ and}$$

$$D_6 = \{21, 22, 23, 28\}.$$

There is no D_7 since the block B_7 is empty. Note that these are exactly the maximal cliques in S that intersect N_1 . The next proposition proves that this is the case for all DS-free matrices with structural zeros.

We note that when D_α is the clique induced by B_α , all of the nonzero rows of D_α lie in $[r]$ by definition of an induced clique. We continue to use the notation $\text{rows}^{j_0}(D_\alpha)$ since the formation of the set B_α depends on the specified column j_0 . For any clique C , let $\text{cols}(C) = \{j \mid (i, j) \in C \text{ for some } i\}$.

Proposition 3.4.5. *For all $\alpha \in [h]$, D_α is a maximal clique. Furthermore, any maximal clique that has nonempty intersection with N_{j_0} is induced by some B_α .*

Proof. We will show that D_α is maximal by showing that we cannot add any rows or columns to it. We cannot add any columns to D_α by definition. We cannot add any of rows $1, \dots, r$ to D_α since all nonzero rows of B_α are already contained in D_α . We cannot add any of rows $r+1, \dots, m$ to D_α since j_0 is a column of D_α whose entries in rows $r+1, \dots, m$ are zero. Note that if we can add one element (i, j) to D_α , then by definition of a clique, we must either be able to add all of $\{i\} \times \text{cols}(D_\alpha)$ or $\text{rows}^{j_0}(D_\alpha) \times \{j\}$ to the clique. Since we cannot add any rows or columns to D_α , it is a maximal clique.

Now let D be a maximal clique that intersects N_{j_0} . For the sake of contradiction, suppose that $D \neq D_\alpha$ for each $\alpha \in [h]$.

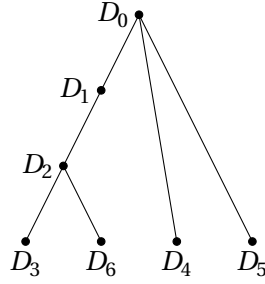


Figure 3.3 The poset $P(j_0)$ for Example 3.4.7

Let j_1 be a column in D such that $\text{rows}^{j_0}(j_1)$ is minimal among all columns of D . We must have that $(i, j_1) \in B_\alpha$ for some $i \in [r]$ and $\alpha \in [h]$. Since $D \neq D_\alpha$, it must be the case that column j_1 has a nonzero row $i_1 \in [r]$ that is not in D . Since D is maximal, there must exist another column j_2 in D that has a zero in row i_1 . Therefore, we have that $\text{rows}^{j_0}(j_1) \not\subset \text{rows}^{j_0}(j_2)$. Furthermore, $\text{rows}^{j_0}(j_2) \not\subset \text{rows}^{j_0}(j_1)$ by the minimality of j_1 . But since D is nonempty, the intersection of $\text{rows}^{j_0}(j_1)$ and $\text{rows}^{j_0}(j_2)$ must be nonempty. This contradicts Proposition 3.4.2, as needed. \square

Proposition 3.4.5 shows that the maximal cliques that intersect N_{j_0} are exactly the cliques that are induced by some B_α . The DS-free condition gives a poset structure on the set of these maximal cliques D_0, \dots, D_h that intersect N_{j_0} nontrivially.

Definition 3.4.6. Let $P(j_0)$ denote the poset with ground set $\{D_0, \dots, D_h\}$ and $D_\alpha \leq D_\beta$ if and only if $\text{rows}^{j_0}(D_\alpha) \subset \text{rows}^{j_0}(D_\beta)$.

Recall that for a poset P and two elements of its ground set, $p, q \in P$, we say that q covers p if $p < q$ and for any $r \in P$, if $p \leq r \leq q$, then $r = p$ or $r = q$. We denote such a cover relation by $p \triangleleft q$. The *Hasse diagram* of a poset is a directed acyclic graph on P with an edge from p to q whenever $p \triangleleft q$. In the case of $P(j_0)$, the Hasse diagram of this poset is a tree since the DS-free condition implies that any D_α is covered by at most one maximal clique.

Example 3.4.7. In our running example S with $j_0 = 1$ and blocks B_0, \dots, B_6 and associated cliques D_0, \dots, D_6 , the Hasse diagram of the poset $P(j_0)$ is pictured in Figure 3.3.

The next proposition shows that the cover relations in this poset, denoted $D_\alpha \triangleleft D_\beta$ correspond to maximal intersections of maximal cliques that intersect N_{j_0} nontrivially. Denote by $\text{cols}(D_\alpha)$ the nonzero columns of the clique D_α . We note that if $D_\alpha \triangleleft D_\beta$, then $\text{cols}(D_\beta) \subset \text{cols}(D_\alpha)$. In particular, this means that if $C = D_\alpha \cap D_\beta$, then $C = \text{rows}^{j_0}(D_\alpha) \times \text{cols}(D_\beta)$.

Proposition 3.4.8. Let $C = D_\alpha \cap D_\beta$. Then C is maximal among all pairwise intersections of maximal cliques if and only if $D_\alpha \triangleleft D_\beta$ or $D_\beta \triangleleft D_\alpha$ in $P(j_0)$.

Proof. Suppose without loss of generality that $D_\alpha \triangleleft D_\beta$ in $P(j_0)$. For the sake of contradiction, suppose that $D_\alpha \cap D_\beta \notin \text{Int}(S)$. Then there exists another maximal clique that contains $D_\alpha \cap D_\beta$. By Proposition

3.4.5 and the fact that $D_\alpha \cap D_\beta$ intersects N_{j_0} nontrivially, we can write this maximal clique as D_γ for some $\gamma \in [h]$.

Note that we have $\text{rows}^{j_0}(C) = \text{rows}^{j_0}(D_\alpha)$ and $\text{cols}(C) = \text{cols}(D_\beta)$. Therefore $C = \text{rows}^{j_0}(D_\alpha) \times \text{cols}(D_\beta)$. So $\text{rows}^{j_0}(D_\alpha) \subsetneq \text{rows}^{j_0}(D_\gamma)$ and $\text{cols}(D_\beta) \subsetneq \text{cols}(D_\gamma)$. In particular, this second inclusion implies that $\text{rows}^{j_0}(D_\gamma) \subsetneq \text{rows}^{j_0}(D_\beta)$. Indeed, suppose that $i \leq r$ is a row of D_γ that is not a row of D_β . Then there exists a column j of D_β for which $(i, j) \notin S$. But since j is also a column of D_γ , this contradicts that D_γ is a clique. So we have the proper containments

$$\text{rows}^{j_0}(D_\alpha) \subsetneq \text{rows}^{j_0}(D_\gamma) \subsetneq \text{rows}^{j_0}(D_\beta),$$

which contradicts that $D_\alpha \triangleleft D_\beta$ in $P(j_0)$. So $D_\alpha \cap D_\beta$ must be maximal.

Now let $C = D_\alpha \cap D_\beta \in \text{Int}(S)$. For the sake of contradiction, suppose that D_α does not cover D_β or vice versa. Since C is nonempty, without loss of generality we must have $\text{rows}^{j_0}(D_\alpha) \subset \text{rows}^{j_0}(D_\beta)$ by the DS-free condition. So there exists a D_γ such that $D_\alpha < D_\gamma < D_\beta$ in $P(j_0)$. Therefore we have that

$$\text{rows}^{j_0}(D_\alpha) \subsetneq \text{rows}^{j_0}(D_\gamma) \subsetneq \text{rows}^{j_0}(D_\beta).$$

Let $(i, j) \in C$. Then i is a row of D_α , so it is a row of D_γ . Furthermore, since j is a column of D_β , $\text{rows}^{j_0}(D_\beta) \subset \text{rows}^{j_0}(j)$. So $\text{rows}^{j_0}(D_\gamma) \subset \text{rows}^{j_0}(j)$ and j is a column of D_γ . Therefore, $C \subsetneq D_\gamma \cap D_\beta$. This containment is proper since $\text{rows}^{j_0}(D_\alpha) \subsetneq \text{rows}^{j_0}(D_\gamma)$. So we have contradicted that C is maximal. \square

We can now state the key lemma regarding the sum of the x_{i, j_0} s over $\{i : (i, j_0) \in D_\alpha\}$ for any $\alpha \in [h]$.

Lemma 3.4.9. *Let S be DS-free and let D_α be a maximal clique that intersects N_{j_0} . Then*

$$\sum_{i \in \text{rows}^{j_0}(D_\alpha)} x_{i, j_0} = u_{+, j_0} \left(\prod_{\substack{C \in \text{Int}(S) \\ C \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D_\alpha) \subset \text{rows}^{j_0}(C)}} C^+ \right) \left(\prod_{\substack{D \in \text{Max}(S) \\ D \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D) \subset \text{rows}^{j_0}(D_\alpha)}} D^+ \right) \left(\prod_{\substack{E \in \text{Max}(S) \\ N_{j_0} \cap D_\alpha \cap E = \emptyset}} E^+ \right) \quad (3.6)$$

In order to prove this, we will sum the entries x_{i, j_0} over all $i \in \text{rows}^{j_0}(D_\beta)$ for each β . We will do this inductively from the bottom of $P(j_0)$. The key idea of this induction is as follows.

Remark 3.4.10. If $D_{\alpha_1}, \dots, D_{\alpha_\ell}$ are covered by D_β in $P(j_0)$, then the rows of $N_{j_0} \cap D_\beta$ are partitioned by each $N_{j_0} \cap D_{\alpha_k}$ along with the set of rows that are in D_β and not in any D_{α_k} . The fact that this is a partition follows from the DS-free condition. Therefore, summing the x_{i, j_0} that belong to each clique covered by D_β and adding in the x_{i, j_0} s for rows i that are not in any clique covered by D_β will give us the sum of x_{i, j_0} over all $i \in \text{rows}^{j_0}(D_\beta)$.

The next proposition focuses on the factors of the right-hand side of Equation (3.6) that correspond to elements of $\text{Int}(S)$. It will be used to show that when we perform the induction and move upwards by one cover relation from D_α to D_β in the poset $P(j_0)$, all but one of these factors stays the

same. The only one that no longer appears in the product corresponds to the maximal intersection $D_\alpha \cap D_\beta$.

Proposition 3.4.11. *Let $D_\alpha \triangleleft D_\beta$ in $P(j_0)$. Let $C \in \text{Int}(S)$ intersect N_{j_0} nontrivially so that $\text{rows}^{j_0}(D_\alpha) \subset \text{rows}^{j_0}(C)$. Then either $C = D_\alpha \cap D_\beta$ or $\text{rows}^{j_0}(D_\beta) \subset \text{rows}^{j_0}(C)$.*

Proof. Without loss of generality, let $C = D_1 \cap D_2$. Proposition 3.4.8 tells us that C must be of this form. By the same proposition, we may assume without loss of generality that $D_2 \triangleleft D_1$, so $\text{rows}^{j_0}(C) = \text{rows}^{j_0}(D_2)$. Suppose that $\text{rows}^{j_0}(D_\beta) \not\subset \text{rows}^{j_0}(D_2)$. Since $\text{rows}^{j_0}(D_\alpha) \subset \text{rows}^{j_0}(D_2)$ and $\text{rows}^{j_0}(D_\alpha) \subset \text{rows}^{j_0}(D_\beta)$, we must have that $\text{rows}^{j_0}(D_2) \cap \text{rows}^{j_0}(D_\beta)$ is nonempty. So by the DS-free condition, $\text{rows}^{j_0}(D_2) \subsetneq \text{rows}^{j_0}(D_\beta)$. So we have the chain of inclusions,

$$\text{rows}^{j_0}(D_\alpha) \subset \text{rows}^{j_0}(D_2) \subsetneq \text{rows}^{j_0}(D_\beta).$$

But since D_β covers D_α in $P(j_0)$, and every element of $P(j_0)$ is covered by at most one element, this implies that $\alpha = 1$ and $\beta = 2$, so $C = D_\alpha \cap D_\beta$, as needed. \square

The following proposition focuses on the factors of the right-hand side of Equation (3.6) that correspond to elements of $\text{Max}(S)$. It gives a correspondence between the factors of this product for D_α and all but one of the factors of this product for D_β when we have the cover relation $D_\alpha \triangleleft D_\beta$ in $P(j_0)$.

Proposition 3.4.12. *For any $D_\alpha, D_\beta \in P(j_0)$, define the following sets:*

$$\begin{aligned} R_\alpha &= \{D \in \text{Max}(S) \mid D \cap N_{j_0} \neq \emptyset, \text{rows}^{j_0}(D) \subset \text{rows}^{j_0}(D_\alpha)\} \cup \{E \in \text{Max}(S) \mid N_{j_0} \cap D_\alpha \cap E = \emptyset\} \\ \bar{R}_\beta &= \{D \in \text{Max}(S) \mid D \cap N_{j_0} \neq \emptyset, \text{rows}^{j_0}(D) \subsetneq \text{rows}^{j_0}(D_\beta)\} \cup \{E \in \text{Max}(S) \mid N_{j_0} \cap D_\beta \cap E = \emptyset\}. \end{aligned}$$

If $D_\alpha \triangleleft D_\beta$ in $P(j_0)$, then $R_\alpha = \bar{R}_\beta$.

Proof. First let $D \in R_\alpha$. If $\text{rows}^{j_0}(D) \subset \text{rows}^{j_0}(D_\alpha)$ and $D \cap N_{j_0} \neq \emptyset$, then since $\text{rows}^{j_0}(D_\alpha) \subsetneq \text{rows}^{j_0}(D_\beta)$, we have that $\text{rows}^{j_0}(D) \subsetneq \text{rows}^{j_0}(D_\beta)$. So $D \in \bar{R}_\beta$.

Otherwise, we have $N_{j_0} \cap D_\alpha \cap D = \emptyset$. There are now two cases.

Case 1: If $N_{j_0} \cap D = \emptyset$, then $N_{j_0} \cap D_\beta \cap D = \emptyset$ as well. So $D \in \bar{R}_\beta$.

Case 2: Suppose that $N_{j_0} \cap D \neq \emptyset$ and $D_\alpha \cap D = \emptyset$. If $D_\beta \cap D$ is empty as well, then $D \in \bar{R}_\beta$.

Otherwise, suppose $D_\beta \cap D \neq \emptyset$. Then we must have that $\text{rows}^{j_0}(D) \subsetneq \text{rows}^{j_0}(D_\beta)$ by the fact that $\text{rows}^{j_0}(D_\beta) \not\subset \text{rows}^{j_0}(D)$ and the DS-free condition. So $D \in \bar{R}_\beta$ in this case as well. Note that it is never the case that $N_{j_0} \cap D \neq \emptyset$ and $D_\alpha \cap D \neq \emptyset$ but $N_{j_0} \cap D \cap D_\alpha = \emptyset$ since j_0 is a column of D_α . So we have shown that $R_\alpha \subset \bar{R}_\beta$.

Now let $D \in \bar{R}_\beta$. We have two cases again.

Case 1: First, consider the case in which $\text{rows}^{j_0}(D) \subsetneq \text{rows}^{j_0}(D_\beta)$ and $D \cap N_{j_0} \neq \emptyset$. If $\text{rows}^{j_0}(D) \subset \text{rows}^{j_0}(D_\alpha)$, then $D \in R_\alpha$, as needed. Otherwise, by the DS-free condition, there are two cases.

Case 1a: If $\text{rows}^{j_0}(D_\alpha) \subsetneq \text{rows}^{j_0}(D)$, then we have the chain of containments,

$$\text{rows}^{j_0}(D_\alpha) \subsetneq \text{rows}^{j_0}(D) \subsetneq \text{rows}^{j_0}(D_\beta),$$

which contradicts that $D_\alpha \triangleleft D_\beta$ in $P(j_0)$. So this case cannot actually occur.

Case 1b: If $\text{rows}^{j_0}(D_\alpha) \cap \text{rows}^{j_0}(D) = \emptyset$, then we have that $N_{j_0} \cap D_\alpha \cap D = \emptyset$. Therefore, $D \in R_\alpha$, as needed.

Case 2: The final case is when $N_{j_0} \cap D_\beta \cap D = \emptyset$. In this case, since $\text{rows}^{j_0}(D_\alpha) \subset \text{rows}^{j_0}(D_\beta)$, we have that $N_{j_0} \cap D_\alpha \cap D = \emptyset$ as well. So $D \in R_\alpha$. So we have shown that $\bar{R}_\beta \subset R_\alpha$, as needed. \square

Remark 3.4.13. Note that Proposition 3.4.12 implies that whenever D_α and D_β are covered by the same element of $P(j_0)$, we have that $R_\alpha = R_\beta$. This shows that the left-hand side of Equation (3.6) for D_α and D_β consist of the same terms that come from cliques in $\text{Max}(S)$.

Let $D_{\alpha_1}, \dots, D_{\alpha_\ell} \triangleleft D_\beta$. As we discussed in Remark 3.4.10, in order to sum the values of $x_{i j_0}$ over $N_{j_0} \cap D_\beta$, we must understand the sum over $x_{i j_0}$ for those rows i such that $i \in \text{rows}^{j_0}(D_\beta)$ but $i \notin \text{rows}^{j_0}(D_{\alpha_k})$ for all k . The following proposition concerns the sum of the $x_{i j_0}$ over these values of i .

Proposition 3.4.14. *Let $D_{\alpha_1}, \dots, D_{\alpha_\ell} \triangleleft D_\beta$. Let r_1, \dots, r_a be the rows of D_β that are not in any D_{α_k} for $k = 1, \dots, \ell$. Then*

$$\sum_{i=1}^a x_{r_i j_0} = u_{+j_0} \left(\prod_{\substack{C \in \text{Int}(S) \\ C \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D_\beta) \subset \text{rows}^{j_0}(C)}} C^+ \right) \left(\prod_{\substack{D \in \text{Max}(S) \\ D \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D) \subsetneq \text{rows}^{j_0}(D_\beta)}} D^+ \right) \left(\prod_{\substack{E \in \text{Max}(S) \\ N_{j_0} \cap D_\beta \cap E = \emptyset}} E^+ \right) \left(\sum_{i=1}^a u_{r_i} \right) \quad (3.7)$$

Proof. Without loss of generality, we will let $D_2, \dots, D_\ell \triangleleft D_1$, and let rows $1, \dots, a$ be the rows of D_1 that are not rows of any D_α for $\alpha = 2, \dots, \ell$. Let $i \in [a]$. Recall that

$$x_{i j_0} = u_{i+} u_{+j_0} \prod_{C \in \text{Int}(i j_0)} C^+ \prod_{D \in \text{Max}(S) \setminus \text{Max}(i j_0)} D^+.$$

We first consider the maximum cliques D with $(i, j_0) \notin D$. If $N_{j_0} \cap D_1 \cap D = \emptyset$, then D^+ is a term of $x_{i j_0}$ for all $i = 1, \dots, a$. So D^+ is a factor of both the left-hand and right-hand sides of Equation (3.7).

Otherwise, we have $N_{j_0} \cap D_1 \cap D \neq \emptyset$. In particular, this means that $N_{j_0} \cap D \neq \emptyset$. Since $\text{rows}^{j_0}(D_1) \cap \text{rows}^{j_0}(D) \neq \emptyset$, and $(i, j_0) \notin D$, we must have $\text{rows}^{j_0}(D) \subsetneq \text{rows}^{j_0}(D_1)$ by the DS-free condition. Furthermore, for all maximal cliques D with $\text{rows}^{j_0}(D) \subsetneq \text{rows}^{j_0}(D_1)$, we have $(i, j_0) \notin D$. Indeed, $D \cap N_{j_0} \neq \emptyset$, so by Proposition 3.4.5, we have $D = D_\gamma$ for some γ with $D_\gamma \triangleleft D_1$ in $P(j_0)$. So $\text{rows}^{j_0}(D_\gamma) \subset \text{rows}^{j_0}(D_\alpha)$ for some $\alpha \in \{2, \dots, \ell\}$. Since $(i, j_0) \notin D_\alpha$, we have $(i, j_0) \notin D_\gamma$ as well.

Therefore, the factors D^+ corresponding to maximal cliques in each $x_{i j_0}$ are the same for all $i \in [a]$, and are exactly those with $\text{rows}^{j_0}(D) \subsetneq \text{rows}^{j_0}(D_1)$ or $N_{j_0} \cap D_1 \cap D = \emptyset$.

Now let $(i, j_0) \in C$ where $C \in \text{Int}(S)$ and $C \cap N_{j_0} \neq \emptyset$. By Proposition 3.4.8, we have $C = D_\gamma \cap D_\delta$ where $D_\gamma \triangleleft D_\delta$ in $P(j_0)$. Since $C \cap D_1$ is nonempty, we must have that $\text{rows}^{j_0}(D_\gamma) \subset \text{rows}^{j_0}(D_1)$ or

$\text{rows}^{j_0}(D_1) \subset \text{rows}^{j_0}(D_\gamma)$, and similarly for D_δ . But since $i \in [a]$, $(i, j_0) \notin D_\alpha$ for any $D_\alpha < D_1$. So we must have $\text{rows}^{j_0}(D_1) \subset \text{rows}^{j_0}(D_\gamma) \subset \text{rows}^{j_0}(D_\delta)$. Therefore, $\text{rows}^{j_0}(D_1) \subset \text{rows}^{j_0}(C)$.

Furthermore, we have that $(i, j_0) \in C$ for all $C \in \text{Int}(S)$ with $\text{rows}^{j_0}(D_1) \subset \text{rows}^{j_0}(C)$. So the factors C^+ corresponding to maximal intersections of maximal cliques are exactly those with $\text{rows}^{j_0}(D_1) \subset \text{rows}^{j_0}(C)$ in each $x_{i j_0}$.

Therefore, we have that

$$\begin{aligned} \sum_{i=1}^a x_{i j_0} &= \sum_{i=1}^a u_{i+} u_{+j_0} \left(\prod_{C \in \text{Int}(i j_0)} C^+ \right) \left(\prod_{D \in \text{Max}(S) \setminus \text{Max}(i j_0)} D^+ \right) \\ &= \left(\prod_{\substack{C \in \text{Int}(S) \\ C \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D_1) \subset \text{rows}^{j_0}(C)}} C^+ \right) \left(\prod_{\substack{D \in \text{Max}(S) \\ D \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D) \subsetneq \text{rows}^{j_0}(D_1)}} D^+ \right) \left(\prod_{\substack{E \in \text{Max}(S) \\ N_{j_0} \cap D_1 \cap E = \emptyset}} E^+ \right) \left(\sum_{i=1}^a u_{i+} u_{+j_0} \right) \\ &= u_{+j_0} \left(\prod_{\substack{C \in \text{Int}(S) \\ C \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D_1) \subset \text{rows}^{j_0}(C)}} C^+ \right) \left(\prod_{\substack{D \in \text{Max}(S) \\ D \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D) \subsetneq \text{rows}^{j_0}(D_1)}} D^+ \right) \left(\prod_{\substack{E \in \text{Max}(S) \\ N_{j_0} \cap D_1 \cap E = \emptyset}} E^+ \right) \left(\sum_{i=1}^a u_{i+} \right), \end{aligned}$$

as needed. \square

Finally, the following proposition gives a way to write D_β^+ as a sum over its intersections with the elements of $P(j_0)$ that it covers, along with the rows of D_β that are not rows of any clique that it covers.

Proposition 3.4.15. *Let $D_{\alpha_1}, \dots, D_{\alpha_\ell} < D_\beta$. Let r_1, \dots, r_a be the rows of D_β that are not in any D_{α_i} for $i = 1, \dots, \ell$. Then*

$$D_\beta^+ = \sum_{i=1}^a u_{r_i+} + \sum_{i=1}^{\ell} (D_{\alpha_i} \cap D_\beta)^+. \quad (3.8)$$

Proof. Without loss of generality, we will let $D_2, \dots, D_\ell < D_1$, and let rows $1, \dots, a$ be the rows of D_1 that are not rows of any D_α for $\alpha = 2, \dots, \ell$. First note that each $u_{i j}$ that appears on the right-hand side of Equation (3.8) is a term of D_1^+ . Indeed, if $(i, j) \in D_\alpha \cap D_1$ for some $\alpha = 2, \dots, \ell$, this is clear.

Otherwise, we have $i \in [a]$. For the sake of contradiction, suppose that there exists a column j so that $(i, j) \notin D_1$ but $(i, j) \in S$. But then $\text{rows}^{j_0}(D_1) \cap \text{rows}^{j_0}(j)$ is non-empty. So by the DS-free condition, either $\text{rows}^{j_0}(D_1) \subset \text{rows}^{j_0}(j)$ or $\text{rows}^{j_0}(j) \subsetneq \text{rows}^{j_0}(D_1)$. If $\text{rows}^{j_0}(D_1) \subset \text{rows}^{j_0}(j)$, then j is a column of D_1 by definition, which is a contradiction. If $\text{rows}^{j_0}(j) \subsetneq \text{rows}^{j_0}(D_1)$, then column j belongs to some block B_η with $\text{rows}^{j_0}(D_\eta) \subsetneq \text{rows}^{j_0}(D_1)$. But this contradicts that row i is not in any D_α for $\alpha = 2, \dots, \ell$.

Now it remains to show that all the terms in D_1^+ appear in the right-hand side of Equation (3.8). Let $(i, j) \in D_1$. If $i \in [a]$, then $u_{i j}$ is a term in the right-hand side, as needed. Otherwise, $i \in \text{rows}^{j_0}(D_\alpha)$ for some $\alpha \in \{2, \dots, \ell\}$. Since $\text{cols}(D_1) \subset \text{cols}(D_\alpha)$ by definition, we must have $j \in \text{cols}(D_\alpha)$. So $(i, j) \in D_\alpha$. Therefore, $u_{i j}$ is a term in $(D_\alpha \cap D_1)^+$. Finally, since $D_\gamma \cap D_\delta = \emptyset$ for all $\gamma, \delta \in \{2, \dots, \ell\}$ with $\gamma \neq \delta$, no term is repeated. \square

We can now use these propositions to prove Lemma 3.4.9.

Proof of Lemma 3.4.9. We will induct over the poset $P(j_0)$. For the base case, we let D_α be minimal in $P(j_0)$. First, by Proposition 3.4.15 we have that

$$D_\alpha^+ = \sum_{i \in \text{rows}^{j_0}(D_\alpha)} u_{i^+}$$

since D_α does not cover any element of $P(j_0)$.

Let $C \in \text{Int}(S)$ with $\text{rows}^{j_0}(D_\alpha) \subset \text{rows}^{j_0}(C)$ such that $C \cap N_{j_0} \neq \emptyset$. Then for any $i \in \text{rows}^{j_0}(D_\alpha)$, $(i, j_0) \in C$. So $C \in \text{Int}(i j_0)$ and C^+ is a factor of $x_{i j_0}$.

If $C \in \text{Int}(i j_0)$, then $N_{j_0} \cap C \neq \emptyset$. It remains to be shown that all factors of $x_{i j_0}$, C^+ corresponding to maximal intersections of maximal cliques have $\text{rows}^{j_0}(D_\alpha) \subset \text{rows}^{j_0}(C)$.

Let $(i, j_0) \in D_\alpha$. Let D_β and D_γ be maximal cliques such that $C = D_\beta \cap D_\gamma \in \text{Int}(i j_0)$. Then we have $\text{rows}^{j_0}(D_\alpha) \cap \text{rows}^{j_0}(D_\beta)$ and $\text{rows}^{j_0}(D_\alpha) \cap \text{rows}^{j_0}(D_\gamma)$ nonempty. Since D_α is minimal, this implies that $\text{rows}^{j_0}(D_\alpha) \subset \text{rows}^{j_0}(D_\beta)$, $\text{rows}^{j_0}(D_\alpha) \subset \text{rows}^{j_0}(D_\gamma)$. So $\text{rows}^{j_0}(D_\alpha) \subset \text{rows}^{j_0}(C)$. Therefore the factors of each $x_{i j_0}$ with $(i, j_0) \in D_\alpha$ that correspond to maximal intersections of maximal cliques are

$$\prod_{\substack{C \in \text{Int}(S) \\ C \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D_\alpha) \subset \text{rows}^{j_0}(C)}} C^+,$$

as needed. The other factors of each $x_{i j_0}$ for $(i, j_0) \in D_\alpha$ are of the form

$$\prod_{\substack{E \in \text{Max}(S) \\ (i, j_0) \notin E}} E^+.$$

Since all $(i, j_0) \in D_\alpha$ are contained in the same maximal cliques when D_α is minimal in $P(j_0)$, the terms corresponding to maximal cliques in each $x_{i j_0}$ are of the form

$$\prod_{\substack{E \in \text{Max}(S) \\ N_{j_0} \cap D_\alpha \cap E = \emptyset}} E^+,$$

as needed. So we have that

$$\begin{aligned}
\sum_{(i,j_0) \in D_\alpha} x_{i j_0} &= \sum_{(i,j_0) \in D_\alpha} u_{i+} u_{+j_0} \left(\prod_{\substack{C \in \text{Int}(S) \\ C \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D_\alpha) \subset \text{rows}^{j_0}(C)}} C^+ \right) \left(\prod_{\substack{E \in \text{Max}(S) \\ N_{j_0} \cap D_\alpha \cap E = \emptyset}} E^+ \right) \\
&= u_{+j_0} \left(\sum_{i \in \text{rows}^{j_0}(D_\alpha)} u_{i+} \right) \left(\prod_{\substack{C \in \text{Int}(S) \\ C \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D_\alpha) \subset \text{rows}^{j_0}(C)}} C^+ \right) \left(\prod_{\substack{E \in \text{Max}(S) \\ N_{j_0} \cap D_\alpha \cap E = \emptyset}} E^+ \right) \\
&= u_{+j_0} D_\alpha^+ \left(\prod_{\substack{C \in \text{Int}(S) \\ C \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D_\alpha) \subset \text{rows}^{j_0}(C)}} C^+ \right) \left(\prod_{\substack{E \in \text{Max}(S) \\ N_{j_0} \cap D_\alpha \cap E = \emptyset}} E^+ \right).
\end{aligned}$$

Since D_α is the only maximal clique whose rows are contained in D_α , we have that

$$D_\alpha^+ = \prod_{\substack{D \in \text{Max}(S) \\ D \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D) \subset \text{rows}^{j_0}(D_\alpha)}} D^+.$$

So the lemma holds for the base case.

Without loss of generality, let $D_2, \dots, D_\ell < D_1$ in $P(j_0)$. Let rows $1, \dots, a$ be the rows of D_1 that are not in any D_α with $\alpha = 2, \dots, \ell$. We have the following chain of equalities.

$$\begin{aligned}
\sum_{(i,j_0) \in D_1} x_{i j_0} &= u_{+j_0} \sum_{i=1}^a x_{i j_0} + \sum_{\alpha=2}^{\ell} \sum_{(i,j_0) \in D_\alpha} x_{i j_0} \\
&= u_{+j_0} \left(\prod_{\substack{C \in \text{Int}(S) \\ C \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D_\beta) \subset \text{rows}^{j_0}(C)}} C^+ \right) \left(\prod_{\substack{D \in \text{Max}(S) \\ D \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D) \subsetneq \text{rows}^{j_0}(D_\beta)}} D^+ \right) \left(\prod_{\substack{E \in \text{Max}(S) \\ N_{j_0} \cap D_\beta \cap E = \emptyset}} E^+ \right) \left(\sum_{i=1}^p u_{i+} \right) \\
&\quad + \sum_{\alpha=2}^{\ell} \sum_{(i,j_0) \in D_\alpha} x_{i j_0} \\
&= u_{+j_0} \left(\prod_{\substack{C \in \text{Int}(S) \\ C \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D_\beta) \subset \text{rows}^{j_0}(C)}} C^+ \right) \left(\prod_{\substack{D \in \text{Max}(S) \\ D \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D) \subsetneq \text{rows}^{j_0}(D_\beta)}} D^+ \right) \left(\prod_{\substack{E \in \text{Max}(S) \\ N_{j_0} \cap D_\beta \cap E = \emptyset}} E^+ \right) \left(\sum_{i=1}^a u_{i+} \right) \\
&\quad + \sum_{\alpha=2}^{\ell} u_{+j_0} \left(\prod_{\substack{C \in \text{Int}(S) \\ C \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D_\alpha) \subset \text{rows}^{j_0}(C)}} C^+ \right) \left(\prod_{\substack{D \in \text{Max}(S) \\ D \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D) \subset \text{rows}^{j_0}(D_\alpha)}} D^+ \right) \left(\prod_{\substack{E \in \text{Max}(S) \\ N_{j_0} \cap D_\alpha \cap E = \emptyset}} E^+ \right) \\
&= u_{+j_0} \left(\prod_{\substack{D \in \text{Max}(S) \\ D \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D) \subsetneq \text{rows}^{j_0}(D_1)}} D^+ \right) \left(\prod_{\substack{E \in \text{Max}(S) \\ N_{j_0} \cap D_1 \cap E = \emptyset}} E^+ \right) \\
&\quad \times \left(\left(\sum_{i=1}^a u_{i+} \times \prod_{\substack{C \in \text{Int}(S) \\ C \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D_1) \subset \text{rows}^{j_0}(C)}} C^+ \right) + \left(\sum_{\alpha=2}^{\ell} \prod_{\substack{C \in \text{Int}(S) \\ \text{rows}^{j_0}(D_\alpha) \subset \text{rows}^{j_0}(C)}} C^+ \right) \right) \\
&= u_{+j_0} \left(\prod_{\substack{D \in \text{Max}(S) \\ D \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D) \subsetneq \text{rows}^{j_0}(D_1)}} D^+ \right) \left(\prod_{\substack{E \in \text{Max}(S) \\ N_{j_0} \cap D_1 \cap E = \emptyset}} E^+ \right) \\
&\quad \times \left(\prod_{\substack{C \in \text{Int}(S) \\ C \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D_1) \subset \text{rows}^{j_0}(C)}} C^+ \right) \left(\sum_{i=1}^a u_{i+} + \sum_{\alpha=2}^{\ell} (D_\alpha \cap D_1)^+ \right) \\
&= u_{+j_0} \left(\prod_{\substack{D \in \text{Max}(S) \\ D \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D) \subsetneq \text{rows}^{j_0}(D_1)}} D^+ \right) \left(\prod_{\substack{E \in \text{Max}(S) \\ N_{j_0} \cap D_1 \cap E = \emptyset}} E^+ \right) \left(\prod_{\substack{C \in \text{Int}(S) \\ C \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D_1) \subset \text{rows}^{j_0}(C)}} C^+ \right) (D_1^+) \\
&= u_{+j_0} \left(\prod_{\substack{D \in \text{Max}(S) \\ D \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D) \subset \text{rows}^{j_0}(D_1)}} D^+ \right) \left(\prod_{\substack{E \in \text{Max}(S) \\ N_{j_0} \cap D_1 \cap E = \emptyset}} E^+ \right) \left(\prod_{\substack{C \in \text{Int}(S) \\ C \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D_1) \subset \text{rows}^{j_0}(C)}} C^+ \right)
\end{aligned}$$

The second equality follows from Proposition 3.4.14. The third equality is an application of the inductive hypothesis. The fourth equality follows from Proposition 3.4.12 along with Remark 3.4.13. The fifth equality follows from Proposition 3.4.11. The sixth equality follows from Proposition

3.4.15. The seventh inequality follows from the fact that D_1 is the only clique whose rows are equal to $\text{rows}^{j_0}(D_1)$. This completes our proof by induction. \square

3.5 Checking the Conditions of Birch's Theorem

In the previous section, we wrote a formula for the sum of $x_{i j_0}$ where i ranges over the rows of some maximal clique D_α . Since the block B_0 induces its own maximal clique, Lemma 3.4.9 allows us to write the sum of the $x_{i j_0}$ s for $1 \leq i \leq r$ in the following concise way. This in turn verifies that the proposed maximum likelihood estimate \hat{p} has the same sufficient statistics as the normalized data u/u_{++} , which is one of the conditions of Birch's theorem.

Corollary 3.5.1. *Let S be DS-free. Then for any column j_0 ,*

$$\sum_{i=1}^r x_{i j_0} = u_{+ j_0} \prod_{D \in \text{Max}(S)} D^+.$$

Proof. The poset $P(j_0)$ has a unique maximal element D_0 with $\text{rows}^{j_0}(D_0) = \text{rows}^{j_0}(j_0)$. Note that D_0 may include more columns than j_0 since it may have columns whose nonzero rows are the same as or contain those of j_0 .

By Proposition 3.4.8, there are no maximal intersections of maximal cliques C with $\text{rows}^{j_0}(D_0) \subset \text{rows}^{j_0}(C)$, since D_0 is maximal in $P(j_0)$. It follows from Proposition 3.4.5 that a maximal clique D intersects N_{j_0} if and only if it has $\text{rows}^{j_0}(D) \subset \text{rows}^{j_0}(D_0)$.

Since $N_{j_0} \subset D_0$, we have that $N_{j_0} \cap D_0 \cap E = N_{j_0} \cap E$ for any clique E . By Lemma 3.4.9, we have

$$\begin{aligned} \sum_{i=1}^r x_{i j_0} &= u_{+ j_0} \left(\prod_{\substack{D \in \text{Max}(S) \\ D \cap N_{j_0} \neq \emptyset \\ \text{rows}^{j_0}(D) \subset \text{rows}^{j_0}(D_0)}} D^+ \right) \left(\prod_{\substack{E \in \text{Max}(S) \\ N_{j_0} \cap D_0 \cap E = \emptyset}} E^+ \right) \\ &= u_{+ j_0} \left(\prod_{\substack{D \in \text{Max}(S) \\ D \cap N_{j_0} \neq \emptyset}} D^+ \right) \left(\prod_{\substack{E \in \text{Max}(S) \\ E \cap N_{j_0} = \emptyset}} E^+ \right) \\ &= u_{+ j_0} \prod_{D \in \text{Max}(S)} D^+, \end{aligned}$$

as needed. \square

Now we will address the condition of Birch's theorem which states that the maximum likelihood estimate must satisfy the equations defining \mathcal{M}_S .

Lemma 3.5.2. *Let S be doubly chordal bipartite. Let $u \in \mathbb{R}^S$ be a generic matrix of counts. Then the point $(\hat{p}_{ij} \mid (i, j) \in S)$ specified in Theorem 3.3.4 is in the Zariski closure of \mathcal{M}_S .*

In order to prove this lemma, we must first describe the vanishing ideal of \mathcal{M}_S . We denote this

ideal $\mathcal{I}(\mathcal{M}_S)$. It is a subset of the polynomial ring in $\#S$ variables,

$$R = \mathbb{C}[p_{ij} \mid ij \in S].$$

Proposition 3.5.3. *Let S be chordal bipartite. Then $\mathcal{I}(\mathcal{M}_S)$ is generated by the 2×2 minors of the matrix form of S that contain no zeros. That is, $\mathcal{I}(\mathcal{M}_S)$ is generated by all binomials of the form*

$$p_{ij}p_{k\ell} - p_{i\ell}p_{kj},$$

such that $(i, j), (k, \ell), (i, \ell), (k, j) \in S$.

Proof. This follows from results in [3, Chapter 10.1]. The *loops* on S correspond to cycles in G_S . The *df 1 loops* as defined in [3, Chapter 10.1] are those whose support does not properly contain the support of any other loop; that is, they correspond to cycles in G_S with no chords. Since G_S is chordal bipartite, each of these cycles contain exactly four edges. Therefore the df 1 loops on S all have degree two, and each corresponds to a 2×2 minor of S by definition. Theorem 10.1 of [3] states that the df 1 loops form a Markov basis for \mathcal{M}_S . Therefore, by the Fundamental Theorem of Markov Bases [15, Theorem 3.1], the 2×2 minors of S form a generating set for $\mathcal{I}(\mathcal{M}_S)$. \square

Example 3.5.4. Consider the matrix S from Example 3.3.3. In Figure 3.2, we see that $G(S)$ has exactly one cycle. This cycle corresponds to the only 2×2 minor in S that contains no zeros, which is the $\{2, 3\} \times \{1, 2\}$ submatrix. Therefore the (complex) Zariski closure of \mathcal{M}_S is the variety of the ideal generated by the polynomial $p_{21}p_{32} - p_{31}p_{22}$.

Proposition 3.5.5. *Let S be set of indices such that G_S is doubly chordal bipartite. Let $\{i_1, i_2\} \times \{j_1, j_2\}$ be a set of indices that corresponds to a 2×2 minor of S that contains no zeros. Let $\hat{p}_{i_1 j_1}, \hat{p}_{i_2 j_2}, \hat{p}_{i_1 j_2}, \hat{p}_{i_2 j_1}$ be as defined in Theorem 3.3.4. Then*

$$\hat{p}_{i_1 j_1} \hat{p}_{i_2 j_2} = \hat{p}_{i_1 j_2} \hat{p}_{i_2 j_1} \tag{3.9}$$

Proof. The terms $u_{i_1+}, u_{i_2+}, u_{+j_1}$ and u_{+j_2} each appear once in the numerator on each side of Equation (3.9), and u_{++}^2 appears in both denominators. Furthermore if (i_1, j_1) and (i_2, j_2) are both contained in any clique in S , then (i_1, j_2) and (i_2, j_1) are also in the clique by definition. So any term that is squared in the numerator or denominator on one side of Equation (3.9) is also squared on the other side. Therefore it suffices to show that $\text{Max}(i_1 j_1) \cup \text{Max}(i_2 j_2) = \text{Max}(i_1 j_2) \cup \text{Max}(i_2 j_1)$ and $\text{Int}(i_1 j_1) \cup \text{Int}(i_2 j_2) = \text{Int}(i_1 j_2) \cup \text{Int}(i_2 j_1)$.

First, we will show that $\text{Max}(i_1 j_1) \cup \text{Max}(i_2 j_2) = \text{Max}(i_1 j_2) \cup \text{Max}(i_2 j_1)$. Let $D \in \text{Max}(i_1 j_1)$. If $(i_2, j_1) \in D$, then we are done.

Now suppose that $(i_2, j_1) \notin D$. Since D intersects column j_1 , by Proposition 3.4.5 we know that D has the form D_α for some block of columns B_α that are identical on rows $^{j_1}(j_1)$. Let $\text{rows}^{j_1}(D_\alpha)$ denote the set of nonzero rows of D_α that are also nonzero rows of j_1 . Since $(i_2, j_1) \notin D$, we have that $i_2 \notin \text{rows}^{j_1}(D_\alpha)$ while $i_1 \in \text{rows}^{j_1}(D_\alpha)$. Since $\text{rows}^{j_1}(j_2) \cap \text{rows}^{j_1}(D_\alpha)$ is nonempty, and since $\text{rows}^{j_1}(j_2) \not\subseteq$

rows ^{j_1} (D_α), we must have that rows ^{j_1} (D_α) \subset rows ^{j_1} (j_2) by the DS-free condition. Therefore $(i_1, j_2) \in D_\alpha$ by definition of D_α . So $D_\alpha = D \in \text{Max}(i_1, j_2)$, as needed.

Switching the roles of i_1 and i_2 or the roles of j_1 and j_2 yields the desired equality.

Now let $C \in \text{Int}(i_1, j_1)$. Then $C = D_\alpha \cap D_\beta$ where $D_\beta \triangleleft D_\alpha$ in the poset $P(j_1)$ by Proposition 3.4.8. If $(i_2, j_1) \in C$, then we are done.

Now suppose that $(i_2, j_1) \notin C$. Then we have that $i_2 \notin \text{rows}^{j_1}(D_\beta)$, whereas $i_1 \in \text{rows}^{j_1}(D_\alpha)$ and $i_1 \in \text{rows}^{j_1}(D_\beta)$. So we must have that rows ^{j_1} (D_β) \subsetneq rows ^{j_1} (j_2) by the DS-free condition. Since rows ^{j_1} (D_α) \cap rows ^{j_1} (j_2) is nonempty, we must have that rows ^{j_1} (D_α) \subset rows ^{j_1} (j_2). This follows from the DS-free condition and the fact that D_α covers D_β in the poset $P(j_1)$. Therefore $(i_1, j_2) \in D_\alpha, D_\beta$ by definition of these cliques. So $C \in \text{Int}(i_1, j_2)$, as needed.

Again, switching the roles of i_1 and i_2 or the roles of j_1 and j_2 in the above proof yields the desired equality. \square

Proof of Lemma 3.5.2. By Proposition 3.5.3, the vanishing ideal of \mathcal{M}_S consists of all fully-observed 2×2 minors of S . By Proposition 3.5.5, each of these 2×2 minors vanishes when evaluated on $\hat{\mathbf{p}}$. \square

We can now prove Theorem 3.3.4.

Proof of Theorem 3.3.4. Let G_S be doubly chordal bipartite. Let $\mathbf{u} \in \mathbb{R}_+^S$ be a matrix of counts. By Corollary 3.5.1, the column marginals of $u_{++}\hat{\mathbf{p}}$ are equal to those of \mathbf{u} . Switching the roles of rows and columns in all of the proofs used to obtain this corollary shows that the row marginals are also equal. Corollary 3.5.1 also implies that $\hat{p}_{++} = 1$ since the vector of all ones is in the rowspan of $A(S)$. So by Lemma 3.5.2 and the fact each $\hat{\mathbf{p}}$ is positive, $\hat{\mathbf{p}} \in \mathcal{M}_S$. Hence by Birch's theorem, $\hat{\mathbf{p}}$ is the maximum likelihood estimate for \mathbf{u} . The other direction is exactly the contrapositive of Theorem 3.2.5. \square

THE CAVENDER-FARRIS-NEYMAN MODEL WITH A MOLECULAR CLOCK

Let T be an n -leaf, rooted, binary phylogenetic tree with labeled edge lengths. Recall from Section 1.5 that the CFN model on T arises as a two-state continuous-time Markov process along T with transition matrices M^e associated to each edge e in T . Recall further that the probability of observing the states (v_1, \dots, v_n) at the leaves of T under the CFN model is

$$p(v_1, \dots, v_n) = \frac{1}{2} \sum_{(v_{n+1}, \dots, v_{2n-1}) \in \mathbb{Z}_2^{n-1}} \prod_{e \in E} M^e(v_{a(e)}, v_{d(e)}).$$

4.1 The Discrete Fourier Transform

In the following discussion, we perform a linear change of coordinates on the probability coordinates and introduce new free parameters in terms of the entries of the transition matrices. This allows us to realize this parametrization as a monomial map. In order to accomplish this, we first provide some background concerning group-based models and the discrete Fourier transform. We always assume that G is a finite abelian group.

Definition 4.1.1. Let $M^e = \exp(Q t_e)$ be a transition matrix arising from a continuous-time Markov process along a tree. Let G be a finite abelian group under addition with order equal to the number of states of the model, and identify the set of states with elements of G . The model is *group-based* with respect to G if for each transition matrix M^e arising from the model, there exists a function $f^e : G \rightarrow \mathbb{R}$ such that $M^e(g, h) = f^e(g - h)$ for all $g, h \in G$.

In particular, note that the CFN-MC model is group-based with respect to \mathbb{Z}_2 with function $f^e : \mathbb{Z}_2 \rightarrow \mathbb{R}$ defined by

$$f^e(0) = (1 + \exp(-2\alpha t_e))/2 \quad \text{and} \quad f^e(1) = (1 - \exp(-2\alpha t_e))/2.$$

Definition 4.1.2. The *dual group* $\hat{G} = \text{Hom}(G, \mathbb{C}^\times)$ of a group G is the group of all homomorphisms $\chi : G \rightarrow \mathbb{C}^\times$, where \mathbb{C}^\times denotes the group of non-zero complex numbers under multiplication. Elements of the dual group are called *characters*. Let $\mathbb{1}$ denote the constant character that maps all elements of G to 1.

Throughout this section, we will make use of the following classical theorems. Proofs of these can be found in [32].

Proposition 4.1.3. *Let G be a finite abelian group. Its dual group \hat{G} is isomorphic to G . Furthermore, for two finite abelian groups G_1 and G_2 , $\widehat{G_1 \times G_2} \cong \hat{G}_1 \times \hat{G}_2$ via $\chi((g_1, g_2)) = \chi_1(g_1)\chi_2(g_2)$ for $g_1 \in G_1$ and $g_2 \in G_2$ and some $\chi_1 \in \hat{G}_1$ and $\chi_2 \in \hat{G}_2$.*

Definition 4.1.4. Let $f : G \rightarrow \mathbb{C}$ be a function. The *discrete Fourier transform* of f is the function

$$\hat{f} : \hat{G} \rightarrow \mathbb{C}, \quad \chi \mapsto \sum_{g \in G} \chi(g) f(g).$$

The discrete Fourier transform is the linear change of coordinates that allows us to view Equation (1.2) as a monomial parametrization. We can write the Fourier transform of p over \mathbb{Z}_2^n in equation (1.2) as

$$\hat{p}(\chi_1, \dots, \chi_n) = \sum_{(g_1, \dots, g_n) \in \mathbb{Z}_2^n} p(g_1, \dots, g_n) \prod_{i=1}^n \chi_i(g_i).$$

Let $\hat{\mathbb{Z}}_2 = \{\mathbb{1}, \phi\}$, where ϕ denotes the only nontrivial homomorphism from \mathbb{Z}_2 to \mathbb{C} . Then \mathbb{Z}_2 and $\hat{\mathbb{Z}}_2$ are isomorphic via the map that identifies 0 to $\mathbb{1}$ and 1 with ϕ . Using this fact, we can write \hat{p} as a function of n elements of \mathbb{Z}_2 as

$$\hat{p}(i_1, \dots, i_n) = \sum_{(j_1, \dots, j_n) \in \mathbb{Z}_2^n} (-1)^{i_1 j_1 + \dots + i_n j_n} p(j_1, \dots, j_n)$$

for all $(i_1, \dots, i_n) \in \mathbb{Z}_2^n$.

The following theorem, independently discovered by Evans and Speed in [17] and Hendy and Penny in [23], describes the monomial parametrization obtained from the discrete Fourier transform. A detailed account can also be found in Chapter 15 of [42].

Theorem 4.1.5. *Let $p(g_1, \dots, g_n)$ be the polynomial describing the probability of observing states $(g_1, \dots, g_n) \in G^n$ at the leaves of phylogenetic tree T under a group-based model. Denote by π the distribution of states at the root of the tree. Let $f^e : G \rightarrow \mathbb{R}$ denote the function associated to edge e by*

the definition of a group based model. Then the Fourier transform of p is

$$\hat{p}(\chi_1, \dots, \chi_n) = \hat{\pi}\left(\prod_{i=1}^n \chi_i\right) \prod_{e \in E(T)} \hat{f}^e\left(\prod_{l \in \lambda(e)} \chi_l\right),$$

where $\lambda(e)$ is the set of all leaves that are descended from edge e .

Note that since π is the uniform distribution,

$$\hat{\pi}(\chi) = \frac{1}{2}(\chi(0) + \chi(1)) = \begin{cases} 1, & \text{if } \chi = \mathbb{1}, \\ 0, & \text{if } \chi = \phi. \end{cases}$$

Interpreting this in the context of the Fourier transform of p and using the isomorphism of \mathbb{Z}_2 and $\hat{\mathbb{Z}}_2$ gives that

$$\hat{p}(\mathbf{g}_1, \dots, \mathbf{g}_n) = \begin{cases} \prod_{e \in E(T)} \hat{f}^e(\sum_{l \in \lambda(e)} \mathbf{g}_l), & \text{if } \sum_{i=1}^n \mathbf{g}_i = 0 \\ 0, & \text{if } \sum_{i=1}^n \mathbf{g}_i = 1. \end{cases} \quad (4.1)$$

Consider the Fourier transform of each $f^e(g)$. In the case of the CFN-MC model, we can think of the discrete Fourier transform as a simultaneous diagonalization of the transition matrices via a 2×2 Hadamard matrix. Indeed, letting $H = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ gives that

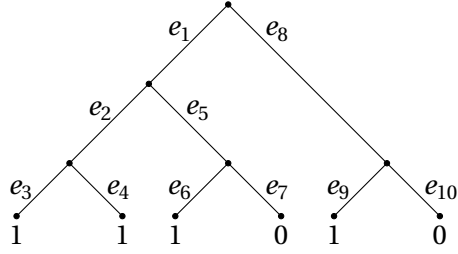
$$H^{-1} M^e H = \begin{bmatrix} a_0^e & 0 \\ 0 & a_1^e \end{bmatrix},$$

where $a_1^e = \exp(-2\alpha t_e)$ and $a_0^e = 1$. (Although $a_0^e = 1$ for all e , it is useful to think of a_0^e as a free parameter for the following discussion.) Note that these values are exactly those obtained by performing the discrete Fourier transform on f^e :

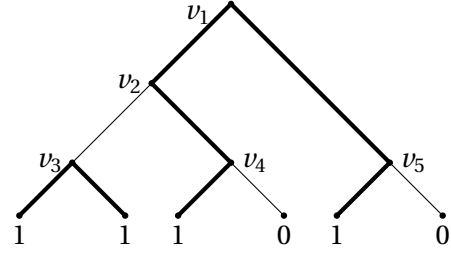
$$\begin{aligned} \hat{f}^e(\mathbb{1}) &= f^e(0)\mathbb{1}(0) + f^e(1)\mathbb{1}(1) \\ &= \frac{1 + \exp(-2\alpha t_e)}{2} + \frac{1 - \exp(-2\alpha t_e)}{2} \\ &= 1 \\ \hat{f}^e(\phi) &= f^e(0)\phi(0) + f^e(1)\phi(1) \\ &= \frac{1 + \exp(-2\alpha t_e)}{2} - \frac{1 - \exp(-2\alpha t_e)}{2} \\ &= \exp(-2\alpha t_e). \end{aligned}$$

Using these new parameters, the isomorphism of \mathbb{Z}_2 and $\hat{\mathbb{Z}}_2$ and equation (4.1), we can see that

$$\hat{p}(\mathbf{g}_1, \dots, \mathbf{g}_n) = \begin{cases} \prod_{e \in E(T)} a_{i(e)}^e, & \text{if } \sum_{i=1}^n \mathbf{g}_i = 0 \\ 0, & \text{if } \sum_{i=1}^n \mathbf{g}_i = 1, \end{cases} \quad (4.2)$$



(a) A leaf labeling of T with elements of \mathbb{Z}_2 that sum to 0



(b) The path system associated to this labeling

Figure 4.1 The tree T referenced in Example 4.1.6

where $i(e)$ denotes the sum in \mathbb{Z}_2 of the group elements at all leaves descended from e . Note that this is, in fact, a monomial parametrization, as desired.

Example 4.1.6. Consider the tree T in Figure 4.1a. We compute $\hat{p}(1, 1, 1, 0, 1, 0)$. Using Equation (4.2), we see that

$$\hat{p}(g_1, \dots, g_n) = a_1^{e_1} a_0^{e_2} a_1^{e_3} a_1^{e_4} a_1^{e_5} a_1^{e_6} a_0^{e_7} a_1^{e_8} a_1^{e_9} a_0^{e_{10}}.$$

Let e_1, e_2, e_3 be edges of rooted binary tree T that are adjacent to a single node v , where $v = d(e_1)$ and $v = a(e_2) = a(e_3)$. Then $i(e_1) = i(e_2) + i(e_3)$, so $i(e_1) + i(e_2) + i(e_3) = 0$. In particular, this means that at any internal node, the edges adjacent to that node have an even number of 1's. Since the labels at the leaves of the tree sum to 0, if e_1, e_2 are the edges adjacent to the root, then $i(e_1) + i(e_2) = 0$. Therefore, to each labeling of the leaves of T with elements of \mathbb{Z}_2 that sum to 0 we may associate a set of disjoint paths, or path systems, between leaves of T . Furthermore, given a set of disjoint paths between leaves of T , we obtain a labeling of the leaves that sums to 0 by assigning a 1 to each leaf included in one of the paths and a 0 elsewhere. So labelings of the leaves of T that sum to 0 and sets of disjoint paths between leaves of T are in bijection with one another.

Definition 4.1.7. Let $\mathbb{Z}_2^{n, \text{even}}$ denote the set of all labelings of the leaves of T with elements of \mathbb{Z}_2 that sum to 0. The *path system* associated to a labeling $(i_1, \dots, i_n) \in \mathbb{Z}_2^{n, \text{even}}$ is the unique set of paths in T that connect the leaves of T that are labeled with 1 and do not use any of the same edges. We often denote a path system by \mathfrak{P} .

In this context, the edges for which a_1^e appears in the parametrization (4.2) of $\hat{p}(i_1, \dots, i_n)$ for $(i_1, \dots, i_n) \in \mathbb{Z}_2^{n, \text{even}}$ are exactly those that appear in the path system associated to (i_1, \dots, i_n) .

Example 4.1.8. The path system associated to the labeling $(1, 1, 1, 0, 1, 0)$ from Example 4.1.6 is pictured in Figure 4.1b. Notice that the bold edges e in T are exactly those for which a_1^e appears in the parametrization of $\hat{p}(1, 1, 1, 0, 1, 0)$.

We now restrict the parametrization of the CFN model to trees that satisfy the molecular clock condition, which restricts us to a lower-dimensional subspace of the parameter space and provides

a new combinatorial way of interpreting the Fourier coordinates. The molecular clock condition imposes that if e_1, \dots, e_s and f_1, \dots, f_r are two paths from an internal node v to leaves descended from v , then $t_{e_1} + \dots + t_{e_s} = t_{f_1} + \dots + t_{f_r}$. On the level of transition matrices, this means that

$$\begin{aligned} M^{e_1} \dots M^{e_s} &= \exp(Q t_{e_1}) \dots \exp(Q t_{e_s}) \\ &= \exp(Q(t_{e_1} + \dots + t_{e_s})) \\ &= \exp(Q(t_{f_1} + \dots + t_{f_r})) \\ &= M^{f_1} \dots M^{f_r}. \end{aligned}$$

Since we can apply the Fourier transform to diagonalize the resulting matrices on both sides of this equation, we see that the products of the new parameters satisfy the identities:

$$a_0^{e_1} \dots a_0^{e_s} = a_0^{f_1} \dots a_0^{f_r} \quad \text{and} \quad a_1^{e_1} \dots a_1^{e_s} = a_1^{f_1} \dots a_1^{f_r}.$$

In particular, this means that we may define new parameters a_0^v and a_1^v for each internal node v by

$$a_i^v = a_i^{e_1} \dots a_i^{e_s}$$

for $i = 0, 1$ where e_1, \dots, e_s is a path from v to any leaf descended from v . Note that if v is a leaf, then $a_i^v = 1$. So we may exclude it from the parametrization, and restrict to a parametrization by a_0^v and a_1^v where v is an internal node. Furthermore, note that for any edge e_1 and $i = 0, 1$, we have the relations

$$\begin{aligned} a_i^{a(e_1)} (a_i^{d(e_1)})^{-1} &= a_i^{e_1} \dots a_i^{e_s} (a_i^{e_s})^{-1} \dots (a_i^{e_2})^{-1} \\ &= a_i^{e_1}, \end{aligned} \tag{4.3}$$

where e_1, \dots, e_s is a path from the ancestral node $a(e_1)$ to a leaf descended from $d(e_1)$.

Let $(i_1, \dots, i_n) \in \mathbb{Z}_2^{n, \text{even}}$. Let \mathfrak{P} be the path system associated to i_1, \dots, i_n . Denote by $\text{Int}(T)$ the set of all internal nodes of tree T .

Definition 4.1.9. We say that v is the *top-most node* of a path in \mathfrak{P} if both of the edges descended from v are in a path in \mathfrak{P} . In other words, v is the node of the path that includes it that is closest to the root. The *top-set* $\text{Top}(i_1, \dots, i_n)$ is the set of all top-most nodes of paths in \mathfrak{P} . The *top-vector* is the vector in $\mathbb{R}^{\text{Int}(T)}$ with v component equal to 1 if $v \in \text{Top}(i_1, \dots, i_n)$ and 0 otherwise. We denote the top-vector of a particular path system \mathfrak{P} by $[\mathfrak{P}]$ or $\mathbf{x}^{\mathfrak{P}}$, depending upon the context.

Example 4.1.10. Consider the path system associated to labeling $(1, 1, 1, 0, 1, 0) \in \mathbb{Z}_2^{n, \text{even}}$ pictured in Figure 4.1b. The top-set of this path system is $\text{Top}(1, 1, 1, 0, 1, 0) = \{v_1, v_3\}$. The top-vector is $(1, 0, 1, 0, 0) \in \mathbb{R}^{\text{Int}(T)}$.

For each internal node ν , we define two new parameters b_0^ν and b_1^ν by

$$\begin{aligned} b_0^\nu &= \begin{cases} (a_0^\nu)^2 & \text{if } \nu \text{ is the root, and} \\ a_0^\nu & \text{otherwise,} \end{cases} \\ b_1^\nu &= \begin{cases} (a_1^\nu)^2 & \text{if } \nu \text{ is the root, and} \\ (a_1^\nu)^2 (a_0^\nu)^{-1} & \text{otherwise.} \end{cases} \end{aligned} \quad (4.4)$$

We now rewrite the parametrization of $\hat{p}(i_1, \dots, i_n)$ from Equation (4.2) in terms of these new parameters using Equation (4.3). Let \mathfrak{P} be the path system associated to (i_1, \dots, i_n) . The choice of b_0^ν or b_1^ν in the new parametrization will depend upon the position of ν in \mathfrak{P} .

Let ν be an internal node of T that is not the root. Let e_1, e_2, e_3 be the edges adjacent to ν so that $\nu = d(e_1)$ and $\nu = a(e_2) = a(e_3)$. By Equation (4.3), $a_i^{e_1}$, $a_i^{e_2}$ and $a_i^{e_3}$ are the only parameters in which some a_0^ν or a_1^ν appear.

If ν is not in any path in \mathfrak{P} , then \mathfrak{P} does not use edges e_1, e_2 or e_3 . So the factors of $\hat{p}(i_1, \dots, i_n)$ in Equation (4.2) associated to these edges are

$$\begin{aligned} a_0^{e_1} a_0^{e_2} a_0^{e_3} &= a_0^{a(e_1)} (a_0^\nu)^{-1} \cdot a_0^\nu (a_0^{d(e_2)})^{-1} \cdot a_0^\nu (a_0^{d(e_3)})^{-1} \\ &= a_0^{a(e_1)} (a_0^{d(e_2)})^{-1} (a_0^{d(e_3)})^{-1} \cdot (a_0^\nu)^2 (a_0^\nu)^{-1} \\ &= a_0^{a(e_1)} (a_0^{d(e_2)})^{-1} (a_0^{d(e_3)})^{-1} \cdot a_0^\nu, \end{aligned}$$

by Equation (4.3). So if ν is not in any path in \mathfrak{P} , a_0^ν is the only factor of $\hat{p}(i_1, \dots, i_n)$ involving ν .

Similarly, consider the case where ν is in a path in \mathfrak{P} , but it is not the top-most node of a path in \mathfrak{P} . Then this path includes e_1 , and without loss of generality, we may assume it includes e_2 and not e_3 . Then the factors of $\hat{p}(i_1, \dots, i_n)$ involving e_1, e_2 and e_3 are

$$\begin{aligned} a_1^{e_1} a_1^{e_2} a_0^{e_3} &= a_1^{a(e_1)} (a_1^\nu)^{-1} \cdot a_1^\nu (a_1^{d(e_2)})^{-1} \cdot a_0^\nu (a_0^{d(e_3)})^{-1} \\ &= a_1^{a(e_1)} (a_1^{d(e_2)})^{-1} (a_0^{d(e_3)})^{-1} \cdot (a_0^\nu) (a_1^\nu) (a_1^\nu)^{-1} \\ &= a_0^{a(e_1)} (a_0^{d(e_2)})^{-1} (a_0^{d(e_3)})^{-1} \cdot a_0^\nu, \end{aligned}$$

by Equation (4.3). So in this case, we have again that a_0^ν is the only factor of $\hat{p}(i_1, \dots, i_n)$ involving ν .

Finally, consider the case where ν is the top-most node of a path in \mathfrak{P} . Then this path includes e_2 and e_3 but not e_1 . Then the factors of $\hat{p}(i_1, \dots, i_n)$ involving e_1, e_2 and e_3 are

$$\begin{aligned} a_0^{e_1} a_1^{e_2} a_1^{e_3} &= a_0^{a(e_1)} (a_0^\nu)^{-1} \cdot a_1^\nu (a_1^{d(e_2)})^{-1} \cdot a_1^\nu (a_1^{d(e_3)})^{-1} \\ &= a_0^{a(e_1)} (a_1^{d(e_2)})^{-1} (a_1^{d(e_3)})^{-1} \cdot (a_0^\nu)^{-1} (a_1^\nu)^2 \end{aligned}$$

by Equation (4.3). So in this case, $(a_0^v)^{-1}(a_1^v)^2$ is exactly the factor of $\hat{p}(i_1, \dots, i_n)$ involving v .

An analogous argument shows that when v is the root, the factor of $\hat{p}(i_1, \dots, i_n)$ involving v is $(a_0^v)^2$ when v is not the top-most node of a path in \mathfrak{P} , and $(a_1^v)^2$ when v is the top-most node of a path in \mathfrak{P} . Note that it can never be the case that the root is in a path and is not the top-most node.

So the new parameters b_0^v and b_i^v for each internal node v of T allow us to rewrite the parametrization in Equation (4.2) as

$$\hat{p}(i_1, \dots, i_n) = \prod_{v \in \text{Top}(i_1, \dots, i_n)} b_1^v \times \prod_{v \in \overline{\text{Top}(i_1, \dots, i_n)}} b_0^v, \quad (4.5)$$

where $\overline{\text{Top}(i_1, \dots, i_n)}$ denotes the compliment of $\text{Top}(i_1, \dots, i_n)$ in the set of all internal nodes of T .

Example 4.1.11. Consider the parametrization of $\hat{p}(1, 1, 1, 0, 1, 0)$ given in Example 4.1.6 for the tree T pictured in 4.1a. First, we verify the identity in Equation (4.3) for $a_1^{e_1}$. We have that $a(e_1) = v_1$ and $d(e_1) = v_2$. We have defined $a_1^{v_1} = a_1^{e_1} a_1^{e_2} a_1^{e_3}$ and $a_1^{v_2} = a_1^{e_2} a_1^{e_3}$. Therefore

$$\begin{aligned} a_1^{v_1} (a_1^{v_2})^{-1} &= a_1^{e_1} a_1^{e_2} a_1^{e_3} (a_1^{e_3})^{-1} (a_1^{e_2})^{-1} \\ &= a_1^{e_1}. \end{aligned}$$

Note that while the choices of paths from v_1 and v_2 to leaves descended from them was not unique, the molecular clock condition implies that the above holds for any such choice of paths.

Substituting the identities in Equation (4.3) into $\hat{p}(1, 1, 1, 0, 1, 0)$, and applying the fact that if l is a leaf of T then $a_i^l = 1$ for $i = 0, 1$ yields

$$\begin{aligned} \hat{p}(1, 1, 1, 0, 1, 0) &= a_1^{v_1} (a_1^{v_2})^{-1} a_0^{v_2} (a_0^{v_3})^{-1} a_1^{v_3} a_1^{v_3} a_1^{v_2} (a_1^{v_4})^{-1} a_1^{v_4} a_0^{v_4} a_1^{v_1} (a_1^{v_5})^{-1} a_1^{v_5} a_0^{v_5} \\ &= (a_1^{v_1})^2 a_0^{v_2} (a_1^{v_3})^2 (a_0^{v_3})^{-1} a_0^{v_4} a_0^{v_5}. \end{aligned}$$

Substituting the new parameters, b_0^v and b_1^v as defined in Equation (4.4) yields

$$\hat{p}(1, 1, 1, 0, 1, 0) = b_1^{v_1} b_0^{v_2} b_1^{v_3} b_0^{v_4} b_0^{v_5},$$

as needed.

Note that two labelings of the leaves with group elements (i_1, \dots, i_n) and (j_1, \dots, j_n) have the same top-sets if and only if $\hat{p}(i_1, \dots, i_n) = \hat{p}(j_1, \dots, j_n)$. Therefore, Equation (4.5) allows us to define new coordinates that are indexed by valid top-sets of path systems in T . These coordinates are in the polynomial ring

$$\mathbb{C}[\mathbf{r}] := \mathbb{C}[r_{k_1, \dots, k_{n-1}} : (k_1, \dots, k_{n-1}) = [\mathfrak{P}]] \text{ for some path system } \mathfrak{P}$$

where (i_1, \dots, i_n) ranges over all elements of $\mathbb{Z}_2^{n, \text{even}}$. By applying this change of coordinates, we effectively quotient by the linear relations among the \hat{p} coordinates that arise from the fact that

their parametrizations in terms of the b_i^ν parameters are equal. This restricts our attention to equivalence classes of labelings in $\mathbb{Z}_2^{n,\text{even}}$ with the same top-sets.

Definition 4.1.12. Label the internal nodes of T with v_1, \dots, v_{n-1} . The *CFN-MC ideal* I_T is the kernel of the map

$$\begin{aligned} \mathbb{C}[\mathbf{r}] &\longrightarrow \mathbb{C}[b_i^\nu \mid i = 0, 1, \nu \in \text{Int}(T)] \\ r_{k_1, \dots, k_{n-1}} &\longmapsto \prod_{i=1}^{n-1} b_{k_i}^{v_i}, \end{aligned}$$

where (k_1, \dots, k_{n-1}) ranges over all indicator vectors corresponding to top-sets of path systems in T .

Note that the polynomials in the ideal I_T evaluate to zero for every choice of parameters in the CFN-MC model for the tree T . In particular, these polynomials are *phylogenetic invariants* of the CFN-MC model. Another important observation is that I_T is the kernel of a monomial map. This implies that I_T is a toric ideal and can be analyzed from a combinatorial perspective.

An equivalent way to define the CFN-MC ideal is as the kernel of the map

$$\begin{aligned} \mathbb{C}[\mathbf{r}] &\longrightarrow \mathbb{C}[t_0, \dots, t_{n-1}] \\ r_{k_1, \dots, k_{n-1}} &\longmapsto t_0 \prod_{k_i=1} t_i, \end{aligned} \tag{4.6}$$

where t_0 is a homogenizing indeterminate. Note that these indeterminates t_i are not related to the branch lengths in T . From this perspective, we define the matrix A_T associated to this monomial map to be the matrix whose columns are the indicator vectors of top-sets of path systems in T with an added homogenizing row of ones. Our goal in the next two chapters is to study the ideals I_T for binary trees and the corresponding polytopes R_T (to be defined in detail in Section 4.2).

Example 4.1.13. Let T be the tree pictured in Figure 1.2a. The CFN-MC ideal I_T is in the polynomial ring $\mathbb{C}[\mathbf{r}] = \mathbb{C}[r_{0000}, r_{1000}, r_{0100}, r_{0010}, r_{0001}, r_{1010}, r_{1001}, r_{0011}]$ where each subscript is the indicator vector of a top-set of a path system in T indexed alphabetically by the internal nodes of T . Therefore, the parametrization in Equation (4.6) is given by

$$\begin{array}{llll} r_{0000} \mapsto t_0 & r_{0100} \mapsto t_0 t_b & r_{0001} \mapsto t_0 t_d & r_{1001} \mapsto t_0 t_a t_d \\ r_{1000} \mapsto t_0 t_a & r_{0010} \mapsto t_0 t_c & r_{1010} \mapsto t_0 t_a t_c & r_{0011} \mapsto t_0 t_c t_d. \end{array}$$

The matrix A_T associated to this monomial map is obtained by taking its columns to be all of the subscripts of an indeterminate in $\mathbb{C}[\mathbf{r}]$ and adding a homogenizing row of ones. In this case, this

matrix is

$$A_T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}.$$

We can write I_T implicitly from its parametrization using standard elimination techniques [38, Algorithm 4.5]. This ideal is generated by the binomials

$$\begin{array}{ll} r_{0000}r_{0011} - r_{0010}r_{0001} & r_{0000}r_{1010} - r_{1000}r_{0010} \\ r_{1000}r_{0011} - r_{1010}r_{0001} & r_{0000}r_{1001} - r_{1000}r_{1010} \\ r_{1000}r_{0011} - r_{0010}r_{1001} & r_{0010}r_{1001} - r_{1010}r_{0001}. \end{array}$$

In fact, these are exactly the binomials described in the proof of Proposition 4.2.29; in this setting, the first column are the elements of the ‘‘Lift’’ set and the second column are the elements of the ‘‘Swap’’ set.

We conclude this section by remarking that the combinatorial interpretation for the parametrization of the Fourier coordinates described in this section relies upon having a model with only two states. One starting point for future work towards applying the molecular clock condition to models with three or more states may be to give a comparable combinatorial description of the non-zero Fourier coordinates in these models.

4.2 The CFN-MC Polytope

In this section we give a description of the combinatorial structure of the polytope associated to the CFN-MC model. In particular, we show that the number of vertices of the CFN-MC polytope is a Fibonacci number and we give a complete facet description of the polytope. One interesting feature of these polytopes is that while the facet structure varies widely depending on the structure of the tree (e.g. some trees with n leaves have exponentially many facets, while others only have linearly many facets), the number of vertices is fixed. Similarly, we will see in Chapter 5 that the volume also does not depend on the number of leaves.

Let T be a rooted binary tree on n leaves. For any path system \mathfrak{P} in T , let $\mathbf{x}^{\mathfrak{P}} \in \mathbb{R}^{n-1}$ have i th component $x_i^{\mathfrak{P}} = 1$ if i is the highest internal node in some path in \mathfrak{P} and $x_i^{\mathfrak{P}} = 0$ otherwise. Hence $\mathbf{x}^{\mathfrak{P}}$ is the top-vector of \mathfrak{P} as discussed in the previous section.

Definition 4.2.1. Let T be a rooted binary tree on n leaves. The *CFN-MC polytope* R_T is the convex hull of all $\mathbf{x}^{\mathfrak{P}}$ for \mathfrak{P} a path system in T .

Example 4.2.2. For the tree in Figure 1.2a, the polytope R_T is the convex hull of the column vectors of the matrix A_T in Example 4.1.13.

We note that the convex hull of the column vectors of A_T is actually a subset of the hyperplane $\{\mathbf{x} \in \mathbb{R}^n \mid x_0 = 1\}$. To obtain R_T , we identify this hyperplane with \mathbb{R}^{n-1} by deleting the first coordinate. We write $\text{conv}(A_T)$ to mean the convex hull of the column vectors of A_T after we have deleted the first coordinate.

Recall that the *Fibonacci numbers* are defined by the recurrence $F_n = F_{n-1} + F_{n-2}$ subject to initial conditions $F_0 = F_1 = 1$.

Proposition 4.2.3. *Let T be a rooted binary tree on $n \geq 2$ leaves. The number of vertices of R_T is F_n , the n -th Fibonacci number.*

Proof. We proceed by induction on n . For the base cases, we note that if T is the 2-leaf tree, then $R_T = \text{conv}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, and if T is the 3-leaf tree, then $R_T = \text{conv}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Let T be an n -leaf tree with $n \geq 4$. Let l_1 and l_2 be leaves of T that are adjacent to the same internal node a , so that a is a cherry node. Leaves l_1 and l_2 exist because every rooted binary tree with $n \geq 2$ leaves has a cherry.

Let T' be the tree obtained from T by deleting leaves l_1 and l_2 and their adjacent edges so that a becomes a leaf. If \mathfrak{P} is a path system in T with $a \notin \text{Top}(\mathfrak{P})$, then we can realize the top-vector of \mathfrak{P} without the a -coordinate as the top-vector of a path system in T' . Furthermore, any path system in T' can be extended to a path system in T without a in its top-set. So the number of vertices of R_T with a -coordinate equal to 0 is the number of vertices of $R_{T'}$, which is F_{n-1} by induction.

Let a' be the direct ancestor of a in T . Let T'' be the tree obtained from T by deleting l_1, l_2 and a , and all edges adjacent to a , and merging the two remaining edges incident to a' so that a' is no longer a node. In the case where a' is the root of T , we simply delete it and its other incident edge to form T'' . If \mathfrak{P} is a path system in T with $a \in \text{Top}(\mathfrak{P})$, then note that $a' \notin \text{Top}(\mathfrak{P})$. Furthermore, edge aa' is not an edge in any path in \mathfrak{P} . Therefore, we can realize the top-vector of \mathfrak{P} without the a - and a' -coordinates as the top-vector of a path system in T'' . Furthermore, any path system in T'' can be extended to a path system in T with a in its top-set. So the number of vertices of R_T with a -coordinate equal to 1 is the number of vertices of $R_{T''}$, which is F_{n-2} by induction.

Therefore, the total number of vertices of R_T is $F_{n-2} + F_{n-1} = F_n$, as needed. \square

In order to give a facet description of the CFN-MC polytope of a tree, we define several intermediary polytopes between the CFN polytope and the CFN-MC polytope, along with linear maps between them. The CFN polytope is the analogue of the CFN-MC polytope for the CFN model; it is obtained by taking the convex hull of the indicator vectors of path systems \mathfrak{P} in T indexed by the *edges* in the path system. We will trace the known description of the facets of the CFN polytope through these linear maps via Fourier-Motzkin elimination to arrive at the facet description of the CFN-MC polytope. See [45, Chapter 1] for background on Fourier-Motzkin elimination.

Let T be a rooted binary tree with n leaves, oriented with the root as the highest node and the leaves as the lowest nodes. Then T has $n - 1$ internal nodes. The internal nodes will now be labeled by $1, \dots, n - 1$.

Remark 4.2.4. For the remainder of the thesis, we have changed the convention of labeling the nodes so that $1, \dots, n-1$ label the internal nodes of the trees, whereas in Section 4.1 we used $1, \dots, n$ to denote the leaves. This is because of the importance that the internal nodes now play in the combinatorics of the CFN-MC model, whereas in Section 4.1 the leaves were the main objects of interest in analyzing and simplifying the parametrization.

Let v be a non-root node in T . Denote by $e(v)$ the unique edge that has $v = d(e(v))$. Introduce a poset $\text{Int}(T)$ whose elements are the internal nodes of T and with relations $v \leq w$ if v is a descendant of w . The Haase diagram of $\text{Int}(T)$ is the tree T with leaf and edge incident to a leaf removed. Recall that an order ideal of $\text{Int}(T)$ is a subset of $\text{Int}(T)$ that is downwards closed. Let I be an order ideal of $\text{Int}(T)$ with s elements. Then the number of edges not below an element of I is $2(n-s-1)$, since T has $2n-2$ edges and each node in I has exactly two edges directly beneath it.

Definition 4.2.5. Let T be a tree, $\text{Int}(T)$ the associated poset, and I an order ideal of $\text{Int}(T)$. Denote by $T-I$ the tree obtained by removing all nodes and edges descended from any node in I . The *edge set* of $T-I$, denoted $\mathcal{E}(T-I)$, is the set of all edges in T that are not descended from an element of I . Notice $T-I$ includes all maximal nodes of I , and all edges that join a node in I with an internal node outside of I .

Definition 4.2.6. Let \mathfrak{P} be a path system between leaves of T . Let $[\mathbf{x}, \mathbf{y}]_I^{\mathfrak{P}}$ be the point in $\mathbb{R}^I \oplus \mathbb{R}^{\mathcal{E}(T-I)}$ defined by

$$x_i = \begin{cases} 1 & \text{if } i \text{ is the top-most node in some path in } \mathfrak{P} \\ 0 & \text{otherwise,} \end{cases}$$

for all $i \in I$ and

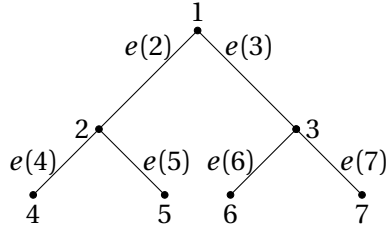
$$y_j = \begin{cases} 1 & \text{if } e(j) \text{ is an edge in some path in } \mathfrak{P} \\ 0 & \text{otherwise,} \end{cases}$$

for all $e(j) \in \mathcal{E}(T-I)$. The polytope $R_T(I)$ is the convex hull of all $[\mathbf{x}, \mathbf{y}]_I^{\mathfrak{P}}$ for all path systems \mathfrak{P} in T . If I is the set of all internal nodes of T , then this polytope is exactly R_T , and $[\mathbf{x}, \mathbf{y}]_I^{\mathfrak{P}} = \mathbf{x}^{\mathfrak{P}} = [\mathfrak{P}]$.

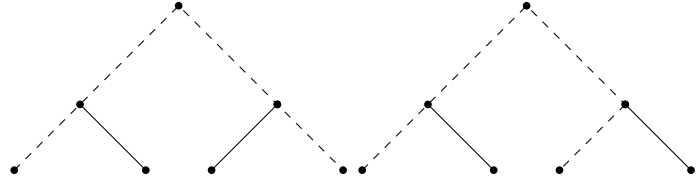
Example 4.2.7. Let T be the 4-leaf tree pictured in Figure 4.2. Let the distinguished order ideal in the set of internal nodes of T be $I = \{3\}$. Then $R_T(I)$ is the convex hull of the following 6 vertices with coordinates corresponding to the labeled edge or node.

$$\begin{array}{l} e(2) \\ e(3) \\ e(4) \\ e(5) \\ 3 \end{array} \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

The dotted lines in Figure 4.2b shows the paths through T that realize the vertex $[1 \ 1 \ 1 \ 0 \ 0]^T$ in $R_T(I)$.



(a) Leaf and edge labels in the tree T



(b) Two paths in T that correspond to the same vertex of $R_T(I)$.

Figure 4.2 The four-leaf tree in Example 4.2.7

For any order ideal I with maximal node r , let a and b be the direct descendants of r . This scenario is pictured in Figure 4.3. Then we can define a linear map

$$\phi_{I,r} : \mathbb{R}^{(I-\{r\})} \oplus \mathbb{R}^{\mathcal{E}(T-(I-\{r\}))} \rightarrow \mathbb{R}^I \oplus \mathbb{R}^{\mathcal{E}(T-I)},$$

sending $\phi_{I,r}(\mathbf{x}', \mathbf{y}') = (\mathbf{x}, \mathbf{y})$ where

$$\begin{cases} x_i = x'_i & \text{if } i \in I - \{r\} \\ y_j = y'_j & \text{if } e(j) \in \mathcal{E}(T-I) \\ x_r = \frac{-y'_r + y'_a + y'_b}{2}. \end{cases}$$

Note that if r is the root, then y'_r is undefined. So we interpret the formula for x_r as if $y'_r = 0$, and we have $x_r = \frac{y'_a + y'_b}{2}$. But in this case, $R_{I-\{r\}}$ lies in the hyperplane defined by $y'_a = y'_b$. So $x_r = y'_a = y'_b$. Here, \mathbf{x}' has elements indexed by nodes in $I - \{r\}$ and \mathbf{y}' has elements indexed by nodes in $\mathcal{E}(T - (I - \{r\}))$. Then \mathbf{x} has elements indexed by nodes in I and \mathbf{y} has elements indexed by nodes in $\mathcal{E}(T - I)$.

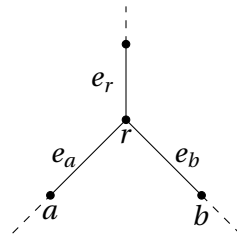


Figure 4.3 The edges and nodes surrounding node r .

Proposition 4.2.8. *The function $\phi_{I,r}$ maps $R_T(I - \{r\})$ onto $R_T(I)$.*

Proof. We show that for all path systems \mathfrak{P} in T , the image of $[\mathbf{x}', \mathbf{y}']_{I-\{r\}}^{\mathfrak{P}}$ under $\phi_{I,r}$ is $[\mathbf{x}, \mathbf{y}]_I^{\mathfrak{P}}$. If r is a node in a path in \mathfrak{P} , then the path includes exactly two edges about r . So we have the following

cases.

Case 1: Suppose that $y'_a = y'_b = y'_r = 0$. Then $e(a)$, $e(b)$ and $e(r)$ are not edges in any path in \mathfrak{P} , so r cannot be the highest node in any path in \mathfrak{P} . After applying $\phi_{I,r}$, we have $x_r = 0$, $x_i = x'_i$ for all $i \in I - \{r\}$, and $y_j = y'_j$ for all $e(j) \in \mathcal{E}(T - I)$. So, the image of $[\mathbf{x}', \mathbf{y}']_{I - \{r\}}^{\mathfrak{P}}$ under $\phi_{I,r}$ is $[\mathbf{x}, \mathbf{y}]_I^{\mathfrak{P}}$ in this case.

Case 2: Suppose that $y'_a = y'_r = 1$ and $y'_b = 0$. In this case, the path in \mathfrak{P} containing r passes through r along $e(a)$ and then upwards out of r along $e(r)$. So, r is not the highest node in this path. Since all paths in \mathfrak{P} are disjoint, r is not the highest node in *any* path in \mathfrak{P} . Applying $\phi_{I,r}$ gives

$$x_r = \frac{y'_a + y'_b - y'_r}{2} = 0,$$

as needed. So, the image of $[\mathbf{x}', \mathbf{y}']_{I - \{r\}}^{\mathfrak{P}}$ under $\phi_{I,r}$ is $[\mathbf{x}, \mathbf{y}]_I^{\mathfrak{P}}$ in this case. The case where $y'_b = y'_r = 1$ and $y'_a = 0$ is analogous.

Case 3: Suppose that $y'_a = y'_b = 1$ and $y'_r = 0$. In this case, the path in \mathfrak{P} containing r comes up to r along $e(a)$ and then back downwards along $e(b)$. So, r is the highest node in this path. Applying $\phi_{I,r}$ gives

$$x_r = \frac{y'_a + y'_b - y'_r}{2} = 1,$$

as needed. So, the image of $[\mathbf{x}', \mathbf{y}']_{I - \{r\}}^{\mathfrak{P}}$ under $\phi_{I,r}$ is $[\mathbf{x}, \mathbf{y}]_I^{\mathfrak{P}}$ in this case.

So, every vertex of $R_T(I - \{r\})$ maps to a vertex of $R_T(I)$ under $\phi_{I,r}$. Furthermore, every vertex $[\mathbf{x}, \mathbf{y}]_I^{\mathfrak{P}}$ of $R_T(I)$ is the image of $[\mathbf{x}', \mathbf{y}']_{I - \{r\}}^{\mathfrak{P}}$. The result holds by linearity of the map $\phi_{I,r}$. \square

Definition 4.2.9. Let I be an order ideal in the poset consisting of all internal nodes of T . A node $v \in I$ is called a *cluster node* if v is connected by edges to three other internal nodes. A connected set of cluster nodes of T is called a *cluster*. Given a cluster $C \subseteq I$, $N_I(C)$ denotes the *neighbor set* of C , which is the set of all internal nodes of T that lie in $I - C$ and are adjacent to some node in C . When I is the set of all internal nodes of T , we denote the neighbor set by $N(C)$. Denote by $m(C)$ the maximal node of C .

Note that the maximal node of a cluster always exists since the cluster is a connected subset of the rooted tree T .

Example 4.2.10. Consider the tree T in Figure 4.4. Then the set of nodes marked with triangles, $\{b, c\}$ forms a cluster since b and c are both cluster nodes and are adjacent. The neighbor set of this cluster, $N(\{b, c\})$, is the set of nodes marked with squares, $\{a, d, e, f\}$. The maximal element is $m(\{b, c\}) = b$.

The main result of this section is Corollary 4.2.13, which gives a list of the facet defining inequalities of the polytopes R_T . This result is obtained by proving the following more general results for the polytopes $R_T(I)$. This facet description depends on the underlying structure of the clusters in T .

Theorem 4.2.11. *The polytope $R_T(I)$ is the solution to the following set of constraints:*

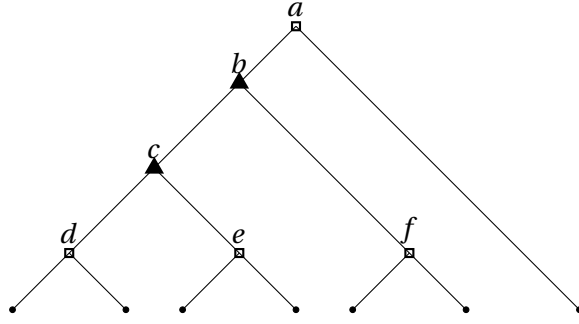


Figure 4.4 An example of a cluster b, c , which are marked with triangles, and the elements of their neighbor set, which are marked with squares.

- $y_s = y_t$, where edges $e(s)$ and $e(t)$ are joined to the root.
- $-y_i \leq 0$, i maximal in I
- $y_i - y_j - y_k \leq 0$, where $e(i), e(j), e(k)$ are three distinct edges that meet at a single node not in I ,
- $y_i + y_j + y_k \leq 2$, where $e(i), e(j), e(k)$ are three distinct edges that meet at a single node not in I ,
- $-x_i \leq 0$, for all $i \in I$,
- $x_i + x_j \leq 1$ for all $i, j \in I$ with i and j adjacent
- $x_i + y_i \leq 1$ for i maximal in I
- $2 \sum_{i \in C} x_i + \sum_{j \in N_I(C)} x_j + y_{m(C)} \leq \#C + 1$ for all clusters $C \subset I$.

Note that if $m(C)$ is not a maximal node of I , then there is no coordinate $y_{m(C)}$. In this case, the final cluster inequality in Theorem 4.2.11 reduces to

$$2 \sum_{i \in C} x_i + \sum_{j \in N_I(C)} x_j \leq \#C + 1.$$

Note that we have chosen to write all of our inequalities with all indeterminates on the left side and using all \leq inequalities, as this will facilitate our proof of Theorem 4.2.11.

Example 4.2.12. Consider the tree T in Figure 4.5. Note that the only cluster in T is $\{c\}$. Let $I \subset \text{Int}(T)$ be the order ideal $\{b, c, d, e\}$. Then $R_T(I)$ lies in the hyperplane $y_b = y_f$ and has facets:

$$\begin{array}{ll} y_f - y_g - y_h \leq 0, & x_b + x_c \leq 1, \\ -y_f + y_g - y_h \leq 0, & x_c + x_d \leq 1, \\ -y_f - y_g + y_h \leq 0, & x_c + x_e \leq 1, \\ y_f + y_g + y_h \leq 2, & x_b + y_b \leq 1, \\ x_b + 2x_c + x_d + x_e \leq 2 & \end{array}$$

and $-x_i \leq 0$ for all $i \in I$.

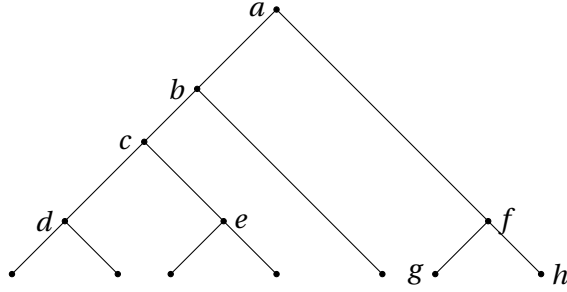


Figure 4.5 The tree T in Example 4.2.12

Proof of Theorem 4.2.11. We proceed by induction on the size of the order ideal I .

When $\#I = 0$, $R_T(\emptyset)$ is the polytope associated to the CFN model, as described in [39]. It follows from the results in [8, 39] that $R_T(\emptyset)$ has facets defined by $y_i - y_j - y_k \leq 0$ and $y_i + y_j + y_k \leq 2$ for all distinct i, j, k such that $e(i), e(j)$, and $e(k)$ that meet at the same internal node.

Let $\#I \geq 1$ and let $1, \dots, r$ be the maximal nodes of I . Suppose that $R_T(I - \{r\})$ has its facets defined by the proposed inequalities. About node r , we have the edges and nodes depicted in Figure 4.3. Note that it is possible that a, b or both are leaves. In the case that a is a leaf, the inequalities below in which x'_a is a term would not exist, and similarly for b and x'_b .

We use Fourier-Motzkin elimination along with the linear map $\phi_{I,r}$ to show that the facets of $R_T(I)$ are defined by a subset of the proposed inequalities.

Recall that primed coordinates such as y'_a indicate coordinates of $R_T(I - \{r\})$. In order to project $R_T(I - \{r\})$ onto $R_T(I)$, we “contract” onto r by replacing y'_b with $2x_r + y'_r - y'_a$, since under $\phi_{I,r}$,

$$x_r = \frac{-y'_r + y'_a + y'_b}{2}.$$

Then we use Fourier-Motzkin elimination to project out y'_a .

By the inductive hypothesis, the following are the inequalities in $R_T(I - \{r\})$ that involve y'_a or y'_b . Note that these are the only types of inequalities that we need to consider, since any inequalities not involving y'_a or y'_b remain unchanged by Fourier-Motzkin elimination.

$$\begin{array}{ll}
-y'_a \leq 0, & y'_a - y'_b - y'_r \leq 0, \\
-y'_b \leq 0, & -y'_a + y'_b - y'_r \leq 0, \\
x'_a + y'_a \leq 1, & -y'_a - y'_b + y'_r \leq 0, \\
x'_b + y'_b \leq 1, & y'_a + y'_b + y'_r \leq 2, \\
2 \sum_{i \in C} x'_i + \sum_{j \in N_{I-\{r\}}(C)} x'_j + y'_a \leq \#C + 1, & 2 \sum_{i \in D} x'_i + \sum_{j \in N_{I-\{r\}}(D)} x'_j + y'_b \leq \#D + 1,
\end{array}
\quad \text{where } C$$

ranges over all clusters that contain a and are contained in the subtree beneath a , and D ranges overall clusters that contain b and are contained in the subtree beneath b . The same is true of C

and D throughout the following discussion. Note that if a (resp. b) is not a cluster node, then no such C (resp. D) exists.

Applying $\phi_{I,r}$ yields the following inequalities, labeled by whether the coefficient of y_a is positive or negative in order to facilitate Fourier-Motzkin elimination.

$$\begin{aligned}
& -x_r \leq 0 && (0) \\
& x_r + y_r \leq 1 && (00) \\
& y_a - y_r - 2x_r \leq 0 && (1_+) \\
& y_a + x_a \leq 1 && (2_+) \\
& y_a - y_r - x_r \leq 0 && (3_+) \\
& y_a + 2 \sum_{i \in C} x_i + \sum_{j \in N_{I-\{r\}}(C)} x_j \leq \#C + 1 && (4_+) \\
& -y_a \leq 0 && (1_-) \\
& -y_a + y_r + x_b + 2x_r \leq 1 && (2_-) \\
& -y_a + x_r \leq 0 && (3_-) \\
& -y_a + y_r + 2x_r + 2 \sum_{i \in D} x_i + \sum_{j \in N_{I-\{r\}}(D)} x_j \leq \#D + 1 && (4_-)
\end{aligned}$$

If, without loss of generality, a is an internal node and b is a leaf, then inequalities 2_- and 4_- do not exist. If both a and b are leaves, then inequalities 2_+ , 4_+ , 2_- and 4_- do not exist.

We perform Fourier-Motzkin elimination to obtain the following 17 types of inequalities, labeled by which of the above inequalities were combined to obtain them. The inequalities from $R_T(I - \{r\})$ that did not contain y'_a or y'_b also remain facet-defining inequalities for $R_T(I)$.

$$\begin{aligned}
& -x_r \leq 0 & (0) \\
& x_r + y_a \leq 1 & (00) \\
& -2x_r - y_r \leq 0 & (1_+1_-) \\
& x_b \leq 1 & (1_+2_-) \\
& -x_r - y_r \leq 0 & (1_+3_-) \\
& 2 \sum_{i \in D} x_i + \sum_{j \in N_{I-(r)}(D)} x_j \leq \#D + 1 & (1_+4_-) \\
& x_a \leq 1 & (2_+1_-) \\
& 2x_r + x_a + x_b + y_r \leq 2 & (2_+2_-) \\
& x_r + x_a \leq 1 & (2_+3_-) \\
& 2 \sum_{i \in DU\{r\}} x_i + y_r + \sum_{j \in N_I(DU\{r\})} x_j \leq \#D + 2 & (2_+4_-) \\
& -x_r - y_r \leq 0 & (3_+1_-) \\
& x_r + x_b \leq 1 & (3_+2_-) \\
& -y_r \leq 0 & (3_+3_-) \\
& 2 \sum_{i \in D} x_i + \sum_{j \in N_I(D)} x_j \leq \#D + 1 & (3_+4_-) \\
& 2 \sum_{i \in C} x_i + \sum_{j \in N_{I-(r)}(C)} x_j \leq \#C + 1 & (4_+1_-) \\
& 2 \sum_{i \in CU\{r\}} x_i + y_r + \sum_{j \in N_I(CU\{r\})} x_j \leq \#C + 2 & (4_+2_-) \\
& 2 \sum_{i \in C} x_i + x_r + \sum_{j \in N_I(C)} x_j \leq \#C + 1 & (4_+3_-) \\
& 2x_r + 2 \sum_{i \in C} x_i + 2 \sum_{j \in D} x_j + y_r + \sum_{k \in N_{I-(r)}(C)} x_k + \sum_{l \in N_{I-(r)}(D)} x_l \leq \#C + \#D + 2 & (4_+4_-)
\end{aligned}$$

The inequalities encompassed by 4₊2₋ (resp. 2₊4₋) give the proposed inequalities for all clusters of size greater than or equal to 2 that contain r and for which all other nodes are contained in the a -subtree (resp. b -subtree). The inequalities given by 4₊3₋ (resp. 1₊4₋) are all of the proposed inequalities for clusters containing a (resp. b) and not r . Inequality 2₊2₋ gives the inequality for the cluster $\{r\}$. Finally, the inequalities given by 4₊4₋ encompass all clusters with the highest node r that contain nodes in both the a - and b -subtrees.

Note also that inequalities 1₊1₋, 1₊2₋, 1₊3₋, 1₊4₋, 2₊1₋, 3₊1₋, 3₊4₋ and 4₊1₋ are all redundant as they are positive linear combinations of other inequalities on the list. For instance, inequality 1₊1₋ can be obtained by adding together two copies of inequality 0 and 3₊3₋. Inequality 1₊4₋ can be obtained by adding together inequalities 3₊4₋ and 0.

Note that if, without loss of generality, a is an internal node and b is a leaf, then the irredundant

inequalities $2_+2_-, 2_+4_-, 3_+2_-, 3_+4_-, 4_+2_-$ and 4_+4_- would not exist. This is because b is not an internal node and r is not a cluster node in this case. Similarly, if a and b are both leaves, then the irredundant inequalities $2_+2_-, 2_+3_-, 2_+4_-, 3_+2_-, 3_+4_-, 4_+2_-$ and 4_+4_- would not exist.

The remaining inequalities, along with the others that are unchanged because they did not involve y'_a and y'_b are exactly those that we claimed would result from contracting onto r , as needed. \square

Corollary 4.2.13. *The facet-defining inequalities of R_T are:*

- $x_i \geq 0$, for all $1 \leq i \leq n-1$,
- $x_i + x_j \leq 1$, for all pairs of adjacent nodes, i and j , and
- $2 \sum_{i \in C} x_i + \sum_{j \in N_T(C)} x_j \leq \#C + 1$ for all clusters C in T .

Proof. Let r be the root of T . Since r is not a descendent of any edge, we interpret the coordinate y_r in R_T to be zero. The fact that the facet defining inequalities of R_T are a subset of the inequalities given in Corollary 4.2.13 along with the inequality $x_r \leq 1$ follows directly from Theorem 4.2.11, since $x_r \leq 1$ is obtained by setting $y_r = 0$ in $x_r + y_r \leq 1$. Note further that $x_r \leq 1$ is a redundant inequality, since for any node a adjacent to r , this inequality can be obtained by summing $x_r + x_a \leq 1$ and $-x_a \leq 0$. So indeed, the facet defining inequalities of R_T are a subset of the proposed inequalities.

Now we must show that none of the proposed inequalities are redundant. To do this, we find $n-1$ affinely independent vertices of R_T that lie on each of the proposed facets.

For all facets of the form $\{\mathbf{x} \mid x_i = 0\}$, the 0 vector, along with each of the standard basis vectors \mathbf{e}_j such that $j \neq i$ are $n-1$ affinely independent vertices that lie on the face. So, $\{\mathbf{x} \mid x_i = 0\}$ is a facet of R_T .

Consider a face of the form $F = \{\mathbf{x} \mid x_i + x_j = 1\}$ where i and j are adjacent nodes of T . Without loss of generality, let i be a descendant of j . First, note that $\mathbf{e}_i, \mathbf{e}_j \in F$.

Let $k \neq i, j$ be an internal node of T . If k is not a node in the i -subtree, then $\mathbf{e}_i + \mathbf{e}_k \in F$, since either k is in the subtree of T rooted at the descendant of j not equal to i , or k lies above j . In the first case, since the i - and k -subtrees are disjoint, we may choose any paths with highest nodes i and k , which yield the desired vertex. In the second case, picking a path with highest node i , and a path with highest node k that passes through the descendant of j not equal to i yields that $\mathbf{e}_i + \mathbf{e}_k$ is a vertex of R_T . Similarly, for all k in the i -subtree, $\mathbf{e}_j + \mathbf{e}_k \in F$. Since every standard basis vector is in the linear span of

$$\{\mathbf{e}_i, \mathbf{e}_j\} \cup \{\mathbf{e}_i + \mathbf{e}_k \mid k \neq j, k \text{ not in the } i\text{-subtree}\} \cup \{\mathbf{e}_j + \mathbf{e}_k \mid k \neq i, k \text{ in the } i\text{-subtree}\},$$

these $n-1$ vectors are linearly independent.

Finally, consider a face of the form $F = \{\mathbf{x} \mid 2 \sum_{c \in C} x_c + \sum_{i \in N(C)} x_i = \#C + 1\}$, for some cluster C in T . Then $\#N(C) = \#C + 2$. First note that $\mathbf{u}_j = \sum_{i \in N(C)} \mathbf{e}_i - \mathbf{e}_j$ is a vertex of R_T for all $j \in N(C)$. If j is the highest node of $N(C)$, then the i -subtrees for $i \in N(C), i \neq j$ are disjoint. So any two paths

with highest nodes $i, k \in N(C)$, $i \neq k \neq j$ are disjoint. If j is not the highest node in $N(C)$, let k be the highest node. Then we may use any paths with highest nodes i for all $i \in N(C)$ with $i \neq j, k$, and then a path with highest node k that passes through j . Since the path between k and j contains only j, k and elements of C , this path does not pass through any i -subtree for $i \in N(C)$, $i \neq j, k$. So, these paths are disjoint, as needed. So, $\{\mathbf{u}_j \mid j \in N(C)\}$ is a set of $\#C + 2$ linearly independent vertices of R_T that lie on F .

For all $c \in C$, let $\mathbf{w}_c = \mathbf{e}_c + \sum_{i \in A_c} \mathbf{e}_i$ where A_c is a set of $\#C - 1$ elements of $N(C)$ such that (1) if $i \in N(C)$ is adjacent to c , then $i \notin A_c$, (2) there exist $i, j \notin A_c$ that are in the left and right subtrees beneath c , respectively, and (3) if i is the highest node in $N(C)$, then $i \notin A_c$. Note that at least one such set exists for all $c \in C$. Then \mathbf{w}_c is a vertex of R_T since it results from the path system containing a path with highest node c that passes through i and j , where i, j are the descendants of c not in A_c that exist by condition (2), and a path with highest node k for all $k \in A_c$. Furthermore, $\mathbf{w}_c \in F$ for all $c \in C$.

Note that $\{\mathbf{u}_j \mid j \in N(C)\} \cup \{\mathbf{w}_c \mid c \in C\}$ is a linearly independent set, since $\{\mathbf{u}_j \mid j \in N(C)\}$ is a linearly independent set of vectors that have all coordinates corresponding to elements of C equal to 0, and each \mathbf{w}_c has a unique nonzero coordinate corresponding to $c \in C$.

Let k be an internal node of T such that $k \notin C \cup N(C)$. If k is a descendant of j for some $j \in N(C)$ that is not the highest node of $N(C)$, then $\mathbf{z}_k = \mathbf{u}_j + \mathbf{e}_k$ is a vertex of R_T that lies on F . Otherwise, k is either a descendant of only the highest node, i , of $N(C)$, or not a descendant of any element of $N(C)$. In either of these cases, $\mathbf{z}_k = \mathbf{u}_i + \mathbf{e}_k$ is a vertex of R_T that lies on F .

Also, $\{\mathbf{u}_j \mid j \in N(C)\} \cup \{\mathbf{w}_c \mid c \in C\} \cup \{\mathbf{z}_k \mid k \notin C \cup N(C)\}$ is a linearly independent set as each element of $\{\mathbf{u}_j \mid j \in N(C)\} \cup \{\mathbf{w}_c \mid c \in C\}$ has coordinates corresponding to nodes not in C or $N(C)$ equal to 0, and each \mathbf{z}_k has a unique nonzero coordinate corresponding to $k \notin N(C) \cup C$. This set also has cardinality $\#C + 2 + \#C + n - 2\#C - 3 = n - 1$. So, since we have found $n - 1$ linearly independent vertices of R_T that lie on F , F is a facet of R_T . \square

We conclude this section with the remark that the number of facets of R_T varies widely for different tree topologies. For a tree with n leaves and no cluster nodes, there are $2n - 3$ facets of R_T corresponding to each non-negativity condition and each of the facets arising from adjacent nodes. In contrast, the following is an example of a construction of trees with exponentially many facets.

Example 4.2.14. Let m be a positive integer. We construct a tree T_m with $4m + 5$ leaves as follows. Begin with a path, or "spine", of length m . To the top node of this spine, attach a single pendant leaf; this top node becomes the root of T_m . Attach a balanced 4-leaf tree descended from every node of the spine, with two attached to the node at the bottom of the spine. There are $2m + 1$ cluster nodes in T_m : the nodes that are in the spine and the root of each of the balanced 4-leaf trees descended from the spine. Figure 4.6 depicts this tree for $m = 3$.

Let S be the set of all nodes in the spine, and let A be any set of nodes immediately descended from a spine node. Then $S \cup A$ is a cluster. Clusters of this form account for 2^{m+1} facets of R_T for this $(4m + 5)$ -leaf tree.

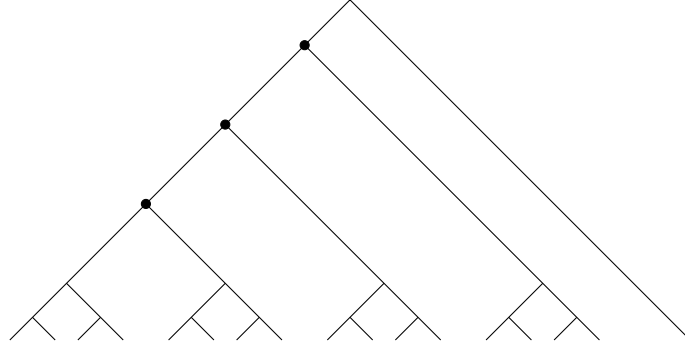


Figure 4.6 The tree construction for T_3 described in Example 4.2.14. “Spine” nodes are marked with a circle.

The aim of this section is to prove the following theorem.

Theorem 4.2.15. *For any rooted binary phylogenetic tree T , the CFN-MC ideal I_T has a Gröbner basis consisting of homogeneous quadratic binomials with squarefree initial terms.*

To accomplish this, we show that for most trees T , the CFN-MC ideal is the toric fiber product of the ideals of two smaller trees. In these cases, we can use results from [41] to describe the generators of I_T in terms of the generators of the ideals of these smaller trees. We then handle the case of trees for which I_T is not a toric fiber product; such trees are called *cluster trees*.

For simplicity of notation, we switch to denoting the top-vector associated to a path system \mathfrak{P} by $[\mathfrak{P}]$. As before, note that it is possible to have two different path systems \mathfrak{P} and \mathfrak{Q} for which $[\mathfrak{P}] = [\mathfrak{Q}]$. We often make use of the following notion of restriction of a path system to a subtree.

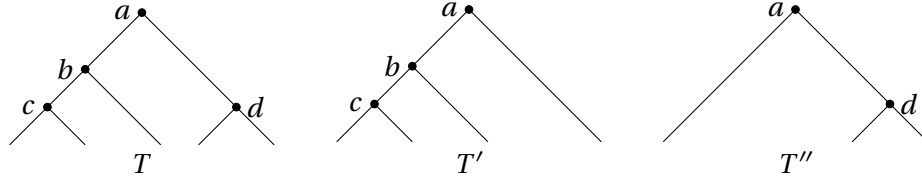
Definition 4.2.16. Let T be a tree and let T' be a subtree of T . Let \mathfrak{P} be a path system in T . Then the *restriction* of \mathfrak{P} to T' is the path system \mathfrak{P}' in T' obtained by the following procedure for each path $P \in \mathfrak{P}$. If the top-most node of P is not in $\text{Int}(T')$, delete P . Otherwise, intersect the edges of P with the edges of T' to obtain a path P' , and add P' to \mathfrak{P}' .

Note that if \mathfrak{P}' is the restriction of \mathfrak{P} to T' , then $[\mathfrak{P}']$ is equal to $[\mathfrak{P}]$ on each coordinate in $\text{Int}(T')$.

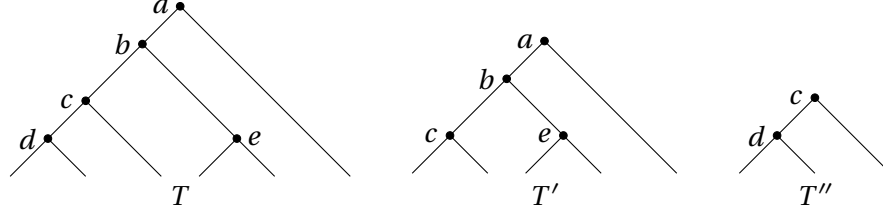
4.2.1 Toric Fiber Products

Let T be a tree that has an internal node v that is adjacent to exactly two other internal nodes. There are two cases for the position of v within T , both of which provide a natural way to divide T into two smaller trees, T' and T'' .

If v is the root, then let T' be the tree with v as a root in which the right subtree of v is equal to the right subtree of T and the left subtree of v is a single edge. Let T'' be the tree with v as a root in which the left subtree of v is equal to the left subtree of T and the right subtree of v is a single edge. This decomposition is pictured in Figure 4.7a.



(a) Splitting T into T' and T'' where distinguished node (a) is the root.



(b) Splitting T into T' and T'' where distinguished node (c) is not the root.

Figure 4.7 Decomposition of T into T' and T'' via a node adjacent to exactly two internal nodes

If v is not the root, then v is adjacent to two internal nodes and a leaf in T . Let T' be the tree consisting of all non-descendants of v (including v itself) with a cherry added below v . Let T'' be the tree consisting of v and all of its descendants. This decomposition is pictured in Figure 4.7b.

In either case, notice that $[\mathfrak{P}]$ is the top-vector of a path system in T if and only if the restrictions of $[\mathfrak{P}]$, $[\mathfrak{P}']$ and $[\mathfrak{P}'']$ to T' and T'' respectively are top-vectors of path systems in T' and T'' that agree on v . The following lemma makes this observation precise.

Lemma 4.2.17. *Suppose that T has an internal node v that is adjacent to exactly two other internal nodes. Let T' and T'' be the induced trees defined above depending upon the position of v within T . Let \mathfrak{P} be a path system in T' and let \mathfrak{R} be a path system in T'' . Then there exists a path system $\mathfrak{P} \vee \mathfrak{R}$ in T such that $[\mathfrak{P} \vee \mathfrak{R}]_i = [\mathfrak{P}]_i$ for each $i \in \text{Int}(T')$ and $[\mathfrak{P} \vee \mathfrak{R}]_j = [\mathfrak{R}]_j$ for each $j \in \text{Int}(T'')$.*

Proof. First, consider the case where v is the root. Then T' is the tree with root v whose left subtree is equal to that of T and whose right subtree is a single leaf. Similarly, T'' is the tree with root v whose right subtree is equal to that of T and whose left subtree is a single leaf. Let \mathfrak{P} be a path system in T' and let \mathfrak{R} be a path system in T'' .

If $[\mathfrak{P}]_v = [\mathfrak{R}]_v = 0$, then no path in \mathfrak{P} or \mathfrak{R} passes through v . So each path in \mathfrak{P} and \mathfrak{R} is also a path in T . So we let $\mathfrak{P} \vee \mathfrak{R} = \mathfrak{P} \cup \mathfrak{R}$, where the edge set of each path is a subset of the edges of T .

If $[\mathfrak{P}]_v = [\mathfrak{R}]_v = 1$, then let \bar{P} be the path of \mathfrak{P} whose top-most node is v and let \bar{R} be the path of \mathfrak{R} whose top-most node is v . Let $\overline{P\bar{R}}$ be the path in T with edge set equal to that of P on the left subtree of T and that of R on the right subtree of T . This is a path in T with top-most node v . In this case, let $\mathfrak{P} \vee \mathfrak{R} = (\mathfrak{P} \cup \mathfrak{R} \cup \{\overline{P\bar{R}}\}) \setminus \{P, R\}$, where the edge set of each path is a subset of the edges of T .

Now consider the case where v is not the root. Let ℓ be the leaf of T that is adjacent to v . Then T' consists of all non-descendants of v and two leaves below v . One of these leaves is ℓ ; let m be the other leaf below v in T' . The tree T'' consists of v and all of its descendants. Let \mathfrak{P} be a path system in T' and let \mathfrak{R} be a path system in T'' .

First, suppose $[\mathfrak{P}]_v = [\mathfrak{A}]_v = 0$. Since $[\mathfrak{P}]_v = 0$, we may assume that any path in \mathfrak{P} that passes through v passes through ℓ and not m . Since v is the root of T' , and $[\mathfrak{A}]_v = 0$, no path in \mathfrak{A} passes through v or ℓ . So we let $\mathfrak{P} \vee \mathfrak{A} = \mathfrak{P} \cup \mathfrak{A}$, where the edge set of each path is a subset of the edges of T .

Now suppose $[\mathfrak{P}]_v = [\mathfrak{A}]_v = 1$. Then \mathfrak{P} contains the path in T' between leaves m and ℓ . So we let $\mathfrak{P} \vee \mathfrak{A} = (\mathfrak{P} \cup \mathfrak{A}) \setminus \{P\}$, where the edge set of each path is a subset of the edges of T . \square

One implication of Lemma 4.2.17 is that the matrix A_T of I_T can be obtained by pairing together all columns in $A_{T'}$ and $A_{T''}$ that agree on v , and consolidating the rows corresponding to v and the homogenizing rows of ones from each. This translates exactly to the operation on toric ideals known as the *toric fiber product*, which was introduced in [41].

Let $I_T \subset \mathbb{C}[\mathbf{r}]$, $I_{T'} \subset \mathbb{C}[\mathbf{x}]$, $I_{T''} \subset \mathbb{C}[\mathbf{y}]$. Consider the map $\xi_{I_{T'}, I_{T''}}$ from $\mathbb{C}[\mathbf{r}]$ to $\mathbb{C}[\mathbf{x}] \otimes_{\mathcal{A}} \mathbb{C}[\mathbf{y}]$ defined as follows. For any path system \mathfrak{P} in T , let \mathfrak{P}' and \mathfrak{P}'' be the restrictions of \mathfrak{P} to T' and T'' respectively. Then

$$\xi_{I_{T'}, I_{T''}}(r_{[\mathfrak{P}]}) = x_{[\mathfrak{P}']} \otimes y_{[\mathfrak{P}'']}.$$

Following the notation of [41], the kernel of $\xi_{I_{T'}, I_{T''}}$ is the *toric fiber product* $I_{T'} \times_{\mathcal{A}} I_{T''}$. Here, \mathcal{A} is the matrix

$$\mathcal{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},$$

where the first row corresponds to the homogenizing row of ones, and the second row corresponds to the shared node v of T' and T'' . By Lemma 4.2.17, we can join any path systems in T' and T'' that whose top-vectors agree on v to create a path system in T .

Proposition 4.2.18. *Suppose that T has an internal node v that is adjacent to exactly two other internal nodes. Let T' and T'' be the induced trees defined above depending upon the position of v within T . Then $I_T \cong I_{T'} \times_{\mathcal{A}} I_{T''}$.*

Proof. The monomial map of which I_T is the kernel is given by

$$\begin{aligned} \psi_T : \mathbb{C}[\mathbf{r}] &\rightarrow \mathbb{C}[t_0, \dots, t_{n-1}] \\ r_{[\mathfrak{P}]} &\mapsto t_0 \prod_{[\mathfrak{P}]_i=1} t_i \end{aligned}$$

Let S be the two-leaf tree rooted at v , and let $\mathbb{C}[\mathbf{z}] = \mathbb{C}[z_0, z_1]$ be its associated polynomial ring. Denote by $\overline{\mathfrak{P}}$ the restriction of \mathfrak{P} to S , by \mathfrak{P}' the restriction of \mathfrak{P} to T' and by \mathfrak{P}'' the restriction of \mathfrak{P} to T'' . Then we have the identity

$$\psi_S(z_{[\overline{\mathfrak{P}}]}) \psi_T(r_{[\mathfrak{P}]}) = \psi_{T'}(x_{[\mathfrak{P}']}) \psi_{T''}(y_{[\mathfrak{P}'']}). \quad (4.7)$$

The map defining the toric fiber product $I_{T'} \times_{\mathcal{A}} I_{T''}$ can be written

$$\begin{aligned} \xi_{I_{T'}, I_{T''}} : \mathbb{C}[\mathbf{r}] &\rightarrow \mathbb{C}[t_0, \dots, t_{n-1}] \\ r_{[\mathfrak{P}]} &\mapsto \left(t_0 \prod_{[\mathfrak{P}'_i]=1} t_i \right) \left(t_0 \prod_{[\mathfrak{P}''_i]=1} t_i \right) \end{aligned}$$

Notice that t_0 and t_v can only appear in the image of $\xi_{I_{T'}, I_{T''}}$ with exponent 2. Therefore we may replace these variables by their square roots in the formula for the image of each $r_{[\mathfrak{P}]}$ in the map $\xi_{I_{T'}, I_{T''}}$ without changing the kernel. This yields the same map as ψ_T . Therefore,

$$I_T = \ker \psi_T \cong \ker \xi_{I_{T'}, I_{T''}} = I_{T'} \times_{\mathcal{A}} I_{T''}. \quad \square$$

Let \mathcal{G}_1 be a Gröbner basis for $I_{T'}$ with weight vector ω_1 , and let \mathcal{G}_2 be a Gröbner basis for $I_{T''}$ with weight vector ω_2 . From these, we define several sets of polynomials in $\mathbb{C}[\mathbf{r}]$ that together form a Gröbner basis for I_T with respect to some weighted monomial order. Let $\mathfrak{P}_1, \dots, \mathfrak{P}_d$ and $\Omega_1, \dots, \Omega_d$ be path systems in T' such that $f = \prod_{i=1}^d x_{[\mathfrak{P}_i]} - \prod_{i=1}^d x_{[\Omega_i]} \in \mathcal{G}_1$. We arrange these so that $[\mathfrak{P}_i]_v = [\Omega_i]_v$ for all i . Note that f can always be written in this form since the parameter t_v must appear with the same power in the image of each monomial under $\psi_{T'}$ in order for f to be in its kernel. Let $R(f)$ denote the set of all d -tuples $(\mathfrak{R}'_1, \dots, \mathfrak{R}'_d)$ of path systems in T'' such that $[\mathfrak{R}'_i]_v = [\mathfrak{P}_i]_v$.

By Lemma 4.2.17, for any path systems \mathfrak{P} in T' and \mathfrak{R} in T'' with $[\mathfrak{P}]_v = [\mathfrak{R}]_v$, we can find a path system $\mathfrak{P} \vee \mathfrak{R}$ in T such that $[\mathfrak{P} \vee \mathfrak{R}]_i = [\mathfrak{P}]_i$ for each $i \in \text{Int}(T')$ and $[\mathfrak{P} \vee \mathfrak{R}]_j = [\mathfrak{R}]_j$ for each $j \in \text{Int}(T'')$.

Define the set

$$\text{Lift}(f) = \left\{ \prod_{i=1}^d r_{[\mathfrak{P}_i \vee \mathfrak{R}'_i]} - \prod_{i=1}^d r_{[\Omega_i \vee \mathfrak{R}'_i]} : (\mathfrak{R}'_1, \dots, \mathfrak{R}'_d) \in R(f) \right\}.$$

Then let

$$\text{Lift}(\mathcal{G}_1) = \cup_{f \in \mathcal{G}_1} \text{Lift}(f),$$

and similarly define $\text{Lift}(\mathcal{G}_2)$.

We now define another family of polynomials that is contained in the Gröbner basis for I_T . Let $[\mathfrak{P}_1], \dots, [\mathfrak{P}_r]$ be the top-vectors of paths in T' with v -coordinate 0 and let $[\Omega_1], \dots, [\Omega_s]$ be the top-vectors of paths in T'' with v -coordinate 0. (Note that these \mathfrak{P}_i and Ω_i are unrelated to those in the previous paragraph). Define the set $\text{Quad}_0(T)$ to be the set of all 2×2 minors of the matrix $M_0(T)$ with (i, j) th entry equal to $r_{[\mathfrak{P}_i \vee \Omega_j]}$. Define $\text{Quad}_1(T)$ and $M_1(T)$ analogously over all top-vectors in T' and T'' with v -coordinate equal to 1. Elements of Quad_k are of the form

$$r_{[\mathfrak{P}_i \vee \Omega_j]} r_{[\mathfrak{P}'_{i'} \vee \Omega_{j'}]} - r_{[\mathfrak{P}_i \vee \Omega_{j'}]} r_{[\mathfrak{P}'_{i'} \vee \Omega_j]},$$

where $[\mathfrak{P}_i], [\Omega_j], [\mathfrak{P}_{i'}]$ and $[\Omega_{j'}]$ all take value k on their v -coordinate. Define

$$\text{Quad}(T) = \text{Quad}_0(T) \cup \text{Quad}_1(T).$$

Let ω be a weight vector on $\mathbb{C}[\mathbf{r}]$ so that $\text{Quad}(T)$ is a Gröbner basis for the ideal generated by all elements of $\text{Quad}(T)$. Such a weight vector exists by the proof of Proposition 2.6 in [41]. Since the \mathcal{A} -matrix of the toric fiber product is invertible, Theorem 2.9 of [41] implies the following proposition. Denote by $\xi_{I_{T'}, I_{T''}}^*$ the pullback of $\xi_{I_{T'}, I_{T''}}$. In other words, $\xi_{I_{T'}, I_{T''}}^*$ is a map from the Cartesian products of the affine spaces associated to $\mathbb{C}[\mathbf{x}]$ and $\mathbb{C}[\mathbf{y}]$ to the affine space associated to $\mathbb{C}[\mathbf{r}]$. If \mathfrak{P} and Ω are path systems in T' and T'' respectively whose top-vectors agree on v , then the $[\mathfrak{P} \vee \Omega]$ coordinate of $\xi_{I_{T'}, I_{T''}}^*(\alpha, \beta)$ is $\alpha_{[\mathfrak{P}]} + \beta_{[\Omega]}$.

Proposition 4.2.19. *Suppose that T has an internal node v that is adjacent to exactly two other nodes. Let T' and T'' be the induced trees defined above depending upon the position of v within T . Then $\text{Lift}(\mathcal{G}_1) \cup \text{Lift}(\mathcal{G}_2) \cup \text{Quad}(T)$ is a Gröbner basis for I_T with respect to weight vector $\xi_{I_{T'}, I_{T''}}^*(\omega_1, \omega_2) + \epsilon \omega$ for sufficiently small $\epsilon > 0$.*

In particular, note that since the Lift operation preserves degree, and since the elements of $\text{Quad}(T)$ are quadratic, if \mathcal{G}_1 and \mathcal{G}_2 consist of quadratic binomials, then I_T has a Gröbner basis consisting of quadratic binomials.

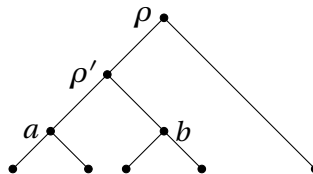
4.2.2 Cluster Trees

Trees that do not have an internal node that is adjacent to exactly two other internal nodes do not have the toric fiber product structure described in the previous section. These are the trees whose internal nodes are comprised of one large cluster and its neighbor nodes. In this case, we exploit the toric fiber product structure of a subtree, and describe a method for lifting the Gröbner basis for the subtree to a Gröbner basis for the entire tree that maintains the degree of the Gröbner basis elements.

Definition 4.2.20. A rooted binary tree T is called a *cluster tree* if there exists a cluster C in T such that every internal node of T is either in C or in $N(C)$. Note that if T is a cluster tree, then C is uniquely determined by T .

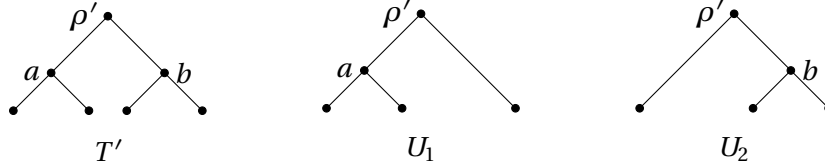
Equivalently, if T has n leaves, then T is a cluster tree if and only if T has a cluster of size $(n-3)/2$. Note that this implies that if T is a cluster tree, then T has an odd number of leaves. It also follows from the definition of a cluster tree that the root of T must be adjacent on one side to a single leaf.

Example 4.2.21. The following tree is an example of a cluster tree with $\{\rho'\}$ as its distinguished cluster.



Let T be a cluster tree with root ρ . Consider the tree T' obtained from T by deleting ρ and its adjacent edges. Let ρ' be the root of T' . Then by Proposition 4.2.18, the CFN-MC ideal of T' , $I_{T'}$ is the toric fiber product $I_{U_1} \times_{\mathcal{A}} I_{U_2}$ where U_1 and U_2 are the cluster trees with root ρ' and maximal clusters given by the left and right subtrees of ρ' respectively. So we call T' a *bicluster tree*.

Example 4.2.22. From the previous example, T' , U_1 and U_2 are as follows.



We are interested in defining when we can add a path with highest node at ρ to a path system in the larger cluster tree, T . This motivates the following definition of *root-augmentability*.

Definition 4.2.23. A path system \mathfrak{P} is *root-augmentable* if there exists a path P' between the leaves of T that has the root as its top-most node and is disjoint from all paths in \mathfrak{P} . In other words, $[\mathfrak{P}]$ has root-coordinate equal to 0, but setting it equal to 1 would still yield a valid top-vector.

Let T be a cluster tree and \mathfrak{P} be a path system T . Then \mathfrak{P} is root-augmentable if and only if this path system does not already have a path with highest node ρ and the restriction of the path system to T' has the following property.

Definition 4.2.24. A path system \mathfrak{P} in a bicluster tree T' is *root-leaf traversable* if there exists a path from the root to some leaf that does not include any internal node that is the top-most node of some path in \mathfrak{P} .

Since T' is a bicluster tree, in order for a path system in T' to be root-leaf traversable, one must be able to add a path from ρ' through the clusters of either U_1 or U_2 to a leaf. Note that a path system is root-leaf traversable if and only if removing all of the maximal nodes of paths in the path system leaves ρ' in the same connected component as some leaf of T' . Therefore, root-leaf traversability is well-defined over classes of path systems with the same top-set. We often say that $[\mathfrak{P}]$ is root-leaf traversable if \mathfrak{P} is root-leaf traversable.

Note that we cannot use the same definition for root-augmentability and root-leaf traversability. Indeed, in a cluster tree, any path system all of whose paths do not contain the root must be root-leaf traversable. Root-augmentability should be thought of as the non-trivial notion of root-leaf traversability for cluster trees.

We can now define a special type of term order on the polynomial ring of the CFN-MC ideal of a cluster tree, and its analogue for that of a bicluster tree.

Definition 4.2.25. Let S be a cluster tree with CFN-MC ideal $I_S \subset \mathbb{C}[\mathbf{x}]$. A term order $<$ on $\mathbb{C}[\mathbf{x}]$ is *liftable* if

1. I_S has a $<$ -Gröbner basis consisting of degree 2 binomials, and

2. $<$ is a block order on I_S with blocks

$$\{x_{[\mathfrak{P}]} \mid \mathfrak{P} \text{ not root augmentable}\} > \{x_{[\mathfrak{P}]} \mid \mathfrak{P} \text{ root augmentable}\}$$

where the order induced on each block is graded.

Definition 4.2.26. Let S be a bicluster tree with CFN-MC ideal $I_S \subset \mathbb{C}[\mathbf{x}]$. A term order $<$ on $\mathbb{C}[\mathbf{x}]$ is *liftable* if

1. I_S has a $<$ -Gröbner basis consisting of degree 2 binomials, and
2. $<$ is a block order on I_S with blocks

$$\{x_{[\mathfrak{P}]} \mid \mathfrak{P} \text{ not root-leaf traversable}\} > \{x_{[\mathfrak{P}]} \mid \mathfrak{P} \text{ root-leaf traversable}\}$$

where the order induced on each block is graded.

If \mathcal{G} is the $<$ -Gröbner basis for the CFN-MC ideal of a cluster or bicluster tree for a liftable turn order $<$, then \mathcal{G} is *liftable*.

Definition 4.2.27. Let S be a cluster tree with CFN-MC ideal $I_S \subset \mathbb{C}[\mathbf{x}]$. Let $<$ be a term order on $\mathbb{C}[\mathbf{x}]$. Let $f = \prod_{i=1}^d x_{[\mathfrak{P}_i]} - \prod_{i=1}^d x_{[\mathfrak{Q}_i]}$ be a binomial in I_S whose leading term is $\prod_{i=1}^d x_{[\mathfrak{P}_i]}$. We say that f satisfies the *liftability property* with respect to $<$ if

$$\#\{\mathfrak{P} \mid \mathfrak{P} \text{ not root-augmentable}\} \geq \#\{\mathfrak{Q} \mid \mathfrak{Q} \text{ root augmentable}\}.$$

We define the liftability property when S is a bicluster tree analogously with respect to root-leaf traversability. When the monomial order has been previously specified, we just say that the polynomial satisfies the liftability property.

Note that a term order $<$ is liftable if and only if it induces a quadratic Gröbner basis all of whose elements satisfy the liftability property with respect to $<$.

Let $I_{U_1} \subset \mathbb{C}[\mathbf{x}]$, $I_{U_2} \subset \mathbb{C}[\mathbf{y}]$, $I_{T'} \subset \mathbb{C}[\mathbf{z}]$ and $I_T \subset \mathbb{C}[\mathbf{r}]$. Note that if T is the smallest cluster tree with five leaves, then U_1 and U_2 are both trees with three leaves, so I_{U_1} and I_{U_2} are the zero ideal. Therefore, they vacuously have liftable Gröbner bases. By induction, let ω_1, ω_2 be weight vectors that induce liftable orders on I_{U_1} and I_{U_2} , respectively. Let \mathcal{G}_1 be the liftable Gröbner basis for I_{U_1} and \mathcal{G}_2 the liftable Gröbner basis for I_{U_2} . Let \mathbf{a} be the weight vector on $\mathbb{C}[\mathbf{z}]$ defined by $\mathbf{a}(z_{[\mathfrak{P}]}) = 1$ if \mathfrak{P} is not root-leaf traversable and $\mathbf{a}(z_{[\mathfrak{P}]}) = 0$ if \mathfrak{P} is root-leaf traversable.

Proposition 4.2.28. Let T' be a bicluster tree. Let U_1 and U_2 be the unique cluster trees obtained from the left and right subtrees of T' such that $I_{T'} = I_{U_1} \times_{\mathcal{A}} I_{U_2}$. There exist a weight vector ω on $\mathbb{C}[\mathbf{z}]$ and $\epsilon, k > 0$ such that $\xi_{U_1, U_2}^*(\omega_1, \omega_2) + \epsilon\omega + k\mathbf{a}$ induces a liftable order on $I_{T'}$.

Proof. By Lemma 4.2.17, we can write any path system in the bicluster tree T' as $\mathfrak{P} \vee \mathfrak{Q}$ where \mathfrak{P} is a path system in U_1 , \mathfrak{Q} is a path system in U_2 , and the top-vectors of \mathfrak{P} and \mathfrak{Q} agree on the

root of T' . Let f be an element of the Gröbner basis for $I_{T'}$ described in Proposition 4.2.19. If $f = z_{[\mathfrak{P}_1 \vee \Omega_1]} z_{[\mathfrak{P}_2 \vee \Omega_2]} - z_{[\mathfrak{P}_1 \vee \Omega_2]} z_{[\mathfrak{P}_2 \vee \Omega_1]} \in \text{Quad}_i(T')$, then

$$\begin{aligned} \xi_{U_1, U_2}^*(\omega_1, \omega_2)(z_{[\mathfrak{P}_1 \vee \Omega_1]} z_{[\mathfrak{P}_2 \vee \Omega_2]}) &= \omega_1(x_{[\mathfrak{P}_1]}) + \omega_2(y_{[\Omega_1]}) + \omega_1(x_{[\mathfrak{P}_2]}) + \omega_2(y_{[\Omega_2]}) \\ &= \xi_{U_1, U_2}^*(\omega_1, \omega_2)(z_{[\mathfrak{P}_1 \vee \Omega_2]} z_{[\mathfrak{P}_2 \vee \Omega_1]}). \end{aligned}$$

We find $\epsilon > 0$ and weight vector ω to “break ties” for leading monomials of each element of $\text{Quad}_i(T')$. If f is a 2×2 minor of $M_1(T')$, then every variable in f is not root-leaf traversable. So the choice of leading monomial does not affect the liftability property. If f is a 2×2 minor of $M_0(T')$, then there is only one case in which the number of root-leaf traversable variables in the two monomials of f varies. Without loss of generality, let \mathfrak{P}_1, Ω_1 be root augmentable and \mathfrak{P}_2, Ω_2 not. Then $\mathfrak{P}_2 \vee \Omega_2$ is not root-leaf traversable, while $\mathfrak{P}_1 \vee \Omega_1, \mathfrak{P}_1 \vee \Omega_2$ and $\mathfrak{P}_2 \vee \Omega_1$ are. So, we must select an ω so that $z_{[\mathfrak{P}_1 \vee \Omega_1]} z_{[\mathfrak{P}_2 \vee \Omega_2]}$ is chosen as the leading monomial of f .

We define this weight vector ω by assigning its values on the entries of $M_0(T')$. Arrange path systems $\mathcal{A}_1, \dots, \mathcal{A}_r$ in U_1 so that if \mathcal{A}_i is root augmentable and \mathcal{A}_j is not, then $i < j$. Arrange path systems $\mathcal{B}_1, \dots, \mathcal{B}_s$ in U_2 so that if \mathcal{B}_i is root augmentable and \mathcal{B}_j is not, then $i < j$.

Define $\omega(z_{[\mathcal{A}_i \vee \mathcal{B}_j]}) = 2^{i+j}$ for all i and j . Let $i_1 < j_1$ and $i_2 < j_2$. Then

$$\begin{aligned} \omega(z_{[\mathcal{A}_{i_1} \vee \mathcal{B}_{j_2}]} z_{[\mathcal{A}_{j_1} \vee \mathcal{B}_{i_2}]}) &= 2^{i_1+j_2} + 2^{i_2+j_1} \\ &\leq 2^{j_1+j_2-1} + 2^{j_1+j_2-1} \\ &= 2(2^{j_1+j_2-1}) \\ &= 2^{j_1+j_2} \\ &< 2^{i_1+i_2} + 2^{j_1+j_2} \\ &= \omega(z_{[\mathcal{A}_{i_1} \vee \mathcal{B}_{i_2}]} z_{[\mathcal{A}_{j_1} \vee \mathcal{B}_{j_2}]}) \end{aligned}$$

So ω chooses the correct leading term of $f \in \text{Quad}_1(T)$. We can allow ω to be any weight vector on the entries of $M_1(T')$ that chooses leading terms as in Proposition 2.6 of [41]. Pick ϵ to be small enough so that for all $g \in \text{Lift}(\mathcal{G}_1) \cup \text{Lift}(\mathcal{G}_2)$,

$$LT_{\xi_{U_1, U_2}^*}^*(\omega_1, \omega_2)(g) = LT_{\xi_{U_1, U_2}^*}^*(\omega_1, \omega_2) + \epsilon \omega(g).$$

Now we must add $k\mathbf{a}$ for some $k \geq 0$ to ensure that the correct leading term is chosen for each $f \in \text{Lift}(\mathcal{G}_1) \cup \text{Lift}(\mathcal{G}_2)$. Without loss of generality, let

$$f = z_{[\mathfrak{P}_1 \vee \mathfrak{R}_1]} z_{[\mathfrak{P}_2 \vee \mathfrak{R}_2]} - z_{[\Omega_1 \vee \mathfrak{R}_1]} z_{[\Omega_2 \vee \mathfrak{R}_2]} \in \text{Lift}(\mathcal{G}_1).$$

An analysis of all possible cases shows that the only instance in which the terms of f have a varying number of root-leaf traversable indices but $\xi_{U_1, U_2}^*(\omega_1, \omega_2)$ may not select the correct leading term occurs when, without loss of generality, $\mathfrak{P}_1, \mathfrak{R}_1$ and Ω_2 are not root augmentable and $\mathfrak{P}_2, \mathfrak{R}_2$ and Ω_1 are root augmentable. In this case, $\mathfrak{P}_1 \vee \mathfrak{R}_1$ is not root-leaf traversable and $\mathfrak{P}_2 \vee \mathfrak{R}_2, \Omega_1 \vee \mathfrak{R}_1$

and $\Omega_2 \vee \mathfrak{R}_2$ are, but $x_{[\mathfrak{P}_1]}x_{[\mathfrak{P}_2]} - x_{[\Omega_1]}x_{[\Omega_2]}$ may not have $x_{[\mathfrak{P}_1]}x_{[\mathfrak{P}_2]}$ as its leading term. Suppose that under the weight vector $\xi_{U_1, U_2}^*(\omega_1, \omega_2)$, $z_{[\Omega_1 \vee \mathfrak{R}_1]}z_{[\Omega_2 \vee \mathfrak{R}_2]}$ is the leading term of f . Adding sufficiently many copies of \mathbf{a} will change this, but we must show that adding \mathbf{a} does not change the Gröbner basis $\mathcal{G} = \text{Lift}(\mathcal{G}_1) \cup \text{Lift}(\mathcal{G}_2) \cup \text{Quad}(T')$.

Define the binomial $\bar{f} = z_{[\Omega_1 \vee \mathfrak{R}_1]}z_{[\Omega_2 \vee \mathfrak{R}_2]} - z_{[\mathfrak{P}_1 \vee \mathfrak{R}_2]}z_{[\mathfrak{P}_2 \vee \mathfrak{R}_1]}$. Note that the polynomial $\bar{f} \in \text{Lift}(\mathcal{G}_1)$ with $z_{[\Omega_1 \vee \mathfrak{R}_1]}z_{[\Omega_2 \vee \mathfrak{R}_2]}$ as the leading term under $\xi_{U_1, U_2}^*(\omega_1, \omega_2) + \epsilon\omega$, and both terms of \bar{f} have all root-leaf traversable indices. Since f and \bar{f} have the same leading term, $\mathcal{G} - \{f\}$ is still a Gröbner basis under $\xi_{U_1, U_2}^*(\omega_1, \omega_2) + \epsilon\omega$. Let \mathcal{G}' be \mathcal{G} with all such $f \in \text{Lift}(\mathcal{G}_1) \cup \text{Lift}(\mathcal{G}_2)$ that violate the liftability property removed. Then every binomial in \mathcal{G}' satisfies the liftability property, and \mathcal{G}' is still a Gröbner basis.

Let $g = m_1 - m_2 \in I_{T'}$ be a binomial. Then there exists a sequence $g_1, \dots, g_r \in \mathcal{G}'$ so that g reduces to 0 upon division by the elements of this sequence in order. Suppose that m_1 is the leading term of g in the order induced by $\xi_{U_1, U_2}^*(\omega_1, \omega_2) + \epsilon\omega$, but m_2 is the leading term in the order induced by $\xi_{U_1, U_2}^*(\omega_1, \omega_2) + \epsilon\omega + \mathbf{a}$. Then m_2 has more variables whose indices are not root-leaf traversable than m_1 . We claim that division by the same g_1, \dots, g_r , possibly in a different order, still reduces g to 0. To divide g by one of g_1, \dots, g_r , pick a g_i whose leading term divides m_2 . One must exist because all of the g_i satisfy the liftability property, so it is impossible to divide m_1 by any g_i and decrease the number of root-leaf traversable variables in it. So, we may choose an element of \mathcal{G}' to proceed with the reduction of g , and \mathcal{G}' is still a Gröbner basis for the weight order induced by $\xi_{U_1, U_2}^*(\omega_1, \omega_2) + \epsilon\omega + \mathbf{a}$. \square

For any path system \mathfrak{P} in T , let \mathfrak{P}' be the path system in T' obtained from \mathfrak{P} by deleting any path in \mathfrak{P} that contains the root ρ of T . Define two maps $\psi, \psi' : \mathbb{C}[\mathbf{r}] \rightarrow \mathbb{C}[\mathbf{z}]$ by

$$\psi(r_{[\mathfrak{P}]}) = z_{[\mathfrak{P}']}$$

and

$$\psi'(r_{[\mathfrak{P}]}) = \begin{cases} z_{[\mathfrak{P}']} & \text{if } [\mathfrak{P}]_\rho = 1, \text{ and} \\ 1 & \text{if } [\mathfrak{P}]_\rho = 0. \end{cases}$$

Let $<$ be a monomial order on $\mathbb{C}[\mathbf{z}]$ whose existence is established by Proposition 4.2.28 and let $\mathcal{G}_<$ be the liftable Gröbner basis that it induces on $I_{T'}$. Then define a monomial order $<$ on $\mathbb{C}[\mathbf{r}]$ by $\mathbf{r}^{\mathbf{b}} < \mathbf{r}^{\mathbf{c}}$ if and only if

- $\psi(\mathbf{r}^{\mathbf{b}}) < \psi(\mathbf{r}^{\mathbf{c}})$, or
- $\psi(\mathbf{r}^{\mathbf{b}}) = \psi(\mathbf{r}^{\mathbf{c}})$ and $\psi'(\mathbf{r}^{\mathbf{b}}) < \psi'(\mathbf{r}^{\mathbf{c}})$.

In words, to determine which of two monomials is bigger, we delete the root and see which is bigger in the order on T' . If those are equal, then we only look at the indices with the root-coordinate equal to 1, and then delete the root from those and see which is bigger in the order on T' .

Denote by \mathcal{F} the Gröbner basis for $I_{T'}$ induced by the term order $<$ on $\mathbb{C}[\mathbf{x}]$. Let $f = z_{[\mathfrak{P}_1]}z_{[\mathfrak{P}_2]} - z_{[\Omega_1]}z_{[\Omega_2]} \in \mathcal{F}$. Define the set $\text{Root}(f)$ to be the set of all possible binomials in I_T that result from

treating $\mathfrak{P}_1, \mathfrak{P}_2, \Omega_1$ and Ω_2 as path systems in T , with or without an added path with top-most node ρ . Specifically,

$$\text{Root}(f) = \{r_{i_1[\mathfrak{P}_1]}r_{i_2[\mathfrak{P}_2]} - r_{j_1[\Omega_1]}r_{j_2[\Omega_2]}\}$$

where $i_1, i_2, j_1, j_2 \in \{0, 1\}$ are such that

- $i_1 + i_2 = j_1 + j_2$, and
- $i_1 = 1$ only if \mathfrak{P}_1 is root-leaf traversable, and similarly for i_2, j_1, j_2 .

Denote by $\text{Root}(T) = \bigcup_{f \in G_{\prec}} \text{Root}(f)$. Define the set

$$\text{Swap}(\rho) = \{r_{1[\mathfrak{P}]}r_{0[\Omega]} - r_{0[\mathfrak{P}]}r_{1[\Omega]}\}$$

where \mathfrak{P} and Ω range over all root-leaf traversable path systems in T' . Define the set $\mathcal{G}_{\prec} = \text{Root}(T) \cup \text{Swap}(\rho)$. For the sake of brevity, we use an underline to indicate the leading term of a polynomial.

Proposition 4.2.29. *The term order \prec described above is liftable. In particular, G_{\prec} is a Gröbner basis for I_T with respect to \prec .*

Proof. First, we show that G_{\prec} constitutes a Gröbner basis. Since, as explained in Section 3, I_T is toric, it suffices to show that every binomial in I_T can be reduced via the elements of G_{\prec} . Let $\prod_{i=1}^d r_{[\mathfrak{P}_i]} - \prod_{i=1}^d r_{[\Omega_i]} \in I_T$. Then if we arrange the terms in each monomial as a table with the vector representing each $[\mathfrak{P}_i]$ (resp. $[\Omega_i]$) as a row, the column sums of each of these tables are equal. By the definition of \prec , we have

$$\prod_{i=1}^d \psi(r_{[\mathfrak{P}_i]}) \geq \prod_{i=1}^d \psi(r_{[\Omega_i]}).$$

For all \mathfrak{P}_i, Ω_i , we have $\psi(r_{[\mathfrak{P}_i]}) = z_{[\mathfrak{P}'_i]}$ and $\psi(r_{[\Omega_i]}) = z_{[\Omega'_i]}$.

We can use the elements of \mathcal{F} to reduce $\prod_{i=1}^d z_{[\mathfrak{P}'_i]} - \prod_{i=1}^d z_{[\Omega'_i]}$ in $I_{T'}$. The properties of the order on $I_{T'}$ induced by \prec guarantee that (without loss of generality) if we divide by $z_{[\mathfrak{P}'_1]}z_{[\mathfrak{P}'_2]} - z_{[\mathfrak{P}'_1]}z_{[\mathfrak{P}'_2]}$ in this reduction, then the number of \mathfrak{P}'_1 and \mathfrak{P}'_2 that are root-leaf traversable is at least the number of \mathfrak{P}'_1 and \mathfrak{P}'_2 that are root-leaf traversable. Therefore, there is a corresponding element $\underline{r_{i_1[\mathfrak{P}'_1]}r_{i_2[\mathfrak{P}'_2]} - r_{j_1[\mathfrak{P}'_1]}r_{j_2[\mathfrak{P}'_2]}} \in \text{Root}(z_{[\mathfrak{P}'_1]}z_{[\mathfrak{P}'_2]} - z_{[\mathfrak{P}'_1]}z_{[\mathfrak{P}'_2]})$ with $r_{i_1[\mathfrak{P}'_1]}r_{i_2[\mathfrak{P}'_2]} = r_{[\mathfrak{P}_1]}r_{[\mathfrak{P}_2]}$. Note that by definition of \prec , $\underline{r_{i_1[\mathfrak{P}'_1]}r_{i_2[\mathfrak{P}'_2]}}$ is indeed the leading term of this binomial.

This Gröbner basis reduction using elements $\text{Root}(T)$ ends in a binomial of the form

$$\prod_{k=1}^d r_{i_k[\mathfrak{P}_k]} - \prod_{k=1}^d r_{j_k[\mathfrak{P}_k]}$$

where $\sum_{k=1}^d i_k = \sum_{k=1}^d j_k$. At this point, we can use elements of $\text{Swap}(\rho)$ to match the columns of each monomial that correspond to the root. Note that it follows from the multiplicative property of monomial orders that we can always reduce the leading term this way by dividing by some element of $\text{Swap}(\rho)$.

Now we can check that \prec is a liftable term order on the elements of G_\prec . Any binomial in $\text{Swap}(\rho)$ has one term that is root-augmentable and one that is not. So the choice of leading term of elements of $\text{Swap}(\rho)$ does not affect the liftability property.

Let $f = r_{[\mathfrak{P}_1]}r_{[\mathfrak{P}_2]} - r_{[\Omega_1]}r_{[\Omega_2]} \in \text{Root}(T)$. Then in particular, $\psi(r_{[\mathfrak{P}_1]}r_{[\mathfrak{P}_2]}) \neq \psi(r_{[\Omega_1]}r_{[\Omega_2]})$. There are several cases.

If $[\mathfrak{P}_1]_\rho = [\mathfrak{P}_2]_\rho = [\Omega_1]_\rho = [\Omega_2]_\rho = 1$, then neither monomial in f has a root-augmentable term. So the leading term of f does not affect the liftability property.

If $[\mathfrak{P}_1]_\rho = [\Omega_1]_\rho = 1$ and $[\mathfrak{P}_2]_\rho = [\Omega_2]_\rho = 0$, without loss of generality, then $[\mathfrak{P}'_1]$ and $[\Omega'_1]$ are both not root-augmentable, and $[\mathfrak{P}'_1]$ and $[\Omega'_1]$ both *are* root-leaf traversable. If $[\mathfrak{P}_2]$ and $[\Omega_2]$ are both root-augmentable or are both not root-augmentable, then the choice of leading term of f does not affect the liftability property. Suppose that $[\mathfrak{P}_2]$ is not root-augmentable and $[\Omega_2]$ is. Then under ψ , $z_{[\mathfrak{P}'_1]}z_{[\mathfrak{P}'_2]} > z_{[\Omega'_1]}z_{[\Omega'_2]}$ since $z_{[\mathfrak{P}'_1]}z_{[\mathfrak{P}'_2]}$ has one root-augmentable term and $z_{[\Omega'_1]}z_{[\Omega'_2]}$ has two root-augmentable terms.

If $[\mathfrak{P}_1]_\rho = [\mathfrak{P}_2]_\rho = [\Omega_1]_\rho = [\Omega_2]_\rho = 0$, then the number of root-augmentable terms in either monomial in f is the same as the number of root-leaf traversable terms in each under ψ . So, since \prec is liftable, the monomial with the fewest root-augmentable terms is chosen as the leading term of f , as needed. \square

Proof of Theorem 4.2.15. If T is the tree with three leaves, then $I_T = \langle 0 \rangle$ and the result holds vacuously. Let T have $n > 3$ leaves. If T is a cluster tree, then by induction on n , we may apply Proposition 4.2.29, and I_T has a liftable Gröbner basis. By definition of a liftable term order, this Gröbner basis consists of quadratic binomials. Otherwise, I_T splits as a toric fiber product. So Proposition 4.2.19 and induction on n imply that I_T has a quadratic Gröbner basis with squarefree initial terms. \square

Corollary 4.2.30. *The CFN-MC polytope has a regular unimodular triangulation and is normal.*

Proof. By Theorem 4.2.15, the CFN-MC ideal has a quadratic Gröbner basis. Elements of this Gröbner basis correspond to elements of the kernel of a 0/1 matrix. The only quadratic binomials that could be generators of a toric ideal have the form $a^2 - bc$ or $ab - cd$ for some indeterminates a, b, c, d in the polynomial ring. However, the type $a^2 - bc$ is not possible in a toric ideal whose associated matrix is a 0/1 matrix. Since Theorem 4.2.15 shows that I_T has a quadratic Gröbner basis, and the leading term of each element of the Gröbner basis is squarefree, so the leading term ideal of I_T with respect to the liftable term order \prec is generated by squarefree monomials. Therefore, it is the Stanley-Reisner ideal of a regular unimodular triangulation of R_T [38, Theorem 8.3]. \square

CHAPTER

5

EHRHART THEORY OF THE CFN-MC MODEL

In this section, we compute the Hilbert series of the CFN-MC ideal for each rooted binary tree T . In particular, we show that this Hilbert depends only on the number of leaves of T and not on the topology of T . To accomplish this, we use Ehrhart theory of the polytopes R_T . Our approach is inspired by the work of Buczynska and Wisniewski, who proved a similar result for ideals arising from the CFN model without the molecular clock [8], and of Kubjas, who gave a combinatorial proof of the same result [28].

The Ehrhart theory of the CFN-MC polytopes is determined by permutation statistics on alternating permutations. In the following section, we provide some introductory definitions related to the combinatorics of alternating permutations.

5.1 Alternating Permutations and Order Polytopes

5.1.1 Alternating Permutations

The *zig-zag poset* \mathcal{Z}_n on ground set $\{z_1, \dots, z_n\}$ is the poset with exactly the cover relations $z_1 < z_2 > z_3 < z_4 > \dots$. That is, this partial order satisfies $z_{2i-1} < z_{2i}$ and $z_{2i} > z_{2i+1}$ for all i between 1 and $\lfloor \frac{n-1}{2} \rfloor$. The order polytope of \mathcal{Z}_n , denoted $\mathcal{O}(\mathcal{Z}_n)$ is the set of all n -tuples $(x_1, \dots, x_n) \in \mathbb{R}^n$ that satisfy $0 \leq x_i \leq 1$ for all i and $x_i \leq x_j$ whenever $z_i < z_j$ in \mathcal{Z}_n . In Section 5.2, we will show that $\mathcal{O}(\mathcal{Z}_n)$ is affinely isomorphic to the CFN-MC polytope on a caterpillar tree; this will allow us to study the Hilbert series of a CFN-MC ideal by analyzing the Ehrhart theory of $\mathcal{O}(\mathcal{Z}_n)$.

Definition 5.1.1. An *alternating permutation* on n letters is a permutation σ such that $\sigma(1) < \sigma(2) > \sigma(3) < \sigma(4) > \dots$. That is, an alternating permutation satisfies $\sigma(2i-1) < \sigma(2i)$ and $\sigma(2i) > \sigma(2i+1)$ for $1 \leq i \leq \lfloor \frac{n}{2} \rfloor$.

We denote by A_n the set of all alternating permutations. Notice that alternating permutations coincide with order-preserving bijections from $[n]$ to \mathcal{Z}_n .

The number of alternating permutations of length n is the n th *Euler zig-zag number* E_n . The sequence of Euler zig-zag numbers starting with E_0 begins $1, 1, 1, 2, 5, 16, 61, 272, \dots$. This sequence can be found in the Online Encyclopedia of Integer Sequences with identification number A000111 [33]. The exponential generating function for the Euler zig-zag numbers satisfies

$$\sum_{n \geq 0} E_n \frac{x^n}{n!} = \tan x + \sec x.$$

Furthermore, the Euler zig-zag numbers satisfy the recurrence

$$2E_{n+1} = \sum_{k=0}^n \binom{n}{k} E_k E_{n-k}$$

for $n \geq 1$ with initial values $E_0 = E_1 = 1$. A thorough background on the combinatorics of alternating permutations can be found in [36]. The following new permutation statistic on alternating permutations is central to our results.

Definition 5.1.2. Let σ be an alternating permutation. The permutation statistic $\text{swap}(\sigma)$ is the number of $i < n$ such that $\sigma^{-1}(i) < \sigma^{-1}(i+1) - 1$. Equivalently, this is the number of $i < n$ such that i is to the left of $i+1$ and swapping i and $i+1$ in σ yields another alternating permutation. The swap-set $\text{Swap}(\sigma)$ is the set of all $i < n$ for which we can perform this operation. We say that σ *swaps to* τ if τ can be obtained from σ by performing this operation a single time.

We will also make use of the following two features which can be defined for any permutation. Let $\sigma \in S_n$.

Definition 5.1.3. A *descent* of σ is an index $i \in [n-1]$ such that $\sigma(i) > \sigma(i+1)$. An *inversion* of σ is any pair (i, j) for $1 \leq i < j \leq n$ such that $\sigma^{-1}(j) < \sigma^{-1}(i)$.

When we write σ in one-line notation, a descent is a position on σ where the value of σ drops. An inversion is any pair of values in $[n]$ where the larger number appears before the smaller number in σ .

5.1.2 Order Polytopes

To every finite poset on n elements one can associate a polytope in \mathbb{R}^n by viewing the cover relations on the poset as inequalities on Euclidean space.

Definition 5.1.4. The *order polytope* $\mathcal{O}(P)$ of any poset P on ground set p_1, \dots, p_n is the set of all $\mathbf{v} \in \mathbb{R}^n$ that satisfy $0 \leq v_i \leq 1$ for all i and $v_i \leq v_j$ if $p_i < p_j$ is a cover relation in P .

Order polytopes for arbitrary posets have been the object of considerable study, and are discussed in detail in [35]. In the case of $\mathcal{O}(\mathcal{Z}_n)$, the facet defining inequalities are those of the form

$$\begin{aligned}
 -v_i &\leq 0 \text{ for } i \leq n \text{ odd} & (5.1) \\
 v_i &\leq 1 \text{ for } i \leq n \text{ even} \\
 v_i - v_{i+1} &\leq 0 \text{ for } i \leq n-1 \text{ odd, and} \\
 -v_i + v_{i+1} &\leq 0 \text{ for } i \leq n-1 \text{ even.}
 \end{aligned}$$

Note that the inequalities of the form $-v_i \leq 0$ for i even and $v_i \leq 1$ for i odd are redundant. The order polytope of \mathcal{Z}_n is also the convex hull of all $(v_1, \dots, v_n) \in \{0, 1\}^n$ that correspond to labelings of \mathcal{Z}_n that are weakly consistent with the partial order on $\{p_1, \dots, p_n\}$.

In [35], Stanley gives the following canonical unimodular triangulation of the order polytope of any poset P on ground set $\{p_1, \dots, p_n\}$. Let $\sigma : P \rightarrow [n]$ be a linear extension of P . Denote by \mathbf{e}_i the i th standard basis vector in \mathbb{R}^n . The simplex Δ^σ is the convex hull of $\mathbf{v}_0^\sigma, \dots, \mathbf{v}_n^\sigma$ where \mathbf{v}_0^σ is the all 1's vector and $\mathbf{v}_i^\sigma = \mathbf{v}_{i-1}^\sigma - \mathbf{e}_{\sigma^{-1}(i)}$. Letting σ range over all linear extensions of P yields a unimodular triangulation of $\mathcal{O}(P)$. Hence, the normalized volume of $\mathcal{O}(P)$ is the number of linear extensions of P . In particular, this means that the volume of $\mathcal{O}(\mathcal{Z}_n)$ is the Euler zig-zag number, E_n .

Example 5.1.5. Consider the case when $n = 4$. The zig-zag poset \mathcal{Z}_4 is pictured in Figure 5.1. The order polytope $\mathcal{O}(\mathcal{Z}_4)$ has facet defining inequalities

$$\begin{aligned}
 -v_1 &\leq 0 & v_2 &\leq 1 \\
 -v_3 &\leq 0 & v_4 &\leq 1 \\
 v_1 - v_2 &\leq 0 & -v_2 + v_3 &\leq 0 \\
 v_3 - v_4 &\leq 0.
 \end{aligned}$$

The vertices of $\mathcal{O}(\mathcal{Z}_4)$ are the columns of the matrix

$$\begin{bmatrix}
 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\
 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\
 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1
 \end{bmatrix}.$$

The alternating permutations on 4 elements, which correspond to linear extensions of \mathcal{Z}_4 are 1324, 1423, 2314, 2413, and 3412. Note that there are $E_4 = 5$ such alternating permutations, so the normalized volume of $\mathcal{O}(\mathcal{Z}_4)$ is 5. The simplex in the canonical triangulation of $\mathcal{O}(\mathcal{Z}_n)$ corresponding to 1423 is

$$\Delta^{1324} = \text{conv} \begin{bmatrix}
 1 & 0 & 0 & 0 & 0 \\
 1 & 1 & 1 & 1 & 0 \\
 1 & 1 & 0 & 0 & 0 \\
 1 & 1 & 1 & 0 & 0
 \end{bmatrix}.$$

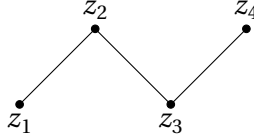


Figure 5.1 The zig-zag poset \mathcal{Z}_4

Consider this order of the simplices of the canonical triangulation:

$$\Delta^{3412}, \Delta^{2413}, \Delta^{2314}, \Delta^{1423}, \Delta^{1324}.$$

This particular ordering of the facets is one of the shelling orders that will be established and proved in the next section. The fact that this is a shelling order can be checked directly in this example, for instance:

$$\Delta^{2314} \cap (\Delta^{3412} \cup \Delta^{2413}) = \text{conv} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

which is a facet of Δ^{2314} . Since the intersection consists of a single facet, it will contribute a 1 to the coefficient of t in $h_{\partial(\mathcal{Z}_4)}^*(t) = 1 + 3t + t^2$.

5.2 Ehrhart Function of the CFN-MC Polytope

The goal of this section is to prove the following theorem.

Theorem 5.2.1. *For any rooted binary tree T with n leaves, the normalized volume of R_T is E_{n-1} , the $(n-1)$ st Euler zig-zag number.*

The proof of Theorem 5.2.1 has two parts. First, we give a unimodular affine isomorphism between the CFN-MC polytope associated to the *caterpillar tree* and the order polytope of the so-called “zig-zag poset”, which is known to have the desired normalized volume [36]. The second, and more difficult, part is to show that the volume and Ehrhart polynomial of the CFN-MC polytope are the same for any n -leaf tree by giving a bijection between the lattice points in $(mR_T \cap \mathbb{Z}^{n-1})$ and $(mR_{T'} \cap \mathbb{Z}^{n-1})$ where T and T' are related by a single tree rotation. Since any two binary trees on n leaves are connected by a sequences of rotations, this proves the theorem.

5.2.1 Caterpillar Trees

For a class of trees known as caterpillar trees, we can find a unimodular affine map between the CFN-MC polytope and the order polytope of a well-understood poset.

Definition 5.2.2. A *caterpillar tree* C_n on n leaves is the unique rooted tree topology with exactly one cherry.

Proposition 5.2.3. Let D be the $n \times n$ diagonal matrix with $D_{ii} = 1$ if i is odd and $D_{ii} = -1$ if i is even. Let \mathbf{a} be the vector in \mathbb{R}^n with $a_i = 0$ if i is odd and $a_i = 1$ if i is even. The rigid motion of \mathbb{R}^n defined by

$$\phi(\mathbf{x}) = D\mathbf{x} + \mathbf{a}$$

is a unimodular affine isomorphism from $R_{C_{n+1}}$ to $\mathcal{O}(\mathcal{Z}_n)$.

Proof. First note that $\det D = \pm 1$, so ϕ is a unimodular affine isomorphism. The image of \mathbf{x} under ϕ is coordinate-wise by

$$\phi(\mathbf{x})_i = \begin{cases} x_i & \text{if } i \text{ is odd, and} \\ 1 - x_i & \text{if } i \text{ is even.} \end{cases}$$

By Corollary, 4.2.13, the facet-defining inequalities of $R_{C_{n+1}}$ are of the form $-x_i \leq 0$ for $i \leq n$ and $x_i + x_{i+1} \leq 1$ for $i \leq n-1$. Substitution $\phi(\mathbf{x})$ into each of these equations yields exactly the inequalities in Equation (5.1), as needed. \square

Corollary 5.2.4. The Ehrhart functions of the CFN-MC polytope $R_{C_{n+1}}$ and the order polytope $\mathcal{O}(\mathcal{Z}_n)$ are equal for all n . This further implies that the normalized volume of $R_{C_{n+1}}$ is the n th Euler zig-zag number, E_n .

Proof. The Ehrhart functions $i_{R_{C_{n+1}}}(m)$ and $i_{\mathcal{O}(\mathcal{Z}_n)}(m)$ are equal because ϕ is a lattice-point preserving transformation from $mR_{C_{n+1}}$ to $m\mathcal{O}(\mathcal{Z}_n)$. The leading coefficient of the Ehrhart polynomial of a polytope is the volume of that polytope. This volume is $\frac{E_n}{n!}$ for $\mathcal{O}(\mathcal{Z}_n)$ and so, for $R_{C_{n+1}}$ as well [36]. So the normalized volume of $R_{C_{n+1}}$ is E_n . \square

5.2.2 The Ehrhart Function and Rotations

We give an explicit bijection between the lattice points in the m -th dilate of R_T and of $R_{T'}$ where $m \in \mathbb{Z}_+$ and T and T' differ by one rotation. This shows that the Ehrhart polynomials of R_T and $R_{T'}$ are the same. Since any tree can be obtained from any other tree by a finite sequence of rotations, this along with Corollary 5.2.4 proves Theorem 5.2.1.

Let b, c, e be three consecutive nodes of a tree T , where c is a descendant of b , and e is a descendant of c . Note that node e need not be an internal node of T . There is a unique *rotation* associated to the triple (b, c, e) ; namely, this move prunes c , the edge ce and the e -subtree from below b , and reattaches these on the other edge immediately below b to yield a new tree, T' . This rotation, and its effect on the internal structure of T is depicted in Figure 5.2. Note that it is also possible for b to be the root, in which case node a in this figure does not exist. A rotation splits the vertices of R_T into two natural categories: the ones that are also vertices of $R_{T'}$ and the ones that are not.

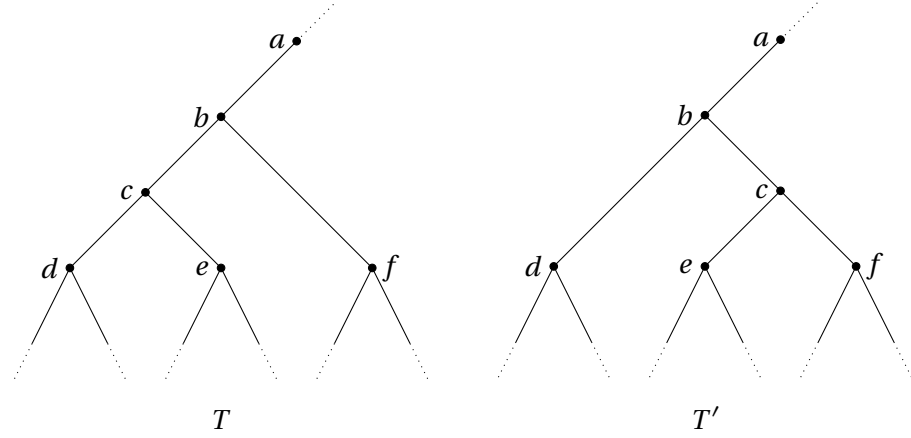


Figure 5.2 A rotation performed by pruning c and its right subtree and reattaching in the right subtree of b .

Definition 5.2.5. Let the tree T' be obtained from T by the rotation associated to (b, c, e) as pictured in Figure 5.2. A vertex of R_T is *maintaining* if it is also a vertex of $R_{T'}$. A vertex of R_T is *nonmaintaining* if it is not a vertex of $R_{T'}$.

We use the following definition to give a characterization of the maintaining and nonmaintaining vertices of R_T .

Definition 5.2.6. Let S be the top-set of a path system in T . Let x be an internal node of T . Then x is *blocked* in S if for every path from x to a leaf descended from x , there exists a $y \in S$ that lies on this path.

Note that if x is blocked in S , then for every path system \mathfrak{P} that realizes S , we cannot add another path to \mathfrak{P} with top-most node above x that passes through x .

Example 5.2.7. Let T be the tree pictured in Figure 5.3 with a path system \mathfrak{P} drawn in bold. Note that up to a swap of the leaves below node h , \mathfrak{P} is the only path system in T that realizes top-set $\{a, d, e, g\}$.

By definition, a, d, e and g are all blocked in $\{a, d, e, g\}$. Furthermore, node c is blocked in $\{a, d, e, g\}$ since any path from c to a leaf descended from c passes through either d or e , but $d, e \in \{a, d, e, g\}$. On the other hand, f is not blocked in $\{a, d, e, g\}$, since there is a path from f to leaf l that only passes through h , and $h \notin \{a, d, e, g\}$. Similarly, b is not blocked in $\{a, d, e, g\}$.

Proposition 5.2.8. Let $[\mathfrak{P}]$ be a vertex of R_T with associated top-set V . Let a, b, c, d, e and f be as in tree T in Figure 5.2.

- (i) If $b, c \notin V$, then $[\mathfrak{P}]$ is maintaining.
- (ii) If $b \in V$, then $[\mathfrak{P}]$ is maintaining if and only if d is not blocked in V .

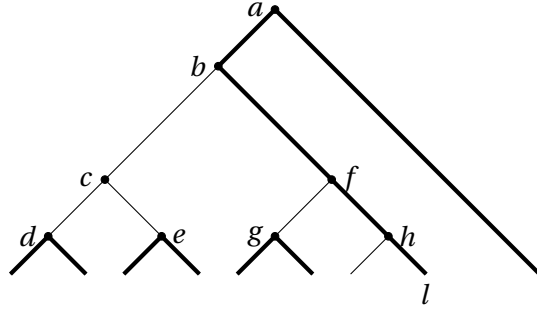


Figure 5.3 Nodes a, c, d, e and g are the nodes that are blocked in top-set $\{a, d, e, g\}$.

(iii) If $c \in V$, then $[\mathfrak{P}]$ is maintaining if and only if f is not blocked in V .

Proof. To prove (i), let $b, c \notin V$. Then since all paths in \mathfrak{P} are disjoint, there can be at most one $P \in \mathfrak{P}$ that is not contained entirely in the d -, e - or f -subtrees, or in T without the b -subtree. If no such P exists, then \mathfrak{P} is still a path system that realizes V in T' , as needed. In particular, if b is the root of T , then no such path can exist since in this case, every path is contained in the b -subtree. Now suppose that such a $P \in \mathfrak{P}$ does exist. Then it must be the case that b is the direct descendent of some node a . We can modify P to be a path P' in T' as follows.

If $ab, bc, cd \in P$, then let P' be the path in T' obtained from P by replacing edges bc and cd with edge bd , and leaving all others the same. If $ab, bc, ce \in P$, then $P' = P$ is also a path in T' , and we do not need to modify it. If $ab, bf \in P$, then let P' be the path in T' obtained by replacing edge bf in P with edges bc and cf , and leaving all others the same. Since $b, c \notin V$, these are the only cases. Then $(\mathfrak{P} - \{P\}) \cup \{P'\}$ is a path system in T' and $[(\mathfrak{P} - \{P\}) \cup \{P'\}] = [\mathfrak{P}]$. So $[\mathfrak{P}]$ is maintaining.

To prove (ii), let $b \in V$. Suppose that $[\mathfrak{P}]$ is maintaining. Then there exists a path system \mathfrak{P}' in T' such that $[\mathfrak{P}'] = [\mathfrak{P}]$. Let $P \in \mathfrak{P}'$ have top-most node b . Then $bd \in P$ and P includes a path \bar{P} from d to a leaf descended from d . Since all paths in \mathfrak{P}' are disjoint, no path in \mathfrak{P}' has its top-most node along \bar{P} . So d is not blocked in V .

Suppose that d is not blocked in V . Then there exists a path \hat{P} from d to a leaf descended from d with none of its nodes in V . Let \mathfrak{P} be the path system in T that realizes V , and let $P \in \mathfrak{P}$ have top-most node b . We may assume that \hat{P} is contained in P . P also includes a path \bar{P} from f to a leaf descended from f . Let $P' = \hat{P} \cup \bar{P} \cup \{bd, bc, cf\}$. Then $\mathfrak{P}' = (\mathfrak{P} - \{P\}) \cup \{P'\}$ is a path system in T' with $[\mathfrak{P}'] = [\mathfrak{P}]$. So $[\mathfrak{P}]$ is maintaining.

To prove (iii), let $c \in V$. Suppose that $[\mathfrak{P}]$ is maintaining. Then there exists a path system \mathfrak{P}' in T' such that $[\mathfrak{P}'] = [\mathfrak{P}]$. Let $P \in \mathfrak{P}'$ have top-most node c . Then $cf \in P$ and P includes a path \bar{P} from f to a leaf descended from f . Since all paths in \mathfrak{P}' are disjoint, no path in \mathfrak{P}' has top-most node along \bar{P} . So f is not blocked in V .

Suppose that f is not blocked in V . Let $P_1 \in \mathfrak{P}$ have top-most node c . Since f is not blocked in V , there exists a path \bar{P} from f to a leaf descended from f such that no node on \bar{P} is in V . Note that this path may be contained in some path $P_2 \in \mathfrak{P}$. If such P_2 exists, it must have top-most node above

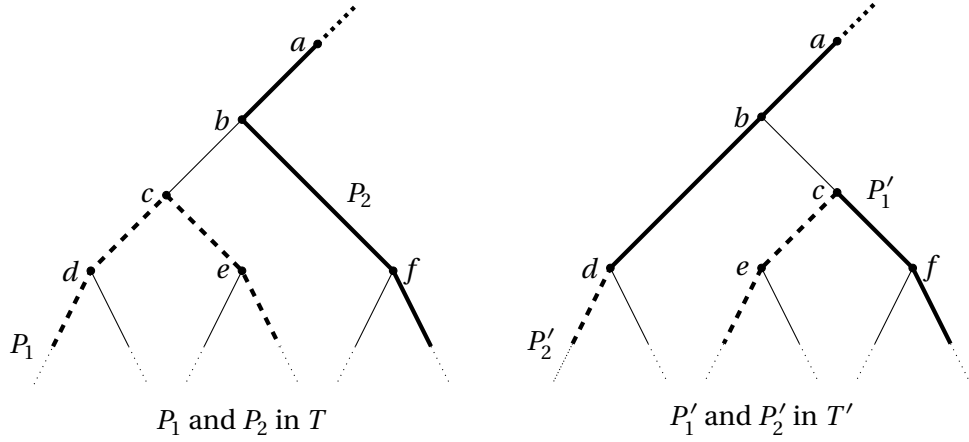


Figure 5.4 Proposition 5.2.8 (iii): In this case, $\overline{P_1}$ contains all dashed edges descended from d . \hat{P}_1 contains all dotted edges descended from e . \overline{P} contains all thick solid edges descended from f . \hat{P}_2 contains all thick solid edges above b .

b . So $bf \in P_2$ in this case. Let $\hat{P}_2 = P_2 - (\overline{P} \cup \{bf\})$. These paths are illustrated in Figure 5.4.

Furthermore, P_1 contains a path $\overline{P_1}$ from d to a leaf descended from d , and \hat{P}_1 from e to a leaf descended from e .

Let P_1' be the path in T' with top-most node c ,

$$P_1' = \hat{P}_1 \cup \overline{P} \cup \{ce, cf\}.$$

If there exists a $P_2 \in \mathfrak{P}$ that contains f , let P_2' be the path in T' ,

$$P_2' = \overline{P_1} \cup \hat{P}_2 \cup \{bd\}.$$

Then $\mathfrak{P}' = (\mathfrak{P} - \{P_1, P_2\}) \cup \{P_1', P_2'\}$, or $(\mathfrak{P} - \{P_1\}) \cup \{P_1'\}$ if no such P_2 exists, is a path system in T' $[\mathfrak{P}'] = [\mathfrak{P}]$. So \mathbf{v} is maintaining. \square

For simplicity, if the b -coordinate of $[\mathfrak{P}]$ is equal to 1 (ie. $[\mathfrak{P}]_b = 1$) and $[\mathfrak{P}]$ is nonmaintaining, we say that $[\mathfrak{P}]$ is b -nonmaintaining, and similarly for node c .

Proposition 5.2.9. *The b -nonmaintaining vertices of R_T are in bijection with the c -nonmaintaining vertices of $R_{T'}$. Similarly, the c -nonmaintaining vertices of R_T are in bijection with the b -nonmaintaining vertices of $R_{T'}$.*

Proof. Let $[\mathfrak{P}]$ be a b -nonmaintaining vertex of R_T . Then by Proposition 5.2.8, d is blocked in the top-set V of $[\mathfrak{P}]$. So the path $P \in \mathfrak{P}$ with top-most node b passes through the e -subtree of T . Let P' be the path in T' given by

$$P' = (P - \{bc, bf\}) \cup \{cf\}.$$

Then $\mathfrak{P}' = (\mathfrak{P} - \{P\}) \cup P'$ is a path system in T' that matches $[\mathfrak{P}]$ on all coordinates other than the

b - and c -coordinates, and that has b -coordinate equal to 0 and c -coordinate equal to 1. Since d is blocked in \mathfrak{P}' , $[\mathfrak{P}']$ is c -nonmaintaining in T' . (Note that the node labels do not match those of Proposition 5.2.8 since we are applying the result to the tree obtained after the rotation has been performed.) Performing the reverse operation on a c -nonmaintaining vertex $[\mathfrak{P}']$ of T' shows that this is a bijection. \square

Definition 5.2.10. If $[\mathfrak{P}]$ is nonmaintaining, let \mathfrak{P}' be the path system described in the proof of Proposition 5.2.9, so that $[\mathfrak{P}']_b = [\mathfrak{P}]_c$, $[\mathfrak{P}']_c = [\mathfrak{P}]_b$, and $[\mathfrak{P}']$ matches $[\mathfrak{P}]$ for all other nodes of T . Proposition 5.2.9 allows us to define the following involution between the vertices of R_T and $R_{T'}$:

$$\begin{aligned} \phi^{T,T'} : \text{vert}(R_T) &\rightarrow \text{vert}(R_{T'}) \\ [\mathfrak{P}] &\mapsto \begin{cases} [\mathfrak{P}], & \text{if } [\mathfrak{P}] \text{ is maintaining} \\ [\mathfrak{P}'], & \text{if } [\mathfrak{P}] \text{ is nonmaintaining.} \end{cases} \end{aligned}$$

We now turn our attention to the integer lattice points in the m^{th} dilates of R_T and $R_{T'}$ for $m \in \mathbb{Z}_+$. Let $\mathbf{v} \in \mathbb{Z}^{n-1} \cap mR_T$. Recall that by Corollary 4.2.30, R_T is normal. So, we may write $\mathbf{v} = [\mathfrak{P}_1] + \dots + [\mathfrak{P}_m]$ for some $[\mathfrak{P}_1], \dots, [\mathfrak{P}_m] \in \text{vert}(R_T)$. We call $[\mathfrak{P}_1] + \dots + [\mathfrak{P}_m]$ a *representation* of \mathbf{v} . Such a representation is *minimal* if it uses the smallest number of nonmaintaining vertices over all representations of \mathbf{v} .

For each vertex $[\mathfrak{P}_i]$ of R_T , let V_i denote the top-set associated to $[\mathfrak{P}_i]$.

Definition 5.2.11. A representation $[\mathfrak{P}_1] + \dots + [\mathfrak{P}_m] = \mathbf{v} \in mR_T \cap \mathbb{Z}^{n-1}$ is *d-compressed* if

- all of $[\mathfrak{P}_1], \dots, [\mathfrak{P}_m]$ with b -coordinate equal to 1 are maintaining, or
- for all $[\mathfrak{P}_i]$ with $b, c \notin V_i$, d is blocked in V_i .

Similarly, this representation is *f-compressed* if

- all of $[\mathfrak{P}_1], \dots, [\mathfrak{P}_m]$ with c -coordinate equal to 1 are maintaining, or
- for all $[\mathfrak{P}_i]$ with $b, c \notin V_i$, f is blocked in V_i .

If $[\mathfrak{P}_1] + \dots + [\mathfrak{P}_m]$ is both d -compressed and f -compressed, then we say that the representation is *df-compressed*.

Consider the map $\phi_m^{T,T'} : (mR_T \cap \mathbb{Z}^{n-1}) \rightarrow (mR_{T'} \cap \mathbb{Z}^{n-1})$ defined by

$$\phi_m^{T,T'}(\mathbf{v}) = \sum_{i=1}^m \phi^{T,T'}([\mathfrak{P}_i])$$

where $\sum_{i=1}^m [\mathfrak{P}_i]$ is a minimal representation of \mathbf{v} . Both the well-definedness of this map, as well as the fact that it is a bijection follow Lemmas 5.2.12 and 5.2.13 below.

Lemma 5.2.12. *If $[\mathfrak{P}_1] + \dots + [\mathfrak{P}_m]$ is a minimal representation of $\mathbf{v} \in mR_T \cap \mathbb{Z}^{n-1}$, then $[\mathfrak{P}_1] + \dots + [\mathfrak{P}_m]$ is df -compressed.*

In the proofs of Lemmas 5.2.12 and 5.2.13, we use the following notation. For all top-sets S and all nodes x of T , denote by S^x the intersection of S with the set of internal nodes of the x -subtree. For all $\mathbf{w} \in \mathbb{R}^{n-1}$, denote by \mathbf{w}^x the restriction of \mathbf{w} to the coordinates corresponding to nodes in the x -subtree. For all path systems \mathfrak{P} in T , denote by \mathfrak{P}^x the set of all paths in \mathfrak{P} that are contained in the x -subtree.

Proof of Lemma 5.2.12. We prove the contrapositive. Suppose that the representation $[\mathfrak{P}_1] + \dots + [\mathfrak{P}_m]$ is not df -compressed. Then this representation is either not d -compressed or not f -compressed. We show that $[\mathfrak{P}_1] + \dots + [\mathfrak{P}_m]$ is not minimal in both cases.

If $[\mathfrak{P}_1] + \dots + [\mathfrak{P}_m]$ is not d -compressed, then it must be the case that both conditions in the definition of a d -compressed representation fail. Since the first condition fails, there exists a $[\mathfrak{P}_i]$ with b -coordinate equal to 1 which is nonmaintaining. Without loss of generality, we may assume that $[\mathfrak{P}_1]$ is this b -nonmaintaining vertex. Since the second condition fails, there exists a $[\mathfrak{P}_i]$ with $b, c \notin V_i$ and where d is not blocked in V_i . Without loss of generality, we may also assume that $[\mathfrak{P}_2]$ is this vertex with $b, c \notin V_2$ but where d is not blocked in $[\mathfrak{P}_2]$. We claim that $\overline{V}_1 = (V_1 - (V_1^d \cup V_1^e)) \cup V_2^d \cup V_2^e$ and $\overline{V}_2 = (V_2 - (V_2^d \cup V_2^e)) \cup V_1^d \cup V_1^e$ are valid top-sets in T with associated path systems $\overline{\mathfrak{P}}_1$ and $\overline{\mathfrak{P}}_2$ respectively. We further claim that $[\overline{\mathfrak{P}}_1]$ and $[\overline{\mathfrak{P}}_2]$ are maintaining. The operation described in this proof is illustrated for an example tree T and top-sets V_1 and V_2 in Figure 5.5.

Let $P \in \mathfrak{P}_1$ with top-most node b . Let \hat{P} be the path from b to a leaf below f contained in P . Let \overline{P} be a path from d to a leaf descended from d that does not contain any nodes in V_2 ; this exists since d is not blocked in V_2 . Let $P' = \hat{P} \cup \{bc, cd\} \cup \overline{P}$. Then

$$\overline{\mathfrak{P}}_1 = (\mathfrak{P}_1 - (\{P\} \cup \mathfrak{P}_1^d \cup \mathfrak{P}_1^e)) \cup \{P'\} \cup \mathfrak{P}_2^d \cup \mathfrak{P}_2^e$$

realizes \overline{V}_1 . Also, d is not blocked in \overline{V}_1 , so $[\overline{\mathfrak{P}}_1]$ is maintaining.

If there is no path in \mathfrak{P}_2 that contains edges cd or ce , then it is clear that

$$\overline{\mathfrak{P}}_2 = (\mathfrak{P}_2 - (\mathfrak{P}_2^d \cup \mathfrak{P}_2^e)) \cup \mathfrak{P}_1^d \cup \mathfrak{P}_1^e$$

realizes \overline{V}_2 . Otherwise, suppose that $Q \in \mathfrak{P}_2$ is a path that contains cd or ce . Let \hat{Q} be the path contained in Q without the edges in the c -subtree. Let P be the path in \mathfrak{P}_1 with top-most node b , and let \tilde{P} be the path contained in P from e to a leaf descended from e ; this exists since d is blocked in V_1 . Let $Q' = \hat{Q} \cup \{ce\} \cup \tilde{P}$. Then

$$\overline{\mathfrak{P}}_2 = (\mathfrak{P}_2 - (\{Q\} \cup \mathfrak{P}_2^d \cup \mathfrak{P}_2^e)) \cup \{Q'\} \cup \mathfrak{P}_1^d \cup \mathfrak{P}_1^e$$

realizes \overline{V}_2 , as needed. Since b and $c \notin \overline{V}_2$, $[\overline{\mathfrak{P}}_2]$ is maintaining.

This operation preserves the number of times each internal node is a top-most node. So, $\mathbf{v} = [\overline{\mathfrak{P}}_1] + [\overline{\mathfrak{P}}_2] + [\mathfrak{P}_3] + \dots + [\mathfrak{P}_m]$ is a representation of \mathbf{v} using fewer nonmaintaining vertices, and $[\mathfrak{P}_1] + \dots + [\mathfrak{P}_m]$ is not minimal.

If $[\mathfrak{P}_1] + \dots + [\mathfrak{P}_m]$ is not f -compressed, we proceed by a similar argument. Without loss of

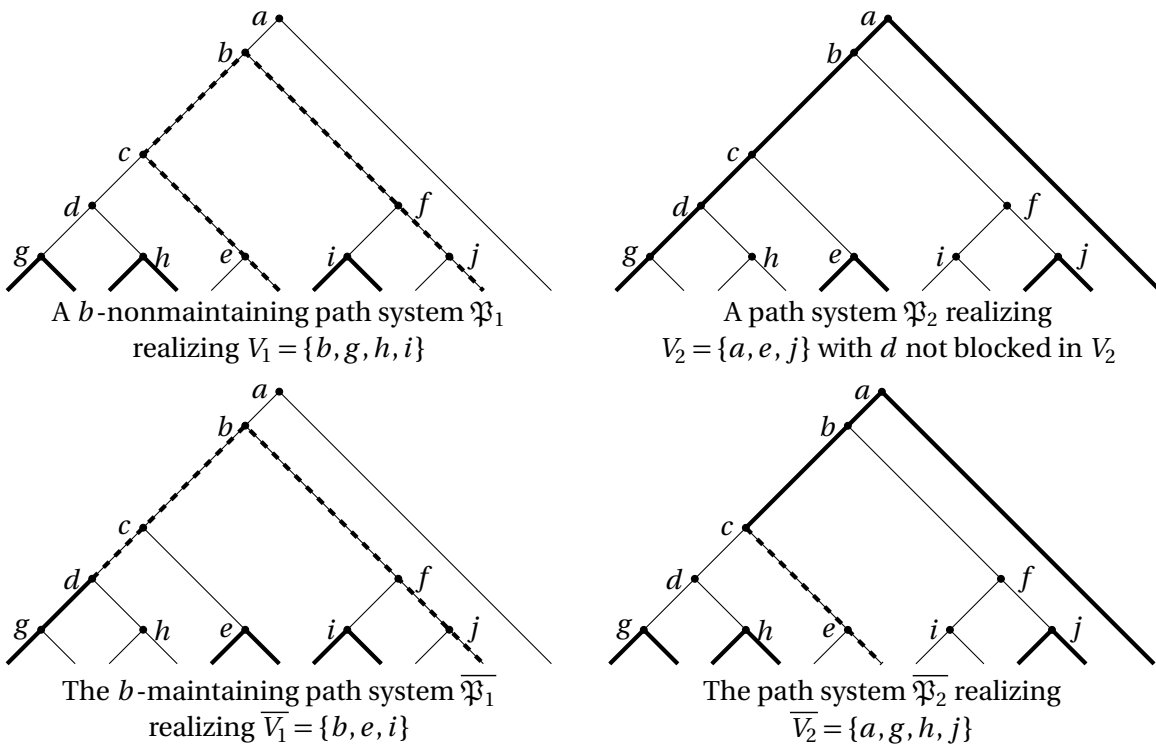


Figure 5.5 Proof of Lemma 5.2.12, The first row of trees contain path systems – one whose top-set is b -nonmaintaining, and one without b or c in its top-set, but with d not blocked in its top-set. The second row of trees are the path systems obtained by performing the operation in the proof of Lemma 5.2.12. Note that the representation given by the path systems in the second row is df -compressed.

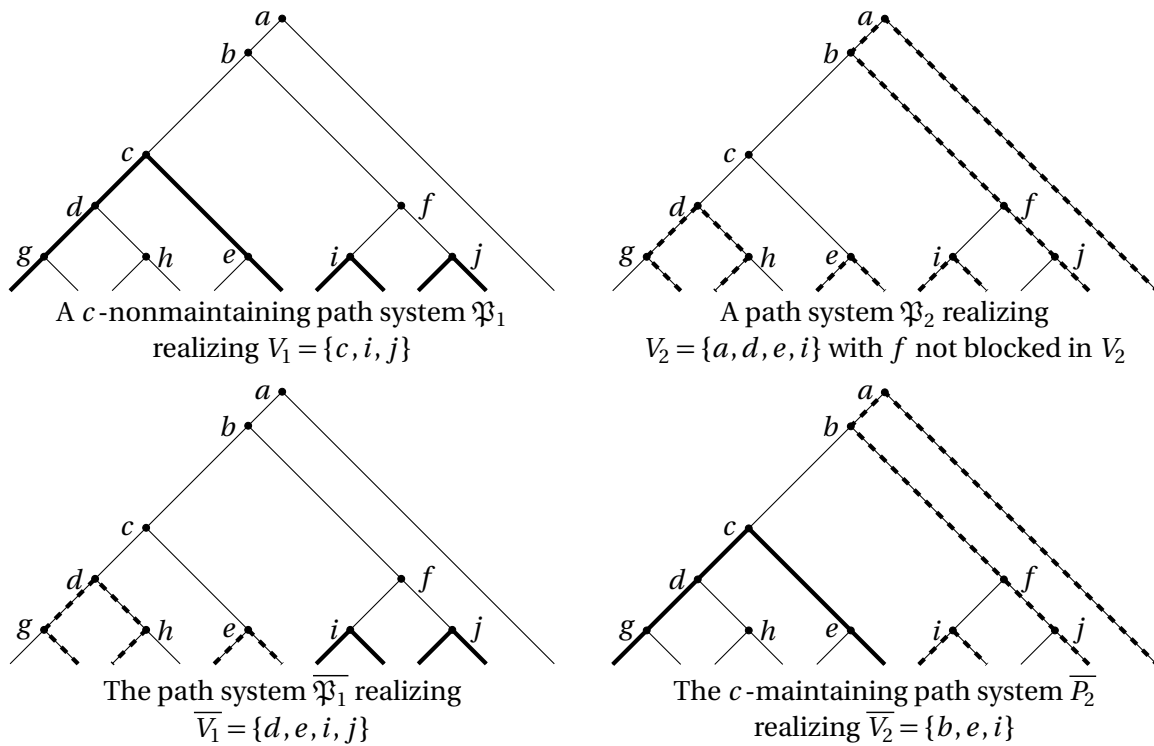


Figure 5.6 Proof of Lemma 5.2.12, The first row of trees contain path systems – one whose top-set is c -nonmaintaining, and one without b or c in its top-set, but with f not blocked in its top-set. The second row of trees are the path systems obtained by performing the operation in the proof of Lemma 5.2.12. Note that the representation given by the path systems in the second row is df -compressed.

generality, we may assume that $[\mathfrak{P}_1]$ is c -nonmaintaining and that f is not blocked in $[\mathfrak{P}_2]$. Then we claim that $\overline{V}_1 = (V_1 - (V_1^c)) \cup V_2^c$ and $\overline{V}_2 = (V_2 - (V_2^c)) \cup V_1^c$ are valid top-sets in T with associated path systems $\overline{\mathfrak{P}}_1$ and $\overline{\mathfrak{P}}_2$ respectively. Furthermore, we claim that $[\overline{\mathfrak{P}}_1]$ and $[\overline{\mathfrak{P}}_2]$ are maintaining. Figure 5.6 shows an example of a path system that is not f -compressed and the path system obtained from it by performing this operation.

Since f is not blocked in V_2 , we may assume that for all $P \in \mathfrak{P}_2$, $b, c \notin P$. This is because $b \notin V_2$, and any $P \in \mathfrak{P}_2$ with top-most node above b may pass through the f -subtree instead of the c -subtree since f is not blocked in V_2 . So, in both \mathfrak{P}_1 and \mathfrak{P}_2 , all paths that intersect the c -subtree are contained entirely within the c -subtree. So $\overline{\mathfrak{P}}_1 = (\mathfrak{P}_1 - \mathfrak{P}_1^c) \cup \mathfrak{P}_2^c$ and $\overline{\mathfrak{P}}_2 = (\mathfrak{P}_2 - \mathfrak{P}_2^c) \cup \mathfrak{P}_1^c$ are path systems that realize \overline{V}_1 and \overline{V}_2 , respectively. Since $b, c \notin \overline{V}_1$, $[\overline{\mathfrak{P}}_1]$ is maintaining. Furthermore, since f is not blocked in V_2 , and since $V_2^f = V_2^f$, f is not blocked in \overline{V}_2 and $[\overline{\mathfrak{P}}_2]$ is maintaining.

This operation preserves the number of times each internal node is used as a top-most node. So $\mathbf{v} = [\overline{\mathfrak{P}}_1] + [\overline{\mathfrak{P}}_2] + [\mathfrak{P}_3] + \dots + [\mathfrak{P}_m]$ is a representation of $[\mathfrak{P}]$ using fewer c -nonmaintaining vertices. \square

Lemma 5.2.13. *Let $\mathbf{v}, \mathbf{u} \in mR_T \cap \mathbb{Z}^{n-1}$ such that $v_b + v_c = u_b + u_c$ and $v_x = u_x$ for all $x \neq b, c$. Let $\mathbf{v} = [\mathfrak{P}_1] + \dots + [\mathfrak{P}_m]$ be a d -compressed representation of \mathbf{v} and let $\mathbf{u} = [\Omega_1] + \dots + [\Omega_m]$ be any representation of \mathbf{u} . If the multiset $\{[\Omega_1], \dots, [\Omega_m]\}$ contains fewer b -nonmaintaining or c -nonmaintaining vertices than the multiset $\{[\mathfrak{P}_1], \dots, [\mathfrak{P}_m]\}$, then $[\mathfrak{P}_1] + \dots + [\mathfrak{P}_m]$ is not a minimal representation of \mathbf{v} .*

Proof. First suppose that the set $\{[\Omega_1], \dots, [\Omega_m]\}$ contains fewer b -nonmaintaining vertices than $\{[\mathfrak{P}_1], \dots, [\mathfrak{P}_m]\}$. Then without loss of generality, let $[\mathfrak{P}]_1$ be b -nonmaintaining. For all i , let V_i denote the top-set corresponding to \mathfrak{P}_i and let U_i denote the top-set corresponding to Ω_i . Figure 5.7 depicts an example of this case and of the procedure that we describe in the following proof.

Since $[\mathfrak{P}_1] + \dots + [\mathfrak{P}_m]$ is d -compressed, for all V_i with $b, c \notin V_i$, d is blocked in V_i . Without loss of generality, let $[\mathfrak{P}_1], \dots, [\mathfrak{P}_{r'}]$ and $[\Omega_1], \dots, [\Omega_{r'}]$ be the b -nonmaintaining vertices where $r' < r$. Let $[\mathfrak{P}_{r'+1}], \dots, [\mathfrak{P}_s]$ and $[\Omega_{r'+1}], \dots, [\Omega_s]$ be the rest of the vertices with b or c coordinate equal to 1. Note that by assumption, there are the same number of vertices summed in the representations of \mathbf{u} and \mathbf{v} .

Let $\overline{V}_i = (V_i - V_i^d) \cup U_i^d$ for all i . We claim that each of the \overline{V}_i are valid top-sets, and that the collection of all corresponding $[\overline{\mathfrak{P}}_i]$ has the same number of b -nonmaintaining vertices as the $[\Omega_i]$, and the same number of c -nonmaintaining vertices as the $[\mathfrak{P}_i]$.

First, let $i \leq r'$. Then since $[\mathfrak{P}]_i$ and $[\Omega]_i$ are both b -nonmaintaining, d is blocked in both V_i and U_i . So $cd \notin \mathfrak{P}_i, \Omega_i$. Therefore, all paths in \mathfrak{P}_i and Ω_i that intersect the d -subtree are contained entirely within the d -subtree. So $\overline{\mathfrak{P}}_i = (\mathfrak{P}_i - \mathfrak{P}_i^d) \cup \Omega_i^d$ is a path system that realizes \overline{V}_i , as needed.

Next, let $r' < i \leq s$. Then $[\Omega]_i$ either is b -maintaining or has c -coordinate equal to 1. In either case, d is not blocked in U_i . So there exists a path Q from d to a leaf descended from d with no node along Q in U_i . Let $P \in \mathfrak{P}_i$ be the path with either b or c as its top-most node.

If P has b as its top-most node, then let \hat{P} be the path from b to a node below f that is contained in P . Let $P' = \hat{P} \cup \{bc, cd\} \cup Q$. Then

$$\overline{P}_i = (\mathfrak{P}_i - (\{P\} \cup \mathfrak{P}_i^d)) \cup \{P'\} \cup \Omega_i^d$$

realizes \overline{V}_i .

If P has c as its top-most node, then let \hat{P} be the path from c to a leaf below e that is contained in P . Let $P' = \hat{P} \cup \{cd\} \cup Q$. Then

$$\overline{P}_i = (\mathfrak{P}_i - (\{P\} \cup \mathfrak{P}_i^d)) \cup \{P'\} \cup \mathfrak{Q}_i^d$$

realizes \overline{V}_i .

Note that in all cases when $r' < i \leq s$, d is not blocked in \overline{V}_i . Since $r' < r$, this means that there are fewer b -nonmaintaining vertices in $\{\overline{[\mathfrak{P}]_1}, \dots, \overline{[\mathfrak{P}]_s}\}$ than in $\{[\mathfrak{P}]_1, \dots, [\mathfrak{P}]_s\}$. Furthermore, since the paths in the f -subtrees remain unchanged, this operation cannot create new c -nonmaintaining vertices.

Finally, let $i > s$. Then $b, c \notin V_i, U_i$. Since d is blocked in every V_i , all paths in \mathfrak{P}_i that intersect the d -subtree are contained entirely in the d -subtree. So $\overline{\mathfrak{P}} = (\mathfrak{P}_i - \mathfrak{P}_i^d) \cup \mathfrak{Q}_i^d$ is a path system that realizes \overline{V}_i .

Since

$$\sum_{i=1}^m [\mathfrak{P}_i]^d = \sum_{i=1}^m [\mathfrak{Q}_i]^d,$$

and since $[\overline{\mathfrak{P}}_i]^d = [\mathfrak{Q}_i]^d$ for all i , $[\overline{\mathfrak{P}}_1] + \dots + [\overline{\mathfrak{P}}_m]$ is a representation of \mathbf{v} using fewer nonmaintaining vertices than $[\mathfrak{P}_1] + \dots + [\mathfrak{P}_m]$. So $[\mathfrak{P}_1] + \dots + [\mathfrak{P}_m]$ is not minimal. An example of the operation used to obtain $[\overline{\mathfrak{P}}_1], \dots, [\overline{\mathfrak{P}}_m]$ is illustrated in Figure 5.7.

Now suppose that the set $\{[\mathfrak{Q}]_1, \dots, [\mathfrak{Q}]_m\}$ contains fewer c -nonmaintaining vertices than the set $\{[\mathfrak{P}]_1, \dots, [\mathfrak{P}]_m\}$. Figure 5.8 depicts an example of this case and of the procedure that we describe in the following proof. Without loss of generality, let $[\mathfrak{P}]_1$ be c -nonmaintaining. Since $[\mathfrak{P}]_1 + \dots + [\mathfrak{P}]_m$ is f -compressed, for all V_i with $b, c \notin V_i$, f is blocked in V_i . Without loss of generality, let $[\mathfrak{P}]_1, \dots, [\mathfrak{P}]_{r'}$ and $[\mathfrak{Q}]_1, \dots, [\mathfrak{Q}]_{r'}$ be the c -nonmaintaining vertices where $r' < r$. Let $[\mathfrak{P}]_{r'+1}, \dots, [\mathfrak{P}]_s$ and $[\mathfrak{Q}]_{r'+1}, \dots, [\mathfrak{Q}]_s$ be the rest of the vertices with b or c coordinate equal to 1. Note that by assumption, there are the same number of vertices summed in the representations of \mathbf{u} and \mathbf{v} .

Let $\overline{V}_i = (V_i - V_i^f) \cup U_i^f$ for all i . We claim that each of the \overline{V}_i are valid top-sets, and that the collection of all \overline{V}_i has the same number of c -nonmaintaining vertices as the U_i , and the same number of b -nonmaintaining vertices as the V_i .

First, let $i \leq r'$. Then since $[\mathfrak{P}]_i$ and $[\mathfrak{Q}]_i$ are both c -nonmaintaining, f is blocked in both V_i and U_i . Therefore, all paths in \mathfrak{P}_i and \mathfrak{Q}_i that intersect the f -subtree are contained entirely within the f -subtree. So $\overline{\mathfrak{P}}_i = (\mathfrak{P}_i - \mathfrak{P}_i^f) \cup \mathfrak{Q}_i^f$ is a path system that realizes \overline{V}_i , as needed.

Next, let $r' < i \leq s$. Then $[\mathfrak{Q}]_i$ is either c -maintaining or has b -coordinate equal to 1. In either case, f is not blocked in U_i . So there exists a path Q from f to a leaf descended from f with no node along Q in U_i .

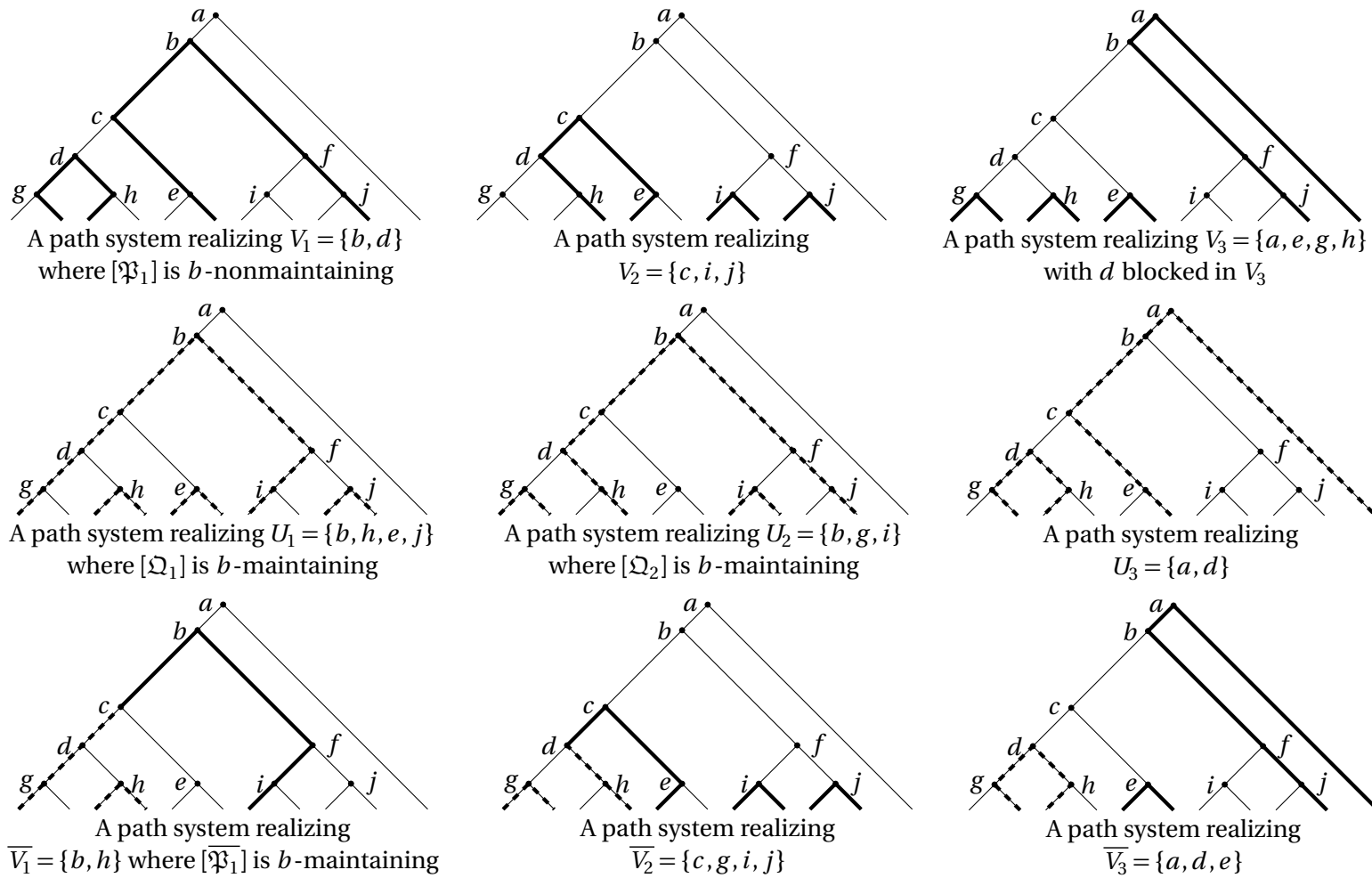


Figure 5.7 Proof of Lemma 5.2.13. This figure illustrates the case where $\{[\Omega_1], \dots, [\Omega_m]\}$ contains fewer b -nonmaintaining vertices than $\{[\mathfrak{P}_1], \dots, [\mathfrak{P}_m]\}$. The first row of trees are path systems that realize $\mathbf{v} = [\mathfrak{P}_1] + [\mathfrak{P}_2] + [\mathfrak{P}_3] \in 3R_T$. The second row of trees are path systems that realize the vector $\mathbf{u} = [\Omega_1] + [\Omega_2] + [\Omega_3] \in 3R_T$ that satisfies the assumptions of the lemma. The third row of trees are a new set of path systems that realize \mathbf{v} using fewer b -nonmaintaining vertices, which we obtained by applying the procedure discussed in the proof of the lemma.

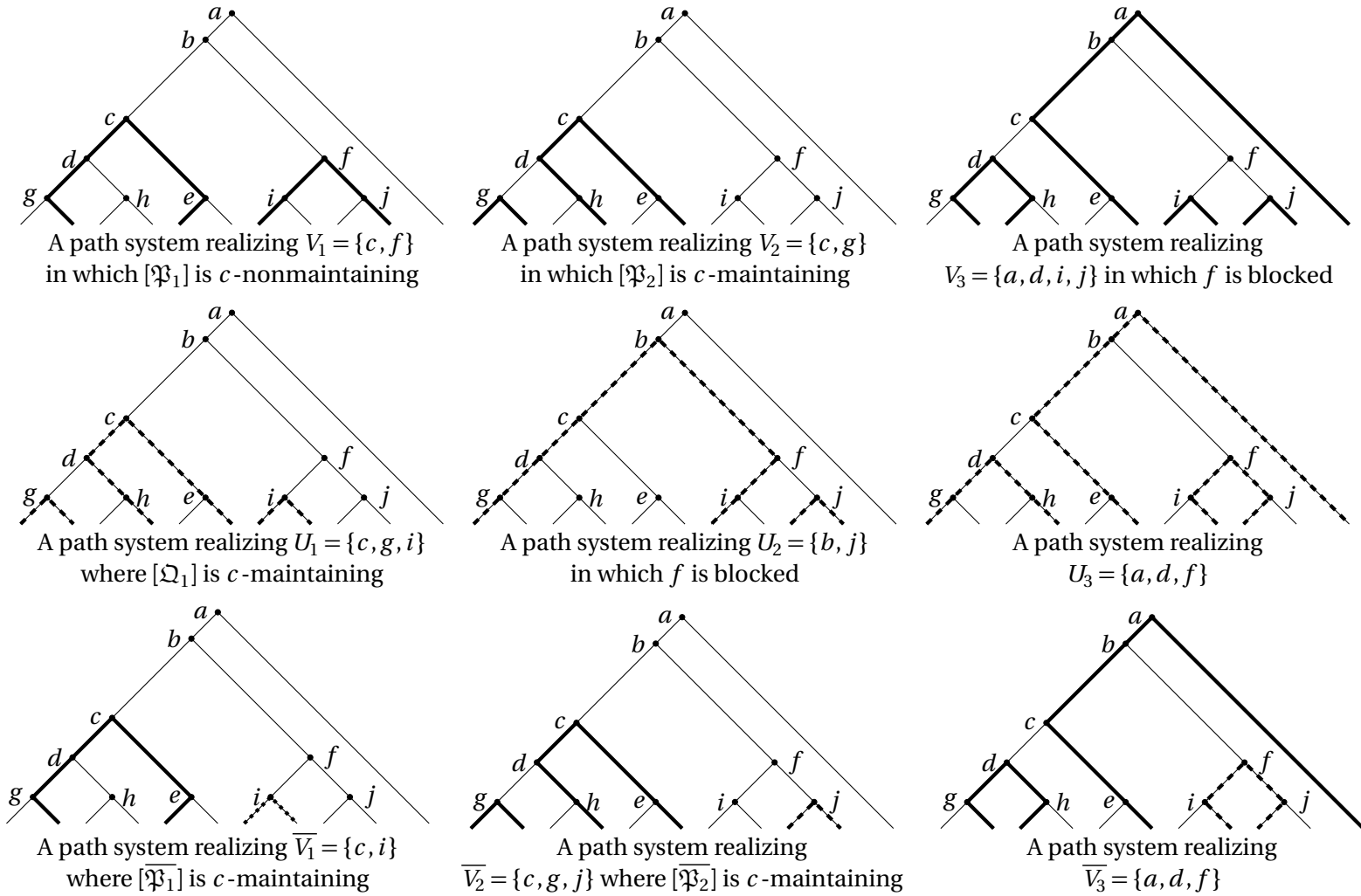


Figure 5.8 Proof of Lemma 5.2.13. $\{[\mathfrak{Q}_1], \dots, [\mathfrak{Q}_m]\}$ contains fewer c -nonmaintaining vertices than $\{[\mathfrak{P}_1], \dots, [\mathfrak{P}_m]\}$. The first row of trees are path systems that realize $\mathbf{v} = [\mathfrak{P}_1] + [\mathfrak{P}_2] + [\mathfrak{P}_3] \in 3R_T$. The second row of trees are path systems that realize $\mathbf{u} = [\mathfrak{Q}_1] + [\mathfrak{Q}_2] + [\mathfrak{Q}_3] \in 3R_T$, which satisfies the assumptions of the lemma. The third row of trees are a new set of path systems that realize \mathbf{v} using fewer c -nonmaintaining vertices, which we obtained by applying the procedure discussed in the proof of the lemma.

Consider the case when $b \in V_i$. Let $P \in \mathfrak{P}_i$ with b as its top-most node, and let \hat{P} be the path from b to a leaf below c that is contained in P . Let $P' = \hat{P} \cup \{bf\} \cup Q$. Then

$$\overline{\mathfrak{P}}_i = (\mathfrak{P}_i - (\{P\} \cup \mathfrak{P}_i^f)) \cup \{P'\} \cup \Omega_i^f$$

realizes \overline{V}_i .

Now suppose that $c \in V_i$. If there does not exist $P \in \mathfrak{P}_i$ with a node above b as its top-most node that passes through the f -subtree, then all paths in \mathfrak{P}_i that intersect the f -subtree are contained in the f -subtree. So $\overline{\mathfrak{P}}_i = (\mathfrak{P}_i - \mathfrak{P}_i^d) \cup \Omega_i^d$ is a path system that realizes \overline{V}_i .

If there does exist $P \in \mathfrak{P}_i$ with top-most node above b that passes through the f -subtree, let \hat{P} denote P without the portion of P that lies in the f -subtree. Let $P' = \hat{P} \cup Q$. Then

$$\overline{\mathfrak{P}}_i = (\mathfrak{P}_i - (\{P\} \cup \mathfrak{P}_i^f)) \cup \{P'\} \cup \Omega_i^f$$

is a path system that realizes \overline{V}_i .

Note that in all cases when $r' < i \leq s$, f is not blocked in \overline{V}_i . Since $r' < r$, this means that there are fewer c -nonmaintaining vertices in $\{[\overline{\mathfrak{P}}_1], \dots, [\overline{\mathfrak{P}}_s]\}$ than in $\{[\mathfrak{P}_1], \dots, [\mathfrak{P}_s]\}$. Furthermore, since the paths in the d -subtrees remain unchanged, this operation cannot create new b -nonmaintaining vertices.

Finally, let $i > s$. Then $b, c \notin V_i, U_i$. Since f is blocked in every V_i , all paths in \mathfrak{P}_i that intersect the f -subtree are contained entirely in the f -subtree. So $\overline{\mathfrak{P}}_i = (\mathfrak{P}_i - \mathfrak{P}_i^f) \cup \Omega_i^f$ is a path system that realizes \overline{V}_i .

Since

$$\sum_{i=1}^m [\mathfrak{P}_i]^f = \sum_{i=1}^m [\Omega_i]^f,$$

and since $[\overline{\mathfrak{P}}_i]^f = [\Omega_i]^f$ for all i , $[\overline{\mathfrak{P}}_1] + \dots + [\overline{\mathfrak{P}}_m]$ is a representation of \mathbf{v} using fewer nonmaintaining vertices than $[\mathfrak{P}_1] + \dots + [\mathfrak{P}_m]$. So $[\mathfrak{P}_1] + \dots + [\mathfrak{P}_m]$ is not minimal. An example of the operation used to obtain $[\overline{\mathfrak{P}}_1], \dots, [\overline{\mathfrak{P}}_m]$ is illustrated in Figure 5.8. \square

Corollary 5.2.14. *The map $\phi_m^{T, T'}$ is well-defined.*

Proof. It suffices to show that all minimal representations of $v \in mR_T \cap \mathbb{Z}^{n-1}$ have the same number of b - and c -nonmaintaining vertices. This follows from Lemma 5.2.13. \square

Corollary 5.2.15. *The map $\phi_m^{T, T'}$ is a bijection.*

Proof. It suffices to show that $\phi_m^{T', T}$ is the inverse map of $\phi_m^{T, T'}$. Suppose that it is not. Then there exists some $\mathbf{v} \in mR_T \cap \mathbb{Z}^{n-1}$ such that $\mathbf{v} = [\mathfrak{P}_1] + \dots + [\mathfrak{P}_m]$ is a minimal representation, but $\phi^{T, T'}([\mathfrak{P}_1]) + \dots + \phi^{T, T'}([\mathfrak{P}_m]) = \phi_m^{T, T'}(\mathbf{v})$ is not a minimal representation of $\phi_m^{T, T'}(\mathbf{v})$.

Let $[\Omega_1] + \dots + [\Omega_m] = \phi_m^{T, T'}(\mathbf{v})$ be minimal. Consider the image $\phi_m^{T', T}(\phi_m^{T, T'}(\mathbf{v})) = \phi^{T', T}([\Omega_1]) + \dots + \phi^{T', T}([\Omega_m])$. The set $\{\phi^{T', T}([\Omega_1]), \dots, \phi^{T', T}([\Omega_m])\}$ contains fewer nonmaintaining vertices than $\{[\mathfrak{P}_1], \dots, [\mathfrak{P}_m]\}$ because $\phi^{T', T}$ maps maintaining vertices to maintaining vertices by the proof of

Proposition 5.2.9. The set $\{\phi^{T',T}([\Omega_1]), \dots, \phi^{T',T}([\Omega_m])\}$ also satisfies all of the assumptions of Lemma 5.2.13. So, $[\mathfrak{P}_1] + \dots + [\mathfrak{P}_m]$ was not actually a minimal representation of \mathbf{v} and we have reached a contradiction. \square

Theorem 5.2.16. *For all rooted binary trees T with n leaves, the Hilbert series of I_T is equal to the Hilbert series of I_{C_n} .*

Proof. Every rooted binary tree can be obtained from the caterpillar tree by a finite sequence of rotations. So, it follows from Corollary 5.2.15 that the number of lattice points in the m^{th} dilates of R_T is equal to that of R_{C_n} for all trees T with n leaves. So, the Ehrhart polynomials and hence, the Ehrhart series of R_T and R_{C_n} are equal. The Ehrhart series of R_T is equal to the Hilbert series of I_T . \square

Proof of Theorem 5.2.1. The leading coefficient of the Ehrhart polynomial of a polytope is the (un-normalized) volume of the polytope. We have shown the equality of the Ehrhart polynomials of R_{C_n} and R_T for any n -leaf tree T . So, R_{C_n} and R_T have the same normalized volumes. This is the $(n-1)$ st Euler zig-zag number by Corollary 5.2.4. \square

5.3 The h^* -Polynomial of the Order Polytope of the Zig-Zag Poset

Using Proposition 5.2.3 and Theorem 5.2.16, we can now study the Ehrhart theory of $\mathcal{O}(\mathcal{Z}_n)$ in order to compute the Hilbert series of each CFN-MC ideal on a rooted binary tree with $n+1$ leaves. The goal of this section is to prove the following theorem relating the h^* -polynomial of $\mathcal{O}(\mathcal{Z}_n)$ and the swap statistic.

Theorem 5.3.1. *The numerator of the Ehrhart series of $\mathcal{O}(\mathcal{Z}_n)$ is*

$$h_{\mathcal{O}(\mathcal{Z}_n)}^*(t) = \sum_{\sigma \in A_n} t^{\text{swap}(\sigma)}.$$

Remark 5.3.2. Alternate formulas for the h^* -polynomial of the order polytope of a poset P exist, as described in [37, Chapter 13.3]. However, many of these formulas refer to the *Jordan-Hölder set* of P and in particular, descents in the permutations in this set, which we will discuss in more detail in Section 5.4. In the case of \mathcal{Z}_n , the elements of this Jordan-Hölder set are not alternating or inverse alternating permutations because they arise from linear extensions with respect to a natural labeling of \mathcal{Z}_n . The elements of the Jordan-Hölder set do not have as nice of a combinatorial description as the alternating permutations, and there is not an obvious bijection between swaps in alternating permutations and descents in elements of the Jordan-Hölder set.

In this section we describe a family of shelling orders on the simplices of the canonical triangulation of $\mathcal{O}(\mathcal{Z}_n)$. Let σ be an alternating permutation. We will denote by $\text{vert}(\sigma)$ the set of all vertices of the simplex Δ^σ . Note that this is the set of all 0/1 vectors \mathbf{v} of length n that have $v_i \leq v_j$ whenever $\sigma(i) < \sigma(j)$.

Proposition 5.3.3. *The simplices Δ^σ and Δ^τ are joined along a facet if and only if σ swaps to τ or τ swaps to σ .*

Proof. Simplices Δ^σ and Δ^τ are joined along a facet if and only if $\text{vert}(\sigma)$ and $\text{vert}(\tau)$ differ by a single element. Since every simplex in the canonical triangulation of $\mathcal{O}(\mathcal{X}_n)$ has exactly one vertex with the sum of its components equal to i for $0 \leq i \leq n$ and the all 0's and all 1's vector are in every simplex in this triangulation, this occurs if and only if there exists an i with $1 \leq i \leq n-1$ such that $\text{vert}(\sigma) - \{\mathbf{v}_i^\sigma\} = \text{vert}(\tau) - \{\mathbf{v}_i^\tau\}$. By definition of each \mathbf{v}_j^σ and \mathbf{v}_j^τ , this occurs if and only if $\sigma^{-1}(j) = \tau^{-1}(j)$ for all $j \neq i, i+1$ and $\mathbf{e}_{\sigma^{-1}(i)} + \mathbf{e}_{\sigma^{-1}(i+1)} = \mathbf{e}_{\tau^{-1}(i)} + \mathbf{e}_{\tau^{-1}(i+1)}$. This is true if and only if swapping the positions of i and $i+1$ in σ yields τ , as needed. \square

Denote by $\text{inv}(\sigma)$ the number of inversions of a permutation σ ; that is, $\text{inv}(\sigma)$ is the number of pairs $i < j$ such that $\sigma(i) > \sigma(j)$. We similarly define a *non-inversion* to be a pair $i < j$ with $\sigma(i) < \sigma(j)$. We call an inversion or non-inversion (i, j) *relevant* if $i < j-1$; that is, if it is not required by the structure of an alternating permutation. Note that performing a swap on an alternating permutation always increases its inversion number by exactly one.

The following lemma relates relevant non-inversions to swaps in between them.

Lemma 5.3.4. *Let σ be an alternating permutation. Let $a, b \in [n]$ such that $(\sigma^{-1}(a), \sigma^{-1}(b))$ is a relevant non-inversion of σ . Then there exists a k with $a \leq k < b$ such that k is a swap of σ .*

Proof. We proceed by induction on $b-a$. If $b-a=1$, then since (i, j) is a relevant non-inversion, a is a swap in σ .

Let $b-a > 1$. Consider the position of $a+1$ in σ . There are three cases. If $\sigma^{-1}(a+1) < \sigma^{-1}(b)-1$, then $(\sigma^{-1}(a+1), \sigma^{-1}(b))$ is a relevant non-inversion, and we are done by induction. If $\sigma^{-1}(a+1) > \sigma^{-1}(b)$, then a is a swap in σ . If $\sigma^{-1}(a+1) = \sigma^{-1}(b)-1$, then note that $\sigma^{-1}(a) < \sigma^{-1}(a+1)-1$ since otherwise, $a, a+1, b$ would be an adjacent increasing sequence in σ , which would contradict that σ is alternating. So a is a swap in σ , as needed. \square

Theorem 5.3.1 follows as a corollary of Theorem 1.2.8, Proposition 5.3.3 and the following theorem.

Theorem 5.3.5. *Let $\sigma_1, \dots, \sigma_{E_n}$ be an order on the alternating permutations such that if $i < j$ then $\text{inv}(\sigma_i) \geq \text{inv}(\sigma_j)$. Then the order $\Delta^{\sigma_1}, \dots, \Delta^{\sigma_{E_n}}$ on the simplices of the canonical triangulation of $\mathcal{O}(\mathcal{X}_n)$ is a shelling order.*

Note that since performing a swap increases inversion number by exactly one, the condition of Theorem 5.3.5 implies that if σ_j swaps to σ_i , then $i < j$. For any alternating permutation σ , define the *exclusion set* of σ , $\text{excl}(\sigma)$ to be the set of all $\mathbf{v}_k^\sigma \in \text{vert}(\sigma)$ such that k is a swap in σ . In other words,

$$\text{excl}(\sigma) = \{\mathbf{v} \mid \mathbf{v} \in \Delta^\sigma - \Delta^\tau \text{ for some } \tau \text{ such that } \sigma \text{ swaps to } \tau\}.$$

In the proof of Theorem 5.3.5, we will show that Proposition 5.3.3 implies that in order to prove Theorem 5.3.5, it suffices to check that if $\text{inv}(\sigma) \leq \text{inv}(\tau)$, then $\text{excl}(\sigma) \not\subset \text{vert}(\tau)$. This fact follows from the next two propositions.

Proposition 5.3.6. *An alternating permutation σ maximizes inversion number over all alternating permutations τ with $\text{excl}(\sigma) \subset \text{vert}(\tau)$.*

Proof. Consider a vertex $\mathbf{v}_k^\sigma \in \text{vert}(\sigma)$. Note that we may read all of the non-inversions (i, j) with $\sigma(i) \leq k < \sigma(j)$ from \mathbf{v}_k^σ since these correspond to pairs of positions in \mathbf{v}_k^σ with a 0 in the first position and a 1 in the second. That is to say, we have $\mathbf{v}_k^\sigma(i) = 0$, $\mathbf{v}_k^\sigma(j) = 1$, and $i < j$.

We claim that every relevant non-inversion of σ can be read from an element of $\text{excl}(\sigma)$ in this way. By Lemma 5.3.4, there exists a swap k in σ with $\sigma(i) \leq k < \sigma(j)$, and the relevant non-inversion (i, j) can be read from \mathbf{v}_k^σ in the manner described above.

Therefore, all relevant non-inversions in σ can be found as a non-adjacent 0–1 pair in a vertex in $\text{excl}(\sigma)$. In particular, we can count the number of relevant non-inversions in σ from the vertices in $\text{excl}(\sigma)$. Furthermore, if $\text{excl}(\sigma) \subset \text{vert}(\tau)$, then all non-inversions in σ must also be non-inversions in τ , though τ can contain more non-inversions as well. So σ minimizes the number of non-inversions, and therefore maximizes the number of inversions, over all τ with $\text{excl}(\sigma) \subset \text{vert}(\tau)$. \square

Proposition 5.3.7. *Let $S \subset \text{vert}(\mathcal{O}(\mathcal{X}_n))$ be contained in $\text{vert}(\sigma)$ for some alternating σ . Then there exists a unique alternating $\hat{\sigma}$ that maximizes inversion number over all alternating permutations whose vertex set contains S .*

Proof. Let $S = \{\mathbf{s}_0, \mathbf{s}_1, \dots, \mathbf{s}_r\}$ ordered by decreasing coordinate sum. We can assume that S contains both the all zeroes and all ones vectors since those vectors belong to the simplex Δ^σ for any alternating permutation σ . Since $S \subset \text{vert}(\sigma)$ for some alternating σ , if $\mathbf{s}_i(j) = 0$, then $\mathbf{s}_k(j) = 0$ for all $k > i$. For $i = 1, \dots, r$, let m_i be the number of positions in \mathbf{s}_i that are equal to zero, and let $n_i = m_i - m_{i-1}$ (with $n_1 = m_1$).

Let τ be any alternating permutation such that $S \subseteq \text{vert}(\tau)$. The 0-pattern of each \mathbf{s}_i partitions the entries of all τ with $S \subset \text{vert}(\tau)$ as follows: For $1 \leq k \leq r$, the n_k positions j such that $\mathbf{s}_k(j) = 0$ and $\mathbf{s}_{k-1}(j) = 1$ are the positions of τ such that $\tau(j) \in \{m_{k-1} + 1, \dots, m_k\}$.

The positions of inversions and non-inversions across these groups are fixed for all τ with $S \subset \text{vert}(\tau)$. We can build an alternating permutation $\hat{\sigma}$ that maximizes the inversions within each group as follows. For $1 \leq k \leq r$, let $j_1^k, \dots, j_{n_k}^k$ be the positions of $\hat{\sigma}$ that must take values in $\{m_{k-1} + 1, \dots, m_k\}$, as described above. We place these values in reverse; i.e. map j_l^k to $m_k - l + 1$. The permutation obtained in this way need not be alternating, so we switch adjacent positions that need to contain non-descents in order to make the permutation alternating. Note that we never need to make such a switch between groups, since the partition given by S respects the structure of an alternating permutation.

This permutation is unique because within the k th group, arranging the values in this way is equivalent to finding the permutation on n_k elements with some fixed non-descent positions that maximizes inversion number. To obtain this permutation, we begin with the permutation $m_k \ m_k - 1 \dots m_{k-1} + 1$ and switch all the positions that must be non-descents. The alternating structure of the original permutation implies that none of these non-descent positions can be adjacent, so these transpositions commute and give a unique permutation. \square

Example 5.3.8. Let $n = 7$ and let

$$S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

We will construct $\hat{\sigma}$, the alternating permutation that maximizes inversion number overall alternating permutations whose vertex set contains S . The second and third vertices in S are the only one that gives information about the position of each character; we will denote them \mathbf{w}_1 and \mathbf{w}_2 , respectively. Since \mathbf{w}_1 has 0's in exactly the first, third and seventh positions, we know that 1, 2 and 3 are in these positions. We insert them into these positions in decreasing order, so that $\hat{\sigma}$ has the form

$$3 _ 2 _ _ _ 1.$$

The zeros added in \mathbf{w}_2 are in the fourth, fifth and sixth positions. Placing them in decreasing order yields the permutation

$$3 _ 2 6 5 4 1.$$

However, this permutation cannot be alternating, since there must be an ascent from position 5 to position 6. To create this ascent, we switch the entries in these positions, yielding a permutation of the form

$$3 _ 2 6 4 5 1.$$

Finally, the only character missing is 7, which must go in the remaining space. This gives the permutation

$$\hat{\sigma} = 3 7 2 6 4 5 1.$$

Proof of Theorem 5.3.5. First, we claim that it suffices to show that for any alternating permutations σ and τ , if $\text{inv}(\tau) \geq \text{inv}(\sigma)$ then $\text{excl}(\sigma) \not\subset \text{vert}(\tau)$. Indeed, for any ρ with $\text{inv}(\rho) \geq \text{inv}(\sigma)$, by Proposition 5.3.3 we have that $\Delta_\sigma \cap \Delta_\rho$ is a facet of Δ_σ if and only if σ swaps to ρ . This is the case if and only if $\Delta_\sigma \cap \Delta_\rho = \Delta_\sigma \setminus \{\mathbf{v}_i\}$ for some $\mathbf{v}_i \in \text{excl}(\sigma)$. So if $\Delta_\sigma \cap \Delta_\tau \not\subset \Delta_\sigma \cap \Delta_\rho$ for any ρ such that $\text{inv}(\rho) \geq \text{inv}(\sigma)$ with $\Delta_\sigma \cap \Delta_\rho$ a facet of Δ_σ , then we must have $\text{excl}(\sigma) \subset \text{vert}(\tau)$. The contrapositive of this statement shows that if $\text{excl}(\sigma) \not\subset \text{vert}(\tau)$, then the given order on the facets of the triangulation is a shelling.

If $\text{inv}(\tau) > \text{inv}(\sigma)$, then since σ maximizes inversion number over all alternating permutations that contain the exclusion set of σ by Proposition 5.3.6, $\text{excl}(\sigma) \not\subset \text{vert}(\tau)$. Furthermore, Proposition 5.3.7 implies that if $\text{inv}(\tau) = \text{inv}(\sigma)$, then $\text{excl}(\sigma) \not\subset \text{vert}(\tau)$ because σ is the *unique* permutation that maximizes inversion number of all alternating permutation that contain its exclusion set. \square

Proof of Theorem 5.3.1. Let $\Delta^{\sigma_1}, \dots, \Delta^{\sigma_{E_n}}$ be a shelling order as described in Theorem 5.3.5. Then

by Proposition 5.3.3, each Δ^{σ_i} is added in the shelling along exactly $\text{swap}(\sigma_i)$ facets. Therefore, by Theorem 1.2.8,

$$h_{\mathcal{O}(\mathcal{Z}_n)}^*(t) = \sum_{\sigma \in A_n} t^{\text{swap}(\sigma)},$$

as needed. □

Corollary 5.3.9. *Let T be a rooted binary tree on n leaves. Then the Hilbert series of I_T has numerator,*

$$\sum_{\sigma \in A_{n-1}} t^{\text{swap}(\sigma)}$$

Proof. By Theorem 5.2.16, the Hilbert series of I_T is equal to that of I_{C_n} . By Proposition 5.2.3, the Hilbert series of I_{C_n} is equal to the Ehrhart series of $\mathcal{O}(\mathcal{Z}_{n-1})$. So the corollary follows directly from Theorem 5.3.1. □

We conclude this section by remarking that not all of the shellings described in Theorem 5.3.5 can be obtained from EL- or CL-labelings of the lattice of order ideals of \mathcal{Z}_n . Denote by $J(\mathcal{Z}_n)$ the distributive lattice of order ideals of \mathcal{Z}_n ordered by inclusion. Saturated chains in $J(\mathcal{Z}_n)$ are in bijection with elements of A_n via the map that sends an alternating permutation σ to the chain of order ideals,

$$I_0 \subsetneq I_1 \subsetneq I_2 \subsetneq \cdots \subsetneq I_n$$

where $I_j = \{\sigma^{-1}(1), \dots, \sigma^{-1}(j)\}$ [37, Chapter 3.5].

Definition 5.3.10. Let P be a graded bounded poset and let $E(P)$ be the set of cover relations of P . An *EL-labeling* of P is a labeling λ of $E(P)$ with integers such that

- each closed interval $[a, b]$ of P has a unique λ -increasing saturated chain, and
- this λ -increasing chain lexicographically precedes all other saturated chains from a to b .

A poset that has an EL-labeling is called *EL-shellable*.

For more details on poset shellability, we refer the reader to [43]. If P is EL-shellable with EL-labeling λ , then lexicographic order on the saturated chains of P with respect to λ gives a shelling of the order complex of P [6]. In the case of $J(\mathcal{Z}_n)$, its order complex is isomorphic to the canonical triangulation of the order polytope $\mathcal{O}(\mathcal{Z}_n)$ via the bijection described above. So finding EL-labelings of $J(\mathcal{Z}_n)$ is one way to construct shellings of the canonical triangulations of $\mathcal{O}(\mathcal{Z}_n)$. However, not all of the shellings described in Theorem 5.3.5 can be obtained in this way.

Proposition 5.3.11. *There exist shelling orders on the canonical triangulation of $\mathcal{O}(\mathcal{Z}_n)$ given by the conditions of Theorem 5.3.5 that cannot be obtained from EL-labelings of $J(\mathcal{Z}_n)$.*

Proof. For the sake of readability, we discuss these shellings on the level of alternating permutations. The “position of σ in a shelling order” is taken to mean the position of Δ^σ in that shelling order

for the canonical triangulation of $\mathcal{O}(\mathcal{Z}_n)$. All of the shellings described in Theorem 5.3.5 begin with the unique alternating permutation $\bar{\sigma}$ that maximizes inversion number over all alternating permutations; this permutation exists by Proposition 5.3.7. Note that $\bar{\sigma}^{-1}(1)$ is $n-1$ or n , depending upon the parity of n .

We address the case where $n \geq 5$ is odd. Then $\bar{\sigma}$ is of the form

$$\bar{\sigma} = (n-1) n (n-3) (n-2) \dots 4 5 2 3 1.$$

Let λ be an EL-labeling of the cover relations of $J(\mathcal{Z}_n)$ that induces a shelling with $\bar{\sigma}$ is its first element. (Note that if no such EL-labeling exists, the proposition holds trivially.)

The sets $\emptyset, \{p_n\}, \{p_{n-2}\}$ and $\{p_{n-2}, p_n\}$ are the order ideals of $J(\mathcal{Z}_n)$ that comprise the interval $[\emptyset, \{p_{n-2}, p_n\}]$. The chain corresponding to $\bar{\sigma}$ begins with $\emptyset \lessdot \{p_n\} \lessdot \{p_{n-2}, p_n\}$. So this must be the unique λ -increasing chain in the interval $[\emptyset, \{p_{n-2}, p_n\}]$. As such, it lexicographically precedes the chain $\emptyset \lessdot \{p_{n-2}\} \lessdot \{p_{n-2}, p_n\}$. This implies that any permutation σ with $\sigma^{-1}(1) = n$ and $\sigma^{-1}(2) = n-2$ will precede any permutation τ with $\tau^{-1}(2) = n$ and $\tau^{-1}(1) = n-2$.

In particular, let σ be obtained from $\bar{\sigma}$ by switching the positions of 3 and 4. Let τ be obtained from $\bar{\sigma}$ by switching the positions of 1 and 2. Then in any EL-shelling, σ will come before τ . However, σ and τ have the same inversion number and have exactly one swap position, so they are interchangeable in any order given by the conditions of Theorem 5.3.5.

An analogous argument works when $n \geq 6$ is even, and can be adapted for the case when $n = 4$. □

This proposition and proof can also be adapted to show that not all shellings arising from Theorem 5.3.5 can be obtained from CL-labelings of $J(\mathcal{Z}_n)$.

5.4 The Swap Statistic Via Rank Selection

An alternate proof of Theorem 5.3.1 relies heavily on the concepts of rank selection and flag f -vectors developed for general posets in Sections 3.13 and 3.15 of [37]. We will focus our attention to the zig-zag poset, \mathcal{Z}_n . Denote by $J(\mathcal{Z}_n)$ the distributive lattice of order ideals in \mathcal{Z}_n ordered by inclusion. Let $S = \{s_1, \dots, s_k\} \subset [0, n]$, where $[0, n] = \{0, \dots, n\}$. We always assume that $s_1 < s_2 < \dots < s_k$. Denote by $\alpha_n(S)$ the number of chains of order ideals $I_1 \subsetneq \dots \subsetneq I_k$ in $J(\mathcal{Z}_n)$ such that $\#I_j = s_j$ for all j . Define

$$\beta_n(S) = \sum_{T \subset S} (-1)^{\#(S-T)} \alpha_n(T).$$

By the Principle of Inclusion-Exclusion, or equivalently, via Möbius inversion on the Boolean lattice,

$$\alpha_n(S) = \sum_{T \subset S} \beta_n(T).$$

In Section 3.13 of [37], the function $\alpha_n : 2^{[0, n]} \rightarrow \mathbb{Z}$ is called the *flag f -vector* of \mathcal{Z}_n and $\beta_n :$

$2^{[0,n]} \rightarrow \mathbb{Z}$ is called the *flag h -vector* of \mathcal{Z}_n . For any poset P of size n , let $\omega : P \rightarrow [n]$ be an order-preserving bijection that assigns a label to each element of P ; in this case, ω is called a *natural labeling*. Then for any linear extension $\sigma : P \rightarrow [n]$, we may define a permutation of the labels by $\omega(\sigma^{-1}(1)), \dots, \omega(\sigma^{-1}(n))$. The *Jordan-Hölder set* $\mathcal{L}(P, \omega)$ is the set of all permutations obtained in this way. The following result for arbitrary finite posets can be found in chapter 3.13 of [37].

Theorem 5.4.1 ([37], Theorem 3.13.1). *Let $S \subset [n-1]$. Then $\beta_n(S)$ is equal to the number of permutations $\tau \in \mathcal{L}(P, \omega)$ with descent set S .*

The *order polynomial* of a poset P , $\Omega_P(m)$ is the number of order preserving maps from P to $[m]$. The Ehrhart polynomial of $\mathcal{O}(\mathcal{Z}_n)$ evaluated at m is equal to the order polynomial of \mathcal{Z}_n evaluated at $m+1$ [35]. We also have the following equality of generating functions from Theorem 3.15.8 of [37]. We restate the relevant special case of this theorem here.

Theorem 5.4.2 ([37], Theorem 3.15.8). *Let $\omega : P \rightarrow [n]$ be an order-preserving bijection. Then*

$$\sum_{m \geq 0} \Omega_P(m) x^m = \frac{\sum_{\sigma \in \mathcal{L}(P, \omega)} x^{1+\text{des}(\sigma)}}{(1-x)^{p+1}},$$

where p is the cardinality of P .

Therefore, since $i_{\mathcal{O}(\mathcal{Z}_n)}(m) = \Omega_{\mathcal{O}(\mathcal{Z}_n)}(m+1)$, we have that

$$\text{Ehr}_{\mathcal{O}(\mathcal{Z}_n)}(t) = \frac{\sum_{\sigma \in \mathcal{L}(\mathcal{O}(\mathcal{Z}_n), \omega)} x^{\text{des}(\sigma)}}{(1-x)^{n+1}}.$$

It follows that the h^* -polynomial of $\mathcal{O}(\mathcal{Z}_n)$ is

$$h_{\mathcal{O}(\mathcal{Z}_n)}^*(t) = \sum_{S \subset [n-1]} \beta_n(S) t^{\#S}. \quad (5.2)$$

So, Theorem 5.3.1 will follow from Equation 5.2 and the following theorem, which is analogous to Theorem 3.13.1 in [37].

Theorem 5.4.3. *Let $S \subset [n-1]$. Then $\beta_n(S)$ is the number of alternating permutations ω with $\text{Swap}(\omega) = S$.*

To prove this theorem, for every $S = \{s_1, \dots, s_n\} \subset [n-1]$, we will find define a function ϕ_S that maps chains of order ideals of sizes s_1, \dots, s_k to alternating permutations whose swap set is contained in S . Let I_1, \dots, I_k be a chain of order ideals in $J(\mathcal{Z}_n)$ with sizes $\#I_j = s_j$. Let \mathbf{w}_i be the vertex of $\mathcal{O}(\mathcal{Z}_n)$ that satisfies

$$\mathbf{w}_i(j) = \begin{cases} 0 & \text{if } j \in I_i \\ 1 & \text{if } j \notin I_i. \end{cases}$$

Define $\phi_S(I_1, \dots, I_k)$ to be the unique alternating permutation that maximizes inversion number over all alternating permutations whose vertex set contains $\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$. This map is well-defined by Proposition 5.3.7.

Let ψ_S be the map that sends an alternating permutation ω with $\text{Swap}(\omega) \subset S$ to the chain of order ideals (I_1, \dots, I_k) where each $I_j = \{\omega^{-1}(1), \dots, \omega^{-1}(s_j)\}$. Since every alternating permutation ω is a linear extension of \mathcal{Z}_n , each I_j obtained in this way is an order ideal. They form a chain by construction, so the map ψ_S is well-defined. We will show that ψ_S is the inverse of ϕ_S in the proof of Theorem 5.4.3.

Example 5.4.4. Consider the zig-zag poset on seven elements \mathcal{Z}_7 pictured in Figure 5.9. Let $S = \{3, 6\}$, and let $I_1 = \{a, c, g\}$ and $I_2 = \{a, c, d, e, f, g\}$ be the given order ideals of sizes 3 and 6 respectively. Then the vectors \mathbf{w}_1 and \mathbf{w}_2 are

$$\mathbf{w}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{w}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} .$$

Notice that these are the same vectors \mathbf{w}_1 and \mathbf{w}_2 as in Example 5.3.8. So the unique alternating permutation $\phi_S(I_1, I_2)$ that maximizes inversion number over all alternating permutations whose vertex set contains $\{\mathbf{w}_1, \mathbf{w}_2\}$ is the same permutation as in Example 5.3.8,

$$\phi_S(I_1, I_2) = 3\ 7\ 2\ 6\ 4\ 5\ 1.$$

Note that $\text{Swap}(3726451) = \{3\} \subset \{3, 6\} = S$.

Now let $\omega = 3726451$. We will recover our original order ideals I_1 and I_2 by finding $\psi_S(\omega)$. For clarity, we will treat ω as a map from $\{a, \dots, g\}$ to $\{1, \dots, 7\}$. The first order ideal of $\psi_S(\omega)$ consists of the inverse images of 1, 2, and 3 in ω . That is,

$$I_1 = \{\omega^{-1}(1), \omega^{-1}(2), \omega^{-1}(3)\} = \{a, c, g\}.$$

The second order ideal of $\psi_S(\omega)$ consists of the inverse images of 1 through 6 in ω . So we obtain

$$I_2 = \{\omega^{-1}(1), \dots, \omega^{-1}(6)\} = \{a, c, d, e, f, g\}.$$

Note that this is, in fact, the chain of order ideals with which we began.

Proof of Theorem 5.4.3. Let $S = \{s_1, \dots, s_k\} \subset [n-1]$. We will show that $\alpha_n(S)$ is the number of alternating permutations whose swap set is contained in S by showing that the map ϕ_S described above is a bijection.

Let I_1, \dots, I_k be a chain of order ideals in $J(\mathcal{Z}_n)$ with sizes $\#I_j = s_j$. It is clear from the definitions

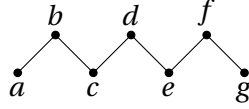


Figure 5.9 The zig-zag poset \mathcal{Z}_7

of ϕ_S and ψ_S that

$$\psi_S(\phi_S(I_1, \dots, I_k)) = (I_1, \dots, I_k).$$

Since ϕ_S is injective, it suffices to show that ψ_S is also injective. We will show that $\phi_S(I_1, \dots, I_k)$ is the only alternating permutation that maps to (I_1, \dots, I_k) under ψ_S .

Since $\omega = \phi_S(I_1, \dots, I_k)$ is the unique alternating permutation that maximizes inversion number over all alternating permutations with $\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ in their vertex sets, any other alternating permutation σ that maps to (I_1, \dots, I_k) under ψ_S must have fewer inversions than ω .

Let σ be such a permutation. Since each inversion between the sets $I_1, \mathcal{Z}_n - I_k$ and $I_j - I_{j-1}$ for all $1 < j \leq k$ are fixed, the additional non-inversion must be contained in one of these sets. Without loss of generality, let this be $R = I_j - I_{j-1}$. Denote by $\sigma|_R$ the restriction of σ to the domain R . Let $(\sigma^{-1}(a), \sigma^{-1}(b))$ be the non-inversion of $\sigma|_R$ that is *not* required by the alternating structure. Then by Lemma 5.3.4, there exists a k such that $a \leq k < b$ and k is a swap in σ . Since $a \leq k < b$, $\sigma^{-1}(k)$ and $\sigma^{-1}(k+1)$ are in R , so k is also a swap in $\sigma|_R$ as well. So the swap set of σ is not contained in S and we have reached a contradiction.

Therefore, ω is the only alternating permutation that can map to (I_1, \dots, I_k) under ψ_S , and ψ_S is the inverse map of ϕ_S . So $\alpha_n(S)$ is equal to the number of alternating permutations whose swap set is contained in S . By the Principle of Inclusion-Exclusion, $\beta_n(S)$ is the number of alternating permutations whose swap set is equal to S . \square

Theorem 5.3.1 follows as a corollary of Theorem 5.4.3.

Proof of Theorem 5.3.1. Equation 5.2 states that

$$h_{\mathcal{O}(\mathcal{Z}_n)}^*(t) = \sum_{S \subseteq [n-1]} \beta_n(S) t^{\#S}.$$

Theorem 5.4.3 tells us that $\beta_n(S)$ is the number of alternating permutations with swap set S . So the sum $\sum_{\#S=k} \beta_n(S)$ is the number of alternating permutations σ with $\text{swap}(\sigma) = k$. So

$$h_{\mathcal{O}(\mathcal{Z}_n)}^*(t) = \sum_{\sigma} t^{\text{swap}(\sigma)},$$

as needed. \square

We conclude this section with an equidistribution result that follows as a corollary of Theorem 5.3.1.

Corollary 5.4.5. *Let ω be a natural labeling of \mathcal{Z}_n . Then*

$$\sum_{\sigma \in A_n} t^{\text{swap}(\sigma)} = \sum_{\sigma \in \mathcal{L}(\mathcal{Z}_n, \omega)} t^{\text{des}(\sigma)}.$$

BIBLIOGRAPHY

- [1] AGRESTI, A. Modelling patterns of agreement and disagreement. *Statistical methods in medical research* 1, 2 (1992), 201–218.
- [2] AMENDOLA, C., BLISS, N., BURKE, I., GIBBONS, C. R., HELMER, M., HOŞTEN, S., NASH, E. D., RODRIGUEZ, J. I., AND SMOLKIN, D. The maximum likelihood degree of toric varieties. *J. Symbolic Comput.* 92 (2019), 222–242.
- [3] AOKI, S., HARA, H., AND TAKEMURA, A. *Markov bases in algebraic statistics*, vol. 199. Springer Science & Business Media, 2012.
- [4] BECK, M., AND ROBINS, S. *Computing the continuous discretely*. Springer, 2007.
- [5] BISHOP, Y. M. M., FIENBERG, S. E., AND HOLLAND, P. W. *Discrete multivariate analysis: theory and practice*. Springer, New York, 2007. With the collaboration of Richard J. Light and Frederick Mosteller, Reprint of the 1975 original.
- [6] BJORNER, A. Shellable and Cohen-Macaulay partially ordered sets. *Trans. Amer. Math. Soc.* 260, 1 (1980), 159–183.
- [7] BOCCI, C., AND RAPALLO, F. Exact tests to compare contingency tables under quasi-independence and quasi-symmetry. *J. Algebr. Stat.* 10, 1 (2019), 13–29.
- [8] BUCZYNSKA, W., AND WISNIEWSKI, J. A. On the geometry of binary symmetric models of phylogenetic trees. *Journal of the European Mathematical Society* 9, 3 (2007), 609–635.
- [9] BUNEMAN, P. A note on the metric properties of trees. *J. Combinatorial Theory Ser. B* 17 (1974), 48–50.
- [10] COLOMBO, A., AND IHM, P. A quasi-independence model to estimate failure rates. *Reliability Engineering & System Safety* 21, 4 (1988), 309–318.
- [11] COONS, J. I., AND SULLIVANT, S. The h^* -polynomial of the order polytope of the zig-zag poset. *arXiv preprint arXiv:1901.07443* (2019).
- [12] COONS, J. I., AND SULLIVANT, S. Quasi-independence models with rational maximum likelihood estimator. *J. Symbolic Comput.* 104 (2021), 917–941.
- [13] COONS, J. I., AND SULLIVANT, S. Toric geometry of the Cavender-Farris-Neyman model with a molecular clock. *Adv. in Appl. Math.* 123 (2021), 102119, 54.
- [14] COX, D. A., LITTLE, J. B., AND SCHENCK, H. K. *Toric varieties*, vol. 124 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2011.

- [15] DIACONIS, P., AND STURMFELS, B. Algebraic algorithms for sampling from conditional distributions. *The Annals of Statistics* 26, 1 (1998), 363–397.
- [16] DUARTE, E., MARIGLIANO, O., AND STURMFELS, B. Discrete statistical models with rational maximum likelihood estimator. arXiv:1903.06110, 2019.
- [17] EVANS, S. N., AND SPEED, T. P. Invariants of some probability models used in phylogenetic inference. *The Annals of Statistics* (1993), 355–377.
- [18] FELSENSTEIN, J. *Inferring phylogenies*, vol. 2. Sinauer associates Sunderland, MA, 2004.
- [19] GEIGER, D., MEEK, C., AND STURMFELS, B. On the toric algebra of graphical models. *Ann. Statist.* 34, 3 (2006), 1463–1492.
- [20] GOODMAN, L. A. On quasi-independence and quasi-dependence in contingency tables, with special reference to ordinal triangular contingency tables. *Journal of the American Statistical Association* 89, 427 (1994), 1059–1063.
- [21] GOUVEIA, J. A., GRAPPE, R., KAIBEL, V., PASHKOVICH, K., ROBINSON, R. Z., AND THOMAS, R. R. Which nonnegative matrices are slack matrices? *Linear Algebra Appl.* 439, 10 (2013), 2921–2933.
- [22] HASSETT, B. *Introduction to algebraic geometry*. Cambridge University Press, Cambridge, 2007.
- [23] HENDY, M. D., AND PENNY, D. Spectral analysis of phylogenetic data. *Journal of classification* 10, 1 (1993), 5–24.
- [24] HUH, J. Varieties with maximum likelihood degree one. *J. Algebr. Stat.* 5, 1 (2014), 1–17.
- [25] HUH, J., AND STURMFELS, B. Likelihood geometry. In *Combinatorial algebraic geometry*, vol. 2108 of *Lecture Notes in Math*. Springer, Cham, 2014, pp. 63–117.
- [26] JUKES, T. H., CANTOR, C. R., ET AL. Evolution of protein molecules. *Mammalian protein metabolism* 3, 21 (1969), 132.
- [27] KIMURA, M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of molecular evolution* 16, 2 (1980), 111–120.
- [28] KUBJAS, K. *Algebraic and combinatorial aspects of group-based models*. PhD thesis, Freie Universität Berlin, 2013.
- [29] RAPALLO, F. Algebraic exact inference for rater agreement models. *Stat. Methods Appl.* 14, 1 (2005), 45–66.
- [30] RAUH, J., KAHLE, T., AND AY, N. Support sets in exponential families and oriented matroid theory. *Internat. J. Approx. Reason.* 52, 5 (2011), 613–626.

- [31] SEMPLE, C., AND STEEL, M. A. *Phylogenetics*, vol. 24. Oxford University Press on Demand, 2003.
- [32] SIMON, B. *Representations of finite and compact groups*. No. 10. American Mathematical Soc., 1996.
- [33] SLOANE, N. The online encyclopedia of integer sequences, 2005. *Published electronically at <http://oeis.org>*.
- [34] SMITH, P. W., AND McDONALD, J. W. Exact conditional tests for incomplete contingency tables: estimating attained significance levels. *Statistics and Computing* 5, 3 (1995), 253–256.
- [35] STANLEY, R. P. Two poset polytopes. *Discrete & Computational Geometry* 1, 1 (1986), 9–23.
- [36] STANLEY, R. P. A survey of alternating permutations. *Contemporary Mathematics* 531 (2010), 165–196.
- [37] STANLEY, R. P. *Enumerative combinatorics. Volume 1*, second ed., vol. 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2012.
- [38] STURMFELS, B. *Gröbner bases and convex polytopes*, vol. 8. American Mathematical Society, 1996.
- [39] STURMFELS, B., AND SULLIVANT, S. Toric ideals of phylogenetic invariants. *Journal of Computational Biology* 12, 2 (2005), 204–228.
- [40] STURMFELS, B., AND XU, Z. Sagbi bases of Cox-Nagata rings. *arXiv preprint arXiv:0803.0892* (2008).
- [41] SULLIVANT, S. Toric fiber products. *Journal of Algebra* 316, 2 (2007), 560–577.
- [42] SULLIVANT, S. *Algebraic statistics*, vol. 194 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2018.
- [43] WACHS, M. L. Poset topology: tools and applications. In *Geometric combinatorics*, vol. 13 of *IAS/Park City Math. Ser.* Amer. Math. Soc., Providence, RI, 2007, pp. 497–615.
- [44] WEST, D. B. *Introduction to Graph Theory*, 2 ed. Pearson Education Inc., 2001.
- [45] ZIEGLER, G. M. *Lectures on polytopes*, vol. 152 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995.