The Cavender-Farris-Neyman Model with a Molecular Clock

Jane Ivy Coons and Seth Sullivant

May 14, 2018

Abstract

We give a combinatorial description of the toric ideal of invariants of the Cavender-Farris-Neyman model with a molecular clock (CFN-MC) 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. Finally, 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. 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.

1 Introduction

The field of phylogenetics is concerned with reconstructing evolutionary histories of different species or other taxonomical units (*taxa* for short), such as genes or bacterial strains (see [5, 10] 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 organisms, 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*, or *mutations*, within the DNA of an organism, wherein one nucleotide gets swapped out for another. Phylogenetic models often 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 [7, 8]. The entries of the transition matrices in this Markov process are the probabilities of observing a substitution from one base to another in the genome at the end of the given time interval.

In the case of the Cavender-Farris-Neyman, or CFN model we are interested in modeling the substitutions at a single site in the gene sequences of the taxa of interest. 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 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 [16], Gröbner bases in many cases [16], polyhedral description [2], the Hilbert series [2], and connections to the Hilbert scheme and toric degenerations [17].

In this paper, we study the CFN model with an added *molecular clock* condition, or the CFN-MC model. The molecular clock condition adds the requirement that the time elapsed from the root of the tree to any leaf of the tree is the same for all leaves. For a fixed tree, the set of all probability distributions in the CFN-MC model along that tree is a semialgebraic set in the probability simplex Δ_{2^n-1}, where n is the number of taxa. The CFN-MC model belongs to a special class of phylogenetic models called *group-based* models. This means that under a linear change of coordinates called the discrete Fourier transform, we can view the Zariski closure of the CFN-MC model as a toric variety [4, 6]. This allows us to study it from the point of view of polyhedral geometry. Our main results are summarized by the following:

Theorem. *Let T be a binary rooted tree with n leaves. Let I_T be the toric vanishing ideal of the CFN-MC model on the tree T, and let R_T be the associated polytope.*

1. *The toric vanishing ideal I_T has a generating set of quadratic and linear binomials. These binomials form a Gröbner basis that has squarefree initial terms. (Theorem 5.1)*

2. *The polytope R_T has F_n vertices where F_n denotes the nth Fibonacci number. (Proposition 4.3)*

3. *The normalized volume of R_T is E_{n-1}, where E_{n-1} is the Euler zig-zag number. (Theorem 6.2)*

4. The Hilbert series of I_T only depends on n, not the specific tree T. (Theorem 6.20)

These results are analogues of the major results in [2, 16, 17] for the CFN model, but the molecular clock assumption presents new challenges.

The outline of the paper is as follows. In Section 2, we give preliminary definitions regarding phylogenetic trees. In Section 3, we describe the CFN-MC model in detail. We describe how the discrete Fourier transform allows us to view the CFN-MC model as a toric variety. In Section 4, we study the combinatorics of the polytope associated to the toric ideal of phylogenetic invariants of the CFN-MC model. In particular, we give vertex and facet descriptions for this polytope, and show that the number of vertices of the CFN-MC polytope is equal to a Fibonacci number. In Section 5, we study the generators of the toric ideal of phylogenetic invariants of the CFN-MC model and give a quadratic Gröbner basis for this ideal. To accomplish this, we make use of the theory of toric fiber products [18], plus new tools to handle the case of cluster trees where the toric fiber product structure is not present. In Section 6, we prove our results on the Hilbert series of the CFN-MC ideal which implies the results on the volume of the polytope R_T.

2 Preliminaries on Trees

This section provides a brief background on combinatorial trees and metric trees. A more detailed description can be found in [5, 10].

Definition 2.1. 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 vertex 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 2.2. Consider the rooted binary tree in Figure 2.1a. 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 $\{1, \ldots, n\}$ 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 a. Furthermore, a rooted binary tree on n leaves has $n-1$ interior vertices and $2n-2$ edges. Proofs of these facts can be found in Chapter 2.1 of [20].

(a) An example of a rooted binary tree

(b) Branch lengths that make this tree equidistant.

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 vertices in the tree, where the distance between a pair of vertices is the sum of the branch lengths on the unique path in the tree connecting those vertices. Not every metric on a finite set arises in this way; for example, the resulting *tree metrics* must satisfy the four-point condition [3]. In this paper we are interested in the following restricted class of tree metrics.

Definition 2.3. 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 2.1b 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*.

3 The CFN-MC Model

In this section, we will review the discrete Fourier transform, as well as known results concerning the toric structure of the CFN model without the molecular clock [16]. 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 for the present paper.

The Cavender-Farris-Neyman (or CFN) model models 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 will often abbreviate purine with a U and pyrimidine with a 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

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

for some parameter α 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.

Let $t_e > 0$ be the branch length of an edge e in the rooted binary tree T. To obtain the transition matrix M_e associated to the edge e, we take the matrix exponential,

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

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 $1, \ldots, 2n-1$. For the following section, we will label these nodes so that $1, \ldots, n$ 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 will 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 with states in the state space, and let u_i denote the ith coordinate of $\mathbf{u}$, which is the labeling of node i. Then the probability of observing the set of states $\mathbf{u}$ is

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

Example 3.1. For the tree in Figure 3.1, 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 over all possible labelings of the interior nodes of the tree. Let $\mathbf{v} \in \mathbb{Z}_2^n$ be a

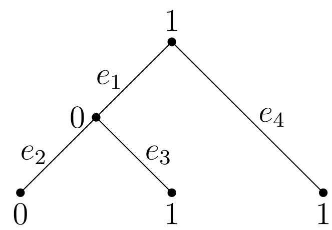


Figure 3.1: A labeling of all nodes of tree T with elements of $\mathbb{Z}_2$.

labeling of the leaves of T. Assuming a uniform distribution of states at the root, the probability of observing the set of states $\mathbf{v}$ at the leaves is

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

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

Example 3.2. Consider the tree from Figure 3.1. We will use Equation (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 interior nodes of T yields

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

In the following discussion, we will perform a linear change of coordinates on the variables and on the parameter space in order to realize this parametrization as a monomial map. In order to accomplish this, we will first provide some background concerning group-based models and the discrete Fourier transform. We will always assume that G is a finite abelian group.

Definition 3.3. Let $M^e = \exp(Qt_e)$ be a transition matrix arising from a continuous-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 $f^e : G \to \mathbb{R}$ 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 $f^e : \mathbb{Z}_2 \to \mathbb{R}$ 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 3.4. The *dual group* $\hat{G} = \mathrm{Hom}(G, \mathbb{C}^\times)$ of a group G is the group of all homomorphisms $\chi : G \to \mathbb{C}^\times$, 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.

We will make use of the following classical theorems, proofs of which can be found in [11].

Proposition 3.5. *Let G be a finite abelian group. Its dual $\hat{G}$ is a group that 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 3.6. Let $f : G \to \mathbb{C}$ be a function. The *discrete Fourier transform* of f is the function

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

The discrete Fourier transform is the linear change of coordinates that allows us to view equation (2) as a monomial parametrization. We can write the Fourier transform of p over $\mathbb{Z}_2^n$ in equation (2) as

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

Using the fact that G is isomorphic to $\hat{G}$, we can write $\hat{p}$ as a function of n elements of $\mathbb{Z}_2$ as

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

for all $(i_1, \ldots, i_n) \in \mathbb{Z}_2^n$. The following theorem, independently discovered by Evans and Speed in [4] and Hendy and Penny in [6], describes the monomial parametrization obtained from the discrete Fourier transform. A detailed account can also be found in Chapter 15 of [19].

Theorem 3.7. *Let $p(g_1, \ldots, g_n)$ be the polynomial describing the probability of observing states $(g_1, \ldots, g_n) \in G^n$ at the leaves of phylogenetic tree T under a group-based model with root distribution π. For each edge e, let f^e denote the function guaranteed to exist for each edge by definition of a group based model. Then the Fourier transform of p is*

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

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

Let $\hat{\mathbb{Z}}_2 = \{\mathbb{1}, \phi\}$, where ϕ denotes the only nontrivial homomorphism from $\mathbb{Z}_2$ to $\mathbb{C}$. 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 G and $\hat{G}$ which identifies $\mathbb{1}$ with 0 and ϕ with 1 gives that

$$\hat{p}(g_1, \ldots, g_n) = \begin{cases} \prod_{e \in E(T)} \hat{f}^e\big(\sum_{l \in \lambda e} g_l\big), & \text{if } \sum_{i=1}^n g_i = 0 \\ 0, & \text{if } \sum_{i=1}^n g_i = 1. \end{cases} \tag{3}$$

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 will be 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:

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

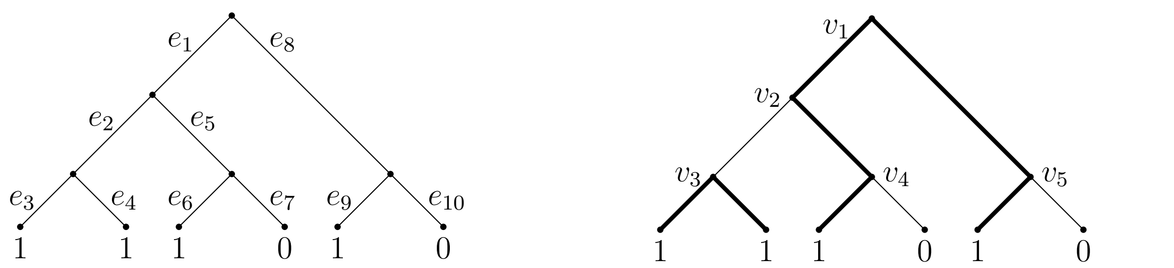
Using these new parameters, the isomorphism of G and $\hat{G}$ and equation (3), we can see that

$$\hat{p}(g_1, \ldots, g_n) = \begin{cases} \prod_{e \in E(T)} a_{i(e)}^e, & \text{if } \sum_{i=1}^n g_i = 0 \bmod 2 \\ 0, & \text{if } \sum_{i=1}^n g_i = 1 \bmod 2, \end{cases} \tag{4}$$

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 3.8. Consider the tree T in Figure 3.2a. We will compute $\hat{p}(1, 1, 1, 0, 1, 0)$. Using Equation (4), we see that

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



(a) A leaf labeling of T with elements of $\mathbb{Z}_2$ that sum to 0 mod 2

(b) The system of disjoint paths associated to this labeling

Figure 3.2: The tree T referenced in Example 3.8

Let e_1, e_2, e_3 be edges of rooted binary tree T that are adjacent to a single vertex v, where $v = t(e_1)$ and $v = h(e_2) = h(e_3)$. Then $i(e_1) = i(e_2) + i(e_3)$, so $i(e_1) + i(e_2) + i(e_3) = 0 \bmod 2$. In particular, this means that at any internal vertex, the edges adjacent to that vertex have an even number of 1's. This, along with the fact that if e_1, e_2 are the edges adjacent to the root, then $i(e_1) + i(e_2) = 0$, implies that labelings of the leaves of T with elements of $\mathbb{Z}_2$ that sum to 0 are equivalent to systems of disjoint paths between leaves of T.

Definition 3.9. 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 $0 \bmod 2$. The *system of disjoint paths* associated to a labeling $(i_1, \ldots, i_n) \in \mathbb{Z}_2^{n,\text{even}}$ is the unique collection 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 system of disjoint paths by $\mathfrak{P}$.

In this context, the edges for which a_1^e appears in the parametrization (4) of $\hat{p}(i_1, \ldots, i_n)$ for $(i_1, \ldots, i_n) \in \mathbb{Z}_2^{n,\text{even}}$ are exactly those that appear in the system of disjoint paths associated to $(i_1, \ldots, i_n)$.

Example 3.10. The system of disjoint paths associated to the labeling $(1, 1, 1, 0, 1, 0)$ from Example 3.8 is pictured in Figure 3.2b. 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 will 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. Recall that the molecular clock condition imposes that if $e_1, \ldots, e_s$ and $f_1, \ldots, f_r$ are two paths from an interior node v to leaves descended from v, then

$t_{e_1} + \cdots + t_{e_s} = t_{f_1} + \cdots + t_{f_s}$. On the level of transition matrices, this means that

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

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 ignore it. Furthermore, note that for any edge e_1 and $i = 0, 1$, we have the relations

$$\begin{aligned} a_i^{a(e_1)} \big(a_i^{d(e_1)}\big)^{-1} &= a_i^{e_1} \dots a_i^{e_s} (a_i^{e_s})^{-1} \dots (a_i^{e_s})^{-1} \\ &= a_i^{e_1}, \end{aligned} \tag{5}$$

where $e_1, \dots, e_s$ is a path from the ancestral vertex $a(e_1)$ 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 system of disjoint paths associated to $i_1, \dots, i_n$. Denote by $\text{Int}(T)$ the set of all interior nodes of tree T.

Definition 3.11. 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 will often denote the top-vector of a particular collection of paths $\mathfrak{P}$ by $[\mathfrak{P}]$ or $\mathbf{x}^{\mathfrak{P}}$, depending upon the context.

Example 3.12. Consider the system of disjoint paths associated to labeling $(1, 1, 1, 0, 1, 0) \in \mathbb{Z}_2^{n,\text{even}}$ pictured in Figure 3.2b. The top-set of this system of disjoint paths is $\text{Top}(1, 1, 1, 0, 1, 0) = \{v_1, v_3\}$. The top vector is $(1, 0, 1, 0, 0) \in \mathbb{R}^{\text{Int}(T)}$.

We will consider what terms appear in the parametrization of $\hat{p}(i_1, \dots, i_n)$ when we use Equation (5) to substitute the new parameters into Equation (4) based on the position of each internal node in the system of disjoint paths $\mathfrak{P}$ associated to $(i_1, \dots, i_n)$.

If v is not used in any path in $\mathfrak{P}$, then it appears in a factor of $(a_0^v)^2 (a_0^v)^{-1} = a_0^v$. If v is used in a path in $\mathfrak{P}$, but it is not the top-most vertex in that path, then v

appears in a factor of $a_0^v a_1^v (a_1^v)^{-1} = a_0^v$. If v is the top-most vertex of a path in $\mathfrak{P}$, then v appears in a factor of $(a_1^v)^2 (a_0^v)^{-1}$. So, defining new parameters

$$\begin{aligned} b_0^v &= a_0^v \quad \text{and} \\ b_1^v &= \begin{cases} (a_1^v)^2 & \text{if } v \text{ is the root, and} \\ (a_1^v)^2 (a_0^v)^{-1} & \text{otherwise} \end{cases} \end{aligned} \tag{6}$$

allows us to rewrite the parametrization in (4) as

$$\hat{p}(i_1, \ldots, i_n) = \prod_{v \in \mathrm{Top}(i_1,\ldots,i_n)} b_1^v \times \prod_{v \in \overline{\mathrm{Top}(i_1,\ldots,i_n)}} b_0^v. \tag{7}$$

Example 3.13. Consider the parametrization of $\hat{p}(1,1,1,0,1,0)$ given in Example 3.8 for the tree T pictured in 3.2a. First, we will verify the identity in Equation (5) for $a_1^{e_1}$. We have that $h(e_1) = v_1$ and $t(e_1) = v_2$. 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 (5) 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 (6) 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 in this parametrization of the Fourier coordinates, two labelings of the leaves with group elements $(i_1, \ldots, i_n)$ and $(j_1, \ldots, j_n)$ have the same top-sets if and only if $\hat{p}(i_1, \ldots, i_n) = \hat{p}(j_1, \ldots, j_n)$. Therefore, we can pass to new coordinates that are indexed by valid top-sets of collections of disjoint paths in T. These coordinates are in the polynomial ring

$$\mathbb{K}[\underline{r}] := \mathbb{K}[r_{k_1,\ldots,k_{n-1}} : (k_1, \ldots, k_{n-1}) = \mathrm{top}(i_1, \ldots, i_n) \text{ for some } (i_1, \ldots i_n) \in \mathbb{Z}_2^{n,\mathrm{even}}]$$

where $(i_1, \ldots, i_n)$ ranges over all elements of $\mathbb{Z}_2^{n,\mathrm{even}}$. 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^v parameters are equal, and restrict our attention to equivalences classes of labelings in $\mathbb{Z}_2^{n,\mathrm{even}}$ with the same top-sets.

Definition 3.14. Label the interior vertices of T with $v_1, \ldots, v_{n-1}$. The *CFN-MC ideal* I_T is the kernel of the map

$$\begin{aligned} \mathbb{K}[\underline{r}] &\longrightarrow \mathbb{K}[b_i^v \mid i = 0, 1, v \in \mathrm{Int}(T)] \\ r_{k_1,\ldots,k_{n-1}} &\longmapsto \prod_{i=1}^{n-1} b_{k_i}^{v_i}, \end{aligned}$$

where $(k_1, \ldots, k_{n-1})$ ranges over all indicator vectors corresponding to top-sets of collections of disjoint paths 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.

Definition 3.15. A *toric ideal* is the kernel of a monomial map. Equivalently, it is a prime ideal that is generated by binomials.

To every monomial map $\mathbb{K}[x_1, \ldots, x_m] \to \mathbb{K}[y_1, \ldots, y_d]$, we can associate a $d \times m$ integer matrix. The entry in the (i, j)th position of this matrix is the exponent of y_j in the image of x_i under this map.

Background on toric ideals can be found beginning in Chapter 4 of [15]. Some applications of toric ideals to phylogenetics are detailed in [16].

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

$$\begin{aligned} \mathbb{K}[\underline{r}] &\longrightarrow \mathbb{K}[t_0, \ldots, t_{n-1}] \\ r_{k_1,\ldots,k_{n-1}} &\longmapsto t_0 \prod_{i:k_i=1} t_i, \end{aligned} \tag{8}$$

where t_0 is a homogenizing indeterminate. 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 collections of disjoint paths in T with an added homogenizing row of ones. The convex hull of these indicator vectors gives a polytope in $\mathbb{R}^{n-1}$ that encodes important information about the ideal I_T [15, Chapter 4]. Our goal in this paper is to study the ideals I_T for binary trees and the corresponding polytopes R_T (to be defined in detail in Section 4.

Example 3.16. Let T be the tree pictured in Figure 2.1a. The CFN-MC ideal I_T is in the polynomial ring $\mathbb{K}[\underline{r}] = \mathbb{K}[r_{0000}, r_{1000}, r_{0100}, r_{0010}, r_{0001}, r_{1010}, r_{1001}, r_{0011}]$ where each subscript is the indicator vector of a top-set of a collection of disjoint paths in T indexed alphabetically by the internal nodes of T. Therefore, the parametrization in Equation (8) is given by

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

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 $\mathbb{K}[\underline{r}]$ 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 [15, 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 5.12; in this setting, the first column are the elements of the "Lift" set and the second column are the elements of the "Swap" set.

4 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 Section 6 that the volume also does not depend on the number of leaves.

Let T be a rooted binary tree on n leaves. For any collection of disjoint paths $\mathfrak{P}$ in T, let $\mathbf{x}^{\mathfrak{P}} \in \mathbb{R}^{n-1}$ have ith component $x_i^{\mathfrak{P}} = 1$ if i is the highest interior vertex in some path in $\mathfrak{P}$ and $x_i^{\mathfrak{P}} = 0$ otherwise. Hence $\mathbf{x}^{\mathfrak{P}}$ is the indicator vector of the tops of the paths in $\mathfrak{P}$ as discussed in the previous section.

Definition 4.1. 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 collection of disjoint paths in T.

Example 4.2. For the tree in Figure 2.1a, the polytope R_T is the convex hull of the column vectors of the matrix A_T in Example 3.16. For completeness, 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 will write $\operatorname{conv}(A_T)$ 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 4.3. *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 $R_T = \operatorname{conv}\begin{bmatrix} 0 & 1 \end{bmatrix}$, and if T is the 3-leaf tree, then $R_T = \operatorname{conv}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Let T be an n-leaf tree. Let l_1 and l_2 be leaves of T that are adjacent to the same interior 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 contracting leaves l_1 and l_2 to a single leaf and suppressing node a. If $\mathfrak{P}$ is a collection of paths in T with $a \notin \operatorname{Top}(\mathfrak{P})$, then we can realize the top-vector of $\mathfrak{P}$ without the a-coordinate as the top-vector of a collection of paths in T'. Furthermore, any collection of paths in T' can be extended to a collection of paths 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 suppressing node a'. If $\mathfrak{P}$ is a collection of paths in T with $a \in \operatorname{Top}(\mathfrak{P})$, then note that $a' \notin \operatorname{Top}(\mathfrak{P})$. 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 collection of paths in T''. Furthermore, any collection of paths in T'' can be extended to a collection of paths 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 will 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 [21, 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 $n-1$ interior nodes. The interior nodes will now be labeled by $1, \ldots, n-1$.

Remark 4.4. Note that for the remainder of the paper we have changed the convention of labeling the the vertices so that $1, \ldots, n-1$ now label the interior vertices of the trees, whereas in Section 3 we used $1, \ldots, n$ to denote the leaves. This change is because of the importance that the interior nodes now play in the combinatorics of the CFN-MC model, whereas in Section 3 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 edge that points from v towards the root of T. Introduce a poset $\mathrm{Int}(T)$ whose elements are the interior vertices of T and with relations $v \leq w$ if v is a descendant of w. Let I be an order ideal of $\mathrm{Int}(T)$ 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 vertex in I has exactly two edges directly beneath it.

Definition 4.5. Let T be a tree, $\mathrm{Int}(T)$ the associated poset, and I an order ideal of $\mathrm{Int}(T)$. Denote by $T-I$ the tree obtained by removing all nodes and edges descended from any node in I. The *edge set of* $T-I$, $\mathcal{E}(T-I)$, 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 4.6. Let $\mathfrak{P}$ be a collection of disjoint paths 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

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

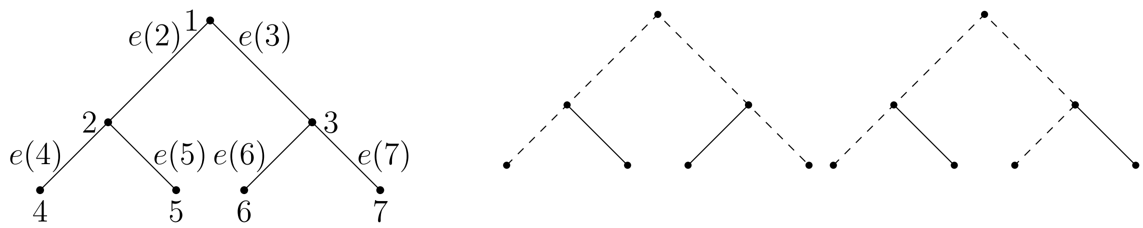
for all $i \in I$ and

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

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 collections of disjoint paths between leaves of T, $\mathfrak{P}$. If I is the set of all interior vertices of T, then this polytope is exactly R_T, and $[\mathbf{x}, \mathbf{y}]_I^{\mathfrak{P}} = \mathbf{x}^{\mathfrak{P}} = [\mathfrak{P}]$.

Example 4.7. Let T be the 4-leaf tree pictured below.

Let the distinguished order ideal in the set of interior vertices 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 vertex.



(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 4.1: The four-leaf tree in Example 4.7

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

Figure 4.1b shows the paths through T that realize the vertex $\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \end{bmatrix}^{\mathrm{T}}$ in $R_T(I)$.

For any order ideal I with maximal vertex r, let a and b be the descendants of r. Then we can define a linear map

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

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

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

Here, $\mathbf{x}'$ has elements indexed by vertices in $I - \{r\}$ and $\mathbf{y}'$ has elements indexed by vertices in $\mathcal{E}(T - (I - \{r\}))$. Then $\mathbf{x}$ has elements indexed by vertices in I and $\mathbf{y}$ has elements index by vertices in $\mathcal{E}(T - I)$.

Note that if r is the root, then there is no y'_r. So we let $x_r = \frac{y'_a + y'_b}{2}$. But in this case, $R_{I-\{r\}}$ lies in the hyperplane defined by $y'_a = y'_b$. So $x_r = y'_a = y'_b$.

Proposition 4.8. *The function $\phi_{I,r}$ maps $R_T(I - \{r\})$ onto $R_T(I)$.*

Proof. We will show that for all collections of disjoint paths through $\mathfrak{P}$ in T, the image of $[\mathbf{x}', \mathbf{y}']^{\mathfrak{P}}_{I-\{r\}}$ under $\phi_{I,r}$ is $[\mathbf{x}, \mathbf{y}]^{\mathfrak{P}}_{I}$. Since $\mathfrak{P}$ is a collection of disjoint paths and, if r is a vertex in one of these paths, then the path includes exactly two edges about r, 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 $[\mathbf{x}', \mathbf{y}']^{\mathfrak{P}}_{I-\{r\}}$ under $\phi_{I,r}$ is $[\mathbf{x}, \mathbf{y}]^{\mathfrak{P}}_I$ 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 $[\mathbf{x}', \mathbf{y}']^{\mathfrak{P}}_{I-\{r\}}$ under $\phi_{I,r}$ is $[\mathbf{x}, \mathbf{y}]^{\mathfrak{P}}_I$ 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 $[\mathbf{x}', \mathbf{y}']^{\mathfrak{P}}_{I-\{r\}}$ under $\phi_{I,r}$ is $[\mathbf{x}, \mathbf{y}]^{\mathfrak{P}}_I$ in this case.

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

Definition 4.9. Let I be an order ideal in the poset consisting of all interior vertices of T. An element $v \in I$ is called a *cluster node* if v is connected by edges to three other interior 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 interior nodes of T that lie in $I - C$ and are adjacent to some node in I. When I is the set of all interior vertices of T, we denote the neighbor set by $N(C)$. Denote by $m(C)$ the maximal vertex of C.

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

Example 4.10. Consider the tree T in Figure 4.2. Then the set of nodes marked with triangles, $\{b, c\}$ forms a cluster, since both b and c are 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 $m(\{b, c\}) = b$.

The main result of this section is a list of the facet defining inequalities of the polytopes R_T, 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.

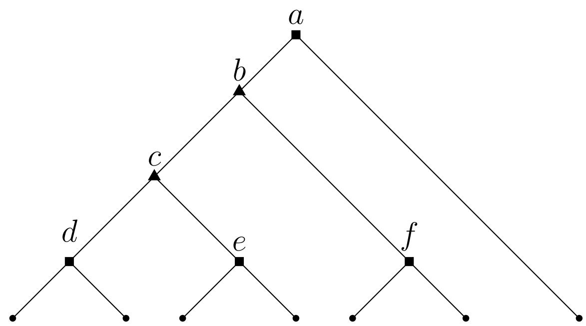


Figure 4.2: An example of a cluster

Theorem 4.11. *The polytope $R_T(I)$ is the solution to the following set of constraints:*

- $y_s = y_t$, *where edges $e(s)$ and $e(t)$ are joined to the root.*
- $y_i - y_j - y_k \leq 0$, *where $e(i), e(j), e(k)$ are three distinct edges that meet at a single vertex 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 vertex 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*
- $2\sum_{i \in C} x_i + \sum_{j \in N_I(C)} x_j + y_{m(C)} \leq |C| + 1$ *for all clusters $C \subset I$.*

Note that if $m(C)$ is not a maximal vertex of I, then there is no coordinate $y_{m(C)}$. In the final cluster inequality in Theorem 4.11, this reduces to the inequality

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

in this case. 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 4.11.

Example 4.12. Consider the tree T in Figure 4.3. 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:

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

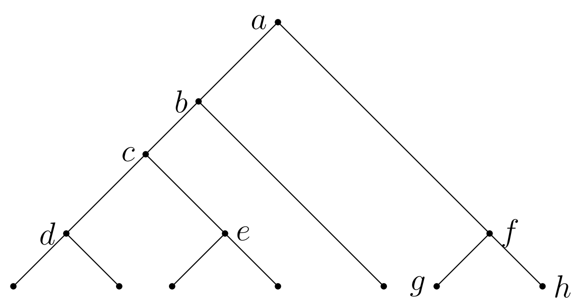


Figure 4.3: The tree T in Example 4.12

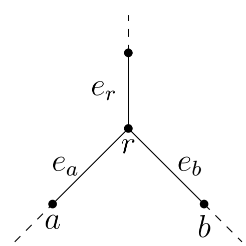
$$
\begin{aligned}
y_f - y_g - y_h &\leq 0, & x_b + x_c &\leq 1,\\
-y_f + y_g - y_h &\leq 0, & x_c + x_d &\leq 1,\\
-y_f - y_g + y_h &\leq 0, & x_c + x_e &\leq 1,\\
y_f + y_g + y_h &\leq 2, & x_b + 2x_c + x_d + x_e + y_b &\leq 2,
\end{aligned}
$$

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

Proof of Theorem 4.11. We proceed by induction on m, the size of the order ideal I.

When $m = 0$, $R_T(\emptyset)$ is the polytope associated to the CFN model (or binary Jukes-Cantor model, or binary symmetric model), as described in [16]. It follows from the results in [16, 2] 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 interior vertex.

Let I have size m, and let $1, \ldots, r$ be the maximal vertices of I. Suppose that $R_T(I - \{r\})$ has its facets defined by the proposed inequalities. About vertex r, we have the following picture.



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

In order to project $R_T(I - \{r\})$ onto $R_T(I)$, we will "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 will 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.

$$\begin{aligned} -y'_a &\le 0, & y'_a - y'_b - y'_r &\le 0,\\ -y'_b &\le 0, & -y'_a + y'_b - y'_r &\le 0,\\ x'_a + y'_a &\le 1, & -y'_a - y'_b + y'_r &\le 0,\\ x'_b + y'_b &\le 1, & y'_a + y'_b + y'_r &\le 0,\\ 2\sum_{i\in C} x'_i + \sum_{j\in N_{I-\{r\}}(C)} x'_j + y'_a &\le |C|+1, & 2\sum_{i\in D} x'_i + \sum_{j\in N_{I-\{r\}}(D)} x'_j + y'_b &\le |D|+1, \end{aligned}$$

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 will be true of C and D throughout the following discussion

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.

$$-x_r \le 0 \tag{0}$$
$$y_a - y_r - 2x_r \le 0 \tag{1_+}$$
$$y_a + x_a \le 1 \tag{2_+}$$
$$y_a - y_r - x_r \le 0 \tag{3_+}$$
$$y_a + 2\sum_{i\in C} x_i + \sum_{j\in N_{I-\{r\}}(C)} x_j \le |C|+1 \tag{4_+}$$
$$-y_a \le 0 \tag{1_-}$$
$$-y_a + y_r + x_b + 2x_r \le 1 \tag{2_-}$$
$$-y_a + x_r \le 0 \tag{3_-}$$
$$-y_a + y_r + 2x_r + 2\sum_{i\in D} x_i + \sum_{j\in N_{I-\{r\}}(D)} x_j \le |D|+1 \tag{4_-}$$

We perform Fourier-Motzkin elimination to obtain the following 17 types of inequalities, labeled by which of the above inequalities where combined to obtain them, in addition to all of the inequalities from $R_T(I-\{r\})$ that did not contain y'_a or y'_b.

$$-x_r \leq 0 \tag{0}$$

$$-2x_r - y_r \leq 0 \tag{1_+1_-}$$

$$x_b \leq 1 \tag{1_+2_-}$$

$$-x_r - y_r \leq 0 \tag{1_+3_-}$$

$$2\sum_{i\in D} x_i + \sum_{j\in N_{I-\{r\}}(D)} x_j \leq |D| + 1 \tag{1_+4_-}$$

$$x_a \leq 1 \tag{2_+1_-}$$

$$2x_r + x_a + x_b + y_r \leq 2 \tag{2_+2_-}$$

$$x_r + x_a \leq 1 \tag{2_+3_-}$$

$$2\sum_{i\in D\cup\{r\}} x_i + y_r + \sum_{j\in N_I(D\cup\{r\})} x_j \leq |D| + 2 \tag{2_+4_-}$$

$$-x_r - y_r \leq 0 \tag{3_+1_-}$$

$$x_r + x_b \leq 1 \tag{3_+2_-}$$

$$-y_r \leq 0 \tag{3_+3_-}$$

$$2\sum_{i\in D} x_i + \sum_{j\in N_I(D)} x_j \leq |D| + 1 \tag{3_+4_-}$$

$$2\sum_{i\in C} x_i + \sum_{j\in N_{I-\{r\}}(C)} x_j \leq |C| + 1 \tag{4_+1_-}$$

$$2\sum_{i\in C\cup\{r\}} x_i + y_r + \sum_{j\in N_I(C\cup\{r\})} x_j \leq |C| + 2 \tag{4_+2_-}$$

$$2\sum_{i\in C} x_i + x_r \sum_{j\in N_I(C)} x_j \leq |C| + 1 \tag{4_+3_-}$$

$$2x_r + 2\sum_{i\in C} x_i + 2\sum_{j\in D} x_j + y_r + \sum_{k\in N_{I-\{r\}}(C)} x_k + \sum_{l\in N_{I-\{r\}}(D)} x_l \leq |C| + |D| + 2 \tag{4_+4_-}$$

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 plus another copy of inequality 3_+3_-. Inequality 1_+4_- can be obtained by adding together inequalities 3_+4_- and 0. 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 4.13. *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. The fact that the facet defining inequalities of R_T are a subset of the proposed inequalities follows directly from the theorem.

Now we must show that none of the proposed inequalities are redundant. To do this, we will 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 interior 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 $\mathbf{e}_i + \mathbf{e}_k$ as 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 $F = \{\mathbf{x} \mid 2\sum_{c \in C} x_c + \sum_{i \in N(C)} = |C| + 1\}$, 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 $i, k \in N(C)$, $i \neq k \neq j$ will be 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, $\{\mathbf{y}_j \mid j \in N(C)\}$ is a collection 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 collection 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 descendants of C 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 collection exists for all $c \in C$. Then $\mathbf{w}_c$ is a vertex of R_T since it results from the collection of paths 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 guaranteed 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 interior 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. □

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 4.14. 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 the root of the tree with a single pendant leaf in its other subtree. Attach a balanced 4-leaf tree descended from ever 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 4.4 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.

5 Generators of the CFN-MC Ideal

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

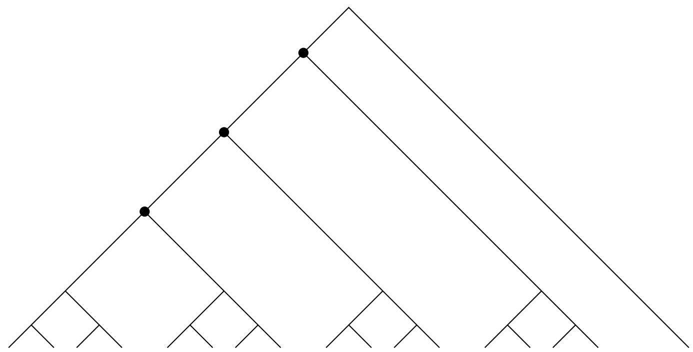


Figure 4.4: The tree construction for T_3 described in Example 4.14. "Spine" nodes are marked with a circle.

Theorem 5.1. *For any rooted binary phylogenetic tree T, the CFN-MC ideal I_T has a Gröbner basis consisting of homogeneous quadratic binomials.*

To accomplish this, we will show that for most trees T, the CFN-MC ideal is the toric fiber product of two smaller trees. In these cases, we can use results from [18] to describe the generators of I_T in terms of the generators of these smaller trees. We will then handle the case of trees for which I_T is not a toric fiber product, called *cluster trees*.

For simplicity of notation, we will switch to denoting the top-vector associated to a collection of paths $\mathfrak{P}$ by $[\mathfrak{P}]$. As before, note that it is possible to have two different collections of paths $\mathfrak{P}$ and $\mathfrak{Q}$ for which $[\mathfrak{P}] = [\mathfrak{Q}]$.

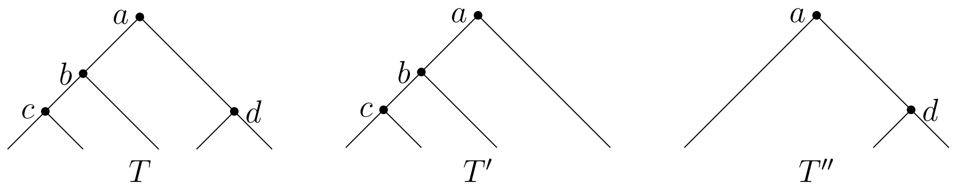
5.1 Toric Fiber Products

Let T be a tree that has an interior node v that is adjacent to exactly two other interior 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''.

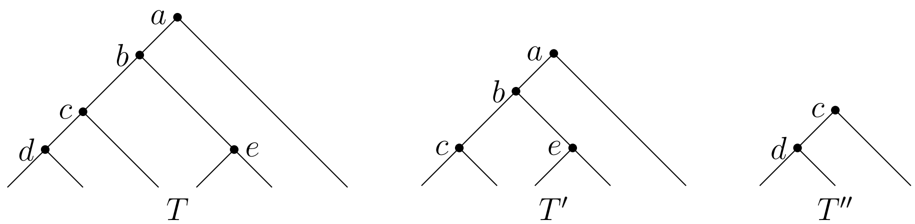
If v is the root, then let T' be the tree with v as a root in which the right subtree of v is equal to the right subtree of T and the left subtree of v is a single edge. Let T'' be the tree with v as a root in which the left subtree of v is equal to the left subtree of T and the right subtree of v is a single edge. This decomposition is pictured in Figure 5.1a.

If v is not the root, then v is adjacent to two interior vertices and a leaf. 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 5.1b.

In either case, notice that $[\mathfrak{P}]$ is the top-vector of a collection of disjoint paths 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 collections of disjoint paths in T' and T'' that agree on v. In other



(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 5.1: Decomposition of T into T' and T'' via a node adjacent to exactly two interior nodes

words, 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 from each, 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 [18].

Let $I_T \subset \mathbb{K}[\underline{r}]$, $I_{T'} \subset \mathbb{K}[\underline{x}]$, $I_{T''} \subset \mathbb{K}[\underline{y}]$. Consider the map $\xi_{I_{T'},I_{T''}}$ from $\mathbb{K}[\underline{r}]$ to $\mathbb{K}[\underline{x}] \otimes_{\mathbb{K}} \mathbb{K}[\underline{y}]$ defined as follows. For any collection of disjoint paths $\mathfrak{P}$ 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 [18], 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''. The fact that v is either the root or is adjacent to a leaf ensures that we can join any systems of disjoint paths in T' and T'' that agree on v to create a system of disjoint paths in T. Let S be the two-leaf tree rooted at v, and let $\mathbb{K}[\underline{z}] = \mathbb{K}[z_0, z_1]$ be its associated polynomial ring.

Proposition 5.2. *Suppose that T has an interior 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 $I_T \cong I_{T'} \times_{\mathcal{A}} I_{T''}$.*

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

$$\begin{aligned} \psi_T : k[\underline{r}] &\to k[t_0, \ldots, t_{n-1}] \\ r_{[\mathfrak{P}]} &\mapsto t_0 \prod_{[\mathfrak{P}]_i = 1} t_i \end{aligned}$$

Denote by $\overline{\mathfrak{P}}$ the restriction of $\mathfrak{P}$ to S. Then we have the identity

$$\psi_S(z_{[\overline{\mathfrak{P}}]})\psi_T(r_{[\mathfrak{P}]}) = \psi_{T'}(x_{[\mathfrak{P}']})\psi_{T''}(y_{[\mathfrak{P}'']}). \tag{9}$$

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

$$\begin{aligned} \xi_{I_{T'}, I_{T''}} : k[\underline{r}] &\to k[t_0, \ldots, t_{n-1}] \\ r_{[\mathfrak{P}]} &\mapsto \big(t_0 \prod_{[\mathfrak{P}']_i = 1} t_i\big)\big(t_0 \prod_{[\mathfrak{P}'']_i = 1} t_i\big) \end{aligned}$$

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 map $\phi_{I_{T'}, I_{T''}}$ without changing the kernel. This yields the same map as ψ_T. Therefore,

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

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 will define several sets of polynomials in $\mathbb{K}[\underline{r}]$ that together form a Gröbner basis for I_T with respect to some weighted monomial order. Let $f = \prod_{i=1}^d x_{[\mathfrak{P}_i]} - \prod_{i=1}^d x_{[\mathfrak{Q}_i]} \in G_1$ arranged so that $[\mathfrak{P}_i]_v = [\mathfrak{Q}_i]_v$ for all i. Let $R(f)$ denote the set of all d-tuples $(\mathfrak{R}_1, \ldots, \mathfrak{R}_d)$ of paths in T'' such that $\mathfrak{R}_i$ agreeswith $\mathfrak{P}_i$ at node v. Let $\mathfrak{P}_i \cup \mathfrak{R}_i$ denote the collection of paths in T obtained by viewing $\mathfrak{P}_i$ and $\mathbb{R}_i$ as paths in T. Define the set

$$\mathrm{Lift}(f) = \left\{ \prod_{i=1}^d r_{[\mathfrak{P}_i \cup \mathfrak{R}_i]} - \prod_{i=1}^d r_{[\mathfrak{Q}_i \cup \mathfrak{R}_i]} : (\mathfrak{R}_1, \ldots, \mathfrak{R}_d) \in R(f) \right\}.$$

Then let

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

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

Let $[\mathfrak{P}_1], \ldots, [\mathfrak{P}_r]$ be the top-vectors of paths in T' with v-coordinate equal to 0. let $[\mathfrak{Q}_1], \ldots, [\mathfrak{Q}_s]$ be the top vectors of paths in T'' with v-coordinate equal to 0. Define the set $\mathrm{Quad}_0(T)$ to be the set of all 2×2 minors of the matrix M_0 with (i,j)th entry equal to $r_{[\mathfrak{P}_i \cup \mathfrak{Q}_j]}$. Define $\mathrm{Quad}_1(T)$ and M_1 analogously over all top-vectors in T' and T'' with v-coordinate equal to 1. Elements of Quad_k are of the form

$$r_{[\mathfrak{P}_i \cup \mathfrak{Q}_j]} r_{[\mathfrak{P}_{i'} \cup \mathfrak{Q}_{j'}]} - r_{[\mathfrak{P}_i \cup \mathfrak{Q}_{j'}]} r_{[\mathfrak{P}_{i'} \cup \mathfrak{Q}_j]},$$

where $[\mathfrak{P}_i], [\mathfrak{Q}_j], [\mathfrak{P}_{i'}]$ and $[\mathfrak{Q}_{j'}]$ all take value k on their v-coordinate. Let ω be a weight vector on $\mathbb{K}[\underline{r}]$ so that Quad(T) is a Gröbner basis for the ideal generated by all elements of Quad(T). Since the $\mathcal{A}$-matrix of the toric fiber product is invertible, Theorem 2.9 of [18] 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 $\mathbb{K}[\underline{x}]$ and $\mathbb{K}[\underline{y}]$ to the affine space associated to $\mathbb{K}[\underline{r}]$. If $\mathfrak{P}$ and $\mathfrak{Q}$ are collections of disjoint paths in T' and T'' respectively that agree on v, then the $[\mathfrak{P} \cup \mathfrak{Q}]$ coordinate of $\xi^*_{I_{T'},I_{T''}}(\alpha, \beta)$ is $\alpha_{[\mathfrak{P}]} + \beta_{[\mathfrak{Q}]}$.

Proposition 5.3. *Suppose that T has an interior 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 Lift$(\mathcal{G}_1) \cup$ Lift$(\mathcal{G}_2) \cup$ 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 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.

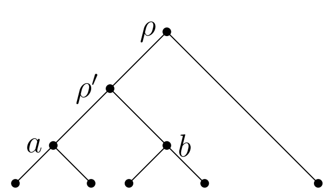
5.2 Cluster Trees

Trees that do not have an interior node that is adjacent to exactly two other interior nodes do not have the toric fiber product structure described in the previous section. These are the trees that are comprised of one large cluster and its neighbor vertices. In this case, we will 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 5.4. A rooted binary tree T is called a *cluster tree* if there exists a cluster C in T such that every non-leaf vertex of T is either in C or in $N(C)$.

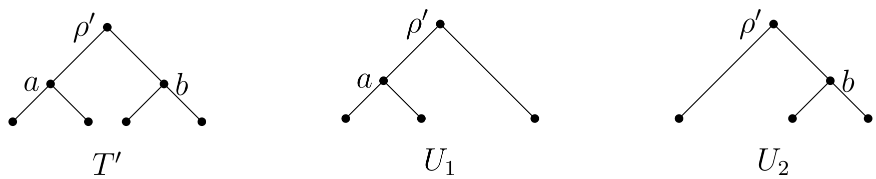
Equivalently, if T has n leaves, then T is a cluster tree if and only if T has a cluster of size $(n-3)/2$. Note that this implies that if T is a cluster tree, then T has an odd number of leaves. It also follows from the definition of a cluster tree that the root of T must be adjacent on one side to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

Example 5.5. 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 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.

Example 5.6. From the previous example, T', U_1 and U_2 are as follows.



Notice that by the results of the previous section, $I_{T'}$ is the toric fiber product of the CFN-MC ideals of two cluster trees, I_{U_1} and I_{U_2}. For this reason, we will call T' a *bicluster tree.*

We are interested in defining when we can add a path with highest node at ρ to a collection of paths in the larger cluster tree, T. This occurs if and only if this collection of paths does not already have a path with highest node ρ and the restriction of the collection of paths to T' has the following property.

Definition 5.7. A collection of disjoint paths $\mathfrak{P}$ 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 interior vertex that is the top-most vertex of some path in $\mathfrak{P}$.

Note that a collection of paths is root-leaf traversable if and only if removing all of the maximal vertices of paths in the collection leaves ρ' in the same connected component as some leaf of T'. Therefore, root-leaf traversability is well-defined over classes of collections of disjoint paths with the same top-set. We will often say that $[\mathfrak{P}]$ is root-leaf traversable if $\mathfrak{P}$ is root-leaf traversable.

Since T' is a bicluster tree, in order for a collection of paths 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. Therefore, we define the following condition of *root augmentability* on cluster trees. Note that we cannot use the same definition as that of root-leaf traversability, since in a cluster tree, any collection of paths 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.

Definition 5.8. A collection of disjoint paths $\mathfrak{P}$ is *root-augmentable* 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.

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 5.9. Let S be a cluster tree and let the CFN-MC ideal of S, $I_S \subset k[\underline{x}]$. A term order $<$ on $k[\underline{x}]$ 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 5.10. Let S be a bicluster tree and let the CFN-MC ideal of S, $I_S \subset k[\underline{x}]$. A term order $<$ on $k[\underline{x}]$ 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.

Let $I_{U_1} \subset k[\underline{x}]$, $I_{U_2} \subset k[\underline{y}]$, $I_{T'} \subset k[\underline{z}]$ and $I_T \subset k[\underline{r}]$. 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 $k[\underline{z}]$ 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. Let ω_1, ω_2 be weight vectors that induce liftable orders on I_{U_1} and I_{U_2}, respectively.

Proposition 5.11. *There exist a weight vector ω on $k[\underline{z}]$ and $\epsilon, k > 0$ such that $\phi^*_{n+1}(\omega_1, \omega_2) + \epsilon\omega + k\mathbf{a}$ induces a liftable order on $I_{T'}$*

Proof. We can write any collection of paths in the bicluster tree T' as $\mathfrak{P} \cup \mathfrak{Q}$ where $\mathfrak{P}$ is a collection of paths in U_1, $\mathfrak{Q}$ is a collection of paths in U_2, and $\mathfrak{P}$ and $\mathfrak{Q}$ agree on the root of T'. Note that if $f = z_{[\mathfrak{P}_1 \cup \mathfrak{Q}_1]} z_{[\mathfrak{P}_2 \cup \mathfrak{Q}_2]} - z_{[\mathfrak{P}_1 \cup \mathfrak{Q}_2]} z_{[\mathfrak{P}_2 \cup \mathfrak{Q}_1]} \in \mathrm{Quad}_i$, then

$$\begin{aligned}\phi^*_{n+1}(\omega_1, \omega_2)(z_{[\mathfrak{P}_1 \cup \mathfrak{Q}_1]} z_{[\mathfrak{P}_2 \cup \mathfrak{Q}_2]}) &= \omega_1(x_{[\mathfrak{P}_1]}) + \omega_2(y_{[\mathfrak{Q}_1]}) + \omega_1(x_{[\mathfrak{P}_2]}) + \omega_2(y_{[\mathfrak{Q}_2]}) \\ &= \phi^*_{n+1}(\omega_1, \omega_2)(z_{[\mathfrak{P}_1 \cup \mathfrak{Q}_2]} z_{[\mathfrak{P}_2 \cup \mathfrak{Q}_1]}).\end{aligned}$$

We will find $\epsilon > 0$ and weight vector ω to "break ties" for leading monomials in each Quad_i. If f is a 2×2 minor of Quad_1, 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 Quad_0, 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 $\mathfrak{P}_1, \mathfrak{Q}_1$ be root augmentable and $\mathfrak{P}_2, \mathfrak{Q}_2$ not. Then $\mathfrak{P}_2 \cup \mathfrak{Q}_2$ is not root-leaf traversable, while $\mathfrak{P}_1 \cup \mathfrak{Q}_1, \mathfrak{P}_1 \cup \mathfrak{Q}_2$ and $\mathfrak{P}_2 \cup \mathfrak{Q}_1$ are. So, we must select an ω so that $z_{[\mathfrak{P}_1 \cup \mathfrak{Q}_1]} z_{[\mathfrak{P}_2 \cup \mathfrak{Q}_2]}$ is chosen as the leading monomial of f.

We will define this weight vector ω by assigning its values on the entries of Quad_0. Arrange collections of paths $\mathcal{A}_1, \ldots, \mathcal{A}_r$ in U_1 so that if $\mathcal{A}_i$ is root augmentable and $\mathcal{A}_j$ is not, then $i < j$. Arrange collections of paths $\mathcal{B}_1, \ldots, \mathcal{B}_s$ in U_2 so that if $\mathcal{B}_i$ is root augmentable and $\mathcal{B}_j$ is not, then $i < j$.

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

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

So ω choose the correct leading term of f. We can allow ω to be any weight vector on Quad_1 that chooses leading terms as in Proposition 2.6 of [18]. Pick ϵ to be small enough so that for all $g \in \mathrm{Lift}(\mathcal{G}_1) \cup \mathrm{Lift}(\mathcal{G}_2)$,

$$LT_{\phi^*_{n+1}(\omega_1,\omega_2)}(g) = LT_{\phi^*_{n+1}(\omega_1,\omega_2)+\epsilon\omega}(g).$$

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

$$f = z_{[\mathfrak{P}_1 \cup \mathfrak{R}_1]} z_{[\mathfrak{P}_2 \cup \mathfrak{R}_2]} - z_{[\mathfrak{Q}_1 \cup \mathfrak{R}_1]} z_{[\mathfrak{Q}_2 \cup \mathfrak{R}_2]} \in \mathrm{Lift}(\mathcal{G}_1).$$

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 $\phi^*_{n+1}(\omega_1, \omega_2)$ may not select the correct leading term occurs when, without loss of generality, $\mathfrak{P}_1, \mathfrak{R}_1$ and $\mathfrak{Q}_2$ are not root augmentable and $\mathfrak{P}_2, \mathfrak{R}_2$ and $\mathfrak{Q}_1$ are root augmentable. In this case, $\mathfrak{P}_1 \cup \mathfrak{R}_1$ is not root-leaf traversable and $\mathfrak{P}_2 \cup \mathfrak{R}_2, \mathfrak{Q}_1 \cup \mathfrak{R}_1$ and $\mathfrak{Q}_2 \cup \mathfrak{R}_2$ are, but $x_{[\mathfrak{P}_1]} x_{[\mathfrak{P}_2]} - x_{[\mathfrak{Q}_1]} x_{[\mathfrak{Q}_2]}$ may not have $x_{[\mathfrak{P}_1]} x_{[\mathfrak{P}_2]}$ as its leading term. Suppose that under the weight vector $\phi^*_{n+1}(\omega_1, \omega_2)$, $z_{[\mathfrak{Q}_1 \cup \mathfrak{R}_1]} z_{[\mathfrak{Q}_2 \cup \mathfrak{R}_2]}$ 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 $\mathcal{G} = \mathrm{Lift}(\mathcal{G}_1) \cup \mathrm{Lift}(\mathcal{G}_2) \cup \mathrm{Quad}_B$.

First note that $\overline{f} = z_{[\mathfrak{Q}_1 \cup \mathfrak{R}_1]} z_{[\mathfrak{Q}_2 \cup \mathfrak{R}_2]} - z_{[\mathfrak{P}_1 \cup \mathfrak{R}_2]} z_{[\mathfrak{P}_2 \cup \mathfrak{R}_1]} \in \mathrm{Lift}(\mathcal{G}_1)$ with $z_{[\mathfrak{Q}_1 \cup \mathfrak{R}_1]} z_{[\mathfrak{Q}_2 \cup \mathfrak{R}_2]}$ as the leading term under $\phi^*_{n+1}(\omega_1, \omega_2) + \epsilon\omega$, and both terms of $\overline{f}$ have all root-leaf traversable indices. Since f and $\overline{f}$ have the same leading term, $\mathcal{G} - \{f\}$ is still a Gröbner basis under $\phi^*_{n+1}(\omega_1, \omega_2) + \epsilon\omega$. Let $\mathcal{G}'$ be $\mathcal{G}$ with all such $f \in \mathrm{Lift}(\mathcal{G}_1) \cup \mathrm{Lift}(\mathcal{G}_2)$ 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 $g_1, \ldots, g_r \in \mathcal{G}'$ 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 $\phi^*_{n+1}(\omega_1, \omega_2) + \epsilon\omega$, but m_2 is the leading term in the order induced by $\phi^*_{n+1}(\omega_1, \omega_2) + \epsilon\omega + \mathbf{a}$. Then m_2 has more variables whose indices are not root-leaf traversable than m_1. We claim that division by the same $g_1, \ldots, g_r$, possibly in a different order, still reduces g to 0. To divide g by one of $g_1, \ldots, g_r$, 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 $\phi^*_{n+1}(\omega_1, \omega_2) + \epsilon\omega + \mathbf{a}$. $\square$

For any collection of disjoint paths $\mathfrak{P}$ in T, let $\mathfrak{P}'$ be the collection of disjoint paths in T' obtained from $\mathfrak{P}$ by deleting any path in $\mathfrak{P}$ that contains the root, ρ of T. Define two maps $\psi, \psi' : k[\underline{r}] \to k[\underline{z}]$ 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 the monomial order on $k[\underline{z}]$ guaranteed by Proposition 5.11. Then we can define the monomial order $\prec$ on $k[\underline{r}]$ by $\underline{r}^{\mathbf{b}} \prec \underline{r}^{\mathbf{c}}$ if and only if

- $\psi(\underline{r}^{\mathbf{b}}) < \psi(\underline{r}^{\mathbf{c}})$, or
- $\psi(\underline{r}^{\mathbf{b}}) = \psi(\underline{r}^{\mathbf{c}})$ and $\psi'(\underline{r}^{\mathbf{b}}) < \psi'(\underline{r}^{\mathbf{c}})$.

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 $I_{T'}$ guaranteed by Proposition 5.11. Let $f = z_{[\mathfrak{P}_1]}z_{[\mathfrak{P}_2]} - z_{[\mathfrak{Q}_1]}z_{[\mathfrak{Q}_2]} \in \mathcal{F}$. Define the set $\mathrm{Root}(f)$ to be the set of all possible binomials in I_T that result from treating $\mathfrak{P}_1, \mathfrak{P}_2, \mathfrak{Q}_1$ and $\mathfrak{Q}_2$ as collections of paths in T, with or without an added path with top-most vertex ρ. In other words,

$$\mathrm{Root}(f) = \{r_{i_1[\mathfrak{P}_1]}r_{i_2[\mathfrak{P}_2]} - r_{j_1[\mathfrak{Q}_1]}r_{j_2[\mathfrak{Q}_2]}\}$$

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 $\mathrm{Root}(T) = \bigcup_{f \in G_\phi} \mathrm{Root}(f)$. Define the set

$$\mathrm{Swap}(\rho) = \{r_{1[\mathfrak{P}]}r_{0[\mathfrak{Q}]} - r_{0[\mathfrak{P}]}r_{1[\mathfrak{Q}]}\}$$

where $\mathfrak{P}$ and $\mathfrak{Q}$ range over all root-leaf traversable collections of disjoint paths in T'.

Define the set $\mathcal{G}_\prec = \mathrm{Root}(T) \cup \mathrm{Swap}(\rho)$.

Proposition 5.12. *The term order $\prec$ is liftable. In particular, $G_\prec$ is a Gröbner basis for I_T with respect to $\prec$.*

Proof. First, we will show that $G_\prec$ constitutes a Gröbner basis. Since I_T is toric, it suffices to show that every binomial in I_T can be reduced via the elements of $G_\prec$. Let $\underline{\prod_{i=1}^d r_{[\mathfrak{P}_i]}} - \prod_{i=1}^d r_{[\mathfrak{Q}_i]} \in I_T$. Then if we arrange the terms in each monomial as a table with the vector representing each $[\mathfrak{P}_i]$ (resp. $[\mathfrak{Q}_i]$) as a row, the column sums of each of these tables are equal. By the definition of $\prec$, we have

$$\prod_{i=1}^{d} \psi(r_{[\mathfrak{P}_i]}) \geq_\phi \prod_{i=1}^{d} \psi(r_{[\mathfrak{Q}_i]}).$$

For all $\mathfrak{P}_i, \mathfrak{Q}_i$, let $\psi(r_{[\mathfrak{P}_i]}) = z_{[\mathfrak{P}'_i]}$ and $\psi(r_{[\mathfrak{Q}_i]}) = z_{[\mathfrak{Q}'_i]}$.

We can use the elements of $\mathcal{F}$ to reduce $\prod_{i=1}^d z_{[\mathfrak{P}'_i]} - \prod_{i=1}^d z_{[\mathfrak{Q}'_i]}$ in $I_{T'}$. The properties of the order on $I_{T'}$ induced by ϕ guarantee that (without loss of generality) if we divide by $z_{[\mathfrak{P}'_1]}z_{[\mathfrak{P}'_2]} - z_{[\mathfrak{R}'_1]}z_{[\mathfrak{R}'_2]}$ in this reduction, then the number of $\mathfrak{R}'_1$ and $\mathfrak{R}'_2$ 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 $\underline{r_{i_1[\mathfrak{P}'_1]}r_{i_2[\mathfrak{P}'_2]}} - r_{j_1[\mathfrak{R}'_1]}r_{j_2[\mathfrak{R}'_2]} \in \mathrm{Root}(z_{[\mathfrak{P}'_1]}z_{[\mathfrak{P}'_2]} - z_{[\mathfrak{R}'_1]}z_{[\mathfrak{R}'_2]})$ 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 $\prec$, $r_{i_1[\mathfrak{P}'_1]}r_{i_2[\mathfrak{P}'_2]}$ is in fact the leading term of this binomial.

This Gröbner basis reduction using elements Root(T) ends in a binomial of the form

$$\prod_{k=1}^{d} r_{i_k[\mathfrak{R}_k]} - \prod_{k=1}^{d} r_{j_k[\mathfrak{R}_k]}$$

where $\sum_{k=1}^d i_k = \sum_{k=1}^d j_k$. At this point, we can use elements of Swap(ρ) 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 Swap(ρ).

Now we can check that $\prec$ is a liftable term order on the elements of $G_\prec$. Any binomial in Swap(ρ) has one term that is root-augmentable and one that is not. So the choice of leading term of elements of Swap(ρ) does not affect the liftability property.

Let $f = r_{[\mathfrak{P}_1]}r_{[\mathfrak{P}_2]} - r_{[\mathfrak{Q}_1]}r_{[\mathfrak{Q}_2]} \in \mathrm{Root}(T)$. Then in particular, $\psi(r_{[\mathfrak{P}_1]}r_{[\mathfrak{P}_2]}) \neq \psi(r_{[\mathfrak{Q}_1]}r_{[\mathfrak{Q}_2]})$. There are several cases.

If $[\mathfrak{P}_1]_\rho = [\mathfrak{P}_2]_\rho = [\mathfrak{Q}_1]_\rho = [\mathfrak{Q}_2]_\rho = 1$, then neither monomial in f has a root-augmentable term. So the leading term of f does not affect the liftability property.

If $[\mathfrak{P}_1]_\rho = [\mathfrak{Q}_1]_\rho = 1$ and $[\mathfrak{P}_2]_\rho = [\mathfrak{Q}_2]_\rho = 0$, without loss of generality, then $[\mathfrak{P}_1]$ and $[\mathfrak{Q}_1]$ are both not root-augmentable, and $[\mathfrak{P}'_1]$ and $[\mathfrak{Q}'_1]$ both *are* root-leaf traversable. If $[\mathfrak{P}_2]$ and $[\mathfrak{Q}_2]$ 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 $[\mathfrak{Q}_2]$ is. Then under ψ, $z_{[\mathfrak{P}'_1]}z_{[\mathfrak{P}'_2]} >_\phi z_{[\mathfrak{Q}'_1]}z_{[\mathfrak{Q}'_2]}$ since $z_{[\mathfrak{P}'_1]}z_{[\mathfrak{P}'_2]}$ has one root-augmentable term and $z_{[\mathfrak{Q}'_1]}z_{[\mathfrak{Q}'_2]}$ has two root-augmentable terms.

If $[\mathfrak{P}_1]_\rho = [\mathfrak{P}_2]_\rho = [\mathfrak{Q}_1]_\rho = [\mathfrak{Q}_2]_\rho = 0$, 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 $<$ is liftable, the monomial with the fewest root-augmentable terms is chosen as the leading term of f, as needed. $\square$

Proof of Theorem 5.1. 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 5.12, 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 5.3 and induction on n imply that I_T has a quadratic Gröbner basis with squarefree initial terms. $\square$

Corollary 5.13. *The CFN-MC polytope has a regular unimodular triangulation and is normal.*

Proof. By Theorem 5.1, 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 $a^2 - bc$ are not possible in a toric ideal whose associated matrix is a 0/1 matrix. Since Theorem 5.1 shows that I_T has a quadratic Gröbner basis, and the leading term of each element of the Gröbner basis is square-free, so the leading term ideal of I_T with respect to the liftable term order $\prec$ is generated by square-free monomials. Therefore, it is the Stanley-Reisner ideal of a regular unimodular triangulation of R_T [15, Theorem 8.3]. $\square$

6 Ehrhart Function of the CFN-MC Polytope

In this section, we will show that for a rooted binary tree T, the Hilbert series of the CFN-MC ideal I_T depends only on the number of leaves of T and not on the topology of T. To accomplish this, we will 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 [2], and of Kubjas, who gave a combinatorial proof of the same result [9]. We provide a brief review of some definitions in Ehrhart theory below, and refer the reader to [1] for a more complete treatment of this material.

Let $P \subset \mathbb{R}^n$ be any polytope with integer vertices. Recall that the *Ehrhart function*, $i_P(m)$, counts in the integer points in dilates of P; that is,

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

where mP denotes the mth dilate of P. The Ehrhart function is, in fact, a polynomial in m. We further define the *Ehrhart series* of P to be the generating function

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

When P is full-dimensional in the ambient space $\mathbb{R}^n$, the Ehrhart series is of the form

$$\mathrm{Ehr}_P(t) = \frac{h^*(t)}{(1-t)^{n+1}},$$

where $h^*(t)$ is a polynomial in t of degree at most n. Furthermore, recall that since R_T has a regular unimodular triangulation, the Ehrhart series of R_T is equal to the Hilbert series of I_T [15, Chapter 8]. We will now introduce *alternating permuations* and the *Euler zig-zag numbers*, which enumerate these combinatorial objects. We will show that these numbers give the normalized volume of the CFN-MC polytopes.

Definition 6.1. A permutation on n letters $a_1 \dots a_n$ is *alternating* if $a_1 < a_2 > a_3 < a_4 > \dots$. The *Euler zig-zag number* E_n is the number of alternating permutations on n letters.

In other words, a permutation is alternating if its descent set is exactly the set of even numbers less than n. For example, in one-line notation, the permutation 13254 is alternating, while 13245 is not since its fourth position is not a descent. The reader should note that some texts refer to these as reverse-alternating permutations and to permutations with descent set equal to the *odd* numbers less than n as alternating. However, the relevant results are not affected by which terminology we choose.

Alternating permutations are fascinating combinatorial objects in their own right. For instance, 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 $E_0 = E_1 = 1$. The sequence of Euler zig-zag numbers beginning with E_0 begins $1, 1, 1, 2, 5, 16, 61, 272, \dots$ and can be found in the Online Encyclopedia of Integer Sequences with identification number A000111 [12]. For a thorough treatment of topics related to alternating permutations and the Euler zig-zag numbers, we refer the reader to Stanley's "Survey of Alternating Permutations" [14]. The goal of this section is to prove the following theorem.

Theorem 6.2. *For any rooted binary tree T with n leaves, the normalized volume of I_T is E_{n-1}, the $(n-1)$st Euler zig-zag number.*

The proof of Theorem 6.2 has two parts. First, we will 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 [14]. 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 nearest neighbor interchange. Since any two binary trees on n leaves are connected by a sequences of nearest neighbor interchanges, this will prove the theorem.

6.1 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 6.3. A *caterpillar tree* C_n on n leaves is the unique rooted tree topology with exactly one cherry.

Definition 6.4. The *zig-zag poset* P_n is the poset on underlying set $\{p_1, \dots, p_n\}$ with the cover relations $p_i < p_{i+1}$ for i odd and $p_i > p_{i+1}$ for i even. Note that these are exactly the inequalities that appear in the definition of an alternating permutation. The *order polytope* of the zig-zag poset $\mathcal{O}(P_n)$ 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 $p_i < p_j$ in P_n.

Order polytopes for arbitrary posets have been the object of considerable study, and are discussed in detail in [13]. For instance, the order polytope of P_n is also the convex hull of all $(v_1, \dots, v_n) \in \{0,1\}^n$ that correspond to labelings of P_n that are weakly consistent with the partial order on $\{p_1, \dots, p_n\}$.

In the case of $\mathcal{O}(P_n)$, the facet defining inequalities are those of the form

$$\begin{aligned} -v_i &\leq 0 \text{ for } i \leq n \text{ odd} \\ v_i &\leq 1 \text{ for } i \leq n \text{ even} \\ v_i - v_{i+1} &\leq 0 \text{ for } i \leq n-1 \text{ odd, and} \\ -v_i + v_{i+1} &\leq 0 \text{ for } i \leq n-1 \text{ even.} \end{aligned} \tag{10}$$

Note that the inequalities of the form $-v_i \leq 0$ for i even and $v_i \leq 1$ for i odd are redundant.

Every order polytope has a unimodular triangulation whose simplices are in bijection with linear extensions of the underlying poset [13]. In the case of the zig-zag poset P_n, the linear extensions of P_n are in bijection with alternating permutations, since we can simply take each alternating permutation to be a labeling of the poset [14]. These facts together imply the following proposition.

Proposition 6.5. *The order polytope $\mathcal{O}(P_n)$ has normalized volume E_n, then nth Euler zig-zag number.*

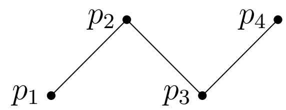


Figure 6.1: The zig-zag poset P_4

Example 6.6. Consider the zig-zag poset P_4 pictured in Figure 6.1. The matrix whose columns are the vertices of $\mathcal{O}(P_4)$ and the facet-defining hyperplanes of $\mathcal{O}(P_4)$ are given below. The volume of $\mathcal{O}(P_4)$ is $E_4 = 5$.

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \end{bmatrix} \qquad \begin{aligned} -v_1 &\leq 0 \\ v_2 &\leq 1 \\ -v_3 &\leq 0 \\ v_4 &\leq 1 \end{aligned} \qquad \begin{aligned} v_1 - v_2 &\leq 0 \\ -v_2 + v_3 &\leq 0 \\ v_3 - v_4 &\leq 0 \end{aligned}$$

Proposition 6.7. *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 $\mathcal{O}(P_n)$.

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

$$\phi(\mathbf{x})_i = \begin{cases} x_i & \text{if } i \text{ is odd, and} \\ 1 - x_i & \text{if } i \text{ is even.} \end{cases}$$

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 (10), as needed. □

Corollary 6.8. *The Ehrhart functions of the CFN-MC polytope $R_{C_{n+1}}$ and the order polytope $\mathcal{O}(P_n)$ are equal for all n. This further implies that the normalized volume of $R_{C_{n+1}}$ is the nth Euler zig-zag number, E_n.*

Proof. The Ehrhart functions $i_{R_{C_{n+1}}}(m)$ and $i_{\mathcal{O}(P)}(m)$ are equal because ϕ is a lattice-point preserving transformation from $mR_{C_{n+1}}$ to $m\mathcal{O}(P_n)$. The leading coefficient of the Ehrhart polynomial of a polytope is the volume of that polytope. This volume is $\frac{E_n}{n!}$ for $\mathcal{O}(P_n)$ and so, for $R_{C_{n+1}}$ as well [14]. So the normalized volume of $R_{C_{n+1}}$ is E_n. □

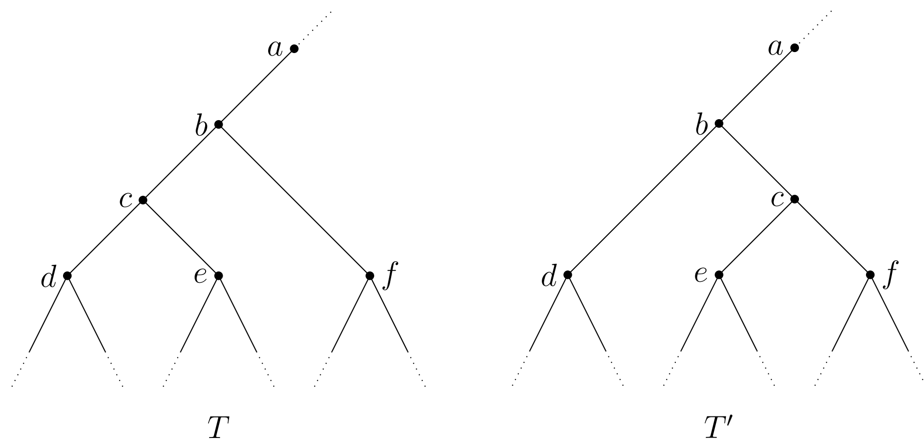


Figure 6.2: An NNI move

6.2 The Ehrhart Function and NNI Moves

We will 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 nearest neighbor interchange, or NNI move. This shows that the Ehrhart polynomials of R_T and $R_{T'}$ are the same. Since the Ehrhart series of R_T is equal to the Hilbert series of I_T, and since any tree can be obtained from any other tree by a finite sequence of NNI moves, this will prove Theorem 6.2.

Let b, c, e be three consecutive vertices of a tree T, where c is a descendant of b, and e is a descendant of c. Note that vertex e need not be an interior vertex of T. There is a unique *nearest neighbor interchange* or *NNI move* 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 NNI move, and its effect on the internal structure of T is depicted in Figure 6.2. An NNI move 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.

Definition 6.9. Let the tree T' be obtained from T by the NNI move associated to (b, c, e) as pictured in Figure 6.2. 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 will use the following definition to give a characterization of the maintaining and nonmaintaining vertices of R_T.

Definition 6.10. Let S be the top-set of a collection of disjoint paths in T. Let x be an interior 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 collection of paths $\mathfrak{P}$ that realizes S, we cannot add another path to $\mathfrak{P}$ with top-most vertex above x that passes through x.

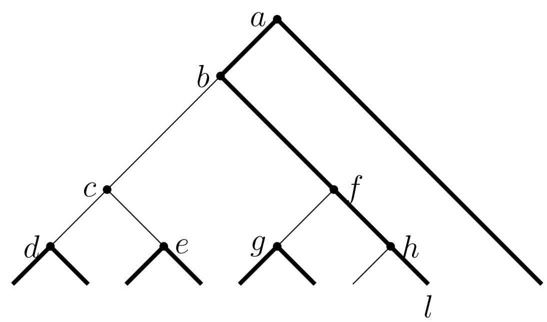


Figure 6.3: Nodes a, c, d, e and g are the nodes that are blocked in top-set $\{a, d, e, g\}$.

Example 6.11. Let T be the tree pictured in Figure 6.3 with a collection of paths $\mathfrak{P}$ drawn in bold. Note that up to a swap of the leaves below node h, $\mathfrak{P}$ is the only collection of paths 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 h to leaf l, and $h \notin \{a, d, e, g\}$. Similarly, b is not blocked in $\{a, d, e, g\}$.

Proposition 6.12. *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 6.2.*

(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.

(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 $b, c \notin V$. Since $b, c, \notin V$, and 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 collection of paths that realizes V in T', as needed. If such a $P \in \mathfrak{P}$ does exist, then we can modify it 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 collection of paths 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 collection of paths $\mathfrak{P}'$ 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 $\overline{P}$ from d to a leaf descended from d. Since all

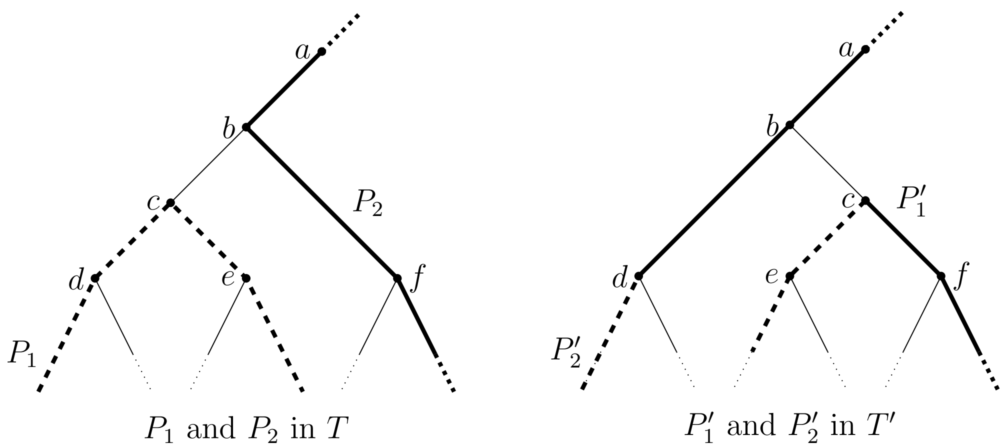


Figure 6.4: Proposition 6.12 (iii)

paths in $\mathfrak{P}'$ are disjoint, no path in $\mathfrak{P}'$ has its top-most node along $\overline{P}$. 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 collection of paths in T that realizes V, and let $P \in \mathfrak{P}$ have top-most vertex b. We may assume that $\hat{P}$ is contained in P. P also includes a path $\overline{P}$ from f to a leaf descended from f. Let $P' = \hat{P} \cup \overline{P} \cup \{bd, bc, cf\}$. Then $\mathfrak{P}' = (\mathfrak{P} - \{P\}) \cup \{P'\}$ is a collection of paths 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 collection of paths $\mathfrak{P}'$ 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 $\overline{P}$ 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 $\overline{P}$. So f is not blocked in V.

Suppose that f is not blocked in V. Let $P_1 \in \mathfrak{P}$ have top-most vertex c. There exists a path from f to a leaf descended from f none of whose vertices are in V. Note that this path may be contained in some path $P_2 \in \mathfrak{P}$. So, we will call this path from f to a leaf descended from f $\overline{P_2}$. Since P_2 must have top-most node above b, $bf \in P_2$. If such a P_2 exists, let $\hat{P}_2 = P_2 - (\overline{P_2} \cup \{bf\})$.

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,

$$P_1' = \hat{P}_1 \cup \overline{P_2} \cup \{ce, cf\}.$$

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 collection of paths in T' $[\mathfrak{P}'] = [\mathfrak{P}]$. This is illustrated in Figure 6.4, where P_1 is drawn with a dashed line, and P_2 is bolded. 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 6.13. *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 6.12, 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 collection of paths 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 6.12 since we are applying the result to the tree obtained after the NNI has been performed.) Performing the reverse operation on a c-nonmaintaining vertex $[\mathfrak{P}']$ of T' shows that this is a bijection. $\square$

Definition 6.14. For any vertex $[\mathfrak{P}]$ of R_T and node a of T, let v_a denote that a-coordinate of $[\mathfrak{P}]$. If $[\mathfrak{P}]$ is nonmaintaining, let $\mathfrak{P}'$ be the collection of paths described in the proof of the previous proposition, 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. The previous proposition allows us to define the following involution between the vertices of R_T and $R_{T'}$:

$$\begin{aligned} \phi^{T,T'} : \operatorname{vert}(R_T) &\to \operatorname{vert}(R_{T'}) \\ \mathbf{v} \quad &\mapsto \begin{cases} [\mathfrak{P}], & \text{if } [\mathfrak{P}] \text{ is maintaining} \\ [\mathfrak{P}'], & \text{if } [\mathfrak{P}] \text{ is nonmaintaining.} \end{cases} \end{aligned}$$

We will 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 5.13, R_T is normal. So, we may write $\mathbf{v} = [\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$ for some $[\mathfrak{P}_1], \ldots, [\mathfrak{P}_m] \in \operatorname{vert}(R_T)$. We call $[\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$ 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 6.15. A representation $[\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m] = \mathbf{v} \in mR_T \cap \mathbb{Z}^{n-1}$ is *d-compressed* if

- for all $[\mathfrak{P}_i]$ with b and $c \notin V_i$, d is blocked in V_i, or

- all of $[\mathfrak{P}_1], \ldots, [\mathfrak{P}_m]$ with b-coordinate equal to 1 are maintaining.

Similarly, this representation is *f-compressed* if

- for all $[\mathfrak{P}_i]$ with b and $c \notin V_i$, f is blocked in V_i, or
- all of $[\mathfrak{P}_1], \ldots, [\mathfrak{P}_m]$ with c-coordinate equal to 1 are maintaining.

If $[\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$ is both d-compressed and f-compressed, then we say that the representation is *df-compressed.*

Consider the map $\phi_m^{T,T'} : (mR_T \cap \mathbb{Z}^{n-1}) \to (mR_{T'} \cap \mathbb{Z}^{n-1})$ 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 will follow Lemmas 6.16 and 6.17 below.

Lemma 6.16. *If $[\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$ is a minimal representation of $\mathbf{v} \in mR_T \cap \mathbb{Z}^{n-1}$, then $[\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$ is df-compressed.*

In the proofs of lemmas 6.16 and 6.17, we will 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 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 collections of disjoint paths $\mathfrak{P}$ 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 6.16. We will prove the contrapositive. Suppose that $[\mathfrak{P}_1]+\cdots+[\mathfrak{P}_m]$ is not df-compressed. Then this representation is either not d-compressed or not f-compressed. We will show that $[\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$ is not minimal in both cases.

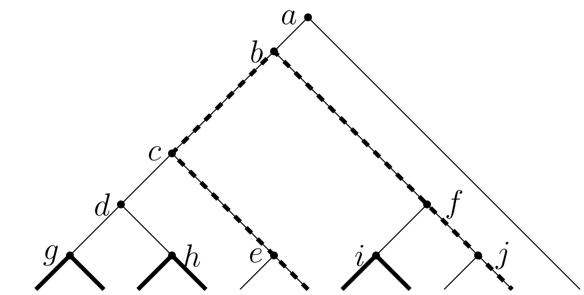
If $[\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$ is not d-compressed, then without loss of generality, we may assume that $[\mathfrak{P}_1]$ is b-nonmaintaining. Furthermore, we may assume that $[\mathfrak{P}_2]$ has b and c-coordinates equal to 0 and that d is not blocked in $[\mathfrak{P}_2]$. We 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 collections $\overline{\mathfrak{P}_1}$ and $\overline{P_2}$ respectively. We further claim that $[\overline{\mathfrak{P}_1}]$ and $[\overline{\mathfrak{P}_2}]$ are maintaining.

Let $P \in \mathfrak{P}_1$ with top-most node b. Let $\hat{P}$ be the path from b to a node below f contained in P. Let $\overline{P}$ be the path from d to a leaf descended from d that does not contain any nodes in V_2; this is guaranteed to exist since d is not blocked in V_2. Let $P' = \hat{P} \cup \{bc, cd\} \cup \overline{P}$. Then

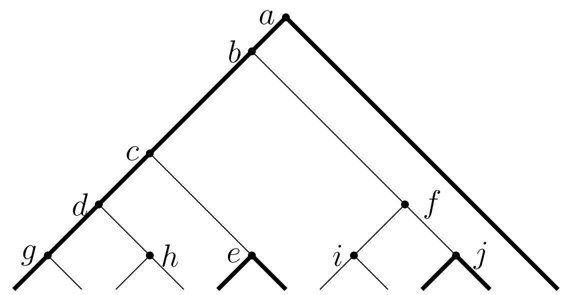
$$\overline{\mathfrak{P}_1} = (\mathfrak{P}_1 - (\{P\} \cup \mathfrak{P}_1^d \cup \mathfrak{P}_1^e)) \cup \{P'\} \cup \mathfrak{P}_2^d \cup \mathfrak{P}_2^e$$

realizes $\overline{V_1}$. Also, d is not blocked in $\overline{V_1}$, so $[\overline{\mathfrak{P}_1}]$ is maintaining.

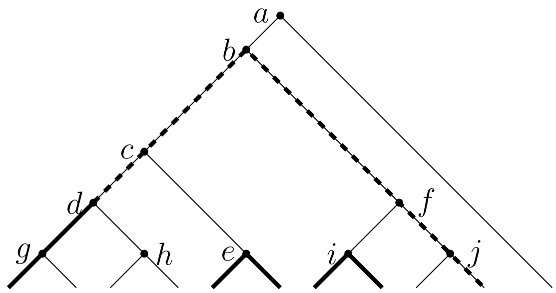
If there is no path in $\mathfrak{P}_2$ that contains edges cd or ce, then it is clear that $(\mathfrak{P}_2 - (\mathfrak{P}_2^d \cup \mathfrak{P}_2^e)) \cup \mathfrak{P}_1^d \cup \mathfrak{P}_2^d$ realizes $\overline{V_2}$. Otherwise, suppose that $Q \in \mathfrak{P}_2$ is a path that contains



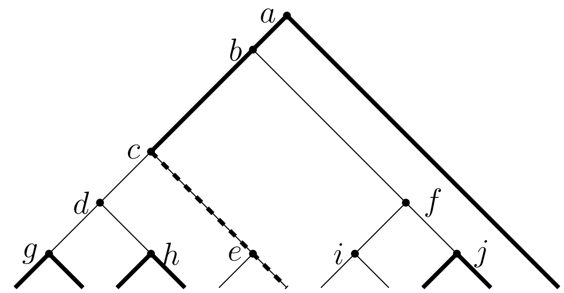
A b-nonmaintainging collection of paths realizing$V_1 = \{b, g, h, i\}$ with d blocked in V_1



A collection of paths realizing $V_2 = \{a, e, j\}$ with d not blocked in V_2



A b-maintainging collection of paths realizing $\overline{V_1} = \{b, e, i\}$ with d not blocked in $\overline{V_1}$



A collection of paths realizing $\overline{V_2} = \{a, g, h, j\}$

Figure 6.5: Proof of Lemma 6.16

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 vertex b, and let $\tilde{P}$ be the path contained in P from e to a leaf descended from e; this is guaranteed to exist since d is blocked in V_1. Let $Q' = \hat{Q} \cup \{ce\} \cup \tilde{P}$. Then

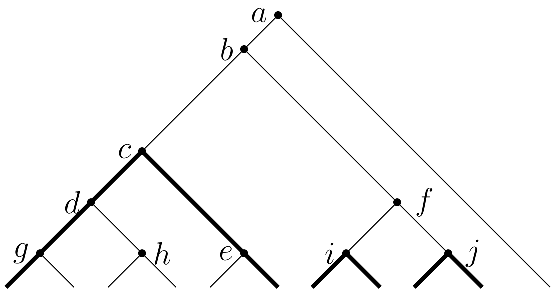
$$\overline{\mathfrak{P}_2} = (\mathfrak{P}_2 - (\{Q\} \cup \mathfrak{P}_2^d \cup \mathfrak{P}_2^e)) \cup \{Q'\} \cup \mathfrak{P}_1^d \cup \mathfrak{P}_1^e$$

realizes $\overline{V_2}$, as needed. Since b and $c \notin \overline{V_2}$, $[\overline{\mathfrak{P}_2}]$ is maintaining. This operation is illustrated for an example tree T and top-sets V_1 and V_2 in Figure 6.5.

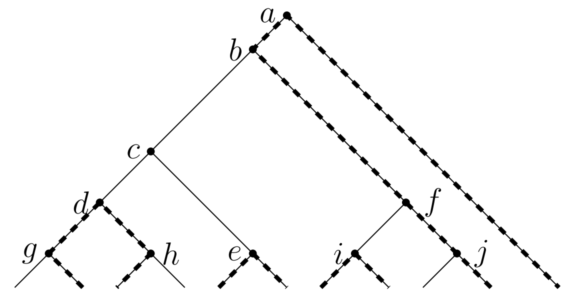
This operation preserves the number of times each interior node is a top-most vertex. So, $\mathbf{v} = [\overline{\mathfrak{P}_1}] + [\overline{\mathfrak{P}_2}] + [\mathfrak{P}_3] + \cdots + [\mathfrak{P}_m]$ is a representation of $\mathbf{v}$ using fewer nonmaintaining vertices, and $[\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$ is not minimal.

If $[\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$ is not f-compressed, we will proceed by a similar argument. Without loss of generality, we may assume that $[\mathfrak{P}_1]$ is c-nonmaintaining and that d 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 collections $\overline{\mathfrak{P}_1}$ and $\overline{\mathfrak{P}_2}$ respectively. Furthermore, we claim that $[\overline{\mathfrak{P}_1}]$ and $[\overline{\mathfrak{P}_2}]$ are maintaining.

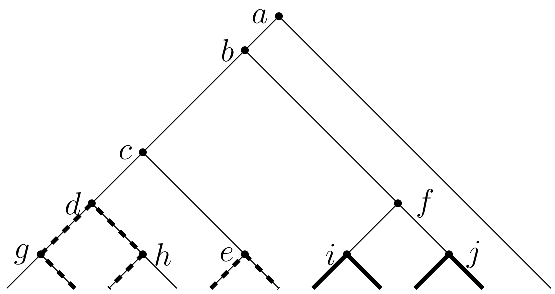
Since f is not blocked in V_2, we may assume that for all $P \in \mathfrak{P}_2$, $bc \notin P$. This is because $b \notin V_2$, and any $P \in \mathfrak{P}_2$ with top-most node above b may pass through the f-subtree instead of the c-subtree since f is not blocked in V_2. So, in both $\mathfrak{P}_1$ and $\mathfrak{P}_2$, all paths that intersect the c-subtree are contained entirely within the c-subtree. So



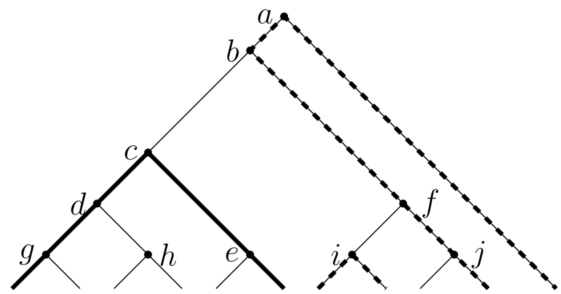
A collection of paths realizing
$V_1 = \{c, i, j\}$ where $[\mathfrak{P}_1]$ is c-nonmaintaining.



A collection of paths realizing
$V_2 = \{a, d, e, i\}$



A collection of paths realizing
$\overline{V_1} = \{d, e, i, j\}$



A collection of paths realizing
$\overline{V_2} = \{a, c, i\}$, where $[\overline{\mathfrak{P}_2}]$ is c-maintaining

Figure 6.6: Proof of Lemma 6.16, The first row of trees contain path collections – one whose top-set is c-nonmaintaining, and one without b or c in its top-set, but with f blocked in its top-set. The second row of trees are the path collections obtained by performing the operation in the proof of Lemma 6.16

$\overline{\mathfrak{P}_1} = (\mathfrak{P}_1 - \mathfrak{P}_1^c) \cup \mathfrak{P}_2^c$ and $\overline{\mathfrak{P}_2} = (\mathfrak{P}_2 - \mathfrak{P}_2^c) \cup \mathfrak{P}_1^c$ are collections of disjoint paths that realize $\overline{V_1}$ and $\overline{V_2}$, respectively. Since b and $c \notin \overline{V_1}$, $[\overline{\mathfrak{P}_1}]$ is maintaining. Furthermore, since f is not blocked in V_2, and since $\overline{V_2^d} = V_2^d$, f is not blocked in $\overline{V_2}$ and $[\overline{\mathfrak{P}_2}]$ is maintaining.

This operation preserves the number of times each interior node is used as a topmost node. So $\mathbf{v} = [\overline{\mathfrak{P}_1}] + [\overline{\mathfrak{P}_2}] + [\mathfrak{P}_3] + \cdots + [\mathfrak{P}_m]$ is a representation of $[\mathfrak{P}]$ using fewer c-nonmaintaining vertices. This operation is illustrated in Figure 6.6. $\square$

Lemma 6.17. *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 $\mathbf{v} = [\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$ be a df-compressed representation of $\mathbf{v}$ and let $\mathbf{u} = [\mathfrak{Q}_1] + \cdots + [\mathfrak{Q}_m]$ be any representation of $\mathbf{u}$.*

(i) If the multiset $\{[\mathfrak{Q}_1], \ldots, [\mathfrak{Q}_m]\}$ contains fewer b-nonmaintaining vertices than the multiset $\{[\mathfrak{P}_1], \ldots, [\mathfrak{P}_m]\}$, then $[\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$ is not a minimal representation of $\mathbf{v}$.

(ii) *If the multiset* $\{[\mathfrak{Q}_1], \ldots, [\mathfrak{Q}_m]\}$ *contains fewer* c*-nonmaintaining vertices than the multiset* $\{[\mathfrak{P}_1], \ldots, [\mathfrak{P}_m]\}$*, then* $[\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$ *is not a minimal representation of* $\mathbf{v}$.

Proof. To prove (i), suppose that $\{[\mathfrak{Q}_1], \ldots, [\mathfrak{Q}_m]\}$ contains fewer b-nonmaintaining vertices than $\{[\mathfrak{P}_1], \ldots, [\mathfrak{P}_m]\}$. 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 $\mathfrak{Q}_i$.

Since $[\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$ is d-compressed, for all V_i with $b, c \notin V_i$, d is blocked in V_i. Without loss of generality, let $[\mathfrak{P}_1], \ldots, [\mathfrak{P}_r]$ and $[\mathfrak{Q}_1], \ldots, [\mathfrak{Q}_{r'}]$ be the b-nonmaintaining vertices where $r' < r$. Let $[\mathfrak{P}_{r+1}], \ldots, [\mathfrak{P}_s]$ and $[\mathfrak{Q}_{r'+1}], \ldots, [\mathfrak{Q}_s]$ be the rest of the vertices with b or c coordinate equal to 1. Note that by assumption, there are the same number of these 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 $[\mathfrak{Q}_i]$, and the same number of c-nonmaintaining vertices as the V_i.

First, let $i \leq r'$. Then since $[\mathfrak{P}]_i$ and $[\mathfrak{Q}]_i$ are both b-nonmaintaining, d is blocked in both V_i and U_i. So $cd \notin \mathfrak{P}_i, \mathfrak{Q}_i$. Therefore, all paths in $\mathfrak{P}_i$ and $\mathfrak{Q}_i$ that intersect the d-subtree are contained entirely within the d-subtree. So $\overline{\mathfrak{P}_i} = (\mathfrak{P}_i - \mathfrak{P}_i^d) \cup \mathfrak{Q}_i^d$ is a collection of disjoint paths that realizes $\overline{V_i}$, as needed.

Next, let $r' < i \leq s$. Then $[\mathfrak{Q}]_i$ 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

$$\overline{P_i} = (\mathfrak{P}_i - (\{P\} \cup \mathfrak{P}_i^d)) \cup \{P'\} \cup \mathfrak{Q}_i^d$$

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 $\{[\overline{\mathfrak{P}_1}], \ldots, [\overline{\mathfrak{P}_s}]\}$ than in $\{[\mathfrak{P}_1], \ldots, [\mathfrak{P}_s]\}$. 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 collection of disjoint paths 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, $\overline{[\mathfrak{P}_1]} + \cdots + \overline{[\mathfrak{P}_m]}$ is a representation of $\mathbf{v}$ using fewer nonmaintaining vertices than $[\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$. So $[\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$ is not minimal. An example of the operation used to obtain $\overline{[\mathfrak{P}_1]}, \ldots, \overline{[\mathfrak{P}_m]}$ is illustrated in Figure 6.7.

To prove (ii), suppose that $\{[\mathfrak{Q}_1], \ldots, [\mathfrak{Q}_m]\}$ contains fewer c-nonmaintaining vertices than $\{[\mathfrak{P}_1], \ldots, [\mathfrak{P}_m]\}$. Then without loss of generality, let $[\mathfrak{P}_1]$ be c-nonmaintaining. Since $[\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$ is f-compressed, for all V_i with $b, c \notin V_i$, f is blocked in V_i. Without loss of generality, let $[\mathfrak{P}_1], \ldots, [\mathfrak{P}_r]$ and $[\mathfrak{Q}_1], \ldots, [\mathfrak{Q}_{r'}]$ be the c-nonmaintaining vertices where $r' < r$. Let $[\mathfrak{P}_{r+1}], \ldots, [\mathfrak{P}_s]$ and $[\mathfrak{Q}_{r'+1}], \ldots [\mathfrak{Q}_s]$ be the rest of the vertices with b or c coordinate equal to 1. Note that by assumption, there are the same number of these 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 collection of disjoint paths 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 from f to a leaf descended from f with no node along Q in U_i.

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

$$\overline{\mathfrak{P}_i} = (\mathfrak{P}_i - (\{P\} \cup \mathfrak{P}_i^f)) \cup \{P'\} \cup \mathfrak{Q}_i^f$$

realizes $\overline{V_i}$.

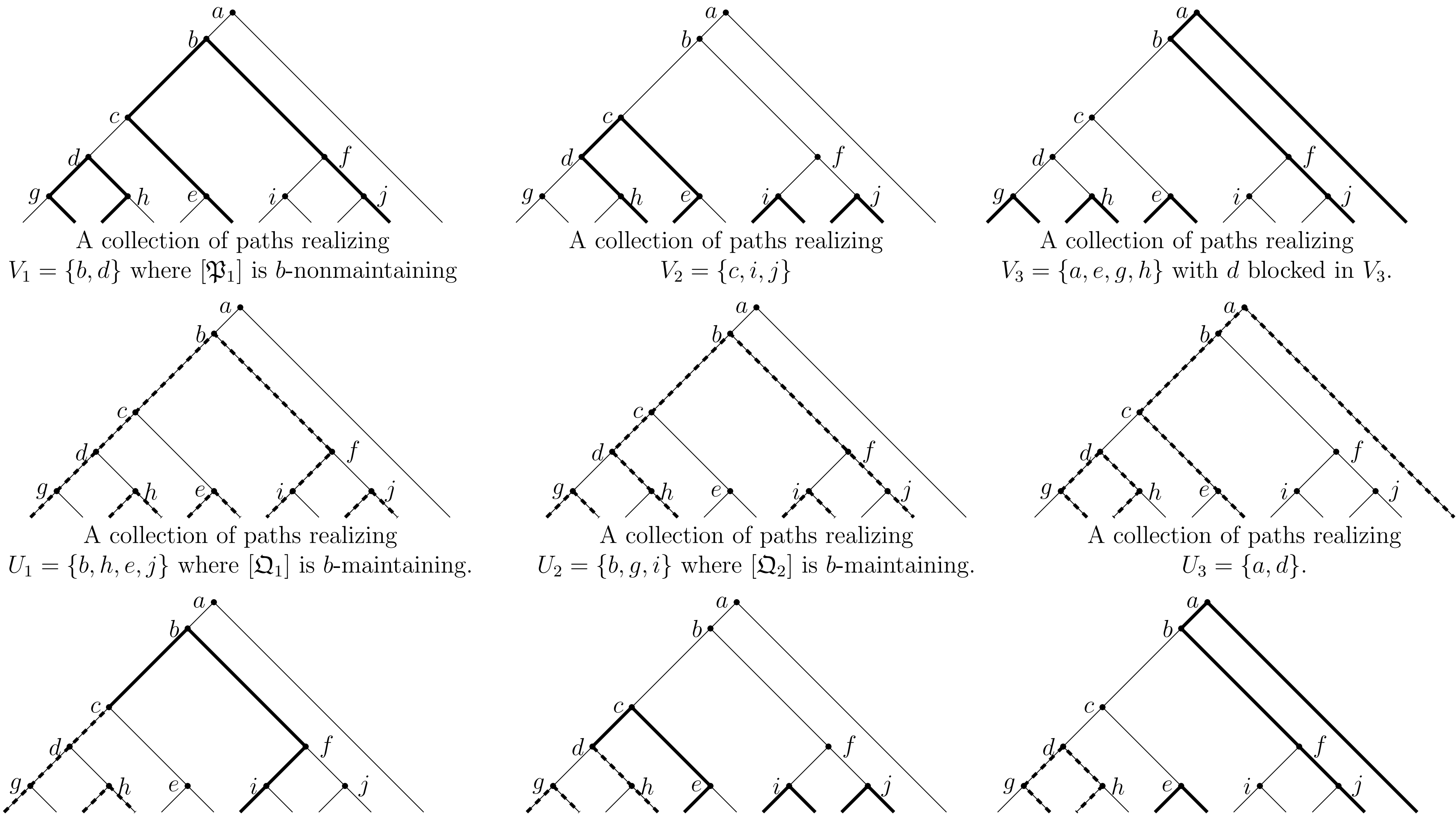
Now suppose that $c \in V_i$. If there does not exist $P \in \mathfrak{P}_i$ with a node above b as its top-most node that passes through the f-subtree, then all paths in $\mathfrak{P}_i$ that intersect the f-subtree are contained in the f-subtree. So $\overline{\mathfrak{P}_i} = (\mathfrak{P}_i - \mathfrak{P}_i^d) \cup \mathfrak{Q}_i^d$ is a collection of disjoint paths 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

$$\overline{\mathfrak{P}_i} = (\mathfrak{P}_i - (\{P\} \cup \mathfrak{P}_i^f)) \cup \{P\} \cup \mathfrak{Q}_i^f$$

is a collection of disjoint paths 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 $\{\overline{[\mathfrak{P}_1]}, \ldots, \overline{[\mathfrak{P}_s]}\}$ than in $\{[\mathfrak{P}_1], \ldots, [\mathfrak{P}_s]\}$. Furthermore, since the paths in the d-subtrees remain unchanged, this operation cannot create new b-nonmaintaining vertices.



A collection of paths realizing $V_1 = \{b, d\}$ where $[\mathfrak{P}_1]$ is b-nonmaintaining

A collection of paths realizing $V_2 = \{c, i, j\}$

A collection of paths realizing $V_3 = \{a, e, g, h\}$ with d blocked in V_3.

A collection of paths realizing $U_1 = \{b, h, e, j\}$ where $[\mathfrak{Q}_1]$ is b-maintaining.

A collection of paths realizing $U_2 = \{b, g, i\}$ where $[\mathfrak{Q}_2]$ is b-maintaining.

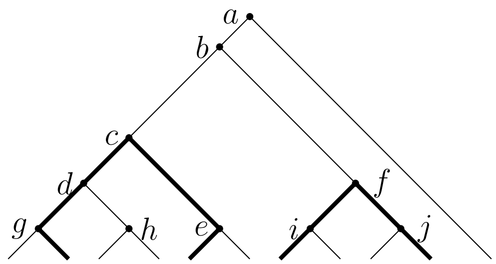
A collection of paths realizing $U_3 = \{a, d\}$.

A collection of paths realizing $\overline{V_1} = \{b, h\}$ where $[\overline{\mathfrak{P}_1}$ is b-maintaining.

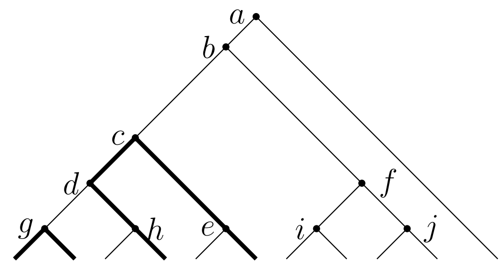
A collection of paths realizing $\overline{V_2} = \{c, g, i, j\}$

A collection of paths realizing $\overline{V_3} = \{a, d, e\}$

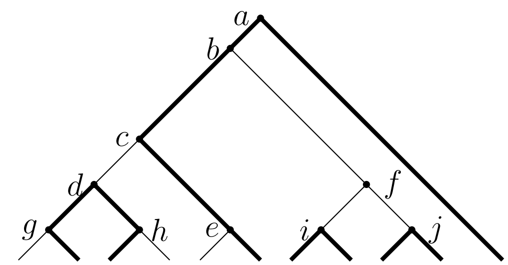
Figure 6.7: Proof of Lemma 6.17 (i). The first row of trees are path collections that realize some $\mathbf{v} = [\mathfrak{P}_1] + [\mathfrak{P}_2] + [\mathfrak{P}_3] \in 3R_T$. The second row of trees are path collections that realize $\mathbf{u} = [\mathfrak{Q}_1] + [\mathfrak{Q}_2] + [\mathfrak{Q}_3] \in 3R_T$ that satisfies the assumptions of the lemma. The third row of trees are a new set of path collections that realize $\mathbf{v}$ using fewer b-nonmaintaining vertices, which we obtained by applying the procedure discussed in the proof of the lemma.



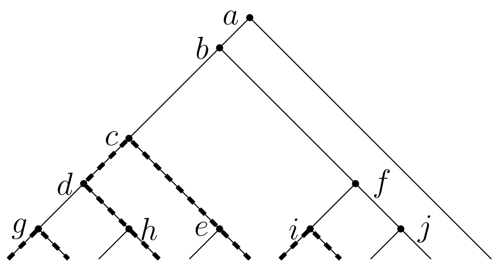
A collection of paths realizing $V_1 = \{c, f\}$ in which $[\mathfrak{P}_1]$ is c-nonmaintaining



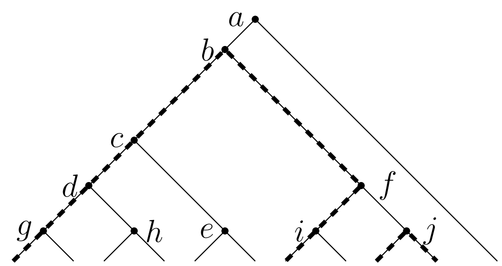
A collection of paths realizing $V_2 = \{c, g\}$ in which $[\mathfrak{P}_2]$ is c-maintaining



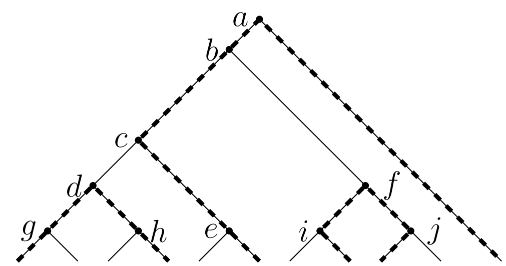
A collection of paths realizing $V_3 = \{a, d, i, j\}$ in which f is blocked



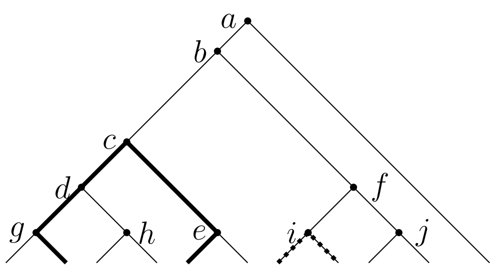
A collection of paths realizing $U_1 = \{c, g, i\}$ where $[\mathfrak{Q}_1]$ is c-maintaining



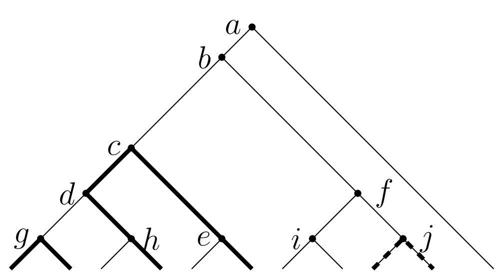
A collection of paths realizing $U_2 = \{b, j\}$ in which f is blocked



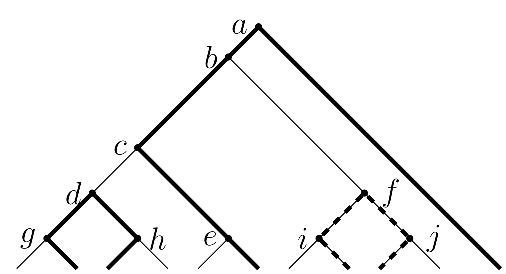
A collection of paths realizing $U_3 = \{a, d, f\}$



A collection of paths realizing $\overline{V_1} = \{c, i\}$ where $[\overline{\mathfrak{P}}_1]$ is c-maintaining



A collection of paths realizing $\overline{V_2} = \{c, g, j\}$ where $[\overline{\mathfrak{P}}_2]$ is c-maintaining



A collection of paths realizing $\overline{V_3} = \{a, d, f\}$

Figure 6.8: Proof of Lemma 6.17 (ii). The first row of trees are path collections that realize some $\mathbf{v} = [\mathfrak{P}_1] + [\mathfrak{P}_2] + [\mathfrak{P}_3] \in 3R_T$. The second row of trees are path collections 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 collections that realize $\mathbf{v}$ using fewer c-nonmaintaining vertices, which we obtained by applying the procedure discussed in the proof of the lemma.

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 $\overline{\mathfrak{P}} = (\mathfrak{P}_i - \mathfrak{P}_i^f) \cup \mathfrak{Q}_i^f$ is a collection of disjoint paths that realizes $\overline{V_i}$.

Since

$$\sum_{i=1}^{m} [\mathfrak{P}_i]^f = \sum_{i=1}^{m} [\mathfrak{Q}_i]^f,$$

and since $[\overline{\mathfrak{P}_i}]^f = [\mathfrak{Q}_i]^f$ for all i, $[\overline{\mathfrak{P}_1}] + \cdots + [\overline{\mathfrak{P}_m}]$ is a representation of $\mathbf{v}$ using fewer nonmaintaining vertices than $[\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$. So $[\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$ is not minimal. An example of the operation used to obtain $[\overline{\mathfrak{P}_1}], \ldots, [\overline{\mathfrak{P}_m}]$ is illustrated in Figure 6.8. $\square$

Corollary 6.18. *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 6.17 since if a minimal representation $\mathbf{v} = [\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$ uses more b-nonmaintaining vertices than another minimal representation $\mathbf{v} = [\mathfrak{Q}_1] + \cdots + [\mathfrak{Q}_m]$, then by the lemma, $[\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$ was not actually minimal. $\square$

Corollary 6.19. *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 $\mathbf{v} = [\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$ is a minimal representation, but $\phi^{T,T'}([\mathfrak{P}_1]) + \cdots + \phi^{T,T'}([\mathfrak{P}_m]) = \phi_m^{T,T'}(\mathbf{v})$ is not a minimal representation of $\phi_m^{T,T'}(\mathbf{v})$.

Let $[\mathfrak{Q}_1] + \cdots + [\mathfrak{Q}_m] = \phi_m^{T,T'}(\mathbf{v})$ be minimal. Consider the image $\phi_m^{T',T}(\phi_m^{T,T'}(\mathbf{v})) = \phi^{T',T}([\mathfrak{Q}_1]) + \cdots + \phi^{T',T}([\mathfrak{Q}_m])$. The set $\{\phi^{T',T}([\mathfrak{Q}_1]), \ldots, \phi^{T',T}([\mathfrak{Q}_m])\}$ contains fewer nonmaintaining vertices than $\{[\mathfrak{P}_1], \ldots, [\mathfrak{P}_m]\}$, and satisfies all of the assumptions of Lemma 6.17. So, $[\mathfrak{P}_1] + \cdots + [\mathfrak{P}_m]$ was not actually a minimal representation of $\mathbf{v}$ and we have reached a contradiction. $\square$

Theorem 6.20. *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 nearest neighbor interchanges. So, it follows from Corollary 6.19 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 6.2. The leading coefficient of the Ehrhart polynomial of a polytope is the (unnormalized) volume of the polytope; that is, it is the normalized volume of the polytope divided by the factorial of the dimension. 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 6.8. $\square$

Acknowledgments

Jane Coons was partially supported by the US National Science Foundation (DGE-1746939) and by the David and Lucille Packard Foundation. Seth Sullivant was partially supported by the US National Science Foundation (DMS 1615660) and by the David and Lucille Packard Foundation.

References

[1] Beck, M., and Robins, S. *Computing the continuous discretely*. Springer, 2007.

[2] 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.

[3] Buneman, P. A note on the metric properties of trees. *Journal of Combinatorial Theory, Series B 17*, 1 (1974), 48–50.

[4] Evans, S. N., and Speed, T. P. Invariants of some probability models used in phylogenetic inference. *The Annals of Statistics* (1993), 355–377.

[5] Felsenstein, J. *Inferring phylogenies*, vol. 2. Sinauer associates Sunderland, MA, 2004.

[6] Hendy, M. D., and Penny, D. Spectral analysis of phylogenetic data. *Journal of classification 10*, 1 (1993), 5–24.

[7] Jukes, T. H., Cantor, C. R., et al. Evolution of protein molecules. *Mammalian protein metabolism 3*, 21 (1969), 132.

[8] 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.

[9] Kubjas, K. *Algebraic and combinatorial aspects of group-based models*. PhD thesis, Freie Universität Berlin, 2013.

[10] Semple, C., and Steel, M. A. *Phylogenetics*, vol. 24. Oxford University Press on Demand, 2003.

[11] Simon, B. *Representations of finite and compact groups*. No. 10. American Mathematical Soc., 1996.

[12] Sloane, N. The online encyclopedia of integer sequences, 2005. *Published electronically at http://oeis. org*.

[13] STANLEY, R. P. Two poset polytopes. *Discrete & Computational Geometry 1*, 1 (1986), 9–23.

[14] STANLEY, R. P. A survey of alternating permuations. *Contemporary Mathematics 531* (2010), 165–196.

[15] STURMFELS, B. *Gröbner bases and convex polytopes*, vol. 8. American Mathematical Society, 1996.

[16] STURMFELS, B., AND SULLIVANT, S. Toric ideals of phylogenetic invariants. *Journal of Computational Biology 12*, 2 (2005), 204–228.

[17] STURMFELS, B., AND XU, Z. Sagbi bases of Cox-Nagata rings. *J. Eur. Math. Soc. (JEMS) 12*, 2 (2010), 429–459.

[18] SULLIVANT, S. Toric fiber products. *Journal of Algebra 316*, 2 (2007), 560–577.

[19] SULLIVANT, S. *Algebraic Statistics*. American Mathematical Society, 2018. in preparation.

[20] WEST, D. B. *Introduction to Graph Theory*, 2 ed. Pearson Education Inc., 2001.

[21] ZIEGLER, G. M. *Lectures on polytopes*, vol. 152 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995.