H-Representation of the Kimura-3 Polytope

Marie Mauhar\textsuperscript{a}, Joseph Rusinko\textsuperscript{b,*}, Zoe Vernon\textsuperscript{b}

\textsuperscript{a}Lenoir-Rhyne University
\textsuperscript{b}Winthrop University

Abstract
Given a group-based Markov model on a tree, one can compute the vertex representation of a polytope which describes the associated toric variety. The half-space representation, however, is not easily computable. In the case of $\mathbb{Z}_2$ or $\mathbb{Z}_2 \times \mathbb{Z}_2$, these polytopes have applications in the field of phylogenetics. We provide a half-space representation for the $m$-claw tree where $G = \mathbb{Z}_2 \times \mathbb{Z}_2$, which corresponds to the Kimura-3 model of evolution.

Keywords: Phylogenetics, Group-based models

1. Introduction

1.1. Phylogenetic Varieties

Phylogenetic trees depict evolutionary relationships between proteins, genes or organisms. Many tree reconstruction methods rely on an underlying model of evolution described as a Markov process along the edges of the tree. The probability of observing a particular site pattern can be computed as a polynomial in the model parameters of the transition matrices. Many biological models of evolution, including the Kimura-3 model, can be viewed as restrictions on the allowable structure of these underlying transition matrices. For an overview of this algebraic statistical viewpoint on phylogenetics we refer the reader to [1].

For a fixed tree and evolutionary model, invariants are polynomials that vanish on all joint distributions of bases at the leaves that arise from the model [2]. Algebraic geometry provides a framework for computing the complete set of phylogenetic invariants as the elements of the prime ideal that define the phylogenetic variety. Many classical varieties arise in the study of phylogenetic models [3] as do modern objects such as conformal blocks [4] and Berenstein and Zelevinsky triangles [5].

One such variety arises for many biological models of evolution where a finite group acts freely and transitively on the set of states in the transition matrix. For these group-based models the transition matrices can be simultaneously diagonalized [6]. After the change of coordinates induced by these diagonalizations the variety is seen to be toric [1, 7]. As

\*corresponding author

Email addresses: marie.mauhar@my.lr.edu (Marie Mauhar), rusinkoj@winthrop.edu (Joseph Rusinko), vernonz2@winthrop.edu (Zoe Vernon)

Preprint submitted to Elsevier September 16, 2014
toric varieties, there is an associated polytope that encapsulates many characteristics of the variety. We refer the reader to [8, 9, 10] for background on toric varieties and polytopes.

The work of [11] and [12] provide an algorithm for computing the vertices of this polytope for any tree and group. For the group \( \mathbb{Z}_2 \), or binary symmetric model of evolution, both the polytope and the toric variety are well understood [13]. Relatively little is known about the polytopes associated to other models. The main result of this paper is to provide the half-space or H-representation for the polytopes corresponding to the Kimura-3 parameter model of evolution, which corresponds to the group \( \mathbb{Z}_2 \times \mathbb{Z}_2 \).

1.2. Kimura-3 Polytopes and Varieties

The Kimura-3 model of evolution is the simplest model which accounts for a difference in mutation rates among transitions and both types of transversions [14]. All biologically meaningful points of the phylogenetic variety are smooth, and contained in a local complete intersection generated in degrees two and four [15]. Also, for any tree the projective scheme associated to the Kimura-3 model of that tree can be generated by an ideal generated in degree at most four [16].

The geometry of the Kimura-3 model is more complex than the binary symmetric model [17, 16]. The Kimura-3 varieties are known to be normal in the case of a trivalent tree [18]. However, this is unknown in the case of a multifurcating tree. In particular, this is unknown even in the case of a tree with one interior node and \( m \) leaves, a tree referred to as the \( m \)-claw tree, which will be the tree shape utilized in this article.

For group-based models, the variety associated to any phylogenetic tree can be computed by a sequence of toric fiber products of the variety associated to claw trees [7, 19]. For bifurcating trees, one need only use the three claw tree. However, to account for multifurcations one must understand the variety associated to an \( m \)-claw tree. See [18] for a broader survey of the relevance of claw trees.

While the vertex description of a polytope for a claw tree is straightforward, little more is known about the polytope. For instance, it is only known to be normal for the binary symmetric model [20, 18]. The binary symmetric model is also the only case for which a H-representation is known. Translating between the vertex and facet description of a polytope is an NP-complete problem in general [21] and is challenging even in classes of examples where a plausible set of facets can be proposed.

In this article we give a complete list of facet inequalities or the H-representation associated to the claw tree of the Kimura-3 model. This description builds off a well known identification of the claw tree polytope for the binary symmetric model with the demihypercube (a hypercube with alternating vertices removed).

1.3. Outline of paper

In Section 2 we review the vertex description of the polytope associated to the binary symmetric and Kimura-3 models. We introduce a polytope \( \Delta(m) \) in Section 3, which we propose as the H-representation for the polytope associated to the Kimura-3 model for an arbitrary claw tree. We show that if \( \Delta(m) \) is integral then it is the H-representation. We introduce a linear change of coordinates which transforms \( \Delta(m) \) into an isomorphic polytope
This polytope is described in terms of a $3 \times m$ matrix whose row and column coordinates satisfy a set of inequalities reminiscent of those that define the demihypercube. In Section 5 we identify a connection between the number of integral coordinates in rows of this matrix and the number of facets on which the point lies. This connection is utilized in Section 6 to prove the main result of our paper, that $\Delta(m)$ is an integral polytope, and thus provides an H-representation of the Kimura-3 polytope associated to a claw tree. We conclude by listing a collection of open problems we feel may help extend the rapidly growing body of knowledge about the combinatorics and geometry of group-based models.

2. Background

A polytope can be represented as the convex hull of a finite set of points in $\mathbb{R}^n$ (V-representation) or the set of points satisfying a set of linear inequalities defining its facets (H-representation). For a fixed tree and group, the V-representation of the polytope associated to a group-based model can be computed using the algorithm described by [11, 12]. For the $m$-claw tree with a group $G$, the vertices of the polytope are in a bijection with collections of $m$ elements of $G$ which sum to the identity. Throughout this paper we assume that $m \geq 3$ as smaller values of $m$ are simple to describe geometrically, and have no biological meaning.

For each claw we define a bijection from $G \to \mathbb{Z}^{|G|-1}$ where the identity is identified with $(0,0,\cdots,0)$ and each non-identity element with a unit basis vector for of the integral lattice $\mathbb{Z}^{|G|-1}$. The polytope associated to the $m$-claw tree is the convex hull in $\mathbb{R}^{|G|-1}$ of the image collections of group elements which sum to the identity under this bijection.

As a motivating example we review the construction for the binary symmetric model where $G = \mathbb{Z}_2$ was described in [13]. For the claw tree, this polytope is the demihypercube.

**Definition 1.** The demihypercube, $DH(m)$, is the $m$-dimensional polytope constructed by implementing the process of alternation on the vertices of the $m$-dimensional-hypercube $[0,1]^m$, and taking the convex hull of the remaining vertices. [22].

The following is a well known H-representation for the demihypercube.

**Definition 2.** The H-representation of the demihypercube is given by

$$DH(m) = x \in [0,1]^m : \sum_{i \in A} x_i \leq |A| - 1 + \sum_{j \notin A} x_j,$$

where $A$ ranges over all subsets of $\{1,2,\cdots,m\}$ of odd cardinality.

The inequalities defining the demihypercube will be instrumental in the construction of the H-representation of the Kimura-3 polytope.

The transition matrices of the Kimura-3 model of evolution are acted on by the group $\mathbb{Z}_2 \times \mathbb{Z}_2$. Thus the associated polytope $K(m)$ lives in $\mathbb{R}^{3m}$, which we view as $m$ triplets $(a_i, b_i, c_i)$ for $1 \leq i \leq m$. When reference to a particular triplet is not needed we refer to the coordinates as $x_j$ for $1 \leq j \leq 3m$. 

3
Theorem 2.1 ([7] as described in [12]). Vertices of $K(m)$ are in bijection with collections of $m$ elements of $\mathbb{Z}_2 \times \mathbb{Z}_2$ such that the sum of these elements is the neutral element in the group. If $C = \{g_1, g_2, \cdots, g_m\}$ is such a collection, we define a function $f : (\mathbb{Z}_2 \times \mathbb{Z}_2)^m \to \mathbb{R}^{3m}$ which sends $g_i$ to coordinates $(a_i, b_i, c_i)$ as described in the following table:

| $g_i$   | $a_i$ | $b_i$ | $c_i$ |
|---------|-------|-------|-------|
| $(0,0)$ | 0     | 0     | 0     |
| $(0,1)$ | 0     | 1     | 0     |
| $(1,1)$ | 0     | 0     | 1     |

In developing an H-representation for $K(m)$ we are motivated by the observation that an element of $\mathbb{Z}_2 \times \mathbb{Z}_2$ is the identity if and only if its image is the identity under all homomorphisms from $\mathbb{Z}_2 \times \mathbb{Z}_2$ to $\mathbb{Z}_2$.

3. H-Representation for Kimura-3 Polytope: $\Delta(m)$

The homomorphisms suggest a connection between the polytopes for Kimura-3 models and those for the binary symmetric models, where $G = \mathbb{Z}_2$. The main result of this article is to confirm this connection, by showing the following set of inequalities provides an H-representation for the Kimura-3 polytope.

Definition 3. The polytope $\Delta(m)$ is the set of points $x \in \mathbb{R}^{3m}$ satisfying $x_j \geq 0$ for all $1 \leq j \leq 3m$, and $a_i + b_i + c_i \leq 1$ for all $1 \leq i \leq m$, and the following collection of $A$-inequalities:

$$\sum_{i \in A}(a_i + b_i) \leq |A| - 1 + \sum_{j \notin A}(a_j + b_j),$$

$$\sum_{i \in A}(a_i + c_i) \leq |A| - 1 + \sum_{j \notin A}(a_j + c_j), \text{ and}$$

$$\sum_{i \in A}(b_i + c_i) \leq |A| - 1 + \sum_{j \notin A}(b_j + c_j).$$

Where $A$ ranges over all cardinality subsets of $I = \{1, 2, 3, \cdots, m\}$.

We first demonstrate that $\Delta(m)$ contains $K(m)$.

Theorem 3.1. The polytope $K(m)$ is a subset of $\Delta(m)$.

Proof. Let $v$ be a vertex of $K(m)$ corresponding to a choice $(g_1, g_2, \cdots, g_m)$ of elements of $\mathbb{Z}_2 \times \mathbb{Z}_2$ which sum to the identity. By Theorem 2.1 $v$ satisfies $x_j \geq 0$ and $a_i + b_i + c_i \leq 1$.

Without loss of generality we prove that $v$ satisfies

$$\sum_{i \in A}(a_i + b_i) \leq |A| - 1 + \sum_{j \notin A}(a_j + b_j).$$


Letting \( X = \{ l \in A | g_l = (1, 0) \} \) and \( Y = \{ l \in A | g_l = (0, 1) \} \) we have \( |X| + |Y| \leq |A| \). We divide the proof into two cases according to the parity of \(|X| \) and \(|Y|\).

If \(|X| \) and \(|Y|\) are of the same parity, then since \(|A|\) is odd, we have \(|X| + |Y| < |A|\). This implies

\[
\sum_{i \in A} (a_i + b_i) \leq |A| - 1 \\
\leq |A| - 1 + \sum_{j \notin A} (a_j + b_j).
\]

If \(|X| \) and \(|Y|\) are of opposite parity, then without loss of generality we assume \(|X|\) is odd, which implies

\[
\sum_{i \in X \cup Y} g_i = (1, 0).
\]

In order to neutralize \((1, 0)\) in \( \mathbb{Z}_2 \times \mathbb{Z}_2 \) we must either add \((1, 0)\), or add both \((1, 1)\) and \((0, 1)\) from an element with index in \( A' \). In either case there must exists an \( l \in A' \) such that \( a_l + b_l = 1 \). Therefore,

\[
\sum_{i \in A} (a_i + b_i) \leq |A| - 1 + \sum_{j \notin A} (a_j + b_j).
\]

Since \( v \) satisfies all of the defining inequalities of \( \Delta(m) \), we have \( v \in \Delta(m) \). Containment of \( K(m) \) in \( \Delta(m) \) follows from the convexity of \( K(m) \).

The opposite inclusion that \( \Delta(m) \subseteq K(m) \) is not immediately clear. However, if \( \Delta(m) \) is integral then it is the H-representation of \( K(m) \).

**Theorem 3.2.** If \( \Delta(m) \) is integral, then \( \Delta(m) = K(m) \).

**Proof.** Let \( v \) be an integral vertex of \( \Delta(m) \). Using Theorem 2.1 we identify \( v \) with a sequence of group elements \( g_1, \ldots, g_m \). If \( v \) is not a vertex of \( K(m) \), then \( \sum_{i=1}^m g_i \neq (0, 0) \).

Where possible, we pair nontrivial \( g_i \) with an identical group element \( g_j \). Since the sum is not the identity there must be one remaining nontrivial element \( g_l \) or a pair of distinct nontrivial elements \( g_l, g'_l \). Without loss of generality we assume \( g_l = (1, 0) \) and, if an additional remaining element exists, that \( g'_l = (0, 1) \).

It follows that \( A = \{ i \in I | g_i = (1, 0) \text{ or } (1, 1) \} \) has odd cardinality. By construction we have \( \sum_{i \in A} a_i + c_i = |A| \) and \( \sum_{j \notin A} a_i + c_i = 0 \). This contradicts the hypothesis that \( v \) was an element of \( \Delta(m) \) since \( v \) does not satisfy

\[
\sum_{i \in A} (a_i + c_i) \leq |A| - 1 + \sum_{j \notin A} (a_j + c_j).
\]

By convexity this shows that if \( \Delta(m) \) is integral, then it is a subset of \( K(m) \). Combining this with Theorem 3.1 yields the equality of polytopes. \( \square \)
While having a tentative facet description pending the integrality of a polytope seems promising, we have simply exchanged the NP-complete problem of switching from a V-representation to an H-representation with another NP complete problem: recognizing when a rational polytope is integral \[23\].

As \( \Delta(m) \) is not totally uni-modular, the standard strategy for proving integrality will not work. However, the relationship between the \( A \)-inequalities of \( \Delta(m) \) and those of the demihypercube suggest a change of coordinates which allow us to demonstrate integrality.

4. Coordinate Change from \( \Delta(m) \) to \( \Delta'(m) \)

The \( A \)-inequalities for both the demihypercube and \( \Delta(m) \) share the same underlying structure. However, a more subtle connection with the demihypercube is revealed only after a change of coordinates motivated by the group homomorphisms from \( \mathbb{Z}_2 \times \mathbb{Z}_2 \rightarrow \mathbb{Z}_2 \). This change of coordinates is best understood by viewing the image as coordinates in a matrix.

**Definition 4.** Define a map \( f : \mathbb{R}^m \rightarrow \mathbb{R}^m \) by \( f(x) = M \), where \( M \) is a \( 3 \times m \) matrix with \( M_{1,i} = a_i + b_i, \ M_{2,i} = a_i + c_i \) and \( M_{3,i} = b_i + c_i \).

**Definition 5.** Define the polytope \( \Delta'(m) = f(\Delta(m)) \). Applying the change of coordinates to the inequalities defining \( \Delta(m) \), the facets of \( \Delta'(m) \) are given by:

\[
\sum_{i \in A} (x_{i,j}) \leq |A| - 1 + \sum_{i \not\in A} x_{i,j} \quad \text{(row inequalities)},
\]

where \( A \) is any subset of \( \{1, 2, \cdots m\} \) of odd cardinality and \( 1 \leq i \leq 3 \), and

\[
\sum_{j \in B} (x_{i,j}) \leq |B| - 1 + \sum_{j \not\in B} x_{i,j} \quad \text{(column inequalities)}
\]

where \( B \) is a subset of \( \{1, 2, 3\} \) of odd cardinality and \( 1 \leq j \leq m \).

We refer to the facets of \( \Delta'(m) \) defined by the row (resp. column) inequalities as row facets (resp. column facets). The following computations show \( \Delta(m) \) and \( \Delta'(m) \) are isomorphic as lattice polytopes.

**Theorem 4.1.** Let \( f_i : \mathbb{R}^3 \rightarrow \mathbb{R}^3 \) be the restriction of \( f \) to a map from the coordinates \((a_i, b_i, c_i) \rightarrow M|_{x_1, x_2, x_3} \). Then \( f_i \) is an isomorphism which maps unit 3-simplex in \( \mathbb{R}^3 \) to \( DH(3) \).

**Proof.** The map \( f_i : \mathbb{R}^3 \rightarrow \mathbb{R}^3 \) is given by \( f_i(a_i, b_i, c_i) = (a_i + b_i, a_i + c_i, b_i + c_i) \). The function \( f_i \) is an isomorphism with inverse \( f_i^{-1}(x, y, z) = (\frac{x+y-z}{2}, \frac{x+z-y}{2}, \frac{y+z-x}{2}) \). The second part of the claim follows from applying the change of coordinates \( f_i \) to the vertices of the 3-simplex \( \{(0,0,0),(1,0,0),(0,1,0),(0,0,1)\} \).

**Corollary 4.2.** The polytopes \( \Delta(m) \) and \( \Delta'(m) \) are isomorphic.

**Corollary 4.3.** The polytope \( \Delta(m) \) is integral if and only if \( \Delta'(m) \) is integral.

Before proving the integrality of \( \Delta'(m) \) in Section 6, we must first examine the facet structure of \( \Delta'(m) \). Given the local nature of the defining inequalities we rely on projections of \( \Delta'(m) \) onto individual rows for our analysis in Section 5.
5. Subpolytopes of $\Delta'(m)$

5.1. Pseudo-Demihypercubes

In this section we examine the connection between the number of integer coordinates in a row (resp. column) of $M$, and the number row (resp. column) facets the point in lies on.

Let $P|_r$ be the projection of $P$ into $\mathbb{R}^m$ corresponding to particular row of the matrix coordinates. We introduce the notion of a pseudo-demihypercube to describe the projection of $\Delta'(m)$ onto a particular row space.

**Definition 6.** For any row $r$, the pseudo-demihypercube is defined by

$$ PDH(m) = \{ P|_r | P \in \Delta'(m) \}. $$

The hyperplanes corresponding to the $A$-inequalities define pseudo-facets of the polytope.

We do not assume that $PDH(m) = DH(m)$ as there may be additional restrictions from the interactions between the row and column inequalities. We note that coordinates of $PDH(m)$ must lie in $[0,1]$, which indicates that pseudo-facets form a proper subset of the facets of $PDH(m)$.

5.2. Relationship between Facets and Integrality of Coordinates in the Pseudo-demihypercube

In this subsection we show that points of $PDH(m)$ that lie on more pseudo-facets are forced to contain more integral coordinates.

**Theorem 5.1.** If a point $P$ lies on two pseudo-facets of $PDH(m)$, then at most two coordinates of $P$ are non-integral.

**Proof.** Assume $P$ is a point in $PDH(m)$ which lies on facets $H_A$ and $H_B$ corresponding to sets $A$ and $B$. We label $x_i$ for the coordinate of $P$ with indexed by $i \in \{1, 2, \ldots, m\}$.

Then

$$ H_A : \sum_{i \in (A \setminus B)} x_i + \sum_{j \in (A \cap B)} x_j = |A \setminus B| + |A \cap B| - 1 + \sum_{k \in (B \setminus A)} x_k + \sum_{l \in (A \cup B)'} x_l, $$

and

$$ H_B : \sum_{k \in (B \setminus A)} x_k + \sum_{j \in (A \cap B)} x_j = |B \setminus A| + |A \cap B| - 1 + \sum_{i \in (A \setminus B)} x_i + \sum_{l \in (A \cup B)'} x_l. $$

Combining these two conditions gives the following equation:

$$ 2 \sum_{j \in (A \cap B)} x_j = |A \setminus B| + |B \setminus A| + 2 |A \cap B| - 2 + 2 \sum_{l \in (A \cup B)'} x_l. $$

Since $A$ and $B$ are distinct sets of odd cardinality, it follows that $|A \setminus B| + |B \setminus A| \geq 2$. In order for the above equation to hold for a point $P \in PDH(m)$, $x_j = 1$ for all $j \in A \cap B$, $x_l = 0$ for all $l \in (A \cup B)'$, and $|A \setminus B| + |B \setminus A| = 2$. Consequently there are at most two non-integer coordinates, and they would have to be indexed by the elements in $(A \setminus B) \cup (B \setminus A)$. \qed
If instead, $P$ lies on three pseudo-facets, not only must it be integral, it must lie on additional pseudo-facets as well.

**Theorem 5.2.** Let $P$ be a point in $PDH(m)$. If $P$ lies on three pseudo-facets, then $P$ is integral and lies on $m$ pseudo-facets.

**Proof.** Assume $P$ lies on pseudo-facets of $PDH(M)$ corresponding to odd cardinality sets $A$, $B$ and $C$. Repeated application of Theorem 5.1 yield:

$$x_i = \begin{cases} 1 & i \in \{A \cap B\} \cup \{A \cap C\} \cup \{B \cap C\} \\ 0 & \text{otherwise} \end{cases}$$

The integral point $P$ cannot have an odd number of coordinates with the value one, or it would not satisfy the $A$-inequality where $I = \{i | x_i = 1\}$.

Let $I = \{i_1, i_2, \ldots, i_{2k}\}$ denote the indices of coordinates of $P$ where $x_i = 1$, and $J = \{j_{2k+1}, \ldots, j_m\}$ denote the indices of coordinates where $x_j = 0$. One checks that $P$ lies on the $2k$ pseudo-facets corresponding the set $S_l = I \setminus \{i_l\}$ for $1 \leq l \leq 2k$, and $T_n = I \cup \{j_n\}$ for $2k + 1 \leq n \leq m$.

We conclude our discussion of the relationship between the pseudo-facets and integrality, by noting that no point can have a single non-integer coordinate.

**Theorem 5.3.** Every non-integral point $P$ of $PDH(m)$ has at least two non-integer coordinates. This equality is met only when $P$ lies on exactly two pseudo-facets.

**Proof.** Let $P$ be a non-integral point of $PDH(m)$, and let $I = \{i_1, i_2, \ldots, i_{2l}\}$ denote the indices of coordinates of $P$ where $x_i = 1$. The proof that there is a second non-integer coordinate is broken into cases based on the parity $|I|$. Assume $P$ has a single non-integer coordinate $y_j$.

If $|I|$ is odd then the $A$ inequality corresponding to $I$ is $|I| \leq |I| - 1 + y$. It follows that $y \geq 1$ which contradicts the existence of a non-integer coordinate $P \in PDH(m)$.

If $|I|$ is even then let $A = I \cup \{j\}$. Then the $A$ inequalities is $|I| + y \leq |I|$ so $y \leq 0$. This contradicts $y_j$ being a non-integer coordinate of $P \in PDH(m)$. Consequently $P$ can not have a single non-integer coordinate.

The fact that the equality is only met when $P$ lies on exactly two faces follows from the proofs of Theorem 5.1 and Theorem 5.2.

### 5.3. Pseudo-Facet Classification

In addition to knowing how the pseudo-facet structure constrains integrality, the proof of the integrality of $\Delta'(m)$ requires an additional classification of the pseudo-facets of $PDH(m)$ in terms of pairs of non-integral coordinates.

**Definition 7.** Let $P$ be a point in $PDH(m)$ with exactly two non-integral coordinates, $y_i$ and $y_j$, which lies on a pseudo-facet $H$ corresponding to a set $A$. If $i$ and $j$ are both elements of $A$ or both elements of $A'$ we label $H$ a same facet or $S$-facet. Otherwise we label $H$ an opposite or $O$-facet.
A facet’s S-O classification is a function of a particular pair of non-integer coordinates. We suppress the function notation since we always use a fixed point and pair of coordinates.

**Theorem 5.4.** Let $P$ be a point in $PDH(m)$ with exactly two non-integral coordinates, $y_1$ and $y_2$, which lies on a pseudo-facet $H$. Let $I = \{i|x_i = 1\}$. If $H$ is a S-facet, then

$$|I| \begin{cases} 	ext{odd} & \text{if } H \text{ is a S-facet} \\ 	ext{even} & \text{if } H \text{ is an O-facet.} \end{cases}$$

**Proof.** Let $P \in PDH(m)$ have exactly two non-integral coordinates $y_1$ and $y_2$, that lie on a pseudo-facet $H$ with index set $A$.

$$H : \sum_{i \in A} x_i \leq |A| - 1 + \sum_{j \notin A} x_j,$$

**Case 1.** If $H$ is a S-facet, then to satisfy $H$, we must have $y_1 + y_2 \in \mathbb{Z}$. Since $\{y_1, y_2\} \in (0, 1)$, we must have $y_1 + y_2 = 1$.

Now assume $\{y_1, y_2\} \in A$, then following equation represents $H$:

$$y_1 + y_2 + |A \cap I| = |A| - 1 + |A' \cap I|.$$

Since $|A \cap I| \leq |A| - 2$, this equation can only be satisfied when $|A \cap I| = |A| - 2$ and $A' \cap I = \emptyset$. This means $|I| = |A| - 2$, which is odd because $|A|$ is odd.

If we assume $\{y_1, y_2\} \notin A$, we obtain the following equation representing $H$:

$$|A \cap I| = |A| - 1 + y_1 + y_2 + |A' \cap I|,$$

which reduces to

$$|A \cap I| = |A| + |A' \cap I|.$$

This equation can only be satisfied when $I = A$ which forces $|I|$ to be odd.

**Case 2.** If $H$ is an O-facet, then we must have $y_1 - y_2 \in \mathbb{Z}$. Since $y_1, y_2 \in \{0, 1\}$, we this forces have $y_1 = y_2$. After canceling the non-integral coordinates, the equation for defining $H$ reduces to,

$$y_1 + |A \cap I| = |A| - 1 + y_2 + |A' \cap I|.$$

Since $|A \cap I| \leq |A| - 1$ this equation can only be satisfied if $|A \cap I| = |A| - 1$ and $A' \cap I = \emptyset$. This demonstrates that $|I| = |A| - 1$, and is therefore odd.

\[\square\]

**Corollary 5.5.** If $P \in PDH(m)$ has exactly two non-integral coordinates, then it cannot lie on both an O-facet and an S-facet.

In Section 6, we use the relationship relationships developed in this section to prove the integrality of $\Delta'(m)$. To do this, we ensure that all theorems from this section also apply to $DM(3)$, replacing pseudo-facet with facet when needed.
6. Integrality of $\Delta'(m)$

The properties of the pseudo-demi-hypercubes discussed in Section 5.2 will be used to show that $\Delta'(m)$ is an integral polytope. We will prove integrality by demonstrating there exists an open interval containing any non-integral point in $\Delta'(m)$, therefore showing a non-integral point cannot be a vertex in $\Delta'(m)$.

**Definition 8.** A point $P$ is in the interior of $\Delta'(m)$ if there exists a vector $v \in \mathbb{R}^{3m}$ and an $\epsilon > 0$ such that $P + \lambda v \in \Delta'(m)$ whenever $\lambda \in (-\epsilon, \epsilon)$.

The proof of integrality utilizes the following three lemmas which serve as tools for demonstrating that $P$ is in the interior of $\Delta'(m)$

**Lemma 6.1.** Let $P \in \Delta'(m)$, $k$ be the number of non-integral coordinates in $P$, and $\omega(P)$ the sum of the number of non-integral rows and columns. If $k$ is greater than $\omega(P)$, then $P$ is in the interior of $\Delta'(m)$.

**Proof.** Let $P$ be a point in $\Delta'(m)$ such that $k > \omega(P)$. We construct a vector $v$, and $\epsilon > 0$ such that $P + \lambda v \in \Delta'(m)$ for all $\lambda \in (-\epsilon, \epsilon)$.

If $x_{i,j}$ is an integer coordinate of $P$, then we set $v_{i,j} = 0$. We setup a linear system of equations to solve for the remaining coordinates of $v$. For convenience, we linearly reorder the coordinates of $P$ such that $x_1, x_2, \ldots, x_k$ are non-integral.

For each row or column containing non-integral coordinates, $P$ may lie on zero, one or two of the associated facets. If it lies on zero, then there exist an $\epsilon$ such that $P + \lambda v$ will satisfy the corresponding inequalities for any choice of $v$.

If $P$ lies on a single facet $H$ corresponding to a particular row or column, this introduces a single homogeneous linear relation on $v_1, v_2, \ldots, v_k$ by assuming the integral coordinates and constant term in the equation defining $H$ must remain fixed. Additionally, if $P$ lies on two facets in a particular row or column, by Theorem 5.4 $v_1, \ldots, v_k$ must satisfy a single linear homogeneous relationship.

Therefore we get a system of up to $\omega(P)$ homogeneous linear equations in $k$ unknowns. By the hypothesis if $k > \omega(P)$ there exists a nontrivial solution for $v$. Then $\lambda v + P \in \Delta'(m)$ for all $\lambda \in (-\epsilon, \epsilon)$ where $\epsilon$ is the minimum of the smallest $x_i$ and one minus the largest $x_i$ for $1 \leq i \leq k$.

In order to apply Lemma 6.1, we first ensure that there are points $P \in \Delta'(m)$ which meet the hypothesis. To help do this, we compare the number of non-integral coordinates with the number of row and columns which contain non-integral coordinates. For any row or column, $RC$, we define $n_{RC}(P)$ to be the number of non-integral coordinates of $P$ in $RC$.

**Lemma 6.2.** If $P \in \Delta'(m)$ and there exists a row or column for which $P$ lies on fewer than two facets, then $k(P) > \omega(P)$.
\[ P_1 = \begin{pmatrix} x_{1,1} & x_{1,2} & x_{1,3} & \cdots & x_{1,m} \\ x_{2,1} & x_{2,2} & x_{2,3} & \cdots & x_{2,m} \\ x_{3,1} & x_{3,2} & x_{3,3} & \cdots & x_{3,m} \end{pmatrix}, \quad P_2 = \begin{pmatrix} x_{1,1} & x_{1,2} & x_{1,3} & \cdots & x_{1,m} \\ x_{2,1} & x_{2,2} & x_{2,3} & \cdots & x_{2,m} \\ x_{3,1} & x_{3,2} & x_{3,3} & \cdots & x_{3,m} \end{pmatrix} \]

Figure 1: Configurations of non-integral coordinates for Lemma 6.3. Coordinates in bold are the only non-integral coordinates in the \( M \).

**Proof.** Let \( P \) be a non-integral point of \( \Delta'(m) \). Since each non-integral coordinate is on exactly one row and one column we have

\[ k(P) = \frac{1}{2} \sum_{RC_i \in M} n_{RC_i}(P). \]

It follows from Theorem 5.3 that \( n_{RC}(P) \geq 2 \) with equality met only if \( P \) lies on at least two \( RC \)-facets. By substitution, this yields \( k \geq \omega(P) \) with equality met only if \( P \) lies on at least two \( RC \)-facets. The hypothesis states that there exists a row or column for which \( P \) does not lie on more than two \( RC \)-facets, therefore we have \( k(P) > \omega(P) \). \( \square \)

If we assume every non-integral coordinate is on a row and column for which \( P \) lies on exactly two facets, then the hypothesis in Lemma 6.2 will not be met. However, up to reordering of the rows and columns, only the two configurations of non-integral coordinates, shown in Figure 1, allow for \( P \) to meet this condition. We explicitly construct an interval containing \( P \) in these cases.

**Lemma 6.3.** If \( P \in \Delta'(m) \) has one of the configurations of non-integral coordinates shown in Figure 1 then \( P \) is an interior point of \( \Delta'(m) \).

**Proof.** Given a point \( P \in \Delta'(m) \) with a configuration of non-integral coordinates as displayed in Figure 1 we demonstrate that \( P \) is in the interior of \( \Delta'(m) \).

In these two cases each row and column has at most two non-integral coordinates, so we use the S and O-facet information to assist in the construction. Let \( |S| \) be the number of S-facets that \( P \) lies on, and \( |O| \) be the number of O-facets that \( P \) lies on. Let \( I = \{ i | x_{i,j} = 1 \} \).

We note we can compute \( |I| \) by taking half of the number of coordinates with value one in each row, plus half of the number of coordinates with value one in each column.

By Theorem 5.4 every S-column and S-row contains an odd number of coordinates with value one, while every O-row and O-column must contain an even number of coordinates with value one. Therefore since \( |I| \) is an integer, we know that the number of S-facets must be even.

For each configuration in Figure 1 there exists a Hamiltonian cycle:

\[ (x_1, x_2, \cdots, x_k) = x_{1,1}, x_{2,1}, \cdots, x_{1,1} \]

in the graph with vertices corresponding to non-integral coordinates, and edges connecting non-integral coordinates in the same row or column.
In the case of \( P_1 \) (resp. \( P_2 \)) we reorder the coordinates of \( v = (v_1, v_2, v_3, v_4, \ldots, v_{3m}) \) (resp. \( v = (v_1, \ldots, v_6, \ldots, v_{3m}) \)) with the first 4 (resp. 6) coordinates corresponding to the non-integral coordinates of \( P_1 \) (resp. \( P_2 \)) in the order of the Hamiltonian cycle.

Using the reordered coordinates we construct a vector \( v \) as follows. First set \( v_1 = 1 \), and for \( v_1 \) through \( v_4 \) (resp. \( v_6 \)) assign:

\[
v_{i+1} = \begin{cases} v_i & \text{if } x_i \text{ and } x_{i+1} \text{ lie on an O-facet} \\ -v_i & \text{if } x_i \text{ and } x_{i+1} \text{ lie on an S-facet} \end{cases}
\]

Such a collection is consistent since \( P \) is on an even number of S-facets. We set \( v_i = 0 \) for all remaining coordinates.

We let \( \epsilon \) be the minimum of the smallest non-integral coordinate of \( P \) and one minus the largest non-integral of \( P \). Then it follows from Theorem 5.4 that \( P + \lambda v \in \Delta'(m) \) for all \( \lambda \in (-\epsilon, \epsilon) \). Thus \( P \) is in the interior of \( \Delta'(m) \).

We now utilize these lemmas to prove that \( \Delta'(m) \) is integral.

**Theorem 6.4.** The polytope \( \Delta'(m) \) is integral.

**Proof.** Let \( P \) be a non-integral point in \( \Delta'(m) \). Then by Theorem 5.2 there exists a row or column for which \( P \) lies on two or fewer row or column facets.

If we assume every non-integral coordinate is on a row and column for which \( P \) lies on exactly two facets, then by Theorem 5.1 there are exactly two non-integral coordinates in each such row and column. By reordering the coordinates we can thus assume \( P \) has one of the configurations in Figure 1. Thus, by Lemma 6.3 \( P \) is not a vertex.

Otherwise, assume there exists a row or column for which \( P \) lies on fewer than two facets. By Lemma 6.1 \( P \) is an interior point of \( \Delta'(m) \).

In summary, if \( P \) is a non-integral point in \( \Delta'(m) \), then it cannot be a vertex.

**Theorem 6.5.** The polytope \( \Delta(m) \) is an H-representation of \( K(m) \).

**Proof.** By Theorem 6.4 the polytope \( \Delta'(m) \) is integral. Applying corollary 4.3 shows \( \Delta(m) \) is also integral. Finally by Theorem 3.2 \( K(m) = \Delta(m) \).

### 7. Conclusion

The H-representation provides a new vantage point for understanding the Kimura-3 variety. The authors hope this will lead to a better understanding of the geometry and associated biology of the Kimura-3 variety, and of group based models in general. To this end, we describe open problems of both mathematical and biological interest.

In biology, phylogenetic varieties have been used to answer identifiability questions [24, 25] and to develop tree reconstruction algorithms [26, 27, 28, 29]. Advances in the understanding of the structure of phylogenetic varieties have improved the speed and accuracy of these methods [30, 31]. A deeper understanding of phylogenetic varieties is also of interest to algebraic geometers in light of recent connections with conformal blocks and Berenstein and Zelevinsky triangles [5]. This motivates our first open problem.
Open Problem 1. Classify the singularity structure of the Kimura-3 varieties.

We are also interested in the structure of the family of all Kimura-3 varieties with a fixed number of leaves. In the binary symmetric case, the variety associated to any tree with \( m \) leaves is deformation equivalent to the variety associated to the \( m \)-claw tree \([13]\). While such a relationship does not hold in the Kimura-3 case \([17]\), one would still like to understand the geometric relationship among varieties associated to different \( n \) leaf trees. We pose this problem from a biological perspective.

Open Problem 2. Describe the relationship between two Kimura-3 varieties whose trees differ by a single nearest neighbor interchange.

In addition these geometric questions, there are also open questions about the combinatorics of the Kimura-3 polytopes.

Open Problem 3. Compute the \( f \)-vector, and Hilbert polynomial of the polytope associated to the Kimura-3 polytope for the \( m \)-claw tree.

Finally, the proof strategies in this paper, along with computations made by Polymake suggest it may be possible to provide a general formula for the H-representation of a group-based model for finite abelian groups.

Open Problem 4. Describe the H-representation for the polytope associated to the \( m \)-claw tree for an arbitrary finite abelian group.

Remark. The first author has a conjectured H-representation of the \( \mathbb{Z}_3 \) \( m \)-claw polytope available upon request.

8. Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. DMS-1358534. Research reported in this publication was supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number P20GM12345.

The authors would like to thank Valery Alexeev, Wayne Anderson, Kais Kubjas and Mateusz Michalek for their comments and suggestions during the development of this work.

References

[1] L. Pachter, B. Sturmfels, Algebraic statistics for computational biology, Vol. 13, Cambridge University Press, 2005.
[2] J. Cavender, J. Felsenstein, Invariants of phylogenies in a simple case with discrete states, J Classification 4 (1987) 57–71.
[3] N. Eriksson, K. Ranestad, B. Sturmfels, S. Sullivant, Phylogenetic algebraic geometry, Projective varieties with unexpected properties (2005) 237–255.
[4] C. A. Manon, The algebra of conformal blocks, arXiv preprint arXiv:0910.0577.
[5] K. Kubjas, C. Manon, Conformal blocks, berenstein–zelevinsky triangles, and group-based models, Journal of Algebraic Combinatorics (2013) 1–26.
[6] S. Evans, T. Speed, Invariants of some probability models used in phylogenetic inferences, Ann Stat. 21 (1993) 355–377.
[7] S. Sturmfels, S. Sullivant, Toric ideals of phylogenetic invariants, J Comput Biol 12 (2005) 204–228.
[8] T. Oda, Convex bodies and algebraic geometry.
[9] W. Fulton, Introduction to toric varieties, no. 131, Princeton University Press, 1993.
[10] D. A. Cox, J. B. Little, H. K. Schenck, Toric varieties, American Mathematical Soc., 2011.
[11] M. Casanellas, J. Fernández-Sánchez, M. Michałek, Local description of phylogenetic group-based models, arXiv preprint arXiv:1402.6945.
[12] M. Donten-Bury, M. Michałek, Phylogenetic invariants for group-based models, Arxiv: 1011.3236.
[13] W. Buczynska, J. A. Wisniewski, On the geometry of binary symmetric models of phylogenetic trees, Journal of the European Mathematical Society 9 (3) (2007) 609–635.
[14] M. Kimura, 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.
[15] M. Casanellas, J. Fernández-Sánchez, Geometry of the kimura 3-parameter model, Advances in Applied Mathematics 41 (3) (2008) 265–292.
[16] M. Michałek, Constructive degree bounds for group-based models, Journal of Combinatorial Theory, Series A 120 (7) (2013) 1672–1694.
[17] K. Kubjas, Hilbert polynomial of the kimura 3-parameter model, arXiv preprint arXiv:1007.3164.
[18] M. Michałek, Toric geometry of the 3-kimura model for any tree, Advances in Geometry 14 (1) (2014) 11–30.
[19] S. Sullivant, Toric fiber products, Journal of Algebra 316 (2) (2007) 560–577.
[20] W. Buczynska, J. A. Wisniewski, On the geometry of binary symmetric models of phylogenetic trees, Journal of the European Mathematical Society 9 (3) (2007) 609–635.
[21] L. Khachiyan, E. Boros, K. Borys, K. Elbassioni, V. Gurvich, Generating all vertices of a polyhedron is hard, Discrete & Computational Geometry 39 (1-3) (2008) 174–190.
[22] H. S. M. Coxeter, Regular polytopes, Courier Dover Publications, 1973.
[23] C. H. Papadimitriou, M. Yannakakis, On recognizing integer polyhedra, Combinatorica 10 (1) (1990) 107–109.
[24] E. S. Allman, S. Petrovic, J. A. Rhodes, S. Sullivant, Identifiability of two-tree mixtures for group-based models, IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB) 8 (3) (2011) 710–722.
[25] C. Long, S. Sullivant, Identifiability of 3-class jukes-cantor mixtures, arXiv preprint arXiv:1406.7256.
[26] N. Eriksson, Tree construction using singular value decomposition, Algebraic Statistics for computational biology (2005) 347–358.
[27] M. Casanellas, J. Fernández-Sánchez, Performance of a new invariants method on homogeneous and nonhomogeneous quartet trees, Molecular biology and evolution 24 (1) (2007) 288–293.
[28] J. P. Rusinko, B. Hipp, Invariant based quartet puzzling., Algorithms for Molecular Biology 7 (1) (2012) 35.
[29] J. Fernández-Sánchez, M. Casanellas, Invariant versus classical approaches when evolution is heterogeneous across sites and lineages, arXiv preprint arXiv:1405.6546.
[30] F. A. Matsen, Fourier transform inequalities for phylogenetic trees, IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB) 6 (1) (2009) 89–95.
[31] M. Casanellas, J. Fernández-Sánchez, Relevant phylogenetic invariants of evolutionary models, Journal de mathématiques pures et appliquées 96 (3) (2011) 207–229.