Gröbner Bases and Polyhedral Geometry of Reducible and Cyclic Models

Serkan Hoşten and Seth Sullivant

5 April 2002

Abstract

This article studies the polyhedral structure and combinatorics of polytopes that arise from hierarchical models in statistics, and shows how to construct Gröbner bases of toric ideals associated to a subset of such models. We study the polytopes for cyclic models, and we give a complete polyhedral description of these polytopes in the binary cyclic case. Further we show how to build Gröbner bases of a reducible model from the Gröbner bases of its pieces. This result also gives a different proof that decomposable models have quadratic Gröbner bases. Finally, we present the solution of a problem posed by Vlach [13] concerning the dimension of fibers coming from models corresponding to the boundary of a simplex.

1 Introduction and Notation

This article is motivated by two fundamental questions on contingency tables that arise in statistics. These two questions can be illustrated with a simple example for 3×3 tables. Such tables are 3×3 arrays of nonnegative real numbers such as

$$\begin{array}{|c|c|c|} \hline 19 & 55 & 88 \\ \hline 28 & 20 & 27 \\ \hline 21 & 32 & 67 \\ \hline \end{array} \quad \text{or} \quad \begin{array}{|c|c|c|} \hline 14 & 63 & 85 \\ \hline 33 & 15 & 27 \\ \hline 21 & 29 & 70 \\ \hline \end{array} \tag{1}$$

Given these tables we can compute their *marginals*, i.e. the row and column sums of their entries. For instance, the marginals of both tables in (1) are $(r_1, r_2, r_3; c_1, c_2, c_3) = (162, 75, 120; 68, 107, 182)$. The first question poses the inverse problem.

Question 1.1 *Given a vector u, does there exist a table whose marginal vector is equal to u?*

Fortunately, this question has an answer for 3×3 (and in general for $m \times n$) tables: $u = (r_1, r_2, r_3; c_1, c_2, c_3)$ is a vector of marginals if and only if $u \in \mathbb{R}^6_+$ and $r_1+r_2+r_3 = c_1 + c_2 + c_3$. In this article, we will answer Question 1.1 for a family of higher-dimensional tables where marginals are computed in a specific way, namely, for *binary cyclic models*.

The second question is about tables with integer entries and identical marginals:

Question 1.2 *Is there a fast way of generating a finite simple set of "moves" that will connect any two tables with integer entries and the marginal vector u such that each intermediate table has the same marginals ?*

In order to give an idea about what we mean by a finite set of moves, let us consider 3×3 tables again. In this case a choice of moves consists of 3×3 tables with entries $x_{ij} = x_{k\ell} = 1$ and $x_{i\ell} = x_{kj} = -1$ (and 0 otherwise) for every choice of $1 \le i < k \le 3$ and $1 \le j < \ell \le 3$. For instance, in order to connect the tables in (1), we can use the following two moves:

$$\begin{array}{|r|r|r|} \hline 1 & -1 & 0 \\ \hline -1 & 1 & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array} \qquad \begin{array}{|r|r|r|} \hline 0 & 1 & -1 \\ \hline 0 & 0 & 0 \\ \hline 0 & -1 & 1 \\ \hline \end{array} \tag{2}$$

By substracting the first move five consecutive times starting from the first table in (1) and then adding the second move three consecutive times to the resulting table we obtain the second table in (1). In this process, the intermediate tables all have nonnegative entries, and the marginals are never altered. Moreover, the finite set of moves we described are simple in the sense that the corresponding tables are sparse and their non-zero entries are small, and they are easy to write down. Such moves were described for all $m \times n$ tables by [11, Proposition 5.4], and they correspond to the elements of a *reduced Gröbner basis* of a certain *toric ideal*. We will answer Question 1.2 for an important family of higher-dimensional tables, namely for *reducible* and *decomposable* models. Now we will describe the objects and summarize the results of this paper more precisely.

An n-dimensional table Γ is a $d_1 \times \cdots \times d_n$ array of nonnegative real numbers where $d_i \geq 2$, $i = 1, \ldots, n$. We will consider such tables as the elements of the vector space $\bigotimes_{i=1}^n \mathbb{R}^{d_i} = \mathbb{R}^{\prod d_i}$. We will denote the coordinate functions and the standard unit vectors in this vector space by x_K and e_K where $K = (k_1, \ldots, k_n)$ with $0 \le k_i \le d_i - 1$. Given such tables there are many different ways of computing marginals. A popular method in statistics is to compute the marginals according to a *hierarchical model* [1], [9]. A hierarchical model consists of a simplical complex Δ on n vertices and the maximal faces of Δ determine how the marginals are computed: for each facet $F = \{j_1, \ldots, j_t\} \subset \Delta$ and each $I = (i_1, \ldots, i_t)$ where $0 \le i_s \le d_{j_s} - 1$, the corresponding component of the marginal vector is

$$u[F(I)] = \sum_{K\,:\,K_F = I} x_K \tag{3}$$

where K_F denotes the restriction of $K = (k_1, \ldots, k_n)$ to the facet F. This computation is induced by the linear map

$$\pi_\Delta : \bigotimes_{i=1}^n \mathbb{R}^{d_i} \longrightarrow \bigoplus_F \bigotimes_{j \in F} \mathbb{R}^{d_j}, \qquad e_K \mapsto \bigoplus_F e_{K_F} \tag{4}$$

where the direct sums are taken over all facets F of Δ.

Remark 1.3 If Δ is the simplicial complex containing just the empty set $\Delta = \{\emptyset\}$ then π_Δ maps to $\mathbb{R}$ where every e_K is mapped to the unique standard unit vector $e = 1$ of $\mathbb{R}$. $\square$

Remark 1.4 In the standard literature (see [1], [9]) a different object called an *independence graph* is used instead of a simplicial complex to represent a model. In this case, the models are called *graphical*, and the corresponding simplicial complex consists of all cliques of the independence graph. By allowing Δ to be any simplicial complex, we can prove results for more general models. Moreover, hierarchical models that are not graphical occur frequently in practice. For example, the simplicial complex Δ that is the boundary of a simplex corresponds to computing all $(n-1)$-marginals of an n-way table, one of the most common models studied. The notation we are using is first introduced in [7]. $\square$

Example 1.5 The example of $m \times n$ tables where the marginals are the row and column sums of the entries is given by the hierarchical model $\Delta = \{\{1\}, \{2\}\}$ of two isolated vertices where $d_1 = m$ and $d_2 = n$. $\square$

Definition 1.6 A hierarchical model on Δ with $d_1 = \cdots = d_n = 2$ is called a binary model.

Example 1.7 Here we introduce the binary four-cycle. Let Δ be the complex $\{\{1,2\}, \{2,3\}, \{3,4\}, \{4,1\}\}$ on four vertices, namely a four-cycle. Suppose $d_i = 2$, i.e. this is a binary model. We will denote the vectors in $\bigoplus_{i=1}^{4} \mathbb{R}^2 \bigotimes \mathbb{R}^2$ by $(a_{00}, a_{01}, a_{10}, a_{11}; b_{00}, \ldots, b_{11}; c_{00}, \ldots, c_{11}; d_{00}, \ldots, d_{11})$. If we order the indices of the unit vectors $e_K = e_{i,j,k,\ell}$ in $\mathbb{R}^2 \bigotimes \mathbb{R}^2 \bigotimes \mathbb{R}^2 \bigotimes \mathbb{R}^2$ where $0 \leq i, j, k, \ell \leq 1$ in *increasing* lexicographic order, then the columns of the following matrix span the image of π_Δ:

$$\begin{pmatrix}
1&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&1&1&1&1&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&1&1&1&1&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&1\\
1&1&0&0&0&0&0&0&1&1&0&0&0&0&0&0\\
0&0&1&1&0&0&0&0&0&0&1&1&0&0&0&0\\
0&0&0&0&1&1&0&0&0&0&0&0&1&1&0&0\\
0&0&0&0&0&0&1&1&0&0&0&0&0&0&1&1\\
1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0\\
0&1&0&0&0&1&0&0&0&1&0&0&0&1&0&0\\
0&0&1&0&0&0&1&0&0&0&1&0&0&0&1&0\\
0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1\\
1&0&1&0&1&0&1&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&1&0&1&0&1&0&1&0\\
0&1&0&1&0&1&0&1&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&1&0&1&0&1&0&1
\end{pmatrix}$$

$\square$

Now we can go back to the questions that motivate this article. In the light of the set-up we developed, Question 1.1 asks about a characterization of all marginal vectors of tables with respect to a hierarchical model defined by Δ and the vector $(d_1, \ldots, d_n)$.

Proposition 1.8 *Let Δ be a simplicial complex on n vertices which defines a hierarchical model together with the vector $(d_1, \ldots, d_n)$. The set of marginal vectors of the tables in this model is a polyhedral cone which is equal to the image of the nonnegative orthant in $\bigotimes_{i=1}^{n} \mathbb{R}^{d_i}$ under the map π_Δ.*

In other words, Question 1.1 points to the study of the cone in Proposition 1.8. Equivalently, we will study the polytope P_Δ which is the convex hull of $\{\pi_\Delta(e_K) : K = \{k_1, \ldots, k_n\},\ 0 \leq k_i \leq d_i - 1\}$. The definition of the linear map π_Δ in (4) implies that P_Δ is a 0/1-polytope whose vertices are precisely $\pi_\Delta(e_K)$. We will first compute the dimension of P_Δ, equivalently, the dimension of kernel of π_Δ. We will do this in Section 2 by constructing an explicit basis for $\ker(\pi_\Delta)$. The structure in this explicit basis will allow us to count the number of facets of P_Δ of a binary model when Δ is the boundary of the n-simplex. Even in this simple case the number of facets grows exponentially in n. Indeed much work has been put to study $m \times n \times p$ tables following the model given by the boundary of a triangle [13] [10], and the partial answers are far from characterizing P_Δ. We will focus in Section 3 on those models where Δ is an n-cycle. In the case of a binary cyclic model we will completely characterize P_Δ by giving explicit facet defining inequalities. Section 4 will address Question 1.2 for reducible and decomposable models. Following the framework developed in [4] we will study the Gröbner bases of the toric ideal given by the vertices of P_Δ where Δ defines a reducible model. We will show how to obtain a Gröbner basis of a reducible model from Gröbner bases of its pieces. This will generalize the results in [6] about the Markov bases of such models. As a corollary we will construct quadratic Gröbner bases for decomposable models, giving a different proof of a result that appears in [7] (see also [5], [12]). In the last section, we will go back to studying the polyhedral geometry of hierarchical models, and present the solution to a question [13, Question 12] about the dimension of fibers, i.e. the dimension of $\pi_\Delta^{-1}(u)$ in the positive orthant for various marginals u, when Δ is the boundary of the n-simplex.

2 A Basis for Vertex Dependencies

For the rest of the paper we let Δ be a simplicial complex on n vertices, and we assume that it determines a hierarchical model together with the vector $(d_1, \ldots, d_n) \in \mathbb{N}^n$ where $d_i \geq 2$. In this section we will prove a formula for the dimension of P_Δ. Equivalently, we will compute the dimension of $\ker(\pi_\Delta)$ by constructing an explicit canonical basis.

Lemma 2.1 *A vector $(x_K) \in \bigotimes_{i=1}^{n} \mathbb{R}^{d_i}$ is in $\ker(\pi_\Delta)$ if and only if for every facet $F = \{j_1, \ldots, j_t\} \subset \Delta$ and each $I = (i_1, \ldots, i_t)$ where $0 \leq i_s \leq d_{j_s} - 1$ we have*

$$\sum_{K:\, K_F = I} x_K = 0 \tag{5}$$

Proof. Fixing F and I means fixing a coordinate in $\bigotimes_{j \in F} \mathbb{R}^{d_j}$. Then the sum in (5) evaluates the value of this coordinate in $\pi_\Delta(x_K)$. □

Definition 2.2 Let S be a nonempty subset of $[n] := \{1, \ldots, n\}$. A vector (x_K) is called an *adjacent minor supported on* S if the following holds for some indices $0 \leq i_j \leq d_j - 2$ where $j \in S$:

1. $x_{k_1,\ldots,k_n} = (-1)^{\sum_{j \in S} \epsilon_j}$ whenever $k_j = i_j + \epsilon_j$ with $\epsilon_j \in \{0, 1\}$ for $j \in S$ and $k_j = 0$ for $j \in [n] \setminus S$.

2. $x_K = 0$ otherwise.

We denote all adjacent minors supported on S by $\beta(S)$.

Remark 2.3 When $n = 2$ and $S = \{1, 2\}$ with $d_1 = m$ and $d_2 = n$, the adjacent minors supported on S correspond to the 2×2 adjacent minors of a generic $m \times n$ matrix. See [3] and [8] for more on 2×2 adjacent minors. $\square$

Example 2.4 Let $n = 3$, $S = \{2, 3\}$ and $d_1 = d_2 = d_3 = 3$. We identify a vector (x_{ijk}) with the three matrices

$$\begin{pmatrix} & & \\ & x_{ij0} & \\ & & \end{pmatrix}\begin{pmatrix} & & \\ & x_{ij1} & \\ & & \end{pmatrix}\begin{pmatrix} & & \\ & x_{ij2} & \\ & & \end{pmatrix}.$$

Then $\beta(S)$ consists of the following four elements:

$$\begin{aligned}
F_1 &= \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\
F_2 &= \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\
F_3 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\
F_4 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.
\end{aligned}$$

$\square$

Lemma 2.5 *The set of all adjacent minors $\bigcup_{S \subseteq [n]} \beta(S)$ where $S \neq \emptyset$ is a linearly independent set. Moreover, this set is a basis over $\mathbb{Z}$ for ker(π_Δ) when $\Delta = \{\emptyset\}$ (for arbitrary $(d_1, \ldots, d_n)$).*

Proof. If we order the indices K of the vector (x_K) reverse lexicographically, the last non-zero coordinate of every adjacent minor is unique, hence the vectors in $\bigcup_{S \subseteq [n]} \beta(S)$ are linearly independent. For every subset $S = \{j_1, \ldots, j_k\}$, the number of the adjacent minors supported on S is $|\beta(S)| = \prod_{i=1}^{k}(d_{j_k} - 1)$, and therefore there are a total of $(\prod_{i=1}^n d_i) - 1$ adjacent minors. If $\Delta = \{\}$, Remark 1.3 implies that the image of any vector (x_K) under π_Δ is $\sum_K x_K$. Therefore the kernel of this map has dimension $(\prod_{i=1}^n d_i) - 1$. By our definition, all adjacent minors are in the kernel of π_Δ, and this shows that the set of all adjacent minors is a basis for the subspace $ker(\pi_\Delta)$. Moreover, the value of the unique last non-zero coordinate of these adjacent minors is either 1 or -1, and hence an integral vector in $ker(\pi_\Delta)$ is an integer linear combination of the adjacent minors. We conclude that the set of adjacent minors is actually a basis over $\mathbb{Z}$. $\square$

Theorem 2.6 *For any hierarchical model, a canonical $\mathbb{Z}$-basis for $ker(\pi_\Delta)$ is given by*

$$\bigcup_{F \notin \Delta} \beta(F),$$

and hence the dimension of this kernel is

$$\sum_{F \notin \Delta} \prod_{F=\{j_1,\ldots,j_t\}} (d_{j_i} - 1).$$

Proof. The second statement follows from the first one. We observe that $ker(\pi_\Delta) \subset ker(\pi_{\Delta'})$ where Δ' is obtained by removing any facet of Δ. In particular, $ker(\pi_\Delta)$ is contained in $ker(\pi_{\{\emptyset\}})$. If (x_K) is an integer vector in the kernel of π_Δ, then by Lemma 2.5 this vector is a unique integer linear combination of all adjacent minors. However, Lemma 2.1 implies that those adjacent minors supported on a face of Δ cannot contribute to this linear combination. $\square$

Corollary 2.7 *The dimension of the polytope P_Δ is*

$$\prod_{i=1}^n d_i - \left(\sum_{F \notin \Delta} \prod_{F=\{j_1,\ldots,j_t\}} (d_{j_i} - 1) \right) - 1 = \left(\sum_{G \in \Delta} \prod_{G=\{j_1,\ldots,j_t\}} (d_{j_i} - 1) \right) - 1.$$

The last result in this section shows that for even very simple hierarchical models the complexity of P_Δ grows fast.

Theorem 2.8 *Suppose Δ gives rise to a binary hierarchical model where Δ is the boundary of an n-simplex. Then the dimension of P_Δ is $2^n - 2$ and this polytope has 4^{n-1} facets.*

Proof. Since $d_i = 2$ for $i = 1, \ldots, n$, the result on the dimension of P_Δ follows from Corollary 2.7. The kernel of π_Δ is generated by a single adjacent minor supported on $S = [n]$. The coordinates of this adjacent minor form the Gale transform (see [14, Chapter 6]) of the vertices of P_Δ. Since 2^{n-1} of the total 2^n coordinates are $+1$ and the rest of them are -1, there are $2^{n-1}2^{n-1}$ positive circuits of this Gale transform. Because positive circuits are in bijection with the facets of P_Δ we conclude that P_Δ has 4^{n-1} facets. $\square$

3 Cyclic Models

In this section we consider simplicial complexes that are n-cycles, i.e $\Delta = \{\{1,2\},\{2,3\}, \ldots, \{n-1,n\},\{n,1\}\}$. We will denote the coordinates in $\bigoplus_{i=1}^{n}(\mathbb{R}^{d_i} \bigotimes \mathbb{R}^{d_{i+1}})$ by y^i_{st} with $1 \le i \le n$, $0 \le s \le d_i - 1$ and $0 \le t \le d_{i+1} - 1$, and where we assume that $d_{n+1} = d_1$. We also identify the indices $K = (k_1, \ldots, k_n)$ with the vertex of P_Δ given by $\pi_\Delta(e_K)$.

Theorem 3.1 *Let V be the set of vertices $(k_1, \ldots, k_n)$ of P_Δ where either $k_1 = k_2 = \cdots = k_j = 0$ and $k_{j+1}, \ldots, k_n > 0$, or $k_1, \ldots, k_j > 0$ and $k_{j+1} = \cdots = k_n = 0$ for some $0 \le j \le n$. Then V forms a facet of P_Δ supported by*

$$y^1_{00} - \sum_{t=1}^{d_3-1} y^2_{0t} - \sum_{t=1}^{d_4-1} y^3_{0t} - \cdots - \sum_{t=1}^{d_n-1} y^{n-1}_{0t} - y^n_{00} \le 0 \tag{6}$$

Proof. It is clear that all vertices (and hence all points) of P_Δ satisfy (6), and the vertices in V are the only vertices on the hyperplane defined by (6). Corollary 2.7 implies that $dim(P_\Delta) = \sum_{i=1}^{n}(d_i - 1)(d_{i+1} - 1) + \sum_{i=1}^{n}(d_i - 1)$, so we will show that the dimension of the subspace generated by the vertices in V is equal to $dim(P_\Delta)$. We show this by induction on the vector $d = (d_1, \ldots, d_n)$ while Δ is the fixed simplicial complex of an n-cycle. In order to emphasize the dependence of P_Δ on d, we will denote this polytope by $P_{\Delta,d}$ when necessary. When $d = (2, \ldots, 2)$ then $y^i_{01} = 1$ and $y^i_{10} = 1$ for $i = 1, \ldots, n-1$, and $y^n_{00} = 1$ and $y^n_{11} = 1$ all appear in a unique vertex in V respectively, and hence all these $2n = dim(P_{\Delta,(2,\ldots,2)})$ vertices are linearly independent. Now suppose the claim is true for $2 \le d_i \le m_i$ for all i. Let $d = (m_1, \ldots, m_n)$ and $d' = (m_1 + 1, m_2, \ldots, m_n)$, and let V and V' be the set of vertices of $P_{\Delta,d}$ and $P_{\Delta,d'}$ respectively supported on (6). We observe that the set of indices of the vertices in V is contained in that of the vertices in V'. By induction the dimension of the subspace generated by V is $dim(P_{\Delta,d})$. We compute that the dimension of the subspace generated by V' is $dim(P_{\Delta,d'}) = dim(P_{\Delta,d}) + (m_2 - 1) + (m_n - 1) + 1$. Among those in $V' \setminus V$ there are the vertices $(m_1, 0, 0, \ldots, 0), (m_1, 1, 0, \ldots, 0), \ldots, (m_1, m_2 - 1, 0, \ldots, 0)$ and $(m_1, 1, \ldots, 1, 1), (m_1, 1, \ldots, 1, 2), \ldots, (m_1, 1, \ldots, 1, m_n - 1)$ which have a distinct "new" coordinate that is equal to one, namely, $y^1_{m_1,t} = 1$ for $t = 0, \ldots, m_2 - 1$ and $y^n_{s,m_1} = 1$ for $s = 1, \ldots, m_n - 1$ respectively. So these $m_2 + m_n - 1$ vertices are linearly independent, as well as they are linearly independent from those vertices in V. □

Example 3.2 Let Δ be the simplicial complex of a cycle on four vertices with $(d_1, d_2, d_3, d_4) = (3, 2, 3, 2)$. In this case the facet (6) reads as

$$y^1_{00} - y^2_{01} - y^2_{02} - y^3_{01} - y^4_{00} \le 0 \tag{7}$$

Using Theorem 3.1 we see that the following 18 vertices are on this facet:

$$\begin{array}{ccc}(2,1,2,1) & (1,1,2,1) & (0,1,2,1)\\ (2,1,1,1) & (1,1,1,1) & (0,1,1,1)\\ (2,1,2,0) & (1,1,2,0) & (0,0,2,1)\\ (2,1,1,0) & (1,1,1,0) & (0,0,1,1)\\ (2,1,0,0) & (1,1,0,0) & (0,0,0,1)\\ (2,0,0,0) & (1,0,0,0) & (0,0,0,0)\end{array}$$

For instance the vertex $(0,0,2,1)$ is on this facet since $y^1_{00} = y^2_{02} = y^3_{21} = y^4_{10} = 1$ and otherwise $y^i_{st} = 0$. □

Corollary 3.3 *Under the natural action of the group $S_{d_1} \times \cdots \times S_{d_n}$, the vertices V of the facet described in Theorem 3.1 are mapped to a (not necessarily different) facet. The number of such facets is equal to $(d_1!)(d_2!)\cdots(d_n!)/|H_V|$ where H_V is the stabilizer of V.*

Corollary 3.4 *Let P_Δ be the polytope for a binary cyclic model on n vertices. Then the facet (6) is a simplicial facet, and there are 2^{n-1} such facets.*

Proof. In the proof of Theorem 3.1, we showed that the set of vertice V of the facet (6) consists of $(0,0,\ldots,0),(0,0,\ldots,1),(0,0,\ldots,1,1),\ldots,(0,1,\ldots,1,1)$, and $(1,0,\ldots,0),(1,1,0,\ldots,0),\ldots,(1,1,\ldots,1)$, and these $2n = dim(P_\Delta)$ vertices are linearly independent. This proves the first statement. For the second statement we observe that the only non-trivial group element that stabilizes V is the one that maps $(0,0,\ldots,0)$ to $(1,1,\ldots,1)$ □

Theorem 3.5 *The polytope P_Δ of a binary cyclic model on n vertices is equal to intersection of the affine subspace given by the equations*

$$\begin{aligned} y^k_{00} + y^k_{10} &= y^{k+1}_{00} + y^{k+1}_{01} \\ y^k_{01} + y^k_{11} &= y^{k+1}_{10} + y^{k+1}_{11} \end{aligned}$$

for $1 \le k \le n$, and

$$\sum_{k=1}^{n} (y^k_{00} + y^k_{10} + y^k_{01} + y^k_{11}) = n$$

with the $4n$ facets $y^k_{ij} \ge 0$ where $1 \le k \le n$, and the 2^{n-1} facets described in (6).

Proof. We proceed by induction on n, where the case for $n = 3$ follows by direct computation, and we assume the statement is true for $< n$. We let $P_n = P_\Delta$. It is not hard to see that affine subspace defined by the above set of equations is $\mathrm{aff}(\pi_\Delta(e_K))$, the affine hull of the vertices of P_n, and we denote the polytope defined by $\mathrm{aff}(\pi_\Delta(e_K))$ and the $4n + 2^{n-1}$ inequalities in the statement of the theorem by Q_n. We need to show that $P_n = Q_n$. The polytope P_n is contained in Q_n since $P_n = \mathrm{conv}(\pi_\Delta(e_K)) \subset \mathrm{aff}(\pi_\Delta(e_K))$ and the vertices of P_n satisfy the inequalities defined by $y^k_{ij} \ge 0$ (trivially) and those defined by (6) (by Theorem 3.1). In order to prove the equality we will first define a linear map $\phi : Q_n \longrightarrow Q_{n-1} = P_{n-1}$. This implies that $\phi(x)$ for any $x \in Q_n$ is a convex combination of the vertices of P_{n-1}. Then we will lift this convex combination to a convex combination of the vertices of P_n to express x, and this will finish the proof. As above we identify the indices $K = (k_1, \ldots, k_n)$ with the vertex $\pi_\Delta(e_K)$ of P_n. Now the map ϕ is a linear map that satisfies $\phi(y^i_{st}) = \bar{y}^i_{st}$ for all $1 \leq i \leq n-2$, and $\phi(y^{n-1}_{s0} + y^n_{0t}) = \bar{y}^{n-1}_{st}$ and $\phi(y^{n-1}_{s1} + y^n_{1t}) = \bar{y}^{n-1}_{st}$. Let $x \in Q_n$ which is an affine combination of the vertices of P_n:

$$x = \sum_{*} \left(\sum_{s,t} \mu_{t,*,s}(t,*,s,0) + \lambda_{t,*,s}(t,*,s,1) \right)$$

where the outer sum runs over the indices $* = (k_2, \ldots, k_{n-2})$, and the sum of all the coefficients $\sum_{t,*,s}(\mu_{t,*,s} + \lambda_{t,*,s})$ is equal to one. Then using the definition of the map ϕ we conclude that

$$\phi(x) = \sum_{*} \left(\sum_{s,t} (\mu_{t,*,s} + \lambda_{t,*,s})(t,*,s) \right)$$

and hence $\phi(x)$ is an affine combination of the vertices of P_{n-1}. This also shows that since x satisfies all the inequalities $y^i_{st} \geq 0$, so does $\phi(x)$ satisfy all of $\bar{y}^i_{st} \geq 0$. By Theorem 3.1, x also satisfies

$$y^1_{00} - y^2_{01} - y^3_{01} - \cdots - y^{n-1}_{01} - y^n_{00} \leq 0 \quad \text{and} \quad y^1_{00} - y^2_{01} - y^3_{01} - \cdots - y^{n-1}_{00} - y^n_{10} \leq 0$$

Combining these two we see that $\phi(x)$ satisfies

$$\bar{y}^1_{00} - \bar{y}^2_{01} - \bar{y}^3_{01} - \cdots - \bar{y}^{n-1}_{00} \leq 0$$

This means, by symmetry, $\phi(x)$ satisfies all the 2^{n-2} facets as in (6), and we conclude that $\phi(x)$ is in P_{n-1}. Hence $\phi(x)$ is a convex combination of the vertices of P_{n-1}, i.e.,

$$\phi(x) = \sum_{*} \left(\sum_{s,t} \bar{\mu}_{t,*,s}(t,*,s) \right).$$

Therefore, we can write x as the convex combination of the vertices of P_n:

$$x = \sum_{*} \left(\sum_{s,t} \frac{1}{2}\bar{\mu}_{t,*,s}(t,*,s,0) + \frac{1}{2}\bar{\mu}_{s,*,t}(t,*,s,1) \right)$$

□

4 Gröbner Bases of Reducible Models

In this section, we use computational algebraic geometry to study hierarchical models in order to answer Question 1.2. In particular, we will show how to build special Gröbner bases for the toric ideal of a hierarchical model whenever the underlying simplicial complex is reducible.

Definition 4.1 A simplicial complex Δ is called *reducible* if there is a nontrivial decomposition (Δ_1, S, Δ_2) such that the following three properties hold:

1. $\Delta = \Delta_1 \cup \Delta_2$,
2. S is a face of Δ_1 and Δ_2,
3. $(|\Delta_1| \setminus S) \cap (|\Delta_2| \setminus S) = \emptyset$

where $|\Delta| = \cup_{F \in \Delta} F$. The face S is called a *separator*.

Example 4.2 The union of the boundary of two triangles that share an edge is reducible. We can take $\Delta = \{\{1,2\},\{1,3\},\{2,3\},\{2,4\},\{3,4\}\}$, and we have the decomposition (Δ_1, S, Δ_2) where $\Delta_1 = \{\{1,2\},\{1,3\},\{2,3\}\}$, $S = \{2,3\}$, and $\Delta_2 = \{\{2,3\},\{2,4\},\{3,4\}\}$. $\square$

The idea we will pursue is to build a Gröbner basis for the toric ideal corresponding to Δ from the toric ideals of the pieces corresponding to Δ_1 and Δ_2. Before we present our results we review the basics of Gröbner bases of toric ideals in connection to Question 1.2. For the details on Gröbner bases and toric ideals see [2] and [11].

Let $R = \mathbb{Q}[x_1, \ldots, x_n]$ be a polynomial ring in n variables over $\mathbb{Q}$. A *monomial* is a product of powers of the variables $x_1^{u_1} \cdots x_n^{u_n}$ which we simply denote by $\mathbf{x}^{\mathbf{u}}$. A *term order* $\prec$ is a well ordering of all the monomials in R (with the minimum element $\mathbf{x}^{\mathbf{0}} = 1$) that is compatible with multiplication; that is, $\mathbf{x}^{\mathbf{u}_1} \prec \mathbf{x}^{\mathbf{u}_2}$ implies that $\mathbf{x}^{\mathbf{u}_3}\mathbf{x}^{\mathbf{u}_1} \prec \mathbf{x}^{\mathbf{u}_3}\mathbf{x}^{\mathbf{u}_2}$ for any monomial $\mathbf{x}^{\mathbf{u}_3}$. Given a nonzero polynomial f and a term order $\prec$, we let $\mathsf{in}_\prec(f)$, the *initial term* of f, to be the largest monomial of f with respect $\prec$. If I is an ideal in R, the *initial ideal* of I with respect to the term order $\prec$ is the monomial ideal generated by all the initial terms of polynomials in I:

$$\mathsf{in}_\prec(I) = \langle \mathsf{in}_\prec(f) \;:\; f \in I \rangle.$$

Definition 4.3 Given an ideal I and a term order $\prec$, a finite subset of polynomials G of I is called a *Gröbner basis* for I with respect to $\prec$ if the initial terms of the polynomials of G generate $\mathsf{in}_\prec(I)$,

$$\mathsf{in}_\prec(I) = \langle \mathsf{in}_\prec(g) \;:\; g \in G \rangle.$$

We now turn our attention to a special class of ideals constructed as follows.

Definition 4.4 Let A be a $p \times r$ matrix with integer entries. The *toric ideal* I_A is the ideal generated by binomials of the form $\mathbf{x}^{\mathbf{u}_1} - \mathbf{x}^{\mathbf{u}_2}$ where $A(\mathbf{u}_1 - \mathbf{u}_2) = \mathbf{0}$, in other words,

$$I_A = \langle \mathbf{x}^{\mathbf{u}_1} - \mathbf{x}^{\mathbf{u}_2} \;:\; A\mathbf{u}_1 = A\mathbf{u}_2 \text{ and } \mathbf{u}_1, \mathbf{u}_2 \in \mathbb{N}^r \rangle.$$

Toric ideals coming from hierarchical models arise in a natural way. Given a simplicial complex Δ on n vertices and a vector of integers $d = (d_1, \ldots, d_n)$ where $d_i \geq 2$, we get a hierarchical model and the map π_Δ defined in (4). Recall that the vectors $\pi_\Delta(e_K)$ are 0/1-vectors which are the vertices of the polytope P_Δ. We let A_Δ be the matrix whose columns are these vertices. The toric ideal of A_Δ, which we denote by I_Δ, lives in the polynomial ring $R = \mathbb{Q}[x_K \ : K = (k_1, \ldots, k_n)]$ with $\prod_{i=1}^n d_i$ variables.

Example 4.5 The matrix in Example 1.7 is A_Δ where Δ is a cycle on four vertices and $d_1 = d_2 = d_3 = d_4 = 2$. The toric ideal I_Δ is an ideal in $\mathbb{Q}[x_{0000}, x_{0001}, \ldots, x_{1111}]$ with 16 variables. A simple computation shows it is generated by the following 16 binomials:

$$
\begin{aligned}
I_\Delta = \langle \quad & x_{1011}x_{1110} - x_{1010}x_{1111}, x_{0111}x_{1101} - x_{0101}x_{1111}, \\
& x_{1001}x_{1100} - x_{1000}x_{1101}, x_{0110}x_{1100} - x_{0100}x_{1110}, \\
& x_{0011}x_{1001} - x_{0001}x_{1011}, x_{0010}x_{1000} - x_{0000}x_{1010}, \\
& x_{0011}x_{0110} - x_{0010}x_{0111}, x_{0001}x_{0100} - x_{0000}x_{0101}, \\
& x_{0000}x_{0011}x_{1101}x_{1110} - x_{0001}x_{0010}x_{1100}x_{1111}, \\
& x_{0000}x_{0111}x_{1001}x_{1110} - x_{0001}x_{0110}x_{1000}x_{1111}, \\
& x_{0000}x_{0110}x_{1011}x_{1101} - x_{0010}x_{0100}x_{1001}x_{1111}, \\
& x_{0001}x_{0110}x_{1010}x_{1101} - x_{0010}x_{0101}x_{1001}x_{1110}, \\
& x_{0000}x_{0111}x_{1011}x_{1100} - x_{0011}x_{0100}x_{1000}x_{1111}, \\
& x_{0010}x_{0101}x_{1011}x_{1100} - x_{0011}x_{0100}x_{1010}x_{1101}, \\
& x_{0001}x_{0111}x_{1010}x_{1100} - x_{0011}x_{0101}x_{1000}x_{1110}, \\
& x_{0100}x_{0111}x_{1001}x_{1010} - x_{0101}x_{0110}x_{1000}x_{1011} \qquad \rangle
\end{aligned}
$$

□

Given a hierarchical model Δ with $d = (d_1, \ldots, d_n)$, a monomial $x_{K_1}^{\alpha_1} \cdots x_{K_t}^{\alpha_t}$ in $R = \mathbb{Q}[x_K]$ corresponds to a unique $d_1 \times \cdots \times d_n$ table Γ where the entries of Γ can be read off from the exponents of the variables: $\Gamma(K_j) = \alpha_j$ for $j = 1, \ldots, t$ and $\Gamma(K) = 0$ otherwise.

Notation 4.6 To make the exposition easier we introduce tableau notation for monomials in the ring R. To each monomial $x_{K_1} \cdots x_{K_t}$ where $K_j = (k_{j1}, \ldots, k_{jn})$ for $j = 1, \ldots, t$, we associate the $t \times n$ tableau

$$
\begin{bmatrix} k_{11} & \cdots & k_{1n} \\ \vdots & \ddots & \vdots \\ k_{t1} & \cdots & k_{tn} \end{bmatrix}.
$$

If a variable occurs to its p-th power in a monomial, its corresponding index set occurs p times in the tableau. Also, when we write $T = T'$ for two tableaus we mean that they are equal up to a permutation of the rows, since the rows of the tableaus are indices for commuting variables.

Example 4.7 In tableau notation, the last generator in Example 4.5 is written as

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix}.$$

□

Recall that Question 1.2 asks whether we can find simple moves (or local alterations) to obtain a target table from a starting one while visiting only those nonnegative tables with the same marginals as the initial and final table. Any Gröbner basis of the toric ideal I_Δ provides the answer:

Proposition 4.8 *Let G be a Gröbner basis of I_Δ for a hierarchical model defined by Δ. Let $\mathbf{x}^\alpha$ and $\mathbf{x}^\beta$ be two monomials corresponding to two tables Γ_1 and Γ_2 with the same marginals. Then*

$$\mathbf{x}^\alpha + \sum_{i=1}^{t} \mathbf{x}^{\gamma_i}(\mathbf{x}^{\alpha_i} - \mathbf{x}^{\beta_i}) = \mathbf{x}^\beta$$

where $\mathbf{x}^{\alpha_i} - \mathbf{x}^{\beta_i}$ are in G, and each partial sum

$$\mathbf{x}^\alpha + \sum_{i=1}^{j} \mathbf{x}^{\gamma_i}(\mathbf{x}^{\alpha_i} - \mathbf{x}^{\beta_i}) \quad j = 1, \ldots, t$$

is a monomial that corresponds to a table with the same marginals as Γ_1 and Γ_2.

The above proposition still holds if we replace "Gröbner basis" with "minimal generating set" and this is proved in [11, Chapter 5]. Although a minimal generating set is a subset of a Gröbner basis, the latter provides more information about I_Δ and the other objects related to it such as P_Δ. Moreover it is more useful algorithmically.

The last thing we need before we construct Gröbner bases of reducible models is a quick way of checking when a binomial is in I_Δ.

Lemma 4.9 *Let Γ be a $d_1 \times \cdots \times d_n$ table and T its expression in tableau notation. Let $F = \{j_1, \ldots, j_t\} \subset [n]$. The F-marginal of Γ*

$$\left(\sum_{K \,:\, K_F = I} \Gamma(K) : \; I = (i_1, \ldots, i_t), \; 0 \leq i_s \leq d_{j_s} - 1 \right)$$

is a $d_{j_1} \times \cdots \times d_{j_t}$ table, and its tableau is computed by deleting the columns of T corresponding to those indices in $[n] \setminus F$.

Taking the F-marginal of a table Γ is an operation that we will use often, and using Lemma 4.9 we will denote the corresponding tableau by $T|_F$ where T is a tableau that goes with Γ.

Proposition 4.10 *Let $g = T - T'$ be a binomial in tableau notation. Then g is in I_Δ if and only if $T|_F = T'|_F$ for every facet F of Δ.*

Proof. The tableaus T and T' correspond to tables Γ and Γ', and by the definition of I_Δ, the binomial g is in I_Δ if and only if for every facet F of Δ the F-marginals of Γ and Γ' are the same. Now Lemma 4.9 finishes the proof. $\square$

We now concentrate our attention on a very special type of reducible model and construct Gröbner bases for this class of models. These Gröbner bases will form the building blocks for everything that follows.

Theorem 4.11 *Consider the model whose simplicial complex $\Delta = \{\{1\},\{2\}\}$ consists of two isolated points with $d = (d_1, d_2)$. Then the set of quadratic binomials of the form*

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

where $0 \leq i_1 < i_2 < d_1$ and $0 \leq j_1 < j_2 < d_2$, are a Gröbner basis for I_Δ with respect to the lexicographic ordering induced by the index vectors of variables x_{ij}.

Proof. See [11, Proposition 5.4]. $\square$

A simple generalization of the two-point model is one with Δ having exactly two facets.

Definition 4.12 Suppose that Δ is a simplicial complex with exactly two facets F_1 and F_2 and let $d = (d_1, \dots, d_n)$. Let $S = F_1 \cap F_2$ and suppose that $F_1 \setminus S = \{1, \dots, a\}$, $S = \{a+1, \dots, b\}$, and $F_2 \setminus S = \{b+1, \dots, n\}$. We define $\mathcal{F}(F_1, F_2)$ as the set of all binomials of the form

$$\begin{bmatrix} p_1 & q & r_1 \\ p_2 & q & r_2 \end{bmatrix} - \begin{bmatrix} p_1 & q & r_2 \\ p_2 & q & r_1 \end{bmatrix}$$

where $p_1 \neq p_2$ range over $\{0, \dots, d_1 - 1\} \times \cdots \times \{0, \dots d_a - 1\}$, q ranges over $\{0, \dots, d_{a+1} - 1\} \times \cdots \times \{0, \dots, d_b - 1\}$, and $r_1 \neq r_2$ range over $\{0, \dots, d_{b+1} - 1\} \times \cdots \times \{0, \dots, d_n - 1\}$.

$\mathcal{F}(F_1, F_2)$ will also be used in models with more than two facets. We will use the same construction whenever F_1 and F_2 are two subsets of $[n]$ such that $F_1 \cup F_2 = [n]$. The set $\mathcal{F}(F_1, F_2)$ was introduced in a different form in [5].

Theorem 4.13 *Let Δ be a hierarchical model with two facets F_1 and F_2. Then the set $\mathcal{F}(F_1, F_2)$ is a Gröbner basis for I_Δ with respect to a lexicographic term order.*

Proof. There are two cases to consider. When $S = \emptyset$, the model is just a two-point model $\Delta' = \{\{1'\},\{2'\}\}$ where $d'_1 = d_1 d_2 \cdots d_a$ and $d'_2 = d_{a+1} d_{a+2} \cdots d_n$, and the statement follows from Theorem 4.11. When $S \neq \emptyset$, the rows and columns of the matrix A_Δ can be rearranged so that it is block diagonal with identical blocks. The columns of a fixed block correspond to the variables x_K with $K = (k_1, \dots, k_n)$ where $k_{a+1}, \dots, k_b$ are fixed, and such a block is the matrix of a two-point model. Hence $\mathcal{F}(F_1, F_2)$ is the union of the set of binomials coming from these two-point models which do not share variables. The statement follows again from Theorem 4.11. $\square$

Assume henceforth that Δ is a reducible simplicial complex with decomposition (Δ_1, S, Δ_2) and that $|\Delta_1| \setminus S = \{1, \ldots, a\}$, $S = \{a+1, \ldots, b\}$, and $|\Delta_2| \setminus S = \{b+1, \ldots, n\}$. We call the models Δ_1 and Δ_2 induced submodels of the hierarchical model Δ.

Any binomial in the ideal I_Δ (for *any* model Δ) is homogeneous; this means, if $f = T - T'$ then the number of rows of T and T' are the same. Now, writing each binomial of I_{Δ_1} in tableau notation as

$$f = T - T' = \begin{bmatrix} p_1 & q_1 \\ \vdots & \vdots \\ p_m & q_m \end{bmatrix} - \begin{bmatrix} p_1{}' & q_1{}' \\ \vdots & \vdots \\ p_m{}' & q_m{}' \end{bmatrix}$$

one can permute the rows of T' so that $q_i = q_i{}'$. This is because S is a face of Δ_1 (and Δ and Δ_2) so the tables corresponding to the tableau T and T' must have the same S marginals. We will make the assumption henceforth that every binomial in I_{Δ_1} is written in the form

$$f = T - T' = \begin{bmatrix} p_1 & q_1 \\ \vdots & \vdots \\ p_m & q_m \end{bmatrix} - \begin{bmatrix} p_1{}' & q_1 \\ \vdots & \vdots \\ p_m{}' & q_m \end{bmatrix}$$

and make the same pairing of S index sets for binomials in I_Δ and I_{Δ_2} whenever we write binomials in these ideals.

Definition 4.14 Let $H \subset I_{\Delta_1}$ be a set of binomials

$$f = \begin{bmatrix} p_1 & q_1 \\ \vdots & \vdots \\ p_m & q_m \end{bmatrix} - \begin{bmatrix} p_1{}' & q_1 \\ \vdots & \vdots \\ p_m{}' & q_m \end{bmatrix}$$

where the p_i, $p_i{}'$, and q_i range over the usual index sets. Let $r_1, \ldots, r_m \in \{0, \ldots, d_{b+1} - 1\} \times \cdots \times \{0, \ldots, d_n - 1\}$. We create a new binomial f' from f by appending the r_i as

$$f' = \begin{bmatrix} p_1 & q_1 & r_1 \\ \vdots & \vdots & \vdots \\ p_m & q_m & r_m \end{bmatrix} - \begin{bmatrix} p_1{}' & q_1 & r_1 \\ \vdots & \vdots & \vdots \\ p_m{}' & q_m & r_m \end{bmatrix}$$

and we define $Ext(H \to I_\Delta) \subset R$ to be the set of all f' as f ranges over H and $r_1, \ldots, r_m$ range over $\{0, \ldots, d_{b+1} - 1\} \times \cdots \times \{0, \ldots, d_n - 1\}$. Similarly, we define $Ext(H' \to I_\Delta)$ for H' a set of binomials in I_{Δ_2}.

Lemma 4.15 *Let H be a set of binomials in I_{Δ_1} or I_{Δ_2}. Then $Ext(H \to I_\Delta)$ is a subset of I_Δ.*

Proof. It suffices to show this for $H \subset I_{\Delta_1}$. Let $f = T - T' \in Ext(H \to I_\Delta)$. For f to be in I_Δ we must have the F-marginals of T and T' equal for every facet F of the simplicial complex Δ. Equivalently, we must show that $f|_F = 0$ for all facets F of Δ.

If F is a facet of Δ then it is also a facet of either Δ_1 or Δ_2. Suppose first that F is a facet of Δ_1. Since $f|_F = (f|_{\Delta_1})|_F$ and $f|_{\Delta_1} \in I_{\Delta_1}$ we have $(f|_{\Delta_1})|_F = 0$. On the other hand, if F is a facet of Δ_2 then $f|_{\Delta_2} = 0$ and hence $f|_F = 0$ as well. $\square$

Lemma 4.16 *Let $\prec_1$ and $\prec_2$ be term orders for the induced submodels Δ_1 and Δ_2 respectively and let $\prec$ be a term order for the model Δ. Then the relation $\prec_*$ on the set of monomials (tableaus) for the model Δ given by*

$$T \prec_* T' \quad \textit{if} \left\{ \begin{array}{ccc} T|_{\Delta_1} \prec_1 T'|_{\Delta_1} & & \textit{or} \\ T|_{\Delta_1} = T'|_{\Delta_1} \quad \textit{and} \quad T|_{\Delta_2} \prec_2 T'|_{\Delta_2} & & \textit{or} \\ T|_{\Delta_1} = T'|_{\Delta_1} \quad \textit{and} \quad T|_{\Delta_2} = T'|_{\Delta_2} & \textit{and} & T \prec T' \end{array} \right.$$

is a term order.

Proof. For any distinct tableau T and T', it is clear that either $T \prec_* T'$ or $T' \prec_* T$. Additionally, the relation $\prec_*$ is preserved under multiplication by monomials. To show that $\prec_*$ is a term order, it remains to show that $\prec_*$ is a transitive relation. We must show that if $T \prec_* T'$ and $T' \prec_* T''$ then $T \prec_* T''$. Verifying this amounts to checking nine cases coming from the three different possible ways that two tables could be compared. We include just one case since all nine cases run in the same fashion.

Suppose $T \prec_* T'$ because $T|_{\Delta_1} = T'|_{\Delta_1}$ but $T|_{\Delta_2} \prec_2 T'|_{\Delta_2}$ and also suppose $T' \prec_* T''$ because $T'|_{\Delta_1} \prec_1 T''|_{\Delta_1}$. Then

$$T|_{\Delta_1} = T'|_{\Delta_1} \prec_1 T''|_{\Delta_1}$$

hence $T \prec_* T''$. $\square$

We now come to the main theorem of this section.

Theorem 4.17 *Let G_1 and G_2 be Gröbner bases for I_{Δ_1} and I_{Δ_2} with respect to $\prec_1$ and $\prec_2$ respectively. Suppose that $\prec$ is a lexicographic term order described in Theorem 4.13 for $\mathcal{F}(|\Delta_1|, |\Delta_2|)$. Then the set*

$$G = Ext(G_1 \to I_\Delta) \cup Ext(G_2 \to I_\Delta) \cup \mathcal{F}(|\Delta_1|, |\Delta_2|)$$

is a Gröbner basis for I_Δ with respect to the term order $\prec_*$ described in Lemma 4.16.

Proof. It suffices to show that for any binomial $f = T - T' \in I_\Delta$ where $\mathsf{in}_{\prec_*}(f) = T$, there is a binomial $g \in G$ with a leading term dividing T. We write f explicitly in tableau notation as

$$f = T - T' = \begin{bmatrix} p_1 & q_1 & r_1 \\ \vdots & \vdots & \vdots \\ p_m & q_m & r_m \end{bmatrix} - \begin{bmatrix} {p_1}' & q_1 & {r_1}' \\ \vdots & \vdots & \vdots \\ {p_m}' & q_m & {r_m}' \end{bmatrix}$$

There are three cases to consider. First, suppose that $T|_{\Delta_1} \neq T'|_{\Delta_1}$. Note that the leading term of $f|_{\Delta_1}$ is $T|_{\Delta_1}$ with respect to $\prec_1$. By assumption, there is a $g \in G_1$ whose leading term divides $T|_{\Delta_1}$. After rearranging the rows of f, we may assume

$$g = U - U' = \begin{bmatrix} p_1 & q_1 \\ \vdots & \vdots \\ p_i & q_i \end{bmatrix} - \begin{bmatrix} p_1'' & q_1 \\ \vdots & \vdots \\ p_i'' & q_i \end{bmatrix}$$

with leading term U. Then the binomial

$$g' = V - V' = \begin{bmatrix} p_1 & q_1 & r_1 \\ \vdots & \vdots & \vdots \\ p_i & q_i & r_i \end{bmatrix} - \begin{bmatrix} p_1'' & q_1 & r_1 \\ \vdots & \vdots & \vdots \\ p_i'' & q_i & r_i \end{bmatrix}$$

is in $Ext(G_1 \to I_\Delta)$ and has the leading term V dividing T. The case where $T|_{\Delta_1} = T'|_{\Delta_1}$ but $T|_{\Delta_2} \neq T'|_{\Delta_2}$ is handled in a similar way. If both $T|_{\Delta_1} = T'|_{\Delta_1}$ and $T|_{\Delta_2} = T'|_{\Delta_2}$ then the binomial f lies not only in I_Δ, but also in the ideal I_Σ, where Σ is the simplicial complex with exactly two facets, $|\Delta_1|$ and $|\Delta_2|$. But $\mathcal{F}(|\Delta_1|, |\Delta_2|)$ is a Gröbner basis for I_Σ and so T is divisible by the leading term of a binomial in $\mathcal{F}(|\Delta_1|, |\Delta_2|)$. This shows that G is a Gröbner basis with respect to $\prec_*$. $\square$

Example 4.18 As an example consider the model described in Example 4.2 in the binary case $d = (2, 2, 2, 2)$. Then I_{Δ_1} and I_{Δ_2} are principal ideals, and every Gröbner basis G_1 and G_2 consist of one binomial each, namely the binomial

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

In this case $Ext(G_1 \to I_\Delta)$ consists of sixteen binomials of the form

$$\begin{bmatrix} 0 & 0 & 0 & r_1 \\ 0 & 1 & 1 & r_2 \\ 1 & 0 & 1 & r_3 \\ 1 & 1 & 0 & r_4 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 1 & r_3 \\ 0 & 1 & 0 & r_4 \\ 1 & 0 & 0 & r_1 \\ 1 & 1 & 1 & r_2 \end{bmatrix}.$$

with all $r_i \in \{0, 1\}$. Similarly, $Ext(G_2 \to I_\Delta)$ consists of 16 binomials of the form

$$\begin{bmatrix} r_1 & 0 & 0 & 0 \\ r_2 & 0 & 1 & 1 \\ r_3 & 1 & 0 & 1 \\ r_4 & 1 & 1 & 0 \end{bmatrix} - \begin{bmatrix} r_1 & 0 & 0 & 1 \\ r_2 & 0 & 1 & 0 \\ r_3 & 1 & 0 & 0 \\ r_4 & 1 & 1 & 1 \end{bmatrix}.$$

where $r_i \in \{0, 1\}$. Finally $\mathcal{F}(\{1, 2, 3\}, \{2, 3, 4\})$ consists of the four binomials of the form

$$\begin{bmatrix} 0 & r_1 & r_2 & 0 \\ 1 & r_1 & r_2 & 1 \end{bmatrix} - \begin{bmatrix} 0 & r_1 & r_2 & 1 \\ 1 & r_1 & r_2 & 0 \end{bmatrix}.$$

with $r_i \in \{0, 1\}$. The union of these three sets is a (not necessarily reduced) Gröbner basis for I_Δ. $\square$

Remark 4.19 We point out that Theorem 4.17 can be applied even when one of Δ_1 or Δ_2 does not give rise to a model in the usual sense. For example, the simplicial complex $\Delta = \{\{1,2\},\{1,3\},\{2,3\},\{2,4\}\}$ is reducible with decomposition $\Delta_1 = \{\{1,2\},\{1,3\},\{2,3\}\}$, $\Delta_2 = \{\{2,4\}\}$, and $S = \{2\}$. In such cases (when one of Δ_1 or Δ_2 is a simplex) we think of such a model as the "trivial model" with $\emptyset$ as its Gröbner basis. □

The following corollaries to Theorem 4.17 show that certain properties of the induced submodels are passed on to the complete model Δ.

Corollary 4.20 *Let Δ be a reducible model with induced submodels Δ_1 and Δ_2, and Gröbner bases G_1 and G_2. If m_1 and m_2 are a bound on the maximum degree elements in G_1 and G_2 respectively, then the degree of any element in G, where G is the Gröbner basis of I_Δ constructed in Theorem 4.17, is at most $m = \max(m_1, m_2, 2)$.*

Proof. If both Δ_1 and Δ_2 are trivial models, then $m = 2$ by Theorem 4.13. So suppose that at least one of the induced submodels is not trivial. The maximum degree of any binomial in $Ext(G_i \to I_\Delta)$ is the same as the maximum degree in any binomial in G_i and all the binomials of $\mathcal{F}(|\Delta_1|, |\Delta_2|)$ have degree two, and Theorem 4.17 implies the corollary. □

For the next corollary, we assume the reader is familiar with triangulations of vector configurations and their connection to toric algebra as outlined in [11, Chapter 8].

Corollary 4.21 *Let Δ be a reducible model with induced submodels Δ_1 and Δ_2. Suppose that the vector configurations A_{Δ_1} and A_{Δ_2} have regular unimodular triangulations. Then A_Δ also has a regular unimodular triangulation.*

Proof. A_{Δ_1} and A_{Δ_2} have regular unimodular triangulations if and only if there are respective reduced Gröbner bases G_1 and G_2 with the property that all the leading terms are square-free. The Ext operator preserves this square-free property: all the binomials in $Ext(G_1 \to I_\Delta) \cup Ext(G_2 \to I_\Delta) \cup \mathcal{F}(|\Delta_1|, |\Delta_2|)$ also have square-free initial terms. Hence A_Δ has a regular unimodular triangulation. □

At this point, a word should be said about decomposable models. Decomposable models are a special class of reducible model which have received much attention because of their simple structure.

Definition 4.22 A simplicial complex Δ is called *decomposable* if it has a decomposition (Δ_1, S, Δ_2) as in Definition 4.1, where Δ_1 and Δ_2 are either decomposable or simplices. We call a model decomposable if its underlying simplicial complex is decomposable.

Example 4.23 Clearly, any simplicial complex with exactly two facets is decomposable. Consider the simplicial complex $\Delta = \{\{1,2\},\{2,3\},\{3,4\}\}$. This simplicial complex has the decomposition $(\{\{1,2\},\{2,3\}\},\{3\},\{\{3,4\}\})$. Observe that $\{\{1,2\},\{2,3\}\}$ is decomposable since it has only two facets and $\{\{3,4\}\}$ is a simplex. This implies that Δ is decomposable.

On the other hand, the simplicial complex given by $\Delta = \{\{1,2\},\{1,3\},\{2,3\},\{3,4\}\}$ is reducible but not decomposable since the subcomplex $\{\{1,2\},\{1,3\},\{2,3\}\}$ has no nontrivial decomposition and is not a simplex. □

Theorem 4.17 and its corollaries have important consequences for decomposable models. In particular, we give an alternate proof of a result that appears in [7].

Theorem 4.24 *If Δ is a decomposable simplicial complex, then I_Δ has a quadratic square-free initial ideal.*

Proof. The proof is by induction on the number of facets of Δ. If Δ has exactly two facets the result follows by Theorem 4.13. Assume the result is true for all decomposable simplicial complexes with $r-1$ facets, and let Δ be a decomposable simplicial complex with r facets. Δ has a decomposition (Δ_1, S, Δ_2) and Δ_1 and Δ_2 both have fewer than r facets. The corresponding ideals I_{Δ_1} and I_{Δ_2} must have quadratic square-free initial ideals by assumption, and Corollary 4.20 guarantees that I_Δ also has this property. □

Corollary 4.25 *The vector configuration A_Δ is* normal *whenever Δ is a decomposable simplicial complex. In other words, if u is a vector of integer marginals lying in the cone spanned by the vertices of P_Δ, then there exists a nonnegative integer table with these marginals.*

Proof. Since I_Δ has a square-free initial ideal, A_Δ has a regular unimodular triangulation and hence it is normal. The second statement is simply a restatement of the definition of normality. □

5 The Dimension of Fibers

In this section, we answer a question posed by Vlach [13, Question 12]. Suppose Δ is the boundary of an $(n-1)$-simplex and $d = (d_1, \ldots, d_n)$. Every $b \in P_\Delta$ gives the fiber $Q(b) := \pi_\Delta^{-1}(b) \cap \mathbb{R}_+^{d_1 d_2 \cdots d_n}$ which is a polytope. The dimension of this polytope is at most $\prod_{i=1}^n (d_i - 1)$. It is not clear whether one can find $b \in P_\Delta$ such that all integers $0 \le k \le \prod_{i=1}^n (d_i - 1)$ are realized as the dimension of $Q(b)$. In fact, in general this is not true.

Example 5.1 Let $\pi : \mathbb{R}^4 \longrightarrow \mathbb{R}^2$ be given by the matrix $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix}$. Then $dim(Q(b))$ is at most two, and we can only find fibers of dimension zero or two, but not one. □

Theorem 5.2 *Let Δ be the boundary of the $(n-1)$-simplex and $d = (d_1, \ldots, d_n)$. Then for each $0 \leq k \leq \prod_{i=1}^{n}(d_i - 1)$ there exists $b \in P_\Delta$ such that dim$(Q(b)) = k$.*

We will prove this theorem by induction on n. We will prove the base case for $n = 2$ separately. In this case Δ consists of just two isolated vertices, and we assume $(d_1, d_2) = (m, n)$. Under this model, the fibers $Q(b)$ are *transportation polytopes*. We first recall some facts about the transportation polytopes and their vertices. The vertices of a transportation polytope $Q(b)$ correspond to the feasible spanning trees of the complete bipartite graph $K_{m,n}$. In other words, those spanning trees T such that

$$\sum_{\{j:(i,j)\in T\}} x_{ij} = b_i,\ 1 \leq i \leq m \quad \text{and} \quad \sum_{\{i:(i,j)\in T\}} x_{ij} = b_{m+j},\ 1 \leq j \leq n$$

where $x_{ij} \geq 0$, correspond to the vertices of $Q(b)$. If T corresponds to a degenerate vertex, i.e. $x_{pq} = 0$ for some $(p,q) \in T$, then more than one spanning tree will correspond to this vertex. Also recall that if v is a vertex of an arbitrary polytope Q, then the dimension of the subspace generated by

$$\{v - v_i :\ v_i \text{ is connected to } v \text{ with an edge of } Q\}$$

is the dimension of Q. Given a feasible spanning tree T_v corresponding to a vertex v of $Q(b)$, we obtain the neighboring vertices of v as follows. Every edge $(p,q) \in K_{m,n} \setminus T$ determines a unique cycle C that uses only (p,q) and edges from T. This cycle has even number of edges, and we assume we traverse this cycle starting with (p,q). Now let δ be the minimum of x_{ij} where (i,j) is an even edge of C. We modify $x'_{ij} = x_{ij}$ if $(i,j) \notin C$, $x'_{ij} = x_{ij} + \delta$ if (i,j) is an odd edge of C and $x'_{ij} = x_{ij} - \delta$ if (i,j) is an even edge of C. This gives us a new feasible spanning tree T', and T' corresponds to a neighboring vertex of v if and only if $\delta > 0$. It is also known that transportation polytopes are simple polytopes and hence the dimension of a transportation polytope is *equal* to the number of vertices that are neighbors of a fixed vertex v. With all this background we can now prove the base case for our induction.

Proposition 5.3 *For any fixed positive integers m and n there are transportation polytopes $Q(b)$ such that dim$(Q(b)) = k$ for all $0 \leq k \leq (m-1)(n-1)$.*

Proof. For any given k we will construct a $Q(b)$ and a vertex v of $Q(b)$ with exactly k neighboring vertices. We assume $min\{m,n\} = m$. First we consider $(m-1)(n-1) \geq k \geq (m-1)(n-2)+1$. For each $0 \leq t \leq m-2$, we construct a spanning tree T_t by giving the edges of this tree: $\{(1, n-t), (2, n-t), \ldots, (m-t, n-t), (m-t, 1), (m-t, 2), \ldots, (m-t, n-t-1)\}$ together with $\{(m-t+1, n-t), (m-t+1, n-t+1), (m-t+2, n-t+1), (m-t+2, n-t+2), \ldots, (m, n-1), (m, n)\}$. We will set $x_{ij} = 0$ for all $(i,j) \notin T$, $x_{ij} = 1$ for all $(i,j) \in T$ except $(m-t, n-t)$ and $x_{m-t,n-t} = 0$. We refer to Figure 1 to clarify this construction for the case $m = 4$, $n = 5$, and $t = 2$. A cycle C that is created by adding an edge $(p,q) \notin T$ contains $(m-t, n-t)$ if and only if $q \leq n-t$ or, $p = m-t$ and $q \geq n-t+1$, and only in the second type of cycles the edge $(m-t, n-t)$ is an even edge. Hence except

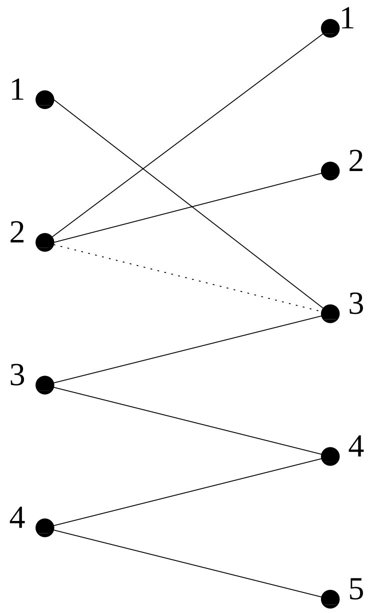

Figure 1: The tree construction in the proof of Proposition 5.3

this last kind of cycles any other cycle will give a neighboring vertex. Since there are $k = (m-1)(n-1) - t$ such good cycles, we get transportation polytopes of dimension k for each $(m-1)(n-1) \geq k \geq (m-1)(n-2)+1$. Now we can assume that we have a $m \times (n-1)$ transportation problem, since such transportation problems are also $m \times n$ transportation problems with $b_{m+n} = 0$. And we repeat the above to get all transportation polytopes of appropriate dimensions. $\square$

Proof of Theorem 5.2. When $n = 2$ the above proposition gives the result. Suppose Δ is the boundary of the $(n-1)$-simplex with $d = (d_1, \ldots, d_n)$. By induction, we have the result for the boundary of the $(n-2)$-simplex with $d' = (d_1, \ldots, d_{n-1})$. In order to get fibers of dimension $\prod_{i=1}^{n}(d_i - 1) - t$ where $0 \leq t \leq \prod_{i=1}^{n-1}(d_i - 1)$ we create the following tables Γ_t for each t. We let Γ to be the $d_1 \times \cdots \times d_{n-1} \times (d_n - 1)$ table with all cell entries equal to one. The fiber that Γ is contained in has dimension $D = (d_n - 2)\prod_{i=1}^{n-1}(d_i - 1)$. In other words, there are D linearly independent vectors that could be added to Γ without leaving this fiber. We will look at Γ as $(d_n - 1)$ copies of a $d_1 \times \cdots \times d_{n-1}$ table "stacked" on top of each other, i.e. $\Gamma = \bigoplus_{i=1}^{d_n-1} \Gamma^i$. Now we let Γ' be a $d_1 \times \cdots \times d_{n-1}$ table where the fiber that Γ' is contained in has dimension $D' = \prod_{i=1}^{n-1}(d_i - 1) - t$. Similarly this means there are D' linearly independent vectors that could be added to Γ' without leaving this fiber. Such a Γ' exists by induction. Now Γ_t will be the table obtained by stacking Γ' on top of Γ, i.e. $\Gamma_t = \Gamma' \oplus \Gamma = \Gamma' \oplus (\bigoplus_{i=1}^{d_n-1} \Gamma^i)$. We observe that any vector V that is among the D vectors used for Γ could be extended to the vector $0 \oplus V$ for Γ_t. Similarly, any vector V' among the D' vectors used for Γ' could be extended to the vector $V' \oplus (-V') \oplus (\bigoplus_{i=1}^{d_n-2} 0)$. By our construction these $D + D' = \prod_{i=1}^{n}(d_i - 1) - t$ vectors are linearly independent, and adding any of these vectors to Γ_t will not result in a table outside the fiber of Γ_t. This means the dimension of the fiber of Γ_t is at least $\prod_{i=1}^{n}(d_i-1)-t$. However, if $V_1 \oplus V_2 \cdots \oplus V_{d_n}$ is a vector which could be added to Γ_t without leaving its fiber, then V_1 is a vector that could be added to Γ' without leaving the fiber of Γ'. In particular, V_1 is a linear combination of vectors from among the D' vectors above. Substracting the "extended" version of this linear combination from $V_1 \oplus V_2 \cdots \oplus V_{d_n}$ we get a vector of the form $0 \oplus V_2' \oplus V_3 \cdots \oplus V_{d_n}$. We observe that $V_2' \oplus V_3 \cdots \oplus V_{d_n}$ is a vector that could be added to Γ without leaving its own fiber, and hence is a linear combination of the D vectors.

When we construct fibers of dimension $\prod_{i=1}^{n}(d_i-1) - k\prod_{i=1}^{n-1}(d_i-1) - t$ where $0 \leq t \leq \prod_{i=1}^{n-1}(d_i-1)$, the table Γ_t will consist of $d_n - k - 1$ copies of a $d_1 \times \cdots d_{n-1}$ table with entries equal to one, one copy of Γ' as above, and k copies of a $d_1 \times \cdots \times d_{n-1}$ table of all zeros as its entries. When $k = d_n - 1$, we again refer to the inductive step. $\square$

References

[1] R. Christensen. *Log-linear Models.* Springer-Verlag, New York, 1990.

[2] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms.* Springer-Verlag, New York, 1996. Second edition.

[3] P. Diaconis, D. Eisenbud, and B. Sturmfels. Lattice walks and primary decomposition. In *Mathematical essays in honor of Gian-Carlo Rota (Cambridge, MA, 1996)*, pages 173–193. Birkhäuser Boston, Boston, MA, 1998.

[4] P. Diaconis and B. Sturmfels. Algebraic algorithms for sampling from conditional distributions. *Annals of Statistics*, 26:363–397, 1998.

[5] A. Dobra. *Statistical Tools for Disclosure Limitation in Multi-way Contingency Tables.* PhD thesis, Department of Statistics, Carnegie Mellon University, 2002.

[6] A. Dobra and S. Sullivant. Markov bases for reducible sets of tables. Manuscript, 2002.

[7] D. Geiger, C. Meek, and B. Sturmfels. On the toric algebra of graphical models. Manuscript, 2002.

[8] S. Hoşten and J. Shapiro. Primary decomposition of lattice basis ideals. *Journal of Symbolic Computation*, 29(4-5):625–639, 2000.

[9] S. Lauritzen. *Graphical Models.* Oxford University Press, New York, 1996.

[10] J. De Loera and S. Onn. The complexity of three-way statistical tables. Manuscript, 2002.

[11] B. Sturmfels. *Gröbner Bases and Convex Polytopes.* American Mathematical Society, Providence, RI, 1995.

[12] A. Takken. *Monte Carlo Goodness-of-Fit Tests for Discrete Data.* PhD thesis, Stanford University, 1999.

[13] M. Vlach. Conditions for the existence of solutions of the three-dimensional planar transportation problem. *Discrete Applied Mathematics*, 13:61–78, 1986.

[14] G. Ziegler. *Lectures on Polytopes.* Springer Verlag, Heidelberg, 1995.