MATHEMATICAL ENGINEERING
TECHNICAL REPORTS

Markov Degree of the Three-State Toric Homogeneous Markov Chain Model

David HAWS, Abraham MARTIN DEL CAMPO,
Akimichi TAKEMURA and Ruriko YOSHIDA

METR 2012–09 May 2012

DEPARTMENT OF MATHEMATICAL INFORMATICS
GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF TOKYO
BUNKYO-KU, TOKYO 113-8656, JAPAN

WWW page: http://www.keisu.t.u-tokyo.ac.jp/research/techrep/index.html

Markov degree of the three-state toric homogeneous Markov chain model

David Haws, Abraham Martín del Campo, Akimichi Takemura, Ruriko Yoshida*

University of Kentucky, Department of Statistics

Abstract

Markov chain models had proved to be useful tools in many fields, such as physics, chemistry, information sciences, economics, finances, mathematical biology, social sciences, and statistics for analyzing data. A discrete time Markov chain is often used as a statistical model from a random physical process to fit the observed data. A time-homogeneous Markov chain is a process that each transition probability from a state to a state does not depend on time. It is important to test if the assumption of the time-homogeneity of the chain fits the observed data. In 2011, Hara and Takemura suggested a Markov Chain Monte Carlo (MCMC) approach to a goodness-of-fit test using Markov bases on the *toric homogeneous Markov chain (THMC) model* and gave a full description of the Markov bases for the two-state THMC model which does not depend on time T. In this paper, we provide a bound on the degree of the Markov bases for the three-state THMC model (without loops and initial parameters), when the transition probabilities of the Markov chains are assumed to be independent of the time. Our proof is based on a result due to Sturmfels, who gave a bound on the degree for the generators of toric ideals, provided the normality of the corresponding toric variety. In our setting, we proved the normality of the semigroup generated by the columns of the design matrix associated to the THMC model by studying the geometric properties of the polytope associated to the design matrix of the model. Moreover, we give a complete description of the facets of this polytope, which does not depend on the time.

Keywords: Markov bases, toric homogeneous Markov chains, polyhedrons, semigroups

1. Introduction

A discrete time Markov chain, X_t for $t = 1, 2, \ldots$, is a stochastic process with the Markov property, that is $P(X_{t+1} = y|X_1 = x_1, \ldots, X_{t-1} = x_{t-1}, X_t = x) = P(X_{t+1} = y|X_t = x)$ for any states x, y. Discrete time Markov chains have applications in several fields, such as physics, chemistry, information sciences, economics, finances, mathematical biology, social sciences, and statistics [8]. In this paper, we consider a discrete time Markov chain X_t, with $t = 1, \ldots, T$ ($T \geq 3$), over a finite set of states $[S] = \{1, \ldots, S\}$. In this paper we focus on $S = 3$.

Discrete time Markov chains are often used in statistical models to fit the observed data from a random physical process. Sometimes, in order to simplify the model, it is convenient to consider time-homogeneous Markov chains, where the transition probabilities do not depend on the time, in other words, when

$$P(X_{t+1} = y|X_t = x) = P(X_2 = y|X_1 = x) \text{ for all states } x, y \text{ and for any } t = 1 \ldots, T.$$

*ruriko.yoshida@uky.edu

In order for a statistical model to reflect the observed data, it is necessary to verify if the model fits the date via the goodness-of-fit test. For instance, for the time-homogeneous Markov chain model, it is necessary to test if the assumption of time-homogeneity fits the observed data. In this paper, we study some properties of these algebraic relations for the toric homogeneous Markov chain (THMC) model, which is a slight generalization of the time-homogeneous Markov chain model, and which we explain as follows.

Let $\mathbf{w} = s_1 \cdots s_T$ denote a word of length T on states $[S]$. Let $p(\mathbf{w})$ denote the likelihood of observing the word $\mathbf{w}$. Since the time-homogeneous Markov chain model assumes that the transition probabilities do not depend on time, we can write the likelihood as the product of probabilities

$$p(\mathbf{w}) = \pi_{s_1} p_{s_1,s_2} \cdots p_{s_{T-1},s_T}, \tag{1.1}$$

where, π_{s_i} indicates the initial distribution at the first state, and p_{s_i,s_j} are the transition probabilities from state s_i to s_j. In usual time-homogeneous Markov chain model it is assumed that the row sums of the transition probabilities are one: $\sum_{j=1}^{S} p_{i,j} = 1$, $\forall i \in [S]$. On the other hand, in the toric homogeneous Markov model (1.1), the parameters $p_{i,j}$ are free and we do not assume that the row sums of the transition probabilities are one. This simplifies the model as we can disregard some information from it. Also in many cases the parameters π_{s_1} for the initial distribution are known, or sometimes these parameters are all constant, namely $\pi_1 = \pi_2 = \cdots = \pi_S = c$; in this situation it is no longer necessary to take them in consideration for the model. Another simplification that arises from practice is when the only transition probabilities considered are those between two different states, i.e. when $p_{i,j} = 0$ whenever $i = j$; this situation is referred as the THMC model without self-loops.

In [11], Takemura and Hara suggested a Markov Chain Monte Carlo (MCMC) approach for the goodness-of-fit test using *Markov bases* of the THMC model. A Markov basis is a set of moves between data sets with the same sufficient statistics so that the transition graph for the MCMC is guaranteed to be connected for any observed value of the sufficient statistics (see Section 2.1 and [8]). A Markov basis can be encoded by polynomial relations among the probabilities $p_{i,j}$. The set of all these algebraic relations is the *toric ideal* associated to the model.

In [11], the authors provided a full description of the Markov bases for the THMC model in two states (i.e. when $S = 2$) that does not depend on T. Inspired by their work, we study the algebraic and polyhedral properties of the Markov bases of the three-state THMC model without initial parameters and without self-loops for any time $T \geq 3$. As a result, we proved that the Markov bases of the model consist of binomials of degree at most 6. Our proof relies on the normality of the semigroup generated by the columns of the design matrix for the three-state THMC model for any time $T \geq 3$, which we settle using a complete description of the facets of the convex hull generated by the columns of the design matrix.

The outline of this paper is as follows. In Section 2, we recall some definitions from Markov bases theory. In Section 3, we explicitly describe the hyperplane representation of the convex hull generated by the columns of the design matrix for the three-state THMC model for any time $T \geq 4$. Then, in Section 4 we show the normality of the semigroup generated by the columns of the design matrix for the three-state THMC model for any time $T \geq 3$. Finally, using these results, we present the proof of our bound on the degree of the Markov basis in Section 5, and we finalize that section with some observations based on the analysis of our computational experiments.

2. Notation

Let $\langle S\rangle^T$ be the set of all words of length T on states $[S]$ such that every word has no self-loops; that is, if $\mathbf{w} = (s_1, \ldots, s_T) \in \langle S\rangle^T$ then $s_i \neq s_{i+1}$ for $i = 1, \ldots, T-1$. We define $\mathcal{P}^*(\langle S\rangle^T)$ to be the set of all multisets of words in $\langle S\rangle^T$.

Let $\mathbb{V}\left(\langle S\rangle^T\right)$ be the real vector space with basis $\langle S\rangle^T$ and note that $\mathbb{V}\left(\langle S\rangle^T\right) \cong \mathbb{R}^{S(S-1)^T}$. We recall some definitions from the book of Pachter and Sturmfels [7]. Let $A = (a_{ij})$ be a non-negative integer $d \times m$ matrix with the property that all column sums are equal:

$$\sum_{i=1}^{d} a_{i1} = \sum_{i=1}^{d} a_{i2} = \cdots = \sum_{i=1}^{d} a_{im}.$$

Write $A = [a_1\ a_2\ \cdots\ a_m]$ where a_j are the column vectors of A and define $\theta^{\mathbf{a}_j} = \prod_{i=1}^{d} \theta_i^{a_{ij}}$ for $j = 1, \ldots, m$. The *toric model* of A is the image of the orthant $\mathbb{R}^d_{\geq 0}$ under the map

$$f : \mathbb{R}^d \to \mathbb{R}^m, \quad \theta \mapsto \frac{1}{\sum_{j=1}^{m} \theta^{a_j}} \left(\theta^{a_1}, \ldots, \theta^{a_m}\right).$$

Here we have d parameters $\theta = (\theta_1, \ldots, \theta_d)$ and a discrete state space of size m. In our setting, the discrete space will be the set of all possible words on $[S]$ of length T without self-loops ($\langle S\rangle^T$) and we can think of $\theta_1, \ldots, \theta_d$ as the probabilities $p_{1,2}, p_{1,3}, \ldots, p_{S-1,S}$.

In this paper, we focus on the THMC model without initial parameters and with no self-loops in three states, (i.e., $S = 3$), which is parametrized by 6 positive real variables: p_{12}, p_{13}, p_{21}, p_{23}, p_{31}, $p_{3,2}$. Thus, the number of parameters is $d = 6$ and the size of the discrete space is $m = 6^{T-1}$, which is precisely the number of words in $\langle 3\rangle^T$. The model we study is thus the toric model represented by the $6 \times 6^{T-1}$ matrix A^T, which will be referred to as the *design matrix* for the model on 3 states with time T. The rows of A^T are indexed by elements in $\langle 3\rangle^2$ and the columns are indexed by words in $\langle 3\rangle^T$. The entry of A^T indexed by row $\sigma_1\sigma_2 \in \langle 3\rangle^2$, and column $\mathbf{w} = (s_1, \ldots, s_T) \in \langle 3\rangle^T$ is equal to the cardinality of the set $\{\, i \in \{1, \ldots, T-1\} \mid \sigma_1\sigma_2 = s_i s_{i+1} \,\}$.

Example 2.1. *Ordering $\langle 3\rangle^2$ and $\langle 3\rangle^T$ lexicographically, and letting $T = 4$, the matrix A^4 is:*

	1212	1213	1231	1232	1312	1313	1321	1323	2121	2123	2131	2132	2312	2313	2321	2323	3121	3123	3131	3132	3212	3213	3231	3232
12	2	1	1	1	1	0	0	0	1	1	1	0	0	0	0	0	1	1	1	0	0	0	0	0
13	0	1	1	0	0	2	1	1	0	0	0	1	1	1	0	0	0	0	0	1	1	1	0	0
21	1	1	0	0	0	0	1	0	2	1	0	1	1	0	1	0	1	1	0	1	0	0	0	0
23	0	0	0	1	1	0	0	1	0	1	1	0	0	1	1	2	0	0	1	0	0	0	1	1
31	0	0	1	1	0	1	0	0	0	0	1	1	0	1	0	0	1	0	1	0	2	1	1	0
32	0	0	0	0	1	0	1	1	0	0	0	0	1	0	1	1	0	1	0	1	0	1	1	2

2.1. Sufficient statistics, ideals, and Markov basis

Let A^T be the design matrix for the THMC model without initial parameters and with no self-loops. The column of A^T indexed by $\mathbf{w} \in \langle 3\rangle^T$ is denoted by $\mathbf{a}^T_{\mathbf{w}}$. Thus, by extending linearly, the map $A^T : \mathbb{V}(\langle 3\rangle^T) \to \mathbb{R}^6$ is well-defined.

Let $W = \{w_1, \ldots, w_N\} \in \mathcal{P}^*(\langle 3\rangle^T)$ where we regard W as observed data which can be summarized in the *data vector* $\mathbf{u} \in \mathbb{N}^{6^{T-1}}$. We index $\mathbf{u}$ by words in $\langle 3\rangle^T$, so the coordinate representing for the word $\mathbf{w}$ in the vector $\mathbf{u}$ is denoted by $u_{\mathbf{w}}$, and its value is the number of words in W equal to $\mathbf{w}$. Note since A^T is linear then $A^T\mathbf{u}$ is well-defined. We also adopt the notation $A^T(W) := A^T\mathbf{u}$. For W from $\mathcal{P}^*(\langle 3\rangle^T)$, let $\mathbf{u}$ be its data vector, the *sufficient statistics* for the model are stored in the vector $A^T\mathbf{u}$. Often the data vector $\mathbf{u}$ is also referred to as a *contingency table*, in which case $A^T\mathbf{u}$ is referred to as the *marginals*.

The design matrix A^T above defines a toric ideal which is of central interest in this paper, as their set of generators are in bijection with the Markov bases. The toric ideal I_{A^T} is defined as the kernel of the homomorphism of polynomial rings $\psi : \mathbb{C}[\{\, p(\mathbf{w}) \mid \mathbf{w} \in \langle S\rangle^T \,\}] \to \mathbb{C}[\{\, p_{ij} \mid i, j \in [3],\ i \neq j \,\}]$ defined by $\psi(p(\mathbf{w})) = p_{s_1,s_2} \cdots p_{s_{T-1},s_T}$, where $\{\, p(\mathbf{w}) \mid \mathbf{w} \in \langle S\rangle^T \,\}$ is regarded as a set of indeterminates.

The set of all contingency tables (data vectors) satisfying a given set of marginals $\mathbf{b} \in \mathbb{Z}^d_{\geq 0}$ is called a *fiber* which we denote by $\mathcal{F}_\mathbf{b} = \{\, \mathbf{x} \in \mathbb{Z}^m_{\geq 0} \mid A^T(\mathbf{x}) = \mathbf{b} \,\}$. A *move* $\mathbf{z} \in \mathbb{Z}^m$ is an integer vector satisfying $A^T(\mathbf{z}) = 0$. A *Markov basis* for our model defined by the design matrix A^T is defined as a finite set $\mathcal{Z}$ of moves satisfying that for all $\mathbf{b}$ and all pairs $\mathbf{x}, \mathbf{y} \in \mathcal{F}_\mathbf{b}$ there exists a sequence $\mathbf{z}_1, \ldots, \mathbf{z}_K \in \mathcal{Z}$ such that

$$\mathbf{y} = \mathbf{x} + \sum_{k=1}^{K} \mathbf{z}_k, \quad \text{with } \mathbf{x} + \sum_{k=1}^{l} \mathbf{z}_k \geq \mathbf{0}, \quad \text{for all } l = 1, \ldots, K.$$

A *minimal Markov basis* is a Markov basis which is minimal in terms of inclusion. See Diaconis and Sturmfels[2] for more details on Markov bases and their toric ideals.

2.2. State Graph

We give here a useful tool to visualize multisets of $\mathcal{P}^*(\langle 3\rangle^T)$. Given any multiset $W \in \mathcal{P}^*(\langle 3\rangle^T)$ we consider the directed multigraph called the *state graph* $G(W)$. The vertices of $G(W)$ are given by the states [3] and the directed edges $i \to j$ are given by the transitions from state i to j in $\mathbf{w} \in W$. Thus, we regard $\mathbf{w} \in W$ as a path with $T-1$ edges (steps, transitions) in $G(W)$.

See Figure 1 for an example of the state graph $G(W)$ of the multiset $W = \{(12132), (12321)\}$ of paths with length 4. Notice that the state graph in Figure 1 is the same for another multiset of paths $\overline{W} = \{(13212), (21232)\}$.

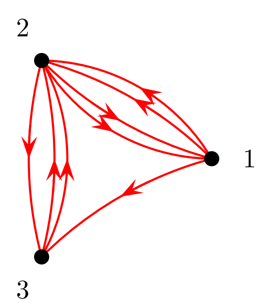

Figure 1: The state graph $G(W)$ of $W = \{(12132), (12321)\}$. Also the state graph $G(\overline{W})$ where $\overline{W} = \{(13212), (21232)\}$.

From the definition of state graph it is clear that it records the transitions in a given multiset of words and we state the following proposition.

Proposition 2.2 (Proposition 2.1 in [5]). *Let A be the design matrix for the THMC, and $W, \overline{W} \in \mathcal{P}^*(\langle S\rangle^T)$. Then $A(W) = A(\overline{W})$ if and only if $G(W) = G(\overline{W})$.*

Throughout this paper we alternate between terminology of the multisets of words W and the graph it defines $G(W)$.

2.3. Semigroup and Smith Normal Form

Given an integer matrix $A \in \mathbb{Z}^{d\times m}$ we associate an integer lattice $\mathbb{Z}A = \{n_1\mathbf{a}_1 + \cdots + n_m\mathbf{a}_m \mid n_i \in \mathbb{Z}\}$. We can also associate the semigroup $\mathbb{N}A := \{n_1\mathbf{a}_1 + \cdots + n_m\mathbf{a}_m \mid n_i \in \mathbb{N}\}$. We say that the semigroup $\mathbb{N}A$ is *normal* when $\mathbf{x} \in \mathbb{N}A$ if and only if there exist $\mathbf{y} \in \mathbb{Z}^d$ and $\alpha \in \mathbb{R}^d_{\geq 0}$ such that $\mathbf{x} = A\mathbf{y}$ and $\mathbf{x} = A\alpha$. The set of vectors $A\alpha$ is called the *saturation* of $\mathbb{N}A$. See [6, 10] for more details on normality.

For an integer matrix $A \in \mathbb{Z}^{d\times m}$, we consider the Smith normal form D of A, which is a diagonal matrix D for which there exist unimodular matrices $U \in \mathbb{Z}^{d\times d}$ and $V \in \mathbb{Z}^{m\times m}$, such that $UAV = D$. The Smith normal form encodes the $\mathbb{Z}$-module structure of the abelian group $\mathbb{Z}A := \{n_1A_1 + \cdots + n_mA_m \mid n_i \in \mathbb{Z}\}$. Some additional material about the Smith normal form for matrices with entries over a PID can be found in the book of C. Yap [12]. The Smith normal form is important for studying the normality of the toric ideal associated to the model.

Proposition 2.3 (Proposition 3.2 in [5])**.** *Let $A^{S,T}$ be the design matrix for the THMC without initial parameters and no self-loops on $S > 1$ states with time $T > 0$. For $S \geq 3$ and $T \geq 4$, the Smith normal form of the design matrix $A^{S,T}$ is $D = \mathrm{diag}(1, \ldots, 1, T-1)$.*

3. Facets of the design polytope

3.1. Polytopes

We recall some necessary definitions from polyhedral geometry and we refer the reader to the book of Schrijver [9] for more details. The *convex hull* of $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\} \subset \mathbb{R}^n$ is defined as

$$\mathrm{conv}(\mathbf{a}_1, \ldots, \mathbf{a}_m) := \left\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \sum_{i=1}^{m} \lambda_i \mathbf{a}_i,\ \sum_{i=1}^{m} \lambda_i = 1,\ \lambda_i \geq 0 \right\}.$$

A *polytope* P is the convex hull of finitely many points. We say $F \subseteq P$ is a *face* of the polytope P if there exists a vector $\mathbf{c}$ such that $F = \arg\max_{\mathbf{x}\in P} \mathbf{c}\cdot\mathbf{x}$. Every face F of P is also a polytope. If the dimension of P is d, a face F is a *facet* if it is of dimension $d-1$. For $k \in \mathbb{N}$, we define the k-th dilation of P as $kP := \{\, k\mathbf{x} \mid \mathbf{x} \in P, \}$. A point $\mathbf{x} \in P$ is a *vertex* if and only if it can not be written as a convex combination of points from $P\backslash\{\mathbf{x}\}$.

The *cone* of $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\} \subset \mathbb{R}^n$ is defined as

$$\mathrm{cone}(\mathbf{a}_1, \ldots, \mathbf{a}_m) := \left\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \sum_{i=1}^{m} \lambda_i \mathbf{a}_i,\ \lambda_i \geq 0 \right\}.$$

Thus, cone(A) denote the cone over the columns of the matrix A. We are interested in the polytope given by the convex hull of the columns of the design matrices of our model. We define the *design polytope* P^T as the convex hull $\mathrm{conv}(A^T)$ of the columns of the design matrix A^T. Notice that in this case, P^T has dimension 6.

If $\mathbf{x} \in \mathbb{R}^6$, we index $\mathbf{x}$ by $\{\, ij \mid 1 \leq i, j \leq 3,\ i \neq j \,\}$. We define $\mathbf{e}_{ij} \in \mathbb{R}^6$ to be the vector of all zeros, except 1 at index ij. We also adopt the notation $x_{i+} := \sum_j x_{ij}$ and $x_{+i} := \sum_j x_{ji}$. For any $\mathbf{x} \in \mathbb{N}^6$ we can define a directed multigraph $G(\mathbf{x})$ on three vertices, where there are x_{ij} directed edges from vertex i to vertex j. One would like to identify the vectors $\mathbf{x} \in \mathbb{N}^6$ for which the graph $G(\mathbf{x})$ is a state graph. Nevertheless, observe that x_{i+} is the out-degree of vertex i and x_{+i} is the in-degree of vertex i with respect to $G(\mathbf{x})$.

We now give some properties which will be used later for describing the facets of the design polytope P^T given by the design matrix for our model, and to prove normality of the semigroup associated with the design matrix.

Proposition 3.1 (Proposition 5.1 in [5])**.** *Let A^T be the design matrix for the THMC without loops and initial parameters. If $\mathbf{x} \in \mathbb{Z}A^T \cap \text{cone}(A^T)$ then $\sum_{i\neq j} x_{ij} = k(T-1)$ for some $k \in \mathbb{N}$ and $|x_{i+} - x_{+i}| \leq k$ for all $i \in \{1, 2, 3\}$.*

Proposition 3.1 states that for $\mathbf{x} \in \mathbb{Z}A^T \cap \text{cone}(A^T)$ the multigraph $G(\mathbf{x})$ will have in-degree and out-degree bounded by $\|\mathbf{x}\|_1/(T-1)$ at every vertex. This implies nice properties when $\|\mathbf{x}\|_1 = (T-1)$. Recall that a path in a directed multigraph is *Eulerian* if it visits every edge only once.

Proposition 3.2 (Proposition 5.2 in [5])**.** *If G is a directed multigraph on three vertices, with no self-loops, $T-1$ edges, and satisfying*

$$|G_{i+} - G_{+i}| \leq 1 \qquad i = 1, 2, 3;$$

then, there exists an Eulerian path in G.

Note that every word $\mathbf{w} \in \mathcal{P}^*(\langle 3 \rangle^T)$ gives an Eulerian path in $G(\{\mathbf{w}\})$ containing all edges. Conversely, for every multigraph G with an Eulerian path containing all edges, there exists $\mathbf{w} \in \mathcal{P}^*(\langle 3 \rangle^T)$ such that $G(\{\mathbf{w}\})= G$. More specifically, $\mathbf{w}$ is the Eulerian path in $G(\{\mathbf{w}\})$. Throughout this paper we use the terms *path* and *word* interchangeably.

Lemma 3.3 (Lemma 5.2 in [5])**.** *Let A^T be the design matrix for the THMC. If $T \geq 4$, then* $\text{conv}(A^T) \cap \mathbb{Z}^6 = A^T$*, where the right hand side is taken as the set of columns of the matrix A^T.*

We define

$$H_{k(T-1)} := \left\{ \mathbf{x} \in \mathbb{R}^6 \mid \sum_{i\neq j} x_{ij} = k(T-1) \right\}.$$

Proposition 3.4 (Proposition 5.3 in [5])**.** *Let A^T be the design matrix for the THMC without initial parameters and no loops.*

1. *For $T \geq 4$ and $k \in \mathbb{N}$,*
$$k\,\text{conv}(A^T) = \text{cone}(A^T) \cap H_{k(T-1)}.$$
2. *For $T \geq 4$,*
$$\text{cone}(A^T) \cap \mathbb{Z}A^T = \bigoplus_{k=0}^{\infty} \left(k\,\text{conv}(A^T) \cap \mathbb{Z}^6 \right).$$

3.2. Facets

In this section we give the facets of the design polytope P^T for $S = 3$ and arbitrary T. Recall that vectors $\mathbf{c} \in \mathbb{R}^6$ are indexed as $[c_{12}, c_{13}, c_{21}, c_{23}, c_{31}, c_{32}]$. The proofs for the facets of P^T rely heavily on the state graph. Note that we give the facets of the design polytope in terms of equivalence classes under permutations of the labels $\{1, 2, 3\}$. For example, if $\mathcal{S}_3$ denote the set of permutations of the set $\{1, 2, 3\}$, and the vector $(c_{12}, c_{13}, c_{21}, c_{23}, c_{31}, c_{32})$ defines a facet, for any $\sigma \in \mathfrak{S}_3$, the vector $(c_{\sigma(1)\sigma(2)}, c_{\sigma(1)\sigma(3)}, c_{\sigma(2)\sigma(1)}, c_{\sigma(2)\sigma(3)}, c_{\sigma(3)\sigma(1)}, c_{\sigma(3)\sigma(2)})$ also defines a facet.

Notice that by definition, the transition count between states should be a non-negative integer, thus we have the following facet.

Proposition 3.5. *For any $T \geq 5$, the row vector*

$$\mathbf{c} = [1, 0, 0, 0, 0, 0]$$

defines a facet of P^T modulo $\mathfrak{S}_3$.

Recall that in general, to show that a vector $\mathbf{c}$ forms a facet of P^T, we need to show the following two things:

i) The elements of $\mathbf{c}\,\mathbf{a}_{\mathbf{w}}^T$ are non-negative for all $\mathbf{w} \in \langle 3 \rangle^T$.

ii) The dimension of linear subspace spanned by $\{\mathbf{a}_{\mathbf{w}}^T \mid \mathbf{c}\,\mathbf{a}_{\mathbf{w}}^T = 0\}$ is 5;

where $\mathbf{c}\,\mathbf{a}$ denotes the vector multiplication of the row vector $\mathbf{c}$ times the column vector $\mathbf{a}$.

Proposition 3.6. *For any $T \geq 5$, the row vector*

$$\mathbf{c} = [T, T, -(T-2), 1, -(T-2), 1]$$

defines a facet of P^T modulo $\mathfrak{S}_3$.

Proof. We now check i). Consider any particular word (path) $\mathbf{w} \in \langle 3 \rangle^T$ of length T with transition counts $x_{12}, x_{13}, x_{21}, x_{23}, x_{31}, x_{32}$. We need to show

$$T(x_{12} + x_{13}) + x_{23} + x_{32} \geq (T-2)(x_{21} + x_{31}). \tag{3.1}$$

Note that $x_{12} + x_{13}$ is the out-degree of vertex 1 and $x_{21} + x_{31}$ is the in-degree of vertex 1 with respect to the graph $G(\mathbf{w})$. By Proposition 3.1 the out-degree and the in-degree can differ by at most 1. Note that (3.1) trivially holds when $x_{12} + x_{13} \geq x_{21} + x_{31}$. Hence we only need to check the case $a = x_{12} + x_{13} = x_{21} + x_{21} - 1$. Now

$$T - 1 = x_{12} + x_{13} + x_{21} + x_{23} + x_{31} + x_{32} = 2a + 1 + x_{23} + x_{32}$$

or

$$2a + 2 + x_{23} + x_{32} - T = 0.$$

Hence the difference of two sides of (3.1) is written as

$$Ta + x_{23} + x_{32} - (T-2)(a+1) = x_{23} + x_{32} + 2(a+1) - T = 0. \tag{3.2}$$

This proves i).

Next we consider ii). Equation (3.1) can not hold with equality in the case $x_{12} + x_{13} > x_{21} + x_{31}$. Also (3.1) can not hold with equality in the case $x_{12} + x_{13} = x_{21} + x_{31} > 0$. Furthermore if $0 = x_{12} + x_{13} = x_{21} + x_{31}$, then the path entirely consists of edges between 2 and 3. Then $T - 1 = x_{32} + x_{23} > 0$ and (3.1) does not hold with equality. Hence the only remaining case is $x_{12} + x_{13} = x_{21} + x_{31} - 1$. But then from (3.2) we see that (3.1) holds with equality. Hence all paths $\mathbf{w}$ such that $x_{12} + x_{13} = x_{21} + x_{31} - 1$ satisfies $\mathbf{c}\,\mathbf{a}_{\mathbf{w}}^T = 0$. We now give five such paths with linearly independent sufficient statistics, which depends on T mod 3 and T mod 2.

If T even, consider

$$3131\cdots 131,\ 2121\cdots 121,\ 3232\cdots 3231$$

If T odd, consider

$$23131\cdots 31,\ 32121\cdots 21,\ 2323\cdots 231$$

If $T \equiv 0 \pmod 3$, consider

$$321321\cdots 321,\ 231231\cdots 231$$

If $T \equiv 2 \pmod 3$, consider

$$213213\cdots 2131,\ 312312\cdots 3121$$

Finally if $T \equiv 1 \pmod 3$, put the loop 232 or 323 in front of the word above for the value of $T \equiv 2 \pmod 3$.

We need to show that the sufficient statistics of these paths are linearly independent. For example, consider the case $T = 6k$. Then the sufficient statistics are given by the vectors

$$[0, 3k-1, 0, 0, 3k, 0],\ [3k-1, 0, 3k, 0, 0, 0],\ [0, 0, 0, 3k-1, 1, 3k-1],\\ [0, 2k-1, 2k, 0, 0, 2k],\ [0, 2k, 0, 2k-1, 2k, 0].$$

For $k \geq 2$, the linear independence of these five vectors can be easily verified. Other cases $T \equiv r \pmod 6$ can be similarly handled.n □

In a similar fashion, we prove the following propositions.

Proposition 3.7. *For any $T = 2k+1$ odd, $T \geq 5$, the row vector*

$$\mathbf{c} = [1, 1, -1, -1, 1, 1]$$

defines a facet of P^T modulo $\mathfrak{S}_3$.

Proof. Consider

$$x_{12} + x_{13} + x_{31} + x_{32} \geq x_{21} + x_{23}. \tag{3.3}$$

We can "merge" two vertices 1 and 3 as a virtual vertex 4 and consider 4 as a single vertex. Then the resulting graph has only two vertices (2 and 4). Then $x_{12} + x_{32}$ is the out-degree of this vertex 4. If $x_{12} + x_{32} \geq x_{21} + x_{23}$ then the inequality is trivial. Consider the case $x_{12} + x_{32} = x_{21} + x_{23} - 1$. We need to show that in this case we have $x_{13} + x_{31} \geq 1$. By contradiction assume that $x_{13} + x_{31} = 0$. Then the path of odd length is like

$$2424242.$$

However in this case $x_{12} + x_{32} = x_{21} + x_{23}$, which is a contradiction. Therefore we have proved that (3.3) holds for any path $\mathbf{w}$.

We now check for which path (3.3) holds with equality. The first case is $0 = x_{13} + x_{31}$. Then as we saw above we have $x_{12} + x_{32} = x_{21} + x_{23}$. Other case is $1 = x_{13} + x_{31}$, i.e. either $x_{13} = 1, x_{31} = 0$ or $x_{31} = 1, x_{13} = 0$. In the former case the path is like

$$2121323$$

and in the latter case the path is like 32323121. We claim that for $T = 2k+1 \geq 5$ the paths

$$121\cdots121,\ 232\cdots232,\ 212\cdots2123,\ 1323\cdots232,\ 3121\cdots212$$

give sufficient statistics that are linearly independent and hold with equality for Equation (3.3). The sufficient statistics for the paths above are

$$[k,0,k,0,0,0],\ [0,0,0,k,0,k],\ [k-1,1,k,0,0,0],\ [0,1,0,k-1,0,k],\ [k,0,k-1,0,1,0].$$

For $k \geq 2$, one can check the linear independence of these 5 vectors. □

We now consider any three consecutive transitions, or a path for $T = 4$. Let $\tilde{x}_{ij}$, $1 \leq i \neq j \leq 3$, be transition counts of these three transitions.

Lemma 3.8. *Let i, j, k be distinct (i.e. $\{i,j,k\} = \{1,2,3\}$). Then*

$$\tilde{x}_{ij} + \tilde{x}_{jk} + \tilde{x}_{ik} \geq 1.$$

Proof. Suppose that $\tilde{x}_{ij} + \tilde{x}_{jk} + \tilde{x}_{ik} = 0$. Then in the three transitions, we can not use the directed edges ij, jk, ik. Then the available edges are ji, ki, kj. By drawing a state graph, it is obvious that by the edges ji, ki, kj only, we can not form a path of three transitions. □

Proposition 3.9. *For any $T \geq 4$ of the form $T = 3k+1$ (for $k \geq 2$), the row vector*

$$\mathbf{c} = [2, -1, -1, -1, 2, 2]$$

defines a facet of P^T modulo $\mathfrak{S}_3$.

Proof. Consider the inequality

$$2(x_{12} + x_{31} + x_{32}) \geq x_{13} + x_{21} + x_{23} \tag{3.4}$$

Since $x_{13} + x_{21} + x_{23} = T - 1 - (x_{12} + x_{31} + x_{32}) = 3k - (x_{12} + x_{31} + x_{32})$, (3.4) is equivalent to

$$x_{12} + x_{31} + x_{32} \geq k. \tag{3.5}$$

If we consider paths in triples of transitions, (3.5) follows from Lemma 3.8.

Now we consider the case when the inequality (3.5) becomes an equality. By the induction above, if we divide a path into triples of transitions (edges), then in each triple only one of $\tilde{x}_{12}$, $\tilde{x}_{31}$, $\tilde{x}_{32}$ has to be 1. That is, we proved above that for every three transitions (edges) the left-hand-side of equation (3.5) increases by one. For equality to hold, the LHS can only increase by exactly one. Knowing this, we consider the cases for which three transitions increases the LHS of Equation (3.5) by exactly one. In three transitions, a path can either come back to the same vertex or move to another vertex. In the former case (say $ijki$), the following three loops 1321,3213,2132 increases x_{32} by 1. Another case going from i to j in three transitions are of the form

$$ijij, ijkj, ikij,$$

where i, j, k are different. Then appropriate ones are only the following ones:

$$2121, 1313, 2323.$$

Therefore in three transitions, we go $2 \to 1$, $1 \to 3$ or $2 \to 3$. Among these three, the only possible connection is

$$2 \to 1 \to 3$$

(or $2 \to 3$ alone). Then the loops are inserted at any point. Thus, we can consider the following paths

$$2121321321 \cdots 1321,\ 1313213 \cdots 3213,\ 2323213 \cdots 3213,\ 2132132 \cdots 2132,\ 2121313 \cdots 1313, \tag{3.6}$$

with sufficient statistics given by the vectors

$$[1, k-1, k+1, 0, 0, k-1],\ [0, k+1, k-1, 0, 1, k-1],\ [0, k-1, k-1, 2, 0, k],$$
$$[0, k, k, 0, 0, k],\ [1, 2(k-1), 2, 0, k-1, 0];$$

respectively. For $k \geq 2$ it is easily checked that these vectors are linearly independent and satisfy Equation (3.5) with equality. □

Proposition 3.10. *For any $T \geq 5$ of the form $T = 3k + 2$, $k \geq 1$, the row vector*

$$\mathbf{c} = [2k+1, -k, -k, -k, 2k+1, 2k+1]$$

defines a facet of P^T modulo $\mathfrak{S}_3$.

Proof. Consider

$$(2k+1)(x_{12} + x_{31} + x_{32}) \geq k(x_{13} + x_{21} + x_{23}). \tag{3.7}$$

Since $x_{13} + x_{21} + x_{23} = T - 1 - (x_{12} + x_{31} + x_{32}) = 3k + 1 - (x_{12} + x_{31} + x_{32})$, the inequality (3.7) is equivalent to

$$(2k+1)(x_{12} + x_{31} + x_{32}) \geq k(3k + 1 - (x_{12} + x_{31} + x_{32})),$$

which simplifies to

$$x_{12} + x_{31} + x_{32} \geq k. \tag{3.8}$$

For a path of length $3k + 2$, consider omitting the last transition. Then we have a path of length $3k + 1$. The inequality already holds for this shortened path by Proposition 3.9. Since the last transition only increases the transition counts, the same inequality holds for $3k + 2$.

Now we can find five paths by adding one of the transitions $2 \to 1$, $1 \to 3$, or $2 \to 3$ either at the end or at the beginning of the paths in Proposition 3.6. In this way, we obtain the following paths

$$2121321321 \cdots 1321\underline{3},\ 2321321321 \cdots 1321\underline{3},\ \underline{2}1313213 \cdots 3213,\ \underline{1}3213 \cdots 3213,\ 2132132 \cdots 2132\underline{1},$$

with the following vectors of frequencies:

$$[1, k, k+1, 0, 0, k-1],\ [0, k, k+1, 1, 0, k-1],\ [0, k+1, k, 0, 1, k-1],\ [0, k+1, k, 0, 0, k],$$
$$[0, k, k+1, 0, 0, k];$$

which are easily checked to be linearly independent and satisfy (3.8) with equality.

□

The following lemma will be useful to show the facets of P^T when $T \geq 6$ is even.

Lemma 3.11. *Let $T = 2k$, with $k \geq 1$. Then, the inequality $2x_{12} + x_{13} + x_{32} \geq k - 1$ holds for every path. Moreover, the inequality is strict for every path ending at $s_T = 2$.*

Proof. We prove the lemma by induction on k.

For $k = 1$, this is a path with a single transition; thus, i) and ii) are obvious, because $2x_{12} + x_{13} + x_{32}$ is non-negative and for two paths 12, 32 ending at $s_2 = 2$, we have $2x_{12} + x_{13} + x_{32} = 1$ or 2.

Assume that the proposition holds for k. Now we prove that it holds for $k + 1$. For a path $\mathbf{w}$ let

$$\mathbf{x} = (x_{12}, x_{13}, x_{21}, x_{23}, x_{31}, x_{32}) = (x_{12}(\mathbf{w}), x_{13}(\mathbf{w}), x_{21}(\mathbf{w}), x_{23}(\mathbf{w}), x_{31}(\mathbf{w}), x_{32}(\mathbf{w}))$$

denote the transition counts of $\mathbf{w}$. Consider a path $\mathbf{w}$ of length $T = 2k + 2$

$$\mathbf{w} = s_1 \dots s_{2k} s_{2k+1} s_{2k+2}.$$

Denote $\mathbf{w}^0 = s_1 \dots s_{2k}$ and $\mathbf{w}^1 = s_{2k} s_{2k+1} s_{2k+2}$. Then

$$2x_{12}(\mathbf{w}) + x_{13}(\mathbf{w}) + x_{32}(\mathbf{w}) = (2x_{12}(\mathbf{w}^0) + x_{13}(\mathbf{w}^0) + x_{32}(\mathbf{w}^0)) + (2x_{12}(\mathbf{w}^1) + x_{13}(\mathbf{w}^1) + x_{32}(\mathbf{w}^1)).$$

The inductive assumption is that

$$2x_{12}(\mathbf{w}^0) + x_{13}(\mathbf{w}^0) + x_{32}(\mathbf{w}^0) \geq k - 1$$

and

$$s_{2k} = 2 \;\Rightarrow\; 2x_{12}(\mathbf{w}^0) + x_{13}(\mathbf{w}^0) + x_{32}(\mathbf{w}^0) \geq k.$$

We prove the first statement of the proposition. If

$$2x_{12}(\mathbf{w}^1) + x_{13}(\mathbf{w}^1) + x_{32}(\mathbf{w}^1) \geq 1,$$

then the inequality holds for $\mathbf{w}$. On the other hand it is easily seen that if $2x_{12}(\mathbf{w}^1) + x_{13}(\mathbf{w}^1) + x_{32}(\mathbf{w}^1) = 0$ then the only possible case is $\mathbf{w}^1 = 231$. Then we have $s_{2k} = 2$. Hence by the second part of the inductive assumption we also have the inequality.

We now prove the second statement of the proposition. Let $s_{2k+2} = 2$. Note that s_{2k+1} is either 1 or 3. If $s_{2k+1} = 1$, then

$$2x_{12}(\mathbf{w}^1) + x_{13}(\mathbf{w}^1) + x_{32}(\mathbf{w}^1) = 2$$

and the inequality for $\mathbf{w}$ is strict. On the other hand let $s_{2k+1} = 3$, then there are two cases:

$$\mathbf{w}^1 = 232 \quad \text{or} \quad = 132.$$

In the former case $2x_{12}(\mathbf{w}^1) + x_{13}(\mathbf{w}^1) + x_{32}(\mathbf{w}^1) = 1$, but $s_{2k} = 2$. Hence by the inductive assumption the inequality is strict. In the latter case $2x_{12}(\mathbf{w}^1) + x_{13}(\mathbf{w}^1) + x_{32}(\mathbf{w}^1) = 2$ and the inequality is strict. □

Proposition 3.12. *For any even $T \geq 6$, the row vector*

$$\mathbf{c} = [\frac{3}{2}T - 1, \frac{T}{2}, -\frac{T}{2} + 1, -\frac{T}{2} + 1, -\frac{T}{2} + 1, \frac{T}{2}]$$

defines a facet of P^T modulo $\mathfrak{S}_3$.

Proof. Write $T = 2k$, $k \geq 3$. Consider

$$(3k-1)x_{12} + k(x_{13} + x_{32}) \geq (k-1)(x_{21} + x_{23} + x_{31}).$$

Substituting

$$x_{21} + x_{23} + x_{31} = T - 1 - (x_{12} + x_{13} + x_{32}) = 2k - 1 - (x_{12} + x_{13} + x_{32})$$

into the above and collecting terms, we have

$$(4k-2)x_{12} + (2k-1)(x_{13} + x_{32}) \geq (k-1)(2k-1)$$

or equivalently

$$2x_{12} + x_{13} + x_{32} \geq k - 1, \tag{3.9}$$

which holds by Lemma 3.11

It remains to show that the first inequality for $T \geq 6$ defines a facet. Consider the following 5 paths.

$$31\ldots31,\ 32\ldots3231,\ 2131\ldots31,\ 2313\ldots13,\ 23123131\ldots31$$

The sufficient statistics for these paths are

$$[0,k-1,0,0,k,0],\ [0,0,0,k-1,1,k-1],\ [0,k-1,1,0,k-1,0],\ [0,k-1,0,1,k-1,0],\ [1,k-3,0,2,k-1,0]$$

For $k \geq 3$, linear independence of these vectors can be easily checked. □

We now state some results that will be useful to treat the remaining case when $T = 3l$.

Lemma 3.13. *Suppose that $\mathbf{u}_1, \ldots, \mathbf{u}_k \in \mathbb{R}^m$ are linearly independent vectors such that $[1, 1, \ldots, 1] \cdot \mathbf{a}_j = c > 0$ is a positive constant for $j = 1, \ldots, k$. Then for any non-negative vector $\mathbf{w} \in \mathbb{R}^m$, $\mathbf{u}_1 + \mathbf{w}, \ldots, \mathbf{u}_k + \mathbf{w}$ are linearly independent.*

Proof. For scalars $\alpha_1, \ldots, \alpha_k$ consider

$$0 = (\mathbf{u}_1 + \mathbf{w})\alpha_1 + \cdots + (\mathbf{u}_k + \mathbf{w})\alpha_k = \mathbf{u}_1\alpha_1 + \cdots + \mathbf{u}_k\alpha_k + \mathbf{w}(\alpha_1 + \cdots + \alpha_k) \tag{3.10}$$

Taking the inner product with $[1, 1, \ldots, 1]$ we have

$$0 = (c + [1, 1, \ldots, 1] \cdot \mathbf{w})(\alpha_1 + \cdots + \alpha_k).$$

Here $c + [1, 1, \ldots, 1] \cdot \mathbf{w} > 0$. Hence we have $0 = \alpha_1 + \cdots + \alpha_k$. But then (3.10) reduces to

$$0 = \mathbf{u}_1\alpha_1 + \cdots + \mathbf{u}_k\alpha_k$$

By linear independence of $\mathbf{u}_1, \ldots, \mathbf{u}_k$ we have $\alpha_j = 0$, $j = 1, \ldots, k$. □

Lemma 3.14. *Let $\tilde{x}_{ij}$, $1 \leq i \neq j \leq 3$, be transition counts of three consecutive transitions and let i, j, k be distinct. Then*

$$2\tilde{x}_{ij} + \tilde{x}_{ik} + \tilde{x}_{kj} \geq \tilde{x}_{ji}. \tag{3.11}$$

The equality holds for the following paths: $jiji$, $jkji$, $jiki$. These three paths start from j and ends at i. Furthermore if the difference of both sides is 1 then the possible transitions in three steps are $j \to k$, $k \to i$, and self-loops $i \to i$, $j \to j$, $k \to k$. Finally if the difference of both sides is 2, then the possible transitions in three steps are $k \to j$, and $i \to k$, and self-loops $i \to i$, $j \to j$, $k \to k$.

Proof. If $\tilde{x}_{ji} \le 1$, then the inequality is obvious from Lemma 3.8. If $\tilde{x}_{ji} = 2$, then the only possible path is $jiji$, for which the equality holds in (3.11).

We now consider the values of the difference of both sides. First we determine paths, where the equality holds. $jiji$ is the unique solution for $\tilde{x}_{jk} = 2$. Consider the case $\tilde{x}_{ji} = 1$. Then the equality holds only if $\tilde{x}_{ij} = 0$ and one of $\tilde{x}_{ik}$ and $\tilde{x}_{kj}$ is 1. It is easy to check that the former case corresponds only to $jiki$ and the latter case corresponds only to $jkji$.

We now enumerate the cases that the difference is 1. If $\tilde{x}_{ji} = 0$, then $\tilde{x}_{ij} = 0$ and one of $\tilde{x}_{ik}$ or $\tilde{x}_{kj}$ is zero. This is only possible for the paths $jkik$ or $kjki$. (Recall that $\tilde{x}_{ji} = 2$ leads to $jiji$, for which the difference is zero, as treated above.) Now consider $\tilde{x}_{ji} = 1$. The case $\tilde{x}_{ij} = 1$ corresponds to $jijk$ or $kiji$. The case $\tilde{x}_{ij} = 0$ corresponds the loop $jikj$, $ikji$, or $kjik$. We now see that the transitions in three steps are $j \to k$, $k \to i$, or the self-loops $i \to i$, $j \to j$, $k \to k$.

Finally we enumerate the cases that the difference is 2. It is easy to see that $\tilde{x}_{ji} \le 1$. First suppose $\tilde{x}_{ji} = 0$. If $\tilde{x}_{ij} = 1, \tilde{x}_{ik} = \tilde{x}_{kj} = 0$, then this corresponds to loops (in reverse direction than in the previous case) $ijki$, $jkij$, $kijk$, resulting in self-loops in three steps. If $\tilde{x}_{ij} = 0, \tilde{x}_{ik} = \tilde{x}_{kj} = 1$, then this corresponds to $kikj$ or $ikjk$. If $\tilde{x}_{ij} = 0, \tilde{x}_{ik} = 2, \tilde{x}_{kj} = 0$, then this corresponds to $ikik$. Similarly $\tilde{x}_{ij} = 0, \tilde{x}_{ik} = 0, \tilde{x}_{kj} = 2$ corresponds to $kjkj$. Second suppose $\tilde{x}_{ji} = 1$. LHS has to be 3. Then $\tilde{x}_{ij} = 1$ and one of $\tilde{x}_{ik}$ and $\tilde{x}_{kj}$ is 1. It is easy to see that these correspond to paths $kjij$ and $ijik$. Then we see that in three steps, the possible transitions are $k \to j$, $i \to k$ or the three self-loops. □

Using Lemma 3.14 we now consider 6 consecutive transitions (in two triples). Let $\bar{x}_{ij}$, $1 \le i \ne j \le 3$, denote transition counts of 6 consecutive transitions. Then we have the following lemma.

Lemma 3.15. *Let i, j, k be distinct. Then*

$$2\bar{x}_{ij} + \bar{x}_{ik} + \bar{x}_{kj} \ge \bar{x}_{ji} + 1. \tag{3.12}$$

When the equality holds, then the path has to start from j and end at i in 6 steps. When the difference of both sides is 1, then the possible transitions in 6 steps are $j \to i$, $j \to k$ and $k \to i$.

Proof. From the previous lemma,

$$2\bar{x}_{ij} + \bar{x}_{ik} + \bar{x}_{kj} \ge \bar{x}_{ji}$$

but the equality is impossible, because then the path has to go from j to i in three steps twice. Hence (3.12) holds.

Now consider the case of equality. Then differences of two triples are 0 and 1. By the previous lemma the order of 0 before 1 only corresponds to $j \to i \to i$. The order of 1 before 0 corresponds to $j \to j \to i$. Hence in both cases, the paths have to start from j and end at i in 6 steps.

Now consider the case of difference of 1, i.e.

$$2\bar{x}_{ij} + \bar{x}_{ik} + \bar{x}_{kj} = \bar{x}_{ji} + 2.$$

The two differences of two triples are $(0, 2)$, $(1, 1)$ or $(2, 0)$. In the case of $(0, 2)$, by the previous lemma, the transitions in 6 steps are $j \to i$, $j \to k$. In the case of $(1, 1)$, the transitions in 6 steps are $j \to k$, $k \to i$ or $j \to i$. In the case of $(2, 0)$, the transitions are $k \to i$ or $j \to i$. In summary, the possible transitions in 6 steps are $j \to i$, $j \to k$, $k \to i$.

□

The final lemma is as follows.

Lemma 3.16. *Consider a path of length $T = 6k + 1$, i.e., path with $6k$ steps. Then*

$$2x_{ij} + x_{ik} + x_{kj} \geq x_{ji} + 2k - 1. \tag{3.13}$$

If the equality holds, then the path has to be at i at time T.

Proof. We divide a path into k subpaths of length 6. Suppose that there exists a block for which

$$2\bar{x}_{ij} + \bar{x}_{ik} + \bar{x}_{kj} = \bar{x}_{ji} + 1. \tag{3.14}$$

In this block the path goes from j to i. Before another block of this type, the path has to come back to j. But then there has to be some block of $i \to k$ or $i \to j$. For these blocks

$$2\bar{x}_{ij} + \bar{x}_{ik} + \bar{x}_{kj} \geq \bar{x}_{ji} + 3. \tag{3.15}$$

Therefore the deficit of 1 in (3.14) is compensated by the gain of 1 in (3.15). The lemma follows from this observation. The condition for equality also follows from this observation. □

Now we will show a facet for the cases $T = 6k$ and $T = 6k + 3$, $k = 1, 2, \ldots$.

Proposition 3.17. *For $T = 6k + 3$, the row vector*

$$\mathbf{c} = [5k + 2, 2k + 1, -4k - 1, -k, -k, 2k + 1]$$

defines a facet of P^T modulo $\mathfrak{S}_3$.

Proof. Consider

$$(5k + 2)x_{12} + (2k + 1)x_{13} + (2k + 1)x_{32} \geq (4k + 1)x_{21} + kx_{23} + kx_{31}. \tag{3.16}$$

The RHS is written as

$$(3k + 1)x_{21} + k(x_{21} + x_{23} + x_{31}) = (3k + 1)x_{21} + k(6k + 2 - x_{12} - x_{13} - x_{32}).$$

Hence (3.16) is equivalent to

$$(6k + 2)x_{12} + (3k + 1)x_{13} + (3k + 1)x_{32} \geq (3k + 1)x_{21} + 2k(3k + 1).$$

Dividing by $3k + 1 > 0$, this is equivalent to

$$2x_{12} + x_{13} + x_{32} \geq x_{21} + 2k. \tag{3.17}$$

A path has $T - 1 = 6k + 2$ steps. Consider the first $6k$ steps divided into triples of steps and apply Lemma 3.16 to the LHS. If the inequality in (3.13) is strict, then we only need to check that

$$2x_{12} + x_{13} + x_{32} \geq x_{21}$$

for the remaining two steps. This is obvious, because a transition 21 has to be preceded or followed by some term on the left-hand side.

If equality holds in (3.13), at time $6k+1$, the path is at state 1 and then the penultimate step is either 12 or 13. In either case we easily see that

$$2x_{12} + x_{13} + x_{32} > x_{21}.$$

It remains to show that the first inequality for $T = 6k+3$ defines a facet. Consider the following 5 paths.

$$23132132132132\ldots132132,\ 321321\ldots321,\ 213232321321321\ldots21,$$
$$213213213\ldots213,\ 21321321321\ldots3213212121.$$

The sufficient statistics for these paths are

$$[0, 2k, 2k, 1, 1, 2k],\ [0, 2k, 2k+1, 0, 0, 2k+1],\ [0, 2k-1, 2k, 2, 0, 2k+1]$$
$$[0, 2k+1, 2k+1, 0, 0, 2k],\ [2, 2k-1, 2k+2, 0, 0, 2k-1].$$

This proves the proposition. □

Now we consider $T = 6k$.

Proposition 3.18. *For $T = 6k$, the row vector*

$$\mathbf{c} = [10k-1, 4k, -8k+2, -2k+1, -2k+1, 4k]$$

defines a facet of P^T modulo $\mathfrak{S}_3$.

Proof. Consider

$$(10k-1)x_{12} + 4kx_{13} + 4kx_{32} \geq (8k-2)x_{21} + (2k-1)x_{23} + (2k-1)x_{31}. \qquad (3.18)$$

The RHS is written as

$$(6k-13)x_{21} + (2k-1)(6k-1-x_{12}-x_{13}-x_{32})$$

Hence (3.18) is equivalent to

$$(12k-2)x_{12} + (6k-1)x_{13} + (6k-1)x_{32} \geq (6k-1)x_{21} + (2k-1)(6k-1)$$

or

$$2x_{12} + x_{13} + x_{32} \geq x_{21} + (2k-1).$$

The rest of the proof is similar to that of Proposition 3.17.

It remains to show that the first inequality for $T = 6k$ defines a facet. Consider the following 5 paths.

$$232321321321\ldots321,\ 213231321321\ldots321,\ 321321321321\ldots321,$$
$$213213\ldots213,\ 212321321321\ldots321.$$

The sufficient statistics for these paths are

$$[0, 2k-2, 2k-1, 2, 0, 2k],\ [0, 2k-1, 2k-1, 1, 1, 2k-1],\ [0, 2k-1, 2k, 0, 0, 2k]$$
$$[0, 2k, 2k, 0, 0, 2k-1],\ [1, 2k-2, 2k, 1, 0, 2k-1].$$

□

Here we summarize all the inequalities in their original form and in their inhomogeneous form. Below we only present one of six of the inequalities with the understanding that for each case that any permutation of the labels $\{1, 2, 3\}$ gives another facet. The inhomogeneous form is derived by substituting the equality $n(T-1) = x_{12}+x_{13}+x_{21}+x_{23}+x_{31}+x_{32}$ into the original form. Inhomogeneous form is essential for proving the normality of semigroup associated with the design matrix A^T.

For any $T \geq 5$ homogeneous

$$\mathbf{c} = [1, 0, 0, 0, 0, 0] \cdot \mathbf{x} \geq 0$$

For any $T \geq 5$, homogeneous

$$\mathbf{c} = [T, T, -(T-2), 1, -(T-2), 1)] \cdot \mathbf{x} \geq 0$$

inhomogeneous

$$\mathbf{c} = [1, 1, -1, 0, -1, 0] \cdot \mathbf{x} \geq -n.$$

For any T odd, $T \geq 5$, homogeneous

$$\mathbf{c} = [1, 1, -1, -1, 1, 1] \cdot \mathbf{x} \geq 0.$$

For any $T \geq 4$ of the form $T = 3k+1$, $k \geq 1$, homogeneous

$$\mathbf{c} = [2, -1, -1, -1, 2, 2] \cdot \mathbf{x} \geq 0.$$

For any $T \geq 5$ of the form $T = 3k+2$, $k \geq 1$, homogeneous

$$\mathbf{c} = [2k+1, -k, -k, -k, 2k+1, 2k+1] \cdot \mathbf{x} \geq 0$$

inhomogeneous

$$x_{12} + x_{31} + x_{32} \geq kn = \frac{T-2}{3}n.$$

For any $T \geq 6$, T:even, homogeneous

$$\mathbf{c} = [\frac{3}{2}T - 1, \frac{T}{2}, -\frac{T}{2} + 1, -\frac{T}{2} + 1, -\frac{T}{2} + 1, \frac{T}{2}] \cdot \mathbf{x} \geq 0$$

inhomogeneous

$$3x_{12} + x_{13} - x_{21} - x_{23} - x_{31} + x_{32} \geq -n.$$

For $T = 6k+3$, homogeneous

$$\mathbf{c} = [5k+2, 2k+1, -4k-1, -k, -k, 2k+1] \cdot \mathbf{x} \geq 0$$

inhomogeneous

$$2x_{12} + x_{13} - x_{21} + x_{32} \geq \frac{T-3}{3}n.$$

For $T = 6k$, homogeneous

$$\mathbf{c} = [10k-1, 4k, -8k+2, -2k+1, -2k+1, 4k] \cdot \mathbf{x} \geq 0$$

inhomogeneous

$$2x_{12} + x_{13} - x_{21} + x_{32} \geq \frac{T-3}{3}n.$$

3.3. There are only 24 facets

In the previous section, we give 24 facets of the polytope P^T for every $T \geq 3$, where death of the 24 facets depend on $T \mod 6$. Here, we discuss how these 24 facets are enough to describe the polytope P^T (the convex hull of the columns of A^T), depending on T. Let $C^T := \text{cone}(A^T)$.

Recall that the columns of A^T are on the following hyperplane

$$H_T = \{(x_{12}, \dots, x_{32}) \mid T - 1 = x_{12} + \cdots + x_{32}\}.$$

Then it is clear by Proposition 3.4 that

$$P^T = C^T \cap H_T.$$

Let $\mathcal{F}_T$ denote the set of facets of the pointed cone C^T. Then the facets F of P^T (within H_T) are of the form $F \subseteq H_T, F \subseteq \mathcal{F}_T$.

For every T, let $\tilde{\mathcal{F}}_T$ denote the 24 facets prescribed in the previous section, and let $\mathcal{F}_T$ denote the set of all facets of $P^{3,T}$. Therefore we have a certain subset $\tilde{\mathcal{F}}_T \subset \mathcal{F}_T$ and we need to show that $\tilde{\mathcal{F}}_T = \mathcal{F}_T$. Let $\tilde{\mathcal{C}}_T$ denote the polyhedral cone defined by $\tilde{\mathcal{F}}_T$. It follows that $\tilde{\mathcal{C}}_T \supset \mathcal{C}_T$. Note that $\tilde{\mathcal{F}}_T = \mathcal{F}_T$ if and only if $\tilde{\mathcal{C}}_T = \mathcal{C}_T$. Also let

$$\tilde{P}_T = \tilde{\mathcal{C}}_T \cap H_T.$$

Then $\tilde{P}_T \supset P^T$ and $\tilde{P}_T = P^T$ if and only if $\tilde{\mathcal{C}}_T = \mathcal{C}_T$.

The above argument shows that to prove $\tilde{\mathcal{F}}_T = \mathcal{F}_T$ it suffices to show that

$$\tilde{P}_T \subset P^T. \tag{3.19}$$

Let $\tilde{V}_T$ be the set of vertices of $\tilde{P}_T$. Then in order to show (3.19), it suffices to show that

$$\tilde{V}_T \subset P^T.$$

Hence, if we can obtain explicit expressions of the vertices of $\tilde{V}_T$ and can show that each vertex belongs to P^T, we are done.

In the previous section, we used only the condition $T - 1 = x_{12} + \cdots + x_{32}$ to settle the equivalence between the homogeneous and inhomogeneous inequalities defining the 24 the facets of P^T. Hence the homogeneous and the inhomogeneous inequalities are equivalent on H_T. Therefore, for each $r = 0, \dots, 5$, there exists a polyhedral region defined by 24 fixed affine half-spaces, say Q^r, such that

$$\tilde{\mathcal{P}}_T = Q^r \cap H_T, \quad T = 6k + r, k = 1, 2, \dots$$

Since Q^r is a polyhedral region it can be written as a Minkowski sum of a polytope P^r and a cone C^r:

$$Q^r = P^r + C^r.$$

Please note that r is modulo 6, but T is not. Recall the Minkowski sum of two sets $A, B \subseteq \mathbb{R}^n$ is simply $\{\, a + b \mid a \in A, b \in B \,\}$. The six cones and polytopes defining Q^r for $r = 0, \dots, 5$ are given in the Appendix and were computed using `Polymake` [3]. For each vertex $\mathbf{v}$ of P^r and each extreme ray $\mathbf{e}$ of C^r let $l_{\mathbf{v},\mathbf{e}}$ denote the half-line emanating from $\mathbf{v}$ in the direction $\mathbf{e}$:

$$l_{\mathbf{v},\mathbf{e}} = \{\mathbf{v} + t\mathbf{e} \mid t \geq 0\}$$

Given the explicit expressions of v and e we can solve

$$[1,1,1,1,1,1] \cdot (\mathbf{v} + t\mathbf{e}) = T - 1$$

for t and get

$$t := t(T, \mathbf{v}, \mathbf{e}) = \frac{T - 1 - [1, \dots, 1]\mathbf{v}}{(1, \dots, 1)\mathbf{e}}.$$

Then $v + t(T, \mathbf{v}, \mathbf{e})\mathbf{e} \in H_T$. Note that

$$\tilde{V}_T \subset \{\mathbf{v} + t(T, \mathbf{v}, \mathbf{e})\mathbf{e} \mid \mathbf{v} : \text{vertex of } P^r,\ \mathbf{e} : \text{extreme ray of } C^r\}.$$

Also clearly

$$\{v + t(T, \mathbf{v}, \mathbf{e})\mathbf{e} \mid \mathbf{v} : \text{vertex of } P^r,\ \mathbf{e} : \text{extreme ray of } C^r\} \in \tilde{P}_T = \text{conv}(\tilde{V}_T).$$

The above argument shows that for proving $\tilde{\mathcal{F}}_T = \mathcal{F}_T$ it suffices to show that

$$\{\mathbf{v} + t(T, \mathbf{v}, \mathbf{e})\mathbf{e} \mid \mathbf{v} : \text{vertex of } P^r, \mathbf{e} : \text{extreme ray of } C^r\} \in P^T. \tag{3.20}$$

For proving (3.20) the following lemma is useful.

Lemma 3.19. *Let $\mathbf{v} \in P^r$ and $\mathbf{e} \in C$. If $\mathbf{v} + t(T, \mathbf{v}, \mathbf{e})\mathbf{e} \in P^T \cap \mathbb{Z}^6$ for some T, then $\mathbf{v} + t(T + 6k, \mathbf{v}, \mathbf{e})\mathbf{e} \in P^{T+6k}$ for all $k \geq 0$.*

Proof. If $\mathbf{x} := \mathbf{v} + t(T, \mathbf{v}, \mathbf{e})\mathbf{e} \in P^T \cap \mathbb{Z}^6$ for some T then $\mathbf{x}$ corresponds to a path of length T on three states with no loops (word in $\langle 3 \rangle^T$). Suppose $\mathbf{e}$ is a two-loop (three-loop) e.g. 121 (1231). Then $\mathbf{x} + (3k)\mathbf{e} \in P^{T+6k}$ ($\mathbf{x} + (2k)\mathbf{e} \in P^{T+6k}$). That is, since $\mathbf{x}$ is an integer point (a path) contained in P^T, we can simply add three (or two depending on the loop) copies of the loop $\mathbf{e}$ and we will be guaranteed to have a path of the correct length meaning it will be contained in P^{T+6}. □

By this lemma we need to compute C^r only for some special small T's. We computed all vertices and all rays for the cases $T = 12, 7, 20, 9, 16, 11$. The software to generate the design matrices can be found at `https://github.com/dchaws/GenWordsTrans` and the design matrices and some other material can be found at `http://www.davidhaws.net/THMC.html`. By our computational result and Lemma 3.19 we verified the following proposition.

Proposition 3.20. *The rays of the cones C^r for $r = 0, \dots, 5$ are ((1,0,1,0,0,0), (1,0,0,1,1,0), (0,1,1,0,0,1), (0,1,0,0,1,0), (0,0,0,1,0,1)). In terms of the state graph, the rays correspond to the five loops 121, 131, 232, 1231, and 1321.*

Note that C^r, $r = 0, \dots, 5$ are common and we denote them as C hereafter. Also note that the rays of the cone C^r are very simple. Proposition 3.20 implies the following theorem.

Theorem 3.21. *The 24 facets given in Propositions 3.6, 3.7, 3.9, 3.10, 3.12, 3.17, 3.18 (depending on $T \mod 6$) are all the facets of $P^T = \text{conv}(A^T)$.*

4. Normality of the semigroup

From the definition of normality of a semigroup defined in Section 2.3, the semigroup $\mathbb{N}A^T$ defined by the design matrix is normal if it coincides with the elements in both, the integer lattice $\mathbb{Z}A^T$ and the cone $\mathrm{cone}(A^T)$.

In this section, we provide an inductive prove on the normality of the semigroup $\mathbb{N}A^T$ for arbitrary T. For the first 135 values of T, we computationally verified the normality. These cases served as an inductive base for a proof of normality for a general T.

Lemma 4.1. *The semigroup* $\mathbb{N}A^T$ *is normal for* $1 \leq T \leq 135$.

Proof. The normality of the design matrices A^T for $1 \leq T \leq 135$ was confirmed computationally using the software `Normaliz` [1]. The software to generate the design matrices and the scripts to run the computations are available at `https://github.com/dchaws/GenWordsTrans`. □

Using Lemma 4.1, we have prove normality in the general case in what follows.

Theorem 4.2. *The semigroup* $\mathbb{N}A^T$ *is normal for any* $T \in \mathbb{N}$.

Proof. We need to show that given any transition counts $x_{12}, \ldots, x_{32}$, such that their sum is divisible by $T-1$ and the counts lie in $\mathrm{cone}(A^T)$, there exists a set of paths having these transition counts. Write $\mathbf{x} = [x_{12}, x_{13}, x_{21}, x_{23}, x_{31}, x_{32}]^T$ and $1_6 = [1,1,1,1,1,1]$. Let

$$n = 1_6 \cdot \mathbf{x}/(T-1)$$

denote the number of the paths. Note that n is determined from T and $\mathbf{x}$.

We listed above inhomogeneous forms of inequalities defining facets. In all cases $T = 6k+r$, $r = 0,1,\ldots,5$, the inhomogeneous form of inequalities for n paths can be put in the form

$$c_{12}x_{12} + \cdots + c_{32}x_{32} \geq a_T n, \tag{4.1}$$

where $c_{12}, \ldots, c_{32}, a_T$ do not depend on n.

Let Q^r denote the polyhedral region defined by the inequalities for $T = 6k+r$ and just one path. We have computed the V-representation

$$Q^r = P^r + C,$$

where P^r is a polytope and C is the common cone. The n-th dilation of Q^r is

$$Q^r_n := nQ^r = nP^r + nC = nP^r + C.$$

Then from (4.1) we have

$$\mathrm{cone}(A^T) \cap \{\mathbf{x} \mid 1_6 \cdot \mathbf{x} = n(T-1)\} = Q^r_n \cap \{\mathbf{x} \mid 1_6 \cdot \mathbf{x} = n(T-1)\}.$$

We now look at vertices of P^r from Appendix. The vertex [0,3,4,3,0,7] for Q^2 has the largest L_1-norm, which is 17. Hence the sum of elements of these vertices P^r_n are at most $17n$.

Any non-negative integer vector $\mathbf{x}$ with $1_6 \cdot \mathbf{x} = n(T-1)$, $T = 6k+r$, belonging to $\mathrm{cone}(A^T)$ is written as

$$\mathbf{x} = \mathbf{b} + \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \alpha_3\mathbf{e}_3 + \alpha_4\mathbf{e}_4 + \alpha_5\mathbf{e}_5, \quad \mathbf{b} \in nP^r_1,\ \alpha_i \geq 0,\ i = 1,\ldots,5.$$

Taking the inner product with 1_6 (i.e. the L_1-norm) we have

$$1_6 \cdot \mathbf{b} \le 17n. \tag{4.2}$$

Hence

$$n(T-1) = 1_6 \cdot \mathbf{x} \le 17n + 2(\alpha_1 + \alpha_2 + \alpha_3) + 3(\alpha_4 + \alpha_5).$$

Consider the case that

$$\alpha_1, \alpha_2, \alpha_3 \le 3n, \quad \alpha_4, \alpha_5 \le 2n. \tag{4.3}$$

Then

$$n(T-1) = 1_6 \cdot \mathbf{x} \le 17n + 18n + 12n = 47n$$

or $T \le 48$. Hence if $T > 48$ we have at least one of

$$\alpha_1 > 3n,\ \alpha_2 > 3n,\ \alpha_3 > 3n,\ \ \alpha_4 > 2n,\ \alpha_5 > 2n. \tag{4.4}$$

Now we employ induction on k for $T = 6k + r$. Note that we arbitrarily fix $n \ge 1$ and use induction on k. For $k \le 21$ we have $T = 6k + r \le 126 + r \le 131 < 135$ and the normality holds by the computational results.

Now consider $k > 22$ and let $T = 6k + r$. In this case at least one inequality of (4.4) holds. Let

$$\mathbf{x} \in \text{cone}(A^T) \cap \{x \mid 1_6 \cdot \mathbf{x} = n(T-1)\}$$

First consider $\mathbf{x}$ such that $\alpha_1 > 3n$. (The argument for α_2 and α_3 is the same.) Let

$$\tilde{\mathbf{x}} = \mathbf{x} - 3n\mathbf{e}_1 \in \text{cone}(A^T) \cap \{\mathbf{x} \mid 1_6 \cdot \mathbf{x} = n(T-1-6)\}$$

Our inductive assumption is that there exists a set of paths $\mathbf{w}_1, \dots, \mathbf{w}_n$ of length $T-6$ having $\tilde{x}$ as the transition counts. We now form n partial paths of length 6:

$$n \text{ times } ijijij$$

Note that instead of $ijijij$ we can also use $jijiji$. We now argue that these n partial paths can be appended (at the end or at the beginning) of each path $\mathbf{w}_1, \dots, \mathbf{w}_n$.

Let $\mathbf{w}_1 = s_1 \dots s_{T-6}$. If $s_1 \ne s_{T-6}$, then

$$\{s_1, s_{T-6}\} \cap \{i, j\} \ne \emptyset,$$

since $|S| = 3$. In this case we see that at least one of the following 4 operations is possible

1. put $ijijij$ at the end of $\mathbf{w}_1$
2. put $jijiji$ at the end of $\mathbf{w}_1$
3. put $ijijij$ in front of $\mathbf{w}_1$
4. put $jijiji$ in front of $\mathbf{w}_1$

Hence $\mathbf{w}_1$ can be extended to a path of length T. Now consider the case that $s_1 = s_{T-6}$, i.e., $\mathbf{w}_1$ is a cycle. It may happen that $s_1 \ne i, j$. But a cycle can be rotated, i.e., instead of $\mathbf{w}_1 = s_1 \dots s_{T-6}$ we can take

$$\mathbf{w}_1' = s_2 s_3 \dots s_{T-6} s_1$$

where $s_2 \ne s_1$, hence $s_2 = i$ or j. Then either $ijijij$ or $jijiji$ can be put in front of $\mathbf{w}_1'$ and $\mathbf{w}_1'$ can be extended. Therefore we see that $\mathbf{w}_1$ can be extended in any case. Similarly $\mathbf{w}_2, \dots, \mathbf{w}_n$ can be extended.

The case of $\alpha_4 > 2n$ is trivial. The path $ijkijk$ can be rotated as $jkijki$ or $kijkij$. Therefore one of them can be appended to each of $\mathbf{w}_1, \dots, \mathbf{w}_n$. □

5. Discussion

In this paper, we considered only the situation of the toric homogeneous Markov chain (THMC) model (1.1) for $S = 3$, with the extra assumption of having non-zero transition probabilities only when the transition is between two different states. In this setting, we described the hyperplane representations of the design polytope for any $T \geq 4$, and from this representation we showed that the semigroup generated by the columns of the design matrix A^T is normal.

We recall from Lemma 4.14 in [10], that a given set of integer vectors $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ is a graded set, if there exists $\mathbf{w} \in \mathbb{Q}^{S^2}$ such that $\mathbf{a}_i \cdot \mathbf{w} = 1$. In our setting, the set of columns of the design matrix A^T is a graded set, as each of its columns add up to $T-1$, so we let $\mathbf{w} = (\frac{1}{T-1}, \ldots, \frac{1}{T-1})$.

In his same book, Sturmfels provided a way to bound the generators of the toric ideal associated to an integer matrix A, the precise statement is the following.

Theorem 5.1 (Theorem 13.14 in[10])**.** *Let $A \subset \mathbb{Z}^d$ be a graded set such that the semigroup generated by the elements in A is normal. Then the toric ideal I_A associate with the set A is generated by homogeneous binomials of degree at most d.*

In our setting, Theorem 4.2 demonstrates the normality of the semigroup generated by the columns of the design matrix A^T, so as a consequence we obtain the following theorem:

Theorem 5.2. *For $S = 3$ and for any $T \geq 4$, a Markov basis for the toric ideal I_{A^T} associated to the THMC model (without loops and initial parameters) consists of binomials of degree at most 6.*

The bound provided by Theorem 5.2 seems not to be sharp, in the sense that there exists Markov basis whose elements have degree strictly less than 6. In our computational experiments, we found evidence that more should be true. Our observations hold in a more general setting. For any S and T, let $A^{S,T}$ denote the design matrix for the THMC model in S states.

Conjecture 5.3. *Fix $S \geq 3$; then, for every $T \geq 4$, there is a Markov basis for the toric ideal $I_{A^{S,T}}$ consisting of binomials of degree at most $S-1$, and there is a Gröbner basis with respect to some term ordering consisting of binomials of degree at most S.*

We also noted that for $S \geq 4$, the semigroup generated by the columns of $A^{S,T}$ is not normal. For example, $S = 4$, $T = 8$, $\frac{1}{2}\mathbf{a}^{4,4}_{12121212} + \frac{1}{2}\mathbf{a}^{4,4}_{34343434}$ is an integral solution in the intersection between the cone and the integer lattice. However this does not form a path. Thus it is interesting to investigate that for any $S \geq 4$ and for any $T \geq 5$ what the necessary and sufficient condition is for the semigroup generated by the columns of the design matrix $A^{S,T}$ to be normal.

6. appendix

$C^r = C := \text{cone}\big([1,0,1,0,0,0],[1,0,0,1,1,0],[0,1,1,0,0,1],[0,1,0,0,1,0],[0,0,0,1,0,1]\big)$. for $r = 0, \ldots, 5$.

vert(Q^0) := ([3, 0, 0, 4, 4, 0], [4, 0, 0, 3, 4, 0], [2, 0, 0, 4, 3, 2], [2, 2, 0, 2, 5, 0], [4, 0, 2, 2, 3, 0], [0, 4, 3, 0, 0, 4], [0, 2, 2, 2, 0, 5], [0, 4, 2, 0, 2, 3], [2, 2, 4, 0, 0, 3], [3, 0, 0, 4, 2, 2], [3, 2, 0, 2, 4, 0], [5, 0, 2, 2, 2, 0], [0, 2, 2, 1, 0, 2], [1, 2, 2, 0, 0, 2], [0, 1, 1, 0, 0, 1], [1, 0, 0, 1, 1, 0], [2, 1, 0, 2, 2, 0], [2, 0, 0, 2, 2, 1], [0, 2, 2, 0, 1, 2], [2, 0, 1, 2, 2, 0], [2, 3, 4, 0, 0, 2], [0, 5, 2, 0, 2, 2], [0, 3, 2, 2, 0, 4], [4, 0, 2, 3, 2, 0], [2, 2, 0, 3, 4, 0], [2, 0, 0, 5, 2, 2], [4, 0, 0, 4, 3, 0], [2, 2, 5, 0, 0, 2], [0, 4, 3, 0, 2, 2], [0, 2, 3, 2, 0, 4], [0, 4, 4, 0, 0, 3], [0, 3, 4, 0, 0, 4], [2, 0, 2/3, 2, 7/3, 0], [4/3, 2/3, 2/3, 4/3, 7/3, 2/3], [8/5, 2/5, 4/5, 6/5, 11/5, 0], [2, 0, 0, 2, 7/3, 2/3], [6/5, 2/5, 0, 8/5, 11/5, 4/5], [3/2, 0, 0, 3/2, 2, 0], [2/3, 4/3, 4/3, 2/3, 2/3, 7/3], [0, 8/5, 6/5, 2/5, 4/5, 11/5], [0, 2, 2, 0, 2/3, 7/3], [2/3, 2, 2, 0, 0, 7/3], [4/5, 6/5, 8/5, 2/5, 0, 11/5], [0, 3/2, 3/2, 0, 0, 2], [7/3, 2/3, 2/3, 4/3, 4/3, 2/3], [11/5, 0, 2/5, 8/5, 6/5, 4/5], [7/3, 0, 0, 2, 2, 2/3], [7/3, 2/3, 0, 2, 2, 0], [11/5, 4/5, 2/5, 6/5, 8/5, 0], [2, 0, 0, 3/2, 3/2, 0], [1/2, 1, 2, 1, 0, 3/2], [3/2, 0, 1, 2, 1, 1/2], [1, 1, 3/2, 1/2, 0, 2], [2, 0, 1/2, 3/2, 1, 1], [2, 1, 1/2, 1, 3/2, 0], [1, 2, 3/2, 0, 1/2, 1], [0, 2, 1, 1, 1/2, 3/2], [1, 1, 0, 2, 3/2, 1/2], [0, 3/2, 1, 1/2, 1, 2], [1, 1/2, 0, 3/2, 2, 1], [1/2, 3/2, 2, 0, 1, 1], [3/2, 1/2, 1, 1, 2, 0], [0, 2, 3/2, 0, 0, 3/2], [2/3, 7/3, 2, 0, 0, 2], [4/5, 11/5, 8/5, 0, 2/5, 6/5], [0, 7/3, 2, 2/3, 0, 2], [0, 11/5, 6/5, 4/5, 2/5, 8/5], [2/3, 7/3, 4/3, 2/3, 2/3, 4/3], [3/2, 0, 0, 2, 3/2, 0], [6/5, 4/5, 0, 11/5, 8/5, 2/5], [2, 2/3, 0, 7/3, 2, 0], [2, 0, 2/3, 7/3, 2, 0], [8/5, 0, 4/5, 11/5, 6/5, 2/5], [4/3, 2/3, 2/3, 7/3, 4/3, 2/3], [0, 3/2, 2, 0, 0, 3/2], [2/5, 6/5, 11/5, 4/5, 0, 8/5], [0, 2, 7/3, 2/3, 0, 2], [2/5, 8/5, 11/5, 0, 4/5, 6/5], [0, 2, 7/3, 0, 2/3, 2], [2/3, 4/3, 7/3, 2/3, 2/3, 4/3]).

vert$[Q^1]$:= [([2, 0, 1, 2, 1, 0], [1, 0, 0, 3, 1, 1], [1, 1, 0, 2, 2, 0], [2, 1, 3, 0, 0, 0], [0, 3, 1, 0, 2, 0], [0, 1, 1, 2, 0, 2], [3, 0, 1, 1, 1, 0], [2, 0, 0, 2, 1, 1], [2, 1, 2, 0, 0, 1], [2, 1, 0, 1, 2, 0], [0, 1, 0, 2, 0, 3], [0, 3, 0, 0, 2, 1], [0, 0, 0, 0, 0, 0], [0, 3, 1, 0, 1, 1], [0, 2, 1, 1, 0, 2], [1, 2, 2, 0, 0, 1], [1, 2, 0, 1, 2, 0], [1, 0, 0, 3, 0, 2], [3, 0, 2, 1, 0, 0], [3, 0, 2, 0, 1, 0], [1, 0, 0, 2, 1, 2], [1, 2, 0, 0, 3, 0], [1, 1, 2, 0, 0, 2], [0, 1, 1, 1, 0, 3], [0, 2, 1, 0, 1, 2], [2, 0, 1, 1, 2, 0], [1, 0, 0, 2, 2, 1], [1, 1, 0, 1, 3, 0], [2, 0, 3, 0, 0, 1], [0, 0, 1, 2, 0, 3], [0, 2, 1, 0, 2, 1], [1, 1, 3, 0, 0, 1], [0, 2, 2, 0, 1, 1], [0, 1, 2, 1, 0, 2], [0, 2, 0, 1, 3, 0], [0, 0, 0, 3, 1, 2], [2, 0, 2, 1, 1, 0], [1/2, 1/2, 1/2, 1, 1/2, 0], [0, 1/2, 1/2, 1, 1/2, 1/2], [0, 1, 1/2, 1/2, 1, 0], [1/2, 1/2, 1, 1/2, 1/2, 0], [1, 1/2, 0, 1/2, 1/2, 1/2], [1/2, 1, 0, 0, 1, 1/2], [1/2, 1/2, 0, 1/2, 1/2, 1], [1, 1/2, 1/2, 0, 1/2, 1/2], [1/2, 1/2, 0, 1, 0, 1], [1/2, 1, 0, 1/2, 1/2, 1/2], [1, 1/2, 1/2, 1/2, 0, 1/2], [1/2, 1, 1/2, 1/2, 1/2, 0], [1/2, 1/2, 1/2, 1, 0, 1/2], [1, 1/2, 1, 1/2, 0, 0], [1/2, 1, 1/2, 1/2, 0, 1/2], [1/2, 1/2, 1/2, 0, 1, 1/2], [1/2, 0, 1/2, 1/2, 1/2, 1], [1, 0, 1, 0, 1/2, 1/2], [1/2, 1/2, 1/2, 0, 1/2, 1], [0, 1/2, 1/2, 1/2, 1, 1/2], [0, 0, 1/2, 1, 1/2, 1], [1/2, 0, 1, 1/2, 1/2, 1/2], [1/2, 0, 1/2, 1/2, 1, 1/2], [0, 1/2, 1, 1/2, 1/2, 1/2]).

vert$[Q^2]$:=([2, 2, 0, 3, 4, 0], [4, 0, 2, 3, 2, 0], [2, 0, 0, 5, 2, 2], [0, 2, 2, 3, 0, 4], [0, 7, 4, 0, 3, 3], [0, 4, 2, 1, 2, 2], [3, 4, 7, 0, 0, 3], [2, 2, 4, 1, 0, 2], [3, 0, 0, 4, 2, 2], [5, 0, 2, 2, 2, 0], [3, 2, 0, 2, 4, 0], [3, 2, 4, 0, 0, 2], [0, 7, 3, 0, 3, 4], [1, 4, 2, 0, 2, 2], [0, 4, 3, 3, 0, 7], [1, 2, 2, 2, 0, 4], [0, 2, 1, 0, 1, 1], [0, 1, 1, 1, 0, 2], [1, 1, 0, 1, 2, 0], [1, 1, 2, 0, 0, 1], [2, 0, 1, 1, 1, 0], [1, 0, 0, 2, 1, 1], [2, 1, 0, 4, 2, 2], [4, 1, 2, 2, 2, 0], [7, 0, 3, 4, 3, 0], [4, 0, 0, 7, 3, 3], [0, 5, 2, 0, 2, 2], [0, 3, 2, 2, 0, 4], [2, 3, 0, 2, 4, 0], [2, 3, 4, 0, 0, 2], [2, 2, 0, 2, 4, 1], [4, 3, 0, 3, 7, 0], [4, 0, 2, 2, 2, 1], [7, 0, 3, 3, 4, 0], [0, 4, 2, 0, 2, 3], [0, 2, 2, 2, 0, 5], [2, 0, 0, 4, 2, 3], [2, 2, 4, 0, 0, 3], [2, 2, 4, 0, 1, 2], [0, 3, 4, 3, 0, 7], [0, 2, 2, 2, 1, 4], [3, 3, 7, 0, 0, 4], [3, 3, 0, 4, 7, 0], [3, 0, 0, 7, 4, 3], [2, 0, 1, 4, 2, 2], [2, 2, 1, 2, 4, 0], [2, 2, 0, 2, 5, 0], [4, 0, 2, 2, 3, 0], [2, 0, 0, 4, 3, 2], [0, 4, 2, 0, 3, 2], [0, 4, 3, 0, 2, 2], [0, 2, 3, 2, 0, 4], [4, 0, 3, 2, 2, 0], [2, 2, 5, 0, 0, 2], [4/3, 2/3, 2/3, 7/3, 4/3, 2/3], [2/3, 4/3, 4/3, 5/3, 2/3, 4/3], [7/6, 1/3, 4/3, 5/3, 2/3, 5/6], [5/6, 2/3, 5/3, 4/3, 1/3, 7/6], [1/2, 1, 2, 2, 0, 5/2], [3/2, 1, 3, 1, 0, 3/2], [5/2, 0, 2, 2, 1, 1/2], [3/2, 0, 1, 3, 1, 3/2], [7/3, 2/3, 2/3, 4/3, 4/3, 2/3], [5/3, 4/3, 4/3, 2/3, 2/3, 4/3], [4/3, 2/3, 7/6, 5/6, 1/3, 5/3], [5/3, 1/3, 5/6, 7/6, 2/3, 4/3], [1, 1, 3/2, 3/2, 0, 3], [2, 1, 5/2, 1/2, 0, 2], [3, 0, 3/2, 3/2, 1, 1], [2, 0, 1/2, 5/2, 1, 2], [1/3, 2/3, 2/3, 1/3, 1/3, 2/3], [2/3, 1/3, 1/3, 2/3, 2/3, 1/3], [5/3, 4/3, 5/6, 2/3, 7/6, 1/3], [4/3, 5/3, 7/6, 1/3, 5/6, 2/3], [3, 1, 3/2, 1, 3/2, 0], [2, 2, 1/2, 1, 5/2, 0], [1, 3, 3/2, 0, 3/2, 1], [2, 2, 5/2, 0, 1/2, 1], [0, 3, 1, 1, 3/2, 3/2], [0, 2, 1, 2, 1/2, 5/2], [1, 2, 0, 2, 5/2, 1/2], [1, 1, 0, 3, 3/2, 3/2], [1/3, 5/3, 2/3, 4/3, 5/6, 7/6], [2/3, 4/3, 1/3, 5/3, 7/6, 5/6], [4/3, 5/3, 2/3, 4/3, 4/3, 2/3], [2/3, 7/3, 4/3, 2/3, 2/3, 4/3], [0, 5/2, 1, 1/2, 2, 2], [0, 3/2, 1, 3/2, 1, 3], [1, 3/2, 0, 3/2, 3, 1], [1, 1/2, 0, 5/2, 2, 2], [1/3, 7/6, 2/3, 5/6, 4/3, 5/3], [2/3, 5/6, 1/3, 7/6, 5/3, 4/3], [4/3, 2/3, 2/3, 4/3, 4/3, 5/3], [2/3, 4/3, 4/3, 2/3, 2/3, 7/3], [3/2, 3/2, 3, 0, 1, 1], [1/2, 5/2, 2, 0, 2, 1], [3/2, 3/2, 1, 1, 3, 0], [5/2, 1/2, 2, 1, 2, 0], [5/6, 7/6, 5/3, 1/3, 4/3, 2/3], [7/6, 5/6, 4/3, 2/3, 5/3, 1/3], [2/3, 4/3, 4/3, 2/3, 5/3, 4/3], [4/3, 2/3, 5/3, 4/3, 4/3, 2/3], [4/3, 2/3, 2/3, 4/3, 7/3, 2/3], [2/3, 4/3, 7/3, 2/3, 2/3, 4/3]).

vert$[Q^3]$:= ([1, 0, 0, 0, 1, 0], [1, 2, 0, 1, 4, 0], [1, 0, 0, 3, 2, 2], [3, 0, 2, 1, 2, 0], [0, 1, 0, 0, 0, 1], [0, 3, 1, 0, 2, 2], [0, 1, 1, 2, 0, 4], [2, 1, 3, 0, 0, 2], [2, 2, 0, 1, 3, 0], [2, 0, 0, 3, 1, 2], [4, 0, 2, 1, 1, 0], [0, 0, 0, 0, 0, 0], [2, 2, 3, 0, 0, 1], [0, 2, 1, 2, 0, 3], [0, 4, 1, 0, 2, 1], [3, 0, 2, 2, 1, 0], [1, 0, 0, 4, 1, 2], [1, 2, 0, 2, 3, 0], [1, 0, 0, 1, 0, 0], [2, 1, 4, 0, 0, 1], [0, 1, 2, 2, 0, 3], [0, 3, 2, 0, 2, 1], [0, 1, 1, 0, 0, 0], [0, 0, 1, 0, 0, 1], [0, 0, 0, 1, 1, 0]).

vert$[Q^4]$:= ([0, 2, 2, 2, 0, 3], [3, 0, 1, 3, 2, 0], [2, 1, 0, 3, 3, 0], [2, 0, 0, 4, 2, 1], [0, 6, 4, 0, 2, 3], [0, 3, 2, 1, 1, 2], [2, 4, 6, 0, 0, 3], [1, 2, 3, 1, 0, 2], [3, 0, 0, 3, 2, 1], [4, 0, 1, 2, 2, 0], [3, 1, 0, 2, 3, 0], [2, 2, 3, 0, 0, 2], [0, 4, 3, 2, 0, 6], [0, 6, 3, 0, 2, 4], [1, 3, 2, 0, 1, 2], [1, 2, 2, 1, 0, 3], [0, 1, 1, 0, 0, 1], [1, 0, 0, 1, 1, 0], [2, 1, 0, 3, 2, 1], [3, 1, 1, 2, 2, 0], [6, 0, 2, 4, 3, 0], [4, 0, 0, 6, 3, 2], [0, 4, 2, 0, 1, 2], [0, 3, 2, 1, 0, 3], [1, 3, 3, 0, 0, 2], [2, 2, 0, 2, 3, 0], [2, 1, 0, 2, 3, 1], [4, 2, 0, 3, 6, 0], [3, 0, 1, 2, 2, 1], [6, 0, 2, 3, 4, 0], [0, 2, 2, 1, 0, 4], [0, 3, 2, 0, 1, 3], [1, 2, 3, 0, 0, 3], [2, 0, 0, 3, 2, 2], [1, 2, 3, 0, 1, 2], [0, 3, 4, 2, 0, 6], [0, 2, 2, 1, 1, 3], [2, 3, 6, 0, 0, 4], [3, 2, 0, 4, 6, 0], [3, 0, 0, 6, 4, 2], [2, 0, 1, 3, 2, 1], [2, 1, 1, 2, 3, 0], [2, 1, 0, 2, 4, 0], [2, 0, 0, 3, 3, 1], [3, 0, 1, 2, 3, 0], [0, 3, 2, 0, 2, 2], [0, 2, 3, 1, 0, 3], [0, 3, 3, 0, 1, 2], [1, 2, 4, 0, 0, 2], [3, 0, 2, 2, 2, 0], [5/3, 1/3, 1/3, 8/3, 5/3, 1/3], [1/3, 5/3, 5/3, 4/3, 1/3, 5/3], [1/2, 1, 2, 1, 0, 3/2], [3/2, 0, 1, 2, 1, 1/2], [8/3, 1/3, 1/3, 5/3, 5/3, 1/3], [4/3, 5/3, 5/3, 1/3, 1/3, 5/3], [1, 1, 3/2, 1/2, 0, 2], [2, 0, 1/2, 3/2, 1, 1], [2, 1, 1/2, 1, 3/2, 0], [1, 2, 3/2, 0, 1/2, 1], [0, 2, 1, 1, 1/2, 3/2], [1, 1, 0, 2, 3/2, 1/2], [5/3, 4/3, 1/3, 5/3, 5/3, 1/3], [1/3, 8/3, 5/3, 1/3, 1/3, 5/3], [0, 3/2, 1, 1/2, 1, 2], [1, 1/2, 0, 3/2, 2, 1], [5/3, 1/3, 1/3, 5/3, 5/3, 4/3], [1/3, 5/3, 5/3, 1/3, 1/3, 8/3], [1/2, 3/2, 2, 0, 1, 1], [3/2, 1/2, 1, 1, 2, 0], [1/3, 5/3, 5/3, 1/3, 4/3, 5/3], [5/3, 1/3, 4/3, 5/3, 5/3, 1/3], [5/3, 1/3, 1/3, 5/3, 8/3, 1/3], [1/3, 5/3, 8/3, 1/3, 1/3, 5/3]).

vert$[Q^5]$:= ([1, 2, 0, 2, 3, 0], [3, 0, 2, 2, 1, 0], [1, 0, 0, 4, 1, 2], [0, 1, 1, 3, 0, 3], [0, 4, 1, 0, 3, 0], [3, 1, 4, 0, 0, 0], [0, 4, 0, 0, 3, 1], [0, 1, 0, 3, 0, 4], [4, 0, 2, 1, 1, 0], [2, 0, 0, 3, 1, 2], [3, 1, 3, 0, 0, 1], [2, 2, 0, 1, 3, 0], [0, 1, 0, 0, 1, 0], [0, 0, 0, 1, 0, 1], [1, 0, 1, 0, 0, 0], [1, 0, 0, 4, 0, 3], [4, 0, 3, 1, 0, 0], [0, 2, 1, 2, 0, 3], [0, 4, 1, 0, 2, 1], [1, 3, 0, 1, 3, 0], [2, 2, 3, 0, 0, 1], [1, 3, 0, 0, 4, 0], [1, 0, 0, 3, 1, 3], [0, 3, 1, 0, 2, 2], [0, 1, 1, 2, 0, 4], [4, 0, 3, 0, 1, 0], [2, 1, 3, 0, 0, 2], [3, 0, 2, 1, 2, 0], [1, 0, 0, 3, 2, 2], [1, 2, 0, 1, 4, 0], [0, 3, 1, 0, 3, 1], [0, 0, 1, 3, 0, 4], [3, 0, 4, 0, 0, 1], [3, 0, 3, 1, 1, 0], [0, 0, 0, 4, 1, 3], [0, 3, 0, 1, 4, 0], [2, 1, 4, 0, 0, 1], [0, 3, 2, 0, 2, 1], [0, 1, 2, 2, 0, 3], [1/2, 1/2, 1/2, 2, 1/2, 1], [3/2, 1/2, 3/2, 1, 1/2, 0], [1/2, 3/2, 1/2, 1, 3/2, 0], [0, 1/2, 1/2, 2, 1/2, 3/2], [0, 2, 1/2, 1/2, 2, 0], [3/2, 1/2, 2, 1/2, 1/2, 0], [1, 1/2, 0, 3/2, 1/2, 3/2], [1, 3/2, 0, 1/2, 3/2, 1/2], [2, 1/2, 1, 1/2, 1/2, 1/2], [2, 1/2, 3/2, 0, 1/2, 1/2], [1/2, 1/2, 0, 3/2, 1/2, 2], [1/2, 2, 0, 0, 2, 1/2], [2, 1/2, 3/2, 1/2, 0, 1/2], [1/2, 2, 0, 1/2, 3/2, 1/2], [1/2, 1/2, 0, 2, 0, 2], [1/2, 1/2, 1/2, 2, 0, 3/2], [2, 1/2, 2, 1/2, 0, 0], [1/2, 2, 1/2, 1/2, 3/2, 0], [1/2, 1, 1/2, 3/2, 0, 3/2], [1/2, 2, 1/2, 1/2, 1, 1/2], [3/2, 1, 3/2, 1/2, 0, 1/2], [1/2, 3/2, 1/2, 0, 2, 1/2], [1/2, 0, 1/2, 3/2, 1/2, 2], [2, 0, 2, 0, 1/2, 1/2], [3/2, 1/2, 3/2, 0, 1/2, 1], [1/2, 1/2, 1/2, 1, 1/2, 2], [1/2, 3/2, 1/2, 0, 3/2, 1], [3/2, 0, 2, 1/2, 1/2, 1/2], [0, 0, 1/2, 2, 1/2, 2], [0, 3/2, 1/2, 1/2, 2, 1/2], [1/2, 1, 1/2, 1/2, 2, 1/2], [1/2, 0, 1/2, 3/2, 1, 3/2], [3/2, 0, 3/2, 1/2, 1, 1/2], [0, 1/2, 1, 3/2, 1/2, 3/2], [0, 3/2, 1, 1/2, 3/2, 1/2], [1, 1/2, 2, 1/2, 1/2, 1/2]).

[1] Bruns, W., Ichim, B., Söger, C., 2011. `Normaliz`, a tool for computations in affine monoids, vector configurations, lattice polytopes, and rational cones.

[2] Diaconis, P., Sturmfels, B., 1998. Algebraic algorithms for sampling from conditional distributions. The Annals of Statistics 26 (1), 363–397.

[3] Gawrilow, E., Joswig, M., 2000. polymake: a framework for analyzing convex polytopes. In: Kalai, G., Ziegler, G. M. (Eds.), Polytopes — Combinatorics and Computation. Birkhäuser, pp. 43–74.

[4] Hara, H., Takemura, A., 2011. A markov basis for two-state toric homogeneous markov chain model without initial paramaters. Journal of Japan Statistical Society 41, 33–49.

[5] Haws, D., Martin del Campo, A., Yoshida, R., 2011. Degree bounds for a minimal markov basis for the three-state toric homogeneous markov chain model. Proceedings of the Second CREST–SBM International Conference, "Harmony of Grobner Bases and the Modern Industrial Society".

[6] Miller, E., Sturmfels, B., 2005. Combinatorial commutative algebra. Graduate texts in mathematics. Springer.
URL `http://books.google.com/books?id=CqEHpxbKgv8C`

[7] Pachter, L., Sturmfels, B., 2005. Algebraic Statistics for Computational Biology. Cambridge University Press, Cambridge, UK.

[8] Pardoux, E., 2009. Markov Processes and Applications: Algorithms, Networks, Genome and Finance. Wiley Series in Probability and Statistics Series. Wiley, John & Sons, Incorporated.

[9] Schrijver, A., 1986. Theory of Linear and Integer Programming. John Wiley & Sons, Inc., New York, NY, USA.

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

[11] Takemura, A., Hara, H., 2012. Markov chain monte carlo test of toric homogeneous markov chains. Statistical Methodology. doi:10.1016/j.stamet.2011.10.004. 9, 392–406.

[12] Yap, C.-K., 2000. Fundamental problems of algorithmic algebra. Oxford University Press. URL `http://books.google.com/books?id=YzjvQ-TSQ4gC`