MATHEMATICAL ENGINEERING TECHNICAL REPORTS

A Markov Basis for Two-State Toric Homogeneous Markov Chain Model without Initial Parameters

Hisayuki HARA and Akimichi TAKEMURA

METR 2010–13 May 2010

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

A Markov basis for two-state toric homogeneous Markov chain model without initial parameters

Hisayuki Hara[1] and Akimichi Takemura[2,3]

[1]Department of Technology Management for Innovation, University of Tokyo
[2]Graduate School of Information Science and Technology, University of Tokyo
[3]JST, CREST

Abstract. We derive a Markov basis consisting of moves of degree at most three for two-state toric homogeneous Markov chain model of arbitrary length without parameters for initial states. Our basis consists of moves of degree three and degree one, which alter the initial frequencies, in addition to moves of degree two and degree one for toric homogeneous Markov chain model with parameters for initial states.

Keywords: algebraic statistics, Gröbner basis, indispensable move, toric ideal

1 Introduction

Consider a Markov chain X_t, $t = 1, \dots, T\,(\geq 3)$, over finite state space $\mathcal{S}$. Let $\omega = (s_1, \dots, s_T) \in \mathcal{S}^T$ denote a path of a Markov chain. In this paper we discuss Markov bases of toric ideals arising from the following statistical model

$$p(\omega) = c\beta_{s_1 s_2} \dots \beta_{s_{T-1} s_T}, \tag{1}$$

where c is a normalizing constant. In [3], the model (1) is called a toric Markov chain model. In [6] we considered a model with additional parameters γ_s, $s \in \mathcal{S}$, for the initial states:

$$p(\omega) = c\gamma_{s_1}\beta_{s_1 s_2} \dots \beta_{s_{T-1} s_T} \tag{2}$$

and derived a Markov basis of toric ideals arising from (2) for the case of $\mathcal{S} = \{1, 2\}$ (arbitrary T) and the case of $T = 3$ (arbitrary $\mathcal{S}$). In [6] we called the model (2) a toric homogeneous Markov chain(THMC) model. The model (1) corresponds to THMC model with $\gamma_1 = \cdots = \gamma_{|\mathcal{S}|}$. For distinguishing two models, we call (1) a toric homogeneous Markov chain model without initial parameters.

In the present paper, we generalize the result in [6] to the model (1) and derive a Markov basis for the case $\mathcal{S} = \{1, 2\}$ and arbitrary T.

From a statistical viewpoint, Markov bases are used to test goodness-of-fit of the model based on the exact distribution of a test statistic. A data set of paths is summarized in an $|\mathcal{S}|^T$ contingency table. The set of contingency tables sharing sufficient statistic is called a fiber. A Markov basis is defined as a set of moves connecting every fiber. In this paper we consider the problem in the framework of contingency table analysis and derive a Markov basis as a set of moves which guarantees the connectivity of every fiber.

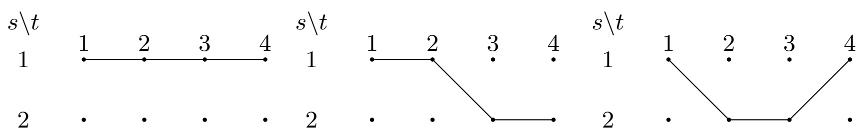
The organization of the paper is as follows. In the rest of this section we introduce some notations and terminologies and give some preliminary results. In Section 2, we state the main theorem and give a Markov basis for the model (1) with $\mathcal{S} = \{1, 2\}$. We give a proof of the theorem in Section 3. In Section 4 we end the paper with some concluding remarks.

1.1 Notation and terminology

Let $\omega = (s_1, \ldots, s_T)$ be a path. For notational simplicity we sometimes write $\omega = (s_1 \ldots s_T)$ or $\omega = s_1 \ldots s_T$. If $s_1 = \cdots = s_T = i$, we call ω a flat path at state i. If for some t $(1 \le t < T)$, $s_1 = \cdots = s_t = i \neq j = s_{t+1} = \cdots = s_T$, then we call ω a single-step path from i to j at time t. We say that a path ω starts at i and ends at j if $s_1 = i, s_T = j$. The set of such paths is denoted by $W_{i*\cdots*j}$. We call ω a non-flat cycle, if it starts from i, visits $j \neq i$ on the way and ends at i. We denote the set of non-flat cycles starting from i by

$$W_{i*j*i} = \{\omega \mid s_1 = i, s_t = j, s_T = i, \text{ for some } 1 < t < T\}. \tag{3}$$

In Figure 1, we depict examples of a flat path, a single-step path and a non-flat cycle.



(i) a flat path (ii) a single-step path (iii) a non-flat cycle

Fig. 1. A flat path, a single-step path and a non-flat cycle for $T = 4$

A data set of n paths is summarized in an $|\mathcal{S}|^T$ contingency table $\boldsymbol{x} = \{x(\omega), \omega \in \mathcal{S}^T\}$ of total frequency n, where $x(\omega)$ denotes the frequency of the path ω. Let $x_{ij}^t = \sum_{\omega: s_t = i, s_{t+1} = j} x(\omega)$ denote the number of transitions from $s_t = i$ to $s_{t+1} = j$ in $\boldsymbol{x}$ and let $x_i^t = \sum_{\omega: s_t = i} x(\omega)$ denote the frequency of the state $s_t = i$ in $\boldsymbol{x}$. In particular x_i^1 is the frequency of the initial state $s_1 = i$. Let

$$x_{ij}^+ = \sum_{t=1}^{T-1} x_{ij}^t$$

denote the total number of transitions from i to j in $\boldsymbol{x}$.

The sufficient statistic for the model (1) is given by the frequencies of transitions

$$\boldsymbol{b} = \{x_{ij}^{+}, i, j \in \mathcal{S}\}. \tag{4}$$

The sufficient statistic for the model (2) is given by the union of (4) and the set of initial frequencies,

$$\boldsymbol{b}' = \{x_i^1, i \in \mathcal{S}\} \ \cup \ \{x_{ij}^{+}, i, j \in \mathcal{S}\} \supset \boldsymbol{b}.$$

For $\mathcal{S} = \{1, 2\}$ we write the elements of $\boldsymbol{b}$ in (4) as $b_{11}, b_{12}, b_{21}, b_{22}$.

If we order paths appropriately and write $\boldsymbol{x}$ as a column vector according to the order, $\boldsymbol{b}$ in (4) is written in a matrix form

$$\boldsymbol{b} = A\boldsymbol{x},$$

where A is an $|\mathcal{S}|^2 \times |\mathcal{S}|^T$ matrix consisting of non-negative integers with the rows indexed by $|\mathcal{S}|^2$ transitions and the columns indexed by $|\mathcal{S}|^T$ paths. The $((i, j), \omega)$ element A is the number of occurrences of the transition from i to j in the path ω.

A is the configuration defining toric ideal I_A arising from the model (1). A toric ideal I_A is the kernel of the homomorphism of polynomial rings $\psi : k[\{p(\omega), \omega \in \mathcal{S}^T\}] \to k[\{\beta_{ij}, i, j \in \mathcal{S}\}]$ defined by

$$\psi : p(\omega) \mapsto \beta_{s_1 s_2} \ldots \beta_{s_{T-1} s_T},$$

where $\{p(\omega), \omega \in \mathcal{S}^T\}$ is regarded as a set of indeterminates. Algebraically a Markov basis is defined as a set of generators of I_A ([1]).

The set of all contingency tables sharing $\boldsymbol{b}$ is called a *fiber* and denoted by $\mathcal{F}_{\boldsymbol{b}} = \{\boldsymbol{x} \in \mathbb{Z}_{\geq 0}^{|\mathcal{S}|^T} \mid A\boldsymbol{x} = \boldsymbol{b}\}$, where $\mathbb{Z}_{\geq 0} = \{0, 1, \ldots\}$. A move $\boldsymbol{z}$ for the model (1) is an integer array satisfying $A\boldsymbol{z} = 0$. In contingency table analysis a Markov basis is defined as a finite set of moves $\mathcal{Z}$ satisfying that for all $\boldsymbol{b}$ and all pairs $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathcal{F}_{\boldsymbol{b}}$ there exists a sequence $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_K \in \mathcal{Z}$ such that

$$\boldsymbol{y} = \boldsymbol{x} + \sum_{k=1}^{K} \boldsymbol{z}_k, \quad \boldsymbol{x} + \sum_{k=1}^{l} \boldsymbol{z}_k \geq 0, \ l = 1, \ldots, K. \tag{5}$$

In this article we provide a Markov basis as a set of moves satisfying (5) for the model (1) with $\mathcal{S} = \{1, 2\}$.

A move $\boldsymbol{z}$ is expressed by a difference of two contingency tables $\boldsymbol{x}$ and $\boldsymbol{y}$ in the same fiber:

$$\boldsymbol{z} = \boldsymbol{x} - \boldsymbol{y}, \qquad z(\omega) = x(\omega) - y(\omega), \ \omega \in \mathcal{S}^T.$$

We write $z_{ij}^t = x_{ij}^t - y_{ij}^t$.

1.2 Some classes of moves

In this section we introduce some classes of moves for the model (2) discussed in [6]. The moves introduced in this section are also moves for the model (1).

Denote $\boldsymbol{s}_{t:t'} = s_t \dots s_{t'}$ and $\boldsymbol{s}'_{t:t'} = s'_t \dots s'_{t'}$ for $t \le t'$. If $t > t'$, let $\boldsymbol{s}_{t:t'}$ be an empty sequence. Let $\bar{\omega} = (s_1, \dots, s_T)$ be a path satisfying $s_{t_0} = s_{t_1} = s_{t_2} = i$ for $1 \le t_0 < t_1 < t_2 \le T$ and $s_t = j \neq i$ for some $t_0 < t < t_2$. Then consider the following swapping

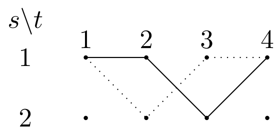
$$\bar{\omega} = (\boldsymbol{s}_{1:t_0-1}, \boldsymbol{s}_{t_0:t_2}, \boldsymbol{s}_{t_2+1:T}) \quad \leftrightarrow \quad \bar{\omega}' := (\boldsymbol{s}_{1:t_0-1}, \boldsymbol{s}_{t_1:t_2-1}, \boldsymbol{s}_{t_0:t_1}, \boldsymbol{s}_{t_2+1:T}).$$

Then an integer array $\boldsymbol{z} = \{z(w), w \in \mathcal{S}^T\}$

$$z(\omega) = \begin{cases} \;\;1 & \text{if } \omega = \bar{\omega} \\ -1 & \text{if } \omega = \bar{\omega}' \\ \;\;0 & \text{otherwise} \end{cases} \tag{6}$$

forms a move for (2). We call this move a degree one move.

We depict a degree one move with $T = 4$, $t_0 = 1$, $t_1 = 2$ and $t_2 = 4$ as



, (7)

where a solid line from (i, t) to $(j, t+1)$ represents $z^t_{ij} = 1$ and a dotted line from (i, t) to $(j, t+1)$ represents $z^t_{ij} = -1$. We call a graph like (7) a move graph. A node of a move graph is a pair (i, t) of state i and time t and an edge from (i, t) to $(j, t+1)$ represents the value of z^t_{ij}. If $|z^t_{ij}| = 0$, there is no corresponding edge in the graph. If $|z^t_{ij}| \ge 2$, we write the value of $|z^t_{ij}|$ beside the edge.

We say that two paths $\omega = (s_1, \dots, s_T)$, $\omega' = (s'_1, \dots, s'_T)$ meet (or cross) at the node (i, t) if $i = s_t = s'_t$. If ω and ω' cross at the node (i, t), consider the swapping of these two paths like

$$\begin{aligned} \{\bar{\omega}, \bar{\omega}'\} &= \{ \; (\boldsymbol{s}_{1:t-1}, i, \boldsymbol{s}_{t+1:T}), \; (\boldsymbol{s}', i, \boldsymbol{s}'_{t+1:T}) \; \} \\ &\leftrightarrow \{ \; (\boldsymbol{s}_{1:t-1}, i, \boldsymbol{s}'_{t+1:T}), \; (\boldsymbol{s}', i, \boldsymbol{s}_{t+1:T}) \; \} =: \{\tilde{\omega}, \tilde{\omega}'\}. \end{aligned}$$

Then the integer array $\boldsymbol{z}$

$$z(\omega) = \begin{cases} \;\;1 & \text{if } \omega = \bar{\omega} \text{ or } \bar{\omega}' \\ -1 & \text{if } \omega = \tilde{\omega} \text{ or } \tilde{\omega}' \\ \;\;0 & \text{otherwise} \end{cases} \tag{8}$$

forms a move for (2). We call this move a crossing path swapping. As shown in [6], a crossing path swapping is expressed by the difference of two tables with the same edge-sign pattern and hence the move graph for a crossing path swapping has no edge. As shown in [6], if $\boldsymbol{x}$ and $\boldsymbol{y}$ are in the same fiber such that the move graph of $\boldsymbol{z}$ has no edge, $\boldsymbol{x}$ and $\boldsymbol{y}$ are connected by crossing path swappings.

Suppose that $t_0 \neq t_1$. Choose (not necessarily distinct) four paths $\omega_1, \omega_2, \omega_3, \omega_4$ as

$$\begin{aligned} \omega_1 &= (\boldsymbol{s}_{1,1:t_0-1}, 1, 1, \boldsymbol{s}_{1,t_0+2:T}),\ \omega_2 = (\boldsymbol{s}_{2,1:t_0-1}, 2, 2, \boldsymbol{s}_{2,t_0+2:T}), \\ \omega_3 &= (\boldsymbol{s}_{3,1:t_1-1}, 1, 2, \boldsymbol{s}_{3,t_1+2:T}),\ \omega_4 = (\boldsymbol{s}_{4,1:t_1-1}, 2, 1, \boldsymbol{s}_{4,t_1+2:T}), \end{aligned}$$

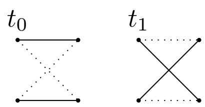
where $\boldsymbol{s}_{k,t:t'} = s_{k,t}, \ldots, s_{k,t'}$. Then we consider the swapping $\{\omega_1, \omega_2, \omega_3, \omega_4\} \leftrightarrow \{\tilde{\omega}_1, \tilde{\omega}_2, \tilde{\omega}_3, \tilde{\omega}_4\}$, where

$$\begin{aligned} \tilde{\omega}_1 &= (\boldsymbol{s}_{1,1:t_0-1}, 1, 2, \boldsymbol{s}_{2,t_0+2:T}),\ \tilde{\omega}_2 = (\boldsymbol{s}_{2,1:t_0-1}, 2, 1, \boldsymbol{s}_{1,t_0+2:T}), \\ \tilde{\omega}_3 &= (\boldsymbol{s}_{3,1:t_1-1}, 1, 1, \boldsymbol{s}_{4,t_1+2:T}),\ \tilde{\omega}_4 = (\boldsymbol{s}_{4,1:t_1-1}, 2, 2, \boldsymbol{s}_{3,t_1+2:T}). \end{aligned}$$

Then the integer array $\boldsymbol{z}$

$$z(\omega) = \begin{cases} \ \ 1 \text{ if } \omega = \omega_1, \ldots, \omega_4, \\ -1 \text{ if } \omega = \tilde{\omega}_1, \ldots, \tilde{\omega}_4, \\ \ \ 0 \text{ otherwise.} \end{cases}$$

is a move for (2) and is called 2 by 2 swap. The corresponding move graph is depicted as



(9)

For $\mathcal{S} = \{1, 2\}$ it can be shown that we can always choose either $\omega_1 = \omega_3$, $\omega_2 = \omega_4$ or $\omega_1 = \omega_4$, $\omega_2 = \omega_3$. Therefore for $\mathcal{S} = \{1, 2\}$ a 2 by 2 swap always corresponds to a degree two move $\boldsymbol{z}$.

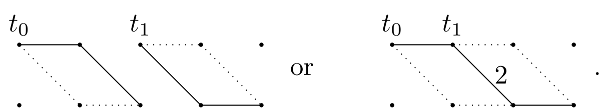
Next we consider the following swapping for $T \geq 4$:

$$\begin{aligned} (\omega_1, \omega_2) &:= \left\{ (\boldsymbol{s}_{t_0-1}, 1, 1, 2, \boldsymbol{s}_{t_0+3:T}), (\boldsymbol{s}'_{1:t_1-1}, 1, 2, 2, \boldsymbol{s}'_{t_1+3:T}) \right\} \\ &\leftrightarrow (\tilde{\omega}_1, \tilde{\omega}_2) := \left\{ (\boldsymbol{s}_{1:t_0-1}, 1, 2, 2, \boldsymbol{s}_{t_0+3:T}), (\boldsymbol{s}'_{1:t_1-1}, 1, 1, 2, \boldsymbol{s}'_{t_1+3:T}) \right\}. \end{aligned}$$

Then the integer array $\boldsymbol{z}$

$$z(\omega) = \begin{cases} \ \ 1 & \text{if } \omega = \omega_1, \omega_2, \\ -1 & \text{if } \omega = \tilde{\omega}_1, \tilde{\omega}_2, \\ \ \ 0 & \text{otherwise.} \end{cases} \tag{10}$$

also forms a move for (2), which can be depicted as



2 Main result

In [6] we provided a Markov basis for the model (2) as follows.

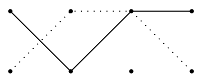
Proposition 1. *A Markov basis for $\mathcal{S} = \{1, 2\}$, $T \geq 4$, for (2) consists of (i) degree one moves (6), (ii) crossing path swappings, (iii) 2 by 2 swaps (9), (iv) moves in (10). For $T = 3$ a Markov basis consists of the first three types of moves.*

Since 2 by 2 swaps for $\mathcal{S} = \{1, 2\}$ corresponds to degree two moves, (2) has a Markov basis consisting of moves of degree at most two.

We call degree one moves in (6) *type I degree one moves*. We note that the moves introduced in the previous section are moves for (2) and hence do not alter the initial frequencies. For connecting fibers for the model (1) we need moves altering initial frequencies. First we present a degree one move altering the initial frequency. Consider a non-flat cycle $\omega \in W_{i*j*i}$ in (3), vising j at time t. Consider the following degree one move:

$$W_{i*j*i} \ni \omega = (s_1, \ldots, s_T) \ \leftrightarrow \ \omega' = (s_t, \ldots, s_{T-1}, s_1, \ldots, s_t) \in W_{j*i*j}. \tag{11}$$

ω and ω' has the same number of transitions, but the initial states are different. We call this move a *type II degree one move*. An example of a type II degree one move for $T = 4$ is depicted as

 . (12)

Next we consider a degree three move. Let $1 \leq a \leq b \leq T-1$, $a+b \leq T-1$. Choose $1 \leq u \leq T-1$ such that

$$a + u \leq T - 1, \quad b + (T - 1 - u) \leq T - 1.$$

The range of such u is

$$b \leq u \leq T - 1 - a.$$

Consider the following set of three paths:

$$W_1 = \{11\ldots1,\ \underbrace{1\ldots1}_{a}\underbrace{2\ldots2}_{T-a},\ \underbrace{1\ldots1}_{b}\underbrace{2\ldots2}_{T-b}\},$$

which consists of a flat path at $i = 1$, a single-step path going from 1 to 2 at time a, and a single-step path going from 1 to 2 at time b. Choose $u \in \{b, \ldots, T-1-a\}$ and let $a' = a + u$, $b' = b + T - 1 - u$. Consider

$$W_2 = \{22\ldots2,\ \underbrace{1\ldots1}_{a'}\underbrace{2\ldots2}_{T-a'},\ \underbrace{1\ldots1}_{b'}\underbrace{2\ldots2}_{T-b'}\}.$$

Then integer array $\boldsymbol{z}$

$$z(\omega) = \begin{cases} 1 & \text{if } \omega \in W_1, \\ -1 & \text{if } \omega \in W_2, \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

is a move for (2). We call this move a degree 3 sliding move. An example of a degree 3 sliding move with $T = 4$, $a = 1$, $b = 2$ and $u = 3$ is depicted in Figure 2 (i). As a degree 3 sliding move, we also consider the time reversal of (13) which involves single-step paths going from 2 to 1 as depicted in Figure 2 (ii).

Now we state the main theorem of this paper.

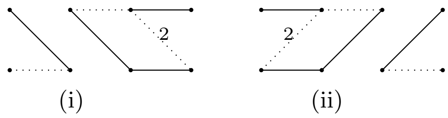


Fig. 2. Sliding moves

Theorem 1. *The set of type II degree one moves of* (11) *and the set of degree 3 sliding moves in* (13)*, in addition to moves in Proposition 1, constitutes a Markov basis for* (2)*.*

We give a proof of this theorem for $T \geq 4$ in the next section. Note that the above statement also covers the case $T = 3$, namely, for $T = 3$, we do not need moves of type (iv) in Proposition 1.

At this point we briefly discuss the case of $T = 3$. Although we can prove Theorem 1 for the case $T = 3$ along the lines of the next section, it is trivial to use 4ti2 ([7]) and verify Theorem 1. It is interesting to note that the degree 3 sliding move for $T = 3$

$$\{111, 122, 122\} \leftrightarrow \{112, 112, 222\}$$

is indispensable ([4], [2]). Hence every Markov basis has to contain a degree three move for $T = 3$.

3 Proof of the main theorem

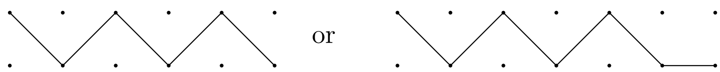
In this section we give a proof of Theorem 1 for $T \geq 4$. We first take care of two types of special fibers. Then we employ a distance reduction argument as in [6] and [5].

Consider a fiber with $b_{11} = 0$ or $b_{22} = 0$. We have the following lemma.

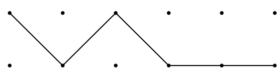
Lemma 1. *Fibers with* $b_{11} = 0$ *or* $b_{22} = 0$ *are connected by type I degree one moves and crossing path swappings.*

Proof. By symmetry consider the case $b_{11} = 0$. Consider an arbitrary fiber $\mathcal{F}_{\boldsymbol{b}}$ with $b_{11} = 0$ and arbitrary $\boldsymbol{x} \in \mathcal{F}_{\boldsymbol{b}}$. It suffices to show that we can transform $\boldsymbol{x}$ to a unique $\boldsymbol{x}^* \in \mathcal{F}_{\boldsymbol{b}}$ by type I degree one moves and crossing path swappings. This can be accomplished as follows. We can easily check that by applying type I degree one moves and crossing path swappings, we can transform paths $w \in W_{1*\cdots*2}$ in $\boldsymbol{x}$ to following three types of paths,

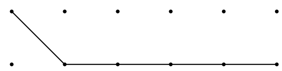
1. $(1212\cdots12)$ when T is even or $(1212\cdots122)$ when T is odd;



2. $(1212\cdots1222\cdots2)$;



3. $(1222\cdots 2)$;



so that the number of type 2 paths is at most one. In the same way we can transform paths $w \in W_{2*\cdots*2}$ in $\boldsymbol{x}$ to

1. $(2121\cdots 2122)$ when T is even and $(2121\cdots 212)$ when T is odd;
2. $(2121\cdots 2122\cdots 2)$;
3. $(22\cdots 2)$;

so that the number of type 2 paths is at most one. The paths $w \in W_{1*\cdots*1}$ in $\boldsymbol{x}$ are transformed to

1. $(1212\cdots 12)$ when T is even and $(1212\cdots 1221)$ when T is odd;
2. $(1212\cdots 1222\cdots 21)$;
3. $(122\cdots 21)$;

so that the number of type 2 paths is at most one. The paths $w \in W_{2*\cdots*1}$ in $\boldsymbol{x}$ are transformed to

1. $(2121\cdots 21)$ when T is even and $(2121\cdots 21221)$ when T is odd;
2. $(2121\cdots 2122\cdots 21)$;
3. $(22\cdots 21)$;

so that the number of type 2 paths is at most one. Then the resulting contingency table is uniquely defined.

Next consider a fiber with $b_{12} = 0$ or $b_{21} = 0$.

Lemma 2. *Fibers with $b_{12} = 0$ or $b_{21} = 0$ are connected by moves of type (iv) in Proposition 1, crossing path swappings and degree 3 sliding moves.*

Proof. By symmetry consider the case $b_{21} = 0$. In this case every path of a contingency table is either a flat path or a single-step path going from 1 to 2. In particular the number of single-step paths is b_{12}. Note that a contingency table of a fiber is determined by choosing positions of transitions for the single-step paths from 1 to 2. A single-step path going from 1 to 2 at time t contains $t-1$ transitions $1 \to 1$ and $T-t-1$ transitions $2 \to 2$. Therefore once we choose positions of the transitions, the number of transitions $1 \to 1$ and $2 \to 2$ contained in these single-step paths are determined. The remaining number of transitions $1 \to 1$ and $2 \to 2$ belong to flat paths. Therefore the remaining numbers have to be multiples of $T-1$. This implies that fiber is partitioned into subsets by the number of flat paths at 1 or at 2. Note that two contingency tables $\boldsymbol{x}, \boldsymbol{y}$ with the same number of flat paths at 1 and at 2 have the same initial frequencies. Then these $\boldsymbol{x}$, $\boldsymbol{y}$ are connected by moves of type (iv) in Proposition 1 and crossing path swappings.

Suppose that initial frequencies of $\boldsymbol{x}$ and $\boldsymbol{y}$ are different. We note that $b_{11} + b_{22} + b_{12} = n(T-1)$. Let $\alpha \in \{0, 1, \ldots, T-2\}$ be the remainder of b_{11} when it is divided by $T-1$ and β be the remainder of b_{22} when it is divided by $T-1$. Then $\alpha+\beta+b_{12}=n'(T-1)$ for some integer $n'\le n$.

Since $0\le\alpha,\beta\le T-2$, $\alpha+\beta+b_{12}=(T-1)$ for $b_{12}=1$. Then we can see that a fiber contains a single contingency table when $b_{12}=1$.

Now consider the case that $b_{12}=2$. It can be easily verified that if $\alpha=T-2$, $\beta=T-2$ and the two transitions from 1 to 2 have to occur at t and $T-t$ $(1\le t\le T-1)$. Then the number of transitions $1\to 1$ and $2\to 2$ in the single-step paths are both $T-2$. Since $\boldsymbol{x}$ and $\boldsymbol{y}$ are in the same fiber, the number of flat paths are the same. Hence the initial frequency of $\boldsymbol{x}$ and $\boldsymbol{y}$ is uniquely determined. Therefore $\boldsymbol{x}$ and $\boldsymbol{y}$ in the same fiber are connected by moves of type (iv) in Proposition 1. If $\alpha<T-2$ or $\beta<T-2$, then the number of flat paths at 1 in the fiber takes two consecutive integer values. Without loss of generality consider the case $x(11\dots1)=y(11\dots1)+1$. Then we can apply a degree 3 sliding move $\boldsymbol{z}$ to $\boldsymbol{x}$ such that $\boldsymbol{x}'(11\dots1)=\boldsymbol{x}(11\dots1)-1$ where $\boldsymbol{x}'=\boldsymbol{x}+\boldsymbol{z}$. Then $\boldsymbol{x}'$ and $\boldsymbol{y}$ are connected by moves of type (iv) in Proposition 1.

For the case $b_{12}\ge 3$, there are more than two single-step paths going from 1 to 2. In this case we pick two such paths and apply moves of type (iv) of Proposition 1, so that the locations of transitions from 1 to 2 of these two paths are far apart from those of other paths. Then we can slide these two paths to alter the initial frequency. This proves the lemma.

After proving above two lemmas, it suffices to consider fibers satisfying

$$b_{11}>0,\ b_{12}>0,\ b_{21}>0,\ b_{22}>0. \tag{14}$$

We now employ distance reduction argument. Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two contingency tables of the same fiber and $\boldsymbol{z}:=\boldsymbol{x}-\boldsymbol{y}$. Define $|\boldsymbol{z}|:=\sum_t\sum_{i,j}|z^t_{ij}|$. If $\boldsymbol{x}$ and $\boldsymbol{y}$ have the same initial frequencies, then they are connected by moves in Proposition 1. We note that

$$b_{21}-b_{12}=x^T_1-x^1_1=y^T_1-y^1_1,$$

which implies $z^1_1=z^T_1$. In the same way we have $z^1_2=z^T_1$. Therefore, without loss of generality we assume

$$z^1_1=a,\ z^T_1=a,\ z^1_2=-a,\ z^T_2=-a,\quad a>0.$$

Lemma 3. *Assume (14) and $T>3$. Suppose that $a=z^1_1>0$ can not be decreased by type II degree one moves. Then $x(11\dots1)>0$ and $y(22\dots2)>0$ or by these moves we can transform $\boldsymbol{x}$ and $\boldsymbol{y}$ to $\boldsymbol{x}'$ and $\boldsymbol{y}'$ such that $x'(11\dots1)>0$ and $y'(22\dots2)>0$.*

Proof. Suppose that there exists a path $w=(s_1,\dots,s_T)\in W_{1*2*1}$ in $\boldsymbol{x}$ such that $s_t=2$ for $1<t<T$. Then by a type II degree one move (11) with $i=1$ and $j=2$, we can reduce a. Therefore we can assume that all paths in $W_{1*\cdots*1}$ are $(11\cdots1)$. By symmetry, we can also assume that all paths of $\boldsymbol{y}$ in $W_{2*\cdots*2}$ are $(22\cdots2)$.

Suppose that $x(11\cdots1)=0$. By the above argument, $\boldsymbol{x}$ has no path in $W_{1*\cdots*1}$. Hence for any path in $\boldsymbol{x}$, $s_1=1$ implies $s_T=2$ and $s_T=1$ implies $s_1=2$. Then we note that $z_2^1=z_2^T=-a$ implies $y(22\cdots2)>0$. Therefore either $x(11\cdots1)>0$ or $y(22\cdots2)>0$ is satisfied. We now assume $x(11\cdots1)>0$ without loss of generality.

Suppose that $y(22\cdots2)=0$. Then for any path in $\boldsymbol{y}$, $s_1=2$ implies $s_T=1$ and $s_T=2$ implies $s_1=1$. Suppose that $\bar{\omega}=(s_1,\ldots,s_T)\in W_{2*\cdots*1}$ is a path in $\boldsymbol{y}$. Since $z_2^1=z_2^T=-a<0$, there has to exist a path $\bar{\omega}'=(s_1',\ldots,s_T')\in W_{1*\cdots*2}$ in $\boldsymbol{y}$.

If $\bar{\omega}$ and $\bar{\omega}'$ meet at t, by applying a crossing path swapping $\boldsymbol{z}$ in (8), $\boldsymbol{y}$ is transformed to $\boldsymbol{y}'=\{y'(w),w\in\{1,2\}^T\}\in\mathcal{F}_{\boldsymbol{b}}$ such that $y'(\tilde{\omega})>0$ and $\tilde{\omega}=(s_1,\ldots,s_t,s_{t+1}',\ldots,s_T')\in W_{2*\cdots*2}$. Then by the above argument we can assume that $y'(22\cdots2)>0$.

Next we consider the case where $\bar{\omega}$ does not meet any $\bar{\omega}'\in W_{1*\cdots*2}$ in $\boldsymbol{y}$. Then $\boldsymbol{y}$ has only one path in each of $W_{2*\cdots*1}$ and $W_{1*\cdots*2}$. Let $\bar{\omega}$ and $\bar{\omega}'$ be such paths. Suppose that $\bar{\omega}=(2121\cdots21)$ and $\bar{\omega}'=(1212\cdots12)$. Then from the assumptions that $b_{11}>0$ and $b_{22}>0$ we have $y'(22\cdots2)>0$.

Suppose that both $\bar{\omega}$ and $\bar{\omega}'$ are single-step paths. If $b_{22}=T-2$,

$$y(\omega)=\begin{cases}1, & \text{if } \omega=\bar{\omega}\text{ or }\bar{\omega}'\\ 0, & \text{otherwise}\end{cases}$$

and $b_{12}=b_{21}=1$. Then $\boldsymbol{x}$ has to contain two single-step paths $\tilde{\omega}$ and $\tilde{\omega}'$ which does not meet each other. $(\bar{\omega},\bar{\omega}')$ is transformed to $(\tilde{\omega},\tilde{\omega}')$ by a 2 by 2 swap (9). Hence $|\boldsymbol{z}|$ is reduced.

If $b_{22}>T-2$, there has to exist another path $\bar{\omega}''=(s_1'',\ldots,s_T'')$ in $\boldsymbol{y}$ such that $s_t''=s_{t+1}''=2$ for some t. By a 2 by 2 swap, $\bar{\omega}$ and $\bar{\omega}'$ are transformed to single-step paths at t. Denote them by $\hat{\omega}=(22\cdots211\cdots1)$ and $\hat{\omega}'=(11\cdots122\cdots2)$. Then apply crossing path swappings as follows,

$$\begin{aligned}\{\hat{\omega},\bar{\omega}''\}&=\{(22\cdots211\cdots1),(\boldsymbol{s}_{1:t}'',\boldsymbol{s}_{t+1:T}'')\}\\&\leftrightarrow\{(22\cdots2,\boldsymbol{s}_{t+1:T}''),(\boldsymbol{s}_{1:t}'',11\cdots1),\}\end{aligned}$$

$$\begin{aligned}\{\hat{\omega}',(22\cdots2,\boldsymbol{s}_{t+1:T}'')\}&=\{(22\cdots22,\boldsymbol{s}_{t+2:T}''),(11\cdots122\cdots2)\}\\&\leftrightarrow\{(22\cdots2),(11\cdots12\boldsymbol{s}_{t+2:T}''),\}\end{aligned}$$

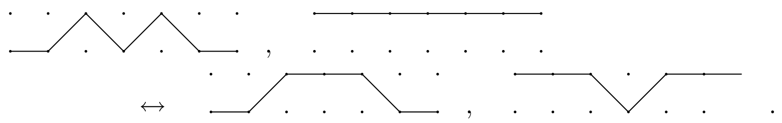
Therefore $\boldsymbol{y}$ is transformed to $\boldsymbol{y}'$ such that $y'(22\cdots2)>0$.

Suppose that both $\bar{\omega}$ and $\bar{\omega}'$ are not single-step paths. Then we can easily see that $\bar{\omega}$ is transformed to another path $\bar{\omega}''=(s_1'',\ldots,s_T'')$ by a type I degree one move. Then ω' and ω'' meet somewhere. Therefore in the same way as the above argument by a crossing path swapping to $\boldsymbol{y}$, $\boldsymbol{y}$ is transformed to $\boldsymbol{y}'$ such that $y'(22\cdots2)>0$.

Lemma 4. *Assume (14) and $T>3$. Suppose that $a=z_1^1>0$ can not be decreased by the moves of Theorem 1 except for degree 3 sliding moves. Then by these moves we can transform $\boldsymbol{x}$ and $\boldsymbol{y}$ to $\boldsymbol{x}'$ and $\boldsymbol{y}'$ which consist of flat paths and single-step paths only.*

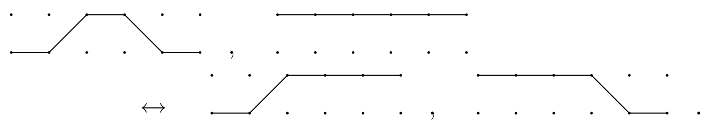
Proof. By Lemma 3, we can assume that $x(11\cdots 1) > 0$. By the argument in the proof of Lemma 3, we can assume that paths in $\boldsymbol{x}$ which start and end at 1 are flat paths at 1 $(11\cdots 1)$. In the same way we can easily show that paths in $\boldsymbol{x}$ which start at 1 and end at 2 are assumed to be single-step paths $(11\cdots 12\cdots 2)$. We can also assume that paths in $\boldsymbol{x}$ which ends at 1 are flat paths at 1 $(11\cdots 1)$ or single-step paths $(22\cdots 21\cdots 1)$.

Next we consider a path ω in $\boldsymbol{x}$ which starts and ends at 2. Suppose that ω is not a flat path at 2. Then there exists $1 < t_0 \le t_1 < T$ such that $\boldsymbol{s}_{1:t_0-1} = (22\cdots 2)$, $s_{t_0} = s_{t_1} = 1$, $\boldsymbol{s}_{t_1+1:T} = (22\cdots 2)$. We now suppose that there exists $t_0 < t_2 < t_1$ such that $s_{t_2} = 2$. Then by crossing path swapping of ω and a flat path ω' at 1, we can transform (ω, ω') to $\tilde{\omega} \in W_{2*1*2}$ and $\tilde{\omega}' \in W_{1*2*1}$ as follows.



Then we can reduce a by applying a type II degree one move to $\tilde{\omega}'$. Hence we can assume that $\boldsymbol{s}_{t_0:t_1} = (11\cdots 1)$.

Since $x(11\cdots 1) > 0$, we can apply a crossing path swapping of ω and a flat path $(11\cdots 1)$, we can transform them to two single-step paths as



Therefore we can transform $\boldsymbol{x}$ to $\boldsymbol{x}'$. In the same way we can show that $\boldsymbol{y}$ is transformed to $\boldsymbol{y}'$.

By Lemma 4 we assume that $\boldsymbol{x}$ and $\boldsymbol{y}$ consist of flat paths and single-step paths only. Now by applying degree 3 sliding moves and adjusting the number of flat paths at 1 as in Lemma 2, we can adjust the initial frequencies of $\boldsymbol{x}$ and $\boldsymbol{y}$. This proves Theorem 1.

4 Concluding remarks

We derived a Markov basis for THMC model without initial parameters (1) for $\mathcal{S} = \{1, 2\}$ and arbitrary $T \ge 3$. The basis consists of moves of degree at most three and the types of moves are common for all $T \ge 4$. For the model (2) we had similar "finiteness" result in [6] with a Markov basis consisting of moves of degree at most two.

Each fiber of (2) is a subset of a fiber in (1). This corresponds to the fact that (1) is a submodel of (2), such that the sufficient statistic for (1) is a linear function of the sufficient statistic for (2). Type II degree one moves and the degree 3 sliding moves are needed to connect fibers of (2) in each fiber of (1). It is of interest to consider other nested toric statistical models and identify moves which are needed to cross fibers of a larger model within each fiber of a smaller model.

References

1. Diaconis, P., Sturmfels, B.: Algebraic algorithms for sampling from conditional distributions. Ann. Statist. 26(1), 363–397 (1998)
2. Ohsugi, H., Hibi, T.: Indispensable binomials of finite graphs. J. Algebra Appl. 4(4), 421–434 (2005)
3. Pachter, L., Sturmfels, B.: Algebraic Statistics for Computational Biology. Cambridge University Press, Cambridge, UK (2005)
4. Takemura, A., Aoki, S.: Some characterizations of minimal Markov basis for sampling from discrete conditional distributions. Ann. Inst. Statist. Math. 56(1), 1–17 (2004)
5. Takemura, A., Aoki, S.: Distance reducing Markov bases for sampling from a discrete sample space. Bernoulli 11(5), 793–813 (2005)
6. Takemura, A., Hara, H.: Markov chain Monte Carlo test of toric homogeneous Markov chains. arXiv:1004.3599v1 (2010)
7. 4ti2 team: 4ti2—a software package for algebraic, geometric and combinatorial problems on linear spaces. Available at www.4ti2.de