MATHEMATICAL ENGINEERING TECHNICAL REPORTS

Distance reducing Markov bases for sampling from discrete sample space

Akimichi TAKEMURA and Satoshi AOKI

METR 2004–15 March 2004

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.i.u-tokyo.ac.jp/mi/mi-e.htm

Distance reducing Markov bases for sampling from discrete sample space

Akimichi TAKEMURA and Satoshi AOKI

Graduate School of Information Science and Technology
University of Tokyo

SUMMARY

We study Markov bases for sampling from discrete sample space, which is equipped with some convenient metric. Starting from any two states in the sample space, we ask whether we can always move closer by an element of a Markov basis. We call a Markov basis *distance reducing* if this is the case. Particular metric we consider in this paper is the 1-norm on the sample space. Some characterizations of 1-norm reducing Markov bases are derived.

1 Introduction

Markov basis for sampling from a discrete conditional distribution is usually studied in the framework of toric ideal and its Gröbner basis (Sturmfels (1995), Diaconis and Sturmfels (1998), Dinwoodie (1998)). For the case of $3 \times 3 \times K$ contingency tables with fixed two-dimensional marginals, in Aoki and Takemura (2003a) we used more elementary approach to derive a unique minimal Markov basis. The approach was based on exhaustive consideration of sign patterns when the 1-norm (L_1-norm) between two contingency tables with the same marginals is minimized. In order to prove that a candidate set $\mathcal{B}$ of moves is a Markov basis, we have shown that the 1-norm between two contingency tables can always be decreased by an element of $\mathcal{B}$.

In order to study minimal Markov basis and its uniqueness for other models, in Takemura and Aoki (2004) we considered whether two elements of the same fiber (reference set) are mutually accessible by a set of lower degree moves and derived some results on the characterization of minimal Markov bases. Note that the notion of mutual accessibility is not directly related to any metric on the fibers. Therefore although the approaches in the above two papers were similar, they were different in explicit consideration of metric on the fibers.

In this paper we explicitly consider 1-norm on the fibers in a general framework of Takemura and Aoki (2004) and derive some characterizations of 1-norm reducing Markov bases. If a Markov basis is 1-norm reducing, then the diameter of each fiber, which is regarded as a transition graph of the Markov chain with respect to the Markov basis, is easily bounded from above. The diameter of the graph is an important factor in various results on the convergence rate of Markov chains (e.g. Section 2.3 of Diaconis and Sturmfels (1998), Section 4 of Diaconis and Saloff-Coste (1998), Bubley and Dyer (1997)). Therefore distance reducing property of Markov bases is also relevant from the viewpoint of the convergence rate of the Markov chain.

The organization of this paper is as follows. In Section 2 we summarize some preliminary material on moves and Markov bases. In Section 3.1 we study general properties of distance reducing Markov basis. In Section 3.2 we derive some properties of 1-norm reducing Markov bases. In Section 4 we give examples and conclude with some discussions.

2 Preliminaries on Markov bases

In this section we summarize preliminary material on Markov bases. We employ the framework of Takemura and Aoki (2004), although we also use well known results on toric ideals and Gröbner basis (e.g. Sturmfels (1995), Hibi (2003)).

2.1 Notations

Here we summarize notations of Takemura and Aoki (2004). Let $\mathcal{I}$ be a finite set with $|\mathcal{I}|$ elements. With multi-way contingency tables in mind, an element of $\mathcal{I}$ is called a *cell* and denoted by $\boldsymbol{i} \in \mathcal{I}$. A non-negative integer $x_{\boldsymbol{i}} \in \mathbb{N} = \{0, 1, \ldots\}$ denotes the frequency of cell $\boldsymbol{i}$. $n = \sum_{\boldsymbol{i}\in\mathcal{I}} x_{\boldsymbol{i}}$ denotes the sample size. Let $a(\boldsymbol{i}) \in \mathbb{N}^\nu, \boldsymbol{i} \in \mathcal{I}$, denote ν-dimensional fixed column vectors consisting of non-negative integers. A ν-dimensional sufficient statistic $\boldsymbol{t}$ is given by $\boldsymbol{t} = \sum_{\boldsymbol{i}\in\mathcal{I}} a(\boldsymbol{i}) x_{\boldsymbol{i}}$.

Let the cells and the vectors $a(\boldsymbol{i})$ be appropriately ordered. Then

$$\boldsymbol{x} = \{x_{\boldsymbol{i}}\}_{\boldsymbol{i}\in\mathcal{I}} \ \in \mathbb{N}^{|\mathcal{I}|}$$

denotes an $|\mathcal{I}|$-dimensional column vector of cell frequencies (*frequency vector*) and

$$A = \{a(\boldsymbol{i})\}_{\boldsymbol{i}\in\mathcal{I}} = (a_{j\boldsymbol{i}})_{j=1,\ldots,\nu,\boldsymbol{i}\in\mathcal{I}}$$

denotes a $\nu \times |\mathcal{I}|$ matrix. Then the sufficient statistic $\boldsymbol{t}$ is written as $\boldsymbol{t} = A\boldsymbol{x}$.

We use the notation $|\boldsymbol{x}| = n = \sum_{\boldsymbol{i}} x_{\boldsymbol{i}}$ to denote the sample size of $\boldsymbol{x}$. It is the 1-norm of $\boldsymbol{x}$. $\boldsymbol{x} \geq 0$ means that the elements of $\boldsymbol{x}$ are non-negative and $\boldsymbol{x} \geq \boldsymbol{y}$ means $\boldsymbol{x} - \boldsymbol{y} \geq 0$. For a given frequency vector $\boldsymbol{x}$, its support is the set of cells with positive frequencies: $\mathrm{supp}(\boldsymbol{x}) = \{\boldsymbol{i} \mid x_{\boldsymbol{i}} > 0\}$. Following Sturmfels (1995) we call the set of $\boldsymbol{x}$'s for a given $\boldsymbol{t}$

$$\mathcal{F}_{\boldsymbol{t}} = \{\boldsymbol{x} \geq 0 \mid A\boldsymbol{x} = \boldsymbol{t}\}.$$

a $\boldsymbol{t}$*-fiber* in this paper.

As in Takemura and Aoki (2004) we assume that the $|\mathcal{I}|$-dimensional row vector $(1, 1, \ldots, 1)$ is a linear combination of the rows of A. This assumption is standard in the theory of toric ideals (Section 4.1 of Hibi (2003)). Under this assumption the sample size n is determined from the sufficient statistic $\boldsymbol{t}$ and all elements of $\mathcal{F}_{\boldsymbol{t}}$ have the same sample size. Somewhat abusing the notation, we write $n = |\boldsymbol{t}|$ to denote the sample size of elements of $\mathcal{F}_{\boldsymbol{t}}$.

Let $\mathbb{Z} = \{0, \pm 1, \ldots\}$. An $|\mathcal{I}|$-dimensional vector of integers $\boldsymbol{z} \in \mathbb{Z}^{|\mathcal{I}|}$ is called a *move* if it is in the kernel of A:

$$A\boldsymbol{z} = 0.$$

For a move $\boldsymbol{z}$, the positive part $\boldsymbol{z}^+$ and the negative part $\boldsymbol{z}^-$ are defined by

$$z_{\boldsymbol{i}}^+ = \max(z_{\boldsymbol{i}}, 0), \qquad z_{\boldsymbol{i}}^- = -\min(z_{\boldsymbol{i}}, 0),$$

respectively. Then $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^-$. The positive part $\boldsymbol{z}^+$ and the negative part $\boldsymbol{z}^-$ have the same value of sufficient statistic $\boldsymbol{t} = A\boldsymbol{z}^+ = A\boldsymbol{z}^-$. The sample size of $\boldsymbol{z}^+$ (or $\boldsymbol{z}^-$) is called the *degree* of $\boldsymbol{z}$ and denoted by

$$\deg \boldsymbol{z} = |\boldsymbol{z}^+| = |\boldsymbol{z}^-|.$$

We also write $|\boldsymbol{z}| = \sum_{\boldsymbol{i}} |z_{\boldsymbol{i}}| = 2 \deg \boldsymbol{z}$, which is the 1-norm of $\boldsymbol{z}$.

We say that a move $\boldsymbol{z}$ is *applicable* to $\boldsymbol{x} \in \mathcal{F}_{\boldsymbol{t}}$ if $\boldsymbol{x} + \boldsymbol{z} \in \mathcal{F}_{\boldsymbol{t}}$, i.e., adding $\boldsymbol{z}$ to $\boldsymbol{x}$ does not produce a negative cell. Clearly $\boldsymbol{z}$ is applicable to $\boldsymbol{x}$ if and only if $\boldsymbol{x} \geq \boldsymbol{z}^-$. Therefore if $\boldsymbol{z}$ is applicable to $\boldsymbol{x}$ then $|\boldsymbol{z}^+| \leq |\boldsymbol{x}|$.

Let $\mathcal{B} = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_L\}$ be a finite set of moves. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{F}_{\boldsymbol{t}}$. We say that $\boldsymbol{y}$ is *accessible* from $\boldsymbol{x}$ by $\mathcal{B}$ and denote it by

$$\boldsymbol{x} \sim \boldsymbol{y} \qquad (\mathrm{mod}\ \mathcal{B}),$$

if there exists a sequence of moves $\boldsymbol{z}_{i_1}, \ldots, \boldsymbol{z}_{i_k}$ of $\mathcal{B}$ and $\epsilon_j = \pm 1$, $j = 1, \ldots, k$, such that $\boldsymbol{y} = \boldsymbol{x} + \sum_{j=1}^{k} \epsilon_j \boldsymbol{z}_{i_j}$ and

$$\boldsymbol{x} + \sum_{j=1}^{h} \epsilon_j \boldsymbol{z}_{i_j} \in \mathcal{F}_{\boldsymbol{t}}, \qquad h = 1, \ldots, k-1, \tag{1}$$

i.e., we can move from $\boldsymbol{x}$ to $\boldsymbol{y}$ by moves of $\mathcal{B}$ without causing negative cells on the way. Obviously the notion of accessibility is symmetric and transitive. Therefore accessibility by $\mathcal{B}$ is an equivalence relation and each $\mathcal{F}_{\boldsymbol{t}}$ is partitioned into disjoint equivalence classes by moves of $\mathcal{B}$. We call these equivalence classes $\mathcal{B}$-equivalence classes of $\mathcal{F}_{\boldsymbol{t}}$. Since the notion of accessibility is symmetric, we also say that $\boldsymbol{x}$ and $\boldsymbol{y}$ are mutually accessible by $\mathcal{B}$ if $\boldsymbol{x} \sim \boldsymbol{y}$ (mod $\mathcal{B}$). Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be elements from two different $\mathcal{B}$-equivalence classes of $\mathcal{F}_{\boldsymbol{t}}$. We say that a move

$$\boldsymbol{z} = \boldsymbol{x} - \boldsymbol{y}$$

connects these two equivalence classes.

$\mathcal{B}$-equivalence classes can be considered in terms of a graph. Consider an undirected graph $G = G_{\boldsymbol{t},\mathcal{B}}$ with the set of vertices $V = \mathcal{F}_{\boldsymbol{t}}$ and the edges between $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{F}_{\boldsymbol{t}}$ if $\boldsymbol{y} = \boldsymbol{x} \pm \boldsymbol{z}$ for some $\boldsymbol{z} \in \mathcal{B}$. $\mathcal{B}$-equivalence classes are connected components of G.

In the following let

$$\mathcal{B}_n = \{\boldsymbol{z} \mid \deg \boldsymbol{z} \leq n\},$$

denote the set of moves with degree less than or equal to n.

2.2 Markov bases

A set of finite moves $\mathcal{B} = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_L\}$ is a *Markov basis* if for all $\boldsymbol{t}$, $\mathcal{F}_{\boldsymbol{t}}$ itself constitutes one $\mathcal{B}$-equivalence class, i.e. the graph $G_{\boldsymbol{t},\mathcal{B}}$ is connected for all $\boldsymbol{t}$. A Markov basis $\mathcal{B}$ is *minimal* if no proper subset of $\mathcal{B}$ is a Markov basis. A minimal Markov basis always exists, because from any Markov basis, we can remove redundant elements one by one, until none of the remaining elements can be removed any further. Minimal basis is *unique* if all minimal bases coincide except for sign changes of their elements.

A move $\boldsymbol{z}$ is *indispensable* if $\boldsymbol{z}$ or $-\boldsymbol{z}$ belongs to every Markov basis. $\boldsymbol{z}$ is indispensable if and only if it is the difference of elements of a two-elements fiber

$$\boldsymbol{z} = \boldsymbol{x} - \boldsymbol{y}, \qquad \{\boldsymbol{x}, \boldsymbol{y}\} = \mathcal{F}_{\boldsymbol{t}} \text{ for some } \boldsymbol{t}.$$

Minimal basis is unique if and only if the set of indispensable moves form a Markov basis. Recently Ohsugi and Hibi (2003) have shown the important fact that the set of indispensable moves is the intersection of all reduced Gröbner bases.

Let $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^-$ be a move of degree n. We call $\boldsymbol{z}$ *non-replaceable by lower degree moves* if

$$\boldsymbol{z}^+ \not\sim \boldsymbol{z}^- \pmod{\mathcal{B}_{n-1}}, \tag{2}$$

i.e., if $\boldsymbol{z}$ connects different $\mathcal{B}_{n-1}$-equivalence classes of $\mathcal{F}_{\boldsymbol{t}} \ni \boldsymbol{z}^+$. Clearly an indispensable move is non-replaceable by lower degree moves. From the argument in Takemura and Aoki (2004) we see that the set

$$\mathcal{B}_{\mathrm{MF}} = \{\boldsymbol{z} \mid \boldsymbol{z} \text{ is non-replaceable by lower degree moves}\}$$

form a Markov basis. We call $\mathcal{B}_{\mathrm{MF}}$ *minimum fiber Markov basis*, in view of the fact that for all minimal Markov bases $\mathcal{B}$, the set of sufficient statistics

$$\{\boldsymbol{t} \mid \boldsymbol{t} = A\boldsymbol{x}, \boldsymbol{x} \in \mathcal{B}\} = \{\boldsymbol{t} \mid \mathcal{F}_{\boldsymbol{t}} \text{ is not a single } \mathcal{B}_{|\boldsymbol{t}|-1}\text{-equivalence class}\}$$

is common. $\mathcal{B}_{\mathrm{MF}}$ is the union of all minimal Markov bases. In the case of contingency tables with fixed marginals, the minimum fiber Markov basis is invariant in the sense of Aoki and Takemura (2003b).

Consider a sum of two moves $\boldsymbol{z} = \boldsymbol{z}_1 + \boldsymbol{z}_2$. We say that there is no cancellation of signs in this sum if

$$\mathrm{supp}(\boldsymbol{z}^+) = \mathrm{supp}(\boldsymbol{z}_1^+) \cup \mathrm{supp}(\boldsymbol{z}_2^+), \quad \mathrm{supp}(\boldsymbol{z}^-) = \mathrm{supp}(\boldsymbol{z}_1^-) \cup \mathrm{supp}(\boldsymbol{z}_2^-).$$

In this case we also say that $\boldsymbol{z}$ is a conformal sum of $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$. Similarly we say that there is no cancellation of signs in the sum of m moves $\boldsymbol{z} = \boldsymbol{z}_1 + \cdots + \boldsymbol{z}_m$ (or $\boldsymbol{z}$ is a conformal sum of m moves) if

$$\mathrm{supp}(\boldsymbol{z}^+) = \mathrm{supp}(\boldsymbol{z}_1^+) \cup \cdots \cup \mathrm{supp}(\boldsymbol{z}_m^+), \quad \mathrm{supp}(\boldsymbol{z}^-) = \mathrm{supp}(\boldsymbol{z}_1^-) \cup \cdots \cup \mathrm{supp}(\boldsymbol{z}_m^-).$$

A move $\boldsymbol{z}$ is *primitive* if it can not be written as sum of two non-zero moves $\boldsymbol{z} = \boldsymbol{z}_1 + \boldsymbol{z}_2$ with no cancellation of signs. Clearly a move $\boldsymbol{z}$, which is non-replaceable by lower degree moves, is primitive. Note that $\boldsymbol{z}$ is primitive if it can not be written as $\boldsymbol{z} = \boldsymbol{z}_1 + \boldsymbol{z}_2$ such that $|\boldsymbol{z}| = |\boldsymbol{z}_1| + |\boldsymbol{z}_2|$ and $0 < |\boldsymbol{z}_1|, |\boldsymbol{z}_2| < |\boldsymbol{z}|$. The set of primitive moves is called the *Graver basis* and it is a Markov basis. If a move $\boldsymbol{z}$ is not primitive, then we can recursively decompose $\boldsymbol{z}$ into a conformal sum of moves. This implies that any move $\boldsymbol{z}$ can be written as a conformal sum

$$\boldsymbol{z} = \boldsymbol{z}_1 + \cdots + \boldsymbol{z}_m, \tag{3}$$

where $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m$ are (not necessarily distinct) non-zero elements of the Graver basis. Because of no cancellation of signs, whenever $\boldsymbol{z}$ is applicable to some $\boldsymbol{x}$, $\boldsymbol{z}$ can be replaced by $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m$ in arbitrary order without causing negative cells on the way.

Let $\{u_{\boldsymbol{i}}\}_{\boldsymbol{i} \in \mathcal{I}}$ be the set of indeterminates and let $K[\{u_{\boldsymbol{i}}\}_{\boldsymbol{i} \in \mathcal{I}}]$ denote the polynomial ring in the indeterminates $\{u_{\boldsymbol{i}}\}_{\boldsymbol{i} \in \mathcal{I}}$ over a field K. A frequency vector $\boldsymbol{x} = \{x_{\boldsymbol{i}}\}_{\boldsymbol{i} \in \mathcal{I}}$ can be identified with the monomial

$$\prod_{\boldsymbol{i} \in \mathcal{I}} u_{\boldsymbol{i}}^{x_{\boldsymbol{i}}}.$$

Let $\{t_j\}_{j=1,\dots,\nu}$ denote the set of indeterminates corresponding to the sufficient statistic $\boldsymbol{t}$ and let $K[\boldsymbol{t}]$ denote the polynomial ring in $t_1, \dots, t_\nu$ over K. The toric ideal I_A is the kernel of the ring homomorphism

$$u_{\boldsymbol{i}} \mapsto \prod_{j=1}^{\nu} t_j^{a_{ji}}.$$

Given a term order $\prec$ on $K[\{u_{\boldsymbol{i}}\}_{\boldsymbol{i}\in\mathcal{I}}]$ let $\mathcal{B}_\prec$ denote the reduced Gröbner basis of I_A with respect to $\prec$. It is well known (Sturmfels (1995), Diaconis and Sturmfels (1998)) that $\mathcal{B}_\prec$ is a Markov basis. The union of all reduced Gröbner bases for all possible term orders is called the universal Gröbner basis and it is contained in the Graver basis.

3 Distance reducing Markov bases

In this section we first define the notion of distance reduction by a set of moves when the sample space is equipped with appropriate metric. Then we derive some characterizations of 1-norm reducing Markov bases.

3.1 Distance reduction by a set of moves

Consider a metric $d(\boldsymbol{x}, \boldsymbol{y})$ on a fiber $\mathcal{F}_{\boldsymbol{t}}$. Although we are mainly concerned with 1-norm in the following, here we consider a general metric. A metric $d = d_{\boldsymbol{t}}$ on $\mathcal{F}_{\boldsymbol{t}}$ can be defined in various ways. If a metric d is defined on the whole sample space $\mathbb{N}^{|\mathcal{I}|}$, we consider the restriction of d onto each $\mathcal{F}_{\boldsymbol{t}}$, i.e., $d_{\boldsymbol{t}}(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{x}, \boldsymbol{y})$, $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{F}_{\boldsymbol{t}}$. If d is a norm on the set $\mathbb{Z}^{|\mathcal{I}|}$ of integer frequency vectors, such as the 1-norm $|\boldsymbol{z}|$, d is defined by $d_{\boldsymbol{t}}(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{x} - \boldsymbol{y})$. For notational simplicity we suppress the subscript $\boldsymbol{t}$ in $d_{\boldsymbol{t}}$ hereafter.

Now we introduce the notion of distance reduction by a set of moves. Let $\mathcal{B}$ denote a set of moves. We call $\mathcal{B}$ *d-reducing for $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{F}_{\boldsymbol{t}}$* if there exists an element $\boldsymbol{z} \in \mathcal{B}$ and $\epsilon = \pm 1$ such that $\epsilon\boldsymbol{z}$ is applicable to $\boldsymbol{x}$ or $\boldsymbol{y}$ and we can decrease the distance, i.e.,

$$\begin{aligned} &\boldsymbol{x} + \epsilon\boldsymbol{z} \in \mathcal{F}_{\boldsymbol{t}} \quad \text{and} \quad d(\boldsymbol{x} + \epsilon\boldsymbol{z}, \boldsymbol{y}) < d(\boldsymbol{x}, \boldsymbol{y}), \qquad \text{or} \\ &\boldsymbol{y} + \epsilon\boldsymbol{z} \in \mathcal{F}_{\boldsymbol{t}} \quad \text{and} \quad d(\boldsymbol{x}, \boldsymbol{y} + \epsilon\boldsymbol{z}) < d(\boldsymbol{x}, \boldsymbol{y}). \end{aligned} \tag{4}$$

We simply call $\mathcal{B}$ *d-reducing* if $\mathcal{B}$ is d-reducing for every $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{F}_{\boldsymbol{t}}$ and for every $\boldsymbol{t}$.

We call $\mathcal{B}$ *strongly d-reducing for $\boldsymbol{x}, \boldsymbol{y}$* if there exist elements $\boldsymbol{z}_1, \boldsymbol{z}_2 \in \mathcal{B}$ and $\epsilon_1, \epsilon_2 = \pm 1$ such that $\boldsymbol{x} + \epsilon_1\boldsymbol{z}_1 \in \mathcal{F}_{\boldsymbol{t}}$, $\boldsymbol{y} + \epsilon_2\boldsymbol{z}_2 \in \mathcal{F}_{\boldsymbol{t}}$ and

$$d(\boldsymbol{x} + \epsilon_1\boldsymbol{z}_1, \boldsymbol{y}) < d(\boldsymbol{x}, \boldsymbol{y}) \quad \text{and} \quad d(\boldsymbol{x}, \boldsymbol{y} + \epsilon_2\boldsymbol{z}_2) < d(\boldsymbol{x}, \boldsymbol{y}) \tag{5}$$

and call $\mathcal{B}$ *strongly d-reducing* if $\mathcal{B}$ is strongly d-reducing for every $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{F}_{\boldsymbol{t}}$ and for every $\boldsymbol{t}$. Clearly if $\mathcal{B}$ is strongly d-reducing, then $\mathcal{B}$ is d-reducing.

A basic argument we have employed in Aoki and Takemura (2003a) is the following obvious fact.

Proposition 1 *Let a metric d is given on each fiber $\mathcal{F}_{\boldsymbol{t}}$. A set of finite moves $\mathcal{B}$ is a Markov basis if it is d-reducing.*

If $\mathcal{B}$ is d-reducing, then for every $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{F}_{\boldsymbol{t}}$, there exist $k > 0$, $\epsilon_l = \pm 1$, $\boldsymbol{z}_l \in \mathcal{B}$, $\boldsymbol{x}_l \in \mathcal{F}_{\boldsymbol{t}}$, $\boldsymbol{y}_l \in \mathcal{F}_{\boldsymbol{t}}$, $l = 1, \ldots, k$, with the following properties: i) $\boldsymbol{x}_k = \boldsymbol{y}_k$,

$$\text{ii)} \qquad d(\boldsymbol{x}_l, \boldsymbol{y}_l) < d(\boldsymbol{x}_{l-1}, \boldsymbol{y}_{l-1}), \qquad l = 1, \ldots, k,$$

where $\boldsymbol{x}_0 = \boldsymbol{x}$ and $\boldsymbol{y}_0 = \boldsymbol{y}$, and

$$\text{iii)} \qquad (\boldsymbol{x}_l, \boldsymbol{y}_l) = (\boldsymbol{x}_{l-1} + \epsilon_l \boldsymbol{z}_l, \boldsymbol{y}_{l-1}) \quad \text{or} \quad (\boldsymbol{x}_l, \boldsymbol{y}_l) = (\boldsymbol{x}_{l-1}, \boldsymbol{y}_{l-1} + \epsilon_l \boldsymbol{z}_l), \qquad l = 1, \ldots, k.$$

Given the above sequence of frequency vectors, we can move from $\boldsymbol{x}$ to $\boldsymbol{x}_k = \boldsymbol{y}_k$ and then reversing the moves we can move from $\boldsymbol{y}_k$ to $\boldsymbol{y}$. Thus $\boldsymbol{y}$ is accessible from $\boldsymbol{x}$ by $\mathcal{B}$. Note that in this sequence of moves the distances

$$d(\boldsymbol{x}, \boldsymbol{x}_1), \ldots, d(\boldsymbol{x}, \boldsymbol{x}_k), d(\boldsymbol{x}, \boldsymbol{y}_{k-1}) \ldots, d(\boldsymbol{x}, \boldsymbol{y})$$

might not be monotone increasing.

On the other hand if $\mathcal{B}$ is strongly d-reducing, then starting from $\boldsymbol{y}$, we can always decrease the distance by moving from the side of $\boldsymbol{y}$, i.e., we can find $k > 0$ and $\boldsymbol{y} = \boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{k-1}, \boldsymbol{y}_k = \boldsymbol{x}$ in $\mathcal{F}_{\boldsymbol{t}}$ such that $\boldsymbol{y}_l = \boldsymbol{y}_{l-1} + \epsilon_l \boldsymbol{z}_l$, $\epsilon_l = \pm 1$, $\boldsymbol{z}_l \in \mathcal{B}$, $l = 1, \ldots, k$, and

$$d(\boldsymbol{x}, \boldsymbol{y}_{k-1}), d(\boldsymbol{x}, \boldsymbol{y}_{k-2}), \ldots, d(\boldsymbol{x}, \boldsymbol{y})$$

is strictly increasing.

Given a Markov basis $\mathcal{B}$, each fiber $\mathcal{F}_{\boldsymbol{t}}$ can be considered as a connected graph $G_{\boldsymbol{t},\mathcal{B}}$ as discussed in Section 2. Associated with a Markov basis $\mathcal{B}$, the *shortest path metric* $d_{\mathcal{B}}(\boldsymbol{x}, \boldsymbol{y})$ is defined to be the minimum number of steps needed to move from $\boldsymbol{x}$ to $\boldsymbol{y}$ by moves of $\mathcal{B}$. Trivially $\mathcal{B}$ is itself $d_{\mathcal{B}}$-reducing. As discussed in Introduction, the diameter of $G_{\boldsymbol{t},\mathcal{B}}$ is relevant for the convergence rate of the Markov chain based on the Markov basis $\mathcal{B}$. However direct computation of the diameter is generally difficult, whereas the largest 1-norm between two states in $\mathcal{F}_{\boldsymbol{t}}$ is clearly bounded from above by $2|\boldsymbol{t}|$. This implies that if $\mathcal{B}$ is 1-norm reducing, then the diameter of $G_{\boldsymbol{t},\mathcal{B}}$ is bounded from above by $2|\boldsymbol{t}|$.

3.2 Some results on 1-norm reducing Markov bases

The 1-norm $|\boldsymbol{z}| = \sum_{\boldsymbol{i} \in \mathcal{I}} |z_{\boldsymbol{i}}|$ on the set $\mathbb{Z}^{|\mathcal{I}|}$ of integer frequency vectors is a natural norm to consider for Markov bases. In this section we investigate Markov bases from the viewpoint of 1-norm reduction.

Here we briefly summarize results of this section. First we discuss the basic importance of the Graver basis in the investigation of 1-norm reduction (Propositions 2, 3, 4). Then we introduce three closely related notions of 1-norm irreducibility of a move and discuss implications among them (Proposition 5). Based on these notions, conditions for unique minimality of (strongly) 1-norm reducing Markov bases are given (Propositions 6, 7). In Proposition 8 we introduce a Markov basis $\mathcal{B}_{\mathrm{LDI}}$ containing $\mathcal{B}_{\mathrm{MF}}$ and in Proposition 9 we discuss the case that $\mathcal{B}_{\mathrm{LDI}}$ consists of the indispensable moves.

We now start with the Graver basis.

Proposition 2 *Graver basis is strongly 1-norm reducing.*

Proof. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{F}_{\boldsymbol{t}}$ be in the same fiber. As in (3), express $\boldsymbol{x} - \boldsymbol{y}$ as a conformal sum of non-zero elements of the Graver basis:

$$\boldsymbol{x} - \boldsymbol{y} = \boldsymbol{z}_1 + \cdots + \boldsymbol{z}_m.$$

Then $|\boldsymbol{x} - \boldsymbol{y}| = |\boldsymbol{z}_1| + \cdots + |\boldsymbol{z}_m|$. Now $\boldsymbol{z}_1$ can be subtracted from $\boldsymbol{x}$ and at the same time $\boldsymbol{z}_1$ can be added to $\boldsymbol{y}$ to give

$$|(\boldsymbol{x} - \boldsymbol{z}_1) - \boldsymbol{y}| = |\boldsymbol{x} - (\boldsymbol{y} + \boldsymbol{z}_1)| = |\boldsymbol{z}_2| + \cdots + |\boldsymbol{z}_m| < |\boldsymbol{x} - \boldsymbol{y}|.$$

Q.E.D.

Note that the Graver basis is rich enough that we can take $\boldsymbol{z}_1 = \boldsymbol{z}_2$ in the definition of strong distance reduction in (5)

Proposition 3 *A set of moves $\mathcal{B}$ is 1-norm reducing if and only if for every element $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^-$ of the Graver basis, $\mathcal{B}$ is 1-norm reducing for $\boldsymbol{z}^+, \boldsymbol{z}^-$.*

Proof. We only have to prove sufficiency. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{F}_{\boldsymbol{t}}$ be arbitrarily given and let $\boldsymbol{x} - \boldsymbol{y} = \boldsymbol{z}_1 + \cdots + \boldsymbol{z}_m$ be a conformal sum of elements of the Graver basis. By assumption $\mathcal{B}$ is 1-norm reducing for $\boldsymbol{z}_1^+, \boldsymbol{z}_1^-$. Among four possible cases, without loss of generality, consider the case that $\boldsymbol{z} \in \mathcal{B}$ is applicable to $\boldsymbol{z}_1^+$ and

$$|(\boldsymbol{z}_1^+ + \boldsymbol{z}) - \boldsymbol{z}_1^-| < |\boldsymbol{z}_1^+ - \boldsymbol{z}_1^-| = |\boldsymbol{z}_1|. \tag{6}$$

Since $\boldsymbol{z}$ is applicable to $\boldsymbol{z}_1^+$, $\boldsymbol{z}^- \le \boldsymbol{z}_1^+ \le (\boldsymbol{x} - \boldsymbol{y})^+$. Furthermore (6) implies that

$$\emptyset \ne \mathrm{supp}(\boldsymbol{z}^+) \cap \mathrm{supp}(\boldsymbol{z}_1^-) \subset \mathrm{supp}(\boldsymbol{z}^+) \cap \mathrm{supp}((\boldsymbol{x} - \boldsymbol{y})^-).$$

It follows that $\boldsymbol{z}$ is applicable to $\boldsymbol{x}$ and $|(\boldsymbol{x} + \boldsymbol{z}) - \boldsymbol{y}| < |\boldsymbol{x} - \boldsymbol{y}|$. Q.E.D.

Note that the same statement holds for strong 1-norm reduction with exactly the same proof:

Proposition 4 *A set of moves $\mathcal{B}$ is strongly 1-norm reducing if and only if for every element $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^-$ of the Graver basis, $\mathcal{B}$ is strongly 1-norm reducing for $\boldsymbol{z}^+, \boldsymbol{z}^-$.*

Suppose that $\mathcal{B}$ is a 1-norm reducing Markov basis. Then any $\mathcal{B}' \supset \mathcal{B}$ is a 1-norm reducing Markov basis as well. In view of this, it is of interest to consider minimality of 1-norm reducing Markov bases. A 1-norm reducing Markov basis $\mathcal{B}$ is minimal if every proper subset of $\mathcal{B}$ is not a 1-norm reducing Markov basis. For 1-norm reducing Markov basis $\mathcal{B}$, we can examine each element $\boldsymbol{z}$ of $\mathcal{B}$ one by one, whether $\mathcal{B} - \{\boldsymbol{z}\}$ remains to be a 1-norm reducing Markov basis. If $\mathcal{B} - \{\boldsymbol{z}\}$ remains to be 1-norm reducing, we remove $\boldsymbol{z}$, recursively, until none of the remaining elements can be removed any further. Then we arrive at a minimal 1-norm reducing Markov basis. Therefore every 1-norm reducing Markov basis $\mathcal{B}$ contains a minimal 1-norm reducing Markov basis.

Exactly the same argument holds concerning minimality of strongly 1-norm reducing Markov bases. Every strongly 1-norm reducing Markov basis contains a minimal strongly 1-norm reducing Markov basis.

In Takemura and Aoki (2004) we considered minimality of Markov bases. Similar argument can be applied to the question of minimality of 1-norm reducing Markov bases.

In order to study this minimality question we introduce three closely related notions of degree reduction of a move $\boldsymbol{z}$ by other moves. We say that a move $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^-$ is *1-norm reducible by another move* $\boldsymbol{z}' \neq \pm\boldsymbol{z}$ if $\boldsymbol{z}'$ is applicable to $\boldsymbol{z}^+$ and $|\boldsymbol{z}+\boldsymbol{z}'| < |\boldsymbol{z}|$ or $\boldsymbol{z}'$ is applicable to $\boldsymbol{z}^-$ and $|-\boldsymbol{z}+\boldsymbol{z}'| = |\boldsymbol{z}-\boldsymbol{z}'| < |\boldsymbol{z}|$. We say that a move $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^-$ is *strongly 1-norm reducible by a pair of (other) moves* $\boldsymbol{z}_1, \boldsymbol{z}_2 \neq \pm\boldsymbol{z}$ if $\boldsymbol{z}_1$ is applicable to $\boldsymbol{z}^+$ and $|\boldsymbol{z}+\boldsymbol{z}_1| < |\boldsymbol{z}|$ and furthermore $\boldsymbol{z}_2$ is applicable to $\boldsymbol{z}^-$ and $|\boldsymbol{z}-\boldsymbol{z}_2| < |\boldsymbol{z}|$. Finally we say $\boldsymbol{z}$ is *1-norm reducible by a lower degree move* $\boldsymbol{z}'$ if $|\boldsymbol{z}'| < |\boldsymbol{z}|$ and $\boldsymbol{z}$ is 1-norm reducible by $\boldsymbol{z}'$.

Consider the implications among these notions. If $\boldsymbol{z}$ is strongly 1-norm reducible by $\boldsymbol{z}_1, \boldsymbol{z}_2$, then $\boldsymbol{z}$ is clearly 1-norm reducible by $\boldsymbol{z}_1$ (or $\boldsymbol{z}_2$). Now we show that if $\boldsymbol{z}$ is 1-norm reducible by a lower degree move $\boldsymbol{z}'$, then $\boldsymbol{z}$ is strongly 1-norm reducible either by the pair $\boldsymbol{z}', \boldsymbol{z}+\boldsymbol{z}'$ or by the pair $\boldsymbol{z}'-\boldsymbol{z}, \boldsymbol{z}'$. To show this first consider the case that $\boldsymbol{z}'$ is applicable to $\boldsymbol{z}^+$ and $|\boldsymbol{z}+\boldsymbol{z}'| < |\boldsymbol{z}|$. Let $\boldsymbol{z}'' = \boldsymbol{z}+\boldsymbol{z}'$. Then $|\boldsymbol{z}-\boldsymbol{z}''| = |\boldsymbol{z}'| < |\boldsymbol{z}|$ and we only need to check that $\boldsymbol{z}''$ is applicable to $\boldsymbol{z}^-$. In fact

$$\begin{aligned} \boldsymbol{z}'' &= \boldsymbol{z}^+ - \boldsymbol{z}^- + (\boldsymbol{z}')^+ - (\boldsymbol{z}')^- = (\boldsymbol{z}^+ - (\boldsymbol{z}')^-) + (\boldsymbol{z}')^+ - \boldsymbol{z}^- \\ &\geq (\boldsymbol{z}')^+ - \boldsymbol{z}^-. \end{aligned}$$

This implies that $(\boldsymbol{z}'')^- \leq \boldsymbol{z}^-$ and $\boldsymbol{z}''$ is applicable to $\boldsymbol{z}^-$. Similarly if $\boldsymbol{z}'$ is applicable to $\boldsymbol{z}^-$, we can check that $\boldsymbol{z}$ is strongly 1-norm reducible by the pair $\boldsymbol{z}'-\boldsymbol{z}, \boldsymbol{z}'$.

Based on the above observation, we define three notions of irreducibility of a move. We call $\boldsymbol{z}$ *1-norm irreducible* if it is not 1-norm reducible by any other move $\boldsymbol{z}' \neq \boldsymbol{z}$. We call $\boldsymbol{z}$ *strongly 1-norm irreducible* if it is not strongly 1-norm reducible by any pair of other moves. Finally we call $\boldsymbol{z}$ *1-norm lower degree irreducible* if it is not 1-norm reducible by any lower degree move. We state the above implications of the properties of moves, as well as further implications among indispensability and primitiveness, in the following proposition.

Proposition 5 *For a move $\boldsymbol{z}$, the following implications hold.*

$$\begin{aligned} \textit{indispensable} &\Rightarrow \textit{1-norm irreducible} \\ &\Rightarrow \textit{strongly 1-norm irreducible} \\ &\Rightarrow \textit{1-norm lower degree irreducible} \\ &\Rightarrow \textit{primitive.} \end{aligned} \tag{7}$$

Proof. If $\boldsymbol{z}$ is not primitive, then $\boldsymbol{z}$ is clearly 1-norm reducible by a lower degree move. This proves the last implication.

If $\boldsymbol{z}$ is 1-norm reducible by $\boldsymbol{z}' \neq \pm\boldsymbol{z}$, then $\boldsymbol{z}^- \neq \boldsymbol{z}^+ + \boldsymbol{z}' \in \mathcal{F}_{\boldsymbol{t}}$ or $\boldsymbol{z}^+ \neq \boldsymbol{z}^- + \boldsymbol{z}' \in \mathcal{F}_{\boldsymbol{t}}$, where $\boldsymbol{t} = A\boldsymbol{z}^+$. Therefore $\mathcal{F}_{\boldsymbol{t}}$ is not a two-elements fiber. Therefore $\boldsymbol{z}$ is not indispensable. This proves the first implication. Q.E.D.

We now prepare several lemmas.

Lemma 1 *If $\boldsymbol{z}$ is 1-norm reducible by another move $\boldsymbol{z}' \neq \pm\boldsymbol{z}$, then there exists a primitive move $\boldsymbol{z}'' \neq \boldsymbol{z}$, $|\boldsymbol{z}''| \leq |\boldsymbol{z}'|$, such that $\boldsymbol{z}$ is 1-norm reducible by $\boldsymbol{z}''$.*

Proof. If $\boldsymbol{z}'$ is itself primitive just let $\boldsymbol{z}'' = \boldsymbol{z}'$. If $\boldsymbol{z}'$ is not primitive write $\boldsymbol{z}'$ as a conformal sum $\boldsymbol{z}' = \boldsymbol{z}_1 + \cdots + \boldsymbol{z}_m$ of non-zero elements of the Graver basis. Among two possible cases, without loss of generality, consider the case that $\boldsymbol{z}'$ is applicable to $\boldsymbol{z}^+$ and $|\boldsymbol{z} + \boldsymbol{z}'| < |\boldsymbol{z}|$. In this case $(\boldsymbol{z}')^- \le \boldsymbol{z}^+$ and $\mathrm{supp}((\boldsymbol{z}')^+) \cap \mathrm{supp}(\boldsymbol{z}^-) \neq \emptyset$. Since $\mathrm{supp}((\boldsymbol{z}')^+) = \mathrm{supp}(\boldsymbol{z}_1^+) \cup \cdots \cup \mathrm{supp}(\boldsymbol{z}_m^+)$, there exists some l such that $\mathrm{supp}((\boldsymbol{z}_l)^+) \cap \mathrm{supp}(\boldsymbol{z}^-) \neq \emptyset$. Furthermore $\boldsymbol{z}_l^- \le (\boldsymbol{z}')^- \le \boldsymbol{z}^+$ and $|\boldsymbol{z}_l| < |\boldsymbol{z}'| \le |\boldsymbol{z}|$. This implies that $\boldsymbol{z}$ is 1-norm reducible by $\boldsymbol{z}'' = \boldsymbol{z}_l \neq \boldsymbol{z}$. Q.E.D.

The same argument proves the following lemma.

Lemma 2 *If $\boldsymbol{z}$ is strongly 1-norm reducible by a pair of moves $\boldsymbol{z}_1, \boldsymbol{z}_2 \neq \pm\boldsymbol{z}$ then there exists a pair of primitive moves $\boldsymbol{z}_1', \boldsymbol{z}_2' \neq \pm\boldsymbol{z}$, $|\boldsymbol{z}_1'| \le |\boldsymbol{z}_1|$, $|\boldsymbol{z}_2'| \le |\boldsymbol{z}_2|$, such that $\boldsymbol{z}$ is strongly 1-norm reducible by the pair $\boldsymbol{z}_1', \boldsymbol{z}_2'$.*

We now state some results on minimality of 1-norm reducing Markov bases. First we show that 1-norm reducing basis has to contain all 1-norm irreducible moves.

Lemma 3 *Let $\boldsymbol{z}$ be a 1-norm irreducible move. Then either $\boldsymbol{z}$ or $-\boldsymbol{z}$ belongs to every 1-norm reducing Markov basis.*

Proof. We argue by contradiction. Let $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^-$ be 1-norm irreducible and let $\mathcal{B}$ be a 1-norm reducing Markov basis not containing $\boldsymbol{z}$ nor $-\boldsymbol{z}$. Since $\mathcal{B}$ be 1-norm reducing, $\mathcal{B}$ is 1-norm reducing for $\boldsymbol{z}^+, \boldsymbol{z}^-$. But this contradicts the 1-norm irreducibility of $\boldsymbol{z}$ in view of (4). Q.E.D.

We say that there exists a unique minimal 1-norm reducing Markov basis if all minimal 1-norm reducing Markov bases coincide except for sign changes of their elements.

Proposition 6 *There exists a unique minimal 1-norm reducing Markov basis if and only if 1-norm irreducible moves form a 1-norm reducing Markov basis.*

Proof. Since every 1-norm irreducible move (or its sign change) belongs to every 1-norm reducing Markov basis, if the set of 1-norm irreducible moves is a 1-norm reducing Markov basis, then it is clearly the unique minimal 1-norm reducing Markov basis ignoring the sign of each move.

Conversely suppose that 1-norm irreducible moves do not form a 1-norm reducing Markov basis. Then every 1-norm reducing Markov basis contains a 1-norm reducible move. Let $\mathcal{B}$ be a minimal 1-norm reducing Markov basis and let $\boldsymbol{z}_0 \in \mathcal{B}$ be 1-norm reducible. Consider

$$\tilde{\mathcal{B}} = (\mathcal{B} \cup \mathcal{B}_{\text{Graver}}) - \{\boldsymbol{z}_0, -\boldsymbol{z}_0\},$$

where $\mathcal{B}_{\text{Graver}}$ is the Graver basis. We show that $\tilde{\mathcal{B}}$ is a 1-norm reducing Markov basis. If this is the case $\tilde{\mathcal{B}}$ contains a minimal 1-norm reducing Markov basis different from $\mathcal{B}$ even if we change the signs the elements.

Now by Propositions 1 and 3, it suffices to show that for every $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^- \in \mathcal{B}_{\text{Graver}}$, $\tilde{\mathcal{B}}$ is 1-norm reducing for $\boldsymbol{z}^+, \boldsymbol{z}^-$. If $\boldsymbol{z}_0$ is not primitive, $\tilde{\mathcal{B}} \supset \mathcal{B}_{\text{Graver}}$ and $\tilde{\mathcal{B}}$ is 1-norm reducing. Therefore let $\boldsymbol{z}_0$ be primitive. Each primitive $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^- \neq \boldsymbol{z}_0$ is already in $\tilde{\mathcal{B}}$ and $\tilde{\mathcal{B}}$ is 1-norm reducing for $\boldsymbol{z}^+, \boldsymbol{z}^-$. The only remaining case is $\boldsymbol{z} = \boldsymbol{z}_0$ itself, but by Lemma 1, $\boldsymbol{z}_0$ is 1-norm reducible by a primitive $\boldsymbol{z}' \neq \pm\boldsymbol{z}_0$, $\boldsymbol{z}' \in \tilde{\mathcal{B}}$. Q.E.D.

Remark 1 *In many examples we expect that the set of 1-norm irreducible moves form a Markov basis. However it seems difficult to state a simple sufficient condition to guarantee this. The difficulty lies in eliminating the following possibility. Suppose that for some* $\boldsymbol{t}$ *with* $|\boldsymbol{t}| = n$*, the fiber* $\mathcal{F}_{\boldsymbol{t}} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3\}$ *is a three element set with*

$$\mathrm{supp}(\boldsymbol{x}_3) \cap (\mathrm{supp}(\boldsymbol{x}_1) \cup \mathrm{supp}(\boldsymbol{x}_2)) = \emptyset, \quad \mathrm{supp}(\boldsymbol{x}_1) \cap \mathrm{supp}(\boldsymbol{x}_2) \neq \emptyset.$$

Then $\boldsymbol{z}_1 = \boldsymbol{x}_1 - \boldsymbol{x}_3$ *is 1-norm reducible by* $\boldsymbol{z}_2 = \boldsymbol{x}_2 - \boldsymbol{x}_3$ *and vice versa. However* $\boldsymbol{z}_1, \boldsymbol{z}_2 \in \mathcal{B}_{\mathrm{MF}}$ *and if we throw away both* $\boldsymbol{z}_1$ *and* $\boldsymbol{z}_2$*, then* $\mathcal{F}_{\boldsymbol{t}}$ *is not connected. See the following Example 1. Also note that even if the set of 1-norm irreducible moves form a Markov basis, the Markov basis itself might not be 1-norm reducing.*

In Proposition 6 we considered uniqueness of minimal 1-norm reducing Markov bases. A parallel argument can be carried out concerning uniqueness of minimal strongly 1-norm reducing Markov bases in terms of strongly 1-norm irreducible moves. The following proposition can be proved in exactly the same way as Proposition 6 using Lemma 2.

Proposition 7 *Let* $\boldsymbol{z}$ *be a strongly 1-norm irreducible move. Then either* $\boldsymbol{z}$ *or* $-\boldsymbol{z}$ *belongs to every strongly 1-norm reducing Markov basis. There exists a unique minimal strongly 1-norm reducing Markov basis if and only if strongly 1-norm irreducible moves form a strongly 1-norm reducing Markov basis.*

Remark 2 *In many examples we expect the set of strongly 1-norm irreducible moves form a Markov basis. However as in the case of 1-norm reducing Markov bases, it seems difficult to state a simple sufficient condition to guarantee this. Suppose that for some* $\boldsymbol{t}$ *with* $|\boldsymbol{t}| = n$*, the fiber* $\mathcal{F}_{\boldsymbol{t}} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4\}$ *is a four element set with*

$$(\mathrm{supp}(\boldsymbol{x}_1) \cup \mathrm{supp}(\boldsymbol{x}_2)) \cap (\mathrm{supp}(\boldsymbol{x}_3) \cup \mathrm{supp}(\boldsymbol{x}_4)) = \emptyset,$$
$$\mathrm{supp}(\boldsymbol{x}_1) \cap \mathrm{supp}(\boldsymbol{x}_2) \neq \emptyset, \quad \mathrm{supp}(\boldsymbol{x}_3) \cap \mathrm{supp}(\boldsymbol{x}_4) \neq \emptyset.$$

Then $\boldsymbol{z}_{13} = \boldsymbol{x}_1 - \boldsymbol{x}_3$ *is strongly 1-norm reducible by the pair* $-\boldsymbol{z}_{14} = \boldsymbol{x}_4 - \boldsymbol{x}_1$*,* $\boldsymbol{z}_{23} = \boldsymbol{x}_2 - \boldsymbol{x}_3$*. Similarly* $\boldsymbol{x}_2 - \boldsymbol{x}_3$*,* $\boldsymbol{x}_1 - \boldsymbol{x}_4$*,* $\boldsymbol{x}_2 - \boldsymbol{x}_4$ *are not strongly 1-norm irreducible. However these four moves belong to* $\mathcal{B}_{\mathrm{MF}}$ *and if we throw away these four moves, then* $\mathcal{F}_{\boldsymbol{t}}$ *is not connected. See the following Example 1.*

In Remarks 1 and 2 we noted difficulty of establishing that the set of 1-norm irreducible moves or the set of strongly 1-norm irreducible moves form a Markov basis. It corresponds to the fact that there is no general implications between the non-replaceability in (2) and (strong) 1-norm irreducibility. The following examples illustrates this difficulty.

Example 1 *As examples of Remarks 1 and 2, we consider a hierarchical model for* $2 \times 2 \times 2 \times 2$ *contingency tables, where the generating set is* 12/13/23/34*. In this case, a frequency vector is*

$$\boldsymbol{x} = (x_{1111}, x_{1112}, x_{1121}, x_{1122}, \ldots, x_{2211}, x_{2212}, x_{2221}, x_{2222})'$$

and the sufficient statistic is

$$\boldsymbol{t} = (\{x_{ij\cdot\cdot}\}, \{x_{i\cdot k\cdot}\}, \{x_{\cdot jk\cdot}\}, \{x_{\cdot\cdot kl}\})'.$$

The corresponding matrix A is given as

$$
A = \begin{bmatrix}
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1
\end{bmatrix}.
$$

In Aoki and Takemura (2003b), we give a minimal Markov basis for this problem. We consider the case that the values of the sufficient statistic are given as follows.

$$
\begin{aligned}
& x_{ij\cdot\cdot} = x_{i\cdot k\cdot} = x_{\cdot jk\cdot} = 1,\ 1 \le i,j,k \le 2, \\
& x_{\cdot\cdot 11} = x_{\cdot\cdot 12} = 1, x_{\cdot\cdot 21} = 0, x_{\cdot\cdot 22} = 2.
\end{aligned}
$$

In this case, the fiber corresponding to the above $\boldsymbol{t}$ contains four elements,

$$
\begin{aligned}
& \boldsymbol{x}_1 = u_{1111}u_{1222}u_{2122}u_{2212},\ \boldsymbol{x}_2 = u_{1112}u_{1222}u_{2122}u_{2211}, \\
& \boldsymbol{x}_3 = u_{1122}u_{1211}u_{2112}u_{2222},\ \boldsymbol{x}_4 = u_{1122}u_{1212}u_{2111}u_{2222},
\end{aligned}
$$

where $\{u_{ijkl}\}$ is the set of indeterminates. This corresponds to the situation in Remark 2.

Modifying the above example, we can make an example for Remark 1. In the above example, we restrict that $\boldsymbol{i} = 1212$ is a structural zero cell ($x_{1212} \equiv 0$) and omit the cell. In this case, a frequency vector $\boldsymbol{x}$ is treated as a 15 dimensional vector and A is 16×15. The corresponding fiber contains three elements,

$$
\boldsymbol{x}_1 = u_{1111}u_{1222}u_{2122}u_{2212},\ \boldsymbol{x}_2 = u_{1112}u_{1222}u_{2122}u_{2211},\ \boldsymbol{x}_3 = u_{1122}u_{1211}u_{2112}u_{2222},
$$

which corresponds to the situation of Remark 1.

We now show that the set of 1-norm lower degree irreducible moves form a Markov basis. In the following let $\mathcal{B}_{\mathrm{LDI}}$ denote the set of 1-norm lower degree moves. In the case of contingency tables with fixed marginals, $\mathcal{B}_{\mathrm{LDI}}$ is invariant in the sense of Aoki and Takemura (2003b). We state the following proposition.

Proposition 8 *If a move $\boldsymbol{z}$ is non-replaceable by lower degree moves, then $\boldsymbol{z}$ is 1-norm lower degree irreducible. Hence $\mathcal{B}_{\mathrm{MF}} \subset \mathcal{B}_{\mathrm{LDI}}$ and $\mathcal{B}_{\mathrm{LDI}}$ is a Markov basis.*

Proof. If $\boldsymbol{z}$ is 1-norm reducible by a lower degree move $\boldsymbol{z}'$, then we can move from $\boldsymbol{z}^+$ to $\boldsymbol{z}^-$ either by two lower degree moves $\boldsymbol{z}'$ and $\boldsymbol{z}+\boldsymbol{z}'$ or by $\boldsymbol{z}'-\boldsymbol{z}$ and $\boldsymbol{z}'$. Therefore $\boldsymbol{z}^+ \sim \boldsymbol{z}^- \pmod{\mathcal{B}_{\deg \boldsymbol{z}-1}}$. This implies that if $\boldsymbol{z}$ is non-replaceable by lower degree moves, then $\boldsymbol{z}$ is 1-norm lower degree irreducible. Q.E.D.

Remark 3 *Together with indispensability and primitiveness, the result of Proposition 8 can be summarized as follows (cf. Proposition 5). For a move $\boldsymbol{z}$, the following implications hold*

$$\begin{aligned} \textit{indispensable} &\Rightarrow \textit{non-replaceable by lower degree moves} \\ &\Rightarrow \textit{1-norm lower degree irreducible} \\ &\Rightarrow \textit{primitive.} \end{aligned}$$

Finally we consider the case that the set of indispensable moves is 1-norm reducing.

Proposition 9 *Suppose that the set $\mathcal{B}_{\rm IDP}$ of indispensable moves is 1-norm reducing. Then $\mathcal{B}_{\rm IDP} = \mathcal{B}_{\rm LDI}$.*

Proof. Let $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^-$ is a move with $\boldsymbol{t} = A\boldsymbol{z}^+$. It suffice to prove that if $\boldsymbol{z}$ is not indispensable, then $\boldsymbol{z}$ is reducible by a lower degree move $\boldsymbol{z}'$. By assumption $\mathcal{B}_{\rm IDP}$ is 1-norm reducing. Therefore there exists an indispensable move $\boldsymbol{z}'$, such that $\boldsymbol{z}$ is 1-norm reducible by $\boldsymbol{z}'$. Note that $\boldsymbol{z}' \notin \mathcal{F}_{\boldsymbol{t}}$. By Lemma 2.1 of Takemura and Aoki (2004), it follows that $|\boldsymbol{z}'| < |\boldsymbol{z}|$. Therefore $\boldsymbol{z}$ is 1-norm reducible by a lower degree move $\boldsymbol{z}'$. Q.E.D.

Remark 4 *Many of the results of this section hold for any metric d on $\mathbb{N}^{|\mathcal{I}|}$ with the following property. If $\mathrm{supp}(\boldsymbol{x}) \cap \mathrm{supp}(\boldsymbol{y}) = \emptyset$, $\mathrm{supp}(\boldsymbol{x}') \subset \mathrm{supp}(\boldsymbol{x})$, $\mathrm{supp}(\boldsymbol{y}') \subset \mathrm{supp}(\boldsymbol{y})$, $\deg \boldsymbol{x}' < \deg \boldsymbol{x}$ and $\deg \boldsymbol{y}' < \deg \boldsymbol{y}$, then*

$$d(\boldsymbol{x}', \boldsymbol{y}') < d(\boldsymbol{x}, \boldsymbol{y}).$$

4 Examples and some discussion

4.1 Examples

In this section, we consider some standard models of contingency tables. In Section 3 of Takemura and Aoki (2004), we have considered minimal Markov bases and their uniqueness for these models. In this paper, we investigate minimal 1-norm reducing Markov bases and their uniqueness. To display each move $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^-$, we write a binomial $\prod_{\boldsymbol{i}} u_{\boldsymbol{i}}^{z_{\boldsymbol{i}}^+} - \prod_{\boldsymbol{i}} u_{\boldsymbol{i}}^{z_{\boldsymbol{i}}^-}$ in this section, where $\{u_{\boldsymbol{i}}\}_{\boldsymbol{i}\in\mathcal{I}}$ is the set of indeterminates. For some cases, we use 4ti2 (Hemmecke and Hemmecke, 2003) to compute the reduced Gröbner basis and the Graver basis.

One-way contingency tables. First we consider the simplest case of one-way contingency tables. Let $\boldsymbol{x} = \{x_i\}$ be an I dimensional frequency vector and $A = (1, \ldots, 1)$. The sufficient statistic $\boldsymbol{t}$ is the sample size n. This corresponds to testing the homogeneity of mean parameters for I independent Poisson variables conditional on the total sample size n. In this case, a minimum fiber Markov basis is the set of degree 1 moves, i.e.,

$$\mathcal{B}_{\rm MF} = \{\boldsymbol{z} \mid \boldsymbol{z} = u_i - u_j,\ i \neq j\}.$$

As shown in Section 3 of Takemura and Aoki (2004), each element of $\mathcal{B}_{\rm MF}$ is dispensable. However, it is obvious that these degree 1 moves are 1-norm irreducible. Furthermore, the Graver basis in this case obviously coincides with $\mathcal{B}_{\rm MF}$. Therefore $\mathcal{B}_{\rm MF}$ is the unique minimal strongly 1-norm reducing Markov basis.

Now we consider the reduced Gröbner basis with respect to purely lexicographic or degree lexicographic order. Let $u_1 \succ \cdots \succ u_I$. Then the reduced Gröbner basis is easily shown to be

$$\mathcal{B}_{\prec} = \{\boldsymbol{z} \mid \boldsymbol{z} = u_i - u_I,\ 1 \le i \le I-1\}.$$

We see that in order to move from one frequency vector to another, the frequencies have to pass through the last cell I. For example to move from $(n, 0, \dots, 0)$ to $(0, n, 0, \dots, 0)$ we need $2n$ steps of $\mathcal{B}_{\prec}$ and $\mathcal{B}_{\prec}$ is not 1-norm reducing. In the case of $\mathcal{B}_{\rm MF}$, n steps of $u_1 - u_2 \in \mathcal{B}_{\rm MF}$ are sufficient for moving from $(n, 0, \dots, 0)$ to $(0, n, 0, \dots, 0)$. We see that the diameter of $G_{n,\mathcal{B}_{\prec}}$ is $2n$, whereas the diameter of $G_{n,\mathcal{B}_{\rm MF}}$ is n.

Two-way contingency tables. Next we consider an $I \times J$ two-way contingency tables with fixed row and column sums. Let $\boldsymbol{x} = \{x_{ij}\}$ be an IJ-dimensional column vector of cell frequencies and

$$A = \left[\begin{array}{c} 1_I' \otimes E_J \\ E_I \otimes 1_J' \end{array} \right].$$

This is a standard example of testing the hypothesis that the rows and the columns are independent. In this case, it is known that the set of degree 2 moves,

$$\mathcal{B} = \{u_{ij}u_{i'j'} - u_{ij'}u_{i'j},\ i \ne i',\ j \ne j'\},$$

is the unique minimal Markov basis (see Section 3 of Takemura and Aoki, 2004). Here we have the following proposition.

Proposition 10 *$\mathcal{B}$ is the unique minimal strongly 1-norm reducing Markov basis.*

Proof. In this case, it is well known that the Graver basis consists of all the moves which are written as

$$\boldsymbol{z} = u_{i_1j_1}u_{i_2j_2}\cdots u_{i_sj_s} - u_{i_2j_1}u_{i_3j_2}\cdots u_{i_1j_s}, \qquad (8)$$

where $(i_1, j_1), (j_1, i_2), \dots, (i_s, j_s), (j_s, i_1)$ is a circuit in the complete bipartite graph $K_{I,J}$. Since $z^+_{i_1j_1}, z^+_{i_2j_2} > 0$ and $z^-_{i_2j_1}, z^-_{i_3j_2} > 0$, $\boldsymbol{z}$ is strongly 1-norm reducible by a pair of moves $\boldsymbol{z}', \boldsymbol{z}'' \in \mathcal{B}$ where

$$\begin{aligned} \boldsymbol{z}' &= u_{i_1j_1}u_{i_2j_2} - u_{i_2j_1}u_{i_1j_2}, \\ \boldsymbol{z}'' &= u_{i_3j_1}u_{i_2j_2} - u_{i_2j_1}u_{i_3j_2}. \end{aligned}$$

Therefore, for any element $\boldsymbol{z}$ of the Graver basis, $\mathcal{B}$ is strongly 1-norm reducing for $\boldsymbol{z}^+, \boldsymbol{z}^-$. Moreover, since each element of $\mathcal{B}$ is indispensable, it is also 1-norm irreducible from Proposition 5. Uniqueness also follows from Proposition 8. Q.E.D.

In Aoki and Takemura (2004) we obtained the unique minimal Markov basis for two-way contingency tables with arbitrary patterns of structural zeros. Since the Graver basis for this case consists of moves of the form (8), which does not involve structural zeros, it can be easily proved that the unique minimal Markov basis is at the same time the unique minimal strongly 1-norm reducing Markov basis.

Three-way contingency tables with fixed one-dimensional marginals. Next we consider three-way contingency tables with fixed one-dimensional marginals. In this case, $\boldsymbol{x} = \{x_{ijk}\}$ is an IJK-dimensional frequency vector and

$$A = \begin{bmatrix} 1'_I \otimes 1'_J \otimes E_K \\ 1'_I \otimes E_J \otimes 1'_K \\ E_I \otimes 1'_J \otimes 1'_K \end{bmatrix}.$$

This corresponds to testing the hypothesis that the three factors are completely independent, i.e., $p_{ijk} = \alpha_i \beta_j \gamma_k$. As Takemura and Aoki (2004) have shown, a minimal Markov basis for this case is not unique. The minimum fiber Markov basis can be written as follows.

$$\begin{aligned}
\mathcal{B}_{\rm MF} &= \mathcal{B}_{\rm IDP} \cup \mathcal{B}^*,\\
\mathcal{B}_{\rm IDP} &= \{u_{ij_1k_1}u_{ij_2k_2} - u_{ij_1k_2}u_{ij_2k_1},\ j_1 \neq j_2,\ k_1 \neq k_2\}\\
&\quad \cup \{u_{i_1jk_1}u_{i_2jk_2} - u_{i_1jk_2}u_{i_2jk_1},\ i_1 \neq i_2,\ k_1 \neq k_2\}\\
&\quad \cup \{u_{i_1j_1k}u_{i_2j_2k} - u_{i_1j_2k}u_{i_2j_1k},\ i_1 \neq i_2,\ j_1 \neq j_2\},\\
\mathcal{B}^* &= \{u_{i_1j_1k_1}u_{i_2j_2k_2} - u_{i_1j_1k_2}u_{i_2j_2k_1},\ u_{i_1j_1k_1}u_{i_2j_2k_2} - u_{i_1j_2k_1}u_{i_2j_1k_2},\\
&\qquad u_{i_1j_1k_1}u_{i_2j_2k_2} - u_{i_1j_2k_2}u_{i_2j_1k_1},\ u_{i_1j_1k_2}u_{i_2j_2k_1} - u_{i_1j_2k_1}u_{i_2j_1k_2},\\
&\qquad u_{i_1j_1k_2}u_{i_2j_2k_1} - u_{i_1j_2k_2}u_{i_2j_1k_1},\ u_{i_1j_2k_1}u_{i_2j_1k_2} - u_{i_1j_2k_2}u_{i_2j_1k_1},\\
&\qquad\qquad i_1 \neq i_2, j_1 \neq j_2, k_1 \neq k_2\}.
\end{aligned}$$

Here, $\mathcal{B}_{\rm IDP}$ is the set of indispensable moves. $\mathcal{B}^*$ is the set of all degree 2 moves which connect all the elements of the four-elements fiber

$$\begin{aligned}
\mathcal{F}_{i_1i_2j_1j_2k_1k_2} &= \{\boldsymbol{x} = \{x_{ijk}\} \mid x_{i_1\cdot\cdot} = x_{i_2\cdot\cdot} = x_{\cdot j_1 \cdot} = x_{\cdot j_2 \cdot} = x_{\cdot\cdot k_1} = x_{\cdot\cdot k_2} = 1\}\\
&= \{u_{i_1j_1k_1}u_{i_2j_2k_2},\ u_{i_1j_1k_2}u_{i_2j_2k_1},\ u_{i_1j_2k_1}u_{i_2j_1k_2},\ u_{i_1j_2k_2}u_{i_2j_1k_1}\}.
\end{aligned}$$

The minimal Markov basis in this case consists of $\mathcal{B}_{\rm IDP}$ and three moves for each $(i_1, i_2, j_1, j_2, k_1, k_2)$, which connects four elements of $\mathcal{F}_{i_1i_2j_1j_2k_1k_2}$ into a tree.

Now, we show that, for the $2 \times 2 \times 2$ case, (a) each minimal Markov basis is not 1-norm reducing and (b) the minimum fiber Markov basis is at the same time the unique minimal strongly 1-norm reducing Markov basis. By 4ti2, the Graver basis for the $2 \times 2 \times 2$ case is calculated as

$$\mathcal{B}_{\rm Graver} = \mathcal{B}_{\rm MF}\ \cup\ \{u_{ijk}^2 u_{i'j'k'} - u_{ijk'}u_{ij'k}u_{i'jk},\ i \neq i', j \neq j'\ k \neq k'\}.$$

It is seen that each dispensable move in $\mathcal{B}_{\rm MF}$ is 1-norm irreducible, and therefore each minimal Markov basis is not 1-norm reducing. On the other hand, a degree 3 move $u_{ijk}^2 u_{i'j'k'} - u_{ijk'}u_{ij'k}u_{i'jk}$ is strongly 1-norm reducible by a pair of moves, $u_{ijk}u_{i'j'k'} - u_{ij'k'}u_{i'jk}$ and $u_{ijk}u_{ij'k'} - u_{ijk'}u_{ij'k}$ for example, which are in $\mathcal{B}_{\rm MF}$.

Again by 4ti2, the reduced Gröbner basis for the $2\times2\times2$ case under the purely lexicographic or degree lexicographic term order

$$u_{111} \succ u_{112} \succ \cdots \succ u_{222}$$

is found to be

$$\mathcal{B}_{\prec} = \mathcal{B}_{\rm IDP} \cup \{u_{121}u_{212} - u_{111}u_{222},\ u_{112}u_{221} - u_{111}u_{222},\ u_{122}u_{211} - u_{111}u_{222}\}.$$

This is a minimal Markov basis. Note that the four elements of $\mathcal{F}_{121212}$ are connected via $u_{111}u_{222}$. This is similar to the case of one-way contingency tables and the diameter of $G_{\boldsymbol{t},\mathcal{B}_{\prec}}$ is generally larger than the diameter $G_{\boldsymbol{t},\mathcal{B}_{\mathrm{MF}}}$.

Note that two-way contingency tables and three-way contingency tables with fixed on-dimensional marginals are special cases of decomposable models considered in Dobra (2003). It is of interest to investigate the Markov basis given by Dobra (2003) in the framework of this paper.

Three-way contingency tables with fixed two-dimensional marginals. Next we consider three-way contingency tables with fixed two-dimensional marginals. In this case, $\boldsymbol{x} = \{x_{ijk}\}$ is an IJK-dimensional frequency vector and

$$A = \begin{bmatrix} 1_I' \otimes E_J \otimes E_K \\ E_I \otimes 1_J' \otimes E_K \\ E_I \otimes E_J \otimes 1_K' \end{bmatrix}.$$

This corresponds to testing no three-factor interactions of the log-linear model. As is stated in many works, it is surprisingly difficult to obtain a Markov basis for this problem, except for small I, J, K.

For the case of $2 \times J \times K$ tables, Diaconis and Sturmfels (1998) have shown that the set of degree $4, 6, \ldots, \min(J, K)$ moves, where the degree $2s$ move is written as

$$u_{1j_1k_1}u_{1j_2k_2}\cdots u_{1j_sk_s}u_{2j_2k_1}u_{2j_3k_2}\cdots u_{2j_1k_s} - u_{1j_2k_1}u_{1j_3k_2}\cdots u_{1j_1k_s}u_{2j_1k_1}u_{2j_2k_2}\cdots u_{2j_sk_s},$$

constitutes a Markov basis. Takemura and Aoki (2004) have shown that this is the unique minimal Markov basis for this problem. Our argument here is that this is also the unique minimal strongly 1-norm reducing Markov basis. This is obvious from the fact that the above unique minimal Markov basis is also the Graver basis. See Corollary 14.12 of Sturmfels (1995).

For the next simpler case, we consider $3 \times 3 \times 3$ tables. Aoki and Takemura (2003a) have shown that the unique minimal Markov basis for this problem consists of two types of moves,

$$u_{111}u_{122}u_{212}u_{221} - u_{112}u_{121}u_{211}u_{222}$$

and

$$u_{111}u_{123}u_{132}u_{212}u_{221}u_{233} - u_{112}u_{121}u_{133}u_{211}u_{223}u_{232}.$$

We show that the above basis $\mathcal{B}$ is also the unique minimal strongly 1-norm reducing Markov basis. To show this, we have to check that, for each element $\boldsymbol{z}$ of the Graver basis, $\mathcal{B}$ is strongly 1-norm reducing for $\boldsymbol{z}^+, \boldsymbol{z}^-$. From Theorem 14.13 of Sturmfels (1995), augmenting the above two types by the following five types of moves gives the Graver basis.

$$\begin{gathered} u_{111}u_{123}u_{132}u_{222}u_{231}u_{313}u_{321} - u_{113}u_{122}u_{131}u_{221}u_{232}u_{311}u_{323}, \\ u_{112}u_{121}u_{133}u_{222}u_{231}u_{311}u_{323}u_{332}^2 - u_{111}u_{123}u_{132}u_{221}u_{232}u_{312}u_{322}u_{331}u_{333}, \\ u_{112}u_{123}u_{131}u_{213}u_{221}u_{232}u_{311}u_{322}u_{333} - u_{113}u_{121}u_{132}u_{211}u_{222}u_{233}u_{312}u_{323}u_{331}, \\ u_{111}u_{123}u_{131}u_{213}^2u_{221}u_{231}u_{311}u_{322}u_{333} - u_{113}u_{122}u_{132}u_{211}^2u_{223}u_{233}u_{313}u_{321}u_{332}, \\ u_{111}u_{123}^2u_{132}^2u_{213}u_{222}^2u_{231}u_{312}u_{321}u_{333} - u_{113}u_{122}^2u_{131}u_{133}u_{212}u_{221}u_{223}u_{232}u_{311}u_{323}u_{332}. \end{gathered}$$

For each move $\boldsymbol{z}$ of the Graver basis, it is easy to see that $\mathcal{B}$ is strongly 1-norm reducing for $\boldsymbol{z}^+, \boldsymbol{z}^-$, i.e., we can apply elements of $\mathcal{B}$ to $\boldsymbol{z}^+$ and $\boldsymbol{z}^-$ and decrease $|\boldsymbol{z}^+ - \boldsymbol{z}^-|$.

Poisson regression. Here we consider a simple example of Poisson regression discussed in Diaconis, Eisenbud and Sturmfels (1998). Let $\boldsymbol{x} = (x_0, x_1, \dots, x_4)'$ and

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \end{bmatrix}.$$

For this problem, Takemura and Aoki (2004) have shown that a minimal Markov basis is not unique. The minimum fiber Markov basis is given as follows.

$$\begin{aligned} &\mathcal{B}_{\mathrm{MF}} = \mathcal{B}_{\mathrm{IDP}} \cup \mathcal{B}^*, \\ &\mathcal{B}_{\mathrm{IDP}} = \{u_0u_3 - u_1u_2,\ u_0u_2 - u_1^2,\ u_1u_4 - u_2u_3,\ u_2u_4 - u_3^2\}, \\ &\mathcal{B}^* = \{u_0u_4 - u_1u_3,\ u_0u_4 - u_2^2,\ u_1u_3 - u_2^2\}. \end{aligned}$$

Here, $\mathcal{B}^*$ is the set of all degree 2 moves which connect all the elements of the three-elements fiber

$$\mathcal{F} = \{u_0u_4,\ u_1u_3,\ u_2^2\}.$$

Therefore there are three minimal Markov bases for this case since any one move in $\mathcal{B}^*$ is not needed to construct a Markov basis.

Similarly to the $2 \times 2 \times 2$ tables with fixed one-dimensional marginals discussed above, $\mathcal{B}_{\mathrm{MF}}$ is the unique minimal strongly 1-norm reducing Markov basis, and each minimal Markov basis is not 1-norm reducing. The later part is obvious since each element of $\mathcal{B}^*$ is 1-norm irreducible. To show that $\mathcal{B}_{\mathrm{MF}}$ is strongly 1-norm reducing, we compute the Graver basis for this problem by 4ti2. It is given as follows.

$$\begin{aligned} \mathcal{B}_{\mathrm{Graver}} = \mathcal{B}_{\mathrm{MF}} \ \cup\ \{&u_0^3u_4 - u_1^4,\ u_0^2u_3 - u_1^3,\ u_0^2u_4 - u_1^2u_2,\ u_0u_3^2 - u_1^2u_4,\ u_2^3 - u_1^2u_4, \\ &u_0u_3^2 - u_2^3,\ u_0u_4^2 - u_2u_3^2,\ u_3^3 - u_1u_4^2,\ u_0u_4^3 - u_3^4\}. \end{aligned}$$

It is seen that each move above is strongly 1-norm reducible by a pair of moves in $\mathcal{B}_{\mathrm{MF}}$.

Hardy-Weinberg model. Finally we consider the Hardy-Weinberg model for 4 alleles, i.e.,

$$\boldsymbol{x} = (x_{11}, x_{12}, x_{13}, x_{14}, x_{22}, x_{23}, x_{24}, x_{33}, x_{34}, x_{44})'$$

and

$$A = \begin{bmatrix} 2 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 2 \end{bmatrix}.$$

As is stated in Takemura and Aoki (2004), a minimal Markov basis for this case is not unique, and the minimum fiber Markov basis is given as follows.

$$\begin{aligned} &\mathcal{B}_{\mathrm{MF}} = \mathcal{B}_{\mathrm{IDP}} \cup \mathcal{B}^*, \\ &\mathcal{B}_{\mathrm{IDP}} = \{u_{i_1i_1}u_{i_2i_3} - u_{i_1i_2}u_{i_1i_3},\ u_{i_1i_1}u_{i_2i_2} - u_{i_1i_2}^2\}, \\ &\mathcal{B}^* = \{u_{12}u_{34} - u_{13}u_{24},\ u_{12}u_{34} - u_{14}u_{23},\ u_{13}u_{24} - u_{14}u_{23}\}, \end{aligned}$$

where i_1, i_2, i_3 are all distinct, and $u_{ij} = u_{ji}$ for $i > j$. Here, $\mathcal{B}^*$ is the set of all degree 2 moves which connect all the elements of the three-elements fiber

$$\mathcal{F} = \{u_{12}u_{34},\ u_{13}u_{24},\ u_{14}u_{23}\}.$$

There are three minimal Markov bases for this case since any one move in $\mathcal{B}^*$ is not needed to construct a Markov basis. In this case, again, the minimum fiber Markov basis is the unique minimal strongly 1-norm reducing Markov basis, and each minimal Markov basis is not 1-norm reducing. To show this, we give the Graver basis for this problem, which is computed by 4ti2. In this case, augmenting $\mathcal{B}_{\mathrm{MF}}$ by the following six types of moves gives the Graver basis.

$$\begin{gathered}
12 \text{ relations of } u_{12}^2 u_{33} - u_{13}^2 u_{22},\\
12 \text{ relations of } u_{12}^2 u_{34} - u_{13} u_{14} u_{22},\\
12 \text{ relations of } u_{12} u_{13} u_{24} - u_{11} u_{22} u_{34},\\
4 \text{ relations of } u_{11} u_{22} u_{33} - u_{12} u_{13} u_{24},\\
6 \text{ relations of } u_{11} u_{23} u_{24} - u_{22} u_{13} u_{14},\\
3 \text{ relations of } u_{12} u_{13} u_{24} u_{34} - u_{11} u_{22} u_{33} u_{44},\\
12 \text{ relations of } u_{12}^2 u_{34}^2 - u_{14}^2 u_{22} u_{33},\\
12 \text{ relations of } u_{12}^2 u_{33} u_{44} - u_{11} u_{23} u_{24} u_{34}.
\end{gathered}$$

It is easily seen that each move above is strongly 1-norm reducible by a pair of moves in $\mathcal{B}_{\mathrm{MF}}$.

4.2 Some discussions

In this paper some results on 1-norm reducing Markov basis are obtained. Actually there remain more unsolved questions than answers in the framework of this paper. For example one can ask under what conditions the minimum fiber Markov basis $\mathcal{B}_{\mathrm{MF}}$ is 1-norm reducing and similarly under what conditions $\mathcal{B}_{\mathrm{LDI}}$ is 1-norm reducing. One can also ask when $\mathcal{B}_{\mathrm{LDI}}$ coincides with $\mathcal{B}_{\mathrm{MF}}$ or the universal Gröbner basis.

Another set of questions can be asked on reducing the distance in more than one steps. We may call a Markov basis $\mathcal{B}$ *1-norm reducing in k-steps* if for every pair of states in the same fiber $\mathcal{F}_{\boldsymbol{t}}$, we can reduce the 1-norm by at most k moves from $\mathcal{B}$. Since the Graver basis is finite and for 1-norm reduction it suffices to move from the positive part to the negative part of each primitive move, every Markov basis $\mathcal{B}$ is 1-norm reducing in k-steps for some finite k. Then the natural question to ask is what is the minimum k such that $\mathcal{B}$ is 1-norm reducing in k-steps.

According to Proposition 3, we have to consider all the primitive moves to check whether a given Markov basis is 1-norm reducing or not. However it is generally difficult to compute the Graver basis even for a problem of moderate size. For example, in the case of $I \times J \times K$ three-way contingency tables with fixed one-dimensional marginals, we have shown in Takemura and Aoki (2004) that minimal Markov bases and the minimum fiber Markov basis consist of two types of $2 \times 2 \times 2$ degree 2 moves only. In Section 4.1, for the simplest $2 \times 2 \times 2$ case we have checked that the minimum fiber Markov basis is the unique minimal strongly 1-norm reducing Markov basis, but we do not know whether the same result holds for the general $I \times J \times K$ case, since the general form of the Graves bases is not known at present. In fact, we have found by 4ti2 that the Graver basis for the $2 \times 2 \times 3$ problem contains primitive moves of degree 4 such as

$$u_{111} u_{122}^2 u_{213} - u_{221} u_{112}^2 u_{123}.$$

Similarly, the Graver basis for the Hardy-Weinberg model of 4 alleles problem suggests a complicated structure of Graver basis for the general r alleles problem, while the minimum fiber Markov basis is the same as the 4 alleles case as shown in Takemura and Aoki (2004).

References

[1] Aoki, S. and Takemura, A. (2003a). Minimal basis for connected Markov chain over $3\times 3\times K$ contingency tables with fixed two-dimensional marginals. *Australian and New Zealand Journal of Statistics.* **45**, 229–249.

[2] Aoki, S. and Takemura, A. (2003b). Invariant minimal Markov basis for sampling contingency tables with fixed marginals. Technical Report METR 03-25, Department of Mathematical Engineering and Information Physics, The University of Tokyo. Submitted for publication.

[3] Aoki, S. and Takemura, A. (2004). Markov chain Monte Carlo exact tests for incomplete two-way contingency tables. To appear in *Journal of Statistical Computation and Simulation.*

[4] Bubley, R. and Dyer, M. (1997). Path coupling: a technique for proving rapid mixing in Markov chains. *Proceedings of the 38th Annual Symposium on Foundations of Computer Science*, IEEE, Miami Beach, Florida, 223–231.

[5] Diaconis, P., Eisenbud, D. and Sturmfels, B. (1998). Lattice walks and primary decomposition. in *Mathematical Essays in Honor of Gian-Carlo Rota.* Sagan, B. E. and Stanley, R. P., editors, pp. 173–193, Birkhauser, Boston.

[6] Diaconis, P. and Saloff-Coste, L. (1998). What do we know about the Metropolis algorithm? *Journal of Computer and System Sciences*, **57**, 20–36.

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

[8] Dinwoodie, J. H. (1998). The Diaconis-Sturmfels algorithm and rules of succession. *Bernoulli*, **4**, 401–410.

[9] Dobra, A. (2003). Markov bases for decomposable graphical models. *Bernoulli*, **9**, 1093–1108.

[10] Hemmecke, R. and Hemmecke, R. (2003). 4ti2 Version 1.1 — Computation of Hilbert bases, Graver bases, toric Gröbner bases, and more, Available at www.4ti2.de.

[11] Hibi, T. (2003). *Gröbner Bases.* Asakura Shoten, Tokyo. (in Japanese).

[12] Ohsugi, H. and Hibi, T. (2003). Indispensable binomials of toric ideals, submitted for publication.

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

[14] Takemura, A. and Aoki, S. (2004). Some characterizations of minimal Markov basis for sampling from discrete conditional distributions, *Annals of the Institute of Statistical Mathematics*, **56**, 1–17.