JOURNAL OF ALGEBRAIC STATISTICS
Vol. 6, No. 2, 2015, 80-107
ISSN 1309-3452 – www.jalgstat.com

Markov degree of configurations defined by fibers of a configuration

Takayuki Koyama[1], Mitsunori Ogawa[2,*], Akimichi Takemura [1]

[1] *Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, Tokyo, Japan*
[2] *Department of Business Administration, Graduate School of Social Sciences, Tokyo Metropolitan University, Tokyo, Japan*

Abstract. We consider a series of configurations defined by fibers of a given base configuration. We prove that Markov degree of the configurations is bounded from above by the Markov complexity of the base configuration. As important examples of base configurations we consider incidence matrices of graphs and study the maximum Markov degree of configurations defined by fibers of the incidence matrices. In particular we give a proof that the Markov degree for two-way transportation polytopes is three.

Key Words and Phrases: Algebraic Statistics, Markov basis, Transportation polytopes

---

1. Introduction

The study of Markov bases has been developing rapidly since the seminal paper of Diaconis and Sturmfels [6], which established the equivalence of a Markov basis for a discrete exponential model in statistics and a generating set of a corresponding toric ideal. See [2, 8, 12] for terminology of algebraic statistics and toric ideals used in this paper.

When we study Markov bases for a specific problem, usually we are not faced with a single configuration, but rather with a series of configurations, possibly parameterized by a few parameters. For example, Markov bases associated with complete bipartite graphs $K_{I,J}$ (in statistical terms, independence model of $I \times J$ two-way contingency tables) are parameterized by I and J. In this case, Markov bases consist of moves of degree two irrespective of I and J. In more general cases, some measure of complexity of Markov bases grows with the parameter and we are interested in bounding the growth.

There are some typical procedures to generate a series of configurations based on a given set of configurations. Perhaps the most important construction is the higher Lawrence lifting of a configuration, for which Santos and Sturmfels [16] described the growth by the notion of Graver

---

*Corresponding author.

Email addresses: `takayuki_koyama@mist.i.u-tokyo.ac.jp` (T. Koyama), `mitsunori_ogawa@tmu.ac.jp` (M. Ogawa), `takemura@stat.t.u-tokyo.ac.jp` (A. Takemura)

complexity. Another important construction is the nested configuration ([15]), where generated series of configurations basically inherit nice properties of original configurations. In this paper we define a new procedure to generate a series of configurations using fibers of a given configuration, which we call the *base* configuration. This construction is closely related to the higher Lawrence lifting of the base configuration and using this fact we prove that *Markov degree* of the configurations is bounded from above by the *Markov complexity* of the base configuration.

There are some nice problems, such as the independence model of two-way contingency tables (corresponding to complete bipartite graphs), where the moves of degree two forms a Markov basis. When a minimal Markov basis contains a move of degree three or higher, it is usually very hard to control measures of complexity of Markov bases. A notable exception is the conjecture by [5] that the Markov degree associated with the Birkhoff polytope is three, i.e., the toric ideal associated with the Birkhoff polytope is generated by binomials of degree at most three. This conjecture was proved in [19]. In view of [9] and [19], Christian Haase (personal communication, 2013) suggested that the Markov degree associated with two-way transportation polytopes and flow polytopes is three. Very recently Domokos and Joó [7] gave a proof of this general conjecture. Adapting the arguments in [19], we give a proof that the Markov degree associated with two-way transportation polytopes is three in Section 4.1. Two-way transportation polytopes are important examples in our framework, since they are fibers of the incidence matrix of a complete bipartite graph.

The organization of this paper is as follows. In Section 2 we set up the framework of this paper and prove the main theorem that the Markov degree of the configurations defined by fibers of a base configuration is bounded from above by the Markov complexity of the base configuration. In the remaining sections of this paper we investigate the maximum Markov degree and the Markov complexity of some important base configurations. In Section 3 we study incidence matrices of complete graphs and in Section 4 we study those of complete bipartite graphs as base configurations. We end the paper with some discussions in Section 5.

2. Main result

Let A be a $d \times n$ configuration matrix. Elements of the integer kernel $\ker_{\mathbb{Z}} A$ of A are called *moves* for A. As in Section 1.5.1 of [12] we assume that there exists a d-dimensional row vector $\boldsymbol{v}$ such that $\boldsymbol{v}A = (1, 1, \dots, 1)$. Let $\mathbb{N} = \{0, 1, 2, \dots\}$ denote the set of non-negative integers and let $\mathbb{N}A = \{A\boldsymbol{x} \mid \boldsymbol{x} \in \mathbb{N}^n\}$. For $\boldsymbol{b} \in \mathbb{N}A$

$$\mathcal{F}_{A,\boldsymbol{b}} = \{\boldsymbol{x} \in \mathbb{N}^n \mid A\boldsymbol{x} = \boldsymbol{b}\}$$

is the $\boldsymbol{b}$*-fiber* of A. Each fiber $\mathcal{F}_{A,\boldsymbol{b}}$ is a finite set and non-empty for $\boldsymbol{b} \in \mathbb{N}A$. We denote the size of $\mathcal{F}_{A,\boldsymbol{b}}$ by $\nu(\boldsymbol{b}) = |\mathcal{F}_{A,\boldsymbol{b}}|$. Hence with an appropriate order the elements of $\mathcal{F}_{A,\boldsymbol{b}}$ are enumerated as

$$\mathcal{F}_{A,\boldsymbol{b}} = \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_{\nu(\boldsymbol{b})}\}.$$

We look at $\boldsymbol{x}_i$, $i = 1, \dots, \nu(\boldsymbol{b})$, as n-dimensional column vectors and define an $n \times \nu(\boldsymbol{b})$ matrix as

$$A_{\boldsymbol{b}} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_{\nu(\boldsymbol{b})}).$$

Note that $\boldsymbol{x} \in \mathcal{F}_{A,\boldsymbol{b}}$ implies

$$\boldsymbol{v}\boldsymbol{b} = \boldsymbol{v}A\boldsymbol{x} = |\boldsymbol{x}| = x_1 + \cdots + x_n > 0$$

if $\boldsymbol{x} \neq 0$. Hence for $\tilde{\boldsymbol{v}} = \boldsymbol{v}A/(\boldsymbol{v}\boldsymbol{b})$

$$\tilde{\boldsymbol{v}}A_{\boldsymbol{b}} = (1, 1, \ldots, 1), \tag{1}$$

and $A_{\boldsymbol{b}}$ is a configuration.

Remark 1. *The configuration $A_{\boldsymbol{b}}$ corresponds to the generalized hypergeometric distribution or the "A-hypergeometric distribution" (see [18], [14]) over the fiber $\mathcal{F}_{A,\boldsymbol{b}}$. Hence our construction is natural also from the viewpoint of statistics.*

Consider the set of moves for $A_{\boldsymbol{b}}$ of degree at most m. The *Markov degree* $\mathrm{MD}(A_{\boldsymbol{b}})$ of $A_{\boldsymbol{b}}$ is the minimum value of m such that the moves of degree at most m form a Markov basis (cf. [19, 11]). We are interested in the maximum of $\mathrm{MD}(A_{\boldsymbol{b}})$ when $\boldsymbol{b}$ ranges over $\mathbb{N}A$:

$$\max_{\boldsymbol{b}\in\mathbb{N}A} \mathrm{MD}(A_{\boldsymbol{b}}).$$

Let $A^{(N)}$ denote the N-th Lawrence lifting of A (cf. [16]). The moves for $A^{(N)}$ are written as $z = (z_1, \ldots z_N)$, such that $\sum_{k=1}^{N} z_k = 0$ and $z_k \in \ker_{\mathbb{Z}} A$, $k = 1, \ldots, N$. In this paper, we call z_k the k-th *layer* or *slice* of z. The *type* of z is the number of non-zero layers among $z_1, \ldots, z_N$:

$$\mathrm{type}(z) = |\{k \mid z_k \neq 0\}|.$$

Let $\mathcal{G}\left(A^{(N)}\right)$ denote the Graver basis of $A^{(N)}$. Then the *Graver complexity* of A is defined (cf. [16, 4, 13]) as

$$\mathrm{GC}(A) = \sup\left(\{0\} \cup \left\{ \mathrm{type}(x) \;\middle|\; x \in \bigcup_{N\geq 1} \mathcal{G}\left(A^{(N)}\right) \right\}\right),$$

where $\{0\}$ is needed for the case that the columns of A are linearly independent. Santos and Sturmfels [16] gave an explicit expression for the Graver complexity, which we will use for computing the Graver complexity of some configurations. The *Markov complexity* $\mathrm{MC}(A)$ of A is defined as the minimum value of m such that the moves of type at most m form a Markov basis for every $A^{(N)}$. Note that $\mathrm{MC}(A) \leq \mathrm{GC}(A)$ since a minimal Markov basis is contained in the Graver basis.

Now we are ready to state our main theorem.

Theorem 1. *The Markov degree of $A_{\boldsymbol{b}}$ is bounded from above by the Markov complexity of A:*

$$\max_{\boldsymbol{b}\in\mathbb{N}A} \mathrm{MD}(A_{\boldsymbol{b}}) \leq \mathrm{MC}(A). \tag{2}$$

Before giving a proof, we discuss how a fiber of $A_{\boldsymbol{b}}$ is embedded in a fiber of some $A^{(N)}$. For $\boldsymbol{c} \in \mathbb{N}A_{\boldsymbol{b}}$ consider an element $\boldsymbol{y} = (y_1, \ldots, y_{\nu(\boldsymbol{b})})$ of $\mathcal{F}_{A_{\boldsymbol{b}},\boldsymbol{c}}$. By

$$\boldsymbol{c} = A_{\boldsymbol{b}}\boldsymbol{y} = \boldsymbol{x}_1 y_1 + \cdots + \boldsymbol{x}_{\nu(\boldsymbol{b})} y_{\nu(\boldsymbol{b})}$$

and by (1), we see that $|\boldsymbol{y}| = y_1 + \cdots + y_{\nu(\boldsymbol{b})} = \tilde{\boldsymbol{\nu}}\boldsymbol{c}$ is common for all $\boldsymbol{y} \in \mathcal{F}_{A_{\boldsymbol{b}},\boldsymbol{c}}$. Let $N = |\boldsymbol{y}|, \boldsymbol{y} \in \mathcal{F}_{A_{\boldsymbol{b}},\boldsymbol{c}}$. Then $\boldsymbol{y} \in \mathcal{F}_{A_{\boldsymbol{b}},\boldsymbol{c}}$ is identified with a multiset $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_N\}$ of elements (columns) of $A_{\boldsymbol{b}}$, where $\boldsymbol{x}_i$ is repeated y_i times, e.g.:

$$\boldsymbol{x}_1 = \boldsymbol{w}_1 = \cdots = \boldsymbol{w}_{y_1}, \ \ \boldsymbol{x}_2 = \boldsymbol{w}_{y_1+1} = \cdots = \boldsymbol{w}_{y_1+y_2}, \ \dots$$

In this notation

$$\boldsymbol{w}_k \in \mathcal{F}_{A,\boldsymbol{b}}, \ k = 1, \dots, N, \ \text{ and } \ \boldsymbol{c} = \boldsymbol{w}_1 + \cdots + \boldsymbol{w}_N. \tag{3}$$

Define a $(dN + n)$-dimensional integer vector $(\boldsymbol{b}^{(N)}, \boldsymbol{c})$ as

$$(\boldsymbol{b}^{(N)}, \boldsymbol{c}) = \begin{pmatrix} \boldsymbol{b} \\ \vdots \\ \boldsymbol{b} \\ \boldsymbol{c} \end{pmatrix}, \tag{4}$$

where $\boldsymbol{b}$ is repeated N times on the right-hand side. For $(\boldsymbol{b}^{(N)}, \boldsymbol{c}) \in \mathbb{N}A^{(N)}$, an element of the fiber $\mathcal{F}_{A^{(N)},(\boldsymbol{b}^{(N)},\boldsymbol{c})}$ of $A^{(N)}$ is written as $\boldsymbol{w} = (\boldsymbol{w}_1, \dots, \boldsymbol{w}_N)$, where $\boldsymbol{w}_k \in \mathcal{F}_{A,\boldsymbol{b}}, k = 1, \dots, N$, and $\boldsymbol{w}_1 + \cdots + \boldsymbol{w}_N = \boldsymbol{c}$. This is the same as (3). Hence any element of the fiber $\mathcal{F}_{A_{\boldsymbol{b}},\boldsymbol{c}}$ of $A_{\boldsymbol{b}}$ corresponds to an element of the fiber $\mathcal{F}_{A^{(N)},(\boldsymbol{b}^{(N)},\boldsymbol{c})}$ of $A^{(N)}$. This correspondence between $\mathcal{F}_{A_{\boldsymbol{b}},\boldsymbol{c}}$ and $\mathcal{F}_{A^{(N)},(\boldsymbol{b}^{(N)},\boldsymbol{c})}$ is one-to-one except for the permutation of vectors $\boldsymbol{w}_1, \dots, \boldsymbol{w}_N$. Note that the same N $\boldsymbol{b}$'s on the right-hand side of (4) may be different for general fibers of $A^{(N)}$. Hence the set of fibers $\mathbb{N}A_{\boldsymbol{b}}$ for $A_{\boldsymbol{b}}$ is a subset of the set of fibers $\cup_{N \geq 1}\mathbb{N}A^{(N)}$. As discussed in [10], Markov bases for a subset of fibers may be smaller than the full Markov bases. This fact is reflected in the inequality in (2).

Now we give a proof of Theorem 1.

Proof. [Proof of Theorem 1] Define a map $f_{\boldsymbol{b}} : \mathcal{F}_{A^{(N)},(\boldsymbol{b}^{(N)},\boldsymbol{c})} \to \mathcal{F}_{A_{\boldsymbol{b}},\boldsymbol{c}}$ by

$$f_{\boldsymbol{b}}(\boldsymbol{w}) = \boldsymbol{y} = (y_1, \dots, y_{\nu(\boldsymbol{b})}), \quad y_i = |\{k \mid \boldsymbol{w}_k = \boldsymbol{x}_i\}|.$$

Then $f_{\boldsymbol{b}}$ is a surjection and furthermore

$$f_{\boldsymbol{b}}(\boldsymbol{w}) = \sum_{k=1}^{N} f_{\boldsymbol{b}}((\boldsymbol{0}, \dots, \boldsymbol{0}, \boldsymbol{w}_k, \boldsymbol{0}, \dots, \boldsymbol{0})) = \sum_{k=1}^{N} (0, \dots, 0, \underset{i:\boldsymbol{x}_i = \boldsymbol{w}_k}{1}, 0, \dots, 0).$$

For any $\boldsymbol{y}^{(s)}, \boldsymbol{y}^{(t)} \in \mathcal{F}_{A_{\boldsymbol{b}},\boldsymbol{c}}$ we choose

$$\boldsymbol{w}^{(s)} \in f_{\boldsymbol{b}}^{-1}(\boldsymbol{y}^{(s)}), \quad \boldsymbol{w}^{(t)} \in f_{\boldsymbol{b}}^{-1}(\boldsymbol{y}^{(t)})$$

and we connect $\boldsymbol{w}^{(s)}$ and $\boldsymbol{w}^{(t)}$ by a Markov basis consisting of moves of type at most MC(A) of $A^{(N)}$. Denote the path from $\boldsymbol{w}^{(s)}$ to $\boldsymbol{w}^{(t)}$ in $\mathcal{F}_{A^{(N)},(\boldsymbol{b}^{(N)},\boldsymbol{c})}$ as

$$\boldsymbol{w}^{(s)} = \boldsymbol{w}^{(0)} \to \boldsymbol{w}^{(1)} \to \cdots \to \boldsymbol{w}^{(T)} = \boldsymbol{w}^{(t)}.$$

Let $\boldsymbol{y}^{(l)} = f_{\boldsymbol{b}}(\boldsymbol{w}^{(l)})$, $l = 0, 1 \dots, T$. Then

$$A_{\boldsymbol{b}}\boldsymbol{y}^{(l)} = \sum_{i=1}^{\nu(\boldsymbol{b})} y_i^{(l)} \boldsymbol{x}_i = \boldsymbol{w}_1^{(l)} + \cdots + \boldsymbol{w}_N^{(l)} = \boldsymbol{c}$$

and $\boldsymbol{y}^{(l)} \in \mathcal{F}_{A_{\boldsymbol{b}},\boldsymbol{c}}$. Hence $\boldsymbol{y}^{(l+1)} - \boldsymbol{y}^{(l)}$ is a move for $A_{\boldsymbol{b}}$. Its degree is bounded as

$$
\begin{aligned}
\frac{1}{2}\,|\boldsymbol{y}^{(l+1)} - \boldsymbol{y}^{(l)}| &= \frac{1}{2}\,|f_{\boldsymbol{b}}(\boldsymbol{w}^{(l+1)}) - f_{\boldsymbol{b}}(\boldsymbol{w}^{(l)})| \\
&= \frac{1}{2}\,|\sum_{k:\boldsymbol{w}_k^{(l+1)} \neq \boldsymbol{w}_k^{(l)}} f_{\boldsymbol{b}}((\boldsymbol{0},\dots,\boldsymbol{0},\boldsymbol{w}_k^{(l+1)},\boldsymbol{0},\dots,\boldsymbol{0})) - f_{\boldsymbol{b}}((\boldsymbol{0},\dots,\boldsymbol{0},\boldsymbol{w}_k^{(l)},\boldsymbol{0},\dots,\boldsymbol{0}))| \\
&\leq \frac{1}{2} \sum_{k:\boldsymbol{w}_k^{(l+1)} \neq \boldsymbol{w}_k^{(l)}} |f_{\boldsymbol{b}}((\boldsymbol{0},\dots,\boldsymbol{0},\boldsymbol{w}_k^{(l+1)},\boldsymbol{0},\dots,\boldsymbol{0})) - f_{\boldsymbol{b}}((\boldsymbol{0},\dots,\boldsymbol{0},\boldsymbol{w}_k^{(l)},\boldsymbol{0},\dots,\boldsymbol{0}))| \\
&= |\{k \mid \boldsymbol{w}_k^{(l+1)} \neq \boldsymbol{w}_k^{(l)}\}| = \mathrm{type}(\boldsymbol{w}^{(l+1)} - \boldsymbol{w}^{(l)}) \\
&\leq \mathrm{MC}(A).
\end{aligned}
$$

Thus $\boldsymbol{y}^{(s)}$ and $\boldsymbol{y}^{(t)}$ can be connected by moves of degree less than or equal to MC(A).

In Theorem 1 an interesting question is when (2) holds with equality. At this point we give a simple but important example. As the base configuration consider a $1 \times n$ row vector $A = (1, 1, \dots, 1)$. Then for any positive integer b, the fiber A_b is the configuration of Veronese-type (Chapter 14 of [17]), whose Markov degree is two. Hence $\max_{\boldsymbol{b}\in\mathbb{N}A} \mathrm{MD}(A_{\boldsymbol{b}}) = 2$. On the other hand, $A^{(N)}$ is the configuration matrix of the complete bipartite graph $K_{n,N}$. Since $A^{(N)}$, $N \geq 2$, has a Markov basis consisting of moves of degree two, we have $\mathrm{MC}(A) = 2$. Hence the equality in (2) holds for this case. Also note that $\mathrm{GC}(A) = n$, since the elements of Graver basis corresponds to cycles of $K_{n,N}$.

For bounding the Markov complexity MC(A) from below, we will find an indispensable move for the higher Lawrence lifting $A^{(N)}$ of A. The following proposition is useful for this purpose. We use the notation $[N] = \{1, 2, \dots, N\}$.

Proposition 1. *Let $\boldsymbol{z} = (\boldsymbol{z}_1, \dots, \boldsymbol{z}_N)$ be a move for $A^{(N)}$ such that each slice $\boldsymbol{z}_k$ is a non-zero indispensable move for A. Then $\boldsymbol{z}$ is indispensable if and only if*

$$\sum_{k\in M} \boldsymbol{z}_k \neq 0$$

for every non-empty proper subset M of $[N]$.

Proof. Write $\boldsymbol{z}$ by its positive part and negative part as $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^-$ and let $\boldsymbol{b}^{(N)} = A^{(N)}\boldsymbol{z}^+$. $\boldsymbol{z}$ is an indispensable move if and only if $\mathcal{F}_{A^{(N)},\boldsymbol{b}^{(N)}} = \{\boldsymbol{z}^+, \boldsymbol{z}^-\}$ is a two-element set. Also write each slice $\boldsymbol{z}_k$ as $\boldsymbol{z}_k = \boldsymbol{z}_k^+ - \boldsymbol{z}_k^-$ and let $\boldsymbol{b}_k = A\boldsymbol{z}_k^+$. We are assuming that $\mathcal{F}_{A,\boldsymbol{b}_k} = \{\boldsymbol{z}_k^+, \boldsymbol{z}_k^-\}$ is a two-element set for each k. Let $\boldsymbol{x} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_N) \in \mathcal{F}_{A^{(N)},\boldsymbol{b}^{(N)}}$. Then $A\boldsymbol{x}_k = \boldsymbol{b}_k$ for each k and hence $\boldsymbol{x}_k$ is either $\boldsymbol{z}_k^+$ or $\boldsymbol{z}_k^-$. Let $M = \{k \mid \boldsymbol{x}_k = \boldsymbol{z}_k^+\}$. Then $\boldsymbol{x}$ is different from both $\boldsymbol{z}^+$ and $\boldsymbol{z}^-$ if and only if M is a non-empty proper subset of $[N]$. Now $\sum_{k=1}^N \boldsymbol{x}_k = \sum_{k=1}^N \boldsymbol{z}_k^- = \boldsymbol{c}$ (say) implies

$$0 = \sum_{k=1}^{N} (\boldsymbol{x}_k - \boldsymbol{z}_k^-) = \sum_{k\in M} (\boldsymbol{z}_k^+ - \boldsymbol{z}_k^-) = \sum_{k\in M} \boldsymbol{z}_k. \tag{5}$$

Hence $\boldsymbol{z}$ is indispensable if and only if (5) hold only for $M = \emptyset$ or $M = [N]$.

Note that $\sum_{k\in M} z_k = 0$ if and only if $\sum_{k\in M^C} z_k = 0$ and any slice k is either in M or in M^C. Hence in order to prove that z is indispensable, we can start from arbitrary slice z_k and show that any sum of slices including k does not vanish except for the sum of all slices.

3. Complete graphs as base configurations

In this section we study the maximum Markov degree and the Markov complexity when the base configuration A is an incidence matrix of a small complete graph without self-loops (Section 3.1) or with self-loops (Section 3.2).

In $\boldsymbol{b} = A\boldsymbol{x}$, the elements of $\boldsymbol{x}$ are the non-negative integer weights of the edges and the elements of $\boldsymbol{b}$ are degrees of vertices, where the degree of a vertex v is the sum of weights of the edges having v as an endpoint. Note that one self-loop $\{v, v\}$ gives two degrees to the vertex v.

In the following, by $\boldsymbol{g}$ we denote a graph with non-negative weights attached to the edges. The elements of a fiber $\mathcal{F}_{A,\boldsymbol{b}}$ are the graphs $\boldsymbol{g}$ with the same degree sequence $\boldsymbol{b}$. See Figure 1 below for an example.

Elements of a fiber $\mathcal{F}_{A_{\boldsymbol{b}},\boldsymbol{c}}$ can be identified with multisets of graphs $\boldsymbol{g}$ such that the sum of weights of each edge is common. A move of degree k for the configuration $A_{\boldsymbol{b}}$ corresponds to replacing k graphs $\boldsymbol{g}_1, \ldots, \boldsymbol{g}_k \in \mathcal{F}_{A,\boldsymbol{b}}$ with $\hat{\boldsymbol{g}}_1, \ldots, \hat{\boldsymbol{g}}_k \in \mathcal{F}_{A,\boldsymbol{b}}$ such that the sum of weights of each edge is preserved.

3.1. Complete graph on four vertices without self-loops

In this section we take the incidence matrix of the complete graph K_4 on four vertices without self-loops as the base configuration A. At the end of this section we give some comments on larger complete graphs. In particular we present a conjecture on K_5.

Let

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}. \tag{6}$$

We prove that both sides of (2) are two and the equality holds for this A.

Theorem 2. *For A in* (6)

$$\max_{\boldsymbol{b}\in\mathbb{N}A} \mathrm{MD}(A_{\boldsymbol{b}}) = \mathrm{MC}(A) = 2.$$

By `4ti2`([1]) we easily obtain $\mathrm{GC}(A) = 3$, which equals the maximum 1-norm of $\mathcal{G}(\mathcal{G}(A))$ (Theorem 3 of [16]).

Denote the four vertices as a, b, c, d, corresponding to the rows of A. There are six edges corresponding to the columns of A. Let $E = \{ab, ac, ad, bc, bd, cd\}$ denote the edge set. A graph $\boldsymbol{g}$ is identified with a 6-dimensional non-negative integer vector

$$\boldsymbol{g} = (g(ab), g(ac), g(ad), g(bc), g(bd), g(cd)) \in \mathbb{N}^6,$$

whose elements represent weights of the edges. For two graphs $\boldsymbol{g}, \hat{\boldsymbol{g}}$ in the same fiber of A, we write $\boldsymbol{z} = \boldsymbol{g} - \hat{\boldsymbol{g}} = (z(ab), \ldots, z(cd))$, which is a move for A.

We prove two lemmas.

Lemma 1. *Let $\boldsymbol{g}, \hat{\boldsymbol{g}}$ be graphs in the same fiber of A and let $\boldsymbol{z} = \boldsymbol{g} - \hat{\boldsymbol{g}}$. Then*

$$z(ab) = z(cd),\ z(ac) = z(bd),\ z(ad) = z(bc).$$

Proof. By symmetry it suffices to prove $z(ab) = z(cd)$. Let $\deg(a)$ denote the degree of vertex a. We have

$$\begin{aligned}\deg(a) &= g(ab) + g(ac) + g(ad) = \hat{g}(ab) + \hat{g}(ac) + \hat{g}(ad),\\ \deg(b) &= g(ab) + g(bc) + g(bd) = \hat{g}(ab) + \hat{g}(bc) + \hat{g}(bd).\end{aligned}$$

Hence

$$\begin{aligned}\deg(a) + \deg(b) &= 2g(ab) + g(ac) + g(ad) + g(bc) + g(bd)\\ &= 2\hat{g}(ab) + \hat{g}(ac) + \hat{g}(ad) + \hat{g}(bc) + \hat{g}(bd).\end{aligned}$$

Similarly

$$\begin{aligned}\deg(c) + \deg(d) &= 2g(cd) + g(ac) + g(ad) + g(bc) + g(bd)\\ &= 2\hat{g}(cd) + \hat{g}(ac) + \hat{g}(ad) + \hat{g}(bc) + \hat{g}(bd).\end{aligned}$$

Then

$$\deg(a) + \deg(b) - (\deg(c) + \deg(d)) = 2(g(ab) - g(cd)) = 2(\hat{g}(ab) - \hat{g}(cd))$$

and

$$g(ab) - \hat{g}(ab) = g(cd) - \hat{g}(cd).$$

Lemma 2. *Let $\boldsymbol{g}, \hat{\boldsymbol{g}}$ in the same fiber of A and let $g(e_1) \neq \hat{g}(e_1)$ for some $e_1 \in E$. Then there exists a loop (e_1, e_2, e_3, e_4) of length 4 passing each vertex, such that $g(e_i) \neq \hat{g}(e_i)$, $i = 1, \ldots, 4$, and the signs of $g(e_i) - \hat{g}(e_i)$ alternate.*

Proof. By symmetry we may assume that $e_1 = ab$ and $g(ab) - \hat{g}(ab) > 0$. Then by the previous lemma $g(cd) - \hat{g}(cd) > 0$. Since $\deg(a)$ is common in $\boldsymbol{g}$ and $\hat{\boldsymbol{g}}$, by symmetry we may assume that $g(ad) - \hat{g}(ad) < 0$. Again by the previous lemma $g(bc) - \hat{g}(bc) < 0$. Then (ab, bc, cd, ad) is the required loop.

We now give a proof of Theorem 2 based on the idea of distance reduction (cf. Chapter 6 of [2]).

Proof. [Proof of Theorem 2] Obviously $\max_{\boldsymbol{b} \in \mathbb{N}A} \mathrm{MD}(A_{\boldsymbol{b}}) > 1$. Hence by Theorem 1 it suffices to prove that $\mathrm{MC}(A) = 2$. Let $\{\boldsymbol{g}_1, \ldots, \boldsymbol{g}_N\}$ and $\{\hat{\boldsymbol{g}}_1, \ldots, \hat{\boldsymbol{g}}_N\}$ be two elements of the same fiber for $A^{(N)}$. Let

$$S = \sum_{k=1}^{N} |\boldsymbol{z}_k|, \qquad \boldsymbol{z}_k = \boldsymbol{g}_k - \hat{\boldsymbol{g}}_k,$$

where $|\cdot|$ denotes the 1-norm of a 6-dimensional vector.

Suppose $S > 0$. By symmetry we may assume that $\boldsymbol{g}_1 \neq \hat{\boldsymbol{g}}_1$. By Lemma 2 we may assume

$$z_1(ab) > 0,\ z_1(bc) < 0,\ z_1(cd) > 0,\ z_1(ad) < 0.$$

Because $\{\boldsymbol{g}_1, \dots, \boldsymbol{g}_N\}$ and $\{\hat{\boldsymbol{g}}_1, \dots, \hat{\boldsymbol{g}}_N\}$ belong to the same fiber, we have

$$\sum_{k=1}^{N} z_k(e) = 0$$

for each $e \in E$ (in particular for $e = bc$). Hence there exists k such that $z_k(bc) > 0$. Let $k = 2$ without loss of generality. By Lemma 1 $g_2(ad) > \hat{g}_2(ad)$. Let

$$\boldsymbol{e}_{ab} = (1, 0, 0, 0, 0, 0) \tag{7}$$

denote the graph with weight 1 only on the edge ab. Similarly define $\boldsymbol{e}_{bc}, \boldsymbol{e}_{cd}, \boldsymbol{e}_{ad}$. Now consider the move

$$(\boldsymbol{g}_1, \boldsymbol{g}_2) \to (\boldsymbol{g}_1 + \boldsymbol{e}_{ab} - \boldsymbol{e}_{bc} + \boldsymbol{e}_{cd} - \boldsymbol{e}_{ad}, \boldsymbol{g}_2 - \boldsymbol{e}_{ab} + \boldsymbol{e}_{bc} - \boldsymbol{e}_{cd} + \boldsymbol{e}_{ad}). \tag{8}$$

Then the vectors on the right-hand side are non-negative and S is strictly decreased. This proves $\mathrm{MC}(A) = 2$.

Remark 2. *Hidefumi Ohsugi gave a simple direct proof of* $\max_{\boldsymbol{b}\in\mathbb{N}A} \mathrm{MD}(A_{\boldsymbol{b}}) = 2$ *by identifying* $A_{\boldsymbol{b}}$ *with a Segre–Veronese configuration. See [3] for the definition of Segre–Veronese configurations and their application to algebraic statistics.*

The move in (8) can be understood as an *exchange* or *swap* of edges between two graphs $\boldsymbol{g}_1, \boldsymbol{g}_2$, i.e., edges bc and ad are given from $\boldsymbol{g}_1$ to $\boldsymbol{g}_2$, and edges ab and cd are taken from $\boldsymbol{g}_2$ to $\boldsymbol{g}_1$. A move of degree two for $A_{\boldsymbol{b}}$ and a move of type two for $A^{(N)}$ is an exchange of edges between two graphs. Similarly a move of degree k for $A_{\boldsymbol{b}}$ and a move of type k for $A^{(N)}$ is an exchange of edges among k graphs.

At this point, we make some remarks on larger complete graphs without self-loops. Consider the complete graph K_5 of five vertices without self-loops and let

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix} \tag{9}$$

be its incidence matrix. By `4ti2` we can check

$$\mathrm{MC}(A) \geq 6, \quad \mathrm{GC}(A) = 15.$$

Concerning $\max_{\boldsymbol{b}\in\mathbb{N}A} \mathrm{MD}(A_{\boldsymbol{b}})$ we make the following conjecture. For A in (9)

$$\max_{\boldsymbol{b}\in\mathbb{N}A} \mathrm{MD}(A_{\boldsymbol{b}}) = 2. \tag{10}$$

Our conjecture is based on the numbers of moves of degrees two and three or higher in minimal Markov bases for various $A_{\boldsymbol{b}}$ in Table 1 computed with `4ti2`. Note that there are no moves of degree three or higher, as far as we could compute with `4ti2`.

For the case K_6 of 6 vertices, we can easily check that $\max_{\boldsymbol{b}\in\mathbb{N}A} \mathrm{MD}(A_{\boldsymbol{b}}) \geq 4$.

Table 1: Number of moves in minimal Markov bases for $A_{\boldsymbol{b}}$ in the case of K_5

$\boldsymbol{b}$	# moves of deg 2	# moves of deg ≥3
(2,2,2,1,1)	9	0
(2,2,2,2,2)	95	0
(3,2,2,2,1)	39	0
(3,3,2,1,1)	9	0
(3,3,2,2,2)	16	0
(3,3,3,2,1)	105	0
(3,3,3,3,2)	741	0
(4,2,2,2,2)	105	0
(4,3,2,2,1)	39	0
(4,3,3,1,1)	9	0
(4,3,3,2,2)	413	0
(4,3,3,3,1)	225	0
(4,3,3,3,3)	1893	0
(4,4,2,1,1)	9	0
(4,4,2,2,2)	216	0
(4,4,3,2,1)	105	0
(4,4,3,3,2)	1179	0
(4,4,4,2,2)	710	0
(4,4,4,3,1)	420	0
(4,4,4,3,3)	4032	0
(4,4,4,4,2)	2718	0
(4,4,4,4,4)	10581	0

3.2. Complete graph on three vertices with self-loops

We consider the incidence matrix of the complete graph on three vertices with self-loops as the base configuration A:

$$A = \begin{pmatrix} 2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 2 \end{pmatrix}. \tag{11}$$

The following theorem holds.

Theorem 3. *For A in* (11)

$$\max_{\boldsymbol{b}\in\mathbb{N}A} \mathrm{MD}(A_{\boldsymbol{b}}) = 3, \quad \mathrm{MC}(A) = 5. \tag{12}$$

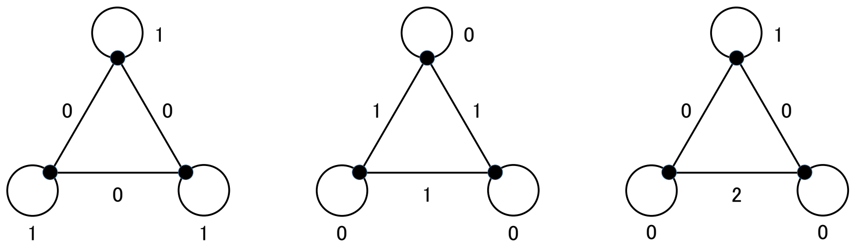
Furthermore $\max_{\boldsymbol{b}\in\mathbb{N}A\setminus\{(2,2,2)\}} \mathrm{MD}(A_{\boldsymbol{b}}) = 2$.

Incidentally we obtained $\mathrm{GC}(A) = 8$ by `4ti2` and Theorem 3 of [16].

As stated in Theorem 3, the fiber with $\boldsymbol{b} = (2,2,2)$ is special. $\mathcal{F}_{A,(2,2,2)}$ consists of five vectors and $A_{(2,2,2)}$ is given as

$$A_{(2,2,2)} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 2 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{pmatrix} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4, \boldsymbol{a}_5).$$

Columns of $A_{(2,2,2)}$ are displayed in Figure 1.



Column 1 of $A_{(2,2,2)}$ Column 4 of $A_{(2,2,2)}$ Columns 2,3,5 of $A_{(2,2,2)}$

Figure 1: Graphs of the fiber $\mathcal{F}_{A,(2,2,2)}$

In this case rank $A_{(2,2,2)} = 4$ and the toric ideal $I_{A_{(2,2,2)}}$ associated with $A_{(2,2,2)}$ is a principal ideal generated by the relation

$$\boldsymbol{a}_1 + 2\boldsymbol{a}_4 = \boldsymbol{a}_2 + \boldsymbol{a}_3 + \boldsymbol{a}_5.$$

Hence

$$\mathrm{MD}(A_{(2,2,2)}) = 3. \tag{13}$$

To express vertices and edges, we label the vertices as Figure 2. Then for example we express the

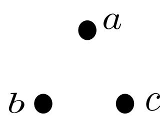


Figure 2: Labels of vertices

self-loop $\{a, a\}$ by an edge aa.

For the rest of this subsection we give a proof of Theorem 3.

It is easy to see that $\mathrm{MD}(A_{\boldsymbol{b}}) \leq 2$ if $\min(\deg(a), \deg(b), \deg(c)) \leq 1$. Hence from now on we assume that the degrees of three vertices are at least two. For our proof we utilize the Graver basis $\mathcal{G}(A)$ of A in (11). By `4ti2` or by checking the moves for A, it is easily verified that $\mathcal{G}(A)$ consists of ten column vectors in (14) and those with the minus sign. Hence $|\mathcal{G}(A)| = 20$. There are four patterns of moves and patterns $\boldsymbol{B}$ and $\boldsymbol{C}$ are indispensable moves.

$$
\begin{array}{c|cccccccccc}
 & \boldsymbol{A} & \boldsymbol{B}(a) & \boldsymbol{B}(b) & \boldsymbol{B}(c) & \boldsymbol{C}(a) & \boldsymbol{C}(b) & \boldsymbol{C}(c) & \boldsymbol{D}(a) & \boldsymbol{D}(b) & \boldsymbol{D}(c) \\
\hline
aa & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & -1 \\
ab & -1 & -1 & -1 & 1 & 0 & 0 & -2 & 2 & -2 & 0 \\
ac & -1 & -1 & 1 & -1 & 0 & -2 & 0 & -2 & 0 & 2 \\
bb & 1 & 0 & 1 & 0 & 1 & 0 & 1 & -1 & 0 & 1 \\
bc & -1 & 1 & -1 & -1 & -2 & 0 & 0 & 0 & 2 & -2 \\
cc & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & -1 & 0
\end{array}
\tag{14}
$$

By using the notation in (7), the move $\boldsymbol{A}$ is written as

$$\boldsymbol{A} = \boldsymbol{e}_{aa} + \boldsymbol{e}_{bb} + \boldsymbol{e}_{cc} - \boldsymbol{e}_{ab} - \boldsymbol{e}_{bc} - \boldsymbol{e}_{ac}.$$

We denote 20 moves of $\mathcal{G}(A)$ by $\boldsymbol{A}, \boldsymbol{B}(a), \ldots, \boldsymbol{D}(c)$ and $-\boldsymbol{A}, -\boldsymbol{B}(a), \ldots, -\boldsymbol{D}(c)$. Moves $\boldsymbol{A}$, $\boldsymbol{B}(a)$, $\boldsymbol{C}(a)$, $\boldsymbol{D}(a), \boldsymbol{D}(b), \boldsymbol{D}(d)$ are displayed in Figure 3. For checking our proof of Theorem 3 it is

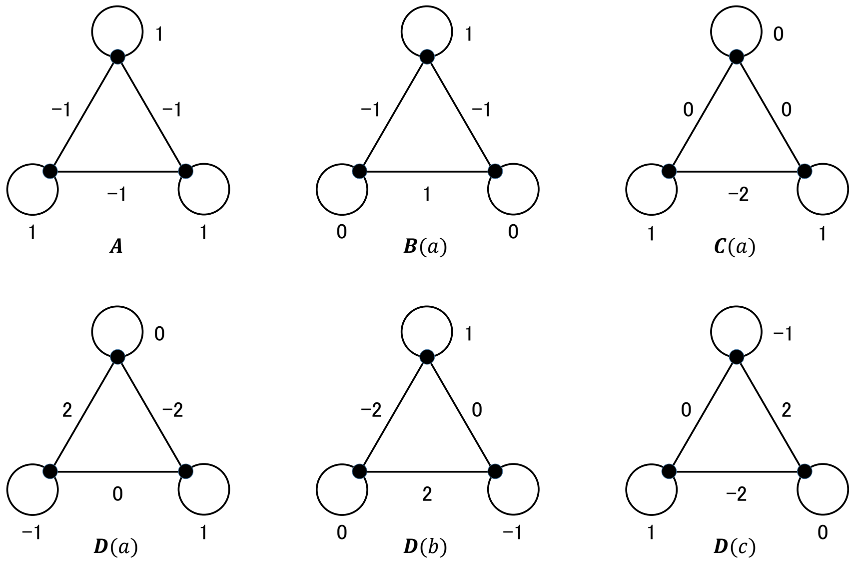


Figure 3: Moves $\boldsymbol{A}$, $\boldsymbol{B}(a)$, $\boldsymbol{C}(a)$, $\boldsymbol{D}(a)$, $\boldsymbol{D}(b)$, $\boldsymbol{D}(c)$

convenient to have graphs for $\boldsymbol{B}(b)$, $\boldsymbol{B}(c)$, $\boldsymbol{C}(b)$, $\boldsymbol{C}(c)$ in Figure 4.

For a move $\boldsymbol{z} \in \ker_{\mathbb{Z}} A \subset \mathbb{Z}^6$, $\boldsymbol{z} \neq 0$, there exists $\boldsymbol{w} \in \mathcal{G}(A)$ such that $\boldsymbol{w} + (\boldsymbol{z} - \boldsymbol{w}) = \boldsymbol{z}$ is a conformal sum, i.e., there is no cancellation of signs in this sum. In this case we write

$$\boldsymbol{w} \sqsubseteq \boldsymbol{z}.$$

Here we are allowing the case $\boldsymbol{z} = \boldsymbol{w}$.

Let $\boldsymbol{g}, \hat{\boldsymbol{g}} \in \mathbb{N}^6$ be two graphs in the same fiber $\mathcal{F}_{A,\boldsymbol{b}}$ of A. Then $\boldsymbol{z} = \boldsymbol{g} - \hat{\boldsymbol{g}}$ is a move and there exists $\boldsymbol{w} \in \mathcal{G}(A)$ such that $\boldsymbol{w} \sqsubseteq \boldsymbol{g} - \hat{\boldsymbol{g}}$. In this case we say that "$(\boldsymbol{g}, \hat{\boldsymbol{g}})$ *contains* $\boldsymbol{w}$". Note that $(\boldsymbol{g}, \hat{\boldsymbol{g}})$

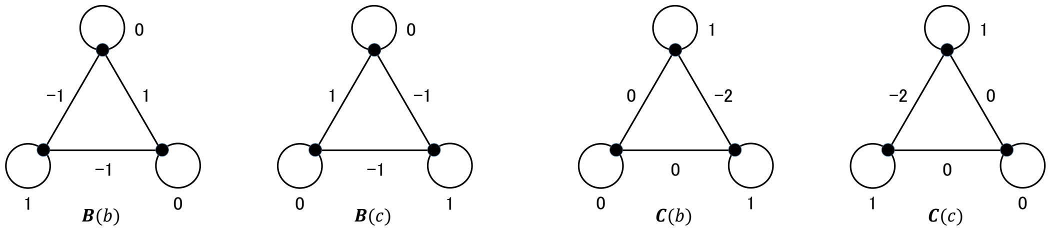


Figure 4: Moves $\boldsymbol{B}(b)$, $\boldsymbol{B}(c)$, $\boldsymbol{C}(b)$, $\boldsymbol{C}(c)$

contains $\boldsymbol{w}$ if and only if $(\hat{\boldsymbol{g}}, \boldsymbol{g})$ contains $-\boldsymbol{w}$. Also if $(\boldsymbol{g}, \hat{\boldsymbol{g}})$ contains $\boldsymbol{w}$ then $\boldsymbol{g}-\boldsymbol{w} \geq 0$ (elementwise) and

$$|(\boldsymbol{g}-\boldsymbol{w})-\hat{\boldsymbol{g}}| = |\boldsymbol{g}-\hat{\boldsymbol{g}}| - |\boldsymbol{w}|.$$

When $(\boldsymbol{g}, \hat{\boldsymbol{g}})$ contains $\boldsymbol{w}$, we denote $\boldsymbol{g}$ by $\boldsymbol{g}_{\boldsymbol{w}}$, provided that there is no confusion about $\hat{\boldsymbol{g}}$. For example $\boldsymbol{g}_{-\boldsymbol{A}}$ denotes a graph $\boldsymbol{g}$ in $(\boldsymbol{g}, \hat{\boldsymbol{g}})$ which contains the negative of the first column of (14). Now we begin proving $\max_{\boldsymbol{b}\in\mathbb{N}A} \mathrm{MD}(A_{\boldsymbol{b}}) = 3$.

I. Proof of $\max_{\boldsymbol{b}\in\mathbb{N}A} \mathrm{MD}(A_{\boldsymbol{b}}) = 3$.

We choose two arbitrary elements of $\mathcal{F}_{A_{\boldsymbol{b}},\boldsymbol{c}} = \{\boldsymbol{y} \mid A_{\boldsymbol{b}}\boldsymbol{y} = \boldsymbol{c}\}$ and denote them by $\boldsymbol{y}$ and $\hat{\boldsymbol{y}}$. Although $\boldsymbol{y}$ and $\hat{\boldsymbol{y}}$ are multisets of graphs, by the embedding of a fiber of $A_{\boldsymbol{b}}$ into a fiber of $A^{(N)}$ discussed after Theorem 1, we index the graphs of $\boldsymbol{y}$ as $\boldsymbol{g}_1, \boldsymbol{g}_2, \dots, \boldsymbol{g}_N$ and graphs of $\hat{\boldsymbol{y}}$ as $\hat{\boldsymbol{g}}_1, \hat{\boldsymbol{g}}_2, \dots, \hat{\boldsymbol{g}}_N$. Then

$$A\boldsymbol{g}_k = A\hat{\boldsymbol{g}}_k = \boldsymbol{b}, \quad k = 1, \dots, N, \tag{15}$$

$$\boldsymbol{g}_1 + \boldsymbol{g}_2 + \cdots + \boldsymbol{g}_N = \hat{\boldsymbol{g}}_1 + \hat{\boldsymbol{g}}_2 + \cdots + \hat{\boldsymbol{g}}_N = \boldsymbol{c}. \tag{16}$$

As in the proof of Theorem 2, let

$$S = \sum_{k=1}^{N} |\boldsymbol{g}_k - \hat{\boldsymbol{g}}_k|.$$

Then, $S = 0$ implies $\boldsymbol{y} = \hat{\boldsymbol{y}}$. We will show that if $S > 0$ there exists an exchange of edges among some fixed number graphs in $\boldsymbol{y}$ or in $\hat{\boldsymbol{y}}$ such that S is decreased.

If $S > 0$, there exists a layer k satisfying $\boldsymbol{g}_k \neq \hat{\boldsymbol{g}}_k$. By $A(\boldsymbol{g}_k - \hat{\boldsymbol{g}}_k) = 0$, there exists $\boldsymbol{w} \in \mathcal{G}(A)$ such that $(\boldsymbol{g}_k, \hat{\boldsymbol{g}}_k)$ contains $\boldsymbol{w}$. In this case we say that there exists a *pattern* $\boldsymbol{w}$ among $\boldsymbol{z}_k = \boldsymbol{g}_k - \hat{\boldsymbol{g}}_k$, $k = 1, \dots, N$. For example, suppose that the pattern $\boldsymbol{B}(a)$ exists. Then for some k, $z_k(aa) > 0$ and $z_k(ab) < 0$. By $0 = \sum_{k=1}^{N} \boldsymbol{z}_k$, there have to be some other layers k', k'' such that $z_{k'}(aa) < 0$ and $z_{k''}(ab) > 0$. In this case we say that the edge aa is "in shortage" and the edge ab is "in excess" on some layers other than k.

At this point we consider an easy case to decrease S, where there are $\boldsymbol{g}_{\boldsymbol{A}}$ and $\boldsymbol{g}_{-\boldsymbol{A}}$, i.e., there are k and k' such that $(\boldsymbol{g}_k, \hat{\boldsymbol{g}}_k)$ contains the move $\boldsymbol{A}$ and $(\boldsymbol{g}_{k'}, \hat{\boldsymbol{g}}_{k'})$ contains the move $-\boldsymbol{A}$. Then we can apply an exchange of edges $(\boldsymbol{g}_{\boldsymbol{A}}, \boldsymbol{g}_{-\boldsymbol{A}}) \to (\boldsymbol{g}'_{\boldsymbol{A}}, \boldsymbol{g}'_{-\boldsymbol{A}})$, where

$$\boldsymbol{g}'_A = \boldsymbol{g}_A - \boldsymbol{e}_{aa} - \boldsymbol{e}_{bb} - \boldsymbol{e}_{cc} + \boldsymbol{e}_{ab} + \boldsymbol{e}_{bc} + \boldsymbol{e}_{ca} = \boldsymbol{g}_A - \boldsymbol{A},$$

$$\boldsymbol{g}'_{-A} = \boldsymbol{g}_{-A} + \boldsymbol{e}_{aa} + \boldsymbol{e}_{bb} + \boldsymbol{e}_{cc} - \boldsymbol{e}_{ab} - \boldsymbol{e}_{bc} - \boldsymbol{e}_{ca} = \boldsymbol{g}_{-A} + \boldsymbol{A}.$$

By this degree-two move (15) and (16) are conserved. Obviously $\boldsymbol{g}'_A$ and $\boldsymbol{g}'_{-A}$ are non-negative and S is immediately decreased. Similar consideration applies to other nine pairs of moves $(\boldsymbol{B}(a), -\boldsymbol{B}(a))$, $\dots$, $(\boldsymbol{D}(c), -\boldsymbol{D}(c))$. Therefore, from now on, we ignore the case that there are two layers containing any of these 10 pairs. Also note that by symmetry between $\boldsymbol{y}$ and $\hat{\boldsymbol{y}}$, we only need to consider one of $\boldsymbol{A}$ or $-\boldsymbol{A}$.

We now distinguish various cases. We first consider the case that the pattern $\boldsymbol{A}$ (or $-\boldsymbol{A}$) exists.

Case 1 $\boldsymbol{A}$ exists.

We are assuming that there exists some k such that $(\boldsymbol{g}_k, \hat{\boldsymbol{g}}_k)$ contains $\boldsymbol{A}$. There are three subcases depending on whether the pattern $\boldsymbol{B}$ exists or not on some other layer $k' \neq k$. By symmetry among a, b, c, we only need to consider $\boldsymbol{B}(a)$.

Case 1-1 $\boldsymbol{B}(a)$ exists.

Because of the existence of $\boldsymbol{A}$ and $\boldsymbol{B}(a)$, the edge aa is in shortage on some other layer k''. The possible patterns are $-\boldsymbol{C}(c), -\boldsymbol{C}(b), -\boldsymbol{D}(b)$, or $\boldsymbol{D}(c)$. By symmetry between b and c, we only need to consider $-\boldsymbol{C}(c)$ or $-\boldsymbol{D}(b)$. If $-\boldsymbol{C}(c)$ exists then S is decreased by

$$\boldsymbol{g}'_{\boldsymbol{A}} = \boldsymbol{g}_{\boldsymbol{A}} - \boldsymbol{C}(c), \qquad \boldsymbol{g}'_{-\boldsymbol{C}(c)} = \boldsymbol{g}_{-\boldsymbol{C}(c)} + \boldsymbol{C}(c)$$

and if $-\boldsymbol{D}(b)$ exists then S is decreased by

$$\boldsymbol{g}'_{\boldsymbol{A}} = \boldsymbol{g}_{\boldsymbol{A}} - \boldsymbol{C}(c), \qquad \boldsymbol{g}'_{-\boldsymbol{D}(b)} = \boldsymbol{g}_{-\boldsymbol{D}(b)} + \boldsymbol{C}(c).$$

Here note that $\boldsymbol{g}_{-\boldsymbol{D}(b)} + \boldsymbol{C}(c) \geq 0$. We omit this kind of remark on non-negativity for the rest this proof.

Case 1-2 $-\boldsymbol{B}(a)$ exists.

In this case we look at $\hat{\boldsymbol{y}}$. S is decreased by

$$\hat{\boldsymbol{g}}'_{-A} = \hat{\boldsymbol{g}}_{-A} + \boldsymbol{B}(a), \qquad \hat{\boldsymbol{g}}'_{B(a)} = \hat{\boldsymbol{g}}_{B(a)} - \boldsymbol{B}(a).$$

Case 1-3 None of $\boldsymbol{B}(a)$, $-\boldsymbol{B}(a)$ exists.

This case can be handled as in Case 1-1, since the edge aa is in shortage.

From now on, we assume that pattern $\pm\boldsymbol{A}$ does not exist. For Case 2, we consider the existence of the pattern $\pm\boldsymbol{D}$.

Case 2 $\boldsymbol{D}$ exists.

By symmetry we consider the case that there is some layer containing $\boldsymbol{D}(a)$. Since there is $\boldsymbol{D}(a)$, the edge bb is in excess on some other layer. The possible patterns for this excess are $\boldsymbol{C}(a), \boldsymbol{C}(c), \boldsymbol{D}(c)$, or $\boldsymbol{B}(b)$. Also the edge cc is in shortage. The possible patterns for this shortage are $-\boldsymbol{C}(a), -\boldsymbol{C}(b), \boldsymbol{D}(b)$, or $-\boldsymbol{B}(c)$.

Note that we are assuming that $\boldsymbol{C}(a)$ and $-\boldsymbol{C}(a)$ do not simultaneously exist, i.e., at least one of $\boldsymbol{C}(a)$ and $-\boldsymbol{C}(a)$ does not exist. If $\boldsymbol{C}(a)$ does not exist, then at least one of $\boldsymbol{C}(c), \boldsymbol{D}(c)$, or $\boldsymbol{B}(b)$ exist. Similarly if $-\boldsymbol{C}(a)$ dos not exist at least one of $-\boldsymbol{C}(b), \boldsymbol{D}(b)$ or $-\boldsymbol{B}(c)$ exist. Hence at least one of $\boldsymbol{C}(c), \boldsymbol{D}(c), \boldsymbol{B}(b), -\boldsymbol{C}(b), \boldsymbol{D}(b), -\boldsymbol{B}(c)$ exist.

Now by simultaneous symmetry $(b, \boldsymbol{y}, \boldsymbol{D}(a)) \leftrightarrow (c, \hat{\boldsymbol{y}}, -\boldsymbol{D}(a))$, we only need to consider one of $\boldsymbol{C}(c)$ and $-\boldsymbol{C}(b)$, one of $\boldsymbol{D}(c)$ and $\boldsymbol{D}(b)$, and one of $-\boldsymbol{B}(c)$ and $\boldsymbol{B}(b)$. Hence we will examine the cases $\boldsymbol{C}(c)$, $\boldsymbol{D}(c)$, $-\boldsymbol{B}(c)$, in turn.

Case 2-1 $\boldsymbol{D}(a)$ and $\boldsymbol{C}(c)$ exist.
S is decreased by

$$\boldsymbol{g}'_{\boldsymbol{D}(a)} = \boldsymbol{g}_{\boldsymbol{D}(a)} + \boldsymbol{C}(c), \qquad \boldsymbol{g}'_{\boldsymbol{C}(c)} = \boldsymbol{g}_{\boldsymbol{C}(c)} - \boldsymbol{C}(c).$$

Case 2-2 $\boldsymbol{D}(a)$ and $\boldsymbol{D}(c)$ exist.
S is decreased by

$$\boldsymbol{g}'_{\boldsymbol{D}(a)} = \boldsymbol{g}_{\boldsymbol{D}(a)} - \boldsymbol{D}(a), \qquad \boldsymbol{g}'_{\boldsymbol{D}(c)} = \boldsymbol{g}_{\boldsymbol{D}(c)} + \boldsymbol{D}(a).$$

Case 2-3 $\boldsymbol{D}(a)$ and $-\boldsymbol{B}(c)$ exist.
S is decreased by

$$\boldsymbol{g}'_{\boldsymbol{D}(a)} = \boldsymbol{g}_{\boldsymbol{D}(a)} - \boldsymbol{B}(c), \qquad \boldsymbol{g}'_{-\boldsymbol{B}(c)} = \boldsymbol{g}_{-\boldsymbol{B}(c)} + \boldsymbol{B}(c).$$

We have now examined all possible cases where $\pm\boldsymbol{D}$ exists. From now on, we may assume that pattern $\pm\boldsymbol{D}$ does not exist.

We now consider the case that the pattern $\pm\boldsymbol{B}$ exists.

Case 3 $\boldsymbol{B}$ exists.
By symmetry we assume that $\boldsymbol{B}(a)$ exists. Then because of the shortage of aa on other layers, there exists pattern $-\boldsymbol{C}(c)$ or $-\boldsymbol{C}(b)$. Because of symmetry of vertices b and c, it is enough to consider $-\boldsymbol{C}(c)$ only. Then because of the excess of bb, there exists pattern $\boldsymbol{B}(b)$ or $\boldsymbol{C}(a)$.

Case 3-1 $\boldsymbol{B}(a)$, $-\boldsymbol{C}(c)$ and $\boldsymbol{B}(b)$ exist.
S is decreased by

$$\boldsymbol{g}'_{B(a)} = \boldsymbol{g}_{B(a)} - \boldsymbol{B}(a), \qquad \boldsymbol{g}'_{-C(c)} = \boldsymbol{g}_{-C(c)} + \boldsymbol{C}(c), \qquad \boldsymbol{g}'_{B(b)} = \boldsymbol{g}_{B(b)} - \boldsymbol{B}(b).$$

Case 3-2 $\boldsymbol{B}(a)$, $-\boldsymbol{C}(c)$ and $\boldsymbol{C}(a)$ exist.
Note that already $\boldsymbol{D}(c)$ and $-\boldsymbol{D}(a)$ do not exist by our assumption. Also in Case 3-1 we considered the existence of $\boldsymbol{B}(b)$. Hence here we consider the case that $\boldsymbol{D}(c)$, $-\boldsymbol{D}(a)$ and $\boldsymbol{B}(b)$ do not exist, but $\boldsymbol{C}(a)$ exists. Then by the shortage of cc, there is a pattern $-\boldsymbol{C}(b)$ or $-\boldsymbol{B}(c)$.

Case 3-2-1 $\boldsymbol{B}(a)$, $-\boldsymbol{C}(c)$, $\boldsymbol{C}(a)$ and $-\boldsymbol{C}(b)$ exist.
This case is difficult. We renumber this case as Case 4 and will discuss this case below.

Case 3-2-2 $\boldsymbol{B}(a)$, $-\boldsymbol{C}(c)$, $\boldsymbol{C}(a)$ and $-\boldsymbol{B}(c)$ exist.

This case is also difficult. We renumber this case as Case 5 and will discuss this case below.

So far we did not use the fact that all graphs $\boldsymbol{g}_1, \dots, \boldsymbol{g}_N$ belong to the same fiber $\mathcal{F}_{A,\boldsymbol{b}}$ of A. Our argument before Case 3-2-1 apply not only to $A_{\boldsymbol{b}}$, but also to the higher Lawrence lifting $A^{(N)}$. However there is a gap between two sides of (12). In order to show the left-hand side $\max_{\boldsymbol{b}\in\mathbb{N}A} \mathrm{MD}(A_{\boldsymbol{b}}) = 3$ we need to use that fact that $\boldsymbol{g}_1, \dots, \boldsymbol{g}_N$ belong to the same fiber.

We now look at Case 4 from this viewpoint.

Case 4 $\boldsymbol{B}(a)$, $-\boldsymbol{C}(c)$, $\boldsymbol{C}(a)$ and $-\boldsymbol{C}(b)$ exist.

First note that the existence $\boldsymbol{B}(a)$ implies $\deg(a) \geq 2$. Also the existence of $\boldsymbol{C}(a)$ implies $\deg(b) \geq 2$, $\deg(c) \geq 2$. Hence the degree of each vertex is at least two. Then $\boldsymbol{g}_{\boldsymbol{B}(a)}$ has additional edges connecting to b and to c. The possible combinations of edges are

1) bc alone, 2) the pair (ab, ac), 3) the pair (bb, ac),
4) the pair (bc, cc), 5) the pair (bb, cc), or 6) the case that $\boldsymbol{g}_{\boldsymbol{B}(a)}$ has two $\boldsymbol{B}(a)$.

These six cases are depicted in Figure 5. Existence of an additional edge is shown as the weight of the form $+p-q$ in Figure 5. $+p$ means that we can subtract p edges without producing a negative weight.

Consider the edge aa. The weight of aa in $\boldsymbol{C}(a)$ is zero. On the other hand in both $-\boldsymbol{C}(c)$ and $-\boldsymbol{C}(b)$ its weight is -1. This extra shortage of aa implies that there exists another pattern $\boldsymbol{B}(a)$ in addition to the already existing $\boldsymbol{B}(a)$, possibly on the same layer as the already existing one or on another layer. The former case corresponds to 6) above.

Also note that $-\boldsymbol{C}(c)$ and $-\boldsymbol{C}(b)$ may be on the same layer, but in this case the weight of the self-loop aa on the layer is less than or equal to -2 and our proof is not affected.

Case 4-1 $\boldsymbol{B}(a)$, $-\boldsymbol{C}(c)$, $\boldsymbol{C}(a)$, $-\boldsymbol{C}(b)$ and $\boldsymbol{B}(a)_1$ exist.

S is decreased by

$$\boldsymbol{g}'_{\boldsymbol{B}(a)_1} = \boldsymbol{g}_{\boldsymbol{B}(a)_1} + \boldsymbol{C}(a), \qquad \boldsymbol{g}'_{\boldsymbol{C}(a)} = \boldsymbol{g}_{\boldsymbol{C}(a)} - \boldsymbol{C}(a).$$

Case 4-2 $\boldsymbol{B}(a)$, $-\boldsymbol{C}(c)$, $\boldsymbol{C}(a)$, $-\boldsymbol{C}(b)$ and $\boldsymbol{B}(a)_2$ exist.

By

$$\boldsymbol{g}'_{\boldsymbol{B}(a)_2} = \boldsymbol{g}_{\boldsymbol{B}(a)_2} + \boldsymbol{B}(a), \qquad \boldsymbol{g}'_{\boldsymbol{B}(a)} = \boldsymbol{g}_{\boldsymbol{B}(a)} - \boldsymbol{B}(a),$$

S is not changed, but $\boldsymbol{g}_{\boldsymbol{B}(a)_2}$ now has $\boldsymbol{B}(a)_6$. Then we will decrease S in Case 4-6 below.

Case 4-3 $\boldsymbol{B}(a)$, $-\boldsymbol{C}(c)$, $\boldsymbol{C}(a)$, $-\boldsymbol{C}(b)$ and $\boldsymbol{B}(a)_3$ exist.

S is decreased by

$$\boldsymbol{g}'_{\boldsymbol{B}(a)_3} = \boldsymbol{g}_{\boldsymbol{B}(a)_3} - \boldsymbol{C}(c), \qquad \boldsymbol{g}'_{-\boldsymbol{C}(c)} = \boldsymbol{g}_{-\boldsymbol{C}(c)} + \boldsymbol{C}(c).$$

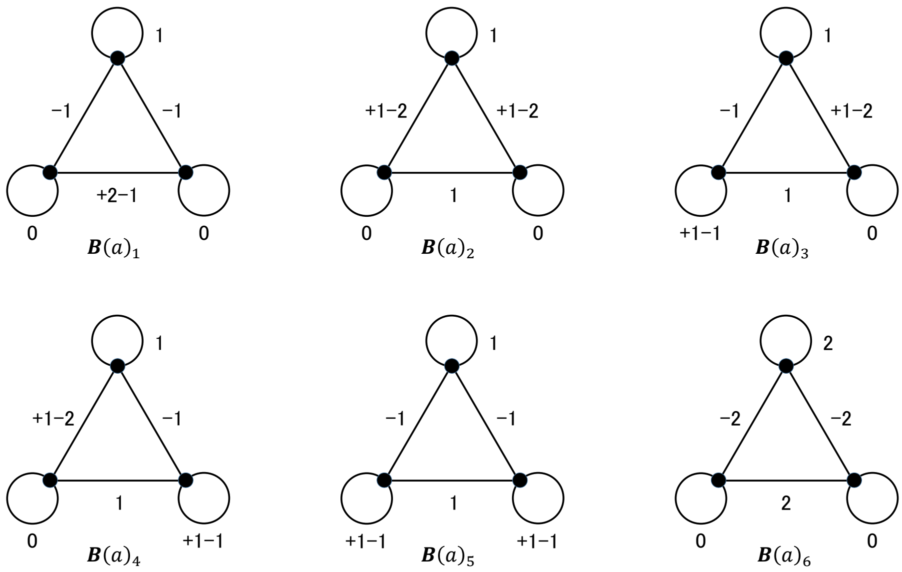


Figure 5: $\boldsymbol{B}(a)_1$, $\boldsymbol{B}(a)_2$, $\boldsymbol{B}(a)_3$, $\boldsymbol{B}(a)_4$, $\boldsymbol{B}(a)_5$, $\boldsymbol{B}(a)_6$

Case 4-4 $\boldsymbol{B}(a)$, $-\boldsymbol{C}(c)$, $\boldsymbol{C}(a)$, $-\boldsymbol{C}(b)$ and $\boldsymbol{B}(a)_4$ exist.

Because of the symmetry of b and c, we can decrease S as in Case 4-3.

Case 4-5 $\boldsymbol{B}(a)$, $-\boldsymbol{C}(c)$, $\boldsymbol{C}(a)$, $-\boldsymbol{C}(b)$ and $\boldsymbol{B}(a)_5$ exist.

S is decreased by

$$\boldsymbol{g}'_{\boldsymbol{B}(a)_5} = \boldsymbol{g}_{\boldsymbol{B}(a)_5} - \boldsymbol{C}(c), \qquad \boldsymbol{g}'_{-\boldsymbol{C}(c)} = \boldsymbol{g}_{-\boldsymbol{C}(c)} + \boldsymbol{C}(c).$$

Case 4-6 $\boldsymbol{B}(a)_6$, $-\boldsymbol{C}(c)$, $\boldsymbol{C}(a)$ and $-\boldsymbol{C}(b)$ exist.

S is decreased by

$$\boldsymbol{g}'_{\boldsymbol{B}(a)_6} = \boldsymbol{g}_{\boldsymbol{B}(a)_6} + \boldsymbol{C}(a), \qquad \boldsymbol{g}'_{\boldsymbol{C}(a)} = \boldsymbol{g}_{\boldsymbol{C}(a)} - \boldsymbol{C}(a).$$

Now we look at Case 5.

Case 5 $\boldsymbol{B}(a)$, $-\boldsymbol{C}(c)$, $\boldsymbol{C}(a)$ and $-\boldsymbol{B}(c)$ exist.

As in Case 4 $\deg(c) \geq 2$ by the existence of $\boldsymbol{C}(a)$. Then $\boldsymbol{g}_{-\boldsymbol{C}(d)}$ has additional edges connecting to c. The possible cases are, 1) cc alone, 2) at least one ac, or 3) 2 bc's. These three cases are depicted in Figure 6.

Case 5-1 $\boldsymbol{B}(a)$, $-\boldsymbol{C}(c)_1$, $\boldsymbol{C}(a)$ and $-\boldsymbol{B}(c)$ exist.

S is decreased by

$$\boldsymbol{g}'_{-C(c)_1} = \boldsymbol{g}_{-C(c)_1} - \boldsymbol{B}(c), \qquad \boldsymbol{g}'_{-B(c)} = \boldsymbol{g}_{-B(c)} + \boldsymbol{B}(c).$$

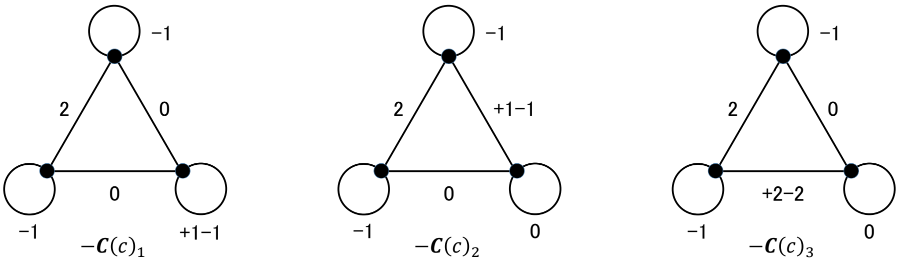


Figure 6: $-\boldsymbol{C}(c)_1$, $-\boldsymbol{C}(c)_2$, $-\boldsymbol{C}(c)_3$

Case 5-2 $\boldsymbol{B}(a)$, $-\boldsymbol{C}(c)_2$, $\boldsymbol{C}(a)$ and $-\boldsymbol{B}(c)$ exist.
S is decreased by

$$\boldsymbol{g}'_{-\boldsymbol{C}(c)_2} = \boldsymbol{g}_{-\boldsymbol{C}(c)_2} + \boldsymbol{B}(a), \qquad \boldsymbol{g}'_{\boldsymbol{B}(a)} = \boldsymbol{g}_{\boldsymbol{B}(a)} - \boldsymbol{B}(a).$$

Case 5-3 $\boldsymbol{B}(a)$, $-\boldsymbol{C}(c)_3$, $\boldsymbol{C}(a)$ and $-\boldsymbol{B}(c)$ exist.
S is decreased by

$$\boldsymbol{g}'_{-\boldsymbol{C}(c)_3} = \boldsymbol{g}_{-\boldsymbol{C}(c)_3} + \boldsymbol{C}(a), \qquad \boldsymbol{g}'_{\boldsymbol{C}(a)} = \boldsymbol{g}_{\boldsymbol{C}(a)} - \boldsymbol{C}(a).$$

We have eliminated patterns $\boldsymbol{A}$, $\boldsymbol{D}$ and $\boldsymbol{B}$. The remaining pattern is $\boldsymbol{C}$.

Case 6 $\boldsymbol{C}$ exists.
Suppose that $\boldsymbol{C}(a)$ exists. In the absence of $\pm\boldsymbol{A}$, $\pm\boldsymbol{D}$ and $\pm\boldsymbol{B}$ and the pair $(\boldsymbol{C}(a), -\boldsymbol{C}(a))$, the excess of bc can not be canceled. Hence this case is impossible.

We have now eliminated all the patterns. We now review the moves we needed to decrease S. Except for Case 3-1, all the moves were exchanges of edges between two graphs, which correspond to moves of degree two. In Case 3-1 we needed a move of degree three. Hence $\max_{\boldsymbol{b}\in\mathbb{N}A} \mathrm{MD}(A_{\boldsymbol{b}}) \le 3$. Together with (13) we have $\max_{\boldsymbol{b}\in\mathbb{N}A} \mathrm{MD}(A_{\boldsymbol{b}}) = 3$.

II. Proof of $\mathrm{MC}(A) = 5$.

Next we show $\mathrm{MC}(A) = 5$. As discussed above, our argument before Case 3-2 applies also to the higher Lawrence lifting $A^{(N)}$. $\boldsymbol{b}$'s can be different in different layers in (15). Therefore we need to check Case 4 and Case 5 again for higher Lawrence lifting. The argument is actually simple. In Case 4, we consider at most five patterns (at most five graphs) which consist of two $\boldsymbol{B}(a)$'s, $-\boldsymbol{C}(c)$, $\boldsymbol{C}(a)$ and $-\boldsymbol{C}(b)$, whose sum is the zero vector. This shows that a move of type at most five decreases S in the Case 4 for $A^{(N)}$. In Case 5 we consider at most four patterns (at most four graphs), whose sum is the zero vector. Hence a move of type at most four decreases S in the Case 5 for $A^{(N)}$. This proves $\mathrm{MC}(A) \le 5$.

To establish the equality, we construct an indispensable move whose type is five. Let $\boldsymbol{g}_1, \dots, \boldsymbol{g}_5$ be graphs displayed in the upper row and let $\hat{\boldsymbol{g}}_1, \dots, \hat{\boldsymbol{g}}_5$ be graphs displayed in the lower row of Figure 7. We show that

$$\boldsymbol{z} = (\boldsymbol{z}_1, \dots, \boldsymbol{z}_5) = (\boldsymbol{g}_1, \dots, \boldsymbol{g}_5) - (\hat{\boldsymbol{g}}_1, \dots, \hat{\boldsymbol{g}}_5)$$

is an indispensable move $A^{(5)}$ by Proposition 1. First, z_i, $i = 1, \ldots, 5$, are patterns $\boldsymbol{B}$ or $\boldsymbol{C}$ and they are indispensable moves for A. By the argument after Proposition 1 we can start from arbitrary slice z_k.

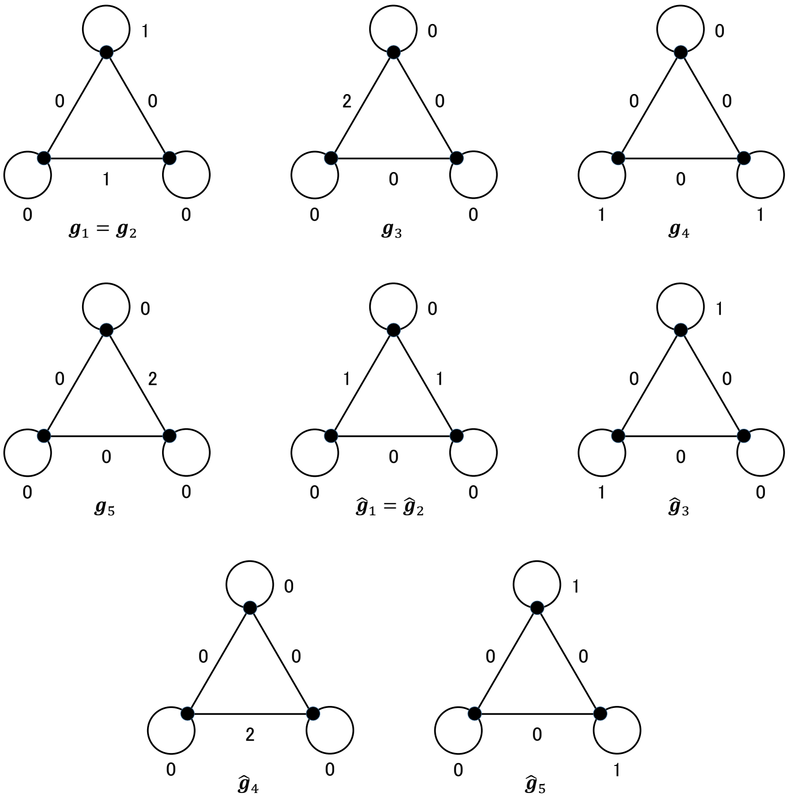


Figure 7: Graphs $\boldsymbol{g}_1 = \boldsymbol{g}_2, \boldsymbol{g}_3, \boldsymbol{g}_4, \boldsymbol{g}_5$ and $\hat{\boldsymbol{g}}_1 = \hat{\boldsymbol{g}}_2, \hat{\boldsymbol{g}}_3, \hat{\boldsymbol{g}}_4, \hat{\boldsymbol{g}}_5$

We start with z_3. Since edges aa, bb in z_3 have to be canceled, we need z_1 and z_4. Since the edge bc in z_4 has to be canceled, we need z_2. Also, since the edge ac in z_1 has to be canceled we need z_5. Hence we need all slices and this proves that z is indispensable.

III. Proof of $\mathrm{MD}(A_{\boldsymbol{b}}) = 2$ for $\boldsymbol{b} \neq (2, 2, 2)$.

Recall that only Case 3-1 needed a degree-three move. We show that this move is not needed if $\boldsymbol{b} \neq (2, 2, 2)$, by a series of lemmas.

We write elements of the Graver basis by their positive part and their negative part, e.g., $\boldsymbol{A} = \boldsymbol{A}^+ - \boldsymbol{A}^-$. We only need to consider the condition on $\boldsymbol{b}$ such that we need degree-three moves to decrease S for the case

$$S = \sum_{i=1}^{3} |\boldsymbol{g}_i - \hat{\boldsymbol{g}}_i|, \quad \hat{\boldsymbol{g}}_1 = \boldsymbol{g}_1 - \boldsymbol{B}(a),\ \hat{\boldsymbol{g}}_2 = \boldsymbol{g}_2 + \boldsymbol{C}(c),\ \hat{\boldsymbol{g}}_3 = \boldsymbol{g}_3 - \boldsymbol{B}(b) \tag{17}$$

and $\boldsymbol{b} = A\boldsymbol{g}_i = A\hat{\boldsymbol{g}}_i$, $i = 1, 2, 3$. Note there is the symmetry of vertices a and b.

Lemma 3. *If degree-two moves do not decrease S in* (17)*, then*

$$\boldsymbol{e}_{cc} \le \boldsymbol{g}_2 - \boldsymbol{C}(c)^-, \quad \boldsymbol{e}_{bc} \nleq \boldsymbol{g}_2 - \boldsymbol{C}(c)^-, \quad \boldsymbol{e}_{ca} \nleq \boldsymbol{g}_2 - \boldsymbol{C}(c)^-.$$

Proof. Since the degree of vertex c of $\boldsymbol{C}^-$ is less than that of $\boldsymbol{B}(a)^+$ by one, the degree of vertex c of $\boldsymbol{g}_2 - \boldsymbol{C}(c)^-$ is greater than one. Then $\boldsymbol{e}_{cc} \le \boldsymbol{g}_2 - \boldsymbol{C}(c)^-$, $\boldsymbol{e}_{bc} \le \boldsymbol{g}_2 - \boldsymbol{C}(c)^-$, or $\boldsymbol{e}_{ca} \le \boldsymbol{g}_2 - \boldsymbol{C}(c)^-$. If $\boldsymbol{e}_{bc} \le \boldsymbol{g}_2 - \boldsymbol{C}(c)^-$, then S is decreased by the following exchange of edges:

$$\boldsymbol{g}_2' = \boldsymbol{g}_2 + \boldsymbol{B}(b), \quad \boldsymbol{g}_3' = \boldsymbol{g}_3 - \boldsymbol{B}(b).$$

Hence $\boldsymbol{e}_{bc} \nleq \boldsymbol{g}_2 - \boldsymbol{C}(c)^-$. We also have $\boldsymbol{e}_{ca} \nleq \boldsymbol{g}_2 - \boldsymbol{C}(c)^-$ by the symmetry between a and b.

Lemma 4. *If degree-two moves do not decrease S in* (17)*, then*

$$\boldsymbol{e}_{bc} \le \boldsymbol{g}_1 - \boldsymbol{B}(a)^+, \quad \boldsymbol{e}_{bb} \nleq \boldsymbol{g}_1 - \boldsymbol{B}(a)^+, \quad \boldsymbol{e}_{ab} \nleq \boldsymbol{g}_1 - \boldsymbol{B}(a)^+.$$

Proof. Since the degree of vertex b of $\boldsymbol{B}(a)^+$ is less than that of $\boldsymbol{C}(a)^-$ by one, the degree of vertex b of $\boldsymbol{g}_1 - \boldsymbol{B}(a)^+$ is greater than one. Then $\boldsymbol{e}_{bc} \le \boldsymbol{g}_1 - \boldsymbol{B}(a)^+$, $\boldsymbol{e}_{bb} \le \boldsymbol{g}_1 - \boldsymbol{B}(a)^+$, or $\boldsymbol{e}_{ab} \le \boldsymbol{g}_1 - \boldsymbol{B}(a)^+$. If $\boldsymbol{e}_{bb} \le \boldsymbol{g}_1 - \boldsymbol{B}(a)^+$, then S is decreased by the following exchange of edges:

$$\boldsymbol{g}_1' = \boldsymbol{g}_1 - \boldsymbol{C}(c), \quad \boldsymbol{g}_2' = \boldsymbol{g}_2 + \boldsymbol{C}(c).$$

Similarly if $\boldsymbol{e}_{ab} \le \boldsymbol{g}_1 - \boldsymbol{B}(a)^+$, S is decreased by the following exchange of edges:

$$\boldsymbol{g}_1' = \boldsymbol{g}_1 + \boldsymbol{B}(b), \quad \boldsymbol{g}_3' = \boldsymbol{g}_3 - \boldsymbol{B}(b).$$

By the symmetry of a and b, the following lemma also holds.

Lemma 5. *If degree-two moves do not decrease S in* (17)*, then*

$$\boldsymbol{e}_{ca} \le \boldsymbol{g}_3 - \boldsymbol{B}(b)^+, \quad \boldsymbol{e}_{aa} \nleq \boldsymbol{g}_3 - \boldsymbol{B}(b)^+, \quad \boldsymbol{e}_{ab} \nleq \boldsymbol{g}_3 - \boldsymbol{B}(b)^+.$$

Lemma 6. *Suppose that degree-two moves do not decrease S in* (17) *and* $\deg(a) \ge 3$ *or* $\deg(b) \ge 3$. *Then* $\deg(c) \ge 3$.

Proof. By symmetry let $\deg(a) \ge 3$. By Lemma 5, in this case, $\boldsymbol{2e}_{ca} \le \boldsymbol{g}_3 - \boldsymbol{B}(b)^+$. Hence $\deg(c) \ge 3$.

By this lemma we can assume that $\deg(c) \ge 3$ if $\boldsymbol{b} \ne (2,2,2)$. Hence our proof is completed by the following lemma.

Lemma 7. *If* $\deg(c) \ge 3$*, then S in* (17) *can be decreased by degree-two moves.*

Proof. By Lemma 3, if $\deg(c) \ge 3$, then $\boldsymbol{2e}_{cc} \le \boldsymbol{g}_3 - \boldsymbol{B}(b)^+$. Then the following series of exchanges of edges decreases S:

$$\begin{aligned}
\boldsymbol{g}_2' &= \boldsymbol{g}_2 - \boldsymbol{D}(a), \quad \boldsymbol{g}_3' = \boldsymbol{g}_3 + \boldsymbol{D}(a),\\
\boldsymbol{g}_1' &= \boldsymbol{g}_1 + \boldsymbol{C}(a), \quad \boldsymbol{g}_2' = \boldsymbol{g}_2 - \boldsymbol{C}(a),\\
\boldsymbol{g}_2' &= \boldsymbol{g}_2 + \boldsymbol{B}(c), \quad \boldsymbol{g}_3' = \boldsymbol{g}_3 - \boldsymbol{B}(c),\\
\boldsymbol{g}_1' &= \boldsymbol{g}_1 - \boldsymbol{A}, \quad \boldsymbol{g}_2' = \boldsymbol{g}_2 + \boldsymbol{A}.
\end{aligned}$$

4. Complete bipartite graphs as base configurations

In this section we take incidence matrices $A(I, J)$ of complete bipartite graphs $K_{I,J}$ as base configurations and study the maximum Markov degree of the configurations defined by their fibers. The fibers correspond to two-way transportation polytopes. In algebraic statistics, $A(I, J)$ is the design matrix specifying the row sums and the column sums of an $I\times J$ two-way contingency table and the N-th Lawrence lifting $A(I, J)^{(N)}$ is the design matrix for no-three-factor interaction model for $I \times J \times N$ three-way contingency tables.

A remarkable fact for the case of complete bipartite graphs is that the maximum Markov degree is three irrespective of I and J as we show in Section 4.1. On the other hand the Markov complexity grows with I and J. Lower bound for the Graver complexity has been obtained by [4, 13]. In Section 4.2 we give a lower bound for the Markov complexity, which appears on the right-hand side of (2) in our main theorem.

4.1. Markov degree for two-way transportation polytopes

In this section we prove that the Markov degree of configurations for two-way transportation polytopes is at most three. As discussed in Section 1, recently this fact was proved by Domokos and Joó [7] in a more general setting. However in this section we give a proof, which is a direct extension of a proof in [19].

Let $\boldsymbol{r} \in \mathbb{N}^I$ and $\boldsymbol{c} \in \mathbb{N}^J$ be two non-negative integer vectors with $\sum_{i=1}^{I} r_i = \sum_{j=1}^{J} c_j$. The two-way transportation polytope is the set of all non-negative matrices $\boldsymbol{x} = (x_{ij})$ whose row sum vector is $\boldsymbol{r}$ and column sum vector is $\boldsymbol{c}$. Let $T_{\boldsymbol{r},\boldsymbol{c}}$ be the set of integral matrices in the transportation polytope. Then

$$T_{\boldsymbol{r},\boldsymbol{c}} = \mathcal{F}_{A(I,J),(\boldsymbol{r},\boldsymbol{c})}$$

is the the $(\boldsymbol{r}, \boldsymbol{c})$-fiber for the incidence matrix $A(I, J)$ of the complete bipartite graph $K_{I,J}$. We regard an element in $T_{\boldsymbol{r},\boldsymbol{c}}$ as complete bipartite graph with non-negative integral weights on edges, which is denoted by $\boldsymbol{g} = (g(ij) \mid (i, j) \in [I] \times [J])$. Set $\boldsymbol{e} = (e_{ij}) \in \mathbb{N}A(I, J)_{(\boldsymbol{r},\boldsymbol{c})}$ arbitrarily. Then an element of the corresponding fiber $\mathcal{F}_{A(I,J)_{(\boldsymbol{r},\boldsymbol{c})},\boldsymbol{e}}$ can be identified with some multiset $\{\boldsymbol{g}_1, \dots, \boldsymbol{g}_N\}$ satisfying $\boldsymbol{g}_k \in T_{\boldsymbol{r},\boldsymbol{c}}, k = 1, \dots, N$, and $\sum_k g_k(ij) = e_{ij}, (i, j) \in [I] \times [J]$. Haase and Paffenholz [9] studied the 3×3 transportation polytopes. When $I = J$ and $\boldsymbol{r} = \boldsymbol{c} = (1, \dots, 1)^\top$, the corresponding transportation polytope is the Birkhoff polytope.

Theorem 4. *The toric ideal associated with the transportation polytope is generated by binomials of degree two and three, i.e.,* $\max_{(\boldsymbol{r},\boldsymbol{c})\in\mathbb{N}A(I,J)} \mathrm{MD}(A(I, J)_{(\boldsymbol{r},\boldsymbol{c})}) = 3$.

The rest of this subsection is devoted to the proof of Theorem 4. Our proof is a direct extension of the proof for the Birkhoff polytope in [19]. We modify the terminologies in [19] to be suitable for our setting.

Definition 1. *An $I \times J$ integer matrix $\boldsymbol{g} = (g(ij))$ is a* proper graph *if $\boldsymbol{g}$ is an element of $T_{\boldsymbol{r},\boldsymbol{c}}$. A multiset $\{\boldsymbol{g}_1, \dots, \boldsymbol{g}_N\}$ is* proper *if each $\boldsymbol{g}_k, k = 1, \dots, N$, is a proper graph.*

For two proper graphs $\boldsymbol{g}$ and $\hat{\boldsymbol{g}}$, we call $D_{\boldsymbol{g},\hat{\boldsymbol{g}}} := \sum_{i,j} |g(ij) - \hat{g}(ij)|$ the *size of differences*.

Definition 2. *An $I \times J$ integer matrix $\boldsymbol{g} = (g(ij))$ is an* improper graph *if $\boldsymbol{g}$ has the row sum $\boldsymbol{r}$ and column sum $\boldsymbol{c}$, and there exists a unique edge $(i^*, j^*) \in [I] \times [J]$ such that*

$$g(i^* j^*) = -1, \quad g(ij) \geq 0, \ \forall (i, j) \neq (i^*, j^*).$$

We call $g(i^* j^*)$ an improper edge *of $\boldsymbol{g}$. A multiset $\{\boldsymbol{g}_1, \dots, \boldsymbol{g}_N\}$ is* improper *if one of $\{\boldsymbol{g}_1, \dots, \boldsymbol{g}_N\}$ is an improper graph, the others are proper graphs, and $\sum_{k=1}^N g_k(ij) \geq 0, \forall i, j$.*

Definition 3. *An $I \times J$ integer matrix $\boldsymbol{g} = (g(ij))$ is a* graph with collision *if $g(ij) \geq 0, \forall i, j$, the column sum of $\boldsymbol{g}$ is $\boldsymbol{c}$ and there exists $i^* \in [I]$ such that*

$$\sum_{j=1}^J g(i^* j) = r_{i^*} + 1, \quad \sum_{j=1}^J g(ij) \leq r_i + 1, \ \forall i \neq i^*.$$

In this case we also say that the graph $\boldsymbol{g}$ contains a collision *or the vertex i^** collides *in $\boldsymbol{g}$.*

We often denote a multiset $\{\boldsymbol{g}_1, \dots, \boldsymbol{g}_N\}$ of $I \times J$ integer matrices by $\mathcal{S}$ if $\sum_{k=1}^N g_k(ij) \geq 0, \forall i, j$, and each element $\boldsymbol{g}_k, \forall k$, is one of the graphs defined in Definitions 1–3. The multiset $\mathcal{S}$ is denoted by $\mathcal{P}$ (resp. $\mathcal{I}$) when $\mathcal{S}$ is proper (resp. improper) and we want to emphasize it.

We now introduce some operations. Let $\mathcal{S} = \{\boldsymbol{g}_1, \dots, \boldsymbol{g}_N\}$ be a multiset of graphs in Definitions 1–3. Consider a pair of distinct graphs in $\mathcal{S}$, say $\boldsymbol{g}_{k_1} = (g_{k_1}(ij))$ and $\boldsymbol{g}_{k_2} = (g_{k_2}(ij))$. Fix $i_1, i_2 \in [I]$ and $j^* \in [J]$ arbitrarily and set the two matrices $z_{k_1} = (z_{k_1}(ij))$ and $z_{k_2} = (z_{k_2}(ij))$ as

$$z_{k_1}(ij) = \begin{cases} +1, & (i, j) = (i_2, j^*), \\ -1, & (i, j) = (i_1, j^*), \\ 0, & \text{otherwise}, \end{cases} \qquad z_{k_2}(ij) = \begin{cases} +1, & (i, j) = (i_1, j^*), \\ -1, & (i, j) = (i_2, j^*), \\ 0, & \text{otherwise}. \end{cases}$$

The *swap* $\{k_1, k_2\} : i_1 \overset{j^*}{\leftrightarrow} i_2$ for $\mathcal{S}$ is an operation transforming $\mathcal{S}$ into another multiset $\mathcal{S}'$ of matrices defined by

$$\mathcal{S}' = (\mathcal{S} \setminus \{\boldsymbol{g}_{k_1}, \boldsymbol{g}_{k_2}\}) \cup \{\boldsymbol{g}_{k_1} + z_{k_1}, \boldsymbol{g}_{k_2} + z_{k_2}\}.$$

Note that the resulting $\mathcal{S}'$ has the same sums of weights of each edge as the original $\mathcal{S}$, although the elements $\mathcal{S}'$ may not be graphs in Definitions 1–3.

Let us consider $n \in \mathbb{N}$ swaps on the same pair of graphs $g_{k_1}, g_{k_2} \in \mathcal{S}$ and denote them as

$$(\boldsymbol{z}_{k_1}^{(1)}, \boldsymbol{z}_{k_2}^{(1)}), \dots, (\boldsymbol{z}_{k_1}^{(n)}, \boldsymbol{z}_{k_2}^{(n)}).$$

Consider the following operation, which transforms a multiset $\mathcal{S}$ into another multiset $\mathcal{S}'$ without changing sums of weights of each edge:

$$\mathcal{S}' = (\mathcal{S} \setminus \{\boldsymbol{g}_{k_1}, \boldsymbol{g}_{k_2}\}) \cup \{\boldsymbol{g}_{k_1} + \sum_{l=1}^n \boldsymbol{z}_{k_1}^{(l)}, \boldsymbol{g}_{k_2} + \sum_{l=1}^n \boldsymbol{z}_{k_2}^{(l)}\}.$$

We call this operation a *swap operation among two graphs of* $\mathcal{S}$ and denote it as $\mathcal{S} \overset{\{k_1,k_2\}}{\longleftrightarrow} \mathcal{S}'$ or merely $\mathcal{S} \longleftrightarrow \mathcal{S}'$. If both of $\mathcal{S}$ and $\mathcal{S}'$ are proper, the operation is nothing but the move of degree two.

Lemma 8. *Let $\mathcal{S} = \{\boldsymbol{g}_1, \ldots, \boldsymbol{g}_N\}$ be a multiset of graphs without any improper edge and suppose that the kth and the k'th graphs contain some collisions. If $\sum_{j=1}^{J}(g_k(ij) + g_{k'}(ij)) = 2r_i$ for each $i \in [I]$, we can resolve all the collisions by a swap operation among these two graphs.*

Proof. We may assume $g_k(ij) = 0$ or $g_{k'}(ij) = 0$ for each $i \in [I], j \in [J]$. Let $\bar{\boldsymbol{g}}_k = (\bar{g}_k(j))$ and $\bar{\boldsymbol{g}}_{k'} = (\bar{g}_{k'}(j))$ be the J-dimensional row vectors whose jth elements $\bar{g}_k(j)$ and $\bar{g}_{k'}(j)$ are the multisets of symbols defined by

$$\bar{g}_k(j) := \{\underbrace{1, \ldots, 1}_{g_k(1j)}, \ldots, \underbrace{I, \ldots, I}_{g_k(Ij)}\}, \quad \bar{g}_{k'}(j) := \{\underbrace{1, \ldots, 1}_{g_{k'}(1j)}, \ldots, \underbrace{I, \ldots, I}_{g_{k'}(Ij)}\}, \qquad j \in [J].$$

Suppose that the vertex $i \in [I]$ collides in $\boldsymbol{g}_k$. This means that the symbol i appears $(r_i + 1)$ times in $\bar{\boldsymbol{g}}_k$ and $(r_i - 1)$ times in $\bar{\boldsymbol{g}}_{k'}$. To resolve the collision of i, we temporarily assign the different labels to vertices as follows. First, we assign $i_1, \ldots, i_{r_i-1}$ to $(r_i - 1)$ i's in each of $\bar{\boldsymbol{g}}_k$ and $\bar{\boldsymbol{g}}_{k'}$. Second, for each vertex not colliding in these two graphs, say i', we assign $i'_1, \ldots, i'_{r_{i'}}$ to $r_{i'}$ i''s in each of $\bar{\boldsymbol{g}}_k$ and $\bar{\boldsymbol{g}}_{k'}$. Finally, for each colliding vertex different from i, say i'', we assign $i''_1, \ldots, i''_{r_{i''}-1}$ to $(r_{i''} - 1)$ i'''s in each graph and $\hat{i}''_1, \hat{i}''_2$ to the remaining two i'''s. At this point, each symbol except i appears once in each of $\bar{\boldsymbol{g}}_k$ and $\bar{\boldsymbol{g}}_{k'}$.

Let $s = \sum_{j=1}^{J} c_j$ and define the $2 \times s$ matrix $D = (d_{K\alpha})$ satisfying the following equations as multisets:

$$\{d_{K\alpha}, \ldots, d_{K(\alpha+c_j-1)}\} = \bar{g}_K(j), \quad K = k, k',\ \alpha = \sum_{m=1}^{j-1} c_m + 1,\ j = 1, \ldots, J.$$

Let G be a graph on the vertex set $[s]$ defined as follows: a directed edge (α,β) exists for $\alpha,\beta \in [s]$ if and only if $d_{k\beta} = d_{k'\alpha}$. The graph G consists of disjoint directed paths and cycles. Then, there exists a path starting from a vertex $\gamma \in [s]$ with $d_{k\gamma} = i$. This path defines the swap operation among $\boldsymbol{g}_k$ and $\boldsymbol{g}_{k'}$ with the original labels of vertices, which resolves the collision of i without causing any new collision. Repeatedly applying this discussion, we obtain the sequence of the swap operations resolving collisions among the two graphs. Combining them into one swap operation, we obtain the desired swap operation among $\boldsymbol{g}_k$ and $\boldsymbol{g}_{k'}$.

Lemma 9. *Let $\mathcal{I} = \{\boldsymbol{g}_1, \ldots, \boldsymbol{g}_N\}$ be an improper multiset with $g_k(ij) = -1$. Then, by a swap operation among two graphs, $\mathcal{I}$ can be transformed to a proper multiset.*

Proof. Choose $i' \in [I]$ with $i' \neq i$ and $g_k(i'j) > 0$. Since $\sum_{l=1}^{N} g_l(ij) \geq 0$, there exists $k' \in [N]$ with $g_{k'}(ij) > 0$. Perform a swap $\{k, k'\} : i \overset{j}{\leftrightarrow} i'$ to resolve the improper element. Then, i collides in $\boldsymbol{g}_{k'}$ and i' collides in $\boldsymbol{g}_k$. Since i (resp. i') appears $2r_i$ (resp. $2r_{i'}$) times in the first two graphs in total, we can resolve these collisions by Lemma 8 by a swap operation among these two graph. Combining the process, we obtain a swap operation among two graphs transforming $\mathcal{I}$ to a proper multiset.

Definition 4. *We call the pair of two graphs $\boldsymbol{g}_k$ and $\boldsymbol{g}_{k'}$ in Lemma 9 a* resolvable pair *and denote it as* $[k_{\mathrm{im}}, k_{\mathrm{pr}}]$.

Definition 5. *A swap operation among two graphs labeled by* $A = \{k, k'\}$ *in* $\mathcal{I} \overset{A}{\longleftrightarrow} \mathcal{I}'$ *is* compatible *with improper multisets* $\mathcal{I}$ *and* $\mathcal{I}'$ *if there exists a common resolvable pair* $[k_{\mathrm{im}}, k_{\mathrm{pr}}]$ *of* $\mathcal{I}$ *and* $\mathcal{I}'$ *such that* $A \cap \{k_{\mathrm{im}}, k_{\mathrm{pr}}\} \neq \emptyset$.

Lemma 10. *Let* $\mathcal{P} = \{\boldsymbol{g}_1, \dots, \boldsymbol{g}_N\}, \hat{\mathcal{P}} = \{\hat{\boldsymbol{g}}_1, \dots, \hat{\boldsymbol{g}}_N\}$ *be two proper multisets in* $\mathcal{F}_{A(I,J)_{(r,c)}, \boldsymbol{e}}$ *and suppose* $\boldsymbol{g}_k \neq \hat{\boldsymbol{g}}_{k'}$ *for some* k, k'. *Then, their size* $D_{\boldsymbol{g}_k, \hat{\boldsymbol{g}}_{k'}}$ *of differences can be decreased by a swap operation among two graphs of* $\mathcal{P}$, *such that if the resulting multiset is not proper, it is improper and its improper graph and the kth graph form a resolvable pair.*

Proof. Since $\boldsymbol{g}_k \neq \hat{\boldsymbol{g}}_{k'}$, there exist $i \in [I]$ and $j, j' \in [J]$ satisfying $g_k(ij) < \hat{g}_{k'}(ij)$ and $g_k(ij') > \hat{g}_{k'}(ij')$. Since $\mathcal{P}$ and $\mathcal{P}'$ belong to $\mathcal{F}_{A(I,J)_{(r,c)}, \boldsymbol{e}}$, there exists $k'' \in [N]$ with $k'' \neq k$ and $g_{k''}(ij) > 0$. Choose $i' \in [I]$ satisfying $i' \neq i$ and $g_k(i'j) > 0$ and consider a swap operation $\{k, k''\} : i' \overset{j}{\leftrightarrow} i \overset{j'}{\leftrightarrow} i'$ to $\mathcal{P}$. This operation decreases $D_{\boldsymbol{g}_k, \hat{\boldsymbol{g}}_{k'}}$. When $g_{k''}(i'j') > 0$, the resulting multiset is proper. Otherwise, the resulting multiset is improper where $[k'', k]$ forms a resolvable pair. This proves the claim.

Lemma 11. *Let* $\mathcal{I} = \{\boldsymbol{g}_1, \dots, \boldsymbol{g}_N\}$ *be an improper multiset and* $\hat{\mathcal{P}} = \{\hat{\boldsymbol{g}}_1, \dots, \hat{\boldsymbol{g}}_N\}$ *be a proper multiset with the same sums* $\boldsymbol{e}$ *of weights of edges. Consider the* k'*th graph* $\hat{\boldsymbol{g}}_{k'}$ *of* $\hat{\mathcal{P}}$ *and choose any resolvable pair* $[k_{\mathrm{im}}, k_{\mathrm{pr}}]$ *of* $\mathcal{I}$. *Then, by at most two swap operations among two graphs of* $\mathcal{I}$, *we can (i) decrease the size* $D_{\boldsymbol{g}_{k_{\mathrm{pr}}}, \hat{\boldsymbol{g}}_{k'}}$ *of differences, or (ii) make* $\mathcal{I}$ *proper without changing* $\boldsymbol{g}_{k_{\mathrm{pr}}}$. *Furthermore, if the resulting multiset is not proper, then it is an improper multiset with a resolvable pair consisting of its improper graph and the* k_{pr}*th graph, and each intermediate swap operation between two consecutive improper multisets is compatible with them.*

Proof. We may suppose $g_{k_{\mathrm{im}}}(ij) = -1$ and $g_{k_{\mathrm{pr}}}(ij) > 0$ for some $i \in [I]$ and $j \in [J]$. In the cases below, when a resulting multiset is improper, $[k_{\mathrm{im}}, k_{\mathrm{pr}}]$ will be a resolvable pair.

Case 1 $\hat{g}_{k'}(ij) \geq g_{k_{\mathrm{pr}}}(ij)$.

Since $\sum_{l=1}^{N} g_l(ij) = \sum_{l=1}^{N} \hat{g}_l(ij) > g_{k_{\mathrm{pr}}}(ij)$, there exists $k \in [N]$ such that $k \neq k_{\mathrm{im}}, k_{\mathrm{pr}}$ and $g_k(ij) > 0$. Then, $[k_{\mathrm{im}}, k]$ is a resolvable pair and $\mathcal{I}$ can be transformed to a proper multiset without changing $\boldsymbol{g}_{k_{\mathrm{pr}}}$ by Lemma 9. This corresponds to (ii) of the lemma and summarized as $\mathcal{I} \overset{\{k_{\mathrm{im}}, k\}}{\longleftrightarrow} \mathcal{P}$.

Case 2 $\hat{g}_{k'}(ij) < g_{k_{\mathrm{pr}}}(ij)$.

Since $\sum_{t=1}^{J} \hat{g}_{k'}(it) = \sum_{t=1}^{J} \hat{g}_{k_{\mathrm{pr}}}(it)$, there exists $j' \in [J]$ with $\hat{g}_{k'}(ij') > g_{k_{\mathrm{pr}}}(ij')$. Fix some $i' \in [I]$ with $g_{k_{\mathrm{pr}}}(i'j') > \hat{g}_{k'}(i'j')$ arbitrarily.

Case 2-1 $g_{k_{\mathrm{im}}}(ij') > 0$.

We perform the swap operations $\{k_{\mathrm{pr}}, k_{\mathrm{im}}\} : i \overset{j}{\leftrightarrow} i'$ and $\{k_{\mathrm{im}}, k_{\mathrm{pr}}\} : i' \overset{j'}{\leftrightarrow} i$ to $\mathcal{I}$ at the same time, which decrease $D_{\boldsymbol{g}_{k_{\mathrm{pr}}}, \hat{\boldsymbol{g}}_{k'}}$. If $g_{k_{\mathrm{im}}}(i'j) > 0$, the resulting multiset is proper. Otherwise, the resulting multiset is improper. This corresponds to (i) of the lemma and is summarized as $\mathcal{I} \overset{\{k_{\mathrm{pr}}, k_{\mathrm{im}}\}}{\longleftrightarrow} \mathcal{P}$ or $\mathcal{I} \overset{\{k_{\mathrm{pr}}, k_{\mathrm{im}}\}}{\longleftrightarrow} \mathcal{I}$.

Case 2-2 $g_{k_{\mathrm{im}}}(ij') = 0$.
Since $\hat{g}_{k'}(ij') > g_{k_{\mathrm{pr}}}(ij')$, there exists $k'' \in [N]$ such that $k'' \neq k_{\mathrm{pr}}, k_{\mathrm{im}}$ and $\hat{g}_{k'}(ij') > 0$. Fix $i'' \in [I]$ with $x_{i''k_{\mathrm{im}}} > 0$ arbitrarily. Consider the swap $\{k'', k_{\mathrm{im}}\} : i \overset{j'}{\leftrightarrow} i''$. Then, i collides in the k_{im}th graph and i'' collides in the k''th graph. By the similar argument as the proof of Lemma 8, we can resolve these collisions by a swap operation among the k_{im}th and the k''th graphs, which leaves $g_{k_{\mathrm{im}}}(ij) = -1$ and makes $g_{k_{\mathrm{im}}}(ij')$ positive. These operation can be done by a singe swap operation among the two graphs. After that, this case reduces to Case 2-1. Together with the subsequent operation of Case 2-1, Case 2-2 is summarized as $\mathcal{I} \overset{\{k_{\mathrm{im}},k''\}}{\longleftrightarrow} \mathcal{I} \overset{\{k_{\mathrm{pr}},k_{\mathrm{im}}\}}{\longleftrightarrow} \mathcal{P}$ or $\mathcal{I} \overset{\{k_{\mathrm{im}},k''\}}{\longleftrightarrow} \mathcal{I} \overset{\{k_{\mathrm{pr}},k_{\mathrm{im}}\}}{\longleftrightarrow} \mathcal{I}$.

We now give a proof of Theorem 4 by the similar argument as [19]. Let $\mathcal{P}$ and $\hat{\mathcal{P}}$ be two proper multisets belonging to the same fiber $\mathcal{F}_{A(I,J)_{(r,c)},\boldsymbol{e}}$. Choose any kth graph $\boldsymbol{g}_k$ of $\mathcal{P}$ and any k'th graph $\hat{\boldsymbol{g}}_{k'}$ of $\hat{\mathcal{P}}$ with $\boldsymbol{g}_k \neq \hat{\boldsymbol{g}}_{k'}$. Thanks to Lemmas 10 and 11, allowing some intermediate improper multisets, we can make $\boldsymbol{g}_k$ identical with $\hat{\boldsymbol{g}}_{k'}$ by a sequence of swap operations among two graphs of $\mathcal{P}$. We throw away this common graph from the two multisets and repeat the procedure. In the end, $\mathcal{P}$ can be fully transformed to $\hat{\mathcal{P}}$. Let us decompose the whole process of transforming $\mathcal{P}$ to $\hat{\mathcal{P}}$ into segments that consist of transformations from a proper multiset to another proper multiset with improper intermediate steps. One segment is depicted as $\mathcal{P}_1 \longleftrightarrow \mathcal{I}_1 \longleftrightarrow \cdots \longleftrightarrow \mathcal{I}_m \longleftrightarrow \mathcal{P}_m$ where each $\longleftrightarrow$ denotes a swap operation among two graphs in Lemmas 10 or 11. Then, for any consecutive multisets $\mathcal{I}_l$ and $\mathcal{I}_{l+1}, l = 1, \ldots, m-1$, there exist proper multisets $\mathcal{P}_l, \mathcal{P}'_l, l = 1, \ldots, m-1$, satisfying

$$\begin{aligned}
&\mathcal{P}_l \longleftrightarrow \mathcal{I}_l \longleftrightarrow \mathcal{I}_{l+1} \longleftrightarrow \mathcal{P}'_{l+1}, \\
&\mathcal{P}'_l \longleftrightarrow \mathcal{I}_l \longleftrightarrow \mathcal{P}_l.
\end{aligned}$$

By the compatibility of the swap operation in $\mathcal{I}_l \longleftrightarrow \mathcal{I}_{l+1}$, $\mathcal{P}_l$ can be transformed to $\mathcal{P}'_{l+1}$ by a swap operation among three graphs. Since $\mathcal{P}'_l \longleftrightarrow \mathcal{I}_l$ and $\mathcal{I}_l \longleftrightarrow \mathcal{P}_l$ involve a common improper graph, $\mathcal{P}'_l$ can also be transformed to $\mathcal{P}_l$ by a swap operation among three graphs. Therefore, the process from $\mathcal{P}_1$ to $\mathcal{P}_m$ is realized by swap operations among three graphs as

$$\begin{tikzcd}
 & \mathcal{P}'_l \arrow[dr, leftrightarrow] & & \mathcal{P}_l \arrow[dl, leftrightarrow] \arrow[r, bend right=60] & \mathcal{P}'_{l+1} \arrow[dr, leftrightarrow] & & \mathcal{P}_{l+1} \arrow[dl, leftrightarrow] & \\
\cdots \arrow[rr] & & \mathcal{I}_l \arrow[rrr, leftrightarrow] & & & \mathcal{I}_{l+1} & & \cdots \arrow[ll]
\end{tikzcd}$$

This proves Theorem 4.

4.2. Lower bound of Markov complexity for complete bipartite graphs

In this section we give a lower bound for $\mathrm{MC}(A(I, J))$, $3 \leq I \leq J$.

Proposition 2. *For* $3 \leq I \leq J$,

$$\mathrm{MC}(A(I, J)) \geq (I-2)(J^2-1)/4 + J - 1. \tag{18}$$

For the rest of this subsection we give a proof of Proposition 2. Let $d = \lfloor J/2 \rfloor$. We display $I \times J$ two-dimensional slice as follows:

$$\begin{array}{c|ccccccc|}
 & 1 & \cdots & d & (r) & J-d+1 & \cdots & J \\
\hline
1 & z_{1,1} & \cdots & z_{1,d} & (z_{1,r}) & z_{1,J-d+1} & \cdots & z_{1,J} \\
\vdots & \vdots & & & & & & \vdots \\
I & z_{I,1} & \cdots & z_{I,d} & (z_{I,r}) & z_{I,J-d+1} & \cdots & z_{I,J} \\
\hline
\end{array}, \tag{19}$$

where, $r = d+1$ if J is odd and r does not exist if J is even. We define $z(i_1, i_2; j_1, , j_2)$ as the following $I \times J$ table:

$$\begin{array}{c|cc|}
 & j_1 & j_2 \\
\hline
i_1 & +1 & -1 \\
i_2 & -1 & +1 \\
\hline
\end{array}, \tag{20}$$

where other entries are 0.

We give an indispensable move $z^* = \{z^*(i,j,k)\}$ of for $A(I,J)^{(N)}$, where $N = (I-2)(J-d)d+2d$ is the type of z^*. The $I \times J$ slices of z^* as follows:

$$\begin{aligned}
&z(1, I; j, J-d+j), \quad j = 1, \cdots, d, \\
&z(I-1, I; J-d+j+1, j), \quad j = 1, \cdots, d-1, \\
&z(I-1, I; j+1, j), \quad j = d, (r), \\
&z(i, i+1; j+1, j) \times j, \quad i = 1, \ldots, I-2,\ j = 1, \ldots, d, \\
&z(i, i+1; r+1, r) \times d, \quad i = 1, \ldots I-2, \\
&z(i, i+1; J-j+1, J-j) \times j, \quad i = 1, \ldots, I-2,\ j = 1, \ldots, d-1.
\end{aligned}$$

It is easy checked that $\sum_{k=1}^{N} z^*(i,j,k) = 0$ for all i, j and z^* is a move for $A(I,J)^{(N)}$. Also all slices of z^* are indispensable. Therefore, if we can show that z^* is indispensable move, then

$$\mathrm{MC}(A(I,J)) \geq (I-2)(J-d)d + 2d \geq (I-2)(J^2-1)/4 + J - 1.$$

Now we again use the argument after Proposition 1. We start with the slice $z(1, I; 1, J-d+1)$. Since the (sum of) $(I,1)$-element is -1, we need a slice whose $(I,1)$-element is $+1$. Therefore we need $z(I-1, I; J-d+2, 1)$. Since the sum of $(I, J-d+2)$-elements is -1, we need $z(1, I; 2, J-d+2)$. In the same way, we find that $z(1, I; j, J-d+j)$, $j = 1, \ldots, d$, $z(I-1, I; J-d+j+1, j)$, $j = 1, \ldots, d-1$, and $z(I-1, I; j+1, j)$, $j = d, r$, are needed.

The sum of slices so far is as follows:

$$\begin{array}{c|ccccccc|}
 & 1 & \cdots & d & (r) & J-d+1 & \cdots & J \\
\hline
1 & +1 & \cdots & +1 & (0) & -1 & \cdots & -1 \\
I-1 & -1 & \cdots & -1 & (0) & +1 & \cdots & +1 \\
\hline
\end{array}.$$

Since the sum of $(I-1,1)$-elements is -1, we need $z(I-2, I-1; 2, 1)$. Since the sum of $(I-1,2)$-elements is -2, we need $z(I-2, I-1; 3, 2) \times 2$. In the same way, we find that $z(I-2, I-1; j+1, j) \times j$, $j = 1, \ldots, r$, and $z(I-2, I-1; J-j+1, J-j) \times j$, $j = 1, \ldots, d-1$, are needed.

The sum of slices so far is as follows:

$$\begin{array}{c|ccccccc|}
 & 1 & \cdots & d & (r) & J-d+1 & \cdots & J \\ \hline
1 & +1 & \cdots & +1 & (0) & -1 & \cdots & -1 \\
I-2 & -1 & \cdots & -1 & (0) & +1 & \cdots & +1 \\ \hline
\end{array}.$$

In the way, we find that $z(i, i+1, j+1, j) \times j$ $(i = 1, \dots, I-3,\ j = 1, \dots, r)$ and $z(i, i+1; J-j+1, J-j) \times j$ $(i = 1, \dots, I-3,\ j = 1, \dots, d-1)$ are needed. Hence all slices are needed for cancellation and this implies that z^* is an indispensable move.

Remark 3. *There are indispensable moves whose types are larger than the one in* (18) *for specific I and J. One example is the following move z of* $5 \times 5 \times 32$ *table for the case* $I = J = 5$. *Each* 5×5 *slice is a move of degree two of the form* $z(i_1, i_2; j_1, j_2)$. *We now list these 32 slices. In the list,* $-(i_1, i_2; j_1, j_2)$ *denotes* $-z(i_1, i_2; j_1, j_2) = z(i_1, i_2; j_2, j_1)$.

slice	1	2	3	4	5	6	7	8
move	(1,5;1,5)	−(1,2;1,2)	−(1,3;2,3)	−(1,2;3,4)	−(1,3;4,5)	−(2,3;1,3)	(2,4;2,4)	−(2,4;3,5)
slice	9	10	11	12	13	14	15	16
move	−(2,4;3,5)	(2,5;4,5)	(2,5;4,5)	−(3,4;1,2)	−(3,5;2,4)	−(3,5;2,4)	(3,5;3,5)	(3,5;3,5)
slice	17	18	19	20	21	22	23	24
move	−(3,4;4,5)	−(3,4;4,5)	−(3,4;4,5)	−(4,5;1,3)	(4,5;2,5)	(4,5;2,5)	−(4,5;3,4)	−(4,5;3,4)
slice	25	26	27	28	29	30	31	32
move	−(4,5;3,4)	−(4,5;4,5)	−(4,5;4,5)	−(4,5;4,5)	−(4,5;4,5)	−(4,5;4,5)	−(4,5;4,5)	−(4,5;4,5)

- *Since* $(1, 1)$*-element of slice 1 is* $+1$, *we need slice 2.*
- *Since the sum of* $(1, 2)$*-elements from slice 1 and 2 is* $+1$, *we need slice 3.*

  $\cdots$
- *Since the sum of* $(2, 2)$*-elements from slice 1 through slice 6 is* -1, *we need slice 7.*
- *Since the sum of* $(2, 3)$*-elements from slice 1 through slice 7 is* $+2$, *we need slice 7 and 8.*

  $\cdots$
- *Since the sum of* $(4, 4)$*-elements from slice 1 through slice 25 is* $+7$, *we need slice 26 through slice 32.*

Therefore, this move is indispensable.

Remark 4. *The case of* $I = 2$ *is relatively easy with the known explicit Graver basis, as already discussed in Section 4 of [6]. The question of when the equality holds in* (18) *seems to be difficult, partly because currently there is no algorithm to compute the Markov complexity.*

5. Discussion

In this paper we investigated a series of configurations A_b defined by fibers of a given base configuration A. We proved that the maximum Markov degree of the configurations is bounded from above by the Markov complexity of A. From our examples, the equality between the maximum Markov degree and the Markov complexity seems to hold only in special simple cases. As discussed after the statement of Theorem 1, the equality holds because the set of fibers for $A_{\boldsymbol{b}}$ is a subset of fibers of the higher Lawrence lifting $A^{(N)}$ of A. The strict inequality suggests that the former is a small subset of the latter. In particular for the case of incidence matrix complete bipartite graphs $K_{m,n}$, the maximum Markov degree for A_b is three independently of m and n, whereas the Markov complexity grows at least polynomially in m and n as shown in Section 4.2. Hence the discrepancy is large for this case.

Another interesting topic to investigate is the dependence of the Markov degree of $A_{\boldsymbol{b}}$ on $\boldsymbol{b}$. The results of Haase and Paffenholz [9] suggest that for generic $\boldsymbol{b}$, the Markov degree of $A_{\boldsymbol{b}}$ may be smaller than the maximum Markov degree. The result of Theorem 3 on the specific $\boldsymbol{b} = (2, 2, 2)$ suggests that this may a general phenomenon.

Acknowledgment

This work is partially supported by Grant-in-Aid for JSPS Fellows (No. 12J07561) from Japan Society for the Promotion of Science (JSPS) and by JSPS Grant-in-Aid for Scientific Research No. 25220001. We are grateful to Hidefumi Ohsugi for very useful suggestions and a proof mentioned in Remark 2.

References

[1] 4ti2 team. 4ti2—a software package for algebraic, geometric and combinatorial problems on linear spaces. Available at www.4ti2.de.

[2] S. Aoki, H. Hara, and A. Takemura. *Markov Bases in Algebraic Statistics*, volume 199 of *Springer Series in Statistics*. Springer, 2012.

[3] S. Aoki, T. Hibi, H. Ohsugi, and A. Takemura. Markov basis and Gröbner basis of Segre-Veronese configuration for testing independence in group-wise selections. *Annals of the Institute of Statistical Mathematics*, 62(2):299–321, 2010.

[4] Y. Berstein and S. Onn. The Graver complexity of integer programming. *Annals of Combinatorics*, 13(3):289–296, 2009.

[5] P. Diaconis and N. Eriksson. Markov bases for noncommutative Fourier analysis of ranked data. *Journal of Symbolic Computation*, 41(2):182–195, 2006.

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

[7] M. Domokos and D. Joó. On the equations and classification of toric quiver varieties. arXiv:1402.5096v1, 2014.

[8] M. Drton, B. Sturmfels, and S. Sullivant. *Lectures on Algebraic Statistics*, volume 39 of *Oberwolfach Seminars*. Birkhäuser Verlag, Basel, 2009.

[9] C. Haase and A. Paffenholz. Quadratic Gröbner bases for smooth 3×3 transportation polytopes. *Journal of Algebraic Combinatorics*, 30(4):477–489, 2009.

[10] H. Hara, A. Takemura, and R. Yoshida. On connectivity of fibers with positive marginals in multiple logistic regression. *Journal of Multivariate Analysis*, 101:909–925, 2010.

[11] D. Haws, A. Martin del Campo, A. Takemura, and R. Yoshida. Markov degree of the three-state toric homogeneous Markov chain model. *Beiträge zur Algebra und Geometrie*, 55:161–188, 2014.

[12] T. Hibi, editor. *Gröbner Bases: Statistics and Software Systems*. Springer, Tokyo, Japan, 2013.

[13] T. Kudo and A. Takemura. A lower bound for the Graver complexity of the incidence matrix of a complete bipartite graph. *Journal of Combinatorics*, 3(4):695–708, 2012.

[14] K. Ohara and N. Takayama. Pfaffian systems of A-hypergeometric systems II — holonomic gradient method. arXiv:1505.02947 [cs.SC], 2015.

[15] H. Ohsugi and T. Hibi. Toric rings and ideals of nested configurations. *Journal of Commutative Algebra*, 2:187–208, 2010.

[16] F. Santos and B. Sturmfels. Higher Lawrence configurations. *Journal of Combinatorial Theory, Series A*, 103(1):151–164, 2003.

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

[18] N. Takayama. Gröbner basis for rings of differential operators and applications. In *Gröbner Bases: Statistics and Sofware Systems*, pages 279–344. Springer, Tokyo, 2013.

[19] T. Yamaguchi, M. Ogawa, and A. Takemura. Markov degree of the Birkhoff model. *Journal of Algebraic Combinatorics*, 40(1):293–311, 2014.