AISM (2008) 60:229–256
DOI 10.1007/s10463-006-0089-x

Minimal invariant Markov basis for sampling contingency tables with fixed marginals

Satoshi Aoki · Akimichi Takemura

Received: 20 January 2006 / Revised: 12 May 2006 / Published online: 16 December 2006

Abstract In this paper we define an invariant Markov basis for a connected Markov chain over the set of contingency tables with fixed marginals and derive some characterizations of minimality of the invariant basis. We also give a necessary and sufficient condition for uniqueness of minimal invariant Markov bases. By considering the invariance, Markov bases can be presented very concisely. As an example, we present minimal invariant Markov bases for all $2 \times 2 \times 2 \times 2$ hierarchical models. The invariance here refers to permutation of indices of each axis of the contingency tables. If the categories of each axis do not have any order relations among them, it is natural to consider the action of the symmetric group on each axis of the contingency table. A general algebraic algorithm for obtaining a Markov basis was given by Diaconis and Sturmfels (*The Annals of Statistics*, **26**, 363–397, 1998). Their algorithm is based on computing Gröbner basis of a well-specified polynomial ideal. However, the reduced Gröbner basis depends on the particular term order and is not symmetric. Therefore, it is of interest to consider the properties of invariant Markov basis.

Keywords Exact tests · Hierarchical models · Markov chain Monte Carlo · Orbit · Symmetric group · Transformation group

S. Aoki (✉)
Department of Mathematics and Computer Science, Kagoshima University,
1-21-35, Korimoto, Kagoshima-shi, Kagoshima 890-0065, Japan
e-mail: aoki@sci.kagoshima-u.ac.jp

A. Takemura
Graduate School of Information Science and Technology, University of Tokyo,
7-3-1, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

1 Introduction

In performing exact conditional tests in discrete exponential families given sufficient statistics, the p values are usually calculated by large sample approximations. However, when the sample size is small compared to the size of the sample space, the large sample approximation may not be sufficiently accurate. When the sample size and the sample space are relatively small, enumeration of the sample space may be feasible with some ingenious enumeration schemes. For the case of two-way contingency tables with fixed row and column sums, Mehta and Patel (1983) proposed a network algorithm, which incorporates appropriate trimming in the enumeration. Aoki (2002), Suzuki et al. (2005) and Aoki (2003) extended this trimming for Fisher's exact test in two-way contingency tables and for the conditional tests of the Hardy–Weinberg proportions (triangular two-way contingency tables). However, the problem of computing p values by enumeration for k-way contingency tables, $k > 2$, seems to be largely open among researchers.

As another approach, a Markov chain Monte Carlo approach is extensively used in various settings of contingency tables, for example, Besag and Clifford (1989) for performing significance tests for the Ising model; Smith et al. (1996) for tests of independence, quasi-independence and quasi-symmetry for square contingency tables; Aoki and Takemura (2005) for tests of quasi-independence for two-way contingency tables containing some structural zeros; Guo and Thompson (1992) for tests of the Hardy-Weinberg proportions; Diaconis and Saloff-Coste (1995) for two-way contingency tables; Hernek (1998), Dyer and Greenhill (2000) for $2 \times J$ contingency tables; Forster et al. (1996) for 2^k contingency tables. Most of these works deal with various two-way settings.

Diaconis and Sturmfels (1998) proposed a general algorithm for generating random samples from a conditional distribution given sufficient statistics for general discrete exponential family of distributions. They suggest computing a Markov basis by finding a Gröbner basis of a well-specified polynomial ideal. Their approach is extremely appealing because it can be used for problems of any dimension. In addition, rapid progress of Gröbner bases computation has made it possible to carry out Markov chain Monte Carlo methods in various statistical problems efficiently. See Sturmfels (1995) and De Loera et al. (2004) for example. Nowadays, we can use some nice systems of algebraic computations to obtain Gröbner bases and Markov bases of toric ideals such as (4ti2 team). Quite recently, the Markov basis of the very difficult and complicated problem of $4 \times 4 \times 4$ contingency tables with fixed two-dimensional marginals has been solved using Gröbner basis technology. For this problem, Lauritzen (2005) reported that there is a 15th basis element that was unknown so far, and Hemmecke and Malkin (2005) actually computed the Markov basis completely. According to Hemmecke and Malkin (2005), this problem has been completely solved at last, showing that the list in Aoki and Takemura (2003b) lacks one basis element.

On the other hand, in this paper, we focus on the theoretical properties of minimal Markov bases from the viewpoint of *invariance*. This work is motivated by the following facts; (1) a reduced Gröbner basis does not coincide with the minimal Markov basis in general, and (2) a reduced Gröbner basis lacks symmetry in general even when the original problem has obvious symmetry in the indeterminates. If we only consider Gröbner bases, then under a given monomial ordering the reduced Gröbner basis is unique minimal and Diaconis and Sturmfels (1998) suggest computing a reduced Gröbner basis under some given monomial ordering to obtain a Markov basis. However, any system of generators of the ideal constitutes a Markov basis and a minimal Markov basis is in general a proper subset of a Gröbner basis. Although the theoretical characterizations of Markov bases such as minimality and symmetry are not necessarily needed for performing Markov chain Monte Carlo sampling in actual problems, the invariance is one of the essential and fundamental features of many statistical models and we believe that this work clarifies some fundamental aspects of minimal Markov bases.

By utilizing invariance as much as possible, some interesting by-product is obtained in this paper, namely, *concise description of Markov bases by orbit lists*. To illustrate this, we list the numbers of the elements of the unique minimal Markov basis, along with the numbers of the reduced Gröbner basis elements calculated by 4ti2 and the number of the orbits with respect to an action of a direct product of symmetric groups for the problem of $3 \times 3 \times K$ ($K \leq 7$) contingency tables with fixed two-dimensional marginals in Table 1. As we show in Sect. 3, a set of moves is partitioned into orbits which are equivalence classes by the action of the group. As is shown in Aoki and Takemura (2003a), there are at most six orbits of indispensable moves for these problems. We should mention that here we are only considering permuting the levels of each axis and not considering permutation of the axes (see Remark 1 below). In these examples, a minimal Markov basis is unique. Furthermore it is minimal invariant in the sense of present paper. Therefore, the representative basis elements for each orbit contain all the information of the minimal Markov basis. To perform the Markov chain Monte Carlo simulations using these orbit lists, ordinary users can first choose an orbit, and then apply the symmetric group action to the representative basis element for each step of the chain. See Sect. 2.3 for details.

It should be noted that the minimality of Markov bases is not needed for performing MCMC simulations and the minimality may be not desirable from the viewpoint of convergence rate. However, given a minimal Markov basis, it is easy to extend it to a larger basis by combining its moves. In fact,

Table 1 Number of the unique minimal Markov bases elements and reduced Gröbner bases elements for $3 \times 3 \times K, K \leq 7$, tables with fixed two-dimensional marginals

K	3	4	5	6	7
No. of the elements in the unique minimal Markov basis	81	450	2,670	10,665	31,815
No. of the elements in the reduced Gröbner basis	110	622	3,240	12,085	34,790
No. of orbits in the unique minimal Markov basis	4	5	6	6	6

Diaconis and Sturmfels (1998) suggested an MCMC algorithm, in which such an extended set of moves was implicitly considered. Since it seems obviously efficient to consider the extended set of moves in view of the convergence rate, our algorithm in Sect. 2.3 also takes into account this extension.

Another interesting consideration is how to choose a minimal Markov basis if it is not unique. For such cases, different minimal Markov bases contain different numbers of orbits in general, and each basis element in these orbits is not necessarily needed in general. As an example, we consider the $2 \times 2 \times 2$ contingency tables with fixed one-dimensional marginals, i.e., the complete independence model. As is shown in Sect. 3 of Takemura and Aoki (2004), the minimal Markov basis for this problem is not unique. Each minimal Markov basis contains six indispensable elements and three dispensable elements. For example, the reduced Gröbner basis with respect to the graded reverse lexicographic order contains three dispensable moves (binomials) such as

$$(121)(212) - (111)(222),\ (122)(211) - (111)(222),\ (112)(221) - (111)(222),$$

where $(121)(212) - (111)(222)$ denotes the move with $+1$ in cells (121), (212) and with -1 in cells (111), (222). A formal definition of this notation is given in Sect. 2.1. It is seen that these three dispensable basis elements are in different orbits, respectively. On the other hand, from the argument in Sect. 3 of this paper, another minimal basis is constructed from three dispensable basis elements such as

$$(121)(212) - (111)(222),\ (122)(211) - (111)(222),\ (112)(221) - (121)(212).$$

In this basis, the second and the third binomials, $(122)(211) - (111)(222)$ and $(112)(221) - (121)(212)$, are in the same orbit. In fact, we see that $(112)(221) - (121)(212)$ can be produced from $(122)(211) - (111)(222)$ by interchanging the cell indices $1, 2$ in the second axis. Accordingly, if we consider an action of the symmetric group, only two basis elements such as

$$(121)(212) - (111)(222),\ (122)(211) - (111)(222)$$

have to be included in our list because the third basis element can be produced by permuting the second axis. This example indeed corresponds to a minimal invariant Markov basis defined in this paper. Such an invariance and minimality is not necessarily needed for irreducibility and the convergence behavior of the chain. Our argument in this paper is focused on the conciseness of the list. In Takemura and Aoki (2004), we derived some characterizations of a minimal Markov basis and gave a necessary and sufficient condition for uniqueness of a minimal Markov basis. We combine these results with the theory of transformation groups to study the minimality of invariant Markov bases and give some characterizations of an invariant Markov basis and its minimality. We also give a necessary and sufficient condition for uniqueness of a minimal invariant Markov basis.

The construction of this paper is as follows. Definitions and notations of contingency tables, Markov basis and invariance are given in Sect. 2. Structures of a minimal invariant Markov basis are derived in Sect. 3. All hierarchical $2 \times 2 \times 2 \times 2$ models are systematically investigated in Sect. 4.

2 Preliminaries

In this section, we give necessary notations and definitions on Markov basis in Sect. 2.1 and group actions on contingency tables in Sect. 2.2. We also describe the Metropolis–Hastings sampling algorithms using the orbit list of an invariant Markov basis in Sect. 2.3.

2.1 Contingency tables and Markov basis

Consider an $I_1 \times \cdots \times I_k$ k-way contingency table $\boldsymbol{x}$. We denote a cell of the contingency table by $\boldsymbol{i} = (i_1 \dots i_k)$ or $\boldsymbol{i} = (i_1, \dots, i_k)$. The set of cells is denoted by

$$\mathcal{I} = \mathcal{I}^1 \times \cdots \times \mathcal{I}^k,$$

where $\mathcal{I}^\ell = \{1, \dots, I_\ell\}$, $\ell = 1, \dots, k$. We write $\boldsymbol{x} = \{x(\boldsymbol{i})\}_{\boldsymbol{i} \in \mathcal{I}}$ where $x(\boldsymbol{i})$ is a frequency of the cell $\boldsymbol{i}$. Let $\mathcal{X}$ denote the set of all k-way contingency tables given by

$$\mathcal{X} = \{\boldsymbol{x} = \{x(\boldsymbol{i})\}_{\boldsymbol{i} \in \mathcal{I}} \mid x(\boldsymbol{i}) \in \{0, 1, 2, \dots\} \text{ for } \boldsymbol{i} \in \mathcal{I}\}.$$

We define the sample size of $\boldsymbol{x} \in \mathcal{X}$ as $|\boldsymbol{x}| = \sum_{\boldsymbol{i} \in \mathcal{I}} x(\boldsymbol{i})$. Then $\mathcal{X}$ is partitioned by the sample sizes as

$$\mathcal{X} = \bigcup_{n=0}^{\infty} \mathcal{X}_n, \quad \mathcal{X}_n = \{\boldsymbol{x} \in \mathcal{X} \mid |\boldsymbol{x}| = n\}.$$

Let $K = \{1, \dots, k\}$ and let D denote a subset of K. The D-*marginal* $\boldsymbol{x}_D = \{x_D(\boldsymbol{i}_D)\}_{\boldsymbol{i}_D \in \mathcal{I}_D}$ of $\boldsymbol{x}$ is the contingency table with *marginal cells* $\boldsymbol{i}_D \in \prod_{\ell \in D} \mathcal{I}^\ell$ and entries given by

$$x_D(\boldsymbol{i}_D) = \sum_{\boldsymbol{j}_{K \setminus D} \in \mathcal{I}_{K \setminus D}} x(\boldsymbol{i}_D, \boldsymbol{j}_{K \setminus D}).$$

Note that $\boldsymbol{x}_D$ is an m-way contingency table if $D = \{i_1, \dots, i_m\}$.

Let $D_1, \dots, D_r \subset K$. Throughout this paper we assume that $D_1 \cup \cdots \cup D_r = K$ and there does not exist $i \neq j$ such that $D_i \subseteq D_j$. Note that $\{D_1, \dots, D_r\}$ corresponds to the generating class of a hierarchical log-linear model for the contingency tables. The set of D-marginal frequencies

$$\boldsymbol{t} = \boldsymbol{t}(\boldsymbol{x}) = (\boldsymbol{x}_{D_1}, \dots, \boldsymbol{x}_{D_r})$$

is the sufficient statistic under the hierarchical log-linear model. Note that if the cells and the elements of the sufficient statistic are ordered appropriately, we can write $\boldsymbol{t}$ in a matrix form as $\boldsymbol{t} = A\boldsymbol{x}$ as in Sect. 2.1 of Takemura and Aoki (2004).

We define the *reference set* of all the contingency tables having the same $(D_1, \ldots, D_r)$-marginals as

$$\mathcal{F}_{\boldsymbol{t}} = \mathcal{F}_{\boldsymbol{t}}(D_1, \ldots, D_r) = \{\boldsymbol{x} \in \mathcal{X} \mid \boldsymbol{t}(\boldsymbol{x}) = \boldsymbol{t}\}.$$

Since all the contingency tables in the same reference set $\mathcal{F}_{\boldsymbol{t}}$ have the same sample size, we define the sample size of $\boldsymbol{t}$ by $|\boldsymbol{t}| = |\boldsymbol{x}|, \boldsymbol{x} \in \mathcal{F}_{\boldsymbol{t}}$. Then the set $\mathcal{T}$ of possible values of the sufficient statistic $\boldsymbol{t}$, i.e., $\mathcal{T} = \{\boldsymbol{t}(\boldsymbol{x}) \mid \boldsymbol{x} \in \mathcal{X}\}$, is partitioned as

$$\mathcal{T} = \bigcup_{n=0}^{\infty} \mathcal{T}_n, \quad \mathcal{T}_n = \{\boldsymbol{t} \mid |\boldsymbol{t}| = n\}.$$

Let $\mathcal{Z} \supset \mathcal{X}$ be the set of k-way arrays $\boldsymbol{z} = \{z(\boldsymbol{i})\}_{\boldsymbol{i} \in \mathcal{I}}$ containing integer entries

$$\mathcal{Z} = \{\boldsymbol{z} = \{z(\boldsymbol{i})\}_{\boldsymbol{i} \in \mathcal{I}} \mid z(\boldsymbol{i}) \in \{\ldots, -1, 0, 1, \ldots\} \text{ for } \boldsymbol{i} \in \mathcal{I}\}.$$

Similarly, to the D-marginal $\boldsymbol{x}_D$ of $\boldsymbol{x}$, the D-marginal of $\boldsymbol{z}$ is defined and denoted by $\boldsymbol{z}_D$. An array $\boldsymbol{z} \in \mathcal{Z}$ is a *move* for $D_1, \ldots, D_r$ if $\boldsymbol{z}_{D_j} = \boldsymbol{0}$ for $j = 1, \ldots, r$. Here $\boldsymbol{0}$ denotes the zero array. Let $\mathcal{M}(D_1, \ldots, D_r)$ denote the set of all moves for $D_1, \ldots, D_r$ given by

$$\mathcal{M}(D_1, \ldots, D_r) = \{\boldsymbol{z} \in \mathcal{Z} \mid \boldsymbol{z}_{D_j} = \boldsymbol{0}, \ \ j = 1, \ldots, r\} \subset \mathcal{Z}.$$

For a move $\boldsymbol{z}$ for $D_1, \ldots, D_r$, the positive part $\boldsymbol{z}^+ = \{z^+(\boldsymbol{i})\}_{\boldsymbol{i} \in \mathcal{I}}$ and the negative part $\boldsymbol{z}^- = \{z^-(\boldsymbol{i})\}_{\boldsymbol{i} \in \mathcal{I}}$ are defined by

$$z^+(\boldsymbol{i}) = \max(z(\boldsymbol{i}), 0), \ \ z^-(\boldsymbol{i}) = \max(-z(\boldsymbol{i}), 0),$$

respectively. Then $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^-$ and $\boldsymbol{z}^+, \boldsymbol{z}^- \in \mathcal{X}$. Moreover, $\boldsymbol{z}^+$ and $\boldsymbol{z}^-$ have the same sufficient statistic, i.e., $\boldsymbol{t}(\boldsymbol{z}^+) = \boldsymbol{t}(\boldsymbol{z}^-)$, and are in the same reference set:

$$\boldsymbol{z}^+, \boldsymbol{z}^- \in \mathcal{F}_{\boldsymbol{t}(\boldsymbol{z}^+)}(D_1, \ldots, D_r) = \mathcal{F}_{\boldsymbol{t}(\boldsymbol{z}^-)}(D_1, \ldots, D_r).$$

Note that if $\boldsymbol{z}$ is a move, then $-\boldsymbol{z}$ is also a move with $(-\boldsymbol{z})^+ = \boldsymbol{z}^-$ and $(-\boldsymbol{z})^- = \boldsymbol{z}^+$. Furthermore non-zero elements of $\boldsymbol{z}^+$ and $\boldsymbol{z}^-$ do not share a common cell. We define the set of moves with the same value of the sufficient statistic of their positive and negative part $\boldsymbol{t} = \boldsymbol{t}(\boldsymbol{z}^+) = \boldsymbol{t}(\boldsymbol{z}^-)$ as

$$\mathcal{M}_{\boldsymbol{t}}(D_1, \ldots, D_r) = \{\boldsymbol{z} \in \mathcal{M}(D_1, \ldots, D_r) \mid \boldsymbol{t}(\boldsymbol{z}^+) = \boldsymbol{t}(\boldsymbol{z}^-) = \boldsymbol{t}\}.$$

We define the *degree* of $\boldsymbol{z} \in \mathcal{M}(D_1, \ldots, D_r)$ as the sample size of its positive and negative part, i.e., $\deg(\boldsymbol{z}) = |\boldsymbol{z}^+| = |\boldsymbol{z}^-|$. We also define the set of moves with degree *less than or equal to* n as

$$\mathcal{M}_n(D_1,\ldots,D_r) = \{\boldsymbol{z} \in \mathcal{M}(D_1,\ldots,D_r) \mid \deg(\boldsymbol{z}) \le n\}. \quad (1)$$

In this paper, we occasionally write simply $\mathcal{M}$, $\mathcal{M}_{\boldsymbol{t}}$ or $\mathcal{M}_n$ instead of $\mathcal{M}(D_1,\ldots, D_r)$, $\mathcal{M}_{\boldsymbol{t}}(D_1,\ldots,D_r)$ or $\mathcal{M}_n(D_1,\ldots,D_r)$, respectively, for convenience.

Let $\mathcal{B} \subset \mathcal{M}(D_1,\ldots,D_r)$ be a set of moves for $D_1,\ldots,D_r$. Let $\boldsymbol{x},\boldsymbol{x}' \in \mathcal{F}_{\boldsymbol{t}}(D_1, \ldots, D_r)$. We say that $\boldsymbol{x}'$ is *accessible* from $\boldsymbol{x}$ by $\mathcal{B}$ if there exists a sequence of moves $\boldsymbol{z}_1,\ldots,\boldsymbol{z}_A \in \mathcal{B}$ and $\varepsilon_s \in \{-1,1\}$, $s = 1,\ldots,A$, such that

$$\begin{aligned} \boldsymbol{x}' &= \boldsymbol{x} + \sum_{s=1}^{A} \varepsilon_s \boldsymbol{z}_s, \\ \boldsymbol{x} + \sum_{s=1}^{a} \varepsilon_s \boldsymbol{z}_s &\in \mathcal{F}_{\boldsymbol{t}}(D_1,\ldots,D_r) \quad \text{for } 1 \le a \le A, \end{aligned} \quad (2)$$

i.e., we can apply moves from $\mathcal{B}$ to $\boldsymbol{x}$ one by one and go from $\boldsymbol{x}$ to $\boldsymbol{x}'$, without causing negative cell frequencies on the way. It should be noted that the accessibility by $\mathcal{B}$ is an equivalence relation and each reference set is partitioned into disjoint equivalence classes by $\mathcal{B}$. We call these equivalence classes *$\mathcal{B}$-equivalence classes* of the reference set. If $\boldsymbol{x}$ and $\boldsymbol{x}'$ are elements from two different $\mathcal{B}$-equivalence classes of the same reference set, we say that a move $\boldsymbol{z} = \boldsymbol{x} - \boldsymbol{x}'$ *connects* these two equivalence classes (see Sect. 2 of Takemura and Aoki 2004).

Here we define a *Markov basis*.

Definition 1 *A finite set $\mathcal{B} \subset \mathcal{M}(D_1,\ldots,D_r)$ is a Markov basis for $D_1,\ldots,D_r$ if for all $\boldsymbol{t} \in \mathcal{T}$, $\mathcal{F}_{\boldsymbol{t}}(D_1,\ldots,D_r)$ itself constitutes one $\mathcal{B}$-equivalence class.*

A logically important point here is the existence of a finite Markov basis for any $D_1,\ldots,D_r$, which is guaranteed by the Hilbert basis theorem (see Sect. 3.1 of Diaconis and Sturmfels 1998). In this definition, if $\mathcal{B}$ is a Markov basis and $\boldsymbol{z}, -\boldsymbol{z} \in \mathcal{B}$, then $\mathcal{B}\backslash\{\boldsymbol{z}\}$ and $\mathcal{B}\backslash\{-\boldsymbol{z}\}$ are also Markov bases, respectively. Moreover, if we replace any element $\boldsymbol{z}$ of a Markov basis $\mathcal{B}$ with $-\boldsymbol{z}$, the resulting set is again a Markov basis. In other words, there is a freedom of the signs of the elements of a Markov basis. In this paper, we identify an element $\boldsymbol{z}$ of a Markov basis with its sign change $-\boldsymbol{z}$ for convenience.

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. However, a minimal Markov basis is not always unique. Takemura and Aoki (2004) gives some characterizations of a minimal Markov basis. An important fact is that for $\boldsymbol{t} \in \mathcal{T}$ such that $\mathcal{F}_{\boldsymbol{t}}(D_1,\ldots,D_r) = \{\boldsymbol{x},\boldsymbol{x}'\}$ is a two-element set, a move $\boldsymbol{z} = \boldsymbol{x} - \boldsymbol{x}'$ belongs to each Markov basis for $D_1,\ldots,D_r$ (see Lemma 2.3 of Takemura and Aoki 2004). We call such a move an *indispensable move*. Furthermore, the unique minimal Markov basis exists if and only if the set of indispensable moves forms a Markov basis. In this case, the set of indispensable moves is the unique minimal Markov basis (see Corollary 2.2 of Takemura and Aoki 2004).

Our moves contain many zero cells. Furthermore, often the non-zero cells of a move contain either 1 or -1. Therefore, a move can be concisely denoted by locations of its non-zero cells. We express a move $\boldsymbol{z}$ of degree n as $\boldsymbol{z} = \boldsymbol{i}_1 \dots \boldsymbol{i}_n - \boldsymbol{j}_1 \dots \boldsymbol{j}_n$, where $\boldsymbol{i}_1, \dots, \boldsymbol{i}_n$ are the cells of positive frequencies of $\boldsymbol{z}$ and $\boldsymbol{j}_1, \dots, \boldsymbol{j}_n$ are the cells of negative frequencies of $\boldsymbol{z}$. In the case $z(\boldsymbol{i}) > 1$, $\boldsymbol{i}$ is repeated $z(\boldsymbol{i})$ times. Similarly $\boldsymbol{j}$ is repeated $-z(\boldsymbol{j})$ times if $z(\boldsymbol{j}) < -1$. We use a similar notation for contingency tables as well. $\boldsymbol{x} \in \mathcal{X}_n$ is simply denoted as $\boldsymbol{x} = \boldsymbol{i}_1 \dots \boldsymbol{i}_n$.

2.2 Symmetric group and its action

Here we define an action of a direct product of symmetric groups on cells. From the action on cells, further actions are induced on contingency tables, marginal cells, marginal frequencies and moves.

First we give a brief list of definitions and notations of a group action. Let a group G acts on a set $\mathcal{U}$. Define $G(u) = \{gu \mid g \in G\}$ as the orbit through u. For a subset A of $\mathcal{U}$, we write $G(A) = \{gu \mid u \in A,\ g \in G\}$. Let $\mathcal{U}/G$ denote the orbit space, i.e., the set of orbits. Let $G_u = \{g \mid gu = u\}$ denote the isotropy subgroup of u in G. If G acts on $\mathcal{U}$, the action of G on the set of functions f on $\mathcal{U}$ is induced by $(gf)(u) = f(g^{-1}u)$. Let $h : \mathcal{U} \to \mathcal{V}$ be a surjection. If $h(u) = h(u') \Rightarrow h(gu') = h(gu), \forall g \in G$, then the action of G on $\mathcal{V}$ is induced by defining $gv = h(gu)$, where $v = h(u)$. Throughout the rest of this paper, the number of elements of a finite set A is denoted by $|A|$.

In our problem G is the direct product of symmetric groups, which acts on the index set $\mathcal{I}$. Let G^ℓ denote the symmetric group of order I_ℓ for $\ell = 1, \dots, k$ and let

$$G = G^1 \times G^2 \times \cdots \times G^k$$

be the direct product. We write an element of $g \in G$ as

$$g = g_1 \times \cdots \times g_k = \begin{pmatrix} 1 & \dots & I_1 \\ \sigma_1(1) & \dots & \sigma_1(I_1) \end{pmatrix} \times \cdots \times \begin{pmatrix} 1 & \dots & I_k \\ \sigma_k(1) & \dots & \sigma_k(I_k) \end{pmatrix}.$$

G acts on $\mathcal{I}$ by

$$\begin{aligned} \boldsymbol{i}' &= g\boldsymbol{i} \\ &= (g_1 i_1, \dots, g_k i_k) \\ &= (\sigma_1(i_1), \dots, \sigma_k(i_k)). \end{aligned}$$

Then the action of G on $\mathcal{X}$ is induced by

$$\begin{aligned} \boldsymbol{x}' &= g\boldsymbol{x} \\ &= \{x(g^{-1}\boldsymbol{i})\}_{\boldsymbol{i} \in \mathcal{I}}. \end{aligned}$$

G also acts on the marginal cells by

$$\begin{aligned} \boldsymbol{i}'_D &= g\boldsymbol{i}_D \\ &= (g_{s_1} i_{s_1}, \ldots, g_{s_m} i_{s_m}) \\ &= (\sigma_{s_1}(i_{s_1}), \ldots, \sigma_{s_m}(i_{s_m})), \end{aligned}$$

where $D = \{s_1, \ldots, s_m\}$. Hence G acts on the marginal tables by

$$\begin{aligned} \boldsymbol{x}'_D &= g\boldsymbol{x}_D \\ &= \{x_D(g^{-1}\boldsymbol{i}_D)\}_{\boldsymbol{i}_D \in \mathcal{I}_D}. \end{aligned}$$

Considering this action simultaneously for $D_1, \ldots, D_r$, the action of G on the sufficient statistic $\boldsymbol{t} = (\boldsymbol{x}_{D_1}, \ldots, \boldsymbol{x}_{D_r})$ is defined by

$$g\boldsymbol{t} = (g\boldsymbol{x}_{D_1}, \ldots, g\boldsymbol{x}_{D_r}).$$

An important point here is that the action of G on $\boldsymbol{t}$ is induced from the action of G on $\boldsymbol{x}$, because the calculation of D-marginals and the action of G on $\mathcal{X}$ are commutative. We state this as a lemma. Proof is easy and omitted.

Lemma 1 $(g\boldsymbol{x})_D = g\boldsymbol{x}_D$ *for all* $g \in G$ *and* $\boldsymbol{x} \in \mathcal{X}$.

By this lemma, if $\boldsymbol{x}_{D_i} = \boldsymbol{y}_{D_i}$, $i = 1, \ldots, r$, then $(g\boldsymbol{x})_{D_i} = (g\boldsymbol{y})_{D_i}$, $i = 1, \ldots, r$, $\forall g \in G$. In terms of the sufficient statistic this can be equivalently written as $\boldsymbol{t}(\boldsymbol{x}) = \boldsymbol{t}(\boldsymbol{y}) \Rightarrow \boldsymbol{t}(g\boldsymbol{x}) = \boldsymbol{t}(g\boldsymbol{y})$, $\forall g \in G$. Therefore, the action of G on $\mathcal{T}$ is induced from the action of G on $\mathcal{X}$. Also it is important to note that the isotropy subgroup $G_{\boldsymbol{t}}$ of $\boldsymbol{t}$ acts on the reference set $\mathcal{F}_{\boldsymbol{t}}$.

So far we have only considered non-negative frequencies. However, clearly the above consideration can also be applied to the set $\mathcal{Z}$ of integer arrays. In particular, Lemma 1 holds for the action of G on $\mathcal{Z}$, i.e., taking marginals of integer arrays commutes with the action of G. Therefore, if $\boldsymbol{z}$ is a move, then $g\boldsymbol{z}$ is a move as well. Therefore,

$$G(\mathcal{M}(D_1, \ldots, D_r)) = \mathcal{M}(D_1, \ldots, D_r).$$

and G acts on $\mathcal{M}(D_1, \ldots, D_r)$. More concretely, in terms of the positive part and the negative part we can write

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

We also define that a move $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^-$ is *symmetric* if $\boldsymbol{z}^+ = g\boldsymbol{z}^-$ for some $g \in G$. Conversely, a move $\boldsymbol{z}$ is *asymmetric* if $G(\boldsymbol{z}^+) \cap G(\boldsymbol{z}^-) = \emptyset$.

Now we can define an invariant set of moves. $\mathcal{B} \subset \mathcal{M}(D_1, \ldots, D_r)$ is *G-invariant* if $G(\mathcal{B}) = \mathcal{B}$. Note that here we are identifying a move $\boldsymbol{z} \in \mathcal{B}$

with its sign change $-\boldsymbol{z}$. Therefore, $\mathcal{B}$ is G-invariant if and only if

$$\forall g \in G, \forall \boldsymbol{z} \in \mathcal{B} \implies g\boldsymbol{z} \in \mathcal{B} \text{ or } -g\boldsymbol{z} \in \mathcal{B}.$$

In other words, $\mathcal{B}$ is G-invariant if and only if it is a union of orbits $\mathcal{B} = \bigcup_{\boldsymbol{z}\in A} G(\boldsymbol{z})$ for some subset $A \subset \mathcal{M}(D_1,\ldots,D_r)$ of moves.

A finite set $\mathcal{B} \subset \mathcal{M}(D_1,\ldots,D_r)$ is an *invariant Markov basis* for $D_1,\ldots,D_r$ if it is a Markov basis and it is G-invariant. An invariant Markov basis is *minimal* if no proper G-invariant subset of $\mathcal{B}$ is a Markov basis. A minimal invariant Markov basis always exists because from any invariant Markov basis, we can remove orbits one by one, until none of the remaining orbits can be removed any further.

Remark 1 In the formulation above we have considered permutation of the levels for each axis. If the number of levels of the axes is common and if in addition the hierarchical log-linear model considered is symmetric with respect to permutations of axes, we can further consider the permutation of the axes. For example in the case of the $3 \times 3 \times 3$ contingency tables with no three-factor interactions, we can consider the permutation of the axes. As is shown in Aoki and Takemura (2003a), if this additional symmetry of axes is considered, there are only 2 orbits corresponding to moves of degree 4 and 6, whereas if this additional symmetry is not considered there are 4 orbits as indicated in Table 1. In this paper we only consider a permutation of the levels for each axis, because it is applicable to all hierarchical models and numbers of levels.

2.3 Metropolis–Hastings sampling using the orbit lists

To perform the exact tests of various hierarchical models, our approach is to generate samples from the conditional distribution $f(\boldsymbol{x} \mid \boldsymbol{x}_{D_1},\ldots,\boldsymbol{x}_{D_r})$, where $\{D_1,\ldots,D_r\}$ is the generating class of the model considered, and calculate the null distribution of various test statistics. If a connected Markov chain over $\mathcal{F}_{\boldsymbol{t}}(D_1,\ldots,D_r)$ is constructed, the chain can be modified to give a connected and aperiodic Markov chain with stationary distribution $f(\boldsymbol{x} \mid \boldsymbol{x}_{D_1},\ldots,\boldsymbol{x}_{D_r})$ by the usual Metropolis procedure (Hastings 1970, for example).

As we have mentioned, we consider the sampling algorithms based on the invariant Markov basis in this paper. Our algorithm is described as follows. Let $\mathcal{B}$ be an invariant Markov basis and let $\boldsymbol{x}$ be the current state in $\mathcal{F}_{\boldsymbol{t}}$. To define the next state, choose $\boldsymbol{z} \in \mathcal{B}$ and $g \in G$ at random, where G is the symmetric group defined in Sect. 2.2, and calculate $\boldsymbol{z}' = g\boldsymbol{z}$. We also calculate the set of integers, $\{\epsilon_i\}$ so that $\boldsymbol{x} + \epsilon_i \boldsymbol{z}' \in \mathcal{F}_{\boldsymbol{t}}$. Then the next state is selected amongst these points with the probability

$$p_i = f(\boldsymbol{x} + \epsilon_i \boldsymbol{z}') \Big/ \sum_j f(\boldsymbol{x} + \epsilon_j \boldsymbol{z}').$$

3 Characterizations of a minimal invariant Markov basis and its uniqueness

In this section, we characterize the structure of a minimal invariant Markov basis and its uniqueness.

3.1 Structure of a minimal invariant Markov basis

In considering the orbits of G acting on $\mathcal{X}$, we note that $|\boldsymbol{x}| = |g\boldsymbol{x}|, \forall g \in G$, and hence $G(\mathcal{X}_n) = \mathcal{X}_n$. Therefore, we can consider the action of G on each $\mathcal{X}_n$ separately. Similarly we can consider the action of G on each $\mathcal{T}_n$ separately since $|\boldsymbol{t}| = |g\boldsymbol{t}|, \forall g \in G$. Consider a particular sufficient statistic $\boldsymbol{t} \in \mathcal{T}_n$. Let $G(\boldsymbol{t}) \in \mathcal{T}_n/G$ be the orbit through $\boldsymbol{t}$. Let

$$\mathcal{M}_{G(\boldsymbol{t})}(D_1,\ldots,D_r) = \bigcup_{\boldsymbol{t}' \in G(\boldsymbol{t})} \mathcal{M}_{\boldsymbol{t}'}(D_1,\ldots,D_r)$$

denote the union of the set of moves $\mathcal{M}_{\boldsymbol{t}'}$ over the orbit $G(\boldsymbol{t})$ through $\boldsymbol{t}$. Hereafter, we write $\mathcal{M}_{G(\boldsymbol{t})}$ instead of $\mathcal{M}_{G(\boldsymbol{t})}(D_1,\ldots,D_r)$ for simplicity.

Let $\mathcal{B} \subset \mathcal{M}$ be a finite set of moves. An important observation is that $\mathcal{B}$ is partitioned as

$$\mathcal{B} = \bigcup_{n} \bigcup_{\alpha \in \mathcal{T}_n/G} \mathcal{B}_{n,\alpha}, \tag{3}$$

where we define

$$\mathcal{B}_{n,\alpha} = \mathcal{B} \cap \mathcal{M}_\alpha, \quad \alpha \in \mathcal{T}_n/G.$$

Since $\mathcal{B}$ is invariant if and only if it is a union of orbits $G(\boldsymbol{z})$, the following lemma holds.

Lemma 2 *$\mathcal{B}$ is invariant if and only if $\mathcal{B}_{n,\alpha}$ is invariant for each n and $\alpha \in \mathcal{T}_n/G$.*

Proof Let $\boldsymbol{z} \in \mathcal{B}_{n,\alpha}$ and $\boldsymbol{t} = \boldsymbol{t}(\boldsymbol{z}^+) \in \alpha$. Then it follows that $g\boldsymbol{z} \in \mathcal{M}_{g\boldsymbol{t}} \subset \mathcal{M}_\alpha$ and the lemma is proved. □

This lemma shows that we can restrict our attention to each $\mathcal{B}_{n,\alpha}$ in studying the invariance of a Markov basis.

In characterizing a Markov basis and its minimality, Takemura and Aoki (2004) showed that it is essential to consider $\mathcal{M}_{|\boldsymbol{t}|-1}$-equivalence classes of $\mathcal{F}_{\boldsymbol{t}}$, where $\mathcal{M}_n$ is given in (1). Considering the appropriate group actions on the set of moves and each reference set, we characterize the structure of a minimal invariant Markov basis in this section. As we will show in the following, the relation between the action of the isotropy subgroup $G_{\boldsymbol{t}}$ and $\mathcal{M}_{|\boldsymbol{t}|-1}$-equivalence classes of $\mathcal{F}_{\boldsymbol{t}}$ is important. In this paper, we write the set of $\mathcal{M}_{|\boldsymbol{t}|-1}$-equivalence classes of $\mathcal{F}_{\boldsymbol{t}}$ as $\mathcal{H}_{\boldsymbol{t}}$ for simplicity, i.e., $\mathcal{H}_{\boldsymbol{t}} = \mathcal{F}_{\boldsymbol{t}}/\mathcal{M}_{|\boldsymbol{t}|-1}$.

Now we state the main theorem.

Theorem 1 *Let $\mathcal{B}$ be a minimal G-invariant Markov basis and Let $\mathcal{B} = \bigcup_n \bigcup_{\alpha \in \mathcal{T}_n/G} \mathcal{B}_{n,\alpha}$ be the partition in (3). Then each $\mathcal{B}_{n,\alpha}$, $\alpha \in \mathcal{T}_n/G$, is a minimal invariant set of moves, where $\mathcal{B}_{n,\alpha} \cap \mathcal{M}_{\boldsymbol{t}}$, $\boldsymbol{t} \in \alpha$, connects $\mathcal{M}_{|\boldsymbol{t}|-1}$-equivalence classes of $\mathcal{F}_{\boldsymbol{t}}$ and*

$$\mathcal{B}_{n,\alpha} = G(\mathcal{B}_{n,\alpha} \cap \mathcal{M}_{\boldsymbol{t}}) \tag{4}$$

for any $\boldsymbol{t} \in \alpha$.

Conversely, from each $\alpha \in \mathcal{T}_n/G$ with $|\mathcal{H}_{\boldsymbol{t}}| \geq 2$, where $\boldsymbol{t} \in \alpha$ is a representative sufficient statistic, choose a minimal $G_{\boldsymbol{t}}$-invariant set of moves $\mathcal{B}_{\boldsymbol{t}} \subset \mathcal{M}_{\boldsymbol{t}}$ connecting $\mathcal{M}_{|\boldsymbol{t}|-1}$-equivalence classes of $\mathcal{F}_{\boldsymbol{t}}$, where $G_{\boldsymbol{t}} \subset G$ is the isotropy subgroup of $\boldsymbol{t}$, and extend $\mathcal{B}_{\boldsymbol{t}}$ to $G(\mathcal{B}_{\boldsymbol{t}})$. Then

$$\mathcal{B} = \bigcup_n \bigcup_{\substack{\alpha \in \mathcal{T}_n/G \\ |\mathcal{H}_{\boldsymbol{t}}| \geq 2, \boldsymbol{t} \in \alpha}} G(\mathcal{B}_{\boldsymbol{t}})$$

is a minimal G-invariant Markov basis.

This theorem only adds a statement of minimal G-invariance to the structure of a minimal Markov basis considered in Theorem 1 of Takemura and Aoki (2004).

In principle this theorem can be used to construct a minimal invariant Markov basis by considering $\bigcup_{\alpha \in \mathcal{T}_n/G} \mathcal{B}_{n,\alpha}, n = 1, 2, 3, \ldots$ step by step. By the Hilbert basis theorem, there exists some n_0 such that for $n \geq n_0$ no new moves need to be added. Then a minimal invariant Markov basis is written as $\bigcup_{n=1}^{n_0} \bigcup_{\alpha \in \mathcal{T}_n/G} \mathcal{B}_{n,\alpha}$. Obviously, there is a considerable difficulty in implementing this procedure directly.

To prove this theorem, we prepare some lemmas in the following.

First, we derive some basic properties of orbits of G acting on each reference set. As we stated before, we consider the action of G on each $\mathcal{X}_n$ separately. Let

$$\mathcal{F}_{G(\boldsymbol{t})} = \bigcup_{\boldsymbol{t}' \in G(\boldsymbol{t})} F_{\boldsymbol{t}'}$$

denote the union of reference sets over the orbit $G(\boldsymbol{t})$ through $\boldsymbol{t}$. Let $\boldsymbol{x} \in \mathcal{F}_{\boldsymbol{t}}$. Because $\boldsymbol{t}(g\boldsymbol{x}) = g\boldsymbol{t}$, it follows that

$$g\boldsymbol{x} \in \mathcal{F}_{g\boldsymbol{t}} \subset \mathcal{F}_{G(\boldsymbol{t})}.$$

Therefore, $G(\mathcal{F}_{G(\boldsymbol{t})}) = \mathcal{F}_{G(\boldsymbol{t})}$. This implies that $\mathcal{X}_n$ is partitioned as

$$\mathcal{X}_n = \bigcup_{\alpha \in \mathcal{T}_n/G} \mathcal{F}_\alpha, \tag{5}$$

where α runs over the set of different orbits and we can consider the action of G on each $\mathcal{F}_{G(\boldsymbol{t})}$ separately.

Consider a particular $\mathcal{F}_{G(\boldsymbol{t})}$. An important observation is that there is a direct product structure in $\mathcal{F}_{G(\boldsymbol{t})}$. Write

$$G(\boldsymbol{t}) = \{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_a\}, \tag{6}$$

where $a = a(\boldsymbol{t}) = |G(\boldsymbol{t})|$ is the number of elements of the orbit $G(\boldsymbol{t}) \subset \mathcal{T}_n$. Let $b = b(\boldsymbol{t}) = |\mathcal{F}_{G(\boldsymbol{t})}/G|$ be the number of orbits of G acting on $\mathcal{F}_{G(\boldsymbol{t})}$ and let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_b$ be representative elements of different orbits, i.e.,

$$\mathcal{F}_{G(\boldsymbol{t})} = G(\boldsymbol{x}_1) \cup \cdots \cup G(\boldsymbol{x}_b) \tag{7}$$

gives a partition of $\mathcal{F}_{G(\boldsymbol{t})}$. Then we have the following lemma.

Lemma 3 *We use the notations (6) and (7). Then $\mathcal{F}_{G(\boldsymbol{t})}$ is partitioned as*

$$\mathcal{F}_{G(\boldsymbol{t})} = \bigcup_{i=1}^{a} \bigcup_{j=1}^{b} \mathcal{F}_{\boldsymbol{t}_i} \cap G(\boldsymbol{x}_j), \tag{8}$$

where each $\mathcal{F}_{\boldsymbol{t}_i} \cap G(\boldsymbol{x}_j)$ is non-empty. Furthermore, if $\boldsymbol{t}'_i = g\boldsymbol{t}_i$, then $\boldsymbol{x} \in \mathcal{F}_{\boldsymbol{t}_i} \mapsto g\boldsymbol{x} \in \mathcal{F}_{\boldsymbol{t}'_i}$ gives a bijection between $\mathcal{F}_{\boldsymbol{t}_i} \cap G(\boldsymbol{x})$ and $\mathcal{F}_{\boldsymbol{t}'_i} \cap G(\boldsymbol{x})$.

Proof Let $\mathcal{F}_{G(\boldsymbol{t})} = \mathcal{F}_{\boldsymbol{t}_1} \cup \cdots \cup \mathcal{F}_{\boldsymbol{t}_a}$ is a partition. Intersecting this partition with $\mathcal{F}_{G(\boldsymbol{t})} = \bigcup_{j=1}^{b} G(\boldsymbol{x}_j)$ gives the partition (8). Let $\boldsymbol{x} \in \mathcal{F}_{\boldsymbol{t}}$. Then the orbit $G(\boldsymbol{x})$ intersects each reference set, i.e. $G(\boldsymbol{x}) \cap \mathcal{F}_{\boldsymbol{t}_i} \neq \emptyset$ for $i = 1, \ldots, a$. Since every $g \in G$ is a bijection of $\mathcal{F}_{G(\boldsymbol{t})}$ to itself and

$$g(\mathcal{F}_{\boldsymbol{t}} \cap G(\boldsymbol{x})) = \mathcal{F}_{g\boldsymbol{t}} \cap G(\boldsymbol{x}),$$

g gives a bijection between $\mathcal{F}_{\boldsymbol{t}_i} \cap G(\boldsymbol{x})$ and $\mathcal{F}_{\boldsymbol{t}'_i} \cap G(\boldsymbol{x})$. □

In particular, for each j, $\mathcal{F}_{\boldsymbol{t}_i} \cap G(\boldsymbol{x}_j)$, $i = 1, \ldots, a$, have the same number of elements

$$|\mathcal{F}_{\boldsymbol{t}_1} \cap G(\boldsymbol{x}_j)| = \cdots = |\mathcal{F}_{\boldsymbol{t}_a} \cap G(\boldsymbol{x}_j)|.$$

In addition, for $\boldsymbol{t}_i, \boldsymbol{t}'_i \in G(\boldsymbol{t})$ such that $\boldsymbol{t}'_i = g\boldsymbol{t}_i$, the map $g : G_{\boldsymbol{t}_i} \to gG_{\boldsymbol{t}_i}g^{-1}$ gives an isomorphism between $G_{\boldsymbol{t}_i}$ and $G_{\boldsymbol{t}'_i} = gG_{\boldsymbol{t}_i}g^{-1}$, where $G_{\boldsymbol{t}_i}$ and $G_{\boldsymbol{t}'_i}$ are the the isotropy subgroup of $\boldsymbol{t}_i$ and $\boldsymbol{t}'_i$ in G, respectively. Therefore, there are the following isomorphic structures in $\mathcal{F}_{\boldsymbol{t}_i}$,

$$(G_{\boldsymbol{t}_i}, \mathcal{F}_{\boldsymbol{t}_i}) \simeq (G_{\boldsymbol{t}'_i}, \mathcal{F}_{\boldsymbol{t}'_i}). \tag{9}$$

Considering the isomorphic structure of (9), now we can focus our attention on each reference set. Consider a particular reference set $\mathcal{F}_{\boldsymbol{t}}$. Here we can restrict our attention to the action of $G_{\boldsymbol{t}}$ on $\mathcal{F}_{\boldsymbol{t}}$. As we have stated before, the relation between the action of $G_{\boldsymbol{t}}$ and $\mathcal{H}_{\boldsymbol{t}} = \mathcal{F}_{\boldsymbol{t}}/\mathcal{M}_{|\boldsymbol{t}|-1}$ (the $\mathcal{M}_{|\boldsymbol{t}|-1}$-equivalence classes of $\mathcal{F}_{\boldsymbol{t}}$) is essential. First we show the following lemma.

Lemma 4 *For any integer n, if $\boldsymbol{x}'$ is accessible from $\boldsymbol{x}$ by $\mathcal{M}_n$, then $g\boldsymbol{x}'$ is accessible from $g\boldsymbol{x}$ by $\mathcal{M}_n$.*

We give a proof of Lemma 4 in Appendix.

This lemma holds for all $g \in G$. In particular, $g\boldsymbol{x} \in \mathcal{F}_{\boldsymbol{t}(\boldsymbol{x})}$ if $g \in G_{\boldsymbol{t}}$. This implies that an action of $G_{\boldsymbol{t}}$ is induced on $\mathcal{H}_{\boldsymbol{t}}$. In the sequel let $X_\gamma \in \mathcal{H}_{\boldsymbol{t}}$ denote each equivalence class:

$$\mathcal{H}_{\boldsymbol{t}} = \{X_\gamma\}_{1 \le \gamma \le |\mathcal{H}_{\boldsymbol{t}}|}.$$

Let $\pi : \boldsymbol{x} \mapsto X_\gamma$ denote the natural projection of $\boldsymbol{x}$ to its equivalence class, then Lemma 4 states

$$\pi(\boldsymbol{x}) = \pi(\boldsymbol{x}') \Rightarrow \pi(g\boldsymbol{x}) = \pi(g\boldsymbol{x}').$$

Let $\boldsymbol{x} \in X_\gamma$ and $g \in G_{\boldsymbol{t}}$. Then $g\boldsymbol{x}$ belongs to some $\mathcal{M}_{n-1}$-equivalence class $X_{\gamma'}$. By Lemma 4, this γ' does not depend on the choice of $\boldsymbol{x} \in X_\gamma$ and we may write $\gamma' = g\gamma$. Since by definition a group action is bijective the following lemma holds.

Lemma 5 *$g \in G_{\boldsymbol{t}} : X_\gamma \mapsto X_{\gamma'}$ is a bijection of $\mathcal{H}_{\boldsymbol{t}}$ to itself.*

Here we give an illustration of a direct product structure of $\mathcal{F}_{G(\boldsymbol{t})}$. Figure 1 shows a structure of $\mathcal{F}_{G(\boldsymbol{t})}$ where $a = a(\boldsymbol{t}) = |G(\boldsymbol{t})| = 3$ and $b = b(\boldsymbol{t}) = |\mathcal{F}_{G(\boldsymbol{t})}/G| = 2$. In each $\mathcal{F}_{\boldsymbol{t}} \subset \mathcal{F}_{G(\boldsymbol{t})}$, there are two $\mathcal{M}_{|\boldsymbol{t}|-1}$-equivalence classes, i.e., $|\mathcal{H}_{\boldsymbol{t}}| = 2$. Figure 1 also shows $G_{\boldsymbol{t}}$ orbits in each $\mathcal{F}_{\boldsymbol{t}}$, which we consider in Sect. 3.2. In fact, Figure 1 is derived from an example of $2 \times 2 \times 2 \times 3$ contingency tables of the model

$$D_1 = \{1, 2\},\ D_2 = \{1, 3\},\ D_3 = \{2, 3\},\ D_4 = \{3, 4\}.$$

We see the above structure by considering $\boldsymbol{x} = (1111)(1221)(2122)(2212)$, for example. In this case, $\mathcal{F}_{\boldsymbol{t}(\boldsymbol{x})}$ is an eight elements set as follows.

$$\begin{array}{l}
\left.\begin{array}{l}(1111)(1221)(2122)(2212),\ (1111)(1222)(2121)(2212),\\ (1112)(1222)(2121)(2211),\ (1112)(1221)(2122)(2211),\end{array}\right\} X_\gamma\,(\ni \boldsymbol{x})\\
(1121)(1211)(2112)(2222),\ (1121)(1212)(2111)(2222),\\
\underbrace{(1122)(1212)(2111)(2221)}_{G_{\boldsymbol{t}}(\boldsymbol{x})},\ (1122)(1211)(2112)(2221).
\end{array}$$

Now we give a proof of Theorem 1.

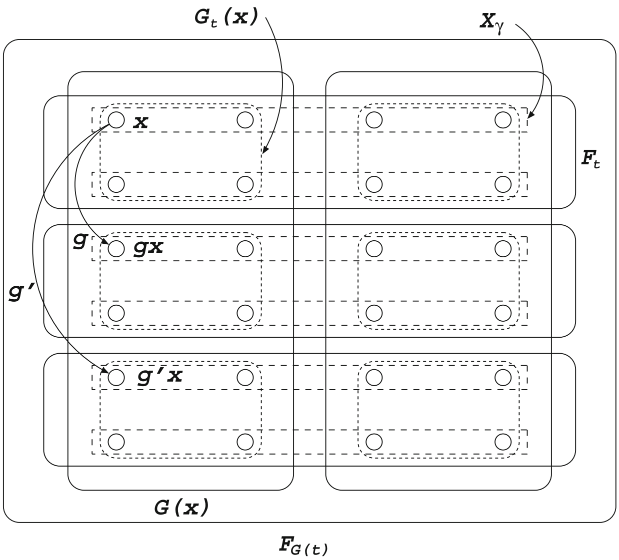

Fig. 1 A direct product structure of $\mathcal{F}_{G(\boldsymbol{t})}$ $(a = 3, b = 2, p = 1, q_i = 2, r_i = 2)$

Proof of Theorem 1 Let $\mathcal{B}$ be a minimal invariant set of moves and consider the partition (3). Then each $\mathcal{B}_{n,\alpha}, \alpha \in \mathcal{T}_n/G$ is G-invariant from Lemma 2. Moreover, from the argument of Takemura and Aoki (2004), each $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^- \in \mathcal{B}_{n,\alpha}$ is a move connecting $X_\gamma \in \mathcal{H}_{\boldsymbol{t}}$ and $X_{\gamma'} \in \mathcal{H}_{\boldsymbol{t}}, \gamma \neq \gamma'$, i.e., $\boldsymbol{z}^+ \in X_\gamma$ and $\boldsymbol{z}^- \in X_{\gamma'}$, from the minimality of $\mathcal{B}$. In this case, $g\boldsymbol{z} = g\boldsymbol{z}^+ - g\boldsymbol{z}^-$ is a move connecting $X_{g\gamma}$ and $X_{g\gamma'}$. Applying g^{-1} the converse is also true. This implies that the way $\mathcal{B}_{n,\alpha} \cap \mathcal{M}_{\boldsymbol{t}}$ connects the $\mathcal{M}_{n-1}$-equivalence classes $\mathcal{H}_{\boldsymbol{t}}$ is the same for all $\boldsymbol{t} \in \alpha$ and hence the relation (4) holds.

Conversely, to construct a minimal invariant Markov basis, we only have to consider sets of moves connecting $\mathcal{M}_{|\boldsymbol{t}|-1}$-equivalence classes of each $\mathcal{F}_{\boldsymbol{t}}$ from the argument of Takemura and Aoki (2004). Considering the isomorphic structure (9) of Lemma 3 and Lemma 5, we see that the structure of $\mathcal{H}_{\boldsymbol{t}'}$ is common for all $\boldsymbol{t}' \in G(\boldsymbol{t})$, and therefore it suffices to consider $G_{\boldsymbol{t}}$-invariant set of moves $\mathcal{B}_{\boldsymbol{t}}$ for some representative sufficient statistic $\boldsymbol{t} \in \alpha$ satisfying $|\mathcal{H}_{\boldsymbol{t}}| \geq 2$ for each $\alpha \in \mathcal{T}_n/G$. $\square$

3.2 Conditions for the uniqueness of a minimal invariant Markov basis

Now we derive a necessary and sufficient condition for the existence of a unique minimal invariant Markov basis. As is shown in Takemura and Aoki (2004), a minimal Markov basis is unique if and only if the set of the indispensable moves constitutes a Markov basis. Since the set of the indispensable moves is G-invariant, a minimal invariant Markov basis and a minimal Markov basis differ only in dispensable moves. In view of this we here state the following obvious fact.

Lemma 6 *If there exists a unique minimal Markov basis, then it is a unique minimal invariant Markov basis.*

To make the arguments clear, we give the following corollary to Theorem 1 without a proof.

Corollary 1 *A minimal invariant Markov basis $\mathcal{B}$ is unique if and only if for each n and $\alpha \in \mathcal{T}_n/G$ with $|\mathcal{H}_{\boldsymbol{t}}| \geq 2$ for a representative sufficient statistic $\boldsymbol{t} \in \alpha$, $\mathcal{B}_{\boldsymbol{t}} = \mathcal{B} \cap \mathcal{M}_{\boldsymbol{t}}$ is a unique minimal $G_{\boldsymbol{t}}$-invariant set of moves connecting $\mathcal{M}_{|\boldsymbol{t}|-1}$-equivalence classes of $\mathcal{F}_{\boldsymbol{t}}$.*

Considering this corollary, we consider $\mathcal{F}_{\boldsymbol{t}}$ for each $\boldsymbol{t}$ separately. We need to understand the action of $G_{\boldsymbol{t}}$ on $\mathcal{F}_{\boldsymbol{t}}$ in more detail. An important structure of $\mathcal{F}_{\boldsymbol{t}}$ is derived by considering the orbit space $\mathcal{H}_{\boldsymbol{t}}/G_{\boldsymbol{t}}$. Write

$$\mathcal{H}_{\boldsymbol{t}}/G_{\boldsymbol{t}} = \{\Gamma_1, \ldots, \Gamma_p\}, \tag{10}$$

where $p = p(\boldsymbol{t}) = |\mathcal{H}_{\boldsymbol{t}}/G_{\boldsymbol{t}}|$ is the number of orbits of $G_{\boldsymbol{t}}$ acting on $\mathcal{H}_{\boldsymbol{t}}$. We also write

$$\Gamma_i = \bigcup_{j=1}^{q_i} X_{\gamma_{i,j}}, \; i = 1, \ldots, p, \tag{11}$$

where $q_i = |\Gamma_i/\mathcal{M}_{|\boldsymbol{t}|-1}|$ is the number of different equivalence classes X_γ in Γ_i. By definition, Γ_i is $G_{\boldsymbol{t}}$-invariant and $G_{\boldsymbol{t}}$ acts on Γ_i for $i = 1, \ldots, p$. Therefore we consider each Γ_i separately. An important observation is that there is a direct product structure in each Γ_i, which is similar to Lemma 3. Let $r_i = |\Gamma_i/G_{\boldsymbol{t}}|$ be the number of orbits of $G_{\boldsymbol{t}}$ acting on Γ_i and let $\boldsymbol{x}_{i,1}, \ldots, \boldsymbol{x}_{i,r_i}$ be representative elements of different $G_{\boldsymbol{t}}$-orbits, i.e.,

$$\Gamma_i = G_{\boldsymbol{t}}(\boldsymbol{x}_{i,1}) \cup \cdots \cup G_{\boldsymbol{t}}(\boldsymbol{x}_{i,r_i}) \tag{12}$$

gives a partition of Γ_i. Then we have the following lemma.

Lemma 7 *$\mathcal{F}_{\boldsymbol{t}}$ is partitioned as*

$$\mathcal{F}_{\boldsymbol{t}} = \bigcup_{i=1}^{p} \Gamma_i = \bigcup_{i=1}^{p} \left(\bigcup_{j=1}^{q_i} \bigcup_{k=1}^{r_i} X_{\gamma_{i,j}} \cap G_{\boldsymbol{t}}(\boldsymbol{x}_{i,k}) \right), \tag{13}$$

where each $X_{\gamma_{i,j}} \cap G_{\boldsymbol{t}}(\boldsymbol{x}_{i,k})$ is non-empty for $i = 1, \ldots, p$, $j = 1, \ldots, q_i$, $k = 1, \ldots, r_i$. Furthermore, if $\gamma'_{i,j} = g\gamma_{i,j}, g \in G_{\boldsymbol{t}}$, then $\boldsymbol{x} \in X_{\gamma_{i,j}} \mapsto g\boldsymbol{x} \in X_{g\gamma_{i,j}}$ gives a bijection between $X_{\gamma_{i,j}} \cap G_{\boldsymbol{t}}(\boldsymbol{x})$ and $X_{\gamma'_{i,j}} \cap G_{\boldsymbol{t}}(\boldsymbol{x})$.

Proof Similarly to the proof of Lemma 3, intersecting the partition (11) with the partition (12) gives the partition (13). For each $\boldsymbol{x} \in \Gamma_i$, the orbit $G_{\boldsymbol{t}}(\boldsymbol{x})$ intersects all the equivalence classes $X_{\gamma_{i,j}}, j = 1, \ldots, q_i$, i.e., $G_{\boldsymbol{t}}(\boldsymbol{x}) \cap X_{\gamma_{i,j}} \neq \emptyset$ for all

$j = 1, \dots, q_i$. From Lemma 5 and the definition of Γ_i, every $g \in G_{\boldsymbol{t}}$ is a bijection of Γ_i to itself and

$$g(X_{\gamma_{i,j}} \cap G_{\boldsymbol{t}}(\boldsymbol{x})) = X_{g\gamma_{i,j}} \cap G_{\boldsymbol{t}}(\boldsymbol{x}).$$

Therefore, $g \in G_{\boldsymbol{t}}$ gives a bijection between $X_{\gamma_{i,j}} \cap G_{\boldsymbol{t}}(\boldsymbol{x})$ and $X_{\gamma'_{i,j}} \cap G_{\boldsymbol{t}}(\boldsymbol{x})$. □

Figure 1 shows a case that $p = |\mathcal{H}_{\boldsymbol{t}}/G_{\boldsymbol{t}}| = 1$. In fact, all the examples considered in Sect. 4 correspond to the case of $p = 1$. For theoretical interest, we present the following complicated example of $p = 2$.

Example 1 Consider the case of $k = 6$ and $D_1 = \{1,2\}, D_2 = \{1,3\}, D_3 = \{2,3\}, D_4 = \{4,5\}, D_5 = \{4,6\}, D_6 = \{5,6\}$. This is a direct product model of two three-way models with all two-dimensional marginals fixed. As for the $1, 2, 3$ axes, we consider $I_1 = 3, I_2 = 5, I_3 = 6$ and define $\boldsymbol{x}_{D_1}, \boldsymbol{x}_{D_2}, \boldsymbol{x}_{D_3}$ as

$$\begin{aligned}
\boldsymbol{x}_{D_1} &= (11)(13)(14)(15)(22)(23)(24)(25)(31)(32)(33)(34)(35)(35),\\
\boldsymbol{x}_{D_2} &= (11)(12)(13)(16)(23)(24)(25)(26)(31)(32)(34)(35)(36)(36),\\
\boldsymbol{x}_{D_3} &= (11)(16)(24)(26)(32)(33)(34)(41)(43)(45)(52)(55)(56)(56).
\end{aligned}$$

In this case, it can be easily verified that there are only two possible patterns of $\boldsymbol{x}_{\{1,2,3\}}$ as

$$\begin{aligned}
\boldsymbol{x}_1 &= (111)(132)(143)(156)(224)(233)(245)(256)(316)(326)(334)(341)(352)(355),\\
\boldsymbol{x}_2 &= (116)(133)(141)(152)(226)(234)(243)(255)(311)(324)(332)(345)(356)(356).
\end{aligned}$$

Note that there is a frequency of 2 at the cell (356) in $\boldsymbol{x}_2$. This implies that there is no $g \in G$ satisfying $\boldsymbol{x}_1 = g\boldsymbol{x}_2$, i.e., $G(\boldsymbol{x}_1) \cap G(\boldsymbol{x}_2) = \emptyset$. Therefore, $\boldsymbol{x}_1 - \boldsymbol{x}_2$ is an asymmetric indispensable move in $\{1, 2, 3\}$-marginal tables. As for the $4, 5, 6$ axes, we consider $I_1 = 2, I_2 = 7, I_3 = 7$ and define $\boldsymbol{x}_{D_4}, \boldsymbol{x}_{D_5}, \boldsymbol{x}_{D_6}$ as

$$\begin{aligned}
\boldsymbol{x}_{D_4} = \boldsymbol{x}_{D_5} &= (11)(12)(13)(14)(15)(16)(17)(21)(22)(23)(24)(25)(26)(27),\\
\boldsymbol{x}_{D_6} &= (11)(12)(21)(23)(32)(34)(43)(45)(54)(56)(65)(67)(76)(77).
\end{aligned}$$

In this case, again there are two possible patterns of $\boldsymbol{x}_{\{4,5,6\}}$ as

$$\begin{aligned}
\boldsymbol{x}'_1 &= (111)(123)(132)(145)(154)(167)(176)(212)(221)(234)(243)(256)(265)(277),\\
\boldsymbol{x}'_2 &= (112)(121)(134)(143)(156)(165)(177)(211)(223)(232)(245)(254)(267)(276).
\end{aligned}$$

Therefore, $\boldsymbol{x}'_1 - \boldsymbol{x}'_2$ is a symmetric indispensable move in $\{4, 5, 6\}$-marginal tables.

For the sufficient statistic $\boldsymbol{t} = \{\boldsymbol{x}_{D_1}, \boldsymbol{x}_{D_2}, \boldsymbol{x}_{D_3}, \boldsymbol{x}_{D_4}, \boldsymbol{x}_{D_5}, \boldsymbol{x}_{D_6}\}$ defined above, consider the structure of $\mathcal{F}_{\boldsymbol{t}}$. $\mathcal{F}_{\boldsymbol{t}}$ is written as

$$\mathcal{F}_{\boldsymbol{t}} = \{\boldsymbol{x} \mid \boldsymbol{x}_{\{1,2,3\}} = \boldsymbol{x}_1 \text{ or } \boldsymbol{x}_2 \text{ and } \boldsymbol{x}_{\{4,5,6\}} = \boldsymbol{x}'_1 \text{ or } \boldsymbol{x}'_2\} = X_{11} \cup X_{12} \cup X_{21} \cup X_{22},$$

where $X_{ij} = \{\boldsymbol{x} \mid \boldsymbol{x}_{\{1,2,3\}} = \boldsymbol{x}_i,\ \boldsymbol{x}_{\{4,5,6\}} = \boldsymbol{x}'_j\}$. Consider $\mathcal{M}_{13}$-equivalence classes of $\mathcal{F}_{\boldsymbol{t}}$. Note that the above four sets $X_{11}, X_{12}, X_{21}, X_{22}$ are $\mathcal{M}_2$-equivalence classes of $\mathcal{F}_{\boldsymbol{t}}$ since each set contains all combinations of permutations of $\{1, 2, 3\}$- and $\{4, 5, 6\}$-marginal patterns. Furthermore, any two elements in the different sets are not mutually accessible by $\mathcal{M}_{13}$ since $\boldsymbol{x}_1 - \boldsymbol{x}_2$ and $\boldsymbol{x}'_1 - \boldsymbol{x}'_2$ are indispensable moves in $\{1, 2, 3\}$- and $\{4, 5, 6\}$-marginal tables, respectively. From these considerations, we see that $|\mathcal{H}_{\boldsymbol{t}}| = 4$ and $\mathcal{H}_{\boldsymbol{t}} = \{X_{11}, X_{12}, X_{21}, X_{22}\}$. Considering the $G_{\boldsymbol{t}}$-orbit space of $\mathcal{H}_{\boldsymbol{t}}$, we have

$$\mathcal{H}_{\boldsymbol{t}}/G_{\boldsymbol{t}} = \{\{X_{11}, X_{12}\}, \{X_{21}, X_{22}\}\}$$

since $\boldsymbol{x}_1 - \boldsymbol{x}_2$ is an asymmetric move in $\{1, 2, 3\}$-marginal tables, whereas $\boldsymbol{x}'_1 - \boldsymbol{x}'_2$ is a symmetric move in $\{4, 5, 6\}$-marginal tables. Therefore, $p = |\mathcal{H}_{\boldsymbol{t}}/G_{\boldsymbol{t}}| = 2$ and $q_i = |\Gamma_i| = 2$ for each $\Gamma_i \in \mathcal{H}_{\boldsymbol{t}}/G_{\boldsymbol{t}}$, and we have the union of the direct products structure in (13).

Using this direct products structure of $\mathcal{F}_{\boldsymbol{t}}$ as shown in (13), first we summarize the structure of a minimal invariant set of moves connecting different Γ_i's.

Lemma 8 *$\mathcal{B}$ is a minimal $G_{\boldsymbol{t}}$-invariant set of moves that connects $\Gamma_1, \ldots, \Gamma_p$ in (10) if and only if $\mathcal{B}$ is written as*

$$\mathcal{B} = G_{\boldsymbol{t}}(\boldsymbol{z}_1) \cup \cdots \cup G_{\boldsymbol{t}}(\boldsymbol{z}_{p-1}), \tag{14}$$

where the set of the representative moves $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{p-1}$ connects $\Gamma_1, \ldots, \Gamma_p$ into a tree.

Proof Let $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^-$ is a move that connects Γ_i and $\Gamma_j, i \neq j$, i.e., $\boldsymbol{z}^+ \in \Gamma_i$ and $\boldsymbol{z}^- \in \Gamma_j$. Then $g\boldsymbol{z}$ also connects Γ_i and Γ_j for any $g \in G_{\boldsymbol{t}}$, since $g\boldsymbol{z}^+ \in \Gamma_i, g\boldsymbol{z}^- \in \Gamma_j$. □

This lemma implies the following necessary condition for the existence of a unique minimal invariant Markov basis.

Corollary 2 *If a minimal invariant Markov basis is unique, then the following conditions hold for all $\boldsymbol{t}$ such that $|\mathcal{H}_{\boldsymbol{t}}| \geq 2$.*

(i) *$p = p(\boldsymbol{t}) = |\mathcal{H}_{\boldsymbol{t}}/G_{\boldsymbol{t}}|$ is at most 2.*
(ii) *For $\mathcal{F}_{\boldsymbol{t}}$ such that $p(\boldsymbol{t}) = 2$, $G_{\boldsymbol{t}}(\boldsymbol{z})$ is the same for all $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^-, \boldsymbol{z}^+ \in \Gamma_1, \boldsymbol{z}^- \in \Gamma_2$, where $\mathcal{F}_{\boldsymbol{t}} = \Gamma_1 \cup \Gamma_2$.*

Combining the above results on the structure of a minimal invariant set of moves connecting the equivalence classes in each Γ_i, we can derive a necessary and sufficient condition that a minimal invariant Markov basis is unique.

First we define an *orbit graph* $\mathcal{G}_{\gamma_{i,j}} = \mathcal{G}(\Gamma_i, E_{\gamma_{i,j}})$ for $j = 2, \ldots, q_i$, where the edge set $E_{\gamma_{i,j}}$ is defined as

$$\begin{aligned} E_{\gamma_{i,j}} = \{(X_{\gamma_{i,j'}}, X_{\gamma_{i,j''}}) \mid (g\boldsymbol{z}^+, g\boldsymbol{z}^-) \in (X_{\gamma_{i,j'}}, X_{\gamma_{i,j''}}) \text{ for some } g \in G_{\boldsymbol{t}} \\ \text{where } \boldsymbol{z}^+ \in X_{\gamma_{i,1}}, \boldsymbol{z}^- \in X_{\gamma_{i,j}}\}. \end{aligned} \tag{15}$$

Here we are considering a $G_{\boldsymbol{t}}$-orbit $G_{\boldsymbol{t}}(\boldsymbol{z})$ of a move $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^-$ connecting $X_{\gamma_{i,1}}, X_{\gamma_{i,j}} \in \Gamma_i$, i.e., $\boldsymbol{z}^+ \in X_{\gamma_{i,1}}, \boldsymbol{z}^- \in X_{\gamma_{i,j}}$. We assume $\boldsymbol{z}^+ \in X_{\gamma_{i,1}}$ without loss of generality because $G_{\boldsymbol{t}}$ acts transitively on Γ_i. $E_{\gamma_{i,j}}$ does not depend on the choice of $\boldsymbol{z}$ with $\boldsymbol{z}^+ \in X_{\gamma_{i,1}}, \boldsymbol{z}^- \in X_{\gamma_{i,j}}$, although the orbit $G_{\boldsymbol{t}}(\boldsymbol{z})$ might depend on $\boldsymbol{z}$.

We also define that the orbit graph $\mathcal{G}_{\gamma_{i,j}}$ is *indispensable* if the graph $\mathcal{G}(\Gamma_i, \bigcup_{j' \neq j} E_{\gamma_{i,j'}})$ is not connected. An important point here is that if the set of indispensable orbit graphs connects all the equivalence classes in Γ_i, then this corresponds to the unique minimal invariant set of moves for Γ_i. Using these definitions, we have the following result.

Theorem 2 *A minimal invariant Markov basis is unique if and only if the following conditions hold for all $\boldsymbol{t}$ such that $|\mathcal{H}_{\boldsymbol{t}}| \geq 2$, in addition to (i) and (ii) of Corollary 2.*

(iii) $r_i = |\Gamma_i / G_{\boldsymbol{t}}| = 1$ *for all* $i = 1, \ldots, p$.
(iv) *The set of indispensable orbit graphs connects all* $X_{\gamma_{i,j}} \in \Gamma_i$ *for all* $i = 1, \ldots, p$.
(v) *For each orbit graph of (iv),* $G_{\boldsymbol{t}}(\boldsymbol{z})$ *is common for all* $\boldsymbol{z}$ *defining the edge set in* (15).

We give a proof of Theorem 2 in Appendix.

In Sect. 3 of Takemura and Aoki (2004), minimal Markov bases and their uniqueness are shown for some examples. We see that for some examples a minimal Markov basis is unique, and for other examples it is not unique. Since a unique minimal Markov basis is also the unique minimal invariant Markov basis, a logically interesting case is that, a minimal invariant Markov basis is unique, nevertheless a minimal Markov basis is not unique. The Hardy-Weinberg model is such an example, if we define a symmetric group acting on the upper triangular tables appropriately. See Sect. 3 of Takemura and Aoki (2004). Except for this peculiar example, the only example that we have found so far is one-way contingency tables.

Example 2 Consider the case of $k = 1$ and $D = \{1\}$. As is stated in Takemura and Aoki (2004), a minimal Markov basis for this case is not unique, and consists of $I_1 - 1$ degree 1 moves that connect I elements in $\mathcal{X}_1$ into a tree. By Cayley's theorem, there are $I_1^{I_1 - 2}$ ways of choosing a minimal Markov basis. On the other hand, the set of all degree 1 moves,

$$\mathcal{B} = \{\boldsymbol{x} - \boldsymbol{x}' \mid \boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}_1, \boldsymbol{x} \neq \boldsymbol{x}'\}$$

is a G-orbit in $\mathcal{M}(D)$. Therefore, $\mathcal{B}$ is the unique minimal invariant Markov basis. $\mathcal{B}$ consists of $\binom{I_1}{2}$ degree 1 moves.

4 Orbit list of minimal invariant Markov bases for all hierarchical 2^4 models

In this section, we present an orbit list of minimal invariant Markov bases for all hierarchical $2 \times 2 \times 2 \times 2$ models. There are 20 different such models. Figure 2 is the list of independence graphs of these models.

For each model, we can derive minimal Markov bases by 4ti2. Our purpose of this section is to characterize the outputs by 4ti2, by considering the direct product structure given in Sect. 3 and derive the orbit lists for minimal invariant Markov bases. Before presenting an orbit list of minimal invariant Markov bases, we first summarize the numbers of different minimal and minimal invariant Markov bases and numbers of elements in each basis. Table 2 shows the numbers ("kinds" in the table) of the minimal and minimal invariant bases and the number of elements in each minimal basis and minimal invariant basis together with their degrees. In this table, we specify each model by their generating set. For example, a model 123/24/34 means $D_1 = \{1,2,3\}, D_2 = \{2,4\}, D_3 = \{3,4\}$.

As we have stated, if the set of indispensable moves constitutes a Markov basis, this is a unique minimal (invariant) Markov basis. On the other hand, if a minimal Markov basis is not unique, uniqueness of a minimal invariant Markov basis is important. In all of 2^4 hierarchical models, however, we found that a minimal invariant Markov basis is also not unique when a minimal Markov basis is not unique. We discuss this point in Sect. 5. Furthermore, even the number of

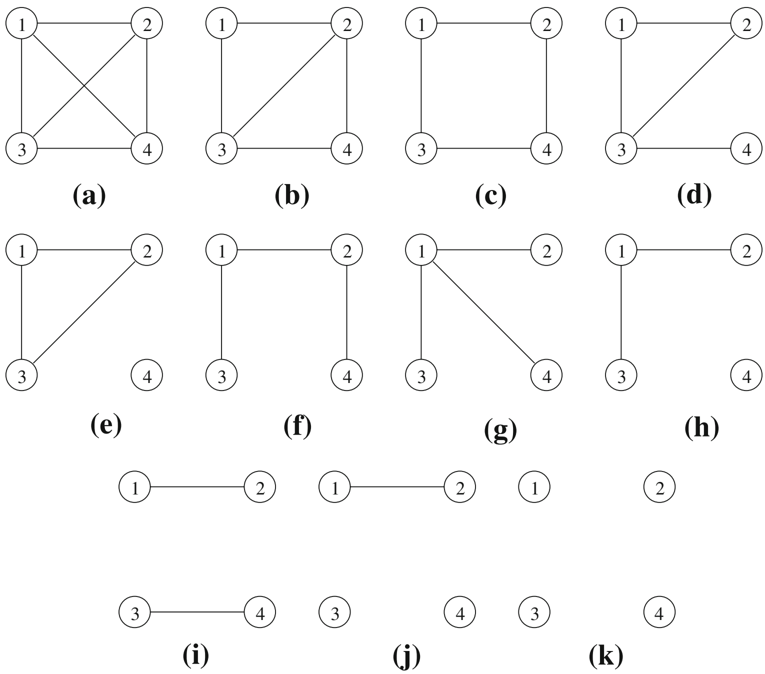

Fig. 2 Independence graphs for four-way contingency tables

Table 2 Number of minimal basis elements and minimal invariant basis elements for 2^4 hierarchical models

Graph	Generating set	Number of basis
(a)	1234	∅
	123/124/134/234	Unique minimal basis (1 move of deg 8)
	123/124/134	Unique minimal basis (2 moves of deg 4)
	123/124/34	Unique minimal basis (6 moves of deg 4)
	123/14/24/34	Unique minimal basis (12 moves of deg 4 and 8 moves of deg 6)
	12/13/14/23/24/34	Unique minimal basis (20 moves of deg 4 and 40 moves of deg 6)
(b)	123/234	Unique minimal basis (4 moves of deg 2)
	123/24/34	Unique minimal basis (4 moves of deg 2 and 16 moves of deg 4)
	12/13/23/24/34	Indispensable moves: 4 moves of deg 2 and 28 moves of deg 4
		Dispensable moves of a minimal basis: 16 kinds of 3 moves of deg 4
		Dispensable moves of a minimal invariant basis:
		3 kinds of 4 moves of deg 4
(c)	12/13/24/34	Unique minimal basis (8 moves of deg 2 and 8 moves of deg 4)
(d)	123/34	Unique minimal basis (12 moves of deg 2)
	12/13/23/34	Indispensable moves: 12 moves of deg 2 and 4 moves of deg 4
		Dispensable moves of a minimal basis:
		4,096 kinds of 5 moves of deg 4
		Dispensable moves of a minimal invariant basis:
		8 kinds of 10 moves of deg 4 or 2 kinds of 16 moves of deg 4
(e)	123/4	Unique minimal basis (28 moves of deg 2)
	12/13/23/4	Indispensable moves: 28 moves of deg 2 and 2 moves of deg 4
		Dispensable moves of a minimal basis:
		9216 kinds of 3 moves of deg 4
		Dispensable moves of a minimal invariant basis:
		24 kinds of 10 moves of deg 4 or 12 kinds of 16 moves of deg 4
(f)	12/13/24	Unique minimal basis (20 moves of deg 2)
(g)	12/13/14	Indispensable moves: 12 moves of deg 2
		Dispensable moves of a minimal basis:
		256 kinds of 6 moves of deg 2
		Dispensable moves of a minimal invariant basis:
		3 kinds of 8 moves of deg 2
(h)	12/13/4	Indispensable moves: 28 moves of deg 2
		Dispensable moves of a minimal basis:
		256 kinds of 6 moves of deg 2
		Dispensable moves of a minimal invariant basis:
		3 kinds of 8 moves of deg 2
(i)	12/34	Unique minimal basis (36 moves of deg 2)
(j)	12/3/4	Indispensable moves: 28 moves of deg 2
		Dispensable moves of a minimal basis:
		$16^6 = 16,777,216$ kinds of 18 moves of deg 2
		Dispensable moves of a minimal invariant basis:
		27 kinds of 24 moves of deg 2
(k)	1/2/3/4	Indispensable moves: 24 moves of deg 2
		Dispensable moves of a minimal basis:
		$16^8 \times 8^6 = 1.1259 \times 10^{15}$ kinds of 31 moves of deg 2
		Dispensable moves of a minimal invariant basis:
		2,268 kinds of 44 moves of deg 2

elements in a minimal invariant basis is not unique for the model 12/13/23/34 as discussed in Sect. 3. In order to illustrate construction of our orbits list, we discuss minimal invariant Markov basis for this complicated model.

Example 3 (Minimal invariant Markov bases for the model 12/13/23/34) For this model, there are 12 indispensable moves of degree 2 and 4 indispensable moves of degree 4, and the set of indispensable moves does not constitute a Markov basis. To construct a Markov basis, we have to consider the following two reference sets:

$$\mathcal{F}_{\boldsymbol{t}(\boldsymbol{x}_1^1)} = \left\{\boldsymbol{x}_1^1, \boldsymbol{x}_2^1, \boldsymbol{x}_3^1, \boldsymbol{x}_4^1\right\},$$
$$\mathcal{F}_{\boldsymbol{t}(\boldsymbol{x}_1^2)} = \{\boldsymbol{x}_1^2, \boldsymbol{x}_2^2, \boldsymbol{x}_3^2, \boldsymbol{x}_4^2, \boldsymbol{x}_5^2, \boldsymbol{x}_6^2, \boldsymbol{x}_7^2, \boldsymbol{x}_8^2\},$$

where

$$\boldsymbol{x}_1^1 = (1111)(1221)(2121)(2212),\ \boldsymbol{x}_2^1 = (1112)(1221)(2121)(2211),$$
$$\boldsymbol{x}_3^1 = (1121)(1211)(2112)(2221),\ \boldsymbol{x}_4^1 = (1121)(1212)(2111)(2221),$$
$$\boldsymbol{x}_1^2 = (1111)(1221)(2122)(2212),\ \boldsymbol{x}_2^2 = (1112)(1222)(2121)(2211),$$
$$\boldsymbol{x}_3^2 = (1111)(1222)(2121)(2212),\ \boldsymbol{x}_4^2 = (1112)(1221)(2122)(2211),$$
$$\boldsymbol{x}_5^2 = (1121)(1211)(2112)(2222),\ \boldsymbol{x}_6^2 = (1122)(1212)(2111)(2221),$$
$$\boldsymbol{x}_7^2 = (1121)(1212)(2111)(2222),\ \boldsymbol{x}_8^2 = (1122)(1211)(2112)(2221).$$

The isomorphic structures of these reference sets are given as

$$\mathcal{F}_{G(\boldsymbol{t}(\boldsymbol{x}_1^1))} = G(\boldsymbol{x}_1^1),$$
$$\left|G\left(\boldsymbol{t}\left(\boldsymbol{x}_1^1\right)\right)\right| = 4, \quad |\mathcal{F}_{G(\boldsymbol{t}(\boldsymbol{x}_1^1))}/G| = 1, \quad |\mathcal{F}_{\boldsymbol{t}(\boldsymbol{x}_1^1)}| = 4,$$
$$\mathcal{F}_{\boldsymbol{t}(\boldsymbol{x}_1^2)} = G(\boldsymbol{x}_1^2) \cup G\left(\boldsymbol{x}_3^2\right),$$
$$\left|G\left(\boldsymbol{t}\left(\boldsymbol{x}_1^2\right)\right)\right| = 1, \quad |\mathcal{F}_{G(\boldsymbol{t}(\boldsymbol{x}_1^2))}/G| = 2, \quad |\mathcal{F}_{\boldsymbol{t}(\boldsymbol{x}_1^2)}| = 8.$$

To make a minimal invariant Markov basis, we have to consider the direct product structures for $\mathcal{F}_{\boldsymbol{t}(\boldsymbol{x}_1^1)}$ and $\mathcal{F}_{\boldsymbol{t}(\boldsymbol{x}_1^2)}$. Since $p = |\mathcal{H}_{\boldsymbol{t}}/G_{\boldsymbol{t}}| = 1$ for both cases, we can write $\mathcal{F}_{\boldsymbol{t}} = \bigcup_{j=1}^{q} X_{\gamma_j}$ for simplicity. The direct product structures of $\mathcal{F}_{\boldsymbol{t}(\boldsymbol{x}_1^1)}$ is given as

$$\mathcal{F}_{\boldsymbol{t}(\boldsymbol{x}_1^1)} = X_{\gamma_1} \cup X_{\gamma_2} = G_{\boldsymbol{t}(\boldsymbol{x}_1^1)}(\boldsymbol{x}_1^1),$$
$$X_{\gamma_1} = \{\boldsymbol{x}_1^1, \boldsymbol{x}_2^1\}, \quad X_{\gamma_2} = \{\boldsymbol{x}_3^1, \boldsymbol{x}_4^1\},$$
$$q = 2, \quad r = 1,$$
$$|X_{\gamma_1} \cap G_{\boldsymbol{t}(\boldsymbol{x}_1^1)}(\boldsymbol{x}_1^1)| = |\{\boldsymbol{x}_1^1, \boldsymbol{x}_2^1\}| = 2.$$

Similarly, the direct product structures of $\mathcal{F}_{\boldsymbol{t}(\boldsymbol{x}_1^2)}$ is given as

$$\mathcal{F}_{\boldsymbol{t}(\boldsymbol{x}_1^2)} = X_{\gamma_1} \cup X_{\gamma_2} = G_{\boldsymbol{t}(\boldsymbol{x}_1^2)}(\boldsymbol{x}_1^2) \cup G_{\boldsymbol{t}(\boldsymbol{x}_1^2)}(\boldsymbol{x}_3^2),$$
$$X_{\gamma_1} = \{\boldsymbol{x}_1^2, \boldsymbol{x}_2^2, \boldsymbol{x}_3^2, \boldsymbol{x}_4^2\},\ X_{\gamma_2} = \{\boldsymbol{x}_5^2, \boldsymbol{x}_6^2, \boldsymbol{x}_7^2, \boldsymbol{x}_8^2\},$$
$$q = 2,\ r = 2,$$
$$|X_{\gamma_1} \cap G_{\boldsymbol{t}(\boldsymbol{x}_1^2)}(\boldsymbol{x}_1^2)| = |\{\boldsymbol{x}_1^2, \boldsymbol{x}_3^2\}| = 2$$

Since $p = 2$ for both cases, one move for each reference set suffices to construct a minimal Markov basis. However, to construct a minimal invariant Markov basis, we have to consider the orbit graph. The orbit graph for $\mathcal{F}_{\boldsymbol{t}(\boldsymbol{x}_1^1)}$ consists of two vertices, $\{X_{\gamma_1}, X_{\gamma_2}\}$, with edge, and is indispensable. However, there are two ways of choosing moves

$$\mathcal{B}_1 = \{\boldsymbol{x}_1^1 - \boldsymbol{x}_3^1,\ \boldsymbol{x}_2^1 - \boldsymbol{x}_4^1\},$$
$$\mathcal{B}_2 = \{\boldsymbol{x}_1^1 - \boldsymbol{x}_4^1,\ \boldsymbol{x}_2^1 - \boldsymbol{x}_3^1\},$$

which derive the above orbit graph. Therefore, the minimal invariant Markov basis has two kinds of two moves, $\mathcal{B}_1$ or $\mathcal{B}_2$, for $\mathcal{F}_{\boldsymbol{t}(\boldsymbol{x}_1^1)}$, and is not uniquely determined. On the other hand, the orbit graph for $\mathcal{F}_{\boldsymbol{t}(\boldsymbol{x}_1^2)}$ consists of two vertices, $\{X_{\gamma_1}, X_{\gamma_2}\}$, with edge, and is also indispensable. In this case, there are five ways of choosing moves

$$\mathcal{B}_1 = \{\boldsymbol{x}_1^2 - \boldsymbol{x}_5^2,\quad \boldsymbol{x}_2^2 - \boldsymbol{x}_6^2\},\ \mathcal{B}_2 = \{\boldsymbol{x}_1^2 - \boldsymbol{x}_6^2,\ \boldsymbol{x}_5^2 - \boldsymbol{x}_2^2\},$$
$$\mathcal{B}_3 = \{\boldsymbol{x}_3^2 - \boldsymbol{x}_7^2,\ \boldsymbol{x}_4^2 - \boldsymbol{x}_8^2\},\quad \mathcal{B}_4 = \{\boldsymbol{x}_3^2 - \boldsymbol{x}_8^2,\ \boldsymbol{x}_7^2 - \boldsymbol{x}_4^2\},$$
$$\mathcal{B}_5 = \{\boldsymbol{x}_1^2 - \boldsymbol{x}_7^2,\ \boldsymbol{x}_1^2 - \boldsymbol{x}_8^2\ \boldsymbol{x}_3^2 - \boldsymbol{x}_5^2,\ \boldsymbol{x}_3^2 - \boldsymbol{x}_6^2\ \boldsymbol{x}_4^2 - \boldsymbol{x}_5^2,\ \boldsymbol{x}_4^2 - \boldsymbol{x}_6^2\ \boldsymbol{x}_2^2 - \boldsymbol{x}_7^2,\ \boldsymbol{x}_2^2 - \boldsymbol{x}_8^2\},$$

which derive the above orbit graph. Therefore, there are five kinds of moves, i.e., four kinds of two moves, $\mathcal{B}_1, \ldots, \mathcal{B}_4$, or one kind of eight moves, $\mathcal{B}_5$, in the minimal invariant Markov basis.

Now we present an orbit list of minimal invariant Markov bases. For the models where the minimal invariant bases are not unique, we present one of the minimal orbits. Table 3 shows the representative basis elements for each orbit. This list includes 1 indispensable degree 8 move, 5 indispensable degree 6 moves $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_5)$, 21 indispensable degree 4 moves $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{21})$, 11 indispensable degree 2 moves $(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{11})$ and sets of dispensable moves constructing minimal invariant Markov bases.

Though our result is restricted to the case of $2 \times 2 \times 2 \times 2$, if a set of moves whose supports are contained in $2 \times 2 \times 2 \times 2$ array constitutes a Markov basis for a general $I_1 \times I_2 \times I_3 \times I_4$ case, we can derive a minimal and a minimal invariant Markov basis for the general case, by considering the orbits $\mathcal{T}_n/G$.

Table 3 Orbits list of minimal invariant bases for 2^4 hierarchical models

Graph	Generating set	Representative elements
(a)	1234	$\emptyset$
	123/124/134/234	(1111)(1122)(1212)(1221)(2112)(2121)(2211)(2222) −(1112)(1121)(1211)(1222)(2111)(2122)(2212)(2221)
	123/124/134	(1111)(1122)(1212)(1221) − (1112)(1121)(1211)(1222) = $\boldsymbol{x}_1$
	123/124/34	$\boldsymbol{x}_1$, (1111)(1122)(2112)(2121) − (1112)(1121)(2111)(2122) = $\boldsymbol{x}_2$, (1111)(1122)(2212)(2221) − (1112)(1121)(2211)(2222) = $\boldsymbol{x}_3$
	123/14/24/34	$\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$, (1111)(1212)(2112)(2211) − (1112)(1211)(2111)(2212) = $\boldsymbol{x}_4$, (1111)(1212)(2122)(2221) − (1112)(1211)(2121)(2222) = $\boldsymbol{x}_5$, (1111)(1222)(2112)(2221) − (1112)(1221)(2111)(2222) = $\boldsymbol{x}_6$, (1111)(1111)(1122)(1212)(2112)(2221) − (1112)(1112)(1121)(1211)(2111)(2222) = $\boldsymbol{w}_1$
	12/13/14/23/24/34	$\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$, $\boldsymbol{x}_4$, $\boldsymbol{x}_5$, $\boldsymbol{x}_6$, $\boldsymbol{w}_1$, (1111)(1221)(2121)(2211) − (1121)(1211)(2111)(2221) = $\boldsymbol{x}_7$, (1111)(1221)(2122)(2212) − (1121)(1211)(2112)(2222) = $\boldsymbol{x}_8$, (1111)(1222)(2121)(2212) − (1121)(1212)(2111)(2222) = $\boldsymbol{x}_9$, (1111)(1222)(2122)(2211) − (1122)(1211)(2111)(2222) = $\boldsymbol{x}_{10}$, (1111)(1111)(1222)(2122)(2212)(2221) − (1112)(1121)(1211)(2111)(2222)(2222) = $\boldsymbol{w}_2$, (1111)(1111)(1122)(1221)(2121)(2212) − (1112)(1121)(1121)(1211)(2111)(2222) = $\boldsymbol{w}_3$, (1111)(1111)(1212)(1221)(2122)(2211) − (1112)(1121)(1211)(1211)(2111)(2222) = $\boldsymbol{w}_4$, (1111)(1111)(1222)(2112)(2121)(2211) − (1112)(1121)(1211)(2111)(2111)(2222) = $\boldsymbol{w}_5$
(b)	123/234	(1111)(2112) − (1112)(2111) = $\boldsymbol{y}_1$
	123/24/34	$\boldsymbol{y}_1$, $\boldsymbol{x}_1$, $\boldsymbol{x}_3$, (1111)(1122)(1212)(2221) − (1112)(1121)(1211)(2222) = $\boldsymbol{x}_{11}$, (1111)(1122)(1221)(2212) − (1112)(1121)(1222)(2211) = $\boldsymbol{x}_{12}$, (1111)(1212)(1221)(2122) − (1112)(1211)(1222)(2121) = $\boldsymbol{x}_{13}$, (1111)(1212)(2122)(2221) − (1112)(1211)(2121)(2222) = $\boldsymbol{x}_{14}$, (1111)(1221)(2122)(2212) − (1112)(1222)(2121)(2211) = $\boldsymbol{x}_{15}$, (1111)(2122)(2212)(2221) − (1112)(2121)(2211)(2222) = $\boldsymbol{x}_{16}$
	12/13/23/24/34	$\boldsymbol{y}_1$, $\boldsymbol{x}_1$, $\boldsymbol{x}_3$, $\boldsymbol{x}_7$, $\boldsymbol{x}_9$, $\boldsymbol{x}_{11}$, $\boldsymbol{x}_{12}$, $\boldsymbol{x}_{13}$, $\boldsymbol{x}_{14}$, $\boldsymbol{x}_{16}$, (1111)(1221)(2121)(2212) − (1121)(1211)(2111)(2222) = $\boldsymbol{x}_{17}$, (1111)(1221)(2122)(2211) − (1121)(1211)(2112)(2221) = $\boldsymbol{x}_{18}$, (1111)(1222)(2121)(2211) − (1121)(1212)(2111)(2221) = $\boldsymbol{x}_{19}$, (1111)(1222)(2122)(2212) − (1121)(1212)(2112)(2222) = $\boldsymbol{x}_{20}$, (1111)(1222)(2122)(2211) − (1121)(1212)(2112)(2221) = $\boldsymbol{x}_{21}$, (1111)(1221)(2122)(2212) − (1121)(1211)(2112)(2222), (1111)(1221)(2122)(2212) − (1112)(1222)(2121)(2211)
(c)	12/13/24/34	$\boldsymbol{y}_1$, $\boldsymbol{x}_3$, $\boldsymbol{x}_5$, $\boldsymbol{x}_9$, $\boldsymbol{x}_{10}$, (1111)(1221) − (1121)(1211) = $\boldsymbol{y}_2$
(d)	123/34	$\boldsymbol{y}_1$, (1111)(1212) − (1112)(1211) = $\boldsymbol{y}_3$, (1111)(2212) − (1112)(2211) = $\boldsymbol{y}_4$
	12/13/23/34	$\boldsymbol{y}_1$, $\boldsymbol{y}_3$, $\boldsymbol{y}_4$, $\boldsymbol{x}_7$, $\boldsymbol{x}_{10}$, (1111)(1221)(2121)(2212) − (1121)(1211)(2112)(2221), (1111)(1221)(2122)(2212) − (1121)(1211)(2112)(2222)
(e)	123/4	$\boldsymbol{y}_1$, $\boldsymbol{y}_3$, $\boldsymbol{y}_4$, (1111)(1122) − (1112)(1121) = $\boldsymbol{y}_5$, (1111)(1222) − (1112)(1221) = $\boldsymbol{y}_6$, (1111)(2122) − (1112)(2121) = $\boldsymbol{y}_7$, (1111)(2222) − (1112)(2221) = $\boldsymbol{y}_8$
	12/13/23/4	$\boldsymbol{y}_1$, $\boldsymbol{y}_3$, $\boldsymbol{y}_4$, $\boldsymbol{y}_5$, $\boldsymbol{y}_6$, $\boldsymbol{y}_7$, $\boldsymbol{y}_8$, $\boldsymbol{x}_7$, (1111)(1221)(2121)(2212) − (1121)(1211)(2111)(2222), (1111)(1221)(2122)(2212) − (1121)(1211)(2112)(2222)
(f)	12/13/24	$\boldsymbol{y}_1$, $\boldsymbol{y}_2$, $\boldsymbol{y}_5$, $\boldsymbol{y}_6$, $\boldsymbol{y}_7$
(g)	12/13/14	$\boldsymbol{y}_2$, $\boldsymbol{y}_3$, $\boldsymbol{y}_5$, (1111)(1222) − (1112)(1221), (1111)(1222) − (1121)(1212)
(h)	12/13/4	$\boldsymbol{y}_1$, $\boldsymbol{y}_2$, $\boldsymbol{y}_3$,$\boldsymbol{y}_4$, $\boldsymbol{y}_5$, $\boldsymbol{y}_7$, $\boldsymbol{y}_8$, (1111)(1222) − (1112)(1221), (1111)(1222) − (1121)(1212)
(i)	12/34	$\boldsymbol{y}_1$, $\boldsymbol{y}_2$, $\boldsymbol{y}_3$, $\boldsymbol{y}_4$, $\boldsymbol{y}_6$, $\boldsymbol{y}_7$, $\boldsymbol{y}_8$, (1111)(2121) − (1121)(2111) = $\boldsymbol{y}_9$, (1111)(2221) − (1121)(2211) = $\boldsymbol{y}_{10}$

Table 3 (continued)

Graph	Generating set	Representative elements
(j)	12/3/4	$\boldsymbol{y}_1$, $\boldsymbol{y}_2$, $\boldsymbol{y}_3$, $\boldsymbol{y}_4$, $\boldsymbol{y}_5$, $\boldsymbol{y}_9$, $\boldsymbol{y}_{10}$, (1111)(1222) − (1112)(1221), (1111)(1222) − (1121)(1212), (1111)(2122) − (1112)(2121), (1111)(2122) − (1121)(2112), (1111)(2222) − (1112)(2221), (1111)(2222) − (1121)(2212)
(k)	1/2/3/4	$\boldsymbol{y}_1$, $\boldsymbol{y}_2$, $\boldsymbol{y}_3$, $\boldsymbol{y}_5$, $\boldsymbol{y}_9$, (1111)(2211) − (1211)(2111) = $\boldsymbol{y}_{11}$, (1111)(1222) − (1112)(1221), (1111)(1222) − (1121)(1212), (1111)(2122) − (1112)(2121), (1111)(2122) − (1121)(2112), (1111)(2212) − (1112)(2211), (1111)(2212) − (1211)(2112), (1111)(2221) − (1121)(2211), (1111)(2221) − (1211)(2121), (1111)(2222) − (1112)(2221), (1111)(2222) − (1121)(2212), (1111)(2222) − (1211)(2122)

5 Discussion

In this paper we define a minimal invariant Markov basis and derive its basic characteristics. Of course, we can construct an invariant Markov basis from any Markov basis as the union of all the orbits of the basis elements. However, even if we start with a minimal Markov basis, the union of all the orbits of the basis elements is not necessarily a minimal invariant basis. For example, consider again the complete independence model of the three-way case of Sect. 1. A set of moves,

$$(121)(212) - (111)(222),\ (122)(211) - (111)(222),\ (112)(221) - (111)(222),$$

i.e., the reduced Gröbner basis with respect to the graded reverse lexicographic order, connects the four elements $\{(111)(222), (112)(221), (121)(212), (122)(211)\}$ into a tree, and thus is a minimal basis elements for the reference set $\{(111)(222), (112)(221), (121)(212), (122)(211)\}$. However, it is seen that the union of the orbits of these three moves contains six moves, and hence not minimal invariant. From these considerations, the structure of a minimal invariant Markov basis is important, if we want to make an orbit list *as concise as possible*.

Theorem 1 states how to construct a minimal invariant Markov basis. This theorem is an extension of Theorem 1 of Takemura and Aoki (2004). To construct a minimal Markov basis, we can add basis elements step by step from the low degree, by considering all the reference sets as stated in Theorem 1 of Takemura and Aoki (2004). On the other hand, to construct a minimal invariant Markov basis, we have to add the orbit of moves step by step from the low degree. Similar to the construction of a minimal Markov basis, it is difficult to construct a minimal invariant Markov basis by applying Theorem 1 directly. But if a minimal Markov basis is available, we can construct a minimal invariant Markov basis relatively easily, by considering all the reference sets one by one, which is covered by the dispensable moves in the minimal Markov basis. The results of Sect. 4 is obtained in such a way.

It seems also difficult to give a simple necessary and sufficient conditions on $D_1, \ldots, D_r$ such that a minimal invariant Markov basis is unique. It is of interest to derive conditions such that a minimal invariant Markov basis is unique even if a minimal Markov basis is not unique. As stated in Sect. 3, such an example we have found so far is the obvious one-way contingency table, except for the peculiar case of the Hardy–Weinberg model. We would like to remind the readers that the investigation of uniqueness of minimal Markov basis in Takemura and Aoki (2004) led to the important notion of indispensable moves. Similarly, we expect that the investigation of uniqueness of minimal invariant Markov bases leads to some interesting facts on Markov bases and the actions of symmetric groups.

In this paper, we restrict our attention to the situation that the number of axes is fixed for a given problem. On the other hand, there are some situations that the cardinality of Markov bases increases linearly in the number of axes. One of such examples is the chain model described as

$$D_1 = \{1,2\}, D_2 = \{2,3\}, \ldots, D_r = \{r, r+1\},$$

which is suggested by a referee. In this case, our minimal invariant Markov basis will require exponentially many orbits as a function on r and it might be true that our minimal invariant Markov basis lacks the universality at this point. As the universality of minimal invariant Markov basis, we only consider the linearity in the number of levels of each axis. The linearity of Markov bases for r seems to be quite difficult condition. In fact, although it is true that this linearity exists for the Markov bases of decomposable models as the chain model, it is quite difficult to investigate the similar linearity for general non-decomposable models. Though this is a very interesting problem, we leave it to the future work.

Finally, the work in this paper is motivated from the theoretical interest for characterizations of invariant Markov bases, rather than practical values of invariant Markov bases. Our procedure for obtaining a minimal invariant Markov in Sect. 4 is rather complicated from practical viewpoint. Moreover, at present there is no efficient algorithm to obtain minimal invariant Markov bases. The authors believe, however, that symmetry is one of the important and essential features in most algebraic topics, and there is a possibility that this work may serve as one step towards a breakthrough in developing more efficient algorithms, namely, a *symmetric Buchberger algorithm*. This is an attractive topic, which is left to future work.

Appendix: Proofs of Lemma 4 and Theorem 2

Proof of Lemma 4 Note that $\deg(\boldsymbol{z}) \le n$ if and only if $\deg(g\boldsymbol{z}) \le n$. If $\boldsymbol{x}'$ is accessible from $\boldsymbol{x}$ by $\mathcal{M}_n$, then there exists $\boldsymbol{z}_1, \ldots \boldsymbol{z}_A \in \mathcal{M}_n$ satisfying

$$\begin{aligned} \boldsymbol{x}' &= \boldsymbol{x} + \sum_{s=1}^{A} \varepsilon_s \boldsymbol{z}_s, \\ \boldsymbol{x} + \sum_{s=1}^{a} \varepsilon_s \boldsymbol{z}_s &\in \mathcal{F}_{\boldsymbol{t}} \quad \text{for } 1 \le a \le A \end{aligned}$$

by (2). Applying g to the both sides of the equations we get

$$gx' = gx + \sum_{s=1}^{A} \varepsilon_s g z_s,$$
$$gx + \sum_{s=1}^{a} \varepsilon_s g z_s \in \mathcal{F}_{gt} \quad \text{for } 1 \leq a \leq A.$$

Since $gz_s \in \mathcal{M}_n$ for $s = 1, \ldots, A$, the lemma is proved. □

Proof of Theorem 2 Let $\mathcal{B}$ be a unique minimal invariant Markov basis. (iv) holds since $\mathcal{B}$ is an invariant Markov basis, and (v) also holds from the uniqueness of $\mathcal{B}$. To show (iii), suppose $r_i \geq 2$ for some $i = 1, \ldots, p$. In this case, the orbits $G_t(z)$ differ for a different choice of (δ_1, δ_2), where $z^+ \in X_{\gamma_{i,1}} \cap G_t(x_{\delta_1}), z^- \in X_{\gamma_{i,j}} \cap G_t(x_{\delta_2}), j \geq 2$, which contradicts the uniqueness of minimal invariant Markov basis.

Conversely, suppose the conditions (i), ..., (v) holds. Let z be an element of a minimal invariant Markov basis connecting different Γ_i's. By (i) we can assume that $p = 2$ and $\mathcal{F}_t = \Gamma_1 \cup \Gamma_2$. By ($ii$) $G_t(z)$ is the same for all $z = z^+ - z^-, z^+ \in \Gamma_1, z^- \in \Gamma_2$. Now let z be an element of a minimal invariant Markov basis connecting different $X_{\gamma_{i,j}}$'s for all $i = 1, \ldots, p$. We see that all the orbit graphs derived from z are indispensable from (iv) and minimality of invariant Markov basis. Moreover, for each orbit graph $G_t(z)$ is common for all z from (v). Therefore, we have shown that every minimal invariant Markov basis has the same set of orbits. □

References

Aoki, S. (2002). Improving path trimming in a network algorithm for Fisher's exact test in two-way contingency tables. *Journal of Statistical Computation and Simulation*, *72*, 205–216.

Aoki, S. (2003). Network algorithm for the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrical Journal*, *45*, 471–490.

Aoki, S., 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.

Aoki, S., Takemura, A. (2003b). The list of indispensable moves of the unique minimal Markov basis for $3 \times 4 \times K$ and $4 \times 4 \times 4$ contingency tables with fixed two-dimensional marginals. Technical Report METR 03-38, Department of Mathematical Engineering and Information Physics, The University of Tokyo (submitted for publication).

Aoki, S., Takemura, A. (2005). Markov chain Monte Carlo exact tests for incomplete two-way contingency tables. *Journal of Statistical Computation and Simulation*, *75*, 787–812.

Besag, J., Clifford, P. (1989). Generalized Monte Carlo significance tests. *Biometrika*, *76*, 633–642.

De Loera, J. A., Haws, D., Hemmecke, R., Huggins, P., Sturmfels, B., Yoshida, R. (2004). Short rational functions for toric algebra and applications. *Journal of Symbolic Computation*, *38*, 959–973.

Diaconis, P., Saloff-Coste, L. (1995). Random walk on contingency tables with fixed row and column sums. Technical Report, Department of Mathematics, Harvard University.

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

Dyer, M., Greenhill, C. (2000). Polynomial-time counting and sampling of two-rowed contingency tables. *Theoretical Computer Sciences*, *246*, 265–278.

Forster, J. J., McDonald, J. W., Smith, P. W. F. (1996). Monte Carlo exact conditional tests for log-linear and logistic models. *Journal of the Royal Statistical Society, Series B*, *58*, 445–453.

Guo, S. W., Thompson, E. A. (1992). Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrika*, *48*, 361–372.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, *57*, 97–109.

4ti2 team (2005). 4ti2 – A software package for algebraic, geometric and combinatorial problems on linear spaces. Available at www.4ti2.de

Hemmecke, R., Malkin, P. (2005). Computing generating sets of lattice ideals, arXiv: math.CO/0508359.

Hernek, D. (1998). Random generation of $2 \times n$ contingency tables. *Random Structures and Algorithms*, *13*, 71–79.

Lauritzen, N. (2005). Homogeneous Buchberger algorithms and Sullivant's computational commutative algebra challenge, arXiv: math.AC/0508287.

Mehta, C. R., Patel, N. R. (1983). A network algorithm for performing Fisher's exact test in $r \times c$ contingency tables. *Journal of the American Statistical Association*, *78*, 427–434.

Smith, P. W. F., Forster, J. J., McDonald, J. W. (1996). Monte Carlo exact tests for square contingency tables. *Journal of the Royal Statistical Society, Series A*, *159*, 309–321.

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

Suzuki, T., Aoki, S., Murota, K. (2005). Use of primal-dual technique in the network algorithm for two-way contingency tables. *Japan Journal of Industrial and Applied Mathematics*, *22*, 133–145.

Takemura, A., 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.