Bernoulli **9**(6), 2003, 1093–1108

Markov bases for decomposable graphical models

ADRIAN DOBRA

Institute of Statistics and Decision Sciences and Department of Molecular Genetics & Microbiology, Duke University, Durham, NC 27708-0251, USA. E-mail: adobra@stat.duke.edu

We show that primitive data swaps or moves are the only moves that have to be included in a Markov basis that links all the contingency tables having a set of fixed marginals when this set of marginals induces a decomposable independence graph. We give formulae that fully identify such Markov bases and show how to use these formulae to dynamically generate random moves.

Keywords: contingency tables; decomposable graphs; disclosure limitation; Gröbner bases; Markov bases; Markov chain Monte Carlo

1. Introduction

The problem of describing sets of multi-way contingency tables that are induced by fixing a number of lower-dimensional marginals has been the focus of many research efforts in recent years. These sets arise in a variety of contexts such as disclosure limitation (Dobra 2002; Fienberg *et al*. 1998; 2001) and the calibration of test statistics (Agresti 1992; Diaconis and Efron 1985; Mehta 1994). Diaconis and Sturmfels (1998) proposed a general algorithm for generating random draws from a set of tables with given fixed marginals. Their approach is extremely appealing because, in theory, it can be used for arrays of any dimension. Despite its generality, the power of this sampling procedure is limited because it requires access to a Markov basis – a finite set of data swaps which allow any two tables with the same fixed marginals to be connected. In addition to sampling, Markov bases can be employed to enumerate all the integer tables with a given set of marginals. As a consequence, Markov bases allow one to create a ‘replacement’ for a database consisting of a k-way contingency table, when such a replacement is needed to protect the individuals with rare characteristics whose identity might be disclosed by the release of a number of marginals (Willenborg and de Waal 2001).

Diaconis and Sturmfels (1998) and Dinwoodie (1998) suggest computing a Markov basis by finding a Gröbner basis (Cox *et al*. 1992) of a well-specified polynomial ideal, but their method is difficult to employ even for tables with three dimensions because of the computational complexity of computing Gröbner bases.

The statistical theory on graphical models (Madigan and York 1995; Whittaker 1990; Lauritzen 1996) shows that the conditional dependencies induced by a set of fixed marginals among the variables cross-classified in a table of counts can be visualized by means of an independence graph. In particular, a lot of attention has been given to

1350–7265 © 2003 ISI/BS

decomposable graphs (Lauritzen 1996), a special class of graphs that can be 'broken' into components such that (i) every component is associated with exactly one fixed marginal, and (ii) no information is lost in the decomposition process, that is, no marginal is 'split' between two components.

Our aim is to show how graphical models could help us identify special settings in which we could develop efficient techniques for considerably reducing and possibly eliminating the amount of computations needed to identify a Markov basis. After presenting some notation and definitions in Section 2, in Section 3 we introduce decomposable sets of marginals and discuss some of their properties. In Section 4 we prove that primitive data swaps or moves are the only moves that have to be included in a Markov basis associated with a set of decomposable marginals and give explicit formulae for dynamically generating such bases. In the last section we make some concluding remarks.

2. Data swapping and Markov bases

A table of counts $\mathbf{n}$ is a k-dimensional array of non-negative integer numbers. Each variable X_j, $j = 1, 2, \ldots, k$, recorded in such a table can take a finite number of values $x_j \in \mathcal{I}_j := \{1, 2, \ldots, I_j\}$. Let $\mathcal{I} = \mathcal{I}_1 \times \mathcal{I}_2 \times \ldots \times \mathcal{I}_k$. A *cell entry* $n(i)$, $i \in \mathcal{I}$, in table $\mathbf{n}$ is a non-negative integer representing the number of units or individuals sharing the same attributes i. By considering an ordering of the cell indices in $\mathcal{I}$ (e.g., lexicographic), the multi-way array $\mathbf{n}$ can be transformed into a linear list of counts (i.e., vector) $\bar{\mathbf{n}}$.

Let $D = \{i_1, i_2, \ldots, i_l\}$ denote an arbitrary subset of $K = \{1, 2, \ldots, k\}$. The D-marginal $\mathbf{n}_D$ of $\mathbf{n}$ is the contingency table with *marginal cells* $i_D \in \mathcal{I}_D := \mathcal{I}_{i_1} \times \ldots \times \mathcal{I}_{i_l}$ and entries given by

$$n_D(i_D) = \sum_{i \in \mathcal{I}_{K \setminus D}} n(i_D, i).$$

The marginals $\mathbf{n}_{D_1}$ and $\mathbf{n}_{D_2}$ are called *overlapping* if $D_1 \cap D_2 = \varnothing$, otherwise they are *non-overlapping*.

Two tables $\mathbf{n}^1$ and $\mathbf{n}^2$ are *equal* if all their cell entries are equal, and in this case we write $\mathbf{n}^1 = \mathbf{n}^2$. If all the counts in table $\mathbf{n}^1$ are zero, we write $\mathbf{n}^1 = \mathbf{0}$. The *sum* of two tables $\mathbf{n}^1$ and $\mathbf{n}^2$ is another table $\mathbf{n}^3 := \mathbf{n}^1 + \mathbf{n}^2$ with entries $n^3(i) = n^1(i) + n^2(i)$. Similarly, the *difference* between $\mathbf{n}^1$ and $\mathbf{n}^2$ is an array $\mathbf{n}^4 := \mathbf{n}^1 - \mathbf{n}^2$ with entries $n^4(i) = n^1(i) - n^2(i)$.

When moving table entries from one cell to the other, some of the cell entries could be increased and other cell entries could be decreased, hence a *data swap* or *move* associated with $\mathbf{n}$ is an array $\mathbf{f} = \{f(i)\}_{i \in \mathcal{I}}$ containing integer entries, that is, $f(i) \in \{\ldots, -2, -1, 0, 1, 2, \ldots\}$, for all $i \in \mathcal{I}$. A *primitive* move has two entries equal to 1, two entries equal to -1, while the remaining entries are 0. Intuitively, a move can be viewed as the difference between the post-swapped and the pre-swapped tables. The table created by repeatedly applying data swaps to the original table is sometimes required to be consistent with the marginals that were previously made public (Willenborg and de Waal

2001). Consequently, we are interested in data swaps that leave a number of marginals unchanged.

Definition 2.1. *Let $D_1, D_2, \ldots, D_r$ be subsets of K. A move $\mathbf{f}$ for $D_1, D_2, \ldots, D_r$ is a data swap that preserves the marginal tables specified by the index sets $D_1, D_2, \ldots, D_r$, that is, $\mathbf{f}_{D_j} = \mathbf{0}$ for all $j = 1, 2, \ldots, r$.*

Denote by $\mathbf{T}^{(\mathbf{n})}(D_1, \ldots, D_r)$ the set of all tables with non-negative elements that have their $\{D_1, D_2, \ldots, D_r\}$-marginals equal to the corresponding marginals of $\mathbf{n}$. A move $\mathbf{f}$ is *admissible* for $\mathbf{n}$ if $\mathbf{n} + \mathbf{f}$ belongs to $\mathbf{T}^{(\mathbf{n})}(D_1, \ldots, D_r)$. Since $\mathbf{f}$ preserves the $\{D_1, D_2, \ldots, D_r\}$-marginals of $\mathbf{n}$, we have $\mathbf{n} + \mathbf{f} \in \mathbf{T}^{(\mathbf{n})}(D_1, \ldots, D_r)$ if and only if $(\mathbf{n} + \mathbf{f})(i) \geqslant 0$, for all $i \in \mathcal{I}$.

Definition 2.2. *A* Markov basis *$\mathcal{M}$ is a finite collection of moves that preserve the $\{D_1, \ldots, D_r\}$-marginals and connect any two k-way tables that have the same $\{D_1, \ldots, D_r\}$-marginals. In other words, for any table $\mathbf{x}$ that belongs to $\mathbf{T}^{(\mathbf{n})}(D_1, \ldots, D_r)$, there exists a sequence of moves $\mathbf{f}^1, \mathbf{f}^2, \ldots, \mathbf{f}^s$ in $\mathcal{M}$ such that*

$$\mathbf{x} - \mathbf{n} = \sum_{j=1}^{s} \mathbf{f}^j, \text{ and } \quad \mathbf{n} + \sum_{j=1}^{s'} \mathbf{f}^j \in \mathbf{T}^{(\mathbf{n})}(D_1, \ldots, D_r), \tag{2.1}$$

for $1 \leqslant s' \leqslant s$. What (2.1) *says is that the table* $\mathbf{n}$ *is* transformed *into* $\mathbf{x}$ *by employing moves in $\mathcal{M}$. Since a Markov basis $\mathcal{M}$ depends only on the index sets $\mathcal{I}_{D_1}, \ldots, \mathcal{I}_{D_r}$, we will say that $\mathcal{M}$ is a Markov basis for $\mathbf{T}(D_1, \ldots, D_r)$, where*

$$\mathbf{T}(D_1, \ldots, D_r) = \{\mathbf{T}^{(\mathbf{n})}(D_1, \ldots, D_r): \mathbf{n} \text{ } \textit{is a table of counts}\}.$$

Diaconis and Sturmfels (1998) prove that a Markov basis $\mathcal{M}$ for $\mathbf{T}(D_1, \ldots, D_r)$ always exists.

3. Special configurations of marginals

In this section we closely follow the notation and definitions relating to graph theory introduced in Lauritzen (1996) and Dobra and Fienberg (2000). A brief introduction with the basic graph terminology needed to understand the results that follow is given in the Appendix.

Consider a set of marginals $\mathbf{n}_{D_1}, \mathbf{n}_{D_2}, \ldots, \mathbf{n}_{D_r}$ such that their index sets cover K, that is, $K = D_1 \cup D_2 \cup \ldots \cup D_r$. We always assume that there are no redundant configurations in this sequence, that is, there are no r_1, r_2 such that $D_{r_1} \subset D_{r_2}$. We visualize the dependency patterns induced by $D_1, D_2, \ldots, D_r$ by constructing an independence graph. Each vertex in this graph represents a variable X_j, $j \in K$. We draw an edge between two vertices if and only if the two-dimensional array defined by the variables associated with these vertices is a marginal of some $\mathbf{n}_{D_l}$.

Definition 3.1. *The* independence graph $\mathcal{G} = \mathcal{G}(D_1, D_2, \ldots, D_r)$ *associated with* $\mathbf{n}_{D_1}, \mathbf{n}_{D_2}, \ldots, \mathbf{n}_{D_r}$ *is a graph with vertex set* $K = D_1 \cup D_2 \cup \ldots \cup D_r$ *and edge set E given by*

$$E := \big\{(u, v) : \{u, v\} \subset D_j, \text{ for some } j \in \{1, \ldots, r\}\big\}.$$

Log-linear models are the usual way of representing and studying contingency tables with fixed marginals (Bishop *et al*., 1975). If the minimal sufficient statistics of a log-linear model define a decomposable independence graph, the model is said to be *decomposable*. By analogy with log-linear models theory, we introduce decomposable sets of marginals.

Definition 3.2. *The set of marginals* $\mathbf{n}_{D_1}, \mathbf{n}_{D_2}, \ldots, \mathbf{n}_{D_r}$ *is called* decomposable *if its corresponding independence graph* $\mathcal{G} = \mathcal{G}(D_1, D_2, \ldots, D_r)$ *is decomposable and the cliques* $\mathcal{C}(\mathcal{G})$ *of* $\mathcal{G}$ *are the index sets associated with* $\mathbf{n}_{D_1}, \mathbf{n}_{D_2}, \ldots, \mathbf{n}_{D_r}$, *that is,*

$$\mathcal{C}(\mathcal{G}) = \{D_1, D_2, \ldots, D_r\}.$$

Therefore a decomposable set of marginals could represent the minimal sufficient statistics of a decomposable log-linear model.

Definition 3.3. *The marginals* $\mathbf{n}_{D_1}, \mathbf{n}_{D_2}, \ldots, \mathbf{n}_{D_r}$ *are* consistent *if, for any* r_1, r_2, *the* $(D_{r_1} \cap D_{r_2})$*-marginal of* $\mathbf{n}_{D_{r_1}}$ *is equal to the* $(D_{r_1} \cap D_{r_2})$*-marginal of* $\mathbf{n}_{D_{r_2}}$.

The consistency of a set of marginals does not necessarily imply the existence of a table having this particular set of marginal totals – see, for example, Vlach (1986). To be more precise, $\mathbf{T}^{(\mathbf{n})}(D_1, \ldots, D_r)$ could be empty even if $\mathbf{n}_{D_1}, \mathbf{n}_{D_2}, \ldots, \mathbf{n}_{D_r}$ are consistent. In the special case of consistent and decomposable marginals, however, $\mathbf{T}^{(\mathbf{n})}(D_1, D_2, \ldots, D_r)$ is never empty (Dobra 2002).

Decomposable sets of marginals have many other exceptional properties that have been well documented in the literature. For example, the corresponding maximum likelihood estimates can be expressed in closed form (Lauritzen 1996; Whittaker 1990). Additionally, Dobra and Fienberg (2000) obtain formulae for calculating sharp upper and lower bounds for cell entries of tables in $\mathbf{T}^{(\mathbf{n})}(D_1, D_2, \ldots, D_r)$ given that the marginals are consistent and decomposable.

4. Markov bases for decomposable sets of marginals

The special structural properties of decomposable graphs can be further exploited to derive a Markov basis of primitive moves for $\mathbf{T}(D_1, \ldots, D_r)$ if $\mathbf{n}_{D_1}, \mathbf{n}_{D_2}, \ldots, \mathbf{n}_{D_r}$ is a decomposable set of marginals.

The simplest decomposable graph has two vertices and no edges. This graph is the independence graph associated with the two one-way marginals of a two-way table. It turns out that it is straightforward to describe a Markov basis in this case (Diaconis and Gangolli, 1995; Diaconis and Sturmfels, 1998).

Proposition 4.1. *Consider a two-way contingency table* $\mathbf{n} = \{n(i, j) : (i, j) \in \mathcal{I}_1 \times \mathcal{I}_2\}$ *with fixed row sums* $\mathbf{n}_1 := \{n_{i+} : i \in \mathcal{I}_1\}$ *and column sums* $\mathbf{n}_2 := \{n_{+j} : j \in \mathcal{I}_2\}$. *For some indices* i_1, i_2, j_1, j_2 *chosen such that* $1 \leqslant i_1 < i_2 \leqslant I_1$ *and* $1 \leqslant j_1 < j_2 \leqslant I_2$, *we define a table* $\mathbf{f}^{i_1 i_2; j_1 j_2} = \{f^{i_1 i_2; j_1 j_2}(i, j) : (i, j) \in \mathcal{I}_1 \times \mathcal{I}_2\}$ *by*

$$f^{i_1 i_2; j_1 j_2}(i, j) = \begin{cases} 1, & \text{if } (i, j) \in \{(i_1, j_1), (i_2, j_2)\}, \\ -1, & \text{if } (i, j) \in \{(i_1, j_2), (i_2, j_1)\}, \\ 0, & \textit{otherwise}. \end{cases} \tag{4.1}$$

Then

$$\left\{\pm \mathbf{f}^{i_1 i_2; j_1 j_2} : 1 \leqslant i_1 < i_2 \leqslant I_1, 1 \leqslant j_1 < j_2 \leqslant I_2\right\} \tag{4.2}$$

is a Markov basis for the class of tables with fixed row sums $\mathbf{n}_1$ *and column sums* $\mathbf{n}_2$.

Proof. The one-way marginals of $\mathbf{f}^{i_1 i_2; j_1 j_2}$ are zero, hence $\mathbf{f}^{i_1 i_2; j_1 j_2}$ will leave $\mathbf{n}_1$ and $\mathbf{n}_2$ unchanged. By making use of computational algebra techniques, Sturmfels (1995) gives a complete proof of the fact that the set of moves described in (4.2) is indeed a Markov basis. The number of moves in this Markov basis is $2 \cdot \binom{I_1}{2} \cdot \binom{I_2}{2}$. □

The set of primitive moves we described above allows one to transform a given two-way table into any other two-way table with the same row and column totals. Proposition 4.1 is the starting point for developing Markov bases for an arbitrary decomposable graphical structure. Consider the case of a k-way contingency table $\mathbf{n}$ with two fixed marginals $\mathbf{n}_{D_1}$ and $\mathbf{n}_{D_2}$. The corresponding independence graph $\mathcal{G}(D_1, D_2)$ is decomposable since it has exactly two cliques, D_1 and D_2.

We show that the Markov basis $\mathcal{F}(D_1, D_2)$ for $\mathbf{T}(D_1, D_2)$ is the union of the Markov bases of one or more two-way tables with fixed one-way marginals. We distinguish two cases. If $D_1 \cap D_2 = \emptyset$, the two fixed marginals are non-overlapping. Introduce two new variables Y_1 and Y_2 with level sets $\mathcal{I}_{D_1}$ and $\mathcal{I}_{D_2}$, respectively. Take the two-way table $\mathbf{n}'$ that cross-classifies Y_1 and Y_2. This table has fixed row sums $\overline{\mathbf{n}}_{D_1}$ and fixed column sums $\overline{\mathbf{n}}_{D_2}$. The basis $\mathcal{F}(D_1, D_2)$ for $\mathbf{T}(D_1, D_2)$ will be the Markov basis of moves for the two-way table $\mathbf{n}'$ as described in Proposition 4.1.

Otherwise, if the two fixed marginals are overlapping, we have $D_1 \cap D_2 \neq \emptyset$. For every $i^0_{D_1 \cap D_2} \in \mathcal{I}_{D_1 \cap D_2}$, define a table $\mathbf{n}^{i^0_{D_1 \cap D_2}} = \{n^{i^0_{D_1 \cap D_2}}(i)\}_{i \in \mathcal{I}_{K \setminus (D_1 \cap D_2)}}$ with entries $n^{i^0_{D_1 \cap D_2}}(i) = n(i, i^0_{D_1 \cap D_2})$. This table has two fixed marginals: $\mathbf{n}^{i^0_{D_1 \cap D_2}}_{D_1 \setminus D_2} = \left\{n_{D_1}(i, i^0_{D_1 \cap D_2})\right\}_{i \in \mathcal{I}_{D_1 \setminus D_2}}$ and $\mathbf{n}^{i^0_{D_1 \cap D_2}}_{D_2 \setminus D_1} = \left\{n_{D_2}(i, i^0_{D_1 \cap D_2})\right\}_{i \in \mathcal{I}_{D_2 \setminus D_1}}$.

The $D_1 \cap D_2 = \emptyset$ shows how to construct a Markov basis $\mathcal{F}^{i^0_{D_1 \cap D_2}}$ for $\mathbf{n}^{i^0_{D_1 \cap D_2}}$ that preserves the two non-overlapping marginals $\mathbf{n}^{i^0_{D_1 \cap D_2}}_{D_1 \setminus D_2}$ and $\mathbf{n}^{i^0_{D_1 \cap D_2}}_{D_2 \setminus D_1}$. It follows that a Markov basis of moves for table $\mathbf{n}$ that preserves the marginals $\mathbf{n}_{D_1}$ and $\mathbf{n}_{D_2}$ is given by

$$\mathcal{F}(D_1, D_2) = \bigcup_{i^0_{D_1 \cap D_2} \in \mathcal{I}_{D_1 \cap D_2}} \mathcal{F}^{i^0_{D_1 \cap D_2}}. \tag{4.3}$$

Therefore $\mathcal{F}(D_1, D_2)$ contains only primitive moves and represents a Markov basis for $\mathbf{T}(D_1, D_2)$.

Example 1. Consider a four-way table $\mathbf{n}$ with fixed three-way marginals $\mathbf{x} := \mathbf{n}_{\{1,2,3\}}$ and $\mathbf{y} := \mathbf{n}_{\{2,3,4\}}$. The corresponding independence graph $\mathcal{G}$ is represented in Figure 1. The edge $\{2, 3\}$ is a separator for $\{1, 2, 3\}$ and $\{2, 3, 4\}$. In addition, $\{1, 2, 3\}$ and $\{2, 3, 4\}$ are complete in $\mathcal{G}$, hence $\mathcal{G}$ is a decomposable graph with two cliques. Consider the set of contingency tables

$$\{\mathbf{n}^{i_2^0,i_3^0} = \{n^{i_2^0,i_3^0}(i_1, i_4) : (i_1, i_4) \in \mathcal{I}_1 \times \mathcal{I}_4\} : i_2^0 \in \mathcal{I}_2,\, i_3^0 \in \mathcal{I}_3\},$$

where $n^{i_2^0,i_3^0}(i_1, i_4) = n(i_1, i_2^0, i_3^0, i_4)$. For every table $\mathbf{n}^{i_2^0,i_3^0}$, we know its row and column sums: $\mathbf{n}_1^{i_2^0,i_3^0} := \{x(i_1, i_2^0, i_3^0) \colon i_1 \in \mathcal{I}_1\}$ and $\mathbf{n}_2^{i_2^0,i_3^0} := \{y(i_2^0, i_3^0, i_4) \colon i_4 \in \mathcal{I}_4\}$, respectively. The Markov basis $\mathcal{F}^{i_2^0,i_3^0}$ that leaves unchanged the one-way marginals of the table $\mathbf{n}^{i_2^0,i_3^0}$ can be obtained as in Proposition 4.1. A Markov basis of primitive moves that preserves the marginals $\mathbf{x}$ and $\mathbf{y}$ is the union $\mathcal{F}(\{1, 2, 3\}, \{2, 3, 4\}) = \{\mathcal{F}^{i_2^0,i_3^0} : (i_2^0, i_3^0) \in \mathcal{I}_2 \times \mathcal{I}_3\}$.

We introduce the set of primitive moves associated with an arbitrary decomposable graph $\mathcal{G}$.

Definition 4.1. *Let $\mathcal{C}(\mathcal{G}) = \{D_1, D_2, \ldots, D_r\}$ be the set of cliques of a decomposable graph $\mathcal{G}$. We let $\mathcal{T} = (\mathcal{C}(\mathcal{G}), \mathcal{E}_{\mathcal{T}})$ be a tree having the star property on the set of cliques of $\mathcal{G}$. For every edge $(D_j, D_i) \in \mathcal{E}_{\mathcal{T}}$, we consider the vertex sets V_j and V_i as in* (A.1). *The set of primitive moves associated with the decomposable graph $\mathcal{G}$ is given by*

$$\mathcal{F}(\mathcal{G}) = \mathcal{F}(D_1, D_2, \ldots, D_r) := \bigcup_{(D_j,D_i)\in\mathcal{E}_{\mathcal{T}}} \mathcal{F}(V_j, V_i), \tag{4.4}$$

where $\mathcal{F}(V_j, V_i)$ was defined in (4.3).

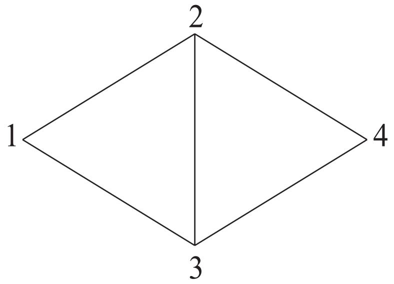

Figure 1. A decomposable graph with four vertices and two cliques

By removing an edge (D_j, D_i) from $\mathcal{T}$, we create two connected components $\mathcal{T}(V_j)$ and $\mathcal{T}(V_i)$. We think about V_j and V_i as being the cliques of a graph $\mathcal{G}^{ij}$ with vertices $V_j \cup V_i = K$ and edges

$$E_{ij} := \{(u, v) : \{u, v\} \subset V_j \text{ or } \{u, v\} \subset V_i\}.$$

The tree $\mathcal{T}$ has the star property, hence $S_{ij} := D_j \cap D_i$ separates $V_j \backslash S_{ij}$ from $V_i \backslash S_{ij}$ in $\mathcal{G}^{ij}$. As a result, $\mathcal{G}^{ij}$ is the independence graph of a decomposable model with two cliques and we know that the set of primitive moves corresponding to $\mathcal{G}^{ij}$ is $\mathcal{F}(V_j, V_i)$. Equation (4.4) essentially says that the set of primitive moves for a decomposable model with independence graph $\mathcal{G}$ is just the union of the sets of primitive moves associated with the two-clique models induced by each minimal vertex separator of $\mathcal{G}$. We have to show that Definition 4.1 is correct.

Proposition 4.2. *The set of primitive moves defined in* (4.4) *is indeed a set of moves for the class of tables* $\mathbf{T}(D_1, D_2, \ldots, D_r)$.

Proof. Let $\mathbf{f} \in \mathcal{F}(D_1, D_2, \ldots, D_r)$. Then $\mathbf{f} \in \mathcal{F}(V_j, V_i)$ for some $(D_j, D_i) \in \mathcal{E}_{\mathcal{T}}$. For any arbitrary clique $D \in \mathcal{C}(\mathcal{G})$, we have either $D \subset V_j$ or $D \subset V_i$. Since $\mathbf{f}_{V_j} = \mathbf{0}$ and $\mathbf{f}_{V_i} = \mathbf{0}$, it follows that we also have $\mathbf{f}_D = \mathbf{0}$. □

Next we will state and prove a series of results that will help us prove the main theorem of the paper. Most of these propositions should be self-explanatory. However, it is worth mentioning the intuition that triggered them: if we delete a vertex that belongs to exactly one clique from a decomposable graph, along with the edges incident to it, we obtain a graph that is still decomposable (Blair and Barry 1993). Consequently, by collapsing across a variable associated with such a vertex, all the conditional dependencies existing among the remaining variables are preserved.

The set of primitive moves associated with a two-clique model induces a set of primitive moves for a two-clique model embedded in it. Collapsing across some of the variables not contained in both cliques preserves the structure of the moves in (4.4).

Proposition 4.3. *Let* $\mathbf{n}$ *be a table with two fixed marginals* $\mathbf{n}_{D_1}$ *and* $\mathbf{n}_{D_2}$. *The corresponding independence graph* $\mathcal{G}$ *is decomposable and has two cliques* D_1, D_2. *The separator of* $\mathcal{G}$ *is* $S := D_1 \cap D_2$. *Consider a vertex set* D *such that* $S \subset D \subset D_1$. *Define a map* ϕ *which assigns to every* $\mathbf{f} \in \mathcal{F}(D_1, D_2)$ *its* $(D \cup D_2)$*-marginal, that is,* $\phi(\mathbf{f}) = \mathbf{f}_{D \cup D_2}$. *Then the following are true:*

(a) *For any* $\mathbf{f} \in \mathcal{F}(D_1, D_2)$, $\phi(\mathbf{f}) \in \mathcal{F}(D, D_2)$ *or* $\phi(\mathbf{f}) = \mathbf{0}$.
(b) *The map* $\phi : \mathcal{F}(D_1, D_2) \rightarrow \mathcal{F}(D, D_2)$ *is surjective.*
(c) *For every table* $\mathbf{x} \in \mathbf{T}^{(\mathbf{n})}(D_1, D_2)$ *and every move* $\mathbf{g} \in \mathcal{F}(D, D_2)$ *such that*

$$\mathbf{x}_{D \cup D_2} + \mathbf{g} \in \mathbf{T}^{(\mathbf{n})}(D, D_2), \tag{4.5}$$

there exists $\mathbf{f} \in \mathcal{F}(D_1, D_2)$ *with* $\phi(\mathbf{f}) = \mathbf{g}$ *and*

$$\mathbf{x} + \mathbf{f} \in \mathbf{T}^{(\mathbf{n})}(D_1, D_2). \tag{4.6}$$

Proof. To simplify the notation, assume that $S = \varnothing$. We consider the marginals $\mathbf{n}_{D_1}$, $\mathbf{n}_{D_2}$ and $\mathbf{n}_D$, along with their associated vectors $\overline{\mathbf{n}}_{D_1}$, $\overline{\mathbf{n}}_{D_2}$ and $\overline{\mathbf{n}}_D$. The table $\mathbf{n}_D$ can be obtained from $\mathbf{n}_{D_1}$ by collapsing across the variables in $D_1 \setminus D$.

(a) In Proposition 4.1, we constructed $\mathcal{F}(D_1, D_2)$ by considering the two-way table with row marginal $\overline{\mathbf{n}}_{D_1}$ and column marginal $\overline{\mathbf{n}}_{D_2}$. A primitive move $\mathbf{f} \in \mathcal{F}(D_1, D_2)$ was obtained by choosing two 'row' indices $i^1_{D_1}$ and $i^2_{D_1}$, and two 'column' indices $i^1_{D_2}$ and $i^2_{D_2}$. Then the table $\mathbf{f}$ is given by

$$f(i_{D_1}, i_{D_2}) = \begin{cases} \pm 1, & \text{if } (i_{D_1}, i_{D_2}) \in \{(i^1_{D_1}, i^1_{D_2}), (i^2_{D_1}, i^2_{D_2})\}, \\ \mp 1, & \text{if } (i_{D_1}, i_{D_2}) \in \{(i^1_{D_1}, i^2_{D_2}), (i^2_{D_1}, i^1_{D_2})\}, \\ 0, & \text{otherwise.} \end{cases}$$

Let $\mathbf{f}_1 = \phi(\mathbf{f})$. We have

$$f_1(i_{D \cup D_2}) = f_1(i_D, i_{D_2}) = \sum_{j \in \mathcal{I}_{D_1 \setminus D}} f(j, i_D, i_{D_2}).$$

We distinguish two cases.

(i) $i^1_D = i^2_D$. Since $i^1_{D_1} \neq i^2_{D_1}$, we need to have $i^1_{D_1 \setminus D} \neq i^2_{D_1 \setminus D}$. It follows that

$$f_1(i^1_D, i^r_{D_2}) = f(i^1_{D_1}, i^r_{D_2}) + f(i^2_{D_1}, i^r_{D_2}) = 0, \qquad \text{for } r = 1, 2.$$

Clearly, $f_1(i^1_D, i_{D_2}) = 0$ if $i_{D_2} \notin \{i^1_{D_2}, i^2_{D_2}\}$. Moreover, for $i_D \neq i^1_D$, $f_1(i_D, i_{D_2}) = 0$. Hence $\phi(\mathbf{f}) = \mathbf{f}_1 = \mathbf{0}$.

(ii) $i^1_D \neq i^2_D$. It follows that

$$f_1(i_D, i_{D_2}) = \begin{cases} f(i^{r_1}_{D_1}, i^{r_2}_{D_2}), & \text{if } (i_D, i_{D_2}) = (i^{r_1}_D, i^{r_2}_{D_2}), \text{ where } r_1, r_2 \in \{1, 2\}, \\ 0, & \text{otherwise.} \end{cases}$$

Thus $\phi(\mathbf{f}) = \mathbf{f}_1 \in \mathcal{F}(D, D_2)$.

(b) In order to prove that ϕ is surjective, we pick an arbitrary move $\mathbf{g} \in \mathcal{F}(D, D_2)$. We choose an index $i^0_{D_1 \setminus D} \in \mathcal{I}_{D_1 \setminus D}$ and define the move $\mathbf{f} = \{f(i)\}_{i \in \mathcal{I}_{D_1 \cup D_2}}$ by

$$f(i) = f(i_{D_1 \setminus D}, i_{D \cup D_2}) := \begin{cases} g(i_{D \cup D_2}), & \text{if } i_{D_1 \setminus D} = i^0_{D_1 \setminus D}, \\ 0, & \text{otherwise.} \end{cases}$$

It is easy to see that $\mathbf{f} \in \mathcal{F}(D_1, D_2)$ and $\phi(\mathbf{f}) = \mathbf{g}$.

(c) The move $\mathbf{g} \in \mathcal{F}(D, D_2)$ is given by

$$g(i_D, i_{D_2}) = \begin{cases} 1, & \text{if } (i_D, i_{D_2}) \in \{(i^1_D, i^1_{D_2}), (i^2_D, i^2_{D_2})\}, \\ -1, & \text{if } (i_D, i_{D_2}) \in \{(i^1_D, i^2_{D_2}), (i^2_D, i^1_{D_2})\}, \\ 0, & \text{otherwise,} \end{cases}$$

where $i^1_D, i^2_D \in \mathcal{I}_D$ and $i^1_{D_2}, i^2_{D_2} \in \mathcal{I}_{D_2}$. A move $\mathbf{f} \in \mathcal{F}(D_1, D_2)$ such that $\mathbf{f}_{D \cup D_2} = \mathbf{g}$ is obtained by choosing two indices $i^1_{D_1 \setminus D}$ and $i^2_{D_1 \setminus D}$ in $\mathcal{I}_{D_1 \setminus D}$. Then

$$f(i) = f(i_{D_1}, i_{D_2}) = \begin{cases} 1, & \text{if } i \in \{(i^1_{D_1\setminus D}, i^1_D, i^1_{D_2}), (i^2_{D_1\setminus D}, i^2_D, i^2_{D_2})\}, \\ -1, & \text{if } i \in \{(i^1_{D_1\setminus D}, i^1_D, i^2_{D_2}), (i^2_{D_1\setminus D}, i^2_D, i^1_{D_2})\}, \\ 0, & \text{otherwise.} \end{cases} \tag{4.7}$$

For any $i^1_{D_1\setminus D}$, $i^2_{D_1\setminus D}$ in $\mathcal{I}_{D_1\setminus D}$, the corresponding move $\mathbf{f}$ defined in (4.7) satisfies $(\mathbf{x}+\mathbf{f})_{D_l} = \mathbf{x}_{D_l} = \mathbf{n}_{D_l}$ for $l = 1, 2$, and $(x+f)(i) \geqslant 0$ for every $i \in \mathcal{I}\setminus\{(i^1_{D_1\setminus D}, i^1_D, i^2_{D_2}), (i^2_{D_1\setminus D}, i^2_D, i^1_{D_2})\}$. Therefore we have to choose $i^1_{D_1\setminus D}$, $i^2_{D_1\setminus D}$ such that

$$\begin{aligned} (x+f)(i^1_{D_1\setminus D}, i^1_D, i^2_{D_2}) &= x(i^1_{D_1\setminus D}, i^1_D, i^2_{D_2}) - 1 \geqslant 0, \\ (x+f)(i^2_{D_1\setminus D}, i^2_D, i^1_{D_2}) &= x(i^2_{D_1\setminus D}, i^2_D, i^1_{D_2}) - 1 \geqslant 0. \end{aligned} \tag{4.8}$$

In this case, (4.6) holds. From (4.5), we obtain that

$$(x_{D\cup D_2} + g)(i^1_D, i^2_{D_2}) \geqslant 0, \qquad (x_{D\cup D_2} + g)(i^2_D, i^1_{D_2}) \geqslant 0. \tag{4.9}$$

But

$$\begin{aligned} (x_{D\cup D_2} + g)(i^1_D, i^2_{D_2}) &= x_{D\cup D_2}(i^1_D, i^2_{D_2}) - 1, \\ &= \sum_{j_{D_1\setminus D}\in\mathcal{I}_{D_1\setminus D}} x(j_{D_1\setminus D}, i^1_D, i^2_{D_2}) - 1. \end{aligned} \tag{4.10}$$

Inequalities (4.9) and equation (4.10) imply that

$$\sum_{j_{D_1\setminus D}\in\mathcal{I}_{D_1\setminus D}} x(j_{D_1\setminus D}, i^1_D, i^2_{D_2}) \geqslant 1,$$

hence there has to exist an index $i^1_{D_1\setminus D} \in \mathcal{I}_{D_1\setminus D}$ with $x(i^1_{D_1\setminus D}, i^1_D, i^2_{D_2}) \geqslant 1$. Similarly, there has to exist another index $i^2_{D_1\setminus D} \in \mathcal{I}_{D_1\setminus D}$ with $x(i^2_{D_1\setminus D}, i^2_D, i^1_{D_2}) \geqslant 1$. With this choice, (4.8) holds. □

The following proposition extends Proposition 4.3 to an arbitrary decomposable model.

Proposition 4.4. *Let* $\mathbf{n}$ *be a table with fixed marginals* $\mathbf{n}_{D_1}, \ldots, \mathbf{n}_{D_r}$ *such that* $\mathcal{C}(\mathcal{G}) = \{D_1, \ldots, D_r\}$ *is the set of cliques of a decomposable graph* $\mathcal{G}$. *Consider a tree* $\mathcal{T} = (\mathcal{C}(\mathcal{G}), \mathcal{E}_{\mathcal{T}})$ *having the star property for* $\mathcal{G}$. *Assume that the clique* D_r *is terminal in* $\mathcal{T}$ *and let* $A := \bigcup_{j=1}^{r-1} D_j$. *Define a map* ϕ *which assigns to every* $\mathbf{f} \in \mathcal{F}(D_1, D_2, \ldots, D_r)$ *its* A*-marginal, i.e.* $\phi(\mathbf{f}) = \mathbf{f}_A$. *Then the following are true:*

(a) *For any* $\mathbf{f} \in \mathcal{F}(D_1, D_2, \ldots, D_r)$, $\phi(\mathbf{f}) \in \mathcal{F}(D_1, \ldots, D_{r-1})$ *or* $\phi(\mathbf{f}) = \mathbf{0}$.
(b) *The map* ϕ *is surjective on* $\mathcal{F}(D_1, \ldots, D_{r-1})$.
(c) *For every table* $\mathbf{x} \in \mathbf{T}^{(\mathbf{n})}(D_1, D_2, \ldots, D_r)$ *and every move* $\mathbf{g} \in \mathcal{F}(D_1, \ldots, D_{r-1})$ *such that*

$$\mathbf{x}_A + \mathbf{g} \in \mathbf{T}^{(\mathbf{n})}(D_1, D_2, \ldots, D_{r-1}), \tag{4.11}$$

there exists $\mathbf{f} \in \mathcal{F}(D_1, D_2, \ldots, D_r)$ *with* $\phi(\mathbf{f}) = \mathbf{g}$ *and*

$$\mathbf{x} + \mathbf{f} \in \mathbf{T}^{(\mathbf{n})}(D_1, D_2, \ldots, D_r). \tag{4.12}$$

Proof. (a) Since the clique D_r is terminal in $\mathcal{T}$, there exists a unique clique in $\mathcal{C}(\mathcal{G})$, say D', such that $(D_r, D') \in \mathcal{E}_{\mathcal{T}}$. The set of primitive moves corresponding to the edge (D_r, D') is $\mathcal{F}(A, D_r)$, and we assume that $\mathbf{f} \in \mathcal{F}(A, D_r)$. By definition, $\mathbf{f}_A = \mathbf{0}$, hence $\phi(\mathbf{f}) = \mathbf{0}$.

The subgraph $\mathcal{G}' = \mathcal{G}(D_1 \cup \ldots \cup D_{r-1})$ is decomposable, and $\mathcal{C}(\mathcal{G}') = \{D_1, \ldots, D_{r-1}\}$. Let $\mathcal{T}'$ be the subtree obtained by removing D_r from $\mathcal{T}$, that is, $\mathcal{T}' = (\mathcal{C}(\mathcal{G}'), \mathcal{E}_{\mathcal{T}} \backslash \{(D_r, D')\})$. Consider an arbitrary edge $(D_j, D_i) \in \mathcal{E}_{\mathcal{T}} \backslash \{(D_r, D')\}$. Let $\mathcal{T}_j = (\mathcal{K}_j, \mathcal{E}_j)$ and $\mathcal{T}_i = (\mathcal{K}_i, \mathcal{E}_i)$ be the two subtrees obtained by removing the edge (D_j, D_i) from $\mathcal{T}$, with $D_j \in \mathcal{K}_j$ and $D_i \in \mathcal{K}_i$. Without loss of generality, we assume that we always have $D_r \in \mathcal{K}_j$.

By removing the same edge from the tree $\mathcal{T}'$, we obtain the subtrees $\mathcal{T}'_j = (\mathcal{K}_j \backslash \{D_r\}, \mathcal{E}_j \backslash \{(D_r, D')\})$ and $\mathcal{T}_i$. We define the vertex sets V_j, V'_j and V_i by

$$V_j := \bigcup_{D \in \mathcal{K}_j} D, \qquad V'_j := \bigcup_{D \in \mathcal{K}_j \backslash \{D_r\}} D, \qquad V_i := \bigcup_{D \in \mathcal{K}_i} D.$$

With this notation, according to Lemma A.1, the tree $\mathcal{T}'$ will have the star property for the graph $\mathcal{G}'$, and consequently the set of primitive primitive associated with $\mathcal{G}'$ is

$$\mathcal{F}(\mathcal{G}') = \mathcal{F}(D_1, \ldots, D_{r-1}) = \bigcup_{(D_j, D_i) \in \mathcal{E}_{\mathcal{T}} \backslash \{(D_r, D')\}} \mathcal{F}(V'_j, V_i). \tag{4.13}$$

Consider an arbitrary move $\mathbf{f} \in \mathcal{F}(D_1, D_2, \ldots, D_r)$ such that $\mathbf{f} \notin \mathcal{F}(A, D_r)$. From (4.4), we see that there must exist some edge $(D_j, D_i) \in \mathcal{E}_{\mathcal{T}} \backslash \{(D_r, D')\}$ such that $\mathbf{f} \in \mathcal{F}(V_j, V_i)$. We have $D_j \neq D_r$ and $D_j \subset V_j$, thus $V'_j \neq \emptyset$. In addition, we have $V'_j \subset V_j$ and $A = V'_j \cup V_i$. By employing Proposition 4.3, we obtain that $\phi(\mathbf{f}) \in \mathcal{F}(V'_j, V_i) \subset \mathcal{F}(D_1, \ldots, D_{r-1})$ or $\phi(\mathbf{f}) = \mathbf{0}$.

(b) In order to prove that ϕ is surjective on $\mathcal{F}(D_1, \ldots, D_{r-1})$, we pick an arbitrary move $\mathbf{g}$ in $\mathcal{F}(D_1, \ldots, D_{r-1})$. From (4.13), we see that there is an edge $(D_j, D_i) \in \mathcal{E}_{\mathcal{T}} \backslash \{(D_r, D')\}$ such that $\mathbf{g} \in \mathcal{F}(V'_j, V_i)$. Since $V'_j \subset V_j$, Proposition 4.3 tells us that there must exist some $\mathbf{f} \in \mathcal{F}(V_j, V_i) \subset \mathcal{F}(D_1, \ldots, D_r)$ such that $\phi(\mathbf{f}) = \mathbf{g}$.

(c) Again, (4.13) shows that it is possible to find an edge $(D_j, D_i) \in \mathcal{E}_{\mathcal{T}} \backslash \{(D_r, D')\}$ such that $\mathbf{g} \in \mathcal{F}(V'_j, V_i)$. This means that $\mathbf{x}_A + \mathbf{g} \in \mathbf{T}^{(\mathbf{x})}(V'_j, V_i)$. We have $V'_j \subset V_i$ and $V'_j \cap V_i = V_j \cap V_i$. From Proposition 4.3, we learn that there exists a move $\mathbf{f} \in \mathcal{F}(V_j, V_i) \subset \mathcal{F}(D_1, D_2, \ldots, D_r)$ such that $\mathbf{x} + \mathbf{f} \in \mathbf{T}^{(\mathbf{x})}(V_j, V_i)$. But we also have $\mathbf{T}^{(\mathbf{x})}(V_j, V_i) \subset \mathbf{T}^{(\mathbf{x})}(D_1, D_2, \ldots, D_r) = \mathbf{T}^{(\mathbf{n})}(D_1, D_2, \ldots, D_r)$, hence (4.12) is true. □

We are now ready to present and prove the main theorem of the paper.

Theorem 4.5. *Let $\mathcal{G}$ be a decomposable graph with cliques $\mathcal{C}(\mathcal{G}) = \{D_1, D_2, \ldots, D_r\}$. Then the set of primitive moves $\mathcal{F}(\mathcal{G}) = \mathcal{F}(D_1, D_2, \ldots, D_r)$ defined in* (4.4) *is a Markov basis for the class of tables $\mathbf{T}(D_1, D_2, \ldots, D_r)$.*

Proof. The proof is by induction. If $\mathcal{G}$ decomposes in $r = 2$ cliques, then we know from

Proposition 4.1 that $\mathcal{F}(D_1, D_2)$ is a Markov basis for $\mathbf{T}(D_1, D_2)$. Suppose the theorem holds for any decomposable graph with $r-1$ cliques. We want to prove that the theorem is true for a decomposable graph with r cliques.

The original table $\mathbf{n}$ is in the set $\mathbf{T}^{(\mathbf{n})} = \mathbf{T}^{(\mathbf{n})}(D_1, D_2, \ldots, D_r)$. Take an arbitrary table $\mathbf{x} \in \mathbf{T}^{(\mathbf{n})}$. We have to show that there exist $\mathbf{f}^1, \ldots, \mathbf{f}^l \in \mathcal{F}(D_1, D_2, \ldots, D_r)$ such that $\mathbf{x} - \mathbf{n} = \sum_{i=1}^{l} \mathbf{f}^i$, and

$$\mathbf{n} + \sum_{i=1}^{l'} \mathbf{f}^i \in \mathbf{T}^{(\mathbf{n})}(D_1, D_2, \ldots, D_r), \tag{4.14}$$

for $1 \leqslant l' \leqslant l$. Let $\mathcal{T} = (\mathcal{C}(\mathcal{G}), \mathcal{E}_{\mathcal{T}})$ be a tree having the star property for $\mathcal{G}$, and assume that the clique D_r is terminal in $\mathcal{T}$. Denote $A := \bigcup_{j=1}^{r-1} D_j$. Consider the map ϕ which assigns to every $\mathbf{f} \in \mathcal{F}(D_1, D_2, \ldots, D_r)$ its A-marginal, i.e. $\phi(\mathbf{f}) = \mathbf{f}_A$.

The marginals $\mathbf{n}_A$ and $\mathbf{x}_A$ lie in the set $\mathbf{T}^{(\mathbf{n})}(D_1, \ldots, D_{r-1})$. From the induction hypothesis we know that $\mathcal{F}(D_1, \ldots, D_{r-1})$ is a Markov basis for $\mathbf{T}^{(\mathbf{n})}(D_1, \ldots, D_{r-1})$, so there exists a sequence of moves $\mathbf{g}^1, \ldots, \mathbf{g}^{l_1} \in \mathcal{F}(D_1, \ldots, D_{r-1})$ such that

$$\mathbf{x}_A - \mathbf{n}_A = \sum_{i=1}^{l_1} \mathbf{g}^i \quad \text{and} \quad \mathbf{n}_A + \sum_{i=1}^{l_1'} \mathbf{g}^i \in \mathbf{T}^{(\mathbf{n})}(D_1, D_2, \ldots, D_{r-1}),$$

for $1 \leqslant l_1' \leqslant l_1$. Proposition 4.4 tells us that the sequence of moves $\mathbf{g}^1, \ldots, \mathbf{g}^{l_1}$ translates into another sequence of moves $\mathbf{f}^1, \ldots, \mathbf{f}^{l_1}$ in $\mathcal{F}(D_1, D_2, \ldots, D_r)$ such that, for every $1 \leqslant l_1' \leqslant l_1$, we have $\mathbf{f}_A^{l_1'} = \mathbf{g}^{l_1'}$, and

$$\mathbf{n} + \sum_{i=1}^{l_1'} \mathbf{f}^i \in \mathbf{T}^{(\mathbf{n})}(D_1, D_2, \ldots, D_r). \tag{4.15}$$

We obtain a table $\mathbf{x}' \in \mathbf{T}^{(\mathbf{n})}(D_1, D_2, \ldots, D_r)$, given by

$$\mathbf{x}' - \mathbf{n} = \sum_{i=1}^{l_1} \mathbf{f}^i, \tag{4.16}$$

such that the marginals $\mathbf{x}_A'$ and $\mathbf{x}_A$ are the same. Moreover, since we employed moves in $\mathcal{F}(D_1, \ldots, D_r)$, the marginals $\mathbf{x}_{D_r}'$ and $\mathbf{n}_{D_r}$ are also equal, and hence $\mathbf{x}' \in \mathbf{T}^{(\mathbf{x})}(A, D_r)$. This implies that we can find a series of moves $\mathbf{f}^{l_1+1}, \ldots, \mathbf{f}^l$ in $\mathcal{F}(A, D_r)$ which transform the table $\mathbf{x}'$ into $\mathbf{x}$, that is,

$$\mathbf{x} - \mathbf{x}' = \sum_{i=l_1+1}^{l} \mathbf{f}^i, \qquad \mathbf{x}' + \sum_{i=l_1+1}^{l'} \mathbf{f}^i \in \mathbf{T}^{(\mathbf{x})}(A, D_r) \subset \mathbf{T}^{(\mathbf{n})}(D_1, D_2, \ldots, D_r), \tag{4.17}$$

for $1 \leqslant l' \leqslant l$. From (4.15)–(4.17) we obtain (4.14), which completes the proof. □

Example 2. The graph $\mathcal{G}$ in Figure 2 has 11 vertices and 28 edges. This is a decomposable graph with the set of cliques $\mathcal{C}(\mathcal{G}) = \{D_1, D_2, D_3, D_4\}$, where $D_1 := \{1, 3, 4, 11\}$, $D_2 := \{3, 4, 7, 8, 9, 11\}$, $D_3 := \{2, 3, 9, 10\}$ and $D_4 := \{4, 5, 6, 7\}$. The tree $\mathcal{T}$ on $\mathcal{C}(\mathcal{G})$ with edge set

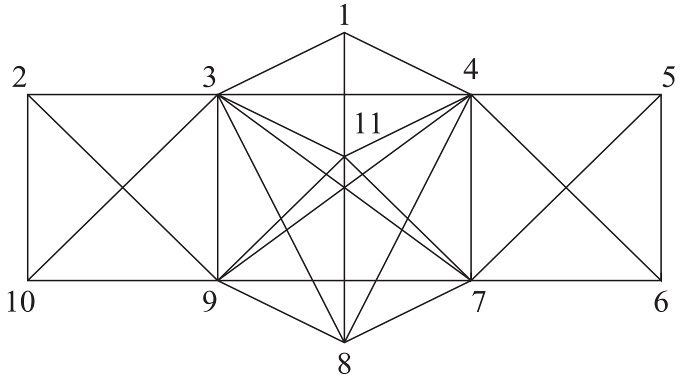

Figure 2. A decomposable graph with eleven vertices and four cliques

$$\mathcal{E}_{\mathcal{T}} = \{(D_2, D_1), (D_3, D_2), (D_4, D_2)\},$$

has the star property, therefore the separators of $\mathcal{G}$ are $S_2 := D_2 \cap D_1 = \{3, 4, 11\}$, $S_3 := D_3 \cap D_2 = \{3, 9\}$, and $S_4 := D_4 \cap D_2 = \{4, 7\}$. The set of primitive moves associated with $\mathcal{G}$ is

$$\mathcal{F}(\mathcal{G}) = \mathcal{F}(D_1, D_2 \cup D_3 \cup D_4) \cup \mathcal{F}(D_3, D_1 \cup D_2 \cup D_4) \cup \mathcal{F}(D_4, D_1 \cup D_2 \cup D_3).$$

Assume we are given an eleven-way table $\mathbf{n}$ with fixed marginals $\mathbf{n}_{D_1}$, $\mathbf{n}_{D_2}$, $\mathbf{n}_{D_3}$ and $\mathbf{n}_{D_4}$. The independence graph associated with these marginals is $\mathcal{G}$. Theorem 4.5 shows that $\mathcal{F}(\mathcal{G})$ is a Markov basis for $\mathbf{T}(D_1, D_2, D_3, D_4)$.

The family of Markov bases we identified is extremely appealing to the potential user since one does not even need to actually list the set of moves $\mathcal{F}(D_1, D_2, \dots, D_r)$. Any Markov basis could grow extremely large due to the size of the original table $\mathbf{n}$, hence handling it might become quite problematic. The procedure we outline below gets around this obstacle by dynamically generating moves in $\mathcal{F}(D_1, D_2, \dots, D_r)$. The first step consists of computing the number of moves associated with every edge of the tree $\mathcal{T}$. We uniformly generate a primitive move in $\mathcal{F}(D_1, D_2, \dots, D_r)$ by choosing an edge in $\mathcal{E}_{\mathcal{T}}$ with probability proportional to the number of primitive moves associated with it, then uniformly selecting a move from the set of primitive moves corresponding to the edge we picked.

Algorithm 4.6. *Let $\mathcal{T} = (\mathcal{C}(\mathcal{G}), \mathcal{E}_{\mathcal{T}})$ be a tree having the star property for $\mathcal{G}$. The set of separators $\mathcal{S}(\mathcal{G}) = \{S_2, \dots, S_r\}$ associated with $\mathcal{C}(\mathcal{G}) = \{D_1, \dots, D_r\}$ will be given by $\mathcal{S}(\mathcal{G}) = \{D_j \cap D_j : (D_j, D_i) \in \mathcal{E}_{\mathcal{T}}\}$.*

1. *For every $S_l \in \mathcal{S}(\mathcal{G})$:*

   (a) *Let $(D_j, D_i) \in \mathcal{E}_{\mathcal{T}}$ with $S_l = D_j \cap D_i$. Consider the subtrees $\mathcal{T}_j$ and $\mathcal{T}_i$ obtained* *by removing the edge (D_j, D_i) from $\mathcal{T}$, and let V_j and V_i be the vertex sets associated with these subtrees, as defined in (A.1).*

(b) *Calculate the weight w_l representing the number of primitive moves corresponding to the edge (D_j, D_i):*

$$w_l \leftarrow \left[2 \cdot \prod_{v \in V_j \setminus S_l} \binom{I_v}{2} \cdot \prod_{v \in V_i \setminus S_l} \binom{I_v}{2} \right]^{\prod_{v \in S_l} I_v}.$$

2. *Normalize the weights $w_2, \ldots, w_r$.*
3. *To uniformly select a move in $\mathcal{F}(\mathcal{G})$:*
   (a) *Randomly select an edge $(D_j, D_i) \in \mathcal{E}_{\mathcal{T}}$ with probability $P(S_l) = w_l$, where $S_l = D_j \cap D_i$.*
   (b) *Uniformly pick a move in $\mathcal{F}(V_j, V_i)$, where V_j and V_i were defined in* (A.1).

5. Conclusions

Many techniques that work well for low-dimensional examples are almost impossible to use for problems that arise in practice due to the huge computational effort they usually require. This paper demonstrates that graphical modelling is a very powerful tool for effectively overcoming major issues related to scaling up algorithms to make them suitable for use in high-dimensional applications. We represented the dependency patterns induced by a number of fixed marginals by means of graphs and, by doing so, we identified Markov bases for an entire family of sets of tables. We proved that a Markov basis for a decomposable model with r cliques can be expressed as a union of Markov bases associated with $r-1$ models with two cliques. Since the Markov basis of a model with two cliques is the set of primitive moves corresponding with one or more two-way tables with fixed one-way marginals, we deduce that the general decomposable case essentially reduces to the two-way case.

It seems important to point out that more results can be derived by exploiting techniques borrowed from the graphical models literature, namely decompositions of graphs by means of separators. Dobra and Sullivant (2003) have developed a divide-and-conquer algorithm for significantly reducing the time needed to find a Markov basis when the underlying independence graph is not decomposable, but can be at least partially decomposed, though the resulting components of the decomposition may correspond to more than one fixed marginal.

Appendix. Graph Theory

A *graph* $\mathcal{G}$ is a pair (K, E), where $K = \{1, 2, \ldots, k\}$ is a finite set of vertices and $E \subseteq K \times K$ is a set of edges linking the vertices. Our interest is in *undirected* graphs for which $(u, v) \in E$ implies $(v, u) \in E$. For any vertex set $A \subseteq K$, we define the edge set associated with it as

$$E(A) := \{(u, v) \in E | u, v \in A\}.$$

Let $\mathcal{G}(A) = (A, E(A))$ denote the subgraph of $\mathcal{G}$ induced by A. Two vertices $u, v \in K$ are *adjacent* if $(u, v) \in E$. A set of vertices in $\mathcal{G}$ is *independent* if no two of its elements are adjacent. An induced subgraph $\mathcal{G}(A)$ is *complete* if the vertices in A are pairwise adjacent in $\mathcal{G}$. We also say that A is *complete* in $\mathcal{G}$. A complete vertex set A in $\mathcal{G}$ that is maximal is a *clique*.

Let $u, v \in K$. A *path* (or *chain*) from u to v is a sequence $u = v_0, \ldots, v_n = v$ of distinct vertices such that $(v_{i-1}, v_i) \in E$ for all $i = 1, 2, \ldots, n$. The path is a *cycle* if the end points are allowed to be the same, $u = v$. If there is a path from u to v we say that u and v are *connected*. The sets $A, B \subset K$ are *disconnected* if u and v are not connected for all $u \in A$, $v \in B$. The *connected component* of a vertex $u \in K$ is the set of all vertices connected with u. A graph is *connected* if all the pairs of vertices are connected.

The set $C \subset K$ is a *uv-separator* if all paths from u to v intersect C. The set $C \subset K$ *separates* A from B if it is a uv-separator for every $u \in A$, $v \in B$. C is a *separator* of $\mathcal{G}$ if two vertices in the same connected component of $\mathcal{G}$ are in two distinct connected components of $\mathcal{G} \backslash C$ or, equivalently, if $\mathcal{G} \backslash C$ is disconnected. In addition, C is a *minimal* separator of $\mathcal{G}$ if C is a separator and no proper subset of C separates the graph. Unless otherwise stated, the separators we work with will be complete.

A *tree* is a connected graph with no cycles. In a tree, there is a unique path between any two vertices. The vertex u is called *terminal* in a tree if there is only one edge linking u with the remaining vertices.

Definition A.1. *The partition (A_1, S, A_2) of K is said to form a* decomposition *of $\mathcal{G}$ if S is a minimal separator of A_1 and A_2.*

In this case (A_1, S, A_2) *decomposes* $\mathcal{G}$ into the *components* $\mathcal{G}(A_1 \cup S)$ and $\mathcal{G}(S \cup A_2)$. The decomposition is *proper* if A_1 and A_2 are not empty.

Definition A.2. *The graph $\mathcal{G}$ is* decomposable *if it is complete or if there exists a proper decomposition (A_1, S, A_2) into decomposable graphs $\mathcal{G}(A_1 \cup S)$ and $\mathcal{G}(S \cup A_2)$.*

Assume that $\mathcal{G}$ is decomposable and let $\mathcal{C}(\mathcal{G}) := \{D_1, D_2, \ldots, D_r\}$ be the set of cliques of $\mathcal{G}$. Since $\mathcal{G}$ is decomposable, it is possible to order the vertex sets in $\mathcal{C}(\mathcal{G})$ in a perfect sequence (Blair and Barry 1993). If we denote $H_j := D_1 \cup D_2 \cup \ldots \cup D_j$ and $S_j := H_{j-1} \cap D_j$, it follows that, for every $j = 2, \ldots, r$, $(H_{j-1} \backslash S_j, S_j, D_j \backslash S_j)$ is a decomposition of $\mathcal{G}(H_j)$ (Lauritzen 1996). We let $\mathcal{S}(\mathcal{G}) := \{S_2, \ldots, S_r\}$ be the set of separators of the graph $\mathcal{G}$ associated with $\mathcal{C}(\mathcal{G})$.

Let $\mathcal{T} = (\mathcal{C}(\mathcal{G}), \mathcal{E}_{\mathcal{T}})$ be a tree defined on the set of cliques of the decomposable graph $\mathcal{G}$.

Definition A.3 *The Star Property. Take $D_j \in \mathcal{C}(\mathcal{G})$ and let $S = D_j \cap D_i$ for some $(D_j, D_i) \in \mathcal{E}_{\mathcal{T}}$. Let $\mathcal{T}_j = (\mathcal{K}_j, \mathcal{E}_j)$ and $\mathcal{T}_i = (\mathcal{K}_i, \mathcal{E}_i)$ be the two subtrees obtained by removing the edge (D_j, D_i) from $\mathcal{T}$, with $D_j \in \mathcal{K}_j$ and $D_i \in \mathcal{K}_i$. Consider the vertex sets*

$$V_j := \bigcup_{D \in \mathcal{K}_j} D \quad \text{and} \quad V_i := \bigcup_{D \in \mathcal{K}_i} D. \tag{A.1}$$

The tree $\mathcal{T}$ is said to have the star property *for $\mathcal{G}$ if, for every edge $(D_j, D_i) \in \mathcal{E}_\mathcal{T}$, $(V_j \backslash S, S, V_i \backslash S)$ is a decomposition of $\mathcal{G}$.*

Blair and Barry (1993) show that it is always possible to construct a tree $\mathcal{T}$ that has the star property. In addition, they show that such $S \subset V$ is a minimal separator of $\mathcal{G}$ if and only if $S = D_j \cap D_i$ for some edge $(D_j, D_i) \in \mathcal{E}_\mathcal{T}$. The set of separators $\mathcal{S}(\mathcal{G})$ associated with $\mathcal{C}(\mathcal{G})$ will be given by $\mathcal{S}(\mathcal{G}) = \{D_j \cap D_i : (D_j, D_i) \in \mathcal{E}_\mathcal{T}\}$.

By removing a terminal clique from such a tree, the star property is preserved, as shown in the next result.

Lemma 1. *Let $\mathcal{T} = (\mathcal{C}(\mathcal{G}), \mathcal{E}_\mathcal{T})$ be a tree defined on the set of cliques of a decomposable graph $\mathcal{G}$. Assume that $\mathcal{T}$ has the star property for $\mathcal{G}$. Let D be a terminal clique in $\mathcal{T}$ and let D' be the the unique clique in $\mathcal{C}(\mathcal{G})$ such that $(D, D') \in \mathcal{E}_\mathcal{T}$. We consider $\mathcal{T}' = (\mathcal{C}(\mathcal{G}) \backslash \{D\}, \mathcal{E}_\mathcal{T} \backslash \{(D, D')\})$ to be the tree obtained by removing D from $\mathcal{T}$. Then $\mathcal{T}'$ is a tree with the star property for the decomposable graph $\mathcal{G}'$ defined by the set of cliques $\mathcal{C}(\mathcal{G}) \backslash \{D\}$.*

Proof. Consider an arbitrary edge $(D_j, D_i) \in \mathcal{E}_\mathcal{T} \backslash \{(D, D')\}$. As before, we let $\mathcal{T}_j = (\mathcal{K}_j, \mathcal{E}_j)$ and $\mathcal{T}_i = (\mathcal{K}_i, \mathcal{E}_i)$ be the two subtrees obtained by removing the edge (D_j, D_i) from $\mathcal{T}$, with $D_j \in \mathcal{T}_j$ and $D_i \in \mathcal{T}_i$. Let V_j and V_i the vertex sets defined in (A.1).

We can assume that $D \in \mathcal{K}_j$. If we were to remove the edge (D_j, D_i) from $\mathcal{T}'$, we would obtain the subtrees $\mathcal{T}_j' = (\mathcal{K}_j \backslash \{D\}, \mathcal{E}_j \backslash \{(D, D')\})$ and $\mathcal{T}_i = (\mathcal{K}_i, \mathcal{E}_i)$. The vertex set associated with $\mathcal{T}_j'$ is

$$V_j' := \bigcup_{D'' \in \mathcal{K}_j \backslash \{D\}} D''.$$

Since D is terminal in $\mathcal{E}_\mathcal{T}$, we have $D_j \neq D$, hence $V_j' \neq \varnothing$. The vertex set $S := D_j \cap D_i$ which is a separator for $\mathcal{G}$, is also a separator for $\mathcal{G}'$. Moreover, $(V_j \backslash S, S, V_i \backslash S)$ is a decomposition of $\mathcal{G}$. From $V_j' \subset V_j$, it follows that $(V_j' \backslash S, S, V_i \backslash S)$ will be a decomposition of $\mathcal{G}'$. Therefore the tree $\mathcal{T}'$ has the star property for $\mathcal{G}'$. □

Acknowledgements

Preparation of this paper was supported in part by the National Science Foundation under grant EIA-9876619 to the National Institute of Statistical Sciences. The author is extremely grateful to Dr Stephen E. Fienberg and Dr Stephen F. Roehrig for all their support, comments and valuable discussions. The author would like to thank Seth Sullivant who provided comments on an earlier version of the paper.

References

Agresti, A. (1992) A survey of exact inference for contingency tables (with discussion). *Statist. Sci.*, **7**, 131–177.

Bishop, Y.M.M., Fienberg, S.E. and Holland, P.W. (1975) *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, MA: MIT Press.

Blair, J.R.S. and Barry, P. (1993) An introduction to chordal graphs and clique trees. In A. George, J.R. Gilbert and J.W.H. Liu (eds), *Graph Theory and Sparse Matrix Computation*, IMA Vol. Math. Appl. 56, pp. 1–30. Berlin: Springer-Verlag.

Cox, D., Little, J. and O'Shea, D. (1992) *Ideals, Varieties and Algorithms*. New York: Springer-Verlag.

Diaconis, P. and Efron, B. (1985) Testing for independence in a two-way table: New interpretations of the chi-square statistic. *Ann. Statist.*, **13**, 845–874.

Diaconis, P. and Gangolli, A. (1995) Rectangular arrays with fixed margins. In D. Aldouns, P. Diaconis, J. Spencer and J.M. Steele (eds), *Discrete Probability and Algorithms*, IMA Vol. Math. Appl. 72, pp. 15–41. New York: Springer-Verlag.

Diaconis, P. and Sturmfels, B. (1998) Algebraic algorithms for sampling from conditional distributions. *Ann. Statist.*, **26**, 363–397.

Dinwoodie, I.H. (1998) The Diaconis–Sturmfels algorithm and rules of succession. *Bernoulli*, **4**, 401–410.

Dobra, A. (2002) Statistical tools for disclosure limitation in multi-way contingency tables. Ph.D. thesis, Department of Statistics, Carnegie Mellon University.

Dobra, A. and Fienberg, S.E. (2000) Bounds for cell entries in contingency tables given marginal totals and decomposable graphs. *Proc. Nat. Acad. Sci. USA*, **97**, 11 885–11 892.

Dobra, A. and Sullivant, S. (2003) A divide-and-conquer algorithm for generating Markov bases of multi-way tables. *Comput. Statist.*. To appear.

Fienberg, S.E., Makov, U.E. and Steele, R.J. (1998) Disclosure limitation using perturbation and related methods for categorical data. *J. Official Statist.*, **14**, 485–511.

Fienberg, S.E., Makov, U.E., Meyer, M.M. and Steele, R.J. (2001) Computing the exact distribution for a multi-way contingency table conditional on its marginals totals. In A. Saleh (ed.), *Data Analysis from Statistical Foundations: A Festschrift in Honour of the 75th Birthday of D.A.S. Fraser*, pp. 145–165. Huntington, NY: Nova Science.

Lauritzen, S.L. (1996) *Graphical Models*. Oxford: Clarendon Press.

Madigan, D. and York, J. (1995) Bayesian graphical models for discrete data. *Internat. Statist. Rev.*, **63**, 215–232.

Mehta, C. (1994) The exact analysis of contingency tables in medical research. *Statist. Methods Medical Res.*, **3**, 135–156.

Sturmfels, B. (1995) *Gröbner Bases and Convex Polytopes*, Univ. Lecture Ser. 8. Providence, RI: American Mathematical Society.

Vlach, M. (1986) Conditions for the existence of solutions of the three-dimensional planar transportation problem. *Discrete Appl. Math.*, **13**, 61–78.

Whittaker, J. (1990) *Graphical Models in Applied Multivariate Statistics*. New York: Wiley.

Willenborg, L. and de Waal, T. (2001) *Elements of Statistical Disclosure Control*, Lecture Notes in Statist. 155. New York: Springer-Verlag.

Received September 2001 and revised June 2003