ON MARGINALIZATION, COLLAPSIBILITY AND PRECOLLAPSIBILITY

M. STUDENÝ
Institute of Information Theory and Automation
Academy of Sciences of Czech Republic
Pod vodárenskou věží 4, 18208 Prague, Czech Republic

Abstract. It is shown that for every undirected graph G over a finite set N and for every nonempty $T \subset N$ there exists an undirected graph G^T over T, called the *marginal graph* of G *for* T, such that the class of marginal distributions for T of (discrete) G-Markovian distributions coincides with the class of G^T-Markovian distributions. An example shows that this is not true within the framework of strictly positive probability distributions. However, an analogous positive result holds for hypergraphs and classes of strictly positive factorizable distributions.

1. Introduction

Frydenberg [2] characterized in graphical terms the situation when an undirected graph G over a set of nodes N is *collapsible onto* a set $T \subset N$, that is when the class of marginals (for T) of Markovian distributions (with respect to G) coincides with the class of Markovian distributions with respect to the induced subgraph G_T. An analogous problem was treated in [1] where collapsibility of loglinear models onto a set of variables was characterized.

In this paper a more general point of view on these results is presented. It is shown that both undirected graphs within the framework of all discrete probability distributions without fixed domain, and hypergraphs (that is loglinear models) within the framework of strictly positive discrete probability distributions (without fixed domain) possess *precollapsibility* property. This means that for any undirected graph G (resp. hypergraph $\mathcal{C}$) over N and any nonempty subset $T \subset N$ there exists a so-called *marginal* graph G^T (resp. marginal hypergraph $\mathcal{C}^T$) such that the class of marginals for T of all discrete Markovian distributions with respect to G (resp. of strictly positive discrete distributions factorizable with respect to $\mathcal{C}$) is exactly the class of discrete Markovian distributions with respect to the marginal graph (resp. the class of strictly positive distributions factorizable with respect to the marginal hypergraph). Thus, the operation of *marginalization* on classes of 'structured' distributions can be performed by simple change of the object, describing the structure. Evidently, the collapsibility occurs iff the marginal graph G^T coincides with the induced graph G_T (similarly for hypergraphs).

On the other hand, an example shows that precollapsibility property does not hold for undirected graphs within the framework of strictly positive probability distributions. In the final discussion further possible mathematical objects (describing structure of distributions) for which the concept of precollapsibilty can be considered, are mentioned: directed acyclic graphs and structural imsets.

2. Basic concepts

2.1. DISCRETE DISTRIBUTIONS

Throughout the paper N will denote a nonempty finite set of factors. For its subsets $A, B \subset N$ the juxtaposition AB will be used to shorten $A \cup B$. A *potential* over a set $B \subset N$ is specified by two entities: by a collection of nonempty finite sets $\{\mathbf{X}_i \,;\, i \in B\}$ and by a nonnegative real function R on the cartesian product $\prod_{i\in B} \mathbf{X}_i$. If we wish to make the domain of R explicit, then we say that R is (a potential) *on* $\prod_{i\in B} \mathbf{X}_i$. If $R(\mathbf{x}) > 0$ for all $\mathbf{x} \in \prod_{i\in B} \mathbf{X}_i$ then R is called *strictly positive.*

The *marginal* of a potential R over B *for* a set $\emptyset \neq A \subset B$ is a potential R^A over A defined as follows: if R is on $\prod_{i\in B} \mathbf{X}_i$, then R^A is defined on $\prod_{i\in A} \mathbf{X}_i$, given by the formula $R^A(\mathbf{a}) = \sum\{R(\mathbf{a}, \mathbf{d});\ \mathbf{d} \in \prod_{i\in B\setminus A} \mathbf{X}_i\}$, where $\mathbf{a} \in \prod_{i\in A} \mathbf{X}_i$.

A discrete *probability distribution* over N is a potential P over N such that $\sum\{P(\mathbf{x});\ \mathbf{x} \in \prod_{i\in N} \mathbf{X}_i\} = 1$ (supposing P is on $\prod_{i\in N} \mathbf{X}_i$). The class of discrete probability distributions over N will be denoted by $\mathcal{P}(N)$, the class of strictly positive discrete probability distributions over N by $\mathcal{P}_+(N)$. Note that they involve distributions on all possible domains $\prod_{i\in N} \mathbf{X}_i$.

Let us extend the notation of marginalization to classes of distributions. Having $\mathcal{L} \subset \mathcal{P}(N)$ and $\emptyset \neq T \subset N$ we denote $\mathcal{L}^T = \{P^T \,;\, P \in \mathcal{L}\}$. Clearly, the operation of *marginalization* treated for classes of distributions respects inclusion, that is $\mathcal{K} \subset \mathcal{L} \subset \mathcal{P}(N)$, $\emptyset \neq T \subset N$ implies $\mathcal{K}^T \subset \mathcal{L}^T$, and is idempotent, that is $\mathcal{L}^T = (\mathcal{L}^S)^T$ whenever $\mathcal{L} \subset \mathcal{P}(N)$, $\emptyset \neq T \subset S \subset N$.

Let $\mathcal{T}(N)$ denote the collection of triplets $\langle A, B|C\rangle$ of disjoint subsets of N whose first two components A and B are nonempty. Having $P \in \mathcal{P}(N)$, defined on $\prod_{i\in N} \mathbf{X}_i$, and $\langle A, B|C\rangle \in \mathcal{T}(N)$ we say that A is *conditionally* *independent of* B *given* C *with respect to* P and write $A \perp\!\!\!\perp B \,|\, C\,(P)$ if $\forall\ \mathbf{a} \in \prod_{i\in A} \mathbf{X}_i \quad \mathbf{b} \in \prod_{i\in B} \mathbf{X}_i \quad \mathbf{c} \in \prod_{i\in C} \mathbf{X}_i$

$$P^{ABC}(\mathbf{a},\mathbf{b},\mathbf{c}) \cdot P^{C}(\mathbf{c}) = P^{AC}(\mathbf{a},\mathbf{c}) \cdot P^{BC}(\mathbf{b},\mathbf{c}).$$

Evidently the statement depends on the marginal P^{ABC} only, and therefore for every $ABC \subset T \subset N$ one has $A \perp\!\!\!\perp B \,|\, C\,(P^T)$ iff $A \perp\!\!\!\perp B \,|\, C\,(P)$.

2.2. UNDIRECTED GRAPHS

An *undirected graph* over a set N is specified by a set of two-element subsets of N called *lines*. The symbol $u - v$ denotes that $\{u, v\}$ is a line. The class of undirected graphs over N will be denoted by $\mathcal{U}(N)$.

A path in $G \in \mathcal{U}(N)$ is a sequence $u_1, \ldots, u_n$, $n \geq 1$ of elements of N such that $u_i - u_{i+1}$ for $i = 1, \ldots, n-1$. Having $G \in \mathcal{U}(N)$ and $\langle A, B|C\rangle \in \mathcal{T}(N)$ we say that C *separates* A from B in G and write $A \perp\!\!\!\perp B \,|\, C\,(G)$ if every path in G from a node in A to a node in B contains a node in C.

We say that $P \in \mathcal{P}(N)$ is *Markovian* with respect to $G \in \mathcal{U}(N)$ (or shortly G-Markovian) if $A \perp\!\!\!\perp B \,|\, C\,(G)$ implies $A \perp\!\!\!\perp B \,|\, C\,(P)$ for every $\langle A, B|C\rangle \in \mathcal{T}(N)$. The class of (discrete) Markovian distributions with respect to G will be denoted by $\mathcal{M}(G)$, the symbol $\mathcal{M}_+(G)$ denotes the class of strictly positive G-Markovian distributions, that is $\mathcal{P}_+(N) \cap \mathcal{M}(G)$.

The *induced subgraph* of $G \in \mathcal{U}(N)$ *for* a nonempty set $T \subset N$ is the graph G_T over T such that $u - v$ in G_T iff $u - v$ in G for every $u, v \in T$.

2.3. HYPERGRAPHS

By a *hypergraph* over N is understood a system of incomparable subsets $\mathcal{C}$ of N, that is for every couple of different sets $A, B \in \mathcal{C}$ neither $A \subset B$ nor $B \subset A$, such that the union of sets in $\mathcal{C}$ is N. The class of hypergraphs over N will be denoted by $\mathcal{H}(N)$.

Having $\mathcal{C} \in \mathcal{H}(N)$ we say that a potential R on $\prod_{i\in N} \mathbf{X}_i$ is *factorizable* with respect to $\mathcal{C}$ (or shortly $\mathcal{C}$-factorizable) if there exists a collection of real functions f_L on $\prod_{i\in L} \mathbf{X}_i$, $L \in \mathcal{C}$ such that $R([\mathbf{x}_i]_{i\in N}) = \prod_{L\in\mathcal{C}} f_L([\mathbf{x}_i]_{i\in L})$. Note that equivalently functions f_L can be demanded to be strictly positive real functions on $\prod_{i\in L} \mathbf{X}_i$. The symbol $\mathcal{F}_+(\mathcal{C})$ denotes the class of all strictly positive discrete probability distributions over N which are factorizable with respect to $\mathcal{C} \in \mathcal{H}(N)$.

The *induced* hypergraph of $\mathcal{C} \in \mathcal{H}(N)$ *for* $\emptyset \neq T \subset N$ consists of maximal sets (with respect to inclusion) of the system $\{\, L \cap T\,;\, L \in \mathcal{C}\,\}$.

3. Precollapsibility for undirected graphs

In this section the precollapsibility property for undirected graphs within the framework of $\mathcal{P}(N)$ is verified.

Having $G \in \mathcal{U}(N)$ and $\emptyset \neq T \subset N$ we introduce the *marginal graph* of *G for T*, denoted by G^T, as a graph over T such that
$u - v$ in $G^T \;\Leftrightarrow\; \neg[\,\{u\} \perp\!\!\!\perp \{v\}\,|\,T \setminus \{u,v\}\,(G)\,]$.
Thus, u and v form a line in G^T iff they are connected by a path in G outside $T \setminus \{u,v\}$. Especially, G_T is a subgraph of G^T. Moreover, the operation of marginalization for undirected graphs is also idempotent, that is $G^T = (G^S)^T$ whenever $\emptyset \neq T \subset S \subset N$.

Lemma 3.1 *$A \perp\!\!\!\perp B\,|\,C\,(G)$ iff $A \perp\!\!\!\perp B\,|\,C\,(G^T)$ for any $\langle A,B|C\rangle \in \mathcal{T}(T)$.*

Proof. Having a path in G from A to B outside C every its section between nodes $u, v \in T$ whose all internal nodes are outside T can be replaced by an edge $u - v$ in G^T. Conversely, having a path in G^T from A to B outside C every its line $w_i - w_{i+1}$ can be replaced by a roundabout way from w_i to w_{i+1} in G which is outside $T \setminus \{w_i, w_{i+1}\}$ and therefore outside C. $\square$

Lemma 3.2 *Supposing $G \in \mathcal{U}(N)$, $\emptyset \neq T \equiv N \setminus \{w\}$ and $Q \in \mathcal{M}(G^T)$ let us denote $Z = \{v \in T\,;\, v - w \text{ in } G\}$. Then there exists $P \in \mathcal{P}(N)$ such that*

- *$P^T = Q$,*
- *$\{w\} \perp\!\!\!\perp (T \setminus Z)\,|\,Z\,(P)$,*
- *whenever $\langle \tilde{A}, \tilde{B}\,|\,\tilde{C}\rangle \in \mathcal{T}(T)$ with $(\tilde{A} \setminus Z) \perp\!\!\!\perp (\tilde{B} \setminus Z)\,|\,\tilde{C}Z\,(G^T)$, then $\tilde{A} \perp\!\!\!\perp \tilde{B}\,|\,\tilde{C} \cup \{w\}\,(P)$.*

Proof. Supposing the distribution Q is defined on a cartesian product $\prod_{i\in T} \mathbf{X}_i$ let us put $\mathbf{X}_w = \prod_{i \in Z} \mathbf{X}_i$ and define
$P([\mathbf{x}_i]_{i\in N}) = Q([\mathbf{x}_i]_{i\in T})$ in case $\mathbf{x}_w = [\mathbf{x}_i]_{i\in Z}$, $\;P([\mathbf{x}_i]_{i\in N}) = 0$ otherwise.
We leave it to the reader to verify that P satisfies the desired conditions. $\square$

Theorem 3.1 *If $G \in \mathcal{U}(N)$, $\emptyset \neq T \subset N$, then $\mathcal{M}(G^T) = \mathcal{M}(G)^T$.*

Remark In our set-up the inclusion $\mathcal{M}(G^T) \subset \mathcal{M}(G)^T$ means that there exists a G-Markovian distribution having prescribed G^T-Markovian marginal for T but it is not claimed that it has also prescribed domain $\prod_{i\in N} \mathbf{X}_i$! In fact, the sets $\mathbf{X}_i$ for $i \in N \setminus T$ are a part of the constuction, they depend both on G and $\prod_{i\in T} \mathbf{X}_i$. An analogous remark holds for Theorem 4.1.

Proof. To see $\mathcal{M}(G)^T \subset \mathcal{M}(G^T)$ let us consider $P \in \mathcal{M}(G)$. Then for every $\langle A,B|C\rangle \in \mathcal{T}(T)$ the statement $A \perp\!\!\!\perp B\,|\,C\,(G^T)$ implies $A \perp\!\!\!\perp B\,|\,C\,(G)$ by Lemma 3.1 and hence $A \perp\!\!\!\perp B\,|\,C\,(P)$, that is $A \perp\!\!\!\perp B\,|\,C\,(P^T)$. Thus, the fact $P^T \in \mathcal{M}(G^T)$ was verified.

To prove $\mathcal{M}(G^T) \subset \mathcal{M}(G)^T$ let us consider $Q \in \mathcal{M}(G^T)$ and try to find $P \in \mathcal{M}(G)$ with $P^T = Q$. However, owing to the fact that the operation of marginalization is idempotent (both for graphs and classes of distributions) and respects inclusion, it suffices to prove this statement only for the case $T = N \setminus \{w\}$ where $w \in N$. Let us use the construction from Lemma 3.2. One needs to verify that $P \in \mathcal{M}(G)$.

Thus, suppose $\langle A, B|C\rangle \in \mathcal{T}(N)$ such that $A \perp\!\!\!\perp B \,|\, C\,(G)$. In case $ABC \subset T$ one derives by Lemma 3.1 and from the fact $P^T = Q \in \mathcal{M}(G^T)$ that $A \perp\!\!\!\perp B \,|\, C\,(Q)$, that is $A \perp\!\!\!\perp B \,|\, C\,(P)$.
If $w \in C$, then $A \perp\!\!\!\perp B \,|\, C\,(G)$ implies that $(A\setminus Z) \perp\!\!\!\perp (B\setminus Z) \,|\, CZ\setminus\{w\}\,(G^T)$. Indeed, otherwise the considered path in G^T from $A \setminus Z$ to $B \setminus Z$ outside $CZ\setminus\{w\}$ belongs to $T\setminus Z$ and will remain in G. Thus, by the third property of P from Lemma 3.2 ($\tilde{C} = C \setminus \{w\}$) one has $A \perp\!\!\!\perp B \,|\, C\,(P)$.
If $w \in A$ (the case $w \in B$ is dual), then $B \cap Z = \emptyset$ (otherwise a line $w - v$ for $v \in B \cap Z$ is a path in G from A to B outside C) and $AD\setminus\{w\} \perp\!\!\!\perp B \,|\, C\,(G^T)$ for $D = Z \setminus AC$. Indeed, otherwise consider a path in G^T from $u \in B$ to $AD \setminus \{w\}$ outside C and take its possible first node v in Z. Its part from u to v can be lenghthened by the line $v - w$ into a path in G from B to $w \in A$ outside C. If the considered path from $u \in B$ to $AD \setminus \{w\}$ is outside Z, then it leads to A what also contradicts the assumption $A \perp\!\!\!\perp B \,|\, C\,(G)$. The fact $Q \in \mathcal{M}(G^T)$ implies $B \perp\!\!\!\perp AD \setminus \{w\} \,|\, C\,(Q)$ and therefore $B \perp\!\!\!\perp AD \setminus \{w\} \,|\, C\,(P)$. Moreover, we know from the second condition of Lemma 3.2 that $\{w\} \perp\!\!\!\perp (T \setminus Z) \,|\, Z\,(P)$. Now, well-known semigraphoid properties of conditional independence [5] can be used to derive $\{w\} \perp\!\!\!\perp B \,|\, ACD \setminus \{w\}\,(P)$, that is $B \perp\!\!\!\perp \{w\} \,|\, ACD\setminus\{w\}\,(P)$ what with $B \perp\!\!\!\perp AD\setminus\{w\} \,|\, C\,(P)$ gives $B \perp\!\!\!\perp AD \,|\, C\,(P)$ and hence $A \perp\!\!\!\perp B \,|\, C\,(P)$. $\square$

4. Precollapsiblity for hypergraphs

In this section the precollapsibility property for hypergraphs within the framework $\mathcal{P}_+(N)$ is treated. Note for explanation that the loglinear model with a generating class (hypergraph) $\mathcal{C}$ can be equivalently introduced as the class of $\mathcal{C}$-factorizable strictly positive distributions.

Having $\mathcal{C} \in \mathcal{H}(N)$ and nonempty $T \subset N$, the marginal hypergraph is constructed as follows. We say that $S, R \in \mathcal{C}$ are *connected outside* T if there exists a sequence $S = L_1, \ldots, L_n = R$, $n \geq 1$ in $\mathcal{C}$ such that $L_i \cap L_{i+1} \setminus T \neq \emptyset$ for $i = 1, \ldots, n-1$. The relation 'be connected outside T' decomposes $\mathcal{C}$ into equivalence classes, let us consider the collection $\mathcal{D}$ of sets which are unions of the equivalence classes. Of course, $\mathcal{D}$ *covers* $\mathcal{C}$ in sense that $\forall L \in \mathcal{C} \;\; \exists K \in \mathcal{D} \;\; L \subset K$.

Then the *marginal hypergraph* of $\mathcal{C}$ *for* T, denoted by $\mathcal{C}^T$, consists of maximal sets (with respect to inclusion) of the class $\{D \cap T\,;\, D \in \mathcal{D}\}$. The reader can verify that the operation of marginalization on hypergraphs is also idempotent.

Lemma 4.1 *Suppose* $\emptyset \neq T \equiv N \setminus \{w\}$, $\mathcal{C} = \{\,\{w,t\}\,; t \in T\,\}$ *and* R *is a strictly positive potential over* T. *Then there exists a strictly positive* $\mathcal{C}$*-factorizable potential* P *over* N *such that* $P^T = R$.

Proof. Without loss of generality one can supose that R is a probability distribution defined on $\prod_{i\in T} \mathbf{X}_i$. Then we put $\mathbf{X}_w = \prod_{i\in T} \mathbf{X}_i$. The constructed probability distribution has the form
$P([\mathbf{x}_i]_{i\in T}, \mathbf{z}) = \alpha(\mathbf{z}) \cdot \prod_{i\in T} S_{i,\mathbf{z}}(\mathbf{x}_i)$ for $\mathbf{z} = [\mathbf{z}_j]_{j\in T} \in \mathbf{X}_w$, $[\mathbf{x}_i]_{i\in T} \in \prod_{i\in T} \mathbf{X}_i$, where $\alpha(\mathbf{z}) \geq 0$ are searched parameters summing to 1 and $S_{i,\mathbf{z}}$ are probability distributions (determined by parameter $\varepsilon > 0$) on $\mathbf{X}_i$ defined $S_{i,\mathbf{z}}(\mathbf{x}_i) = \delta_i = 1 - (\mathrm{card}\,\mathbf{X}_i - 1) \cdot \varepsilon$ in case $\mathbf{x}_i = \mathbf{z}_i$ and $S_{i,\mathbf{z}}(\mathbf{x}_i) = \varepsilon$ otherwise.

One can show that for sufficiently small choice of ε one can find parameters $\alpha(\mathbf{z})$ such that $P^T = R$. To comply with the requirement that P should be strictly positive the set $\mathbf{X}_w$ can be reduced to the set $\{\,\mathbf{z}\,; \alpha(\mathbf{z}) > 0\,\}$. Evidently, P is $\mathcal{C}$-factorizable. To prove the existence of parameters $\alpha(\mathbf{z})$ one can use the Banach fixed-point theorem. The idea is to consider a certain mapping T on the metric space $\prod_{\mathbf{z}\in\mathbf{X}_w} [0, \prod_{i\in T} \delta_i^{-1} \cdot R(\mathbf{z})]$ which is contractive for sufficiently small ε. The solution α of the equation $T(\alpha) = \alpha$ then forms the desired collection of parameters. For limited scope of a conference contribution I omit technically complicated details. $\square$

Theorem 4.1 *If* $\mathcal{C} \in \mathcal{H}(N)$, $\emptyset \neq T \subset N$, *then* $\mathcal{F}_+(\mathcal{C}^T) = \mathcal{F}_+(\mathcal{C})^T$.

Proof. The inclusion $\mathcal{F}_+(\mathcal{C})^T \subset \mathcal{F}_+(\mathcal{C}^T)$ is easy as the system $\mathcal{D}$ (see the definition of marginal hypergraph) covers $\mathcal{C}$ and therefore every $P \in \mathcal{F}_+(\mathcal{C})$ factorizes with respect to $\mathcal{D}$. Because of $D \cap E \setminus T = \emptyset$ for different $D, E \in \mathcal{D}$ the marginal P^T factorizes with respect to $\mathcal{C}^T$ and therefore $P^T \in \mathcal{F}_+(\mathcal{C}^T)$.

To prove $\mathcal{F}_+(\mathcal{C}^T) \subset \mathcal{F}_+(\mathcal{C})^T$ let us take $Q \in \mathcal{F}_+(\mathcal{C}^T)$ and try to find $P \in \mathcal{F}_+(\mathcal{C})$ with $P^T = Q$. By the same reasons as explained in the proof of Theorem 3.1 one can suppose without loss of generality that $T = N \setminus \{w\}$ where $w \in N$. In this special situation, the union of the class $\mathcal{C}_w$ of sets in $\mathcal{C}$ containing w forms one set in $\mathcal{D}$, the remaining elements of $\mathcal{C} \setminus \mathcal{C}_w$ also belong to $\mathcal{D}$. Let us denote by Z the union of sets in $\mathcal{C}_w$ with removed w. In case $Z \in \mathcal{C}^T$ (otherwise trivial) one considers a factorization of Q with respect to $\mathcal{C}^T$ and the only problem is to show that the potential corresponding to Z (in this factorization) is marginal of a potential over $Z \cup \{w\}$ which is $\mathcal{C}_w$-factorizable. However, it follows from Lemma 4.1 (for $T = Z$) since the hypergraph $\{\,\{w, z\}\,;\, z \in Z\,\}$ is covered by $\mathcal{C}_w$. $\square$

5. Counterexample

On the other hand, the precollapsibility property for undirected graphs does not hold within the framework of $\mathcal{P}_+(N)$ as the following example shows.

Example 5.1 Take $N = \{1,2,3,4\}$, the graph $G \in \mathcal{U}(N)$ consisting of the cycle $1 - 2 - 3 - 4 - 1$ and $T = \{1,2,3\}$. No other graph over T except the complete graph K over T can be considered as a candidate for a 'marginal graph' within the framework $\mathcal{P}_+(N)$. The reason is that one can find $P \in \mathcal{M}_+(G)$ such that $\neg\,[\, A \perp\!\!\!\perp B \,|\, C\,(P)\,]$ for every $\langle A, B|C\rangle \in \mathcal{T}(T)$ and therefore $P^T \notin \mathcal{M}_+(H)$ for any other $H \in \mathcal{U}(T)$ except K. Consider the distribution Q over T defined on $\prod_{i\in T}\{0,1\}$ where $Q(0,0,0) = 1/12$, $Q(0,1,0) = 1/6$ and $Q(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3) = 1/8$ for remaining points $[\mathbf{x}_i]_{i\in T}$ of $\prod_{i\in T}\{0,1\}$. Since $Q(0,x_2,0)\cdot Q(1,x_2,1)/Q(0,x_2,1)\cdot Q(1,x_2,0)$ does depend on $\mathbf{x}_2$, the distribution Q cannot be factorized with respect to the hypergraph $\mathcal{C} = \{\,\{1,2\},\{1,3\},\{2,3\}\,\}$. However, every $P \in \mathcal{M}_+(G)$ is factorizable with respect to the system of its cliques $\mathcal{S} = \{\,\{1,2\},\{2,3\},\{3,4\},\{1,4\}\,\}$ by well-known Hammersley-Clifford theorem (see [6] or [3]) and therefore its marginal P^T is factorizable with respect to its marginal hypergraph $\mathcal{S}^T = \mathcal{C}$ by Theorem 4.1. So, $Q \notin \mathcal{M}_+(G)^T$ although $Q \in \mathcal{P}_+(T) = \mathcal{M}_+(K)$.

6. Discussion

Note that the concept of marginal hypergraph is equivalent to the concept of derivative of a generating class of a loglinear model treated in [4]. To clarify the connection with that result, let us recall that Malvestuto [4] showed that the inclusion $\mathcal{F}_+(\mathcal{C})^T \subset \mathcal{F}_+(\mathcal{B})$ holds just for those hypergraphs $\mathcal{B}$ over T which cover $\mathcal{C}^T$ and no other hypergraphs over T. Here, a more special fact $\mathcal{F}_+(\mathcal{C}^T) \setminus \mathcal{F}_+(\mathcal{C})^T = \emptyset$ is showed.

The precollapsibility condition can be also considered for other mathematical objects describing structures of discrete probability distributions. For example, *directed acyclic graphs* define the class of recursively factorizable distributions [3] or *structural imsets* [7] allow to describe all structures of probabilistic conditional independence for discrete distributions. I conjecture that the precollapsibility property does not hold for directed acyclic graphs since there is no 'reasonable candidate' for the marginal directed acyclic graph. On the other hand, I have some reasons to hope that the precollapsibility property holds for structural imsets. Well, I believe that the precollapsibility property is a very natural demand which should be one of the important criteria for choice of methods of structural description.

Acknowledgements

I would like to thank an anonymous reviewer for a helpful comment concerning Theorems 3.1 and 4.1. This work was supported by the grant GAČR n. 201/94/0471, the grant CEC n. CIPA3511 CT930053 and the grant MŠMT n. VS96008.

References

1. Asmussen, S. and Edwards, D. (1983) Collapsibility and response variables in contingency tables, *Biometrika*, **70**, pp. 567–578.
2. Frydenberg, M. (1990) Marginalization and collapsibility in graphical interaction models, *The Annals Statistics*, **18**, pp. 790–805.
3. Lauritzen, S.L., Dawid, A.P., Larsen B.N. and Leimer, H.-G. (1990) Independence properties of directed Markov fields, *Networks*, **20**, pp. 491–505.
4. Malvestuto, F.M. (1996) Testing implication of hierarchical log-linear models for probability distributions, *Statistics and Computing*, **6**, pp. 169–176.
5. Pearl, J. (1988) *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann, San Mateo.
6. Speed, T.P. (1979) A note on nearest-neighbour Gibbs and Markov probabilities, *Sakhya A*, **41**, pp. 184–197.
7. Studený, M. (1994/1995) Description of structures of stochastic conditional independence by means of faces and imsets (a series of 3 papers), *International Journal of General Systems*, **23**, pp. 123–137, 201–219, 323–341.