Persegrams of Compositional Models Revisited: conditional independence

Radim Jiroušek
Faculty of management, University of Economics
and
Institute of Information Theory and Automation, Acad. Sciences
Czech Republic
radim@utia.cas.cz

Abstract

The paper gives instructions how to read conditional independence relations for multidimensional probability distributions represented in a form of a compositional model.

Keywords: Conditional independence, Probability, Multidimensional model.

1 Introduction

Most of the graphical Markov models offer a way how to read conditional independence relations from their underlying graphs. For Bayesian networks one can do it either using Pearl's *d-separation criterion* [6, 1], or with the help of moralization criterion of Lauritzen et al. [5]. For the same purpose we introduced in [4] *persegrams*, special tables representing structures of *compositional models*, i.e. multidimensional distributions assembled from a system of low-dimensional distributions by iterative application of an operator of composition. In [4] we also proved theorem saying that two groups of variables are (unconditionally) independent if there does not exist a simple trail between the corresponding variables. In the present paper we present a new term: *L-active trail* (hopefully a more transparent modification of the formerly introduced notion of an avoiding trail), which is a generalization of a simple trail enabling us to read conditional independence relations. We show that persegrams can also be used to recognize whether two permutations of a generating sequence define the same model or not. Both the important notions, persegram and L-active trail, are abundantly illustrated with examples.

2 Probabilistic compositional models

In the whole paper we shall deal with a finite number of variables $X_1, X_2, \ldots, X_n$ each of which is specified by a finite set $\mathbf{X}_i$ of its values. A *projection* of $x = (x_1, x_2, \ldots, x_n) \in \mathbf{X}_N = \mathbf{X}_1 \times \ldots \times \mathbf{X}_n$ into $\mathbf{X}_K = \times_{i \in K} \mathbf{X}_i$ is denoted $x^{\downarrow K}$. $\pi(K)$ denotes probability distribution defined for the group of variables $X_K = \{X_i\}_{i \in K}$. $\pi(x)$ for $x \in \mathbf{X}_K$ denotes the value of this distribution for the vector $x \in \mathbf{X}_K$. For $L \subset K$, symbol $\pi^{\downarrow L}$ denotes the *marginal* distribution defined for variables X_L, i.e. for each $x \in \mathbf{X}_L$

$$\pi^{\downarrow L}(x) = \sum_{y \in \mathbf{X}_K : y^{\downarrow L} = x} \pi(y).$$

(Realize that $\pi^{\downarrow \emptyset} = 1$.) Consider three disjoint sets $I, J, K \subset N$ ($I \neq \emptyset \neq J$). We say that for distribution $\kappa(N)$ groups of variables X_I and X_J are *conditionally independent given variables* X_K (in symbol $X_I \perp\!\!\!\perp X_J | X_K[\kappa]$) if for all $x \in \mathbf{X}_{I \cup J \cup K}$ the following equality holds true

$$\kappa^{\downarrow I \cup J \cup K}(x) \cdot \kappa^{\downarrow K}(x^{\downarrow K}) = \kappa^{\downarrow I \cup K}(x^{\downarrow I \cup K}) \cdot \kappa^{\downarrow J \cup K}(x^{\downarrow J \cup K}).$$

It is well known that this is equivalent to the fact that for all $x \in \mathbf{X}_{I \cup J \cup K}$

$$\kappa^{\downarrow I \cup J \cup K}(x) = \kappa^{\downarrow I \cup K}(x^{\downarrow I \cup K}) \cdot \kappa^{\downarrow J \cup K}(x^{\downarrow J} | x^{\downarrow K}).$$

L. Magdalena, M. Ojeda-Aciego, J.L. Verdegay (eds): Proceedings of IPMU'08, pp. 915–922
Torremolinos (Málaga), June 22–27, 2008

From two low-dimensional distributions π_1 and π_2 one can get a distribution of a higher dimension with the help of the following operator of composition.

Definition 1 Consider arbitrary two distributions $\pi(K_1)$ and $\pi(K_2)$ ($K_1 \neq \emptyset \neq K_2$). If $\pi_1^{\downarrow K_1 \cap K_2}$ is dominated by $\pi_2^{\downarrow K_1 \cap K_2}$, i.e. for all $z \in \mathbf{X}_{K_1 \cap K_2}$

$$\pi_2^{\downarrow K_1 \cap K_2}(z) = 0 \Longrightarrow \pi_1^{\downarrow K_1 \cap K_2}(z) = 0,$$

then $\pi_1 \triangleright \pi_2$ is for all $x \in \mathbf{X}_{K_1 \cup K_2}$ defined by the expression

$$(\pi_1 \triangleright \pi_2)(x) = \frac{\pi_1(x^{\downarrow K_1}) \cdot \pi_2(x^{\downarrow K_2})}{\pi_2^{\downarrow K_1 \cap K_2}(x^{\downarrow K_1 \cap K_2})}.$$

($\frac{0 \cdot 0}{0} = 0$.) Otherwise the composition $\pi_1 \triangleright \pi_2$ remains undefined.

We proved it in the paper [2] that the result of composition, if defined, is a probability distribution of variables $X_{K_1 \cup K_2}$. Therefore, if the operator is applied iteratively to a sequence of distributions $\pi_1(K_1), \pi_2(K_2), \ldots \pi_n(K_n)$ (we will call it a *generating sequence* in the sequel), and if the resulting distribution

$$\pi_1 \triangleright \pi_2 \triangleright \ldots \triangleright \pi_n = (\ldots(\pi_1 \triangleright \pi_2) \triangleright \ldots \triangleright \pi_n)$$

is defined, it is a probability distribution for variables $X_{K_1 \cup K_2 \cup \ldots \cup K_n}$. Remember that the operators are always, if not specified by brackets otherwise, applied from left to right.

In the rest of the paper we will consider a generating sequence $\pi_1(K_1), \pi_2(K_2), \ldots \pi_n(K_n)$, for which $\pi_1 \triangleright \pi_2 \triangleright \ldots \triangleright \pi_n$ is defined, and will deal with the problem how to read conditional independence relations for this distribution.

3 Persegrams

Definition 2 *Persegram* of a generating sequence is a table in which rows correspond to variables (in an arbitrary order) and columns to low-dimensional distributions; ordering of the columns corresponds to the generating sequence ordering. A position in the table is marked if the respective variable is among the arguments of the corresponding distribution. Markers for the first occurrence of each variable (i.e. the leftmost markers in rows) are squares (we will call them *box-markers*) and for other occurrences they are *bullets*.

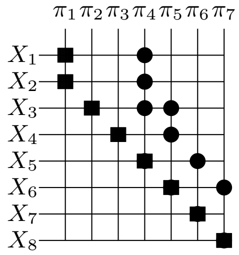
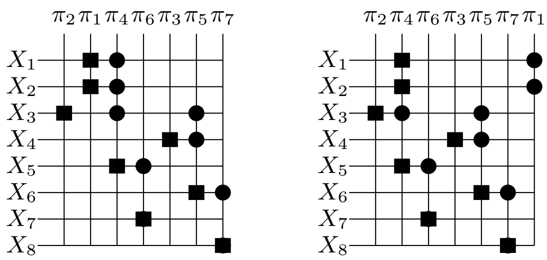


Figure 1: Persegram

Example 1 In Figure 1 we can see a persegram for the sequence

$$\pi_1(\{1,2\}), \pi_2(\{3\}), \pi_3(\{4\}), \pi_4(\{1,2,3,5\}), \pi_5(\{3,4,6\}), \pi_6(\{5,7\}), \pi_7(\{6,8\}).$$



(a) (b)

Figure 2: Persegram

Taking another permutations of this generating sequence $\pi_2, \pi_1, \pi_4, \pi_6, \pi_3, \pi_5, \pi_7$ and $\pi_2, \pi_4, \pi_6, \pi_3, \pi_5, \pi_7, \pi_1$ we get different persegrams presented in Figure 2(a) and (b), respectively. Notice the difference between these persegrams. Whilst the only difference between persegrams in Figures 1 and 2(a) is the ordering of distributions $\pi_1, \pi_2, \ldots, \pi_7$, the difference between persegram in Figure 2(b) and the other two ones is more fundamental. Examine, for example, the markers of the distribution π_4. In Figures 1 and 2(a) this distribution has only one box-marker: $X_5\pi_4$. On the other hand, in persegram in Figure 2(b) there are 3 box-markers for this distribution: $X_1\pi_4, X_2\pi_4$ and $X_5\pi_4$. Importance of this difference will be clear from the following assertion.

Theorem 1 *Consider a generating sequence $\pi_1, \pi_2, \ldots, \pi_n$ and its permutation $\pi_{i_1}, \pi_{i_2}, \ldots, \pi_{i_n}$. If the corresponding persegrams have the same box-markers (i.e. $X_i\pi_j$ is a box-marker in the persegram of sequence $\pi_1, \ldots, \pi_n$ if and only if it is a box marker also in the persegram of $\pi_{i_1}, \ldots, \pi_{i_n}$), then these two generating sequences represent the same multidimensional distribution:*

$$\pi_1 \triangleright \pi_2 \triangleright \ldots \triangleright \pi_n = \pi_{i_1} \triangleright \pi_{i_2} \triangleright \ldots \triangleright \pi_{i_n}.$$

Proof Consider the persegram of the generating sequence $\pi_1, \pi_2, \ldots, \pi_n$ and denote for each $i = 1, \ldots, n$ by B_i the set of those indices j from K_i, for which $X_j\pi_i$ is a box-marker.

Generating sequence $\pi_1, \pi_2, \ldots, \pi_n$ represents multidimensional distribution

$$\pi_1 \triangleright \ldots \triangleright \pi_n = \pi_1 \cdot \prod_{i=2}^{n} \frac{\pi_i}{\pi_i^{\downarrow K_i \cap (K_1 \cup \ldots \cup K_{i-1})}}.$$

From the definition of a persegram it is obvious that $j \in K_i \cap (K_1 \cup \ldots \cup K_{i-1})$ if and only if the corresponding marker $X_j\pi_i$ is a bullet. Since all markers corresponding to π_1 (in the persegram of $\pi_1, \pi_2, \ldots, \pi_n$) are box-markers, and $\pi_1^{\downarrow\emptyset} = 1$, we see that

$$\pi_1 \triangleright \ldots \triangleright \pi_n = \frac{\prod_{i=1}^{n} \pi_i}{\prod_{i=1}^{n} \pi_i^{\downarrow K_i \setminus B_i}}.$$

An analogous expression can be deduced also for generating sequence $\pi_{i_1}, \pi_{i_2}, \ldots, \pi_{i_n}$. Due to the assumption of this assertion, sets B_i are the same for both the considered generating sequences and therefore also the corresponding multidimensional distributions must coincide. □

Remark Let us stress that this assertion holds true only under the implicit assumption that both $\pi_1 \triangleright \ldots \triangleright \pi_n$ and $\pi_{i_1} \triangleright \ldots \triangleright \pi_{i_n}$ are defined.

Definition 3 Consider a persegram of a generating sequence $\pi_1, \ldots, \pi_n$ and $L \subset K_1 \cup \ldots \cup K_n$. A sequence of markers $m_0, m_1, \ldots, m_t$ of a persegram is called an *L-active trail ($L \subset K_1 \cup K_2 \cup \ldots \cup K_n$) that connects m_0 and m_t* if it meets the following 4 conditions:

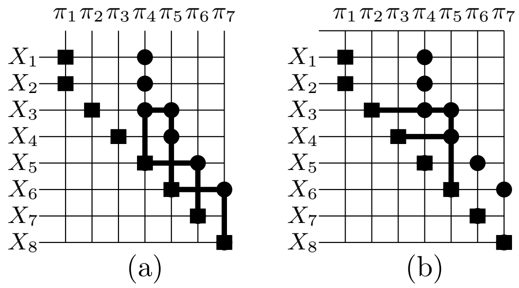


Figure 3: Active trails: (a) $X_7\pi_6, X_5\pi_6, X_5\pi_4, X_3\pi_4, X_3\pi_5, X_6\pi_5, X_6\pi_7, X_8\pi_7$;
(b) $X_3\pi_2, X_3\pi_5, X_6\pi_5, X_4\pi_5, X_4\pi_3$

1. for each $s = 1, \ldots, t$ a couple (m_{s-1}, m_s) is in the same row (i.e. horizontal connection) or in the same column (vertical connection);
2. each vertical connection must be adjacent to a box-marker (one of the markers is a box-marker);
3. no horizontal connection corresponds to a variable from X_L;
4. vertical and horizontal connections regularly alternate with the following possible exception: two vertical connections may be in a direct succession if their common adjacent marker is a box-marker of a variable from X_L.

If an L-active trail connects two box-markers corresponding to variables X_j and X_k, $j \notin L$, $k \notin L$, we also say that these *variables are connected by an L-active trail.* This situation will be denoted $X_j \rightsquigarrow_L X_k$.

Remark Notice, that in an L-active trail one marker may appear several times.

Example 2 An example of a $\{2, 4\}$-active trail is the trail in Figure 3(a); horizontal connections of this trail correspond to variables X_3, X_5 and X_6, so all the conditions of Definition 3 are fulfilled. (Notice, it is also an $\emptyset$-active trail, which is also called a *simple trail.*) However, this trail is not a $\{3, 4\}$-active trail because there is a horizontal connection $(X_3\pi_4, X_3\pi_5)$ corresponding to variable X_3.

A little bit more complex example of an active trail is in Figure 3(b): it is a $\{6\}$-active trail. It starts with a horizontal connection $(X_3\pi_2, X_3\pi_5)$, after which two vertical connections $(X_3\pi_5, X_6\pi_5)$ and $(X_6\pi_5, X_4\pi_5)$ go in

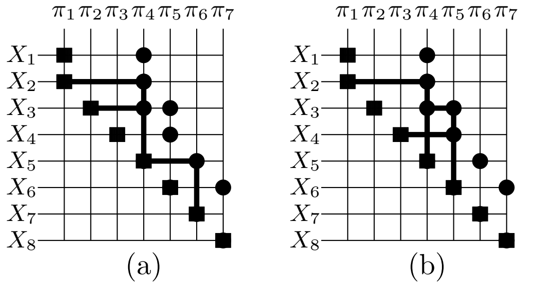


Figure 4: Active trails: (a) $\{7\}$-active trail $X_2\pi_1$, $X_2\pi_4$, $X_5\pi_4, X_5\pi_6$, $X_7\pi_6$, $X_5\pi_6$, $X_5\pi_4$, $X_3\pi_4, X_3\pi_2$; (b) $\{5,6\}$-active trail $X_2\pi_1$, $X_2\pi_4$, $X_5\pi_4, X_3\pi_4$, $X_3\pi_5, X_6\pi_5, X_4\pi_5, X_4\pi_3$

a direct succession. This is possible because both of them are adjacent to a box-marker $X_6\pi_5$.

Other examples of active trails can be seen in Figure 4. The trail in Figure 4(a) contains two consecutive vertical connections $X_5\pi 6, X_7\pi_6, X_5\pi 6$ with the common box-marker $X_7\pi_6$. This is possible because the trail is 7-active. Notice also that in this trail there appear some connections twice, which is not forbidden by the definition.

The trail in Figure 4(b) is $\{5,6\}$-active.In this trail there are two consecutive vertical connections $X_2\pi_4, X_5\pi_4, X_3\pi_4$, which is allowed since the common adjacent marker $X_5\pi_4$ correspond to variable X_5 and the trail is $\{5,6\}$-active. An analogous property holds also for the other couple of consecutive vertical connections $X_3\pi_5, X_6\pi_5, X_4\pi_5$.

Let us now present the main result of this contribution.

Theorem 2 *Consider a generating sequence $\pi_1,\ldots,\pi_n$, and three disjoint subsets $I,J,L \subset K_1 \cup \ldots \cup K_n$ such that $I \neq \emptyset \neq J$. If there does not exist an L-active trail $X_i \rightsquigarrow_L X_j$ in the corresponding persegram with $i \in I$ and $j \in J$ then the groups of variables X_I and X_J are conditionally independent given variables X_L under the distribution $\pi_1 \triangleright \ldots \triangleright \pi_n$:*

$$X_I \perp\!\!\!\perp X_J | X_L [\pi_1 \triangleright \ldots \triangleright \pi_n].$$

The proof of this assertion is rather technical and requires some lemmas proved in previous papers and therefore we adjourn it to the appendix.

Remark Let us say that conditional independence relations determined from a persegram are those, which are necessary for any distribution represented by a generating sequence with the given persegram. This system of conditional independence relations is also maximal in the sense that if there exists an active trail $X_j \rightsquigarrow_L X_k$ then there exists a distribution represented by a generating sequence with the given persegram, and variables X_j, X_k are conditionally dependent given variables X_L under this distribution.

Example 3 Consider a generating sequence

$$\pi_1(x_1), \pi_2(x_2), \pi_3(x_1,x_2,x_3),$$
$$\pi_4(x_2,x_3,x_4), \pi_5(x_3,x_5),$$

and show how to read all the (conditional) independence relations from its persegram (see Figure 5(a)). Let us stress that we do not present here a general algorithm; this should be based on the principles employed in algorithms for seeking all paths in graphs.

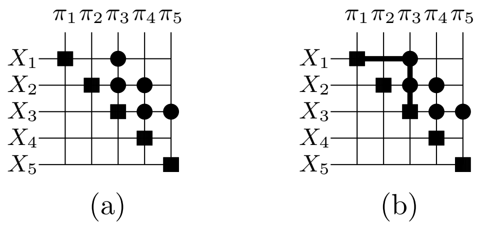
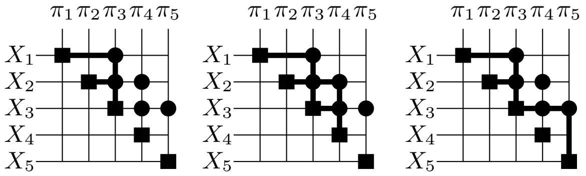


Figure 5: Persegram of a sequence from Example 3

It is obvious that the trail connecting X_1 and X_3 (see Figure 5(b)) is an L-active trail for any $L \subseteq \{2,4,5\}$ (including $\emptyset$). Therefore, variables X_1 and X_3 cannot be (conditionally) independent. The same holds also for couples (X_2,X_3), (X_2,X_4), (X_3,X_4), (X_3,X_5). Therefore, in what follows we shall investigate only the remaining couples: (X_1,X_2), (X_1,X_4), (X_1,X_5), (X_2,X_5), (X_4,X_5).

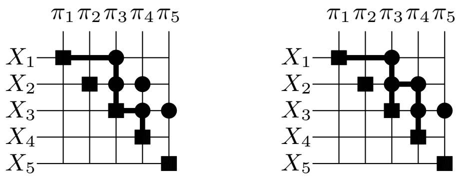
Let us examine for which L there exist L-active trails connecting X_1 and X_2. One can easily verify that there is no such a trail with $L = \emptyset$.

The trail in Figure 6(a) is L-active for any L equaling to $\{3\}$, $\{3,4\}$, $\{3,5\}$, $\{3,4,5\}$. The trail in Figure 6(b) is L-active for $L = \{4\}$ and



(a) (b) (c)

Figure 6: Active trails connecting X_1 and X_2: (a) $X_1\pi_1, X_1\pi_3, X_3\pi_3, X_2\pi_3, X_2\pi_2$; (b) $X_1\pi_1, X_1\pi_3, X_3\pi_3, X_3\pi_4,\ X_4\pi_4,\ X_2\pi_4,\ X_2\pi_2$; (c) $X_1\pi_1,\ X_1\pi_3,\ X_3\pi_3,\ X_3\pi_5,\ X_5\pi_5,\ X_3\pi_5,\ X_3\pi_3,\ X_2\pi_3,\ X_2\pi_2$



(a) (b)

Figure 7: Active trails connecting X_1 and X_4: (a) $X_1\pi_1, X_1\pi_3, X_3\pi_3, X_3\pi_4, X_4\pi_4$, (b) $X_1\pi_1, X_1\pi_3, X_3\pi_3, X_2\pi_3, X_2\pi_4, X_4\pi_4$

$L = \{4, 5\}$. The trail in Figure 6(c) is L-active for $L = \{5\}$ (and also $L = \{4, 5\}$, for which the previous trail was also L-active). Summarizing the up to now achieved results we get that there exist L-active trails connecting X_1 and X_2 whenever L is a non-empty subset of $\{3, 4, 5\}$. Therefore X_1 and X_2 are (unconditionally) independent but not conditionally independent given any (non-empty) subset of the remaining variables.

How is it with the couple X_1 and X_4? Examining the persegram in Figure 7 one can see that all L-active trails connecting X_1 and X_4, must contain at least one horizontal connection corresponding either to X_2 or to X_3. Therefore, there exists neither a $\{2, 3\}$-active trail nor $\{2, 3, 5\}$-active trail connecting variables X_1 and X_4. On the other hand, for all the remaining subsets there exists at least one L-active trail connecting X_1 and X_4: trail in Figure 7(a) is L-active for $L = \emptyset$, $\{2\}$, $\{5\}$ and $\{2, 5\}$, whereas the trail in Figure 7(b) is L-active for $L = \{3\}$ and $\{3, 5\}$. Summarizing this we get that variables X_1 and X_4 are conditionally independent only for conditioning sets $\{X_2, X_3\}$ and $\{X_2, X_3, X_5\}$.

The rest of the example is simple since all the remaining couples contain variable X_5 and all trails connecting this variable with any other must contain marker $X_3\pi_5$. Therefore we leave it to the reader to show that

$$
\begin{array}{ll}
X_2 \perp\!\!\!\perp X_5 | X_3, & X_2 \perp\!\!\!\perp X_5 | X_1, X_3, \\
X_2 \perp\!\!\!\perp X_5 | X_3, X_4, & X_2 \perp\!\!\!\perp X_5 | X_1, X_3, X_4, \\
X_4 \perp\!\!\!\perp X_5 | X_3, & X_4 \perp\!\!\!\perp X_5 | X_1, X_3, \\
X_4 \perp\!\!\!\perp X_5 | X_2, X_3, & X_4 \perp\!\!\!\perp X_5 | X_1, X_2, X_3.
\end{array}
$$

4 Conclusions

In the paper we presented a new notion of an L-active trail. It enabled us to read from a persegram corresponding to a generating sequence all the conditional independence relations guaranteed by a structure of a compositional model. We also showed in Theorem 1 that persegrams can be used to uncover that two permutations of a system of low-dimensional distributions represent the same multidimensional model.

Acknowledgements

This work was partially supported by MŠMT ČR under grants 1M0572 and 2C06019.

References

[1] D. Geiger, T. Verma, J. Pearl (1989) d-Separation: From Theorems to Algorithms. In *Proceedings of the conference UAI'89*, pages 139-148, Windsor, Ontario, Canada, August 1989.

[2] R. Jiroušek (1997) Composition of Probability Measures on Finite Spaces. In *Proceedings of the conf. UAI'97*, pages 274-281, San Francisco, U.S.A, August 1997.

[3] R. Jiroušek (2001) An anticipating operator and its role for computation of marginal distributions. In *Proceedings of the 3rd Czech-Japan Seminar on Data Analysis and Decision Making under Uncertainty*, pages 69-75, Osaka, Japan, September 2001.

[4] R. Jiroušek (2002) Detection of independence relations from persegrams. In *Proceedings of the conference IPMU 2002*,

volume C, pages 1261-1267, Annecy, France, July 2002.

[5] S.L. Lauritzen, A.P. Dawid, B.N. Larsen, H.G. Leimer (1990) Independence properties of directed Markov fields. *Networks*, 20, pages 491-505.

[6] T. Verma, J. Pearl (1988) Influence Diagrams and d-Separation, UCLA Cognitive Systems Laboratory, Technical Report 880052 (R-101), March 1988.

Appendix: Proof of Theorem 2

Lemma 1 *Let $K, L, M \subseteq N$. If $K \cup L \supseteq M \supseteq K \cap L$ then for any probability distributions $\pi \in \Pi^{(K)}$ and $\kappa \in \Pi^{(L)}$*

$$(\pi \triangleright \kappa)^{\downarrow M} = \pi^{\downarrow K \cap M} \triangleright \kappa^{\downarrow L \cap M}.$$

Lemma 2 *Let $\nu(x_{K \cup L}) = \pi(x_K) \triangleright \kappa(x_L)$ be defined. Then*

$$X_{K \setminus L} \perp\!\!\!\perp X_{L \setminus K} | X_{K \cap L}[\nu].$$

Lemma 3 *Consider a distribution $\pi(x_K)$ and two subsets $L_1, L_2 \subset K$ such that $L_1 \setminus L_2 \neq \emptyset \neq L_2 \setminus L_1$. Then*

$$X_{L_1 \setminus L_2} \perp\!\!\!\perp X_{L_2 \setminus L_1} | X_{L_1 \cap L_2}[\pi] \iff \pi^{\downarrow L_1 \cup L_2} = \pi^{\downarrow L_1} \triangleright \pi^{\downarrow L_2}.$$

Proof of the preceding Lemmas can be found in [2], the next Lemma was proved in [3].

Lemma 4 *If $K_2 \supseteq (K_1 \cap K_3)$ then*

$$\pi_1 \triangleright \pi_2 \triangleright \pi_3 = \pi_1 \triangleright (\pi_2 \triangleright \pi_3) = \pi_2 \triangleright \pi_3 \triangleleft \pi_1.$$

Proof of Theorem 2. The proof will be performed with the help of mathematical induction with respect to the length n of the generating sequence in question.

Let us start considering a generating sequence $\pi_1(x_{K_1}), \pi_2(x_{K_2})$. We know that any $X_i, X_j \in \mathbf{X}_{K_1}$ are connected by a trail consisting of a single vertical connection $(X_i\pi_1, X_j\pi_1)$ (see Figure 8(a)). The same holds also for $i, j \in K_2 \setminus K_1$. Realize that these single-connection trails $(X_i\pi_\ell, X_j\pi_\ell)$ are L-active trails for any

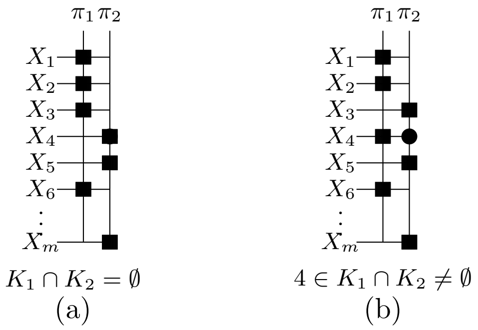


Figure 8: Examples of persegrams for a generating sequence π_1, π_2

$L \subseteq K_1 \cup K_2 \setminus \{i, j\}$. From Figure 8(b) we can also see that if $k \in K_1 \cap K_2 \neq \emptyset$ then

$$X_i\pi_1, X_k\pi_1, X_k\pi_2, X_j\pi_2$$

is a an $\emptyset$-active trail $X_i \rightsquigarrow_\emptyset X_j$ for $i \in K_1 \setminus K_2$ and $j \in K_2 \setminus K_1$. Moreover, it is also an L-active trail $X_i \rightsquigarrow_L X_j$ for any

$$L \subseteq (K_1 \cup K_2) \setminus \{i, j, k\}.$$

Therefore, we can easily answer the question when there does not exist a trail $X_i \rightsquigarrow_L X_j$. It happens if and only if i and j are not simultaneously in one of the sets K_1 or K_2, and if $L \supseteq K_1 \cap K_2$. It means that if I, J, L meets all the assumptions of the theorem, we know that (because there does not exist an L-active trail $X_i \rightsquigarrow_L X_j$ for $i \in I$ and $j \in J$) I must be a subset of one of the sets $K_1 \setminus K_2$ or $K_2 \setminus K_1$, J must be a subset of the other one from these two sets, and $L \supseteq K_1 \cap K_2$. Without loss of generality assume that $I \subseteq K_1 \setminus K_2$ and $J \subseteq K_2 \setminus K_1$, and, applying Lemma 1, compute

$$(\pi_1 \triangleright \pi_2)^{\downarrow I \cup J \cup L} = \pi_1^{\downarrow I \cup (L \cap K_1)} \triangleright \pi_2^{\downarrow J \cup (L \cap K_2)},$$

from which, using Lemma 2, we get

$$X_{I \cup (L \cap K_1) \setminus K_2} \perp\!\!\!\perp X_{J \cup (L \cap K_2) \setminus K_1} | X_{K_1 \cap K_2}[\pi_1 \triangleright \pi_2],$$

which, when marginalized, yields the required conditional independence

$$X_I \perp\!\!\!\perp X_J | X_L[\pi_1 \triangleright \pi_2],$$

which finishes the proof for $n = 2$.

Now, assume the assertion holds for all generating sequences of length less or equal $n \geq 2$.

We have to prove that it also holds for a generating sequence $\pi_1(x_{K_1}), \ldots, \pi_{n+1}(x_{K_{n+1}})$. This part of the proof, in which

$$M = K_{n+1} \cap (K_1 \cup \ldots \cup K_n),$$

will be performed in four successive steps:

A we will show that the assertion holds in case that $I \cup J \cup L \subseteq K_1 \cup \ldots \cup K_n$;

B under the assumption that $I \cup J \subseteq K_1 \cup \ldots \cup K_n$ and $L \cap (K_{n+1} \setminus M) \neq \emptyset$ we will prove validity of the *extended property*:
there is no L-active trail $X_i \rightsquigarrow_L X_j$ in the corresponding persegram with $i \in I \cup (M \setminus L)$ and $j \in J$;

C we will show that the extended property holds also in case that $J \subseteq K_1 \cup \ldots \cup K_n$ and $I \cap (K_{n+1} \setminus M) \neq \emptyset$;

D we will finish the proof by showing that the required conditional independence can be deduced from the extended property.

Notice that we need not consider the case with $I \cap (K_{n+1} \setminus M) \neq \emptyset \neq J \cap (K_{n+1} \setminus M)$, because in this situation there exists an L-active trail ($X_i \rightsquigarrow_L X_j$ with $i \in I$ and $j \in J$) consisting of one vertical connection, which violates assumptions of the theorem. Situation when $I \subseteq K_1 \cup \ldots \cup K_n$ and $J \cap (K_{n+1} \setminus M) \neq \emptyset$ is covered by step **C** after exchanging denotation of sets I and J.

Step A. So, let us assume the simplest situation when $I \cup J \cup L \subseteq K_1 \cup \ldots \cup K_n$, i.e. the box-markers of all the variables $X_{I \cup J \cup L}$ are not in the last column of the persegram corresponding to the generating sequence $\pi_1, \ldots, \pi_{n+1}$. Regarding the assumption that for any $i \in I$ and $j \in J$ there is no L-active trail $X_i \rightsquigarrow_L X_j$ in the persegram corresponding to $\pi_1, \ldots, \pi_{n+1}$, no such an L-active trail can exist in the persegram of $\pi_1, \ldots, \pi_n$. Therefore, due to the induction assumption,

$$X_I \perp\!\!\!\perp X_J | X_L[\pi_1 \triangleright \ldots \triangleright \pi_n].$$

Since $\pi_1 \triangleright \ldots \triangleright \pi_n$ is marginal to $\pi_1 \triangleright \ldots \triangleright \pi_{n+1}$,

$$X_I \perp\!\!\!\perp X_J | X_L[\pi_1 \triangleright \ldots \triangleright \pi_{n+1}]$$

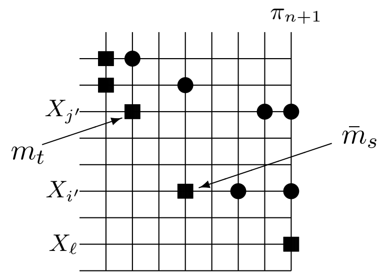


Figure 9: Construction of a trail $X_i \rightsquigarrow_L X_j$

holds also true.

Step B. Consider the situation when $I, J \subseteq K_1 \cup \ldots \cup K_n$ and $L \cap (K_{n+1} \setminus M) \neq \emptyset$. We will show that the set $M \setminus L$ can be added either to I or to J without violating the assumption of the theorem; we will show that either there does not exist an L-active trail $X_i \rightsquigarrow_L X_j$ for $i \in I \cup (M \setminus L)$ and $j \in J$, or there does not exist such a trail for $i \in I$ and $j \in J \cup (M \setminus L)$. Assume the opposite. Since there are not L-active trails from I to J, this assumption means that there are two L-active trails $X_i \rightsquigarrow_L X_{j'}$ and $X_j \rightsquigarrow_L X_{i'}$ for $i \in I, j \in J, i', j' \in M \setminus L$. Now we will show that from these two L-active trails it is always possible to construct an L-active trail $X_i \rightsquigarrow_L X_j$.

Let $m_0, \ldots, m_t$ and $\bar{m}_0, \ldots, \bar{m}_s$ denote trails $X_i \rightsquigarrow_L X_{j'}$ and $X_j \rightsquigarrow_L X_{i'}$, respectively. Choose any $\ell \in L \cap (K_{n+1} \setminus M)$. From Figure 9 it is obvious that:

1. if both (m_{t-1}, m_t) and $(\bar{m}_{s-1}, \bar{m}_s)$ are vertical connections, then
$$m_0, \ldots m_t, X_{j'}\pi_{n+1}, X_\ell\pi_{n+1}, X_{i'}\pi_{n+1}, \bar{m}_s, \bar{m}_{s-1}, \ldots, \bar{m}_0$$
is a required L-active trail $X_i \rightsquigarrow_L X_j$;
2. if (m_{t-1}, m_t) is a vertical and $(\bar{m}_{s-1}, \bar{m}_s)$ is a horizontal connection, then the required L-active trail is
$$m_0, \ldots m_t, X_{j'}\pi_{n+1}, X_\ell\pi_{n+1}, X_{i'}\pi_{n+1}, \bar{m}_{s-1}, \bar{m}_{s-2} \ldots, \bar{m}_0;$$
3. if (m_{t-1}, m_t) is a horizontal and $(\bar{m}_{s-1}, \bar{m}_s)$ is a vertical connection, then
$$m_0, \ldots m_{t-1}, X_{j'}\pi_{n+1}, X_\ell\pi_{n+1}, X_{i'}\pi_{n+1}, \bar{m}_s, \bar{m}_{s-1} \ldots, \bar{m}_0$$
is $X_i \rightsquigarrow_L X_j$;

4. if both (m_{t-1}, m_t) and $(\bar{m}_{s-1}, \bar{m}_s)$ are horizontal connections, then one can consider L-active trail

$$m_0, \ldots m_{t-1}, X_{j'}\pi_{n+1}, X_\ell\pi_{n+1}, \\ X_{i'}\pi_{n+1}, \bar{m}_{s-1}, \bar{m}_{s-2}, \ldots, \bar{m}_0,$$

which connects X_i and X_j.

Thus we have proved that $M \setminus L$ can always be added either to I or to J without violating the assumptions on non-existence of an L-active trail from I to J. Without loss of generality assume we can add it to I. So there does not exist an L-active trail $X_i \rightsquigarrow_L X_j$ for $i \in I \cup (M \setminus L)$ and $j \in J$ in the persegram corresponding to $\pi_1, \ldots, \pi_{n+1}$.

Step C. Now, we will show that the same property (*extended property*) holds also in the last case we have not considered yet. This step will again be performed by contradiction. Assume $J \subseteq K_1 \cup \ldots \cup K_n$ and $I \cap (K_{n+1} \setminus M) \neq \emptyset$, and assume there is an L-active trail $m_0, m_1, \ldots, m_t$, which is $X_j \rightsquigarrow_L X_{i'}$ for $j \in J$ and $i' \in I \cup (M \setminus L)$. Since we assume that there is no such a trail between I and J we know that the assumed trail must connect $j \in J$ with $i' \in M \setminus L$. However, again, this trail can be prolonged in a simple way to get an L-active trail $X_j \rightsquigarrow_L X_i$ for any $i \in I \cap (K_{n+1} \setminus M)$. If (m_{t-1}, m_t) is a vertical connection then such a trail is $m_0, m_1, \ldots, m_t, X_{i'}\pi_{n+1}, X_i\pi_{n+1}$. If (m_{t-1}, m_t) is a horizontal connection then the required trail is $m_0, m_1, \ldots, m_{t-1}, X_{i'}\pi_{n+1}, X_i\pi_{n+1}$.

Step D. So, up to now we have proved that if $J \subseteq K_1 \cup \ldots \cup K_n$ and either I or L (or both) has a nonempty intersection with $(K_{n+1} \setminus M)$, then there does not exists an L-active trail $X_i \rightsquigarrow_L X_j$ for $i \in I \cup (M \setminus L)$ and $j \in J$ in the persgram corresponding to $\pi_1, \pi_2, \ldots, \pi_{n+1}$. The more there does not exist an L-active trail $X_i \rightsquigarrow_{L \cap (K_1 \cup \ldots \cup K_n)} X_j$ in the presegram of $\pi_1, \ldots, \pi_n$ (for $i \in I \cap (K_1 \cup \ldots \cup K_n) \cup (M \setminus L)$ and $j \in J$).

In the rest of the proof we will use the following symbols: $\kappa_n = \pi_1 \triangleright \ldots \triangleright \pi_n$, $I^- = I \cap (K_1 \cup \ldots \cup K_n)$, $I^+ = I \setminus (K_1 \cup \ldots \cup K_n)$, $L^- = L \cap (K_1 \cup \ldots \cup K_n)$ and $L^+ = L \setminus (K_1 \cup \ldots \cup K_n)$. Using them the above expressed nonexistence of an L-active trail says that in the persegram of $\pi_1, \ldots, \pi_n$ there is no L-active trail $X_i \rightsquigarrow_{L^-} X_j$ for $i \in I^- \cup (M \setminus L)$ and $j \in J$. According to the induction assumption we can deduce that

$$X_{I^- \cup (M \setminus L)} \perp\!\!\!\perp X_J | X_{L^-} [\pi_1 \triangleright \ldots \triangleright \pi_n],$$

or, expressing this equivalently (due to Lemma 3)

$$\begin{aligned} \kappa_n^{\downarrow J \cup I^- \cup L^- \cup M} &= \kappa_n^{\downarrow J \cup I^- \cup L^- \cup (M \setminus L)} \\ &= \kappa_n^{J \cup L^-} \triangleright \kappa_n^{\downarrow I^- \cup (M \setminus L) \cup L^-} \\ &= \kappa_n^{J \cup L^-} \triangleright \kappa_n^{\downarrow I^- \cup M \cup L^-}. \quad (1) \end{aligned}$$

Since κ_n is marginal to $\pi_1 \triangleright \ldots \triangleright \pi_{n+1}$, it is evident that

$$\kappa_n^{\downarrow J \cup L^-} = (\pi_1 \triangleright \ldots \triangleright \pi_{n+1})^{\downarrow J \cup L^-}. \quad (2)$$

In the next computations we will also need the following equality (which is deduced with the help of Lemma 1)

$$\begin{aligned} (\pi_1 \triangleright \ldots \triangleright \pi_{n+1})^{\downarrow I \cup L \cup M} &= (\kappa_n \triangleright \pi_{n+1})^{\downarrow I \cup L \cup M} \\ &= \kappa_n^{\downarrow I^- \cup L^- \cup M} \triangleright (\pi_{n+1})^{\downarrow I^+ \cup L^+ \cup M}. \quad (3) \end{aligned}$$

In the following computation we use in successive steps Lemma 1, equality (1), Lemma 4 and finally equalities (2) and (3).

$$\begin{aligned} &(\pi_1 \triangleright \ldots \triangleright \pi_{n+1})^{\downarrow I \cup J \cup L \cup M} \\ &\quad = (\kappa_n \triangleright \pi_{n+1})^{\downarrow I \cup J \cup L \cup M} \\ &\quad = \kappa_n^{\downarrow I^- \cup J \cup L^- \cup M} \triangleright \pi_{n+1}^{\downarrow I^+ \cup L^+ \cup M} \\ &\quad = \kappa_n^{J \cup L^-} \triangleright \kappa_n^{\downarrow I^- \cup M \cup L^-} \triangleright \pi_{n+1}^{\downarrow I^+ \cup L^+ \cup M} \\ &\quad = \kappa_n^{J \cup L^-} \triangleright \left(\kappa_n^{\downarrow I^- \cup M \cup L^-} \triangleright \pi_{n+1}^{\downarrow I^+ \cup L^+ \cup M} \right) \\ &\quad = (\pi_1 \triangleright \ldots \triangleright \pi_{n+1})^{\downarrow J \cup L^-} \\ &\qquad\qquad \triangleright (\pi_1 \triangleright \ldots \triangleright \pi_{n+1})^{\downarrow I \cup L \cup M}. \end{aligned}$$

This yields (see Lemma 3)

$$X_J \perp\!\!\!\perp X_{I \cup L^+ \cup (M \setminus L^-)} | X_{L^-} [\pi_1 \triangleright \ldots \triangleright \pi_{n+1}],$$

from which the required independence

$$X_J \perp\!\!\!\perp X_I | X_L [\pi_1 \triangleright \ldots \triangleright \pi_{n+1}]$$

can be received by marginalization. □