Acta Math. Hungar., **154** (**2**) (2018), 511–524
https://doi.org/10.1007/s10474-018-0799-6
First published online February 22, 2018

ON PATTERNS OF CONDITIONAL INDEPENDENCES AND COVARIANCE SIGNS AMONG BINARY VARIABLES

F. MATÚŠ

Institute of Information Theory and Automation, Czech Academy of Sciences, Prague, Czech Republic
e-mail: matus@utia.cas.cz

(Received October 16, 2017; accepted December 7, 2017)

Abstract. Given finitely many events in a probability space, conditional independences among the indicators of events are considered simultaneously with the signs of covariances. Resulting discrete structures are studied restricting attention mostly to all couples and triples of events. Necessary and sufficient conditions for such structures to be represented by events are found. Consequences of the results for the patterns of conjunctive forks are discussed.

1. Introduction

Let N be a finite set. Elements i, j, k of N are not distinguished from singletons and the symbol for union of subsets I, J, K of N is omitted for simplicity. The set of couples $(ij|K)$ where i and j are not necessarily different and $i, j \notin K$ is denoted by $\mathcal{T}(N)$. A subset $\mathcal{L}$ of $\mathcal{T}(N)$ is called *probabilistically (p-) representable* if there exist random variables ξ_i, $i \in N$, taking finite numbers of values such that $(ij|K) \in \mathcal{L}$ if and only if ξ_i and ξ_j are conditionally independent given $\xi_K = (\xi_k)_{k\in K}$, in symbols $\xi_i \perp\!\!\!\perp \xi_j|\xi_K$. In particular, $\xi_i \perp\!\!\!\perp \xi_i|\xi_K$ means that ξ_i is a function of ξ_K. A set $\mathcal{L}$ p-represented by random variables is the *pattern* of conditional independences.

A subset $\mathcal{L}$ of $\mathcal{T}(N)$ is *binary (bin-) representable* if there exist events A_i in a probability space such that the indicator functions $\mathbf{1}_{A_i}$, $i \in N$, p-represent $\mathcal{L}$. If additionally $(\mathbf{1}_{A_i})_{i\in N}$ takes $2^{|N|}$ values with positive probabilities then $\mathcal{L}$ is *bin*$^+$*-representable*. This condition on the events is referred to as the positivity assumption. It excludes functional dependence and forces $\mathcal{L}$ to be contained in $\mathcal{R}(N)$, defined as the set of $(ij|K) \in \mathcal{T}(N)$ with $i \neq j$.

This research was supported by Grant Agency of the Czech Republic, Grant 16-12010S.
Key words and phrases: conditional independence, covariance, correlation, binary variable, ternary relation, conjunctive fork, forkness, betweenness, semigraphoid.
Mathematics Subject Classification: primary 62H20, secondary 62A01, 60A99, 62H05.

A mapping $\sigma\colon \binom{N}{2} \to \{-1,0,1\}$ is *sign-representable* by events A_i, $i \in N$, if σ_{ij} is equal to the sign of the covariance between $\mathbf{1}_{A_i}$ and $\mathbf{1}_{A_j}$, $i,j \in N$. Under the positivity assumption, it is *sign$^+$-representable.* A mapping σ sign-represented by events is the *pattern* of signs.

In this work, the pattern of conditional independences is studied simultaneously with the pattern of signs. A pair $[\mathcal{L},\sigma]$ is *event-representable* if there exist events A_i, $i \in N$, such that they bin-represent $\mathcal{L}$ and sign-represent σ. It is *event$^+$-representable* when the events bin$^+$-represent $\mathcal{L}$ and sign$^+$-represent σ.

Let $\mathcal{R}_0(N)/\mathcal{R}_1(N)$denote the set of couples $(ij|K) \in \mathcal{R}(N)$ with K empty/singleton and $\mathcal{R}_{01}(N)$be the union of $\mathcal{R}_0(N)$and $\mathcal{R}_1(N)$. Theorem 1 presents necessary and sufficient conditions for the event- and event$^+$-representability of any pair $[\mathcal{L},\sigma]$ with $\mathcal{L} \subseteq \mathcal{R}_{01}(N)$. The assertion is based on the following three notions.

For $\mathcal{L} \subseteq \mathcal{T}(N)$ and $M \subseteq N$ let $\mathcal{L}_{|M} = \mathcal{L} \cap \mathcal{T}(M)$ be the restriction of $\mathcal{L}$ to M. If M has three elements then it is a 3-restriction.

A subset $\mathcal{L}$ of $\mathcal{T}(N)$ is called *solvable* if the system of equations

$$x_{ij} = x_{ik} + x_{jk}\,, \quad (ij|k) \in \mathcal{L} \cap \mathcal{R}_1(N)\,, \tag{1}$$

in real indeterminates x_{ij}, $ij \in \binom{N}{2}$, has a solution with all coordinates positive, cf. [1]. Let $\mathcal{L}^\circ$ denote the set of $(ij|k) \in \mathcal{L} \cap \mathcal{R}_1(N)$ such that $(ik|j) \notin \mathcal{L}$ and $(jk|i) \notin \mathcal{L}$.

A sign mapping σ on $\binom{N}{2}$ is *adjusted* to $\mathcal{L} \subseteq \mathcal{T}(N)$ if

$$\sigma_{ij} = 0 \;\Leftrightarrow\; (ij|\emptyset) \in \mathcal{L} \cap \mathcal{R}_0(N)\,, \tag{2}$$

$$\sigma_{ij} = \sigma_{ik}\,\sigma_{jk}\,, \quad (ij|k) \in \mathcal{L} \cap \mathcal{R}_1(N)\,. \tag{3}$$

THEOREM 1. *For $\mathcal{L} \subseteq \mathcal{R}_{01}(N)$ and $\sigma\colon \binom{N}{2} \mapsto \{-1,0,1\}$, the following are equivalent*:

(i) *$[\mathcal{L},\sigma]$ is event$^+$-representable*,

(ii) *$[\mathcal{L},\sigma]$ is event-representable*,

(iii) *the 3-restrictions of $\mathcal{L}$ are bin-representable, $\mathcal{L}^\circ$ is solvable and σ is adjusted to $\mathcal{L}$.*

The assertion of (iii) can be algorithmically verified in polynomial time. In fact, the bin-representable sets over N with three elements are of seven types, by Lemma 3. The solvability reduces to the question whether an instance of the linear programming is feasible.

EXAMPLE 1. Let N be the vertex set of a tree and let $\mathcal{L}$ consist of the couples $(ij|k)$ of distinct vertices such that k is on the path between i and j. Let $\sigma_{ij} = (-1)^{d_{ij}}$ where d_{ij} is the length of the path between i and j. Then, $\mathcal{L} = \mathcal{L}^\circ$ has the 3-restrictions isomorphic to $\emptyset$ or $\{(12|3)\}$, which are both bin-representable. It is solvable, taking $x_{ij} = d_{ij}$. Since σ is adjusted to $\mathcal{L}$

the pair $[\mathcal{L}, \sigma]$ is event$^+$-representable by Theorem 1. This example settles an open question formulated after [11, Example 1].

Section 3 presents a constructive proof of Theorem 1. In Section 4, $\mathcal{L} \subseteq \mathcal{T}(N)$ features in Theorem 2, covering more general situations that can be reduced to that of Theorem 1. Section 5 discusses consequences, other related material and literature. Section 6 works out consequences of the results for the patterns of conjunctive forks.

2. Preliminaries

This section presents technicalities and four elementary lemmas.

Events live in a probability space $(\Omega, \mathcal{S}, P)$. The complement to an event $A \in \mathcal{S}$ is $\bar{A} = \Omega \setminus A$. The event is P-trivial if $P(A) \in \{0,1\}$. The indicator function $\mathbf{1}_A$ of A is equal to one if $\omega \in A$ and zero otherwise. The mean $E_P \mathbf{1}_A$ of $\mathbf{1}_A$ is $P(A)$. The centering $\zeta_A \triangleq \mathbf{1}_A - P(A)$ has the mean zero. The variance $E_P \zeta_A^2$ of ζ_A, or $\mathbf{1}_A$, or A, is $P(A) - P(A)^2$.

Events A, B are P-equal if $P(A \vartriangle B) = 0$, in symbols $A \stackrel{P}{=} B$. They are P-complementary if $A \stackrel{P}{=} \bar{B}$. The covariance $E_P(\zeta_A \zeta_B)$ of A and B equals $\langle A, B \rangle_P \triangleq P(A \cap B) - P(A)P(B)$. If A, B are not P-trivial then it has the same sign as the correlation between A and B

$$\frac{E_P(\zeta_A \zeta_B)}{\sqrt{E_P(\zeta_A^2)}\sqrt{E_P(\zeta_B^2)}} = \frac{\langle A, B \rangle_P}{\sqrt{\langle A, A \rangle_P}\sqrt{\langle B, B \rangle_P}}\,.$$

LEMMA 1. (i) $\langle A, B \rangle_P \leqslant \langle A, A \rangle_P$ *with equality if only if* A *is* P*-trivial or* $A \stackrel{P}{=} B$,

(ii) $|\langle A, B \rangle_P| \leqslant \langle A, A \rangle_P$ *with equality if only if* A *is* P*-trivial or* $A \stackrel{P}{=} B$ *or* $A \stackrel{P}{=} \bar{B}$.

(iii) $\langle A, B \rangle_P^2 \leqslant \langle A, A \rangle_P \langle B, B \rangle_P$ *with equality if only if* A *or* B *is* P*-trivial, or* $A \stackrel{P}{=} B$ *or* $A \stackrel{P}{=} \bar{B}$.

PROOF. (i) The inequality is equivalent to

$$P(A)P(\bar{A} \cap B) + P(\bar{A})P(A \cap \bar{B}) \geqslant 0.$$

(ii) Since $\langle \bar{A}, B \rangle_P = -\langle A, B \rangle_P$ the inequality is obtained from *(i)* and its instance with A replaced by $\bar{A}$, by $-\langle A, B \rangle_P = \langle \bar{A}, B \rangle_P \leqslant \langle \bar{A}, \bar{A} \rangle_P = \langle A, A \rangle_P$.

(iii) This follows from (ii) and its instance with A interchanged with B. $\square$

Actually, (iii) is the Cauchy inequality $E_P(\zeta_A \zeta_B)^2 \leqslant E_P(\zeta_A^2) E_P(\zeta_B^2)$. It is tight if and only if one of ζ_A, ζ_B is proportional to the other, P-a.s.

A modern approach to the conditional stochastic independence is summarized in [10]. For events A and B, the unconditional independence $\mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_B$ is equivalent to $\langle A, B\rangle_P = 0$, vanishing of the covariance $E_P(\zeta_A \zeta_B)$. For events A, B and C, the conditional independence $\mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_B|\mathbf{1}_C$ is equivalent to $A \perp\!\!\!\perp B|C$ and $A \perp\!\!\!\perp B|\bar{C}$, thus to

$$P(A\cap B\cap C)\,P(C) = P(A\cap C)\,P(B\cap C)\,, \tag{4}$$

$$P(A\cap B\cap \bar{C})\,P(\bar{C}) = P(A\cap \bar{C})\,P(B\cap \bar{C})\,, \tag{5}$$

respectively. For P-trivial event C this reduces to the unconditional independence $\mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_B$. An event A is P-trivial if and only if $\mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_A$. By (4) and (5),

$$\mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_A|\mathbf{1}_B \;\Leftrightarrow\; \big[A \text{ is } P\text{-trivial or } A \overset{P}{=} B \text{ or } A \overset{P}{=} \bar{B}\big]\,. \tag{6}$$

The two equalities in disjunction express that $\mathbf{1}_A$ is a function of $\mathbf{1}_B$ and vice versa, P-a.s.

The following implication goes back at least to [13, p. 160].

LEMMA 2. *If* $\mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_B|\mathbf{1}_C$ *then* $\langle A,B\rangle_P \langle C,C\rangle_P = \langle A,C\rangle_P \langle B,C\rangle_P$.

PROOF. The difference $\langle A,B\rangle_P \langle C,C\rangle_P - \langle A,C\rangle_P \langle B,C\rangle_P$ equals

$$\big[P(A\cap B\cap C)\,P(C) - P(A\cap C)\,P(B\cap C)\big]\,P(\bar{C})$$
$$+\big[P(A\cap B\cap \bar{C})\,P(\bar{C}) - P(A\cap \bar{C})\,P(B\cap \bar{C})\big]\,P(C)$$

[6, p. 28, eq. (10)]. By the assumption, the brackets vanish, see (4) and (5). □

The most basic property of conditional independence

$$\big[\mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_B|\mathbf{1}_C \text{ and } \mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_C\big] \;\Leftrightarrow\; \big[\mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_C|\mathbf{1}_B \text{ and } \mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_B\big] \tag{7}$$

holds even for σ-algebras [2]. By Lemma 2, $\mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_B|\mathbf{1}_C$ and $\mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_B$ imply $\mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_C$ or $\mathbf{1}_B \perp\!\!\!\perp \mathbf{1}_C$. On account of (7),

$$\begin{aligned} &\big[\mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_B|\mathbf{1}_C \text{ and } \mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_B\big] \\ &\Leftrightarrow\; \big[\mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_C|\mathbf{1}_B,\ \mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_C \text{ or } \mathbf{1}_B \perp\!\!\!\perp \mathbf{1}_C|\mathbf{1}_A,\ \mathbf{1}_B \perp\!\!\!\perp \mathbf{1}_C\big]\,. \end{aligned} \tag{8}$$

If $\mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_B|\mathbf{1}_C$ and $\mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_C|\mathbf{1}_B$ then two instances of Lemma 2 combine to

$$\langle A,B\rangle_P \cdot \langle A,C\rangle_P \cdot \big[\langle B,B\rangle_P \langle C,C\rangle_P - \langle B,C\rangle_P^2\big] = 0\,.$$

By Lemma 1(iii) and (6),

$$\begin{aligned} &\big[\mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_B|\mathbf{1}_C \text{ and } \mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_C|\mathbf{1}_B\big] \\ &\Rightarrow\; \big[\mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_B \text{ or } \mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_C \text{ or } \mathbf{1}_B \perp\!\!\!\perp \mathbf{1}_B|\mathbf{1}_C \text{ or } \mathbf{1}_C \perp\!\!\!\perp \mathbf{1}_C|\mathbf{1}_B\big]\,. \end{aligned} \tag{9}$$

LEMMA 3. *If $\mathcal{L} \subseteq \mathcal{R}(N)$ with $N = \{1,2,3\} = 123$ is bin-representable then it is isomorphic to one of the following*

$$\begin{cases} \emptyset \quad \{(12|\emptyset)\} \quad \{(12|\emptyset),(13|\emptyset)\} \quad \{(12|\emptyset),(13|\emptyset),(23|\emptyset)\} \\ \{(12|3)\} \quad \{(12|3),(13|2),(12|\emptyset),(13|\emptyset)\} \quad \mathcal{R}(123)\,. \end{cases} \tag{10}$$

Each of these sets is bin$^+$-representable.

PROOF. For $\mathcal{L} \subseteq \mathcal{R}(123)$, if $\{(12|3)\} \subsetneq \mathcal{L}$ then $(13|\emptyset) \in \mathcal{L}$ or $(12|\emptyset) \in \mathcal{L}$ or $(13|2) \in \mathcal{L}$, up to a permutation. Applying (7)–(9), bin-representable $\mathcal{L}$ is isomorphic to one of the last two sets in (10). In (9), functional dependences played no role because $(2|3) \notin \mathcal{L}$ and $(3|2) \notin \mathcal{L}$. Otherwise, $\mathcal{L}$ is isomorphic to one of the first five sets in (10).

The second assertion can be verified case by case, which is omitted. □

For $\mathcal{L} \subseteq \mathcal{T}(N)$ let $E_{\mathcal{L}}$ denote union of the sets $\{ij, ik, jk\}$ over $(ij|k) \in \mathcal{L} \cap \mathcal{R}_1(N)$. This union indexes the variables x_{ij} that effectively occur in the definition of solvability.

LEMMA 4. *If $\mathcal{L} \subseteq \mathcal{T}(N)$ is bin-representable then $\mathcal{L}^\circ$ is solvable.*

PROOF. Let A_i, $i \in N$, bin-represent $\mathcal{L}$. For $ij \in E_{\mathcal{L}^\circ}$ there exists $k \in N$ different from i and j such that $\mathcal{L} \cap \mathcal{R}_1(ijk)$ is isomorphic to $\{(12|3)\}$. By (7) and (8), $\mathcal{L} \cap \mathcal{R}(ijk)$ is also isomorphic to $\{(12|3)\}$. Hence, all terms in

$$x_{ij} \triangleq \ln \frac{\sqrt{\langle A_i, A_i\rangle_P}\,\sqrt{\langle A_j, A_j\rangle_P}}{|\langle A_i, A_j\rangle_P|}\,, \quad ij \in E_{\mathcal{L}^\circ}\,,$$

are positive. The ratio is greater than one by Lemma 1(iii) where the inequality is strict. In fact, A_i and A_j are not P-equal or P-complementary because this implies that $(ik|j)$ and $(jk|i)$ be in $\mathcal{L}$. Thus, $x_{ij} > 0$. Eqs. (1) with $\mathcal{L}$ replaced by $\mathcal{L}^\circ$ follow from Lemma 2, applied to A_i, A_j and A_k. Thus, $\mathcal{L}^\circ$ is solvable. □

3. Constructing event$^+$-representations

In this section a constructive proof of Theorem 1 is worked out.

When considering only events A_i, $i \in N$, in a probability space $(\Omega, \mathcal{S}, P)$ there is no loss of generality in assuming that Ω is the Abelian group $\mathbb{Z}_2^N$ with the addition coordinatewise, all subsets of Ω are $\mathcal{S}$-measurable and the event A_i consists of all $\omega = (\omega_j)_{j \in N}$ with the j-th coordinate ω_j equal to 1. Thus, P is the very distribution of $(\mathbf{1}_{A_i})_{i \in N}$. The positivity of P means that $P(\omega) > 0$ for all $\omega \in \Omega$.

Elementary instances of general results from the harmonic analysis have the form

$$P(\omega) = 2^{-|N|} \sum_{L \subseteq N} \hat{P}(L)\, \chi_L(\omega) \quad \text{and} \quad \hat{P}(L) = \sum_{\omega \in \mathbb{Z}_2^N} P(\omega)\, \chi_L(\omega) \tag{11}$$

where $\hat{P}$ is the Fourier–Stieltjes transform of P and χ_L are the characters of the group $\mathbb{Z}_2^N$. Recall that $\chi_L(\omega)$ equals $+1$ if ω has even number of 1's within its coordinates indexed by L and -1 otherwise. For $I \subseteq N$ let A_I denote the intersection of A_i over $i \in I$. Then

$$P(A_I) = 2^{-|I|} \sum_{L \subseteq I} (-1)^{|L|} \hat{P}(L) \tag{12}$$

follows from (11) by summing over ω having $\omega_i = 1$ for all $i \in I$.

Proof of Theorem 1. (i) $\Rightarrow$ (ii). This holds by definitions.

(ii) $\Rightarrow$ (iii). For $\mathcal{L} \subseteq \mathcal{R}(N)$ and $\sigma\colon \binom{N}{2} \mapsto \{-1, 0, 1\}$, let $[\mathcal{L}, \sigma]$ be event-representable by A_i, $i \in N$. Any 3-restriction of $\mathcal{L}$ is bin-representable. Lemma 4 implies that $\mathcal{L}^\circ$ is solvable. As σ is the pattern of signs, (2) holds. The implication (3) follows from Lemma 2 by taking signs. Hence, σ is adjusted to $\mathcal{L}$.

(iii) $\Rightarrow$ (i). Let $\mathcal{L} \subseteq \mathcal{R}_{01}(N)$, every 3-restriction of $\mathcal{L}$ be bin$^+$-representable, $\mathcal{L}^\circ$ be solvable and σ be adjusted to $\mathcal{L}$. There exist $x_{ij} > 1$ with $ij \in E_{\mathcal{L}^\circ}$ that witness the solvability. For $ij \in \binom{N}{2} \setminus E_{\mathcal{L}^\circ}$ let $x_{ij} = 1$.

The space $\Omega = \mathbb{Z}_2^N$ and events A_i, $i \in N$, are taken as above. Let $F_{\mathcal{L}}$ denote the family of $ijk \in \binom{N}{3}$ such that $\mathcal{L}_{|ijk} \subseteq \mathcal{R}_0(N)$. A probability measure P on Ω is constructed by (11)

$$2^{|N|} P = 1 + \sum_{ij \in \binom{N}{2}} \sigma_{ij}\, \gamma^{x_{ij}} \cdot \chi_{ij} + \sum_{ijk \in F_{\mathcal{L}}} \delta_{ijk} \cdot \chi_{ijk} + \sum_{L \subseteq N, |L| \geqslant 4} \varepsilon^{|L|^2} \cdot \chi_L$$

where $0 < \gamma < \frac{1}{3}\, |N|^{-2}$, $0 < \delta_{ijk} < \frac{1}{3}\, |N|^{-3}$, $\varepsilon > 0$ and $\varepsilon^4 (1 + \varepsilon)^{|N|} < \frac{1}{3}$. The strict inequalities imply that P is positive. Additionally, let δ_{ijk} be different from the difference $\sigma_{ij}\gamma^{x_{ij}} - \sigma_{ik}\sigma_{jk}\gamma^{x_{ik}+x_{jk}}$ and from the two differences obtained by permuting for i, j, k.

Equation (12) implies that $P(A_i) = \frac{1}{2}$ and $P(A_{ij}) = \frac{1}{4}[1 + \sigma_{ij}\gamma^{x_{ij}}]$, $ij \in \binom{N}{2}$. Then, $\mathrm{sign}\langle A_i, A_j \rangle_P$ equals σ_{ij} so that the events sign$^+$-represent σ. Since σ is adjusted to $\mathcal{L}$ it follows from (2) that $(ij|\emptyset) \in \mathcal{L}$ if and only if $\mathbf{1}_{A_i} \perp\!\!\!\perp \mathbf{1}_{A_j}$.

By (12),

$$8P(A_{ijk}) = 1 + \sigma_{ij}\gamma^{x_{ij}} + \sigma_{ik}\gamma^{x_{ik}} + \sigma_{jk}\gamma^{x_{jk}} - \begin{cases} \delta_{ijk}\,, & ijk \in F_{\mathcal{L}}\,, \\ 0\,, & ijk \in \binom{N}{3} \setminus F_{\mathcal{L}}\,. \end{cases}$$

The independence $\mathbf{1}_{A_i} \perp\!\!\!\perp \mathbf{1}_{A_j}|\mathbf{1}_{A_k}$ is expressed, in accordance with (4) and (5), as

$$P(A_{ijk}) \cdot \tfrac{1}{2} = \tfrac{1}{16} \cdot \left[1 + \sigma_{ik}\gamma^{x_{ik}}\right] \cdot \left[1 + \sigma_{jk}\gamma^{x_{jk}}\right], \tag{13}$$

$$[P(A_{ij}) - P(A_{ijk})] \cdot \tfrac{1}{2} = \tfrac{1}{16} \cdot \left[1 - \sigma_{ik}\gamma^{x_{ik}}\right] \cdot \left[1 - \sigma_{jk}\gamma^{x_{jk}}\right]. \tag{14}$$

To verify that $(ij|k) \in \mathcal{L}$ is equivalent to $\mathbf{1}_{A_i} \perp\!\!\!\perp \mathbf{1}_{A_j}|\mathbf{1}_{A_k}$, the 3-restriction $\mathcal{L}_{|ijk}$ is considered. It is bin-representable by assumption, thus isomorphic to some of the sets in (10).

If $\mathcal{L}_{|ijk}$ is contained in $\mathcal{R}_0(N)$ then (13) rewrites to $\delta_{ijk} = \sigma_{ij}\gamma^{x_{ij}} - \sigma_{ik}\sigma_{jk}\gamma^{x_{ik}+x_{jk}}$ which fails by the choice of δ_{ijk}. Thus, $\mathbf{1}_{A_i} \perp\!\!\!\perp \mathbf{1}_{A_j}|\mathbf{1}_{A_k}$ fails and, similarly, $\mathbf{1}_{A_i} \perp\!\!\!\perp \mathbf{1}_{A_k}|\mathbf{1}_{A_j}$ and $\mathbf{1}_{A_j} \perp\!\!\!\perp \mathbf{1}_{A_k}|\mathbf{1}_{A_i}$ do as well. Hence, $\mathcal{L}_{|ijk}$ is bin$^+$-representable by A_i, A_j and A_k.

If $\mathcal{L}_{|ijk} = \mathcal{R}(ijk)$ then (13) and (14) reduce to $\frac{1}{16} = \frac{1}{16}$ because $ijk \notin F_{\mathcal{L}}$, and σ_{ij}, σ_{ik} and σ_{jk} vanish, remembering the events sign-represent σ and σ is adjusted to $\mathcal{L}$. Again, $\mathcal{L}_{|ijk}$ is bin$^+$-representable by A_i, A_j and A_k.

If $\mathcal{L}_{|ijk}$ is isomorphic to $\{(12|3), (13|2), (12|\emptyset), (13|\emptyset)\}$ with i corresponding to 1 then (13) and (14) reduce to $\frac{1}{16}[1 \pm \sigma_{jk}\,\gamma^{x_{jk}}] = \frac{1}{16}[1 \pm \sigma_{jk}\,\gamma^{x_{jk}}]$. Therefore, $\mathbf{1}_{A_i} \perp\!\!\!\perp \mathbf{1}_{A_j}|\mathbf{1}_{A_k}$. By symmetry, $\mathbf{1}_{A_i} \perp\!\!\!\perp \mathbf{1}_{A_k}|\mathbf{1}_{A_j}$. However, $\mathbf{1}_{A_j} \perp\!\!\!\perp \mathbf{1}_{A_k}|\mathbf{1}_{A_i}$ fails because (13) with ijk permuted to jki takes the form

$$\tfrac{1}{16}[1 + \sigma_{jk}\,\gamma^{x_{jk}}] = \tfrac{1}{16}\,.$$

Here, $\sigma_{jk} \neq 0$ because σ is adjusted to $\mathcal{L}$. Hence, $\mathcal{L}_{|ijk}$ is bin$^+$-representable by A_i, A_j and A_k.

If $\mathcal{L}_{|M} = \{(ij|k)\}$ then (13) and (14) are equivalent to

$$1 + \sigma_{ij}\,\gamma^{x_{ij}} \pm \sigma_{ik}\,\gamma^{x_{ik}} \pm \sigma_{jk}\,\gamma^{x_{jk}} = [1 \pm \sigma_{ik}\,\gamma^{x_{ik}}][1 \pm \sigma_{jk}\,\gamma^{x_{jk}}]\,.$$

Since $(ij|k) \in \mathcal{L}^\circ$ the equations hold by the solvability (1) and adjustment (3) assumptions. Transposing i and k, eq. (13) rewrites to $\sigma_{jk}\,\gamma^{x_{jk}} = \sigma_{ik}\,\gamma^{x_{ik}}\sigma_{ij}\,\gamma^{x_{ij}}$ and fails. In fact, σ_{ij}, σ_{ik} and σ_{jk} are nonzero by adjustment. Then, the equation $x_{ij} = x_{ik} + x_{jk}$, $x_{ik} > 0$, from the definition of solvability, implies $1 = \gamma^{2x_{ik}}$ where $\gamma < 1$ and $x_{ik} \geqslant 1$, by construction. By symmetry, $\mathcal{L}_{|ijk}$ is bin$^+$-representable by A_i, A_j and A_k.

So far it is verified that $(ij|K) \in \mathcal{L}$ is equivalent to $\mathbf{1}_{A_i} \perp\!\!\!\perp \mathbf{1}_{A_j}|(\mathbf{1}_{A_k})_{k\in K}$ whenever $|K| \leqslant 1$.

The probability $P(A_I)$, $|I| \leqslant 3$, is independent of ε. Remembering that the characters are orthogonal, $\hat{P}(L) = \varepsilon^{|L|^2}$, $|L| \geqslant 4$, by (11). The leading monomial of $P(A_I)$ is $\varepsilon^{|I|^2}$, $|I| \geqslant 4$, by (12). Hence, if $(ij|K) \in \mathcal{R}(N)$ and $|K| \geqslant 4$ then $P(A_{ijK})P(A_K) - P(A_{iK})P(A_{jK})$ has the degree $(|K|+2)^2 + |K|^2$. If $|K|$ equals 2/3, it has the degree 16/256. When ε avoids the roots of those nonzero polynomials, $\mathbf{1}_{A_i}$ depends on $\mathbf{1}_{A_j}$ given $(\mathbf{1}_{A_k})_{k\in K}$, $|K| \geqslant 2$, in accordance with $(ij|K) \notin \mathcal{L} \subseteq \mathcal{R}_{01}(N)$. Therefore, $\mathcal{L}$ is bin$^+$-representable. In turn, $[\mathcal{L}, \sigma]$ is event$^+$-representable. $\square$

The discrete Fourier–Stieltjes transform [3] featured previously in the proof of [8, Theorem 2], dealing with unconditional independence among binary variables. The above construction of the probability measure P is inspired by the one that was proposed by the author in the proof of the unpublished [1, Theorem 1]. The measure can be arbitrarily close to the uniform probability measure on Ω.

4. Event-representations

In this section, Theorem 1 is extended to situations where events are allowed to be P-trivial, P-equal or P-complementary.

An element $\ell \in N$ is a *loop* of $\mathcal{L} \subseteq \mathcal{T}(N)$ if $(\ell|\emptyset) \in \mathcal{L}$. Let $\lambda(\mathcal{L})$ denote the set of loops. If ξ_i, $i \in N$, p-represent $\mathcal{L}$ then to be a loop ℓ means that ξ_ℓ is constant, a.s. Hence,

$$i \in \lambda(\mathcal{L}) \text{ and } i,j \notin K \quad \Rightarrow \quad (ij|K) \in \mathcal{L}, \tag{15}$$

$$(ij|K) \in \mathcal{L} \quad \Leftrightarrow \quad (ij|K \setminus \lambda(\mathcal{L})) \in \mathcal{L} \text{ and } i,j \notin K \cap \lambda(\mathcal{L}). \tag{16}$$

Elements i and j from $N \setminus \lambda(\mathcal{L})$ are *parallel* in $\mathcal{L}$, in symbols $i \|_{\mathcal{L}} j$, if $i = j$, or both $(i|j)$ and $(j|i)$ belong to $\mathcal{L}$. When $\mathcal{L}$ is p-representable by ξ_i, $i \in N$, to be parallel means that ξ_i is a function of ξ_j and vice versa, a.s. Therefore,

$$\text{the relation } \|_{\mathcal{L}} \text{ is an equivalence on } N \setminus \lambda(\mathcal{L}), \tag{17}$$

$$\begin{array}{l} (ij|K) \in \mathcal{L},\ i\|_{\mathcal{L}} i',\ i', j \notin K',\ K \text{ and } K' \text{ intersect} \\ \text{the same blocks of } \|_{\mathcal{L}} \Rightarrow (i'j|K') \in \mathcal{L}, \end{array} \tag{18}$$

$$\begin{array}{l} (i|k) \in \mathcal{L},\ i,j \notin K,\ K \text{ intersects the block of } \|_{\mathcal{L}} \\ \text{containing } k \Rightarrow (ij|K) \in \mathcal{L}. \end{array} \tag{19}$$

A set $\mathcal{L} \subseteq \mathcal{T}(N)$ is *regular* if it enjoys (15)–(19). The p-representable sets are regular. Every $\mathcal{L} \subseteq \mathcal{R}(N)$ is regular, having $\lambda(\mathcal{L}) = \emptyset$ and $i \|_{\mathcal{L}} j$ equivalent to $i = j$. A *crossing* of a regular set $\mathcal{L}$ is a set $M \subseteq N \setminus \lambda(\mathcal{L})$ that contains exactly one element of each block of $\|_{\mathcal{L}}$.

LEMMA 5. *Every regular set $\mathcal{L} \subseteq \mathcal{T}(N)$ is uniquely given by $\lambda(\mathcal{L})$, $\|_{\mathcal{L}}$ and the restriction of $\mathcal{L}$ to any crossing M of $\mathcal{L}$.*

PROOF. If $\mathcal{L}$ satisfies (15)–(16) then it recovers uniquely from its restriction to $N \setminus \lambda(\mathcal{L})$. If the restriction satisfies (17)–(19) then it recovers uniquely from $\mathcal{L}_{|M}$ for any crossing M of $\mathcal{L}$. In fact, let $(ij|K) \in \mathcal{T}(N \setminus \lambda(\mathcal{L}))$. There exist unique $i', j' \in M$ such that $i \|_{\mathcal{L}} i'$ and $j \|_{\mathcal{L}} j'$. Let $K' \subseteq M$ intersect the same blocks of $\|_{\mathcal{L}}$ as K. In the case $i', j' \notin K'$, $(ij|K) \in \mathcal{L}$ is equivalent to $(i'j'|K') \in \mathcal{L}_{|M}$, by (18). Otherwise, $i' \in K'$ or $j' \in K'$ whence $i\|_{\mathcal{L}}k$ or $j\|_{\mathcal{L}}k$ for some $k \in K$, due to (17). Then (19) implies $(ij|K) \in \mathcal{L}$. $\square$

Let $\sigma_{|M}$ denote the restriction of σ to $\binom{M}{2}$.

THEOREM 2. *For $\mathcal{L} \subseteq \mathcal{T}(N)$ regular, $\sigma\colon \binom{N}{2} \to \{-1,0,1\}$ adjusted to $\mathcal{L}$ and a crossing M of $\mathcal{L}$, $[\mathcal{L},\sigma]$ is event-representable if and only if $[\mathcal{L}_{|M},\sigma_{|M}]$ is event-representable. If additionally $\mathcal{L}_{|M} \subseteq \mathcal{R}_{01}(M)$ then this is equivalent to having the 3-restrictions of $\mathcal{L}_{|M}$ bin-representable, $\mathcal{L}^{\circ} \cap \mathcal{R}(M)$ solvable and $\sigma_{|M}$ adjusted to $\mathcal{L}_{|M}$.*

PROOF. The second assertion follows by Theorem 1. The forward implication of the first one is trivial. Let $\mathcal{L}$ be regular, σ adjusted to $\mathcal{L}$ and $[\mathcal{L}_{|M},\sigma_{|M}]$ be event-representable by A_i, $i \in M$. Let $A_\ell = \emptyset$ for $\ell \in \lambda(\mathcal{L})$.

The block of $\|_{\mathcal{L}}$ containing $i \in M$ can be uniquely decomposed to a disjoint union $K \cup L$ such that $i \in K$, $\sigma_{k\ell} = -1$ for $k \in K$ and $\ell \in L$, and $\sigma_{k\ell} = 1$ otherwise. In fact, by regularity (19) holds and all $(jk|\ell) \in \mathcal{R}_1(N)$ with j, k, ℓ in the block belong to $\mathcal{L}$. It suffices to assign k to $K \setminus i$ if and only if $\sigma_{ik} = 1$, since σ is adjusted to $\mathcal{L}$. Having a decomposed block $K \cup L$ of $\|_{\mathcal{L}}$, let $A_k = A_i$ for $k \in K$, and $A_\ell = \bar{A}_i$ for $\ell \in L$.

It remains to prove that A_i, $i \in N$, event-represent $[\mathcal{L},\sigma]$. Let $\mathcal{K}$ denote the pattern of conditional independences and τ the pattern of signs among the events. The sets $\mathcal{L}$ and $\mathcal{K}$ are regular, $\lambda(\mathcal{L}) = \lambda(\mathcal{K})$, $\|_{\mathcal{L}} = \|_{\mathcal{K}}$ and the restrictions of $\mathcal{L}$ and $\mathcal{K}$ to the crossing M coincide, by construction. Lemma 5 implies that $\mathcal{L} = \mathcal{K}$ is bin-representable by the events.

The values of σ and τ coincide at ij for ij contained in any block of $\|_{\mathcal{L}}$, by the construction of the events. By (15) with $K = \emptyset$, $\sigma_{ij} = 0$ if i or j is a loop of $\mathcal{L}$. For $i, j \in N \setminus \lambda(\mathcal{L})$ not in the same block of $\|_{\mathcal{L}}$, there are $k, \ell \in M$ different such that $i\|_{\mathcal{L}}k$ and $j\|_{\mathcal{L}}\ell$. If $i \neq k$ and $j \neq \ell$ then $\sigma_{ij} = \sigma_{ik}\sigma_{jk} = \sigma_{ik}\sigma_{j\ell}\sigma_{k\ell}$, using that σ is adopted to M and $(ij|k) \in \mathcal{L}$ and $(jk|\ell) \in \mathcal{L}$ by (19). If $i \neq k$ and $j = \ell$ then $\sigma_{ij} = \sigma_{ik}\sigma_{k\ell}$ using that σ is adopted to M and $(i\ell|k) \in \mathcal{L}$. The remaining case is symmetric. The same argumentation goes through with σ replaced by τ as τ is adjusted to $\mathcal{L} = \mathcal{K}$. Hence, $\sigma = \tau$, and $[\mathcal{L},\sigma]$ is event-representable. $\square$

EXAMPLE 2. Over $N = 1234$, let $\mathcal{L}$ be the union

$$\big\{ (12|\emptyset), (13|\emptyset), (23|\emptyset), (14|\emptyset), (24|\emptyset) \big\}$$
$$\cup \big\{ (ij|K)\colon 3 \in ij,\ 4 \in K \text{ or} 4 \in ij,\ 3 \in K \big\}$$

Then, $\lambda(\mathcal{L}) = \emptyset$, $3\|_{\mathcal{L}}4$ and $\mathcal{L}$ is regular. Let $\sigma_{ij} = 0$ up to $\sigma_{34} \neq 0$. Then, σ is adjusted to $\mathcal{L}$. For the crossing $M = 123$, the restriction $\mathcal{L}_{|M} = \{(12|\emptyset), (13|\emptyset), (23|\emptyset)\}$ is bin^{+}-representable. By Theorem 2, $[\mathcal{L},\sigma]$ is event-representable.

Theorem 2, however, does not cover some desirable cases. For example, it does not allow to recognize whether $\mathcal{L} = \{(1|23), (2|13), (3|12)\}$ over $N = 123$ is bin-representable.

5. Discussion

The combination of patterns of conditional independences with the signs of covariances has not been studied before, up to the special case of conjunctive forks that is discussed separately in the following section. One could speak about a decoration of the patterns of conditional independence.

The question which $\mathcal{L} \subseteq \mathcal{T}(N)$ is p-representable covers majority of approaches to patterns of the stochastic conditional independence [14]. Even when N has four elements a solution is far from trivial [9], and the bin- or bin$^+$-representability remain open [16]. For patterns among Gaussian variables see [5,7,15].

The definition of p-representability adopted in this work includes functional dependences. A frequently studied, but different, question is what are intersections of the patterns of conditional independences $\mathcal{L} \subseteq \mathcal{T}(N)$ with $\mathcal{R}(N)$. For example, when $\mathbf{1}_{A_1} = \mathbf{1}_{A_2} = \mathbf{1}_{A_3}$ are nonconstant the pattern is $\{(ij|K) \in \mathcal{T}(N)\colon K \neq \emptyset\}$. It intersects $\mathcal{R}(N)$ in $\{(12|3), (13|2), (23|1)\}$ which is not bin-representable. This question is not treated here.

The following trivial consequence of Theorem 1 is surely well-known. For every sign mapping $\sigma\colon \binom{N}{2} \mapsto \{-1, 0, 1\}$ there exist events A_i, $i \in N$, such that σ_{ij} equals the sign of the covariance between A_i and A_j, $ij \in \binom{N}{2}$. In fact, if $\mathcal{L} \subseteq \mathcal{R}_0(N)$ is constructed from σ by (2) then the 3-restrictions of $\mathcal{L}$ are bin-representable, $\mathcal{L}^\circ = \emptyset$ is solvable and (3) is void. Thus, σ is adjusted to $\mathcal{L}$. By Theorem 1, $[\mathcal{L}, \sigma]$ is event$^+$-representable whence σ is sign$^+$-representable.

As another consequence of Theorem 1, any $\mathcal{L} \subseteq \mathcal{R}_0(N)$ is bin$^+$-representable. For the general result on a binary representability of unconditional independence structures with the positivity assumption see [8, Theorem 2].

Even the following two consequences of Theorem 1 are new.

Corollary 1. *A set $\mathcal{L} \subseteq \mathcal{R}_{01}(N)$ is bin$^+$-representable if and only if it is bin-representable which is equivalent to having the 3-restrictions of $\mathcal{L}$ bin-representable and $\mathcal{L}^\circ$ solvable.*

Proof. Let the 3-restrictions of $\mathcal{L}$ be bin-representable and $\mathcal{L}^\circ$ be solvable. A sign mapping σ is defined by $\sigma_{ij} = 0$ when $(ij|\emptyset) \in \mathcal{L} \cap \mathcal{R}_0(N)$ and by $\sigma_{ij} = 1$ otherwise. Then, σ is adjusted to $\mathcal{L}$. In fact, (2) holds by construction, and (3) holds as $1 = 1$ or $0 = 0$, due to the assumption on the 3-restrictions. It suffices to apply Theorem 1 to $[\mathcal{L}, \sigma]$. $\square$

Corollary 2. *A set $\mathcal{L} \subseteq \mathcal{R}_1(N)$ is bin-representable if and only if $\mathcal{L}$ is solvable.*

Proof. This follows from Corollary 1 as the 3-restrictions of $\mathcal{L}$ are empty or isomorphic to $\{(12|3)\}$, thus $\mathcal{L} = \mathcal{L}^\circ$. $\square$

Conditions for p-representability of $\mathcal{L} \subseteq \mathcal{R}_{01}(N)$, or even $\mathcal{L} \subseteq \mathcal{R}_1(N)$, seem to be elusive.

6. Patterns of conjunctive forks

The notion of conjunctive fork goes back to [13]. A historical definition and references are in [1]. In a probability space $(\Omega, \mathcal{S}, P)$, events $A, B, C \in \mathcal{S}$ create a *conjunctive fork* $(A, B \vdots C)_P$ if $\mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_B | \mathbf{1}_C$, $\langle A, C\rangle_P > 0$ and $\langle B, C\rangle_P > 0$. By Lemma 2, $\langle A, B\rangle_P > 0$. Events A_i, $i \in N$, induce the pattern of forks $\{(i, k, j) \in N^3 \colon (A_i, A_j \vdots A_k)_P\}$. A ternary relation $\mathfrak{r} \subseteq N^3$ is *fork-representable* if it coincides with such a pattern. In a *fork*$^+$*-representability* the positivity is required as well. In this section, both kinds of representability are exposed as consequences of Theorem 1.

Lemma 6. *Any fork-representable relation* $\mathfrak{r} \subseteq N^3$ *satisfies*

(20) $$\big[(i, j, k) \in \mathfrak{r} \text{ and } (i, k, j) \in \mathfrak{r}\big] \;\Rightarrow\; (j, k, j) \in \mathfrak{r}\,,$$

(21) $$(i, j, k) \in \mathfrak{r} \;\Rightarrow\; \big\{(i, j, j), (j, k, k), (k, i, i)\big\} \subseteq \mathfrak{r}\,,$$

(22) $$(i, j, i) \in \mathfrak{r} \;\Rightarrow\; (j, i, j) \in \mathfrak{r}\,,$$

(23) $$\big[(i, j, i) \in \mathfrak{r} \text{ and } (j, k, j) \in \mathfrak{r}\big] \;\Rightarrow\; (i, k, i) \in \mathfrak{r}\,.$$

Proof. By (9) and positivity of correlations, if $(A,C \vdots B)_P$ and $(A,B \vdots C)_P$ then $B \overset{P}{=} C$ whence $(B, B \vdots C)_P$. Then, (20) holds.

The fork $(A, B \vdots B)_P$ occurs if and only if $\langle A, B\rangle_P > 0$ which implies (21).

The fork $(A, A \vdots B)_P$ is equivalent to $A \overset{P}{=} B$ not P-trivial. By symmetry, $(A, A \vdots B)_P$ if and only if $(B, B \vdots A)_P$ which implies (22). Likewise, (23) holds. $\square$

A relation $\mathfrak{r} \subseteq N^3$ is a *forkness* [1] if it satisfies (20)–(23) and the symmetry condition

(24) $$(i, k, j) \in \mathfrak{r} \;\;\Leftrightarrow\;\; (j, k, i) \in \mathfrak{r}\,.$$

For $\mathfrak{r} \subseteq N^3$ let $\mathcal{K}_\mathfrak{r} \triangleq \{(ij|k) \in \mathcal{R}_1(N) \colon \{(i, k, j), (j, k, i)\} \subseteq \mathfrak{r}\}$. Theorem 1 has the following consequence.

Corollary 3. *A relation* $\mathfrak{r}$ *is fork*$^+$*-representable if and only if it is a forkness,* $(i, i, i) \in \mathfrak{r}$, $i \in N$, *and* $\mathcal{K}_\mathfrak{r}$ *is solvable.*

Proof. Let $\mathfrak{r}$ be fork$^+$-represented by events. By Lemma 6, it is a forkness. The positivity assumption implies that the representing events are not P-trivial which entails $(i,i,i) \in \mathfrak{r}$, $i \in N$. If $(ij|k) \in \mathcal{K}_\mathfrak{r}$ then $(i,k,j) \in \mathfrak{r}$, and then (i,j,k) and (j,i,k) are not in $\mathfrak{r}$, using (9), positivity of correlations and positivity of the distribution of the events. Hence $(ij|K)$ belongs to $\mathcal{K}_\mathfrak{r}^\circ$. Thus, $\mathcal{K}_\mathfrak{r} = \mathcal{K}_\mathfrak{r}^\circ$. Since $\mathcal{K}_\mathfrak{r}$ is contained in the pattern $\mathcal{L}$ of conditional independences among indicators and $\mathcal{L}^\circ$ solvable by Lemma 4, $\mathcal{K}_\mathfrak{r}$ is solvable.

Let $\mathfrak{r}$ be a forkness that contains $\{(i,i,i) \colon i \in N\}$ and $\mathcal{K}_\mathfrak{r}$ be solvable. The latter implies $\mathcal{K}_\mathfrak{r} = \mathcal{K}_\mathfrak{r}^\circ$. Any 3-restriction of $\mathcal{K}_\mathfrak{r}$ is empty or isomorphic to $\{(12|3)\}$. Let σ be the sign mapping with $\sigma_{ij} = 1$ for $ij \in E_{\mathcal{K}_\mathfrak{r}}$. Otherwise, let $\sigma_{ij} = \pm 1$ according to whether $(i,j,j) \in \mathfrak{r}$. By construction, σ is adapted to $\mathcal{K}_\mathfrak{r}$. It follows from Theorem 1 that $[\mathcal{K}_\mathfrak{r}, \sigma]$ is event$^+$-representable. Then, $\mathfrak{r}$ is fork$^+$-representable, using $\{(i,i,i) \colon i \in N\} \subseteq \mathfrak{r}$ and the construction of σ outside of $E_{\mathcal{K}_\mathfrak{r}}$. $\square$

For $\mathfrak{r} \subseteq N^3$ let $\lambda(\mathfrak{r}) = \{i \in N \colon (i,i,i) \notin \mathfrak{r}\}$ and $\stackrel{\mathfrak{r}}{\sim}$ be the binary relation on $N \setminus \lambda(\mathcal{L})$ with $i \stackrel{\mathfrak{r}}{\sim} j$ if and only if $(i,j,i) \in \mathfrak{r}$. The relation $\stackrel{\mathfrak{r}}{\sim}$ is reflexive. In a forkness, it is symmetric by (22), and transitive by (23). Thus, it is an equivalence relation. A forkness is *regular* if

$$(i,j,k) \in \mathfrak{r}, \quad i \stackrel{\mathfrak{r}}{\sim} i', \quad j \stackrel{\mathfrak{r}}{\sim} j', \quad k \stackrel{\mathfrak{r}}{\sim} k' \ \Rightarrow \ (i',j',k') \in \mathfrak{r}\,.$$

LEMMA 7. *A regular forkness $\mathfrak{r} \subseteq N^3$ is fork-representable if and only if a restriction of $\mathfrak{r}$ to M^3 is fork-representable for any set $M \subseteq N \setminus \lambda(\mathfrak{r})$ that contains exactly one element of each block of $\stackrel{\mathfrak{r}}{\sim}$.*

PROOF. One implication is trivial. For the set M, let a restriction of $\mathfrak{r}$ to M^3 be fork-representable by A_i, $i \in M$. Let $A_\ell = \emptyset$, $\ell \in \lambda(\mathfrak{r})$. Having a block K of $\stackrel{\mathfrak{r}}{\sim}$ let $A_k = A_i$ for $k \in K$ and the unique $i \in M$. The forkness $\mathfrak{r}$ is regular and so is the pattern of forks within the events. Their restrictions to M^3 coincide by construction. Then, their restrictions to $(N \setminus \lambda(\mathfrak{r}))^3$ coincide by regularity. Hence, $\mathfrak{r}$ is fork-representable. $\square$

The main result of [1, Theorem 1] is a consequence of Corollary 3 of Theorem 1 and auxiliary Lemmas 4 and 7.

COROLLARY 4. *A ternary relation $\mathfrak{r}$ on N is fork-representable if and only if it is a regular forkness and for any set $M \subseteq N \setminus \lambda(\mathfrak{r})$ that contains exactly one element of each block of $\stackrel{\mathfrak{r}}{\sim}$ the restriction $\mathcal{K}_\mathfrak{r} \cap \mathcal{R}_1(M)$ is solvable.*

PROOF. The pattern $\mathfrak{r}$ of conjunctive forks among events A_i, $i \in N$, is a forkness by Lemma 6. It is regular because $i \stackrel{\mathfrak{r}}{\sim} j$ means that $A_i \stackrel{P}{=} A_j$ is not P-trivial. Hence, if $(ij|k) \in \mathcal{K}_\mathfrak{r} \cap \mathcal{R}_1(M)$ then $(ik|j)$ does not belong to this restriction. Thus, $\mathcal{K}_\mathfrak{r} \cap \mathcal{R}_1(M)$ is contained in $\mathcal{L}^\circ$ where $\mathcal{L}$ is the pattern of conditional independences among the indicators of the events. It is solvable since $\mathcal{L}^\circ$ is solvable by Lemma 4.

For the set M as above with $\mathcal{K}_{\mathfrak{r}} \cap \mathcal{R}_1(M)$ solvable, the restriction of $\mathfrak{r}$ to M^3 is fork$^+$-representable by Corollary 3. It follows from Lemma 7 that $\mathfrak{r}$ is fork-representable. $\square$

The last part of this section discusses fork- and fork$^+$- representability in the restricted framework of the traditional notion of betweenness [12, p. 96]. They are equivalent to solvability.

A relation $\mathfrak{r} \subseteq N^3$ satisfying the symmetry (24) and

$$\big[\,(i,k,j) \in \mathfrak{r} \text{ and } (i,j,k) \in \mathfrak{r}\,\big] \;\Leftrightarrow\; j = k \tag{25}$$

is called *betweenness*. The implication $\Leftarrow$ in (25) says that $(i,j,j) \in \mathfrak{r}$ for every $i,j \in N$. In particular, $\mathfrak{s}^* = \bigcup_{i,j\in N}\{(i,j,j),(j,j,i)\}$ is the inclusion smallest betweenness. No betweenness contains a triple (i,j,i) with $i \neq j$, by $\Rightarrow$ in (25).

LEMMA 8. *The mapping $\mathfrak{r} \mapsto \mathcal{K}_{\mathfrak{r}}$ is a bijection between the family of betweennesses $\mathfrak{r}$ on N and the family of sets $\mathcal{L} \subseteq \mathcal{R}_1(N)$ satisfying $\mathcal{L} = \mathcal{L}^\circ$.*

PROOF. If $\mathfrak{r}$ is a betweenness then $\mathcal{K}_{\mathfrak{r}} = \mathcal{K}_{\mathfrak{r}}^\circ$, by $\Rightarrow$ in (25). Given $\mathcal{L} \subseteq \mathcal{R}_1(N)$, let

$$\mathfrak{s}_{\mathcal{L}} \triangleq \mathfrak{s}^* \cup \bigcup_{(ij|k)\in\mathcal{L}} \{(i,k,j),(j,k,i)\}\,.$$

If $\mathcal{L} = \mathcal{L}^\circ$ then (24)–(25) hold for $\mathfrak{s}_{\mathcal{L}}$, thus $\mathfrak{s}_{\mathcal{L}}$ is a betweenness. Since $\mathfrak{r} = \mathfrak{s}_{\mathcal{K}_{\mathfrak{r}}}$ for a betweenness $\mathfrak{r}$, and $\mathcal{L} = \mathcal{K}_{\mathfrak{s}_{\mathcal{L}}}$ once $\mathcal{L} = \mathcal{L}^\circ$, the mapping $\mathfrak{r} \mapsto \mathcal{K}_{\mathfrak{r}}$ is a bijection and $\mathcal{L} \mapsto \mathfrak{s}_{\mathcal{L}}$ is its inverse. Thus, $\overset{\mathfrak{r}}{\sim}$ is the equality relation. $\square$

LEMMA 9. *Every betweenness is a forkness.*

PROOF. Let $\mathfrak{r}$ be a betweenness. The assumption in (20) is valid only if two of i, j, k coincide. If $i = j$ then $(i,k,i) \in \mathfrak{r}$ implies $i = k$. By symmetry, (20) holds with $j = k$. The implication (21) says that $\mathfrak{r}$ contains $\mathfrak{s}^*$, and (22) is void unless $i = j$. Also, (23) is void unless i,j,k coincide. $\square$

Corollaries 3 and 4 together with Lemma 9 imply the following.

COROLLARY 5. *A betweenness $\mathfrak{r}$ is fork$^+$-representable if and only if it is fork-representable which is equivalent to the solvability of $\mathcal{K}_{\mathfrak{r}}$.*

References

[1] V. Chvátal, F. Matúš and Y. Zwólš, Patterns of conjunctive forks, arXiv:1608.03949 [math.PR] (2016).

[2] A. P. Dawid, Conditional independence in statistical theory (with discussion), *J. R. Statist. Soc. B*, **41** (1979), 1–31.

[3] A. Deitmar, *A First Course in Harmonic Analysis*, Springer-Verlag (New York, 2015).

[4] M. Drton, B. Sturmfels and S. Sullivant, *Lectures on Algebraic Statistics*, Oberwolfach Seminars, Birkhäuser Applied Probability and Statistics, Birkhäuser (2009).

[5] M. Drton and H. Xiao, Smoothness of Gaussian conditional independence models, *Contemp. Math.*, **516** (2010), 155–177.

[6] P. F. Lazarsfeld and N. W. Henry, *Latent Structure Analysis*, Houghton Mifflin Company (Boston, 1968).

[7] R. Lněnička and F. Matúš, On Gaussian conditional independence structures, *Kybernetika*, **43** (2007), 327–342.

[8] F. Matúš, Stochastic independence, algebraic independence and abstract connectedness, *Theoret. Comput. Sci.*, **134** (1994), 455–471.

[9] F. Matúš, Conditional independences among four random variables III: final conclusion, *Combin. Probab. Comput.*, **8** (1999) 269–276.

[10] O. Kallenberg, *Foundations of Modern Probability*, Springer-Verlag (New York, 1997).

[11] S. Lauritzen and K. Sadeghi, Unifying Markov properties for graphical models. arXiv: 1608. 05810v3 [math.ST] (2017).

[12] E. Pitcher and M.F. Smiley, Transitivities of betweenness, *Trans. Amer. Math. Soc.*, **52** (1942), 95–114.

[13] H. Reichenbach, *The Direction of Time*, University of California Press (Berkeley, Los Angeles, CA, 1956).

[14] M. Studený, *Probabilistic Conditional Independence Structures*, Springer-Verlag (New York, 2005).

[15] S. Sullivant, Gaussian conditional independence has no finite complete axiom system, *J. Pure Appl. Algebra*, **213** (2009), 1502–1506.

[16] P. Šimeček, Classes of Gaussian, discrete and binary representable independence models have no finite characterization, in: *Prague Stochastics* (M. Hušková and M. Janžura, eds.), Matfyzpress, Charles University (Prague, 2006), pp. 622–632.