On finite exchangeability and conditional independence

Kayvan Sadeghi

Department of Statistical Science, University College London, Gower Street, London, WC1E 6BT, United Kingdom

Abstract: We study the independence structure of finitely exchangeable distributions over random vectors and random networks. In particular, we provide necessary and sufficient conditions for an exchangeable vector so that its elements are completely independent or completely dependent. We also provide a sufficient condition for an exchangeable vector so that its elements are marginally independent. We then generalize these results and conditions for exchangeable random networks. In this case, it is demonstrated that the situation is more complex. We show that the independence structure of exchangeable random networks lies in one of six regimes that are two-fold dual to one another, represented by undirected and bidirected independence graphs in graphical model sense with graphs that are complement of each other. In addition, under certain additional assumptions, we provide necessary and sufficient conditions for the exchangeable network distributions to be faithful to each of these graphs.

Keywords and phrases: Conditional independence, exchangeability, faithfulness, random networks.

1. Introduction

The concept of exchangeability has been a natural and convenient assumption to impose in probability theory and for simplifying statistical models. As defined originally for random sequences, it states that any order of a finite number of samples is equally likely. The concept was later generalized for binary random arrays [1], and consequently for random networks in statistical network analysis. In this context, exchangeability is translated into invariance under relabeling of the nodes of the network, whereby isomorphic graphs have the same probabilities; see, e.g., [14].

Exchangeability is closely related to the concept of independent and identically distributed random variables. It is an immediate consequence of the definition that independent and identically distributed random variables are exchangeable, but the converse is not true. For infinite sequences, the converse is established by the well-known deFinetti's Theorem [5], which implies that in any infinite sequence of exchangeable random variables, the random variables are conditionally independent and identically-distributed given the underlying distributional form. Other versions of deFinetti's Theorem exist for the generalized definitions of exchangeability [27, 7].

However, for finitely exchangeable random sequences (vectors) and arrays (matrices), the converse does not hold, and the available results basically provide approximations of the infinite case; see, e.g., [6, 21, 15]. In this paper, we utilize a completely different approach to study the relationship between finite exchangeability and (conditional) independence. We employ the theory of graphical models (see e.g. [13]) in order to provide the independence structure of exchangeable distributions.

In particular, we exploit the necessary and sufficient conditions, provided in [26], for faithfulness of probability distributions and graphs, which determine when the conditional independence structure of the distribution is exactly the same as that of a graph in graphical model sense. Thus, we, in practice, work on the induced independence model of an exchangeable probability distribution rather than the distribution itself. The specialization to finitely exchangeable random vectors leads to necessary and sufficient conditions for an exchangeable vector to be completely independent or completely dependent, meaning that two elements of the vector are conditionally independent or dependent, respectively, given any subset of the remaining elements of the vector. These conditions are namely intersection and composition properties [22, 28]. We also use the results to provide a sufficient condition for an exchangeable vector to be marginally independent.

For random networks, we follow a similar procedure. It was shown in [14] that if a distribution over an exchangeable random network could be faithful to a graph then the skeleton of the graph is only one of the four possible types: the empty graph, the complete graph, and the, so-called incidence graph, and its complement (see Proposition 8). A question is then which graphs that emerge from these skeleta are faithful to the exchangeable distribution. We show that, other than complete independence or dependence, there are four other independence structures that can arise for exchangeable networks. These are faithful to a pair of dual graphs with incidence graph skeleton, and a pair of dual graphs with the complement of the incidence graph skeleton. We then provide assumptions under which intersection and composition properties are necessary and sufficient for the exchangeable random networks to be faithful to each of these six cases.

The structure of the paper is as follows: In the next section, we provide some definitions and known results needed for this paper from graph theory, random networks, and graphical models. In Section 3, we provide the results for exchangeable random vectors. In Section 4, we provide the results for exchangeable random networks by starting with providing the aforementioned possible cases in Theorem 2, and then introducing them in separate subsections. We end with a short discussion on these results in Section 5. We will provide technical definitions and results needed for the proof of Theorem 2, and its proof, in Appendix A.

2. Definitions and preliminary results

2.1. Graph-theoretic concepts

A *(labeled) graph* is an ordered pair $G = (V, E)$ consisting of a *vertex* set V, which is non-empty and finite, an *edge* set E, and a relation that with each edge associates two vertices, called its *endpoints*. We omit the term labeled in this paper since the context does not give rise to ambiguity. When vertices u and v are the endpoints of an edge, these are *adjacent* and we write $u \sim v$; we denote the corresponding edge as uv.

In this paper, we will restrict our attention to *simple* graphs, i.e. graphs without loops (this assumption means that the endpoints of each edge are distinct) or multiple edges (each pair of vertices are the endpoints of at most one edge). Further, three *types* of edges, denoted by *arrows*, *arcs* (solid lines with two-headed arrows) and *lines* (solid lines without arrowheads) have been used in the literature of graphical models. Arrows can be represented by ordered pairs of vertices, while arcs and lines by 2-subsets of the vertex set. However, for our purpose, except for Section 4.2, we will distinguish only lines and arcs, which respectively form *undirected* and *bidirected* graphs.

The graphs $F = (V_F, E_F)$ and $G = (V_G, E_G)$ are considered equal if and only if $(V_F, E_F) = (V_G, E_G)$.

A *subgraph* of a graph $G = (V_G, E_G)$ is graph $F = (V_F, E_F)$ such that $V_F \subseteq V_G$ and $E_F \subseteq E_G$ and the assignment of endpoints to edges in F is the same as in G.

The *line graph* $L(G)$ of a graph $G = (V, E)$ is the intersection graph of the edge set E, i.e. its vertex set is E and $e_1 \sim e_2$ if and only if e_1 and e_2 have a common endpoint [32, p. 168]. We will in particular be interested in the line graph of a complete graph, which we will refer to as the *incidence graph*. Figure 1 displays the incidence graph for $V = \{1, 2, 3, 4\}$. We denote by $L_-(n)$
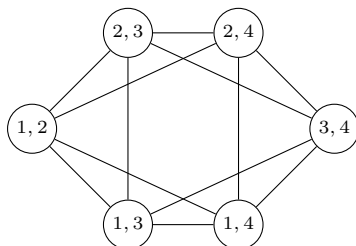

Fig 1. *The incidence graph for $V = \{1, 2, 3, 4\}$.*

and $L_\leftrightarrow(n)$, the undirected incidence graph and the bidirected incidence graph for n nodes, respectively; and by $L_-^c(n)$ and $L_\leftrightarrow^c(n)$, the undirected complement of the incidence graph and the bidirected complement of the incidence graph for n nodes, respectively, where the *complement* of graph G refers to a graph with the same node set as G, but with the edge set that is the complement set of the edge set of G.

The *skeleton* of a graph is the undirected graph where all arrowheads are removed from the graphs, i.e., all edges are replaced by lines. We denote the skeleton of a graph G by $\mathrm{sk}(G)$.

A *walk* ω is a list $\omega = \langle i_0, e_1, i_1, \ldots, e_n, i_n \rangle$ of vertices and edges such that for $1 \leq m \leq n$, the edge e_m has endpoints i_{m-1} and i_m. When indicating a specific walk, we may skip the edges, and only write the nodes, when the walk we are considering is clear from the context. A *path* is a walk with no repeated nodes.

2.2. Random networks

Given a finite node set $\mathcal{N}$ — representing *individuals* or *actors* in a given population of interest — we define a *random network* over $\mathcal{N}$ to be a collection $X = (X_d, d \in \mathcal{D}(\mathcal{N}))$ of binary random variables taking values 0 and 1 indexed by a set $\mathcal{D}(\mathcal{N})$, which is a collection of unordered pairs ij of nodes in $\mathcal{N}$. The binary random variables X_d are called *dyads*, and nodes i and j are said to have a *tie* if the random variable X_{ij} takes the value 1, and no tie otherwise. Thus, a random network is a random variable taking value in $\{0,1\}^{\binom{\mathcal{N}}{2}}$ and can, therefore, be seen as a random *simple, undirected* graph with node set $\mathcal{N}$, whereby the ties form the random edges of the graphs.

We use the terms network, node, and tie rather than graph, vertex, and edge to differentiate from the terminology used in the graphical model sense. Indeed, as we shall discuss graphical models for networks, we will also consider each dyad d as a vertex in a graph $G = (\mathcal{D}, E)$ representing the dependence structure of the random variables associated with the dyads, with the edge set of such graph representing Markov properties of the distribution of X.

2.3. Probabilistic independence models and their properties

An *independence model* $\mathcal{J}$ over a finite set V is a set of triples $\langle X, Y \mid Z \rangle$ (called *independence statements*), where X, Y, and Z are disjoint subsets of V; Z may be empty, but $\langle \varnothing, Y \mid Z \rangle$ and $\langle X, \varnothing \mid Z \rangle$ are always included in $\mathcal{J}$. The independence statement $\langle X, Y \mid Z \rangle$ is read as "X is independent of Y given Z". Independence models may in general have a probabilistic interpretation, but not necessarily. Similarly, not all independence models can be easily represented by graphs. For further discussion on general independence models, see [28].

In order to define probabilistic independence models, consider a set V and a collection of random variables $\{X_\alpha\}_{\alpha \in V}$ with state spaces $\mathcal{X}_\alpha, \alpha \in V$ and joint distribution P. We let $X_A = \{X_v\}_{v \in A}$ etc. for each subset A of V. For disjoint subsets A, B, and C of V we use a short notation $A \perp\!\!\!\perp B \mid C$ to denote that X_A is *conditionally independent of* X_B *given* X_C [4, 13], i.e. that for any measurable $\Omega \subseteq \mathcal{X}_A$ and P-almost all x_B and x_C,

$$P(X_A \in \Omega \mid X_B = x_B, X_C = x_C) = P(X_A \in \Omega \mid X_C = x_C).$$

We can now induce an independence model $\mathcal{J}(P)$ by letting

$$\langle A, B \mid C \rangle \in \mathcal{J}(P) \text{ if and only if } A \perp\!\!\!\perp B \mid C \text{ w.r.t. } P.$$

Similarly we use the notation $A \not\perp\!\!\!\perp B \,|\, C$ for $\langle A, B \,|\, C \rangle \notin \mathcal{J}(P)$.

We say that non-empty A and B are *completely independent* if, for every $C \subseteq V \setminus (A \cup B)$, $A \perp\!\!\!\perp B \,|\, C$. Similarly, we say that A and B are *completely dependent* if, for any $C \subseteq V \setminus (A \cup B)$, $A \not\perp\!\!\!\perp B \,|\, C$.

If A, B, or C has only one member $\{u\}$, $\{v\}$, or $\{w\}$, for better readability, we write $u \perp\!\!\!\perp v \,|\, w$. We also write $A \perp\!\!\!\perp B$ when $C = \varnothing$, which denotes the *marginal independence* of A and B.

A probabilistic independence model $\mathcal{J}(P)$ over a set V is always a *semi-graphoid* [22], i.e., it satisfies the four following properties for disjoint subsets A, B, C, and D of V:

1. $A \perp\!\!\!\perp B \,|\, C$ if and only if $B \perp\!\!\!\perp A \,|\, C$ (*symmetry*);
2. if $A \perp\!\!\!\perp B \cup D \,|\, C$ then $A \perp\!\!\!\perp B \,|\, C$ and $A \perp\!\!\!\perp D \,|\, C$ (*decomposition*);
3. if $A \perp\!\!\!\perp B \cup D \,|\, C$ then $A \perp\!\!\!\perp B \,|\, C \cup D$ and $A \perp\!\!\!\perp D \,|\, C \cup B$ (*weak union*);
4. if $A \perp\!\!\!\perp B \,|\, C \cup D$ and $A \perp\!\!\!\perp D \,|\, C$ then $A \perp\!\!\!\perp B \cup D \,|\, C$ (*contraction*).

Notice that the reverse implication of contraction clearly holds by decomposition and weak union. A semi-graphoid for which the reverse implication of the weak union property holds is said to be a *graphoid*; that is, it also satisfies

5. if $A \perp\!\!\!\perp B \,|\, C \cup D$ and $A \perp\!\!\!\perp D \,|\, C \cup B$ then $A \perp\!\!\!\perp B \cup D \,|\, C$ (*intersection*).

Furthermore, a graphoid or semi-graphoid for which the reverse implication of the decomposition property holds is said to be *compositional*, that is, it also satisfies

6. if $A \perp\!\!\!\perp B \,|\, C$ and $A \perp\!\!\!\perp D \,|\, C$ then $A \perp\!\!\!\perp B \cup D \,|\, C$ (*composition*).

If, for example, P has strictly positive density, the induced probabilistic independence model is always a graphoid; see e.g. Proposition 3.1 in [13]. See also [24] for a necessary and sufficient condition for P in order for the intersection property to hold. If the distribution P is a regular multivariate Gaussian distribution, $\mathcal{J}(P)$ is a compositional graphoid; e.g. see [28]. Probabilistic independence models with positive densities are not in general compositional; this only holds for special types of multivariate distributions such as, for example, Gaussian distributions and the symmetric binary distributions used in [31].

Another important property that is not necessarily satisfied by probabilistic independence models is *singleton-transitivity* (also called *weak transitivity* in [22], where it is shown that for Gaussian and binary distributions P, $\mathcal{J}(P)$ always satisfies it). For u, v, and w, single elements in V,

7. if $u \perp\!\!\!\perp v \,|\, C$ and $u \perp\!\!\!\perp v \,|\, C \cup \{w\}$ then $u \perp\!\!\!\perp w \,|\, C$ or $v \perp\!\!\!\perp w \,|\, C$ (*singleton-transitivity*).

In addition, we have the two following properties:

8. if $u \perp\!\!\!\perp v \,|\, C$ then $u \perp\!\!\!\perp v \,|\, C \cup \{w\}$ for every $w \in V \setminus \{u, v\}$ (*upward-stability*);
9. if $u \perp\!\!\!\perp v \,|\, C$ then $u \perp\!\!\!\perp v \,|\, C \setminus \{w\}$ for every $w \in V \setminus \{u, v\}$ (*downward-stability*).

Henceforth, instead of saying that "$\mathcal{J}(P)$ satisfies these properties", we simply say that "P satisfies these properties". First we provide the following well-known result [26]:

Lemma 1. *For a probability distribution P the following holds:*

1. *If P satisfies upward-stability then P satisfies composition.*
2. *If P satisfies downward-stability then P satisfies intersection.*

2.4. Exchangeability for random vectors and networks

A probability distribution P over a finite vector $(X_1, X_2, X_3, \ldots, X_n)$ of random variables with the same shared sample space is *(finitely) exchangeable* if for any permutation $\pi \in S(n)$ of the indices $1, 2, 3, \ldots, n$, the probability distribution of the permuted vector $(X_{\pi(1)}, X_{\pi(2)}, X_\pi(3), \ldots, X_{\pi(n)})$ is the same as P; see [1]. We shall for brevity say that the sequence X is exchangeable in the meaning that its distribution is.

We are also concerned with probability distributions on networks that are finitely exchangeable. A distribution P of a random matrix $X = (X_{ij})_{i,j \in \mathcal{N}}$ over a finite node set $\mathcal{N}$ with the same shared sample space is said to be *(finitely) weakly exchangeable* [27, 10] if for all permutations $\pi \in S(\mathcal{N})$ we have that

$$\mathbb{P}\{(X_{ij} = x_{ij})_{i,j \in \mathcal{N}}\} = \mathbb{P}\{(X_{ij} = x_{\pi(i)\pi(j)})_{i,j \in \mathcal{N}}\}. \tag{2.1}$$

If the matrix X is symmetric — i.e. $X_{ij} = X_{ji}$, we say it is *symmetric weakly exchangeable.* Again, we shall for brevity say that X is weakly or symmetric weakly exchangeable in the meaning that its distribution is.

A symmetric binary array with zero diagonal can be interpreted as a matrix of ties (the *adjacency matrix*) of a random network and, thus, the above concepts can be translated into networks. A random network is *exchangeable* if its adjacency matrix is symmetric weakly exchangeable. Then it is easy to observe that a random network is exchangeable if and only if its distribution is invariant under relabeling of the nodes of the network.

2.5. Undirected and bidirected graphical models

Graphical models [see, e.g. 13] are statistical models expressing conditional independence statements among a collection of random variables $X_V = (X_v, v \in V)$ indexed by a finite set V. A graphical model is determined by a graph $G = (V, E)$ over the indexing set V, and the edge set E (which may include edges of undirected, directed or bidirected type) encodes conditional independence relations among the variables, or *Markov properties.*

We say that C *separates* A and B in an undirected graph G, denoted by $A \perp_u B \mid C$, if every path between A and B has a vertex in C, that is there is no path between A and B outside C. For a bidirected graph G, we say that C *separates* A and B, denoted by $A \perp_b B \mid C$, if every path between A and B has

a vertex outside $C \cup A \cup B$, that is there is no path between A and B within $A \cup B \cup C$. Note the obvious duality between this and separation for undirected graphs. We might skip the subscripts u and b in $\perp_u$ and $\perp_b$ when it is apparent from the context with which separation we are dealing.

A joint probability distribution P for X_V is *Markovian* with respect to an undirected graph [3] with the vertex set V if $A \perp_u B \mid C$ implies $A \perp\!\!\!\perp B \mid C$. P is *Markovian* with respect to a bidirected graph [2, 12] if $A \perp_b B \mid C$ implies $A \perp\!\!\!\perp B \mid C$.

For example, in the undirected graph of Figure 2(a), the global Markov property implies that $\{u, x\} \perp\!\!\!\perp w \mid v$, whereas in the bidirected graph of Figure 2(b), the global Markov property implies that $\{u, x\} \perp\!\!\!\perp w$.
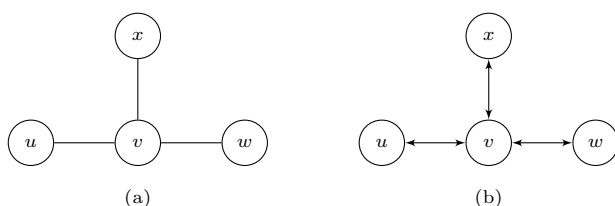

FIG 2. *(a) An undirected dependence graph. (b) A bidirected dependence graph.*

If, for P and undirected G, $A \perp_u B \mid C \iff A \perp\!\!\!\perp B \mid C$ then we say that P and G are *faithful*; Similarly, if, for P and bidirected G, $A \perp_b B \mid C \iff A \perp\!\!\!\perp B \mid C$ then P and G are *faithful*. Hence, faithfulness implies being Markovian, but not the other way around.

For a given probability distribution P, we define the *skeleton* of P, denoted by $\mathrm{sk}(P)$, to be the undirected graph with the vertex set V such that vertices u and v are not adjacent if and only if there is *some* subset C of V so that $u \perp\!\!\!\perp v \mid C$. Thus, if P is Markovian with respect to an undirected graph G then $\mathrm{sk}(P)$ would be a subgraph of G (since for every missing edge ij in G, $i \perp\!\!\!\perp j \mid V \setminus \{i, j\}$); and if P is Markovian with respect to a bidirected graph G then $\mathrm{sk}(P)$ is a subgraph of $\mathrm{sk}(G)$ (since for every missing edge ij in G, $i \perp\!\!\!\perp j$).

In general, a graph $G(P)$ is induced by P with skeleton $\mathrm{sk}(P)$. For undirected graphs, let $G_u(P) = \mathrm{sk}(P)$, whereas for bidirected graphs, let $G_b(P)$ be $\mathrm{sk}(P)$ with all edges being bidirected. We shall need the following results from [26] (where the first part was first shown in [23]):

Proposition 1. *Let P be a probability distribution defined over $\{X_\alpha\}_{\alpha \in V}$. It then holds that*

1. *P and $G_u(P)$ are faithful if and only if P satisfies intersection, singleton-transitivity, and upward-stability.*
2. *P and $G_b(P)$ are faithful if and only if P satisfies composition, singleton-transitivity, and downward-stability.*

2.6. Duality in independence models and graphs

The *dual* of an independence model $\mathcal{J}$ (defined in [20] under the name of *dual relation*) is the independence model defined by $\mathcal{J}^d = \{\langle A, B \,|\, V \setminus (A \cup B \cup C)\rangle :$ $\langle A, B \,|\, C\rangle \in \mathcal{J}\}$; see also [9]. We will need the following lemma regarding the duality of independence models:

Lemma 2. *For an independence model $\mathcal{J}$ and its dual $\mathcal{J}^d$,*

1. *$\mathcal{J}$ is semi-graphoid if and only if $\mathcal{J}^d$ is semi-graphoid;*
2. *$\mathcal{J}$ satisfies intersection if and only if $\mathcal{J}^d$ satisfies composition; and vice versa;*
3. *$\mathcal{J}$ satisfies singleton-transitivity if and only if $\mathcal{J}^d$ satisfies singleton-transitivity;*
4. *$\mathcal{J}$ satisfies upward-stability if and only if $\mathcal{J}^d$ satisfies downward-stability; and vice versa.*

Proof. 1., 2., and 3. are proven in [18], which showed that if $\mathcal{J}$ is singleton-transitive compositional graphoid, so is the dual couple of $\mathcal{J}$ (although this is proven based on the so-called *Gaussoid* formulation). An alternative proof for these, with the same formulation as in this paper, is found in [19]. (In fact the statement in the mentioned article was proven for a generalization of singleton-transitivity, called *dual decomposable transitivity*).

In order to prove 4., suppose that $\mathcal{J}$ satisfies upward-stability, and assume that $\langle i, j \,|\, C\rangle \in \mathcal{J}^d$ and $k \in C$. Therefore, $\langle i, j \,|\, V \setminus (\{i, j\} \cup C)\rangle \in \mathcal{J}$. Hence, $\langle i, j \,|\, V \setminus (\{i, j\} \cup C) \cup \{k\}\rangle \in \mathcal{J}$ because of upward-stability. Therefore $\langle i, j \,|\, C \setminus \{k\}\rangle \in \mathcal{J}^d$, which implies downward-stability of $\mathcal{J}^d$. The other direction is similar. □

As an additional statement to the lemma, it can also be shown that $\mathcal{J}$ is closed under *marginalization* if and only if $\mathcal{J}^d$ is closed under *conditioning*; and vice versa; but this result is not needed in this paper.

In addition, for separation in graphs, we will use the following lemma related to duality:

Lemma 3. *Let G_u and G_b be an undirected and a bidirected graph such that $\mathrm{sk}(G_u) = \mathrm{sk}(G_b)$. Then, the independence model $\mathcal{J}(G_b)$ induced by G_b is $\mathcal{J}^d(G_u)$ and vice versa.*

Proof. Since the separation satisfies the composition property, it is sufficient to prove the statement for singletons. Suppose that there is a connecting path ω between i and j given C in G_u. This means that no inner vertex of ω is in C; thus they are all in $V \setminus (\{i, j\} \cup C)$. Therefore, in G_b, i and j are connecting given $V \setminus (\{i, j\} \cup C)$. The other direction (where $i \not\perp j \,|\, C$ in $G_b \Rightarrow i \not\perp j \,|\, V \setminus (A \cup B \cup C)$ in G_u) is proven in a similar way. □

In fact, the second part of Proposition 1, could be implied by the first part, and vice versa, using Lemmas 2 and 3; and the second part of Lemma 1, could be implied by the first part, and vice versa, using Lemma 2.

3. Results for vector exchangeability

We shall study the relationship between vector exchangeability and conditional independence by using the definitions and results in the previous section. In the entire section, we assume that P is a probability distribution defined over the vector $(X_v)_{v \in V}$. First, notice that exchangeability is closed under marginalization and conditioning:

Proposition 2. *If X_V is an exchangeable random vector then so are the marginal vectors X_A, for $A \subseteq V$, and the conditional vectors $X_A \mid X_C = x_C^*$, for disjoint A, C where $V = A \cup C$, if they exist.*

Proof. First we prove closedness under marginalization: For a permutation matrix π of A, let the permutation matrix over V be $\pi^*(a) = \pi(a)$, for $a \in A$, and $\pi^*(b) = b$ for $b \in V \setminus A$. The proof then follows from exchangeability of X_V.

Now we prove closedness under conditioning: By the definition of conditioning, we need to prove that $P(x_A, x_C^*) = P(x_{\pi(A)}, x_C^*)$ for every π. This is again true by considering $\pi^*(a) = \pi(a)$, for $a \in A$, and $\pi^*(c) = c$ for $c \in C$. $\quad\square$

Notice that the above result implies that the marginal/conditional $X_A \mid X_C$, (i.e., when $A \cup C \subset V$) is also exchangeable. We now have the following results:

Proposition 3. *If P satisfies vector exchangeability then the following holds:*

1. *P satisfies upward-stability if and only if it satisfies composition.*
2. *P satisfies downward-stability if and only if it satisfies intersection.*

Proof. 1. ($\Rightarrow$) follows from Lemma 1. To prove ($\Leftarrow$), let $i \perp\!\!\!\perp j \mid C$. By exchangeability, for an arbitrary $k \notin C \cup \{i, j\}$, we have $i \perp\!\!\!\perp k \mid C$. Composition implies $i \perp\!\!\!\perp j \cup k \mid C$. Weak union implies $i \perp\!\!\!\perp j \mid C \cup \{k\}$.

2. follows from 1. and the consequence of duality provided in Lemma 2. $\quad\square$

Proposition 4. *If P is exchangeable then it satisfies singleton-transitivity.*

Proof. If $i \perp\!\!\!\perp j \mid C$ and $k \notin C \cup \{i, j\}$ then clearly $i \perp\!\!\!\perp k \mid C$ and $k \perp\!\!\!\perp j \mid C$. $\quad\square$

We see that the skeleton of an exchangeable distribution can only take a very specific form:

Proposition 5. *If P is exchangeable then $\mathrm{sk}(P)$ is either an empty or a complete graph.*

Proof. If there is any independence statement of form $i \perp\!\!\!\perp j \mid C$ then by permutation for all variables, we obtain $k \perp\!\!\!\perp l \mid C'$ for all k and l. Therefore, $\mathrm{sk}(P)$ is empty. If there is no independence statement of this form then $\mathrm{sk}(P)$ is complete. $\quad\square$

Notice that for faithfulness to empty or complete graphs, it is immaterial whether one considers undirected or bidirected interpretation of graphs.

Corollary 1. *Let P satisfy vector exchangeability. If P is faithful to a graph (both under undirected interpretation and under bidirected interpretation) then the graph is empty or complete.*

Hence, there are two regimes available: if there is no independence statement implied by P then we are in the complete graph regime; and if there is at least one conditional independence statement implied by P then we are in the empty graph regime. The following also provides conditions for the opposite direction of the above result:

Theorem 1. *If P is exchangeable then P is faithful to a graph (both under undirected interpretation and under bidirected interpretation) if and only if P satisfies the intersection and composition properties. The graph must then be either empty or complete.*

Proof. The first result follows from Propositions 1, 3, and 4. The second follows from Proposition 5. □

Indeed other exchangeable distributions may exist but they are not faithful to a graph. We then have the following corollaries:

Corollary 2. *Let P be exchangeable and there exists an independence statement induced by P. It then holds that all variables X_v are completely independent of each other if and only if P satisfies intersection and composition.*

Corollary 3. *Let P be a regular exchangeable Gaussian distribution. If there is a zero element in its covariance matrix then all X_v are completely independent; and otherwise they are completely dependent.*

Proof. The proof follows from the fact that a regular Gaussian distribution satisfies the intersection and composition properties. □

Notice that the above statement could be shown otherwise since a zero off-diagonal entry of the covariance matrix can be permuted by exchangeability to every other off-diagonal entry, making the covariance matrix diagonal.

Proposition 6. *If P is exchangeable, satisfies intersection (this holds when P has a positive density), and if there exists one independence statement induced by P then all variables X_v are marginally independent of each other.*

Proof. By Proposition 3, P satisfies downward-stability. An independence statement $A \perp\!\!\!\perp B \mid C$, by the use of decomposition implies $i \perp\!\!\!\perp j \mid C$ for an arbitrary $i \in A$ and $j \in B$. Downward-stability implies $i \perp\!\!\!\perp j$. Exchangeability implies all pairwise marginal independences. □

Example 1. *Consider an exchangeable distribution P over four variables (i, j, k, l) with $i \perp\!\!\!\perp j \mid k$. By exchangeability, all independences of form $\pi(i) \perp\!\!\!\perp \pi(j) \mid \pi(k)$ hold for any permutation π on (i, j, k, l). It is easy to see that none of the semi-graphoid axioms can generate new independence statements from these. If intersection holds then, for example, from $i \perp\!\!\!\perp j \mid k$ and $i \perp\!\!\!\perp k \mid j$, we obtain $i \perp\!\!\!\perp \{j, k\}$, which by decomposition implies $i \perp\!\!\!\perp j$, and hence all marginal independences between singletons. If composition holds then, for example, from $i \perp\!\!\!\perp j \mid k$ and $i \perp\!\!\!\perp l \mid k$, we obtain $i \perp\!\!\!\perp \{j, l\} \mid k$, which by weak union implies $i \perp\!\!\!\perp j \mid \{k, l\}$, and hence all conditional independences between singletons given the remaining variables.*

4. Results for exchangeability for random networks

Henceforth, in the context of random networks, we consider vectors whose components are indexed by dyads (i.e. two-element subsets of the node set $\mathcal{N}$). Thus, for a conditional independence statement of the form $A \perp\!\!\!\perp B \mid C$, $A, B, C \subset \mathcal{D}(\mathcal{N})$ are pairwise disjoint subsets of dyads. (Later on, we particularly write the conditioning sets as simply C, which should be considered a subset of dyads.) Notice that we simply use the notation ij for a dyad whose endpoints are nodes i and j. This is different from the notation $\{i, j\}$, which indicates the set of two nodes i and j as used in the previous section.

4.1. Marginalization and conditioning for exchangeable random networks

Here we focus on marginalization over and conditioning on arbitrary sets of dyads. Notice that by *marginalizing over* a set M, we mean we marginalize the set M out, which results in a distribution over $\mathcal{D}(\mathcal{N}) \backslash M$. However, exchangeable networks are not always closed under marginalization over or conditioning on an arbitrary set of dyads:

For a marginal network X_A, where A is a subset of dyads, exchangeability and summing up all probabilities over values of the dyads that are marginalized over imply that $P(X_A = x_A) = P((X_{\pi(i)\pi(j)})_{ij \in A} = x_A)$, for any permutation π. However, this is not necessarily equal to $P(X_A = (x_{\pi(i)\pi(j)})_{ij \in A})$, which is what we need for exchangeability of the marginal to hold. For conditioning, in fact, $X_A \mid X_C$ is not exchangeable if a node appears in a dyad in A and a dyad in C (e.g. i appearing in $ij \in A$ and $ik \in C$). This is because, for a permutation π that maps i to a node other than i, exchangeability of $X_A \mid X_C$ is equivalent to $P(x_A \mid x_C) = P((x_{\pi(i)\pi(j)})_{ij \in A} \mid x_C)$, which itself is equivalent to $P(x_A, x_C) = P((x_{\pi(i)\pi(j)})_{ij \in A}, x_C)$. But, this does not necessarily hold as i is mapped to another node in A but not in C.

However, we have the following:

Proposition 7. *Let A and C be disjoint subsets of dyads of an exchangeable random network X such that A and C do not share any nodes, i.e. if $ij \in A$ then there is no dyad ik or jk in C for any node $k \in \mathcal{N}$. It then holds that the conditional/marginal random network $X_A \mid X_C$ is exchangeable.*

Proof. Let $\mathcal{N}(A \cup C)$ be the set of all endpoints of dyads in A and C, and define similarly $\mathcal{N}(A)$ and $\mathcal{N}(C)$. Define the permutation $\pi^* \in S(\mathcal{N}(A \cup C))$ such that $\pi^*(i) = \pi(i)$ for $i \in \mathcal{N}(A)$ and $\pi^*(k) = k$ for $k \in \mathcal{N}(C)$. Notice that this is well-defined since A and C do not share any nodes. Using π^* and by exchangeability of X, we conclude that $P(x_A, x_C) = P((x_{\pi(i)\pi(j)})_{ij \in A}, x_C)$, which, as mentioned before, is equivalent to the exchangeability of $X_A \mid X_C$. $\qquad\square$

4.2. Types of graphs faithful to exchangeable distributions

An analogous result to that of vector exchangeability, concerning the skeleton of an exchangeable probability distribution, was proven in [14]:

Proposition 8. *If a distribution P over a random network X is exchangeable then $\mathrm{sk}(P)$ is one of the following:*

1. *the empty graph;*
2. *the incidence graph;*
3. *the complement of the incidence graph;*
4. *the complete graph.*

Notice that a pairwise independence statement for an exchangeable P over a random network X is of form $ij \perp\!\!\!\perp kl \,|\, C$, where i, j, k, l are nodes of X. Depending on the type of $\mathrm{sk}(P)$, these statements take different forms.

Lemma 4. *Suppose that there exists an independence statement of form $ij \perp\!\!\!\perp kl \,|\, C$, $i \neq j, k \neq l$, for an exchangeable P over a random network X. It is then in one of the following forms depending on the type of $\mathrm{sk}(P)$:*

1. *empty graph $\Rightarrow$ no constraints on i, j, k, l;*
2. *incidence graph $\Rightarrow i \neq l, k$ and $j \neq l, k$;*
3. *complement of the incidence graph $\Rightarrow i = k$ or $i = l$ or $j = k$ or $j = l$;*

if $\mathrm{sk}(P)$ is the complete graph then it is not possible to have such a conditional independence statement.

Proof. The proof follows from the fact if there is an edge between ij and kl in $G(P)$ then there is no statement of form $ij \perp\!\!\!\perp kl \,|\, C$. $\square$

It is clear that if $\mathrm{sk}(P)$ is the complete graph then every dyad is completely dependent on every other dyad. Thus we consider the cases of the incidence graph skeleton and the complement of the incidence graph skeleton separately. However, before this, we show that among the known graphical models, only undirected and bidirected graphs can be faithful to an exchangeable distribution.

In order to do so, we can start off by considering any class of *mixed graphs*, i.e. graphs with simultaneous undirected, directed, or bidirected edges that use the (unifying) separation criterion introduced in [16]. To the best of our knowledge, the largest class of such graphs is the class of *chain mixed graphs* [16], which includes the classes of *ancestral graphs* [25], *LWF* [17] and *regression chain graphs* [30], and several others; see [16]. We require some definitions and results, including the formulation of the separation criterion, which we only need for the results in this subsection. We provide these together with the proof of the main result (Theorem 2) in Appendix A. If the reader is only interested in the statement of the theorem and not the proof, we suggest that they skip the material in the Appendix.

Two graphs are called *Markov equivalent* if they induce the same independence model.

Theorem 2. *If a distribution P over an exchangeable random network with n nodes is faithful to a chain mixed graph G then G is Markov equivalent to one of the following graphs:*

1. *the empty graph;*
2. *the undirected incidence graph, $L_-(n)$;*
3. *the bidirected incidence graph, $L_\leftrightarrow(n)$;*
4. *the undirected complement of the incidence graph, $L_-^c(n)$;*
5. *the bidirected complement of the incidence graph, $L_\leftrightarrow^c(n)$;*
6. *the complete graph.*

Hence, for exchangeable random networks, there are six regimes available, where these are three pairs that are complement of each other. Within the two non-trivial pairs, the two undirected and bidirected cases act as dual of each other as described in Section 2.6. We will make use of this duality to simplify the results and proofs.

Although the following method is not unique, here we provide a simple test to decide in which regime a given exchangeable distribution lies:

Algorithm 1. *For arbitrary fixed nodes i, j, k, l, m of a given exchangeable network, test the following:*

- $ij \perp\!\!\!\perp kl \mid C$, *for some* C, *and* $ij \perp\!\!\!\perp ik \mid C'$, *for some* $C' \Rightarrow$ *Empty graph;*
- $ij \perp\!\!\!\perp kl \mid C$, *for some* C, *and* $ij \not\!\perp\!\!\!\perp ik \mid C'$, *for any* $C' \Rightarrow L(n)$:

  - $ik \in C \Rightarrow L_-(n)$;

  - $ik \notin C \Rightarrow L_\leftrightarrow(n)$;

- $ij \not\!\perp\!\!\!\perp kl \mid C$, *for any* C, *and* $ij \perp\!\!\!\perp ik \mid C'$, *for some* $C' \Rightarrow L^c(n)$:

  - $lm \in C' \Rightarrow L_-^c(n)$;

  - $lm \notin C' \Rightarrow L_\leftrightarrow^c(n)$;

- $ij \not\!\perp\!\!\!\perp kl \mid C$, *for any* C, *and* $ij \not\!\perp\!\!\!\perp ik \mid C'$, *for any* $C' \Rightarrow$ *Complete graph.*

Proposition 9. *If a distribution P over an exchangeable random network is faithful to a graph G then Algorithm 1 determines the Markov equivalence class of G.*

Proof. First, we show that this test covers all the possible cases: The first level tests ($ij \perp\!\!\!\perp kl \mid C$ and $ij \perp\!\!\!\perp ik \mid C'$) clearly cover all the cases concerning the skeleton of the graph. But, the second level test ($ik \in C$ or $lm \in C'$), which only concerns the non-trivial skeletons, might not be consistent: For example, first consider the $L(n)$ case, and assume that $ij \perp\!\!\!\perp kl \mid C_1$ and $ij \perp\!\!\!\perp kl \mid C_2$, for some C_1, C_2, but $ik \in C_1$ and $ik \notin C_2$. However, in such cases, it is easy to show that P cannot be faithful to either of the two undirected or bidirected graphs. The case of $L^c(n)$ is similar.

The algorithm also outputs the correct regimes: The tests of $ij \not\!\perp\!\!\!\perp kl \mid C$ and $ij \perp\!\!\!\perp ik \mid C'$ clearly determine the skeleton $\mathrm{sk}(P)$. The test $ik \in C$ then determines the type of edges since, in $\langle ij, ik, kl \rangle$, if $ik \notin C$ then ij and kl cannot

be separated in the undirected graph; and if $ik \in C$ then ij and kl cannot be separated in the bidirected graph. The test $lm \in C'$ can be proven similarly. $\quad\square$

Unlike the vector exchangeability case, the intersection and composition properties are not in general sufficient for faithfulness of exchangeable network distributions to the graphs provided in Theorem 2. However, in special cases this holds. We detail this below for each regime mentioned above.

4.3. The incidence graph case

In this section, we assume that $\mathrm{sk}(P) = L_-(n)$. The following example shows that, in principle, intersection and composition are not sufficient for faithfulness of P and $L_-(n)$ (and P and the bidirected incidence graph $L_\leftrightarrow(n)$).

Example 2. *Suppose that there is an exchangeable P that induces $ij \perp\!\!\!\perp kl \,|\, C_{ij,kl}$, where $C_{ij,kl} = \{ik, il, jk, jl\}$. Exchangeability implies that $ij \perp\!\!\!\perp kl \,|\, C_{ij,kl}$ for all i, j, k, l. Suppose, in addition, that P satisfies upward-stability. Notice that here we do not show that such a probability distribution necessarily exists – one can treat this defined independence model as an example of "a network-exchangeable semi-graphoid".*

It is easy to see that $\mathrm{sk}(P) = L_-(n)$. Moreover, by Lemma 1, P satisfies composition. In addition, P satisfies intersection: Notice that none of the semi-graphoid axioms plus upward-stability and composition imply an independence statement of form $A \perp\!\!\!\perp B \,|\, C$ where A, B both contain a node i. In addition, it can be seen that these axioms imply that if there is $A \perp\!\!\!\perp B \,|\, C$ then for every $ij \in A$ and $kl \in B$, it holds that $C_{ij,kl} \subseteq C$. Let $C_{A,B} = \bigcup_{ij \in A, kl \in B} C_{ij,kl}$. By these two observations, we conclude that if $A \perp\!\!\!\perp B \,|\, C \cup D$ and $A \perp\!\!\!\perp D \,|\, C \cup B$ then $C_{A,D} \subseteq C$. By upward-stability all statements of form $ij \perp\!\!\!\perp kl \,|\, C$ where $C_{ij,kl} \subseteq C$ hold. Hence, by composition, we have that $A \perp\!\!\!\perp B \,|\, C$. Hence, by contraction, $A \perp\!\!\!\perp B \cup D \,|\, C$.

However, P does not satisfy singleton-transitivity: Consider $ij \perp\!\!\!\perp kl \,|\, C_{ij,kl}$ and $ij \perp\!\!\!\perp kl \,|\, C_{ij,kl} \cup \{km\}$. If, for contradiction, singleton-transitivity holds then, because $\mathrm{sk}(P) = L_-(n)$, we have that $ij \perp\!\!\!\perp km \,|\, C_{ij,kl}$. But, it can be seen that this statement is not in the independence model since no compositional graphoid axioms can generate this statement from $\{ij \perp\!\!\!\perp kl \,|\, C_{ij,kl}\}$.

By Proposition 1 (1.), it is implied that, although intersection and composition are satisfied, P is not faithful to $L_-(n)$.

If we now suppose that P induces $ij \perp\!\!\!\perp kl \,|\, C_{ij,kl}^d$, where $C_{ij,kl}^d = V \setminus (C_{ij,kl} \cup \{ij, kl\})$, P is exchangeable, and it satisfies downward-stability then, by using the duality (Lemmas 2 and 3), we conclude that although intersection and composition are satisfied, P is not faithful to $L_\leftrightarrow(n)$.

However, under certain assumptions, intersection and composition are sufficient for faithfulness. We will consider the two dual regimes within the incidence graph case related to the cases of Theorem 2: the undirected incidence graph

case and the bidirected incidence graph case. We make use of the duality between these to immediately extend the results in the undirected case to the bidirected case.

As in Example 2, let $C_{ij,kl} = \{ik, il, jk, jl\}$. For the faithfulness results to the undirected case, one assumption that is used is that for some (and because of exchangeability for all) i, j, k, l, $ij \perp\!\!\!\perp kl \,|\, C$ implies that $C_{ij,kl} \subseteq C$. For the faithfulness results to the bidirected case, one assumption is that for some (and because of exchangeability for all) i, j, k, l, $ij \perp\!\!\!\perp kl \,|\, C$ implies that $C_{ij,kl} \cap C = \varnothing$.

For a separation statement $A \perp B \,|\, C$, we define C to be a *minimal* separator in the case that if we remove any vertex from C, the separation does not hold; we define a *maximal* separator similarly. Let also $C_{ij} = \{ir, jr : \forall r \neq i, j\}$ and $C_{ij}^d = V \setminus (C_{ij} \cup \{ij\}) = \{lm : \forall l, m \notin \{i, j\}\}$. It holds that C_{ij} is a minimal separator of ij, kl in $L_-(n)$ and $C_{ij}^d \setminus \{kl\}$ is a maximal separator in $L_\leftrightarrow(n)$:

Proposition 10.

1. In $L_-(n)$, it holds that $ij \perp kl \,|\, C_{ij}$, and if $ij \perp kl \,|\, C$ then $|C_{ij}| \leq |C|$. In fact, if $C \neq C_{ij}$ and $C \neq C_{kl}$ then $|C_{ij}| < |C|$.
2. In $L_\leftrightarrow(n)$, it holds that $ij \perp kl \,|\, C_{ij}^d \setminus \{kl\}$, and if $ij \perp kl \,|\, C$ then $|C_{ij}^d| \geq |C| + 1$. In fact, if $C \neq C_{ij}^d \setminus \{kl\}$ and $C \neq C_{kl}^d \setminus \{ij\}$ then $|C_{ij}^d| > |C| + 1$.

Proof. 1. The first claim is straightforward to prove since every path from kl to ij must pass through an adjacent vertex of ij, which contains i or j. To prove the second statement, notice that if $im \notin C$ then at least km and lm must be in C. The same vertices km and lm appear if $jm \notin C$ too, but for no other vertices of C_{ij} missing in C. Thus, for every missing member of C_{ij} in C, there is at least a member of C_{kl} that should be in C. Hence, $|C_{ij}| \leq |C|$.

To prove the third statement, suppose, for contradiction, that $C \neq C_{ij}$ and $C \neq C_{kl}$ and $|C| = |C_{ij}|$. Using the fact that, for every missing member of C_{ij} in C, there is at least a member of C_{kl} that should be in C, there cannot be any vertex outside $C_{ij} \cup C_{kl}$ in C. Hence, $C \subset C_{ij} \cup C_{kl}$. If $n > 5$ and, say, $im, jm \notin C$ but $km, lm \in C$ then consider the vertices ih, jh, kh, lh. Without loss of generality, assume that $ih, jh \in C$ but $kh, lh \notin C$. Then the path $\langle kl, kh, mh, jm, ij \rangle$ connects kl and ij, a contradiction. The cases where $n = 3, 4, 5$ are easy to check.

2. The proof follows from 1. by using the duality (Lemma 3), and observing that $|C_{ij}^d \setminus \{kl\}| = |C_{ij}^d| - 1$. $\qquad\square$

However, not all minimal separators of ij, kl in $L_-(n)$ are of the form above. For example, in $L_-(6)$, consider the set $C = \{ik, il, im, jk, jl, jm, hk, hl, mh\}$. It holds that $ij \perp kl \,|\, C$. However, C is not of the form C_{pq} for any pair $p, q \in \{i, j, k, l, m, h\}$. The same can be said about $L_-(n)$: Consider $C^d = V \setminus (C \cup \{ij, kl\})$. By Lemma 3, we have that $ij \perp kl \,|\, C^d$ in $L_\leftrightarrow(6)$, and, in addition, C^d is maximal.

Proposition 11. *Let a distribution P be defined over an exchangeable random network and $\mathrm{sk}(P) = L_-(n)$, and consider some nodes i, j, k, l.*

1. *Suppose that for every minimal C such that $ij \perp\!\!\!\perp kl \,|\, C$, it holds that $C_{ij,kl} \subseteq C$ and C is invariant under swapping k and m, and l and h, for every m, h, $m \neq h$, $mh \notin C \cup \{ij, kl\}$. It then holds that if P satisfies composition then it satisfies upward-stability and singleton-transitivity.*

2. *Suppose that for every maximal C such that $ij \perp\!\!\!\perp kl \,|\, C$, it holds that $C_{ij,kl} \cap C = \varnothing$ and C is invariant under swapping k and m, and l and h, for every m, h, $m \neq h$, $mh \in C$. It then holds that if P satisfies intersection then it satisfies downward-stability and singleton-transitivity.*

Proof. By Lemma 4, we know that i, j, k, l are all different.

1. First, we prove upward-stability. Suppose that $ij \perp\!\!\!\perp kl \,|\, C$. Notice that, because of exchangeability, the assumptions of the statement hold for every i, j, k, l, m, h. For a variable $mh \notin C \cup \{ij, kl\}$, we prove that $ij \perp\!\!\!\perp kl \,|\, C \cup \{mh\}$ by induction on $|C|$. Notice that, by assumption, $mh \notin C_{ij,kl}$. In addition, without loss of generality, we can assume that $m, h \neq i, j$ since otherwise we can swap i and k and j and l and proceed as follows, and finally swap them back.

The base case is when C is minimal. Because of exchangeability, by the invariance of C under the aforementioned swaps, and since $mh \notin C_{ij,kl}$, we have that $ij \perp\!\!\!\perp mh \,|\, C$. By composition $ij \perp\!\!\!\perp \{kl, mh\} \,|\, C$, which, by weak union, implies $ij \perp\!\!\!\perp kl \,|\, C \cup \{mh\}$.

Now suppose that $ij \perp\!\!\!\perp kl \,|\, C$, C is not minimal, and, for every C' such that $|C'| < |C|$, $[ij \perp\!\!\!\perp kl \,|\, C' \Rightarrow ij \perp\!\!\!\perp kl \,|\, C' \cup \{op\}]$, for every op. Consider the independence statement $ij \perp\!\!\!\perp kl \,|\, C_0$ such that $C_0 \subset C$ and C_0 is minimal. We have that $ij \perp\!\!\!\perp mh \,|\, C_0$. By induction hypothesis, we can add vertices to the conditioning set in order to obtain $ij \perp\!\!\!\perp mh \,|\, C$. Now, again composition and weak union imply the result.

Singleton-transitivity also follows from the above argument. We need to show that $ij \perp\!\!\!\perp kl \,|\, C$ and $ij \perp\!\!\!\perp kl \,|\, C \cup \{mh\}$ imply $ij \perp\!\!\!\perp mh \,|\, C$ or $mh \perp\!\!\!\perp kl \,|\, C$. (Notice that because of upward-stability $ij \perp\!\!\!\perp kl \,|\, C \cup \{mh\}$ is immaterial.) Again without loss of generality, we can assume that $m, h \neq i, j$, and the above argument showed that $ij \perp\!\!\!\perp mh \,|\, C$.

2. The proof follows from part 1. and the duality (Lemma 2). □

Theorem 3. *Let a distribution P be defined over an exchangeable random network and $\mathrm{sk}(P) = L_-(n)$, and consider some nodes i, j, k, l.*

1. *Suppose that for every minimal C such that $ij \perp\!\!\!\perp kl \,|\, C$, it holds that $C_{ij,kl} \subseteq C$ and C is invariant under swapping k and m, and l and h, for every m, h, $m \neq h$, $mh \notin C \cup \{ij, kl\}$. Then P is faithful to $L_-(n)$ if and only if P satisfies the intersection and composition properties.*

2. *Suppose that for every maximal C such that $ij \perp\!\!\!\perp kl \,|\, C$, it holds that $C_{ij,kl} \cap C = \varnothing$ and C is invariant under swapping k and m, and l and h, for every m, h, $m \neq h$, $mh \in C$. Then P is faithful to $L_{\leftrightarrow}(n)$ if and only if P satisfies the intersection and composition properties.*

Proof. The proof follows from Propositions 11 and 1. □

4.4. The complement of the incidence graph case

In this section, we assume that $\mathrm{sk}(P) = L_{-}^{c}(n)$. We show again that, under certain assumptions, intersection and composition are sufficient for faithfulness. We again utilize the duality of the two additional regimes in the complement of the incidence graph case related to the cases of Theorem 2: the undirected complement of the incidence graph case and the bidirected complement of incidence graph case. Let $C_{ijk}^{d} = \{lm : \forall l, m \notin \{i, j, k\}\}$. For the faithfulness results to the undirected and bidirected case, an assumption that is used is that for some (and because of exchangeability for all) i, j, k, $ij \perp\!\!\!\perp ik \,|\, C$ implies that $C_{ijk}^{d} \subseteq C$ for undirected case and $C_{ijk}^{d} \cap C = \varnothing$ for the bidirected case.

Let $C_j = \{jr : \forall r \neq j\}$. Recall also that $C_{ij} = \{ir, jr : \forall r \neq i, j\}$ and $C_{ij}^{d} = \{lm : \forall l, m \notin \{i, j\}\}$. C_{ij}^{d} is a minimal separator of ij, ik in $L_{-}^{c}(n)$, and $C_{ij} \setminus \{ik\}$ is a maximal separator in $L_{\leftrightarrow}^{c}(n)$:

Proposition 12. *Let $n > 4$.*

1. *In $L_{-}^{c}(n)$, it holds that $ij \perp ik \,|\, C_{ij}^{d}$, and if $ij \perp ik \,|\, C$ then $|C_{ij}^{d}| \leq |C|$. In fact, if $C \neq C_{ij}^{d}$ and $C \neq C_{ik}^{d}$ then $|C_{ij}^{d}| < |C|$.*
2. *In $L_{\leftrightarrow}^{c}(n)$, $n > 4$, it holds that $ij \perp ik \,|\, C_{ij} \setminus \{ik\}$, and if $ij \perp ik \,|\, C$ then $|C_{ij}| \geq |C| + 1$. In fact, if $C \neq C_{ij} \setminus \{ik\}$ and $C \neq C_{ik}^{d} \setminus \{ij\}$ then $|C_{ij}^{d}| > |C| + 1$.*

Proof. 1. The first claim is straightforward to prove since every path from ik to ij must pass through an adjacent vertex of ij, which does not contain i or j. This set is C_{ij}^{d}. To prove the second statement, consider the sets $C_{ij}^{d} \setminus C_{ik}^{d} = C_k \setminus \{ik, jk\}$ and $C_{ik}^{d} \setminus C_{ij}^{d} = C_j \setminus \{ij, jk\}$, i.e., the neighbours of ij and ik with the joint neighbours removed. For every subset S of $C_k \setminus \{ik, jk\}$, clearly there are at least the same number of vertices in $C_j \setminus \{ij, jk\}$ adjacent to members of S. Hence, the Hall's marriage theorem [11] implies the result.

To prove the third statement, suppose, for contradiction, that $C \neq C_{ij}^{d}$ and $C \neq C_{ik}^{d}$ and $|C| = |C_{ij}^{d}|$. Using the fact that, for every missing member of C_{ij}^{d} in C, there is at least a member of C_{ik}^{d} that should be in C, there cannot be any vertex outside $C_{ij}^{d} \cup C_{ik}^{d}$ in a C. Hence, $C \subset C_{ij}^{d} \cup C_{ik}^{d}$. If $n > 4$ and, $km, jo \notin C$, where $o \neq m$, the path $\langle ik, jo, km, ij \rangle$ connects ik and ij, a contradiction. Thus, say, $km, jm \notin C$ (which implies $jl, kl \in C$). Then the path $\langle ik, jm, il, km, ij \rangle$ connects ik and ij, a contradiction.

2. The proof follows from the previous part and Lemma 3, and observing that $|C_{ij} \setminus \{ik\}| = |C_{ij}| - 1$. $\square$

However, not all minimal separators of ij, ik in $L_{-}^{c}(n)$ are of the form above. For example, in $L_{-}^{c}(5)$, consider the set $C = \{kl, jl, il, lm\}$. It holds that $ij \perp ik \,|\, C$. In addition, C is minimal. However, C is not of the form in the above proposition. In $L_{\leftrightarrow}^{c}(5)$, for the set $C^{d} = V \setminus (C \cup \{ij, ik\})$, by using Lemma 3, we see that $ij \perp ik \,|\, C^{d}$, and C^{d} is maximal.

Proposition 13. *Let a distribution P be defined over an exchangeable random network and $\mathrm{sk}(P) = L_{-}^{c}(n)$, and consider some nodes i, j, k.*

1. Suppose that for every minimal C such that $ij \perp\!\!\!\perp ik \,|\, C$, $C^d_{ijk} \subseteq C$ and C is invariant under swapping k and m for every m with an l such that $lm \notin C$. It then holds that if P satisfies composition then it satisfies upward-stability and singleton-transitivity.

2. Suppose that for every maximal C such that $ij \perp\!\!\!\perp ik \,|\, C$, $C^d_{ijk} \cap C = \varnothing$ and C is invariant under swapping k and m for every m with an l such that $lm \in C$. It then holds that if P satisfies intersection then it satisfies downward-stability and singleton-transitivity.

Proof. By Lemma 4, we know that the form of independencies for $\mathrm{sk}(P) = L^c_-(n)$ is $ij \perp\!\!\!\perp ik \,|\, C$, as provided in the statement of the proposition.

1. First, we prove upward-stability. Suppose that $ij \perp\!\!\!\perp ik \,|\, C$. Notice that, because of exchangeability, the assumptions of the statement hold for every i, j, k, l, m. For a variable $lm \notin C \cup \{ij, ik\}$, we prove that $ij \perp\!\!\!\perp ik \,|\, C \cup \{lm\}$ by induction on $|C|$. Notice that, by assumption, $lm \in C_{ijk}$. Without loss of generality, we can assume that $l \in \{i, j, k\}$, and further, $l = i$ or $l = k$ since otherwise we can swap j and k and proceed as follows, and finally swap them back.

The base case is when C is minimal. We have two cases: If $l = i$ then because of exchangeability, by the invariance of C under swapping k and m, and since $im \notin C^d_{ijk}$, we have that $ij \perp\!\!\!\perp im \,|\, C$. By composition $ij \perp\!\!\!\perp \{ik, im\} \,|\, C$, which, by weak union, implies $ij \perp\!\!\!\perp ik \,|\, C \cup \{im\}$. If $l = k$ then by swapping k and i, we have that $jk \perp\!\!\!\perp ik \,|\, C$. Now, notice that the statement $ij \perp\!\!\!\perp ik \,|\, C$ does not change under the kj-swap, which implies that C is also invariant under swapping j and m. By this swap, we have $km \perp\!\!\!\perp ik \,|\, C$. By this, $ij \perp\!\!\!\perp ik \,|\, C$, and the use of composition, we obtain $\{ij, km\} \perp\!\!\!\perp ik \,|\, C$, which, by weak union, implies $ij \perp\!\!\!\perp ik \,|\, C \cup \{km\}$.

Now suppose that $ij \perp\!\!\!\perp ik \,|\, C$, C is not minimal, and, for every C' such that $|C'| < |C|$, $[ij \perp\!\!\!\perp ik \,|\, C' \Rightarrow ij \perp\!\!\!\perp ik \,|\, C' \cup \{op\}]$, for every op. Consider the independence statement $ij \perp\!\!\!\perp ik \,|\, C_0$ such that $C_0 \subset C$ and C_0 is minimal. We have that $ij \perp\!\!\!\perp im \,|\, C_0$ or $km \perp\!\!\!\perp ik \,|\, C_0$. By induction hypothesis, we can add vertices to the conditioning set in order to obtain $ij \perp\!\!\!\perp im \,|\, C$ or $km \perp\!\!\!\perp ik \,|\, C$. Now, again composition and weak union imply the result.

Singleton-transitivity also follows from the above argument. We need to show that $ij \perp\!\!\!\perp ik \,|\, C$ and $ij \perp\!\!\!\perp ik \,|\, C \cup \{lm\}$ imply $ij \perp\!\!\!\perp lm \,|\, C$ or $lm \perp\!\!\!\perp ik \,|\, C$. (Notice that because of upward-stability $ij \perp\!\!\!\perp ik \,|\, C \cup \{lm\}$ is immaterial.) Again without loss of generality, we can assume that $l = i$ or $l = k$, and the above argument showed that $ij \perp\!\!\!\perp im \,|\, C$ or $km \perp\!\!\!\perp ik \,|\, C$.

2. The proof follows from part 1. and the duality (Lemma 2). □

Theorem 4. *Let a distribution P be defined over an exchangeable random network and $\mathrm{sk}(P) = L^c_-(n)$, and consider some nodes i, j, k.*

1. *Suppose that for every minimal C such that $ij \perp\!\!\!\perp ik \,|\, C$, $C^d_{ijk} \subseteq C$ and C is invariant under swapping k and m for every m with an l such that $lm \notin C$. Then P is faithful to $L^c_-(n)$ if and only if P satisfies the intersection and composition properties.*

2. *Suppose that for every maximal C such that $ij \perp\!\!\!\perp ik \mid C$, $C_{ijk}^d \cap C = \varnothing$ and C is invariant under swapping k and m for every m with an l such that $lm \in C$. Then P is faithful to $L_{\leftrightarrow}^c(n)$ if and only if P satisfies the intersection and composition properties.*

Proof. The proof follows from Propositions 13 and 1. $\qquad\square$

4.5. The empty graph case

Clearly every minimal separator in the empty graph is the empty set, and the maximal separator is all the remaining vertices. In the case where $\mathrm{sk}(P)$ is the empty graph, we have the following.

Proposition 14. *Let a distribution P be defined over an exchangeable random network, and $\mathrm{sk}(P)$ be the empty graph, and consider some nodes i, j, k, l.*

1. *Suppose that $ij \perp\!\!\!\perp kl$ and $ij \perp\!\!\!\perp ik$ hold. It then holds that if P satisfies composition then it satisfies upward-stability and singleton-transitivity.*
2. *Suppose that $ij \perp\!\!\!\perp kl \mid V \setminus \{ij, kl\}$ and $ij \perp\!\!\!\perp ik \mid V \setminus \{ij, ik\}$ hold. It then holds that if P satisfies intersection then it satisfies downward-stability and singleton-transitivity.*

Proof. By Lemma 4, we know that i, j, k, l are disjoint.

1. First, we prove upward-stability. Suppose that $ij \perp\!\!\!\perp kl \mid C$ or $ij \perp\!\!\!\perp ik \mid C$. Notice that, because of exchangeability, the assumptions of the statement hold for i, j, k, l that are all different. We prove that $ij \perp\!\!\!\perp kl \mid C \cup \{mh\}$ or $ij \perp\!\!\!\perp ik \mid C \cup \{mh\}$, for every $mh \notin C \cup \{ij, kl\}$ or $mh \notin C \cup \{ij, ik\}$, respectively, by induction on $|C|$.

The base case is when $C = \varnothing$. First, consider the case where $ij \perp\!\!\!\perp kl$. If $mh \notin C_{ij,kl}$ then by swapping k and m and l and h, we obtain $ij \perp\!\!\!\perp mh$. If $mh \in C_{ij,kl}$ then say $mh = jl$. We use $ij \perp\!\!\!\perp ik$ and first swap i and j to obtain $ij \perp\!\!\!\perp jk$. Now we swap k and l to obtain $ij \perp\!\!\!\perp jl$. (The other three cases of mh are similar.) Now composition and weak-union imply the result.

Now, consider the case where $ij \perp\!\!\!\perp ik$. First suppose that $mh \in C_{ij,kl}$. If $mh = il$ then by swapping j and l, we obtain $il \perp\!\!\!\perp ik$. If $mh = jk$ then by swapping i and k, we obtain $jk \perp\!\!\!\perp ik$. If $mh = jl$ then by swapping i and j and then k and l, we obtain $ij \perp\!\!\!\perp jl$. If $mh \notin C_{ij,kl}$ then use $ij \perp\!\!\!\perp kl$ and swap k and m and l and h to obtain $ij \perp\!\!\!\perp mh$. Now composition and weak-union imply the result.

The inductive step is similar to the inductive step of Proposition 11 or Proposition 13 depending on the form $ij \perp\!\!\!\perp kl \mid C$ or $ij \perp\!\!\!\perp ik \mid C$.

Singleton-transitivity also follows from the above argument.

2. The proof follows from the first part and the duality (Lemma 2). $\qquad\square$

Theorem 5. *Let a distribution P be defined over an exchangeable random network, and $\mathrm{sk}(P)$ be empty. Suppose also that one of the following cases holds for i, j, k, l that are all different:*

a) $ij \perp\!\!\!\perp kl$ *and* $ij \perp\!\!\!\perp ik$;
b) $ij \perp\!\!\!\perp kl \,|\, V \setminus \{ij, kl\}$ *and* $ij \perp\!\!\!\perp ik \,|\, V \setminus \{ij, ik\}$.

Then P is faithful to the empty graph if and only if P satisfies the intersection and composition properties.

Proof. The proof follows from Propositions 14 and 1. □

5. Summary and discussion

Our results concern the conditional independence models induced by exchangeable distributions for random vectors and random networks. We have shown that exchangeable random vectors are completely independent of each other or completely dependent of each other if they satisfy intersection and composition properties. In addition, they are marginally independent if there exists at least one independence statement and the intersection property is satisfied. The intersection property is well-understood, and we know that a positive joint density is a sufficient condition for it to hold; thus it is particularly important to study the composition property for exchangeable random vectors, which, in this case, is simplified to $A \perp\!\!\!\perp B \,|\, C \Rightarrow A \perp\!\!\!\perp (B \cup D) \,|\, C$, for every disjoint D.

For exchangeable random networks, as it turned out, the situation is much more complicated. As an important extension of our results in [14], we showed that the independence structures of exchangeable random networks that can be represented by a graph in graphical model sense are one of the six possible cases: completely dyadic-independent, faithful to the undirected or bidirected incidence graphs, faithful to the undirected or bidirected complement of the incidence graph, or completely dyadic-dependent.

The undirected and bidirected versions of the incidence graph and its complement are in fact dual to each other, so in a sense there are four regimes available. In other words, with exchangeability and duality factored in, one passes from the skeleton to the graphical Markov equivalence class. Exploiting this duality, all the results for the undirected case can be extended to the bidirected case. In addition, the remaining four cases are just two cases modulo graph complement as well. Although we failed to do so, it would be especially nice if this duality could be understood, so that one can simply present the results with two-fold duality arguments.

We have provided a simple test to decide in which of the six regimes an exchangeable random network lies in cases when it has a "structured" independence structure. The main two elements of the four "non-trivial" cases is whether an independence is of form $ij \perp\!\!\!\perp kl \,|\, C$ or $ij \perp\!\!\!\perp ik \,|\, C$; and whether $C_{ij,kl} = \{ik, il, jk, jl\}$ is in C or is disjoint from C.

We, in fact, do not have "necessary" and sufficient conditions for whether an exchangeable random network is structured, but rather sufficient conditions that, in addition to the expected intersection and composition properties, are mainly based on whether a minimal (maximal) separator is invariant under the mentioned node swaps of the network. For testing purposes, it is important to

stress that the only conditioning set that needs to be tested are the minimal ones, which would significantly improve the computational complexity of any relevant algorithms.

Indeed, in practice, it is more important to understand the independence structures of exchangeable statistical network models for random networks with (in most situations) binary dyads. One point is that binary distributions always satisfy singleton-transitivity [8], a necessary condition for faithfulness, although under our sufficient assumptions this condition is automatically satisfied. In general, however, it would be useful to study which actual exchangeable models for networks (such as exchangeable exponential random graph models [29]) satisfy the provided conditions, both when we deal with binary random networks or weighted ones.

Appendix A: Proof of Theorem 2

As mentioned in Section 4.2, we focus on the class of chain mixed graphs, which contains simultaneous undirected, directed, or bidirected edges, with the separation criterion introduced in [16]. We refrain from defining this class explicitly as it is not needed for our purpose. However, we note that we can focus only on simple graphs since it was shown in [16] that for any (non-simple) chain mixed graph there is a Markov equivalent simple graph (the collection of which constitutes the class of *anterial graphs*). We also only focus on *maximal graphs*, which are graphs where a missing edge between vertices u and v implies that there exists a separation statement of form $u \perp v \,|\, C$, for some C – again it was shown in [16] that for any (non-maximal) chain mixed graph there is a Markov equivalent maximal graph.

First, we need the following additional definitions: A *section* ρ of a walk is a maximal subwalk consisting only of lines, meaning that there is no other subwalk that only consists of lines and includes ρ. Thus, any walk decomposes uniquely into sections; these are not necessarily edge-disjoint and sections may also be single vertices. A section ρ on a walk ω is called a *collider section* if one of the following walks is a subwalk of ω: $i \longrightarrow \rho \longleftarrow j$, $i \longleftrightarrow \rho \longleftarrow j$, $i \longleftrightarrow \rho \longleftrightarrow j$. All other sections on ω are called *non-collider* sections. A *trisection* is a walk $\langle i, \rho, j \rangle$, where ρ is a section. If in the trisections, i and j are distinct and not adjacent then the trisection is called *unshielded*. We say that a trisection is collider or non-collider if its section ρ is collider or non-collider respectively.

We say that a walk ω in a graph is *connecting* given C if all collider sections of ω intersect C and all non-collider sections are disjoint from C. For pairwise disjoint subsets A, B, C, we say that A and B are separated by C if there are no connecting walks between A and B given C, and we use the notation $A \perp B \,|\, C$.

Lemma 5. *If two maximal graphs G and H are Markov equivalent then G and H have the same unshielded collider trisections.*

Proof. Because of maximality, G and H have the same skeleton. An unshielded trisection in these graphs cannot be a collider in one and a non-collider in the

other. This is because if that is the case (say an unshielded trisection between i and j and separation $i \perp j \mid C$), by Markov equivalence, it implies that an inner vertex of the corresponding section is not in C in one, but in C in the other, which is a contradiction. □

We will not need the converse of the above lemma, but only a weaker result:

Lemma 6. *If there are no unshielded collider trisections in G then G is Markov equivalent to $\mathrm{sk}(G)$.*

Proof. First, we show that if $A \not\perp B \mid C$ in G then $A \not\perp B \mid C$ in $\mathrm{sk}(G)$: It holds that there is a connecting walk ω in G between A and B given C. If there are no collider sections on ω then no vertex on ω is in C. Therefore, ω is a connecting walk in $\mathrm{sk}(G)$. If there is a collider section with endpoints i and j on ω then it has to be shielded. We can now replace this collider section with the ij edge. Applying this method repeatedly, we obtain a walk that has no vertex in C, and is, therefore, connecting in $\mathrm{sk}(G)$.

Now, we show that if $A \perp B \mid C$ in G then $A \perp B \mid C$ in $\mathrm{sk}(G)$: Consider an arbitrary path between A and B in $\mathrm{sk}(G)$. We need to show that there is a vertex on this path that is in C. Consider this path in G and call it ω. If all sections on ω are non-collider then there must be a vertex on ω that is in C, and we are done. Hence, consider a collider section with endpoints i and j on ω. This section is shielded; thus, replace the section with the ij edge. By repeating this procedure, we either obtain a path, with a subset of vertices of ω, whose sections are all non-collider; or we eventually obtain an edge between the endpoints of ω, which is impossible. □

The following lemma extends the concept of exchangeability for random networks to graphs in graphical models:

Lemma 7. *Suppose that a distribution P over an exchangeable random network with node set $\mathcal{N}$ is faithful to a graph G, and let π be a permutation function on $\mathcal{N}$. Let also H be the graph obtained by permuting the vertices of G by vertex ij being mapped to $\pi(i)\pi(j)$. Then P is faithful to H; hence, G and H are Markov equivalent.*

Proof. Let $\mathcal{J}_\pi(P)$ be the independence model obtained from $\mathcal{J}(P)$ by mapping independence statements $A \perp\!\!\!\perp B \mid C$ to $\pi(A) \perp\!\!\!\perp \pi(B) \mid \pi(C)$, where $\pi(A) = \{\pi(i)\pi(j) : ij \in A\}$, etc. It is obvious that $\mathcal{J}_\pi(P)$ is faithful to H. Because of exchangeability, we also have that $\mathcal{J}(P) = \mathcal{J}_\pi(P)$. Therefore, P is faithful to H. Hence, G and H are Markov equivalent. □

We are now ready to provide the proof of Theorem 2:

Proof of Theorem 2. Cases 1 and 6 are trivial. Thus, we need to consider cases 2 and 3 of Proposition 8. By Lemma 6, if there are no unshielded collider trisections in G then G is Markov equivalent to $L_-(n)$ or $L_-^c(n)$, respectively. Thus, suppose that there is an unshielded collider trisection in G.

For case 2 of Proposition 8, we can assume that there is an edge $12, 13$ in an unshielded collider trisection, such that there is an arrowhead at 13 on $12, 13$.

Now, consider an arbitrary edge ij, ik in G. Let π be a permutation that only swaps i and 1, j and 2, and k and 3, and call the resulted graph H. (Notice that it is possible that $j = 3$ and $k = 2$.) By Lemma 7, H and G are Markov equivalent. In addition, ik is in an unshielded collider trisection in H with an arrowhead at vertex ik on ij, ik. Hence, by Lemma 5, ik is in an unshielded collider trisection in G with an arrowhead at vertex ik on ij, ik. Since i, j, k are arbitrary, and in particular, every edge could be mapped to ij, ik by a permutation, we conclude that there is an arrowhead at every vertex on every edge in G. Therefore, G is a bidirected graph (and Markov equivalent to $L_{\leftrightarrow}(n)$).

For case 3 of Proposition 8, we assume that there is an edge $12, 34$ in an unshielded collider trisection, such that there is an arrowhead at vertex 34 on $12, 34$. In this case, we apply a similar method to the previous case, but by a permutation that only swaps i and 1, j and 2, k and 3, and l and 4. □

Acknowledgements

References

[1] ALDOUS, D. (1985). Exchangeability and Related Topics. In *École d'Été de Probabilités de Saint–Flour XIII — 1983* (P. Hennequin, ed.) 1–198. Springer-Verlag, Heidelberg. Lecture Notes in Mathematics 1117. MR0883646

[2] COX, D. R. and WERMUTH, N. (1993). Linear dependencies represented by chain graphs (with discussion). *Stat. Sci.* **8** 204–218; 247–277. MR1243593

[3] DARROCH, J. N., LAURITZEN, S. L. and SPEED, T. P. (1980). Markov fields and log-linear interaction models for contingency tables. *Ann. Statist.* **8** 522–539. MR0568718

[4] DAWID, A. P. (1979). Conditional independence in statistical theory (with discussion). *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **41** 1–31. MR0535541

[5] DEFINETTI, B. (1931). Funzione Caratteristica Di un Fenomeno Aleatorio. *6. Memorie* 251–299. Academia Nazionale del Linceo.

[6] DIACONIS, P. and FREEDMAN, D. (1980). Finite exchangeable sequences. *Ann. Probab.* **8** 745–764. MR0577313

[7] DIACONIS, P. and JANSON, S. (2008). Graph limits and exchangeable random graphs. *Rendiconti* **28** 33–61. MR2463439

[8] DRTON, M., STURMFELS, B. and SULLIVANT, S. (2009). *Lectures on Algebraic Statistics. Oberwolfach Seminars* **39**. Springer. MR2723140

[9] DRTON, M. and XIAO, H. (2010). Smoothness of Gaussian conditional independence models. *Contemp. Math.* **516** 155–177. MR2730747

[10] EAGLESON, G. K. and WEBER, N. C. (1978). Limit theorems for weakly exchangeable arrays. *Mathematical Proceedings of the Cambridge Philosophical Society* **84** 123–130. MR0501275

[11] HALL, P. (1935). On representatives of subsets. *J. London Math. Soc.* **10** 26–30.

[12] KAUERMANN, G. (1996). On a dualization of graphical Gaussian models. *Scand. J. Stat.* **23** 105–116. MR1380485

[13] LAURITZEN, S. L. (1996). *Graphical Models*. Clarendon Press, Oxford, United Kingdom. MR1419991

[14] LAURITZEN, S. L., RINALDO, A. and SADEGHI, K. (2018). Random networks, graphical models, and exchangeability. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 481–508. MR3798875

[15] LAURITZEN, S. L., RINALDO, A. and SADEGHI, K. (2019). On exchangeability in network models. *J. Algebr. Stat.* **10** 85–113. MR3947125

[16] LAURITZEN, S. L. and SADEGHI, K. (2018). Unifying Markov properties for graphical models. *Ann. Statist.* **46**. MR3845017

[17] LAURITZEN, S. L. and WERMUTH, N. (1989). Graphical models for association between variables, some of which are qualitative and some quantitative. *Ann. Statist.* **17** 31–57. MR0981437

[18] LNĚNIČKA, R. and MATÚŠ, F. (2007). On Gaussian conditional independence structures. *Kybernetika* **43** 327–342. MR2362722

[19] MALOUCHE, D., RAJARATNAM, B. and ROLFS, B. T. (2013). Duality in Graphical Models.

[20] MATÚŠ, F. (1992). Ascending and Descending Conditional Independence Relations. In *Trans. 11th Prague Conf. Information Theory, Statistical Decision Functions and Random Processes , vol. B.* 181–200.

[21] MATÚŠ, F. (1995). Finite partially exchangeable arrays. Technical Report No. 1856, Inst. of Inf. Theory and Autom., `http://staff.utia.cas.cz/matus/CsiMaGSI13.pdf`.

[22] PEARL, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers, San Mateo, CA, USA. MR0965765

[23] PEARL, J. and PAZ, A. (1985). *Graphoids: A Graph-Based Logic for Reasoning about Relevance Relations*. University of California (Los Angeles). Computer Science Department.

[24] PETERS, J. (2015). On the intersection property of conditional independence and its application to causal discovery. *J. Causal Inference* **3** 97–108.

[25] RICHARDSON, T. S. and SPIRTES, P. (2002). Ancestral graph Markov models. *Ann. Statist.* **30** 962–1030. MR1926166

[26] SADEGHI, K. (2017). Faithfulness of probability distributions and graphs. *J. Mach. Learn. Res.* **18**. MR3763782

[27] SILVERMAN, B. W. (1976). Limit theorems for dissociated random variables. *Adv. in Appl. Probab.* **8** 806–819. MR0431348

[28] STUDENÝ, M. (2005). *Probabilistic Conditional Independence Structures*. Springer-Verlag, London, United Kingdom. MR3183760

[29] WASSERMAN, S. and PATTISON, P. (1996). Logit models and logistic re-

gressions for social networks: I. An introduction to Markov graphs and p^*. *Psychometrika* **61** 401–425. [MR1424909](#)

[30] WERMUTH, N., COX, D. R. and PEARL, J. (1994). Explanation for multivariate structures derived from univariate recursive regressions. Technical Report No. 94(1), Univ. Mainz, Germany.

[31] WERMUTH, N., MARCHETTI, G. M. and COX, D. R. (2009). Triangular systems for symmetric binary variables. *Electron. J. Stat.* **3** 932–955. [MR2540847](#)

[32] WEST, D. B. (2001). *Introduction to Graph Theory.* Prentice Hall, Upper Saddle River, NJ, USA. [MR1367739](#)