Two generalizations of Markov blankets

Victor Cohen, Axel Parmentier
Ecole des Ponts Paristech, CERMICS, Université Paris-Est, Marne-la-Vallée, France
March 11, 2019

Abstract

In a probabilistic graphical model on a set of variables \( V \), the Markov blanket of a random vector \( B \) is the minimal set of variables conditioned to which \( B \) is independent from the remaining of the variables \( V \setminus B \). We generalize Markov blankets to study how a set \( C \) of variables of interest depends on \( B \). Doing that, we must choose if we authorize vertices of \( C \) or vertices of \( V \setminus C \) in the blanket. We therefore introduce two generalizations. The Markov blanket of \( B \) in \( C \) is the minimal subset of \( C \) conditionally to which \( B \) and \( C \) are independent. It is naturally interpreted as the inner boundary through which \( C \) depends on \( B \), and finds applications in feature selection. The Markov blanket of \( B \) in the direction of \( C \) is the nearest set to \( B \) among the minimal sets conditionally to which ones \( B \) and \( C \) are independent, and finds applications in causality. It is the outer boundary of \( B \) in the direction of \( C \). We provide algorithms to compute them that are not slower than the usual algorithms for finding a d-separator in a directed graphical model. All our definitions and algorithms are provided for directed and undirected graphical models.

Keywords Markov blanket, probabilistic graphical models, feature selection, causality

1 Introduction

Markov blanket, probabilistic graphical models, feature selection, causality

A distribution on a set of variables \( V \) factorizes as a probabilistic graphical model on a graph \( G = (V, A) \) if variables in \( V \) satisfy some independences that are encoded by \( G \). Given a set \( B \) of variables in \( V \), the Markov blanket of \( B \) is the boundary in \( V \setminus B \) through which \( B \) and \( V \setminus B \) are dependent. More formally, it is the smallest subset \( M \) of \( V \setminus B \) such that

\[
B \perp \perp V \setminus (B \cup M) \mid M
\]

for any distribution that factorizes as a probabilistic graphical model on \( G \), where, given three random vectors \( X, Y, \) and \( Z \), we denote by

\[
X \perp \perp Y \mid Z
\]

the fact that \( X \) is independent from \( Y \) given \( Z \). As illustrated on Figure 1, \( \text{mb}(B) \) corresponds to the “outer boundary” of \( B \), and \( \text{mb}(V \setminus B) \) to its “inner boundary”. The Markov Blanket of \( B \) is the smallest set of variables of \( V \setminus B \) containing all the information about \( B \) that is in \( V \setminus B \) [Pellet and Elisseeff, 2008].

In this paper, we introduce two generalizations of Markov blankets to model how a subset of variables depends on another. The first is the Markov blanket of \( B \) in \( C \), which we denote by \( \text{mb}_C(B) \). It is the smallest subset \( M \) of \( C \) such that \( B \perp \perp C \setminus M \mid M \). The second is the Markov blanket of \( B \) in the direction of \( D \), which we denote by \( \text{mb}(B \to D) \). Among the sets \( M \) in \( V \setminus B \) such that \( B \perp \perp D \mid M \) and that are minimal for inclusion, it is the “nearest” to \( B \). Figure 2 illustrated how these notions can be interpreted as inner and outer boundaries.

We introduce \( \text{mb}_C(B) \) and \( \text{mb}(B \to D) \) in directed and undirected graphical models. We characterize \( \text{mb}_C(B) \) and \( \text{mb}(B \to D) \) in terms of separation and d-separation, which provides
Indeed, if a patient suffering from disease $D$ might cause $B$, it might also be that $B$ and $D$ are both caused by another factor. Fixing $B$ will cure the patient from $D$ only if $B$ is a cause of $D$. Counting the number of patients suffering from $D$ among those having $B$ indicates the correlation of $B$ and $D$, i.e., the conditional probability $\mathbb{P}(D|B)$ of $D$ given $B$, but not the causal effect of $B$ and $D$. To measure this causal effect, we need to compute the conditional probability of $D$ given $B$ in an experiment where, all other things being equal, parameter $B$ is controlled. We denote it by $\mathbb{P}(D|\text{do}(B))$. If $B$ and $D$ are random variables of a probabilistic graphical model, causality theory enables to identify if the causal effect $\mathbb{P}(D|\text{do}(B))$ can be computed from historical data without setting up a new experiment, and to compute it when it is possible. Shpitser and Pearl [2012] introduce an algorithm which returns all the causal effects $\mathbb{P}(D|\text{do}(B))$ that can be computed in a directed graphical model. This algorithm, which uses the back-door criterion [Pearl, 1993], requires to compute a d-separator between $(\text{dsc}(B) \cap \text{asc}(D)) \cup D$ and $B$ in the graph where we remove arcs outgoing from $B$, where asc($M$) and dsc($M$) respectively denote the ascendants and descendants of a set of vertices $M$. Let $S$ be such d-separator. Computing the causal effect of $B$ on $D$ becomes equivalent to computing conditional probabilities and marginals in a directed graphical model [Lauritzen 1999, e.g. Theorem 1.14]):

$$\mathbb{P}(D|\text{do}(B = b)) = \sum_s \mathbb{P}(D = s, B = b)\mathbb{P}(S = s)$$

Hence, we need to perform an inference task to compute the probabilities in the sum above. This latter inference problem is easier if the d-separator is small and near to $B$. The Markov Blanket of $(\text{dsc}(B) \cap \text{asc}(D)) \cup D$ in the direction of $B$ is therefore an excellent candidate as d-separator $S$: it is the nearest from $(\text{dsc}(B) \cap \text{asc}(D)) \cup D$ among all the minimal d-separator between $(\text{dsc}(B) \cap \text{asc}(D)) \cup D$ and $B$.

Figure 1: Markov blanket of $B$ and $V\setminus B$

Section 2 introduces the notions and notations we need on directed and undirected graphical models, as well as a literature review on Markov blankets. Section 3 the introduces the Markov blanket of $B$ in $C$, and Section 4 the Markov blanket of $B$ in the direction of $D$. 

...
2 Preliminaries on probabilistic graphical models

2.1 Graphs

A graph is a pair \( G = (V, A) \) where \( V \) is a finite set and \( A \) is a family of unordered pairs from \( V \). A vertex \( v \) is an element of \( V \). In an undirected graph, the pairs \( e = (u, v) \) in \( A \) are unordered and called edges. In a directed graph, the pairs \( a = (u, v) \) in \( A \) are ordered and called arcs.

A \( u \)-\( v \) path \( P \) in a graph is a sequence of vertices \( v_0, \ldots, v_k \) such that \( v_0 = u, v_k = v \), and \( (v_i, v_{i+1}) \) belongs to \( A \) for each \( i \) in \([k]\). Remark that if \( v_1, \ldots, v_k \) is a path in an undirected graph, then \( v_k, \ldots, v_1 \) is also a path. But if \( v_1, \ldots, v_k \) is a path in a directed graph, then \( v_k, \ldots, v_1 \) is generally not a path. A cycle in a graph is a path \( v_0, \ldots, v_k \) such that \( k > 0 \) and \( v_0 = v_k \). An directed graph is acyclic if it has no cycle. A \( u \)-\( v \) trail in an acyclic directed graph is a sequence of vertices \( v_1, \ldots, v_k \) such that \( v_0 = u, v_k = v \), and either \((v_{i-1}, v_i) \) or \((v_i, v_{i+1}) \) belongs to \( A \) for each \( i \) in \([i]\). A vertex \( v \) in a trail \( v_0, \ldots, v_k \) is a v-structure if \( 0 < i < k \) and \((v_{i-1}, v_i) \) and \((v_{i+1}, v_i) \) belong to \( A \). A clique in an undirected graph is a subset \( C \) of vertices of \( V \) such that, if \( u \) and \( v \) are two distinct elements of \( V \), then \((u, v) \) belongs to \( A \).

Let \( G \) be an acyclic directed graph. A parent of a vertex \( v \) is a vertex \( u \) such that \((u, v) \) belongs to \( A \); we denote by \( \text{prt}(v) \) the set of parents of \( v \). A vertex \( u \) is an ascendant (resp. a descendant) of \( v \) if there exists a \( u \)-\( v \) path (resp. a \( v \)-\( u \) path). We denote respectively \( \text{asc}(v) \) and \( \text{dsc}(v) \) the set of ascendants and descendants of \( v \). Finally, let \( \overline{\text{asc}}(v) = \{v\} \cup \text{asc}(v) \), and \( \overline{\text{dsc}}(v) = \{v\} \cup \text{dsc}(v) \). For a set of vertices \( C \), the parent set of \( C \), again denoted by \( \text{prt}(C) \), is the set of vertices \( u \) that are parents of a vertex \( v \in C \). We define similarly \( \text{asc}(C) \), and \( \text{dsc}(C) \).

We associate with each vertex \( v \) in \( V \) a random variable \( X_v \) taking its value in a finite set \( \mathcal{X}_v \). For any subset \( A \) of \( V \), we define \( X_A \) as the subvector \((X_v)_{v \in A}\), and \( \mathcal{X}_A \) as the Cartesian product \( \bigotimes_{v \in A} \mathcal{X}_v \).

2.2 Undirected graphical model

Given an undirected graph \( G = (V, A) \), a probability distribution \( P \) on \( \mathcal{X}_V \) factorizes as an undirected graphical model on \( G \) if there exists a collection \( C \) of cliques of \( G \), and mappings \( \psi_C : \mathcal{X}_C \rightarrow \mathbb{R}^+ \) for each \( C \) in \( C \) such that

\[
P(X_V = x_V) = \frac{1}{Z} \prod_{C \in C} \psi_C(x_C),
\]

where \( Z \) is a constant ensuring that \( P \) is a probability distribution. Vertices of a graphical model corresponds to random variables, and sets of vertices to random vectors.

A \( u \)-\( v \) path \( P \) is active given a subset of vertices \( M \) if no vertex of \( P \) is in \( M \). A set of vertices \( M \) separates two sets of vertices \( X \) and \( Y \) if there is no active path between a vertex of \( X \) and a vertex of \( Y \), which we denote by

\[X \perp Y|M.\]
Given three random vectors $X$, $Y$, and $M$, graphical model theory tells us that $X$ is independent from $Y$ given $M$ for any distribution that factorizes as a graphical model on $G$ if and only if $M$ separates $X$ and $Y$ (see e.g. Theorem 4.3 of [Koller and Friedman 2009]).

We are interested in independences of probabilistic graphical models $G$, that is, independences that are true for any distribution that factorizes as a graphical models. Such independences must therefore be characterized only in terms of the structure of $G$, that is, in terms of separation and d-separating.

### 2.3 Directed graphical models

Let $G = (V, A)$ be an acyclic directed graph. A conditional distribution of $v$ given its parent is a mapping $p_{v|\text{pt}(v)} : \mathcal{X}_v \times \mathcal{X}_{\text{pt}(v)} \rightarrow \mathbb{R}_+$ such that, for each $x_{\text{pt}(v)}$ in $\mathcal{X}_{\text{pt}(v)}$, the mapping $x_v \mapsto p_{v|\text{pt}(v)}(x_v, x_{\text{pt}(v)})$ is a probability distribution. A distribution $\mathbb{P}$ on $\mathcal{X}_v$ factorizes as a directed graphical model on $G$ if there exists conditional distributions $p_{v|\text{pt}(v)}$ such that

$$\mathbb{P}(x_v) = \prod_{v \in V} p_{v|\text{pt}(v)}(x_v, x_{\text{pt}(v)}).$$

Given a subset $M$ in $V$, a $u$-$v$ trail $P$ is active if and only if any vertex $v$ in $P$ that is not a v-structure does not belong to $P$, and any vertex $v$ in $P$ that is a v-structure is such that $\text{dsc}(v) \cap M \neq \emptyset$. Given three random vectors $X$, $Y$, and $M$, then $M$ d-separates $X$ and $Y$ if there is no active trail between $X$ and $Y$ that is active given $M$, which we again denote by

$$X \perp Y|M.$$

Three random vectors $X$, $Y$, and $M$ are such that $X$ is independent from $Y$ given $M$ for any distribution that factorizes as a graphical model on $G$ is and only if $X$ is d-separated from $Y$ given $M$ (see e.g. Theorems 3.4 and 3.5 of [Koller and Friedman 2009]).

### 2.4 Markov blankets and separators

A separator (resp. a d-separator) between two set of vertices $B$ and $D$ given an evidence set $E$ in an undirected (resp. directed) graphical model $G$ is a set of vertices $M$ that separates (resp. d-separates) $B$ and $D$. A (d-)separator $M$ between two sets of vertices $B$ and $D$ given an evidence set $E$ is minimal if for any strict subset $M'$ of $M$, $M' \cup E$ does not (d-)separate $C$ and $D$.

The Markov blanket $\text{mb}(B)$ of $B$ is the smallest (d-)separator $M \subseteq V \setminus B$ of $B$ and $V \setminus B$. By smallest, we mean that any (d-)separator $M \subseteq V \setminus B$ of $B$ and $V \setminus B$ contains $\text{mb}(B)$.

### 2.5 Literature review

Markov blankets are built on the fact that independences in a graphical model are characterized in terms of separation and d-separation. Lauritzen et al. [1990] introduces the notion separation in a undirected graphical model, which coincides with the separation in graph theory. The author also introduces the notion of d-separation in a directed graphical model. Geiger et al. [1990] presents the Bayes-ball algorithm that checks if two vertices in a directed graph $G = (V, A)$ are d-separated by a given set of vertices in $O(|V| + |A|)$. Pearl [1988] introduced the notion of Markov Blanket in the context of causal structure learning, under the name Markov boundary. Given samples of a set of random variables, causal structure learning aims at learning a directed graphical model that represents the causal structure of the random variables. Pearl [1988] and Spirtes et al. [2000] characterize graphically the Markov blanket: in undirected graphical model, it is the set of neighbors of $B$, while in directed graphical models, it is the set of parents, co-parents, and children of $B$.
Figure 3: The Markov blanket of $t$ is $\{u, v\}$, and its Markov blanket in $C$ is $\{u, w\}$.

Our generalizations of Markov blankets are minimal d-separators between two sets $B$ and $D$. As we mentioned in Example 2, minimal d-separators play a role in causality theory. In that context, Tian and Paz [1998] prove that a minimal d-separator between two subsets of variables can be found with a polynomial algorithm in $O(|V| \cdot |A|)$.

3 Markov blanket in a set

We now introduce the notion of Markov blanket in a set.

**Definition 1.** Let $B$, $C$ and $E$ be three set of vertices in a graph $G = (V, A)$. The Markov blanket of $B$ in $C$ given $E$, denoted by $mb_C(B|E)$, is the smallest subset $M \subseteq C$ of vertices satisfying

$$X_B \perp X_{C \setminus (B \cup M)} | X_{M \setminus E}$$

for any distribution that factorizes on $G$, (2)

where smallest means that a set $M \subseteq C$ satisfies (2) if and only if $mb_C(B|E) \subseteq M$.

Note that this definition holds both in directed and undirected graphical models. When $E = \emptyset$, we use the simpler notation $mb_C(B)$. The Markov blanket $mb_C(B)$ coincides with $mb(B)$ if $C = V$. Figure 3 illustrates the difference between the usual Markov blanket and the Markov blanket in a set.

The next theorem shows the existence and uniqueness of the Markov Blanket in a set and provides a graphical characterization in directed and undirected graphical models.

**Theorem 1.** Let $B$, $C$ and $E$ be three sets of vertices in a graph $G = (V, A)$. The Markov blanket of $B$ in $C$ given $E$ exists, is unique, and equal to

$$mb_C(B|E) = \left\{ v \in C : v \text{ is not (d-)separated from } B \text{ given } E \cup (C \setminus (B \cup \{v\})) \right\},$$

(3)

where “d-separated” and “separated” apply in directed and undirected graphical models respectively.

The Markov blanket in a set no longer admits a characterization in terms of parents, coparents, children and neighbor vertices. However, thanks to the characterization in [3], $mb_C(B|E)$ can be computed in $O(|C|(|A| + |V|))$ using a (d-)separation algorithm [Geiger et al. 1990].

**Proof of Theorem 1.** In undirected graphical models. Let $B$, $C$, and $E$ be three sets of vertices, and $M$ as in (3).

We start by proving that $B$ is separated from $C \setminus (B \cup M)$ given $M \cup E$. Let $v$ be a vertex in $C \setminus (B \cup M)$, and $P$ be a $B-v$ path. As $v$ does not belong to $M$, path $P$ is not active given $E \cup (C \setminus (B \cup \{v\}))$, and there is a vertex in $E \cup (C \setminus (B \cup \{v\}))$ on $P \setminus \{v\}$. Let $w$ be the first vertex of $P$ in that set, starting from $B$. If $w$ is in $E$, path $P$ is not active given $E \cup M$. Otherwise, the $B-w$ restriction of $P$ is active given $E \cup (C \setminus (B \cup \{w\}))$. Vertex $w$ thus belongs to $M$ and $P$ is not active given $E \cup M$, which gives the result.

Let $N$ be a subset of $C$ such that $B$ is separated from $C \setminus (N \cup B)$ given $N \cup E$. Let $v$ be a vertex in $M$. By definition of $v$, there exists a $B-v$ path that is active given $E \cup C \setminus (B \cup \{v\})$
with a minimum number of arcs. Let \( P \) be such a path. The only intersection of \( P \) with \( E \cup C \) is \( \{v\} \). Path \( P \) is therefore not active given \( E \cup N \) if and only if \( v \) belongs to \( N \). Hence \( v \) belongs to \( N \), and we obtain \( M \subseteq N \). 

The proof for directed graphical models is similar but more technical due to d-separation.

**Proof of Theorem 1** directed graphical models. Let \( B, C, \) and \( E \) be three sets of vertices, and \( M \) as in (3).

We start by proving that \( B \) is d-separated from \( C \setminus (B \cup M) \) given \( M \cup E \). Let \( P \) be a trail between a vertex \( b \in B \) and a vertex \( v \in C \setminus (B \cup M) \). We prove that \( P \) is not active. Without loss of generality, we can suppose that \( P \cap B = \{b\} \). Indeed, if \( P \) is active, then any of its subtrails whose extremities are not in \( M \) must be active. As \( B \cap M = \emptyset \), it suffices to show that the subtrail \( Q \) between the last vertex of \( P \) in \( B \) (starting from \( b \)) is not active. If \( P \) has a v-structure that is not active given \( E \cup M \), or if \( P \) has a vertex that is not a v-structure in \( E \cup M \), then \( P \) is not active. Suppose now that we are not in one of those cases. Starting from \( b \), let \( w \) be the first vertex of \( P \) in \( C \) that is not the middle of a v-structure in \( P \), and let \( Q \) be the \( b-w \) subtrail of \( P \). By definition of \( w \), any vertex of \( Q \) that is not in the middle of a v-structure is not in \( C \), and by hypothesis it is not in \( E \), hence it is not in \( E \cup (C \setminus (B \cup \{v\})) \). Furthermore, by hypothesis, any v-structure of \( Q \) is active given \( E \cup M \). Suppose that \( w \) is not in \( M \): we obtain \( M \subseteq E \cup (C \setminus (B \cup \{w\})) \), and hence, any v-structure of \( Q \) is active given \( E \cup (C \setminus (B \cup \{w\})) \). Therefore \( Q \) is active given \( E \cup (C \setminus (B \cup \{w\})) \) and \( w \in M \), which is a contradiction. We deduce that \( w \in M \). Hence \( w \neq v \). As \( w \in M \) is not in the middle of a v-structure, \( P \) is not active given \( M \cup E \), which gives the result.

Conversely, let \( N \subseteq C \) be a set of vertices such that \( B \) is d-separated from \( C \setminus (N \cup B) \) given \( N \cup E \). We now prove that \( M \subseteq N \). This part of the proof is illustrated on Figure 4. Let \( v \) be a vertex in \( M \). As \( v \) is in \( M \), there is an active trail between \( B \) and \( v \) given \( E \cup (C \setminus \{v \cup B\}) \). Let \( P \) be such a trail. Without loss of generality, we can suppose \( B \cap P = \{b\} \). As \( P \) is active given \( E \cup (C \setminus \{v \cup B\}) \) and \( B \cap P = \{b\} \), any vertex of \( P \setminus \{b,v\} \) that is not in the middle of a v-structure is not in \( C \setminus \{v \cup B\} \), and hence not in \( C \), and not in \( N \). Starting from \( b \), let \( u_{1}, \ldots, u_{k} \) be an indexation of the vertices of \( P \) that are in the middle of v-structures in \( P \). We prove by iteration on \( i \) that \( \text{desc}(u_{i}) \cap (E \cup N) \neq \emptyset \). Suppose the result true up to \( i - 1 \), and \( P_{i} \) be the subtrail of \( P \) from \( b \) to \( u_{i} \). Suppose that \( u_{i} \) is not in \( \text{desc}(E) \). As \( P \) is an active trail given \( E \cup (C \setminus \{v \cup B\}) \) and \( u_{i} \) is in the middle of a v-structure, \( u_{i} \) has a descendant \( w \) in \( C \setminus \{v \cup B\} \), and there is a directed path \( Q \) from \( u_{i} \) to \( w \). Let \( w' \) be the first vertex of \( Q \) in \( C \setminus \{v \cup B\} \) and \( Q' \) the \( u_{i}-w' \) restriction of \( Q \). Note that we may have \( u_{i} = w \) or \( u_{i} = w' \). Suppose that \( w' \notin N \). It implies that \( w' \in C \setminus (N \cup E) \). By induction hypothesis, the trail \( P_{i} \) followed by \( Q' \) is active given \( N \cup E \) between \( B \) and \( C \setminus (N \cup E) \). It contradicts Equation (2) for \( N \). We deduce that
Proposition 1. Let $B$, $C$, $C'$ and $E$ be four sets of vertices. Then $mb_{C\cup C'}(B|E) = mb_{C}(B|E)$ if and only if $C' \perp B|C \cup E$.

Proof of Proposition 1 for undirected graphical models. Suppose that $C' \perp B|C \cup E$. Let $v \in mb_{C\cup C'}(B|E)$, there exists an active path $Q$ between $B$ and $v$ such that $Q \cap (C \cup C' \cup E) = \emptyset$. Therefore $Q \cap (C \cup E) = \emptyset$. If $v \in C'$, then the assumption $C' \perp B|C \cup E$ implies that $Q \cap (C \cup E) \neq \emptyset$, which contradicts our assumption. We deduce that $v \in C$ and $v$ is not separated from $B$ by $C \cup E$. Therefore, $v \in mb_{C}(B|E)$. Let $u \in mb_{C}(B|E)$, there exists a path $Q$ from $B$ to $u$ such that $Q \cap (C \cup E) = \emptyset$. If $Q \cap C' \neq \emptyset$, the assumption $C' \perp B|C \cup E$ implies that $C \cup E$ intersects $Q$ which contradicts our assumption on $Q$. Therefore, $Q \cap (C \cup C' \cup E) = \emptyset$. We deduce that $v \in mb_{C\cup C'}(B|E)$. It achieves the proof.

Proof of Proposition 1 for directed graphical models. Let $B$, $C$, and $C'$ be such that $C' \perp B|C \cup E$. We only have to show that, given a vertex $v$ in $C$ and a $B$-$v$ trail $P$, then $P$ is active given $(C \cup C')(\{B \cup \{v\}\}) \cup E$ if and only if $P$ is active given $(C \setminus (B \cup \{v\})) \cup E$. Let $v$ be a vertex in $C$ and $P$ be a $B$-$v$ trail. W.l.o.g., we suppose that it intersects $B$ at most once, and $v$ at most once.

Suppose that $P$ is active given $(C \setminus (B \cup \{v\})) \cup E$. Then $P$ does not intersect $C'$. Indeed, suppose it intersects $C'$ in a vertex $w$. Then, the $B$-$w$ subtrail is active given $C \setminus (B \cup \{v\}) \cup E$, which contradicts $B \perp C'|C \cup E$. Furthermore, all the v-structures of $P$ are active given $(C \cup C')(\{B \cup \{v\}\}) \cup E$, as they have a descendant in $(C \setminus (B \cup \{v\})) \cup E$. Hence $P$ is active given $(C \cup C')(\{B \cup \{v\}\}) \cup E$.

Suppose now that $P$ is active given $(C \cup C')(\{B \cup \{v\}\}) \cup E$. It intersects $C \setminus (B \cup \{v\}) \cup E$ only on v-structures, and all these v-structures are active given $(C \cup C')(\{B \cup \{v\}\}) \cup E$. Suppose that there is a v-structure that is not active given $(C \setminus (B \cup \{v\})) \cup E$, and let $s$ be the first one starting from $B$. Then $s$ has a descendant $w$ in $C \setminus (C \cup E)$, and the $B$-$s$ subtrail of $P$ followed by the $s$-$w$ path is active given $C \cup E$, which contradicts $B \perp C'|C \cup E$. Hence $P$ is active given $(C \cup C')(\{B \cup \{v\}\}) \cup E$.

4 Directional Markov blanket

We write “a $(d)$-separator $S$” when we make statement that hold both in directed and undirected graphical models. Set $S$ is a then a $d$-separator in directed graphical models, and a separator in undirected graphical models.

Definition 2. Let $B,D$, and $E$ be three sets of vertices in a graph $G = (V,A)$. The Markov blanket of $B$ in the direction of $D$ given $E$, denoted $mb(B \rightarrow D|E)$, is the minimal $(d)$-separator $M$ of $B$ and $D$ such that

$$D \perp M \mid M' \cup E \quad \text{for any $(d)$-separator } M' \text{ between } B \text{ and } D \text{ given } E.$$  (4)
Figure 5: Example of the directional Markov Blanket from $B$ to $D$ given an evidence set $E$. In this case $\text{mb}(B \rightarrow D|E) = \{u, v\}$

Figure 5 shows an example of the Markov Blanket of $B$ in the direction $D$ given an evidence $E$. Note that in this definition, the evidence set $E$ can be such that $E \cap B \neq \emptyset$. The Markov blanket of $B$ in the direction of $D$ is the d-separator between $B$ and $D$ that is the nearest to $B$. Furthermore, the following proposition provides an alternative definition.

**Proposition 2.** Let $B, D,$ and $E$ be three sets of vertices in a graph $G = (V, A)$. Let $M$ be a minimal (d-)separator between $B$ and $D$ given $E$.

$M$ satisfies \((4)\) if, and only if:

\[ B \perp M' | M \cup E \quad \text{for any minimal (d-)separator } M' \text{ between } B \text{ and } D \text{ given } E. \quad (5) \]

**Proof of Proposition 2 in undirected graphical models.** Let $M$ be a minimal d-separator. We start by proving \((4)\) implies \((5)\). Let $M'$ be a minimal separator between $B$ and $D$ given $E$, and let $P$ be a path between $B$ and $x \in M'$, where $M'$ is minimal, there exists a path $Q$ from $x$ to $D$ such that $Q \cap (M' \cup E) \neq \emptyset$. The path $R$ composed of $P$ followed by $Q$ is a $B$-$D$ path. Since $M$ is a d-separator, there exists $v \in R \cap M$. If $v \in Q$, then \((4)\) implies that $Q \cap (M' \setminus \{x\}) \neq \emptyset$, which contradicts the assumption on $Q$. Therefore, $v \in P$. We deduce that all path from $B$ to $M'$ is intersected by $M \cup E$, which implies that $B \perp M' | M \cup E$.

Suppose now that \((5)\) holds. Let $Q$ be a path from $u \in M$ to $D$ and $M'$ be a separator between $B$ and $D$ given $E$. Since $M$ is minimal, there exists a path $P$ from $B$ to $u$ such that $P \cap (M \setminus \{u\}) = \emptyset$. The path $R$ composed of $P$ followed by $Q$ is a $B$-$D$ path, there exists $v \in R \cap (M' \cup E)$. Using the same arguments as above, $v \in Q$, which implies that $x \perp D | M' \cup E$.

The proof of Proposition 2 in directed graphical model is more involved and postponed to Section 4.3. Similarly to the Markov Blanket in a set, we need to prove that $\text{mb}(B \rightarrow D|E)$ in Definition 2 exists. The following theorem states the existence and uniqueness of the Directional Markov Blanket.

**Theorem 2.** Let $B, D,$ and $E$ be three sets of vertices in a graph $G = (V, A)$. If there exists a (d-)separator between $B$ and $D$ given $E$, the Markov blanket of $B$ in the direction of $D$ given $E$ exists, is unique, and is given by

\[
\text{mb}(B \rightarrow D|E) = \text{mb}_{\text{mb}(B|E)}(D|E) \quad \text{in undirected graphical models, and by}
\]

\[
\text{mb}(B \rightarrow D|E) = \text{mb}_{\text{mb}(B \cup D \cup E)(B|E)}(D|E) \quad \text{in directed graphical models.}
\]

The rest of the section is dedicated to the proofs of Proposition 2 in directed graphical models and of Theorem 2.
Remark 1. Using Definition 2, the Markov blanket of $B$ in the direction of $D$ given $E$ exists if and only if there exists a d-separator between $B$ and $D$ given $E$. We can extend the definition of the Markov blanket of $B$ in the direction of $D$ given $E$ as the set $M$ of $V \setminus B$ satisfying

(i) $B \perp D|M \cup E$,
(ii) $B \perp D|M' \cup E$ implies $D \perp M'|M \cup E$,
(iii) $B \perp D|M' \cup E$ and $D \perp M'|M \cup E$ implies $M \subseteq M'$.

It is immediate that the two definitions coincide when there exists a d-separator between $B$ and $D$. With this new definition, even without the existence of a d-separator, it follows from Theorem 4 in Section 4.2 that $mb(B \rightarrow D|E)$ exists and admits the following updated characterization

$$mb(B \rightarrow D|E) = \overline{D} \cup mb_{mb(B|E)}(D|E)$$

in undirected graphical models, and by

$$mb(B \rightarrow D|E) = \overline{D} \cup mb_{\overline{asc}(B \cup D \cup E)}(B|E)(\hat{D}|E)$$

in directed graphical models, where

$$\overline{D} = \begin{cases} 
D \cap mb(B|E) & \text{in undirected graphical models,} \\
D \cap mb_{\overline{asc}(B \cup D \cup E)}(B|E) & \text{in directed graphical models,}
\end{cases}$$

and $\hat{D} = D \setminus \overline{D}$.

\section{4.1 Preliminary lemmas in directed graphical models}

In this section we present some technical results on d-Separators in directed graphical models. In the remaining of this section, $B$, $D$ and $E$ denote three sets of vertices in a graph $G = (V, A)$.

Lemma 1. Let $M$ be a d-separator between $B$ and $D$ given $E$. Then any $B$-$D$ trail in $\overline{asc}(B \cup D \cup M \cup E)$ intersects $M \cup E$ in a vertex $x$ that is not a $v$-structure.

Proof. Let $P$ be a $B$-$D$ trail in $\overline{asc}(B \cup D \cup M \cup E)$. Starting from $B$, let $v$ be the last $v$-structure of $P$ that is not active given $M \cup E$ and that is in $asc(B)$, with $v$ being equal to the first vertex of $P$ if there is no such $v$-structure. Starting from $v$, let $w$ be equal to the first $v$-structure of the $v$-$d$ subpath of $P$ that is not active given $M \cup E$, and to the last vertex of $P$ if there is no such $v$-structure. By definition of $v$, vertex $w$ has necessarily a descendant in $D$. Taking a $B$-$w$ path followed by the $v$-$w$ subtrail of $P$ and then a $w$-$D$ path, we obtain an active trail given $M \cup E$, which gives a contradiction.

Lemma 2. Let $M$ be a d-separator between $B$ and $D$ given $E$, and $N \subseteq \overline{asc}(B \cup D \cup M \cup E)$. Then $M \cup N$ is a d-separator between $B$ and $D$ given $E$.

Proof. Suppose that there exists an active trail between $B$ and $D$ given $M \cup E \cup N$. Let $P$ be such a trail. Since $N \in \overline{asc}(B \cup D \cup M \cup E)$, we deduce that $P$ is a trail in $\overline{asc}(B \cup D \cup M \cup E)$ because all $v$-structures have a descendant in $M \cup E \cup N$ and $N \subseteq \overline{asc}(B \cup D \cup M \cup E)$. Lemma 1 ensures that $P$ intersects $M \cup E$ in a vertex that is not a $v$-structure. It contradicts the assumption on $P$.

The following lemma is an extension of Theorem 6 of Tian and Paz [1998] where we allow an evidence $E$.

Lemma 3. If $M$ is a d-separator between $B$ and $D$ given $E$, then $M \cap \overline{asc}(B \cup D \cup E)$ is also a d-separator between $B$ and $D$ given $E$. 

\end{document}
Proof. Any trail that intersects $V \setminus \operatorname{asc}(B \cup D \cup E)$ is not active given $(M \cap \operatorname{asc}(B \cup D \cup E)) \cup E$. And by Lemma 1, any trail in $\operatorname{asc}(B \cup D \cup E)$ intersects $(M \cap \operatorname{asc}(B \cup D \cup E)) \cup E$ on a non v-structure, which gives the result.

Corollary 1. Let $M$ be a set of vertices. Then there exists a subset of $M$ that $d$-separates $B$ and $D$ given $E$ if and only if

$$B \perp D | (M \cap \operatorname{asc}(B \cup D \cup E)) \cup E$$

Proof. An immediate corollary of the two previous lemmas.

Lemma 4. Let $M$ be a $d$-separator between $B$ and $D$ given $E$, and $x \in \operatorname{asc}(B \cup D \cup M \cup E)$. Then at least one of the following statement is true: $x \perp B | M \cup E$ or $x \perp D | M \cup E$.

Proof. Suppose that none of the independences are satisfied. Then $x \notin M$, and there is a $B$-$x$ trail $Q$ that is active given $M \cup E$, and an $x$-$D$ trail $R$ that is active given $M \cup E$. As $x \in \operatorname{asc}(B \cup D \cup M \cup E)$, if trails $Q$ and $R$ intersect $V \setminus \operatorname{asc}(B \cup D \cup M \cup E)$, they are not active given $M \cup E$. As $x \notin M \cup E$, the trail composed of $Q$ followed by $R$ is a $B$-$D$ trail that intersects $M \cup E$ only on v-structures. This contradicts Lemma 1 and gives the result.

4.2 Proof of Theorem 2

In this section we prove Theorem 2.

Lemma 5. Let $M$ be a $(d)$-separator between $B$ and $D$ given $E$, then $\operatorname{mb}_M(B | E)$ is a $(d)$-separator between $B$ and $D$ given $E$.

Proof of Lemma 5 in undirected graphical models. Consider a path $Q$ from $B$ to $D$. Since $B \perp D | M \cup E$, we have $Q \cap (M \cup E) \neq \emptyset$. Starting from $B$, consider the first vertex $x$ of $M \cup E$ on the path $Q$. By Theorem 1, $x \in \operatorname{mb}_M(B | E)$. It implies that $Q \cap \operatorname{mb}_M(B | E) \neq \emptyset$. We conclude that $B$ and $D$ are separated by $\operatorname{mb}_M(B | E) \cup E$.

Proof of Lemma 5 in directed graphical models. Suppose that $B \perp D | \operatorname{mb}_M(B | E) \cup E$. Let $P$ be a trail between $B$ and $D$ that is active given $\operatorname{mb}_M(B | E) \cup E$. Since $\operatorname{mb}_M(B | E) \cup E \subseteq M \cup E$, all the v-structures of $P$ are active given $M \cup E$. Since $P$ is not active given $M \cup E$, there exists at least one element in $(M \cup E) \cap P$, which is not in a v-structure of $P$. Starting from $B$, consider the first element $x$ on $P$ such that $x \in (M \setminus \{x\}) \cup E$. The subtrail of $P$ from $B$ to $x$ is active given $(M \setminus \{x\}) \cup E$. Therefore, $x \in \operatorname{mb}_M(B | E)$, which contradicts our assumption on $P$.

Corollary 2. Let $M$ be a minimal $(d)$-separator between $B$ and $D$, then $\operatorname{mb}_M(B | E) = M$.

Proof. Lemma 5 ensures that $\operatorname{mb}_M(B | E)$ is a $d$-separator (resp. separator) between $B$ and $D$ given $E$. Since $\operatorname{mb}_M(B | E) \subseteq M$ and $M$ is minimal, we deduce that $\operatorname{mb}_M(B | E) = M$.

Lemma 6. Let $B$ and $D$ given $E$ be three sets of vertices of an undirected graphical model (resp. directed graphical model) $G = (V, E)$. Let $M$ be a separator between $B$ and $D$ given $E$ (resp. a $d$-separator between $B$ and $D$ given $E$ in $\operatorname{asc}(B \cup D \cup E)$). Then $\operatorname{mb}_M(B | E)$ is a $(d)$-separator between $B$ and $D$ given $E$, and $\operatorname{mb}_{\operatorname{mb}_M(B | E)}(D | E)$ is a minimal $(d)$-separator between $B$ and $D$ given $E$.

Proof of Lemma 6 in undirected graphical models. Let $M' = \operatorname{mb}_M(B | E)$ and $M''$ be equal to $\operatorname{mb}_{\operatorname{mb}_M(B | E)}(D | E)$. Lemma 5 ensures that $M'$ and $M''$ are separators between $B$ and $D$ given $E$. We prove that $M''$ is minimal. Let $v \in M''$. There exists a path $P$ from $B$ to $v$ such that $P \cap (M \cup E) \setminus \{x\} = \emptyset$ and there exists a path $Q$ from $v$ to $D$ such that $Q \cap (M' \cup E) \setminus \{x\} = \emptyset$. Consider the path $R$ composed of $P$ followed by $Q$. Then $R$ is a $B$-$D$ path with $R \cap (M'' \cup E \setminus \{v\}) = \emptyset$. We deduce that $R$ is not separated by $M'' \cup \{v\} \cup E$, which implies that $M'' \setminus \{v\}$ is not a separator given $E$. It achieves the proof.
Proof of Lemma 2 in directed graphical models. Let \( M' = \text{mb}_M(B|E) \) and \( M'' \) be defined as \( \text{mb}_{\text{mb}_M(B|E)}(D|E) \). We prove that \( M'' = \text{mb}_{\text{mb}_M(B|E)}(D|E) \) is a minimal d-separator. Lemma 3 ensures that \( M' \) and \( M'' \) are d-separators between \( B \) and \( D \) given \( E \). Let \( v \) be a vertex in \( M'' \). Let \( Q \) be a \( B-v \) trail active given \( M \cup E \setminus \{v\} \), and \( R \) be a \( v-D \) trail active given \( M' \cup E \setminus \{v\} \), and \( P \) the trail composed of \( R \) followed by \( Q \). Then \( P \) is a \( B-D \) trail in \( \sec(B \cup D \cup E) \) that intersects \( M'' \cup E \setminus \{v\} \) only on v-structures. Hence, Lemma 1 ensures that \( M'' \setminus \{v\} \) is not a d-separator, and Corollary 1 enables to conclude that \( M'' \) is a minimal d-separator.

The following theorem is a stronger version of Theorem 3.

**Theorem 3.** Let \( B \) and \( D \) given \( E \) be three sets of vertices of an undirected graphical model (resp. directed graphical model) \( G = (V,E) \). Let \( M \) be a separator between \( B \) and \( D \) given \( E \) (resp. a d-separator between \( B \) and \( D \) given \( E \) in \( \sec(B \cup D \cup E) \)). Then \( M_1 = \text{mb}_{\text{mb}_M(B|E)}(D|E) \) is the unique minimal (d-)separator between \( B \) and \( D \) given \( E \) such that \( M_1 \perp D|E \) for any (d-)separator \( M_2 \) in \( M \).

**Proof of uniqueness in Theorem 3.** Suppose that \( M_1 \) and \( M_1' \) are minimal (d-)separator between \( B \) and \( D \) given \( E \) such that \( M_1 \perp D|E \cup E \) for any (d-)separator \( M_2 \) in \( M \). Then \( M_1' \perp D|E \cup E \cup \{v\} \) for any (d-)separator \( M_2 \) in \( M \). Then \( M_1' \perp D|M_1 \cup \{v\} \) since \( M_1 \) is a minimal d-separator. We prove that \( M_1 \setminus \{v\} \) is a minimal d-separator between \( B \) and \( D \) given \( E \). Let \( M_2 \subseteq M \) be a separator between \( B \) and \( D \) given \( E \). We prove that \( M_1 \setminus \{v\} \subseteq M_2 \cup E \). There exists an active path between \( v \in M_1 \) and \( E \cup \{v\} \). Since \( v \in \text{mb}_M(B|E) \), there exists an active path between \( B \) and \( v \) given \( M \cup E \setminus \{v\} \). Let \( P \) be such a path. Therefore we have \( P \cap (M \cup E \setminus \{v\}) = \emptyset \). Let \( R \) be the path composed of \( P \) followed by \( Q \). \( R \) is a \( B-D \) path and \( R \cap (M_2 \cup E \setminus \{v\}) = \emptyset \), which contradicts the assumption on \( M_2 \).

**Proof of Theorem 3 in directed graphical models.** Lemma 1 ensures that \( M_1 \) is a minimal d-separator between \( B \) and \( D \) given \( E \). Let \( M_2 \subseteq M \) be a d-separator between \( B \) and \( D \) given \( E \). We prove that \( M_1 \cup M_2 \subseteq M \cup E \setminus \{v\} \). There exists an active path between \( v \in M_1 \) and \( D \cup \{v\} \). Let \( Q \) be such a path. Since \( v \in \text{mb}_M(B|E) \), there exists an active path between \( B \) and \( v \) given \( M \cup E \setminus \{v\} \). Let \( R \) be the trail composed of \( P \) followed by \( Q \). \( R \) is a \( D-B \) path and \( R \cap (M_2 \cup E \setminus \{v\}) = \emptyset \), which contradicts the assumption on \( M_2 \).

**4.3. Proof of Proposition 2 in directed graphical models**

The two following lemmas are intermediary technical results for the proof of the alternative definition of the directional Markov Blanket in directed graphical models in Proposition 2.

**Lemma 7.** Let \( M \) be a minimal d-separator between \( B \) and \( D \) given \( E \). Let \( N \subseteq \sec(B \cup D \cup E) \). Let \( L = \text{mb}_{\text{mb}_M}(B|E), \text{and } O = \text{mb}_L(D|E). \) Then \( L \cap M = O \cap M \).

**Proof.** Remark that \( M \subseteq \sec(B \cup D \cup E) \) because \( M \) is a minimal d-separator between \( B \) and \( D \) given \( E \). Inclusion \( O \subseteq L \) gives \( O \cap M \subseteq L \cap M \). Suppose that \( O \cap M \neq L \cap M \). Since \( L \cap M \) contains strictly \( O \cap M \), it ensures the existence of \( x \) in \( (L \cap M) \setminus O \). By definition of \( L \) there exists a \( B-x \) trail \( Q \) in \( \sec(B \cup D \cup E) \) that is active given \( (L \setminus \{x\}) \cup E \). Since \( O \subseteq L \), any vertex of \( Q \) in \( O \cup E \) is a v-structure. As \( M \) is minimal there is a \( x-D \) trail \( R \) that is active given \( M \cup E \). Since \( Q \) followed by \( R \) is a \( B-D \) trail in \( \sec(B \cup D \cup O \cup E) \), and \( Q \) does not intersect \( O \cup E \) on a vertex which is not a v-structure, by Lemma 1 there is a non v-structure of \( R \) in
O. Starting from $x$, let $y$ be the last such vertex. Let $T$ be the $y$-$D$ subtrail of $R$. Note that $R$ can intersect $M \cup E$ only on v-structures, and hence $y \notin M$ and $T$ can intersect $M$ only on v-structures. As $y \in L = \text{mb}_{M \cup N}(B|E)$, there is a $B$-$y$ trail $S$ in $\text{asc}(B \cup D \cup E)$ that intersects $M$ only on v-structures. Hence, $S$ followed by $T$ is a $B$-$D$ trail in $\text{asc}(B \cup D \cup E \cup M)$ that intersects $M \cup E$ only on v-structures, and Lemma \ref{lem:intersect} gives a contradiction.

**Lemma 8.** Let $M$ and $N$ be two d-separators between $B$ and $D$ given $E$. If $N$ is minimal and $M \perp D|N \cup E$, then

$$B \perp N|M \cup E \quad (6)$$

**Proof.** Suppose that $B \perp N|M \cup E$. Let $x$ be a vertex of $N \setminus M$ that is not d-separated from $B$ given $M \cup E$, and $Q$ be a $B$-$x$ trail that is active given $M \cup E$. As $N$ is minimal, $N \subseteq \text{asc}(B \cup D \cup E)$ and there is an $x$-$D$ trail $R$ that is active given $N \cup E$. This trail does not intersect $M$ as this would contradict $M \perp D|N \cup E$. Hence $Q$ followed by $R$ is a $B$-$D$ trail in $\text{asc}(B \cup D \cup E \cup M)$ that intersects $M \cup E$ only on v-structures, which gives a contradiction.

**Proof of Proposition \ref{prop:intersect} in directed graphical models.** Let $M$ be a minimal d-separator between $B$ and $D$ given $E$.

We start by proving “not (3)” implies “not (1)”. Suppose that there exists a minimal d-separator $M'$ such that $B \perp M'|M \cup E$. Since $M'$ is minimal, Lemma \ref{lem:intersect} ensures that $D \perp M|M \cup E$. There exists a d-separator $M'$ such that $D \perp M'|M \cup E$.

We now prove “not (1)” implies “not (3)”. Let $M$ be a minimal d-separator, and $M'$ be a d-separator such that $D \perp M'|M \cup E$. Let $M'' = M' \cap \text{asc}(B \cup D \cup E)$. Let $M_1 = \text{mb}_{\text{mb}_{M''}}(B|E)(D|E)$. Lemma \ref{lem:intersect} ensures that $B \perp D|M'' \cup E$. Since $M'' \subseteq \text{asc}(B \cup D \cup E)$, Lemma \ref{lem:intersect} ensures that $M \cap M''$ is a d-separator between $B$ and $D$ given $E$. Hence Lemma \ref{lem:intersect} ensures that $M_1$ is a minimal d-separator between $B$ and $D$ given $E$. To prove “not (3)”, we prove $B \perp M_1|M \cup E$.

Let $x$ be a vertex of $M$ such that $x \perp D|M'' \cup E$. We start by proving $x \perp B|M'' \cup E$. Let $Q$ be a $B$-$x$ trail. We prove that $Q$ is not active given $M'' \cup E$. Since, $x \in \text{asc}(B \cup D \cup E)$ and $B \subseteq \text{asc}(B \cup D \cup E)$, if $Q$ intersects $V \setminus \text{asc}(B \cup D \cup E)$, then it contains a v-structure in $V \setminus \text{asc}(B \cup D \cup E)$ which cannot be active given $M'' \cup E$ because $M'' \subseteq \text{asc}(B \cup D \cup E)$. Suppose now that $Q$ is in $\text{asc}(B \cup D \cup E)$, and let $R$ be an $x$-$D$ trail that is active given $M'' \cup E$. As $M''$ d-separates $B$ and $D$ given $E$, Lemma \ref{lem:intersect} ensures that $Q$ followed by $R$ intersects $M''$ on a non-v-structure. This intersection is necessarily in $Q$ and in $M''$. Hence $Q$ is not active given $M'' \cup E$. And we have proved $x \perp B|M'' \cup E$.

We now prove that $x$ does not belong to $M_1$. By Lemma \ref{lem:intersect}, it suffices to prove that $x$ does not belong to $\text{mb}_{M \cup M''}(B|E)$. Suppose that there is a $B$-$x$ trail active given $(M \cup M'' \setminus \{x\}) \cup E$. Let $P$ be such a trail with a minimal number of v-structure. Remark that $P$ is in $\text{asc}(B \cup D \cup E)$. Let $b_0$ be the first vertex of the trail starting from $B$. Let $s_1, \ldots, s_k$ be the v-structure of $P$ that have no descendants in $M'' \cup E$. We prove recursively that $s_i$ has a descendant $b_i$ in $B$. Indeed, $s_i$ has either a descendant in $B$ or in $D$. By iteration hypothesis, it cannot have a descendant in $D$ as otherwise we would have a $b_{i-1}$-$D$ trail that is active given $M'' \cup E$. Hence it has a descendant $b_i$ in $B$, with gives the iteration hypothesis. Hence there is a $b_k$-$x$ path that is active given $M'' \cup E$, which gives a contradiction.

The set $M \setminus M_1$ contains $x$ and is therefore non-empty. Theorem \ref{thm:intersect} ensures that $M_1$ satisfies $M_1 \perp D|M \cup E$. Since $M$ is a minimal d-separator between $B$ and $D$ given $E$, Lemma \ref{lem:intersect} ensures that $B \perp M|M_1 \cup E$. Proposition \ref{prop:intersect} ensures that $\text{mb}_{M \cup M_1}(B|E) = \text{mb}_{M_1}(B|E)$. As $M_1$ is a minimal d-separator between $B$ and $D$ given $E$, Corollary \ref{cor:intersect} ensures that $\text{mb}_{M_1}(B|E) = M_1$. We deduce that $\text{mb}_{M \cup M_1}(B|E) = M_1$. We therefore cannot have $M_1 \perp B|M \cup E$, as this would imply $M_1 = \text{mb}_{M \cup M_1}(B|E) = \text{mb}_{M}(B|E) = M$, which gives “not (3).”
References

Dan Geiger, Thomas Verma, and Judea Pearl. d-separation: From theorems to algorithms. In Proceedings of the Fifth Annual Conference on Uncertainty in Artificial Intelligence, UAI ’89, pages 139–148, Amsterdam, The Netherlands, The Netherlands, 1990. North-Holland Publishing Co.

Ron Kohavi and George H. John. Wrappers for feature subset selection. Artif. Intell., 97(1-2):273–324, December 1997.

Daphne Koller and Nir Friedman. Probabilistic graphical models: principles and techniques. MIT press, 2009.

S. L. Lauritzen, A. P. Dawid, B. N. Larsen, and H.-G. Leimer. Independence properties of directed markov fields. Networks, 20(5):491–505, 1990.

Steffen L. Lauritzen. Causal inference from graphical models. In In Complex Stochastic Systems, pages 63–107. Chapman and Hall/CRC Press, 1999.

Judea Pearl. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1988.

Judea Pearl. Graphical models, causality, and intervention, 1993.

Jean-Philippe Pellet and André Elisseeff. Using markov blankets for causal structure learning. J. Mach. Learn. Res., 9:1295–1342, June 2008.

Ilya Shpitser and Judea Pearl. Identification of conditional interventional distributions. CoRR, abs/1206.6876, 2012.

P. Spirtes, C. Glymour, and R. Scheines. Causation, Prediction, and Search. MIT press, 2nd edition, 2000.

Jin Tian and Azaria Paz. Finding minimal d-separators, 1998.