Reading Dependencies from Polytree-Like Bayesian Networks Revisited

Jose M. Peña
ADIT, IDA, Linköping University, Sweden

Abstract

We present a graphical criterion for reading dependencies from the minimal directed independence map G of a graphoid p, under the assumption that G is a polytree and p satisfies weak transitivity. We prove that the criterion is sound and complete. We argue that assuming weak transitivity is not too restrictive.

1 Introduction

A minimal directed independence map G of an independence model p is typically used to read independencies holding in p. However, G can also be used to read dependencies holding in p. For instance, if p is a graphoid that is faithful to G, then lack of vertex separation is a sound and complete graphical criterion for reading dependencies from G. If p is simply a graphoid, then there also exists a sound and complete graphical criterion for reading dependencies from G (Bouckaert, 1995). In (Peña, 2007), we present a further sound and complete graphical criterion for reading dependencies from G under the assumption that G is a polytree and p is a graphoid that satisfies composition and weak transitivity. In this paper, we revisit the latter work and drop the assumption that p satisfies composition. In general, the more assumptions a criterion makes about G and p the more powerful it is (i.e. the more dependencies it can read from G) but the less applicable it is (i.e. the smaller the set of independence models it can be applied to). Then, our new criterion may be seen as being in between the criteria in (Bouckaert, 1995) and (Peña, 2007): It is more (resp. less) powerful but less (resp. more) applicable than the former (resp. latter) criterion. See Section 5 for an example.

The rest of the paper is organized as follows. Section 2 is devoted to the preliminaries, Section 3 to our assumptions, Section 4 to our contribution, and Section 5 to the discussion.

2 Preliminaries

Let U denote a set of random variables. The elements of U are not distinguished from singletons, and the union of the sets $U_1, \ldots, U_n \subseteq U$ is written as the juxtaposition $U_1 \ldots U_n$. When evaluating an expression, the union of sets precedes the set difference. Let X, Y, Z and W denote four mutually disjoint subsets of U. An independence model p is a set of independencies of the form X is independent of Y given Z. We denote that an independence is in p by $X \perp_p Y|Z$ and that an independence is not in p by $X \not\perp_p Y|Z$. In the latter case, we say that the dependence $X \not\perp_p Y|Z$ is in p. An independence model is a graphoid if it satisfies the following properties: Symmetry $X \perp_p Y|Z \Rightarrow Y \perp_p X|Z$, decomposition $X \perp_p YW|Z \Rightarrow X \perp_p Y|Z$, weak union $X \perp_p YW|Z \Rightarrow X \perp_p Y|ZW$, contraction $X \perp_p Y|ZW \wedge X \perp_p W|Z \Rightarrow X \perp_p YW|Z$, and intersection $X \perp_p Y|ZW \wedge X \perp_p W|ZY \Rightarrow X \perp_p YW|Z$.

We say that a node C is a collider in a route in a directed and acyclic graph (DAG) if $A \rightarrow C \leftarrow B$ is a subroute of the route. Note that A and B may coincide since we are dealing with a route and not with a path. A route in a DAG is said to be superactive wrt Z when (i) every collider node in the route is in Z, and (ii) every non-collider node in the route is outside Z. When there is no route in a DAG G between a node in X and a node in Y that is superactive wrt Z, we say that X is separated from Y given Z in G and denote it as $X \perp_G Y|Z$.

This definition of separation in DAGs is equivalent to other more common definitions (Studený, 1998). Given an undirected graph (UG) G, we say that X is separated from Y given Z in G and denote it as $X \perp_G Y|Z$ when every path in G between a node in X and a node in Y contains a node in Z. An independence model p is faithful to an UG or DAG G when $X \perp_p Y|Z$ iff $X \perp_G Y|Z$. A DAG G is a directed independence map of an independence model p when $X \perp_p Y|Z$ if $X \perp_G Y|Z$. Moreover, G is a minimal directed independence (MDI) map of p when removing any edge from G makes it cease to be an independence map of p. If G is a MDI map of p, then the parents of a node A in G, $Pa(A)$, are the smallest subset of the nodes preceding A in a given total ordering of U, $Pre(A)$, such that $A \perp_p Pre(A) \setminus Pa(A)|Pa(A)$. We denote the children of A in G by $Ch(A)$. Finally, recall that a polytree is a directed graph without undirected cycles.

3 WT Graphoids

Let X, Y and Z denote three mutually disjoint subsets of U. Let $V \in U \setminus XYZ$. We call WT graphoid to any graphoid p that satisfies weak transitivity $X \perp_p Y|Z \wedge X \perp_p Y|ZV \Rightarrow X \perp_p V|Z \vee V \perp_p Y|Z$. This paper studies WT graphoids. We regard WT graphoids as worth studying because important families of probability distributions are WT graphoids. For instance, any probability distribution that is Gaussian or faithful to some UG or DAG is a WT graphoid (Pearl, 1988). The following theorem implies that there also exist probability distributions that are WT graphoids although they are neither Gaussian nor faithful to any UG or DAG. See (Peña et al., 2009) for the proof and examples.

Theorem 1. *Let p be a probability distribution that is a WT graphoid and let $W \subseteq U$. Then, $p(U \setminus W)$ is a WT graphoid. If $p(U \setminus W|W = w)$ has the same independencies for all value w of W, then $p(U \setminus W|W = w)$ for any w is a WT graphoid.*

The following theorem introduces a new property that every WT graphoid satisfies.

Theorem 2. *Let p be a WT graphoid. Then, p satisfies the following property: Intersectional weak transitivity $X \perp_p Y|Z \wedge X \perp_p Y|ZV \Rightarrow X \perp_p V|ZY \vee V \perp_p Y|ZX$.*

Proof. Assume to the contrary that $X \not\perp_p V|ZY$ and $V \not\perp_p Y|ZX$. Then,

1. $X \not\perp_p VY|Z$ and $VX \not\perp_p Y|Z$ by the contrapositive form of weak union on $X \not\perp_p V|ZY$ and $V \not\perp_p Y|ZX$

2. $X \not\perp_p V|Z$ and $V \not\perp_p Y|Z$ by the contrapositive form of contraction on (1) and $X \perp_p Y|ZV$

3. $X \not\perp_p Y|ZV$ by the contrapositive form of weak transitivity on (2) and $X \perp_p Y|Z$.

However, (3) contradicts the antecedent of the property. □

4 Reading Dependencies

If G is a MDI map of a WT graphoid p then we know, by construction of G, that $A(Pre(B) \setminus Pa(B)) \not\perp_p B|Pa(B) \setminus A$ for all the edges $A \to B$ in G. We call these dependencies the dependence base of p for G. Further dependencies in p can be derived from the dependence base via the WT graphoid properties. For this purpose, we rephrase the WT graphoid properties in their contrapositive form as follows. Symmetry $Y \not\perp_p X|Z \Rightarrow X \not\perp_p Y|Z$. Decomposition $X \not\perp_p Y|Z \Rightarrow X \not\perp_p YW|Z$. Weak union $X \not\perp_p Y|ZW \Rightarrow X \not\perp_p YW|Z$. Contraction $X \not\perp_p YW|Z \Rightarrow X \not\perp_p Y|ZW \vee X \not\perp_p W|Z$ is problematic for deriving new dependencies because it contains a disjunction in the consequent and, thus, we split it into two properties: Contraction1 $X \not\perp_p YW|Z \wedge X \perp_p Y|ZW \Rightarrow X \not\perp_p W|Z$, and contraction2 $X \not\perp_p YW|Z \wedge X \perp_p W|Z \Rightarrow X \not\perp_p Y|ZW$. Likewise, intersection gives rise to intersection1 $X \not\perp_p YW|Z \wedge X \perp_p Y|ZW \Rightarrow X \not\perp_p W|ZY$, and intersection2 $X \not\perp_p YW|Z \wedge X \perp_p W|ZY \Rightarrow X \not\perp_p Y|ZW$. Note that intersection1 and intersection2 are equivalent and, thus, we refer to them simply as intersection. Similarly, weak transitivity gives rise to weak transitivity1 $X \not\perp_p V|Z \wedge V \not\perp_p Y|Z \wedge X \perp_p Y|Z \Rightarrow X \not\perp_p Y|ZV$, and weak transitivity2 $X \not\perp_p V|Z \wedge V \not\perp_p Y|Z \wedge X \perp_p Y|ZV \Rightarrow X \not\perp_p Y|Z$. Finally, intersectional weak transitivity gives rise to intersectional weak transitivity1 $X \not\perp_p V|ZY \wedge V \not\perp$

$_pY|ZX \wedge X \perp_p Y|Z \Rightarrow X \not\perp_p Y|ZV$, and intersectional weak transitivity2 $X \not\perp_p V|ZY \wedge V \not\perp_p Y|ZX \wedge X \perp_p Y|ZV \Rightarrow X \not\perp_p Y|Z$. The independence in the antecedent of any of the properties above holds if the corresponding separation statement holds in G. This is the best solution we can hope for because separation is sound and complete. Separation is sound in the sense that it only identifies independencies in p. Moreover, separation is complete in the sense that it identifies all the independencies in p that can be identified by studying G alone (Peña, 2007). We call the WT (resp. IWT) graphoid closure of the dependence base of p for G to the set of dependencies that are in the dependence base of p for G plus those that can be derived from it by applying the first eight (resp. all the ten) properties above. The following example shows that the WT and IWT graphoid closures of a dependence base do not coincide in general.

Example 1. Let p be a probability distribution over $U = \{A, B, C\}$ where A, B and C are binary random variables. Let $p(A, B)$ be uniform and $C = XOR(A, B)$. Note that $A \perp_p B$, $A \perp_p C$ and $B \perp_p C$ are the only independencies in p. Then, p is a WT graphoid. Let G denote the DAG $A \to C \leftarrow B$. Note that G is a MDI map of p. Now, note that $A \not\perp_p B|C$ is in the IWT graphoid closure of the dependence base of p for G: The dependence base of p for G is $\{A \not\perp_p C|B, B \not\perp_p C|A\}$, which implies $A \not\perp_p B|C$ by intersectional weak transitivity1 and $A \perp_G B$. However, the WT graphoid closure of the dependence base of p for G is $\{A \not\perp_p C|B, B \not\perp_p C|A, A \not\perp_p BC, AB \not\perp_p C, B \not\perp_p AC, C \not\perp_p A|B, C \not\perp_p B|A, BC \not\perp_p A, C \not\perp_p AB, AC \not\perp_p B\}$ which does not contain $A \not\perp_p B|C$.

Hereinafter, we use $A : B$ to denote a route between two nodes A and B in a DAG G. We also use $A : B$ to denote the nodes in the route. It should be clear from the context which of the two meanings is being used. We define the parents of a route $A : B$ as $Pa(A : B) = [\cup_{C \to D \in A:B} Pa(D)] \setminus (A : B)$. We say that a route $A : B$ is minimally superactive wrt X, Y and Z in G if (i) $A \in X$ and $B \in Y$, (ii) $A : B$ is superactive wrt Z, and (iii) no proper subroute of $A : B$ is minimally superactive wrt X, Y and Z in G. Finally, we introduce our graphical criterion for reading dependencies from a polytree-like MDI map of a WT graphoid.

Definition 1. Let G be a polytree. Let X, Y and Z denote three mutually disjoint subsets of U. We say that $X \sim_G Y|Z$ holds if

- there exist two nodes $A \in X$ and $B \in Y$ and a single route $A : B$ between them that is minimally superactive wrt X, Y and Z in G, and
- for all $A' \in Pa(A : B)$, $A' \in XYZ \setminus AB$ or $A' \sim_{G_{\setminus A'}} XYZ \setminus AB$ where $G_{\setminus A'}$ is the DAG resulting from removing from G the edge between A' and its child in $A : B$.

Note the recursive flavor of the definition above: A base case ($A' \in XYZ \setminus AB$) and a recursive call ($A' \sim_{G_{\setminus A'}} XYZ \setminus AB$). The next theorem proves that the criterion in Definition 1 is sound. We prove first some auxiliary lemmas.

Lemma 1. *Let G be a polytree-like MDI map of a WT graphoid p. Let A and B be two nodes such that $A \sim_G B|Z$ holds due to a route $A : B$ with no collider node. Let $Pa(A : B) \subseteq Z$. Then, $A \not\perp_p B|Z$.*

Proof. We prove the lemma by induction over the length of $A : B$. We first prove the lemma for length one, i.e. $A : B$ is $A \to B$ or $A \leftarrow B$. Assume without loss of generality that $A : B$ is $A \to B$. Let Z^A denote the nodes in Z that are in $Pa(A)$ or connected to A by an undirected path that passes through $Pa(A)$. Let Z_A denote the nodes in Z that are in $Ch(A) \setminus B$ or connected to A by an undirected path that passes through $Ch(A) \setminus B$. Let Z^B denote the nodes in Z that are in $Pa(B) \setminus A$ or connected to B by an undirected path that passes through $Pa(B) \setminus A$. Note that $Pa(B) \setminus A \subseteq Z^B$ because we have assumed that $Pa(A : B) \subseteq Z$. Then,

1. $A(Pre(B) \setminus Pa(B)) \not\perp_p B|Pa(B) \setminus A$ from the dependence base of p for G
2. $A \not\perp_p B|Pa(B) \setminus A$ by contraction1 on (1) and $Pre(B) \setminus Pa(B) \perp_G B|(Pa(B) \setminus A)A$
3. $AZ^A Z_A \not\perp_p B|Pa(B) \setminus A$ by decomposition on (2)
4. $A \not\perp_p B|(Pa(B) \setminus A)Z^A Z_A$ by intersection on (3) and $Z^A Z_A \perp_G B|(Pa(B) \setminus A)A$

5. $A \not\perp_p B(Z^B \setminus (Pa(B) \setminus A))|(Pa(B) \setminus A)Z^A Z_A$ by decomposition on (4)

6. $A \not\perp_p B|Z^A Z_A Z^B$ by contraction2 on (5) and $A \perp_G Z^B \setminus (Pa(B) \setminus A)|(Pa(B) \setminus A)Z^A Z_A$

7. $A \not\perp_p B(Z \setminus Z^A Z_A Z^B)|Z^A Z_A Z^B$ by decomposition on (6)

8. $A \not\perp_p B|Z$ by intersection on (7) and $A \perp_G Z \setminus Z^A Z_A Z^B|Z^A Z_A Z^B B$.

Assume as induction hypothesis that the lemma holds when the length of $A : B$ is smaller than l. We now prove the lemma for length l. Let C be any node in $A : B$ except A and B. Recall that $A : B$ has no collider node. Thus, $C \notin Z$ because $A : B$ is minimally superactive wrt A, B and Z in G. Moreover, $C \notin Z$ implies $A \perp_G B|ZC$, $A \sim_G C|Z$, and $B \sim_G C|Z$. The latter two statements imply $A \not\perp_p C|Z$ and $B \not\perp_p C|Z$ by the induction hypothesis, which together with $A \perp_G B|ZC$ imply $A \not\perp_p B|Z$ by weak transitivity2. □

Lemma 2. *Let G be a polytree-like MDI map of a WT graphoid p. Let A and B be two nodes such that $A \sim_G B|Z$ holds due to a route $A : B$ of the form $A \to C \to \ldots \to D \leftarrow \ldots \leftarrow C \leftarrow B$ with possibly $C = D$. Let $Pa(A : B) \subseteq Z$. Then, $A \not\perp_p B|Z$.*

Proof. Let Z_D be the descendants of D that are in Z. Note that $A \perp_G B|Z \setminus Z_D D$ as $A : B$ is the only route between A and B that is minimally superactive wrt A, B and Z in G. Then,

1. $A \not\perp_p D|(Z \setminus Z_D D)B$ by Lemma 1
2. $B \not\perp_p D|(Z \setminus Z_D D)A$ by Lemma 1
3. $A \not\perp_p B|(Z \setminus Z_D D)D$ by intersectional weak transitivity1 on (1), (2), and $A \perp_G B|Z \setminus Z_D D$
4. $A \not\perp_p B Z_D|(Z \setminus Z_D D)D$ by decomposition on (3)
5. $A \not\perp_p B|Z$ by contraction2 on (4) and $A \perp_G Z_D|(Z \setminus Z_D D)D$. □

Lemma 3. *Let G be a polytree-like MDI map of a WT graphoid p. Let $X \sim_G Y|Z$ hold due to a route $A : B$ with $A \in X$ and $B \in Y$. Let $Pa(A : B) \subseteq XYZ \setminus AB$. Then, $X \not\perp_p Y|Z$.*

Proof. Let $W = XYZ \setminus AB$. Let $D_1, \ldots, D_n$ be the collider nodes in $A : B$. Then, for all i, $A : B$ has a subroute of the form $A_i \to C_i \to \ldots \to D_i \leftarrow \ldots \leftarrow C_i \leftarrow B_i$ with $A_i \neq B_i$ but possibly $C_i = D_i$. Let W_{D_i} denote the descendants of D_i that are in W. Let W_{C_i} denote the descendants of C_i that are in $W \setminus W_{D_i} D_i$. Let $W' = \cup_i W_{C_i}$. We first prove $A \not\perp_p B|W \setminus W'$. Note that $Pa(A : B) \subseteq W \setminus W'$ and, thus, that $A \sim_G B|W \setminus W'$ holds due to $A : B$. Then, we can divide $A : B$ into subroutes such that each of them is of the form of the route in Lemma 1 or 2. We prove $A \not\perp_p B|W \setminus W'$ by induction over the number of such subroutes. If the number of subroutes is one, then the result is immediate by Lemma 1 or 2. Assume as induction hypothesis that the result holds when the number of subroutes is smaller than l. We now prove the result when this number is l. Let E be any node in $A : B$ where two of the subroutes meet. Note that E is a non-collider node in $A : B$ and, thus, $E \notin W \setminus W'$. Moreover, $E \notin W \setminus W'$ implies $A \perp_G B|(W \setminus W')E$, $A \sim_G E|W \setminus W'$, and $E \sim_G B|W \setminus W'$. Then,

1. $A \not\perp_p E|W \setminus W'$ and $E \not\perp_p B|W \setminus W'$ by $A \sim_G E|W \setminus W'$, $E \sim_G B|W \setminus W'$ and the induction hypothesis
2. $A \not\perp_p B|W \setminus W'$ by weak transitivity2 on (1) and $A \perp_G B|(W \setminus W')E$.

Finally, let $X' \subseteq X \setminus A$ and $Y' \subseteq Y \setminus B$ contain the nodes in W' that are not descendant of another node in X' or Y'. Note that X' and Y' must exist for $A : B$ to be the only route between A and B that is minimally superactive wrt X, Y and Z in G. Let $W_{X'}$ (resp. $W_{Y'}$) contain the descendants of X' (resp. Y') that are in W'. Then,

3. $AX'W_{X'} \not\perp_p B|W \setminus W'$ by decomposition on (2)
4. $AX' \not\perp_p B|(W \setminus W')W_{X'}$ by intersection on (3) and $W_{X'} \perp_G B|(W \setminus W')AX'$
5. $AX' \not\perp_p BY'W_{Y'}|(W \setminus W')W_{X'}$ by decomposition on (4)
6. $AX' \not\perp_p BY'|(W \setminus W')W_{X'}W_{Y'}$ by intersection on (5) and $AX' \perp_G W_{Y'}|(W \setminus W')W_{X'}BY'$
7. $X \not\perp_p Y|Z$ by decomposition and weak union on (6). □

If a route $A : B$ in a DAG has a subroute of the form $C \to D \to \ldots \to E \leftarrow \ldots \leftarrow D \leftarrow F$ with $C \neq F$ but possibly $D = E$, the subroute $D \to \ldots \to E \leftarrow \ldots \leftarrow D$ is called a rope.

Theorem 3. *Let G be a polytree-like MDI map of a WT graphoid p. If $X \sim_G Y|Z$, then $X \not\perp_p Y|Z$ is in the IWT graphoid closure of the dependence base of p for G.*

Proof. Let $X \sim_G Y|Z$ hold due to a route $A : B$ with $A \in X$ and $B \in Y$. Let $W = XYZ \setminus AB$. We prove the theorem by induction over the total number of recursive calls performed by $X \sim_G Y|Z$. If this number is zero, then the theorem is immediate by Lemma 3. Assume as induction hypothesis that the theorem holds when the total number of recursive calls is smaller than l. We now prove the theorem when this number is l. Let $A' \sim_{G_{\setminus A'}} W$ with $A' \in Pa(E)$ for some $E \in A : B$ be any recursive call performed by $X \sim_G Y|Z$. Let $A' \sim_{G_{\setminus A'}} W$ hold due to a route $A' : B'$ with $B' \in W$. We consider two scenarios. The first scenario is when E is outside every rope in $A : B$. Then, $A : B$ must have a subroute of the form $C \rightarrow E$ or $E \leftarrow D$ for the recursive call $A' \sim_{G_{\setminus A'}} W$ to be performed. Assume without loss of generality that the subroute is of the form $C \rightarrow E$. Let W_E denote the descendants of E that are in W. Note that $X \sim_G Y|Z$ implies that $A \sim_G A'|(W \setminus W_E B')E$ holds. To see it, note that the subroute of $A : B$ between A and E followed by $E \leftarrow A'$, here denoted $A : A'$, is the only route between A and A' that is minimally superactive wrt A, A' and $(W \setminus W_E B')E$ in G. Moreover, every recursive call that $A \sim_G A'|(W \setminus W_E B')E$ performs is of the form $A'' \sim_{G_{\setminus A''}} (W \setminus W_E B')E$ with $A'' \in Pa(A : A')$. This recursive call holds because $A'' \in Pa(A : B)$ and, thus, $X \sim_G Y|Z$ performs the recursive call $A'' \sim_{G_{\setminus A''}} W$ and W_E, B' and E are not used in it. By a similar reasoning, one can prove that $A' \sim_{G_{\setminus A'}} W$ implies that $A' \sim_G B'|(W \setminus W_E B')E$ holds. Then,

1. $A \not\perp_p A'|(W \setminus W_E B')E$ by $A \sim_G A'|(W \setminus W_E B')E$ and the induction hypothesis
2. $A' \not\perp_p B'|(W \setminus W_E B')E$ by $A' \sim_G B'|(W \setminus W_E B')E$ and the induction hypothesis
3. $A \not\perp_p B'|(W \setminus W_E B')E$ by weak transitivity2 on (1), (2), and $A \perp_G B'|(W \setminus W_E B')EA'$
4. $A \not\perp_p B'E|W \setminus W_E B'$ by weak union on (3)
5. $A \not\perp_p E|W \setminus W_E$ by contraction2 on (4) and $A \perp_G B'|W \setminus W_E B'$
6. $A \not\perp_p EW_E|W \setminus W_E$ by decomposition on (5)
7. $A \not\perp_p E|W$ by intersection on (6) and $A \perp_p W_E|(W \setminus W_E)E$
8. $E \not\perp_p B|W$ by $E \sim_G B|W$ and the induction hypothesis
9. $A \not\perp_p B|W$ by weak transitivity2 on (7), (8), and $A \perp_G B|WE$
10. $X \not\perp_p Y|Z$ by weak union and decomposition on (9).

The second scenario that we consider in the proof is when E is in a rope in $A : B$. In this case, $A : B$ has a subroute of the form $C \rightarrow D \rightarrow \ldots \rightarrow E \rightarrow \ldots \rightarrow F \leftarrow \ldots \leftarrow E \leftarrow \ldots \leftarrow D \leftarrow H$ with $C \neq H$ but possibly $D = E$ and/or $E = F$. Let W_F denote the descendants of F that are in W. Let W' be as in the proof of Lemma 3. Note that $C \perp_G H|W \setminus W'W_F F$ because $A : B$ is the only route between A and B that is minimally superactive wrt X, Y and Z in G. Then,

11. $C \not\perp_p F|(W \setminus W'W_F F)H$ by considering the first scenario for $C \sim_G F|(W \setminus W'W_F F)H$
12. $H \not\perp_p F|(W \setminus W'W_F F)C$ by considering the first scenario for $H \sim_G F|(W \setminus W'W_F F)C$
13. $C \not\perp_p H|(W \setminus W'W_F F)F$ by intersectional weak transitivity1 on (11), (12), and $C \perp_G H|W \setminus W'W_F F$
14. $C \not\perp_p HW_F|(W \setminus W'W_F F)F$ by decomposition on (13)
15. $C \not\perp_p H|W \setminus W'$ by contraction2 on (14) and $C \perp_G W_F|(W \setminus W'W_F F)F$
16. $A \not\perp_p C|W \setminus W'$ by $A \sim_G C|W \setminus W'$ and the induction hypothesis
17. $A \not\perp_p H|W \setminus W'$ by weak transitivity2 on (15), (16), and $A \perp_G H|(W \setminus W')C$
18. $H \not\perp_p B|W \setminus W'$ by $H \sim_G B|W \setminus W'$ and the induction hypothesis
19. $A \not\perp_p B|W \setminus W'$ by weak transitivity2 on (17), (18), and $A \perp_G B|(W \setminus W')H$
20. $X \not\perp_p Y|Z$ follows from (19) by repeating the steps (3)-(7) in the proof of Lemma 3.

Finally, note that we have derived (20) from the dependence base of p for G by using only the ten properties introduced at the beginning of Section 4. Thus, $X \not\perp_p Y|Z$ is in the IWT graphoid closure of the dependence base of p for G. $\square$

The theorem below proves that the criterion in Definition 1 is complete in certain sense.

Theorem 4. *Let G be a polytree-like MDI map of a WT graphoid p. If $X \not\perp_p Y|Z$ is in the IWT graphoid closure of the dependence base of p for G, then $X \sim_G Y|Z$.*

Proof. Clearly, all the dependencies in the dependence base of p for G are identified by the criterion in Definition 1. It only remains to prove that the criterion satisfies the ten properties introduced at the beginning of Section 4.

• Symmetry $Y \sim_G X|Z \Rightarrow X \sim_G Y|Z$.

Trivial.

• Weak union $X \sim_G Y|ZW \Rightarrow X \sim_G YW|Z$.

We prove a simplified version of the property: We assume that W contains a single node. Repeated application of this simplified property proves the original property. Let $X \sim_G Y|ZW$ hold due to a route $A : B$ with $A \in X$ and $B \in Y$. We prove the simplified property by induction over the number of collider nodes in $A : B$. If this number is zero, then the proof is immediate because W cannot be in $A : B$ for $A : B$ to be minimally superactive wrt X, Y and ZW in G. Assume as induction hypothesis that the simplified property holds when the number of collider nodes in $A : B$ is smaller than l. We now prove the simplified property when this number is l. The proof is immediate unless W is in $A : B$. If the latter occurs, then $X \sim_G YW|Z$ holds due to $A : W$, i.e. the subroute of $A : B$ between A and W. To see it, note that $A : W$ is the only route between A and W that is minimally superactive wrt X, YW and Z in G. Note also that W must be a collider node in $A : B$ for $A : B$ to be minimally superactive wrt X, Y and ZW in G. Thus, $A : B$ has a subroute of the form $C \rightarrow D \rightarrow \ldots \rightarrow W \leftarrow \ldots \leftarrow D \leftarrow E$ with $C \neq E$ but possibly $D = W$. Then, every recursive call that $X \sim_G YW|Z$ performs belongs to one of the following two groups. The first group consists of the recursive calls $A' \sim_{G_{\setminus A'}} XYZW \setminus AW$ with $A' \in Pa(A : W) \setminus E$. These recursive calls hold because $A' \in Pa(A : B)$ and, thus, $X \sim_G Y|ZW$ performs the recursive calls $A' \sim_{G_{\setminus A'}} XYZW \setminus AB$ and B and W are not used in them. The second group consists of the recursive call $E \sim_{G_{\setminus E}} XYZW \setminus AW$. We prove that $E \sim_{G_{\setminus E}} Y|XZW \setminus A$ holds, which implies that $E \sim_{G_{\setminus E}} XYZW \setminus A$ holds by repeated application of the induction hypothesis and, since W is not used in the recursive call, $E \sim_{G_{\setminus E}} XYZW \setminus AW$ holds too. To see it, note that the subroute of $A : B$ between E and B, here denoted $E : B$, is the only route between E and B that is minimally superactive wrt E, Y and $XZW \setminus A$ in $G_{\setminus E}$. Moreover, every recursive call that $E \sim_{G_{\setminus E}} Y|XZW \setminus A$ performs is of the form $A' \sim_{(G_{\setminus E})_{\setminus A'}} XYZW \setminus AB$ with $A' \in Pa(E : B)$. This recursive call holds because $A' \in Pa(A : B)$ and, thus, $X \sim_G Y|ZW$ performs the recursive call $A' \sim_{G_{\setminus A'}} XYZW \setminus AB$.

• Decomposition $X \sim_G Y|Z \Rightarrow X \sim_G YW|Z$.

We prove a simplified version of the property: We assume that W contains a single node. Repeated application of this simplified property proves the original property. Let $X \sim_G Y|Z$ hold due to a route $A : B$ with $A \in X$ and $B \in Y$. The proof is immediate unless W is in $A : B$. If the latter occurs, then, $X \sim_G YW|Z$ holds due to $A : W$, i.e. the subroute of $A : B$ between A and W. To see it, note that $A : W$ is the only route between A and W that is minimally superactive wrt X, YW and Z in G. Now, consider the following two scenarios. The first scenario is when W is outside every rope in $A : B$. In this case, every recursive call that $X \sim_G YW|Z$ performs is of the form $A' \sim_{G_{\setminus A'}} XYZW \setminus AW$ with $A' \in Pa(A : W)$. This recursive call holds because $A' \in Pa(A : B)$ and, thus, $X \sim_G Y|Z$ performs the recursive call $A' \sim_{G_{\setminus A'}} XYZ \setminus AB$ and B and W are not used in it. The second scenario is when W is in some rope in $A : B$. In this case, $A : B$ has a subroute of the form $C \rightarrow D \rightarrow \ldots \rightarrow W \rightarrow \ldots \rightarrow E \leftarrow \ldots \leftarrow W \leftarrow \ldots \leftarrow D \leftarrow F$ with $C \neq F$ but possibly $D = W$ and/or $W = E$. However, in this case, $X \sim_G Y|ZW$ holds due to the route $(A : B) \setminus (W \rightarrow \ldots \rightarrow E \leftarrow \ldots \leftarrow W)$, here denoted ρ. To see it, note that ρ is the only route between A and B that is minimally superactive wrt X, Y and ZW in G. Moreover, every recursive call that $X \sim_G Y|ZW$ performs is of the form $A' \sim_{G_{\setminus A'}} XYZW \setminus AB$ with $A' \in Pa(\rho)$. This recursive call holds because $A' \in Pa(A:B)$ and, thus, $X \sim_G Y|Z$ performs the recursive call $A' \sim_{G_{\setminus A'}} XYZ \setminus AB$ and W is not used in it. Finally, note that if $X \sim_G Y|ZW$ holds, then $X \sim_G YW|Z$ holds by weak union.

• Contraction1 $X \sim_G YW|Z \wedge X \perp_G Y|ZW \Rightarrow X \sim_G W|Z$.

In the proof of this property, we make use of the fact that separation in DAGs is a WT graphoid (Pearl, 1988) and, thus, it satisfies the ten properties introduced at the beginning of Section 4. Let $X \sim_G YW|Z$ hold due to a route $A : B$ with $A \in X$ and $B \in YW$. Then, $A \not\perp_G B|XYZW \setminus AB$ and, thus, $X \not\perp_G B(Y \setminus B)|WZ \setminus B$ by weak union. This implies that $B \notin Y$ because, otherwise, it would contradict $X \perp_G Y|ZW$. Likewise, for all $A' \in Pa(A : B)$, $A \not\perp_G A'|XYZW \setminus AA'$ and, thus, $X \setminus A' \not\perp_G A'(Y \setminus A')|WZ \setminus A'$ by weak union. This implies that $A' \notin Y$ because, otherwise, it would contradict $X \perp_G Y|ZW$. Furthermore, note that every recursive call that $X \sim_G YW|Z$ performs is of the form $A' \sim_{G_{\setminus A'}} XYZW \setminus AB$ with $A' \in Pa(A : B)$ and $A' \notin XYZW$. Assume that this recursive call holds due to a route $A' : B'$ with $B' \in XYZW \setminus AB$. By reasoning as above, we can conclude that $A' \not\perp_G B'|XYZW \setminus ABB'$ and $A' \not\perp_G A''|XYZW \setminus ABA''$ with $A'' \in Pa(A' : B')$. Then,

1. $A \not\perp_G A'|XYZW \setminus AB'$ by $A \not\perp_G A'|XYZW \setminus AA'$, $A' \notin XYZW$, and B' is not involved

2. $A' \not\perp_G B'|XYZW \setminus AB'$ by $A' \not\perp_G B'|XYZW \setminus ABB'$ and B is not involved

3. $A \not\perp_G B'|XYZW \setminus AB'$ by weak transitivity on (1), (2), and $A \not\perp_G B'|(XYZW \setminus AB')A'$

4. $X \setminus B' \not\perp_G B'(Y \setminus B')|WZ \setminus B'$ by weak union on (3).

Note that (4) implies that $B' \notin Y$ because, otherwise, it would contradict $X \perp_G Y|ZW$. Moreover,

5. $A \not\perp_G A'|XYZW \setminus AA''$ by $A \not\perp_G A'|XYZW \setminus AA'$, $A' \notin XYZW$, and A'' is not involved

6. $A' \not\perp_G A''|XYZW \setminus AA''$ by $A' \not\perp_G A''|XYZW \setminus ABA''$ and B is not involved

7. $A \not\perp_G A''|XYZW \setminus AA''$ by weak transitivity on (5), (6), and $A \not\perp_G A''|(XYZW \setminus AA'')A'$

8. $X \setminus A'' \not\perp_G B'(Y \setminus A'')|WZ \setminus A''$ by weak union on (7).

Note that (8) implies that $A'' \notin Y$ because, otherwise, it would contradict $X \perp_G Y|ZW$. Therefore, we have proven that $X \sim_G W|Z$ holds if the recursive calls performed by $X \sim_G YW|Z$ do not perform other recursive calls because, in this case, none of the key nodes is in Y and, thus, Y can be dropped. When a recursive call performed by $X \sim_G YW|Z$ performs another recursive call and this possibly another and so on, one just needs to repeat the reasoning above for each of these recursive calls.

• Contraction2 $X \sim_G YW|Z \wedge X \perp_G W|Z \Rightarrow X \sim_G Y|ZW$.

Let $X \sim_G YW|Z$ hold due to the route $A : B$ with $A \in X$ and $B \in YW$. We prove that $X \sim_G Y|ZW$ holds due to $A : B$. Since $A : B$ is minimally superactive wrt X, YW and Z in G, no node in $XYW \setminus AB$ can be in $A : B$. Then, $B \in Y$ by $X \perp_G W|Z$ and, thus, $A : B$ is minimally superactive wrt X, Y and ZW in G. Moreover, $A : B$ is the only such route between A and B. To see it, assume to the contrary that there is a second such route between A and B. Note that this second route must have some collider node in $C \in W$ for $A : B$ to be the only route between A and B that is minimally superactive wrt X, YW and Z in G. Then, this second route must also have some collider node $D \in Y$ between A and C by $X \perp_G W|Z$. However, this is a contradiction. Finally, note that every recursive call that $X \sim_G Y|ZW$ performs holds because $X \sim_G YW|Z$ also performs that recursive call because, as shown, $B \in Y$.

• Intersection $X \sim_G YW|Z \wedge X \perp_G Y|ZW \Rightarrow X \sim_G W|ZY$.

Let $X \sim_G YW|Z$ hold due to the route $A : B$ with $A \in X$ and $B \in YW$. We prove that $X \sim_G W|ZY$ holds due to $A : B$. Since $A : B$ is minimally superactive wrt X, YW and Z in G, no node in $XYW \setminus AB$ can be in $A : B$. Then, $B \in W$ by $X \perp_G Y|ZW$ and, thus, $A : B$ is minimally superactive wrt X, W and ZY in G. Moreover, $A : B$ is the only such route between A and B. To see it, assume to the contrary that there is a second such route between A and B. Note that this second route must have some collider node in $C \in Y$ for $A : B$ to be the only route between A and B that is minimally superactive wrt X, YW and Z in G. Then, this second route must also have some non-collider node $D \in W$ between A and C by $X \perp_G Y|ZW$. However, this is a contradiction. Finally, note that every recursive call that $X \sim_G W|ZY$ performs holds because $X \sim_G YW|Z$ also performs that recursive call because, as shown, $B \in W$.

• Intersectional weak transitivity1 $X \sim_G V|ZY \wedge V \sim_G Y|ZX \wedge X \perp_G Y|Z \Rightarrow X \sim_G Y|ZV$.

Let $X \sim_G V|ZY$ and $V \sim_G Y|ZX$ hold due to the routes $A : V$ with $A \in X$ and $V : B$ with $B \in Y$, respectively. We prove that $A \sim_G B|XYZV \setminus AB$ holds, which implies $X \sim_G Y|ZV$ by weak union and decomposition. Note first that $X \perp_G Y|Z$ implies that $A : V$ followed by $V : B$, here denoted $A : B$, is the only route between A and B that is minimally superactive wrt A, B and $XYZV \setminus AB$ in G. Note also that every recursive call that $A \sim_G B|XYZV \setminus AB$ performs is of the form $A' \sim_{G_{\setminus A'}} XYZV \setminus AB$ with $A' \in Pa(A : B)$. Note also that $A' \in Pa(A : V)$ or $A' \in Pa(V : B)$. In the former case, we know that $X \sim_G V|ZY$ performs the recursive call $A' \sim_{G_{\setminus A'}} XYZ \setminus A$ which does not use B or V and, thus, $A' \sim_{G_{\setminus A'}} XYZV \setminus AB$ holds. In the latter case, we know that $V \sim_G Y|ZX$ performs the recursive call $A' \sim_{G_{\setminus A'}} XYZ \setminus B$ which does not use A or V and, thus, $A' \sim_{G_{\setminus A'}} XYZV \setminus AB$ holds.

• Intersectional weak transitivity2 $X \sim_G V|ZY \wedge V \sim_G Y|ZX \wedge X \perp_G Y|ZV \Rightarrow X \sim_G Y|Z$.

Let $X \sim_G V|ZY$ and $V \sim_G Y|ZX$ hold due to the routes $A : V$ with $A \in X$ and $V : B$ with $B \in Y$, respectively. We prove that $A \sim_G B|XYZ \setminus AB$ holds, which implies $X \sim_G Y|Z$ by weak union and decomposition. Note first that $X \perp_G Y|ZV$ implies that $A : V$ followed by $V : B$, here denoted $A : B$, is the only route between A and B that is minimally superactive wrt A, B and $XYZ \setminus AB$ in G. Note also that every recursive call that $A \sim_G B|XYZ \setminus AB$ performs is of the form $A' \sim_{G_{\setminus A'}} XYZ \setminus AB$ with $A' \in Pa(A : B)$. Note also that $A' \in Pa(A : V)$ or $A' \in Pa(V : B)$. In the former case, we know that $X \sim_G V|ZY$ performs the recursive call $A' \sim_{G_{\setminus A'}} XYZ \setminus A$ which does not use B and, thus, $A' \sim_{G_{\setminus A'}} XYZ \setminus AB$ holds. In the latter case, we know that $V \sim_G Y|ZX$ performs the recursive call $A' \sim_{G_{\setminus A'}} XYZ \setminus B$ which does not use A and, thus, $A' \sim_{G_{\setminus A'}} XYZ \setminus AB$ holds.

• Weak transitivity1 $X \sim_G V|Z \wedge V \sim_G Y|Z \wedge X \perp_G Y|Z \Rightarrow X \sim_G Y|ZV$.

1. $X \sim_G VY|Z$ and $VX \sim_G Y|Z$ by decomposition on $X \sim_G V|Z$ and $V \sim_G Y|Z$

2. $X \sim_G V|ZY$ and $V \sim_G Y|ZX$ by contraction2 on (1) and $X \perp_G Y|Z$

3. $X \sim_G Y|ZV$ by intersectional weak transitivity1 on (1), (2), and $X \perp_G Y|Z$.

• Weak transitivity2 $X \sim_G V|Z \wedge V \sim_G Y|Z \wedge X \perp_G Y|ZV \Rightarrow X \sim_G Y|Z$.

Just replace (3) in the proof of weak transitivity1 by

3. $X \sim_G Y|Z$ by intersectional weak transitivity2 on (1), (2), and $X \perp_G Y|ZV$. □

5 Discussion

As discussed in Section 1, the new criterion introduced in this paper is more (resp. less) powerful but less (resp. more) applicable than the criterion in (Bouckaert, 1995) (resp. (Peña, 2007)). To see it, consider Example 1. The new criterion and the criterion in (Bouckaert, 1995) can be applied but the criterion in (Peña, 2007) cannot, because p does not satisfy composition $X \perp_p Y|Z \wedge X \perp_p W|Z \Rightarrow X \perp_p YW|Z$. However, the new criterion reads $A \not\perp_p B|C$ from G but the criterion in (Bouckaert, 1995) does not, because A and B are not adjacent in G and this is necessary for that criterion to be conclusive.

References

Bouckaert, R. R. *Bayesian Belief Networks: From Construction to Inference.* PhD Thesis, University of Utrecht, 1995.

Pearl, J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufmann, 1988.

Peña, J. M. Reading Dependencies from Polytree-Like Bayesian Networks. In *UAI 2007*, 303-309.

Peña, J. M., Nilsson, R., Björkegren, J. and Tegnér, J. An Algorithm for Reading Dependencies from the Minimal Undirected Independence Map of a Graphoid that Satisfies Weak Transitivity. Journal of Machine Learning Research, 10, 1071-1094.

Studený, M. Bayesian Networks from the Point of View of Chain Graphs. In *UAI 1998*, 496-503.