Proof supplement to:
A Logical Characterization of Constraint-Based Causal Discovery

Tom Claassen and Tom Heskes
Institute for Computer and Information Science
Radboud University Nijmegen
The Netherlands

Sections and lemmas refer to the corresponding entry in (Claassen and Heskes, 2011). Note: X, Y, Z, $\mathbf{W}$, etc. are disjoint (sets of) observed nodes in a causal DAG $\mathcal{G}_C$, and $\mathbf{S}$ represents the (possibly empty) set of selection nodes.

2 Background

The definition of a causal relation in a causal DAG, rewritten in terms of standard logical properties:

Proposition 1. Causal relations in a DAG $\mathcal{G}_C$ are:

$$\begin{array}{llll} \textit{irreflexive} & : X \Rightarrow X & \vdash & \textit{false} \\ \textit{acyclic} & : X \Rightarrow Y & \vdash & Y \not\Rightarrow X \\ \textit{transitive} & : (X \Rightarrow Y) \wedge (Y \Rightarrow Z) & \vdash & X \Rightarrow Z \end{array}$$

Proof. As the edges in $\mathcal{G}_C$ represent causal relations, a path of length 0 (no edge) is not considered a causal relation, and existence of a directed path from a variable back to itself would contravene the causal DAG assumption, hence: irreflexive and acyclic (or, more accurate, *asymmetric*). Transitivity follows immediately, by concatenation, from the sequence $\langle X, .., Y, .., Z \rangle$, in which each node is parent of its successor, hence a directed path from X to Z in $\mathcal{G}_C$. □

3 Rules from minimal independencies

The first lemma is a familiar result, see (Spirtes et al., 1999; Claassen and Heskes, 2010), that brings out the symmetry between a node that creates an independence, and a node that destroys an independence. For illustration purposes we include the proof below.

Lemma 2 If a node Z changes an (in)dependence relation between X and Y in a causal DAG, then:

1. $X \perp\!\!\!\perp Y \mid \mathbf{W} \cup [Z] \ \vdash \ Z \Rightarrow (X \cup Y \cup \mathbf{W} \cup \mathbf{S})$,

2. $X \not\perp\!\!\!\perp Y \mid \mathbf{W} \cup [Z] \ \vdash \ Z \not\Rightarrow (X \cup Y \cup \mathbf{W} \cup \mathbf{S})$.

with special case $X \perp\!\!\!\perp Y \mid [\mathbf{W} \cup Z] \ \vdash \ Z \Rightarrow (X \cup Y \cup \mathbf{S})$.

Proof. (1.) To block, node Z must be a noncollider on a path $\pi = \langle X, .., Z, .., Y \rangle$ in $\mathcal{G}_C$ that is unblocked given $\mathbf{W} \cup \mathbf{S}$. As Z is a noncollider it has at least one outgoing arc along π. Follow π in this direction until either a collider is encountered or the end of π is reached. Every collider along π has to be an ancestor of $(\mathbf{W} \cup \mathbf{S})$, which implies that in either case Z has a nonzero directed path in $\mathcal{G}_C$ (=causal relation) to at least one node from $(X \cup Y \cup \mathbf{W} \cup \mathbf{S})$.

(2.) To unblock, node Z must be (a descendant of) at least one collider on a path $\pi = \langle X, .., Y \rangle$ in $\mathcal{G}_C$ that is blocked given $\mathbf{W} \cup \mathbf{S}$. Any directed path in $\mathcal{G}_C$ from Z to a node in $\mathbf{W} \cup \mathbf{S}$ implies that the collider(s) would already be unblocked when conditioning on just $\mathbf{W} \cup \mathbf{S}$. No directed paths from Z to $(\mathbf{W} \cup \mathbf{S})$ implies that if there existed a directed path from Z to X or Y, then it could not be blocked by any node $(\mathbf{W} \cup \mathbf{S})$. But then such a path would make Z a noncollider on an unblocked path between X and Y given $(\mathbf{W} \cup \mathbf{S})$: starting from X, let θ_X be the first collider encountered along π that is unblocked by conditioning on Z, and similarly θ_Y the first collider along π starting from Y, (possibly $\theta_X = \theta_Y$, but $\{\theta_X, \theta_Y\} \notin \mathbf{W}$ (otherwise Z not needed)); then the paths $\langle X, \theta_X, Z \rangle$ and $\langle Z, \theta_Y, Y \rangle$ are into Z and unblocked given $\mathbf{W} \cup Z$, so a directed path $Z \Rightarrow X$ would make Z a noncollider on unblocked path $\langle X, \theta_X, Z, Y \rangle$ given $\mathbf{W}$, contradicting $X \perp\!\!\!\perp Y \mid \mathbf{W}$; idem for $Z \Rightarrow Y$.

The special case for minimal $X \perp\!\!\!\perp Y \mid [\mathbf{W} \cup Z]$ follows from (1.) and acyclicity. By contradiction: suppose $\exists U_1 \in (\mathbf{W} \cup Z) : U_1 \not\Rightarrow (X \cup Y \cup \mathbf{S})$, then, as (1.) applies to all nodes $(\mathbf{W} \cup Z)$, there must be a node $U_2 \in (\mathbf{W} \cup Z) \setminus U_1$ (acyclicity) such that $U_1 \Rightarrow U_2$. But (transitivity) U_2 also cannot have a directed path to $(X \cup Y \cup \mathbf{S})$, and so there must be a node $U_3 \in (\mathbf{W} \cup Z) \setminus \{U_1, U_2\}$ (acyclicity) such that $U_2 \Rightarrow U_3$. This can continue until all nodes in $(\mathbf{W} \cup Z)$ have been allocated at which stage the last node cannot have a directed path to any $(X \cup Y \cup \mathbf{W} \cup Z \cup \mathbf{S})$, in contradiction with (1.). □

Observed minimal independences can be converted into logical statements on causal relations:

Lemma 3. For observed minimal (in)dependencies between nodes in a causal DAG $\mathcal{G}_C$:

1. $X \perp\!\!\!\perp Y \,|\, [\mathbf{W} \cup Z] \;\vdash\; Z \Rightarrow X \;\vee\; Z \Rightarrow Y \;\vee\; Z \Rightarrow \mathbf{S}$

2. $X \not\perp\!\!\!\perp Y \,|\, \mathbf{W} \cup [Z] \;\vdash\; Z \nRightarrow X \;\wedge\; Z \nRightarrow Y \;\wedge$
$Z \nRightarrow \mathbf{W} \wedge Z \nRightarrow \mathbf{S}$

Proof. Follows immediately from lemma 2, together with the (absence of) causal relations to sets of nodes:
(1.) $\equiv \exists U \in (X \cup Y \cup \mathbf{S}) : Z \Rightarrow U$,
(2.) $\equiv \forall U \in (X \cup Y \cup \mathbf{W} \cup \mathbf{S}) : Z \nRightarrow U$. $\qquad\square$

Next two lemmas used in the proof of lemma 4.

Lemma 4.1 For two observed nodes, X and Y, in a causal DAG $\mathcal{G}_C$: $X \not\perp\!\!\!\perp Y$, iff they are connected by a trek in $\mathcal{G}_C$ or they both have treks into $\mathbf{S}$

Proof. Almost by definition. Assuming causal Markov and faithfulness, two observed nodes X and Y are dependent given a set $\mathbf{Z}$, iff they are connected by a path π in $\mathcal{G}_C$ on which all noncolliders are not in $\mathbf{Z}$ and all colliders are (ancestor of) nodes in $(\mathbf{Z} \cup \mathbf{S})$. For $\mathbf{Z} = \varnothing$ this reduces to a path π on which all colliders are in $An(\mathbf{S})$. Starting from X, follow π until the first collider. Then X has a colliderless path to a node with a directed path to $\mathbf{S}$, which implies a trek from X to $\mathbf{S}$. If Y is reached, then π is by definition a colliderless path, or trek, to Y. Idem for Y. $\qquad\square$

Lemma 4.2 In a causal DAG $\mathcal{G}_C$, if $X \not\perp\!\!\!\perp Y$, then identifiable absence of a causal relation $X \nRightarrow Y$ implies absence of selection bias $X \nRightarrow \mathbf{S}$.

Proof. For adjacent nodes in a PAG $\mathcal{P}$, the proof is trivial: identifiable absence of a causal relation means identifiable non-ancestorship, and so an invariant arrowhead $X \leftarrow\!\!* \, Y$ in $\mathcal{P}$. By definition of the MAG, see §4.2 in (Richardson and Spirtes, 2002), this means that $X \notin An(Y \cup \mathbf{S})$.
For nonadjacent nodes we can use Theorem 2 from (Claassen and Heskes, 2010), which states that there is identifiable absence of a causal relation $X \nRightarrow Y$, iff it is impossible to go from X to Y in the graph $\mathcal{P}$, without going against an arrowhead. By contradiction: suppose that $X \Rightarrow \mathbf{S}$. This implies $An(X) \Rightarrow \mathbf{S}$, so nodes that are ancestor of X have no (invariant) arrowheads (only tails). By lemma 4.1, dependent nodes either have a trek between them, or both have treks to $\mathbf{S}$. But if there is a trek between X and Y, then no node between X and the source of that trek can have an arrowhead, and all nodes between the source and Y are going 'with' the arrowhead, so then not all

paths go against an arrowhead. Similarly for treks to $\mathbf{S}$. Therefore also for nonadjacent nodes $X \nRightarrow \mathbf{S}$. $\qquad\square$

Lemma 4 (Inferred blocking node). In a causal system $\mathcal{G}_C$, if $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}]$, and there is a subset $\{Z_1, \ldots, Z_k, Z\} \subseteq \mathbf{Z}$, such that in the sequence $[\mathbf{U}] \equiv [U_0, \ldots, U_{k+2}] = [X, Z_1, \ldots, Z_k, Z, Y]$ it holds that:

- $U_i \nRightarrow \{U_{i-1}, U_{i+1}\}$,
- $U_j \not\perp\!\!\!\perp U_{j+1} \,|\, \mathbf{Z}'$,

with $i = 1..k$, and with $j = 0..(k+1)$ and $\forall \mathbf{Z}' \subseteq \mathbf{Z} \setminus \{U_j, U_{j+1}\}$, then $Z \Rightarrow (Z_k \cup Y \cup \mathbf{S})$.

Proof. In words: if no node Z_i in the minimal independence $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}]$ has a causal relation[1] (directed path in $\mathcal{G}_C$) to either of its neighbors in the sequence $[X, Z_1, \ldots, Z_k, Z, Y]$, and all neighboring nodes in the sequence are dependent given any subset of $\mathbf{Z}$, then Z has a causal relation to Z_k, Y, and/or $\mathbf{S}$.

First we show that there is an unblocked path from X to Z in $\mathcal{G}_C$ relative to $\mathbf{Z}_{\setminus Z}$. The first item, in combination with lemma 4.1 implies that there is no selection bias on any of the nodes Z_i. By lemma 2, this, together with the given $Z_1 \nRightarrow X$, implies $Z_1 \Rightarrow Y$, and so it also follows that there is no selection bias on Y (otherwise $Z_1 \Rightarrow (Y) \Rightarrow \mathbf{S}$).
By the second item, all neighbors in the sequence $[\mathbf{U}]$ are dependent (given empty set), and so by the previous observation in combination with lemma 4.1 this implies that each successive pair is connected by a trek (but not a directed path, by item 1) in $\mathcal{G}_C$, with the possible exception of the edges to X and Z, that can still correspond to directed paths and/or treks to $\mathbf{S}$.
As each successive pair in the sequence is connected by an unblocked (sub)path given $\mathbf{Z}_{\setminus Z}$ that is *into* both Z_i and Z_{i+1}, it follows (by concatenating them) that there is also an unblocked path from X to Z in $\mathcal{G}_C$ relative to $\mathbf{Z}_{\setminus Z}$. Nodes Z and Y are also not separated by any subset from $\mathbf{Z}$, and so are connected by an unblocked subpath relative to $\mathbf{Z}_{\setminus Z}$.

In conclusion, by construction there are unblocked paths from X (via Z_k) and Y to Z in $\mathcal{G}_C$, given $\mathbf{Z}_{\setminus Z}$. If both paths from Z_k and Y are into Z, then the sequence $[\mathbf{U}]$ would represent an unblocked path between X and Y given $\mathbf{Z}$ in $\mathcal{G}_C$, which would make X and Y dependent, contrary the given. Therefore Z must be an ancestor of Z_k and/or Y, and/or have a directed path to $\mathbf{S}$ in $\mathcal{G}_C$. In other words, then: $Z \Rightarrow Z_k \vee Z \Rightarrow Y \vee Z \Rightarrow \mathbf{S}$. $\qquad\square$

---

[1]The lemma is stated in terms of *identifiable* absence of causal relations $U_i \nRightarrow \{U_{i-1}, U_{i+1}\}$, which, by lemma 4.2, implies that in the underlying $\mathcal{G}_C$ it holds that $U_i \nRightarrow \{U_{i-1}, U_{i+1}\} \cup \mathbf{S}$.

Lemma 5. In a causal system $\mathcal{G}_C$, a (minimal) conditional independence $X \perp\!\!\!\perp Y \,|\, \mathbf{Z}$ implies that all direct causal paths $X \Rightarrow Y$ or $X \Leftarrow Y$, or common causes of X and Y in $\mathcal{G}_C$ are mediated by nodes in $\mathbf{Z}$.

Proof. See (Spirtes et al., 2000). Assumes that no causal paths are blocked by selection nodes, which is implicitly covered by the faithfulness assumption. $\square$

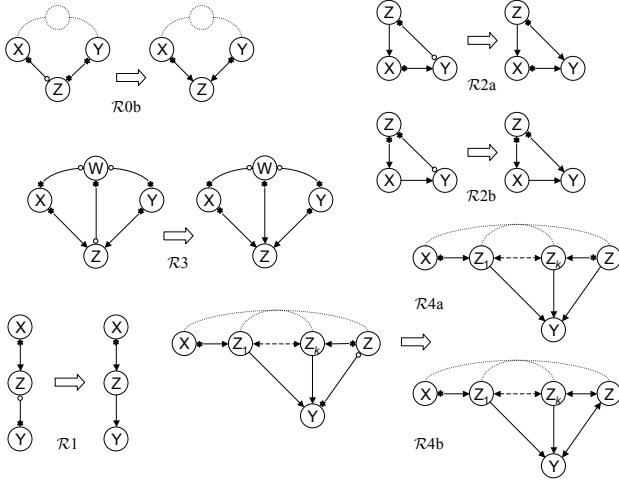# 4 A logical characterization of causal information

Figure 1: Rules $\mathcal{R}$0b$-\mathcal{R}$4b, arrowhead orientation rules

Lemma 7 (Invariant arrowheads). In a PAG $\mathcal{P}$, all invariant arrowheads $Z \ast\!\!\rightarrow Y$ are instances of

(1): $U \not\perp\!\!\!\perp V \,|\, \mathbf{W} \cup [Y]$, created from $U \perp\!\!\!\perp V \,|\, [\mathbf{W}]$, with $Z \in \{U, V, \mathbf{W}\}$,

(2): $X \perp\!\!\!\perp Y \,|\, [\mathbf{W} \cup Z]$, with $Z \not\Rightarrow (X \cup \mathbf{S})$ from either case (1) or case (2).

Proof sketch. Both cases are sound:
(1.) By lemma 1.2, the first gives $(Y \not\Rightarrow Z) \wedge (Y \not\Rightarrow \mathbf{S})$, which, by definition, implies that if Y has an edge to Z in $\mathcal{P}$, then the mark at Y is an (invariant) arrowhead.
(2.) The second is an application of corollary 2, giving $(Z \Rightarrow X) \vee (Z \Rightarrow Y) \vee (Z \Rightarrow \mathbf{S})$, where the first and third are eliminated by the arrowhead at $X \ast\!\!\rightarrow Z$ (def). Therefore $Z \Rightarrow Y$, and so (acyclicity) also $Y \not\Rightarrow Z$, but also $Y \not\Rightarrow \mathbf{S}$, otherwise (transitivity) $Z \Rightarrow \mathbf{S}$. Therefore, if Y has an edge to Z in $\mathcal{P}$, then it has an arrowhead mark at Y.

The proof that they are also complete follows from the lemmas below, by induction on the graphical orientation rules $\mathcal{R}$0b$-\mathcal{R}$4b, showing that none of them introduces a violation of Lemma 7. As these rules are sufficient for arrowhead completeness (Ali et al., 2005;

Zhang, 2008), it follows that the theorem holds for all invariant arrowheads. $\square$

Lemma 7.1 The arrowheads at Z from rules $\mathcal{R}$0b, $\mathcal{R}$3, and $\mathcal{R}$4b are covered by case (1) and the arrowhead at Y from rule $\mathcal{R}$1 is covered by case (2).

Proof. Implied directly by the corresponding patterns in Figure 1:

$\mathcal{R}$0b: If this rule fires, then it implies $X \perp\!\!\!\perp Y \,|\, [\mathbf{W}]$ for some set $\mathbf{W}$ (possibly empty), with $X \not\perp\!\!\!\perp Y \,|\, \mathbf{W} \cup Z$. Therefore case (1) applies and Z gets arrowheads on the edges from X and Y in $\mathcal{G}$, just as in the consequent of $\mathcal{R}$0b in fig.1.

$\mathcal{R}$1: Implies $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}]$ with $Z \in \mathbf{Z}$ and an arrowhead at Z from a rule that fired before. If no violations before $\mathcal{R}$1 fires, then case (2) applies, and there is an arrowhead at $Z \longrightarrow Y$ in $\mathcal{G}$, just as in the consequent of $\mathcal{R}$1.

$\mathcal{R}$3: Implies $X \perp\!\!\!\perp Y \,|\, [\mathbf{W}]$ with $W \in \mathbf{W}$, and $X \not\perp\!\!\!\perp Y \,|\, \mathbf{W} \cup Z$. Therefore case (1) applies, to give $W \ast\!\!\rightarrow Z$ in $\mathcal{G}$, just as in $\mathcal{R}$3.

$\mathcal{R}$4b: By construction of the discriminating path, $\mathcal{R}$4b implies $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}]$, with $\{Z_1, \ldots, Z_k\} \in \mathbf{Z}$, but $Z \notin \mathbf{Z}$ as $X \not\perp\!\!\!\perp Y \,|\, \mathbf{Z} \cup Z$. Therefore case (1) applies, resulting in the addition of $Z_k \ast\!\!\rightarrow Z \leftarrow\!\ast Y$ to $\mathcal{G}$, just as in $\mathcal{R}$4b.

$\square$

Lemma 7.2 The arrowheads at Y from rules $\mathcal{R}$2b, $\mathcal{R}$4a, and $\mathcal{R}$4b are covered by cases (1) and (2).

Proof. First $\mathcal{R}$2b. If no violations before this rule fires, then the arrowhead at $Z \ast \rightarrow X$ either appeared by case (1) as a node X that creates the dependency $U \not\perp\!\!\!\perp V \,|\, \mathbf{W} \cup X$ from $U \perp\!\!\!\perp V \,|\, [\mathbf{W}]$, with $Z \in \{U, V, \mathbf{W}\}$ (case 1a), or by case (2), as a minimal conditional independence $X \perp\!\!\!\perp U \,|\, [\mathbf{W} \cup Z]$, with a (somehow) established $Z \not\Rightarrow (U \cup \mathbf{S})$, for which either U and Y are also independent given $\mathbf{W} \cup Z$ (case 2a), or not (case 2b). (Note: $Y \notin \mathbf{W}$ in case (2), otherwise (from $Y \not\Rightarrow X$, lemma 2) $Y \Rightarrow U$, which, together with $Z \Rightarrow X$ and $X \Rightarrow Y$, would imply $Z \Rightarrow U$). For these three instances:

1a) If conditioning on X creates $U \not\perp\!\!\!\perp V \,|\, \mathbf{W} \cup X$, then conditioning on Y as a descendant of X implies $U \not\perp\!\!\!\perp V \,|\, \mathbf{W} \cup Y$, and so case (1) also applies to $Z \ast\!\!\rightarrow Y$.

2a) If $Y \perp\!\!\!\perp U \,|\, \mathbf{W} \cup Z$, then also $Y \perp\!\!\!\perp U \,|\, [\mathbf{W} \cup Z]$, as no subset can block the path between Y and U via X, and so case (2) applies to $Z \longrightarrow Y$.

2b) If $Y \not\perp\!\!\!\perp U \,|\, \mathbf{W} \cup Z$, then there is an unblocked path π between U and Y given $\mathbf{W} \cup Z$. The path π is

into Y, since otherwise the path $\langle X, Y \rangle + \pi$ would be an unblocked path between X and U given $\mathbf{W} \cup Z$, contrary to $X \perp\!\!\!\perp U \,|\, [\mathbf{W} \cup Z]$. Therefore, conditioning on collider Y on the path creates the dependency $X \not\perp\!\!\!\perp U \,|\, \mathbf{W} \cup Z \cup Y$, and so case (1) applies.

In $\mathcal{R}4a$ and $\mathcal{R}4b$, the arrowhead at Y is simply an instance of $\mathcal{R}2b$ with $Z_k = X$. $\qquad\square$

This leaves rule $\mathcal{R}2a$ as the only remaining case to prove. For that we use the observation:

Lemma 7.3 If two nodes X and Y are conditionally independent given a set of nodes $\mathbf{Z}$, $X \perp\!\!\!\perp Y \,|\, \mathbf{Z}$, then an arbitrary node V is either:

(a) part of the conditional independence, i.e. $V \in (X \cup Y \cup \mathbf{Z})$,

(b) conditionally independent of X and/or Y given $\mathbf{Z}$, i.e. $(V \perp\!\!\!\perp X \,|\, \mathbf{Z}) \vee (V \perp\!\!\!\perp Y \,|\, \mathbf{Z})$, or

(c) (descendant of) a collider between U and V such that $X \not\perp\!\!\!\perp Y \,|\, \{\mathbf{Z} \cup V\}$.

Proof. If neither (a) nor (b), i.e. $V \notin (X \cup Y \cup \mathbf{Z})$ *and* $V \not\perp\!\!\!\perp \{X, Y\} \,|\, \mathbf{Z}$, then there are paths $\pi_X = \langle X, \ldots, V \rangle$ and $\pi_Y = \langle Y, \ldots, V \rangle$ in the corresponding graph that are unblocked given $\mathbf{Z}$. Node V has to be a collider on the path $\pi = \pi_X + \pi_Y$, otherwise π would be unblocked given $\mathbf{Z}$ (as $V \notin \mathbf{Z}$), contrary to $X \perp\!\!\!\perp Y \,|\, \mathbf{Z}$. But then conditioning on $\mathbf{Z} \cup V$ will make them dependent, i.e. then (c). $\qquad\square$

Note that if $\mathbf{Z}$ is a *minimal* set that makes X and Y independent, then case (b) does not imply that it is also minimal for $V \perp\!\!\!\perp X/Y \,|\, \mathbf{Z}$, as shown by the example in fig.4: from $X \perp\!\!\!\perp Y \,|\, [\{Z_1, Z_2\}]$, for node V we find $V \perp\!\!\!\perp X \,|\, \{Z_1, Z_2\}$ (as none of the other options in lemma 7.3 applies), ... but this is only *minimal* for subset $V \perp\!\!\!\perp X \,|\, [Z_2]$.
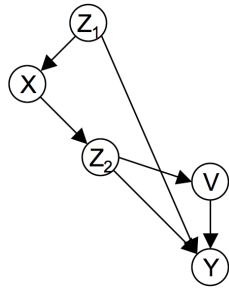

Figure 2: Example of case (b) in lemma 7.2 with 'minimal' only for subset

For the proof of $\mathcal{R}2a$ we also use:

Lemma 7.4 In an ancestral graph $\mathcal{G}$, if a node Z unblocks a blocked path $\pi = \langle U, \ldots, V \rangle$ between two nodes U and V given some set $\mathbf{W}$, then there are unblocked paths from both U and V into Z relative to $\mathbf{W}$, and so $Z \not\perp\!\!\!\perp \{U, V\} \,|\, \mathbf{W}$.

Proof. By definition, a path π is unblocked relative to $\mathbf{W}$ if all noncolliders on the path are not in $\mathbf{W}$ and all colliders are in $An(\mathbf{W})$. Adding a node Z to the conditioning set can never remove a noncollider, so it can only unblock on a collider that is in $(\mathbf{W} \cup Z)$, but not in $(\mathbf{W})$. So the node Z must be (a descendant of) a collider C on the path (possibly $C = Z$). No node $W \in \mathbf{W}_{\backslash C}$ blocks the path $X \Rightarrow Z$ (otherwise conditioning on Z would not be needed), therefore if π is unblocked relative to $\mathbf{W}$, then so are the two paths $\pi_U = \langle U, ..(, X, ..), Z \rangle$ and $\pi_V = \langle Z, (.., X,).., V \rangle$, which implies $Z \not\perp\!\!\!\perp \{U, V\} \,|\, \mathbf{W}$. $\qquad\square$

Finally we need the following result:

Lemma 7.5 In an ancestral graph $\mathcal{G}$, if there are (sets of) nodes U, Y, Z and $\mathbf{W}$, such that $U \perp\!\!\!\perp Z \,|\, \mathbf{W}$ and $U \not\perp\!\!\!\perp Z \,|\, \mathbf{W} \cup Y$, with $Z \ast\!\!\rightarrow Y$ in $\mathcal{G}$, then there is a node $W \in (U \cup \mathbf{W})$, and a set $\mathbf{Q} \subseteq \mathbf{W}$, such that $W \perp\!\!\!\perp Z \,|\, [\mathbf{Q}]$ and $W \not\perp\!\!\!\perp Z \,|\, \mathbf{Q} \cup Y$.
In words: if conditioning on a node Y destroys (unblocks) some conditional independence for a neighbouring node Z, then the same holds for at least some *minimal* conditional independence between Z and one of the other nodes involved.

Proof. By definition, there is a $\mathbf{W}' \subseteq \mathbf{W}$ such that $U \perp\!\!\!\perp Z \,|\, [\mathbf{W}']$. If then also $U \not\perp\!\!\!\perp Z \,|\, \mathbf{W}' \cup Y$, then the lemma applies with $W = U$ and $\mathbf{Q} = \mathbf{W}'$. If not, i.e. if $U \perp\!\!\!\perp Z \,|\, \mathbf{W}' \cup Y$, then we can show that there is a node $W \in \mathbf{W}$ for which the lemma holds.
Let $\mathcal{G}'$ be the MAG obtained from $\mathcal{G}$ by marginalizing out all nodes in $\mathcal{G}$ that are not in $\{U, Y, Z\} \cup \mathbf{W}$. From the original $U \not\perp\!\!\!\perp Z \,|\, \mathbf{W} \cup Y$, by lemma 7.4, there is an unblocked path $\pi = \langle U, \ldots, Y \rangle$ in $\mathcal{G}'$ that is into Y given $\mathbf{W}$. The path π contains one or more (say k) colliders in $\mathcal{G}'$, some of which are (ancestors of) nodes from $\mathbf{W}$, but not from $\mathbf{W}'$ (otherwise the path to Y would also be unblocked given $\mathbf{W}'$, which, together with edge $Z \ast\!\!\rightarrow Y$, would imply $U \not\perp\!\!\!\perp Z \,|\, \mathbf{W}' \cup Y$, contrary to the assumed). Number the colliders as $W_1, \ldots, W_k$, as they are encountered along π when starting from Y, such that $\pi = U \ast\!\!\rightarrow W_k \longleftrightarrow \ldots \longleftrightarrow W_2 \longleftrightarrow W_1 \longleftrightarrow Y$. By induction: if there is no edge between W_1 and Z in $\mathcal{G}'$, then they are (minimally) conditionally independent given some set $\mathbf{Q}_1 \subset \mathbf{W}$ (possibly empty), but dependent given Y, as the paths from both W_1 and Z into Y are not blocked by any node from $\mathbf{W}$ (as a bi-directed edge in $\mathcal{G}'$, resp. (direct) edge into Y), and so the lemma is

satisfied. If not, i.e. if there is an edge in $\mathcal{G}'$, then this edge is *out of* W_1, otherwise the path $\langle U, W_k, .., W_1, Z \rangle$ would be unblocked relative to $\mathbf{W}$, making U and Z dependent given $\mathbf{W}$, contrary the given. But then for W_2, if there is no edge between W_2 and Z in $\mathcal{G}'$, then $W_2 \perp\!\!\!\perp Z \,|\, [\mathbf{Q}_2]$, with $W_1 \in \mathbf{Q}_2$, because it is the only node from $\mathbf{W}$ that blocks the trek $W_2 \longleftrightarrow W_1 \longrightarrow Z$. But that also means that the path from W_2 to Y is unblocked given $\mathbf{Q}_2$, and so $W_2 \not\perp\!\!\!\perp Z_2 \,|\, \mathbf{Q}_2 \cup Y$. If not, then the edge to Z is (again) *out of* W_2, otherwise $U \not\perp\!\!\!\perp Z \,|\, \mathbf{W}$, contrary the given. This applies to all successive colliders W_i on the path π. But if all, up to and including W_k, have an edge in $\mathcal{G}'$ into Z, then no unblocked path between U and Z implies that W_k is needed to block $U \ast\!\!\rightarrow W_k \longrightarrow Z$, and so *all* W_i on π are in $\mathbf{W}'$, implying an unblocked path to Y, and so $U \not\perp\!\!\!\perp Z \,|\, \mathbf{W}' \cup Y$. $\qquad \square$

Now we can finally show:

Lemma 7.6 The arrowhead at Y from $\mathcal{R}$2a is covered by cases (1) and (2).

Proof. Assuming no violations before $\mathcal{R}$2a fires, then if the arrowhead at $X \ast\!\!\rightarrow Y$ originates from case (2), then the edge appears as $X \longrightarrow Y$, and is therefore also an instance of $\mathcal{R}$2b, which we already found to be valid. If $X \ast\!\!\rightarrow Y$ originates from case (1), then there is a minimal $U \perp\!\!\!\perp V \,|\, [\mathbf{W}]$, with $X \in (U \cup V \cup \mathbf{W})$, and the node Y creates $U \not\perp\!\!\!\perp V \,|\, \mathbf{W} \cup Y$. By lemma 7.3 there are now three cases for node Z:

(a) $Z \in (U \cup V \cup \mathbf{W})$,

(b) $Z \perp\!\!\!\perp U \,|\, \mathbf{W}$, (and/or $Z \perp\!\!\!\perp V \,|\, \mathbf{W}$)

(c) $U \not\perp\!\!\!\perp V \,|\, \mathbf{W} \cup Z$.

For case (a), both X and Z are in $(U \cup V \cup \mathbf{W})$, and so if rule (1) applies to $X \ast\!\rightarrow Y$ it also applies to $Z \ast\!\rightarrow Y$. Case (c) cannot occur, as that would imply $Z \not\!\Rrightarrow (U \cup V \cup \mathbf{W} \cup \mathbf{S})$ by lemma 1.2, with $X \in (U \cup V \cup \mathbf{W})$, while $\mathcal{R}$2a has $Z \longrightarrow X$.

For the remaining case (b), w.l.o.g. we assume $U \perp\!\!\!\perp Z \,|\, \mathbf{W}$. Lemma 7.4 implies $U \not\perp\!\!\!\perp Y \,|\, \mathbf{W}$ which, together with $Z \ast\!\!\rightarrow Y$, implies $U \not\perp\!\!\!\perp Z \,|\, \mathbf{W} \cup Y$, because the unblocked path from U to Y given $\mathbf{W}$ cannot contain Z, as that would create an unblocked path from U
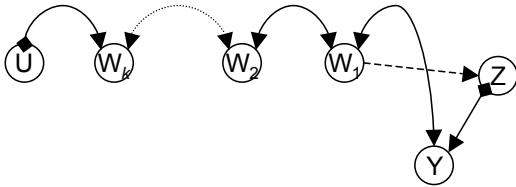

Figure 3: Configuration for Lemma 7.5

via Z to X given $\mathbf{W}$, contrary $U \perp\!\!\!\perp Z \,|\, \mathbf{W}$. Then from lemma 7.5 it follows that there is at least one minimal conditional independence between Z and some node from $(U \cup \mathbf{W})$ that is destroyed by conditioning on Y. Therefore, the arrowhead $Z \ast\!\!\rightarrow Y$ is then covered by case (1). $\qquad \square$
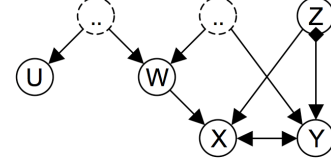

Figure 4: Example of non-minimal case (b) in lemma 7.5.

Example. Figure 4 shows an instance of case (b) for $\mathcal{R}$2a where the initial separating set is not minimal. Here, $\mathcal{R}$2a applies to $Z \ast\!\!\rightarrow Y$, after $X \ast\!\!\rightarrow Y$ is derived via case (1) from $U \perp\!\!\!\perp X \,|\, [W]$ with $U \not\perp\!\!\!\perp X \,|\, W \cup Y$ (the origin of the edge $Z \longrightarrow X$ is not depicted). By lemma 7.3, for node Z indeed $U \perp\!\!\!\perp Z \,|\, W$ holds (case b), but not as a minimal independence, as $U \perp\!\!\!\perp Z \,|\, [\varnothing]$. As a result, edge $Z \ast\!\!\rightarrow Y$ does not follow from case (1) applied to this combination of nodes as conditioning on Y does not make U and Z dependent, i.e. $U \perp\!\!\!\perp Z \,|\, Y$. However, as in the proof of lemma 7.6, Z is minimally conditionally independent of 'eliminated' node W, but dependent when conditioning on Y. Therefore, case (1) applies to $W \perp\!\!\!\perp Z \,|\, [\varnothing]$ and $W \not\perp\!\!\!\perp Z \,|\, Y$, from which follows that $Z \ast\!\!\rightarrow Y$.

We can now complete the proof of the main invariant arrowhead lemma:

Proof of lemma 7. Follows from the arrowhead completeness of rules $\mathcal{R}$0b-$\mathcal{R}$4b, the fact that after $\mathcal{R}$0a the theorem holds (no arrowheads), in combination with the proof in lemmas 7.1-7.6 that none of the rules $\mathcal{R}$0b-$\mathcal{R}$4b introduces a violation of the lemma, if there was no violation prior to firing of the rule. $\qquad \square$

Next we continue with the invariant tails.

Lemma 9 (Invariant tails). In a PAG $\mathcal{P}$, all invariant tails $Z \longrightarrow\!\ast Y$ from graphical orientation rules $\mathcal{R}$4a, $\mathcal{R}$5, $\mathcal{R}$7, $\mathcal{R}$9, and $\mathcal{R}$10 are instances of:

(2b): $X \perp\!\!\!\perp Y \,|\, [\mathbf{W} \cup Z]$, with $X \Rightarrow (Z \cup \mathbf{S})$ from either case (3) or another instance of (2b),

(3): $U \perp\!\!\!\perp V \,|\, [\mathbf{W} \cup W]$, with two transitive relations $[W, U, .., Y] + [W, V, .., Y]$, and $Z \in \{U, V, W\}$,

(4): $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z} \cup Z]$, with a Z an inferred blocking node in $[X, Z_1, .., Z_k, Z, Y]$, together with $Z_k \Rightarrow (Y \cup \mathbf{S})$ from either case (2) or case (4).

5

Proof. Case (2b) covers rule $\mathcal{R}7$, and is so named because of its similarity/overlap with case(2) for $\mathcal{R}1$. Case (3) covers all instances of rules $\mathcal{R}5$, $\mathcal{R}9$, and $\mathcal{R}10$, and case (4) accounts for tails from orientation rule $\mathcal{R}4a$.

All three cases are sound:
(2b): By lemma 3, $X \perp\!\!\!\perp Y \,|\, [\mathbf{W} \cup Z]$ gives $(Z \Rightarrow X) \vee (Z \Rightarrow Y) \vee (Z \Rightarrow \mathbf{S})$. Combined with $(X \Rightarrow Z) \vee (X \Rightarrow \mathbf{S})$ this gives $(Z \Rightarrow Y) \vee (Z \Rightarrow \mathbf{S})$, and so a tail at Z if it has an edge to Y in $\mathcal{P}$.
(3): Idem, $U \perp\!\!\!\perp V \,|\, [\mathbf{W} \cup W]$ gives $(W \Rightarrow U) \vee (W \Rightarrow V) \vee (W \Rightarrow \mathbf{S})$. From corollary 8, the transitive relations give $(W \Rightarrow \{U \cup \mathbf{S}\}) \vdash (W \Rightarrow \{Y \cup \mathbf{S}\})$, and $(W \Rightarrow \{V \cup \mathbf{S}\}) \vdash (W \Rightarrow \{Y \cup \mathbf{S}\})$. Substituting these two in the first then gives $(W \Rightarrow Y) \vee (W \Rightarrow \mathbf{S})$. This holds for all nodes on the two transitive chains, hence if $Z \in \{U, V, W\}$, then $(Z \Rightarrow Y) \vee (Z \Rightarrow \mathbf{S})$, and therefore a tail $Z \relbar\!\!* Y$, if they are connected in $\mathcal{P}$.
(4): By lemma 4, as Z is an inferred blocking node between X and Y given $\mathbf{Z}$, there is a $Z_k \in \mathbf{Z}$ such that $Z \Rightarrow Z_k \vee Z \Rightarrow Y \vee Z \Rightarrow \mathbf{S}$. Together with the given $Z_k \Rightarrow Y \vee Z_k \Rightarrow \mathbf{S}$, this reduces to $Z \Rightarrow Y \vee Z \Rightarrow \mathbf{S}$, and hence an invariant tail $Z \longrightarrow Y$.

In rule $\mathcal{R}7$, X and Y are nonadjacent, so conditionally independent given some set, and Z as a noncollider between the two is needed in all such sets, hence $X \perp\!\!\!\perp Y \,|\, [\mathbf{W} \cup Z]$. Only rules $\mathcal{R}6$ and $\mathcal{R}7$ can produce the required $X \relbar\!\!\circ Z$ edge to trigger $\mathcal{R}7$, however, every (chain of) $\mathcal{R}7$ orientations needs to start from an instance of $\mathcal{R}6$. Rule $\mathcal{R}6$ implies identifiable selection bias on X (undirected edges to other nodes), and so, if it triggers $\mathcal{R}7$ then this satisfies case (2b), and therefore any subsequent tail oriented by $\mathcal{R}7$ as well.

Rule $\mathcal{R}5$ triggers on an uncovered circle path. In Figure 4, let U be the node next to X on the circle path (U could be W), so that we have $Z \perp\!\!\!\perp U \,|\, [X \cup ..]$,

as Z and U by definition not adjacent. Furthermore, there are two transitive relations $[X, U, .., W, Y]$ and $[X, Z, Y]$ that are both from X to Y, and so $\mathcal{R}5$ satisfies case (3) and gives (among others) $Z \relbar Y$.
Rule $\mathcal{R}9$ similar: now $Z \perp\!\!\!\perp W \,|\, [X \cup ..]$, with two transitive relations $[X, W, .., Y]$ and $[X, Z, Y]$ that are both from X into Y, and so $\mathcal{R}9$ satisfies case (3) and will orient $Z \relbar Y$.
Rule $\mathcal{R}10$ idem: now $V \perp\!\!\!\perp S \,|\, [Z \cup ..]$, with two transitive relations $[Z, S, .., X, Y]$ and $[Z, V, .., W, Y]$ that are both from Z into Y, and so $\mathcal{R}10$ satisfies case (3) and will orient $Z \relbar Y$.

In rule $\mathcal{R}4a$, from the description of the graphical orientation, it follows that X and Y are non-adjacent in $\mathcal{P}$, and that all nodes $Z_1, .., Z_k, Z$, see also Figure 1, are needed to make them independent, and hence appear in the set $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z} \cup Z]$. Furthermore, each neighboring node in the sequence is adjacent in the graph, so not separated by any set, let alone a subset from $\mathbf{Z}$. All nodes Z_i have arrowheads at edges to their neighbors in the sequence, implying non-ancestorship, so no causal relation to either. Therefore Z is an inferred blocking node. The tail at $Z_k \longrightarrow Y$ implies $Z_k \Rightarrow (Y \cup \mathbf{S})$, and so $\mathcal{R}4a$ satisfies case (4) (in fact, even stronger, as identifiable $Z_k \Rightarrow Y$)
By construction of the discriminating path, all nodes Z_i in the sequence, except perhaps Z_1, also satisfy the conditions in case (4). For Z_2, the arc $Z_1 \longrightarrow Y$ follows from case (2). For Z_3, the invariant arc $Z_2 \longrightarrow Y$ therefore satisfies case (4) (although it may also be derived in other ways as well). Similar for all subsequent nodes up to Z_k. $\square$

Corollary 10 (Identifiable selection). In a PAG $\mathcal{P}$, all identifiable selection nodes $X \Rightarrow \mathbf{S}$ are covered by case (3), in the form of a minimal independence with two transitive relations back to itself.

Figure 5: Rules $\mathcal{R}5 - \mathcal{R}10$, tail orientation rules

Proof. Identifiable selection bias $X \Rightarrow \mathbf{S}$ corresponds to a node with an undirected edge. Only $\mathcal{R}5$, $\mathcal{R}6$, and $\mathcal{R}7$ can produce undirected edges.
When the transitive relations reach all the way back to the node Z from the initial minimal conditional independence, then the conclusion becomes $(Z \Rightarrow Z) \vee (Z \Rightarrow \mathbf{S})$, which (irreflexivity) reduces to $(Z \Rightarrow \mathbf{S})$. In other words, then there is *identifiable selection bias* on Z, and therefore also on all other nodes involved in the transitive relation (including the $\mathbf{U}_i$s). This is what happens in $\mathcal{R}5$. Afterwards, if $\mathcal{R}6$ is always executed before $\mathcal{R}7$ when a new undirected edge is found, then $\mathcal{R}6$ will never (need to) identify a new selection node, as it produces only tails on nodes that are already established selection nodes. That leaves just $\mathcal{R}7$. It is possible that part of the transitive relation in case (3) is traversed in both ways (that is where
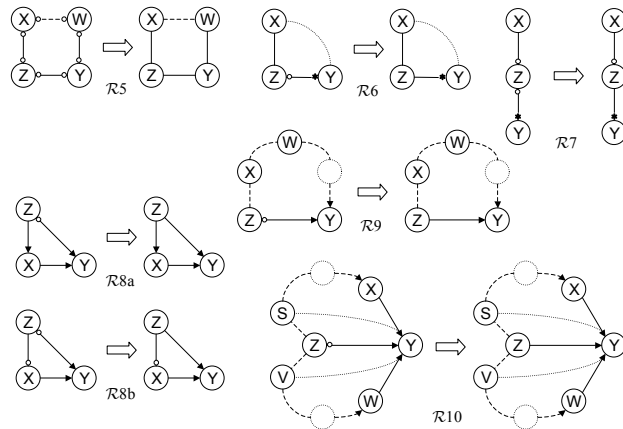
the 'not necessarily disjoint' part in the definition of transitive relation comes in). This occurs for single nodes that separate two nonchordal undirected subgraphs in $\mathcal{P}$. Then $\mathcal{R}7$ will orient $Z \,{-}{*}\, Y$ in the direction away from one undirected subgraph, and $Y \,{-}{*}\, Z$ when orienting in the direction away from the other subgraph, resulting in $Y - Z$, and so identifiable selection bias on both Y and Z. These are also the *only* ways in which undirected edges can be created by the orientation rules. $\square$

6 Reconstructing the PAG

Lemma 12 (Single minimal independence). In the structure independent approach, finding a *single*, arbitrary $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}]$, for each pair of nodes (X,Y) in the graph (if it exists) is sufficient to find all invariant features of the PAG.

Proof sketch. This stems from the fact that the graphical orientation rules are defined on sets of adjacent nodes, which ensures that most nodes are almost always needed to separate two nonadjacent nodes in the same rule, and so will be found as part of the separating set, no matter how large/variable the set of nodes to block all paths between the two can be. Note: once a minimal set is found for a pair of nodes, then all remaining nodes are checked to see if including them destroys the independence (so lemma 3, item 1 applies). $\square$

In the detailed proof per rule, we use the following result for rule $\mathcal{R}2a$:

Lemma 12.1. In an ancestral graph $\mathcal{G}$, if there are (sets of) nodes U, Y, Z and $\mathbf{W}$, such that $U \perp\!\!\!\perp Z \,|\, [\mathbf{W}]$ and $U \not\!\perp\!\!\!\perp Z \,|\, \mathbf{W} \cup Y$, with $Z \,{*}{\rightarrow}\, Y$ in $\mathcal{G}$, then there is a node V (possibly $V = U$), such that for all sets $\mathbf{Q}$ for which $V \perp\!\!\!\perp Z \,|\, [\mathbf{Q}]$ it holds that $V \not\!\perp\!\!\!\perp Z \,|\, \mathbf{Q} \cup Y$.
In words: if conditioning on a node Y destroys (unblocks) some minimal conditional independence for a neighbouring node Z, then it does so in *all* minimal independencies between Z and at least one node in $\mathcal{G}$.

Proof. Follows along the lines of lemma 7.5. From that proof, there is an unblocked path π in $\mathcal{G}'$ (the graph $\mathcal{G}$, marginalized over $\{U, Y, Z\} \cup \mathbf{W}$, of the form $\pi = U \,{*}{\rightarrow}\, W_k \longleftrightarrow \ldots \longleftrightarrow W_2 \longleftrightarrow W_1 \longleftrightarrow Y$ that is into Y given $\mathbf{W}$, for some $k \geq 0$. Here we consider the corresponding path(s) θ in $\mathcal{G}$, where the colliders on the path are now indicated by $\{U_1, .., U_m\}$. (Note that edges in $\mathcal{G}'$ may correspond to multiple unblocked paths relative to $\mathbf{W}$ in $\mathcal{G}$, and that θ may contain different nodes than π, including other (ancestors of) colliders from $\mathbf{W}$).

So, let $\theta = U \,{*}{\rightarrow}\, U_m \longleftrightarrow \ldots \longleftrightarrow U_2 \longleftrightarrow U_1 \longleftrightarrow Y$ be an unblocked path in $\mathcal{G}$ that is into Y given $\mathbf{W}$. We look at nodes V along θ, starting from Y, and try to find one that does not have a link to Z. Suppose $V = V_1$ is encountered on the first leg (trek) $U_1 \leftarrow\!{*} .. {*}\!\rightarrow Y$. If V_1 does not have an edge to Z in $\mathcal{G}$, then there is some $\mathbf{Q}$ such that $V_1 \perp\!\!\!\perp Z \,|\, [\mathbf{Q}]$, while also $V_1 \not\!\perp\!\!\!\perp Z \,|\, \mathbf{Q} \cup Y$ (both edges to Y), and so the lemma is satisfied. If there is a link, then it can only be of the form $U_1 \leftarrow\!{*} .. {*}\!\rightarrow V_1 \leftarrow\!{*}\, Z$, otherwise there would be an unblocked path between U and Z given $\mathbf{W}$.
For a second node, V_2, a similar story holds: if there is no edge $V_2 - Z$ in $\mathcal{G}$, then there is some $\mathbf{Q}$ such that $V_2 \perp\!\!\!\perp Z \,|\, [\mathbf{Q}]$; but note that now $V_1 \notin \mathbf{Q}$, as the edges from both V_2 and Z are into V_1, and therefore also/again $V_2 \not\!\perp\!\!\!\perp Z \,|\, \mathbf{Q} \cup Y$ (unblocked path resp. edge into Y given $\mathbf{Q}$), and so the lemma is satisfied. If there is an edge, then again it must be of the form $U_1 \leftarrow\!{*} .. {*}\!\rightarrow V_2 \leftarrow\!{*}\, Z$, otherwise there would be an unblocked path to Z, and we can continue until we reach U_1. At that point, again, if U_1 has no edge to Z, then there is some $U_1 \perp\!\!\!\perp Z \,|\, [\mathbf{Q}]$, with none of the V_i encountered on the first leg in $\mathbf{Q}$ (as V_j is not ancestor of either U_1 or Z, from which, by $V_j \longrightarrow .. \longrightarrow V1$, follows that neither are any of the other V_i), and so there are unblocked paths from U_1 and Z into Y given $\mathbf{Q}$. If it has an edge, then, contrary the edges from Z into V_i (if any), it must be an edge *out of* $U_1 \longrightarrow Z$, otherwise there would be an unblocked path given $\mathbf{W}$. Continuing with the second leg, $U_2 \leftarrow\!{*} .. {*}\!\rightarrow U_1$, we now find that a node V encountered in going from U_1 to U_2 along θ (if any) *cannot* have a direct edge to Z without creating an unblocked path from U to Z relative to $\mathbf{W}$ (either as noncollider between U_2 and Z, or as collider that is ancestor of U_1). So then $V \perp\!\!\!\perp Z \,|\, [\mathbf{Q}]$, with $U_1 \in \mathbf{Q}$ (as it is the only node blocking the path $V \,{*}\!\rightarrow U_1 \longrightarrow Z$). But like before, none of the previous V_i (if any) are part of $\mathbf{Q}$: by contradiction, if $V_i \in \mathbf{Q}$, then $V_i \longrightarrow V$ (because $V_i \notin An(Z)$), so if $U_2 \longleftrightarrow V \longrightarrow U_1$ along θ, then the 'top' V_j would be ancestor of U_1, contrary the arrowhead $U_1 \longleftrightarrow V_j$, and if $U_2 \longleftarrow V \,{*}\!\rightarrow U_1$ along θ, then the path via $U_2 \longleftarrow V \longleftarrow V_i \leftarrow\!{*}\, Z$ would be unblocked (end-of-by-contradiction). So, if none of the V_i are part of $\mathbf{Q}$, then the paths from both V and Z to Y are unblocked given $\mathbf{Q}$, and the lemma is satisfied.
If no node V on $U_2 \longleftrightarrow U_1$, then again, for U_2 there is either a minimal independence that satisfies the lemma or an edge $U_2 \longrightarrow Z$. This can be repeated along θ until a node is found or the final node U is reached. At that point, if no other node has been found before, it can be applied to any set for which $U \perp\!\!\!\perp Z \,|\, [\mathbf{Q}]$, as all other colliders $U_i \in \mathbf{Q}$, are needed to separate U and Z, but none of the V_i, and so the path to Y is always unblocked. $\square$

It turns out it is actually easier to use a more restricted variant of lemma 7 to prove lemma 12:

Lemma 12.2 (Restricted lemma 7). In a PAG $\mathcal{P}$, all invariant arrowheads $Z \ast\!\!\rightarrow Y$ are instances of

(1′) $U \not\!\!\perp V \mid \mathbf{W} \cup [Y]$, created from $U \perp\!\!\!\perp V \mid [\mathbf{W}]$, with $Z \in \{U, V, \mathbf{W}\}$, and where for all sets $\mathbf{W}' : U \perp\!\!\!\perp V \mid [\mathbf{W}']$ the paths from U and V to Y are unblocked relative to $\mathbf{W}'$, and either $Z \in \{U, V\}$ or (necessarily) $Z \in \mathbf{W}'$.

(2′) $X \perp\!\!\!\perp Y \mid [\mathbf{W}]$ with $Z \in \mathbf{W}$, and $Z \not\Rrightarrow (X \cup \mathbf{S})$ from either case (1′) or case (2′), and Z in all sets $\mathbf{W}' : X \perp\!\!\!\perp Y \mid [\mathbf{W}']$.

In words: case (1) only needs to be applied to instances where Z is always part of the minimal conditional independence, and that will always be unblocked when conditioning on Y. Case (2) also only needs to be applied to instances where Z is always part of the minimal conditional independence.

Proof. For each arrowhead rule:

$\mathcal{R}0b$ fires on any (minimal) conditional independence $X \perp\!\!\!\perp Y \mid \mathbf{W}$ between X and Y, and for any such $\mathbf{W}$, including Z will unblock the path $\langle X, Z, Y \rangle$, so case (1′) applies,

$\mathcal{R}1$ node Z is part of *any* set (minimal or not) that separates X and Y, and so case (2′) applies,

$\mathcal{R}3$ similar to $\mathcal{R}0b$, fires on a node W that is part of all sets separating X and Y, and including Z will unblock the path $\langle X, Z, Y \rangle$, and so case (1′) applies,

$\mathcal{R}4b$ (arrowheads at Z) all nodes $Z_1, \ldots, Z_k$ are part of all sets separating X and Y, and including Z then makes them dependent, so case (1′) applies,

$\mathcal{R}2b$ for instance (1a) in lemma 7.2, if case (1′) applies to $Z \ast\!\!\rightarrow X$, then it also applies to $Z \ast\!\!\rightarrow Y$, as X is never part of the minimal conditional independence involving Z, and so unblocked paths to X imply unblocked paths to Y; for instance (2a), Z is present in all sets that make X and U independent, and so also in all sets that make Y and U independent (as it implies $Z \longrightarrow Y$), and so case (2′) applies; for instance (2b), if $Y \not\!\!\perp U \mid [\mathbf{W} \cup Z]$ holds for all sets for which $X \perp\!\!\!\perp U \mid [\mathbf{W} \cup Z]$, then it is an instance of case (1′) with $V \equiv X$ and $\mathbf{W} \equiv (\mathbf{W} \cup Z)$, and so for all $\mathbf{W} \cup Z$ there are unblocked paths from X and U into Y, resulting in $X \not\!\!\perp U \mid \mathbf{W} \cup Z \cup Y$. If not, then there is *some* $\mathbf{W}'$ for which $X \perp\!\!\!\perp U \mid [\mathbf{W}' \cup Z]$ and *not* $Y \not\!\!\perp U \mid [\mathbf{W}' \cup Z]$. But as Z is needed in all sets that block a path $\pi = U .. \ast\!\!\rightarrow Z \longrightarrow X$ between U and X, it means that Z is also needed in all sets that separate U and Y, because if there is any remaining unblocked path π from U to either X or Z, then $\pi+$ either $X \longrightarrow Y$ or $Z \longrightarrow Y$ is

an unblocked path from U into Y. Therefore Z is also needed in all sets that separate U and Y, which, together with $Z \not\Rrightarrow (U \cup \mathbf{S})$, implies that it is an instance of case (2′).

$\mathcal{R}4a/b$ (arrowhead at Y) instances of $\mathcal{R}2b$ with $Z_k = Y$,

$\mathcal{R}2a$ if the arrowhead between X and Y originates from case (2′) then $X \longrightarrow Y$, and so is an instance of $\mathcal{R}2b$. If not, then the arrowhead $X \ast\!\!\rightarrow Y$ originates from case (1′) with node X in $\mathcal{R}2a$ in the role of Z in (1′).

Now, if Z, like X, is also a *necessary* member of the minimal independence $U \perp\!\!\!\perp V \mid [\mathbf{W}]$, i.e. $Z \in \{U, V\}$ or $\forall \mathbf{W}', U \perp\!\!\!\perp V \mid [\mathbf{W}'] : Z \in \mathbf{W}'$, then case (1′) also applies immediately to Z.

If Z is *not* necessary, then there is some $U \perp\!\!\!\perp V \mid [\mathbf{W}']$, with $Z \notin (U \cup V \cup \mathbf{W}')$, for which $U \not\!\!\perp V \mid \mathbf{W}' \cup Y$. For node Z then instance (b) in lemma 7.3 applies, say as $U \perp\!\!\!\perp Z \mid \mathbf{W}'$, as instance (a) is excluded by the assumed $Z \notin (U \cup V \cup \mathbf{W}')$, and instance (c) still cannot occur, as $Z \longrightarrow X$. But then also $U \not\!\!\perp Z \mid \mathbf{W}' \cup Y$, as both U and Z have unblocked paths into Y relative to $\mathbf{W}'$ (by lemma 7.4, applied on the $U \not\!\!\perp V \mid \mathbf{W}' \cup Y$ from case (1′), together with edge $Z \ast\!\!\rightarrow Y$). But then by lemma 7.5 there is also a $W \perp\!\!\!\perp Z \mid [\mathbf{Q}]$ with $W \not\!\!\perp Z \mid \mathbf{Q} \cup Y$, and therefore, by lemma 12.1 there is also some node Q for which for every set $Q \perp\!\!\!\perp Z \mid [\mathbf{Q}']$ also $Q \not\!\!\perp Z \mid \mathbf{Q}' \cup Y$, i.e. then case (1′) is also satisfied for the arrowhead $Z \ast\!\!\rightarrow Y$. $\square$

(Note that $\mathcal{R}2b$ is not simply an instance of lemma 12.1, as that only says that X has a conditional independence with some node $V \in \mathbf{W}$ that will always be destroyed by conditioning on Y. It does not guarantee that Z is also always a part of this set.)

So all arrowhead rules are covered by lemma 12.2. As the two cases in lemma 12.2 are just a restricted form of the cases in lemma 7, it follows that all rules are also covered by lemma 7 if just a (any) single minimal independence is found between each pair (if it exists). We can now complete the proof that a single minimal independence suffices to find all invariant marks:

Proof of lemma 12. By lemma 12.2, all invariant arrowheads are instances of cases (1′) and (2′), and so will always be found if at least one minimal conditional independence is found between each pair of nodes (if it exists). For the remaining invariant tails: *For each rule:*

$\mathcal{R}5$ The circle path corresponds to a transitive chain of minimal conditional independencies where each node is necessarily part of any set that separates its two neighbors.

$\mathcal{R}7$ Similar to $\mathcal{R}1$.

$\mathcal{R}9$ Node X is part of every conditional independence

between Z and W; the same holds for the successive nodes in both transitive relations from X via Z to Y and from X via W to Y.

$\mathcal{R}10$ Similar to $\mathcal{R}5/\mathcal{R}9$.

Remaining rules $\mathcal{R}6$, $\mathcal{R}8a$, and $\mathcal{R}8b$ do not require any separate independence statement. Therefore all orientation rules that trigger on an instance of case (1) do so for at least one that is part of a set that is present in all minimal conditional independencies for a given pair of nodes. $\qquad\square$

Theorem 1. The Logical Causal Inference (LoCI) algorithm is sound and complete.

Proof. Soundness follows from the validity of the lemmas 3 and 4, that produce the logical statements in the list L, in combination with the causal logic rules in proposition 1. Completeness follows from the fact that all rules are instancess of cases (1)-(4) (lemmas 7 and 9), for a single, arbitrary minimal independence between nodes, in combination with subsequent dependencies (lemma 12), the fact that all logical inference in each of the cases (1)-(4) is covered by proposition 1, the fact that case (1) and (2) will find all required non-ancestor relations / invariant arrowheads (lemma 7, (Zhang, 2008)), needed to obtain the only remaining piece of information (inferred blocking node for case (4) from lemma 4). After running the logical rules on this set of statements to completion, all invariant edge marks have been found and can be transferred to the PAG. $\qquad\square$

As a final step we demonstrate that for each case (1)-(4), the logical inference steps in the LoCI algroithm do indeed keep the simple form of the list of statements in section 3.1 in (Claassen and Heskes, 2011), i.e. either a statement on the absence of a specific causal causal relation, or a disjunction of possible causal relations from one variable to at most two other variables, and/or the selection set $\mathbf{S}$.

Case (1): Follows directly from lemma 3, item 2. (Note: the restriction to $U \perp\!\!\!\perp V \mid [\mathbf{W}]$ plays no role, as it is only states that the (crucial) dependency by Y will be found from a *minimal* conditional independence).

$$1: \quad Y \nRightarrow Z \quad \wedge\ Y \nRightarrow .. \quad \wedge\ Y \nRightarrow \mathbf{S}$$

Case (2): A minimal independence in combination with already inferred information on the absence of a causal relation.

$$
\begin{array}{llll}
1: & Z \Rightarrow X & \vee\ Z \Rightarrow Y & \vee\ Z \Rightarrow \mathbf{S} \\
2: & Z \nRightarrow X & \wedge & \wedge\ Z \nRightarrow \mathbf{S} \\
1': & & Z \Rightarrow Y & \\
3: & & Y \nRightarrow Z &
\end{array}
$$

Case (2b): Idem, but now in combination with an already inferred causal relation and/or selection bias on a specific node.

$$
\begin{array}{llll}
1: & Z \Rightarrow X & \vee\ Z \Rightarrow Y & \vee\ Z \Rightarrow \mathbf{S} \\
2: & X \Rightarrow Z & \wedge & \wedge\ X \Rightarrow \mathbf{S} \\
\vdash & Z \Rightarrow Z & \vee\ Z \Rightarrow Y & \vee\ Z \Rightarrow \mathbf{S} \\
1': & & Z \Rightarrow Y & \vee\ Z \Rightarrow \mathbf{S}
\end{array}
$$

Case (3): A minimal independence involving Z, with both alternatives leading to Y.

$$
\begin{array}{llll}
1: & W \Rightarrow U & \vee\ W \Rightarrow V & \vee\ W \Rightarrow \mathbf{S} \\
2: & U \Rightarrow W & \vee\ U \Rightarrow Y & \vee\ U \Rightarrow \mathbf{S} \\
3: & V \Rightarrow W & \vee\ V \Rightarrow Y & \vee\ V \Rightarrow \mathbf{S} \\
\vdash & W \Rightarrow Y & \vee\ W \Rightarrow V & \vee\ W \Rightarrow \mathbf{S} \\
\vdash & W \Rightarrow Y & \vee\ W \Rightarrow Y & \vee\ W \Rightarrow \mathbf{S} \\
4: & & W \Rightarrow Y & \vee\ W \Rightarrow \mathbf{S}
\end{array}
$$

This for $Z = W$; cases $Z = U/V$ go the same.

Case (4): An inferred blocking node Z, with an earlier/afterwards established causal relation to Y or selection bias.

$$
\begin{array}{lll}
1: & Z \Rightarrow Z_k & \vee\ Z \Rightarrow Y \quad \vee\ Z \Rightarrow \mathbf{S} \\
2: & & Z_k \Rightarrow Y \quad \vee\ Z_k \Rightarrow \mathbf{S} \\
1' & & Z \Rightarrow Y \quad \vee\ Z \Rightarrow \mathbf{S}
\end{array}
$$

References

R.A. Ali, T. Richardson, P. Spirtes, and J. Zhang. Towards characterizing markov equivalence classes for directed acyclic graphs with latent variables. In *Proc. of the 21st Conference on Uncertainty in Artificial Intelligence*, pages 10–17, 2005.

T. Claassen and T. Heskes. Causal discovery in multiple models from different experiments. In *NIPS-23*, pages 415–423. 2010.

T. Claassen and T. Heskes. A logical characterization of constraint-based causal discovery. In *Proc. of the 27th Conference on Uncertainty in Artificial Intelligence*, 2011.

T. Richardson and P. Spirtes. Ancestral graph Markov models. *Ann. Stat.*, 30(4):962–1030, 2002.

P. Spirtes, C. Meek, and T. Richardson. An algorithm for causal inference in the presence of latent variables and selection bias. In *Computation, Causation, and Discovery*, pages 211–252. 1999.

P. Spirtes, C. Glymour, and R. Scheines. *Causation, Prediction, and Search*. The MIT Press, Cambridge, Massachusetts, 2nd edition, 2000.

J. Zhang. On the completeness of orientation rules for causal discovery in the presence of latent confounders and selection bias. *Artificial Intelligence*, 172(16-17):1873 − 1896, 2008.