ALTERNATIVE MARKOV AND CAUSAL PROPERTIES FOR ACYCLIC DIRECTED MIXED GRAPHS - SUPPLEMENTARY MATERIAL

This supplement contains the proofs of the theorems in the manuscript.

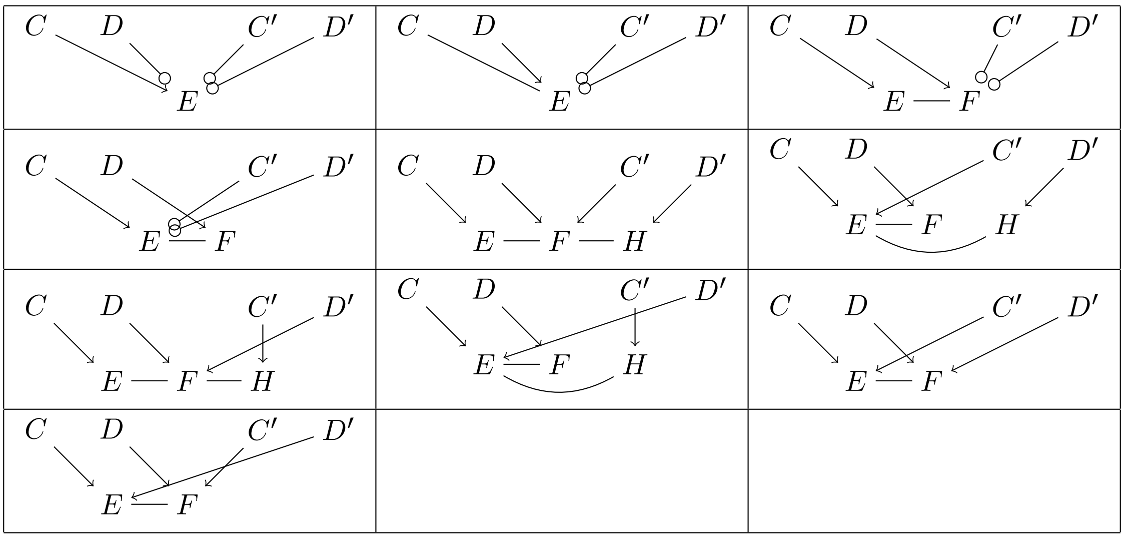
Lemma 1. *If there is a path ρ in an ADMG G between $A \in X$ and $B \in Y$ such that (i) no non-collider C on ρ is in Z unless $A - C - B$ is a subpath of ρ and $Pa_G(C) \setminus Z \neq \emptyset$, and (ii) every collider on ρ is in $An_G(X \cup Y \cup Z)$, then there is a path in G connecting a node in X and a node in Y given Z.*

Proof. Suppose that ρ has a collider C such that $C \in An_G(D) \setminus An_G(Z)$ with $D \in X$, or $C \in An_G(E) \setminus An_G(Z)$ with $E \in Y$. Assume without loss of generality that $C \in An_G(D) \setminus An_G(Z)$ with $D \in X$ because, otherwise, a symmetric argument applies. Then, replace the subpath of ρ between A and C with $D \leftarrow \ldots \leftarrow C$. Note that the resulting path (i) has no non-collider in Z unless $A - C - B$ is a subpath of ρ and $Pa_G(C) \setminus Z \neq \emptyset$, and (ii) has every collider in $An_G(X \cup Y \cup Z)$. Note also that the resulting path has fewer colliders than ρ that are not in $An_G(Z)$. Continuing with this process until no such collider C exists produces the desired result. □

Lemma 2. *Given an ADMG G, let ρ denote a shortest path in $G[A \cup B \cup Z]^a$ connecting two nodes A and B given Z. Then, a path in G between A and B can be obtained as follows. First, replace every augmented edge on ρ with an associated collider path in $G[A \cup B \cup Z]$. Second, replace every non-augmented edge on ρ with an associated edge in $G[A \cup B \cup Z]$. Third, replace any configuration $C - D \leftarrow F \rightarrow D - E$ produced in the previous steps with $C - D - E$.*

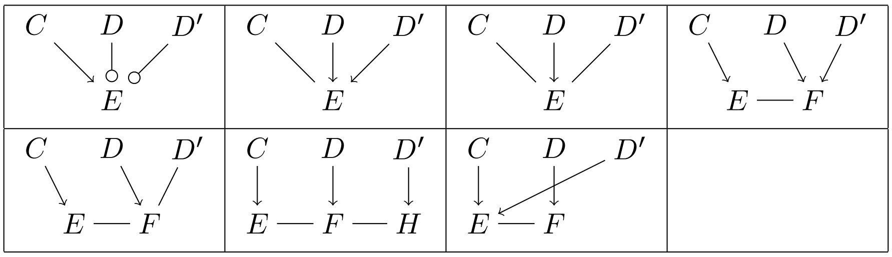
Proof. We start by proving that the collider paths added in the first step of the lemma either do not have any node in common except possibly one of the endpoints, or the third step of the lemma removes the repeated nodes. Suppose for a contradiction that $C - D$ and $C' - D'$ are two augmented edges on ρ such that their associated collider paths have in common a node which is not an endpoint of these paths. Consider the following two cases.

Case 1: Suppose that $D \neq C'$. Then, one of the following configurations must exist in $G[A \cup B \cup Z]$.



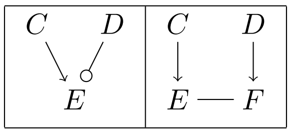
However, the first case implies that $C - D'$ is in $G[A \cup B \cup Z]^a$, which implies that replacing the subpath of ρ between C and D' with $C - D'$ results in a path in $G[A \cup B \cup Z]^a$ connecting A and B given Z that is shorter than ρ. This is a contradiction. Similarly for the fourth, sixth, seventh, eighth, ninth and tenth cases. And similarly for the rest of the cases by replacing the subpath of ρ between D and D' with $D - D'$.

Case 2: Suppose that $D = C'$. Then, one of the following configurations must exist in $G[A \cup B \cup Z]$.



However, the first case implies that $C - D'$ is in $G[A \cup B \cup Z]^a$, which implies that replacing the subpath of ρ between C and D' with $C - D'$ results in a path in $G[A \cup B \cup Z]^a$ connecting A and B given Z that is shorter than ρ. This is a contradiction. Similarly for the second, fourth and seventh cases. For the third, fifth and sixth cases, the third step of the lemma removes the repeated nodes. Specifically, it replaces $C - E \leftarrow D \rightarrow E - D'$ with $C - E - D'$ in the third case, $E - F \leftarrow D \rightarrow F - D'$ with $E - F - D'$ in the fifth case, and $E - F \leftarrow D \rightarrow F - H$ with $E - F - H$ in the sixth case.

It only remains to prove that the collider paths added in the first step of the lemma have no nodes in common with ρ except the endpoints. Suppose that ρ has an augmented edge $C - D$. Then, one of the following configurations must exist in $G[A \cup B \cup Z]$.



Consider the first case and suppose for a contradiction that E occurs on ρ. Note that $E \notin Z$ because, otherwise, ρ would not be connecting. Assume without loss of generality that E occurs on ρ before C and D because, otherwise, a symmetric argument applies. Then, replacing the subpath of ρ between E and D with $E - D$ results in a path in $G[A \cup B \cup Z]^a$ connecting A and B given Z that is shorter than ρ. This is a contradiction. Similarly for the second case. Specifically, assume without loss of generality that E occurs on ρ because, otherwise, a symmetric argument with F applies. Note that $E \notin Z$ because, otherwise, ρ would not be connecting. If E occurs on ρ after C and D, then replace the subpath of ρ between C and E with $C - E$. This results in a path in $G[A \cup B \cup Z]^a$ connecting A and B given Z that is shorter than ρ, which is a contradiction. If E occurs on ρ before C and D, then replace the subpath of ρ between E and D with $E - D$. This results in a path in $G[A \cup B \cup Z]^a$ connecting A and B given Z that is shorter than ρ, which is a contradiction. ☐

Lemma 3. *Let ρ denote a path in an ADMG G connecting two nodes A and B given Z. The sequence of non-colliders on ρ forms a path in $G[A \cup B \cup Z]^a$ between A and B.*

Proof. Consider the maximal undirected subpaths of ρ. Note that each endpoint of each subpath is ancestor of a collider or endpoint of ρ, because ρ is connecting. Thus, all the nodes on ρ are in $G[A \cup B \cup Z]^a$. Suppose that C and D are two successive non-colliders on ρ. Then, the subpath of ρ between C and D consists entirely of colliders. Specifically, the subpath is of the form $C \multimap D$, $C \rightarrow D$, $C \rightarrow E \multimap D$ or $C \rightarrow E - F \leftarrow D$. Then, C and D are adjacent in $G[A \cup B \cup Z]^a$. ☐

Theorem 1. *There is a path in an ADMG G connecting a node in X and a node in Y given Z if and only if there is a path in $G[X \cup Y \cup Z]^a$ connecting a node in X and a node in Y given Z.*

Proof. We start by proving the only if part. Let ρ denote a path in G connecting $A \in X$ and $B \in Y$ given Z. By Lemma 3 the non-colliders on ρ form a path ρ^a between A and B in $G[X \cup Y \cup Z]^a$. Since ρ is connecting, every non-collider C on ρ is outside Z unless $D - C - E$ is a subpath of ρ and $Pa_G(C) \setminus Z \neq \emptyset$. To address the latter case, consider in turn each maximal subpath of ρ of the form $D - C_1 - \ldots - C_n - E$ such that $C_i \in Z$ for all $1 \leq i \leq n$. Note that ρ^a has a subpath $D' - C_1 - \ldots - C_n - E'$

where $D' = D$ unless D is a collider on ρ, i.e. $D' \to D - C$ is on ρ. Similarly for E and E'. Therefore, we replace this subpath of ρ^a with $D' - F_1 - \ldots - F_n - E'$ where $F_i \in Pa_G(C_i) \setminus Z$ for all $1 \le i \le n$. Then, ρ^a is connecting given Z. Note that ρ^a may be a route (e.g. if $F_1 = F_n$). Obtaining the desired path is trivial though. Finally, note that $D' - F_1$ is guaranteed to be in $G[X \cup Y \cup Z]^a$, because $F_1 \to C_1$ is in G with $C_1 \in Z$, and $D - C_1$ is in G when $D' = D$ or otherwise $D' \to D - C_1$ is in G with $D \in An_G(Z)$. Similarly for $F_n - E'$. Likewise, $F_i - F_{i+1}$ is guaranteed to be in $G[X \cup Y \cup Z]^a$ for all $1 \le i < n$, because $F_i \to C_i$ and $F_{i+1} \to C_{i+1}$ are in G with $C_i, C_{i+1} \in Z$.

To prove the if part, let ρ^a denote a shortest path in $G[X \cup Y \cup Z]^a$ connecting $A \in X$ and $B \in Y$ given Z. We can transform ρ^a into a path ρ in G as described in Lemma 2. Since ρ^a is connecting, no node on ρ^a is in Z and, thus, no non-collider on ρ is in Z. Finally, since all the nodes on ρ are in $G[X \cup Y \cup Z]^a$, it follows that every collider on ρ is in $An_G(X \cup Y \cup Z)$. To see it, note that if $C - D$ is an augmented edge in $G[X \cup Y \cup Z]^a$ then the colliders on any collider path associated with $C - D$ are in $An_G(X \cup Y \cup Z)$. Thus, by Lemma 1 there exist a node in X and a node in Y which are connected given Z in G. □

Theorem 2. *There is a path in an ADMG G connecting A and B given Z if and only if there is a route in G connecting A and B given Z.*

Proof. We prove the theorem for the following separation criterion, which is equivalent to criterion 2: A route is said to be connecting given $Z \subseteq V$ when

- every collider on the route is in Z, and
- every non-collider C on the route is outside Z unless $A - C - B$ is a subroute and $Pa_G(C) \setminus Z \neq \emptyset$.

The only if part is trivial. To prove the if part, let ρ denote a route in G connecting A and B given Z. Let C denote a node that occurs more than once in ρ. Consider the following cases.

Case 1: Assume that ρ is of the form $A \ldots D \to C \ldots C \to E \ldots B$. Then, $C \notin Z$ for ρ to be connecting given Z. Then, removing the subroute between the two occurrences of C from ρ results in the route $A \ldots D \to C \to E \ldots B$, which is connecting given Z.

Case 2: Assume that ρ is of the form $A \ldots D \to C \ldots C \multimap E \ldots B$. Then, $C \in An_G(Z)$ for ρ to be connecting given Z. Then, removing the subroute between the two occurrences of C from ρ results in the route $A \ldots D \to C \multimap E \ldots B$, which is connecting given Z.

Case 3: Assume that ρ is of the form $A \ldots D \leftarrow C \ldots C \ldots B$. Then, $C \notin Z$ for ρ to be connecting given Z. Then, removing the subroute between the two occurrences of C from ρ results in the route $A \ldots D \leftarrow C \ldots B$, which is connecting given Z.

Case 4: Assume that ρ is of the form $A \ldots D - C \ldots C \to E \ldots B$. Then, $C \notin Z$ for ρ to be connecting given Z. Then, removing the subroute between the two occurrences of C from ρ results in the route $A \ldots D - C \to E \ldots B$, which is connecting given Z.

Case 5: Assume that ρ is of the form $A \ldots D - C \ldots C \leftarrow E \ldots B$. Then, $C \in An_G(Z)$ for ρ to be connecting given Z. Then, removing the subroute between the two occurrences of C from ρ results in the route $A \ldots D - C \leftarrow E \ldots B$, which is connecting given Z.

Case 6: Assume that ρ is of the form $A \ldots D - C \ldots C - E \ldots B$ and $C \notin Z$. Then, removing the subroute between the two occurrences of C from ρ results in the route $A \ldots D - C - E \ldots B$, which is connecting given Z.

Case 7: Assume that ρ is of the form $A \ldots D - C \ldots C - E \ldots B$ and $C \in Z$. Then, ρ must actually be of the form $A \ldots D - C \leftarrow F \ldots C - E \ldots B$ or $A \ldots D - C - F \ldots C - E \ldots B$. Note that in the former case $F \notin Z$ for ρ to be connecting given Z. For the same reason, $Pa_G(C) \setminus Z \neq \emptyset$ in the latter case. Then, $Pa_G(C) \setminus Z \neq \emptyset$ in either case. Then, removing the subroute between the two occurrences of C from ρ results in the route $A \ldots D - C - E \ldots B$, which is connecting given Z.

Repeating the process above until no such node C exists produces the desired path. □

Theorem 3. *Given an ADMG G, there is a path in $G[X \cup Y \cup Z]^a$ connecting a node in X and a node in Y given Z if and only if there is a path in $(G[X \cup Y \cup Z]^m)^a$ connecting a node in X and a node in Y given Z.*

Proof. We start by proving the only if part. Suppose that there is path in $G[X \cup Y \cup Z]^a$ connecting a node in X and a node in Y given Z. We can then obtain a path in G connecting $A \in X$ and

$B \in Y$ given Z as shown in the proof of Theorem 1. In this path, replace with $C - D$ every subpath $C - V_1 - \ldots - V_n - D$ such that $C, D \in An_G(X \cup Y \cup Z)$ and $V_1, \ldots, V_n \notin An_G(X \cup Y \cup Z)$. The result is a path in $G[X \cup Y \cup Z]^m$. Moreover, the path connects A and B given Z. To see it, note that the resulting and original paths have the same colliders, and the non-colliders on the resulting path are a subset of the non-colliders on the original path. Then, there is path in $(G[X \cup Y \cup Z]^m)^a$ connecting A and B given Z.

To prove the if part, suppose that there is path in $(G[X \cup Y \cup Z]^m)^a$ connecting $A \in X$ and $B \in Y$ given Z. Suppose that the path contains an edge $C - D$ that is not in $G[X \cup Y \cup Z]^a$. This is due to one the following reasons.

Case 1: $C - V_1 - \ldots - V_n - D$ is in $G[X \cup Y \cup Z]$ with $V_1, \ldots, V_n \notin An_G(X \cup Y \cup Z)$. Then, $C - V_1 - \ldots - V_n - D$ is in $G[X \cup Y \cup Z]^a$.

Case 2: $C \to E - D$ is in $G[X \cup Y \cup Z]^m$, which means that $C \to E - V_1 - \ldots - V_n - D$ is in $G[X \cup Y \cup Z]$ with $V_1, \ldots, V_n \notin An_G(X \cup Y \cup Z)$. Then, $C - V_1 - \ldots - V_n - D$ is in $G[X \cup Y \cup Z]^a$.

Case 3: $C \to E - F \leftarrow D$ is in $G[X \cup Y \cup Z]^m$, which means that $C \to E - V_1 - \ldots - V_n - F \leftarrow D$ is in $G[X \cup Y \cup Z]$ with $V_1, \ldots, V_n \notin An_G(X \cup Y \cup Z)$. Then, $C - V_1 - \ldots - V_n - D$ is in $G[X \cup Y \cup Z]^a$.

Either case implies that there is a path in $G[X \cup Y \cup Z]^a$ connecting $A \in X$ and $B \in Y$ given Z. □

Theorem 4. *Given a probability distribution p satisfying the intersection property, p satisfies the global Markov property with respect to an ADMG if and only if it satisfies the ordered local Markov property with respect to the ADMG and a consistent ordering of its nodes.*

Proof. We start by proving the only if part. It suffices to note that every node that is adjacent to B in $G[S]^a$ is in $Mb_{G[S]}(B)$, hence B is separated from $S \setminus (B \cup Mb_{G[S]}(B))$ given $Mb_{G[S]}(B)$ in $G[S]^a$. Thus, $B \perp_p S \setminus (B \cup Mb_{G[S]}(B)) | Mb_{G[S]}(B)$ by the global Markov property.

To prove the if part, let A be the node in $X \cup Y \cup Z$ that occurs the latest in $\prec$, and let $S = X \cup Y \cup Z$. Note that for all $B \in S$, the set of nodes that are adjacent to B in $G[S]^a$ is precisely $Mb_{G[S]}(B)$. Then, the ordered local Markov property implies the global Markov property (Lauritzen, 1996, Theorem 3.7). □

Theorem 5. *Given a probability distribution p satisfying the intersection property, p satisfies the global Markov property with respect to an ADMG if and only if it satisfies the ordered pairwise Markov property with respect to the ADMG and a consistent ordering of its nodes.*

Proof. We start by proving the only if part. It suffices to note that if B and C are not adjacent in $G[S]^a$, then they are separated in $G[S]^a$ given $V(G[S]) \setminus (B \cup C)$. Thus, $B \perp_p C | V(G[S]) \setminus (B \cup C)$ by the global Markov property.

To prove the if part, let A be the node in $X \cup Y \cup Z$ that occurs the latest in $\prec$, and let $S = X \cup Y \cup Z$. Then, the ordered pairwise Markov property implies the global Markov property (Lauritzen, 1996, Theorem 3.7). □

Theorem 6. *C1, C2 and C3* hold if and only if the following two properties hold:*

- *C1*:* $D \perp_p Nd_G(D) \setminus Pa_G(D) | Pa_G(D)$ *for all* $D \subseteq C$.
- *C2*:* $p(C | Pa_G(C))$ *satisfies the global Markov property with respect to* G_C.

Proof. First, C1* implies C3* by decomposition. Second, C1* implies C1 by taking $D = C$ and applying weak union. Third, C1 and the fact that $Nd_G(D) = Nd_G(C)$ imply $D \perp_p Nd_G(D) \setminus Cc_G(Pa_G(C)) | Cc_G(Pa_G(C))$ by symmetry and decomposition, which together with C3* imply C1* by contraction. Finally, C2 and C2* are equivalent because $p(C | Pa_G(C)) = p(C | Cc_G(Pa_G(C)))$ by C1* and decomposition. □

Theorem 7. *A probability distribution p satisfying the intersection property satisfies the global Markov property with respect to an AMP CG G if and only if the following two properties hold for all* $C \in Cc(G)$:

- *L1:* $A \perp_p C \setminus (A \cup Ne_G(A)) | Nd_G(C) \cup Ne_G(A)$ *for all* $A \in C$.
- *L2*:* $A \perp_p Nd_G(C) \setminus Pa_G(A \cup S) | S \cup Pa_G(A \cup S)$ *for all* $A \in C$ *and* $S \subseteq C \setminus A$.

Proof. To see the only if part, note that C1* with $D = C$ implies that $p(C|Nd_G(C)) = p(C|Pa_G(C))$. This implies that $p(C|Nd_G(C))$ satisfies the global Markov property with respect to G_C by C2*. This implies L1 (Lauritzen, 1996, Theorem 3.7). Moreover, let $D = A \cup S$ and note that $Nd_G(D) = Nd_G(C)$. Then, L2* follows from C1* by symmetry and weak union.

To see the if part, note that L2* with $S = Ne_G(A)$ implies that $A \perp_p Nd_G(C) \setminus Pa_G(C)|Ne_G(A) \cup Pa_G(C)$ by weak union. This together with L1 imply $A \perp_p C \setminus (A \cup Ne_G(A))|Ne_G(A) \cup Pa_G(C)$ by contraction and decomposition. This implies C2* (Lauritzen, 1996, Theorem 3.7). Moreover, let $D = \{D_1, \ldots, D_n\}$. Then

(1) $D_1 \perp_p Nd_G(C) \setminus Pa_G(D)|(D \setminus D_1) \cup Pa_G(D)$ by L2* with $A = D_1$ and $S = D \setminus D_1$.
(2) $D_2 \perp_p Nd_G(C) \setminus Pa_G(D)|(D \setminus D_2) \cup Pa_G(D)$ by L2* with $A = D_2$ and $S = D \setminus D_2$.
(3) $D_1 \cup D_2 \perp_p Nd_G(C) \setminus Pa_G(D)|(D \setminus (D_1 \cup D_2)) \cup Pa_G(D)$ by symmetry and intersection on (1) and (2).
(4) $D_3 \perp_p Nd_G(C) \setminus Pa_G(D)|(D \setminus D_3) \cup Pa_G(D)$ by L2* with $A = D_3$ and $S = D \setminus D_3$.
(5) $D_1 \cup D_2 \cup D_3 \perp_p Nd_G(C) \setminus Pa_G(D)|(D \setminus (D_1 \cup D_2 \cup D_3)) \cup Pa_G(D)$ by symmetry and intersection on (3) and (4).

Continuing with this for $D_4, \ldots, D_n$ leads to C1*. □

Theorem 8. *A probability distribution p satisfying the intersection property satisfies the global Markov property with respect to an AMP CG G if and only if the following two properties hold for all $C \in Cc(G)$:*

- *P1: $A \perp_p B|Nd_G(C) \cup C \setminus (A \cup B)$ for all $A \in C$ and $B \in C \setminus (A \cup Ne_G(A))$.*
- *P2*: $A \perp_p B|S \cup Nd_G(C) \setminus B$ for all $A \in C$, $S \subseteq C \setminus A$ and $B \in Nd_G(C) \setminus Pa_G(A \cup S)$.*

Proof. To see the only if part, note that L1 and L2* imply P1 and P2* by weak union. To see the if part, let $Nd_G(C) \setminus Pa_G(A \cup S) = \{B_1, \ldots, B_n\}$. Then

(1) $A \perp_p B_1|S \cup Nd_G(C) \setminus B_1$ by P2* with $B = B_1$.
(2) $A \perp_p B_2|S \cup Nd_G(C) \setminus B_2$ by P2* with $B = B_2$.
(3) $A \perp_p B_1 \cup B_2|S \cup Nd_G(C) \setminus (B_1 \cup B_2)$ by intersection on (1) and (2).
(4) $A \perp_p B_3|S \cup Nd_G(C) \setminus B_3$ by P2* with $B = B_3$.
(5) $A \perp_p B_1 \cup B_2 \cup B_3|S \cup Nd_G(C) \setminus (B_1 \cup B_2 \cup B_3)$ by intersection on (3) and (4).

Continuing with this for $B_4, \ldots, B_n$ leads to L2*. Finally, let $C \setminus (A \cup Ne_G(A)) = \{B_1, \ldots, B_n\}$. Then

(6) $A \perp_p B_1|Nd_G(C) \cup C \setminus (A \cup B_1)$ by P1 with $B = B_1$.
(7) $A \perp_p B_2|Nd_G(C) \cup C \setminus (A \cup B_2)$ by P1 with $B = B_2$.
(8) $A \perp_p B_1 \cup B_2|Nd_G(C) \cup C \setminus (A \cup B_1 \cup B_2)$ by intersection on (6) and (7).
(9) $A \perp_p B_3|Nd_G(C) \cup C \setminus (A \cup B_3)$ by P1 with $B = B_3$.
(10) $A \perp_p B_1 \cup B_2 \cup B_3|Nd_G(C) \setminus (B_1 \cup B_2 \cup B_3)$ by intersection on (8) and (9).

Continuing with this for $B_4, \ldots, B_n$ leads to L1. □

Theorem 9. *Let X, Y and Z denote three disjoint subsets of V. Then, $X \perp_G Y|Z$ if and only if $X \perp_{G'} Y|Z$.*

Proof. It suffices to show that every path in G connecting α and β given Z can be transformed into a path in G' connecting α and β given Z and vice versa, with $\alpha, \beta \in V$ and $Z \subseteq V \setminus (\alpha \cup \beta)$.

Let ρ denote a path in G connecting α and β given Z. We can easily transform ρ into a path ρ' in G' between α and β: Simply, replace every maximal subpath of ρ of the form $V_1 - V_2 - \ldots - V_{n-1} - V_n$ ($n \geq 2$) with $V_1 \leftarrow \epsilon_{V_1} - \epsilon_{V_2} - \ldots - \epsilon_{V_{n-1}} - \epsilon_{V_n} \rightarrow V_n$. We now show that ρ' is connecting given Z.

First, if $B \in V$ is a collider on ρ', then ρ' must have one of the following subpaths:

$A \longrightarrow B \longleftarrow C$	$A \longrightarrow B \longleftarrow \epsilon_B — \epsilon_C$

with $A, C \in V$. Therefore, ρ must have one of the following subpaths:

$A \longrightarrow B \longleftarrow C$	$A \longrightarrow B — C$

In either case, B is a collider on ρ and, thus, $B \in An_G(Z)$ for ρ to be connecting given Z. Then, $B \in An_{G'}(Z)$ by construction of G' and, thus, $B \in An_{G'}(Dt(Z))$.

Second, if $B \in V$ is a non-collider on ρ', then ρ' must have one of the following subpaths:

$A \longrightarrow B \longrightarrow C$	$A \longleftarrow B \longrightarrow C$	$A \longleftarrow B \longleftarrow \epsilon_B — \epsilon_C$

with $A, C \in V$. Therefore, ρ must have one of the following subpaths:

$A \longrightarrow B \longrightarrow C$	$A \longleftarrow B \longrightarrow C$	$A \longleftarrow B — C$

In either case, B is a non-collider on ρ and, thus, $B \notin Z$ for ρ to be connecting given Z. Since Z contains no error node, Z cannot determine any node in V that is not already in Z. Then, $B \notin Dt(Z)$.

Third, if ϵ_B is a non-collider on ρ' (note that ϵ_B cannot be a collider on ρ'), then ρ' must have one of the following subpaths:

$A \longrightarrow B \longleftarrow \epsilon_B — \epsilon_C$	$\epsilon_A — \epsilon_B \longrightarrow B \in \{\alpha, \beta\}$	$A \longleftarrow B \longleftarrow \epsilon_B — \epsilon_C$	$\epsilon_A — \epsilon_B — \epsilon_C$

with $A, C \in V$. Recall that $\epsilon_B \notin Z$ because $Z \subseteq V \setminus (\alpha \cup \beta)$. In the first case, if $A = \alpha$ or $A = \beta$ then $A \notin Z$, else $A \notin Z$ for ρ to be connecting given Z. Then, $\epsilon_B \notin Dt(Z)$. In the second case, $B \notin Z$ and, thus, $\epsilon_B \notin Dt(Z)$. In the third case, $B \notin Z$ for ρ to be connecting given Z. Then, $\epsilon_B \notin Dt(Z)$. The last case implies that ρ has the following subpath:

$A — B — C$

Thus, B is a non-collider on ρ, which implies that $B \notin Z$ or $Pa_G(B) \setminus Z \neq \emptyset$ for ρ to be connecting given Z. In either case, $\epsilon_B \notin Dt(Z)$ (recall that $Pa_{G'}(B) = Pa_G(B) \cup \epsilon_B$ by construction of G').

Finally, let ρ' denote a path in G' connecting α and β given Z. We can easily transform ρ' into a path ρ between α and β: Simply, replace every maximal subpath of ρ' of the form $V_1 \leftarrow \epsilon_{V_1} - \epsilon_{V_2} - \ldots - \epsilon_{V_{n-1}} - \epsilon_{V_n} \rightarrow V_n$ ($n \geq 2$) with $V_1 - V_2 - \ldots - V_{n-1} - V_n$. We now show that ρ is connecting given Z.

First, note that all the nodes in ρ are in V. Moreover, if B is a collider on ρ, then ρ must have one of the following subpaths:

$A \longrightarrow B \longleftarrow C$	$A \longrightarrow B — C$

with $A, C \in V$. Therefore, ρ' must have one of the following subpaths:

$A \longrightarrow B \longleftarrow C$	$A \longrightarrow B \longleftarrow \epsilon_B — \epsilon_C$

In either case, B is a collider on ρ' and, thus, $B \in An_{G'}(Dt(Z))$ for ρ' to connecting given Z. Since Z contains no error node, Z cannot determine any node in V that is not already in Z. Then, $B \in Dt(Z)$ if and only if $B \in Z$. Since no error node is a descendant of B, then any node $D \in Dt(Z)$ that is a descendant of B must be in V which, as seen, implies that $D \in Z$. Then, $B \in An_{G'}(Dt(Z))$ if and only if $B \in An_{G'}(Z)$. Moreover, $B \in An_{G'}(Z)$ if and only if $B \in An_G(Z)$ by construction of G'. These results together imply that $B \in An_G(Z)$.

Second, if B is a non-collider on ρ, then ρ must have one of the following subpaths:

$A \longrightarrow B \longrightarrow C$	$A \longleftarrow B \longrightarrow C$	$A \longleftarrow B — C$	$A — B — C$

with $A, C \in V$. Therefore, ρ' must have one of the following subpaths:

$A \longrightarrow B \longrightarrow C$	$A \longleftarrow B \longrightarrow C$	$A \longleftarrow B \longleftarrow \epsilon_B — \epsilon_C$	$\epsilon_A — \epsilon_B — \epsilon_C$

In the first three cases, B is a non-collider on ρ' and, thus, $B \notin Dt(Z)$ for ρ' to be connecting given Z. Since Z contains no error node, Z cannot determine any node in V that is not already in Z. Then, $B \notin Z$. In the last case, ϵ_B is a non-collider on ρ' and, thus, $\epsilon_B \notin Dt(Z)$ for ρ' to be connecting given Z. Then, $B \notin Z$ or $Pa_{G'}(B) \setminus (\epsilon_B \cup Z) \neq \emptyset$. Then, $B \notin Z$ or $Pa_G(B) \setminus Z \neq \emptyset$ (recall that $Pa_{G'}(B) = Pa_G(B) \cup \epsilon_B$ by construction of G'). □

Theorem 10. *Every probability distribution $p(V)$ specified by Equations (1) and (2) is Gaussian.*

Proof. Modify the equation $A = \beta_A \cdot Pa_G(A) + \epsilon_A$ by replacing each $B \in V$ in the right-hand side of the equation with the right-hand side of the equation of B, i.e. $\beta_B \cdot Pa_G(B) + \epsilon_B$. Since G is directed acyclic, repeating this process results in a set of equations for the elements of V whose right-hand sides are linear combinations of the elements of ϵ. In other words, $V = \delta\epsilon$ with $\epsilon \sim \mathcal{N}(0, \Lambda)$. Then, $V \sim \mathcal{N}(0, \delta\Lambda\delta^T)$. □

Theorem 11. *Every probability distribution $p(V)$ specified by Equations (1) and (2) satisfies the global Markov property with respect to G.*

Proof. Equation (1) implies for any $A \in V$ that

$$A \perp_{p(V \cup \epsilon)} (V \cup \epsilon) \setminus (A \cup Pa_{G'}(A)) | Pa_{G'}(A)$$

and thus

$$A \perp_{p(V \cup \epsilon)} Nd_{G'}(A) \setminus Pa_{G'}(A) | Pa_{G'}(A) \tag{3}$$

by decomposition.

Moreover, Equation (1) implies for any $\epsilon_C \in Cc(G'_\epsilon)$ that

$$p(\epsilon_C \cup Nd_{G'}(\epsilon_C)) = p(\epsilon_C) p(Nd_{G'}(\epsilon_C)) \tag{4}$$

and thus

$$\epsilon_C \perp_{p(V \cup \epsilon)} Nd_{G'}(\epsilon_C) | \emptyset$$

and thus

$$\epsilon_A \perp_{p(V \cup \epsilon)} Nd_{G'}(\epsilon_C) | \epsilon_Z \tag{5}$$

where $\epsilon_A \in \epsilon_C$ and $\epsilon_Z \subseteq \epsilon_C \setminus \epsilon_A$, by decomposition and weak union.

Finally, Equation (2) implies for any $\epsilon_A \in \epsilon_C$ and $\epsilon_C \in Cc(G'_\epsilon)$ that

$$\epsilon_A \perp_{p(V \cup \epsilon)} \epsilon_C \setminus (\epsilon_A \cup Ne_{G'}(\epsilon_A)) | Ne_{G'}(\epsilon_A)$$

by Lauritzen (1996, Theorem 3.7 and Proposition 5.2), and thus

$$p(\epsilon_C) = h(\epsilon_A \cup Ne_{G'}(\epsilon_A)) k(\epsilon_C \setminus \epsilon_A)$$

by Lauritzen (1996, Equation 3.6). This together with Equation (4) imply that

$$p(\epsilon_C \cup Nd_{G'}(\epsilon_C)) = h(\epsilon_A \cup Ne_{G'}(\epsilon_A)) k(\epsilon_C \setminus \epsilon_A) p(Nd_{G'}(\epsilon_C))$$

and thus

$$\epsilon_A \perp_{p(V \cup \epsilon)} \epsilon_C \setminus (\epsilon_A \cup Ne_{G'}(\epsilon_A)) | Nd_{G'}(\epsilon_C) \cup Ne_{G'}(\epsilon_A) \tag{6}$$

by Lauritzen (1996, Equation 3.6).

Consequently, Equations (3), (5) and (6) imply that $p(V \cup \epsilon)$ satisfies the global Markov property with respect to G' by Theorem 7 because (i) G' is actually an AMP CG over $V \cup \epsilon$, (ii) A is the only node in the connectivity component of G' that contains A, and (iii) ϵ_C has no parents in G'. Then, $p(V)$ satisfies the global Markov property with respect to G, because G and G' represent the same separations over V by Theorem 9. □

References

Lauritzen, S. L. *Graphical Models.* Oxford University Press, 1996.