Supplementary Materials For: Acyclic Linear SEMs Obey the Nested Markov Property

Ilya Shpitser
Department of Computer Science
Johns Hopkins University
ilyas@cs.jhu.edu

Robin J. Evans
Department of Statistics
University of Oxford
evans@stats.ox.ac.uk

Thomas S. Richardson
Department of Statistics
University of Washington
thomasr@u.washington.edu

Appendix

m-separation and the Global Markov Property

Let $\mathcal{G}$ be an ADMG with vertex set V, and let $a, b \in V$ and $C \subseteq V \setminus \{a, b\}$, with C possibly empty. We say a path π from a to b is *open* if no noncollider on π is in C, and every collider on π is in $\text{an}_{\mathcal{G}}(C)$. A path which is not open is said to be *blocked* by C.

Now let A, B, C be disjoint subsets of V (again, C may be empty). We say that A is *m-separated* from B by C in $\mathcal{G}$, if every path from any $a \in A$ to any $b \in B$ is blocked by C.

We say a density p obeys the *global Markov property* with respect to $\mathcal{G}$ if whenever A and B are m-separated by C, the conditional independence $X_A \perp\!\!\!\perp X_B \mid X_C$ holds in p.

Fixing and Conditional Independence

Let $q_V(x_V \mid x_W)$ be a kernel. The usual notion of conditional independence in distributions is naturally extended to kernels by saying that $X_A \perp\!\!\!\perp X_B \mid X_C$ if $A \subseteq V$ and $q_V(x_A \mid x_B, x_C, x_{W \setminus (B \cup C)})$ is a function only of x_A and x_C (or otherwise with the roles of A and B interchanged). See [1] for more details.

Proof of Proposition 32. First note that for any $A, B \subseteq S$ we have

$$p^*_S(x_A \mid x_B, x_{V \setminus S}) = q_S(x_A \mid x_B, x_{V \setminus S}).$$

Let $W = V \setminus (S \cup B \cup C)$, so that

$$\begin{aligned}
&p^*_S(x_A \mid x_B, x_C)\\
&= \sum_{x_W} p^*_S(x_A, x_W \mid x_B, x_C)\\
&= \sum_{x_W} p^*_S(x_A \mid x_B, x_C, x_W) \cdot p^*_S(x_W \mid x_B, x_C)\\
&= \sum_{x_W} q_S(x_A \mid x_B, x_C, x_W) \cdot p^*_S(x_W \mid x_B, x_C).
\end{aligned}$$

Then by definition of conditional independence in q_S, the first factor depends only on x_A and x_C, so

$$\begin{aligned}
p^*_S(x_A \mid x_B, x_C) &= f(x_A, x_C) \sum_{x_W} p^*_S(x_W \mid x_B, x_C)\\
&= f(x_A, x_C).
\end{aligned}$$

Hence $X_A \perp\!\!\!\perp X_B \mid X_C$ in p^*_S. □

Technical Proofs

Proof of Proposition 17. If $a \in \text{dis}_{\mathcal{G}^\dagger}(b)$, then fix a bidirected path $a \leftrightarrow w_1 \leftrightarrow \ldots \leftrightarrow w_k \leftrightarrow b$ in $\mathcal{G}^\dagger$. Each bidirected edge on this path from c to d is due to $\langle \{c, d\} \rangle_{\mathcal{G}}$ being bidirected-connected in $\mathcal{G}$. But this implies the existence of a bidirected path from c to d in $\mathcal{G}$. Thus, there is a bidirected path from a to b in $\mathcal{G}$.

Suppose $a, b \in S$ such that there is a bidirected edge $a \leftrightarrow b$ in $\mathcal{G}^\dagger$ (and hence also in $\phi_{V \setminus S}(\mathcal{G}^\dagger)$). By the construction of $\mathcal{G}^\dagger$, a and b are bidirected-connected in the closure of $\{a, b\}$ in $\mathcal{G}$. If $a, b \in S$, then by definition of closure and fixing (in $\mathcal{G}$), every vertex in the closure of $\{a, b\}$ is in S. Hence a and b are bidirected-connected in $\mathcal{G}$ by a path on which every vertex is in S. Hence a and b are bidirected-connected (in $\phi_{V \setminus S}(\mathcal{G})$). Consequently, the districts of $\phi_{V \setminus S}(\mathcal{G}^\dagger)$ form a sub-partition of the districts in $\phi_{V \setminus S}(\mathcal{G})$. □

Proof of Lemma 18. If there is such a path in $\mathcal{G}$ then the same path exists in $\mathcal{G}^\dagger$ by Proposition 14. If there is a path in $\mathcal{G}^\dagger$ then consider an edge $c \to d$; since this exists in $\mathcal{G}^\dagger$ then $c \in \mathrm{pa}_\mathcal{G}(\langle d \rangle_\mathcal{G})$, so there is a directed path in $\mathcal{G}$ from c to d whose internal vertices (if any) are all in $\langle d \rangle_\mathcal{G} \setminus \{b\}$. Such vertices are not fixable in $\mathcal{G}$ by definition and therefore do not include v. Hence we have constructed a path that does not intersect v. □

Proof of Theorem 19. If v is fixable in $\mathcal{G}$, then it is also fixable in $\mathcal{G}^\dagger$ by application of Propositions 15 and 17.

By Proposition 17 the districts of $\phi_v(\mathcal{G}^\dagger)$ forms a subpartition of the districts in $\phi_v(\mathcal{G})$. If a is an ancestor of b in $\phi_v(\mathcal{G})$, then this is because there is a directed path in $\mathcal{G}$ from a to b that does not intersect v; by Lemma 18 this happens if and only if there is such a path in $\mathcal{G}^\dagger$, and hence a is an ancestor of b in $\phi_v(\mathcal{G}^\dagger)$. It follows that any vertex fixable in $\phi_v(\mathcal{G})$ is also fixable in $\phi_v(\mathcal{G}^\dagger)$, so a simple induction gives the first result.

Now let $S \in \mathcal{R}(\mathcal{G})$. We have $a \to b$ in $(\phi_{V\setminus S}(\mathcal{G}))^\dagger$ if and only if $b \in S$ and $a \in \mathrm{pa}_{\phi_{V\setminus S}(\mathcal{G})}(\langle b \rangle_{\phi_{V\setminus S}(\mathcal{G})})$. Since S is reachable, $\langle b \rangle_{\phi_{V\setminus S}(\mathcal{G})} = \langle b \rangle_\mathcal{G}$ by Proposition 4; and since $\langle b \rangle_{\phi_{V\setminus S}(\mathcal{G})} \subseteq S$, then $\mathrm{pa}_{\phi_{V\setminus S}(\mathcal{G})}(\langle b \rangle_\mathcal{G}) = \mathrm{pa}_\mathcal{G}(\langle b \rangle_\mathcal{G})$. Hence $a \in \mathrm{pa}_\mathcal{G}(\langle b \rangle_\mathcal{G})$, which happens if and only if $a \to b$ in $\mathcal{G}^\dagger$. But since $S \ni b$ this happens if and only if $a \to b$ in $\phi_{V\setminus S}(\mathcal{G}^\dagger)$. The directed edges are therefore the same.

$a \leftrightarrow b$ in $(\phi_{V\setminus S}(\mathcal{G}))^\dagger$ if and only if $a, b \in S$ and $\langle a, b \rangle_{\phi_{V\setminus S}(\mathcal{G})}$ is bidirected-connected. By Proposition 4, $\langle a, b \rangle_{\phi_{V\setminus S}(\mathcal{G})} = \langle a, b \rangle_\mathcal{G}$, so this happens if and only if $a \leftrightarrow b$ in $\mathcal{G}^\dagger$, which occurs if and only if $a \leftrightarrow b$ in $\phi_{V\setminus S}(\mathcal{G}^\dagger)$, since $a, b \in S$. □

Proof of Proposition 21. Suppose for contradiction that $\mathcal{G}^\dagger$ is not arid, so there exists v and $t \in \langle v \rangle_{\mathcal{G}^\dagger} \setminus \{v\}$ such that $t \leftrightarrow v$ in $\mathcal{G}^\dagger$, by Proposition 9.

Now $t \in \langle v \rangle_{\mathcal{G}^\dagger}$ implies $t \in \langle v \rangle_\mathcal{G}$ by Corollary 20, and since $t \neq v$ we have $t \in \mathrm{pa}_\mathcal{G}(\langle v \rangle_\mathcal{G})$ by Lemma 5.

But by construction of $\mathcal{G}^\dagger$ this implies that graph should contain $t \to v$; this is a contradiction since $t \leftrightarrow v$ was assumed to exist in $\mathcal{G}^\dagger$. This establishes $\mathcal{G}^\dagger$ is arid.

Now suppose a and b are densely connected in $\mathcal{G}^\dagger$. If this is because $a \in \mathrm{pa}_{\mathcal{G}^\dagger}(\langle b \rangle_{\mathcal{G}^\dagger})$, then $\mathcal{G}^\dagger$ being arid implies $\langle b \rangle_{\mathcal{G}^\dagger} = \{b\}$, and thus $a \in \mathrm{pa}_{\mathcal{G}^\dagger}(b)$; hence a and b are adjacent.

Alternatively, suppose $\langle \{a, b\} \rangle_{\mathcal{G}^\dagger}$ is a bidirected-connected set. Note that $\langle \{a, b\} \rangle_{\mathcal{G}^\dagger} \subseteq \langle \{a, b\} \rangle_\mathcal{G}$ by Corollary 20, so a and b are also bidirected-connected in $\langle \{a, b\} \rangle_\mathcal{G}$ in $\mathcal{G}^\dagger$. By Proposition 17, the districts in $\mathcal{G}$ form a superpartition of those in $\mathcal{G}^\dagger$, and therefore a and b are also bidirected-connected in $\langle \{a, b\} \rangle_\mathcal{G}$ in $\mathcal{G}$. Hence they satisfy the condition to add an edge $a \leftrightarrow b$ in the definition of $\mathcal{G}^\dagger$, and hence are adjacent in $\mathcal{G}^\dagger$. □

Proof of Lemma 22. The only possibly fixable vertices in $\langle \{v, w\} \rangle_\mathcal{G}$ are v and w by definition. But $w \in \mathrm{pa}_\mathcal{G}(\langle v \rangle_\mathcal{G})$ implies that w is an ancestor of v, and since $\langle \{v, w\} \rangle_\mathcal{G}$ is bidirected-connected w is also in the same district. Hence w is not fixable in $\langle \{v, w\} \rangle_\mathcal{G}$, giving the result. □

Proof of Lemma 23. Suppose that such a path exists in $\mathcal{G}$. Each adjacent pair $\{a, b\}$ on the path satisfies the criterion for insertion of an edge in $\mathcal{G}^\dagger$ of the same type as (one of) the edge(s) between a and b in $\mathcal{G}$. The only potential concern is that a bidirected edge in $\mathcal{G}$ might instead be replaced by a directed edge in $\mathcal{G}^\dagger$.

Let $\pi^\dagger$ be the path in $\mathcal{G}^\dagger$ with the same vertices as π (this is unique since $\mathcal{G}^\dagger$ is simple). Then $\pi^\dagger$ is such that all adjacent nodes a, b satisfy the condition that $\langle \{a, b\} \rangle_\mathcal{G}$ is bidirected-connected in $\mathcal{G}$, except possibly for the end pairs which either satisfy this or $v \in \mathrm{pa}_\mathcal{G}(\langle a \rangle_\mathcal{G})$.

If all the internal nodes on $\pi^\dagger$ are colliders in $\mathcal{G}^\dagger$ then we are done; otherwise we claim we can find a strict subpath of $\pi^\dagger$ (still from v to w), such that any adjacent nodes still satisfy the condition above.

Suppose a is an internal non-collider on $\pi^\dagger$ because $a \leftrightarrow b$ has been replaced by $a \to b$. The replacement implies that $a \in \mathrm{pa}_\mathcal{G}(\langle b \rangle_\mathcal{G})$ which with $a \leftrightarrow b$ in $\mathcal{G}$ implies that $a \in \langle b \rangle_\mathcal{G}$ by Lemma 22. Let c be the other neighbour of a on the path.

First suppose $c \leftrightarrow a$ on π; then $a \in \langle b \rangle_\mathcal{G}$ implies that $a \in \langle \{b, c\} \rangle_\mathcal{G}$, so $\langle \{b, c\} \rangle_\mathcal{G}$ is bidirected-connected. Hence we can remove a from $\pi^\dagger$ and repeat the argument. Alternatively, if $c \to a$ on π (i.e. c is the end vertex on the path) then $c \in \mathrm{pa}_\mathcal{G}(\langle b \rangle_\mathcal{G})$ since $a \in \langle b \rangle_\mathcal{G}$, so start the path with $c \to b$. In either case, we have reduced the number of vertices on the path being considered, and this process will eventually terminate. □

Proof of Proposition 26. Let $S \in \mathcal{I}(\mathcal{G})$. Using Theorem 19 it is sufficient to consider the case in which S is the set of all random vertices in $\mathcal{G}$. Let $H \subseteq S$ be the set of childless vertices in $\mathcal{G}$. Since S is bidirected-connected in $\mathcal{G}$, every pair of vertices in H is connected by a path of bidirected edges within S in $\mathcal{G}$, and hence is also connected by a collider path in $\mathcal{G}^\dagger$ by Lemma 23. By Proposition 15, vertices in H are also childless in $\mathcal{G}^\dagger$, so the collider paths consist entirely of bidirected edges; hence H is bidirected-connected by paths in S in $\mathcal{G}^\dagger$.

Let the district of $\mathcal{G}^\dagger$ containing H be $S^\dagger \subseteq S$. It then follows that $\langle H \rangle_{\mathcal{G}^\dagger} = S^\dagger$; see below for a proof. Further, since S is reachable in $\mathcal{G}$, S is reachable in $\mathcal{G}^\dagger$ by Theorem 19. Since $S^\dagger$ is a district in a reachable subgraph $\phi_{V\setminus S}(\mathcal{G}^\dagger)$ of $\mathcal{G}^\dagger$, $S^\dagger \in \mathcal{I}(\mathcal{G}^\dagger)$.

Conversely, let $S^\dagger \in \mathcal{I}(\mathcal{G}^\dagger)$ and $H^\dagger \subseteq S^\dagger$ the set of childless vertices in $S^\dagger$ in $\phi_{V\setminus S^\dagger}(\mathcal{G})$. Since only elements of $H^\dagger$ are fixable in $\phi_{V \setminus S^\dagger}(\mathcal{G})$, $S^\dagger = \langle H^\dagger \rangle_{\mathcal{G}^\dagger}$. Let $S = \langle H^\dagger \rangle_{\mathcal{G}}$, which is a superset of $S^\dagger = \langle H^\dagger \rangle_{\mathcal{G}^\dagger}$ by Corollary 20.

For every pair a, b in $S^\dagger \subseteq S$ with $a \in \mathrm{sib}_{\mathcal{G}^\dagger}(b)$, $\langle \{a, b\} \rangle_{\mathcal{G}}$ must be in $S = \langle H^\dagger \rangle_{\mathcal{G}}$. Consequently H is bidirected connected in $\phi_{V \setminus S}(\mathcal{G})$, and thus so is S. Then S is intrinsic in $\mathcal{G}$, and $H^\dagger$ is the set of childless vertices in S in $\phi_{V \setminus S}(\mathcal{G})$. This establishes the correspondence. □

Proof that $\langle H \rangle_{\mathcal{G}^\dagger} = S^\dagger$ used in Proof of Proposition 26. Since H is bidirected-connected in $\mathcal{G}^\dagger$, clearly $\langle H \rangle_{\mathcal{G}^\dagger} \subseteq S^\dagger$. Let v be any vertex in $S^\dagger \setminus H$. Since $v \notin H$, $\mathrm{ch}_{\mathcal{G}}(v) \neq \emptyset$. Since, by hypothesis $\mathcal{G}$ consists of a single district it follows that there is a collider path: $v \to c \cdots \leftrightarrow h$ with $h \in H$ in $\mathcal{G}$. Hence by Lemma 23 and the fact that $\mathrm{ch}_{\mathcal{G}^\dagger}(H) = \emptyset$, there is a path $v \to c^\dagger \cdots \leftrightarrow h$ in $\mathcal{G}^\dagger$, so $c^\dagger \in S^\dagger$. Thus from every vertex in $S^\dagger \setminus H$ there is a directed path to a vertex $h \in H$ on which every vertex is in $S^\dagger \setminus H$. Hence $\langle H \rangle_{\mathcal{G}^\dagger} = S^\dagger$. □

References

[1] T. S. Richardson, R. J. Evans, J. M. Robins, and I. Shpitser. Nested Markov properties for acyclic directed mixed graphs. Working paper, `https://arxiv.org/abs/1701.06686v2`, 2017.