Supplementary Material to Segregated Graphs and Marginals of Chain Graph Models

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

1 Proofs

Lemma 3.1. *The Markov properties defined by superactive routes (walks) [10] in CGs, m-separation [8] in ADMGs, and d-separation [6] in DAGs are special cases of the Markov property defined by s-separation in SGs.*

Proof: An argument that d-separation in DAGs is a special case of the separation criterion based on superactive routes appears in [10]. An argument that d-separation in DAGs is a special case of m-separation in ADMGs trivially follows by definition. That separation based on superactive routes is a special case of s-separation follows from the fact that CGs are a special case of SGs with no $\leftrightarrow$ edges, which implies only directed edges can result in collider sections in CGs. That m-separation is a special case of s-separation follows by extension of the argument in [10]. $\square$

Lemma 4.1. *For V sensitive in a SG $\mathcal{G}$, let $\mathcal{G}^{\langle V \rangle}$ be the graph be obtained from $\mathcal{G}$ by replacing all $-$ edges adjacent to V by $\rightarrow$ edges pointing away from V. Then $\mathcal{G}^{\langle V \rangle}$ is an SG, and $\mathcal{P}(\mathcal{G}) = \mathcal{P}(\mathcal{G}^{\langle V \rangle})$.*

Proof: Since $\mathcal{G}^{\langle V \rangle}$ is constructed from an SG by replacing certain $-$ edges by $\rightarrow$, then if $\mathcal{G}$ does not contain $\circ \leftrightarrow \circ - \circ$, then neither does $\mathcal{G}^{\langle V \rangle}$. If $\mathcal{G}^{\langle V \rangle}$ contains a partially directed cycle not including V, so does $\mathcal{G}$, which is a contradiction. If $\mathcal{G}^{\langle V \rangle}$ contains a partially directed cycle including V, then it must be via a subpath $\circ \rightarrow V \rightarrow \circ$, with all other edges on the path present in $\mathcal{G}$. But either the outgoing edge from V that is on the cycle is also present in $\mathcal{G}$ or it is undirected. In both cases, there is still a partially directed cycle in $\mathcal{G}$, which is a contradiction. Thus $\mathcal{G}^{\langle V \rangle}$ is a SG. If V has no adjacent $-$ edges, $\mathcal{G} = \mathcal{G}^{\langle V \rangle}$.

Assume $(\mathbf{A} \not\perp\!\!\!\perp \mathbf{B} \mid \mathbf{C})_{\mathcal{G}^{\langle V \rangle}}$. Fix a walk α from $\mathbf{A}$ to $\mathbf{B}$ s-connected by $\mathbf{C}$ in $\mathcal{G}^{\langle V \rangle}$. We will construct an s-connected walk α^* from $\mathbf{A}$ to $\mathbf{B}$ given $\mathbf{C}$ in $\mathcal{G}$. By definition, every collider section in α intersects $\mathbf{C}$ and every non-collider section in α is free of $\mathbf{C}$. Any section of α where V does not occur either remains a section of α in $\mathcal{G}$, and retains its open status (if its neighboring edges do not change status in $\mathcal{G}$), or is subsumed by the argument for the following case. We now consider all sections β_i of α where V occurs. Note that β_i is a singleton section. If β_i is a collider section, $V \in \mathbf{C}$, and β_i exists in $\mathcal{G}$. Assume β_i is a non-collider section. Then $V \notin \mathbf{C}$. If β_i is in $\mathcal{G}$, we are done. Otherwise, consider a section β_j in α^* containing sections $\beta_{i-l}, \ldots, \beta_{i+k}$ in α. By definition of $\mathcal{G}^{\langle V \rangle}$, all sections except possibly β_{i-l} and β_{i+k} are either of the form $\leftarrow V \rightarrow$ or collider sections. Note that since α is open, all collider sections intersect $\mathbf{C}$.

If β_j is a collider section, we are done. Otherwise, we have two cases. If both neighboring edges along α^* into β_j are not into β_j, then $\beta_i \leftarrow V \rightarrow \beta_{i+k}$ shares the same endpoint behavior as β_j and is open, since β_i, $\leftarrow V \rightarrow$, and β_{i+k} are non-collider sections in α and thus do not intersect $\mathbf{C}$. If a single neighboring edge along α^* into β_j is into β_j (say into β_{i-l}), then either that edge is from V or not. If it is from V, the section $V \rightarrow \beta_{i+k}$ shares the same endpoint behavior as β_j and is open. If it is not from V, but another edge W, then since V is sensitive, $W \rightarrow V$ exists in $\mathcal{G}$, and the section $W \rightarrow V \rightarrow \beta_{i+k}$ shares the same endpoint behavior as β_j and is open.

Assume $(\mathbf{A} \not\perp\!\!\!\perp \mathbf{B} \mid \mathbf{C})_{\mathcal{G}}$. Fix a walk α from $\mathbf{A}$ to $\mathbf{B}$ s-connected by $\mathbf{C}$ in $\mathcal{G}$. We will construct an s-connected walk from $\mathbf{A}$ to $\mathbf{B}$ given $\mathbf{C}$ in $\mathcal{G}^{\langle V \rangle}$. By definition, every collider section in α intersects $\mathbf{C}$ and every non-collider section in α is free of $\mathbf{C}$. Any section of α where V does not occur remains a section of α in $\mathcal{G}^{\langle V \rangle}$, and retains its open status. We now consider all sections β_i of α where V occurs.

Assume β_i is a collider section with end points Z, W. If $V \in \mathbf{C}$, then since V is sensitive, $Z \to V \leftarrow W$ is present in $\mathcal{G}^{\langle V \rangle}$. Then we can construct a walk α' which shares all sections with α except β_i is replaced by $Z \to V \leftarrow W$, which is open since $V \in \mathbf{C}$. If $V \notin \mathbf{C}$, then there must be some section β_j in β_i in $\mathcal{G}^{\langle V \rangle}$ intersecting $\mathbf{C}$. This section either has V as both endpoints, or V and an endpoint Z of β_i with an arrowhead into β_i. We can then replace α with another walk α' which shares all sections with α except β_i is replaced either by $W \to V \beta_j V \leftarrow Z$, or $W \to V \beta_j \leftarrow Z$, which is open since β_j intersects $\mathbf{C}$. In either case, we then repeat the argument for other sections of α'.

Assume β_i is a non-collider section with end points Z, W, and does not intersect $\mathbf{C}$. This means there is at most one arrowhead into β_i, say from Z, or no arrowheads into β_i. In the former case, fix the section β_j (possibly of length 0 if $V = W$) in $\mathcal{G}^{\langle V \rangle}$ between the last occurrence of V and W in β_i. Replace α by a walk α' sharing all sections with α except β_i is replaced with $Z \to V \beta_j W$, which is open. If no arrowheads are into β_i, let β_j be the part of β_i from Z to first occurrence of V, and β_k be the part of β_i from the last occurrence of V to W. Replace α by a walk α' sharing all sections with α except β_i is replaced by $Z \beta_j V \beta_k W$. In all cases, the newly added sections to α' are open and share end edge behavior with sections they are replacing. We then repeat the argument for other sections of α'. Thus, $(\mathbf{A} \not\perp\!\!\!\perp \mathbf{B} \mid \mathbf{C})_{\mathcal{G}^{\langle V \rangle}}$. □

Lemma 4.2. *Let $\mathcal{G}$ be an SG, and $\mathcal{G}'$ a graph obtained from adding an edge $W \to V$ for two non-adjacent vertices W, V where $W \to \circ - \ldots - \circ - V$ exists in $\mathcal{G}$. Then $\mathcal{G}'$ is an SG.*

Proof: Since $\mathcal{G}$ is an SG, and we are adding only $\to$ edges to $\mathcal{G}'$, then there is no $\circ \leftrightarrow \circ - \circ$ structure in $\mathcal{G}'$. If there were a partially directed cycle involving $W \to V$ in $\mathcal{G}'$, then replacing $W \to V$ by $W \to \circ - \ldots - \circ - V$ in the cycle would still result in a partially directed cycle, which would also be present in $\mathcal{G}$. But this is a contradiction. □

Lemma 4.3. *For any V in an SG $\mathcal{G}$, let $\mathcal{G}^{\overline{V}}$ be obtained from $\mathcal{G}$ by adding $W \to Z$, whenever $W \to \circ - \ldots - \circ - Z \leftarrow V$ exists in $\mathcal{G}$. Then $\mathcal{G}^{\overline{V}}$ is an SG, and $\mathcal{P}(\mathcal{G})^V = \mathcal{P}(\mathcal{G}^{\overline{V}})^V$.*

Proof: $\mathcal{G}^{\overline{V}}$ is an SG by an inductive application of Lemma 4.2. If $\mathbf{A} \perp\!\!\!\perp \mathbf{B} \mid \mathbf{C}$ holds in $\mathcal{G}^{\overline{V}}$, then $\mathbf{A} \perp\!\!\!\perp \mathbf{B} \mid \mathbf{C}$ holds in $\mathcal{G}$, since $\mathcal{G}$ is an edge subgraph of $\mathcal{G}^{\overline{V}}$.

Assume $(\mathbf{A} \perp\!\!\!\perp \mathbf{B} \mid \mathbf{C})_{\mathcal{G}}$, where $V \notin \mathbf{A} \cup \mathbf{B} \cup \mathbf{C}$. Fix a walk α from $\mathbf{A}$ to $\mathbf{B}$ in $\mathcal{G}^{\overline{V}}$. If α exists in $\mathcal{G}$, then it retains the same edges in $\mathcal{G}^{\overline{V}}$, which implies if α is s-separated by $\mathbf{S}$ in $\mathcal{G}$, it is also in $\mathcal{G}^{\overline{V}}$. Assume α does not exist in $\mathcal{G}$ and is s-connected given $\mathbf{C}$. This means α contains a set of edges of the form $W \to Z$ which do not exist in $\mathcal{G}$. We will repeatedly replace edges $W \to Z$ in α by sections that exist in $\mathcal{G}$ while preserving the open status of the resulting walk. In this way, we will construct a new walk that is s-connected given $\mathbf{C}$ and exists in $\mathcal{G}$, deriving a contradiction.

Pick an edge $W \to Z$ in α that does not exist in $\mathcal{G}$, let β_j be the section of α starting at Z with $W \to Z$ pointing into it. By definition of $\mathcal{G}^{\overline{V}}$, there exists $\beta_i \equiv W \to \circ - \ldots - \circ - Z$ in $\mathcal{G}$. If β_j is a collider section, then replace $W \to \beta_j$ by $\beta_i \beta_j$. The new extended section is thus also a collider section intersecting $\mathbf{C}$, and exists in $\mathcal{G}$. If β_j is not a collider section, then either β_i intersects $\mathbf{C}$ or not. If it does, replace $W \to \beta_j$ by $\beta_i \leftarrow V \to \beta_j$. This results in three new sections which are all open given $\mathbf{C}$, exist in $\mathcal{G}$, and have same endpoint behavior as β_j. If it does not, replace $W \to \beta_j$ by $\beta_i \beta_j$. This results in a new extended section which is a non-collider section that does not intersect $\mathbf{C}$, exists in $\mathcal{G}$, and has same endpoint behavior as β_j.

Repeating the argument for every $W \to V$ that does not exist in $\mathcal{G}$ gives us the contradiction. □

Theorem 4.1. *If $\mathcal{G}$ is an SG with at least 2 vertices $\mathbf{V}$, and $V \in \mathbf{V}$, there exists an SG $\mathcal{G}^V$ with vertices $\mathbf{V} \setminus \{V\}$ such that $\mathcal{P}(\mathcal{G})^V = \mathcal{P}(\mathcal{G}^V)^V$.*

Proof: Construct $\mathcal{G}^{\underline{V}}$ as in Lemma 4.4. Construct $\mathcal{G}^V$ from $\mathcal{G}^{\underline{V}}$ as follows. Retain all vertices in $\mathbf{V} \setminus \{V\}$ and edges between them. For any two vertices W, Z: if $W \to V \to Z$, add $W \to Z$; if

$W \leftarrow V \rightarrow Z$, add $W \leftrightarrow Z$; if $W - V - Z$, add $W - Z$; if $W - V \rightarrow Z$, add $W \rightarrow Z$; and if $W \rightarrow V - Z$, add $W \rightarrow Z$.

Because $\mathcal{G}^{\underline{V}}$ is a SG with no $V \rightarrow \circ - \circ$, there is no $\circ \leftrightarrow \circ - \circ$ structure in $\mathcal{G}^V$. Assume there exists a partially directed cycle in $\mathcal{G}^V$ involving new edges. Then we can systematically replace them by the two edge paths in $\mathcal{G}$ to yield a partially directed cycle in $\mathcal{G}$, giving a contradiction.

Let $\mathcal{G}^{V^\dagger}$ be an edge supergraph of $\mathcal{G}^{\underline{V}}$ where we add all edges in $\mathcal{G}^V$ that do not exist in $\mathcal{G}$. We first show $\mathcal{P}(\mathcal{G}^{V^\dagger})^V = \mathcal{P}(\mathcal{G}^{\underline{V}})^V$. If $(\mathbf{A} \perp\!\!\!\perp \mathbf{B} \mid \mathbf{C})_{\mathcal{G}^{V^\dagger}}$, then $(\mathbf{A} \perp\!\!\!\perp \mathbf{B} \mid \mathbf{C})_{\mathcal{G}^{\underline{V}}}$ because $\mathcal{G}^{V^\dagger}$ is an edge supergraph of $\mathcal{G}^{\underline{V}}$. Assume $(\mathbf{A} \perp\!\!\!\perp \mathbf{B} \mid \mathbf{C})_{\mathcal{G}^{\underline{V}}}$, and fix a walk α from $\mathbf{A}$ to $\mathbf{B}$ that is s-connected given $\mathbf{C}$ in $\mathcal{G}^{V^\dagger}$. If α exists in $\mathcal{G}^{\underline{V}}$, we have a contradiction. Otherwise, since $V \notin \mathbf{A} \cup \mathbf{B} \cup \mathbf{C}$, it is easy to construct a walk α' that is s-connected given $\mathbf{C}$ and exists in $\mathcal{G}^{\underline{V}}$ by replacing edges in α that do not exist in $\mathcal{G}^{\underline{V}}$ by their corresponding two edges used in the construction of $\mathcal{G}^V$.

Finally, we show that $\mathcal{P}(\mathcal{G}^{V^\dagger})^V = \mathcal{P}(\mathcal{G}^V)^V$. Since $\mathcal{G}^{V^\dagger}$ is an edge supergraph of $\mathcal{G}^V$, if $\mathbf{A} \perp\!\!\!\perp \mathbf{B} \mid \mathbf{C}$ in $\mathcal{G}^{V^\dagger}$, then $\mathbf{A} \perp\!\!\!\perp \mathbf{B} \mid \mathbf{C}$ in $\mathcal{G}^V$. If $\mathbf{A} \perp\!\!\!\perp \mathbf{B} \mid \mathbf{C}$ in $\mathcal{G}^V$, and there is a s-connecting walk α from $\mathbf{A}$ to $\mathbf{B}$ given $\mathbf{C}$ in $\mathcal{G}^{V^\dagger}$, it must involve V. But we can construct a walk α' that does not contain V by replacing V containing segments by edges connecting nodes adjacent to V following above rules used to construct $\mathcal{G}^V$. It is easy to see α' is s-connecting given $\mathbf{C}$ if α is. This is a contradiction. □

Corollary 4.1. *Let $\mathcal{G}$ be an SG with vertices $\mathbf{V}$. Then for any $\mathbf{W} \subset \mathbf{V}$, there exists an SG $\mathcal{G}^*$ with vertices $\mathbf{V} \setminus \mathbf{W}$ such that $\mathcal{P}(\mathcal{G})^{\mathbf{W}} = \mathcal{P}(\mathcal{G}^*)$.*

Proof: Follows by an inductive application of Theorem 4.1 for any ordering of vertices in $\mathbf{W}$. □

Lemma 5.1. *If $p(\mathbf{V})$ factorizes with respect to $\mathcal{G}$ then $f_{\mathbf{S}}(\mathbf{S} \mid \mathrm{pa}^s_{\mathcal{G}}(\mathbf{S})) = p(\mathbf{S} \mid \mathrm{pa}^s_{\mathcal{G}}(\mathbf{S}))$ for every $\mathbf{S} \in \mathcal{B}^*(\mathcal{G})$, and $f_{\mathbf{S}}(\mathbf{S} \mid \mathrm{pa}^s_{\mathcal{G}}(\mathbf{S})) = \prod_{V \in \mathbf{S}} p(V \mid \mathrm{pre}_{\mathcal{G},\prec}(V) \cap \mathrm{ant}_{\mathcal{G}}(\mathbf{S}))$ for every $\mathbf{S} \in \mathcal{D}^a(\mathcal{G})$ and any topological ordering $\prec$ on $\mathcal{G}$.*

Proof: We will proceed by induction on anterial subgraphs. We will add either a singleton vertex that will be become a new singleton district or a part of an existing district, or a block of vertices $\mathbf{S}$ to construct $\mathcal{G}$ with vertices $\mathbf{V}$, such that $\mathbf{V} \setminus \mathbf{S} \in \mathcal{A}(\mathcal{G})$. For the base case, the conclusion clearly holds for $\mathcal{G}$ with a single vertex. Assume the inductive hypothesis holds for $\mathcal{G}^i$, and we added a block $\mathbf{S}$ to $\mathcal{G}^i$ to yield $\mathcal{G}$, where $\mathbf{S} \in \mathcal{B}^*(\mathcal{G})$. By the inductive hypothesis, $p(\mathbf{V}) = f_{\mathbf{S}}(\mathbf{S} \mid \mathrm{pa}^s_{\mathcal{G}}(\mathbf{S})) \cdot \prod_{\mathbf{S} \in \mathcal{D}(\mathcal{G}^i) \cup \mathcal{B}^*(\mathcal{G}^i)} p(\mathbf{S} \mid \mathrm{pa}^s_{\mathcal{G}^i}(\mathbf{S}))$. This implies our conclusion. Assume the inductive hypothesis holds for $\mathcal{G}^i$, and we added V to $\mathcal{G}^i$ to yield $\mathcal{G}$, where $V \in \mathbf{S} \in \mathcal{D}(\mathcal{G})$. Then the conclusion follows by a simple extension of the argument used to prove Lemma 1 in [11]. □

Theorem 5.1. *If $p(\mathbf{V})$ factorizes with respect to a SG $\mathcal{G}$, then $p(\mathbf{V}) \in \mathcal{P}^a(\mathcal{G})$.*

Proof: Implied by the fact that the UG factorization implies the UG global Markov property [5]. □

Lemma 5.2. *If there exists a walk α in $\mathcal{G}$ between $A \in \mathbf{A}, B \in \mathbf{B}$ with all non-collider sections not intersecting $\mathbf{C}$, and all collider sections in $\mathrm{ant}_{\mathcal{G}}(\mathbf{A} \cup \mathbf{B} \cup \mathbf{C})$, then there exist $A^* \in \mathbf{A}, B^* \in \mathbf{B}$ such that A^* and B^* are s-connected given $\mathbf{C}$ in $\mathcal{G}$.* [1]

Proof: Let D be the last vertex on α in $\mathrm{ant}_{\mathcal{G}}(\mathbf{A}) \setminus \mathrm{ant}_{\mathcal{G}}(\mathbf{C})$ if such a vertex exists, or $D \equiv A$ otherwise. Let E be the first vertex in $\mathrm{ant}_{\mathcal{G}}(\mathbf{B}) \setminus \mathrm{ant}_{\mathcal{G}}(\mathbf{C})$ which occurs between the last occurrence of D in α and B, if such a vertex exists, or $E \equiv B$ otherwise. If $D \neq A$, let A^* be any vertex such that $D \in \mathrm{ant}_{\mathcal{G}}(A^*)$, otherwise let $A^* \equiv A$. Similarly, if $E \neq B$, let B^* be any vertex such that $E \in \mathrm{ant}_{\mathcal{G}}(B^*)$, otherwise let $B^* \equiv B$.

Let α^* be the subwalk of α between the last occurrence of D and the first occurrence of E. Then: (a) every vertex in α^* is in $\mathrm{ant}_{\mathcal{G}}(\mathbf{C})$; (b) there is a partially directed path δ from D to A^*, and ϵ from E to B^*; (c) other than possibly D or E, no vertex in δ or ϵ is in $\mathrm{ant}_{\mathcal{G}}(\mathbf{C})$; and (d) no vertex in ϵ other than possibly E is an ancestor of A^*.

It follows from (a) and (c) that α^* and ϵ only intersect at E, and α^* and δ only intersect at D. Let β be a walk obtained by concatenating δ, α^*, and γ. By construction, every collider section in α^* is in $\mathrm{ant}_{\mathcal{G}}(\mathbf{C})$, every non-collider section in α^* does not intersect $\mathbf{C}$. Furthermore, every section in δ and ϵ is non-collider and does not intersect $\mathbf{C}$. Thus β is s-connecting given $\mathbf{C}$. □

[1] The proof follows the proof of lemma 1 in [8].

Theorem 5.2. $\mathcal{P}(\mathcal{G}) = \mathcal{P}^a(\mathcal{G})$. [2]

Proof: Fix disjoint $\mathbf{A}, \mathbf{B}, \mathbf{C}$, and consider the smallest anterial set $\mathbf{A}^\dagger$ containing $\mathbf{A}, \mathbf{B}, \mathbf{C}$. By definition of s-separation, it suffices to restrict our attention to walks contained in $\mathbf{A}^\dagger$. Fix a walk α from $A \in \mathbf{A}$ to $B \in \mathbf{B}$ open in $\mathcal{G}_{\mathbf{A}^\dagger}$ given $\mathbf{C}$. We will construct a path β from A to B in $(\mathcal{G}_{\mathbf{A}^\dagger})^a$ which does not intersect $\mathbf{C}$. Since α is open, every section $\alpha_1, \ldots, \alpha_k$ in α is open. We will first construct a walk $\alpha^\dagger$ in $(\mathcal{G}_{\mathbf{A}^\dagger})^a$ consisting of fragments corresponding to sections in α_i, and then simplify this walk to a path that does not intersect $\mathbf{C}$. If α_i is a non-collider section, let $\alpha_i^\dagger$ consist of the undirected edges corresponding to those in α_i. If α_i is a collider section with end points C, D, let $\alpha_i^\dagger$ consist of $C - D$. Then the starting vertex of $\alpha_1^\dagger$ is A, the ending vertex of $\alpha_k^\dagger$ is B, $\alpha_1^\dagger, \ldots, \alpha_k^\dagger$ are undirected walks that do not intersect $\mathbf{C}$ by construction, and for each $i \in 1, \ldots, k-1$ either $\alpha_i^\dagger$ shares the ending vertex with the starting vertex of $\alpha_{i+1}^\dagger$, or the ending vertex of $\alpha_i^\dagger$ and the starting vertex of $\alpha_{i+1}^\dagger$ are neighbors. Thus, we can construct a walk from these walks with a starting vertex A, ending vertex B, and which does not intersect $\mathbf{C}$. But this means we can construct a path β with the same property.

Fix a minimal path β from $A \in \mathbf{B}$ to $B \in \mathbf{B}$ that does not intersect $\mathbf{C}$ in $(\mathcal{G}_{\mathbf{A}^\dagger})^a$. We will construct a walk α from A to B s-connected given $\mathbf{C}$ in $\mathcal{G}_{\mathbf{A}^\dagger}$. Let the edges of β be $b_1, \ldots b_k$. We will construct α by replacing all b_i that do not exist in $\mathcal{G}_{\mathbf{A}^\dagger}$ by a witnessing collider walk, and all other b_i between C, D by the (possibly directed or bidirected) edge between C, D in $\mathcal{G}_{\mathbf{A}^\dagger}$. The result is clearly a walk. Furthermore, all non-collider sections on this walk do not intersect $\mathbf{C}$, and all collider sections are in $\mathbf{A}^\dagger$, so in the anterior of $\mathbf{A} \cup \mathbf{B} \cup \mathbf{C}$. By lemma 5.2, there exists a walk from A to B s-connected given $\mathbf{C}$ in $\mathbf{A}^\dagger$. $\square$

Theorem 5.3. *For a SG $\mathcal{G}$, if $p(\mathbf{V}) \in \mathcal{P}(\mathcal{G})$ and is positive, then $p(\mathbf{V})$ factorizes with respect to $\mathcal{G}$.*

Proof: Fix any $\mathbf{D} \in \mathcal{A}(\mathcal{G})$, and a topological ordering $\prec$. By the chain rule of probabilities, $p(\mathbf{D}) = \prod_{V \in \mathbf{D}} p(V \mid \text{pre}_{\mathcal{G},\prec}(V) \cap \mathbf{D})$ which is equal to $\prod_{\mathbf{S} \in \mathcal{D}^a(\mathcal{G}_\mathbf{D}) \cup \mathcal{B}^*(\mathcal{G}_\mathbf{D})} \prod_{V \in \mathbf{S}} p(V \mid \text{pre}_{\mathcal{G},\prec}(V) \cap \mathbf{D})$ since non-trivial blocks and districts partition $\mathbf{V}$. This in turn is equal to $\prod_{\mathbf{S} \in \mathcal{D}^a(\mathcal{G}_\mathbf{D}) \cup \mathcal{B}^*(\mathcal{G}_\mathbf{D})} \prod_{V \in \mathbf{S}} p(V \mid \text{pre}_{\mathcal{G},\prec}(V) \cap \text{pa}^*_{\mathcal{G}}(\mathbf{S}))$, by assumption. This implies that we obtain the outer level factorization: $p(\mathbf{D}) = \prod_{\mathbf{S} \in \mathcal{D}^a(\mathcal{G}_\mathbf{D}) \cup \mathcal{B}^*(\mathcal{G}_\mathbf{D})} f_\mathbf{S}(\mathbf{S} \mid \text{pa}^s_{\mathcal{G}}(\mathbf{S}))$. That the inner factorization holds for any $f_\mathbf{S}(\mathbf{S} \mid \text{pa}^s_{\mathcal{G}}(\mathbf{S}))$ for $\mathbf{S} \in \mathcal{B}^*(\mathcal{G})$ for a positive $p(\mathbf{V})$ follows from Theorem 3.36 in [5] (and ultimately the Hammersley Clifford theorem for UG models). $\square$

References

[1] J. Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1(3):195–200, 1964.

[2] Z. Cai, M. Kuroki, J. Pearl, and J. Tian. Bounds on direct effects in the presence of confounded intermediate variables. *Biometrics*, 64:695 – 701, 2008.

[3] M. Drton. Discrete chain graph models. *Bernoulli*, 15(3):736–753, 2009.

[4] R. J. Evans and T. S. Richardson. Markovian acyclic directed mixed graphs for discrete data. *Annals of Statistics*, pages 1–30, 2014.

[5] S. Lauritzen. *Graphical Models*. Oxford, U.K.: Clarendon, 1996.

[6] J. Pearl. *Probabilistic Reasoning in Intelligent Systems*. Morgan and Kaufmann, San Mateo, 1988.

[7] T. Richardson and P. Spirtes. Ancestral graph Markov models. *Annals of Statistics*, 30:962–1030, 2002.

[8] T. S. Richardson. Markov properties for acyclic directed mixed graphs. *Scandinavial Journal of Statistics*, 30(1):145–157, 2003.

[9] I. Shpitser, R. J. Evans, T. S. Richardson, and J. M. Robins. Introduction to nested Markov models. *Behaviormetrika*, 41(1):3–39, 2014.

[10] M. Studeny. Bayesian networks from the point of view of chain graphs. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence (UAI-98)*, pages 496–503. Morgan Kaufmann, San Francisco, CA, 1998.

[11] J. Tian and J. Pearl. A general identification condition for causal effects. In *Eighteenth National Conference on Artificial Intelligence*, pages 567–573, 2002.

[2] This proof follows lemma 3 in [8].

[12] T. J. VanderWeele, E. J. T. Tchetgen, and M. E. Halloran. Components of the indirect effect in vaccine trials: identification of contagion and infectiousness effects. *Epidemiology*, 23(5):751–761, 2012.

[13] T. S. Verma and J. Pearl. Equivalence and synthesis of causal models. Technical Report R-150, Department of Computer Science, University of California, Los Angeles, 1990.