Submitted to the Annals of Statistics

SUPPLEMENT TO "MARGINALIZATION AND CONDITIONING FOR LWF CHAIN GRAPHS"

By Kayvan Sadeghi

University of Cambridge

In the Appendix, we provide proofs of non-trivial lemmas, propositions, and theorems in [1] as well as some more technical and yet less informative lemmas that are used in the proofs.

Appendix: proofs.

Proof of Lemma 1. ($\Rightarrow$) Suppose that there is a c-connecting walk π between i and j given C. Consider the shortest subpath ρ_0 of the section ρ of π between k and l. If ρ is a collider then a node of ρ is in C, and since all the nodes on ρ (including those on ρ_0) are connected by lines, they are all in $C \cup \mathrm{ant}(C)$. If ρ is a non-collider then all the nodes on ρ (including those on ρ_0) are outside C. Hence, by replacing all such ρ by ρ_0 we obtain the desired walk.

($\Leftarrow$) Suppose that there is a walk π between i and j whose sections are all paths and nodes of every collider section are in $C \cup \mathrm{ant}(C)$, and non-collider sections are outside C. We keep all non-collider sections of π intact. For a collider section ρ between k and l, if there is a node of ρ in C, we keep it intact. Otherwise we replace ρ with $\rho_4 = \langle k, \rho_1, \rho_2, c, \rho_2^r, \rho_3, l\rangle$, where ρ_1 is a subpath of ρ between k and h , ρ_2 is a semi-directed path from h to a member c of C, ρ_2^r is ρ_2 in the reverse direction, and ρ_3 is a subpath of ρ between h and l. It is easy to observe that ρ_4 is c-connecting given C. (If there is an arrow on ρ_2 then ρ_4 consists of non-collider sections containing ρ_1 and ρ_3, and a collider section containing c; otherwise ρ_4 is a collider section containing c.) In addition, ρ and ρ_4 are endpoint-identical. Hence, by this replacement for all such ρ on π, we obtain a c-connecting walk given C between i and j.

Finally, from the construction of walks that we have in both directions of the proof, it is seen that the walks are endpoint-identical. $\square$

Proof of Proposition 1. The resulting graphs have obviously the three desired types of edges, thus it is enough to prove that there is no semi-directed cycle that contains an arrow in the graph. Suppose, for contradiction, that there exists such a cycle. It is easy to observe that by replacing a generated line or arrow with the generating tripaths (cases 1, 2, 6, and 7 of Table 1) or trislide (case 8), a semi-directed path remains semi-directed. Therefore, it is implied inductively that there is a semi-directed path in the original chain graph. This also contains an arrow since an arrow can only be replaced by a tripath or a trislide that contains an arrow. This is a contradiction. □

PROOF OF LEMMA 2. (⇐) Suppose that there exists a walk π between i and j in the graph generated after applying step 1 of Algorithm 1 to G whose inner sections are all non-collider and whose inner nodes are all in M. By Algorithm 1, for a section between k and l, a line between k and l is generated, and then, for a tripath $\langle h, q, r\rangle$ consisting of a line hq with $q \in M$, the same edge as qr is generated. Therefore, a walk is generated between i and j whose inner nodes are in M, and on which lines may only be adjacent to i and j, and every section is a non-collider. By applying steps of Table 1, we trivially obtain an endpoint-identical edge between i and j.

(⇒) Suppose that there is an edge between i and j in $\alpha_{CMG}(G; M, \varnothing)$. We are only interested in the case where this edge does not exist after applying step 1 of Algorithm 1. In this case, this edge is generated by step 2 by one of the tripaths in steps 1 to 7 of Table 1 in an iteration of step 2. Each edge in the tripath may have now been generated by a tripath with the inner node in M. By an inductive argument, we imply that in the graph generated after applying step 1 of Algorithm 1 to G, there is a walk π (because of possible self-intersections) between i and j whose inner nodes are in M. We show that there is no collider section on π: If, for contradiction, there is a collider section ρ with endpoints $\langle k, \rho, l\rangle$ then it is easy to observe that, in some iteration of the algorithm, we obtain a collider tripath with endpoints k and l, but no edge can be generated between k and l by the algorithm. Hence, there is no edge between i and j in $\alpha_{CMG}(G; M, \varnothing)$, a contradiction. Since in every iteration of the algorithm, the existence of an arrowhead at sections containing i and j does not change, π remains endpoint-identical to the ij edge. □

LEMMA 12. *Let G be a CMG and M a subset of its node set. If there is a path $i - \cdots - k \leftarrow j$ or $i - \cdots - k \leftrightarrow j$ in G, and there is a semi-directed path of form $m_1 \rightarrow m_2 - \dots - m_r - i$ with $m_s \in M$, $1 \leq s \leq r$ then Algorithm 1 generates an arrow from j to i or an arc between i and j, respectively.*

PROOF. Consider the section between k, i, and m_2. By step 1 of Algorithm 1, an arrow from j to m_2 or a jm_2 arc is generated. Now by Lemma

2, when we apply step 2 of the algorithm, an arrow from j to i or an ij arc is generated. □

LEMMA 13. *Let G be a CMG and M a subset of its node set. There is a walk π in G with sections $\{\rho_1, \dots, \rho_r\}$ if and only if there is an endpoint-identical walk π' in the graph generated after applying step 1 of Algorithm 1 for M with sections $\{\rho'_1, \dots, \rho'_r\}$ such that ρ'_q is a subsection of ρ_q for $1 \leq q \leq r$. In addition, every node on π that is not on π' is on a subsection of π with endpoints l and k such that l exists on π' and is a child of a member of M, and there is an arrowhead to k on π.*

PROOF. The result follows from the fact that by replacing arrows and arcs on π' by paths in cases 8 and 9 of Table 1 (the replacements that have occurred in step 1 of Algorithm 1), sections become larger and no new section is generated; and vice versa. □

PROOF OF THEOREM 1. **(⇒)** Suppose that in $\alpha_{CMG}(\alpha_{CMG}(G; M, \varnothing); M_1, \varnothing)$, there is an edge between i and j. Notice that $i, j \notin M \cup M_1$. We prove that there is the same edge in $\alpha_{CMG}(G; M \cup M_1, \varnothing)$. Starting from an edge between i and j, we discuss the type of path or walk that exists between i and j in every graph generated by different steps of Algorithm 1:

In the graph generated before applying step 2 of Algorithm 1 to $\alpha_{CMG}(G; M, \varnothing)$ for M_1: By Lemma 2, there exists an endpoint-identical walk π between i and j whose inner sections are all non-collider and inner nodes are all in M_1.

In $\alpha_{CMG}(G; M, \varnothing)$: By Lemma 13, there is a new walk, denoted by π_1. Define l also as defined in the lemma, and notice that in this case l is both in M_1 and a child of $m_1 \in M_1$.

In the graph generated before applying step 2 of Algorithm 1 to G for M: For every edge of π_1, again by Lemma 2, there exists an endpoint-identical walk between its endpoints, but with inner nodes in M. Denote the new walk generated by replacing all edges of π_1 by endpoint-identical walks at this stage by π_2. Notice that, because of endpoint-identicality, all nodes on π_1 remain non-collider on π_2. In addition, the $m_1 l$ arrow might turn into a walk that contains a subwalk of from $m \rightarrow m_2 — \dots — m_r — l$ with $m \in M \cup M_1$ and $m_s \in M$, $2 \leq s \leq r$.

In G: Again by Lemma 13, there is a new walk, denoted by π_3. Notice that the arrow from m to m_2 might be replaced by a path, but nevertheless, by possibly changing the node m to m', there is the same type of walk from m' to l with $m' \in M \cup M_1$. In addition, $l \in M_1$ remains the same as an endpoint of subsections on which there are nodes on π_3 that are not in $M \cup M_1$.

In $\alpha_{CMG}(G; M \cup M_1, \varnothing)$: By Lemma 12, all subpaths of π_3 of form π' are replaced by the $k'l$ arrows or arcs respectively. Therefore, there is an endpoint-identical walk whose inner sections are all non-collider and whose inner nodes are all in $M \cup M_1$. By Lemma 2, we conclude that there is an endpoint-identical (i.e. the same type of) edge between i and j.

($\Leftarrow$) Suppose that there is an edge between i and j in $\alpha_{CMG}(G; M \cup M_1, \varnothing)$. Starting from this edge, we discuss the type of path or walk that exists between i and j in every graph generated by different steps of Algorithm 1:

In the graph generated before applying step 2 of Algorithm 1 to G for $M \cup M_1$: By Lemma 2, there exists an endpoint-identical walk π between i and j whose inner sections are all non-collider and inner nodes are all in $M \cup M_1$.

In G: By Lemma 13, there is a new walk, denoted by π_1. Define l also as defined in the lemma, and notice that in this case l is both in $M \cup M_1$ and a child of $m_1 \in M \cup M_1$.

In the graph generated after applying step 1 of Algorithm 1 to G for M: All subpaths of π_1 of the mentioned form and properties π' where l is a child of M can be replaced by kl arrows or lines respectively.

In $\alpha_{CMG}(G; M, \varnothing)$: Now the generated walk can be partitioned into subwalks with endpoints in outside M and all inner nodes in M (there might be single edges in the partition). All these subwalks with lengths more than two satisfy the conditions of Lemma 2 for M. Hence, there exist endpoint-identical edges between the endpoints of the subwalks. These edges form a walk, which is denoted by π_2.

In the graph generated after applying step 1 of Algorithm 1 to $\alpha_{CMG}(G; M, \varnothing)$ for M_1: Since there are no collider sections on π_1, and because of endpoint-identicality, there are no collider sections on π_2. In addition, the endpoints l (as defined) of subpaths of π_2 whose members may not be in M_1, are children of M_1. Therefore, again by applying step 1 of the algorithm for M_1 we obtain a walk with all inner nodes in M_1.

In $\alpha_{CMG}(\alpha_{CMG}(G; M, \varnothing); M_1, \varnothing)$: Now by applying Lemma 2 to the generated walk, we obtain an endpoint-identical (and hence the same type ij edge as the original ij edge). $\square$

PROOF OF PROPOSITION 2. **First, we prove that every CG G is mapped into $\mathcal{H}$:** By proposition 1, we know that the generated graphs are CMGs. We consider each case separately:

Suppose that there is a collider trislide of form $k \leftrightarrow i - \dots - j \leftarrow l$ *in the generated graph* $\alpha_{CMG}(G; M, \varnothing)$*.* We go through how this trislide has been generated by steps of Algorithm 1.

In the graph generated before applying step 2 of Algorithm 1: Since by step 2 of Algorithm 1 only case 7 of Table 1 can generate lines, by an inductive argument it is clear that between i and j there is a section. By Lemma 2, instead of the arrow from l to j, there is a walk with non-collider sections and inner nodes in M such that there is an arrowhead at the endpoint section containing j (say from node r, which may be l).

In addition, notice that G is a CG and by step 1 of Algorithm 1, no arc is generated from trislides that do not contain arcs. This fact together with Lemma 2 implies that there is a walk between i and k that only contains lines and arrows, and, on this walk, there is an arrowhead at the endpoint section containing i (say at node o, which may be i and has a parent in M).

By considering the path between r and o, we conclude that by step 1 of Algorithm 1 (case 8 of Table 1), an arrow from r to o is generated.

In $\alpha_{CMG}(G; M, \varnothing)$*:* Now by Lemma 2, and considering the walk with non-collider sections and inner node in M that connects l, r, o, and i, an arrow from l to i is generated.

Suppose that there is a collider trislide of form $k \leftrightarrow i - \dots - j \leftrightarrow l$ *in the generated graph:* r and o can be defined in the same way as in the previous case. Notice that in this case on the walk (obtained by Lemma 2) there is an arrowhead at the section containing l. By a similar argument to that in the previous case, we conclude that there is an arc generated between l and i in the generated graph. By the symmetry in the path, one can similarly obtain an arc between k and j. Furthermore, by Lemma 2, and considering the walk with non-collider sections and inner node in M that connects j, r, o, and i, there exists an arc between i and j in the generated graph, since, on this walk, there are arrowheads at both sections that contain i an j.

Now we prove that the function is surjective: Consider an arbitrary chain mixed graph H in $\mathcal{H}$. Define a chain graph G from H as follows: keep all arrows and lines of H in G and replace arcs ij with $i \leftarrow m \rightarrow j$; and define a subset M of the node set of G as the set of all such m.

We first prove that G is a CG: It only contains the two desired types of edges. In addition, it does not contain semi-directed cycles that contains an arrow since if, for contradiction, it does then it must contain the tripath $i \leftarrow m \rightarrow j$, which is impossible.

We now prove that $\alpha_{CMG}(G; M, \varnothing) = H$*:* The changes that might occur by step 1 of Algorithm 1 are only when, in H, there are the two types of collider trislides in properties 1 and 2, which correspond to the walks $k \leftarrow\!\!\!\!\prec m \succ\!\!\!\!\rightarrow i \,\text{———}\, \dots \,\text{———}\, j \leftarrow\!\!\!\!\prec l$ and $k \leftarrow\!\!\!\!\prec m_1 \succ\!\!\!\!\rightarrow i \,\text{———}\, \dots \,\text{———}\, j \leftarrow\!\!\!\!\prec m_2 \succ\!\!\!\!\rightarrow l$ in G. In the former case, the generated arrow from l to i exists in H. In the latter case, an arrow from m_2 to i is generated, but since m_2 is only adjacent to j and l, in the next step of the algorithm, it can only generate il and ij arcs, both of which exist in H; the same argument also works for the generated arrow from m_1 to j. In addition, step 9 is not applied since there are no arcs in G. By step 2 of the algorithm, the only type of tripath with inner node in M is case 4 of Table 1 (except those that are already discussed). These tripaths obviously turn into the arcs existing in H, and no other edge is generated. $\square$

PROOF OF THEOREM 2. We need to prove that $A \perp_c B \,|\, C_1$ in $G \iff A \perp_c B \,|\, C_1$ in $\alpha_{CMG}(G; M, \varnothing)$.

($\Rightarrow$) Suppose that there is a c-connecting walk π given C_1 between i and j in G. Consider all maximal subwalks of π whose inner sections are all non-collider, endpoints are not in M, and inner nodes are all in M. Notice that all nodes of π that are in M are included in these subwalks since no collider section on π_1 has all nodes in M. Denote such a subwalk by ϖ.

In the generated graph after applying step 1 of Algorithm 1: First consider the case where the endpoints of ϖ are the same node l. Sections on ϖ are non-collider, and hence, the edge between l and an endpoint of ϖ (call it m) is an arrow from m to l. We can easily obtain a shorter c-connecting walk by removing ϖ from π if, by doing so, l is on a collider section or on a non-collider section with no node in C_1. If that is not the case then there exists $l \leftarrow\!\!\!\!\prec m \succ\!\!\!\!\rightarrow l \,\text{———}\, \cdots \,\text{———}\, \circ \leftarrow\!\!\!\!\prec k$ or $l \leftarrow\!\!\!\!\prec m \succ\!\!\!\!\rightarrow l \,\text{———}\, \cdots \,\text{———}\, \circ \leftarrow\!\!\!\!\prec\!\succ\!\!\!\!\rightarrow k$, where $l \notin C_1$ but an inner node of the section containing l is in C_1. (Notice that if l is i or j then one can easily remove m from the walk.) By step 1, there is a generated lk edge. We replace all these walks with the generated edge and call the resulting walk π_1. Because the generated edges are endpoint-identical to the subwalks, π_1 is c-connecting.

In the generated graph after applying step 2 of Algorithm 1: The subwalks of π_1 with the property mentioned above now have distinct endpoints. By Lemma 2, instead of these subwalks, there are endpoint-identical edges in $\alpha_{CMG}(G; M, \varnothing)$. By replacing all the subpaths with these edges, we obtain a walk π_2. Walk π_2 exists in $\alpha_{CMG}(G; M, \varnothing)$ since there are no members of M on π_2. In addition, π_2 is c-connecting given C_1 since, because of endpoint-identicality of the generated edges to the subwalks, every node that is an inner node of a collider or a non-collider section on π_2 is an inner node of a collider or a non-collider section on π_1, and no node in C_1 on π_1 has been taken out.

($\Leftarrow$) Suppose that there is a c-connecting walk π given C_1 between i and j in $\alpha_{CMG}(G; M, \varnothing)$. We show what types of walks generated π at each step of Algorithm 1.

In the graph before applying step 2 of Algorithm 1: By Lemma 2, for every edge kl on π, there is an endpoint-identical walk π' between k and l with the stated properties in the lemma. By replacing every edge on π by such π', we obtain a walk π_1. We prove that π_1 is c-connecting given C_1: Notice that π' is obviously c-connecting. In addition, because of endpoint-identicality, for a replaced edge kl, if l is an inner node of a collider or a non-collider section, after the replacement, it remains an inner node of a collider or non-collider section respectively, and all added nodes are in M.

In G, before applying step 1 of Algorithm 1: Now a uv edge on π_1 might have been replaced by a path by step 1 of the algorithm, where u is a child of $m \in M$. By all such replacements, we obtain a larger walk π_2. Again, because of endpoint-identicality, if u is on a collider section or a non-collider section ρ_1 on π_1 then it remains on a (possibly larger) collider section or a non-collider section ρ_2 on π_2 respectively. If ρ_2 is non-collider and all inner nodes of the new path are outside C_1 then it is clearly open on π_2. If ρ_2 is non-collider with a node in C_1 then we modify π_2 by adding the subwalk $\langle u, m, u\rangle$ (i.e. the arrow from m to u in both directions) to π_2. Now the subpath of ρ_2 between v and u becomes a collider section and open on π_2, and the rest of ρ_2 (with an arrow pointing to it from m) remains a non-collider section and open. If ρ_2 is a collider, it is clearly open since there is already a node in C_1 on ρ_1. Therefore, by an inductive argument, π_2 is a c-connecting walk. ☐

PROOF OF LEMMA 3. ($\Leftarrow$) Suppose that there exists a walk π between i and j in the generated graph after step 2 whose inner sections are all collider and in $C \cup \mathrm{ant}(C)$, and endpoint sections contain a single node. We prove the result by induction on the number of edges of π. If it is 1 then we are clearly done. If it is $n > 1$ then consider the trislide $\tau = \langle i\rho k\rangle$ on π, where ρ is a section. By step 3 of the algorithm, an endpoint-identical edge ik is generated. Notice that ik is either an arrow or an arc unless possibly $k = j$. Now by replacing τ by the ik edge, we obtain a shorter walk with the same properties. By the induction hypothesis, we obtain the result.

($\Rightarrow$) Suppose that there is an edge between i and j in the graph generated after step 3 of Algorithm 2. If this edge were generated by step 3 of Algorithm 2 then it would be generated by one of the first three trislides in Table 2 in an iteration of step 3 of the algorithm. Each arrow or arc on the trislide may now have been generated by a trislide with inner nodes in $C \cup \mathrm{ant}(C)$ (since no generated line can be used in the iterations). Since the trislides are endpoint-identical to the generated edge, it is implied that all sections remain collider. By an inductive argument, we imply that, in the graph generated after applying step 2 of the algorithm, there is an endpoint-identical walk between i and j whose inner nodes are in $C \cup \mathrm{ant}(C)$ and all sections are collider. In addition, i and j are clearly not adjacent to a line on this walk, i.e., endpoint sections contain a single node. $\square$

PROOF OF LEMMA 4. The first result for step 4 is trivial, and for step 3 follows directly from Lemma 3. This implies that if a generated line lies on a collider section after step 3 then since $j \in S$, by step 4, all arrowheads at the section will be removed. $\square$

PROOF OF LEMMA 5. One direction of the proof is obvious since steps 1, 2, and 3 of Algorithm 2 do not remove or replace any edges, and by removing an arrowhead at an arrow pointing to i by step 4, no new node can become an anterior of i. Thus, suppose that $i \in \mathrm{ant}(C)$ after step 4 of the algorithm. We go back on the steps of the algorithm in order to show that i has been in $\mathrm{ant}(C)$.

Before applying step 4 of Algorithm 2: Suppose that there is a node k on the semi-directed path π from i to C such that, on π, there is an arrowhead at k on the opposite direction of π. In addition, suppose that this arrowhead has been removed by step 4. It then holds that $k \in C \cup \mathrm{ant}(C)$. By considering the closest of such nodes to i on π, i is an anterior of k, and consequently C.

Before applying step 3 of Algorithm 2: Consider the closest arrow to i on π that is generated by step 3. The result then follows from Lemma 4.

Before applying step 2 of Algorithm 2: The only possible arrow on π (say from k to l) can be generated by step 2 (case 4 of Table 2). This implies that $k \in \mathrm{ant}(l)$. By an inductive argument, this implies the result. $\square$

PROOF OF PROPOSITION 3. Graphs generated by Algorithm 2 have the three desired types of edges. We prove that there is no semi-directed cycle with an arrow in a generated chain mixed graph from G. Suppose, for contradiction, that a generated graph does contain a semi-directed cycle π with an arrow. Since π does not exist in G, at least one arrow, say from j to i, or a line, say between k and l has been generated by Algorithm 2. If ij or lk has been generated by steps 3 or 4 of the algorithm then by Lemma 4, $j, k, l \in S$ in G. This implies that there should be no arrow on π, a contradiction.

Thus, the only option that is left is that ij has been generated by step 2, case 4 of Table 2. In this case $j \in \mathrm{ant}(i)$ in G with an arrow existing on the directed path from j to i. By considering all arrows generated by this step of the algorithm on π, we conclude that there is a semi-directed cycle with an arrow in G, a contradiction. □

PROOF OF LEMMA 6. **We first prove that there is an ij edge in $\alpha_{CMG}(G; \varnothing, C)$ if and only if there is a walk as described in the lemma in G:**

(⇒) Suppose that in $\alpha_{CMG}(G; \varnothing, C)$ there is an edge between i and j. We will follow how this edge might have been generated by the steps of Algorithm 2.

In the graph generated before applying step 4: It is clear that there is an ij edge.

In the graph generated before applying step 3: Now, by Lemma 3, there exists an endpoint-identical walk π between i and j to the edge ij whose inner sections are all collider and in S, and whose endpoint sections contain a single node. Notice that this means that all edges on π are either lines or arcs except possibly those containing i and j.

In G: By replacing arcs or arrows on π by endpoint-identical paths (provided in cases 4 and 5 of Table 2), only collider sections on π become larger. The newly added nodes to the sections will obviously be in S since they are anteriors of the rest of the section, which is in S – the only exception is when there is an arrowhead at i (or j) and the section containing i gets larger. In this case, $i \in \mathrm{sp}(k)$, for $k \in S$.

(⇐) Suppose that in G, there exists a walk π as described in the lemma. The edges of π are all arcs and lines except possibly those including i and j. We will go through how this walk changes by the steps of Algorithm 2.

In the graph generated after applying step 2: The endpoint sections turn into single nodes, and other sections may get shortened, but since the generated edges are endpoint-identical to the generating paths (provided in cases 4 and 5 of Table 2), inner sections of the resulting walk are still collider. Lemma 5 implies that the inner sections stay in S.

In $\alpha_{CMG}(G; \varnothing, C)$: The generated walk in the previous step satisfies the conditions of Lemma 3. Hence, there is an ij edge generated by step 3, which keeps existing after step 4.

We now prove the second claim in the lemma: Since all generated edges by steps 2 and 3 of the algorithm (all cases of Table 2) are endpoint identical to their generating paths, the generated edge after step 3 and the walk in G are endpoint-identical. Step 4 changes endpoint-identicality only when it removes the arrowhead at i, which always and only happens when $i \in \mathrm{ant}(C)$. □

Proof of Theorem 3. Notice that $i, j \notin C \cup C_1$. **We first prove that there is an ij edge in $\alpha_{CMG}(\alpha_{CMG}(G;\varnothing,C);\varnothing,C_1)$ if and only if there is an ij edge in $\alpha_{CMG}(G;\varnothing,C\cup C_1)$:**

By Lemma 6, there is an edge between i and j in $\alpha_{CMG}(\alpha_{CMG}(G;\varnothing,C);\varnothing,C_1)$ if and only if there is a walk π as described in the lemma between i and j in $\alpha_{CMG}(G;\varnothing,C)$ with inner sections in $S_1 = C_1 \cup \mathrm{ant}(C_1)$.

Notice that by Lemma 4, lines on the inner sections of π exist in G. In addition, there at most two arrows might exist on π, which are from the endpoints i and j. Now again by Lemma 6, instead of a kl arc on π, in G, there is an endpoint-identical walk π' as described in the lemma between k and l with inner sections in $S = C \cup \mathrm{ant}(C)$. By replacing kl by π', one obtains a walk with the same properties as in Lemma 6 for $S \cup S_1$. Inductively, we replace all such kl arcs. We also replace a possible arrow (say from i to h) by a walk with properties as described in Lemma 6, where there might be an arrowhead at i with $i \in \mathrm{ant}(C)$. By all these replacements, one obtains a walk π_1 in G. Since conditions of Lemma 6 are both necessary and sufficient, it holds that there is the walk π in $\alpha_{CMG}(G;\varnothing,C)$ if and only if there is the walk π_1 in G.

Walk π_1 satisfies the properties in Lemma 6 for $S \cup S_1$. Again by Lemma 6, there is the walk π_1 in G if and only if there is an ij edge in $\alpha_{CMG}(G;\varnothing,C\cup C_1)$.

We now prove that the ij edge is the same in both graphs: We only need to show that there is an arrowhead at i on the ij edge in $\alpha_{CMG}(\alpha_{CMG}(G;\varnothing,C);\varnothing,C_1)$ if and only if there is an arrowhead at i on the ij edge in $\alpha_{CMG}(G;\varnothing,C\cup C_1)$. This follows from the second part of Lemma 6 and the fact that if $i \in \mathrm{ant}(C) \cup \mathrm{ant}(C_1)$ in G then there is no arrowhead at the ij edge in $\alpha_{CMG}(\alpha_{CMG}(G;\varnothing,C);\varnothing,C_1)$ or $\alpha_{CMG}(G;\varnothing,C\cup C_1)$. Below we prove the latter claim:

The result for $\alpha_{CMG}(G;\varnothing,C\cup C_1)$ is again clear by Lemma 6. We now condider $\alpha_{CMG}(\alpha_{CMG}(G;\varnothing,C);\varnothing,C_1)$. If $i \in \mathrm{ant}(C)$ then there is no arrowhead at i on π in $\alpha_{CMG}(G;\varnothing,C)$. If $i \in \mathrm{ant}(C_1) \setminus \mathrm{ant}(C)$ then consider the semi-directed path from i to a member of C_1 in G. This path remains intact in $\alpha_{CMG}(G;\varnothing,C)$ since $i \notin \mathrm{ant}(C)$. Hence, the arrowhead at i on the ij edge will be removed in $\alpha_{CMG}(\alpha_{CMG}(G;\varnothing,C);\varnothing,C_1)$.

□

Proof of Theorem 4. We prove that $A \perp_c B \mid C \cup C_1$ in G if and only if $A \perp_c B \mid C_1$ in $\alpha_{CMG}(G; \varnothing, C)$.

($\Leftarrow$) Suppose that there is a c-connecting walk π given $C \cup C_1$ between i and j in G. We apply the steps of Algorithm 2 to this walk. Consider all maximal subwalks of π whose inner sections are all collider and in C, and endpoints are single nodes and not in C. Notice that all nodes of π that are in C are included in these subwalks since no non-collider section on π has a node in C. Denote such a subwalk by ϖ.

After applying step 2: First consider the case where the endpoints of ϖ are the same node l. Sections on ϖ are collider, and hence, the edge between l and an endpoint of ϖ (call it c) has an arrowhead at c. We can easily obtain a shorter c-connecting walk by removing ϖ from π if, by doing so, l is on a collider section or on a non-collider section with no node in $C \cup C_1$. First, this implies that the cl edge is an arc. In addition, if that is not the case then there exists $l \leftarrow\!\!\!\rightarrow c \leftarrow\!\!\!\rightarrow l$ —— $\cdots$ —— $\circ \leftarrow k$ or $l \leftarrow\!\!\!\rightarrow c \leftarrow\!\!\!\rightarrow l$ —— $\cdots$ —— $\circ \leftarrow\!\!\!\rightarrow k$, where $l \notin C \cup C_1$ but an inner node of the section containing l is in C_1. (Notice that if l is i or j then one can easily remove c from the walk.) By step 2, there is a generated lk edge. We replace all these walks with the generated edge and call the resulting walk π_1. Because the generated edges are endpoint-identical to the subwalks, π_1 is c-connecting.

After applying step 3: By Lemma 1, there is an alternative c-connecting walk π_1' to π_1, where all sections are paths and inner nodes of collider sections are in $C \cup C_1 \cup (\mathrm{ant}(C) \cup \mathrm{ant}(C_1))$. Consider all maximal subwalks of π_1' whose inner sections are all collider and in $C \cup \mathrm{ant}(C)$, and endpoints are single nodes and not in C. Because of the previous step, the endpoints of such subwalks are distinct nodes. Now, by Lemma 3, instead of these subwalks, there are endpoint-identical edges. By replacing all the subwalks with these edges, we obtain a walk π_2. Walk π_2 is c-connecting given C_1 since generated edges on π_2 are endpoint-identical to the subpaths on π_1' that have been replaced.

After applying step 4: By this step, no collider sections turn into a non-collider one on π_2 since if an arrowhead on a node k is removed then $k \in \mathrm{ant}(C)$ in G and so are all inner nodes of the section that contains k. Hence, k cannot be on π_2 by how π_2 is generated. Therefore, π_2 is a c-connecting walk given C_1 in $\alpha_{CMG}(G; \varnothing, C)$.

($\Rightarrow$) Suppose that there is a c-connecting walk π given C_1 between i and j in $\alpha_{CMG}(G; \varnothing, C)$. In G, we obtain a walk π' by replacing every edge on π with the corresponding walks described in Lemma 6. All these generated walks by edges of π are c-connecting given $C \cup C_1$ themselves. Hence, if their endpoints are open then π' would be c-connecting given $C \cup C_1$.

If a generated subwalk on π' is endpoint-identical to the generating edge on π with endpoint sections containing a single node then it is open. Hence, we need to consider two cases where this does not happen for a generated subwalk:

1) When the endpoint sections of the generated walks contain more than a node, we know that there is an arrowhead at the section, and the endpoint k is a spouse of $s \in S$. It is possible that the endpoint section ρ is not open on π' (but the corresponding edge is open on π) if it is a non-collider with a node in C. In this case add $\langle k, s, k\rangle$ (i.e., repeating the ks edge twice) instead of k to π'. This makes ρ collider and also adds a collider section s (containing a single node) and one non-collider section containing k, which are all open.

2) We know that the generated walks on π' and the generating edges on π are endpoint-identical except when there is an arrowhead at the endpoint section ρ' containing l and there is a semi-directed path ϖ from l to $c \in C$ in G. In this case, add $\langle \varpi, \varpi^r\rangle$ instead of l to π' (i.e. go from l to c and come back to l on ϖ). By this method, we split the collider ρ' at l into two subpaths, both of which are non-colliders, and obtain other open non-collider sections along ϖ and a collider section c. □

Proof of Proposition 4. The generated graphs obviously contain only lines and arrows, thus it is enough to prove that they do not contain semi-directed cycles with an arrow. Suppose, for contradiction, that a generated graph does contain a semi-directed cycle π with an arrow. If a line ij on π has been generated by step 4 then $i, j \in S$ in G and, therefore, all nodes on π are in S . This implies that there is no arrow on π, a contradiction. If a line kl has been generated by step 3 then it is easy to see that both $k, l \in S$, and again there is no arrow on π, a contradiction. Therefore, all lines on π exist in the original graph, and no arrows are generated by the algorithm. Hence, π exists in the original graph, a contradiction. □

Proof of Lemma 7. We show that for any choice of C, $i \perp j \mid C$ dos not hold: Suppose that there is an arrow from an inner node k to j. If any of the inner nodes is in C then i and j are dependent given C. If no inner node is in C then the subwalk between i and k in addition to the kj arrow constitutes a connecting walk given C. □

Proof of Lemma 8. **We prove the first claim:**

(⇒) Suppose that in $\alpha_{CMG}(\alpha_{CMG}(G; \varnothing, C); M, \varnothing)$ there is an edge between i and j.

In the graph generated before applying step 2 of Algorithm 1 to $\alpha_{CMG}(G;\varnothing,C)$: By lemma 2, there exists a walk π between i and j whose inner sections are all non-collider and inner nodes are all in M.

In $\alpha_{CMG}(G;\varnothing,C)$: By Lemma 13, there is a walk π_0 between i and j with the same non-collider sections. In addition, every node on π_0 on section ρ that is not in M is on a subsection with an endpoint that is the endpoint of ρ as well with an arrowhead pointing to it from the other adjacent node on π_0. The other endpoint h is in M and a child of a member of M.

In G: For every edge kl on π_0, by Lemma 6, there exists a walk π' between k and l whose inner sections are all collider and in $C \cup \mathrm{ant}(C)$. We denote the walk in this graph that consists of all such adjacent π' of π_0 by π_1. Even if the endpoint sections of π' are not single elements or π' is not endpoint-identical to the kl edge, all the existing non-collider sections remain non-collider (although some sections might become larger). It is then observed that all non-collider sections on π_1 have all inner nodes outside C, and all collider sections have inner nodes in $C \cup \mathrm{ant}(C)$. In addition, every node on π_1 on section ρ' that is not in M is on a subsection with an endpoint that is the endpoint of ρ' as well with an arrowhead pointing to it from the other adjacent node on π_1. The other endpoint h is in M and either a child of a member of M or a spouse of a member of $C \cup \mathrm{ant}(C)$.

($\Leftarrow$) Suppose that there is a walk between i and j in G with the two mentioned properties. In place of this walk, we have the following walks in the following graphs:

After applying step 1 of Algorithm 1 to G: By this step it can bee seen that all subwalks containing non-collider sections outside M with an endpoint that is a child of M get closed, and, therefore, we obtain a walk on which (i) all nodes on collider sections are in $C \cup \mathrm{ant}(C)$; (ii) (a) all nodes on non-collider sections are in M or (b) on the non-collider section one endpoint is in M and a spouse of a node in $C \cup \mathrm{ant}(C)$, and the other endpoint has an arrowhead at it from the adjacent node on the walk.

In $\alpha_{CMG}(G;M,\varnothing)$: By Lemma 2, we obtain a walk on which all sections are collider and in $C \cup \mathrm{ant}(C) \cup \mathrm{ant}(j)$. Notice that the spouses of the endpoints of non-collider sections in the previous walk, which are in $C \cup \mathrm{ant}(C)$, appear on the generated walk.

In $\alpha_{CMG}(G;M,C)$: By Lemma 6, we obtain an edge.

We now prove the second claim: We go through the corresponding walks in the intermediate graph, provided above. By lemma 6, the ij edge in $\alpha_{CMG}(G;M,C)$ and the corresponding walk in $\alpha_{CMG}(G;M,\varnothing)$ remain endpoint-identical except when there is an arrowhead at the endpoint section containing, say, i, and $i \in \mathrm{ant}(C)$ in $\alpha_{CMG}(G;M,\varnothing)$. This walk, by Lemma

2, is endpoint-identical to the corresponding walk in the graph generated after applying step 1 of Algorithm 1 to G. Since the anterior set does not change at this step and the next step in G, and since step 1 of Algorithm 1 generates endpoint-identical edges, the result follows for the corresponding walk in G. $\square$

LEMMA 14. *For a chain mixed graph G and M and C subsets of its node set, if $i \in \mathrm{ant}(j)$ in $\alpha_{CMG}(\alpha_{CMG}(G;\varnothing,C);M,\varnothing)$ then $i \in \mathrm{ant}(C \cup \{j\})$ in G.*

PROOF. The proof follows from Lemma 8 by the following observations: A line between k and l or an arrow from k to l on the semi-directed walk from i to j in $\alpha_{CMG}(\alpha_{CMG}(G;\varnothing,C);M,\varnothing)$ is not endpoint-identical to the corresponding walk π in G if and only if $k \in \mathrm{ant}(C)$ in G. If they are endpoint-identical then start from k and move towards l on π. At each step we either reach a collider section and conclude that $k \in \mathrm{ant}(C)$, or we finally reach l and conclude that $k \in \mathrm{ant}(l)$. By an inductive argument on the nodes of π, we obtain the result. $\square$

PROOF OF PROPOSITION 6. **We first prove that there is an ij edge in $\alpha_{CMG}(\alpha_{CMG}(G;M,\varnothing);\varnothing,C)$ if and only if there is an ij edge in $\alpha_{CMG}(\alpha_{CMG}(G;\varnothing,C);M,\varnothing)$:** We go through Algorithms 1 and 2 to follow the types of walks corresponding to the ij edge in any of these graphs in each step of the algorithms.

(⇒) Suppose that in $\alpha_{CMG}(\alpha_{CMG}(G;M,\varnothing);\varnothing,C)$ there is an edge between i and j.

In $\alpha_{CMG}(G;M,\varnothing)$: By Lemma 6, there is a walk π between i an j with the properties described in the lemma.

In the graph generated before applying step 2 of Algorithms 1 to G: For every edge kl on π, by Lemma 2, there exists an endpoint-identical walk π' between k and l whose inner sections are all non-collider and inner nodes are all in M. We denote the walk that consists of all such adjacent π' by π_0. It is easy to observe that all collider sections are in $C \cup \mathrm{ant}(C)$. In addition, either the endpoint sections of π_0 still satisfy the conditions of Lemma 6, or the endpoints that are not single elements become children of members of M.

In G: By Lemma 13, there exists another walk π_1, on which, all collider sections are in $C \cup \mathrm{ant}(C)$. In addition, collider and non-collider sections remain intact. In addition, it can be seen that on π_1, the conditions for endpoint sections described in the previous paragraph still hold.

In $\alpha_{CMG}(\alpha_{CMG}(G;\varnothing,C);M,\varnothing)$: The walk described in the previous paragraph in G satisfies the conditions of Lemma 8. Hence, by this lemma, we obtain the result.

($\Leftarrow$) Suppose that in $\alpha_{CMG}(\alpha_{CMG}(G;\varnothing,C);M,\varnothing)$ there is an edge between i and j. By Lemma 8, there is a walk π_1 as described in the lemma in G. We now continue to check how this walk alters along the steps of the relevant algorithms:

In the graph generated after applying step 1 of Algorithm 1 to G: All maximal subsections of non-collider sections whose nodes are outside M, but an endpoint l is in M and a child of M can be replaced by an endpoint-identical edge. By all such replacements, we obtain a walk π_2, which contains collider sections in $C \cup \mathrm{ant}(C)$ and non-collider sections outside C. In addition, every node on π_2 on section ρ that is not in M is on a subsection with an endpoint that is the endpoint of ρ as well with an arrowhead pointing to it from the other adjacent node on π_2. The other endpoint h is in M and a spouse of a member of $C \cup \mathrm{ant}(C)$.

In $\alpha_{CMG}(G;M,\varnothing)$: First consider a non-collider trislide $\langle r,\rho',q\rangle$ where ρ' has members outside M. In addition, say r is the endpoint of ρ' with an arrowhead pointing to it from the other adjacent node on π_2. Consider the node h as defined in the above paragraph, which is a spouse of $s \in C \cup \mathrm{ant}(C)$. Denote the adjacent node to h closer to r by t and the adjacent node to h closer to q by v. By this step, an edge between t and v as well as ts and sv arcs are generated.

In addition, by using Lemma 2, we replace the maximal subwalks of π_2 that contain only non-collider sections and in which all nodes are in M, but endpoints are outside M, by the generated endpoint-identical edges. By all these replacements, we obtain a walk π_3 that contains collider sections with nodes in $C \cup \mathrm{ant}(C)$ and non-collider sections outside C. In particular, we obtain an sq arc as well as an arrow from t to q.

In $\alpha_{CMG}(\alpha_{CMG}(G;M,\varnothing);\varnothing,C)$: By Lemma 6, instead of all subwalks of π_3 that contain inner collider sections, there exists an edge. In addition, for non-collider sections, the collider tripath $\langle t,s,q\rangle$ (described in the above paragraph) generates a tq arc. Because of the arrow from t to q and the subwalk of the trislide between r and t, and by Lemma 7, we conclude that the graph is not maximal except when there is an endpoint-identical edge between r and q. Therefore, by an inductive argument, there is an edge between the endpoints of π_3.

We now prove that the ij edge is of the same type in both graphs: For every graph generated by a step of the algorithm, we discussed a walk between i and j in both directions of the proof above. We focus on the arrowhead pointing to i on these walks:

By Lemma 6, there is no arrowhead pointing to i on the ij edge in $\alpha_{CMG}(\alpha_{CMG}(G;M,\varnothing);\varnothing,C)$ if and only if there is no arrowhead pointing to i or there is an arrowhead at i and $i \in \mathrm{ant}(C)$ in $\alpha_{CMG}(G;M,\varnothing)$.

By Lemma 2 and the fact that the anterior sets do not change at this step, the statement above is equivalent to no arrowhead pointing to i or an arrowhead pointing to i only when $i \in \mathrm{ant}(C)$ in the graph generated before applying step 2 of Algorithms 1 to G.

The result then follows from Lemma 8 for the corresponding walk in $\alpha_{CMG}(\alpha_{CMG}(G;\varnothing,C);M,\varnothing)$.

□

PROOF OF PROPOSITION 7. **We first prove that every CG G is mapped into $\mathcal{H}$:** By propositions 1, 3, and 6, we conclude that the generated graphs are CMGs. By Proposition 2, we know that $H = \alpha_{CMG}(G;M,\varnothing)$ is in $\mathcal{H}$. We need to prove that H is mapped into $\mathcal{H}$ by conditioning.

Suppose that there is a collider trislide π of form $k \leftarrow\!\!\!\!\rightarrow i$ —— ... —— $j \leftarrow l$ in the generated graph $\alpha_{CMG}(G;M,C)$. By Lemma 4, the lines on π exist in H. By Lemma 6, instead of the lj arrow and the ki arc, there are walks π_1 and π_2, respectively, as described in the lemma, in H. Consider the node r adjacent to the endpoint section containing j on π_1, and the node h that is the other endpoint of the endpoint section containing i on π_2. (Notice that r may be j and h may be i.)

Since H is in $\mathcal{H}$, there is an arc (or an arrow if possibly $h = l$) between r and h. Now the walk containing the subwalk of π_1 between l and r, the rh arc, and the subsection on π_2 between h and i satisfies the conditions of the walk described in Lemma 6. Hence, by this lemma, there is an arrow from l to i in $\alpha_{CMG}(G;M,C)$.

If there is a collider trislide of form $k \leftarrow\!\!\!\!\rightarrow i$ —— ... —— $j \leftarrow\!\!\!\!\rightarrow l$ in the generated graph then by the same argument as that in the previous paragraph (and considering the fact that $k, l \notin S$), there are il and kj arcs in the generated graph. In addition, this time the walk containing the subwalk of π_1 between j and r, the rh arc, and the subsection on π_2 between h and i satisfies the conditions of the walk described in Lemma 6. Hence, there is an arc between j and i in $\alpha_{CMG}(G;M,C)$.

We now prove that the function is surjective: by Proposition 2, after marginalization, CGs are surjectively mapped onto $\mathcal{H}$. Thus, by letting $C = \varnothing$, Proposition 6, and the fact that $\alpha_{CMG}(G;\varnothing,\varnothing) = G$, CGs are surjectively mapped onto $\mathcal{H}$ after marginalization and conditioning. □

Proof of Proposition 8. By Propositions 1 and 3, we know that, after step 2 of Algorithm 3, we obtain a CMG. Steps 3 and 4 do not generate a semi-directed cycle with an arrow by generating an arrow from j to i: This is because if, for contradiction, that is the case then in the previous iteration of step 4, $j \in \text{ant}(k)$ and $k \in \text{ant}(i)$ which imply that $j \in \text{ant}(i)$, and, in the previous iteration of step 3, $j \in \text{ant}(i)$. This is a contradiction since it means by induction that the semi-directed cycle with an arrow exists in the generated graph after applying step 2.

Step 5 obviously removes all arcs with one endpoint that is an anterior of the other endpoint. This step also does not generate semi-directed cycles with an arrow by replacing an arc ij by an arrow from j to i or an ij line: this is because if, for contradiction, that is the case then $j \in \text{ant}(i)$ in the generated graph after applying step 4, which is a contradiction since it means by induction that the semi-directed cycle with an arrow exists in this graph. □

Proof of Lemma 9. We show that at every step of Algorithm 3, a semi-directed path from i to j remain semi-directed and vice versa. For step 3 of the algorithm, the result is clear since the generating path of an arrow from h to l is semi-directed from h to l. For step 4, this is correct as well since there is a node k on the generating path such that $k \in \text{ant}(l)$, and, on the generating path, $h \in \text{ant}(k)$. This is also true for step 5 since if an arc turns into an arrow from h to l then h is already an anterior of l. □

Proof of Lemma 10. **First, we prove the first claim:**

(⇒) Suppose that there is an ij edge in $\alpha_{CMG.AnG}(H)$. We see how this edge changes by steps of Algorithm 3:

Before applying step 5: There is still an edge between i and j.

Before applying step 4: Instead of an arrow or an arc ij at some iteration of this step of the algorithm, there may be a path between i and j, consisting of one inner collider section and with inner nodes, say, in $\text{ant}(i)$. By any other iteration, the arrow or the arc kl might be replaced by another such path. By this replacement, we obtain a path (by discarding the intersection of lines) with all inner sections to be collider. Notice that by Lemma 9, at no iteration the anterior set of the endpoints changes. In addition, regardless of whether inner nodes of the path between k and l are anteriors of k or l, all inner nodes are anteriors of i. By an inductive argument, we finally obtain a subprimitive inducing path from j to i.

In H: By replacements of the arrow and arcs in step 3 of the algorithm, only sections become larger and inner nodes remain anteriors of an endpoint. If an endpoint of the arrow or arc is i or j then an endpoint section of the generated walk is not a single element and there is a node h such that $h \in \mathrm{ant}(i) \cap \mathrm{sp}(i)$ or $h \in \mathrm{ant}(j) \cap \mathrm{sp}(j)$ respectively; otherwise the endpoint sections are single elements. In the former case, we add $\langle i, h, i\rangle$ to the walk; and similarly for j.

($\Leftarrow$) Suppose that there is a subprimitive inducing walk π from j to i in H. Consider the trislide ρ containing i. First suppose that the endpoints of ρ are a single element i (i.e. $\rho = \langle i, l, i\rangle$, where $l \in \mathrm{ant}(i)$). Consider the path $\langle k, \rho'\rangle$, where i is an endpoint of the section ρ' adjacent to ρ and there is an arc between k and the other endpoint of ρ' (or possibly an arrow if $k = j$). By step 3 of Algorithm 3, we can replace this path by an arc (or an arrow).

By step 4 of the algorithm we obtain an arc instead of this trislide. By considering the trislide containing i after the replacement, we have that inner nodes of the trislide are in $\mathrm{ant}(i)$. By repeating this argument we obtain an ij edge.

We now prove the second claim: If $j \in \mathrm{ant}(i)$ in H then, by step 5 of the algorithm, there is no arrowhead at j on the ij edge in $\alpha_{CMG.AnG}(H)$. If $j \notin \mathrm{ant}(i)$ in H then, by Lemma 9, $j \notin \mathrm{ant}(i)$ after applying step 4 of the algorithm. Hence, step 5 is not applicable. The result then follows from the fact that steps 3 and 4 generate endpoint-identical edges. □

Proof of Lemma 11. By Lemma 10, it is enough to prove that (1) there is a subprimitive inducing walk from i to j in $\alpha_{CMG}(\alpha_{CMG.AnG}(H); M, C)$ with single-element endpoint sections if and only if there is an endpoint-identical walk of the same type from i to j in $\alpha_{CMG}(H; M, C)$; (2) $j \in \mathrm{ant}(i)$ in $\alpha_{CMG}(\alpha_{CMG.AnG}(H); M, C)$ if and only if $j \in \mathrm{ant}(i)$ in $\alpha_{CMG}(H; M, C)$.

Proving (1): By Lemma 8, every edge on the subprimitive inducing walk π from i to j in $\alpha_{CMG}(H; M, C)$ can be replaced by the described walk in the lemma. Denote the new walk by π' in H. Notice that if a replaced subwalk is not endpoint-identical to the original edge then an endpoint k of the edge should be in $\mathrm{ant}(C)$ in H, which means that k is on a non-collider inner section on π (or is an endpoint with no arrowheads pointing to it), but this is impossible. Therefore, all such edge-replacements are endpoint identical. In addition, by Lemma 14, if a node h is in $\mathrm{ant}(j)$ in $\alpha_{CMG}(H; M, C)$ then $h \in \mathrm{ant}(C \cup \{j\})$ in H.

These imply that there is a subprimitive inducing walk from i to j with the mentioned properties in $\alpha_{CMG}(H; M, C)$ if and only if in H there is a walk between i and j on which (i) all nodes on collider sections are in $C \cup \mathrm{ant}(C) \cup \{j\}$; (ii) (a) all nodes on non-collider sections are in M, or (b) on non-collider sections, one endpoint is in M and also either a child of a node in M or a spouse of a node in $C \cup \mathrm{ant}(C)$, and the other endpoint has an arrowhead at it from the adjacent node on the walk. In addition, the two walks are endpoint-identical except when there is an arrowhead at the endpoint section containing i (or j), and $i \in \mathrm{ant}(C)$ (or $j \in \mathrm{ant}(C)$) in H.

Now by using Lemma 9, we have that $i \in \mathrm{ant}(C)$ in H if and only if $i \in \mathrm{ant}(C)$ in $\alpha_{CMG.AnG}(H)$. Therefore, since the same statements as above hold also for $\alpha_{CMG}(\alpha_{CMG.AnG}(H); M, C)$ and $\alpha_{CMG.AnG}(H)$, and in order to complete the proof, we need to show that there is a walk between i and j in H with the two mentioned properties if and only if there is an endpoint-identical walk π_0 of the same type between i and j in $\alpha_{CMG.AnG}(H)$:

To prove this, it is enough to show that by placing the walks described in Lemma 10 in place of the edges of π_0, the form of π_0 does not change: Without loss of generality, suppose that π_0 is a shortest walk of the described form, and an rs edge on π_0 has been replaced by a subprimitive inducing walk ϖ from r to s. The newly added sections are all collider. Because of transitivity of anteriors, and since the inner nodes of ϖ are anteriors of s, they stay is $\mathrm{ant}(C \cup \{j\})$. It is now enough to only check the sections containing r and s on π_0. Firstly, it is easy to see by Lemma 10 that the type of these sections do not change regardless of whether they are single elements on ϖ.

Secondly, if the rs edge and ϖ are endpoint-identical then theses sections remain of the same type. This completes the proof by using Lemma 9.

If these are not endpoint-identical then $s \in \mathrm{ant}(r)$. A problem only may arise when the section containing s is a non-collider in $\alpha_{CMG.AnG}(H)$ but a collider in H. If, for contradiction, this is the case then there is an arrow to s from the other adjacent node q to s on π_0. In addition, since all inner nodes of ϖ are anteriors of s, they are anteriors of r, and hence in H, $\langle \varpi, q \rangle$ is a subprimitive inducing walk from q to r, and hence π_0 is not a shortest walk, a contradiction. This completes the proof of this section.

Proving (2): Consider a semi-directed walk π in $\alpha_{CMG}(\alpha_{CMG.AnG}(H); M, C)$ from j to i. Since every edge is a subprimitive inducing walk, lines on π remain the same, and instead of an arrow from k to l on π we may have a subprimitive inducing walk from k to l. It is easy to observe that $k \in \mathrm{ant}(l)$, and by an inductive argument, we obtain the result.

The proof of other direction uses exactly the same argument (although, in fact, edges remain edges in this case). ☐

PROOF OF PROPOSITION 9. **First we prove that every CG G is mapped into $\mathcal{K}$:** By Proposition 8, we know that α_{AnG} maps CGs into $\mathcal{ANG}$. By Proposition 7, we know that after applying steps 1 and 2 of Algorithm 3, a CG G is mapped into $\mathcal{H}$, defined in Proposition 2. We need to prove that after applying steps 3, 4, and 5 of Algorithm 3, a CMG $H \in \mathcal{H}$ is mapped into $\mathcal{K}$.

Suppose that there is a trislide $\pi = k \leftrightarrow i \,—\, \dots \,—\, j \leftarrow l$ **in the generated graph**: By Lemma 10, there is a subprimitive inducing walk from l to j in H. Denote the node on this walk adjacent to j by q. The jq edge is an arc unless $l = q$, in which case it is an arrow from q to j. Since lines are not generated by Algorithm 3, and since $H \in \mathcal{H}$, there is an iq arc or an arrow from l to i.

In the generated graph, $j \in \mathrm{ant}(i)$, and there is a subprimitive inducing walk from l to i that goes through the subprimitive inducing walk from l to j, the section from j to i, the iq edge, the jq edge, and again the section between j and i. Hence, again by Lemma 10, there is an edge between l and i. This edge can only be an arrow from l to i since otherwise there is a semi-directed cycle or an arc with one endpoint that is an anterior of the other endpoint in the generated anterial graph.

Suppose that there is a trislide $\pi = k \leftrightarrow i \,—\, \dots \,—\, j \leftrightarrow l$ **in the generated graph:** It holds that $l \notin \mathrm{ant}(i)$ since otherwise $l \in \mathrm{ant}(j)$, which is impossible due to the existence of an arrowhead at l. This fact together with the same argument as that in the previous paragraphs implies that there is an il arc in the generated graph. By the symmetry on the trislide we also conclude that there is a jk arc in the generated graph. In addition, by what we proved in the previous paragraphs, there is a tripath $q' \leftrightarrow i \,—\, \dots \,—\, j \leftrightarrow q$ in H, which implies that there is an ij arc in H. This arc turns into a line by step 5 since i and j are anteriors of one another.

We now prove that the function is surjective: Consider an arbitrary graph $K \in \mathcal{K}$. We prove that there exists an $H \in \mathcal{H}$ such that $\alpha_{CMG.AnG}(H) = K$, i.e. by applying steps 3, 4, and 5 of Algorithm 3 to H, we obtain K. This completes the proof since α_{CMG} is surjective onto $\mathcal{H}$, and $\alpha_{AnG} = \alpha_{CMG.AnG} \circ \alpha_{CMG}$.

If K does not contain a trislide of form $\pi = k \leftrightarrow i \,—\, \dots \,—\, j \leftrightarrow l$ then $K \in \mathcal{H}$, and we simply let $H = K$. Since α_{AnG} does not change anterial graphs, we are done.

If K does contain a trislide π of the mentioned form then there is the ij line in K. Now let H be K, but with an arc between i and j instead of the existing line. We have that $H \in \mathcal{H}$. Denote also the section between i and j by ρ.

By Lemma 10, the ij arc turns into a line and clearly no other edge changes its type in $\alpha_{CMG.AnG}(H)$. Hence, it is enough to show that no other edge is generated. If the ij arc is part of any subprimitive inducing walk except when i or j is an endpoint then it can be replaced by ρ to obtain another primitive inducing walk. If i or j is an endpoint then, by how H is constructed, the possible arrows or lines that can be generated already exist in H. This completes the proof. $\square$

PROOF OF THEOREM 8. By Theorem 6, it is enough to prove that $A \perp_c B \mid C_1$ in $\alpha_{AnG}(G; M, C)$ if and only if $A \perp_c B \mid C_1$ in $\alpha_{CMG}(G; M, C)$.

Since Steps 1 and 2 of Algorithm 3 generate $\alpha_{CMG}(G; M, C)$, we need to prove that there is a c-connecting walk in a chain mixed graph H if and only if there is a c-connecting walk after applying steps 3, 4, and 5 of the algorithm to H.

(⇒) Suppose that there is a c-connecting walk π given C_1 between i and j in H. After applying steps 3 and 4, π is intact. If an arc kl is replaced by an arrow from k to l or a kl line, in step 5 of the algorithm then we have the two following cases:

1) If k is on a non-collider section on π by using the kl arrow or line instead of arc, one obtains a c-connecting walk.

2) Suppose that k is an endpoint of a collider section ρ and there is $\pi_1 = \langle h, \rho, l\rangle$ on π. By Lemma 1, one can assume that ρ is a path. By Lemma 9, $k \in \mathrm{ant}(l)$. If $h \neq l$ then by step 4, there is an endpoint-identical hl edge to π_1. One can now use the hl edge instead of π_1 to obtain a c-connecting walk. If $h = l$ then ρ can be considered to be the single node k. Now if h is on a non-collider section then we can easily skip k to obtain a c-connecting path. If h is an endpoint of a collider section ρ' then from $\pi_2 = \langle q, \rho', k\rangle$ and by using step 3 of the algorithm, we obtain an endpoint-identical qh edge, which can be replaced by π_2 to obtain a c-connecting path. This, by an inductive argument, implies the result.

(⇐) Suppose that there is a c-connecting walk π given C_1 between i and j in $\alpha_{CMG.AnG}(H)$, which is graph H after applying steps 3, 4, and 5 of Algorithm 3.

For every edge on π, by Lemma 10, there exists a subprimitive inducing walk in H between the same endpoints. We replace all the edges on π by these walks and call the generated walk π'. Notice that it can be shown that regardless of the choice of C, a subprimitive inducing walk is c-connecting itself. Hence, if the replaced subwalk of π' by an edge is endpoint-identical to the original edge then it does not affect the c-connectivity of π'. We, therefore, need to check the case where the generated walk is not endpoint-identical to the edge.

Suppose that this is the case for the edge ij in $\alpha_{CMG.AnG}(H)$ replaced by a subprimitive inducing walk ϖ from j to i. By the lemma, we have that either $j \in \text{ant}(i)$ or $i \in \text{ant}(j)$ in H, in which cases there is no arrowhead at j or i on the ij edge respectively.

Assume that $j \in \text{ant}(i)$. We need to consider the case where ij is an arrow from j to i, and j is not in C, but there is an arrowhead at j on ϖ. Denote the semi-directed walk from j to i by τ. If no node on τ is in C then we replace ϖ by τ to obtain a c-connecting walk. Otherwise, consider the closest node $k \in C$ on τ to j. The walk consisting of the subwalk of τ from j to k, the same subwalk in the reverse direction (from k to j), and ϖ is now c-connecting since j is on non-collider sections, except when j and k are on the same subsection of τ (which is still fine).

The case where $i \in \text{ant}(j)$ follows the exact same argument. □

References.

[1] Sadeghi, K. (2015). Marginalization and conditioning for LWF chain graphs. *submitted*.

Statistical Laboratory
Centre for Mathematical Studies
Wilberforce Road
Cambridge, CB3 0WB
United Kingdom.
E-mail: k.sadeghi@statslab.cam.ac.uk