



Separation and Completeness Properties for AMP Chain Graph Markov Models

Author(s): Michael Levitz, Michael D. Perlman and David Madigan

Source: *The Annals of Statistics*, Dec., 2001, Vol. 29, No. 6 (Dec., 2001), pp. 1751-1784

Published by: Institute of Mathematical Statistics

Stable URL: <https://www.jstor.org/stable/2699950>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve and extend access to *The Annals of Statistics*

SEPARATION AND COMPLETENESS PROPERTIES FOR AMP CHAIN GRAPH MARKOV MODELS¹

BY MICHAEL LEVITZ, MICHAEL D. PERLMAN AND DAVID MADIGAN

University of Washington

Pearl's well-known d -separation criterion for an acyclic directed graph (ADG) is a pathwise separation criterion that can be used to efficiently identify all valid conditional independence relations in the Markov model determined by the graph. This paper introduces p -separation, a pathwise separation criterion that efficiently identifies all valid conditional independences under the Andersson–Madigan–Perlman (AMP) alternative Markov property for chain graphs (= adicyclic graphs), which include both ADGs and undirected graphs as special cases. The equivalence of p -separation to the augmentation criterion occurring in the AMP global Markov property is established, and p -separation is applied to prove completeness of the global Markov property for AMP chain graph models. Strong completeness of the AMP Markov property is established, that is, the existence of Markov perfect distributions that satisfy those and only those conditional independences implied by the AMP property (equivalently, by p -separation). A linear-time algorithm for determining p -separation is presented.

1. Introduction. Over the past eight decades, graphical Markov models based on undirected graphs (UG) and acyclic directed graphs (ADG, also DAG), also called Bayesian networks, have been used by computer scientists, econometricians, geneticists, statisticians and many others to represent multivariate dependences among statistical variables in a parsimonious and readily interpretable manner [cf. Pearl (1988), Whittaker (1990), Cowell, David, Lauritzen and Spiegelhalter (1999)]. Chain graphs (= adicyclic graphs), which may have both undirected and directed edges, were introduced by Lauritzen, Wermuth and Frydenberg [LWF] to combine the properties of UGs and ADGs, representing dependences which may be both associative and directional [cf. Lauritzen and Wermuth (1989), Frydenberg (1990a), Wermuth and Lauritzen (1990), Whittaker (1990), Buntine (1995), Bouckaert and Studený (1995), Højsgaard and Thiesen (1995), Cox and Wermuth (1996), Lauritzen (1996), Andersson, Madigan, and Perlman [AMP] (1997a), Studený (1998), Studený and Bouckaert (1998)]. D. R. Cox (1999) states that chain graphs represent “a minimal level of complexity needed to model empirical data.”

Initially, expert system builders focussed on ADG models because of the ease of their interpretation and analysis [cf. Spiegelhalter, Dawid, Lauritzen and Cowell (1993), Heckerman, Geiger and Chickering (1995)]. Adopting the larger class of chain graph (CG) models can be advantageous, however, not

Received June 2000; revised May 2001.

¹Supported in part by the NIH and the NSF.

AMS 2000 subject classifications. Primary 62M45, 60K99; secondary 68R10, 68T30.

Key words and phrases. Graphical Markov model, acyclic directed graph, Bayesian network, d -separation, chain graph, AMP model, p -separation, completeness, efficient algorithm.

only because of the greater modeling flexibility they provide but also because each Markov equivalence class of ADGs can be uniquely represented by a certain Markov-equivalent CG, the *essential graph* ([AMP] (1997b)). Such CGs can therefore be used to avoid the practical difficulties associated with managing equivalent ADG models [cf. Heckerman, Geiger and Chickering (1995), Madigan, Andersson, Perlman and Volinsky (1996)].

The LWF Markov property for CGs is an extension of the Markov properties of both ADGs and UGs. Recently, [AMP] (1996, 2001) have proposed an alternative Markov property for CGs that also extends the ADG and UG properties but that in some ways more closely retains the recursive character of ADG models. For example, unlike the LWF property, the AMP property is satisfied by a block-recursive normal linear system naturally associated with the graph. In this case the AMP Markov property, like the ADG Markov property, corresponds to the assumption that certain regression coefficients are zero, directly indicating an absence of (conditional) dependence between the two variables; this is not the case for the LWF property; compare [AMP] (2001).

Furthermore, AMP Markov equivalence of CGs, as for ADGs, is determined by their *triplexes*, which contain three vertices, whereas LWF Markov equivalence of CGs is determined by their *complexes*, which can contain arbitrarily many vertices [Frydenberg (1990a), [AMP] (1996, 1997a, 2001)].

Pearl's (1988) *d-separation* criterion is the standard method for identifying all valid conditional independences (CI) in the Markov model associated with an ADG D . Lauritzen, Dawid, Larsen and Leimer (1990) established the equivalence of *d-separation* and the *moralization* criterion occurring in the global Markov property for ADGs. Bouckaert and Studený (1995), Studený (1996, 1997, 1998), and Studený and Bouckaert (1996) introduced *c-separation*, a more complex graphical criterion for identifying all valid CIs under the LWF Markov property for a general CG G .

In this paper we introduce *p-separation*, a simpler graphical separation criterion that identifies all valid CIs under the AMP Markov property for G . Like *d-separation* but unlike *c-separation*, *p-separation* requires consideration only of non-self-intersecting routes, here called "trails" (see the Appendix), and the active/blocked status of a trail is determined only by its individual vertices, not by all of its subtrails). Thus, algorithmic implementation for *p-separation* is somewhat simpler than for *c-separation*.

As noted above, the Markov equivalence class $[D]$ for an ADG D is uniquely represented by its essential graph D^* , a CG having a special form characterized in Theorem 4.1 of [AMP] (1997b). By Theorem 4.3 of [AMP] (2001), the AMP and LWF global Markov properties coincide for D^* , so the simpler *p-separation* criterion, rather than the *c-separation* criterion, can be applied to D^* in order to determine the valid CIs for all ADGs that are Markov equivalent to D .

The UG and ADG global Markov properties, the *d-separation* criterion for ADGs, and the LWF and AMP global Markov properties for CGs are reviewed in Section 2. In Section 3, the *p-separation* criterion for CGs is introduced

in a form that shows its similarity to d -separation. Theorem 4.1 establishes the equivalence of the p -separation criterion and the augmentation criterion occurring in the AMP global Markov property for CGs.

Theorem 5.1 establishes the *completeness* of the AMP global Markov property for a general CG G . Completeness asserts that the CIs specified by the AMP global Markov property for G (equivalently, by the p -separation criterion applied to G) are the *only* CIs that are simultaneously satisfied by all distributions in the AMP Markov model determined by G . Our proof extends the elegant construction of Geiger and Pearl (1988) for completeness of the ADG global Markov property using d -separation. Strong completeness of the AMP global Markov property is established in Section 6. It is shown that in the Gaussian case, almost all AMP G -Markovian distributions are *Markov perfect* \equiv *faithful* for G , that is, almost every such distribution satisfies exactly those CIs specified by the AMP property. These results insure that the CG G is a valid mathematical object for representing the independence/dependence structure of the AMP Markov model that it defines. A linear-time algorithm for determining p -separation and all valid CIs entailed by an AMP model is presented in Section 7.

In the Appendix we briefly review the graph-theoretic terminology given more fully in, for example, Cowell, Dawid, Lauritzen and Spiegelhalter (1999) or [AMP] (2001). Some additional terminology and notation will be introduced in the present paper as needed.

2. UG, ADG, LWF and AMP graphical Markov models. In the remainder of this paper, unless otherwise specified, $G \equiv (V, E)$ shall denote a chain graph (CG) with finite vertex (= node) set $V \equiv V(G)$ and edge set $E \equiv E(G)$. We shall consider multivariate probability distributions P on a product probability space $\mathbf{X} \equiv \times(\mathbf{X}_v | v \in V)$, where each \mathbf{X}_v is sufficiently regular to ensure the existence of regular conditional probabilities. A distribution $P \in \mathcal{P}$ will be represented by a random vector $X \equiv (X_v | v \in V) \in \mathbf{X}$. For $A \subseteq V$, P_A is the marginal probability distribution represented by $X \equiv (X_v | v \in A)$ and is defined as a measure on $\mathbf{X}_A \equiv \times(\mathbf{X}_v | v \in A)$.

Unless otherwise specified, A, B, S will denote three mutually disjoint subsets of V such that $A, B \neq \emptyset$. In this case, the conditional independence relation $X_A \perp\!\!\!\perp X_B | X_S[P]$ often will be abbreviated as $A \perp\!\!\!\perp B | S[P]$.

First we recall the global Markov properties for UG and ADG models, then review Pearl's d -separation criterion, the standard graphical pathwise criterion for identifying valid CIs in ADG models (\equiv Bayesian networks); cf. Lauritzen, Dawid, Larsen and Leimer (1990), Lauritzen (1996). An ADG is denoted by $D \equiv (V, E)$.

If $G \equiv (V, E)$ is a UG, we say that S separates A and B in G , denoted by $A \bowtie B | S[G]$, if every path between A and B in G necessarily intersects S . In particular, $A \bowtie B | \emptyset[G]$ iff there exist no paths between A and B in G .

DEFINITION 2.1 (The global Markov property for UGs). Let $G \equiv (V, E)$ be an undirected graph. A probability measure P on \mathbf{X} is *global G -Markovian* if $A \perp\!\!\!\perp B \mid S [P]$ whenever $A \bowtie B \mid S [G]$.

DEFINITION 2.2 (The global Markov property for ADGs). Let $D \equiv (V, E)$ be an acyclic directed graph. A probability measure P on \mathbf{X} is *global D -Markovian* if $A \perp\!\!\!\perp B \mid S [P]$ whenever

(2.1)
$$A \bowtie B \mid S [(D_{\text{An}(ABS)})^m],$$

where $(D_{\text{An}(ABS)})^m$ is the moralized ancestral graph for $ABS \equiv A \dot{\cup} B \dot{\cup} S$.

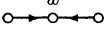
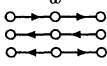
Condition (2.1) is called the *ADG moralization criterion*. It is not a separation criterion in D itself, because the ancestral graph appearing in (2.1) varies with A, B, S . For fixed A and S , Geiger, Verma, and Pearl (1990) show that algorithms for finding $B \subseteq D \setminus \{A \dot{\cup} S\}$ such that $A \bowtie B \mid S [(D_{\text{An}(ABS)})^m]$ have complexity $O(|V|^3)$ if based on the moralization criterion, but have complexity $O(|E|) = O(|V|^2)$ if based on the d -separation criterion (see below). Thus the gain in computational efficiency afforded by d -separation may be substantial, especially for sparse graphs.

Let $\pi \equiv (a \equiv v_0, \dots, v_n \equiv b)$ be a trail between A and B in D . An interior vertex v_i is a *head-to-head* vertex (= node) in π if $v_{i-1} \rightarrow v_i \leftarrow v_{i+1}$ occurs as a subgraph (not necessarily induced) of D . An interior vertex w is *active relative to S* ($\equiv S$ -active) in π if either:

- w is a head-to-head node in π and $w \in \text{An}_D(S)$, or
- w is not a head-to-head node in π and $w \notin S$.

If an interior vertex is not S -active in π , it is said to be *blocking relative to S* ($\equiv S$ -blocking) in π . It is convenient to refer to Table 1 to determine the active/blocking status of an interior vertex w .

TABLE 1
The d -separation criterion for an ADG D

$w \in \pi^\circ \equiv \pi \setminus \{a, b\}$	$w \in S$	$w \notin S$	
		$w \in \text{an}_D(S)$	$w \notin \text{an}_D(S)$
Either w is a head-to-head node in π : 	w is ACTIVE	ACTIVE	BLOCKING
Or w is not a head-to-head node in π : 	w is BLOCKING	ACTIVE	ACTIVE

The trail π is said to be *blocked relative to S* ($\equiv S$ -*blocked*) if it contains at least one S -blocked interior vertex; otherwise π is *active relative to S* ($\equiv S$ -*active*).

DEFINITION 2.3 (Pearl's d -separation criterion). S d -separates A and B in the ADG D , denoted by $A \bowtie_d B \mid S \ [D]$, if every trail π between A and B in D is S -blocked in D .

THEOREM 2.1 [Lauritzen, Dawid, Larsen and Leimer (1990)]. S d -separates A and B in D if and only if S separates A and B in $(D_{\text{An}(ABS)})^m$. That is,

$$A \bowtie_d B \mid S \ [D] \iff A \bowtie B \mid S \ [(D_{\text{An}(ABS)})^m].$$

REMARK 2.1. As a special case of results of Studeny (1998) and Koster (1999), we note that the d -separation criterion can be stated in a simpler form if the definition of trail is extended to allow self-intersection. A possibly self-intersecting (p.s.i.) trail π between A and B in D is called S -open if:

- Every head-to-head node in π lies in S , and
- Every other node in π lies in $V \setminus S$.

(Note that $\text{An}_D(S)$ does not appear here.) Then it can be shown that d -separation is equivalent to the condition that no S -open p.s.i. trail exists between A and B in D . Of course, the disadvantage of this formulation of d -separation is that there are infinitely many p.s.i. trails in D , compared to only finitely many trails in the original sense.

Next we review the LWF and AMP Markov properties for CGs. Frydenberg (1990a) defined the LWF global Markov property for CGs as follows [cf. Lauritzen and Wermuth (1989), Lauritzen (1996), [AMP] (2001)].

DEFINITION 2.4 (The LWF global Markov property for CGs). Let $G \equiv (V, E)$ be a chain graph. A probability measure P on \mathbf{X} is *LWF global G -Markovian* if $A \perp\!\!\!\perp B \mid S \ [P]$ whenever

$$(2.2) \quad A \bowtie B \mid S \ [(G_{\text{At}(ABS)})^m],$$

where $(G_{\text{At}(ABS)})^m$ is the moralized anterior graph for $ABS \equiv A \cup B \cup S$.

Condition (2.2) is called the *LWF moralization criterion* for CGs (again, it is not a separation criterion in G itself). To avoid this complication, Studeny and Bouckaert (1998) developed the *c-separation criterion*, which we denote by \bowtie_c , and established its equivalence to the LWF global Markov property,

$$A \bowtie_c B \mid S \ [G] \iff A \bowtie B \mid S \ [(G_{\text{At}(ABS)})^m].$$

We shall not discuss c -separation here, except to say that although c -separation for CGs is an extension of d -separation for ADGs, it loses the

pathwise nature of d -separation, instead involves possibly self-intersecting trails and blocking by subtrails rather than by single vertices only [cf. Studený and Bouckaert (1998), Studený (1998), Section 5].

Andersson, Madigan and Perlman [AMP] (1996, 2001) defined an alternative Markov property (AMP) for CGs. They first defined this via a natural block-recursive Markov property for CGs, then showed this to be equivalent to the following global condition.

DEFINITION 2.5 (The AMP global Markov property for CGs). Let $G \equiv (V, E)$ be a chain graph. A probability measure P on \mathbf{X} is *AMP global G -Markovian* if $A \perp\!\!\!\perp B \mid S \mid P$ whenever

$$(2.3) \quad A \bowtie B \mid S \mid (G[ABS])^a,$$

where $(G[ABS])^a$ is the augmented extended subgraph for $ABS \equiv A \cup B \cup S$. The set of all *AMP global G -Markovian* P on \mathbf{X} is denoted by $\mathcal{P}_{\text{AMP}}^g(G; X)$.

Condition (2.3) is called the *AMP augmentation criterion* (reviewed below). Like the two moralization criteria, it is not a separation criterion in G itself. In the next section we shall introduce *p-separation*, an equivalent pathwise separation criterion in G itself that requires only non-self-intersecting trails and blocking by single vertices only, hence which offers computational efficiency in determining valid CIs in an AMP CG model.

For the convenience of the reader, we conclude this section by briefly reviewing the definition of augmentation for CGs and related graphical terminology [see [AMP] (2001, Section 2) for further details].

Let u, v, w be distinct vertices of G . An *immorality* $(u, v; w)$ in G is an induced subgraph of the form $u \rightarrow w \leftarrow v$ [Figure 1(a)]. A *flag* $[u, v; w]$ is an induced subgraph of the form $u \rightarrow w - v$ [Figure 1(b)]. A *triplex* is an ordered pair $(\{u, v\}, w)$ such that *either* $(u, v; w)$ is an immorality *or else* $[u, v; w]$ or $[v, u; w]$ is a flag. Thus, the triplex $(\{u, v\}, w)$ occurs in G iff one of the three graphs indicated in Figures 1a or 1b occurs as an induced subgraph of G .

Let u, v, w', w'' be distinct vertices of G . A *2-biflag* $[u, v; w', w'']$ in G is an induced subgraph $G_{\{u, v, w', w''\}}$ such that $[u, w''; w']$ and $[v, w'; w'']$ are flags in G . The four possible forms of the 2-biflag $[u, v; w', w'']$ are indicated in Figure 1c, where the “?” indicates that either $u - v \in G$, $u \rightarrow v \in G$, $\leftarrow v \in G$, or $u \cdot v$ in G .

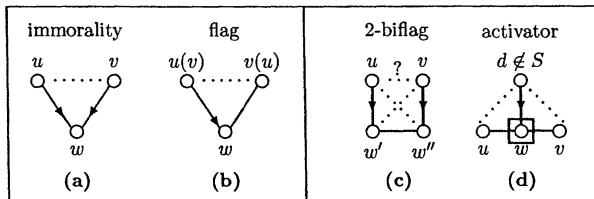


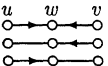
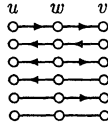

FIG. 1. Augmentation in an AMP chain graph.

The *augmented* triplex $(\{u, v\}, w)^a$ is the complete UG with vertices u, v, w , that is, an undirected triangle. The *augmented* 2-biflag $[u, v; w', w'']^a$ is the complete UG with vertices u, v, w', w'' , that is, a complete undirected square. The lines added in these augmentations are indicated by dotted lines in Figure 1. (The line between u and v in Figure 1(c) may already be present in G .) The *augmented graph* G^a derived from a CG G is the UG obtained by augmenting all triplexes and 2-biflags in G , then converting all remaining arrows of G into lines.

3. The p -separation criterion for AMP chain graphs. Let $\pi \equiv (a \equiv v_0, \dots, v_n \equiv b)$ be a trail between A and B in a chain graph $G \equiv (V, E)$, where, as before, A, B, S are mutually disjoint subsets of V with $A, B \neq \emptyset$. An interior vertex v_i is a *head-no-tail* vertex (= node) in π if either $v_{i-1} \rightarrow v_i \leftarrow v_{i+1} \subseteq G$, $v_{i-1} \rightarrow v_i - v_{i+1} \subseteq G$, or $v_{i-1} - v_i \leftarrow v_{i+1} \subseteq G$, that is, if either occurs as a subgraph (not necessarily induced) of G . Each interior vertex v is either *active relative to S* ($\equiv S$ -active) in π , or *blocking relative to S* ($\equiv S$ -blocking) in π , according to Table 2.

In Table 2, u, w, v denote consecutive vertices in π . In the 4-node subgraph appearing in Table 2, a box is placed around the node of w to emphasize that $w \in S$; unboxed vertices may either belong to S or not, unless specifically indicated. The vertex $d \in \text{pa}_G(w) \setminus S$ in this subgraph is called an *S -activator of w for π* , or, simply, an *activator*, because w would be S -blocking in π if no such d existed. (The possibility that $d \in \pi$ is allowed.) Note that for a

TABLE 2
The p -separation criterion for a chain graph G

$w \in \pi^\circ \equiv \pi \setminus \{a, b\}$	$w \in S$	$w \notin S$	
		$w \in \text{an}_G(S)$	$w \notin \text{an}_G(S)$
Either w is a head-no-tail node in π :  or w is not a head-no-tail node in π : 	<div>w is ACTIVE</div> <div>w is BLOCKING</div> <div> unless $\exists d \in \text{pa}_G(w) \setminus S$ such that  occurs as a subgraph of G, whence w is ACTIVE </div>	ACTIVE	BLOCKING
		ACTIVE	ACTIVE

fixed trail π , such a w may have more than one S -activator, and a given d can activate (i.e., serve as an S -activator of) more than one w . Each of the two dashed arrows in the subgraph may be either present or absent. (Also see Figure 1(d).)

The trail π is *blocked relative to $S \equiv S$ -blocked*, if it contains at least one S -blocking interior vertex; otherwise, π is *active relative to $S \equiv S$ -active*.

DEFINITION 3.1 (The p -separation criterion). S p -separates A and B in the CG G , denoted by $A \bowtie_p B \mid S [G]$, if every trail π between A and B in G is S -blocked in G .

Clearly p -separation reduces to ordinary pathwise graphical separation when G is a UG, while, by comparing Tables 1 and 2, it is seen that p -separation reduces to d -separation when G is an ADG.

REMARK 3.1. As in Remark 2.1, the p -separation criterion can be stated in a simpler form if the definition of trail is extended to allow self-intersection. A possibly self-intersecting (p.s.i.) trail π between A and B in G is called S -open if:

- Every head-no-tail node in π lies in S , and
- Every other node in π lies in $V \setminus S$.

Then it can be shown that p -separation is equivalent to the requirement that no S -open p.s.i. trail exists between A and B in G . Again, this formulation of p -separation has the disadvantage that there are infinitely many p.s.i. trails in G . (We thank Michael Eichler for this observation.)

4. Equivalence of p -separation and the AMP global Markov property.

Our first main result establishes the equivalence of the p -separation and augmentation criteria for AMP chain graph (CG) models.

THEOREM 4.1. Let $G \equiv (V, E)$ be a chain graph and let $A, B, S \subseteq V$ be mutually disjoint with $A, B \neq \emptyset$. Then S p -separates A and B in G if and only if S separates A and B in $(G[ABS])^a$. That is,

$$A \bowtie_p B \mid S [G] \iff A \bowtie B \mid S [(G[ABS])^a].$$

Lemma 4.1 is needed for the proofs of Theorems 4.1 and 5.1. Suppose that $\pi \equiv (v_0, \dots, v_n)$ ($n \geq 1$) is an S -active trail between A and B in G . The set S_π of S -activated vertices in π is defined as follows:

$$(4.1) \quad S_\pi := \{v_i \in \pi^\circ \cap S \mid v_{i-1} - v_i - v_{i+1} \sqsubseteq G\}.$$

That is, S_π is the set (possibly empty) of all interior vertices of π that lie in S and are connected to their predecessor and successor in π by lines in G .

We shall consider an S -active trail $\pi \equiv (v_0, \dots, v_n)$ between A and B in G that satisfies the following minimum cardinality condition:

(M1) π minimizes $|S_\pi|$ over all S -active trails π' between A and B in G .

If $m := |S_\pi| > 0$, denote the members of S_π in their order of occurrence in π by s_1, \dots, s_m and let k_i be that index such that $v_{k_i} = s_i$. Thus $1 \leq k_1 < \dots < k_m \leq n-1$ and

$$(4.2) \quad S_\pi = \{s_1, \dots, s_m\} = \{v_{k_1}, \dots, v_{k_m}\}.$$

Because π is S -active, each $s_i \in S_\pi$ is S -active, so arbitrarily specify one S -activator $d_i \in \text{pa}_G(s_i) \setminus S$ and define

$$(4.3) \quad D_\pi := \{d_1, \dots, d_m\}.$$

LEMMA 4.1. *If π satisfies M1, then:*

- (a) $\pi \cap D_\pi = \emptyset$;
- (b) d_1, \dots, d_m are mutually distinct, that is, each $d_i \in D_\pi$ activates only s_i .

PROOF. (a) Suppose to the contrary that $d_i = v_j$ for some $1 \leq i \leq m$ and $0 \leq j \leq n$. Because $v_j \rightarrow v_{k_i} \in G$ and $v_{k_i-1} - v_{k_i} - v_{k_i+1} \subseteq G$, necessarily $|k_i - j| \geq 2$. If $j < k_i$, then

$$\pi' := (v_0, \dots, v_j, v_{k_i}, v_{k_i+1}, \dots, v_n)$$

is a trail between A and B in G that is again S -active (see Table 2). However, $v_{k_i} \equiv s_i$ is not an S -activated vertex in π' , so $|S_{\pi'}| \leq |S_\pi| - 1$, which contradicts M1. A similar argument applies if $k_i < j$. Therefore $\pi \cap D_\pi = \emptyset$.

(b) Suppose to the contrary that $d_i = d_j =: d (\notin \pi)$ for some $1 \leq i < j \leq m$. Because $v_{k_i} \leftarrow d \rightarrow v_{k_j} \subseteq G$,

$$\pi' := (v_0, \dots, v_{k_i-1}, v_{k_i}, d, v_{k_j}, v_{k_j+1}, \dots, v_n)$$

is a trail between A and B in G that is again S -active (see Table 2). However, $v_{k_i} \equiv s_i$, d , and $v_{k_j} \equiv s_j$ are not S -activated in π' , so $|S_{\pi'}| \leq |S_\pi| - 2$, again contradicting M1. Thus, each $d_i \in D_\pi$ activates only $s_i \in \pi$, so d_1, \dots, d_m are mutually distinct. \square

PROOF OF THEOREM 4.1 ("if"). Suppose that $\pi \equiv (a \equiv v_0, \dots, v_n \equiv b)$ is an S -active trail between A and B in G that satisfies M1. By Lemma 4.1, the S -activators in $D_\pi \equiv \{d_i \mid i = 1, \dots, m\}$ are distinct. We shall show that π can be modified to obtain a path $\tilde{\pi}$ between A and B in the UG H^a that bypasses S , where $H := G[ABS]$.

CLAIM 4.1. $\pi \dot{\cup} D_\pi \subseteq \text{Co}(\text{An}(ABS))$.

PROOF. Clearly $\{v_0, v_n\} \subseteq AB$. Next consider $v_i \in \pi^\circ$. If v_i is a head-no-tail node in π , then $v_i \in \text{An}(S)$ since v_i is S -active. If v_i is not a head-no-tail node, at least one of the following edges must occur in G : (a) $v_i \rightarrow v_{i+1}$, (b) $v_i - v_{i+1}$, (c) $v_{i-1} \leftarrow v_i$ or (d) $v_{i-1} - v_i$. If (a), then the subtrail (v_i, \dots, v_n) is either a directed path from v_i to $v_n \equiv b$, in which case $v_i \in \text{an}(B)$, or this subtrail begins at v_i as a directed outgoing path and first encounters a line

or opposing arrow at some $v_j \in \pi^\circ$ ($i < j < n$) before reaching b . In this case v_j is a head-no-tail node in π , hence $v_j \in \text{An}(S)$ since v_j is S -active, so $v_i \in \text{an}(S)$. If (b), then the subtrail (v_i, \dots, v'_n) is *either* an undirected path from v_i to $v_n \equiv b$, in which case $v_i \in \text{Co}(B)$, or this subtrail begins at v_i as an undirected path and first encounters an arrow $v_j \leftarrow v_{j+1}$ or $v_j \rightarrow v_{j+1}$ at some $v_j \in \pi^\circ$ ($i < j < n$). If $v_j \leftarrow v_{j+1}$, then v_j is an S -active head-no-tail node in π , hence $v_j \in \text{An}(S)$ so $v_i \in \text{Co}(\text{An}(S))$. If $v_j \rightarrow v_{j+1}$, then apply the argument in case (a) to the subtrail (v_j, \dots, v_n) to deduce that $v_j \in \text{an}(BS)$, whence $v_i \in \text{Co}(\text{An}(BS))$. Cases (c) and (d) are similar to (a) and (b), respectively, so in all four cases, $v_i \in \text{Co}(\text{An}(ABS))$, hence $\pi \subseteq \text{Co}(\text{An}(ABS))$. Clearly $D_\pi \subseteq \text{an}(S)$, so Claim 4.1 is established. \square

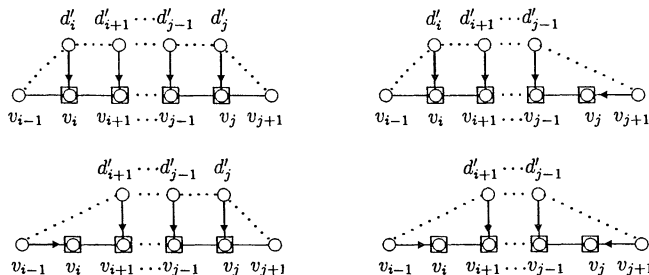
Because $\pi \equiv (v_0, v_1, \dots, v_n)$ is S -active, if $v_i \in \pi^\circ \cap S$ then either v_i is a head-no-tail node in π or else $v_i \in S_\pi$. Therefore, if $v_i, v_{i+1} \in \pi^\circ \cap S$, necessarily $v_i - v_{i+1} \in G$. We say that $v_i \in \pi^\circ \cap S$ is *S -isolated* if $v_{i-1}, v_{i+1} \notin S$. We say that the consecutive subsequence $(v_i, \dots, v_j) \subseteq \pi^\circ \cap S$ ($1 \leq i < j \leq n-1$) is a *maximal S -run* in π° if $v_{i-1}, v_{j+1} \notin S$; necessarily $v_{i+1}, \dots, v_{j-1} \in S_\pi$ if $j-i \geq 2$.

The vertices of the requisite S -bypassing path $\tilde{\pi}$ between A and B in H^a are obtained from π as follows:

1. Retain each $v_i \in \pi \setminus S$.
2. Remove each $v_i \in \pi^\circ \cap S$ that is head-no-tail in π .
3. Remove each $s_i \equiv v_{k_i} \in S_\pi$ and replace it by d_i .

Thus the vertices of $\tilde{\pi}$ consist exactly of those in $(\pi \setminus S) \cup D_\pi$. These vertices are now linked by lines as follows:

1. $v_i - v_{i+1}$ for each consecutive pair $v_i, v_{i+1} \in \pi \setminus S$;
2. $v_{i-1} - v_{i+1}$ for each S -isolated $v_i \in \pi^\circ \cap S$ that is head-no-tail in π ;
3. $v_{k_{i-1}} - d_i - v_{k_{i+1}}$ for each S -isolated $s_i \equiv v_{k_i} \in S_\pi$;
4. If $(v_i, v_{i+1}, \dots, v_j) \subseteq \pi^\circ \cap S$ is a maximal S -run in π° , then one of the following four configurations must occur as a (not necessarily induced) subgraph of G (boxed nodes are in S , unboxed nodes are not in S):



Each d'_h denotes the S -activator for the corresponding v_h . In each configuration, the dotted lines between unboxed nodes indicate where lines are to be

included in $\tilde{\pi}$. For example, if the first configuration occurs in G then the lines $v_{i-1} - d'_i - d'_{i+1} - \dots - d'_{j-1} - d'_j - v_{j+1}$ are included in $\tilde{\pi}$.

Note: If $j = i + 1$ then in the fourth configuration no S -activators d'_h occur and only the single line $v_{i-1} - v_{j+1}$ is to be included in $\tilde{\pi}$.

By Lemma 4.1, Claim 4.1, and the construction of $\tilde{\pi}$, it is clear that $\tilde{\pi}$ is an S -bypassing path between A and B in the complete undirected graph with vertex set $Co := Co(An(ABS))$. It remains to show that each edge in $\tilde{\pi}$, as given in (1), (2), (3), (4), actually occurs in H^a .

(1) Because $v_i, v_{i+1} \in \pi \setminus S$, necessarily $v_i \dots v_{i+1} \in G_{Co}$ by Claim 4.1. (a) If $v_i - v_{i+1} \in G_{Co}$ then $v_i - v_{i+1} \in G_{Co}^\wedge \subseteq H^a$. (b) Suppose that $v_i \rightarrow v_{i+1} \in G_{Co}$. Two possibilities arise: either $v_{i+1} \in An := An(ABS)$ or $v_{i+1} \in Co \setminus An$. If $v_{i+1} \in An$ then also $v_i \in An$, hence $v_i \rightarrow v_{i+1} \in G_{An} \subseteq H$, so $v_i - v_{i+1} \in H^a$. The second possibility cannot occur: if $v_{i+1} \in Co \setminus An$, then $v_{i+1} \notin B$, hence $v_{i+1} \in \pi^\circ$. Because $v_{i+1} \notin An(S)$ and is S -active in π , necessarily $v_{i+1} \rightarrow v_{i+2} \in G_{Co}$. Thus also $v_{i+2} \in Co \setminus An$, hence $v_{i+2} \notin B$. Repeating this argument shows that all $v_{i+3}, v_{i+4}, \dots \notin B$. Since π is finite, this leads to a contradiction. (c) If $v_i \leftarrow v_{i+1} \in G_{Co}$, argue as in (b) (with B replaced by A) to conclude again that $v_i - v_{i+1} \in H^a$.

(2) (a) Suppose that $v_{i-1} \rightarrow v_i \leftarrow v_{i+1} \subseteq G$. Since $v_i \in S$, it follows that $v_{i-1} \rightarrow v_i \leftarrow v_{i+1} \subseteq G_{An(S)} \subseteq H$, hence $v_{i-1} - v_{i+1} \in H^a$. (b) Suppose that $v_{i-1} \rightarrow v_i - v_{i+1} \subseteq G$. Because $v_i \in S$, it follows that $v_{i-1} \rightarrow v_i \in G_{An(S)}$ and $v_i - v_{i+1} \in G_{Co(S)}$. Thus $v_{i-1} \rightarrow v_i - v_{i+1} \subseteq H$, hence $v_{i-1} - v_{i+1} \in H^a$. A similar argument applies if $v_{i-1} - v_i \leftarrow v_{i+1} \subseteq G$.

(3) Apply case (2b) with $v_{i-1} \rightarrow v_i - v_{i+1}$ replaced by $d_i \rightarrow v_{k_i} - v_{k_{i-1}}$ and $d_i \rightarrow v_{k_i} - v_{k_{i+1}}$ to obtain $v_{k_{i-1}} - d_i \in H^a$ and $d_i - v_{k_{i+1}} \in H^a$, respectively [refer to Figure 1(d)].

(4) Consider the first configuration. As in (2b) we see that $v_{i-1} - d'_i \in H^a$ and $d'_j - v_{j+1} \in H^a$. Next, because $v_i, v_{i+1} \in S$ and because d'_i and d'_{i+1} activate only v_i and v_{i+1} , respectively, $[d'_i, d'_{i+1}; v_i, v_{i+1}]$ is a 2-biflag in $G_{An(S)} \subseteq H$, hence $d'_i - d'_{i+1} \in H^a$. Similarly, $d'_{i+1} - d'_{i+2} \in H^a, \dots, d'_{j-1} - d'_j \in H^a$. Now consider the second configuration. It follows as in (a) that $v_{i-1} - d'_i \in H^a$, $d'_i - d'_{i+1} \in H^a, \dots$, and $d'_{j-2} - d'_{j-1} \in H^a$. Furthermore, $v_{j-1} \not\rightarrow v_{j+1}$ in G ; otherwise, $v_{j-1} \leftarrow v_{j+1} \in G$ so $\pi' := (v_0, \dots, v_{j-1}, v_{j+1}, \dots, v_n)$ would be an S -active path between A and B in G such that $|S_{\pi'}| = |S_\pi| - 1$, contradicting M1. Therefore, $[d'_{j-1}, v_{j+1}; v_{j-1}, v_j]$ is a 2-biflag in $G_{An(S)} \subseteq H$, so $d'_{j-1} - v_{j+1} \in H^a$. The third configuration is similar to the second. If $j \geq i + 2$, the fourth configuration can be treated as a combination of the second and third configurations. If $j = i + 1$, then by again appealing to M1, $[v_{i-1}, v_{j+1}; v_i, v_j]$ is seen to be a 2-biflag in G , hence $v_{i-1} - v_{j+1} \in H^a$. This completes the verification that each edge in $\tilde{\pi}$ actually occurs in H^a .

PROOF OF THEOREM 4.1 ("only if"). Suppose that $\tilde{\pi} \equiv (v_0, \dots, v_n)$ is an S -bypassing path of *minimal length* between A and B in $H^a \equiv G[ABS]^a$. We shall show that $\tilde{\pi}$ can be modified to obtain an S -active trail π between A and B in G . Several facts about $\tilde{\pi}$ are needed.

FACT 1. For each $i = 1, \dots, n-1$, neither $v_{i-1} \rightarrow v_i \leftarrow v_{i+1}$, $v_{i-1} \rightarrow v_i - v_{i+1}$, nor $v_{i-1} - v_i \leftarrow v_{i+1}$ occurs as a subgraph of H .

PROOF. If any of these configurations occurred as a subgraph of H , then $v_{i-1} - v_{i+1} \in H^a$ so $(v_0, \dots, v_{i-1}, v_{i+1}, \dots, v_n)$ would also be an S -bypassing path in H^a , contradicting the minimality of $\tilde{\pi}$. \square

If no edges in $\tilde{\pi}$ occur in H^a due to augmentation in H , then by Fact 1 and Table 2, $\tilde{\pi}$ itself is the desired S -active trail π between A and B in G , so we assume that at least one such augmentation edge is present in $\tilde{\pi}$.

Suppose that the edge $v_{i-1} - v_i$ in $\tilde{\pi}$ occurs in H^a due to augmentation in H ($i = 1, \dots, n$). That is, either $(\{v_{i-1}, v_i\}, w)$ is a triplex in H for some (not necessarily unique) $w \in An := An(ABS)$, or, if not, then $[v_{i-1}, v_i; w', w'']$ is a 2-biflag in H for some (not necessarily unique) $w', w'' \in An$. Note: In the next two figures, the vertices v_{i-1}, v_i are indicated by open circles and w, w', w'' by dark circles.

FACT 2. $w, w', w'' \neq v_j$ for any $j = 0, \dots, n$.

PROOF. First suppose that $(\{v_{i-1}, v_i\}, w)$ is a triplex in H such that $w = v_j$, so $v_{i-1} \cdots v_j \in H \subseteq H^a$ and $v_i \cdots v_j \in H \subseteq H^a$. Clearly, $j \neq i-1, i$. If $j \leq i-2$ then $(v_0, \dots, v_j, v_i, \dots, v_n)$ is an S -bypassing path in H^a , while if $j \geq i+1$ then $(v_0, \dots, v_{i-1}, v_j, \dots, v_n)$ is an S -bypassing path in H^a ; in either case the minimality of $\tilde{\pi}$ is contradicted. Next suppose that $[v_{i-1}, v_i; w', w'']$ is a 2-biflag in H such that $w' = v_j$. Here, $v_{i-1} \rightarrow v_j \in H \subseteq H^a$ and $j \neq i-1, i$. If $j \leq i-3$ then $(v_0, \dots, v_j, v_{i-1}, \dots, v_n)$ is an S -bypassing path in H^a ; if $j \geq i+1$ then $(v_0, \dots, v_{i-1}, v_j, \dots, v_n)$ is an S -bypassing path in H^a ; if $j = i-2$ then $(\{v_{i-2}, v_i\}, w'')$ is a triplex in H , so $v_{i-2} - v_i \in H^a$, hence $(v_0, \dots, v_{i-2}, v_i, \dots, v_n)$ is an S -bypassing path in H^a ; in all three cases the minimality of $\tilde{\pi}$ is contradicted. The supposition that $[v_{i-1}, v_i; w', w'']$ is a 2-biflag in H such that $w'' = v_j$ similarly leads to a contradiction. \square

Suppose that exactly m edges occur in $\tilde{\pi}$ due to augmentation in H^a ($1 \leq m \leq n$):

$$v_{i_1-1} - v_{i_1}, \dots, v_{i_m-1} - v_{i_m}, \quad 1 \leq i_1 < i_2 < \dots < i_m \leq n.$$

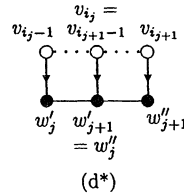
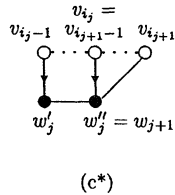
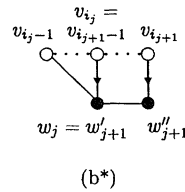
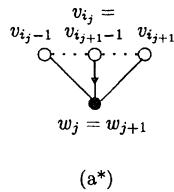
Define the sequence π' as follows:

$$\begin{aligned} \pi' := & (v_0, \dots, v_{i_1-1}, [w]_1, v_{i_1}, \dots, v_{i_2-1}, [w]_2, v_{i_2}, \dots, v_{i_3-1}, \\ & \dots, [w]_{m-1}, v_{i_{m-1}}, \dots, v_{i_m-1}, [w]_m, v_{i_m}, \dots, v_n), \end{aligned}$$

where $[w]_j$ denotes w_j if the triplex $(\{v_{i_j-1}, v_{i_j}\}, w_j)$ occurs in H , or (w'_j, w''_j) if no such triplex occurs but the 2-biflag $[v_{i_j-1}, v_{i_j}; w'_j, w''_j]$ occurs in H . (If more than one choice for w_j , w'_j , and/or w''_j is possible, the choice is made arbitrarily, then fixed.)

The vertices in $[w]_1, \dots, [w]_m$ are not necessarily distinct, but coincidences among them can occur in only four restricted ways. The nine possible coincidences are (a) $w_j = w_k$; (b) $w_j = w'_k$; (c) $w''_j = w_k$; (d) $w''_j = w'_k$; (e) $w_j = w''_k$; (f) $w'_j = w_k$; (g) $w'_j = w'_k$; (h) $w''_j = w''_k$; (i) $w'_j = w''_k$ ($1 \leq j < k \leq m$).

FACT 3. Of these nine possible coincidences, only the first four can actually occur, and then only when $k = j + 1$ and $i_{j+1} = i_j + 1$. If any of (a), (b), (c) or (d) does occur then it must assume the form (a*) $w_j = w_{j+1}$, (b*) $w_j = w'_{j+1}$, (c*) $w''_j = w_{j+1}$, or (d*) $w''_j = w'_{j+1}$, respectively, and the triplexes and 2-biflags associated with these vertices are configured in H as follows:



PROOF. (a) If $w_j = w_k$ then, because the triplexes $(\{v_{i_j-1}, v_{i_j}\}, w_j)$ and $(\{v_{i_k-1}, v_{i_k}\}, w_k)$ occur in H , it must be that either $v_{i_j-1} \rightarrow w_j \leftarrow v_{i_k} \subseteq H$, $v_{i_j-1} \rightarrow w_j - v_{i_k} \subseteq H$, $v_{i_j-1} - w_j \leftarrow v_{i_k} \subseteq H$, or $v_{i_j-1} - w_j - v_{i_k} \subseteq H$. In the first three cases $v_{i_j-1} - v_{i_k} \in H^a$, hence $(v_0, \dots, v_{i_j-1}, v_{i_k}, \dots, v_n)$ is an S -bypassing path in H^a that is shorter than $\tilde{\pi}$ because $i_k \geq i_j + 1$, contradicting the minimality of $\tilde{\pi}$. In the fourth case, $v_{i_j-1} - v_{i_k-1} \in H^a$, hence $(v_0, \dots, v_{i_j-1}, v_{i_k-1}, \dots, v_n)$ is an S -bypassing path in H^a . This path is shorter than $\tilde{\pi}$ unless $i_k = i_j + 1$, which occurs iff $k = j + 1$ and $v_{i_j} = v_{i_{j+1}-1}$, in which case $v_{i_j-1}, v_{i_j} \equiv v_{i_{j+1}-1}$, and $v_{i_{j+1}}$ are consecutive vertices in $\tilde{\pi}$, so configuration (a*) occurs as a subgraph of H .

(b) If $w_j = w'_k$ then, because the triplex $(\{v_{i_j-1}, v_{i_j}\}, w_j)$ and the 2-biflag $[v_{i_k-1}, v_{i_k}, w'_k, w''_k]$ occur in H , either $v_{i_j-1} \rightarrow w'_k - w''_k \leftarrow v_{i_k} \subseteq H$ or $v_{i_j-1} - w'_k \leftarrow v_{i_k-1} \subseteq H$. In the first case $v_{i_j-1} - v_{i_k} \in H^a$, hence $(v_0, \dots, v_{i_j-1}, v_{i_k}, \dots, v_n)$ is an S -bypassing path in H^a that is shorter than $\tilde{\pi}$ because $i_k \geq i_j + 1$, contradicting the minimality of $\tilde{\pi}$. In the second case $v_{i_j-1} - v_{i_k-1} \in H^a$, hence $(v_0, \dots, v_{i_j-1}, v_{i_k-1}, \dots, v_n)$ is an S -bypassing path in H^a . This path is shorter than $\tilde{\pi}$ unless $i_k = i_j + 1$, which occurs iff $k = j + 1$ and $v_{i_j} = v_{i_{j+1}-1}$, in which

case $v_{i_j-1}, v_{i_j} \equiv v_{i_{j+1}-1}$, and $v_{i_{j+1}}$ are consecutive vertices in $\tilde{\pi}$, so configuration (b*) occurs as a subgraph of H .

(c) This case is analogous to (b) and leads to configuration (c*).

(d) If $w'_j = w'_k$ then, because the 2-biflags $[v_{i_j-1}, v_{i_j}; w'_j, w'_j]$ and $[v_{i_k-1}, v_{i_k}; w'_k, w'_k]$ occur in H , $v_{i_j-1} \rightarrow w'_j - w'_k \leftarrow v_{i_k-1} \subseteq H$. Therefore $v_{i_j-1} - v_{i_k-1} \in H^a$, so $(v_0, \dots, v_{i_j-1}, v_{i_k-1}, \dots, v_n)$ is an S -bypassing path in H^a . This path is shorter than $\tilde{\pi}$ unless $i_k = i_j + 1$, which occurs iff $k = j + 1$ and $v_{i_j} = v_{i_{j+1}-1}$, whence $v_{i_j-1}, v_{i_j} \equiv v_{i_{j+1}-1}$, and $v_{i_{j+1}}$ are consecutive vertices in $\tilde{\pi}$. Furthermore, $w'_j \neq w''_k$ by the impossibility of case (i) (see below), hence configuration (d*) must occur as a subgraph of H .

(e) If $w_j = w''_k$ then, because the triplex $(\{v_{i_j-1}, v_{i_j}\}, w_j)$ and the 2-biflag $[v_{i_k-1}, v_{i_k}; w'_k, w''_k]$ occur in H , it must be that either $v_{i_j-1} \rightarrow w_j \leftarrow v_{i_k} \subseteq H$ or $v_{i_j-1} - w_j \leftarrow v_{i_k} \subseteq H$. Thus $v_{i_j-1} - v_{i_k} \in H^a$, hence $(v_0, \dots, v_{i_j-1}, v_{i_k}, \dots, v_n)$ is an S -bypassing path in H^a that is shorter than $\tilde{\pi}$ because $i_k \geq i_j + 1$, contradicting the minimality of $\tilde{\pi}$.

(f) By an argument similar to (e), this case is also impossible.

(g) If $w'_j = w'_k$ then, because the 2-biflags $[v_{i_j-1}, v_{i_j}; w'_j, w'_j]$ and $[v_{i_k-1}, v_{i_k}; w'_k, w''_k]$ occur in H , necessarily $v_{i_j-1} \rightarrow w'_j \leftarrow v_{i_k-1} \subseteq H$. Thus $v_{i_j-1} - v_{i_k-1} \in H^a$, hence $(v_0, \dots, v_{i_j-1}, v_{i_k-1}, \dots, v_n)$ is an S -bypassing path in H^a . Because $i_k - 1 \geq i_j$ and $v_{i_j} \cdot \cdot w'_j$ in H but $v_{i_k-1} \rightarrow w'_j \in H$, in fact $i_k - 1 > i_j$, hence this path is shorter than $\tilde{\pi}$, contradicting minimality.

(h) By an argument similar to (g), this case is also impossible.

(i) If $w'_j = w''_k$ then, because the 2-biflags $[v_{i_j-1}, v_{i_j}; w'_j, w'_j]$ and $[v_{i_k-1}, v_{i_k}; w'_k, w''_k]$ occur in H , necessarily $v_{i_j-1} \rightarrow w'_j \leftarrow v_{i_k} \subseteq H$. Thus $v_{i_j-1} - v_{i_k} \in H^a$, so $(v_0, \dots, v_{i_j-1}, v_{i_k}, \dots, v_n)$ is an S -bypassing path in H^a . Since $i_k \geq i_j + 1$, this path is shorter than $\tilde{\pi}$, contradicting minimality. \square

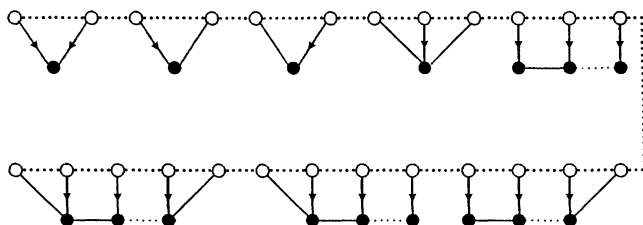
We now modify π' to produce a *trail* $\bar{\pi}$ in H . Define

$$\begin{aligned} \bar{\pi} := & (v_0, \dots, v_{i_1-1}, [\bar{w}]_1, \bar{v}_{i_1}, \dots, \bar{v}_{i_2-1}, [\bar{w}]_2, \bar{v}_{i_2}, \dots, \bar{v}_{i_3-1}, \\ & \dots, [\bar{w}]_{m-1}, \bar{v}_{i_{m-1}}, \dots, \bar{v}_{i_m-1}, [w]_m, v_{i_m}, \dots, v_n), \end{aligned}$$

where, for $1 \leq j \leq m-1$,

$$([\bar{w}]_j, \bar{v}_{i_j}, \dots, \bar{v}_{i_{j+1}-1}) = \begin{cases} \emptyset, & w_j = w_{j+1} \text{ or} \\ & w_j = w'_{j+1}, \\ w'_j, & w'_j = w_{j+1} \text{ or} \\ & w'_j = w'_{j+1}, \\ ([w]_j, v_{i_j}, \dots, v_{i_{j+1}-1}), & \text{otherwise.} \end{cases}$$

The form of $\bar{\pi}$ is illustrated by the following figure:



The original path $\tilde{\pi}$ is indicated by the heavy dotted lines, both horizontal and descending. The trail $\bar{\pi}$ departs from $\tilde{\pi}$ by following the outermost solid arrows and/or solid lines and proceeding along the dark circles (which indicate the vertices w_j, w'_j, w''_j that occur in $[\bar{w}]_j$). By Facts 2 and 3, all vertices in $\bar{\pi}$ are distinct, all vertices in $[w]_1, \dots, [w]_m$ occur exactly once in $\bar{\pi}$, and all edges between consecutive vertices in $\bar{\pi}$ occur in H , so $\bar{\pi}$ is a trail in H and therefore in G . Recall that $v_0 \in A$, $v_n \in B$, $v_1, \dots, v_{n-1} \in \text{Co}(\text{An}(\text{ABS})) \setminus (\text{ABS})$ and $w_j, w'_j, w''_j \in \text{An}(\text{ABS})$ for all $j = 1, \dots, m$. We shall use $\bar{\pi}$ to construct an S -active trail π between A and B in G .

We begin by examining the S -active/blocking status of the interior vertices in the trail $\bar{\pi}$ itself. Relabel the vertices in $\bar{\pi}$ as follows:

$$\bar{\pi} =: (x_0, x_1, \dots, x_p),$$

where $p \geq 2$, $x_0 = v_0 \in A$, $x_p = v_n \in B$. We may assume that $\bar{\pi}$ is a trail between A and B in G , that is, $\bar{\pi}^\circ \cap AB = \emptyset$; otherwise simply replace $\bar{\pi}$ by any subtrail that does lie between A and B in G . Each $x_r \equiv \hat{x}$ ($1 \leq r \leq p-1$) in $\bar{\pi}^\circ$ is adjacent in G to its two neighbors $x_{r-1} \equiv \dot{x}$ and $x_{r+1} \equiv \ddot{x}$ in $\bar{\pi}$, forming a linked triple

$$T := \dot{x} \cdots \hat{x} \cdots \ddot{x}.$$

Either $\hat{x} = v_i$ for some $v_i \in \tilde{\pi}$, or $\hat{x} = w_j$ for some triplex $(\{v_{i_{j-1}}, v_{i_j}\}, w_j)$ such that $w_j \in \tilde{\pi}$, or: $\hat{x} = w'_j$ or $\hat{x} = w''_j$ for some 2-biflag $[v_{i_{j-1}}, v_{i_j}; w'_j, w''_j]$ such that $w'_j, w''_j \in \tilde{\pi}$. If $1 < r < p-1$ then, since $\dot{x}, \hat{x}, \ddot{x}$ each must be either “ v ” $\equiv v_i \in \tilde{\pi}$ or “ w ” $\equiv w_j, w'_j, w''_j \in \tilde{\pi}$ and since each edge “ \dots ” in T must be either “ \rightarrow ”, “ \leftarrow ”, or “ $-$ ”, there are $2^3 3^2 = 72$ possibilities for the form of T . (But only 36 if $r = 1 < p-1$ or $1 < r = p-1$, and only 18 if $r = 1 = p-1$.) However, by the construction of $\bar{\pi}$ [recall Figures 1(a), 1(b), 1(c) and (a*), (b*), (c*), (d*)], many of these possibilities cannot occur in T :

- (i) An edge of the form $v \cdots w$ in T must be either $v \rightarrow w$ or $v - w$.
- (ii) An edge of the form $w \cdots w$ in T must be $w - w$.
- (iii) T cannot have the form $\dot{v} \rightarrow \hat{v} - \ddot{w}$ or $\dot{w} - \hat{v} \leftarrow \ddot{v}$. If the former occurred then $\dot{v} \rightarrow \hat{v} - \ddot{w} \leftarrow \ddot{v} \subseteq H$ (see preceding figure) where $\dot{v}, \hat{v}, \ddot{v}$ are consecutive vertices in $\tilde{\pi}$. This implies that $\dot{v} - \ddot{v} \in H^a$, contradicting the minimality of $\tilde{\pi}$. Similarly, $\dot{w} - \hat{v} \leftarrow \ddot{v}$ is also impossible.

Furthermore,

(iv) An edge of the form $v \rightarrow v$ or $v \leftarrow v$ in T must have occurred in $\bar{\pi}$, hence both v 's $\in \text{An}(ABS)$.

(v) If T has the form $\dot{x} - \hat{w} - \ddot{x}$ with $\hat{w} \in S$, then \hat{w} is S -activated in $\bar{\pi}$.

We now examine the 72 possibilities for $T \equiv \dot{x} \cdots \hat{x} \cdots \ddot{x}$ in detail.

- (1) $\dot{v} \cdots \hat{v} \cdots \ddot{v}$. Here \hat{v} cannot be head-no-tail in $\bar{\pi}$ by Fact 1 so, because $\hat{v} \notin S$ ($\bar{\pi} \cap S = \emptyset$), \hat{v} is S -active in $\bar{\pi}$.
- (2) $\dot{v} \cdots \hat{v} \cdots \ddot{w}$. By (i) and (iii), this must occur in one of five forms: $\dot{v} \cdots \hat{v} \rightarrow \ddot{w}$ (three forms), $\dot{v} \leftarrow \hat{v} - \ddot{w}$, $\dot{v} - \hat{v} - \ddot{w}$. Here $\hat{v} \notin S$ is not head-no-tail in $\bar{\pi}$, so \hat{v} is S -active in $\bar{\pi}$.
- (3) $\dot{w} \cdots \hat{v} \cdots \ddot{v}$. As in (2), \hat{v} is S -active in $\bar{\pi}$.
- (4) $\dot{v} \cdots \hat{w} \cdots \ddot{v}$. By (i), this must occur in one of four forms:

(a) $\dot{v} \rightarrow \hat{w} \leftarrow \ddot{v}$, $\dot{v} \rightarrow \hat{w} - \ddot{v}$, $\dot{v} - \hat{w} \leftarrow \ddot{v}$, where $\hat{w} \in \text{An}(ABS) \setminus AB$. If $\hat{w} \in \text{An}(S)$ then \hat{w} is S -active in $\bar{\pi}$. If $\hat{w} \notin \text{An}(S)$ then $\hat{w} \in \text{an}(AB) \setminus \text{An}(S)$ and \hat{w} is S -blocking.

(b) $\dot{v} - \hat{w} - \ddot{v}$. If $\hat{w} \notin S$ then \hat{w} is S -active in $\bar{\pi}$. If $\hat{w} \in S$ then by (v), \hat{w} is S -activated in $\bar{\pi}$, hence is again S -active.

- (5) $\dot{w} \cdots \hat{w} \cdots \ddot{v}$. By (i) and (ii), this must occur in one of two forms:

(a) $\dot{w} - \hat{w} \leftarrow \ddot{v}$, where $\hat{w} \in \text{An}(ABS) \setminus AB$. If $\hat{w} \in \text{An}(S)$ then \hat{w} is S -active in $\bar{\pi}$. If $\hat{w} \notin \text{An}(S)$ then $\hat{w} \in \text{an}(AB) \setminus \text{An}(S)$ and \hat{w} is S -blocking.

(b) $\dot{w} - \hat{w} - \ddot{v}$. If $\hat{w} \notin S$ then \hat{w} is S -active in $\bar{\pi}$. If $\hat{w} \in S$ then by (v), \hat{w} is S -activated in $\bar{\pi}$, hence is again S -active.

- (6) $\dot{v} \cdots \hat{w} \cdots \ddot{w}$. As in (5), \hat{w} is S -active in $\bar{\pi}$ unless $\dot{v} \rightarrow \hat{w} - \ddot{w}$ and $\hat{w} \in \text{an}(AB) \setminus \text{An}(S)$, in which case \hat{w} is S -blocking.
- (7) $\dot{w} \cdots \hat{v} \cdots \ddot{w}$. By (i), this must occur in one of four forms:

(a) $\dot{w} \leftarrow \hat{v} \rightarrow \ddot{w}$, $\dot{w} \leftarrow \hat{v} - \ddot{w}$, $\dot{w} - \hat{v} \rightarrow \ddot{w}$. Here $\hat{v} \notin S$ is not head-no-tail, so \hat{v} is S -active in $\bar{\pi}$.

(b) $\dot{w} - \hat{v} - \ddot{w}$. Since $\hat{v} \notin S$, \hat{v} is S -active in $\bar{\pi}$.

- (8) $\dot{w} \cdots \hat{w} \cdots \ddot{w}$. By (ii), this must occur as $\dot{w} - \hat{w} - \ddot{w}$. If $\hat{w} \notin S$ then \hat{w} is S -active in $\bar{\pi}$. If $\hat{w} \in S$ then by (v), \hat{w} is S -activated in $\bar{\pi}$, hence is again S -active.

Thus, S -blocking vertices \hat{w} occur in $\bar{\pi}^\circ$ in exactly five ways:

(4a) $\dot{v} \rightarrow \hat{w} \leftarrow \ddot{v}$, $\dot{v} \rightarrow \hat{w} - \ddot{v}$, $\dot{v} - \hat{w} \leftarrow \ddot{v}$, $\hat{w} \in \text{an}(AB) \setminus \text{An}(S)$;

(5a) $\dot{w} - \hat{w} \leftarrow \ddot{v}$, $\hat{w} \in \text{an}(AB) \setminus \text{An}(S)$;

(6) $\dot{v} \rightarrow \hat{w} - \ddot{w}$, $\hat{w} \in \text{an}(AB) \setminus \text{An}(S)$.

Let $\Delta \subseteq \text{an}(AB) \setminus \text{An}(S)$ denote the set of all such S -blocking vertices $\hat{w} \in \bar{\pi}^\circ$. If $\Delta = \emptyset$ then $\bar{\pi}$ is the desired S -active path between A and B in G . If $\Delta \neq \emptyset$, without loss of generality assume that $\Delta_A := \Delta \cap \text{an}(A) \neq \emptyset$, set

$\Delta_B := \Delta \setminus \text{an}(A) \subseteq \text{an}(B)$, and define

$$r^* := \max\{r \mid x_r \in \bar{\pi}^\circ \cap \text{an}(A) \cap \text{De}(\Delta_A)\},$$

$$t^* := \begin{cases} \min\{t \mid t > r^*, x_t \in \bar{\pi}^\circ \cap \text{an}(B) \cap \text{De}(\Delta_B)\}, & \text{if } \Delta_B \neq \emptyset, \\ p, & \text{if } \Delta_B = \emptyset, \end{cases}$$

so $1 \leq r^* < t^* \leq p$. Let

$$\pi_y := (y_l \leftarrow \cdots \leftarrow y_1 \leftarrow y_0 \equiv x_{r^*}),$$

$$\pi_z := (x_{t^*} \equiv z_0 \rightarrow z_1 \rightarrow \cdots \rightarrow z_q),$$

denote directed paths in G between x_{r^*} and A and between x_{t^*} and B , respectively ($l \geq 1, q \geq 0, y_l \in A, z_q \in B$). Finally, define

$$\pi := (y_l, \dots, y_0 \equiv x_{r^*}, x_{r^*+1}, \dots, x_{t^*-1}, x_{t^*} \equiv z_0, \dots, z_q)$$

$$=: (\pi_y, \bar{\pi}_x^\circ, \pi_z),$$

where $\bar{\pi}_x := (x_{r^*}, \dots, x_{t^*})$. (Note that $\bar{\pi}_x^\circ = \emptyset$ if $t^* = r^* + 1$.)

It is readily verified that $\pi_y \cap \bar{\pi}_x^\circ = \bar{\pi}_x^\circ \cap \pi_z = \pi_y \cap \pi_z = \emptyset$. Thus π itself is not self-intersecting and hence is a trail between A and B in G . Because $\bar{\pi}_x$ is a subtrail of $\bar{\pi}$ and $\bar{\pi}_x^\circ \cap \Delta = \emptyset$, each $x_r \in \bar{\pi}_x^\circ$ (if any) is S -active in $\bar{\pi}$ and therefore in π . Because $\pi_y \dot{\cup} \pi_z \subseteq \text{De}(\Delta)$, necessarily $\pi_y \cap S = \pi_z \cap S = \emptyset$. Since each $y_r \in \pi_y$ and each $z_t \in \pi_z$ is not head-no-tail in π , each is also S -active in π , hence π is an S -active trail between A and B in G . This completes the proof of Theorem 4.1.

COROLLARY 4.1. *If P is AMP global G -Markovian, then*

$$A \bowtie_p B \mid S[G] \implies A \perp\!\!\!\perp B \mid S[P].$$

COROLLARY 4.2. *Let $G' \equiv (V, E')$ be a subgraph of $G \equiv (V, E)$ with the same vertex set V . Then*

$$\mathcal{P}_{\text{AMP}}^G(G'; X) \subseteq \mathcal{P}_{\text{AMP}}^G(G; X) \quad \forall X.$$

PROOF. By Definition 2.5 and Theorem 4.1, it suffices to show that

$$(4.4) \quad A \bowtie_p B \mid S[G] \implies A \bowtie_p B \mid S[G'].$$

For this, it suffices to show that any S -active trail π between A and B in G' is also an S -active trail in G . Because $G' \subseteq G$, π is also a trail in G with exactly the same edges as in G' and (refer to Table 2) $\text{an}_{G'}(S) \subseteq \text{an}_G(S)$ and $\text{pa}_{G'}(w) \subseteq \text{pa}_G(w)$. Thus if π were S -blocked in G then it would be S -blocked in G' , a contradiction. \square

5. Completeness of the AMP global Markov property. Completeness of the global Markov property for UG models can be proved in a straightforward manner [see Frydenberg (1990b), Theorem 2.3 for a proof in a more general context]. For ADG models, the d -separation criterion was applied elegantly by Geiger and Pearl (1988) to establish completeness of the global Markov property. For CGs, Studený and Bouckaert (1998) established the equivalence of their c -separation criterion to the LWF moralization criterion, then applied this to prove completeness of the LWF global Markov property.

For our second main result, we apply the equivalence of p -separation and the AMP augmentation criterion (Theorem 4.1) to prove completeness of the AMP global Markov property for CGs. As in Geiger and Pearl (1988), completeness is established by the construction of a nonsingular Gaussian G -Markovian distribution on $\mathbf{X} := \mathbb{R}^V$ that violates a CI not specified by the p -separation criterion.

THEOREM 5.1. *Let $G \equiv (V, E)$ be a chain graph and let $A, B, S \subseteq V$ be mutually disjoint with $A, B \neq \emptyset$. If $A \not\bowtie_p B \mid S[G]$, then there exists a Gaussian $P \equiv P_{A, B, S} \in \mathcal{P}_{\text{AMP}}^G(G; \mathbb{R}^V)$ such that $A \not\perp B \mid S[P]$.*

For the proof of Theorem 5.1 we shall need both Lemma 4.1 and the following Lemma 5.1. For any S -active trail π between A and B in G , let H_π denote the set of all head-no-tail nodes in π° . (Note that $S_\pi \cap H_\pi = \emptyset$.)

Let $\pi \equiv (a \equiv v_0, \dots, v_n \equiv b)$ satisfy the following minimality condition:

(M2) π minimizes $|H_{\pi'}|$ over all S -active trails π' between A and B in G that satisfy M1.

If $r := |H_\pi| > 0$, denote the members of H_π in their order of occurrence in π by h_1, \dots, h_r and let l_i be that index such that $v_{l_i} = h_i$. Thus $1 \leq l_1 < \dots < l_r \leq n - 1$ and

$$(5.1) \quad H_\pi = \{h_1, \dots, h_r\} = \{v_{l_1}, \dots, v_{l_r}\}.$$

Because π is S -active, each $h_i \in \text{An}_G(S)$ so we can find a directed path,

$$(5.2) \quad \pi_i := (h_i \equiv y_{i0} \rightarrow y_{i1} \rightarrow \dots \rightarrow y_{in_i}) \quad (y_{in_i} \in S)$$

of length $n_i \geq 0$ between h_i and S in G . Define

$$(5.3) \quad \pi'_i := \pi_i \setminus \{h_i\} \equiv (y_{i1} \rightarrow \dots \rightarrow y_{in_i}).$$

LEMMA 5.1. *If π satisfies M2, then π , D_π , and π'_1, \dots, π'_r are mutually disjoint. That is, $\pi \cap D_\pi = \emptyset$ and:*

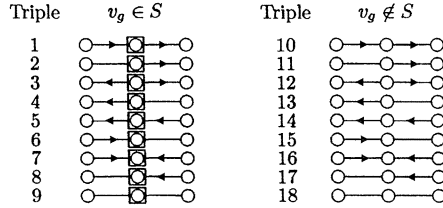
- (a) $\pi'_i \cap \pi = \emptyset$ for all $1 \leq i \leq r$.
- (b) $\pi'_i \cap \pi'_j = \emptyset$ for all $1 \leq i < j \leq r$.
- (c) $\pi'_i \cap D_\pi = \emptyset$ for all $1 \leq i \leq r$.

PROOF. Since M2 implies M1, Lemma 4.1 applies to π , so $\pi \cap D_\pi = \emptyset$. First suppose that (a) fails for some i . Let $p := \min\{e \geq 1 \mid y_{ie} \in \pi\}$. Thus $y_{ip} = v_g$ for some $g \in \{0, \dots, l_i - 1, l_i + 1, \dots, n\}$. Assume that $l_i < g$ (the case $l_i > g$ is similar). Then

$$\pi' := (v_0, \dots, v_{l_i} \equiv h_i \equiv y_{i0} \rightarrow \dots \rightarrow y_{ip} \equiv v_g, \dots, v_n)$$

is a trail between A and B in G such that $|S_{\pi'}| \leq |S_\pi|$. Because $y_{i0} \in H_\pi \setminus H_{\pi'}$ and $y_{i1}, \dots, y_{i(p-1)} \notin H_{\pi'}$, necessarily $|H_{\pi'}| \leq |H_\pi|$, with $|H_{\pi'}| < |H_\pi|$ if $g < n$ and either $v_g \in H_\pi$ or $v_g \notin H_{\pi'}$. (Note that if $g < n$, v_g may or may not be head-no-tail in π and/or π' .) Because π is S -active and $y_{i0}, \dots, y_{i(p-1)} \notin S$, all interior vertices of π' are S -active except possibly v_g if $g < n$. Thus, if $g = n$ then π' is S -active and also $y_{ip} \equiv v_g \notin H_{\pi'}$, hence $|H_{\pi'}| < |H_\pi|$, contradicting M2.

If $g < n$, the linked triple $v_{g-1} \cdots v_g \cdots v_{g+1}$ might occur in 18 possible forms in π : each edge might occur as either \rightarrow , \leftarrow , or $-$, and either $v_g \in S$ or $v_g \notin S$. These 18 forms are shown here:



Because π is S -active, forms 1–5 cannot occur. Because $y_{i(p-1)} \rightarrow v_g \in G$, v_g is S -active in π' for forms 6–9; because $v_g \equiv y_{ip} \in \text{an}_G(S)$ for forms 10–18, v_g is S -active in π' for these forms as well; hence π' is S -active for forms 6–18. But $v_g \in H_\pi$ for forms 6–8 and 15–17, while $v_g \notin H_{\pi'}$ for forms 10–12, so $|H_{\pi'}| < |H_\pi|$ for forms 6–8, 10–12, and 15–17, contradicting M2.

For each form 9, 13, 14 and 18, $v_g \in H_{\pi'} \setminus H_\pi$. In these four cases, however, necessarily $g \geq l_i + 2$ and $v_k \rightarrow v_{k+1} \in G$ for at least one $k = l_i, \dots, g - 2$; otherwise $y_{i0}, \dots, y_{ip} \equiv v_g, \dots, v_{l_i} \equiv y_{i,0}$ would be a semidirected cycle in G , contradicting its adicyclicity. Therefore $H_\pi \cap \{v_{l+1}, \dots, v_{g-1}\} \neq \emptyset$, which implies that $|H_{\pi'}| < |H_\pi|$ for forms 9, 13, 14, 18, again contradicting M2. Thus, (a) holds.

Next, suppose that (b) fails to hold for some $1 \leq i < j \leq r$. Define $p := \min\{e \geq 1 \mid y_{ie} \in \pi'_j\}$, and let $q \in \{1, \dots, n_j\}$ be the unique index such that $y_{ip} = y_{jq}$. Then by (a),

$$\pi' := (v_0, \dots, v_{l_i} \equiv h_i \equiv y_{i0} \rightarrow \dots \rightarrow y_{ip} \equiv y_{jq} \leftarrow \dots \leftarrow y_{j0} \equiv h_j \equiv v_{l_j}, \dots, v_n)$$

is a trail between A and B in G such that $|S_{\pi'}| \leq |S_\pi|$. Because $y_{i0}, \dots, y_{i(p-1)}, y_{j(q-1)}, \dots, y_{j0} \notin S$ and each of these vertices is not head-no-tail in π' , each is S -active in π' . Furthermore, $y_{ip} \equiv y_{jq} \in \text{an}_G(S)$ and is head-no-tail in π' , so is also S -active in π' . Therefore, because π is S -active, so too is π' . But $h_i, h_j \in H_\pi \setminus H_{\pi'}$ while $y_{i1}, \dots, y_{i(p-1)}, y_{j(q-1)}, \dots, y_{j1} \notin H_{\pi'}$, so,

although $y_{ip} \equiv y_{jq} \in H_{\pi'}$, necessarily $|H_{\pi'}| < |H_{\pi}|$, contradicting M2. Thus, (b) holds.

Last, suppose that (c) fails for some i . Define $p := \min\{e \geq 1 \mid y_{ie} \in D_{\pi}\}$. Thus $y_{ip} = d_j$ for some $j = 1, \dots, m$ [recall (4.3)]. Assume that $l_i < k_j$, where, as in Section 4, $s_j = v_{k_j}$. (The case $l_i > k_j$ is similar). Then by (a),

$$\pi' := (v_0, \dots, v_{l_i} \equiv h_i \equiv y_{i0} \rightarrow \dots \rightarrow y_{ip} \equiv d_j \rightarrow s_j \equiv v_{k_j} - v_{k_{j+1}}, \dots, v_n)$$

is a trail between A and B in G . Because $y_{i0}, \dots, y_{i(p-1)} \notin S$ and $y_{ip} \equiv d_j \notin S$, and because each of these vertices is not head-no-tail in π' , each is S -active in π' . Furthermore, $v_{k_j} \equiv s_j \in S$ and is head-no-tail in π' , hence is S -active in π' , so π' is an S -active trail. But $s_j \in S_{\pi} \setminus S_{\pi'}$, so $|S_{\pi'}| < |S_{\pi}|$, contradicting M1 and therefore M2. Thus, (c) holds. \square

PROOF OF THEOREM 5.1. Suppose that $\pi \equiv (a \equiv v_0, \dots, v_n \equiv b)$ is an S -active trail between A and B in G that satisfies M2 (and therefore M1). Let $G^* \equiv (V^*, E^*)$ be the subgraph (not necessarily induced) of G consisting of the trail π , the S -activating arrows $d_1 \rightarrow s_1, \dots, d_m \rightarrow s_m$ and the directed paths π_1, \dots, π_r . By Lemmas 4.1 and 5.1, G^* has the form illustrated in Figure 2. Here, s_1, \dots, s_m are the S -activated nodes in π and d_1, \dots, d_m the corresponding S -activators; h_1, \dots, h_r are the head-no-tail nodes in π ; boxed nodes belong to S while unboxed nodes do not. (In Figure 2 the path π_2 has length $n_2 = 0$.) For notational simplicity we have set $t_i = y_{in_i} \in S, i = 1, \dots, r$, $t_0 = a, t_{r+1} = b$. Note that G^* is *singly connected*: there exists *exactly one* trail between any pair of its vertices.

Let $G' \equiv (V, E^*)$ be the subgraph of G consisting of G^* together with the remaining vertices in $V \setminus V^*$ but no additional edges, so that each $v \in V \setminus V^*$ is an isolated vertex in G' . Thus:

(*) G' is a chain graph such that there exists at most one trail between any pair of its vertices.

Undirected edges (if any) of G^* and G' occur only within the trail π . For each $i = 0, \dots, r$, the subtrail

(5.4)
$$\chi_i := \{h_i \equiv v_{l_i}, \dots, v_{l_{i+1}} \equiv h_{i+1}\}$$

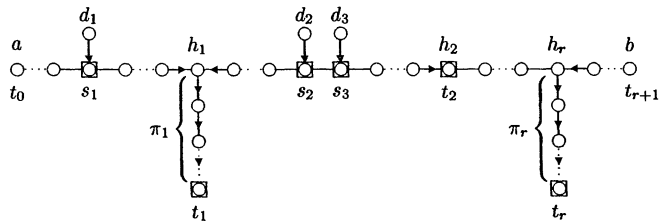


FIG. 2. The singly connected subgraph $G^* \equiv (V^*, E^*)$.

either contains exactly one maximal undirected subtrail $\tau_i \subseteq \chi_i$, or else contains no undirected edges, in which case we take $\tau_i = \emptyset$. (We have set $l_0 = 0$, $h_0 = v_{l_0} \equiv a$, $l_{r+1} = n$, and $h_{r+1} = v_{l_{r+1}} \equiv b$.) If $\tau_i \neq \emptyset$, then τ_i has the form

$$(5.5) \quad \tau_i = \{v_{e_i}, v_{e_i+1}, \dots, v_{f_i-1}, v_{f_i}\},$$

where $l_i \leq e_i < f_i \leq l_{i+1}$. Note that τ_0, \dots, τ_r are mutually disjoint and that

$$(5.6) \quad S_\pi \subseteq \tau_0^\circ \cup \dots \cup \tau_r^\circ,$$

where $\tau_j^\circ \neq \emptyset \Leftrightarrow f_j \geq e_j + 2$. Thus

$$(5.7) \quad S_\pi = \sigma_0 \cup \sigma_1 \cup \dots \cup \sigma_r,$$

where, for $0 \leq i \leq r$,

$$(5.8) \quad \sigma_i := S_\pi \cap \tau_i^\circ.$$

These features of G^* are illustrated in Figure 3, where $n = 12$, $r = 3$, $a = h_0 = t_0$, $\tau_0 = \sigma_0 = \delta_0 = \emptyset$, $\pi'_2 = \emptyset$, $e_1 = 1$, $f_1 = 4$, $e_2 = 5$, $f_2 = 9$, $e_3 = 10$, $f_3 = 12 (= n)$, $t_4 = v_{f_3} = h_4 = b$.

We now define the required Gaussian $P \equiv P_{A,B,S} \in \mathcal{P}_{\text{AMP}}^g(G; \mathbb{R}^V)$. Let $(\zeta_v | v \in V)$ be a family of independent, identically distributed $\mathcal{N}(0, 1)$ random variables and define

$$(5.9) \quad \varepsilon_v := \begin{cases} \zeta_v, & \text{if } v \notin \tau_0 \cup \dots \cup \tau_r, \\ \zeta_{v_{e_i}} + \dots + \zeta_{v_{e_i+q}}, & \text{if } v = v_{e_i+q} \in \tau_i \neq \emptyset. \end{cases}$$

Define a Gaussian random vector $X \equiv (X_v | v \in V)$ according to the explicit (not recursive) formula

$$(5.10) \quad X_v := \sum (\varepsilon_w | w \in \text{An}_{G'}(v)), \quad v \in V.$$

However,

$$\text{An}_{G'}(v) = \left(\bigcup (\text{An}_{G'}(w) | w \in \text{pa}_{G'}(v)) \right) \cup \{v\}$$

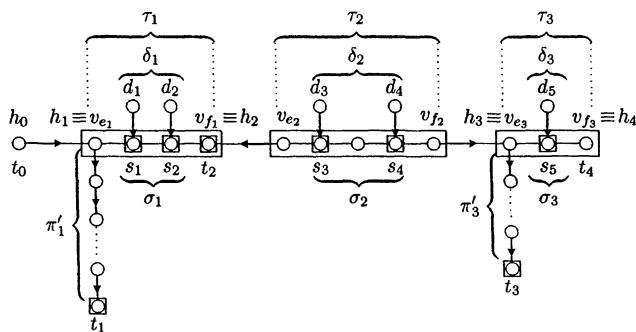


FIG. 3.

and, by (*), $(\text{An}_{G'}(w) \mid w \in \text{pa}_{G'}(v))$ is a *disjoint* family, hence

$$X_v = \sum (X_w \mid w \in \text{pa}_{G'}(v)) + \varepsilon_v, \quad v \in V.$$

Therefore, for each $\tau \in \mathcal{T}(G')$ (the set of *chain components* of G'),

$$X_\tau = \beta_\tau X_{\text{pa}_{G'}(\tau)} + \varepsilon_\tau,$$

where $\beta_\tau \equiv (\beta_{vw} \mid v \in \tau, w \in \text{pa}_{G'}(\tau))$ is the $\tau \times \text{pa}_{G'}(\tau)$ matrix given by

$$\beta_{vw} := \begin{cases} 1, & \text{if } w \in \text{pa}_{G'}(v), \\ 0, & \text{if } w \notin \text{pa}_{G'}(v). \end{cases}$$

The nonsingleton chain components of G' are exactly the nonempty members of $\{\tau_0, \dots, \tau_r\}$, and if $\tau_i \neq \emptyset$ then ε_{τ_i} is global G'_{τ_i} -Markovian by (5.9). Also by (5.9), the random variates $\{\varepsilon_\tau \mid \tau \in \mathcal{T}(G')\}$ are mutually independent, Gaussian, and nonsingular. Thus, if we denote the joint distribution of X by $P \equiv P_{A,B,S}$, it follows from Remark 5.1 of [AMP] (2001) and Corollary 4.2 that P is Gaussian, nonsingular and

$$P \in \mathcal{P}_{\text{AMP}}^g(G'; \mathbb{R}^V) \subseteq \mathcal{P}_{\text{AMP}}^g(G; \mathbb{R}^V),$$

as required.

To complete the proof of Theorem 5.1 we will show that $A \not\perp\!\!\!\perp B \mid S[P]$, in particular $a \not\perp\!\!\!\perp b \mid S[P]$. Since $a, b \in V^*$ and $V^* \perp\!\!\!\perp (V \setminus V^*)[P]$, it suffices to show that $a \not\perp\!\!\!\perp b \mid S^*[P]$, where

$$(5.11) \quad S^* := S \cap V^* = S_\pi \cup \{t_1, \dots, t_r\}.$$

By normality, it suffices to show (recall that $t_0 \equiv a, t_{r+1} \equiv b$) that

$$(5.12) \quad \text{Cov}(X_{t_0}, X_{t_{r+1}} \mid X_{S^*}) \neq 0.$$

It follows from (5.7) and (5.11) (recall that some σ_i may be empty) that

$$\{t_0\} \dot{\cup} S^* \dot{\cup} \{t_{r+1}\} = \{t_0\} \dot{\cup} \sigma_0 \dot{\cup} \{t_1\} \dot{\cup} \sigma_1 \dot{\cup} \{t_2\} \dot{\cup} \dots \dot{\cup} \{t_{r-1}\} \dot{\cup} \sigma_{r-1} \dot{\cup} \{t_r\} \dot{\cup} \sigma_r \dot{\cup} \{t_{r+1}\}.$$

CLAIM 5.1. *The covariance matrix Δ of $(X_{t_0}, X_{S^*}, X_{t_{r+1}})$ has the form*

$$\begin{array}{cccccccccccc} & t_0 & \sigma_0 & t_1 & \sigma_1 & t_2 & \cdots & t_{r-1} & \sigma_{r-1} & t_r & \sigma_r & t_{r+1} \\ \begin{array}{l} t_0 \\ \sigma_0 \\ t_1 \\ \sigma_1 \\ t_2 \\ \vdots \\ t_{r-1} \\ \sigma_{r-1} \\ t_r \\ \sigma_r \\ t_{r+1} \end{array} & \left(\begin{array}{cccccccccccc} \alpha_0 & \gamma_0 & \beta_0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ \gamma'_0 & \Delta_0 & \eta_0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ \beta_0 & \eta'_0 & \alpha_1 & \gamma_1 & \beta_1 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \gamma'_1 & \Delta_1 & \eta_1 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \beta_1 & \eta'_1 & \alpha_2 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & \alpha_{r-1} & \gamma_{r-1} & \beta_{r-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & \gamma'_{r-1} & \Delta_{r-1} & \eta_{r-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & \beta_{r-1} & \eta'_{r-1} & \alpha_r & \gamma_r & \beta_r & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \gamma'_r & \Delta_r & \eta_r & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \beta_r & \eta'_r & \alpha_{r+1} & 0 \end{array} \right), \end{array}$$

where $\alpha_i := \text{Var}(X_{t_i})$, $\beta_i := \text{Cov}(X_{t_i}, X_{t_{i+1}})$, $\gamma_i := \text{Cov}(X_{t_i}, X_{\sigma_i})$, $\eta_i := \text{Cov}(X_{\sigma_i}, X_{t_{i+1}})$, $\Delta_i := \text{Cov}(X_{\sigma_i})$, and $(')$ denotes transpose.

PROOF. We must verify that the matrix elements and nonempty blocks indicated by “0” are in fact zero. That is,

$$(5.13) \quad \text{Cov}(X_{\sigma_i}, X_{\sigma_{i'}}) = 0 \quad \text{if } i \neq i' \text{ and } \sigma_i, \sigma_{i'} \neq \emptyset,$$

$$(5.14) \quad \text{Cov}(X_{t_i}, X_{t_{i'}}) = 0 \quad \text{if } |i - i'| \geq 2,$$

$$(5.15) \quad \text{Cov}(X_{\sigma_i}, X_{t_{i'}}) = 0 \quad \text{if } i \neq i', i' - 1 \text{ and } \sigma_i \neq \emptyset.$$

If $r = 0$ then the result is vacuously true, so assume that $r \geq 1$. For any subset $\omega \subseteq V$ let $\zeta_\omega := (\zeta_v | v \in \omega)$, a collection of iid $\mathcal{N}(0, 1)$ random variables. By normality,

$$(5.16) \quad \omega_1, \dots, \omega_k \text{ mutually disjoint} \implies \zeta_{\omega_1}, \dots, \zeta_{\omega_k} \text{ mutually independent.}$$

For $i = 0, \dots, r$ define [recall (4.3) and see Figure 3]

$$(5.17) \quad \delta_i := D_\pi \cap \text{pa}_G(\sigma_i) = D_\pi \cap \text{pa}_{G'}(\sigma_i),$$

$$(5.18) \quad \tilde{\tau}_i := \{v_{e_i}\} \dot{\cup} \tau_i^\circ \equiv \tau_i \setminus \{v_{f_i}\},$$

$$(5.19) \quad \tilde{\chi}_i := \{v_{l_i}\} \dot{\cup} \chi_i^\circ \equiv \chi_i \setminus \{v_{l_{i+1}}\}.$$

Note that $\tilde{\chi}_0, \dots, \tilde{\chi}_{r-1}, \chi_r$ are mutually disjoint and that for $i = 0, \dots, r$,

$$(5.20) \quad \tilde{\chi}_i \supseteq \tilde{\tau}_i \supseteq \sigma_i.$$

By (5.8), (5.9), (5.10) and the topology of G^* and G' , X_{σ_i} is a linear function of ζ_{δ_i} and $\zeta_{\tilde{\tau}_i}$, while X_{t_i} is a linear function of $\zeta_{\pi'_i}$ and $\zeta_{\tilde{\chi}_i}$, where $\pi'_0 := \emptyset$, $\pi'_{r+1} := \emptyset$, and

$$(5.21) \quad \bar{\chi}_i := \begin{cases} \tilde{\chi}_0, & \text{if } i = 0, \\ \tilde{\chi}_{i-1} \dot{\cup} \tilde{\chi}_i, & \text{if } 1 \leq i \leq r-1, \\ \tilde{\chi}_{r-1} \dot{\cup} \chi_r, & \text{if } i = r, \\ \chi_r, & \text{if } i = r+1. \end{cases}$$

[The asymmetries in this definition are necessitated by the form (5.9) of ε_v for $v \in \tau_i$.] If $i \neq i'$ then $\delta_i, \tilde{\tau}_i, \delta_{i'}, \tilde{\tau}_{i'}$ are mutually disjoint (recall Lemma 4.1b), so (5.13) follows from (5.16). If $|i - i'| \geq 2$ then $\pi'_i, \bar{\chi}_i, \pi'_{i'}, \bar{\chi}_{i'}$ are mutually disjoint, whence (5.14) follows from (5.16). If $i \neq i', i' - 1$ then $\delta_i, \tilde{\tau}_i, \pi'_{i'}, \bar{\chi}_{i'}$ are mutually disjoint, whence (5.15) follows from (5.16). \square

It is well known [cf. Lauritzen (1996), Proposition 5.2] that (5.12) holds iff $(\Delta^{-1})_{t_0, t_{r+1}} \neq 0$. This holds in turn iff the cofactor of $\Delta_{t_0, t_{r+1}}$, equivalently, the determinant of the submatrix of Δ obtained by deleting its t_0 th row and t_{r+1} th

column, is nonzero. This submatrix is given by

$$\begin{matrix} & t_0 & \sigma_0 & t_1 & \sigma_1 & t_2 & \cdots & t_{r-1} & \sigma_{r-1} & t_r & \sigma_r \\ \begin{matrix} \sigma_0 \\ t_1 \\ \sigma_1 \\ t_2 \\ \vdots \\ t_{r-1} \\ \sigma_{r-1} \\ t_r \\ \sigma_r \\ t_{r+1} \end{matrix} & \left(\begin{matrix} \gamma'_0 & \Delta_0 & \eta_0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \beta_0 & \eta'_0 & \alpha_1 & \gamma_1 & \beta_1 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & \gamma'_1 & \Delta_1 & \eta_1 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & \beta_1 & \eta'_1 & \alpha_2 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & \alpha_{r-1} & \gamma_{r-1} & \beta_{r-1} & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & \gamma'_{r-1} & \Delta_{r-1} & \eta_{r-1} & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & \beta_{r-1} & \eta'_{r-1} & \alpha_r & \gamma_r \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \gamma'_r & \Delta_r \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \beta_r & \eta'_r \end{matrix} \right), \end{matrix}$$

a block-triangular matrix with determinant $\prod_{i=0}^r \det(\Lambda_i)$, where

$$\Lambda_i := \begin{pmatrix} \gamma'_i & \Delta_i \\ \beta_i & \eta'_i \end{pmatrix} : (p_i + 1) \times (1 + p_i),$$

with $p_i := |\sigma_i| \geq 0$. However, because of the explicit representation of X given by (5.9) and (5.10), and by the topology of G^* and G' , we can obtain the following explicit expressions (recall that $\text{Var}(\zeta_v) = 1 \ \forall v \in V$):

$$\begin{aligned} \beta_i &= 1, \\ \gamma_i &= (1, \dots, 1) : 1 \times p_i, \\ \eta'_i &= (q_{i1}, \dots, q_{ip_i}) : 1 \times p_i, \\ \Delta_i &= 1_{p_i} + (q_{i, \min\{e, f\}} \mid e, f = 1, \dots, p_i) : p_i \times p_i, \end{aligned}$$

where, if $p_i \geq 1$, $q_{i1} < \dots < q_{ip_i}$ are the positive integers that satisfy

$$\sigma_i = \{v_{e_i+q_{i1}}, \dots, v_{e_i+q_{ip_i}}\} \quad (\subseteq \tau_i^\circ)$$

and where 1_p denotes the $p \times p$ identity matrix. (In Figure 3, $p_1 = 2$, $q_{11} = 1$, $q_{12} = 2$; $p_2 = 2$, $q_{21} = 1$, $q_{22} = 3$; $p_3 = 1$, $q_{31} = 1$.) Thus, Λ_i has the form

$$\begin{pmatrix} 1 & 1 + q_1 & q_1 & q_1 & \cdots & q_1 & q_1 & q_1 \\ 1 & q_1 & 1 + q_2 & q_2 & \cdots & q_2 & q_2 & q_2 \\ 1 & q_1 & q_2 & 1 + q_3 & \cdots & q_3 & q_3 & q_3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 1 & q_1 & q_2 & q_3 & \cdots & q_{p-2} & 1 + q_{p-1} & q_{p-1} \\ 1 & q_1 & q_2 & q_3 & \cdots & q_{p-2} & q_{p-1} & 1 + q_p \\ 1 & q_1 & q_2 & q_3 & \cdots & q_{p-2} & q_{p-1} & q_p \end{pmatrix},$$

where $q_j \equiv q_{ij}$. Successively subtract the i th row from the $(i + 1)$ th row, $i = p, p - 1, \dots, 1$, preserving the determinant at each step, finally obtaining

an upper-triangular matrix with diagonal entries $1, -1, \dots, -1$, whose determinant is therefore $\pm 1 \neq 0$. Thus, $\det(\Lambda_i) \neq 0$ for $i = 0, \dots, r$, so (5.12) holds and the proof of Theorem 5.1 is complete. \square

6. Strong completeness of the AMP global Markov property. In the proof of Theorem 5.1, the Gaussian probability distribution $P \equiv P_{A,B,S}$ constructed to satisfy $P \in \mathcal{P}_{\text{AMP}}^g(G; \mathbb{R}^V)$ and $A \not\perp\!\!\!\perp B \mid S[P]$ depended on the specified non- p -separated triple A, B, S . In Theorem 6.1 we show that *almost all Gaussian* $P \in \mathcal{P}_{\text{AMP}}^g(G; \mathbb{R}^V)$ are AMP Markov perfect for G , that is, satisfy those and only those CIs specified by the AMP global Markov property $\equiv p$ -separation. This shows that in the Gaussian case, the CG G is a *faithful* representation of the independence–dependence structure of the AMP Markov model that it defines. Our proof is based on the methods of Spirtes, Glymour, and Scheines (1993) and Meek (1995) for strong completeness of ADG models.

DEFINITION 6.1. Let $G \equiv (V, E)$ be a chain graph. A probability distribution P on $\mathbf{X} \equiv \times(\mathbf{X}_v \mid v \in V)$ is AMP Markov perfect for G if $P \in \mathcal{P}_{\text{AMP}}^g(G; \mathbf{X})$ and

$$(6.1) \quad A \not\perp\!\!\!\perp B \mid S[G] \implies A \not\perp\!\!\!\perp B \mid S[P].$$

For any CG $G \equiv (V, E)$ define

$$\begin{aligned} \mathbf{N}(0, V) &:= \{\mathcal{N}(0, \Sigma) \mid \Sigma \in \mathbf{P}(V)\}, \\ \mathbf{N}(0, G) &\equiv \mathbf{N}_{\text{AMP}}(0, G) := \mathbf{N}(0, V) \cap \mathcal{P}_{\text{AMP}}^g(G; \mathbb{R}^V), \\ \mathbf{P}(G) &\equiv \mathbf{P}_{\text{AMP}}(G) := \{\Sigma \mid \mathcal{N}(0, \Sigma) \in \mathbf{N}(0, G)\}, \end{aligned}$$

where $\mathcal{N}(0, \Sigma)$ is the normal distribution with mean 0 and covariance matrix Σ , and for $A \subseteq V$, $\mathbf{P}(A)$ denotes the set of all $A \times A$ positive definite symmetric matrices. It is shown in [AMP] [(2001), Section 5] that

$$(6.2) \quad X \equiv (X_\tau \mid \tau \in \mathcal{T}(G)) \sim \mathcal{N}_V(0, \Sigma) \in \mathbf{N}(0, G),$$

if and only if X can be uniquely represented as a *constrained block-recursive normal linear system*:

$$(6.3) \quad X_\tau = \beta_\tau X_{\text{pa}_{\mathcal{D}}(\tau)} + \varepsilon_\tau, \quad \tau \in \mathcal{T}.$$

Here $\mathcal{D} = \mathcal{D}(G)$, $\mathcal{T} = \mathcal{T}(G)$, $\beta_\tau := \Sigma_{\tau, \text{pa}_{\mathcal{D}}(\tau)} \Sigma_{\text{pa}_{\mathcal{D}}(\tau)}^{-1}$ is the $\tau \times \text{pa}_{\mathcal{D}}(\tau)$ matrix of regression coefficients for X_τ given $X_{\text{pa}_{\mathcal{D}}(\tau)}$, $\varepsilon_\tau \sim \mathcal{N}(0, \Lambda_\tau)$ where $\Lambda_\tau := \Sigma_{\tau, \text{pa}_{\mathcal{D}}(\tau)}$ is the (nonsingular) $\tau \times \tau$ conditional covariance matrix of X_τ given $X_{\text{pa}_{\mathcal{D}}(\tau)}$, the ε_τ are mutually independent, and β_τ and Λ_τ are constrained as follows:

$$(6.4) \quad \beta_\tau \in \mathbf{B}_\tau(G) := \{\beta_\tau \mid u \in \tau, v \in \text{pa}_{\mathcal{D}}(\tau) \setminus \text{pa}_G(u) \Rightarrow (\beta_\tau)_{uv} = 0\},$$

$$(6.5) \quad \Lambda_\tau \in \mathbf{P}(G_\tau) = \{\Lambda_\tau \mid u, v \in \tau, u - v \notin G_\tau \Rightarrow (\Lambda_\tau^{-1})_{uv} = 0\}.$$

Thus, the parameter space of the model $\mathbf{N}(0, G)$ factors into the product of the conditional parameter spaces according to the bijective diffeomorphism

$$\begin{aligned} \lambda_G : \mathbf{P}(G) &\rightarrow \times(\mathbf{B}_\tau(G) \times \mathbf{P}(G_\tau) \mid \tau \in \mathcal{T}) =: \Lambda(G), \\ \Sigma &\mapsto ((\beta_\tau, \Lambda_\tau) \mid \tau \in \mathcal{T}). \end{aligned}$$

The family $((\beta_\tau, \Lambda_\tau) \mid \tau \in \mathcal{T}) \equiv \lambda_G(\Sigma)$ is the family of G -parameters of Σ ; bijectivity implies that these parameters are variation independent.

Next, set $\Psi_\tau := \Lambda_\tau^{-1}$ and $\mathbf{Q}(G_\tau) := (\mathbf{P}(G_\tau))^{-1}$, $\tau \in \mathcal{T}$. The mapping

$$\begin{aligned} \psi_G : \Lambda(G) &\rightarrow \times(\mathbf{B}_\tau(G) \times \mathbf{Q}(G_\tau) \mid \tau \in \mathcal{T}) =: \Psi(G), \\ ((\beta_\tau, \Lambda_\tau) \mid \tau \in \mathcal{T}) &\mapsto ((\beta_\tau, \Psi_\tau) \mid \tau \in \mathcal{T}) \end{aligned}$$

is also a bijective diffeomorphism, hence the composition mapping

$$\begin{aligned} \psi_G \circ \lambda_G : \mathbf{P}(G) &\rightarrow \Psi(G), \\ \Sigma &\mapsto ((\beta_\tau, \Psi_\tau) \mid \tau \in \mathcal{T}) \end{aligned}$$

is itself a bijective diffeomorphism and determines an alternative parameterization for $\mathbf{N}(0, G)$. The family $((\beta_\tau, \Psi_\tau) \mid \tau \in \mathcal{T}) \equiv (\psi_G \circ \lambda_G)(\Sigma)$ is called the family of *inverse G -parameters* of Σ ; bijectivity implies that these are also variation independent. Because

$$\mathbf{Q}(G_\tau) = \{\Psi_\tau \in \mathbf{P}(\tau) \mid u, v \in \tau, \ u - v \notin G_\tau \Rightarrow (\Psi_\tau)_{uv} = 0\},$$

$\mathbf{Q}(G_\tau)$ is a relatively open subcone of the open convex cone $\mathbf{P}(\tau)$, so the alternative parameter space $\Psi(G)$ has a simple form: it is a Cartesian product of real Euclidean vector spaces $\mathbf{B}_\tau(G)$ and open convex cones $\mathbf{Q}(G_\tau)$, hence admits a Lebesgue measure μ .

THEOREM 6.1 (Strong completeness of normal AMP models). *Let $G \equiv (V, E)$ be a chain graph. Then almost every $P \in \mathbf{N}(0, G)$ is AMP Markov perfect for G ; that is,*

$$(6.6) \qquad \mu((\psi_G \circ \lambda_G)(\mathbf{P}(G) \setminus \mathbf{M})) = 0,$$

where $\mathbf{M} := \{\Sigma \in \mathbf{P}(G) \mid \mathcal{N}(0, \Sigma) \text{ is AMP Markov perfect for } G\}$.

PROOF. Define $T(G)$ to be the set of all disjoint triples (A, B, S) of subsets of V such that $A, B \neq \emptyset$ and $A \not\propto B \mid S[G]$. In the proof of Theorem 5.1 it was shown that if $(A, B, S) \in T(G)$ then there exist $a \equiv a(A, B, S) \in A$, $b \equiv b(A, B, S) \in B$, and $\Sigma_{A,B,S} \in \mathbf{P}(G)$ such that $a \not\perp b \mid S[P]$, or equivalently, such that $C_{a,b,S}(\Sigma_{A,B,S}) \neq 0$, where $C_{a,b,S}(\Sigma)$ denotes the cofactor of the (a, b) -element in the (a, b, S) -submatrix of Σ . [Take a, b to be the endpoints of the trail π satisfying the minimality condition M2 and $\Sigma_{A,B,S}$ to be the covariance

matrix of the normal distribution $P_{A,B,S}$ determined by (5.9) and (5.10).] By Definition 6.1 and normality,

$$\begin{aligned} \mathbf{P}(G) \setminus \mathbf{M} &= \bigcup_{\substack{(A,B,S) \\ \in T(G)}} \bigcap_{\substack{a' \in A \\ b' \in B}} \{\Sigma \in \mathbf{P}(G) \mid C_{a',b',S}(\Sigma) = 0\} \\ &\subseteq \bigcup_{\substack{(A,B,S) \\ \in T(G)}} \{\Sigma \in \mathbf{P}(G) \mid C_{a,b,S}(\Sigma) = 0\}, \end{aligned}$$

so

$$(6.7) \quad \mu((\psi_G \circ \lambda_G)(\mathbf{P}(G) \setminus \mathbf{M})) \leq \sum_{\substack{(A,B,S) \\ \in T(G)}} \mu(\{\psi \in \Psi_G \mid C_{a,b,S}(\lambda_G^{-1}(\psi_G^{-1}(\psi))) = 0\}).$$

The cofactor $C_{a,b,S}(\cdot)$ is a determinant, hence is a polynomial function of its arguments. The mapping $\lambda_G^{-1}: \Lambda(G) \rightarrow \mathbf{P}(G)$ is given by the *reconstruction algorithm* of Andersson and Perlman [(1998), Section 5]; examination of this algorithm shows that each component of $\lambda_G^{-1}(\cdot)$ is a polynomial function of its arguments. The mapping

$$\begin{aligned} \psi_G^{-1}: \Psi(G) &\rightarrow \Lambda(G), \\ ((\beta_\tau, \Psi_\tau) \mid \tau \in \mathcal{T}) &\mapsto ((\beta_\tau, \Psi_\tau^{-1}) \mid \tau \in \mathcal{T}) \end{aligned}$$

has both linear and nonlinear components, the latter given by matrix inversion. Because each element of Ψ_τ^{-1} is the quotient of the corresponding cofactor in Ψ_τ (a polynomial in the elements of Ψ_τ) and $\det(\Psi_\tau)$, each component of $\psi_G^{-1}(\psi)$ is either a linear function of β_τ for some $\tau \in \mathcal{T}$ or else is such a quotient for some $\tau \in \mathcal{T}$. Therefore $C_{a,b,S}(\lambda_G^{-1}(\psi_G^{-1}(\psi)))$ can be expressed as a *rational function* $p_{a,b,S}(\psi)/q_{a,b,S}(\psi)$ of $\psi \in \Psi(G)$, where $p_{a,b,S}(\psi)$ is a polynomial in the elements of ψ and $q_{a,b,S}(\psi)$ is a product of powers of $\det(\Psi_\tau^{-1})$, $\tau \in \mathcal{T}$. Thus $q_{a,b,S}(\psi)$ does not vanish on $\Psi(G)$, so

$$(6.8) \quad \{\psi \in \Psi_G \mid C_{a,b,S}(\lambda_G^{-1}(\psi_G^{-1}(\psi))) = 0\} = \{\psi \in \Psi_G \mid p_{a,b,S}(\psi) = 0\}.$$

Because $C_{a,b,S}(\Sigma_{A,B,S}) \neq 0$, the polynomial $p_{a,b,S}$ is not identically zero, so by Okamoto's lemma [an application of Fubini's theorem; see Okamoto (1973)],

$$(6.9) \quad \mu(\{\psi \in \Psi_G \mid p_{a,b,S}(\psi) = 0\}) = 0.$$

Now (6.6) follows from (6.7), (6.8) and (6.9). \square

EXAMPLE 6.1. Let G be the chain graph $1 \rightarrow 2 - 3$ (a single flag). Here $\mathcal{T}(G) = \{\{1\}, \{2, 3\}\}$ and $\mathcal{D}(G)$ is the ADG $\{1\} \rightarrow \{2, 3\}$, so by (6.3) – (6.5) every distribution in $\mathbf{N}(0, G)$ has the linear representation

$$\begin{aligned} X_1 &= \varepsilon_1, \\ X_2 &= \beta_{21}\varepsilon_1 + \varepsilon_2, \\ X_3 &= \varepsilon_3, \end{aligned}$$

where $\varepsilon_1 \perp\!\!\!\perp (\varepsilon_2, \varepsilon_3)$, $\varepsilon_1 \sim \mathcal{N}(0, \Lambda_{\{1\}})$, and $(\varepsilon_2, \varepsilon_3) \sim \mathcal{N}((0, 0), \Lambda_{\{2,3\}})$. The covariance matrix Σ of (X_1, X_2, X_3) is given by

$$(6.10) \quad \Sigma = \begin{pmatrix} \lambda_{11} & \beta_{21}\lambda_{11} & 0 \\ \beta_{21}\lambda_{11} & \beta_{21}^2\lambda_{11} + \lambda_{22} & \lambda_{23} \\ 0 & \lambda_{32} & \lambda_{33} \end{pmatrix}$$

so $1 \perp\!\!\!\perp 3$, the sole CI specified by the AMP global Markov property for G . The G -parameters of Σ are

$$\begin{aligned} \Lambda_{\{1\}} &\equiv \lambda_{11} \in \mathbb{R}^+, & \beta_{\{2,3\}} &\equiv \beta_{21} \in \mathbb{R}, \\ \Lambda_{\{2,3\}} &\equiv \begin{pmatrix} \lambda_{22} & \lambda_{23} \\ \lambda_{32} & \lambda_{33} \end{pmatrix} \in \mathbf{P}(\{2, 3\}), \end{aligned}$$

and are variation independent. The inverse G -parameters are

$$\begin{aligned} \Psi_{\{1\}} &\equiv \psi_{11} = \lambda_{11}^{-1} \in \mathbb{R}^+, \\ \beta_{\{2,3\}} &\equiv \beta_{21} \in \mathbb{R}, \\ \Psi_{\{2,3\}} &\equiv \begin{pmatrix} \psi_{22} & \psi_{23} \\ \psi_{32} & \psi_{33} \end{pmatrix} = \Lambda_{\{2,3\}}^{-1} \in \mathbf{P}(\{2, 3\}), \end{aligned}$$

and are also variation independent. By considering the appropriate elements and cofactors of Σ in (6.10), it is readily seen that (X_1, X_2, X_3) satisfies no CI other than $1 \perp\!\!\!\perp 3$ unless $\beta_{21} = 0$ or $\psi_{23} \equiv \psi_{32} = 0$. These exceptions determine a Lebesgue-null set in the space of inverse G -parameters, so almost every $P \in \mathbf{N}(0, G)$ is AMP Markov perfect for G . Theorem 6.1 shows that this is true for all CGs. \square

7. A linear time p -separation algorithm. The algorithm presented in this section combines and extends Algorithms 1 and 2 in Section 5 of Geiger, Verma, and Pearl (1990) for identifying conditional independences entailed in ADG models. Only a few changes are needed to adapt these algorithms to AMP chain graphs. Moreover, the linear time $O(\max\{|V|, |E|\})$ complexity of the original algorithms remains unchanged for identifying CIs in AMP CGs.

Let $G \equiv (V, E)$ be a CG, $G^\vee \equiv (V, E^\vee)$ its underlying undirected graph, and A, S disjoint subsets of V with $A \neq \emptyset$. For $v \in V$ define

$$(7.1) \quad a(v) := \begin{cases} 1, & \text{if } v \in S \text{ and } \text{pa}_G(v) \setminus S \neq \emptyset, \\ 0, & \text{otherwise,} \end{cases}$$

$$(7.2) \quad d(v) := \begin{cases} 1, & \text{if } v \in \text{An}_G(S), \\ 0, & \text{otherwise;} \end{cases}$$

$a(v)$ will indicate whether or not v is S -activated relative to a specified trail.

DEFINITION 7.1. A linked pair $((u, v), (v, w)) \in E^\vee \times E^\vee$ is **legal** if $u \neq w$ and either:

- (i) v is a head-no-tail node in the trail $u \cdots v \cdots w$ in G and $d(v) = 1$, or

- (ii) v is not a head-no-tail node in the trail $u \cdots v \cdots w$ in G and $v \notin S$, or
- (iii) $u \cdots v \cdots w = u - v - w$ and $a(v) = 1$.

If neither (i), (ii) nor (iii) hold, $((u, v), (v, w)) \in E^\vee \times E^\vee$ is **illegal**.

A path (v_0, \dots, v_n) in G^\vee is *legal* if each consecutive linked pair $((v_{i-1}, v_i), (v_i, v_{i+1}))$ is legal. A vertex $b \in V \setminus (A \dot{\cup} S)$ is *reachable* from $a \in A$ if it lies on a legal path in G^\vee emanating from a , equivalently, if it lies on an S -active trail in G emanating from a . Algorithm I finds all vertices reachable from A .

The proof of the following proposition is similar to those of Lemma 10 and Theorem 13 in Geiger, Verma and Pearl (1990) and therefore omitted.

PROPOSITION 7.1. *Algorithm I is valid: the set B returned is exactly $\{b \in V \setminus (A \dot{\cup} S) \mid A \bowtie_p b \mid S[G]\}$.*

Thus, by Theorems 4.1 and 5.1, Algorithm I finds the maximal B such that $A \perp\!\!\!\perp B \mid S[P]$ for every AMP G -Markovian P .

Algorithm I. A linear-time algorithm for determining p -separation in a chain graph.

Input. A chain graph $G \equiv (V, E)$; disjoint subsets $A, S \subseteq V$ such that $A \neq \emptyset$.

Data Structure. For each $v \in V$, an in-list $IL(v)$ of all arrows pointing to v in G .

Output. The maximal set $B \subseteq V \setminus (A \dot{\cup} S)$ such that $A \bowtie_p B \mid S[G]$.

Procedure.

- (1) Construct $(a(v) \mid v \in V)$ and $(d(v) \mid v \in V)$ according to (7.1) and (7.2).
- (2) Add a *new* vertex x to V ; set $R := \{x\}$; set $i := 1$.
- (3) **For each** $a \in A$: **add** (a, x) to E and **label** it 1; **set** $R := R \dot{\cup} \{a\}$. (Note: now $G := (V \cup \{x\}, E \cup \{(a, x) \mid a \in A\})$.)
- (4) Construct $E^\vee := \{(v, w) \mid (v, w) \in E \vee (w, v) \in E\}$.
- (5) **Find all unlabeled** $(v, w) \in E^\vee$ **adjacent** to at **least** one (u, v) labeled i , such that $w \notin R$ and $((u, v), (v, w))$ is legal; label each such (v, w) with $i + 1$; set $R := R \cup \{w\}$. If no such (v, w) exist, **go to** (7).
- (6) Set $i := i + 1$; **go to** (5).
- (7) **Return** $B := V \setminus (R \cup A \cup S)$.

Algorithm I is a variant of Breadth First Search. As noted by Geiger, Verma and Pearl (1990), the complexity of Algorithm I is $O(\max\{|V|, |E|\})$ for ADG models when legality is determined by d -separation. By a similar argument, the complexity of Algorithm I is also $O(\max\{|V|, |E|\})$ when legality is determined by p -separation for AMP models. Some details are now provided.

To determine $(a(v)|v \in V)$ in Step (1), assign $a(s) = 1$ for each $s \in S$, then reassign $a(s) = 0$ if $\text{IL}(s) \setminus S = \emptyset$. To determine $(d(v)|v \in V)$, use the in-lists to determine all parents of S , then all parents of parents, and so on, assigning $d(v) = 1$ to each vertex so encountered. Because each arrow of G is encountered at most once in each of these two determinations, Step (1) requires $O(|E|)$ operations. Step (2) is trivial and requires only constant time. Step (3) requires at most $O(|V|)$ operations. Step (4) converts each edge in G to a line, which requires $O(|E|)$ operations.

For step (5), we now show that for each $(u, v) \in E^\vee$ labeled i , the decisions concerning the labeling of all unlabeled adjacent $(v, w) \in E^\vee$ require only constant time in toto, hence step (5) requires at most $O(|E|)$ operations. If $v \in \text{IL}(u)$ (so $u \leftarrow v \in G$) and $v \in S$, all linked pairs $((u, v), (v, w))$ are illegal [so all $(v, w) \in E^\vee$ remain unlabeled]; if $v \in \text{IL}(u)$ and $v \notin S$, all $((u, v), (v, w))$ are legal (so all $(v, w) \in E^\vee$ are labeled $i + 1$). If $u \in \text{IL}(v)$ (so $u \rightarrow v \in G$) and $v \in \text{IL}(w)$, then $((u, v), (v, w))$ is legal iff $v \notin S$; if $u \in \text{IL}(v)$ and $v \notin \text{IL}(w)$, then $((u, v), (v, w))$ is legal iff $d(v) = 1$. If $v \notin \text{IL}(u)$ and $u \notin \text{IL}(v)$ (so $u - v \in G$), then: if $w \in \text{IL}(v)$ then $((u, v), (v, w))$ is legal iff $d(v) = 1$; if $v \in \text{IL}(w)$ then $((u, v), (v, w))$ is legal iff $v \notin S$; if $w \notin \text{IL}(v)$ and $v \notin \text{IL}(w)$, then $((u, v), (v, w))$ is legal iff $v \notin S$ or $a(v) = 1$.

Steps (6) and (7) are trivial, requiring only constant time. Thus, the complexity of Algorithm I is at most $O(\max\{|V|, |E|\})$.

If Algorithm I is to be applied to several subsets A_1, \dots, A_k of V for the same separating subset S , then the determination of the set of legal linked pairs $((u, v), (v, w)) \in E^\vee \times E^\vee$ according to Definition 7.1 can be done first, not repeated for each A_i .

APPENDIX

Graph-theoretic terminology. A graph G is a pair (V, E) , where V is a finite set of *vertices* and $E \subseteq \{(v, w) \in V \times V \mid v \neq w\}$ is a set of *edges*, that is, a set of ordered pairs of distinct vertices. An edge $(v, w) \in E$ whose opposite $(w, v) \in E$ is called an *undirected* edge and appears as a line $v - w$ in our figures; in the text we write $v - w \in G$. An edge $(v, w) \in E$ whose opposite $(w, v) \notin E$, is called a *directed* edge and appears as an arrow $v \rightarrow w$ in our figures; in the text we write $v \rightarrow w \in G$. A graph with only undirected edges is called an *undirected graph* (UG). A graph with only directed edges is a *directed graph* (\equiv *digraph*). Only simple graphs are considered, that is, graphs without loops or multiple edges between any pair of vertices.

Two vertices $v, w \in V$ are *adjacent* in G , written as $v \cdots w \in G$, if $(v, w) \in E$ or $(w, v) \in E$ or both. A graph is *complete* if all vertices are adjacent. A graph

$G' \equiv (V', E')$ is a *subgraph* of $G \equiv (V, E)$, denoted by $G' \sqsubseteq G$, if $V' \subseteq V$, $v \rightarrow w \in G' \Rightarrow v \rightarrow w \in G$, and $v - w \in G' \Rightarrow v - w \in G$. A subset $A \subseteq V$ *induces* the subgraph $G_A := (A, E_A)$, where $E_A := E \cap (A \times A)$; that is, E_A is obtained from E by retaining all edges with both endpoints in A .

A graph $G = (V, E)$ determines two UGs $G^\vee \equiv (V, E^\vee)$, $G^\wedge \equiv (V, E^\wedge)$, where

$$E^\vee := \{(v, w) \mid (v, w) \in E \vee (w, v) \in E\},$$

$$E^\wedge := \{(v, w) \mid (v, w) \in E \wedge (w, v) \in E\},$$

respectively. Thus, G^\vee is the *skeleton* of G , that is, the underlying UG obtained by converting all arrows of G into lines, while G^\wedge is obtained by deleting all arrows of G , so $G^\wedge \sqsubseteq G$.

A *path* π of length $n \geq 1$ from a to b in G is a sequence of distinct vertices ($a \equiv v_0, \dots, v_n \equiv b$) such that $(v_{i-1}, v_i) \in E$ for all $i = 1, \dots, n$. An n -*cycle* is a path of length $n \geq 3$ such that $v_0 = v_n$. A path or cycle is undirected if $v_{i-1} - v_i \in G$ for all $i = 1, \dots, n$. A path or cycle is directed or semidirected if $v_{i-1} \rightarrow v_i \in G$ for, respectively, all or at least one of $i = 1, \dots, n$, and $v_{i-1} \leftarrow v_i \notin G$ for all $i = 1, \dots, n$. A directed graph with no directed cycles is an *acyclic digraph* (ADG). A *chain graph* (CG) is an *adicyclic* graph, that is, contains no semidirected cycles. UGs and ADGs are special cases of CGs.

A *trail* π of length $n \geq 1$ between distinct vertices $a, b \in V$ in G is a sequence ($a \equiv v_0, \dots, v_n \equiv b$) of $n + 1$ *distinct* vertices such that each consecutive pair (v_{i-1}, v_i) is adjacent in G . Equivalently, a trail corresponds to a path in the UG G^\vee . A vertex w is an *interior vertex* of π if $w = v_i$ for some $i = 1, \dots, n - 1$, equivalently, if $w \in \pi^\circ := \pi \setminus \{a, b\}$. We say that π is a trail (or path) *between* A and B in G if $a \in A$, $b \in B$, and $\pi^\circ \cap (A \cup B) = \emptyset$.

A UG G is *connected* if, for every distinct $v, w \in V$, there is a path between v and w in G . A subset $A \subseteq V$ is *connected* in G if G_A is connected. The maximal connected subsets are called the *connected components* of G , and V can be uniquely partitioned into the disjoint union of the connected components of G . For pairwise disjoint subsets $A (\neq \emptyset)$, $B (\neq \emptyset)$, and S of V , A and B are *separated* by S in the UG G if all paths in G between A and B intersect S . Note that if $S = \emptyset$, then A and B are separated by S in G if and only if there are no paths connecting A and B in G .

If $u \rightarrow v \in G$, then u is a *parent* of v . The set of parents of v in G is denoted by $\text{pa}(v) \equiv \text{pa}_G(v)$. For any $A \subseteq V$, $\text{de}(A) := \text{de}_G(A)$ denotes the *descendants* of A in G , that is, the set of all $v \in G$ such that there is a directed path from A to v , and $\text{De}(A) := A \cup \text{de}(A)$.

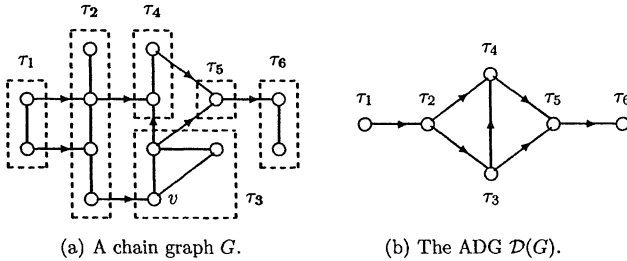
Hereafter, let $G \equiv (V, E)$ be a CG. A subset $A \subseteq V$ is called *G-anterior* if $v \in A$ whenever there is a path from v to some $a \in A$. For any subset $A \subseteq V$, $\text{At}(A) \equiv \text{At}_G(A) :=$ the smallest *G-anterior* set containing A . A subset $A \subseteq V$ is called *G-ancestral* if $v \in A$ whenever there is a directed path from v to some $a \in A$. For any subset $A \subseteq V$, $\text{An}(A) \equiv \text{An}_G(A) :=$ the smallest *G-ancestral* set containing A , and $\text{an}(A) := \text{An}(A) \setminus A$.

Let $\mathcal{T} \equiv \mathcal{T}(G)$ denote the set of *chain components* of G , that is, the connected components of G^\wedge . The chain components of G , when regarded as a set of vertices, form a graph $\mathcal{D}(G) \equiv (\mathcal{T}(G), \mathcal{E}(G))$ in which $\mathcal{E}(G)$ is defined as

$$(7.3) \quad \mathcal{E}(G) := \{(\tau, \tau') \in \mathcal{T} \times \mathcal{T} \mid \tau \neq \tau', \exists v \in \tau, v' \in \tau' \ni v \rightarrow v' \in G\}.$$

It follows from the acyclicity of G that $\mathcal{D}(G)$ is an ADG.

An example of a chain graph appears below. In (a), dashed lines indicate the set of chain components $\mathcal{T}(G) = \{\tau_1, \tau_2, \dots, \tau_6\}$ of G . In (b), the chain components are the vertices of the ADG $\mathcal{D}(G)$.



A subset $A \subseteq V$ is *G-coherent* if A is a union of chain components of G . For $A \subseteq V$, define $\text{Co}(A) \equiv \text{Co}_G(A) :=$ the smallest G -coherent set containing A . For $A \subseteq V$, the *extended subgraph* $G[A]$ is defined by

$$G[A] := G_{\text{An}(A)} \cup G_{\text{Co}(\text{An}(A))}^\wedge.$$

Note that $G[A] \subseteq G$; that is, $G[A]$ is in fact a subgraph of G , and that a *directed edge* occurs in $G[A]$ iff it occurs in $G_{\text{An}(A)}$.

Acknowledgments. We wish to thank Steen Andersson, Babette Brumback, Michael Eichler, Dan Geiger, Jan Koster, Steffen Lauritzen, Thomas Richardson, Anthony Rossini, Milan Studený, the Editors and two referees for helpful comments and suggestions.

REFERENCES

- ANDERSSON, S. A., MADIGAN, D. and PERLMAN, M. D. (1996). An alternative Markov property for chain graphs. In *Uncertainty in Artificial Intelligence: Proceedings of the Twelfth Conference* (F. Jensen and E. Horvitz, eds.) 40–48. Morgan Kaufmann, San Francisco.
- ANDERSSON, S. A., MADIGAN, D. and PERLMAN, M. D. (1997a). On the Markov equivalence of chain graphs, undirected graphs and acyclic digraphs. *Scand. J. Statist.* **24** 81–102.
- ANDERSSON, S. A., MADIGAN, D. and PERLMAN, M. D. (1997b). A characterization of Markov equivalence classes for acyclic digraphs. *Ann. Statist.* **25** 505–541.
- ANDERSSON, S. A., MADIGAN, D. and PERLMAN, M. D. (2001). Alternative Markov properties for chain graphs. *Scand. J. Statist.* **28** 33–85.
- ANDERSSON, S. A. and PERLMAN, M. D. (1998). Normal linear regression models with recursive graphical Markov structure. *J. Multivariate Anal.* **66** 133–187.
- BOUCKAERT, R. R. and STUDENÝ, M. (1995). Chain graphs: semantics and expressiveness. In *Symbolic and Quantitative Approaches to Reasoning and Uncertainty. Lecture Notes in AI* **946** 67–76. Springer, New York.

- BUNTINE, W. L. (1995). Chain graphs for learning. In *Uncertainty in Artificial Intelligence: Proceedings of the Eleventh Conference* (P. Besnard and S. Hanks, eds.) 87–98. Morgan Kaufmann, San Francisco.
- COWELL, R. G., DAWID, A. P., LAURITZEN, S. L. and SPIEGELHALTER, D. J. (1999). *Probabilistic Networks and Expert Systems*. Springer, New York.
- COX, D. R. (1999). Graphical Markov models: statistical motivation. Paper presented at the Workshop on Conditional Independence Structures and Graphical Models, Fields Institute for Research in Mathematical Sciences, Toronto, Canada.
- COX, D. R. and WERMUTH, N. (1996). *Multivariate Dependencies: Models, Analysis and Interpretation*. Chapman and Hall, London.
- FRYDENBERG, M. (1990a). The chain graph Markov property. *Scand. J. Statist.* **17** 333–353.
- FRYDENBERG, M. (1990b). Marginalization and collapsibility in graphical interaction models. *Ann. Statist.* **18** 790–805.
- GEIGER, D. and PEARL, J. (1988). On the logic of causal models. In *Proceedings of the Fourth Workshop on Uncertainty in AI* (R. Shachter, T. Levitt, L. Kanal and J. Lemmer, eds.) 136–147. North-Holland, Amsterdam.
- GEIGER, D., VERMA, T. and PEARL, J. (1990). Identifying independence in Bayesian networks. *Networks* **20** 507–534.
- HECKERMAN, D., GEIGER, D. and CHICKERING, D. M. (1995). Learning Bayesian networks: the combination of knowledge and statistical data. *Machine Learning* **20** 197–243.
- HØJSGAARD, S. and THIESSEN, B. (1995). BIFROST: Block recursive models induced from relevant knowledge, observations, and statistical techniques. *Comput. Statist. Data Anal.* **19** 155–175.
- KOSTER, J. T. A. (1999). *Linear Structural Equations and Graphical Models*. The Fields Institute, Toronto, Canada.
- LAURITZEN, S. L. (1996). *Graphical Models*. Oxford Univ. Press.
- LAURITZEN, S. L., DAWID, A. P., LARSEN, B. N. and LEIMER, H. G. (1990). Independence properties of directed Markov fields. *Networks* **20** 491–505.
- LAURITZEN, S. L. and WERMUTH, N. (1989). Graphical models for association between variables, some of which are qualitative and some quantitative. *Ann. Statist.* **17** 31–57.
- MADIGAN, D., ANDERSSON, S. A., PERLMAN, M. D. and VOLINSKY, C. M. (1996). Bayesian model averaging and model selection for Markov equivalence classes of acyclic diagraphs. *Comm. Statist. Theory Methods Ser. A* **25** 2493–2519.
- MEEK, C. (1995). Strong completeness and faithfulness in Bayesian networks. In *Uncertainty in Artificial Intelligence: Proceedings of the Eleventh Conference* (P. Besnard and S. Hanks, eds.) 403–410. Morgan Kaufmann, San Mateo, CA.
- OKAMOTO, M. (1973). Distinctness of the eigenvalues of a quadratic form in a multivariate sample. *Ann. Statist.* **1** 763–765.
- PEARL, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann, San Mateo, CA.
- SPIEGELHALTER, D. J., DAWID, A. P., LAURITZEN, S. L. and COWELL, R. G. (1993). Bayesian analysis in expert systems (with discussion). *Statist. Sci.* **8** 219–283.
- SPIRITES, P., GLYMOUR, C. and SCHEINES, R. (1993). *Causation, Prediction, and Search. Lecture Notes in Statist.* **81** Springer, New York.
- STUDENÝ, M. (1996). On separation criterion and recovery algorithm for chain graphs. In *Uncertainty in Artificial Intelligence: Proceedings of the Twelfth Conference* (F. Jensen and E. Horvitz, eds.) 509–516. Morgan Kaufmann, San Francisco.
- STUDENÝ, M. (1997). A recovery algorithm for chain graphs. *Internat. J. Approx. Reason.* **17** 265–293.
- STUDENÝ, M. (1998). Bayesian networks from the point of view of chain graphs. In *Uncertainty in Artificial Intelligence: Proceedings of the Fourteenth Conference* (G. F. Cooper and S. Moral, eds.) 496–503. Morgan Kaufmann, San Francisco.
- STUDENÝ, M. and BOUCKAERT, R. R. (1998). On chain graph models for description of conditional independence structure. *Ann. Statist.* **26** 1434–1495.

WERMUTH, N. and LAURITZEN, S. L. (1990). On substantive research hypotheses, conditional independence graphs, and graphical chain models (with discussion). *J. Roy. Statist. Soc. Ser. B* **52** 21–72.

WHITTAKER, J. (1990). *Graphical Models in Applied Multivariate Statistics*. Wiley, New York.

M. LEVITZ
DEPARTMENT OF BIOSTATISTICS
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON, 98195-7232
E-mail: levitz@biostat.washington.edu

M. D. PERLMAN
DEPARTMENT OF STATISTICS
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON, 98195-4322
E-mail: michael@stat.washington.edu

D. MADIGAN
DEPARTMENT OF STATISTICS
RUTGERS UNIVERSITY
PISCATAWAY, NEW JERSEY 08855
E-mail: madigan@stat.rutgers.edu