Proceedings of Machine Learning Research 323:1–56, 2026 5th Conference on Causal Learning and Reasoning

Cross-Sectional Conditional Independence in Stationary Time Series: Graphical Equivalence and Completeness of Collider Separation

Hubert Drazkowski HUBERT.DRAZKOWSKI@DI.KU.DK
University of Copenhagen

Editors: Bijan Mazaheri and Niels Richard Hansen

Abstract

We study cross-sectional conditional independence (CI) in strictly stationary time series represented by summary (local-dependence) graphs. Our focus is on the completeness of separation rules, i.e., when—and in what sense—graphical separation on the summary graph is not only sufficient, but also necessary for CI at a time slice. Prior work has established soundness of collider-based c-separation in discrete time and trek-graph separation in continuous time with respect to cross-sectional CI. We complement these results twofold. First, we show that, on summary graphs, c-separation and trek-graph separation induce the same equivalence classes of graphs. This identifies a single underlying graphical notion governing cross-sectional CI and reconciles discrete- and continuous-time separation rules. In addition, we translate c-separation to space-time graphs and introduce space-time trek separation completing the graphical separation picture. In essence, it is the collider structure that governs what shared past information travels to cross section through common ancestors and governs CI. Second, we prove the completeness of the collider separation for the discrete-time case. For any given summary graph, there exists a strictly stationary process whose cross-sectional law realizes all and only the c-separation statements. Moreover, within the family of stationary Gaussian VAR(1) models, in the absence of c-separation, parameter values that nonetheless yield independence form a Lebesgue measure-zero set. Together, these results sharpen the theoretical underpinnings of graphical reasoning about cross-sectional CI in stationary time series and clarify when faithfulness-type assumptions for causal discovery are generically justified for snapshots of dynamic Bayesian networks.

Keywords: Summary graphs, Dynamic Bayesian networks, Cross section data, Faithfulness, Time series, Graphical separation

1. Introduction

Stochastic processes are the natural language for systems that evolve in time, yet in many applications we do not observe trajectories—we observe *snapshots*. In some applications, it is difficult—or impossible—to observe a dynamical system more than once. In biology, measurement may destroy the observed collection of cells, and in social sciences, we routinely observe agents through cross-sectional data collection. Crucially, a snapshot—a sample from a single time slice—is the marginal of a dynamical system. Past common causes induce contemporaneous dependencies that have to be adjusted for when studying *conditional independence* (CI).

To reason about the dynamical system, we represent it with a *summary (local-dependence) graph* (Niemiro and Rajkowski, 2023). Summary graphs encode sparsity relations between variables and allow asymmetric relations between them, so their graphical structure also carries a causal interpretation (e.g., Peters et al., 2017). They provide a compact representation of the stochastic process and simplify the analysis of complex systems. Crucially, snapshots should not be analyzed with d/m-separation on the summary graph, because they do not characterize cross-sectional independences. They are neither *sound* ("separation implies CI"), nor *complete* ("CI implies separation", in some sense)

Understanding CI from summary graphs matters thrice over for causality. First, for causal inference, ignoring the time dimension and analysing snapshots as if they were from a static distribution can invalidate the standard exogeneity assumptions (e.g., Thams et al., 2024). Second, for reasoning, CI tells us which modelling constraints or control objectives are justified given a known underlying graph of the data generating process (e.g., Meek, 1995). Third, for discovery, constraint-based methods attempt to recover the structure from many CI tests, and their reliability hinges on when the graph's separation rule truly forces, or cannot force independence (e.g., Vowels et al., 2022; Assaad et al., 2022). These methods typically assume *faithfulness*—graphical separations coincide with probabilistic independences—an untestable assumption that is, nevertheless, supported by *typicality* in i.i.d. settings (Boeken et al., 2024).

In this paper, we study the cross-sectional analog of typicality for *dynamic Bayesian networks* in discrete time and focus on the completeness (necessity) of the separation criteria. To work with cross-sectional CI for a stochastic process, we assume temporal stability in the form of strict stationarity. Stationarity turns the cross-sectional problem into a well-posed target distribution (the marginal at any t). In this setting, the *trek-graph separation* is sound and at least *weakly complete* (see below) for cross-sectional CI diffusion models (Boege et al., 2025). In practice, however, workhorses for inference, simulations, and control are often discrete-time models such as *vector autoregressions* (VARs). There is no simple translation between the two regimes: sampling a diffusion process in discrete time induces a different distribution than that of the diffusion (e.g., Aït-Sahalia and Mykland, 2003). In the discrete-time setting, Niemiro and Rajkowski (2023) introduced a collider-based separation—*c-separation*—and proved soundness for cross-sectional CI, together with a positive conjecture about its completeness.

Main results We address two questions about cross-sectional conditional independence in *strictly stationary* time series represented by summary graphs:

1. *Graphical equivalence.* Do trek-based and collider-based separations give the same semantics for cross-sectional CI?
   We show that for any summary graph, the equivalence class induced by *c-separation* (Niemiro and Rajkowski, 2023) coincides with the equivalence class induced by *trek-graph separation* (Boege et al., 2025). This reconciles the discrete-time and diffusion-based criteria and identifies a single underlying graph-theoretic notion for cross-sectional CI. We then translate c-separation further to the unrolled space-time DAG by introducing *space-time trek separation*. Together, these results identify a single graphical notion governing cross-sectional CI and make explicit how the collider structure on the summary graph translates into information flow in space time. (Section 3)

2. *Completeness.* Is c-separation a necessary condition for conditional independence?
   We prove *completeness* resolving the conjecture of Niemiro and Rajkowski (2023). For each summary graph, there exists a strictly stationary process on that graph whose snapshot distribution obeys all c-separations and exhibits conditional dependence for every triple that is not c-separated. Within the family of stable Gaussian VAR(1) models Markov to a given summary graph if c-separation fails, the parameter values that nonetheless yield conditional independence form a Lebesgue measure-zero subset of the stable parameter space. These results

sharpen cross-sectional CI reasoning for stationary time series and clarify when faithfulness-type assumptions are generically justified for dynamic Bayesian networks. (Section 4)

In addition, in Section 2 we review summary (local-dependence) graphs, c-separation and trek-graph separation rules, and the stationary Gaussian VAR(1) framework for cross-sectional CI. The appendices collect background on notions mentioned in the main text, a more in-depth discussion, some examples, and detailed proofs. Discussion of the assumptions is in Appendix A; background on separation properties, Gaussian VARs are in Appendix B; equivalence proofs are in Appendix C; all proofs of constructions for completeness are in Appendix D.

2. Setting

Summary graphs Let $X = \{X_v(t) : v \in \mathcal{V},\ t \in \mathcal{T} \subset \mathbb{Z}\}$ be a multivariate discrete-time stochastic process, with components indexed by a finite index set $\mathcal{V}$. For each $v \in \mathcal{V}$, $X_v(t)$ takes values in a measurable space $(\mathcal{X}_v, \mathscr{X}_v)$. For $A \subseteq \mathcal{V}$ define $X_A(t) := (X_v(t))_{v\in A}$ and, for $s < t$, let $X_A(< t) := \{X_A(u) : u \in \mathcal{T},\ u < t\}$.

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a directed graph that may contain self-loops and reciprocal pairs of arrows. For $v \in \mathcal{V}$, write $\mathrm{pa}_{\mathcal{G}}(v) := \{u \in \mathcal{V} : u \to v \in \mathcal{E}\}$ for the parent set of v. In our notation it is possible that $v \in \mathrm{pa}_{\mathcal{G}}(v)$, meaning that there can be autoregression, expressed by a *self-loop* in the summary graph. We call $\mathcal{G}$ a summary (local-dependence) graph for X when the law of X factorizes with respect to $\mathcal{G}$ as follows. For measurable rectangles $E = \prod_v E_v$ and $F = \prod_v F_v$ with $E_v, F_v \in \mathscr{X}_v$, and for $t > t_0$, where $t_0 = \min \mathcal{T}$

$$\mathbb{P}\big(X(t) \in F \mid X(< t) \in E\big) = \prod_{v\in\mathcal{V}} \mathbb{P}\Big(X_v(t) \in F_v \mid X_{\mathrm{pa}_{\mathcal{G}}(v)}(< t) \in E_{\mathrm{pa}_{\mathcal{G}}(\mathrm{v})}\Big)\,. \tag{1}$$

Equation 1 encodes local conditional independence: given the past, $X_v(t)$ depends only on the past of its parents. The summary graph has edges whenever the past influences the future of subprocesses represented by nodes, summarizing the causal pattern of the process. We do not allow for *instantaneous effects* (as in Niemiro and Rajkowski, 2023) (in contrast to González-Pérez, 2025). A self-loop in the summary graph means that past realizations of a variable influence the current moment. A bidirectional path (two paths between the same nodes with opposite arrows) would indicate a feedback loop between two variables, past of both variables influencing the present moment of the other variable.

For snapshot separation questions, we will also use the *trek graph* (Boege et al., 2025). Given a summary graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, its trek graph is the graph $\widehat{\mathcal{G}} = (\mathcal{V}, \widehat{\mathcal{E}})$ with a bidirected edge between i and j, if and only if $\mathrm{An}_{\mathcal{G}}(i) \cap \mathrm{An}_{\mathcal{G}}(j) \neq \varnothing$; that is, i and j have a common ancestor connected to each by a directed path in $\mathcal{G}$.

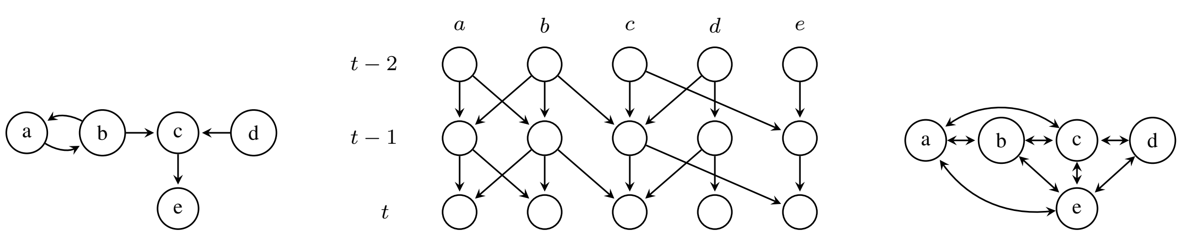
Space-time graph Summary graphs can have *cycles*. Although static graphical models with cycles may fail to induce a (unique) joint law without additional fixed-point or solvability assumptions (Bongers et al., 2021), *temporal unrolling* yields an *acyclic representation*. Fix a time index set $\mathcal{T} \subset \mathbb{Z}$. Given a summary graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with order-1 dynamics (no instantaneous edges), define its *space-time graph*

$$\mathcal{G}_{\mathrm{T}} = (\mathcal{V}_{\mathrm{T}}, \mathcal{E}_{\mathrm{T}}), \qquad \mathcal{V}_{\mathrm{T}} := \mathcal{V} \times \mathcal{T}, \quad \mathcal{E}_{\mathrm{T}} := \{(u, t{-}1) \to (v, t) \ : \ u \in \mathrm{pa}_{\mathcal{G}}(v),\ t, t{-}1 \in \mathcal{T}\}.$$

By construction, for each finite time window, the induced subgraph is a finite DAG $\mathcal{G}_{\mathrm{T}}$ (this would not be guaranteed if instantaneous effects were allowed). Equation 1 encodes a dynamic Bayesian network factorization on $\mathcal{G}_{\mathrm{T}}$. In this sense, the process indexed by a summary graph is well-defined. However, throughout the work, the full space-time graph is used as a graphical representation of an already specified process. Existence and uniqueness on the full countably infinite graph are assumed separately, or in the Gaussian VAR(1) case follow from construction of the coefficient matrix [1]. We make use of space-time graphs heavily in Section 4.

Standing assumptions Throughout the completeness section, we make four standing assumptions about the nature of temporal dynamics (relevant for completeness). First, we assume that all the vertices of $\mathcal{G}$ have *self-loops*. This corresponds to the seemingly natural existence of autoregression, widely accepted in the time series literature (e.g., Hamilton, 2020). Second, we assume *Markovianity of order 1*—the realization of the process depends on the past only through the previous moment. We see this assumption as technical convenience, however one that is not essential to our proof strategy at the conceptual level. Third, we work with a *time-homogeneous* dependence pattern: the parental sets $\mathrm{pa}(v)$ do not depend on t. Fourth, we specialize to *strictly stationary processes*. Precisely, we assume that there exists X that is ergodic, has unique stationary distribution π, and $X(t_0) \sim \pi$, where t_0 is a starting point. We see this as a feature of the analysis that isolates the snapshot properties. Relaxing this assumption would require different conceptualization and analysis tools as the snapshots would potentially become time and history-dependent.

We graph $\mathcal{G}$, three slices $(t-2, t-1, t)$ of the space-time graph $\mathcal{G}_T$, and the trek graph of the summary graph $\widehat{\mathcal{G}}$ in Figure 1.



(a) Summary graph (b) Space-time graph (c) Trek graph

Figure 1: Summary graph, the corresponding three-slice space-time DAG, and the trek graph. Self-loops are implicit and suppressed for summary and trek graphs.

2.1. Separation rules on summary graphs

We next recall the two separation notions that have been proposed for stationary processes: c-separation on the summary graph (Niemiro and Rajkowski, 2023) and separation on the associated trek graph (Boege et al., 2025).

A *path* in a graph is a sequence of pairwise distinct vertices $(v_0, \ldots, v_m)$ such that each consecutive pair is adjacent (with arbitrary edge orientations). A non-endpoint v_i on a path is a collider if both incident edges have arrowheads into v_i: $\to v_i \leftarrow$, $\leftrightarrow v_i \leftrightarrow$, where $\leftrightarrow$ is a shorthand for a reciprocal pair; otherwise, it is a non-collider. For $K \subseteq \mathcal{V}$, write $\mathrm{An}_{\mathcal{G}}(K)$ for the set of graph-

1. For more general discussion on existence and uniqueness of processes that "start" at minus infinity see Hochsprung et al. (2024).

theoretic *ancestors* of K in $\mathcal{G}$ (including K itself in the presence of self-loops), and $\mathrm{De}_{\mathcal{G}}(K)$ for the set of *descendants*.

Definition 1 (C-separation) *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a summary graph. A path is* c-open given $K \subseteq \mathcal{V}$ *if every collider on the path lies in $\mathrm{An}_{\mathcal{G}}(K)$; otherwise it is* c-blocked by K*. For disjoint $I, J, K \subseteq \mathcal{V}$ we say that I and J are* c-separated *by K in $\mathcal{G}$, written $(I \perp_c J \mid K)_{\mathcal{G}}$, if every I–J path in $\mathcal{G}$ is c-blocked by K.*

Unlike d-separation, conditioning does not block via non-colliders; only collider ancestry matters. The trek-graph separation of Boege et al. (2025) is defined on an auxiliary bidirected graph that records shared ancestors.

Definition 2 (Trek graph separation) *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a summary graph, and $\widehat{\mathcal{G}}$ it's trek graph. For disjoint $I, J, K \subseteq V$ we say that I and J are trek graph separated, written $(I \perp J \mid K)_{\widehat{\mathcal{G}}}$ if every I–J path in $\widehat{\mathcal{G}}$ meets K in at least one internal vertex, i.e. trek graph separation is an undirected graph separation on the skeleton of the trek graph.*

One of our main results (Theorem 5) shows that for any summary graph $\mathcal{G}$ with the trek graph $\widehat{\mathcal{G}}$, c-separation in $\mathcal{G}$ and separation in $\widehat{\mathcal{G}}$ encode the same cross-sectional CI structure in a precise sense. One can find an example of a summary graph, space-time graph and a trek graph in Appendix C.

We informally recap the main properties of the separations of interest. The formal definitions and their relation to *faithfulness* and *typicality* are given in Appendix B.1. A separation rule is *sound* if separation always implies CI in the model class, *weakly complete* if every failure of separation is witnessed by a model showing dependence, and *strongly complete* if, in a parametrised family, non-separation implies dependence for almost all parameter values (in a measure-theoretic sense). Weak completeness is a corollary to strong completeness.

2.2. Stationary Gaussian VAR(1)

To make completeness questions concrete and tractable, we focus on linear Gaussian VAR(1) processes Markov to a given summary graph. This gives a natural parametrized model class in which snapshot conditional independence can be expressed through covariance minors and in which we establish strong completeness.

VAR(1). Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a summary graph on $\mathcal{V} = [n]$. Consider an $\mathbb{R}^n$-valued process $X(t) = (X_1(t), \dots, X_n(t))^\top$ satisfying

$$X(t+1) = M\,X(t) + \varepsilon(t), \qquad t \in \mathbb{Z}, \tag{2}$$

where $M \in \mathbb{R}^{n \times n}$ is a *coefficient matrix*, $\varepsilon(t) \sim \mathcal{N}(0, D)$ are i.i.d. with $D \succ 0$, and M respects the sparsity pattern of $\mathcal{G}$: $\mathbb{R}^{\mathcal{E}} := \{M \in \mathbb{R}^{n \times n} : M_{ij} = 0 \text{ whenever } (j \to i) \notin \mathcal{E}\}$. Thus edges $j \to i$ in $\mathcal{E}$ are precisely the positions where M_{ij} may be nonzero; in particular $M_{ii} \neq 0$ (autoregression) is enforced.

Lyapunov equation. If the spectral radius $\rho(M) < 1$, then Equation 2 admits a unique solution (strictly stationary process) with mean zero and covariance matrix Σ solving the discrete *Lyapunov equation*

$$\Sigma = M\Sigma M^\top + D.$$

Equivalently, writing $\Sigma_t = \mathrm{Cov}(X(t))$, we have the recursion $\Sigma_{t+1} = M\Sigma_t M^\top + D$, and under $\rho(M) < 1$, the series $\Sigma = \sum_{k=0}^{\infty} M^k D (M^\top)^k$ is convergent. The series representation admits a graph pathwise interpretation: entries of M^k are sums over directed paths of length k in the space-time graph—if D is diagonal, $(M^k D (M^\top)^k)_{ij} = \sum_{s=1}^{n} (M^k)_{is} D_{ss} (M^k)_{js}$ only receives contributions from nodes that are common ancestors of i and j at lag k. This already suggests why trek-based and collider-based descriptions govern snapshot dependence—if there is a collider on all paths between a and b the unconditional covariance entry Σ_{ab} will be zero. In the interest of clarity, we assume that the noise covariance is diagonal. A more detailed discussion on the relation between the summary graph and covariance of Gaussian VAR is in Appendix B.

In Gaussian models, conditional independence is equivalent to algebraic constraints on Σ. We will express CI as vanishing determinants of covariance minors for the completeness proof.

2.3. Related work

Historically, in the i.i.d. case, Verma and Pearl (1990); Lauritzen et al. (1990) showed the soundness of the d-separation, and then later a separation for mixed (with both directed and undirected edges possible) graphs was proposed—m-separation Richardson (2003). A stronger separation for Gaussian SEMs—t-separation was proven to be sound Sullivant et al. (2010). Local-dependence (summary) graphs have been developed and analysed for several classes of structured processes, including point processes (Didelez, 2008; Mogensen and Hansen, 2022), diffusions (Mogensen and Hansen, 2020; Manten et al., 2025; Boege et al., 2025), and discrete-time stochastic processes (Eichler, 2007; Eichler and Didelez, 2007; Niemiro and Rajkowski, 2023; Gerhardus, 2024). These works introduce sound criteria tailored to temporal structure, governing conditional independences from past to present, or future. Two works are particularly relevant here: Niemiro and Rajkowski (2023) define c-separation ("c" for the collider) on the summary graph and prove its soundness for cross-sectional CI in strictly stationary discrete-time models; and, for stationary diffusion processes, Boege et al. (2025) establish trek-graph separation as a sound criterion for cross-sectional CI.

A spectrum of completeness results links graphical separation to (in)dependence with different quantifiers. Historically, in i.i.d. DAGs, (Geiger and Pearl, 1990a) established weak completeness of d-separation by showing that when there is no d-separation, a Gaussian based dependent Armstrong relation witness exists. Later on in this setting Meek (1995) has proven that d-separation is strongly complete for discrete BNs; Spirtes et al. (2000) showed that unfaithful parameters have Lebesgue measure zero in linear-Gaussian BNs. Generalizing the above results, Boeken et al. (2024) show that, without restriction to any parametric or nonparametric family of distributions, the faithful distributions are typical in topological sense. In discrete time, Niemiro and Rajkowski (2023) establish weak completeness for ε-and δ-separation notions which relate past of subprocesses to presence or future of other subprocesses. Moreover, in Niemiro and Rajkowski (2023) a conjecture is made about the weak completeness of the c-separation. For continuous-time diffusions Boege et al. (2025) prove completeness for trek-graph separation for cross-sectional CI statements.

3. Equivalence between graphical separations

In this section, we relate the three graphical separation notions relevant for cross-sectional conditional independence: c-separation on the summary graph, trek graph separation separation on the trek graph, and a space-time trek separation criterion on the unrolled space-time DAG. The first equivalence identifies the snapshot CI semantics already encoded by the summary graph: c-separation and trek-graph separation induce the same equivalence classes of graphs. The second equivalence translates c-separation directly to the space-time graph, which is the representation used in the completeness proof. Proofs of the results in this section, together with illustrative examples, are given in Appendix C.

3.1. Summary graphs and treks

We first formalize what it means for two graphs to have the same c-separation structure.

Definition 3 (C-equivalence) *Let $\mathcal{G}_1 = (\mathcal{V}, \mathcal{E}_1)$ and $\mathcal{G}_2 = (\mathcal{V}, \mathcal{E}_2)$ be two summary graphs on the same vertex set $\mathcal{V}$. We say that $\mathcal{G}_1$ and $\mathcal{G}_2$ are* c-equivalent*, written $\mathcal{G}_1 \equiv_c \mathcal{G}_2$, if for every triple of pairwise disjoint subsets $I, J, K \subset \mathcal{V}$,*

$$(I \perp_c J \mid K)_{\mathcal{G}_1} \quad \Longleftrightarrow \quad (I \perp_c J \mid K)_{\mathcal{G}_2}.$$

The following lemma captures the relationship between c-equivalence and trek graphs. The intuition is that the trek graph only records which vertices share a common ancestor (Definition 2); in particular, it distinguishes colliders and their descendants from non-colliders.

Lemma 4 (Summary graphs c-equivalence through trek graphs) *Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be summary graphs on V with their trek graphs $\widehat{\mathcal{G}}_1$ and $\widehat{\mathcal{G}}_2$ respectively. Then*

$$\mathcal{G}_1 \equiv_c \mathcal{G}_2 \quad \Longleftrightarrow \quad \widehat{\mathcal{G}}_1 = \widehat{\mathcal{G}}_2.$$

This shows already that the two notions encode the same structure of the summary graph. However, it remains to translate the directed and undirected graph separations. We can now state the main equivalence result.

Theorem 5 (Equivalence between c-separation and trek separation) *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a summary graph with trek graph $\widehat{\mathcal{G}}$, and let $I, J, K \subset V$ be pairwise disjoint. Then*

$$(I \perp_c J \mid K)_{\mathcal{G}} \iff \big(I \perp J \mid V \setminus (I \cup J \cup K)\big)_{\widehat{\mathcal{G}}}.$$

Trek graph separation is about existence of paths, as it says that there is no $I - J$ path in the subgraph of the trek graph induced by the $I \cup J \cup K$. Combine the fact with the fact that the trek graph encodes colliders and their descendants and observe that these act as "bottlenecks/bridges" through which a path can be formed (recall Figure 1 for illustration).

Together, Lemma 4 and Theorem 5 show that graphical conditions phrased in terms of c-separation (Niemiro and Rajkowski, 2023) can equivalently be checked by trek graph separation on the associated trek graph, paralleling the perspective of Boege et al. (2025). Intuitively, collider structures govern the flow of information from the past that travels to the current cross-section through common ancestors.

3.2. Summary graphs and space-time treks

To complement the picture of collider separations and for convenience in proving completeness with the unrolled space-time DAG $\mathcal{G}_T$ we translate c-separation to space-time. We formulate a separation directly on $\mathcal{G}_T$ and show that it is equivalent to c-separation on the summary graph.

We define the basic graphical objects on the space-time graph. First, the notion of space-time treks is an application of the idea found in Sullivant et al. (2010) that defined treks on DAGs relevant for static setting.

Definition 6 (Space-time trek) *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a summary graph and let $\mathcal{G}_T = (\mathcal{V}_T, \mathcal{E}_T)$ be its space-time graph. Fix a time t and an integer $k \geq 1$. A* space-time trek *(st-trek) of lag k between (i, t) and (j, t) is an ordered pair of directed paths (P_1, P_2) together with a space-time vertex $(u, t-k)$ such that*

1. *P_1 runs from $(u, t-k)$ to (i, t) and P_2 runs from $(u, t-k)$ to (j, t),*
2. *every internal vertex of P_1 and P_2 has time coordinate strictly between $t-k$ and t.*

The vertex $(u, t-k)$ is the top *of the trek.*

Because every edge in $\mathcal{G}_T$ advances time by one step, each leg of an st-trek is time-monotone. In particular, directed paths in $\mathcal{G}_T$ cannot contain collider turns. Thus st-treks are precisely the space-time counterparts of collider-free paths in the summary graph.

Definition 7 (Multi-trek) *Let $a, b \in \mathcal{V}$. A* multi-trek *from a to b at time t is a sequence of distinct spatial vertices $\mu(a, b) = (c_0, \ldots, c_m)$ with $c_0 = a$ and $c_m = b$ such that for each $r = 1, \ldots, m$ there exists an st-trek between (c_{r-1}, t) and (c_r, t). The vertices $c_1, \ldots, c_{m-1}$ are the* junctions.

The multi-trek records how information can pass between present-time variables through successive common-ancestor links. To match the collider-based conditioning rule of c-separation, we require that these junctions lie in the ancestor set of the conditioning variables.

Definition 8 (Space-time separation) *Let $I, J, K \subseteq \mathcal{V}$ be pairwise disjoint. A multi-trek $\mu(i, j) = (c_0, \ldots, c_m)$ is* active *given C if every junction $c_1, \ldots, c_{m-1}$ lies in $\mathrm{An}_{\mathcal{G}}(K)$. We say that I is* space-time separated *from J by K, written*

$$(I \perp_{st} J \mid K)_{\mathcal{G}_T},$$

if there is no active multi-trek from any $i \in I$ to any $j \in J$ given K.

The following theorem is the space-time translation of c-separation.

Theorem 9 (Equivalence between c-separation and space-time separation) *Let $\mathcal{G}$ be a summary graph with self loops and let $\mathcal{G}_T$ be its space-time graph. For pairwise disjoint node sets $I, J, K \subseteq \mathcal{V}$,*

$$(I \perp_c J \mid K)_{\mathcal{G}} \iff (I \perp_{st} J \mid K)_{\mathcal{G}_T}.$$

Theorem 9 provides the graphical bridge used later in the completeness proof. It allows us to translate a failed c-separation on the summary graph into the existence of an active multi-trek in the acyclic space-time DAG. This is exactly the form needed for the covariance-minor and path-expansion arguments in Section 4.

4. Completeness

We now turn to the completeness question for collider separation. Our main result is a strong completeness theorem for the family of stable Gaussian VAR(1) models Markov to a given summary graph. In this formulation, whenever c-separation fails, the corresponding snapshot conditional independence can hold only on a Lebesgue measure-zero subset of the parameter space. As a consequence, weak completeness follows immediately. Throughout this section we work with diagonal innovation covariance and with the standing assumptions from Section 2. We relegate the proofs to the Appendix D

Theorem 10 (Strong completeness within Gaussian VAR(1)) *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a summary graph on $\mathcal{V} = [n]$, and let $\Theta_{\mathcal{G}} := \{(M, D) : \rho(M) < 1,\ D = \operatorname{diag}(d_v) \succ 0,\ M \in \mathbb{R}^{\mathcal{E}}\}$ be the parameter domain, equipped with Lebesgue measure on its coordinates. For $(M, D) \in \Theta_{\mathcal{G}}$, let X be the corresponding strictly stationary Gaussian VAR(1) process with stationary covariance matrix $\Sigma(M, D)$. Then for every triple of pairwise disjoint node sets $I, J, K \subseteq \mathcal{V}$:*

1. Soundness. *If $(I \perp_c J \mid K)_{\mathcal{G}}$, then $X_I(t) \perp\!\!\!\perp X_J(t) \mid X_K(t) \qquad$ for every $(M, D) \in \Theta_{\mathcal{G}}$.*
2. Strong completeness. *If $(I \not\perp_c J \mid K)_{\mathcal{G}}$, then $\{(M, D) \in \Theta_{\mathcal{G}} : X_I(t) \perp\!\!\!\perp X_J(t) \mid X_K(t)\}$ has Lebesgue measure zero.*

Corollary 11 (Weak completeness within VAR(1)) *Under the assumptions of Theorem 10, c-separation is weakly complete for Markov order 1 processes Markov to $\mathcal{G}$. Let*

$$\mathcal{I}_{\mathcal{G}} := \{(I, J, K) : I, J, K \subseteq \mathcal{V} \text{ pairwise disjoint and } (I \perp_c J \mid K)_{\mathcal{G}}\}.$$

Then there exists a strictly stationary stochastic process $X = \big(X_v(t)\big)_{v \in \mathcal{V},\, t \in \mathbb{Z}}$ on a common probability space such that, for all pairwise disjoint $I, J, K \subseteq \mathcal{V}$ and for all $t \in \mathbb{Z}$,

1. Soundness. *For* every *triple $(I, J, K) \in \mathcal{I}_{\mathcal{G}}$: $X_I(t) \perp\!\!\!\perp X_J(t) \mid X_K(t)$,*
2. Weak Completeness. *Conversely, for any triple $(I, J, K) \notin \mathcal{I}_{\mathcal{G}}$: $X_I(t) \not\perp\!\!\!\perp X_J(t) \mid X_K(t)$.*

Weak completeness parallels the result on the classical role of d-separation in the static setting (Geiger and Pearl, 1990a) and other time series separation criteria (Niemiro and Rajkowski, 2023).

Proof idea. The proof idea is inspired by Sullivant et al. (2010). The proof proceeds by translating conditional independence into an algebraic statement and then interpreting that algebraic statement graphically on the space-time DAG. By Theorem 9, we may replace c-separation on the summary graph by the equivalent space-time separation on the unrolled graph. In the Gaussian VAR(1) setting, snapshot conditional independence is equivalent to the vanishing of suitable covariance minors. Thus, the problem becomes: when is a certain determinant of the stationary covariance matrix identically zero as a function of the VAR parameters?

To analyse these determinants, we factor the stationary covariance as $\Sigma = XX^\top$ and expand the relevant minors by the Cauchy–Binet theorem over top-lag source sets. For a fixed source set, a Lindström-Gessel–Viennot argument on a finite time truncation of the space-time graph expresses the corresponding minor as a signed sum over clean systems of vertex-disjoint directed space-time paths. This identifies the precise graphical objects that contribute to each Cauchy–Binet summand.

If space-time separation holds, no such clean system exists, so every summand vanishes; this gives soundness.

For the strong-completeness direction, we start from a failed space-time separation. We compress that failure to a shortest path in the trek graph, extract from it a canonical clean system of positive-lag space-time treks, and adjoin lag-zero innovation columns for the remaining conditioned vertices. This yields a distinguished Cauchy–Binet summand. Inside that summand we isolate a monomial determined by the chosen witness system and show that no other summand can produce the same monomial. Hence, the covariance minor is not identically zero. Since the covariance depends real-analytically on the VAR parameters on the domain, the zero set of that minor has Lebesgue measure zero. This proves strong completeness, and weak completeness follows immediately as a corollary.

4.1. Supporting results

Covariance and space-time treks Before turning to determinants, we record the basic link between the stationary covariance matrix and the space-time graph. It is the dynamic analogue of the usual trek rule: every off-diagonal covariance entry is a sum over pairs of directed space-time paths with a common top.

Recall that a for stable Gaussian VAR(1) with diagonal innovation covariance $D = \operatorname{diag}(d_1, \dots, d_n)$, we have $\Sigma = \sum_{s\geq 0} M^s D (M^\top)^s = \sum_{s\geq 0}\sum_u (M^s)_{iu}\, d_u\, (M^s)_{ju}$. Let $\tau = (P_1, P_2)$ be an st-trek of lag $s \geq 1$ between (i,t) and (j,t) with top $(u, t-s)$. Because every edge in $\mathcal{G}_T$ advances time by one step, a directed path from $(u, t-s)$ to (i,t) has exactly s edges. Hence, for $s \geq 1$,

$$(M^s)_{iu} = \sum_{P:(u,t-s)\rightsquigarrow(i,t)} \; \prod_{(k,r-1)\to(l,r)\in P} M_{lk},$$

where the sum ranges over all directed paths in the space-time graph from $(u, t-s)$ to (i,t).

Define the *weight* of τ by

$$\omega(M, D, \tau) := d_u \prod_{(k,r-1)\to(l,r)\in P_1} M_{lk} \prod_{(k,r-1)\to(l,r)\in P_2} M_{lk}.$$

Equivalently, the weight is the innovation variance at the top multiplied by the products of the VAR coefficients along the two directed legs of the trek. Combining the above, we obtain

$$\begin{aligned}
\Sigma_{ij} &= \mathbf{1}_{\{i=j\}}\, d_i + \sum_{s\geq 1}\sum_{u=1}^{n} (M^s)_{iu}\, d_u\, (M^s)_{ju} \\
&= \mathbf{1}_{\{i=j\}}\, d_i + \sum_{s\geq 1}\sum_{u=1}^{n} \sum_{P_1:(u,t-s)\rightsquigarrow(i,t)} \; \sum_{P_2:(u,t-s)\rightsquigarrow(j,t)} \omega(M, D, (P_1, P_2))
\end{aligned}$$

Thus every off-diagonal covariance entry is a sum over st-treks between (i,t) and (j,t), indexed by their lag and top.

Gaussian CI For pairwise disjoint $i, j, K \subseteq \mathcal{V}$, define

$$I' := \{i\} \cup K, \qquad J' := \{j\} \cup K, \qquad f_{i,j,K}(M, D) := \det \Sigma_{I',J'}(M, D).$$

The following is standard for Gaussian vectors.

Lemma 12 (Gaussian minor criterion) *Let $X \sim \mathcal{N}_n(0, \Sigma)$ and let I, J, K be pairwise disjoint,*

$$X_I \perp\!\!\!\perp X_J \mid X_K \quad \Longleftrightarrow \quad \det \Sigma_{\{i\}\cup K, \{j\}\cup K} = 0 \quad \textit{for all } i \in I,\ j \in J.$$

Thus the pairwise problem is to understand when $f_{i,j,K}$ vanishes identically and when it is a nontrivial analytic function. The block statement of Theorem 10 will follow by applying the pairwise criterion to all $i \in I$, $j \in J$.

From determinants to series In order to apply Cauchy-Binet lemma one has to be able to express a matrix as a product of two matrices. We write $\Sigma = XX^\top$, with

$$X = [D^{1/2},\ MD^{1/2},\ M^2D^{1/2}, \dots] = Y\Delta, \qquad Y = [I, M, M^2, \dots],$$

where columns are indexed by top-lag pairs $(u, k) \in \mathcal{V} \times \mathbb{N}_0$, and Δ is diagonal with $\Delta_{(u,k),(u,k)} = \sqrt{d_u}$.

Lemma 13 (Cauchy–Binet decomposition by top-lag sets) *Let $I, J \subseteq \mathcal{V}$ with $|I| = |J| = \ell$. For a finite column set $S = \{(v_1, k_1), \dots, (v_\ell, k_\ell)\} \subseteq \mathcal{V} \times \mathbb{N}_0$, define the top-multiplicity vector $m(S) = (m_v(S))_{v\in\mathcal{V}}$, with $m_v(S) := \big|\{p :\ v_p = v\}\big|$. Then*

$$\det \Sigma_{I,J} = \sum_{\substack{S\subseteq\mathcal{V}\times\mathbb{N}_0\\ |S|=\ell}} \det X_{I,S}\, \det X_{J,S} = \sum_{\substack{m\in\mathbb{N}_0^{\mathcal{V}}\\ \sum_v m_v=\ell}} \left(\sum_{\substack{S\subseteq\mathcal{V}\times\mathbb{N}_0\\ |S|=\ell,\ m(S)=m}} \det Y_{I,S}\, \det Y_{J,S} \right) \prod_{v\in\mathcal{V}} d_v^{m_v}.$$

Hence, if the d_v are algebraically independent, then

$$\det \Sigma_{I,J} \equiv 0 \quad \Longleftrightarrow \quad \sum_{\substack{S\subseteq\mathcal{V}\times\mathbb{N}_0\\ |S|=\ell,\ m(S)=m}} \det Y_{I,S}\, \det Y_{J,S} \equiv 0 \quad \textit{for every } m.$$

We next interpret the fixed-source minors in terms of the space-time graph and show that there exists a unique monomial that is a summand in the Cauchy-Binet expansion.

Lemma 14 (Lindström–Gessel–Viennot for X-minors) *Fix a set of sinks $S = \{s_1, \dots, s_\ell\} \subseteq \mathcal{V}$ and a set of distinct top-lag pairs $R = \{(v_1, k_1), \dots, (v_\ell, k_\ell)\} \subseteq \mathcal{V} \times \mathbb{N}$. Let $\mathcal{P}_{k_p}(v_p \rightsquigarrow s)$ denote the set of directed space-time paths from $(v_p, t - k_p)$ to (s, t), and let $\mathcal{N}(R, S)$ be the set of vertex-disjoint systems of such paths from the tops in R to the sinks in S. Then*

$$\det X_{S,R} = \sum_{\mathbf{P}\in\mathcal{N}(R,S)} \operatorname{sgn}\big(\sigma(\mathbf{P})\big) \prod_{p=n}^{\ell} \sqrt{d_{v_n}}\, M_{P_n},$$

where $\sigma(\mathbf{P})$ is the permutation induced by the routing of the tops to the sinks, and M_{P_n} is a product over vertices on the whole path P_n.

In particular, if $\mathcal{N}(R, S) = \varnothing$, then $\det X_{S,R} \equiv 0$.

Lemma 15 (Canonical clean chain from failed st-separation) *Assume $i \not\perp_{st} j \mid K$. Choose a shortest path $q_0 = i,\ q_1, \dots, q_r = j$ in the trek graph $\widehat{\mathcal{G}}$ whose internal vertices lie in K. Then for each $p = 1, \dots, r$ there exist a common ancestor $u_p \in \mathrm{An}_{\mathcal{G}}(q_{p-1}) \cap \mathrm{An}_{\mathcal{G}}(q_p)$ and directed paths $L_p : u_p \rightsquigarrow q_{p-1}$, $R_p : u_p \rightsquigarrow q_p$, such that, after padding them to equal lengths by self-loops at the top, the resulting positive-lag st-treks form a clean ordered system: the left legs are pairwise vertex-disjoint and the right legs are pairwise vertex-disjoint.*

Lemma 16 (Isolation of a distinguished monomial) *Under the assumptions of Lemma 15, let $K_0 := K \setminus \{q_1, \dots, q_{r-1}\}$, $S^\star := \{(u_1, k_1), \dots, (u_r, k_r)\} \cup \{(k, 0) : k \in K_0\}$, where the k_i are the padded lags from Lemma 15. Assign algebraically independent variables to parametrization of M, then*

1. *the minor $\det Y_{\{i\}\cup K,\, S^\star}$ contains a monomial $\mu_L^\star$, the minor $\det Y_{\{j\}\cup K,\, S^\star}$ contains a monomial $\mu_R^\star$, arising from the chosen clean system, and these monomials occur only once in that minor;*

2. *the $S^\star$-summand in the Cauchy–Binet expansion of $\det \Sigma_{\{i\}\cup K,\, \{j\}\cup K}$ therefore contains the product monomial*
$$\widetilde{\mu}^\star := \Big(\prod_{v \in \mathcal{V}} d_v^{m_v(S^\star)}\Big) \mu_L^\star \mu_R^\star$$
with nonzero coefficient;

3. *no other Cauchy–Binet summand can produce the same monomial $\widetilde{\mu}^\star$.*

The $S^\star$-summand in the Cauchy–Binet expansion of $\det \Sigma_{i\cup K,\, j\cup K}$ contains a monomial with nonzero coefficient that is produced by no other Cauchy–Binet summand. Consequently, the corresponding covariance minor is not identically zero, $\det \Sigma_{i\cup K,\, j\cup K} \not\equiv 0$, as a function of the VAR parameters.

Finally the map $(M, D) \mapsto \Sigma(M, D)$ is real-analytic on the domain $\Theta_{\mathcal{G}} = \{(M, D) : \rho(M) < 1,\ D \succ 0,\ M \in \mathbb{R}^{\mathcal{E}}\}$. Hence $f_{i,j,K}(M, D) := \det \Sigma_{i\cup K,\, j\cup K}(M, D)$ is real-analytic on $\Theta_{\mathcal{G}}$. By Lemma 16, it is not identically zero. Therefore, we conclude by noticing that its zero set has Lebesgue measure zero.

4.2. Examples

Example 1 (Collider) *Consider the summary graph $a \to b \leftarrow c$ with self-loops at all vertices.*

Closed *Consider conditional independence query $I = \{a\}$, $\quad J = \{c\}$, $\quad K = \varnothing$. Then $(a \perp_c c)_{\mathcal{G}}$, since the only a-c path has a collider at b, and $b \notin \mathrm{An}_{\mathcal{G}}(\varnothing)$. By Theorem 9, this is equivalent to $a \perp_{st} c$ on the space-time graph. To see how the soundness proof works, write*

$$A' = \{a\}, \qquad B' = \{c\}, \qquad \Sigma_{ac} = \sum_{\substack{S \subseteq \mathcal{V}\times\mathbb{N}_0 \\ |S|=1}} \det X_{A',S} \det X_{B',S}.$$

For any singleton source set $S = \{(u, k)\}$, if both $\det X_{A',S}$ and $\det X_{B',S}$ were nonzero, then by Lemma 14 there would exist a directed space-time path from $(u, t-k)$ to (a, t) and another from $(u, t-k)$ to (c, t), i.e. an st-trek between a and c. This contradicts $a \perp_{st} c$. Hence every Cauchy–Binet summand vanishes and $\Sigma_{ac} = 0$. Therefore $X_a(t) \perp\!\!\!\perp X_c(t)$.

Opened *Now let* $K = \{b\}$. *Then* $A' = \{a, b\}$, $\quad B' = \{c, b\}$. *Take* $S^\star = \{(a,1),(b,1)\}$. *With rows ordered as* (a, c) *and* (c, b), *and columns ordered as* $((a,1),(b,1))$,

$$X_{A',S^\star} = \begin{pmatrix} \sqrt{d_a} M_{aa} & 0 \\ \sqrt{d_a} M_{ca} & \sqrt{d_b} M_{cb} \end{pmatrix}, \qquad X_{B',S^\star} = \begin{pmatrix} \sqrt{d_a} M_{ca} & \sqrt{d_b} M_{cb} \\ 0 & \sqrt{d_b} M_{bb} \end{pmatrix}.$$

Therefore

$$\det X_{A',S^\star} = \sqrt{d_a d_b}\, M_{aa} M_{cb}, \qquad \det X_{B',S^\star} = \sqrt{d_a d_b}\, M_{ca} M_{bb}.$$

So the $S^\star$*-summand contributes the monomial* $d_a d_b\, M_{aa} M_{bb} M_{ca} M_{cb}$, *and thus*

$$\det \Sigma_{a\cup c,\, b\cup c} \not\equiv 0.$$

The contrast with the unconditional case is the basic graphical mechanism behind the proof: without conditioning, the collider blocks all clean witness systems; conditioning on the collider creates a canonical clean system and therefore a nontrivial covariance minor.

5. Conclusions

In this paper, we reconcile two previously separate graphical viewpoints of summary graphs: trek-graph separation (Boege et al., 2025) and collider-based c-separation (Niemiro and Rajkowski, 2023). We complement the picture with a translation of collider separation to space-time graphs with space-time trek separation. We show that, graphically, at the snapshot level, it is precisely the collider structure that governs if information flows from past to cross-section through shared ancestors. Moreover, we resolve the conjecture of Niemiro and Rajkowski (2023) on the necessity of c-separation for cross-sectional CI in strictly stationary discrete-time processes. In particular, we show that c-separation is at least weakly complete and strongly complete within the Gaussian VAR(1) family. These results, from a causal perspective, identify precise conditions under which algorithms that operate only on snapshot CI tests are sound and complete for recovering the c-separation / trek-equivalence class of the underlying summary graph of a dynamic data-generating process. On the technical side, our arguments adapt trek-based techniques from Sullivant et al. (2010) to space-time graphs.

Several extensions appear natural. For completeness, it would be valuable to relax the Markov-order-1 assumption, and for modelling, to understand how far strong completeness extends beyond Gaussian. A further generalization of CI to non-stationary stochastic processes seems of great utility, but most probably requires a different approach than presented in this article. At last, we speculate that alternative proofs for completeness results that use trek rules defined directly on summary graphs might be possible (see Recke et al., 2026).

A more detailed discussion of the practical implications for causal discovery, as well as of the role and scope of our main assumptions—including strict stationarity, order-1 dynamics, absence of instantaneous effects, and the Gaussian VAR(1) setting—is deferred to Appendix A. For a dessert, we leave an alternative, but only partial proof of weak completeness in Appendix E, where we adapt strategy of Geiger and Pearl (1990a) and extend Armstrong product to strictly stationary processes.

Acknowledgments

Part of this work was completed while the author was at the University of Warsaw. I am grateful to Lukasz Rajkowski for many helpful discussions about weak completeness, and to CoCaLa members for valuable discussion about the exposition of the results. Moreover, I would like to thank anonymous CLeaR reviewers for actionable and detailed feedback on the previous version of the paper. Any remaining errors are mine. This paper was written within the scope of Novo Nordisk Foundation grant Project Grants in the Natural and Technical Sciences 2024 (#0096126), whose support is gratefully acknowledged.

References

Yacine Aït-Sahalia and Per A Mykland. The effects of random and discrete sampling when estimating continuous–time diffusions. *Econometrica*, 71(2):483–549, 2003.

Charles K Assaad, Emilie Devijver, and Eric Gaussier. Discovery of extended summary graphs in time series. In *Uncertainty in Artificial Intelligence*, pages 96–106. Pmlr, 2022.

Abraham Berman and Robert J Plemmons. *Nonnegative matrices in the mathematical sciences*. SIAM, 1994.

Tobias Boege, Mathias Drton, Benjamin Hollering, Sarah Lumpp, Pratik Misra, and Daniela Schkoda. Conditional independence in stationary diffusions. *Stochastic Processes and their Applications*, 2025.

Philip Boeken, Patrick Forré, and Joris M Mooij. Are bayesian networks typically faithful? *arXiv preprint arXiv:2410.16004*, 2024.

Stephan Bongers, Patrick Forré, Jonas Peters, and Joris M Mooij. Foundations of structural causal models with cycles and latent variables. *The Annals of Statistics*, 49(5):2885–2915, 2021.

Vanessa Didelez. Graphical models for marked point processes based on local independence. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 70(1):245–264, 2008.

Michael Eichler. Granger causality and path diagrams for multivariate time series. *Journal of Econometrics*, 137(2):334–353, 2007.

Michael Eichler and Vanessa Didelez. Causal reasoning in graphical time series models. In *Proceedings of the Twenty-Third Conference on Uncertainty in Artificial Intelligence*, UAI'07, page 109–116. AUAI Press, 2007. ISBN 0974903930.

Dan Geiger and Judea Pearl. On the logic of causal models. In *Machine intelligence and pattern recognition*, volume 9, pages 3–14. Elsevier, 1990a.

Dan Geiger and Judea Pearl. Logical and algorithmic properties of independence and their application to bayesian networks. *Annals of Mathematics and Artificial Intelligence*, 2(1):165–178, 1990b.

Dan Geiger and Judea Pearl. Logical and algorithmic properties of conditional independence and graphical models. *The Annals of Statistics*, pages 2001–2021, 1993.

Andreas Gerhardus. Characterization of causal ancestral graphs for time series with latent confounders. *The Annals of Statistics*, 52(1):103–130, 2024.

Ignacio González-Pérez. Causality for varma processes with instantaneous effects: The global markov property, faithfulness and instrumental variables. *arXiv preprint arXiv:2501.13037*, 2025.

James D Hamilton. *Time series analysis*. Princeton university press, 2020.

Tom Hochsprung, Jakob Runge, and Andreas Gerhardus. A global markov property for solutions of stochastic difference equations and the corresponding full time graphs. In *The 40th Conference on Uncertainty in Artificial Intelligence*, 2024.

Steffen L Lauritzen, A Philip Dawid, Birgitte N Larsen, and H-G Leimer. Independence properties of directed markov fields. *Networks*, 20(5):491–505, 1990.

Georg Manten, Cecilia Casolo, Søren Wengel Mogensen, and Niki Kilbertus. An asymmetric independence model for causal discovery on path spaces. *Proceedings of the Fourth Conference on Causal Learning and Reasoning, PMLR*, 2025.

Christopher Meek. Strong completeness and faithfulness in bayesian networks. *UAI'95: Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*, pages 411–418, 1995.

Søren Wengel Mogensen and Niels Richard Hansen. Markov equivalence of marginalized local independence graphs. *The Annals of Statistics*, 48(1):539–559, 2020.

Søren Wengel Mogensen and Niels Richard Hansen. Graphical modeling of stochastic processes driven by correlated noise. *Bernoulli*, 28(4):3023–3050, 2022.

Wojciech Niemiro and Łukasz Rajkowski. Local dependence graphs for discrete time processes. In *Conference on Causal Learning and Reasoning*, pages 772–790. PMLR, 2023.

Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. *Elements of causal inference: foundations and learning algorithms*. The MIT Press, 2017.

Cecilie Olesen Recke, Sarah Lumpp, Nataliia Kushnerchuk, Janike Oldekop, Jiayi Li, Jane Ivy Coons, and Elina Robeva. Identifiability in graphical discrete lyapunov models. *arXiv preprint arXiv:2601.21818*, 2026.

Thomas Richardson. Markov properties for acyclic directed mixed graphs. *Scandinavian Journal of Statistics*, 30(1):145–157, 2003.

Peter Spirtes, Clark N Glymour, and Richard Scheines. *Causation, prediction, and search*. MIT press, 2000.

Seth Sullivant, Kelli Talaska, and Jan Draisma. Trek separation for gaussian graphical models. *The Annals of Statistics*, 38(3):1665–1685, 2010.

Nikolaj Thams, Rikke Søndergaard, Sebastian Weichwald, and Jonas Peters. Identifying causal effects using instrumental time series: Nuisance iv and correcting for the past. *Journal of Machine Learning Research*, 25(302):1–51, 2024.

Thomas Verma and Judea Pearl. Causal networks: Semantics and expressiveness. In *Machine intelligence and pattern recognition*, volume 9, pages 69–76. Elsevier, 1990.

Matthew J Vowels, Necati Cihan Camgoz, and Richard Bowden. D'ya like dags? a survey on structure learning and causal discovery. *ACM Computing Surveys*, 55(4):1–36, 2022.

Appendix A. Further discussion

A.1. What are the practical implications for causal discovery from this work?

Primarily, we interpret the position of our work as complementing part of the foundations of causality with graphical models. However, we believe a couple of observations of practical usage could be made.

If one applies constraint-based procedures to snapshot data generated by a stationary time series, our results justify treating c-separation (equivalently, the trek-graph separation) as the correct graphical conditional independence language for the snapshots. We clarify the limitation: snapshot CI tests can, at best, recover the trek graph / collider separation-equivalence class. The separation depends on collider/ancestor structure in a way that can be shared by multiple directed summary graphs. Orienting edges in the summary graph requires extra sources of information. Moreover, for this setting, our results are a first step and evidence for justification of the use of snapshot conditional independence constraints in constraint-based procedures and suggest that faithfulness is a rather weak, plausibly satisfied assumption. The strong completeness theorem justifies faithfulness-type assumptions in the mentioned model class.

A.2. What is the role of time index when it comes to existence and uniqueness of the stochastic process and how do the results depend on the time index?

Throughout the main results we work with a process $X = (X_v(t))_{v\in\mathcal{V},\, t\in\mathcal{T}\subset\mathbb{Z}}$. The associated space-time graph

$$\mathcal{G}_{\mathrm{T}} = (\mathcal{V}\times\mathbb{Z}, \mathcal{E}_{\mathrm{T}}), \qquad (u, s-1) \to (v, s) \in \mathcal{E}_{\mathrm{T}} \iff u \in \mathrm{pa}_{\mathcal{G}}(v),$$

is therefore possibly countably infinite. In the absence of instantaneous effects, every edge advances time by one step. Hence the full space-time graph is acyclic, and for every finite window $W \subset \mathbb{Z}$ the induced subgraph $\mathcal{G}_{\mathrm{T}\subseteq W}$ is a finite DAG [2]. Equation (1) therefore induces the usual dynamic Bayesian network factorization on each finite window.

However, in contrast to the finite-DAG case, acyclicity of a countably infinite space-time graph does not by itself imply existence or uniqueness of a joint law Markov to $\mathcal{G}_{\mathrm{T}}$ (see Hochsprung et al., 2024, for extensive discussion). Accordingly, in this paper the space-time graph is used as a graphical representation of an already existing process. In the general setting we assume existence of a strictly stationary process satisfying the local Markov property with respect to the summary graph. We assume that there exists X that is ergodic, has unique stationary distribution π, and $X(t_0) \sim \pi$, where t_0 is a starting point. In the Gaussian VAR(1) sections, existence and uniqueness are obtained separately from the stability condition $\rho(M) < 1$. Whenever we refer to separation statements on the space-time graph, these are statements about finite sets of nodes.

A.3. What is the role of the assumption about strict stationarity?

Our focus is explicitly on a well-posed snapshot law that does not depend on the time index. Without stationarity (or any other assumption of this type), cross-sectional conditional independence of the process becomes time-dependent, and the object we aim to characterize is no longer a single

2. For relevant marginalization for time series see Thams et al. (2024).

conditional independence structure. On the one hand, one might see this as a limitation of the generality of the results. On the other hand, this focus helps to better understand the foundations of reasoning with conditional independence for cross-sectional regime. In this light, we don't see it as a limitation, rather a feature.

A.4. How strong is the assumption about first order Markov dynamics?

In order to introduce higher order processes, one would need to carefully redefine the summary graph or introduce a similar notion, such that it is clear what the edges in the summary graph mean. This would require more discussion, but technically should be in reach.

For the strong completeness result the use of space-time treks and strategy to express the determinantal criterion as a power series are robust to extensions. The technical bottleneck lies in covariance representation provided by the Lyapunov equation and the convergent series that allows us to use Cauchy Binet and dynamic Lindström-Gessel–Viennot lemmas. For higher-order VAR(p), for example one can in principle use a companion-form state augmentation to obtain a VAR(1) on an enlarged state, but (a) the relevant conditional independence snapshots are not the same as the conditional independences in the augmented state, and (b) the space-time combinatorics becomes more involved. So while in principle it seems this proof technique may generalize, it requires careful representation of the Lyapunov equation covariance for higher order VARs.

To put this assumption in the context, the technical difficulty of working with stochastic processes is higher than static, fixed settings, we are not outliers to assume VAR(1) setting (e.g. Recke et al., 2026).

A.5. How strong is the assumption about absence of instantaneous effects?

If instantaneous effects are allowed, then one has to carefully study conditions when does there exist a process that is well defined, as the space-time graph might become cyclic. In general, if instantaneous effects are introduced, the class of admissible processes changes substantially. In particular, one must specify a contemporaneous structural model (or an equivalent solvability condition), and the cross-sectional conditional independence constraints can change even for the same lagged structure.

In an instantaneous-effect model (e.g., a structural VAR with a contemporaneous mixing matrix), cross-sectional conditional independences can depend on contemporaneous coefficients and innovation correlations in ways that are not captured by the purely lagged summary graph. Consequently, both the completeness and the appropriate separation rule would need to be reconsidered.

In sum, we view instantaneous effects as a genuinely different modelling regime that would require (a) an explicit contemporaneous structural model (or equivalent well-posedness assumptions), and (b) a revised graphical separation for snapshot conditional independences. Our current results are not intended to cover that regime.

To put this assumption about absence of instantaneous effects in the context, it is also present in Niemiro and Rajkowski (2023), who proposed and conjectured the completeness of collider separation.

A.6. How strong is the assumption on Gaussian VAR(1)?

Our proof of strong completeness explicitly relies on Gaussianity of the process. Gaussianity works when translating conditional independence into algebraic constraints on the covariance matrix—

vanishing of certain determinants/minors. This is the bridge that turns conditional independence into a real-analytic condition in the model parameters, enabling the ”nontrivial analytic function measure-zero zero set” argument. Without Gaussianity, conditional independence is not characterized by covariance minors, so the entire determinantal / trek-expansion route does not apply directly. This is why our strong completeness proof currently does not extend to non-Gaussian settings without substantially different tools

To put this assumption in larger context, many works that attempt to prove strong completeness also resort to making assumption of Gaussianity, for example, Spirtes et al. (2000) does that for d-separartion, Sullivant et al. (2010) for static graphs and t-separation, and Boege et al. (2025) does that for diffusions.

Appendix B. Background and notation

Definition 17 (Walks, paths, and directed paths) *Let $G = (V, E)$ be a directed graph, possibly with self-loops.*

- *A* walk *in $\mathcal{G}$ is a finite sequence of vertices $(v_0, \ldots, v_m)$ such that for every $r = 1, \ldots, m$, the vertices v_{r-1} and v_r are adjacent.*
- *A* path *is a walk whose vertices are pairwise distinct.*
- *A* directed walk *is a walk such that $v_{r-1} \to v_r \in E \qquad$ for all $r = 1, \ldots, m$.*
- *A* directed path *is a directed walk whose vertices are pairwise distinct.*

Collider/non-collider status and c-openness are extended from paths to walks in the obvious way: a walk is c-open given K if every internal collider on the walk lies in $An_G(K)$.

B.1. Separation criteria properties

A *separation* d is a graph-theoretic relation $A \perp_{\mathrm{d}} B \mid C$ between three disjoint node sets $A, B, C \subseteq \mathcal{V}$ intended to correspond to probabilistic conditional independence statements through factorisation between random variables indexed by these node sets: $X_A \perp\!\!\!\perp X_B \mid X_C$. The interest lies in finding which sparsity structures on the summary graph correspond (both directions) to the probabilistic space CI relations.

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph, and let $X = (X_v)_{v \in \mathcal{V}}$ be random variables in a common probability space. Fix a graphical separation rule d (e.g., d-separation for DAGs (Geiger and Pearl, 1990a), m-separation for ADMGs (Richardson, 2003), or a temporal rule for local dependence graphs (Niemiro and Rajkowski, 2023)). For disjoint $A, B, C \subseteq \mathcal{V}$ write

$$(A \perp_{\mathsf{d}} B \mid C)_{\mathcal{G}} \quad \text{for “}A\text{ and }B\text{ are separated by }C\text{ in }\mathcal{G}\text{ under rule d,”}$$

and

$$(X_A \perp\!\!\!\perp X_B \mid X_C)_{\mathbb{P}} \quad \text{for probabilistic conditional independence under law }\mathbb{P}.$$

Properties of separations In the literature there are many different statements about how the graphical statements relate to the probabilistic statements in the (dynamic) Bayesian networks. Therefore, we explicitly mention them below to avoid confusion when navigating the results of related works.

Let $\mathfrak{BN}(\mathcal{G})$ denote a class of Bayesian network models Markov to $\mathcal{G}$ (i.e., distributions $\mathbb{P}$ that satisfy the Bayesian network factorisation/local Markov property with respect to $\mathcal{G}$).[3]

Definition 18 (Soundness for a model class) *The rule* d *is* sound *for* $\mathfrak{BN}(\mathcal{G})$ *if, for all disjoint* $A, B, C \subseteq \mathcal{V}$*,*

$$(A \perp_{\mathsf{d}} B \mid C)_{\mathcal{G}} \implies \forall\, \mathbb{P} \in \mathfrak{BN}(\mathcal{G}):\ (X_A \perp\!\!\!\perp X_B \mid X_C)_{\mathbb{P}}.$$

Definition 19 (Weak completeness (witness property)) *The rule* d *has* weak completeness *for* $\mathfrak{BN}(\mathcal{G})$ *if, for all disjoint* $A, B, C \subseteq \mathcal{V}$*,*

$$\neg(A \perp_{\mathsf{d}} B \mid C)_{\mathcal{G}} \implies \exists\, \mathbb{P} \in \mathfrak{BN}(\mathcal{G}):\ \neg(X_A \perp\!\!\!\perp X_B \mid X_C)_{\mathbb{P}}.$$

Soundness with weak completeness implies class-level completeness (by contraposition).

Definition 20 (Strong completeness) *Fix a parameterisation* $\{\mathbb{P}_\theta : \theta \in \Theta(\mathcal{G})\}$ *of* $\mathfrak{BN}(\mathcal{G})$ *with Lebesgue measure* λ *on* $\Theta(\mathcal{G})$*. The rule* d *is* strongly complete *if, for all disjoint* $A, B, C \subseteq \mathcal{V}$*,*

$$\neg(A \perp_{\mathsf{d}} B \mid C)_{\mathcal{G}} \implies \lambda\big(\big\{\theta \in \Theta(\mathcal{G}) : (X_A \perp\!\!\!\perp X_B \mid X_C)_{\mathbb{P}_\theta}\big\}\big) = 0.$$

Strong completeness implies weak completeness and formalises the idea that, absent graphical separation, any remaining independence is a non-generic (measure-zero) coincidence of parameters.

Definition 21 (Class-level completeness (Markov-perfectness)) *The rule* d *is* class-level complete *for* $\mathfrak{BN}(\mathcal{G})$ *if, for all disjoint* $A, B, C \subseteq \mathcal{V}$*,*

$$\Big(\forall\, \mathbb{P} \in \mathfrak{BN}(\mathcal{G}):\ (X_A \perp\!\!\!\perp X_B \mid X_C)_{\mathbb{P}}\Big) \iff (A \perp_{\mathsf{d}} B \mid C)_{\mathcal{G}}.$$

Definition 22 (Faithfulness (distribution-level)) *A distribution* $\mathbb{P}$ *is* faithful *to* $(\mathcal{G}, \mathsf{d})$ *if, for all disjoint* $A, B, C \subseteq \mathcal{V}$*,*

$$(X_A \perp\!\!\!\perp X_B \mid X_C)_{\mathbb{P}} \iff (A \perp_{\mathsf{d}} B \mid C)_{\mathcal{G}}.$$

Faithfulness is an ubiquitous, idealistic assumption made in causal discovery problems (see Boeken et al., 2024, for a discussion).

Definition 23 (Typicality of faithful laws) *Fix a parametrisation* $\{\mathbb{P}_\theta : \theta \in \Theta(\mathcal{G})\}$ *of* $\mathfrak{BN}(\mathcal{G})$ *with Lebesgue measure* λ*. The faithful laws are* measure-theoretically typical *if*

$$\lambda\big(\{\theta :\ \mathbb{P}_\theta \textit{ is unfaithful to } (\mathcal{G}, \mathsf{d})\}\big) = 0,$$

and topologically typical *if faithful laws form a dense open subset of* $\mathfrak{BN}(\mathcal{G})$ *(in a chosen topology) (Boeken et al., 2024).*

It is important to interpret precisely the notation and semantics of the above statement. The class-level completeness does not assert that for each distribution $(X_A \perp\!\!\!\perp X_B \mid X_C)_{\mathbb{P}} \Leftrightarrow (A \perp_{\mathsf{d}} B \mid C)_{\mathcal{G}}$. That per-distribution equivalence is exactly faithfulness. Class level is even weaker than the strong completeness. Class-level completeness uses the "if an only if" in the sense that it is false for all the class and not for each model in the class.

3. For Dynamic Bayesian Networks, replace $\mathcal{G}$ by the (space–time) DAG and the same definitions apply verbatim.

B.2. Interpretation of the covariance matrix and paths through graph

Lyapunov equation If the spectral radius satisfies $\rho(M) < 1$, then Equation 2 admits a unique strictly stationary solution with mean zero and covariance $\Sigma \in \mathrm{PD}_n$ solving the discrete Lyapunov equation

$$\Sigma = M\Sigma M^\top + D.$$

Equivalently, if $\Sigma_t := \mathrm{Cov}(X(t))$, then

$$\Sigma_{t+1} = M\Sigma_t M^\top + D, \qquad \Sigma_t = M^t\Sigma_0(M^\top)^t + \sum_{k=0}^{t-1} M^k D(M^\top)^k,$$

so under $\rho(M) < 1$,

$$\Sigma = \lim_{t\to\infty} \Sigma_t = \sum_{k=0}^{\infty} M^k\, D\, (M^\top)^k. \tag{3}$$

Since $D \succ 0$ and each term $M^k D(M^\top)^k$ is positive semidefinite. A simple sufficient condition for $\rho(M) < 1$ is $\|M\| < 1$ for any submultiplicative matrix norm; e.g., $\|M\|_\infty \le \max_i \sum_j |M_{ij}|$. If row i of M has at most d_i nonzeros and all allowed entries have $|M_{ij}| \le \eta$, then $\|M\|_\infty \le \max_i d_i\, \eta$, so choosing $\eta < (\max_i d_i)^{-1}$ guarantees $\rho(M) < 1$.

Pathwise interpretation. Observe that, informally, M^k accumulates information sent from the initial distribution $X(t_0)$ across all directed paths in space time diagram by k-moments further in time. Heuristically, the matrix M^k encodes causal dependencies in the covariance matrix. Therefore, if there were no direct paths from a node $(a, \cdot) \in \mathcal{V}_T$ to a node $(b, \cdot) \in \mathcal{V}_T$ whenever up to k moments in time from the start, then M^k_{ab} will be zero. This offers two powerful interpretations, one structural informing about zeros of the steady-state covariance matrix, the other combinational in illuminating how paths in the space-time diagram relate to the covariance matrix. Write the kth term of Equation 3 as

$$S_k \;:=\; M^k\, C\, (M^\top)^k, \qquad \Sigma = \sum_{k\ge 0} S_k, \qquad (S_k)_{ij} = \sum_{r,s=1}^{n} (M^k)_{ir}\, C_{rs}\, (M^k)_{js}.$$

When C is diagonal, $C = \mathrm{diag}(\sigma_1^2, \dots, \sigma_n^2)$,

$$(S_k)_{ij} = \sum_{s=1}^{n} \sigma_s^2\, (M^k)_{is}(M^k)_{js}. \tag{4}$$

For diagonal C, $(S_k)_{ij} = 0$ for all $M \in \mathbb{R}^E \iff$ for every node s there is no directed length-k walk from s to i and from s to j.

For the remainder of the paper, we make the assumption that the innovation matrix is diagonal to focus on the combinatorics in Equation 4, or set it to I all together. All completeness arguments in the following extend to general $C \succeq 0$ with the obvious modification of Equation 4.

First, the structural interpretation. The term $(S_k)_{ij}$ can be positive if after k time steps, there is at least one common ancestor s in the space-time diagram that influences both i and j along the directed paths in the space-time diagram. Therefore, from our interpretation, it is easy to see that the presence of a common ancestor in the full-time graph produces a nonzero contribution. Conversely, if every path between i and j has a collider between them, then every $(S_k)_{ij} = 0$ and $(S_k)_{ij} = 0$ and hence the sum yields $\Sigma_{ij} = 0$ and $\Sigma_{ji} = 0$.

Second, the combinatorial interpretation. Observation about $(S_k)_{ij}$ comes to light once we write matrix M in the masked form (Hadamaard product) $M = W \odot A$, with $A_{ij} \in \{0, 1\}$ encoding the incidence matrix. Then $\sum_s (A^k)_{is}(A^k)_{js}$ is the number of different paths in the space diagram from $(i, t_0 + k)$ to $(j, t_0 + k)$ of the following structure $(i, t_0 + k)\underbrace{\leftarrow \cdots \leftarrow}_{k}(l, t_0)\underbrace{\rightarrow \cdots \rightarrow}_{k}(j, t_0 + k)$ or $(j, t_0 + k)\underbrace{\leftarrow \cdots \leftarrow}_{k}(l, t_0)\underbrace{\rightarrow \cdots \rightarrow}_{k}(i, t_0 + k)$ for all $(l, t_0 + k)$. One could use such fact to control the covariance terms with combinatorial expressions, but in our experience the difficulty of expressing complex conditional independencies arises pretty fast.

B.3. Independence for Gaussian random variables

The following proposition gathers well-known results on the Gaussian variables characterization of (in)dependence.

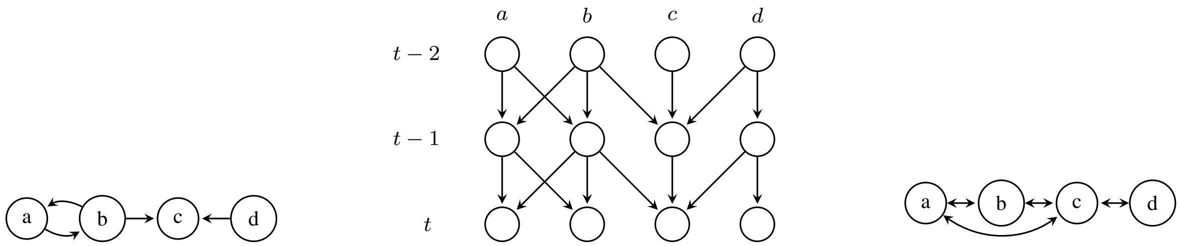
Proposition 24 (Independence) *Let $X \sim \mathcal{N}_n(\mu, \Sigma)$ be a Gaussian random vector. For disjoint subsets $I, J, K \subseteq [n]$, the following are equivalent:*

1. $X_I \perp X_J \mid X_K$.
2. $X_i \perp X_j \mid X_K$ *for all* $i \in I$ *and* $j \in J$.
3. $\det\big(\Sigma_{iK,jK}\big) = 0$ *for all* $i \in I$ *and* $j \in J$, *where* $\Sigma_{iK,jK}$ *denotes the submatrix of* Σ *with row indices* $\{i\} \cup K$ *and column indices* $\{j\} \cup K$.
4. $\operatorname{rank}(\Sigma_{A\cup K, B\cup K}) = |K|$.
5. $\Sigma_{ij} - \Sigma_{iK}\,\Sigma_{KK}^{-1}\,\Sigma_{Kj} = 0,$ *for all* $i \in I$ *and* $j \in J$.

Appendix C. Proofs and discussion for Section 3

We illustrate summary, space-time, and trek graphs on a small example, which will be useful for building intuition for Lemma 25 and Lemma 26.

Example 2 (Summary and space-time graphs) *Let G have $V = \{a, b, c, d\}$ and edges $a \to b$, $b \to a$, $b \to c$, $d \to c$ (self-loops implicit). The associated space-time graph $\mathcal{G}_T$ has vertices (v, s) and edges $(u, s) \to (v, s + 1)$ whenever $u \to v$ in $\mathcal{G}$. Therefore, these are consistent. Observe Figure 2.*



(a) Summary graph (b) Space-time DAG (c) Trek graph

Figure 2: Summary graph $\mathcal{G}$, the corresponding three-slice space-time DAG, and the trek graph. Self-loops are implicit and suppressed.

It helps to see that c–equivalence forces identical answers to all singleton queries, which (by Proposition 25) pins down all edges of the trek graph.

Proposition 25 (Bidirected edge $\iff$ failure of singleton c–separation) *For distinct $i, j \in V$ and any summary graph $\mathcal{G} = (V, E)$, and its respective trek graph $\widehat{\mathcal{G}} = (\mathcal{V}, \widehat{\mathcal{E}})$,*

$$i \leftrightarrow j \in \widehat{E} \quad\iff\quad \big(\{i\} \not\perp_c \{j\} \mid \varnothing\big)_{\mathcal{G}}.$$

Using these, we can spell out the lemma that proves the main lemma—Lemma 26.

Lemma 26 (Failure of c–separation $\iff$ path in trek graph avoiding S) *Let $I, J, K \subset V$ be pairwise disjoint and put $S := V \setminus (I \cup J \cup K)$. For any summary graph $\mathcal{G}$, and a corresponding trek graph $\widehat{\mathcal{G}}$*

$$(I \not\perp_c J \mid K)_{\mathcal{G}} \quad\iff\quad \text{there exists an } I\text{–}J \text{ path in } \widehat{\mathcal{G}} \text{ that avoids } S,$$

where avoiding S means that it only goes through $I \cup J \cup K$.

From the above, the main theorem about equivalence Theorem 5 follows. The theorem is reformulation really of Lemma 26, as it remains only to see the contrapositive of both sides.

Proposition 27 (Trek-graph edge $\iff$ existence of an st-trek) *Assume that every vertex of the summary graph $\mathcal{G}$ has a self-loop. Let $u, v \in \mathcal{V}$ be distinct and fix $t \in \mathcal{T}$. Then*

$$u \leftrightarrow v \in \widehat{\mathcal{G}} \quad\iff\quad \text{there exists an } st\text{-trek between } (u, t) \text{ and } (v, t) \text{ in } \mathcal{G}_T.$$

C.1. Deferred proofs of lemmas

Proof [Proof of Proposition 25]

($\Rightarrow$) If $i \leftrightarrow j \in \widehat{\mathcal{G}}$, then by definition there exists $a \in An_G(i) \cap An_G(j)$. Concatenating a directed path $a \to \cdots \to i$ with a directed path $a \to \cdots \to j$ yields a collider-free path between i and j. Hence

$$(\{i\} \not\perp_c \{j\} \mid \emptyset)_{\mathcal{G}}.$$

($\Leftarrow$) Assume

$$(\{i\} \not\perp_c \{j\} \mid \emptyset)_{\mathcal{G}}.$$

Then there exists a collider-free path

$$p = (v_0 = i, v_1, \dots, v_m = j)$$

in $\mathcal{G}$. Along a collider-free path, the orientation pattern can change from "towards the left" to "towards the right" at most once; otherwise an interior collider would occur. Therefore either

$$i \leftarrow \cdots \leftarrow a \rightarrow \cdots \rightarrow j$$

for some vertex a on p, or the whole path is directed from i to j, or the whole path is directed from j to i. In the first case $a \in An_G(i) \cap An_G(j)$; in the second case $i \in An_G(i) \cap An_G(j)$; in the third case $j \in An_G(i) \cap An_G(j)$. Hence $i \leftrightarrow j \in \widehat{\mathcal{G}}$. ■

Proof [Proof of Lemma 26] We use the following convention inside the proof: if the walk uses the oriented edge $u \rightarrow v$, then that edge contributes a *tail* at u and a *head* at v. If both $u \rightarrow v$ and $v \rightarrow u$ are present, we write $u \leftrightarrow v$ as shorthand for the reciprocal pair.

($\Rightarrow$) Assume

$$(I \not\perp_c J \mid K)_{\mathcal{G}}.$$

Then there exists a c-open path

$$\tau = (v_0, \dots, v_r)$$

in $\mathcal{G}$ from $v_0 = i \in I$ to $v_r = j \in J$ given K.

If τ has no collider, then it is collider-free, so

$$i \leftrightarrow j \in \widehat{\mathcal{G}}$$

by Proposition 25, and we are done.

Assume now that τ has colliders. Let

$$c_1, \dots, c_m$$

be the colliders on τ in the order in which they appear from i to j. Since τ is c-open given K, each $c_q \in \mathrm{An}_{\mathcal{G}}(K)$. Choose

$$k_q \in K \cap \mathrm{De}_{\mathcal{G}}(c_q), \qquad q = 1, \dots, m.$$

We claim that

$$i \leftrightarrow k_1, \qquad k_q \leftrightarrow k_{q+1} \;\; (1 \le q < m), \qquad k_m \leftrightarrow j$$

are edges of $\widehat{\mathcal{G}}$.

For the first edge, the segment of τ from i to c_1 has no collider. Appending a directed path $c_1 \rightarrow \cdots \rightarrow k_1$ gives a collider-free path from i to k_1. Hence $i \leftrightarrow k_1$ by Proposition 25.

For $k_q \leftrightarrow k_{q+1}$, concatenate

$$k_q \leftarrow \cdots \leftarrow c_q,$$

the segment of τ from c_q to c_{q+1}, and

$$c_{q+1} \rightarrow \cdots \rightarrow k_{q+1}.$$

Because c_q and c_{q+1} are consecutive colliders on τ, the middle segment has no collider. At c_q, one incident edge has a tail and the other a head, so c_q is not a collider on the concatenated path; the same holds at c_{q+1}. Thus the whole concatenation is collider-free, so $k_q \leftrightarrow k_{q+1} \in \widehat{\mathcal{G}}$ by Proposition 25.

The proof of $k_m \leftrightarrow j$ is analogous. Therefore

$$i, k_1, \ldots, k_m, j$$

is an I-J walk in $\widehat{\mathcal{G}}$ whose vertices all lie in $I \cup J \cup K$. Deleting repeated vertices yields an I-J path in $\widehat{\mathcal{G}}$ that avoids

$$S = V \setminus (I \cup J \cup K).$$

($\Leftarrow$) Assume that $\widehat{\mathcal{G}}$ contains an I-J path avoiding S. Among all such paths, choose one of minimal length:

$$\pi = (d_0, \ldots, d_m), \qquad d_0 \in I,\ d_m \in J, \qquad d_0, \ldots, d_m \in I \cup J \cup K.$$

Then every internal vertex lies in K. Indeed, if some internal $d_r \in I$, then $(d_r, \ldots, d_m)$ would be a shorter I-J path; if some internal $d_r \in J$, then $(d_0, \ldots, d_r)$ would be a shorter one.

Thus

$$d_1, \ldots, d_{m-1} \in K.$$

For each edge $d_{r-1} \leftrightarrow d_r$ of π, choose a collider-free path γ_r in $\mathcal{G}$ from d_{r-1} to d_r, which exists by Proposition 25. Concatenate these paths:

$$W := \gamma_1 \cdot \gamma_2 \cdots \gamma_m.$$

This is a walk from d_0 to d_m.

Each γ_r is collider-free, so any collider of W can occur only at a stitching vertex d_r, $1 \leq r \leq m-1$. Since each such $d_r \in K \subseteq \mathrm{An}_{\mathcal{G}}(K)$, the walk W is c-open given K.

It remains to show that a c-open walk contains a c-open path. Among all c-open walks from d_0 to d_m given K, choose one with the smallest number of vertices; we continue to call it $W = (w_0, \ldots, w_n)$. We claim that W has no repeated vertex.

Suppose $w_i = w_j$ for some $i < j$.

If $i = 0$, then

$$(w_j, w_{j+1}, \ldots, w_n)$$

is a shorter c-open walk with the same endpoints, because $w_j = w_0$ is now an endpoint and endpoint status is irrelevant for c-openness. This contradicts minimality. Similarly, if $j = n$, then

$$(w_0, \ldots, w_i)$$

is a shorter c-open walk, contradiction.

Hence we may assume

$$0 < i < j < n.$$

Delete the closed subwalk from the first occurrence of w_i to the second and form

$$W' = (w_0, \ldots, w_i, w_{j+1}, \ldots, w_n).$$

This is again a walk from d_0 to d_m. Every internal vertex of W' other than w_i has the same collider/non-collider status as in W, so only w_i needs checking.

If w_i is not a collider on W', then W' is a shorter c-open walk, contradiction. So suppose w_i is a collider on W'.

If w_i was already a collider at the first occurrence or at the second occurrence in W, then $w_i \in \mathrm{An}_{\mathcal{G}}(K)$ because W is c-open, and again W' is c-open, contradiction. Therefore w_i was a non-collider at both occurrences in W.

Since w_i is a collider on W', the edge used between w_{i-1} and w_i has a head at w_i, and the edge used between w_i and w_{j+1} also has a head at w_i. Because the first occurrence of w_i was a non-collider in W, the edge used between w_i and w_{i+1} must have a tail at w_i. Because the second occurrence of $w_i = w_j$ was a non-collider in W, the edge used between w_{j-1} and $w_j = w_i$ must also have a tail at w_i.

Now consider the deleted closed subwalk

$$C = (w_i, w_{i+1}, \dots, w_j),$$

whose first and last vertices are both w_i. In C, the first used edge has a tail at the initial w_i, and the last used edge has a tail at the final w_i. We show that C contains an internal collider z such that $w_i \in \mathrm{An}_{\mathcal{G}}(z)$.

Indeed, let $s \in \{0, \dots, j-i-1\}$ be minimal such that the s-th used edge of C has a head at its left endpoint. Such an s exists, because otherwise every used edge of C would have a tail at its left endpoint, so the last used edge would have a head, not a tail, at the final w_i, contradiction. By minimality, all earlier used edges of C have tails at their left endpoints, so the initial segment of C is a directed path starting at w_i and ending at the corresponding left endpoint z. The preceding used edge has a head at z, and by choice of s the next used edge also has a head at z. Thus z is an internal collider on C, and

$$w_i \in \mathrm{An}_{\mathcal{G}}(z).$$

Since C is part of the c-open walk W, every collider on C lies in $\mathrm{An}_{\mathcal{G}}(K)$. In particular $z \in \mathrm{An}_{\mathcal{G}}(K)$, hence also

$$w_i \in \mathrm{An}_{\mathcal{G}}(K).$$

Therefore W' is c-open, contradiction.

So the minimal c-open walk W has no repeated vertices. Hence W is a c-open path from $d_0 \in I$ to $d_m \in J$ given K. Therefore

$$(I \not\perp_c J \mid K)_{\mathcal{G}}.$$

This completes the proof. ∎

Proof [Proof of Lemma 4]
$(\Rightarrow)$ If $\mathcal{G}_1 \equiv_c \mathcal{G}_2$, then for every distinct $i, j \in V$ we have $(\{i\} \perp_c \{j\} \mid \varnothing)_{\mathcal{G}_1} \iff (\{i\} \perp_c \{j\} \mid \varnothing)_{\mathcal{G}_2}$. By Proposition 25, $i \leftrightarrow j$ is present in $\widehat{\mathcal{G}}_1$ if and only if it is present in $\widehat{\mathcal{G}}_2$, hence $\widehat{\mathcal{G}}_1 = \widehat{\mathcal{G}}_2$.
$(\Leftarrow)$ If $\widehat{\mathcal{G}}_1 = \widehat{\mathcal{G}}_2$, then for every disjoint triple (I, J, K) the existence of an I–J path in the common trek graph that avoids $S = V \setminus (I \cup J \cup K)$ is the same for $\mathcal{G}_1$ and $\mathcal{G}_2$. Lemma 26 then yields $(I \perp_c J \mid K)_{\mathcal{G}_1} \iff (I \perp_c J \mid K)_{\mathcal{G}_2}$. ∎

Proof [Proof of Theorem 5]
By Lemma 26, $(I \not\perp_c J \mid K)_G \iff \exists$ an I–J path in $\widehat{\mathcal{G}}$ avoiding S. Taking contrapositives on both sides gives $(I \perp_c J \mid K)_G \iff$ every I–J path in $\widehat{\mathcal{G}}$ meets S, which is exactly $(I \perp J \mid S)_{\widehat{\mathcal{G}}}$ in the trek graph (coinciding with undirected) sense. ∎

Proof [Proof of Proposition 27] ($\Rightarrow$) Assume $u \leftrightarrow v \in \widehat{\mathcal{G}}$. By definition of the trek graph, there exists

$$a \in \mathrm{An}_{\mathcal{G}}(u) \cap \mathrm{An}_{\mathcal{G}}(v).$$

Hence there are directed paths in $\mathcal{G}$

$$P_u : a \rightsquigarrow u, \qquad P_v : a \rightsquigarrow v.$$

Let $|P_u| = \ell_u$ and $|P_v| = \ell_v$. Set

$$k := \max\{\ell_u, \ell_v, 1\}.$$

Because every vertex has a self-loop, we may prepend $k - \ell_u$ copies of the self-loop $a \to a$ to P_u and $k - \ell_v$ copies of the self-loop $a \to a$ to P_v. This yields two directed paths in the space-time graph

$$\widetilde{P}_u : (a, t-k) \rightsquigarrow (u, t), \qquad \widetilde{P}_v : (a, t-k) \rightsquigarrow (v, t),$$

both of lag k. Therefore $(\widetilde{P}_u, \widetilde{P}_v)$ is an st-trek between (u, t) and (v, t).

($\Leftarrow$) Assume there exists an st-trek of lag k between (u, t) and (v, t), with top $(a, t-k)$. By definition, the two legs are directed paths in $\mathcal{G}_T$ from $(a, t-k)$ to (u, t) and $(a, t-k)$ to (v, t). Projecting these two legs to the spatial graph gives directed walks in $\mathcal{G}$ from a to u and from a to v. By deleting repeated vertices if necessary, each directed walk contains a directed path with the same endpoints. Hence

$$a \in \mathrm{An}_{\mathcal{G}}(u) \cap \mathrm{An}_{\mathcal{G}}(v),$$

so $u \leftrightarrow v \in \widehat{\mathcal{G}}$. ■

Proof [Proof of Theorem 9] It suffices to prove the claim pairwise. Indeed, both c-separation and space-time separation between sets are defined by the absence of a witness for every pair $i \in I$, $j \in J$. Thus it is enough to show that for all distinct $i, j \in \mathcal{V}$ and all $K \subseteq \mathcal{V} \setminus \{i, j\}$,

$$(i \perp_c j \mid K)_{\mathcal{G}} \quad \Longleftrightarrow \quad (i \perp_{st} j \mid K)_{\mathcal{G}_T}.$$

We prove the equivalent failure statement.

($\Rightarrow$) Assume

$$(i \not\perp_c j \mid K)_{\mathcal{G}}.$$

By Lemma 26, there exists an i-j path

$$q_0 = i,\ q_1, \dots, q_r = j$$

in the trek graph $\widehat{\mathcal{G}}$ whose internal vertices lie in K. For each edge $q_{p-1} \leftrightarrow q_p$, Proposition 27 gives an st-trek between (q_{p-1}, t) and (q_p, t). Therefore

$$\mu(i, j) := (q_0, \dots, q_r)$$

is a multi-trek from i to j. Its junctions are $q_1, \dots, q_{r-1}$, and these lie in $K \subseteq \mathrm{An}_{\mathcal{G}}(K)$. Hence $\mu(i, j)$ is active given K, so

$$(i \not\perp_{st} j \mid K)_{\mathcal{G}_T}.$$

($\Leftarrow$) Assume

$$(i \not\perp_{st} j \mid K)_{\mathcal{G}_T}.$$

Then there exists an active multi-trek

$$\mu(i,j) = (d_0 = i, d_1, \dots, d_m = j)$$

such that each consecutive pair d_{r-1}, d_r is joined by an st-trek, and each junction $d_1, \dots, d_{m-1}$ lies in $\mathrm{An}_{\mathcal{G}}(K)$.

If $m = 1$, then i and j are joined by a single st-trek. By Proposition 27,

$$i \leftrightarrow j \in \widehat{\mathcal{G}}.$$

Hence there is an i-j path in $\widehat{\mathcal{G}}$ with no internal vertices, and Lemma 26 gives

$$(i \not\perp_c j \mid K)_{\mathcal{G}}.$$

Assume now that $m \geq 2$. For each $r = 1, \dots, m$, the consecutive pair d_{r-1}, d_r is joined by an st-trek, so by Proposition 27,

$$d_{r-1} \leftrightarrow d_r \in \widehat{\mathcal{G}}.$$

For each junction d_r with $1 \leq r \leq m-1$, choose

$$k_r \in K \cap \mathrm{De}_{\mathcal{G}}(d_r),$$

which is possible because $d_r \in \mathrm{An}_{\mathcal{G}}(K)$.

We claim that

$$i \leftrightarrow k_1, \qquad k_r \leftrightarrow k_{r+1} \;\; (1 \leq r \leq m-2), \qquad k_{m-1} \leftrightarrow j$$

are edges of $\widehat{\mathcal{G}}$.

For the first edge, let a_1 be any common ancestor witnessing $d_0 = i \leftrightarrow d_1$. Since $d_1 \in \mathrm{An}_{\mathcal{G}}(k_1)$, the same vertex a_1 is a common ancestor of i and k_1. Hence $i \leftrightarrow k_1 \in \widehat{\mathcal{G}}$.

For the middle edges, let a_{r+1} be any common ancestor witnessing $d_r \leftrightarrow d_{r+1}$. Since

$$d_r \in \mathrm{An}_{\mathcal{G}}(k_r), \qquad d_{r+1} \in \mathrm{An}_{\mathcal{G}}(k_{r+1}),$$

the same vertex a_{r+1} is a common ancestor of k_r and k_{r+1}. Hence

$$k_r \leftrightarrow k_{r+1} \in \widehat{\mathcal{G}}.$$

The proof of $k_{m-1} \leftrightarrow j$ is analogous.

Therefore

$$i, \; k_1, \dots, k_{m-1}, \; j$$

is an i-j walk in $\widehat{\mathcal{G}}$ whose internal vertices all lie in K. Deleting repeated vertices yields an i-j path in $\widehat{\mathcal{G}}$ whose internal vertices lie in K. By Lemma 26,

$$(i \not\perp_c j \mid K)_{\mathcal{G}}.$$

We have proved

$$(i \not\perp_c j \mid K)_{\mathcal{G}} \quad \Longleftrightarrow \quad (i \not\perp_{st} j \mid K)_{\mathcal{G}_T}.$$

Taking contrapositives gives

$$(i \perp_c j \mid K)_{\mathcal{G}} \quad \Longleftrightarrow \quad (i \perp_{st} j \mid K)_{\mathcal{G}_T}.$$

Since this holds for every $i \in I$, $j \in J$, the set-valued statement follows. ■

Appendix D. Proofs for Section 4

Throughout this appendix, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a summary graph on $\mathcal{V} = [n]$, the innovation covariance is diagonal,

$$D = \operatorname{diag}(d_1, \dots, d_n) \succ 0,$$

and

$$\Theta_{\mathcal{G}} := \{(M, D) : \rho(M) < 1,\ D = \operatorname{diag}(d_v) \succ 0,\ M \in \mathbb{R}^{\mathcal{E}}\}.$$

For $(M, D) \in \Theta_{\mathcal{G}}$, let

$$\Sigma(M, D) = \sum_{k=0}^{\infty} M^k D (M^\top)^k$$

denote the stationary covariance matrix of the corresponding stable Gaussian VAR(1) process.

We also use the matrices

$$X = [D^{1/2},\ MD^{1/2},\ M^2 D^{1/2}, \dots], \qquad Y = [I, M, M^2, \dots],$$

whose columns are indexed by top–lag pairs $(u, k) \in \mathcal{V} \times \mathbb{N}_0$. Thus

$$X_{v,(u,k)} = \sqrt{d_u}\,(M^k)_{vu}, \qquad Y_{v,(u,k)} = (M^k)_{vu}.$$

The lag-zero columns satisfy

$$X_{v,(u,0)} = \sqrt{d_u}\,\mathbf{1}_{\{u=v\}}.$$

D.1. Gaussian minors and lag-zero columns

Proof [Proof of Lemma 12] This is a standard result, but we copy the proof for reference. Let $X \sim \mathcal{N}_n(0, \Sigma)$, and let $I, J, K \subseteq \mathcal{V}$ be pairwise disjoint.

For fixed $i \in I$, $j \in J$, define

$$I' = \{i\} \cup K, \qquad J' = \{j\} \cup K.$$

Write the submatrix $\Sigma_{I',J'}$ in block form as

$$\Sigma_{I',J'} = \begin{pmatrix} \Sigma_{ij} & \Sigma_{iK} \\ \Sigma_{Kj} & \Sigma_{KK} \end{pmatrix}.$$

Since $\Sigma \succ 0$, the principal submatrix Σ_{KK} is invertible. Hence, by the Schur complement formula,

$$\det \Sigma_{I',J'} = \det(\Sigma_{KK})\,\big(\Sigma_{ij} - \Sigma_{iK}\Sigma_{KK}^{-1}\Sigma_{Kj}\big).$$

Because $\Sigma_{KK} \succ 0$, $\det(\Sigma_{KK}) > 0$. Therefore

$$\det \Sigma_{I',J'} = 0 \quad\Longleftrightarrow\quad \Sigma_{ij} - \Sigma_{iK}\Sigma_{KK}^{-1}\Sigma_{Kj} = 0.$$

For Gaussian vectors, the right-hand side is exactly the condition

$$X_i \perp\!\!\!\perp X_j \mid X_K.$$

Thus

$$X_i \perp\!\!\!\perp X_j \mid X_K \quad \Longleftrightarrow \quad \det \Sigma_{\{i\}\cup K,\{j\}\cup K} = 0.$$

Finally, for Gaussian vectors,

$$X_I \perp\!\!\!\perp X_J \mid X_K \quad \Longleftrightarrow \quad X_i \perp\!\!\!\perp X_j \mid X_K \quad \text{for all } i \in I,\ j \in J,$$

which proves the lemma. ■

Lemma 28 (Lag-zero reduction for X-minors) *Let $T \subseteq \mathcal{V}$, let*

$$S = S_+ \sqcup S_0 \subseteq \mathcal{V} \times \mathbb{N}_0,$$

where

$$S_0 = \{(u,0) : u \in U\}$$

for some $U \subseteq \mathcal{V}$, and $S_+ \subseteq \mathcal{V} \times \mathbb{N}$. Then:

1. *If $U \not\subseteq T$, then* $\det X_{T,S} = 0$.

2. *If $U \subseteq T$, then after ordering the rows as $(T \setminus U,\, U)$ and the columns as $(S_+,\, S_0)$,*

$$X_{T,S} = \begin{pmatrix} X_{T\setminus U,S_+} & 0 \\ * & \operatorname{diag}(\sqrt{d_u})_{u\in U} \end{pmatrix}.$$

Consequently,

$$\det X_{T,S} = \pm\Big(\prod_{u\in U} \sqrt{d_u}\Big) \det X_{T\setminus U,S_+}.$$

Proof For a lag-zero column $(u,0)$, we have

$$X_{v,(u,0)} = \sqrt{d_u}\,\mathbf{1}_{\{u=v\}}.$$

Hence the column $(u,0)$ has exactly one nonzero entry, namely $\sqrt{d_u}$ in row u. If $u \notin T$, then this column is zero in the submatrix $X_{T,S}$, and therefore $\det X_{T,S} = 0$.

Assume now $U \subseteq T$. Ordering rows and columns as stated, every lag-zero column contributes only in the corresponding U-row, and the upper-right block is zero. The lower-right block is diagonal with diagonal entries $\sqrt{d_u}$, $u \in U$. The determinant formula follows from block triangularity. ■

D.2. Cauchy–Binet and path-system minors

Proof [Proof of Lemma 13] Fix $I, J \subseteq \mathcal{V}$ with $|I| = |J| = \ell$.

Step 1: finite truncations. For $N \geq 0$, define

$$X^{(N)} = [D^{1/2},\, MD^{1/2},\, \dots,\, M^N D^{1/2}], \qquad \Sigma^{(N)} := X^{(N)}(X^{(N)})^\top.$$

Then

$$\Sigma^{(N)} = \sum_{k=0}^{N} M^k D (M^\top)^k.$$

By finite-dimensional Cauchy–Binet,

$$\det \Sigma^{(N)}_{I,J} = \sum_{\substack{S \subseteq \mathcal{V} \times \{0,\dots,N\} \\ |S|=\ell}} \det X^{(N)}_{I,S}\, \det X^{(N)}_{J,S}.$$

Step 2: passage to the infinite sum. We show that

$$\det \Sigma_{I,J} = \sum_{\substack{S \subseteq \mathcal{V} \times \mathbb{N}_0 \\ |S|=\ell}} \det X_{I,S}\, \det X_{J,S},$$

and that the series is absolutely convergent.

Since $\rho(M) < 1$, there exists a matrix norm $\|\cdot\|$ and some $r \in (0,1)$ such that $\|M\| \leq r$. Therefore, for every $k \geq 0$,

$$\|M^k\| \leq r^k.$$

As all norms are equivalent in finite dimension, there exists $C > 0$ such that

$$|(M^k)_{vu}| \leq C r^k \qquad \text{for all } u, v \in \mathcal{V},\ k \geq 0.$$

Hence every column (u,k) of X has Euclidean norm bounded by

$$\|X_{\cdot,(u,k)}\|_2 \leq \sqrt{n}\, C\, \sqrt{d_u}\, r^k.$$

By Hadamard's inequality, for every ℓ-element column set S,

$$|\det X_{I,S}| \leq C_1 \prod_{(u,k)\in S} r^k, \qquad |\det X_{J,S}| \leq C_1 \prod_{(u,k)\in S} r^k$$

for some constant C_1 depending only on I, J, M, D. Therefore

$$|\det X_{I,S} \det X_{J,S}| \leq C_1^2 \prod_{(u,k)\in S} r^{2k}.$$

The series

$$\sum_{\substack{S \subseteq \mathcal{V} \times \mathbb{N}_0 \\ |S|=\ell}} \prod_{(u,k)\in S} r^{2k}$$

converges because $\sum_{k\geq 0} r^{2k} < \infty$. Hence the Cauchy–Binet series converges absolutely, and letting $N \to \infty$ in the truncated expansion yields

$$\det \Sigma_{I,J} = \sum_{\substack{S \subseteq \mathcal{V} \times \mathbb{N}_0 \\ |S|=\ell}} \det X_{I,S}\, \det X_{J,S}.$$

Step 3: grouping by multiplicity vectors. Write

$$X = Y\Delta,$$

where Δ is the diagonal matrix indexed by $\mathcal{V} \times \mathbb{N}_0$ with

$$\Delta_{(u,k),(u,k)} = \sqrt{d_u}.$$

For a finite column set $S = \{(v_1, k_1), \dots, (v_\ell, k_\ell)\}$, we have

$$\det X_{I,S} = \det Y_{I,S} \det \Delta_S, \qquad \det X_{J,S} = \det Y_{J,S} \det \Delta_S.$$

Thus

$$\det X_{I,S} \det X_{J,S} = \det Y_{I,S} \det Y_{J,S} \prod_{p=1}^{\ell} d_{v_p}.$$

If

$$m_v(S) := \big|\{p : \ v_p = v\}\big|,$$

then

$$\prod_{p=1}^{\ell} d_{v_p} = \prod_{v \in \mathcal{V}} d_v^{m_v(S)}.$$

Grouping all column sets with the same multiplicity vector gives

$$\det \Sigma_{I,J} = \sum_{\substack{m \in \mathbb{N}_0^{\mathcal{V}} \\ \sum_v m_v = \ell}} \left(\sum_{\substack{S \subseteq \mathcal{V} \times \mathbb{N}_0 \\ |S| = \ell,\ m(S) = m}} \det Y_{I,S} \ \det Y_{J,S} \right) \prod_{v \in \mathcal{V}} d_v^{m_v}.$$

If the d_v are treated as algebraically independent, then the monomials $\prod_v d_v^{m_v}$ are linearly independent, so

$$\det \Sigma_{I,J} \equiv 0 \quad \iff \quad \sum_{\substack{S \subseteq \mathcal{V} \times \mathbb{N}_0 \\ |S| = \ell,\ m(S) = m}} \det Y_{I,S} \ \det Y_{J,S} \equiv 0 \quad \text{for every } m.$$

■

Proof [Proof of Lemma 14] The Lindström-Gessel–Viennot applies directly. For completeness we give a proof sketch. Let $R = \{(v_1, k_1), \dots, (v_\ell, k_\ell)\} \subseteq \mathcal{V} \times \mathbb{N}$ be a set of distinct positive-lag top–lag pairs, and let $S = \{s_1, \dots, s_\ell\} \subseteq \mathcal{V}$.

We first prove the formula for Y, then multiply by the $\sqrt{d_{v_p}}$-factors.

For each $p \in \{1, \dots, \ell\}$ and each sink $s \in S$,

$$Y_{s,(v_p,k_p)} = (M^{k_p})_{s v_p}.$$

Because the space-time graph is acyclic and every directed path from $(v_p, t - k_p)$ to (s, t) has exactly k_p edges, we may expand

$$(M^{k_p})_{s v_p} = \sum_{P \in \mathcal{P}_{k_p}(v_p \rightsquigarrow s)} M_P,$$

where

$$M_P := \prod_{(u,r-1)\to(w,r)\in P} M_{wu}.$$

Therefore, by Leibniz,

$$\det Y_{S,R} = \sum_{\sigma\in\mathfrak{S}_\ell} \operatorname{sgn}(\sigma) \prod_{p=1}^{\ell} \left(\sum_{P_p\in\mathcal{P}_{k_p}(v_p\rightsquigarrow s_{\sigma(p)})} M_{P_p} \right).$$

Distributing the product gives

$$\det Y_{S,R} = \sum_{\sigma\in\mathfrak{S}_\ell} \operatorname{sgn}(\sigma) \sum_{\mathbf{P}\in\mathcal{S}_\sigma(R,S)} \prod_{p=1}^{\ell} M_{P_p},$$

where $\mathcal{S}_\sigma(R,S)$ is the set of all path systems that route (v_p, k_p) to $s_{\sigma(p)}$.

We now cancel all intersecting systems by the usual tail-swapping involution. If a system $\mathbf{P} = (P_1, \dots, P_\ell)$ has an intersection, let (u^*, r^*) be the first space-time vertex at which two paths meet, and among all such pairs choose the lexicographically smallest $p < q$. Write

$$P_p = \alpha_p \cdot \beta_p, \qquad P_q = \alpha_q \cdot \beta_q,$$

where α_p, α_q end at (u^*, r^*), and β_p, β_q are the remaining suffixes. Define

$$P'_p = \alpha_p \cdot \beta_q, \qquad P'_q = \alpha_q \cdot \beta_p,$$

and keep all other paths unchanged.

The new system is again valid because the two suffixes begin at the same space-time vertex. It remains intersecting, the map is an involution, and it has no fixed points. The sink assignment of paths p and q is swapped, so the permutation sign changes by a transposition:

$$\operatorname{sgn}(\sigma') = -\operatorname{sgn}(\sigma).$$

The product of path weights is unchanged:

$$\prod_{p=1}^{\ell} M_{P'_p} = \prod_{p=1}^{\ell} M_{P_p},$$

since the same two suffixes are merely reassigned. Thus all intersecting systems cancel in pairs, leaving only the vertex-disjoint systems:

$$\det Y_{S,R} = \sum_{\mathbf{P}\in\mathcal{N}(R,S)} \operatorname{sgn}(\sigma(\mathbf{P})) \prod_{p=1}^{\ell} M_{P_p}.$$

Finally, each column (v_p, k_p) of X is $\sqrt{d_{v_p}}$ times the corresponding column of Y. Hence

$$\det X_{S,R} = \left(\prod_{p=1}^{\ell} \sqrt{d_{v_p}} \right) \det Y_{S,R} = \sum_{\mathbf{P}\in\mathcal{N}(R,S)} \operatorname{sgn}(\sigma(\mathbf{P})) \prod_{p=1}^{\ell} \sqrt{d_{v_p}}\, M_{P_p}.$$

In particular, if $\mathcal{N}(R,S) = \varnothing$, then $\det X_{S,R} = 0$. ■

D.3. Canonical clean chains and monomial isolation

Lemma 29 (Shortest trek-graph paths are chordless) *Let*

$$q_0, \dots, q_r$$

be a shortest path in the trek graph $\widehat{\mathcal{G}}$. Then no two nonconsecutive vertices on this path are adjacent in $\widehat{\mathcal{G}}$.

Proof If q_p and q_q with $|p-q| \geq 2$ were adjacent, then the subpath $q_p, \dots, q_q$ could be replaced by the single edge $q_p \leftrightarrow q_q$, producing a shorter path between q_0 and q_r, contradiction. ■

Proof [Proof of Lemma 15] Assume

$$i \not\perp_{st} j \mid K.$$

By Theorem 9, this is equivalent to

$$i \not\perp_{c} j \mid K.$$

By Theorem 5, there exists a path

$$q_0 = i,\ q_1, \dots, q_r = j$$

in the trek graph $\widehat{\mathcal{G}}$ with all internal vertices in K. Fix one of minimal length. By Lemma 29, it is chordless.

For each $p = 1, \dots, r$, choose a top

$$u_p \in \mathrm{An}_{\mathcal{G}}(q_{p-1}) \cap \mathrm{An}_{\mathcal{G}}(q_p).$$

Choose directed paths

$$L_p : u_p \rightsquigarrow q_{p-1}, \qquad R_p : u_p \rightsquigarrow q_p$$

with $|L_p| + |R_p|$ minimal among all such pairs of branches from u_p to q_{p-1} and q_p.

We verify the required properties.

(1) L_p and R_p meet only at u_p. Suppose they share a vertex $v \neq u_p$. Let v be the last common vertex encountered when moving away from u_p. Then the subpaths from v to q_{p-1} and q_p have strictly smaller total length than $|L_p| + |R_p|$, contradicting minimality. Hence $L_p \cap R_p = \{u_p\}$.

(2) Top isolation. We claim that

$$u_p \notin \mathrm{An}_{\mathcal{G}}(q_s) \qquad \text{for every } s \notin \{p-1, p\}.$$

If $s < p-1$, then u_p is an ancestor of both q_s and q_p, so $q_s \leftrightarrow q_p$ in the trek graph. Since $p - s \geq 2$, this is a chord, contradicting Lemma 29.

If $s > p$, then u_p is an ancestor of both q_{p-1} and q_s, so $q_{p-1} \leftrightarrow q_s$, again a chord. This proves top isolation.

(3) Distinct tops. Suppose $u_p = u_q$ with $p < q$. If $q = p+1$, then $u_p = u_q$ is a common ancestor of q_{p-1} and q_{p+1}, so $q_{p-1} \leftrightarrow q_{p+1}$, a chord. If $q \geq p+2$, then $u_p = u_q$ is a common ancestor of q_p and q_q, again a chord. Thus all tops are distinct.

(4) Branch interiors avoid the rest of the trek-graph path. Suppose some $q_s \notin \{q_{p-1}, q_p\}$ lies in the interior of L_p. Then $u_p \in \mathrm{An}_{\mathcal{G}}(q_s)$, contradicting top isolation. The same argument applies to R_p.

(5) Left-disjointness. Suppose L_p and L_q share a vertex v with $p < q$. Since $v \in L_q$, we have $u_q \in \mathrm{An}(v)$. Since $v \in L_p$, we have $v \in \mathrm{An}(q_{p-1})$. Therefore

$$u_q \in \mathrm{An}(q_{p-1}).$$

But $u_q \in \mathrm{An}(q_q)$ as well, so $q_{p-1} \leftrightarrow q_q$ in the trek graph. Because $q - (p-1) \geq 2$, this is a chord, contradiction.

(6) Right-disjointness. Suppose R_p and R_q share a vertex v with $p < q$. Since $v \in R_p$, we have $u_p \in \mathrm{An}(v)$. Since $v \in R_q$, we have $v \in \mathrm{An}(q_q)$. Therefore

$$u_p \in \mathrm{An}(q_q).$$

But $u_p \in \mathrm{An}(q_{p-1})$ as well, so $q_{p-1} \leftrightarrow q_q$, again a chord, contradiction.

Thus the L_p are pairwise vertex-disjoint and the R_p are pairwise vertex-disjoint.

(7) Padding. Choose integers

$$k_p \geq \max\{|L_p|, |R_p|, 1\} \qquad (p = 1, \ldots, r).$$

Pad each branch by self-loops at the top u_p to obtain a k_p-padded lift in the space-time graph. Because the tops u_p are distinct and same-sided branches are disjoint, the resulting left legs are pairwise vertex-disjoint and the resulting right legs are pairwise vertex-disjoint. Hence the padded lifts form a clean ordered system of positive-lag st-treks. ■

Proof [Proof of Lemma 16] Assume the hypotheses of Lemma 15. Write

$$K_0 := K \setminus \{q_1, \ldots, q_{r-1}\}, \qquad I_{\mathrm{ch}} := \{q_0, \ldots, q_{r-1}\}, \qquad J_{\mathrm{ch}} := \{q_1, \ldots, q_r\}.$$

Then

$$\{i\} \cup K = I_{\mathrm{ch}} \sqcup K_0, \qquad \{j\} \cup K = J_{\mathrm{ch}} \sqcup K_0.$$

Define

$$S^\star_+ := \{(u_1, k_1), \ldots, (u_r, k_r)\}, \qquad S^\star_0 := \{(k, 0) : k \in K_0\}, \qquad S^\star := S^\star_+ \cup S^\star_0.$$

Order the rows and columns as

$$(\{i\} \cup K) = (I_{\mathrm{ch}}, K_0), \qquad (\{j\} \cup K) = (J_{\mathrm{ch}}, K_0), \qquad S^\star = (S^\star_+, S^\star_0).$$

Step 1: block form and triangularity. By Lemma 28,

$$X_{\{i\}\cup K, S^\star} = \begin{pmatrix} X_{I_{\mathrm{ch}}, S^\star_+} & 0 \\ * & \mathrm{diag}(\sqrt{d_k})_{k \in K_0} \end{pmatrix}, \qquad X_{\{j\}\cup K, S^\star} = \begin{pmatrix} X_{J_{\mathrm{ch}}, S^\star_+} & 0 \\ * & \mathrm{diag}(\sqrt{d_k})_{k \in K_0} \end{pmatrix}.$$

We show that the chain blocks are triangular.

The (s,p)-entry of $X_{I_{\mathrm{ch}},S_+^\star}$ is

$$X_{q_{s-1},(u_p,k_p)} = \sqrt{d_{u_p}}\,(M^{k_p})_{q_{s-1},u_p}.$$

If $s < p$, then $s-1 < p-1$, so by top isolation

$$u_p \notin \mathrm{An}_{\mathcal{G}}(q_{s-1}),$$

and therefore

$$(M^{k_p})_{q_{s-1},u_p} = 0.$$

Hence $X_{I_{\mathrm{ch}},S_+^\star}$ is lower triangular.

Similarly, the (s,p)-entry of $X_{J_{\mathrm{ch}},S_+^\star}$ is

$$X_{q_s,(u_p,k_p)} = \sqrt{d_{u_p}}\,(M^{k_p})_{q_s,u_p}.$$

If $s > p$, then by top isolation $u_p \notin \mathrm{An}(q_s)$, so the entry is zero. Hence $X_{J_{\mathrm{ch}},S_+^\star}$ is upper triangular.

Therefore

$$\det X_{\{i\}\cup K,S^\star} = \Big(\prod_{k\in K_0}\sqrt{d_k}\Big)\prod_{p=1}^{r} X_{q_{p-1},(u_p,k_p)},$$

$$\det X_{\{j\}\cup K,S^\star} = \Big(\prod_{k\in K_0}\sqrt{d_k}\Big)\prod_{p=1}^{r} X_{q_p,(u_p,k_p)}.$$

Step 2: restriction to a sparse parameter subspace. To prove nonvanishing, it suffices to restrict M to a sparse subspace: if the full determinant were identically zero on $\Theta_{\mathcal{G}}$, then so would be its restriction to any linear subspace.

We therefore keep only:

- the non-loop edges appearing on some branch L_p or R_p,
- the self-loops $u_p \to u_p$ at the selected tops.

Set all other entries of M equal to 0.

Introduce independent indeterminates:

x_e for each selected non-loop edge e, $\qquad z_p$ for the self-loop $u_p \to u_p$.

For each p, define

$$\lambda_p := z_p^{k_p-|L_p|}\prod_{e\in L_p} x_e, \qquad \rho_p := z_p^{k_p-|R_p|}\prod_{e\in R_p} x_e.$$

These are exactly the monomials of the k_p-padded lifts of L_p and R_p.

Step 3: the $S^\star$-summand contains a distinguished monomial. Consider the diagonal entry

$$X_{q_{p-1},(u_p,k_p)} = \sqrt{d_{u_p}}\,(M^{k_p})_{q_{p-1},u_p}.$$

The quantity $(M^{k_p})_{q_{p-1},u_p}$ is the sum of the monomials of all directed length-k_p walks from u_p to q_{p-1} in the restricted graph. The k_p-padded lift of L_p is one such walk, and it contributes the monomial λ_p. Every walk appears with coefficient $+1$, so the coefficient of λ_p is positive. Likewise,

$$X_{q_p,(u_p,k_p)} = \sqrt{d_{u_p}}\,(M^{k_p})_{q_p,u_p}$$

contains ρ_p with positive coefficient.

Thus

$$\det X_{\{i\}\cup K,S^\star}\,\det X_{\{j\}\cup K,S^\star}$$

contains the monomial

$$\mathfrak{m}^\star := \Big(\prod_{k\in K_0} d_k\Big)\prod_{p=1}^{r} d_{u_p}\lambda_p\rho_p$$

with positive coefficient.

Step 4: no other Cauchy–Binet summand can produce $\mathfrak{m}^\star$. Let $S \neq S^\star$ be any other $(|K|+1)$-element column set. Every monomial in the summand

$$\det X_{\{i\}\cup K,S}\,\det X_{\{j\}\cup K,S}$$

has d_v-exponents given by the multiplicity vector $m(S)$. Therefore, if $m(S) \neq m(S^\star)$, then no monomial in this summand can equal $\mathfrak{m}^\star$.

So suppose $m(S) = m(S^\star)$. Then S uses the same spatial tops as $S^\star$, counted with multiplicity. There are two cases.

Case 1: the lag-zero part differs. Suppose some $k \in K_0$ is represented in S by a positive-lag column (k,κ) with $\kappa > 0$, instead of $(k,0)$. Every monomial from that column then contains at least one M-variable, because every positive-length walk uses at least one edge. But in $\mathfrak{m}^\star$, the contribution of that k-column is only the factor d_k, with no M-variable attached. Hence the summand cannot contain $\mathfrak{m}^\star$.

Thus any possible competitor must have the same lag-zero part:

$$S_0 = S_0^\star.$$

Case 2: the lag-zero part agrees, but some positive lag changes. Then

$$S = \{(u_1,\kappa_1),\ldots,(u_r,\kappa_r)\} \cup S_0^\star$$

for some $\kappa_p \in \mathbb{N}$, and $(\kappa_1,\ldots,\kappa_r) \neq (k_1,\ldots,k_r)$.

Fix p. In order to reproduce $\mathfrak{m}^\star$, the non-loop edge factors coming from the u_p-column on the left and right must be exactly the edge factors of L_p and R_p. Since the x_e are independent indeterminates, this forces the contributing left walk to use the same non-loop edges as L_p and the contributing right walk to use the same non-loop edges as R_p. The remaining steps must then be realized by self-loops at u_p, so the total exponent of z_p becomes

$$(\kappa_p - |L_p|) + (\kappa_p - |R_p|) = 2\kappa_p - |L_p| - |R_p|.$$

In $\mathfrak{m}^\star$, the exponent of z_p is

$$2k_p - |L_p| - |R_p|.$$

Since $\kappa_p \neq k_p$, the two exponents differ. Hence this summand cannot contain $\mathfrak{m}^\star$.

Therefore no summand with $S \neq S^\star$ contributes $\mathfrak{m}^\star$.

Step 5: no cancellation within the $S^\star$-summand. Inside the $S^\star$-summand itself, the chain blocks are triangular, so in the Leibniz expansion only the diagonal permutation contributes. Thus no second term from the same summand can produce $\mathfrak{m}^\star$ with opposite sign.

Hence the full Cauchy–Binet expansion of

$$\det \Sigma_{\{i\}\cup K, \{j\}\cup K}$$

contains $\mathfrak{m}^\star$ with nonzero coefficient. Therefore

$$\det \Sigma_{\{i\}\cup K, \{j\}\cup K} \not\equiv 0.$$

■

D.4. Real-analyticity

Lemma 30 (Real-analyticity of covariance minors) *The map*

$$(M, D) \longmapsto \Sigma(M, D)$$

is real-analytic on $\Theta_\mathcal{G}$. Consequently, every minor

$$(M, D) \longmapsto \det \Sigma_{I,J}(M, D)$$

is real-analytic on $\Theta_\mathcal{G}$.

Proof Vectorizing the Lyapunov equation gives

$$\operatorname{vec}(\Sigma) = (I - M \otimes M)^{-1} \operatorname{vec}(D).$$

Since $\rho(M) < 1$, we have $\rho(M \otimes M) = \rho(M)^2 < 1$, so $I - M \otimes M$ is invertible on $\Theta_\mathcal{G}$. Its inverse is a rational matrix-valued function of the entries of M, analytic wherever the determinant is nonzero; here it is nonzero throughout $\Theta_\mathcal{G}$. Hence $\Sigma(M, D)$ is real-analytic. Determinants of finite submatrices are polynomials in their entries, so all covariance minors are real-analytic as well. ■

Proposition 31 (Zero set of a nontrivial real-analytic function) *Let $U \subseteq \mathbb{R}^N$ be a connected open set and $f : U \to \mathbb{R}$ a real-analytic function that is not identically zero. Then $f^{-1}(0)$ has Lebesgue measure zero in U.*

D.5. Main theorem and corollary

Proof [Proof of Theorem 10] Let $I, J, K \subseteq \mathcal{V}$ be pairwise disjoint.

(i) Soundness. Assume

$$(I \perp_c J \mid K)_{\mathcal{G}}.$$

By Theorem 9,

$$(I \perp_{st} J \mid K)_{\mathcal{G}_T}.$$

We must show that for every $(M, D) \in \Theta_{\mathcal{G}}$,

$$X_I(t) \perp\!\!\!\perp X_J(t) \mid X_K(t).$$

By Lemma 12, it is enough to prove that for every $i \in I$, $j \in J$,

$$\det \Sigma_{\{i\}\cup K, \{j\}\cup K} = 0.$$

Fix $i \in I$, $j \in J$. Set

$$I' = \{i\} \cup K, \qquad J' = \{j\} \cup K.$$

By Lemma 13,

$$\det \Sigma_{I',J'} = \sum_{\substack{S \subseteq \mathcal{V}\times\mathbb{N}_0 \\ |S|=|K|+1}} \det X_{I',S} \det X_{J',S}.$$

We show that every summand is zero.

Suppose, towards contradiction, that for some S,

$$\det X_{I',S} \neq 0 \qquad \text{and} \qquad \det X_{J',S} \neq 0.$$

Write

$$S = S_+ \sqcup S_0, \qquad S_0 = \{(u, 0) : u \in U\}.$$

By Lemma 28, nonvanishing of both minors implies

$$U \subseteq I' \cap J' = K.$$

Applying Lemma 28 again, we obtain

$$\det X_{I',S} = \pm\Big(\prod_{u\in U} \sqrt{d_u}\Big) \det X_{I'\setminus U, S_+},$$

$$\det X_{J',S} = \pm\Big(\prod_{u\in U} \sqrt{d_u}\Big) \det X_{J'\setminus U, S_+}.$$

Hence

$$\det X_{I'\setminus U, S_+} \neq 0, \qquad \det X_{J'\setminus U, S_+} \neq 0.$$

Since $S_+ \subseteq \mathcal{V} \times \mathbb{N}$, Lemma 14 gives:

- a clean positive-lag path system from the sources S_+ to the sinks $I' \setminus U$,
- a clean positive-lag path system from the same sources S_+ to the sinks $J' \setminus U$.

Let

$$\alpha : S_+ \to I' \setminus U, \qquad \beta : S_+ \to J' \setminus U$$

be the corresponding sink bijections. Define

$$\pi := \beta \circ \alpha^{-1} : I' \setminus U \to J' \setminus U.$$

Since $i \in I' \setminus U$ (because $i \notin K$), start with

$$q_0 = i, \qquad q_{m+1} := \pi(q_m)$$

as long as $q_m \neq j$. If $q_{m+1} \neq j$, then $q_{m+1} \in K \setminus U$, hence $q_{m+1} \in I' \setminus U$ and the iteration continues.

We claim that the sequence cannot repeat before hitting j. Suppose $q_r = q_s$ with $0 < r < s$. Since β is injective,

$$\alpha^{-1}(q_{r-1}) = \alpha^{-1}(q_{s-1}),$$

hence $q_{r-1} = q_{s-1}$. Repeating backwards yields

$$q_0 = q_{s-r}.$$

But $q_0 = i \notin K$, whereas $q_{s-r} \in K \setminus U$, contradiction. Since $I' \setminus U$ is finite, the repetition-free sequence must eventually reach j. Thus we obtain

$$q_0 = i,\ q_1, \dots, q_m = j$$

with

$$q_1, \dots, q_{m-1} \in K.$$

For each $r = 1, \dots, m$, the common source $\alpha^{-1}(q_{r-1}) \in S_+$ defines one positive-lag path to q_{r-1} in the first clean system and one positive-lag path to q_r in the second. These two paths form an st-trek between q_{r-1} and q_r. Therefore

$$(q_0, \dots, q_m)$$

is an active multi-trek from i to j given K, contradicting

$$(I \perp_{st} J \mid K)_{\mathcal{G}_T}.$$

Hence every Cauchy–Binet summand vanishes, and so

$$\det \Sigma_{\{i\} \cup K, \{j\} \cup K} = 0.$$

By Lemma 12, this proves

$$X_I(t) \perp\!\!\!\perp X_J(t) \mid X_K(t).$$

(ii) Strong completeness. Assume

$$(I \not\perp_c J \mid K)_{\mathcal{G}}.$$

Then there exist $i \in I$, $j \in J$ such that

$$i \not\perp_c j \mid K.$$

By Theorem 9,

$$i \not\perp_{st} j \mid K.$$

Define

$$f_{i,j,K}(M,D) := \det \Sigma_{\{i\}\cup K,\{j\}\cup K}(M,D).$$

By Lemma 16, $f_{i,j,K}$ is not identically zero on $\Theta_{\mathcal{G}}$. By Lemma 30, it is real-analytic on $\Theta_{\mathcal{G}}$. Therefore, by Proposition 31, its zero set

$$Z_{i,j,K} := \big\{(M,D) \in \Theta_{\mathcal{G}} : f_{i,j,K}(M,D) = 0\big\}$$

has Lebesgue measure zero.

By Lemma 12,

$$\big\{(M,D) \in \Theta_{\mathcal{G}} : X_I(t) \perp\!\!\!\perp X_J(t) \mid X_K(t)\big\} \subseteq Z_{i,j,K},$$

because block conditional independence implies all pairwise Gaussian minor constraints, in particular the one for the chosen $i \in I$, $j \in J$. Therefore

$$\big\{(M,D) \in \Theta_{\mathcal{G}} : X_I(t) \perp\!\!\!\perp X_J(t) \mid X_K(t)\big\}$$

has Lebesgue measure zero. This proves strong completeness. ■

Proof [Proof of Corollary 11] Let

$$\mathcal{D} := \big\{(I,J,K) : I,J,K \subseteq \mathcal{V} \text{ pairwise disjoint}\big\}.$$

Since $\mathcal{V}$ is finite, $\mathcal{D}$ is finite. Let

$$\mathcal{D}_{\text{fail}} := \big\{(I,J,K) \in \mathcal{D} : (I \not\perp_c J \mid K)_{\mathcal{G}}\big\}.$$

For each $(I,J,K) \in \mathcal{D}_{\text{fail}}$, define the exceptional set

$$E_{IJK} := \big\{(M,D) \in \Theta_{\mathcal{G}} : X_I(t) \perp\!\!\!\perp X_J(t) \mid X_K(t)\big\}.$$

By Theorem 10(ii), each E_{IJK} has Lebesgue measure zero. Hence the finite union

$$E := \bigcup_{(I,J,K)\in\mathcal{D}_{\text{fail}}} E_{IJK}$$

also has Lebesgue measure zero.

Because $\Theta_{\mathcal{G}}$ is a nonempty open subset of Euclidean space, it has positive Lebesgue measure, so

$$\Theta_{\mathcal{G}} \setminus E \neq \varnothing.$$

Choose $(M^\star, D^\star) \in \Theta_{\mathcal{G}} \setminus E$, and let X be the corresponding stable Gaussian VAR(1) process. This process is strictly stationary by construction.

We verify the two claims.

Soundness. If $(I, J, K) \in \mathcal{I}_{\mathcal{G}}$, then

$$(I \perp_c J \mid K)_{\mathcal{G}},$$

so by Theorem 10(i),

$$X_I(t) \perp\!\!\!\perp X_J(t) \mid X_K(t).$$

Weak completeness. If $(I, J, K) \notin \mathcal{I}_{\mathcal{G}}$, then

$$(I \not\perp_c J \mid K)_{\mathcal{G}},$$

so $(I, J, K) \in \mathcal{D}_{\text{fail}}$. Because $(M^\star, D^\star) \notin E_{IJK}$, we have

$$X_I(t) \not\perp\!\!\!\perp X_J(t) \mid X_K(t).$$

Thus this single strictly stationary Gaussian VAR(1) process satisfies exactly the c-separation conditional independences of $\mathcal{G}$, which proves the corollary. ∎

Appendix E. Altenrative partial proof of weak completeness

In case where there is only one collider we show that we are able to construct a probability distribution that satisfies the list of independencies defined by c-separations in the summary graph and none other through Armstrong product.

Theorem 32 (Weak completeness) *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a summary graph on a finite node set $\mathcal{V}$, but only one collider can have a non empty set of it's descendants and let*

$$\mathcal{I}_{\mathcal{G}} := \big\{(A, B, S) : A, B, S \subseteq \mathcal{V} \text{ pairwise disjoint and } (A \perp_c B \mid S)_{\mathcal{G}} \big\}.$$

Then there exists a strictly stationary stochastic process $X = \big(X_v(t)\big)_{v\in\mathcal{V},\, t\in\mathbb{Z}}$ on a common probability space such that, for all pairwise disjoint $A, B, S \subseteq \mathcal{V}$ and for all $t \in \mathbb{Z}$,

1. Soundness. *For* every *triple $(A, B, S) \in \mathcal{I}_{\mathcal{G}}$: $X_A(t) \perp\!\!\!\perp X_B(t) \mid X_S(t)$,*
2. Weak Completeness. *Conversely, for any triple $(A, B, S) \notin \mathcal{I}_{\mathcal{G}}$: $X_A(t) \not\perp\!\!\!\perp X_B(t) \mid X_S(t)$.*

This justifies using c-separation as a "complete" logical description of what snapshot CIs the graph can ever generate paralleling the classical role of d-separation in the static setting (Geiger and Pearl, 1990a). We note, however, that our construction does not guarantee preserving δ- nor ϵ-separations of the original graph.

The proof of the above theorem is "local-to-global". We do not try to build one faithful model on the whole summary graph at once. Instead, for each failed c-separation we construct a small local witness supported on a minimal c-open subgraph, chosen so that a Gaussian Schur-complement calculation forces the relevant conditional covariance to be nonzero. We then assemble all these local witnesses by an Armstrong product, whose conditional-independence model is the intersection of the conditional-independence models of its factors. This proof strategy is inspired by how weak necessity of d-separation was established in Geiger and Pearl (1990a), although adapted significantly to stationary time series. The three constructions behind this strategy are introduced below.

The first step is to localize a failed c-separation to a smallest subgraph that still remembers why the path is c-open.

Definition 33 (Minimal c-open path) *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a summary graph and let $\sigma = (\{a, b\}, S)$ with $a, b \in \mathcal{V}$ and $S \subseteq \mathcal{V}$. Among all c-open a-b paths, choose one with the smallest number of vertices (break ties arbitrarily) and denote it by q. List the colliders on q in the order they appear when traversing q from a to b as $c_1, \ldots, c_k$. For each i, since q is c-open given S we have $c_i \in \mathrm{An}_{\mathcal{G}}(S)$, so there exists at least one $s \in S$ with a directed path from c_i to s. Choose $s_i \in S \cap \mathrm{De}_{\mathcal{G}}(c_i)$ minimizing the directed-path length among such descendants, and let p_i be a shortest directed path from c_i to s_i (trivial if $c_i \in S$). The* minimal c-open path *associated with σ is the subgraph $\mathcal{G}_\sigma$ induced by all vertices on q and on the paths p_i. By construction, $\mathcal{G}_\sigma$ still contains a c-open path from a to b given S.*

The second step is to modify such a witness subgraph without changing its collider structure, so that when constructing the process we can control the terms of the covariance and it's inverse when checking the Schur condition for (in)dependence.

Definition 34 (Fraternization family and collider-invariant fraternization) *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a summary graph. A path $p = (v_0, \ldots, v_k)$ in $\mathcal{G}$ is* collider-free *if every internal vertex v_i ($1 \leq i \leq k-1$) is a non-collider on p, i.e., the two edges of p incident to v_i do* not *both have arrowheads at v_i. For distinct $u, v \in \mathcal{V}$, write $u \rightsquigarrow_{\mathrm{nc}} v$ if there exists a collider-free path between u and v in $\mathcal{G}$. Let $\mathcal{M} := \Big\{\{u, v\} \subseteq \mathcal{V} :\ u \neq v,\ u \rightsquigarrow_{\mathrm{nc}} v,$ and u and v are not adjacent in $\mathcal{G}\Big\}$. A* fraternization orientation *is any map $\omega : \mathcal{M} \to \mathcal{V} \times \mathcal{V}$, $\omega(\{u, v\}) \in \{(u, v), (v, u)\}$, interpreted as adding the directed edge $u \to v$ or $v \to u$. For such ω, define $\mathcal{G}_\omega := \Big(\mathcal{V},\ \mathcal{E} \cup \{\omega(\{u, v\}) : \{u, v\} \in \mathcal{M}\}\Big)$, $\mathrm{Frat}(\mathcal{G}) := \{\mathcal{G}_\omega : \omega\}$.*

A collider triple *in a summary graph $\mathcal{H}$ is an ordered triple (a, v, b) of distinct vertices such that $a \to v$ and $b \to v$ are edges of $\mathcal{H}$. Write $\mathrm{Coll}(\mathcal{H})$ for the set of collider triples. A fraternization $\mathcal{F} \in \mathrm{Frat}(\mathcal{G})$ is* collider-invariant *if $\mathrm{Coll}(\mathcal{F}) = \mathrm{Coll}(\mathcal{G})$.*

Collider-invariant fraternization preserves the set of collider triples, and hence c-separation statements on $\mathcal{V}$ are unchanged. To construct local witnesses, we reduce any non-separated triple to a smallest "witness subgraph" that still has a c-open path. In our proofs we only fraternize carefully chosen minimal c-open subgraphs $\mathcal{G}_\sigma$, and on those one can always choose an orientation that preserves all collider/non-collider statuses ($\mathcal{G}_\sigma \equiv_c \mathcal{F}$). That is because global c-invariant fraternization does not always exist for arbitrary summary graphs (Remark 41. Fraternization makes an edge between all nodes that do not have a collider on a path between them. This makes the local covariance structure strongly connected in the "right places", which in turn forces conditional covariances along these segments to be nonzero.

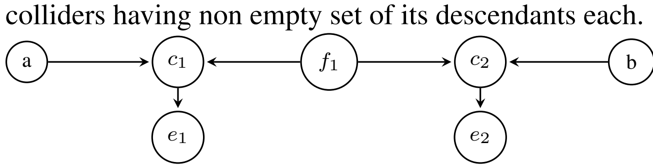
The final step is an assembly operation that combines local witnesses into one global strictly stationary process.

Definition 35 (Armstrong product (for stationary processes).) *Let $(X_t^{(1)})_{t\in\mathbb{Z}}$ and $(X_t^{(2)})_{t\in\mathbb{Z}}$ be strictly stationary Markov chains with state spaces $\mathcal{X}_1$ and $\mathcal{X}_2$, transition kernels K_1, K_2, and* *stationary laws* π_1, π_2. *The* Armstrong product *is the process* $(X_t)_{t\in\mathbb{Z}}$ *on* $\mathcal{X}_1 \times \mathcal{X}_2$ *with the transition kernel* $K\big((x_1, x_2) \to (y_1, y_2)\big) := K_1(x_1 \to y_1)\, K_2(x_2 \to y_2)$, *and* $\pi(x_1, x_2) := \pi_1(x_1)\pi_2(x_2)$.

The Armstrong product π is stationary for K (Lemma 48). Moreover, this construction behaves like a logical "AND" over CI statements: a conditional independence holds in the Armstrong product if and only if it holds in every process that forms it. For any disjoint coordinate sets A, B, C and any conditional independence statement $X_A(t) \perp\!\!\!\perp X_B(t) \mid X_C(t)$, this statement holds for the product process if and only if it holds for each process. We use it to glue together many local witnesses into a single global stationary process that satisfies exactly the intersection of their CI statements.

Proof idea. The construction proceeds in four steps.
1. **Reduce to singleton pairs.** Introduce all admissible disjoint triples of the vertices in $\mathcal{V}$ of $\mathcal{G}$ (unordered pair plus conditioning set) of this form $\mathcal{D} = \big\{\big(\{A, B\}, S\big) : A, B, S \subset \mathcal{V}\big\}$. From the summary graph $\mathcal{G}$ construct a list of all true graphical c-separation statements which imply the respective probabilistic independence and it's complement L_D. To witness failure of $X_A(t) \perp\!\!\!\perp X_B(t) \mid X_S(t)$, it suffices to find $a \in A$ and $b \in B$ with $X_a(t) \not\perp\!\!\!\perp X_b(t) \mid X_S(t)$. Thus every failed block statement can be reduced to a failed singleton-pair statement.
2. **Local witnesses on fraternized minimal c-open paths.** For each triple $\sigma = (\{a, b\}, S)$ in the list L_D we construct a minimal c-open path by taking a subgraph of $\mathcal{G}$. We do so by taking the path between a and b given S, together with paths from its colliders to the closest descendants in S. This yields a subgraph $\mathcal{G}_\sigma$ that still contains a c-open a-b path (Proposition 43). On $\mathcal{G}_\sigma$ we perform a collider-invariant fraternization that makes an edge between all vertices in non-collider segments without changing the set of collider triples (Lemma 44), obtaining $F_\sigma \in \mathrm{Frat}(\mathcal{G}_\sigma)$. In particular, all c-separations on the vertices of $\mathcal{G}_\sigma$ are preserved and the a-b path remains c-open given S.
3. **Gaussian VAR(1) with Z-covariance and nonzero Schur complement.** On each fraternized subgraph F_σ we build a local Gaussian VAR(1) model $X^{(\sigma)}$ whose stationary covariance matrix $\Sigma^{(\sigma)}$ is a Z-matrix (positive diagonal, non-positive off-diagonals), (Lemma 45). For such matrices, the inverse $(\Sigma^{(\sigma)})^{-1}$ is entrywise nonnegative and strictly positive on irreducible blocks (Lemma 46). In turn, using the Gaussian Schur complement formula, $\Sigma^{(\sigma)}_{ab|S} := \Sigma^{(\sigma)}_{ab} - \Sigma^{(\sigma)}_{aS}\big(\Sigma^{(\sigma)}_{SS}\big)^{-1}\Sigma^{(\sigma)}_{Sb}$, one shows that $\Sigma^{(\sigma)}_{ab|S} \neq 0$ (Lemma 47), hence $X^{(\sigma)}_a(t) \not\perp\!\!\!\perp X^{(\sigma)}_b(t) \mid X^{(\sigma)}_S(t)$ in the local model, while c-separations inside F_σ are respected. Every failure of c-separation is witnessed by a stationary process with dependence at the snapshot.
4. **Armstrong product for stationary processes.** Finally, we combine all local witnesses by an Armstrong product. We prove that the Armstrong relation works also for stationary stochastic processes (Lemma 48). Since a conditional independence holds in the product if and only if it holds in every factor, every c-separation statement is preserved, while every failed c-separation is violated in at least one factor and hence also in the product. This yields a strictly stationary process whose snapshot CI structure is exactly the one implied and necessitated by c-separation, establishing soundness and weak completeness.

Unfortunately this proof does not work in the current form for summary graphs with multiple colliders having non empty set of its descendants each. Observe this example:



Fraternization demands making an arrow between c_1 and c_2. Such arrow changes the c separations.

E.1. Supporting results and proofs

We recall the fraternization operation and spell out a collider-invariant orientation on minimal c-open subgraphs.

Lemma 36 (Singleton reduction) *Let $A, B, S \subseteq \mathcal{V}$ be pairwise disjoint. If there exist $a \in A$ and $b \in B$ with $X_a(t) \not\perp\!\!\!\perp X_b(t) \mid X_S(t)$, then $X_A(t) \not\perp\!\!\!\perp X_B(t) \mid X_S(t)$.*

Definition 37 (Fraternization) *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a summary graph. For distinct $u, v \in \mathcal{V}$, write $u \rightsquigarrow_{\mathrm{nc}} v$ if there exists a non-collider (collider-free) path between u and v in $\mathcal{G}$ (i.e., a path whose internal vertices are all non-colliders). Let*

$$\mathcal{M} := \big\{\{u, v\} \subseteq \mathcal{V} : \ u \neq v, \ u \rightsquigarrow_{\mathrm{nc}} v, \ u \text{ and } v \text{ are not adjacent in } \mathcal{G}\big\}.$$

A fraternization orientation *is any choice of directed edges*

$$\omega : \mathcal{M} \to \mathcal{V} \times \mathcal{V}, \qquad \omega(\{u, v\}) \in \{(u, v), (v, u)\},$$

and $\mathcal{G}_\omega$, $\mathrm{Frat}(\mathcal{G})$ are defined as in Definition 34.

Remark 38 *The relation $\rightsquigarrow_{\mathrm{nc}}$ is symmetric but not transitive in general: in $u \to v \leftarrow w$ we have $u \rightsquigarrow_{\mathrm{nc}} v$ and $v \rightsquigarrow_{\mathrm{nc}} w$, but $u \not\rightsquigarrow_{\mathrm{nc}} w$. Hence we do not form equivalence classes from it. Accordingly, we work with the set of fraternal pairs $\mathcal{M}(\mathcal{G})$ rather than equivalence classes.*

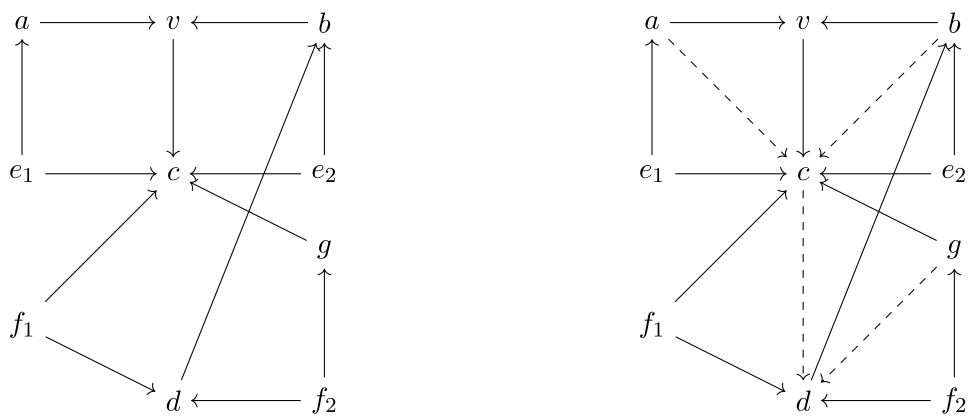
Definition 39 (Collider triple) *Let $\mathcal{G} = (V, E)$ be a summary graph. A* collider triple *is an ordered triple (a, v, b) of pairwise distinct vertices such that $a \to v \in E$ and $b \to v \in E$. Write $\mathrm{Coll}(\mathcal{G})$ for the set of all collider triples of $\mathcal{G}$.*

Definition 40 (Collider-invariant fraternization) *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and let $\mathrm{Frat}(\mathcal{G})$ be the fraternization family. A graph $\mathcal{F} \in \mathrm{Frat}(\mathcal{G})$ is a* collider-invariant fraternization *of $\mathcal{G}$ if*

$$\mathrm{Coll}(\mathcal{F}) = \mathrm{Coll}(\mathcal{G}).$$

Such a c-invariant orientation, if possible, can be constructed as follows. For each fraternity C_ℓ and for every unordered pair $\{v^i, v^j\} \subseteq C_\ell$ that were nonadjacent in $\mathcal{G}$, choose one of the two directed edges $v^i \to v^j$ or $v^j \to v^i$ with the constraint that its addition does not alter the set of collider triples. (At least one of the two orientations is safe because v^i and v^j are connected by a non-collider path in $\mathcal{G}$.) If both orientations are safe, tie-break in a consistent way, for example, prefer orienting the edge into a vertex that is a descendant of a collider in $\mathcal{G}$, so as not to inadvertently create a new head in a collider configuration. Denote by $\mathcal{E}^{\mathrm{new}}$ the set of all such newly oriented edges; then $\mathcal{E}' = \mathcal{E} \cup \mathcal{E}^{\mathrm{new}}$, and $\mathrm{Frat}(\mathcal{G}) = (\mathcal{V}, \mathcal{E}')$ is a collider-invariant fraternization of $\mathcal{G}$.

Remark 41 (C-invariant fraternization does not always exist) *Not every summary graph admits a collider-invariant fraternization that orients all fraternal pairs in $\mathcal{M}(\mathcal{G})$: for some fraternal pair $\{u, v\}$, both orientations $u \to v$ and $v \to u$ may create a new collider triple (see Figure 3 for an illustration). For our purposes, fraternization is applied only to the carefully chosen minimal witness subgraphs $\mathcal{G}_\sigma$; Lemma 44 shows that in that restricted setting, a collider-invariant fraternization exists.*



(a) Original graph

(b) Completing fraternal pairs may break collider-invariance

Figure 3: Illustration of counterexample.

Definition 42 (Minimal c-open path) *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a summary graph, and let $\sigma = (\{a, b\}, S)$ be a triple, where $a, b \in \mathcal{V}$, $S \subseteq \mathcal{V}$. Let q be a minimal c-open path between a and b given S, and write its colliders (in order from a to b) as $c_1, \dots, c_k$. For each i, let $s_i \in S$ be the closest descendant of c_i in S, and let p_i be a path from c_i to s_i (trivial if $c_i \in S$). The minimal c-open path is a subgraph induced by the vertices and edges on q together with all vertices and edges on the p_i's. We use $\mathcal{G}_\sigma$ to denote the graph induced by σ.*

A similar construction of a subgraph and its properties can be found in Geiger and Pearl (1990a).

Proposition 43 (Minimal c-open path subgraph) *Let $\sigma = (\{a, b\}, C)$ be such that $(a \not\perp_c b \mid C)_\mathcal{G}$. here exists a minimal c-open path q from a to b given C, with colliders $c_1, \dots, c_k$, and for each i a path $p_i : c_i \rightsquigarrow s_i$ to the closest descendant $s_i \in C$. The subgraph H induced by q and $\{p_i\}$ is still c- open between a and b given C.*

We now apply a collider-invariant fraternization to $\mathcal{G}_\sigma$. By Lemma 44, the witness path q remains c-open given S.

Lemma 44 (Collider-invariant fraternization along a minimal witness) *Let $H = \mathcal{G}_\sigma$ be a minimal c-open witness subgraph for $\sigma = (\{a, b\}, S)$ with witness path q. Then there exists $F_\sigma \in \mathrm{Frat}(H)$ such that:*

- *every pair $\{u, v\} \in \mathcal{M}(H)$ becomes adjacent in F_σ, and*
- $\mathrm{Coll}(F_\sigma) = \mathrm{Coll}(H)$.

In particular, the witness path q remains c-open between a and b given S in F_σ.

Let $\mathcal{G}_\sigma$ be the minimal c-open witness subgraph for σ, and fix a collider-invariant fraternization $F_\sigma \in \mathrm{Frat}(\mathcal{G}_\sigma)$ as in Lemma 44. On F_σ we construct a local stable Gaussian VAR(1) witness process $X^{(\sigma)}$ (equivalently, its stationary law P_σ under the Lyapunov model) whose VAR coefficient matrix respects the directed edges of F_σ. This local model is chosen so that: (i) it satisfies all snapshot conditional independences implied by c-separation on F_σ (hence it does not violate any statement from L_I restricted to V_σ), and (ii) it witnesses the desired dependence for σ, i.e. $X_a^{(\sigma)}(t) \not\perp\!\!\!\perp X_b^{(\sigma)}(t) \mid X_S^{(\sigma)}(t)$, by ensuring a nonzero conditional covariance (via the Z-covariance / Schur-complement argument).

We now apply collider-invariant fraternization to this subgraph; by Lemma 44, q remains c-open given C.

Lemma 45 (Local VAR(1) with Z-covariance) *On F_σ there exists a Gaussian VAR(1) model $X_{t+1} = MX_t + \varepsilon_t$ with $\rho(M) < 1$ and $\varepsilon_t \sim \mathcal{N}(0, I)$ such that its stationary covariance $\Sigma = \sum_{k\geq 0} M^k (M^\top)^k$ is positive definite and a Z-matrix (positive diagonal, non-positive off-diagonals).*

From $\Sigma \succ 0$ and the Z-structure, we obtain an entrywise nonnegative inverse, and strict positivity within irreducible (connected) blocks.

Lemma 46 (Inverse of SPD Z-matrices) *If $\Sigma \succ 0$ is a symmetric Z-matrix, then $\Sigma^{-1} \geq 0$ entrywise. Furthermore, for any index set C, each irreducible block of Σ_{CC} has a strictly positive inverse; entries across different blocks are zero.*

Apply Lemma 46 to Σ_{CC}. It shows that $\Sigma_{CC}^{-1} \geq 0$ entrywise and strictly positive on irreducible blocks. We now certify $\Sigma_{ab|C} \neq 0$. We now certify a nonzero Schur complement for (a, b) given C.

Lemma 47 (Nonzero conditional covariance) *For the local model of Lemma 45 and index sets $\{a\}, \{b\}, C$, the Schur complement identity is used for an (in)dependence check on Gaussian variables*

$$\Sigma_{ab|C} := \Sigma_{ab} - \Sigma_{aC}\, \Sigma_{CC}^{-1}\, \Sigma_{Cb} \neq 0.$$

Consequently $X_a(t) \not\perp\!\!\!\perp X_b(t) \mid X_C(t)$.

To avoid juggling a single joint distribution for all possible triplets (a, b, C), we invoke the Armstrong relation Geiger and Pearl (1990b, 1993). Concretely, for each triple (a,b,C) we build a "local" Gaussian VAR(1) model whose stationary law satisfies exactly the list of independencies, whenever a and b are c-separated by C, and violates it otherwise. The Armstrong relation then prescribes a product-measure construction that stitches these local models into a single global process. By a fundamental property of the Armstrong product, a conditional-independence holds in the product exactly when it holds in *every* factor. Thus we automatically obtain one overarching stationary distribution that realizes *all* and only the c-separation independencies entailed by the summary graph—without ever writing down one joint density in closed form.

We can create separate probability distributions for all triplets (a, b, C) and through the Armstrong relation Geiger and Pearl (1990b) we can then join them so that the Armstrong relation between them satisfies all the conditional independence statements that were true for all the probability distributions for all the triplets (intersection). In this sense we do not need to care for the joint distribution at all. The existence of a distribution that satisfies all independencies of the c-separation kind from the summary graph and all its dependencies (not merely a single dependency) is guaranteed Geiger and Pearl (1990a).

We give a more careful treatment and probabilistic interpretation of the Armstrong relation in Appendix E.2.

Lemma 48 (Armstrong Product for Stationary Markov Chains) *Let $\mathcal{X}$ be a finite dimensional state space. For $i = 1, 2$, let K_i be a transition kernel on $\mathcal{X}$ and let π_i be a strictly stationary distribution for K_i, i.e.,*

$$\pi_i K_i = \pi_i,$$

or equivalently $\sum_{x \in \mathcal{X}} \pi_i(x) K_i(x \to y) = \pi_i(y)$ for all $y \in \mathcal{X}$. Consider the product chain on $\mathcal{X} \times \mathcal{X}$ with kernel K and measure π defined by

$$K\big((x_1, x_2) \to (y_1, y_2)\big) := K_1(x_1 \to y_1)\, K_2(x_2 \to y_2), \qquad \pi(x_1, x_2) := \pi_1(x_1)\, \pi_2(x_2),$$

named Armstrong relation. Then:

1. *π is stationary for K; i.e., $\pi K = \pi$. In particular, if $X_0 \sim \pi$ then the resulting chain $(X_t)_{t \in \mathbb{Z}}$ with transition kernel K is strictly stationary with time-t marginals π for all t.*

2. *Suppose in addition that $\mathcal{X}$ carries a product structure $\mathcal{X} = \prod_{v \in V} \mathcal{X}_v$ for some finite index set V, and that each factor chain $(X_t^{(i)})_{t \in \mathbb{Z}}$ takes values in $\mathcal{X}$ with the same coordinate structure:$X_t^{(i)} = \big(X_v^{(i)}(t)\big)_{v \in V}, \qquad i = 1, 2$. Let $(X_t)_{t \in \mathbb{Z}}$ be the Armstrong product chain, so that $X_t = \big(X_t^{(1)}, X_t^{(2)}\big),\ X_t^{(i)} = \big(X_v^{(i)}(t)\big)_{v \in V}$. For any pairwise disjoint subsets $A, B, C \subseteq V$, write $X_A(t) := (X_v(t))_{v \in A},\ X_A^{(i)}(t) := \big(X_v^{(i)}(t)\big)_{v \in A}$, and let $I(U, W, V)$ denote the conditional independence statement $X_U \perp\!\!\!\perp X_V \mid X_W$. Then, for every $t \in \mathbb{Z}$, $I\big(X_A(t), X_C(t), X_B(t)\big)$ holds under the law of the Armstrong product chain if and only if, for each $i = 1, 2$, $I\big(X_A^{(i)}(t), X_C^{(i)}(t), X_B^{(i)}(t)\big)$ holds under the law of the ith factor chain.*

   Equivalently, the time-t marginal π of the Armstrong product satisfies

   $$X_A(t) \perp\!\!\!\perp X_B(t) \mid X_C(t) \quad \iff \quad X_A^{(i)}(t) \perp\!\!\!\perp X_B^{(i)}(t) \mid X_C^{(i)}(t) \ \textit{for } i = 1, 2.$$

Remark 49 (Type of the assembled process) *The Armstrong product is used as a logical device to assemble local models. The global process need not be Gaussian; the theorem requires existence, not Gaussianity. Lemma 48 states the product construction for stationary Markov chains; the same "CI holds in the product if and only if in each factor" argument applies to independent products of stationary processes at a fixed time t (by factorization of the time-t marginal).*

Corollary 50 (Global assembly) *Let $\{P_\sigma : \sigma \in L_D\}$ be the local stationary models constructed on G_σ. Their Armstrong product $P = \bigotimes_{\sigma \in L_D} P_\sigma$ is strictly stationary and satisfies exactly the c-separation independencies of $\mathcal{G}$; for each $\sigma \in L_D$ it preserves the dependence $X_a \not\perp\!\!\!\perp X_b \mid X_C$.*

Assembling all the models P_σ for all F_σ over all the triple vertices that form G_σ from L_D, we can construct the Armstrong relation over them. To this end, we construct $P = \otimes_\sigma P_\sigma$. P has a set of dependencies being a union over dependencies from each model and a set of independencies being an intersection of independencies from each sub model. By Lemma 48, a CI holds in P if and only if it holds in each P_σ, hence independencies are exactly those implied by c-separation. For each $\sigma \in L_D$, Lemma 47 gives a dependence that persists in P, establishing weak necessity. This ends the claim about weak necessity and soundness. In other words, c-separation is Markov-perfect for cross-sectional CI in strictly stationary stochastic processes.

Proof [Proof of Theorem 32] Introduce the summary graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Fix a triple (A, B, C) with $A, B, C \subseteq \mathcal{V}$ pairwise disjoint. Introduce the catalogue of unordered pairs with a conditioning set

$$\mathcal{D} = \big\{ \, (\{A', B'\}, S) \ : \ A', B', S \subseteq \mathcal{V} \text{ pairwise disjoint} \, \big\}.$$

Form

$$L_I = \big\{(\{A', B'\}, S) \in \mathcal{D} : \ (A' \perp_c B' \mid S)_{\mathcal{G}}\big\}, \qquad L_D = \mathcal{D} \setminus L_I.$$

By Lemma 36 (singleton reduction), to prove weak necessity it suffices to exhibit $a \in A$ and $b \in B$ such that $(\{a, b\}, C) \in L_D$ and

$$X_a(t) \not\perp\!\!\!\perp X_b(t) \;\big|\; X_C(t).$$

Step 1: isolate a minimal c-open witness and keep c-separations invariant. Choose $\sigma = (\{a, b\}, S)$. By Proposition 43 there exists a minimal c-open path q from a to b given C, with colliders $c_1, \ldots, c_k$ on q and, for each i, a path $p_i : c_i \rightsquigarrow s_i$ to the closest descendant $s_i \in C$. Let G_σ be the subgraph induced by q and the family $\{p_i\}$. By construction H remains c-open between a and b given C. Apply Lemma 44 to obtain a *collider-invariant fraternization* $F_\sigma \in \mathrm{Frat}(G_\sigma)$: it makes an edge between all vertices without altering the set of collider triples. In particular, every c-separation on nodes of (G_σ) (hence on nodes of (F_σ)) is preserved, and q stays c-open given C.

Step 2: build a local stationary Gaussian model on F_σ with a Z-covariance. On F_σ construct a VAR(1) model

$$X_{t+1} = M X_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, I) \text{ i.i.d.}$$

Choose M with positive diagonal and non-positive off-diagonals aligned with the (oriented) fraternization edges; then replace M by αM with $\alpha \in (0, 1)$ small so that $\rho(M) < 1$ (this scaling preserves the sign pattern). By Lemma 45 the stationary covariance

$$\Sigma = \sum_{k=0}^{\infty} M^k (M^\top)^k$$

exists, is *symmetric positive definite*, and is a *Z-matrix* (diagonal > 0, all off-diagonals ≤ 0). Moreover, if two vertices are collider-away in F_σ, then $\Sigma_{ij} = 0$; this ensures that fraternization has not introduced spurious covariances across c-separated parts of F_σ.

Step 3: certify a nonzero (indeed negative) conditional covariance along the witness. By Lemma 46, $\Sigma_{CC}^{-1} \geq 0$ entrywise and is strictly positive on irreducible blocks (determined by the sign-graph of Σ_{CC}). Consider the Gaussian Schur complement

$$\Sigma_{ab|C} = \Sigma_{ab} \; - \; \Sigma_{aC} \, \Sigma_{CC}^{-1} \, \Sigma_{Cb}.$$

Two cases arise: (i) If a and b are connected by a collider-free segment in F_σ, then $\Sigma_{ab} < 0$ by the Z-pattern and $\Sigma_{aC} \Sigma_{CC}^{-1} \Sigma_{Cb} \geq 0$, whence $\Sigma_{ab|C} < 0$. (ii) If every a-b path in F_σ has a collider, the C-openness of q provides a collider c_i with a descendant in C; then $\Sigma_{ac_i} < 0$, $(\Sigma_{CC}^{-1})_{c_i c_i} > 0$, and $\Sigma_{c_i b} < 0$, giving $\Sigma_{aC} \Sigma_{CC}^{-1} \Sigma_{Cb} > 0$ while $\Sigma_{ab} = 0$, so again $\Sigma_{ab|C} < 0$. In either case Lemma 47 yields

$$X_a(t) \not\perp\!\!\!\perp X_b(t) \;\big|\; X_C(t)$$

for the local model on F_σ.

Step 4: assemble the global strictly stationary process via an Armstrong product. Repeat the local construction for every $\sigma' = (\{a', b'\}, C') \in L_D$, producing a family $\{P_{\sigma'}\}$. For triples in L_D we use the dependent construction above (on the corresponding $F_{\sigma'}$). For triples in L_I we choose a local stationary model that satisfies the required conditional independence on its vertex set (e.g., by taking block-diagonal M that separates A' and B' given C'). By Lemma 48 and Corollary 50, the Armstrong product $P = \bigotimes_{\sigma' \in \mathcal{D}} P_{\sigma'}$ is strictly stationary and has the key property that a time-t conditional independence holds in P if and only if it holds in every factor. Hence, if $(A, B, C) \in L_I$, then for any factor $P_{\sigma'}$ the triple (A, B, C) is either outside of the set of vertices σ of $P_{\sigma'}$ or c-separated within $F_{\sigma'}$, so the CI holds in each $P_{\sigma'}$ and therefore in P (soundness). If $(A, B, C) \in L_D$, pick $\sigma' = (\{a, b\}, C)$ from Step 3. The corresponding factor $P_{\sigma'}$ violates $X_a \perp\!\!\!\perp X_b \mid X_C$, so the product P violates it as well (weak necessity).

Since the argument is time-t marginal and each factor is strictly stationary, the same statements hold for all $t \in \mathbb{Z}$. This establishes the equivalence in the theorem and completes the proof sketch. ■

E.2. Armstrong relation

Definition 51 (Armstrong relation from Geiger and Pearl (1993)) *A probability model over a finite set of attributes $U = \{u_1, \dots, u_n\}$ is a pair (d, P), where d is a domain mapping that maps each u_i to a finite domain $d(u_i)$, and*

$$P \colon d(u_1) \times \cdots \times d(u_n) \to [0, 1]$$

is a probability distribution having the Cartesian product of these domains as its sample space.

Definition 52 *The* (binary) direct product *also called* Armstrong relation *for $\mathcal{F}$ is a mapping, $\otimes : \mathcal{F} \times \mathcal{F} \to \mathcal{F}$, where $\mathcal{F}$ is a class of probability models over a finite set of attributes $\{u_1, \dots, u_n\}$, and $(d, P) = (d_1, P_1) \otimes (d_2, P_2)$ is defined as follows:*

Let $d_1(u_i)$ and $d_2(u_i)$ be the domains associated with u_i in (d_1, P_1) and in (d_2, P_2), respectively. Let a_i and b_i be values drawn respectively from these domains. Set the domain of u_i in (d, P) to be the Cartesian product $d_1(u_i) \times d_2(u_i)$, and let

$$P(a_1 b_1, a_2 b_2, \dots, a_n b_n) = P_1(a_1, a_2, \dots, a_n) \cdot P_2(b_1, b_2, \dots, b_n), \tag{5}$$

where $a_i b_i$ denotes a value of u_i in (d, P).

A notable property of $\otimes$ is the assignment of a new domain, $d_1(u_i) \times d_2(u_i)$, to each u_i. Thus u_i is treated as an attribute rather than a variable with a fixed domain.

Lemma 53 *Let P_1 and P_2, and P be probability models over U. Then, for any three disjoint subsets X, Y, and Z of U, and conditional independence statements $I(\cdot, \cdot, \cdot)$*

$$I(X, Z, Y) \text{ holds for } P = P_1 \otimes P_2 \quad \text{if and only if} \quad I(X, Z, Y) \text{ holds for } P_1 \text{ and for } P_2.$$

For every set of independence statements L over the attributes of U, there exists a probability model P in $\mathcal{P}$ such that P satisfies every statement in L and none other.

E.3. Deferred proofs of propositions and lemmas

Proof [Sketch of the proof of Lemma 36]
This stems from semi-graphoid axioms Geiger and Pearl (1990a). The properties of d-separation and conditional independence apply also to c separation and the independence it implies, in particular decomposition:

$$I(X, Z, Y \cup W) \implies I(X, Z, Y) \wedge I(X, Z, W).$$

Negating the above we obtain the property that is useful for us:

$$\neg I(X, Z, Y) \vee \neg I(X, Z, W) \implies \neg I(X, Z, Y \cup W).$$

■

Proof [Proof of Lemma 44]
Make each set of vertices that do not have a collider on a path between one another restricted to $q \setminus \{h_i\}$ and each $p_i \setminus \{h_i\}$ into having an edge between all vertices. Orient those added edges so that the set of c-separated triples are preserved.

- From the minimal path consisting k colliders, enumerate all fraternities $F_1, \dots F_{k+1}$.
- The intersection of each two fraternities F_i, F_{i+1} is the path q_i.
- Orient all the edges between vertices in the fraternity that do not fall in the intersection with other fraternities as bidirected.
- Orient the edges from F_i and F_{i+1} with arrows to q_i.
- All the edges in q_i except of h_i and its direct descendant can be also oriented with bidirected edges.
- The descendant and the collider have to have all incoming edges from all other vertices.

Each added edge is oriented along a non-collider class; at least one of the two possible orientations preserves collider status at every triple, hence c-separations are unaffected. The minimal c-open path was arbitrary. By this construction we have shown that c-invariant fraternization exists. ■

Proof [Proof of Lemma 45]
We want to start from M and show that in special cases we can make the $\Sigma = \sum_k M^k (M^k)^T$ such that it has all nonnegative entries of its inverse. We can show that it is possible to pick M for any one path in the summary graph such that the covariance is the Z-matrix - i.e., they have non-positive off-diagonal and positive diagonal elements. Knowing that the solution to the Lyapunov equation is positive definite, we can show that the inverse will have non-negative entries.

We are interested in particular not in Σ^{-1}, but in Σ^{-1}_{CC}. Any principal submatrix of a symmetric positive definite matrix is also symmetric positive definite. Therefore, any principal submatrix Σ_{CC} of the original Σ matrix will also retain the property for the elements of the inverse.

We will construct the Gaussian VAR(1) in Lyapunov model:

$$X_{t+1} = MX(t) + \epsilon, \quad \epsilon \sim N(0, I).$$

We begin by setting the values of the $n \times n$ matrix M (possibly not dense, and not symmetric) such that it aligns with the fraternized path subgraph G_σ. We set

1. every off-diagonal element in a given row is non-positive:

$$M_{ij} \leq 0, \quad j \neq i.$$

2. the diagonal entry is positive:

$$M_{ii} > 0.$$

An analogous statement holds for $M^\top$ then $(M^\top)_{ij} = M_{ji} \leq 0, \quad i \neq j$, because row j (which corresponds to column j of $M^\top$) satisfies the same property.

For simplicity let us set two parameters

- Starting diagonal $\forall i$: $M_{ii} = d > 0$
- Starting off-diagonal: $M_{ij} = -\epsilon_{ij} \leq 0$, where

$$\epsilon_{ij} = \begin{cases} \epsilon & \text{if } i \leftarrow j \text{ in } G_\sigma \\ 0 & \text{otherwise} \end{cases},$$

  where G_σ is the graph that has all the vertices - that are no further than collider away in the original G - connected by bidirected edges.
- Number of starting non-zero elements in row i in the matrix: $0 \leq m_i = |\{j \neq i : \epsilon_{ij} \neq 0\}|$

Let $S_k = M^k(M^k)^\top$ denote the k-th element of the sum that constitutes the covariance $\Sigma = \sum_k S_k$.

Now, for $k = 1$,

$$S_1 = MM^\top.$$

The diagonal elements are

$$(S_1)_{ii} = \sum_{l=1} M_{il}^2 = d^2 + \sum_l \epsilon_{il}\epsilon_{il} = d^2 + n_{i,1}\epsilon^2 > 0$$

This is always satisfied, no matter what the structure of M is.

The off-diagonal elements for i, j, $i \neq j$ are

$$(S_1)_{ij} = (MM^\top)_{ij} = \sum_{l=1}^{n} M_{il}M_{jl} = -d\epsilon_{ji} - d\epsilon_{ij} + \sum_{l \notin \{i,j\}} \epsilon_{il}\epsilon_{jl}.$$

$$(S_1)_{ij} \leq 0 \iff d(\epsilon_{ji} + \epsilon_{ij}) \geq \sum_{l \notin \{i,j\}} \epsilon_{il}\epsilon_{jl}$$

The terms on the left hand side are non zero if the i and j nodes are less than a collider away. However, if these are further than a collider away then they will also not have a common ancestor, which in fact is a sufficient condition for any term in the right hand side sum to be zero. In the case of i and j being colliders away both terms are equal to zero and the off-diagonal element is also equal to zero.

Choose

$$\frac{d}{\epsilon} \geq \max_i m_i.$$

Then for every pair $i \neq j$ that is not a collider away we satisfy the condition:

$$d(\epsilon_{ji} + \epsilon_{ij}) \geq d\epsilon \geq \max_i m_i \epsilon^2 \geq \sum_{l \notin \{i,j\}} \epsilon_{il}\epsilon_{jl}.$$

Moreover, since all variables have at least one child or one parent, for a pair (i, j) there is at least one non zero entry in row i or row j in M. Therefore, there will be at least one non zero entry in S_1 in each row in each off-diagonal guaranteed by the operation $MM^\top$.

Therefore we ensure that

- $(S_1)_{ii} > 0$
- $(S_1)_{ij} \leq 0$
- S_1 is symmetric and for each i there exists $j : (S_1)_{ij} < 0$

What about the stationary covariance matrix $\Sigma = \sum_{k=0}^{\infty} S_k$? Since $\rho(M) < 1$, the discrete Lyapunov series converges and $\Sigma \succ 0$, the diagonal of the matrix will be positive. Moreover, the diagonal terms of all S_k can be guaranteed to be always positive. Surely they are non-negative as those are the variance terms of non-degenerate random variables. The positivity can be proven explicitly algebraically by induction:

Observe that $S_1 = MM^\top$ is positive definite, since the product is semi-definite and M is invertible. Assume S_k is symmetric positive definite and let us show S_{k+1} is also symmetric and positive definite. Naturally the matrix S_{k+1} is symmetric

$$S_{k+1}^\top = \left(MS_kM^\top\right)^\top = (M^\top)^\top S_k^\top M^\top = MS_kM^\top = S_{k+1},$$

since S_k is symmetric and $S_k^\top = S_k$. Hence S_{k+1} is symmetric.

The positive-definiteness can be seen in the following. For any non-zero $x \in \mathbb{R}^n$ we must show $x^\top S_{k+1} x > 0$.

$$x^\top S_{k+1} x = x^\top \left(MS_kM^\top\right)x = (x^\top M)S_k(M^\top x).$$

Set $y := M^\top x$. As M is invertible, so is $M^\top$; thus $x \neq 0 \implies y \neq 0$. Therefore

$$x^\top S_{k+1} x = y^\top S_k y.$$

By the inductive hypothesis S_k is positive definite, so $y^\top S_k y > 0$ for every non-zero y.

Hence $x^\top S_{k+1} x > 0$.

The sum of positive definite matrices is also positive definite. Thus Σ also retains this property.

The off diagonals are

$$\Sigma_{ij} = \sum_{k=0}^{\infty} (S_k)_{ij} \quad \text{where } i \neq j.$$

- Case A (collider away): For the indices that were collider away, $A = \{(i, j) : (S_1)_{ij} = 0\}$, each term $\{(i, j) \in A : (S_k)_{i,j} = 0\}$, because not having common ancestors produce zero entry in the covariance. This follows because the support of $M^k(M^\top)^k$ is contained in pairs that share a length-k bidirected walk in F_σ; collider-away vertices share none in F_σ by construction.

- Case B (less than collider away): $(S_1)_{ij} = -s_{ij} = -\left(d\epsilon_{ji} + d\epsilon_{ij} - \sum_{l\notin\{i,j\}} \epsilon_{il}\epsilon_{jl}\right) < 0$
  For $k \geq 2$

  $$|(S_k)_{ij}| \leq \|S_k\|_2 = \|M^k M^{\top k}\|_2 \leq ||M^k||_2^2 \leq ||M||_2^{2k} =: \gamma^{2k}$$

  Therefore, the tail is bounded by a geometric series

  $$\left|\sum_{k=2}^{\infty}(S_k)_{ij}\right| \leq \sum_{k=2}^{\infty}\gamma^{2k} = \frac{\gamma^4}{1-\gamma^2}$$

  Putting it together

  $$\Sigma_{ij} = (S_1)_{ij} + \sum_{k=2}^{\infty}(S_k)_{ij} \leq -s_{ij} + \frac{\gamma^4}{1-\gamma^2} \leq 0.$$

  To satisfy that every off diagonal has to be non-positive $\Sigma_{ij} \leq 0$ we need

  $$\frac{\gamma^4}{1-\gamma^2} < \min_{i,j}\{s_{ij}\}$$

  But we can control the γ term by scaling down both d and ϵ. Introduce scaling factor $\alpha \in (0,1)$: such that we replace M with αM. This would lead to $\|\alpha M\|_2 = \alpha\gamma$, and since every term in s_{ij} is a product of two entries, the same scaling would occur $s_{ij} \mapsto \alpha^2 s_{ij}$. Then our inequality behaves

  $$\frac{\alpha^4\gamma^4}{1-\alpha^2\gamma^2} \leq \alpha^2 \min_{i,j}\{s_{ij}\}$$
  $$\frac{\alpha^2\gamma^4}{1-\alpha^2\gamma^2} \leq \min_{i,j}\{s_{ij}\}$$

  We can always choose $\alpha \in (0,1)$ such that the left hand side becomes arbitrarily small

  $$\lim_{\alpha\to 0}\frac{\alpha^2\gamma^4}{1-\alpha^2\gamma^2} = 0$$

Therefore $\exists(d,\epsilon)$ such that

1. The stationary distribution exists

   $$\rho(M) < 1$$

2. The covariance matrix is a Z-matrix

   $$\Sigma_{ij} \leq 0 \quad \text{for } i \neq j$$

■

Proof [Proof of Lemma 46]
Inverse of positive definite Z matrix has non negative elements. Write

$$\Sigma = D - N,$$

where D is the diagonal part,

$$D = \operatorname{diag}(\Sigma_{11}, \ldots, \Sigma_{nn}) \quad \text{(all positive)},$$

and N is the nonnegative off-diagonal part, defined by

$$N_{ij} = -\Sigma_{ij}.$$

Note that the solution to Lyapunov equation is positive definite. Observe

$$x^\top S x = x^\top D x - x^\top N x > 0.$$

Set $y = D^{1/2}x$. Then:

$$x^\top D x = \|y\|^2, \quad x^\top N x = y^\top (D^{-1/2} N D^{-1/2}) y.$$

Hence,

$$y^\top (D^{-1/2} N D^{-1/2}) y < \|y\|^2 \quad \forall\, y \neq 0,$$

which implies that the symmetric matrix $D^{-1/2}ND^{-1/2}$ has spectral radius less than 1.

But since the two matrices are similar

$$\rho(D^{-1/2} N D^{-1/2}) = \rho(D^{-1}N),$$

so

$$\rho(D^{-1}N) < 1.$$

The inverse of Σ can be expressed via a Neumann series:

$$\Sigma^{-1} = (D - N)^{-1} = (I - D^{-1}N)^{-1} D^{-1} = \sum_{k=0}^{\infty} (D^{-1}N)^k D^{-1}.$$

Since $D^{-1} \geq 0$ and $D^{-1}N \geq 0$, each term in the sum is entry-wise nonnegative, so

$$\Sigma^{-1}_{i,j} \geq 0 \quad \text{for all } i, j.$$

■

Theorem 54 (Inverse of an irreducible symmetric M-matrix Berman and Plemmons (1994)) *Let Σ be symmetric, positive definite, and a Z-matrix.*

1. *Σ^{-1} is entry-wise* ***non-negative****.*
2. *Moreover, Σ^{-1} is* ***strictly positive*** *(all entries > 0) if and only if the undirected graph $\mathcal{G}(\Sigma) = \{\{i, j\} : \Sigma_{ij} < 0\}$ is* strongly connected *(equivalently, Σ is matrix-irreducible).*

Proof [Proof of Lemma 47]
We want to check conditional (in)dependence of X_a and X_b given X_C. What Lemma 46 gives us is

$$\Sigma_{ab} - \sum_{ij} \left(\Sigma_{c_i a} \left(\Sigma_{CC}^{-1} \right)_{c_i c_j} \Sigma_{b c_j} \right) \leq 0.$$

That is, because

- Case I: At least one path without a collider between a,b:
  $\Sigma_{ab} < 0$ and $\sum_{ij}(\Sigma_{c_i a}\Sigma^{-1}_{c_i c_j}\Sigma_{bc_j}) \geq 0$
- Case II: All paths with a collider between a,b, among which $\exists(i,j)$ such that c_{ij} is the collider or descendant of the collider
  $\Sigma_{ab} = 0$ and $\sum_{ij}(\Sigma_{c_i a}\Sigma^{-1}_{c_i c_j}\Sigma_{bc_j}) \geq 0$
  We need to prove that there exist (i,j) for which we have

$$(\Sigma_{c_i a}(\Sigma^{-1}_{CC})_{c_i c_j}\Sigma_{bc_j}) > 0$$

A non-singular M-matrix that is irreducible has a strictly positive inverse. A matrix is irreducible if and only if its digraph is strongly connected, i.e. every vertex is reachable from every other vertex. In our case, Σ is an M-matrix, that has a digraph defined by its non-zero entries. Each entry (i,j) of Σ is non zero if i and j are not collider away. Within block Σ_{CC} we can always reach each vertex via the path in the summary graph that is c-open, so whenever arbitrary c_i and c_j are collider away, the collider is connected to both c_i and c_j making the link in Σ. Therefore, some subset of vertices from Σ_{CC} is a strongly connected on the c-active path. However, not all of the components in C need to lie in the c-connected component. There will still exist at least one pair i,j for which $\Sigma_{ac_i}\Sigma^{-1}_{c_i c_j}\Sigma_{c_j b} > 0$. Let G_C denote the undirected sign-graph of Σ_{CC} with edges $\{i,j\}$ whenever $(\Sigma_{CC})_{ij} < 0$. Decompose Σ_{CC} according to the connected components of its sign-graph G_C. Permute C so that vertices belonging to the same connected component are contiguous:

$$\Sigma_{CC} = P^{\top}\begin{bmatrix} B_1 & 0 & \cdots & 0 \\ 0 & B_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & B_r \end{bmatrix} P.$$

Each B_ℓ is irreducible $\implies$ B_ℓ^{-1} is strictly positive. Off-block blocks in Σ_{CC} are zero, so the inverse keeps the same block structure:

$$\Sigma^{-1}_{CC} = P^{\top}\begin{bmatrix} B_1^{-1} & 0 & \cdots & 0 \\ 0 & B_2^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & B_r^{-1} \end{bmatrix} P.$$

Therefore

$$\left(\Sigma^{-1}_{CC}\right)_{ij} = \begin{cases} > 0, & \text{if } c_i, c_j \text{ lie in the same component of } G_C, \\ 0, & \text{if } c_i, c_j \text{ lie in different components.} \end{cases}$$

There exists a pair (i,j) such that $(\Sigma_{c_i a})(\Sigma^{-1}_{CC})_{c_i c_j}(\Sigma_{bc_j}) > 0$. Take any $a - b$ path and let v be its first collider. Because the path is C-open, v has a descendant in C; call it c_i. The collider-free treks $a \rightsquigarrow c_i$ and $c_i \rightsquigarrow b$ contribute negative off-diagonal covariances in a Z-matrix, so $\Sigma_{ac_i} < 0$ and $\Sigma_{c_i b} < 0$; Vertex c_i is, by construction, in the same connected component of G_C as itself, so pick $j = i$; we have $(\Sigma^{-1}_{CC})_{c_i c_i} > 0$. Multiplying the three positives yields the claim:

$$\Sigma_{ab} - \sum_{ij}\left(\Sigma_{c_i a}(\Sigma^{-1}_{CC})_{c_i c_j}\Sigma_{bc_j}\right) < 0.$$

■

Proof [Proof of Proposition 48]
(i) Direct computation gives:

$$(\pi K)(y_1, y_2) = \sum_{(x_1,x_2)} \pi(x_1, x_2)\, K\big((x_1, x_2) \to (y_1, y_2)\big) = \sum_{x_1,x_2} \pi_1(x_1)\pi_2(x_2)\, K_1(x_1 \to y_1)K_2(x_2 \to y_2)$$

$$= \left(\sum_{x_1} \pi_1(x_1)K_1(x_1 \to y_1)\right) \left(\sum_{x_2} \pi_2(x_2)K_2(x_2 \to y_2)\right) = \pi_1(y_1)\,\pi_2(y_2) = \pi(y_1, y_2).$$

Thus $\pi K = \pi$.

(ii) Because the time-t marginal of the product chain factorizes as the product measure $\pi_1 \times \pi_2$, any conditional independence among subsets of coordinates at time t holds in the product if and only if it holds in each factor. More concretely, writing the joint distribution over $(X^{(1)}(t), X^{(2)}(t))$ as $\pi_1 \otimes \pi_2$, the standard Armstrong lemma for static variables implies

$$I(X_A(t), X_B(t) \mid X_C(t)) \text{ holds under } \pi_1{\otimes}\pi_2 \iff I(X_A(t), X_B(t) \mid X_C(t)) \text{ holds under each } \pi_i.$$

Therefore the same equivalence of conditional independence statements carries over to the stationary Markov chain setting. ■