WHEN CAUSAL STRUCTURE REPLACES REPLICATION: REFERENCE-FREE FEDERATED CAUSAL AUDITING

Anonymous authors
Paper under double-blind review

ABSTRACT

In federated causal analysis, different data holders may answer different fixed conditional-independence (CI) questions without a common reference answer, so the same reports can support different explanations of the true CI answers and source alignments. We ask when relationships among CI statements can resolve this ambiguity without additional cross-source answers and when such answers remain necessary. For any specified set of possible CI answer patterns, we exactly characterize the surviving ambiguities, when relative source alignment is identified, and, under any fixed set of allowed cross-source requests, the minimum number of additional answers needed, including when identification is impossible. Holding the queries, sources, reporting graph, reporting mechanism, and allowed requests fixed, we prove that a particular four-variable pattern of CI changes and its complement cannot be the exact difference between two DAG-consistent causal explanations, even inside a larger DAG, while the same change is realizable by regular Gaussians. Repeating the construction across disjoint four-variable blocks gives a linear separation with only a linear number of fixed CI questions: DAG structure needs no additional cross-source answers, whereas regular Gaussians need exactly one per complete block, even when every missing source–query answer may be requested. Controlled exact comparisons over the complete ambient-DAG quartet family show that the replication advantage is common across broader reporting patterns, while a separate repeated-task analysis shows an analogous identification benefit under asymmetric stochastic source errors.

1 INTRODUCTION

Federated causal analysis often receives CI answers reported separately by different data holders without a common reference answer. For example, consider four hospitals, East, South, West, and North, that audit the same standardized conditional-independence (CI) questions. These questions have the same intended Yes/No answer at every hospital; the hospital-specific variation studied here is the persistent binary orientation in which that answer is reported, as can arise when local and central interfaces use opposite Yes/No conventions. The local Yes/No answers are already formed; our audit begins from those reports. Suppose East and South share answered queries, while West and North form a second group with no query shared between the groups. The same reports can then support different explanations of the correct CI answers and of whether the two hospital groups report them in the same or reversed direction. Yet CI answers are not arbitrary: one CI statement can constrain which combinations of others are possible under a DAG or another structured model. Structure can therefore sometimes avoid asking one group to answer a CI question already answered by the other, because an otherwise indistinguishable explanation may itself be structurally impossible. *When is structure alone sufficient to determine the relative alignment of the hospital groups, and when it is insufficient, how many additional cross-hospital answers are actually necessary?*

Work without gold labels already shows that when sources answer only partially overlapping tasks, their reports alone may not reveal how the sources align with the truth (Ma et al., 2018; Jaffe et al., 2015; Pilliat, 2025; Ratner et al., 2016). Federated causal-learning methods provide the distributed CI setting (Hahn et al., 2026), while a separate line studies imperfect causal experts or priors (Peng et al., 2026; Shan & Zhou, 2026). Separately, work on CI reasoning, Gaussian CI structure, and graph-specific redundancy shows that CI statements cannot vary independently (Amini et al., 2026; Kim & Lee, 2026; Boege et al., 2019; Faller & Janzing, 2026). Together, these lines supply the

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

distributed and structural ingredients of the hospital setting. The missing link is exact: for a fixed pattern of who answered which query, exactly which source ambiguities are removed by structural restrictions, and how does any ambiguity that remains determine the minimum additional cross-source evidence required?

We represent who answered which CI question by a bipartite source–query graph. In the hospital example, one disconnected group can be paired with a compatible change in the underlying CI truths without changing any reported answer; structural CI constraints determine whether such an alternative explanation is actually possible. Our first result exactly characterizes the explanations that remain indistinguishable and gives a necessary-and-sufficient condition for determining their relative orientation. If ambiguity remains, replication asks another hospital to answer an already-fixed CI question; our second result gives the exact minimum number of allowed additional answers needed to eliminate all remaining alternatives, including cases in which the permitted actions can never succeed. Finally, applying the four-variable construction across disjoint blocks gives a matched linear separation under complete cross-source replication access: ambient DAG structure requires no additional answers, whereas the regular-Gaussian family requires exactly one additional cross-source answer per complete block, using only a linear number of fixed CI questions.

The separation survives when the queried variables sit inside arbitrarily larger DAGs, including external variables and causal connections across the queried groups, because reversing only some reporting groups would require a pattern of CI changes that no ambient DAG can produce. Within controlled four-variable reporting partitions that mirror the same sparse hospital-style interface, the strict ambient-DAG replication advantage over regular Gaussians becomes more frequent as the number of reporting components increases. A separate repeated-task model shows an analogous structural identification advantage under asymmetric stochastic source errors. Together, these results show that structural restrictions on possible CI truths have operational evidence value: they can remove indistinguishable explanations before new data are collected and thereby replace cross-source evidence that would otherwise be necessary for identification. Our contributions are:

- **Exact ambiguity-to-evidence framework.** For an arbitrary feasible CI family and source–query observation graph, we characterize the surviving source–truth aliases and, under any fixed set of allowed replication actions, the exact minimum number of additional cross-source answers needed to remove them.
- **Causal exclusion under arbitrary ambient DAGs.** We prove a dimension-minimal four-variable causal exclusion: the source-group reorientation required by an ambiguous explanation cannot be realized by any pair of ambient DAGs, even with unqueried external variables and causal edges across the queried groups.
- **Complete-action linear evidence separation.** Across disjoint four-variable blocks, we construct matched audits with complete cross-source replication access and a linear number of fixed CI queries in which the ambient DAG family needs no additional answers, while the regular-Gaussian family needs exactly one additional answer per complete block.

2 Related Work

Reliability without labeled ground truth spans observer-error models, crowdsourcing, and weak supervision. Representative methods recover latent labels or source quality from repeated noisy judgments, sparse worker–task graphs, and structured weak supervision (Dawid & Skene, 1979; Raykar et al., 2010; Karger et al., 2011; Jaffe et al., 2015; Ma et al., 2018; Pilliat, 2025; Ratner et al., 2016; Polo et al., 2024). Ma et al. show in particular how sparse worker interactions govern worker-skill identifiability. Our audit starts from a fixed source–query reporting pattern: the latent object is relative source orientation under a feasible CI family, and the operational quantity is the exact additional replication required to resolve the aliases that remain.

Federated causal discovery provides the distributed setting, including privacy-aware structure learning, heterogeneous client populations or variable sets, and federated CI construction (Huang et al., 2023; Wang et al., 2023; Li et al., 2024; Wang et al., 2025; Hahn et al., 2026; Guo et al., 2026; Xu et al., 2026). fedCI is particularly close at the interface level because it constructs CI tests from distributed heterogeneous data, whereas our audit characterizes the residual source–truth equivalence class induced by sparse query coverage for a fixed set of CI answers. Work with imperfect causal experts

or priors studies how external knowledge should be trusted or combined with data (Long et al., 2023; Clivio et al., 2025; Peng et al., 2026; Shan & Zhou, 2026); the reference-free audit studied here determines what is identifiable, and what additional cross-source evidence is required, before such weighting or graph aggregation. This also separates the audit from methods whose primary object is learning or aggregating the causal graph itself.

CI reasoning and graphical redundancy exploit relations among CI statements for compact representation, deduction, consistency checking, and error detection (Amini et al., 2026; Kim & Lee, 2026; Ma et al., 2024; Faller & Janzing, 2026). Faller and Janzing are a relevant structural comparison because graph-specific CI redundancy can detect or correct errors in CI-based discovery. We use structural restrictions at a different latent level: they eliminate source–truth aliases across disconnected reporting components, and the replication theorem converts the surviving ambiguity into an exact evidence requirement. The distinction is therefore one of information interface and latent object. Regular-Gaussian CI models provide a structured comparator between unrestricted binary consistency and the additional restrictions imposed by causal-DAG semantics (Boege et al., 2019).

3 Problem Setup

We consider a finite set of conditional-independence (CI) propositions, denoted by Q, and a finite collection of reporting sources, denoted by $\mathcal{K}$. When we specialize to elementary CI queries, these have the form $X_i \perp\!\!\!\perp X_j \mid S$, with singleton endpoints and S contained in the remaining variables. The shared canonical CI truth of query $q \in Q$ is denoted by $Y_q \in \{0,1\}$, with $Y_q = 1$ meaning that the corresponding CI statement holds. Thus Y_q is the hidden Yes/No truth of query q, not a report observed by the server. We collect these truths into the complete profile $Y = (Y_q)_{q\in Q} \in \{0,1\}^Q$. All sources refer to this same canonical truth profile; within this auditing model, source-specific variation enters through the reporting mechanism defined below. The audit begins after the semantics and local answerability of these fixed queries have been established. The nonempty set of jointly possible CI truth profiles is denoted by the *feasible CI family* $\mathcal{F} \subseteq \{0,1\}^Q$. The generic identification and replication results below are stated directly for this shared feasible family. For DAG specializations, the queried variables may be embedded in a larger ambient causal system. We use the following assumption only when interpreting a DAG d-separation profile as statistical CI truth.

Assumption 1 (Shared faithful CI semantics for DAG specializations)**.** The shared canonical truth profile is induced by a population distribution that is Markov and faithful to an ambient DAG whose variable set contains every variable appearing in the queried propositions.

For a fixed n-variable ambient set containing every variable appearing in Q, $\mathcal{F}^{(n)}_{\mathrm{DAG}}(Q)$ denotes the queried d-separation profiles induced by all DAGs on that ambient set. Under Assump. 1, statistical CI truth on Q coincides with the restricted DAG d-separation pattern. The unrestricted binary family on the same query set is $\mathcal{F}_{\mathrm{Bin}}(Q) = \{0,1\}^Q$. Throughout the deterministic core, the server observes only source-local answers, has no source-independent truth reference, and each source reuses one answer polarity across all queries it answers. Sec. 4.2 introduces the fixed set of allowed replication actions; App. F gives the repeated-task extension with stochastic sensitivity and specificity channels.

The observed source–query pairs form an edge set, denoted by E, and the resulting active bipartite observation graph is $G = (\mathcal{K} \cup Q, E)$. An edge $(k,q) \in E$ means that the server observes source k's answer to query q. Every active source and query is incident to at least one observed edge. The active graph has c connected components, denoted by $C_1, \ldots, C_c$, and we write $[c] = \{1, \ldots, c\}$. The source and query vertices in component C_j form the blocks $K_j = C_j \cap \mathcal{K}$ and $B_j = C_j \cap Q$, respectively. For any query subset $A \subseteq Q$, $\mathbf{1}_A \in \{0,1\}^Q$ denotes its binary indicator mask; when no confusion is possible, we identify a query subset with this mask inside XOR and difference-set expressions.

The deterministic answer behavior of source k is represented by a source-polarity bit, denoted by $t_k \in \{0,1\}$. We collect these source polarities into $t = (t_k)_{k\in\mathcal{K}} \in \{0,1\}^{\mathcal{K}}$. The answer reported by source k to the server for query q is denoted by Z_{kq} and satisfies

$$Z_{kq} = Y_q \oplus t_k, \qquad (k,q) \in E, \tag{1}$$

where $\oplus$ denotes XOR: $t_k = 0$ preserves the canonical reporting orientation and $t_k = 1$ reverses it. Each source reuses its polarity across queries, yielding component-level joint-flip ambiguities;

query-specific polarities define a different model. The deterministic analysis assumes that the reports admit some $(Y, t) \in \mathcal{F} \times \{0,1\}^{\mathcal{K}}$. Reports admitting no such pair indicate model inconsistency, rather than unresolved relative polarity.

Within each connected component, the observed equations determine relative source polarities. We choose one representative source in each component and denote it by $a_j \in K_j$. The remaining cross-component orientation is summarized, for $c > 1$, by a vector denoted by $R(t)$:

$$R(t) = \big(t_{a_2} \oplus t_{a_1}, \ldots, t_{a_c} \oplus t_{a_1}\big), \tag{2}$$

and $R(t)$ is the empty vector when $c = 1$. Together with within-component relations, this vector determines every pairwise relative polarity $t_k \oplus t_\ell$. Our main identification target is therefore the relative orientation $R(t)$ across the initially disconnected reporting components; the next section develops the ambiguity geometry needed to determine when the observed reports and the feasible CI family fix this target.

4 Structural Constraints for Reference-Free Identification and Replication

The source–query graph determines which component-level source–truth changes are observationally invisible, while the feasible CI family determines which of those changes are compatible with any underlying truth. This separates the audit into two questions: which aliases survive structure, and what additional evidence is required to remove the survivors? We first characterize the exact alias set and the exact action-constrained replication cost. We then exploit a causal-specific rigidity of ambient DAG marginals to eliminate a local two-world change that remains realizable for regular Gaussians, yielding the matched linear evidence separation.

4.1 Exact Observational Alias Geometry

Two ingredients determine whether the reference-free ambiguity survives. The source–query graph determines which simultaneous component flips cannot be seen from the local answers, while the feasible CI family $\mathcal{F}$ determines which of those flips still correspond to feasible latent truth profiles. Connectivity therefore creates candidate source–truth symmetries, and CI geometry removes the candidates that cannot occur under the assumed structural family.

Ambiguity geometry. A latent configuration that respects the feasible CI family is called a *feasible world* and is written $w = (Y, t) \in \mathcal{F} \times \{0,1\}^{\mathcal{K}}$. The complete vector of answers observed on E is denoted by $Z_E(Y, t) = \big(Y_q \oplus t_k\big)_{(k,q)\in E}$. Two feasible worlds are observationally equivalent when they produce the same observed vector; the equivalence class of w is denoted by $[w]_{\mathrm{obs}}$.

For a component-selection vector $h = (h_1, \ldots, h_c) \in \{0,1\}^c$, define the query mask $h_Q = \bigoplus_{j:h_j=1} \mathbf{1}_{B_j}$ and the source mask $h_{\mathcal{K}} \in \{0,1\}^{\mathcal{K}}$ by $h_{\mathcal{K}}(k) = h_j$ for every $k \in K_j$. The simultaneous component flip is $\Phi_h(Y, t) = \big(Y \oplus h_Q, t \oplus h_{\mathcal{K}}\big)$. It leaves every observed source–query answer unchanged. The feasible truth-difference set is

$$\Delta(\mathcal{F}) = \{y \oplus y' : y, y' \in \mathcal{F}\}. \tag{3}$$

Thus, $\Delta(\mathcal{F})$ records exactly which coordinate-wise truth changes can separate two feasible latent truth profiles. A component flip can survive the structural constraints only when its induced query mask belongs to this set. Relative source polarity is *universally identified over* $\mathcal{F}$ when the reports determine all pairwise source orientations at every feasible world.

Proposition 1 (Exact observational alias geometry)**.** *For every nonempty feasible CI family $\mathcal{F} \subseteq \{0,1\}^Q$ and every feasible world $w = (Y, t)$,*

$$[w]_{\mathrm{obs}} = \{\Phi_h(Y, t) : h \in \{0,1\}^c,\ Y \oplus h_Q \in \mathcal{F}\}\,. \tag{4}$$

Consequently, relative source polarity is universally identified over $\mathcal{F}$ if and only if

$$\mathbf{1}_{\cup_{j\in S} B_j} \notin \Delta(\mathcal{F}) \qquad \text{for every } \varnothing \neq S \subsetneq [c]. \tag{5}$$

Proof sketch. Equal reports imply, edge by edge, that the truth discrepancy and the source-polarity discrepancy coincide. Along each connected component this forces one common discrepancy bit, so the observation graph alone generates exactly the simultaneous component flips Φ_h. The feasible family then filters those candidates: a flip survives precisely when its induced truth mask maps one feasible CI profile to another. Relative polarity is therefore identified exactly when no nonempty proper component union survives this filter, which gives Eq. (5). The complete proof is in App. B.

Interpretation for general targets. For any target space $\mathcal{G}$ and functional $g : \mathcal{F} \times \{0,1\}^{\mathcal{K}} \to \mathcal{G}$, the value $g(w)$ is identified at a feasible world w if and only if g is constant on the exact observational orbit $[w]_{\mathrm{obs}}$ in Eq. (4). Thus the orbit characterization applies to arbitrary world functionals, while relative source polarity is the principal target studied here.

Remark 1 (Structure can replace cross-group overlap). In the four-hospital example, suppose East and South form one observed component and West and North another, with no query shared across the two groups. The observation graph alone permits flipping one group's latent truths and reporting orientations together. If the corresponding query mask cannot occur as the difference between two feasible CI profiles, the alternative explanation is structurally impossible and the two hospital groups are oriented relative to one another without any new repeated answer. Structure and cross-group overlap are therefore two distinct sources of identification.

For the unrestricted task family $\mathcal{F}_{\mathrm{Bin}}(Q) = \{0,1\}^Q$, every component-union mask belongs to $\Delta(\mathcal{F}_{\mathrm{Bin}}(Q))$, so disconnected observation graphs retain the usual reference-free ambiguity. The most direct specialization is the two-group setting used in the hospital example.

Two-group specialization. When G has exactly two connected components with query blocks B_1 and B_2, Prop. 1 reduces to the direct test

$$\mathbf{1}_{B_1} \notin \Delta(\mathcal{F}) \quad \text{and} \quad \mathbf{1}_{B_2} \notin \Delta(\mathcal{F}). \tag{6}$$

Thus, in the East–South versus West–North example, the server only needs to ask whether flipping either group's entire queried block can occur between two feasible CI profiles.

Interface-level consequence. If a nonempty proper component-union mask does lie in $\Delta(\mathcal{F})$, then two feasible worlds with different relative source polarities induce the same observed report vector. No downstream processing of those unchanged reports can universally distinguish them, so resolving the alias requires new information, for example through a cross-component replication. A simultaneous global source flip leaves relative polarity unchanged; whether the corresponding complemented truth profile is also feasible depends on $\mathcal{F}$.

4.2 Exact Action-Constrained Replication

Prop. 1 identifies the ambiguities that remain after CI structure is used. Replication can remove such ambiguity by adding an answer that links previously disconnected source–query components. We therefore ask for the minimum number of replications allowed by a fixed audit interface that reaches an observation graph whose remaining flips that change relative source polarity are all structurally excluded.

An *allowed replication action* adds one previously unobserved source–query edge (k, q) and reveals source k's answer to the already-fixed proposition q. The fixed action set is $\mathcal{R} \subseteq (\mathcal{K} \times Q) \setminus E$. Each added answer has unit cost, so replication cost throughout the deterministic core means the number of source–query edges added beyond the initial observation graph E. The component pairs joined by at least one action form an action graph $H_{\mathcal{R}} = ([c], L_{\mathcal{R}})$, whose edge set $L_{\mathcal{R}}$ contains component pairs connected by at least one allowed action. For a selected set of allowed replications $E_{\mathrm{add}} \subseteq \mathcal{R}$, let $G + E_{\mathrm{add}}$ denote the observation graph obtained after adding those source–query edges. The exact minimum replication cost is

$$\mathsf{A}_{\mathcal{F},\mathcal{R}}(G) = \min\left\{|E_{\mathrm{add}}| : E_{\mathrm{add}} \subseteq \mathcal{R},\ G + E_{\mathrm{add}} \text{ universally identifies relative polarity}\right\}, \tag{7}$$

with the minimum of an empty set equal to $+\infty$.

A partition $\mathcal{P}$ of $[c]$ is called a *coarsening* of the initial observation components when each block $U \in \mathcal{P}$ specifies a collection of initial connected components to be merged. For each block $U \in \mathcal{P}$, define its associated query block by $B_U := \bigcup_{j \in U} B_j$. The coarsening is *$\mathcal{F}$-identifying* when

$$\mathbf{1}_{\cup_{U \in \mathcal{S}} B_U} \notin \Delta(\mathcal{F}) \quad \text{for every } \varnothing \neq \mathcal{S} \subsetneq \mathcal{P}, \tag{8}$$

and $\mathcal{R}$-*attainable* when each block U induces a connected subgraph $H_{\mathcal{R}}[U]$. The sets of $\mathcal{F}$-identifying and $\mathcal{R}$-attainable coarsenings are denoted by $\mathfrak{I}_{\mathcal{F}}(G)$ and $\mathfrak{C}_{\mathcal{R}}(G)$, respectively.

Theorem 1 (Exact action-constrained replication complexity)**.** *For every nonempty feasible family $\mathcal{F}$ and every fixed action set $\mathcal{R}$, the value is $+\infty$ when no attainable identifying coarsening exists; otherwise,*

$$\mathsf{A}_{\mathcal{F},\mathcal{R}}(G) = \min_{\mathcal{P}\in\mathfrak{I}_{\mathcal{F}}(G)\cap\mathfrak{C}_{\mathcal{R}}(G)} \big(c - |\mathcal{P}|\big). \tag{9}$$

Proof sketch. Every successful audit ends at some coarsening $\mathcal{P}$ of the initial observation components. Because one added answer can merge at most two current components, reaching $\mathcal{P}$ costs at least $c - |\mathcal{P}|$ actions. Conversely, if $\mathcal{P}$ is attainable, a spanning tree inside each block of $H_{\mathcal{R}}$ realizes exactly that coarsening with $\sum_{U\in\mathcal{P}}(|U| - 1) = c - |\mathcal{P}|$ allowed replications; if it is also identifying, Prop. 1 guarantees that no observational alias changing relative source polarity remains. Minimizing over the attainable identifying coarsenings proves the formula. App. C gives the full proof and the two monotonicity consequences.

Remark 2 (Replication cost counts necessary component mergers). Return to the four-hospital picture. Repeating a question within the East–South group does not help align that group with West–North; a useful repetition must bridge components whose relative flip is still structurally feasible. If causal constraints already rule out that flip, no bridge is needed. Starting from c disconnected observation components, each useful cross-component replication can reduce the component count by at most one, so reaching a coarsening with $|\mathcal{P}|$ blocks requires at least $c - |\mathcal{P}|$ component mergers; an attainable coarsening achieves this bound with a spanning tree inside each block. Thus $c - |\mathcal{P}|$ counts the physical cross-component links whose information cannot be supplied by structure, while $\mathcal{R}$ records which such links the auditor is actually allowed to request.

Corollary 1 (Reachability and complete actions)**.** *Let $\mathcal{P}(H_{\mathcal{R}})$ be the connected-component partition of the action graph. Then $\mathsf{A}_{\mathcal{F},\mathcal{R}}(G) < \infty$ if and only if $\mathcal{P}(H_{\mathcal{R}})$ is $\mathcal{F}$-identifying. Under the complete action interface $\mathcal{R}_{\text{complete}} := (\mathcal{K}\times Q)\setminus E$, every coarsening is attainable and*

$$\mathsf{A}_{\mathcal{F}}(G) := \mathsf{A}_{\mathcal{F},\mathcal{R}_{\text{complete}}}(G) = \min_{\mathcal{P}\in\mathfrak{I}_{\mathcal{F}}(G)} (c - |\mathcal{P}|). \tag{10}$$

Remark 3 (Structural ambiguity and action constraints are distinct bottlenecks). The feasible family and the allowed action set restrict different parts of the audit. Suppose East–South and West–North still admit a relative flip after all CI constraints are used. If the protocol permits no repeated query capable of connecting the relevant components, the ambiguity is structurally real and operationally unreachable, so the exact cost is $+\infty$. Under complete actions the reachability bottleneck disappears; the remaining cost then measures only how many component mergers are still required after structure has removed all aliases it can. Thm. 1 therefore separates "what the model leaves ambiguous" from "what evidence the protocol allows us to collect."

The characterizations above are information-theoretic. Constructing $\Delta(\mathcal{F})$ and optimizing identifying coarsenings for a large implicitly specified feasible family are separate computational tasks; in the structured causal construction below, the relevant alias is instead ruled out by a local analytic rigidity argument.

4.3 All-Ambient Causal Two-World Rigidity

The preceding identification and replication characterizations are dimension-agnostic. We now use four queried variables to construct an exact causal–Gaussian witness and show that the resulting local exclusion survives arbitrary ambient DAG structure. The key is a two-world difference-set exclusion that remains valid even when the queried variables interact arbitrarily with variables outside the quartet. Let Q_4 be the 24 elementary CI propositions on $U = \{A, B, C, D\}$, and define the four-query blocking set $M_{\text{blk}} \subset Q_4$ by

$$M_{\text{blk}} = \{A \perp\!\!\!\perp D \mid BC,\ B \perp\!\!\!\perp C,\ B \perp\!\!\!\perp C \mid A,\ B \perp\!\!\!\perp C \mid D\}. \tag{11}$$

This four-coordinate pattern couples three $B \perp\!\!\!\perp C$ statements across different conditioning contexts with the statement $A \perp\!\!\!\perp D \mid BC$. The theorem below shows that this joint flip is incompatible with ambient DAG semantics, while the regular-Gaussian construction used for the matched separation can realize the same exact four-coordinate change. Let $\mathcal{F}_{\text{DAG-marg}}(Q_4) := \bigcup_{m\geq 4} \mathcal{F}^{(m)}_{\text{DAG}}(Q_4)$ denote all complete Q_4 profiles obtained by restricting the d-separation model of an arbitrary finite

ambient DAG to the quartet U. The proof uses standard CI-consistency rules together with ancestral-graph marginalization, ordered stability, and DAG-marginal chordality (Richardson & Spirtes, 2002; Sadeghi, 2017; Sadeghi & Soo, 2022; de Waal, 2009); the theorem below establishes the two-world exclusion these ingredients imply for the auditing interface.

Theorem 2 (All-ambient causal two-world rigidity)**.** *Neither M_{blk} nor its quartet complement can be the exact difference between two ambient-DAG quartet profiles:*

$$M_{\mathrm{blk}} \notin \Delta(\mathcal{F}_{\text{DAG-marg}}(Q_4)), \qquad Q_4 \setminus M_{\mathrm{blk}} \notin \Delta(\mathcal{F}_{\text{DAG-marg}}(Q_4)). \tag{12}$$

The statement allows arbitrary ambient variables and DAG edges involving or crossing the queried quartet; a census of four-node DAGs would not establish it, because the quartet restriction of a larger DAG need not coincide with the CI model of any DAG on those four variables alone.

Proof sketch. Assume two candidate ambient-DAG worlds differ exactly by one of the two masks in Eq. (12). For M_{blk}, generic CI consistency first synchronizes the three disputed $B \perp\!\!\!\perp C$ statements: one world must realize the pattern $(0, 1, 1)$ across $B \perp\!\!\!\perp C$, $B \perp\!\!\!\perp C \mid A$, and $B \perp\!\!\!\perp C \mid D$, while the other realizes $(1, 0, 0)$. The shared outer statement $B \perp\!\!\!\perp C \mid AD$ then forces one of the two conditioning squares to be 0110 or 1001; both patterns are forbidden for an ancestral-graph marginal. For the complementary mask, the three-variable restrictions in the two worlds are truth-value complements, so the trivariate antipode rule forces a full-versus-empty configuration. After composition and singleton transitivity fix the shared quartet relations, Pearl chordality requires at least one of $A \perp\!\!\!\perp D \mid B$ or $A \perp\!\!\!\perp D \mid C$, whereas the complementary mask forces both to be false. Thus neither exact two-world change can arise from an ambient DAG, regardless of external variables. App. D gives the full analytic argument.

Remark 4 (Local certificate with arbitrary ambient scope). Suppose East–South and West–North are two reporting groups and an alternative explanation reverses one group while preserving every reported answer. On a queried quartet, that explanation must replace the latent CI truth by a second profile that toggles exactly the four statements in M_{blk}, or exactly the complementary twenty. Thm. 2 says that neither replacement can arise from two ambient DAGs, even when unqueried variables connect arbitrarily through the quartet. The quartet is therefore a local certificate inside an unrestricted ambient causal system, rather than a restriction of the causal model to four variables. As shown below, four variables are moreover the minimal local dimension at which this matched DAG–Gaussian distinction can occur.

Minimality. On three variables, the elementary-CI profiles realized by faithful DAGs and by positive-definite regular Gaussians coincide. Hence four queried variables are the minimal local dimension at which this matched DAG–Gaussian separation can occur. In the hospital interpretation, three queried variables therefore cannot expose a matched DAG–Gaussian structural distinction of this kind; four are the first possible local witness. App. D gives the analytic three-variable argument.

4.4 Complete-Action Linear Evidence Separation

Thm. 2 is a local impossibility statement; by itself it does not yet quantify how much evidence that impossibility saves. We now repeat the same local interface across disjoint quartets and compare ambient DAGs with a broader regular-Gaussian family that can realize the blocked change. For a fixed query set Q over n ambient variables, let $\mathcal{F}^{(n)}_{\mathrm{DAG}}(Q)$ denote the queried d-separation profiles induced by all DAGs on those variables, and let $\mathcal{F}^{(n)}_{\mathrm{G}}(Q)$ denote the queried CI profiles realizable by positive-definite regular Gaussian distributions on the same variables. Faithful regular-Gaussian realizations of DAGs give $\mathcal{F}^{(n)}_{\mathrm{DAG}}(Q) \subseteq \mathcal{F}^{(n)}_{\mathrm{G}}(Q)$ (Spirtes et al., 2000). For compactness in this subsection, write $\mathsf{A}_{\mathrm{DAG},\mathcal{R}}(G) := \mathsf{A}_{\mathcal{F}^{(n)}_{\mathrm{DAG}}(Q),\mathcal{R}}(G)$ and $\mathsf{A}_{\mathrm{G},\mathcal{R}}(G) := \mathsf{A}_{\mathcal{F}^{(n)}_{\mathrm{G}}(Q),\mathcal{R}}(G)$ whenever the common ambient size and query set are clear.

Fix $n \geq 4$ and set $s = \lfloor n/4 \rfloor$. Choose s disjoint quartets $U_i = \{A_i, B_i, C_i, D_i\}$. Let $Q(U_i)$ denote the 24 elementary CI propositions internal to quartet U_i, and set the query universe to the disjoint union $Q = \bigsqcup_{i=1}^{s} Q(U_i)$. Let M_i be the relabelled copy of M_{blk} on U_i. One central source answers all queries in $\bigcup_{i=1}^{s}(Q(U_i) \setminus M_i)$; each satellite $i \in [s]$ answers the four queries in M_i. The resulting observation graph G_n has one central component and s satellite components, uses $24s \leq 6n$ queries, and has conditioning order at most two. Both structural families receive the same complete replication interface: any source may answer any already-fixed queried proposition it did not initially answer.

Theorem 3 (Complete-action linear evidence separation)**.** *For every $n \geq 4$,*

$$\mathsf{A}_{\mathrm{DAG},\mathcal{R}_{\mathrm{complete}}}(G_n) = 0, \qquad \mathsf{A}_{\mathrm{G},\mathcal{R}_{\mathrm{complete}}}(G_n) = \lfloor n/4 \rfloor. \tag{13}$$

The DAG side ranges over all ambient n-variable DAGs, while the two sides use the same CI propositions, reporting sources, initial observation graph, reporting mechanism, and complete replication interface.

Proof sketch. The observation graph has one central component and s satellite components. Any reorientation that changes relative polarity flips at least one satellite against the center, which on that quartet requires exactly M_i or $Q(U_i) \setminus M_i$; Thm. 2 rules out both for ambient DAGs, so no replication is needed. For regular Gaussians, an explicit rational positive-definite pair realizes exactly M_{blk}, and block products allow any subset of satellites to flip independently. All s relative satellite alignments therefore survive. Under complete actions, Thm. 1 says that resolving $s+1$ components requires exactly s mergers, achieved by a spanning tree. App. D gives the full construction and proof.
Remark 5 (Structure replaces evidence even under unrestricted replication). In the hospital construction, one central hospital and s satellite hospitals form $s+1$ initially disconnected reporting components. Even if the auditor may ask any hospital any already-fixed CI question, the regular-Gaussian family leaves each satellite's alignment with the center unresolved, so one effective cross-component bridge per satellite is still necessary. Ambient DAG semantics rule out those same alternative alignments before any new answer is collected. Thus the linear gap measures physical cross-hospital evidence replaced by causal structure even when every cross-source replication action is available.

Whenever these families are compared on the same fixed query set, observation graph, and action set, family inclusion gives $\mathsf{A}_{\mathcal{F}^{(n)}_{\mathrm{DAG}}(Q),\mathcal{R}}(G) \leq \mathsf{A}_{\mathcal{F}^{(n)}_{\mathrm{G}}(Q),\mathcal{R}}(G) \leq \mathsf{A}_{\mathcal{F}_{\mathrm{Bin}}(Q),\mathcal{R}}(G)$. Thm. 3 strengthens this generic monotonicity into an exact linear matched gap under complete replication access. The matched gap is driven by the additional graph-specific CI restrictions of ambient DAGs, through which component-union masks occur in $\Delta(\mathcal{F})$, rather than by family cardinality alone.

Alternative structural mechanism. App. E gives a second arbitrary-dimensional route to zero-replication identification using standard compositional-graphoid closure. When source–query components are the elementary-CI conditioning-cardinality layers, two feasible profiles cannot differ by a nonconstant pattern of complete layer flips; every nontrivial component-union alias is therefore excluded. This conditioning-layer mechanism complements the four-coordinate blocking-mask mechanism used in the main causal separation.

Repeated-task stochastic extension. App. F asks whether structural restrictions can still change identifiability when deterministic source polarity is replaced by reusable asymmetric sensitivity/specificity channels across repeated tasks. This comparison is distinct from the four-variable DAG–Gaussian minimality result above: it contrasts a three-variable faithful-DAG family with the unrestricted truth family under a repeated-task stochastic reporting model. Under the unrestricted six-bit truth family, interior channel changes can be absorbed by changes in the latent truth-profile mixture, producing local non-identifiability. Under the complete 11-profile three-variable faithful-DAG CI family, the same repeated-task interface is generically locally identifiable on a nonempty interior set; an exact nonzero 14×14 Jacobian minor supplies the local-rank witness. Its scope is population-level local identifiability for repeated fixed-format tasks with reusable source-channel parameters. This supporting result shows that the structural identification effect extends beyond deterministic source polarity to a repeated-task stochastic setting.

5 Experiments

The theory gives worst-case and exact identification statements; the experiments complement them by measuring how much replication structure saves on controlled reporting interfaces and how often ambient-DAG structure improves on the already structured regular-Gaussian family. All replication costs use complete actions and are computed exactly, while a supporting finite-sample check evaluates the repeated-task extension.

5.1 Controlled structural comparison

We compare the complete 248-profile ambient-DAG quartet family on the 24 elementary CI queries with the 629 regular-Gaussian profiles and the unrestricted binary family. For each component

count $c \in \{2, 3, 4, 6, 8\}$, all three families use the same 5,000 partitions (25,000 total), sampled through uniform labeled surjections onto c nonempty blocks, then canonicalized and deduplicated within each c. We write the corresponding complete-action costs as $\mathsf{A}_{\mathrm{Amb}}$, A_{G}, and $\mathsf{A}_{\mathrm{Bin}}$; on every design, $\mathsf{A}_{\mathrm{Amb}} \leq \mathsf{A}_{\mathrm{G}} \leq \mathsf{A}_{\mathrm{Bin}} = c - 1$. App. G retains the 185-profile four-node DAG family as a narrower-scope sensitivity comparison.

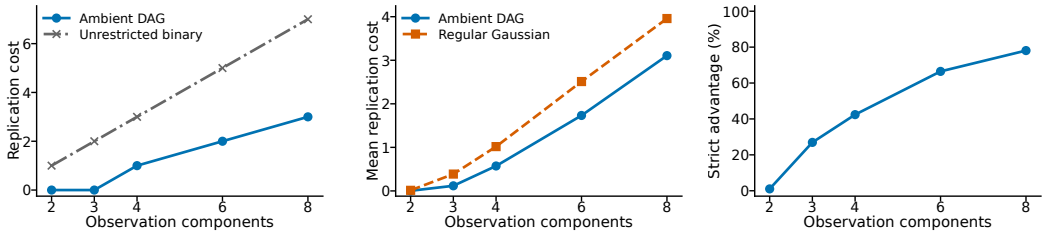


Figure 1: Complete-action replication costs on the same 25,000 sampled reporting partitions. (a) Ambient-DAG structure reduces replication relative to unrestricted binary truths. (b) Regular Gaussians have higher mean costs than ambient DAGs. (c) The empirical frequency of $\mathsf{A}_{\mathrm{Amb}} < \mathsf{A}_{\mathrm{G}}$ increases with c, reaching 78.08% at $c = 8$. Table 1 in App. G gives the summary statistics and four-node DAG sensitivity comparison.

Fig. 1 shows that the ambient-DAG family remains strictly cheaper than the regular-Gaussian family on an increasing fraction of reporting designs, reaching 78.08% at $c = 8$. At that point the mean costs are 3.1062 and 3.9580, respectively. Allowing arbitrary ambient variables weakens the advantage relative to the four-node DAG specialization but does not remove it; App. G gives the full comparison.

5.2 Targeted validation

The targeted checks complement the analytic results with finite constructions and a supporting numerical evaluation. As an independent finite cross-check of Thm. 2, we exhaustively enumerate the labeled five-variable DAGs, treat each variable in turn as unqueried, and restrict the remaining quartet to its 24 elementary CI propositions. Across all five choices, neither M_{blk} nor its quartet complement occurs as an exact two-world difference. This finite check corroborates the all-ambient exclusion on the smallest nontrivial ambient extension.

For the Gaussian side, exact rational arithmetic verifies that the two displayed covariance matrices in App. D are positive definite and that their 24-query profiles differ on exactly the four coordinates in M_{blk}; for Prop. 3, finite-sample fits yield mean source-channel root-mean-square error (RMSE) values of 0.0993 at $R = 100$ repeated tasks and 0.00757 at $R = 100{,}000$. Thus the controlled ambient-quartet study quantifies the replication-cost gap, the five-variable check probes its first ambient extension, the repeated-task check complements the supporting stochastic identification result, and Thm. 3 gives the analytic arbitrary-dimensional scaling. App. G reports the controlled-partition details and targeted validation results.

6 Conclusion

Reference-free federated causal auditing of fixed CI reports can leave source–truth explanations compatible with the same sparse reports. We showed that the assumed CI structure exactly determines which relative source ambiguities remain and that any surviving ambiguity induces an exact minimum number of additional cross-source answers under fixed admissible replication actions. For arbitrary ambient DAGs, a causal two-world rigidity theorem removes a local ambiguity that remains feasible for regular Gaussians; on a matched complete-action interface, this yields zero additional replications versus $\lfloor n/4 \rfloor$. Controlled exact comparisons show that the ambient-DAG evidence advantage recurs across sampled reporting partitions. A supporting repeated-task stochastic extension shows an analogous identification advantage under asymmetric source errors. Causal CI structure therefore has operational evidence value: it can replace cross-source evidence that would otherwise be necessary for identification.

AI USE STATEMENT

AI was used to assist with research refinement, theoretical formulation and proof checking, experimental design, analysis and interpretation of experimental results, literature organization, and manuscript polishing. All AI-assisted mathematical claims, proofs, experimental protocols, reported results, and manuscript content were reviewed and validated by the authors. The authors take full responsibility for the correctness, originality, and integrity of the work.

REPRODUCIBILITY STATEMENT

Complete proofs and theoretical specifications are provided in the appendices. The accompanying code package provides the code needed to regenerate the controlled corpus and reproduce the computational experiments, exact finite validations, reported figures, and tables.

REFERENCES

Arash A. Amini, Bryon Aragam, and Qing Zhou. Learning general conditional independence structures via the neighbourhood lattice. *Journal of Machine Learning Research*, 27(117):1–42, 2026.

Tobias Boege. Gaussoids are two-antecedental approximations of Gaussian conditional independence structures. *Annals of Mathematics and Artificial Intelligence*, 90:645–673, 2022.

Tobias Boege, Alessio D'Alì, Thomas Kahle, and Bernd Sturmfels. The geometry of gaussoids. *Foundations of Computational Mathematics*, 19(4):775–812, 2019.

Oscar Clivio, Divyat Mahajan, Perouz Taslakian, Sara Magliacane, Ioannis Mitliagkas, Valentina Zantedeschi, and Alexandre Drouin. Learning to defer for causal discovery with imperfect experts. In *ICLR 2025 Workshop on Reasoning and Planning for Large Language Models*, 2025.

A. P. Dawid and A. M. Skene. Maximum likelihood estimation of observer error-rates using the EM algorithm. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 28(1):20–28, 1979. doi: 10.2307/2346806.

Peter R. de Waal. Marginals of DAG-isomorphic independence models. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, volume 5590 of *Lecture Notes in Computer Science*, pp. 192–203. Springer, 2009. doi: 10.1007/978-3-642-02906-6_18.

Philipp M. Faller and Dominik Janzing. On different notions of redundancy in conditional-independence-based discovery of graphical models. In *Proceedings of the 29th International Conference on Artificial Intelligence and Statistics*, volume 300 of *Proceedings of Machine Learning Research*. PMLR, 2026.

Xianjie Guo, Yuwei Wang, Guodu Xiang, Xiaoli Tang, Kui Yu, Han Yu, and Qiang Yang. A survey on federated causal discovery and inference. *arXiv preprint arXiv:2606.23741*, 2026.

Maximilian Hahn, Alina Zajak, Dominik Heider, and Adèle Helena Ribeiro. Federated causal discovery across heterogeneous datasets under latent confounding. *arXiv preprint arXiv:2603.05149*, 2026.

Jianli Huang, Xianjie Guo, Kui Yu, Fuyuan Cao, and Jiye Liang. Towards privacy-aware causal structure learning in federated setting. *IEEE Transactions on Big Data*, 9(6):1525–1535, 2023. doi: 10.1109/TBDATA.2023.3285477.

Ariel Jaffe, Boaz Nadler, and Yuval Kluger. Estimating the accuracies of multiple classifiers without labeled data. In *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics*, volume 38 of *Proceedings of Machine Learning Research*, pp. 407–415. PMLR, 2015.

David R. Karger, Sewoong Oh, and Devavrat Shah. Iterative learning for reliable crowdsourcing systems. In *Advances in Neural Information Processing Systems*, volume 24, pp. 1953–1961, 2011.

Jonghwan Kim and Sanghack Lee. Don't test what you can deduce: Causal discovery with logical inference. In *Proceedings of the Conference on Uncertainty in Artificial Intelligence*, 2026. Oral presentation.

Loka Li, Ignavier Ng, Gongxu Luo, Biwei Huang, Guangyi Chen, Tongliang Liu, Bin Gu, and Kun Zhang. Federated causal discovery from heterogeneous data. In *International Conference on Learning Representations*, 2024.

Radim Lněnička and František Matúš. On Gaussian conditional independence structures. *Kybernetika*, 43(3):327–342, 2007.

Stephanie Long, Alexandre Piché, Valentina Zantedeschi, Tibor Schuster, and Alexandre Drouin. Causal discovery with language models as imperfect experts. In *ICML 2023 Workshop on Structured Probabilistic Inference and Generative Modeling (SPIGM)*, 2023.

Pingchuan Ma, Zhenlan Ji, Peisen Yao, Shuai Wang, and Kui Ren. Enabling runtime verification of causal discovery algorithms with automated conditional independence reasoning. In *Proceedings of the 46th IEEE/ACM International Conference on Software Engineering*. IEEE Computer Society, 2024. doi: 10.1145/3597503.3623348.

Yao Ma, Alexander Olshevsky, Csaba Szepesvari, and Venkatesh Saligrama. Gradient descent for sparse rank-one matrix completion for crowd-sourced aggregation of sparsely interacting workers. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 3335–3344. PMLR, 2018.

Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, 1988.

Bo Peng, Kaiwen Wu, Sirui Chen, Zhiheng Wang, Yu Qiao, and Chaochao Lu. CauTion: Knowing when to trust LLMs for ensemble causal discovery. *arXiv preprint arXiv:2606.03602*, 2026.

Emmanuel Pilliat. Recovering labels from crowdsourced data: an optimal and polynomial-time method. In *Proceedings of the Thirty-Eighth Conference on Learning Theory*, volume 291 of *Proceedings of Machine Learning Research*, pp. 4565–4595. PMLR, 2025.

Felipe Maia Polo, Subha Maity, Mikhail Yurochkin, Moulinath Banerjee, and Yuekai Sun. Weak supervision performance evaluation via partial identification. In *Advances in Neural Information Processing Systems*, volume 37, 2024.

Alexander J. Ratner, Christopher M. De Sa, Sen Wu, Daniel Selsam, and Christopher Ré. Data programming: Creating large training sets, quickly. In *Advances in Neural Information Processing Systems*, volume 29, 2016.

Vikas C. Raykar, Shipeng Yu, Linda H. Zhao, Gerardo Hermosillo Valadez, Charles Florin, Luca Bogoni, and Linda Moy. Learning from crowds. *Journal of Machine Learning Research*, 11(43): 1297–1322, 2010.

Thomas Richardson and Peter Spirtes. Ancestral graph Markov models. *The Annals of Statistics*, 30 (4):962–1030, 2002. doi: 10.1214/aos/1031689015.

Kayvan Sadeghi. Faithfulness of probability distributions and graphs. *Journal of Machine Learning Research*, 18(148):1–29, 2017.

Kayvan Sadeghi and Terry Soo. Conditions and assumptions for constraint-based causal structure learning. *Journal of Machine Learning Research*, 23:1–34, 2022.

Xihang Shan and Da Zhou. PRCD-MAP: Learning how much to trust imperfect priors in causal discovery. *arXiv preprint arXiv:2605.01669*, 2026.

Peter Spirtes, Clark Glymour, and Richard Scheines. *Causation, Prediction, and Search*. MIT Press, Cambridge, MA, 2 edition, 2000.

Yunxia Wang, Fuyuan Cao, Kui Yu, and Jiye Liang. Federated causal structure learning with non-identical variable sets. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 62445–62466. PMLR, 2025.

Zhaoyu Wang, Pingchuan Ma, and Shuai Wang. Towards practical federated causal structure learning. In *Machine Learning and Knowledge Discovery in Databases: Research Track*, pp. 351–367. Springer, 2023. doi: 10.1007/978-3-031-43415-0_21.

Jiegang Xu, Fuyuan Cao, and Jiye Liang. Credibility-aware weighting federated causal discovery for time series. In *Forty-third International Conference on Machine Learning*, 2026.

Appendix

A Notation and Proof Roadmap

This appendix provides a compact notation reference for the proofs; the model and assumptions are defined in the main text.

A.1 Notation at a glance

The tables below collect notation used across the deterministic core, replication result, DAG–Gaussian construction, and repeated-task extension. Formal definitions and assumptions remain in the main text.

Core CI and source objects.

Symbol	Meaning
Q	Fixed finite set of conditional-independence (CI) queries.
$\mathcal{K}$	Finite set of reporting sources.
Y_q	Shared canonical CI truth of query q, with $Y_q = 1$ meaning that the CI statement holds.
$Y = (Y_q)_{q\in Q}$	Complete canonical CI truth profile.
$\mathcal{F}$	Feasible CI family, a nonempty subset of $\{0,1\}^Q$.
$\mathcal{F}_{\mathrm{Bin}}(Q)$	Unrestricted binary family $\{0,1\}^Q$.
t_k	Deterministic source polarity of source k.
$t = (t_k)_{k\in\mathcal{K}}$	Complete deterministic source-polarity vector.
Z_{kq}	Observed answer from source k to query q, equal to $Y_q \oplus t_k$ in the deterministic core.
$w = (Y, t)$	Feasible latent world.

Observation graph and ambiguity.

Symbol	Meaning
$G = (\mathcal{K} \cup Q, E)$	Bipartite source–query observation graph.
E	Observed source–query pairs.
$C_1, \dots, C_c$	Connected components of the active observation graph.
c	Number of connected components of the active observation graph.
K_j	Source block $C_j \cap \mathcal{K}$ of component j.
B_j	Query block $C_j \cap Q$ of component j.
$[c]$	Index set $\{1, \dots, c\}$.
$h \in \{0,1\}^c$	Selection vector specifying which connected components are flipped.
h_Q	Query-coordinate mask induced by h.
$h_{\mathcal{K}}$	Source-coordinate mask defined by $h_{\mathcal{K}}(k) = h_j$ for $k \in K_j$.
Φ_h	Simultaneous component flip $(Y, t) \mapsto (Y \oplus h_Q, t \oplus h_{\mathcal{K}})$.
$\Delta(\mathcal{F})$	XOR difference set $\{y \oplus y' : y, y' \in \mathcal{F}\}$.
$Z_E(Y, t)$	Complete observed answer vector on E.
$\sim_{\mathrm{obs}}$	Observational-equivalence relation induced by equality of Z_E.
$[w]_{\mathrm{obs}}$	Observational equivalence class of world w.
a_j	Chosen anchor source in component C_j, used to define $R(t)$.
$R(t)$	Vector of cross-component relative source polarities defined from the anchors a_j.

Replication and component coarsening.

Symbol	Meaning
$\mathcal{R}$	Fixed set of allowed replication actions.
$H_{\mathcal{R}} = ([c], L_{\mathcal{R}})$	Action graph on the initial observation components.
$L_{\mathcal{R}}$	Component pairs connected by at least one allowed replication action.
E_{add}	Selected additional replication edges, with $E_{\mathrm{add}} \subseteq \mathcal{R}$.
$G + E_{\mathrm{add}}$	Observation graph after adding the selected replications.
$\mathcal{P}$	Partition of $[c]$ whose blocks specify a coarsening of the initial observation components.
$U \in \mathcal{P}$	One block of a coarsening.
B_U	Associated query block $\bigcup_{j \in U} B_j$ for $U \in \mathcal{P}$.
$\mathfrak{I}_{\mathcal{F}}(G)$	Set of $\mathcal{F}$-identifying coarsenings.
$\mathfrak{C}_{\mathcal{R}}(G)$	Set of coarsenings attainable under $\mathcal{R}$.
$\mathsf{A}_{\mathcal{F},\mathcal{R}}(G)$	Exact minimum replication cost under feasible CI family $\mathcal{F}$ and action set $\mathcal{R}$.
$\mathcal{P}(H_{\mathcal{R}})$	Connected-component partition of the action graph.
$\mathcal{R}_{\mathrm{complete}}$	Complete-action interface $(\mathcal{K} \times Q) \setminus E$.
$s = \lfloor n/4 \rfloor$	Number of queried four-variable blocks in the all-ambient separation.
G_n	Matched all-ambient-n source–query construction used for the linear separation.

DAG–Gaussian construction.

Symbol	Meaning
n	Number of ambient variables in a fixed-size DAG–Gaussian comparison.
Q_4	Complete elementary-CI query universe on four labeled variables.
$Q(U_i)$	The 24 elementary-CI propositions internal to quartet U_i.
M_{blk}	Four-query blocking set in Eq. (11).
M_i	Relabelled copy of M_{blk} on quartet U_i.
U_i	Queried four-variable set $\{A_i, B_i, C_i, D_i\}$ in block i.
$\mathcal{F}^{(n)}_{\mathrm{DAG}}(Q)$	Queried d-separation profiles induced by all DAGs on a fixed n-variable ambient set.
$\mathcal{F}_{\text{DAG-marg}}(Q_4)$	Quartet profiles induced by arbitrary finite ambient DAGs, equivalently $\bigcup_{m \geq 4} \mathcal{F}^{(m)}_{\mathrm{DAG}}(Q_4)$.
$\mathcal{F}^{(n)}_{\mathrm{G}}(Q)$	Queried CI profiles induced by all positive-definite regular Gaussians on the ambient variables.

Repeated-task stochastic extension.

Symbol	Meaning
r	Repeated-task index.
$Y^{(r)}$	Canonical CI truth profile for task r.
$Z^{(r)}_{kq}$	Observed response of source k to query q in repeated task r.
π	Mixture distribution over feasible truth profiles across repeated tasks.
π_y	Mixture mass assigned to truth profile y.
s_k	Sensitivity of source k.
c_k	Specificity of source k.
$A_k(s_k, c_k)$	2×2 stochastic answer-channel matrix of source k.
θ	Collection of stochastic source-channel parameters.
$p_\theta(z \mid y)$	Conditional probability of observed response pattern z given truth profile y.
$\Psi_{\mathcal{F}}(\theta, \pi)$	Population observation law induced by feasible CI family $\mathcal{F}$, channel parameters θ, and mixture π.

Operators and conventions.

Symbol	Meaning
$\oplus$	XOR, applied coordinatewise to binary vectors.
$\mathbf{1}_A$	Binary indicator vector of a query subset $A \subseteq Q$.
$\lvert\mathcal{P}\rvert$	Number of blocks in partition or coarsening $\mathcal{P}$.
$\bigsqcup$	Disjoint union.
$\perp\!\!\!\perp$	Conditional independence.
K_c	Complete graph on the c initial observation components.

A.2 Proof dependencies

App. B proves the observational alias proposition, and App. C proves the exact action-constrained replication theorem. App. D establishes the causal two-world rigidity result, gives the explicit Gaussian witness, and proves the complete-action linear separation. App. E gives the independent conditioning-layer mechanism, App. F proves the repeated-task stochastic proposition, and App. G gives the controlled-partition details and targeted validation results. The deterministic dependency order is

$$\begin{aligned} \text{alias geometry} &\longrightarrow \text{replication complexity}, \\ \text{causal rigidity} + \text{replication complexity} &\longrightarrow \text{linear evidence separation}. \end{aligned} \tag{14}$$

B Proof of Source–Truth Identification

Proof of Prop. 1. Suppose feasible worlds (Y, t) and (Y', t') induce the same observations. For every observed edge (k, q),

$$Y_q \oplus t_k = Y'_q \oplus t'_k, \tag{15}$$

so

$$Y_q \oplus Y'_q = t_k \oplus t'_k. \tag{16}$$

The query-coordinate discrepancy is denoted by δ_Q, and the source-coordinate discrepancy by $\delta_{\mathcal{K}}$. They satisfy

$$\delta_Q(q) = Y_q \oplus Y'_q, \qquad \delta_{\mathcal{K}}(k) = t_k \oplus t'_k. \tag{17}$$

The discrepancies agree across every observed edge. Connectivity therefore forces one common bit h_j on every vertex of component C_j, giving

$$Y' = Y \oplus h_Q, \qquad t' = t \oplus h_{\mathcal{K}}. \tag{18}$$

Feasibility of Y' gives the orbit in Eq. (4); the converse follows because simultaneous truth/source flips cancel on every observed edge. For the relative-polarity criterion, a nonempty proper component subset S changes at least one cross-component relative polarity and has truth discrepancy

$$\mathbf{1}_{\cup_{j \in S} B_j}. \tag{19}$$

If this mask belongs to $\Delta(\mathcal{F})$, choose $y, y' \in \mathcal{F}$ with that XOR difference and any polarity vector t. Flipping exactly the components in S converts (y, t) into a distinct feasible world with the same observations but a different relative orientation, so universal identification fails. Conversely, if universal relative polarity fails, the orbit characterization in Eq. (4) gives an observational alias generated by some component flip. Because the alias changes relative polarity, its selected component set is nonempty and proper, and the corresponding component-union mask is therefore in $\Delta(\mathcal{F})$. The zero and all-component flips preserve every relative polarity. This proves Prop. 1. □

C Proof of Exact Replication Complexity

This appendix proves Thm. 1 and the reachability/complete-action statement in Cor. 1, and then records two monotonicity consequences of the same partition formula.

C.1 Exact cost under fixed action constraints

Consider any successful replication set $E_{\mathrm{add}} \subseteq \mathcal{R}$. Its final component coarsening is denoted by $\mathcal{P}_{E_{\mathrm{add}}}$. By Prop. 1, success implies that $\mathcal{P}_{E_{\mathrm{add}}}$ is $\mathcal{F}$-identifying. Because every added cross-component edge belongs to $\mathcal{R}$, each final block is connected in $H_{\mathcal{R}}$, so the coarsening is $\mathcal{R}$-attainable. One added edge reduces the number of connected components by at most one; hence

$$|E_{\mathrm{add}}| \geq c - |\mathcal{P}_{E_{\mathrm{add}}}| \geq \min_{\mathcal{P} \in \mathfrak{I}_{\mathcal{F}}(G) \cap \mathfrak{C}_{\mathcal{R}}(G)} (c - |\mathcal{P}|). \tag{20}$$

Conversely, fix any $\mathcal{P} \in \mathfrak{I}_{\mathcal{F}}(G) \cap \mathfrak{C}_{\mathcal{R}}(G)$. For every block $U \in \mathcal{P}$, choose a spanning tree of $H_{\mathcal{R}}[U]$. By definition of the action graph, each tree edge has at least one allowed source–query action in $\mathcal{R}$ connecting the corresponding initial components. Add one such action per tree edge. Block U is merged with exactly $|U| - 1$ actions, so the total is

$$\sum_{U \in \mathcal{P}} (|U| - 1) = c - |\mathcal{P}|. \tag{21}$$

The final partition is exactly $\mathcal{P}$, which is identifying; Prop. 1 then gives universal relative-polarity identification. This proves Eq. (9), including $+\infty$ when the intersection of identifying and attainable coarsenings is empty.

C.2 Reachability criterion

The connected-component partition of the action graph is denoted by $\mathcal{P}(H_{\mathcal{R}})$. It is the coarsest partition reachable under the allowed action set. If an attainable identifying coarsening exists, every further coarsening remains identifying because a union of coarser blocks is also a union of the finer blocks; therefore $\mathcal{P}(H_{\mathcal{R}})$ is identifying. Conversely, if $\mathcal{P}(H_{\mathcal{R}})$ is identifying, it is attainable by a spanning forest of the action graph. This proves Cor. 1's finite-cost equivalence.

If $H_{\mathcal{R}} = K_c$, every coarsening is attainable and the exact formula reduces to Eq. (10). For $\mathcal{F}_{\mathrm{Bin}}(Q) = \{0,1\}^Q$ every nontrivial component-union mask is feasible, so the only identifying complete-action coarsening has one block and $\mathsf{A}_{\mathrm{Bin}}(G) = c - 1$.

C.3 Two monotonicities

If $\mathcal{F}_1 \subseteq \mathcal{F}_2$, then $\Delta(\mathcal{F}_1) \subseteq \Delta(\mathcal{F}_2)$. Every coarsening identifying for $\mathcal{F}_2$ is therefore identifying for $\mathcal{F}_1$, while attainability is unchanged. Taking minima in Eq. (9) gives structural monotonicity.

If $\mathcal{R}_1 \subseteq \mathcal{R}_2$, every $\mathcal{R}_1$-attainable coarsening is also $\mathcal{R}_2$-attainable, while the identifying family is unchanged. Taking minima gives action-set monotonicity. □

D Proofs of All-Ambient Causal Rigidity and Linear Separation

This appendix proves Thms. 2 and 3. The first part establishes the analytic two-world exclusion for a quartet embedded in an arbitrary ambient DAG. The second gives an explicit regular-Gaussian witness and amplifies that local difference to the complete-action linear replication gap.

D.1 Graphical preliminaries for an arbitrary ambient DAG

Fix a quartet $U = \{A, B, C, D\}$ inside an arbitrary DAG on a larger ambient variable set. Relative to CI queries whose endpoints and conditioning variables lie in U, the remaining ambient variables may be marginalized. The induced CI model on U has an ancestral-graph representation whose m-separation relations agree with the original DAG queries (Richardson & Spirtes, 2002). We use two standard properties of such graphical CI models: ordered upward/downward stability (Sadeghi, 2017; Sadeghi & Soo, 2022), and Pearl's chordality condition, which remains necessary for DAG-isomorph marginals (de Waal, 2009). Compositional-graphoid and singleton-transitivity rules are used only as established graphical CI inference rules (Sadeghi, 2017).

Lemma 1 (Three-variable DAG–Gaussian coincidence)**.** *On three variables, the complete elementary-CI profile families realized by faithful DAGs and by positive-definite regular Gaussians coincide. Consequently, their two-world difference sets also coincide.*

Proof. After standardizing variances, write the Gaussian correlation matrix as

$$\Sigma = \begin{pmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{pmatrix} \succ 0. \tag{22}$$

The six elementary CI statements satisfy

$$A \perp\!\!\!\perp B \iff a = 0, \qquad A \perp\!\!\!\perp B \mid C \iff a = bc, \tag{23}$$

with the two cyclic analogues. Positive definiteness gives $|a|, |b|, |c| < 1$. If a, b, c are all nonzero, no marginal CI holds and two distinct conditional CIs cannot hold simultaneously: for example, $a = bc$ and $b = ac$ imply $a(1 - c^2) = 0$, a contradiction. Hence the profile is empty or contains exactly one conditional CI. If exactly one of a, b, c is zero, exactly the corresponding marginal CI holds. If exactly two are zero, one variable is independent of the other two, giving the four elementary CIs involving that isolated variable. If all three vanish, all six CIs hold. These are exactly the faithful three-variable DAG patterns: a complete DAG, a chain or fork, a collider, an isolated vertex plus one edge, and the empty DAG, including all relabellings. Conversely, every DAG has a faithful regular-Gaussian realization (Spirtes et al., 2000), so every one of these DAG profiles is also attained by a positive-definite Gaussian. Thus the Gaussian and faithful-DAG profile families coincide, and so do their difference sets. □

Lemma 2 (Trivariate antipode)**.** *Let two three-variable gaussoids have complementary truth values on all six elementary CI propositions. Then one contains all six propositions and the other contains none.*

Proof. On three variables the gaussoid axioms have the standard local forms: the empty form, one marginal CI, one conditional CI, an isolated-variable form with four CIs, or the full six-CI form (Boege, 2022). Every nonempty proper form therefore has one or four statements, whose bitwise complement has five or two and is not a gaussoid. Only the empty and full forms can be truth-value complements. □

Lem. 1 proves the minimal-dimension remark stated after Thm. 2: a matched DAG–Gaussian separation cannot first occur on fewer than four queried variables.

Lemma 3 (Conditioning-square unateness)**.** *Let $\mathcal{I}$ be an independence model induced by a directed ancestral graph on distinct variables X, Y, a, b. Define*

$$f(S) = \mathbf{1}\{X \perp\!\!\!\perp Y \mid S\}, \qquad S \subseteq \{a, b\}. \tag{24}$$

Then the ordered four-tuple $(f(\varnothing), f(\{a\}), f(\{b\}), f(\{a, b\}))$ is neither 0110 nor 1001.

Proof. Use the minimal partial order $<$ of the directed ancestral graph: $u < v$ means that v is a strict ancestor of u, and distinct vertices unrelated by ancestry are incomparable. Ordered upward-stability states that $X \perp\!\!\!\perp Y \mid C$ implies $X \perp\!\!\!\perp Y \mid C \cup \{k\}$ whenever $k \notin C \cup \{X, Y\}$ and $X < k$ or $Y < k$. Thus an ancestor of either endpoint may be added to the conditioning set. Ordered downward-stability states that $X \perp\!\!\!\perp Y \mid C$ implies $X \perp\!\!\!\perp Y \mid C \setminus \{k\}$ whenever $k \in C$, $X \not< k$, $Y \not< k$, and $\ell \not< k$ for every $\ell \in C \setminus \{k\}$ (Sadeghi & Soo, 2022). The latter conditions mean that k is not an ancestor of either endpoint or any remaining conditioner. These order conventions apply throughout the following two cases.

For 0110, $X \perp\!\!\!\perp Y \mid b$ but $X \not\perp\!\!\!\perp Y \mid ab$ implies, by the contrapositive of upward-stability, $X \not< a$ and $Y \not< a$. Downward-stability applied to $X \perp\!\!\!\perp Y \mid a$ can then remove a, forcing $X \perp\!\!\!\perp Y$, a contradiction.

For 1001, failure of independence after adding a or b to the empty conditioning set gives $X \not< a, Y \not< a, X \not< b, Y \not< b$. Starting from $X \perp\!\!\!\perp Y \mid ab$, if $b \not< a$, downward-stability removes a and forces $X \perp\!\!\!\perp Y \mid b$, contradiction; hence $b < a$. Symmetrically $a < b$, which contradicts antisymmetry. □

D.2 Two-world synchronization and analytic local blocking

On the generic quartet define

$$M_{\mathrm{blk}} = \{A \perp\!\!\!\perp D \mid BC,\ B \perp\!\!\!\perp C,\ B \perp\!\!\!\perp C \mid A,\ B \perp\!\!\!\perp C \mid D\}. \tag{25}$$

Write the corresponding truth bits as

$$a = [A \perp\!\!\!\perp D \mid BC], \quad b = [B \perp\!\!\!\perp C], \quad c = [B \perp\!\!\!\perp C \mid A], \quad d = [B \perp\!\!\!\perp C \mid D]. \tag{26}$$

Lemma 4 (Two-world synchronization)**.** *Suppose two singleton-transitive compositional graphoids agree on the other 20 elementary Q4 CI propositions and differ on all four of a, b, c, d. Then in either world*

$$c = d = 1 - b. \tag{27}$$

Proof. Assume first that $b = c$. Since both bits flip, one world W has $B \perp\!\!\!\perp C$ and $B \perp\!\!\!\perp C \mid A$. Singleton transitivity gives $A \perp\!\!\!\perp B$ or $A \perp\!\!\!\perp C$; by symmetry take $A \perp\!\!\!\perp B$. Composition of $B \perp\!\!\!\perp A$ and $B \perp\!\!\!\perp C$, followed by weak union, gives the outer statement $A \perp\!\!\!\perp B \mid C$, so both $A \perp\!\!\!\perp B$ and $A \perp\!\!\!\perp B \mid C$ also hold in the other world W'. There $B \not\perp\!\!\!\perp C$, so singleton transitivity forces $A \perp\!\!\!\perp C$, another shared outer statement. Hence both worlds satisfy $A \perp\!\!\!\perp B$ and $A \perp\!\!\!\perp C$, and composition gives $A \perp\!\!\!\perp BC$.

Exactly one world has $A \perp\!\!\!\perp D \mid BC$. In that world contraction with $A \perp\!\!\!\perp BC$ yields $A \perp\!\!\!\perp BCD$, hence the shared outer statement $A \perp\!\!\!\perp D$. In the other world, $A \perp\!\!\!\perp BC$ and $A \perp\!\!\!\perp D$ imply $A \perp\!\!\!\perp BCD$ by composition and then $A \perp\!\!\!\perp D \mid BC$ by weak union, contradicting the flip of a. Thus $b \neq c$. Interchanging A and D gives $b \neq d$. Since the bits are binary, $c = d = 1 - b$. □

Recall that $\mathcal{F}_{\text{DAG-marg}}(Q_4) = \bigcup_{m \geq 4} \mathcal{F}^{(m)}_{\text{DAG}}(Q_4)$ is the family of complete Q_4 elementary-CI profiles obtained by restricting an arbitrary finite ambient DAG d-separation model to a four-variable quartet.

Proof of Thm. 2. For the first mask, suppose the two worlds agree on the outer 20 coordinates and flip exactly a, b, c, d. Lem. 4 forces one world to have $(b, c, d) = (0, 1, 1)$ and the other $(1, 0, 0)$. The outer bit $e = [B \perp\!\!\!\perp C \mid AD]$ is shared. If $e = 0$, the first world's conditioning square for endpoints B, C and conditioners A, D is 0110; if $e = 1$, the second world's square is 1001. Both contradict Lem. 3. Hence M_{blk} is impossible.

For the complementary mask, the worlds share a, b, c, d and flip the other 20 coordinates. Restrict to the six elementary CIs on A, B, D: all six flip, so Lem. 2 says that one world is full on this triple and the other empty. The same holds for A, C, D. The coordinate $A \perp\!\!\!\perp D$ occurs in both triple restrictions. Within either world it has one common truth value, so the two restrictions cannot have opposite full-versus-empty assignments. Consequently, one world is full on both triples and the other is empty on both; call them W_+ and W_-, respectively. Here the coordinate is common to the two restrictions, while its truth value flips between the two worlds. In W_+, the three marginal independences $A \perp\!\!\!\perp B, A \perp\!\!\!\perp C, A \perp\!\!\!\perp D$ imply $A \perp\!\!\!\perp BCD$ by composition and hence $a = 1$ by weak union. Also $A \perp\!\!\!\perp B$ gives $b \Leftrightarrow c$ by composition/weak union in one direction and contraction/decomposition in the other; $B \perp\!\!\!\perp D$ similarly gives $b \Leftrightarrow d$. Thus $b = c = d = s_0$. If $s_0 = 1$, then in W_- both $B \perp\!\!\!\perp C$ and $B \perp\!\!\!\perp C \mid A$ hold while $A \not\perp\!\!\!\perp B$ and $A \not\perp\!\!\!\perp C$, violating singleton transitivity. Therefore the shared core is $(a, b, c, d) = (1, 0, 0, 0)$.

Finally set

$$e = [B \perp\!\!\!\perp C \mid AD], \quad p = [A \perp\!\!\!\perp D \mid B], \quad q = [A \perp\!\!\!\perp D \mid C]. \tag{28}$$

These outer bits flip. The full triple restrictions give $p = q = 1$ in W_+. Moreover $B \perp\!\!\!\perp A$ and $B \perp\!\!\!\perp D$ imply $B \perp\!\!\!\perp AD$. If $e = 1$ in W_+, contraction with $B \perp\!\!\!\perp C \mid AD$ would imply $B \perp\!\!\!\perp C$, contradicting $b = 0$. Hence $e = 0$ in W_+, so in W_- we have $a = e = 1$ and $p = q = 0$. Pearl's chordality condition (Pearl, 1988) for a DAG-isomorph marginal (de Waal, 2009), applied as

$$A \perp\!\!\!\perp D \mid BC, \quad B \perp\!\!\!\perp C \mid AD \quad \Longrightarrow \quad A \perp\!\!\!\perp D \mid B \text{ or } A \perp\!\!\!\perp D \mid C, \tag{29}$$

contradicts $p = q = 0$. The complementary mask is therefore impossible as well. □

D.3 Explicit regular-Gaussian local witness

The two rational covariance matrices

$$\Sigma_0 = \begin{pmatrix} 1 & 1/5 & 3/10 & 2/5 \\ 1/5 & 1 & 0 & 1/4 \\ 3/10 & 0 & 1 & 7/20 \\ 2/5 & 1/4 & 7/20 & 1 \end{pmatrix}, \qquad \Sigma_1 = \begin{pmatrix} 1 & 1/2 & 1/2 & 2/5 \\ 1/2 & 1 & 1/4 & 1/2 \\ 1/2 & 1/4 & 1 & 1/2 \\ 2/5 & 1/2 & 1/2 & 1 \end{pmatrix} \tag{30}$$

have leading principal minors

$$1,\ 24/25,\ 87/100,\ 25961/40000 \quad \text{and} \quad 1,\ 3/4,\ 9/16,\ 27/80, \tag{31}$$

respectively, so both are positive definite. By the standard almost-principal-minor characterization of Gaussian CI (Boege et al., 2019), direct exact determinant evaluation gives

$$\mathcal{I}(\Sigma_0) = \{B \perp\!\!\!\perp C\}, \tag{32}$$

and

$$\mathcal{I}(\Sigma_1) = \{A \perp\!\!\!\perp D \mid BC,\ B \perp\!\!\!\perp C \mid A,\ B \perp\!\!\!\perp C \mid D\}. \tag{33}$$

All remaining elementary Q4 almost-principal minors are nonzero. Hence the two profiles differ exactly by M_{blk}.

D.4 All-ambient construction and complete-action lower bound

Fix $n \geq 4$, let $s = \lfloor n/4 \rfloor$, and use the query/source construction from Sec. 4.4. Consider any nonconstant component-flip vector. Some satellite differs from the central component, so on that quartet the two candidate ambient-DAG profiles differ by either M_{blk} or $Q_4 \setminus M_{\mathrm{blk}}$. Thm. 2 excludes both without assumptions on cross-quartet or leftover-variable edges. Therefore every surviving component flip is constant and preserves relative source polarity. The zero flip always retains the original feasible world. The all-component flip is feasible only when the complemented truth profile belongs to the feasible family; its possible presence does not affect relative polarity. Hence

$$\mathsf{A}_{\mathrm{DAG},\mathcal{R}_{\mathrm{complete}}}(G_n) = 0. \tag{34}$$

For the Gaussian side, fix the central component flip to zero. For every satellite subset $S \subseteq [s]$, use (Σ_0, Σ_1) on quartet i when $i \in S$ and identical local covariance blocks when $i \notin S$. The block-diagonal product across quartets, with independent leftover variables, is positive definite and realizes exactly the corresponding union of local masks. Thus all 2^s relative satellite-flip patterns are feasible.

After fewer than s added replication edges, a graph on the $s+1$ initial source components is disconnected. Choose a connected component not containing the central source and flip exactly its satellites. The Gaussian block-product construction realizes that nontrivial alias while every added edge lies within a constant-flip component, so relative polarity is not identified. Hence at least s replications are necessary. Since $\mathcal{R}_{\mathrm{complete}}$ allows a spanning tree on the $s+1$ components, s replications suffice. Therefore

$$\mathsf{A}_{\mathrm{G},\mathcal{R}_{\mathrm{complete}}}(G_n) = s = \lfloor n/4 \rfloor. \tag{35}$$

Together with the DAG side, this proves Thm. 3. □

E Alternative Structural Mechanism

The main all-ambient result uses a local causal two-world exclusion. A distinct zero-replication mechanism follows from standard compositional-graphoid closure, showing that the general alias framework can exploit structural restrictions beyond the four-variable blocking mask. A compositional graphoid satisfies the semigraphoid axioms together with intersection and composition (Amini et al., 2026; Sadeghi, 2017).

Proposition 2 (Zero replication from conditioning-cardinality layers)**.** *Consider $n \geq 3$ variables $X_1, \ldots, X_n$. For each conditioning cardinality $r \in \{0, \ldots, n-2\}$, let*

$$L_r = \{X_i \perp\!\!\!\perp X_j \mid S :\ |S| = r\} \tag{36}$$

be the set of elementary CI queries whose conditioning set has size r. If the feasible family consists of elementary compositional-graphoid CI profiles and each L_r is one connected component of the source–query graph, then every nonempty proper component-union mask lies outside $\Delta(\mathcal{F})$. Hence $\mathsf{A}_{\mathcal{F}}(G) = 0$ under complete actions.

Proof. Take any two feasible profiles $\mathcal{I}$ and $\mathcal{J}$ whose difference is a union of conditioning-cardinality layers, and let $f : \{0, \ldots, n-2\} \to \{0, 1\}$ indicate which layers are flipped. Fix $r \leq n-3$, a conditioning set K with $|K| = r$, and distinct $i, j, k \notin K$. Writing $Kk = K \cup \{k\}$ and $Kj = K \cup \{j\}$, symmetry together with composition, weak union, intersection, and decomposition gives

$$[(i \perp\!\!\!\perp j \mid K) \wedge (i \perp\!\!\!\perp k \mid K)] \iff [(i \perp\!\!\!\perp j \mid Kk) \wedge (i \perp\!\!\!\perp k \mid Kj)]. \tag{37}$$

Write the four truth bits in this order as x, y, u, v, so $x \wedge y = u \wedge v$. If $f(r) = 0$ and $f(r+1) = 1$, applying the same equivalence to $\mathcal{J}$ gives $x \wedge y = \neg u \wedge \neg v$, and hence $u \wedge v = \neg u \wedge \neg v$, so $u \neq v$. Repeating the argument with the other pivots introduces a third upper-layer bit that must differ from both u and v, impossible for binary values. The case $f(r) = 1$ and $f(r+1) = 0$ is symmetric on the lower-layer bits. Therefore $f(r) = f(r+1)$ for every adjacent pair, so f is constant. A nonempty proper union of layer components would require a nonconstant f, which is impossible. Prop. 1 therefore identifies relative polarity without additional replication, giving $\mathsf{A}_{\mathcal{F}}(G) = 0$. □

F Stochastic Repeated-Task Extension

This appendix gives the formal repeated-task model, proposition, and complete proof of Prop. 3. The result is a supporting repeated-task extension of the deterministic structural core.

F.1 Formal repeated-task model and identification proposition

The deterministic theory models each source with a fixed polarity. We next replace that reporting mechanism by asymmetric stochastic errors shared across repeated tasks and ask whether CI structure still changes identifiability. Each task has the same query format, while the sources reuse the same channel parameters across tasks. This isolates whether the structural advantage persists when source errors are probabilistic rather than deterministic.

Repeated tasks are indexed by r. Let Q_3 denote the six elementary CI propositions on the three task variables. The canonical six-query truth profile of task r is denoted by $Y^{(r)} \in \mathcal{F} \subseteq \{0,1\}^{Q_3}$. These profiles are independent and identically distributed according to an unknown mixture π on $\mathcal{F}$, whose mass on profile y is denoted by π_y. In the causal case, every task satisfies Assump. 1, so $Y^{(r)} \in \mathcal{F}^{(3)}_{\mathrm{DAG}}(Q_3)$; the unrestricted comparator uses $\mathcal{F}_{\mathrm{Bin}}(Q_3) = \{0,1\}^{Q_3}$. Source 0 answers the three marginal queries and source 1 the corresponding three one-variable-conditioned queries in every task. The observed response of source k to query q in task r is denoted by $Z^{(r)}_{kq}$.

The task-invariant sensitivity and specificity of source k are denoted by s_k and c_k, respectively:

$$s_k = P(Z^{(r)}_{kq} = 1 \mid Y^{(r)}_q = 1), \qquad c_k = P(Z^{(r)}_{kq} = 0 \mid Y^{(r)}_q = 0). \tag{38}$$

The corresponding binary answer-channel matrix is denoted by $A_k(s_k, c_k)$ and equals

$$A_k(s_k, c_k) = \begin{pmatrix} c_k & 1 - s_k \\ 1 - c_k & s_k \end{pmatrix}. \tag{39}$$

Conditional on $Y^{(r)}$ and the shared channel parameters, source responses are independent across queried coordinates, and tasks are independent across r. This gives the stochastic extension its standing repeated-task conditions before the population map is introduced.

The four source-channel parameters are collected into the vector $\theta = (s_0, c_0, s_1, c_1)$. The interior parameter domain used below has $0 < s_k, c_k < 1$ for both sources and strictly positive mixture weights $\pi_y > 0$ for every $y \in \mathcal{F}$, subject to the simplex constraint. For a fixed truth profile y and observed response pattern $z \in \{0,1\}^6$, the conditional response probability under channel parameters θ is denoted by $p_\theta(z \mid y)$. The resulting one-task population observation law is denoted by $\Psi_{\mathcal{F}}(\theta, \pi)$, with coordinates

$$\Psi_{\mathcal{F}}(\theta, \pi)_z = \sum_{y \in \mathcal{F}} \pi_y p_\theta(z \mid y). \tag{40}$$

Local identifiability refers to the parameter pair (θ, π) from this repeated-task population law; each draw represents a distinct causal subproblem with the same query format and shared source parameters.

Proposition 3 (Matched stochastic identification gap across repeated tasks)**.** *Under the fixed two-source six-query repeated-task interface:*

1. *For $\mathcal{F}_{\mathrm{Bin}}(Q_3) = \{0,1\}^{Q_3}$, every interior parameter point satisfying $s_k + c_k \neq 1$ for each $k \in \{0,1\}$ is jointly locally non-identifiable: nearby source-channel changes can be exactly compensated by changes in the unrestricted truth-profile mixture.*

2. *For the complete 11-profile three-node faithful-DAG CI family, the 14-dimensional free parameterization is locally identifiable on the intersection of the interior parameter domain with a nonempty Zariski-open set. At the interior reference point with a uniform mixture over the 11 feasible profiles,*

$$\theta^\star = \left(\frac{9}{10}, \frac{4}{5}, \frac{7}{10}, \frac{3}{5}\right), \qquad \pi_y^\star = \frac{1}{11}, \tag{41}$$

an exact rational 14×14 *Jacobian minor is nonzero.*

3. *If all truth bits observed by one source are almost surely* 0 *under* π*, that source's sensitivity is not identifiable; if they are almost surely* 1*, its specificity is not identifiable.*

Proof sketch. For the unrestricted binary family, the one-task observation map factors as an invertible Kronecker product of source-channel matrices applied to the latent truth-profile mixture. A sufficiently small perturbation of the channel parameters can therefore be canceled exactly by an inverse transformation of the mixture, which remains in the interior and creates a local continuum of observational aliases. For the 11-profile faithful-DAG CI family, the feasible mixture support is restricted; at the stated interior point an exact 14×14 Jacobian minor is nonzero, so the inverse function theorem gives local injectivity on a nonempty generic interior set. Finally, if a source sees only truth value 0 or only truth value 1, one of its two channel parameters drops out of the likelihood, giving the excitation boundary in item 3. The complete proof follows below.

The comparison keeps the repeated-task interface and asymmetric channel parameterization fixed. With an unrestricted truth family, nearby channel changes can be absorbed by changes in the latent truth-profile mixture. The 11-profile faithful-DAG CI family leaves fewer mixture degrees of freedom, and the exact Jacobian calculation gives full local rank on a nonempty generic interior set. Thus feasible CI structure changes local population identifiability even when source errors are stochastic. The complete reparameterization proof, exact rational Jacobian calculation, and excitation boundary follow below.

F.2 Unrestricted-family local non-identifiability

For one query answered by source k, the corresponding channel matrix is denoted by A_k and equals

$$A_k = \begin{pmatrix} c_k & 1 - s_k \\ 1 - c_k & s_k \end{pmatrix}. \tag{42}$$

Its determinant is $s_k + c_k - 1$. Under the stated assumptions both source channels are invertible.

For the unrestricted family, the one-task population operator is denoted by T_θ and, up to a fixed coordinate ordering, equals

$$T_\theta = A_0 \otimes A_1 \otimes A_0 \otimes A_1 \otimes A_0 \otimes A_1. \tag{43}$$

The resulting unrestricted-family observable law is denoted by $\Psi_{\mathrm{Bin}}(\theta, \pi)$, so

$$\Psi_{\mathrm{Bin}}(\theta, \pi) = T_\theta \pi. \tag{44}$$

Choose any distinct θ' sufficiently close to θ for which the source channels remain invertible. The compensating mixture associated with this perturbation is

$$\pi' = T_{\theta'}^{-1} T_\theta \pi. \tag{45}$$

The observable law is unchanged. Because each T_θ is column stochastic, $\mathbf{1}^\top T_\theta = \mathbf{1}^\top$ and therefore $\mathbf{1}^\top T_{\theta'}^{-1} = \mathbf{1}^\top$, so π' has total mass one. At an interior mixture $\pi > 0$, continuity keeps every coordinate of π' strictly positive for all sufficiently small channel perturbations. Thus the unrestricted model has a local continuum of distinct exact channel/mixture aliases.

F.3 Faithful-DAG CI local identifiability

For the 11-profile faithful-DAG CI family, the 14-dimensional free parameter vector is denoted by x:

$$x = (s_0, c_0, s_1, c_1, \pi_1, \ldots, \pi_{10}), \qquad \pi_{11} = 1 - \sum_{j=1}^{10} \pi_j. \tag{46}$$

Name the three task variables A, B, C and order the six truth and output coordinates as $A \perp\!\!\!\perp B$, $A \perp\!\!\!\perp B \mid C$, $A \perp\!\!\!\perp C$, $A \perp\!\!\!\perp C \mid B$, $B \perp\!\!\!\perp C$, and $B \perp\!\!\!\perp C \mid A$. A truth-profile bit equals one exactly when the corresponding CI statement holds. Encode each six-bit pattern as an integer from 0 to 63, with the first listed coordinate as the most significant bit.

The mixture weight π_j corresponds to profile $y^{(j)}$ in the following fixed order. The last column lists all elementary CI statements that hold in that profile; every unlisted statement is false.

j	Integer code	$y^{(j)}$	CI statements that hold
1	0	`000000`	None
2	1	`000001`	$B \perp\!\!\!\perp C \mid A$
3	2	`000010`	$B \perp\!\!\!\perp C$
4	4	`000100`	$A \perp\!\!\!\perp C \mid B$
5	8	`001000`	$A \perp\!\!\!\perp C$
6	15	`001111`	$A \perp\!\!\!\perp C,\ A \perp\!\!\!\perp C \mid B,\ B \perp\!\!\!\perp C,\ B \perp\!\!\!\perp C \mid A$
7	16	`010000`	$A \perp\!\!\!\perp B \mid C$
8	32	`100000`	$A \perp\!\!\!\perp B$
9	51	`110011`	$A \perp\!\!\!\perp B,\ A \perp\!\!\!\perp B \mid C,\ B \perp\!\!\!\perp C,\ B \perp\!\!\!\perp C \mid A$
10	60	`111100`	$A \perp\!\!\!\perp B,\ A \perp\!\!\!\perp B \mid C,\ A \perp\!\!\!\perp C,\ A \perp\!\!\!\perp C \mid B$
11	63	`111111`	All six statements

Thus π_{11}, the mass of the all-independent profile 111111, is the eliminated simplex coordinate. Output patterns z use the same six-bit encoding, and Jacobian columns follow the displayed order of x. At the interior reference point $\theta^\star = (9/10, 4/5, 7/10, 3/5)$ and $\pi_j^\star = 1/11$ for every j, the 14×14 Jacobian minor using output-pattern rows $[0, 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 16, 17]$, in that order, has determinant

$$-\frac{8003806668923692288815 09}{5201528097131813410669565200805664062500000000} \neq 0. \tag{47}$$

Thus the Jacobian has full column rank. Restricting the observable map to the 14 output coordinates used by this minor gives a square nonsingular Jacobian, so the inverse function theorem yields local injectivity of that submap and hence of the full observation map. Since the selected determinant is a nonzero polynomial in the free parameters, the same conclusion holds on the intersection of the interior parameter domain with a nonempty Zariski-open set.

This conclusion concerns local identification of the parameter pair (θ, π) from the population observation law. At each point in the stated set, a sufficiently small parameter neighborhood contains no distinct pair producing the same law. The argument does not establish global injectivity: distinct parameter points outside that neighborhood are not ruled out as observational aliases. It also provides no finite-sample estimation rate or optimization-convergence guarantee. The result uses the repeated-task assumptions stated above, including task-invariant source channels and conditional independence of responses; it does not establish identification under arbitrary query-dependent or task-dependent reporting biases.

F.4 Excitation boundary

If every truth coordinate seen by one source is almost surely 0, its likelihood does not depend on sensitivity; if every such coordinate is almost surely 1, it does not depend on specificity. This proves the excitation boundary and completes Prop. 3. □

G Experimental Details and Targeted Validation

G.1 Complete ambient-DAG quartet family

The ambient-DAG comparison uses $\mathcal{F}_{\text{DAG-marg}}(Q_4)$, the complete quartet-level CI family induced by arbitrary finite ambient DAGs. Every such latent projection is represented by a four-vertex acyclic directed mixed graph; conversely, each enumerated mixed graph is realized by retaining its acyclic directed part and introducing one fresh latent root for each bidirected edge. Enumerating the 543 acyclic directed parts and the $2^6 = 64$ bidirected-edge masks yields 34,752 mixed graphs

and 248 distinct complete Q_4 CI profiles. Exact profile comparison verifies that the 185 four-node DAG profiles form a strict subset of this ambient family, which is itself contained in the 629-profile regular-Gaussian family.

G.2 Controlled-partition comparison

The headline experiment compares the 248 ambient-DAG quartet profiles, the 629 regular-Gaussian profiles reconstructed exhaustively from the four-variable characterization of Lněnička & Matúš (2007), and the unrestricted binary family. Applying that characterization exhaustively to the full 2^{24} binary-profile universe in our reconstruction recovers 679 labeled gaussoids and 629 regular-Gaussian-representable profiles. For each $c \in \{2, 3, 4, 6, 8\}$, seeds $42, \ldots, 51$ each contribute 500 unique nonempty query partitions, for 5,000 designs per component count and 25,000 designs in total. Partitions are generated by uniform labeled surjections onto c nonempty blocks, canonicalized to unlabeled set partitions, and deduplicated within each component count. The same controlled designs are evaluated for the ambient-DAG, regular-Gaussian, unrestricted binary, and four-node DAG families under complete replication actions.

Table 1 reports the complete summary statistics for the controlled comparison.

Table 1: Controlled-partition replication costs for the four-node DAG, ambient-DAG quartet, regular-Gaussian, and unrestricted binary families. For each c, means and probability columns are empirical averages over the same 5,000 sampled partitions; each partition's replication cost is computed exactly.

c	mean $\mathsf{A}_{\mathrm{DAG4}}$	mean $\mathsf{A}_{\mathrm{Amb}}$	mean A_{G}	$\mathsf{A}_{\mathrm{Bin}}$	$\Pr(\mathsf{A}_{\mathrm{Amb}} = 0)$	$\Pr(\mathsf{A}_{\mathrm{Amb}} < \mathsf{A}_{\mathrm{G}})$
2	0.0022	0.0036	0.0140	1	0.9964	0.0104
3	0.0602	0.1174	0.3868	2	0.8826	0.2694
4	0.3554	0.5726	1.0178	3	0.4326	0.4240
6	1.3148	1.7326	2.5114	5	0.0018	0.6648
8	2.5576	3.1062	3.9580	7	0.0000	0.7808

Across all 25,000 designs, $\mathsf{A}_{\mathrm{DAG4}} \leq \mathsf{A}_{\mathrm{Amb}} \leq \mathsf{A}_{\mathrm{G}} \leq \mathsf{A}_{\mathrm{Bin}}$. Expanding from the four-node DAG family to the complete ambient-DAG quartet family increases the mean replication cost by 0.0014, 0.0572, 0.2172, 0.4178, and 0.5486 for $c = 2, 3, 4, 6, 8$, respectively. Thus arbitrary ambient structure weakens part of the four-node advantage without eliminating the systematic gap to the regular-Gaussian family.

G.3 Targeted validation details

For Thm. 2, all 29,281 labeled five-variable DAGs are enumerated. Each variable is treated in turn as an unqueried ambient variable, and the remaining quartet is restricted to its 24 elementary CI propositions. For all five quartet choices, neither M_{blk} nor $Q_4 \setminus M_{\mathrm{blk}}$ belongs to the exact restricted two-world difference set, providing a finite check of the theorem's ambient-variable scope.

For the Gaussian side of Thm. 3, exact rational arithmetic verifies that the two covariance matrices in App. D are positive definite and that their 24-query profiles differ on exactly M_{blk}. For Prop. 3, we additionally simulate finite samples under the same two-source, six-query repeated-task model. At the reference point in Prop. 3, constrained maximum-likelihood fits are evaluated over 10 independent repetitions for each task count. Table 2 reports the resulting source-channel recovery error.

Table 2: Finite-sample recovery under the repeated-task model of Prop. 3. RMSE is computed over the four source-channel parameters; SD denotes the standard deviation over 10 independent repetitions.

Repeated tasks R	Mean RMSE	SD
100	0.0993	0.0495
300	0.0857	0.0431
1,000	0.0454	0.0251
3,000	0.0368	0.0158
10,000	0.0276	0.0252
30,000	0.0132	0.0058
100,000	0.00757	0.00397

The mean source-channel error decreases over the tested task-count range. As a robustness check away from the reference point, across 10 additional interior generating settings the mean RMSE decreases from 0.0234 at $R = 3{,}000$ to 0.00727 at $R = 30{,}000$. These simulations provide a finite-sample numerical counterpart to the population-level local-identification result in Prop. 3; they do not establish global injectivity or a finite-sample convergence rate.

G.4 Relation to statistical CI aggregation

Bootstrap resampling can assess the sampling stability of a local CI decision, while majority voting combines multiple answers to the same query under assumptions connecting source errors and reporting conventions to the canonical truth. These procedures address the statistical construction or aggregation of CI answers. The deterministic audit begins with a fixed vector of binary reports and asks whether the feasible CI family makes relative source polarity identifiable from those reports.

The distinction is operational. If two feasible worlds produce the same report vector but different relative polarities, any aggregation rule using only that vector receives identical inputs in the two worlds. Even a randomized rule has the same output distribution and cannot identify the relative polarity correctly in both worlds with certainty. Resolving such an alias requires a structural restriction that excludes one world or information beyond the unchanged reports. Bootstrap analysis of underlying observations accesses a richer information interface, whose statistical guarantees depend on the sampling and reporting model.

Thus statistical aggregation and structural auditing can play complementary roles, but aggregation alone does not justify treating finite-sample answers as exact inputs to the deterministic theory. That theory requires compatibility with a shared feasible truth profile and persistent source polarities. App. F studies a specified repeated-task alternative with reusable asymmetric source channels and establishes a population-level local-identification result.

G.5 Application to benchmark networks

A benchmark application on a network such as ALARM, or on a specified reference DAG for a Sachs-based study, would first fix the canonical CI queries, the source–query reporting graph, and the allowed replication actions. Different source variable sets can restrict query coverage, while every reported query must retain the same canonical truth across sources. For an oracle evaluation, the reference DAG would generate CI labels by d-separation in the full network. Restricting attention to a quartet means retaining queries whose endpoints and conditioning variables lie in that quartet, rather than computing separation in the induced four-node subgraph. Paths through unqueried variables can still affect these answers.

The reference DAG should serve as evaluation ground truth, while the auditor receives the reports and a declared feasible family that does not reveal the reference graph. Otherwise, supplying the exact graph would already supply the canonical CI answers and remove the reference-free problem. To isolate the structural effect, feasible-family comparisons should retain the same queries, reporting graph, reporting mechanism, and allowed actions.

Thm. 2 provides one way to use local structure within such a larger network. A candidate component flip is excluded whenever its restriction to a fully queried quartet equals M_{blk} or its quartet complement, regardless of paths through external variables. Merely finding four variables in the network is insufficient: the reporting pattern must induce one of these exact discrepancy masks. Eliminating all relative-polarity aliases requires that every candidate alias be excluded, potentially using different quartets. The exact cost characterization in Thm. 1 additionally requires determining which aliases remain and which component mergers are attainable.

Such an evaluation would measure relative-polarity identification and additional source–query answers under the chosen interface. A finite-sample version would separately assess how estimated CI answers and reporting errors affect these outcomes. Benchmark-scale runtime and finite-sample performance are not evaluated by the controlled-partition and exact checks reported here.