I-FLOP: Fast Learning of Order and Parents from Interventional Data

Liuting Chen[†] JNK758@ALUMNI.KU.DK

Alex Markham[†] ALEX.MARKHAM@CAUSAL.DEV

[†] *Department of Mathematical Sciences*
University of Copenhagen
Denmark

Editors: Gustau Camps-Valls, Manuele Leonelli and Gherardo Varando

Abstract

We extend the FLOP (fast learning of order and parents) algorithm recently developed by Wienöbst et al. (2026) from observational to interventional data. In particular, we use the interventional BIC score of Hauser and Bühlmann (2012), adapting it to be used with the iterative Cholesky-based score updates that are partly responsible for FLOP's speed. We show that, in the sample limit, I-FLOP recovers a DAG in the same interventional Markov equivalence class as the data-generating DAG. We compare I-FLOP to existing causal structure learning algorithms on simulated and real interventional data, where it performs favorably in terms of both performance and run time.

Keywords: Causal Structure Learning, Interventional Data, Discrete Search

1. Introduction

With observational data, we can typically only determine a Markov equivalence class (MEC) (Verma and Pearl, 1990)—a set of causal graphs that cannot be distinguished from such data alone. Interventions with known targets break some of these ambiguities, and with multiple interventional environments we obtain a finer equivalence class, the interventional Markov equivalence class (I-MEC) (Hauser and Bühlmann, 2012), which determines the direction of more edges than the MEC alone.

To learn causal structures from interventional data, researchers have proposed various algorithms. These algorithms can be broadly categorized into three types based on how they utilize intervention information. The first type is score-based methods, such as Greedy Interventional Equivalence Search (GIES) (Hauser and Bühlmann, 2012): by defining a scoring function that incorporates intervention information, candidate graph structures are scored, and the optimal solution is searched within the graph space. The second type is permutation or ordering-based methods, such as Interventional Greedy Sparsest Permutation (I-GSP) (Wang et al., 2017): using intervention invariance constraints to guide the inference of node order, followed by selecting parent sets under the order constraints. The third type is invariance-based methods, such as Invariant Causal Prediction (ICP) (Peters et al., 2016), which leverage the principle of mechanism invariance where true causal mechanisms remain stable across multiple interventional environments while non-causal associations change. By identifying this invariance, causal relationships are pinpointed.

© 2026 L. Chen & A. Markham.

Among these approaches, ICP primarily tests invariance of predictive relationships for a specified response variable but is not designed as an interventional completed partially directed acyclic graph (I-CPDAG) search procedure. We therefore focus on GIES and I-GSP, which are more directly aligned with our goal of learning a global graph or an equivalence-class representative under known intervention targets. These two methods serve as the main reference points for our setting, but they face different challenges in practical use.

GIES uses a local greedy strategy to search the space of I-MEC, gradually adjusting the graph structure through three phases: forward, backward, and turning, which respectively consider score-improving local modifications of the current equivalence-class representative. At each step within a phase, GIES selects the admissible local move that yields the largest improvement in the score. While this greedy approach is computationally efficient, it does not guarantee convergence to the globally optimal structure (Wang et al., 2017). More critically, the candidate search at each step may have exponential complexity (Hauser and Bühlmann, 2012) in the worst case, making it difficult to fully explore the search space in large-scale problems or scenarios with limited computational resources.

I-GSP uses conditional independence (CI) tests under observational and interventional distributions to guide its permutation search. Its consistency result is asymptotic, whereas in finite samples the CI decisions can be sensitive to sample size, test choice, and significance threshold.

Motivated by these challenges, we adopt an order-based score-and-search approach for known-target interventional data. Searching over node orderings provides an attractive alternative to direct structure search, as acyclicity is guaranteed by construction and large neighborhoods can be explored efficiently (Bouckaert, 1992; Singh and Valtorta, 1993; Larranaga et al., 1996; Friedman and Koller, 2003; Teyssier and Koller, 2005; Scanagatta et al., 2017). Building on this paradigm, FLOP (Wienöbst et al., 2026) combines decomposable scoring with efficient reinsertion-based search and local parent-set updates.

We extend FLOP to known-target interventional data, yielding interventional FLOP (I-FLOP)[1]. Our main contributions are threefold:

1) We introduce a target-filtered pooled local score that extends order-based search to multi-environment interventional data with known targets.

2) We establish both local and global theoretical consistency guarantees: we prove that the target-filtered grow–shrink procedure consistently recovers the optimal prefix-restricted parent sets from any initialization, and that exact global minimization of our target-filtered pooled score asymptotically identifies the true I-MEC.

3) We conduct synthetic and real-world experiments, showing that I-FLOP achieves competitive scalability and superior I-CPDAG accuracy compared with established interventional baselines, including GIES and I-GSP.

1. Code is available at https://github.com/L2-T2/I-FLOP

2. Preliminaries

Let $G = (V, E_G)$ be a directed acyclic graph (DAG) with node set $V = \{1, \dots, p\}$ and edge set $E_G \subsetneq \{i \to j \mid i, j \in V, i \neq j\}$. We write $\mathrm{Pa}_G(v)$ for the parent set of node v. A linear Gaussian structural equation model (SEM) (Peters and Bühlmann, 2014) associated with G is given by *structural equations* $X_v = \sum_{u \in \mathrm{Pa}_G(v)} W_{uv} X_u + \varepsilon_v$ where the noise variables ε_v are jointly independent Gaussian random variables with $\varepsilon_v \sim \mathcal{N}(0, \sigma_v^2)$ and $\sigma_v^2 > 0$.

We consider multiple environments $\mathcal{E} = \{0, 1, \dots, m\}$ associated with a known intervention target family $\mathcal{I} = \{I_e : e \in \mathcal{E}\}$, where $I_0 = \emptyset$ denotes the observational environment. For a target set $I_e \subseteq V$, a known-target hard intervention replaces the structural equations of the targeted variables so that they are instead drawn jointly from an intervention distribution $Q_{I_e}^{(e)}$, i.e., $X_{I_e}^{(e)} \sim Q_{I_e}^{(e)}$, and leaves the structural equations of all non-targeted variables $v \in V \setminus I_e$ unchanged.

Every DAG can be associated to at least one topological order τ such that $u \to v \in E_G$ implies that u precedes v in τ. Given τ, the candidate parent set P_v of node v consists only of nodes that precede v in the order, which can be represented as $P_v \subseteq \mathrm{Pred}_\tau(v) = \{u \in V : u \text{ precedes } v \text{ in } \tau\}$. The ordered Markov property states that $X_v \perp\!\!\!\perp X_{\mathrm{Pred}_\tau(v) \setminus \mathrm{Pa}_G(v)} \mid X_{\mathrm{Pa}_G(v)}$ for every $v \in V$.

Since the linear Gaussian SEM above has strictly positive noise variances, the induced joint Gaussian distributions are non-degenerate and hence positive. Therefore, the ordered, local, and global Markov properties are equivalent in our setting (Lauritzen, 2026). A joint distribution $\mathbb{P}$ over X_V is Markov with respect to G if it satisfies these equivalent Markov properties.

Observational conditional independence relations do not generally identify a unique DAG: different DAGs may encode the same set of conditional independence relations and are therefore observationally indistinguishable. Such DAGs form a Markov equivalence class (MEC), represented by a completed partially directed acyclic graph (CPDAG) (Andersson et al., 1997). Following Hauser and Bühlmann (2012), two DAGs G and H are interventionally Markov equivalent with respect to the target family $\mathcal{I}$ if they induce the same interventional Markov model $M_{\mathcal{I}}(G) = M_{\mathcal{I}}(H)$. The corresponding equivalence class is the I-MEC. For the true DAG $G^\star$, we denote the true I-MEC by $\mathcal{G}_{\mathcal{I}}^\star = \{G \in \mathcal{G}_{\mathrm{DAG}} : M_{\mathcal{I}}(G) = M_{\mathcal{I}}(G^\star)\}$, where $\mathcal{G}_{\mathrm{DAG}}$ denotes the set of all DAGs on V. The unique graphical representative of an I-MEC is the interventional essential graph, also called the I-CPDAG.

Let $\mathcal{D} = \{X^{(e)}\}_{e \in \mathcal{E}}$ denote the multi-environment dataset. We propose I-FLOP as an extension of FLOP that retains FLOP's search framework for optimizing a decomposable score $S(G; \mathcal{D}) = \sum_{v \in V} s_v(\mathrm{Pa}_G(v); \mathcal{D})$ over the space of topological orders (Algorithm 1). Given a topological order τ, each node v selects its parent set $\mathrm{Pa}_G(v)$ only from its prefix $\mathrm{Pred}_\tau(v)$, thereby inducing a DAG G. Candidate orders are explored by reinsertion moves (Algorithm 2). During the evaluation of a reinsertion move, the affected parent sets and local scores are updated locally using a warm-start grow-shrink procedure (Algorithm 3), avoiding a full recomputation for all nodes. This reinsertion-move and warm-start grow-shrink machinery is unchanged from FLOP; the component that adapts it to interventional data is the target-filtered pooled local score s_v, introduced next, which is the score these procedures optimize. After the reinsertion-based local search reaches a local optimum, Iterated Local Search (ILS) applies a random perturbation to the current best order and

restarts local search from the perturbed order; the best-scoring order found across these restarts is retained (Lourenço et al., 2018; De Campos et al., 2003).

3. I-FLOP: order-based search for interventional data

Algorithm 1: Target-filtered order search for known-target interventional data

Input: Multi-environment data $\{X^{(e)}\}_{e\in\mathcal{E}}$, target family $\mathcal{I} = \{I_e\}_{e\in\mathcal{E}}$, initial order τ_0
Output: Estimated DAG $\widehat{G}$ and its induced I-CPDAG $\widehat{C}_{\mathcal{I}}$

`// Phase 1: Precompute environment statistics`
foreach $e \in \mathcal{E}$ **do**
  Center $X^{(e)}$ within environment e and compute scatter matrix $C^{(e)}$
end

`// Phase 2: Construct target-filtered pooled scores`
foreach $v \in V$ **do**
  $\mathcal{E}_v \leftarrow \{e \in \mathcal{E} : v \notin I_e\}$
  $N_v \leftarrow \sum_{e\in\mathcal{E}_v} n_e$
  $C_v^{\text{pool}} \leftarrow \sum_{e\in\mathcal{E}_v} C^{(e)}$
  Define $s_v := s_v^{\text{tf-pool}}$ from C_v^{pool} and N_v
end

`// Phase 3: Initialize parent sets and local scores under` τ_0
$\tau \leftarrow \tau_0$
foreach $v \in V$ **do**
  $Z_v \leftarrow \text{Pred}_\tau(v)$
  $\ell_{\emptyset,v} \leftarrow s_v(\emptyset)$
  $(P_v, \ell_v) \leftarrow \text{GrowShrink}_{\text{tf}}(v, Z_v, \emptyset, \ell_{\emptyset,v}, 0)$ `// Algorithm 3`
end
Let $P := (P_v)_{v\in V}$ and $\ell := (\ell_v)_{v\in V}$

`// Phase 4: Reinsertion-based local order search`
improved $\leftarrow$ true
while improved **do**
  $(\tau^{\text{old}}, \ell^{\text{old}}) \leftarrow (\tau, \ell)$
  foreach $v \in \tau^{\text{old}}$ **do**
    $(\tau, P, \ell) \leftarrow \text{Reinsert}_{\text{tf}}(\tau, P, \ell, v)$ `// Algorithm 2`
  end
  improved $\leftarrow \left(\sum_{v\in V} \ell_v < \sum_{v\in V} \ell_v^{\text{old}}\right)$
end

`// Phase 5: Graph construction and output`
$\widehat{G} \leftarrow \left(V, \{u \to v : u \in P_v,\ v \in V\}\right)$
$\widehat{C}_{\mathcal{I}} \leftarrow \text{I-CPDAG}(\widehat{G}, \mathcal{I})$

return $(\widehat{G}, \widehat{C}_{\mathcal{I}})$

Algorithm 1 presents the core reinsertion-based local order search. In our implementation, the optional ILS wrapper described above repeatedly perturbs the locally optimal order and reruns the reinsertion search, retaining the best-scoring solution across restarts.

We first state two assumptions for the known-target hard-intervention setting: one on the data-generating process and one on the conservativeness of the intervention target family.

Assumption 1 (Data-generating process and causal sufficiency)
We assume causal sufficiency over V and a linear Gaussian SEM with true DAG $G^\star = (V, E)$. In each environment $e \in \mathcal{E}$, the known hard-intervention targets $I_e \subseteq V$ satisfy $X_{I_e}^{(e)} \sim Q_{I_e}^{(e)}$, while each non-targeted node $v \notin I_e$ retains its structural equation $X_v^{(e)} = \sum_{u \in \mathrm{pa}_{G^\star}(v)} W_{uv}^\star X_u^{(e)} + \varepsilon_v^{(e)}$, where $\varepsilon_v^{(e)} \sim \mathcal{N}(0, (\sigma_v^\star)^2)$. The noise variables are mutually independent, have positive variances, and are invariant across environments.

Assumption 2 (Conservative family of targets)
Following Hauser and Bühlmann (2012), we assume the intervention target family $\mathcal{I} = \{I_e : e \in \mathcal{E}\}$ is conservative, meaning that for every node $v \in V$, there exists at least one environment $e \in \mathcal{E}$ such that $v \notin I_e$. Intuitively, this assumption ensures that no variable is perpetually perturbed across all experimental settings, so that the unperturbed, natural causal mechanism $P(v \mid \mathrm{pa}(v))$ of each variable can be observed and learned from at least one environment.

3.1. Valid environments under interventions

In the observational setting, FLOP scores each candidate parent set by fitting a single local regression on samples assumed to come from one observational regime. This is no longer appropriate for interventional multi-environment data: different environments may induce different joint distributions, and samples from environments where node v is directly intervened upon do not follow the original mechanism of v. The invariant object is instead the local mechanism of v across environments in which v is not targeted. This leads to response-target filtering:

Definition 1 *For each node $v \in V$, we define the valid environment set $\mathcal{E}_v := \{e \in \mathcal{E} : v \notin I_e\}$, which consists of the environments in which v is not directly intervened upon, and the corresponding total sample size is $N_v := \sum_{e \in \mathcal{E}_v} n_e$.*

We use N to denote the total sample size over all environments. The valid environment set $\mathcal{E}_v$ is node-specific: different response nodes may use different subsets of environments for local scoring. Importantly, the validity of an environment for node v depends only on whether v itself is directly intervened upon. In particular, let $P \subseteq V \setminus \{v\}$ be a candidate parent set for v. If $u \in P$ is directly intervened upon in environment e, while $v \notin I_e$, then $e \in \mathcal{E}_v$ remains valid for scoring v. By Assumption 1, intervening on u may change the distribution of $X_u^{(e)}$, but it does not replace the structural equation of v when $v \in V \setminus I_e$. Therefore, observations from all environments in $\mathcal{E}_v$ can be pooled to estimate the same structural equation for v. We next use these node-specific valid environments to define the target-filtered pooled local score.

3.2. Target-filtered pooled local score

Definition 2 *For node $v \in V$ and candidate parent set $P \subseteq V \setminus \{v\}$, define the target-filtered pooled local score by*

$$s_v^{\text{tf-pool}}(P) = \frac{N_v}{2}\left(1 + \log \widehat{\sigma}^2_{v|P}\right) + \frac{1}{2}\log N\,(|P|+1).$$

Here, $\widehat{\sigma}^2_{v|P}$ is estimated only from the valid environments $\mathcal{E}_v$.

For each environment, the centered data and scatter matrix are

$$\widetilde{X}^{(e)} = X^{(e)} - \bar{X}^{(e)}, \qquad C^{(e)} = (\widetilde{X}^{(e)})^\top \widetilde{X}^{(e)},$$

and the pooled scatter matrix for node v over its valid environments is $C_v^{\text{pool}} = \sum_{e\in\mathcal{E}_v} C^{(e)}$. We impose the following non-singularity assumption on the relevant principal submatrices of C_v^{pool}.

Assumption 3 (Local non-singularity) *For every node $v \in V$ and every candidate parent set $P \subseteq V \setminus \{v\}$ considered in the local score evaluation, the principal submatrix $(C_v^{\text{pool}})_{PP}$ is non-singular.*

Under Assumption 3, the pooled residual variance estimator $\widehat{\sigma}^2_{v|P}$ for node v and candidate parent set P is defined as

$$\widehat{\sigma}^2_{v|P} = \frac{1}{N_v}\Big[(C_v^{\text{pool}})_{vv} - (C_v^{\text{pool}})_{vP}((C_v^{\text{pool}})_{PP})^{-1}(C_v^{\text{pool}})_{Pv}\Big],$$

where the second term is set to zero when $P = \emptyset$ (the Schur complement of the pooled scatter matrix; see the Frisch–Waugh–Lovell identity used in Appendix A). Thus, $\widehat{\sigma}^2_{v|P}$ corresponds to the residual variance obtained by fitting a shared local linear mechanism for v across its valid environments, rather than across all environments.

By construction, this score preserves node-wise decomposability and is compatible with order-restricted parent sets. For a candidate DAG G', the total score is

$$S_{\text{tf-pool}}(G'; \mathcal{D}, \mathcal{I}) = \sum_{v\in V} s_v^{\text{tf-pool}}(\text{Pa}_{G'}(v)).$$

Given a fixed order τ, score-based I-FLOP selects parent sets by

$$P_v(\tau) \in \operatorname*{arg\,min}_{P\subseteq \text{Pred}_\tau(v)} s_v^{\text{tf-pool}}(P),$$

and induces the order-restricted DAG $G_\tau = \{u \to v : u \in P_v(\tau)\}$.

The corresponding order objective is

$$Q_{\text{score}}(\tau) = \sum_{v\in V} s_v^{\text{tf-pool}}(P_v(\tau)).$$

3.3. Warm-start parent-set selection

For each node v and prefix Z, I-FLOP updates the candidate parent set using a grow–shrink procedure with the target-filtered pooled local score (Algorithm 3).[2] This strategy is inherited from FLOP, which recovers the restricted Markov boundary under a single observational distribution. In our setting, however, valid environments preserve the response mechanism but may exhibit different joint distributions due to interventions on other nodes. Consequently, the restricted Markov boundary characterization does not directly apply to multiple data sources. To characterize the asymptotic objective optimized by the grow–shrink procedure, we introduce a pooled population residual risk over the valid environments. The following assumption specifies the asymptotic sample proportions contributed by each environment.

Assumption 4 (Environment proportions) *For every environment $e \in \mathcal{E}$, the sample proportion satisfies $n_e/N \to \pi_e$, where $\pi_e > 0$. Consequently, for every node v, $N_v/N \to \alpha_v := \sum_{e \in \mathcal{E}_v} \pi_e > 0$.*

For each environment $e \in \mathcal{E}$, let $\Sigma^{(e)} = \mathrm{Cov}\left(\widetilde{X}^{(e)}\right)$ denote the covariance matrix of the environment-centered variables. Under Assumptions 1, 2, and 4, the normalized valid-environment pooled scatter matrix $N_v^{-1} C_v^{\mathrm{pool}}$ converges in probability to $\Sigma_v^{\mathrm{pool}} = \frac{1}{\alpha_v} \sum_{e \in \mathcal{E}_v} \pi_e \Sigma^{(e)}$ (by the weak law of large numbers applied within each valid environment, combined via Assumption 4's sample-proportion limits). For a candidate parent set $P \subseteq Z$, the corresponding population pooled residual risk is

$$\sigma^2_{v|P,\mathrm{pool}} = (\Sigma_v^{\mathrm{pool}})_{vv} - (\Sigma_v^{\mathrm{pool}})_{vP} \big((\Sigma_v^{\mathrm{pool}})_{PP}\big)^{-1} (\Sigma_v^{\mathrm{pool}})_{Pv}.$$

This quantity is the population counterpart of $\widehat{\sigma}^2_{v|P}$ and characterizes the asymptotic fit of the shared response mechanism across the valid environments. It therefore induces the population target of the grow–shrink procedure:

Definition 3 *The pooled local parent target of v relative to Z is defined as the sparsest candidate parent set in Z that gives the optimal valid-environment structural-mechanism fit:*

$$M_v^{\mathrm{pool}}(Z) \in \arg\min_{P \subseteq Z} \sigma^2_{v|P,\mathrm{pool}},$$

where ties are resolved in favor of a sparsest minimizer.

Although single-variable greedy search procedures could in principle be susceptible to local optima, the underlying linear Gaussian model and positive definiteness of the pooled covariance guarantee that the pooled residual-risk landscape has no local traps:

Proposition 4 *Fix $v \in V$ and $Z \subseteq V \setminus \{v\}$. Under Assumptions 1, 2, and 4, the pooled residual risk criterion $R(P) = \sigma^2_{v|P,\mathrm{pool}}$ has a unique sparsest minimizer $M_v^{\mathrm{pool}}(Z) = \mathrm{supp}(w_Z)$, where $w_Z = (\Sigma_{ZZ}^{\mathrm{pool}})^{-1} \Sigma_{Zv}^{\mathrm{pool}}$. Furthermore, for every candidate set $P \subseteq Z$ with $M_v^{\mathrm{pool}}(Z) \not\subseteq P$, there exists $u \in M_v^{\mathrm{pool}}(Z) \setminus P$ satisfying $\sigma^2_{v|P\cup\{u\},\mathrm{pool}} < \sigma^2_{v|P,\mathrm{pool}}$.*

2. Proofs of all theoretical results in this section are provided in Appendix A.

Proposition 4 establishes that whenever the current candidate parent set P does not fully cover the optimal parent set $M_v^{\text{pool}}(Z)$, there always exists at least one missing parent whose addition strictly reduces the pooled residual risk. Consequently, target-filtered grow–shrink provably converges to the global parent target from any initialization:

Theorem 5 *Fix $v \in V$ and $Z \subseteq V \setminus \{v\}$. Let $\widehat{P}_{v,N}(Z)$ be the parent set returned by target-filtered grow–shrink using $s_v^{\text{tf-pool}}$. Under Assumptions 1–4, starting from any initial parent set $P_0 \subseteq Z$,*

$$\mathbb{P}\left(\widehat{P}_{v,N}(Z) = M_v^{\text{pool}}(Z)\right) \longrightarrow 1 \quad \text{as } N \to \infty.$$

Since a fixed order τ induces a fixed-prefix problem for each node, applying Theorem 5 node-wise immediately yields the following fixed-order recovery guarantee:

Corollary 6 *Fix an order τ, and let $Z_v(\tau) = \text{Pred}_\tau(v)$ for each $v \in V$. For each node v, let $\widehat{P}_{v,N}(\tau)$ be the parent set returned by target-filtered grow–shrink with prefix $Z_v(\tau)$, and define the induced graph $\widehat{G}_N(\tau) = \{u \to v : u \in \widehat{P}_{v,N}(\tau),\ v \in V\}$. Then, under Assumptions 1–4,*

$$\mathbb{P}\left(\text{Pa}_{\widehat{G}_N(\tau)}(v) = M_v^{\text{pool}}(Z_v(\tau)) \text{ for all } v \in V\right) \longrightarrow 1 \quad \text{as } N \to \infty.$$

Theorem 5 and Corollary 6 establish the large-sample behavior of the target-filtered grow–shrink updates for a fixed prefix and a fixed order, respectively. These results justify the local parent-set selection used in I-FLOP. We next show that the target-filtered pooled score admits a valid likelihood interpretation under known interventions and use this connection to state its global recovery property.

3.4. Global recovery from the target-filtered pooled score

Under known hard interventions, the replacement mechanism of a directly targeted node does not depend on its candidate parent set in the original structural mechanism and therefore contributes only a graph-independent nuisance term. After profiling out environment-specific intercepts, equivalently by working with the environment-centered observations, let

$$\ell_N(G, \theta_G) = \sum_{v \in V} \sum_{e \in \mathcal{E}_v} \ell_{v,e}\big(\text{Pa}_G(v); \theta_v\big),$$

denote the resulting log-likelihood contribution of the original structural mechanisms, evaluated only over the valid environments of each response node. Here, $\theta_G = \{\theta_v : v \in V\}$, where θ_v contains the regression coefficients associated with $\text{Pa}_G(v)$ and the residual variance of the original mechanism of node v.

Lemma 7 *Under Assumptions 1–3, the target-filtered pooled graph score is the decomposable score associated with $\ell_N(G, \theta_G)$, up to terms independent of the candidate DAG. In particular, for any candidate DAG G',*

$$S_{\text{tf-pool}}(G'; \mathcal{D}, \mathcal{I}) = -\sup_{\theta_{G'}} \ell_N(G', \theta_{G'}) + \frac{1}{2} \log N \sum_{v \in V} \big(|\,\text{Pa}_{G'}(v)| + 1\big) + C_N.$$

where $C_N = -\frac{1}{2} \sum_{v \in V} N_v \log(2\pi)$ does not depend on G'.

We now characterize the global target of exact score minimization in terms of I-MEC. For a candidate DAG G, define the population counterpart of the graph-dependent fit component of the target-filtered pooled score as

$$L_\infty(G) = \sum_{v \in V} \frac{\alpha_v}{2} \log \sigma^2_{v|\mathrm{Pa}_G(v),\mathrm{pool}}, \qquad G \in \mathcal{G}_{\mathrm{DAG}}.$$

Let

$$\mathcal{A} = \arg\min_{G \in \mathcal{G}_{\mathrm{DAG}}} L_\infty(G), \qquad \mathcal{A}_{\min} = \arg\min_{G \in \mathcal{A}} |E_G|.$$

The restriction to $\mathcal{A}_{\min}$ corresponds to the graph-dependent complexity term

$$\sum_{v \in V} \big(|\,\mathrm{Pa}_G(v)| + 1\big) = |E_G| + |V|.$$

By definition, $\mathcal{A}$ contains the DAGs attaining the optimal population fit of the valid-environment structural mechanisms, and $\mathcal{A}_{\min}$ retains the sparsest DAGs among these representations. To conclude that the members of $\mathcal{A}_{\min}$ belong to the true $\mathcal{I}$-MEC, we require the conditional independence relations induced by the known intervention targets to coincide with those encoded by the true DAG $G^\star$. We therefore impose the following interventional faithfulness assumption.

Assumption 5 (Interventional faithfulness assumption)
Let $\mathcal{G}^\star_{\mathcal{I}} := \{G \in \mathcal{G}_{\mathrm{DAG}} : M_{\mathcal{I}}(G) = M_{\mathcal{I}}(G^\star)\}$ denote the true I-MEC. The interventional distribution family generated by $G^\star$ is faithful to $G^\star$ with respect to the known intervention target family $\mathcal{I}$.

Theorem 8 *Under Assumptions 1, 2, 3, 4 and 5, the sparsest global residual-risk minimizers belong to the true I-MEC:*

$$\mathcal{A}_{\min} \subseteq \mathcal{G}^\star_{\mathcal{I}} = \{G \in \mathcal{G}_{\mathrm{DAG}} : M_{\mathcal{I}}(G) = M_{\mathcal{I}}(G^\star)\}\,.$$

Theorem 8 is a global statement. For the empirical score, uniform convergence of the target-filtered residual risks and the BIC sparsity penalty imply that an exact global minimizer selects a DAG in $\mathcal{A}_{\min}$ in the limit $N \to \infty$. Combining this with $\mathcal{A}_{\min} \subseteq \mathcal{G}^\star_{\mathcal{I}}$ gives the following I-CPDAG recovery result:

Corollary 9 *Let $\widehat{G}_N$ be an exact global minimizer of $S_{\text{tf-pool}}(G; \mathcal{D}, \mathcal{I})$ over DAG G. Under Assumptions 1, 2, 3, 4, and 5, $\widehat{G}_N$ belongs to the true I-MEC with probability tending to one. If the target family $\mathcal{I}$ fully identifies the DAG, so that $\mathcal{G}^\star_{\mathcal{I}} = \{G^\star\}$, then $\Pr(\widehat{G}_N = G^\star) \to 1$.*

4. Experiments

In this section, we evaluate I-FLOP for interventional causal discovery. We compare it with two external baselines, GIES and I-GSP, and two FLOP-based variants: FLOP-obs, the original observational method, and FLOP-envwise, which uses environment-wise variance estimation without target-based environment filtering. We generate synthetic data

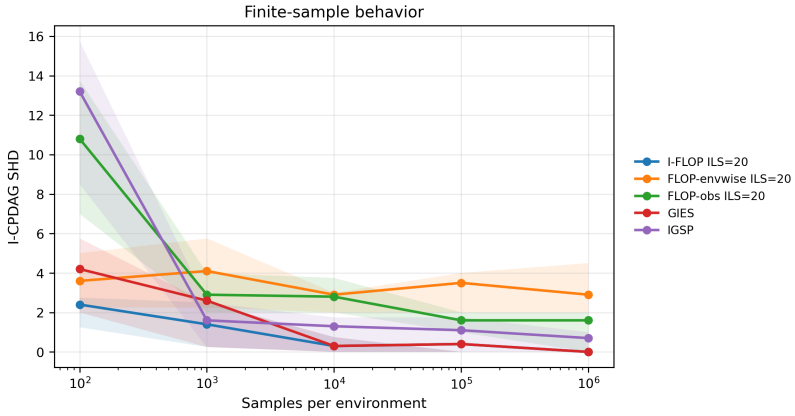


Figure 1: I-CPDAG SHD decreases as sample size increases from 10^2 to 10^6 per environment with 20 nodes and expected degree 2 for one observational and five known-target singleton hard intervention environments.

by Erdős–Rényi (ER) DAGs (Erdős and Rényi, 1959) under linear Gaussian SEMs. Each dataset contains one observational environment and multiple known-target hard interventional environments.

The experiment can be divided into three main parts. First, we assess finite-sample convergence, specifically whether the algorithms recover the I-CPDAG (Section 4.1). Second, we evaluate the algorithms' performances under well-specified synthetic settings (Section 4.2). Finally, we demonstrate performance on real interventional data from the Causal Chamber Light Tunnel system (Section 4.3).

4.1. Finite sample convergence toward the asymptotic guarantee

We first investigate the finite-sample convergence behavior of I-FLOP as the sample size increases. Our theoretical results establish that, in the large-sample limit, optimizing the target-filtered pooled score recovers the true I-CPDAG. This experiment therefore examines whether the structural estimates produced by I-FLOP become increasingly accurate and stable with more data. Figure 1 shows that I-FLOP recovers the I-CPDAG in large sample limit, which supports our claim.

4.2. Main accuracy and runtime comparison

Figure 2 compares accuracy and runtime in a representative well-specified synthetic setting. I-FLOP achieves the lowest I-CPDAG SHD in both ER graph regimes, with a clearer advantage on denser graphs, where GIES and I-GSP have substantially larger structural errors. I-FLOP is also markedly faster than GIES in the denser setting while remaining substantially more accurate than the faster I-GSP baseline, yielding a favorable accuracy–runtime trade-off. Table 1 shows the decomposition comparison of I-CPDAG SHD between our I-FLOP and the external baselines.

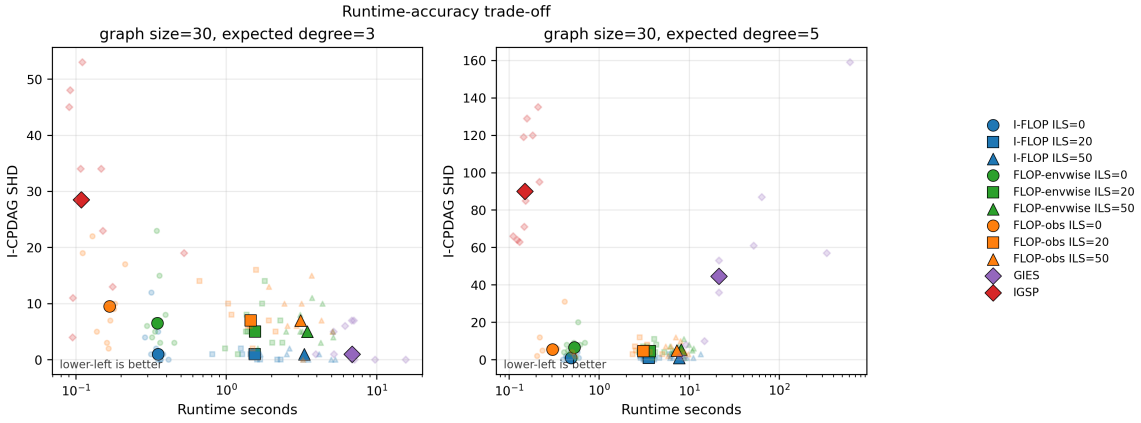


Figure 2: The I-CPDAG Structural Hamming Distance (SHD) is plotted against runtime for linear-Gaussian ER DAGs. Each setting uses one observational environment and five known-target singleton hard intervention environments, with 1,000 samples per environment.

Table 1: SHD decomposition of experiment in well-specified synthetic setting

	Expected degree $d = 3$				Expected degree $d = 5$			
Algorithm	Mean	Std.	Skel.	Orient.	Mean	Std.	Skel.	Orient.
I-FLOP (ILS=0)	2.50	3.78	1.30	1.20	1.60	1.58	0.90	0.70
I-FLOP (ILS=20)	0.70	0.67	0.60	0.10	1.50	1.27	1.10	0.40
I-FLOP (ILS=50)	0.70	0.67	0.60	0.10	1.50	1.27	1.10	0.40
GIES	2.70	3.13	2.00	0.70	47.40	49.02	39.10	8.30
I-GSP	28.40	16.93	19.60	8.80	94.70	28.79	73.10	21.60

4.3. Real-data evaluation on Causal Chamber

We evaluate algorithms on real interventional data from the Causal Chamber Light Tunnel system (Gamella et al., 2025). The view we use contains 20 variables, including one observational environment (10,000 samples) and five singleton hard-intervention environments with known targets (red, green, blue, `pol_1`, `pol_2`, each containing 1,000 samples), all derived from `lt_interventions_standard_v1`. Figure 3 shows that I-FLOP exhibits the best accuracy–runtime trade-off on this real interventional dataset.

5. Conclusion

We introduce I-FLOP, an order-based score search method for causal structure learning from multi-environment interventional data with known targets. By replacing FLOP's observational local score with a target-filtered pooled interventional score, I-FLOP retains the efficiency of warm-start grow–shrink updates and reinsertion-based order search while incorporating intervention information in a principled manner. Under Assumptions 1–

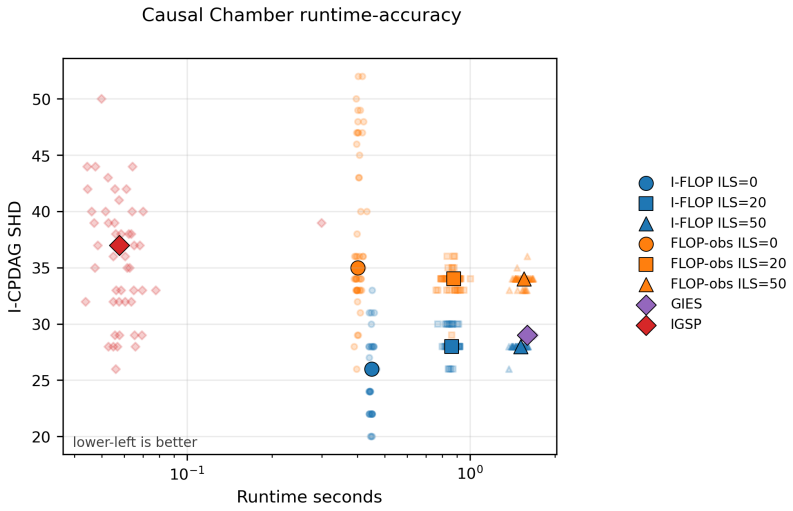


Figure 3: I-FLOP (ILS = 0) attains the lowest mean I-CPDAG SHD while remaining faster than GIES. All I-FLOP configurations outperform FLOP-obs and I-GSP, and positive ILS budgets reduce variability (details in Appendix D).

5, exact global minimization of the target-filtered pooled score asymptotically recovers the true I-MEC, while under Assumptions 1–4, the fixed-prefix target-filtered grow–shrink update recovers the pooled local parent target in the large-sample limit. Empirically, I-FLOP achieves a favorable accuracy–runtime trade-off on synthetic interventional data and competitive performance on the Causal Chamber benchmark.

The main limitations of the current framework stem from its reliance on known hard interventions, linear-Gaussian mechanisms with strictly positive error variances, and conservative intervention designs. These assumptions restrict the model class to stochastic linear systems and exclude deterministic or functional relationships with degenerate noise, as well as nonlinear or uncertain intervention mechanisms. This mismatch is particularly relevant for physical systems such as the Causal Chamber, where variables may obey deterministic conservation laws or bounded nonlinear responses, and helps explain why more aggressive score optimization does not necessarily yield lower SHD when the data-generating process departs from the assumed model class. Future work could therefore extend target-filtered order search to deterministic or functional relationships, nonlinear additive noise models, and soft, uncertain, or partially known intervention targets (i.e., settings where the identity of the intervened variable(s) in a given environment is not fully or reliably known, as considered by Squires et al. (2020) and Gamella et al. (2022)), potentially through more flexible non-parametric or generalized scoring criteria.

Acknowledgments

A.M. was partially supported by Novo Nordisk Foundation grant NNF20OC0062897 and also by the Pioneer Centre for Statistical and computational Methods for Advanced Re-

search to Transform Biomedicine (SMARTbiomed), DNRF grant number P4. We thank Ming Cai from the Graduate School of Informatics, Kyoto University, for helpful discussions and comments on an earlier version of the manuscript.

References

Steen A. Andersson, David Madigan, and Michael D. Perlman. A characterization of Markov equivalence classes for acyclic digraphs. *The Annals of Statistics*, 25(2):505–541, 1997. doi: 10.1214/aos/1031833662.

Remco R. Bouckaert. Optimizing causal orderings for generating DAGs from data. In *Uncertainty in Artificial Intelligence*, pages 9–16. Morgan Kaufmann, 1992. doi: 10.1016/B978-1-4832-8287-9.50006-2.

Luis M. De Campos, Juan M. Fernández-Luna, and J. Miguel Puerta. An iterated local search algorithm for learning Bayesian networks with restarts based on conditional independence tests. *International Journal of Intelligent Systems*, 18(2):221–235, 2003. doi: 10.1002/int.10085.

Paul Erdős and Alfred Rényi. On random graphs I. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959. doi: 10.5486/PMD.1959.6.3-4.12.

Nir Friedman and Daphne Koller. Being Bayesian about network structure. a Bayesian approach to structure discovery in Bayesian networks. *Machine Learning*, 50:95–125, 2003. doi: 10.1023/A:1020249912095.

Juan L. Gamella, Armeen Taeb, Christina Heinze-Deml, and Peter Bühlmann. Characterization and greedy learning of Gaussian structural causal models under unknown interventions. *arXiv preprint arXiv:2211.14897*, 2022. doi: 10.48550/arXiv.2211.14897.

Juan L. Gamella, Jonas Peters, and Peter Bühlmann. Causal chambers as a real-world physical testbed for AI methodology. *Nature Machine Intelligence*, 7(1):107–118, 2025. doi: 10.1038/s42256-024-00964-x.

Alain Hauser and Peter Bühlmann. Characterization and greedy learning of interventional Markov equivalence classes of directed acyclic graphs. *Journal of Machine Learning Research*, 13(79):2409–2464, 2012. URL http://jmlr.org/papers/v13/hauser12a.html.

Alain Hauser and Peter Bühlmann. Jointly interventional and observational data: Estimation of interventional Markov equivalence classes of directed acyclic graphs. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 77(1):291–318, 01 2015. doi: 10.1111/rssb.12071.

P. Larranaga, C.M.H. Kuijpers, R.H. Murga, and Y. Yurramendi. Learning Bayesian network structures by searching for the best ordering with genetic algorithms. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 26(4):487–493, 1996. doi: 10.1109/3468.508827.

Steffen L. Lauritzen. *Graphical Models*. Oxford University Press, 2nd edition, June 2026. doi: 10.1093/9780191775284.001.0001.

Helena Ramalhinho Lourenço, Olivier C. Martin, and Thomas Stützle. Iterated local search: Framework and applications. In *Handbook of metaheuristics*, pages 129–168. Springer, 2018. doi: 10.1007/978-3-319-91086-4_5.

Michael C. Lovell. A simple proof of the FWL theorem. *Journal of Economic Education*, 39(1):88–91, 2008. doi: 10.3200/JECE.39.1.88-91.

Jonas Peters and Peter Bühlmann. Identifiability of Gaussian structural equation models with equal error variances. *Biometrika*, 101(1):219–228, 2014. doi: 10.1093/biomet/ast043.

Jonas Peters, Peter Bühlmann, and Nicolai Meinshausen. Causal inference by using invariant prediction: Identification and confidence intervals. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 78(5):947–1012, 2016. doi: 10.1111/rssb.12167.

Mauro Scanagatta, Giorgio Corani, and Marco Zaffalon. Improved local search in bayesian networks structure learning. In *Proceedings of The 3rd International Workshop on Advanced Methodologies for Bayesian Networks*, volume 73 of *Proceedings of Machine Learning Research*, pages 45–56. PMLR, 20–22 Sep 2017. URL https://proceedings.mlr.press/v73/scanagatta17a.html.

Moninder Singh and Marco Valtorta. An algorithm for the construction of Bayesian network structures from data. In *Proceedings of the Ninth Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 259–265, 1993. doi: 10.48550/arXiv.1303.1485.

Chandler Squires, Yuhao Wang, and Caroline Uhler. Permutation-based causal structure learning with unknown intervention targets. In *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence (UAI)*, volume 124 of *Proceedings of Machine Learning Research*, pages 1039–1048. PMLR, 2020. URL https://proceedings.mlr.press/v124/squires20a.html.

Marc Teyssier and Daphne Koller. Ordering-based search: A simple and effective algorithm for learning Bayesian networks. In *Proceedings of the Twenty-First Conference on Uncertainty in Artificial Intelligence*, UAI'05, pages 584–590, Arlington, Virginia, USA, 2005. AUAI Press. doi: 10.48550/arXiv.1207.1429.

Thomas Verma and Judea Pearl. Equivalence and synthesis of causal models. In *Proceedings of the Sixth Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 255–270. Elsevier Science Inc., 1990. doi: 10.48550/arXiv.1304.1108.

Yuhao Wang, Liam Solus, Karren Yang, and Caroline Uhler. Permutation-based causal inference algorithms with interventions. In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL https://proceedings.neurips.cc/paper_files/paper/2017/file/275d7fb2fd45098ad5c3ece2ed4a2824-Paper.pdf.

Marcel Wienöbst, Leonard Henckel, and Sebastian Weichwald. Embracing discrete search: A reasonable approach to causal structure learning. In *International Conference on Learning Representations*, volume 2026, pages 54630–54652, 2026. URL https://openreview.net/forum?id=lejOV6j3cj.

Appendix A. Proofs of Theoretical Results

A.1. Single-Move Score Comparisons

We first establish the large-sample score comparisons which will be used in the proof of Theorem 5.

Lemma 10 *Fix $v \in V$ and $Z \subseteq V \setminus \{v\}$. Under Assumptions 1–4, the following statements hold simultaneously over the finitely many candidate parent sets contained in Z with probability tending to one.*

1. *If $P, Q \subseteq Z$ satisfy $\sigma^2_{v|Q,\text{pool}} < \sigma^2_{v|P,\text{pool}}$, then $s_v^{\text{tf-pool}}(Q) < s_v^{\text{tf-pool}}(P)$.*
2. *If $u \in P \subseteq Z$ and $\sigma^2_{v|P\setminus\{u\},\text{pool}} = \sigma^2_{v|P,\text{pool}}$, then $s_v^{\text{tf-pool}}(P \setminus \{u\}) < s_v^{\text{tf-pool}}(P)$.*

Proof By Assumptions 1 and 4, the environment-wise sample covariances converge to their population counterparts. Since $n_e/N \to \pi_e > 0$ for every $e \in \mathcal{E}$, it follows that, for every fixed $P \subseteq Z$,

$$\widehat{\sigma}^2_{v|P} \xrightarrow{p} \sigma^2_{v|P,\text{pool}}.$$

Since Z is fixed, there are only finitely many candidate parent sets, and this convergence holds simultaneously over all $P \subseteq Z$.

For the first statement, suppose that $\sigma^2_{v|Q,\text{pool}} < \sigma^2_{v|P,\text{pool}}$. Then

$$\frac{1}{N_v}\left(s_v^{\text{tf-pool}}(Q) - s_v^{\text{tf-pool}}(P)\right) \xrightarrow{p} \frac{1}{2}\log\frac{\sigma^2_{v|Q,\text{pool}}}{\sigma^2_{v|P,\text{pool}}} < 0.$$

The fit contribution is of order N_v, whereas the difference in complexity penalties is of order $\log N$, and $N_v/N \to \alpha_v > 0$. Therefore, any move that strictly improves the population fit decreases the empirical local score with probability tending to one.

For the second statement, suppose that deleting u leaves the population residual risk unchanged. The two nested local models then attain the same population optimum for the centered valid-environment criterion. Under Assumptions 1–4, a standard fixed-dimensional quadratic expansion of the maximized criterion gives an $O_p(1)$ fit improvement from retaining the redundant parent u. Retaining u, however, incurs the additional penalty $\frac{1}{2}\log N$, which diverges. Therefore,

$$s_v^{\text{tf-pool}}(P \setminus \{u\}) < s_v^{\text{tf-pool}}(P).$$

■

A.2. Proof of Proposition 4

Proof Fix $v \in V$ and $Z \subseteq V \setminus \{v\}$. Under Assumptions 1, 2 and 4, the valid-environment pooled covariance matrix $\Sigma_v^{\text{pool}} = \frac{1}{\alpha_v}\sum_{e\in\mathcal{E}_v}\pi_e\Sigma^{(e)}$ is positive definite, since each $\Sigma^{(e)} > 0$ and $\pi_e/\alpha_v > 0$ for every $e \in \mathcal{E}_v$. Hence every principal submatrix of Σ_v^{pool}, in particular $\Sigma_{ZZ}^{\text{pool}}$, is positive definite and therefore invertible.

Define the pooled population regression coefficient vector of X_v on X_Z by

$$w_Z := (\Sigma_{ZZ}^{\text{pool}})^{-1}\Sigma_{Zv}^{\text{pool}}, \qquad S := \text{supp}(w_Z).$$

Since $\Sigma_{ZZ}^{\text{pool}} > 0$, the least-squares coefficient vector w_Z is unique. For each $P \subseteq Z$, let $R(P) := \sigma^2_{v|P,\text{pool}}$ denote the pooled population residual variance from regressing X_v on X_P.

We first show that $R(P) = R(Z)$ if and only if $S \subseteq P$. If $S \subseteq P$, then w_Z vanishes on $Z \setminus P$, so the optimal predictor based on X_Z depends only on X_P. Hence $R(P) \leq R(Z)$. Since $P \subseteq Z$, enlarging the regressor set cannot increase the least-squares residual variance, so $R(Z) \leq R(P)$. Therefore $R(P) = R(Z)$.

Conversely, suppose that $R(P) = R(Z)$. Extend the least-squares coefficient vector for the regression on X_P to a vector indexed by Z by assigning zero coefficients to $Z \setminus P$. This extended vector achieves residual variance $R(Z)$, and is therefore also a minimizer of the least-squares problem over X_Z. By uniqueness of w_Z, it must coincide with w_Z. Thus $(w_Z)_u = 0$ for every $u \in Z \setminus P$, which implies $S \subseteq P$.

Therefore, the minimizers of $R(P)$ over $P \subseteq Z$ are exactly the supersets of S. In particular, S is the unique minimizer of smallest cardinality. We have $M_v^{\text{pool}}(Z) = S = \text{supp}(w_Z)$. It remains to prove the absence of local traps. Let $P \subseteq Z$ satisfy $M_v^{\text{pool}}(Z) \nsubseteq P$, and set $A := M_v^{\text{pool}}(Z) \setminus P$, so that $A \neq \emptyset$. Suppose, toward a contradiction, that $R(P \cup \{u\}) = R(P)$ for every $u \in A$.

For $u \in A$, let r_v^P and r_u^P denote the population residuals obtained by projecting X_v and X_u, respectively, onto X_P under the pooled distribution. By the Frisch–Waugh–Lovell theorem (Lovell, 2008), equivalently the Schur complement identity,

$$R(P) - R(P \cup \{u\}) = \frac{\text{Cov}_{\text{pool}}(r_v^P, r_u^P)^2}{\text{Var}_{\text{pool}}(r_u^P)}.$$

Since $\Sigma_v^{\text{pool}} > 0$, we have $\text{Var}_{\text{pool}}(r_u^P) > 0$. Hence the assumption $R(P \cup \{u\}) = R(P)$ implies $\text{Cov}_{\text{pool}}(r_v^P, r_u^P) = 0$ for every $u \in A$.

Let r_A^P denote the vector of residuals obtained by projecting X_A onto X_P. Then $\text{Cov}_{\text{pool}}(r_A^P, r_v^P) = \mathbf{0}$. Applying the multivariate Frisch–Waugh–Lovell identity gives

$$R(P) - R(P \cup A) = \text{Cov}_{\text{pool}}(r_v^P, r_A^P)\big[\text{Var}_{\text{pool}}(r_A^P)\big]^{-1}\,\text{Cov}_{\text{pool}}(r_A^P, r_v^P) = 0,$$

where $\text{Var}_{\text{pool}}(r_A^P) > 0$ by positive definiteness of the corresponding Schur complement. Thus $R(P \cup A) = R(P)$.

On the other hand, $P \cup A = P \cup M_v^{\text{pool}}(Z) \supseteq M_v^{\text{pool}}(Z) = S$. By the characterization above, this implies $R(P \cup A) = R(Z)$. Therefore $R(P) = R(Z)$, which in turn implies $S \subseteq P$. Since $S = M_v^{\text{pool}}(Z)$, this contradicts the assumption $M_v^{\text{pool}}(Z) \nsubseteq P$.

Hence there exists at least one $u \in M_v^{\text{pool}}(Z) \setminus P$ such that $\sigma^2_{v|P\cup\{u\},\text{pool}} < \sigma^2_{v|P,\text{pool}}$. ■

A.3. Proof of Theorem 5

Proof Let P_{grow} denote the parent set obtained at the end of the grow phase. We first show that $M \subseteq P_{\text{grow}}$. Suppose otherwise that $M \nsubseteq P_{\text{grow}}$. By Proposition 4, there exists $u \in M \setminus P_{\text{grow}}$ such that

$$\sigma^2_{v|P_{\text{grow}}\cup\{u\},\text{pool}} < \sigma^2_{v|P_{\text{grow}},\text{pool}}.$$

While Lemma 10 gives

$$s_v^{\text{tf-pool}}\big(P_{\text{grow}} \cup \{u\}\big) < s_v^{\text{tf-pool}}\big(P_{\text{grow}}\big),$$

which contradicts the grow phase. Hence, $M \subseteq P_{\text{grow}}$.

We next show inductively that the shrink phase preserves M while removing all variables outside M. Initially, $M \subseteq P_{\text{grow}}$. Suppose that the current parent set P during the shrink phase satisfies $M \subseteq P$. Since M is a global minimizer of the pooled residual-risk criterion and adding nodes cannot increase the optimal population residual risk, then we have

$$\sigma^2_{v|P,\text{pool}} = \sigma^2_{v|M,\text{pool}}.$$

If $M \subsetneq P$, take any $u \in P \setminus M$. Since $M \subseteq P \setminus \{u\} \subseteq P$, the same monotonicity and optimality argument gives

$$\sigma^2_{v|P\setminus\{u\},\text{pool}} = \sigma^2_{v|P,\text{pool}} = \sigma^2_{v|M,\text{pool}}.$$

By Lemma 10, deleting u decreases the empirical local score:

$$s_v^{\text{tf-pool}}(P \setminus \{u\}) < s_v^{\text{tf-pool}}(P).$$

Thus, every strict superset of M admits a score-decreasing deletion of an extraneous parent. It remains to show that no score-decreasing deletion removes a member of M. Let $u \in M$. Since $M \subseteq P$, we have

$$M \setminus (P \setminus \{u\}) = \{u\}.$$

Suppose, for contradiction, that $P \setminus \{u\}$ attains the same minimum population residual risk as M. Since it does not contain M, the no-local-trap assumption implies that adding u strictly reduces its population residual risk:

$$\sigma^2_{v|P,\text{pool}} = \sigma^2_{v|(P\setminus\{u\})\cup\{u\},\text{pool}} < \sigma^2_{v|P\setminus\{u\},\text{pool}}.$$

which contradicts the assumed equality with the minimum risk attained by M. Therefore,

$$\sigma^2_{v|P\setminus\{u\},\text{pool}} > \sigma^2_{v|P,\text{pool}}.$$

Applying the first part of Lemma 10, with $Q = P$ and with P there replaced by $P \setminus \{u\}$, gives

$$s_v^{\text{tf-pool}}(P) < s_v^{\text{tf-pool}}(P \setminus \{u\}).$$

Hence deleting a node in M does not improve the score.

It follows inductively that every accepted shrink move removes a node in $P_{\text{grow}} \setminus M$, and the shrink phase terminates at M.

Therefore, one complete grow–shrink cycle returns $M_v^{\text{pool}}(Z)$ in the limit $N \to \infty$. Since the argument does not depend on the initial parent set $P_0 \subseteq Z$, target-filtered grow–shrink returns $M_v^{\text{pool}}(Z)$ with probability tending to one from any warm start. ■

A.4. Proof of Corollary 6

Proof Fix an order τ. For every node $v \in V$, let $Z_v(\tau) = \mathrm{Pred}_\tau(v)$ be the set of variables preceding v in τ. Then applying Theorem 5 to each node v gives

$$\Pr\Big(\widehat{P}_{v,N}(\tau) = M_v^{\mathrm{pool}}(Z_v(\tau))\Big) \longrightarrow 1.$$

Since V is finite, these events hold simultaneously for all nodes with probability tending to one. Therefore,

$$\Pr\Big(\widehat{P}_{v,N}(\tau) = M_v^{\mathrm{pool}}(Z_v(\tau)) \text{ for all } v \in V\Big) \longrightarrow 1.$$

By construction, the DAG $\widehat{G}_N(\tau)$ returned under the fixed order τ has parent sets

$$\mathrm{Pa}_{\widehat{G}_N(\tau)}(v) = \widehat{P}_{v,N}(\tau), \qquad v \in V.$$

It follows that

$$\mathrm{Pa}_{\widehat{G}_N(\tau)}(v) = M_v^{\mathrm{pool}}(Z_v(\tau)), \qquad v \in V,$$

simultaneously with probability tending to one. ■

A.5. Proof of Lemma 7

Proof The local centered likelihood contribution of the original mechanism of v over its valid environments is

$$\ell_N(P; \beta_v, \sigma_v^2) = \sum_{e \in \mathcal{E}_v} \ell_{v,e}(P; \beta_v, \sigma_v^2) = -\frac{N_v}{2} \log(2\pi\sigma_v^2) - \frac{1}{2\sigma_v^2} \sum_{e \in \mathcal{E}_v} \left\| \widetilde{X}_v^{(e)} - \widetilde{X}_P^{(e)} \beta_v \right\|_2^2.$$

For fixed P, maximizing over β_v yields the pooled least-squares estimator computed from the valid environments of v. Maximizing further over σ_v^2 gives

$$\widehat{\sigma}_{v|P}^2 = \frac{1}{N_v} \sum_{e \in \mathcal{E}_v} \left\| \widetilde{X}_v^{(e)} - \widetilde{X}_P^{(e)} \widehat{\beta}_{v|P} \right\|_2^2.$$

Therefore,

$$-\sup_{\beta_v, \sigma_v^2} \ell_N(P; \beta_v, \sigma_v^2) = \frac{N_v}{2} \left(\log(2\pi) + 1 + \log \widehat{\sigma}_{v|P}^2 \right).$$

Since $\ell_N(G, \theta_G)$ is node-wise decomposable,

$$-\sup_{\theta_G} \ell_N(G, \theta_G) = \sum_{v \in V} \frac{N_v}{2} \left(1 + \log \widehat{\sigma}_{v|\mathrm{Pa}_G(v)}^2 \right) + \frac{1}{2} \sum_{v \in V} N_v \log(2\pi).$$

Adding the complexity penalty in the target-filtered pooled score yields

$$S_{\text{tf-pool}}(G; \mathcal{D}, \mathcal{I}) = -\sup_{\theta_G} \ell_N(G, \theta_G) + \frac{1}{2} \log N \sum_{v \in V} (|\,\mathrm{Pa}_G(v)| + 1) + C_N,$$

where $C_N = -\frac{1}{2} \sum_{v \in V} N_v \log(2\pi)$ does not depend on G. Thus, $S_{\text{tf-pool}}$ is a decomposable criterion for the valid-environment structural-mechanism likelihood after profiling out environment-specific intercepts. By Assumption 1, the likelihood contributions associated with directly targeted replacement mechanisms are independent of the candidate parent sets of the original mechanisms and therefore do not affect graph comparison. ■

A.6. Proof of Theorem 8

Proof By Lemma 7, the target-filtered pooled score agrees, up to terms independent of the candidate DAG, with the negative maximized valid-environment structural-mechanism likelihood after profiling out environment-specific intercepts. Under known hard interventions, the omitted likelihood contributions of directly targeted replacement mechanisms are graph-independent nuisance terms. Hence, $L_\infty(G)$ is the population fit criterion induced by the graph-dependent part of the known-target interventional likelihood.

The true DAG $G^\star$ generates the observational and interventional distribution family under the known target family $\mathcal{I}$. Therefore, it attains the minimum value of L_∞, and $G^\star \in \mathcal{A}$.

Now let $H \in \mathcal{A}$. Since H attains the same minimum value of L_∞ as $G^\star$, the maximized population graph-dependent likelihood under H has no loss relative to the true known-target interventional distribution family. Under the regular finite-dimensional Gaussian model, a misspecified candidate model would incur a strictly positive population Kullback–Leibler loss. Hence, the true observational distribution is Markov with respect to H, and, for every intervention target $I_e \in \mathcal{I}$, the corresponding interventional distribution is Markov with respect to the intervention graph H^{I_e}. Thus, every $H \in \mathcal{A}$ is a causal independence map of the true known-target interventional distribution family.

Fix any $G \in \mathcal{A}_{\min}$. Since $G^\star \in \mathcal{A}$ and $\mathcal{A}_{\min}$ retains the sparsest members of $\mathcal{A}$, we have $|E_G| \le |E_{G^\star}|$. By the interventional faithfulness assumption, Hauser and Bühlmann (2015, Remark 1) show that the minimum causal independence maps of the true known-target interventional distribution family are exactly the DAGs in the $\mathcal{I}$-Markov equivalence class of $G^\star$. In particular, $G^\star$ is a minimum causal independence map, so no causal independence map of this distribution family has fewer edges than $G^\star$.

Since $G \in \mathcal{A}_{\min} \subseteq \mathcal{A}$, the preceding argument shows that G is itself a causal independence map. Therefore, $|E_G| \ge |E_{G^\star}|$.

Combining the two inequalities yields $|E_G| = |E_{G^\star}|$.

Hence, G is a minimum causal independence map of the true known-target interventional distribution family and $G \in \mathcal{G}^\star_{\mathcal{I}}$.

Since $G \in \mathcal{A}_{\min}$ was arbitrary, we conclude that

$$\mathcal{A}_{\min} \subseteq \mathcal{G}^\star_{\mathcal{I}}.$$

■

A.7. Proof of Corollary 9

Proof Let

$$\widehat{G}_N \in \operatorname*{arg\,min}_{G \in \mathcal{G}_{\mathrm{DAG}}} S_{\text{tf-pool}}(G; \mathcal{D}, \mathcal{I})$$

be an exact global minimizer of the empirical target-filtered pooled score. Define its empirical fit component by

$$R_N(G) = \sum_{v \in V} \frac{N_v}{2} \log \widehat{\sigma}^2_{v|\mathrm{Pa}_G(v)}.$$

Up to terms independent of G,

$$S_{\text{tf-pool}}(G;\mathcal{D},\mathcal{I}) = R_N(G) + \frac{1}{2}\log N \sum_{v\in V}(|\operatorname{Pa}_G(v)|+1)\,.$$

For every fixed DAG G, consistency of the valid-environment pooled residual variances and the assumption on the asymptotic environment sample proportions imply

$$\frac{1}{N}R_N(G) \xrightarrow{p} L_\infty(G).$$

Since V is fixed, $\mathcal{G}_{\text{DAG}}$ is finite, and this convergence holds simultaneously over all candidate DAGs. Consider any $G \notin \mathcal{A}$ and any $G_0 \in \mathcal{A}$. By the definition of $\mathcal{A}$, we have $L_\infty(G) > L_\infty(G_0)$. Therefore,

$$R_N(G) - R_N(G_0) = N\,(L_\infty(G) - L_\infty(G_0)) + o_p(N).$$

The difference in complexity penalties is of order $\log N$, whereas the fit difference is of order N. Hence,

$$S_{\text{tf-pool}}(G;\mathcal{D},\mathcal{I}) - S_{\text{tf-pool}}(G_0;\mathcal{D},\mathcal{I}) \xrightarrow{p} +\infty.$$

Since the DAG space is finite, it follows that $\Pr\left(\widehat{G}_N \in \mathcal{A}\right) \longrightarrow 1$.

It remains to distinguish among the members of $\mathcal{A}$. Let $G \in \mathcal{A} \setminus \mathcal{A}_{\min}$ and $G_0 \in \mathcal{A}_{\min}$. Since $G, G_0 \in \mathcal{A}$, both attain the same value of L_∞ and, by the proof of Theorem 8, are causal independence maps of the true known-target interventional distribution family. Under Assumptions 1, 3, and 4, the difference between their maximized valid-environment structural-mechanism likelihood contributions is $O_p(1)$ for the fixed-dimensional Gaussian conditional-mechanism models considered here. Equivalently,

$$R_N(G) - R_N(G_0) = O_p(1).$$

Since $G \notin \mathcal{A}_{\min}$ and $G_0 \in \mathcal{A}_{\min}$, we have $|E_G| > |E_{G_0}|$. Moreover,

$$\sum_{v\in V}(|\operatorname{Pa}_G(v)|+1) = |E_G| + |V|.$$

Hence, the complexity penalty difference between G and G_0 is $\frac{1}{2}\log N\,(|E_G| - |E_{G_0}|)$, which diverges to $+\infty$. Therefore,

$$\Pr(S_{\text{tf-pool}}(G;\mathcal{D},\mathcal{I}) > S_{\text{tf-pool}}(G_0;\mathcal{D},\mathcal{I})) \longrightarrow 1.$$

Again using finiteness of $\mathcal{G}_{\text{DAG}}$, we obtain $\Pr\left(\widehat{G}_N \in \mathcal{A}_{\min}\right) \longrightarrow 1$. By Theorem 8, we also have $\mathcal{A}_{\min} \subseteq \mathcal{G}^\star_{\mathcal{I}}$. Consequently, $\Pr\left(\widehat{G}_N \in \mathcal{G}^\star_{\mathcal{I}}\right) \longrightarrow 1$.

If the known target family fully identifies the true DAG, so that $\mathcal{G}^\star_{\mathcal{I}} = \{G^\star\}$, then

$$\Pr\left(\widehat{G}_N = G^\star\right) \longrightarrow 1.$$

■

Appendix B. Detailed Algorithmic Specifications

Algorithms 2 and 3 detail the reinsertion and warm-start parent-set updates invoked in Algorithm 1. For brevity, let

$$s_v := s_v^{\text{tf-pool}}, \qquad P := (P_x)_{x \in V}, \qquad \ell := (\ell_x)_{x \in V},$$

where P_x is the current parent set of node x and $\ell_x = s_x(P_x)$ is its current local score. We write

$$Q(\ell) := \sum_{x \in V} \ell_x$$

for the total score.

For a prefix change of node x, $+x$ means that x becomes admissible as a parent, whereas $-x$ means that x ceases to be admissible as a parent. To describe the prefix changes induced by an adjacent swap, define

$$\Delta(d, x) := \begin{cases} +x, & d = +1, \\ -x, & d = -1, \end{cases} \qquad d \in \{-1, +1\}.$$

Algorithm 2: Target-filtered reinsertion of node v

Input: Current order τ, parent sets P, local scores ℓ, node v
Output: Best reinsertion state $(\widehat{\tau}, \widehat{P}, \widehat{\ell})$

```
// Initialize the best state with the current state
```
$(\widehat{\tau}, \widehat{P}, \widehat{\ell}) \leftarrow (\tau, P, \ell)$

```
// Explore moving v to the left and to the right
```
foreach $d \in \{-1, +1\}$ **do**
  $(\tau', P', \ell') \leftarrow (\tau, P, \ell)$
  $j \leftarrow \text{position}_{\tau'}(v)$
  while $1 \le j + d \le p$ **do**
    $u \leftarrow \tau'_{j+d}$
    `// Swap` v `with the adjacent node` u
    $(\tau'_j, \tau'_{j+d}) \leftarrow (\tau'_{j+d}, \tau'_j)$
    `// Update only the two parent sets affected by the swap`
    $(P'_v, \ell'_v) \leftarrow \text{GrowShrink}_{\text{tf}}\big(v, \text{Pred}_{\tau'}(v), P'_v, \ell'_v, \Delta(d, u)\big)$
    $(P'_u, \ell'_u) \leftarrow \text{GrowShrink}_{\text{tf}}\big(u, \text{Pred}_{\tau'}(u), P'_u, \ell'_u, \Delta(-d, v)\big)$
    $j \leftarrow j + d$
    `// Retain the best reinsertion position found so far`
    if $Q(\ell') < Q(\widehat{\ell})$ **then**
      $(\widehat{\tau}, \widehat{P}, \widehat{\ell}) \leftarrow (\tau', P', \ell')$
    end
  end
end

return $(\widehat{\tau}, \widehat{P}, \widehat{\ell})$

Algorithm 3: Target-filtered warm-start grow–shrink

Function $\texttt{GrowShrink}_{\mathrm{tf}}(v, Z, P_0, \ell_0, \delta)$**:**

- **if** $\delta = +u$ **then**
  - $P \leftarrow P_0 \cup \{u\}$; $\ell \leftarrow s_v(P)$
  - **if** $\ell \geq \ell_0$ **then**
    - **return** P_0, ℓ_0
  - **end**
- **else if** $\delta = -u$ **then**
  - **if** $u \notin P_0$ **then**
    - **return** P_0, ℓ_0
  - **end**
  - $P \leftarrow P_0 \setminus \{u\}$; $\ell \leftarrow s_v(P)$
- **else**
  - $P \leftarrow \emptyset$; $\ell \leftarrow s_v(\emptyset)$
- **end**
- **foreach** $d \in \{+, -\}$ **do**
  - **repeat**
    - $P^{\mathrm{old}} \leftarrow P$
    - $\mathcal{U} \leftarrow \begin{cases} Z \setminus P^{\mathrm{old}}, & d = +, \\ P^{\mathrm{old}}, & d = -; \end{cases}$
    - **foreach** $u \in \mathcal{U}$ **do**
      - **if** $d = +$ **then**
        - $P' \leftarrow P \cup \{u\}$
      - **end**
      - **else**
        - **if** $u \notin P$ **then**
          - **continue**
        - **end**
        - $P' \leftarrow P \setminus \{u\}$
      - **end**
      - $\ell' \leftarrow s_v(P')$
      - **if** $\ell' < \ell$ **then**
        - $(P, \ell) \leftarrow (P', \ell')$
      - **end**
    - **end**
  - **until** $P = P^{\mathrm{old}}$
- **end**
- **return** P, ℓ

The $\delta = +u$ short-circuit (returning P_0, ℓ_0 immediately when adding u does not improve the score) is justified whenever P_0 is itself already a grow–shrink fixed point over $Z \setminus \{u\}$, which holds at every call site in Algorithm 2: by Proposition 4, if no single admissible addition to such a P_0 improves the score, no further local move can either, so the full grow–shrink procedure would return P_0 unchanged regardless.

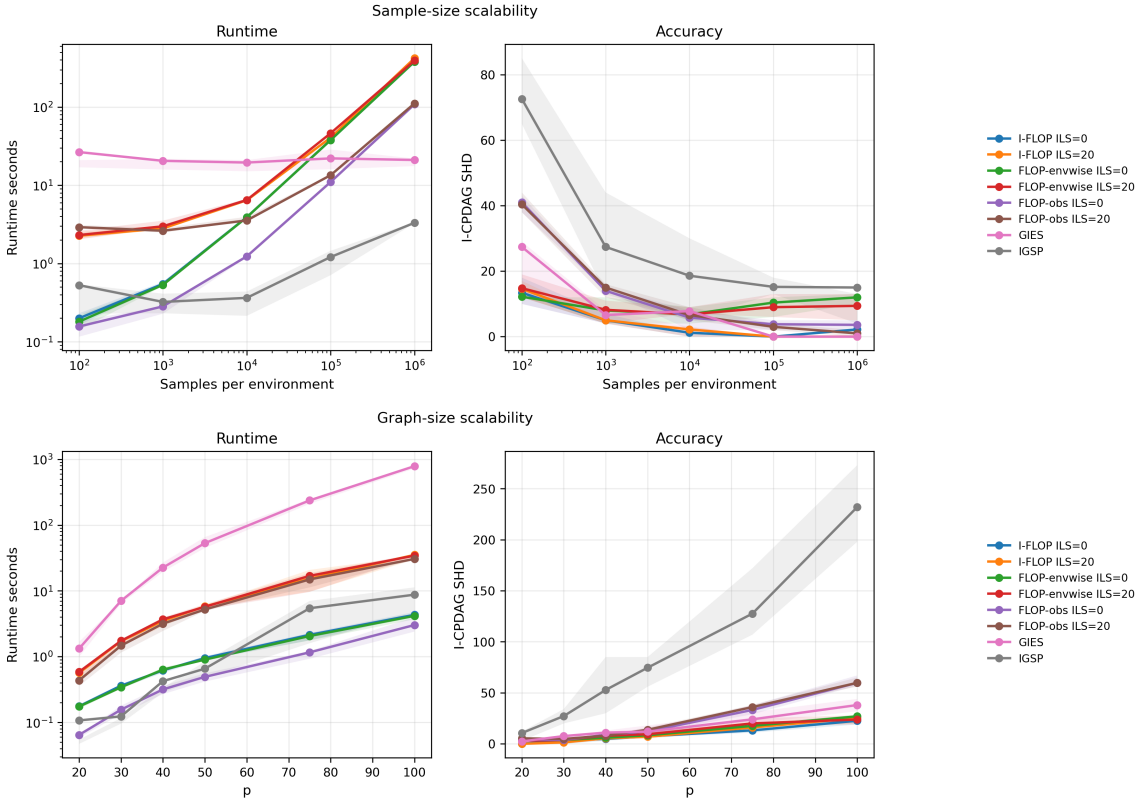


Figure 4: Top: Scaling with the number of samples per environment, with $p = 40$, expected degree 3, and sample size varied from 10^2 to 10^6 in each environment. *Bottom:* Scaling with graph size, with expected degree 3, 1000 samples per environment, and p varied from 20 to 100. Each setting contains one observational environment and five known-target singleton perfect hard-intervention environments. Runtime is shown on a logarithmic scale, and accuracy is measured by I-CPDAG SHD.

Appendix C. Computational Scaling with Graph Size and Sample Size

Figure 4 reports additional scaling results on sparse ER DAGs. We examine two complementary regimes: increasing the number of samples per environment while holding the graph regime fixed, and increasing the number of variables while fixing the sample size, expected degree, and intervention design.

Scaling with sample size. As the number of samples per environment increases, I-FLOP substantially reduces its I-CPDAG SHD, approaching near-perfect recovery. GIES exhibits a comparable reduction in structural error while maintaining nearly constant runtime in this experiment. By contrast, the runtime of the FLOP-based methods increases markedly with sample size, particularly beyond 10^4 samples per environment. FLOP-envwise does not match the accuracy of I-FLOP at large sample sizes, indicating that environment-

wise estimation alone is insufficient without target-based filtering. I-GSP improves with additional samples but retains noticeably higher structural error than I-FLOP and GIES.

Scaling with graph size. When the number of variables increases, I-FLOP maintains the lowest structural error among the evaluated methods while remaining computationally inexpensive relative to GIES. Its runtime grows moderately from $p = 20$ to $p = 100$, whereas GIES becomes substantially more expensive, reaching several hundred seconds at the largest graph size. The accuracy difference is also pronounced for larger graphs: I-FLOP retains comparatively low I-CPDAG SHD, while the errors of GIES, FLOP-obs, and especially I-GSP increase considerably. FLOP-envwise remains computationally close to I-FLOP but incurs consistently higher recovery error, again supporting the contribution of target-based environment filtering.

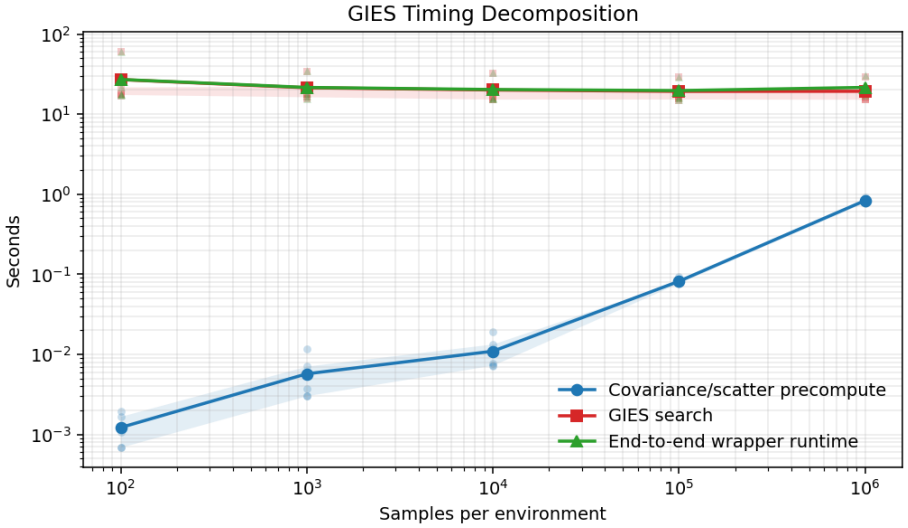


Figure 5: Runtime of GIES is decomposed into covariance/scatter precomputation, graph search, and end-to-end wrapper execution. All the settings are the same as Figure 4.

Explanation of the runtime of GIES. The nearly flat runtime curve of GIES in the sample-size experiment is explained by the timing decomposition in Figure 5. GIES performs its graph search using covariance computation. As the number of samples per environment increases from 10^2 to 10^6, the cost of computing these statistics rises from the millisecond scale to below one second. However, the search stage consistently requires approximately 20 seconds and therefore dominates the end-to-end runtime throughout the evaluated range. End-to-end wrapper execution, covering input conversion into the GIES implementation's required format and output conversion back into a standardized DAG representation for evaluation, contributes only a small, roughly constant overhead relative to the search stage. Consequently, increasing the sample size has little visible effect on total GIES runtime in Figure 4. The small non-monotonic variation at lower sample sizes is

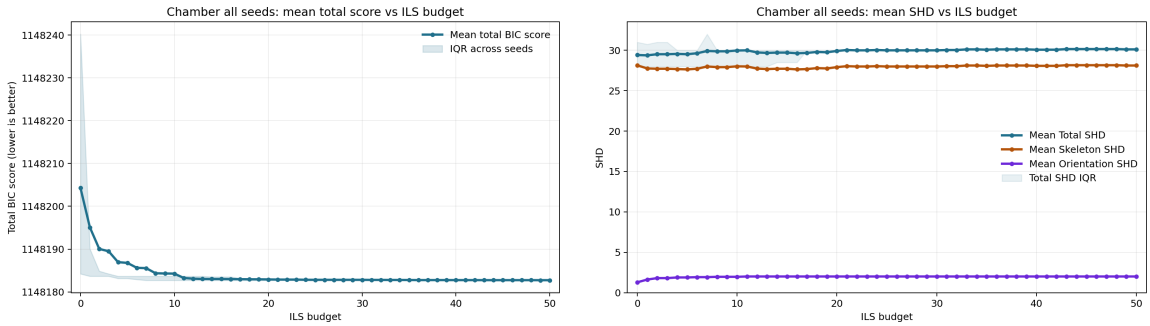


Figure 6: **Left:** Mean optimized total BIC score across 50 seeds as the ILS budget increases from 0 to 50. **Right:** Mean total SHD and its skeleton and orientation components across the same seeds. The shaded regions denote the interquartile range across seeds.

attributable to changes in the estimated scores, which will alter the greedy search trajectory and its number of steps, rather than to a systematic reduction in computational cost.

Appendix D. Diagnostic ILS Trace on the Real Data Experiment

To investigate the observed mismatch between score improvement and graph recovery as the ILS budget increases, we conducted an additional diagnostic analysis on the Chamber dataset from the real-data experiment (Section 4.3). This analysis follows exactly the same experimental setting as the original real-data evaluation and examines how the optimized total BIC score and the corresponding SHD change as the ILS budget increases.

As shown in Figure 6 (left), the mean total score decreases rapidly during the first several ILS steps and then gradually stabilizes. From ILS 0 to ILS 50, the mean total BIC score decreases by approximately 21.55 units, and 49 of the 50 seeds achieve a lower score at ILS 50 than at ILS 0. However, this improvement in the optimized total score does not translate into better graph recovery. Figure 6 (right) shows that the mean total SHD increases from 29.42 at ILS 0 to 30.12 at ILS 50. This increase is almost entirely attributable to the orientation component: the mean skeleton SHD remains essentially unchanged, decreasing slightly from 28.14 to 28.12, whereas the mean orientation SHD increases from 1.28 to 2.00.

However, the aggregate results mask substantial heterogeneity across seeds. As shown in Figure 7, comparing ILS 0 with ILS 50, total SHD worsens for 27 seeds, improves for 14 seeds, and remains unchanged for 9 seeds. Skeleton SHD increases for 20 seeds and decreases for 20 seeds, whereas orientation SHD increases for 20 seeds and decreases for none. Although skeleton changes constitute the larger component-wise changes for 33 seeds, they occur in both directions and therefore largely cancel in the aggregate mean. By contrast, orientation changes are one-sided and consequently account for the net increase in mean total SHD. Thus, in the real data experiment, increasing the ILS budget generally yields lower-scoring solutions, but these solutions are not necessarily closer to the reference graph under SHD.

This score–recovery mismatch is plausibly related to a mismatch between the assumptions of the I-FLOP algorithm and the data-generating process of the Causal Chamber

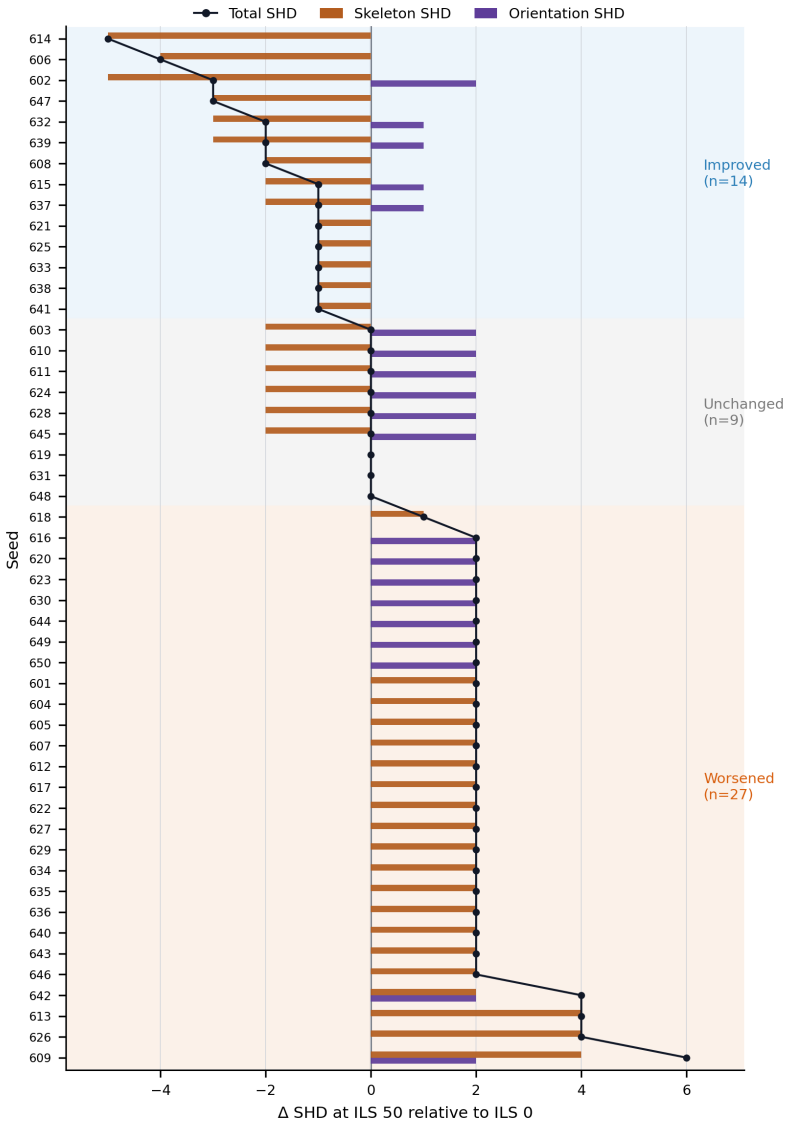


Figure 7: Seed-level changes in SHD on the real data dataset when increasing the ILS budget from 0 to 50. The horizontal axis reports the change in SHD, computed as the value at ILS 50 minus that at ILS 0. Seeds are grouped according to the change in total SHD: improved ($n = 14$), unchanged ($n = 9$), and worsened ($n = 27$). The black line denotes the change in total SHD, while the orange and purple bars denote the changes in skeleton SHD and orientation SHD, respectively.

dataset. In the synthetic experiments, the data are generated from a linear-Gaussian SEM, and interventions are implemented as Gaussian shifts on target variables while preserving the same linear mechanisms for the remaining variables. However, Causal Chamber is obtained from a real optical system rather than from this synthetic SEM. Its intervention

environments are created by changing the operating ranges of physical actuators to bounded intervention-specific ranges. These interventions therefore alter the support and operating regime of the intervened variables, rather than merely applying Gaussian mean shifts within the same assumed model.

These differences affect the interpretation of the optimized score. The I-FLOP score evaluates candidate graphs through linear-Gaussian regressions pooled across the relevant observational and interventional environments. Consequently, a lower total BIC score on the real data experiment represents improved fit under the assumed linear-Gaussian scoring approximation, rather than direct evidence of improved recovery of the physical reference graph. The ILS results suggest that increasing the search budget improves the optimized score, while the corresponding solutions retain essentially unchanged skeleton accuracy and accumulate additional orientation errors.