Towards Structural Causal Bandits with Non-Manipulable Variables and Unknown Causal Structure

Iason Skylitsis[*]
University of Amsterdam
iason.skylitsis@student.uva.nl

Roel Hulsman
University of Amsterdam
r.p.hulsman@uva.nl

Abstract

Sequential decision-making problems, such as clinical trials or online advertising, require agents to learn effective actions while balancing exploration and exploitation. The Structural Causal Bandit (SCM-MAB) framework improves exploration efficiency by leveraging causal structure to prune the action space, but assumes either access to a fully specified causal graph or that all observed variables are manipulable—requirements rarely met in practice. This work integrates causal discovery with structural causal bandits to handle unknown causal graphs and non-manipulable variables. Specifically, we combine constraint-based causal discovery (PC and FCI algorithms) with existing methods for computing Possibly-Optimal Minimal Intervention Sets (POMISs), enumerating graphs in the learned equivalence class and handling non-manipulable variables through latent projection. We evaluate the framework on synthetic benchmarks spanning causally sufficient and insufficient settings. Our experiments show that bandit performance depends critically on the quality of the learned causal structure: when key dependencies are recovered, performance matches oracle baselines with access to the true graph, but degrades otherwise. These results highlight both the potential and the challenges of integrating causal discovery with causal bandits under structural uncertainty.

1 Introduction

Sequential decision-making under uncertainty is prevalent in numerous real-world domains. In medication dosing, for instance, a physician must determine the appropriate dosage for incoming patients, where the optimal amount depends on each patient's genetic profile and medical records (Bastani and Bayati, 2020). Incorrect dosing can lead to adverse consequences such as stroke or bleeding, making it critical to learn effective dosing strategies quickly. This illustrates the fundamental exploration-exploitation trade-off: the agent must gather information about the effects of different actions while simultaneously maximizing cumulative reward over time.

The multi-armed bandit framework provides a principled approach to such problems (Robbins, 1952), and algorithms such as Upper Confidence Bound (UCB) and Thompson Sampling offer provable guarantees on cumulative regret (Auer et al., 2002; Thompson, 1933). However, these guarantees rely on the assumption that the rewards of different arms are independent. In many real-world settings, pulling an arm corresponds to intervening on a set of variables, and the resulting rewards are governed by the underlying causal mechanisms of the environment. When arms exhibit such dependencies, standard regret guarantees no longer hold. The *structural causal bandit* (SCM-MAB) framework introduced by Lee and Bareinboim (2018) addresses this by integrating structural causal models with bandit algorithms, leveraging the causal graph to prune the action space and improve sample efficiency.

[*]Corresponding author: iason.skylitsis@student.uva.nl

However, the original SCM-MAB formulation relies on two strong assumptions that restrict its applicability: (1) that every variable in the system is directly manipulable, and (2) that the true causal directed acyclic graph (DAG) is known a priori. In real-world domains such as medical trials (Bastani and Bayati, 2020; Durand et al., 2018) or finance (Shen et al., 2015; Huo and Fu, 2017), agents can typically intervene on only a subset of variables, and the causal structure is often unknown. Even when causal structure can be learned from data, it is generally identifiable only up to a Markov Equivalence Class (MEC) rather than a unique DAG. Recent work addresses these limitations individually. Lee and Bareinboim (2019) extend the SCM-MAB framework to handle non-manipulable variables through projection-based methods. Park et al. (2025) take a different approach by extending SCM-MAB to accept a MEC as input, developing graphical criteria for identifying *possibly-optimal minimal intervention sets* (POMISs) directly from maximal ancestral graphs (MAGs) and partial ancestral graphs (PAGs). However, their work assumes the equivalence class is given rather than learned from data.

In this work, we take a first step toward learning POMISs while allowing for non-manipulable variables and learning the unknown causal structure from data. We implement and combine the theoretical results of Lee and Bareinboim (2018) and Lee and Bareinboim (2019). To handle structural uncertainty, we enumerate the member graphs of an equivalence class learned via causal discovery and take the union of the resulting POMISs. This enumeration-based approach serves as a proof-of-concept, though it does not scale to large equivalence classes. Integrating the more efficient MEC-based methods of Park et al. (2025) with causal discovery remains an interesting direction for future work.

The primary contribution of this work is empirical. We evaluate our approach on synthetic benchmark structural causal models, including scenarios with non-manipulable variables, spanning both causally sufficient and insufficient settings. We compare against an oracle baseline that has access to the true causal graph, which represents an upper bound on performance since it uses strictly more information than our method. Our experiments show that when causal discovery successfully recovers key dependencies, our method achieves cumulative regret comparable to this oracle. However, performance degrades when discovery fails to recover critical edges, highlighting both the promise and limitations of integrating causal discovery with structural causal bandits.

Contributions. This work makes the following contributions:

1. We propose a method that integrates causal discovery with structural causal bandits to handle unknown causal graphs and non-manipulable variables, combining existing theoretical results with an enumeration-based approach over learned equivalence classes.
2. We empirically evaluate this method on synthetic benchmarks spanning causally sufficient and insufficient settings, showing that performance matches oracle baselines when discovery recovers key dependencies, but degrades otherwise.
3. We provide an open-source implementation to enable reproducibility and future research.[2]

2 Background

2.1 Multi-Armed Bandits (MAB)

The multi-armed bandit problem is a sequential decision-making problem in which an agent repeatedly chooses among K actions (arms) and observes a corresponding reward. Each arm a yields rewards from an unknown probability distribution with some expected value μ_a. The objective is to minimize the cumulative regret, defined over a horizon of T rounds as:

$$\mathrm{Reg}_T = T\mu^* - \sum_{t=1}^{T} \mathbb{E}[Y_{a_t}], \tag{1}$$

where a_t is the arm played at time t, Y_{a_t} is the observed reward, and $\mu^* = \max_{1 \leq a \leq K} \mu_a$ is the maximum expected reward.

[2] `https://github.com/iasonsky/causal-discovery-bandits`

The problem centers on the exploration-exploitation trade-off: the agent must balance exploiting arms with historically high rewards and exploring lesser-known options to gather information. Standard algorithms for managing this trade-off include Upper Confidence Bound (UCB) (Auer et al., 2002; Cappé et al., 2013), Thompson sampling (Thompson, 1933) and ϵ-greedy (Sutton and Barto, 2018). However, these algorithms assume independent arm rewards. When arms share underlying causal structure, particularly involving unobserved confounders, this assumption is violated, and standard guarantees may no longer hold. For instance, Lee and Bareinboim (2018) show that intervening on all variables simultaneously may exclude the optimal arm when a smaller intervention set is optimal, leading to linear regret.

2.2 Structural Causal Models (SCMs)

Structural causal models provide a language for describing systems of cause and effect. An SCM $\mathcal{M} = \langle \mathbf{U}, \mathbf{V}, \mathbf{F}, P(\mathbf{U}) \rangle$, consists of a set of exogenous variables $\mathbf{U}$, a set of endogenous variables $\mathbf{V} = \{V_1, \ldots, V_n\}$, a set of structural functions $\mathbf{F} = \{f_i\}_{i=1}^n$, and a distribution $P(\mathbf{U})$ over the exogenous variables. Each $V_i \in \mathbf{V}$ is determined by a structural assignment $V_i \leftarrow f_i(pa(V_i), \mathbf{U}^i)$, where $pa(V_i) \subseteq \mathbf{V}$ denotes the parents of V_i and $\mathbf{U}^i \subseteq \mathbf{U}$ represent exogenous noise. Each structural assignment describes a causal mechanism for how V_i takes its value. An SCM induces both observational and interventional distributions. The effect of an intervention $do(\mathbf{X} = \mathbf{x})$ on outcome Y is given by $P(Y = y \mid do(\mathbf{x}))$, denoted as $P_{\mathbf{x}}(y)$.

Each SCM is associated with a causal graph $\mathcal{G} = \langle \mathbf{V}, \mathbf{E} \rangle$, where nodes $\mathbf{V}$ correspond to endogenous variables and edges $\mathbf{E}$ encode causal relationships. Directed edges $V_i \rightarrow V_j$ indicate that V_i is a direct parent of V_j in the structural function f_j, while bidirected edges $V_i \leftrightarrow V_j$ represent the presence of an unobserved confounder (UC) in $\mathbf{U}$ that simultaneously influences both V_i and V_j. A causal model is said to be causally sufficient if there are no unobserved confounders among the observed variables. We extend the parent notation $pa(V_i)$ to children $ch(V_i)$, ancestors $an(V_i)$, and descendants $de(V_i)$. When including the argument itself, we write $Pa(V_i) = pa(V_i) \cup \{V_i\}$, and analogously for Ch, An, and De. For a set of variables $\mathbf{W} \subseteq \mathbf{V}$, these operators are applied element-wise and then unioned, e.g. $An(\mathbf{W}) = \bigcup_{W \in \mathbf{W}} An(W)$.

Interventions modify the structural equations of the SCM. For hard interventions, which set variables to fixed values and are the focus of this work, this corresponds to removing incoming edges in the causal graph. For $\mathbf{X} \subseteq \mathbf{V}$, the intervention graph $\mathcal{G}_{\overline{\mathbf{X}}}$ is obtained by removing all edges pointing into $\mathbf{X}$, reflecting that the values of $\mathbf{X}$ are set externally. The interventional distribution $P(y \mid do(\mathbf{x}))$ respects the topology of $\mathcal{G}_{\overline{\mathbf{X}}}$. Do-calculus (Pearl, 1995) provides rules for connecting observational and interventional distributions according to d-separation relations in the causal graph.

2.3 Structural Causal Bandits (SCM-MAB Framework)

Lee and Bareinboim (2018) introduced the structural causal bandit (SCM-MAB) framework, which connects causal models with bandit algorithms. In this setting, each arm corresponds to an intervention $do(\mathbf{X} = \mathbf{x})$ on variables in the causal graph, and the reward distribution is given by $P(y \mid do(\mathbf{x}))$. To exploit dependencies among arms induced by the causal structure, Lee and Bareinboim (2018) introduced minimal intervention sets (MIS) and possibly-optimal minimal intervention sets (POMIS), which prune the action space while ensuring the optimal intervention is not excluded.

Minimal Intervention Sets (MIS). A set $\mathbf{X} \subseteq \mathbf{V} \setminus \{Y\}$ is a minimal intervention set if no strict subset yields the same expected reward across all models consistent with $\mathcal{G}$, i.e., $\mu_{\mathbf{x}'} = \mu_{\mathbf{x}} = \mathbb{E}[Y \mid do(\mathbf{x})]$ for every strict subset $\mathbf{X}' \subset \mathbf{X}$. In other words, $\mathbf{X}$ contains no redundant variables, and removing any element would change its causal effect on Y. Graphically, $\mathbf{X}$ is a MIS if and only if every variable in $\mathbf{X}$ remains an ancestor of Y after intervening on $\mathbf{X}$, i.e., $\mathbf{X} \subseteq an(Y)_{\mathcal{G}_{\overline{\mathbf{X}}}}$ (Lee and Bareinboim, 2018, Proposition 1). An algorithm for enumerating all MISs given $\mathcal{G}$ and Y is provided in Lee and Bareinboim (2018, Appendix A, Algorithm 3). We denote the set of all minimal intervention sets with respect to $\mathcal{G}$ and Y by $\mathbb{M}_{\mathcal{G},Y}$.

Possibly-Optimal MIS (POMIS). An $\mathbf{X} \in \mathbb{M}_{\mathcal{G},Y}$ is a possibly-optimal minimal intervention set if there exists an SCM consistent with $\mathcal{G}$ in which $\mathbf{X}$ achieves the highest expected reward, i.e., $\mu_{\mathbf{x}^*} = \max_{\mathbf{X}' \in \mathbb{M}_{\mathcal{G},Y}} \mu_{\mathbf{x}'}$. Lee and Bareinboim (2018, Theorem 6) provide a graphical characterization based on the interventional border (IB) and minimal unobserved confounders' territory (MUCT), and an algorithm for enumerating all POMISs given $\mathcal{G}$ and Y (Lee and Bareinboim, 2018, Algorithm 1). We use this algorithm as a building block in our method. We denote the set of all POMISs with respect to $\mathcal{G}$ and Y by $\mathbb{P}_{\mathcal{G},Y}$. Crucially, interventions that are not POMISs can be safely discarded from the action space, as they are guaranteed to be suboptimal under every SCM consistent with $\mathcal{G}$.

Non-Manipulability and Latent Projection. Lee and Bareinboim (2019) extended the framework to handle non-manipulable variables $\mathbf{N} \subseteq \mathbf{V} \setminus \{Y\}$, variables that cannot be directly intervened upon. For MISs, the constrained sets are simply a subset of the unconstrained ones: $\mathbb{M}^{\mathbf{N}}_{\mathcal{G},Y} = \{\mathbf{W} \in \mathbb{M}_{\mathcal{G},Y} \mid \mathbf{W} \cap \mathbf{N} = \emptyset\}$ (Lee and Bareinboim, 2019). For POMISs, however, this does not hold, and latent projection is required. They employed latent projection (Verma and Pearl, 1990) to produce an acyclic directed mixed graph (ADMG) $\mathcal{H} = \mathcal{G}[\mathbf{V} \setminus \mathbf{N}]$ over only the manipulable variables. An ADMG may contain both directed and bidirected edges. Crucially, the POMISs are invariant under projection: $\mathbb{P}^{\mathbf{N}}_{\mathcal{G},Y} = \mathbb{P}_{\mathcal{H},Y}$ (Lee and Bareinboim, 2019, Theorem 4).

2.4 Causal Discovery and Equivalence Classes.

Constraint-based causal discovery algorithms recover Markov equivalence classes of causal graphs from data. The PC algorithm (Spirtes et al., 2000) applies to causally sufficient settings and outputs a completed partially directed acyclic graph (CPDAG), representing all DAGs with the same conditional independences. The FCI algorithm (Spirtes et al., 1995) applies to causally insufficient settings and outputs a partial ancestral graph (PAG), representing an equivalence class of ADMGs.

3 Method

Our method takes as input observational data, a reward variable Y, and optionally a set of non-manipulable variables $\mathbf{N}$, and outputs a reduced set of intervention sets. These intervention sets define the action space for a bandit algorithm, which learns the optimal intervention through sequential experimentation.

The method proceeds as follows. First, we learn an equivalence class of causal structures from the observational data using constraint-based causal discovery (Section 2.4). We then enumerate all graphs in the equivalence class and compute POMISs or MISs for each graph using the algorithms from Lee and Bareinboim (2018). When non-manipulable variables are present, we apply latent projection (Section 2.3) before POMIS identification, while MISs can be obtained by filtering out sets containing variables in $\mathbf{N}$. Finally, the union of POMISs (or MISs) across all graphs in the equivalence class defines the action space, ensuring that the optimal intervention is included whenever the true graph is a member of the equivalence class.

3.1 Pipeline

The proposed framework integrates causal discovery, latent projection, and intervention optimization into a single causal bandit pipeline, summarized in Figure 1. We introduce four methods that vary along two dimensions: intervention set type (POMIS vs. MIS) and manipulability constraints (all variables manipulable vs. some non-manipulable). The methods – CD-POMIS, CD-MIS, CD-POMIS-NM, and CD-MIS-NM – follow the same six-stage pipeline, with Step 3 (latent projection) applied only for CD-POMIS-NM.

1. Causal Discovery Given observational data, the framework first learns a causal structure using either the PC or FCI algorithm from the `causal-learn` package (Zheng et al., 2024). PC assumes causal sufficiency and outputs a CPDAG, while FCI accounts for latent confounding and outputs a PAG. These graphs represent equivalence classes of DAGs or ADMGs, respectively.

2. Equivalence-Class Enumeration. From the discovered CPDAG or PAG, we enumerate the set of all fully oriented graphs that are consistent with the observed conditional independences. Let $\mathcal{E} = \{\mathcal{G}^{(1)}, \ldots, \mathcal{G}^{(K)}\}$ denote this equivalence class, where each $\mathcal{G}^{(k)}$ represents a DAG (for CPDAGs) or an ADMG (for PAGs). For CPDAGs, enumeration is performed using the `pdag2allDags` procedure from the `pcalg` package (Kalisch et al., 2025). For PAGs, we use `pag2admg` (Hyttinen et al., 2017). For each $\mathcal{G}^{(k)} \in \mathcal{E}$, we initialize the graph $\mathcal{H}^{(k)} = \mathcal{G}^{(k)}$.

3. Latent Projection for Non-Manipulable Variables (CD-POMIS-NM Only). When non-manipulable variables are present ($\mathbf{N} \neq \emptyset$), CD-POMIS-NM applies latent projection to each graph:

$$\mathcal{H}^{(k)} \leftarrow \mathcal{G}^{(k)}[\mathbf{V} \setminus \mathbf{N}].$$

This step is required for POMIS identification because constrained POMISs are not a subset of unconstrained POMISs (Section 2.3). For MIS variants, this step is skipped since constrained MISs can be obtained by filtering.

4. POMIS and MIS Identification per Graph. For each $\mathcal{H}^{(k)}$, we compute either POMISs or MISs using the algorithms described in Section 2.3. Specifically, we apply Algorithm 1 from Lee and Bareinboim (2018) for POMIS identification or Algorithm 3 for MIS enumeration, with input graph $\mathcal{H}^{(k)}$ and reward variable Y, obtaining $\mathbb{P}_{\mathcal{H}^{(k)},Y}$ or $\mathbb{M}_{\mathcal{H}^{(k)},Y}$ respectively.

5. Aggregation across the Equivalence Class. Because the true causal structure is uncertain, the framework aggregates the per-graph results across all members of the equivalence class. The sets of POMISs and MISs are obtained as

$$\mathbb{P}^{\mathbf{N}}_{\mathcal{E},Y} = \bigcup_{k=1}^{K} \mathbb{P}_{\mathcal{H}^{(k)},Y}, \qquad \mathbb{M}^{\mathbf{N}}_{\mathcal{E},Y} = \bigcup_{k=1}^{K} \mathbb{M}_{\mathcal{H}^{(k)},Y},$$

For CD-MIS-NM, the aggregated MISs are further filtered to exclude any sets containing variables in $\mathbf{N}$. This aggregation ensures that no potentially optimal intervention is excluded under any causal structure consistent with the discovered graph.

6. Bandit Learning. Each intervention $do(\mathbf{X} = \mathbf{x})$ for $\mathbf{X} \in \mathbb{P}^{\mathbf{N}}_{\mathcal{E},Y}$ (for POMIS variants) or $\mathbf{X} \in \mathbb{M}^{\mathbf{N}}_{\mathcal{E},Y}$ (for MIS variants) is treated as an arm in a causal bandit problem. Standard solvers such as Thompson Sampling or UCB are then employed to minimize cumulative regret by sequentially exploring and exploiting these interventions.

4 Results

4.1 Experimental Setup

We evaluate our proposed framework on a suite of benchmark structural causal models (SCMs) adapted from Lee and Bareinboim (2018, 2019). Our experiments demonstrate the impact of structural uncertainty and non-manipulability constraints on intervention selection and regret performance.

SCMs. We consider five representative SCMs covering both causally sufficient and causally insufficient settings. All SCMs use binary variables with XOR or OR functional relationships. We generate $N = 10{,}000$ observational samples for causal discovery. The ground-truth and discovered graphs are presented in Appendix B, while the complete structural definitions are provided in Appendix A.

- **Causally Sufficient Setting:** We evaluate two causally sufficient SCMs. The **Simple Markovian SCM** (Lee and Bareinboim, 2018, Appendix D, Task 1) is a five-variable system (OR-based) used to validate the basic causal bandit setup. The **Chain SCM** is a three-variable variant (OR-based) designed to yield a non-trivial CPDAG with multiple Markov-equivalent DAGs. We evaluate two variants of the Chain SCM—with and without enforcing Y as a sink—to examine how structural constraints affect bandit performance.
- **Causally Insufficient Setting:** The **Instrumental Variable (IV) SCM** (Lee and Bareinboim, 2018, Appendix D, Task 2) is a three-variable model with an unobserved confounder ($X \leftrightarrow Y$). We evaluate both the original XOR-based formulation and a modified OR-based variant that yields stronger statistical dependencies for structure discovery.
- **Non-Manipulable Causally Insufficient Setting:** The **Frontdoor SCM** (Lee and Bareinboim, 2019, Appendix) is a three-variable model (XOR-based) with $\mathbf{N} = \{Z\}$. The **Four-Variable SCM** (Lee and Bareinboim, 2019, Fig. 2a) has $\mathbf{N} = \{A\}$. We evaluate both the original XOR-based formulation and an OR-based variant to compare discovery performance under identical graph topology.

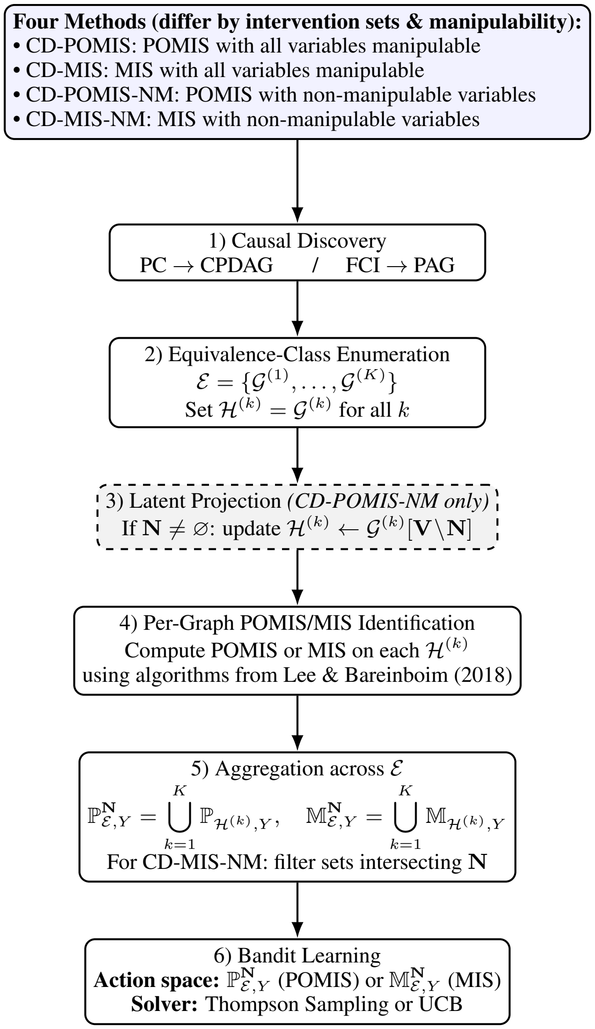

Four Methods (differ by intervention sets & manipulability):
• CD-POMIS: POMIS with all variables manipulable
• CD-MIS: MIS with all variables manipulable
• CD-POMIS-NM: POMIS with non-manipulable variables
• CD-MIS-NM: MIS with non-manipulable variables
1) Causal Discovery
PC → CPDAG / FCI → PAG
2) Equivalence-Class Enumeration
$\mathcal{E} = \{\mathcal{G}^{(1)}, \ldots, \mathcal{G}^{(K)}\}$
Set $\mathcal{H}^{(k)} = \mathcal{G}^{(k)}$ for all k
3) Latent Projection *(CD-POMIS-NM only)*
If $\mathbf{N} \neq \varnothing$: update $\mathcal{H}^{(k)} \leftarrow \mathcal{G}^{(k)}[\mathbf{V} \backslash \mathbf{N}]$
4) Per-Graph POMIS/MIS Identification
Compute POMIS or MIS on each $\mathcal{H}^{(k)}$
using algorithms from Lee & Bareinboim (2018)
5) Aggregation across $\mathcal{E}$
$\mathbb{P}^{\mathbf{N}}_{\mathcal{E},Y} = \bigcup_{k=1}^{K} \mathbb{P}_{\mathcal{H}^{(k)},Y}, \quad \mathbb{M}^{\mathbf{N}}_{\mathcal{E},Y} = \bigcup_{k=1}^{K} \mathbb{M}_{\mathcal{H}^{(k)},Y}$
For CD-MIS-NM: filter sets intersecting $\mathbf{N}$
6) Bandit Learning
Action space: $\mathbb{P}^{\mathbf{N}}_{\mathcal{E},Y}$ (POMIS) or $\mathbb{M}^{\mathbf{N}}_{\mathcal{E},Y}$ (MIS)
Solver: Thompson Sampling or UCB

Figure 1: Overview of the proposed causal bandit pipeline. The framework integrates causal discovery, equivalence-class enumeration, optional latent projection for non-manipulable variables (Step 3, applied only for CD-POMIS-NM), per-graph POMIS/MIS identification, aggregation across the equivalence class, and bandit learning. The four methods differ in two dimensions: intervention set type (POMIS vs. MIS) and whether non-manipulable variables are present ($\mathbf{N} = \varnothing$ vs. $\mathbf{N} \neq \varnothing$).

Causal Discovery. For both PC and FCI we employ the `chisq` independence test with significance level $\alpha = 0.01$, since our data is binary. The discovered graphs, represented as CPDAGs (for PC) or PAGs (for FCI), are then used to enumerate all consistent DAGs or ADMGs in the corresponding equivalence class. These structures form the basis for POMIS/MIS computation and subsequent bandit learning.

Bandit Learning. Following Lee and Bareinboim (2018), each intervention $do(\mathbf{X} = \mathbf{x})$ is treated as a bandit arm. When all variables are manipulable, we compare POMIS/MIS computed from the ground-truth graph against CD-POMIS/CD-MIS computed from the discovered graph. When non-manipulable variables are present, we compare POMIS-NM/MIS-NM computed from the ground-truth graph against CD-POMIS-NM/CD-MIS-NM computed from the discovered graph. All interventions correspond to hard *do*-operations. Bandit experiments are conducted with a horizon of $T = 2{,}000$ and averaged over 50 independent trials. Both Thompson Sampling and UCB solvers are evaluated for regret minimization.

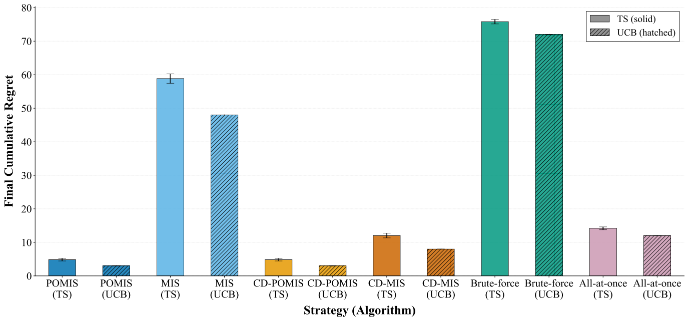


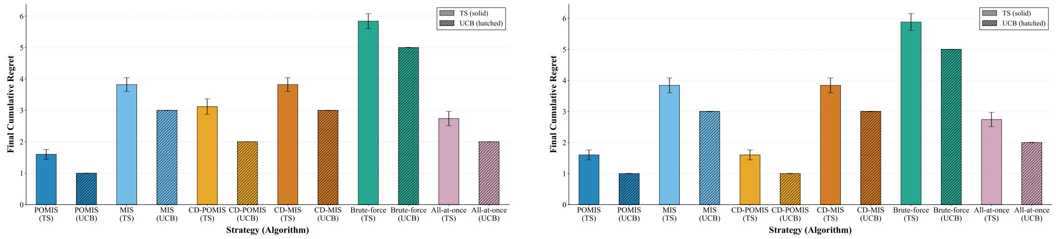
Figure 2: Final regret comparison at $T = 2{,}000$ for the Simple Markovian SCM. Full regret progression is shown in Fig. 12

Evaluation Metrics. We evaluate bandit performance using **cumulative regret**, which measures the difference between the optimal reward and the actual reward accumulated over time. Regret is computed relative to the oracle optimal intervention under the true SCM. Our primary metric is the **final cumulative regret** (i.e., regret at the final timestep), averaged across runs, which provides a quantitative summary for comparing methods across different SCMs. Lower final regret indicates that the algorithm identified and exploited the optimal intervention more frequently. We also plot the full **regret progression** over time in Appendix C to illustrate convergence behavior, analogous to the cumulative regret plots in Lee and Bareinboim (2018). To assess the quality of causal discovery, we measure the **Structural Hamming Distance (SHD)** between the discovered and ground-truth equivalence-class graphs (CPDAG or PAG). For CPDAGs in causally sufficient scenarios, we follow Tsamardinos et al. (2006) and compute SHD as the sum of edge insertions, deletions, and flips required to transform the discovered graph into the true graph. For PAGs in causally insufficient scenarios, we adopt the metric from Jabbari et al. (2017), counting each incorrect edge mark (e.g., arrowhead vs. circle vs. tail) as one error, and each spurious or missing edge as two errors (one for each endpoint). Lower SHD values indicate more accurate structure learning. We also report the **equivalence class size** (|MEC|), which determines the number of graphs enumerated for POMIS computation. In our experiments, |MEC| remains small (at most 6), though this enumeration-based approach may not scale to larger equivalence classes.

4.2 Markovian Setting: From Simple to Chain SCMs

We begin by validating our method on the **Simple Markovian SCM** introduced by Lee and Bareinboim (2018, Appendix D, Task 1). While this model reproduces the basic causal bandit setting, its CPDAG contains only a single DAG, meaning that if causal discovery recovers the correct CPDAG, it recovers the true graph exactly. This makes it a basic sanity check rather than a test of how performance degrades when enumerating multiple graphs in a Markov equivalence class. In our experiments, the discovered CPDAG did not fully match the ground truth: edges from Z_1 and Z_2 to X_1 and X_2 were missing (Fig. 7). However, the edges from X_1 and X_2 to Y were correctly recovered, which sufficed for CD-POMIS to match the performance of the oracle POMIS baseline (Fig. 2).

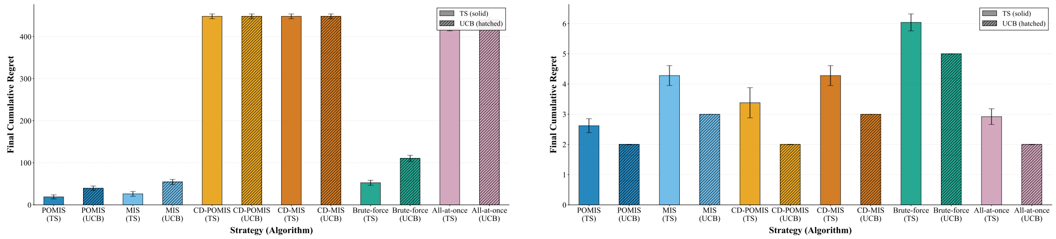
To explore the effects of equivalence-class ambiguity, we next introduced the **Chain SCM**, whose true graph $Z \to X \to Y$ yields a CPDAG with three Markov-equivalent DAGs (Fig. 8). We evaluated two variants of this setup: one without enforcing Y as a sink and another where outgoing edges from Y were prohibited. In the unconstrained case, the presence of DAGs where Y acts as a parent variable led to redundant exploration and higher regret. Enforcing the sink constraint eliminated implausible structures, effectively pruning two of the three DAGs and improving performance (Fig. 3). In this specific case, the constraint was sufficient to resolve all ambiguity, reducing the equivalence class to only the true graph. While this constraint is not guaranteed to resolve all ambiguity in every MEC, it can often reduce the number of possible graphs under consideration. Therefore, all subsequent experiments adopt this sink-enforced version for consistency.



(a) Chain SCM (no Y-sink)

(b) Chain SCM (Y-sink enforced)

Figure 3: Final regret comparison at $T = 2{,}000$ for Chain SCM variants. **Left:** Without Y-sink constraint. **Right:** With Y-sink constraint. Full regret trajectories are shown in Fig. 13.



(a) IV (XOR) variant

(b) IV (OR) variant

Figure 4: Comparison of final regret at $T = 2{,}000$ for the IV SCM under XOR and OR formulations. **Left:** XOR-based variant. **Right:** OR-based variant. Full regret trajectories are shown in Fig. 14.

4.3 Causally Insufficient Setting: Instrumental Variable SCM

We next evaluate the **Instrumental Variable (IV)** SCM from Lee and Bareinboim (2018), which introduces latent confounding between X and Y. In the original XOR-based formulation, the FCI algorithm failed to recover any edges, yielding an empty PAG and consequently poor bandit performance (Fig. 9). To strengthen dependency signals while preserving the underlying causal relationships, we introduced an **IV (OR)** variant that replaces XOR operations with OR functions. Although the discovered PAG did not fully match the true structure (Fig. 9), it recovered key dependencies that enabled significantly lower regret, demonstrating that improved structural recovery—even if partial—can lead to better downstream performance (Fig. 4).

4.4 Non-Manipulable Setting: Frontdoor and Four-Variable SCMs

We finally extend our evaluation to settings involving **non-manipulable variables**, following Lee and Bareinboim (2019).

Frontdoor SCM. For the **Frontdoor** model, we follow Lee and Bareinboim (2019) in defining $\mathbf{N} = \{Z\}$ as the non-manipulable mediator variable. Although the FCI algorithm partially recovered the structure, it identified only edges between Z and Y, failing to capture the complete causal pathway. Because Z cannot be directly intervened upon, this limited discovery prevented the bandit from converging to the optimal intervention, resulting in persistently high regret (Fig. 5).

Four-Variable SCM. For the **Four-Variable** model, we again follow Lee and Bareinboim (2019) in defining $\mathbf{N} = \{A\}$ as the non-manipulable variable, and we evaluate both XOR and OR functional variants. In the XOR-based formulation, the weak statistical dependencies caused the FCI algorithm to output an empty PAG resulting in poor structural recovery and high regret. Replacing XOR with OR functions strengthened the dependency signals, allowing FCI to recover the correct adjacencies and substantially improve bandit performance. Despite minor structural inaccuracies in the discovered PAG, the CD-POMIS strategy achieved the lowest final regret (Fig. 6). This counterintuitive result can be attributed to the fact that, due to the incomplete discovery, CD-POMIS identified only a

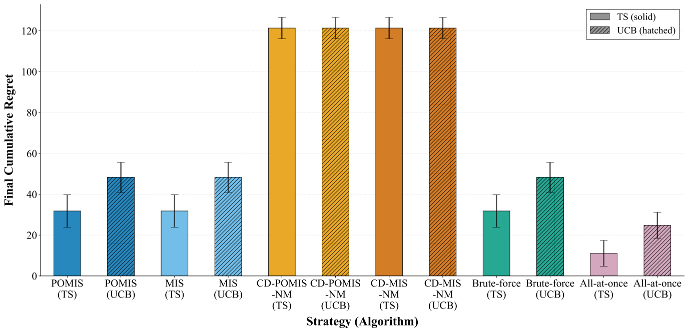
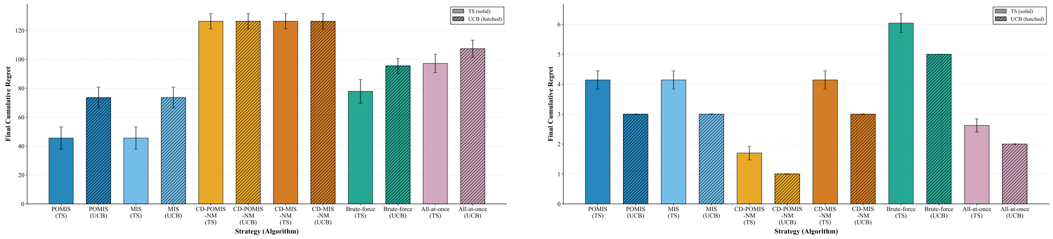


Figure 5: Final regret comparison at $T = 2{,}000$ for the Frontdoor SCM with $\mathbf{N} = \{Z\}$. Full regret progression is shown in Fig. 15



(a) Four-Variable SCM (XOR)

(b) Four-Variable SCM (OR)

Figure 6: Final regret comparison at $T = 2{,}000$ for the Four-Variable SCM with $\mathbf{N} = \{A\}$. **Left:** XOR-based variant. **Right:** OR-based variant. Full regret trajectories are shown in Fig. 16.

subset of the true POMISs. While the ground-truth POMIS included $\emptyset, \{B\}, \{C\}$ resulting in five possible arms, the discovered CD-POMIS contained only the $\{C\}$ intervention set (two arms in total). The smaller action space required less exploration, allowing the bandit to converge faster and thus outperform the oracle baseline despite the structural inaccuracies.

4.5 Structural Hamming Distance (SHD)

To quantify the relationship between causal discovery accuracy and bandit performance, Table 1 reports SHD alongside final regret (for CD-POMIS in causally sufficient settings and CD-POMIS-NM in causally insufficient settings) for all SCMs. While lower SHD generally improves performance for a given causal structure (e.g., IV (OR) with SHD=2 outperforms IV (XOR) with SHD=6), absolute SHD values do not directly predict regret. Four-Variable (OR) achieves near-optimal regret (1.0 ± 0.0) with SHD=6, while Frontdoor achieves only 121.36 ± 5.22 with SHD=4. This demonstrates that SHD uniformly penalizes all edge errors, but what matters for intervention selection is *which* edges are recovered rather than total edge count accuracy. The discovered structure need not be perfect. It must preserve the causal pathways essential for identifying optimal interventions, which may constitute only a subset of the full causal graph. Conversely, severe structural errors such as finding an empty graph inevitably lead to poor performance, as no causal information is available for intervention selection.

5 Discussion

The experiments show that the performance of causal bandit algorithms is highly dependent on the quality of the discovered causal structure. When the recovered graph captures the key causal dependencies, the bandit efficiently identifies optimal interventions and achieves low regret. Conversely, structural errors—such as missing edges or incorrect dependencies—can lead to redundant exploration and higher cumulative regret. In extreme cases where causal discovery fails entirely (e.g., yielding an empty graph), no causal information is available for intervention selection, resulting in poor performance. However, perfect causal discovery is not always required for strong performance.

Table 1: Structural Hamming Distance (SHD), equivalence class size (|MEC|), and final regret across SCMs. |MEC| represents an upper bound on graphs enumerated, as structural constraints may reduce this number.

SCM	Nodes	Type	Algorithm	SHD	|MEC|	Final Regret (UCB)
Simple Markovian	5	Causally sufficient	PC	4	1	3.0 ± 0.0
Chain	3	Causally sufficient	PC	0	3	1.0 ± 0.0
IV (XOR)	3	Causally insufficient	FCI	6	1	448.28 ± 5.63
IV (OR)	3	Causally insufficient	FCI	2	6	2.0 ± 0.00
Frontdoor	3	Causally insufficient, NM	FCI	4	5	121.36 ± 5.22
Four-Variable (XOR)	4	Causally insufficient, NM	FCI	12	1	126.36 ± 5.22
Four-Variable (OR)	4	Causally insufficient, NM	FCI	6	6	1.0 ± 0.0

In both the Simple Markovian and Four-Variable (OR) SCMs, CD-POMIS achieved comparable or even lower regret than the oracle POMIS baseline, showing that partial recovery can be sufficient to find the optimal arm.

Enforcing practical constraints further improves stability and performance. Making the reward variable Y a sink, for instance, eliminated implausible causal directions and led to better performance in the Chain SCM. The patterns observed in causally sufficient settings persist when non-manipulable variables are present: performance does not degrade as long as the critical causal pathways among manipulable variables are recovered.

We also observed that XOR-based functional relationships produced weak statistical dependencies that caused FCI to fail entirely in the IV and Four-Variable SCMs, while replacing them with OR operations enabled successful discovery without altering the causal topology.

Limitations While the proposed framework demonstrates the potential of integrating causal discovery with structural causal bandits, several limitations remain. First, the experiments are limited to small-scale SCMs with binary variables, which may not generalize directly to high-dimensional or continuous systems. Second, our approach depends on the accuracy of constraint-based causal discovery algorithms such as PC and FCI, which can be sensitive to sample size and statistical test thresholds. We acknowledge that other classes of causal discovery methods exist, such as score-based or restricted structural models, which were not evaluated in this work but could be applicable, particularly in causally sufficient settings. Finally, the aggregation of POMIS across all members of an equivalence class, while conservative, may lead to overly large action spaces and increased computation times in more complex graphs.

Conclusions This work takes a first step toward relaxing the assumption of a known causal graph in structural causal bandits by integrating causal discovery. Our experiments show that performance depends critically on the quality of the discovered structure: when key dependencies are recovered, the method matches oracle baselines, but degrades otherwise. Future research could explore more efficient methods for reasoning directly over equivalence classes—without enumerating all member graphs—drawing inspiration from Park et al. (2025). Another promising direction is incorporating the z^{2}ID approach from Lee and Bareinboim (2019), which leverages z-identifiability to determine when an arm's reward distribution can be estimated from other interventions, thereby enabling information sharing and improved sample efficiency.

References

Auer, P., Cesa-Bianchi, N., and Fischer, P. (2002). Finite-time analysis of the multiarmed bandit problem. *Machine Learning*, 47(2-3):235–256.

Bastani, H. and Bayati, M. (2020). Online decision making with high-dimensional covariates. *Operations Research*, 68(1):276–294.

Cappé, O., Garivier, A., Maillard, O.-A., Munos, R., and Stoltz, G. (2013). Kullback–leibler upper confidence bounds for optimal sequential allocation. *The Annals of Statistics*, 41(3):1516–1541.

Durand, A., Achilleos, C., Iacovides, D., Strati, K., Mitsis, G. D., and Pineau, J. (2018). Contextual bandits for adapting treatment in a mouse model of de novo carcinogenesis. In *Machine Learning for Healthcare Conference*, pages 67–82.

Huo, X. and Fu, F. (2017). Risk-aware multi-armed bandit problem with application to portfolio selection. *Royal Society Open Science*, 4(11):171377.

Hyttinen, A., Eberhardt, F., and Järvisalo, M. (2017). Causal discovery with general non-linear relationships using sat. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pages 2404–2410.

Jabbari, F., Ramsey, J., Spirtes, P., and Cooper, G. F. (2017). Discovery of causal models that contain latent variables through Bayesian scoring of independence constraints. In *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2017*, volume 10535 of *Lecture Notes in Computer Science*, pages 142–157. Springer.

Kalisch, M., Mächler, M., Colombo, D., Maathuis, M. H., and Bühlmann, P. (2025). *pcalg: Methods for Graphical Models and Causal Inference*. R package version 2.7-12.

Lee, S. and Bareinboim, E. (2018). Structural causal bandits: Where to intervene? In *Advances in Neural Information Processing Systems*, volume 31.

Lee, S. and Bareinboim, E. (2019). Structural causal bandits with non-manipulable variables. In *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence*, volume 33, pages 7369–7376.

Park, M. W., Arditi, A., Bareinboim, E., and Lee, S. (2025). Structural causal bandits under markov equivalence. Technical Report R-122, Columbia CausalAI Laboratory.

Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, 82(4):669–688.

Robbins, H. (1952). Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*, 58(5):527–535.

Shen, W., Wang, J., Jiang, Y.-G., and Zha, H. (2015). Portfolio choices with orthogonal bandit learning. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*.

Spirtes, P., Glymour, C. N., Scheines, R., and Heckerman, D. (2000). *Causation, Prediction, and Search*. MIT Press, Cambridge, MA, 2nd edition.

Spirtes, P., Meek, C., and Richardson, T. (1995). Causal inference in the presence of latent variables and selection bias. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, pages 499–506. Morgan Kaufmann Publishers Inc.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. MIT Press, 2nd edition.

Thompson, W. R. (1933). On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3-4):285–294.

Tsamardinos, I., Brown, L. E., and Aliferis, C. F. (2006). Max-min hill-climbing Bayesian network structure learning algorithm. *Machine Learning*, 65(1):31–78.

Verma, T. and Pearl, J. (1990). Equivalence and synthesis of causal models. In *Proceedings of the Sixth Conference on Uncertainty in Artificial Intelligence (UAI 1990)*, pages 220–227.

Zheng, Y., Huang, B., Chen, W., Ramsey, J., Gong, M., Cai, R., Shimizu, S., Spirtes, P., and Zhang, K. (2024). Causal-learn: Causal discovery in python. *Journal of Machine Learning Research*, 25(60):1–8.

Appendix

A Structural Causal Model Definitions

This appendix provides the full definitions of all Structural Causal Models (SCMs) used in our experiments, following the formulations in Lee and Bareinboim (2018, 2019). For each SCM, we list the exogenous variable distributions $P(U)$ and structural assignments f_i, along with the non-manipulable variable set $\mathbf{N}$ when applicable. We also provide both the original XOR-based and the modified OR-based functional variants used in our experiments.

A.1 Markovian SCMs

Simple Markovian SCM (Task 1, Lee and Bareinboim (2018))

$$\begin{aligned} &P(U_{X_1} = 1) = 0.54, \quad P(U_{X_2} = 1) = 0.67, \quad P(U_Y = 1) = 0.58, \\ &P(U_{Z_1} = 1) = 0.54, \quad P(U_{Z_2} = 1) = 0.44 \end{aligned}$$

and the structural functions:

$$\begin{aligned} f_{Z_1}(u_{Z_1}) &= u_{Z_1}, \\ f_{Z_2}(u_{Z_2}) &= u_{Z_2}, \\ f_{X_1}(z_1, z_2, u_{X_1}) &= z_1 \oplus z_2 \oplus u_{X_1}, \\ f_{X_2}(z_1, z_2, u_{X_2}) &= 1 \oplus z_1 \oplus z_2 \oplus u_{X_2}, \\ f_Y(x_1, x_2, u_Y) &= (x_1 \wedge x_2) \vee u_Y. \end{aligned}$$

Chain SCM (ours). We define a binary Markovian SCM with structure $Z \to X \to Y$, designed to induce multiple Markov-equivalent DAGs under the PC algorithm.

$$P(U_Z = 1) \in [0.3, 0.7], \quad P(U_X = 1) \in [0.01, 0.1], \quad P(U_Y = 1) \in [0.01, 0.1],$$

and the structural functions:

$$\begin{aligned} f_Z(u_Z) &= u_Z, \\ f_X(z, u_X) &= z \vee u_X, \\ f_Y(x, u_Y) &= x \vee u_Y. \end{aligned}$$

This formulation maintains the canonical chain structure while producing stronger dependencies that allow the PC algorithm to recover a non-trivial CPDAG.

A.2 Non-Markovian SCMs

Instrumental Variable (IV) SCM (Task 2: Lee and Bareinboim (2018))

$$P(U_X = 1) = 0.11, \quad P(U_Y = 1) = 0.15, \quad P(U_Z = 1) = 0.6, \quad P(U_{XY} = 1) = 0.51,$$

and the structural functions (XOR-based):

$$\begin{aligned} f_Z(u_Z) &= u_Z, \\ f_X(z, u_X, u_{XY}) &= u_X \oplus u_{XY} \oplus z, \\ f_Y(x, u_Y, u_{XY}) &= 1 \oplus u_Y \oplus u_{XY} \oplus x. \end{aligned}$$

We additionally define an OR-based variant by replacing $\oplus$ (XOR) with $\vee$:

$$\begin{aligned} f_Z(u_Z) &= u_Z, \\ f_X(z, u_X, u_{XY}) &= u_X \vee u_{XY} \vee z, \\ f_Y(x, u_Y, u_{XY}) &= u_Y \vee u_{XY} \vee x. \end{aligned}$$

The OR formulation preserves the causal structure while producing stronger dependencies that improve FCI recovery.

Frontdoor SCM (Lee and Bareinboim, 2019).

$$\begin{aligned}\mathbf{N} &= \{Z\},\\ P(U_X = 1) = 0.5, &\quad P(U_Y = 1) = 0.4,\\ P(U_Z = 1) = 0.4, &\quad P(U_{XY} = 1) = 0.5\end{aligned}$$

and the structural functions:

$$\begin{aligned}f_X &= u_X \oplus u_{XY},\\ f_Z &= u_Z \oplus x,\\ f_Y &= (u_Y \wedge u_{XY}) \oplus z.\end{aligned}$$

Four-Variable SCM (Lee and Bareinboim, 2019).

$$\begin{aligned}&\mathbf{N} = \{A\},\\ &P(U_B = 1) = 0.5, \quad P(U_C = 1) = 0.25, \quad P(U_Y = 1) = 0.25,\\ &P(U_{BY}) = 0.25, \quad P(U_{AB}) = 0.4\end{aligned}$$

and the structural functions (XOR-based):

$$\begin{aligned}f_A &= u_{AB},\\ f_B &= u_B \oplus u_{AB} \oplus u_{BY},\\ f_C &= u_C \oplus a \oplus b,\\ f_Y &= 1 - (u_{BY} \oplus u_Y \oplus a \oplus c).\end{aligned}$$

We additionally define an OR-based variant by replacing $\oplus$ (XOR) with $\vee$:

$$\begin{aligned}f_A &= u_{AB},\\ f_B &= u_B \vee u_{AB} \vee u_{BY},\\ f_C &= u_C \vee a \vee b,\\ f_Y &= u_{BY} \vee u_Y \vee a \vee c.\end{aligned}$$

As with the IV case, replacing XOR with OR operations enhances detectability for constraint-based causal discovery while preserving the qualitative causal structure.

B Causal Discovery Sanity Checks

This section presents the causal discovery sanity checks for all SCMs, showing both the ground-truth graphs and the corresponding discovered structures (CPDAGs for PC, PAGs for FCI). For causally sufficient settings (PC), we additionally visualize all enumerated DAGs within the Markov equivalence class, along with the corresponding POMIS and MIS sets computed per DAG, confirming correctness under structural uncertainty. For causally insufficient settings (FCI), we instead report the discovered PAG together with the number of consistent ADMGs it entails. While all ADMGs are fully enumerated during computation, their number often grows rapidly, making direct visualization impractical. These figures collectively verify the correctness of our causal discovery, equivalence-class enumeration, and intervention-set computation pipelines.

Simple Markovian SCM. Figure 7 shows the sanity check for the Simple Markovian SCM. The discovered CPDAG did not fully match the ground truth: edges from Z_1 and Z_2 to X_1 and X_2 were missing. However, the CPDAG contains only a single DAG, and the POMIS and MIS sets computed for this DAG confirm the correctness of the implementation.

Chain SCM. Figure 8 shows the sanity check for the Chain SCM. The discovered CPDAG exactly matches the ground truth. The equivalence class contains multiple DAGs, and the POMIS and MIS sets are computed for each, confirming correct enumeration and intervention-set computation.

Instrumental Variable SCM. Figure 9 shows the sanity checks for the IV SCM under both XOR and OR formulations. The XOR variant yields an empty PAG (no discovered edges), while the OR variant strengthens dependencies, enabling FCI to recover part of the true structure and produce a richer equivalence class.

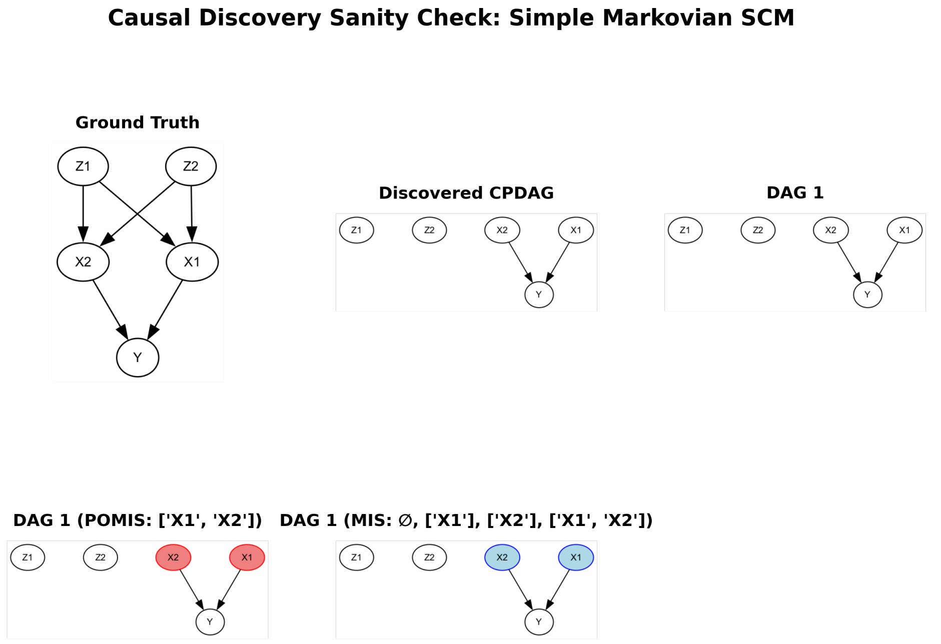

Causal Discovery Sanity Check: Simple Markovian SCM
Ground Truth
Z1
Z2
X2
X1
Y
Discovered CPDAG
DAG 1
DAG 1 (POMIS: ['X1', 'X2'])
DAG 1 (MIS: ∅, ['X1'], ['X2'], ['X1', 'X2'])

Figure 7: Sanity check for the **Simple Markovian SCM**.

Frontdoor SCM. Figure 10 shows the sanity check for the Frontdoor SCM with $\mathbf{N} = \{Z\}$ as the non-manipulable mediator. The FCI algorithm recovered partial structure, identifying edges between Z and Y, while missing the full causal pathway. The number of enumerated ADMGs is shown below the discovered PAG.

Four-Variable SCM. Figure 11 shows the sanity checks for the Four-Variable SCM with $\mathbf{N} = \{A\}$ as the non-manipulable variable. The XOR variant produces an empty PAG, indicating weak statistical dependencies. Replacing XOR with OR functions strengthens dependencies, leading FCI to discover more edges and a larger set of consistent ADMGs.

C Additional Results: Full Regret Progression

This section reports the full cumulative regret trajectories (mean with 95% CIs over 50 trials) for all SCMs and strategy variants discussed in the main text. For each final-regret figure in the paper, the corresponding full progression is provided here.

Simple Markovian SCM. Figure 12 shows the full regret progression for the Simple Markovian SCM. All methods converge quickly, with CD-POMIS matching the oracle POMIS baseline despite minor discovery errors.

Chain SCM. Figure 13 shows the full regret progression for the Chain SCM under two variants. Enforcing Y as a sink removes implausible DAGs and yields better convergence.

Instrumental Variable SCM. Figure 14 shows the full regret progression for the IV SCM. The XOR variant yields an empty PAG, causing all methods to perform poorly. The OR variant strengthens dependencies, improving discovery quality and leading to lower regret.

Frontdoor SCM. Figure 15 shows the full regret progression for the Frontdoor SCM with $\mathbf{N} = \{Z\}$. Discovery was limited to edges connected to Z, preventing convergence under non-manipulability constraints.

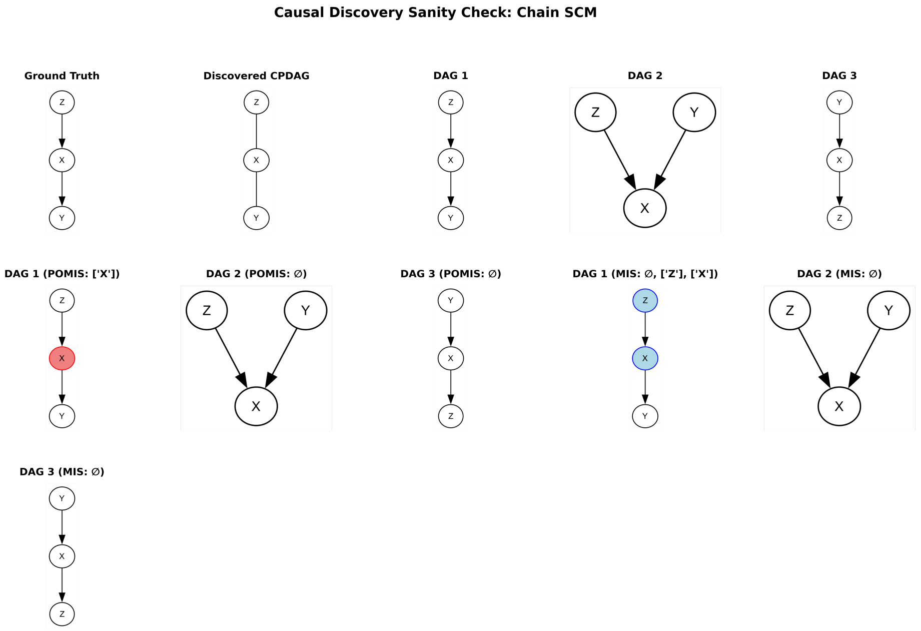
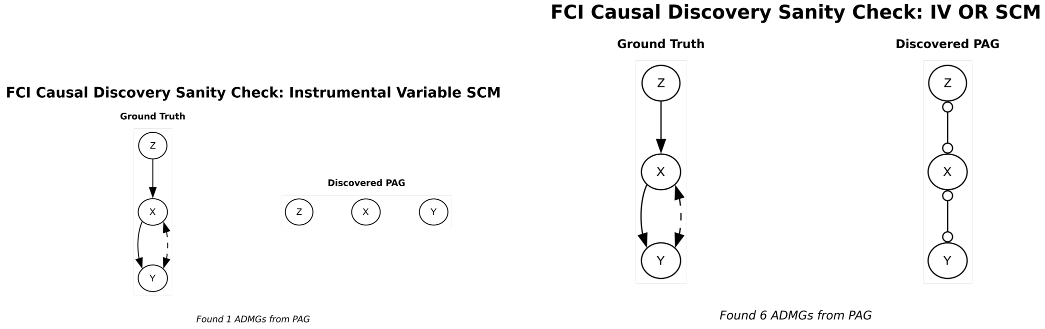


Figure 8: Sanity check for the **Chain SCM**.


FCI Causal Discovery Sanity Check: Instrumental Variable SCM
Ground Truth
Discovered PAG
Z
X
Y
Found 1 ADMGs from PAG
FCI Causal Discovery Sanity Check: IV OR SCM
Ground Truth
Discovered PAG
Z
X
Y
Found 6 ADMGs from PAG

(a) IV SCM (XOR)

(b) IV SCM (OR)

Figure 9: Sanity checks for the **Instrumental Variable (IV)** SCM.

Four-Variable SCM. Figure 16 shows the full regret progression for the Four-Variable SCM with $\mathbf{N} = \{A\}$. The XOR variant produces an empty PAG, resulting in poor performance. The OR variant yields better discovery, and CD-POMIS-NM achieves the lowest regret even with minor discovery errors.

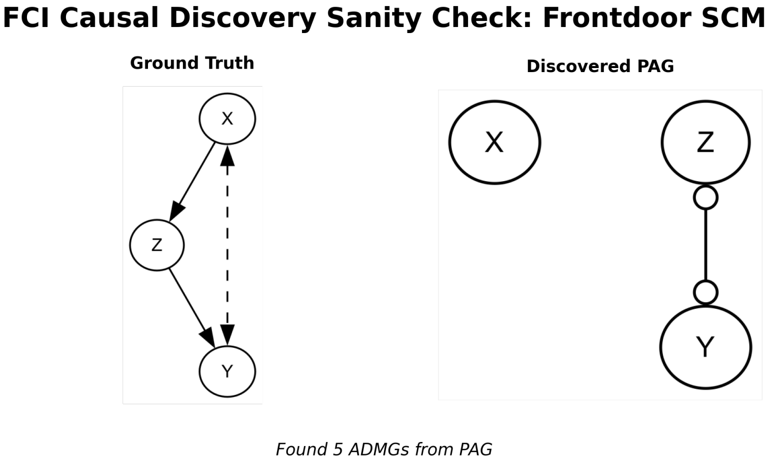
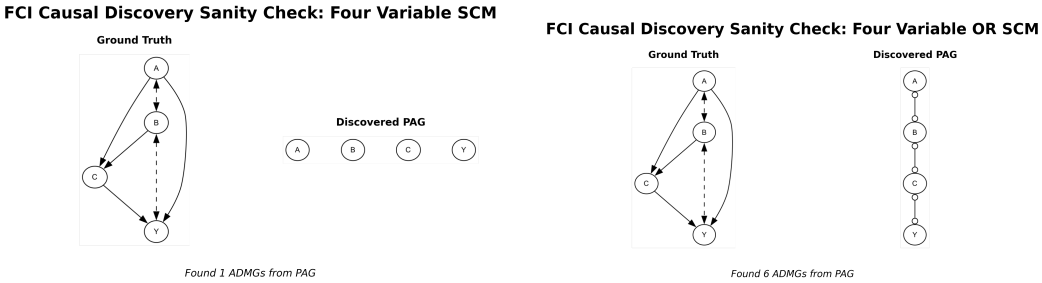


Figure 10: Sanity check for the **Frontdoor SCM** with $\mathbf{N} = \{Z\}$.



(a) Four-Variable SCM (XOR)

(b) Four-Variable SCM (OR)

Figure 11: Sanity checks for the **Four-Variable SCM** with $\mathbf{N} = \{A\}$.

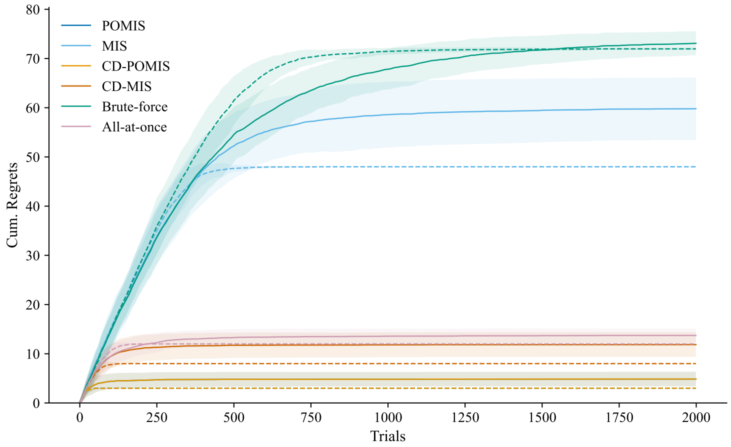
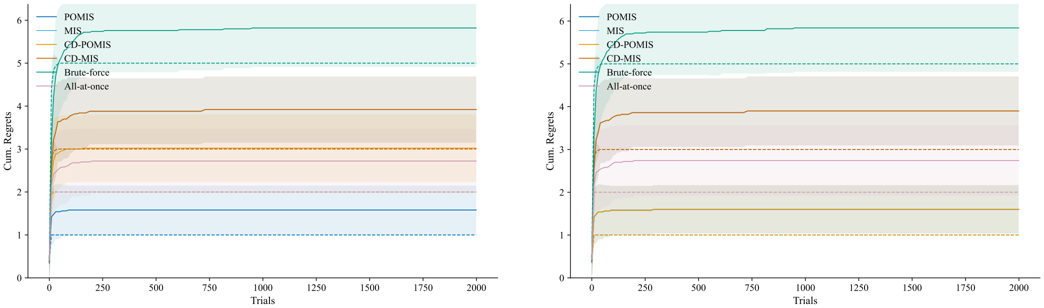


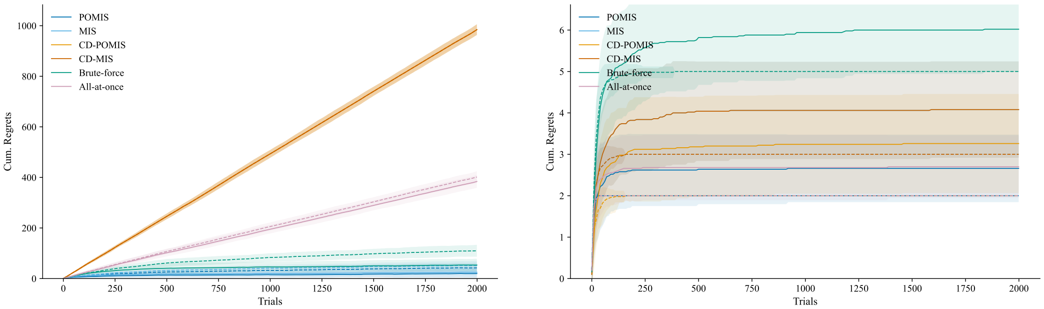
Figure 12: Full regret progression for the **Simple Markovian** SCM (cf. Fig. 2).



(a) No Y-sink

(b) Y-sink enforced

Figure 13: Full regret progression for the **Chain** SCM (cf. Fig. 3).



(a) XOR variant

(b) OR variant

Figure 14: Full regret progression for the **Instrumental Variable** SCM (cf. Fig. 4).

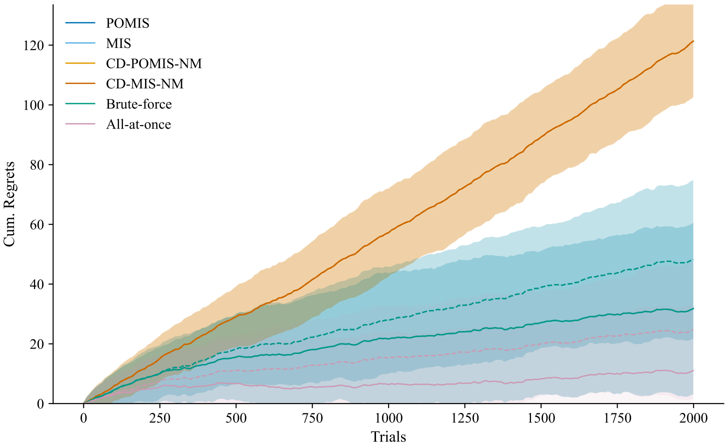
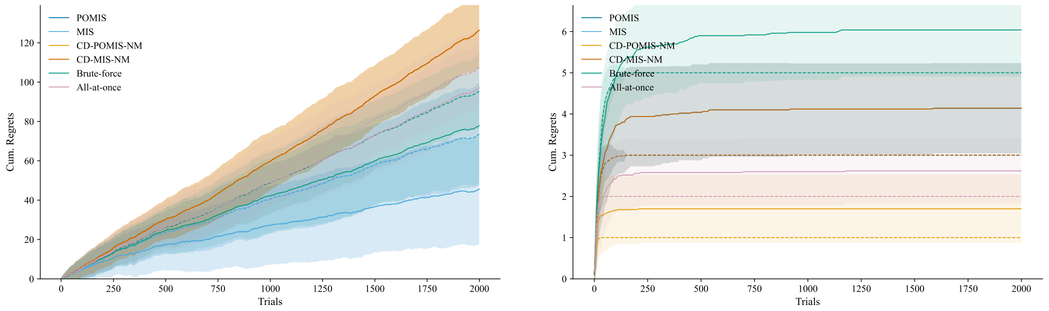


Figure 15: Full regret progression for the **Frontdoor** SCM with $\mathbf{N} = \{Z\}$ (cf. Fig. 5).



(a) XOR variant

(b) OR variant

Figure 16: Full regret progression for the **Four-Variable** SCM with $\mathbf{N} = \{A\}$ (cf. Fig. 6).