UNCERTAINTY-AWARE CAUSAL DECISION MAKING VIA EFFECT BOUND DECOMPOSITION

Anonymous authors
Paper under double-blind review

ABSTRACT

Causal inference from observational data can provide strong evidence for finding the best action in a decision-making scenario without having to perform expensive randomized trials. The causal effect of an action is often not pointwise identifiable even with infinite data due to unobserved confounding factors. Furthermore, having only finitely many samples adds another layer of uncertainty to causal effect estimation. Several existing methods can be used to obtain upper and lower bounds to the causal effect, ranging from symbolic methods to the more recent neural network-based approaches, which implicitly incorporate both sources of uncertainty. However, these methods do not inform whether collecting more samples may or may not help identify the best action from observational data, leaving experts in the dark about their data collection strategies. We address this problem with a novel framework that can distinguish the range of causal effect values that might be eliminated by collecting more samples from the range of values that, with high probability, cannot be eliminated with more observational samples. We show that this partitioning can be obtained by solving max-min and min-max optimization problems. We leverage neural causal models to approximately recover this decomposition in practice. We demonstrate via experiments on synthetic and real-world datasets that our algorithm can determine when collecting more samples will not help determine the best action. Our framework can help practitioners decide when to resort to non-observational studies or seek to measure some of the unmeasured confounders for optimal decision-making.

1 INTRODUCTION

Causal decision-making aims to identify the best action relative to a goal. For example, *which drug maximizes the life expectancy of an average patient? Which candidate government policy should be enacted to reduce unemployment more?* Since actions affect the system relative to its observational state, answering such questions requires causal models. Pearl's structural causal models (SCMs) provide us with a systematic set of complete algorithmic principles to identify the effect of actions, or post-interventional distributions from observational data under well-defined assumptions (Pearl, 1995a; Tian, 2002; Huang & Valtorta, 2006; Shpitser & Pearl, 2008). When the interventional distribution or the causal effect cannot be determined uniquely, the framework can be used to obtain bounds on these quantities (Zhang & Bareinboim, 2021a; Hu et al., 2021; Li & Pearl, 2022).

For decision-making, we may wish to understand if an action, say enforcing $X = x$, increases the probability of a certain event, say $Y = 1$. In SCM language, we want to test if $P(Y = 1|\text{do}(x)) > P(Y = 1)$. Or we may want ot decide which of the two actions, enforcing $X = x_0$ or enforcing $X = x_1$, leads to a higher probability of the said event, which we would test by asking if $P(Y = 1|\text{do}(x_1)) > P(Y = 1|\text{do}(x_0))$. Equivalently, we can directly test whether the average treatment effect (ATE), defined as $\mathbb{E}[Y|\text{do}(x_1)] - \mathbb{E}[Y|\text{do}(x_0)]$, is positive. These can be extended to having more than two actions, for example, by testing if $P(Y = 1|\text{do}(x_0)) > \max_{i\neq 0} P(Y = 1|\text{do}(x_i))$. However, these interventional quantities that are essential for decision-making are not always identifiable even with infinite data, i.e., there may be different SCMs that entail the observational distribution, but lead to different interventional quantities. On top of this, in any realistic scenario, we have a *finite* dataset from the observational distribution, which introduces another source of uncertainty for estimating these interventional quantities.

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

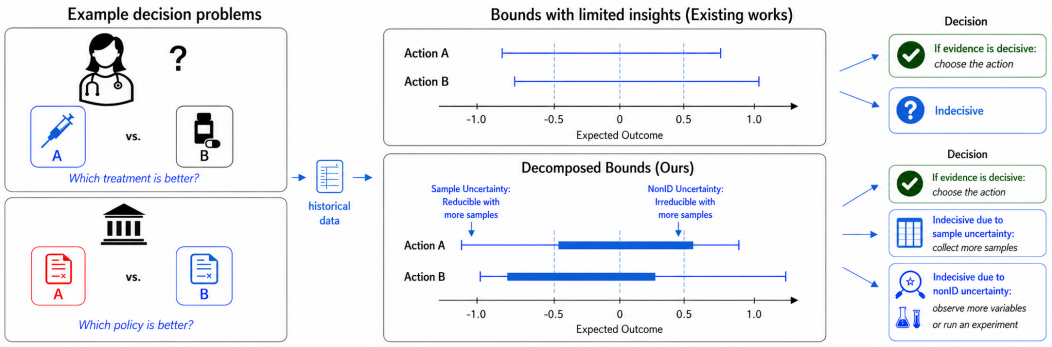


Figure 1: Decision making with decomposed causal effect bounds.

If, despite these sources of uncertainty, the *lower bound* to the utility of Action 1 is greater than the *upper bound* to the utility of all other actions, we can still decide Action 1 is more desirable than the others. However, often, compounding these two sources of uncertainty leads to a wide range of values for the interventional quantities of interest which leaves optimal action ambiguous. Importantly then it is not clear whether bounds were wide due to not having sufficiently many observational samples, or due to structural reasons that even with infinite samples one would not be able to make a decision. In the special cases where interventional quantites are identifiable due to the causal graph, we know collecting more samples will allow decision-making since the range of possible interventional values will converge to a single point (either ATE is positive or negative, or $P(Y = 1|\text{do}(x_1))$ is either greater or less than $P(Y = 1|\text{do}(x_0))$). However, to the best of our knowledge, it is currently unknown how to answer this question for non-identifiable causal queries.

We address this problem through a novel decomposition of the range of interventional values into two subsets: A range of interventional values that will not be removed with high probability by increasing the number of samples in our dataset, called the nonID region, and the rest that may be reduced by collecting more samples (e.g., by waiting for more data to be collected via sensors in an environment). In a causal decision-making scenario, this nonID region crucially helps us to decide the next move: collect more samples from observed distribution, or resort to a different distribution, e.g., by adding new sensors to observe more variables, or resorting to a randomized trial. We show that the proposed decomposition can be computed by solving min-max and max-min problems. Finally, to approximate these bounds in practice we develop an approach that leverages neural/deep causal generative models (Kocaoglu et al., 2018; Balazadeh Meresht et al., 2022; Rahman & Kocaoglu, 2024; Xia et al., 2021). Our contributions are summarized below:

- We propose a novel decomposition of the range of interventional quantities of interest into structural and sample uncertainty. This is, to the best of our knowledge, the first time a method can infer with high probability that collecting more observational samples can not help with finding the optimal action for non-identifiable causal queries.
- We establish a novel algorithm, `UA-DCM`, that approximately recovers our *nonID uncertainty* and *sample uncertainty* in practice.
- We demonstrate the utility of our algorithm through extensive experiments on synthetic data of various complexities and on a real-world parent labor dataset.

2 Background

We formalize the observational data-generating process with a structural causal model.

Definition 1 (Structural causal model, SCM) (Pearl, 2009)**.** An SCM S is a tuple of five elements: $S = (\mathbf{V}, \mathcal{N}, \mathcal{U}, \mathcal{F}, P(.))$. Here, each observed variable $V_i \in \mathbf{V}$ is realized as an evaluation of the function $f_i \in \mathcal{F}$ by taking as input a subset of the remaining observed variables $Pa_i \subset \mathbf{V}$, an exogenous noise variable $E_i \in \mathcal{N}$, and optionally an unobserved confounding variable $U_i \in \mathcal{U}$. $P(.)$ is a product joint over all unobserved $\mathcal{N} \cup \mathcal{U}$. An SCM containing unobserved confounders is called a **Non-Markovian causal model**. A *causal graph*, G, representing the variables as nodes and structural functional relationships encoded in the SCM as edges, is called an Acyclic Directed Mixed Graph (ADMG) and denoted as $S \models G$.

The best action is chosen based on the causal effect of each action on the outcome variable Y, which we formalize using do-interventions.

Definition 2 (Causal effect and do-intervention)**.** A do-intervention $do(x)$ replaces the functional equation $X = f_X$ with a specific value $X = x$ without affecting other equations. The distribution induced on the variable Y after such an intervention is called an interventional distribution $P(Y|\text{do}(x))$. With no intervention, the observational joint distribution of $\mathbf{V}$ is $P(\mathbf{V})$.

To efficiently find the bounds for causal effects, we model the SCM with deep neural networks.

Definition 3 (Deep causal models, DCM) (Kocaoglu et al., 2018; Xia et al., 2021; Rahman & Kocaoglu, 2024))**.** A neural net architecture $\mathbb{G}$ is called a deep causal model (DCM) for an ADMG $G = (\mathbf{V}, \mathcal{E})$ if it consists of a collection of neural nets, each f_i accepts a sufficiently high-dimensional noise vector N_i, ii) the output of f_j is input to f_i iff $V_j \in Pa_G(V_i)$, iii) $N_i = N_j$ iff V_i, V_j share an unobserved confounder. Functions in DCM, $\mathbb{G} = \{f_1, ..., f_n\}$ are parameterized by $\Theta = \{\theta_1, ..., \theta_{|\mathcal{V}|}\}$. Similar to the original data distribution, $P(\mathcal{V})$, we define $P_\theta(\mathcal{V})$ to be the distribution induced by the DCM. Sufficiently high-dimensional noise vectors N_i, e.g., Gaussian, can replace both the exogenous noises and the unobserved confounders in the true SCM. We consider DCM to be *representative enough for an SCM* if the neural networks have sufficiently many parameters to induce the same observed distribution as the true SCM. When a latent confounder U affects two observed variables X and Y, we can match the joint distribution $P(X, Y)$ by feeding the same noise $N_X = N_Y$ (as latents) into both f_X, f_Y.

Definition 4 (Causal effect with DCM)**.** To perform a hard intervention $do(X = x)$, we manually set the values for the intervened variables as $X = x$ instead of using their neural network. Then, we feed forward those values into their children and the rest of the mechanisms will work as usual to generate their corresponding values from which we can estimate the causal effect $P(y|\text{do}(x))$ and ATE. A loss function measuring $d(P(\mathbf{V}, P_\theta(\mathbf{V}))$ can be used to train DCM's end-to-end differentiable networks. When causal effects are not pointwise identifiable, we add an additional loss function to maximize or minimize the causal effects. Please check Theorem 8 for details.

3 Related Work

When the causal effect is not identifiable, one can identify a range of causal effects from the SCMs that are compatible with the observational data. Tian & Pearl (2000) used the response variable to get bounds of counterfactual queries from observational and interventional data. Duarte et al. (2024) proposed an automated method to derive bounds for causal queries in arbitrary graphs. Many researchers have explored the use of neural nets Xia et al. (2021); Balazadeh Meresht et al. (2022); Rahman & Kocaoglu (2024) to design causal models and estimate causal queries by maximizing and minimizing the query under the semi-Markovian setting.

In the uncertainty quantification literature, Melnychuk et al. (2024) proposed a method for quantifying aleatoric uncertainty in individualized treatment effects by deriving sharp bounds on the conditional distributions of the treatment effect. Marmarelis et al. (2024) introduced an approach for constructing causal effect intervals with hidden confounding. Existing causal methods approach the decision-making problem from different perspectives. Frauen et al. (2023) proposed a neural framework for sensitivity analysis that is compatible with a large class of existing sensitivity models. Jesson et al. (2020) studied the decision-making problem under non-overlap. Frauen et al. (2025) propose an optimal decision-making approach with two-stage CATE estimators.

These approaches cannot disentangle the uncertainty in causal effect bounds obtained from finite samples. Distributionally robust optimization (DRO) (Bui et al., 2022; Gao & Kleywegt, 2023) also performs max-min or min-max optimization. However, while DRO searches for a model that is robust to distributions within an ϵ-ball, we search for distributions within the ball that allow a model to achieve the best local minimum or the worst local maximum. Confidence intervals can also be added to existing causal effect estimators via bootstrapping, but neither these intervals nor existing causal effect bounds can distinguish the sources of uncertainty or suggest to a practitioner how to reduce it.

Chernozhukov et al. (2013) propose a precision-corrected estimator for intersection bounds. Although intersection bounds shares the same name as our paper, the object being intersected is different from that in our paper and serves a different purpose. The intersection bounds correction (Chernozhukov et al., 2013) is also applicable when the causal effect can be expressed with symbolic terms (Sachs

et al., 2023). But the number of parameters grows exponentially with the number of variables in the DAG. And more importantly, existing work does not provide the decomposition of the uncertainty. In contrast, our method not only allows us to quantify the uncertainty of causal effects for queries that cannot be expressed as symbolic bounds, but also enables the decomposition of that uncertainty into finite-sample uncertainty and non-id uncertainty, which can improve downstream decision-making.

4 Causal Decision-Making with Finite Data

We first define the average treatment effect that we will use to compare the effect of multiple actions.

Definition 5 (Average Treatment Effect)**.** *The average treatment effect (ATE) is defined for binary treatments as:* $ATE = \mathbb{E}[Y|do(X=1)] - \mathbb{E}[Y|do(X=0)]$. *For multiple actions, we define the ATE of an action* $X = x$ *with respect to the best among remaining actions* x_b *as:* $ATE(x) = \mathbb{E}[Y|do(X=x)] - \max_{x_b \neq x} \mathbb{E}[Y|do(X=x_b)]$.

4.1 Interval for Identifiable and Non-identifiable ATE

When randomized trials are not possible, we need an estimation of an interventional query from observational data. The ID algorithm (Tian & Pearl, 2002; Shpitser & Pearl, 2008) provides us with a complete characterization of causal queries that can be uniquely identified from observational joint distribution. However, this might be challenging in practice, as we do not have access to the true joint distribution from finite observational samples. We propose using a confidence set to quantify the uncertainty of the observational distribution. Specifically, for a probability distribution P, $\mathcal{C}(\hat{P})$ is a random set such that $\mathbb{P}(P \in \mathcal{C}(\hat{P})) \geq 1 - \alpha$, where $\mathbb{P}$ is the empirical measure induced by observing a given number of IID samples from P.

Having a set of possible observational distributions converts a pointwise identifiable query to an interval for the causal effect. We establish this first:

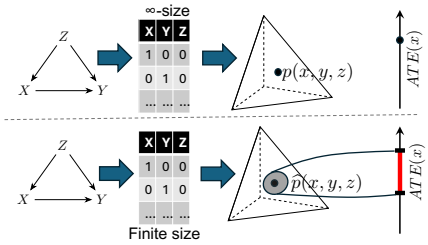


Figure 2: Given finite samples, we cannot pointwise estimate since the true distribution lies in a continuum of a confidence set, but only obtain an interval.

Theorem 1. *Let* $f(P, S)$ *be the estimand of an identifiable causal effect for SCM* S*, with the observational distribution* P*. Let* $\mathcal{P}$ *be any connected set of observational distributions that contains* P*. Let* $a = \min_{p \in \mathcal{P}} f(P, S), b = \max_{P \in \mathcal{P}} f(P, S)$*. Then* $f|_{\mathcal{P}} : \mathcal{P} \to [a, b]$ *is surjective.*

In words, the collection of ATEs for a set of observational distributions forms a contiguous interval in $[0, 1]$. This establishes that it is sufficient to obtain maximum and minimum of ATE, to characterize this range of ATE values. For certain graphs this problem can be solved easily, e.g., backdoor graph, the objective can be reduced to a linear program (Section D.2).

Next, we address non-identifiable queries under finite sample uncertainty. When the causal effect is identifiable, the only source of uncertainty is due to finite sample. With unobserved confounders, however, the effect may not be uniquely determined even with infinite samples: multiple SCMs consistent with the causal graph and the observational distribution may induce different causal effects. This is known as a non-identifiable query. For a fixed causal graph and observational distribution, existing algorithms can search the SCM space for the minimum and maximum causal effect, yielding a bound $P(y|\text{do}(x)) \in [L_x, U_x]$. Similar to Theorem 1, under the conditions of Lemma 3, the attainable values of a non-identifiable scalar causal query form an interval. Therefore, its minimum and maximum are sufficient to characterize the entire identified set.

Theorem 2. *Let* $f(p, S)$ *be the estimand of some scalar non-identifiable causal effect of interest for a finite-state causal graph satisfying the conditions of Lemma 3, and let* p *be a compatible observational distribution. Let* $\mathcal{S}$ *be the set of the SCM such that* $P_{obs}(S) = p$*. Let* $a = \min_{S \in \mathcal{S}} f(p, S), b = \max_{S \in \mathcal{S}} f(p, S)$*. Then* $f|_p : \mathcal{S} \to [a, b]$ *is surjective.*

4.2 Decomposing Causal Effect Uncertainty

Given the uncertainties in estimating the ATE, we now define the challenges in decision-making.

Definition 6 (Causal Decision)**.** *A causal decision with $ATE(x)$, given confidence set $\mathcal{C}(D)$, is unambiguous for action x if all SCM S entailing some $p \in \mathcal{C}(D)$ agree in sign, i.e., either $ATE(x) \geq 0$ for every such S, or $ATE(x) \leq 0$ for every such S. The decision is ambiguous if there exist two SCM S_1, S_2 entailing $p_1, p_2 \in \mathcal{C}(D)$ such that $ATE(x) > 0$ in S_1 and $ATE(x) < 0$ in S_2.*

When a decision is ambiguous, a natural next step is to collect more samples, but for a non-identifiable query, decision might still be ambiguous even with infinite data due to the non-ID uncertainty. In the next section, we decompose the uncertainty for decision-making.

The main question we pose in this paper is: *Can we know when collecting more samples will unlikely lead to informed decision-making?* To answer this question, we need to gain insight into two different types of epistemic uncertainty without knowing the observational distribution exactly: one modeling uncertainty due to unobserved confounding, and one modeling the uncertainty due to finite samples.

Our proposed method distinguishes the two uncertainties and provide a lower bound for the uncertainty that is unlikely to be reduced by collecting more data. This helps us decide the right conclusion about data collection practices for decision-making.

Assumptions: We assume *i)* discrete variables , *ii)* Semi-Markovian SCM and *iii)* access to the acyclic directed mixed graph (ADMG). See Section D for details. In this section, we discuss a general scenario in which we compare the effects of multiple (two or more) actions to choose the best action. We discuss the case of optimizing only a single action in Section D.2. We formally introduce the two types of uncertainties below.

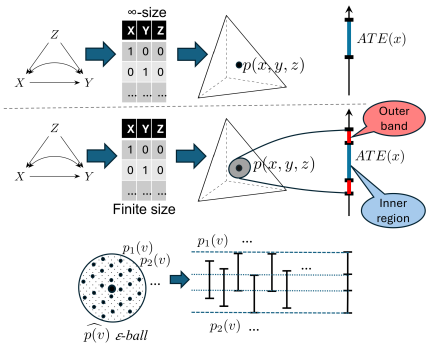


Figure 3: Inner and outer causal effect bounds when the effect is not ID.

Definition 7 (sample and nonID Uncertainty (Figure 3))**.** *Let $\hat{P}$ be an empirical estimation of the joint distribution and a confidence set $\mathcal{C}$ be such that $\mathcal{C}$ covers the true P with probability of at least $1-\alpha$. Define the set $\mathcal{S}_P = \{S \text{ an SCM} \mid P_{obs}(S) = P, S \models G\}$ and $\mathcal{I}_P(x) = \{ATE(x) \mid S \in \mathcal{S}_P\}$. Then define the intersection $\bigcap_{P\in\mathcal{C}} \mathcal{I}_P$ to be the **nonID uncertainty**, and the set difference $\bigcup_{P\in\mathcal{C}} \mathcal{I}_P \setminus \bigcap_{P\in\mathcal{C}} \mathcal{I}_P$ to be the **sample uncertainty**.*

Intuitively, for each distribution P in the confidence set $\mathcal{C}$, we obtain a set of ATE values $\mathcal{I}_P(x)$ from a set of SCMs $\mathcal{S}_P$ that matches P and consistent with causal graph G. We construct nonID uncertainty by taking intersection of $\mathcal{I}_P$ for all $P \in \mathcal{C}$ and claim the following proposition:

Proposition 1. *Let $P^\star$ be the true observational distribution, $\mathcal{C}(D)$ a confidence set satisfying $\Pr(P^\star \in \mathcal{C}(D)) \geq 1-\alpha$, and $\mathcal{I}_P(x)$ the causal-effect identified set for observational distribution P. On the coverage event $P^\star \in \mathcal{C}(D)$, (i) $\bigcap_{P\in\mathcal{C}(D)} \mathcal{I}_P(x) \subseteq \mathcal{I}_{P^\star}(x)$; and (ii) if the inner region crosses the decision boundary, i.e., $L_{\text{inner}} < 0 < U_{\text{inner}}$, then, with probability at least $1-\alpha$, the population identified set $\mathcal{I}_{P^\star}(x)$ also crosses zero.*

Note that (i) should be interpreted as a high-probability inner certificate of population non-identification obtained from the current dataset D. By (ii), if the inner region $\bigcap_{P\in\mathcal{C}(D)} \mathcal{I}_P(x)$ crosses zero, then with probability at least $1-\alpha$, the population identified set $\mathcal{I}_{P^\star}(x)$ crosses zero, and the decision would remain ambiguous even if infinitely many additional observational samples revealed $P^\star$ exactly. Thus, we should ***obtain*** additional variables to reduce this uncertainty. This result does not require empirical distributions or confidence sets obtained from future datasets to be identical or nested. Next, we construct sample uncertainty for ATE values: if $|\bigcup_{P\in\mathcal{C}} \mathcal{I}_P \setminus \bigcap_{P\in\mathcal{C}} \mathcal{I}_P| \neq 0$, it is possible that collecting more data points can disambiguate the causal effect. Thus, we ***collect*** more samples to reduce this uncertainty. Formally:

Proposition 2. *There exists SCMs $S \in \mathcal{S}_P$ where the uncertainty in the causal decision-making problem due to sample uncertainty in $ATE(x)$ can be reduced by increasing the dataset size.*

To represent the uncertainty regions in Proposition 1- 2 numerically, we define four quantities:

Algorithm 1 Explore ϵ-ball

Input: data $\{\mathbf{v}_k\}_{k=1}^n$, graph $\mathcal{G}$, small interval ϵ_s
Output: Decision/Collect/Observe
Initialize $\gamma = 0$ for multiple actions (ATE). ▷ Decision boundary value
while Not Decided **do**
  Candidates = Construct ϵ-net(data, $\mathbf{V}, \epsilon_s$) ▷ Candidate joint distributions consistent with input data
  for $P_\epsilon(\mathbf{V}) \in$ Candidates **do** ▷ Explore the confidence set
    for $x \in \sup(X)$ **do** ▷ Optimize each action in the support
      $U_x, L_x \leftarrow \text{RelaxedDCM}(P_\epsilon(\mathbf{V}), \mathbf{V}, \mathcal{G}, X = x, \epsilon_s)$ ▷ Execute Alg 2 to compute bounds
      $\overline{U}_x, \underline{U}_x \leftarrow \max(\overline{U}_x, U_x), \min(\underline{U}_x, U_x)$ ▷ Update upper bounds (maxmax,minmax)
      $\underline{L}_x, \overline{L}_x \leftarrow \min(\underline{L}_x, L_x), \max(\overline{L}_x, L_x)$ ▷ Update lower bounds (minmin, maxmin)
  if $\gamma \notin [\underline{L}_x, \overline{U}_x]$ **then** ▷ Decision is conclusive
    Return $X = x$ as conclusive; $\text{do}(x)$ is better/worse compared to other actions.
  else if $\gamma \in [\overline{L_x}, \underline{U_x}]$ **then** ▷ Additional samples are futile. Need more variables
    Observe variables
  else if $\gamma \in [\underline{L}_x, \overline{L}_x] \cup [\underline{U}_x, \overline{U}_x]$ **then** ▷ Collect samples to reduce sample uncertainty
    Collect more samples

Definition 8. *For some action $x \in \mathcal{X}$, we define these quantities to compute uncertainty regions.*

$$\begin{aligned} \overline{U_x} &:= \sup_{P \in \mathcal{C}(\hat{P})} \max_{S \in \mathcal{S}_P} ATE(x) & \underline{L_x} &:= \inf_{P \in \mathcal{C}(\hat{P})} \min_{S \in \mathcal{S}_P} ATE(x), \\ \underline{U_x} &:= \inf_{P \in \mathcal{C}(\hat{P})} \max_{S \in \mathcal{S}_P} ATE(x) & \overline{L_x} &:= \sup_{P \in \mathcal{C}(\hat{P})} \min_{S \in \mathcal{S}_P} ATE(x). \end{aligned} \tag{1}$$

Theorem 3. *Given a confidence set $\mathcal{C}$ and ATE values $\mathcal{I}_P$, $\bigcap_{P \in \mathcal{C}} \mathcal{I}_P = [\overline{L_x}, \underline{U_x}]$ if not empty.*

Corollary 1. *Given a confidence set, $\mathcal{C}$ and a set of average treatment effect values, $\mathcal{I}_P$ (Def. 7),*

$$\bigcup_{P \in \mathcal{C}} \mathcal{I}_P \setminus \bigcap_{P \in \mathcal{C}} \mathcal{I}_P = [\underline{L_x}, \overline{L_x}] \cup [\underline{U_x}, \overline{U_x}].$$

Thus, we can access the sample and non-id uncertainty by estimating four quantities for ATE as Equation (1). Now that we have connected uncertainty regions with bounds obtained by mini-max optimization in Equation 1, we can select the best action or suggest how to approach the decision with Theorem 4.

Theorem 4. *For a causal decision-making problem with $ATE(x)$ as the decision metric, let $(\overline{U_x}, \underline{U_x}, \overline{L_x}, \underline{L_x})$ be estimated from the data. The decision of $X = x$ as the best action (or cannot be the best action) is unambiguous if $\underline{L_x} \geq 0$ (or $\overline{U_x} < 0$). The decision is ambiguous and, with probability at least $1 - \alpha$, cannot be improved with more data if $\overline{L_x} < 0 < \underline{U_x}$.*

From ambiguous to unambiguous decision: Using Theorem 4, we can reach the final decision through a sequence of three moves. *1. Return:* if $0 < \underline{L_x}$ then we return $do(X = x)$ as the best action. *2. Observe*: if $\overline{L_x} < 0 < \underline{U_x}$, then we observe additional variables for example instrument variables. Observing additional variables gives us a small causal effect set $\mathcal{I}'_P$, for each $P \in \mathcal{C}$. Thus, for fixed $\mathcal{C}$, we get a narrower inner region, $[\overline{L_x}, \underline{U_x}] = \bigcap_{P \in \mathcal{C}} \mathcal{I}'_P$. Also, the average non-id bound width will reduce since $1/|\mathcal{C}| \sum_{P \in \mathcal{C}} |\mathcal{I}'_P| \leq 1/|\mathcal{C}| \sum_{P \in \mathcal{C}} |\mathcal{I}_P|$. *3. Collect*: In all other scenarios, we are unsure about the source of uncertainty (low sample size or non-identifiability) and we collect more samples, with the assumption that obtaining samples with additional variables is more challenging compared to collecting samples of the same set of variables. More samples typically provide us a smaller confidence set $\mathcal{C}'(\hat{P})$ that concentrates around $P^\star$. By Proposition 1 and 3, the inner region, $[\overline{L}_x, \underline{U}_x] = \bigcap_{P \in \mathcal{C}'} \mathcal{I}_P$ (Theorem 5 and 3), remains contained in $\mathcal{I}_{P^\star}$ with probability at least $1 - \alpha$, typically widens toward $\mathcal{I}_{P^\star}$, while the outer band, $[\underline{L}_x, \overline{L}_x] \cup [\underline{U}_x, \overline{U}_x] = \bigcup_{P \in \mathcal{C}'} \mathcal{I}_P \setminus \bigcap_{P \in \mathcal{C}'} \mathcal{I}_P$ (Corollary 2 and 1), typically narrows. The *collect* and *observe* moves continue until we reach the *return* move and decide the best action. The decomposition of uncertainty methods we discussed is also applicable to other causal queries such as $\mathbb{E}[Y|do(x)]$ which we discuss in Section D.2.

4.3 DCM FOR ESTIMATING SAMPLE AND NON-ID UNCERTAINTY

We introduce three algorithms: Alg. 1 calls Alg. 3 to construct the ϵ-net and then executes Alg. 2: RelaxedDCM for each distribution in the ϵ-net to find the non-id causal effect bound. Finally, Alg. 1: Explore ϵ-ball combines these bounds to construct $[\underline{L}_x, \overline{L}_x, \underline{U}_x, \overline{U}_x]$ to decide the next move.

Algorithm 2: Optimizing Relaxed-DCM: Given a fixed distribution P_ϵ, in Alg. 2, we train deep causal models (see Def. 3) to learn P_ϵ while optimizing the causal effects. Unlike existing works that attempt to match P_ϵ exactly, we use a Lagrangian duality-based optimization (Bui et al., 2022; Gao & Kleywegt, 2023) to keep the observational distribution entailed by the DCM P_θ, within a small region around input distribution, P_ϵ (Alg 2: line 7). The dual update parameter successfully finds the right amount of regularization and maintains the desired distributional distance (Alg 2: line 9). We add the causal effect magnitude with the distributional loss terms and optimize the parameters to maximize or minimize the effects(Alg 2: lines 8, 10).

Algorithm 3: Construct ϵ-net : Since the given dataset might be sampled from any distribution in the confidence set $\mathcal{C}$ of the empirical distribution $\hat{P}$, we need to solve the max-max, max-min, min-max and min-min problems (Equation 4, 1) by searching over $\mathcal{C}$. To obtain the confidence set in practice, we factorize the empirical distribution $\hat{P}(\mathbf{v})$ into conditional distributions, i.e., $\hat{P}(\mathbf{v}) = \Pi_{v\in\mathbf{v}}\hat{P}(v_i|v^{\pi_i-1})$ and construct non-asymptotic confidence interval for variable with m total number of configurations (exponential in alphabet size), from n samples w/ error probability $\alpha = 0.05m$, using Hoeffding inequality (Hoeffding, 1963) and Bonferroni correction, i.e., $\hat{P}_n(v_i = 1|v^{\pi_i-1}) \pm \sqrt{\frac{\ln(2/\alpha)}{2n}}$. See Appendix E.3 for derivation. If we pick one point from each interval to form a joint distribution and optimize the causal effect, we obtain the maximum and minimum effect for that specific distribution. To obtain the four quantities in Equation 4, 1, we have to search over all constructed intervals and optimize the causal effect while matching each formed distribution.

Note that solving a max-min and min-max problem by searching over these confidence intervals is not convex in general. To address this scenario, we resort to a heuristic approach motivated from ϵ-net (Gonzalez, 1985). To implement ϵ-net, we cover the metric space of observational distributions M consistent with the input dataset, with open balls having centers in $X \subset M$ such that every distribution in M is within ϵ distance of at least one distribution in X. For our purpose, we cover each confidence interval with smaller intervals of width ϵ_s. If ϵ is the interval width, this would give us $\lceil \epsilon/\epsilon_s \rceil$ number of centroids. We uniformly pick centroid for each conditional and form a joint distribution. After optimizing the DCM for these distributions, we take the minimum (or maximum) of all optimized values to approximate the minmax (or maxmin).

Algorithm 1: Explore ϵ-ball : After collecting $[L_x, U_x]$ bounds from each distribution in the confidence set, we follow Equation 4 or 1 to obtain the inner $[\overline{L}_x, \underline{U}_x]$ region and the outer $[\underline{L}_x, \overline{L}_x] \cup [\underline{U}_x, \overline{U}_x]$ band. If the bounds are completely separated for $do(x_0)$ and $do(x_1)$, we return the best action(Alg 1,line: 12). If the inner region intersects, we conclude with high probability that additional variables must be observed (Alg 1,line: 14). In this paper, we suggest observing instrument variables I since an instrument can reduce the non-id region (Balke & Pearl, 1996). Whether other variables can reduce the causal effect bounds is an open area to explore. In all other cases, we conclude collecting more samples (Alg 1,line: 16). The conclusion holds with probability at least $1 - \alpha$, provided that the ϵ-net adequately covers the confidence set and the neural networks are optimized well; otherwise, the resulting bounds may be inaccurate. See Appendix Section E for details.

5 EXPERIMENTAL ANALYSIS

In this section, we evaluate our algorithm in three simulated setups and on one real-world Parent labor supply dataset. We demonstrate how our algorithm performs with access to more samples and observed variables and how it adapts to large graphs. Finally, we analyze the behavior of our algorithm for mis-specified causal graphs (Section F.6). In all setups, we obtain $[\underline{L}_x, \overline{L}_x, \underline{U}_x, \overline{U}_x]$ by choosing ~ 100 distributions from the ϵ-net and training the neural nets on 5000 samples from each.

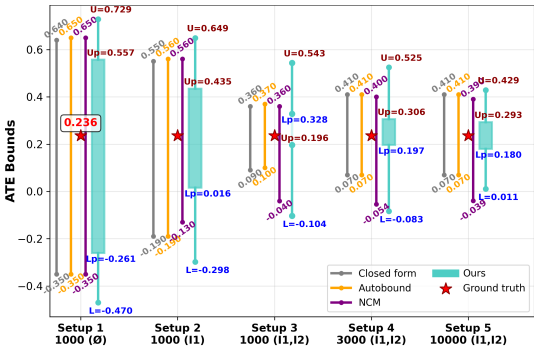


Figure 4: Bounds comparison for the multiple IV.

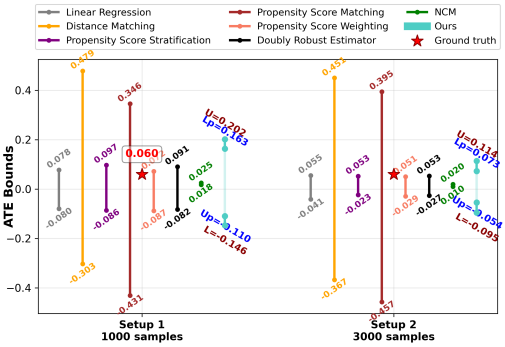


Figure 5: Bounds comparison in the ID setting.

5.1 Performance on Simulated Datasets

SCM 1 `UA-DCM` correctness & utility): In Figure 4, we show the next moves and consequences of our algorithm for different sample sizes. In setup 1, for $N = 1000$ samples of $D[X, Y]$, we observed that all three baselines, i.e., closedform, autobound and NCM output similar ATE bounds: $[-0.35, 0.64]$. They estimate $P(X, Y)$ from data and use it as the true distribution to calculate the bounds. These baselines only indicate that the true ATE lies within a large bound but provide no additional information on how to reduce their bounds to make the best action decision. On the other hand, in `UA-DCM`, since $\gamma \in [\overline{L}, \underline{U}]$, i.e., $0 \in [-0.26, 0.56]$, we conclude that we must *observe* additional variables. With probability at least $1-\alpha$, additional samples will not resolve this ambiguity.

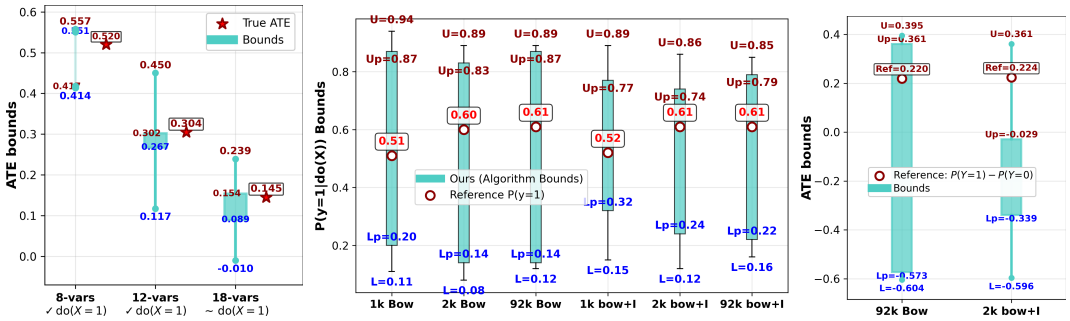
We illustrate the consequence of observing additional variables in setup 2 where we consider an additional instrument variable. We observe that for 1000 samples of $D[I_1, X, Y]$, all baselines bounds reduce to $[-0.19, 0.56]$. For `UA-DCM`, both the inner region and the outer band shrink, especially the inner bound from $[-0.26, 0.56] \rightarrow [0.02, 0.44]$, indicating the reduction in the non-id region. However, as the solid region $[\overline{L}, \underline{U}]$ does not contain 0, we do not know if the current uncertainty is due to unobserved confounders or low sample size. Thus, based on the feasibility of the setup, we have to select the next move. For our experiment, we conclude, *observing* another instrument I_2.

In setup 3, we show bounds obtained from all baselines and our algorithm `UA-DCM` for 1000 samples of the dataset $D[I_1, I_2, X, Y]$, and in setup 4, we show results when we collect 3000 samples. In setup 3, the bounds reduce significantly while changing slightly in setup 4. `UA-DCM` says that the outer band is $[-0.08, 0.20]$, which is very close to zero. Thus, collecting more samples will reveal that $do(X = 1)$ is a better action compared to $do(X = 0)$. Finally, we observe in setup 5 that with 10k samples and 90% confidence, the minmin value is above 0. Now, we can decide safely that $do(X = 1)$ is the best action for all distributions consistent with the finite input samples. We perform additional analysis in Appendix E with various confidence levels and various sample sizes.

SCM 2: We consider only observed confounders (Figure 7c) with no structural uncertainty. We compare with 6 baselines that are designed to provide point-wise estimations. To be fair, we perform bootstrapping and obtain bounds from them. We also compare with NCM which outputs ATE bound.

Results: We observe that for 1000 samples, distance matching and propensity score matching output a very wide bound, ex: $[-0.457, 0.395]$. Other baselines provide tighter bounds and NCM outputs $[0.018, 0.025]$. All of these methods estimate a probability table from the input 1000 samples and estimate the causal effect based on that. Even with bootstrapping, some of the bounds do not contain the true ATE$= 0.06$. As most of the baselines contain 0 within it, they suggest that no action can be decided as the better one. No other insight about uncertainty reduction can be extracted from them. On the other hand, `UA-DCM` provides an outer bound $[-0.14, 0.20]$ for 1000 samples and an outer bound $[-0.09, 0.11]$ for 3000 samples. This indicates that the ATE obtained from the true distribution will be contained in this bound. Also, note that there is no gap between $\underline{U}$ and $\overline{L}$ (Up,Lp in the figure), which implies that there is no non-id uncertainty here, hinting us to collect more samples.

Scaling Experiment: To scale our algorithm for large graphs, we utilize two practical solutions (details in Appendix F.4). 1. We factorize the joint distribution according to the causal graph and match the factorized conditional distributions instead of the full joint. 2. We ignore low-probability



(a) Performance on large graphs (b) $do(x)$ bounds on Parents Labor Supply (PLS). (c) ATE bounds for PLS .

Figure 6: UA−DCM performance on synthetic and real-world datasets.

joint combinations and redistribute their probability mass. This reduces the slicing of the alpha value to a large number of combinations that do not meaningfully change the causal effect.

Observation: In Figure 6a, we show the algorithm's performance on graphs with 8, 12, and 18 variables, each using $10k$ samples. For the 8-variable setup, pointwise identification of the causal effect collapses the inner region; thus, there is no non-identifiable region. For each scenario of 8-vars, 12-vars and 18-vars, our bounds enclosed the true ATE values and the outer bound was completely above zero, implying that $do(X = 1)$ is the best action for these problem instances. Thus, our algorithm scales to a relatively larg number of variables and offers bounds that enclose the true ATE.

5.2 Real Data Experiments: Parents' Labor Supply

Setup: We apply our algorithm on the Parents' Labor Supply dataset(Angrist & Evans, 1996) which contains contains 92k family records of their demographic and work details. Here, the goal is to evaluate the causal effect of having more than two children (treatment X:yes/no) on mother's labor supply (outcome Y: working or not) with respect to the population, i.e., $P(y = 1|\text{do}(x = 1))$ vs $P(y = 1)$. We follow (Angrist & Evans, 1996) and consider whether first two children are of same sex or not as an instrument variable I since this significantly motivates people to have the third child. Please see F.5 for further justification. We illustrate our algorithm performance on 1000, 2000 and 92k (full dataset) with and without an instrument.

Results: In Fig 6b, as we increase sample size from 1k to 92k, the inner region expands as $[0.20, 0.87] \rightarrow [0.14, 0.87]$, and the total width of the outer band shrinks as $0.16 \rightarrow 0.04$. However, $P(y = 1)$ is still contained in the non-id inner region. Thus, with high probability, collecting more samples will not help, and we include the sex of the first two children as an instrument. For 1000 samples, as we move from no instrument to one instrument (1k bow $\rightarrow$ 1k bow+I), the non-id region is reduced as $[0.20, 0.87] \rightarrow [0.32, 0.77]$. Although the bound does not exclude $P(y)$, it shows we need to keep observing variables to reach a conclusion. We compare the *ATE* bounds for the $92k$ bow and $2k$ bow + I setups in Figure 6c. The $92k$ samples eliminate almost all sample uncertainty, leaving only the non-identification region. In contrast, in the $2k$ bow + I setup, the non-id region $[-0.57, 0.36]$ decreases significantly to $[-0.33, 0.22]$ after including the instrumental variable. The outer band is extended due to limited samples. Our results are consistent with Angrist & Evans (1996), who applied Two-Stage Least Squares and found a negative impact: -0.12 on "worked for pay" and -5.68 on "weeks worked per year". Our ATE bounds also suggest such a negative effect.

6 Conclusion

In this paper, we proposed a novel decomposition of epistemic uncertainty in causal effect estimation from finite data. This decomposition can be used to conclude with high probability collecting more samples will not help with deciding the best action. We further develop a neural causal modeling approach to approximate this decomposition in practice. A limitation of the current method is computational scalability: the decomposition requires optimizing causal effect bounds over a confidence set of observational distributions, which we approximate with an ϵ-net. The search over candidate distributions can grow exponentially with the number of variables in the worst case. Developing more efficient optimization procedures is an important direction for future work.

AI USE STATEMENT

We used LLM-based tools in two ways: to polish the writing (grammar, wording, and formatting), and to help search for related work. LLMs were not used to develop the method, the theory, or the experiments. The authors read and checked every reference found this way, and reviewed all edited text. We take full responsibility for the content of this paper.

ETHICS STATEMENT

This work does not involve human subjects or the collection of new data. Our experiments use synthetic data and the publicly available Parents' Labor Supply dataset (Angrist & Evans, 2009). The main risk we see is misuse through a misspecified causal graph, which can lead to a wrong recommendation; we discuss this in Appendix A, and the broader impact of the work in Appendix B.

REPRODUCIBILITY STATEMENT

Assumptions and complete proofs of our theoretical results are given in Appendix D. The algorithms, their complexity, and implementation details are described in Appendix E. Causal graphs, data-generating processes, sample sizes, hyperparameters, and compute resources for all experiments are listed in Appendix F, with details for the real dataset in Appendix F.5. The synthetic data can be regenerated from these descriptions, and the real dataset is public. We will release our code.

REFERENCES

Joaquín Abellán, George J Klir, and Serafín Moral. Disaggregated total uncertainty measure for credal sets. *International Journal of General Systems*, 35(1):29–44, 2006.

Joshua Angrist and William N Evans. Children and their parents' labor supply: Evidence from exogenous variation in family size, 1996.

Joshua D. Angrist and William N. Evans. Replication data for: Children and their parents' labor supply: Evidence from exogenous variation in family size, 2009.

Vahid Balazadeh Meresht, Vasilis Syrgkanis, and Rahul G Krishnan. Partial identification of treatment effects with implicit generative models. *Advances in Neural Information Processing Systems*, 35: 22816–22829, 2022.

Alexander Balke and Judea Pearl. Counterfactual probabilities: Computational methods, bounds and applications. In *Uncertainty in artificial intelligence*, pp. 46–54. Elsevier, 1994.

Alexander Balke and Judea Pearl. Bounds on treatment effects from studies with imperfect compliance. *Journal of the American statistical Association*, 92(439):1171–1176, 1997.

Alexander Balke and Judea Pearl. Probabilistic evaluation of counterfactual queries. In *Probabilistic and causal inference: The works of Judea Pearl*, pp. 237–254. 2022.

Alexander A Balke and Judea Pearl. Universal formulas for treatment effects from noncompliance data. In *Lifetime data: models in reliability and survival analysis*, pp. 39–43. Springer, 1996.

Alexis Bellot, Alan Malek, and Silvia Chiappa. Transportability for bandits with data from different environments. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

Patrick Blöbaum, Peter Götz, Kailash Budhathoki, Atalanti A. Mastakouri, and Dominik Janzing. Dowhy-gcm: An extension of dowhy for causal inference in graphical causal models. *Journal of Machine Learning Research*, 25(147):1–7, 2024. URL `http://jmlr.org/papers/v25/22-1258.html`.

Blai Bonet. Instrumentality tests revisited. In *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence (UAI)*, pp. 48–55, 2001.

Tuan Anh Bui, Trung Le, Quan Tran, He Zhao, and Dinh Phung. A unified wasserstein distributional robustness framework for adversarial training. *arXiv preprint arXiv:2202.13437*, 2022.

Patrick Chao, Patrick Blöbaum, Sapan Patel, and Shiva Prasad Kasiviswanathan. Modeling causal mechanisms with diffusion models for interventional and counterfactual queries. *arXiv preprint arXiv:2302.00860*, 2023.

Victor Chernozhukov, Sokbae Lee, and Adam M Rosen. Intersection bounds: Estimation and inference. *Econometrica*, 81(2):667–737, 2013.

Brian Cho, Dominik Meier, Kyra Gan, and Nathan Kallus. Reward maximization for pure exploration: Minimax optimal good arm identification for nonparametric multi-armed bandits. *arXiv preprint arXiv:2410.15564*, 2024.

Guilherme Duarte, Noam Finkelstein, Dean Knox, Jonathan Mummolo, and Ilya Shpitser. An automated approach to causal inference in discrete settings. *Journal of the American Statistical Association*, 119(547):1778–1793, 2024.

Carlos Fernández-Loría and Foster Provost. Causal decision making and causal effect estimation are not the same… and why it matters. *INFORMS Journal on Data Science*, 1(1):4–16, 2022.

Dennis Frauen, Fergus Imrie, Alicia Curth, Valentyn Melnychuk, Stefan Feuerriegel, and Mihaela van der Schaar. A neural framework for generalized causal sensitivity analysis. *arXiv preprint arXiv:2311.16026*, 2023.

Dennis Frauen, Valentyn Melnychuk, Jonas Schweisthal, Mihaela van der Schaar, and Stefan Feuerriegel. Treatment effect estimation for optimal decision-making. *arXiv preprint arXiv:2505.13092*, 2025.

Yarin Gal et al. Uncertainty in deep learning. *phd thesis, University of Cambridge*, 2016.

Rui Gao and Anton Kleywegt. Distributionally robust stochastic optimization with wasserstein distance. *Mathematics of Operations Research*, 48(2):603–655, 2023.

Teofilo F Gonzalez. Clustering to minimize the maximum intercluster distance. *Theoretical computer science*, 38:293–306, 1985.

Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

Stephen C Hora. Aleatory and epistemic uncertainty in probability elicitation with an example from hazardous waste management. *Reliability Engineering & System Safety*, 54(2-3):217–223, 1996.

Yaowei Hu, Yongkai Wu, Lu Zhang, and Xintao Wu. A generative adversarial framework for bounding confounded causal effects. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 12104–12112, 2021.

Yimin Huang and Marco Valtorta. Identifiability in causal bayesian networks: A sound and complete algorithm. In *Proceedings of the national conference on artificial intelligence*, volume 21, pp. 1149. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2006.

Yuanhanqing Huang and Jianghai Hu. A bandit learning method for continuous games under feedback delays with residual pseudo-gradient estimate. In *2023 62nd IEEE Conference on Decision and Control (CDC)*, pp. 1207–1212. IEEE, 2023.

Eyke Hüllermeier and Willem Waegeman. Aleatoric and epistemic uncertainty in machine learning: An introduction to concepts and methods. *Machine learning*, 110(3):457–506, 2021.

Eyke Hüllermeier, Sébastien Destercke, and Mohammad Hossein Shaker. Quantification of credal uncertainty in machine learning: A critical analysis and empirical comparison. In *Uncertainty in Artificial Intelligence*, pp. 548–557. PMLR, 2022.

Andrew Jesson, Sören Mindermann, Uri Shalit, and Yarin Gal. Identifying causal-effect inference failure with uncertainty-aware models. *Advances in Neural Information Processing Systems*, 33: 11637–11649, 2020.

Wenlong Ji, Yihan Pan, Ruihao Zhu, and Lihua Lei. Multi-armed bandits with machine learning-generated surrogate rewards. *arXiv preprint arXiv:2506.16658*, 2025.

Ziwei Jiang and Murat Kocaoglu. Conditional common entropy for instrumental variable testing and partial identification. In *Forty-first International Conference on Machine Learning*, 2024.

Ziwei Jiang, Lai Wei, and Murat Kocaoglu. Approximate causal effect identification under weak confounding. In *International Conference on Machine Learning*, pp. 15125–15143. PMLR, 2023.

Nathan Kallus and Angela Zhou. Confounding-robust policy improvement. *Advances in neural information processing systems*, 31, 2018.

Murat Kocaoglu, Christopher Snyder, Alexandros G Dimakis, and Sriram Vishwanath. Causalgan: Learning causal implicit generative models with adversarial training. In *International Conference on Learning Representations*, 2018.

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

Sanghack Lee and Elias Bareinboim. Structural causal bandits with non-manipulable variables. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pp. 4164–4172, 2019.

Ang Li and Judea Pearl. Bounds on causal effects and application to high dimensional data. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pp. 5773–5780, 2022.

Ang Li, Scott Mueller, and Judea Pearl. Epsilon-identifiability of causal quantities. *arXiv preprint arXiv:2301.12022*, 2023.

Myrl G Marmarelis, Greg Ver Steeg, Aram Galstyan, and Fred Morstatter. Ensembled prediction intervals for causal outcomes under hidden confounding. In *Causal Learning and Reasoning*, pp. 18–40. PMLR, 2024.

Valentyn Melnychuk, Stefan Feuerriegel, and Mihaela van der Schaar. Quantifying aleatoric uncertainty of the treatment effect: a novel orthogonal learner. *Advances in Neural Information Processing Systems*, 37:105039–105089, 2024.

Min Woo Park, Andy Arditi, Elias Bareinboim, and Sanghack Lee. Structural causal bandits under markov equivalence. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

J. Pearl. *Causality*. Cambridge University Press, 2009. ISBN 9781139643986. URL `https://books.google.com/books?id=LLkhAwAAQBAJ`.

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

Judea Pearl. On the testability of causal models with latent and instrumental variables. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence (UAI)*, pp. 435–443, 1995b.

Judea Pearl. On the testability of causal models with latent and instrumental variables. *arXiv preprint arXiv:1302.4976*, 2013.

Md Musfiqur Rahman and Murat Kocaoglu. Modular learning of deep causal generative models for high-dimensional causal inference. In *Forty-first International Conference on Machine Learning*, 2024. URL `https://openreview.net/forum?id=bOhzU7NpTB`.

Md Musfiqur Rahman, Matt Jordan, and Murat Kocaoglu. Conditional generative models are sufficient to sample from any causal effect estimand. *Advances in Neural Information Processing Systems*, 37:77269–77315, 2024.

Michael C Sachs, Gustav Jonzon, Arvid Sjölander, and Erin E Gabriel. A general method for deriving tight symbolic bounds on causal effects. *Journal of Computational and Graphical Statistics*, 32(2): 567–576, 2023.

Murat Sensoy, Lance Kaplan, and Melih Kandemir. Evidential deep learning to quantify classification uncertainty. *Advances in neural information processing systems*, 31, 2018.

Amit Sharma and Emre Kiciman. Dowhy: An end-to-end library for causal inference. *arXiv preprint arXiv:2011.04216*, 2020.

Ilya Shpitser and Judea Pearl. Complete identification methods for the causal hierarchy. *Journal of Machine Learning Research*, 9:1941–1979, 2008.

Jin Tian. *Studies in causal reasoning and learning*. University of California, Los Angeles, 2002.

Jin Tian and Judea Pearl. Probabilities of causation: Bounds and identification. *Annals of Mathematics and Artificial Intelligence*, 28(1):287–313, 2000.

Jin Tian and Judea Pearl. A general identification condition for causal effects. In *Aaai/iaai*, pp. 567–573, 2002.

Lai Wei, Muhammad Qasim Elahi, Mahsa Ghasemi, and Murat Kocaoglu. Approximate allocation matching for structural causal bandits with unobserved confounders. *Advances in Neural Information Processing Systems*, 36:68810–68832, 2023.

Kevin Xia, Kai-Zhan Lee, Yoshua Bengio, and Elias Bareinboim. The causal-neural connection: Expressiveness, learnability, and inference. *Advances in Neural Information Processing Systems*, 34:10823–10836, 2021.

Kevin Xia, Yushu Pan, and Elias Bareinboim. Neural causal models for counterfactual identification and estimation. *arXiv preprint arXiv:2210.00035*, 2022.

Mengyue Yang, Furui Liu, Zhitang Chen, Xinwei Shen, Jianye Hao, and Jun Wang. Causalvae: Disentangled representation learning via neural structural causal models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 9593–9602, 2021.

Junzhe Zhang and Elias Bareinboim. Transfer learning in multi-armed bandits: A causal approach. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 1340–1346, 2017.

Junzhe Zhang and Elias Bareinboim. Bounding causal effects on continuous outcome. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 12207–12215, 2021a.

Junzhe Zhang and Elias Bareinboim. Non-parametric methods for partial identification of causal effects. *Columbia CausalAI Laboratory Technical Report*, 2021b.

Junzhe Zhang, Jin Tian, and Elias Bareinboim. Partial counterfactual identification from observational and experimental data. In *International Conference on Machine Learning*, pp. 26548–26558. PMLR, 2022.

Zhiheng Zhang. Tight partial identification of causal effects with marginal distribution of unmeasured confounders. In *Forty-first International Conference on Machine Learning*, 2024.

A Limitations and Future Works

We highlight the main limitations of our framework to help practitioners judge when its recommendations can be trusted.

Correct causal graph. Our framework assumes a correctly specified ADMG (Assumption 3). If the graph omits edges that affect identifiability, the framework may misclassify a non-identifiable query as identifiable and consequently produce incorrect predictions.

Conservative confidence set. The Hoeffding and Bonferroni based confidence set (Sec. 4.3) trades tightness for coverage. As a result, it can underestimate the non-identifiable region and delay the "Observe" recommendation when the query is non-identifiable. To mitigate this, we provide a formal analysis of how many additional samples suffice for the predicted non-identifiable region to recommend "Observe", which gives practitioners guidance on the sample complexity required to recover from such underestimation. Moreover, distributions in the confidence set that are far from the empirical distribution may violate the constraints induced by the causal graph (e.g., the instrumental inequalities); for the IV graph, we discard such distributions before exploration, while enforcing these constraints for general graphs with latent variables remains an open problem (Appendix E.6).

Optimization of the bounds. Performing the effect bound decomposition requires solving challenging min-max and max-min problems, which we approach using an ϵ-net. The size of this net grows with the desired resolution, adding computational cost. Moreover, neural network optimization may be imperfect in practice and yield bounds tighter than the true ones.

Future work. We aim to resolve the above limitations i.e. develop more efficient min-max optimization procedures, reduce the complexity of the ϵ-net, and relax the reliance on a fully specified causal graph.

B Broader Impact

In this paper, we propose an algorithm that can determine when with high probability collecting more samples will not help determine the best action. This can guide practitioners to realize when increasing the number of participants in the current setup, i.e., collecting additional data samples is not being helpful anymore. Our algorithm will advice to collect more variables or lean towards a randomized study for best action identification. As a result, our algorithm can be used as a too for strategic planning of randomized control study and reduce cost for experimentation. Additionally, our work can advance the field of AI by improving machine learning methods and make them fair by providing the causal analysis of different actions.

C Related Work

Partial Identification of Causal Effect. When the causal effect is not identifiable from observational data, one can identify a set of causal effects from the SCMs that are compatible with the observational data. For the discrete variables, Balke & Pearl (1994; 2022) proposed a method using response variables to partition the latent variable into finite states and estimate the bounds of causal queries. Tian & Pearl (2000) used the response variable to get bounds of counterfactual queries from observational and interventional data. Many works focus on getting narrower bounds by utilizing additional information from the graph. Balke & Pearl (1997) obtained tighter bounds with instrumental variables. Duarte et al. (2024) proposed an automated method to derive bounds for causal queries in arbitrary graphs.

There are works that are closely related to ours utilizing a posterior sampling approach for partial identification. Zhang & Bareinboim (2021b) introduces canonical nonparametric SCMs for bounding causal effects in arbitrary causal diagrams with categorical observed variables. Zhang et al. (2022) extend this approach to counterfactual queries using observational and experimental data, and provide MCMC-based credible intervals for approximating partial identification bounds. These works are complementary to ours: they provide finite sample guarantees for the MCMC approximation of

credible intervals conditioned on the observed data, whereas our framework studies the decision problem induced by the finite sample uncertainty. This line of work also suggests an alternative but related decision-making problem: given a limited sample budget, determine whether the available data are sufficient to reach an unambiguous decision or whether additional samples within the budget are unlikely to resolve the ambiguity. Developing such a criterion requires determining the number of samples needed for the estimation to reach a given bound width and propagating finite-sample uncertainty through partial identification bounds. This might be an interesting future direction connecting the posterior sampling method with the decision problem.

Another line of work tries to obtain narrower bounds of causal effect utilizing information from the unobserved confounders. Li & Pearl (2022) proposed using nonlinear programming to get bounds when the latent confounder is partially observed. Jiang et al. (2023) proposed a convex programming formulation to get tighter bounds of the causal effect when the entropy of the latent confounder is known. Li et al. (2023) derived bounds in closed form when the marginal distribution of the latent confounders is known. Zhang (2024) proposed a method that finds the tight identifiable region of causal effect given the marginal distribution of latent confounders.

Causal Inference with Neural Models. Many researchers have explored the use of neural networks in causal inference. Kocaoglu et al. (2018) introduced CausalGAN that produces interventional images with a GAN trained with image data. Yang et al. (2021) proposed CausalVAE to learn the causally related representations from the data. Xia et al. (2021; 2022) proposed neural causal models that find the causal and counterfactual queries by maximizing and minimizing the query under the semi-Markovian setting. Chao et al. (2023) proposed a diffusion-based approach to model the causal and counterfactual queries. Rahman & Kocaoglu (2024) proposed a modularized training algorithm to train a causal generative model with latent confounders for high-dimensional variables. Rahman et al. (2024) proposed a recursive algorithm that uses a set of conditional generative models to sample from any identifiable interventional distributions.

Uncertainty Quantification. Uncertainty quantification is a crucial problem in deep learning (Gal et al., 2016), especially in applications where reliable prediction is essential. Hora (1996) introduced the notion of aleatoric and epistemic to distinguish uncertainties for the risk models in applications such as decision making. Abellán et al. (2006) defined the total uncertainty based on the Shannon entropy of a credal set. Sensoy et al. (2018) quantified the epistemic uncertainty of the classification model with the Dirichlet prior. In general, the uncertainty is assumed to have an additive representation between the total uncertainty (TU), epistemic uncertainty (EU), and aleatoric uncertainty (AU), and it is generally hard to disentangle the EU and AU (Hüllermeier & Waegeman, 2021; Hüllermeier et al., 2022). Melnychuk et al. (2024) proposed a method for quantifying aleatoric uncertainty in individualized treatment effects by deriving sharp bounds on the conditional distributions of the treatment effect. Marmarelis et al. (2024) introduced an approach for constructing causal effect intervals with hidden confounding. Jiang & Kocaoglu (2024) proposed using entropy as a sensitivity parameter to obtain bounds in the IV graph with assumption violations.

Decision Making. Decision-making is highly related to the bandit problem, which aims to find the action that optimizes the rewards (Cho et al., 2024; Huang & Hu, 2023; Ji et al., 2025). In the context of causal inference, although the causal effect is closely related to the decision-making problem, Fernández-Loría & Provost (2022) suggested that they are not exactly equivalent, where the main distinction is their estimand of interest. Wei et al. (2023) proposed an algorithm that utilizes the causal graph to balance exploration and exploitation in sequential decision-making problems with latent confounders. Kallus & Zhou (2018) studied the decision-making problem with observational data with the presence of latent confounders. They showed that with a well-specified uncertainty set, the proposed robust policy learning is no worse than the baseline. Jesson et al. (2020) introduced a method to assess the uncertainty in the neural network model, and demonstrated that the proposed method is able to handle the positivity violation. Sensitivity analysis is also closely related to the uncertainty quantification in causal inference. Frauen et al. (2023) proposed a neural framework for sensitivity analysis that is compatible with a large class of existing sensitivity models. Jesson et al. (2020) studied the decision-making problem under the non-overlapping situation. Frauen et al. (2025) propose a optimal decision-making approach based on two-stage CATE estimators.

Causal Bandits. A line of work uses offline causal knowledge to derive bounds for actions first and then proceed to the online phase.Zhang & Bareinboim (2017) derive two-sided bounds on each arm's expected reward from a source agent's causal model and batch data, and feed these bounds into a

modified UCB algorithm (B-kl-UCB) that filters and truncates arms accordingly. Bellot et al. (2023) instead use a known selection diagram, i.e., a causal graph annotated with where source and target environments differ, to build an informative Bayesian prior over reward distributions from batch source data, which then seeds posterior sampling exploration in the target environment. A related line studies where in the causal graph to intervene: Lee & Bareinboim (2018) introduce structural causal bandits, which use graphical criteria derived from do-calculus (possibly-optimal minimal intervention sets, or POMIS) to prune the arm set before play begins; Lee & Bareinboim (2019) extend this framework to non-manipulable variables, Park et al. (2025) generalize it from a known DAG to a Markov equivalence class (PAG/MAG), and Wei et al. (2023) build directly on this framework. Our framework operates one stage earlier than these works. Rather than pruning or selecting arms for an online phase, it determines whether an online phase can resolve the ambiguity in the first place. Given a fixed action, it decomposes why an effect bound is wide, separating finite-sample noise from structural non-identifiability, and issues one of three recommendations: return the best action (RETURN), collect more samples of the same variables (COLLECT), or measure additional variables (OBSERVE). Notably, Park et al. (2025) identify finite-sample uncertainty in causal bounds as future work, which is precisely the source of uncertainty that our non-identifiability versus finite-sample decomposition addresses.

D THEORETICAL ANALYSIS

Assumption 1. *We assume the variables are discrete.*

In the binary case, confidence intervals are defined over single conditional probabilities, which simplifies the exploration within the ϵ-ball. For a non-binary variable X with $|\mathcal{X}|$ states, there will be $|\mathcal{X}| - 1$ confidence intervals due to the simplex constraint. Constructing valid distributions from these confidence intervals becomes a difficult constrained optimization problem.

Assumption 2. *The structural causal model is Semi-Markovian.*

Assumption 3. *We have access to the acyclic directed mixed graph (ADMG).*

D.1 DECISION FOR MULTIPLE ACTIONS

Lemma 1. *Let L_i, U_i be the upper bound and lower bound of ATE(x_0, x_i). Action x_0 is the best action if and only if $L_i \geq 0 \; \forall i \neq 1$.*

Proof. $\implies$ is clear from the definition.

$\impliedby$: Suppose for the sake of contradiction that there exists some SCM such that ATE$(x_0, x_j) < 0$. But this contradicts the assumption that $L_i \geq 0 \; \forall i \neq 1$. □

Lemma 2. *If there exists any x_i such that $U_i > 0$ and $L_i < 0$, then both x_0 and x_i cannot be the best action.*

Proof. If there exists any xi such that $U_i > 0$ and $L_i < 0$, that means there exist two SCM $S1$, $S2$ such that ATE$(x_0, x_i) > 0$ in $S1$, and ATE$(x_0, x_i) < 0$ in $S2$. So both x_0 and x_i cannot be determined to be the best action.

□

To prove our main theorem, we need the following lemma.

Lemma 3. *Let $V_1, \ldots, V_n$ be the set of observed variables in the topological order.*

Suppose the observed variables can be partitioned into groups $D_1, \ldots, D_m$, such that variables within each group may share one common exogenous noise, while exogenous noises associated with distinct groups are mutually independent.

For any compatible observational distribution P, the SCMs entailing P admit a response function representation whose feasible parameter set is nonempty, compact, and convex.

Proof. For each variable V_i, let r_i be a deterministic function assigning a value of V_i to each configuration of its observed parents.

For group D_j, let $R_{D_j} = (R_i : i \in D_j)$ be its collection of response variables with joint distribution $q_j(r_{D_j}) = P(R_{D_j} = r_{D_j})$.

Since the exogenous noise are mutually independent across groups, the response variable vectors are independent

$$P(r_{D_1}, r_{D_2}, \ldots, r_{D_m}) = \prod_{j=1}^{m} q_j(r_{D_j}).$$

For a fixed $v_{1:k} = (v_1, \ldots, v_k)$, define

$$H_{j,k}(v_{1:k}; q_j) := \sum_{r_{D_j}} q_j(r_{D_j}) \prod_{\substack{i \le k \\ i \in D_j}} \mathbf{1}\left\{r_i(v_{pa(i)}) = v_i\right\}.$$

By construction, $H_{j,k}(v_{1:k}; q_j)$ is linear in q_j.

Since $V_1, \ldots, V_n$ are in topological order, all parents of V_i occur before V_i. Therefore, the event $V_{1:k} = v_{1:k}$ occurs if and only if, for every group D_j, all response functions in that group corresponding to variables among $V_1, \ldots, V_k$ agree with the configuration $v_{1:k}$. Hence, by independence across groups,

$$P_q(v_{1:k}) = \prod_{j=1}^{m} H_{j,k}(v_{1:k}; q_j), \tag{2}$$

where P_q denotes the observational distribution induced by $q = (q_1, \ldots, q_m)$.

Let $d(k)$ denote the unique group containing V_k. When we extend from $v_{1:k-1}$ to $v_{1:k}$, only the factor corresponding to $D_{d(k)}$ changes. Thus, whenever $P(v_{1:k-1}) > 0$, requiring $P_q = P$ implies

$$\frac{H_{d(k),k}(v_{1:k}; q_{d(k)})}{H_{d(k),k-1}(v_{1:k-1}; q_{d(k)})} = P(v_k \mid v_{1:k-1}),$$

or equivalently,

$$H_{d(k),k}(v_{1:k}; q_{d(k)}) = P(v_k \mid v_{1:k-1}) H_{d(k),k-1}(v_{1:k-1}; q_{d(k)}). \tag{3}$$

Since $P(v_k \mid v_{1:k-1})$ is fixed, and both H terms are linear in $q_{d(k)}$, each constraint in (3) is linear in a single group distribution.

We now show that these constraints are also sufficient. Suppose that (3) holds for every k and every prefix with $P(v_{1:k-1}) > 0$. We prove by induction on k that

$$P_q(v_{1:k}) = P(v_{1:k}).$$

The claim is trivial for the empty prefix. Assume it holds for $v_{1:k-1}$.

If $P(v_{1:k-1}) > 0$, then by (2) and (3),

$$\begin{aligned} P_q(v_{1:k}) &= P_q(v_{1:k-1}) \frac{H_{d(k),k}(v_{1:k}; q_{d(k)})}{H_{d(k),k-1}(v_{1:k-1}; q_{d(k)})} \\ &= P(v_{1:k-1}) P(v_k \mid v_{1:k-1}) \\ &= P(v_{1:k}). \end{aligned}$$

If $P(v_{1:k-1}) = 0$, then by the induction hypothesis $P_q(v_{1:k-1}) = 0$, and hence every extension $v_{1:k}$ also has probability zero under both P_q and P. Therefore $P_q = P$.

For each group D_j, let

$$\mathcal{Q}_j(P) := \left\{ q_j : q_j(r_{D_j}) \ge 0,\ \sum_{r_{D_j}} q_j(r_{D_j}) = 1,\ q_j \text{ satisfies all constraints (3) involving } D_j \right\}.$$

Each $\mathcal{Q}_j(P)$ is the intersection of a finite-dimensional probability simplex with finitely many linear equalities. Hence $\mathcal{Q}_j(P)$ is a compact convex polytope.

Moreover, the preceding necessity and sufficiency argument shows that the full feasible response-function parameter set is

$$\mathcal{Q}(P) = \prod_{j=1}^{m} \mathcal{Q}_j(P).$$

Since P is assumed compatible with the SCM class, $\mathcal{Q}(P)$ is nonempty. A finite Cartesian product of compact convex sets is compact and convex. Therefore $\mathcal{Q}(P)$ is nonempty, compact, and convex.

□

Proof of Theorem 1

Theorem 1. *Let $f(P, S)$ be the estimand of an identifiable causal effect for SCM S, with the observational distribution P. Let $\mathcal{P}$ be any connected set of observational distributions that contains P. Let $a = \min_{p\in\mathcal{P}} f(P, S), b = \max_{P\in\mathcal{P}} f(P, S)$. Then $f|_{\mathcal{P}} : \mathcal{P} \to [a, b]$ is surjective.*

Proof. Consider any connected set of observed distributions $\mathcal{P}$. For any $P \in \mathcal{P}$, let $\bar{P} \in \bar{\Delta}$ be the distribution including the latent variables U, such that $\sum_U \bar{P} = P$. Assume the causal query is well defined, i.e. $P(pa(y)) > 0$.

Let $adj(X, Y)$ be the set of variables that are the parents of Y but not descendants of X. Then

$$\begin{aligned} h(\bar{P}, G) &= P(y|do(x)) \\ &= \sum_{adj(X,Y)} \bar{P}(y|x, adj(X, Y))\bar{P}(adj(X, Y)) \\ &= \frac{\sum_{v\notin\{x,y,adj(X,Y)\}} \bar{P}(v)}{\sum_{v\notin\{x,adj(X,Y)\}} \bar{P}(v)} \left(\sum_{v\notin\{adj(X,Y)\}} \bar{P}(v) \right). \end{aligned}$$

Since the causal effect is a rational function of $\bar{P}$, it is continuous when the denominator is strictly positive.

Let $f(P, G)$ be the function that maps $P \in \mathcal{P}$ to some causal effect in $[0, 1]$ such that $f(P, G) = h(\bar{P}, G)$ for all $\bar{P}$ such that $\sum_u \bar{P} = P$. Since the query is identifiable, the value of $h(\bar{P}, G)$ is identical for all $\{\bar{P} \mid \sum_U \bar{P} = P, P \in \mathcal{P}\}$. So f is well defined, and $h = f \circ g$.

Now for any open subset $M \subseteq [0, 1]$, since h is continuous, the preimage $h^{-1}(M) \subseteq \{\bar{P} \mid \sum_U \bar{P} = P, P \in \mathcal{P}\}$ is open. Let g be the function mapping full joint distribution to observed distribution, i.e., $g : \bar{P} \mapsto P$. By our construction, for any $P \in \mathcal{P}$, $g : h^{-1}(M) \to P$ is a linear surjective function. With the relative topologies on the simplices, g is an open map. Since $h^{-1}(M)$ is open, $g(h^{-1}(M))$ is open. For any open set $M \subseteq [0, 1]$, $f^{-1}(M) = g(h^{-1}(M))$ is open, so the function f is continuous.

By the generalized intermediate value theorem, the image of $f(\mathcal{P}, G)$ is connected. Since the image is a subset of $[0, 1]$, it is an interval that contains $[a, b]$.

□

Proof of Theorem 2

Theorem 2. *Let $f(p, S)$ be the estimand of some scalar non-identifiable causal effect of interest for a finite-state causal graph satisfying the conditions of Lemma 3, and let p be a compatible observational distribution. Let $\mathcal{S}$ be the set of the SCM such that $P_{obs}(S) = p$. Let $a = \min_{S\in\mathcal{S}} f(p, S), b = \max_{S\in\mathcal{S}} f(p, S)$. Then $f|_p : \mathcal{S} \to [a, b]$ is surjective.*

Proof. Suppose the causal query of interest is $P(y|do(x))$. Consider the graph G that includes both observed and unobserved latent variables. Let $\mathcal{P}$ be the set of distributions over the graph G. For

each $\bar{P} \in \mathcal{P}$, let the set of $\bar{P}$ denote the underlying SCM, i.e. $S = \bar{P}$ for some $\bar{P} \in \mathcal{P}$. Assume the causal query is well defined, i.e. $P(pa(y)) > 0$. Then

$$f(p, S) = P(y|do(x)) = \sum_{pa(y)} \bar{P}(y|pa(y))\bar{P}(pa(y)) = \frac{\sum_{v \notin \{x,y,pa(y)\}} \bar{P}(v)}{\sum_{v \notin \{x,pa(y)\}} \bar{P}(v)} \left(\sum_{v \notin \{pa(y)\}} \bar{P}(v) \right).$$

So the causal effect is a rational function of the full distribution over G. Since the causal query is well defined, the denominator is strictly positive. Thus $f(p, S)$ is continuous over all $\bar{P} \in \mathcal{P}$. For each $\bar{P} \in \mathcal{P}$, the only set of constraints are in form of $\sum_{u \in U} \bar{P}(v, u) = P(v)$. The set $\mathcal{P}$ is connected. Therefore, the image of $f(p, \cdot)$ is connected, and thus it is an interval. Since there is exist $f(p, S) = a$ and $f(p, S) = b$, and the image of $f(p, S)$ is connected, $[a, b] \subseteq f(p, \mathcal{S})$. so the map $f(p, \cdot)$ is surjective.

□

Proof of Proposition 1

Proposition 1. *Let $P^\star$ be the true observational distribution, $\mathcal{C}(D)$ a confidence set satisfying $\Pr(P^\star \in \mathcal{C}(D)) \geq 1 - \alpha$, and $\mathcal{I}_P(x)$ the causal-effect identified set for observational distribution P. On the coverage event $P^\star \in \mathcal{C}(D)$, (i) $\bigcap_{P \in \mathcal{C}(D)} \mathcal{I}_P(x) \subseteq \mathcal{I}_{P^\star}(x)$; and (ii) if the inner region crosses the decision boundary, i.e., $L_{\text{inner}} < 0 < U_{\text{inner}}$, then, with probability at least $1 - \alpha$, the population identified set $\mathcal{I}_{P^\star}(x)$ also crosses zero.*

Proof. Recall the definition 7, $\mathcal{I}_P = \{ATE(x) \mid S \in \mathcal{S}_P\}$, and $\mathcal{S}_P = \{S \text{ an SCM} \mid P_{obs}(S) = P, S \vDash G\}$. Suppose the coverage event $P^\star \in \mathcal{C}(D)$ holds, i.e., $\{P^\star\} \subseteq \mathcal{C}(D)$.

For (i), since an intersection over a larger collection of sets is smaller, we have $\bigcap_{P \in \mathcal{C}(D)} \mathcal{I}_P(x) \subseteq \bigcap_{P \in \{P^\star\}} \mathcal{I}_P(x) = \mathcal{I}_{P^\star}(x)$.

For (ii), if $L_{\text{inner}} < 0 < U_{\text{inner}}$, the inner region contains a value below zero and a value above zero. By (i), both values belong to $\mathcal{I}_{P^\star}(x)$, so $\mathcal{I}_{P^\star}(x)$ also crosses zero. Since the coverage event holds with probability at least $1 - \alpha$, so does this conclusion.

Note that, unlike a comparison between confidence sets obtained from two datasets, this argument does not require their empirical distributions to be identical or the confidence sets to be nested. □

On the arrival of new datasets: We emphasize that the conclusion in Proposition 1 concerns the fixed population quantity $\mathcal{I}_{P^\star}(x)$, not the data-dependent inner region: the inner region computed from a future dataset need not be monotone in the sample size, and need not cross zero again, but this does not change $\mathcal{I}_{P^\star}(x)$. Consequently, the argument does not require future empirical distributions or confidence sets to be identical or nested.

Proof of Theorem 3

Theorem 3. *Given a confidence set $\mathcal{C}$ and ATE values $\mathcal{I}_P$, $\bigcap_{P \in \mathcal{C}} \mathcal{I}_P = [\overline{L_x}, \underline{U_x}]$ if not empty.*

Proof. For each $P \in \mathcal{C}$, write

$$\mathcal{I}_P = [\ell(P), u(P)],$$

where

$$\ell(P) := \min_{S:\, P_{\text{obs}}(S) = P} \text{ATE}_S(x), \qquad u(P) := \max_{S:\, P_{\text{obs}}(S) = P} \text{ATE}_S(x).$$

Assume that

$$\bigcap_{P \in \mathcal{C}} \mathcal{I}_P$$

is nonempty. Since the intersection of intervals is again an interval, we may write

$$\bigcap_{P \in \mathcal{C}} \mathcal{I}_P = [\underline{c}, \overline{c}].$$

We first show that

$$\underline{c} = \sup_{P \in \mathcal{C}} \ell(P).$$

Let

$$\bar{L} := \sup_{P \in \mathcal{C}} \ell(P).$$

Suppose first, for contradiction, that $\bar{L} < \underline{c}$. Choose

$$\theta \in (\bar{L}, \underline{c}).$$

Since $\theta < \underline{c}$, we have

$$\theta \notin \bigcap_{P \in \mathcal{C}} \mathcal{I}_P.$$

Therefore, there exists some $P_0 \in \mathcal{C}$ such that

$$\theta \notin \mathcal{I}_{P_0} = [\ell(P_0), u(P_0)].$$

Hence either $\theta < \ell(P_0)$ or $u(P_0) < \theta$.

If $\theta < \ell(P_0)$, then

$$\bar{L} = \sup_{P \in \mathcal{C}} \ell(P) \geq \ell(P_0) > \theta,$$

contradicting $\theta > \bar{L}$.

If $u(P_0) < \theta$, then since $\theta < \underline{c}$, we have

$$u(P_0) < \underline{c}.$$

Thus $\underline{c} \notin [\ell(P_0), u(P_0)]$, contradicting the fact that

$$\underline{c} \in \bigcap_{P \in \mathcal{C}} \mathcal{I}_P.$$

Therefore, $\bar{L} \geq \underline{c}$.

Conversely, since $\underline{c}$ belongs to every interval $\mathcal{I}_P$, we have

$$\ell(P) \leq \underline{c} \qquad \text{for all } P \in \mathcal{C}.$$

Taking the supremum over $P \in \mathcal{C}$ gives

$$\bar{L} = \sup_{P \in \mathcal{C}} \ell(P) \leq \underline{c}.$$

Therefore,

$$\underline{c} = \sup_{P \in \mathcal{C}} \ell(P).$$

Similarly, we show that

$$\overline{c} = \inf_{P \in \mathcal{C}} u(P).$$

Let

$$\bar{U} := \inf_{P \in \mathcal{C}} u(P).$$

Since $\overline{c}$ belongs to every interval $\mathcal{I}_P$, we have

$$\overline{c} \leq u(P) \qquad \text{for all } P \in \mathcal{C}.$$

Thus

$$\overline{c} \leq \inf_{P \in \mathcal{C}} u(P) = \bar{U}.$$

Now suppose, for contradiction, that $\overline{c} < \bar{U}$. Choose

$$\theta \in (\overline{c}, \bar{U}).$$

Since $\theta > \overline{c}$, we have

$$\theta \notin \bigcap_{P \in \mathcal{C}} \mathcal{I}_P.$$

Therefore, there exists some $P_0 \in \mathcal{C}$ such that

$$\theta \notin \mathcal{I}_{P_0} = [\ell(P_0), u(P_0)].$$

Hence either $\theta < \ell(P_0)$ or $u(P_0) < \theta$.

If $\theta < \ell(P_0)$, then since $\overline{c} < \theta$, we have

$$\overline{c} < \ell(P_0),$$

so $\overline{c} \notin [\ell(P_0), u(P_0)]$, contradicting

$$\overline{c} \in \bigcap_{P \in \mathcal{C}} \mathcal{I}_P.$$

If $u(P_0) < \theta$, then

$$\bar{U} = \inf_{P \in \mathcal{C}} u(P) \leq u(P_0) < \theta,$$

contradicting $\theta < \bar{U}$.

Therefore,

$$\overline{c} = \inf_{P \in \mathcal{C}} u(P).$$

Combining the two endpoint identities gives

$$\bigcap_{P \in \mathcal{C}} \mathcal{I}_P = \left[\sup_{P \in \mathcal{C}} \ell(P), \inf_{P \in \mathcal{C}} u(P) \right].$$

Substituting the definitions of $\ell(P)$ and $u(P)$, we obtain

$$\bigcap_{P \in \mathcal{C}} \mathcal{I}_P = \left[\sup_{P \in \mathcal{C}} \min_{S:\, P_{\text{obs}}(S) = P} \text{ATE}_S(x),\ \inf_{P \in \mathcal{C}} \max_{S:\, P_{\text{obs}}(S) = P} \text{ATE}_S(x) \right].$$

Therefore,

$$\bigcap_{P \in \mathcal{C}} \mathcal{I}_P = [L_{\text{ate}_x}, U_{\text{ate}_x}].$$

If the extrema over $P \in \mathcal{C}$ are attained, the supremum and infimum above can equivalently be written as a maximum and a minimum, respectively. □

Proof of Corollary 1

Corollary 1. *Given a confidence set, $\mathcal{C}$ and a set of average treatment effect values, $\mathcal{I}_P$ (Def. 7),*

$$\bigcup_{P \in \mathcal{C}} \mathcal{I}_P \setminus \bigcap_{P \in \mathcal{C}} \mathcal{I}_P = [\underline{L_x}, \overline{L_x}] \cup [\underline{U_x}, \overline{U_x}].$$

Proof. Since $\bigcap_{P \in \mathcal{C}} \mathcal{I}_P = [\overline{L}_{ate_x}, \underline{U}_{ate_x}]$, we only need to show that $\bigcup_{P \in \mathcal{C}} \mathcal{I}_P = [\underline{L}_{ate_x}, \overline{U}_{ate_x}]$. Let $m_P = \min \mathcal{I}_P = \min_{S \in \{S \text{ an SCM} | P_{\text{obs}}(S) = P\}} ATE(x)$ and $\mathcal{M} = \{m_P \mid P \in \mathcal{C}\}$. Then $\min \bigcup_{P \in \mathcal{C}} \mathcal{I}_P = \min \mathcal{M} = \min \min_{S \in \{S \text{ an SCM} | P_{\text{obs}}(S) = P\}} ATE(x) = \underline{L}_{ate_x}$. Similarly, by a symmetric argument, we have $\max \bigcup_{P \in \mathcal{C}} \mathcal{I}_P = \overline{U}_{ate_x}$ □

Proof of Theorem 4

Theorem 4. *For a causal decision-making problem with $ATE(x)$ as the decision metric, let $(\overline{U_x}, \underline{U_x}, \overline{L_x}, \underline{L_x})$ be estimated from the data. The decision of $X = x$ as the best action (or cannot be the best action) is unambiguous if $\underline{L_x} \geq 0$ (or $\overline{U_x} < 0$). The decision is ambiguous and, with probability at least $1 - \alpha$, cannot be improved with more data if $\overline{L_x} < 0 < \underline{U_x}$.*

Proof. By the definition of $ATE(x)$, for a fixed SCM, an action x is the best action if $ATE(x) \geq 0$. For a fixed distribution P, $\mathcal{I}_P$ corresponds to the values of $ATE(x)$ with the set of SCMs that is compatible with P, where each SCM corresponds to a unique value of $ATE(x)$. And if $ATE(x) \geq 0$ for some x, the action x is the optimal action.

For a confidence set $\mathcal{C}$, if $\underline{L_x} > 0$, then for each $P \in \mathcal{C}$ and all the SCMs that are compatible with P, we have $ATE(x) \geq \underline{L_x} > 0$. Then all $\mathcal{I}_P$ lie above zero, and therefore $\bigcup_{P\in\mathcal{C}} \mathcal{I}_P = [\underline{L_x}, \overline{U_x}]$ is above zero, i.e., $\underline{L_{ate_x}} \geq 0$, and the action x is better than any other actions in all SCMs that are compatible with all distributions in the confidence set.

Therefore, x is the best action unambiguously. Similarly, if $\overline{U_x} < 0$, then for all $P \in \mathcal{C}$ and all the SCMs that are compatible with P, we have $ATE(x) < 0$, which indicates that x is not the best action in any of these SCMs. If $\overline{L_x} < 0 < \underline{U_x}$, by Theorem 3, for all $P \in \mathcal{C}$, the $\mathcal{I}_P$ contains both positive and negative values, so there exist two SCMs S_1 and S_2 such that x is the best action in S_1, but not the best action in S_2, and the decision is ambiguous. Moreover, on the coverage event $P^\star \in \mathcal{C}$, which holds with probability at least $1-\alpha$, this is true in particular for $P^\star$, i.e., S_1 and S_2 can be chosen to entail $P^\star$. Then S_1 and S_2 are compatible with any confidence set that contains $P^\star$, including one constructed from a larger dataset. Therefore, with probability at least $1-\alpha$, the best action remains ambiguous with more data. □

D.2 Decisions for Single Action

In many scenarios, instead of comparing effects of two or more actions, we need to compare the effect of a single action, $P(y|\text{do}(x))$ with a scalar quantity such as the marginal $P(y)$. Suppose we need to make the critical decision of choosing an action $X = x$ for the population level to improve an outcome. We can decide that by comparing the bounds of $P(Y \mid do(x))$ with the observational probability $P(Y)$. Specifically, action $X = x$ is beneficial or harmful if the interval $[\min P(Y \mid do(x)), \max P(Y \mid do(x))]$ is either completely above or below $P(Y)$. Otherwise, we need to collect or observe additional variables until we can make that decision. In this section, we consider all bounds to be constructed from $P(Y|do(x))$.

Backdoor Example: For example, let us consider the backdoor graph example as shown in the top row of Figure 3. Let $\hat{P}(y|x,z_i)$ and $\hat{P}(z)$ be the empirical estimation from the data and l_i^y and u_i^y be the confidence intervals of the conditional probabilities $\hat{P}(y|x,z_i)$. Similarly, l_i^z and u_i^z are the confidence intervals of the conditional probabilities $\hat{P}(z_i)$. The for $\hat{P}$ in the confidence set described by the above constraint, the maximum $P(y|\text{do}(x))$ can be estimated by the following (left):

$$\begin{aligned} &\max P(y|do(x)) = \sum_i \hat{P}(y|x,z_i)\hat{P}(z_i) \\ \text{s.t.}\quad &\hat{P}(y|x,z_i) \in [l_i^y, u_i^y]; \forall_i \\ &\hat{P}(z_i) \in [l_i^z, u_i^z]; \forall_i \\ &\sum_i \hat{P}(z_i) = 1 \end{aligned} \quad \Longrightarrow \quad \begin{aligned} &\max P(y|do(x)) = \sum_i u_i^y \hat{P}(z_i) \\ \text{s.t.}\quad &\hat{P}(z_i) \in [l_i^z, u_i^z]; \forall_i \\ &\sum_i \hat{P}(z_i) = 1 \end{aligned}$$

The objective function in the above optimization is a quadratic function, which could be nontrivial to solve. But in this case, since the $a_i \in [l_i^y, u_i^y]$ are the only constraints for a_i, and the objective function is an affine function of a_i, the maximum is attained at the boundary when $a_i = u_i^y \quad \forall i$. The optimal solution of the above optimization problem can be found by the equivalent problem at the right, which can be solved through linear programming.

Below, we specify the uncertainty for any specific action and the decision-making based on that.

Definition 9 (sample and nonID Uncertainty (Figure 3))**.** *Let an empirical estimation of the joint distribution P be $\hat{P}$ and a confidence set $\mathcal{C}$ be such that $\mathcal{C}$ covers the true P with probability of at least $1-\alpha$. Define the set $\mathcal{S}_P = \{S \text{ an SCM} \mid P_{obs}(S) = P, S \models G\}$ and $\mathcal{I}_P = \{P_S(y|do(x)) \mid S \in \mathcal{S}_P\}$. Then define the intersection $\bigcap_{P\in\mathcal{C}} \mathcal{I}_P$ to be the **nonID uncertainty**, and the set difference $\bigcup_{P\in\mathcal{C}} \mathcal{I}_P \setminus \bigcap_{P\in\mathcal{C}} \mathcal{I}_P$ to be the **sample uncertainty**. The epistemic uncertainty region in the estimation would be the union of these subcomponents. The components are named the inner region and the outer band, respectively.*

Proposition 3. *For a causal decision-making problem with $P(y|do(x))$ as the decision metric for a single action x, let $P^\star$ be the true observational distribution, $\mathcal{C}(D)$ a confidence set satisfying $\Pr(P^\star \in \mathcal{C}(D)) \geq 1-\alpha$, and $\mathcal{I}_P(x)$ the set of values of $P(y|do(x))$ identified for observational distribution P. Then, with probability at least $1-\alpha$, the following hold:*

(i) $\bigcap_{P\in\mathcal{C}(D)} \mathcal{I}_P(x) \subseteq \mathcal{I}_{P^\star}(x)$.

(ii) If the inner region of $P(y|do(x))$ *crosses the decision boundary, i.e.,* $\overline{L_x} < \gamma < \underline{U_x}$*, then the population identified set* $\mathcal{I}_{P^\star}(x)$ *also crosses* γ.

Definition 10 (Inner region: $[\overline{L}_x, \underline{U}_x]$, Outer band $[\underline{L}_x, \overline{L}_x] \cup [\underline{U}_x, \overline{U}_x]$)**.** *Given a confidence set* $\mathcal{C}$ *of an empirical distribution* $\hat{P}$*, Define the set* $\mathcal{S}_P = \{S \text{ an SCM} \mid P_{obs}(S) = P, S \models G\}$*. We access the sample and non-id uncertainty regions by estimating the following quantities:*

$$\begin{aligned}
\overline{U_x} &:= \max_{P\in\mathcal{C}(\hat{P})} \max_{S\in\mathcal{S}_P} P_S(Y|do(X=x)) \\
\underline{U_x} &:= \min_{P\in\mathcal{C}(\hat{P})} \max_{S\in\mathcal{S}_P} P_S(Y|do(X=x)) \\
\overline{L_x} &:= \max_{P\in\mathcal{C}(\hat{P})} \min_{S\in\mathcal{S}_P} P_S(Y|do(X=x)) \\
\underline{L_x} &:= \min_{P\in\mathcal{C}(\hat{P})} \min_{S\in\mathcal{S}_P} P_S(Y|do(X=x)),
\end{aligned} \tag{4}$$

Proposition 4. *There exists SCMs* $S \in \mathcal{S}_P$ *where the uncertainty in the causal decision-making problem due to sample uncertainty in* $P(y|do(x))$ *(i.e., outer band) can be reduced by increasing the number of samples in the data.*

Proof. By the definition, the sample uncertainty are the intervals $[\underline{U}_x, \overline{U}_x]$ and $[\underline{L}_x, \overline{L}_x]$. As the number of samples increases, the radius $r_n = c/\sqrt{n}$ of the epsilon ball that is searched over decreases. Since $\overline{U}_x$ and $\underline{U}_x$ are the supremum and infimum of $u(P) = \max_{S\in\mathcal{S}_P} P_S(y|\text{do}(x))$ over $\mathcal{C}(\hat{P})$, the width of $[\underline{U}_x, \overline{U}_x]$ is at most the variation of u over a ball of radius r_n, which becomes narrower whenever u varies continuously in P. For example, in the graph $X \to Y$ without unobserved confounding, $u(P) = P(y|x)$ and the width is exactly $2r_n$. Similarly for $[\underline{L}_x, \overline{L}_x]$. □

Theorem 5. *Given a confidence set,* $\mathcal{C}$ *and a set of causal effects,* $\mathcal{I}_P$ *(Def. 9),* $\bigcap_{P\in\mathcal{C}} \mathcal{I}_P = [\overline{L_x}, \underline{U_x}]$ *if not empty.*

Proof. Let $m_P = \min \mathcal{I}_P = \min_{S\in\{S \text{ an SCM}|P_{\text{obs}}(S)=P\}} P_S(Y|do(X=x))$ and $\mathcal{M} = \{m_P \mid P \in \mathcal{C}\}$. Then $\min \bigcap_{P\in\mathcal{C}} \mathcal{I}_P = \max \mathcal{M} = \max \min_{S\in\{S \text{ an SCM}|P_{\text{obs}}(S)=P\}} P_S(Y|do(X=x)) = \overline{L_x}$.

By a symmetric argument, we can show that $\max \bigcap_{P\in\mathcal{C}} \mathcal{I}_P = \max \min_{S\in\{S \text{ an SCM}|P_{\text{obs}}(S)=P\}} P_S(Y|do(X=x)) = \underline{U_x}$. □

Corollary 2. *Given a confidence set,* $\mathcal{C}$ *and a set of causal effects,* $\mathcal{I}_P$ *(Def. 9), we have* $\bigcup_{P\in\mathcal{C}} \mathcal{I}_P \setminus \bigcap_{P\in\mathcal{C}} \mathcal{I}_P = [\underline{L_x}, \overline{L_x}] \cup [\underline{U_x}, \overline{U_x}]$

Proof. Since $\bigcap_{P\in\mathcal{C}} \mathcal{I}_P = [\overline{L_x}, \underline{U_x}]$, we only need to show that $\bigcup_{P\in\mathcal{C}} \mathcal{I}_P = [\underline{L_x}, \overline{U_x}]$. Let $m_P = \min \mathcal{I}_P = \min_{S\in\{S \text{ an SCM}|P_{\text{obs}}(S)=P\}} P_S(Y|do(X=x))$ and $\mathcal{M} = \{m_P \mid P \in \mathcal{C}\}$. Then $\min \bigcup_{P\in\mathcal{C}} \mathcal{I}_P = \min \mathcal{M} = \min \min_{S\in\{S \text{ an SCM}|P_{\text{obs}}(S)=P\}} P_S(Y|do(X=x)) = \underline{L_x}$. Similarly, by a symmetric argument, we have $\max \bigcup_{P\in\mathcal{C}} \mathcal{I}_P = \overline{U_x}$ □

Definition 11. *An causal decision with* $P(y|do(x))$ *as decision estimand is ambiguous given confidence set* $\mathcal{C}$ *if there exists two SCMs* S_1, S_2 *where* $p_{S_1}(y_1|do(x)) > P(y)$ *and* $p_{S_2}(y_1|do(x)) < P(y)$ *where* S_1, S_2 *entail* $p_1, p_2 \in \mathcal{C}$.

Theorem 6. *For a causal decision making problem with* $P(y|do(x))$ *bound, let* $(\overline{U_x}, \underline{U_x}, \overline{L_x}, \underline{L_x})$ *be estimated from the data, and let* $\gamma = P(y)$ *be the decision boundary. The decision of* $X = x$ *as the best action (or cannot be the best action) is unambiguous if* $\underline{L_x} > \gamma$ *(or* $\gamma > \overline{U_x}$*). The decision is ambiguous and, with probability at least* $1 - \alpha$*, cannot be improved with more data if* $\overline{L_x} < \gamma < \underline{U_x}$.

Proof. The proof follows that of Theorem 4, replacing $ATE(x)$ with $P(y|\text{do}(x))$, the decision boundary 0 with γ, the quantities $(\overline{U_x}, \underline{U_x}, \overline{L_x}, \underline{L_x})$ with $(\overline{U_x}, \underline{U_x}, \overline{L_x}, \underline{L_x})$, Theorem 3 with Theorem 5, and Proposition 1 with Proposition 3. □

Note that $P(y)$ can be provided by a domain expert. If we estimate it from data, the framework for multiple actions discussed in the previous section can be utilized to make a decision in such cases.

D.3 Approximation with ϵ-net

The following lemma applies even with unobserved confounding between X, Y. The corollary shows that by searching over small enough epsilon net, the output value should find close enough to the true optimal.

Lemma 4. *Consider the causal graph with binary X, Y, Z where Z is an observed confounder with observational distribution $P(X, Y, Z)$. Consider another joint distribution $P'(X, Y, Z)$. If $TVD(P(X, Y|Z = z), P'(X, Y|Z = z)) < \epsilon$ and $TVD(P(Z), P'(Z)) < \epsilon$ then $|f(p) - f(p')| \leq 3\epsilon$ where $f(p) = \max_{\{S|P_{obs}(S)=p\}} p_S(y|do(x))$.*

Proof. Consider the bow-backdoor graph with binary variables X, Y, Z where Z is a confounder. Define the response variables R_y, R_x as follows.

R_x	z_0	z_1
0	y_0	y_0
1	y_0	y_1
2	y_1	y_0
3	y_1	y_1

Table 1: Response function R_x

R_y	z_0, x_0	z_0, x_1	z_1, x_0	z_1, x_1
0	y_0	y_0	y_0	y_0
1	y_0	y_0	y_0	y_1
2	y_0	y_0	y_1	y_0
3	y_0	y_0	y_1	y_1
4	y_0	y_1	y_0	y_0
5	y_0	y_1	y_0	y_1
6	y_0	y_1	y_1	y_0
7	y_0	y_1	y_1	y_1
8	y_1	y_0	y_0	y_0
9	y_1	y_0	y_0	y_1
10	y_1	y_0	y_1	y_0
11	y_1	y_0	y_1	y_1
12	y_1	y_1	y_0	y_0
13	y_1	y_1	y_0	y_1
14	y_1	y_1	y_1	y_0
15	y_1	y_1	y_1	y_1

Table 2: Response function R_y

Let $f(p) = \max p(y_0|do(x_0))$ then we can express $f(p)$ with response variables as following

$$f(p) = \sum_j \left(\left(\sum_{i=0}^{7} q_{i,j} p(z_0) \right) + \left(\sum_{i=\{0,1,4,5,8,9,12,13\}} q_{i,j} p(z_1) \right) \right) \tag{5}$$

where $q_{i,j} = p(R_x = i, R_y = j)$

$$\mathbf{A} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ p(z_0) \\ p(z_0) \\ p(z_0) \\ p(z_0) \\ p(z_1) \\ p(z_1) \\ p(z_1) \\ p(z_1) \\ 1 \\ \vdots \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} q_{0,0} \\ q_{0,1} \\ q_{0,4} \\ q_{0,5} \\ q_{0,2} \\ q_{0,3} \\ q_{0,6} \\ q_{0,7} \\ q_{0,8} \\ q_{0,9} \\ q_{0,12} \\ q_{0,13} \\ q_{1,0} \\ \vdots \end{bmatrix}$$

$$\mathbf{v}_{z_0} = \begin{bmatrix} q_{0,0} \\ q_{0,1} \\ q_{0,2} \\ q_{0,3} \\ q_{0,4} \\ q_{0,5} \\ q_{0,6} \\ q_{0,7} \\ q_{1,0} \\ \vdots \end{bmatrix}, \quad \mathbf{v}_{z_1} = \begin{bmatrix} q_{0,0} \\ q_{0,1} \\ q_{0,4} \\ q_{0,5} \\ q_{0,8} \\ q_{0,9} \\ q_{0,12} \\ q_{0,13} \\ q_{1,0} \\ \vdots \end{bmatrix}$$

The domain of $\mathbf{x}$ is defined by the following constraints

$$\begin{aligned} p(y_0, x_0|z_0) &= q_{0,0} + q_{0,1} + q_{0,2} + q_{0,3} + q_{0,4} + q_{0,5} + q_{0,6} + q_{0,7} \\ &\quad + q_{1,0} + q_{1,1} + q_{1,2} + q_{1,3} + q_{1,4} + q_{1,5} + q_{1,6} + q_{1,7} \\ p(y_0, x_0|z_1) &= q_{0,0} + q_{0,1} + q_{0,4} + q_{0,5} + q_{0,8} + q_{0,9} + q_{0,12} + q_{0,13} \\ &\quad + q_{2,0} + q_{2,1} + q_{2,4} + q_{2,5} + q_{2,8} + q_{2,9} + q_{2,12} + q_{2,13} \\ &\dots \\ p(y_1, x_1|z_1) &= ... \end{aligned}$$

Assume $TVD(P(X,Y|z), P'(X,Y|z)) \leq \epsilon$ and $TVD(P(Z), P'(Z)) \leq \epsilon$ for each z .

Let $k_{xyz} = \{i, j | \sum q_{i,j} = p(x, y|z)\}$. From the construction, for each z, k_{xyz} are non-overlap for each $Z = z$, and $\sum_{xy} k_{xyz} = 1$.

Since $TVD(P(Z), P'(Z)) \leq \epsilon$, $|p(z) - p'(z)| \leq \epsilon$ for all z.

Let $S_z = \{x, y, z | k_{xyz} \cap \mathbf{v}_z \neq \emptyset\}$ Then we have

$$\begin{aligned} \frac{1}{2} \sum_{x,y,z \in S_z} \Big| (p(x, y|z) - p'(x, y|z)) \Big| &\leq \frac{\epsilon}{2} \\ \sum_{x,y,z \in S_z} \Big| \sum_{i,j \in k_{xyz}} (q_{i,j} - q'_{i,j}) \Big| &\leq \epsilon \\ |\sum \mathbf{v}_z - \sum \mathbf{v}'_z| &\leq \epsilon \end{aligned} \tag{6}$$

Hence

$$
\begin{aligned}
|f(p) - f(p')| &= \left|\left(p(z_0)\sum \mathbf{v}_{z_0} + p(z_1)\sum \mathbf{v}_{z_1}\right) - \left(p'(z_0)\sum \mathbf{v}'_{z_0} + p'(z_1)\sum \mathbf{v}'_{z_1}\right)\right| \\
&= \left|\left(p(z_0)\sum \mathbf{v}_{z_0} - p'(z_0)\sum \mathbf{v}'_{z_0}\right) + \left(p(z_1)\sum \mathbf{v}_{z_1} - p'(z_1)\sum \mathbf{v}'_{z_1}\right)\right| \\
&\leq \left|\left(p(z_0)\sum \mathbf{v}_{z_0} - p'(z_0)\sum \mathbf{v}'_{z_0}\right)\right| + \left|\left(p(z_1)\sum \mathbf{v}_{z_1} - p'(z_1)\sum \mathbf{v}'_{z_1}\right)\right| \\
&= \left|\left(p(z_0)\sum \mathbf{v}_{z_0} - p(z_0)\sum \mathbf{v}'_{z_0} + p(z_0)\sum \mathbf{v}'_{z_0} - p'(z_0)\sum \mathbf{v}'_{z_0}\right)\right| \\
&\quad + \left|\left(p(z_1)\sum \mathbf{v}_{z_1} - p(z_1)\sum \mathbf{v}'_{z_1} + p(z_1)\sum \mathbf{v}'_{z_1} - p'(z_1)\sum \mathbf{v}'_{z_1}\right)\right| \qquad (7) \\
&= \left|p(z_0)(\sum \mathbf{v}_{z_0} - \sum \mathbf{v}'_{z_0}) + \sum v'_{z_0}(p(z_0) - p'(z_0))\right| \\
&\quad + \left|p(z_1)(\sum \mathbf{v}_{z_1} - \sum \mathbf{v}'_{z_1}) + \sum v'_{z_1}(p(z_1) - p'(z_1))\right| \\
&\leq p(z_0)\epsilon + \sum v'_{z_0}\epsilon + p(z_1)\epsilon + \sum v'_{z_1}\epsilon \\
&\leq 3\epsilon
\end{aligned}
$$

□

Corollary 3. *For binary variables X, Y with a binary confounder Z and latent confounders. The $\underline{U}_x$ and $\bar{L}_x$ estimated from our algorithm are within 3ϵ distance from the true values.*

Proof. Let $\underline{U}_x$ denote the true $\min\max P(y|do(x))$ and $\hat{\underline{U}}_x$ denote the estimated value of our algorithm. Suppose for the sake of contradiction that $|\hat{\underline{U}}_x - \underline{U}_x| > 3\epsilon$. By Lemma 1, the observed distributions that are compatible with $\underline{U}_x$ and $\hat{\underline{U}}_x$ do not belong to the same ball with ϵ distance. Let B be the ball that contains the observed distribution that is compatible with $\underline{U}_x$, and U_x be the maximum value of $P(y|do(x))$ with a compatible observational distribution inside B. By Lemma 1, $|U - \underline{U}_x| \leq 3\epsilon$. Since $\underline{U}_x, \hat{\underline{U}}_x, U$ are scalers and $\underline{U}_x$ is the smallest $\max P(y|do(x))$ over all observational distributions, we have $U - \underline{U}_x \leq 3\epsilon$, $\hat{\underline{U}}_x - \underline{U}_x > 3\epsilon$. And that implies $\hat{\underline{U}}_x > U$, which contradicts the assumption that $\hat{\underline{U}}_x$ is the smallest $\max P(y|do(x))$ among all epsilon balls. So we can conclude that $|\hat{\underline{U}}_x - \underline{U}_x| \leq 3\epsilon$ □

Lemma 5. *For an observational distribution from IV graph $P(X, Y, Z)$ where X, Y, Z are binary variables, there exist a SCM such that $P(y_1|do(x_1))$ attains the upper bound and $P(y_1|do(x_0))$ attains the lower bound.*

Proof. Let R_x, R_y be the response variables for X, Y, and $q_{ij} := P(R_x = i, R_y = j)$. Then we have the following constraints from the observed data.

$$
\begin{aligned}
P(y_0, x_0|z_0) &= q_{00} + q_{10} + q_{01} + q_{11} \\
P(y_0, x_1|z_0) &= q_{20} + q_{22} + q_{30} + q_{32} \\
P(y_1, x_0|z_0) &= q_{02} + q_{03} + q_{12} + q_{13} \\
P(y_1, x_1|z_0) &= q_{21} + q_{23} + q_{31} + q_{33} \\
P(y_0, x_0|z_1) &= q_{00} + q_{20} + q_{01} + q_{21} \\
P(y_0, x_1|z_1) &= q_{10} + q_{12} + q_{30} + q_{32} \\
P(y_1, x_0|z_1) &= q_{02} + q_{03} + q_{22} + q_{23} \\
P(y_1, x_1|z_1) &= q_{11} + q_{13} + q_{31} + q_{33}
\end{aligned}
$$

$$P(y_1|do(x_1)) = \sum_{i=0}^{3} q_{i1} + q_{i3}$$
$$P(y_1|do(x_0)) = \sum_{i=0}^{3} q_{i2} + q_{i3}$$

The goal is to show an SCM that obtains the minimum of $P(y_1|do(x_0))$ and the maximum of $P(y_1|do(x_1))$. For the SCM minimize $P(y_1|do(x_0))$, the corresponding q_{ij} are minimized. Without loss of generality, assume $P(y_1, x_0|z_1) > P(y_1, x_0|z_0)$. The minimum of $P(y_1|do(x_0))$ can be obtained by let $q_{12}, q_{13}, q_{32}, q_{33}$ equal to zero (because the constraint $P(y_1, x_0|z_1) = q_{02} + q_{03} + q_{22} + q_{23}$, the terms $q_{02}, q_{03}, q_{22}, q_{23}$ do not affect the value of $P(y_1|do(x_0))$).

Similarly, for the maximum of $P(y_1|do(x_1))$, without loss of generality, assume $P(y_0, x_1|z_1) > P(y_0, x_1|z_0)$. To maximize $P(y_1|do(x_1)) = \sum_{i=0}^{3} q_{i1} + q_{i3}$, let $q_{01} + q_{21} = P(y_0, x_0|z_1)$, and $q_{03} + q_{23} = P(y_1, x_0|z_1)$ (due to the constraint $P(y_1, x_1|z_1) = q_{11} + q_{13} + q_{31} + q_{33}$,the terms $q_{11}, q_{13}, q_{31}, q_{33}$ do not affect the value of $P(y_1|do(x_1))$). This is equivalent to making $q_{00}, q_{20}, q_{02}, q_{22}$ equal to zero.

Combining the two derivations above, any SCM with response variables such that $q_{00}, q_{20}, q_{02}, q_{22}, q_{12}, q_{13}, q_{32}, q_{33}$ equals to zero would attains the maximum of $P(y_1|do(x_1))$ and minimum of $P(y_1|do(x_0))$. □

D.4 Properties of the Non-ID region/inner bound

Preliminaries

$r_n := c/\sqrt{n}, \quad c := \sqrt{\frac{1}{2}\ln(2m/\alpha)}$	confidence radius at sample size n
$I_P = [\ell(P), u(P)]$	ATE bound set for distribution P
Λ	Lipschitz constant of ℓ, u in P
$w^* := \lvert I_{P^*}\rvert > 0$	true non-ID width (fixed, population quantity)
$\mathcal{C}_n := B(\hat{P}_n, r_n)$	confidence set at sample size n
$J_n := \bigcap_{P\in\mathcal{C}_n} I_P$	reported non-ID region
$K_n := \big[\ell(P^*) + 2\Lambda r_n,\ u(P^*) - 2\Lambda r_n\big]$	guaranteed core at sample size n
$\kappa := \dfrac{4\Lambda r_n}{w^*} > 1$	resolution-excess factor

On the coverage event $\{P^* \in \mathcal{C}_n\}$, one has $K_n \subseteq J_n \subseteq I_{P^*}$, hence taking widths,

$$w^* - 4\Lambda r_n \ \le\ \omega_n \ \le\ w^*, \tag{6}$$

where $\omega_n := |J_n|$ (signed, so $J_n = \emptyset \iff \omega_n < 0$).

Theorem 7 (Resolving batch size)**.** *Suppose at sample size n the reported non-ID region J_n is empty, while $w^* > 0$. Then, on the coverage event $\{P^* \in \mathcal{C}_n\} \cap \{P^* \in \mathcal{C}_{n'}\}$, of probability at least $1 - 2\alpha$, collecting*

$$k \ >\ k^* \ :=\ \frac{8\Lambda^2 \ln(2m/\alpha)}{(w^*)^2} - n \qquad \Longleftrightarrow \qquad k \ >\ (\kappa^2 - 1)\, n$$

additional samples guarantees

$$J_{n+k} \ \supseteq\ K_{n+k} \ \neq\ \emptyset.$$

Equivalently, in terms of total sample size $n' = n + k$, this is $n' > N^* := \left\lceil \dfrac{8\Lambda^2 \ln(2m/\alpha)}{(w^*)^2} \right\rceil$.

Proof. It suffices to show $K_{n'} \neq \emptyset$, since $K_{n'} \subseteq J_{n'}$ on the coverage event.

Step 1 (nonemptiness of $K_{n'}$). $K_{n'}$ is an interval whose width is

$$|K_{n'}| = \big(u(P^*) - 2\Lambda r_{n'}\big) - \big(\ell(P^*) + 2\Lambda r_{n'}\big) = w^* - 4\Lambda r_{n'}.$$

An interval is nonempty iff its width is positive, so

$$K_{n'} \neq \emptyset \iff w^* - 4\Lambda r_{n'} > 0 \iff w^* > 4\Lambda r_{n'}.$$

Step 2 (solving for k). Substituting $r_{n'} = c/\sqrt{n'}$,

$$w^* > \frac{4\Lambda c}{\sqrt{n'}} \iff \sqrt{n'} > \frac{4\Lambda c}{w^*} \iff n' > \frac{16\Lambda^2 c^2}{(w^*)^2}.$$

Both sides were positive throughout, so squaring preserves the inequality. Substituting $c^2 = \frac{1}{2}\ln(2m/\alpha)$ gives

$$n' > \frac{8\Lambda^2 \ln(2m/\alpha)}{(w^*)^2} =: N^*.$$

Writing $n' = n + k$,

$$n + k > N^* \iff k > \underbrace{\frac{8\Lambda^2 \ln(2m/\alpha)}{(w^*)^2} - n}_{=k^*}.$$

Step 3 (equivalence with κ). Since $c = r_n\sqrt{n}$, so $c^2 = r_n^2 n$,

$$N^* = \frac{16\Lambda^2 c^2}{(w^*)^2} = \frac{16\Lambda^2 r_n^2 n}{(w^*)^2} = \left(\frac{4\Lambda r_n}{w^*}\right)^2 n = \kappa^2 n.$$

Hence $k^* = N^* - n = \kappa^2 n - n = (\kappa^2 - 1)n$, i.e. $k > k^* \iff k > (\kappa^2 - 1)n$.

Combining Steps 1–3: $k > (\kappa^2 - 1)n \implies n' > N^* \implies w^* > 4\Lambda r_{n'} \implies K_{n'} \neq \emptyset \implies J_{n+k} \supseteq K_{n+k} \neq \emptyset$. □

E Algorithm Details and Complexity Analysis

E.1 From Ambiguous to Unambiguous Decision for Single Action

We can reach the final decision through a sequence of three moves. Let $\gamma = P(y)$. *1. Return:* if $\gamma < \underline{L_x}$, we decide the action is beneficial in population level. *2. Observe*: if $\overline{L_x} < \gamma < \underline{U_x}$, then we observe additional variables for example instrument variables. Observing additional variables gives us a small causal effect set $\mathcal{I}'_P$, for each $P \in \mathcal{C}$. Thus, for fixed $\mathcal{C}$, we get a narrower inner region, $[\overline{L}_x, \underline{U}_x] = \bigcap_{P\in\mathcal{C}} \mathcal{I}'_P$ and a narrower/wider outer band, $[\underline{L}_x, \overline{L}_x] \cup [\underline{U}_x, \overline{U}_x] = \bigcup_{P\in\mathcal{C}} \mathcal{I}'_P \setminus \bigcap_{P\in\mathcal{C}} \mathcal{I}'_P$. Also, the average non-id bound width will reduce since $1/|\mathcal{C}| \sum_{P\in\mathcal{C}} |\mathcal{I}'_P| \leq 1/|\mathcal{C}| \sum_{P\in\mathcal{C}} |\mathcal{I}_P|$.

3. Collect: In all other scenarios, we are unsure about the source of uncertainty (low sample size or non-identifiability) and we collect more samples, with the assumption that obtaining samples with additional variables is more challenging compared to collecting samples of the same set of variables. More samples provides us a smaller confidence set $\mathcal{C}'(\hat{P})$. Thus, according to Theorem 5 and 3, we get a wider inner region, $[\overline{L}_x, \underline{U}_x] = \bigcap_{P\in\mathcal{C}'} \mathcal{I}_P$ and according to Corollary 2 and 1, a narrower outer band, $[\underline{L}_x, \overline{L}_x] \cup [\underline{U}_x, \overline{U}_x] = \bigcup_{P\in\mathcal{C}'} \mathcal{I}_P \setminus \bigcap_{P\in\mathcal{C}'} \mathcal{I}_P$. The *collect* and *observe* moves can continue until we reach the *return* move and decide the best action.

E.2 Effect of Sample Size on an Individual Action Interval

Suppose we are given N samples from the true distribution $P(.)$, resulting in an empirical distribution $\hat{P}$. From the samples, we can construct confidence regions $\mathcal{C}_{P(.)}$. Any distribution $P_\epsilon \in \mathcal{C}_{P(.)}$ with $d(P_\epsilon, \hat{P}) < \epsilon$ can be the true distribution. We maximize the causal effect while keeping the empirical distribution in this region. Suppose all the empirical distributions $P_\epsilon \in \mathcal{C}$ provide us a set of max causal effects $M_{P(y|\text{do}(x))}$.

Now, suppose we have a larger confidence region $\mathcal{C}^L \supseteq \mathcal{C}$, e.g., obtained from the same empirical distribution with fewer samples or a higher confidence level. $\mathcal{C}^L$ allows more distributions as potential candidates to be the true distribution. As a result, we will have a new set of max causal effects $M'_{P(y|\text{do}(x))}$ for the distributions $P_\epsilon \in \mathcal{C}^L \setminus \mathcal{C}$. Now, $\overline{U} = \max(M_{P(y|\text{do}(x))}, M'_{P(y|\text{do}(x))}) \geq \max(M_{P(y|\text{do}(x))})$ and $\underline{U} = \min(M_{P(y|\text{do}(x))}, M'_{P(y|\text{do}(x))}) \leq \min(M_{P(y|\text{do}(x))})$. Thus, for a larger confidence region, the bound $[\underline{U}, \overline{U}]$ extends. Similarly, the bound $[\underline{L}, \overline{L}]$ also extends. As the above two bounds extend, the bound $[\overline{L}, \underline{U}]$ shrinks.

Now, suppose we have a smaller confidence region $\mathcal{C}^S \subseteq \mathcal{C}$, allowing fewer distributions as potential candidates to be the true distribution. As a result, we will have a new set of max causal effects $M'_{P(y|\text{do}(x))}$ for the distributions $P_\epsilon \in \mathcal{C}^S$. Now, $\overline{U} = \max(M'_{P(y|\text{do}(x))}) \leq \max(M_{P(y|\text{do}(x))})$ and $\underline{U} = \min(M'_{P(y|\text{do}(x))}) \geq \min(M_{P(y|\text{do}(x))})$. Thus, for a smaller confidence region, the bound $[\underline{U}, \overline{U}]$ shrinks. Similarly, the bound $[\underline{L}, \overline{L}]$ also shrinks. As the above two bounds shrink, the bound $[\overline{L}, \underline{U}]$ extends.

The above comparisons are exact when the confidence regions are nested, e.g., for the same empirical distribution with different confidence levels (Appendix F.2). When additional samples are collected, the empirical distribution itself changes, so the new confidence region is not guaranteed to be nested in the previous one. However, since the confidence radius shrinks at rate $O(1/\sqrt{n})$, the confidence region typically shrinks and concentrates around $P^\star$, and the above behavior typically holds, while the inner region remains contained in $\mathcal{I}_{P^\star}$ with probability at least $1 - \alpha$ (Proposition 3).

The consequence of sample size can also be visualized as imagining $\max p(y|\text{do}(x))$ as an estimand. With finite samples, there is a confidence interval below and above it, which is $\min\max p(y|\text{do}(x))$ and $\max\max p(y|\text{do}(x))$. With fewer samples, this confidence interval typically gets wider. So the $\min\max p(y|\text{do}(x))$ will typically be smaller. Similarly, the $\max\min p(y|\text{do}(x))$ will typically be larger.

Hyper parameters: (α): In Alg 2, determines how fast the model's implicit distribution should return to the ϵ-ball around the empirical distribution (i.e., controls the soft constraints). (ϵ_s): In Alg 1, determines the size of the search space around the empirical distribution. In general, smaller ϵ_s values provide tighter estimates but require a larger number of candidate balls within the search space.

Algorithm 2 RelaxedDCM Training Model

Input: Data distribution $P_\epsilon(\mathbf{V})$, Variables $\mathbf{V}$, graph $\mathcal{G}$, action x, smaller interval ϵ_s
for $j \in [0, 1]$ ▷ 0: minimization, 1: maximization **do**
  Initialize dual parameter $\{\alpha_i\}_{V_i \in \mathbf{V}}$ and ate weight λ.
  Initialize DCM parameters $\Theta \leftarrow \{\theta_i\}_{V_i \in \mathbf{V}}$
  for $epoch = 1 \rightarrow max_epoch$ **do**
    $d_i = dist\left(P_\epsilon(v_i \mid v_{\pi^{(i-1)}}), P_{\Theta_i}(v_i \mid v_{\pi^{(i-1)}})\right), \forall_i : V_i \in \mathbf{V}$
    $L \leftarrow \frac{1}{n} \sum_{k=1}^{n} \sum_{V_i \in \mathbf{V}} \alpha_i \left[d_i - \epsilon_s\right]$
    $\mathcal{L} \leftarrow L + (-1)^j * \lambda * ATE^{(\Theta)}(x)$ ▷ ATE with respect to the best among remaining actions.
    $\alpha_i \leftarrow \max\left(0,\ \alpha_i + lr \cdot (d_i - \epsilon_s)\right), \forall_i : V_i \in \mathbf{V}$
    $\Theta \leftarrow \Theta - \eta \nabla \mathcal{L}$
  $bound[j] = ATE^{(\Theta)}(x)$
Return $\{U_x, L_x\} = \{bound[1], bound[0]\}$

Algorithm 3 Construct ϵ-net

Input: data $\{\mathbf{v}_k\}_{k=1}^n$, Variable set $\mathbf{V}$, graph $\mathcal{G}$, small interval ϵ_s
Factorize $\hat{P}(\mathbf{v})$ in $\Pi_{v \in \mathbf{v}} \hat{P}(v_i | v^{\pi_i - 1})$
for each $V_i \in \mathbf{V}$ **do**
  $\epsilon = \sqrt{\frac{\ln(2/\alpha)}{2n}}$ where $n = count(D[v_\pi^{(i-1)}])$
  $C_i = \lceil \epsilon / \epsilon_s \rceil$ equidistance points in $\hat{P}_n(v_i = 1 | v_\pi^{(i-1)}) \pm \epsilon$
Sample m joints s.t.$\{(p_1^{(j)}, .., p_{|\mathbf{V}|}^{(j)})\}_{j=1}^m \subseteq C_1 \times C_2 \times ... \times C_{|\mathbf{V}|}$
Return m candidate joint distributions $\{P^{(j)}\}_{j=1}^m$.

E.3 Algorithm 3: Construct ϵ-net.

There are various ways to construct a confidence set, extending on the classical confidence intervals. We use a simple set defined as the cartesian product of intervals obtained via concentration inequalities in each coordinate of the joint distribution, given as follows: $\mathcal{S} := \prod_{v \in \mathcal{V}} [l_v, u_v]$, where v is a configuration of the joint variable set and l_v, u_v are the lower and upper bounds that contain true $p(v)$ with probability α/m, where m is the total number of configurations in the observed distributions. Thus, $p \in S$ with probability α through a simple union bound argument.

Calculate Confidence Intervals We calculate the confidence interval for any conditional distribution $P(v_i | v^{\pi_i - 1})$ as below:

First, we need to obtain the empirical estimators $\hat{P}(v_i = 1 | v^{\pi_i - 1})$ for the binary variables from the training data. $\hat{P}_n(v_i = 1 \mid v^{\pi_i - 1}) = \frac{1}{m} \sum_{k=1}^{m} \mathbb{1}(v_i^k(v^{\pi_i - 1}) = 1)$

Now, let $n = count(D[v^{\pi_i - 1}])$. Then according to Hoeffding's inequality:

$$
\begin{aligned}
&Pr(|\hat{P} - P| \geq \epsilon) \leq 2e^{-2n\epsilon^2} \\
&\alpha = 2e^{-2n\epsilon^2} \\
&\alpha/2 = e^{-2n\epsilon^2} \\
&ln(2/\alpha) = 2n\epsilon^2 \\
&\frac{ln(2/\alpha)}{2n} = \epsilon^2 \\
&\epsilon = \sqrt{\frac{ln(2/\alpha)}{2n}}
\end{aligned}
\tag{8}
$$

Thus, the confidence interval for our empirical estimators are:

$$
\mathcal{C}_{P(v_i = 1 | v^{\pi_i - 1})} := \left(\hat{P}_n(v_i = 1 | v^{\pi_i - 1}) - \sqrt{\tfrac{\ln(2/\alpha)}{2n}},\ \hat{P}_n(v_i = 1 | v^{\pi_i - 1}) + \sqrt{\tfrac{\ln(2/\alpha)}{2n}} \right)
$$

$P(v_i = 1|v^{\pi_i - 1}) \in \mathcal{C}_{P(v_i=1|v^{\pi_i-1})}$ with probability greater than $1 - \alpha$

Note that the number of confidence intervals increases with the support size and the number of variables. If we consider $\lceil \epsilon/\epsilon_s \rceil$ centroids in each interval, the number of all possible joint distributions will be extremely high. To deal with such a scenario, we uniformly pick a centroid from each conditional distribution interval and form a valid joint distribution P. For a binary two variable case X, Y, we can split the CI of $P(X = 1)$ into $n + 1$ smaller intervals by placing n centroids. We pick uniformly a value p from these n centroids having $P(X = 0) = 1 - p$. Doing the same for $P(Y = 1|X = 0)$ and $P(Y = 1|X = 1)$ will allow us to construct $P(X, Y)$. However, since we are sampling the joint distribution only a finite number of times, it is possible that we miss the true distribution. We have to accept a small error in the constructed bounds $[\underline{L}_x, \overline{L}_x, \underline{U}_x, \overline{U}_x]$ as well. We leave the development of better search strategies for future work. While exploring the confidence interval, if a candidate distribution becomes inconsistent with the assumed instrument graph, we reject it based on conditions proposed in (Pearl, 2013). Finally, our algorithm can deal with positivity violation of any distribution by considering the whole region $[0, 1]$ as its confidence interval.

Complexity: Let T denote the number of epochs, B denote the batch size, $|D|$ denote the dataset size, $|V|$ denote the number of variables, $\mathcal{X}$ denote the cardinality of each variable. We construct confidence intervals from dataset D for each of the $|V|$ variables, which costs $O(|V||D|)$. Sampling M joint distributions from these intervals adds $O(M)$ cost. Thus,

$$A_3 = O(|V||D| + M).$$

E.4 Algorithm 2: RelaxedDCM Training Model.

For the Deep Causal Models defined in Definition 3, we have the following theorem.

Theorem 8. *(Kocaoglu et al., 2018; Xia et al., 2021; Rahman & Kocaoglu, 2024) Consider any SCM $S = (\mathbf{V}, \mathcal{N}, \mathcal{U}, \mathcal{F}, P(.))$. A DCM $\mathbb{G} = \{f_1, ..., f_n\}$ for G entails the same identifiable interventional distributions as the SCM S if it entails the same observational distribution.*

Complexity: Let T denote the number of epochs, B denote the batch size, $|V|$ denote the number of variables, $\mathcal{X}$ denote the cardinality of each variable, c denote the forward-backward pass per sample cost, and m denote the number of Monte Carlo samples. We use one neural network for each variable. The joint distribution of $|V|$ variables has $\mathcal{X}^{|V|}$ configurations. In each epoch, we process B samples of each configuration, with m Monte Carlo samples, evaluate $|V|$ neural networks, and repeat the procedure for both minimization and maximization.

The total cost is

$$A_2 = T \cdot 2 \cdot \mathcal{X}^{|V|} \cdot B \cdot m \cdot |V| \cdot c = 2T\mathcal{X}^{|V|}B|V|mc = O\Big(T\mathcal{X}^{|V|}B|V|mc\Big).$$

E.5 Algorithm 1: Explore ϵ-ball.

Complexity: Let T denote the number of epochs, B denote the batch size, $|D|$ denote the dataset size, $|V|$ denote the number of variables, $\mathcal{X}$ denote the cardinality of each variable, c denote the forward-backward pass per sample cost, and m denote the number of Monte Carlo samples. We use one neural network for each variable.

After constructing the ϵ-net using Algorithm 3, we execute RelaxedDCM using Algorithm 2 for each of the M candidate joint distributions. Thus,

$$A_1 = O(A_3 + M \cdot A_2) = O\Big(|V||D| + M + M \cdot T\mathcal{X}^{|V|}B|V|mc\Big).$$

The overall complexity is dominated by Algorithm 2 and scales as $\mathcal{X}^{|V|}$ due to operations over the joint space. Therefore, the complexity is exponential in the number of variables $|V|$ and polynomial in the domain size $\mathcal{X}$ with degree $|V|$.

E.6 Graph-Induced Constraints on the Confidence Set

The confidence set $\mathcal{C}(D)$ is constructed from the data and is not, in general, contained in $\mathcal{M}(\mathcal{G})$, the set of observational distributions entailed by some SCM compatible with $\mathcal{G}$. Even when $P^\star \in \mathcal{M}(\mathcal{G})$,

distributions in $\mathcal{C}(D)$ that are far from the empirical distribution may violate the constraints implied by $\mathcal{G}$. For such a distribution $P \notin \mathcal{M}(\mathcal{G})$, no compatible SCM entails P, so $\mathcal{I}_P(x) = \emptyset$. Consequently, whenever $\mathcal{C}(D) \not\subseteq \mathcal{M}(\mathcal{G})$, the inner region $\bigcap_{P \in \mathcal{C}(D)} \mathcal{I}_P(x)$ is empty in exact arithmetic, while in practice the neural optimization returns non-empty approximate intervals for such P, which can distort both the inner and outer regions. Restricting the confidence set to $\mathcal{C}(D) \cap \mathcal{M}(\mathcal{G})$ preserves coverage when $\mathcal{G}$ is correctly specified and removes such problems.

Whether this restriction is needed depends on the graph. For graphs such as the bow graph, $\mathcal{M}(\mathcal{G})$ imposes no constraints on the observational distribution. For causally sufficient graphs, the constraints are enforced by construction, since we build confidence intervals on local conditionals. With latent variables, factorizing over c-components captures part of this structure but does not enforce inequality constraints such as the instrumental inequalities. For the IV graph, we therefore check the instrumental inequalities for the distributions explored in the ϵ-net and discard those that violate them.

For the binary IV graph, these inequalities (Pearl, 1995b) are linear in $p(x, y \mid z)$ and characterize the observational class exactly (Balke & Pearl, 1997; Bonet, 2001), so the restricted confidence set is a polytope over which an ϵ-net can be constructed directly. Constructing confidence sets within the observational class of a general graph with latent variables remains an open problem.

E.7 Computational Complexity:

When the training dataset contains small number of samples for variables with large number of states, the search space of epsilon-net is large and therefore more iterations are needed to find the min-max and max-min. In that case, the size of epsilon-net can be adjusted with a trade-off between accuracy and efficiency. As the number of samples increases, the search space of epsilon-net will become smaller and therefore more efficient when searching for min-max and max-min. Finally, when the empirical distribution converges to the true distribution for a large dataset, and the confidence intervals shrink accordingly, our method recovers the true bounds for non-identifiable queries, similar to existing approaches. Our main contribution lies in handling the uncertainty region where we can only estimate the empirical distributions with finite samples. We can utilize and separate the sources of uncertainty to make optimal decisions or develop data collection strategies to improve the results.

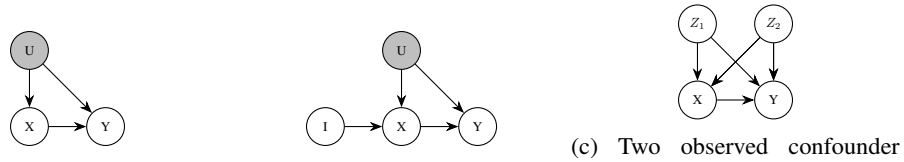
We pick each distribution P_{ϵ_s} and train two models for at least 2000 epochs to maximize and minimize the ATE respectively while keeping the model learned distribution within ϵ_s distance of P_{ϵ_s}. We learn k (~ 200) such distributions and achieve k many $[\min, \max,]$ bounds. We can then construct the $[\underline{L}, \overline{L}, \underline{U}, \overline{U}]$ following Equation 1.

All experiments were conducted on workstations equipped with two NVIDIA GeForce RTX 4090 GPUs, each with 24 GB of memory. We executed the algorithm using 15-20 parallel threads, where each thread explores one candidate joint distribution from the ϵ-ball at a time. The runtime depends on the problem instance, computational resources, and hyperparameter choices. For a dataset with 18 binary variables, we train for 500-1000 epochs with a batch size of 256, processing each meaningful joint configuration. A single run takes approximately 20 minutes to 1 hour, and evaluating 100-200 candidate distributions takes approximately 6-7 hours under parallel execution.

F Experiments

F.1 Causal Graphs

The following causal graphs are used in our experiments.



(a) Bow graph. (b) Instrument variable (IV) graph. graph. (c) Two observed confounder

Figure 7: Causal graphs used in experiments

F.2 Experiment Detail for Section 5.1

Setup: In the first experiments, we consider the causal graph shown in Figure 7b where we have 2 instrument variables (I_1, I_2), one unobserved confounder (U), a treatment (X) and an outcome variable (Y). All of them are considered binary. The ground truth ATE of this causal model is 0.236, i.e, X_1 is the best action. To illustrate the gradual progression of our algorithm we consider four setups: setup 1) observed variables: $\{X, Y\}$ are sampled from $P(x, y)$ as dataset $D[X, Y]$ of size=1000; setup 2) observed: $\{I_1, X, Y\}$, are sampled from $P(i, x, y)$ as dataset $D[I_1, X, Y]$ of size 1000; setup 3) observed $\{I_1, I_2, X, Y\}$ are sampled from $P(\mathbf{V})$ as $D[\mathbf{V}]$ of size 1000; setup 4: same as setup 3 but with 3000 samples. The goal is to choose one of the two actions $do(X = 0)$ and $do(X = 1)$ as the best action.

Baselines: Given the dataset, we estimate the corresponding joint distribution and use closed-form expression such as Tian and Pearl bound (Tian & Pearl, 2000) for bow graph with no instruments and IV bound when we have one or more instrument variables. We also use NCM (Xia et al., 2021) and autobound (Duarte et al., 2024) to obtain bounds for the input distribution.

Confidence levels: We evaluate the behavior of our algorithm for different confidence levels. We plug in the closed-form bound estimator and execute Algorithms 1 and 3 with $10k$ samples but different confidence levels. We observe that for $[70\%, 90\%]$ confidence levels, the outer region is always greater than zero, while for 95% it is not. Thus, for a fixed sample size, we have to reduce our confidence level to find the best action. For $10k$ samples, the estimated interval $[\text{minmin}, \text{maxmax}]$ is $[-0.02, 0.48]$ at 95% confidence level, $[0.02, 0.48]$ at 90% confidence level, and $[0.02, 0.45]$ at 70% confidence level.

Sample size: One metric to evaluate the cost of our decision making might be the **sample size**. We extend Section 5.1 by performing additional experiments for sample sizes $8k$ and $16k$ and show how the bound changes. For $3k$ samples, the outer bound is $[-0.083, 0.52]$, and the resulting strategy is to collect more samples. For $8k$ samples, the outer bound is $[0.0042, 0.43]$, and the resulting strategy is to return $\text{do}(X = 1)$. For $16k$ samples, the outer bound is $[0.04, 0.39]$, and the resulting strategy is to return $\text{do}(X = 1)$. Although our algorithm requires a high number of samples, it follows a principled way to propagate uncertainty using causal inference machinery and offers a confident decision unlike existing baselines.

F.3 Experiment Detail for Section 5.1

Setup: In this experiment, we consider the causal graph shown in Figure 7c where we have 2 observed confounders (Z_1, Z_2), a treatment (X) and an outcome variable (Y). All of them are considered binary. In this setup, we do not have any unobserved confounder and thus no uncertainty from non-identifiability of ATE or $P(Y|\text{do}(X))$. The ground truth ATE=0.06. Here we consider two setups with the same graph: setup 1) observed variables: $\{Z_1, Z_2, X, Y\}$ are sampled from $P(\mathbf{V})$ as dataset $D[Z_1, Z_2, X, Y]$ of size=1000; setup 2) same as setup 1 but with 3000 samples. The goal is to choose one of the two actions $do(X = 0)$ and $do(X = 1)$ as the best action.

Baselines: Since there exists no unobserved confounders, we can compare our algorithm output with existing baselines that assume unconfoundedness. We use the dowhy package (Sharma & Kiciman, 2020; Blöbaum et al., 2024) to execute 6 baselines such as: propensity score matching, doubly robust estimator etc., on our data. These methods provides a point-wise estimation. We also execute the NCM approach which provides a bound for the ATE estimation.

F.4 Scaling Experiment

The overall complexity of the naive implementation of our algorithm is exponential in the number of variables $|V|$ and polynomial in the domain size $\mathcal{X}$ with degree $|V|$. Please see Section E for complexity details. The computation ($\mathcal{X}^{|\mathbf{V}|}$) is required since the method executes Algorithm 2 for each candidate distribution by exploring the ϵ-ball. To scale our algorithm for large graphs, we utilize two practical solutions. 1. **Reduce the number of confidence intervals:** The joint distribution factorizes according to the causal graph $P(V) = \prod_{V_i \in V} P(V_i \mid pa(V_i))$. We match the factorized conditional distributions instead of the full joint. Similar factorization with c-components can be leveraged when latents are present. 2. **Ignore low-probability joint combinations:** For large graphs

with large-cardinality variables, many combinations are assigned a very low probability. A practical solution is to identify these very low-probability combinations and redistribute their probability mass to the rest of the states and fit that distribution. This reduces the slicing of the alpha value to a large number of combinations that do not meaningfully change the causal effect since the resultant perturbation of the observations is small. Accordingly, for an 18-node graph, we end up with just ~ 400 meaningful combinations.

F.5 Experiment Details for Parent Labor Dataset

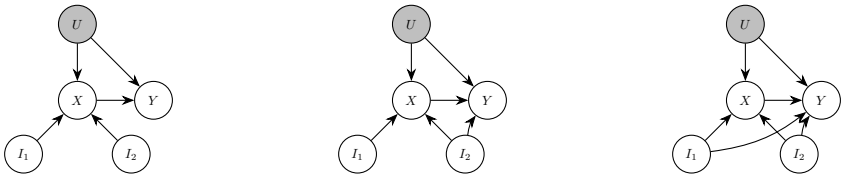
Angrist & Evans (1996) focus on the causal link running from fertility to the work effort of both men and women. According to Angrist & Evans (1996), we consider an instrumental variable I: the sex of the first two children in families with two or more children. This variable can be justified as a valid instrument since it represents the widely observed phenomenon of parental preferences for a mixed sibling-sex composition. Parents of same-sex children are significantly likely interested to have an additional child. Also, sex is randomly assigned, I does not have any causal parent. Thus, it is a plausible instrument for further child-bearing among women with at least two children. We obtain the dataset from Angrist & Evans (2009).

Angrist and Evans Angrist & Evans (1996) estimate the causal effect of fertility on parents' labor supply using an instrumental variable strategy based on the sex composition of the first two children. Their treatment variable is whether a family has more than two children, and their main instrument is an indicator for whether the first two children are of the same sex. The identifying intuition is that parents with two same-sex children are more likely to have an additional child due to preferences for a mixed sibling-sex composition, while the sex composition of the first two births is plausibly as good as randomly assigned. Using two-stage least squares (2SLS), they first predict the probability of having more than two children using the same-sex instrument and then estimate the effect of the predicted fertility variation on labor-supply outcomes. In their main 1980 PUMS specification for women aged 21 to 35 with at least two children, the 2SLS estimates using the same-sex instrument show that having more than two children reduces the probability that a woman worked for pay by 12.1 percentage points, annual weeks worked by 5.68 weeks, weekly hours worked by 4.61 hours, and labor income by approximately $\$1{,}960.5$. They also find little evidence that additional children reduce husbands' labor supply. Overall, their results indicate that additional childbearing causally reduces mothers' labor supply and earnings, while having little effect on fathers' labor-market outcomes. These estimates are reported in Table 5 of Angrist and Evans Angrist & Evans (1996).

F.6 How Model Misspecification in the Causal Graph Affects the Final Results

We analyze the final result for a scenario where incorrect edges are present.

True graph. The true graph contains $I_1 \to X, I_2 \to X, X \to Y$, and an unobserved confounder U between X and Y.



(a) True graph. (b) Wrong graph 1. (c) Wrong graph 2.

Figure 8: Causal graphs used in the misspecification experiment. The true graph is a bow graph with two IVs. Wrong graph 1 adds the extra edge $I_2 \to Y$. Wrong graph 2 adds the extra edges $I_1 \to Y$ and $I_2 \to Y$.

Wrong graph 1 has the extra edge $I_2 \to Y$. Wrong graph 2 has the extra edges $I_2 \to Y$ and $I_1 \to Y$.

Since any estimator can be incorporated in our framework, we pick a closed-form formula to calculate bounds for these graphs. We explore distributions in the ϵ-ball and combine the bounds to obtain the following four quantities.

Table 3: Effect of causal graph misspecification on the final causal effect bounds.

Setup	**Instruments**	**minmin**	**maxmin**	**minmax**	**maxmax**
True graph	I_1, I_2	0.01	0.11	0.333	0.41
Wrong graph 1	I_1	-0.23	-0.16	0.52	0.59
Wrong graph 2	None	-0.35	-0.32	0.64	0.67

We observe that the additional incorrect edge loosens the causal effect bound compared to the correct graph, although it still encloses the correct bound.