Differentiable Multi-Target Causal Bayesian Experimental Design

Panagiotis Tigas [* 1] **Yashas Annadani** [* 2 3] **Desi R. Ivanova** [4] **Andrew Jesson** [1] **Yarin Gal** [1] **Adam Foster** [† 5]
**Stefan Bauer** [† 3 6 7]

Abstract

We introduce a gradient-based approach for the problem of Bayesian optimal experimental design to learn causal models in a batch setting — a critical component for causal discovery from finite data where interventions can be costly or risky. Existing methods rely on greedy approximations to construct a batch of experiments while using black-box methods to optimize over a *single target-state* pair to intervene with. In this work, we completely dispose of the black-box optimization techniques and greedy heuristics and instead propose a conceptually simple end-to-end gradient-based optimization procedure to acquire a set of optimal intervention target-state pairs. Such a procedure enables parameterization of the design space to efficiently optimize over a batch of *multi-target-state* interventions, a setting which has hitherto not been explored due to its complexity. We demonstrate that our proposed method outperforms baselines and existing acquisition strategies in both single-target and multi-target settings across a number of synthetic datasets.

1. Introduction

Imagine a scientist entering a wet lab to conduct experiments in order to discover the underlying causal relations within the system of interest. The scientist first comes up with some hypotheses, based on prior knowledge and past observations. Then, based on the formed hypotheses, an experimentation protocol to disambiguate between the competing hypotheses is devised. Additionally, because of the

---

*Equal contribution ,†Equal supervision [1]OATML, University of Oxford [2]KTH Stockholm, Sweden [3]Helmholtz AI [4]Department of Statistics, University of Oxford [5]Microsoft Research [6]TU Munich [7]CIFAR Azrieli Global Scholar. Correspondence to: Yashas Annadani <yashas.annadani@gmail.com>, Panagiotis Tigas <ptigas@robots.ox.ac.uk>. Code available at: https://github.com/yannadani/DiffCBED.
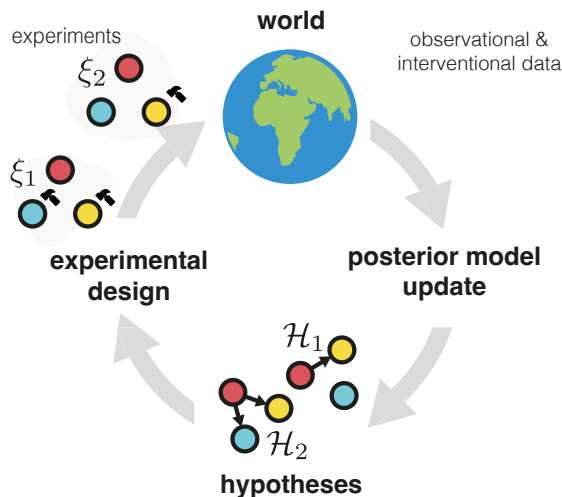
Figure 1: Causal Bayesian Experimental Design optimizes experiments that help disambiguate between competing causal hypotheses.

financial and ethical costs and risks involved in such experimentation, it is in the scientist's interest to minimize the number of batches required. This process is known as experimental design, and assuming that the question of interest concerns discovering the causal structure of the system of interest, the process is known as experimental design for causal discovery. A Bayesian framework for this process has been proposed in prior work (Tong & Koller, 2001; Murphy, 2001; Cho et al., 2016; Agrawal et al., 2019; Sussex et al., 2021; Tigas et al., 2022; Toth et al., 2022) which typically consists of updating an approximate posterior with past experimental data and using this updated posterior to compute experiments that are maximally informative, as evaluated by expected information gain—the objective of interest in *Bayesian Optimal Experimental Design* (BOED) (Lindley, 1956; Chaloner & Verdinelli, 1995).

The problem of Bayesian Optimal Experimental Design for Causal Discovery (BOECD) is hard; computing the Bayesian posterior over Structural Causal Models (SCM)—a common framework for capturing causal relationships—is intractable. More importantly, the estimation and optimization of batch BOED objectives are computationally challenging, which has resulted in heuristics like the greedy

Table 1: Comparison of different BOED for causal discovery methods based on their design space assumptions.

	Design Space Assumptions				
	Target Acquisition (Single Target)	State Acquisition (Single Target)	Target Acquisition (Multi-target)	State Acquisition (Multi-target)	Batch Acquisition
Murphy (2001)	✓				
Tong & Koller (2001)	✓				
Cho et al. (2016)	✓				
Agrawal et al. (2019)	✓				✓
Toth et al. (2022)	✓	✓			
Tigas et al. (2022)	✓	✓			✓
Sussex et al. (2021)	✓		✓		✓
Ours	✓	✓	✓	✓	✓

batch strategy (Agrawal et al., 2019; Sussex et al., 2021) and soft-top-k batch strategy (Tigas et al., 2022). Additionally, in causal discovery, one is interested not only in identifying the variable (target) to intervene *on* but also the state to set the intervention *to*, resulting in a design space which is a product space of discrete targets and continuous states, making experimental design even more challenging. Tigas et al. (2022) proposed to use Bayesian Optimization (BO) to optimize over the continuous state-space of the interventions and a soft-top-k heuristic to select a batch.

In this work, we propose a method for estimating and optimizing the BOED objective in a differentiable end-to-end manner, alleviating the inefficiencies introduced by the heuristics of the batch selection but also the black-box optimization over the intervening states. Specifically, we introduce estimators of mutual information based on nested estimation (Ryan, 2003; Myung et al., 2013; Huan & Marzouk, 2014; Foster et al., 2019) and importance sampling and extend it to the problem of causal discovery where the optimization is over both discrete nodes and continuous states. We cast the problem of batch experiment selection as a policy optimization where the policy uses either the Gumbel-Softmax or relaxed Bernoulli distribution (Jang et al., 2016; Maddison et al., 2016) for single target and multi-target settings respectively. When combined with the straight-through gradient estimator (Bengio et al., 2013) to optimize over the targets and gradient ascent over corresponding states, we can explore the space of optimal designs efficiently with powerful optimizers (Kingma & Ba, 2014). Our proposed method requires very few assumptions about the causal model and can explore wide range of design settings as compared to prior work (see Table 1), thus opening up possibilities of experimental design for causal discovery in a broader range of applications.

2. Related Work

Differentiable Bayesian Optimal Experimental Design. Huan & Marzouk (2014); Foster et al. (2019; 2020); Kleinegesse & Gutmann (2020; 2021) developed a unified framework for estimating expected information gain and opti-

mizing the designs with gradient-based methods. More recently, Ivanova et al. (2022) applied the Gumbel-Softmax relaxation within gradient-based BOED for contextual optimization. In Ivanova et al. (2021); Foster et al. (2021), the authors introduced a policy-based method for performing adaptive experimentation. More recently, work like Blau et al. (2022); Lim et al. (2022) used reinforcement learning to train policies for adaptive experimental design.

Experimental Design for Causal Discovery. One of the earliest works of experimental design for causal discovery in a BOED setting was proposed by (Murphy, 2001) and (Tong & Koller, 2001) in the case of discrete variables for single target acquisition. Since then, a number of works have attempted to address this problem for continuous variables in both the BOED framework (Agrawal et al., 2019; von Kügelgen et al., 2019; Toth et al., 2022; Cho et al., 2016) and other frameworks (Kocaoglu et al., 2017a; Gamella & Heinze-Deml, 2020; Eberhardt et al., 2012; Lindgren et al., 2018; Mokhtarian et al., 2022; Ghassami et al., 2018; Olko et al., 2022; Scherrer et al., 2021). In contrast to the setting studied in this paper, of particular note are the approaches for experimental design for causal discovery in a non-BOED setting in the presence of cycles (Mokhtarian et al., 2022) and latent variables (Kocaoglu et al., 2017b).

Closer to our BOED setting are the approaches of (Tigas et al., 2022) and (Sussex et al., 2021). Specifically, in (Tigas et al., 2022), the authors introduce a method for selecting single target-state pair with stochastic batch acquisition while (Sussex et al., 2021) introduce a method for selecting a batch of multi-target experiments with a greedy strategy, based on a gradient-based approximation to mutual information, without selecting the intervention state. Our presented method in contrast can acquire a batch of multi-target-state pairs.

Bayesian Causal Discovery. Designing experiments involves approximating the posterior of causal models to estimate mutual information. Approximating posterior distributions of causal structures and SCMs is hard due to combinatorially growing graph space (Heinze-Deml et al., 2018).

There are several works which have proposed various methods of approximate inference/ posterior sampling (Friedman et al., 2013; Annadani et al., 2021; Lorch et al., 2021; Cundy et al., 2021; Deleu et al., 2022; Nishikawa-Toomey et al., 2022), which could be used for our proposed design framework.

3. Background

3.1. Causality

Notation. Let $\mathbf{V} = \{1, \ldots, d\}$ be the vertex set of any Directed Acyclic Graph (DAG) $\mathbf{g} = (\mathbf{V}, E)$ and $\mathbf{X_V} = \{X_1, \ldots, X_d\} \subseteq \mathcal{X}$ be the random variables of interest indexed by $\mathbf{V}$.

Structural Causal Model. To deal with questions related to modelling causal relations between variables of interest, we employ the framework of Structural Causal Models (SCM) (Peters et al., 2017). In many fields of empirical sciences like network inference in single cell genomics (Greenfield et al., 2010), inferring protein-signalling networks (Sachs et al., 2005) and medicine (Shen et al., 2020), SCMs provide a framework to model the effects of interventions (Pearl, 2009)– experiments which perturb the state of a variable to a desired state, thereby allowing to study the mechanisms which affect the downstream variables (for example, CRISPR gene knockouts (Meinshausen et al., 2016) in genomics). Under this framework, each variable X_i has an associated *structural equation*, and is assigned a value which is a deterministic function of its direct causes $X_{\text{pa}(i)}$ as well as an exogenous noise variable ϵ_i with a distribution P_{ϵ_i}:

$$X_i := f_i(X_{\text{pa}(i)}, \epsilon_i) \quad \forall i \in \mathbf{V}$$

f_i's are mechanisms that relate how the direct causes affect the variable X_i. If the structural assignments are assumed to be acyclic, these equations induce a DAG $\mathbf{g} = (\mathbf{V}, E)$ whose vertices correspond to the variables and edges indicate direct causes. An intervention on any variable X_i corresponds to changing the structural equation of that variable to the desired state (value), $X_i := s_i$, where $s_i \in \mathcal{X}_i$. It is denoted by the do-operator (Pearl, 2009) as $\mathrm{do}(X_i = s_i)$.

In this work, we assume that the SCM is causally sufficient, i.e. all the variables are measurable, and the noise variables are mutually independent. Though the mechanisms f can be nonparametric in the general case, we assume that there exists a parametric approximation to these mechanisms with parameters $\gamma \in \Gamma$. In the case of linear SCMs, γ corresponds to the weights of the edges in E. We further require that the functions f are differentiable with respect to their parameters. Many classes of SCMs fall under this category, including the most commonly studied one – the Gaussian

additive noise models (ANM)[1]:

$$X_i := f_i(X_{\text{pa}(i)}; \gamma_i) + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \sigma_i^2)$$

Bayesian Causal Discovery. If the SCM for a given set of variables $\mathbf{X_V}$ is unknown, it has to be estimated from a combination of observational data (data obtained in an unperturbed state of a system) and experimental data under an intervention. This problem is called causal induction or causal discovery (Spirtes et al., 2000). This amounts to learning the parameters of the unknown SCM given by DAG $\mathbf{g}$, parameters of mechanisms, $\gamma = [\gamma_1, \ldots, \gamma_d]$, and variances, $\sigma^2 = [\sigma_1^2, \ldots, \sigma_d^2]$. For notational brevity, henceforth we denote $\phi = (\gamma, \sigma^2)$ and all the parameters of interest with $\theta = (\mathbf{g}, \phi)$. In Bayesian causal discovery (Heckerman et al., 1997), the parameters of SCM are treated as random variables whose beliefs are updated according to the Bayes rule. A Bayesian method for causal discovery is preferable to model epistemic uncertainty about the model due to finite data as well as characterize equivalence classes of SCM like Markov Equivalence Class (MEC) in the case of non-identifiability (Peters et al., 2012). Interventions improve identifiability, but they have to be planned carefully. After acquiring interventional data, Bayesian methods update the posterior distribution to reduce the uncertainty of the SCM.

3.2. Bayesian Optimal Experimental Design

Bayesian Optimal Experimental Design (BOED) (Lindley, 1956; Chaloner & Verdinelli, 1995) is an information theoretic approach to the problem of selecting the optimal experiment to estimate any parameter θ. For BOED, the *utility* of the experiment ξ is the mutual information (MI) between the observation $\mathbf{y}$ and θ:

$$\begin{aligned} \mathrm{U_{BOED}}(\xi) &\triangleq \mathcal{I}(\mathbf{Y}; \mathbf{\Theta} \mid \xi) \\ &= \underset{p(\theta)p(\mathbf{y}|\theta,\xi)}{\mathbb{E}} [\log p(\mathbf{y} \mid \theta, \xi) - \log p(\mathbf{y} \mid \xi)] \end{aligned}$$

This objective is also known as the *Expected Information Gain* (EIG). The goal of BOED is to select the experiment that maximizes this objective $\xi^* = \arg\max_\xi \mathrm{U_{BOED}}(\xi)$. Unfortunately, evaluating and optimizing this objective is challenging because of the nested expectations (Rainforth et al., 2018) and several estimators have been introduced (Foster et al., 2019; Kleinegesse & Gutmann, 2019), which can be combined with various optimization methods to select the designs (Foster et al., 2020; Ivanova et al., 2021; Foster et al., 2021; Blau et al., 2022).

A common setting, called *static*, *fixed* or *batch* design, is to optimize B designs $\{\xi_1, \ldots, \xi_B\}$ at the same time. The

---

[1]Note that differentiability of f is the only assumption we require with respect to an SCM. We do not require that the noise is additive. For clarity of exposition, we restrict our focus to an ANM as they are the most commonly studied class of SCMs.

designs are then executed, and the experimental outcomes are collected for a Bayesian update of the model parameters.

3.3. Causal Bayesian Experimental Design

Causal Bayesian Experimental Design is concerned with designing the most informative experiments to identify the true SCM so that the number of experiments required is as few as possible. An experiment in causal discovery corresponds to picking the intervention targets $I \in \mathcal{P}(\mathbf{V})$ and the corresponding states $S^I \in \bigcup_{i \in I} \mathcal{X}_i$ to set those targets to. A key component of such methods is computing a posterior over the parameters of the SCM. However, computing the posterior is a difficult task since the number of DAGs grows exponentially in the number of variables. Nevertheless, a plethora of methods exist (Friedman et al., 2013; Annadani et al., 2021; Lorch et al., 2021; Cundy et al., 2021) which can be used with our approach.

Having access to such posterior models, one can estimate the EIG objective. One difficulty that still remains though is that optimizing the EIG objective over the experiments is a mixed discrete and continuous optimization problem, for which previous work has proposed to find the optimal value per candidate target via the use of black-box methods like *Bayesian Optimization* (BO) (Tigas et al., 2022). Additionally, for the construction of the batch of experimental designs, a greedy approximation is used to incrementally select experiments, a method that is $1 - \frac{1}{\epsilon}$-approximate to the optimal solution (Krause & Golovin, 2014).

4. Differentiable Causal Bayesian Experimental Design

Let $\mathbf{\Theta}$ be a random variable that models the uncertainty in the parameters of the true SCM, of which $\boldsymbol{\theta} := (\mathbf{g}, \phi)$ is a realization. An experiment to identify an intervention is denoted by $\xi := \{(I, S^I)\} := \mathrm{do}(\mathbf{X}_I = S^I)$, where $I \in \mathcal{P}(\mathbf{V})$ is a set of target indices in the multi-target setting, and S^I are the corresponding states of those targets under intervention. The outcome of the experiment is denoted by $\mathbf{y} \sim P\left(X_1 = \mathrm{x}_1, \ldots, X_d = \mathrm{x}_d \mid \mathrm{do}\left(\mathbf{X}_I = S^I\right)\right) = p(\mathbf{y} \mid \xi)$. Here, $\mathbf{y}$ is an instance of the random variable $\mathbf{Y} \subseteq \mathcal{X}$ distributed according to the interventional distribution[2]. Due to causal sufficiency, the likelihood of data for a given $\boldsymbol{\theta}$ satisfies the causal Markov condition:

$$p(\mathbf{y} \mid \boldsymbol{\theta}, \xi) = \prod_{j \in \mathbf{V} \setminus I} p\left(x_j \mid \phi_j, \mathbf{x}_{\mathrm{pa}_{\mathbf{g}}(j)}, \mathrm{do}\left(\mathbf{X}_I = S^I\right)\right) \tag{1}$$

Along with a prior $p(\boldsymbol{\theta})$, the above equation defines a generative model of the data.

---

[2]Note that when $I = \varnothing$, it corresponds to an observational/ non-experimental setting. In this case, $\mathbf{Y} = \mathbf{X}_{\mathbf{V}}$.

Design setting. As in prior work (Tigas et al., 2022; Sussex et al., 2021), we are interested in the setting of batch design where we design B experiments at once before collecting experimental data. In other words, we seek a multiset of intervention targets and corresponding states which are jointly maximally informative about the parameters. We denote this multiset as $\xi_{1:B} := (I_{1:B}, S^I_{1:B})$. After executing a batch of experiments and collecting experimental outcomes, an experimenter might wish to design a new batch of experiments based on collected data (as summarized by the posterior distribution). Let h_t denote experimental history $(\xi^1, \mathbf{y}^1), \ldots, (\xi^t, \mathbf{y}^t)$ after t batches of acquisition. The BOED objective for this batch setting at any point t is given by the joint mutual information:

$$\mathcal{I}(\mathbf{Y}^t_{1:B}; \mathbf{\Theta} \mid \xi^t_{1:B}, h_{t-1})$$
$$= \mathop{\mathbb{E}}_{\substack{p(\boldsymbol{\theta} \mid h_{t-1}) \\ p(\mathbf{y}^t_{1:B} \mid \boldsymbol{\theta}, \xi^t_{1:B})}} \left[\log \frac{p(\mathbf{y}^t_{1:B} \mid \boldsymbol{\theta}, \xi^t_{1:B})}{p(\mathbf{y}^t_{1:B} \mid \xi^t_{1:B}, h_{t-1})} \right] \tag{2}$$

where $\mathbf{Y}^t_{1:B}$ are the random variables corresponding to experimental outcomes for iteration t, $\mathbf{y}^{\mathbf{t}}_{\mathbf{1:B}}$ are the instances of these random variables and $\xi^t_{1:B}$ is the corresponding multiset of experimental designs. We drop the superscript t from these variables for simplicity of exposition. Ideally, we wish to maximize the above objective by obtaining the gradients $\nabla_{\xi_{1:B}} I$ and performing gradient ascent. However, the above objective is doubly intractable (Rainforth et al., 2018) and approximations are required. This usually leads to a two-stage procedure where the above objective is first estimated with respect to an inference network and then maximized with respect to designs (Foster et al., 2019), which can be typically inefficient (Foster et al., 2020).

4.1. Estimators of the Joint Mutual Information

NESTED MONTE CARLO

Following (Huan & Marzouk, 2014; Foster et al., 2020; 2021), we consider an estimator that allows for approximating the EIG objective while *simultaneously* optimizing for the experiment ξ that maximizes the objective via gradient-based methods. This estimator, called Nested Monte Carlo (NMC), is based on contrastive estimation of the experimental likelihood and has been extensively used in Bayesian experimental design (Ryan, 2003; Myung et al., 2013). More precisely, assuming some past observational and interventional data $h_{t-1} = \{(\xi^1, \mathbf{y}^1), \ldots, (\xi^{t-1}, \mathbf{y}^{t-1})\}$, for every parameter sample from the posterior distribution $\boldsymbol{\theta}_0 \sim p(\boldsymbol{\theta} \mid h_{t-1})$, a set of contrastive samples $\boldsymbol{\theta}_{1:L} \sim p(\boldsymbol{\theta} \mid h_{t-1})$ are considered to obtain a unified objective:

$$\mathcal{U}^t_{\mathrm{NMC}}(\xi_{1:B}) = \mathop{\mathbb{E}}_{\substack{p(\boldsymbol{\theta}_{0:L} \mid h_{t-1}) \\ p(\mathbf{y}_{1:B} \mid \boldsymbol{\theta}_0, \xi_{1:B})}} \left[\log \frac{p(\mathbf{y}_{1:B} \mid \xi_{1:B}, \boldsymbol{\theta}_0)}{\frac{1}{L} \sum_{\ell=1}^L p(\mathbf{y}_{1:B} \mid \xi_{1:B}, \boldsymbol{\theta}_\ell))} \right] \tag{3}$$

This estimator converges to the true mutual information as $L \to \infty$ (Rainforth et al., 2018). If the design space is

continuous, the optimal *batch* of experiment $\xi_{1:B}^*$ can be found by *directly* maximizing the NMC objective ($\xi_{1:B}^* \leftarrow \arg\max_{\xi_{1:B}} \mathcal{U}_{\text{NMC}}^t(\xi_{1:B})$) with gradient-based techniques (Huan & Marzouk, 2014).

The above objective requires estimating the posterior distribution $p(\boldsymbol{\theta} \mid h_{t-1})$ after every acquisition. For causal models, while it is generally hard to estimate this posterior due to DAG space of causal structures being discrete and super-exponential in the number of variables (Tong & Koller, 2001), many approaches exist in the literature (Agrawal et al., 2019; Lorch et al., 2021; Cundy et al., 2021). These approximate posteriors can be nevertheless used for estimating the NMC objective.

IMPORTANCE WEIGHTED NESTED MONTE CARLO

To establish an alternative path to estimating the mutual information, we begin by utilizing an observation from Foster et al. (2019) that it is possible to draw the contrastive samples from a distribution other than $p(\boldsymbol{\theta} \mid h_{t-1})$ and obtain an asymptotically exact estimator, up to a constant C that does not depend on $\xi_{1:B}^t$. Drawing samples from the *original* prior $p(\boldsymbol{\theta})$ gives the estimator

$$\mathcal{I}(\mathbf{Y}_{1:B}^t; \boldsymbol{\Theta} \mid \xi_{1:B}^t, h_{t-1}) - C = \lim_{L \to \infty}$$

$$\mathop{\mathbb{E}}_{\substack{p(\boldsymbol{\theta}_0 | h_{t-1}) p(\boldsymbol{\theta}_{1:L}) \\ p(\mathbf{y}_{1:B} | \boldsymbol{\theta}_0, \xi_{1:B})}} \left[\log \frac{p(\mathbf{y}_{1:B} | \xi_{1:B}, \boldsymbol{\theta}_0)}{\frac{1}{L} \sum_{\ell=1}^L p(\mathbf{y}_{1:B} | \xi_{1:B}, \boldsymbol{\theta}_\ell) p(h_{t-1} | \boldsymbol{\theta}_\ell)} \right]$$

The remaining wrinkle is that we must sample $\boldsymbol{\theta}_0$ from $p(\boldsymbol{\theta}_0 | h_{t-1})$. We propose the conceptually simplest approach of applying self-normalized importance sampling (SNIS) to the outer expectation. The resulting objective, based on efficiently re-using samples in a leave-one-out manner, can optimize designs by just sampling parameters from the prior, without having to estimate the posterior:

$$\mathcal{U}_{\text{IWNMC}}^t(\xi_{1:B}) =$$

$$\mathbb{E} \left[\sum_{m=1}^L \omega_m \log \frac{p(\mathbf{y}_{m,1:B} | \boldsymbol{\theta}_m, \xi_{1:B})}{\frac{1}{L-1} \sum_{\ell \neq m} p(\mathbf{y}_{m,1:B} | \boldsymbol{\theta}_\ell, \xi_{1:B}) p(h_{t-1} | \boldsymbol{\theta}_\ell)} \right] \quad (4)$$

where $\boldsymbol{\theta}_{1:L} \sim p(\boldsymbol{\theta}_{1:L})$ are sampled from the original prior, $\mathbf{y}_{m,1:B} \sim p(\mathbf{y}_{1:B} | \boldsymbol{\theta}_m, \xi_{1:B})$ are all the experimental outcomes in the batch for parameter $\boldsymbol{\theta}_m$ and $\omega_m \propto p(h_{t-1} | \boldsymbol{\theta}_m)$ are self-normalized weights.

A full derivation is given in Section A.

As IWNMC does not require any posterior estimation but instead relies entirely on the prior, it completely sidesteps the causal discovery process for designing experiments. This is a paradigm change from the NMC estimator which requires causal discovery through the estimation of the posterior.

However, we note that using IWNMC with just the prior (Eq. 4) as opposed to NMC (Eq. 3) comes with trade-offs. IWNMC typically requires a large L to get a good estimate of the EIG. In high dimensions, this can be computationally infeasible. Having a small L on the other hand might result in a failure case if the effective sample size of importance samples becomes 1. We can alleviate this issue if there is some prior information available which could be leveraged to design better proposal distributions. This might consist of knowledge of certain causal mechanisms of the system under study or access to some initial observational data. In such a case, a proposal distribution which encodes this information (for example with support on graphs which are in the Markov Equivalence Class (MEC) of the observational distribution) can be used instead of the prior. If no prior information is available or a good approximate inference technique is at our disposal, NMC is preferable in high dimensions. Surprisingly, we get good results on variables of size up to 5 with IWNMC from just the prior and up to 40 variables from a proposal distribution which has support on the MEC of observational distribution (see Sec 5.5).

4.2. Optimizing over Targets and States (`DiffCBED`)

While the NMC estimator provides a unified objective to directly optimize over the designs $\xi_{1:B}$, it requires that the design space is continuous so that the gradients $\frac{\partial \mathcal{U}_{\text{NMC}}}{\partial I_{1:B}}$ and $\frac{\partial \mathcal{U}_{\text{NMC}}}{\partial S_{1:B}^I}$ can be computed. However, in the case of designing experiments for causal models, the challenge still remains that optimizing over intervention targets I with gradient-based techniques is not possible because it is a discrete choice.

In order to address this problem, we introduce a *design policy* π_ϕ with learnable parameters ϕ that parameterize a joint distribution over possible intervention targets and corresponding states. Instead of seeking the gradients $\frac{\partial \mathcal{U}_{\text{NMC}}}{\partial I_{1:B}}$ and $\frac{\partial \mathcal{U}_{\text{NMC}}}{\partial S_{1:B}^I}$, the goal now instead is to estimate $\frac{\mathcal{U}_{\text{NMC}}}{\partial \phi}$ so that policy can be updated to be close to optimal. Such a characterization of the design space allows us to use continuous relaxations of discrete distributions (Maddison et al., 2016; Jang et al., 2016) to obtain samples of designs and estimate NMC gradients.

Let $\mathbf{I}$ and $\mathbf{S}$ be the random variables which model all possible intervention target combinations and states for a batch design respectively. While there are many possibilities of instantiating the policy in practice, we consider the simplest case where $\pi_\phi(\mathbf{I}, \mathbf{S}) \triangleq \pi_{\phi_n}(\mathbf{I}) \pi_{\phi_m}(\mathbf{S})$. As the state space is continuous[3], π_{ϕ_m} can be either deterministic (a delta Dirac with $\phi_m \in \mathbb{R}^{B \times d}$) or Gaussian with $\phi_m \in \mathbb{R}^{2 \times B \times d}$ parameterizing its mean and log variance. In this work, we found

---

[3]If the state space is discrete, optimizing π_{ϕ_m} would be similar to π_{ϕ_n} which involves reparameterized gradients.

it sufficient to use a deterministic policy over the state space. For the interventional targets, $\phi_n \in \mathbb{R}^{B \times d}$ parameterizes the logits of different relaxed versions of discrete distributions depending on the setting, which we describe below.


Algorithm 1: Differentiable CBED (`DiffCBED`)

Input : $\mathcal{E}$ SCM Environment, N Initial observational samples, B Batch Size

1 $\mathcal{D}_{\text{obs}} \leftarrow \mathcal{E}.\text{sample}(N), \mathcal{D}_{\text{int}} \leftarrow \varnothing$

2 Train $q(\boldsymbol{\Theta} \mid \mathcal{D}_{\text{obs}}) \approx p(\boldsymbol{\Theta} \mid \mathcal{D}_{\text{int}})$ using appropriate algorithm.

3 **for** *batch* $t = 1 \ldots T$ *Batches* **do**

4   Initialize design policy parameters $\phi = \{\phi_n, \phi_m\}$: trainable logits ϕ_n for the targets; trainable parameters ϕ_m for the states.

5   **for** *update step* $c = 1 \ldots C$ **do**

        ▷ Sample Interventional Targets and States

6         $\{\xi_{1:B}^{(o)}\}_{o=1}^O \sim \pi_\phi(\mathbf{I}, \mathbf{S})$

        ▷ Gradient ascent with straight-through gradient estimator

7   Update
$$\phi \rightarrow \phi + \alpha \frac{\partial}{\partial \phi} \frac{1}{O} \sum_{o=1}^O \left[\mathcal{U}_{\text{NMC}}^t(\xi_{1:B}^{(o)}) \right]$$

    ▷ Intervene with learned policy

8     $\xi_{1:B} \sim \pi_\phi$

9     $\mathcal{D}_{\text{int}} \leftarrow \mathcal{D}_{\text{int}} \cup \mathcal{E}.\text{intervene}(\xi_{1:B})$

10   Update the posterior $q(\boldsymbol{\Theta} \mid \mathcal{D}_{\text{obs}} \cup \mathcal{D}_{\text{int}})$.

SINGLE TARGET ($q = 1$)

In this setting, the intervention targets are one-hot vectors, as demonstrated in Figure 2. To sample one-hot vectors in a differentiable manner, we parametrize π_{ϕ_n} as a Gumbel-Softmax distribution (Maddison et al., 2016; Jang et al.,
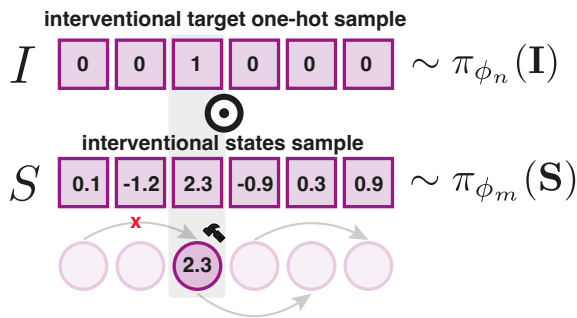
Figure 2: A design sample is obtained by first sampling $I_{1:B} \sim \pi_{\phi_n}(\mathbf{I})$, $S_{1:B} \sim \pi_{\phi_m}(\mathbf{S})$ and then setting states to be $S_{1:B}^I = S_{1:B} \odot I_{1:B}$. To obtain hard samples of I, we use the straight-through estimator (Bengio et al., 2013). Illustration for $B = 1$.

2016) over intervention targets, which is a continuous relaxation of the categorical distribution (in *one-hot* form). Additionally, we use the straight-through (ST) gradient estimator (Bengio et al., 2013).

UNCONSTRAINED MULTI-TARGET ($q \leq d$)

If instead of a continuous relaxation of the categorical distribution, we parametrise the policy π_{ϕ_n} as a continuous relaxation of the Bernoulli distribution (Binary Concrete) (Maddison et al., 2016), we can now sample multi-target experiments. Since each interventional target sample will have at most d non-zero entries, this policy is suitable for multi-target experiments with an unconstrained number of interventions per experiment.

CONSTRAINED MULTI-TARGET ($q = k$)

Finally, when considering a setting where the number of targets per intervention is exactly k. However, this is a significantly more challenging case, since the policy needs to select a subset of k from d nodes. By using a continuous relaxation of subset sampling, as introduced in Xie & Ermon (2019), combined with straight-through gradient estimator, we can efficiently optimize the policy to select a subset of nodes to intervene on. Note that with $k = 1$, this would be equal to the single-target setting.

The `DiffCBED` algorithm is outlined in Algorithm 1.

5. Experiments

We evaluate the performance of our method on synthetic and semi-synthetic datasets and a range of baselines. We aim to investigate the following aspects empirically: (1) To what extent can we design good experiments without performing intermediate causal discovery/ posterior estimation with IWNMC estimator from the prior? (2) Ability to design good experiments with a proposal distribution with IWNMC (3) the performance of our policy-based design in combination with the differentiable NMC estimator in single-target and multi-target settings, as compared to suitable baselines.

5.1. Bivariate Setting

First, we demonstrate the method in a two variable setting to qualitatively assess the objective and the optimization method. Since computing the posterior over graphs and parameters is intractable in the general case, as a first step to study how well we can optimize the EIG objective, we assume a simple two-variable SCM. To compute the posterior we enumerate all DAGs with two nodes and parameterize the conditional distributions as neural network parameterized Gaussian distribution ($\mathcal{N}(X_i; \mu_{\text{NN}}(X_{\text{pa}(i)}), \sigma_{\text{NN}}(X_{\text{pa}(i)}))$). We compute the posterior over the parameters of the conditional distributions via Monte-Carlo dropout (Gal & Ghahra-

6
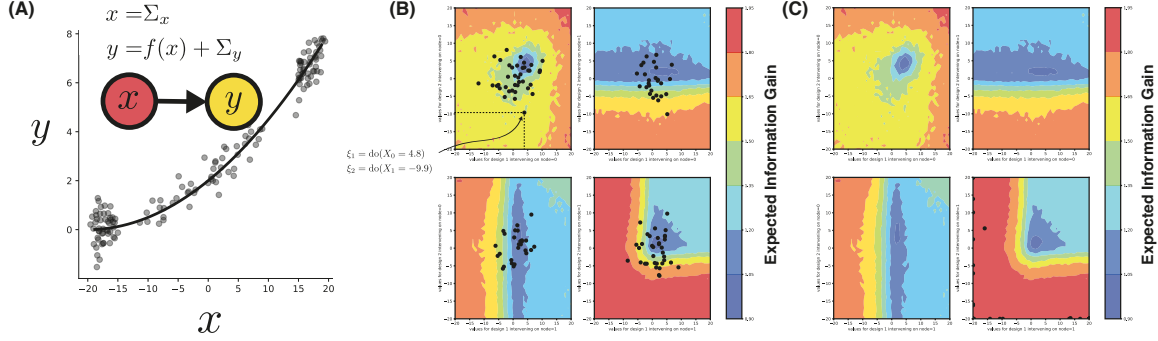
Figure 3: Two variables and two experiments scenario. We assume a ground-truth graph G_T of two nodes $X_1 \to X_2$. (A) The conditional distribution $p(X_2 \mid X_1)$ is shown in **(A)**. The corresponding SCM (Fig. A) is $x_1 = \Sigma_{x_1}$ and $x_2 = f(x_1) + \Sigma_{x_2}$. The four panels represent the EIG of all possible experiments of batch size two, when intervening on nodes $(0,0), (0,1), (1,0), (1,1)$. (B, C) Each panel shows how the EIG change on different interventional states. E.g. right top panel shows how EIG changes when applying interventions with states in ranges $[-20, 20]$. Fig. B Shows the designs before optimizing the objective and Fig. C after. As we can observe that the algorithm successfully places the designs (samples from the policy) on the high EIG (1.95) area of the plot (● on the plot).
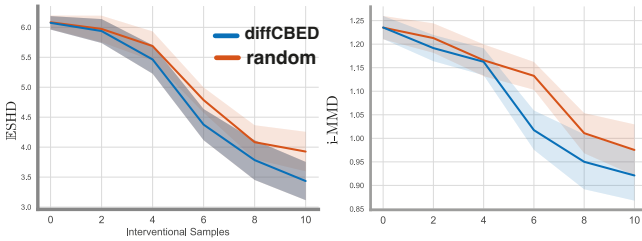
Figure 4: We test the designs acquired with IWNMC estimator with just the prior as opposed to the random policy (with random target and state acquisition) on variables of 5 dimensions. Plots correspond to unconstrained multi-target setting with $B = 2$ (shaded area represents 95% confidence intervals - 60 seeds).

mani, 2016). We parameterize the intervention targets policy with Gumbel-Softmax and interventional states policy with a Gaussian distribution. The final policy consists of the logits of the Gumbel-Softmax and the sufficient statistics of the Gaussian distribution. We use Adam optimizer (Kingma & Ba, 2014) to optimize the parameters of the policy. As we can see in Fig.3(B-C), the optimizer successfully concentrates the policy on the nodes and states that maximize the EIG objective.

5.2. Results

We present experimental results in various settings for the following metrics:

EVALUATION METRICS

Expected SHD: This metric evaluates the expected structural hamming distance between the graphs sampled from the posterior model and the ground truth graph.

**Expected F1-score**: This metric evaluates the expected f1-

score of each edge being present or absent in the graphs sampled from the posterior as compared to the ground truth graph.

i-MMD: interventional MMD distance uses the non-parametric distance metric MMD (Gretton et al., 2012). As compared to the graph evaluation metrics, this metric evaluates the interventional distributions induced by both the graph structure and the corresponding conditional distributions. We provide the full definition in Appendix C.

5.3. Evaluation of the IWNMC estimator

In this section, we consider optimizing the designs with respect to the IWNMC estimator entirely from the prior, introduced in 4.1, sidestepping the causal discovery procedure. As noted before, estimating posteriors of causal models is hard, so it is important to understand to what extent IWNMC can be considered a suitable candidate for designing good experiments in the absence of a posterior. For this setting, we sample from the prior distribution over graphs by first sampling an ordering of nodes at random and then sampling edges with probability $p = 0.25$ which adhere to this topological order. We sample the mechanism parameters and noise variances of ANM at random from a Gaussian distribution with mean 0 and variance 1.

Figure 4 demonstrates results for 5 variable unconstrained multi-target setting with batch size 2. For evaluation, we train DAG Bootstrap (Friedman et al., 2013) with GIES (Hauser & Bühlmann, 2012) on the data acquired from each policy. We can see that we can recover the ground truth SCM faster than a random strategy. This is a surprising, but positive result given that our policy was trained entirely from samples from the prior. We also tested this approach for 10 variables (results in Appendix E). While this resulted in better performance of the policy as opposed to random in
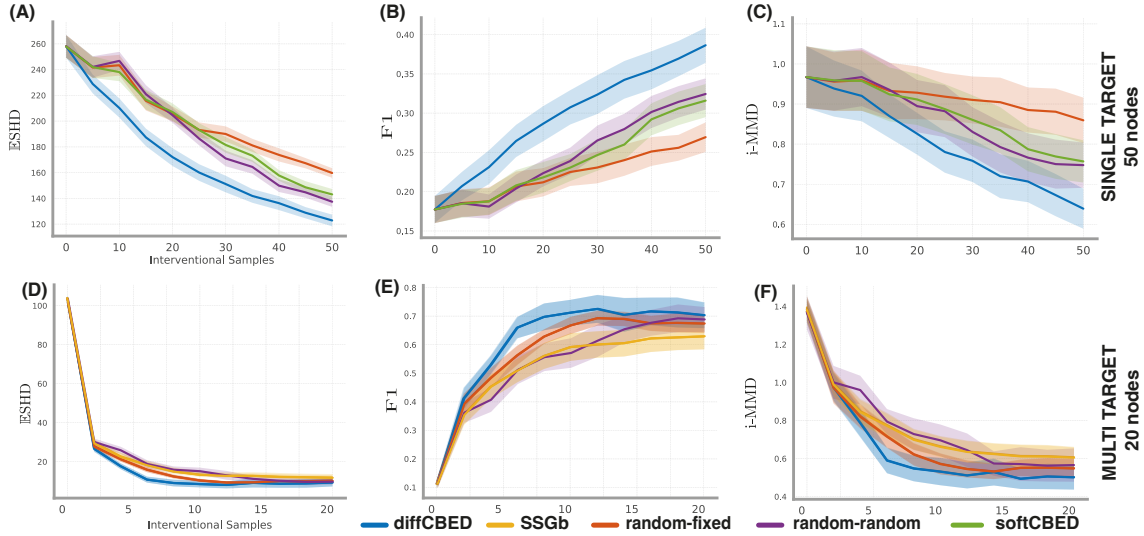
Figure 5: (A,B,C) Single target-state design setting results for Erdős–Rényi (Erdős & Rényi, 1959) graphs with $d = 50$ variables. **(D,E,F)** Multi target-state design setting results for Erdős–Rényi (Erdős & Rényi, 1959) graphs with $d = 20$ variables. Each experiment was run with 30 random seeds (shaded area represents 95% CIs)

terms of downstream metrics, we observed effective sample size reach 1, indicating that for 10 dimensions or higher, we might need a better proposal distribution or a posterior estimate.

5.4. Baselines

Before we evaluate the IWNMC estimator with a proposal distribution more informative than the prior and the NMC estimator with a posterior estimate of SCM, we present the baselines with which we can compare the designs.

SINGLE TARGET

Random-Fixed: Uniform random selection of target, fixing the state to a value of 0 (as introduced in (Agrawal et al., 2019; Tigas et al., 2022)). **Random-Random**: Uniform selection of node, uniform selection of state (introduced in (Toth et al., 2022)). **SoftCBED**: A stochastic approximation of greedy batch selection as introduced in (Tigas et al., 2022).

MULTI-TARGET

Random-Random: Multi-target version of uniform selection of node, uniform selection of value (introduced in (Toth et al., 2022)). **Random-Fixed**: Multitarget version of uniform selection of node, fixed value to 5 (Sussex et al., 2021), as suggested by the authors. **SSGb**: Finite sample baseline from (Sussex et al., 2021) with fixed value equal 5. We emphasize that in contrast to our method, the baselines cannot select states, but they either assume a fixed predefined state or select a state at random.

5.5. Evaluation in Higher Dimensions

EVALUATION OF IWNMC WITH PROPOSAL DISTRIBUTION

In this experiment, we consider 40 variables, constrained ($q = 5$) multi-target and batch size $B = 2$. Further, we use the same setup as Sussex et al. (2021) to make a fair comparison as well as to construct a proposal distribution. For the proposal distribution, we use 800 observational samples to train DAG Bootstrap (Friedman et al., 2013; Agrawal et al., 2019) and augment our posterior samples with samples of DAGs from the MEC of the true graph, to make sure that there is sufficient support within the MEC of the true graph (see Sussex et al. (2021) for details). We then acquire a single batch of experiments from IWNMC estimator for our approach. For the baseline, we acquire a single batch of experiments from the estimator defined in (Sussex et al., 2021).

For random and SSGb baseline, we set the interventional state (value) to 5, as explained in (Sussex et al., 2021). Our approach doesn't fix the state to 5 but optimizes over states to perform the intervention with. In Table 2 we summarize our results. As we can see, our method outperforms random and SSGb by a great margin, indicating that with a good proposal distribution, IWNMC can still be a promising candidate in higher dimensions.

RESULTS WITH NMC ESTIMATOR

For the following results, we use DAG-Bootstrap (Agrawal et al., 2019), an approximate posterior method based on GIES causal discovery method (Hauser & Bühlmann, 2012).
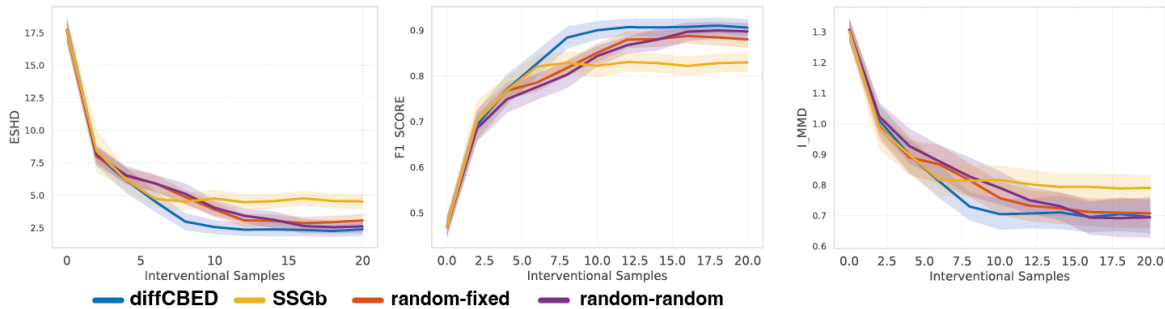
Figure 6: Results on the E. Coli gene interaction network ($d = 10$) of the DREAM (Greenfield et al., 2010) simulator. The data is generated based on this graph by simulating the noise and mechanisms. We find that our method performs better than the baselines. Shaded area represents 95% CI.

Table 2: Results of multi-target experiments on graphs of size 40 (30 seeds $\pm$ s.e.). Similarly to (Sussex et al., 2021), we are using posterior samples trained on observational data and re-weighting them with likelihoods.

Method	$\mathbb{E}$SHD $\downarrow$	F1 $\uparrow$	iMMD $\downarrow$
Random	43.78 ± 46.67	0.91 ± 0.08	0.16 ± 0.07
SSGb	15.59 ± 29.66	0.97 ± 0.05	0.10 ± 0.06
diffCBED	0.44 ± 0.21	0.99 ± 0.00	0.07 ± 0.01

As GIES is not a differentiable method, once we compute the posterior, we transfer the posterior samples (the boot-straps) into JAX (jax) tensors to allow for the gradients to be computed with respect to the experiments.

Single-target synthetic graphs: In this experiment, we test against synthetic graphs of 50 variables and batch size 5, where the graph is sampled from the class of Erdos-Renyi (a common benchmark in the literature (Tigas et al., 2022; Toth et al., 2022; Scherrer et al., 2021)). In Figure 10 **(A,B,C)** we summarize the results. We observe that our method performs significantly better than the baselines.

20 **nodes, unconstrained** ($q \leq 20$)**, batch size $B = 2$:** In this experiment, we evaluate the performance of our method as compared to the baselines, on sparse graphs over several acquisitions. Figure 10 **(D,E,F)** summarizes the results of this setting. We observe strong empirical performance as compared to all the baselines. Additional results are given in Section G.1.

5.6. Evaluation on Semi-Synthetic Data

We evaluate our proposed design framework on semi-synthetic setting based on the DREAM gene networks (Greenfield et al., 2010). In particular, we use the E. Coli gene interaction network, which is real-world inspired gene regulatory network of 10 variables, and simulate the mechanisms (and hence the data generation process). Then

we run a random node with a fixed intervention value, a random node with a random value, SSGb finite sample and our algorithm. The results are presented in Figure 6. We find that our algorithm performs better than the baselines, indicating the potential of gradient-based approach to more realistic settings.

6. Discussion

Limitations: A primary limitation of our method is that it needs to estimate a posterior after every acquisition. While the proposed IWNMC estimator presents an interesting alternative, the designs are still non-adaptive. An interesting direction is to train a policy to design adaptive experiments (Foster et al., 2021; Greenewald et al., 2019).

Conclusion: We present a gradient-based method for differentiable Bayesian optimal experimental design for causal discovery. Our method allows not only for single-target but also various multi-target (constrained and unconstrained) batch design of experiments. While prior work in relies on greedy approximations for the selection of a batch (Agrawal et al., 2019; Tigas et al., 2022) or black-box methods (Toth et al., 2022; Tigas et al., 2022) for optimizing over interventional states, our method utilizes gradient-based optimization procedures to simultaneously optimize for various design choices. Evaluation on different datasets suggests that our method is competitive with baselines.

References

Jax: Autograd and XLA. https://github.com/google/jax. 9

Agrawal, R., Squires, C., Yang, K., Shanmugam, K., and

Uhler, C. Abcd-strategy: Budgeted experimental design for targeted causal structure discovery. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pp. 3400–3409. PMLR, 2019. 1, 2, 5, 8, 9

Annadani, Y., Rothfuss, J., Lacoste, A., Scherrer, N., Goyal, A., Bengio, Y., and Bauer, S. Variational causal networks: Approximate bayesian inference over causal structures. *arXiv preprint arXiv:2106.07635*, 2021. 3, 4

Batagelj, V. and Brandes, U. Efficient generation of large random networks. *Physical Review E*, 71(3):036113, 2005. 16

Bengio, Y., Léonard, N., and Courville, A. Estimating or propagating gradients through stochastic neurons for conditional computation. *arXiv preprint arXiv:1308.3432*, 2013. 2, 6

Blau, T., Bonilla, E. V., Chades, I., and Dezfouli, A. Optimizing sequential experimental design with deep reinforcement learning. In *International Conference on Machine Learning*, pp. 2107–2128. PMLR, 2022. 2, 3

Chaloner, K. and Verdinelli, I. Bayesian experimental design: A review. *Statistical Science*, pp. 273–304, 1995. 1, 3

Cho, H., Berger, B., and Peng, J. Reconstructing causal biological networks through active learning. *PloS one*, 11 (3):e0150611, 2016. 1, 2

Cundy, C., Grover, A., and Ermon, S. Bcd nets: Scalable variational approaches for bayesian causal discovery. *Advances in Neural Information Processing Systems*, 34, 2021. 3, 4, 5

Deleu, T., Góis, A., Emezue, C., Rankawat, M., Lacoste-Julien, S., Bauer, S., and Bengio, Y. Bayesian structure learning with generative flow networks. *arXiv preprint arXiv:2202.13903*, 2022. 3

Eberhardt, F., Glymour, C., and Scheines, R. On the number of experiments sufficient and in the worst case necessary to identify all causal relations among n variables. *arXiv preprint arXiv:1207.1389*, 2012. 2

Erdős, P. and Rényi, A. On random graphs i. publicationes mathematicae (debrecen). 1959. 8, 15, 16

Foster, A., Jankowiak, M., Bingham, E., Horsfall, P., Teh, Y. W., Rainforth, T., and Goodman, N. Variational bayesian optimal experimental design. *arXiv preprint arXiv:1903.05480*, 2019. 2, 3, 4, 5, 13

Foster, A., Jankowiak, M., O'Meara, M., Teh, Y. W., and Rainforth, T. A unified stochastic gradient approach to designing bayesian-optimal experiments. In *International Conference on Artificial Intelligence and Statistics*, pp. 2959–2969. PMLR, 2020. 2, 3, 4

Foster, A., Ivanova, D. R., Malik, I., and Rainforth, T. Deep adaptive design: Amortizing sequential bayesian experimental design. *arXiv preprint arXiv:2103.02438*, 2021. 2, 3, 4, 9

Friedman, N., Goldszmidt, M., and Wyner, A. Data analysis with bayesian networks: A bootstrap approach. *arXiv preprint arXiv:1301.6695*, 2013. 3, 4, 7, 8

Gal, Y. and Ghahramani, Z. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning*, pp. 1050–1059. PMLR, 2016. 6

Gamella, J. L. and Heinze-Deml, C. Active invariant causal prediction: Experiment selection through stability. *arXiv preprint arXiv:2006.05690*, 2020. 2

Ghassami, A., Salehkaleybar, S., Kiyavash, N., and Bareinboim, E. Budgeted experiment design for causal structure learning. In *International Conference on Machine Learning*, pp. 1724–1733. PMLR, 2018. 2

Greenewald, K., Katz, D., Shanmugam, K., Magliacane, S., Kocaoglu, M., Boix Adsera, E., and Bresler, G. Sample efficient active learning of causal trees. *Advances in Neural Information Processing Systems*, 32, 2019. 9

Greenfield, A., Madar, A., Ostrer, H., and Bonneau, R. Dream4: Combining genetic and dynamic information to identify biological networks and dynamical models. *PloS one*, 5(10):e13397, 2010. 3, 9

Gretton, A., Borgwardt, K. M., Rasch, M. J., Schölkopf, B., and Smola, A. A kernel two-sample test. *The Journal of Machine Learning Research*, 13(1):723–773, 2012. 7, 15

Hauser, A. and Bühlmann, P. Characterization and greedy learning of interventional markov equivalence classes of directed acyclic graphs. *The Journal of Machine Learning Research*, 13(1):2409–2464, 2012. 7, 8

Heckerman, D., Meek, C., and Cooper, G. A bayesian approach to causal discovery. Technical report, Technical report msr-tr-97-05, Microsoft Research, 1997. 3

Heinze-Deml, C., Peters, J., and Meinshausen, N. Invariant causal prediction for nonlinear models. *Journal of Causal Inference*, 6(2), 2018. 2

Huan, X. and Marzouk, Y. Gradient-based stochastic optimization methods in bayesian experimental design. *International Journal for Uncertainty Quantification*, 4(6), 2014. 2, 4, 5

Ivanova, D. R., Foster, A., Kleinegesse, S., Gutmann, M. U., and Rainforth, T. Implicit deep adaptive design: policy-based experimental design without likelihoods. *Advances in Neural Information Processing Systems*, 34:25785–25798, 2021. 2, 3

Ivanova, D. R., Jennings, J., Zhang, C., and Foster, A. Efficient real-world testing of causal decision making via bayesian experimental design for contextual optimisation. *ICML 2022 Workshop on Adaptive Experimental Design and Active Learning in the Real World*, 2022. 2

Jang, E., Gu, S., and Poole, B. Categorical reparameterization with gumbel-softmax. *arXiv preprint arXiv:1611.01144*, 2016. 2, 5, 6

Kingma, D. P. and Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 2, 7

Kleinegesse, S. and Gutmann, M. Bayesian experimental design for implicit models by mutual information neural estimation. In *Proceedings of the 37th International Conference on Machine Learning*, Proceedings of Machine Learning Research, pp. 5316–5326. PMLR, 2020. 2

Kleinegesse, S. and Gutmann, M. U. Efficient bayesian experimental design for implicit models. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pp. 476–485. PMLR, 2019. 3

Kleinegesse, S. and Gutmann, M. U. Gradient-based bayesian experimental design for implicit models using mutual information lower bounds. *arXiv preprint arXiv:2105.04379*, 2021. 2

Kocaoglu, M., Dimakis, A., and Vishwanath, S. Cost-optimal learning of causal graphs. In *International Conference on Machine Learning*, pp. 1875–1884. PMLR, 2017a. 2

Kocaoglu, M., Shanmugam, K., and Bareinboim, E. Experimental design for learning causal graphs with latent variables. *Advances in Neural Information Processing Systems*, 30, 2017b. 2

Krause, A. and Golovin, D. Submodular function maximization. *Tractability*, 3:71–104, 2014. 4

Lim, V., Novoseller, E., Ichnowski, J., Huang, H., and Goldberg, K. Policy-based bayesian experimental design for non-differentiable implicit models. *arXiv preprint arXiv:2203.04272*, 2022. 2

Lindgren, E., Kocaoglu, M., Dimakis, A. G., and Vishwanath, S. Experimental design for cost-aware learning of causal graphs. *Advances in Neural Information Processing Systems*, 31, 2018. 2

Lindley, D. V. On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, pp. 986–1005, 1956. 1, 3

Lorch, L., Rothfuss, J., Schölkopf, B., and Krause, A. Dibs: Differentiable bayesian structure learning. *arXiv preprint arXiv:2105.11839*, 2021. 3, 4, 5

Maddison, C. J., Mnih, A., and Teh, Y. W. The concrete distribution: A continuous relaxation of discrete random variables. *arXiv preprint arXiv:1611.00712*, 2016. 2, 5, 6

Meinshausen, N., Hauser, A., Mooij, J. M., Peters, J., Versteeg, P., and Bühlmann, P. Methods for causal inference from gene perturbation experiments and validation. *Proceedings of the National Academy of Sciences*, 113(27): 7361–7368, 2016. 3

Mokhtarian, E., Salehkaleybar, S., Ghassami, A., and Kiyavash, N. A unified experiment design approach for cyclic and acyclic causal models. *arXiv preprint arXiv:2205.10083*, 2022. 2

Murphy, K. P. Active learning of causal bayes net structure. 2001. 1, 2

Myung, J. I., Cavagnaro, D. R., and Pitt, M. A. A tutorial on adaptive design optimization. *Journal of mathematical psychology*, 57(3-4):53–67, 2013. 2, 4

Nishikawa-Toomey, M., Deleu, T., Subramanian, J., Bengio, Y., and Charlin, L. Bayesian learning of causal structure and mechanisms with gflownets and variational bayes. *arXiv preprint arXiv:2211.02763*, 2022. 3

Olko, M., Zając, M., Nowak, A., Scherrer, N., Annadani, Y., Bauer, S., Kuciński, Ł., and Miłoś, P. Trust your ∇: Gradient-based intervention targeting for causal discovery. *arXiv preprint arXiv:2211.13715*, 2022. 2

Pearl, J. *Causality*. Cambridge university press, 2009. 3

Peters, J., Mooij, J., Janzing, D., and Schölkopf, B. Identifiability of causal graphs using functional models. *arXiv preprint arXiv:1202.3757*, 2012. 3

Peters, J., Janzing, D., and Schölkopf, B. *Elements of causal inference: foundations and learning algorithms*. The MIT Press, 2017. 3

Rainforth, T., Cornish, R., Yang, H., Warrington, A., and Wood, F. On nesting monte carlo estimators. In *International Conference on Machine Learning*, pp. 4267–4276. PMLR, 2018. 3, 4

Ryan, K. J. Estimating expected information gains for experimental designs with application to the random fatigue-limit model. *Journal of Computational and Graphical Statistics*, 12(3):585–603, 2003. 2, 4

Sachs, K., Perez, O., Pe'er, D., Lauffenburger, D. A., and Nolan, G. P. Causal protein-signaling networks derived from multiparameter single-cell data. *Science*, 308(5721): 523–529, 2005. 3

Scherrer, N., Bilaniuk, O., Annadani, Y., Goyal, A., Schwab, P., Schölkopf, B., Mozer, M. C., Bengio, Y., Bauer, S., and Ke, N. R. Learning neural causal models with active interventions. *arXiv preprint arXiv:2109.02429*, 2021. 2, 9

Shen, X., Ma, S., Vemuri, P., and Simon, G. Challenges and opportunities with causal discovery algorithms: application to alzheimer's pathophysiology. *Scientific reports*, 10(1):2975, 2020. 3

Spirtes, P., Glymour, C. N., Scheines, R., and Heckerman, D. *Causation, prediction, and search*. MIT press, 2000. 3

Sussex, S., Uhler, C., and Krause, A. Near-optimal multi-perturbation experimental design for causal structure learning. *Advances in Neural Information Processing Systems*, 34:777–788, 2021. 1, 2, 4, 8, 9

Tigas, P., Annadani, Y., Jesson, A., Schölkopf, B., Gal, Y., and Bauer, S. Interventions, where and how? experimental design for causal models at scale. *Advances in Neural Information Processing Systems*, 35, 2022. 1, 2, 4, 8, 9

Tong, S. and Koller, D. Active learning for structure in bayesian networks. In *International joint conference on artificial intelligence*, volume 17, pp. 863–869. Citeseer, 2001. 1, 2, 5

Toth, C., Lorch, L., Knoll, C., Krause, A., Pernkopf, F., Peharz, R., and von Kügelgen, J. Active bayesian causal inference. *arXiv preprint arXiv:2206.02063*, 2022. 1, 2, 8, 9

von Kügelgen, J., Rubenstein, P. K., Schölkopf, B., and Weller, A. Optimal experimental design via bayesian optimization: active causal structure learning for gaussian process networks. *arXiv preprint arXiv:1910.03962*, 2019. 2

Xie, S. M. and Ermon, S. Reparameterizable subset sampling via continuous relaxations. *arXiv preprint arXiv:1901.10517*, 2019. 6

A. Derivation of Importance Weighted Nested Monte Carlo Estimator

In this section, we derive the $\mathcal{U}_{\text{IWNMC}}$ (Eq. 4) estimator. We derive the estimator for a single design with an experiment denoted by ξ, parameters $\boldsymbol{\theta}$ and experimental outcome random variable $\mathbf{Y}$ and its instance $\mathbf{y}$. Since it is a static design, all the steps of the derivation hold if we replace ξ with $\xi_{1:B}$, $\mathbf{Y}$ with $\mathbf{Y}_{1:B}$ and $\mathbf{y}$ with $\mathbf{y}_{1:B}$. We begin from the variational NMC (VNMC) estimator, introduced by Foster et al. (2019)

$$\mathcal{I}(\mathbf{Y};\boldsymbol{\Theta} \mid \xi) \le \mathcal{U}_{\text{VNMC}}(\xi) = \mathop{\mathbb{E}}_{\substack{p(\boldsymbol{\theta}_0|h_{t-1}) \\ p(\mathbf{y}|\boldsymbol{\theta}_0,\xi) \\ q(\boldsymbol{\theta}_{1:L}|h_{t-1},\mathbf{y})}} \left[\log \frac{p(\mathbf{y} \mid \xi,\boldsymbol{\theta}_0)}{\frac{1}{L}\sum_{\ell=1}^{L} \frac{p(\mathbf{y}|\xi,\boldsymbol{\theta}_\ell)p(\boldsymbol{\theta}_\ell|h_{t-1})}{q(\boldsymbol{\theta}_{1:L}|h_{t-1},\mathbf{y})}} \right]. \tag{5}$$

This can be rewritten as

$$\mathcal{U}_{\text{VNMC}}(\xi) = \mathop{\mathbb{E}}_{\substack{p(\boldsymbol{\theta}_0|h_{t-1}) \\ p(\mathbf{y}|\boldsymbol{\theta}_0,\xi) \\ q(\boldsymbol{\theta}_{1:L}|h_{t-1},\mathbf{y})}} \left[\log \frac{p(\mathbf{y} \mid \xi,\boldsymbol{\theta}_0)}{\frac{1}{L}\sum_{\ell=1}^{L} \frac{p(\mathbf{y}|\xi,\boldsymbol{\theta}_\ell)p(\boldsymbol{\theta}_\ell)p(h_{t-1}|\boldsymbol{\theta}_\ell)}{q(\boldsymbol{\theta}_{1:L}|h_{t-1},\mathbf{y})}} + \log p(h_{t-1}) \right] \tag{6}$$

and Foster et al. (2019) observed that $\log p(h_{t-1})$ is a constant that does not depend on ξ and so can be safely neglected when optimizing over designs. If we take the original prior $p(\boldsymbol{\theta}_\ell)$ as our proposal distribution q, then we arrive at

$$\mathcal{U}_{\text{VNMC-prior}}(\xi) = \mathop{\mathbb{E}}_{\substack{p(\boldsymbol{\theta}_0|h_{t-1})p(\boldsymbol{\theta}_{1:L}) \\ p(\mathbf{y}|\boldsymbol{\theta}_0,\xi)}} \left[\log \frac{p(\mathbf{y} \mid \xi,\boldsymbol{\theta}_0)}{\frac{1}{L}\sum_{\ell=1}^{L} p(\mathbf{y} \mid \xi,\boldsymbol{\theta}_\ell)p(h_{t-1} \mid \boldsymbol{\theta}_\ell)} \right] + C \tag{7}$$

where $C = \log p(h_{t-1})$. This allows us to sample contrastive samples from any distribution, but does not account for $\boldsymbol{\theta}_0$. If we were to sample $\boldsymbol{\theta}_0$ from $p(\boldsymbol{\theta}_0)$, we can correct using an importance weight

$$\mathcal{U}_{\text{VNMC-prior}}(\xi) = \mathop{\mathbb{E}}_{\substack{p(\boldsymbol{\theta}_{0:L}) \\ p(\mathbf{y}|\boldsymbol{\theta}_0,\xi)}} \left[\frac{p(\boldsymbol{\theta}_0 \mid h_{t-1})}{p(\boldsymbol{\theta}_0)} \log \frac{p(\mathbf{y} \mid \xi,\boldsymbol{\theta}_0)}{\frac{1}{L}\sum_{\ell=1}^{L} p(\mathbf{y} \mid \xi,\boldsymbol{\theta}_\ell)p(h_{t-1} \mid \boldsymbol{\theta}_\ell)} \right] + C, \tag{8}$$

but unfortunately, this relies on knowing the density of the posterior or using the fact that $p(\boldsymbol{\theta}_0 \mid h_{t-1})/p(\boldsymbol{\theta}_0) = p(h_{t-1} \mid \boldsymbol{\theta}_0)/p(h_{t-1})$, knowing the marginal likelihood of the data h_{t-1}. Neither of these is usually tractable. Instead, we can use a self-normalized importance sampling approach, which amounts to estimating $p(h_{t-1})$ by a sum over $\boldsymbol{\theta}_{0:L}$, giving the *approximation* IWNMC:

$$\mathcal{U}_{\text{IWNMC}}(\xi) = \mathop{\mathbb{E}}_{\substack{p(\boldsymbol{\theta}_{0:L}) \\ p(\mathbf{y}|\boldsymbol{\theta}_0,\xi)}} \left[\frac{p(h_{t-1} \mid \boldsymbol{\theta}_0)}{\frac{1}{L+1}\sum_{k=0}^{L} p(h_{t-1} \mid \boldsymbol{\theta}_k)} \log \frac{p(\mathbf{y} \mid \xi,\boldsymbol{\theta}_0)}{\frac{1}{L}\sum_{\ell=1}^{L} p(\mathbf{y} \mid \xi,\boldsymbol{\theta}_\ell)p(h_{t-1} \mid \boldsymbol{\theta}_\ell)} \right] + C. \tag{9}$$

The form that is given in (4) is obtained by first relabelling the $\boldsymbol{\theta}$ samples to start from 1

$$\mathcal{U}_{\text{IWNMC}}(\xi) = \mathop{\mathbb{E}}_{\substack{p(\boldsymbol{\theta}_{1:L}) \\ p(\mathbf{y}|\boldsymbol{\theta}_1,\xi)}} \left[\frac{p(h_{t-1} \mid \boldsymbol{\theta}_1)}{\frac{1}{L}\sum_{k=1}^{L} p(h_{t-1} \mid \boldsymbol{\theta}_k)} \log \frac{p(\mathbf{y} \mid \xi,\boldsymbol{\theta}_1)}{\frac{1}{L-1}\sum_{\ell=2}^{L} p(\mathbf{y} \mid \xi,\boldsymbol{\theta}_\ell)p(h_{t-1} \mid \boldsymbol{\theta}_\ell)} \right] + C, \tag{10}$$

noting that the role of $\boldsymbol{\theta}_1$ is arbitrary and can be replaced by any $m \in \{1,\ldots,L\}$

$$\mathcal{U}_{\text{IWNMC}}(\xi) = \mathop{\mathbb{E}}_{\substack{p(\boldsymbol{\theta}_{1:L}) \\ p(\mathbf{y}|\boldsymbol{\theta}_m,\xi)}} \left[\frac{p(h_{t-1} \mid \boldsymbol{\theta}_m)}{\frac{1}{L}\sum_{k=1}^{L} p(h_{t-1} \mid \boldsymbol{\theta}_k)} \log \frac{p(\mathbf{y} \mid \xi,\boldsymbol{\theta}_m)}{\frac{1}{L-1}\sum_{\ell\neq m}^{L} p(\mathbf{y} \mid \xi,\boldsymbol{\theta}_\ell)p(h_{t-1} \mid \boldsymbol{\theta}_\ell)} \right] + C, \tag{11}$$
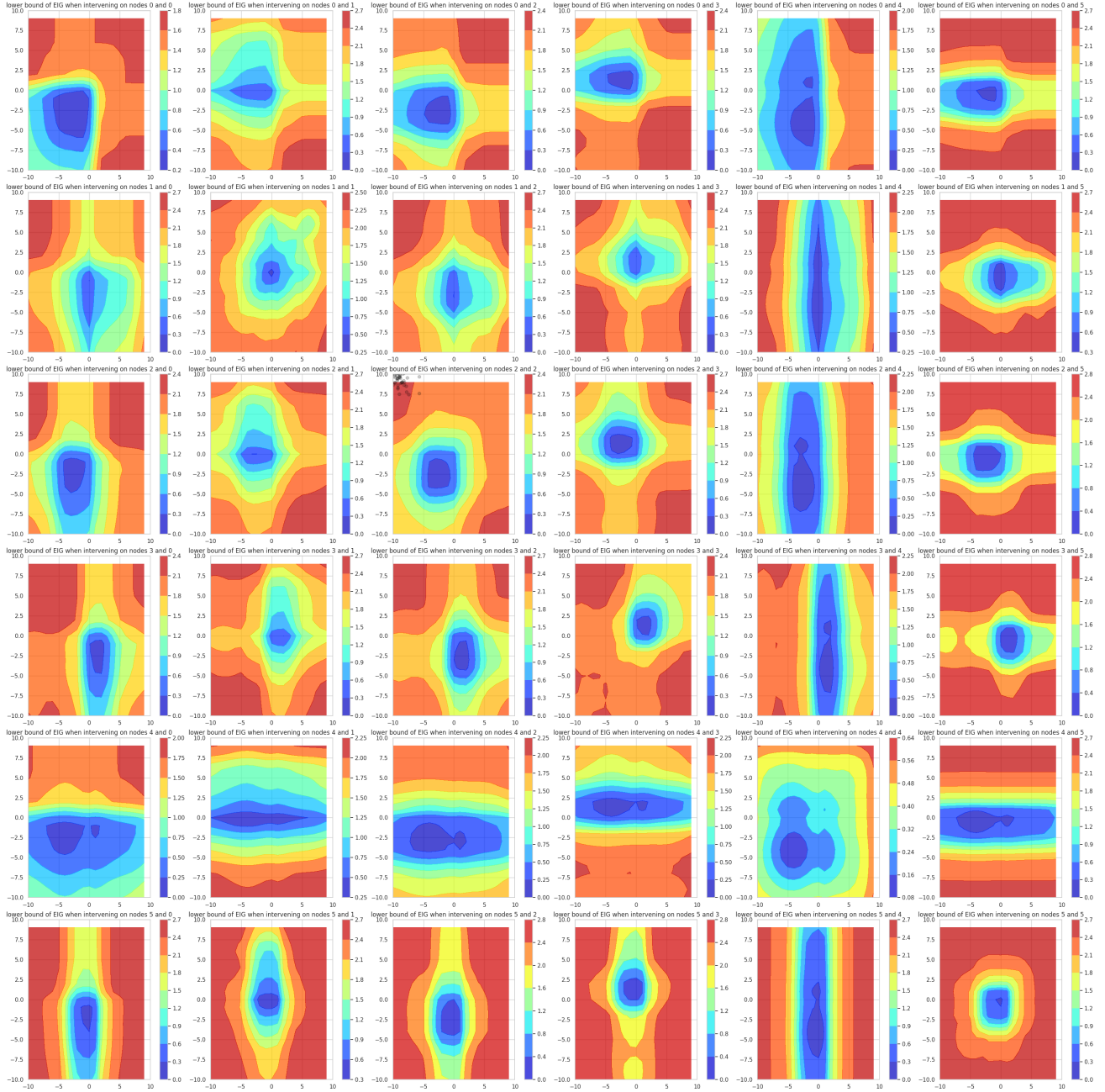
and finally taking the mean over m, noting that this does not change the expected value due to linearity

$$\mathcal{U}_{\text{IWNMC}}(\xi) = \mathop{\mathbb{E}}_{\substack{p(\boldsymbol{\theta}_{1:L}) \\ p(\mathbf{y}|\boldsymbol{\theta}_m,\xi)}} \left[\sum_{m=1}^{L} \frac{p(h_{t-1} \mid \boldsymbol{\theta}_m)}{\sum_{k=1}^{L} p(h_{t-1} \mid \boldsymbol{\theta}_k)} \log \frac{p(\mathbf{y} \mid \xi,\boldsymbol{\theta}_m)}{\frac{1}{L-1}\sum_{\ell\neq m}^{L} p(\mathbf{y} \mid \xi,\boldsymbol{\theta}_\ell)p(h_{t-1} \mid \boldsymbol{\theta}_\ell)} \right] + C. \tag{12}$$

We finally drop the constant C as it is independent of ξ and take

$$\omega_m = \frac{p(h_{t-1} \mid \boldsymbol{\theta}_m)}{\sum_{k=1}^{L} p(h_{t-1} \mid \boldsymbol{\theta}_k)}. \tag{13}$$

B. Expected Information Gain for 6 Variables and Batch Size 2

Figure 7: Here we visualize the Expected Information Gain of batch size two, on two nodes over different interventional states of the range $[-10, 10]$.

C. Metrics

.

$\mathbb{E}$**-SHD**: Defined as the *expected structural hamming distance* between samples from the posterior model over graphs and the true graph $\mathbb{E}\text{-SHD} := \mathbb{E}_{\mathbf{g}\sim p(\mathcal{G}|\mathcal{D})}\big[\text{SHD}(\mathbf{g}, \tilde{\mathbf{g}})\big]$

Expected edges F1: The expected F1 score of the binary classification task of predicting the presence/ absence of all edges.

The expectation is taken over multiple posterior samples.

i-MMD: Interventional MMD is defined as MMD distance (Gretton et al., 2012) between the true interventional distribution and the interventional distribution induced by θ and $\mathbf{g}$ (posterior sample). We take an expectation over different posterior samples, interventional targets and interventional states. For the kernel choice, we use the median heuristic as described in (Gretton et al., 2012).

D. DAG Bootstrap

The DAG bootstrap bootstraps observations and interventions to infer a different causal structure per bootstrap. We used GIES as the causal inference algorithm because of the adaptation of GES on interventional data as well. In our experiments, we used the pcalg R implementation `https://github.com/cran/pcalg/blob/master/R/gies.R` to discover 100 graphs. Each graph can be seen as a posterior sample from $p(\mathbf{G} \mid h_{t-1})$. For each of the sampled graphs G_i we compute the appropriate θ_{MLE} under linear Gaussian assumption for the conditional distributions.
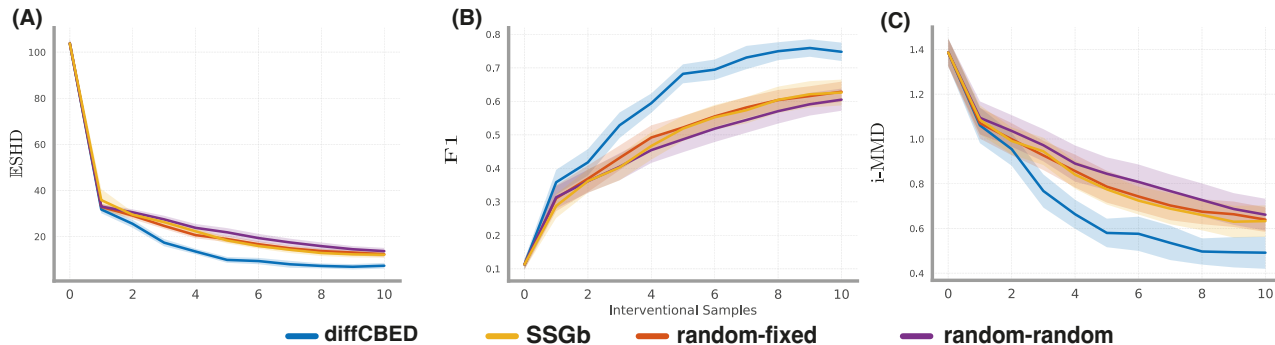
E. Importance Weighted Nested Monte Carlo Full Results
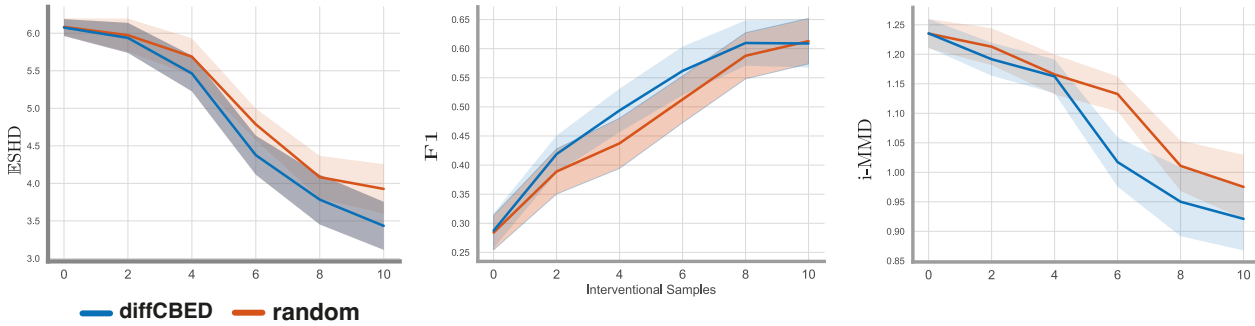
Figure 8: Multi target-state design setting results for Erdős–Rényi (Erdős & Rényi, 1959) graphs with $d = 5$ variables. Each experiment was run with 60 random seeds (shaded area represents 95% CIs)
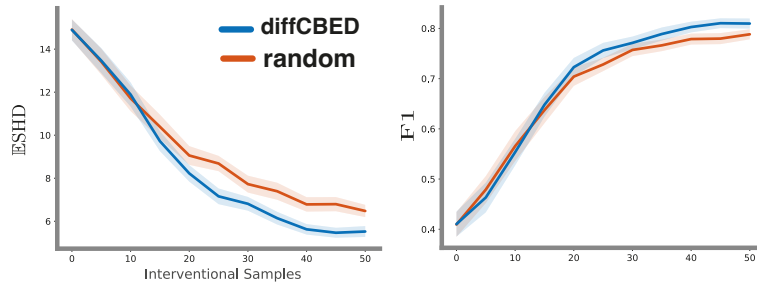
Figure 9: Single target-state design setting results for Erdős–Rényi (Erdős & Rényi, 1959) graphs with $d = 10$ variables. Each experiment was run with 60 random seeds (shaded area represents 95% CIs)

F. 20 **nodes, unconstrained** $(q \leq 20)$, **batch size** $B = 1$:

Figure 10: Multi target-state design setting results for Erdős–Rényi (Erdős & Rényi, 1959) graphs with $d = 50$ variables. Each experiment was run with 30 random seeds (the shaded area represents 95% CIs). We observe that for batch size 1, the difference between the methods becomes more significant.

G. Datasets and Experiment Details

G.1. Synthetic Graphs Experiments

In the synthetic data experiments, we focus on Erdős-Rényi graph model. We used `networkx`[4] and method `fast_gnp_random_graph` (Batagelj & Brandes, 2005) to generate graphs based on the Erdős-Rényi model. We set the expected number of edges per vertex to 1.

H. Optimizer Settings

**Table 3:** Table indicating the hyperparameters and optimizer settings for different experimental results.

	Optimization settings			
	Single Target NMC	Multi-Target NMC	Multi-Target IWNMC with prior	Multi-Target IWNMC with proposal
L	30	30	1000	60
Number of outer DAGs N_o	30	30	1000	60
Batch Size	5	2	2	2
Relaxation temperature	$5 \rightarrow .5$	$5 \rightarrow .5$	0.1	$5 \rightarrow .5$
Optimizer	Adam	Adam	Adam	Adam
Learning rate of optimizer	0.1	0.1	0.01	0.1
Number of starting samples (observational)	60	60	2	800
Number of batches	10	10	5	1
Number of DAG Bootstraps	30	30	-	60
Number of training steps per batch	100	100	100	100

I. Relaxed Distribution Temperature Sensitivity Analysis

We perform ablation where we empirically test the sensitivity to temperature hyperparameter for the relaxed distribution. Results are presented in Figure 11 for the unconstrained multi-target setting ($d = 20$). We find that our approach is fairly robust to different temperatures. Note that for the reported results, we anneal the temperature during the course of training and hence a single choice of temperature is not necessary.

---

[4]https://networkx.org/documentation/networkx-1.10/reference/generated/networkx.generators.random_graphs.fast_gnp_random_graph.html
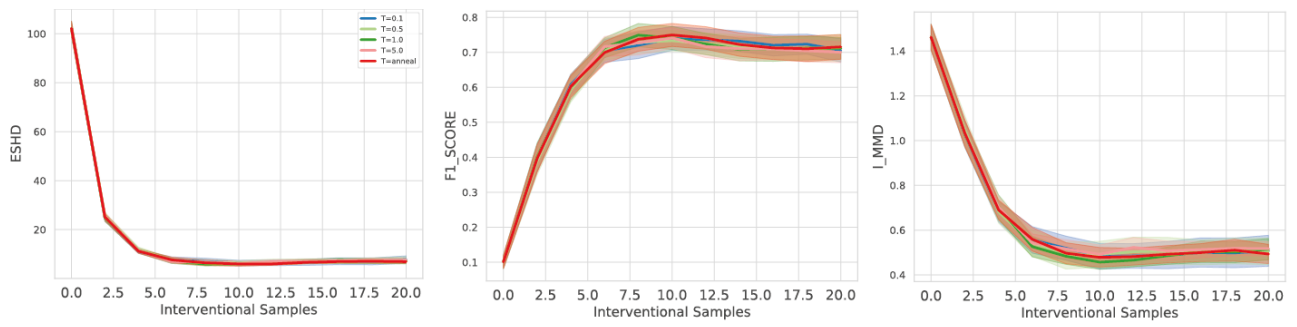
Figure 11: Performance of our proposed design strategy for different temperatures of the relaxed distribution for the unconstrained multi-target setting $(d = 20)$. We find that our method is fairly robust.