

---

# Causal Bayesian Optimization

---

**Virginia Aglietti**  
University of Warwick  
The Alan Turing Institute  
V.Aglietti@warwick.ac.uk

**Xiaoyu Lu**  
Amazon  
Cambridge, UK  
luxiaoyu@amazon.com

**Andrei Paleyes**  
Amazon  
Cambridge, UK  
paleyes@amazon.com

**Javier González**  
Amazon  
Cambridge, UK  
gojav@amazon.com

## Abstract

This paper studies the problem of globally optimizing a variable of interest that is part of a causal model in which a sequence of interventions can be performed. This problem arises in biology, operational research, communications and, more generally, in all fields where the goal is to optimize an output metric of a system of interconnected nodes. Our approach combines ideas from causal inference, uncertainty quantification and sequential decision making. In particular, it generalizes Bayesian optimization, which treats the input variables of the objective function as independent, to scenarios where causal information is available. We show how knowing the causal graph significantly improves the ability to reason about optimal decision making strategies decreasing the optimization cost while avoiding suboptimal solutions. We propose a new algorithm called Causal Bayesian Optimization (CBO). CBO automatically balances two trade-offs: the classical exploration-exploitation and the new observation-intervention, which emerges when combining real interventional data with the estimated intervention effects computed via *do*-calculus. We demonstrate the practical benefits of this method in a synthetic setting and in two real-world applications.

## 1 Introduction

Decision making problems in a variety of domains, such as biological systems, modern industrial processes or social systems, require implementing interventions and

manipulating variables in order to optimize an outcome of interest. For instance, in strategic planning, companies need to decide how to allocate scarce resources across different projects or business units in order to achieve performance goals. In biology, it is common to change the phenotype of organisms by acting on individual components of complex gene networks. This paper describes how to find such interventions or policies.

Focusing on a specific example, consider a setting in which  $\mathbf{Y}$  denotes the crop yields for different agricultural products,  $X$  denotes soil fumigants and  $\mathbf{Z} = \{Z_1, Z_3, Z_4\}$  represents the eel-worm population at different times (Cochran and Cox, 1957). Given a causal graph (Pearl, 1995) representing the investigator’s understanding of the major causal influences among the variables (Fig. 1, left), she aims at finding the highest yielding intervention in a limited number of seasons and subject to a budget constraint. Each intervention has a cost which is determined by the manipulated variables’ costs and the implemented intervention levels.

In order to solve this problem, the investigator could resort to Bayesian Optimization (BO). BO is an efficient heuristic to optimize objective functions whose evaluations are costly and when no explicit functional form is available (Jones et al., 1998). In the example above, BO would try to find the global optima by making a series of function evaluations in which all variables are manipulated. BO would thus break the dependency structure existing among  $X$  and  $\mathbf{Z}$ , potentially leading to suboptimal solutions. Indeed, as described later in detail, depending on the structural relationships between variables, intervening on a subgroup might lead to a propagation of effects in the causal graph and a higher final yield. In addition, intervening on all variables is cost-ineffective in cases when the same yield can be obtained by setting only a subgroup of them.

### 1.1 Approach and Contributions

The framework proposed in this work combines BO and causal inference, offering a novel approach for decision making under uncertainty. Probabilistic causal

models are commonly used in disciplines where explicit experimentation may be difficult such as social science or economics. In particular, the *do*-calculus (Pearl, 1995) relates observational distributions to interventional ones. It allows to predict the effect of an intervention without explicitly performing it and by solely using observational data. We develop a model which integrates observational and interventional data so as to further reduce the uncertainty around the optimal value and the number of interventions required to find it. Particularly, we make the following contributions:

- We formulate a new class of optimization problems called *Causal Global Optimization* (CGO) where the causal structure existing among the input variables is accounted for in the objective functions.
- We solve CGO problems by combining ideas from BO and causal calculus. We propose a Gaussian process (GP) surrogate model, the causal GP, that integrates observational and interventional data via the definition of a causal prior distribution computed through *do*-calculus.
- We propose an acquisition function, the causal expected improvement (EI), which drives the exploration of different interventions.
- We develop an algorithm, henceforth named *Causal Bayesian Optimization* (CBO), that exploits the structural characteristics of the graph, the causal EI and the proposed GP prior to find an optimal intervention. In doing that, it balances the emerging trade-off between observation and intervention via an  $\epsilon$ -greedy policy.
- We show the benefits of the proposed approach in a variety of experimental settings featuring different dependency structures, unobserved confounders and non-manipulative variables.

## 1.2 Related Work

While there exists an extensive literature on BO (see Shahriari et al. (2015) for a review) and causality (Guo et al., 2018), the literature on causal decision making is limited. Recent works have focused on multi-armed bandit (MAB) problems and reinforcement learning (RL) settings where actions or arms correspond to interventions on an arbitrary causal graph and there exists complex links between the agent’s decisions and the received rewards. Bareinboim et al. (2015) and Lu et al. (2018) focus on settings with unobserved confounders. Lee and Bareinboim (2018) identify a set of possibly-optimal arms that an agent should play in order to maximize its expected reward in a MAB problem. Lee and Bareinboim (2019) extend this work to graphs with non-manipulable variables. Lattimore

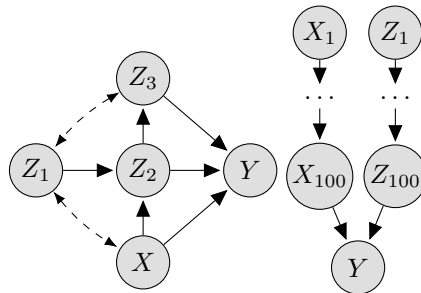


Figure 1: Examples of causal graphs. Nodes denote variables, arrows represent causal effects and dashed edges indicate unobserved confounders. *Left*: Yield optimization example.  $Y$  is the crop yield,  $X$  denotes soil fumigants and  $\mathbf{Z}$  represents the eel-worm population. *Right*: A 200-dimensional optimization problem with causal intrinsic dimensionality equal to 2.

et al. (2016) study a specific family of MAB problems called parallel bandit problems. Finally, Ortega and Braun (2014) focus on causal discovery in Causal MAB. In the RL literature, Buesing et al. (2018) leverage structural causal models for counterfactual evaluation of arbitrary policies on individual off-policy episodes. Foerster et al. (2018) focus on the multi-agents setting and propose a framework in which each agent learns from a shaped reward that compares the global reward to the counterfactual reward received when that agent’s action is replaced with a default action.

## 2 Background and Problem Statement

In this paper, random variables and observations are denoted in upper case and lower case respectively. Vectors are represented in bold.  $do(\mathbf{X} = \mathbf{x})$  represents an intervention on  $\mathbf{X}$  whose value is set to  $\mathbf{x}$ .  $P(\mathbf{Y}|\mathbf{X} = \mathbf{x})$  represents an *observational distribution* and  $P(\mathbf{Y}|do(\mathbf{X} = \mathbf{x}))$  represents an *interventional distribution*.  $\mathcal{D}^O$  and  $\mathcal{D}^I$  denote observational and interventional datasets respectively.

### 2.1 Structural Equation Models (SEM) and *do*-calculus

Consider a probabilistic causal model (Pearl, 2000) consisting of a directed acyclic graph  $\mathcal{G}$  (DAG) and a four-tuple  $\langle \mathbf{U}, \mathbf{V}, F, P(\mathbf{U}) \rangle$ , where  $\mathbf{U}$  is a set of independent *exogenous* background variables distributed according to the probability distribution  $P(\mathbf{U})$ ,  $\mathbf{V}$  is a set of observed *endogenous* variables and  $F = \{f_1, \dots, f_{|\mathbf{V}|}\}$  is a set of functions such that  $v_i = f_i(pa_i, u_i)$  with  $pa_i$  denoting the parents of  $V_i$ .  $\mathcal{G}$  encodes our knowledge of the existing causal mechanisms among  $\mathbf{V}$ . Within  $\mathbf{V}$ , we distinguish between three different types of variables: non-manipulative variables  $\mathbf{C}$ , which cannot be modified, treatment variables  $\mathbf{X}$  that can be set

to specific values and output variables  $\mathbf{Y}$  that represents the agent’s outcomes of interest. Given the conditional independence relationships encoded in  $\mathcal{G}$ , we can exploit the causal Markov condition (Pearl, 2000) to write the joint observational distribution as  $P(\mathbf{V}) = \prod_{i=1}^{|\mathbf{V}|} p(V_i|pa_i)$ . The interventional distribution for two disjoint sets in  $\mathbf{V}$ , say  $\mathbf{X}$  and  $\mathbf{Y}$ , is  $P(\mathbf{Y}|\text{do}(\mathbf{X} = \mathbf{x}))$ . This is the distribution of  $\mathbf{Y}$  obtained by intervening on  $\mathbf{X}$  and fixing its value to  $\mathbf{x}$  in the data generating mechanism, irrespective of the values of its parents and keeping  $\mathbf{C}$  unchanged. Note the difference between  $P(\mathbf{Y}|\text{do}(\mathbf{X} = \mathbf{x}))$  that requires “real” interventions and  $P(\mathbf{Y}|\mathbf{X} = \mathbf{x})$  that only requires “observing” the system. In this work we assume  $\mathcal{G}$  to be known. Causal discovery (Glymour et al., 2019) is a complex topic and analysing what happens when the graph is unknown goes beyond the scope of this paper. We leave this for future work.

*Do*-calculus offers a powerful tool to estimate interventional distributions and causal effects from observational distributions. If the causal effects are identifiable, we can apply the three rules of *do*-calculus to link interventional distributions with observational distributions which can be approximated with e.g. Monte Carlo estimates. The first formula, called “back-door adjustment”, applies to cases in which we have a set of observed confounders between cause and effect. For instance, in Fig. 1 (left),  $p(y|\text{do}(Z_2 = z_2))$  can be computed by adjusting for  $Z_1$  and  $X$ :

$$p(y|\text{do}(Z_2 = z_2)) = \int_{x, z_1} p(y|Z_2 = z_2, z_1, x)p(x, z_1)dx dz_1.$$

The second formula, called “front-door adjustment” applies to cases with unobserved confounders. When computing  $p(y|\text{do}(X = x))$  in Fig. 1 (left), the adjustment is given by:

$$p(y|\text{do}(X = x)) = \int_{z_1, z_2, z_3} p(y|z_2, z_3, X = x)p(z_2|z_1, X = x) \\ \times \int_{x'} p(z_3|z_1, z_2, x')p(z_1, x')dx' dz_1 dz_2 dz_3.$$

When the above integrals are not tractable, observational data can be used to get a Monte Carlo estimate  $\hat{P}(\mathbf{Y}|\text{do}(\mathbf{X} = \mathbf{x})) \approx P(\mathbf{Y}|\text{do}(\mathbf{X} = \mathbf{x}))$ , which is consistent when the number of samples drawn from  $P(\mathbf{V})$  is sufficiently large.

## 2.2 Problem Setup

A novel class of global optimization problems called *Causal Global Optimization* (CGO) is introduced in this section. Given  $\mathcal{G}$  and  $\langle \mathbf{U}, \mathbf{V}, F, P(\mathbf{U}) \rangle$ , the goal is to select the set of intervention variables  $\mathbf{X}_s^*$  and intervention levels  $\mathbf{x}_s^*$  that optimize the expected target

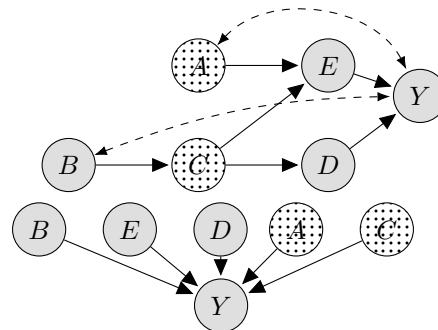


Figure 2: DAG representation of a CGO problem (top) and the DAG considered when using BO (bottom) to address the same problem. Shaded nodes represent  $\mathbf{X}$  while dotted nodes give  $\mathbf{C}$ . Dashed edges indicate unobserved confounders.

outcomes  $\mathbf{Y}$ . Formally, the goal is to find:

$$\mathbf{X}_s^*, \mathbf{x}_s^* = \arg \min_{\mathbf{X}_s \in \mathcal{P}(\mathbf{X}), \mathbf{x}_s \in D(\mathbf{X}_s)} \mathbb{E}_{P(\mathbf{Y}|\text{do}(\mathbf{X}_s = \mathbf{x}_s))}[\mathbf{Y}], \quad (1)$$

where  $\mathcal{P}(\mathbf{X})$  is the power set of  $\mathbf{X}$  and  $D(\mathbf{X}_s) = \times_{X \in \mathbf{X}_s} (D(\mathbf{X}))$  with  $D(\mathbf{X})$  denoting the interventional domain of  $\mathbf{X}$  and the expectation is computed according to the interventional distribution. For notational convenience, we denote  $\mathbb{E}_{P(\mathbf{Y}|\text{do}(\mathbf{X}_s = \mathbf{x}_s))}[\mathbf{Y}] \doteq \mathbb{E}[\mathbf{Y}|\text{do}(\mathbf{X}_s = \mathbf{x}_s)]$  and  $\mathbb{E}_{\hat{P}(\mathbf{Y}|\text{do}(\mathbf{X}_s = \mathbf{x}_s))}[\mathbf{Y}] \doteq \hat{\mathbb{E}}[\mathbf{Y}|\text{do}(\mathbf{X}_s = \mathbf{x}_s)]$ . The optimal subset of intervention variables  $\mathbf{X}_s$  belongs to  $\mathcal{P}(\mathbf{X})$  which includes the empty set  $\emptyset$  and  $\mathbf{X}$  itself. When  $\mathbf{X}_s = \emptyset$ , no intervention is implemented in the system and the target expected values correspond to the observational expected outcomes. When  $\mathbf{X}_s = \mathbf{X}$ , all variables are intervened upon except for the context variables  $\mathbf{C}$  that can only be observed.

The problem given in Eq. (1) is challenging because of two reasons. Firstly, the cardinality of  $\mathcal{P}(\mathbf{X})$  grows exponentially with  $|\mathbf{X}|$  and finding the optimal set requires, in principle, a combinatorial search. Secondly, for every set  $\mathbf{X}_s$ , finding  $\mathbf{x}_s^*$  requires evaluating the objective function and thus implementing interventions in the system several times. In most settings, the number of function evaluations, whose cost is assumed to be given by some cost function  $Co(\mathbf{X}_s, \mathbf{x}_s)$ , needs to be kept low. We thus want to find the optimal configuration with the minimal cost,  $\sum_{i=1}^T Co(\mathbf{X}_s, \mathbf{x}_i)$  for a sequence of interventions  $1, \dots, T$ .

## 2.3 Connections and Generalisations

**Bayesian Optimization:** Consider the DAG in Fig. 2 (top). The problem in Eq. (1) can be solved through the BO method which breaks the input variables dependencies (Fig. 2 bottom) and intervenes simultaneously on all of them thus setting  $\mathbf{X}_s = \mathbf{X}$ . To this aim, BO considers a surrogate probabilistic model for the

objective function  $\mathbf{Y} = f(\mathbf{X})$  which is given by a GP  $p(f) = \mathcal{GP}(m, k)$  with mean function  $m$  and covariance function  $k$ . Given a dataset  $\mathcal{D}_n = \{\mathbf{x}_i, y_i\}_{i=1}^n$ , the posterior distribution of  $f$  under Gaussian likelihood is also a GP with posterior mean and variance given by  $m_n(\mathbf{x}) = \mathbf{k}_n(\mathbf{x})^T [\mathbf{K}_n + \sigma_n^2 \mathbf{I}]^{-1} \mathbf{y}_n$  and  $\sigma_n^2(\mathbf{x}) = k(\mathbf{x}, \mathbf{x}) - \mathbf{k}_n(\mathbf{x})^T [\mathbf{K}_n + \sigma_n^2 \mathbf{I}]^{-1} \mathbf{k}_n(\mathbf{x})$  where  $\mathbf{K}_n$  is the matrix such that  $(\mathbf{K}_n)_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$ ,  $\mathbf{k}_n(\mathbf{x}) = [k(x_1, \mathbf{x}), \dots, k(x_n, \mathbf{x})]^T$  (Rasmussen, 2003) and  $\mathbf{x}$  is the point where the GP is evaluated. This posterior is used to form the acquisition function  $\alpha(\mathbf{x}, \mathcal{D}_n)$ , which can be chosen among popular acquisition functions such as expected improvement or entropy search. The next evaluation is placed at the (numerically estimated) global maximum of  $\alpha(\mathbf{x}, \mathcal{D}_n)$ .

**Causal Dimensionality:** It is well known (Wang et al., 2016) that the performance of standard BO algorithms deteriorates in high dimensional problems as the number of evaluations needed to find the global optimum increases exponentially with the space dimensionality. Interestingly, knowing the causal graph allows to reason about the effective dimensionality of the problem. We formalize this idea by defining the notion of *causal intrinsic dimensionality*:

**Definition 2.1.** The causal intrinsic dimensionality of a causal function  $\mathbb{E}_{P(Y|\text{do}(\mathbf{X}=\mathbf{x}))}[Y]$  is given by the number of parents of the target variable, that is  $|Pa(Y)|$ .

In Fig. 1 (right), the input space dimensionality is 200. However,  $\mathbb{E}[Y|\text{do}(X_1, \dots, X_{100}, Z_1, \dots, Z_{100})] = \mathbb{E}[Y|\text{do}(X_{100}, Z_{100})]$ , thus  $X_{100}$  and  $Z_{100}$  are the only 2 relevant variables and the intrinsic dimensionality of the problem is 2. For the general problem in Eq. (1) we have  $\mathbb{E}[\mathbf{Y}|\text{do}(\mathbf{X})] = \mathbb{E}[\mathbf{Y}|\text{do}(Pa(\mathbf{Y}))]$ .

Related to the concept of causal dimensionality, Wang et al. (2016) proposed to perform Bayesian optimization in a low-dimensional space which reflects the *intrinsic dimensionality* of a function. Provided that the objective function has low intrinsic dimensionality, Wang et al. (2016) use random embeddings to reduce the problem dimensionality without knowing which dimensions are important. This idea can be formalized and made explicit by taking a causal perspective on the optimization problem. The causal graph allows to determine not only if the function has low intrinsic dimensionality but also to identify which dimensions are important.

**Causal Bandits:** There is a significant link between our problem setup and the settings tackled by causal MAB algorithms. Causal MAB algorithms interpret decisions as interventions, target a causal effect function and account for complex dependency structure between actions which are encoded in the causal graph. Indeed, when all intervention variables  $\mathbf{X}$  are binary, the CGO

setting reduces to the causal MAB setting. However, Eq. (1) gives a more general formulation of the problem where variables can be continuous or categorical and, more importantly, where the intervention values need to be determined together with the intervention set.

### 3 Methodology

This section details a new methodology, which we call *Causal Bayesian Optimization*, addressing the problem in Eq. (1). The elements of this approach are: (i) an exploration set (§3.1) determining a set of variables which is worth intervening on based on the topology of  $\mathcal{G}$ , (ii) a surrogate model (§3.2), called Causal GP model, that enables the integration of observational and interventional data, (iii) an acquisition function (§3.3) solving the exploration/exploitation trade off *across* interventions, (iv) an  $\epsilon$ -greedy policy (§3.4) solving the observation/intervention trade-off. This paper considers settings where a data set  $\mathcal{D}^O = \{(\mathbf{v}^n, y^n)\}_{n=1}^N$  from an observational study is available. Here  $\mathbf{v}^n \in \mathbb{R}^{|\mathbf{V}|}$ ,  $y^n \in \mathbb{R}$  and the joint distribution follows the conditional independence assumptions encoded in  $\mathcal{G}$ . In the following discussion we consider a single output and leave the integration of our framework with multi-output schemes as future work.

#### 3.1 Selecting the Optimal Exploration Set

A naive approach to find  $\mathbf{X}_s^*$  would be to explore the  $2^{|\mathbf{X}|}$  sets in  $\mathcal{P}(\mathbf{X})$ . Albeit this is a valid strategy, its complexity grows exponentially with  $|\mathbf{X}|$ . However, exploiting the rules of *do*-calculus and the partial orders among subsets, Lee and Bareinboim (2018) identify invariances in the interventional space and potentially optimal intervention set which we define below.

**Definition 3.1. Minimal Intervention set (MIS).** Given  $\langle \mathcal{G}, \mathbf{Y}, \mathbf{X}, \mathbf{C} \rangle$ , a set of variables  $\mathbf{X}_s \in \mathcal{P}(\mathbf{X})$  is said to be a MIS if there is no  $\mathbf{X}'_s \subset \mathbf{X}_s$  such that  $\mathbb{E}[Y|\text{do}(\mathbf{X}_s = \mathbf{x}_s)] = \mathbb{E}[Y|\text{do}(\mathbf{X}'_s = \mathbf{x}'_s)]$ .

We denote by  $\mathbb{M}_{\mathcal{G}, \mathbf{Y}}^{\mathbf{C}}$  the set of MISs for  $\langle \mathcal{G}, \mathbf{Y}, \mathbf{X}, \mathbf{C} \rangle$  where each MIS represents a set of variables that is worth intervening on. When  $\mathbf{C} = \emptyset$ , we use  $\mathbb{M}_{\mathcal{G}, \mathbf{Y}}$ . Incorporating into MIS the partial orderedness among subsets of  $\mathcal{P}(\mathbf{X})$  we define the so-called POMIS.

**Definition 3.2. Possibly-Optimal Minimal Intervention set (POMIS).** Given  $\langle \mathcal{G}, \mathbf{Y}, \mathbf{X}, \mathbf{C} \rangle$ , let  $\mathbf{X}_s \in \mathbb{M}_{\mathcal{G}, \mathbf{Y}}^{\mathbf{C}}$ .  $\mathbf{X}_s$  is a POMIS if there exists a SEM conforming to  $\mathcal{G}$  such that  $\mathbb{E}[Y|\text{do}(\mathbf{X}_s = \mathbf{x}^*)] > \forall \mathbf{w} \in \mathbb{M}_{\mathcal{G}, \mathbf{Y}}^{\mathbf{C}} \setminus \mathbf{X}_s \mathbb{E}[Y|\text{do}(\mathbf{W} = \mathbf{w}^*)]$  where  $\mathbf{x}^*$  and  $\mathbf{w}^*$  denote the optimal intervention values.

We denote by  $\mathbb{P}_{\mathcal{G}, \mathbf{Y}}^{\mathbf{C}}$  the set of POMIS for  $\langle \mathcal{G}, \mathbf{Y}, \mathbf{X}, \mathbf{C} \rangle$  where each POMIS represents a variable on which inter-

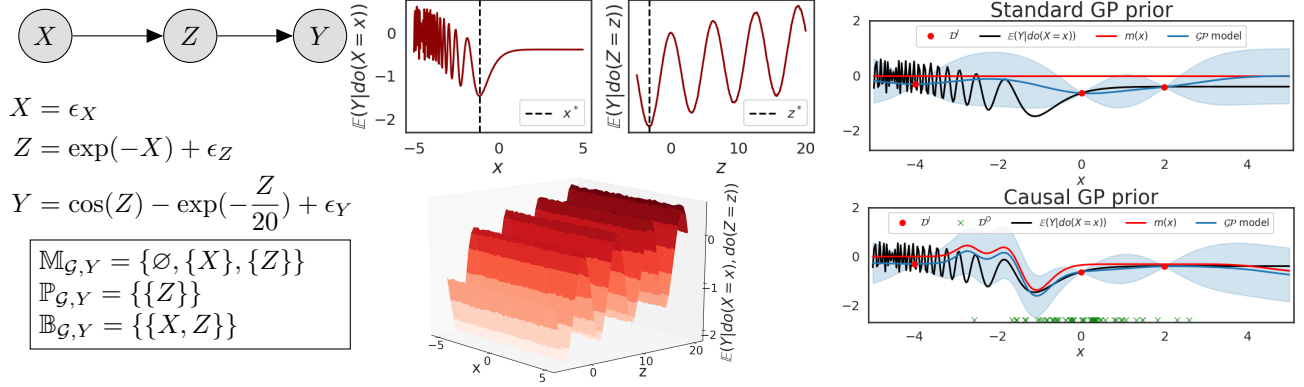


Figure 3: Toy example illustrating the elements of CBO. *Left panel:* DAG, SEM and optimal sets. *Central panel:* Objective functions for different intervention sets. *Right panel:* Posterior GP obtained with two different prior.

vening always improves  $Y$  with respect to the remaining elements in  $\mathbb{M}_{\mathcal{G}, Y}^C$ . For completeness, we also use  $\mathbb{B}_{\mathcal{G}, Y}^C$  to denote the unique set on which BO performs interventions that includes all manipulative variables  $\mathbf{X}$ . In Fig. 3 we give an example in which  $|\mathbb{M}_{\mathcal{G}, Y}| < |\mathcal{P}(\mathbf{X})|$  and intervening on  $\mathbb{B}_{\mathcal{G}, Y}$  is suboptimal. Indeed, the central panel shows how  $\mathbb{E}[Y|\text{do}(X=x), \text{do}(Z=z)] = \mathbb{E}[Y|\text{do}(Z=z)]$  and the causal intrinsic dimensionality of this problem is  $|\text{Pa}(Y)| = 1$ . In addition,  $\mathbb{E}[Y|\text{do}(X=x^*)] > \mathbb{E}[Y|\text{do}(Z=z^*)]$  thus  $Z$  is optimal with respect to  $X$  and is the only set in  $\mathbb{P}_{\mathcal{G}, Y}$ . For notational convenience, we refer to the exploration set, which can be  $\mathbb{M}_{\mathcal{G}, Y}^C$  or  $\mathbb{P}_{\mathcal{G}, Y}^C$ , as  $\mathbf{ES}$ . The choice of the  $\mathbf{ES}$  depends on the causal graph and the next sections are agnostic to the choice of the  $\mathbf{ES}$ .

### 3.2 Causal GP Model

To integrate experimental and observational data, for each  $\mathbf{X}_s \in \mathbf{ES}$ , we place a GP prior on  $f(\mathbf{x}_s) = \mathbb{E}[Y|\text{do}(\mathbf{X}_s = \mathbf{x}_s)]$  with prior mean and kernel function computed via do-calculus:

$$f(\mathbf{x}_s) \sim \mathcal{GP}(m(\mathbf{x}_s), k_C(\mathbf{x}_s, \mathbf{x}'_s)) \quad (2)$$

$$m(\mathbf{x}_s) = \hat{\mathbb{E}}[Y|\text{do}(\mathbf{X}_s = \mathbf{x}_s)] \quad (3)$$

$$k_C(\mathbf{x}_s, \mathbf{x}'_s) = k_{RBF}(\mathbf{x}_s, \mathbf{x}'_s) + \sigma(\mathbf{x}_s)\sigma(\mathbf{x}'_s) \quad (4)$$

where  $\sigma(\mathbf{x}_s) = \sqrt{\hat{\mathbb{V}}(Y|\text{do}(\mathbf{X}_s = \mathbf{x}_s))}$  with  $\hat{\mathbb{V}}$  representing the variance estimated from observational data and  $k_{RBF}$  representing the radial basis function kernel defined as  $k_{RBF}(\mathbf{x}_s, \mathbf{x}'_s) := \exp(-\frac{\|\mathbf{x}_s - \mathbf{x}'_s\|^2}{2l^2})$ . Fig. 3 (right) illustrates the difference between the posterior GP distribution obtained with a zero mean prior distribution and RBF kernel (upper) and with the proposed causal prior (lower). Notice how the prior mean function captures the behaviour on the target function in area where observations are available (crosses at the bottom) despite the lack of interventional data. In addition, the posterior variance is higher in area where

$\sqrt{\hat{\mathbb{V}}(Y|\text{do}(\mathbf{X}_s = \mathbf{x}_s))}$  is inflated due to the lack of observational data, which enables proper calculation of the uncertainties about the causal effects.

### 3.3 Causal Acquisition Function

For each  $\mathbf{X}_s \in \mathbf{ES}$ , we define the acquisition function as the expected improvement (EI) with respect to the current best observed interventional setting across all sets in  $\mathbf{ES}$ . At every step of the optimization, denote by  $y_s = \mathbb{E}[Y|\text{do}(\mathbf{X}_s = \mathbf{x})]$  and  $y^*$  the optimal value of  $y_s$ ,  $s = 1, \dots, |\mathbf{ES}|$  observed so far. The EI is given by:

$$EI^s(\mathbf{x}) = \mathbb{E}_{p(y_s)}[\max(y_s - y^*, 0)]/Co(\mathbf{x}). \quad (5)$$

Let  $\alpha_1, \dots, \alpha_{|\mathbf{ES}|}$  be solutions of the optimization of  $EI^s(\mathbf{x})$  for each set in  $\mathbf{ES}$  and  $\alpha^* := \max\{\alpha_1, \dots, \alpha_{|\mathbf{ES}|}\}$ . The best new intervention set is given by  $s^* = \text{argmax}_{s \in \{1, \dots, |\mathbf{ES}|\}} \alpha_s$ . Therefore, the set-value pair to intervene on is  $(s^*, \alpha^*)$ . Fig. 4 (left and center) shows the acquisition functions for  $\mathbf{ES} = \mathbb{M}_{\mathcal{G}, Y}^C$  in the toy example. The new intervention is selected by comparing the maxima of the acquisition functions across interventions (red and back dots).

### 3.4 $\epsilon$ -greedy Policy

For some graph structures, such as in Fig 2 (top), the empty set, which represents the observational case, is part of  $\mathbf{ES}$ . A mechanism in the optimization process is then needed to observe the system when that is the optimal strategy. Inspired by the  $\epsilon$ -greedy policies in RL (Tokic, 2010), we propose an  $\epsilon$ -greedy approach where  $\epsilon$  determines the probability of observing and trades off exploration and exploitation. The value of  $\epsilon$ , which is a parameter of CBO, can be selected in several different ways and needs to balance the *observation-intervention trade-off* emerging in CBO. On the one hand, collecting observational data allows to reliably estimate causal effects via *do*-calculus. On the other

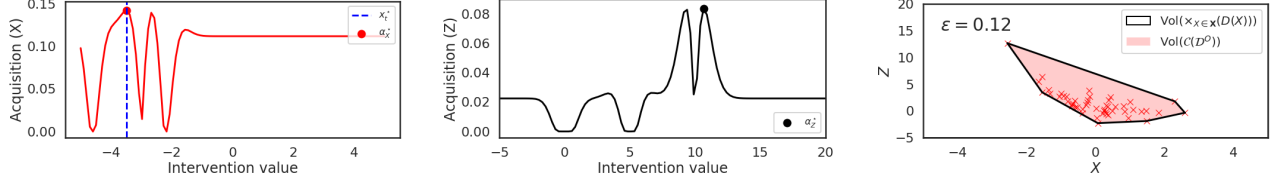


Figure 4: Toy example. *Left and central panels:* Acquisition functions for the variables in  $\mathbb{M}_{\mathcal{G}, Y}^{\mathcal{C}}$ . The dashed blue line gives the next optimal intervention. *Right panel:* Convex hull in the  $X$ - $Z$  space used to calculate  $\epsilon$ . Red crosses give  $\mathcal{D}^O$  while the boundaries of the plot correspond to the interventional domain.

hand, computing consistent causal effects for values outside of the observational range, requires intervening. The agent needs to find the right combination of the two actions so as to exploit observational information while intervening in areas where the uncertainty is higher. In this paper, we define  $\epsilon$  as:

$$\epsilon = \frac{\text{Vol}(\mathcal{C}(\mathcal{D}^O))}{\text{Vol}(\times_{X \in \mathbf{X}}(D(X)))} \times \frac{N}{N_{\max}}, \quad (6)$$

where  $\text{Vol}(\mathcal{C}(\mathcal{D}^O))$  represents the volume of the convex hull for the observational data and  $\text{Vol}(\times_{X \in \mathbf{X}}(D(X)))$  gives the volume of the interventional domain (see Fig. 4 (right)).  $N_{\max}$  represents the maximum number of observations the agent is willing to collect and  $N$  is the current size of  $\mathcal{D}^O$ . When  $\text{Vol}(\mathcal{C}(\mathcal{D}^O))$  is small with respect to  $N$ , the interventional space is bigger than the observational space. We thus intervene and explore part of the interventional space not covered by  $\mathcal{D}^O$ . On the contrary, if  $\text{Vol}(\mathcal{C}(\mathcal{D}^O))$  is large with respect to  $N$ , we obtain consistent estimates of the causal effects by collecting more observations. We thus observe and update the prior GP in Eqs. (3) - (4).

Other  $\epsilon$ -greedy policies can be formulated in order to solve the trade-off differently. For instance, the agent could define an adaptive  $\epsilon$  which favours observations in the first stages of the optimization procedure and interventions as  $N$  increases. Alternatively, the value of  $\epsilon$  could depend on the agent’s budget and favours interventions when their cost is low.

### 3.5 The CBO Algorithm

We give the complete CBO algorithm in Alg. 1). The time complexity of CBO is dominated by algebraic operations on  $k_C(\mathbf{x}_s, \mathbf{x}'_s)$  which are  $\mathcal{O}(P^3)$  where  $P$  denotes the number of function evaluations of the BO algorithm. The space complexity is dominated by storing  $k_C(\mathbf{x}_s, \mathbf{x}'_s)$  which is  $\mathcal{O}(P^2)$ . Given the acquisition function and the surrogate model, the theoretical guarantees of CBO are limited and follow directly from the theoretical properties of *do*-calculus. However, one could extend CBO to use a GP-UCB acquisition function for which a cumulative regret bound has been derived (Srinivas et al., 2012). We leave this for future work.

---

#### Algorithm 1: Causal Bayesian Optimization - CBO

---

**Data:**  $\mathcal{D}^O, \mathcal{D}^I, \mathcal{G}, \mathbf{ES}$ , number of steps  $T$

**Result:**  $\mathbf{X}_s^*, \mathbf{x}_s^*, \hat{\mathbb{E}}[\mathbf{Y}^* | \text{do}(\mathbf{X}_s^* = \mathbf{x}_s^*)]$

**Initialise:** Set  $\mathcal{D}_0^I = \mathcal{D}^I$  and  $\mathcal{D}_0^O = \mathcal{D}^O$

**for**  $t=1, \dots, T$  **do**

Compute  $\epsilon$  and sample  $u \sim \mathcal{U}(0, 1)$

**if**  $\epsilon > u$  **then**

(Observe)

1. Observe new observations  $(\mathbf{x}_t, c_t, \mathbf{y}_t)$ .

2. Augment  $\mathcal{D}^O = \mathcal{D}^O \cup \{(\mathbf{x}_t, c_t, \mathbf{y}_t)\}$ .

3. Update prior of the causal GP (Eq. (2)).

**end**

**else**

(Intervene)

1. Compute  $EI^s(\mathbf{x})/Co(\mathbf{x})$  for each element  $s \in \mathbf{ES}$  (Eq. (5)).

2. Obtain the optimal interventional set-value pair  $(s^*, \alpha^*)$ .

3. Intervene on the system.

4. Update posterior of the causal GP.

**end**

**end**

Return the optimal value  $\hat{\mathbb{E}}[\mathbf{Y}^* | \text{do}(\mathbf{X}_s^* = \mathbf{x}_s^*)]$  in  $\mathcal{D}_T^I$  and the corresponding  $\mathbf{X}_s^*, \mathbf{x}_s^*$ .

---

## 4 Experiments

We test our algorithm on a synthetic setting and on two real-world applications for which a DAG is available and can be used as a simulator. We run CBO to explore both  $\mathbb{M}_{\mathcal{G}, Y}^{\mathcal{C}}$  and  $\mathbb{P}_{\mathcal{G}, Y}^{\mathcal{C}}$  and show how the optimal intervention set, intervention values and cost incurred to achieve the optimum change depending on the DAG and the SEM. For all variables in  $\mathbf{X}$  and their combinations, we assume to have data from previous interventions which we denote by  $\mathcal{D}^I = \{(\mathbf{x}_s^i, \mathbb{E}[Y | \text{do}(\mathbf{X}_s^i = \mathbf{x}_s^i)])\}_{i=1, s=1}^{P, |\mathbf{ES}|}$ . Typically  $P$  is very small and interventions are prohibitive to implement.

**Baselines:** We compare CBO against a standard BO algorithm, in which all variables are intervened upon, and a CBO version where a standard GP prior given by  $p(f(\mathbf{x}_s)) = \mathcal{GP}(0, k_{\text{RBF}}(\mathbf{x}_s, \mathbf{x}'_s))$  is used.

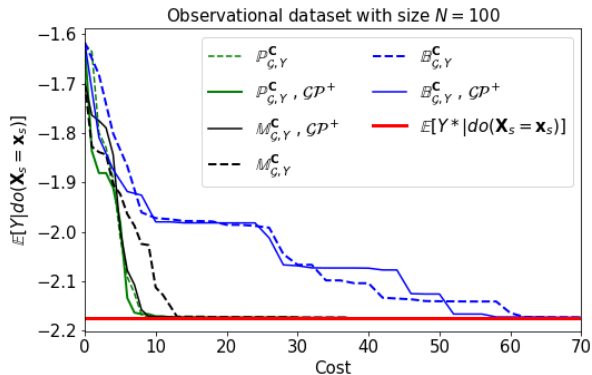


Figure 5: Toy example. Convergence of CBO and standard BO across different initializations of  $\mathcal{D}^I$ . The red line gives the optimal  $Y^*$  when intervening on sets in  $\mathbb{M}_{\mathcal{G},Y}^C$ ,  $\mathbb{P}_{\mathcal{G},Y}^C$  or  $\mathbb{E}_{\mathcal{G},Y}^C$ . Solid lines give CBO results when using the causal GP model which is denoted by  $\mathcal{GP}^+$ . Dotted line correspond to CBO with a standard GP prior model  $p(f(\mathbf{x}_s)) = \mathcal{GP}(0, k_{\text{RBF}}(\mathbf{x}_s, \mathbf{x}'_s))$ . See the supplement for standard deviations (Fig. (1)).

**Performance measures:** We run CBO with different initializations of  $\mathcal{D}^I$  and report the average convergence performances together with standard errors. In the synthetic setting, we consider three different cost configurations: equal unit cost per node, different fix costs per node and variable costs per node. The total cost at each optimization step is computed as the sum of the cost for each intervened node. We show the results for equal unit cost per node and report the full comparison in the supplement.

#### 4.1 Toy Experiment

We show the convergence results for CBO and competing algorithms for the toy example described in the text. For this experiment we set  $N = 100$  and  $P = 3$ . Given the SEM in Fig. 3, the optimal configuration is  $(X_s^*, x_s^*) = (Z, -3.20)$ . CBO converges to the optimum faster than BO which requires to intervene on all nodes and it is thus twice more expensive (Fig. 5).

#### 4.2 Synthetic Experiment

We test the algorithm on the DAG given in Fig. 2 (top). This DAG includes unobserved confounders, non-manipulative variables and requires to apply both front-door and back-door adjustment formulas to estimate the causal effects. We set  $P = 10$  and test different values of  $N$ . The ESS are  $\mathbb{M}_{\mathcal{G},Y}^C = \{\emptyset, \{B\}, \{D\}, \{E\}, \{B, D\}, \{B, E\}, \{D, E\}\}$ ,  $\mathbb{P}_{\mathcal{G},Y}^C = \{\emptyset, \{B\}, \{D\}, \{E\}, \{B, D\}, \{D, E\}\}$  and  $\mathbb{E}_{\mathcal{G},Y}^C = \{B, D, E\}$ . All the intervention sets in  $\mathbb{M}_{\mathcal{G},Y}^C$  and  $\mathbb{P}_{\mathcal{G},Y}^C$  include a maximum of two variables. On the contrary, BO only considers interventions on three

variables thus increasing the dimensionality of the problem by one. The SEM, the *do*-calculus computations and further details about the experimental settings are given in the supplement.

Fig. 6 shows how CBO outperforms standard BO and achieves the best performance when the causal GP model is used. There are two main reasons why BO leads to a suboptimal solution. Firstly,  $\mathbb{E}[Y|\text{do}(B = b), \text{do}(D = d), \text{do}(E = e)] = \mathbb{E}[Y|\text{do}(D = d), \text{do}(E = e)]$ . This means that the same outcome can be achieved by intervening on  $\{D, E\}$  at a significantly lower cost.

Secondly, although it may seem counterintuitive, intervening only on a subset of variables leads to better outcomes. Manipulating all variables breaks the causal links between them and blocks the propagation of causal effects in the graph. In this example, intervening on  $B, E, D$  blocks the causal effect of  $B$  on  $Y$ . Manipulating only  $B$  leads to a propagation of its causal effect through  $D$  and  $E$ . Given the SEM,  $\mathbb{E}[Y|\text{do}(B = b, D = d, E = e)] < \mathbb{E}[Y|\text{do}(B = b, D = d)] \forall b \in D(B), d \in D(D), e \in D(E)$ . Indeed, setting the level of  $B$  makes  $D$  and  $E$  take values outside of their interventional domains  $D(D)$  and  $D(E)$  thus leading to function values not achievable in BO. Furthermore, the causal GP prior determines the locations of the function evaluations thus reducing the number of steps required to find the optimum. As expected, the benefit of incorporating  $\mathcal{D}^O$  into the prior becomes more evident when  $N$  increases. The optimal configuration for this setting is  $(\mathbf{X}_s^*, \mathbf{x}_s^*) = (\{B, D\}, (-5.0, 3.28))$ .

#### 4.3 Example in Ecology

We apply CBO to a large-scale optimization problem in ecology. We consider the issue of maximizing the net coral ecosystem calcification (NEC) in the Bermuda given a set of environmental variables. The causal graph (Fig. 4 in the supplement) is taken from Courtney et al. (2017) and modified so as to avoid directed cycles. We consider a subset of 5 variables as manipulative, that is  $\mathbf{X} = \{\text{Nut}, \Omega_A, \text{Chl}\alpha, \text{TA}, \text{DIC}\}$ , and assume to be able to intervene *contemporaneously* on a maximum of 3 variables. Given these assumptions and the DAG,  $\mathbb{M}_{\mathcal{G},Y}^C$  includes the single variable interventions and all the 2 and 3 variables interventions that can be performed selecting variables in  $\mathbf{X}$ . The cardinality of  $\mathbb{M}_{\mathcal{G},Y}^C$  is thus 25. Notice that the size of  $\mathbb{E}_{\mathcal{G},Y}^C = \{\text{Nut}, \Omega_A, \text{Chl}\alpha, \text{TA}, \text{DIC}\}$  is greater than 3 thus BO is not a viable strategy for this application. We first construct a simulator by fitting a linear SEM with the 50 observations provided by Andersson (2018). We then use the simulator to generate  $N = 500$  observations

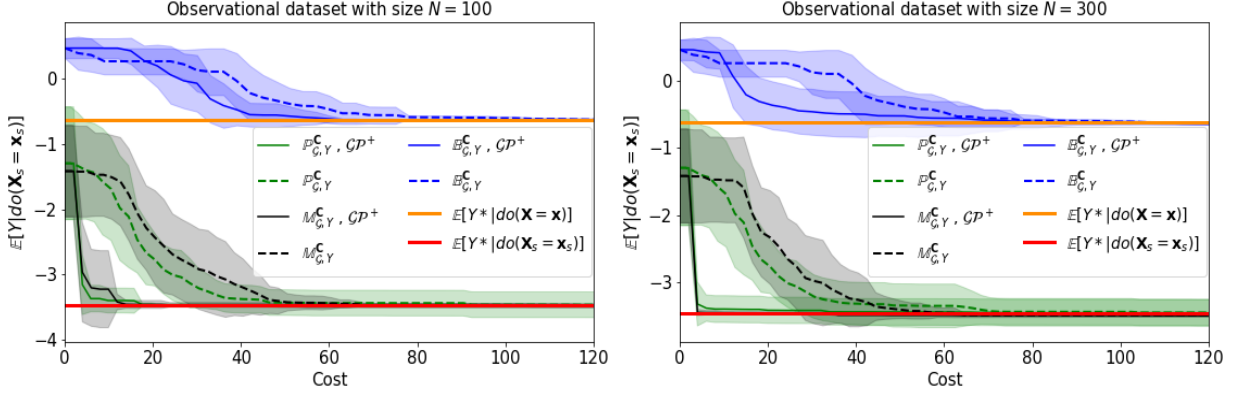


Figure 6: Synthetic example. Convergence of CBO and BO across different initialization of  $\mathcal{D}^I$ . The orange line gives the optimal  $Y^*$  when intervening on  $\mathbb{B}_{\mathcal{G},Y}^C$ . The red line gives the optimal  $Y^*$  when intervening on sets in  $\mathbb{M}_{\mathcal{G},Y}^C$  or  $\mathbb{P}_{\mathcal{G},Y}^C$ . Solid lines give CBO results when using the causal GP model, denoted by  $\mathcal{G}\mathcal{P}^+$ , while dotted lines correspond to CBO with a standard GP prior model. Shaded areas are  $\pm$  standard deviation.

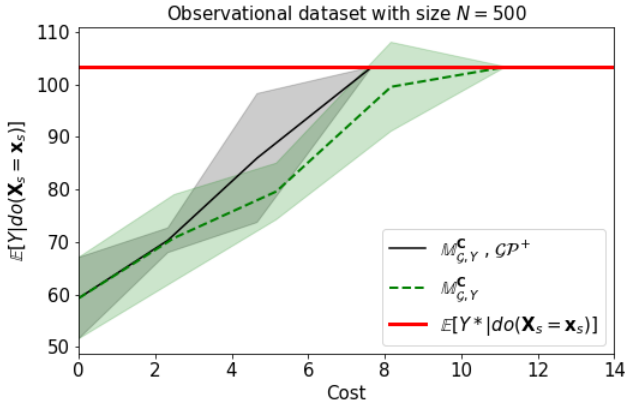


Figure 7: NEC example. Convergence of CBO across different initialization of  $\mathcal{D}^I$ , with and without causal GP prior. The red line gives the optimal  $Y^*$  when intervening on  $\mathbb{M}_{\mathcal{G},Y}^C$ .

and  $P = 1$  initial interventional data points. We set the interventional domains to  $D(\text{Nut}) = [-2, 5]$ ,  $D(\Omega_A) = [2, 4]$ ,  $D(\text{Chl}\alpha) = [0.3, 0.4]$ ,  $D(\text{TA}) = [2200, 2550]$  and  $D(\text{DIC}) = [1950, 2150]$ . We run CBO on  $\mathbb{M}_{\mathcal{G},Y}^C$ , with and without the causal GP prior. We found CBO to successfully explore  $\mathbb{M}_{\mathcal{G},Y}^C$ , especially when the causal GP prior is used (Fig. 7). The optimal intervention is  $(\mathbf{X}_s^*, \mathbf{x}_s^*) = (\{\Omega_A, \text{TA}, \text{DIC}\}, (2, 2550, 1950))$ .

#### 4.4 Example in Healthcare

Finally, we apply our method to an example in healthcare. The DAG (Fig. 3 in the supplement) is taken from Thompson (2019) and Ferro et al. (2015) which is used to model the causal effect of statin drugs on the levels of prostate specific antigen (PSA). Our goal is to minimize PSA by intervening on statin and aspirin usage.  $\mathcal{D}^O$  consists of  $N = 500$  instances sampled from the simulator while  $P = 3$ . Given the causal struc-

ture,  $\mathbb{M}_{\mathcal{G},Y}^C = \{\emptyset, \{\text{aspirin}\}, \{\text{statin}\}, \{\text{aspirin}, \text{statin}\}\}$  while  $\mathbb{P}_{\mathcal{G},Y}^C = \{\{\text{aspirin}, \text{statin}\}\}$ . We set the domain  $D(\text{aspirin}) = D(\text{statin}) = [0.0, 1.0]$  and run CBO on both  $\mathbb{M}_{\mathcal{G},Y}^C$  and  $\mathbb{P}_{\mathcal{G},Y}^C$ . We found the optimal intervention to be  $(\mathbf{X}_s^*, \mathbf{x}_s^*) = (\{\text{aspirin}, \text{statin}\}, (0.0, 1.0))$ . This experiment shows how CBO can help doctors, and decision-makers in general, to find optimal interventions in real-life scenarios based on simulators and thus avoiding expensive and invasive interventions.

## 5 Conclusions

This paper formalizes the problem of globally optimizing a variable that is part of a causal model in which a sequence of interventions can be performed. We propose a Causal Bayesian Optimization (CBO) algorithm which solves the global optimization problem exploring a set of potentially optimal sets. This is achieved via a causal expected improvement acquisition function and an  $\epsilon$ -greedy policy solving the emerging observation-intervention trade off. In addition, we formulate a causal GP model which allows to integrate observational and interventional data via *do*-calculus. We show the benefits of the proposed approach in a variety of settings characterized by different causal graph structures. Our results demonstrate how CBO outperforms BO and reaches the global optimum after a significantly lower number of optimization steps. Future work will focus on reducing the number of causal GPs used by CBO to explore  $\mathbf{ES}$ . For instance, this could be achieved through a multi-output GP model where different outputs correspond to different interventional sets. In addition, we will work on extending CBO to multi-objective optimization. Finally, we will focus on combining the proposed framework with a causal discovery algorithm so as to account for uncertainty in the graph structure.



## References

- Andersson, A., B. N. (2018). In situ measurements used for coral and reef-scale calcification structural equation modeling including environmental and chemical measurements, and coral calcification rates in bermuda from 2010 to 2012 (beacon project). *Biological and Chemical Oceanography Data Management Office (BCO-DMO). Dataset version 2018-03-02*. <http://lod.bco-dmo.org/id/dataset/720788>.
- Bareinboim, E., Forney, A., and Pearl, J. (2015). Bandits with unobserved confounders: A causal approach. In *Advances in Neural Information Processing Systems*, pages 1342–1350.
- Buesing, L., Weber, T., Zwols, Y., Racaniere, S., Guez, A., Lespiau, J.-B., and Heess, N. (2018). Woulda, coulda, shoulda: Counterfactually-guided policy search. *arXiv preprint arXiv:1811.06272*.
- Cochran, W. and Cox, G. (1957). *Experimental design*. John Wiley and Sons, Inc., New York, NY.
- Courtney, T. A., Lebrato, M., Bates, N. R., Collins, A., De Putron, S. J., Garley, R., Johnson, R., Molinero, J.-C., Noyes, T. J., Sabine, C. L., et al. (2017). Environmental controls on modern scleractinian coral and reef-scale calcification. *Science advances*, 3(11):e1701356.
- Ferro, A., Pina, F., Severo, M., Dias, P., Botelho, F., and Lunet, N. (2015). Use of statins and serum levels of prostate specific antigen. *Acta Urológica Portuguesa*, 32(2):71–77.
- Foerster, J. N., Farquhar, G., Afouras, T., Nardelli, N., and Whiteson, S. (2018). Counterfactual multi-agent policy gradients. In *Thirty-Second AAAI Conference on Artificial Intelligence*.
- Glymour, C., Zhang, K., and Spirtes, P. (2019). Review of causal discovery methods based on graphical models. *Frontiers in Genetics*, 10.
- Guo, R., Cheng, L., Li, J., Hahn, P. R., and Liu, H. (2018). A survey of learning causality with data: Problems and methods. *arXiv preprint arXiv:1809.09337*.
- Jones, D. R., Schonlau, M., and Welch, W. J. (1998). Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 13(4):455–492.
- Lattimore, F., Lattimore, T., and Reid, M. D. (2016). Causal bandits: Learning good interventions via causal inference. In *Advances in Neural Information Processing Systems*, pages 1181–1189.
- Lee, S. and Bareinboim, E. (2018). Structural causal bandits: where to intervene? In *Advances in Neural Information Processing Systems*, pages 2568–2578.
- Lee, S. and Bareinboim, E. (2019). Structural causal bandits with non-manipulable variables. Technical report, Technical Report R-40, Purdue AI Lab, Department of Computer Science, Purdue.
- Lu, C., Schölkopf, B., and Hernández-Lobato, J. M. (2018). Deconfounding reinforcement learning in observational settings. *arXiv preprint arXiv:1812.10576*.
- Ortega, P. A. and Braun, D. A. (2014). Generalized thompson sampling for sequential decision-making and causal inference. *Complex Adaptive Systems Modeling*, 2(1):2.
- Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, 82(4):669–688.
- Pearl, J. (2000). *Causality: models, reasoning and inference*, volume 29. Springer.
- Rasmussen, C. E. (2003). Gaussian processes in machine learning. In *Summer School on Machine Learning*, pages 63–71. Springer.
- Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., and De Freitas, N. (2015). Taking the human out of the loop: A review of bayesian optimization. *Proceedings of the IEEE*, 104(1):148–175.
- Srinivas, N., Krause, A., Kakade, S. M., and Seeger, M. W. (2012). Information-theoretic regret bounds for gaussian process optimization in the bandit setting. *IEEE Transactions on Information Theory*, 58(5):3250–3265.
- Thompson, C. (2019). Causal graph analysis with the causalgraph procedure. <https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2019/2998-2019.pdf>.
- Tokic, M. (2010). Adaptive  $\epsilon$ -greedy exploration in reinforcement learning based on value differences. In *Annual Conference on Artificial Intelligence*, pages 203–210. Springer.
- Wang, Z., Hutter, F., Zoghi, M., Matheson, D., and de Freitas, N. (2016). Bayesian optimization in a billion dimensions via random embeddings. *Journal of Artificial Intelligence Research*, 55:361–387.