

# Bayesian Optimization for Unknown Cost-Varying Variable Subsets with No-Regret Costs

Vu Viet Hoang<sup>1\*</sup>, Quoc Anh Hoang Nguyen<sup>1\*</sup>, Hung The Tran<sup>2\*†</sup>

<sup>1</sup>FPT Software AI Center

<sup>2</sup>Hanoi University of Science and Technology

HoangVV18@fpt.com, AnhNQH1@fpt.com, Hungtt@soict.hust.edu.vn

## Abstract

Bayesian Optimization (BO) is a widely-used method for optimizing expensive-to-evaluate black-box functions. Traditional BO assumes that the learner has full control over all query variables without additional constraints. However, in many real-world scenarios, controlling certain query variables may incur costs. Therefore, the learner needs to balance the selection of informative subsets for targeted learning against leaving some variables to be randomly sampled to minimize costs. This problem is known as Bayesian Optimization with cost-varying variable subsets (BOCVS). While the goal of BOCVS is to identify the optimal solution with minimal cost, previous works have only guaranteed finding the optimal solution without considering the total costs incurred. Moreover, these works assume precise knowledge of the cost for each subset, which is often unrealistic. In this paper, we propose a novel algorithm for the extension of the BOCVS problem with random and unknown costs that separates the process into exploration and exploitation phases. The exploration phase will filter out low-quality variable subsets, while the exploitation phase will leverage high-quality ones. Furthermore, we theoretically demonstrate that our algorithm achieves a sub-linear rate in both quality regret and cost regret, addressing the objective of the BOCVS problem more effectively than previous analyses. Finally, we show that our proposed algorithm outperforms comparable baselines across a wide range of benchmarks.

## Introduction

Bayesian optimization (BO) is a powerful technique designed for the sample-efficient optimization of expensive-to-evaluate black-box objective functions. It has been effectively used in various important experimental design problems, such as hyperparameter optimization (Snoek, Larochelle, and Adams 2012), chemical synthesis (Shields et al. 2021), and particle accelerator control (Kirschner et al. 2022). Typically, BO has access to all query variables of the objective function, allowing it to re-evaluate any variable without additional hamper. However, in many practical situations, not all query variables can be easily modified. Some variables may follow specific distributions, and altering their

values deterministically may incur costs. This setting is known as *Bayesian Optimization with Cost-varying variable subsets* (BOCVS), as introduced in (Tay et al. 2023). The aim is to identify optimal candidates for the objective function while minimizing incurred costs as much as possible. Naturally, greater control over the query variables leads to higher costs. Therefore, in each iteration, the learner must balance between the set of controllable query variables and the associated costs.

We give an illustrative example with the BOCVS setting. Other concrete scenarios can be found in (Tay et al. 2023). In the context of adaptive control for manufacturing systems (Kruger et al. 2011), the goal is to optimize the production process to maximize output quality while managing the associated costs of process adjustments. The operator must adjust various control parameters, such as machine settings, temperature, and operational schedules, to enhance production quality. However, these adjustments can incur costs like energy consumption, maintenance or operating costs. To mitigate these costs, the operator can selectively modify certain control parameters while allowing others to vary randomly within a controlled range. The primary challenge lies in the fact that the cost constraints are not fully known or predictable, complicating the evaluation of the impact of randomized versus deterministic adjustments. These costs are often uncertain because they depend on fluctuating factors such as labor, or machine conditions. The operator must strike a balance between managing the risk of high, unknown costs and maintaining acceptable production quality.

To the best of our knowledge, UCB-CVS (Tay et al. 2023) is the first and only work that addresses this problem precisely. While UCB-CVS has shown promising theoretical and empirical results, there are some drawbacks. Firstly, the definition of their objective function focuses solely on finding the optimal solution and neglects the cost constraints. To incorporate cost constraints, the authors introduced the cost-varying cumulative regret, a cost-weighted version of the regular cumulative regret. While it can be proven that UCB-CVS achieves sublinear cost-varying cumulative regret, it is unclear whether it minimizes the total cost. For example, if there exists two control sets with similar performance while the cost is significantly difference, the sub-linear rate for cost-varying regret stated in Theorem 4.1 (Tay

\*These authors contributed equally.

†Corresponding author.

Copyright © 2025, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

et al. 2023) does not emphasize the advantage of choosing cheaper control set. Therefore, we need a new definition of the objective function and regret to better reflect the goals of the BOCVS setting. Secondly, UCB-CVS assumes that the cost of every control set is deterministically known, which is a strict assumption that limits the algorithm’s practicality. It would be more desirable to have an algorithm that does not require precise knowledge of the cost of every control sets. Thirdly, UCB-CVS uses an  $\epsilon$ -schedule as a relaxation to enable exploration with cheaper control sets. However, choosing an appropriate  $\epsilon$ -schedule is challenging. In practice, UCB-CVS uses a simple *explore-then-commit* approach instead of an  $\epsilon$ -schedule. This creates a mismatch between the theoretical algorithm and its practical implementation.

In this paper, we address an extension of the BOCVS problem where the cost of playing a subset is unknown. We introduce an efficient novel approach to analyzing the effectiveness of an algorithm by defining both quality regret and cost regret, which is a more rigorous analysis compared to previous work. An efficient algorithm should aim to minimize both types of regret. Consequently, we develop an algorithm designed to minimize these regrets simultaneously. The main contributions of this work are as follows:

- We introduce an extension of the BOCVS problem in which the costs of the subset may be random and unknown.
- We develop a cost-sensitive algorithm for the extension version. We derive an upper bound on the cumulative regret for our algorithm and theoretically show that the algorithm achieves sub-linear cumulative regret of both the objective function value and the cost.
- We conducted an empirical evaluation of our proposed algorithm’s performance against baseline methods across a variety of experimental conditions. This included testing on both synthetic and real-world datasets, specifically a plant growth dataset and an airfoil self-noise dataset, which are relevant to the precision agriculture and advanced manufacturing applications discussed earlier.

## Preliminaries

### Gaussian Process and Bayesian Optimization

We consider a query set  $\mathcal{X}$  and an objective function  $f : \mathcal{X} \rightarrow \mathbb{R}$  where  $\mathcal{X} \subset \mathbb{R}^d$ . The learner’s goal is to find the query that maximizes  $f$ , denoted as  $\mathbf{x}^* = \operatorname{argmax}_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x})$ . However,  $f$  is a black-box function, meaning it is not available in a closed form and can only be evaluated by submitting a query  $\mathbf{x}_t \in \mathcal{X}$  at each iteration  $t$ , and receiving a noisy observation  $y_t = f(\mathbf{x}_t) + \xi_t$ , where each  $\xi_t$  is i.i.d.  $\sigma$ -sub-Gaussian noise with zero mean.

Gaussian processes (Williams and Rasmussen 2006) (GPs) are a popular choice as a surrogate model for optimizing  $f$  in BO because they enable exact posterior inference: The GP posterior estimate of  $f$  at any query  $\mathbf{x} \in \mathcal{X}$  after  $t$  iterations is a Gaussian with posterior mean and variance given by

$$\begin{aligned} \mu_t(\mathbf{x}) &:= \mathbf{k}_t(\mathbf{x})^\top (\mathbf{K}_t + \lambda \mathbf{I})^{-1} \mathbf{y}_t, \\ \sigma_t^2(\mathbf{x}) &:= k(\mathbf{x}, \mathbf{x}) - \mathbf{k}_t(\mathbf{x})^\top (\mathbf{K}_t + \lambda \mathbf{I})^{-1} \mathbf{k}_t(\mathbf{x}) \end{aligned} \quad (1)$$

where  $\mathbf{y}_t := (y_j)_{j=1}^t \in \mathbb{R}^t$ ,  $k$  is a positive semidefinite kernel (covariance function),  $\mathbf{k}_t(\mathbf{x}) := (k(\mathbf{x}, \mathbf{x}_j))_{j=1}^t \in \mathbb{R}^t$ ,  $\mathbf{K}_t := (k(\mathbf{x}_j, \mathbf{x}_{j'}))_{j, j'=1}^t \in \mathbb{R}^{t \times t}$ , and  $\lambda$  is an algorithm parameter; if the noise is a Gaussian with variance  $\sigma^2$ , then the true posterior is recovered with  $\lambda = \sigma^2$ . The kernel  $k$  is an important modeling choice as the GP posterior mean will reside in the reproducing kernel Hilbert space (RKHS) associated with  $k$ . For simplicity, we assume w.l.o.g. that  $k(\mathbf{x}, \mathbf{x}') \leq 1$  for all  $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$ . Kernel  $k$  affects the maximum information gain (MIG) defined as

$$\gamma_T(\mathcal{X}) := \max_{\{\mathbf{x}_t\}_{t=1}^T \subseteq \mathcal{X}} 0.5 \log |\mathbf{I} + \lambda^{-1} \mathbf{K}_T|.$$

The MIG characterizes the statistical complexity of a problem and plays an integral role in the theoretical analysis. For the commonly used squared exponential kernel,  $\gamma_T(\mathcal{X}) = \mathcal{O}((\log T)^{d+1})$ , while for the Matérn kernel with  $\nu > 1$ ,  $\gamma_T(\mathcal{X}) = \mathcal{O}(T^{d(d+1)/(2\nu+d(d+1))}(\log T))$  (Srinivas et al. 2009). Importantly,  $\gamma_T(\mathcal{X})$  is increasing in the volume of  $\mathcal{X}$  (Srinivas et al. 2009).

### Multi-Armed Bandits

In more formal terms, the multi-armed bandit (MAB) problem can be described as follows: given a set of  $k$  arms (or actions), each arm  $i$  provides a reward from an unknown probability distribution. The goal is to develop a strategy or policy that maximizes the cumulative reward over a sequence of  $n$  trials.

One effective approach to solving the MAB problem is the Upper Confidence Bound (UCB) algorithm. The UCB algorithm addresses the exploration-exploitation dilemma by using a principled method to select which arm to pull. The core idea is to choose the arm that maximizes the upper confidence bound of its estimated reward, thereby balancing the potential for high rewards and the uncertainty associated with each arm (Slivkins et al. 2019).

The UCB (Slivkins et al. 2019) algorithm can be summarized in the following steps:

1. **Initialization:** Start by playing each arm once to gather initial data.
2. **Estimation:** After each round  $t$ , calculate the estimated average reward for each arm  $i$ , denoted as  $\hat{r}_i(t)$
3. **Confidence Bound Calculation:** Compute the upper confidence bound for each arm  $i$ , given by:

$$r_i^{UCB}(t) = \hat{r}_i(t) + \sqrt{2 \ln T / N_i(t)}$$

where  $N_i(t)$  is the number of times arm  $i$  has been played up to round  $t$ , and  $\ln T$  is the natural logarithm of the current round number, which accounts for the total number of trials and ensures logarithmic growth.

4. **Selection:** Select the arm  $i$  with the highest  $r_i^{UCB}(t)$  value for the next trial.

The term  $\sqrt{2 \ln T / N_i(t)}$  represents the confidence interval around the estimated reward  $\hat{\mu}_i(t)$ . As more data is gathered for a particular arm (i.e., as  $N_i(t)$  increases), the confidence interval narrows, reflecting increased certainty about

the arm's true reward distribution. Conversely, for arms that have been played fewer times, the wider confidence interval encourages exploration by allowing the possibility of higher potential rewards.

### Problem Definition

We introduce extension of BOCVS setting in this section. The problem involves a compact query set  $\mathcal{X} \subset \mathbb{R}^d$  and an objective function  $f : \mathcal{X} \rightarrow \mathbb{R}$  that resides in the Reproducing Kernel Hilbert Space (RKHS) associated with a kernel  $k$ , with the RKHS norm constrained by an upper bound  $B$ . Without loss of generality, we assume that  $\mathcal{X} = [0, 1]^d$ . Let  $[d] = \{1, 2, \dots, d\}$  represent the indices of the  $d$  variables.

The learner has access to a collection  $\mathcal{I} \subseteq 2^{[d]}$  of control sets, indexed by  $1, 2, \dots, m$ , where  $m = |\mathcal{I}|$ . Each control set  $i \in [m]$ , denoted as  $\mathcal{I}_i \subseteq [d]$ , specifies the subset of variables that the learner can choose to control. The complement of  $\mathcal{I}_i$ , denoted as  $\bar{\mathcal{I}}_i = [d] \setminus \mathcal{I}_i$ , indicates the subset of variables that will be randomly sampled from a known distribution.

A query  $\mathbf{x} \in \mathcal{X}$  can thus be decomposed into two parts: the control partial query  $\mathbf{x}^i$  and the random partial query  $\mathbf{x}^{-i}$ . The control partial query is defined as  $\mathbf{x}^i = (x_\ell)_{\ell \in \mathcal{I}_i}$ , which includes the variables indexed by  $\mathcal{I}_i$ . The random partial query is defined as  $\mathbf{x}^{-i} = (x_\ell)_{\ell \in \bar{\mathcal{I}}_i}$ , which includes the variables indexed by  $\bar{\mathcal{I}}_i$ . Therefore, we can write the query  $x \in \mathcal{X}$  as  $x = [\mathbf{x}^i, \mathbf{x}^{-i}]$ . Furthermore, let  $\mathcal{X}^i = \{\mathbf{x}^i \mid \mathbf{x} \in \mathcal{X}\}$  denote the set of all possible control partial queries.

In each iteration  $t$ , the learner selects a control set  $i_t \in \mathcal{I}$  and specifies the values for the control partial query  $\mathbf{x}_t^{i_t}$ . The values for the random partial query  $\mathbf{x}_t^{-i_t}$  are then sampled from the environment according to a known distribution. To simplify, the complete query for iteration  $t$  is  $\mathbf{x}_t = [\mathbf{x}_t^{i_t}, \mathbf{x}_t^{-i_t}] = [\mathbf{x}^{i_t}, \mathbf{x}^{-i_t}] = (x_{t,\ell})_{\ell \in [d]}$ . For each variable  $x_{t,\ell}$  where  $\ell \in \bar{\mathcal{I}}_{i_t}$ , it is a realization of a random variable  $X_{t,\ell} \sim \mathcal{P}_\ell$ . Consequently, the observed random partial query  $\mathbf{x}^{-i_t}$  is a realization of the random vector  $\mathbf{X}^{-i_t} = (X_{t,\ell})_{\ell \in \bar{\mathcal{I}}_{i_t}} \sim \mathbb{P}^{-i_t}$ , where  $\mathbb{P}^{-i_t}$  is the product measure  $\prod_{\ell \in \bar{\mathcal{I}}_{i_t}} \mathcal{P}_\ell$ . This means each variable in the random partial query is independently sampled from its respective probability distribution. All distributions  $\mathcal{P}_\ell$  are assumed to be known. The learner then observes the output  $y_t = f(\mathbf{x}_t) + \xi_t$ , where  $\xi_t$  is independently and identically distributed  $\sigma$ -sub-Gaussian noise with a zero mean.

The learner wishes to find the optimal control set  $i^+$  and specified values in control partial query  $\mathbf{x}^{i^+}$  that maximize the expected value of  $f([\mathbf{x}^i, \mathbf{X}^{-i}])$  where the expectation is w.r.t.  $\mathbf{X}^{-i} \sim \mathbb{P}^{-i}$ :

$$\left(i^+, \mathbf{x}^{i^+}\right) \in \underset{(i, \mathbf{x}^i) \in [m] \times \mathcal{X}^i}{\operatorname{argmax}} \mathbb{E} [f([\mathbf{x}^i, \mathbf{X}^{-i}])]. \quad (2)$$

In every iteration  $t$ , the learner pays a cost  $c_{i_t}^{(t)}$ , which is sampled from a unknown distribution with mean  $c_{i_t}$ . The distribution of costs for each subset may be different. The learning procedure ends after  $T$  iterations when  $C - \sum_{t=1}^T c_{i_t}^{(t)} < c_{i_{T+1}}^{(t)}$ , where  $C$  is the budget that can be paid.

Note that different from the original BOCVS setting (Tay et al. 2023), we do not require to know exactly the cost of every control subsets or the ordering of cost for control subsets, i.e.  $c_1 \leq c_2 \leq \dots \leq c_m$ .

In some scenarios, it is acceptable for the learner to sacrifice the quality of the objective value in order to manage costs more efficiently. For instance, in the adaptive control problem for manufacturing systems (Kruger et al. 2011), it is more practical for the operator to keep production quality above a certain predefined threshold rather than constantly striving to maximize product quality, which would result in higher costs. Therefore, different from (Tay et al. 2023), to manage effectively costs spent, assume  $f(x) \geq 0$  (can be achieved by subtracting the worst value that  $f$  can achieve), we allow the learner to be agnostic between subsets, whose expected reward is greater than  $1 - \alpha$  fraction of the highest expected value, for a fixed and known value of  $\alpha$ . The learner's objective is to learn and exploit the cheapest subset among these high-quality subsets as frequently as possible to specify the optimal value for the corresponding partial control query. More specifically, we denote the set of subsets  $\mathcal{C}_*$  whose expected value of  $f([\mathbf{x}^i, \mathbf{X}^{-i}])$  is within  $1 - \alpha$  factor of the highest expected value, i.e.,

$$\mathcal{C}_* = \left\{ i \in \mathcal{I} \mid \exists \mathbf{x}_i \in \mathcal{X}_i : \mathbb{E} [f([\mathbf{x}^i, \mathbf{X}^{-i}])] \geq (1 - \alpha) \max_{\substack{(i', \mathbf{x}^{i'}) \\ \in [m] \times \mathcal{X}^{i'}}} \mathbb{E} [f([\mathbf{x}^{i'}, \mathbf{X}^{-i'}])] \right\}$$

The learner's goal is to design an algorithm that will exploit the cheapest subset whose expected value is at least as large as the smallest tolerated reward. In other words, the learner needs to simultaneously identify and maximize the number of choices of subsets  $i_* = \operatorname{argmin}_{i \in \mathcal{C}_*} c_i$ . To measure the performance of any algorithms, we use two notions of regret:

- Quality regret:

$$R_T^f := \sum_{t=1}^T ((1 - \alpha) \mathbb{E} [f([\mathbf{x}^{i^+}, \mathbf{X}^{-i^+}])]) - \mathbb{E} [f([\mathbf{x}^{i_t}, \mathbf{X}^{-i_t}])]$$

- Cost regret

$$R_T^c := \sum_{t=1}^T \max\{(c_{i_t} - c_{i^*}), 0\}$$

The definitions of quality regret and cost regret in the paper show a notable resemblance to those outlined in (Sinha et al. 2021). In (Sinha et al. 2021), the authors employ the concepts of rewards and costs associated with arms in the MAB problem to characterize these types of regret. The objective then is to design an algorithm that simultaneously minimizes both the cost and quality regret.

**Discussion** Our definition of the objective function and regret introduces a new factor,  $\alpha$ , which represents the extent

to which the learner is willing to forgo the quality of the objective value in order to reduce costs. When  $\alpha$  approximates 0, the problem focuses solely on optimizing the objective function, and the cost can be neglected (if there are multiple optimal subsets, the focus is on finding the one with the smallest cost among them). In contrast, as  $\alpha$  approaches 1, the cost is optimized more effectively, but the quality of the objective function solution can decrease with  $T$  given. Therefore, the goal of optimizing both regrets is to find a balanced solution with the parameter  $\alpha$ . This makes our framework more generalized and applicable to a broader range of real-world scenarios. Furthermore, our regret analysis is more thorough and better aligned with the goals of BOCVS compared to previous work. Tay et al. (2023) define the cost-varying cumulative regret for the BOCVS problem as follows:

$$R_T := \sum_{t=1}^T c_{i_t} \left( \mathbb{E} \left[ f(\mathbf{x}^{i_t^*}, \mathbf{X}^{-i_t^*}) \right] - \mathbb{E} \left[ f(\mathbf{x}^{i_t}, \mathbf{X}^{-i_t}) \right] \right)$$

In the case where  $\alpha = 0$ , if an algorithm achieves sublinear  $R_T^f$ , it will also achieve sublinear  $R_T$ . However, the reverse is not true. When  $\alpha = 0$  and multiple optimal subsets exist, an algorithm that achieves sublinear  $R_T$  may not simultaneously achieve sublinear  $R_T^f$  and  $R_T^c$ . Analyzing both types of regret provides a more comprehensive assessment of the algorithm's performance and its cost management.

## Method

In this section, we introduce an explore-then-commit algorithm to address Bayesian Optimization with unknown cost-varying subsets. At the high-level idea, this algorithm initially explores each subset for a predetermined number of rounds before transitioning to the exploitation phase. During the exploitation phase, the algorithm identifies a feasible set of subsets based on the upper and lower confidence bounds (UCB, LCB) of their maximum values in each round. This set includes all subsets where the upper confidence bound of their maximum expected values exceeds the minimum acceptable value. From this feasible set, the subset with the lowest estimated cost is selected for exploitation. The detailed procedure is outlined in Algorithm 1. A crucial part of this algorithm is determining the appropriate number of exploration rounds before moving to the exploitation phase, which will be discussed in the next section.

At iteration  $t$ , the Gaussian Process (GP) posterior estimate of  $f$  is used to define an upper confidence bound  $u_{t-1}$  and a lower confidence bound for  $l_{t-1}$  for  $f$ , represented as:

$$u_{t-1}(\mathbf{x}) = \mu_{t-1}(\mathbf{x}) + \beta_t \sigma_{t-1}(\mathbf{x}),$$

$$l_{t-1}(\mathbf{x}) = \mu_{t-1}(\mathbf{x}) - \beta_t \sigma_{t-1}(\mathbf{x}),$$

where  $\mu_{t-1}(\mathbf{x})$  is the posterior mean,  $\sigma_{t-1}(\mathbf{x})$  is the posterior standard deviation, and  $\beta_t$  is a parameter controlling the exploration-exploitation trade-off. The goal of the exploration phase is to filter out low-quality control subsets by setting an initial threshold for the exploitation phase. This is achieved by defining the intersected lower confidence bound

$\overline{lc}b$  for  $\mathbb{E} \left[ f \left( \left[ \mathbf{x}^{i^+}, \mathbf{X}^{-i^+} \right] \right) \right]$  and the intersected upper confidence bound  $\overline{uc}b_i$  for  $\max_{\mathbf{x}^i \in \mathcal{X}^i} \mathbb{E} \left[ f \left( \left[ \mathbf{x}^i, \mathbf{X}^{-i} \right] \right) \right]$  in lines 10 and 11. High-quality control sets  $i$  must meet the condition  $\overline{uc}b_i \geq (1 - \alpha)\overline{lc}b$ , forming the set  $\mathcal{S}_1$  that includes only the high-quality subsets. Additionally, during the exploitation phase, the intersected confidence bounds  $\overline{lc}b$  and  $\overline{uc}b_i$  are updated for each control subset  $i$  in lines 14, 15, 18, and 19. This update mechanism ensures that, with high probability, the set  $\mathcal{S}_1$  gradually narrows over time while always containing the optimal set  $\mathcal{C}_*$ . As a result, the condition  $\mathcal{S}_1 = \emptyset$  in line 17 is likely to occur with low probability. However, if this condition does arise, the intersected confidence bounds is updated so that  $\mathcal{S}_1$  is revised in line 20 to ensure it is not empty. Without considering cost regret, selecting any subset within  $\mathcal{S}_1$  would provide satisfactory performance by minimizing quality regret. However, to optimize both cost and quality regret, it is essential to distinguish between the subsets and identify the one with the lowest cost in set  $\mathcal{S}_1$ . Since the exact cost is unknown, it is approximated using  $c_i^{\text{LCB}}$ , as defined in line 24. This approximation is then used to balance between exploring various subsets and exploiting the currently cheapest subset. Finally, the promising subset is selected, and the values in the control partial query are specified in lines 25 and 26.

## Main Theorem

**Theorem 1.** Assume that the distribution of the random cost of each subset has support  $[0, 1]$ ,  $B$  is the upper bound of the RKHS norm of  $f$ ,  $M^f = \max_{\mathbf{x} \in \mathcal{X}} \|f_{\mathbf{x}}\|$ . With probability at least  $1 - 3\delta - 2/T^2$ , Alg. 1 incurs an objective cumulative regret and a cost cumulative regret bounded by

$$\begin{aligned} R_T^f &= \mathcal{O}((m-1)\tau M^f) + \\ &\quad \mathcal{O}\left(\beta_t \left( \sqrt{4(T+2-m\tau)\gamma_{T-m\tau}(\mathcal{X})} + m \left( \log \frac{m}{\delta} \right) \right)\right) \\ &\quad + \mathcal{O}\left(\frac{T-(m-1)\tau}{\tau} \beta_{m\tau} \left( \sqrt{\tau\gamma_{\tau}(\mathcal{X})} + \log \frac{m}{\delta} \right)\right) \\ R_T^c &= \mathcal{O}(m\tau) + \mathcal{O}\left((T-m\tau)\sqrt{\frac{2\log T}{\tau}}\right) \end{aligned}$$

by setting  $\beta_t = B + \sigma\sqrt{2(\gamma_{t-1}(\mathcal{X}) + 1 + \log(1/\delta))}$ .

For any appropriately chosen kernel such that  $\gamma_T(\mathcal{X}) \leq \mathcal{O}(T^{\frac{1}{2}-\epsilon})$  where  $\epsilon > 0$  (e.g., commonly used squared exponential kernel) and we choose  $\tau$  satisfying that  $\mathcal{O}(T^{\epsilon_0}) \leq \tau < \mathcal{O}(T)$  where  $\epsilon_0 > 0$ , Theorem 1 implies that both the quality regret  $R_T^f$  and the cost regret  $R_T^c$  will be sublinear in  $T$ . The best control set and specified values in the control partial query in the algorithm's choices eventually converge to the cheapest acceptable solution. This ensures the algorithm still achieves good performance while still managing costs effectively. The main idea of the theorem's proof is based on analyzing the components in different phases of the regrets.

In case we explore subsets with a sufficiently large number of evaluations  $\tau$ , we can be assured that with high probability we will only focus on subsets in the set  $\mathcal{C}_*$ . We can describe this property by the following lemma:

---

**Algorithm 1: Proposed method**


---

1: **Input:**  $T$  evaluations, GP with kernel  $k$ , control sets  $\mathcal{I}$ , costs  $(c_i)_{i=1}^m$ ,  $\tau$ , tolerated parameter  $\alpha$

2: **Exploration phase:**

3: **for** for  $t \in [1, m\tau]$  **do**

4:  $i := t \bmod m$ ;

5:  $\mathbf{x}_t^i := \arg \max_{\mathbf{x}^i \in \mathcal{X}^i} \mathbb{E} [u_{t-1}([\mathbf{x}^i, \mathbf{X}^{-i}])]$

6: Observe  $\mathbf{x}^{-it}$  drawn from  $\mathbb{P}^{-it}$

7: Observe  $y_t := f(\mathbf{x}_t^i) + \xi_t$  and  $c_i^{(t)}$

8:  $T_i(t+1) := T_i(t) + \mathbf{1}\{I_t = i\} \forall i \in [m]$

9: **end for**

10:  $\overline{lcb} := \max_{(i, \mathbf{x}^i) \in [m] \times \mathcal{X}^i, t \leq m\tau} \mathbb{E} [l_{t-1}([\mathbf{x}^i, \mathbf{X}^{-i}])]$

11:  $\overline{ucb}_i := \min_{k \leq m} \max_{\mathbf{x}^i \in \mathcal{X}^i} \mathbb{E} [u_{k\tau+i-1}([\mathbf{x}^i, \mathbf{X}^{-i}])], \forall i \in [m]$

12: **Exploitation phase:**

13: **for** iteration  $t = m\tau + 1$  to  $T$  **do**

14:  $\overline{lcb} := \max \left\{ \overline{lcb}, \max_{(i, \mathbf{x}^i) \in [m] \times \mathcal{X}^i} \mathbb{E} [l_{t-1}([\mathbf{x}^i, \mathbf{X}^{-i}])] \right\}$

15:  $\overline{ucb}_i := \min \left\{ \overline{ucb}_i, \max_{\mathbf{x}^i \in \mathcal{X}^i} \mathbb{E} [u_{t-1}([\mathbf{x}^i, \mathbf{X}^{-i}])] \right\},$   
 $\forall i \in [m]$

16:  $\mathcal{S}_1 := \{i \in [m] \mid \overline{ucb}_i > (1 - \alpha)\overline{lcb}\}$

17: **if**  $\mathcal{S}_1 = \emptyset$  **then**

18:  $\overline{lcb} := \max_{(i, \mathbf{x}^i) \in [m] \times \mathcal{X}^i} \mathbb{E} [l_{t-1}([\mathbf{x}^i, \mathbf{X}^{-i}])]$

19:  $\overline{ucb}_i := \max_{\mathbf{x}^i \in \mathcal{X}^i} \mathbb{E} [u_{t-1}([\mathbf{x}^i, \mathbf{X}^{-i}])], \forall i \in [m]$

20:  $\mathcal{S}_1 := \{i \in [m] \mid \overline{ucb}_i > (1 - \alpha)\overline{lcb}\}$

21: **end if**

22:  $\hat{c}_i(t) := \frac{\sum_{k=1}^{t-1} c_i^{(k)} \mathbf{1}\{I_k=i\}}{T_i(t)} \forall i \in [m];$

23:  $\beta_i(t) := \sqrt{\frac{2 \log T}{T_i(t)}} \forall i \in [m];$

24:  $c_i^{\text{LCB}}(t) := \max \{\hat{c}_i(t) - \beta_i(t), 0\} \forall i \in [m]$

25:  $\mathcal{S}_3 := \{i \in \mathcal{S}_1 \mid c_i^{\text{LCB}} = \min_{j \in \mathcal{S}_1} c_j^{\text{LCB}}\}$

26:  $(i_t, \mathbf{x}^{it}) := \operatorname{argmax}_{(i, \mathbf{x}^i) \in \mathcal{S}_3 \times \mathcal{X}^i} \mathbb{E} [u_{t-1}([\mathbf{x}^i, \mathbf{X}^{-i}])]$

27: Observe  $\mathbf{x}^{-it}$  drawn from  $\mathbb{P}^{-it}$

28: Observe  $y_t := f(\mathbf{x}_t^{it}) + \xi_t$  and  $c_{i_t}^{(t)}$

29: Update  $\mathcal{D}_t := \{(\mathbf{x}_\tau, y_\tau)\}_{\tau=1}^t$

30: **end for**

31: **return**  $\mathcal{D}_t$

---

**Lemma 2.** *If there exists a  $\tilde{\epsilon} > 0$  s.t. for all  $i' \notin \mathcal{C}_*$ ,  $\mathbb{E} [f([\mathbf{x}^{i'}, \mathbf{X}^{-i'}])] \leq (1 - \alpha) \max_{\mathbf{x}^i \in \mathcal{X}^i} \mathbb{E} [f([\mathbf{x}^i, \mathbf{X}^{-i}])] - \tilde{\epsilon}$ , and  $\gamma_T(\mathcal{X}) < \mathcal{O}(\sqrt{T})$ , with probability at least  $1 - \delta$ , if  $\tau \geq \min \left\{ t : \beta_{tm} \frac{(2\sqrt{4(t+2)}\gamma_t(\mathcal{X}) + 4 \log \frac{m}{\delta} + 8 \log(4) + 1)}{t} \leq \frac{\tilde{\epsilon}}{2 - \sigma} \right\}$ , then  $\mathcal{S}_1 = \mathcal{C}_*$  for all iteration in Exploitation phase.*

This lemma's proof is perhaps the trickiest and crucially relies on  $\tau$  and distance gap  $\tilde{\epsilon}$ . From Lemma 2, we can improve the upper bound of  $R_T^f$  in theorem 1 as follows:

**Theorem 3.** *Suppose that the assumption of both Theorem 1 and Lemma 2 hold, with probability at least  $1 - 3\delta - 2/T^2$ ,*

*Alg. 1 incurs an objective cumulative regret and a cost cumulative regret bounded by*

$$R_T^f = \mathcal{O}(m\tau M^f) + \mathcal{O}\left(\beta_t \left(\sqrt{4(T+2-m\tau)\gamma_{T-m\tau}(\mathcal{X})} + m(\log(m/\delta))\right)\right)$$

$$R_T^c = \mathcal{O}(m\tau) + \mathcal{O}\left((T-m\tau)\sqrt{2\log(T/\tau)}\right)$$

**Practical Considerations** Theorems 1 and 3 give us an idea of how much evaluation is needed during the exploration phase depending on  $T$  and  $\tilde{\epsilon}$  for the algorithm to achieve the best performance. However, in practice, we are only given a budget without knowing these two parameters in advance. Therefore, we suggest using a larger budget for the exploration phase than for the exploitation phase. Moreover, if the application does not specify the tolerated parameter  $\alpha$ , we can allow it to decrease as the remaining budget nears zero. This gradual reduction allows us to work with many subsets that are sufficient to increase information gain while still ensuring low costs, before moving on to focus on playing better but more expensive subsets.

## Experiments

This section empirically evaluates the performance of the tested algorithms with 4 objective functions: (a) Hartmann ( $d = 6$ ) and (b) Ackley ( $d = 6$ ) as the synthetic benchmark functions, adding 6 unrelated variables, which leads to the existence of many subsets containing optimal points, (c) a plant growth simulator built from real-world data where the variables are nutrients such as  $\text{NH}_3$  and  $\text{pH}(5-D)$ , and (d) a simulator built from the airfoil self-noise dataset (5-D) from the UCI Machine Learning Repository (Dua, Graff et al. 2017). For the first two objective functions, there are 7 control sets, of which 2 are optimal sets: the full set and the subset containing valid variables. For the plant growth objective function, we pick 7 control sets including the full query control set. For the airfoil self-noise objective function, similar to that of (Hayashi, Honda, and Kashima 2022a), we pick 7 control sets of 2 variables each that are not subsets of each other. We use 2 different sets of means of costs  $c$  for the 7 control sets: cheap ( $\{0.01, 0.01, 0.01, 0.1, 0.1, 0.1, 1\}$ ), moderate ( $\{0.1, 0.1, 0.1, 0.2, 0.2, 0.2, 1\}$ ) and add Gaussian noises  $\mathcal{N}(0, 0.02)$  to the cost observation  $c_{i_t}^{(t)}$  of the subset  $i_t$  if  $c_{i_t} \geq 0.1$ . Using these sets of costs, the control sets are ordered such that  $\mathcal{I}_i \subset \mathcal{I}_j \Rightarrow c_i < c_j$ . While these cost sets may seem arbitrary, it is the algorithms' relative performance across these cost sets rather than the absolute performance on a single cost set that allows us to understand the conditions under which particular algorithms perform better or worse. Real-world applications will come with their own cost sets defined by real-world constraints. Every probability distribution  $\mathcal{P}_\ell$  is a truncated normal distribution with mean 0.5 and the same variance which is one of 0.02, 0.04 (the uniform distribution on  $[0, 1]$  has variance  $1/12$ ). We set the maximum budget for any experiment to 100 units. We compare the performance of our algorithm against that of the baseline Thompson sampling (TS-PSQ) and UCB-PSQ de-

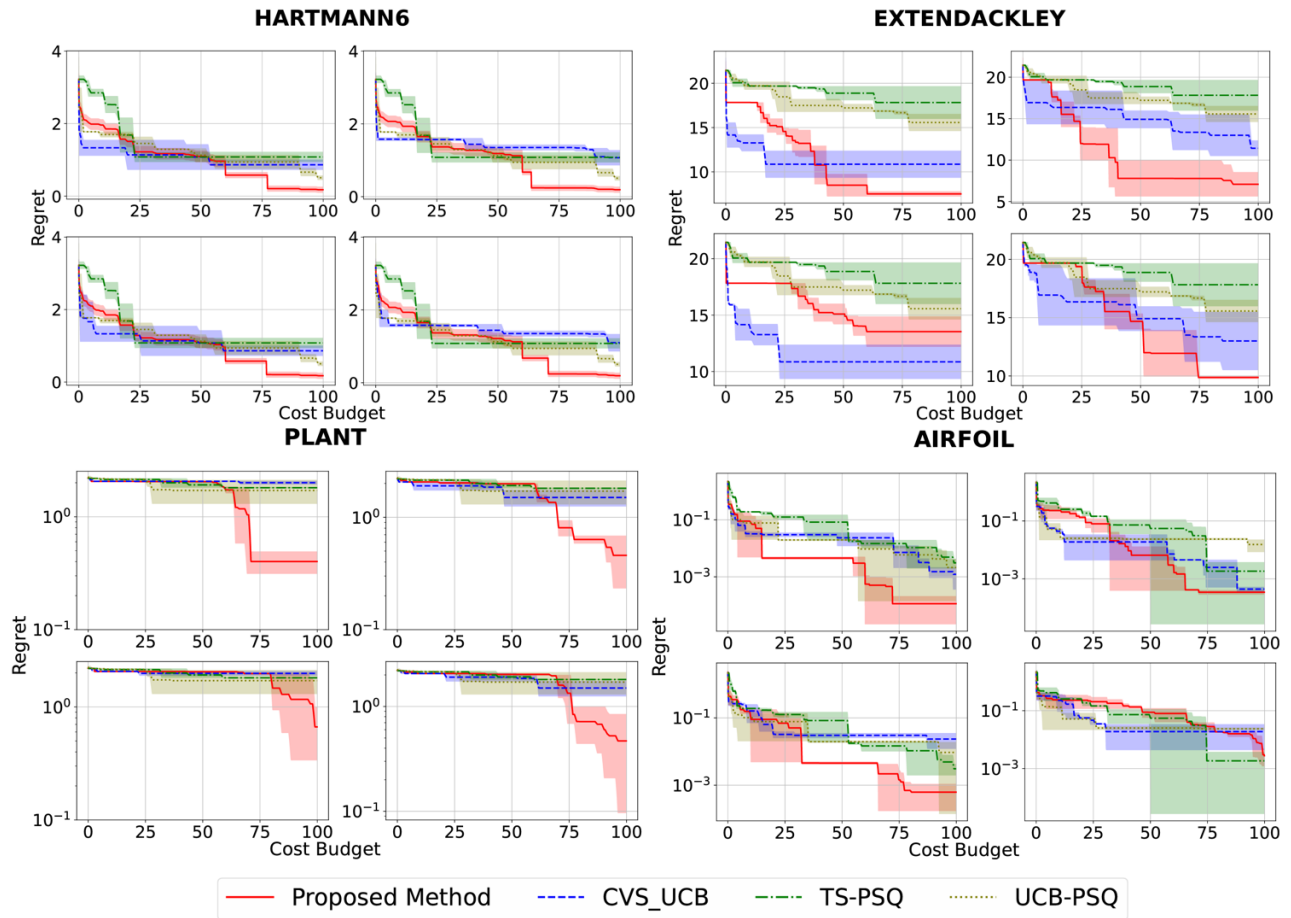


Figure 1: Comparison with the BO baselines. Each function is represented in a set of subplots, where the variance of the variable distributions increases from left to right, and the cost set grows from top to bottom.

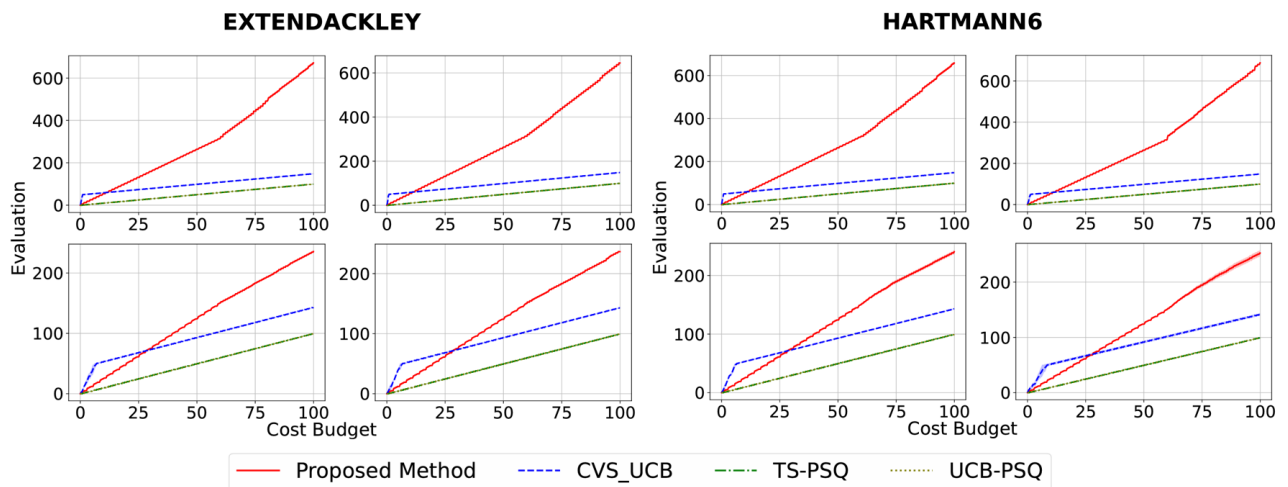


Figure 2: Number of evaluation against cost spent.

veloped in (Hayashi, Honda, and Kashima 2022a). We evaluate a variant of UCB-CVS (Tay et al. 2023) called ETC-50,

which involves 50 plays per group, with each group containing subsets having the same number of variables. This ap-

proach differs slightly from (Tay et al. 2023), where groups are defined by subsets having the same cost due to the unknown cost setting. In the proposed method, we spend 60 units of cost for the exploration phase. The parameter  $\alpha$  at the beginning of the exploitation phase is set to 0.1 and is halved after  $d$  function evaluations.

Figure 1 shows the mean and standard error of the simple regret  $\min_{1 \leq t \leq \mathcal{T}(C)} \mathbb{E} \left[ f \left( \left[ \mathbf{x}^{i^+}, \mathbf{X}^{-i^+} \right] \right) \right] - \mathbb{E} \left[ f \left( \left[ \mathbf{x}^{i_t}, \mathbf{X}^{-i_t} \right] \right) \right]$  (lower is better) incurred against cost spent (budget)  $C$  by each algorithm with varying objective functions, cost sets, and variances of distributions where  $\mathcal{T}(C)$  denotes the maximum iteration reached after spending  $C$ . The simple regret encodes the value of the best solution an algorithm has chosen within a certain budget and is a measure of cost efficiency. We observed that the naive methods are often less effective than the proposed method. The efficiency of the proposed algorithm stems from its strategy of prioritizing more plays on cost-effective feasible subsets, rather than solely targeting subsets likely to contain optimal solutions that are very expensive. Figure 2 compares the number of evaluations by the algorithms relative to the costs spent. The results indicate that the proposed algorithm conducts significantly more evaluations than the other algorithms once the exploitation phase starts. This implies that the baseline algorithms consistently evaluate the entire set, which is costly, rather than using the subset of valid variables, which is not only optimal but also less expensive. As a result, it leads to a lower number of evaluations compared to the proposed algorithm.

## Related Work

To the best of our knowledge, UCB-CVS (Tay et al. 2023) is the only work that addresses the BOCVS setting at the time of writing. UCB-CVS uses an  $\epsilon$ -schedule as a relaxation technique, allowing cheaper control query sets to be selected for exploitation. The final control set is then chosen from this relaxed set by comparing their known costs. While UCB-CVS achieves a sublinear rate for cost-varying cumulative regret, it does not guarantee the minimization of the total incurred cost. Additionally, the practicality of the  $\epsilon$ -schedule is questionable due to the difficulty in selecting an appropriate value. Our proposed method addresses these issues, even in settings with unknown costs, making it more applicable to real-world problems.

The BOCVS setting can be viewed as a specific instance of the Causal Bayesian Optimization (CBO) framework (Aglietti et al. 2020), where the causal relationships among input variables are incorporated into the objective functions. In CBO, a subset of variables is selected for intervention, and their values are set to achieve optimal results while minimizing the overall intervention cost. Depending on the causal structure, intervening on certain variables may lead to a propagation of effects in the causal graph, affecting both the outcomes and associated costs. Therefore, the causal relationships in CBO can be seen as a generalized representation of input distribution in the BOCVS setting. CBO is a challenging problem, and to the best of our knowledge, most previous works (Aglietti et al. 2021, 2023; Gultchin

et al. 2023; Branchini et al. 2023) have not succeeded in establishing useful theoretical regret bounds. The only work addressing regret bounds in CBO is Sussex, Makarova, and Krause (2023), which introduces such bounds under additional assumptions. However, their theoretical analysis does not account for minimizing the total incurred cost.

In BO field, there are also some settings where the learner is not only concerned about optimizing the objective function but also considers the trade-offs associated with multiple objectives that are intrinsic to several practical applications. Hayashi, Honda, and Kashima (2022b) introduced the setting *with partially specified queries* (BOPSQ) where the learner simultaneously selects a control set and specifies their values to maximize the objective function. Hence, the learner is required to consider the extent to which the query variables contribute to the function’s output as well as which values the unspecified input variables will take. The key distinction between BOPSQ and the BOCVS setting is that BOPSQ does not incorporate cost constraints, which can significantly influence the learner’s decisions. On the other hand, multi-fidelity Bayesian Optimization (MFBO) (Kandasamy et al. 2016; Poloczek, Wang, and Frazier 2017; Takeno et al. 2020) addresses scenarios where inexpensive, low-fidelity approximations of the true objective function are available. Each low-fidelity approximation is associated with a different level of model discrepancy and querying cost, with larger discrepancies generally resulting in lower costs. The goal of MFBO is to achieve an optimal objective value while minimizing the total cost of queries. Although the principle of paying less for potentially less informative queries is conceptually similar, the BOCVS setting differs in that the uncertainty arises from the randomness of inputs not included in the control set, rather than from querying a lower-cost approximation as in MFBO.

BOCVS can also be viewed as a specialized version of multi-armed bandits (MAB) with cost constraints (Cayci, Eryilmaz, and Srikant 2020; Ding et al. 2013; Sinha et al. 2021). In this context, each control set functions as an arm, and pulling an arm yields a reward along with an associated cost. The goal of the algorithm is to select a sequence of arms to maximize the expected total reward while ensuring that the costs of pulling these arms stay within a budget constraint. However, unlike in the standard cost-constrained MAB setting, BOCVS requires the learner not only to pull an arm but also to specify parameters, such as the values for the control set associated with that arm. This simultaneous decision-making process adds a layer of complexity, making BOCVS more challenging than traditional cost-constrained MAB problems.

## Conclusion

We introduce an extension of the BOCVS problem where the costs of subsets are random and unknown, and we develop a cost-sensitive algorithm with theoretical guarantees, demonstrating sub-linear cumulative regret for both objective function value and cost. Our empirical evaluation compared the proposed algorithm’s performance against baseline methods across various experimental conditions, including both synthetic and real-world datasets.

## References

- Aglietti, V.; Dhir, N.; González, J.; and Damoulas, T. 2021. Dynamic Causal Bayesian Optimization. In Ranzato, M.; Beygelzimer, A.; Dauphin, Y. N.; Liang, P.; and Vaughan, J. W., eds., *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021, December 6-14, 2021, virtual*, 10549–10560.
- Aglietti, V.; Lu, X.; Paleyes, A.; and González, J. 2020. Causal Bayesian Optimization. In Chiappa, S.; and Calandra, R., eds., *The 23rd International Conference on Artificial Intelligence and Statistics, AISTATS 2020, 26-28 August 2020, Online [Palermo, Sicily, Italy]*, volume 108 of *Proceedings of Machine Learning Research*, 3155–3164. PMLR.
- Aglietti, V.; Malek, A.; Ktena, I.; and Chiappa, S. 2023. Constrained Causal Bayesian Optimization. In Krause, A.; Brunskill, E.; Cho, K.; Engelhardt, B.; Sabato, S.; and Scarlett, J., eds., *International Conference on Machine Learning, ICML 2023, 23-29 July 2023, Honolulu, Hawaii, USA*, volume 202 of *Proceedings of Machine Learning Research*, 304–321. PMLR.
- Branchini, N.; Aglietti, V.; Dhir, N.; and Damoulas, T. 2023. Causal Entropy Optimization. In Ruiz, F. J. R.; Dy, J. G.; and van de Meent, J., eds., *International Conference on Artificial Intelligence and Statistics, 25-27 April 2023, Palau de Congressos, Valencia, Spain*, volume 206 of *Proceedings of Machine Learning Research*, 8586–8605. PMLR.
- Cayci, S.; Eryilmaz, A.; and Srikant, R. 2020. Budget-Constrained Bandits over General Cost and Reward Distributions. In Chiappa, S.; and Calandra, R., eds., *The 23rd International Conference on Artificial Intelligence and Statistics, AISTATS 2020, 26-28 August 2020, Online [Palermo, Sicily, Italy]*, volume 108 of *Proceedings of Machine Learning Research*, 4388–4398. PMLR.
- Ding, W.; Qin, T.; Zhang, X.; and Liu, T. 2013. Multi-Armed Bandit with Budget Constraint and Variable Costs. In desJardins, M.; and Littman, M. L., eds., *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence, July 14-18, 2013, Bellevue, Washington, USA*, 232–238. AAAI Press.
- Dua, D.; Graff, C.; et al. 2017. UCI machine learning repository, 2017. URL <http://archive.ics.uci.edu/ml>, 7(1): 62.
- Gultchin, L.; Aglietti, V.; Bellot, A.; and Chiappa, S. 2023. Functional causal Bayesian optimization. In Evans, R. J.; and Shpitser, I., eds., *Uncertainty in Artificial Intelligence, UAI 2023, July 31 - 4 August 2023, Pittsburgh, PA, USA*, volume 216 of *Proceedings of Machine Learning Research*, 756–765. PMLR.
- Hayashi, S.; Honda, J.; and Kashima, H. 2022a. Bayesian optimization with partially specified queries. *Machine Learning*, 1–30.
- Hayashi, S.; Honda, J.; and Kashima, H. 2022b. Bayesian optimization with partially specified queries. *Mach. Learn.*, 111(3): 1019–1048.
- Kandasamy, K.; Dasarathy, G.; Oliva, J. B.; Schneider, J. G.; and Póczos, B. 2016. Gaussian Process Bandit Optimisation with Multi-fidelity Evaluations. In Lee, D. D.; Sugiyama, M.; von Luxburg, U.; Guyon, I.; and Garnett, R., eds., *Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, December 5-10, 2016, Barcelona, Spain*, 992–1000.
- Kirschner, J.; Mutný, M.; Krause, A.; Coello de Portugal, J.; Hiller, N.; and Snuverink, J. 2022. Tuning particle accelerators with safety constraints using Bayesian optimization. *Phys. Rev. Accel. Beams*, 25: 062802.
- Kruger, G. H.; Shih, A. J.; Hattingh, D. G.; and van Niekerk, T. I. 2011. Intelligent machine agent architecture for adaptive control optimization of manufacturing processes. *Advanced Engineering Informatics*, 25(4): 783–796. Special Section: Advances and Challenges in Computing in Civil and Building Engineering.
- Poloczek, M.; Wang, J.; and Frazier, P. I. 2017. Multi-Information Source Optimization. In Guyon, I.; von Luxburg, U.; Bengio, S.; Wallach, H. M.; Fergus, R.; Vishwanathan, S. V. N.; and Garnett, R., eds., *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*, 4288–4298.
- Shields, B. J.; Stevens, J. M.; Li, J.; Parasram, M.; Damani, F. N.; Alvarado, J. I. M.; Janey, J. M.; Adams, R. P.; and Doyle, A. G. 2021. Bayesian reaction optimization as a tool for chemical synthesis. *Nat.*, 590(7844): 89–96.
- Sinha, D.; Sankararaman, K. A.; Kazerouni, A.; and Avadhanula, V. 2021. Multi-Armed Bandits with Cost Subsidy. In Banerjee, A.; and Fukumizu, K., eds., *The 24th International Conference on Artificial Intelligence and Statistics, AISTATS 2021, April 13-15, 2021, Virtual Event*, volume 130 of *Proceedings of Machine Learning Research*, 3016–3024. PMLR.
- Slivkins, A.; et al. 2019. Introduction to multi-armed bandits. *Foundations and Trends® in Machine Learning*, 12(1-2): 1–286.
- Snoek, J.; Larochelle, H.; and Adams, R. P. 2012. Practical Bayesian Optimization of Machine Learning Algorithms. In Pereira, F.; Burges, C.; Bottou, L.; and Weinberger, K., eds., *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc.
- Srinivas, N.; Krause, A.; Kakade, S. M.; and Seeger, M. 2009. Gaussian process optimization in the bandit setting: No regret and experimental design. *arXiv preprint arXiv:0912.3995*.
- Sussex, S.; Makarova, A.; and Krause, A. 2023. Model-based Causal Bayesian Optimization. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net.
- Takeno, S.; Fukuoka, H.; Tsukada, Y.; Koyama, T.; Shiga, M.; Takeuchi, I.; and Karasuyama, M. 2020. Multi-fidelity Bayesian Optimization with Max-value Entropy Search and its Parallelization. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event*, volume 119 of *Proceedings of Machine Learning Research*, 9334–9345. PMLR.

Tay, S.; Foo, C. S.; Urano, D.; Leong, R.; and Low, B. K. H. 2023. Bayesian Optimization with Cost-varying Variable Subsets. In Oh, A.; Naumann, T.; Globerson, A.; Saenko, K.; Hardt, M.; and Levine, S., eds., *Advances in Neural Information Processing Systems*, volume 36, 3008–3031. Curran Associates, Inc.

Williams, C. K.; and Rasmussen, C. E. 2006. *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA.