Stochastic Bayesian Optimization with Unknown Continuous Context Distribution via Kernel Density Estimation

Xiaobin Huang, Lei Song, Ke Xue, Chao Qian*

National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China
School of Artificial Intelligence, Nanjing University, Nanjing 210023, China
{huangxb, songl, xuek, qianc}@lamda.nju.edu.cn

Abstract

Bayesian optimization (BO) is a sample-efficient method and has been widely used for optimizing expensive black-box functions. Recently, there has been a considerable interest in BO literature in optimizing functions that are affected by context variable in the environment, which is uncontrollable by decision makers. In this paper, we focus on the optimization of functions' expectations over continuous context variable, subject to an unknown distribution. To address this problem, we propose two algorithms that employ kernel density estimation to learn the probability density function (PDF) of continuous context variable online. The first algorithm is simpler, which directly optimizes the expectation under the estimated PDF. Considering that the estimated PDF may have high estimation error when the true distribution is complicated, we further propose the second algorithm that optimizes the distributionally robust objective. Theoretical results demonstrate that both algorithms have sub-linear Bayesian cumulative regret on the expectation objective. Furthermore, we conduct numerical experiments to empirically demonstrate the effectiveness of our algorithms.

1 Introduction

Bayesian optimization (BO) (Shahriari et al. 2015; Frazier 2018) is a popular and sample-efficient method for optimizing expensive black-box functions. BO has shown excellent performance in various fields, such as chemical molecular design (Gómez-Bombarelli et al. 2018; Griffiths and Hernández-Lobato 2020), neural architecture search (Klein et al. 2017; Kandasamy et al. 2018b; Song et al. 2022), and hyper-parameter tuning (Chen et al. 2018; Qian, Xiong, and Xue 2020). The typical process of BO involves approximating the objective function by a Gaussian process (GP) surrogate model (Rasmussen and Williams 2006), and then selecting the most valuable point for evaluation by optimizing an acquisition function based on the posterior of the surrogate model.

In some practical scenarios, uncontrollable context variable from the environment can impact the objective function, such as customer demand in inventory management (Dai, Chen, and Birge 2000; Hannah, Powell, and Blei 2010), bid-ask spread and borrowing cost in portfolio optimization (Boyd et al. 2017; Cakmak et al. 2020), and temperature in crop size optimization (Tay et al. 2022). There has been some BO literature taking context variable into account with different objectives. Robust optimization aims to find solutions that perform well in the worst-case scenario (Marzat, Walter, and Piet-Lahanier 2013; Bogunovic et al. 2018), while stochastic optimization (SO) focuses on finding solutions that perform well in expectation (Williams 2000; Xie et al. 2012; Beland and Nair 2019; Kirschner and Krause 2019; Toscano-Palmerin and Frazier 2022). Risk optimization considers risk measures such as mean-variance (Iwazaki, Inatsu, and Takeuchi 2021), value at risk (VaR) (Cakmak et al. 2020; Nguyen et al. 2021) or conditional VaR (Cakmak et al. 2020). However, robust optimization ignores the distribution information of the context, and most existing works on SO and risk optimization assume that the distribution of context is known.

Distributionally robust BO (DRBO) (Kirschner et al. 2020; Nguyen et al. 2020; Inatsu et al. 2022; Tay et al. 2022) is proposed to address problems with unknown context distribution by optimizing the worst expectation over a set of distributions. Existing DRBO works focus on finite context. When the context variable is in a continuous space, which is common (e.g., temperature in crop size optimization, and energy output in wind power prediction (Tay et al. 2022)) in practice, they usually discretize the space. However, the computational complexity of the inner convex optimization in DRBO is at least cubed to the size $|\mathcal{C}|$ of context space $\mathcal{C}$ (Tay et al. 2022). Smaller $|\mathcal{C}|$ leads to poor approximation to the expectation over the continuous space, thus poor performance, while larger $|\mathcal{C}|$ leads to unacceptable computational complexity. While Tay et al. (2022) have developed a method based on fast worst case sensitivity to efficiently approximate and accelerate the inner optimization problem, it suffers from linear regret due to approximation errors. Nguyen et al. (2020) used Lagrange multipliers to accelerate the optimization, but it is limited to the simulator setting, where the decision makers can select the context.

In this paper, we consider the problem of maximizing the SO objective $\max_{\boldsymbol{x}} \mathbb{E}_{\boldsymbol{c}\sim p(\boldsymbol{c})}[f(\boldsymbol{x},\boldsymbol{c})]$ over the decision variable $\boldsymbol{x} \in \mathcal{X}$, where f is a black-box function, and the distribution p of context variable $\boldsymbol{c} \in \mathcal{C}$ is continuous and unknown. The context is observable after making decisions.

*Chao Qian is the corresponding author.

To avoid the drawbacks of discretization, we propose two algorithms to directly address this problem. The first algorithm employs kernel density estimation (KDE) to estimate the unknown context distribution and maximizes the objective function f's expectation under the estimated PDF, which is simple and time-efficient. Considering that the estimated PDF may have high estimation error when the true distribution is complicated, we propose the second algorithm, which also uses KDE for PDF estimation but maximizes a distributionally robust objective function, i.e., optimizes the worst-case expectation across a set of distributions around the estimated one. We provide theoretical analyses for both algorithms, proving that they have sub-linear Bayesian cumulative regret on the SO objective. The experiments on synthetic functions and two real-world optimization tasks (i.e., newsvendor problem and portfolio optimization) demonstrate that our proposed algorithms achieve better performance.

2 Background

2.1 Bayesian Optimization

BO is a popular algorithm for black-box optimization, which consists of two main components: a surrogate model and an acquisition function. GP (Rasmussen and Williams 2006) is the most commonly used surrogate model. The function f is assumed to be a sample path from a GP, denoted as $\mathcal{GP}(\mathbf{0}, k(\cdot,\cdot))$, where $\mathbf{0}$ is the prior mean and $k(\cdot,\cdot)$ is a kernel function. Given observed data set $\mathcal{D}_{t-1} = \{(\boldsymbol{x}_i, \boldsymbol{c}_i, y_i)\}_{i=1}^{t-1}$, where $y_i = f(\boldsymbol{x}_i, \boldsymbol{c}_i) + \epsilon_i$ is the noisy observation and $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$, we can calculate the posterior distribution of the function $f \mid \mathcal{D}_{t-1} \sim \mathcal{GP}(\mu_t(\boldsymbol{x}, \boldsymbol{c}), k_t((\boldsymbol{x}, \boldsymbol{c}), (\boldsymbol{x}', \boldsymbol{c}')))$, where the posterior mean $\mu_t(\boldsymbol{x}, \boldsymbol{c}) = \boldsymbol{k}_{t-1}(\boldsymbol{x}, \boldsymbol{c})^\top (\mathbf{K}_{t-1} + \sigma^2\mathbf{I})^{-1} \boldsymbol{y}_{t-1}$, and posterior covariance $k_t((\boldsymbol{x}, \boldsymbol{c}), (\boldsymbol{x}', \boldsymbol{c}')) = k((\boldsymbol{x}, \boldsymbol{c}), (\boldsymbol{x}', \boldsymbol{c}')) - \boldsymbol{k}_{t-1}(\boldsymbol{x}, \boldsymbol{c})^\top (\mathbf{K}_{t-1} + \sigma^2\mathbf{I})^{-1} \boldsymbol{k}_{t-1}(\boldsymbol{x}', \boldsymbol{c}')$. Here, $\boldsymbol{k}_{t-1}(\boldsymbol{x}, \boldsymbol{c}) = [k((\boldsymbol{x}_i, \boldsymbol{c}_i), (\boldsymbol{x}, \boldsymbol{c}))]_{i=1,\dots,t-1}^\top$, $\mathbf{K}_{t-1} \in \mathbb{R}^{(t-1)\times(t-1)}$ is the positive semi-definite kernel matrix with $[\mathbf{K}_{t-1}]_{ij} = k((\boldsymbol{x}_i, \boldsymbol{c}_i), (\boldsymbol{x}_j, \boldsymbol{c}_j))$, and $\boldsymbol{y}_{t-1} = [y_1, \dots, y_{t-1}]^\top$.

Based on the posterior distribution obtained from GP, various acquisition functions can be used to determine the next query point, e.g., Probability of Improvement (PI) (Kushner 1964), Expected Improvement (EI) (Jones, Schonlau, and Welch 1998) and Upper Confidence Bound (UCB) (Srinivas et al. 2012). In this work, we use the UCB acquisition function for both our algorithms, defined as $\mathrm{ucb}_t(\boldsymbol{x}, \boldsymbol{c}) = \mu_t(\boldsymbol{x}, \boldsymbol{c}) + \sqrt{\beta_t}\sigma_t(\boldsymbol{x}, \boldsymbol{c})$, where β_t is a hyper-parameter to balance the exploitation and exploration. We use $\sigma_t^2(\boldsymbol{x}, \boldsymbol{c}) = k_t((\boldsymbol{x}, \boldsymbol{c}), (\boldsymbol{x}, \boldsymbol{c}))$ to represent the posterior variance at $(\boldsymbol{x}, \boldsymbol{c})$. Note that under deterministic environments, the context variable $\boldsymbol{c} \in \mathcal{C}$ can be neglected.

2.2 Stochastic Bayesian Optimization

In real-world problems, the objective function may be affected by context variable, which is uncontrollable by the decision makers. The problem can be formalized as a black-box function $f(\boldsymbol{x}, \boldsymbol{c})$ over a convex and compact domain $\mathcal{X} \times \mathcal{C} \subset \mathbb{R}^{D_x} \times \mathbb{R}^{D_c}$, where $\mathcal{X}$ is a D_x-dimensional decision space controlled by decision makers, and $\mathcal{C}$ is a D_c-dimensional context space controlled by environment. In this paper, we consider the setting that $\mathcal{C}$ is continuous. At iteration t, a decision $\boldsymbol{x}_t$ is made, followed by the observation of a context $\boldsymbol{c}_t \sim p(\boldsymbol{c})$ provided by the environment and observed by the decision maker. Note that the context distribution $p(\boldsymbol{c})$ is unknown here. Next, we observe the noisy evaluation $y_t = f(\boldsymbol{x}_t, \boldsymbol{c}_t) + \epsilon_t$, where $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$. We consider the SO setting aiming to identify the optimum $\boldsymbol{x}^* \in \arg\max_{\boldsymbol{x}\in\mathcal{X}} \mathbb{E}_{\boldsymbol{c}\sim p(\boldsymbol{c})}[f(\boldsymbol{x}, \boldsymbol{c})]$. Given the evaluation budget T, the goal is to minimize the cumulative regret of SO objective, i.e.,

$$R_T := \sum_{t=1}^{T} \left(\mathbb{E}_{\boldsymbol{c}\sim p(\boldsymbol{c})}[f(\boldsymbol{x}^*, \boldsymbol{c})] - \mathbb{E}_{\boldsymbol{c}\sim p(\boldsymbol{c})}[f(\boldsymbol{x}_t, \boldsymbol{c})] \right). \quad (1)$$

There has been plentiful research considering context variable in the BO literature. For instance, Krause and Ong (2011) considered the case that the context $\boldsymbol{c}_t$ is given before decision. When the context cannot be known beforehand, numerous approaches have been proposed with different optimization objectives.

Robust optimization considers a worst-case objective, formulated as $\max_{\boldsymbol{x}} \min_{\boldsymbol{c}\in\mathcal{C}} f(\boldsymbol{x}, \boldsymbol{c})$, and has been studied in (Marzat, Walter, and Piet-Lahanier 2013; Bogunovic et al. 2018). However, robust optimization is too pessimistic, ignoring the distribution information of context.

Stochastic optimization (SO) considers an average-case optimization, i.e., $\max_{\boldsymbol{x}} \mathbb{E}_{\boldsymbol{c}\sim p(\boldsymbol{c})}[f(\boldsymbol{x}, \boldsymbol{c})]$. A special case of SO is optimizing the expectation under input perturbation $\max_{\boldsymbol{x}} \mathbb{E}_{\boldsymbol{c}\sim p(\boldsymbol{c})}[f(\boldsymbol{x} \diamond \boldsymbol{c})]$, where $\diamond$ denotes the perturbation of the input $\boldsymbol{x}$ by $\boldsymbol{c}$, which has been discussed in (Nogueira et al. 2016; Beland and Nair 2019; Oliveira, Ott, and Ramos 2019; Fröhlich et al. 2020). The general case of SO has also been studied using different acquisition functions in (Williams 2000; Xie et al. 2012; Beland and Nair 2019; Kirschner and Krause 2019; Toscano-Palmerin and Frazier 2022), which, however, assume that the distribution of context is known.

Risk optimization uses risk measures, such as mean-variance (Iwazaki, Inatsu, and Takeuchi 2021), value at risk (VaR) (Cakmak et al. 2020; Nguyen et al. 2021) and conditional VaR (Cakmak et al. 2020), as the objectives when dealing with contextual uncertainty. For example, Cakmak et al. (2020) and Nguyen et al. (2021) studied $\mathrm{VaR}_\delta(\boldsymbol{x}) := \sup\{s : \mathbb{P}(f(\boldsymbol{x}, \boldsymbol{c}) \geq s) \geq 1 - \delta\}$, which measures the risk under a specified level of confidence $1 - \delta$.

Distributionally robust optimization (DRO). The above methods usually ignore the context distribution or assume the distribution of the context variable is known. When the distribution is unknown, the DRO objective can be adopted, considering the worst-case expectation over a set of distributions. The DRO objective is formulated as $\max_{\boldsymbol{x}} \min_{q\in\mathcal{Q}} \mathbb{E}_{\boldsymbol{c}\sim q(\boldsymbol{c})}[f(\boldsymbol{x}, \boldsymbol{c})]$, where $\mathcal{Q}$ is a given distribution set on the context space $\mathcal{C}$. Different approaches have been proposed to optimize the DRO objective. For instance, Nguyen et al. (2020) considered the simulator setting where the decision makers can select the context $\boldsymbol{c}$, while Inatsu et al. (2022) considered DRO under chance constraints. Kirschner et al. (2020) used maximum mean discrepancy (MMD) to construct the distribution set $\mathcal{Q}$ given a reference distribution. However, they only considered the case where the context space is discrete with a size of $|\mathcal{C}|$, and the inner optimization is a $|\mathcal{C}|$-dimensional optimization problem, which can be computationally expensive when the size $|\mathcal{C}|$ is large. Tay et al. (2022) proposed using worst-case sensitivity for approximation and acceleration of the inner optimization. However, the approximation error can lead to a decrease in the performance, and the regret bound they derived is linear even when the distribution distance $\epsilon_t = 0$ in Theorem 4 in (Tay et al. 2022). Continuous DRO has been discussed in (Husain, Nguyen, and Hengel 2022), but the algorithms they proposed only hold under certain conditions, which will be discussed in Section 3.

In this paper, we consider the SO objective and assume that the distribution of context variable is continuous and unknown. The setting is similar to the data-driven setting in (Kirschner et al. 2020), which, however, used a discrete context space. A similar setting has also been discussed in bandit problems (Lamprier, Gisselbrecht, and Gallinari 2018), where the context distributions for each arm are unknown and estimated online.

2.3 Kernel Density Estimation

Kernel density estimation (KDE) (Silverman 1986; Scott 2015; Chen 2017) is a non-parametric method used for estimating the probability density function (PDF) of a random variable, and is widely used in machine learning communities due to its flexibility (Elgammal et al. 2002; Pérez, Larrañaga, and Inza 2009). The basic idea of KDE is to estimate the PDF by aggregating the density assigned around each sample. Given the i.i.d. samples $\{\boldsymbol{c}_i\}_{i=1}^t$ drawn from $p(\boldsymbol{c})$, the estimated distribution $\hat{p}(\boldsymbol{c})$ of $p(\boldsymbol{c})$ is calculated as

$$\hat{p}(\boldsymbol{c}) = \sum\nolimits_{i=1}^{t} K\left(\mathbf{H}_t^{-1}(\boldsymbol{c}-\boldsymbol{c}_i)\right)/(t|\mathbf{H}_t|), \tag{2}$$

where $\mathbf{H}_t = \mathrm{diag}([h_t^{(1)}, h_t^{(2)}, \ldots, h_t^{(D_c)}])$ is a diagonal positive definite bandwidth matrix, $|\mathbf{H}_t|$ and $\mathbf{H}_t^{-1}$ denote the determinant and inverse of $\mathbf{H}_t$, respectively, and $K(\cdot)$ is a kernel function satisfying $K(\boldsymbol{c}) = K(-\boldsymbol{c}), \forall \boldsymbol{c} \in \mathcal{C}$, $\int_{\mathcal{C}} \boldsymbol{c}\boldsymbol{c}^{\mathrm{T}} K(\boldsymbol{c})\, d\boldsymbol{c} = m_2(K)\mathbf{I}_{D_c}$ for some constant $m_2(K) > 0$, and $\int_{\mathcal{C}} K(\boldsymbol{c})\, d\boldsymbol{c} = 1$.

Besides flexibility, KDE has good theoretical convergence properties for different error functions, e.g., uniform error (Jiang 2017), ℓ_1 error (Deroye and Gyorfi 1985) and mean integrated square error (MISE) (Wand and Jones 1994; Chen 2017). In this work, we primarily focus on MISE, which is one of the most well-known error measurements. Lemma 1 gives an upper bound on the MISE. It can be shown that by choosing $h_t^{(i)}\, \forall i$ to be of order $\Theta\left(t^{-1/(4+D_c)}\right)$, the MISE can be upper bounded by $\mathcal{O}\left(t^{-4/(D_c+4)}\right)$ (Wand and Jones 1994), which will play a crucial role in deriving the regret bound for our algorithms in Section 3.

Lemma 1 (Wand and Jones, 1994)**.** *Suppose $K(\cdot)$ is a bounded kernel for KDE, and $p(\boldsymbol{c})$ is a twice-differentiable PDF over $\mathcal{C}$. Let $J = \int_{\mathcal{C}} (p(\boldsymbol{c}) - \hat{p}(\boldsymbol{c}))^2\, d\boldsymbol{c}$ with $\hat{p}$ defined as Eq. (2). Then the MISE $\mathbb{E}[J]$ has an order of*

$$\mathcal{O}\left(\frac{1}{t|\mathbf{H}_t|}R(K) + \frac{1}{4}m_2(K)^2(\mathrm{vec}^{\mathrm{T}}(\mathbf{H}_t^2))\Psi_4(\mathrm{vec}(\mathbf{H}_t^2))\right),$$

where $R(K) = \int_{\mathcal{C}} K(\boldsymbol{c})^2\, d\boldsymbol{c}$, $\mathrm{vec}(\cdot)$ is the vector operator that vectorizes a matrix into a column vector, $\Psi_4 = \int_{\mathcal{C}} (\mathrm{vec}(\nabla^2 p(\boldsymbol{c})))(\mathrm{vec}^{\mathrm{T}}(\nabla^2 p(\boldsymbol{c})))\, d\boldsymbol{c}$ is a $D_c^2 \times D_c^2$ matrix of integrated second order partial derivatives of the PDF p, and the expectation $\mathbb{E}[J]$ is taken over the randomness of samples $\{\boldsymbol{c}_i\}_{i=1}^t$ from $p(\boldsymbol{c})$.

3 Stochastic Bayesian Optimization with Kernel Density Estimation

We propose two algorithms to optimize the SO objective with unknown continuous context PDF. The main idea is to estimate the PDF of the context online using KDE. The only difference lies in the design and optimization of their acquisition functions. The first algorithm, SBO-KDE, is directly based on an acquisition function of SO objective, which takes the expectation under the PDF estimated by KDE. The second one, DRBO-KDE, is based on a distributionally robust acquisition function, which accounts for distribution discrepancy between the true and estimated PDF, by taking the worst-case expected value in the distribution set centered around the estimated PDF.

3.1 SBO-KDE

SBO-KDE optimizes the SO objective $\mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t(\boldsymbol{c})}[f(\boldsymbol{x}, \boldsymbol{c})]$ directly by using the estimated distribution $\hat{p}_t$ by KDE. The key component of the algorithm is the acquisition function $\alpha_t(\boldsymbol{x}) = \mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t(\boldsymbol{c})}[\mathrm{ucb}_t(\boldsymbol{x}, \boldsymbol{c})]$, which can be interpreted as the expectation of the UCB acquisition function under the estimated distribution $\hat{p}_t$. The algorithm procedure is described in Algorithm 1. In line 1, the initial data set $\mathcal{D}_{n_0} = \{(\boldsymbol{x}_i, \boldsymbol{c}_i, y_i)\}_{i=1}^{n_0}$ is sampled using Sobol sequence (Owen 2003), where n_0 is the number of initial points. The optimization procedure is shown in lines 2–8. At iteration t, with the observed context $\mathcal{C}_{t-1} = \{\boldsymbol{c}_i\}_{i=1}^{t-1}$, we estimate the unknown context distribution $p(\boldsymbol{c})$ using KDE in line 3, and the estimated distribution is denoted as $\hat{p}_t(\boldsymbol{c})$. Then we fit a GP model based on the current data set $\mathcal{D}_{t-1}$ in line 4. With the estimated PDF and the posterior information, we optimize the acquisition function using sample average approximation (SAA) to get the next query point $\boldsymbol{x}_t$ in line 5. SAA uses the average of sample values to estimate the value of acquisition function $\alpha_t(\boldsymbol{x})$. When evaluating $\boldsymbol{x}_t$, the context $\boldsymbol{c}_t$ provided by the environment and the noisy function value y_t is observed in line 6. The data set is then augmented with the new triple $(\boldsymbol{x}_t, \boldsymbol{c}_t, y_t)$ in line 7. The whole process is repeated for $T - n_0$ iterations.

To optimize the acquisition function $\alpha_t(\boldsymbol{x})$, any technique from traditional SO can be employed. In this work, we adopt the SAA method (Homem-de Mello 2008; Kim, Pasupathy, and Henderson 2015), which optimizes the average function value of Monte Carlo samples. Specifically, we draw M samples $\{\hat{c}_i\}_{i=1}^M$ from $\hat{p}_t(\boldsymbol{c})$ and estimate the value of acquisition function as $\hat{\alpha}_t^M(\boldsymbol{x}) = \frac{1}{M}\sum_{i=1}^M \mathrm{ucb}_t(\boldsymbol{x}, \hat{\boldsymbol{c}}_i)$. We then

Algorithm 1: SBO-KDE

Parameters: number n_0 of initial points, budget T

Process:

1: Obtain the initial data set $\mathcal{D}_{n_0} = \{(\boldsymbol{x}_i, \boldsymbol{c}_i, y_i)\}_{i=1}^{n_0}$ and context $\mathcal{C}_{n_0} = \{\boldsymbol{c}_i\}_{i=1}^{n_0}$ using Sobol sequence;

2: **for** $t = n_0 + 1$ to T **do**

3: Use KDE to obtain $\hat{p}_t$ based on $\mathcal{C}_{t-1} = \{\boldsymbol{c}_i\}_{i=1}^{t-1}$;

4: Fit a GP model using $\mathcal{D}_{t-1} = \{(\boldsymbol{x}_i, \boldsymbol{c}_i, y_i)\}_{i=1}^{t-1}$;

5: Optimize $\boldsymbol{x}_t = \arg\max_{\boldsymbol{x}\in\mathcal{X}} \mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t(\boldsymbol{c})}[\text{ucb}_t(\boldsymbol{x}, \boldsymbol{c})]$ using SAA;

6: Evaluate $\boldsymbol{x}_t$, and then observe $\boldsymbol{c}_t \sim p(\boldsymbol{c})$ and $y_t = f(\boldsymbol{x}_t, \boldsymbol{c}_t) + \epsilon_t$;

7: $\mathcal{D}_t = \mathcal{D}_{t-1} \cup \{(\boldsymbol{x}_t, \boldsymbol{c}_t, y_t)\}$

8: **end for**

optimize $\boldsymbol{x}_t = \arg\max_{\boldsymbol{x}\in\mathcal{X}} \hat{\alpha}_t^M(\boldsymbol{x})$ using L-BFGS (Nocedal 1980). SAA is a popular technique for optimizing acquisition functions in BO (Balandat et al. 2020; Cakmak et al. 2020), due to its exponential convergence property in Proposition 1. The property can be derived based on the theoretical results from (Homem-de Mello 2008; Balandat et al. 2020), and the proof is provided in (Huang et al. 2023). The exponential convergence rate of SAA enables us to obtain good acquisition function optimization quality of $\alpha_t(\boldsymbol{x})$.

Proposition 1. *Suppose that (i) $\{\hat{\boldsymbol{c}}_i\}_{i=1}^M$ is i.i.d., and (ii) f is a GP with continuously differentiable prior mean and kernel function. Then, $\forall\delta > 0$, there exist $Q < \infty$ and $\eta > 0$ such that $\mathbb{P}(dist(\hat{\boldsymbol{x}}_M^*, \mathcal{X}_t^*) > \delta) \le Qe^{-\eta M}$ for all $M \ge 1$, where $dist(\hat{\boldsymbol{x}}_M^*, \mathcal{X}_t^*) = \inf_{\boldsymbol{x}\in\mathcal{X}_t^*} \|\boldsymbol{x} - \hat{\boldsymbol{x}}_M^*\|_2$, $\hat{\boldsymbol{x}}_M^* \in \arg\max_{\boldsymbol{x}\in\mathcal{X}} \hat{\alpha}_t^M(\boldsymbol{x})$ and $\mathcal{X}_t^* = \arg\max_{\boldsymbol{x}\in\mathcal{X}} \alpha_t(\boldsymbol{x})$.*

Although the acquisition function $\alpha_t(\boldsymbol{x})$ uses an estimated context distribution $\hat{p}_t$ by KDE, we prove that the algorithm SBO-KDE can still achieve a sub-linear bound under the true distribution p, as shown in Theorem 1. The sub-linear bound is on the commonly used Bayesian cumulative regret (BCR) (Russo and Van Roy 2014; Kandasamy et al. 2018a; Nguyen et al. 2020) in Definition 1, which is actually the expectation of R_T in Eq. (1).

Definition 1 (BCR)**.** *Let $r_t := \mathbb{E}_{\boldsymbol{c}\sim p}[f(\boldsymbol{x}^*, \boldsymbol{c})] - \mathbb{E}_{\boldsymbol{c}\sim p}[f(\boldsymbol{x}_t, \boldsymbol{c})]$ denote the regret at iteration t, where $\boldsymbol{x}^* \in \arg\max_{x\in\mathcal{X}} \mathbb{E}_{\boldsymbol{c}\sim p}[f(\boldsymbol{x}, \boldsymbol{c})]$ is the optimum. Then, the Bayesian cumulative regret is defined as*

$$BCR(T) := \mathbb{E}\left[\sum_{t=1}^{T} r_t\right] \tag{3}$$
$$= \mathbb{E}\left[\sum_{t=1}^{T} \left(\mathbb{E}_{\boldsymbol{c}\sim p}[f(\boldsymbol{x}^*, \boldsymbol{c})] - \mathbb{E}_{\boldsymbol{c}\sim p}[f(\boldsymbol{x}_t, \boldsymbol{c})]\right)\right],$$

where the outer expectation is taken over the GP f, the randomness of samples from $p(\boldsymbol{c})$ and the observation noise ϵ_t.

As in (Srinivas et al. 2012), we assume that the input space $\mathcal{Z} = \mathcal{X}\times\mathcal{C} \subset [0, r]^{D_x+D_c} = [0, r]^D$ is convex and compact, and f satisfies the following Lipschitz assumption.

Assumption 1. *The function f is a GP sample path from $\mathcal{GP}(\boldsymbol{0}, k(\cdot,\cdot))$ with $k(\boldsymbol{z}, \boldsymbol{z}') \le 1$, $\forall \boldsymbol{z}, \boldsymbol{z}' \in \mathcal{Z}$. Let $[D] = \{1, 2, \ldots, D\}$. For some $a, b > 0$, $\forall L > 0$, the partial derivatives of f satisfy*

$$\forall i \in [D],\ \mathbb{P}(\sup_{\boldsymbol{z}\in\mathcal{Z}} \mid \partial f(\boldsymbol{z})/\partial z_i \mid > L) \le ae^{-(L/b)^2}.$$

Theorem 1 provides an upper bound on BCR(T) of SBO-KDE. Ignoring the log factors from $\beta_T\gamma_T$, compared with the bound of stochastic BO (SBO) with $\mathcal{O}(T^{1/2})$ in (Kirschner and Krause 2019), our bound increases to $\mathcal{O}(T^{(2+D_c)/(4+D_c)})$, which comes from the estimation error of KDE under the unknown context distribution setting. However, the bound is still sub-linear, i.e., $\lim_{T\to\infty} \text{BCR}(T)/T = 0$.

Theorem 1. *Let $\beta_t = 2\log(t^2/\sqrt{2\pi}) + 2D_x \log(t^2 D_x abr\sqrt{\pi}/2)$. With the underlying PDF $p(\boldsymbol{c})$ satisfying the condition in Lemma 1, $\hat{p}_t(\boldsymbol{c})$ defined as Eq. (2) and $h_t^{(i)} = \Theta\left(t^{-1/(4+D_c)}\right) \forall i \in [D_c]$, the BCR of SBO-KDE satisfies*

$$\text{BCR}(T) \le \frac{\pi^2}{3} + \sqrt{\beta_T\gamma_T C_2}\left(\sqrt{T^{D_c/(4+D_c)}} + \sqrt{T}\right) + 2C_1 T^{\frac{2+D_c}{4+D_c}}, \tag{4}$$

where $C_1, C_2 > 0$ are constants, $\gamma_T = \max_{|\mathcal{D}|=T} I(\boldsymbol{y}_{\mathcal{D}}, \boldsymbol{f}_{\mathcal{D}})$, $I(\cdot,\cdot)$ is the information gain, and $\boldsymbol{y}_{\mathcal{D}}, \boldsymbol{f}_{\mathcal{D}}$ are the noisy and true observations of a data set $\mathcal{D}$, respectively.

We present only a proof sketch here, and the detailed proof can be found in (Huang et al. 2023). The proof idea is to decompose the instantaneous regret r_t into the following three terms using the estimated PDF $\hat{p}_t$. Specifically, $r_t = \mathbb{E}_{\boldsymbol{c}\sim p}[f(\boldsymbol{x}^*, \boldsymbol{c})] - \mathbb{E}_{\boldsymbol{c}\sim p}[f(\boldsymbol{x}_t, \boldsymbol{c})] = (\mathbb{E}_{\boldsymbol{c}\sim p}[f(\boldsymbol{x}^*, \boldsymbol{c})] - \mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t}[f(\boldsymbol{x}^*, \boldsymbol{c})]) + (\mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t}[f(\boldsymbol{x}^*, \boldsymbol{c})] - \mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t}[f(\boldsymbol{x}_t, \boldsymbol{c})]) + (\mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t}[f(\boldsymbol{x}_t, \boldsymbol{c})] - \mathbb{E}_{\boldsymbol{c}\sim p}[f(\boldsymbol{x}_t, \boldsymbol{c})])$. The first and last terms can be upper bounded using the error bound between the true distribution p and the estimated distribution $\hat{p}_t$ as given in Lemma 1. The second term $\mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t}[f(\boldsymbol{x}^*, \boldsymbol{c})] - \mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t}[f(\boldsymbol{x}_t, \boldsymbol{c})]$ is the UCB regret under expectation over $\hat{p}_t$, which can be upper bounded using the similar idea as in (Kandasamy et al. 2018a). That is, using the fact that $\boldsymbol{x}_t = \arg\max_{\boldsymbol{x}\in\mathcal{X}} \mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t}[\text{ucb}_t(\boldsymbol{x}_t, \boldsymbol{c})]$, we further decompose the second term as follows: $\mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t}[f(\boldsymbol{x}^*, \boldsymbol{c})] - \mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t}[f(\boldsymbol{x}_t, \boldsymbol{c})] \le (\mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t}[f(\boldsymbol{x}^*, \boldsymbol{c})] - \mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t}[\text{ucb}_t(\boldsymbol{x}^*, \boldsymbol{c})]) + (\mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t}[\text{ucb}_t(\boldsymbol{x}_t, \boldsymbol{c})] - \mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t}[f(\boldsymbol{x}_t, \boldsymbol{c})])$, which can be upper bounded under the posterior information of GP. By summing up the upper bound on r_t from $t = 1$ to T, the upper bound on BCR(T), i.e., Eq. (4), can be derived.

3.2 DRBO-KDE

Considering that the true PDF of context variable in practice may be complicated, leading to a non-negligible distribution discrepancy between the estimated PDF and the true PDF, we further propose the second algorithm, DRBO-KDE. The optimization objective is the worst-case expectation $\min_{q\in\mathcal{B}(\hat{p}_t,\delta_t)} \mathbb{E}_{\boldsymbol{c}\sim q}[f(\boldsymbol{x}, \boldsymbol{c})]$ within a distribution set $\mathcal{B}(\hat{p}_t, \delta_t) = \{q : d(q, \hat{p}_t) \le \delta_t\}$, where $d(\cdot,\cdot)$ is a distance measure over the distribution space, and $\delta_t > 0$ is the radius of the ball centered around $\hat{p}_t$. We choose the total variation metric (Tsybakov 2008) here, which is a type of ϕ-divergence (Bayraksan and Love 2015). The total variation between two distributions Q and P, with PDFs $dQ = q$ and $dP = p$ respectively, is defined as $d_{TV}(Q, P) = d(q, p) = \int_{\mathcal{C}} p(\boldsymbol{c})\phi\left(\frac{q(\boldsymbol{c})}{p(\boldsymbol{c})}\right) d\boldsymbol{c}$ with $\phi(x) = |x - 1|$ and $Q \ll P$. The total variation metric can be thought of as the ℓ_1 distance between two PDFs when $Q \ll P$.

While DRBO with ϕ-divergence over continuous space has been discussed in (Husain, Nguyen, and Hengel 2022), the proposed algorithms under χ^2-divergence and total variation hold only if the variance of the function value is sufficiently high, which can be derived from the result of Theorem 1 in (Namkoong and Duchi 2017). Therefore, we redevelop an algorithm for DRBO over continuous context space using total variation. Proposition 2 derives an equivalent form of the DRO objective under total variation, which transforms the inner convex minimization problem, which has infinite dimensions, into a two-dimensional convex SO problem. This transformation has been commonly utilized in DRO literature (Bayraksan and Love 2015; Rahimian and Mehrotra 2022). For the sake of completeness, we also provide a detailed derivation in (Huang et al. 2023).

Proposition 2 (DRO under Total Variation)**.** *Given a bounded function $f(\boldsymbol{x}, \boldsymbol{c})$ over $\mathcal{X} \times \mathcal{C}$, a radius $\delta > 0$ and a PDF $p(\boldsymbol{c})$, we have*

$$\begin{aligned} &\min_{q\in\mathcal{B}(p,\delta)} \mathbb{E}_{\boldsymbol{c}\sim q(\boldsymbol{c})}\left[f(\boldsymbol{x}, \boldsymbol{c})\right] \qquad (5) \\ &= \max_{(\alpha,\beta)\in S(f)} \mathbb{E}_{\boldsymbol{c}\sim p(\boldsymbol{c})}\left[-\beta - \delta\alpha + \min\{f(\boldsymbol{x}, \boldsymbol{c}) + \beta, \alpha\}\right], \end{aligned}$$

where $\mathcal{B}(p, \delta) = \{q : d(q, p) \leq \delta\}$, and $S(f) := \{(\alpha, \beta) : \beta \in \mathbb{R}, \alpha \geq 0, \alpha + \beta \geq -\inf_{\boldsymbol{c}\in\mathcal{C}} f(\boldsymbol{x}, \boldsymbol{c})\}$.

Based on Proposition 2, we propose DRBO-KDE as presented in Algorithm 2. The only difference between SBO-KDE and DRBO-KDE is the definition and optimization of the acquisition function in line 5. For DRBO-KDE, the acquisition function is defined as $\alpha_t(\boldsymbol{x}) = \min_{q\in\mathcal{B}(\hat{p}_t,\delta_t)} \mathbb{E}_{\boldsymbol{c}\sim q(\boldsymbol{c})}\left[\text{ucb}_t(\boldsymbol{x}, \boldsymbol{c})\right]$, which is the worst expectation of UCB in the distribution set $\mathcal{B}(\hat{p}_t, \delta_t)$. By applying Proposition 2, we can equivalently transform the optimization into a two-dimensional SO problem: $\alpha_t(\boldsymbol{x}) = \max_{(\alpha,\beta)\in S(\text{ucb}_t)} \mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t(\boldsymbol{c})}[-\beta - \delta_t\alpha + \min\{\text{ucb}_t(\boldsymbol{x}, \boldsymbol{c}) + \beta, \alpha\}]$. We also use SAA to optimize the acquisition function, i.e., we sample the Monte Carlo samples $\{\hat{\boldsymbol{c}}_i\}_{i=1}^M$ from $\hat{p}_t(\boldsymbol{c})$ to find the next query point

$$\begin{aligned} \boldsymbol{x}_t = \arg\max_{\boldsymbol{x}\in\mathcal{X}} \max_{(\alpha,\beta)\in S(\text{ucb}_t)} \frac{1}{M} \sum\nolimits_{i=1}^{M} (-\beta - \delta_t\alpha \\ + \min\{\text{ucb}_t(\boldsymbol{x}, \hat{\boldsymbol{c}}_i) + \beta, \alpha\}). \qquad (6) \end{aligned}$$

This is a two-stage optimization, where the inner optimization problem is a two-dimensional convex optimization problem which can be solved efficiently, and the outer optimization is solved by using L-BFGS (Nocedal 1980). Note that $S(\text{ucb}_t)$ can be calculated by numerical optimization.

Theorem 2 gives an upper bound on $\text{BCR}(T)$ of DRBO-KDE. Compared with the bound of SBO-KDE in Theorem 1, the bound of DRBO-KDE is higher by the additional terms related to δ_t. This is because in our proof, we used the estimated distribution $\hat{p}_t$ to establish the connection between the true distribution p and the distribution set $\mathcal{B}(\hat{p}_t, \delta_t)$

Algorithm 2: DRBO-KDE

Parameters: number n_0 of initial points, evaluation budget T, radius $\delta_t > 0$

Process:

1: Obtain the initial data set $\mathcal{D}_{n_0} = \{(\boldsymbol{x}_i, \boldsymbol{c}_i, y_i)\}_{i=1}^{n_0}$ and context $\mathcal{C}_{n_0} = \{\boldsymbol{c}_i\}_{i=1}^{n_0}$ using Sobol sequence;
2: **for** $t = n_0 + 1$ to T **do**
3: Use KDE to obtain $\hat{p}_t$ based on $\mathcal{C}_{t-1} = \{\boldsymbol{c}_i\}_{i=1}^{t-1}$;
4: Fit a GP model using $\mathcal{D}_{t-1} = \{(\boldsymbol{x}_i, \boldsymbol{c}_i, y_i)\}_{i=1}^{t-1}$;
5: Optimize $\boldsymbol{x}_t = \arg\max_{\boldsymbol{x}\in\mathcal{X}} \min_{q\in\mathcal{B}(\hat{p}_t,\delta_t)} \mathbb{E}_{\boldsymbol{c}\sim q(\boldsymbol{c})}$ $[\text{ucb}_t(\boldsymbol{x}, \boldsymbol{c})]$ using Eq. (5) and SAA;
6: Evaluate $\boldsymbol{x}_t$, and then observe $\boldsymbol{c}_t \sim p(\boldsymbol{c})$ and $y_t = f(\boldsymbol{x}_t, \boldsymbol{c}_t) + \epsilon_t$;
7: $\mathcal{D}_t = \mathcal{D}_{t-1} \cup \{(\boldsymbol{x}_t, \boldsymbol{c}_t, y_t)\}$
8: **end for**

around $\hat{p}_t$. If we could directly build the connection between p and $\mathcal{B}(\hat{p}_t, \delta_t)$, we might obtain a better bound, which is, however, more challenging and is left for future work. Nevertheless, by setting $\delta_t = \mathcal{O}\left(t^{-2/(4+D_c)}\right)$, the bound of DRBO-KDE can also be sub-linear with the same order of $\mathcal{O}\left(T^{(2+D_c)/(4+D_c)}\right)$ as SBO-KDE.

Theorem 2. *Let $\beta_t = 2\log(t^2/\sqrt{2\pi}) + 2D_x \log(t^2 D_x abr\sqrt{\pi}/2)$. With the underlying PDF $p(\boldsymbol{c})$ satisfying the condition in Lemma 1, $\hat{p}_t(\boldsymbol{c})$ defined as Eq. (2) and $h_t^{(i)} = \Theta\left(t^{-1/(4+D_c)}\right) \forall i \in [D_c]$, the BCR of DRBO-KDE satisfies*

$$\begin{aligned} \text{BCR}(T) \leq &\frac{\pi^2}{3} + \sqrt{\beta_T\gamma_T C_2}\Bigg(\sqrt{T^{D_c/(4+D_c)}} + \sqrt{T} \\ &+ \sqrt{C_3 \sum_{t=1}^{T} \delta_t^2}\Bigg) + 2C_1 T^{\frac{2+D_c}{4+D_c}} + \sum_{t=1}^{T} C_4\delta_t, \end{aligned}$$

where $C_1, C_2, C_3, C_4 > 0$ are constants, $\gamma_T = \max_{|\mathcal{D}|=T} I(\boldsymbol{y}_\mathcal{D}, \boldsymbol{f}_\mathcal{D})$, $I(\cdot,\cdot)$ is the information gain, and $\boldsymbol{y}_\mathcal{D}, \boldsymbol{f}_\mathcal{D}$ are the noisy and true observations of a data set $\mathcal{D}$, respectively.

We also only present a proof sketch here, and the detailed proof is provided in (Huang et al. 2023). The idea is similar to that of Theorem 1 for SBO-KDE. Specifically, for any function g, let $q_{\boldsymbol{x}}^g := \arg\min_{q\in\mathcal{B}(\hat{p}_t,\delta_t)} \mathbb{E}_{\boldsymbol{c}\sim q}[g(\boldsymbol{x}, \boldsymbol{c})]$. Then, we decompose the instantaneous regret at iteration t as $r_t = \mathbb{E}_{\boldsymbol{c}\sim p}[f(\boldsymbol{x}^*, \boldsymbol{c})] - \mathbb{E}_{\boldsymbol{c}\sim p}[f(\boldsymbol{x}_t, \boldsymbol{c})] = (\mathbb{E}_{\boldsymbol{c}\sim p}[f(\boldsymbol{x}^*, \boldsymbol{c})] - \mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t}[f(\boldsymbol{x}^*, \boldsymbol{c})]) + (\mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t}[f(\boldsymbol{x}^*, \boldsymbol{c})] - \mathbb{E}_{\boldsymbol{c}\sim q_{\boldsymbol{x}^*}^f}[f(\boldsymbol{x}^*, \boldsymbol{c})]) + (\mathbb{E}_{\boldsymbol{c}\sim q_{\boldsymbol{x}^*}^f}[f(\boldsymbol{x}^*, \boldsymbol{c})] - \mathbb{E}_{\boldsymbol{c}\sim q_{\boldsymbol{x}_t}^f}[f(\boldsymbol{x}_t, \boldsymbol{c})]) + (\mathbb{E}_{\boldsymbol{c}\sim q_{\boldsymbol{x}_t}^f}[f(\boldsymbol{x}_t, \boldsymbol{c})] - \mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t}[f(\boldsymbol{x}_t, \boldsymbol{c})]) + (\mathbb{E}_{\boldsymbol{c}\sim\hat{p}_t}[f(\boldsymbol{x}_t, \boldsymbol{c})] - \mathbb{E}_{\boldsymbol{c}\sim p}[f(\boldsymbol{x}_t, \boldsymbol{c})])$. All terms, with the except of the third one, can be bounded using the error bound between the estimated PDF $\hat{p}_t$ and the true distribution p or the radius δ_t of the distribution ball $\mathcal{B}(\hat{p}_t, \delta_t)$. The third term is the UCB regret, but under the DRO objective at each iteration, which can be bounded using the posterior information of GP and the fact that $q_{\boldsymbol{x}}^g$ belongs to $\mathcal{B}(\hat{p}_t, \delta_t)$ for any $\boldsymbol{x}$ and g.

4 Experiments

In order to empirically evaluate the effectiveness of SBO-KDE and DRBO-KDE, we conduct numerical experiments on synthetic functions and two real-world problems, i.e., the Newsvendor and portfolio problem. We use five identical random seeds (100–104) for all problems and methods. The code is available at https://github.com/lamda-bbo/sbokde.

4.1 Experimental Setting

We adopt the frequently-used cumulative regret (or reward) in BO literature considering uncertainty (Kirschner et al. 2020; Tay et al. 2022) as performance evaluation metric. The experimental setting of our algorithms and compared baselines are summarized as follows.

SBO-KDE. We choose the Gaussian kernel $K(\boldsymbol{x}) = (2\pi)^{-D_c/2}e^{-\|\boldsymbol{x}\|_2^2}$ for KDE. The bandwidth $h_t^{(i)} = (4/(D_c+2))^{1/(4+D_c)}\hat{\sigma}_t^{(i)}t^{-1/(4+D_c)}$ based on the rule of thumb (Silverman 1986), where $\hat{\sigma}_t^{(i)}$ is the standard deviation of the ith dimension of observed context.

DRBO-KDE. The radius of the distribution set is set as $\delta_t = t^{-2/(4+D_c)}$, which can guarantee a sub-linear regret as we introduced in Section 3.2. The kernel and bandwidth for KDE are the same as those used in SBO-KDE.

DRBO-MMD (Kirschner et al. 2020) discretizes the continuous context space $\mathcal{C}$, and selects the next query point $\boldsymbol{x}_t = \arg\max_{\boldsymbol{x}\in\mathcal{X}} \min_{Q\in\mathcal{B}(\hat{P}_t,\delta_t)} \mathbb{E}_{\boldsymbol{c}\sim Q}[\mathrm{ucb}_t(\boldsymbol{x},\boldsymbol{c})]$ with mean maximum discrepancy (MMD) as the distribution distance on the discretized context space $\tilde{\mathcal{C}}$. The reference distribution $\hat{P}_t$ is the empirical distribution in $\tilde{\mathcal{C}}$, and we set $\delta_t = (2+\sqrt{2\log(1/\gamma)})/\sqrt{t}$ with $\gamma = 0.1$, as suggested by Lemma 3 in (Kirschner et al. 2020). Due to its high computational complexity of the inner convex optimization, we use a small discretization size $|\tilde{C}| = \lceil 100^{1/D_c}\rceil^{D_c}$.

DRBO-MMD-MinimaxApprox (Tay et al. 2022) accelerates DRBO-MMD with minimax approximation. Thus, we can use a larger discretization size $|\tilde{C}| = \lceil 1024^{1/D_c}\rceil^{D_c}$.

StableOpt (Bogunovic et al. 2018) selects $\boldsymbol{x}_t = \arg\max_{\boldsymbol{x}\in\mathcal{X}} \min_{\boldsymbol{c}\in\mathcal{C}_t} \mathrm{ucb}_t(\boldsymbol{x},\boldsymbol{c})$. There is no standard way to choose $\mathcal{C}_t$. Instead of setting $\mathcal{C}_t = \mathcal{C}$, which is overly conservative and assumes worst-case scenario under the full context space, we set each dimension of $\mathcal{C}_t$ to $[\mu_{\boldsymbol{c}}^{(i)} - \sigma_{\boldsymbol{c}}^{(i)}, \mu_{\boldsymbol{c}}^{(i)} + \sigma_{\boldsymbol{c}}^{(i)}]$, where $\mu_{\boldsymbol{c}}^{(i)}$ and $\sigma_{\boldsymbol{c}}^{(i)}$ are the mean and standard deviation of the ith dimension of observed context, respectively.

GP-UCB (Srinivas et al. 2012) ignores the context variable. That is, it builds GP only on the decision space $\mathcal{X}$ and selects the next query point $\boldsymbol{x}_t$ by maximizing the UCB acquisition function.

Further detailed descriptions and hyper-parameters of these algorithms are provided in (Huang et al. 2023).

4.2 Synthetic Functions

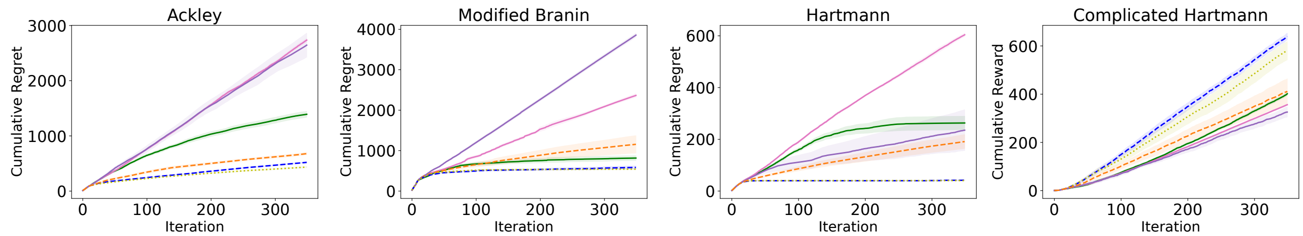
We conduct experiments on four commonly used synthetic test functions (Surjanovic and Bingham 2013), in which we follow the approach of setting some dimensions as context variable from (Williams 2000; Cakmak et al. 2020). The functions include the Ackley function with one dimension set as a context variable following a normal distribution, the Modified Branin function with two dimensions set as context variable following a normal distribution, the Hartmann function with one dimension set as a context variable following a normal distribution, and the Complicated Hartmann function whose context variable follows a more complicated distribution (a mixture of six normal and two Cauchy distributions). More details can be found in (Huang et al. 2023). For the first three functions, we use the cumulative regret in Eq. (1) as the metric,[1] and we calculate the expectation $\mathbb{E}_{\boldsymbol{c}\sim p(\boldsymbol{c})}[f(\boldsymbol{x},\boldsymbol{c})]$ by averaging 2^{21} quasi-Monte Carlo (QMC) samples.[2] For the last function, due to the complexity of the context distribution, it is difficult to perform QMC sampling, so we only report the observed cumulative reward, that is $\sum_{t=1}^{T} f(\boldsymbol{x}_t,\boldsymbol{c}_t)$.

The results are shown in Figure 1(a). We can observe that the proposed algorithms, SBO-KDE and DRBO-KDE, outperform all the other baselines on the synthetic functions. In the Ackley function, SBO-KDE performs slightly better than DRBO-KDE, which is because DRBO-KDE is more conservative. However, in the Complicated Hartmann function with a more complex context distribution, DRBO-KDE performs better. This is because KDE may suffer from a higher estimation error for complicated distributions, while DRBO-KDE takes into account the discrepancy between the estimated distribution and the true distribution, thus exhibiting a more robust performance. To observe the higher estimation error between the KDE and the true context distribution under the complicated distribution, we report the discrepancy measured by total variation between PDF estimated by KDE and the true context distribution under the two distributions on Hartmann function in (Huang et al. 2023). In the Modified Branin and original Hartmann functions, the performance of SBO-KDE and DRBO-KDE is very close. It is interesting to note that GP-UCB, which does not consider context variable, has an acceptable performance. This may be because GP-UCB can model the impact of context variable on function evaluations as evaluation noise. DRBO-MMD performs well in the Modified Branin function but performs worse in the other functions, which is related to the quality of the discretization space approximation. StableOpt performs poorly because the robust objective is too pessimistic. DRBO-MMD-MinimaxApprox also has poor performance, which could be due to the minimax approximation error being too large for these problems.

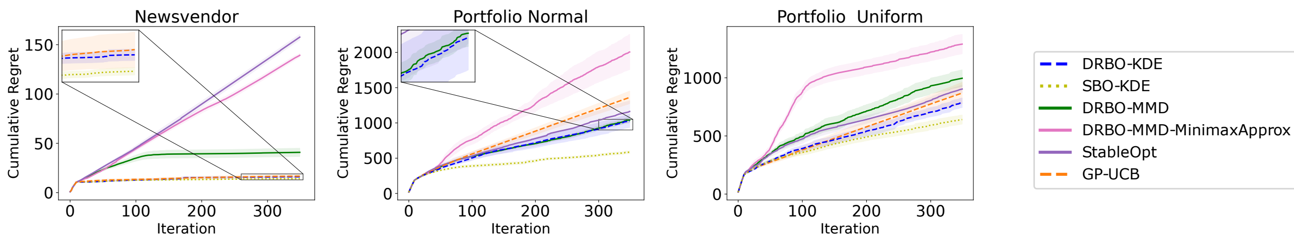
In addition to the performance of optimization, we also provide the computational complexity comparison of the algorithms in (Huang et al. 2023). Besides, although we use a small dimension of context variable by following the experiments in DRBO literature, where the dimension of context variable tends to be relatively low (at most three) (Kirschner et al. 2020; Tay et al. 2022), we conduct experiments on

[1] The optimal solution is approximately obtained by optimizing the average of QMC samples using multi-restart L-BFGS.

[2] We set it arbitrarily. Using more QMC samples leads to more accurate estimation. Due to limitations in computing resources, we use 2^{21} QMC samples, which, however, can guarantee the estimation accurate enough.



(a) Synthetic Functions



(b) Real-World Problems

Figure 1: Mean and standard error of cumulative regret (the lower the better) or cumulative reward (the higher the better).

one more problem with four-dimensional context in (Huang et al. 2023), to show the performance of algorithms on problems with higher dimensional context variable.

4.3 Real-World Problems

We further examine the effectiveness of our algorithms on two real-world optimization tasks, including newsvendor problem and portfolio optimization. Newsvendor problem is a classic inventory management problem in stochastic optimization, where a seller pre-determines the inventory to satisfy customer demand, and portfolio optimization is the process of adjusting trading strategies to maximize the returns on investment.

Newsvendor problem. The first real-world problem we consider is a continuous newsvendor problem (Eckman et al. 2021), where a vendor purchases a certain amount of liquid denoted by x, and sells to customers with a demand of c units. The vendor incurs a cost of s_0 per unit for the initial inventory and sells the liquid to customers at a price of s_1 per unit. Any unsold liquid at the end of the day can be sold for a salvage value of w per unit. The decision variable is the purchase quantity x, and the context variable is the customer demand c. The goal is to maximize the vendor's profit, which is defined as $f(x, c) = s_1 \min\{x, c\} + w \max\{0, x - c\} - s_0 x$. We use the default setting of (Eckman et al. 2021), where $s_0 = 5$, $s_1 = 9$, and $w = 1$. Additionally, the customer demand c follows a Burr Type XII distribution with PDF $p(c; \alpha, \beta) = \alpha\beta \frac{c^{\alpha-1}}{(1+c^{\alpha})^{\beta+1}}$, where $\alpha = 2$ and $\beta = 20$.

Portfolio optimization. The second real-world problem is portfolio optimization (Cakmak et al. 2020; Nguyen et al. 2021). The problem involves three-dimensional decision variable (risk and trade aversion parameters, and holding cost multiplier), and two-dimensional context variable (bid-ask spread and borrowing cost). The objective function is the posterior mean of a GP trained on $3,000$ samples, which are generated by (Cakmak et al. 2020) from the CVXPortfolio problem (Boyd et al. 2017). We define two tasks by setting the distribution of the context variable as normal distribution or uniform distribution. For a more detailed description, please refer to (Huang et al. 2023). We modify the problem to the setting where the distribution is unknown and the context can be observed after each evaluation.

The results are shown in Figure 1(b). Newsvendor problem uses 2^{21} QMC samples for calculating the expectation, while portfolio optimization uses 2^{16} QMC samples because the evaluation is more time-consuming. SBO-KDE and DRBO-KDE still outperform all the other baselines, with SBO-KDE demonstrating better performance. Among the methods that consider context variable in the newsvendor problem, only SBO-KDE and DRBO-KDE outperform GP-UCB. For portfolio optimization with a normal context distribution, DRBO-MMD is competitive with DRBO-KDE, which is because the expectation under the discretized context space has a good approximation over the continuous context space.

5 Conclusion

In this paper, we consider the stochastic optimization problem with an unknown continuous context distribution, and propose the two algorithms, SBO-KDE and DRBO-KDE. The former directly optimizes the SO objective using the estimated density from KDE. The latter optimizes the distributionally robust objective considering the discrepancy between the true and estimated PDF, which is more suitable when the KDE approximation error might be high due to the complexity of the true distribution. We prove sublinear Bayesian cumulative regret bounds for both algorithms. Furthermore, we conduct numerical experiments on synthetic functions and two real-world problems to empirically demonstrate the effectiveness of the proposed algorithms. One limitation of this work is that we assume that the distribution of context variable remains static over time. We will investigate scenarios where distributional shifts occur in our future work.

Acknowledgements

This work was supported by the National Key R&D Program of China (2022ZD0116600) and National Science Foundation of China (62022039, 62276124).

References

Balandat, M.; Karrer, B.; Jiang, D. R.; Daulton, S.; Letham, B.; Wilson, A. G.; and Bakshy, E. 2020. BoTorch: a framework for efficient Monte-Carlo Bayesian optimization. In *Advances in Neural Information Processing Systems 33*, 21524–21538. Virtual.

Bayraksan, G.; and Love, D. K. 2015. Data-driven stochastic programming using phi-divergences. *The operations research revolution*, 1–19.

Beland, J. J.; and Nair, P. B. 2019. Bayesian optimization under uncertainty. In *Proceedings of the 22rd International Conference on Artificial Intelligence and Statistics*, volume 89, 1177–1184. Naha, Japan.

Bogunovic, I.; Scarlett, J.; Jegelka, S.; and Cevher, V. 2018. Adversarially robust optimization with Gaussian processes. In *Advances in Neural Information Processing Systems 31*, 5765–5775. Montréal, Canada.

Boyd, S.; Busseti, E.; Diamond, S.; Kahn, R. N.; Koh, K.; Nystrup, P.; Speth, J.; et al. 2017. Multi-period trading via convex optimization. *Foundations and Trends® in Optimization*, 3(1): 1–76.

Cakmak, S.; Astudillo Marban, R.; Frazier, P.; and Zhou, E. 2020. Bayesian Optimization of Risk Measures. In *Advances in Neural Information Processing Systems 33*, 20130–20141. Virtual.

Chen, Y.; Huang, A.; Wang, Z.; Antonoglou, I.; Schrittwieser, J.; Silver, D.; and de Freitas, N. 2018. Bayesian Optimization in AlphaGo. *CoRR*, abs/1812.06855.

Chen, Y.-C. 2017. A tutorial on kernel density estimation and recent advances. *Biostatistics & Epidemiology*, 1(1): 161–187.

Dai, L.; Chen, C.-H.; and Birge, J. R. 2000. Convergence Properties of Two-Stage Stochastic Programming. *Journal of Optimization Theory and Applications*, 106(3): 489–509.

Deroye, L.; and Gyorfi, L. 1985. *Nonparametric Density Estimation*. Wiley Series in Probability and Mathematical Statistics.

Eckman, D. J.; Henderson, S. G.; Shashaani, S.; and Pasupathy, R. 2021. SimOpt. https://github.com/simopt-admin/simopt. Accessed: 2023-06-12.

Elgammal, A.; Duraiswami, R.; Harwood, D.; and Davis, L. S. 2002. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance. *Proceedings of the IEEE*, 90(7): 1151–1163.

Frazier, P. I. 2018. A tutorial on Bayesian optimization. *CoRR*, abs/1807.02811.

Fröhlich, L.; Klenske, E.; Vinogradska, J.; Daniel, C.; and Zeilinger, M. 2020. Noisy-input entropy search for efficient robust Bayesian optimization. In *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics*, 2262–2272. Virtual.

Griffiths, R.-R.; and Hernández-Lobato, J. M. 2020. Constrained Bayesian optimization for automatic chemical design using variational autoencoders. *Chemical Science*, 11(2): 577–586.

Gómez-Bombarelli, R.; Wei, J. N.; Duvenaud, D.; Hernández-Lobato, J. M.; Sánchez-Lengeling, B.; Sheberla, D.; Aguilera-Iparraguirre, J.; Hirzel, T. D.; Adams, R. P.; and Aspuru-Guzik, A. 2018. Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules. *ACS Central Science*, 4(2): 268–276.

Hannah, L.; Powell, W.; and Blei, D. 2010. Nonparametric Density Estimation for Stochastic Optimization with an Observable State Variable. In *Advances in Neural Information Processing Systems 23*, 820–828. Vancouver, Canada.

Homem-de Mello, T. 2008. On rates of convergence for stochastic optimization problems under non-independent and identically distributed sampling. *SIAM Journal on Optimization*, 19(2): 524–551.

Huang, X.; Song, L.; Xue, K.; and Qian, C. 2023. Stochastic Bayesian optimization with unknown continuous context distribution via kernel density estimation. arXiv:2312.10423.

Husain, H.; Nguyen, V.; and Hengel, A. v. d. 2022. Distributionally Robust Bayesian Optimization with ϕ-divergences. *CoRR*, abs/2203.02128.

Inatsu, Y.; Takeno, S.; Karasuyama, M.; and Takeuchi, I. 2022. Bayesian optimization for distributionally robust chance-constrained problem. In *Proceedings of the 39th International Conference on Machine Learning*, 9602–9621. Baltimore, Maryland.

Iwazaki, S.; Inatsu, Y.; and Takeuchi, I. 2021. Mean-variance analysis in Bayesian optimization under uncertainty. In *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics*, 973–981. Virtual.

Jiang, H. 2017. Uniform convergence rates for kernel density estimation. In *Proceedings of the 34th International Conference on Machine Learning*, 1694–1703. Sydney, Australia.

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.

Kandasamy, K.; Krishnamurthy, A.; Schneider, J.; and Póczos, B. 2018a. Parallelised Bayesian optimisation via Thompson sampling. In *Proceedings of the 21st International Conference on Artificial Intelligence and Statistics*, 133–142. Playa Blanca, Spain.

Kandasamy, K.; Neiswanger, W.; Schneider, J.; Poczos, B.; and Xing, E. 2018b. Neural Architecture Search with Bayesian Optimisation and Optimal Transport. In *Advances in Neural Information Processing Systems 31*, 2020–2029. Montréal, Canada.

Kim, S.; Pasupathy, R.; and Henderson, S. G. 2015. A guide to sample average approximation. *Handbook of Simulation Optimization*, 207–243.

Kirschner, J.; Bogunovic, I.; Jegelka, S.; and Krause, A. 2020. Distributionally robust Bayesian optimization. In *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics*, 2174–2184. Virtual.

Kirschner, J.; and Krause, A. 2019. Stochastic bandits with context distributions. In *Advances in Neural Information Processing Systems 32*, 14090–14099. Vancouver, Canada.

Klein, A.; Falkner, S.; Bartels, S.; Hennig, P.; and Hutter, F. 2017. Fast Bayesian Optimization of Machine Learning Hyperparameters on Large Datasets. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, 528–536. Fort Lauderdale, FL.

Krause, A.; and Ong, C. 2011. Contextual Gaussian Process Bandit Optimization. In *Advances in Neural Information Processing Systems 24*, 2447–2455. Granada, Spain.

Kushner, H. J. 1964. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. *Journal of Basic Engineering*, 86(1): 97–106.

Lamprier, S.; Gisselbrecht, T.; and Gallinari, P. 2018. Profile-based bandit with unknown profiles. *Journal of Machine Learning Research*, 19(1): 2060–2099.

Marzat, J.; Walter, E.; and Piet-Lahanier, H. 2013. Worst-case global optimization of black-box functions through Kriging and relaxation. *Journal of Global Optimization*, 55(4): 707–727.

Namkoong, H.; and Duchi, J. C. 2017. Variance-based regularization with convex objectives. In *Advances in Neural Information Processing Systems 30*, 2971–2980. Long Beach, CA.

Nguyen, Q. P.; Dai, Z.; Low, B. K. H.; and Jaillet, P. 2021. Value-at-Risk Optimization with Gaussian Processes. In *Proceedings of the 38th International Conference on Machine Learning*, 8063–8072. Virtual.

Nguyen, T.; Gupta, S.; Ha, H.; Rana, S.; and Venkatesh, S. 2020. Distributionally robust Bayesian quadrature optimization. In *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics*, 2174–2184. Virtual.

Nocedal, J. 1980. Updating quasi-Newton matrices with limited storage. *Mathematics of Computation*, 35(151): 773–782.

Nogueira, J.; Martinez-Cantin, R.; Bernardino, A.; and Jamone, L. 2016. Unscented Bayesian optimization for safe robot grasping. In *Proceedings of the 26th IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1967–1972. Daejeon, Korea.

Oliveira, R.; Ott, L.; and Ramos, F. 2019. Bayesian optimisation under uncertain inputs. In *Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics*, 1177–1184. Okinawa, Japan.

Owen, A. B. 2003. Quasi-Monte Carlo sampling. *Monte Carlo Ray Tracing: Siggraph*, 1: 69–88.

Pérez, A.; Larrañaga, P.; and Inza, I. 2009. Bayesian classifiers based on kernel density estimation: Flexible classifiers. *International Journal of Approximate Reasoning*, 50(2): 341–362.

Qian, C.; Xiong, H.; and Xue, K. 2020. Bayesian optimization using pseudo-points. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence*, 3044–3050. Yokohama, Japan.

Rahimian, H.; and Mehrotra, S. 2022. Frameworks and Results in Distributionally Robust Optimization. *Open Journal of Mathematical Optimization*, 3: 1–85.

Rasmussen, C. E.; and Williams, C. K. I. 2006. *Gaussian Processes for Machine Learning*. The MIT Press.

Russo, D.; and Van Roy, B. 2014. Learning to optimize via posterior sampling. *Mathematics of Operations Research*, 39(4): 1221–1243.

Scott, D. W. 2015. *Multivariate Density Estimation: Theory, Practice, and Visualization*. John Wiley & Sons.

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.

Silverman, B. W. 1986. *Density Estimation for Statistics and Data Analysis*, volume 26. CRC press.

Song, L.; Xue, K.; Huang, X.; and Qian, C. 2022. Monte Carlo tree search based variable selection for high dimensional Bayesian optimization. In *Advances in Neural Information Processing Systems 35*, 28488–28501. New Orleans, LA.

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.

Surjanovic, S.; and Bingham, D. 2013. Virtual Library of Simulation Experiments: Test Functions and Datasets. Retrieved May 15, 2023, from http://www.sfu.ca/~ssurjano.

Tay, S. S.; Foo, C. S.; Daisuke, U.; Leong, R.; and Low, B. K. H. 2022. Efficient distributionally robust Bayesian optimization with worst-case sensitivity. In *Proceedings of the 39th International Conference on Machine Learning*, 21180–21204. Baltimore, Maryland.

Toscano-Palmerin, S.; and Frazier, P. I. 2022. Bayesian Optimization with Expensive Integrands. *SIAM Journal on Optimization*, 32(2): 417–444.

Tsybakov, A. B. 2008. *Introduction to Nonparametric Estimation*. Springer.

Wand, M. P.; and Jones, M. C. 1994. *Kernel smoothing*. CRC press.

Williams, B. J. 2000. *Sequential Design of computer experiments to minimize integrated response functions*. The Ohio State University.

Xie, J.; Frazier, P. I.; Sankaran, S.; Marsden, A.; and Elmohamed, S. 2012. Optimization of computationally expensive simulations with Gaussian processes and parameter uncertainty: Application to cardiovascular surgery. In *Proceedings of the 50th Annual Allerton Conference on Communication, Control, and Computing*, 406–413. Champaign, IL.