Journal of Machine Learning Research 26 (2025) 1-33 Submitted 5/22; Revised 12/23; Published 3/25

Adjusted Expected Improvement for Cumulative Regret Minimization in Noisy Bayesian Optimization

Shouri Hu HUSHOURI@UESTC.EDU.CN
School of Mathematical Sciences
University of Electronic Science and Technology of China, China

Haowei Wang HAOWEI_WANG@U.NUS.EDU
Department of Industrial Systems Engineering and Management
National University of Singapore, Singapore

Zhongxiang Dai DAIZHONGXIANG@CUHK.EDU.CN
School of Data Science
The Chinese University of Hong Kong, Shenzhen, China

Bryan Kian Hsiang Low LOWKH@COMP.NUS.EDU.SG
Department of Computer Science
National University of Singapore, Singapore

Szu Hui Ng ISENSH@NUS.EDU.SG
Department of Industrial Systems Engineering and Management
National University of Singapore, Singapore

Editor: Tor Lattimore

Abstract

The expected improvement (EI) is one of the most popular acquisition functions for Bayesian optimization (BO) and has demonstrated good empirical performances in many applications for the minimization of simple regret. However, under the evaluation metric of cumulative regret, the performance of EI may not be competitive, and its existing theoretical regret upper bound still has room for improvement. To adapt the EI for better performance under cumulative regret, we introduce a novel quantity called the *evaluation cost* which is compared against the acquisition function, and with this, develop the *expected improvement-cost* (EIC) algorithm. In each iteration of EIC, a new point with the largest acquisition function value is sampled, only if that value exceeds its evaluation cost. If none meets this criteria, the current best point is resampled. This evaluation cost quantifies the potential downside of sampling a point, which is important under the cumulative regret metric as the objective function value in every iteration affects the performance measure. We establish in theory a high-probability regret upper bound of EIC based on the maximum information gain, which is tighter than the bound of existing EI-based algorithms. It is also comparable to the regret bound of other popular BO algorithms such as Thompson sampling (GP-TS) and upper confidence bound (GP-UCB). We further perform experiments to illustrate the improvement of EIC over several popular BO algorithms.

Keywords: Cumulative regret, expected improvement, Gaussian processes, maximum information gain, noisy Bayesian optimization, regret upper bound

1. Introduction

Bayesian optimization (BO) is a sequential design framework for the global optimization of black-box functions with the following key characteristics: Firstly, the (objective) function has an unknown structure, with the input vectors residing in low- to moderate-dimensional Euclidean space. Secondly, the evaluation of such functions is expensive, hence it is impossible to search the entire domain in high precision before exhausting the budget. Thirdly, the derivative information of the objective function is unavailable or impractical to estimate, therefore classical gradient-based methods are not applicable. BO was initially studied by Kushner (1964), Mŏckus (1975), Žilinskas (1975) and Mŏckus et al. (1978), and later popularized by the work of Jones et al. (1998). Recently it has gained substantial attention in a large number of important areas such as engineering systems optimization and hyper-parameter tuning for machine learning algorithms (Torun et al., 2018; Kirschner et al., 2019; Letham et al., 2019; Sim et al., 2021).

Within the BO framework, the evaluated points are sequentially selected by maximizing an *acquisition function*, whose calculation requires a surrogate function to model the objective function using the sequentially collected points and their (either noisy or noise-free) observations. The most commonly used surrogate functions are the Gaussian processes (GPs). A GP is specified a priori by the mean function and the covariance kernel. Given a set of training data, the posterior of a GP remains a GP, with closed-form expressions for the posterior distribution at any point. This nice property makes the GPs become a powerful tool for statistical modelling. Refer to Williams and Rasmussen (2006), Osborne et al. (2009) and Malkomes et al. (2016) for a more detailed introduction to GPs.

BO has been extensively applied in numerous fields, and the goal of many of them is to find the best final solution, i.e., to find the point that maximizes the objective function. Under this goal, points with good chances to achieve the maximum function value are evaluated sequentially, and after the budget is exhausted, the evaluated point with the largest observed function value is usually reported as the final solution. Application fields with this goal include engineering system optimization (Jones et al., 1998; Torun et al., 2018), materials science design (Frazier and Wang, 2016; Packwood, 2017; Fukazawa et al., 2019), and pharmaceutical product development (Ban et al., 2017; Sano et al., 2020). This goal in BO is typically known as the minimization of the *simple regret*, and there has been several theoretical results on its properties in the literature (Grünewälder et al., 2010; Bull, 2011; Ryzhov, 2016; Wüthrich et al., 2021).

To minimize the simple regret of BO, a number of acquisition functions have been proposed. Among them, one of the most widely used acquisition functions is *expected improvement* (EI). EI was firstly proposed by Mŏckus (1975) under the noise-free BO setting, and then received further attention due to the work of Jones et al. (1998) who successfully incorporated GPs into the calculation of EI. As a conceptually intuitive method, EI has demonstrated impressive empirical performances in various applications. In every iteration, EI calculates the expected gain (over the best observed function value so far) from every point in the domain based on the posterior mean and variance of the GP model, and evaluates the point that maximizes this expected gain. Besides EI, other popular acquisition functions for simple regret minimization include *knowledge gradient* which evaluates the point that maximizes the increment of posterior mean function (Frazier et al., 2009; Wu

and Frazier, 2016), and *entropy search* which selects the point that is most informative about the location of the global optimum (Hennig and Schuler, 2012; Hernández-Lobato et al., 2014; Wang and Jegelka, 2017).

Despite the popularity of BO methods based on simple regret minimization, many common applications of BO are also concerned with the overall performance of BO throughout the entire experiment, instead of only finding the best final solution (i.e., minimizing the simple regret). One representative example of such applications is the recommendation system (Koren et al., 2009; Kawale et al., 2015; Galuzzi et al., 2020), where algorithms (such as matrix-factorization) are deployed to recommend items (e.g., news, movies and songs) to customers in order to increase their stickiness or likelihood to make a purchase. Like many machine learning algorithms, these recommendation algorithms require a hyper-parameter tuning process in order to enhance their effectiveness, for which BO is usually a prominent choice. The performance of recommendation algorithms, which is summarized as a numerical score to quantify the customers' utility within a particular time period, is evaluated on a regular time basis. As the experience of every customer over time matters, it is therefore inappropriate to only aim at finding a good recommendation algorithm for final future customers as this may deteriorate the utility of current customers. Hence, minimizing the simple regret is not a suitable objective in this case. Instead, companies usually aim to boost the total utility scores of all customers throughout the entire hyper-parameter tuning process. Another example of such applications is the development of combination therapies through clinical trials, where BO is often used to sequentially choose combinations of therapies for a series of patients, in order to improve the treatment efficacy (Shahriari et al., 2016; Kharkovskii et al., 2020; Takahashi and Suzuki, 2021). In these applications, the focus is on the overall efficacy of combination therapies for all patients, and not only on the final patient.

This objective of maximizing the overall performance throughout the entire experiment originated from the multi-armed bandit (MAB) literature, in which it is known as *cumulative regret* minimization (Lai and Robbins, 1985; Agrawal, 2019; Lattimore and Szepesvári, 2020). Popular MAB approaches for this objective include upper confidence bound (UCB, Lai and Robbins, 1985; Auer et al., 2002; Cappé et al., 2013; Lattimore, 2018) and Thompson sampling (TS, Thompson, 1933; Agrawal and Goyal, 2012; Korda et al., 2013), which are based on the frequentist and Bayesian perspectives, respectively. Both approaches have been extended into the BO framework to derive the GP-UCB and GP-TS acquisition functions, and have been shown to perform well analytically under the cumulative regret setting (Srinivas et al., 2010; Chowdhury and Gopalan, 2017; Kandasamy et al., 2018; Berkenkamp et al., 2019; Vakili et al., 2021). The acquisition function of EI has also been analyzed analytically under the cumulative regret evaluation metric (Wang and De Freitas, 2014; Nguyen et al., 2017), and its existing performance bound is shown to be not as tight compared to those of GP-UCB and GP-TS (Chowdhury and Gopalan, 2017). This theoretical gap in the cumulative regret of EI is also reflected in practice, which we illustrate using a numerical example in Figure 1. Here, we examine the performances of EI and GP-UCB on the commonly used Ackley function, which is a two-dimensional test function equipped with many local extrema (refer to Table 1 for more details). The figure shows that GP-UCB (green curve), which is designed for the cumulative regret setting, incurs smaller cumulative regret than EI (orange curve).

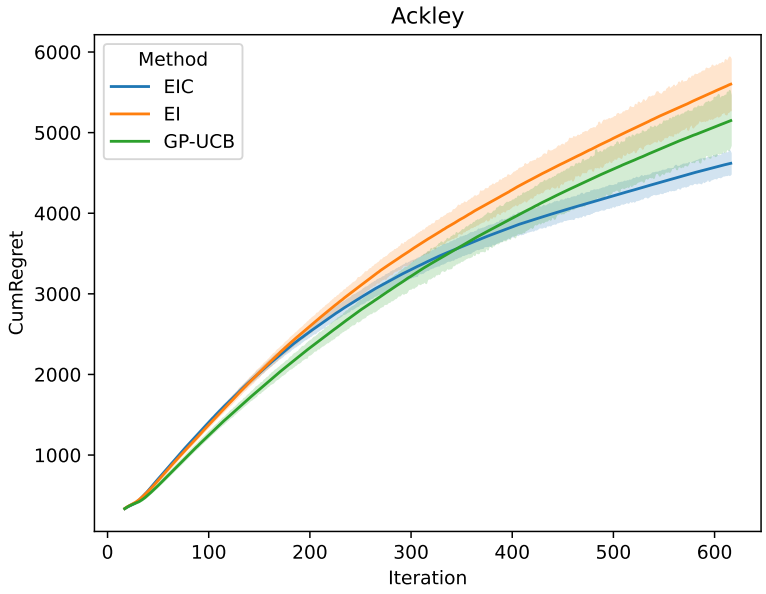


Figure 1: Cumulative regret of common BO methods and our proposed EIC algorithm on the Ackley test function. The solid line represents the cumulative regret (averaged over 100 independent runs) and the shaded area is the corresponding 95% confidence region.

In view of the above-mentioned theoretical and empirical gaps of the EI in cumulative regret minimization, the following question arises: *can EI be adapted to achieve a tight upper bound on its cumulative regret which can then enable it to also perform well in applications focused on maximizing the overall performance?* This is in fact an important open problem in BO because of the impressive real-world performances and wide adoption of EI. To this end, we adapt the traditional EI algorithm to suit the objective of cumulative regret minimization, and propose the *expected improvement-cost* (EIC) algorithm (Section 3). We plot in Figure 1 the cumulative regret of EIC (blue curve), which shows that our proposed EIC achieving a smaller cumulative regret than traditional EI and performing comparably with GP-UCB.

The contributions of this paper are as follows. **Firstly**, we propose the EIC algorithm (Section 3) for cumulative regret minimization. EIC is designed to consciously balance the evaluation gains and losses in every iteration, so that a smaller cumulative regret is achieved. The algorithm starts with a systematic and budget-dependent initial experimental design, which ensures that the global model fitting is reasonably good. After the initialization, to choose a point to evaluate, EIC firstly calculates the evaluation cost of every point in the domain based on the expected loss (and the number of remaining iterations). This serves as a criterion to decide the worthiness of a point for evaluation, i.e., a point is worth evaluating only if its expected gain (i.e., the EI acquisition function value) is larger than its evaluation cost. As a result, in every iteration, we sample the point that has the largest EI acquisition function value, provided that it is not lower than its evaluation cost. If no point in the domain has an acquisition function value not lower than its evaluation cost, then the

previous evaluated point with the best observed function value is sampled. **Secondly**, we analyze the cumulative regret of our EIC algorithm and establish a high-probability finite-time regret upper bound. Importantly, we show that EIC can achieve an upper bound of $O\big(\sqrt{N}\gamma_N(\log N)^{1/2}\big)$ under some mild regularity conditions, where N is the total number of iterations and γ_N is the maximum information gain (see Section 3.2). This regret upper bound is tighter than the bound of existing EI-based algorithms in the literature (Wang and De Freitas, 2014; Nguyen et al., 2017).

The layout of the paper is as follows. In Section 2, we present some background information about BO and the GP model. We also give a brief review of the EI acquisition function and its related theoretical results. In Section 3, we describe our proposed EIC algorithm and explain some interesting insights into the development of EIC, including why the evaluation cost is applied. In Section 4, we establish a cumulative regret upper bound for EIC. In Section 5 we perform several numerical experiments to demonstrate the practical effectiveness of our EIC algorithm. Finally, Section 6 concludes the paper.

2. Problem Statement and Background

Let A^T denote the transpose of a vector or matrix A and let I_n denote the identity matrix of size n. Let $\text{diag}(a_1, a_2, \ldots, a_n)$ denote the diagonal matrix of size n with the (i,i) entry equal to a_i. Let $\|\mathbf{x}\| := \sqrt{x_1^2 + \cdots + x_d^2}$ denote the Euclidean norm of vector $\mathbf{x} = (x_1, \ldots, x_d)^T$. Let a^+ denote $\max(0, a)$ and let $\phi(\cdot)$ and $\Phi(\cdot)$ denote the density and cumulative distribution functions of standard normal distribution, respectively. Let $\mathcal{N}_k(\boldsymbol{\mu}, \Sigma)$ denote the k-dimensional multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ. Let $a_n \sim b_n$ if $\lim_{n\to\infty}(a_n/b_n) = 1$, $a_n = O(b_n)$ if $\limsup_{n\to\infty} |a_n/b_n| < \infty$ and $a_n = \Omega(b_n)$ if $\liminf_{n\to\infty} |a_n/b_n| > 0$.

BO aims to sequentially maximize an unknown objective function $f : D \to \mathbb{R}$, where $D = [0,1]^d \subseteq \mathbb{R}^d$. In each iteration n, BO selects a point $\mathbf{x}_n$ to evaluate, and receives a noisy observation $y_n = f(\mathbf{x}_n) + \epsilon_n$. As discussed in Section 1, we consider the objective of minimizing the cumulative regret after N total iterations:

$$R_N = \sum_{n=1}^{N} (f(\mathbf{x}^*) - f(\mathbf{x}_n)), \tag{1}$$

where $\mathbf{x}^* = \arg\max_{\mathbf{x}\in D} f(\mathbf{x})$ denotes the location where f attains the global maximum. In order to choose the sequential points $\mathbf{x}_n$'s intelligently, BO usually models the objective function using a GP model. We briefly introduce GPs in Section 2.1, and refer the readers to Williams and Rasmussen (2006) and Kanagawa et al. (2018) for more complete introduction to GPs.

2.1 Gaussian Processes

The objective function in BO is typically modelled as a stationary Gaussian process. We say a random function f follows a time-varying prior distribution $\mathcal{GP}(\mu, \omega_n^2 k)$ with mean function $\mu : D \to \mathbb{R}$, covariance kernel $k : D \times D \to \mathbb{R}$ and signal variance ω_n^2 if and only if the following condition holds: For every finite set of points $X = (\mathbf{x}_1, \ldots, \mathbf{x}_n)^T$, the values

$\big(f(\mathbf{x}_1),\ldots,f(\mathbf{x}_n)\big)^T \sim \mathcal{N}_n(\mu_X, \omega_n^2 K_{XX})$, which is an n-dimensional multivariate normal distribution with

$$\mu_X := \big(\mu(\mathbf{x}_1),\ldots,\mu(\mathbf{x}_n)\big)^T, \text{ and} \tag{2}$$

$$K_{XX} := \begin{pmatrix} k(\mathbf{x}_1,\mathbf{x}_1) & k(\mathbf{x}_1,\mathbf{x}_2) & \cdots & k(\mathbf{x}_1,\mathbf{x}_n) \\ k(\mathbf{x}_2,\mathbf{x}_1) & k(\mathbf{x}_2,\mathbf{x}_2) & \cdots & k(\mathbf{x}_2,\mathbf{x}_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(\mathbf{x}_n,\mathbf{x}_1) & k(\mathbf{x}_n,\mathbf{x}_2) & \cdots & k(\mathbf{x}_n,\mathbf{x}_n) \end{pmatrix}. \tag{3}$$

One useful result of the GP model is that the posterior distribution of f, conditioned on the sampled data, is still a GP. Let $X = (\mathbf{x}_1,\ldots,\mathbf{x}_n)^T$ and $Y = (y_1,\ldots,y_n)^T$ denote the sampled points and the corresponding noisy observations up to iteration n. At each iteration, the noise terms are assumed bo be independent and identically distributed (i.i.d) normal random variables $\epsilon_1,\ldots,\epsilon_n \overset{i.i.d}{\sim} \mathrm{N}(0,\lambda^2\omega_n^2)$. Consequently, we have

$$f|(X,Y) \sim \mathcal{GP}(\mu_n, \omega_n^2 k_n), \tag{4}$$

with $\mu_n : D \to \mathbb{R}$ and $k_n : D \times D \to \mathbb{R}$ given by

$$\mu_n(\mathbf{x}) = k_{\mathbf{x}X}(K_{XX} + \lambda^2 I_n)^{-1}(Y_n - \mu_X), \tag{5}$$

$$k_n(\mathbf{x},\mathbf{x}') = k(\mathbf{x},\mathbf{x}') - k_{\mathbf{x}X}(K_{XX} + \lambda^2 I_n)^{-1} k_{X\mathbf{x}'}, \tag{6}$$

where $k_{\mathbf{x}X} = k_{X\mathbf{x}}^T = \big(k(\mathbf{x}_1,\mathbf{x}),\ldots,k(\mathbf{x}_n,\mathbf{x})\big)$. Note that ω_n^2 and λ^2 are algorithm-specific parameters, which could possibly depend on n. Based on (4), it can be deduced that the predictive distribution at any point follows a Gaussian distribution. That is, for any $\mathbf{x} \in D$,

$$f(\mathbf{x})|(X,Y) \sim \mathrm{N}(\mu_n(\mathbf{x}), \omega_n^2\sigma_n^2(\mathbf{x})), \text{ with } \sigma_n^2(\mathbf{x}) = k_n(\mathbf{x},\mathbf{x}). \tag{7}$$

The mean function and the covariance kernel serve as the prior for the GP model, and they reflect the initial belief about f. Without loss of generality, the mean function $\mu(\mathbf{x})$ is usually set to be zero, indicating there is no prior knowledge on the global maxima location, whereas the choice of covariance kernel is more varied. In this paper, we consider positive definite covariance kernel functions that are isotropic and bounded, which is a quite general setting. Typical examples that satisfy this condition include the squared exponential (SE) and the Matérn kernel, which are perhaps the most popular covariance kernels in practice for Bayesian optimization (Snoek et al., 2012; Shahriari et al., 2016; Teckentrup, 2020). Let $\mathbf{h} := (h_1,\ldots,h_d)^T$ with $h_i > 0, \forall 1 \le i \le d$ be the length-scale parameter. The Euclidean distance between $\mathbf{x}$ and $\mathbf{x}'$, adjusted by the length-scale parameter $\mathbf{h}$, is given by

$$\|\mathbf{x}-\mathbf{x}'\|_{\mathbf{h}} := \sqrt{\Big(\frac{x_1-x_1'}{h_1}\Big)^2 + \cdots + \Big(\frac{x_d-x_d'}{h_d}\Big)^2}.$$

Define the SE covariance kernel as

$$k_{\mathrm{SE}}(\mathbf{x},\mathbf{x}') = \exp\left(-\frac{\|\mathbf{x}-\mathbf{x}'\|_{\mathbf{h}}^2}{2}\right), \tag{8}$$

and the Matérn kernel

$$k_{\text{Matérn}}(\mathbf{x}, \mathbf{x}') = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu}\|\mathbf{x} - \mathbf{x}'\|_{\mathbf{h}}\right)^{\nu} B_{\nu} \left(\sqrt{2\nu}\|\mathbf{x} - \mathbf{x}'\|_{\mathbf{h}}\right), \tag{9}$$

where $\nu > 0$ is the smoothness parameter, and B_ν is the modified Bessel function. The length-scale $\mathbf{h}$ and smoothness ν are the hyper-parameters of the GP model. Standard methods can be applied to estimate them, such as the maximum likelihood method by Santner et al. (2018) and the maximum a posteriori method by Ng and Yin (2012).

2.2 Expected Improvement

EI is one of the most popular acquisition functions under BO framework. We review here EI as proposed by Mŏckus (1975) and Jones et al. (1998) when the observations are evaluated without noise, i.e., $y_m = f(\mathbf{x}_m)$. The noisy version will be discussed in more detail in Section 3. Let $E_n[\cdot]$ denote the expectation with respect to the posterior distribution of $f(\mathbf{x})$ described in (7). After n iteration, the acquisition function of EI is defined as

$$\alpha_n^{EI}(\mathbf{x}) := E_n([f(\mathbf{x}) - f(\mathbf{x}_n^*)]^+), \tag{10}$$

where $f(\mathbf{x}_n^*) = \max_{1 \le m \le n} f(\mathbf{x}_m)$ is the current best observation value. This acquisition function (10) quantifies the expected gain of sampling $\mathbf{x}$ over the current best observation value, and a point that maximizes (10) will be sampled in the next iteration, i.e.,

$$\mathbf{x}_{n+1} = \arg\max_{\mathbf{x} \in D} \alpha_n^{EI}(\mathbf{x}).$$

EI has been repeatedly shown to perform competitively under the simple regret evaluation metric. Moreover, it is computationally convenient because the acquisition function (10) has a closed-form expression.

2.3 Related Works

In the theoretical analyses of BO algorithms, the objective function is usually assumed to be fixed and belongs to the reproducing kernel Hilbert space (RKHS) associated with the covariance kernel in the GP model. Based on the different focus or interest in the application, the analyses in literature can be classified into two categories: one which focuses on the analysis of the asymptotic properties of the obtained optimum or optimal solution (i.e., the simple regret), and the other which focuses on analysing the cumulative regret. Intuitively, the analysis of asymptotic properties (i.e., simple regret) can be used to assess if the algorithm converges and how fast it converges to the true optimum or optimal solution, whereas the cumulative regret analysis informs on the performance of the algorithm throughout the entire experiment.

The theoretical analyses of the EI started from studying its asymptotic properties (i.e., simple regret). Under the noise-free BO setting, Vazquez and Bect (2010) showed that when the covariance kernel function is fixed and has finite smoothness, the EI converges to the global maxima almost surely for any objective function. Subsequently, Bull (2011) proved that for any d-dimensional objective function, the convergence rate of the EI is of the order $O(N^{-(\nu \wedge 1)/d})$ with $\nu > 0$ being the smoothness parameter of the covariance kernel. The

author also showed that a combination of EI and the ϵ-greedy method converges at the near-optimal rate of $O(N^{-\nu/d})$. Under the noisy setting, Ryzhov (2016) showed that a variant of EI can achieve the same convergence rate as the optimal computing budget allocation (OCBA) algorithm, which is an algorithm that has been shown to be asymptotically near-optimal.

Srinivas et al. (2010) did the first cumulative regret analysis on BO algorithms. They showed that when the objective function belongs to the RKHS associated with certain covariance kernel, the acquisition function of GP-UCB can achieve a finite-time regret upper bound of $R_N = O\Big(\gamma_N\sqrt{N}(\log N)^{3/2}\Big)$ with high probability. The term γ_N is the maximum information gain about the objective function f that can be obtained from any set of N sampled points. Its value is algorithm-independent and closely related to the effective dimension associated with the kernel. The regret bound of GP-UCB was later improved by Chowdhury and Gopalan (2017) to $R_N = O\Big(\gamma_N\sqrt{N}\Big)$. In addition, they proposed the acquisition function of GP-TS and showed that it can attain a regret upper bound of $R_N = O\Big(\gamma_N\sqrt{N}(\log N)^{1/2}\Big)$. As for the EI acquisition function under the noisy BO setting, Wang and De Freitas (2014) replaced $f(\mathbf{x}_n^*)$ in (10) with the maximum of the posterior mean $\max_{\mathbf{x}\in D}\mu_n(\mathbf{x})$, and showed that their algorithm achieves $R_N = O\Big(\gamma_N^{3/2}\sqrt{N}(\log N)\Big)$ with high probability. Nguyen et al. (2017) further considered replacing $f(\mathbf{x}_n^*)$ with the current best observation $\max_{1\leq m\leq n} y_m$, and reduced the regret upper bound to $R_N = O\Big(\gamma_N\sqrt{N}(\log N)^{3/2}\Big)$. However, their analysis depends on a pre-defined constant $\kappa > 0$ which is set to be small for good performances of EI, yet the constant term in their regret upper bound will explode quickly as $\kappa \to 0$. The recent work of Tran-The et al. (2022) also performed theoretical analyses of EI, but they considered a regret definition that is different from our paper and these previous works.

Scarlett et al. (2017) and its update Scarlett et al. (2018) derived a universal regret lower bound that applies to all algorithms under the noisy BO setting. The authors showed that for the SE and Matérn kernel respectively, an expected regret of $E(R_N) = \Omega\Big(\sqrt{N}(\log N)^{d/4}\Big)$ and $E(R_N) = \Omega\Big(N^{\frac{\nu+d}{2\nu+d}}\Big)$ is unavoidable. This strong result motivates us to examine whether the current regret upper bounds of EI algorithms are tight. As shown from these previous results in the literature, for the EI, there exists a gap between the regret lower bound and the current regret upper bounds. In order to reduce this gap, we introduce a novel quantity called the evaluation cost, and compare it with the EI acquisition function. At each iteration, a new point will be sampled only if its acquisition function exceeds its evaluation cost. We name this new algorithm the expected improvement-cost (EIC) algorithm. Moreover, we adopt an analysis framework from the BO literature and establish in Section 4 a finite-time regret upper bound of EIC as $O\Big(\sqrt{N}\gamma_N(\log N)^{1/2}\Big)$, which is tighter than the regret upper bound of Wang and De Freitas (2014) and Nguyen et al. (2017).

3. The Expected Improvement-Cost (EIC) algorithm

In this section we describe in detail our expected improvement-cost (EIC) algorithm. We first introduce in Section 3.1 an experiment scheme for choosing the initial design points. This scheme is to ensure that the initial GP model fit is good so that the posterior variance is not too large uniformly over the entire domain. Next, we discuss in Section 3.2 how the incumbent function of EI should be selected in the setting of noisy BO, i.e., when the observations are corrupted by homogenous noises. The EI has a built-in mechanism to trade-off between sampling points with high expected value (posterior mean) versus high uncertainty (posterior variance), which is a desirable property under the cumulative regret objective. However, the EI acquisition function, as it was originally designed, only quantifies the potential upside of evaluating a point and overlooks the potential downside. Under the objective of cumulative regret, it is important to also account for the potential downside of evaluating a point, because the value of the objective function in every iteration contributes to the performance measure (i.e., the cumulative regret). Hence, this requires the algorithm to be more conservative than under the simple regret objective. To this end, in Section 3.3, we propose to quantify this potential downside using an *evaluation cost* function, which helps determine whether our EIC algorithm should evaluate existing good points or explore new points.

3.1 Initial Experiment Scheme

In the theoretical framework, the role of the initial experiment scheme is less significant, provided that the total number of initial design points does not substantially contribute to the regret. However, from a practical perspective, a well-structured initial experiment scheme is crucial for the success of Bayesian Optimization. As highlighted in Bull (2011), inappropriate selection of the initial design points may cause the algorithm to fail completely. Therefore, when practitioners use our algorithm, it is essential to establish a proper initial design. In this paper, we adopt an initial experimental design that evenly spaces design points across the domain $D = [0, 1]^d$. Define the collection of the initial design points B_M, indexed by $M \in \mathbb{N}$, as

$$B_M = \left\{ \mathbf{x} \in D : x_i = \frac{2k_i - 1}{2M} \text{ for } k_i \in \{1, \dots, M\}, i = 1, \dots, d \right\}. \tag{11}$$

Following this scheme, given a user-specified parameter $M \in \mathbb{N}$, the interval $[0, 1]$ along each dimension is divided into M equal-sized segments. As a result, the domain D is partitioned into M^d hyper-cubes, and the initial design points in B_M are located at the centre of these hyper-cubes. According to (11), the corresponding total number of initial design points is $n_0 = M^d$.

The reason for adopting such an initial experiment scheme is to control the overall posterior variance of the GP model by the following: For any point $\mathbf{x}$, it can be shown that $\sigma_n^2(\mathbf{x})$ increases with the Euclidean distance $\min_{1 \le j \le n} \|\mathbf{x} - \mathbf{x}_j\|$. That is, the posterior variance has a positive correlation with the Euclidean distance to its nearest sampled point. With the initial experiment scheme (11), it can be seen that its nearest sampled point is the centre of hyper-cube to which it belongs. As each hyper-cube has a diagonal length of $\sqrt{d}M^{-1}$, it is guaranteed that the distance is not more than $\frac{1}{2}\sqrt{d}M^{-1}$ for all points.

Therefore after our initial experiment design scheme, the posterior variance of the GP model is uniformly controlled by M, which is the number of segments along each dimension.

The next issue is to determine the value of M. Setting M too large will make the algorithm focus too much on global modelling, leaving little budget for optimization; On the other hand, setting M too small will result in a bad global model which can substantially diminish the sample efficiency. Hence, we need to seek a balance in choosing M so that the algorithm will converge fast enough without utilizing too much budget for the initial design. We suggest setting $M = cN^{1/2d}$ for some constant $c > 0$. This choice ensures that

$$n_0 = c^d N^{1/2}. \tag{12}$$

3.2 Acquisition Function

The general form of EI acquisition function, after n observations has been evaluated is

$$\alpha_n^{EI}(\mathbf{x}) = E_n([f(\mathbf{x}) - \xi_n]^+) = (\mu_n(\mathbf{x}) - \xi_n)\Phi\big(\tfrac{\mu_n(\mathbf{x})-\xi_n}{\omega_n\sigma_n(\mathbf{x})}\big) + \omega_n\sigma_n(\mathbf{x})\phi\big(\tfrac{\mu_n(\mathbf{x})-\xi_n}{\omega_n\sigma_n(\mathbf{x})}\big), \tag{13}$$

where ξ_n is the incumbent value and ω_n^2 is the signal variance parameter. In the noise-free BO setting, the incumbent ξ_n is usually selected as the current best observed function value: $\xi_n = \max_{1\le i\le n} f(\mathbf{x}_i)$. When the observations are corrupted by homogenous noises, the noise-free function values are not observable, and a natural replacement is the current best (noisy) observation: $\xi_n = \max_{1\le i\le n} y_i$. However, this can make the acquisition function very unstable due to the observation noises. To address this issue, Huang et al. (2006) proposed to multiply (13) by a factor of $\sigma_n(\mathbf{x})/\sqrt{\sigma_n^2(\mathbf{x}) + \lambda^2}$, so that the acquisition function is discounted according to the parameter λ^2. Another possible remedy is to rely on the GP model, for example Brochu et al. (2010) and Wang and De Freitas (2014) suggested to use the best posterior mean $\max_{\mathbf{x}\in D} \mu_n(\mathbf{x})$ as the incumbent, and Picheny et al. (2013) recommended a quantile-based incumbent $\max_{1\le i\le n} \mu_n(\mathbf{x}_i) - \Phi^{-1}(\beta)\sigma_n(\mathbf{x}_i)$ with quantile level parameter $\beta = 0.5$ or 0.9. In our EIC algorithm, we adopt the approach of Brochu et al. (2010) and Wang and De Freitas (2014), but adapt it by considering the current best posterior mean at only the observed locations as the incumbent. Specifically, we define the incumbent value after n observations as:

$$\xi_n = \max_{1\le i\le n} \mu_n(\mathbf{x}_i). \tag{14}$$

This choice of incumbent is intuitive and allows us to derive our theoretical guarantee. Intuitively, if the GP model fitting is good, then ξ_n will be a good estimation of the current best function value with low variability. Moreover, unlike the incumbent used in Wang and De Freitas (2014), (14) does not require optimizing the GP posterior mean over the entire domain and hence does not introduce excessive computational cost to the algorithm and uncertainty.

The implementation of BO algorithms and their regret bounds often depends on the maximal information gain γ_n between n observations and the underlying GP model, which is defined as

$$\gamma_n := \sup_{A_n\subset D} I(Y_{A_n}; F_{A_n}), \tag{15}$$

where $A_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n)^T$ represents **any** set of points in D. The term $I(Y_{A_n}; F_{A_n})$ refers to the mutual information between $F_{A_n} = (f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n))^T$ and $Y_{A_n} = (y_1, \ldots, y_n)^T$ with each $y_i = f(\mathbf{x}_i) + \boldsymbol{\epsilon}_i$. As discussed in Section 2.1, F_{A_n} follows a multivariate normal distribution, and the noise terms are modelled as $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n \overset{i.i.d}{\sim} N(0, \lambda^2\omega_n^2)$, this allows the mutual information to be expressed in closed form as

$$I(Y_{A_n}; F_{A_n}) = \tfrac{1}{2} \log \det(\mathbf{I}_n + \lambda^{-2} K_{A_n A_n}). \tag{16}$$

There are two remarks regrading the maximal information gain. First, from definition (15) and equation (16), it is evident that the mutual information, and consequently the γ_n, does not depend on ω_n. Second, it is important to note that identifying the set A_n that maximizes the information gain is an NP-hard problem. However, γ_n can be upper bounded by a known quantity. For instance, as demonstrated in Srinivas et al. (2010), if the set A'_n is generated using the greedy algorithm, then $\gamma_n \leq (1 - 1/e)^{-1} I(Y_{A'_n}; F_{A'_n})$.

The signal variance parameter ω_n^2 in the GP model is used to balance the exploration and exploitation of the EI algorithm. When ω_n is large, EI tends to explore regions which have less points, whereas a small ω_n makes EI prefer exploiting regions which are predicted to have good points based on the observations so far. The signal variance parameter has been studied in many previous literature (Agrawal and Goyal, 2012; Wang and De Freitas, 2014; Chowdhury and Gopalan, 2017; Tran-The et al., 2022). Inspired by Theorem 1, we found that setting

$$\omega_n = c_0 \sqrt{\gamma_n + 1 + \log(1/\delta)} \tag{17}$$

with user-specified constants $c_0 > 0$ and $0 < \delta < 1$ guarantees the convergence of EIC.

3.3 Evaluation Cost and the EIC Algorithm

Under the traditional EI framework, a point with the largest acquisition function value (13) will be sampled in each iteration. The acquisition function (13) quantifies the expected gain over ξ_n if the point $\mathbf{x}$ is sampled in the next iteration, which is the potential upside of evaluating this point. However, this traditional strategy does not take into account the potential downside of sampling $\mathbf{x}$, which is important under cumulative regret considerations. This is because the function value in every iteration contributes to the cumulative regret performance measure. As a result, if the algorithm samples a point with a significantly inferior function value (i.e., with substantial downside), the cumulative regret will increase significantly. In this regard, in order to achieve a smaller cumulative regret, our EIC takes an additional step of comparing the acquisition function value against an *evaluation cost*, which allows us to better balance between the evaluation gains and losses. Specifically, we define the evaluation cost after n observations as:

$$\begin{aligned} L_n(\mathbf{x}) &:= E_n([\xi_n - f(\mathbf{x})]^+)/(N-n) \\ &= \left[(\xi_n - \mu_n(\mathbf{x}))\Phi\left(\tfrac{\xi_n - \mu_n(\mathbf{x})}{\omega_n \sigma_n(\mathbf{x})}\right) + \omega_n \sigma_n(\mathbf{x}) \phi\left(\tfrac{\xi_n - \mu_n(\mathbf{x})}{\omega_n \sigma_n(\mathbf{x})}\right)\right]/(N-n). \end{aligned} \tag{18}$$

The numerator here quantifies the expected loss of sampling $\mathbf{x}$ if its function value is less than the incumbent ξ_n, and the denominator is the number of remaining iterations. We will first describe our complete EIC algorithm in the next paragraph, following which we will explain the intuitions behind the evaluation cost (18).

Algorithm 1 Expected improvement-cost (EIC) algorithm

Require: N, $\mathcal{GP}(\mu, k)$, n_0, ξ_n, ω_n^2.
  Sample n_0 initial design points as described in (11). Each point is sampled with one replication.
  for $n = n_0, \ldots, N-1$ **do**
    Update the GP posterior model $\mathcal{GP}(\mu_n, k_n)$ using the history of observations.
    if $\exists \mathbf{x} \in D : \alpha_n^{EI}(\mathbf{x}) \geq L_n(\mathbf{x})$ **then**
      Let $B_n = \{\mathbf{x} \in D : \alpha_n^{EI}(\mathbf{x}) \geq L_n(\mathbf{x})\}$.
      Select the point $\mathbf{x}_{n+1} = \arg\max_{x:x \in B_n} \alpha_n^{EI}(\mathbf{x})$.
    else
      Select the point $\mathbf{x}_{n+1} = \arg\max_{1 \leq i \leq n} \mu_n(\mathbf{x}_i)$.
    end if
    Evaluate the selected point $\mathbf{x}_{n+1}$ with one replication and observe y_{n+1}.
    Add the newly collected $(\mathbf{x}_{n+1}, y_{n+1})$ to the history of observations.
  end for

Algorithm 1 presents the pseudo-code of our EIC algorithm. It has 5 input parameters: the total budget N, the prior GP model, the total number of initial design points n_0 and the incumbent function ξ_n and the signal variance parameter ω_n^2. Note that ω_n is a pre-determined whole sequence, whereas ξ_n is a sequence of functions that depend on the sampled data. EIC starts with the initial experiment scheme (Section 3.1), where a total number of n_0 different points are sampled. These points are pre-determined by the initial design scheme (11) and each point is evaluated with one replication. After the initial experiment, the subsequent points are sampled based on the acquisition function and evaluation cost. In each iteration n, if there exists at least one point $\mathbf{x}$ which satisfies the condition of $\alpha_n^{EI}(\mathbf{x}) \geq L_n(\mathbf{x})$ (i.e., its acquisition function value is not smaller than its evaluation cost), then we select the point with the largest acquisition function value among all points that satisfy this condition, and evaluate it with one replication. If no point satisfies this condition, we select the previously sampled point with the largest posterior mean and add one more replication run to that point.

Intuitions behind the Evaluation Cost. We use the case where the observations are noise-free to better illustrate the intuitions. In the noise-free case, the incumbent $\xi_n = \max_{1 \leq m \leq n} f(\mathbf{x}_m)$ is the current best function value. In iteration n, suppose we decide not to evaluate any new point but instead continue sampling at the current best point until the budget is exhausted, then the cumulative regret of the remaining iterations will be $(N-n)(f(\mathbf{x}^*) - \xi_n)$. On the other hand, we can choose to sample a new point $\mathbf{x}_{n+1}$ in the next iteration, and then stop evaluating any new point afterwards. For this newly sampled point $\mathbf{x}_{n+1}$, its function value $f(\mathbf{x}_{n+1})$ may be larger or smaller than ξ_n. In the first case where $f(\mathbf{x}_{n+1}) \geq \xi_n$, an improvement is achieved and its expected value is given as $E_n([f(\mathbf{x}_{n+1}) - \xi_n]^+)$. In this case, we can choose to keep sampling at $\mathbf{x}_{n+1}$ instead of $\arg\max_{1 \leq m \leq n} f(\mathbf{x}_m)$ in all the $(N-n)$ remaining iterations and this will provide a higher reward than with the continued sampling at ξ_n. Hence, the total expected reduction of the

cumulative regret (i.e., total expected gain) in all $(N-n)$ remaining iterations will be

$$(N-n)E_n([f(\mathbf{x}_{n+1})-\xi_n]^+) = (N-n)\alpha_n^{EI}(\mathbf{x}_{n+1}).$$

In the second case where $f(\mathbf{x}_{n+1}) < \xi_n$, a loss will be incurred and the expected value will be $E_n([\xi_n - f(\mathbf{x}_{n+1})]^+)$. Unlike the improvement which can be exploited for all $(N-n)$ remaining iterations as mentioned above, we only suffer this loss once (in iteration $n+1$) since we can switch back to the strategy of sampling $\arg\max_{1\leq m\leq n} f(\mathbf{x}_m)$ for all future iterations to avoid further this loss. This provides us a balance, and we see that $\mathbf{x}_{n+1}$ is therefore worth sampling if and only if its *total* expected gain is larger than its *total* expected loss:

$$(N-n)\alpha_n^{EI}(\mathbf{x}_{n+1}) \geq E_n([\xi_n - f(\mathbf{x}_{n+1})]^+),$$

or equivalently: $\alpha_n^{EI}(\mathbf{x}_{n+1}) \geq L_n(\mathbf{x}_{n+1})$.

Another interesting insight is that the evaluation cost $L_n(\mathbf{x}_{n+1})$ (18) increases with the iteration number n. Therefore, our EIC algorithm is endowed with a built-in mechanism which allows it to more aggressively explore new points at the beginning and then gradually become more exploitative as the budget gets exhausted. Overall, our EIC algorithm is more conservative (in sampling new points) than the traditional EI as it evaluates the potential gains and losses by comparing the acquisition function with the evaluation costs. Further note that the traditional EI can be regarded as a special case of EIC algorithm with zero evaluation cost.

4. Regret Analysis of EIC

In this section we perform regret analysis of our proposed EIC algorithm. We adopt the frequentist view by assuming that f is an arbitrary function from the reproducing kernel Hilbert space (RKHS) associated with the covariance kernel in the GP model. In Section 4.1, we present a brief introduction to RKHS and its connections with GP models. A complete overview of RKHS can be found in Berlinet and Thomas-Agnan (2011). In Section 4.2, we establish a finite-time cumulative regret upper bound for EIC.

4.1 Reproducing Kernel Hilbert Space

Let $\mathcal{X}$ be a non-empty set and $k(\cdot,\cdot) : \mathcal{X}\times\mathcal{X} \to \mathbb{R}$ be a symmetric positive definite kernel. Common examples of symmetric positive definite kernels are the SE and Matérn covariance kernel. A Hilbert space $\mathcal{H}_k$ of functions on $\mathcal{X}$ equipped with an inner-product $\langle\cdot,\cdot\rangle_{\mathcal{H}_k}$ is called a *reproducing kernel Hilbert space* (RKHS) with reproducing kernel k if the following conditions are satisfied:

1. For all $\mathbf{x}\in\mathcal{X}$, we have $k(\cdot,\mathbf{x}) \in \mathcal{H}_k$;
2. For all $\mathbf{x}\in\mathcal{X}$ and all $f\in\mathcal{H}_k$, we have

$$f(\mathbf{x}) = \langle f, k(\cdot,\mathbf{x})\rangle_{\mathcal{H}_k}. \tag{19}$$

Based on the definition of RKHS, it can be deduced that for all $\mathbf{x},\mathbf{x}'\in\mathcal{X}$, the kernel function $k(\mathbf{x},\mathbf{x}')$ can be written as:

$$k(\mathbf{x},\mathbf{x}') = \langle k(\cdot,\mathbf{x}), k(\cdot,\mathbf{x}')\rangle_{\mathcal{H}_k}. \tag{20}$$

(20) suggests that every RKHS defines a reproducing kernel function k that is both symmetric and positive definite. The other direction also holds as shown by the Moore-Aronszajn theorem (Aronszajn, 1950), which states that given a positive definite kernel k, we can construct a unique RKHS of real-valued functions with k as its reproducing kernel function. That is, RKHSs and positive definite kernels are one-to-one: for every kernel k, there exits a unique associated RKHS, and vice versa. When performing regret analysis under the frequentist view, it is natural to assume that the objective functions f belongs to the RKHS associated with the covariance kernel k specified in the GP model.

4.2 Regret Upper Bound of EIC

4.2.1 Regret Upper Bound

In order to derive the regret bound of EIC (Theorem 1), we make the following assumptions. These are standard assumptions commonly used in the analysis of BO algorithms (Srinivas et al., 2010; Chowdhury and Gopalan, 2017).

(A1) The objective function f belongs to the RKHS $\mathcal{H}_k$ associated with the positive semi-definite kernel function $k(\cdot,\cdot)$. Moreover, k is isotropic (that is $k(\mathbf{x},\mathbf{x}')$ depends only on $\|\mathbf{x}-\mathbf{x}'\|$ and decreases with $\|\mathbf{x}-\mathbf{x}'\|$) and $k(\mathbf{x},\mathbf{x}) \leq 1, \forall \mathbf{x} \in D$.

(A2) The RKHS norm of the objective function $\|f\|_{\mathcal{H}_k} := \sqrt{\langle f, f\rangle_{\mathcal{H}_k}}$ satisfies

$$\|f\|_{\mathcal{H}_k} \leq B \text{ for some } B > 0.$$

(A3) The noise sequence $\{\epsilon_n\}_{n\geq 1}$ is conditionally R-sub-Gaussian for a fixed constant $R > 0$:

$$\forall n \geq 0, \forall \lambda \in \mathbb{R}, E[e^{\lambda \epsilon_n} | \mathcal{F}_{n-1}] \leq \tfrac{\lambda^2 R^2}{2},$$

where $\mathcal{F}_{n-1}$ is the σ-algebra generated by random variables $\{\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}, \epsilon_1, \ldots, \epsilon_{n-1}, \mathbf{x}_n\}$.

Note that these assumptions do not directly reflect the design of our EIC algorithm: Although the objective function f is fixed, unknown, and a member of $\mathcal{H}_k$, and the noise random variables ϵ_n are conditionally R-sub-Gaussian martingale difference sequence, we still run EIC algorithm under the Gaussian process model as described in Section 2.1. In general, this represents a mis-specified prior and noise model, which is known as the agnostic setting in Srinivas et al. (2010).

Theorem 1 *Assume* (A1)–(A3). *Let* $0 < \delta < 1$. *With probability at least* $1-\delta$, *running EIC algorithm under a GP model with prior mean function* $\mu(\mathbf{x}) \equiv 0$, *and with parameters (12), (14), (17),* $\lambda^2 = 1 + 2/N$ *achieves*

$$\begin{aligned} R_N \;\leq\; & c^d B\sqrt{2N} + C'\sqrt{2}B + C'(\sqrt{2}B + 2\beta_N + \omega_N)(\log N + 1) \\ & + 2\lambda\big[\big(C' + \sqrt{2\log N}\big)\omega_N + (2C'+1)\beta_N\big]\sqrt{\gamma_N N}, \end{aligned}$$

where $\beta_N = B + R\sqrt{2(\gamma_N + 1 + \log(1/\delta))}$, $\omega_N = c_0\sqrt{\gamma_N + 1 + \log(1/\delta)}$ *and* C' *is a universal constant that depends only on* B, R *and* c_0. *Hence*

$$R_N = O(\sqrt{N}\gamma_N (\log N)^{1/2}). \tag{21}$$

Remark 1 *Under the noisy BO setting, Wang and De Freitas (2014) proved that the EI algorithm can achieve a regret upper bound of* $R_N = O\big(\sqrt{N}\gamma_N^{3/2}(\log N)\big)$. *Thereafter, Nguyen et al. (2017) derived an improved regret upper bound of EI. However, to the best of our knowledge, there is an issue in their proofs as highlighted in Tran-The et al. (2022). Compared with Wang and De Freitas (2014), the regret upper bound in Theorem 1 is much tighter, which improves upon traditional EI by a factor of* $\sqrt{\gamma_N \log N}$. *This significant improvement is achievable due to the evaluation cost feature in our EIC algorithm.*

High-probability regret upper bounds which hold for every finite budget N *have been established for a number of BO algorithms, such as GP-UCB and GP-TS. Under the frequentist setting, the tightest regret upper bound for vanilla GP-UCB and GP-TS are* $R_N = O\big(\sqrt{N}\gamma_N\big)$ *and* $R_N = O\big(\sqrt{N}\gamma_N(\log N)^{1/2}\big)$ *respectively as shown by Chowdhury and Gopalan (2017). To the best of our knowledge, our Theorem 1 is the first regret bound for EI-based algorithms which matches the regret bound of GP-UCB and GP-TS.*

Vakili et al. (2021) provided general upper bounds of the maximum information gain, based on the decay rate of the eigenvalues of the covariance kernel function. Using their results for the SE and Matérn kernels, we can derive the following corollary.

Corollary 1 *Let* $0 < \delta < 1$. *With probability at least* $1-\delta$, *the regret of EIC algorithm satisfies*

$$\begin{aligned} R_N &= O\Big(\sqrt{N}(\log N)^{d+\frac{3}{2}}\Big) \text{ for the SE kernel,} \\ R_N &= O\Big(N^{\frac{2\nu+3d}{4\nu+2d}}(\log N)^{\frac{6\nu+d}{4\nu+2d}}\Big) \text{ for the Matérn kernel.} \end{aligned}$$

Comparing with Scarlett et al. (2018), the regret upper bound of EIC for the SE covariance kernel is tight up to logarithmic factors.

4.2.2 Proof of Theorem 1

We preface the proof of Theorem 1 with Lemmas 1–7. These lemmas are proved in the appendix. The first lemma is borrowed from Chowdhury and Gopalan (2017), which provides a uniform confidence interval for the objective function based on maximum information gain.

Lemma 1 (Theorem 2 of Chowdhury and Gopalan (2017)) *Let* $\{\mathbf{x}_n\}_{n\geq 1}$ *be a discrete time stochastic process which is measurable with respect to the* σ*-algebra* $\{\mathcal{F}_{n-1}\}_{n\geq 1}$. *Let* $\{\epsilon_n\}_{n\geq 1}$ *be a real-valued stochastic process such that is (a)* $\mathcal{F}_n$*-measurable, and (b)* R*-sub-Gaussian conditionally on* $\mathcal{F}_{n-1}$ *for some* $R \geq 0$. *Let* f *be a member of the RKHS of real-valued functions with RKHS norm bounded by* B. *For any* $\delta \in (0,1)$, *define* $\beta_n = B + R\sqrt{2(\gamma_n + 1 + \log(1/\delta))}$. *Then, with probability at least* $1-\delta$, *the following holds:*

$$\text{uniformly over all } \mathbf{x} \in D \text{ and } n \geq 0, |f(\mathbf{x}) - \mu_n(\mathbf{x})| \leq \beta_n \sigma_n(\mathbf{x}), \tag{22}$$

where $\mu_n(\mathbf{x}), \sigma_n^2(\mathbf{x})$ *are posterior mean and variance kernel as defined in (7), with* $\lambda^2 = 1 + \frac{2}{N}$.

The second lemma plays a key role in the proof of Theorem 1. It states that if the acquisition function of point a $\mathbf{x}$ is larger than its evaluation cost, then the quantity $z_n(\mathbf{x}) = \frac{\mu_n(\mathbf{x})-\xi_n}{\sigma_n(\mathbf{x})}$ must be lower bounded. This explains why EIC can achieve a smaller cumulative regret than traditional EI. Lemma 2 is a crucial technical contribution of the paper. Moreover, the lower bound of $z_n(\mathbf{x})$ in (23) is asymptotically tight and hence cannot be further improved.

Lemma 2 *Let* $z_n(\mathbf{x}) := \frac{\mu_n(\mathbf{x})-\xi_n}{\sigma_n(\mathbf{x})}$. *For* $1 \leq n \leq N-1$,

$$\alpha_n^{EI}(\mathbf{x}) \geq L_n(\mathbf{x}) \text{ only if } z_n(\mathbf{x}) \geq -\omega_n\sqrt{2\log(N-n)}. \tag{23}$$

The third lemma provides upper and lower bounds on the EI acquisition function, which is based on Bull (2011) and Wang and De Freitas (2014).

Lemma 3 *Let* $I_n(\mathbf{x}) = \max\{0, f(\mathbf{x}) - \xi_n\}$. *Under the situation where event (22) occurs, the following holds uniformly for all* $\mathbf{x} \in D$ *and* $n \geq 0$*:*

$$\max\left\{I_n(\mathbf{x}) - \beta_n\sigma_n(\mathbf{x}), \tfrac{h(-\beta_n/\omega_n)}{h(\beta_n/\omega_n)}I_n(\mathbf{x})\right\} \leq \alpha_n^{EI}(\mathbf{x}) \leq I_n(\mathbf{x}) + (\beta_n+\omega_n)\sigma_n(\mathbf{x}).$$

It can be seen from Lemma 3 that the upper bound of acquisition function depends on $I_n(\mathbf{x})$ and $\sigma_n(\mathbf{x})$. In Theorem 1, we further bound these two quantities based on Lemma 4 and 5.

Lemma 4 *Let* $I_n(\mathbf{x}) = \max\{0, f(\mathbf{x}) - \xi_n\}$ *with incumbent* $\xi_n := \max_{1\leq i\leq n}\mu_n(\mathbf{x}_i)$. *Under the situation where event (22) occurs, the following holds:*

$$\sum_{n=1}^{N-1} I_n(\mathbf{x}_{n+1}) \leq \sqrt{2}B + \beta_N\sum_{n=0}^{N-1}\sigma_n(\mathbf{x}_{n+1}).$$

Lemma 5 *For any* $N \geq 1$, *the points* $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$ *selected by any algorithm satisfy*

$$\sum_{n=0}^{N-1}\sigma_n^2(\mathbf{x}_{n+1}) \leq 4\lambda^2\gamma_N.$$

Lemma 6 bounds the function value difference for any two points within the domain based on RKHS norm.

Lemma 6 *Assume* (A1) *and* (A2). $\forall \mathbf{x}, \mathbf{x}' \in D$, *we have* $f(\mathbf{x}) - f(\mathbf{x}') \leq \sqrt{2}B$.

The last lemma shows that the posterior variance kernel is monotonically decreasing in the number of iterations.

Lemma 7 $\forall \mathbf{x} \in D$ *and* $n \geq 1$, *we have*

$$\sigma_n^2(\mathbf{x}) = \sigma_{n-1}^2(\mathbf{x}) - \frac{k_{n-1}^2(\mathbf{x}, \mathbf{x}_n)}{\sigma_{n-1}^2(\mathbf{x}_n) + \lambda^2}, \tag{24}$$

where $k_n(\cdot,\cdot)$ *are posterior covariance kernel as defined in (6), and by convention,* $k_0(\cdot,\cdot) = k(\cdot,\cdot)$ *and* $\sigma_0^2(\mathbf{x}) = k_0(\mathbf{x},\mathbf{x})$. *Hence* $\sigma_n^2(\mathbf{x})$ *is monotonically decreasing in* n.

PROOF OF THEOREM 1. Let $z_n(\mathbf{x}) = \frac{\mu_n(\mathbf{x})-\xi_n}{\sigma_n(\mathbf{x})}$ and let $h(x) = x\Phi(x) + \phi(x)$. Express the acquisition function and the evaluation cost function as:

$$\alpha_n^{EI}(\mathbf{x}) = \omega_n\sigma_n(\mathbf{x})h\big(\tfrac{z_n(\mathbf{x})}{\omega_n}\big), \tag{25}$$

$$L_n(\mathbf{x}) = \omega_n\sigma_n(\mathbf{x})h\big(-\tfrac{z_n(\mathbf{x})}{\omega_n}\big)/(N-n). \tag{26}$$

Let $r_n = f(\mathbf{x}^*) - f(\mathbf{x}_{n+1})$, then the cumulative regret can be written as

$$R_N = \sum_{n=0}^{n_0-1} r_n + \sum_{n=n_0}^{N-1} r_n.$$

By (12) and Lemma 6,

$$\sum_{n=0}^{n_0-1} r_n \le \sqrt{2}Bn_0 = c^d B\sqrt{2N}. \tag{27}$$

For $n \ge n_0$, we break the regret at iteration n as

$$r_n = f(\mathbf{x}^*) - f(\mathbf{x}_{n+1}) = [f(\mathbf{x}^*) - \xi_n] + [\xi_n - f(\mathbf{x}_{n+1})] = [\mathrm{I}] + [\mathrm{II}].$$

Bounding term [I]. Since $\omega_n = c_0\sqrt{\gamma_n + 1 + \log(1/\delta)}$ and given that $\sqrt{\gamma_n + 1 + \log(1/\delta)} \ge 1$, we have $\beta_n/\omega_n \le (B + \sqrt{2}R)/c_0$. Hence by Lemma 3,

$$[\mathrm{I}] \le I_n(x^*) \le \tfrac{h(\beta_n/\omega_n)}{h(-\beta_n/\omega_n)}\alpha_n^{EI}(\mathbf{x}^*) \le C'\alpha_n^{EI}(\mathbf{x}^*) \text{ for some } C' > 0. \tag{28}$$

Here, C' is a universal constant that depends only on B, R and c_0. To bound $\alpha_n^{EI}(\mathbf{x}^*)$, we consider two cases:

- $\alpha_n^{EI}(\mathbf{x}^*) \ge L_n(\mathbf{x}^*)$ at iteration n. By Lemma 3,

  $$\alpha_n^{EI}(\mathbf{x}^*) \le \alpha_n^{EI}(\mathbf{x}_{n+1}) \le I_n(\mathbf{x}_{n+1}) + (\beta_n + \omega_n)\sigma_n(\mathbf{x}_{n+1}), \tag{29}$$

  where the first inequality follows because $\mathbf{x}_{n+1}$ has the largest acquisition function value.

- $\alpha_n^{EI}(\mathbf{x}^*) < L_n(\mathbf{x}^*)$ at iteration n. Since $\frac{dh}{dx} = \Phi(x) > 0$ and $\lim_{x\to-\infty} h(x) = 0$, we have $h(x) > 0$. By (25), (26) and $h(x) > 0$, this implies that

  $$h\big(\tfrac{z_n(\mathbf{x}^*)}{\omega_n}\big) < h\big(-\tfrac{z_n(\mathbf{x}^*)}{\omega_n}\big)/(N-n) \le h\big(-\tfrac{z_n(\mathbf{x}^*)}{\omega_n}\big). \tag{30}$$

  Since $h(z) = z + h(-z)$, we can deduce from (30) that $z_n(\mathbf{x}^*) \le 0$ when $\alpha_n^{EI}(\mathbf{x}^*) < L_n(\mathbf{x}^*)$. Let $\mathbf{x}_n^* = \arg\max_{1\le i\le n}\mu_n(\mathbf{x}_i)$. We have

  $$\begin{aligned} h\big(-\tfrac{z_n(\mathbf{x}^*)}{\omega_n}\big) &\le \frac{\xi_n - \mu_n(\mathbf{x}^*)}{\omega_n\sigma_n(\mathbf{x}^*)} + 1 \\ &\le \frac{f(\mathbf{x}_n^*) + \beta_n\sigma_n(\mathbf{x}_n^*) - f(\mathbf{x}^*) + \beta_n\sigma_n(\mathbf{x}^*)}{\omega_n\sigma_n(\mathbf{x}^*)} + 1 \\ &\le \frac{\sqrt{2}B + \beta_n\sigma_n(\mathbf{x}_n^*) + \beta_n\sigma_n(\mathbf{x}^*)}{\omega_n\sigma_n(\mathbf{x}^*)} + 1, \end{aligned} \tag{31}$$

where the first inequality follows from $h(z) \leq z+1$ when $z \geq 0$, the second inequality follows from Lemma 1 and the last inequality follows from Lemma 6. Substitute (31) back into (26) yields:

$$\alpha_n^{EI}(\mathbf{x}^*) < L_n(\mathbf{x}^*) \leq \frac{\sqrt{2}B + \beta_n\sigma_n(\mathbf{x}_n^*) + (\beta_n+\omega_n)\sigma_n(\mathbf{x}^*)}{N-n} \leq \frac{\sqrt{2}B + 2\beta_n + \omega_n}{N-n}, \tag{32}$$

where the last inequality holds because, by assumption (A1) and Lemma 7, it follows that $0 \leq \sigma_n^2(\mathbf{x}) \leq \sigma_0^2(\mathbf{x}) = k(\mathbf{x},\mathbf{x}) \leq 1$ for all $\mathbf{x} \in D$.

By combining (28), (29) and (32), we can deduce that

$$[\mathrm{I}] \leq C'[I_n(\mathbf{x}_{n+1}) + (\beta_n+\omega_n)\sigma_n(\mathbf{x}_{n+1}) + (\sqrt{2}B + 2\beta_n + \omega_n)/(N-n)]. \tag{33}$$

Bounding term [II]. Suppose there exist an $\mathbf{x} \in D$ at iteration n such that $\alpha_n^{EI}(\mathbf{x}) \geq L_n(\mathbf{x})$, then by Lemma 1 and Lemma 2,

$$\begin{aligned} [\mathrm{II}] &= \xi_n - f(\mathbf{x}_{n+1}) \\ &= \big(-\tfrac{\mu_n(\mathbf{x}_{n+1})-\xi_n}{\sigma_n(\mathbf{x}_{n+1})} + \tfrac{\mu_n(\mathbf{x}_{n+1})-f(\mathbf{x}_{n+1})}{\sigma_n(\mathbf{x}_{n+1})}\big)\sigma_n(\mathbf{x}_{n+1}) \\ &\leq \big[\omega_n\sqrt{2\log(N-n)} + \beta_n\big]\sigma_n(\mathbf{x}_{n+1}). \end{aligned} \tag{34}$$

On the other hand, suppose there exists no such point, then $\mathbf{x}_{n+1} = \arg\max_{1\leq i\leq n}\mu_n(\mathbf{x}_i)$. By Lemma 1, $\xi_n - f(\mathbf{x}_{n+1}) = \mu_n(\mathbf{x}_{n+1}) - f(\mathbf{x}_{n+1}) \leq \beta_n\sigma_n(\mathbf{x}_{n+1})$ and (34) still holds. Combining (33), (34) and Lemma 5:

$$\begin{aligned} \sum_{n=n_0}^{N-1} r_n &\leq \sum_{n=n_0}^{N-1} \Big\{C' I_n(\mathbf{x}_{n+1}) + C'(\sqrt{2}B + 2\beta_n + \omega_n)(N-n)^{-1} \\ &\qquad + \big[\big(C' + \sqrt{2\log(N-n)}\big)\omega_n + (C'+1)\beta_n\big]\sigma_n(\mathbf{x}_{n+1})\Big\} \\ &\leq C'\sum_{n=1}^{N-1} I_n(\mathbf{x}_{n+1}) + C'(\sqrt{2}B + 2\beta_N + \omega_N)\sum_{n=1}^{N-1}(N-n)^{-1} \\ &\qquad + \big[\big(C' + \sqrt{2\log N}\big)\omega_N + (C'+1)\beta_N\big]\sum_{n=0}^{N-1}\sigma_n(\mathbf{x}_{n+1}) \\ &\leq C'\sqrt{2}B + C'(\sqrt{2}B + 2\beta_N + \omega_N)(\log N + 1) \\ &\qquad + \big[\big(C' + \sqrt{2\log N}\big)\omega_N + (2C'+1)\beta_N\big]\sum_{n=0}^{N-1}\sigma_n(\mathbf{x}_{n+1}), \end{aligned} \tag{35}$$

where the last inequality follows from Lemma 4 and $\sum_{i=1}^{N}(N-n)^{-1} \leq \log N + 1$. By the Cauchy-Schwarz inequality and Lemma 5,

$$\sum_{n=0}^{N-1}\sigma_n(\mathbf{x}_{n+1}) \leq \sqrt{N\sum_{n=0}^{N-1}\sigma_n^2(\mathbf{x}_{n+1})} \leq 2\lambda\sqrt{N\gamma_N}. \tag{36}$$

Finally, combining (27), (35) and (36), we have

$$\begin{aligned} R_N \;\leq\; & c^d B\sqrt{2N} + C'\sqrt{2}B + C'(\sqrt{2}B + 2\beta_N + \omega_N)(\log N + 1) \\ & +2\lambda\big[\big(C' + \sqrt{2\log N}\big)\omega_N + (2C'+1)\beta_N\big]\sqrt{\gamma_N N}. \end{aligned} \tag{37}$$

Applying $\beta_N = B + R\sqrt{2(\gamma_N + 1 + \log(1/\delta))}$, $\omega_N = c_0\sqrt{\gamma_N + 1 + \log(1/\delta)}$ and $\lambda^2 = 1 + \frac{2}{N}$ to (37) gives us (21). Theorem 1 follows. □

Remark 2 *The main challenge in bounding the cumulative regret of EI lies in establishing an upper bound for term [II]. Our proposed EIC algorithm provides a significant advantage here, as it ensures that the inequality in (34) holds. This allows us to derive a tighter cumulative regret upper bound than what is currently achievable with EI.*

The EI is designed to minimize the simple regret. Related theoretical analysis can be found in Bull (2011), where it is referred to as the convergence rate. EI is inherently a myopic strategy, selecting points under the assumption that the experiment will conclude after the next observation. In contrast, EIC is designed to minimize cumulative regret, which is why we adjust it by incorporating the evaluation costs. Consequently, EI and EIC are suited for different scenarios: EI focuses on the final suggested value, while EIC emphasizes overall performance throughout the entire procedure. They evaluate performance using different metrics, targeting on simple regret and cumulative regret, respectively.

5. Experiments

In the experiments, we examine the numerical performance of our EIC algorithm using synthetic test functions (Section 5.1) as well as a real-world dataset (Section 5.2). For each test function/dataset, we generate $R = 100$ independent experiment trials, and the results are summarized in Figures 2 and 3. The solid line represents the average cumulative regret over these independent trials, with the shaded area showing the corresponding 95% confidence region. We use the GP prior with the mean function of $\mu \equiv 0$ and the SE covariance kernel in all experiments. The parameters are estimated using the maximum likelihood method by Santner et al. (2018). The total budget N for each test function/dataset is set to $N = 200 + n_0$, where the number of initial design points is set at $n_0 = 16, 36, 64$ for the $2, 4, 6$–dimensional objective functions respectively.

We compare our proposed EIC algorithm with the acquisition functions of the traditional EI, GP-UCB (Srinivas et al., 2010), and GP-TS (Chowdhury and Gopalan, 2017). We also compare with the algorithm proposed by Nguyen et al. (2017) which we refer to as *EI-Nyugen*. The EI-Nyugen algorithm requires a user-specified threshold parameter κ, and selects the point that maximizes the acquisition function $\alpha_n(\mathbf{x})$ if the largest acquisition function value is not smaller than κ. Otherwise, the point with the largest observed sample mean is selected. We follow their suggestion and choose $\kappa = 10^{-4}$ in our experiments.

5.1 Synthetic Experiments

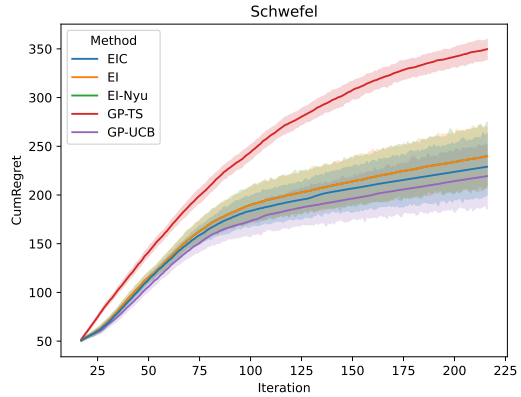
In the synthetic experiments, we consider six commonly used test functions: Schwefel-2, Eggholder-2, Ackley-2, Levy-4, Griewank-6, and Hartmann-6. Table 1 lists the mathematical expressions of these functions. Before implementing any BO algorithm, all test functions

are standardised so that the function values have mean zero and standard deviation one. The standard deviation of the homogenous noise σ is set as 0.1.

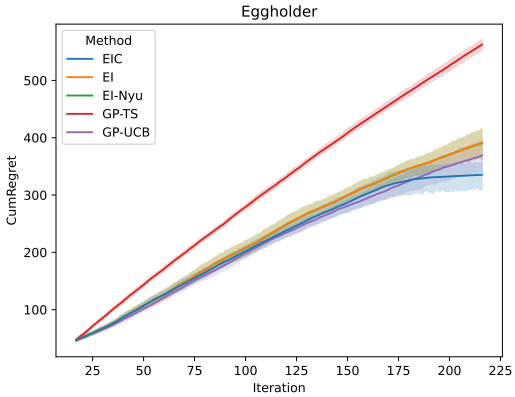
Figure 2 summarizes the results. We first observe that the traditional EI and EI-Nyugen have nearly identical performance for all test functions. This is probably due to the pre-specified threshold κ in EI-Nyugen being too small, which leads to the acquisition function seldom, if ever, falling below κ throughout the entire experiment. Therefore, EI-Nyugen behaves identically to the traditional EI. The GP-TS does not peform well across different functions probably due to its high randomness. Our proposed EIC algorithm has the smallest mean cumulative regret for the Eggholder-2, Griewank-6 and Harmann-6 test

Functions	d	Equation
Schwefel-2	2	$f(\mathbf{x}) = -\frac{1}{274.3}\left(418.9829 * 2 - \sum_{i=1}^{2} w_i \sin\left(\sqrt{\lvert w_i \rvert}\right) - 838.57\right)$ $w_i = 500x_i, i = 1, 2$ $x_i \in [-1, 1], i = 1, 2$ $\mathbf{x}* = (0.8419, 0.8419), f^* = 3.057$
Eggholder-2	2	$f(\mathbf{x}) = -\frac{1}{347.31}\left(-(w_2 + 47)\sin\left(\sqrt{\left\lvert w_2 + \frac{w_1}{2} + 47\right\rvert}\right) - w_1 \sin\left(\sqrt{\lvert w_1 - (w_2 + 47)\rvert}\right) - 1.96\right)$ $w_i = 512x_i, i = 1, 2$ $x_i \in [-1.17, 1.17], i = 1, 2$ $\mathbf{x}* = (1, 0.7895), f^* = 2.769$
Ackley-2	2	$f(\mathbf{x}) = 20\exp\left(-0.2\sqrt{\frac{1}{2}\sum_{i=1}^{2} x_i^2}\right) - \exp\left(\frac{1}{2}\sum_{i=1}^{2}\cos(2\pi x_i)\right) - 20 - \exp(1)$ $x_i \in [-32.768, 32.768], i = 1, 2$ $\mathbf{x}* = (0, 0), f^* = 0$
Levy-4	4	$f(\mathbf{x}) = -\frac{1}{27.9}\left(\sin^2(\pi w_1) + \sum_{i=1}^{3}(w_i - 1)^2\left[1 + 10\sin^2(\pi w_i + 1)\right] + (w_4 - 1)^2\left[1 + \sin^2(2\pi w_4)\right] - 42.55\right),$ $w_i = 1 + \frac{x_i - 1}{4}, \text{ for all } i = 1, \cdots, d$ $x_i \in [-10, 10], i = 1, \cdots, 4$ $\mathbf{x}* = (1, 1, 1, 1), f^* = 1.525$
Griewank-6	6	$f(\mathbf{x}) = -\frac{1}{0.47}\left(\sum_{i=1}^{6}\frac{x_i^2}{4000} - \prod_{i=1}^{6}\cos\left(\frac{x_i}{\sqrt{i}}\right) + 1 - 2.25\right)$ $x_i \in [-50, 50], i = 1, \cdots, 6$ $\mathbf{x}* = (0, 0, 0, 0, 0, 0), f^* = 4.787$
Hartmann-6	6	$f(\mathbf{x}) = -\frac{1}{0.38}\left(-\sum_{i=1}^{4}\alpha_i \exp\left(-\sum_{j=1}^{6} A_{ij}(x_j - P_{ij})^2\right) + 0.26\right)$ $\alpha = (1.0, 1.2, 3.0, 3.2)^T$ $\mathbf{A} = \begin{pmatrix} 10 & 3 & 17 & 3.50 & 1.7 & 8 \\ 0.05 & 10 & 17 & 0.1 & 8 & 14 \\ 3 & 3.5 & 1.7 & 10 & 17 & 8 \\ 17 & 8 & 0.05 & 10 & 0.1 & 14 \end{pmatrix}$ $\mathbf{P} = 10^{-4}\begin{pmatrix} 1312 & 1696 & 5569 & 124 & 8283 & 5886 \\ 2329 & 4135 & 8307 & 3736 & 1004 & 9991 \\ 2348 & 1451 & 3522 & 2883 & 3047 & 6650 \\ 4047 & 8828 & 8732 & 5743 & 1091 & 381 \end{pmatrix}$ $x_i \in [0, 1], i = 1, \cdots, 6$ $\mathbf{x}* = (0.20169, 0.150011, 0.476874, 0.275332, 0.311652, 0.6573), f^* = 8.059$

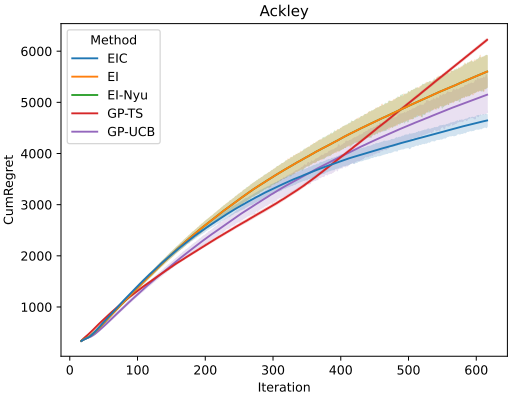
Table 1: List of test functions



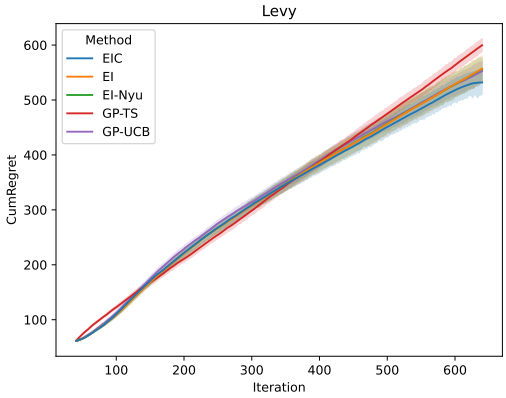
(a) Schwefel-2



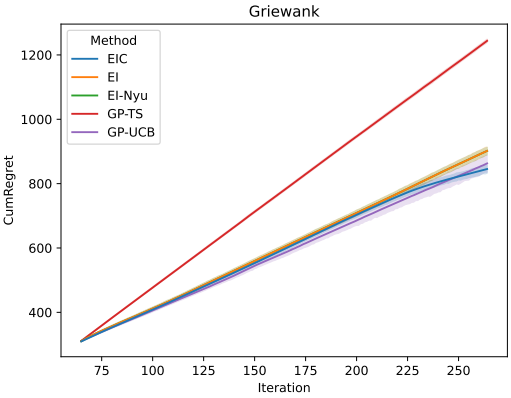
(b) Eggholder-2



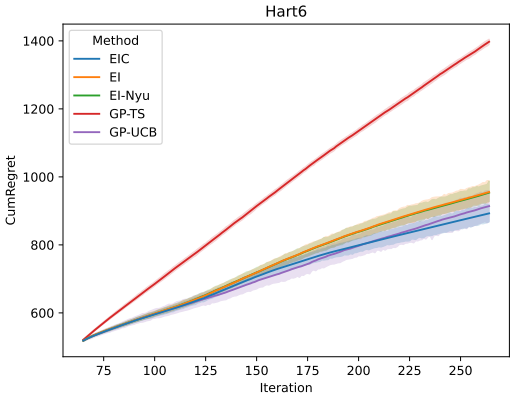
(c) Ackley-2



(d) Levy-4



(e) Griewank-6



(f) Hartmann-6

Figure 2: Cumulative regret of different BO algorithms on six test functions. EI and EI-Nyu lines overlap, causing the green line (EI-Nyu) to be obscured by the yellow line (EI).

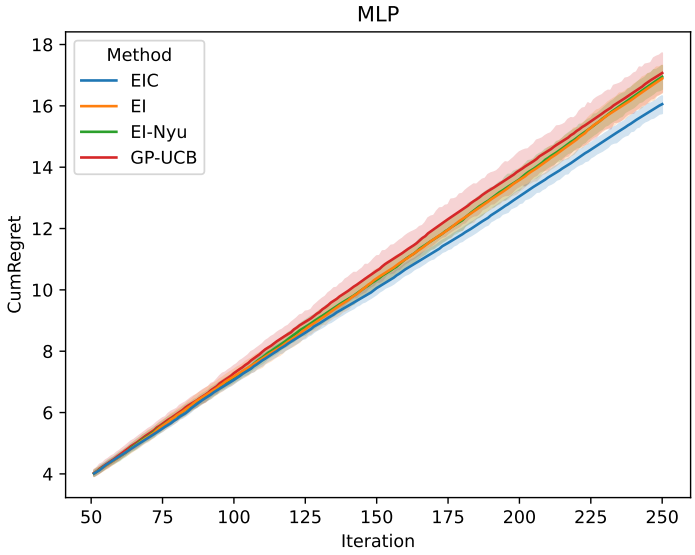


Figure 3: Cumulative regret of different BO algorithms in the neural network hyper-parameter tuning experiment.

functions after $200+n_0$ iterations. In addition, for these functions, the confidence regions of CumRegret for EIC overlap with those of UCB, showing EIC is competitive with UCB. The confidence regions of CumRegret for EIC does not overlapped with EI, EI-Nyugen and GP-TS, hence the reduction on cumulative regret is statistically significant. For the Schwefel-2, GP-UCB has the smallest cumulative regret. However the confidence regions of GP-UCB, EIC, EI-Nyugen and traditional EI overlap, indicating there is no statistical significance among these algorithms under 95% confidence level. For the Ackley-2 and the Levy-4 function, we use a bigger budget of $N = 600 + n_0$, since none of the compared algorithms is able to converge after $200 + n_0$ iterations. For the Ackley-2, GP-UCB has the smallest cumulative regret at the beginning, but it is outperformed by EIC after 350 iterations. For Levy-4 function, EI-Nyugen, EI, GP-UCB, and EIC have similar performance, but EIC shows a tendency to converge after 600 iterations while other algorithm's cumulative regret values continue to grow. These results show that our EIC algorithm performs competitively and consistently across a variety of test functions.

5.2 Real-world Experiment

In this section, we test the performance of various BO algorithms on a hyper-parameter tuning experiment for a neural network model, which is used in a healthcare application to classify whether a cancer is malignant or benign. We adopt the breast cancer Wisconsin dataset (https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+(diagnostic)) which contains 569 patients. Among them, 212 patients have malignant cancer and the other 357 have benign cancer. The dataset is further split into training and testing datasets with a ratio of 7:3. A one-hidden-layer multilayer perception (MLP) neural network model

is trained to classify whether the patients have malignant or benign cancer based on 30-dimensional covariates. We consider the tuning of four hyper-parameters: the number of hidden units, batch size, learning rate, and learning rate decay coefficients. To evaluate a point (i.e., a 4-dimensional vector specifying the values of the four hyper-parameters) selected by BO algorithms, we first train the MLP on the training dataset using the corresponding selected values of the hyper-parameters, and then evaluate the trained MLP on the testing dataset and report the classification accuracy as the corresponding observation.

In summary, this experiment involves a 4-dimensional optimization problem with the objective function f representing the classification accuracy. Since the true maximizer $\mathbf{x}^*$ and the maximum value $f(\mathbf{x}^*)$ are unknown, we set the maximum value $f(\mathbf{x}^*)$ to be 1 (i.e., an accuracy of 100%) in the calculation of the cumulative regret. Figure 3 summarizes the results for this experiment. We observe that for GP-UCB, EI-Nyugen and traditional EI, their confidence regions overlap and hence have very similar performance. The EIC achieves the smallest cumulative regret after 150 iterations and the difference is significant as the its confidence regions do not overlap with those of other algorithms.

6. Conclusion

In this article, we propose the EIC algorithm which aims to enhance the performance of traditional EI under the evaluation metric of cumulative regret. This is achieved by introducing an evaluation cost function that is compared against the acquisition function to balance the potential improvement to the cost of evaluation at each point. The algorithm is then designed to sample a point only if its acquisition function value exceeds its evaluation cost. We show that EIC can achieve a finite-time regret upper bound of $O\big(\sqrt{N}\gamma_N(\log N)^{1/2}\big)$ with high probability, and use both synthetic and real-world experiments to demonstrate that EIC indeed achieves smaller cumulative regret compared to traditional EI as well as other commonly used BO algorithms.

We suggest here two extensions of BO for future work. The first is to consider the optimization in more complex domains, such as graphs, discrete sequences and trees. The second is to handle the extra domain specific knowledge. For example in some nuclear safety applications we may know the optimum function value in advance, and how to incorporate such extra information still remains an open problem.

Acknowledgments

The authors would like to thank the Action Editor and the anonymous reviewers for their constructive suggestions, which greatly improved the quality of this manuscript. Shouri Hu's work is supported in part by the National Natural Science Foundation of China [Grant No.12401336], by Natural Science Foundation of Sichuan Province [Grant No.2023NSFSC1357], and the Fundamental Research Funds for the Central Universities [Grant No.Y030222059002059]. Szu Hui Ng's work is supported in part by the Singapore Ministry of Education (MOE) Academic Research Fund (AcRF) [R-266-149–114]. Bryan Kian Hsiang Low's research is supported by the National Research Foundation, Prime

Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme. The authors declare no competing interests.

Appendix A. Proof of Supporting Lemmas

Proof of Lemma 1. See Theorem 2 of Chowdhury and Gopalan (2017). □

Proof of Lemma 2. Let $h(z) := z\Phi(z) + \phi(z)$. We can write

$$\begin{aligned} \alpha_n^{EI}(\mathbf{x}) &= \omega_n \sigma_n(\mathbf{x}) h\big(\tfrac{z_n(\mathbf{x})}{\omega_n}\big) \text{ and} \\ L_n(\mathbf{x}) &= \omega_n \sigma_n(\mathbf{x}) h\big(-\tfrac{z_n(\mathbf{x})}{\omega_n}\big)/(N-n). \end{aligned}$$

Since $h(z) = z + h(-z)$, $\alpha_n^{EI}(\mathbf{x}) \geq L_n(\mathbf{x})$ is equivalent to

$$z_n(\mathbf{x}) \geq -(N-n-1)\omega_n h\big(\tfrac{z_n(\mathbf{x})}{\omega_n}\big). \tag{38}$$

It suffices to show that (38) does not holds for $z_n(\mathbf{x}) < -\omega_n\sqrt{2\log(N-n)}$, that is to show

$$\tfrac{z_n(\mathbf{x})}{\omega_n} < -(N-n-1)h\big(\tfrac{z_n(\mathbf{x})}{\omega_n}\big) \text{ when } z_n(\mathbf{x}) < -\omega_n\sqrt{2\log(N-n)}. \tag{39}$$

Check that (39) holds for $n = N-1$. Since $\Phi(x) > (-1/x + 1/x^3)\phi(x)$ for $x < 0$, we have

$$h\big(\tfrac{z_n(\mathbf{x})}{\omega_n}\big) = \tfrac{z_n(\mathbf{x})}{\omega_n}\Phi\big(\tfrac{z_n(\mathbf{x})}{\omega_n}\big) + \phi\big(\tfrac{z_n(\mathbf{x})}{\omega_n}\big) < \phi\big(\tfrac{z_n(\mathbf{x})}{\omega_n}\big)/\big(\tfrac{z_n(\mathbf{x})}{\omega_n}\big)^2.$$

Hence to show (39), it suffices to show that for $1 \leq n \leq N-2$,

$$(N-n-1)\phi\big(\tfrac{z_n(\mathbf{x})}{\omega_n}\big) < -\big(\tfrac{z_n(\mathbf{x})}{\omega_n}\big)^3 \text{ when } z_n(\mathbf{x}) < -\omega_n\sqrt{2\log(N-n)}. \tag{40}$$

Taking logarithm on both side of (40) give us

$$\log\big(\tfrac{N-n-1}{\sqrt{2\pi}}\big) - \tfrac{1}{2}\big(\tfrac{z_n(\mathbf{x})}{\omega_n}\big)^2 < 3\log\big(-\tfrac{z_n(\mathbf{x})}{\omega_n}\big).$$

Let $g(y) = \log\big(\tfrac{N-n-1}{\sqrt{2\pi}}\big) - \tfrac{1}{2}y^2 - 3\log(-y)$. Check that $g'(y) > 0$ for $y < 0$ and

$$g\big(-\sqrt{2\log(N-n)}\big) = \log\big(\tfrac{N-n-1}{N-n}\big) - \log(\sqrt{2\pi}) - \tfrac{3}{2}\log(2\log(N-n)) < 0.$$

Hence (40) holds for for $1 \leq n \leq N-2$. □

Proof of Lemma 3. Define $h(z) := z\Phi(z) + \phi(z)$. Let $z_n = \frac{\mu_n(\mathbf{x}) - \xi_n}{\sigma_n(\mathbf{x})}$ and $q_n = \frac{f(\mathbf{x}) - \xi_n}{\sigma_n(\mathbf{x})}$. Under the situation where event (22) occurs, we have $|z_n - q_n| = |\frac{\mu_n(\mathbf{x}) - f(\mathbf{x})}{\sigma_n(\mathbf{x})}| \leq \beta_n$.
To show the upper bound, we have

$$\begin{aligned} \alpha_n^{EI}(\mathbf{x}) &= \omega_n \sigma_n(\mathbf{x}) h(\tfrac{z_n}{\omega_n}) \\ &\leq \omega_n \sigma_n(\mathbf{x}) h(\tfrac{q_n + \beta_n}{\omega_n}) \\ &\leq \omega_n \sigma_n(\mathbf{x}) h(\tfrac{\max\{0, q_n\} + \beta_n}{\omega_n}) \\ &\leq \omega_n \sigma_n(\mathbf{x}) [\tfrac{\max\{0, q_n\}}{\omega_n} + \tfrac{\beta_n}{\omega_n} + 1] \\ &\leq I_n(\mathbf{x}) + (\beta_n + \omega_n)\sigma_n(\mathbf{x}), \end{aligned}$$

where the third inequality follows from $h(z) \leq z + 1$ for $z \geq 0$.
To show the lower bound, note that

$$\begin{aligned} \alpha_n^{EI}(\mathbf{x}) &= \omega_n \sigma_n(\mathbf{x}) h(\tfrac{z_n}{\omega_n}) \\ &\geq z_n \sigma_n(\mathbf{x}) \\ &\geq (q_n - \beta_n)\sigma_n(\mathbf{x}) \\ &= f(\mathbf{x}) - \xi_n - \beta_n \sigma_n(\mathbf{x}) \\ &\geq I_n(\mathbf{x}) - \beta_n \sigma_n(\mathbf{x}), \end{aligned} \tag{41}$$

where the first inequality follows from $h(z) \geq z$ for all z. Also, suppose that $f(\mathbf{x}) - \xi_n \geq 0$, we have

$$\alpha_n^{EI}(\mathbf{x}) \geq \omega_n \sigma_n(\mathbf{x}) h(\tfrac{q_n - \beta_n}{\omega_n}) \tag{42}$$

$$\geq \omega_n \sigma_n(\mathbf{x}) h(-\tfrac{\beta_n}{\omega_n}) \tag{43}$$

Combining (41) and (42) gives us

$$\begin{aligned} \alpha_n^{EI}(\mathbf{x}) &\geq \frac{\omega_n h(-\frac{\beta_n}{\omega_n})}{\omega_n h(-\frac{\beta_n}{\omega_n}) + \beta_n} I_n(\mathbf{x}) \\ &= \frac{h(-\frac{\beta_n}{\omega_n})}{h(\frac{\beta_n}{\omega_n})} I_n(\mathbf{x}), \end{aligned} \tag{44}$$

where the last line follows from the fact that $h(z) = z + h(-z)$. If $f(\mathbf{x}) - \xi_n < 0$, (44) still holds as $\alpha_n^{EI}(\mathbf{x}) > 0$. Hence Lemma 3 follows. □

Proof of Lemma 4. Suppose there exists some $1 \leq n_1 < n_2 < \cdots < n_K < N$ and $K \geq 1$ such that $f(\mathbf{x}_{n_k+1}) - \xi_{n_k} \geq 0$ for $k = 1, \ldots, K$. With a slight abuse of notation, let $n_0 = 0$ (rather than the total number of initial design points). Under the situation where event (22) occurs, we have

$$\begin{aligned} \sum_{n=1}^{N-1} I_n(\mathbf{x}_{n+1}) &= \sum_{n=1}^{N-1} \big(f(\mathbf{x}_{n+1}) - \xi_n\big) \mathbf{1}_{\{f(\mathbf{x}_{n+1}) \geq \xi_n\}} \\ &= \sum_{k=1}^{K} \big(f(\mathbf{x}_{n_k+1}) - \xi_{n_k}\big) \\ &\leq \sum_{k=1}^{K} \big(f(\mathbf{x}_{n_k+1}) - \mu_{n_k}(\mathbf{x}_{n_{k-1}+1})\big) \\ &\leq \sum_{k=1}^{K} \big(f(\mathbf{x}_{n_k+1}) - f(\mathbf{x}_{n_{k-1}+1}) + \beta_{n_k} \sigma_{n_k}(\mathbf{x}_{n_{k-1}+1})\big) \\ &= f(\mathbf{x}_{n_K+1}) - f(\mathbf{x}_{n_0+1}) + \sum_{k=1}^{K} \beta_{n_k} \sigma_{n_k}(\mathbf{x}_{n_{k-1}+1}) \end{aligned} \tag{45}$$

The first inequality of (45) follows from the fact that $\xi_{n_k} = \max_{1\le i\le n_k} \mu_{n_k}(\mathbf{x}_i) \ge \mu_{n_k}(\mathbf{x}_{n_{k-1}+1})$ and the second inequality follows from (22). By Lemma 7, the posterior variance is monotonically decreasing in n, and we have

$$\begin{aligned}\sum_{k=1}^{K} \beta_{n_k}\sigma_{n_k}(\mathbf{x}_{n_{k-1}+1}) &\le \beta_N \sum_{k=1}^{K} \sigma_{n_k}(\mathbf{x}_{n_{k-1}+1}) \\ &\le \beta_N \sum_{k=1}^{K} \sigma_{n_{k-1}}(\mathbf{x}_{n_{k-1}+1}) \\ &\le \beta_N \sum_{n=0}^{N-1} \sigma_n(\mathbf{x}_{n+1})\end{aligned} \tag{46}$$

Hence Lemma 4 follows from (45), (46) and Lemma 6. □

Proof of Lemma 5. Let $y_n = f(\mathbf{x}_n) + \epsilon_n$ for $1 \le n \le N$, with $\epsilon_1, \dots, \epsilon_N \overset{i.i.d}{\sim} N(0, \lambda^2\omega_N^2)$. Let $F_N = \big(f(\mathbf{x}_1), f(\mathbf{x}_2) \dots, f(\mathbf{x}_N)\big)^T$ and $Y_N = (y_1, y_2, \dots, y_N)^T$. The mutual information between Y_N and F_N can be written as

$$I(Y_N; F_N) = H(Y_N) - H(Y_N|F_N), \tag{47}$$

where $H(X)$ denote the differential entropy of X and $H(X|Y)$ denote the conditional (differential) entropy of X given Y. By the chain rule for entropy,

$$H(Y_N) = \sum_{n=1}^{N} H(y_n|y_{n-1}, \dots, y_1) = \frac{1}{2}\sum_{n=1}^{N} \log\big(2\pi e\omega_N^2(\sigma_{n-1}^2(\mathbf{x}_n) + \lambda^2)\big), \tag{48}$$

where the last equality follows since $y_n|y_{n-1}, \dots, y_1 \sim N\big(\mu_{n-1}(\mathbf{x}_n), \omega_N^2(\sigma_{n-1}^2(\mathbf{x}_n) + \lambda^2)\big)$. Moreover, conditioning on F_N, Y_N follows a multivariate normal distribution with covariance matrix $\Sigma = \lambda^2\omega_N^2\mathbf{I}_N$). Hence

$$H(Y_N|F_N) = \frac{1}{2}\log\big((2\pi e)^N \det(\Sigma)\big) = \frac{1}{2}\sum_{n=1}^{N} \log(2\pi e\lambda^2\omega_N^2). \tag{49}$$

Combining (48) and (49) gives us

$$I(Y_N; F_N) = \frac{1}{2}\sum_{n=1}^{N} \log\big(1 + \tfrac{\sigma_{n-1}^2(\mathbf{x}_n)}{\lambda^2}\big) = \frac{1}{2}\sum_{n=0}^{N-1} \log\big(1 + \tfrac{\sigma_n^2(\mathbf{x}_{n+1})}{\lambda^2}\big).$$

By the definition of maximum information gain, $\gamma_N \ge I(Y_N; F_N)$. Using assumption (A1) and Lemma 7, it follows that $0 \le \sigma_n^2(\mathbf{x}_{n+1}) \le \sigma_0^2(\mathbf{x}_{n+1}) = k(\mathbf{x}_{n+1}, \mathbf{x}_{n+1}) \le 1$. Given that $\lambda^2 = 1 + 2/N$, this implies $0 \le \frac{\sigma_n^2(\mathbf{x}_{n+1})}{\lambda^2} \le 1$ for all n. Since $\log(1+x) \ge x/2$ for $x \in [0, 1]$, we have

$$\gamma_N \ge \frac{1}{2}\sum_{n=0}^{N-1} \log\big(1 + \tfrac{\sigma_n^2(\mathbf{x}_{n+1})}{\lambda^2}\big) \ge \frac{1}{4}\sum_{n=0}^{N-1} \tfrac{\sigma_n^2(\mathbf{x}_{n+1})}{\lambda^2}.$$

Rearranging the terms gives us Lemma 5. □

Proof of Lemma 6. By (19), the Cauchy-Schwarz inequality and (4.2.1),

$$\begin{aligned} f(\mathbf{x}) - f(\mathbf{x}') &= \langle f, k(\cdot,\mathbf{x})\rangle_{\mathcal{H}_k} - \langle f, k(\cdot,\mathbf{x}_{n_0})\rangle_{\mathcal{H}_k} \\ &= \langle f, k(\cdot,\mathbf{x}) - k(\cdot,\mathbf{x}')\rangle_{\mathcal{H}_k} \\ &\leq \|f\|_{\mathcal{H}_k}\|k(\cdot,\mathbf{x}) - k(\cdot,\mathbf{x}')\|_{\mathcal{H}_k} \\ &\leq B\|k(\cdot,\mathbf{x}) - k(\cdot,\mathbf{x}')\|_{\mathcal{H}_k}. \end{aligned} \tag{50}$$

Since k is isotropic, $k(\mathbf{x},\mathbf{x}) = 1$ and $k(\mathbf{x},\mathbf{x}') > 0$, we have

$$\begin{aligned} & \|k(\cdot,\mathbf{x}) - k(\cdot,\mathbf{x}')\|_{\mathcal{H}_k} \\ = & \sqrt{\langle k(\cdot,\mathbf{x}) - k(\cdot,\mathbf{x}'), k(\cdot,\mathbf{x}) - k(\cdot,\mathbf{x}')\rangle_{\mathcal{H}_k}} \\ = & \sqrt{2 - k(\mathbf{x},\mathbf{x}')} \leq \sqrt{2}. \end{aligned} \tag{51}$$

Combining (50) and (51) give us the result. □

Proof of Lemma 7. Recall that $X = (\mathbf{x}_1,\ldots,\mathbf{x}_n)^T$, $k_{\mathbf{x}X} = k_{X\mathbf{x}}^T = \big(k(\mathbf{x}_1,\mathbf{x}),\ldots,k(\mathbf{x}_n,\mathbf{x})\big)$ and the (i,j) entry of K_{XX} is $k(\mathbf{x}_i,\mathbf{x}_j)$. Define the leave-one-out vector $X^{(-n)} = (\mathbf{x}_1,\ldots,\mathbf{x}_{n-1})^T$ and $k_{\mathbf{x}X^{(-n)}} = k_{X^{(-n)}\mathbf{x}}^T = \big(k(\mathbf{x}_1,\mathbf{x}),\ldots,k(\mathbf{x}_{n-1},\mathbf{x})\big)$. we can express

$$K_{XX} + \lambda^2 I_n = \begin{pmatrix} K_{X^{(-n)}X^{(-n)}} + \lambda^2 I_{n-1} & k_{X^{(-n)}\mathbf{x}_n} \\ k_{\mathbf{x}_n X^{(-n)}} & k(\mathbf{x}_n,\mathbf{x}_n) + \lambda^2 \end{pmatrix}.$$

Using the block matrix inversion formula, we have

$$(K_{XX} + \lambda^2 I_n)^{-1} = \begin{pmatrix} A^{-1} + \rho^{-1}A^{-1}BB^TA^{-1} & -\rho^{-1}A^{-1}B \\ -\rho^{-1}B^TA^{-1} & \rho^{-1} \end{pmatrix}, \tag{52}$$

where $A^{-1} = (K_{X^{(-n)}X^{(-n)}} + \lambda^2 I_{n-1})^{-1}$, $B = k_{X^{(-n)}\mathbf{x}_n}$ and $\rho = (k(\mathbf{x}_n,\mathbf{x}_n) + \lambda^2) - B^TA^{-1}B = \sigma_{n-1}^2(\mathbf{x}_n) + \lambda^2$ is the Schur complement. By (52) and the block matrix multiplication formula:

$$\begin{aligned} \sigma_n^2(\mathbf{x}) &= k(\mathbf{x},\mathbf{x}) - k_{\mathbf{x}X}(K_{XX} + \lambda^2 I_n)^{-1}k_{X\mathbf{x}} \\ &= k(\mathbf{x},\mathbf{x}) - \begin{pmatrix} k_{\mathbf{x}X^{(-n)}} & k(\mathbf{x},\mathbf{x}_n) \end{pmatrix} \begin{pmatrix} A^{-1} + \rho^{-1}A^{-1}BB^TA^{-1} & -\rho^{-1}A^{-1}B \\ -\rho^{-1}B^TA^{-1} & \rho^{-1} \end{pmatrix} \begin{pmatrix} k_{X^{(-n)}\mathbf{x}} \\ k(\mathbf{x}_n,\mathbf{x}) \end{pmatrix} \\ &= k(\mathbf{x},\mathbf{x}) - k_{\mathbf{x}X^{(-n)}}A^{-1}k_{X^{(-n)}\mathbf{x}} - \rho^{-1}k_{\mathbf{x}X^{(-n)}}A^{-1}BB^TA^{-1}k_{X^{(-n)}\mathbf{x}} \\ &\quad + \rho^{-1}k_{\mathbf{x}X^{(-n)}}A^{-1}Bk(\mathbf{x}_n,\mathbf{x}) + \rho^{-1}k(\mathbf{x},\mathbf{x}_n)B^TA^{-1}k_{X^{(-n)}\mathbf{x}} - \rho k(\mathbf{x},\mathbf{x}_n)k(\mathbf{x}_n,\mathbf{x}) \\ &= \sigma_{n-1}^2(\mathbf{x}) - \rho^{-1}\big(k^2(\mathbf{x},\mathbf{x}_n) - 2k(\mathbf{x},\mathbf{x}_n)B^TA^{-1}k_{X^{(-n)}\mathbf{x}} + (B^TA^{-1}k_{X^{(-n)}\mathbf{x}})^2\big) \\ &= \sigma_{n-1}^2(\mathbf{x}) - \rho^{-1}\big(k(\mathbf{x},\mathbf{x}_n) - B^TA^{-1}k_{X^{(-n)}\mathbf{x}}\big)^2 \\ &= \sigma_{n-1}^2(\mathbf{x}) - \frac{k_{n-1}^2(\mathbf{x},\mathbf{x}_n)}{\sigma_{n-1}^2(\mathbf{x}_n) + \lambda^2}. \ \square \end{aligned}$$

References

S. Agrawal. Recent advances in multiarmed bandits for sequential decision making. *INFORMS TutORials in Operations Research*, pages 167–188, 2019.

S. Agrawal and N. Goyal. Analysis of thompson sampling for the multi-armed bandit problem. In *The 25th Annual Conference on Learning Theory*, volume 23, pages 39.1–39.26, 2012.

N. Aronszajn. Theory of reproducing kernels. *Transactions of the American Mathematical Society*, 68:337–404, 1950.

P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time analysis of the multi-armed bandit problem. *Machine Learning*, 47:235–256, 2002.

T. Ban, M. Ohue, and Y. Akiyama. Efficient hyperparameter optimization by using bayesian optimization for drug-target interaction prediction. In *2017 IEEE 7th International Conference on Computational Advances in Bio and Medical Sciences (ICCABS)*, pages 1–6, 2017.

F. Berkenkamp, A. P. Schoellig, and A. Krause. No-regret bayesian optimization with unknown hyperparameters. *Journal of Machine Learning Research*, 20:1–24, 2019.

A. Berlinet and C. Thomas-Agnan. *Reproducing kernel Hilbert spaces in probability and statistics*. Springer, Boston, MA, 2011.

E. Brochu, V. M. Cora, and N. De Freitas. A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv preprint*, 2010.

A. D. Bull. Convergence rates of efficient global optimization algorithms. *Journal of Machine Learning Research*, 12:2879–2904, 2011.

O. Cappé, A. Garivier, O. A. Maillard, R. Munos, and G. Stoltz. Kullback-leibler upper confidence bounds for optimal sequential allocation. *Annals of Statistics*, 41:1516–1541, 2013.

S. R. Chowdhury and A. Gopalan. On kernelized multi-armed bandits. In *International Conference on Machine Learning*, pages 844–853, 2017.

P. Frazier, W. Powell, and S. Dayanik. The knowledge-gradient policy for correlated normal beliefs. *INFORMS journal on Computing*, 21:599–613, 2009.

P. I. Frazier and J. Wang. Bayesian optimization for materials design. In *Information science for materials discovery and design*, pages 45–75. Springer, Cham, 2016.

T. Fukazawa, Y. Harashima, Z. Hou, and T. Miyake. Bayesian optimization of chemical composition: A comprehensive framework and its application to RFe12-type magnet compounds. *Physical Review Materials*, 3:053807, 2019.

B. G. Galuzzi, I. Giordani, A. Candelieri, R. Perego, and F. Archetti. Hyperparameter optimization for recommender systems through bayesian optimization. *Computational Management Science*, 17:495–515, 2020.

S. Grünewälder, J. Y. Audibert, M. Opper, and J. Shawe–Taylor. Regret bounds for gaussian process bandit problems. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 273–280, 2010.

P. Hennig and C. J. Schuler. Entropy search for information-efficient global optimization. *Journal of Machine Learning Research*, 13:1809–1837, 2012.

J. M. Hernández-Lobato, M. W. Hoffman, and Z. Ghahramani. Predictive entropy search for efficient global optimization of black-box functions. In *Proceedings of the 27th International Conference on Neural Information Processing Systems*, volume 1, pages 918–926, 2014.

D. Huang, T. T. Allen, W. I. Notz, and N. Zeng. Global optimization of stochastic black-box systems via sequential kriging meta-models. *Journal of Global Optimization*, 34:441–466, 2006.

D. R. Jones, M. Schonlau, and W. J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13:455–492, 1998.

M. Kanagawa, P. Hennig, D. Sejdinovic, and B. K. Sriperumbudur. Gaussian processes and kernel methods: A review on connections and equivalences. *arXiv preprint*, 2018.

K. Kandasamy, A. Krishnamurthy, J. Schneider, and B. Póczos. Parallelised bayesian optimisation via thompson sampling. In *Proceedings of Machine Learning Research*, volume 84, pages 133–142, 2018.

J. Kawale, H. H. Bui, B. Kveton, L. Tran-Thanh, and S. Chawla. Efficient thompson sampling for online matrix-factorization recommendation. *Advances in neural information processing systems*, 28, 2015.

D. Kharkovskii, Z. Dai, and B. K. H. Low. Private outsourced bayesian optimization. In *International Conference on Machine Learning*, pages 5231–5242. PMLR, 2020.

J. Kirschner, M. Mutny, N. Hiller, R. Ischebeck, and A. Krause. Adaptive and safe bayesian optimization in high dimensions via one-dimensional subspaces. In *International Conference on Machine Learning*, pages 3429–3438. PMLR, 2019.

N. Korda, E. Kaufmann, and R. Munos. Thompson sampling for 1-dimensional exponential family bandits. *Advances in Neural Information Processing Systems*, 26:1448–1456, 2013.

Y. Koren, R. Bell, and C. Volinsky. Matrix factorization techniques for recommender systems. *Computer*, 42:30–37, 2009.

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

T. L. Lai and H. Robbins. Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics*, 6:4–22, 1985.

T. Lattimore. Refining the confidence level for optimistic bandit strategies. *The Journal of Machine Learning Research*, 19:765–796, 2018.

T. Lattimore and C. Szepesvári. *Bandit Algorithms*. Cambridge University Press, 2020.

B. Letham, B. Karrer, G. Ottoni, and E. Bakshy. Constrained bayesian optimization with noisy experiments. *Bayesian Analysis*, 14:495–519, 2019.

G. Malkomes, C. Schaff, and R. Garnett. Bayesian optimization for automated model selection. *Advances in Neural Information Processing Systems*, 29, 2016.

J. Mŏckus. On bayesian methods for seeking the extremum. In *Optimization techniques IFIP technical conference*, pages 400–404. Springer, 1975.

J. Mŏckus, V. Tiesis, and A. Žilinskas. The application of bayesian methods for seeking the extremum. *Towards Global Optimization*, 2:117–129, 1978.

S. H. Ng and J. Yin. Bayesian kriging analysis and design for stochastic simulations. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 22:1–26, 2012.

V. Nguyen, S. Gupta, S. Rana, C. Li, and S. Venkatesh. Regret for expected improvement over the best-observed value and stopping condition. In *Proceedings of Machine Learning Research*, volume 77, pages 279–294, 2017.

M. A. Osborne, R. Garnett, and S. J. Roberts. Gaussian processes for global optimization. In *3rd international conference on learning and intelligent optimization*, pages 1–15, 2009.

D. Packwood. *Bayesian Optimization for Materials Science*. Springer Singapore, 2017.

V. Picheny, D. Ginsbourger, Y. Richet, and G. Caplin. Quantile-based optimization of noisy computer experiments with tunable precision. *Technometrics*, 55:2–13, 2013.

I. O. Ryzhov. On the convergence rates of expected improvement methods. *Operations Research*, 64:1515–1528, 2016.

S. Sano, T. Kadowaki, K. Tsuda, and S. Kimura. Application of bayesian optimization for pharmaceutical product development. *Journal of Pharmaceutical Innovation*, 15:333–343, 2020.

T. J. Santner, B. J. Williams, and W. I. Notz. *The design and analysis of computer experiments (Second Edition)*. Springer Nature, 2018.

J. Scarlett, I. Bogunovic, and V. Cevher. Lower bounds on regret for noisy gaussian process bandit optimization. In *Conference on Learning Theory*, pages 1723–1742, 2017.

J. Scarlett, I. Bogunovic, and V. Cevher. Lower bounds on regret for noisy gaussian process bandit optimization. *arXiv preprint*, 2018.

B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. De Freitas. Taking the human out of the loop: A review of bayesian optimization. *Proceedings of the IEEE*, 104:148–175, 2016.

R. H. L. Sim, Y. Zhang, B. K. H. Low, and P. Jaillet. Collaborative bayesian optimization with fair regret. In *International Conference on Machine Learning*, pages 9691–9701, 2021.

J. Snoek, H. Larochelle, and R. P. Adams. Practical bayesian optimization of machine learning algorithms. *Advances in Neural Information Processing Systems*, 25, 2012.

N. Srinivas, A. Krause, S. M. Kakade, and M. Seeger. Gaussian process optimization in the bandit setting: no regret and experimental design. In *Proceedings of the 27th International Conference on International Conference on Machine Learning*, pages 1015–1022, 2010.

A. Takahashi and T. Suzuki. Bayesian optimization design for finding a maximum tolerated dose combination in phase i clinical trials. *The International Journal of Biostatistics*, 2021. doi: 10.1515/ijb-2020-0147.

A. L. Teckentrup. Convergence of gaussian process regression with estimated hyper-parameters and applications in bayesian inverse problems. *SIAM/ASA Journal on Uncertainty Quantification*, 8:1310–1337, 2020.

W. Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25:285–294, 1933.

H. M. Torun, M. Swaminathan, A. K. Davis, and M. L. F. Bellaredj. A global bayesian optimization algorithm and its application to integrated system design. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 26:792–802, 2018.

H. Tran-The, S. Gupta, S. Rana, and S. Venkatesh. Regret bounds for expected improvement algorithms in gaussian process bandit optimization. In *International Conference on Artificial Intelligence and Statistics*, pages 8715–8737, 2022.

S. Vakili, K. Khezeli, and V. Picheny. On information gain and regret bounds in gaussian process bandits. In *International Conference on Artificial Intelligence and Statistics*, pages 82–90, 2021.

E. Vazquez and J. Bect. Convergence properties of the expected improvement algorithm with fixed mean and covariance functions. *Journal of Statistical Planning and inference*, 140:3088–3095, 2010.

A. Žilinskas. Single-step bayesian search method for an extremum of functions of a single variable. *Cybernetics and Systems Analysis*, 11:160–166, 1975.

Z. Wang and N. De Freitas. Theoretical analysis of bayesian optimisation with unknown gaussian process hyper-parameters. *arXiv preprint*, 2014.

Z. Wang and S. Jegelka. Max-value entropy search for efficient bayesian optimization. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70, pages 3627–3635, 2017.

C. K. Williams and C. E. Rasmussen. *Gaussian Processes for Machine Learning*. MIT press, Cambridge, MA, 2006.

J. Wu and P. Frazier. The parallel knowledge gradient method for batch bayesian optimization. *Advances in Neural Information Processing Systems*, 29:3126–3134, 2016.

M. Wüthrich, B. Schölkopf, and A. Krause. Regret bounds for gaussian-process optimization in large domains. *Advances in Neural Information Processing Systems*, 34, 2021.