Randomized Gaussian Process Upper Confidence Bound
with Tighter Bayesian Regret Bounds

Shion Takeno [1 2] **Yu Inatsu** [1] **Masayuki Karasuyama** [1]

Abstract

Gaussian process upper confidence bound (GP-UCB) is a theoretically promising approach for black-box optimization; however, the confidence parameter β is considerably large in the theorem and chosen heuristically in practice. Then, randomized GP-UCB (RGP-UCB) uses a randomized confidence parameter, which follows the Gamma distribution, to mitigate the impact of manually specifying β. This study first generalizes the regret analysis of RGP-UCB to a wider class of distributions, including the Gamma distribution. Furthermore, we propose improved RGP-UCB (IRGP-UCB) based on a two-parameter exponential distribution, which achieves tighter Bayesian regret bounds. IRGP-UCB does not require an increase in the confidence parameter in terms of the number of iterations, which avoids over-exploration in the later iterations. Finally, we demonstrate the effectiveness of IRGP-UCB through extensive experiments.

1. Introduction

Bayesian optimization (BO) (Mockus et al., 1978) has become a widely-used framework for expensive black-box optimization problems. To reduce the number of function evaluations, BO sequentially observes the noisy function value using an acquisition function computed from a Bayesian model. BO has been applied to many different fields, including automatic machine learning (Snoek et al., 2012; Kandasamy et al., 2018b), materials informatics (Ueno et al., 2016), and drug design (Korovina et al., 2020; Griffiths & Hernández-Lobato, 2020).

Theoretical guarantees for BO have also been studied extensively (e.g., Srinivas et al., 2010; Russo & Van Roy, 2014).

The Gaussian process upper confidence bound (GP-UCB) (Srinivas et al., 2010), which achieves a strong theoretical guarantee for cumulative regret bounds, is a seminal work in this field. Based on this versatile framework for the analysis, while maintaining the theoretical guarantee, GP-UCB has been extended to various problems, which include multi-objective BO (Paria et al., 2020; Zuluaga et al., 2016), multi-fidelity BO (Kandasamy et al., 2016; 2017), parallel BO (Contal et al., 2013; Desautels et al., 2014), high-dimensional BO (Kandasamy et al., 2015), and cascade BO (Kusakawa et al., 2022).

However, the choice of a *confidence parameter* β_t, which controls the exploitation-exploration trade-off, is a practical challenge for GP-UCB-based methods, where t is the iteration of BO. Since the theoretical choice of β_t is considerably large and increases with t, particularly in the later iterations, GP-UCB with the theoretical β_t causes over-exploration, which reduces the optimization performance. Therefore, in practice, β_t is often specified manually using some heuristics, which strongly affects the performance of GP-UCB. Since this is a common problem with GP-UCB-based methods, mitigating the effect of the manual specification of β_t is an important task.

To resolve this issue, Berk et al. (2020) have proposed a *randomized* GP-UCB (RGP-UCB), which uses a randomized confidence parameter ζ_t, which follows the Gamma distribution. Their experimental results demonstrated that RGP-UCB performed better than the standard GP-UCB. Although Berk et al. (2020) provided the regret analysis, it appears to contain some technical issues, including an unknown convergence rate (See Appendix C for details). Therefore, the theoretical justification of RGP-UCB is still an open problem.

This study demonstrates a generalized regret bound of RGP-UCB, which holds for a wider class of distributions for ζ_t in the Bayesian setting, where an objective function f is assumed to be a sample path from the GP. However, this analysis requires an increase in ζ_t, as $\mathbb{E}[\zeta_t] \approx \beta_t$, which implies the over-exploration cannot be avoided. Furthermore, this generalization does not provide us the choice of the distribution of ζ_t.

[1] Department of Computer Science, Nagoya Institute of Technology, Aichi, Japan [2] RIKEN AIP, Tokyo, Japan. Correspondence to: Masayuki Karasuyama <karasuyama@nitech.ac.jp>.

	GP-UCB (Srinivas et al., 2010)	RGP-UCB	IRGP-UCB (proposed)
BCR (discrete)	$O(\sqrt{T\gamma_T \log T})$	$O(\sqrt{T\gamma_T \log T})$	$* \, O(\sqrt{T\gamma_T})$
BCR (continuous)	$O(\sqrt{T\gamma_T \log T})$	$O(\sqrt{T\gamma_T \log T})$	$O(\sqrt{T\gamma_T \log T})$
Sufficient condition of BSR $< \eta$	$\sqrt{\gamma_T \log T / T} \lesssim \eta$	$\sqrt{\gamma_T \log T / T} \lesssim \eta$	$* \, \sqrt{\gamma_T / T} \lesssim \eta/\sqrt{-\log \eta}$

Table 1. Summary of Bayesian regret of GP-UCB-based algorithms. The first and second rows show the BCR bounds for discrete and continuous input domains, respectively, where γ_T is the maximum information gain defined in Section 2. The third row shows the sufficient conditions to achieve that BSR is lower than the predefined accuracy $\eta \in (0, 1)$ for both continuous and discrete input domains, where $\lesssim$ represents an inequality ignoring factors except for T and η. Star means better bounds than known results with respect to T. Note that the results in the right two columns are derived in this paper.

The aforementioned analyses motivated us to propose an improved RGP-UCB (IRGP-UCB), in which ζ_t follows a *two-parameter exponential distribution*. First, we show the sub-linear Bayesian cumulative regret (BCR) bounds of IRGP-UCB in the Bayesian setting for both finite and infinite input domains. In particular, when the input domain is finite, the BCR for IRGP-UCB achieves the better convergence rate, which shows a $O(\sqrt{\log T})$ multiplicative factor improvement from the known bounds, where T is the entire time horizon. Furthermore, IRGP-UCB achieves a tighter convergence rate with respect to T for the Bayesian simple regret (BSR) than existing analyses for both finite and infinite domains. More importantly, these analyses also reveal that the increase in ζ_t for IRGP-UCB is unnecessary. Therefore, the over-exploration of the original GP-UCB is theoretically alleviated via randomization using the two-parameter exponential distribution. A key finding in our proof is a direct upper bound of $\mathbb{E}[\max f(\boldsymbol{x})]$. We show that $\mathbb{E}[\max f(\boldsymbol{x})]$ can be bounded by the expectation of the UCB with our random non-increasing confidence parameter, which enables us to derive tighter Bayesian regret bounds than known results.

Our main contributions are summarized as follows:

- We provide the Bayesian regret bounds for RGP-UCB, which holds for a wider class of distributions for ζ_t and achieves the same convergence rate as GP-UCB.

- We propose yet another randomized variant of GP-UCB called IRGP-UCB, which sets the confidence parameter ζ_t using the two-parameter exponential distribution. A notable advantage of IRGP-UCB is that increases in the confidence parameter in proportion to t are unnecessary.

- We show Bayesian regret bounds for IRGP-UCB, which achieves the better convergence rate, in which $O(\sqrt{\log T})$ multiplicative factor is improved from known results. This result suggests that the over-exploration of GP-UCB is theoretically alleviated.

- We provide the upper bound for $\mathbb{E}[\max f(\boldsymbol{x})]$ along the way of the proof, which enables us to improve the Bayesian regret bounds.

The theoretical results are summarized in Table 1. Finally, we demonstrate the effectiveness of IRGP-UCB through a wide range of experiments.

1.1. Related Work

This study considers BO with the Bayesian setting, where the objective function f is assumed to be a sample path from the GP. Various BO methods have been developed in the literature, for example, expected improvement (EI) (Mockus et al., 1978), entropy search (ES) (Hennig & Schuler, 2012), and predictive entropy search (PES) (Hernández-Lobato et al., 2014). Although the regret analysis of EI for the noiseless and frequentist setting, in which f is an element of reproducing kernel Hilbert space, is provided in (Bull, 2011), the Bayesian setting has not been considered. Further, although the practical performance of ES and PES has been shown repeatedly, their regret analysis is an open problem.

GP-UCB is one of the prominent studies for theoretically guaranteed BO methods. Srinivas et al. (2010) showed the high probability bound for cumulative regret. Although Bayesian regret analysis for GP-UCB has not been shown explicitly, Paria et al. (2020) have shown the BCR bound for a multi-objective extension of GP-UCB, which contains that of the original single-objective GP-UCB as the special case. For completeness, we show the BCR bound of GP-UCB in Section 2.4. Although many studies (e.g., Srinivas et al., 2010; Chowdhury & Gopalan, 2017; Janz et al., 2020) considered the frequentist setting, this study concentrates on the Bayesian setting.

Berk et al. (2020) attempted to alleviate the GP-UCB hyperparameter tuning issue through randomization using the Gamma distribution. Note that our result shown in Theorem 4.1 differs from (Berk et al., 2020). We believe that these differences come from several technical issues in the proof in (Berk et al., 2020) (See Appendix C for details). In addition, the final convergence rate in (Berk et al., 2020) includes the variables whose convergence rate is not proved. In contrast, our Theorem 4.1 fully clarifies the convergence rate without those unproved variables. Moreover, Theorem 4.1 is generalized to a wider class of distributions for confidence parameters that contain the Gamma distribution.

In addition, the two-parameter exponential distribution used in IRGP-UCB is not considered in (Berk et al., 2020).

Another standard BO method with a theoretical guarantee is Thompson sampling (TS) (Russo & Van Roy, 2014; Kandasamy et al., 2018a). TS achieves the sub-linear BCR bounds by sequentially observing the optimal point of the GP's posterior sample path. Although TS does not require any hyperparameter, TS is often deteriorated by over-exploration, as discussed in (Shahriari et al., 2016).

Wang et al. (2016) have shown the regret analysis of the GP estimation (GP-EST) algorithm, which can be interpreted as GP-UCB with the confidence parameter defined using $\hat{m}$, an estimator of $\mathbb{E}[\max f(\boldsymbol{x})]$. Their analysis requires an assumption $\hat{m} \geq \mathbb{E}[\max f(\boldsymbol{x})]$, whose sufficient condition is provided in (Wang et al., 2016, Corollary 3.5). However, this sufficient condition does not typically hold, as discussed immediately after the corollary in (Wang et al., 2016). Furthermore, the final convergence rate contains $\hat{m}$ itself, whose convergence rate is not clarified. In contrast, our Lemma 4.2 shows the bound for $\mathbb{E}[\max f(\boldsymbol{x})]$ under common regularity conditions. Wang & Jegelka (2017) have shown the regret analysis of max-value entropy search (MES). However, it is pointed out that their proof contains several technical issues (Takeno et al., 2022b).

2. Background

2.1. Bayesian Optimization

We consider an optimization problem $\boldsymbol{x}^* = \arg\max_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x})$, where f is an unknown expensive-to-evaluate objective function, $\mathcal{X} \subset \mathbb{R}^d$ is an input domain, and d is an input dimension. BO sequentially observes the noisy function value aiming to minimize the number of function evaluations. Thus, at each iteration t, we can query $\boldsymbol{x}_t$ and obtain $y_t = f(\boldsymbol{x}_t) + \epsilon_t$, where $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$ is i.i.d. Gaussian noise with a positive variance $\sigma^2 > 0$.

We assume that f is a sample path from a GP (Rasmussen & Williams, 2005) with a zero mean and a stationary kernel function $k : \mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}$ denoted as $f \sim \mathcal{GP}(0, k)$. At each iteration t, a dataset $\mathcal{D}_{t-1} := \{(\boldsymbol{x}_i, y_i)\}_{i=1}^{t-1}$ is already obtained from the nature of BO. Then, the posterior distribution $p(f \mid \mathcal{D}_{t-1})$ is a GP again. The posterior mean and variance of $f(\boldsymbol{x})$ are derived as follows:

$$\mu_{t-1}(\boldsymbol{x}) = \boldsymbol{k}_{t-1}(\boldsymbol{x})^\top (\boldsymbol{K} + \sigma^2 \boldsymbol{I}_{t-1})^{-1} \boldsymbol{y}_{t-1},$$
$$\sigma_{t-1}^2(\boldsymbol{x}) = k(\boldsymbol{x}, \boldsymbol{x}) - \boldsymbol{k}_{t-1}(\boldsymbol{x})^\top (\boldsymbol{K} + \sigma^2 \boldsymbol{I}_{t-1})^{-1} \boldsymbol{k}_{t-1}(\boldsymbol{x}),$$

where $\boldsymbol{k}_{t-1}(\boldsymbol{x}) := (k(\boldsymbol{x}, \boldsymbol{x}_1), \dots, k(\boldsymbol{x}, \boldsymbol{x}_{t-1}))^\top \in \mathbb{R}^{t-1}$, $\boldsymbol{K} \in \mathbb{R}^{(t-1) \times (t-1)}$ is the kernel matrix whose (i, j)-element is $k(\boldsymbol{x}_i, \boldsymbol{x}_j)$, $\boldsymbol{I}_{t-1} \in \mathbb{R}^{(t-1) \times (t-1)}$ is the identity matrix, and $\boldsymbol{y}_{t-1} := (y_1, \dots, y_{t-1})^\top \in \mathbb{R}^{t-1}$. Hereafter, we denote that the probability density function (PDF)

$p(\cdot \mid \mathcal{D}_{t-1}) = p_t(\cdot)$, the probability $\Pr(\cdot \mid \mathcal{D}_{t-1}) = \Pr_t(\cdot)$, and the expectation $\mathbb{E}[\cdot \mid \mathcal{D}_{t-1}] = \mathbb{E}_t[\cdot]$ for brevity.

2.2. Preliminaries for Regret Analysis

When $\mathcal{X}$ is infinite (continuous), the following regularity assumption is used in most analyses (e.g., Srinivas et al., 2010; Kandasamy et al., 2018a; Paria et al., 2020):

Assumption 2.1. *Let $\mathcal{X} \subset [0, r]^d$ be a compact and convex set, where $r > 0$. Assume that the kernel k satisfies the following condition on the derivatives of a sample path f. There exists the constants $a, b > 0$ such that,*

$$\Pr\left(\sup_{\boldsymbol{x} \in \mathcal{X}} \left| \frac{\partial f}{\partial \boldsymbol{x}_j} \right| > L\right) \leq a \exp\left(-\left(\frac{L}{b}\right)^2\right), \text{ for } j \in [d],$$

where $[d] = \{1, \dots, d\}$.

Our analysis also requires this assumption.

The convergence rates of regret bounds are characterized by *maximum information gain* (MIG) (Srinivas et al., 2010; Vakili et al., 2021). MIG γ_T is defined as follows:

Definition 2.1 (Maximum information gain). *Let $f \sim \mathcal{GP}(0, k)$. Let $A = \{\boldsymbol{a}_i\}_{i=1}^T \subset \mathcal{X}$. Let $\boldsymbol{f}_A = (f(\boldsymbol{a}_i))_{i=1}^T$, $\boldsymbol{\epsilon}_A = (\epsilon_i)_{i=1}^T$, where $\forall i, \epsilon_i \sim \mathcal{N}(0, \sigma^2)$, and $\boldsymbol{y}_A = \boldsymbol{f}_A + \boldsymbol{\epsilon}_A \in \mathbb{R}^T$. Then, MIG γ_T is defined as follows:*

$$\gamma_T := \arg\max_{A \subset \mathcal{X}; |A|=T} I(\boldsymbol{y}_A; \boldsymbol{f}_A),$$

where I is the Shanon mutual information.

MIG is known to be sub-linear for commonly used kernel functions, e.g., $\gamma_T = O((\log T)^{d+1})$ for RBF kernels and $\gamma_T = O(T^{\frac{d}{2\nu+d}} (\log T)^{\frac{2\nu}{2\nu+d}})$ for Matèrn-ν kernels (Srinivas et al., 2010; Vakili et al., 2021).

2.3. Bayesian Regret Analysis

In this paper, we evaluate the performance of BO methods by *Bayesian regret* (Russo & Van Roy, 2014; Kandasamy et al., 2018a; Paria et al., 2020). The BSR[1] and BCR are defined as follows:

$$\text{BSR}_T := \mathbb{E}\left[f(\boldsymbol{x}^*) - \max_{t \leq T} f(\boldsymbol{x}_t)\right], \quad (1)$$

$$\text{BCR}_T := \mathbb{E}\left[\sum_{t=1}^T f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t)\right],$$

where the expectation is taken with all randomness, including f, ϵ_t, and the randomness of BO algorithms. We

---

[1]This definition of BSR only requires that some observed input, which may be unknown due to the noise, achieves low regret. We define BSR using the input recommended by the algorithm and show its convergence in Appendix A.

discuss the convergence of BSR by analyzing the BCR. That is, if BCR is sub-linear, BSR converges to zero since $\mathrm{BSR}_T \leq \frac{1}{T}\mathrm{BCR}_T$. Furthermore, we evaluate a required number of function evaluations to achieve $\mathrm{BSR}_T \leq \eta$, where $\eta > 0$ represents the desired accuracy of a solution.

According to (Russo & Van Roy, 2014), Bayesian regret bounds directly provide high probability bounds. Assume $\mathrm{BCR}_T = O(h(T))$ with some non-negative function h. Then, the direct consequence of Markov's inequality implies that $\sum_{t=1}^{T} f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t) = O(h(T)/\delta)$ holds with probability $1 - \delta$, where $\delta \in (0,1)$. Thus, the improvement of the convergence rate of BCR for T shows that of high-probability bound although the rate for δ is worse compared to $\log(1/\delta)$ (Srinivas et al., 2010).

The existing Bayesian regret analyses typically use the technique called *regret decomposition* (Russo & Van Roy, 2014). This approach decomposes BCR as follows:

$$\mathrm{BCR}_T = \sum_{t=1}^{T} \mathbb{E}\left[f(\boldsymbol{x}^*) - u_t(\boldsymbol{x}_t) \right] + \mathbb{E}\left[u_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t) \right],$$

(2)

where $u_t(\boldsymbol{x}) := \mu_{t-1}(\boldsymbol{x}) + \beta_t^{1/2}\sigma_{t-1}(\boldsymbol{x})$. If we use fixed constant for $\{\beta_t\}_{t\geq 1}$ (e.g., $\forall t \geq 1, \beta_t = 2$), bounding the first term $\sum_{t=1}^{T} \mathbb{E}\left[f(\boldsymbol{x}_*) - u_t(\boldsymbol{x}_t) \right]$ is difficult. Thus, in existing analyses, $\{\beta_t\}_{t\geq 1}$ is scheduled so that $\beta_t \propto \log t$. Since β_t remains in the upper bound, the regret decomposition using $u_t(\boldsymbol{x})$ deteriorates the final convergence rate, as we will show an example of GP-UCB in Theorem 2.1. IRGP-UCB avoids this problem by randomizing the confidence parameters, as shown in Section 4.

2.4. GP-UCB

The GP-UCB (Srinivas et al., 2010) selects the next evaluation by maximizing UCB as follows:

$$\boldsymbol{x}_t = \arg\max_{\boldsymbol{x}\in\mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \beta_t^{1/2}\sigma_{t-1}(\boldsymbol{x}).$$

Although BCR bound for the standard GP-UCB has not been shown explicitly, sub-linear regret bounds of GP-UCB can be shown in nearly the same way as (Paria et al., 2020):

Theorem 2.1 (Informal: BCR of GP-UCB). *Suppose that $f \sim \mathcal{GP}(0,k)$, where $k(\boldsymbol{x},\boldsymbol{x}) = 1$. Assume that $\mathcal{X}$ is finite or Assumption 2.1 holds. Then, by running GP-UCB with $\beta_t \propto \log(t)$, BCR can be bounded as follows:*

$$\mathrm{BCR}_T = O(\sqrt{T\beta_T\gamma_T}),$$

which implies sub-linearity when $\gamma_T = o(T/\log T)$.

See Appendix B for the proof[2]. In the theorem, the confidence parameter β_t must be scheduled as $\beta_t \propto \log(t)$.

---

[2]It is worth noting that the convergence rate for a in Assump-

---

Algorithm 1 IRGP-UCB

Require: Input space $\mathcal{X}$, Parameters s and λ for ζ_t, GP prior $\mu = 0$ and k
1: $\mathcal{D}_0 \leftarrow \emptyset$
2: **for** $t = 1, \ldots$ **do**
3:     Fit GP to $\mathcal{D}_{t-1}$
4:     Generate ζ_t by Eq. (3)
5:     $\boldsymbol{x}_t \leftarrow \arg\max_{\boldsymbol{x}\in\mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \zeta_t^{1/2}\sigma_{t-1}(\boldsymbol{x})$
6:     Observe $y_t = f(\boldsymbol{x}_t) + \epsilon_t$ and $\mathcal{D}_t \leftarrow \mathcal{D}_{t-1}\cup(\boldsymbol{x}_t,y_t)$
7: **end for**

---

3. Algorithm

Algorithm 1 shows a simple procedure of IRGP-UCB. The difference from the original GP-UCB is that ζ_t is a random variable, not a constant. Using generated ζ_t, the next evaluation is chosen as follows:

$$\boldsymbol{x}_t = \arg\max_{\boldsymbol{x}\in\mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \zeta_t^{1/2}\sigma_{t-1}(\boldsymbol{x}).$$

Our choice of the distribution of ζ_t is a *two-parameter exponential distribution* (e.g., Beg, 1980; Lam et al., 1994), whose PDF is written as follows:

$$p(\zeta; s, \lambda) = \begin{cases} \lambda \exp\bigl(\lambda(\zeta - s)\bigr) & \text{if } \zeta \geq s, \\ 0 & \text{otherwise,} \end{cases}$$

where λ is a rate parameter. This can be seen as a distribution of the sum of s and $Z \sim \mathrm{Exp}(\lambda)$. Thus, the sampling from the two-parameter exponential distribution can be easily performed using an exponential distribution as follows:

$$\zeta_t \leftarrow s + Z, \text{ where } Z \sim \mathrm{Exp}(\lambda). \qquad (3)$$

The theoretical choice of s and λ will be shown in Section 4.

4. Regret Analysis

First, we show the Bayesian regret analysis for a general RGP-UCB, which is not restricted to the two-parameter exponential distribution. Next, we show the tighter Bayesian regret bounds for IRGP-UCB.

4.1. Regret Bound for General RGP-UCB

Here, we provide a generalized theoretical analysis for RGP-UCB, which holds for a wider class of distributions.

Theorem 4.1 (Informal: BCR of RGP-UCB). *Suppose that $f \sim \mathcal{GP}(0,k)$, where $k(\boldsymbol{x},\boldsymbol{x}) = 1$. Assume that $\mathcal{X}$ is finite or Assumption 2.1 holds. Let $\{\zeta_t\}_{t\geq 1}$ be a sequence of non-negative random variables, which satisfies $\mathbb{E}[\zeta_t] = O(\log t)$*

---

tion 2.1 is tightened in our proof compared with the proofs in (Paria et al., 2020; Kandasamy et al., 2018a). This is applied to all subsequent regret analyses for a continuous domain. See details for Appendixes B and H.

and $\mathbb{E}[\exp(-\zeta_t/2)] = O(1/t^2)$. *By running RGP-UCB with* ζ_t, *BCR can be bounded as follows:*

$$\mathrm{BCR}_T = O(\sqrt{T\gamma_T \mathbb{E}[\zeta_T]}),$$

which implies sub-linearity when $\gamma_T = o(T/\log T)$.

See Appendix D for the proof, in which we also rectify the technical issues in (Berk et al., 2020) listed in Appendix C. A notable difference between Theorem 4.1 and (Berk et al., 2020, Theorem 3) is that Theorem 4.1 is applicable to a wider class of distributions of ζ_t, not only Gamma distribution proposed in (Berk et al., 2020). Roughly speaking, Theorem 4.1 only requires that the distribution has the parameters that can control the distribution so that $\mathbb{E}[\zeta_t] = O(\log t)$ and $\mathbb{E}[\exp(-\zeta_t/2)] = O(1/t^2)$. This condition can be satisfied by many distributions, such as Gamma, two-parameter exponential, and truncated normal distributions, as shown in Appendix D.2. See Appendix D.2 for a detailed discussion.

Although Theorem 4.1 achieves sub-linear BCR bounds, the implication of Theorem 4.1 is restricted. In the proof, ζ_t must behave similar to β_t, that is, $\mathbb{E}[\zeta_t] \propto \beta_t$ and $\mathbb{E}[\exp(-\zeta_t/2)] \propto \exp(-\beta_t/2)$. Therefore, based on this theorem, over-exploration of the GP-UCB cannot be avoided. Furthermore, this theorem does not specify a distribution, that is suitable for the randomization of GP-UCB.

4.2. Regret Bounds for IRGP-UCB

First, for completeness, we provide a modified version of the basic lemma derived in (Srinivas et al., 2010):

Lemma 4.1. *Suppose that* f *is a sample path from a GP with zero mean and a stationary kernel* k *and* $\mathcal{X}$ *is finite. Pick* $\delta \in (0,1)$ *and* $t \geq 1$. *Then, for any given* $\mathcal{D}_{t-1}$,

$$\mathrm{Pr}_t\left(f(\boldsymbol{x}) \leq \mu_{t-1}(\boldsymbol{x}) + \beta_\delta^{1/2}\sigma_{t-1}(\boldsymbol{x}), \forall \boldsymbol{x} \in \mathcal{X}\right) \geq 1 - \delta,$$

where $\beta_\delta = 2\log(|\mathcal{X}|/(2\delta))$.

See Appendix E.1 for the proof. This lemma differs slightly from (Srinivas et al., 2010, Lemma 5.1), since, in Lemma 4.1, the iteration t is fixed, and β_δ does not depend on t.

Based on Lemma 4.1, we show the following key lemma to show the improved convergence rate:

Lemma 4.2. *Let* $f \sim \mathcal{GP}(0,k)$, *where* k *is a stationary kernel and* $k(\boldsymbol{x},\boldsymbol{x}) = 1$, *and* $\mathcal{X}$ *be finite. Assume that* ζ *follows a two-parameter exponential distribution with* $s = 2\log(|\mathcal{X}|/2)$ *and* $\lambda = 1/2$. *Then, the following inequality holds:*

$$\mathbb{E}[f(\boldsymbol{x}^*)] \leq \mathbb{E}\left[\max_{\boldsymbol{x} \in \mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \zeta_t^{1/2}\sigma_{t-1}(\boldsymbol{x})\right],$$

for all $t \geq 1$.

Proof. We here show the short proof of Lemma 4.2 although detailed proof is shown in Appendix E.2. From the tower property of the expectation, it suffices to show the following inequality:

$$\mathbb{E}_t[f(\boldsymbol{x}^*)] \leq \mathbb{E}_t\left[\max_{\boldsymbol{x} \in \mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \zeta_t^{1/2}\sigma_{t-1}(\boldsymbol{x})\right].$$

Since this inequality only considers the expectation given $\mathcal{D}_{t-1}$, we can fix $\mathcal{D}_{t-1}$ in the proof. Furthermore, from Lemma 4.1, we can obtain the following inequality:

$$F_t^{-1}(1-\delta) \leq \max_{\boldsymbol{x} \in \mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \beta_\delta^{1/2}\sigma_{t-1}(\boldsymbol{x}),$$

where $F_t(\cdot) \coloneqq \mathrm{Pr}_t\left(f(\boldsymbol{x}_*) < \cdot\right)$ is a cumulative distribution function of $f(\boldsymbol{x}_*)$, and F_t^{-1} is its inverse function. Then, substituting $U \sim \mathrm{Uni}(0,1)$ into δ and taking the expectation, we obtain the following inequality:

$$\mathbb{E}_t\left[f(\boldsymbol{x}_*)\right] \leq \mathbb{E}_U\left[\max_{\boldsymbol{x} \in \mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \beta_U^{1/2}\sigma_{t-1}(\boldsymbol{x})\right],$$

which can be derived in a similar way to inverse transform sampling. Hence, $\beta_U^{1/2} = 2\log(|\mathcal{X}|/2) - 2\log(U)$ results in a random variable, which follows the two-parameter exponential distribution. Consequently, we can obtain the desired bound. $\square$

Lemma 4.2 shows a direct upper bound of the expectation of $f(\boldsymbol{x}^*)$. A remarkable consequence of Lemma 4.2 is an upper bound of BCR as follows:

$$\begin{aligned}
\mathrm{BCR}_T &= \sum_{t=1}^{T} \mathbb{E}\left[f(\boldsymbol{x}^*) - v_t(\boldsymbol{x}_t)\right] + \mathbb{E}\left[v_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right] \\
&\leq \sum_{t=1}^{T} \mathbb{E}\left[v_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right],
\end{aligned}$$

where $v_t(\boldsymbol{x}) \coloneqq \mu_{t-1}(\boldsymbol{x}) + \zeta_t^{1/2}\sigma_{t-1}(\boldsymbol{x})$. This upper bound eliminates the first term, which cannot be bounded without increasing the confidence parameter in the conventional regret decomposition Eq. (2). Importantly, $\mathbb{E}[\zeta_t] = 2 + s$ is a constant since s and λ for ζ does not depend on t in Lemma 4.2. Thus, using Lemma 4.2, we can obtain the tighter BCR bound for the finite input domain:

Theorem 4.2 (BCR bound for finite domain)**.** *Let* $f \sim \mathcal{GP}(0,k)$, *where* k *is a stationary kernel and* $k(\boldsymbol{x},\boldsymbol{x}) = 1$, *and* $\mathcal{X}$ *be finite. Assume that* ζ_t *follows a two-parameter exponential distribution with* $s = 2\log(|\mathcal{X}|/2)$ *and* $\lambda = 1/2$ *for any* $t \geq 1$. *Then, by running IRGP-UCB with* ζ_t, *BCR can be bounded as follows:*

$$\mathrm{BCR}_T \leq \sqrt{C_1 C_2 T \gamma_T},$$

where $C_1 \coloneqq 2/\log(1 + \sigma^{-2})$ *and* $C_2 \coloneqq 2 + s$.

See Appendix F for detailed proof.

Theorem 4.2 has two important implications. First, Theorem 4.2 shows the convergence rate $O(\sqrt{T\gamma_T})$, in which multiplicative $\sqrt{\log T}$ factor is improved compared with $O(\sqrt{T\gamma_T \log T})$ achieved by GP-UCB (Srinivas et al., 2010), RGP-UCB, and TS (Russo & Van Roy, 2014). Second, more importantly, IRGP-UCB does not need to schedule the parameters of ζ_t, i.e., s and λ, in contrast to GP-UCB and RGP-UCB. Therefore, through randomization, IRGP-UCB essentially alleviates the problem that the well-known $\beta_t \propto \log(t)$ strategy results in a practically too large confidence parameter. Further, over-exploration in the later iterations can be avoided. The IRGP-UCB is the first GP-UCB-based method that enjoys the above technical and practical benefits.

On the other hand, note that the dependence $\sqrt{\log |\mathcal{X}|}$ remains as with the prior works (Srinivas et al., 2010; Russo & Van Roy, 2014; Kandasamy et al., 2018a; Paria et al., 2020). In BO, we are usually interested in the case that the total number of iterations $T < |\mathcal{X}|$. Thus, $\sqrt{\log |\mathcal{X}|}$ is dominant compared with $\sqrt{\log T}$ factor improvement.

It is worth noting that our key lemma (Lemma 4.2) mainly requires Lemma 4.1 only, which is expected to be satisfied in a wide range of exploration problems in which a GP is used as a surrogate model. Therefore, we conjecture that the same proof technique can be applicable to more advanced problem settings, such as multi-objective BO (Paria et al., 2020), for which further analyses are important future directions of our results.

Next, we show the BCR bound for the infinite (continuous) domain:

Theorem 4.3 (BCR bound for infinite domain). *Let $f \sim \mathcal{GP}(0, k)$, where k is a stationary kernel, $k(\boldsymbol{x}, \boldsymbol{x}) = 1$, and Assumption 2.1 holds. Assume that ζ_t follows a two-parameter exponential distribution with $s_t = 2d \log(bdrt^2(\sqrt{\log(ad)} + \sqrt{\pi}/2)) - 2\log 2$ and $\lambda = 1/2$ for any $t \geq 1$. Then, by running IRGP-UCB, BCR can be bounded as follows:*

$$\mathrm{BCR}_T \leq \frac{\pi^2}{6} + \sqrt{C_1 T \gamma_T (2 + s_T)},$$

where $C_1 := 2/\log(1 + \sigma^{-2})$.

See Appendix G for the proof. Unfortunately, in this case, $\mathbb{E}[\zeta_t] = O(\log t)$ and the resulting BCR bound is $O(\sqrt{T\gamma_T \log T})$. This is because the discretization of input domain $\mathcal{X}_t$, which is refined as $|\mathcal{X}_t| = O(t^{2d})$, is required to obtain the BCR bound for the infinite domain. Since we cannot avoid the dependence on $|\mathcal{X}|$ also in Theorem 4.2, the resulting BCR bound in Theorem 4.3 requires the term $\log(|\mathcal{X}_t|) = O(2d \log(t))$.

As discussed above, our proof for BCR needs to refine the discretization as $|\mathcal{X}_t| = O(t^{2d})$ to bound the summation of the discretization error. On the other hand, if we aim to bound $\mathrm{BSR}_T < \eta$, the discretization error at an iteration is only required to be smaller than η. Therefore, the refinement of the discretization is not necessarily needed. Hence, we can obtain the following corollaries for finite and infinite input domains, respectively:

Corollary 4.1 (BSR bound for finite domain). *Fix a required accuracy $\eta > 0$. Assume the same condition as in Theorem 4.2. Then, by running IRGP-UCB, $\mathrm{BSR}_T \leq \eta$ by at most T function evaluations, where T is the smallest positive integer satisfying the following inequality:*

$$\sqrt{\frac{C_1 C_2 \gamma_T}{T}} \leq \eta,$$

where $C_1 := 2/\log(1 + \sigma^{-2})$ and $C_2 := 2 + s$.

Corollary 4.2 (BSR bound for infinite domain). *Fix a required accuracy $\eta > 0$. Assume the same condition as in Theorem 4.3 except for $s_\eta = 2d \log(2bdr(\sqrt{\log(ad)} + \sqrt{\pi}/2)/\eta) - 2\log 2$. Then, by running IRGP-UCB, $\mathrm{BSR}_T \leq \eta$ by at most T function evaluations, where T is the smallest positive integer satisfying the following inequality:*

$$\sqrt{\frac{C_1(2 + s_\eta)\gamma_T}{T}} \leq \frac{\eta}{2},$$

where $C_1 := 2/\log(1 + \sigma^{-2})$.

See Appendixes F and G for the proof, respectively. These corollaries suggest that, for both finite and infinite domains, an arbitrary accuracy solution can be obtained using finite function evaluations, whose rate is shown by the inequalities. In Corollary 4.2, s_η depends on η instead of t. Thus, in both corollaries, the convergence rate of the left-hand side with respect to T is $O(\sqrt{\gamma_T/T})$, which is improved compared with $O(\sqrt{\gamma_T \log T/T})$ achieved by GP-UCB, RGP-UCB, and TS. This faster convergence rate is an essential improvement derived based on Lemma 4.2. Therefore, IRGP-UCB does not need to schedule ζ_t if we require optimization rather than cumulative regret minimization.

5. Experiments

We demonstrate the experimental results on synthetic and benchmark functions and the materials dataset provided in (Liang et al., 2021). As a baseline, we performed EI (Mockus et al., 1978), TS (Russo & Van Roy, 2014), MES (Wang & Jegelka, 2017), joint entropy search (JES) (Hvarfner et al., 2022), and GP-UCB (Srinivas et al., 2010). For the posterior sampling in TS, MES, and JES, we used random Fourier features (Rahimi & Recht, 2008). For
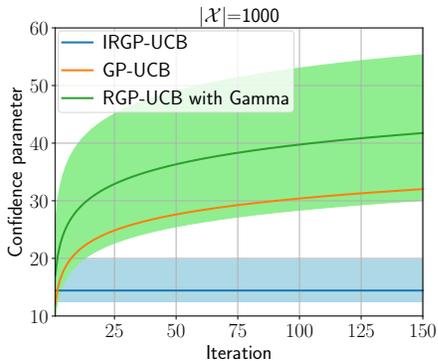
Figure 1. Confidence parameter of GP-UCB-based methods. For GP-UCB, the solid line represents β_t. For RGP-UCB and IRGP-UCB, the solid lines and shaded area represent the expectations $\mathbb{E}[\zeta_t]$ and 95% credible intervals, respectively.
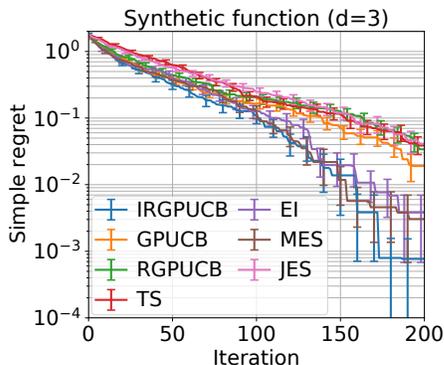

Figure 2. Average and standard errors of simple regret.

Monte Carlo estimation in MES and JES, we used ten samples. We evaluate the practical performance of BO by simple regret $f(\boldsymbol{x}^*) - \max_{t \leq T} f(\boldsymbol{x}_t)$.

5.1. Synthetic Function Experiment

We use synthetic functions generated from $\mathcal{GP}(0, k)$, where k is Gaussian kernel with a length scale parameter $\ell = 0.1$ and the input dimension $d = 3$. We set the noise variance $\sigma^2 = 10^{-4}$. The input domain consists of equally divided points in $[0, 0.9]$, i.e., $\mathcal{X} = \{0, 0.1, \ldots, 0.9\}^d$ and $|\mathcal{X}| = 1000$. Figure 1 shows the theoretical confidence parameters of GP-UCB, RGP-UCB with $\mathrm{Gamma}(\kappa_t = \log(|\mathcal{X}|t^2)/\log(1.5), \theta = 1)$, and IRGP-UCB with a two-parameter exponential distribution. We can see that the confidence parameters for GP-UCB and RGP-UCB are considerably large due to logarithmic increase, particularly in the later iterations. In contrast, the confidence parameter of IRGP-UCB is drastically smaller than that of GP-UCB and not changed with respect to the iteration. Therefore, we can observe that IRGP-UCB alleviates the over-exploration.

Next, we report the simple regret in Figure 2. Ten synthetic

functions and ten initial training datasets are randomly generated. Thus, the average and standard error for 10×10 trials are reported. The hyperparameters for GP are fixed to those used to generate the synthetic functions. The confidence parameters for GP-UCB, RGP-UCB, and IRGP-UCB are set as in Figure 1. We can see that the regrets of GP-UCB, RGP-UCB, TS, and JES converge slowly. We conjecture that this slow convergence came from over-exploration. On the other hand, EI, MES, and IRGP-UCB show fast convergence. In particular, IRGP-UCB achieves the best average in most iterations. These results suggest that IRGP-UCB bridges the gap between theory and practice in contrast to GP-UCB and RGP-UCB.

5.2. Benchmark Function Experiments

We employ three benchmark functions called Holder table ($d = 2$), Cross in tray($d = 2$), and Ackley ($d = 4$) functions, whose analytical forms are shown at https://www.sfu.ca/~ssurjano/optimization.html. For each function, we report the average and standard error for 10 trials using ten random initial datasets $\mathcal{D}_0$, where $|\mathcal{D}_0| = 2^d$. We set the noise variance $\sigma^2 = 10^{-4}$. We used the Gaussian kernel with automatic relevance determination, whose hyperparameter was selected by marginal likelihood maximization per 5 iterations (Rasmussen & Williams, 2005).

In this experiment, since the input domain is continuous, the theoretical choice of the confidence parameter contains an unknown variable. Thus, we use the heuristic choice for confidence parameters. For GP-UCB, we set the confidence parameter as $\beta_t = 0.2d \log(2t)$, which is the heuristics used in (Kandasamy et al., 2015; 2017). For RGP-UCB, we set $\zeta_t \sim \mathrm{Gamma}(\kappa_t, \theta = 1)$ with $\kappa_t = 0.2d \log(2t)$ since $\mathbb{E}[\zeta_t]$ must have the same order as β_t (note that $\mathbb{E}[\zeta_t] = \theta \kappa_t$). For IRGP-UCB, we set $s = d/2$ and $\lambda = 1/2$. Note that λ is equal to the theoretical setting, and s captures the dependence on d, as shown in Corollary 4.2.

Figure 3(a) shows the results. In the Holder table function, JES shows faster convergence until 40 iterations. However, the regrets of all baseline methods stagnate in $[10^{-1}, 10^0]$. In contrast, the regret of IRGP-UCB converges to 10^{-3} until 60 iterations. In the Cross in tray function, IRGP-UCB showed rapid convergence at approximately 50 iterations. In the Ackley function, IRGP-UCB constantly belonged to the top group and showed minimum regret after 125 iterations. Since the input dimension is relatively large, TS and JES, which depends on the sampled maxima $\boldsymbol{x}_*$, deteriorate by over-exploration. Throughout the experiments, IRGP-UCB outperforms TS, GP-UCB, and RGP-UCB, which achieves sub-linear BCR bounds, and EI and MES, which are practically well-used. This result supports the effectiveness of randomization using the two-parameter exponential distribution.
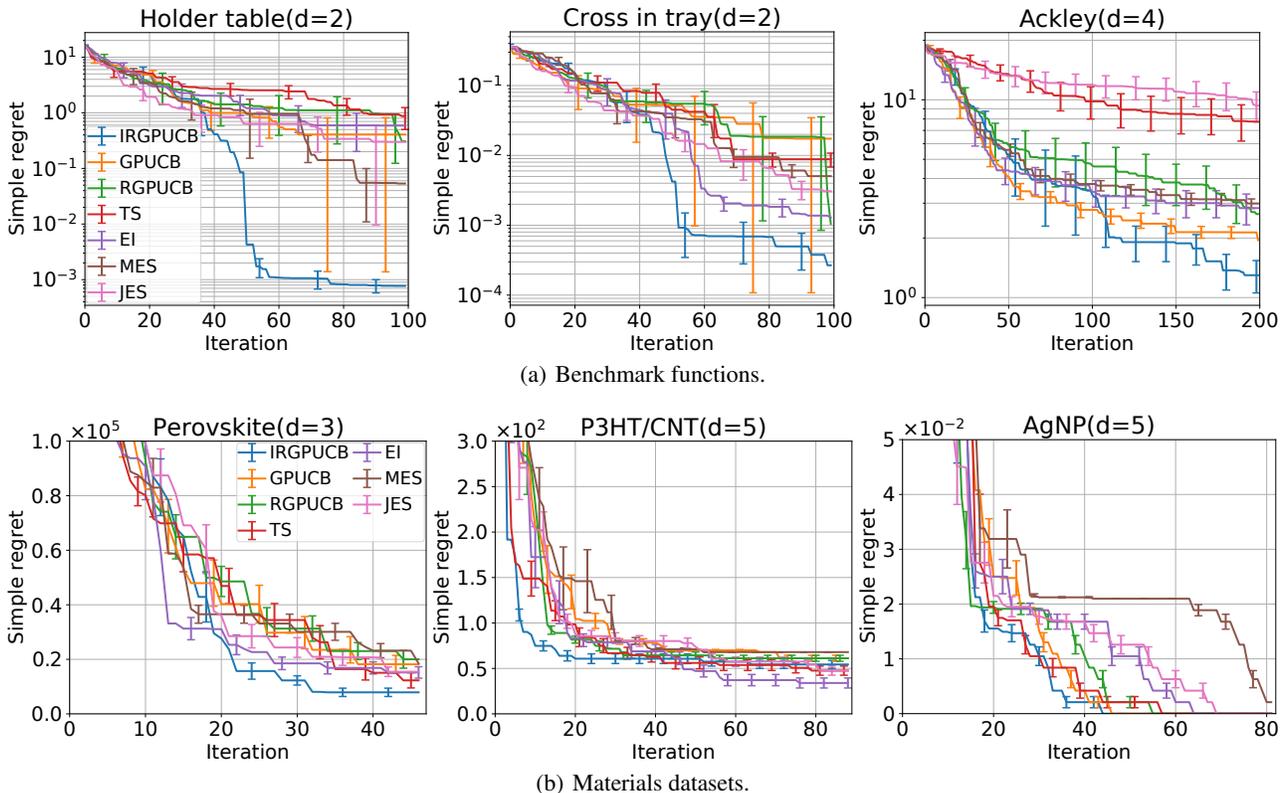
(a) Benchmark functions.

(b) Materials datasets.

Figure 3. Average and standard errors of simple regret.

5.3. Real-World Dataset Experiments

This section provides the experimental results on the materials datasets provided in (Liang et al., 2021). In the perovskite dataset (Sun et al., 2021), we optimize environmental stability with respect to composition parameters for halide perovskite ($d = 3$ and $|\mathcal{X}| = 94$). In the P3HT/CNT dataset (Bash et al., 2021), we optimize electrical conductivity with respect to composition parameters for carbon nanotube polymer blend ($d = 5$ and $|\mathcal{X}| = 178$). In the AgNP dataset (Mekki-Berrada et al., 2021), we optimize the absorbance spectrum of synthesized silver nanoparticles with respect to processing parameters for synthesizing triangular nanoprisms ($d = 5$ and $|\mathcal{X}| = 164$). See (Liang et al., 2021) for more details about each dataset.

We set the initial dataset size $|\mathcal{D}_0| = 2$ as with (Liang et al., 2021). Since the dataset size is small at earlier iterations and the dataset contains fluctuations from real-world experiments, we observed that hyperparameter tuning could be unstable. Thus, we optimized the hyperparameters of the RBF kernel in each iteration to avoid repeatedly obtaining samples using an inappropriate hyperparameter. The other settings matched those used in the benchmark function experiments.

Figure 3(b) shows the results. In the perovskite dataset,

IRGP-UCB constantly belonged to the top group and showed the best performance after 20 iterations. In the P3HT/CNT dataset, EI converged to the smallest value after 60 iterations. On the other hand, IRGP-UCB shows faster convergence during the first 20 iterations. In the AgNP dataset, IRGP-UCB found the optimal point until 42 iterations in all the trials earliest. In this experiment, heuristic methods, EI, MES, and JES, showed worse performance and required at least 60 function evaluations to find the optimal point. Consequently, we can observe the effectiveness of IRGP-UCB against real-world datasets.

6. Conclusion

First, by showing the generalized BCR bounds for RGP-UCB, this study showed that randomization without ingenuity does not improve the regret bounds. Then, this study proposed an improved randomized GP-UCB (IRGP-UCB) using the two-parameter exponential distribution, whose Bayesian regret bound is tighter than known results. Furthermore, IRGP-UCB does not require an increase in the confidence parameter with respect to the number of iterations. Lemma 4.2, which directly bounds the expectation of the maximum value, plays a key role in the proof. Additionally, we demonstrated the practical effectiveness of IRGP-UCB through extensive experiments, including the

application to the materials dataset.

Several directions for future works can be considered. First, whether we can show the tighter BCR bounds for the continuous domain is of interest. Second, since the Bayesian regret analysis of TS depends on the usual UCB, we may improve the BCR bound of TS using randomized UCB. Last, we may be able to extend IRGP-UCB to the other various practical settings, where the usual UCB-based methods have been extended, e.g., multi-objective BO (Paria et al., 2020; Zuluaga et al., 2016; Suzuki et al., 2020), multi-fidelity BO (Kandasamy et al., 2016; 2017; Takeno et al., 2020; 2022a), parallel BO (Contal et al., 2013; Desautels et al., 2014), high-dimensional BO (Kandasamy et al., 2015), cascade BO (Kusakawa et al., 2022), and robust BO(Bogunovic et al., 2018; Iwazaki et al., 2021; Inatsu et al., 2022).

Acknowledgements

References

Bash, D., Cai, Y., Chellappan, V., Wong, S. L., Yang, X., Kumar, P., Tan, J. D., Abutaha, A., Cheng, J. J., Lim, Y.-F., et al. Multi-fidelity high-throughput optimization of electrical conductivity in P3HT-CNT composites. *Advanced Functional Materials*, pp. 2102606, 2021.

Beg, M. A. On the estimation of pr $\{Y < X\}$ for the two-parameter exponential distribution. *Metrika*, 27(1):29–34, 1980.

Berk, J., Gupta, S., Rana, S., and Venkatesh, S. Randomised Gaussian process upper confidence bound for Bayesian optimisation. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence*, pp. 2284–2290. International Joint Conferences on Artificial Intelligence Organization, 2020.

Bogunovic, I., Scarlett, J., Jegelka, S., and Cevher, V. Adversarially robust optimization with Gaussian processes. In *Advances in Neural Information Processing Systems 31*, pp. 5760–5770. Curran Associates, Inc., 2018.

Bull, A. D. Convergence rates of efficient global optimization algorithms. *Journal of Machine Learning Research*, 12(10), 2011.

Chowdhury, S. R. and Gopalan, A. On kernelized multi-armed bandits. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 844–853, 2017.

Contal, E., Buffoni, D., Robicquet, A., and Vayatis, N. Parallel Gaussian process optimization with upper confidence bound and pure exploration. In *Proceedings of the 2013th European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 225–240. Springer-Verlag, 2013.

Desautels, T., Krause, A., and Burdick, J. W. Parallelizing exploration-exploitation tradeoffs in Gaussian process bandit optimization. *Journal of Machine Learning Research*, 15:4053–4103, 2014.

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

Hennig, P. and Schuler, C. J. Entropy search for information-efficient global optimization. *Journal of Machine Learning Research*, 13(57):1809–1837, 2012.

Hernández-Lobato, J. M., Hoffman, M. W., and Ghahramani, Z. Predictive entropy search for efficient global optimization of black-box functions. In *Advances in Neural Information Processing Systems 27*, pp. 918–926. Curran Associates, Inc., 2014.

Hvarfner, C., Hutter, F., and Nardi, L. Joint entropy search for maximally-informed Bayesian optimization. In *Advances in Neural Information Processing Systems*, 2022. To appear.

Inatsu, Y., Takeno, S., Karasuyama, M., and Takeuchi, I. Bayesian optimization for distributionally robust chance-constrained problem. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162, pp. 9602–9621. PMLR, 2022.

Iwazaki, S., Inatsu, Y., and Takeuchi, I. Mean-variance analysis in Bayesian optimization under uncertainty. In *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, volume 130, pp. 973–981. PMLR, 2021.

Janz, D., Burt, D., and Gonzalez, J. Bandit optimisation of functions in the Matérn kernel RKHS. In *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pp. 2486–2495, 2020.

Kandasamy, K., Schneider, J., and Poczos, B. High dimensional Bayesian optimisation and bandits via additive models. In *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pp. 295–304, 2015.

Kandasamy, K., Dasarathy, G., Oliva, J., Schneider, J., and Póczos, B. Gaussian process bandit optimisation with multi-fidelity evaluations. In *Advances in Neural Information Processing Systems 29*, pp. 1000–1008. Curran Associates, Inc., 2016.

Kandasamy, K., Dasarathy, G., Schneider, J., and Póczos, B. Multi-fidelity Bayesian optimisation with continuous approximations. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 1799–1808, 2017.

Kandasamy, K., Krishnamurthy, A., Schneider, J., and Póczos, B. Parallelised Bayesian optimisation via Thompson sampling. In *Proceedings of the 21st International Conference on Artificial Intelligence and Statistics*, volume 84 of *Proceedings of Machine Learning Research*, pp. 133–142, 2018a.

Kandasamy, K., Neiswanger, W., Schneider, J., Poczos, B., and Xing, E. P. Neural architecture search with Bayesian optimisation and optimal transport. *Advances in neural information processing systems 31*, pp. 2016–2025, 2018b.

Korovina, K., Xu, S., Kandasamy, K., Neiswanger, W., Poczos, B., Schneider, J., and Xing, E. ChemBO: Bayesian optimization of small organic molecules with synthesizable recommendations. In *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pp. 3393–3403, 2020.

Kusakawa, S., Takeno, S., Inatsu, Y., Kutsukake, K., Iwazaki, S., Nakano, T., Ujihara, T., Karasuyama, M., and Takeuchi, I. Bayesian optimization for cascade-type multistage processes. *Neural Computation*, 34(12):2408–2431, 2022.

Lam, K., Sinha, B. K., and Wu, Z. Estimation of parameters in a two-parameter exponential distribution using ranked set sample. *Annals of the Institute of Statistical Mathematics*, 46(4):723–736, 1994.

Liang, Q., Gongora, A., Ren, Z., Tiihonen, A., Liu, Z., Sun, S., Deneault, J., Bash, D., Mekki-Berrada, F., Khan, S., et al. Benchmarking the performance of Bayesian optimization across multiple experimental materials science domains. *npj Computational Material*, 7(188), 2021.

Mekki-Berrada, F., Ren, Z., Huang, T., Wong, W. K., Zheng, F., Xie, J., Tian, I. P. S., Jayavelu, S., Mahfoud, Z., Bash, D., et al. Two-step machine learning enables optimized nanoparticle synthesis. *npj Computational Materials*, 7(1):1–10, 2021.

Mockus, J., Tiesis, V., and Zilinskas, A. The application of Bayesian methods for seeking the extremum. *Towards Global Optimization*, 2(117-129):2, 1978.

Paria, B., Kandasamy, K., and Póczos, B. A flexible framework for multi-objective Bayesian optimization using random scalarizations. In *Proceedings of The 35th Uncertainty in Artificial Intelligence Conference*, volume 115 of *Proceedings of Machine Learning Research*, pp. 766–776, 2020.

Rahimi, A. and Recht, B. Random features for large-scale kernel machines. In *Advances in Neural Information Processing Systems 20*, pp. 1177–1184. Curran Associates, Inc., 2008.

Rasmussen, C. E. and Williams, C. K. I. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005.

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

Shahriari, B., Swersky, K., Wang, Z., Adams, R., and De Freitas, N. Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE*, 104 (1):148–175, 2016.

Snoek, J., Larochelle, H., and Adams, R. P. Practical Bayesian optimization of machine learning algorithms. In *Advances in Neural Information Processing Systems 25*, pp. 2951–2959. Curran Associates, Inc., 2012.

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

Sun, S., Tiihonen, A., Oviedo, F., Liu, Z., Thapa, J., Zhao, Y., Hartono, N. T. P., Goyal, A., Heumueller, T., Batali, C., et al. A data fusion approach to optimize compositional stability of halide perovskites. *Matter*, 4(4):1305–1322, 2021.

Suzuki, S., Takeno, S., Tamura, T., Shitara, K., and Karasuyama, M. Multi-objective Bayesian optimization using Pareto-frontier entropy. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119, pp. 9279–9288. PMLR, 2020.

Takeno, S., Fukuoka, H., Tsukada, Y., Koyama, T., Shiga, M., Takeuchi, I., and Karasuyama, M. Multi-fidelity Bayesian optimization with max-value entropy search and its parallelization. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119, pp. 9334–9345. PMLR, 2020.

Takeno, S., Fukuoka, H., Tsukada, Y., Koyama, T., Shiga, M., Takeuchi, I., and Karasuyama, M. A Generalized Framework of Multifidelity Max-Value Entropy Search Through Joint Entropy. *Neural Computation*, 34(10): 2145–2203, 2022a.

Takeno, S., Tamura, T., Shitara, K., and Karasuyama, M. Sequential and parallel constrained max-value entropy search via information lower bound. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 20960–20986, 2022b.

Ueno, T., Rhone, T. D., Hou, Z., Mizoguchi, T., and Tsuda, K. COMBO: An efficient Bayesian optimization library for materials science. *Materials discovery*, 4:18–21, 2016.

Vakili, S., Khezeli, K., and Picheny, V. On information gain and regret bounds in Gaussian process bandits. In *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, volume 130 of *Proceedings of Machine Learning Research*, pp. 82–90, 2021.

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

Wang, Z., Zhou, B., and Jegelka, S. Optimization as estimation with Gaussian processes in bandit settings. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, volume 51 of *Proceedings of Machine Learning Research*, pp. 1022–1031, 2016.

Zuluaga, M., Krause, A., and Püschel, M. e-PAL: An active learning approach to the multi-objective optimization problem. *Journal of Machine Learning Research*, 17 (104):1–32, 2016.

A. BSR with Recommended Input

When the observation is contaminated by the noise, we cannot know the observed input, which minimizes BSR defined in Eq. (1). Thus, it is required that the algorithm explicitly returns low regret input. Then, we can consider the following variant of BSR:

$$\mathrm{BSR}_T := \mathbb{E}\big[f(\boldsymbol{x}^*) - f(\hat{\boldsymbol{x}}_T)\big],$$

where $\hat{\boldsymbol{x}}_t$ is a recommendation point from the algorithm. To bound this BSR, we can consider the following two options for the recommendation:

$$\hat{\boldsymbol{x}}_t = \arg\max_{\boldsymbol{x} \in \mathcal{X}} \mu_{t-1}(\boldsymbol{x}),$$

$$\hat{\boldsymbol{x}}_t = \arg\max_{\boldsymbol{x} \in \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_{t-1}\}} \mu_{t-1}(\boldsymbol{x}).$$

Although we here consider the finite input domain for simplicity, it can be extended to the infinite domain in the same way as our proof of Corollaries 4.2. Then, the BSR defined by the both $\hat{\boldsymbol{x}}_t$ can be bounded as follows:

$$
\begin{aligned}
\mathrm{BSR}_T &= \mathbb{E}\big[f(\boldsymbol{x}^*) - f(\hat{\boldsymbol{x}}_T)\big] \\
&= \mathbb{E}_{\mathcal{D}_{T-1}}\big[\mathbb{E}_f\big[f(\boldsymbol{x}^*)|\mathcal{D}_{T-1}\big] - \mu_{T-1}(\hat{\boldsymbol{x}}_T)\big] \\
&\le \mathbb{E}_{\mathcal{D}_{T-1}}\bigg[\mathbb{E}_f\big[f(\boldsymbol{x}^*)|\mathcal{D}_{T-1}\big] - \frac{1}{T}\sum_{t=1}^{T}\mu_{T-1}(\boldsymbol{x}_t)\bigg] \qquad \big(\because \forall 1 \le t \le T, \mu_{T-1}(\hat{\boldsymbol{x}}_T) \le \mu_{T-1}(\boldsymbol{x}_t)\big) \\
&= \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_{\mathcal{D}_{T-1}}\Big[\mathbb{E}_f\big[f(\boldsymbol{x}^*)|\mathcal{D}_{T-1}\big] - \mathbb{E}_f\big[f(\boldsymbol{x}_t)|\mathcal{D}_{T-1}\big]\Big] \\
&\le \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\big[f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t)\big] \\
&= \frac{1}{T}\mathrm{BCR}_T.
\end{aligned}
$$

Consequently, we can show the upper bound of this variant of BSR, which is the same as that of the BSR defined in Eq. (1).

B. Proof for BCR Bound of GP-UCB

Srinivas et al. (2010) provided the high-probability bounds for cumulative regret, not BCR bounds. On the other hand, although Paria et al. (2020) have shown BCR bounds for the GP-UCB-based multi-objective BO method, they did not provide explicit proof for standard single-objective BO. Thus, although the following result is an almost direct consequence of Theorem 2 in (Paria et al., 2020), we here provide the explicit proof for standard GP-UCB for completeness. On the other hand, for the case of (ii), the bound with respect to a is tightened as $O\big(\log(\log(a))\big)$ compared with $O(\log(a))$ (Kandasamy et al., 2018a; Paria et al., 2020), which is mainly based on Lemma H.1.

Theorem B.1. *Suppose that $f \sim \mathcal{GP}(0, k)$, where k is a stationary kernel and $k(\boldsymbol{x}, \boldsymbol{x}) = 1$.*

(i) *When $\mathcal{X}$ is a finite set, GP-UCB with $\beta_t = 2\log(|\mathcal{X}|t^2/\sqrt{2\pi})$ achieves the following BCR bound:*

$$\mathrm{BCR}_T \le \frac{\pi^2}{6} + \sqrt{C_1 T \beta_T \gamma_T},$$

where $C_1 := 2/\log(1 + \sigma^{-2})$.

(ii) *When $\mathcal{X}$ is an infinite set, let Assumption 2.1 holds. Then, GP-UCB with $\beta_t = 2d\log\big(bdrt^2\big(\sqrt{\log(ad)} + \sqrt{\pi}/2\big)\big) + 2\log(t^2/\sqrt{2\pi})$ achieves the following BCR bound:*

$$\mathrm{BCR}_T \le \frac{\pi^2}{3} + \sqrt{C_1 T \beta_T \gamma_T},$$

where $C_1 := 2/\log(1 + \sigma^{-2})$.

Proof. First, we show the proof of (i). For regret analysis of BCR, regret decomposition (Russo & Van Roy, 2014) with $U_t(\boldsymbol{x}) \coloneqq \mu_{t-1}(\boldsymbol{x}) + \beta_t^{1/2}\sigma_{t-1}(\boldsymbol{x})$ is used as follows:

$$
\begin{aligned}
\mathrm{BCR}_T &= \sum_{t=1}^{T} \mathbb{E}[f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t)] \\
&= \sum_{t=1}^{T} \mathbb{E}[f(\boldsymbol{x}^*) - U_t(\boldsymbol{x}^*) + U_t(\boldsymbol{x}^*) - U_t(\boldsymbol{x}_t) + U_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)] \\
&\leq \underbrace{\sum_{t=1}^{T} \mathbb{E}[f(\boldsymbol{x}^*) - U_t(\boldsymbol{x}^*)]}_{R_1} + \underbrace{\sum_{t=1}^{T} \mathbb{E}[U_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)]}_{R_2}. \qquad \left(\because U_t(\boldsymbol{x}^*) - U_t(\boldsymbol{x}_t) \leq 0\right)
\end{aligned}
$$

First, we consider R_1. We use the following fact for a Gaussian random variable $Z \sim \mathcal{N}(m, s^2)$ with $m \leq 0$:

$$
\mathbb{E}[Z_+] \leq \frac{s}{\sqrt{2\pi}} \exp\left\{-\frac{m^2}{2s^2}\right\}, \tag{4}
$$

where $Z_+ \coloneqq \max\{0, Z\}$. This fact has been frequently used in this literature (Russo & Van Roy, 2014; Kandasamy et al., 2018a; Paria et al., 2020). Then, we can extend R_1 as follows:

$$
\begin{aligned}
R_1 &= \sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}_{t-1}}[\mathbb{E}_t[f(\boldsymbol{x}^*) - U_t(\boldsymbol{x}^*)]] \\
&\leq \sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}_{t-1}}[\mathbb{E}_t[(f(\boldsymbol{x}^*) - U_t(\boldsymbol{x}^*))_+]] \qquad \left(\because f(\boldsymbol{x}^*) - U_t(\boldsymbol{x}^*) \leq (f(\boldsymbol{x}^*) - U_t(\boldsymbol{x}^*))_+\right) \\
&\leq \sum_{t=1}^{T} \sum_{\boldsymbol{x} \in \mathcal{X}} \mathbb{E}_{\mathcal{D}_{t-1}}\left[\mathbb{E}_t\left[(f(\boldsymbol{x}) - U_t(\boldsymbol{x}))_+\right]\right]. \qquad \left(\because (f(\boldsymbol{x}^*) - U_t(\boldsymbol{x}^*))_+ \leq \sum_{\boldsymbol{x} \in \mathcal{X}} (f(\boldsymbol{x}) - U_t(\boldsymbol{x}))_+\right)
\end{aligned}
$$

Then, since $f(\boldsymbol{x}) - U_t(\boldsymbol{x}) \mid \mathcal{D}_{t-1} \sim \mathcal{N}(-\beta_t^{1/2}(\boldsymbol{x})\sigma_{t-1}(\boldsymbol{x}), \sigma_{t-1}^2(\boldsymbol{x}))$, we can apply Eq. (4):

$$
\begin{aligned}
R_1 &\leq \sum_{t=1}^{T} \sum_{\boldsymbol{x} \in \mathcal{X}} \mathbb{E}_{\mathcal{D}_{t-1}}\left[\frac{\sigma_{t-1}(\boldsymbol{x})}{\sqrt{2\pi}} \exp\left\{-\frac{\beta_t}{2}\right\}\right] \qquad (\because \text{Eq. (4)}) \\
&\leq \sum_{t=1}^{T} |\mathcal{X}| \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{\beta_t}{2}\right\} \qquad (\because \sigma_{t-1}(\boldsymbol{x}) \leq \sigma_0(\boldsymbol{x}) = 1) \\
&= \sum_{t=1}^{T} \frac{1}{t^2} \\
&\leq \frac{\pi^2}{6}. \qquad \left(\because \sum_{t=1}^{\infty} \frac{1}{t^2} = \frac{\pi^2}{6}\right) \tag{5}
\end{aligned}
$$

Then, we bound R_2 as follows:

$$R_2 = \sum_{t=1}^{T} \mathbb{E}[U_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)]$$

$$= \sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}_{t-1}}[\mathbb{E}_t[U_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)]]$$

$$= \sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}_{t-1}}[U_t(\boldsymbol{x}_t) - \mu_{t-1}(\boldsymbol{x}_t)]$$

$$= \sum_{t=1}^{T} \mathbb{E}\left[\beta_t^{1/2}\sigma_{t-1}(\boldsymbol{x}_t)\right]$$

$$= \mathbb{E}\left[\sum_{t=1}^{T} \beta_t^{1/2}\sigma_{t-1}(\boldsymbol{x}_t)\right]$$

$$\leq \mathbb{E}\left[\sqrt{\sum_{t=1}^{T}\beta_t \sum_{t=1}^{T}\sigma_{t-1}^2(\boldsymbol{x}_t)}\right] \qquad \left(\because \text{Cauchy–Schwarz inequality}\right)$$

$$= \sqrt{\sum_{t=1}^{T}\beta_t}\,\mathbb{E}\left[\sqrt{\sum_{t=1}^{T}\sigma_{t-1}^2(\boldsymbol{x}_t)}\right]$$

$$\leq \sqrt{T\beta_T}\,\mathbb{E}\left[\sqrt{\sum_{t=1}^{T}\sigma_{t-1}^2(\boldsymbol{x}_t)}\right] \qquad \left(\because \beta_1 < \beta_2 < \cdots < \beta_T\right).$$

Then, from Lemma 5.4 in (Srinivas et al., 2010), $\sqrt{\sum_{t=1}^{T}\sigma_{t-1}^2(\boldsymbol{x}_t)} \leq \sqrt{C_1\gamma_T}$ with probability 1. Thus, we obtain

$$R_2 \leq \sqrt{C_1 T \beta_T \gamma_T}. \qquad (6)$$

Consequently, by combining Eqs. (5) and (6), we obtain the theorem for (i).

Next, we consider (ii). Purely for the sake of analysis, we use a set of discretization $\mathcal{X}_t \subset \mathcal{X}$ for $t \geq 1$. For any $t \geq 1$, let $\mathcal{X}_t \subset \mathcal{X}$ be a finite set with each dimension equally divided into $\tau_t = bdrt^2\left(\sqrt{\log(ad)} + \sqrt{\pi}/2\right)$. Thus, $|\mathcal{X}_t| = \tau_t^d$. In addition, we define $[\boldsymbol{x}]_t$ as the nearest points in $\mathcal{X}_t$ of $\boldsymbol{x} \in \mathcal{X}$.

Then, BCR can be decomposed as follows:

$$\text{BCR}_T = \sum_{t=1}^{T}\mathbb{E}\left[f(\boldsymbol{x}^*) - f([\boldsymbol{x}^*]_t) + f([\boldsymbol{x}^*]_t) - U_t([\boldsymbol{x}^*]_t) + U_t([\boldsymbol{x}^*]_t) - U_t(\boldsymbol{x}_t) + U_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right]$$

$$\leq \sum_{t=1}^{T}\mathbb{E}\left[f(\boldsymbol{x}^*) - f([\boldsymbol{x}^*]_t) + f([\boldsymbol{x}^*]_t) - U_t([\boldsymbol{x}^*]_t) + U_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right],$$

where $\mathbb{E}\left[U_t([\boldsymbol{x}^*]_t) - U_t(\boldsymbol{x}_t)\right] \leq 0$ since $\boldsymbol{x}_t = \arg\max_{\boldsymbol{x}\in\mathcal{X}}U_t(\boldsymbol{x})$. Furthermore, we obtain the following:

$$\sum_{t=1}^{T}\mathbb{E}\left[f(\boldsymbol{x}^*) - f([\boldsymbol{x}^*]_t)\right] \leq \sum_{t=1}^{T}\mathbb{E}\left[\sup_{\boldsymbol{x}\in\mathcal{X}}f(\boldsymbol{x}) - f([\boldsymbol{x}]_t)\right]$$

$$\leq \sum_{t=1}^{T}\frac{1}{t^2} \qquad \left(\because \text{Lemma } H.2\right)$$

$$\leq \frac{\pi^2}{6} \qquad \left(\because \sum_{t=1}^{\infty}\frac{1}{t^2} = \frac{\pi^2}{6}\right)$$

The remaining terms $\sum_{t=1}^{T} \mathbb{E}\big[f([\boldsymbol{x}^*]_t) - U_t([\boldsymbol{x}^*]_t) + U_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\big]$ can be bounded as with the case (i) by setting $\beta_t = 2\log\big(|\mathcal{X}_t|t^2/\sqrt{2\pi}\big)$. That is,

$$\sum_{t=1}^{T} \mathbb{E}\big[f([\boldsymbol{x}^*]_t) - U_t([\boldsymbol{x}^*]_t) + U_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\big] \leq \frac{\pi^2}{6} + \sqrt{C_1 T \beta_T \gamma_T}.$$

By substituting $|\mathcal{X}_t| = \big(bdrt^2\big(\sqrt{\log(ad)} + \sqrt{\pi}/2\big)\big)^d$, we conclude the proof. $\qquad\square$

C. Technical Issues of Proof in (Berk et al., 2020)

The proofs of Theorems 1, 2, and 3 in (Berk et al., 2020) appear to contain many technical issues. We enumerated them using the notation in (Berk et al., 2020) as follows:

- Lemma 1 does not necessarily hold since $\alpha([\boldsymbol{x}_t]_\tau)$ can be smaller than $\alpha([\boldsymbol{x}^*]_\tau)$ although $\alpha(\boldsymbol{x}_t) \geq \alpha(\boldsymbol{x}^*)$.

- To treat $f(\boldsymbol{x}_t)$ as the random variable, which follows $\mathcal{N}(\mu_t(\boldsymbol{x}), \sigma_t^2(\boldsymbol{x}))$, the conditioning by D_t is required. However, they do not perform the conditioning in most equations, such as Eqs. (11) and (12).

- From Eq. (13) to the next equations, the term $|\mathcal{X}_{\mathrm{dis}}|$ is eliminated. This term is imperatively required.

- In Eqs. (11), (12), (16), and (17), variables such as $\sigma_{t-1}([\boldsymbol{x}_t]_r)$ are random variables that depend on D_t and $[\boldsymbol{x}_t]_r$. However, these random variables are taken out from the expectation.

- From Eq.(17) to Eq. (19), a discretization for $\boldsymbol{x}_t$ is eliminated.

- In Eq. (18), an approximately equal sign $\approx$ is used, which is inappropriate for theoretical proof.

- The convergence rate of the term $F^{-1}(1 - 1/T)$ in the resulting bound is not clarified.

Hence, Theorems 1, 2, and 3 in (Berk et al., 2020) are not shown using their proofs. Furthermore, even if Theorem 3 holds, it would be insufficient to show the sub-linear BCR bound.

D. Proof of Theorem for RGP-UCB

Here, we rectify and generalize the theorems in (Berk et al., 2020). First, we show the generalized theorem, which shows the upper bound using the expectation and MGFs of $\{\zeta_t\}_{t\geq1}$. Next, we provide some examples of the distributions, which can achieve the sub-linear regret bound by appropriately controlling the parameters.

D.1. Generalized BCR Bounds for RGP-UCB

The following theorem shows the upper bound of BCR for setting $\{\zeta_t\}_{t\geq1}$ as arbitrary non-negative random variables:

Theorem D.1. *Let $f \sim \mathcal{GP}(0, k)$, where k is a stationary kernel k and $k(\boldsymbol{x}, \boldsymbol{x}) = 1$. Let $\{\zeta_t\}_{t\geq1}$ be a sequence of non-negative random variables and their MGFs be $M_t(\cdot)$, where $M_t(-1/2) < \infty$ at least.*

(i) When $\mathcal{X}$ is a finite set, by running RGP-UCB with ζ_t, BCR can be bounded as follows:

$$\mathrm{BCR}_T \leq \frac{|\mathcal{X}|}{\sqrt{2\pi}} \sum_{t=1}^{T} M_t(-1/2) + \sqrt{\sum_{t=1}^{T} \mathbb{E}[\zeta_t]C_1\gamma_T},$$

where $C_1 := 2/\log(1 + \sigma^{-2})$.

(ii) When $\mathcal{X}$ is an infinite set, let Assumption 2.1 holds. By running RGP-UCB with ζ_t, BCR can be bounded as follows:

$$\mathrm{BCR}_T \leq \frac{\pi^2}{6} + \frac{1}{\sqrt{2\pi}} \sum_{t=1}^{T} (\tau_t)^d M_t(-1/2) + \sqrt{C_1 \sum_{t=1}^{T} \mathbb{E}[\zeta_t]\gamma_T},$$

where $C_1 := 2/\log(1 + \sigma^{-2})$ and $\tau_t = bdrt^2\big(\sqrt{\log(ad)} + \sqrt{\pi}/2\big)$.

Proof. The procedure of the proof is similar to the proof for the standard GP-UCB shown in Appendix B. BCR can be decomposed by $V_t(\boldsymbol{x}) = \mu_{t-1}(\boldsymbol{x}) + \zeta_t^{1/2}\sigma_{t-1}(\boldsymbol{x})$ as follows:

$$
\begin{aligned}
\text{BCR}_T &= \sum_{t=1}^T \mathbb{E}\left[f(\boldsymbol{x}^*) - V_t(\boldsymbol{x}^*) + V_t(\boldsymbol{x}^*) - V_t(\boldsymbol{x}_t) + V_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right] \\
&\leq \underbrace{\sum_{t=1}^T \mathbb{E}\left[f(\boldsymbol{x}^*) - V_t(\boldsymbol{x}^*)\right]}_{=:R_1} + \underbrace{\sum_{t=1}^T \mathbb{E}\left[V_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right]}_{=:R_2}. \qquad \left(\because V_t(\boldsymbol{x}^*) - V_t(\boldsymbol{x}_t) \leq 0\right)
\end{aligned}
$$

Then, we derive the upper bounds of R_1 and R_2, respectively.

First, we consider R_1. We use the following fact for Gaussian random variables $Z \sim \mathcal{N}(m, s^2)$ with $m \leq 0$:

$$
\mathbb{E}[Z_+] \leq \frac{s}{\sqrt{2\pi}}\exp\left\{-\frac{m^2}{2s^2}\right\}, \tag{7}
$$

where $Z_+ := \max\{0, Z\}$. This fact is frequently used in the literature (Russo & Van Roy, 2014; Kandasamy et al., 2018a; Paria et al., 2020). Then, we can extend R_1 as follows:

$$
\begin{aligned}
R_1 &= \sum_{t=1}^T \mathbb{E}\left[f(\boldsymbol{x}^*) - V_t(\boldsymbol{x}^*)\right] \\
&\leq \sum_{t=1}^T \mathbb{E}\left[\left(f(\boldsymbol{x}^*) - V_t(\boldsymbol{x}^*)\right)_+\right] && \left(\because f(\boldsymbol{x}^*) - V_t(\boldsymbol{x}^*) \leq \left(f(\boldsymbol{x}^*) - V_t(\boldsymbol{x}^*)\right)_+\right) \\
&\leq \sum_{t=1}^T \mathbb{E}\left[\sum_{\boldsymbol{x}\in\mathcal{X}}\left(f(\boldsymbol{x}) - V_t(\boldsymbol{x})\right)_+\right] && \left(\because \left(f(\boldsymbol{x}^*) - V_t(\boldsymbol{x}^*)\right)_+ \leq \sum_{\boldsymbol{x}\in\mathcal{X}}\left(f(\boldsymbol{x}) - V_t(\boldsymbol{x})\right)_+\right) \\
&= \sum_{t=1}^T \sum_{\boldsymbol{x}\in\mathcal{X}} \mathbb{E}_{\mathcal{D}_{t-1},\zeta_t}\left[\mathbb{E}\left[\left(f(\boldsymbol{x}) - V_t(\boldsymbol{x})\right)_+ \mid \mathcal{D}_{t-1}, \zeta_t\right]\right].
\end{aligned}
$$

Here, $\left(f(\boldsymbol{x}) - V_t(\boldsymbol{x})\right) \mid \mathcal{D}_{t-1}, \zeta_t$ follows $\mathcal{N}(-\zeta_t^{1/2}\sigma_{t-1}(\boldsymbol{x}), \sigma_{t-1}^2(\boldsymbol{x}))$. Thus, by using Eq. (7), we can obtain the following:

$$
\begin{aligned}
R_1 &\leq \sum_{t=1}^T \sum_{\boldsymbol{x}\in\mathcal{X}} \mathbb{E}_{\mathcal{D}_{t-1},\zeta_t}\left[\frac{\sigma_{t-1}(\boldsymbol{x})}{\sqrt{2\pi}}\exp\left\{-\frac{\zeta_t}{2}\right\}\right] && (\because \text{Eq. (7)}) \\
&\leq \frac{|\mathcal{X}|}{\sqrt{2\pi}}\sum_{t=1}^T \mathbb{E}_{\zeta_t}\left[\exp\left\{-\frac{\zeta_t}{2}\right\}\right] && (\because \sigma_{t-1}(\boldsymbol{x}) \leq \sigma_0(\boldsymbol{x}) = 1) \\
&= \frac{|\mathcal{X}|}{\sqrt{2\pi}}\sum_{t=1}^T M_t(-1/2). && (\because \text{Definition of MGF}) \tag{8}
\end{aligned}
$$

Second, we expand R_2 as follows:

$$
\begin{aligned}
R_2 &= \sum_{t=1}^{T} \mathbb{E}\left[V_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right] \\
&= \sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}_{t-1}, \zeta_t}\left[\mathbb{E}\left[V_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t) \mid \mathcal{D}_{t-1}, \zeta_t\right]\right] \\
&= \sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}_{t-1}, \zeta_t}\left[V_t(\boldsymbol{x}_t) - \mu_{t-1}(\boldsymbol{x}_t)\right] \\
&= \sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}_{t-1}, \zeta_t}\left[\zeta_t^{1/2}\sigma_{t-1}(\boldsymbol{x}_t)\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T} \zeta_t^{1/2}\sigma_{t-1}(\boldsymbol{x}_t)\right] \\
&\leq \mathbb{E}\left[\sqrt{\sum_{t=1}^{T}\zeta_t \sum_{t=1}^{T}\sigma_{t-1}^2(\boldsymbol{x}_t)}\right]. \qquad (\because \text{Cauchy–Schwarz inequality})
\end{aligned}
$$

Then, as with Lemma 5.4 in (Srinivas et al., 2010), $\sqrt{\sum_{t=1}^{T}\sigma_{t-1}^2(\boldsymbol{x}_t)} \leq \sqrt{C_1\gamma_T}$ with probability 1, where $C_1 = 2/\log(1+\sigma^{-2})$. Hence,

$$
\begin{aligned}
R_2 &\leq \mathbb{E}\left[\sqrt{\sum_{t=1}^{T}\zeta_t}\right]\sqrt{C_1\gamma_T} \\
&\leq \sqrt{\mathbb{E}\left[\sum_{t=1}^{T}\zeta_t\right]}\sqrt{C_1\gamma_T} \qquad (\because \text{Jensen's inequality}) \\
&= \sqrt{\sum_{t=1}^{T}\mathbb{E}\left[\zeta_t\right]}\sqrt{C_1\gamma_T}.
\end{aligned}
\qquad (9)
$$

Consequently, combining Eqs. (8) and (9), we can obtain the theorem for (i).

Then, we consider the case (ii). As with the proof of Theorem B.1, purely for the sake of analysis, we used a set of discretization $\mathcal{X}_t \subset \mathcal{X}$ for $t \geq 1$. For any $t \geq 1$, let $\mathcal{X}_t \subset \mathcal{X}$ be a finite set with each dimension equally divided into $\tau_t = bdrt^2\left(\sqrt{\log(ad)} + \sqrt{\pi}/2\right)$. Thus, $|\mathcal{X}_t| = \tau_t^d$. In addition, we define $[\boldsymbol{x}]_t$ as the nearest points in $\mathcal{X}_t$ of $\boldsymbol{x} \in \mathcal{X}$.

Then, we decompose BCR as follows:

$$
\begin{aligned}
\text{BCR}_T &= \sum_{t=1}^{T} \mathbb{E}\left[f(\boldsymbol{x}^*) - f([\boldsymbol{x}^*]_t) + f([\boldsymbol{x}^*]_t) - V_t([\boldsymbol{x}^*]_t) + V_t([\boldsymbol{x}^*]_t) - V_t(\boldsymbol{x}_t) + V_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right] \\
&\leq \sum_{t=1}^{T} \mathbb{E}\left[f(\boldsymbol{x}^*) - f([\boldsymbol{x}^*]_t) + f([\boldsymbol{x}^*]_t) - V_t([\boldsymbol{x}^*]_t) + V_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right],
\end{aligned}
$$

17

Table 2. PDFs, MGFs, and expectations for distributions of ζ, where Γ is the Gamma function, and ϕ and Φ are PDF and cumulative distribution function of standard normal distribution, respectively.

	PDF	MGF: $M(-1/2)$	Expectation
Gamma: (k, θ)	$\frac{\zeta^{k-1}e^{-\zeta/\theta}}{\Gamma(k)\theta^k}$	$(1+\theta/2)^{-k}$	$k\theta$
Two-parameter exponential: (s, λ)	$\lambda e^{-\lambda(\zeta-s)}$ for $\zeta \geq s \geq 0$	$\frac{\lambda}{\lambda+1/2}e^{-s/2}$	$s+1/\lambda$
Truncated normal: $(m, s^2 = 1)$ with truncation $[m-s, m+s]$	$\frac{\phi((\zeta-m)/s)}{\Phi(1)-\Phi(-1)}$ for $m-s \leq \zeta \leq m+s$	$e^{-m/2+1/8}\frac{\Phi(3/2)-\Phi(-1/2)}{\Phi(1)-\Phi(-1)}$	m

where $\mathbb{E}\big[V_t([\boldsymbol{x}^*]_t) - V_t(\boldsymbol{x}_t)\big] \leq 0$ since $\boldsymbol{x}_t = \arg\max_{\boldsymbol{x}\in\mathcal{X}} V_t(\boldsymbol{x})$. Furthermore, we obtain the following:

$$\sum_{t=1}^{T} \mathbb{E}\left[f(\boldsymbol{x}^*) - f([\boldsymbol{x}^*]_t)\right] \leq \sum_{t=1}^{T} \mathbb{E}\left[\sup_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) - f([\boldsymbol{x}]_t)\right]$$

$$\leq \sum_{t=1}^{T} \frac{1}{t^2} \qquad\qquad (\because \text{Lemma } H.2)$$

$$\leq \frac{\pi^2}{6} \qquad\qquad \left(\because \sum_{t=1}^{\infty} \frac{1}{t^2} = \frac{\pi^2}{6}\right)$$

The remaining terms $\sum_{t=1}^{T} \mathbb{E}\big[f([\boldsymbol{x}^*]_t) - V_t([\boldsymbol{x}^*]_t) + V_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\big]$ can be bounded as with the case (i). That is,

$$\sum_{t=1}^{T} \mathbb{E}\big[f([\boldsymbol{x}^*]_t) - V_t([\boldsymbol{x}^*]_t) + V_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\big] \leq \frac{1}{\sqrt{2\pi}}\sum_{t=1}^{T} |\mathcal{X}_t|M_t(-1/2) + \sqrt{C_1\sum_{t=1}^{T} \mathbb{E}[\zeta_t]\gamma_T}.$$

By substituting $|\mathcal{X}_t| = (\tau_t)^d$, we conclude the proof. $\qquad\qquad\square$

D.2. Example of Distributions for Sub-linear BCR Bounds

Next, we show examples of the distributions that can achieve sub-linear regret bound using Theorem D.1. Here, we discuss the case (ii) for infinite $\mathcal{X}$. The derivation for case (i) can be obtained easily by replacing $|\mathcal{X}_t| = (\tau_t)^d$ with $|\mathcal{X}|$.

Theorem D.1 shows that BCR can be bounded by the expectation and MGF. Roughly speaking, if we can control MGF as $M_t(-1/2) = O(1/(t^2|\mathcal{X}_t|))$ and the expectation $\mathbb{E}[\zeta_t] = O(\log t)$ by scheduling the parameter of the distribution of ζ_t, we can obtain sub-linear regret bounds, which is similar to Theorem B.1. This is because we can bound the first term by $\frac{1}{\sqrt{2\pi}}\sum_{t=1}^{T} 1/t^2 = \pi^2/(6\sqrt{2\pi})$ if the following inequality holds:

$$|\mathcal{X}_t|M_t(-1/2) \leq \frac{1}{t^2}.$$

This inequality can be rewritten as,

$$M_t(-1/2) \leq \frac{1}{|\mathcal{X}_t|t^2}.$$

Furthermore, the expectation $\mathbb{E}[\zeta_t] = O(\log t)$ is required to bound the second term.

The above requirements can be satisfied not only by the Gamma distribution proposed in (Berk et al., 2020) but also by many other distributions, such as two-parameter exponential and truncated normal distributions. Their PDFs, MGFs, and expectations are listed in Table 2. Note that although we describe only those three distributions, we can use other distributions, such as chi-squared distribution and truncated normal with other truncation, to obtain sub-linear regret bound from Theorem D.1. Each scheduling of the parameter and resulting regret bound are shown as follows:

Gamma Distribution: If we use ζ_t following Gamma distribution with k_t and θ, we require the following:

$$(1+\theta/2)^{-k_t} \leq \frac{1}{|\mathcal{X}_t|t^2}.$$

Therefore, we obtain the following:

$$k_t \geq \frac{\log(|\mathcal{X}_t|t^2)}{\log(1 + \theta/2)}.$$

The above derivation is mostly the same as the proof of Theorem 1 in (Berk et al., 2020). However, Berk et al. (2020) bounded as $\sqrt{\sum_{t=1}^{T} \mathbb{E}[\zeta_t]} \leq \sqrt{T\mathbb{E}[\max_{t \leq T} \zeta_t]}$. This operation is correct, but the evaluation of $\mathbb{E}[\max_{t \leq T} \zeta_t]$ is slightly complicated. To bound this term, Berk et al. (2020) used nearly equal, which is forbidden. On the other hand, we know the expectation $\mathbb{E}[\zeta_t] = k_t\theta$ analytically. Let

$$\begin{aligned} k_t &= \frac{\log(|\mathcal{X}_t|t^2)}{\log(1 + \theta/2)} \\ &= \frac{d\log\big(bdrt^2\big(\sqrt{\log(ad)} + \sqrt{\pi}/2\big)\big) + 2\log(t)}{\log(1 + \theta/2)}. \end{aligned}$$

Consequently, since $\mathbb{E}[\zeta_t] = k_t\theta$ is monotone increasing, we obtain the following bound:

$$\mathrm{BCR}_T \leq \frac{\pi^2}{6} + \frac{\pi^2}{6\sqrt{2\pi}} + \sqrt{C_1\gamma_t T\mathbb{E}[\zeta_T]} = \frac{\pi^2}{6} + \frac{\pi^2}{6\sqrt{2\pi}} + \sqrt{C_1\gamma_t T\theta k_T}.$$

Two-Parameter Exponential Distribution: If we use ζ_t following two-parameter exponential distribution with s_t and λ, we requires the following:

$$\frac{\lambda}{\lambda + 1/2}e^{-s_t/2} \leq e^{-s_t/2} \leq \frac{1}{|\mathcal{X}_t|t^2},$$

where we bound $\frac{\lambda}{\lambda+1/2} < 1$ so that the condition for λ to maintain $s_t \geq 0$ will be eliminated. Therefore, we obtain the following:

$$s_t \geq 2\log\left(|\mathcal{X}_t|t^2\right).$$

Hence, by setting $s_t = 2\log\left(|\mathcal{X}_t|t^2\right) = 2d\log\big(bdrt^2\big(\sqrt{\log(ad)} + \sqrt{\pi}/2\big)\big) + 4\log(t)$, since $\mathbb{E}[\zeta_t] = s_t + 1/\lambda$ is monotone increasing, we obtain the following bound:

$$\mathrm{BCR}_T \leq \frac{\pi^2}{6} + \frac{\pi^2}{6\sqrt{2\pi}} + \sqrt{C_1\gamma_t T\mathbb{E}[\zeta_T]} = \frac{\pi^2}{6} + \frac{\pi^2}{6\sqrt{2\pi}} + \sqrt{C_1\gamma_t T(s_T + 1/\lambda)}.$$

Truncated Normal Distribution: For brevity, we use a truncated normal distribution with $s^2 = 1$ and truncation $[m - s, m + s]$. We schedule m_t so that the following inequality holds:

$$e^{-m/2+1/8}\frac{\Phi(3/2) - \Phi(-1/2)}{\Phi(1) - \Phi(-1)} < e^{-m_t/2+1/8} < \frac{1}{|\mathcal{X}_t|t^2},$$

where we bound $\frac{\Phi(1)-\Phi(0)}{\Phi(1/2)-\Phi(-1/2)} < 1$ for simplicity. Then, we obtain the following condition:

$$m_t \geq 2\log\left(|\mathcal{X}_t|t^2\right) + 1/4.$$

Furthermore, the lower bound $m_t - s > 0$ should be satisfied. Hence, by setting $m_t = 2\log\left(|\mathcal{X}_t|t^2\right) + 1 = 2d\log\big(bdrt^2\big(\sqrt{\log(ad)} + \sqrt{\pi}/2\big)\big) + 4\log(t) + 1$, since $\mathbb{E}[\zeta_t] = m_t$ is monotone increasing, we obtain the following bound:

$$\mathrm{BCR}_T \leq \frac{\pi^2}{6} + \frac{\pi^2}{6\sqrt{2\pi}} + \sqrt{C_1\gamma_t T\mathbb{E}[\zeta_T]} = \frac{\pi^2}{6} + \frac{\pi^2}{6\sqrt{2\pi}} + \sqrt{C_1\gamma_t Tm_T}.$$

E. Proofs of Lemmas 4.1 and 4.2

E.1. Proof of Lemma 4.1

The main difference between Lemma 4.1 and (Lemma 5.1 in Srinivas et al., 2010) is that Lemma 4.1 does not consider the union bound for all $t \geq 1$. That is, (Lemma 5.1 in Srinivas et al., 2010) bounds the following probability:

$$\Pr\left(f(\boldsymbol{x}) \leq \mu_{t-1}(\boldsymbol{x}) + \beta_t^{1/2}\sigma_{t-1}(\boldsymbol{x}), \forall \boldsymbol{x} \in \mathcal{X}, \forall t \geq 1\right).$$

On the other hand, our Lemma 4.1 bounds the following probability with one fixed t:

$$\Pr_t\left(f(\boldsymbol{x}) \leq \mu_{t-1}(\boldsymbol{x}) + \beta_\delta^{1/2}\sigma_{t-1}(\boldsymbol{x}), \forall \boldsymbol{x} \in \mathcal{X}\right),$$

where $\beta_\delta^{1/2}$ does not depend on t as described below.

Assume the same condition as in Lemma 4.1, and thus, $\beta_\delta = 2\log(|\mathcal{X}|/(2\delta))$. Then, for all $\boldsymbol{x} \in \mathcal{X}$ and $\mathcal{D}_{t-1}$, we see that

$$\Pr_t\left(f(\boldsymbol{x}) > \mu_{t-1}(\boldsymbol{x}) + \beta_\delta^{1/2}\sigma_{t-1}(\boldsymbol{x})\right) \leq \frac{\delta}{|\mathcal{X}|},$$

where we use an elementary result of Gaussian distribution shown in Lemma H.3 in Appendix H. Then, we can obtain the following bound:

$$\Pr_t\left(f(\boldsymbol{x}) > \mu_{t-1}(\boldsymbol{x}) + \beta_\delta^{1/2}\sigma_{t-1}(\boldsymbol{x}), \exists \boldsymbol{x} \in \mathcal{X}\right) \leq \sum_{\boldsymbol{x} \in \mathcal{X}} \Pr_t\left(f(\boldsymbol{x}) > \mu_{t-1}(\boldsymbol{x}) + \beta_\delta^{1/2}\sigma_{t-1}(\boldsymbol{x})\right)$$
$$\leq \delta.$$

Therefore,

$$\Pr_t\left(f(\boldsymbol{x}) \leq \mu_{t-1}(\boldsymbol{x}) + \beta_\delta^{1/2}\sigma_{t-1}(\boldsymbol{x}), \forall \boldsymbol{x} \in \mathcal{X}\right) = 1 - \Pr_t\left(f(\boldsymbol{x}) > \mu_{t-1}(\boldsymbol{x}) + \beta_\delta^{1/2}\sigma_{t-1}(\boldsymbol{x}), \exists \boldsymbol{x} \in \mathcal{X}\right) \geq 1 - \delta,$$

which concludes the proof.

E.2. Detailed Proof of Lemma 4.2

This section provides detailed proof of the following Lemma 4.2:

Lemma 4.2. *Let $f \sim \mathcal{GP}(0, k)$, where k is a stationary kernel and $k(\boldsymbol{x}, \boldsymbol{x}) = 1$, and $\mathcal{X}$ be finite. Assume that ζ follows a two-parameter exponential distribution with $s = 2\log(|\mathcal{X}|/2)$ and $\lambda = 1/2$. Then, the following inequality holds:*

$$\mathbb{E}[f(\boldsymbol{x}^*)] \leq \mathbb{E}\left[\max_{\boldsymbol{x} \in \mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \zeta_t^{1/2}\sigma_{t-1}(\boldsymbol{x})\right],$$

for all $t \geq 1$.

Proof. It suffices to show that, for an arbitrary $\mathcal{D}_{t-1}$,

$$\mathbb{E}_t[f(\boldsymbol{x}^*)|\mathcal{D}_{t-1}] \leq \mathbb{E}_t\left[\max_{\boldsymbol{x} \in \mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \zeta_t^{1/2}\sigma_{t-1}(\boldsymbol{x})\right],$$

since the claim of Lemma 4.2 can be rewritten as follows:

$$\mathbb{E}_{\mathcal{D}_{t-1}}[\mathbb{E}_t[f(\boldsymbol{x}^*)]]$$
$$\leq \mathbb{E}_{\mathcal{D}_{t-1}}\left[\mathbb{E}_t\left[\max_{\boldsymbol{x} \in \mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \zeta^{1/2}\sigma_{t-1}(\boldsymbol{x})\right]\right].$$

Then, fix the dataset $\mathcal{D}_{t-1}$ and $\delta \in (0, 1)$. From Lemma 4.1, we can see that

$$\Pr_t\left(f(\boldsymbol{x}^*) \leq \max_{\boldsymbol{x} \in \mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \beta_\delta^{1/2}\sigma_{t-1}(\boldsymbol{x})\right) \geq 1 - \delta,$$

20

where $\beta_\delta = 2\log(|\mathcal{X}|/(2\delta))$. Note that the probability is taken only with $f(\boldsymbol{x}^*)$ since $\mathcal{D}_{t-1}$ is fixed. Using the cumulative distribution function $F_t(\cdot) := \Pr_t(f(\boldsymbol{x}^*) \leq \cdot)$ and its inverse function F_t^{-1}, we can rewrite

$$F_t\left(\max_{\boldsymbol{x}\in\mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \beta_\delta^{1/2}\sigma_{t-1}(\boldsymbol{x})\right) \geq 1 - \delta$$

$$\iff \max_{\boldsymbol{x}\in\mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \beta_\delta^{1/2}\sigma_{t-1}(\boldsymbol{x}) \geq F_t^{-1}(1-\delta),$$

since F_t^{-1} is a monotone increasing function. Since $\delta \in (0,1)$ is arbitrary, we can see that

$$\Pr\left(\max_{\boldsymbol{x}\in\mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \beta_U^{1/2}\sigma_{t-1}(\boldsymbol{x}) \geq F_t^{-1}(1-U)\right) = 1,$$

where $U \sim \mathrm{Uni}(0,1)$ and probability is taken by the randomness of U. Then, because of the basic property of the expectation $\mathbb{E}[X] \leq \mathbb{E}[Y]$ if $\Pr(X \leq Y) = 1$, we can obtain,

$$\mathbb{E}_U\left[\max_{\boldsymbol{x}\in\mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \beta_U^{1/2}\sigma_{t-1}(\boldsymbol{x})\right] \geq \mathbb{E}_U\left[F_t^{-1}(1-U)\right]$$

$$= \mathbb{E}_U\left[F_t^{-1}(U)\right],$$

where we use the fact that $1 - U$ also follows $\mathrm{Uni}(0,1)$. Since $F_t^{-1}(U)$ and $f(\boldsymbol{x}^*)$ are identically distributed as with the inverse transform sampling, we obtain

$$\mathbb{E}_U\left[\max_{\boldsymbol{x}\in\mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \beta_U^{1/2}\sigma_{t-1}(\boldsymbol{x})\right] \geq \mathbb{E}_t\left[f(\boldsymbol{x}^*)\right].$$

Then, β_U can be decomposed as follows:

$$\beta_U = 2\log(|\mathcal{X}|/2) - 2\log(U).$$

From the inverse transform sampling, $-2\log(U) \sim \mathrm{Exp}(\lambda = 1/2)$. Hence, β_U follows a two-parameter exponential distribution with $s = 2\log(|\mathcal{X}|/2)$ and $\lambda = 1/2$. Therefore, we obtain the following:

$$\mathbb{E}_t\left[\max_{\boldsymbol{x}\in\mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \zeta^{1/2}\sigma_{t-1}(\boldsymbol{x})\right] \geq \mathbb{E}_t\left[f(\boldsymbol{x}^*)\right].$$

$\square$

F. Proof of Theorem 4.2

Using Lemma 4.2, we can obtain the proof of the following theorem for the finite input domain:

Theorem 4.2. *Let $f \sim \mathcal{GP}(0, k)$, where k is a stationary kernel and $k(\boldsymbol{x}, \boldsymbol{x}) = 1$, and $\mathcal{X}$ be finite. Assume that ζ_t follows a two-parameter exponential distribution with $s = 2\log(|\mathcal{X}|/2)$ and $\lambda = 1/2$ for any $t \geq 1$. Then, by running IRGP-UCB with ζ_t, BCR can be bounded as follows:*

$$\mathrm{BCR}_T \leq \sqrt{C_1 C_2 T \gamma_T},$$

where $C_1 := 2/\log(1 + \sigma^{-2})$ and $C_2 := 2 + s$.

Proof. From Lemma 4.2, we obtain

$$\mathrm{BCR}_T = \sum_{t=1}^T \mathbb{E}[f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t)]$$

$$= \sum_{t=1}^T \mathbb{E}[f(\boldsymbol{x}^*) - V_t(\boldsymbol{x}_t) + V_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)]$$

$$\leq \sum_{t=1}^T \mathbb{E}[V_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)].$$

Then, we can easily derive the bound as follows:

$$
\begin{aligned}
\mathrm{BCR}_T &\leq \sum_{t=1}^{T} \mathbb{E}[V_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)] \\
&= \sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}_{t-1}, \zeta_t}[\mathbb{E}[V_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t) | \mathcal{D}_{t-1}, \zeta_t]] \\
&= \sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}_{t-1}}[\mathbb{E}_{\zeta_t}[V_t(\boldsymbol{x}_t)] - \mu_{t-1}(\boldsymbol{x}_t)] \\
&= \sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}_{t-1}, \zeta_t}[\zeta_t^{1/2} \sigma_{t-1}(\boldsymbol{x}_t)] \\
&= \mathbb{E}\left[\sum_{t=1}^{T} \zeta_t^{1/2} \sigma_{t-1}(\boldsymbol{x}_t)\right] \\
&\leq \mathbb{E}\left[\sqrt{\sum_{t=1}^{T} \zeta_t \sum_{t=1}^{T} \sigma_{t-1}^2(\boldsymbol{x}_t)}\right] && (\because \text{Cauchy–Schwarz inequality}) \\
&\leq \mathbb{E}\left[\sqrt{\sum_{t=1}^{T} \zeta_t}\right] \sqrt{C_1 \gamma_T} && (\because \text{Lemma 5.4 in (Srinivas et al., 2010)}) \\
&\leq \sqrt{\sum_{t=1}^{T} \mathbb{E}[\zeta_t]} \sqrt{C_1 \gamma_T} && (\because \text{Jensen's inequality}) \\
&= \sqrt{C_1 C_2 T \gamma_T}, && (\because \text{Definition of } \zeta_t)
\end{aligned}
$$

where $C_2 = \mathbb{E}[\zeta_t] = 2 + 2\log(|\mathcal{X}|/2)$. This concludes the proof. $\qquad \square$

Corollary F.1. *Assume the same condition as in Theorem 4.2. Fix a required accuracy $\eta > 0$. Then, $\mathrm{BSR}_T \leq \eta$ by at most T function evaluations, where T is the smallest positive integer satisfying the following inequality:*

$$
\sqrt{\frac{C_1 C_2 \gamma_T}{T}} \leq \eta,
$$

where $C_1 := 2/\log(1 + \sigma^{-2})$ and $C_2 := 2 + 2\log(|\mathcal{X}|/2)$.

Proof. This BCR bound in Theorem 4.2 immediately suggests that the following:

$$
\mathrm{BSR}_T \leq \frac{\mathrm{BCR}_T}{T} \leq \sqrt{\frac{C_1 C_2 \gamma_T}{T}}.
$$

Therefore, if $\sqrt{C_1 C_2 \gamma_T / T} < \eta$, then $\mathrm{BSR}_T < \eta$. $\qquad \square$

G. Proof of Theorem 4.3

Theorem 4.3. *Let $f \sim \mathcal{GP}(0, k)$, where k is a stationary kernel, $k(\boldsymbol{x}, \boldsymbol{x}) = 1$, and Assumption 2.1 holds. Assume that ζ_t follows a two-parameter exponential distribution with $s_t = 2d\log(bdrt^2(\sqrt{\log(ad)} + \sqrt{\pi}/2)) - 2\log 2$ and $\lambda = 1/2$ for any $t \geq 1$. Then, by running IRGP-UCB, BCR can be bounded as follows:*

$$
\mathrm{BCR}_T \leq \frac{\pi^2}{6} + \sqrt{C_1 T \gamma_T (2 + s_T)},
$$

where $C_1 := 2/\log(1 + \sigma^{-2})$.

Proof. Purely for the sake of analysis, we used a set of discretization $\mathcal{X}_t \subset \mathcal{X}$ for $t \geq 1$. For any $t \geq 1$, let $\mathcal{X}_t \subset \mathcal{X}$ be a finite set with each dimension equally divided into $\tau_t = bdrt^2\left(\sqrt{\log(ad)} + \sqrt{\pi}/2\right)$. Thus, $|\mathcal{X}_t| = \tau_t^d$. In addition, we define $[\boldsymbol{x}]_t$ as the nearest points in $\mathcal{X}_t$ of $\boldsymbol{x} \in \mathcal{X}$.

Then, we decompose BCR as follows:

$$\text{BCR}_T = \sum_{t=1}^{T} \mathbb{E}\left[f(\boldsymbol{x}^*) - f([\boldsymbol{x}^*]_t) + f([\boldsymbol{x}^*]_t) - \max_{\boldsymbol{x} \in \mathcal{X}_t} V_t(\boldsymbol{x}) + \max_{\boldsymbol{x} \in \mathcal{X}_t} V_t(\boldsymbol{x}) - V_t(\boldsymbol{x}_t) + V_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right].$$

Obviously, $\mathbb{E}\left[\max_{\boldsymbol{x} \in \mathcal{X}_t} V_t(\boldsymbol{x}) - V_t(\boldsymbol{x}_t)\right] \leq 0$ from the definition of $\boldsymbol{x}_t$. Furthermore, we observe the following:

$$\mathbb{E}\left[f([\boldsymbol{x}^*]_t) - \max_{\boldsymbol{x} \in \mathcal{X}_t} V_t(\boldsymbol{x})\right] \leq \mathbb{E}\left[\max_{\boldsymbol{x} \in \mathcal{X}_t} f(\boldsymbol{x}) - \max_{\boldsymbol{x} \in \mathcal{X}_t} V_t(\boldsymbol{x})\right],$$

which can be bounded above by 0 by setting $s_t = 2\log(|\mathcal{X}_t|/2)$ and $\lambda = 1/2$ from Lemma 4.2. Therefore, we obtain the following:

$$\text{BCR}_T \leq \sum_{t=1}^{T} \mathbb{E}\left[f(\boldsymbol{x}^*) - f([\boldsymbol{x}^*]_t) + V_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right]$$

$$= \sum_{t=1}^{T} \mathbb{E}\left[f(\boldsymbol{x}^*) - f([\boldsymbol{x}^*]_t)\right] + \sum_{t=1}^{T} \mathbb{E}\left[V_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right]. \tag{10}$$

First, we consider the first term $\sum_{t=1}^{T} \mathbb{E}\left[f(\boldsymbol{x}^*) - f([\boldsymbol{x}^*]_t)\right]$. From Lemma H.2, we can obtain the following:

$$\sum_{t=1}^{T} \mathbb{E}\left[f(\boldsymbol{x}^*) - f([\boldsymbol{x}^*]_t)\right] \leq \sum_{t=1}^{T} \mathbb{E}\left[\sup_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x}) - f([\boldsymbol{x}]_t)\right]$$

$$\leq \sum_{t=1}^{T} \frac{1}{t^2} \qquad\qquad (\because \text{ Lemma H.2})$$

$$\leq \frac{\pi^2}{6}, \qquad\qquad \left(\because \sum_{t=1}^{\infty} \frac{1}{t^2} = \frac{\pi^2}{6}\right) \tag{11}$$

where u_t in Lemma H.2 corresponds to t^2.

The second term is bounded as follows:

$$\sum_{t=1}^{T} \mathbb{E}\left[V_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right] \leq \sqrt{\sum_{t=1}^{T} \mathbb{E}[\zeta_t] C_1 \gamma_T}, \qquad (\because \text{See the proof of Theorem 4.2}) \tag{12}$$

where

$$\mathbb{E}[\zeta_t] = 2 + 2\log(|\mathcal{X}_t|/2)$$

$$= 2 + 2d\log(bdrt^2\left(\sqrt{\log(ad)} + \sqrt{\pi}/2\right)) - 2\log 2,$$

which is monotone increasing. Therefore, we obtain the following:

$$\sum_{t=1}^{T} \mathbb{E}\left[V_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right] \leq \sqrt{C_1 T \gamma_T \mathbb{E}[\zeta_T]},$$

where $\mathbb{E}[\zeta_T] = 2 + s_T = 2 + 2d\log(bdrT^2\left(\sqrt{\log(ad)} + \sqrt{\pi}/2\right)) - 2\log 2$. Consequently, combining Eqs. (10), (11), and (12) concludes the proof. □

Corollary G.1. *Assume the same condition as in Theorem 4.3. Fix a required accuracy $\eta > 0$. Then, by running IRGP-UCB with $s_\eta = 2d \log(2bdr(\sqrt{\log(ad)} + \sqrt{\pi}/2)/\eta) - 2\log 2$ and $\lambda = 1/2$, $\mathrm{BSR}_T \leq \eta$ by at most T function evaluations, where T is the smallest positive integer satisfying the following inequality:*

$$\sqrt{\frac{C_1(2 + s_\eta)\gamma_T}{T}} \leq \frac{\eta}{2},$$

where $C_1 \coloneqq 2/\log(1 + \sigma^{-2})$.

Proof. We obtain the upper bound of BSR via the upper bound of BCR as follows:

$$\mathrm{BSR}_T \leq \frac{\mathrm{BCR}_T}{T}.$$

Let $\tau_t = 2bdr(\sqrt{\log(ad)} + \sqrt{\pi}/2)/\eta$. Until the derivation of Eq. (10), the same derivation as in Theorem 4.3 can be applied. That is,

$$\mathrm{BCR}_T \leq \sum_{t=1}^{T} \mathbb{E}\left[f(\boldsymbol{x}^*) - f([\boldsymbol{x}^*]_t)\right] + \sum_{t=1}^{T} \mathbb{E}\left[V_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right].$$

Then, for the first term, we can obtain the following

$$\sum_{t=1}^{T} \mathbb{E}\left[f(\boldsymbol{x}^*) - f([\boldsymbol{x}^*]_t)\right] \leq \sum_{t=1}^{T} \frac{\eta}{2} \qquad (\because \text{ Lemma H.2 })$$

$$\leq \frac{\eta}{2}T.$$

For the second term,

$$\sum_{t=1}^{T} \mathbb{E}\left[V_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right] \leq \sqrt{C_1 \sum_{t=1}^{T} \mathbb{E}[\zeta_t]\gamma_T},$$

where, from the assumption, $\mathbb{E}[\zeta_t] = 2 + s_\eta$ for any $t \geq 1$. Thus, we can derive the bound of BSR:

$$\mathrm{BSR}_T \leq \frac{\eta}{2} + \sqrt{\frac{C_1(2 + s_\eta)\gamma_T}{T}}.$$

Hence, we obtain the smallest integer T that satisfies $\mathrm{BSR}_T \leq \eta$ by arranging the following inequality:

$$\frac{\eta}{2} + \sqrt{\frac{C_1(2 + s_\eta)\gamma_T}{T}} \leq \eta.$$

$\square$

H. Auxiliary Lemmas

For convenience, we here show the assumption again:

Assumption 2.1. *Let $\mathcal{X} \subset [0, r]^d$ be a compact and convex set, where $r > 0$. Assume that the kernel k satisfies the following condition on the derivatives of sample path f. There exists the constants $a, b > 0$ such that,*

$$\Pr\left(\sup_{\boldsymbol{x} \in \mathcal{X}} \left|\frac{\partial f}{\partial \boldsymbol{x}_j}\right| > L\right) \leq a \exp\left(-\left(\frac{L}{b}\right)^2\right), \text{ for } j \in [d],$$

where $[d] = \{1, \ldots, d\}$.

Then, from Assumption 2.1, we obtain several lemmas. First, we show the upper bound of the supremum of the partial derivatives. This result is tighter than Lemma 12 in (Kandasamy et al., 2018a). Note that, in the proof of Lemma 12 in (Kandasamy et al., 2018a), there is a typo that $\Pr(L \geq t)$ is bounded above by $a \exp\{t^2/b^2\}$, which should be $ad \exp\{-t^2/b^2\}$.

Lemma H.1. *Let* $f \sim \mathcal{GP}(0, k)$ *and Assumption 2.1 holds. Let the supremum of the partial derivatives* $L_{\max} := \sup_{j \in [d]} \sup_{\boldsymbol{x} \in \mathcal{X}} \left| \frac{\partial f}{\partial x_j} \right|$. *Then,* $\mathbb{E}[L_{\max}]$ *can be bounded above as follows:*

$$\mathbb{E}[L_{\max}] \leq b\left(\sqrt{\log(ad)} + \sqrt{\pi}/2 \right).$$

Proof. From the assumption, using union bound, we obtain the bound of the probability,

$$\Pr\left(L_{\max} > L \right) \leq \sum_{j=1}^{d} a \exp\left(-\left(\frac{L}{b} \right)^2 \right) = ad \exp\left(-\left(\frac{L}{b} \right)^2 \right). \tag{13}$$

Then, using expectation integral identity, we can bound $\mathbb{E}[L_{\max}]$ as follows:

$$
\begin{aligned}
\mathbb{E}[L_{\max}] &= \int_0^\infty \Pr\left(L_{\max} > L \right) \mathrm{d}L \\
&\leq \int_0^\infty \min\{1, ade^{-(L/b)^2}\} \mathrm{d}L && (\because \text{Eq. (13)}) \\
&= b\sqrt{\log(ad)} + \int_{b\sqrt{\log(ad)}}^\infty ade^{-(L/b)^2} \mathrm{d}L \\
&= b\sqrt{\log(ad)} + abd\sqrt{\pi} \int_{b\sqrt{\log(ad)}}^\infty \frac{1}{\sqrt{2\pi(b^2/2)}} e^{-(L/b)^2} \mathrm{d}L \\
&= b\sqrt{\log(ad)} + abd\sqrt{\pi} \left(1 - \Phi\left(\frac{b\sqrt{\log(ad)}}{b/\sqrt{2}} \right) \right) \\
&\leq b\sqrt{\log(ad)} + \frac{b\sqrt{\pi}}{2}, && (\because \text{Lemma H.3})
\end{aligned}
$$

where Φ is a cumulative distribution function of the standard normal distribution. $\qquad\square$

Next, we show the bound for the difference of discretized outputs:

Lemma H.2. *Let* $f \sim \mathcal{GP}(0, k)$ *and Assumption 2.1 holds. Let* $\mathcal{X}_t \subset \mathcal{X}$ *be a finite set with each dimension equally divided into* $\tau_t = bdru_t\left(\sqrt{\log(ad)} + \sqrt{\pi}/2 \right)$ *for any* $t \geq 1$. *Then, we can bound the expectation of differences,*

$$\sum_{t=1}^{T} \mathbb{E}\left[\sup_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x}) - f([\boldsymbol{x}]_t) \right] \leq \sum_{t=1}^{T} \frac{1}{u_t},$$

where $[\boldsymbol{x}]_t$ *is the nearest point in* $\mathcal{X}_t$ *of* $\boldsymbol{x} \in \mathcal{X}$.

Proof. From the construction of $\mathcal{X}_t$, we can obtain the upper bound of L1 distance between $\boldsymbol{x}$ and $[\boldsymbol{x}]_t$ as follows:

$$
\begin{aligned}
\sup_{\boldsymbol{x} \in \mathcal{X}} \|\boldsymbol{x} - [\boldsymbol{x}]_t\|_1 &\leq \frac{dr}{bdru_t\left(\sqrt{\log(ad)} + \sqrt{\pi}/2 \right)} \\
&= \frac{1}{bu_t\left(\sqrt{\log(ad)} + \sqrt{\pi}/2 \right)}. \tag{14}
\end{aligned}
$$

Note that this discretization does not depend on any randomness and is fixed beforehand.

Then, we obtain the following:

$$\sum_{t=1}^{T} \mathbb{E}\left[\sup_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) - f([\boldsymbol{x}]_t)\right] \leq \sum_{t=1}^{T} \mathbb{E}\left[L_{\max}\sup_{\boldsymbol{x}\in\mathcal{X}} \|\boldsymbol{x} - [\boldsymbol{x}]_t\|_1\right]$$

$$\leq \sum_{t=1}^{T} \mathbb{E}\left[L_{\max}\right] \frac{1}{bu_t\left(\sqrt{\log(ad)} + \sqrt{\pi}/2\right)} \qquad (\because \text{Eq. (14)})$$

$$\leq \sum_{t=1}^{T} \frac{1}{u_t}. \qquad (\because \text{Lemma H.1})$$

$\square$

We used the following useful lemma:

Lemma H.3 (in Lemma 5.2 of (Srinivas et al., 2010)). *For $c > 0$, the survival function of the standard normal distribution can be bounded above as follows:*

$$1 - \Phi(c) \leq \frac{1}{2}\exp(-c^2/2).$$

Proof. Let $R \sim \mathcal{N}(0, 1)$. Then, we obtain the following:

$$\Pr(R > c) = \int_c^\infty \frac{1}{\sqrt{2\pi}}\exp(-r^2/2)\mathrm{d}r$$

$$= \exp(-c^2/2)\int_c^\infty \frac{1}{\sqrt{2\pi}}\exp(-r^2/2 + c^2/2)\mathrm{d}r$$

$$= \exp(-c^2/2)\int_c^\infty \frac{1}{\sqrt{2\pi}}\exp(-(r-c)^2/2 - rc + c^2)\mathrm{d}r$$

$$= \exp(-c^2/2)\int_c^\infty \frac{1}{\sqrt{2\pi}}\exp(-(r-c)^2/2 - c(r-c))\mathrm{d}r$$

$$\leq \exp(-c^2/2)\int_c^\infty \frac{1}{\sqrt{2\pi}}\exp(-(r-c)^2/2)\mathrm{d}r \qquad \left(\because e^{-c(r-c)} \leq 1 \text{ since } r \geq c > 0\right)$$

$$= \frac{1}{2}\exp(-c^2/2).$$

$\square$