Posterior Sampling-Based Bayesian Optimization with Tighter Bayesian Regret Bounds

Shion Takeno [1 2] **Yu Inatsu** [3] **Masayuki Karasuyama** [3] **Ichiro Takeuchi** [1 2]

Abstract

Among various acquisition functions (AFs) in Bayesian optimization (BO), Gaussian process upper confidence bound (GP-UCB) and Thompson sampling (TS) are well-known options with established theoretical properties regarding Bayesian cumulative regret (BCR). Recently, it has been shown that a randomized variant of GP-UCB achieves a tighter BCR bound compared with GP-UCB, which we call the tighter BCR bound for brevity. Inspired by this study, this paper first shows that TS achieves the tighter BCR bound. On the other hand, GP-UCB and TS often practically suffer from manual hyperparameter tuning and over-exploration issues, respectively. Therefore, we analyze yet another AF called a probability of improvement from the maximum of a sample path (PIMS). We show that PIMS achieves the tighter BCR bound and avoids the hyperparameter tuning, unlike GP-UCB. Furthermore, we demonstrate a wide range of experiments, focusing on the effectiveness of PIMS that mitigates the practical issues of GP-UCB and TS.

1. Introduction

Bayesian optimization (BO) (Mockus et al., 1978) has become a popular tool for an expensive-to-evaluate black-box function optimization problem. BO aims to optimize with fewer function evaluations by adaptively evaluating the objective function based on the Bayesian model and *acquisition functions* (AF). Due to its effectiveness, BO has been applied to a wide range of problems, including drug discovery (Korovina et al., 2020), AutoML (Klein et al., 2017; Falkner et al., 2018), and materials informatics (Ueno et al., 2016).

Gaussian process upper confidence bound (GP-UCB) (Kushner, 1962; Srinivas et al., 2010) is a seminal work of the regret analysis in BO literature. GP-UCB achieves sublinear Bayesian cumulative regret (BCR) bounds by sequentially evaluating the maximizer of the UCB, whose width is controlled by the confidence parameter β_t. However, the theoretical choice of the confidence parameter can be huge. Therefore, GP-UCB practically requires manually tuning the confidence parameter, which strongly affects the optimization performance. Furthermore, GP-UCB with a manually tuned confidence parameter is no longer theoretically guaranteed.

Another well-known theoretically guaranteed BO method is Thompson sampling (TS) (Thompson, 1933), whose same BCR bound as that of GP-UCB has been shown (Russo & Van Roy, 2014). TS is a randomized algorithm that sequentially evaluates the optimal solution of a sample path from a posterior. Therefore, TS does not have hyperparameters like the confidence parameter for the AF, except for those for GPs. Hence, TS does not require any manual hyperparameter tuning regarding AF. However, it is frequently discussed that TS suffers from *over-exploration*, particularly in higher-dimensional optimization problem (e.g., Shahriari et al., 2016; Takeno et al., 2023).

Recently, Takeno et al. (2023) showed improved randomized GP-UCB (IRGP-UCB) achieves tighter BCR bounds for a finite input domain compared with TS and GP-UCB. In particular, it is shown that there is no need to increase the confidence parameter in proportion to the iterations. However, the theoretical confidence parameter still depends on the cardinality of the finite input domain. Since the cardinality of the finite input domain can be huge, manual tuning of the confidence parameter can be required even in IRGP-UCB.

This study provides the same BCR bound of TS as IRGP-UCB, inspired by the analysis of IRGP-UCB, in which randomness plays a key role. Furthermore, against the practical issues of GP-UCB-based methods and TS, we analyze a *probability of improvement from a maximum of sample-path* (PIMS). PIMS, as the name suggests, is a randomized algo-

[1]Department of Engineering, Nagoya University, Aichi, Japan [2]RIKEN AIP, Tokyo, Japan [3]Department of Computer Science, Nagoya Institute of Technology, Aichi, Japan. Correspondence to: Shion Takeno <takeno.shion.m6@f.mail.nagoya-u.ac.jp>.

Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024. Copyright 2024 by the author(s).

	GP-UCB	IRGP-UCB	TS	PIMS
BCR for $\lvert\mathcal{X}\rvert < \infty$	$O(\sqrt{T\gamma_T \log(\lvert\mathcal{X}\rvert T)})$	$O(\sqrt{T\gamma_T \log \lvert\mathcal{X}\rvert})$	* $O(\sqrt{T\gamma_T \log \lvert\mathcal{X}\rvert})$	* $O(\sqrt{T\gamma_T \log \lvert\mathcal{X}\rvert})$
BCR for $\mathcal{X} \subset [0, r]^d$	$O(\sqrt{T\gamma_T \log T})$	$O(\sqrt{T\gamma_T \log T})$	* $O(\sqrt{T\gamma_T \log T})$	* $O(\sqrt{T\gamma_T \log T})$

Table 1. Summary of BCR bounds. The first and second rows show the BCR bounds for the finite and infinite input domains, respectively, where γ_T is the maximum information gain(Srinivas et al., 2010), $\mathcal{X}$ is the input domain, $d > 0$ is the input dimension, and $r > 0$ is a constant. The BCR bounds of GP-UCB and IRGP-UCB are shown in Theorem B.1 and Theorems 4.2 and 4.3 in Takeno et al. (2023), respectively. Stars mean our results.

rithm based on the PI from the maximum of a sample path generated from the posterior. We prove that PIMS achieves the same BCR bound as IRGP-UCB and TS. In addition, we empirically show that PIMS alleviates the over-exploration problem of TS, although PIMS inherits the benefits of TS having no hyperparameter of AF.

One technical challenge to showing the tighter BCR bounds of TS and PIMS is handling their nature of posterior sampling-based randomization. In the regret analyses of TS and PIMS, we consider the randomized confidence parameters, which are the same as IRGP-UCB, as a formality. However, given the randomized confidence parameters and training dataset, although the next evaluation point $\boldsymbol{x}_t$ in IRGP-UCB becomes constant, the next evaluation points $\boldsymbol{x}_t$ in TS and PIMS are still random variables. Therefore, we will show how to deal with these additional random quantities in the regret analyses.

One of the benefits of TS and PIMS is being hyperparameter-free. Most existing analyses (e.g., Srinivas et al., 2010; Takeno et al., 2023) have used discretization-based proofs in the BCR analysis for the continuous input domain. Note that their discretization is only for theoretical proofs, and the algorithms can still select the next point from the original continuous space. For TS and PIMS, to directly use a similar theoretical proof, the discretization is required not only in the proof but also in their algorithms (i.e., optimizing the sample path over finite discretized points). However, the cardinality of the discretization must be tuned manually (e.g., through the number of grid points), which means that the AFs are no longer hyperparameter-free. Therefore, we present an additional analysis for the BCR bound of TS and PIMS that does not require the discretization procedure in the algorithm; nevertheless, the proof is still based on the discretization.

Our contributions are summarized as follows:

1. We show the sub-linear BCR bounds of TS. For the finite input domain, the BCR bound is tightened compared with the known result (Russo & Van Roy, 2014). For the infinite input domain, we show the proof without the discretization in the algorithm.

2. We show the sub-linear BCR bounds of a BO method called PIMS, which is the PI from the maximum of the sample path from the posterior. PIMS is hyperparameter-free and easy to implement, as with TS. For the finite input domain, PIMS achieves the tighter BCR bound as with TS and IRGP-UCB. For the infinite input domain, we analyze PIMS without the discretization in the algorithm.

Theoretical results are summarized in Table 1. Finally, we show broad experiments, particularly focusing on the practical effectiveness of PIMS compared with TS and GP-UCB-based methods.

2. Background

2.1. Problem Statement

We consider an optimization of an unknown expensive-to-evaluate objective function f, which is formalized as $\boldsymbol{x}^* = \arg\max_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x})$ using an input domain $\mathcal{X} \subset \mathbb{R}^d$ and an input dimension d. BO aims to optimize with fewer function evaluations by adaptively observing the function value. Thus, at each iteration t, we sequentially query $\boldsymbol{x}_t$ based on the Bayesian model and AF. We assume that the observation $y_t = f(\boldsymbol{x}_t) + \epsilon_t$ is contaminated by a Gaussian noise $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$ with a positive variance $\sigma^2 > 0$.

We assume the Bayesian setting, i.e., f is a sample path from a GP (Rasmussen & Williams, 2005) with a zero mean and a stationary kernel function $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ denoted as $f \sim \mathcal{GP}(0, k)$. We denote a dataset at the beginning of t-th iteration as $\mathcal{D}_{t-1} \coloneqq \{(\boldsymbol{x}_i, y_i)\}_{i=1}^{n_{t-1}}$ with $n_{t-1} > 0$. Then, the posterior distribution $p(f \mid \mathcal{D}_{t-1})$ is again a GP, whose posterior mean and variance are derived as follows:

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

where $\boldsymbol{k}_{t-1}(\boldsymbol{x}) \coloneqq \left(k(\boldsymbol{x}, \boldsymbol{x}_1), \ldots, k(\boldsymbol{x}, \boldsymbol{x}_{n_{t-1}})\right)^\top \in \mathbb{R}^{n_{t-1}}$, $\boldsymbol{K} \in \mathbb{R}^{n_{t-1}\times n_{t-1}}$ is the kernel matrix whose (i, j)-element is $k(\boldsymbol{x}_i, \boldsymbol{x}_j)$, $\boldsymbol{I}_{n_{t-1}} \in \mathbb{R}^{n_{t-1}\times n_{t-1}}$ is the identity matrix, and $\boldsymbol{y}_{n_{t-1}} \coloneqq (y_1, \ldots, y_{n_{t-1}})^\top \in \mathbb{R}^{n_{t-1}}$. Hereafter, we denote that a probability density function (PDF) $p(\cdot|\mathcal{D}_{t-1}) = p_t(\cdot)$, a probability $\Pr(\cdot|\mathcal{D}_{t-1}) = \Pr_t(\cdot)$, and an expectation $\mathbb{E}[\cdot|\mathcal{D}_{t-1}] = \mathbb{E}_t[\cdot]$ for brevity.

2.2. Performance Mesuare

This paper analyzes the *Bayesian regret* (Russo & Van Roy, 2014) of BO algorithms. The BCR and Bayesian simple regret (BSR) are defined as follows:

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

where the expectation is defined via all the randomness, i.e., $f, \{\epsilon_t\}_{t\ge 1}$, and the randomness of BO algorithms. Since $\arg\max_{t\le T} f(\boldsymbol{x}_t)$ is unknown in practice, the modified BSR with the recommendation, for example $\overline{\mathrm{BSR}}_T \coloneqq \mathbb{E}\left[f(\boldsymbol{x}^*) - f(\hat{\boldsymbol{x}}_T)\right]$, where $\hat{\boldsymbol{x}}_T = \arg\max_{\boldsymbol{x}\in\mathcal{X}} \mu_{T-1}(\boldsymbol{x})$, is often also analyzed (Appendix A in Takeno et al., 2023).

The following lemma shows the property of BSR.

Lemma 2.1. *BSR can be bounded from above as* $\overline{\mathrm{BSR}}_T \le \sum_{t=1}^{T} \overline{\mathrm{BSR}}_t / T \le \mathrm{BCR}_T / T$ *and* $\mathrm{BSR}_T \le \mathrm{BCR}_T / T$.

The proof is shown in Appendix D. Thus, we mainly focus on showing that BCR is sub-linear since it implies BSR converges to 0. In addition, for continuous input domain, we use $\overline{\mathrm{BSR}}_T \le \sum_{t=1}^{T} \overline{\mathrm{BSR}}_t / T$ to show the modified BSR bound of PIMS, which is tighter than the direct consequence of BCR bound. Furthermore, as discussed in Russo & Van Roy (2014); Takeno et al. (2023), although the rate regarding the probability δ becomes worse, the BCR bound implies high probability bounds as a direct consequence of Markov's inequality. Thus, although tighter regret bounds regarding δ is an important future work, we mainly discuss the BCR upper bound in this paper.

For BCR analysis, we use the following widely used assumption for the continuous $\mathcal{X}$ (e.g., Srinivas et al., 2010; Kandasamy et al., 2018; Takeno et al., 2023):

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 $a \ge 1$ and $b > 0$ such that,*

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

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

Further, we define *maximum information gain* (MIG) (Srinivas et al., 2010; Vakili et al., 2021):

Definition 2.1 (Maximum information gain)**.** *Let $f \sim \mathcal{GP}(0, k)$ over $\mathcal{X} \subset [0, r]^d$. Let $A = \{\boldsymbol{a}_i\}_{i=1}^{T} \subset \mathcal{X}$. Let $\boldsymbol{f}_A = \left(f(\boldsymbol{a}_i)\right)_{i=1}^{T}$, $\boldsymbol{\epsilon}_A = \left(\epsilon_i\right)_{i=1}^{T}$, 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:*

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

where I is the Shanon mutual information.

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

2.3. Related Work

Many AFs for BO, such as expected improvement (EI) (Mockus et al., 1978), entropy search (ES) (Hennig & Schuler, 2012; Villemonteix et al., 2009), predictive ES (PES)(Hernández-Lobato et al., 2014), and knowledge gradients (KG)(Frazier et al., 2009), have been proposed. In these AFs, only the regret analysis of EI in the noiseless setting is provided (Bull, 2011). Therefore, particularly in the general case that noises contaminate observations, these AFs are heuristics without regret analysis. Furthermore, although the practical effectiveness of ES, PES, and KG has been reported, they are computationally expensive and hard to implement due to cumbersome approximations.

Max-value entropy search (MES) (Wang & Jegelka, 2017) is a variant of ES based on the entropy of optimal value in contrast to the optimal solution used in ES and PES. MES is easier to compute than ES and PES and empirically shows high practical performance. Wang & Jegelka (2017) claimed that the special case of MES, in which only one Monte Carlo (MC) sample is used, achieves a high-probability simple regret bound. PIMS is equivalent to this variant of MES with one MC sample, as shown in Lemma 3.1 in Wang & Jegelka (2017). However, several technical issues for the proof have been pointed out (Takeno et al., 2022). In addition, since using only one MC sample is hard to interpret in the sense of MC estimation, most experiments in Wang & Jegelka (2017) were performed with 100 MC samples. In contrast, we provide analyses depending on the randomness of the algorithm, which implies that using only one sample is more promising. Furthermore, MES with many MC samples requires heavy computational time.

Probability of improvement (PI) (Kushner, 1962; 1964) is also a well-known BO method, which evaluates the probability that the next evaluated function value improves the best observation at iteration t. To the best of our knowledge, only Wang et al. (2018) have provided the high-probability bound of simple regret for PI. However, the regret analysis in Wang et al. (2018) is based on a strong assumption that the upper bound of the black-box objective function $\hat{f}^* \ge f(\boldsymbol{x}^*)$ is known beforehand. Although GP estimation (GP-EST) (Wang et al., 2016) can also be seen as a PI with adaptively estimated $\hat{f}^*$, its estimation guaranteeing $\hat{f}^* \ge f(\boldsymbol{x}^*)$

is not obvious, as discussed in Takeno et al. (2023). If we admit Theorem 2 in Wang et al. (2018), it only implies that $f(\boldsymbol{x}^*) - \max_{t \le T} f(\boldsymbol{x}_t) = O(\sqrt{\gamma_T/T + \sigma^2})$, i.e., $T\big(f(\boldsymbol{x}^*) - \max_{t \le T} f(\boldsymbol{x}_t)\big) = O(\sqrt{T\gamma_T + T^2\sigma^2})$, which is linear with respect to T. On the other hand, the cumulative regret $\sum_{t \le T} f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t) \ge T\big(f(\boldsymbol{x}^*) - \max_{t \le T} f(\boldsymbol{x}_t)\big)$. Hence, Theorem 2 in Wang et al. (2018) does not show a sub-linear cumulative regret bound. Thus, we believe that deriving the sub-linear cumulative regret upper bound of PI-type AFs is still an open problem.

In contrast, several kinds of research represented by GP-UCB (Srinivas et al., 2010) and TS (Russo & Van Roy, 2014) achieve the sub-linear regret upper bounds. The assumptions of the regret analysis in BO are two-fold: frequentists and Bayesian settings. In the frequentist setting (e.g., Srinivas et al., 2010; Janz et al., 2020; Chowdhury & Gopalan, 2017), f is assumed to be an element of reproducing kernel Hilbert space (RKHS) specified by a kernel function for GP. In the Bayesian setting (e.g., Srinivas et al., 2010; Russo & Van Roy, 2014; Kandasamy et al., 2018), f is assumed to be a sample path from the GP. For TS, Russo & Van Roy (2014); Kandasamy et al. (2018) and Chowdhury & Gopalan (2017) provided the regret analysis in the Bayesian and frequentists settings, respectively. Although the regret analysis of TS and PIMS in the frequentist setting is an important research direction, this paper concentrates on the Bayesian setting.

Although GP-UCB achieves the sub-linear BCR bound, the theoretical $\beta_t \propto \log\big(|\mathcal{X}|t\big)$ can be huge. Thus, manually tuning confidence parameter β_t can be a practically crucial problem. On the other hand, Takeno et al. (2023) shows IRGP-UCB, which uses two-parameter exponential random variables $\{\zeta_t\}_{t \ge 1}$ instead of β_t, achieves the tighter BCR bound and $\mathbb{E}[\zeta_t] \propto \log(|\mathcal{X}|)$, in which increasing $\mathbb{E}[\zeta_t]$ is not required. However, $\log(|\mathcal{X}|)$ can be large depending on the problem. Thus, although the problem of a too-large confidence parameter is alleviated in IRGP-UCB, hyperparameter tuning is still required. PIMS can also be interpreted as a randomized variant of GP-UCB, in which the confidence parameter is defined via the maximum of the sample path from the posterior. See Sec. 4.2 for details.

3. Tighter BCR Bounds for TS

AF of TS is the sample path from the posterior. Algorithm 1 shows an algorithm of TS.

3.1. Regret Analysis

We use the following lemma in Takeno et al. (2023), which is a modified lemma from Srinivas et al. (2010):

Lemma 3.1 (Lemma 4.1 in (Takeno et al., 2023))**.** *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 $t \ge 1$. Then, for any given $\mathcal{D}_{t-1}$,*

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

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

Algorithm 1 TS

Require: Input space $\mathcal{X}$, GP prior $\mu = 0$ and k, and initial dataset $\mathcal{D}_0$

for $t = 1, \dots$ **do**
  Fit GP to $\mathcal{D}_{t-1}$
  Generate a sample path $g_t \sim p(f|\mathcal{D}_{t-1})$
  $\boldsymbol{x}_t \leftarrow \arg\max_{\boldsymbol{x} \in \mathcal{X}} g_t(\boldsymbol{x})$
  Observe $y_t = f(\boldsymbol{x}_t) + \epsilon_t$ and $\mathcal{D}_t \leftarrow \mathcal{D}_{t-1} \cup (\boldsymbol{x}_t, y_t)$
end for

From Lemma 3.1, we show the following lemma:

Lemma 3.2. *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. Let $\eta_t := \frac{g_t(\boldsymbol{x}_t) - \mu_{t-1}(\boldsymbol{x}_t)}{\sigma_{t-1}(\boldsymbol{x}_t)}$, where $g_t \sim p(f|\mathcal{D}_{t-1})$ and $\boldsymbol{x}_t = \arg\max_{\boldsymbol{x} \in \mathcal{X}} g_t(\boldsymbol{x})$. Then, for all $t \ge 1$, the following inequality holds:*

$$\mathbb{E}\left[\eta_t^2 \mathbb{1}\{\eta_t \ge 0\}\right] \le 2 + 2\log\big(|\mathcal{X}|/2\big),$$

where $\mathbb{1}\{\eta_t \ge 0\} = 1$ if $\eta_t \ge 0$, and otherwise 0.

See Appendix B for the proof. Note that $f(\boldsymbol{x}^*)$ and η_t do not follow Gaussian distribution since $\boldsymbol{x}^*$ is chosen as the maximizer.

By using Lemma 3.2, we can obtain a tighter BCR bound, in which $\sqrt{\log T}$ factor is removed compared with Proposition 5 in Russo & Van Roy (2014):

Theorem 3.1. *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. Then, by running TS, BCR can be bounded as follows:*

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

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

short proof. For all $t \ge 1$, the following holds:

$$\begin{aligned}
\mathbb{E}_t\left[f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t)\right] &= \mathbb{E}_t\left[g_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right] \\
&= \mathbb{E}_t\left[\frac{g_t(\boldsymbol{x}_t) - \mu_{t-1}(\boldsymbol{x}_t)}{\sigma_{t-1}(\boldsymbol{x}_t)} \sigma_{t-1}(\boldsymbol{x}_t)\right] \\
&\le \mathbb{E}_t\left[\eta_t \mathbb{1}\{\eta_t \ge 0\} \sigma_{t-1}(\boldsymbol{x}_t)\right],
\end{aligned}$$

where $\eta_t = \frac{g_t(\boldsymbol{x}_t) - \mu_{t-1}(\boldsymbol{x}_t)}{\sigma_{t-1}(\boldsymbol{x}_t)}$. Then, applying the Cauchy–Schwartz and Jensen inequalities and $\sum_{t \le T} \sigma_{t-1}^2(\boldsymbol{x}_t) \le$

$C_1\gamma_T$ (Srinivas et al., 2010) to the summation for all $t \leq T$, we can obtain

$$\begin{aligned}\mathrm{BCR}_T &\leq \sqrt{\sum_{t\leq T} \mathbb{E}\left[\eta_t^2 \mathbb{1}\{\eta_t \geq 0\}\right]}\sqrt{C_1\gamma_T} \\ &\leq \sqrt{C_1 C_2 T \gamma_T},\end{aligned}$$

where we use Lemma 3.2 in the second inequality. □

See Appendix B.1 for the details. Theorem 3.1 shows that TS achieves the same BCR bounds as IRGP-UCB.

Next, we consider the continuous $\mathcal{X}$. In the proof for the continuous $\mathcal{X}$, the discretization of $\mathcal{X}$ denoted as $\mathcal{X}_t$ is commonly used (e.g., Srinivas et al., 2010). We describe the nearst point in $\mathcal{X}_t$ of $\boldsymbol{x}$ as $[\boldsymbol{x}]_t$.

If we modify TS so that $[\boldsymbol{x}_t]_t$ is evaluated instead of $\boldsymbol{x}_t$, the same BCR upper bound of TS as that of IRGP-UCB can be obtained similarly to the existing analysis (Takeno et al., 2023) (See Appendix B.3 for details.). However, the theoretical discretization depends on the generally unknown parameters, such as a and b in Assumption 2.1. Thus, discretization fineness must be set manually in the algorithm. Hence, the benefit of TS having no hyperparameters is lost. Therefore, the regret analysis without the discretization in the algorithm is also important.

Kandasamy et al. (2018) have shown the BCR bound of TS without the discretization for the continuous $\mathcal{X}$. However, in Eq. (8) of Kandasamy et al. (2018), non-obvious result that the $\sum_{t\leq T} \sigma_t^2([\boldsymbol{x}_t]_t)$ is bounded from above by MIG γ_T is used without the proof[1]. Since we cannot show this non-obvious result, we show a more explicit analysis by using the following useful result (Theorem E.4 in Kusakawa et al., 2022).

Lemma 3.3 (Lipschitz constants for posterior standard deviation)**.** *Let $k(\boldsymbol{x}, \boldsymbol{x}') : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be Linear, Gaussian, or Matérn kernel and $k(\boldsymbol{x}, \boldsymbol{x}) = 1$. Moreover, assume that a noise variance σ^2 is positive. Then, for any $t \geq 1$ and $\mathcal{D}_{t-1}$, the posterior standard deviation $\sigma_{t-1}(\boldsymbol{x})$ satisfies that*

$$\forall \boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^d,\ |\sigma_{t-1}(\boldsymbol{x}) - \sigma_{t-1}(\boldsymbol{x}')| \leq L_\sigma \|\boldsymbol{x} - \boldsymbol{x}'\|_1,$$

where L_σ is a positive constant given by

$$L_\sigma = \begin{cases} 1 & \text{if } k(\boldsymbol{x}, \boldsymbol{x}') \text{ is the linear kernel}, \\ \frac{\sqrt{2}}{\ell} & \text{if } k(\boldsymbol{x}, \boldsymbol{x}') \text{ is the Gaussian kernel}, \\ \frac{\sqrt{2}}{\ell}\sqrt{\frac{\nu}{\nu-1}} & \text{if } k(\boldsymbol{x}, \boldsymbol{x}') \text{ is the Matérn kernel}, \end{cases}$$

where ℓ is length scale parameters in Gaussian and Matérn kernels, and ν is a degree of freedom with $\nu > 1$.

[1]Lemma 7 in (Kandasamy et al., 2018) only implies that $\sum_{t\leq T} \sigma_t^2(\boldsymbol{x}_t) \leq C_1\gamma_T$ without the discretization.

For completeness, a more detailed version of this lemma is provided in Appendix B.2. Then, Lemma 3.3 derives the upper bound of $\sigma_t([\boldsymbol{x}_t]_t)$:

$$\sigma_t([\boldsymbol{x}_t]_t) \leq \sigma_t(\boldsymbol{x}_t) + L_\sigma \|\boldsymbol{x}_t - [\boldsymbol{x}_t]_t\|_1.$$

Thus, appropriately refining $\mathcal{X}_t$, we can obtain the following BCR bounds for continuous $\mathcal{X}$ without the discretization in the algorithm:

Theorem 3.2. *Let $f \sim \mathcal{GP}(0, k)$, where k is a stationary kernel, $k(\boldsymbol{x}, \boldsymbol{x}) = 1$, and Assumption 2.1 holds. Let $L = \max\left\{L_\sigma, b\left(\sqrt{\log(ad)} + \sqrt{\pi}/2\right)\right\}$, where L_σ is defined as in Lemma 3.3. Then, by running TS, BCR can be bounded as follows:*

$$\mathrm{BCR}_T \leq \frac{\pi^2}{3} + \frac{\pi^2}{6}\sqrt{s_T} + \sqrt{C_1\gamma_T T s_T},$$

where $C_1 := 2/\log(1+\sigma^{-2})$ and $s_t = 2 - 2\log 2 + 2d\log(\lceil drLt^2 \rceil)$.

See Appendix B.2 for the proof. Main differences from Kandasamy et al. (2018) are (i) constant factors in s_t are essentially tightened as $\log(|\mathcal{X}_t|)$ from $\log(|\mathcal{X}_t|t^2)$ using Lemma 3.2 (ii) the rate regarding a is improved as $\log(\log a)$ from $\log a$ using Lemma H.1 of Takeno et al. (2023) (iii) the discretization error regarding posterior standard deviation is considered as discussed earlier.

4. Tighter Bayesian Regret Bounds for PIMS

AF of PIMS is the PI from the maximum of the sample path from the posterior:

$$\boldsymbol{x}_t = \operatorname*{arg\,max}_{\boldsymbol{x}\in\mathcal{X}} \left\{1 - \Phi\left(\frac{g_t^* - \mu_{t-1}(\boldsymbol{x})}{\sigma_{t-1}(\boldsymbol{x})}\right)\right\},$$

where $g_t^* = \max_{\boldsymbol{x}\in\mathcal{X}} g_t(\boldsymbol{x})$, $g_t \sim p(f|\mathcal{D}_{t-1})$, and Φ is the cumulative distribution function of a standard Gaussian distribution. From the monotonicity of Φ, we can rewrite as follows:

$$\boldsymbol{x}_t = \operatorname*{arg\,min}_{\boldsymbol{x}\in\mathcal{X}} \frac{g_t^* - \mu_{t-1}(\boldsymbol{x})}{\sigma_{t-1}(\boldsymbol{x})}.$$

We employ this equivalent expression due to its simplicity. Algorithm 2 shows an algorithm of PIMS.

Usual PI is based on the current best observation $\max_{i\leq t-1} y_i$ instead of g_t^* and is known to result in over-exploitation frequently. In the small noise variance regime, g_t^* is often larger than $\max_{i\leq t-1} y_i$ and randomly generated. Thus, PIMS explores appropriately compared with the usual PI since the AF of PI tends to explore when the base value (i.e., $\max_{i\leq t-1} y_i$ or g_t^*) is large.

Algorithm 2 PIMS

Require: Input space $\mathcal{X}$, GP prior $\mu = 0$ and k, and initial dataset $\mathcal{D}_0$

for $t = 1, \dots$ **do**
  Fit GP to $\mathcal{D}_{t-1}$
  Generate a sample path $g_t \sim p(f|\mathcal{D}_{t-1})$
  $g_t^* \leftarrow \max_{\boldsymbol{x} \in \mathcal{X}} g_t$
  $\boldsymbol{x}_t \leftarrow \arg\min_{\boldsymbol{x} \in \mathcal{X}} \frac{g_t^* - \mu_{t-1}(\boldsymbol{x})}{\sigma_{t-1}(\boldsymbol{x})}$
  Observe $y_t = f(\boldsymbol{x}_t) + \epsilon_t$ and $\mathcal{D}_t \leftarrow \mathcal{D}_{t-1} \cup (\boldsymbol{x}_t, y_t)$
end for

4.1. Regret Analysis

First, we show a similar lemma as Lemma 3.2 as follows:

Lemma 4.1. *Assume the same condition as in Lemma 3.2. Let* $\xi_t \coloneqq \min_{\boldsymbol{x} \in \mathcal{X}} \frac{g_t^* - \mu_{t-1}(\boldsymbol{x})}{\sigma_{t-1}(\boldsymbol{x})}$. *Then, the following inequality holds:*

$$\mathbb{E}\left[\xi_t^2 \mathbb{1}\{\xi_t \geq 0\}\right] \leq 2 + 2\log\big(|\mathcal{X}|/2\big),$$

for all $t \geq 1$.

See Appendix C.1 for the proof.

Using Lemma 4.1, we can obtain the following BCR bound for the finite input domain:

Theorem 4.1. *Assume the same condition as in Theorem 3.1. Then, by running PIMS, 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* $C_2 \coloneqq 2 + 2\log\big(|\mathcal{X}|/2\big)$.

See Appendix C.1 for the proof. Therefore, PIMS achieves the same BCR bound as TS and IRGP-UCB.

Next, we consider the continuous $\mathcal{X}$. As with TS, if we modify PIMS so that $\tilde{g}_t^* \coloneqq \max_{\boldsymbol{x} \in \mathcal{X}_t} g_t(\boldsymbol{x})$ is used instead of g_t^*, the proof of the sub-linear BCR bound of PIMS is relatively easy (See Appendix C.3 for details). However, for the same reason as TS, the theoretical discretization in the algorithm is not preferable. Therefore, we show the BCR bound using g_t^* without the discretization in the algorithm.

On the other hand, the same proof as TS does not provide the BCR bound of PIMS. This is because $\boldsymbol{x}_t$ can be far away from $\arg\min_{\boldsymbol{x} \in \mathcal{X}} \big(\tilde{g}_t^* - \mu_{t-1}(\boldsymbol{x})\big)/\sigma_{t-1}(\boldsymbol{x})$ in contrast to TS. Thus, instead of Lipshitz continuity of $\sigma_{t-1}(\boldsymbol{x})$ (Lemma 3.3), we use the following lemma, which is repeatedly used in the literature (e.g., Lemma 13 in Mutny & Krause, 2018; Wang & de Freitas, 2014):

Lemma 4.2. *Let* k *be a kernel s.t.* $k(\boldsymbol{x}, \boldsymbol{x}) = 1$. *Then, the posterior variance is bounded from below as,*

$$\sigma_t^2(\boldsymbol{x}) \geq \frac{\sigma^2}{\sigma^2 + n_t},$$

for all $\boldsymbol{x} \in \mathcal{X}$ *and for all* $t \geq 0$, *where* $n_t = |\mathcal{D}_t|$.

For completeness, the proof is shown in Appendix D.

Then, we can obtain the following BCR bound:

Theorem 4.2. *Let* $f \sim \mathcal{GP}(0, k)$, *where* k *is a stationary kernel,* $k(\boldsymbol{x}, \boldsymbol{x}) = 1$, *and Assumption 2.1 holds. Then, by running PIMS, BCR can be bounded as follows:*

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

where $C_1 \coloneqq 2/\log(1+\sigma^{-2})$ *and* $m_t \coloneqq 2d\log\left(\left\lceil t^2 bdr\big(\log(ad) + \sqrt{\pi}/2\big)\sqrt{(\sigma^2+n_t)/\sigma^2}\right\rceil\right) - 2\log 2 + 2$.

See Appendix C.2 for the proof. Since $m_t = O(\log t)$, $\mathrm{BCR}_T = O(\sqrt{T\gamma_T \log T})$, which is the same rate as the existing results of GP-UCB, IRGP-UCB, and TS.

As discussed in Sec. 2, our BCR bound directly implies that BSR converges to 0. On the other hand, for the modified BSR, we can avoid the use of Lemma 4.2 using the inequality $\overline{\mathrm{BSR}}_T \leq \sum_{t=1}^T \overline{\mathrm{BSR}}_t / T$ in Lemma 2.1:

Theorem 4.3. *Assume the same condition as in Theorem 4.2. Then, by running PIMS, BSR can be bounded as follows:*

$$\overline{\mathrm{BSR}}_T \leq \frac{\pi^2}{6T} + \sqrt{\frac{C_1 \gamma_T m_T}{T}},$$

where $C_1 \coloneqq 2/\log(1+\sigma^{-2})$ *and* $m_t \coloneqq 2d\log\left(\left\lceil t^2 bdr\big(\log(ad) + \sqrt{\pi}/2\big)\right\rceil\right) - 2\log 2 + 2$.

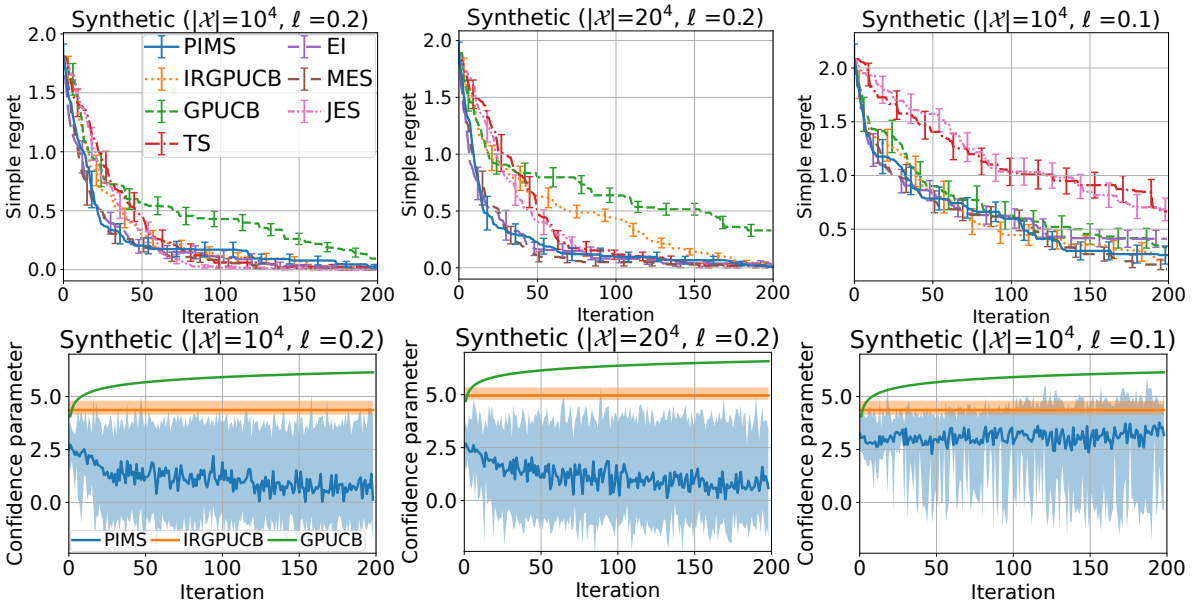
See Appendix C.2 for the proof. Therefore, we can remove the term $\log((\sigma^2+n_t)/\sigma^2)$ in m_t.

4.2. Relationship with GP-UCB

Wang & Jegelka (2017) have also shown the equivalence between PIMS and GP-UCB with $\beta_t^{1/2} = \xi_t$ defined in Lemma 4.1. (We employ this definition of ξ_t since ξ_t can be negative.) Details are shown in Appendix A. Therefore, PIMS can also be interpreted as a random variant of GP-UCB, whose confidence parameter is determined via g_t^*. Thus, we can expect that PIMS defines the more practical confidence parameter ξ_t based on the GP model in contrast to GP-UCB and IRGP-UCB, whose theoretical confidence parameters are defined by the worst-case analysis and are too conservative. We will show this important benefit of PIMS compared with GP-UCB and IRGP-UCB in Sec. 5.

5. Experiments

We demonstrate the experiments on synthetic functions generated from GP, benchmark functions, and emulators derived from real-world datasets. We performed the existing



(a) Default setting. (b) Large $|\mathcal{X}| = 20^4$. (c) Small $\ell = 0.1$.

Figure 1. The results on synthetic function experiments. The top row shows the average and standard error of the simple regret. In all the settings (a-c), we can confirm that several BO methods, including PIMS (blue), achieve the best convergence. In contrast, other theoretically guaranteed methods (GP-UCB, IRGP-UCB, TS) deteriorated in a certain setting. Therefore, we can observe that PIMS flexibly deals with various problem settings while keeping the theoretical guarantee and has superior or comparable performance compared with baselines, including heuristic methods without the theoretical guarantee, such as EI, MES, and JES. The bottom row represents the expectation and quantiles of $\beta_t^{1/2}$, $\zeta_t^{1/2}$, and ξ_t.

BO methods, EI(Mockus et al., 1978), GP-UCB(Srinivas et al., 2010), IRGP-UCB(Takeno et al., 2023), TS(Russo & Van Roy, 2014), MES(Wang & Jegelka, 2017), joint entropy search (JES) (Hvarfner et al., 2022) as baselines. We generated 5 inputs for the initial dataset using the Latin hypercube sampling(Loh, 1996). We fixed the noise variance $\sigma^2 = 10^{-6}$ and used the RBF kernel $k(\boldsymbol{x}, \boldsymbol{x}') = \exp\left(-\|\boldsymbol{x} - \boldsymbol{x}'\|_2^2/(2\ell^2)\right)$, where ℓ is hyperparameter. For the posterior sampling in TS, PIMS, MES, and JES, we used the random Fourier feature (RFF) (Rahimi & Recht, 2008). For MC estimation in MES and JES, we generated 10 MC samples. Unless otherwise noted, we report the average and standard errors of the simple regret $r_t \coloneqq f(\boldsymbol{x}^*) - \max_{t \le T} f(\boldsymbol{x}_t)$ over 20 random trials.

Overall, we focus on the effectiveness of theoretical BO methods with regret analysis. In synthetic function experiments, we consider the purely theoretical setting, in which we set hyperparameter ℓ as one used to generate the objective function. We aim to demonstrate that PIMS shows at least comparable performance to baselines including not theoretically guaranteed methods, in contrast to that GP-UCB-based methods and TS are degraded by over-exploration. In benchmark function and real-world emulator experiments, we performed marginal likelihood maximization (Rasmussen & Williams, 2005) every 5 iteration. Furthermore, in these experiments, we use the manually tuned confidence parameters for GP-UCB-based methods since theoretical confidence parameters contain unknown parameters such as a and b. Thus, GP-UCB-based methods are no longer theoretically guaranteed. Hence, only TS and PIMS match the theoretical algorithm.

5.1. Experiments on Synthetic Functions

We performed the experiments for the GP-derived synthetic function and the finite input domain $\mathcal{X}$. Note that the objective function is also randomly generated for all 20 trials. We show the results of changing the length scale parameter ℓ and the cardinality of $\mathcal{X}$. In particular, note that small ℓ implies that the objective function and the GP model are more complex. The input domain is set as $\{0.1, 0.2, \dots, 1\}^4$ when $|\mathcal{X}| = 10^4$ and $\{0.05, 0.1, \dots, 1\}^4$ when $|\mathcal{X}| = 20^4$, respectively. For this experiment, we used the theoretical confidence parameters, $\beta_t = 2\log\left(|\mathcal{X}|t^2/\sqrt{2\pi}\right)$ for GP-UCB and $\zeta_t \sim \mathrm{Exp}\left(2\log\left(|\mathcal{X}|/2\right), 1/2\right)$ for IRGP-UCB.

The top row of Fig. 1 shows the simple regret. In this experiment, we performed PI (Kushner, 1964), by which we can confirm the over-exploitation tendency of the usual PI. In contrast, we see that PIMS shows the best or comparable performance in all the plots of Fig. 1 compared with the baselines, including the heuristic methods without theoretical guarantees, such as EI, MES, and JES. Consequently,

Table 2. Average and standard deviation of the mean of evaluated posterior standard deviation $\sum_{t \leq T} \sigma_{t-1}(\boldsymbol{x}_t)/T$ for synthetic function experiments.

	Fig. 1a	Fig. 1b	Fig. 1c
PIMS	0.27 ± 0.14	0.26 ± 0.13	0.71 ± 0.13
TS	0.36 ± 0.14	0.36 ± 0.14	0.92 ± 0.09

the effectiveness of PIMS under different $|\mathcal{X}|$ and ℓ can be confirmed.

GP-UCB-based methods deteriorated due to the over-exploration, particularly in Fig. 1b with $|\mathcal{X}| = 20^4$, since the confidence parameters of GP-UCB and IRGP-UCB depend on $\mathcal{X}$ as $\log |\mathcal{X}|$. This can also be confirmed by the plot of the confidence parameters shown in the bottom row of Fig. 1. Note that PIMS can be interpreted as GP-UCB using ξ_t instead of $\beta_t^{1/2}$ as discussed in Sec. 4.2. Then, we can observe that ξ_t is substantially smaller than $\beta_t^{1/2}$ and $\zeta_t^{1/2}$, which is conservatively set via $|\mathcal{X}|$. In addition, comparing Figs. 1a and 1c, we can see that ξ_t becomes large (exploration is emphasized) when ℓ is small (the objective function has many local maxima). These results suggest that PIMS can use more practical confidence parameters based on the posterior sampling.

TS shows poor performance in Fig. 1c with small ℓ. We conjecture that this poor performance is caused by over-exploration. Thus, we show the average and standard deviation of the evaluated posterior standard deviation mean $\sum_{t \leq T} \sigma_{t-1}(\boldsymbol{x}_t)/T$ for PIMS and TS in Table 2. We can observe that TS evaluates the input whose posterior variance is larger than that of PIMS, particularly in Fig. 1c. Therefore, we can confirm that TS results in over-exploration. It is worth noting that $\ell = 0.1$ is ideal in the sense that $\ell = 0.1$ is used to generate the objective function. Considering the existing studies (e.g., Shahriari et al., 2016), in which TS deteriorated in the higher-dimensional problem, TS is empirically sensitive to the complexity of the GP model depending on the parameters such as d and ℓ.

5.2. Experiments on Benchmark Functions

Figure 2 shows the results on the benchmark functions called Ackley and Shekel functions in `https://www.sfu.ca/~ssurjano/optimization.html`. In this experiment, we employed the heuristic confidence parameters for GP-UCB and IRGP-UCB as $\beta_t = 0.2d \log(2t)$ (Kandasamy et al., 2015) and $\zeta_t \sim \mathrm{Exp}(2/d, 1/2)$ (Takeno et al., 2023). TS is inferior to other methods due to the over-exploration since a small ℓ is required to represent both functions, which have many local maxima. On the other hand, PIMS shows comparable or superior performance in both functions compared with all the baselines, including heuristics such as EI, MES, and JES. In particular, in the

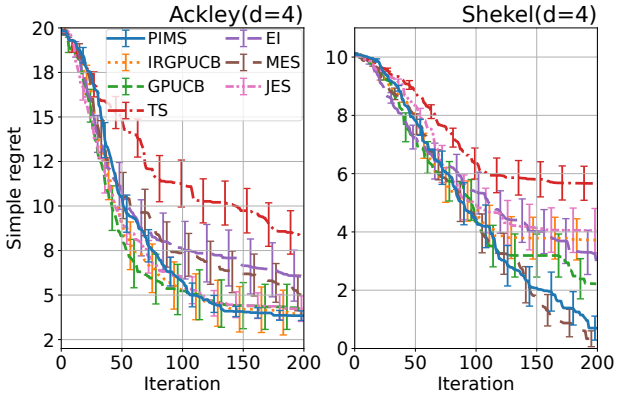


Figure 2. Average and standard error of the simple regret in benchmark function experiments.

Shekel function, PIMS shows faster convergence compared with GP-UCB-based methods with the heuristic confidence parameters. Thus, we conclude that PIMS appropriately controls the exploitation and exploration trade-off via the posterior sampling in this experiment.

5.3. Experiments on Real-World Emulators

We performed the experiments on the real-world data emulators (Häse et al., 2021). The alkox dataset (Häse et al., 2021) is the measurement of alkoxylation reaction with respect to 4 parameters: catalase, horseradish peroxidase, alcohol oxidase, and pH. The Fullerenes dataset (Walker et al., 2017) is the mole fraction of the desired products regarding o-xylenyl adducts of Buckminsterfullerenes with respect to 3 parameters: temperature, reaction time, and the ratio of sultine to C_{60}. We optimized the Bayes neural network trained by each dataset as proposed by Häse et al. (2021). Details are shown in Häse et al. (2021).

Figure 3 shows the best observed value since the optimal value is unknown. Note that GP-UCB-based methods are based on the heuristic confidence parameters as with the benchmark function experiments. We can confirm that EI, JES, and TS are inferior to others in both emulators. In contrast, GP-UCB-based methods and PIMS show superior performance. Thus, we can confirm the effectiveness of PIMS in both emulators derived from the real-world dataset without heuristic tunings, unlike GP-UCB-based methods.

6. Conclusion

First, we derived the tighter BCR bounds of TS. Furthermore, we analyze the randomized variant of PI called PIMS, which alleviates the practical issues of GP-UCB-based methods and TS. We showed the BCR bounds of PIMS, whose rate is the same as TS and IRGP-UCB. Finally, we demonstrated the effectiveness of PIMS, particularly focusing on

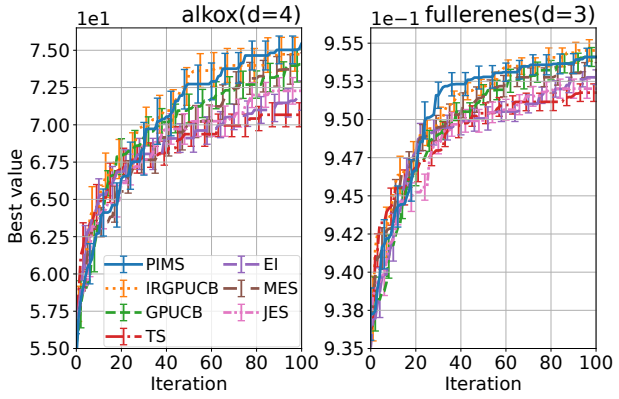


Figure 3. Average and standard error of the obtained best value in real-world emulator experiments.

the comparison of PIMS, TS, and GP-UCB-based methods.

Limitations and Future Works Although the hyperparameters of AF do not exist, the hyperparameter tuning of GP is crucial. Furthermore, the posterior sampling must be approximated by RFF for continuous $\mathcal{X}$. Therefore, regret analysis incorporating the uncertainty of GP hyperparameters (Berkenkamp et al., 2019; Bogunovic & Krause, 2021) and the RFF approximation error (Mutny & Krause, 2018) are important future directions. Moreover, as discussed in Sec. 2.3, the regret analysis for the RKHS setting is also intriguing. On the other hand, various extensions for, e.g., multi-fidelity (Kandasamy et al., 2016), constrained (Gardner et al., 2014), and multi-objective (Paria et al., 2020) optimization, are practically important. In addition, EI based on the posterior sampling may be motivated as with PIMS.

Acknowkedgements

This work was partially supported by MEXT Program: Data Creation and Utilization-Type Material Research and Development Project Grant Number JPMXP1122712807, JSPS KAKENHI Grant Numbers JP20H00601, JP21H03498, JP22H00300, JP23K16943, JP23K19967, JP23K21696, JST CREST Grant Numbers JPMJCR21D3, JPMJCR22N2, JST Moonshot R&D Grant Number JPMJMS2033-05, JST AIP Acceleration Research Grant Number JPMJCR21U2, JST ACT-X Grant Number JPMJAX23CD, NEDO Grant Numbers JPNP18002, JPNP20006, and RIKEN Center for Advanced Intelligence Project.

Impact Statement

This paper focuses on the theoretical and algorithmic aspects of the machine learning method. We consider that some potential societal consequences of this paper need not necessarily be highlighted here.

References

Berkenkamp, F., Schoellig, A. P., and Krause, A. No-regret Bayesian optimization with unknown hyperparameters. *Journal of Machine Learning Research*, 20(50):1–24, 2019.

Bogunovic, I. and Krause, A. Misspecified Gaussian process bandit optimization. *Advances in Neural Information Processing Systems 34*, pp. 3004–3015, 2021.

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.

Falkner, S., Klein, A., and Hutter, F. BOHB: Robust and efficient hyperparameter optimization at scale. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80, pp. 1437–1446. PMLR, 2018.

Frazier, P., Powell, W., and Dayanik, S. The knowledge-gradient policy for correlated normal beliefs. *INFORMS journal on Computing*, 21(4):599–613, 2009.

Gardner, J., Kusner, M., Zhixiang, Weinberger, K., and Cunningham, J. Bayesian optimization with inequality constraints. In *Proceedings of the 31st International Conference on Machine Learning*, volume 32, pp. 937–945. PMLR, 2014.

Häse, F., Aldeghi, M., Hickman, R. J., Roch, L. M., Christensen, M., Liles, E., Hein, J. E., and Aspuru-Guzik, A. Olympus: a benchmarking framework for noisy optimization and experiment planning. *Machine Learning: Science and Technology*, 2(3):035021, 2021.

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 35*, pp. 11494–11506. Curran Associates, Inc., 2022.

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., 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, 2018.

Klein, A., Falkner, S., Bartels, S., Hennig, P., and Hutter, F. Fast Bayesian optimization of machine learning hyperparameters on large datasets. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, volume 54, pp. 528–536. PMLR, 2017.

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.

Kushner, H. J. A versatile stochastic model of a function of unknown and time varying form. *Journal of Mathematical Analysis and Applications*, 5(1):150–167, 1962.

Kushner, H. J. A New Method of Locating the Maximum Point of an Arbitrary Multipeak Curve in the Presence of Noise. *Journal of Basic Engineering*, 86(1):97–106, 1964.

Loh, W.-L. On Latin hypercube sampling. *Ann. Statist.*, 24 (5):2058–2080, 1996.

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

Mutny, M. and Krause, A. Efficient high dimensional Bayesian optimization with additivity and quadrature Fourier features. *Advances in Neural Information Processing Systems 31*, pp. 9005–9016, 2018.

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.

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.

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, 2022.

Takeno, S., Inatsu, Y., and Karasuyama, M. Randomized Gaussian process upper confidence bound with tighter Bayesian regret bounds. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 33490–33515. PMLR, 2023.

Thompson, W. R. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3-4):285–294, 1933.

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.

Villemonteix, J., Vazquez, E., and Walter, E. An informational approach to the global optimization of expensive-to-evaluate functions. *Journal of Global Optimization*, 44(4):509–534, 2009.

Walker, B. E., Bannock, J. H., Nightingale, A. M., and deMello, J. C. Tuning reaction products by constrained optimisation. *React. Chem. Eng.*, 2:785–798, 2017.

Wang, Z. and de Freitas, N. Theoretical analysis of Bayesian optimisation with unknown Gaussian process hyper-parameters. *arXiv:1406.7758*, 2014.

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.

Wang, Z., Kim, B., and Kaelbling, L. P. Regret bounds for meta Bayesian optimization with an unknown Gaussian process prior. In *Advances in Neural Information Processing Systems 31*, pp. 10477–10488. Curran Associates, Inc., 2018.

A. Equivalence between PIMS and Existing Methods

Wang et al. (2016); Wang & Jegelka (2017) claimed the following lemma:

Lemma A.1. *The following methods are equivalent:*

1. *PIMS*
2. *MES with only one MC sample*
3. *(randomized) GP-UCB that maximize* $\mu_{t-1}(\boldsymbol{x}) + \xi_t \sigma_{t-1}(\boldsymbol{x})$*, where* $\xi_t = \min_{\boldsymbol{x} \in \mathcal{X}} \left\{ \left(g_t^* - \mu_{t-1}(\boldsymbol{x}) \right) / \sigma_{t-1}(\boldsymbol{x}) \right\}$.
4. *(GP-EST using* g_t^* *instead of* $\hat{\mathrm{m}}$ *defined in Wang et al. (2016))*

First, the equivalence between PIMS and MES with only one MC sample is shown in Lemma 3.1 in Wang & Jegelka (2017). This equivalence is obvious from the monotonicity of MES's AF when the MC sample size is one. Then, for the equivalence between PIMS and GP-UCB, Wang et al. (2016) provided the proof in Lemma 2.1. However, since this proof is slightly ambiguous, we provide the proof for completeness.

Let $\boldsymbol{x}_{\mathrm{UCB}}$ and $\boldsymbol{x}_{\mathrm{PIMS}}$ be the selected points by GP-UCB and PIMS, i.e.,

$$\boldsymbol{x}_{\mathrm{UCB}} \in \arg\max_{\boldsymbol{x} \in \mathcal{X}} \left\{ \mu_{t-1}(\boldsymbol{x}) + \xi_t \sigma_{t-1}(\boldsymbol{x}) \right\},$$
$$\boldsymbol{x}_{\mathrm{PIMS}} \in \arg\min_{\boldsymbol{x} \in \mathcal{X}} \left\{ \frac{g_t^* - \mu_{t-1}(\boldsymbol{x})}{\sigma_{t-1}(\boldsymbol{x})} \right\}.$$

Fix $g_t^* \in \mathbb{R}$. From the definitions of GP-UCB, PIMS, and ξ_t, the following holds:

$$\begin{aligned}
\mu_{t-1}(\boldsymbol{x}_{\mathrm{UCB}}) + \xi_t \sigma_{t-1}(\boldsymbol{x}_{\mathrm{UCB}}) &\geq \mu_{t-1}(\boldsymbol{x}_{\mathrm{PIMS}}) + \xi_t \sigma_{t-1}(\boldsymbol{x}_{\mathrm{PIMS}}) \\
&= \mu_{t-1}(\boldsymbol{x}_{\mathrm{PIMS}}) + \frac{g_t^* - \mu_{t-1}(\boldsymbol{x}_{\mathrm{PIMS}})}{\sigma_{t-1}(\boldsymbol{x}_{\mathrm{PIMS}})} \sigma_{t-1}(\boldsymbol{x}_{\mathrm{PIMS}}) \\
&= g_t^*.
\end{aligned}$$

Then, if $\mu_{t-1}(\boldsymbol{x}_{\mathrm{UCB}}) + \xi_t \sigma_{t-1}(\boldsymbol{x}_{\mathrm{UCB}}) > g_t^*$, we can see that

$$\xi_t > \frac{g_t^* - \mu_{t-1}(\boldsymbol{x}_{\mathrm{UCB}})}{\sigma_{t-1}(\boldsymbol{x}_{\mathrm{UCB}})}.$$

This clearly contradicts the definition of ξ_t. Therefore,

$$\begin{aligned}
\mu_{t-1}(\boldsymbol{x}_{\mathrm{UCB}}) + \xi_t \sigma_{t-1}(\boldsymbol{x}_{\mathrm{UCB}}) &= \mu_{t-1}(\boldsymbol{x}_{\mathrm{PIMS}}) + \xi_t \sigma_{t-1}(\boldsymbol{x}_{\mathrm{PIMS}}) \\
&= g_t^*.
\end{aligned}$$

In addition, we can transform the inequality as follows:

$$\begin{aligned}
\frac{g_t^* - \mu_{t-1}(\boldsymbol{x}_{\mathrm{UCB}})}{\sigma_{t-1}(\boldsymbol{x}_{\mathrm{UCB}})} &= \frac{g_t^* - \mu_{t-1}(\boldsymbol{x}_{\mathrm{PIMS}})}{\sigma_{t-1}(\boldsymbol{x}_{\mathrm{PIMS}})} \\
&= \xi_t.
\end{aligned}$$

Consequently, since the above equalities hold for all $g_t^* \in \mathbb{R}$, we can obtain

$$\boldsymbol{x}_{\mathrm{UCB}}, \boldsymbol{x}_{\mathrm{PIMS}} \in \arg\max_{\boldsymbol{x} \in \mathcal{X}} \left\{ \mu_{t-1}(\boldsymbol{x}) + \xi_t \sigma_{t-1}(\boldsymbol{x}) \right\},$$
$$\boldsymbol{x}_{\mathrm{UCB}}, \boldsymbol{x}_{\mathrm{PIMS}} \in \arg\min_{\boldsymbol{x} \in \mathcal{X}} \left\{ \frac{g_t^* - \mu_{t-1}(\boldsymbol{x})}{\sigma_{t-1}(\boldsymbol{x})} \right\}.$$

B. Proof regarding TS

First, we show the proof of the following lemma:

Lemma 3.2. *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. Let $\eta_t \coloneqq \frac{g_t(\boldsymbol{x}_t) - \mu_{t-1}(\boldsymbol{x}_t)}{\sigma_{t-1}(\boldsymbol{x}_t)}$, where $g_t \sim p(f|\mathcal{D}_{t-1})$ and $\boldsymbol{x}_t = \arg\max_{\boldsymbol{x} \in \mathcal{X}} g_t(\boldsymbol{x})$. Then, for all $t \geq 1$, the following inequality holds:*

$$\mathbb{E}\left[\eta_t^2 \mathbb{1}\{\eta_t \geq 0\}\right] \leq 2 + 2\log\bigl(|\mathcal{X}|/2\bigr),$$

where $\mathbb{1}\{\eta_t \geq 0\} = 1$ if $\eta_t \geq 0$, and otherwise 0.

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

$$\mathbb{E}_t\left[\eta_t^2 \mathbb{1}\{\eta_t \geq 0\}\right] \leq 2 + 2\log\bigl(|\mathcal{X}|/2\bigr).$$

Then, fix the dataset $\mathcal{D}_{t-1}$ and $\delta \in (0, 1)$. From Lemma 3.1, we can see that

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

where $\beta_\delta = 2\log(|\mathcal{X}|/(2\delta))$. Then, since $g_t, \boldsymbol{x}_t \mid \mathcal{D}_{t-1} \overset{\mathrm{d}}{=} f, \boldsymbol{x}^* \mid \mathcal{D}_{t-1}$, where $\overset{\mathrm{d}}{=}$ implies the equality in distribution, we can arrange the probability as follows:

$$\begin{aligned}\mathrm{Pr}_t\left(f(\boldsymbol{x}^*) \leq \mu_{t-1}(\boldsymbol{x}^*) + \beta_\delta^{1/2}\sigma_{t-1}(\boldsymbol{x}^*)\right) &= \mathrm{Pr}_t\left(\frac{f(\boldsymbol{x}^*) - \mu_{t-1}(\boldsymbol{x}^*)}{\sigma_{t-1}(\boldsymbol{x}^*)} \leq \beta_\delta^{1/2}\right) && \left(\because \sigma_{t-1}(\boldsymbol{x}^*) > 0\right)\\ &= \mathrm{Pr}_t\left(\eta_t \leq \beta_\delta^{1/2}\right) && \left(\because g_t, \boldsymbol{x}_t \mid \mathcal{D}_{t-1} \overset{\mathrm{d}}{=} f, \boldsymbol{x}^* \mid \mathcal{D}_{t-1}\right)\\ &= \mathrm{Pr}_t\left(\eta_t \mathbb{1}\{\eta_t \geq 0\} \leq \beta_\delta^{1/2}\right) && \left(\because \beta_\delta^{1/2} > 0\right).\end{aligned}$$

Therefore, we obtain

$$\mathrm{Pr}_t\left(\eta_t^2 \mathbb{1}\{\eta_t \geq 0\} \leq \beta_\delta\right) \geq 1 - \delta.$$

Using the cumulative distribution function $F_t(\cdot) \coloneqq \mathrm{Pr}_t(\eta_t^2 \mathbb{1}\{\eta_t \geq 0\} \leq \cdot)$ and its generalized inverse function F_t^{-1} ("generalized" is required since $\eta_t^2 \mathbb{1}\{\eta_t \geq 0\}$ is not an absolute continuous random variable), we can rewrite

$$F_t\left(\beta_\delta\right) \geq 1 - \delta \Longleftrightarrow \beta_\delta \geq F_t^{-1}(1 - \delta),$$

since F_t^{-1} is a monotone non-decreasing function. Since $\delta \in (0, 1)$ is arbitrary, we can substitute $U \sim \mathrm{Uni}(0, 1)$ as,

$$\beta_U \geq F_t^{-1}(1 - U).$$

By taking the expectation with respect to U, we can obtain,

$$\begin{aligned}\mathbb{E}_U\left[\beta_U\right] &\geq \mathbb{E}_U\left[F_t^{-1}(1 - U)\right]\\ &= \mathbb{E}_U\left[F_t^{-1}(U)\right],\end{aligned}$$

where we use the fact that $1 - U$ also follows $\mathrm{Uni}(0, 1)$. Since $F_t^{-1}(U)$ and $\eta_t^2 \mathbb{1}\{\eta_t \geq 0\}|\mathcal{D}_{t-1}$ are identically distributed as with the inverse transform sampling, we obtain

$$\mathbb{E}_t\left[\eta_t^2 \mathbb{1}\{\eta_t \geq 0\}\right] \leq \mathbb{E}_U\left[\beta_U\right] = 2 + 2\log\bigl(|\mathcal{X}|/2\bigr),$$

which concludes the proof. □

B.1. Discrete Domain

Theorem 3.1. *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. Then, by running TS, 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 $C_2 \coloneqq 2 + 2\log\bigl(|\mathcal{X}|/2\bigr)$.

Proof. We can obtain the following upper bound:

$$\begin{aligned}
\mathrm{BCR}_T &= \sum_{t=1}^{T} \mathbb{E}\left[f(\boldsymbol{x}^*) - g_t(\boldsymbol{x}_t) + g_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right] \\
&= \sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}_{t-1}}\left[\mathbb{E}_t\left[f(\boldsymbol{x}^*) - g_t(\boldsymbol{x}_t)\right]\right] + \sum_{t=1}^{T} \mathbb{E}\left[g_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right] \\
&= \sum_{t=1}^{T} \mathbb{E}\left[g_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right] \\
&= \sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}_{t-1}}\left[\mathbb{E}_t\left[g_t(\boldsymbol{x}_t) - f(\boldsymbol{x}_t)\right]\right] \\
&= \sum_{t=1}^{T} \mathbb{E}\left[g_t(\boldsymbol{x}_t) - \mu_{t-1}(\boldsymbol{x}_t)\right].
\end{aligned}$$

Then, we apply Lemma 3.2 to $\eta_t \coloneqq \frac{g_t(\boldsymbol{x}_t) - \mu_{t-1}(\boldsymbol{x}_t)}{\sigma_{t-1}(\boldsymbol{x}_t)}$. Therefore, we obtain

$$\begin{aligned}
\mathrm{BCR}_T &= \sum_{t=1}^{T} \mathbb{E}\left[g_t(\boldsymbol{x}_t) - \mu_{t-1}(\boldsymbol{x}_t)\right] \\
&= \sum_{t=1}^{T} \mathbb{E}\left[\eta_t \sigma_{t-1}(\boldsymbol{x}_t)\right] \\
&\leq \sum_{t=1}^{T} \mathbb{E}\left[\eta_t \mathbb{1}\{\eta_t \geq 0\} \sigma_{t-1}(\boldsymbol{x}_t)\right] \\
&\leq \mathbb{E}\left[\sqrt{\sum_{t=1}^{T} \eta_t^2 \mathbb{1}\{\eta_t \geq 0\} \sum_{t=1}^{T} \sigma_{t-1}^2(\boldsymbol{x}_t)}\right] && \bigl(\because \text{Cauchy–Schwarz inequality}\bigr) \\
&\leq \mathbb{E}\left[\sqrt{\sum_{t=1}^{T} \eta_t^2 \mathbb{1}\{\eta_t \geq 0\}}\right] \sqrt{C_1 \gamma_T} && \bigl(\because \text{Lemma 5.4 in (Srinivas et al., 2010)}\bigr) \\
&\leq \sqrt{\mathbb{E}\left[\sum_{t=1}^{T} \eta_t^2 \mathbb{1}\{\eta_t \geq 0\}\right]} \sqrt{C_1 \gamma_T} && \bigl(\because \text{Jensen inequality}\bigr) \\
&\leq \sqrt{C_1 C_2 T \gamma_T} && \bigl(\because \text{Lemma 3.2}\bigr),
\end{aligned}$$

where $C_2 \coloneqq 2 + 2\log\bigl(|\mathcal{X}|/2\bigr)$. □

B.2. Continuous Domain with Continuous Sample Path

First, we provide the rigorous version of Lemma 3.3:

Lemma B.1 (Theorem E.4 in (Kusakawa et al., 2022), Lipschitz constants for posterior standard deviation)**.** *Let $k(\boldsymbol{x}, \boldsymbol{x}') : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be one of the following kernel functions:*

Linear kernel: $k(\boldsymbol{x}, \boldsymbol{x}') = \sigma_f^2 \boldsymbol{x}^\top \boldsymbol{x}'$, *where* σ_f *is a positive parameter.*

Gaussian kernel: $k(\boldsymbol{x}, \boldsymbol{x}') = \sigma_f^2 \exp(-\|\boldsymbol{x} - \boldsymbol{x}'\|^2/(2\ell^2))$, *where* σ_f *and* ℓ *are positive parameters.*

Matérn kernel:

$$k(\boldsymbol{x}, \boldsymbol{x}') = \sigma_f^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu}\frac{\|\boldsymbol{x} - \boldsymbol{x}'\|}{\ell}\right)^\nu K_\nu \left(\sqrt{2\nu}\frac{\|\boldsymbol{x} - \boldsymbol{x}'\|}{\ell}\right),$$

where σ_f *and* ℓ *are positive parameters,* ν *is a degree of freedom with* $\nu > 1$, Γ *is the gamma function, and* K_ν *is the modified Bessel function of the second kind.*

Moreover, assume that a noise variance σ^2 *is positive. Then, for any* $t \geq 1$ *and* $\mathcal{D}_{t-1}$, *the posterior standard deviation* $\sigma_{t-1}(\boldsymbol{x})$ *satisfies that*

$$\forall \boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^d,\ |\sigma_{t-1}(\boldsymbol{x}) - \sigma_{t-1}(\boldsymbol{x}')| \leq L_\sigma \|\boldsymbol{x} - \boldsymbol{x}'\|_1,$$

where L_σ *is a positive constant given by*

$$L_\sigma = \begin{cases} \sigma_f & \text{if } k(\boldsymbol{x}, \boldsymbol{x}') \text{ is the linear kernel}, \\ \frac{\sqrt{2}\sigma_f}{\ell} & \text{if } k(\boldsymbol{x}, \boldsymbol{x}') \text{ is the Gaussian kernel}, \\ \frac{\sqrt{2}\sigma_f}{\ell}\sqrt{\frac{\nu}{\nu-1}} & \text{if } k(\boldsymbol{x}, \boldsymbol{x}') \text{ is the Matérn kernel}, \end{cases}$$

Next, we show the following BCR bounds of TS:

Theorem 3.2. *Let* $f \sim \mathcal{GP}(0, k)$, *where* k *is a stationary kernel,* $k(\boldsymbol{x}, \boldsymbol{x}) = 1$, *and Assumption 2.1 holds. Let* $L = \max\{L_\sigma, b(\sqrt{\log(ad)} + \sqrt{\pi}/2)\}$, *where* L_σ *is defined as in Lemma 3.3. Then, by running TS, BCR can be bounded as follows:*

$$\mathrm{BCR}_T \leq \frac{\pi^2}{3} + \frac{\pi^2}{6}\sqrt{s_T} + \sqrt{C_1 \gamma_T T s_T},$$

where $C_1 \coloneqq 2/\log(1 + \sigma^{-2})$ *and* $s_t = 2 - 2\log 2 + 2d\log(\lceil drLt^2 \rceil)$.

Proof. 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 $\tau_t = \lceil drLt^2 \rceil$. Thus, $|\mathcal{X}_t| = \tau_t^d$. In addition, we define $[\boldsymbol{x}]_t$ as the nearest point in $\mathcal{X}_t$ of $\boldsymbol{x} \in \mathcal{X}$.

Then, we decompose BCR as follows:

$$\begin{aligned}
\mathrm{BCR}_T &= \sum_{t=1}^T \mathbb{E}\left[f(\boldsymbol{x}^*) - f([\boldsymbol{x}^*]_t) + f([\boldsymbol{x}^*]_t) - g_t([\boldsymbol{x}_t]_t) + g_t([\boldsymbol{x}_t]_t) - f([\boldsymbol{x}_t]_t) + f([\boldsymbol{x}_t]_t) - f(\boldsymbol{x}_t)\right] \\
&= \underbrace{\mathbb{E}\left[\sum_{t=1}^T f(\boldsymbol{x}^*) - f([\boldsymbol{x}^*]_t)\right]}_{A_1} + \underbrace{\mathbb{E}\left[\sum_{t=1}^T f([\boldsymbol{x}^*]_t) - g_t([\boldsymbol{x}_t]_t)\right]}_{A_2} \\
&\quad + \underbrace{\mathbb{E}\left[\sum_{t=1}^T g_t([\boldsymbol{x}_t]_t) - f([\boldsymbol{x}_t]_t)\right]}_{A_3} + \underbrace{\mathbb{E}\left[\sum_{t=1}^T f([\boldsymbol{x}_t]_t) - f(\boldsymbol{x}_t)\right]}_{A_4}.
\end{aligned}$$

Then, using Lemma D.3, terms A_1 and A_4 can be bounded above by $\pi^2/6$, respectively. Furthermore, from the definition, $g_t([\boldsymbol{x}_t]_t)|\mathcal{D}_{t-1}$ and $f([\boldsymbol{x}^*]_t)|\mathcal{D}_{t-1}$ are identically distributed. Hence, we can see that

$$A_2 = \sum_{t=1}^T \mathbb{E}_{\mathcal{D}_{t-1}}\left[\mathbb{E}_t\left[f([\boldsymbol{x}^*]_t) - g_t([\boldsymbol{x}_t]_t)\right]\right] = 0.$$

Finally, we need to bound A_3. As with Lemma 3.2, for all $t \geq 0$, we can obtain

$$\mathbb{E}\left[\eta_t^2 \mathbb{1}\{\eta_t \geq 0\}\right] \leq 2 + 2\log\bigl(|\mathcal{X}_t|/2\bigr),$$

where $\eta_t \coloneqq \frac{g_t([\boldsymbol{x}_t]_t) - \mu_{t-1}([\boldsymbol{x}_t]_t)}{\sigma_{t-1}([\boldsymbol{x}_t]_t)}$. Note that the proof for Lemma 3.2 can apply to any distribution over $\mathcal{X}_t$ including $\Pr_t([\boldsymbol{x}_t]_t)$, not only $\Pr_t(\boldsymbol{x}_t)$. Therefore, we can obtain

$$\begin{aligned}
A_3 &= \mathbb{E}\left[\sum_{t=1}^{T} \eta_t \sigma_{t-1}([\boldsymbol{x}_t]_t)\right] \\
&\leq \mathbb{E}\left[\sum_{t=1}^{T} \eta_t \mathbb{1}\{\eta_t \geq 0\} \sigma_{t-1}([\boldsymbol{x}_t]_t)\right] \\
&\leq \underbrace{\mathbb{E}\left[\sum_{t=1}^{T} \eta_t \mathbb{1}\{\eta_t \geq 0\} \sigma_{t-1}(\boldsymbol{x}_t)\right]}_{A_5} + \underbrace{\mathbb{E}\left[\sum_{t=1}^{T} \eta_t \mathbb{1}\{\eta_t \geq 0\} |\sigma_{t-1}([\boldsymbol{x}_t]_t) - \sigma_{t-1}(\boldsymbol{x}_t)|\right]}_{A_6}.
\end{aligned}$$

Then, as with Theorem 3.1, we can see that

$$A_5 \leq \mathbb{E}\left[\sqrt{\sum_{t=1}^{T} \eta_t^2 \mathbb{1}\{\eta_t \geq 0\} \sum_{t=1}^{T} \sigma_{t-1}^2(\boldsymbol{x}_t)}\right] \leq \sqrt{C_1 \gamma_T T s_T},$$

where $s_T = 2 + 2\log\left(|\mathcal{X}_t|/2\right) = 2 - \log 2 + 2d\log(\lceil drLt^2 \rceil)$. Furthermore, we can obtain

$$\begin{aligned}
A_6 &\leq \mathbb{E}\left[\sum_{t=1}^{T} \eta_t \mathbb{1}\{\eta_t \geq 0\} \frac{1}{t^2}\right] && \bigl(\because \text{Lemma D.4 based on Lemma B.1}\bigr) \\
&= \sum_{t=1}^{T} \mathbb{E}[\eta_t \mathbb{1}\{\eta_t \geq 0\}] \frac{1}{t^2} \\
&\leq \sum_{t=1}^{T} \sqrt{\mathbb{E}[\eta_t^2 \mathbb{1}\{\eta_t \geq 0\}]} \frac{1}{t^2} && \bigl(\because \text{Jensen's inequality}\bigr) \\
&\leq \sum_{t=1}^{T} \sqrt{2 + 2\log\bigl(|\mathcal{X}_t|/2\bigr)} \frac{1}{t^2} \\
&\leq \sqrt{2 + 2\log\bigl(|\mathcal{X}_T|/2\bigr)} \sum_{t=1}^{T} \frac{1}{t^2} && \bigl(\because |\mathcal{X}_t| \text{ is monotonically increasing}\bigr) \\
&\leq \frac{\pi^2}{6} \sqrt{s_T}
\end{aligned}$$

which concludes the proof. □

B.3. Continuous Domain with Discretized Sample Path

If we modify TS so that $[\boldsymbol{x}_t]_t$ is evaluated, the following theorem holds:

Theorem B.1. *Let $f \sim \mathcal{GP}(0, k)$, where k is a stationary kernel, $k(\boldsymbol{x}, \boldsymbol{x}) = 1$, and Assumption 2.1 holds. Let $L = b\bigl(\sqrt{\log(ad)} + \sqrt{\pi}/2\bigr)$. Then, by running modified TS that evaluates $[\boldsymbol{x}_t]_t$, BCR can be bounded as follows:*

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

where $C_1 \coloneqq 2/\log(1 + \sigma^{-2})$ and $s_t = 2 - 2\log 2 + 2d\log(\lceil drLt^2 \rceil)$.

Proof. We can decompose BCR as follows:

$$\begin{aligned}\mathrm{BCR}_T &= \sum_{t=1}^{T} \mathbb{E}\left[f(\boldsymbol{x}^*) - f([\boldsymbol{x}^*]_t) + f([\boldsymbol{x}^*]_t) - g_t([\boldsymbol{x}_t]_t) + g_t([\boldsymbol{x}_t]_t) - f([\boldsymbol{x}_t]_t) \right] \\ &= \underbrace{\mathbb{E}\left[\sum_{t=1}^{T} f(\boldsymbol{x}^*) - f([\boldsymbol{x}^*]_t) \right]}_{A_1} + \underbrace{\mathbb{E}\left[\sum_{t=1}^{T} f([\boldsymbol{x}^*]_t) - g_t([\boldsymbol{x}_t]_t) \right]}_{A_2} + \underbrace{\mathbb{E}\left[\sum_{t=1}^{T} g_t([\boldsymbol{x}_t]_t) - f([\boldsymbol{x}_t]_t) \right]}_{A_3}.\end{aligned}$$

Then, as with the proof of Theorem 3.2, we can obtain $A_1 \leq \frac{\pi^2}{6}$ and $A_2 = 0$. Furthermore, we can obtain

$$A_3 \leq \mathbb{E}\left[\sum_{t=1}^{T} \eta_t \mathbb{1}\{\eta_t \geq 0\} \sigma_{t-1}([\boldsymbol{x}_t]_t) \right].$$

Since modified TS evaluates $[\boldsymbol{x}_t]_t$, we can directly bound A_3 using MIG and $L = b\big(\sqrt{\log(ad)} + \sqrt{\pi}/2\big)$, i.e.,

$$A_3 \leq \sqrt{C_1 \gamma_T T s_T}.$$

□

C. Proof regarding PIMS

First, we show the following lemma:

Lemma 4.1. *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. Let $\xi_t := \min_{\boldsymbol{x} \in \mathcal{X}} \frac{g_t^* - \mu_{t-1}(\boldsymbol{x})}{\sigma_{t-1}(\boldsymbol{x})}$. Then, the following inequality holds:*

$$\mathbb{E}\left[\xi_t^2 \mathbb{1}\{\xi_t \geq 0\} \right] \leq 2 + 2\log\big(|\mathcal{X}|/2\big),$$

for all $t \geq 1$.

Proof. From the property of non-negative random variable,

$$\mathbb{E}\Big[\xi_t^2 \mathbb{1}\{\xi_t \geq 0\} \Big] = \int_0^\infty \Pr\Big(\xi_t^2 \mathbb{1}\{\xi_t \geq 0\} > c \Big) \mathrm{d}c. \tag{1}$$

Using Lemmas D.1, D.5, and union bound, for all $\mathcal{D}_{t-1}$ and $c \geq 0$,

$$\begin{aligned}\Pr_t\Big(\xi_t^2 \mathbb{1}\{\xi_t \geq 0\} > c \Big) &= \Pr_t\Big(\xi_t > \sqrt{c} \Big) \\ &= \Pr_t\big(g_t^* > \max_{\boldsymbol{x} \in \mathcal{X}} \big\{ \mu_{t-1}(\boldsymbol{x}) + \sqrt{c}\sigma_{t-1}(\boldsymbol{x}) \big\} \big) && (\because \text{Lemma D.5}) \\ &= \Pr_t\big(\exists \boldsymbol{x}, g_t(\boldsymbol{x}) > \max_{\boldsymbol{x} \in \mathcal{X}} \big\{ \mu_{t-1}(\boldsymbol{x}) + \sqrt{c}\sigma_{t-1}(\boldsymbol{x}) \big\} \big) \\ &= \Pr_t\big(\exists \boldsymbol{x}, f(\boldsymbol{x}) > \max_{\boldsymbol{x} \in \mathcal{X}} \big\{ \mu_{t-1}(\boldsymbol{x}) + \sqrt{c}\sigma_{t-1}(\boldsymbol{x}) \big\} \big) \\ &\leq \Pr_t\big(\exists \boldsymbol{x}, f(\boldsymbol{x}) > \mu_{t-1}(\boldsymbol{x}) + \sqrt{c}\sigma_{t-1}(\boldsymbol{x}) \big) \\ &\leq \sum_{\boldsymbol{x} \in \mathcal{X}} \Pr_t\big(f(\boldsymbol{x}) > \mu_{t-1}(\boldsymbol{x}) + \sqrt{c}\sigma_{t-1}(\boldsymbol{x}) \big) && (\because \text{union bound}) \\ &\leq \frac{|\mathcal{X}|}{2} e^{-c/2}. && (\because \text{Lemma D.1})\end{aligned}$$

Therefore, we obtain

$$\Pr(\xi_t^2 \mathbb{1}\{\xi_t \geq 0\} > c) = \mathbb{E}_{\mathcal{D}_{t-1}}\big[\Pr_t(\xi_t^2 \mathbb{1}\{\xi_t \geq 0\} > c) \big] \leq \frac{|\mathcal{X}|}{2} e^{-c/2}.$$

Since $\Pr\Big(\xi_t^2 \mathbb{1}\{\xi_t \geq 0\} > c\Big) \leq \min\{1, \frac{|\mathcal{X}|}{2} e^{-c/2}\}$ due to $\Pr\Big(\xi_t^2 \mathbb{1}\{\xi_t \geq 0\} > c\Big) \leq 1$, we obtain

$$
\begin{aligned}
(1) &\leq \int_0^{2\log(|\mathcal{X}|/2)} 1 \mathrm{d}c + \int_{2\log(|\mathcal{X}|/2)}^{\infty} \frac{|\mathcal{X}|}{2} e^{-c/2} \mathrm{d}c \\
&= 2\log(|\mathcal{X}|/2) + \frac{|\mathcal{X}|}{2} \int_0^{\infty} e^{-(c+2\log(|\mathcal{X}|/2))/2} \mathrm{d}c \\
&= 2\log(|\mathcal{X}|/2) + \int_0^{\infty} e^{-c/2} \mathrm{d}c \\
&= 2\log(|\mathcal{X}|/2) + 2,
\end{aligned}
$$

which concludes the proof. □

C.1. Discrete Domain

For the discrete input domain, using Lemma 4.1, we show the following theorem:

Theorem 4.1. *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. Then, by running PIMS, 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 $C_2 \coloneqq 2 + 2\log\big(|\mathcal{X}|/2\big)$.

Proof. Since $\boldsymbol{x}_t \perp\!\!\!\perp f | \mathcal{D}_{t-1}$ and $\boldsymbol{x}_t | \mathcal{D}_{t-1}$ depends on only g_t^*, we can transform BCR as follows:

$$
\begin{aligned}
\mathrm{BCR}_T &= \sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}_{t-1}} \Big[\mathbb{E}_t \big[f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t) \big] \Big] \\
&= \sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}_{t-1}} \Big[\mathbb{E}_t \big[f(\boldsymbol{x}^*) - g_t^* + g_t^* - f(\boldsymbol{x}_t) \big] \Big] \\
&= \sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}_{t-1}} \Big[\mathbb{E}_t \big[g_t^* - f(\boldsymbol{x}_t) \big] \Big] && \big(\because \mathbb{E}_t[f(\boldsymbol{x}^*)] = \mathbb{E}_t[g_t^*]\big) \\
&= \sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}_{t-1}} \Big[\mathbb{E}_t \big[g_t^* - \mu_{t-1}(\boldsymbol{x}_t) \big] \Big] && \big(\because f \perp\!\!\!\perp \boldsymbol{x}_t | \mathcal{D}_{t-1}\big) \\
&= \sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}_{t-1}} \Big[\mathbb{E}_t \Big[\frac{g_t^* - \mu_{t-1}(\boldsymbol{x}_t)}{\sigma_{t-1}(\boldsymbol{x}_t)} \sigma_{t-1}(\boldsymbol{x}_t) \Big] \Big].
\end{aligned}
$$

Then, by defining $\xi_t \coloneqq \frac{g_t^* - \mu_{t-1}(\boldsymbol{x}_t)}{\sigma_{t-1}(\boldsymbol{x}_t)}$, we can obtain

$$
\begin{aligned}
\mathrm{BCR}_T &= \sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}_{t-1}} \Big[\mathbb{E}_t \Big[\xi_t \sigma_{t-1}(\boldsymbol{x}_t) \Big] \Big] \\
&= \mathbb{E} \Big[\sum_{t=1}^{T} \xi_t \sigma_{t-1}(\boldsymbol{x}_t) \Big] \\
&\leq \mathbb{E} \Big[\sum_{t=1}^{T} \xi_t \mathbb{1}\{\xi_t \geq 0\} \sigma_{t-1}(\boldsymbol{x}_t) \Big] \\
&\leq \mathbb{E} \Big[\sqrt{\sum_{t=1}^{T} \xi_t^2 \mathbb{1}\{\xi_t \geq 0\} \sum_{t=1}^{T} \sigma_{t-1}^2(\boldsymbol{x}_t)} \Big]. && \big(\because \text{Cauchy-Schwarz inequality}\big)
\end{aligned}
$$

We can use the following well-known bound of MIG:

$$\sum_{t=1}^{T} \sigma_{t-1}^2(\boldsymbol{x}_t) \leq C_1 \gamma_T, \qquad (\because \text{ Lemma 5.4 in (Srinivas et al., 2010)})$$

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

$$\begin{aligned}
\mathrm{BCR}_T &\leq \mathbb{E}\left[\sqrt{\sum_{t=1}^{T} \xi_t^2 \mathbb{1}\{\xi_t \geq 0\}}\right] \sqrt{C_1 \gamma_T} \\
&\leq \sqrt{\mathbb{E}\Big[\sum_{t=1}^{T} \xi_t^2 \mathbb{1}\{\xi_t \geq 0\}\Big]} \sqrt{C_1 \gamma_T} \qquad (\because \text{ Jensen inequality}) \\
&= \sqrt{\sum_{t=1}^{T} \mathbb{E}\Big[\xi_t^2 \mathbb{1}\{\xi_t \geq 0\}\Big]} \sqrt{C_1 \gamma_T}.
\end{aligned}$$

We can apply Lemma 4.1 to ξ_t for all $t \geq 1$. Thus, we obtain the following upper bound of BCR:

$$\begin{aligned}
\mathrm{BCR}_T &\leq \sqrt{\sum_{t=1}^{T} \mathbb{E}\Big[\xi_t^2 \mathbb{1}\{\xi_t(\boldsymbol{x}_t) \geq 0\}\Big]} \sqrt{C_1 \gamma_T} \\
&\leq \sqrt{C_1 C_2 T \gamma_T},
\end{aligned}$$

where $C_2 \coloneqq 2 + 2\log(|\mathcal{X}|/2)$. □

C.2. Continuous Domain with Continuous Sample Path

Theorem 4.2. *Let $f \sim \mathcal{GP}(0, k)$, where k is a stationary kernel, $k(\boldsymbol{x}, \boldsymbol{x}) = 1$, and Assumption 2.1 holds. Then, by running PIMS, BCR can be bounded as follows:*

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

where $C_1 \coloneqq 2/\log(1+\sigma^{-2})$ and $m_t \coloneqq 2d \log\left(\left\lceil t^2 bdr\big(\log(ad) + \sqrt{\pi}/2\big)\sqrt{(\sigma^2+n_t)/\sigma^2}\right\rceil\right) - 2\log 2 + 2$.

Proof. 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 $\tau_t = \lceil t^2 bdr\big(\log(ad) + \sqrt{\pi}/2\big)\sqrt{(\sigma^2+n_t)/\sigma^2}\rceil$. Thus, $|\mathcal{X}_t| = \tau_t^d$. In addition, we define $[\boldsymbol{x}]_t$ as the nearest point in $\mathcal{X}_t$ of $\boldsymbol{x} \in \mathcal{X}$. Note that the discretization does not depend on any randomness and is deterministic.

As with Theorem 4.1, we obtain

$$\mathrm{BCR}_T = \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{E}_t\left[\frac{g_t^* - \mu_{t-1}(\boldsymbol{x}_t)}{\sigma_{t-1}(\boldsymbol{x}_t)} \sigma_{t-1}(\boldsymbol{x}_t)\right]\right].$$

Let $\tilde{g}_t^* \coloneqq \max_{\boldsymbol{x} \in \mathcal{X}_t} g_t(\boldsymbol{x})$, $\tilde{\boldsymbol{x}}_t \coloneqq \arg\min_{\boldsymbol{x} \in \mathcal{X}_t} \{(\tilde{g}_t^* - \mu_{t-1}(\boldsymbol{x}))/\sigma_{t-1}(\boldsymbol{x})\}$, and $\boldsymbol{z}_t^* \coloneqq \arg\max_{\boldsymbol{x} \in \mathcal{X}} g_t(\boldsymbol{x})$. Then, we can

obtain

$$
\begin{aligned}
\mathrm{BCR}_T &\le \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{E}_t\left[\frac{g_t^* - \mu_{t-1}(\tilde{\boldsymbol{x}}_t)}{\sigma_{t-1}(\tilde{\boldsymbol{x}}_t)}\sigma_{t-1}(\boldsymbol{x}_t)\right]\right] && (\because \boldsymbol{x}_t \text{ is minimum}) \\
&= \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{E}_t\left[\frac{g_t^* - g([\boldsymbol{z}_t^*]_t) + g([\boldsymbol{z}_t^*]_t) - \tilde{g}_t^* + \tilde{g}_t^* - \mu_{t-1}(\tilde{\boldsymbol{x}}_t)}{\sigma_{t-1}(\tilde{\boldsymbol{x}}_t)}\sigma_{t-1}(\boldsymbol{x}_t)\right]\right] \\
&\le \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{E}_t\left[\frac{g_t^* - g([\boldsymbol{z}_t^*]_t) + \tilde{g}_t^* - \mu_{t-1}(\tilde{\boldsymbol{x}}_t)}{\sigma_{t-1}(\tilde{\boldsymbol{x}}_t)}\sigma_{t-1}(\boldsymbol{x}_t)\right]\right] && (\because g([\boldsymbol{z}_t^*]_t) \le \tilde{g}_t^*) \\
&= \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{E}_t\left[\frac{g_t^* - g([\boldsymbol{z}_t^*]_t)}{\sigma_{t-1}(\tilde{\boldsymbol{x}}_t)}\sigma_{t-1}(\boldsymbol{x}_t) + \frac{\tilde{g}_t^* - \mu_{t-1}(\tilde{\boldsymbol{x}}_t)}{\sigma_{t-1}(\tilde{\boldsymbol{x}}_t)}\sigma_{t-1}(\boldsymbol{x}_t)\right]\right] \\
&= \underbrace{\sum_{t=1}^{T} \mathbb{E}\left[\frac{g_t^* - g([\boldsymbol{z}_t^*]_t)}{\sigma_{t-1}(\tilde{\boldsymbol{x}}_t)}\sigma_{t-1}(\boldsymbol{x}_t)\right]}_{=:B_1} + \underbrace{\sum_{t=1}^{T} \mathbb{E}\left[\frac{\tilde{g}_t^* - \mu_{t-1}(\tilde{\boldsymbol{x}}_t)}{\sigma_{t-1}(\tilde{\boldsymbol{x}}_t)}\sigma_{t-1}(\boldsymbol{x}_t)\right]}_{=:B_2}
\end{aligned}
$$

For B_2, as with the proof of Theorem 4.1,

$$
B_2 \le \sqrt{C_1 \gamma_T}\sqrt{\sum_{t=1}^{T} \mathbb{E}\left[\tilde{\xi}_t^2 \mathbb{1}\{\tilde{\xi}_t \ge 0\}\right]},
$$

where

$$
\tilde{\xi}_t := \frac{\tilde{g}_t^* - \mu_{t-1}(\tilde{\boldsymbol{x}}_t)}{\sigma_{t-1}(\tilde{\boldsymbol{x}}_t)} = \min_{\boldsymbol{x} \in \mathcal{X}_t} \frac{\tilde{g}_t^* - \mu_{t-1}(\boldsymbol{x})}{\sigma_{t-1}(\boldsymbol{x})}.
$$

Hence, by replacing $\mathcal{X}$ with $\mathcal{X}_t$ compared with Theorem 4.1, we can apply Lemma 4.1 as follows:

$$
\begin{aligned}
\mathbb{E}\left[\tilde{\xi}_t^2(\tilde{\boldsymbol{x}}_t)\mathbb{1}\{\tilde{\xi}_t(\tilde{\boldsymbol{x}}_t) \ge 0\}\right] &\le 2\log(|\mathcal{X}_t|/2) + 2 \\
&= 2d\log\Big(\Big\lceil t^2 bdr\left(\log(ad) + \sqrt{\pi}/2\right)\sqrt{(\sigma^2 + n_t)/\sigma^2}\Big\rceil\Big) - 2\log 2 + 2 \\
&= m_t.
\end{aligned}
$$

Therefore, we obtain $B_2 \le \sqrt{C_1 T \gamma_T m_T}$.

Next, for B_1,

$$
\begin{aligned}
B_1 &= \sum_{t=1}^{T} \mathbb{E}\left[\frac{g_t^* - g([\boldsymbol{z}_t^*]_t)}{\sigma_{t-1}(\tilde{\boldsymbol{x}}_t)}\sigma_{t-1}(\boldsymbol{x}_t)\right] \\
&\le \sum_{t=1}^{T} \mathbb{E}\left[\frac{g_t^* - g([\boldsymbol{z}_t^*]_t)}{\sigma_{t-1}(\tilde{\boldsymbol{x}}_t)}\right] && (\because \sigma_{t-1}(\boldsymbol{x}_t) \le 1) \\
&\le \sum_{t=1}^{T} \mathbb{E}\left[\left(g_t^* - g([\boldsymbol{z}_t^*]_t)\right)\sqrt{\frac{\sigma^2 + n_t}{\sigma^2}}\right] && (\because \text{Lemma 4.2}) \\
&\le \sum_{t=1}^{T} \frac{1}{t^2} && (\because \text{Lemma D.3}) \\
&\le \frac{\pi^2}{6}.
\end{aligned}
$$

Consequently, we obtain the desired result. □

Finally, we show the BSR bound:

Theorem 4.3. *Assume the same condition as in Theorem 4.2. Then, by running PIMS, BSR can be bounded as follows:*

$$\mathrm{BSR}_T \leq \frac{\pi^2}{6T} + \sqrt{\frac{C_1 \gamma_T m_T}{T}},$$

where $C_1 \coloneqq 2/\log(1+\sigma^{-2})$ *and* $m_t \coloneqq 2d \log\left(\lceil t^2 bdr\big(\log(ad) + \sqrt{\pi}/2\big)\rceil\right) - 2\log 2 + 2$.

Proof. 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 $\tau_t = \lceil t^2 bdr\big(\log(ad) + \sqrt{\pi}/2\big)\rceil$. Thus, $|\mathcal{X}_t| = \tau_t^d$. In addition, we define $[\boldsymbol{x}]_t$ as the nearest point in $\mathcal{X}_t$ of $\boldsymbol{x} \in \mathcal{X}$. Note that the discretization does not depend on any randomness and is deterministic.

Let $\tilde{g}_t^* \coloneqq \max_{\boldsymbol{x}\in\mathcal{X}_t} g_t(\boldsymbol{x})$, $\tilde{\boldsymbol{x}}_t \coloneqq \arg\min_{\boldsymbol{x}\in\mathcal{X}_t} \{(\tilde{g}_t^* - \mu_{t-1}(\boldsymbol{x}))/\sigma_{t-1}(\boldsymbol{x})\}$, and $\boldsymbol{z}_t^* \coloneqq \arg\max_{\boldsymbol{x}\in\mathcal{X}} g_t(\boldsymbol{x})$. From Lemma 2.1, BSR can be transformed as follows:

$$\begin{aligned}
\mathrm{BSR}_T &\leq \frac{1}{T}\sum_{t=1}^{T} \mathrm{BSR}_t \\
&= \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\left[f(\boldsymbol{x}^*) - f(\hat{\boldsymbol{x}}_t)\right] \\
&= \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\left[\mathbb{E}_t\left[f(\boldsymbol{x}^*)\right] - \mu_{t-1}(\hat{\boldsymbol{x}}_t)\right] \\
&= \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\left[f(\boldsymbol{x}^*) - f([\boldsymbol{x}^*]_t) + f([\boldsymbol{x}^*]_t) - g_t([\boldsymbol{z}_t^*]_t) + g_t([\boldsymbol{z}_t^*]_t) - \tilde{g}_t^* + \tilde{g}_t^* - \mu_{t-1}(\hat{\boldsymbol{x}}_t)\right] \\
&\leq \frac{\pi^2}{6T} + \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\left[\tilde{g}_t^* - \mu_{t-1}(\hat{\boldsymbol{x}}_t)\right],
\end{aligned}$$

where the final inequality can be obtained by the same proof of Theorem 3.2 and $g_t([\boldsymbol{z}_t^*]_t) \leq \tilde{g}_t^*$.

Then, we show the inequality between $\sigma_{t-1}(\boldsymbol{x}_t)$ and $\sigma_{t-1}(\tilde{\boldsymbol{x}}_t)$. From the definition of $\boldsymbol{x}_t$ and $\tilde{\boldsymbol{x}}_t$,

$$\frac{\tilde{g}_t^* - \mu_{t-1}(\boldsymbol{x}_t)}{\sigma_{t-1}(\boldsymbol{x}_t)} + \frac{g_t^* - \tilde{g}_t^*}{\sigma_{t-1}(\boldsymbol{x}_t)} = \frac{g_t^* - \mu_{t-1}(\boldsymbol{x}_t)}{\sigma_{t-1}(\boldsymbol{x}_t)} \leq \frac{g_t^* - \mu_{t-1}(\tilde{\boldsymbol{x}}_t)}{\sigma_{t-1}(\tilde{\boldsymbol{x}}_t)} = \frac{\tilde{g}_t^* - \mu_{t-1}(\tilde{\boldsymbol{x}}_t)}{\sigma_{t-1}(\tilde{\boldsymbol{x}}_t)} + \frac{g_t^* - \tilde{g}_t^*}{\sigma_{t-1}(\tilde{\boldsymbol{x}}_t)},$$

and

$$\frac{\tilde{g}_t^* - \mu_{t-1}(\tilde{\boldsymbol{x}}_t)}{\sigma_{t-1}(\tilde{\boldsymbol{x}}_t)} \leq \frac{\tilde{g}_t^* - \mu_{t-1}(\boldsymbol{x}_t)}{\sigma_{t-1}(\boldsymbol{x}_t)}.$$

Therefore, by combining the above two inequalities and $g_t^* \geq \tilde{g}_t^*$, we see that

$$\sigma_{t-1}(\tilde{\boldsymbol{x}}_t) \leq \sigma_{t-1}(\boldsymbol{x}_t).$$

Hence, we can bound the remained term as follows:

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\left[\tilde{g}_t^* - \mu_{t-1}(\hat{\boldsymbol{x}}_t)\right] &= \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\left[\frac{\tilde{g}_t^* - \mu_{t-1}(\hat{\boldsymbol{x}}_t)}{\sigma_{t-1}(\boldsymbol{x}_t)}\sigma_{t-1}(\boldsymbol{x}_t)\right] \\
&\leq \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\left[\frac{\tilde{g}_t^* - \mu_{t-1}(\hat{\boldsymbol{x}}_t)}{\sigma_{t-1}(\boldsymbol{x}_t)}\mathbb{1}\{\tilde{g}_t^* - \mu_{t-1}(\hat{\boldsymbol{x}}_t)\}\sigma_{t-1}(\boldsymbol{x}_t)\right] \\
&\leq \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\left[\frac{\tilde{g}_t^* - \mu_{t-1}(\hat{\boldsymbol{x}}_t)}{\sigma_{t-1}(\tilde{\boldsymbol{x}}_t)}\mathbb{1}\{\tilde{g}_t^* - \mu_{t-1}(\hat{\boldsymbol{x}}_t)\}\sigma_{t-1}(\boldsymbol{x}_t)\right] \qquad \left(\because \sigma_{t-1}(\tilde{\boldsymbol{x}}_t) \leq \sigma_{t-1}(\boldsymbol{x}_t)\right) \\
&\leq \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\left[\frac{\tilde{g}_t^* - \mu_{t-1}(\tilde{\boldsymbol{x}}_t)}{\sigma_{t-1}(\tilde{\boldsymbol{x}}_t)}\mathbb{1}\{\tilde{g}_t^* - \mu_{t-1}(\tilde{\boldsymbol{x}}_t)\}\sigma_{t-1}(\boldsymbol{x}_t)\right] \qquad \left(\because \mu_{t-1}(\tilde{\boldsymbol{x}}_t) \leq \mu_{t-1}(\hat{\boldsymbol{x}}_t)\right) \\
&\leq \sqrt{\frac{C_1 \gamma_T m_T}{T}},
\end{aligned}$$

where the final inequality can be obtained as with B_2 in the proof of Theorem 4.2. □

C.3. Continuous Domain with Discretized Sample Path

Theorem C.1. *Let $f \sim \mathcal{GP}(0, k)$, where k is a stationary kernel, $k(\boldsymbol{x}, \boldsymbol{x}) = 1$, and Assumption 2.1 holds. Then, by running PIMS that uses $\tilde{g}_t^*$ instead of g_t^*, i.e., evaluates $\tilde{\boldsymbol{x}}_t = \arg\min_{\boldsymbol{x} \in \mathcal{X}_t} (\tilde{g}_t^* - \mu_{t-1}(\boldsymbol{x}))/\sigma_{t-1}(\boldsymbol{x})$, BCR can be bounded as follows:*

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

where $C_1 \coloneqq 2/\log(1+\sigma^{-2})$ and $m_t \coloneqq 2d \log\left(\lceil t^2 bdr\big(\log(ad) + \sqrt{\pi}/2\big)\rceil\right) - 2\log 2 + 2$.

Proof. We can transform the BCR as follows:

$$\begin{aligned}
\mathrm{BCR}_T &= \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{E}_t\left[g_t^* - g([\boldsymbol{z}_t^*]_t) + g([\boldsymbol{z}_t^*]_t) - \tilde{g}_t^* + \tilde{g}_t^* - \mu_{t-1}(\tilde{\boldsymbol{x}}_t)\right]\right] \\
&\leq \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{E}_t\left[g_t^* - g([\boldsymbol{z}_t^*]_t) + \tilde{g}_t^* - \mu_{t-1}(\tilde{\boldsymbol{x}}_t)\right]\right] \qquad \left(\because \tilde{g}_t^* \geq g([\boldsymbol{z}_t^*]_t)\right) \\
&= \sum_{t=1}^{T} \mathbb{E}\left[g_t^* - g([\boldsymbol{z}_t^*]_t)\right] + \sum_{t=1}^{T} \mathbb{E}\left[\frac{\tilde{g}_t^* - \mu_{t-1}(\tilde{\boldsymbol{x}}_t)}{\sigma_{t-1}(\tilde{\boldsymbol{x}}_t)} \sigma_{t-1}(\tilde{\boldsymbol{x}}_t)\right],
\end{aligned}$$

where $\boldsymbol{z}_t^* = \arg\max_{\boldsymbol{x} \in \mathcal{X}} g_t(\boldsymbol{x})$. Then, the first term can be bounded above by $\pi^2/6$ using Lemma D.3. Furthermore, the second term can be bounded above by $\sqrt{C_1 T \gamma_T m_T}$ as with the proof of Theorem 4.1. □

D. Auxiliary Lemmas

We used the following useful lemmas:

Lemma D.1 (in Lemma 5.2 of (Srinivas et al., 2010) and Lemma H.3 of (Takeno et al., 2023))**.** *For $c > 0$, the survival function of the standard normal distribution can be bounded above as follows:*

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

Lemma D.2 (Lemma H.1 of (Takeno et al., 2023))**.** *Let $f \sim \mathcal{GP}(0, k)$ and Assumption 2.1 holds. Let the supremum of the partial derivatives $L_{\max} \coloneqq \sup_{j \in [d]} \sup_{\boldsymbol{x} \in \mathcal{X}} \left|\frac{\partial f}{\partial x_j}\right|$. Then, $\mathbb{E}[L_{\max}]$ can be bounded above as follows:*

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

Lemma D.3 (Lemma H.2 of (Takeno et al., 2023))**.** *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 $\tau_t \geq bdru_t\big(\sqrt{\log(ad)} + \sqrt{\pi}/2\big)$ 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 $\boldsymbol{x} \in \mathcal{X}$.

Similarly, we can obtain the following lemma:

Lemma D.4. *Let $f \sim \mathcal{GP}(0, k)$ and assume the same condition as in Lemma B.1. Let $\mathcal{X}_t \subset \mathcal{X}$ be a finite set with each dimension equally divided into $\tau_t \geq drL_\sigma u_t$ for any $t \geq 1$. Then, the following inequality holds with probability 1:*

$$\forall t \geq 1, \forall \mathcal{D}_{t-1}, \sup_{\boldsymbol{x} \in \mathcal{X}} |\sigma_{t-1}(\boldsymbol{x}) - \sigma_{t-1}([\boldsymbol{x}]_t)| \leq \frac{1}{u_t},$$

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:

$$\begin{aligned} \sup_{\boldsymbol{x} \in \mathcal{X}} \|\boldsymbol{x} - [\boldsymbol{x}]_t\|_1 &\leq \frac{dr}{drL_\sigma u_t} \\ &= \frac{1}{L_\sigma u_t}. \end{aligned} \tag{2}$$

Note that since this discretization is fixed beforehand, the discretization does not depend on any randomness.

Then, we obtain the following:

$$\begin{aligned} \mathbb{E}\left[\sup_{\boldsymbol{x} \in \mathcal{X}} |\sigma_{t-1}(\boldsymbol{x}) - \sigma_{t-1}([\boldsymbol{x}]_t)|\right] &\leq L_\sigma \sup_{\boldsymbol{x} \in \mathcal{X}} \|\boldsymbol{x} - [\boldsymbol{x}]_t\|_1 && (\because \text{Lemma B.1}) \\ &\leq L_\sigma \frac{1}{L_\sigma u_t} && (\because \text{Eq. (2)}) \\ &\leq \frac{1}{u_t}. && (\because \text{Lemma D.2}) \end{aligned}$$

□

For BSR, based on the discussion in Appendix A of Takeno et al. (2023), we show the following lemma:

Lemma 2.1. *BSR can be bounded from above as follows:*

$$\overline{\mathrm{BSR}}_T \leq \sum_{t=1}^{T} \overline{\mathrm{BSR}}_t / T \leq \mathrm{BCR}_T / T.$$

and

$$\mathrm{BSR}_T \leq \mathrm{BCR}_T / T$$

Proof. For the first inequality, we can bound the modified BSR from above as follows:

$$\begin{aligned} \overline{\mathrm{BSR}}_T &= \mathbb{E}\left[f(\boldsymbol{x}^*) - f(\hat{\boldsymbol{x}}_T)\right] \\ &= \mathbb{E}_{\mathcal{D}_{T-1}}\left[\mathbb{E}_T\left[f(\boldsymbol{x}^*) - f(\hat{\boldsymbol{x}}_T)\right]\right] \\ &= \mathbb{E}_{\mathcal{D}_{T-1}}\left[\mathbb{E}_T\left[f(\boldsymbol{x}^*)\right] - \mu_{T-1}(\hat{\boldsymbol{x}}_T)\right] \\ &\leq \mathbb{E}_{\mathcal{D}_{T-1}}\left[\mathbb{E}_T\left[f(\boldsymbol{x}^*)\right] - \frac{1}{T}\sum_{t=1}^{T} \mu_{T-1}(\hat{\boldsymbol{x}}_t)\right] && (\because \forall t \leq T, \mu_{T-1}(\hat{\boldsymbol{x}}_t) \leq \mu_{T-1}(\hat{\boldsymbol{x}}_T)) \\ &= \mathbb{E}_{\mathcal{D}_{T-1}}\left[\mathbb{E}_T\left[f(\boldsymbol{x}^*) - \frac{1}{T}\sum_{t=1}^{T} f(\hat{\boldsymbol{x}}_t)\right]\right] \\ &= \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\left[f(\boldsymbol{x}^*) - f(\hat{\boldsymbol{x}}_t)\right] \\ &= \frac{1}{T}\sum_{t=1}^{T} \overline{\mathrm{BSR}}_t. \end{aligned}$$

In addition, we can obtain

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\overline{\mathrm{BSR}}_t &= \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t) + f(\boldsymbol{x}_t) - f(\hat{\boldsymbol{x}}_t)\right] \\
&= \mathbb{E}_{\mathcal{D}_{t-1}}\left[\frac{1}{T}\sum_{t=1}^{T}\left\{\mathbb{E}_t\big[f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t)\big] + \mu_{t-1}(\boldsymbol{x}_t) - \mu_{t-1}(\hat{\boldsymbol{x}}_t)\right\}\right] \\
&\leq \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t)\right] && \big(\because \forall t \geq 1, \mu_{t-1}(\boldsymbol{x}_t) \leq \mu_{t-1}(\hat{\boldsymbol{x}}_t)\big) \\
&= \frac{1}{T}\mathrm{BCR}_T.
\end{aligned}$$

The inequality for BSR can be obtained from the relationship between average and the maximum $\max_{t\leq T} f(\boldsymbol{x}_t) \geq \sum_{t=1}^{T} f(\boldsymbol{x}_t)/T$. □

Lemma D.5. *Suppose that $\mathcal{X}$ is a discrete input domain. Let $\xi_t = \min_{\boldsymbol{x}\in\mathcal{X}} \frac{g_t^* - \mu_{t-1}(\boldsymbol{x})}{\sigma_{t-1}(\boldsymbol{x})} = \frac{g_t^* - \mu_{t-1}(\boldsymbol{x}_t)}{\sigma_{t-1}(\boldsymbol{x}_t)}$. Then, for all $t \in \mathbb{N}$ and $c \geq 0$,*

$$\Pr\big(\xi_t > c\big) = \Pr\big(g_t^* > \max_{\boldsymbol{x}\in\mathcal{X}}\left\{\mu_{t-1}(\boldsymbol{x}) + c\sigma_{t-1}(\boldsymbol{x})\right\}\big).$$

Proof. From Lemma A.1, $\boldsymbol{x}_t \in \arg\max_{\boldsymbol{x}\in\mathcal{X}}\left\{\mu_{t-1}(\boldsymbol{x}) + \frac{g_t^* - \mu_{t-1}(\boldsymbol{x}_t)}{\sigma_{t-1}(\boldsymbol{x}_t)}\sigma_{t-1}(\boldsymbol{x})\right\}$. Therefore, if $\xi_t > c$, i.e.,

$$\frac{g_t^* - \mu_{t-1}(\boldsymbol{x}_t)}{\sigma_{t-1}(\boldsymbol{x}_t)} > c,$$

then

$$\begin{aligned}
g_t^* &= \max_{\boldsymbol{x}\in\mathcal{X}}\left\{\mu_{t-1}(\boldsymbol{x}) + \frac{g_t^* - \mu_{t-1}(\boldsymbol{x}_t)}{\sigma_{t-1}(\boldsymbol{x}_t)}\sigma_{t-1}(\boldsymbol{x})\right\} \\
&> \max_{\boldsymbol{x}\in\mathcal{X}}\left\{\mu_{t-1}(\boldsymbol{x}) + c\sigma_{t-1}(\boldsymbol{x})\right\}.
\end{aligned}$$

Inversely, if

$$\begin{aligned}
g_t^* &= \max_{\boldsymbol{x}\in\mathcal{X}}\left\{\mu_{t-1}(\boldsymbol{x}) + \frac{g_t^* - \mu_{t-1}(\boldsymbol{x}_t)}{\sigma_{t-1}(\boldsymbol{x}_t)}\sigma_{t-1}(\boldsymbol{x})\right\} \\
&> \max_{\boldsymbol{x}\in\mathcal{X}}\left\{\mu_{t-1}(\boldsymbol{x}) + c\sigma_{t-1}(\boldsymbol{x})\right\}.
\end{aligned}$$

then

$$\frac{g_t^* - \mu_{t-1}(\boldsymbol{x}_t)}{\sigma_{t-1}(\boldsymbol{x}_t)} > c.$$

Hence, the statement of the lemma holds. □

Finally, we show the following lemma for completeness:

Lemma 4.2. *Let k be a kernel s.t. $k(\boldsymbol{x}, \boldsymbol{x}) = 1$. Then, the posterior variance is bounded from below as,*

$$\sigma_t^2(\boldsymbol{x}) \geq \frac{\sigma^2}{\sigma^2 + n_t},$$

for all $\boldsymbol{x} \in \mathcal{X}$ and for all $t \geq 0$.

Proof. When only one input $\boldsymbol{x}$ is repeatedly evaluated t times, $\sigma_t(\boldsymbol{x})$ become the minimum. For simplicity, let $\mathcal{D}_0 = \emptyset$ without loss of generality. Then, we see that

$$\sigma_0^2(\boldsymbol{x}) = 1,$$
$$\sigma_1^2(\boldsymbol{x}) = \frac{\sigma^2}{1+\sigma^2}.$$

Assume $\sigma_i^2(\boldsymbol{x}) = \frac{\sigma^2}{i+\sigma^2}$. Then, we can obtain

$$\begin{aligned}
\sigma_{i+1}^2(\boldsymbol{x}) &= \sigma_i^2(\boldsymbol{x}) - \frac{\left(\sigma_i^2(\boldsymbol{x})\right)^2}{\sigma_i^2(\boldsymbol{x}) + \sigma^2} \\
&= \frac{\sigma^2 \sigma_i^2(\boldsymbol{x})}{\sigma_i^2(\boldsymbol{x}) + \sigma^2} \\
&= \frac{\frac{\sigma^4}{i+\sigma^2}}{\frac{(i+1)\sigma^2+\sigma^4}{i+\sigma^2}} \\
&= \frac{\sigma^2}{(i+1)+\sigma^2}.
\end{aligned}$$

Since $\sigma_i^2(\boldsymbol{x}) = \frac{\sigma^2}{i+\sigma^2}$ holds when $i = 0$ and 1, the statement $\sigma_t^2(\boldsymbol{x}) = \frac{\sigma^2}{t+\sigma^2}$ holds for all $t \geq 0$. If $|\mathcal{D}_0| = n_0 > 0$, t changes to $n_t = n_0 + t$. □