Bayesian Optimization with Inexact Acquisition: Is Random Grid Search Sufficient?

Hwanwoo Kim[1] **Chong Liu**[2] **Yuxin Chen**[3]

[1]Department of Statistical Science, Duke University, Durham, NC, USA
[2]Department of Computer Science, University at Albany, State University of New York, Albany, NY, USA
[3]Department of Computer Science, University of Chicago, Chicago, IL, USA

Abstract

Bayesian optimization (BO) is a widely used iterative algorithm for optimizing black-box functions. Each iteration requires maximizing an acquisition function, such as the upper confidence bound (UCB) or a sample path from the Gaussian process (GP) posterior, as in Thompson sampling (TS). However, finding an exact solution to these maximization problems is often intractable and computationally expensive. Reflecting realistic scenarios, this paper investigates the effect of using inexact acquisition function maximizers in BO. Defining a measure of inaccuracy in acquisition solutions, we establish cumulative regret bounds for both GP-UCB and GP-TS based on imperfect acquisition function maximizers. Our results show that under appropriate conditions on accumulated inaccuracy, BO algorithms with inexact maximizers can still achieve sublinear cumulative regret. Motivated by such findings, we provide both theoretical justification and numerical validation for random grid search as an effective and computationally efficient acquisition function solver.

1 INTRODUCTION

Bayesian optimization (BO) is a class of machine learning-based black-box optimization strategies for finding a global optimum of a real-valued function f. Typically, f is a function whose analytical form is unavailable, but it can be evaluated with a substantial computational cost. Due to its black-box nature, BO has demonstrated its efficacy across a broad spectrum of practical applications, such as tuning hyperparameters in deep learning [25, 40, 49], searching for neural network architectures [24, 47, 56], designing materials [15, 31, 54], and discovering new drugs [4, 9, 27, 52].

BO follows a sequential decision-making process, where the outcome of each iteration informs the selection of the next evaluation point. After each step, the surrogate model is updated, thereby guiding the selection of the next point for function evaluation. This process is built around two key components: (1) updating a Gaussian process (GP) surrogate model and (2) optimizing the acquisition function to decide the next evaluation point. The Gaussian process surrogate is favored for modeling the objective function f due to its capability to iteratively incorporate all previous observations, possibly corrupted by noise. The acquisition function α_t directs the optimization by solving for the next evaluation location, denoted by x_t,

$$x_t = \operatorname*{argmax}_{x \in \mathcal{X}} \alpha_t(x).$$

This function balances the exploration of the search space with the exploitation of the accurately estimated surrogate model. In BO, this is referred to as *acquisition (inner) optimization* [48], while optimization of the whole objective function is termed *outer optimization*. Various strategies for designing the acquisition function exist, with popular options including Expected Improvement (EI) [20], Knowledge Gradient (KG) [13], Probability of Improvement (PI) [28], Upper Confidence Bound (UCB) [38], and Thompson Sampling (TS) [39], each offering a unique approach to manage the exploration-exploitation balance.

1.1 INEXACT ACQUISITION MAXIMIZATION

Optimizing the acquisition function α_t is generally considered more straightforward and computationally efficient than directly optimizing the objective f, largely because α_t is defined through a Gaussian process surrogate that offers a tractable approximation to the otherwise complex, black-box function. However, theoretical approaches often require an *exact solution* x_t for establishing regret bounds, a task that becomes particularly difficult in practice due to the possible non-convex nature of the acquisition function. This situation raises an important question:

What are the implications of *inexact* acquisition function maximization in BO?

Despite the widespread success of Bayesian optimization, this question has been surprisingly overlooked for an extended period and remains unexplored. Moreover, besides Bayesian optimization, this problem widely exists in bandits. Take the classical LinUCB algorithm [1] as an example, the acquisition function optimization requires an exact optimization solution of context and model parameters, which is highly intractable in practice. And the problem is even worse when it comes to generalized linear bandits and non-linear bandits. Recently, there has been a growing line of research on bandits and global optimization with neural network approximation [10, 53, 55] where the inexact acquisition function optimization problem still exists. In real-world applications, algorithms commonly employ quasi-Newton methods or random search to optimize the acquisition function, resulting in an approximate solution. This practice diverges from the theoretical assumptions of Bayesian optimization, which typically presume that an exact solution x_t is attainable at every iteration t. Therefore, we need a systematic study on this problem to understand what theoretical guarantee one can provide when an exact acquisition function solution cannot be obtained. In Table 1, we summarize existing works that address inexactness in bandit optimization and point out the unique position of our paper in bandit optimization. In the subsection below, we discuss related works.

1.2 RELATED WORKS

Optimization of acquisition functions. The optimization of acquisition functions has been a long-neglected issue in BO. Wilson et al. [48] first studied acquisition function maximization in Bayesian optimization. Since acquisition functions are usually non-convex and high-dimensional, they focused on how to *approximately* optimize acquisition functions. They found that acquisition functions estimated via Monte Carlo integration are consistently amenable to gradient-based optimization, and EI and UCB can be solved by greedy approaches. Although both their work and this paper study the inexact acquisition function optimization problem, our research distinguishes itself by concentrating on the impact that inexact solutions have on Bayesian optimization.

Inexact acquisition function optimization has also been explored in the context of combinatorial semi-bandits, where the decision-making process involves choosing from a combinatorial set of actions to optimize a reward function under constraints. For instance, Xu and Li [51] and Ross et al. [36] investigate the intricacies of making efficient selections amidst a combinatorial structure of actions by leveraging contextual information to inform the selection of action subsets. These works consider an α-approximate oracle [21], focusing on a relaxed version of regret to account for the inexactness of acquisition optimization. Regret is defined as the difference in reward between the pulled (super-) arm with an approximation of the best reward. Wang and Chen [45] show that in general, one cannot achieve no-regret with an approximation oracle in Thompson sampling, even for the classical multi-armed bandit problem. Kong et al. [26] revealed that linear approximation regret for combinatorial Thompson sampling is pathological, while Perrault [33] studies a specific condition on the approximation oracle, allowing a reduction to the exact oracle analysis and thus attaining sublinear regret for combinatorial semi-bandits. In contrast to these works concerning multi-armed bandits, our work focuses on Gaussian process bandits and proposes sufficient conditions that allow popular BO algorithms to achieve sublinear regret.

Random sampling and discretization methods. Random discretization methods have been popular approaches in BO to address the challenges posed by non-convex and high-dimensional acquisition functions. One source of inexactness in acquisition function optimization is due to the discretization inherent in these methods. Bergstra and Bengio [5] demonstrate that random search is often more effective than grid search for hyperparameter optimization, particularly as the dimensionality of the problem increases. They show that random search explores a larger, less promising configuration space, but still finds better models within a smaller fraction of the computation time compared to grid search. This approach is beneficial in BO as well, where the curse of dimensionality makes grid-based methods impractical. Our work provides a theoretical understanding of when random sampling works well in the context of inexact acquisition function maximization. More recently, Gramacy et al. [17] propose using candidates based on a Delaunay triangulation of the existing input design for Bayesian optimization. These "tricands" outperform both numerically optimized acquisitions and random candidate-based alternatives on benchmark synthetic and real simulation experiments. Similarly, Wycoff et al. [50] introduce the use of candidates lying on the boundary of the Voronoi tessellation of current design points. This approach significantly improves the execution time of multi-start continuous search without a loss in accuracy, by efficiently sampling the Voronoi boundary without explicitly generating the tessellation, thus accommodating large designs in high-dimensional spaces. These methods leverage geometric structures to provide more efficient candidate sets for the acquisition optimization process, aligning with our focus on inexact acquisition function optimization.

Misspecified and approximate inference in bandits. Besides acquisition functions (inner optimization), inexactness also occurs in the (outer) optimization loop of bandit problems, where the objective function class is misspecified with respect to the modeling function class. For instance, in the misspecified linear bandits setting [30], the true objective

Problem settings	Model misspecification	Approximate posterior	Inexact acquisition function
Multi-armed bandits	Lattimore et al. [30] Foster et al. [12]	Phan et al. [34] Lu and Van Roy [32]	Wang and Chen [45] Kong et al. [26]; Perrault [33]
Bayesian optimization	Bogunovic and Krause [7]	Vakili et al. [42, 43]	This paper

Table 1: Theoretical study of inexactness in bandit optimization

may be nonlinear, while in misspecified Gaussian process bandit optimization [7], the objective does not necessarily lie in a function class with a bounded RKHS norm.

Phan et al. [34] studied the effects of approximate inference on the performance of Thompson sampling in k-armed bandit problems where only an approximate posterior distribution can be used. With α-divergence governing the difference between approximate and true posterior distributions, they proposed a new algorithm that works with sublinear regret in this challenging scenario. In the context of Bayesian optimization, Vakili et al. [42, 43] introduced inducing-point-based sparse Gaussian process approximations to address the scalability challenges of full Gaussian process models. In contrast, our work assumes an exact posterior and focuses exclusively on the inexactness introduced by imperfect acquisition function optimization.

1.3 OUR CONTRIBUTIONS

In this paper, we address the problem of inexact acquisition function maximization in Bayesian optimization by analyzing its impact on regret and demonstrating that random grid search is a theoretically justified and practical acquisition solver. We focus on two popular Bayesian optimization algorithms, GP-UCB [38] and GP-TS [8], and study their cumulative regret bound when acquisition function maximization problems are not perfectly solved. Our key contributions are summarized as follows.

- To the best of our knowledge, our paper is the first to theoretically study the effect of the inexact acquisition function solutions in BO. Despite its widespread practical use, the effect of inexact acquisition maximization has been largely overlooked in the literature. Our work systematically answers the question: "How do inexact acquisition solutions impact the regret of BO algorithms, and under what conditions do they still guarantee convergence?" More broadly, our results bridge an important gap in inexact bandit optimization, complementing prior studies on model misspecification and approximate inference in multi-armed bandits and Bayesian optimization (see Table 1).
- Formally, we introduce a measure of inaccuracy in acquisition solution, worst-case accumulated inaccuracy, and we establish the cumulative regret bounds of both inexact GP-UCB and GP-TS. Our analysis quantifies how the cumulative impact of inexact acquisition solutions influences regret and establishes sufficient conditions under which inexact BO algorithms can still achieve sublinear regrets. These results generalize classical BO regret bounds to more realistic settings where exact acquisition maximization is infeasible.
- We theoretically justify random grid search as a valid acquisition solver for BO. Our regret bounds show that even with a linearly growing grid size $|\mathcal{X}_t| = \Theta(t)$, random grid search achieves sublinear regret. This significantly relaxes the condition in prior work [8], which required an exponentially larger grid size t^{2d}, demonstrating that random grid search is both computationally efficient and theoretically grounded.
- We empirically validate the efficiency of random grid search over the existing acquisition function solvers in the context of BO. Our experiments confirm that random grid search not only maintains strong regret performance but also offers substantial computational savings when solving acquisition functions, reinforcing its practical viability as an acquisition function solver.
- The Python code to reproduce the numerical experiments in Section 5 is available at `https://github.com/hwkim12/INEXACT_UCB_GRID`.

2 PRELIMINARIES

2.1 PROBLEM SETUP AND NOTATIONS

We consider a global optimization problem where the goal is to maximize an objective function $f : \mathcal{X} \to [0, \infty)$. We use the following notations

$$x^* = \operatorname*{argmax}_{x \in \mathcal{X}} f(x), \quad f^* = f(x^*),$$

where $\mathcal{X} = [-b, b]^d,\ b \in \mathbb{R}_{>0}$ is a search space of interest, which could be viewed as a set of actions or arms. Unlike the first and second-order optimization methods, where one needs an analytic expression or derivative information of f, we allow f to be a blackbox function, whose closed-form expression for the function or the derivative is not necessarily known. Only through evaluations of the function f, which could be contaminated by random noise, we aim to identify the maximum of f.

To facilitate theoretical understanding of BO methods, we assume the objective function f belongs to the reproducing kernel Hilbert space (RKHS) [3, 6, 44] corresponding to a positive semi-definite kernel function $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$, denoted by $\mathcal{H}_k$. Following [8], we restrict our attention to a set of functions f whose RKHS norm is bounded by some constant $B \in \mathbb{R}_{>0}$. Two important assumptions that will be used throughout the paper are stated below.

Assumption 1. *The kernel $k(x, x) \leq 1$ for all $x \in \mathcal{X}$.*

Assumption 2. *We consider the kernel k to be either a square-exponential kernel or a Matérn kernel with a smoothness parameter $\nu \geq 2$.*

At each t^{th} round, we select x_t by maximizing an acquisition function α_t and observe a reward

$$y_t = f(x_t) + \epsilon_t,$$

where the noise ϵ_t is assumed to be conditionally R-sub-Gaussian with respect to the σ-algebra $\mathcal{F}'_{t-1}$ generated by $\{x_1, \cdots, x_t, \epsilon_1, \cdots, \epsilon_{t-1}\}$ [8, 29], i.e., $\forall \lambda \in \mathbb{R}$, $\mathbb{E}[e^{\lambda \epsilon_t} | \mathcal{F}'_{t-1}] \leq \exp(\lambda^2 R^2/2)$ for some $R \geq 0$.

We assess the performance of a BO algorithm over T iterations based on the cumulative regret R_T, given by

$$R_T = \sum_{t=1}^{T} f^* - f(x_t),$$

where $r_t = f^* - f(x_t)$ is referred to as an instantaneous regret at round t. The BO algorithm is said to be a zero-regret algorithm if $\lim_{T\to\infty} R_T/T = 0$, and typical theoretical analyses [8, 38, 41] aim to show sublinear growth of R_T as the zero-regret algorithm will guarantee the convergence of the algorithm to the maximum. This can be seen from the fact that the simple regret, given by $f^* - \max_{t=1,\cdots,T} f(x_t)$, is bounded above by the average cumulative regret R_T/T. Throughout the paper, we use standard big $\mathcal{O}$ notation that hides universal constants, and we use $\tilde{\mathcal{O}}$ to hide all logarithmic factors in T.

2.2 GP-UCB AND GP-TS ALGORITHMS

In the midst of numerous Bayesian optimization methods, two particular strategies of interest are Gaussian Process-Upper Confidence Bound (GP-UCB) introduced in Srinivas et al. [38] and Gaussian Process-Thompson Sampling (GP-TS) proposed by Chowdhury and Gopalan [8]. More formally, to design BO algorithms, one typically imposes a zero-centered GP prior to the target objective function f and models the random noise through a Gaussian random variable with variance τ. Conditioning on all $t-1$ observations prior to obtaining t th observation, the posterior mean μ_{t-1} and standard variance $\sigma^2_{t-1}(x)$ of the Gaussian process are given by

$$\mu_{t-1}(x) = k_{t-1}(x)^T (K_{t-1} + \tau I)^{-1} Y_{t-1},$$
$$\sigma^2_{t-1}(x) = k(x, x) - k_{t-1}(x)^T (K_{t-1} + \tau I)^{-1} k_{t-1}(x),$$

where

$$k_{t-1}(x) = [k(x_1, x), \cdots, k(x_{t-1}, x)]^T,$$
$$Y_{t-1} = [y_1, \cdots, y_{t-1}],$$
$$K_{t-1} = [k(x, x')]_{x,x' \in \{x_1,\cdots,x_{t-1}\}}.$$

In other words, the posterior distribution of f conditioning on $\{(x_i, y_i)\}_{i=1}^{t-1}$ is given by the Gaussian process with mean function μ_{t-1} and posterior variance σ^2_{t-1}.

Algorithm 1 GP-UCB [8, 38]

Input: Kernel k; Total number of iterations T; Initial design points X_0; Initial observations Y_0.
Construct $\mu_0(x)$ and $\sigma_0(x)$ using X_0, Y_0.
for $t = 1, \ldots, T$ **do**
  $\beta_t \leftarrow B + R\sqrt{2(\gamma_{t-1} + 1 + \log(1/\delta))}$.
  $x_t \leftarrow \text{argmax}_x \, \mu_{t-1}(x) + \beta_t \sigma_{t-1}(x)$.
  Observe $y_t = f(x_t) + \epsilon_t$.
  $X_t = X_{t-1} \cup \{x_t\}, Y_t = Y_{t-1} \cup \{y_t\}$.
  Update $\mu_t(x)$ and $\sigma_t(x)$ using X_t, Y_t.
end for

Algorithm 2 GP-TS [8]

Input: Kernel k; Total number of iterations T; Initial design points X_0; Initial observations Y_0.
Construct $\mu_0(x)$ and $\sigma_0(x)$ using X_0, Y_0.
for $t = 1, \ldots, T$ **do**
  $\beta_t \leftarrow B + R\sqrt{2(\gamma_{t-1} + 1 + \log(2/\delta))}$.
  Sample $f_t(\cdot) \sim \mathcal{GP}_D\left(\mu_{t-1}(\cdot), s_t^2\right)$, with $s_t^2(\cdot) = \beta_t^2 \sigma^2_{t-1}(\cdot)$.
  Choose the current decision set $\mathcal{X}_t \subset \mathcal{X}$ of size $|\mathcal{X}_t| = (2BLbdt^2)^d$.
  Set $x_t = \arg\max_{x \in \mathcal{X}_t} f_t(x)$.
  Observe $y_t = f(x_t) + \epsilon_t$.
  $X_t = X_{t-1} \cup \{x_t\}, Y_t = Y_{t-1} \cup \{y_t\}$.
  Update $\mu_t(x)$ and $\sigma_t(x)$ using X_t, Y_t.
end for

For the choice of an acquisition function to facilitate a Bayesian optimization strategy, Srinivas et al. [38] considered a UCB function of the form

$$\alpha_t(x) \coloneqq \mu_{t-1}(x) + \beta_t \sigma_{t-1}(x),$$

where β_t is chosen to balance exploitation (picking points with high function values) and exploration (picking points where the prediction based on the posterior distribution is highly uncertain). In their work, they termed the Bayesian optimization strategy based on the UCB acquisition function

as GP-UCB, which has since become popular in both theoretical analyses and empirical applications. In the meantime, Chowdhury and Gopalan [8] proposed another BO strategy under the name of GP-TS, which leverages the acquisition function of the form

$$\alpha_t(x) \coloneqq f_t(x),$$

where f_t is a sample path from the posterior Gaussian process with the mean function μ_{t-1} and variance function $\beta_t^2\sigma_{t-1}^2$. Such a strategy has been widely used under the name of Thompson sampling in BO and bandit literature [2, 22, 37]. In practice, optimizing a randomly drawn sample path f_t can be just as hard as finding the optimum of the target objective f. In reflection of such difficulty, a more practical version of the aforementioned TS-based BO strategy through discretizing the search space for the acquisition function was proposed in Chowdhury and Gopalan [8].

For the theoretical analysis, Srinivas et al. [38] and Chowdhury and Gopalan [8] considered an increasing sequence for the choice of β_t, and in particular, Chowdhury and Gopalan [8] set β_t to grow at a rate of $\mathcal{O}(\sqrt{\gamma_T})$ where γ_T is a kernel-dependent quantity known as the maximum information gain at time t, given by

$$\gamma_T = \max_{\mathcal{X}_T \subset \mathcal{X}: |\mathcal{X}_T| = T} \frac{1}{2} \log\det(I_T + \tau^{-1} K_T).$$

The formal procedures for Bayesian optimization using the UCB and Thompson sampling acquisition functions are outlined in Algorithms 1 and 2, respectively. Furthermore, the state-of-the-art growth rate of the maximum information gain is provided in Vakili et al. [41], which states that $\gamma_T = \mathcal{O}\left(T^{\frac{d}{d+2\nu}} \log^{\frac{2\nu}{2\nu+d}}(T)\right)$ for Matérn Kernel and $\gamma_T = \mathcal{O}\left(\log^{d+1}(T)\right)$ for SE Kernel. Although we work with UCB and TS in this paper, our analysis can be extended to more acquisition functions[1].

3 REGRET BOUNDS UNDER INEXACT ACQUISITION FUNCTION MAXIMIZATION

Although theoretical developments in BO often assume that the acquisition maximization is solved perfectly, this is rarely the case in practice. For example, when optimizing the UCB acquisition function—whose gradient information is typically available—a popular approach is to use quasi-Newton methods with multiple starting points [14, 16, 35]. While these methods have proven effective for convex problems, the UCB acquisition function is usually nonconvex. Consequently, factors such as the number of starting points, the total number of optimization iterations, and the stopping criteria all impact the quality of the acquisition solution obtained.

In TS-based BO methods, it is common to use a discretized subset of the action space because gradients for the sample paths are difficult to obtain. In these cases, a sorting algorithm is used to select the maximum value from the posterior sample path within the chosen grid [8, 23]. Although theoretical studies have shown that the TS strategy converges to the optimal function value, the grid size must grow exponentially with the number of dimensions [8]. Moreover, the granularity of the discretization heavily influences the computational cost. As a result, typical implementations use the finest discretization possible within the available computational budget, which can lead to inexact acquisition solutions.

In this section, we examine how a series of inexact acquisition function maximizations affects BO strategies. Although we restrict our attention to acquisition functions of GP-UCB and GP-TS, our analysis is not pertained to any specific acquisition function solver. We introduce a measure to quantify the accumulated inaccuracies caused by these inexact maximizations, with the goal of understanding how they impact the existing cumulative regret bounds.

3.1 ACCUMULATED INACCURACY

We first introduce a measure of inaccuracies that arise when solving acquisition function optimization problems. Let $\alpha_t^* = \max \alpha_t(x)$, be the maximum value of the acquisition function at t^{th} iteration, which is at least as large as $\alpha_t(x_t)$ where x_t is the action selected at that iteration. For our analysis, we assume that the acquisition functions $\{\alpha_t\}_{t=1}^T$ are nonnegative and always achieve a strictly positive maximum. This assumption can be easily met by adding a positive constant to the acquisition function at each iteration or by appropriately restricting the search space. A detailed justification for the existence of such a constant is provided in Appendix A.2. We emphasize that this modification does not change the chosen action, nor does it incur any additional computational cost.

At each iteration, we quantify the accuracy of the acquisition optimization solution by the ratio of the acquisition function value at the selected action to its maximum value:

$$\eta_t \coloneqq \frac{\alpha_t(x_t)}{\alpha_t^*} \in [\tilde{\eta}_t, 1],$$

so that $\eta_t = 1$ if the acquisition function is maximized perfectly. In other words, a larger value of η_t indicates a more accurate solution to the t^{th} acquisition optimization problem. Here, $\tilde{\eta}_t$ represents the worst-case accuracy of the t^{th} acquisition function solver. The overall impact of

[1] Wang and Jegelka [46] showed that MES, MVES, and PI can be cast as special cases of UCB with appropriately chosen confidence parameters, and thus fall within the scope of our theoretical framework.

inaccurate solutions across iterations is then captured by the worst-case accumulated inaccuracy,

$$M_T = \sum_{t=1}^{T}(1-\tilde{\eta}_t) \in [0, T],$$

which represents the total inaccuracy permitted over T rounds of Bayesian optimization. Using this measure of accumulated inaccuracy, we establish cumulative regret bounds for GP-UCB and GP-TS in the remainder of this section, in contrast to existing theoretical works that assume perfect acquisition solutions.

3.2 REGRET BOUNDS

Recall that for the kernel function under consideration, we have $|k(x, x')| \leq 1$. Furthermore, the t^{th} action x_t satisfies $\alpha_t^* := \max \alpha_t(x) \geq \alpha_t(x_t) \geq \eta_t \alpha_t^*$, where $\alpha_t(x) = \mu_{t-1}(x) + \beta_t \sigma_{t-1}(x)$. We denote by $\mathcal{H}_k$ the RKHS associated with a chosen kernel k. We then establish cumulative regret bounds for the GP-UCB algorithm in the presence of inexact acquisition function solutions.

Theorem 3. *Under assumptions 1 and 2, suppose the objective function $f \in \mathcal{H}_k$ satisfies $\|f\|_{\mathcal{H}_k} \leq B$. Then, the inexact GP-UCB algorithm with $\beta_t = B + R\sqrt{2(\gamma_{t-1} + 1 + \log(1/\delta))}$ and worst-case accumulated inaccuracy M_T achieves a cumulative regret bound of the form:*

$$R_T = \mathcal{O}\left(\gamma_T \sqrt{T} + M_T \sqrt{\gamma_T}\right),$$

with probability $1 - \delta$.

Note that when $M_T = 0$ (exact maximization), the cumulative bound above recovers standard regret guarantees for GP-UCB [8, 38]. The additive factor of $M_T\sqrt{\gamma_T}$ accounts for the effect of inaccurate acquisition solutions. This implies that if the worst-case accumulated inaccuracy M_T and the maximum information gain γ_T do not grow too quickly so that $\frac{M_T\sqrt{\gamma_T}}{T} \to 0$ as $T \to \infty$, the inexact GP-UCB algorithm will converge to the optimal solution. For example, with a squared-exponential kernel, it has been shown that γ_T grows logarithmically [41]. Theorem 3 thus indicates that the inexact GP-UCB algorithm converges asymptotically to the optimal function value, provided that the worst-case accumulated inaccuracy M_T is sublinear.

One can also establish a similar result for the GP-TS algorithm, as stated below. In contrast to the GP-TS algorithm introduced in [8], our formulation accounts for the extra uncertainty induced by inexact solutions through the use of sample paths obtained from the GP posterior with an enlarged variance.

Theorem 4. *Under assumptions 1 and 2, suppose the objective function $f \in \mathcal{H}_k$ satisfies $\|f\|_{\mathcal{H}_k} \leq B$. Then, the inexact GP-TS algorithm with variance $s_{t-1}^2(x) = (\beta_t \sigma_{t-1}(\cdot) + \tilde{v}_t)^2$, $\tilde{v}_t = (\frac{1}{\tilde{\eta}_t} - 1)B$ and worst-case accumulated inaccuracy M_T achieves a cumulative regret bound of the form:*

$$R_T = \mathcal{O}\left(\gamma_T \sqrt{T \log T} + M_T \sqrt{\gamma_T \log T}\right),$$

with probability $1 - \delta$.

The above regret bound matches the existing bound for GP-TS [8], up to an additive factor of $M_T\sqrt{\gamma_T \log T}$, which captures the inaccuracies in the acquisition function solutions. Similar to the GP-UCB case, if the worst-case accumulated inaccuracy M_T and the maximum information gain γ_T do not increase too rapidly so that $M_T\sqrt{\gamma_T \log T}/T \to 0$ as $T \to \infty$, Theorem 4 shows that the inexact GP-TS algorithm will achieve a sublinear cumulative regret bound.

An overall implication of the theorems established in this section is that BO strategies can retain asymptotic convergence guarantees even without exact acquisition function solutions, provided that the accuracy of these solutions improves over time. This insight naturally motivates the exploration of BO strategies that are computationally more efficient by reducing the effort spent on acquisition function maximizations while preserving convergence guarantees.

4 SOLVING ACQUISITION FUNCTION THROUGH RANDOM GRID SEARCH

Building on the insights from Section 3, we further investigate a simple yet computationally efficient approach: a random grid search for acquisition function maximization. In this method, to maximize an acquisition function α_t, one obtains a set of random points from the search space and selects the one with the highest acquisition function value. This approach has been employed in numerous BO problems and has demonstrated its effectiveness [23, 35]. Importantly, our results provide the first theoretical validation for this approach, showing that even with a linear growth in grid size, one can still achieve sublinear regret.

Concretely, we analyze the GP-UCB and GP-TS algorithms using a random grid search for acquisition maximization and derive their cumulative regret bounds. As noted in the previous section, the accuracy of the acquisition solver must improve over iterations. To achieve this, we employ a sequence of random grids $\{\mathcal{X}_t\}_{t\in\mathbb{N}}$ whose size grows linearly with the iteration count, i.e., $|\mathcal{X}_t| = \Theta(t)$. Such a strategy of increasing the grid size has proven empirically successful [23] in combination with the TS acquisition function.

To precisely define our acquisition function optimization procedure, consider a grid of random samples $\mathcal{X}_t \subset \mathcal{X}$ that serves as the search space for the t^{th} acquisition function

optimization. At each t^{th} iteration, we set

$$x_t \coloneqq \begin{cases} \operatorname{argmax}_{x \in \mathcal{X}_t} \mu_{t-1}(x) + \beta_t \sigma_{t-1}(x) & \text{for GP-UCB} \\ \operatorname{argmax}_{x \in \mathcal{X}_t} f_t(x) & \text{for GP-TS.} \end{cases}$$

We refer to the first strategy as random grid GP-UCB and the second as random grid GP-TS. For the choice of random grid, we make the following assumption and state our cumulative regret bound results.

Assumption 5. *At iteration t, $\mathcal{X}_t$ consists of t independent samples drawn uniformly from $\mathcal{X}$. Additionally, the sequence of random grids $\{\mathcal{X}_t\}_{t \in \mathbb{N}}$ is independent across iterations.*

Remark 6. *Our results extend to any sampling scheme whose density remains strictly positive over the search space. Here, we focus on a uniform random grid because it is the easiest to implement in practice.*

Theorem 7. *Under assumptions 1, 2 and 5, suppose $f \in \mathcal{H}_k$ with $\|f\|_{\mathcal{H}_k} \leq B$. With probability $1-\delta$, the random grid GP-UCB with $\beta_t = B + R\sqrt{2(\gamma_{t-1} + 1 + \log(2/\delta))}$ yields*

$$R_T = \mathcal{O}\left(\gamma_T \sqrt{T}\right) + \tilde{\mathcal{O}}\left(T^{\frac{d-1}{d}}\right).$$

Theorem 8. *Under assumptions 1, 2 and 5, suppose $f \in \mathcal{H}_k$ with $\|f\|_{\mathcal{H}_k} \leq B$. With probability $1-\delta$, the random grid GP-TS with $\beta_t = B + R\sqrt{2(\gamma_{t-1} + 1 + \log(3/\delta))}$ yields*

$$R_T = \mathcal{O}\left(\gamma_T \sqrt{T \log T}\right) + \tilde{\mathcal{O}}\left(T^{\frac{d-1}{d}}\right).$$

Remark 9. *Similar to the regret bounds from Section 3, our cumulative regret bounds now incorporate an additional factor that accounts for the inaccuracies introduced by the random grid search. In both Theorem 7 and Theorem 8, these effects are represented by the term $\tilde{\mathcal{O}}\left(T^{\frac{d-1}{d}}\right)$. Importantly, our results demonstrate that even a simple random grid search is sufficient for acquisition function maximization, ensuring that Bayesian optimization asymptotically converges to the global optimum under a suitable growth rate of the maximum information gain γ_T.*

Remark 10. *As the grid size increases superlinearly, the order of the inaccuracy factor can potentially decrease. For instance, if the grid size grows at an order of t^2, we would instead get the inaccuracy factor of $\tilde{\mathcal{O}}\left(t^{\frac{d-2}{d}}\right)$ or if the grid size grows at an order of t^d, we would approximately get the inaccuracy factor of $\mathcal{O}(\log t)$.*

Remark 11. *In our analyses, we mainly considered the squared-exponential and Matérn kernels, thanks to their popularity and existing growth rate results on their maximum information gain. The key aspect of the maximum information gain growth rate can be characterized by the decaying rate of eigenvalues associated with the kernel [41]. Naturally, as long as the decaying rate of a kernel is known, our analyses can also be extended to other types of kernels. Furthermore, our analysis can also be extended to settings with non-stationary noise variance [19].*

Our theoretical results indicate that random grid search is sufficient for optimizing UCB and TS acquisition functions to achieve sublinear regret bounds for both GP-UCB and GP-TS. Although many successful implementations of random grid search for acquisition function maximization exist, there has been little theoretical justification for this approach. The closest work we are aware of is by [8], which incorporated a grid search within the GP-TS algorithm to solve the acquisition function optimization problem. Their analysis yields sublinear regret for certain kernels but requires the grid size to scale as t^{2d}, where t is the iteration index and d is the problem dimension. This requirement is computationally demanding; indeed, in the same paper, the authors used a fixed grid size for numerical experiments, highlighting the gap between theoretical developments and practical implementations. Unlike their results, our regret bounds justify the empirical success of the TS acquisition function when using a random grid search that grows linearly with the iteration number, as demonstrated in [23]. This result not only demonstrates the theoretical soundness of using random grid search for acquisition maximization but also provides rigorous justification for the empirical findings in [23], which observed that a random grid search of order $\mathcal{O}(t)$ is robust and efficient for TS algorithms. Furthermore, as Section 5 will show, the computational efficiency of random grid search approach offers a significant practical advantage over more complex solvers such as quasi-Newton methods, without sacrificing convergence guarantees.

5 EXPERIMENTS

In this section, we first present experimental results on synthetic functions and then a real-world automated machine learning task. In Appendix B, we provide additional details of the experimental setup and more ablation studies.

5.1 SYNTHETIC EXPERIMENTS

We conduct numerical experiments on six benchmark functions to demonstrate the effectiveness of random grid search combined with the UCB acquisition function. We refer the reader to [23] for a demonstration of the effectiveness of random grid search when using the TS acquisition function. The test functions include Branin (2D on $[-5, 10] \times [0, 15]$), Rastrigin (3D on $[-5.12, 5.12]^3$), Hartmann3 (3D on $[0, 1]^3$), Hartmann4 (4D on $[0, 1]^4$), Levy (5D on $[-10, 10]^5$), and Hartmann6 (6D on $[0, 1]^6$). For each problem, a set of initial design points is generated using a Sobol sequence. These

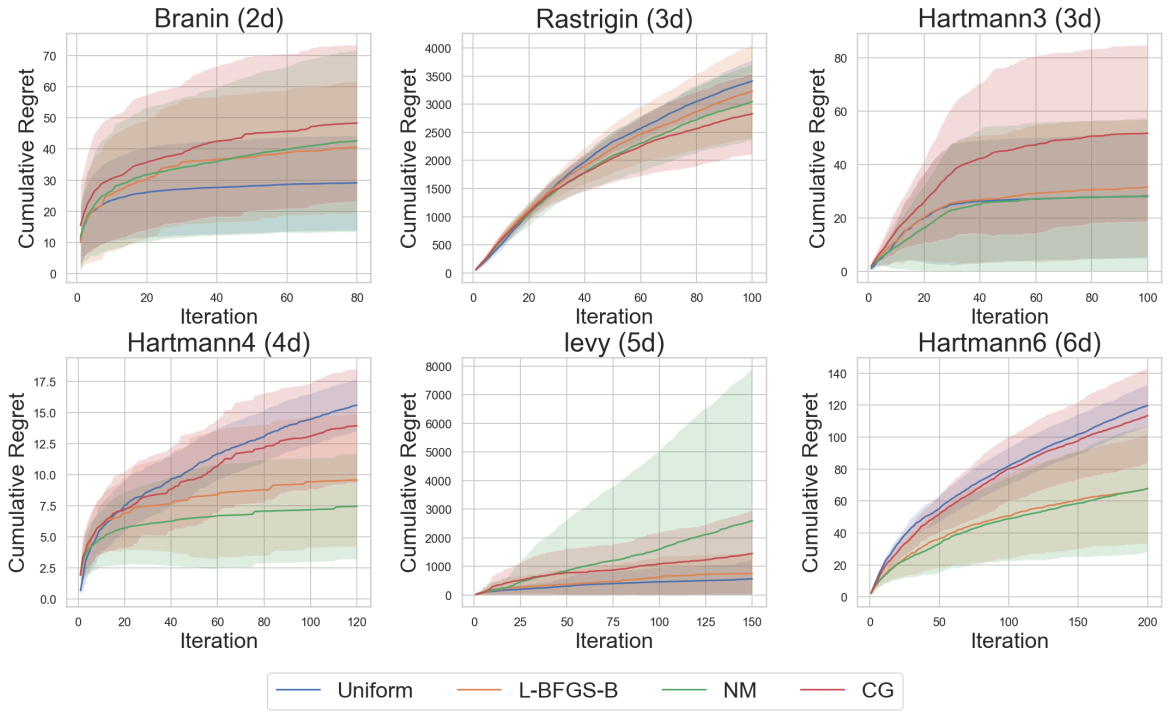

Figure 1: Cumulative regret comparison between acquisition function solvers.

points are scaled to the appropriate domain. We vary the number of initial design points by problem (20 for Branin, 30 for Rastrigin and Hartmann3, 40 for Hartmann4, 50 for Levy, and 60 for Hartmann6), reflecting the dimension of the objective function. After initialization, BO algorithms with different choices of acquisition function solvers are performed for a predetermined number of iterations: 80 iterations for Branin, 100 for Rastrigin, Hartmann3, and Hartmann4, 150 for Levy, and 200 for Hartmann6. In each iteration, a Gaussian process with a Matérn kernel is fitted to the available data, and the UCB acquisition function is optimized with an exploration parameter $\beta_t = \sqrt{\log(t+2)}$ for each iteration t. The optimization of this acquisition function is carried out using one of four acquisition optimization methods: a uniform random grid search (abbreviated as Uniform), popular quasi-Newton methods including Limited-memory Broyden–Fletcher–Goldfarb–Shanno (abbreviated as L-BFGS-B) as well as Nelder–Mead (abbreviated as NM), and the conjugate gradient (abbreviated as CG) method. For the random grid size $|\mathcal{X}_t|$, we set it to be $100t$, which scales linearly in terms of the number of iterations. Each method is evaluated over 20 independent experiments (with varying random seeds) to provide a statistically valid comparison in terms of both cumulative regret and computational time.

Figure 1 shows that, across a range of dimensions and objective landscapes, the uniform sample grid search consistently achieves competitive cumulative regret and keeps pace with, or in some cases outperforms, the popular acquisition optimization tools. For instance, in the 2D Branin problem, the

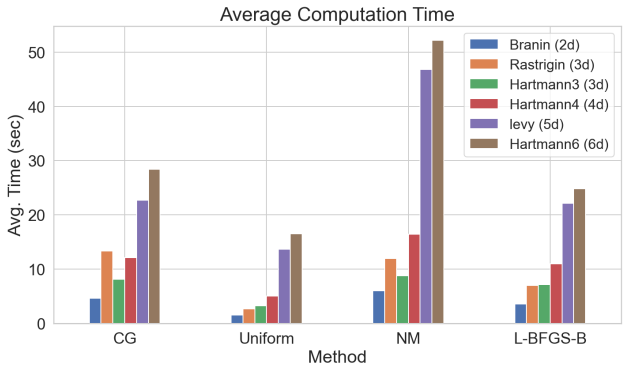

Figure 2: Computational time comparison between acquisition function solvers.

uniform sample grid search's cumulative regret curve is superior to that of other sophisticated optimization algorithms over the entire horizon, demonstrating rapid improvement from the outset. On the more challenging 3D Rastrigin, 3D Hartmann3, and 4D Hartmann4 functions, uniform sample grid search continues to exhibit a steady reduction in regret, closely matching or surpassing other solvers. Even for the higher-dimensional Levy (5D) and Hartmann6 (6D) problems, the uniform sample grid search approach remains impressively competitive, achieving a final cumulative regret that is comparable to the gradient-based methods.

Figure 2 further highlights the uniform sample grid search's practical advantages: it maintains one of the lowest average

runtimes across all problems, standing in stark contrast to NM, which exhibits significantly longer runtimes (especially in the 5D and 6D settings)[2].

While the gap in runtime between uniform random grid search and existing acquisition function solvers narrows in higher dimensions—partly due to the increased number of BO iterations and corresponding grid evaluations—the uniform sample grid search continues to offer a favorable balance between efficiency and regret reduction, making it an appealing choice when balancing rapid progress in regret reduction with efficient use of computational resources.

5.2 REAL-WORLD AUTOML TASK

Similarly, we perform a hyperparameter tuning task focusing on improving the validation accuracy of the gradient boosting model (from the Python `sklearn` package) on the breast-cancer dataset (obtained from the UCI machine learning repository). Details of 11 hyperparameters are shown in Appendix B. Figure 3 shows the performances using different inner optimization methods for GP-UCB. We can see that random grid search ("uniform" in the figure) achieves relatively low cumulative regret as well as low computation time from Figure 3.

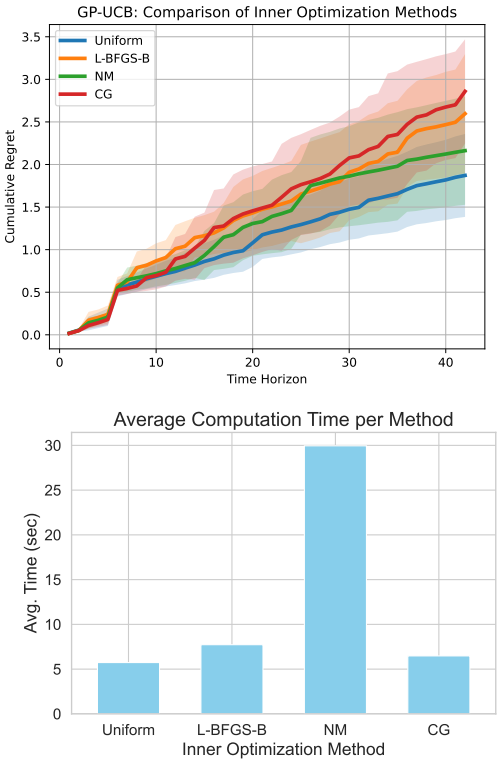

Figure 3: Cumulative regret (upper) and average computational time (lower) of different acquisition optimization tools on the real-world AutoML task.

6 CONCLUSION

In this paper, to the best of our knowledge, we study the inexact acquisition function maximization problem for BO for the first time, a topic that has been largely overlooked. Existing works predominantly operate under the assumption that an exact solution is attainable, thus allowing for the establishment of rigorous theoretical upper bounds. However, in practice, finding exact solutions is difficult due to the multimodal nature of acquisition functions

To address this discrepancy, we define a measure of inaccuracy in acquisition solution, referred to as the worst-case accumulated inaccuracy. We establish cumulative regret bounds for both inexact GP-UCB and GP-TS. Our bounds show that under some conditions on the worst-case accumulated inaccuracy, inexact BO algorithms can still achieve sublinear regrets. Moreover, we extend to provide the first theoretical validation for random grid search, showing that even with a linear growth in grid size, one can still achieve sublinear regret. Our experimental results also validate the effectiveness of random grid search in solving acquisition maximization problems.

Although our work focuses on GP-UCB and GP-TS, our analysis can be extended to other acquisition functions with similar structural properties. Additionally, our framework opens avenues for studying adaptive inexact optimization strategies, where the computational effort allocated to acquisition function maximization is dynamically adjusted based on accumulated inaccuracy. We hope our study will spur systematic theoretical work on inexact acquisition optimization and inspire empirical investigations into efficiently allocating computational resources in BO.

Acknowledgements

This work was partially done when HK and CL were at the University of Chicago. This work was supported in part by the National Science Foundation under Grant No. IIS 2313131, IIS 2332475, and CMMI 2037026. The authors would like to thank the reviewers and the AC for helpful comments that improved the final version of this paper.

[2]We provide a brief comparison of the computational complexity of the acquisition optimizers used in Figure 2. For the random grid search, the grid size scales as $\mathcal{O}(t)$, roughly yielding a total complexity of $\tilde{\mathcal{O}}(T^2)$ over T iterations. In contrast, gradient-based methods (e.g., L-BFGS-B, CG) require iterative optimization with per-iteration costs dependent on the dimensionality. Let I_t denote the number of gradient updates at step t, and let cost(gradient) be the cost of a single gradient evaluation (e.g., $\mathcal{O}(d^3)$ for Newton, $\mathcal{O}(d^2)$ for CG/BFGS). Then the overall complexity becomes $\sum_{t=1}^{T} \mathcal{O}(I_t \times \text{cost(gradient)})$.

References

[1] Yasin Abbasi-yadkori, Dávid Pál, and Csaba Szepesvári. Improved algorithms for linear stochastic bandits. In *Advances in Neural Information Processing Systems 24*, 2011.

[2] Shipra Agrawal and Navin Goyal. Thompson sampling for contextual bandits with linear payoffs. In *International Conference on Machine Learning*, 2013.

[3] Nachman Aronszajn. Theory of reproducing kernels. *Transactions of the American mathematical society*, 68(3):337–404, 1950.

[4] Hugo Bellamy, Abbi Abdel Rehim, Oghenejokpeme I Orhobor, and Ross King. Batched bayesian optimization for drug design in noisy environments. *Journal of Chemical Information and Modeling*, 62(17):3970–3981, 2022.

[5] James Bergstra and Yoshua Bengio. Random search for hyper-parameter optimization. *Journal of machine learning research*, 13(2), 2012.

[6] Alain Berlinet and Christine Thomas-Agnan. *Reproducing Kernel Hilbert Spaces in Probability and Statistics*. Springer Science & Business Media, 2011.

[7] Ilija Bogunovic and Andreas Krause. Misspecified gaussian process bandit optimization. *Advances in Neural Information Processing Systems*, 34:3004–3015, 2021.

[8] Sayak Ray Chowdhury and Aditya Gopalan. On kernelized multi-armed bandits. In *International Conference on Machine Learning*, pages 844–853, 2017.

[9] Lionel Colliandre and Christophe Muller. Bayesian optimization in drug discovery. In *High Performance Computing for Drug Discovery and Biomedicine*, pages 101–136. Springer, 2023.

[10] Zhongxiang Dai, Yao Shu, Bryan Kian Hsiang Low, and Patrick Jaillet. Sample-then-optimize batch neural Thompson sampling. In *Advances in Neural Information Processing Systems 35*, 2022.

[11] Nando de Freitas, Alex Smola, and Masrour Zoghi. Regret bounds for deterministic gaussian process bandits. *arXiv preprint arXiv:1203.2177*, 2012.

[12] Dylan J Foster, Claudio Gentile, Mehryar Mohri, and Julian Zimmert. Adapting to misspecification in contextual bandits. *Advances in Neural Information Processing Systems*, 33:11478–11489, 2020.

[13] Peter Frazier, Warren Powell, and Savas Dayanik. The knowledge-gradient policy for correlated normal beliefs. *INFORMS Journal on Computing*, 21(4):599–613, 2009.

[14] Peter I Frazier. A tutorial on bayesian optimization. *arXiv preprint arXiv:1807.02811*, 2018.

[15] Peter I Frazier and Jialei Wang. Bayesian optimization for materials design. In *Information science for materials discovery and design*, pages 45–75. Springer, 2016.

[16] Robert B Gramacy. *Surrogates: Gaussian process modeling, design, and optimization for the applied sciences*. Chapman and Hall/CRC, 2020.

[17] Robert B Gramacy, Annie Sauer, and Nathan Wycoff. Triangulation candidates for bayesian optimization. *Advances in Neural Information Processing Systems*, 35:35933–35945, 2022.

[18] Tapio Helin, Andrew M Stuart, Aretha L Teckentrup, and Konstantinos C Zygalakis. Introduction to gaussian process regression in bayesian inverse problems, with new results on experimental design for weighted error measures. In *International Conference on Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing*, pages 49–79. Springer, 2022.

[19] Shogo Iwazaki and Shion Takeno. Improved regret analysis in gaussian process bandits: Optimality for noiseless reward, rkhs norm, and non-stationary variance. *arXiv preprint arXiv:2502.06363*, 2025.

[20] Donald R Jones, Matthias Schonlau, and William J Welch. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4):455–492, 1998.

[21] Sham M Kakade, Adam Tauman Kalai, and Katrina Ligett. Playing games with approximation algorithms. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 546–555, 2007.

[22] Kirthevasan Kandasamy, Jeff Schneider, and Barnabás Póczos. High dimensional bayesian optimisation and bandits via additive models. In *International Conference on Machine Learning*, 2015.

[23] Kirthevasan Kandasamy, Akshay Krishnamurthy, Jeff Schneider, and Barnabás Póczos. Parallelised bayesian optimisation via thompson sampling. In *International conference on artificial intelligence and statistics*, pages 133–142. PMLR, 2018.

[24] Kirthevasan Kandasamy, Willie Neiswanger, Jeff Schneider, Barnabas Poczos, and Eric P Xing. Neural architecture search with bayesian optimisation and optimal transport. In *Advances in neural information processing systems 31*, 2018.

[25] Kirthevasan Kandasamy, Karun Raju Vysyaraju, Willie Neiswanger, Biswajit Paria, Christopher R.

Collins, Jeff Schneider, Barnabas Poczos, and Eric P. Xing. Tuning hyperparameters without grad students: Scalable and robust bayesian optimisation with dragonfly. *Journal of Machine Learning Research*, 21(81): 1–27, 2020.

[26] Fang Kong, Yueran Yang, Wei Chen, and Shuai Li. The hardness analysis of thompson sampling for combinatorial semi-bandits with greedy oracle. *Advances in Neural Information Processing Systems*, 34:26701–26713, 2021.

[27] Ksenia Korovina, Sailun Xu, Kirthevasan Kandasamy, Willie Neiswanger, Barnabas Poczos, Jeff Schneider, and Eric Xing. Chembo: Bayesian optimization of small organic molecules with synthesizable recommendations. In *International Conference on Artificial Intelligence and Statistics*, pages 3393–3403, 2020.

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

[29] Tor Lattimore and Csaba Szepesvári. *Bandit Algorithms*. Cambridge University Press, 2020.

[30] Tor Lattimore, Csaba Szepesvari, and Gellert Weisz. Learning with good feature representations in bandits and in rl with a generative model. In *International Conference on Machine Learning*, pages 5662–5670, 2020.

[31] Bowen Lei, Tanner Quinn Kirk, Anirban Bhattacharya, Debdeep Pati, Xiaoning Qian, Raymundo Arroyave, and Bani K Mallick. Bayesian optimization with adaptive surrogate models for automated experimental design. *Npj Computational Materials*, 7(1):194, 2021.

[32] Xiuyuan Lu and Benjamin Van Roy. Ensemble sampling. *Advances in neural information processing systems*, 30, 2017.

[33] Pierre Perrault. When combinatorial thompson sampling meets approximation regret. *Advances in Neural Information Processing Systems*, 35:17639–17651, 2022.

[34] My Phan, Yasin Abbasi Yadkori, and Justin Domke. Thompson sampling and approximate inference. *Advances in Neural Information Processing Systems*, 32, 2019.

[35] Tony Pourmohamad and Herbert KH Lee. *Bayesian optimization with application to computer experiments*. Springer, 2021.

[36] Stephane Ross, Jiaji Zhou, Yisong Yue, Debadeepta Dey, and Drew Bagnell. Learning policies for contextual submodular prediction. In *International Conference on Machine Learning*, pages 1364–1372. PMLR, 2013.

[37] Daniel Russo and Benjamin Van Roy. Learning to optimize via posterior sampling. *Mathematics of Operations Research*, 39(4):1221–1243, 2014.

[38] Niranjan Srinivas, Andreas Krause, Sham Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: no regret and experimental design. In *International Conference on Machine Learning*, 2010.

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

[40] Ryan Turner, David Eriksson, Michael McCourt, Juha Kiili, Eero Laaksonen, Zhen Xu, and Isabelle Guyon. Bayesian optimization is superior to random search for machine learning hyperparameter tuning: Analysis of the black-box optimization challenge 2020. In *NeurIPS 2020 Competition and Demonstration Track*, pages 3–26, 2021.

[41] Sattar Vakili, Kia Khezeli, and Victor Picheny. On information gain and regret bounds in gaussian process bandits. In *International Conference on Artificial Intelligence and Statistics*, pages 82–90. PMLR, 2021.

[42] Sattar Vakili, Henry Moss, Artem Artemev, Vincent Dutordoir, and Victor Picheny. Scalable thompson sampling using sparse gaussian process models. *Advances in neural information processing systems*, 34: 5631–5643, 2021.

[43] Sattar Vakili, Jonathan Scarlett, Da-shan Shiu, and Alberto Bernacchia. Improved convergence rates for sparse approximation methods in kernel-based learning. In *International Conference on Machine Learning*, pages 21960–21983. PMLR, 2022.

[44] Grace Wahba. *Spline Models for Observational Data*. SIAM, 1990.

[45] Siwei Wang and Wei Chen. Thompson sampling for combinatorial semi-bandits. In *International Conference on Machine Learning*, pages 5114–5122. PMLR, 2018.

[46] Zi Wang and Stefanie Jegelka. Max-value entropy search for efficient bayesian optimization. In *International conference on machine learning*, pages 3627–3635. PMLR, 2017.

[47] Colin White, Willie Neiswanger, and Yash Savani. Bananas: Bayesian optimization with neural architectures for neural architecture search. In *AAAI Conference on Artificial Intelligence*, pages 10293–10301, 2021.

[48] James Wilson, Frank Hutter, and Marc Deisenroth. Maximizing acquisition functions for bayesian optimization. *Advances in neural information processing systems*, 31, 2018.

[49] Jian Wu, Saul Toscano-Palmerin, Peter I Frazier, and Andrew Gordon Wilson. Practical multi-fidelity bayesian optimization for hyperparameter tuning. In *Uncertainty in Artificial Intelligence*, pages 788–798, 2020.

[50] Nathan Wycoff, John W Smith, Annie S Booth, and Robert B Gramacy. Voronoi candidates for bayesian optimization. *arXiv preprint arXiv:2402.04922*, 2024.

[51] Haike Xu and Jian Li. Simple combinatorial algorithms for combinatorial bandits: Corruptions and approximations. In *Uncertainty in Artificial Intelligence*, pages 1444–1454. PMLR, 2021.

[52] Kevin K Yang, Yuxin Chen, Alycia Lee, and Yisong Yue. Batched stochastic bayesian optimization via combinatorial constraints design. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 3410–3419. PMLR, 2019.

[53] Weitong Zhang, Dongruo Zhou, Lihong Li, and Quanquan Gu. Neural thompson sampling. In *International Conference on Learning Representations*, 2020.

[54] Yichi Zhang, Daniel W Apley, and Wei Chen. Bayesian optimization for materials design with mixed quantitative and qualitative variables. *Scientific reports*, 10(1):4924, 2020.

[55] Dongruo Zhou, Lihong Li, and Quanquan Gu. Neural contextual bandits with ucb-based exploration. In *International Conference on Machine Learning*, 2020.

[56] Hongpeng Zhou, Minghao Yang, Jun Wang, and Wei Pan. Bayesnas: A bayesian approach for neural architecture search. In *International conference on machine learning*, pages 7603–7613, 2019.

A PROOF FOR THEORETICAL STATEMENTS

A.1 ASSUMPTIONS AND KEY LEMMAS

We first list out necessary assumptions and key lemmas to establish theorems stated in the manuscript.

Assumption 12. *The kernel $k(x, x) \leq 1$ for all $x \in \mathcal{X}$.*

Assumption 13. *We consider the kernel k to be either a square-exponential kernel or a Matérn kernel with smoothness parameter $\nu \geq 2$.*

Assumption 14. *At iteration t, $\mathcal{X}_t$ is a collection of t random samples drawn uniformly from $\mathcal{X}$. Furthermore, we assume that the set of uniform random grids $\{\mathcal{X}_t\}_{t\in\mathbb{N}}$ are independent.*

Lemma 15 (Theorem 2 of Chowdhury and Gopalan [8])**.** *Suppose $f \in \mathcal{H}_k$ with $\|f\|_{\mathcal{H}_k} \leq B$. Then the following statement holds with probability at least $1-\delta$, for all $x \in \mathcal{X}$ and $t \in \mathbb{N}$,*

$$|f(x) - \mu_{t-1}(x)| \leq \beta_t \sigma_{t-1}(x),$$

where $\beta_t = B + R\sqrt{2(\gamma_{t-1} + 1 + \log(1/\delta))}$ and γ_{t-1} is the maximum information gain after $(t-1)$ rounds.

Lemma 16 (Lemma 4 of Chowdhury and Gopalan [8])**.** *Let $\{z_1, \cdots, z_T\}$ be the points selected by arbitrary BO strategy. Then the following holds,*

$$\sum_{t=1}^{T} \sigma_{t-1}(z_t) \leq \sqrt{4(T+2)\gamma_T}$$

where γ_T is the maximum information gain after T rounds.

Throughout the remainder of the appendix, we use the notation $\mathcal{O}(t^{\xi})$, for arbitrarily small $\xi > 0$, to represent logarithmic factors in t.

Lemma 17 (Proposition 4 of Helin et al. [18])**.** *Let $\tilde{x}_1, ..., \tilde{x}_t$ be t uniform random samples from a hyperrectangle $\mathcal{X} \subseteq \mathbb{R}^d$. Define $h_t = \sup_{x\in\mathcal{X}} \inf_{i=1,...t} \|x - \tilde{x}_i\|$, then there exists t^* such that $\forall t \geq t^*$,*

$$\mathbb{E}[h_t] = \mathcal{O}(t^{-\frac{1}{d}+\xi}),$$

where $\xi > 0$ is an arbitrarily small positive constant.

In the following lemma, we establish a high-probability upper bound on the cumulative discrepancy between the random grid $\mathcal{X}_t$ and the search space $\mathcal{X}$.

Lemma 18. *Let $h_t = \sup_{x\in\mathcal{X}} \inf_{\tilde{x}_i\in\mathcal{X}_t} \|x - \tilde{x}_i\|$. For $\delta > 0$ small, with probability at least $1-\delta$, we have*

$$\sum_{t=1}^{T} h_t \leq \sum_{t=1}^{t^*} h_t + \frac{CT^{\frac{d-1}{d}+\xi}}{\delta},$$

for some $C > 0$ and $t^* \in \mathbb{N}$, with arbitrarily small $\xi > 0$.

Proof. From Lemma 17, we know that there exists $C > 0, \xi > 0$ and $t^* \in \mathbb{N}$ such that $\mathbb{E}[h_t] \leq Ct^{-\frac{1}{d}+\xi}$ for all $t \geq t^*$. From the Markov's inequality, we know that

$$\mathbb{P}\left[\sum_{t=t^*}^{T} h_t > \frac{CT^{\frac{d-1}{d}+\xi}}{\delta}\right] \leq \frac{\mathbb{E}\left[\sum_{t=t^*}^{T} h_t\right]}{CT^{\frac{d-1}{d}+\xi}}\delta \leq \delta$$

where the second inequality follows from $\mathbb{E}[\sum_{t=t^*}^{T} h_t] \leq CT^{\frac{d-1}{d}+\xi}$. □

A.2 GP-UCB/GP-TS WITH INEXACTNESS

We first provide justifications for the nonnegativity of the UCB acquisition function after the constant shift. To this end, we show the existence of a constant C such that $\alpha_t + C$ is non-negative on the search space $\mathcal{X}$.

Lemma 19. *For some $\delta > 0$, a constant shifted t^{th} acquisition function of the GP-UCB algorithm, given by*

$$\alpha_t(x) + C = \mu_{t-1}(x) + \beta_t \sigma_{t-1}(x) + C$$

is non-negative on $\mathcal{X}$ with probability $1 - \delta$, for some constant sequence $C \in \mathbb{R}$.

Proof. From Lemma 15, we know that

$$f(x) \leq \mu_{t-1}(x) + \beta_t \sigma_{t-1}(x),$$

for all $x \in \mathcal{X}$ with probability $1 - \delta$. Therefore,

$$0 \leq f(x) + |f(x)| \leq \mu_{t-1}(x) + \beta_t \sigma_{t-1}(x) + |f(x)|.$$

Since $\sup_x |f(x)| \leq \|f\|_{\mathcal{H}_k} \leq B$, we have that $\mu_{t-1}(x) + \beta_t \sigma_{t-1}(x) + B \geq 0$. □

We next provide justifications for the nonnegativity of the TS acquisition function after the constant shift. To this end, we show the existence of a constant C_t such that $\alpha_t + C_t$ is non-negative on the search space $\mathcal{X}_t$.

Lemma 20. *For some $\delta > 0$, the constant shifted t^{th} acquisition function of the robust GP-TS algorithm, given by*

$$\alpha_t(x) + C_t = f_t(x) + C_t$$

is non-negative on $\mathcal{X}_t$ with probability $1 - \delta$, for some sequence C_t.

Proof. From Lemma 5 of [8], conditioning on the history until time t, i.e., $\mathcal{H}_t := \{(x_1, y_1), \cdots, (x_t, y_t)\}$, we know that

$$\forall x \in \mathcal{X}_t, \quad |f_t(x) - \mu_{t-1}(x)| \leq \sqrt{2\log(|\mathcal{X}_t|/\delta)} s_{t-1}(x),$$

with probability $1 - \delta/2$. Therefore, from the definition $s_{t-1}(x) = \beta_t \sigma_{t-1}(\cdot) + v_t$, we have

$$\mu_{t-1}(x) - \sqrt{2\log(|\mathcal{X}_t|/\delta)}(\beta_T + v) \leq \mu_{t-1}(x) - \sqrt{2\log(|\mathcal{X}_t|/\delta)} s_{t-1}(x) \leq f_t(x), \tag{1}$$

where $v_t = \left(\frac{1}{\eta_t} - 1\right) B$. Using Lemma 15 with $\beta_t \sigma_t$ replaced with s_t, with probability at least $1 - \delta/2$

$$-B - (\beta_T + v) \leq f(x) - s_{t-1}(x) \leq \mu_{t-1}(x) \tag{2}$$

holds for all $x \in \mathcal{X}$. Combining (1) and (2), we observe that

$$f_t(x) + (1 + \sqrt{2\log(|\mathcal{X}_t|/\delta)})(\beta_T + v) + B \geq 0,$$

for all $x \in \mathcal{X}_t$ with probability at least $1 - \delta$. □

Theorem 21 (Restatement of Theorem 3)**.** *Under assumptions 12 and 13, suppose an objective function $f \in \mathcal{H}_k$ with $\|f\|_{\mathcal{H}_k} \leq B$. The inexact GP-UCB algorithm with $\beta_t = B + R\sqrt{2(\gamma_{t-1} + 1 + \log(1/\delta))}$ and worst-case accumulated inaccuracy M_T yields a cumulative regret bound of the form,*

$$R_T = \mathcal{O}\left(\gamma_T \sqrt{T} + M_T \sqrt{\gamma_T}\right),$$

with probability $1 - \delta$.

Proof. From Theorem 2 of Chowdhury and Gopalan [8], we know that

$$\mu_{t-1}(x^*) - \beta_t\sigma_{t-1}(x^*) \leq f(x^*) \leq \mu_{t-1}(x^*) + \beta_t\sigma_{t-1}(x^*),$$
$$\mu_{t-1}(x_t) - \beta_t\sigma_{t-1}(x_t) \leq f(x_t) \leq \mu_{t-1}(x_t) + \beta_t\sigma_{t-1}(x_t)$$

hold with probability $1-\delta$. Then

$$\begin{aligned} f(x^*) - f(x_t) &\leq \mu_{t-1}(x^*) + \beta_t\sigma_{t-1}(x^*) - \mu_{t-1}(x_t) + \beta_t\sigma_{t-1}(x_t) \\ &= \alpha_{t-1}(x^*) - \alpha_{t-1}(x_t) + 2\beta_t\sigma_{t-1}(x_t) \\ &= (1-\eta_t)\alpha_t^* + 2\beta_t\sigma_{t-1}(x_t), \end{aligned}$$

where the first equality follows from the definition of the UCB acquisition function, and the second equality is due to the fact that $\alpha_t(x_t) = \eta_t\alpha_t(x^*)$. Since $\mu_{t-1}(x) \leq f(x) + \beta_t\sigma_{t-1}(x)$, for all $x \in \mathcal{X}$, we conclude that, with probability $1-\delta$,

$$\alpha_t^* \leq \max_{x\in\mathcal{X}}[f(x) + 2\beta_T\sigma_{t-1}(x)] \leq B + 2\beta_T,$$

where the second inequality follows from the fact that $\sup_{x\in\mathcal{X}}|f(x)| \leq \|f\|_{\mathcal{H}_k} \leq B$, and for all $x \in \mathcal{X}, \sigma_{t-1}(x) \leq 1$. Therefore, we conclude that

$$R_T = \sum_{t=1}^{T} f(x^*) - f(x_t) \leq (B+2\beta_T)\sum_{t=1}^{T}(1-\eta_t) + 2\beta_T\sum_{t=1}^{T}\sigma_{t-1}(x_t) \leq (B+2\beta_T)\sum_{t=1}^{T}(1-\tilde{\eta}_t) + 2\beta_T\sum_{t=1}^{T}\sigma_{t-1}(x_t).$$

From Lemma 16, we arrive at the conclusion. □

A similar analysis can be established for the GP-TS algorithm with an enlarged variance $s_{t-1}^2(x) = (\beta_t\sigma_{t-1}(\cdot) + v_t)^2$, $v_t = (\frac{1}{\tilde{\eta}_t} - 1)B$. We omit the proof of the following statements as it substantially overlaps with the approach taken in [8] with a variance factor adjustment.

Theorem 22 (Restatement of Theorem 4)**.** *Under assumptions 12 and 13, suppose an objective function $f \in \mathcal{H}_k$ with $\|f\|_{\mathcal{H}_k} \leq B$. The inexact GP-TS algorithm with variance $s_{t-1}^2(x) = (\beta_t\sigma_{t-1}(\cdot) + \tilde{v}_t)^2$, $\tilde{v}_t = (\frac{1}{\tilde{\eta}_t} - 1)B$ and worst-case accumulated inaccuracy M_T yields a cumulative regret bound of the form*

$$R_T = \mathcal{O}\left(\gamma_T\sqrt{T\log T} + M_T\sqrt{\gamma_T\log T}\right),$$

with probability $1-\delta$.

A.3 RANDOM GRID SEARCH BASED GP-UCB

Theorem 23. *Under assumptions 12, 13 and 14, suppose $f \in \mathcal{H}_k$ with $\|f\|_{\mathcal{H}_k} \leq B$. With probability at least $1-\delta$, the random grid GP-UCB with $\beta_t = B + R\sqrt{2(\gamma_{t-1} + 1 + \log(2/\delta))}$ yields*

$$R_T = \mathcal{O}\left(\gamma_T\sqrt{T} + T^{\frac{d-1}{d}+\xi}\right),$$

for arbitrarily small $\xi > 0$.

Proof. Note that

$$\begin{aligned} R_T &= \sum_{t=1}^{T} f(x^*) - f(x_t) \\ &= \sum_{t=1}^{T} f(x^*) - f([x^*]_t) + f([x^*]_t) - f(x_t), \end{aligned}$$

where $[x^*]_t$ is the point closest to x^* in $\mathcal{X}_t$. By Lemma 15, for all possible realizations of $\mathcal{X}_1, \cdots, \mathcal{X}_T$, with a probability $1-\delta/2$, we have

$$\mu_{t-1}([x^*]_t) - \beta_t\sigma_{t-1}([x^*]_t) \leq f([x^*]_t) \leq \mu_{t-1}(x_t^*) + \beta_t\sigma_{t-1}([x^*]_t)$$
$$\mu_{t-1}(x_t) - \beta_t\sigma_{t-1}(x_t) \leq f(x_t) \leq \mu_{t-1}(x_t) + \beta_t\sigma_{t-1}(x_t).$$

From the definition of x_t, we know $\mu_{t-1}([x^*]_t) + \beta_t\sigma_{t-1}([x^*]_t) \leq \mu_{t-1}(x_t) + \beta_t\sigma_{t-1}(x_t)$. Therefore, we have

$$f([x^*]_t) - f(x_t) \leq 2\beta_t\sigma_{t-1}(x_t).$$

Furthermore, from Lemma 1 of [11], we know that $f(x^*) - f([x^*]_t) \leq C\|x^* - [x^*]_t\|$ for some constant $C > 0$. Since $\|x^* - [x^*]_t\| \leq h_t$ for $h_t := \sup_{x\in\mathcal{X}} \inf_{\bar{x}_i\in\mathcal{X}_t} \|x - \bar{x}_i\|$, we have

$$R_T \leq C\sum_{t=1}^{T} h_t + 2\beta_T \sum_{t=1}^{T} \sigma_{t-1}(x_t).$$

Then by Lemma 18, we know that with probability at least $1 - \delta/2$, $C\sum_{t=1}^{T} h_t = \mathcal{O}\left(T^{\frac{d-1}{d}+\xi}\right)$. Combined with Lemma 16 and invoking the union bound, we have the result. □

A.4 RANDOM GRID SEARCH BASED GP-TS

In this section, we list all additional definitions and lemmas we use in proofs for the regret bound of a random grid search based GP-TS. Since our approach closely follows to that of Chowdhury and Gopalan [8], many of the preliminary lemma we list here can be proven in an analogous fashion. We will adjust and restate the Lemma and proof if needed.

Definition 24. *We define the filtration $\mathcal{F}_t = \sigma\{(x_1, y_1, \mathcal{X}_2), \cdots, (x_t, y_t, \mathcal{X}_{t+1})\}$ as the σ-algebra generated by the collection of evaluation locations, function evaluations and search grids observed until time t. Note that conditional on $\mathcal{F}_t$, the search grid is no longer random.*

Lemma 25 (Lemma 5 of [8])**.** *For all $t \in \mathbb{N}$, assume $|\mathcal{X}_t| = ct$ with $c > 0$. Then*

$$\mathbb{P}\left[\forall x \in \mathcal{X}_t, |f_t(x) - \mu_{t-1}(x)| \leq \beta_t\sqrt{2\log(ct^3)}\sigma_{t-1}(x)\middle|\mathcal{F}_{t-1}\right] \geq 1 - 1/t^2,$$

for some constant $c > 0$, $\delta \in (0, 1)$ and $\beta_t = B + R\sqrt{2(\gamma_{t-1} + 1 + \log(3/\delta))}$.

We introduce some definitions here.

Definition 26. *$\forall t \geq 1$, $\tilde{c}_t = \sqrt{6\log t + 2\log c}$ and $c_t = \beta_t(1 + \tilde{c}_t)$.*

Definition 27. *We define the following two events:*

$$E^f(t) = \{\forall x \in \mathcal{X}, |\mu_{t-1}(x) - f(x)| \leq \beta_t\sigma_{t-1}(x)\}$$
$$E^{f_t}(t) = \{\forall x \in \mathcal{X}_t, |f_t(x) - \mu_{t-1}(x)| \leq \beta\tilde{c}_t\sigma_{t-1}(x)\}$$

Definition 28. *Given a grid $\mathcal{X}_t$, define the set of saturated points to be*

$$S_t \coloneqq \{x \in \mathcal{X}_t : \Delta_t(x) > c_t\sigma_{t-1}(x)\},$$

where $[x^*]_t$ is the point closest to x^* in $\mathcal{X}_t$ and $\Delta_t(x) := f([x^*]_t) - f(x)$. Notice that conditioning on $\mathcal{X}_t$, $[x^*]_t \in \mathcal{X}_t \setminus S_t$.

Lemma 29 (Lemma 6 of [8])**.** *Suppose $|\mathcal{X}_t| = ct,\ c > 0$. For $\delta \in (0, 1)$, following statements hold.*

- $\mathbb{P}[\forall\, t \geq 1, E^f(t)] \geq 1 - \delta/3$
- $\mathbb{P}[E^{f_t}(t)|\mathcal{F}_{t-1}] \geq 1 - 1/t^2$

Lemma 30 (Lemma 7 of Chowdhury and Gopalan [8])**.** *For any filtration $\mathcal{F}_{t-1}$ such that $E^f(t)$ is true,*

$$\mathbb{P}\left[f_t(x) > f(x)\middle|\mathcal{F}_{t-1}\right] \geq \eta \coloneqq \frac{1}{4e\sqrt{\pi}} > 0,$$

holds for any $x \in \mathcal{X}$.

Lemma 31 (Lemma 8 of Chowdhury and Gopalan [8])**.** *For any filtration $\mathcal{F}_{t-1}$ such that $E^f(t)$ is true,*

$$\mathbb{P}[x_t \in \mathcal{X}_t \setminus S_t|\mathcal{F}_{t-1}] \geq \eta - 1/t^2.$$

Lemma 32. *For any filtration $\mathcal{F}_{t-1}$ such that $E^f(t)$ is true, we have*

$$\mathbb{E}[\Delta_t(x_t)|\mathcal{F}_{t-1}] \leq \frac{11c_t}{\eta}\mathbb{E}[\sigma_{t-1}(x_t)|\mathcal{F}_{t-1}] + \frac{2B}{t^2},$$

where $\Delta_t(x_t) := f([x^*]_t) - f(x_t)$.

Proof. Given a grid $\mathcal{X}_t$, let $\bar{x}_t = \operatorname{argmin}_{x\in\mathcal{X}_t\setminus S_t} \sigma_{t-1}(x)$. From the law of total expectation and positivity of the σ_{t-1}, we have

$$\begin{aligned}\mathbb{E}[\sigma_{t-1}(x_t)|\mathcal{F}_{t-1}] &\geq \mathbb{E}[\sigma_{t-1}(x_t)|\mathcal{F}_{t-1}, x_t \in \mathcal{X}_t \setminus S_t]\mathbb{P}[x_t \in \mathcal{X}_t \setminus S_t|\mathcal{F}_{t-1}] \\ &\geq \mathbb{E}[\sigma_{t-1}(\bar{x}_t)|\mathcal{F}_{t-1}](\eta - 1/t^2), \end{aligned} \tag{3}$$

where the second inequality follows from the definition of $\bar{x}_t$ and Lemma 31.

To control $\Delta_t(x_t) = f([x^*]_t) - f(x_t)$, recall that if $E^f(t)$ and $E^{f_t}(t)$ are both true, we know that

$$\text{for all } x \in \mathcal{X}_t, \quad f_t(x) - c_t\sigma_{t-1}(x) \leq f(x) \leq f_t(x) + c_t\sigma_{t-1}(x). \tag{4}$$

Notice that

$$\begin{aligned}\Delta_t(x_t) &= f([x^*]_t) - f(x_t) \\ &= f([x^*]_t) - f(\bar{x}_t) + f(\bar{x}_t) - f(x_t) \\ &\leq \Delta_t(\bar{x}_t) + f_t(\bar{x}_t) + c_t\sigma_{t-1}(\bar{x}_t) - f_t(x_t) + c_t\sigma_{t-1}(x_t) \\ &\leq c_t(2\sigma_{t-1}(\bar{x}_t) + \sigma_{t-1}(x_t)) + f_t(\bar{x}_t) - f_t(x_t) \\ &\leq c_t(2\sigma_{t-1}(\bar{x}_t) + \sigma_{t-1}(x_t))\end{aligned}$$

where the first inequality is due to (4), the second inequality follow from the fact $\bar{x}_t \notin S_t$ and the last inequality comes from the definition of x_t.

Therefore, we have

$$\begin{aligned}\mathbb{E}[\Delta_t(x_t)|\mathcal{F}_{t-1}] &\leq 2c_t\mathbb{E}[\sigma_{t-1}(\bar{x}_t)|\mathcal{F}_{t-1}] + c_t\mathbb{E}[\sigma_{t-1}(x_t)|\mathcal{F}_{t-1}] + 2B\mathbb{P}\left[E^{f_t}(t)^c|H_{t-1}\right] \\ &\leq \frac{2c_t}{\eta - 1/t^2}\mathbb{E}[\sigma_{t-1}(x_t)|\mathcal{F}_{t-1}] + c_t\mathbb{E}[\sigma_{t-1}(x_t)|\mathcal{F}_{t-1}] + \frac{2B}{t^2} \\ &\leq \frac{11c_t}{\eta}\mathbb{E}[\sigma_{t-1}(x_t)|\mathcal{F}_{t-1}] + \frac{2B}{t^2},\end{aligned}$$

where we used the fact $\sup_{x\in cX}\Delta_t(x) \leq 2B$ which can be deduced from the fact $\sup|f(x)| \leq \|f\|_{\mathcal{H}_k} \leq B$ in the first inequality. The second inequality follows from the inequality in (3) and Lemma 25. The third inequality is due to the fact $\frac{1}{\eta - 1/t^2} \leq \frac{5}{\eta}$. □

Next, we define random variables and associated filtration to invoke concentration inequality for super-martingales.

Definition 33. *Let $Y_0 = 0$, and for all $t \in \{1, \cdots, T\}$,*

$$\begin{aligned}\bar{\Delta}_t(x_t) &= \Delta_t(x_t)\mathbb{I}\left\{E^f(t)\right\} \\ Z_t &= \bar{\Delta}_t(x_t) - \frac{11c_t}{\eta}\sigma_{t-1}(x_t) - \frac{2B}{t^2} \\ Y_t &= \sum_{s=1}^{t} Z_s\end{aligned}$$

From the definition, and by Lemma 32, we deduce the following result, which we formally state as lemma.

Lemma 34. *$(Y_t)_{t=0}^T$ is a super-martingale process with respect to filtration $\mathcal{F}_t$.*

Proof. It suffices to show that for all $t \in \{1, \cdots, T\}$ and any $\mathcal{F}_{t-1}$, $\mathbb{E}\left[Y_t - Y_{t-1} | \mathcal{F}_{t-1}\right] \leq 0$. Note that

$$\mathbb{E}\left[Y_t - Y_{t-1} | \mathcal{F}_{t-1}\right] = \mathbb{E}\left[Z_t | \mathcal{F}_{t-1}\right] = \mathbb{E}\left[\bar{\Delta}_t(x_t) | \mathcal{F}_{t-1}\right] - \frac{11c_t}{\eta}\mathbb{E}\left[\sigma_{t-1}(x_t) | \mathcal{F}_{t-1}\right] - \frac{2B}{t^2}. \tag{5}$$

If $E^t(t)$ is false, we have $\mathbb{E}\left[\bar{\Delta}_t(x_t) | \mathcal{F}_{t-1}\right] = 0$, which shows that (5) is less than or equal to zero. On the other hand, if $E^t(t)$ is true, from Lemma 32, we can again conclude that (5) is less than or equal to zero. □

Lemma 35. *Given any $\delta > 0$,*

$$\sum_{t=1}^{T} \Delta_t(x_t) \leq \frac{11c_T}{\eta}\sum_{t=1}^{T}\sigma_{t-1}(x_t) + \frac{2B\pi^2}{6} + \frac{4B + 11c_T}{\eta}\sqrt{2T\log(3/\delta)},$$

with probability at least $1 - 2\delta/3$.

Proof. By construction,

$$|Y_t - Y_{t-1}| = |Z_t| \leq \left|\bar{\Delta}_t(x_t)\right| + \frac{11c_t}{\eta}\sigma_{t-1}(x_t) + \frac{2B}{t^2} \leq 2B + \frac{11c_t}{\eta} + \frac{2B}{t^2} \leq \frac{4B + 11c_t}{\eta}$$

where the first inequality is due to the triangle inequality. The second inequality comes from the fact that $\left|\bar{\Delta}_t(x_t)\right| \leq 2\sup_{x\in\mathcal{X}}|f(x)| \leq 2\|f\|_{\mathcal{H}_k} \leq 2B$ and $\sigma_{t-1}(x) \leq 1$ for all $x \in \mathcal{X}$. The third inequality follows from $\eta \leq 1$. From the Azuma-Hoeffding inequality, with at least probability $1 - \delta/3$, we have

$$Y_T - Y_0 = \sum_{t=1}^{T}\bar{\Delta}_t(x_t) - \sum_{t=1}^{T}\frac{11c_t}{\eta}\sigma_{t-1}(x_t) - \sum_{t=1}^{T}\frac{2B}{t^2} \leq \sqrt{2\log(3/\delta)\sum_{t=1}^{T}\frac{(4B + 11c_t)^2}{\eta^2}}.$$

In other words, we have

$$\begin{aligned}\sum_{t=1}^{T}\bar{\Delta}_t(x_t) &\leq \sum_{t=1}^{T}\frac{11c_t}{\eta}\sigma_{t-1}(x_t) + \sum_{t=1}^{T}\frac{2B}{t^2} + \sqrt{2\log(3/\delta)\sum_{t=1}^{T}\frac{(4B + 11c_t)^2}{\eta^2}} \\ &\leq \frac{11c_T}{\eta}\sum_{t=1}^{T}\sigma_{t-1}(x_t) + \frac{2B\pi^2}{6} + \frac{4B + 11c_T}{\eta}\sqrt{2T\log(3/\delta)}\end{aligned}$$

with at least probability $1 - \delta/3$. From Lemma 29, we know $E^f(t)$ holds for all $t \geq 1$ with probability at least $1 - \delta/3$. In other words, by definition, $\Delta_t(x_t) = \bar{\Delta}_t(x_t)$ for all $t \geq 1$ with probability at least $1 - \delta/3$. Applying the union bound, we obtain the statement. □

Theorem 36 (Restatement of Theorem 8)**.** *Under assumptions 12, 13 and 14, suppose $f \in \mathcal{H}_k$ with $\|f\|_{\mathcal{H}_k} \leq B$, for some $B > 0$. With probability at least $1 - \delta$, the random grid GP-TS with $\beta_t = B + R\sqrt{2(\gamma_{t-1} + 1 + \log(3/\delta))}$ yields*

$$R_T = \mathcal{O}\left(\gamma_T\sqrt{T\log T} + T^{\frac{d-1}{d}+\xi}\right),$$

for arbitrarily small $\xi > 0$.

Proof. Note that

$$\begin{aligned}
R_T &= \sum_{t=1}^{T} f(x^*) - f(x_t) \\
&= \sum_{t=1}^{T} f(x^*) - f([x^*]_t) + f([x^*]_t) - f(x_t) \\
&= \sum_{t=1}^{T} f(x^*) - f([x^*]_t) + \sum_{t=1}^{T} \Delta_t(x_t) \\
&\leq C \sum_{t=1}^{T} \|x^* - [x^*]_t\|_2 + \sum_{t=1}^{T} \Delta_t(x_t) \\
&\leq C \sum_{t=1}^{T} h_t + \sum_{t=1}^{T} \Delta_t(x_t),
\end{aligned}$$

where the first inequality follows from the Lipschitzness of f, which was shown in Lemma 1 of [11] and the second inequality is due to the definition of fill distance $h_t = \sup_{x \in \mathcal{X}} \inf_{\tilde{x}_i \in \mathcal{X}_t} \|x - \tilde{x}_i\|$. From Lemma 18, we know that the leading term $C \sum_{t=1}^{T} h_t = \mathcal{O}\left(T^{\frac{d-1}{d}+\xi}\right)$. For the second term, note that from Lemma 35

$$\begin{aligned}
\sum_{t=1}^{T} \Delta_t(x_t) &\leq \frac{11 c_T}{\eta} \sum_{t=1}^{T} \sigma_{t-1}(x_t) + \frac{2B\pi^2}{6} + \frac{4B + 11c_T}{\eta} \sqrt{2T \log(3/\delta)} \\
&= \mathcal{O}\left(c_T \sqrt{4(T+2)\gamma_T} + c_T \sqrt{T}\right)
\end{aligned}$$

where the last equality is due to Lemma 16. From the definition of c_t, we know $c_T = \mathcal{O}(\sqrt{\gamma_T \log T})$. Since the leading term dominates, we have

$$\sum_{t=1}^{T} \Delta_t(x_t) = \mathcal{O}\left(\gamma_T \sqrt{T \log T}\right).$$

Applying union bound and combining everything, we get the result. □

B ADDITIONAL DETAILS OF EXPERIMENTS

B.1 EXPERIMENTAL SETUP OF REAL-WORLD AUTOML TASK

In the 11-dimensional real-world hyperparameter tuning task, we focus on improving the validation accuracy of the gradient boosting model (from the Python `sklearn` package) on the breast-cancer dataset (from the UCI machine learning repository). The 11 hyperparameters are:

- Loss, (string, "logloss" or "exponential").
- Learning rate, (float, (0, 1)).
- Number of estimators, (integer, [20, 200]).
- Fraction of samples to be used for fitting the individual base learners, (float, (0, 1)).
- Function to measure the quality of a split, (string, "friedman mse" or "squared error").
- Minimum number of samples required to split an internal node, (integer, [2, 10]).
- Minimum number of samples required to be at a leaf node, (integer, [1, 10]).
- Minimum weighted fraction of the sum total of weights, (float, (0, 0.5)).
- Maximum depth of the individual regression estimators, (integer, [1, 10]).
- Number of features to consider when looking for the best split, (float, "sqrt" or "log2").

- Maximum number of leaf nodes in best-first fashion, (integer, [2, 10]).

For integer-valued parameters, we conducted optimization on the continuous domain and rounded to the nearest integer value.

B.2 ABLATION STUDIES

B.2.1 Experiments with Different Grid Sizes

We have further added a numerical experiment investigating the effect of linearly varying grid size versus a fixed grid size of 100 in the 11-dimensional machine learning hyperparameter tuning problem. From Figure 4, it can be seen that the linearly varying grid size performs slightly better than the fixed grid size in terms of the cumulative regret plot.

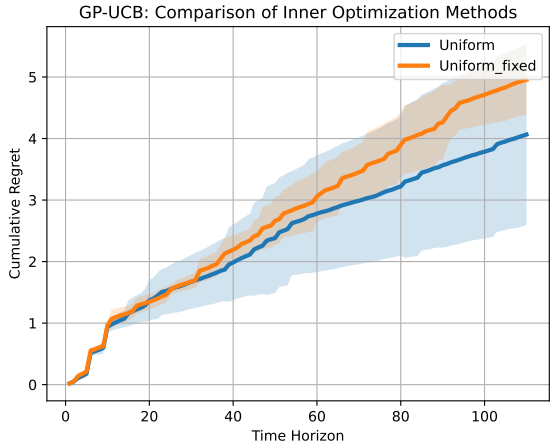

Figure 4: Performances with uniform random and uniform fixed grid sizes.

B.2.2 Experiments with a Smaller Initial Design Points

We further investigated the impact of a smaller initial design. In Section 5, we used $n = 10d$; here, we repeat those simulations with $n = 5d$. Results are shown in Figure 6 and Figure 5. Once again, the random grid search GP-UCB achieved competitive performance while substantially improving computational efficiency.

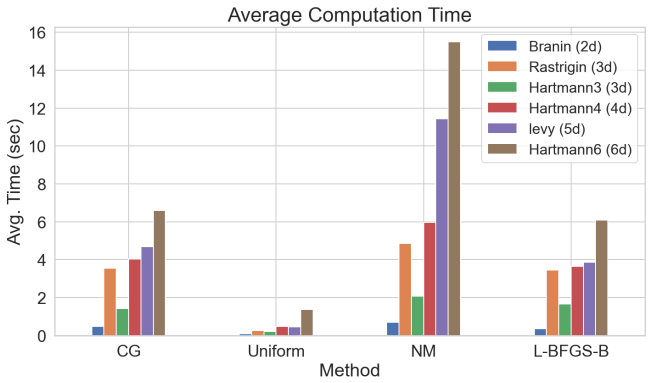

Figure 5: Average computational time of different acquisition optimization methods with a smaller initial design points ($n = 5d$, in the manuscript, we considered $n = 10d$).

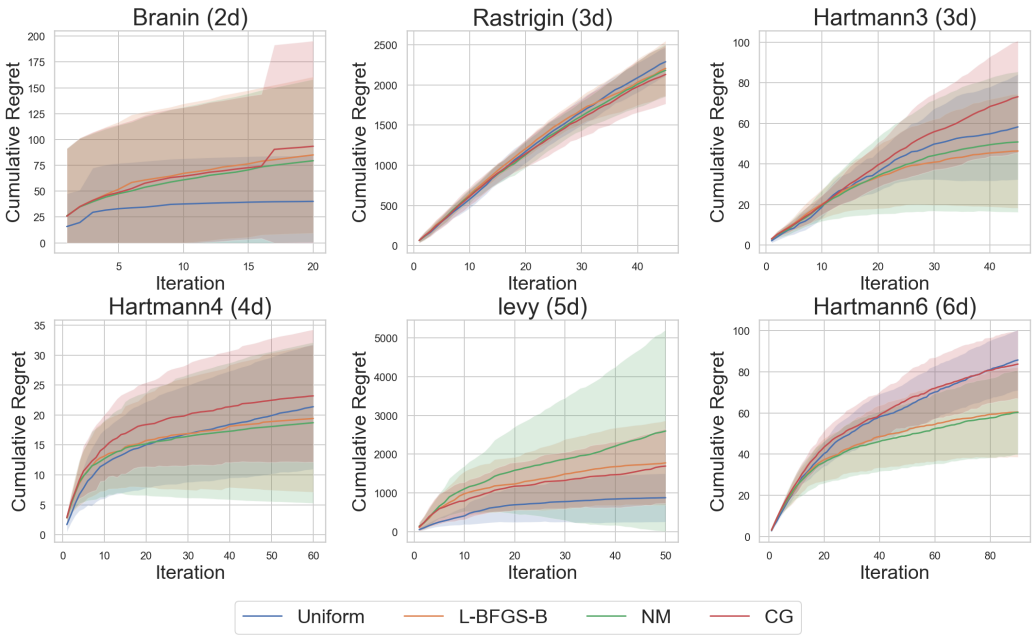

Figure 6: Cumulative regret of different acquisition optimization methods with a smaller initial design points ($n = 5d$, in the manuscript, we considered $n = 10d$).