Abstract

We consider the problem of optimizing a black-box function $f : \mathcal{X} \to \mathbb{R}$ under the assumption that it has bounded norm in the Reproducing Kernel Hilbert Space (RKHS) associated with a given kernel $K$. This problem is known to have an agnostic Gaussian Process (GP) bandit interpretation in which an appropriately constructed GP surrogate model with kernel $K$ is used to obtain an upper confidence bound (UCB) algorithm. In this paper, we propose a new algorithm (LP-GP-UCB) where the usual GP surrogate model is augmented with Local Polynomial (LP) estimators to construct a multi-scale upper confidence bound guiding the search for the optimizer. We analyze this algorithm and derive high probability bounds on its simple and cumulative regret for a practically relevant class of kernels called the Matérn family $(K_{\nu})_{\nu>0}$. We show that for certain ranges of $\nu$, the algorithm achieves near-optimal bounds on simple and cumulative regrets, matching the algorithm-independent lower bounds up to poly-logarithmic factors, and thus closing the large gap between the existing upper and lower bounds for these values of $\nu$.

Keywords: RKHS, GP Bandits, Local Polynomial Estimators

1. Introduction

Consider the problem of maximizing a black-box objective function $f : \mathcal{X} \to \mathbb{R}$ which can be accessed through a noisy zero-order oracle which upon querying the objective function $f$ at an arbitrary point $x \in \mathcal{X}$ provides an observation $y_x = f(x) + \eta_x$. Our goal is to design a query-point selection strategy $\mathcal{A}$, which can efficiently learn about a maximizer $x^*$ of $f$ given a finite query (or evaluation) budget $n$. After the evaluation budget is exhausted, the algorithm $\mathcal{A}$ must recommend a point, denoted by $z_n$. Two commonly used measures for the performance of a sampling strategy are its simple regret $S_n = f(x^*) - f(z_n)$ and its cumulative regret $R_n = \sum_{t=1}^n f(x^*) - f(x_t)$.

This problem is intractable without any regularity assumptions on the objective function $f$. In this paper, we assume that $f$ has a bounded norm in the RKHS associated with a kernel $K$ (denoted by $\mathcal{H}_K$). This formulation is referred to as the agnostic Gaussian Process (GP) bandit problem in literature (Srinivas et al., 2012) where a GP surrogate can be used to derive an upper confidence bound (UCB) on the function which then is utilized to guide the search for the optimizer. We propose a new algorithm which exploits smoothness properties...
of functions lying in the \( \mathcal{H}_K \) of commonly used kernels, to obtain tighter bounds on both \( S_n \) and \( R_n \). Our work is motivated by a large gap, in the agnostic GP bandits setting, between the best known upper bounds and the algorithm-independent lower bounds on the regret for the Matérn family, the most commonly used family of kernels (Scarlett et al., 2017).

**Notations.** Recall that \( f \) is the objective function mapping \( \mathcal{X} = [0, 1]^d \) to \( \mathcal{Y} = \mathbb{R} \), and can be accessed through a noisy evaluations of the form \( y = f(x) + \eta \), where the additive noise \( \eta \) is assumed to be \( \sigma \) sub-Gaussian. We will use the term cell to refer to subsets \( E \) of \( \mathcal{X} \) of the form \( E = \{ x \in \mathcal{X} : \| x - x_E \|_\infty \leq r_E / 2 \} \) for some \( x_E \in \mathcal{X} \) and \( r_E > 0 \). The terms \( x_E \) and \( r_E \) shall be referred to as the center and the side-length of \( E \). Given a kernel \( K \), we shall use the term \( \mathcal{H}_K \) and \( \| \cdot \|_K \) to denote the associated RKHS and the RKHS norm respectively. An important quantity of interest associated with \( K \) is the maximum information gain \( \gamma_n := \max_{S \subseteq \mathcal{X}, |S| = n} I(y_S; f) \), where \( I \) denotes the mutual information between a Gaussian process \( f \) and noisy observation vector \( y_S \). For \( k \in \mathbb{N} \) and \( 0 < \alpha \leq 1 \), we use \( \mathcal{C}^{k,\alpha} \) to denote the Hölder space with parameters \( k \) and \( \alpha \) (see App. A of Shekhar and Javidi (2020) for definition). Let \( \mathcal{D} = \{ (x_i, y_i) : 1 \leq i \leq m \} \subset \mathcal{X} \times \mathcal{Y} \) denote a labelled data set, and introduce \( \mathcal{D}_X := \{ x : \exists y \in \mathcal{Y}, (x, y) \in \mathcal{D} \} \) and \( \mathcal{D}_Y := \{ y : \exists x \in \mathcal{X}, (x, y) \in \mathcal{D} \} \). For a cell \( E \subset \mathcal{X} \), we use \( \mathcal{D}(E) \) to denote the subset of \( \mathcal{D} \) with \( x_i \in E \). The sets \( \mathcal{D}_X(E) \) and \( \mathcal{D}_Y(E) \) are also defined in an analogous manner. For positive integers \( k \) and \( D \), we use \( \mathcal{P}_D^k \) to denote the set of all polynomials in \( D \) variables of degree \( k \). Given \( g : \mathcal{X} \mapsto \mathbb{R} \) and \( E \subset \mathcal{X} \), we define \( \Phi_k(g, E) := \inf_{p \in \mathcal{P}_D^k} \sup_{x \in E} |g(x) - p(x)| \) as the smallest uniform approximation error of \( g \) with \( p \) in \( \mathcal{P}_D^k \).

2. Main Results

We first formally state the main assumptions required in our analysis.

**Assumption 1.** We make the following assumptions: *(A1.1)* \( f \in \mathcal{H}_K \) for some known kernel \( K \) and furthermore \( \| f \|_K \leq B \) for some known constant \( B > 0 \). *(A1.2)* \( f \in \mathcal{C}^{k,\alpha} \) for \( k \in \mathbb{N} \cup \{0\} \) and \( \alpha \in (0, 1] \) with \( \| f \|_{\mathcal{C}^{k,\alpha}} \leq L \) for some known \( L > 0 \). *(A1.3)* the observation noise \( \eta_i \geq 0 \) are i.i.d and \( \sigma^2 \)-sub-Gaussian for some known constant \( \sigma^2 > 0 \).

Assumption (A1.1) and (A1.3) are standard, and have been used in most prior works in agnostic GP bandits. For the case of Matérn kernels, we next show that Assumption (A1.2) follows as a consequence of (A1.1).

**Proposition 1.** If \( f \in \mathcal{H}_K \) with \( K \in \{ K_\nu : \nu > 0 \} \), then there exists constants \( 0 < C_1, C_2 < \infty \) such that we have \( \| f \|_{\mathcal{C}^{k,\alpha}} \leq C_1 C_2 \| f \|_K \) where \( k \in \mathbb{N} \) and \( \alpha \in (0, 1] \) such that \( k + \alpha = \nu \).

The proof of this statement relies on the norm-equivalence of RKHS with certain fractional Sobolev spaces followed by an application of a Sobolev embedding theorem. Details are in App. B.1 of (Shekhar and Javidi, 2020). The constant \( C_1 \) in the proof outline of Prop. 1 can be computed in terms of the parameters \( \nu \) and \( D \). Appropriate bounds on the other term \( C_2 \) also exist for some cases, see for example (Talenti, 1976). For simplicity however, we make the following assumption.

**Assumption 2.** We assume that \( n \) is large enough to ensure that \( C_2 \leq \log(n) \) for \( K = K_\nu \) where \( C_2 \) is the constant introduced in Proposition 1.

We now describe the steps of our algorithm LP-GP-UCB (pseudo-code in Algorithm 1).
Algorithm 1: LP-GP-UCB

Input: $n$, $K$, $B$, $(k, \alpha)$, $L$, $\rho_0$.

1. Initialize: $t = 1$, $n_e = 0$, $\mathcal{P}_t = \{X\}$, $u_X^{(0)} = +\infty$, $\mathcal{D}_t = \emptyset$.
2. while $n_e < n$ do
   3.   for $E \in \mathcal{P}_t$ do
      4.     Draw $x_{t,E} \sim \text{Unif}(E)$, $U_{t,E} = \min\{u_E^{(0)}, u_{t,E}^{(1)}, u_{t,E}^{(2)}\}$
   5.   end
   6.   $E_t \in \arg\max_{E \in \mathcal{P}_t} U_{t,E}$, $x_t = x_{t,E_t}$
   7.   $Q_1 = Q_2 = \{E_t\}$
   8.   if $\beta_n \sigma_t(x_t) < L(\sqrt{Dr_E})^{\alpha_1}$ AND $r_{E_t} \geq \rho_0$ then
      9.      $\text{val} = u_{t,E_t}^{(1)}$, $\text{flag} = 1$, $Q_2 = \text{ExpandAndBound}(E_t, \text{flag}, \text{val}, \mathcal{D}_t, \ldots)$
   10.  else if $b_l(E_t) \leq L(\sqrt{Dr_E})^{\alpha_1}$ AND $r_{E_t} \geq \rho_0$ then
       11.     $\text{val} = u_{t,E_t}^{(2)}$, $\text{flag} = 1$, $Q_2 = \text{ExpandAndBound}(E_t, \text{flag}, \mathcal{D}_t, \ldots)$
   12.  else if $b_l(E_t) \leq L(\sqrt{Dr_E})^{k+\alpha}$ AND $r_{E_t} \in [\frac{1}{2}, \rho_0)$ then
       13.     $\text{val} = +\infty$, $\text{flag} = 2$, $Q_2 = \text{ExpandAndBound}(E_t, \text{flag}, \mathcal{D}_t, \ldots)$
   14.  else
       15.     Observe $y_t = f(x_t) + \eta_t$, $\mu_t, \sigma_t$, $n_e \leftarrow n_e + 1$,
                $\mathcal{D}_t \leftarrow \mathcal{D}_t \cup \{(x_t, y_t)\}$
       16.  end
   17.  $\mathcal{P}_t \leftarrow (\mathcal{P}_t \setminus Q_1) \cup Q_2$, $t \leftarrow t + 1$
end
Output: $z_n$ using Recommend function (Def.1).

Inputs. Algorithm 1 takes in as inputs the query (or evaluation) budget $n$, the kernel $K$, the parameter $B$ which is a bound on $\|f\|_K$, the noise parameter $\sigma$, an integer $k$ and an $\alpha \in (0, 1]$ for the local polynomial estimator, the parameter $L$ which is an upper bound on $\|f\|_{c,k,\alpha}$ and a real-number $\rho_0 \geq (\gamma_n/\sqrt{LnD})^{1/\alpha_1}$ for $\alpha_1 := \max\{\alpha, \min(1, k)\}$.

Algorithm Outline. The LP-GP-UCB algorithm maintains a partition, $\mathcal{P}_t$, of the domain $X$ at any time $t$, and to each cell $E$ in the partition, it assigns a term $u_E^{(0)}$ which is an upper bound on the maximum $f$ value in the cell, calculated using local estimates based on prior observations. At $t = 1$, $\mathcal{P}_t$ is initialized as $\{X\}$ and $u_X^{(0)}$ is set to $+\infty$. As new cell $E$ is added to $\mathcal{P}_t$, the value of $u_E^{(0)}$ are decided by the ExpandAndBound algorithm (pseudo-code in Algorithm 2). For every $t \geq 1$, the LP-GP-UCB algorithm loops through all the cells in $\mathcal{P}_t$, and constructs a UCB denoted by $U_{t,E}$, by taking the minimum of $u_E^{(0)}$ and two other terms: $u_{t,E}^{(1)}$ and $u_{t,E}^{(2)}$ defined as

$$u_{t,E}^{(1)} = \mu_t(x_{t,E}) + \beta_n \sigma_t(x_{t,E}) + L(\sqrt{Dr_E})^{\alpha_1}, \quad u_{t,E}^{(2)} = \hat{\mu}_t(E) + b_l(E) + L(\sqrt{Dr_E})^{\alpha_1},$$

where $x_{t,E}$ is a point drawn uniformly from $E$ and $\mu_t$ and $\sigma_t$ are the posterior mean and variance of the surrogate GP model, and $\beta_t$ is a factor defined in Lemma 2 in App. C.1 of Shekhar and Javidi (2020). The term $\hat{\mu}_t(E)$ is the empirical estimate of $\hat{f}_E$, the average $f$ value in the cell $E$, and $b_l(E) = \sqrt{2\log(n/\delta)/\rho_{E}}$ is the length of the confidence interval of $\hat{f}_E$. 

Improved Regret Bounds Agnostic GP Bandits
(see Lemma 3 in App.C.1 of Shekhar and Javidi (2020) for details). Then, the algorithm selects a candidate cell \( E_t \) and the corresponding point \( x_t \) with the largest value of \( U_{t,E} \), and decides to either expand the partition or evaluate the function at the point \( x_t \). When the budget \( n \) is exhausted, it recommends a point \( z_n \) according to the rule \texttt{Recommend} described in Definition 1 below.

**Definition 1 (Recommend).** Suppose the algorithm stops in round \( t_n \) and let \( T \subset \{1, 2, \ldots, t_n\} \) denote the set of times at which Algorithm 1 performed function evaluations (note that \( |T| = n \)). Define \( E_n = \arg \min_{E \in \mathcal{P}} r_E \), and \( \xi := \min_{t \in T} \beta_t \sigma_t(x_t) \). If \( L(\sqrt{D} r_{E_n})^\alpha \leq \xi \), then return \( z_n = x_{E_n} \), where \( x_{E_n} \) is the center of the cell \( E_n \). Else, return \( z_n = x_\tau \) where \( \tau := \arg \min_{t \in T} \beta_t \sigma_t(x_t) \).

The update of the partition \( \mathcal{P}_t \) in Algorithm 1 requires calls to the \texttt{ExpandAndBound} algorithm (pseudo-code in Algorithm 2). This function takes as inputs a cell \( E \), data \( D \), an integer \( k \), and a confidence parameter \( \delta > 0 \), a variable flag with values in \{1, 2\}, a positive quantity val along with the terms \( k, \alpha, B \) and \( L \), and outputs \( \mathcal{Q}_2 \), which is a partition of the cell \( E \), and assigns an upper bound \((u_0^F)\) on the value of \( f \) in every \( F \in \mathcal{Q}_2 \). \( \mathcal{Q}_2 \) is constructed in \texttt{ExpandAndBound} by a call \texttt{Partition} operation which we define next.

**Definition 2 (Partition).** Given a cell \( E = \times_{i=1}^D [a_i, b_i] \subset X \), the function call \texttt{Partition}(\( E, r \)) for an some \( r < r_E \) returns a partition of \( E \) of cardinality \( \lceil r_E / r \rceil^D \), consisting of sets of the form \( F = \times_{i=1}^D [\tilde{a}_i, \min\{\tilde{a}_i + r, b_i\}] \), where \( \tilde{a}_i = a_i + lr \) for \( l \in \{0, 1, \ldots, \lceil r_E / r \rceil \} \).

To complete the description of the \texttt{ExpandAndBound} algorithm, we next describe the construction of local polynomial estimators, and introduce the functions \texttt{LocalPoly} (in Def. 3) and \texttt{MaxErr} (in Def. 4) which are called by \texttt{ExpandAndBound}.

Given a cell \( E \subset X \) and a point \( z \in E \), we define the LP estimator at \( z \) as \( \hat{f}_E(z, \bar{w}) = \sum_{x \in D(E)} w_x y_x \) where \( \bar{w} = \{w_x : x \in D(E)\} \) are defined as the solution of the following problem (Nemirovski, 2000, Eq. (1.36)):

\[
\min_{\bar{w} = \{v_x : x \in D(E)^\mathcal{X}\}} \sum_{x \in D(E)} |v_x|^2 \text{ s.t. } p(z) = \sum_{x \in D(E)} v_x p(x) \quad \forall p \in \mathcal{P}_D. \quad (LP)
\]

If the number of data points in the cell \( E \), i.e., \( |D(E)^\mathcal{X}| \), is larger than \((k + 2)^D\), then a unique solution to the problem \( (LP) \) is guaranteed to exist. We now introduce the \texttt{LocalPoly} function used in \texttt{ExpandAndBound}.

**Definition 3 (LocalPoly).** Given a cell \( E \), a point \( x \in E \), the function \texttt{LocalPoly} returns the estimated function value \( \hat{f}_E(x, \bar{w}) \) at \( x \), calculated according to the formula stated above.
If $n_E \equiv |D^{(E)}| > (k + 2)^D$, the weights $\bar{w}$ are the solution to (LP), while if $n_E < (k + 2)^D$, the weights are set as $w_x = 1/n_E$ for all $x \in D^{(E)}$.

We next state a result which bounds the estimation error between $\hat{f}_E(x, \bar{w})$ and $f(x)$.

**Lemma 1.** (Nemirovski, 2000, Prop. 1.3.1) Given a labelled dataset $D$, a cell $E \subset \mathcal{X}$, and a point $x_E \in E$, let $\bar{w}_E$ represent the unique solution of the problem (LP) under the assumption that $n_E \equiv |D^{(E)}| \geq (k + 2)^D$. Then, assuming that the observation noise is $\sigma$-subgaussian, for any $\delta > 0$, we have with probability at least $1 - \delta$:

$$|\hat{f}_E(x_E, \bar{w}_E) - f(x_E)| \leq (1 + \|\bar{w}_E\|_1) \Phi_k(f, E) + \sigma\|\bar{w}_E\|_2 \sqrt{2 \log(2/\delta)}.$$  

**Remark 1.** Recall that term $\Phi_k(f, E) := \inf_{p \in P_D^k} \sup_{x \in E} |f(x) - p(x)|$ depends on how well elements of $\mathcal{H}_K$ can be approximated by polynomials in $P_D^k$. For functions $f$ with $\|f\|_{C_{k, \alpha}} \leq L$, it is known that we have $\Phi_k(f, E) \leq L(\sqrt{D_E})^{k+\alpha}$. Using this, we can construct an upper bound on the first term on the RHS of (1), denoted by $e_D(E, \bar{w}_E, K, B) := (1 + \|\bar{w}_E\|_1)L(\sqrt{D_E})^{k+\alpha}$, from the information available to the algorithm. The second term in the RHS of (1), which we shall denote by $e_S(\bar{w}_E, \sigma, \delta)$, depends only on the data and known terms ($\sigma$ and $\delta$) and thus can be computed as well.

The last definition required is the MaxErr operation, which computes the maximum estimation error when a LP estimator is used to estimate $f$ in a cell $E$.

**Definition 4 (MaxErr).** The function MaxErr takes $E$, $D$, $K$, $B$, $k$, $\delta$ and $\sigma$ as inputs, and returns $\text{err}$ defined as $\text{err} := \max_{x \in E} e_D(E, \bar{w}_{E,x}, K, B) + e_S(\bar{w}_{E,x}, \sigma, \delta)$, where $e_S$ and $e_D$ were introduced in Remark 1, and $\bar{w}_{E,x}$ denotes the solution to (LP) at $x \in E$.

This completes our description of LP-GP-UCB algorithm. The reader is referred to App. A.3 of (Shekhar and Javidi, 2020) for some additional details.

### 3. Regret Analysis of LP-GP-UCB

We now state the main result of this section which provides high probability regret bounds for LP-GP-UCB algorithm for Matérn kernels.

**Theorem 1.** Suppose Assumptions 1 and 2 hold, and Algorithm 1 is run with budget $n$, $k = \lceil \nu \rceil - 1$, $\alpha = \nu - k$, $L = BC_1 \log n$ and other inputs as described in Sec. 2. Then the following statement is true with probability at least $1 - \delta$, for $n$ large enough:

$$S_n = \tilde{O} \left( \min \left( n^{\frac{1}{2} - \frac{D(D+1)}{2 + D(D+1)n}}, n^{-a_\nu} \right) \right), \quad \text{and} \quad R_n = \tilde{O} \left( \min \left( n^{\frac{1}{2} + \frac{D(D+1)}{2 + D(D+1)n}}, n^{b_\nu} \right) \right).$$  

The exponent $a_\nu$ is $\alpha_1/(2\alpha_1 + D)$ for $\nu > D(D+1)/2$ and $\nu/(2\nu + D)$ otherwise, while the exponent $b_\nu$ is $(D + \alpha_1)/(D + 2\alpha_1)$ for $\nu > D(D+1)/2$ and $(2\nu - \alpha_1 + D)/(2\nu + D)$ otherwise. Recall that $\alpha_1 = \max\{\alpha, \min\{1,k\}\}$.

The proof of this statement is given in App.C.3 of (Shekhar and Javidi, 2020). Note that under the condition $\nu \leq D(D+1)$, the upper bound on $\gamma_n$ is $\Omega(\sqrt{n})$. In this parameter range, the LP-GP-UCB algorithm achieves near-optimal rates for $S_n$, thus closing the large gap between the upper and lower bounds in the literature. However, the improvement
achieved for $R_n$ is not as significant as that of $S_n$. The main reason is that for $k \geq 1$, the LP-GP-UCB algorithm is more exploratory as it performs many function evaluations before expanding cells with radius smaller than $r_0$. While this ensures that the algorithm can find at least one good point resulting in small $S_n$, due to insufficient exploitation, the bound on $R_n$ suffers. A similar trade-off between obtaining tight bounds for both $S_n$ and $R_n$ occurs in some other bandit problems as well; see for example (Bubeck et al., 2011, §3).

In Proposition 2 next, we show that by using a zeroth degree LP estimators, we can obtain tighter control over $R_n$ for Matérn kernels with $\nu > 1$.

**Proposition 2.** For $f \in \mathcal{H}_{K_\nu}$ with $\nu > 1$, the LP-GP-UCB algorithm with $k = 0$ achieves the following regret bounds (in the display below we have $c_\nu = D(D+3)/(4\nu + D(D+5))$):

$$S_n = \tilde{O} \left( \min \left( n^{-1/2+c_\nu}, n^{-1/(D+2)} \right) \right), \quad \text{and} \quad R_n = \tilde{O} \left( \min \left( n^{1/2+c_\nu}, n^{(D+1)/(D+2)} \right) \right).$$

Combined with the regret bounds derived in Theorem 1, the above result implies that the LP-GP-UCB algorithm (with appropriate choice of $k$) achieves improved regret bounds for the Matérn family of kernels for all values of $\nu$ and $D$.

### 4. Summary of Improvements

To discuss the improvements in regret bounds, we introduce the notations $I_0 = (0, 1], I_1 = (1, D(D+1)/2], I_2 = (D(D+1)/2, D^2+5D+12)/4$ and $I_3 = (e, \infty)$ where $e = \max\{D(D+1)/2, (D^2+5D+12)/4\}$.

Note that $I_2$ is non-empty only for $D \leq 5$. We then have the following as a consequence of Theorem 1 and Prop. 2:

- **Simple Regret.** The best bounds on $S_n$ is $\tilde{O} \left( n^{-\nu/(2+D)} \right)$ for $\nu \in I_0 \cup I_1$ and $\tilde{O} \left( n^{-1/(D+2)} \right)$ for $\nu \in I_2$, both of which are achieved by the LP-GP-UCB algorithm with $k = \lceil \nu \rceil - 1$, i.e., (2). For $\nu \in I_3$, we have $S_n = \tilde{O} \left( n^{-\nu/2} \right)$ for $c_\nu = D(D+3)/(4\nu + 2D+5)$ which is achieved by LP-GP-UCB with $k = 0$. Note that for $\nu \in I_0 \cup I_1$, the regret bound achieved by LP-GP-UCB algorithm is near-optimal, i.e., it matches the lower bound up to poly-log factors. This closes the large gap between the upper and lower bounds for existing algorithms, as in this parameter regime the existing bounds on $S_n$ are $O(1)$ due to $\gamma_n$ being $\Omega(\sqrt{n})$.

- **Cumulative Regret.** The LP-GP-UCB algorithm with $k = 0$ achieves tighter control over $R_n$ for all values of $\nu$. In particular for $\nu \in I_0$ we have $R_n = \tilde{O} \left( n^{(\nu+D)/(2\nu+D)} \right)$, which matches the lower bound of (Scarlett et al., 2017) up to poly-log factors. For $\nu \in I_1 \cup I_2$, $R_n$ is $\tilde{O} \left( n^{(D+1)/(D+2)} \right)$ while for $\nu \in I_3$ it is $O \left( n^{1/2+c_\nu} \right)$ with $c_\nu$ defined in Prop. 2.

The cumulative regret bounds discussed above improve upon the state-of-the-art for all values of $\nu, D$. More specifically, (i) since $\gamma_n$ based bounds are known only for $\nu > 1$, our results provide the first explicit (and near-optimal) regret bounds of Matérn kernels with $\nu \in I_0 = (0, 1]$, and (ii) for $\nu > 1$ our results improve upon the best known bounds of $\tilde{O} \left( n^{\nu/2} \right)$ with $c_\nu = D(D+3) + 2\nu/(D+4)$, recently derived by (Janz et al., 2020).

From the above discussion, we can conclude that the LP-GP-UCB algorithm achieves tighter control over both simple and cumulative regret for all values of $\nu > 0$ than existing results in literature.
References


