Improved Regret Bounds for Agnostic Gaussian Process Bandits using Local Polynomial Estimators

Shubhanshu Shekhar SHSHEKHA@ENG.UCSD.EDU
ECE Department
University of California, San Diego

Tara Javidi TJAVIDI@ENG.UCSD.EDU
ECE Department
University of California, San Diego

Abstract

We consider the problem of optimizing a black-box function $f : \mathcal{X} \mapsto \mathbb{R}$ 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 ν, 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 ν.

Keywords: RKHS, GP Bandits, Local Polynomial Estimators

1. Introduction

Consider the problem of maximizing a black-box objective function $f : \mathcal{X} \mapsto \mathbb{R}$ 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* $\mathcal{S}_n$ and its *cumulative regret* $\mathcal{R}_n$ defined as $\mathcal{S}_n = f(x^*) - f(z_n)$ and $\mathcal{R}_n = \sum_{t=1}^n f(x^*) - f(x_t)$.

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 $\mathcal{S}_n$ and $\mathcal{R}_n$. 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 η is assumed to be σ 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* $\gamma_n \coloneqq \max_{S \subset \mathcal{X}, |S|=n} I(y_S; f)$, 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 α (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 $\mathcal{D}_\mathcal{X} \coloneqq \{x \ : \exists y \in \mathcal{Y}, (x,y) \in \mathcal{D}\}$ and $\mathcal{D}_\mathcal{Y} \coloneqq \{y \ : \ \exists x \in \mathcal{X}, (x,y) \in \mathcal{D}\}$. 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 $\mathcal{D}^{(E)}_\mathcal{X}$ and $\mathcal{D}^{(E)}_\mathcal{Y}$ are also defined in an analogous manner. For positive integers k and D, we use $\mathcal{P}^k_D$ 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 $\Phi_k(g, E) \coloneqq \inf_{p \in \mathcal{P}^k_D} \sup_{x \in E} |g(x) - p(x)|$ as the smallest uniform approximation error of g with p in $\mathcal{P}^k_D$.

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 $(\eta_t)_{t \geq 0}$ are i.i.d and σ^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 ν 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, α), L, ρ_0.

Initialize: $t = 1$, $n_e = 0$, $\mathcal{P}_t = \{\mathcal{X}\}$, $u^{(0)}_{\mathcal{X}} = +\infty$, $\mathcal{D}_t = \emptyset$;

while $n_e < n$ **do**

  for $E \in \mathcal{P}_t$ **do**

    Draw $x_{t,E} \sim \texttt{Unif}(E)$, $\quad U_{t,E} = \min\{u^{(0)}_E,\ u^{(1)}_{t,E},\ u^{(2)}_{t,E}\}$

  end

  $E_t \in \arg\max_{E \in \mathcal{P}_t} U_{t,E}$, $\quad x_t = x_{t,E_t}$

  $\mathcal{Q}_1 = \mathcal{Q}_2 = \{E_t\}$

  if $\beta_n \sigma_t(x_t) < L(\sqrt{D} r_E)^{\alpha_1}$ *AND* $r_{E_t} \geq \rho_0$ **then**

    $\texttt{val} = u^{(1)}_{t,E_t}$, $\quad \texttt{flag} = 1$, $\quad \mathcal{Q}_2 = \texttt{ExpandAndBound}(E_t, \texttt{flag}, \texttt{val}, \mathcal{D}_t, \ldots)$

  else if $b_t(E_t) \leq L(\sqrt{D} r_E)^{\alpha_1}$ *AND* $r_{E_t} \geq \rho_0$ **then**

    $\texttt{val} = u^{(2)}_{t,E_t}$, $\quad \texttt{flag} = 1$, $\quad \mathcal{Q}_2 = \texttt{ExpandAndBound}(E_t, \texttt{flag}, \texttt{val}, \mathcal{D}_t, \ldots)$

  else if $b_t(E_t) \leq L(\sqrt{D} r_E)^{k+\alpha}$ *AND* $r_{E_t} \in [\frac{1}{n}, \rho_0)$ **then**

    $\texttt{val} = +\infty$, $\quad \texttt{flag} = 2$, $\quad \mathcal{Q}_2 = \texttt{ExpandAndBound}(E_t, \texttt{flag}, \texttt{val}, \mathcal{D}_t, \ldots)$

  else

    Observe $y_t = f(x_t) + \eta_t$, $\quad$ Update μ_t, σ_t, $\quad n_e \leftarrow n_e + 1$,
    $\mathcal{D}_t \leftarrow \mathcal{D}_t \cup \{(x_t, y_t)\}$

  end

  $\mathcal{P}_t \leftarrow (\mathcal{P}_t \setminus \mathcal{Q}_1) \cup \mathcal{Q}_2$, $\quad t \leftarrow t + 1$

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 σ, an integer k and an $\alpha \in (0, 1]$ for the local polynomial estimator, the parameter L which is an upper bound on $\|f\|_{\mathcal{C}^{k,\alpha}}$ and a real-number $\rho_0 \geq (\gamma_n / \sqrt{LnD^{\alpha_1}})^{1/\alpha_1}$ for $\alpha_1 \coloneqq \max\{\alpha, \min\{1, k\}\}$.

Algorithm Outline. The `LP-GP-UCB` algorithm maintains a partition, $\mathcal{P}_t$, of the domain $\mathcal{X}$ at any time t, and to each cell E in the partition, it assigns a term $u^{(0)}_E$ 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 $\{\mathcal{X}\}$ and $u^{(0)}_{\mathcal{X}}$ is set to $+\infty$. As new cell E is added to $\mathcal{P}_t$, the value of $u^{(0)}_E$ 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 $u^{(0)}_E$ and two other terms: $u^{(1)}_{t,E}$ and $u^{(2)}_{t,E}$ defined as

$$u^{(1)}_{t,E} = \mu_t(x_{t,E}) + \beta_n \sigma_t(x_{t,E}) + L(\sqrt{D} r_E)^{\alpha_1}, \quad u^{(2)}_{t,E} = \hat{\mu}_t(E) + b_t(E) + L(\sqrt{D} r_E)^{\alpha_1},$$

where $x_{t,E}$ is a point drawn uniformly from E and μ_t and σ_t are the posterior mean and variance of the surrogate GP model, and β_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 $\tilde{f}_E$, the average f value in the cell E, and $b_t(E) = \sqrt{2\log(n/\delta)/n_E}$ is the length of the confidence interval of $\tilde{f}_E$

(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 `Recommend` described in Definition 1 below.

Definition 1 (`Recommend`). *Suppose the algorithm stops in round t_n and let $\mathrm{T} \subset \{1, 2, \ldots, t_n\}$ denote the set of times at which Algorithm 1 performed function evaluations (note that $|\mathrm{T}| = n$). Define $E_n = \arg\min_{E \in \mathcal{P}_{t_n}} r_E$, and $\xi := \min_{t \in \mathrm{T}} \beta_t \sigma_t(x_t)$. If $L(\sqrt{D} r_{E_n})^{\alpha_1} \leq \xi$, 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 $\tau := \arg\min_{t \in \mathrm{T}} \beta_t \sigma_t(x_t)$.*

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

Algorithm 2: `ExpandAndBound`

Input: E, `flag`, `val`, $\mathcal{D}$, k, α, δ

$n_E \leftarrow |\mathcal{D}^{(E)}|$

if *flag==1* **then**

  $\mathcal{Q}_2 =$ `Partition`$(E, r_E/2)$

  for $F \in \mathcal{Q}_2$ **do**

    $u_F^{(0)} =$ `val`

  end

else

  `err` = `MaxErr`$(E, \mathcal{D}, K, B, k, \delta, \sigma)$

  $\tilde{r} = \min\left\{\frac{r_E}{2}, \frac{1}{\sqrt{D}}\left(\frac{\texttt{err}}{L}\right)^{1/\alpha_1}\right\}$

  $\mathcal{Q}_2 =$ `Partition`$(E, \tilde{r})$

  for $F \in \mathcal{Q}_2$ **do**

    $\hat{f}_F =$ `LocalPoly`$(F, \mathcal{D}, x_F)$

    $u_F^{(0)} = \hat{f}_F + 2$`err`

  end

end

Output: $\mathcal{Q}_2$

Definition 2 (`Partition`). *Given a cell $E = \bigtimes_{i=1}^D [a_i, b_i] \subset \mathcal{X}$, the function call `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 $F = \bigtimes_{i=1}^D [\tilde{a}_i, \min\{\tilde{a}_i + r, b_i\}]$, where $\tilde{a}_i = a_i + lr$ for $l \in \{0, 1, \ldots, \lfloor r_E/r \rfloor\}$.*

To complete the description of the `ExpandAndBound` algorithm, we next describe the construction of local polynomial estimators, and introduce the functions `LocalPoly` (in Def. 3) and `MaxErr` (in Def. 4) which are called by `ExpandAndBound`. Given a cell $E \subset \mathcal{X}$ and a point $z \in E$, we define the LP estimator at z as $\hat{f}_E(z, \vec{w}) = \sum_{x \in \mathcal{D}^{(E)}} w_x y_x$, where $\vec{w} = \{w_x : x \in \mathcal{D}_{\mathcal{X}}^{(E)}\}$ are defined as the solution of the following problem (Nemirovski, 2000, Eq. (1.36)):

$$\min_{\vec{v} = \{v_x : x \in \mathcal{D}_{\mathcal{X}}^{(E)}\}} \sum_{x \in \mathcal{D}_{\mathcal{X}}^{(E)}} |v_x|^2 \text{ s.t. } p(z) = \sum_{x \in \mathcal{D}_{\mathcal{X}}^{(E)}} v_x p(x) \quad \forall p \in \mathcal{P}_D^k. \tag{\textbf{LP}}$$

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

Definition 3 (`LocalPoly`). *Given a cell E, a point $x \in E$, the function `LocalPoly` returns the estimated function value $\hat{f}_E(x, \vec{w})$ at x, calculated according to the formula stated above.*

If $n_E \coloneqq |\mathcal{D}^{(E)}| > (k+2)^D$, the weights $\vec{w}$ are the solution to (**LP**)*, while if $n_E < (k+2)^D$, the weights are set as $w_x = 1/n_E$ for all $x \in \mathcal{D}_{\mathcal{X}}^{(E)}$.*

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

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

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

Remark 1. *Recall that term $\Phi_k(f, E) \coloneqq \inf_{p \in \mathcal{P}_D^k} \sup_{x \in E} |f(x) - p(x)|$ depends on how well elements of $\mathcal{H}_K$ can be approximated by polynomials in $\mathcal{P}_D^k$. For functions f with $\|f\|_{\mathcal{C}^{k,\alpha}} \leq L$, it is known that we have $\Phi_k(f, E) \leq L(\sqrt{D} r_E)^{k+\alpha}$. Using this, we can construct an upper bound on the first term on the RHS of* (1)*, denoted by $e_D(E, \vec{w}_E, K, B) \coloneqq (1 + \|\vec{w}_E\|_1) L(\sqrt{D} r_E)^{k+\alpha}$, from the information available to the algorithm. The second term in the RHS of* (1)*, which we shall denote by $e_S(\vec{w}_E, \sigma, \delta)$, depends only on the data and known terms (σ and δ) 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, $\mathcal{D}$, K, B, k, δ and σ as inputs, and returns `err` defined as $\texttt{err} \coloneqq \max_{x \in E}\; e_D(E, \vec{w}_{E,x}, K, B) + e_S(\vec{w}_{E,x}, \sigma, \delta)$, where e_S and e_D were introduced in Remark 1, and $\vec{w}_{E,x}$ 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:*

$$\mathcal{S}_n = \tilde{\mathcal{O}}\left(\min\left(n^{\frac{1}{2} - \frac{D(D+1)}{2\nu + D(D+1)}},\; n^{-a_\nu}\right)\right), \quad \textit{and} \quad \mathcal{R}_n = \tilde{\mathcal{O}}\left(\min\left(n^{\frac{1}{2} + \frac{D(D+1)}{2\nu + D(D+1)}},\; n^{b_\nu}\right)\right). \tag{2}$$

The exponent a_ν is $\alpha_1/(2\alpha_1 + D)$ for $\nu > D(D+1)/2$ and $\nu/(2\nu + D)$ otherwise, while the exponent b_ν 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 γ_n is $\Omega(\sqrt{n})$. In this parameter range, the `LP-GP-UCB` algorithm achieves near-optimal rates for $\mathcal{S}_n$, thus closing the large gap between the upper and lower bounds in the literature. However, the improvement

achieved for $\mathcal{R}_n$ is not as significant as that of $\mathcal{S}_n$. 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 ρ_0. While this ensures that the algorithm can find at least one *good* point resulting in small $\mathcal{S}_n$, due to insufficient *exploitation*, the bound on $\mathcal{R}_n$ suffers. A similar trade-off between obtaining tight bounds for both $\mathcal{S}_n$ and $\mathcal{R}_n$ occurs in some other bandit problems as well; see for example (Bubeck et al., 2011, § 3). In Proposition 2 next, we show that the by using a zeroth degree LP estimators, we can obtain tighter control over $\mathcal{R}_n$ 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 $c_\nu = {}^{D(D+3)}/_{4\nu+D(D+5)}$) :*

$$\mathcal{S}_n = \tilde{\mathcal{O}}\left(\min\left(n^{-1/2+c_\nu}, n^{-1/(D+2)}\right)\right), \qquad \text{and} \quad \mathcal{R}_n = \tilde{\mathcal{O}}\left(\min\left(n^{1/2+c_\nu}, n^{(D+1)/(D+2)}\right)\right). \quad (3)$$

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 ν and D.

4. Summary of Improvements

To discuss the improvements in regret bounds, we introduce the notations $\mathcal{I}_0 = (0, 1]$, $\mathcal{I}_1 = (1, {}^{D(D+1)}/_2]$, $\mathcal{I}_2 = ({}^{D(D+1)}/_2, {}^{(D^2+5D+12)}/_4]$ and $\mathcal{I}_3 = (e, \infty)$ where $e = \max\{{}^{D(D+1)}/_2, {}^{(D^2+5D+12)}/_4\}$. Note that $\mathcal{I}_2$ 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 $\mathcal{S}_n$ is $\tilde{\mathcal{O}}\left(n^{-\nu/(2\nu+D)}\right)$ for $\nu \in \mathcal{I}_0 \cup \mathcal{I}_1$ and $\tilde{\mathcal{O}}\left(n^{-1/(D+2)}\right)$ for $\nu \in \mathcal{I}_2$, both of which are achieved by the LP-GP-UCB algorithm with $k = \lceil \nu \rceil - 1$, i.e., (2). For $\nu \in \mathcal{I}_3$, we have $\mathcal{S}_n = \tilde{\mathcal{O}}\left(n^{-\frac{1}{2}+c_\nu}\right)$ for $c_\nu = \frac{D(D+3)}{4\nu+D(D+5)}$ which is achieved by LP-GP-UCB with $k = 0$. Note that for $\nu \in \mathcal{I}_0 \cup \mathcal{I}_1$, 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 $\mathcal{S}_n$ are $\mathcal{O}(1)$ due to γ_n being $\Omega(\sqrt{n})$.

- *Cumulative Regret.* The LP-GP-UCB algorithm with $k = 0$ achieves tighter control over $\mathcal{R}_n$ for all values of ν. In particular for $\nu \in \mathcal{I}_0$ we have $\mathcal{R}_n = \tilde{\mathcal{O}}\left(n^{(\nu+D)/(2\nu+D)}\right)$, which matches the lower bound of (Scarlett et al., 2017) up to poly-log factors. For $\nu \in \mathcal{I}_1 \cup \mathcal{I}_2$, $\mathcal{R}_n$ is $\tilde{\mathcal{O}}\left(n^{(D+1)/(D+2)}\right)$ while for $\nu \in \mathcal{I}_3$ it is $\tilde{\mathcal{O}}\left(n^{1/2+c_\nu}\right)$ with c_ν defined in Prop. 2.

  The cumulative regret bounds discussed above improve upon the state-of-the-art for all values of ν, D. More specifically, **(i)** since γ_n based bounds are known only for $\nu > 1$, our results provide the first explicit (and near-optimal) regret bounds of Matérn kernels with $\nu \in \mathcal{I}_0 = (0, 1]$, and **(ii)** for $\nu > 1$ our results improve upon the best known bounds of $\tilde{\mathcal{O}}\left(n^{e_\nu}\right)$ with $e_\nu = \frac{D(2D+3)+2\nu}{D(2D+4)+4\nu}$, 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

Sébastien Bubeck, Rémi Munos, and Gilles Stoltz. Pure exploration in finitely-armed and continuous-armed bandits. *Theoretical Computer Science*, 412(19):1832–1852, 2011.

David Janz, David R Burt, and Javier González. Bandit optimisation of functions in the Matérn kernel RKHS. *arXiv preprint arXiv:2001.10396*, 2020.

Arkadi Nemirovski. Topics in Non-parametric Statistics. *Ecole d'Eté de Probabilités de Saint-Flour*, 28:85, 2000.

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

Shubhanshu Shekhar and Tara Javidi. Multi-Scale Zero-Order Optimization of Smooth Functions in an RKHS. *arXiv preprint arXiv:2005.04832*, 2020.

Niranjan Srinivas, Andreas Krause, Sham M Kakade, and Matthias W Seeger. Information-theoretic Regret Bounds for Gaussian Process Optimization in the Bandit Setting. *IEEE Transactions on Information Theory*, 58(5):3250–3265, 2012.

Giorgio Talenti. Best constant in Sobolev inequality. *Annali di Matematica pura ed Applicata*, 110(1):353–372, 1976.