

# Lower Bounds on Regret for Noisy Gaussian Process Bandit Optimization

Jonathan Scarlett, Ilija Bogunovic, and Volkan Cevher

Laboratory for Information and Inference Systems (LIONS), EPFL  
 {jonathan.scarlett, ilija.bogunovic, volkan.cevher}@epfl.ch

## Abstract

In this paper, we consider the problem of sequentially optimizing a black-box function  $f$  based on noisy samples and bandit feedback. We assume that  $f$  is smooth in the sense of having a bounded norm in some reproducing kernel Hilbert space (RKHS), yielding a commonly-considered non-Bayesian form of Gaussian process bandit optimization. We provide algorithm-independent lower bounds on the simple regret, measuring the suboptimality of a single point reported after  $T$  rounds, and on the cumulative regret, measuring the sum of regrets over the  $T$  chosen points.

For the isotropic squared-exponential kernel in  $d$  dimensions, we find that an average simple regret of  $\epsilon$  requires  $T = \Omega(\frac{1}{\epsilon^2}(\log \frac{1}{\epsilon})^{d/2})$ , and the average cumulative regret is at least  $\Omega(\sqrt{T(\log T)^{d/2}})$ , thus matching existing upper bounds up to the replacement of  $d/2$  by  $2d + O(1)$  in both cases. For the Matérn- $\nu$  kernel, we give analogous bounds of the form  $\Omega((\frac{1}{\epsilon})^{2+d/\nu})$  and  $\Omega(T^{\frac{\nu+d}{2\nu+d}})$ , and discuss the resulting gaps to the existing upper bounds.

## I. INTRODUCTION

The problem of sequentially optimizing a black-box function based on noisy bandit feedback has recently attracted a great deal of attention, and finds applications in robotics, environmental monitoring, and hyperparameter optimization in machine learning, just to name a few. In order to make this problem tractable, one must place smoothness assumptions on the function. Gaussian processes (GPs) provide a versatile means for doing this, capturing the smoothness properties through a suitably-chosen kernel.

Within the GP framework, there are two distinct viewpoints: Bayesian and non-Bayesian. In the Bayesian viewpoint, one assumes that the underlying function is random according to a GP distribution with the specified kernel. On the other hand, in the non-Bayesian viewpoint, one treats the function as fixed and unknown, and only assumes that it has a low norm in the reproducing kernel Hilbert space (RKHS) corresponding to the kernel.

Upper bounds on the *regret*, a widely-adopted performance measure (see Section I-A), have been given for a variety of practical GP bandit optimization<sup>1</sup> algorithms, including upper confidence bound (GP-UCB) [21] and various alternatives [4; 19; 22]. However, to our knowledge, there are no existing

<sup>1</sup>We use the terminology *Gaussian process bandit optimization* to refer collectively to the Bayesian and non-Bayesian settings, whereas the former is typically referred to as *Bayesian optimization*.

*algorithm-independent lower bounds* on the regret in the noisy setting, thus making it unclear to what extent the upper bounds can be improved.

In this paper, we provide such lower bounds in the *non-Bayesian* setting, focusing on two specific widely-used kernels: squared exponential (SE) and Matérn. For the SE kernel, our lower bound nearly matches existing upper bounds, whereas the gap is seen to be more significant in the case of the Matérn kernel. As a result, the latter deserves further study both in terms of the lower bounds and the existing algorithms; see Section II for further discussion.

### A. Setup

We consider the optimization of a function  $f$  defined on a compact domain  $D = [0, 1]^d$ . The smoothness is dictated by a *kernel function*  $k(x, x')$  [18]; specifically, we assume that the corresponding RKHS norm is bounded as  $\|f\|_k \leq B$ . The set of all such functions is denoted by  $\mathcal{F}_k(B)$ , and for a given  $f \in \mathcal{F}_k(B)$ , we let  $x^*$  denote an arbitrary maximizer of  $f$ . In the  $t$ -th round, a point  $x_t$  is selected, and a noisy sample  $y_t = f(x_t) + z_t$  is observed, with  $z_t \sim N(0, \sigma^2)$  for some  $\sigma^2 > 0$ . We assume independence among the noise terms at differing time instants.

We consider two widely considered performance metrics:

- **Simple regret:** At the end of  $T$  rounds, an additional point  $x^{(T)}$  is reported, and the simple regret incurred is  $r^{(T)} = f(x^*) - f(x^{(T)})$ . We seek to characterize how large  $T$  must be to permit  $\mathbb{E}[r^{(T)}] \leq \epsilon$ .
- **Cumulative regret:** At the end of  $T$  rounds, the cumulative regret incurred is  $R_T = \sum_{t=1}^T r_t$ , where  $r_t = f(x^*) - f(x_t)$ . We seek to provide lower bounds on  $\mathbb{E}[R_T]$ .

We focus on two commonly-considered kernels, namely, squared exponential (SE) and Matérn [18]:

$$k_{\text{SE}}(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2l^2}\right) \quad (1)$$

$$k_{\text{Matérn}}(x, x') = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}\|x - x'\|}{l}\right)^\nu J_\nu\left(\frac{\sqrt{2\nu}\|x - x'\|}{l}\right), \quad (2)$$

where  $l > 0$  denotes the length-scale,  $\nu > 0$  is an additional parameter that dictates the smoothness, and  $J_\nu$  denotes the modified Bessel function. Although we only consider these specific kernels, the approach that we take can also potentially be applied to other stationary kernels.

### B. Related Work

A recent review of Gaussian process bandit optimization is given in [20]. Upper bounds on the regret for specific algorithms are given for the Bayesian and non-Bayesian settings with cumulative regret in [19; 21; 23], and with simple regret in [4; 8]. Of these, [23] is the only one that did not assume perfect knowledge of the kernel, but this distinction will not be important for our purposes, since our lower bounds hold in either case. Bounded RKHS norm assumptions have also been used in other problems, such as uniform function approximation [16; 17].

We are not aware of any existing lower bounds for the noisy setting. However, [7] is related to our work, providing tight lower bounds on the simple regret for the Matérn kernel in the *noiseless*

setting. While our high-level approach is similar to [7] in the sense of considering a class of *needle-in-haystack* type functions that are difficult to distinguish, there are a number of additional challenges that we need to address:

- 1) Incorporating noise is non-trivial, and it is unclear *a priori* the extent to which it affects the regret bounds. For instance, in the Bayesian setting,  $O(1)$  cumulative regret is possible for noiseless observations [10; 12], whereas the best known bounds for the noisy case are between  $O(\sqrt{T})$  and  $o(T)$  [21].
- 2) We consider not only the simple regret, but also the cumulative regret, which is not captured by the framework of [7].
- 3) While [7, Thm. 1] gives a tight lower bound for the Matérn kernel, no analogous result is given for the widely-adopted SE kernel. Moreover, an inspection of the analysis in [7] reveals that the *compactly supported* functions used therein have an infinite RKHS norm for the SE kernel, and thus cannot be used. To circumvent this issue, we consider a different class of functions that are compactly supported *in frequency domain*, and hence have unbounded support in the spatial domain  $\mathbb{R}^d$ . This introduces additional challenges into the analysis, since we can no longer directly state that most samples are uninformative as a result of the sampled function value being exactly zero.

Another related work on the noiseless setting is [11], which provided upper and lower bounds for the Bayesian case under a Hölder-continuity assumption on the kernel.

While multi-armed bandit problems [2; 3; 5] typically consist of *finite* action spaces, some continuous variants have been proposed [13; 14]. However, these make significantly different smoothness assumptions of the *Lipschitz* variety, and as a result, the results and analysis techniques do not apply to our setting of bounded RKHS norm. In particular, Lipschitz properties do not imply bounds on the RKHS norm, nor vice versa.

## II. MAIN RESULTS

In all of our results, we assume that the dimension  $d$  in the definition of  $D = [0, 1]^d$ , as well as the parameters  $l$  and  $\nu$  in the kernels in (1)–(2), behave as  $\Theta(1)$ . In contrast, we allow the RKHS norm upper bound  $B$  and the noise level  $\sigma$  to scale generally. We let  $O^*(\cdot)$  denote asymptotic expressions up to dimension-independent logarithmic factors, e.g.,  $4\sqrt{T}(\log T)^{d+3} = O^*(\sqrt{T}(\log T)^d)$ .

Our main results are expressed in terms of the *average* regret. However, it is straightforward to modify the proofs in order to show that the scaling laws are the same when considering regret bounds that hold with constant (but arbitrarily close to one) probability. See Section V-D for details.

We first provide our main result for the simple regret.

**Theorem 1.** (Simple Regret) *Fix  $\epsilon \in (0, \frac{1}{2})$ ,  $B > 0$ , and  $T \in \mathbb{Z}$ . Suppose there exists an algorithm that, for any  $f \in \mathcal{F}_k(B)$ , achieves an average simple regret  $\mathbb{E}[r^{(T)}] \leq \epsilon$  after  $T$  rounds. Then, provided that  $\frac{\epsilon}{B}$  is sufficiently small, we have the following:*

1) For  $k = k_{\text{SE}}$ , it is necessary that

$$T = \Omega\left(\frac{\sigma^2}{\epsilon^2} \left(\log \frac{B}{\epsilon}\right)^{d/2}\right). \quad (3)$$

2) For  $k = k_{\text{Matérn}}$ , it is necessary that

$$T = \Omega\left(\frac{\sigma^2}{\epsilon^2} \left(\frac{B}{\epsilon}\right)^{d/\nu}\right). \quad (4)$$

Our main result for the cumulative regret is as follows.

**Theorem 2.** (Cumulative Regret) *Fix  $B > 0$  and  $T \in \mathbb{Z}$ . Suppose there exists an algorithm that, for any  $f \in \mathcal{F}_k(B)$ , achieves an average cumulative regret of  $\mathbb{E}[R_T]$  after  $T$  rounds. Then, we have the following:*

1) For  $k = k_{\text{SE}}$ , it is necessary that

$$\mathbb{E}[R_T] = \Omega\left(\sqrt{T\sigma^2 \left(\log \frac{B^2 T}{\sigma^2}\right)^{d/2}}\right) \quad (5)$$

*provided that  $\frac{\sigma}{B} = O(\sqrt{T})$  with a sufficiently small implied constant.*

2) For  $k = k_{\text{Matérn}}$ , it is necessary that

$$\mathbb{E}[R_T] = \Omega\left(B^{\frac{d}{2\nu+d}} \sigma^{\frac{2\nu}{2\nu+d}} T^{\frac{\nu+d}{2\nu+d}}\right) \quad (6)$$

*provided that  $\frac{\sigma}{B} = O(\sqrt{T})$  with a sufficiently small implied constant.*

We note that the condition  $\frac{\sigma}{B} = O(\sqrt{T})$  is assumed for technical reasons, and is quite mild. In particular, below we will focus primarily on the case that  $\sigma$  and  $B$  are constants that do not scale with  $T$ , in which case this condition is trivially satisfied.

**Comparisons to Upper Bounds:** The best known upper bounds on the cumulative regret were given for the upper confidence bound (GP-UCB) algorithm [21], where bounded noise was assumed. While Gaussian noise has unbounded support, it is bounded by  $O(\log T)$  for all  $t = 1, \dots, T$  with high probability, and hence the bounded noise results transfer to the Gaussian setting at the expense of additional dimension-independent logarithmic factors.

A summary of the comparisons is shown in Table I, and we proceed by discussing the entries in more detail.<sup>2</sup> We first compare our Theorem 2 (cumulative regret) to [21] in the case that  $\sigma^2$  and  $B$  behave as  $\Theta(1)$ ; we will later discuss the dependence on these parameters. For fixed  $\sigma^2$ , the upper bound in [21] is of the form

$$R_T = O^*\left(\sqrt{TB\gamma_T + T\gamma_T^2}\right), \quad (7)$$

where, letting  $I(X; Y)$  denote the mutual information [9], we define

$$\gamma_T = \max_{x_1, \dots, x_T} \max_{S: |S|=T} I(f; \mathbf{y}_S), \quad f \sim \mathcal{GP}(0, k). \quad (8)$$

<sup>2</sup>This version of Table I corrects some minor mistakes in the previous version of this manuscript, where a few of the exponents  $(\log(\cdot))^{d/2}$  vs.  $(\log(\cdot))^d$  vs.  $(\log(\cdot))^{2d}$  were off by a factor of two for the SE kernel.

	Upper bound ([21])	Conjectured upper bound (see Section II)	Lower bound (Theorems 1–2)
SE kernel Cumulative regret	$O^*\left(\sqrt{T(\log T)^{2d}}\right)$	$O^*\left(\sqrt{T(\log T)^d}\right)$	$\Omega\left(\sqrt{T(\log T)^{d/2}}\right)$
Matérn kernel Cumulative regret	$O^*\left(T^{\frac{1}{2} \cdot \frac{2\nu+3d(d+1)}{2\nu+d(d+1)}}\right)$	$O^*\left(T^{\frac{\nu+d(d+1)}{2\nu+d(d+1)}}\right)$	$\Omega\left(T^{\frac{\nu+d}{2\nu+d}}\right)$
SE kernel Time to simple regret $\epsilon$	$O^*\left(\frac{1}{\epsilon^2} \left(\log \frac{1}{\epsilon}\right)^{2d}\right)$	$O^*\left(\frac{1}{\epsilon^2} \left(\log \frac{1}{\epsilon}\right)^d\right)$	$\Omega\left(\frac{1}{\epsilon^2} \left(\log \frac{1}{\epsilon}\right)^{d/2}\right)$
Matérn kernel Time to simple regret $\epsilon$	$O^*\left(\left(\frac{1}{\epsilon}\right)^{\frac{2(2\nu+d(d+1))}{2\nu-d(d+1)}}\right)$ (if $2\nu - d(d+1) > 0$ )	$O^*\left(\left(\frac{1}{\epsilon}\right)^{2+d(d+1)/\nu}\right)$	$\Omega\left(\left(\frac{1}{\epsilon}\right)^{2+d/\nu}\right)$

Table I: Summary of regret bounds for a fixed RKHS norm bound  $B > 0$  and noise level  $\sigma^2 > 0$ .

This represents the maximum amount of information that a set of  $T$  noisy observations  $\mathbf{y}_S = (y_1, \dots, y_T)$  (corresponding to  $\mathbf{x}_S = (x_1, \dots, x_T)$ ) can reveal about a zero-mean Gaussian process  $f$  with kernel  $k$ . We have  $\gamma_T = O((\log T)^{d+1})$  for the squared-exponential kernel, and  $\gamma_T = O^*\left(T^{\frac{d(d+1)}{2\nu+d(d+1)}}\right)$  for the Matérn kernel [21].

Hence, for  $k_{\text{SE}}$ , the upper bound in (7) is  $O^*\left(\sqrt{T(\log T)^{2d}}\right)$  and the lower bound in (3) is  $\Omega\left(\sqrt{T(\log T)^{d/2}}\right)$ , so the two coincide up to the factor of 2 vs.  $\frac{1}{2}$  in the exponent, as well as the (few) extra  $\log T$  factors hidden in the  $O^*(\cdot)$  notation.

While [21] focused on the cumulative regret, one can obtain simple regret bounds by noting that it is always possible to have a simple regret no higher than the normalized cumulative regret [6]; this yields the condition  $\frac{T}{B\gamma_T + \gamma_T^2} \geq \frac{C}{\epsilon^2}$  for  $\epsilon$ -optimality, where  $C = O^*(1)$  (see also [4; 8]). By substituting  $\gamma_T = O((\log T)^{d+1})$  and rearranging, we find that for the SE kernel this condition is of the form  $T \geq \frac{C'}{\epsilon^2} \left(\log \frac{1}{\epsilon}\right)^{2d}$  for some  $C' = O^*(1)$ , again matching (3) up to the factor of 2 vs.  $\frac{1}{2}$  in the exponent.

For the Matérn kernel, the gaps are more significant. In particular, substituting  $\gamma_T = O^*\left(T^{\frac{d(d+1)}{2\nu+d(d+1)}}\right)$ , we find that the cumulative regret upper bound in (7) is only sublinear when  $\frac{d(d+1)}{2\nu+d(d+1)} < \frac{1}{2}$ , or equivalently  $2\nu - d(d+1) > 0$ . A similar statement holds for having a non-void condition on the simple regret. This suggests that these existing upper bounds may be loose, since, at least in the authors' judgment, sublinear regret should be possible for all  $\nu$  and  $d$ . This poses an interesting open problem for future work.

In contrast, if the right-hand side of (7) could be improved to  $O^*\left(\sqrt{TB\gamma_T}\right)$ , then the upper bound would be sublinear for all  $d$  and  $\nu$ . We conjecture that such a bound holds (as is the case in the Bayesian setting [21]), but we do not currently have a proof. We proceed by discussing the bounds that would arise if this conjecture were true, as summarized in the middle column of Table I.

For the squared exponential kernel, the conjecture would partially close the above-mentioned gap on the factor of 2 in the exponent to  $\log T$ . For the Matérn kernel, we would obtain a cumulative regret of  $O^*\left(T^{\frac{\nu+d(d+1)}{2\nu+d(d+1)}}\right)$ , hence matching our lower bound up to the replacement of  $d$  by  $d(d+1)$ . Thus, both bounds would be of the same form, and behave as  $T^c$  for some  $c$  that tends to one as  $d$

grows large. Analogous observations hold for the simple regret, as summarized in Table I.

Finally, we discuss the dependence of the bounds on  $B$  and  $\sigma$ , focusing on the squared exponential kernel with cumulative regret for brevity. Up to the logarithmic term, our bound (5) depends linearly on  $\sigma$ . In contrast, the upper bound in [21] contains a multiplicative term  $O\left(\sqrt{\frac{1}{\log(1+\sigma^{-2})}}\right)$ , which reduces to  $O(\sigma)$  when  $\sigma$  is large, but tends to zero much more slowly when  $\sigma$  is small. However, it is shown in [4] that a linear dependence is attainable for the simple regret, and one should expect that such a dependence is similarly attainable for the cumulative regret. As for the dependence on  $B$ , the optimal scaling as  $B \rightarrow \infty$  is unclear. For fixed  $T$ ,  $d$ , and  $\sigma^2$ , our lower bound has dependence  $(\log B)^d$ , whereas the upper bound of in (7) has dependence  $\sqrt{B}$ , which is stronger since  $d = O(1)$ .

The remainder of the paper is devoted to the proofs of Theorems 1 and 2.

### III. CONSTRUCTION OF A FINITE ENSEMBLE OF FUNCTIONS

We obtain lower bounds on the regret by considering the case that  $f$  is uniform on a finite set  $\mathcal{F}' = \{f_1, \dots, f_M\}$ . If we can lower bound the regret (of an arbitrary algorithm) averaged over  $m \in \{1, \dots, M\}$ , then it follows that there must exist a particular value of  $m$  such that the same lower bound applies to  $f_m$ .

More specifically, we let  $\mathcal{F}'$  have the property that any single point  $x$  can only be  $\epsilon$ -optimal for at most one function in this set. By this property, the optimization problem with simple regret not exceeding  $\epsilon$  is equivalent to the correct identification of the index  $m \in \{1, \dots, M\}$ .

Note that for the simple regret, the parameter  $\epsilon$  is part of the problem statement, whereas for cumulative regret it is a parameter that we can choose to our liking. The starting point of the analysis is the same for both regret notions, and throughout, we assume that  $\frac{\epsilon}{B}$  is sufficiently small. This is true by assumption in Theorem 1, whereas when we prove Theorem 2, we will need to check that our choice of  $\epsilon$  is consistent with this assumption.

#### A. Function Class

The class that we consider corresponds to a ‘‘needle in haystack’’ problem, and is described as follows: Let  $g(x)$  be a function on  $\mathbb{R}^d$  with values in  $[-2\epsilon, 2\epsilon]$ , an RKHS norm not exceeding  $B$ , a peak value of  $2\epsilon$  at  $x = 0$ , and a value strictly less than  $\epsilon$  when  $\|x\|_\infty \geq w$ , for some *width*  $w > 0$  to be chosen later. We let each  $f_m(x)$  be given by  $g(x)$  shifted so that its peak is at a given point, and then cropped to the domain  $[0, 1]^d$ . By forming a  $d$ -dimensional grid of step size  $w$  in each dimension, we can form

$$M = \left\lfloor \left(\frac{1}{w}\right)^d \right\rfloor \quad (9)$$

of these while ensuring that any  $\epsilon$ -optimal point for  $f_m$  fails to be  $\epsilon$ -optimal for any of the other  $f_{m'}$ . See Figure 1 for an illustration.

It remains to choose  $g$ , and to determine small we can allow  $w$  to be without violating the RKHS norm constraint. Intuitively, a smaller  $w$  gives a ‘‘less smooth’’ function and hence a higher RKHS

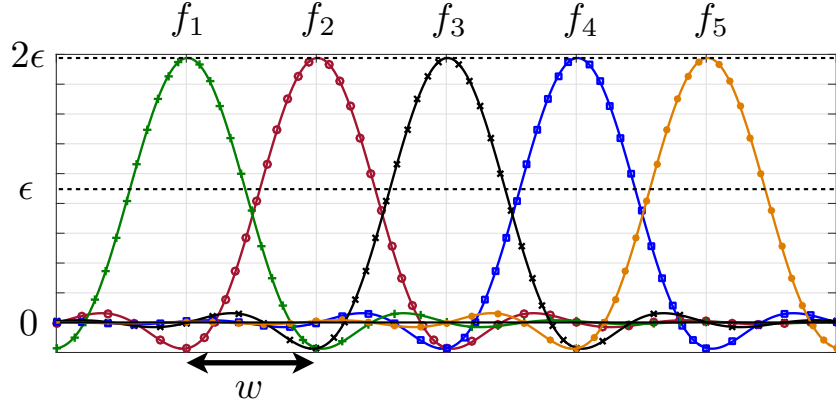


Figure 1: Illustration of functions  $f_1, \dots, f_5$  such that any given point  $x \in [0, 1]$  can only be  $\epsilon$ -optimal for one function. See (10)–(11) for the specific function being plotted.

norm, but a smaller amplitude  $2\epsilon$  also decreases the RKHS norm in a linear fashion. Hence, as  $\epsilon \rightarrow 0$ , we can afford to take  $w \rightarrow 0$ .

Instead of specifying  $g$  directly, we specify the Fourier transform  $H$  of another function  $h$ , and then let  $g$  be a suitably scaled version of  $h$ . Specifically,  $H$  is defined to be the multi-dimensional *bump function*:

$$H(\xi) = \begin{cases} \exp\left(-\frac{1}{1-\|\xi\|_2^2}\right) & \|\xi\|_2^2 \leq 1 \\ 0 & \|\xi\|_2^2 > 1, \end{cases} \quad (10)$$

which is compactly supported and infinitely differentiable. Let  $h(x)$  be the inverse Fourier transform of  $H$ . Since  $H$  is a fixed function with finite energy, the amplitude of  $h(x)$  must vanish as  $\|x\|$  grows large; hence, there exists a constant  $\zeta$  such that  $h(x) < \frac{1}{2}h(0)$  for  $\|x\|_\infty > \zeta$ .

We choose

$$g(x) = \frac{2\epsilon}{h(0)} h\left(\frac{x\zeta}{w}\right), \quad (11)$$

where the parameter  $w$  is the same as that in (9). Note that since  $H$  is real and symmetric, the maximum of  $h$  occurs at zero, and thus the maximum of  $g$  is  $g(0) = 2\epsilon$ , as desired. We consider  $w$  as arbitrary for now, but this will be chosen to ensure that the RKHS norm of  $g$  is upper bounded by  $B$ ; see Section IV. To get an idea of how  $g$  behaves, we note that it is precisely the function that was used in producing Figure 1 (with  $d = 1$ ).

#### IV. BOUNDING THE RKHS NORM

Intuitively, in our function class in Section III-A, we would like  $M$  to be as large as possible, since this means that there are more functions that need to be distinguished, and a higher lower bound can be obtained. In this section, we determine how large  $M$  can be in our construction while still ensuring that our assumptions are valid, namely, the RKHS norm of each function is upper bounded by  $B$ , and any given point can only be  $\epsilon$ -optimal for at most one function.

### A. Properties of RKHS Norms

The following properties of RKHS norms are well-known; see for example [1, Sec. 1.5] and [7, Sec. 3]. The first property explicitly expresses the RKHS norm on  $\mathbb{R}^d$  in terms of the Fourier transform.

**Lemma 1.** [1, Sec. 1.5] *Consider an RKHS  $\mathcal{H}$  for functions on  $\mathbb{R}^d$ , corresponding to a kernel of the form  $k(x, x') = k(\tau)$  with  $\tau = x - x'$ , and let  $K(\xi)$  be the  $d$ -dimensional Fourier transform of  $k(\tau)$ . Then for any  $\tilde{f} \in \mathcal{H}$  with Fourier transform  $\tilde{F}(\xi)$ , we have*

$$\|\tilde{f}\|_{\mathcal{H}}^2 = \int \frac{|\tilde{F}(\xi)|^2}{K(\xi)} d\xi. \quad (12)$$

The next result relates the RKHS norm on  $\mathbb{R}^d$  to that on a compact subset  $D \subseteq \mathbb{R}^d$ , and reveals that the cropping operation described at the start of Section III-A cannot increase the RKHS norm.

**Lemma 2.** [1, Sec. 1.5] *Consider two RKHS  $\mathcal{H}(D)$  and  $\mathcal{H}(\mathbb{R}^d)$  for functions on a compact set  $D \subseteq \mathbb{R}^d$  and  $\mathbb{R}^d$  respectively, corresponding to a kernel of the form  $k(x, x') = k(\tau)$  with  $\tau = x - x'$ . Then for any  $\tilde{f} \in \mathcal{H}(D)$ , we have*

$$\|\tilde{f}\|_{\mathcal{H}(D)} = \inf_{\tilde{g}} \|\tilde{g}\|_{\mathcal{H}(\mathbb{R}^d)}, \quad (13)$$

where the infimum is over all functions  $\tilde{g} \in \mathcal{H}(\mathbb{R}^d)$  that agree with  $\tilde{f}$  when restricted to  $D$ .

### B. Squared Exponential Kernel

We first consider the isotropic squared exponential kernel with lengthscale  $l$ , defined in (1). Writing  $k(x, x') = k(\tau)$  with  $\tau = x - x'$ , the Fourier transform is given by [18, Sec. 4.2]

$$K(\xi) = (2\pi l^2)^{d/2} e^{-2\pi^2 l^2 \|\xi\|_2^2}. \quad (14)$$

Using Lemma 1 and (14), and denoting  $a_0 = \frac{2\epsilon}{h(0)}$  and  $w_0 = \frac{w}{\zeta}$  (cf., (11)) for brevity, the RKHS norm of  $g$  in  $\mathbb{R}^d$  is given by

$$\|g\|_k^2 = \frac{1}{(2\pi l^2)^{d/2}} \int |G(\xi)|^2 e^{2\pi^2 l^2 \|\xi\|_2^2} d\xi \quad (15)$$

$$= \frac{a_0^2 w_0^{2d}}{(2\pi l^2)^{d/2}} \int_{\|w_0 \xi\|_2 \leq 1} \exp\left(2\pi^2 l^2 \|\xi\|_2^2 - \frac{2}{1 - \|w_0 \xi\|_2^2}\right) d\xi \quad (16)$$

$$\leq \frac{a_0^2 w_0^{2d}}{(2\pi l^2)^{d/2}} \int_{\|\xi\|_2 \leq \frac{1}{w_0}} \exp(2\pi^2 l^2 \|\xi\|_2^2) d\xi \quad (17)$$

$$\leq \frac{a_0^2 w_0^{2d}}{(2\pi l^2)^{d/2}} V(w_0^{-1}) \exp\left(\frac{2\pi^2 l^2}{w_0^2}\right) \quad (18)$$

$$\leq \frac{a_0^2}{(2\pi l^2)^{d/2}} \exp\left(\frac{2\pi^2 l^2}{w_0^2}\right), \quad (19)$$

where:

- (16) follows from (10)–(11) and the fact that the  $d$ -dimensional Fourier transform of  $ah(x/b)$  is  $ab^d H(xb)$ ;



- (18) follows by upper bounding  $\|\xi\|_2^2 \leq \frac{1}{w_0^2}$  according to the constraint, and defining  $V(w_0^{-1})$  to be the volume of a ball of radius  $w_0^{-1}$  in  $d$  dimensions:

$$V(w_0^{-1}) = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2} - 1)} (w_0^{-1})^d, \quad (20)$$

where  $\Gamma(\cdot)$  denotes the gamma function;

- (19) holds when  $w_0$  is chosen such that  $w_0^{2d}V(w_0^{-1}) \leq 1$ , as will be done shortly.

Equating (19) with  $B^2$  (so that  $\|g\|_k \leq B$ ), and recalling the definition  $a_0 = \frac{2\epsilon}{h(0)}$ , we find that

$$B(2\pi l^2)^{d/4} \exp\left(-\frac{\pi^2 l^2}{w_0^2}\right) h(0) = 2\epsilon \quad (21)$$

$$\implies w_0 = \frac{\pi l}{\sqrt{\log \frac{B(2\pi l^2)^{d/4} h(0)}{2\epsilon}}}. \quad (22)$$

From this choice, we see that  $w_0 \rightarrow 0$  as  $\frac{\epsilon}{B} \rightarrow 0$ . Hence, and since  $w_0^{2d}V(w_0^{-1}) = O(w_0^d)$ , we find that the above assumption  $w_0^{2d}V(w_0^{-1}) \leq 1$  is indeed true as long as  $\frac{\epsilon}{B}$  is sufficiently small. The latter is assumed in Theorem 1, and will be ensured when we choose  $\epsilon$  to prove Theorem 2.

Recalling that  $h(x) < \frac{1}{2}h(0)$  for  $\|x\|_\infty > \zeta$ , we have from (11) that  $g(x) < \frac{1}{2}g(0) = \epsilon$  for  $\|x\|_\infty > \zeta w_0 = w$ , as was assumed in the construction of the functions in Section III-A. Hence, from (22), we can choose  $w = \zeta w_0 = \frac{\zeta \pi l}{\sqrt{\log \frac{B(2\pi l^2)^{d/4} h(0)}{2\epsilon}}}$  in (9), yielding

$$M = \left\lceil \left( \frac{\sqrt{\log \frac{B(2\pi l^2)^{d/4} h(0)}{2\epsilon}}}{\zeta \pi l} \right)^d \right\rceil. \quad (23)$$

In summary, we have shown that with this choice of  $M$ , we can construct  $M$  functions whose RKHS norm is upper bounded by  $B$ , whose peak values are  $2\epsilon$ , and such that any  $\epsilon$ -optimal point for one function cannot be  $\epsilon$ -optimal for any of the other  $M - 1$  functions.

### C. Matérn Kernel

We now consider the Matérn kernel with parameters  $l$  and  $\nu$ , defined in (2). The Fourier transform of  $k$  (treated as a function of  $\tau = x - x'$ ) is [18, Sec. 4.2]

$$K(\xi) = c_1 \left( \frac{2\nu}{l^2} + 4\pi^2 \|\xi\|_2^2 \right)^{-(\nu+d/2)}, \quad (24)$$

where  $c_1 = \frac{2^d \pi^{d/2} \Gamma(\nu+d/2) (2\nu)^\nu}{\Gamma(\nu) l^{2\nu}}$  is a constant (recall that  $d$ ,  $l$ , and  $\nu$  are assumed to be fixed). Hence, using Lemma 1 and again denoting  $a_0 = \frac{2\epsilon}{h(0)}$  and  $w_0 = \frac{w}{\zeta}$  (cf., (11)), we have the following, analogously to (15):

$$\|g\|_k^2 = c_1^{-1} \int |G(\xi)|^2 \left( \frac{2\nu}{l^2} + 4\pi^2 \|\xi\|_2^2 \right)^{(\nu+d/2)} d\xi \quad (25)$$

$$= c_1^{-1} a_0^2 w_0^{2d} \int_{\|w_0 \xi\|_2 \leq 1} \exp\left(-\frac{2}{1 - \|w_0 \xi\|_2^2}\right) \left( \frac{2\nu}{l^2} + 4\pi^2 \|\xi\|_2^2 \right)^{(\nu+d/2)} d\xi \quad (26)$$

$$\leq c_1^{-1} a_0^2 w_0^{2d} \int_{\|\xi\|_2 \leq \frac{1}{w_0}} \left( \frac{2\nu}{l^2} + 4\pi^2 \|\xi\|_2^2 \right)^{(\nu+d/2)} d\xi \quad (27)$$

$$\leq c_1^{-1} a_0^2 w_0^{2d} V(w_0^{-1}) \left( \frac{2\nu}{l^2} + \frac{4\pi^2}{w_0^2} \right)^{(\nu+d/2)} \quad (28)$$

$$= c_1^{-1} a_0^2 w_0^{d-2\nu} V(w_0^{-1}) \left( \frac{2\nu w_0^2}{l^2} + 4\pi^2 \right)^{(\nu+d/2)} \quad (29)$$

$$\leq c_2 a_0^2 w_0^{-2\nu} \left( \frac{2\nu w_0^2}{l^2} + 4\pi^2 \right)^{(\nu+d/2)} \quad (30)$$

$$\leq c_2 a_0^2 w_0^{-2\nu} (8\pi^2)^{(\nu+d/2)}, \quad (31)$$

where:

- (26) follows from (10)–(11) and the fact that the  $d$ -dimensional Fourier transform of  $ah(x/b)$  is  $ab^d H(xb)$ ;
- (27) follows by upper bounding the exponential term by one;
- (28) follows by upper bounding  $\|\xi\|_2^2 \leq \frac{1}{w_0^2}$  according to the constraint, and using the definition of the volume in (20);
- (30) holds for some  $c_2 > 0$  since  $V(w_0^{-1}) = O(w_0^{-d})$  by (20);
- (31) holds when  $w_0$  is chosen such that  $\frac{2\nu w_0^2}{l^2} \leq 4\pi^2$ , as will be done shortly.

Equating (31) with  $B^2$  (thus ensuring  $\|g\|_k \leq B$ ), and recalling that  $a_0 = \frac{2\epsilon}{h(0)}$ , we obtain

$$\frac{B c_2^{-1/2} h(0) w_0^\nu}{(8\pi^2)^{(\nu+d/2)/2}} = 2\epsilon. \quad (32)$$

We again observe that in the limit as  $\frac{\epsilon}{B} \rightarrow 0$  we have  $w_0 \rightarrow 0$ ; hence, to satisfy the above condition  $\frac{2\nu w_0^2}{l^2} \leq 4\pi^2$ , it suffices that  $\frac{\epsilon}{B}$  is sufficiently small.

Rearranging (32), we obtain

$$w_0 = \left( 2\epsilon \frac{(8\pi^2)^{(\nu+d/2)/2}}{B c_2^{-1/2} h(0)} \right)^{1/\nu}. \quad (33)$$

Recalling that  $w = \zeta w_0$  in (9) and (11), we find that (33) gives the following analog of (23):

$$M = \left\lfloor \left( \frac{B c_3}{\epsilon} \right)^{d/\nu} \right\rfloor, \quad (34)$$

where

$$c_3 := \left( \frac{1}{\zeta} \right)^\nu \cdot \left( \frac{c_2^{-1/2} h(0)}{2(8\pi^2)^{(\nu+d/2)/2}} \right). \quad (35)$$

## V. BOUNDING THE REGRET

In this section, we combine the tools from Sections III and IV to deduce the regret bounds given in Theorems 1 and 2. Throughout the section, we use the fact that for  $P_1$  and  $P_2$  representing the

density functions of Gaussian random variables with means  $(\mu_1, \mu_2)$  and variance  $\sigma^2$ , we have

$$D(P_1\|P_2) = \frac{(\mu_1 - \mu_2)^2}{2\sigma^2}. \quad (36)$$

Here,  $D(P_1\|P_2) = \int_{\mathbb{R}} P_1(z) \log \frac{P_1(z)}{P_2(z)} dz$  denotes the Kullback-Leibler divergence between two density functions [9].

#### A. Auxiliary Lemmas and Definitions

Fix an arbitrary (e.g. optimal) bandit optimization algorithm. Let  $\mathbf{y} = (y_1, \dots, y_T)$  be the observations up to time  $T$ , and for  $m = 1, \dots, M$ , let  $P_m(\mathbf{y})$  be the probability density function of  $\mathbf{y}$  upon running the algorithm on a given function  $f_m$  indexed by  $m$ . Moreover, let  $P_0(\mathbf{y})$  be the probability density of  $\mathbf{y}$  when the algorithm is run with  $f$  being zero everywhere. Let  $\mathbb{E}_m$  and  $\mathbb{P}_m$  denote expectations and probabilities when the underlying function is  $f = f_m$ , and similarly for  $\mathbb{E}_0$  and  $\mathbb{P}_0$ . Finally, let  $\mathbb{E}[\cdot] = \frac{1}{M} \sum_{m=1}^M \mathbb{E}_m[\cdot]$  be the expectation averaged over a uniformly random function index, and similarly for  $\mathbb{P}[\cdot]$ .

The following lemma from [3] relates two expectations  $\mathbb{E}_m$  and  $\mathbb{E}_0$  in terms of the corresponding divergence  $D(P_0\|P_m)$ .

**Lemma 3.** [3, p. 27] *For any function  $a(\mathbf{y})$  taking values in a bounded range  $[0, A]$ , we have*

$$\mathbb{E}_m[a(\mathbf{y})] \leq \mathbb{E}_0[a(\mathbf{y})] + A\sqrt{D(P_0\|P_m)}. \quad (37)$$

To make the paper self-contained, the proof is given in the appendix.

Before proceeding, we introduce some additional notation:

- Let  $\{\mathcal{R}_m\}_{m=1}^M$  be the partition of the domain into  $M$  regions according to (9), i.e., done according to a uniform grid with  $f_m$  taking the maximum in the centre of  $\mathcal{R}_m$ ;
- Let  $j_t$  be the index at time  $t$  such that  $x_t$  falls into  $\mathcal{R}_{j_t}$  – this can be thought of as a quantization of  $x_t$ ;
- Let  $N_j = \sum_{t=1}^T \mathbb{1}\{j_t = j\}$  denote the number of points from  $\mathcal{R}_j$  that are selected throughout the  $T$  rounds;
- Let  $P_m(y_t|\mathbf{y}_{t-1})$  denote the distribution of  $y_t$  given all the observations up to time  $t-1$ , denoted by  $\mathbf{y}_{t-1} = (y_1, \dots, y_{t-1})$  (and  $\mathbf{y}_0 = \emptyset$ ), in the case that  $f = f_m$ ;
- Define the maximum function value within a single region  $\mathcal{R}_j$  as

$$\bar{v}_m^j := \max_{x \in \mathcal{R}_j} f_m(x), \quad (38)$$

and the maximum divergence within the region as

$$\bar{D}_m^j := \max_{x \in \mathcal{R}_j} D(P_0(\cdot|x)\|P_m(\cdot|x)), \quad (39)$$

where  $P_m(y|x)$  is the distribution of an observation  $y$  for a given selected point  $x$  under the function  $f_m$ , and similarly for  $P_0(y|x)$ .

The following lemma provides a simple upper bound on the divergence in (37).

**Lemma 4.** *Under the preceding definitions, we have*

$$D(P_0 \| P_m) \leq \sum_{j=1}^M \mathbb{E}_0[N_j] \bar{D}_m^j \quad (40)$$

*Proof.* We use the chain rule for divergence [9, Sec. 2.5] to write

$$D(P_0 \| P_m) = \sum_{t=1}^T \mathbb{E}_0 \left[ D(P_0(\cdot | \mathbf{y}_{t-1}) \| P_m(\cdot | \mathbf{y}_{t-1})) \right] \quad (41)$$

$$\leq \sum_{t=1}^T \sum_{j=1}^M \mathbb{P}_0[j_t = j] \bar{D}_m^j \quad (42)$$

$$= \sum_{j=1}^M \left( \sum_{t=1}^T \mathbb{P}_0[j_t = j] \right) \bar{D}_m^j \quad (43)$$

$$= \sum_{j=1}^M \mathbb{E}_0[N_j] \bar{D}_m^j, \quad (44)$$

where (42) follows by noting that the argument to the expectation in (41) depends on  $\mathbf{y}_{t-1}$  only through the resulting selected point  $x_t$ , which is chosen based on  $\mathbf{y}_{t-1}$ . When  $x_t$  falls into  $\mathcal{R}_j$  (i.e.,  $j_t = j$ ), the divergence is upper bounded by  $\bar{D}_m^j$  due to (39).  $\square$

Finally, the following technical lemma will also be key to the analysis. It roughly states that if we sum the function values  $f_m(x)$  over  $x$  lying on the grid defined by  $\{\mathcal{R}_j\}$ , then the total is dominated by the largest value, i.e., it behaves as  $O(\epsilon)$ .

**Lemma 5.** *The functions  $\{f_m\}$  constructed in Section III-A are such that  $\bar{v}_m^j$  satisfy the following:*

- 1)  $\sum_{j=1}^M \bar{v}_m^j = O(\epsilon)$  for all  $m$ ;
- 2)  $\sum_{m=1}^M \bar{v}_m^j = O(\epsilon)$  for all  $j$ ;
- 3)  $\sum_{m=1}^M (\bar{v}_m^j)^2 = O(\epsilon^2)$  for all  $j$ .

*Proof.* The three claims in the lemma statement are proved in a nearly identical manner, so we focus on the first. Recall from (9) that each  $f_m$  is a shifted version of  $g(x) = a_0 h(\frac{x}{\zeta w})$ , where the shifts among the  $m$  values differ by integer multiples of  $w$  in one or more of the  $d$  dimensions. That is, the amount by which  $g$  is shifted to produce  $f_m$  is given by  $w \mathbf{i}_m$  for some  $\mathbf{i}_m = (\mathbf{i}_{m,1}, \dots, \mathbf{i}_{m,d}) \in \mathbb{Z}^d$ , and the vectors  $\mathbf{i}_m$  and  $\mathbf{i}_{m'}$  differ in at least one coordinate when  $m \neq m'$ .

In the following, we let  $\mathbf{1}$  denote the vector of ones, and we write  $\mathbf{i} \preceq \mathbf{i}'$  for element-wise inequalities.

Recalling the definitions of  $\mathcal{R}_j$  and  $\bar{v}_m^j$  following Lemma 3, we have

$$\sum_{j=1}^M \bar{v}_m^j = \sum_{j=1}^M \max_{x \in \mathcal{R}_j} f_m(x) \quad (45)$$

$$\leq \sum_{\mathbf{i} \in \mathbb{Z}^m} \max_{i w \preceq x \preceq (\mathbf{i}+1)w} g(x) \quad (46)$$

$$= \frac{2\epsilon}{h(0)} \sum_{\mathbf{i} \in \mathbb{Z}^m} \max_{\mathbf{i} \preceq x' \preceq \mathbf{i}+1} h(\zeta x'), \quad (47)$$

where (46) follows by expanding the sum over  $\mathbf{i} \in \{\mathbf{i}_m : m = 1, \dots, M\}$  to all  $\mathbf{i} \in \mathbb{Z}^d$  and recalling that each  $f_m$  is a shifted version of  $g$ , and (47) follows by substituting (11) and applying the change of variable  $x' = \frac{x}{w}$ .

Since  $\frac{2\epsilon}{h(0)} = O(\epsilon)$ , it only remains to show that the summation in (47) is  $O(1)$ . To do this, we note that since the bump function  $H(\xi)$  in (10) is infinitely differentiable, its inverse Fourier transform  $h(x)$  decays to zero as  $\|x\|_2 \rightarrow \infty$ , at a rate faster than any finite power of  $\frac{1}{\|x\|_2}$  [15]. The  $2d$ -th power suffices for our purposes: There exist constants  $C_1, C_2$  such that  $h(\zeta x') \leq \frac{C_1}{\zeta^2 \|x'\|_2^{2d}}$  whenever  $\|x'\|_2^2 \geq C_2$ . Hence, we bound the sum in (47) by considering two separate cases:

- If the vector  $\mathbf{i}$  is such that  $\|x'\|_2^2 < C_2$  for some  $\mathbf{i} \preceq x' \preceq (\mathbf{i}+1)$ , we upper bound  $h(\zeta x')$  by its maximum value  $h(0)$ . This case can only occur for a finite number of  $\mathbf{i}$ , and hence the total contribution from such  $\mathbf{i}$  is  $O(1)$  (recall that we assume  $d = O(1)$ ).
- For the vectors  $\mathbf{i}$  is such that  $\|x'\|_2^2 \geq C_2$  for all  $\mathbf{i} \preceq x' \preceq (\mathbf{i}+1)$ , we apply the upper bound  $h(\zeta x') \leq \frac{C_1}{\zeta^2 \|x'\|_2^{2d}}$ . The total contribution is again finite, since the sum of  $\frac{1}{\|\mathbf{i}\|_2^{2d}}$  over all  $\mathbf{i} \neq \mathbf{0}$  is finite. For instance, this can be seen by upper bounding  $\frac{1}{\|\mathbf{i}\|_2^{2d}} \leq \prod_{l=1}^d \frac{1}{i_l^2}$  (since  $\|\mathbf{i}\|_2^2 \geq i_l^2$ ), which implies that  $\sum_{\mathbf{i} \geq \mathbf{1}} \frac{1}{\|\mathbf{i}\|_2^{2d}} \leq (\sum_{i=1}^{\infty} \frac{1}{i^2})^d < \infty$ .

Combining the above, we conclude that  $\sum_{j=1}^M \bar{v}_m^j = O(\epsilon)$ .  $\square$

### B. Completion of the Proof of Theorem 1

**Applying Lemma 3:** Letting  $v^{(T)} = f(x^{(T)})$  be the reward of the recommended point, we have

$$\mathbb{E}_m[v^{(T)}] = \mathbb{E}_m[f(x^{(T)})] \quad (48)$$

$$\leq \sum_{j=1}^M \mathbb{P}_m[x^{(T)} \in \mathcal{R}_j] \bar{v}_m^j \quad (49)$$

$$\leq \sum_{j=1}^M \bar{v}_m^j \left( \mathbb{P}_0[x^{(T)} \in \mathcal{R}_j] + \sqrt{\sum_{j'=1}^M \mathbb{E}_0[N_{j'}] \bar{D}_m^{j'}} \right) \quad (50)$$

where (49) follows by the same argument as (42), and (50) follows from Lemma 3 and (40), with the former setting  $a(\mathbf{y}) = \mathbb{1}\{x^{(T)} \in \mathcal{R}_j\}$  and using the fact that  $\mathbb{1}\{\cdot\} \in [0, 1]$  (note that  $x^{(T)}$  is a function

of  $\mathbf{y}$ ). Averaging over  $m$  gives the following bound on the average reward:

$$\mathbb{E}[v^{(T)}] \leq \frac{1}{M} \sum_{m=1}^M \sum_{j=1}^M \bar{v}_m^j \left( \mathbb{P}_0[x^{(T)} \in \mathcal{R}_j] + \sqrt{\sum_{j'=1}^M \mathbb{E}_0[N_{j'}] \bar{D}_m^{j'}} \right). \quad (51)$$

**Bounding the two terms in (51):** Using the second part of Lemma 5, we can bound the first term in (51) as follows:

$$\frac{1}{M} \sum_{m=1}^M \sum_{j=1}^M \bar{v}_m^j \mathbb{P}_0[x^{(T)} \in \mathcal{R}_j] = \frac{1}{M} \sum_{j=1}^M \left( \sum_{m=1}^M \bar{v}_m^j \right) \mathbb{P}_0[x^{(T)} \in \mathcal{R}_j] \quad (52)$$

$$= O\left(\frac{\epsilon}{M}\right) \sum_{j=1}^M \mathbb{P}_0[x^{(T)} \in \mathcal{R}_j] \quad (53)$$

$$= O\left(\frac{\epsilon}{M}\right), \quad (54)$$

since  $\sum_{j=1}^M \mathbb{P}_0[x^{(T)} \in \mathcal{R}_j] = 1$ .

The second term in (51) is bounded as follows:

$$\begin{aligned} & \frac{1}{M} \sum_{m=1}^M \sum_{j=1}^M \bar{v}_m^j \sqrt{\sum_{j'=1}^M \mathbb{E}_0[N_{j'}] \bar{D}_m^{j'}} \\ &= \frac{1}{\sqrt{2} \cdot \sigma} \cdot \frac{1}{M} \sum_{m=1}^M \left( \sum_{j=1}^M \bar{v}_m^j \right) \sqrt{\sum_{j'=1}^M \mathbb{E}_0[N_{j'}] (\bar{v}_m^{j'})^2} \end{aligned} \quad (55)$$

$$= O\left(\frac{\epsilon}{\sigma}\right) \cdot \frac{1}{M} \sum_{m=1}^M \sqrt{\sum_{j'=1}^M \mathbb{E}_0[N_{j'}] (\bar{v}_m^{j'})^2} \quad (56)$$

$$\leq O\left(\frac{\epsilon}{\sigma}\right) \cdot \sqrt{\frac{1}{M} \sum_{m=1}^M \sum_{j'=1}^M \mathbb{E}_0[N_{j'}] (\bar{v}_m^{j'})^2} \quad (57)$$

$$= O\left(\frac{\epsilon}{\sigma}\right) \cdot \sqrt{\frac{1}{M} \sum_{j'=1}^M \mathbb{E}_0[N_{j'}] \left( \sum_{m=1}^M (\bar{v}_m^{j'})^2 \right)} \quad (58)$$

$$= O\left(\frac{\epsilon^2}{\sqrt{M}\sigma}\right) \cdot \sqrt{\sum_{j'=1}^M \mathbb{E}_0[N_{j'}]} \quad (59)$$

$$= O\left(\frac{\sqrt{T}\epsilon^2}{\sqrt{M}\sigma}\right), \quad (60)$$

where (55) follows since the divergence associated with a point  $x$  having value  $v(x)$  is  $\frac{v(x)^2}{2\sigma^2}$  (cf., (36)), (56) follows from the first part of Lemma 5, (57) follows from Jensen's inequality, (59) follows from the third part of Lemma 5, and (60) follows from  $\sum_{j'} N_{j'} = T$ .

**Combining and Simplifying:** Substituting (54) and (60) into (51) gives

$$\mathbb{E}[v^{(T)}] \leq C \cdot \epsilon \left( \frac{1}{M} + \frac{\epsilon}{\sigma} \sqrt{\frac{T}{M}} \right) \quad (61)$$

for some constant  $C$ . Since the maximum function value is  $f(x^*) = 2\epsilon$  by the construction in Section III-A, this implies that the simple regret is lower bounded by

$$\mathbb{E}[r^{(T)}] \geq 2\epsilon - C \cdot \epsilon \left( \frac{1}{M} + \frac{\epsilon}{\sigma} \sqrt{\frac{T}{M}} \right) \quad (62)$$

$$= \epsilon \left( 2 - \frac{C}{M} - \frac{C\epsilon}{\sigma} \sqrt{\frac{T}{M}} \right). \quad (63)$$

Noting that  $M \rightarrow \infty$  as  $\frac{\epsilon}{B} \rightarrow 0$  in both (23) and (34), we have for sufficiently small  $\frac{\epsilon}{B}$  that  $\frac{C}{M} \leq \frac{1}{2}$ , and hence

$$\mathbb{E}[r^{(T)}] \geq \epsilon \left( \frac{3}{2} - \frac{C\epsilon}{\sigma} \sqrt{\frac{T}{M}} \right). \quad (64)$$

By equating the bracketed term with one, it follows that if the time horizon satisfies

$$T \leq \frac{M\sigma^2}{4C^2\epsilon^2}, \quad (65)$$

then the average simple regret at time  $T$  is at least  $\epsilon$ .

Theorem 1 now follows by substituting  $M = \Theta((\log \frac{B}{\epsilon})^{d/2})$  into (65) for the squared exponential kernel (cf., (23)), and  $M = \Theta((\frac{B}{\epsilon})^{d/\nu})$  for the Matérn kernel (cf., (34)).

### C. Completion of the Proof of Theorem 2

We initially follow similar steps to those used above for the simple regret, but the later steps become slightly more involved.

**Applying Lemma 3:** Letting  $V_T = \sum_{t=1}^T f(x_t)$  be the total cumulative reward, we have

$$\mathbb{E}_m[V_T] = \sum_{t=1}^T \mathbb{E}_m[f(x_t)] \quad (66)$$

$$\leq \sum_{t=1}^T \sum_{j=1}^M \mathbb{P}_m[j_t = j] \bar{v}_m^j \quad (67)$$

$$= \sum_{j=1}^M \bar{v}_m^j \mathbb{E}_m[N_j] \quad (68)$$

$$\leq \sum_{j=1}^M \bar{v}_m^j \left( \mathbb{E}_0[N_j] + T \sqrt{\sum_{j'=1}^M \mathbb{E}_0[N_{j'}] \bar{D}_m^{j'}} \right), \quad (69)$$

where (67) follows by the same argument as (42), and (69) follows from Lemma 3 and (40), with the former setting  $a(\mathbf{y}) = N_j$  and using the fact that  $N_j \in [0, T]$  (note that  $\{x_t\}_{t=1}^T$  is a function of  $\mathbf{y}$ ).

Averaging over  $m$  gives the following bound on the average total reward:

$$\mathbb{E}[V_T] \leq \frac{1}{M} \sum_{m=1}^M \sum_{j=1}^M \bar{v}_m^j \left( \mathbb{E}_0[N_j] + T \sqrt{\sum_{j'=1}^M \mathbb{E}_0[N_{j'}] \bar{D}_m^{j'}} \right). \quad (70)$$

**Bounding the two terms in (70):** Using the second part of Lemma 5, we can bound the first term in (70) as follows:

$$\frac{1}{M} \sum_{m=1}^M \sum_{j=1}^M \bar{v}_m^j \mathbb{E}_0[N_j] = \frac{1}{M} \sum_{j=1}^M \left( \sum_{m=1}^M \bar{v}_m^j \right) \mathbb{E}_0[N_j] \quad (71)$$

$$= O\left(\frac{\epsilon}{M}\right) \sum_{j=1}^M \mathbb{E}_0[N_j] \quad (72)$$

$$= O\left(\frac{T\epsilon}{M}\right), \quad (73)$$

since  $\sum_{j=1}^M N_j = T$ . Moreover, the second term in (70) was already bounded in (55)–(60):

$$T \cdot \frac{1}{M} \sum_{m=1}^M \sum_{j=1}^M \bar{v}_m^j \sqrt{\sum_{j'=1}^M \mathbb{E}_0[N_{j'}] \bar{D}_m^{j'}} = O\left(\frac{T\sqrt{T}\epsilon^2}{\sqrt{M}\sigma}\right), \quad (74)$$

**Combining and simplifying:** Substituting (73) and (74) into (70) gives

$$\mathbb{E}[V_T] \leq C' \cdot T\epsilon \left( \frac{1}{M} + \frac{\epsilon}{\sigma} \sqrt{\frac{T}{M}} \right) \quad (75)$$

for some constant  $C'$ . Since the maximum function value is  $f(x^*) = 2\epsilon$  by the construction in Section III-A, this implies that the cumulative regret is lower bounded by

$$\mathbb{E}[R_T] \geq 2T\epsilon - C' \cdot T\epsilon \left( \frac{1}{M} + \frac{\epsilon}{\sigma} \sqrt{\frac{T}{M}} \right) \quad (76)$$

$$= T\epsilon \left( 2 - \frac{C'}{M} - \frac{C'\epsilon}{\sigma} \sqrt{\frac{T}{M}} \right). \quad (77)$$

Noting that  $M \rightarrow \infty$  as  $\frac{\epsilon}{B} \rightarrow 0$  in both (23) and (34), we have for sufficiently small  $\frac{\epsilon}{B}$  that  $\frac{C'}{M} \leq \frac{1}{2}$ , and hence

$$\mathbb{E}[R_T] \geq T\epsilon \left( \frac{3}{2} - \frac{C'\epsilon}{\sigma} \sqrt{\frac{T}{M}} \right). \quad (78)$$

By equating the bracketed term with one, it follows that if the time horizon satisfies

$$T \leq \frac{M\sigma^2}{4(C')^2\epsilon^2}, \quad (79)$$

then the average cumulative regret at time  $T$  is at least  $T\epsilon$ . For convenience, we also note the following equivalent form of (79):

$$\epsilon \leq \sqrt{\frac{M\sigma^2}{4(C')^2T}}, \quad (80)$$



While  $\epsilon$  represented the target simple regret in Section V-B, for cumulative regret it can be treated as a parameter that we can set to our liking, subject to the fact that we assumed in the analysis that  $\frac{\epsilon}{B}$  is sufficiently small. Hence, we simply need to choose  $\epsilon$  to ensure that (79) (or equivalently, (80)) holds, and substitute this value into the lower bound  $T\epsilon$ .

To do this, we choose  $\epsilon$  such that (80) nearly holds with equality:

$$\epsilon \geq \sqrt{\frac{M\sigma^2}{8(C')^2T}}. \quad (81)$$

Note that the right-hand side depends on  $M$ , while  $M$  itself is expressed in terms of  $\epsilon$  in (23) and (34). Hence, it remains to show that these choices are consistent, and in particular, to lower bound  $\epsilon$  (and hence the cumulative regret bound  $T\epsilon$ ) in terms of  $T$  alone.

**Application to the SE kernel:** For the SE kernel, we have from the choice  $M = \Theta\left(\left(\log \frac{B}{\epsilon}\right)^{d/2}\right)$  in (23), along with the upper and lower bounds on  $\epsilon$  in (80) and (81), that

$$\epsilon = \Theta\left(\sqrt{\frac{\sigma^2}{T}\left(\log \frac{B}{\epsilon}\right)^{d/2}}\right). \quad (82)$$

We therefore deduce that

$$\log \frac{B}{\epsilon} = \log \sqrt{\frac{TB^2}{\sigma^2}} - \log \left(\Theta(1)\left(\log \frac{B}{\epsilon}\right)^{d/4}\right) \quad (83)$$

by a direct substitution of (82). Now, since we have assumed that  $d = O(1)$ , we find that the second term behaves as  $O(\log \log \frac{B}{\epsilon})$ , which is at most  $\frac{1}{2} \log \frac{B}{\epsilon}$  when  $\frac{\epsilon}{B}$  is sufficiently small. By moving this term to the left-hand side of (83), we find that  $\log \frac{B}{\epsilon} = \Theta\left(\log \frac{B^2T}{\sigma^2}\right)$ , and substitution into (82) yields  $\epsilon = \Theta\left(\sqrt{\frac{\sigma^2}{T}\left(\log \frac{B^2T}{\sigma^2}\right)^{d/2}}\right)$ , meaning that the lower bound  $T\epsilon$  yields (5).

Note that the behavior  $\epsilon = \Theta\left(\sqrt{\frac{\sigma^2}{T}\left(\log \frac{B^2T}{\sigma^2}\right)^{d/2}}\right)$  reveals that  $\frac{\epsilon}{B}$  is indeed arbitrary small when the implied constant in the theorem assumption  $\frac{\sigma}{B} = O(\sqrt{T})$  is sufficiently small (note that  $z\left(\log \frac{1}{z}\right)^d \rightarrow 0$  as  $z \rightarrow 0$ ). This justifies the assumption of  $\frac{\epsilon}{B}$  being sufficiently small throughout our analysis.

**Application to the Matérn kernel:** For the Matérn kernel, we have from (34) that  $M = \Theta\left(\left(\frac{B}{\epsilon}\right)^{d/\nu}\right)$ , and substitution into (81) gives  $\epsilon^2 = \Theta\left(\frac{\sigma^2 B^{d/\nu}}{T\epsilon^{d/\nu}}\right)$ , or equivalently  $T = \Theta\left(\frac{\sigma^2 B^{d/\nu}}{\epsilon^{2+d/\nu}}\right)$ . This, in turn, implies that  $\epsilon = \Theta\left(\sigma^{\frac{2}{2+d/\nu}} B^{\frac{d/\nu}{2+d/\nu}} T^{\frac{-1}{2+d/\nu}}\right)$ , and hence the lower bound  $T\epsilon$  yields (6). Once again, we find that the assumption  $\frac{\sigma}{B} = O(\sqrt{T})$  (with a sufficiently small implied constant) in the theorem statement ensures that  $\frac{\epsilon}{B}$  is sufficiently small.

#### D. High-Probability Regret Bounds

Here we discuss how to adapt the above proofs to provide lower bounds on the regret that hold with a given probability  $1 - \delta$ .

For the simple regret, recall that we considered functions where all points are at most  $4\epsilon$ -suboptimal (implying  $r^{(T)} \leq 4\epsilon$ ), and showed that  $\mathbb{E}[r^{(T)}] \geq \epsilon$  for all  $T < T_\epsilon$  and suitably-defined  $T_\epsilon$ . For such

$T$ , it follows from the reverse Markov inequality (i.e., Markov's inequality applied to the random variable  $4\epsilon - r^{(T)}$ ) that

$$\mathbb{P}[r^{(T)} \geq \eta\epsilon] \geq \frac{\epsilon - \eta\epsilon}{4\epsilon - \eta\epsilon} = \frac{1 - \eta}{4 - \eta} \quad (84)$$

for any  $\eta \in (0, 1)$ . By choosing  $\eta$  sufficiently small, and renaming  $\eta\epsilon$  as  $\epsilon'$ , it follows that in order to achieve some target simple regret  $\epsilon'$  with any constant probability above  $\frac{3}{4}$ , a lower bound of the same form as the average regret bound holds.

An analogous argument applies for the cumulative regret upon recalling that we chose  $\epsilon$  such that the average cumulative regret was lower bounded by  $T\epsilon$ , while also considering functions such that the cumulative regret can never exceed  $4T\epsilon$ .

## VI. CONCLUSION

We have given, to our knowledge, the first lower bounds on regret for (non-Bayesian) Gaussian process bandit optimization in the presence of noise, considering both the simple regret and the cumulative regret. These bounds nearly match existing upper bounds for the squared exponential kernel, with the gaps being more significant for the Matérn kernel.

An immediate direction for future research is to settle our conjecture in Section II regarding the presence of  $\gamma_T$  vs.  $\gamma_T^2$  in the existing upper bounds, which would close the gaps present for the squared exponential kernel, and bring the bounds for the Matérn kernel closer together. Another interesting direction is to provide lower bounds and improved upper bounds for the *Bayesian* setting. In this setting, the “needle in haystack” type functions considered in the present paper are extremely unlikely to be observed, and it is thus reasonable to expect that the optimal scaling laws may be significantly milder than those for the RKHS setting.

## APPENDIX

We repeat the short proof from [3] here for completeness:

$$\mathbb{E}_m[a(\mathbf{y})] - \mathbb{E}_0[a(\mathbf{y})] = \int a(\mathbf{y})(P_m(\mathbf{y}) - P_0(\mathbf{y}))d\mathbf{y} \quad (85)$$

$$\leq \int_{P_m(\mathbf{y}) \geq P_0(\mathbf{y})} a(\mathbf{y})(P_m(\mathbf{y}) - P_0(\mathbf{y}))d\mathbf{y} \quad (86)$$

$$\leq A \int_{P_m(\mathbf{y}) \geq P_0(\mathbf{y})} (P_m(\mathbf{y}) - P_0(\mathbf{y}))d\mathbf{y} \quad (87)$$

$$= \frac{A}{2} \|P_0 - P_m\|_1 \quad (88)$$

$$\leq A\sqrt{D(P_0\|P_m)}, \quad (89)$$

where (86)–(87) follow from the assumption that  $a(\mathbf{y}) \in [0, A]$ , (88) is a standard property of the  $\ell_1$ -norm, and (89) follows from Pinsker's inequality.

## ACKNOWLEDGMENT

This work was supported in part by the European Commission under Grant ERC Future Proof, SNF 200021-146750 and SNF CRSII2-147633, and the ‘EPFL Fellows’ program (Horizon2020 665667).

## REFERENCES

- [1] N. Aronszajn. Theory of reproducing kernels. *Trans. Amer. Math. Soc.*, 68(3):337–404, 1950.
- [2] J.-Y. Audibert and S. Bubeck. Best arm identification in multi-armed bandits. In *Conf. Learning Theory (COLT)*, 2010.
- [3] P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire. Gambling in a rigged casino: The adversarial multi-armed bandit problem. Technical report, <http://www.dklevine.com/archive/refs4462.pdf>, 1998.
- [4] I. Bogunovic, J. Scarlett, A. Krause, and V. Cevher. Truncated variance reduction: A unified approach to Bayesian optimization and level-set estimation. In *Conf. Neur. Inf. Proc. Sys. (NIPS)*, 2016.
- [5] S. Bubeck and N. Cesa-Bianchi. *Regret Analysis of Stochastic and Nonstochastic Multi-Armed Bandit Problems*. Found. Trend. Mach. Learn. Now Publishers, 2012.
- [6] S. Bubeck, R. Munos, and G. Stoltz. Pure exploration in multi-armed bandits problems. In *Conf. Alg. Learn. Theory (ALT)*, 2009.
- [7] A. D. Bull. Convergence rates of efficient global optimization algorithms. *J. Mach. Learn. Res.*, 12(Oct.):2879–2904, 2011.
- [8] E. Contal, D. Buffoni, A. Robicquet, and N. Vayatis. *Machine Learning and Knowledge Discovery in Databases*, chapter Parallel Gaussian Process Optimization with Upper Confidence Bound and Pure Exploration, pages 225–240. Springer Berlin Heidelberg, 2013.
- [9] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., 2006.
- [10] N. de Freitas, M. Zoghi, and A. J. Smola. Exponential regret bounds for Gaussian process bandits with deterministic observations. In *Proc. Int. Conf. Mach. Learn. (ICML)*, 2012.
- [11] S. Grünewälder, J.-Y. Audibert, M. Opper, and J. Shawe-Taylor. Regret bounds for Gaussian process bandit problems. In *Int. Conf. Art. Intel. Stats. (AISTATS)*, pages 273–280, 2010.
- [12] K. Kawaguchi, L. P. Kaelbling, and T. Lozano-Pérez. Bayesian optimization with exponential convergence. In *Adv. Neur. Inf. Proc. Sys. (NIPS)*, 2015.
- [13] R. Kleinberg, A. Slivkins, and E. Upfal. Multi-armed bandits in metric spaces. In *Proc. ACM Symp. Theory Comp. (STOC)*, pages 681–690, 2008.
- [14] R. D. Kleinberg. Nearly tight bounds for the continuum-armed bandit problem. In *Adv. Neur. Inf. Proc. Sys. (NIPS)*, pages 697–704, 2004.
- [15] K. O. Mead and L. M. Delves. On the convergence rate of generalized Fourier expansions. *IMA J. Appl. Math.*, 12:247–259, 1973.
- [16] G. Raskutti, M. J. Wainwright, and B. Yu. Minimax-optimal rates for sparse additive models over kernel classes via convex programming. *J. Mach. Learn. Res.*, 13(Feb):389–427, 2012.

- [17] G. Raskutti, M. J. Wainwright, and B. Yu. Early stopping and non-parametric regression: An optimal data-dependent stopping rule. *J. Mach. Learn. Res.*, 15(1):335–366, 2014.
- [18] C. E. Rasmussen. Gaussian processes for machine learning. MIT Press, 2006.
- [19] D. Russo and B. Van Roy. An information-theoretic analysis of Thompson sampling. *J. Mach. Learn. Res.*, 17(1):2442–2471, Jan. 2016.
- [20] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas. Taking the human out of the loop: A review of Bayesian optimization. *Proc. IEEE*, 104(1):148–175, 2016.
- [21] N. Srinivas, A. Krause, S. M. Kakade, and M. Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. In *Int. Conf. Mach. Learn. (ICML)*, 2010.
- [22] Z. Wang, B. Zhou, and S. Jegelka. Optimization as estimation with Gaussian processes in bandit settings. In *Int. Conf. Art. Intel. Stats. (AISTATS)*, 2016.
- [23] Z. Wang and N. de Freitas. Theoretical analysis of Bayesian optimisation with unknown Gaussian process hyper-parameters. <http://arxiv.org/abs/1406.7758>, 2014.