Tighter Confidence Bounds for Sequential Kernel Regression

Hamish Flynn
Universitat Pompeu Fabra
Barcelona, Spain
hamishedward.flynn@upf.edu

David Reeb
Bosch Center for Artificial Intelligence
Renningen, Germany
david.reeb@de.bosch.com

Abstract

Confidence bounds are an essential tool for rigorously quantifying the uncertainty of predictions. They are a core component in many sequential learning and decision-making algorithms, with tighter confidence bounds giving rise to algorithms with better empirical performance and better performance guarantees. In this work, we use martingale tail inequalities to establish new confidence bounds for sequential kernel regression. Our confidence bounds can be computed by solving a conic program, although this bare version quickly becomes impractical, because the number of variables grows with the sample size. However, we show that the dual of this conic program allows us to efficiently compute tight confidence bounds. We prove that our new confidence bounds are always tighter than existing ones in this setting. We apply our confidence bounds to kernel bandit problems, and we find that when our confidence bounds replace existing ones, the KernelUCB (GP-UCB) algorithm has better empirical performance, a matching worst-case performance guarantee and comparable computational cost.

1 INTRODUCTION

A confidence sequence for an unknown function $f^\star$ is a sequence of sets $\mathcal{F}_1, \mathcal{F}_2, \ldots$, where each set $\mathcal{F}_t$ contains all functions that could plausibly be $f^\star$, given all data available at time step t. For any point x, the corresponding upper confidence bound, $\max_{f \in \mathcal{F}_t}\{f(x)\}$, is thus the largest value that $f^\star(x)$ could plausibly take, given all data available at time step t.

The uncertainty estimates provided by confidence bounds are exceptionally useful in sequential learning and decision-making problems, such as bandits (Lattimore and Szepesvári, 2020) and reinforcement learning (Sutton and Barto, 2018), where they can be used to design exploration strategies (Auer, 2002), establish stopping criteria (Jamieson et al., 2014) and guarantee safety (Sui et al., 2015; Berkenkamp et al., 2017). In general, tighter confidence bounds give rise to sequential learning and decision-making algorithms with better empirical performance and better performance guarantees. As many modern algorithms use uncertainty estimates for flexible nonparametric function approximators, such as kernel methods, it is vital to establish tight confidence bounds for infinite-dimensional function classes.

In this work, we extend the confidence sequence developed by Flynn et al. (2023) from sequential linear regression to sequential kernel regression. In the kernel case, the corresponding upper confidence bound at any point x is the solution of an *infinite-dimensional* convex program in the kernel Hilbert space. To make this practical, we formulate the optimization problem using the representer theorem, yielding a finite-dimensional convex program, which is the core of our Kernel CMM-UCB method. However, computing the solution becomes expensive quickly, because the number of variables grows with the sample size. One remedy is to use an analytic upper bound on the exact confidence bound, as we do in our Kernel AMM-UCB method. Unfortunately, this analytic confidence bound can be noticeably looser. A better solution is to compute confidence bounds by approximately solving the Lagrangian dual of the original convex program, which can be done cheaply. This results in our Kernel DMM-UCB method, which achieves the best of both worlds: cheap and tight confidence bounds (see Fig. 1).

Our new confidence bounds can be used in kernel bandit algorithms, such as GP-UCB/KernelUCB (Srinivas et al., 2010; Valko et al., 2013). We show that using our confidence bounds preserves the worst-case per-
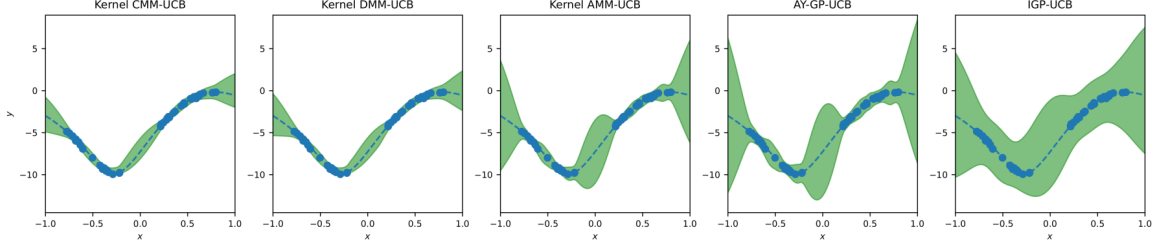
Figure 1: The upper and lower confidence bounds used by our Kernel CMM-UCB method (left), the confidence bounds from (Abbasi-Yadkori, 2012) (AY-GP-UCB) (middle-right), and the confidence bounds used by Improved GP-UCB (IGP-UCB) (Chowdhury and Gopalan, 2017) (right) for a Matérn kernel test function with smoothness $\nu = 3/2$ and lengthscale $\ell = 0.5$. Our Kernel CMM-UCB method – and its relaxed versions DMM-UCB and AMM-UCB – produces confidence bounds that are visibly closer to the ground-truth function (dashed line) than those of AY-GP-UCB and IGP-UCB (cf. also Sec. 6.1).

formance guarantees of KernelUCB. In doing so, we refine the standard KernelUCB regret analysis by using new bounds on the number of large elliptical potentials. With these new tools, we show that in the noiseless setting, KernelUCB with our confidence bounds enjoys a polylog(T) cumulative regret bound when the RBF kernel is used, and a sub-$\sqrt{T}$ regret bound when the Matérn kernel with smoothness $\nu \geq d/\sqrt{2}$ is used. Finally, we demonstrate that the empirical performance of KernelUCB can be vastly improved by using our new confidence bounds.

2 PROBLEM STATEMENT

We consider a sequential kernel regression problem with covariates $x_t \in \mathcal{X}$ and responses $y_t \in \mathbb{R}$, where

$$y_t = f^\star(x_t) + \epsilon_t.$$

The target function $f^\star \in \mathcal{H}$ is an unknown function in a reproducing kernel Hilbert space (RKHS) $\mathcal{H}$ and $\epsilon_1, \epsilon_2, \ldots$ are noise variables. We allow the covariates to be generated in an arbitrary sequential fashion, meaning each x_t can depend on $x_1, y_1, \ldots, x_{t-1}, y_{t-1}$. We assume that the RKHS norm of $f^\star$ is bounded by B, i.e. $\|f^\star\|_{\mathcal{H}} \leq B$. For most of the paper, we assume that, conditioned on $x_1, y_1, \ldots, x_{t-1}, y_{t-1}, x_t$, each ϵ_t is a σ-sub-Gaussian random variable. In Sec. 6.3, we consider the noiseless setting, where $\epsilon_t \equiv 0$, and we observe $y_t = f^\star(x_t)$.

Our first aim is to construct a *confidence sequence* for the function $f^\star$, which we now define formally. For any level $\delta \in (0, 1]$, a $(1 - \delta)$-confidence sequence is a sequence $\mathcal{F}_1, \mathcal{F}_2, \ldots$ of subsets of $\mathcal{H}$, such that each $\mathcal{F}_t$ can be calculated using the data $x_1, y_1, \ldots, x_t, y_t$ and the sequence satisfies the coverage condition

$$\mathbb{P}_{\substack{x_1,x_2,\ldots \\ y_1,y_2,\ldots}} [\forall t \geq 1 : f^\star \in \mathcal{F}_t] \geq 1 - \delta.$$

In words, this coverage condition says that with high probability (over the random draw of $x_1, y_1, x_2, y_2, \ldots$),

$f^\star$ lies in $\mathcal{F}_t$ for all times $t \geq 1$ simultaneously. We call each set in a confidence sequence a confidence set. Our second aim is to construct *confidence bounds* for the value of $f^\star(x)$ at every $x \in \mathcal{X}$. For any level $\delta \in (0, 1]$, an anytime-valid $(1 - \delta)$-upper confidence bound (UCB) is a sequence $u_1, u_2, \ldots$ of functions on $\mathcal{X}$, such that each u_t can be calculated using the data $x_1, y_1, \ldots, x_t, y_t$, and the sequence satisfies

$$\mathbb{P}_{\substack{x_1,x_2,\ldots \\ y_1,y_2,\ldots}} [\forall t \geq 1, \forall x \in \mathcal{X} : f^\star(x) \leq u_t(x)] \geq 1 - \delta.$$

This condition says that with high probability, $u_t(x)$ is an upper bound on $f^\star(x)$ for all $t \geq 1$ and all $x \in \mathcal{X}$ simultaneously. The definition of an anytime-valid lower confidence bound (LCB) is analogous. If we already have a confidence sequence, we can use it to construct upper and lower confidence bounds. In particular, if $\mathcal{F}_1, \mathcal{F}_2, \ldots$ is a $(1 - \delta)$-confidence sequence, then with probability at least $1 - \delta$, for all $t \geq 1$ and all $x \in \mathcal{X}$, we have

$$\min_{f \in \mathcal{F}_t} \{f(x)\} \leq f^\star(x) \leq \max_{f \in \mathcal{F}_t} \{f(x)\}. \qquad (1)$$

Therefore, as long as we can solve the maximisation/minimisation problems in (1), we can compute anytime-valid confidence bounds from a confidence sequence. We use $\text{UCB}_{\mathcal{F}_t}(x)$ to denote the RHS of (1), i.e. $\text{UCB}_{\mathcal{F}_t}(x) := \max_{f \in \mathcal{F}_t} \{f(x)\}$, and $\text{LCB}_{\mathcal{F}_t}(x)$ to denote the LHS of (1).

Notation. For any $n \in \mathbb{N}$, $[n] := \{1, \ldots, n\}$. We use matrix notation to denote the inner product $f_1^\top f_2 := \langle f_1, f_2 \rangle_{\mathcal{H}}$ and outer product $f_1 f_2^\top := f_1 \langle f_2, \cdot \rangle_{\mathcal{H}}$ between two functions $f_1, f_2 \in \mathcal{H}$. We denote the kernel function by $k(\cdot, \cdot)$. For a sequence of observations $x_1, \ldots, x_t$, we define $\boldsymbol{k}_t(x) := [k(x, x_1), \ldots, k(x, x_t)]^\top$. We let $\Phi_t : \mathcal{H} \rightarrow \mathbb{R}^t$ be the linear operator from the RKHS $\mathcal{H}$ to $\mathbb{R}^t$, which is defined by the mapping $f \mapsto [\langle k(x_1, \cdot), f \rangle_{\mathcal{H}}, \ldots, \langle k(x_t, \cdot), f \rangle_{\mathcal{H}}]^\top = [f(x_1), \ldots, f(x_t)]^\top$. This means that we can write

$\boldsymbol{k}_t(x) = \Phi_t k(x, \cdot)$. The transpose (or adjoint) $\Phi_t^\top : \mathbb{R}^t \to \mathcal{H}$ of Φ_t is described by the mapping $\boldsymbol{v} \mapsto \sum_{s=1}^t k(x_s, \cdot) v_s$. We define the kernel matrix as $\boldsymbol{K}_t = \Phi_t \Phi_t^\top$, where $\Phi_t \Phi_t^\top : \mathbb{R}^t \to \mathbb{R}^t$ is described by the mapping $\boldsymbol{v} \mapsto [\boldsymbol{k}_t(x_1)^\top \boldsymbol{v}, \ldots, \boldsymbol{k}_t(x_t)^\top \boldsymbol{v}]^\top$. We can equivalently define $\boldsymbol{K}_t$ as the $t \times t$ matrix whose i, j^{th} element is $k(x_i, x_j)$. Finally, we define the empirical covariance operator $\boldsymbol{V}_t = \Phi_t^\top \Phi_t$, where $\Phi_t^\top \Phi_t : \mathcal{H} \to \mathcal{H}$ is defined by the mapping $f \mapsto \sum_{s=1}^t k(x_s, \cdot) \langle k(x_s, \cdot), f \rangle_\mathcal{H} = \sum_{t=1}^s k(x_s, \cdot) f(x_s)$. For $\alpha \geq 0$, we define

$$\mu_{\alpha,t}(x) := \boldsymbol{k}_t(x)^\top (\boldsymbol{K}_t + \alpha \boldsymbol{I})^{-1} \boldsymbol{y}_t, \qquad (2)$$

$$\rho_{\alpha,t}^2(x) := k(x, x) - \boldsymbol{k}_t(x)^\top (\boldsymbol{K}_t + \alpha \boldsymbol{I})^{-1} \boldsymbol{k}_t(x). \quad (3)$$

$\mu_{\alpha,t}(x)$ is the kernel ridge estimate (evaluated at x) or, equivalently, the predictive mean of a standard Gaussian process (GP) posterior with noise α. $\rho_{\alpha,t}^2(x)$ is the predictive variance of a standard GP posterior.

3 RELATED WORK

Several confidence sequences/bounds have been proposed for the sequential kernel regression problem that we consider. The confidence bounds in (Srinivas et al., 2010) are perhaps the most well-known. The two most widely used were developed by Abbasi-Yadkori et al. (2011); Abbasi-Yadkori (2012) and Chowdhury and Gopalan (2017), using concentration inequalities for self-normalised processes (de la Peña et al., 2004, 2009). Subsequently, Durand et al. (2018) derived empirical Bernstein-type versions of these confidence bounds, and Whitehouse et al. (2023) showed that for certain kernels, these confidence bounds can be tightened by regularising in proportion to the smoothness of the kernel. Confidence sequences for this setting have also been derived using online-to-confidence-set conversions (Abbasi-Yadkori et al., 2012; Abbasi-Yadkori, 2012).

With applications to Bayesian optimisation and kernel bandits in mind, Neiswanger and Ramdas (2021) and Emmenegger et al. (2023) developed confidence sequences using martingale tail bounds for sequential likelihood ratios. However, cumulative regret bounds are not established in (Neiswanger and Ramdas, 2021), and the regret bound in (Emmenegger et al., 2023) is worse than the regret bounds in (Abbasi-Yadkori, 2012) and in this paper.

In the setting where $\epsilon_1, \epsilon_2, \ldots$ are bounded deterministic errors, Scharnhorst et al. (2022) derived confidence bounds using convex programming and Lagrangian duality. However, these confidence bounds are not comparable to ours. In our setting, the random errors $\epsilon_1, \epsilon_2, \ldots$ average out to 0, and our confidence bounds contract around $f^\star$. This is not the case for the confidence bounds in (Scharnhorst et al., 2022).

The work most closely related to ours is (Flynn et al., 2023), in which an improved confidence sequence for linear bandits is developed. While (8) in our paper is a simple kernel generalisation of (5) in (Flynn et al., 2023), deriving confidence bounds is more involved in the kernel setting. In particular, $\max_{f \in \mathcal{F}_t} \{f(x)\}$ is now an infinite-dimensional optimisation problem, which makes our DMM-UCB method vital for balancing tightness and efficiency. Moreover, even in the simpler linear setting, the regret analysis in (Flynn et al., 2023) fails to provide a sub-$\sqrt{T}$ rate as σ goes to zero.

4 CONFIDENCE BOUNDS FOR KERNEL REGRESSION

In this section, we describe our new confidence sequences and bounds for sequential kernel regression.

4.1 Martingale Mixture Tail Bounds

We use a martingale tail bound from (Flynn et al., 2023) to derive a data-dependent constraint for $f^\star$. We begin by describing the general setting in which it holds. There is a stochastic process $(Z_t(g) | t \in \mathbb{N}, g \in \mathbb{R})$, indexed by a time t and a real number g, which is adapted to a filtration $(\mathcal{D}_t | t \in \mathbb{N})$. In this paper, $(\mathcal{D}_t | t \in \mathbb{N})$ is any filtration such that y_t and x_{t+1} are both $\mathcal{D}_t$-measurable, e.g. $\mathcal{D}_t = \sigma(x_1, y_1, \ldots, x_t, y_t, x_{t+1})$. In addition, there is a sequence of predictable "guesses" $(g_t \in \mathbb{R} | t \in \mathbb{N})$, and a sequence of predictable random variables $(\lambda_t \in \mathbb{R} | t \in \mathbb{N})$. We use the shorthand $\boldsymbol{g}_t = [g_1, \ldots, g_t]^\top$ and $\boldsymbol{\lambda}_t = [\lambda_1, \ldots, \lambda_t]^\top$. We define

$$M_t(\boldsymbol{g}_t, \boldsymbol{\lambda}_t) := \exp\left(\sum_{s=1}^t \lambda_s Z_s(g_s) - \psi_s(g_s, \lambda_s)\right), \quad (4)$$

$$\psi_t(g_t, \lambda_t) := \ln\left(\mathbb{E}\left[\exp(\lambda_t Z_t(g_t)) | \mathcal{D}_{t-1}\right]\right), \qquad (5)$$

Flynn et al. (2023) show that $M_t(\boldsymbol{g}_t, \boldsymbol{\lambda}_t)$ is a non-negative martingale, which means Ville's inequality (Ville, 1939) can be applied. In particular, for any level $\delta \in (0, 1]$, this tells us that with probability at least $1 - \delta$, $M_t(\boldsymbol{g}_t, \boldsymbol{\lambda}_t) \leq 1/\delta$ for all time steps $t \geq 1$. To obtain the tightest possible inequality, we employ the method of mixtures. That is, we place a distribution P_t on each vector of guesses $\boldsymbol{g}_t$. Flynn et al. (2023) show that if the sequence of distributions $(P_t | t \in \mathbb{N})$ satisfies: (a) P_t is $\mathcal{D}_{t-1}$-measurable; (b) $\int P_t(\boldsymbol{g}_t) d g_t = P_{t-1}(\boldsymbol{g}_{t-1})$ for all t and $\boldsymbol{g}_{t-1}$, then the mixture $\mathbb{E}_{\boldsymbol{g}_t \sim P_t}[M_t(\boldsymbol{g}_t, \boldsymbol{\lambda}_t)]$ is also a non-negative martingale, and Ville's inequality can still be used.

Theorem 4.1 (Theorem 5.1 in (Flynn et al., 2023))**.**
For any $\delta \in (0, 1)$, any sequence of distributions $(P_t | t \in \mathbb{N})$ satisfying (a) and (b), and any sequence of predictable random variables $(\lambda_t | t \in \mathbb{N})$, it holds with

probability at least $1 - \delta$ that

$$\forall t \geq 1: \quad \ln \left(\underset{\boldsymbol{g}_t \sim P_t}{\mathbb{E}} [M_t(\boldsymbol{g}_t, \boldsymbol{\lambda}_t)] \right) \leq \ln(1/\delta). \quad (6)$$

To turn the inequality in Thm. 4.1 into a constraint for $f^\star$, all that remains is to specify the process $Z_t(g)$, the random variables λ_t and the distributions P_t. Intuitively, each g_t is a guess for the value y_t, given the previous data $x_1, y_1, \ldots, x_{t-1}, y_{t-1}, x_t$. We should therefore choose each P_t such that (a) and (b) are satisfied and P_t assigns high probability to the vector $[y_1, \ldots, y_t]^\top$. The choices of $Z_t(g)$ and λ_t determine the shape of the resulting confidence sequence. We choose $Z_t(g) = (g - f^\star(x_t))\epsilon_t$, which leads to a convex quadratic constraint for $f^\star$. As $Z_t(g)$ is linear in the noise variable ϵ_t, $\psi_t(g, \lambda_t)$ can be upper bounded using the σ-sub-Gaussian property of ϵ_t (cf. (15)).

$$\psi_t(g_t, \lambda_t) \leq \tfrac{1}{2}\sigma^2 \lambda_t^2 (g_t - f^\star(x_t))^2.$$

Using this upper bound, choosing $\lambda_t \equiv 1/\sigma^2$, and rearranging (6) (see App. A.1), we obtain an upper bound on the squared Euclidean distance between the vector of function values $\boldsymbol{f}_t^\star := [f^\star(x_1), \ldots, f^\star(x_t)]^\top$ and the response vector $\boldsymbol{y}_t := [y_1, \ldots, y_t]^\top$. With probability at least $1 - \delta$, for all $t \geq 1$,

$$\begin{aligned} \|\boldsymbol{f}_t^* - \boldsymbol{y}_t\|_2^2 \leq\ & 2\sigma^2 \ln(1/\delta) \\ & - 2\sigma^2 \ln \left(\underset{\boldsymbol{g}_t \sim P_t}{\mathbb{E}} \left[\exp \left(-\tfrac{1}{2\sigma^2} \|\boldsymbol{g}_t - \boldsymbol{y}_t\|_2^2 \right) \right] \right). \end{aligned} \quad (7)$$

For the mixture distributions, we choose $P_t = \mathcal{N}(\boldsymbol{0}, c\boldsymbol{K}_t)$, resembling a zero-mean Gaussian process with covariance scaled by $c > 0$. There are several reasons for this choice of P_t. First, for an appropriate covariance scale $c > 0$, the confidence bounds obtained from (7) with $P_t = \mathcal{N}(\boldsymbol{0}, c\boldsymbol{K}_t)$ are always tighter (by constant factors) than those in (Abbasi-Yadkori, 2012) and (Chowdhury and Gopalan, 2017) (see Sec. 6.1). Second, any Gaussian P_t yields a convenient closed-form expression for the expected value in (7) (see App. A.2). Substituting this expression into (7), we obtain

$$\begin{aligned} \|\boldsymbol{f}_t^\star - \boldsymbol{y}_t\|_2^2 \leq\ & \boldsymbol{y}_t^\top \left(\boldsymbol{I} + \tfrac{c}{\sigma^2}\boldsymbol{K}_t \right)^{-1} \boldsymbol{y}_t \quad (8) \\ & + \sigma^2 \ln \det \left(\boldsymbol{I} + \tfrac{c}{\sigma^2}\boldsymbol{K}_t \right) + 2\sigma^2 \ln \tfrac{1}{\delta} \\ =:\ & R_t^2. \end{aligned}$$

For brevity, we leave the dependence of R_t on parameters such as c implicit. This inequality states that with probability at least $1 - \delta$, at every step t, the vector of ground-truth function values $\boldsymbol{f}_t^\star$ lies within a sphere of radius R_t around the observed response vector $\boldsymbol{y}_t$.

4.2 Confidence Sequences

We now describe how the inequality in (8) can be used to construct confidence sequences for $f^\star$. Let $\boldsymbol{f}_t = [f(x_1), \ldots, f(x_t)]^\top$ denote the values of an arbitrary function $f \in \mathcal{H}$ at the points $x_1, \ldots, x_t$. Since, with probability at least $1 - \delta$, (8) holds for all $t \geq 1$, the sets of functions that (at each t) satisfy (8) form a $1 - \delta$ confidence sequence for $f^\star$. However, there is a problem with this approach. For typical kernels, and all $t \geq 1$, the set of functions that satisfies $\|\boldsymbol{f}_t - \boldsymbol{y}_t\|_2 \leq R_t$ is unbounded in some directions. In other words, we can always find a function $f \in \mathcal{H}$ such that $\|\boldsymbol{f}_t - \boldsymbol{y}_t\|_2$ is arbitrarily large. To fix this, we use the assumption that $\|f^\star\|_{\mathcal{H}} \leq B$. We intersect each set in the sequence with the set of all functions that satisfy $\|f\|_{\mathcal{H}} \leq B$, which results in the following confidence sequence.

Lemma 4.2. *For any $c > 0$ and $\delta \in (0, 1]$, it holds with probability at least $1 - \delta$ that for all $t \geq 1$ simultaneously, the function $f^\star$ lies in the set*

$$\mathcal{F}_t = \{f \in \mathcal{H} : \|\boldsymbol{f}_t - \boldsymbol{y}_t\|_2 \leq R_t, \ \|f\|_{\mathcal{H}} \leq B\}.$$

The two constraints that define $\mathcal{F}_t$ are both ellipsoids in $\mathcal{H}$, meaning that $\mathcal{F}_t$ is the intersection of two ellipsoids. In contrast, most existing confidence sets for functions in RKHS's (e.g. (Abbasi-Yadkori, 2012; Chowdhury and Gopalan, 2017)) are single ellipsoids centred at the kernel ridge estimate $\mu_{\alpha,t}$ (see (2)), for some regularisation parameter $\alpha > 0$. It turns out that by taking a weighted sum of the constraints in $\mathcal{F}_t$, we can obtain a single-ellipsoid confidence set $\widetilde{\mathcal{F}}_t \supseteq \mathcal{F}_t$, which is also centred at the kernel ridge estimate.

Corollary 4.3. *For any $c > 0$, any $\alpha > 0$ and any $\delta \in (0, 1]$, it holds with probability at least $1 - \delta$ that for all $t \geq 1$ simultaneously, $f^\star$ lies in the set*

$$\widetilde{\mathcal{F}}_t = \{f \in \mathcal{H} : \|(\boldsymbol{V}_t + \alpha\boldsymbol{I}_{\mathcal{H}})^{1/2}(f - \mu_{\alpha,t})\|_{\mathcal{H}} \leq \widetilde{R}_{\alpha,t}\},$$

where

$$\widetilde{R}_{\alpha,t}^2 := R_t^2 + \alpha B^2 - \boldsymbol{y}_t^\top \boldsymbol{y}_t + \boldsymbol{y}_t^\top \Phi_t (\boldsymbol{V}_t + \alpha\boldsymbol{I}_{\mathcal{H}})^{-1} \Phi_t^\top \boldsymbol{y}_t.$$

The proof (in App. B.2) involves rearranging

$$\|\boldsymbol{f}_t - \boldsymbol{y}_t\|_2^2 + \alpha \|f\|_{\mathcal{H}}^2 \leq R_t^2 + \alpha B^2,$$

into the constraint that defines $\widetilde{\mathcal{F}}_t$. This confidence sequence can be more easily compared to those in (Abbasi-Yadkori, 2012; Chowdhury and Gopalan, 2017).

4.3 Implicit Confidence Bounds

We now turn our attention to the confidence bounds that can be obtained from our confidence sequences,

and how we can compute them. We start by considering the exact upper confidence bound $\text{UCB}_{\mathcal{F}_t}(x) = \max_{f \in \mathcal{F}_t} \{f(x)\}$, which can be written as

$$\max_{f \in \mathcal{H}} \; f(x) \; \text{ s.t. } \; \|\boldsymbol{f}_t - \boldsymbol{y}_t\|_2 \le R_t, \quad \|f\|_{\mathcal{H}} \le B. \quad (9)$$

The optimisation variable in (9) is a function in the possibly infinite-dimensional RKHS $\mathcal{H}$. However, as (9) contains a constraint on the RKHS norm of f, the (generalised) representer theorem (Kimeldorf and Wahba, 1971; Schölkopf et al., 2001) can be applied. This means that the optimising function f in (9) can be expressed as a finite linear combination of the form $\boldsymbol{k}_{t+1}(x)^\top \boldsymbol{w}$, where $\boldsymbol{k}_{t+1}(x) := [k(x, x_1), \dots, k(x, x_t), k(x, x)]^\top$. When we substitute the functional form $f(\cdot) = \boldsymbol{k}_{t+1}(\cdot)^\top \boldsymbol{w}$ into the objective and both constraints, we obtain the finite-dimensional convex program in Thm. 4.4. (see App. B.3).

Theorem 4.4. $\text{UCB}_{\mathcal{F}_t}(x)$ *equals the solution of the convex program*

$$\max_{\boldsymbol{w} \in \mathbb{R}^{t+1}} \; \boldsymbol{k}_{t+1}(x)^\top \boldsymbol{w} \quad (10)$$

$$\text{s.t.} \quad \|\boldsymbol{K}_{t,t+1} \boldsymbol{w} - \boldsymbol{y}_t\|_2 \le R_t,$$
$$\|\boldsymbol{L}_{t+1} \boldsymbol{w}\|_2 \le B.$$

where $\boldsymbol{K}_{t,t+1} := [\boldsymbol{K}_t, \boldsymbol{k}_t(x)] \in \mathbb{R}^{t \times (t+1)}$ *denotes kernel matrix with added last column* $\boldsymbol{k}_t(x)$*, and* $\boldsymbol{L}_{t+1}$ *is any matrix satisfying* $\boldsymbol{L}_{t+1}^\top \boldsymbol{L}_{t+1} = \boldsymbol{K}_{t+1}$ *(e.g. the right Cholesky factor of* $\boldsymbol{K}_{t+1}$*).*

This is a $(t + 1)$-dimensional second-order cone program. Using interior point methods, the solution can be computed to ϵ accuracy in $\mathcal{O}(t^3 \text{polylog}(1/\epsilon))$ steps (see Sec. 10.2 of Nemirovski, 2004). To obtain cheaper confidence bounds, we consider the Lagrangian dual of (9). We derive the dual in App. B.4 and give a short sketch here. The primal in (9) can be written as $\max_{f \in \mathcal{H}} \min_{\eta_1, \eta_2 > 0} L(f, \eta_1, \eta_2)$, where

$$L(f, \eta_1, \eta_2) = f(x) + \eta_1(R_t^2 - \|\boldsymbol{f}_t - \boldsymbol{y}_t\|_2^2) + \eta_2(B^2 - \|f\|_{\mathcal{H}}^2),$$

is the Lagrangian and η_1, η_2 are the Lagrange multipliers. The dual problem $\min_{\eta_1, \eta_2 > 0} \max_{f \in \mathcal{H}} L(f, \eta_1, \eta_2)$ is obtained by swapping the min and the max. Since the Lagrangian is concave and quadratic in f, the inner maximisation in the dual can be solved analytically, which gives the following expression for the dual.

Theorem 4.5. *The dual of (9) is*

$$\min_{\substack{\eta_1 > 0 \\ \eta_2 > 0}} \left\{ \mu_{\eta_2/\eta_1, t}(x) - \frac{\rho_{\eta_2/\eta_1, t}^2(x)}{4\eta_2} + \eta_1 \widetilde{R}_{\eta_2/\eta_1, t}^2 \right\}. \quad (11)$$

See (2) and (3) for the definitions of $\mu_{\eta_2/\eta_1, t}(x)$ and $\rho_{\eta_2/\eta_1, t}^2(x)$. By weak duality, the solution of (11) is

always an upper bound on the solution of the primal problem in (9). Whenever there exists an f that lies in the interior of the ellipsoids that define $\mathcal{F}_t$, as is usually the case, then strong duality holds and the solutions of (9) and (11) are actually equal. In App. B.4, we show that (11) can be further reduced to a one-dimensional optimisation problem. Using the substitution $\eta_2 = \alpha \eta_1$ (for $\alpha > 0$), the solution of (11) is equal to

$$\min_{\alpha > 0} \left\{ \mu_{\alpha, t}(x) + \tfrac{1}{\sqrt{\alpha}} \widetilde{R}_{\alpha, t} \rho_{\alpha, t}(x) \right\}. \quad (12)$$

The objective function in (12) is not always convex in α, though empirical evidence suggests that it is quasiconvex in α (see App. F.4), and could therefore be solved using the bisection method.

4.4 Explicit Confidence Bounds

In this section, we focus on explicit confidence bounds. These are mainly useful for deriving worst-case regret bounds with explicit dependence on T (see Sec. 6.2). From (12), for any fixed $\alpha > 0$, we have

$$\text{UCB}_{\mathcal{F}_t}(x) \; \le \; \mu_{\alpha, t}(x) + \tfrac{1}{\sqrt{\alpha}} \widetilde{R}_{\alpha, t} \rho_{\alpha, t}(x). \quad (13)$$

It turns out that this upper bound on $\text{UCB}_{\mathcal{F}_t}(x)$ is equal to $\text{UCB}_{\widetilde{\mathcal{F}}_t}(x)$, where $\widetilde{\mathcal{F}}_t$ is the set from Cor. 4.3.

Corollary 4.6. *For all* $t \ge 1$ *and* $x \in \mathcal{X}$*,*

$$\text{UCB}_{\widetilde{\mathcal{F}}_t}(x) \; = \; \mu_{\alpha, t}(x) + \tfrac{1}{\sqrt{\alpha}} \widetilde{R}_{\alpha, t} \rho_{\alpha, t}(x). \quad (14)$$

This follows from the proof of Thm. 4.5 and the derivation of (12). In particular, the dual of $\max_{f \in \widetilde{\mathcal{F}}_t} \{f(x)\}$ is a minimisation problem with a single variable, which can be solved analytically.

4.5 Confidence Bound Maximisation

We now discuss how our upper confidence bounds can be maximised w.r.t. $x \in \mathcal{X}$, which is an important consideration for kernel bandits (see Sec. 5). If $\mathcal{X}$ has finite (and not too large) cardinality, then $\text{UCB}_{\mathcal{F}_t}(x)$ (and upper bounds on it) can be maximised exactly.

If $\mathcal{X}$ is a continuous subset of $\mathbb{R}^d$, maximising $\text{UCB}_{\mathcal{F}_t}(x)$ over $\mathcal{X}$ exactly is intractable in general, though other similar confidence bounds have the same limitation. There are at least two options for approximate maximisation. One option is to use gradient-based methods, as in (Flynn et al., 2023). Gradients (w.r.t. x) of the confidence bounds in (10), (11) and (12) can be computed using differentiable convex optimisation methods (Agrawal et al., 2019) or automatic implicit differentiation methods (Look et al., 2020; Blondel et al., 2022). Gradient-based maximisation works well in practice, though there are no rigorous approximation guarantees

as $\text{UCB}_{\mathcal{F}_t}(x)$ is (in general) a non-concave function of x. A second option is to discretise $\mathcal{X}$ and optimise over the discretisation (see e.g. Sec. 3.1 in (Li and Scarlett, 2022)). If the kernel function $k(x, x')$ is Lipschitz (as the RBF and Matérn kernels are), then the discretisation error can be controlled. However, the size of a sufficiently fine discretisation grows exponentially in the dimension d, so this quickly becomes impractical.

5 KERNEL BANDITS

To demonstrate the utility of our confidence bounds, we apply them to the stochastic kernel bandit problem.

5.1 Stochastic Kernel Bandits

A learner plays a game over a sequence of T rounds. In each round t, the learner observes an action set $\mathcal{X}_t \subseteq \mathcal{X}$ and must choose an action $x_t \in \mathcal{X}_t$. The learner then receives a reward $y_t = f^\star(x_t) + \epsilon_t$. The reward function $f^\star : \mathcal{X} \to \mathbb{R}$ is an unknown function in an RKHS. As before, we assume that $f^\star$ has bounded norm, i.e. $\|f^\star\|_{\mathcal{H}} \leq B$, and that (in the noisy setting) $\epsilon_1, \epsilon_2, \ldots, \epsilon_T$ are conditionally σ-sub-Gaussian.

The goal of the learner is to choose a sequence of actions that minimises the cumulative regret, which is the difference between the total expected reward of the learner and the optimal strategy. The regret in round t is $r_t = f^\star(x_t^\star) - f^\star(x_t)$, where $x_t^\star = \text{argmax}_{x \in \mathcal{X}_t}\{f^\star(x)\}$. After T rounds, the cumulative regret is $r_{1:T} = \sum_{t=1}^{T} r_t$.

5.2 Kernel Bandit Algorithms

In Algorithm 1, we present a recipe for a kernel bandit algorithm that uses confidence bounds to select actions.

Algorithm 1 KernelUCB

Input: Upper confidence bounds $u_0, u_1, u_2, \ldots$
for $t = 1$ **to** T **do**
  Play $x_t = \text{argmax}_{x \in \mathcal{X}_t}\{u_{t-1}(x)\}$
  Observe $y_t = f^\star(x_t) + \epsilon_t$
end for

In each round t, we use all previous observations $\{(x_s, y_s)\}_{s=1}^{t-1}$ to construct the UCB u_{t-1}, and then select the action that maximises u_{t-1}. Kernel bandit algorithms of this form have appeared under different names, such as GP-UCB (Srinivas et al., 2010) and KernelUCB (Valko et al., 2013), based on the confidence bounds that they use. We therefore use the names KernelUCB and GP-UCB synonymously. When we use the confidence bounds from Thm. 3.11 in (Abbasi-Yadkori, 2012), we call Algorithm 1 AY-GP-UCB. If we use the confidence bounds from Thm. 2 in (Chowdhury

and Gopalan, 2017), then Algorithm 1 is the Improved GP-UCB algorithm (IGP-UCB).

5.3 Kernel CMM-UCB

For our first variant of KernelUCB, we use our exact confidence bound $\text{UCB}_{\mathcal{F}_t}(x)$. In particular, we set $u_t(x)$ to be the numerical solution of (10). The resulting algorithm is called Kernel Convex Martingale Mixture UCB (Kernel CMM-UCB).

5.4 Kernel DMM-UCB

For our second variant of KernelUCB, we use the confidence bound in (12) that comes from the dual problem associated with $\max_{f \in \mathcal{F}_t}\{f(x)\}$. In the interest of computational efficiency, we replace $\min_{\alpha > 0}$ in (12) with $\min_{\alpha \in A}$, for a small grid A of values of α. In particular, we set

$$u_t(x) = \min_{\alpha \in A}\left\{\mu_{\alpha,t}(x) + \frac{1}{\sqrt{\alpha}}\widetilde{R}_{\alpha,t}\rho_{\alpha,t}(x)\right\}.$$

We call the resulting algorithm Kernel Dual Martingale Mixture UCB (Kernel DMM-UCB). Because we only optimise α over the grid A, this upper confidence bound is looser than the exact confidence bound used by Kernel CMM-UCB. In return, the cost of computing this UCB is only a factor of $|A|$ more than the cost of computing the analytic UCB in Cor. 4.6.

5.5 Kernel AMM-UCB

For our third variant of KernelUCB, we set $u_t(x)$ to be the analytic UCB from Cor. 4.6 with a fixed value of α. The resulting algorithm is called Kernel Analytic Martingale Mixture UCB (Kernel AMM-UCB).

6 THEORETICAL ANALYSIS

In this section, we show that our analytic confidence bound from Cor. 4.6 is always tighter than the confidence bounds from Thm. 3.11 in (Abbasi-Yadkori, 2012) and Thm. 2 in (Chowdhury and Gopalan, 2017), which hold under the same conditions and assumptions. In addition, we establish cumulative regret bounds for our KernelUCB algorithms. We begin by formally stating the assumptions under which our analysis holds. We assume that $\mathcal{X}$ is a compact subset of $\mathbb{R}^d$. In the noisy setting, we assume that the noise variables are conditionally σ-sub-Gaussian, which means

$$\forall \lambda \in \mathbb{R}, \quad \mathbb{E}\left[\exp(\lambda \epsilon_t)|\mathcal{D}_{t-1}\right] \leq \exp(\lambda^2 \sigma^2 / 2). \quad (15)$$

This implies that $\mathbb{E}[\epsilon_t|\mathcal{D}_{t-1}] = 0$. For some $B, C > 0$, we assume that $\|f^\star\|_{\mathcal{H}} \leq B$ and $\sup_x |f^\star(x)| \leq C$. If the kernel is bounded by $\kappa > 0$, i.e. $\sup_{x,x'}|k(x, x')| \leq \kappa$, then $\|f^\star\|_{\mathcal{H}} \leq B$ implies that $C \leq \kappa B$.

6.1 Tighter Confidence Bounds

The confidence bounds in (Abbasi-Yadkori, 2012) and (Chowdhury and Gopalan, 2017) are the same as our analytic UCB in (14), except that the (scaled) radius $\widetilde{R}_{\alpha,t}/\sqrt{\alpha}$ is replaced with a different quantity. For any values of the regularisation parameters $\lambda > 0$ and $\eta > 0$ used in (Abbasi-Yadkori, 2012) and (Chowdhury and Gopalan, 2017) respectively, we show that we can set c and α such that $\widetilde{R}_{\alpha,t}/\sqrt{\alpha}$ is strictly less than these other radius quantities. This result is stated in Thm. 6.1 and proved in App. C.2.

Theorem 6.1. *For every $\delta \in (0,1]$ and every value of the regularisation parameter $\lambda > 0$, the RHS of (14) with $c = \sigma^2/\lambda$ and $\alpha = \lambda$ is strictly less than the UCB from Thm. 3.11 in (Abbasi-Yadkori, 2012).*

For every $\delta \in (0,1]$ and every value of the regularisation parameter $\eta > 0$, the RHS of (14) with $c = \sigma^2/(1+\eta)$ and $\alpha = 1 + \eta$ is strictly less than the UCB from Thm. 2 in (Chowdhury and Gopalan, 2017).

6.2 Cumulative Regret Bounds

We bound the cumulative regret in terms of two kernel-dependent quantities. The first is called the maximum information gain, and is defined as

$$\gamma_t(\alpha) := \max_{x_1,\dots,x_t} \left\{ \tfrac{1}{2} \ln \left(\det \left(\tfrac{1}{\alpha} \boldsymbol{K}_t + \boldsymbol{I} \right) \right) \right\}.$$

The second is defined as $\tau_t(\alpha) := |\mathcal{T}_t(\alpha)|$, where

$$\mathcal{T}_t(\alpha) = \{s \in [t] : \tfrac{1}{\alpha} \rho_{\alpha,s-1}^2(x_s) \geq 1\}. \quad (16)$$

The quantity $\tau_t(\alpha)$ counts the number of rounds in which the scaled variance $\rho_{\alpha,s-1}^2(x_s)/\alpha$, also known as the elliptical potential, is large. We call $\tau_t(\alpha)$ the elliptical potential count. First, we upper bound our radius quantity $\widetilde{R}_{\alpha,t}$ in terms of the maximum information gain. We show that when $\alpha = \sigma^2/c$, the quadratic terms in $\widetilde{R}_{\alpha,t}$ cancel out (see Lemma C.2), and we are left with

$$\widetilde{R}_{\sigma^2/c,t} \leq \sqrt{2\sigma^2 \gamma_t(\sigma^2/c) + \tfrac{\sigma^2 B^2}{c} + 2\sigma^2 \ln(\tfrac{1}{\delta})}. \quad (17)$$

We now proceed to upper bound the cumulative regret of Kernel AMM-UCB with $\alpha = \sigma^2/c$. Using a standard optimism argument (see e.g. Prop. 1 in Russo and Van Roy 2013 and also Lemma D.12 in App. D.3), the regret of any UCB algorithm at round t is (with high probability) upper bounded by the difference between the upper and lower confidence bounds (at x_t) that it uses. For Kernel AMM-UCB, this bound is

$$r_t \leq \tfrac{2}{\sqrt{\sigma^2/c}} \widetilde{R}_{\sigma^2/c,t-1} \rho_{\sigma^2/c,t-1}(x_t). \quad (18)$$

Because the confidence bounds used by Kernel CMM-UCB (with the same c) and Kernel DMM-UCB (with the same c and with $\alpha = \sigma^2/c \in A$) are never looser than those used by Kernel AMM-UCB, these algorithms also satisfy (18). Next, as is now somewhat common in analyses of (generalised) linear bandit algorithms (Gales et al., 2022; Lee et al., 2024; Janz et al., 2024), we split $r_{1:T}$ into $\sum_{t \in \mathcal{T}_T(\sigma^2/c)} r_t$ and $\sum_{t \notin \mathcal{T}_T(\sigma^2/c)} r_t$. Using $\sup_x |f^\star(x)| \leq C$, the first part can be bounded by $2C\tau_T(\sigma^2/c)$. Using (17), (18) and the Elliptical Potential Lemma (Lemma D.4), we bound the second part in terms of $\gamma_T(\sigma^2/c)$, which yields the regret bound in Thm 6.2 (see App. D.3 for a full proof).

Theorem 6.2. *For any covariance scale $c > 0$, with probability at least $1 - \delta$, for all $T \geq 1$, the cumulative regret of Kernel CMM-UCB, Kernel DMM-UCB (with a grid A containing $\alpha = \sigma^2/c$) and Kernel AMM-UCB (with $\alpha = \sigma^2/c$) satisfies*

$$r_{1:T} \leq 2C\tau_T(\tfrac{\sigma^2}{c}) + \sigma\sqrt{24T\gamma_T(\tfrac{\sigma^2}{c})(\gamma_T(\tfrac{\sigma^2}{c}) + \tfrac{B^2}{2c} + \ln\tfrac{1}{\delta})}.$$

Unless σ is close enough to 0, the dominant term in this regret bound is the second term. Its dependence on T and the maximum information gain matches that of the best existing regret bound for KernelUCB (Whitehouse et al., 2023), although the dependence on γ_T is sub-optimal by a factor of $\sqrt{\gamma_T}$. However, due to the lower bound in (Lattimore, 2023), this cannot be improved by using a better confidence sequence for $f^\star$.

We use bounds on $\gamma_T(\alpha)$ to derive new upper bounds $\tau_T(\alpha)$. Bounds on $\tau_T(\alpha)$ exist for the linear setting (e.g. Gales et al., 2022), where $k(x, x') = x^\top x'$, and $\gamma_T(\alpha)$ admits a relatively simple upper bound. However, for more general kernels, we must resort to more elaborate bounds on $\gamma_T(\alpha)$, which makes bounding $\tau_T(\alpha)$ rather complicated (cf. Lemma D.10 and Lemma D.11). These bounds allow us to: a) make the dependence of Thm. (6.2) on T explicit for specific kernels; b) provide sub-$\sqrt{T}$ regret bounds for the noiseless case (Sec. 6.3).

For the RBF kernel, and with $c \propto 1$, we show in Prop. D.14 that the regret bound in Thm. 6.2 is $\mathcal{O}(\sqrt{T}(\ln(T))^{d+1})$. For the Matérn kernel with smoothness parameter ν and $c \propto T^{-\frac{d}{2d+2\nu}}$, we show in Prop. D.13 that the regret bound in Thm. 6.2 is $\mathcal{O}(T^{\frac{2d+\nu}{2d+2\nu}}(\ln(T))^{\frac{2\nu}{d+2\nu}})$. The same rates are obtained in (Whitehouse et al., 2023), which are the best known rates for (vanilla) KernelUCB.

6.3 The Noiseless Setting

In a 2022 COLT open problem, Vakili (2022) discussed the lack of order-optimal regret bounds for noiseless

Table 1: Cumulative regret of our KernelUCB-style algorithms as well as AY-GP-UCB, IGP-UCB and a random baseline after $T = 1000$ rounds with $d = 3$. We show the mean $\pm$ standard deviation over 10 repetitions.

	RBF Kernel		Matérn Kernel ($\nu = 5/2$)		Matérn Kernel ($\nu = 3/2$)	
	$\ell = 0.5$	$\ell = 0.2$	$\ell = 0.5$	$\ell = 0.2$	$\ell = 0.5$	$\ell = 0.2$
Kernel DMM-UCB (Ours)	**32.2** $\pm$ 20.9	**491.4** $\pm$ 117.1	**129.5** $\pm$ 45.6	**795.1** $\pm$ 206.0	**195.6** $\pm$ 78.0	**814.1** $\pm$ 344.4
Kernel AMM-UCB (Ours)	88.8 $\pm$ 6.1	1206.2 $\pm$ 20.8	197.0 $\pm$ 24.4	1661.5 $\pm$ 90.1	316.1 $\pm$ 51.1	1741.2 $\pm$ 351.2
AY-GP-UCB	136.9 $\pm$ 12.7	1518.4 $\pm$ 38.9	331.7 $\pm$ 45.2	2382.4 $\pm$ 135.4	546.0 $\pm$ 70.0	2421.3 $\pm$ 568.5
IGP-UCB	314.1 $\pm$ 110.5	1433.0 $\pm$ 122.8	553.3 $\pm$ 67.5	1853.1 $\pm$ 105.7	655.6 $\pm$ 67.4	1707.5 $\pm$ 375.5
Random	4282.4 $\pm$ 1015.4	3872.4 $\pm$ 783.7	4264.7 $\pm$ 778.0	3677.5 $\pm$ 559.2	4175.1 $\pm$ 681.0	3442.0 $\pm$ 1080.4
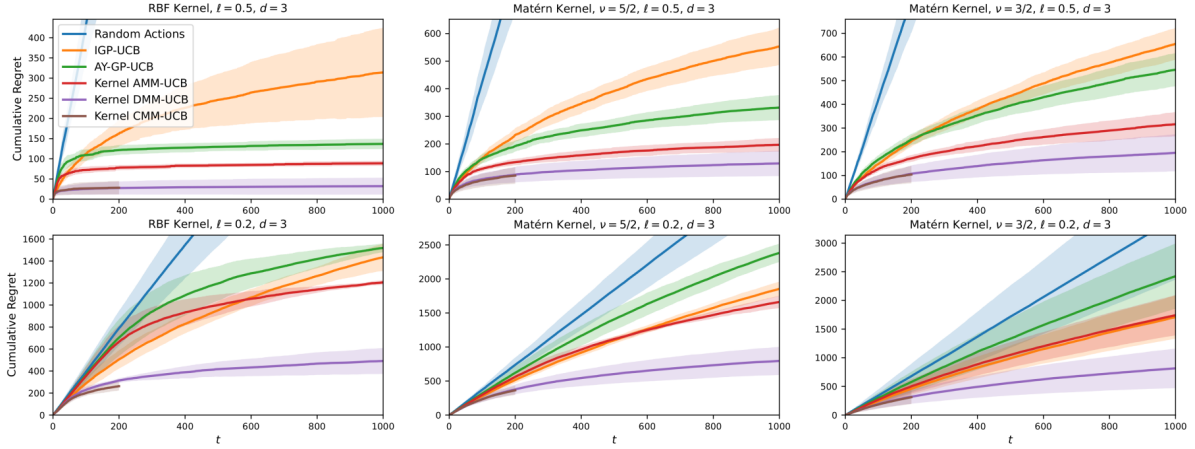
Figure 2: Cumulative regret of our KernelUCB-style algorithms as well as AY-GP-UCB, IGP-UCB and a random baseline over $T = 1000$ rounds for $d = 3$ and various kernels (columns) and length scales (rows). We show the mean $\pm$ standard deviation over 10 repetitions.

kernel bandits. Though the open problem has since been resolved (Salgia et al., 2024), we believe it is still not known whether KernelUCB can achieve sub-$\sqrt{T}$ regret bounds in the noiseless setting (see Sec. 4.1. in Vakili, 2022). We show that this is in fact the case. In the noiseless setting, the constraint $\|\boldsymbol{f}_t - \boldsymbol{y}_t\|_2 \leq R_t$ in Lem. 4.2 can be replaced with $\|\boldsymbol{f}_t - \boldsymbol{y}_t\|_2 = 0$. Since $R_t = 0$, the radius of the analytic confidence bound in Cor. 4.6 becomes

$$\widetilde{R}_{\alpha,t}^2 = \alpha B^2 - \boldsymbol{y}_t^\top (\tfrac{1}{\alpha}\boldsymbol{K}_t + \boldsymbol{I})^{-1}\boldsymbol{y}_t \leq \alpha B^2$$

Following the proof of Thm. 6.2, one can show that, for all $\alpha > 0$ and $T \geq 1$,

$$r_{1:T} \leq 2C\tau_T(\alpha) + \sqrt{12\alpha B^2 T \gamma_T(\alpha)}.$$

From here, in App. D.5, we show that for the RBF kernel, if $\alpha \propto 1/T$, then this bound becomes $r_{1:T} = \mathcal{O}(\ln(T)^{d+1})$. For the Matérn kernel, if $\alpha \propto T^{-\frac{2\nu^2 + 2\nu d}{2\nu^2 + 2\nu d + d^2}}$, then the regret bound becomes $r_{1:T} = \mathcal{O}(T^{\frac{\nu d + d^2}{2\nu^2 + 2\nu d + d^2}} \ln(T))$. Whenever $\nu > d/\sqrt{2}$, this bound grows slower that $\sqrt{T}$. We remark that our regret bounds for the noiseless setting also apply to GP-UCB with the confidence bounds in (Abbasi-Yadkori, 2012) and (Whitehouse et al., 2023).

7 EXPERIMENTS

We aim to verify that when our confidence bounds replace existing ones, KernelUCB shows better empirical performance at a similar computational cost. We compare *Kernel CMM-UCB*, *Kernel DMM-UCB*, *Kernel AMM-UCB*, *AY-GP-UCB*, *IGP-UCB* and *Random*, which chooses actions uniformly at random.

We run each algorithm in several kernel bandit problems. To generate each reward function, we sample 20 inducing points $z_1, \ldots, z_{20}$ uniformly at random from the hypercube $[0,1]^d$ and a 20-dimensional weight vector $\boldsymbol{w} \in \mathbb{R}^{20}$ from a standard normal distribution. The test function is $f^\star(x) = b\sum_{i=1}^{20} w_i k(x, z_i)$, with b being set such that $\|f^\star\|_{\mathcal{H}} = B$. The noise variables $\epsilon_1, \epsilon_2, \ldots$ are drawn independently from a normal distribution with mean 0 and standard deviation σ. In each round t, the action set $\mathcal{X}_t$ consists of 100 d-dimensional vectors drawn uniformly at random from the hypercube $[0,1]^d$. We present results for $d \in \{2, 3, 4\}$ and with the RBF and Matérn kernels with length scale $\ell \in \{0.2, 0.5\}$.

We run each algorithm for $T = 1000$ rounds with $\delta = 0.01$ and the true values $B = 10$ and $\sigma = 0.1$. We run IGP-UCB with the recommend value of $\eta = 2/T$. For each of our methods, we set $c = 1$ for the RBF kernel

and $c = T^{-\frac{d}{2d+2\nu}}$ for the Matérn kernel. Equivalently, for AY-GP-UCB, we set $\lambda = \sigma^2$ for the RBF kernel and $\lambda = \sigma^2 T^{\frac{d}{2d+2\nu}}$ for the Matérn kernel. We run Kernel AMM-UCB with $\alpha = \sigma^2/c$ and Kernel DMM-UCB with $A = \{0.1\sigma^2/c, 0.3\sigma^2/c, \sigma^2/c, 3\sigma^2/c, 10\sigma^2/c\}$. Due to prohibitive cost, we only run Kernel CMM-UCB for 200 rounds. All experiments were run on a single computer with an Intel i5-1145G7 CPU. [1]

Fig. 2 shows the cumulative regret of each algorithm with $d = 3$ (results with $d \in \{2, 4\}$ are shown in App. F.2). The cumulative regret at $T = 1000$ is displayed in Tab. 1. Our Kernel CMM-UCB algorithm consistently achieves the lowest cumulative regret over the first 200 steps, and is closely followed by Kernel DMM-UCB which achieves the lowest cumulative regret over 1000 steps. Fig. 4 (in App. F.1) shows the wall-clock time per step for each method. IGP-UCB, AY-GP-UCB and Kernel AMM-UCB take approximately the same time per step. Since we used a grid A of size 5, the time per step for Kernel DMM-UCB is a little under 5 times that of Kernel AMM-UCB. The time per step for Kernel CMM-UCB grows at the fastest rate in t. At $t = 200$, the time per step for Kernel CMM-UCB is more than 15 seconds, whereas the time per step for every other method is below 0.02 seconds.

8 CONCLUSION

In this paper, we developed new confidence bounds for sequential kernel regression. In our theoretical analysis, we proved that our new confidence bounds are always tighter than existing confidence bounds, and we showed that our KernelUCB algorithms can have sub-$\sqrt{T}$ cumulative regret in the noiseless setting. Our experimental results suggest that our Kernel DMM-UCB algorithm is the preferred choice. It closely matches the empirical performance of our effective but expensive Kernel CMM-UCB algorithm, yet it has roughly the same cost as our Kernel AMM-UCB method.

Though we applied our confidence bounds only to the kernel bandit problem, and the KernelUCB algorithm in particular, they are a generic tool that could be used anywhere that existing confidence bounds are used. For instance, one could use our confidence bounds in the π-GP-UCB algorithm (Janz et al., 2020), which would give a better rate for the Matérn kernel. In addition, one could apply our confidence bounds to kernelised reinforcement learning (Yang et al., 2020; Liu and Su, 2022; Vakili and Olkhovskaya, 2023) or adaptive control (Kakade et al., 2020).

Finally, we comment on some limitations of our work.

We assume that the kernel as well as reasonably good upper bounds on σ and B are known in advance, which is often not the case in practice. While using a loose upper bound on σ appears to have little impact, we found that using a loose bound on B can harm empirical performance (see Fig. 7). In future work, we would like to address this limitation by using model selection methods to estimate $\|f^\star\|_{\mathcal{H}}$ and choose the kernel.

Acknowledgements

References

Abbasi-Yadkori, Y. (2012). *Online learning for linearly parametrized control problems.* PhD thesis, University of Alberta.

Abbasi-Yadkori, Y., Pál, D., and Szepesvári, C. (2011). Improved algorithms for linear stochastic bandits. *Advances in neural information processing systems*, 24.

Abbasi-Yadkori, Y., Pal, D., and Szepesvari, C. (2012). Online-to-confidence-set conversions and application to sparse stochastic bandits. In *Artificial Intelligence and Statistics*, pages 1–9. PMLR.

Agrawal, A., Amos, B., Barratt, S., Boyd, S., Diamond, S., and Kolter, J. Z. (2019). Differentiable convex optimization layers. *Advances in neural information processing systems*, 32.

Auer, P. (2002). Using confidence bounds for exploitation-exploration trade-offs. *Journal of Machine Learning Research*, 3(Nov):397–422.

Belkin, M. (2018). Approximation beats concentration? an approximation view on inference with smooth radial kernels. In *Conference On Learning Theory*, pages 1348–1361. PMLR.

Berkenkamp, F., Turchetta, M., Schoellig, A., and Krause, A. (2017). Safe model-based reinforcement learning with stability guarantees. *Advances in neural information processing systems*, 30.

Blondel, M., Berthet, Q., Cuturi, M., Frostig, R., Hoyer, S., Llinares-López, F., Pedregosa, F., and Vert, J.-P. (2022). Efficient and modular implicit differentiation. *Advances in neural information processing systems*, 35:5230–5242.

Boyd, S. and Vandenberghe, L. (2004). *Convex optimization.* Cambridge university press.

Cesa-Bianchi, N. and Lugosi, G. (2006). *Prediction, learning, and games.* Cambridge university press.

---

[1]Source code is available at https://github.com/hamishflynn/KernelConfidenceBounds.

Chowdhury, S. R. and Gopalan, A. (2017). On kernelized multi-armed bandits. In *International Conference on Machine Learning*, pages 844–853. PMLR.

de la Peña, V. H., Klass, M. J., and Leung Lai, T. (2004). Self-normalized processes: exponential inequalities, moment bounds and iterated logarithm laws. *Annals of Probability*, 32:1902–1933.

de la Peña, V. H., Lai, T. L., and Shao, Q.-M. (2009). *Self-normalized processes: Limit theory and Statistical Applications*. Springer.

Durand, A., Maillard, O.-A., and Pineau, J. (2018). Streaming kernel regression with provably adaptive mean, variance, and regularization. *The Journal of Machine Learning Research*, 19(1):650–683.

Emmenegger, N., Mutnỳ, M., and Krause, A. (2023). Likelihood ratio confidence sets for sequential decision making. In *Advances in Neural Information Processing Systems (NeurIPS)*.

Flynn, H., Reeb, D., Kandemir, M., and Peters, J. (2023). Improved algorithms for stochastic linear bandits using tail bounds for martingale mixtures. In *Advances in Neural Information Processing Systems (NeurIPS)*.

Gales, S. B., Sethuraman, S., and Jun, K.-S. (2022). Norm-agnostic linear bandits. In *International Conference on Artificial Intelligence and Statistics*, pages 73–91. PMLR.

Jamieson, K., Malloy, M., Nowak, R., and Bubeck, S. (2014). lil'ucb: An optimal exploration algorithm for multi-armed bandits. In *Conference on Learning Theory*, pages 423–439. PMLR.

Janz, D., Burt, D., and González, J. (2020). Bandit optimisation of functions in the Matérn kernel RKHS. In *International Conference on Artificial Intelligence and Statistics*, pages 2486–2495. PMLR.

Janz, D., Liu, S., Ayoub, A., and Szepesvári, C. (2024). Exploration via linearly perturbed loss minimisation. In *International Conference on Artificial Intelligence and Statistics*, pages 721–729. PMLR.

Kakade, S., Krishnamurthy, A., Lowrey, K., Ohnishi, M., and Sun, W. (2020). Information theoretic regret bounds for online nonlinear control. *Advances in Neural Information Processing Systems*, 33:15312–15325.

Kimeldorf, G. and Wahba, G. (1971). Some results on Tchebycheffian spline functions. *Journal of mathematical analysis and applications*, 33(1):82–95.

Lattimore, T. (2023). A lower bound for linear and kernel regression with adaptive covariates. In *The Thirty Sixth Annual Conference on Learning Theory*, pages 2095–2113. PMLR.

Lattimore, T. and Szepesvári, C. (2020). *Bandit algorithms*. Cambridge University Press.

Lee, J., Yun, S.-Y., and Jun, K.-S. (2024). Improved regret bounds of (multinomial) logistic bandits via regret-to-confidence-set conversion. In *International Conference on Artificial Intelligence and Statistics*, pages 4474–4482. PMLR.

Li, Z. and Scarlett, J. (2022). Gaussian process bandit optimization with few batches. In *International Conference on Artificial Intelligence and Statistics*, pages 92–107. PMLR.

Liu, S. and Su, H. (2022). Provably efficient kernelized Q-learning. *arXiv preprint arXiv:2204.10349*.

Look, A., Doneva, S., Kandemir, M., Gemulla, R., and Peters, J. (2020). Differentiable implicit layers. In *NeurIPS ML for Engineering Workshop*.

Neiswanger, W. and Ramdas, A. (2021). Uncertainty quantification using martingales for misspecified Gaussian processes. In *Algorithmic Learning Theory*, pages 963–982. PMLR.

Nemirovski, A. (2004). Interior point polynomial time methods in convex programming. *Lecture notes*, 42(16):3215–3224.

Russo, D. and Van Roy, B. (2013). Eluder dimension and the sample complexity of optimistic exploration. *Advances in Neural Information Processing Systems*, 26.

Salgia, S., Vakili, S., and Zhao, Q. (2024). Random exploration in bayesian optimization: Order-optimal regret and computational efficiency. In *International Conference on Machine Learning*.

Santin, G. and Schaback, R. (2016). Approximation of eigenfunctions in kernel-based spaces. *Advances in Computational Mathematics*, 42(4):973–993.

Scharnhorst, P., Maddalena, E. T., Jiang, Y., and Jones, C. N. (2022). Robust uncertainty bounds in reproducing kernel hilbert spaces: A convex optimization approach. *IEEE Transactions on Automatic Control*, 68(5):2848–2861.

Schölkopf, B., Herbrich, R., and Smola, A. J. (2001). A generalized representer theorem. In *International conference on computational learning theory*, pages 416–426. Springer.

Srinivas, N., Krause, A., Kakade, S., and Seeger, M. (2010). Gaussian process optimization in the bandit setting: No regret and experimental design. In *Proc. International Conference on Machine Learning (ICML)*.

Sui, Y., Gotovos, A., Burdick, J., and Krause, A. (2015). Safe exploration for optimization with Gaussian processes. In *International conference on machine learning*, pages 997–1005. PMLR.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Vakili, S. (2022). Open problem: Regret bounds for noise-free kernel-based bandits. In *Conference on Learning Theory*, pages 5624–5629. PMLR.

Vakili, S., Khezeli, K., and Picheny, V. (2021). On information gain and regret bounds in Gaussian process bandits. In *International Conference on Artificial Intelligence and Statistics*, pages 82–90. PMLR.

Vakili, S. and Olkhovskaya, J. (2023). Kernelized reinforcement learning with order optimal regret bounds. In *Advances in Neural Information Processing Systems (NeurIPS)*.

Valko, M., Korda, N., Munos, R., Flaounas, I., and Cristianini, N. (2013). Finite-Time Analysis of Kernelised Contextual Bandits. In *Uncertainty in Artificial Intelligence*.

Ville, J. (1939). Etude critique de la notion de collectif. *Bull. Amer. Math. Soc*, 45(11):824.

Whitehouse, J., Wu, Z. S., and Ramdas, A. (2023). On the sublinear regret of GP-UCB. *Advances in Neural Information Processing Systems*.

Yang, Z., Jin, C., Wang, Z., Wang, M., and Jordan, M. I. (2020). On function approximation in reinforcement learning: optimism in the face of large state spaces. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, pages 13903–13916.

Zhang, F. (2006). *The Schur complement and its applications*, volume 4. Springer Science & Business Media.

Checklist

1. For all models and algorithms presented, check if you include:

   (a) A clear description of the mathematical setting, assumptions, algorithm, and/or model.
   Yes. The setting and assumptions are described in sections 2 and 6. The algorithms are described in Section 5.

   (b) An analysis of the properties and complexity (time, space, sample size) of any algorithm.
   Yes. The regret of the bandit algorithms are analysed in Section 6.

   (c) (Optional) Anonymized source code, with specification of all dependencies, including external libraries.
   Yes.

2. For any theoretical claim, check if you include:

   (a) Statements of the full set of assumptions of all theoretical results.
   Yes. At the beginning of Section 6.

   (b) Complete proofs of all theoretical results.
   Yes. We provide some proof sketches in the main text, and include full proofs in the Supplementary Material.

   (c) Clear explanations of any assumptions.
   Yes. At the beginning of Section 6.

3. For all figures and tables that present empirical results, check if you include:

   (a) The code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL).
   Yes. The setup is described in Sec. 7. Source code is available here.

   (b) All the training details (e.g., data splits, hyperparameters, how they were chosen).
   Yes. In Sec. 7.

   (c) A clear definition of the specific measure or statistics and error bars (e.g., with respect to the random seed after running experiments multiple times).
   Yes. In the captions of the tables/figures.

   (d) A description of the computing infrastructure used. (e.g., type of GPUs, internal cluster, or cloud provider).
   Yes. In Sec. 7.

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets, check if you include:

   (a) Citations of the creator If your work uses existing assets. [Yes/No/Not Applicable]
   Not Applicable.

   (b) The license information of the assets, if applicable.
   Not Applicable.

   (c) New assets either in the supplemental material or as a URL, if applicable.
   Not Applicable.

   (d) Information about consent from data providers/curators.
   Not Applicable.

   (e) Discussion of sensible content if applicable, e.g., personally identifiable information or offensive content.
   Not Applicable.

5. If you used crowdsourcing or conducted research with human subjects, check if you include:

(a) The full text of instructions given to participants and screenshots.

Not Applicable.

(b) Descriptions of potential participant risks, with links to Institutional Review Board (IRB) approvals if applicable.

Not Applicable.

(c) The estimated hourly wage paid to participants and the total amount spent on participant compensation.

Not Applicable.

A Tail Bound Derivations

A.1 Derivation of (7)

For convenience, we first re-state the setting in which the tail bound holds. $(\mathcal{D}_t | t \in \mathbb{N})$ is any filtration such that x_t is $\mathcal{D}_{t-1}$-measurable and y_t is $\mathcal{D}_t$-measurable (e.g. $\mathcal{D}_t = \sigma(x_1, y_1, \ldots, x_t, y_t, x_{t+1})$). $(Z_t(g) | t \in \mathbb{N}, g \in \mathbb{R})$ is a stochastic process, indexed by a time t and a real number g, which is adapted to a filtration $(\mathcal{D}_t | t \in \mathbb{N})$. $(g_t | t \in \mathbb{N})$ is a sequence of predictable guesses and $(\lambda_t | t \in \mathbb{N})$ is a sequence of predictable random variables. $(P_t | t \in \mathbb{N})$ is any sequence of mixture distributions that satisfies: (a) P_t is a distribution over $\boldsymbol{g}_t \in \mathbb{R}^t$; (b) P_t is $\mathcal{D}_{t-1}$-measurable; (c) $\int P_t(\boldsymbol{g}_t) dg_t = P_{t-1}(\boldsymbol{g}_{t-1})$ for all t. Define

$$M_t(\boldsymbol{g}_t, \boldsymbol{\lambda}_t) := \exp\left(\sum_{s=1}^{t} \lambda_s Z_s(g_s) - \psi_s(g_s, \lambda_s)\right), \quad \psi_t(g_t, \lambda_t) := \ln\left(\mathbb{E}\left[\exp(\lambda_t Z_t(g_t)) | \mathcal{D}_{t-1}\right]\right). \tag{19}$$

Due to Thm. 5.1 of (Flynn et al., 2023), for any $\delta \in (0, 1]$, with probability at least $1 - \delta$,

$$\forall t \geq 1: \quad \ln\left(\mathbb{E}_{\boldsymbol{g}_t \sim P_t}[M_t(\boldsymbol{g}_t, \boldsymbol{\lambda}_t)]\right) \leq \ln(1/\delta). \tag{20}$$

We choose $Z_t(g_t) = (g_t - f^\star(x_t))\epsilon_t$. Using the σ-sub-Gaussian property of the noise variables, we have

$$\psi_t(g_t, \lambda_t) = \ln\left(\mathbb{E}\left[\exp(\lambda_t(g_t - f^\star(x_t))\epsilon_t) | \mathcal{D}_{t-1}\right]\right) \leq \frac{\sigma^2 \lambda_t^2 (g_t - f^\star(x_t))^2}{2}.$$

Combining this with (20), we have

$$\ln\left(\mathbb{E}_{\boldsymbol{g}_t \sim P_t}\left[\exp\left(\sum_{s=1}^{t} \lambda(g_s - f^\star(x_s))\epsilon_s - \frac{\sigma^2 \lambda_s^2 (g_s - f^\star(x_s))^2}{2}\right)\right]\right) \tag{21}$$

$$\leq \ln\left(\mathbb{E}_{\boldsymbol{g}_t \sim P_t}\left[\exp\left(\sum_{s=1}^{t} \lambda Z_s(g_s) - \psi_s(g_s, \lambda)\right)\right]\right) = \ln\left(\mathbb{E}_{\boldsymbol{g}_t \sim P_t}[M_t(\boldsymbol{g}_t, \lambda)]\right) \leq \ln(1/\delta).$$

Next, we rearrange the integrand on the LHS of (21). For every $s \in [t]$, we have

$$(f^\star(x_s) - y_s)^2 - (g_s - y_s)^2 = (\epsilon_s)^2 - (g_s - f^\star(x_s) - \epsilon_s)^2$$
$$= (\epsilon_s)^2 - (g_s - f^\star(x_s))^2 + 2(g_s - f^\star(x_s))\epsilon_s - (\epsilon_s)^2$$
$$= 2(g_s - f^\star(x_s))\epsilon_s - (g_s - f^\star(x_s))^2.$$

This means that

$$\lambda_s(g_s - f^\star(x_s))\epsilon_s - \frac{\sigma^2 \lambda_s^2 (g_s - f^\star(x_s))^2}{2} = \frac{\lambda}{2}(f^\star(x_s) - y_s)^2 - \frac{\lambda}{2}(g_s - y_s)^2 + \frac{1}{2}(\lambda_s - \sigma^2 \lambda^2)(g_s - f^\star(x_s))^2.$$

(21) can now be re-written as

$$\ln\left(\mathbb{E}_{\boldsymbol{g}_t \sim P_t}\left[\exp\left(\sum_{s=1}^{t} \frac{\lambda}{2}(f^\star(x_s) - y_s)^2 - \frac{\lambda_s}{2}(g_s - y_s)^2 + \frac{1}{2}(\lambda_s - \sigma^2 \lambda_s^2)(g_s - f^\star(x_s))^2\right)\right]\right) \leq \ln(1/\delta). \tag{22}$$

We set $\lambda_s \equiv 1/\sigma^2$. For this choice, we have $\lambda_s - \sigma^2 \lambda_s^2 = 0$, and (22) becomes

$$\ln\left(\mathbb{E}_{\boldsymbol{g}_t \sim P_t}\left[\exp\left(\sum_{s=1}^{t} \frac{1}{2\sigma^2}(f^\star(x_s) - y_s)^2 - \frac{1}{2\sigma^2}(g_s - y_s)^2\right)\right]\right) \leq \ln(1/\delta). \tag{23}$$

By defining $\boldsymbol{f}_t^\star := [f^\star(x_1), \ldots, f^\star(x_t)]^\top$ and $\boldsymbol{y}_t := [y_1, \ldots, y_t]^\top$, rearranging (23), and then writing the resulting inequality using squared norms rather than sums of squares, we obtain

$$\|\boldsymbol{f}_t^\star - \boldsymbol{y}_t\|_2^2 \leq -2\sigma^2 \ln\left(\mathbb{E}_{\boldsymbol{g}_t \sim P_t}\left[\exp\left(-\frac{1}{2\sigma^2}\|\boldsymbol{g}_t - \boldsymbol{y}_t\|_2^2\right)\right]\right) + 2\sigma^2 \ln(1/\delta).$$

A.2 Derivation of (8)

(8) is derived from (7) using the following Lemma, which is proved in App. B.2 of (Flynn et al., 2023).

Lemma A.1 ((Flynn et al., 2023)). *For any t, any $\boldsymbol{\mu}_t$ and any positive semi-definite $\boldsymbol{T}_t$, we have*

$$\underset{\boldsymbol{g}_t \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{T}_t)}{\mathbb{E}} \left[\exp\left(-\frac{1}{2\sigma^2} \|\boldsymbol{g}_t - \boldsymbol{y}_t\|_2^2 \right) \right] = \sqrt{\frac{1}{\det(\boldsymbol{I} + \boldsymbol{T}_t/\sigma^2)}} \exp\left(-\frac{1}{2\sigma^2} (\boldsymbol{\mu}_t - \boldsymbol{y}_t)^\top \left(\boldsymbol{I} + \frac{\boldsymbol{T}_t}{\sigma^2} \right)^{-1} (\boldsymbol{\mu}_t - \boldsymbol{y}_t) \right)$$

Combining (7) and Lemma A.1, we obtain

$$\|\boldsymbol{f}_t^\star - \boldsymbol{y}_t\|_2^2 \leq -2\sigma^2 \ln \left(\underset{\boldsymbol{g}_t \sim \mathcal{N}(\boldsymbol{0}, c\boldsymbol{K}_t)}{\mathbb{E}} \left[\exp\left(-\frac{1}{2\sigma^2} \|\boldsymbol{g}_t - \boldsymbol{y}_t\|_2^2 \right) \right] \right) + 2\sigma^2 \ln(1/\delta). \tag{24}$$

$$= \boldsymbol{y}_t^\top \left(\boldsymbol{I} + \frac{c\boldsymbol{K}_t}{\sigma^2} \right)^{-1} \boldsymbol{y}_t + \sigma^2 \ln \left(\det \left(\boldsymbol{I} + \frac{c\boldsymbol{K}_t}{\sigma^2} \right) \right) + 2\sigma^2 \ln(1/\delta) = (R_t(\delta))^2.$$

B Confidence Sequence and Confidence Bound Derivations

In this section, we provide the full derivations of our confidence sequences and confidence bounds from Section 4. Before proving our main results, we state and prove some useful lemmas.

B.1 Useful Lemmas

Lemma B.1 allows us to express weighted sums of the quadratic constraints that define $\mathcal{F}_t$ (from Lemma 4.2) as a single quadratic constraint.

Lemma B.1. *For any $\alpha > 0$ and any $f \in \mathcal{H}$,*

$$\|\boldsymbol{f}_t - \boldsymbol{y}_t\|_2^2 + \alpha \|f\|_{\mathcal{H}}^2 = \left\| (\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}})^{1/2} (f - \mu_{\alpha,t}) \right\|_{\mathcal{H}}^2 + \boldsymbol{y}_t^\top \boldsymbol{y}_t - \boldsymbol{y}_t^\top \Phi_t (\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}})^{-1} \Phi_t^\top \boldsymbol{y}_t.$$

Proof. $\|\boldsymbol{f}_t - \boldsymbol{y}_t\|_2^2 + \alpha \|f\|_{\mathcal{H}}^2$ can be rewritten in the form $(f - b)^\top \boldsymbol{A} (f - b) + c$ by completing the square, where $\boldsymbol{A}$ is a self-adjoint (symmetric) linear operator, $b \in \mathcal{H}$ and $c \in \mathbb{R}$. Since $\boldsymbol{A}$ is self-adjoint, we have

$$(f - b)^\top \boldsymbol{A} (f - b) + c = f^\top \boldsymbol{A} f - 2b^\top \boldsymbol{A} f + b^\top \boldsymbol{A} b + c.$$

We also have

$$\|\boldsymbol{f}_t - \boldsymbol{y}_t\|_2^2 + \alpha \|f\|_{\mathcal{H}}^2 = (\Phi_t f - \boldsymbol{y}_t)^\top (\Phi_t f - \boldsymbol{y}_t) + \alpha f^\top f$$
$$= f^\top (\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}}) f - 2\boldsymbol{y}_t^\top \Phi_t f + \boldsymbol{y}_t^\top \boldsymbol{y}_t.$$

By equating coefficients, we have

$$\boldsymbol{A} = \boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}},$$
$$b = (\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}})^{-1} \Phi_t^\top \boldsymbol{y}_t = \Phi_t^\top (\boldsymbol{K}_t + \alpha \boldsymbol{I})^{-1} \boldsymbol{y}_t = \mu_{\alpha,t},$$
$$c = \boldsymbol{y}_t^\top \boldsymbol{y}_t - \boldsymbol{y}_t^\top \Phi_t (\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}})^{-1} \Phi_t^\top \boldsymbol{y}_t.$$

□

Lemma B.2 gives the maximum of certain concave quadratic functionals. We use it in the proof of Theorem 4.5 to calculate the dual function associated with $\max_{f \in \mathcal{F}_t} \{f(x)\}$.

Lemma B.2. *For any functions $a, b \in \mathcal{H}$ and any self-adjoint (symmetric) positive-definite linear operator $\boldsymbol{A}$, we have*

$$\max_{f \in \mathcal{H}} \left\{ a^\top f - (f - b)^\top \boldsymbol{A}(f - b) \right\} = a^\top b + \frac{1}{4} a^\top \boldsymbol{A}^{-1} a.$$

Proof. Let $F(f) = a^\top f - (f - b)^\top \boldsymbol{A}(f - b)$. Since $F(f)$ is a concave quadratic functional, $F'(\bar{f}) = 0$ is a sufficient condition for $\bar{f}$ to be a maximiser of $F(f)$, where $F'(f)$ is the Fréchet-derivative of F at f. For any direction $g \in \mathcal{H}$ and any $s > 0$, we have

$$\frac{\mathrm{d}}{\mathrm{d}s} F(f + sg) = a^\top g - 2sg^\top \boldsymbol{A}g - 2(f - b)^\top \boldsymbol{A}g.$$

Therefore, the directional derivative of F (in the direction g) is

$$\frac{\mathrm{d}}{\mathrm{d}s} F(f + sg) \bigg|_{s=0} = a^\top g - 2(f - b)^\top \boldsymbol{A}g.$$

This means that the Fréchet-derivative of F at f is

$$F'(f) = a - 2\boldsymbol{A}(f - b).$$

There is a unique solution of $F'(\bar{f}) = 0$, which is

$$\bar{f} = b + \frac{1}{2} \boldsymbol{A}a.$$

The maximum is

$$F(\bar{f}) = a^\top b + \frac{1}{4} a^\top \boldsymbol{A}^{-1} a.$$

$\square$

Lemma B.3 provides an alternative expression for $\rho_{\alpha,t}^2(x)$, which is already well-known.

Lemma B.3. *For any $\alpha > 0$,*

$$k(\cdot, x)^\top (\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}})^{-1} k(\cdot, x) = \frac{1}{\alpha} \left(k(x, x) - \boldsymbol{k}_t(x)^\top (\boldsymbol{K}_t + \alpha \boldsymbol{I})^{-1} \boldsymbol{k}_t(x) \right).$$

Proof.

$$
\begin{aligned}
k(\cdot, x)^\top (\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}})^{-1} k(\cdot, x) &= \frac{1}{\alpha} k(\cdot, x)^\top (\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}})^{-1} (\alpha \boldsymbol{I}_{\mathcal{H}}) k(\cdot, x) \\
&= \frac{1}{\alpha} k(\cdot, x)^\top (\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}})^{-1} (\Phi_t^\top \Phi_t + \alpha \boldsymbol{I}_{\mathcal{H}} - \Phi_t^\top \Phi_t) k(\cdot, x) \\
&= \frac{1}{\alpha} \left(k(\cdot, x)^\top \boldsymbol{I}_{\mathcal{H}} k(\cdot, x) - k(\cdot, x)^\top (\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}})^{-1} \Phi_t^\top \Phi_t k(\cdot, x) \right) \\
&= \frac{1}{\alpha} \left(k(\cdot, x)^\top \boldsymbol{I}_{\mathcal{H}} k(\cdot, x) - k(\cdot, x)^\top \Phi_t^\top (\boldsymbol{K}_t + \alpha \boldsymbol{I})^{-1} \Phi_t k(\cdot, x) \right) \\
&= \frac{1}{\alpha} \left(k(x, x) - \boldsymbol{k}_t(x)^\top (\boldsymbol{K}_t + \alpha \boldsymbol{I})^{-1} \boldsymbol{k}_t(x) \right).
\end{aligned}
$$

$\square$

B.2 Proof of Corollary 4.3

Proof of Corollary 4.3. From $\|f^\star\|_{\mathcal{H}} \leq B$ and (8), for any $\delta \in (0,1]$, with probability at least $1 - \delta$, $f^\star$ satisfies

$$\forall t \geq 1, \alpha \geq 0: \quad \|\boldsymbol{f}_t^\star - \boldsymbol{y}_t\|_2^2 + \alpha \|f^\star\|_{\mathcal{H}}^2 \leq R_t^2 + \alpha B^2. \tag{25}$$

Using Lemma B.1, the inequality in (25) can be rewritten as

$$\left\| (\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}})^{1/2} (f^\star - \mu_{\alpha,t}) \right\|_{\mathcal{H}}^2 \leq R_t^2 + \alpha B^2 - \boldsymbol{y}_t^\top \boldsymbol{y}_t + \boldsymbol{y}_t^\top \Phi_t (\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}})^{-1} \Phi_t^\top \boldsymbol{y}_t.$$

This means that $f^\star \in \widetilde{\mathcal{F}}_t$ for all $t \geq 1$ and $\alpha > 0$. From the definition of the confidence set $\mathcal{F}_t$ from Lemma 4.2, every $f \in \mathcal{F}_t$ must satisfy $\|\boldsymbol{f}_t - \boldsymbol{y}_t\|_2^2 + \alpha \|f\|_{\mathcal{H}}^2 \leq R_t^2 + \alpha B^2$, which means $f \in \widetilde{\mathcal{F}}_t$. Therefore, $\mathcal{F}_t \subseteq \widetilde{\mathcal{F}}_t$. $\qquad\square$

B.3 Proof of Theorem 4.4

For convenience, we first restate some of the definitions from Section 4.3. $\boldsymbol{k}_{t+1}(x) = [k(x, x_1), \ldots, k(x, x_t), k(x, x)]^\top$. $\boldsymbol{K}_{t,t+1}$ is the $t \times t + 1$ kernel matrix with i, j^{th} element equal to $k(x_i, x_j)$ if $j < t + 1$ and $k(x_i, x)$ otherwise. $\boldsymbol{L}_{t+1}$ is any matrix satisfying $\boldsymbol{L}_{t+1}^\top \boldsymbol{L}_{t+1} = \boldsymbol{K}_{t+1}$.

Proof of Theorem 4.4. The optimisation problem $\max_{f \in \mathcal{F}_t} \{f(x)\}$ can be written as

$$\max_{f \in \mathcal{H}} \ f(x) \quad \text{s.t.} \quad \|\boldsymbol{f}_t - \boldsymbol{y}_t\|_2 \leq R_t, \quad \|f\|_{\mathcal{H}} \leq B. \tag{26}$$

First, we prove a statement that resembles a representer theorem Kimeldorf and Wahba (1971); Schölkopf et al. (2001). We will show that the solution of (26) must be of the form $f(x) = \boldsymbol{k}_{t+1}(x)^\top \boldsymbol{w}$, for some weight vector $\boldsymbol{w} \in \mathbb{R}^{t+1}$.

All functions $f \in \mathcal{H}$ can be written in the form $f = f_\| + f_\perp$, where $f_\| = \boldsymbol{k}_{t+1}(\cdot)^\top \boldsymbol{w}$ is in the subspace of $\mathcal{H}$ spanned by $k(\cdot, x_1), \ldots, k(\cdot, x_t), k(\cdot, x)$, and $f_\perp$ is orthogonal to this subspace. Since $f_\perp$ is orthogonal to the basis $k(\cdot, x_1), \ldots, k(\cdot, x_t), k(\cdot, x)$, we have

$$\|f\|_{\mathcal{H}}^2 = \langle f_\| + f_\perp, f_\| + f_\perp \rangle_{\mathcal{H}} = \langle f_\|, f_\| \rangle_{\mathcal{H}} + 2 \langle f_\|, f_\perp \rangle_{\mathcal{H}} + \langle f_\perp, f_\perp \rangle_{\mathcal{H}} = \|f_\|\|_{\mathcal{H}}^2 + \|f_\perp\|_{\mathcal{H}}^2,$$

This means that the RKHS norm of f is minimised w.r.t. $f_\perp$ by choosing $f_\perp \equiv 0$. Using the reproducing property of the kernel and the fact that $f_\perp$ is orthogonal to $k(\cdot, x_1), \ldots, k(\cdot, x_t), k(\cdot, x)$, for any $z \in \{x_1, \ldots, x_t, x\}$, we have

$$f(z) = \langle f, k(\cdot, z) \rangle_{\mathcal{H}} = \langle f_\| + f_\perp, k(\cdot, z) \rangle_{\mathcal{H}} = \langle f_\|, k(\cdot, z) \rangle_{\mathcal{H}} = f_\perp(z).$$

This means that the function values $f(x_1), \ldots, f(x_t), f(x)$ are entirely determined by $f_\|$. The objective function is $f(x)$ and the constraint $\|\boldsymbol{f}_t - \boldsymbol{y}_t\|_2 \leq R_t$ depends only on $f(x_1), \ldots, f(x_t)$. Therefore, any $f_\perp \neq 0$ will have no effect on the objective or this constraint, and will only increase the norm of f. This means that we must have $f_\perp \equiv 0$ at the solution of $\max_{f \in \mathcal{F}_t} \{f(x)\}$. We are now free to look for a solution of the form $f(x) = \boldsymbol{k}_{t+1}(x)^\top \boldsymbol{w}$, which means

$$\|\boldsymbol{f}_t - \boldsymbol{y}_t\|_2 = \|\boldsymbol{K}_{t,t+1}\boldsymbol{w} - \boldsymbol{y}_t\|_2,$$

$$\|f\|_{\mathcal{H}} = \sqrt{\left\langle \sum_{s=1}^{t} k(\cdot, x_s) w_s + k(\cdot, x) w_{t+1}, \sum_{s'=1}^{t} k(\cdot, x_{s'}) w_{s'} + k(\cdot, x) w_{t+1} \right\rangle_{\mathcal{H}}}$$

$$= \sqrt{\boldsymbol{w}^\top \boldsymbol{K}_{t+1} \boldsymbol{w}} = \|\boldsymbol{L}_{t+1}\boldsymbol{w}\|_2.$$

Substituting these expressions for the objective and constraints into (26) yields (10). $\qquad\square$

B.4 Proof of Theorem 4.5

The optimisation problem $\max_{f \in \mathcal{F}_t}\{f(x)\}$ can be written as

$$\max_{f \in \mathcal{H}} f(x) \quad \text{s.t.} \quad \|\boldsymbol{f}_t - \boldsymbol{y}_t\|_2^2 \leq R_t^2, \quad \|f\|_{\mathcal{H}}^2 \leq B^2. \tag{27}$$

The Lagrangian for this problem is

$$L(f, \eta_1, \eta_2) = f(x) + \eta_1(R_t^2 - \|\boldsymbol{f}_t - \boldsymbol{y}_t\|_2^2) + \eta_2(B^2 - \|f\|_{\mathcal{H}}^2),$$

where η_1, η_2 are the Lagrange multipliers. The dual function is

$$g(\eta_1, \eta_2) = \max_{f \in \mathcal{H}}\{L(f, \eta_1, \eta_2)\}.$$

By weak duality, the solution of (27) is upper bounded by the solution of the dual problem, which is

$$\min_{\eta_1 \geq 0, \eta_2 \geq 0}\{g(\eta_1, \eta_2)\}.$$

Since the Lagrangian $L(f, \eta_1, \eta_2)$ is linear in the Lagrange multipliers for every $f \in \mathcal{H}$, the dual function is the maximum over a set of linear functions, which is a convex function (see (3.7) in Sec. 3.2.3 of (Boyd and Vandenberghe, 2004)). This means that the dual problem is a convex program. We will now show that, for any $\eta_1 > 0$ and $\alpha > 0$, there is a closed-form expression for the dual function $g(\eta_1, \alpha\eta_1)$. Using Lemma B.1, the Lagrangian evaluated at the Lagrange multipliers η_1 and $\alpha\eta_1$ is

$$L(f, \eta_1, \alpha\eta_1) = f(x) + \eta_1(R_t^2 - \|\boldsymbol{f}_t - \boldsymbol{y}_t\|_2^2 + \alpha B^2 - \alpha \|f\|_{\mathcal{H}}^2)$$
$$= f(x) + \eta_1\left(\widetilde{R}_{\alpha,t}^2 - \left\|(\boldsymbol{V}_t + \alpha\boldsymbol{I}_{\mathcal{H}})^{1/2}(f - \mu_{\alpha,t})\right\|_{\mathcal{H}}^2\right).$$

Using Lemma B.2 and Lemma B.3, the dual function evaluated at η_1 and $\alpha\eta_1$ is

$$g(\eta_1, \alpha\eta_1) = \max_{f \in \mathcal{H}}\{L(f, \eta_1, \alpha\eta_1)\}$$
$$= \max_{f \in \mathcal{H}}\left\{k(\cdot, x)^\top f - \eta_1(f - \mu_{\alpha,t})^\top(\boldsymbol{V}_t + \alpha\boldsymbol{I}_{\mathcal{H}})(f - \mu_{\alpha,t})\right\} + \eta_1\widetilde{R}_{\alpha,t}^2$$
$$= k(\cdot, x)^\top \mu_{\alpha,t} + \frac{1}{4\eta_1}k(\cdot, x)^\top(\boldsymbol{V}_t + \alpha\boldsymbol{I}_{\mathcal{H}})^{-1}k(\cdot, x) + \eta_1\widetilde{R}_{\alpha,t}^2$$
$$= \boldsymbol{k}_t(x)^\top(\boldsymbol{K}_t + \alpha\boldsymbol{I})^{-1}\boldsymbol{y}_t + \frac{k(x, x) - \boldsymbol{k}_t(x)^\top(\boldsymbol{K}_t + \alpha\boldsymbol{I})^{-1}\boldsymbol{k}_t(x)}{4\eta_1\alpha} + \eta_1\widetilde{R}_{\alpha,t}^2.$$

Therefore, the dual problem is

$$\min_{\eta_1 > 0, \alpha > 0}\left\{\boldsymbol{k}_t(x)^\top(\boldsymbol{K}_t + \alpha\boldsymbol{I})^{-1}\boldsymbol{y}_t + \frac{k(x, x) - \boldsymbol{k}_t(x)^\top(\boldsymbol{K}_t + \alpha\boldsymbol{I})^{-1}\boldsymbol{k}_t(x)}{4\eta_1\alpha} + \eta_1\widetilde{R}_{\alpha,t}^2\right\}.$$

Using the substitution $\eta_2 = \alpha\eta_1$, we arrive at the dual problem stated in (11). By minimising this objective function w.r.t. η_1 for any fixed α, we can also express the dual problem as

$$\min_{\alpha > 0}\left\{\boldsymbol{k}_t(x)^\top(\boldsymbol{K}_t + \alpha\boldsymbol{I})^{-1}\boldsymbol{y}_t + \frac{\widetilde{R}_{\alpha,t}}{\sqrt{\alpha}}\sqrt{k(x, x) - \boldsymbol{k}_t(x)^\top(\boldsymbol{K}_t + \alpha\boldsymbol{I})^{-1}\boldsymbol{k}_t(x)}\right\}.$$

C Confidence Bound Tightness

In this section, we prove Theorem 6.1. First, we state and prove some useful lemmas.

C.1 Useful Lemmas

Lemma C.1 gives us alternative expressions for the quadratic terms in $\widetilde{R}_{\alpha,t}$, which we use to prove Lemma C.2.

Lemma C.1. *For any $\alpha > 0$, we have*

$$\boldsymbol{y}_t^\top \boldsymbol{y}_t - \boldsymbol{y}_t^\top \Phi_t \left(\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}}\right)^{-1} \Phi_t^\top \boldsymbol{y}_t = \boldsymbol{y}_t^\top \left(\frac{1}{\alpha}\boldsymbol{K}_t + \boldsymbol{I}\right)^{-1} \boldsymbol{y}_t.$$

Proof. We start with the identity

$$\Phi_t \left(\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}}\right) = \left(\boldsymbol{K}_t + \alpha \boldsymbol{I}\right) \Phi_t.$$

By post-multiplying both sides with $(\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}})^{-1}$ and pre-multiplying both sides with $(\boldsymbol{K}_t + \alpha \boldsymbol{I})^{-1}$, we obtain

$$\left(\boldsymbol{K}_t + \alpha \boldsymbol{I}\right)^{-1} \Phi_t = \Phi_t \left(\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}}\right)^{-1}. \tag{28}$$

Now, using (28), we have

$$
\begin{aligned}
\boldsymbol{y}_t^\top \boldsymbol{y}_t - \boldsymbol{y}_t^\top \Phi_t \left(\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}}\right)^{-1} \Phi_t^\top \boldsymbol{y}_t &= \boldsymbol{y}_t^\top \boldsymbol{y}_t - \boldsymbol{y}_t^\top \left(\boldsymbol{K}_t + \alpha \boldsymbol{I}\right)^{-1} \boldsymbol{K}_t \boldsymbol{y}_t \\
&= \boldsymbol{y}_t^\top \boldsymbol{y}_t - \boldsymbol{y}_t^\top \left(\boldsymbol{K}_t + \alpha \boldsymbol{I}\right)^{-1} \left(\boldsymbol{K}_t + \alpha \boldsymbol{I} - \alpha \boldsymbol{I}\right) \boldsymbol{y}_t \\
&= \boldsymbol{y}_t^\top \boldsymbol{y}_t - \boldsymbol{y}_t^\top \boldsymbol{y}_t + \alpha \boldsymbol{y}_t^\top \left(\boldsymbol{K}_t + \alpha \boldsymbol{I}\right)^{-1} \boldsymbol{y}_t \\
&= \boldsymbol{y}_t^\top \left(\frac{1}{\alpha}\boldsymbol{K}_t + \boldsymbol{I}\right)^{-1} \boldsymbol{y}_t.
\end{aligned}
$$

$\square$

Lemma C.2 gives us a simplified expression for $\widetilde{R}_{\alpha,t}$ when $\alpha = \sigma^2/c$. This allows us to compare $\widetilde{R}_{\alpha,t}$ with the equivalent radius quantities of the confidence bounds in (Abbasi-Yadkori, 2012) and (Chowdhury and Gopalan, 2017).

Lemma C.2. *For any $c > 0$, set $\alpha = \sigma^2/c$. We have*

$$\widetilde{R}_{\alpha,t} = \sqrt{\sigma^2 \ln\left(\det\left(\frac{c}{\sigma^2}\boldsymbol{K}_t + \boldsymbol{I}\right)\right) + \frac{\sigma^2 B^2}{c} + 2\sigma^2 \ln(\frac{1}{\delta})}.$$

Proof. From Lemma C.1, with this c and α, the squared radius is

$$
\begin{aligned}
\widetilde{R}_{\alpha,t}^2 &= \boldsymbol{y}_t^\top \left(\frac{c}{\sigma^2}\boldsymbol{K}_t + \boldsymbol{I}\right)^{-1} \boldsymbol{y}_t - \boldsymbol{y}_t^\top \boldsymbol{y}_t + \boldsymbol{y}_t^\top \Phi_t \left(\boldsymbol{V}_t + \frac{\sigma^2}{c}\boldsymbol{I}_{\mathcal{H}}\right)^{-1} \Phi_t^\top \boldsymbol{y}_t \\
&\quad + \sigma^2 \ln\left(\det\left(\frac{c}{\sigma^2}\boldsymbol{K}_t + \boldsymbol{I}\right)\right) + 2\sigma^2 \ln(1/\delta) + \frac{\sigma^2 B^2}{c} \\
&= \sigma^2 \ln\left(\det\left(\frac{c}{\sigma^2}\boldsymbol{K}_t + \boldsymbol{I}\right)\right) + 2\sigma^2 \ln(1/\delta) + \frac{\sigma^2 B^2}{c}.
\end{aligned}
\tag{29}
$$

$\square$

C.2 Proof of Theorem 6.1

Proof of Theorem 6.1. For any $\delta \in (0, 1]$ and any $\lambda > 0$, the UCB from Thm. 3.11 of (Abbasi-Yadkori, 2012) states that, with probability at least $1 - \delta$,

$$\forall t \geq 1: \quad f^\star(x) \leq \mu_{\lambda,t}(x) + \frac{\widetilde{R}_{\lambda,t}^{\mathrm{AY}}}{\sqrt{\lambda}} \rho_{\lambda,t}(x),$$

where

$$\widetilde{R}_{\lambda,t}^{\mathrm{AY}} = \sigma\sqrt{\ln(\det(\frac{1}{\lambda}\boldsymbol{K}_t + \boldsymbol{I})) + 2\ln(1/\delta)} + \sqrt{\lambda}B.$$

We will now show that when we set the covariance scale to $c = \sigma^2/\lambda$ and set $\alpha = \sigma^2/c = \lambda$, the RHS of (14) is strictly less than this UCB. Starting from the RHS of (14) with $c = \sigma^2/\lambda$ and $\alpha = \sigma^2/c = \lambda$, and then using Lemma C.2 and the inequality $\sqrt{a+b} < \sqrt{a} + \sqrt{b}$ (for $a, b > 0$), we have

$$
\begin{aligned}
\mu_{\lambda,t}(x) + \frac{\widetilde{R}_{\lambda,t}}{\sqrt{\lambda}}\rho_{\lambda,t}(x) &= \mu_{\lambda,t}(x) + \frac{\sqrt{\sigma^2 \ln\left(\det\left(\frac{1}{\lambda}\boldsymbol{K}_t + \boldsymbol{I}\right)\right) + 2\sigma^2 \ln(1/\delta) + \lambda B^2}}{\sqrt{\lambda}}\rho_{\lambda,t}(x) \\
&< \mu_{\lambda,t}(x) + \frac{\sigma\sqrt{\ln\left(\det\left(\frac{1}{\lambda}\boldsymbol{K}_t + \boldsymbol{I}\right)\right) + 2\ln(1/\delta)} + \sqrt{\lambda}B}{\sqrt{\lambda}}\rho_{\lambda,t}(x) \\
&= \mu_{\lambda,t}(x) + \frac{\widetilde{R}_{\lambda,t}^{\mathrm{AY}}}{\sqrt{\lambda}}\rho_{\lambda,t}(x).
\end{aligned}
$$

For any $\delta \in (0, 1]$ and any $\eta > 0$, the UCB from Thm. 2 of (Chowdhury and Gopalan, 2017) states that, with probability at least $1 - \delta$,

$$
\forall t \geq 1: \quad f^\star(x) \leq \mu_{1+\eta,t}(x) + \widetilde{R}_{\eta,t}^{\mathrm{IGP}}\rho_{1+\eta,t}(x),
$$

where

$$
\widetilde{R}_{\eta,t}^{\mathrm{IGP}} = \sigma\sqrt{\ln(\det(\tfrac{1}{1+\eta}\boldsymbol{K}_t + \boldsymbol{I})) + t\eta + 2\ln(1/\delta)} + B.
$$

If the number of rounds T is known in advance, then one can set η to the recommended value of $2/T$, and then the $t\eta$ term in $\widetilde{R}_{\eta,t}^{\mathrm{IGP}}$ is bounded by 2 for all $t \in [T]$. We compare confidence bounds for a general value of $\eta > 0$. Starting from the RHS of (14) with $c = \sigma^2/(1+\eta)$ and $\alpha = \sigma^2/c = (1+\eta)$, we have

$$
\begin{aligned}
\mu_{1+\eta,t}(x) + \frac{\widetilde{R}_{1+\eta,t}}{\sqrt{1+\eta}}\rho_{1+\eta,t}(x) &= \mu_{1+\eta,t}(x) + \frac{\sqrt{\sigma^2 \ln\left(\det\left(\frac{1}{1+\eta}\boldsymbol{K}_t + \boldsymbol{I}\right)\right) + 2\sigma^2 \ln(1/\delta) + (1+\eta)B^2}}{\sqrt{1+\eta}}\rho_{1+\eta,t}(x) \\
&< \mu_{1+\eta,t}(x) + \left(\sigma\sqrt{\ln\left(\det\left(\frac{1}{1+\eta}\boldsymbol{K}_t + \boldsymbol{I}\right)\right) + 2\ln(1/\delta)} + B\right)\rho_{1+\eta,t}(x) \\
&< \mu_{1+\eta,t}(x) + \widetilde{R}_{\eta,t}^{\mathrm{IGP}}\rho_{1+\eta,t}(x).
\end{aligned}
$$

$\square$

C.3 Why Are Our Confidence Bounds Tighter?

Here, we explain why our analytic confidence bounds in Cor. 4.6 are tighter than the confidence bounds in Thm. 3.11 of (Abbasi-Yadkori, 2012). The same reasoning explains why our analytic confidence bounds are tighter than those from Thm. 2 of (Chowdhury and Gopalan, 2017). Both our analytic confidence bounds and the confidence bounds in (Abbasi-Yadkori, 2012) are derived from a constraint of the form

$$
\left\|(\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}})^{1/2}(f^\star - \mu_{\alpha,t})\right\|_{\mathcal{H}} \leq R.
$$

The difference between the confidence bounds is in the radius quantity R. The confidence bounds in (Abbasi-Yadkori, 2012) use a self-normalised concentration inequality (Thm. 3.4 in Abbasi-Yadkori, 2012), which states that, with high probability,

$$
\left\|(\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}})^{-1/2}\Phi_t\boldsymbol{\epsilon}_t\right\|_{\mathcal{H}} \leq \sigma\sqrt{\ln\left(\det\left(\frac{1}{\alpha}\boldsymbol{K}_t + \boldsymbol{I}\right)\right) + 2\ln(1/\delta)}.
$$

Using the definition of $\mu_{\alpha,t}$ and then the triangle inequality, it is shown in (Abbasi-Yadkori, 2012) that

$$
\begin{aligned}
\left\| (\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}})^{1/2} (f^\star - \mu_{\alpha,t}) \right\|_{\mathcal{H}} &= \left\| (\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}})^{-1/2} (\Phi_t \boldsymbol{\epsilon}_t - \alpha f^\star) \right\|_{\mathcal{H}} \\
&\leq \left\| (\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}})^{-1/2} \Phi_t \boldsymbol{\epsilon}_t \right\|_{\mathcal{H}} + \alpha \left\| (\boldsymbol{V}_t + \alpha \boldsymbol{I}_{\mathcal{H}})^{-1/2} f^\star \right\|_{\mathcal{H}} \\
&\leq \sigma \sqrt{\ln \left(\det \left(\frac{1}{\alpha} \boldsymbol{K}_t + \boldsymbol{I} \right) \right) + 2 \ln(1/\delta)} + \sqrt{\alpha} B.
\end{aligned}
$$

This gives us the expression for the radius in (Abbasi-Yadkori, 2012). The triangle inequality step causes the log determinant and the RKHS norm to appear under separate square roots. Instead of the self-normalised concentration inequality, we use a bound on $\| \boldsymbol{f}_t^\star - \boldsymbol{y}_t \|_2^2$ ((8)), which can be combined with the constraint $\| f^\star \|_{\mathcal{H}} \leq B$ more directly via completing the square (Cor. 4.3) or Lagrangian duality.

D   Regret Analysis

In this section, we prove the cumulative regret bound in Theorem 6.2. In addition, we prove cumulative regret bounds with explicit dependence on T for the noisy (Sec. D.4) and noiseless (Sec. D.5) settings. We use upper bounds on $\gamma_T(\sigma^2/c)$ from (Vakili et al., 2021) and derive our own upper bounds on $\tau_T(\sigma^2/c)$. The bounds on both of these quantities apply when the kernel k is a Mercer kernel, and their rates depend on how quickly the eigenvalues of the kernel decay to zero. We therefore begin by describing Mercer kernels, as well as the polynomial and exponential eigendecay conditions.

D.1   Mercer Kernels and Eigendecay Conditions

A positive-definite kernel k on a set $\mathcal{X}$ is called a Mercer kernel if: a) $\mathcal{X}$ is a compact metric space (e.g. a compact subset of $\mathbb{R}^d$); b) the function $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is continuous. Mercer's theorem provides a useful representation for Mercer kernels. Let ρ be a non-degenerate measure on $\mathcal{X}$ and let $L^2(\mathcal{X}, \rho)$ denote the set of square integrable functions on $\mathcal{X}$, i.e.

$$
L^2(\mathcal{X}, \rho) := \left\{ f : \mathcal{X} \to \mathbb{R} : \int_{\mathcal{X}} (f(x))^2 \mathrm{d}\rho(x) < \infty \right\}.
$$

Define the linear operator $L_k : L^2(\mathcal{X}, \rho) \to L^2(\mathcal{X}, \rho)$ as

$$
L_k(f)(x) := \int_{\mathcal{X}} k(x, y) f(y) \mathrm{d}\rho(y).
$$

Theorem D.1 (Mercer's Theorem). *If $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is a Mercer kernel, then there exist non-negative eigenvalues $\xi_1 \geq \xi_2 \geq \cdots \geq 0$ and corresponding eigenfunctions $\phi_1, \phi_2, \ldots$, such that*

$$
L_k(\phi_m) = \xi_m \phi_m, \quad \text{for all } m = 1, 2, \ldots. \tag{30}
$$

In addition, the kernel function has the eigendecomposition

$$
k(x, y) = \sum_{m=1}^{\infty} \xi_m \phi_m(x) \phi_m(y), \tag{31}
$$

where the convergence of the infinite series is absolute for each $x, y \in \mathcal{X}$ and uniform on $\mathcal{X} \times \mathcal{X}$.

The scaled eigenfunctions $\sqrt{\xi_1} \phi_1, \sqrt{\xi_2} \phi_2, \ldots$ form a (countable) orthonormal basis of the RKHS $\mathcal{H}$. Following Vakili et al. (2021), we make the following additional assumption about the kernel k: $\forall m \in \mathbb{N}, \forall x \in \mathcal{X}, |\phi_m(x)| \leq \psi$, for some $\psi > 0$. The rate at which the eigenvalues $\xi_1, \xi_2, \ldots$ of the kernel decay to zero determines both the complexity of the corresponding RKHS and the hardness of regret minmisation in kernel bandits. The two most commonly studied eigendecay conditions are defined as follows.

Definition D.2 (Eigenvalue decay)**.** Consider the eigenvalues $\xi_1, \xi_2, \ldots$ of a Mercer kernel $k(x, x') = \sum_{i=1}^{\infty} \xi_i \phi_i(x) \phi_i(x')$. The polynomial eigenvalue decay condition is satisfied if there exist $C_p > 0$ and $\beta_p > 1$ such that

$$\xi_i \leq C_p i^{-\beta_p}.$$

The exponential eigenvalue decay condition is satisfied if there exist $C_{e_1}, C_{e_2} > 0$ and $\beta_e \in (0, 1]$ such that

$$\xi_i \leq C_{e_1} \exp(-C_{e_2} i^{\beta_e}).$$

If a kernel is known to satisfy one of these eigendecay conditions, then the maximum information gain can be bounded. In particular, Vakili et al. (2021) proved the following bounds.

Lemma D.3 (Corollary 1 in (Vakili et al., 2021))**.** *If the kernel k satisfies the polynomial eigenvalue decay condition, then*

$$\sup_{x_1, \ldots, x_T} \left\{ \frac{1}{2} \ln \det \left(\frac{1}{\alpha} \boldsymbol{K}_T + \boldsymbol{I} \right) \right\} \leq \ln(1 + \kappa T/\alpha) + \left(\frac{C_p \psi^2 T}{\alpha} \right)^{1/\beta_p} \ln(1 + \kappa T/\alpha)^{1 - 1/\beta_p}. \tag{32}$$

If the kernel k satisfies the exponential eigenvalue decay condition, then

$$\sup_{x_1, \ldots, x_T} \left\{ \frac{1}{2} \ln \det \left(\frac{1}{\alpha} \boldsymbol{K}_T + \boldsymbol{I} \right) \right\} \leq \left(\left(\frac{2}{C_{e_2}} (\ln(T) + C_\beta) \right)^{\frac{1}{\beta_e}} + 1 \right) \ln(1 + \kappa T/\alpha), \tag{33}$$

where $C_\beta = \ln(\frac{C_{e_1} \psi^2}{\alpha C_{e_2}})$ if $\beta_e = 1$, and $C_\beta = \ln(\frac{2 C_{e_1} \psi^2}{\alpha \beta_e C_{e_2}}) + (1/\beta_e - 1)(\ln(\frac{2}{C_{e_2}}(1/\beta_e - 1)) - 1)$ if $\beta_e < 1$.

D.2 Elliptical Potential Lemmas

In this subsection, we recall the well-known Elliptical Potential Lemma (see e.g. Lemma 11.11 in Cesa-Bianchi and Lugosi 2006, Lemma 5.3 in Srinivas et al. 2010 or Lemma 11 in Abbasi-Yadkori et al. 2011), which states that the sum of the elliptical potentials $\sum_{t=1}^{T} \frac{1}{\alpha} \rho_{\alpha, t-1}^2(x_t)$ can be upper bounded by the maximum information gain. Then, for each eigendecay condition, we state and prove our bounds on the elliptical potential count $\tau_T(\alpha)$.

The version of the Elliptical Potential Lemma stated in Lemma D.4 is essentially the same as the version in Lemma 5 in (Whitehouse et al., 2023), so we omit the proof. The only real difference is that we start with $\sum_{t=1}^{T} \min \left(1, \frac{1}{\alpha} \rho_{\alpha, t-1}^2(x_t) \right)$ instead of $\sum_{t=1}^{T} \frac{1}{\alpha} \rho_{\alpha, t-1}^2(x_t)$. This change means that Lemma D.4 holds for all $\alpha > 0$, rather than for $\alpha \geq \max(1, \kappa)$, where κ is an upper bound on $\sup_{x, x'} |k(x, x')|$.

Lemma D.4 (Elliptical Potential Lemma)**.** *For any $T \geq 1$ and any $\alpha > 0$,*

$$\sum_{t=1}^{T} \min \left(1, \frac{1}{\alpha} \rho_{\alpha, t-1}^2(x_t) \right) \leq \frac{3}{2} \ln \left(\det \left(\frac{1}{\alpha} \boldsymbol{K}_t + \boldsymbol{I} \right) \right).$$

Now, we turn our attention to the elliptical potential count $\tau_T(\alpha)$. In the linear/finite-dimensional setting, there exist bounds on $\tau_T(\alpha)$. Lattimore and Szepesvári (2020) (Exercise 19.3) and Gales et al. (2022) (Lemma 7), show that if the kernel is $k(x, x') = x^\top x'$, then $\tau_T(\alpha) = \mathcal{O}(d \ln(1/\alpha))$. We follow (as closely as we can) the proof used in (Gales et al., 2022), which uses the following lemma.

Lemma D.5 (Lemma 8 from (Gales et al., 2022))**.** *Let $X, A, B \geq 0$. If $X \leq A \ln(1 + BX)$, then*

$$X \leq \inf_{\eta \in (0, 1)} \left\{ \frac{A}{1 - \eta} \ln \left(\frac{A}{e\eta} \left(\frac{1}{X} + B \right) \right) \right\}.$$

By choosing $\eta = 1/e$, we get the following corollary, which comes from the proof of Lemma 7 in (Gales et al., 2022).

Corollary D.6. *Let* $X, A, B \geq 0$. *If* $X \leq A \ln(1 + BX)$, *then*

$$X \leq \frac{eA}{e-1} \ln(1 + AB).$$

Proof. From Lemma D.5 with $\eta = 1/e$, we have

$$X \leq \frac{eA}{e-1} \ln\left(A\left(\frac{1}{X} + B\right)\right).$$

On the one hand, if $1/X \leq 1/A$, then

$$X \leq \frac{eA}{e-1} \ln(1 + AB).$$

On the other hand, if $1/X \geq 1/A$, then the inequality $X \leq A \ln(1 + BX)$ tells us that

$$X \leq A \ln(1 + AB).$$

$\square$

We also use the following variant of Lemma D.5. This inequality is stated just after Lemma 8 (Gales et al., 2022). For completeness, we give a proof.

Lemma D.7. *Let* $X, A, B \geq 0$. *If* $X \leq A \ln(BX)$, *then*

$$X \leq \inf_{\eta \in (0,1)} \left\{ \frac{1}{1-\eta}\left(A \ln\left(\frac{AB}{e\eta}\right)\right) \right\}.$$

Proof.

$$
\begin{aligned}
X &\leq A \ln(BX) \\
&= A \ln\left(\frac{\eta X}{A}\right) + A \ln\left(\frac{AB}{\eta}\right) \\
&\leq \eta X - A \ln(e) + A \ln\left(\frac{AB}{\eta}\right) \\
&= \eta X + A \ln\left(\frac{AB}{e\eta}\right).
\end{aligned}
$$

$\square$

Finally, we require the following lemma, which gives an expression for $\det\left(\frac{1}{\alpha} \boldsymbol{K}_T + \boldsymbol{I}\right)$ whenever the RKHS $\mathcal{H}$ is separable.

Lemma D.8 (Lemma E.1 in (Abbasi-Yadkori, 2012)). *If* $\mathcal{H}$ *is separable, then*

$$\det\left(\frac{1}{\alpha} \boldsymbol{K}_T + \boldsymbol{I}\right) = \prod_{t=1}^{T}\left(1 + \|k(\cdot, x_t)\|^2_{(\boldsymbol{V}_{t-1} + \alpha \boldsymbol{I}_{\mathcal{H}})^{-1}}\right).$$

Since a Hilbert space is separable if and only if it has a countable orthonormal basis, the RKHS of any Mercer kernel is separable. The next lemma gives an intermediate result that can be specialised to either of the eigendecay conditions.

Lemma D.9 (Elliptical potentials: You cannot have more than $\mathcal{O}(\gamma_T)$ big intervals). *Let $b > 0$, let $x_1, \ldots, x_T$ be any sequence of elements in $\mathcal{X}$ and let $k(x, x')$ be any positive-definite kernel. Let $\mathcal{T}_T(\alpha) = \{t \in [T] : \|k(\cdot, x_t)\|_{(\mathbf{V}_{t-1} + \alpha \mathbf{I}_\mathcal{H})^{-1}} \geq b\}$. For every $T \geq 1$,*

$$|\mathcal{T}_T(\alpha)| \leq \frac{2\gamma_{|\mathcal{T}_T(\alpha)|}(\alpha)}{\ln(1 + b^2)} \leq \frac{2\gamma_T(\alpha)}{\ln(1 + b^2)}. \tag{34}$$

Using Lemma B.3, we have $\|k(\cdot, x_t)\|_{(\mathbf{V}_{t-1} + \alpha \mathbf{I}_\mathcal{H})^{-1}} = \frac{1}{\sqrt{\alpha}} \rho_{\alpha, t-1}(x_t)$, which means the set $\mathcal{T}_T(\alpha)$ defined here is the same as the one in (16), except here we allow b to be any positive number, not just 1.

Proof. Define $\mathbf{G}_T = \sum_{t=1}^{T} \mathbb{I}\{t \in \mathcal{T}_T(\alpha)\} k(\cdot, x_t) k(\cdot, x_t)^\top$ and let $\mathbf{C}_T$ be the $|\mathcal{T}_T(\alpha)| \times |\mathcal{T}_T(\alpha)|$ kernel matrix constructed from the points $\{x_t\}_{t \in \mathcal{T}_T(\alpha)}$. Using Lemma D.8, we have

$$\det\left(\frac{1}{\alpha} \mathbf{C}_T + \mathbf{I}\right) = \prod_{t \in \mathcal{T}_T(\alpha)} (1 + \|k(\cdot, x_t)\|^2_{(\mathbf{G}_{t-1} + \alpha \mathbf{I}_\mathcal{H})^{-1}})$$

$$\geq \prod_{t \in \mathcal{T}_T(\alpha)} (1 + \|k(\cdot, x_t)\|^2_{(\mathbf{V}_{t-1} + \alpha \mathbf{I}_\mathcal{H})^{-1}})$$

$$\geq (1 + b^2)^{|\mathcal{T}_T(\alpha)|}.$$

Taking logarithms of both sides finishes the proof. $\qquad \square$

Relying on the second inequality in (34) can give a rather loose dependence on T. From the first inequality in (34), we can derive tighter bounds on $|\mathcal{T}_T(\alpha)|$ when the kernel satisfies either the polynomial or exponential eigendecay conditions. First, we consider polynomial eigendecay.

Lemma D.10 (Elliptical potentials: polynomial eigendecay). *Let $b > 0$ and let $x_1, \ldots, x_T$ be any sequence of elements in $\mathcal{X}$. Let k be a Mercer kernel that satisfies the polynomial eigendecay condition and let $\mathcal{T}_T(\alpha)$ be the set defined in Lemma D.9. For all $T \geq 1$ and all $\alpha > 0$,*

$$|\mathcal{T}_T(\alpha)| \leq \max\left(\frac{4e}{(e-1)\ln(1 + b^2)} \ln\left(1 + \frac{4\kappa}{\alpha \ln(1 + b^2)}\right),\right.$$

$$\left.\left(\frac{4e}{(e-1)\ln(1 + b^2)}\right)^{\frac{\beta_p}{\beta_p - 1}} \left(\frac{C_p \psi^2}{\alpha}\right)^{\frac{1}{\beta_p - 1}} \ln\left(1 + \frac{\kappa}{\alpha}\left(\frac{4}{\ln(1 + b^2)}\right)^{\frac{\beta_p}{\beta_p - 1}} \left(\frac{C_p \psi^2}{\alpha}\right)^{\frac{1}{\beta_p - 1}}\right)\right).$$

Proof. From Lemma D.9, we have $|\mathcal{T}_T(\alpha)| \leq \frac{2}{\ln(1 + b^2)} \gamma_{|\mathcal{T}_T(\alpha)|}(\alpha)$. The bound in Lemma D.3 tells us that

$$|\mathcal{T}_T(\alpha)| \leq \frac{2}{\ln(1 + b^2)} \left(\left(\frac{C_p \psi^2 |\mathcal{T}_T(\alpha)|}{\alpha}\right)^{1/\beta_p} \ln(1 + \kappa |\mathcal{T}_T(\alpha)| / \alpha)^{-1/\beta_p} + 1\right) \ln(1 + \kappa |\mathcal{T}_T(\alpha)| / \alpha). \tag{35}$$

We consider two cases. If $\frac{C_p \psi^2 |\mathcal{T}_T(\alpha)|}{\alpha \ln(1 + \kappa |\mathcal{T}_T(\alpha)| / \alpha)} \leq 1$, then

$$|\mathcal{T}_T(\alpha)| \leq \frac{4}{\ln(1 + b^2)} \ln(1 + \kappa |\mathcal{T}_T(\alpha)| / \alpha). \tag{36}$$

Using Corollary D.6, with $A = \frac{4}{\ln(1 + b^2)}$, $B = \kappa / \alpha$ and $X = |\mathcal{T}_T(\alpha)|$, we obtain

$$|\mathcal{T}_T(\alpha)| \leq \frac{4e}{(e-1)\ln(1 + b^2)} \ln\left(1 + \frac{4\kappa}{\alpha \ln(1 + b^2)}\right).$$

If $\frac{C_p\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\ln(1+\kappa|\mathcal{T}_T(\alpha)|/\alpha)} \geq 1$, then from (35), we have

$$|\mathcal{T}_T(\alpha)| \leq \frac{4}{\ln(1+b^2)} \left(\frac{C_p\psi^2|\mathcal{T}_T(\alpha)|}{\alpha}\right)^{1/\beta_p} \ln(1+\kappa|\mathcal{T}_T(\alpha)|/\alpha)^{1-1/\beta_p}.$$

After some algebra, we obtain

$$|\mathcal{T}_T(\alpha)| \leq \left(\frac{4}{\ln(1+b^2)}\right)^{\frac{\beta_p}{\beta_p-1}} \left(\frac{C_p\psi^2}{\alpha}\right)^{\frac{1}{\beta_p-1}} \ln(1+\kappa|\mathcal{T}_T(\alpha)|/\alpha).$$

Using Corollary D.6, with $A = \left(\frac{4}{\ln(1+b^2)}\right)^{\beta_p/(\beta_p-1)} \left(\frac{C_p\psi^2}{\alpha}\right)^{1/(\beta_p-1)}$, $B = \kappa/\alpha$ and $X = |\mathcal{T}_T(\alpha)|$, we obtain

$$|\mathcal{T}_T(\alpha)| \leq \left(\frac{4e}{(e-1)\ln(1+b^2)}\right)^{\frac{\beta_p}{\beta_p-1}} \left(\frac{C_p\psi^2}{\alpha}\right)^{\frac{1}{\beta_p-1}} \ln\left(1 + \frac{\kappa}{\alpha}\left(\frac{4}{\ln(1+b^2)}\right)^{\frac{\beta_p}{\beta_p-1}} \left(\frac{C_p\psi^2}{\alpha}\right)^{\frac{1}{\beta_p-1}}\right).$$

Taking the maximum of the two inequalities gives an upper bound $|\mathcal{T}_T(\alpha)|$ that holds for all $T \geq 1$ and $\alpha > 0$. $\quad\square$

Next we consider exponential eigendecay.

Lemma D.11 (Elliptical potentials: exponential eigendecay). *Let $b > 0$ and let $x_1, \ldots, x_T$ be any sequence of elements in $\mathcal{X}$. Let k be a Mercer kernel that satisfies the exponential eigendecay condition and let $\mathcal{T}_T(\alpha)$ be the set defined in Lemma D.9. If $\beta_e = 1$, then for all $T \geq 1$,*

$$|\mathcal{T}_T(\alpha)| \leq \max(A, B, C, D, E),$$

where $A = \frac{\alpha C_{e_2}}{C_{e_1}\psi^2}$,

$$B = \frac{4e}{(e-1)\ln(1+b^2)} \ln\left(1 + \frac{4\kappa}{\alpha\ln(1+b^2)}\right),$$

$$C = \frac{4}{\ln(1+b^2)} \left(\frac{2}{C_{e_2}}\right)^{\frac{1}{\beta_e}} \ln(2)^{1+\frac{1}{\beta_e}},$$

$$D = \frac{4}{\ln(1+b^2)} \left(\frac{2}{C_{e_2}}\right)^{\frac{1}{\beta_e}} \left(\frac{e(\beta_e+1)}{(e-1)\beta_e}\right)^{\frac{\beta_e+1}{\beta_e}} \ln\left(\left(\frac{8\kappa}{\alpha\ln(1+b^2)}\right)^{\beta_e/(\beta_e+1)} \left(\frac{2}{C_{e_2}}\right)^{\frac{1}{\beta_e+1}} \frac{\beta_e+1}{\beta_e}\right)^{\frac{\beta_e+1}{\beta_e}},$$

$$E = \frac{4}{\ln(1+b^2)} \left(\frac{2}{C_{e_2}}\right)^{\frac{1}{\beta_e}} \left(\frac{e(\beta_e+1)}{(e-1)\beta_e}\right)^{\frac{\beta_e+1}{\beta_e}} \ln\left(\left(\frac{4C_{e_1}\psi^2}{\alpha C_{e_2}\ln(1+b^2)}\right)^{\frac{\beta_e}{\beta_e+1}} \left(\frac{2}{C_{e_2}}\right)^{\frac{1}{\beta_e+1}} \frac{\beta_e+1}{\beta_e}\right)^{\frac{\beta_e+1}{\beta_e}}.$$

If $\beta_e < 1$, then

$$|\mathcal{T}_T(\alpha)| \leq \max(A, B, C, D, E, F),$$

$$\text{where} \quad A = \frac{\alpha \beta_e C_{e_2} \exp(-\tilde{C}_\beta)}{2 C_{e_1} \psi^2},$$

$$B = \frac{4e}{(e-1)\ln(1+b^2)} \ln\left(1 + \frac{4\kappa}{\alpha \ln(1+b^2)}\right),$$

$$C = \frac{4}{\ln(1+b^2)}\left(\frac{4}{C_{e_2}}\right)^{\frac{1}{\beta_e}} \ln(2)^{1+\frac{1}{\beta_e}},$$

$$D = \frac{4}{\ln(1+b^2)}\left(\frac{4}{C_{e_2}}\right)^{\frac{1}{\beta_e}}\left(\frac{e(\beta_e+1)}{(e-1)\beta_e}\right)^{\frac{\beta_e+1}{\beta_e}} \ln\left(\left(\frac{8\kappa}{\alpha \ln(1+b^2)}\right)^{\beta_e/(\beta_e+1)}\left(\frac{4}{C_{e_2}}\right)^{\frac{1}{\beta_e+1}}\frac{\beta_e+1}{\beta_e}\right)^{\frac{\beta_e+1}{\beta_e}},$$

$$E = \frac{4}{\ln(1+b^2)}\left(\frac{4}{C_{e_2}}\right)^{\frac{1}{\beta_e}}\left(\frac{e(\beta_e+1)}{(e-1)\beta_e}\right)^{\frac{\beta_e+1}{\beta_e}} \ln\left(\left(\frac{8C_{e_1}\psi^2}{\alpha\beta_e C_{e_2}\ln(1+b^2)}\right)^{\frac{\beta_e}{\beta_e+1}}\left(\frac{4}{C_{e_2}}\right)^{\frac{1}{\beta_e+1}}\frac{\beta_e+1}{\beta_e}\right)^{\frac{\beta_e+1}{\beta_e}},$$

$$F = \frac{2e}{(e-1)\ln(1+b^2)}\left(\left(\frac{4\tilde{C}_\beta}{C_{e_2}}\right)^{\frac{1}{\beta_e}}+1\right)\ln\left(1 + \frac{2\kappa}{\alpha \ln(1+b^2)}\left(\left(\frac{4\tilde{C}_\beta}{C_{e_2}}\right)^{\frac{1}{\beta_e}}+1\right)\right).$$

Component A of both bounds can be very loose when α is large. For every t and x_t, $\|k(\cdot, x_t)\|^2_{(\mathbf{V}_{t-1}+\alpha \mathbf{I}_{\mathcal{H}})^{-1}} \leq \frac{1}{\alpha}k(x_t, x_t) \leq \kappa/\alpha$. Therefore, whenever $\alpha \geq \kappa/b^2$, we have $|\mathcal{T}_T(\alpha)| = 0$. However, since we consider situations in which α is constant or decreasing in T, this bound suffices for our regret analysis.

Proof. From Lemma D.9, we have $|\mathcal{T}_T(\alpha)| \leq \frac{2}{\ln(1+b^2)}\gamma_{|\mathcal{T}_T(\alpha)|}(\alpha)$. The bound in Lemma D.3 tells us that

$$|\mathcal{T}_T(\alpha)| \leq \frac{2}{\ln(1+b^2)}\left(\left(\frac{2}{C_{e_2}}(\ln(|\mathcal{T}_T(\alpha)|)+C_\beta)\right)^{\frac{1}{\beta_e}}+1\right)\ln(1+\kappa|\mathcal{T}_T(\alpha)|/\alpha). \tag{37}$$

First, we consider the case when $\beta_e = 1$, which means $C_\beta = \ln(\frac{C_{e_1}\psi^2}{\alpha C_{e_2}})$, and (37) becomes

$$|\mathcal{T}_T(\alpha)| \leq \frac{2}{\ln(1+b^2)}\left(\left(\frac{2}{C_{e_2}}\ln\left(\frac{C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha C_{e_2}}\right)\right)^{\frac{1}{\beta_e}}+1\right)\ln(1+\kappa|\mathcal{T}_T(\alpha)|/\alpha). \tag{38}$$

Suppose that $\ln\left(\frac{C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha C_{e_2}}\right) \leq 0$. This implies the bound

$$|\mathcal{T}_T(\alpha)| \leq \frac{\alpha C_{e_2}}{C_{e_1}\psi^2}. \tag{39}$$

For the rest of the proof for $\beta_e = 1$, suppose that $\ln\left(\frac{C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha C_{e_2}}\right) \geq 0$. If $\frac{2}{C_{e_2}}\ln\left(\frac{C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha C_{e_2}}\right) \leq 1$, then we obtain

$$|\mathcal{T}_T(\alpha)| \leq \frac{4}{\ln(1+b^2)}\ln(1+\kappa|\mathcal{T}_T(\alpha)|/\alpha).$$

We know from the proof of Lemma D.10 that this inequality implies

$$|\mathcal{T}_T(\alpha)| \leq \frac{4e}{(e-1)\ln(1+b^2)}\ln\left(1 + \frac{4\kappa}{\alpha \ln(1+b^2)}\right). \tag{40}$$

If $\frac{2}{C_{e_2}} \ln\left(\frac{C_{e_1}\psi^2 |\mathcal{T}_T(\alpha)|}{\alpha C_{e_2}}\right) \geq 1$, then from (38), we obtain

$$|\mathcal{T}_T(\alpha)| \leq \frac{4}{\ln(1+b^2)} \left(\frac{2}{C_{e_2}}\right)^{\frac{1}{\beta_e}} \ln\left(\frac{C_{e_1}\psi^2 |\mathcal{T}_T(\alpha)|}{\alpha C_{e_2}}\right)^{\frac{1}{\beta_e}} \ln(1+\kappa|\mathcal{T}_T(\alpha)|/\alpha). \tag{41}$$

From here, if $\ln(\frac{C_{e_1}\psi^2 |\mathcal{T}_T(\alpha)|}{\alpha C_{e_2}}) \leq \ln(1 + \kappa|\mathcal{T}_T(\alpha)|/\alpha)$, then we must have $\ln(\frac{C_{e_1}\psi^2 |\mathcal{T}_T(\alpha)|}{\alpha C_{e_2}})^{1/\beta_e} \leq \ln(1 + \kappa|\mathcal{T}_T(\alpha)|/\alpha)^{1/\beta_e}$, since $1/\beta_e > 0$ and we are in the case where $\ln(\frac{C_{e_1}\psi^2 |\mathcal{T}_T(\alpha)|}{\alpha C_{e_2}}) \geq 0$. In this case, we have

$$|\mathcal{T}_T(\alpha)| \leq \frac{4}{\ln(1+b^2)} \left(\frac{2}{C_{e_2}}\right)^{\frac{1}{\beta_e}} \ln(1+\kappa|\mathcal{T}_T(\alpha)|/\alpha)^{1+\frac{1}{\beta_e}}. \tag{42}$$

We must have either $\kappa|\mathcal{T}_T(\alpha)|/\alpha \leq 1$ or $\kappa|\mathcal{T}_T(\alpha)|/\alpha \geq 1$. If $\kappa|\mathcal{T}_T(\alpha)|/\alpha \leq 1$, then we get

$$|\mathcal{T}_T(\alpha)| \leq \frac{4}{\ln(1+b^2)} \left(\frac{2}{C_{e_2}}\right)^{\frac{1}{\beta_e}} \ln(2)^{1+\frac{1}{\beta_e}}. \tag{43}$$

Otherwise, if $\kappa|\mathcal{T}_T(\alpha)|/\alpha \geq 1$, then

$$|\mathcal{T}_T(\alpha)| \leq \frac{4}{\ln(1+b^2)} \left(\frac{2}{C_{e_2}}\right)^{\frac{1}{\beta_e}} \ln(2\kappa|\mathcal{T}_T(\alpha)|/\alpha)^{1+\frac{1}{\beta_e}}.$$

After some algebra, we obtain

$$|\mathcal{T}_T(\alpha)|^{\beta_e/(\beta_e+1)} \leq \left(\frac{4}{\ln(1+b^2)}\right)^{\beta_e/(\beta_e+1)} \left(\frac{2}{C_{e_2}}\right)^{\frac{1}{\beta_e+1}} \frac{\beta_e+1}{\beta_e} \ln\left(\left(\frac{2\kappa}{\alpha}\right)^{\beta_e/(\beta_e+1)} |\mathcal{T}_T(\alpha)|^{\beta_e/(\beta_e+1)}\right).$$

Now, using Lemma D.7 with $A = (\frac{4}{\ln(1+b^2)})^{\beta_e/(\beta_e+1)}(\frac{2}{C_{e_2}})^{\frac{1}{\beta_e+1}} \frac{\beta_e+1}{\beta_e}$, $B = (\frac{2\kappa}{\alpha})^{\beta_e/(\beta_e+1)}$, $X = |\mathcal{T}_T(\alpha)|^{\beta_e/(\beta_e+1)}$ and $\eta = 1/e$, we get

$$|\mathcal{T}_T(\alpha)|^{\beta_e/(\beta_e+1)} \leq \left(\frac{4}{\ln(1+b^2)}\right)^{\beta_e/(\beta_e+1)} \left(\frac{2}{C_{e_2}}\right)^{\frac{1}{\beta_e+1}} \frac{e(\beta_e+1)}{(e-1)\beta_e} \ln\left(\left(\frac{8\kappa}{\alpha\ln(1+b^2)}\right)^{\beta_e/(\beta_e+1)} \left(\frac{2}{C_{e_2}}\right)^{\frac{1}{\beta_e+1}} \frac{\beta_e+1}{\beta_e}\right),$$

which means

$$|\mathcal{T}_T(\alpha)| \leq \frac{4}{\ln(1+b^2)} \left(\frac{2}{C_{e_2}}\right)^{\frac{1}{\beta_e}} \left(\frac{e(\beta_e+1)}{(e-1)\beta_e}\right)^{\frac{\beta_e+1}{\beta_e}} \ln\left(\left(\frac{8\kappa}{\alpha\ln(1+b^2)}\right)^{\beta_e/(\beta_e+1)} \left(\frac{2}{C_{e_2}}\right)^{\frac{1}{\beta_e+1}} \frac{\beta_e+1}{\beta_e}\right)^{\frac{\beta_e+1}{\beta_e}}. \tag{44}$$

Now, we return to (41). We have considered the case when $\ln(\frac{C_{e_1}\psi^2 |\mathcal{T}_T(\alpha)|}{\alpha C_{e_2}}) \leq \ln(1 + \kappa|\mathcal{T}_T(\alpha)|/\alpha)$. If instead, $\ln(\frac{C_{e_1}\psi^2 |\mathcal{T}_T(\alpha)|}{\alpha C_{e_2}}) \geq \ln(1 + \kappa|\mathcal{T}_T(\alpha)|/\alpha)$, then from (41), we have

$$|\mathcal{T}_T(\alpha)| \leq \frac{4}{\ln(1+b^2)} \left(\frac{2}{C_{e_2}}\right)^{\frac{1}{\beta_e}} \ln\left(\frac{C_{e_1}\psi^2 |\mathcal{T}_T(\alpha)|}{\alpha C_{e_2}}\right)^{1+\frac{1}{\beta_e}}.$$

After some algebra, this becomes

$$|\mathcal{T}_T(\alpha)|^{\frac{\beta_e}{\beta_e+1}} \leq \left(\frac{4}{\ln(1+b^2)}\right)^{\frac{\beta_e}{\beta_e+1}} \left(\frac{2}{C_{e_2}}\right)^{\frac{1}{\beta_e+1}} \frac{\beta_e+1}{\beta_e} \ln\left(\left(\frac{C_{e_1}\psi^2}{\alpha C_{e_2}}\right)^{\frac{\beta_e}{\beta_e+1}} |\mathcal{T}_T(\alpha)|^{\frac{\beta_e}{\beta_e+1}}\right).$$

Using Lemma D.7 with $A = (\frac{4}{\ln(1+b^2)})^{\beta_e/(\beta_e+1)}(\frac{2}{C_{e_2}})^{\frac{1}{\beta_e+1}}\frac{\beta_e+1}{\beta_e}$, $B = (\frac{C_{e_1}\psi^2}{\alpha C_{e_2}})^{\beta_e/(\beta_e+1)}$, $X = |\mathcal{T}_T(\alpha)|^{\beta_e/(\beta_e+1)}$ and $\eta = 1/e$, we get

$$|\mathcal{T}_T(\alpha)|^{\frac{\beta_e}{\beta_e+1}} \leq \left(\frac{4}{\ln(1+b^2)}\right)^{\frac{\beta_e}{\beta_e+1}}\left(\frac{2}{C_{e_2}}\right)^{\frac{1}{\beta_e+1}}\frac{e(\beta_e+1)}{(e-1)\beta_e}\ln\left(\left(\frac{4C_{e_1}\psi^2}{\alpha C_{e_2}\ln(1+b^2)}\right)^{\frac{\beta_e}{\beta_e+1}}\left(\frac{2}{C_{e_2}}\right)^{\frac{1}{\beta_e+1}}\frac{\beta_e+1}{\beta_e}\right),$$

which gives

$$|\mathcal{T}_T(\alpha)| \leq \frac{4}{\ln(1+b^2)}\left(\frac{2}{C_{e_2}}\right)^{\frac{1}{\beta_e}}\left(\frac{e(\beta_e+1)}{(e-1)\beta_e}\right)^{\frac{\beta_e+1}{\beta_e}}\ln\left(\left(\frac{4C_{e_1}\psi^2}{\alpha C_{e_2}\ln(1+b^2)}\right)^{\frac{\beta_e}{\beta_e+1}}\left(\frac{2}{C_{e_2}}\right)^{\frac{1}{\beta_e+1}}\frac{\beta_e+1}{\beta_e}\right)^{\frac{\beta_e+1}{\beta_e}}. \quad (45)$$

By taking the maximum of the bounds in (39), (40), (43), (44) and (45), we obtain the claimed result for $\beta_e = 1$. Now, we consider the case where $\beta_e < 1$, which means $C_\beta = \ln(\frac{2C_{e_1}\psi^2}{\alpha\beta_e C_{e_2}}) + \tilde{C}_\beta$, where $\tilde{C}_\beta = (1/\beta_e - 1)(\ln(\frac{2}{C_{e_2}}(1/\beta_e - 1)) - 1)$, and (37) becomes

$$|\mathcal{T}_T(\alpha)| \leq \frac{2}{\ln(1+b^2)}\left(\left(\frac{2}{C_{e_2}}\left(\ln\left(\frac{2C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\beta_e C_{e_2}}\right) + \tilde{C}_\beta\right)\right)^{\frac{1}{\beta_e}} + 1\right)\ln(1 + \kappa|\mathcal{T}_T(\alpha)|/\alpha). \quad (46)$$

Note that if $\ln(\frac{2C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\beta_e C_{e_2}}) + \tilde{C}_\beta \leq 0$, then we have the bound

$$|\mathcal{T}_T(\alpha)| \leq \frac{\alpha\beta_e C_{e_2}\exp(-\tilde{C}_\beta)}{2C_{e_1}\psi^2}. \quad (47)$$

For the rest of the proof, we consider the case when $\ln(\frac{2C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\beta_e C_{e_2}}) + \tilde{C}_\beta \geq 0$, which means $(\ln(\frac{2C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\beta_e C_{e_2}}) + \tilde{C}_\beta)^{\frac{1}{\beta_e}}$ is a nonnegative real. If $\tilde{C}_{\beta_e} \leq \ln(\frac{2C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\beta_e C_{e_2}})$, then since $\ln(\frac{2C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\beta_e C_{e_2}}) + \tilde{C}_\beta \geq 0$, $\ln(\frac{2C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\beta_e C_{e_2}})$ must be nonnegative and we must have $(\ln(\frac{2C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\beta_e C_{e_2}}) + \tilde{C}_\beta)^{1/\beta_e} \leq (2\ln(\frac{2C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\beta_e C_{e_2}}))^{1/\beta_e}$. From (46), we then have

$$|\mathcal{T}_T(\alpha)| \leq \frac{2}{\ln(1+b^2)}\left(\left(\frac{4}{C_{e_2}}\ln\left(\frac{2C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\beta_e C_{e_2}}\right)\right)^{\frac{1}{\beta_e}} + 1\right)\ln(1 + \kappa|\mathcal{T}_T(\alpha)|/\alpha). \quad (48)$$

If $\frac{4}{C_{e_2}}\ln(\frac{2C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\beta_e C_{e_2}}) \leq 1$, then we have

$$|\mathcal{T}_T(\alpha)| \leq \frac{4}{\ln(1+b^2)}\ln(1 + \kappa|\mathcal{T}_T(\alpha)|/\alpha).$$

Corollary D.6 with $A = \frac{4}{\ln(1+b^2)}$, $B = \kappa/\alpha$, $X = |\mathcal{T}_T(\alpha)|$ then tells us that

$$|\mathcal{T}_T(\alpha)| \leq \frac{4e}{(e-1)\ln(1+b^2)}\ln\left(1 + \frac{4\kappa}{\alpha\ln(1+b^2)}\right). \quad (49)$$

Returning to (48), if $\frac{4}{C_{e_2}}\ln(\frac{2C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\beta_e C_{e_2}}) \geq 1$, then we have

$$|\mathcal{T}_T(\alpha)| \leq \frac{4}{\ln(1+b^2)}\left(\frac{4}{C_{e_2}}\ln\left(\frac{2C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\beta_e C_{e_2}}\right)\right)^{\frac{1}{\beta_e}}\ln(1 + \kappa|\mathcal{T}_T(\alpha)|/\alpha). \quad (50)$$

From here, if $\ln(\frac{2C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\beta_e C_{e_2}}) \leq \ln(1 + \kappa|\mathcal{T}_T(\alpha)|/\alpha)$, then we must have $\ln(\frac{2C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\beta_e C_{e_2}})^{1/\beta_e} \leq \ln(1 + \kappa|\mathcal{T}_T(\alpha)|/\alpha)^{1/\beta_e}$, since $1/\beta_e > 0$ and $\ln(\frac{2C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\beta_e C_{e_2}})$ is nonnegative. In this case, we have

$$|\mathcal{T}_T(\alpha)| \leq \frac{4}{\ln(1+b^2)}\left(\frac{4}{C_{e_2}}\right)^{\frac{1}{\beta_e}}\ln(1 + \kappa|\mathcal{T}_T(\alpha)|/\alpha)^{1+\frac{1}{\beta_e}}.$$

Note that this inequality is the same as (42), except a 2 is replaced by a 4. As before, this inequality implies that

$$|\mathcal{T}_T(\alpha)| \le \max\left(\frac{4}{\ln(1+b^2)}\left(\frac{4}{C_{e_2}}\right)^{\frac{1}{\beta_e}}\ln(2)^{1+\frac{1}{\beta_e}},\right. \tag{51}$$

$$\left.\frac{4}{\ln(1+b^2)}\left(\frac{4}{C_{e_2}}\right)^{\frac{1}{\beta_e}}\left(\frac{e(\beta_e+1)}{(e-1)\beta_e}\right)^{\frac{\beta_e+1}{\beta_e}}\ln\left(\left(\frac{8\kappa}{\alpha\ln(1+b^2)}\right)^{\beta_e/(\beta_e+1)}\left(\frac{4}{C_{e_2}}\right)^{\frac{1}{\beta_e+1}}\frac{\beta_e+1}{\beta_e}\right)^{\frac{\beta_e+1}{\beta_e}}\right).$$

Returning to (50), if $\ln(\frac{2C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\beta_e C_{e_2}}) \ge \ln(1+\kappa|\mathcal{T}_T(\alpha)|/\alpha)$, then

$$|\mathcal{T}_T(\alpha)| \le \frac{4}{\ln(1+b^2)}\left(\frac{4}{C_{e_2}}\right)^{\frac{1}{\beta_e}}\ln\left(\frac{2C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\beta_e C_{e_2}}\right)^{1+\frac{1}{\beta_e}}. \tag{52}$$

After some algebra, we obtain

$$|\mathcal{T}_T(\alpha)|^{\frac{\beta_e}{\beta_e+1}} \le \left(\frac{4}{\ln(1+b^2)}\right)^{\frac{\beta_e}{\beta_e+1}}\left(\frac{4}{C_{e_2}}\right)^{\frac{1}{\beta_e+1}}\frac{\beta_e+1}{\beta_e}\ln\left(\left(\frac{2C_{e_1}\psi^2}{\alpha\beta_e C_{e_2}}\right)^{\frac{\beta_e}{\beta_e+1}}|\mathcal{T}_T(\alpha)|^{\frac{\beta_e}{\beta_e+1}}\right).$$

Using Lemma D.7 with $A = (\frac{4}{\ln(1+b^2)})^{\frac{\beta_e}{\beta_e+1}}(\frac{4}{C_{e_2}})^{\frac{1}{\beta_e+1}}\frac{\beta_e+1}{\beta_e}$, $B = (\frac{2C_{e_1}\psi^2}{\alpha\beta_e C_{e_2}})^{\frac{\beta_e}{\beta_e+1}}$, $X = |\mathcal{T}_T(\alpha)|^{\beta_e/(\beta_e+1)}$ and $\eta = 1/e$, we get

$$|\mathcal{T}_T(\alpha)|^{\frac{\beta_e}{\beta_e+1}} \le \left(\frac{4}{\ln(1+b^2)}\right)^{\frac{\beta_e}{\beta_e+1}}\left(\frac{4}{C_{e_2}}\right)^{\frac{1}{\beta_e+1}}\frac{e(\beta_e+1)}{(e-1)\beta_e}\ln\left(\left(\frac{8C_{e_1}\psi^2}{\alpha\beta_e C_{e_2}\ln(1+b^2)}\right)^{\frac{\beta_e}{\beta_e+1}}\left(\frac{4}{C_{e_2}}\right)^{\frac{1}{\beta_e+1}}\frac{\beta_e+1}{\beta_e}\right),$$

which gives

$$|\mathcal{T}_T(\alpha)| \le \frac{4}{\ln(1+b^2)}\left(\frac{4}{C_{e_2}}\right)^{\frac{1}{\beta_e}}\left(\frac{e(\beta_e+1)}{(e-1)\beta_e}\right)^{\frac{\beta_e+1}{\beta_e}}\ln\left(\left(\frac{8C_{e_1}\psi^2}{\alpha\beta_e C_{e_2}\ln(1+b^2)}\right)^{\frac{\beta_e}{\beta_e+1}}\left(\frac{4}{C_{e_2}}\right)^{\frac{1}{\beta_e+1}}\frac{\beta_e+1}{\beta_e}\right)^{\frac{\beta_e+1}{\beta_e}}. \tag{53}$$

Returning to (46), if $\ln(\frac{2C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\beta_e C_{e_2}}) \le \tilde{C}_\beta$, then since $\ln(\frac{2C_{e_1}\psi^2|\mathcal{T}_T(\alpha)|}{\alpha\beta_e C_{e_2}}) + \tilde{C}_\beta \ge 0$, $\tilde{C}_\beta$ must be nonnegative. From (46), we have

$$|\mathcal{T}_T(\alpha)| \le \frac{2}{\ln(1+b^2)}\left(\left(\frac{4\tilde{C}_\beta}{C_{e_2}}\right)^{\frac{1}{\beta_e}}+1\right)\ln(1+\kappa|\mathcal{T}_T(\alpha)|/\alpha)$$

Using Corollary D.6, with $A = \frac{2}{\ln(1+b^2)}((\frac{4\tilde{C}_\beta}{C_{e_2}})^{\frac{1}{\beta_e}}+1)$, $B = \kappa/\alpha$, $X = |\mathcal{T}_T(\alpha)|$, we obtain

$$|\mathcal{T}_T(\alpha)| \le \frac{2e}{(e-1)\ln(1+b^2)}\left(\left(\frac{4\tilde{C}_\beta}{C_{e_2}}\right)^{\frac{1}{\beta_e}}+1\right)\ln\left(1+\frac{2\kappa}{\alpha\ln(1+b^2)}\left(\left(\frac{4\tilde{C}_\beta}{C_{e_2}}\right)^{\frac{1}{\beta_e}}+1\right)\right). \tag{54}$$

Finally, by taking the maximum of the bounds in (47), (49), (51), (53) and (54), we obtain the claimed result for $\beta_e < 1$. □

D.3 Generic Regret Bound

Before proving the generic regret bound in Thm. 6.2, we state a useful bound on the regret in a given round t. The proof is standard in the analysis of optimistic bandit algorithms (e.g. Prop. 1 in Russo and Van Roy, 2013), but we include it here for completeness.

Lemma D.12 (Per-round regret bound). *For any $c > 0$ and any $\delta \in (0, 1]$, let $x_1, x_2, \ldots$ be the actions chosen by Kernel CMM-UCB, Kernel DMM-UCB (with any grid A that contains α) or Kernel AMM-UCB (with any fixed α). With probability at least $1 - \delta$,*

$$\forall t \geq 1, \qquad r_t \leq \frac{2\widetilde{R}_{\alpha,t-1}}{\sqrt{\alpha}} \rho_{\alpha,t-1}(x_t).$$

Proof. We prove the result for Kernel CMM-UCB. The same argument applies for Kernel DMM-UCB and Kernel AMM-UCB. Using Lemma 4.2, for any $c > 0$, with probability at least $1 - \delta$, we have

$$\forall t \geq 1, \forall x \in \mathcal{X}, \qquad \text{LCB}_{\mathcal{F}_t}(x) \leq f^\star(x) \leq \text{UCB}_{\mathcal{F}_t}(x).$$

Since $x_1, x_2, \ldots$ are the actions chosen by Kernel CMM-UCB, we have $x_t = \text{argmax}_{x \in \mathcal{X}_t} \{\text{UCB}_{\mathcal{F}_{t-1}}(x)\}$. Therefore,

$$\begin{aligned}
r_t &= f^\star(x_t^\star) - f^\star(x_t) \\
&\leq \text{UCB}_{\mathcal{F}_{t-1}}(x_t^\star) - \text{LCB}_{\mathcal{F}_{t-1}}(x_t) \\
&\leq \text{UCB}_{\mathcal{F}_{t-1}}(x_t) - \text{LCB}_{\mathcal{F}_{t-1}}(x_t) \\
&\leq \frac{2\widetilde{R}_{\alpha,t-1}}{\sqrt{\alpha}} \rho_{\alpha,t-1}(x_t).
\end{aligned}$$

The final inequality uses (13) and the analogous lower bound for $\text{LCB}_{\mathcal{F}_t}(x)$. $\square$

We can now prove the generic regret bound in Thm. 6.2.

Proof of Thm. 6.2. Using Lemma D.12 with $\alpha = \sigma^2/c$, with probability at least $1 - \delta$, we have

$$\forall t \geq 1, \qquad r_t \leq \frac{2\widetilde{R}_{\sigma^2/c,t-1}}{\sqrt{\sigma^2/c}} \rho_{\sigma^2/c,t-1}(x_t).$$

From our boundedness assumptions, we also have $r_t \leq 2C$. We define the set $\mathcal{T}_T(\alpha)$ as

$$\mathcal{T}_T(\alpha) = \left\{ t \in [T] : \frac{1}{\alpha} \rho_{\alpha,t-1}^2(x_t) \geq 1 \right\}.$$

Recall that $\tau_T(\alpha) = |\mathcal{T}_T(\alpha)|$. From Lemma B.3, we have $\frac{1}{\alpha} \rho_{\alpha,t-1}^2(x_t) = \|k(\cdot, x_t)\|_{(V_{t-1}+\alpha I_{\mathcal{H}})^{-1}}^2$, which means $\mathcal{T}_T(\alpha)$ is the same as the set defined in Lemma D.9, with $b = 1$. By decomposing the cumulative regret, and then using Cauchy-Schwarz, we obtain

$$\begin{aligned}
\sum_{t=1}^T r_t &= \sum_{t \in \mathcal{T}_T} r_t + \sum_{t \notin \mathcal{T}_T} r_t \\
&\leq 2C\tau_T(\tfrac{\sigma^2}{c}) + \sqrt{T \sum_{t \notin \mathcal{T}_T} r_t^2} \\
&\leq 2C\tau_T(\tfrac{\sigma^2}{c}) + \sqrt{T \sum_{t \notin \mathcal{T}_T} 4\widetilde{R}_{\sigma^2/c,t-1}^2 \frac{c}{\sigma^2} \rho_{\sigma^2/c,t-1}^2(x_t)} \\
&\leq 2C\tau_T(\tfrac{\sigma^2}{c}) + \sqrt{4 \max_{s \in \{0,1,\ldots,T\}} \widetilde{R}_{\sigma^2/c,s}^2 T \sum_{t=1}^T \min\left(1, \frac{c}{\sigma^2} \rho_{\sigma^2/c,t-1}^2(x_t)\right)}.
\end{aligned}$$

Using the bound on the squared radius and Lemma D.4, we have

$$\sum_{t=1}^T r_t \leq 2C\tau_T\left(\frac{\sigma^2}{c}\right) + \sigma\sqrt{24T\gamma_T\left(\frac{\sigma^2}{c}\right)\left(\gamma_T\left(\frac{\sigma^2}{c}\right) + \frac{B^2}{2c} + \ln\left(\frac{1}{\delta}\right)\right)} \tag{55}$$

$\square$

D.4 Regret Bounds By Eigendecay Rate

Here, we determine regret bounds for kernels satisfying the polynomial or exponential eigendecay conditions. We focus on the dependence of the regret on T (and c, since we choose this based on T). First, suppose that the kernel k satisfies the polynomial eigenvalue decay condition.

Proposition D.13 (Regret Bound for Polynomial Decay). *Suppose that k satisfies the polynomial eigendecay condition, and let $c \propto T^{-\frac{1}{\beta_p+1}}$. For any $T \geq 1$ and any $\delta \in (0,1]$, with probability at least $1 - \delta$, the cumulative regret of Kernel CMM-UCB, Kernel DMM-UCB (using a grid A containing $\alpha = \sigma^2/c$) and Kernel AMM-UCB (using $\alpha = \sigma^2/c$) satisfies*

$$r_{1:T} = \mathcal{O}\left(T^{\frac{\beta_p+3}{2\beta_p+2}} \ln(T)^{1-\frac{1}{\beta_p}}\right).$$

Proof. From the bounds in Lemma D.3 and Lemma D.10, we have

$$\gamma_T\left(\frac{\sigma^2}{c}\right) = \mathcal{O}\left((cT)^{\frac{1}{\beta_p}} \ln(cT)^{1-\frac{1}{\beta_p}}\right), \quad \text{and} \quad \tau_T\left(\frac{\sigma^2}{c}\right) = \mathcal{O}\left(c^{\frac{1}{\beta_p-1}} \ln\left(c^{\frac{\beta_p}{\beta_p-1}}\right)\right).$$

If we plug these into the regret bound in (55), we get

$$\sum_{t=1}^{T} r_t = \mathcal{O}\left(c^{\frac{1}{\beta_p-1}} \ln\left(c^{\frac{\beta_p}{\beta_p-1}}\right) + T^{\frac{1}{2}+\frac{1}{\beta_p}} c^{\frac{1}{\beta_p}} \ln(cT)^{1-\frac{1}{\beta_p}} + T^{\frac{1}{2}+\frac{1}{2\beta_p}} c^{\frac{1}{2\beta_p}-\frac{1}{2}} \ln(cT)^{\frac{1}{2}-\frac{1}{2\beta_p}}\right).$$

If we set $c \propto T^q$, for some q to be chosen later, this becomes

$$\sum_{t=1}^{T} r_t = \mathcal{O}\left(T^{\frac{q}{\beta_p-1}} \ln\left(T^{\frac{q\beta_p}{\beta_p-1}}\right) + T^{\frac{1}{2}+\frac{1}{\beta_p}+\frac{q}{\beta_p}} \ln\left(T^{1+q}\right)^{1-\frac{1}{\beta_p}} + T^{\frac{1}{2}+\frac{1}{2\beta_p}+\frac{q}{2\beta_p}-\frac{q}{2}} \ln\left(T^{1+q}\right)^{\frac{1}{2}-\frac{1}{2\beta_p}}\right).$$

If we choose $q = -\frac{1}{\beta_p+1}$, then

$$\frac{1}{2} + \frac{1}{\beta_p} + \frac{q}{\beta_p} = \frac{1}{2} + \frac{1}{2\beta_p} + \frac{q}{2\beta_p} - \frac{q}{2} = \frac{\beta_p+3}{2\beta_p+2}.$$

Therefore, with this choice of q, the regret bound is

$$\sum_{t=1}^{T} r_t = \mathcal{O}\left(T^{\frac{\beta_p+3}{2\beta_p+2}} \ln(T)^{1-\frac{1}{\beta_p}}\right).$$

□

This matches the regret bound in Thm. 2 in (Whitehouse et al., 2023), which we believe to be the best known regret bound for (vanilla) KernelUCB/GP-UCB under polynomial eigendecay. It is known that the Matérn kernel, with smoothness parameter $\nu > 1/2$, satisfies the polynomial eigendecay condition with $\beta_p = \frac{2\nu+d}{d}$ (Santin and Schaback, 2016). Therefore, if we choose $c \propto T^{-\frac{d}{2\nu+2d}}$, the the regret bound is

$$\sum_{t=1}^{T} r_t = \mathcal{O}\left(T^{\frac{\nu+2d}{2\nu+2d}} \ln(T)^{\frac{2\nu}{2\nu+d}}\right).$$

Now, suppose that the kernel k satisfies the exponential eigendecay condition.

Proposition D.14 (Regret Bound for Exponential Eigendecay). *Suppose that k satisfies the exponential eigendecay condition, and let $c \propto 1$. For any $T \geq 1$ and any $\delta \in (0,1]$, with probability at least $1 - \delta$, the cumulative regret of Kernel CMM-UCB, Kernel DMM-UCB (using a grid A containing $\alpha = \sigma^2/c$) and Kernel AMM-UCB (using $\alpha = \sigma^2/c$) satisfies*

$$r_{1:T} = \mathcal{O}\left(\sqrt{T} \ln(T)^{1+\frac{1}{\beta_e}}\right).$$

Proof. From the bounds in Lemma D.3 and Lemma D.11, we have

$$\gamma_T\left(\frac{\sigma^2}{c}\right) = \mathcal{O}\left(\ln\left(cT\right)^{1+\frac{1}{\beta_e}}\right), \quad \text{and} \quad \tau_T\left(\frac{\sigma^2}{c}\right) = \mathcal{O}\left(\max\left(\frac{1}{c}, \ln\left(c\right)^{1+\frac{1}{\beta_e}}\right)\right).$$

If we plug these into (55), we get

$$\sum_{t=1}^{T} r_t = \mathcal{O}\left(\max\left(\frac{1}{c}, \ln\left(c\right)^{1+\frac{1}{\beta_e}}\right) + \sqrt{T}\ln\left(cT\right)^{1+\frac{1}{\beta_e}} + \sqrt{\frac{T}{c}\ln\left(cT\right)^{1+\frac{1}{\beta_e}}}\right).$$

If we set $c \propto 1$, this becomes

$$\sum_{t=1}^{T} r_t = \mathcal{O}\left(\sqrt{T}\ln\left(T\right)^{1+\frac{1}{\beta_e}}\right).$$

$\square$

It is known that the RBF kernel satisfies the exponential eigendecay condition with $\beta_e = 1/d$ (Belkin, 2018). This means that, when $c \propto 1$, the regret bound is

$$\sum_{t=1}^{T} r_t = \mathcal{O}\left(\sqrt{T}\ln\left(T\right)^{d+1}\right).$$

Note that one could also use the regret bound in (55) to determine how c should depend on B. For exponential eigendecay, one can show that the choice $c \propto B^2$ leads to a regret bound which is polylogarithmic in B in the dominant $\sqrt{T}$ term. One can also consider the dependence of the regret on σ. For exponential eigendecay, when $\sigma \propto 1/\sqrt{T}$, the regret bound becomes polylogarithmic in T. Therefore, KernelUCB with our confidence bounds can still enjoy polylog(T) regret even when the problem is only approximately noiseless.

D.5 Regret Bounds for the Noiseless Setting

In Sec. 6.3, we showed that for all $\alpha > 0$ and $T \geq 1$,

$$r_{1:T} \leq 2C\tau_T(\alpha) + \sqrt{12\alpha B^2 T \gamma_T(\alpha)}. \tag{56}$$

First, suppose that the kernel k satisfies the polynomial eigendecay condition. In the noiseless setting, we can prove the following regret bound.

Proposition D.15 (Regret Bound for Polynomial Decay in the Noiseless Setting). *Suppose that k satisfies the polynomial eigendecay condition, and let $\alpha \propto T^{\frac{1-\beta_p^2}{\beta_p^2+1}}$. For any $T \geq 1$, the cumulative regret of Kernel CMM-UCB, Kernel DMM-UCB (using a grid A containing this α) and Kernel AMM-UCB (using this α) satisfies*

$$r_{1:T} = \mathcal{O}\left(T^{\frac{\beta_p+1}{\beta_p^2+1}}\ln\left(T\right)\right).$$

This is always sublinear in T, since $\beta_p > 1$, and has a growth rate slower than $\sqrt{T}$ whenever $\beta_p > 1 + \sqrt{2}$.

Proof. We're interested in the case where $\alpha \propto T^q$ for some $q < 0$. In this regime, the bounds in Lemma D.3 and Lemma D.10 tell us that

$$\tau_T(\alpha) = \mathcal{O}\left(\alpha^{-\frac{1}{\beta_p-1}}\ln\left(\alpha^{-\frac{\beta_p}{\beta_p-1}}\right)\right), \quad \gamma_T(\alpha) = \mathcal{O}\left(\left(\frac{T}{\alpha}\right)^{\frac{1}{\beta_p}}\ln\left(\frac{T}{\alpha}\right)^{1-\frac{1}{\beta_p}}\right).$$

This means that (56) becomes

$$r_{1:T} = \mathcal{O}\left(\alpha^{-\frac{1}{\beta_p-1}}\ln\left(\alpha^{-\frac{\beta_p}{\beta_p-1}}\right) + T^{\frac{1}{2}+\frac{1}{2\beta_p}}\alpha^{\frac{1}{2}-\frac{1}{2\beta_p}}\ln\left(\frac{T}{\alpha}\right)^{\frac{1}{2}-\frac{1}{2\beta_p}}\right).$$

If $\alpha \propto T^q$, for some $q < 0$ to be chosen later, then

$$r_{1:T} = \mathcal{O}\left(T^{-\frac{q}{\beta_p-1}} \ln\left(T^{-\frac{\beta_p q}{\beta_p-1}}\right) + T^{\frac{1}{2}+\frac{1}{2\beta_p}+\frac{q}{2}-\frac{q}{2\beta_p}} \ln\left(T^{1-q}\right)^{\frac{1}{2}-\frac{1}{2\beta_p}}\right).$$

We want to choose q such that the maximum of the powers of T is minimised. We notice that, since $\beta_p > 1$, $-\frac{q}{\beta_p-1}$ is monotonically decreasing in q and $\frac{1}{2}+\frac{1}{2\beta_p}+\frac{q}{2}-\frac{q}{2\beta_p}$ is monotonically increasing in q. If we choose $q = \frac{1-\beta_p^2}{\beta_p^2+1}$, then

$$-\frac{q}{\beta_p-1} = \frac{1}{2}+\frac{1}{2\beta_p}+\frac{q}{2}-\frac{q}{2\beta_p} = \frac{\beta_p+1}{\beta_p^2+1}.$$

Therefore, with this choice of q, we have

$$r_{1:T} = \mathcal{O}\left(T^{\frac{\beta_p+1}{\beta_p^2+1}} \ln(T)\right).$$

$\square$

Consider the the Matérn kernel, where $\beta_p = \frac{2\nu+d}{d}$. If we choose $\alpha \propto T^{-\frac{2\nu^2+2\nu d}{2\nu^2+2\nu d+d^2}}$, then the regret bound for the noiseless case is

$$r_{1:T} = \mathcal{O}\left(T^{\frac{\nu d+d^2}{2\nu^2+2\nu d+d^2}} \ln(T)\right).$$

Whenever $\nu > d/\sqrt{2}$, the growth rate is slower than $\sqrt{T}$. Now, suppose that the kernel satisfies the exponential eigendecay condition. In the noiseless setting, we can prove the following regret bound.

Proposition D.16 (Regret Bound for Exponential Decay in the Noiseless Setting). *Suppose that k satisfies the exponential eigendecay condition, and let $\alpha \propto 1$. For any $T \geq 1$, the cumulative regret of Kernel CMM-UCB, Kernel DMM-UCB (using a grid A containing this α) and Kernel AMM-UCB (using this α) satisfies*

$$r_{1:T} = \mathcal{O}\left(\ln(T)^{1+\frac{1}{\beta_e}}\right).$$

Proof. We're interested in the case where $\alpha \propto T^q$ for some $-1 \leq q \leq 0$. In this regime, the bounds in Lemma D.3 and Lemma D.11 tell us that

$$\tau_T(\alpha) = \mathcal{O}\left(\ln\left(\tfrac{1}{\alpha}\right)^{1+\frac{1}{\beta_e}}\right), \quad \gamma_T(\alpha) = \mathcal{O}\left(\ln\left(\tfrac{T}{\alpha}\right)^{1+\frac{1}{\beta_e}}\right).$$

This means that (56) becomes

$$r_{1:T} = \mathcal{O}\left(\ln\left(\tfrac{1}{\alpha}\right)^{1+\frac{1}{\beta_e}} + T^{\frac{1}{2}}\alpha^{\frac{1}{2}} \ln\left(\tfrac{T}{\alpha}\right)^{\frac{1}{2}+\frac{1}{2\beta_e}}\right).$$

If we choose $\alpha = 1/T$, then

$$r_{1:T} = \mathcal{O}\left(\ln(T)^{1+\frac{1}{\beta_e}}\right).$$

$\square$

For the RBF kernel, where $\beta_e = 1/d$, we get

$$r_{1:T} = \mathcal{O}\left(\ln(T)^{d+1}\right).$$

E Additional Practical Details

In this section, we comment on some practical details to aid with reproducibility.

E.1 Sample CVXPY Code For Solving (10)

In Fig. 3, we provide some sample CVXPY code for computing the solution of the second order cone program in (10). `_w` is the optimisation variable $\boldsymbol{w} \in \mathbb{R}^{t+1}$; `k_t1` is $\boldsymbol{k}_{t+1}(x)$; `k_tt1` is the $t \times (t+1)$ kernel matrix $\boldsymbol{K}_{t,t+1}$; `y_t` is the reward vector $\boldsymbol{y}_t$; `R_t` is the radius R_t; `B` is the norm bound B; `l_t1` is the right Cholesky factor $\boldsymbol{L}_{t+1}$ of the $(t+1) \times (t+1)$ kernel matrix $\boldsymbol{K}_{t+1}$ (so `l_t1` is an upper triangular matrix).

```python
import cvxpy as cp

_w = cp.Variable(t+1)
obj = k_t1.T @ _w
cons = [cp.norm(k_tt1 @ _w - y_t) <= R_t,
        cp.norm(l_t1 @ _w) <= B]
prob = cp.Problem(cp.Maximize(obj), cons)
ucb = prob.solve()
```

Figure 3: Python code for solving (10) with CVXPY.

The variable `ucb` will now be equal to the numerical solution of $\max_{f \in \mathcal{F}_t} \{f(x)\}$. In practice, we find that it is helpful to add a small multiple of the identity matrix to the kernel matrix $\boldsymbol{K}_{t+1}$ before computing the Cholesky decomposition. The reason is that, due to numerical error, small positive eigenvalues of $\boldsymbol{K}_{t+1}$ can appear to be negative. One could avoid the necessity of computing the Cholesky decomposition of $\boldsymbol{K}_{t+1}$ by expressing the second constraint as $\boldsymbol{w}^\top \boldsymbol{K}_{t+1} \boldsymbol{w} \le B^2$ (which is still convex in $\boldsymbol{w}$). However, the conic form of (10) (with norm constraints) is favourable for the conic solvers used by CVXPY.

E.2 Recursive Updates

For a more efficient implementation of our Kernel CMM-UCB, Kernel DMM-UCB and Kernel AMM-UCB algorithms, the inverse $(\boldsymbol{K}_t + \alpha \boldsymbol{I})^{-1}$, the log determinant $\ln(\det(\frac{1}{\alpha}\boldsymbol{K}_t + \boldsymbol{I}))$ and the right Cholesky factor $\boldsymbol{L}_t$ can be updated using the following recursive formulas. These update rules can be derived using properties of the Schur complement (Zhang, 2006), and have been used before in other kernel bandit algorithms (see e.g. App. F of (Chowdhury and Gopalan, 2017)).

For the inverse, we have

$$K_{22} = \frac{1}{k(x_{t+1}, x_{t+1}) + \alpha - \boldsymbol{k}_t(x_{t+1})^\top (\boldsymbol{K}_t + \alpha \boldsymbol{I})^{-1} \boldsymbol{k}_t(x_{t+1})},$$
$$K_{11} = (\boldsymbol{K}_t + \alpha \boldsymbol{I})^{-1} + K_{22} \boldsymbol{k}_t(x_{t+1})^\top (\boldsymbol{K}_t + \alpha \boldsymbol{I})^{-1} \boldsymbol{k}_t(x_{t+1}),$$
$$K_{12} = -K_{22} (\boldsymbol{K}_t + \alpha \boldsymbol{I})^{-1} \boldsymbol{k}_t(x_{t+1}),$$
$$K_{21} = -K_{22} \boldsymbol{k}_t(x_{t+1})^\top (\boldsymbol{K}_t + \alpha \boldsymbol{I})^{-1},$$
$$(\boldsymbol{K}_{t+1} + \alpha \boldsymbol{I})^{-1} = \left[\begin{array}{c|c} K_{11} & K_{12} \\ \hline K_{21} & K_{22} \end{array} \right].$$

For the log determinant, we have

$$\ln\left(\det\left(\frac{1}{\alpha}\boldsymbol{K}_{t+1} + \boldsymbol{I}\right)\right) = \ln\left(\det\left(\frac{1}{\alpha}\boldsymbol{K}_t + \boldsymbol{I}\right)\right) + \ln\left(1 + \frac{k(x_{t+1}, x_{t+1}) - \boldsymbol{k}_t(x_{t+1})^\top (\boldsymbol{K}_t + \alpha \boldsymbol{I})^{-1} \boldsymbol{k}_t(x_{t+1})}{\alpha}\right).$$

For the right Cholesky factor, we have

$$\boldsymbol{L}_{t+1} = \left[\begin{array}{c|c} \boldsymbol{L}_t & \boldsymbol{k}_t(x_{t+1})^\top \boldsymbol{L}_t^{-1} \\ \hline \boldsymbol{0} & \sqrt{k(x_{t+1}, x_{t+1}) - \|\boldsymbol{k}_t(x_{t+1})^\top \boldsymbol{L}_t^{-1}\|_2^2} \end{array} \right].$$

Note that the vector $\boldsymbol{k}_t(x_{t+1})^\top \boldsymbol{L}_t^{-1}$ is the transpose of $\boldsymbol{v}$, where $\boldsymbol{v}$ is the solution of $\boldsymbol{L}_t^\top \boldsymbol{v} = \boldsymbol{k}_t(x_{t+1})$. $\boldsymbol{v}$ can be computed using numerical methods for solving lower triangular systems (see e.g. `linalg.triangular_solve` in the SciPy library), so it is not necessary to invert $\boldsymbol{L}_t$.

F   Additional Experimental Results

In this section, we present some additional experimental results.

F.1   Time Per Step

We run each method in a synthetic kernel bandit problem, as described in Sec. 7. We use the Matérn kernel with smoothness $\nu = 5/2$ and length-scale $\ell = 0.5$.
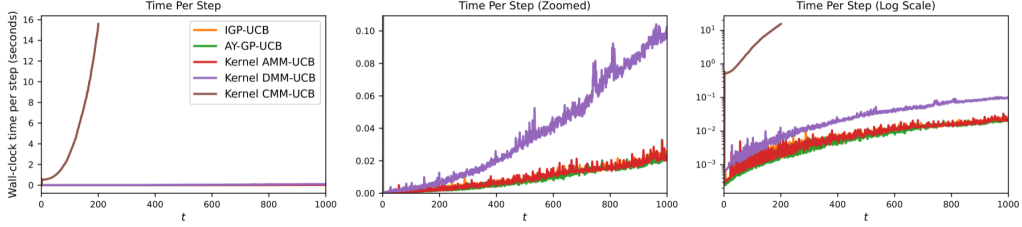

Figure 4: Wall-clock time per step for our kernel UCB-style algorithms as well as AY-GP-UCB, IGP-UCB over $T = 1000$ rounds with: (left) original y-axis; (middle) zoomed-in y-axis; (right) log scale y-axis. We show the mean over 10 repetitions.

We find that Kernel DMM-UCB and Kernel AMM-UCB have a similar computational cost to AY-GP-UCB and IGP-UCB, whereas Kernel CMM-UCB has a much greater computational cost.

F.2   2-Dimensional and 4-Dimensional Actions

We display the cumulative regret of each method in the synthetic kernel bandit problems described in Sec. 7. Here, the action set $\mathcal{X}_t$ in each round t consists of 100 random vectors in the 2-dimensional hypercube $[0, 1]^2$ (Fig. 5) or the 4-dimensional hypercube $[0, 1]^4$ (Fig. 6).
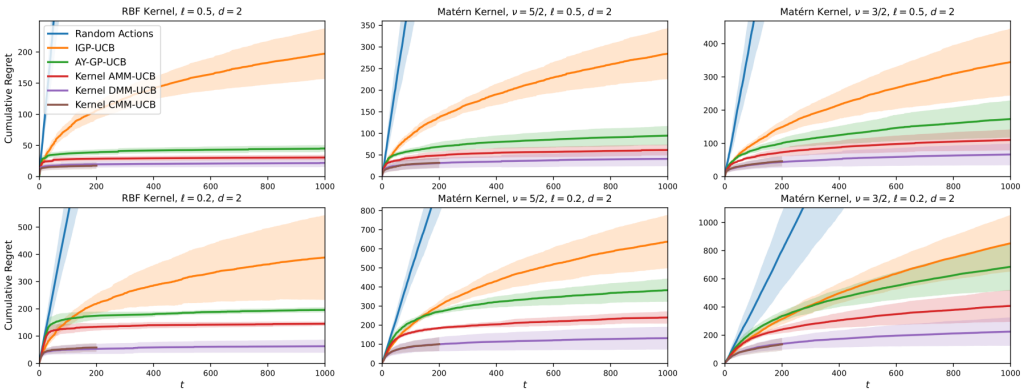

Figure 5: Cumulative regret of our KernelUCB-style algorithms as well as AY-GP-UCB, IGP-UCB and a random baseline over $T = 1000$ rounds for $d = 2$ and various kernels (columns) and length scales (rows). We show the mean $\pm$ standard deviation over 10 repetitions.

In Fig. 5, we observe that our Kernel CMM-UCB and Kernel DMM-UCB algorithms achieve the lowest cumulative regret, followed by Kernel AMM-UCB and the AY-GP-UCB and IGP-UCB. All methods performed noticeably better with 2-dimensional action sets than with 3-dimensional action sets (see Fig. 2).
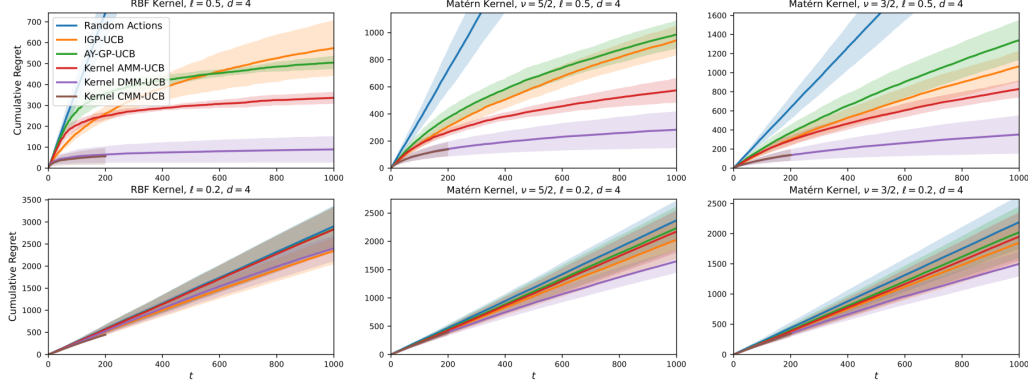
Figure 6: Cumulative regret of our KernelUCB-style algorithms as well as AY-GP-UCB, IGP-UCB and a random baseline over $T = 1000$ rounds for $d = 4$ and various kernels (columns) and length scales (rows). We show the mean $\pm$ standard deviation over 10 repetitions.

In Fig. 6, we observe a similar drop in the performance of all methods when moving up to 4-dimensional action sets. In the top row of Fig. 6, in which the kernel length-scale is $\ell = 0.5$, we observe that each method still achieves much lower cumulative regret than the random baseline, and Kernel CMM-UCB and Kernel DMM-UCB algorithms still achieve the lower cumulative regret. However, in bottom row of Fig. 6, we see that no method performs much better than the random baseline.

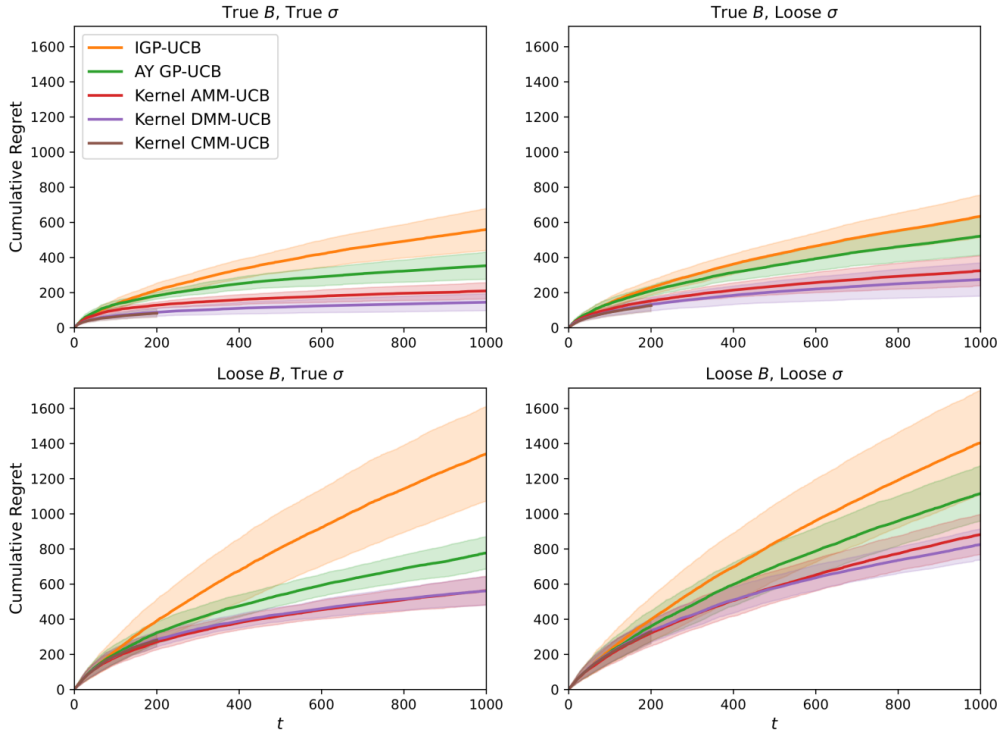### F.3 Loose Upper Bounds on B and σ

Figure 7: Cumulative regret of our kernel UCB-style algorithms as well as AY-GP-UCB, IGP-UCB and a random baseline after $T = 1000$ rounds for various kernels (columns) and length scales (rows). The action sets are subsets of $[0,1]^4$. We show the mean $\pm$ standard deviation over 10 repetitions.

To run each kernel bandit algorithm, we only need to know upper bounds on the sub-Gaussian parameter σ and the bound B on the RKHS norm of the reward function $f^\star$. We now investigate what happens when the known bounds on one or both of these quantities are loose.

We run each algorithm in the same synthetic kernel bandit problem as described in Sec. 7, where the kernel is a Matérn kernel with smoothness $\nu = 5/2$ and length-scale $\ell = 0.5$, and the actions are 3-dimensional vectors in $[0,1]^3$. As before, the noise variable are σ-sub-Gaussian with $\sigma = 0.1$ and $\|f^\star\|_{\mathcal{H}} = 10$. We run each algorithm with all four pairs of values of (σ, B), where $\sigma \in \{0.1, 0.2\}$ and $B \in \{10, 20\}$.

Fig. 7 shows the cumulative regret of each method and with each combination of correct or loose σ and B. We observe that using a loose upper bound on σ has a relatively small effect on each method, whereas a loose bound on B can cause the cumulative regret of each method to grow considerably.

F.4   Quasiconvexity

We plot the objective function in (12) for various values of α to verify that it is (sometimes) quasiconvex in α. We use the Matérn kernel with smoothness $\nu = 5/2$ and length-scale $\ell = 0.5$, and generate a reward function $f^\star$ with norm $\|f^\star\|_{\mathcal{H}} = 10$ following the procedure described in Sec. 7. We generate random data $\{(x_t, y_t)\}_{t=1}^{10}$, where $x_t \sim \mathcal{U}([-1,1])$ and $y_t \sim \mathcal{N}(f^\star(x_t), \sigma^2)$, with $\sigma = 0.1$. We plot the value of $\mu_{\alpha,t}(x) + \frac{1}{\sqrt{\alpha}}\widetilde{R}_{\alpha,t}\rho_{\alpha,t}(x)$ for α between 10^{-4} and 2. The test point x is another independent random draw from the distribution $\mathcal{U}([-1,1])$.
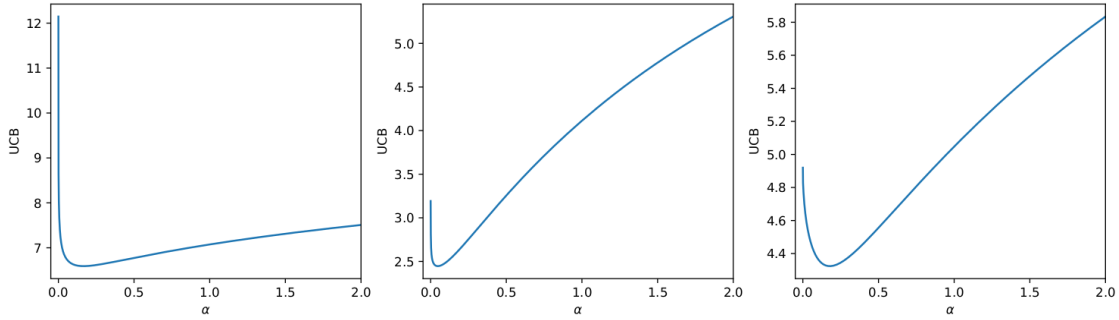

Figure 8: The value of the analytic UCB in (14) for α between 10^{-4} and 2. Each subplot shows the UCB for a different randomly drawn test function f^*, data set $\{(x_t, y_t)\}_{t=1}^{10}$ and test point x.

In Fig. 8, we plot the value of the analytic UCB, $\mu_{\alpha,t}(x) + \frac{1}{\sqrt{\alpha}}\widetilde{R}_{\alpha,t}\rho_{\alpha,t}(x)$, against α, where each subplot shows a different randomly drawn test function f^*, data set $\{(x_t, y_t)\}_{t=1}^{10}$ and test point x. In all three subplots, we observe that $\mu_{\alpha,t}(x) + \frac{1}{\sqrt{\alpha}}\widetilde{R}_{\alpha,t}\rho_{\alpha,t}(x)$ is a quasiconvex function of α.