Adaptation to Misspecified Kernel Regularity in Kernelised Bandits

Yusha Liu
Carnegie Mellon University
yushal@cs.cmu.edu

Aarti Singh
Carnegie Mellon University
aarti@cs.cmu.edu

Abstract

In continuum-armed bandit problems where the underlying function resides in a reproducing kernel Hilbert space (RKHS), namely, the kernelised bandit problems, an important open problem remains of how well learning algorithms can adapt if the regularity of the associated kernel function is unknown. In this work, we study adaptivity to the regularity of translation-invariant kernels, which is characterized by the decay rate of the Fourier transformation of the kernel, in the bandit setting. We derive an adaptivity lower bound, proving that it is impossible to simultaneously achieve optimal cumulative regret in a pair of RKHSs with different regularities. To verify the tightness of this lower bound, we show that an existing bandit model selection algorithm applied with minimax non-adaptive kernelised bandit algorithms matches the lower bound in dependence of T, the total number of steps, except for log factors. By filling in the regret bounds for adaptivity between RKHSs, we connect the statistical difficulty for adaptivity in continuum-armed bandits in three fundamental types of function spaces: RKHS, Sobolev space, and Hölder space.

1 Introduction

We consider the problem of continuum-armed bandit, a sequential decision-making problem, where the goal of a learning algorithm is the optimization of a black-box reward function, by selecting query points and eliciting rewards from the underlying function sequentially. The performance of algorithms is measured by the cumulative regret, which is the sum of differences between the maximum of the underlying function and the reward incurred by the learning algorithm across all the time steps. Optimizing cumulative regret requires from the learning algorithms a delicate exploration-exploitation tradeoff. The learning algorithm needs to simultaneously exploit high-reward regions and explore uncertain regions. The exploration-exploitation tradeoff is often dependent on complexity of the function space to which the reward function belongs. In most theoretical analyses of cumulative regret of algorithms, complexity of the function space is assumed to be known. Many studies use this assumption to design algorithms that achieve minimax optimal performance when the function space is known, for example, for linear functions (Dani et al., 2008; Abbasi-Yadkori et al., 2011), functions residing in reproducing kernel Hilbert spaces (RKHS) (Valko et al., 2013; Janz et al., 2020) or drawn from Gaussian Processes (Srinivas et al., 2009; Chowdhury and Gopalan, 2017), as well as neural networks functions (Zhou et al., 2020; Kassraie and Krause, 2021).

However, despite the theoretical convenience, it is not always realistic to assume access to the underlying function space. For this reason, some recent works in continuum-armed bandits have started to develop adaptive algorithms for when the function space is misspecified (see Section 2 for a summary of related works). The best possible performance of adaptive algorithms is equivalent to algorithms that know the parameter. An algorithm that simultaneously achieves minimax cumulative regret rates without access to the parameter is said to achieve minimax adaptivity. While minimax adaptivity is possible under the simple regret minimization setting, recent works have proved that it is not always achievable for cumulative regret minimzation (Locatelli and Carpentier, 2018), due to the exploration-exploitation dilemma.

When the reward function resides in an RKHS induced of some kernel function k, the problem also is referred to as kernelised bandit. In this work, we focus on an important and open problem in adaptivity in kernelised bandits, precisely, adaptivity to unknown kernel regularity. Recently, there has been a line of theoretical works that study adaptivity under the kernelised bandit setting, such as adaptivity to the length scale parameter and the RKHS norm (Berkenkamp et al., 2019) for a given kernel,

Proceedings of the 26th International Conference on Artificial Intelligence and Statistics (AISTATS) 2023, Valencia, Spain. PMLR: Volume 206.

and adaptivity to ϵ-misspecification, where the underlying function is ϵ-approximated by functions in an RKHS (Bogunovic and Krause, 2021). To the best of our knowledge, the work of Kassraie et al. (2022) is most closely related to our setting. They consider the setting where the underlying function lies in an RKHS but the kernel is unknown. Kassraie et al. (2022) assume that the kernel is a sparse combination of known base kernels and design algorithms with sublinear regret guarantees under this assumption. A more detailed discussion of the prior works on adaptivity in kernelised bandit is continued in Section 2.

Adaptivity of any algorithm with respect to the explicit regularity of the kernel function k, however, remains an unsolved problem. We characterize the regularity of k using a general notion: the decay rate of the Fourier transform of k (Section 3). In contrast to, for example, adapting to the RKHS norm which measures the smoothness of a function with respect to a fixed kernel, we adapt to the regularity of kernels which controls the differentiability of functions in the associated RKHSs. The kernel regularity thus determines the statistical complexity of the associated learning problem in a more fundamental way. In estimation, optimization (including simple regret minimization) (Bull, 2011) and cumulative regret minimization tasks (Srinivas et al., 2009; Kandasamy et al., 2019; Janz et al., 2020), the kernel regularity affects the minimax regret rate through exponential dependence on T, as opposed to the RKHS norm which only affects the rate polynomially. We focus on this fundamental problem of how well bandit algorithms can adapt to the unknown kernel regularity.

The contributions of this work are summarized as follows:

1. We derive the first lower bound on adaptivity to kernel regularity, expressed in terms of the kernel Fourier transformation decay rate, for kernelised bandit problems. This lower bound serves as an impossibility result, that no algorithms can simultaneously achieve minimax optimal performance in RKHSs with different regularities.
2. For RKHSs of the Matérn family (Matern et al., 1960) of kernels, we prove that CORRAL (Pacchiano et al., 2020b), an existing model selection algorithm, applied with (non-adaptive) minimax optimal kernelised bandit algorithms, matches the adaptivity lower bound[1] in the dependence on T. In contrast, another model selection algorithm RBBE Pacchiano et al. (2020a) does not match the lower bound.
3. By comparing the upper and lower bounds derived by this work to existing adaptivity results, we draw connections between the statistical difficulty of adaptivity in three types of function spaces: RKHSs, Sobolev spaces, and Hölder spaces.

[1]Except for log factors.

A summary of our results amongst existing results can be found in Table 1. Our main results (Section 4.2) are stated for more general kernels but in Table 1 only results with Matérn-ν (Definition 4) kernels are shown as an example, for clear comparisons. For adaptive results, the values $\tilde{\nu}$ and R are input parameters to the adaptive algorithms, such that they achieve (non-adaptive) minimax regret rates if the true parameter satisfies $\nu = \tilde{\nu}$ (for Matérn RKHS) or $\alpha = R$ (for Hölder spaces). We use $\tilde{O}$ to denote the asymptotic regret rate of T. $\tilde{O}$ omits dependence on other parameters such as the radius of the RKHS ball B (Section 3), any constant factors, and log factors of T unless otherwise specified.

Relationship with Neural Bandits. The kernelised bandit formulation has implications for optimization of more complex functions under the bandit setting as well, such as neural network functions. The Neural Tangent Kernel (NTK) literature (Jacot et al., 2018; Arora et al., 2019; Lee et al., 2019; Bietti and Bach, 2020; Chen and Xu, 2020) argue that over-parameterized neural networks can be approximated by functions in an RKHS of some composite kernel named the Neural Tangent Kernel, given that the network is sufficiently wide and the training is lazy (Chizat et al., 2019). Recent advances in this field establish interesting connections between the structure of a neural network and the regularity of its corresponding NTK. For example, Vakili et al. (2021a) consider wide fully-connected neural networks with activation functions with smoothness s. The show that the RKHS of the NTK of such a network is norm equivalent to, or embedded in, the RKHS of a Matérn-ν kernel with $\nu = s - \frac{1}{2}$. The value of ν dictates the differentiability of functions in the RKHS. Hence, the neural network functions considered in Vakili et al. (2021a) are approximated by functions in the RKHS of a Matérn-ν kernel.[2] These connections imply that adaptivity to the kernel regularity in kernelised bandits can potentially be extended to adaptivity to the structure of neural networks (such as smoothness of the activation functions considered in Vakili et al. (2021a)) in neural network bandits.

The rest of the paper is structured as follows. In Section 2, we discuss relevant prior works. In Section 3 we state the problem formulation. In Section 4 we present the main result of this paper, a lower bound on adaptivity to kernel regularity. In Secion 5 we discuss upper bounds of existing adaptive algorithms and whether they match the lower bound. In Section 6 we connect adaptivity to kernel regularity and adaptivity to Hölder exponents. Finally, we discuss the limitations and future directions of our work in Section 7.

[2]The result in Corollary 3 of Bietti and Bach (2020) can be thought as a special case of when $s = 1$, since the activation function considered is ReLU.

Table 1: Summary of Our Results and Comparison to Existing Results

<table>
<tr><td colspan="2" rowspan="2">Regret</td><td>RKHS of Matern-ν: $\mathcal{H}_{k,\nu}(\mathcal{X}) = \mathcal{W}^{\nu+\frac{d}{2}}(\mathcal{X})$</td><td>Hölder Space: $\Sigma^{\alpha}(\mathcal{X})$</td></tr>
<tr><td colspan="2">Relationship: $\mathcal{H}_{k,\nu}(\mathcal{X}) = \mathcal{W}^{\nu+\frac{d}{2}}(\mathcal{X}) \subset \Sigma^{\alpha=\nu}(\mathcal{X})$</td></tr>
<tr><td colspan="2">Non-adaptive minimax</td><td>$\tilde{\Theta}(T^{\frac{\nu+d}{2\nu+d}})$
Valko et al. (2013),Scarlett et al. (2017)</td><td>$\tilde{\Theta}\left(T^{\frac{d+\alpha}{d+2\alpha}}\right)$
Liu et al. (2021),Wang et al. (2018)</td></tr>
<tr><td rowspan="2">Adaptive
($d=1$)</td><td>Upper bound</td><td>$\tilde{O}\left(T^{\frac{1+2\tilde{\nu}+\tilde{\nu}\nu}{(1+2\tilde{\nu})(1+\nu)}}\right)$, for $\nu < \tilde{\nu}$
$\tilde{\nu}$: Input to adaptive algorithm.
This work (Theorem 7)</td><td>$\tilde{O}\left(T^{\frac{1+2R+R\alpha}{(1+2R)(1+\alpha)}}\right)$, for $\alpha < R$
R: Input to adaptive algorithm.
Liu et al. (2021, Theorem 8)</td></tr>
<tr><td>Lower bound</td><td>$\tilde{\Omega}\left(T^{\frac{1^2+2\tilde{\nu}+\tilde{\nu}\nu}{(1+2\tilde{\nu})(1+\nu)}}\right)$, for $\nu < \tilde{\nu}$
This work (Corollary 5)</td><td>$\tilde{\Omega}\left(T^{\frac{1^2+2R\alpha+R\alpha}{(1+2R)(1+\alpha)}}\right)$, for $\alpha < R \leq 1$
Locatelli and Carpentier (2018, Theorem 3)</td></tr>
</table>

2 Related Work

Kernelised Bandit In kernelised bandit problems, the reward function lies in a reproducing kernel Hilbert space (RKHS). This problem has been studied by many previous works, under the assumption that the kernel and other parameters (such as the upper bound on the function's RKHS norm) are known. Valko et al. (2013) take a frequentist approach and design a SuperKernelUCB algorithm, based on applying the kernel trick to the (Sup)LinREL and (Sup)LinUCB algorithms (Auer, 2002; Chu et al., 2011). The same technique is later used in extension to neural networks by Kassraie and Krause (2021), who propose SupNTKUCB which works with neural networks. SupKernelUCB achieves $\tilde{O}(\sqrt{T\gamma_T})$ regret where γ_T is the maximum information gain between T total observations and the underlying function. For common kernels such as the Matérn-ν kernels, this regret is minimax optimal in its dependence on T (except for log factors), by the lower bound provided later in Scarlett et al. (2017). However, SupKernelUCB relies on a batching technique that makes the algorithm performs poorly in practice (Calandriello et al., 2019; Janz et al., 2020). In the (parallel) Bayesian setting (the Gaussian Process bandit problem), the underlying function is assumed to be drawn from a GP. GP-UCB algorithm (Srinivas et al., 2009; Chowdhury and Gopalan, 2017; Janz et al., 2020) achieves the same regret bound as SupKernelUCB $\tilde{O}(\sqrt{T\gamma_T})$ in the GP setting but becomes suboptimal (sometimes with linear regret rate) in the RKHS setting with a $\tilde{O}(\gamma_T\sqrt{T})$ regret (Vakili et al., 2021b).

Adaptivity in Kernelised Bandit This problem we consider falls under the scope of model misspecification in bandit setting, which has been studied for linear functions and Hölder-smooth functions (Du et al., 2019; Foster et al., 2019; Lattimore et al., 2020; Zhu and Nowak, 2021; Locatelli and Carpentier, 2018; Liu et al., 2021). For Hölder functions, in particular, Locatelli and Carpentier (2018); Hadiji (2019) provide a lower bound indicating that it is impossible to achieve minimax adaptivity to the Hölder exponent. In this work, we convey a similar message with respect to the regularity of RKHS. For adaptivity in kernelised bandit problems, Berkenkamp et al. (2019) propose an algorithm with sublinear regret for when the lengthscale parameter (Definition 4) and upper bound on the RKHS norm (equation 4) are unknown. Neiswanger and Ramdas (2021) develop robust confidence sequence under the Bayesian framework to use in adaptive methods for GP optimization when the prior mean and/or covariance parameters are unknown. They conduct simulations for optimization on functions drawn from GPs but do not provide explicit regret analyses. Bogunovic and Krause (2021) develop methods for ϵ-misspecification, where the underlying function can be arbitrarily non-smooth, but is approximated by functions in a (known) RKHS with an ϵ-error in infinity norm. They prove a $\Omega(\epsilon T)$ lower bound for this setting and derived a matching upper bound. However, note that between two function spaces, the approximation error is a constant value and does not depend on T. Since a constant ϵ means an inevitable linear regret ($\Omega(\epsilon T)$), the ϵ-misspecification setting (Bogunovic and Krause, 2021) does not directly apply to adaptation to the kernel parameters. In the Meta-learning regime, Kassraie et al. (2022) consider RKHS with unknown kernels that are sparse combinations of known base kernels and proves that a Meta-learned kernel can yield sublinear regret. However, since the kernel is Meta-learned, it relies on offline tasks as training data. We do not assume the availability of offline data in the (fully online) bandit setting.

To summarize, prior works (to the best of our knowledge) only consider parameters that influence the regret rate in polynomial factors while our focus is on the regularity parameter which affects the rate in the exponent of T.

General Model Selection for Bandit Another line of recent works on model selection in bandit settings makes less stringent assumptions on the underlying function. These works consider algorithms based on a "corralling" mechanism, where a master algorithm "corrals" several base algorithms as arms and each base algorithm selects actions with different principles. The base algorithms usually assume different function spaces. Agarwal et al. (2017); Pacchiano et al. (2020b) propose an algorithm named CORRAL where the master algorithm is based on online mirror descent. In certain cases, CORRAL performs comparably to the best base algorithm running standalone.[3] Pacchiano et al. (2020a) propose the Regret Bound Balancing and Elimination (RBBE) which uses a (simpler) stochastic master algorithm and an additional base-algorithm-elimination step. We refer readers to Section 5 for details about these two methods and their performance in our problem setting.

3 Problem Setting

Problem Formulation Consider the problem of zeroth-order black-box optimization under bandit feedback. The learner interacts with a stochastic environment in a sequential manner. This problem is also formulated as stochastic continuum-armed bandit. At time step $t \in \{1, \dots, T\}$, the learner chooses an action x_t from the compact domain $\mathcal{X} = [0,1]^d$, and receives a reward y_t. The reward is a noisy observation of the underlying reward function $f : \mathcal{X} \to \mathbb{R}$:

$$y_t = f(x_t) + \eta_t, \tag{1}$$

where the noise variable η_t follows a zero-mean sub-Gaussian distribution (see Theorem 3). The optimization objective is the cumulative (pseudo-)regret defined as follows.

$$R_T = \sum_{t=1}^{T} f(x^*) - f(x_t), \tag{2}$$

where x^* is the global maximizer of f, unknown to the learner. Results in this paper are expressed in expected cumulative (pseudo-)regret $\mathbb{E}[R_T]$, where the expectation is taken over the randomness of $\{x_t\}_{t=1\dots T}$.

Kernelised Bandit We consider the setting where f is square-integrable and resides in an RKHS $\mathcal{H}_k$ of a symmetric, positive-definite kernel $k : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$. The RKHS is unique given the kernel (Wainwright, 2019, Theorem 12.11). We denote the RKHS of k on domain $\mathcal{X}$ as $\mathcal{H}_k(\mathcal{X})$. In this work, we restrict our attention to *translation-invariant* kernels, precisely, kernels that satisfy the following: $k(x,x') = \kappa(x - x')$, for some function $\kappa : \mathbb{R}^d \to \mathbb{R}$. For a translation-invariant kernel, the regularity of functions in the RKHS is captured by the Fourier transform of the kernel. Precisely, we have the following definition when the domain is $\mathbb{R}^d$. Let $\hat{g}(\omega)$ denote the Fourier transformation (Wendland, 2004; Williams and Rasmussen, 2006) of a function g as $\forall \omega \in \mathbb{R}^d$.

$$\mathcal{H}_k(\mathbb{R}^d) = \{f \in \mathcal{L}_2(\mathbb{R}^d) \cap C(\mathbb{R}^d) : \\ \|f\|_{\mathcal{H}_k} := (2\pi)^{-d/2} \int_{\mathbb{R}^d} \frac{|\hat{f}(\omega)|^2}{\hat{\kappa}(\omega)} d\omega < \infty\}. \tag{3}$$

When the domain $\mathcal{X}$ is a subset of $\mathbb{R}^d$, $\hat{\kappa}$ still captures the regularity of $\mathcal{H}_k(\mathcal{X})$, via a norm equivalency result that holds as long as $\mathcal{X}$ has a Lipschitz boundary. Details can be found in Section 4.1, Lemma 1. We write $\|f\|_k \stackrel{\triangle}{=} \|f\|_{\mathcal{H}_k(\mathcal{X})}$ for simplicity. We apply the common assumption (Srinivas et al., 2009; Valko et al., 2013) that the RKHS norm of f is upper bounded by a value $B, 0 < B < \infty$:

$$f \in \mathcal{H}_k(\mathcal{X}, B) := \{f : f \in \mathcal{H}_k, \|f\|_k \le B\}. \tag{4}$$

We refer to $\mathcal{H}_k(\mathcal{X}, B)$ as a ball in the RKHS with radius B.

4 Main Result: Adaptivity Lower Bound

In this section, we present the main result, a lower bound on adaptivity to the regularity of kernel (Theorem 3). The regularity of a translation-invariant kernel is expressed as the decay rate of its Fourier transformation (equation 9). We next instantiate this idea with a norm equivalency result between an RKHS and a Sobolev space. The norm equivalency result is dependent on the kernel Fourier decay rate. The proof of Theorem 3, in turn, relies on this norm equivalency as well.

4.1 Norm Equivalency Between RKHS and Sobolev Space

Consider integer-order Sobolev space $\mathcal{W}^{m,p}(\mathcal{X})$ where m, p are integers greater or equal to 1. We define the following notions for a multi-index vector $\alpha = (\alpha_1 \dots \alpha_d)$: $|\alpha| = \alpha_1 + \cdots + \alpha_d$, $\alpha! = \alpha_1! \dots \alpha_d!$ and $x^\alpha = x_1^{\alpha_1} \dots x_d^{\alpha_d}$. Let $D^{(\alpha)} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}}$ denote the multivariate mixed partial weak derivative. The Sobolev space and corresponding Sobolev norm ($\|\cdot\|_{m,p,\mathcal{X}}$) are defined as follows.

$$\mathcal{W}^{m,p}(\mathcal{X}) = \{f \in \mathcal{L}_p(\mathcal{X}) : D^{(\alpha)} f \in \mathcal{L}_p(\mathbb{R}^d), \forall |\alpha| \le m\}, \tag{5}$$

$$\|f\|_{m,p,\mathcal{X}} := \left(\sum_{|\alpha| \le m} \int |D^{(\alpha)} f(x)|^p dx \right)^{\frac{1}{p}}. \tag{6}$$

We refer to m as the order of the Sobolev space. Furthermore, define the j-th order seminorm (Adams and Fournier, 2003, Definition 4.11) $|\cdot|_{j,p,\mathcal{X}}$ with integer $j \le m$, which

[3] Arora et al. (2021) also study the problem of corralling bandit algorithms in the stochastic setting, but only finite-armed case is considered.

is the sum of $\mathcal{L}_p$ norm of its j-th weak derivatives.

$$|f|_{j,p,\mathcal{X}} = \left(\sum_{|\alpha|=j} \int |D^{(\alpha)} f(x)|^p dx \right)^{\frac{1}{p}}. \quad (7)$$

In correspondence to the RKHS ball (equation 4), we define a Sobolev ball with radius L as the set of functions whose m-th order *seminorm* are upper bounded by L.

$$\mathcal{W}^{m,p}(\mathcal{X}, L) = \{f \in \mathcal{W}^{m,p}(\mathcal{X}) : |f|_{m,p,\mathcal{X}} \leq L\}. \quad (8)$$

When $p = 2$, the Sobolev space is equivalent to the RKHS of a translation-invariant kernel k. This connection plays an important role in the analysis. We consider only Sobolev spaces with $p = 2$, and hence abbreviate $\mathcal{W}^m(\mathcal{X}) \triangleq \mathcal{W}^{m,2}(\mathcal{X})$. The precise norm equivalency is introduced in the following lemma.

Lemma 1. *Wendland (2004, Corollary 10.48) Let $k : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be a translation-invariant kernel function such that $k(\cdot, \cdot) = \kappa(\cdot - \cdot)$ for $\kappa \in \mathcal{L}_1(\mathbb{R}^d)$. Suppose $\Omega \in \mathbb{R}^d$ is a domain with Lipschitz boundary. Suppose $\hat{\kappa}$ has the following polynomial decay rate of s, for $s > d/2, s \in \mathbb{N}$,*

$$c_1(1 + \|\omega\|_2^2)^{-s} \leq \hat{k}(\omega) \leq c_2(1 + \|\omega\|_2^2)^{-s}, \forall w \in \mathbb{R}^d, \quad (9)$$

for some constants $0 < c_1 \leq c_2$. Then, the associated RKHS $\mathcal{H}_k(\Omega)$ is norm equivalent to the Sobolev space $W^m(\Omega)$ with $m = s$.

Having established the equivalency between RKHS and Sobolev spaces, we further introduce some notions to quantify the relationship between Sobolev seminorm (which is the radius of Sobolev balls) and RKHS norm in the following lemma.

Lemma 2. *Suppose that m is a positive integer larger than $d/2$. Let Ω be a finite-width domain with Lipschitz boundary. Let $\mathcal{W}_0^{m,p}(\Omega)$ denote the closure of $C_0^\infty(\Omega)$ (set of functions that have compact support in Ω and, together with their infinite order of partial derivatives, are continuous) in $\mathcal{W}^{m,p}(\Omega)$ Adams and Fournier (2003). Then, the m-th Sobolev seminorm of f can be bounded by its RKHS norm with respect to a translation-invariant kernel k with Fourier decay rate m. Precisely,*

$$\underline{c}|f|_{m,2} \leq \|f\|_{\mathcal{H}_k} \leq \bar{c}|f|_{m,2}, \quad (10)$$

for some constants $0 < \underline{c} < \bar{c}$.

The constants $\underline{c}, \bar{c}$ are used globally in this work and appear in the lower bound in Section 4.2. The proof of Lemma 2 can be found in Appendix A.1.

4.2 Lower Bound on Adaptivity to Kernel Regularity

Theorem 3 presents our lower bound for adapting between a pair of RKHSs of different (kernel) regularities. An intuitive interpretation of the theorem is as follows. Consider a nested pair of balls in two RKHSs. Suppose both kernels satisfy the conditions in Lemma 1 but with different Fourier decay rates: $m_1 \in \mathbb{N}$ and $m_2 \in \mathbb{N}$ such that $0 < m_1 < m_2$. If an algorithm that is oblivious to the true regularity value somehow achieves a small (for example, minimax optimal) regret on all functions inside the (smoother) RKHS ball with parameter m_2, this algorithm will suffer a price of larger (suboptimal) regret on at least one function inside the (rougher) RKSH ball with parameter m_1. For the lower bound analysis, we consider $d = 1$ and leave the extension of the lower bound to $d > 1$ as a future direction (Section 7).

Theorem 3. *Consider the problem setting in Section 3 with noises $\{\eta_t\}_{t=1...T}$ that are $\frac{1}{4}$-subgaussian. Let $\tilde{R}$ be a positive number, let m_1, m_2 be two positive integers that satisfy $m_1 < m_2$. There exist two positive values B_1 and B_2, such that the following statement is true. Consider an algorithm that achieves in the RKHS of a kernel k_{m_2} with Fourier decay rate m_2 the following regret upper bound.*

$$\sup_{f \in \mathcal{H}_{k_{m_2}}(\mathcal{X}, B_2)} \mathbb{E}[R_T] \leq \tilde{R}. \quad (11)$$

Then, the regret of this algorithm in a (less smooth) RKHS of another kernel k_{m_1} with Fourier decay rate m_1 is lower bounded by the following. Suppose that functions in the function spaces have bounded $\mathcal{L}_2$ norm.[4]

$$\sup_{f \in \mathcal{H}_{k_{m_1}}(\mathcal{X}, B_1)} \mathbb{E}[R_T] \geq \quad (12)$$

$$\frac{1}{8} \left(\frac{C(m_1)}{32} \right)^{\frac{m_1 - 1/2}{m_1 + 1/2}} \left(\frac{B_1}{\bar{c}} \right)^{\frac{1}{m_1 + 1/2}} \tilde{R}^{-\frac{m_1 - 1/2}{m_1 + 1/2}} T.$$

Here, $C(m_1)$ denotes a constant that depends on m_1.

It is worth noting that, although the lower bound has a factor of T, the regret is not necessarily linear in T, because $\tilde{R}$ also depends on T and in fact usually ranges from $\tilde{O}(\sqrt{T})$ to $\tilde{O}(T)$. The full version of this theorem is presented as Theorem 9 in Appendix A.2, where we state the full constraints on the radius values B_1 and B_2. Since B_1 and B_2 are only upper bounds on the RKHS norm and *not* the kernel regularity that we focus on, we present only the concise version here to show the adaptivity difficulty with respect to regularity parameters m_1 and m_2.

4.2.1 Proof Sketch

The proof of Theorem 3 consists of two key parts. The first part is constructing the hypothesis functions, in which we borrow ideas from lower bounds in regression problems (Tsybakov, 2004). The second part is lower bounding the cumulative regret, given the constructed hypothesis functions, where we follow Hadiji (2019, Section 2.2). Intuitively, the second part shows that if any player achieves a

[4] Functions in Sobolev spaces and RKHSs are square-integrable.

small regret on all the smoother functions, then it inevitably incurs large regret on the rougher functions in the space, because of its disproportionally small amount of exploration. The method in Hadiji (2019) is itself an improved version of the adaptivity lower bound for Hölder spaces proposed in Locatelli and Carpentier (2018).

4.2.2 A Sobolev Version of the Lower Bound

It is convenient to construct functions with compact support and finite Sobolev semi-norms from an infinitely-differentiable base function, such as the bump function Tsybakov (2004). On the other hand, directly constructing functions with finite RKHS norms (Scarlett et al., 2017, Section III.A) involves inverse Fourier transformation of the bump function and thus leads to wavelet-like functions with non-compact support. Therefore, it is more natural for us to first consider functions in (integer-order) Sobolev spaces as hypothesis functions, and then use the norm equivalency result between Sobolev spaces and RKHSs to prove the lower bound. More precisely, the hypothesis functions constructed in the proof reside in a Sobolev ball $\mathcal{W}^m(\mathcal{X}, L)$, for some (integer) order m and radius L. Via the norm equivalency (Lemma 2), those functions also resides in a RKHS ball of a kernel with Fourier decay rate m.

As a result, there is a Sobolev version of the adaptivity lower bound. Informally, let m_1, m_2 be two positive integers such that $m_2 > m_1$. If an algorithm achieves a $\tilde{R}$ regret upper bound in the smoother Sobolev space $\mathcal{W}^{m_2}(\mathcal{X})$, then its regret over functions in $\mathcal{W}^{m_1}(\mathcal{X})$ is lower bounded by $\Omega(\tilde{R}^{-\frac{m_1-1/2}{m_1+1/2}}\ T)$. We formally state the Sobolev version of the adaptivity lower bound in Theorem 11 in Appendix B.1. The two lower bounds share the same proof structure, connected via the norm equivalency in Lemma 1.

4.2.3 Impossibility Result for Matérn Kernels

For the Matérn-ν family of kernels (Matern et al., 1960), an implication of Theorem 3 is that no algorithm can achieve minimax adaptivity between two RKHSs if they have different regularity. Therefore, we also refer to this lower bound as an impossibility result for adaptivity to the kernel regularity. We formally define Matérn-ν family of kernels in Definition 4.

Definition 4. The Matérn-ν kernel and its Fourier transformation are defined as follows for dimension d.

$$k_{\text{Matérn},\nu}(x, x') \tag{13}$$

$$= \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}\|x - x'\|_2}{l} \right)^{\nu} J_\nu\left(\frac{\sqrt{2\nu}\|x - x'\|_2}{l}\right), \tag{14}$$

$$\hat{k}_{\text{Matérn},\nu}(\omega) = c_1\left(\frac{2\nu}{l^2} + \|\omega\|_2^2\right)^{-(\nu+\frac{d}{2})}. \tag{15}$$

where $c_1 = \frac{2^d \pi^{d/2}\Gamma(\nu+d/2)(2\nu)^\nu}{\Gamma(\nu)l^{2\nu}}$, J_ν is the modified Bessel function of the second kind, l is the length-scale, and $\nu > 0$ is the regularity parameter. In this work, we assume for simplicity that the length-scale is set to $\propto \sqrt{2\nu}$.

The Fourier transformation of a Matérn kernel with regularity parameter ν decays with a rate of $\nu + \frac{d}{2}$ (equation equation 15). Therefore, we can instantiate the impossibility result for Matérn kernels. The result is presented in Corollary 5. Precisely, for $0 < \nu_1 < \nu_2$, if an adaptive algorithm achieves minimax regret rate on a Matérn RKHS with regularity ν_2, then it has a strictly suboptimal regret rate on the RKHS with ν_1.

Corollary 5. *Suppose the problem is the same as defined in Theorem 3. Let ν_1, ν_2 be real numbers that satisfy $0 < \nu_1 < \nu_2$ and $\nu_1 + \frac{1}{2} \in \mathbb{N}, \nu_2 + \frac{1}{2} \in \mathbb{N}$. There exist two positive values B_1, B_2, such that the following statement is true. Suppose an algorithm oblivious to the true regularity parameter value achieves the following minimax optimal regret* [5] *on $\mathcal{H}_{k_{\text{Matérn},\nu_2}}(\mathcal{X}, B_2)$,*

$$\sup_{f\in\mathcal{H}_{k_{\text{Matérn},\nu_2}}(\mathcal{X},B_2)} \mathbb{E}[R_T] = \tilde{O}\left(T^{\frac{\nu_2+1}{2\nu_2+1}}\right), \tag{16}$$

then the regret of this algorithm on RKHS $\mathcal{H}_{k_{\text{Matérn},\nu_1}}(\mathcal{X}, B_1)$ is lower bounded by the following.

$$\sup_{f\in\mathcal{H}_{k_{\text{Matérn},\nu_1}}(\mathcal{X},B_1)} \mathbb{E}[R_T] = \tilde{\Omega}\left(T^{\frac{\nu_1\nu_2+2\nu_2+1}{(\nu_1+1)(2\nu_2+1)}}\right), \tag{17}$$

The proof of Corollary 5 is an application of Theorem 3 and can be found in Appendix B.2. The cumulative regret rate in 17 is suboptimal compared to the minimax rate which is $\tilde{O}(T^{\frac{\nu_1+1}{2\nu_1+1}})$ (see Section 5.1 for non-adaptive minimax rates). Therefore, Theorem 3 is an impossibility result for adaptivity to kernel regularity with Matérn kernels.

5 Upper Bounds of Adaptive Algorithms

We consider two adaptive algorithms particularly: CORRAL from Agarwal et al. (2017); Pacchiano et al. (2020b) and Regret Bound Balancing and Elimination (RBBE) from (Pacchiano et al., 2020a). The two algorithms (i) can be applied to the problem of adaptation to kernel regularity and (ii) have explicit regret guarantees in this setting.

The adaptive algorithms, however, need base algorithms that are non-adaptive minimax optimal. We first provide an overview of such non-adaptive algorithms for kernelised bandit in Section 5.1. Then, we derive adaptivity upper bounds of CORRAL and RBBE in Section 5.2 and Section 5.3 respectively. For concreteness, we only consider RKHS of Matérn-ν kernel (Definition 4) in this section. To

[5] Omitting the dependence on the upper bound on RKHS norm.

match the lower bound, we set $d = 1$. Comparison of the upper bounds to the lower bound (Theorem 3), shows that CORRAL (coupled with minimax optimal base algorithms) can match the lower bound in dependence on T between certain pairs of values for ν.

5.1 Overview: Non-adaptive Minimax Algorithms

We discuss the theoretical performance of algorithms developed for kernelised bandits in Section 5.1.1. We show that a recent algorithm that is designed for continuum-armed bandit in Hölder spaces (Liu et al., 2021) is also optimal over functions in RKHS of Matérn kernels in Section 5.1.2.

5.1.1 SupKernelUCB and GP-UCB for RKHS

Recall that the lower bound (in terms of T) on cumulative regret for kernelised bandit with Matérn-ν kernels $k_{\text{Matérn},\nu}$ is $\Omega(T^{\frac{\nu+1}{2\nu+1}})$, as proved by Scarlett et al. (2017). There are mainly two types of algorithms applicable for the kernelised bandit problem: (i) GP-UCB (Srinivas et al., 2009) and its variants (Chowdhury and Gopalan, 2017; Janz et al., 2020), and (ii) KernelUCB and its *Sup*-variant Valko et al. (2013). The GP-UCB-style algorithms display a non-trivial empirical advantage over the impractical SupKernelUCB. That being said, GP-UCB is suboptimal theoretical upper bounds for certain types of kernels under the RKHS assumption, including for Matérn-ν kernels. In the RKHS of a Matérn kernel $k_{\text{Matérn},\nu}$, GP-UCB achieves a regret of $\tilde{O}(T^{\frac{\nu+\frac{3}{2}}{2\nu+1}})$.[6]. On the other hand, SupKernelUCB matches the lower bound with a regret rate of $\tilde{O}(T^{\frac{\nu+1}{2\nu+1}})$.[7]

5.1.2 UCB-Meta for Hölder Space

Apart from the kernelised bandit algorithms discussed above, Liu et al. (2021) propose an algorithm for continuum-armed bandits in Hölder space with exponent $\alpha > 1$ with regret upper bound that matches existing lower bounds (Wang et al., 2018; Singh, 2021) except log factors. This algorithm is named UCB-Meta. We show in Theorem 6 that UCB-Meta is naturally minimax optimal in dependence on T over the RKHS of certain kernels.

Theorem 6. *Consider the kernelised bandit problem where $f \in \mathcal{H}_{k_{\text{Matérn},\nu}}(\mathcal{X}, B)$, where $\nu > 0$ and $\nu + \frac{1}{2} \in \mathbb{N}$. Then, UCB-Meta achieves the following regret upper bound,*

$$\sup_{f \in \mathcal{H}_{k_{\text{Matérn},\nu}}(\mathcal{X},B)} \mathbb{E}[R_T] = \tilde{O}\left(T^{\frac{\nu+1}{2\nu+1}}\right), \tag{18}$$

where $\tilde{O}$ omits dependence on radius of the RKHS ball B, constant factors depending on ν, and $\log$ factors of T.

The regret rate shown in Theorem 6 is derived from the result that $\mathcal{H}_{k_{\text{Matérn},\nu}}(\mathcal{X})$ is embedded in a Hölder space $\Sigma^{\alpha}(\mathcal{X})$ with $\alpha = \nu$. The proof can be found in Appendix B.3. Singh (2021) have shown a similar argument while focusing mainly on the connection between Besov and Hölder spaces.

5.2 CORRAL as Adaptive Algorithm

The original CORRAL algorithm for model selection in the bandit setting is first proposed by Agarwal et al. (2017). The original CORRAL requires that modifications be made to each base algorithm for them to satisfy a stability condition (Definition 3 in Agarwal et al. (2017)). These modifications, however, have to be made on a case-by-case basis. Therefore, we use the smoothed version of CORRAL which is proposed by Pacchiano et al. (2020b). The smoothed CORRAL puts a smoothing operation between the master algorithm and base algorithms and thus does not require modifications be made to the base algorithms. Smoothed CORRAL operates only with stochastic environments, which is satisfied by our assumptions (Section 3). For simplicity, we refer to the smoothed version of CORRAL as CORRAL. CORRAL uses an adversarial online mirror descent algorithm as the master algorithm.

Recall that a non-adaptive minimax kernelised bandit algorithm achieves $\tilde{O}(T^{\frac{\nu+1}{2\nu+1}})$ regret (Section 5.1), if instantiated with the correct parameter ν. By plugging in the regret of base kernelised bandit algorithms in the general result in Theorem 5.3 in Pacchiano et al. (2020b), we derive a adaptive upper bound for CORRAL in Theorem 7. CORRAL achieves sublinear $\tilde{o}(T)$ regret on all possible values of ν^* (See Theorem 7). Oppositely, a non-adaptive algorithm instantiated with parameter value $\tilde{\nu}$ does not have sublinear regret guarantees if the true parameter $\nu^* < \tilde{\nu}$, because the underlying function space $\mathcal{H}_{k_{\text{Matérn},\nu^*}}$ is not contained in algorithm's hypothesis space. In Theorem 7, $\tilde{\nu} \in \boldsymbol{u}$ is a parameter that is specified by the user and can be interpreted as the parameter that specifies the space on which the algorithm is configured to achieve minimax regret.

Theorem 7. *Consider the kernelised bandit problem where $f \in \mathcal{H}_{k_{\text{Matérn},\nu^*}}(\mathcal{X}, B^*)$, $\nu^* + \frac{1}{2} \in \mathbb{N}$ and ν^*, B^* unknown to the learner. Let $\boldsymbol{u} = \{(\nu_1, B_1), (\nu_2, B_2), \ldots, (\nu_M, B_M)\}$ be a list of candidate input value pairs such that $\boldsymbol{u}$ specifies a nested set of RKHS: $\mathcal{H}_{k_{\text{Matérn},\nu_1}}(\mathcal{X}, B_1) \subset \mathcal{H}_{k_{\text{Matérn},\nu_2}}(\mathcal{X}, B_2) \subset \ldots \mathcal{H}_{k_{\text{Matérn},\nu_M}}(\mathcal{X}, B_M)$. Suppose that $(\nu^*, B^*) \in \boldsymbol{u}$. Let $\mathbb{A} = \{\mathcal{A}_i, i \in [M]\}$ be a set of (non-adaptive) minimax optimal kernelised bandit algorithms with anytime regret guarantees, each instantiated with the regularity and radius $(\nu_i, B_i) \in \boldsymbol{u}$. The regret from running CORRAL with input total time steps T and learning rate $\eta = \tilde{O}(T^{-\frac{1+\tilde{\nu}}{1+2\tilde{\nu}}})$ applied with base algorithms from $\mathbb{A}$*

[6] The suboptimality of GP-UCB is discussed more extensively in Vakili et al. (2021b)

[7] The analysis of SupKernelUCB was originally for finite-armed setting, but Cai and Scarlett (2021, Appendix A.4) state that it can be extended to the continuum-armed setting where $\mathcal{X} = [0, 1]^d$, suffering only a $O(d \log(T))$ term in the regret.

is as follows.[8]

$$\sup_{f\in\mathcal{H}_{k_{\text{Matérn},\nu^*}}} \mathbb{E}[R_T] = \tilde{O}\left(T^{\max(\frac{1+\tilde{\nu}}{1+2\tilde{\nu}},\frac{1^2+2\tilde{\nu}+\tilde{\nu}\nu^*}{(1+2\tilde{\nu})(1+\nu^*)})}\right). \tag{19}$$

The proof of Theorem 7 can be found in Appendix B.4. This result indicates that CORRAL achieves (i) minimax optimal rate $\tilde{O}\left(T^{\frac{1+\nu^*}{1+2\nu^*}}\right)$ in terms of T, if the underlying kernel regularity $\nu^* = \tilde{\nu}$; (ii) suboptimal rate $\tilde{O}\left(T^{\frac{1+\tilde{\nu}}{1+2\tilde{\nu}}}\right)$ if $\nu^* > \tilde{\nu}$ and (iii) suboptimal rate $\tilde{O}\left(T^{\frac{1+2\tilde{\nu}+\tilde{\nu}\nu^*}{(1+2\tilde{\nu})(1+\nu^*)}}\right)$ when $\nu^* < \tilde{\nu}$. Let ν_1^*, ν_2^* satisfying $\nu_1^* < \nu_2^*$ be two possible values of the true regularity that both satisfy the assumptions in Theorem 7. Suppose $\nu_1^* \leq \tilde{\nu} < \nu_2*$. By Theorem 7, CORRAL achieves regret $\tilde{O}\left(T^{\frac{1^2+2\tilde{\nu}+\tilde{\nu}\nu_1^*}{(1+2\tilde{\nu})(1+\nu_1^*)}}\right)$ if the true parameter is ν_1^* and $\tilde{O}\left(T^{\frac{1+\tilde{\nu}}{1+2\tilde{\nu}}}\right)$ if the true parameter is ν_2^*. By Theorem 3, the lower bound over the rougher RKHS with ν_1^* is $\tilde{\Omega}\left(T^{\frac{1+2\tilde{\nu}+\tilde{\nu}\nu_1^*}{(1+2\tilde{\nu})(1+\nu_1^*)}}\right)$. The lower bound is matched by the upper bound in the exponent of T.

In conclusion, CORRAL matches the adaptivity lower bound in the dependence on T except log factors, between any pair of regularity values (ν_1^*, ν_2^*), such that $\nu_1^* \geq \tilde{\nu}, \nu_1^* + \frac{1}{2} \in \mathbb{N}$ and $\nu_2^* < \tilde{\nu}, \nu_2^* + \frac{1}{2} \in \mathbb{N}$.

Finally, note that in this subsection, the assumption is that the true parameter(s) are contained in the candidate set $\boldsymbol{u}$. Hence, Theorem 7 reflects the cost of adaptation (model selection), which is the difficulty of selecting the best base learner out of all candidates. If, however, the true parameter is not contained in $\boldsymbol{u}$, then adaptive algorithms will incur another type of cost, namely the cost of "discretization". This cost is generated from the difference between the true parameter and the closest value in $\boldsymbol{u}$. Using an exponential (Pacchiano et al., 2020b) or linear (Liu et al., 2021) grid for $\boldsymbol{u}$ can usually incur a small cost of "discretization".

5.3 RBBE as Adaptive Algorithm

The regret bound balancing and elimination (RBBE) algorithm proposed in Pacchiano et al. (2020a) achieves near-optimal regret in several adaptivity problems with linear function spaces. RBBE can be thought of as using a stochastic master algorithm that selects the base algorithm with the smallest candidate cumulative regret at each time. Therefore, it enjoys advantages such as gap-dependent regret bounds and high probability regret bounds. Unlike CORRAL, it does not need a user-specified parameter to control the space over which the algorithm will achieve minimax optimal regret on. Instead, the algorithm achieves simultaneously on all possible values of ν^* the regret upper bound of $\tilde{O}(T^{\frac{1+4\nu^*+2\nu^{*2}}{1+4\nu^*+4\nu^{*2}}})$. If we plug this upper bound in Theorem 3 for $\nu^* = \nu_2^*$, then a lower bound of $\tilde{\Omega}(T^{\frac{(2\nu_2^*+1)^2+2\nu_1^*\nu_2^{*2}}{(2\nu_2^{*2}+1)^2(\nu_1^*+1)}})$ is incurred for when $\nu^* = \nu_1^*$, given that $0 < \nu_1^* < \nu_2^*$. The upper bound of RBBE is larger than the lower bound in the exponent of T. A more detailed description of the RBBE algorithm and a formal statement of its adaptivity upper bound can be found in Appendix A.3 and Theorem 10 therein.

To summarize, although both CORRAL and RBBE as adaptive algorithms can achieve sublinear regret simultaneously on different kernel regularity, CORRAL has a better theoretical adaptivity in this problem. While RBBE fails to match the lower bound, CORRAL achieves the adaptivity lower bound for certain pairs of ν values for Matérn-ν kernels. [9]

6 Connection with Adaptivity to Hölder Exponents

The adaptivity lower bound in Theorem 3 specifies the difficulty of adapting between two RKHSs of kernels with polynomial Fourier decay rate m_1 and m_2, where $0 < m_1 < m_2, m_1 \in \mathbb{N}, m_2 \in \mathbb{N}$. Recall that $\tilde{R}$ is the regret upper bound on the smoother RKHS with parameter m_2. The lower bound on the RKHS specified by m_1 depends inversely on $\tilde{R}$ through an $\Omega(T \cdot \tilde{R}^{-\frac{m_1-1/2}{m_1+1/2}})$ dependence.

Shifting the perspective from RKHS to Hölder spaces, the adaptivity difficulty has been studied by Locatelli and Carpentier (2018); Hadiji (2019), for a subset of values for the Hölder exponent α. Precisely, Theorem 3 in Locatelli and Carpentier (2018) provides an $\Omega(T \cdot \tilde{R}^{-\frac{\alpha_1}{\alpha_1+1}})$ dependence as the lower bound, for adapting between two Hölder spaces with exponents α_1, α_2 satisfying $\alpha_1 < \alpha_2 \leq 1$.[10] Here, $\tilde{R}$ is the upper regret bound on the smoother Hölder space $\Sigma^{\alpha_2}(\mathcal{X})$. We know by Lemma 1 that an RKHS $\mathcal{H}_{k_{m_1}}(\mathcal{X})$ with kernel Fourier decay rate m_1 is norm equivalent to Sobolev space $\mathcal{W}^{m_1}(\mathcal{X})$. Coupled with the Sobolev embedding theorem for integer-order Sobolev spaces (Adams and Fournier, 2003, Theorem 5.4), it is straightforward to see that $\mathcal{H}_{k_{m_1}}(\mathcal{X}) \subset \Sigma^{\alpha}(\mathcal{X})$, where $\alpha = m_1 - \frac{1}{2}$ (Appendix B.3).

Note that we have the following equivalence between the lower bounds if $\alpha_1 = m_1 - \frac{1}{2}$.

$$T\tilde{R}^{-\frac{m_1-1/2}{m_1+1/2}} \propto T\tilde{R}^{-\frac{\alpha_1}{\alpha_1+1}}. \tag{20}$$

[8] $\tilde{O}$ omits dependence on radius of the RKHS ball B, constant factors depending on ν, and log factors of T.

[9] It is our conjecture that the stochastic master used by RBBE (as opposed to the adversarial one in CORRAL) limits its model selection ability in certain cases.

[10] Proving the adaptivity rate for when the exponents are larger than 1 remains an open problem.

Therefore, for continuum-armed bandit problems, the statistical difficulty of adapting to kernel regularity of RKHS is the same as adapting to Hölder exponents, if the Hölder exponents represent the smallest Hölder spaces that the RKHSs embed in.

7 Discussion

We discuss several future directions, stemming from the current limitations of our work. Our current theoretical results are for the domain with $d = 1$,[11] so it is of interest to extend the current results to $d > 1$. Instead of partitioning the domain $\mathcal{X} = [0, 1]$ into M sub-intervals, one needs to partition the hypercube $[0, 1]^d$ into M sub-cubes and construct the hypothesis functions with appropriate Fourier decay correspondingly. Such an extension is possible akin to Scarlett et al. (2017).

Another direction is to derive adaptivity upper bounds in terms of Fourier decay as well and verify the tightness of the lower bound in more cases than Matérn kernels. Since we currently investigate translation-invariant kernels, a more long-term direction is the investigation of adaptivity to rotation-invariant kernels, to connect to NTKs which are usually rotation-invariant dot-product kernels (Bietti and Bach, 2020; Chen and Xu, 2020; Vakili et al., 2021a). Finally, this study is of theoretical nature, so it remains an open problem to empirically study adaptivity to kernel regularity, based on the insights provided by our lower and upper bounds.

Acknowledgements

This work is supported by NSF CCF1763734 and Simons Foundation grant.

References

Abbasi-Yadkori, Y., Pál, D., and Szepesvári, C. (2011). Improved algorithms for linear stochastic bandits. *Advances in neural information processing systems*, 24:2312–2320.

Adams, R. A. and Fournier, J. J. (2003). *Sobolev spaces*. Elsevier.

Agarwal, A., Luo, H., Neyshabur, B., and Schapire, R. E. (2017). Corralling a band of bandit algorithms. In *Conference on Learning Theory*, pages 12–38. PMLR.

Arora, R., Marinov, T. V., and Mohri, M. (2021). Corralling stochastic bandit algorithms. In *International Conference on Artificial Intelligence and Statistics*, pages 2116–2124. PMLR.

Arora, S., Du, S. S., Hu, W., Li, Z., Salakhutdinov, R. R., and Wang, R. (2019). On exact computation with an infinitely wide neural net. *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.

Auer, P., Cesa-Bianchi, N., Freund, Y., and Schapire, R. E. (1995). Gambling in a rigged casino: The adversarial multi-armed bandit problem. In *Proceedings of IEEE 36th annual foundations of computer science*, pages 322–331. IEEE.

Auer, P., Ortner, R., and Szepesvári, C. (2007). Improved rates for the stochastic continuum-armed bandit problem. In *International Conference on Computational Learning Theory*, pages 454–468. Springer.

Berkenkamp, F., Schoellig, A. P., and Krause, A. (2019). No-regret bayesian optimization with unknown hyperparameters. *arXiv preprint arXiv:1901.03357*.

Bietti, A. and Bach, F. (2020). Deep equals shallow for relu networks in kernel regimes. *arXiv preprint arXiv:2009.14397*.

Bogunovic, I. and Krause, A. (2021). Misspecified gaussian process bandit optimization. *Advances in Neural Information Processing Systems*, 34:3004–3015.

Bull, A. D. (2011). Convergence rates of efficient global optimization algorithms. *Journal of Machine Learning Research*, 12(10).

Cai, X. and Scarlett, J. (2021). On lower bounds for standard and robust gaussian process bandit optimization. In *International Conference on Machine Learning*, pages 1216–1226. PMLR.

Calandriello, D., Carratino, L., Lazaric, A., Valko, M., and Rosasco, L. (2019). Gaussian process optimization with adaptive sketching: Scalable and no regret. In *Conference on Learning Theory*, pages 533–557. PMLR.

Chen, L. and Xu, S. (2020). Deep neural tangent kernel and laplace kernel have the same rkhs. *arXiv preprint arXiv:2009.10683*.

Chizat, L., Oyallon, E., and Bach, F. (2019). On lazy training in differentiable programming. *Advances in Neural Information Processing Systems*, 32.

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

Chu, W., Li, L., Reyzin, L., and Schapire, R. (2011). Contextual bandits with linear payoff functions. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pages 208–214. JMLR Workshop and Conference Proceedings.

[11] Note that this is not equivalent to the dimension of the feature map of a kernel.

Dani, V., Hayes, T. P., and Kakade, S. M. (2008). Stochastic linear optimization under bandit feedback. *Conference on Learning Theory*.

Du, S. S., Kakade, S. M., Wang, R., and Yang, L. F. (2019). Is a good representation sufficient for sample efficient reinforcement learning? *arXiv preprint arXiv:1910.03016*.

Foster, D. J., Krishnamurthy, A., and Luo, H. (2019). Model selection for contextual bandits. *arXiv preprint arXiv:1906.00531*.

Hadiji, H. (2019). Polynomial cost of adaptation for x-armed bandits. *Advances in Neural Information Processing Systems*, 32.

Jacot, A., Gabriel, F., and Hongler, C. (2018). Neural tangent kernel: Convergence and generalization in neural networks. *Advances in neural information processing systems*, 31.

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.

Kandasamy, K., Neiswanger, W., Zhang, R., Krishnamurthy, A., Schneider, J., and Poczos, B. (2019). Myopic posterior sampling for adaptive goal oriented design of experiments. In *International Conference on Machine Learning*, pages 3222–3232. PMLR.

Kassraie, P. and Krause, A. (2021). Neural contextual bandits without regret. *arXiv preprint arXiv:2107.03144*.

Kassraie, P., Rothfuss, J., and Krause, A. (2022). Meta-learning hypothesis spaces for sequential decision-making. *arXiv preprint arXiv:2202.00602*.

Lattimore, T., Szepesvari, C., and Weisz, G. (2020). Learning with good feature representations in bandits and in rl with a generative model. In *International Conference on Machine Learning*, pages 5662–5670. PMLR.

Lee, J., Xiao, L., Schoenholz, S., Bahri, Y., Novak, R., Sohl-Dickstein, J., and Pennington, J. (2019). Wide neural networks of any depth evolve as linear models under gradient descent. *Advances in neural information processing systems*, 32.

Liu, Y., Wang, Y., and Singh, A. (2021). Smooth bandit optimization: Generalization to holder space. In *International Conference on Artificial Intelligence and Statistics*, pages 2206–2214. PMLR.

Locatelli, A. and Carpentier, A. (2018). Adaptivity to smoothness in x-armed bandits. In *Conference on Learning Theory*, pages 1463–1492. PMLR.

Matern, B. et al. (1960). Spatial variation. stochastic models and their application to some problems in forest surveys and other sampling investigations. *Meddelanden fran Statens Skogsforskningsinstitut*, 49(5).

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

Pacchiano, A., Dann, C., Gentile, C., and Bartlett, P. (2020a). Regret bound balancing and elimination for model selection in bandits and rl. *arXiv preprint arXiv:2012.13045*.

Pacchiano, A., Phan, M., Abbasi-Yadkori, Y., Rao, A., Zimmert, J., Lattimore, T., and Szepesvari, C. (2020b). Model selection in contextual stochastic bandit problems. *arXiv preprint arXiv:2003.01704*.

Scarlett, J., Bogunovic, I., and Cevher, V. (2017). Lower bounds on regret for noisy gaussian process bandit optimization. In *Conference on Learning Theory*, pages 1723–1742. PMLR.

Shekhar, S. and Javidi, T. (2020). Multi-scale zero-order optimization of smooth functions in an rkhs. *arXiv preprint arXiv:2005.04832*.

Singh, S. (2021). Continuum-armed bandits: A function space perspective. In *International Conference on Artificial Intelligence and Statistics*, pages 2620–2628. PMLR.

Srinivas, N., Krause, A., Kakade, S. M., and Seeger, M. (2009). Gaussian process optimization in the bandit setting: No regret and experimental design. *arXiv preprint arXiv:0912.3995*.

Tsybakov, A. B. (2004). Introduction to nonparametric estimation, 2009. *URL https://doi. org/10.1007/b13794. Revised and extended from the*, 9(10).

Vakili, S., Bromberg, M., Garcia, J., Shiu, D.-s., and Bernacchia, A. (2021a). Uniform generalization bounds for overparameterized neural networks. *arXiv preprint arXiv:2109.06099*.

Vakili, S., Scarlett, J., and Javidi, T. (2021b). Open problem: Tight online confidence intervals for rkhs elements. In *Conference on Learning Theory*, pages 4647–4652. PMLR.

Valko, M., Korda, N., Munos, R., Flaounas, I., and Cristianini, N. (2013). Finite-time analysis of kernelised contextual bandits. *arXiv preprint arXiv:1309.6869*.

Wainwright, M. J. (2019). *High-dimensional statistics: A non-asymptotic viewpoint*, volume 48. Cambridge University Press.

Wang, Y., Balakrishnan, S., and Singh, A. (2018). Optimization of smooth functions with noisy observations: Local minimax rates. *Advances in Neural Information Processing Systems*, 31.

Wendland, H. (2004). *Scattered data approximation*, volume 17. Cambridge university press.

Williams, C. K. and Rasmussen, C. E. (2006). *Gaussian processes for machine learning*. MIT press Cambridge, MA.

Zhou, D., Li, L., and Gu, Q. (2020). Neural contextual bandits with ucb-based exploration. In *International Conference on Machine Learning*, pages 11492–11502. PMLR.

Zhu, Y. and Nowak, R. (2021). Pareto optimal model selection in linear bandits. *arXiv preprint arXiv:2102.06593*.

A AUXILIARY

A.1 Proof of Norm Equivalency Between RKHS Norm and Sobolev Seminorm

Proof of Lemma 2. It is shown by Wendland (2004, Theorem 10.12, Corollary 10.48) that: if a translation-invariant kernel k with Fourier decay rate s (Lemma 1), then the associated RKHS $\mathcal{H}_k$ defined on a Lipschitz domain Ω is norm equivalent to the Sobolev space $\mathcal{W}^{m=s,2}(\Omega)$. The norm equivalency indicates that there exist two constants c_1, c_2, $0 < c_1 < c_2$, such that for $f \in \mathcal{H}_k(\Omega)$, the following statement holds.

$$c_1\|f\|_{m,2,\mathcal{X}} \leq \|f\|_{\mathcal{H}_k} \leq c_2\|f\|_{m,2,\mathcal{X}}. \tag{21}$$

Now, we examine conditions under which the norm equivalency can be extended between the *seminorm* (equation 7) of Sobolev spaces and the RKHS norm. As in Lemma 2, let $W_0^{m,p}(\mathcal{X})$ denote the closure of $C_0^\infty(\mathcal{X})$ in $W^{m,p}(\mathcal{X})$.[12] Adams and Fournier (2003, 6.26) give the following result: if $\mathcal{X}$ has finite width, then for $f \in W_0^{m,p}$, the seminorm $|\cdot|_{m,p}$ is equivalent to the standard norm $\|\cdot\|_{m,p}$. The one-dimensional interval domain $\mathcal{X}$ we consider trivially satisfies the Lipschitz boundary condition, hence we have the following result.

Lemma 8. *If a function lies in $W_0^{m,p}(\mathcal{X})$ where $\mathcal{X} = [0,1]$, then there exists a constant $K < \infty$, such that*

$$|\cdot|_{m,p,\mathcal{X}} \leq \|\cdot\|_{m,p,\mathcal{X}} \leq K|\cdot|_{m,p,\mathcal{X}}. \tag{22}$$

Combining Lemma 8 with the norm equivalency in equation 21, we recover the inequalities in Lemma 2.

$$c_1|f|_{m,2,\mathcal{X}} \leq \|f\|_{\mathcal{H}_k} \leq Kc_2|f|_{m,2,\mathcal{X}}. \tag{23}$$

□

A.2 The Full Statement of Theorem 3

We present the full version of Theorem 3, which fully states the constraints on the radius B_1 and B_2 in Theorem 3. The proof is deferred to Appendix B.1.

Theorem 9. *Consider the bandit problem setting (Section 3) with noises $\{\eta_t\}_{t=1\ldots T}$ that are $\frac{1}{4}$-subgaussian. Further assume that the $\mathcal{L}_2$ norm of functions f we consider is upper bounded by finite value $\gamma_0 < \infty$: $\|f\|_2 \leq \gamma_0$. Let $\tilde{R}$ be a positive number, let $m_2 > m_1 > 0$ be two positive integers, and let B_1, B_2 be two positive variables that satisfy the following conditions.*

$$\begin{aligned}&\bar{c}\max\left\{\frac{3^{m_1+\frac{1}{2}}}{32}C(m_1)^{-m_1+\frac{1}{2}}\tilde{R}^{-1}, K(m_1,m_2,\gamma_0,\mathcal{X})\bar{c}^{\frac{m_2-m_1}{m_2}}B_2^{\frac{m_1}{m_2}}\right\}\\ &\qquad\leq B_1 \leq C'(m_1,m_2)^{-(m_1+\frac{1}{2})}\bar{c}^{(-m_1+\frac{1}{2})}B_2^{m_1+\frac{1}{2}}\tilde{R}^{m_1-\frac{1}{2}}\end{aligned} \tag{24}$$

where $C(m_1)$ and $C'(m_1,m_2)$ are constants whose exact forms are defined in equation 58 and equation 64 in the proof. $K(m_1,m_2,\gamma_0,\mathcal{X})$ is a constant depending on m_1, m_2, the domain and γ_0.[13]

Consider any algorithm that achieves in RKHS ball $\mathcal{H}_{k_{m_2}}(\mathcal{X}, B_2)$ the following regret upper bound, where the kernel k_{m_2} has Fourier decay rate m_2.

$$\sup_{f\in\mathcal{H}_{k_{m_2}}(\mathcal{X},B_2)} \mathbb{E}[R_T] \leq \tilde{R}, \tag{25}$$

then, the regret of this algorithm in a (less smooth) RKHS ball induced by another kernel k_{m_1} with Fourier decay rate m_1 is lower bounded by the following.

$$\sup_{f\in\mathcal{H}_{k_{m_1}}(\mathcal{X},B_1)} \mathbb{E}[R_T] \geq \frac{1}{8}\left(\frac{C(m_1)}{32}\right)^{\frac{m_1-1/2}{m_1+1/2}}\left(\frac{B_1}{\bar{c}}\right)^{\frac{1}{m_1+1/2}}\tilde{R}^{-\frac{m_1-1/2}{m_1+1/2}}T. \tag{26}$$

[12] Here, we borrow the definitions from Adams and Fournier (2003).

[13] The exact value of $K(m_2,\gamma_0,\mathcal{X})$ is deferred to the proof of Theorem 4.14 in Adams and Fournier (2003).

A.3 Adaptivity Upper Bound of RBBE

At each round, RBBE (Pacchiano et al., 2020a) first performs an elimination step to remove misspecified base algorithms, then selects a base algorithm among the remaining ones. The elimination step tests whether each base algorithm is well-specified, that is, whether each base algorithm's hypothesis space contains the underlying function. If a base algorithm fails the test, then it is eliminated. In the selection step, the master algorithm simply chooses the base algorithm with the smallest presumed cumulative pseudo-regret. Therefore, RBBE can be thought of as using a stochastic master algorithm (remarked in Pacchiano et al. (2020a) as well), instead of using an adversarial one as CORRAL (Agarwal et al., 2017; Pacchiano et al., 2020b) does.

The general regret of RBBE is stated in terms of the play ratio, which is the ratio between the number of times a base algorithm is played and the number of times that the best base algorithm is played. To instantiate the play ratio, Pacchiano et al. (2020b) considers only the setting where the regret rates of all base algorithms (if well-specified) have the same exponents on T. That is, the regret rates are T^β with a *fixed* $\beta \in (0, 1]$ across all base algorithms. However, this setting does not align with our setting where, for base algorithm i with input value ν_i, the exponent of T in its (well-specified) regret bound is $\frac{\nu_i+1}{2\nu_i+1}$. Hence, we make changes to the proof in Pacchiano et al. (2020b) to apply it to our problem setting. The result of RBBE is stated in Theorem 10 and the proof is deferred to Appendix B.5.

Theorem 10. *Suppose that the problem setting, the set of candidate values $\boldsymbol{u}$ and the set of base algorithms $\mathbb{A}$ are the same as defined in Theorem 7. The regret of RBBE applied with base algorithms in $\mathbb{A}$ is as follows, with high probability $1 - \delta$.*

$$\sup_{f \in \mathcal{H}_{k_{\text{Matérn},\nu^*}}} R_T = \tilde{O}(T^{\frac{1+4\nu^*+2\nu^{*2}}{1+4\nu^*+4\nu^{*2}}}). \tag{27}$$

B PROOFS OF RESULTS

B.1 Proof of Theorem 9

As explained in Section 4.2.1, the proof of Theorem 9 arises from the proof of a parallel Sobolev version of the adaptivity lower bound. We formally state the Sobolev version of adaptivity lower bound below.

Theorem 11. *Consider the bandit problem setting (Section 3) with noises $\{\eta_t\}_{t=1\ldots T}$ that are $\frac{1}{4}$-subgaussian. Further assume that the $\mathcal{L}_2$ norms of functions f we consider are upper bounded by the finite value $\gamma_0 < \infty$: $\|f\|_2 \le \gamma_0$.* [14] *Let $\tilde{R}$ be a positive number, let $m_2 > m_1 > 0$ be two positive integers, and let L_1, L_2 be two positive variables that satisfy the following conditions:*

$$\max\left\{\frac{3^{m_1+\frac{1}{2}}}{32} C(m_1)^{-m_1+\frac{1}{2}} \tilde{R}^{-1}, K(m_1, m_2, \gamma_0, \mathcal{X}) L_2^{\frac{m_1}{m_2}}\right\} \tag{28}$$
$$\le L_1 \le C'(m_1, m_2)^{-(m_1+\frac{1}{2})} L_2^{m_1+\frac{1}{2}} \tilde{R}^{m_1-\frac{1}{2}}$$

where $C(m_1)$, $C'(m_1, m_2)$ are constants whose exact forms are defined in equation 58 and equation 64 respectively. $K(m_1, m_2, \gamma_0, \mathcal{X})$ is a constant depending on m_1, m_2, the domain and γ_0, the upper bound on the $\mathcal{L}_2$ norm of functions in the Sobolev ball.[15] *Consider an algorithm that achieves in the Sobolev ball $\mathcal{W}^{m_2}(\mathcal{X}, L_2)$ a regret upper bound of $\tilde{R}$.*

$$\sup_{f \in \mathcal{W}^{m_2,2}(\mathcal{X}, L_2)} \mathbb{E}[R_T] \le \tilde{R}, \tag{29}$$

then, the regret of this algorithm in the less-smooth Sobolev ball $\mathcal{W}^{m_1}(\mathcal{X}, L_1)$ is lower bounded by the following.

$$\sup_{f \in \mathcal{W}^{m_1}(\mathcal{X}, L_1)} \mathbb{E}[R_T] \ge \frac{1}{8} \left(\frac{C(m_1)}{32}\right)^{\frac{m_1-1/2}{m_1+1/2}} L_1^{\frac{1}{m_1+1/2}} \tilde{R}^{-\frac{m_1-1/2}{m_1+1/2}} T. \tag{30}$$

In the next part, we present the proof of Theorem 11, which also leads to Theorem 9. The values B_1, B_2 in Theorem 3 should be set as follows.

$$B_1 = \bar{c} L_1, B_2 = \bar{c} L_2, \tag{31}$$

[14]By our assumption on the underlying function f in equation 5, we know that it has bounded $\mathcal{L}_2$ norm.

[15]The exact value of $K(m_2, \gamma_0, \mathcal{X})$ is deferred to the proof of Theorem 4.14 in Adams and Fournier (2003).

where $\bar{c}$ is the global constant in Lemma 2.

Proof. Consider the Sobolev version of the theorem (Theorem 11). Recall that the adaptivity is between balls in two different spaces, the "rougher" space $\mathcal{W}^{m_1}(\mathcal{X}, L_1)$ and the "smoother" space $\mathcal{W}^{m_2}(\mathcal{X}, L_2)$. First, we consider the constraints between L_1 and L_2 such that $\mathcal{W}^{m_2}(\mathcal{X}, L_2) \subset \mathcal{W}^{m_1}(\mathcal{X}, L_1)$. In other words, $f \in \mathcal{W}^{m_2}(\mathcal{X}, L_2)$ should be sufficient condition for $f \in \mathcal{W}^{m_1}(\mathcal{X}, L_1)$. Theorem 4.14 in Adams and Fournier (2003) and references therein give the following interpolation upper bound between orders of smoothness for a function $f \in \mathcal{W}^{m_1}(\mathcal{X})$,

$$|f|_{m_1,2} \leq K(m_2, \mathcal{X})(|f|_{m_2})^{\frac{m_1}{m_2}} \, \|f\|_2^{\frac{m_2-m_1}{m_2}}, \tag{32}$$

where $K(m_2, \mathcal{X})$ is a constant depending only on m_2 and the domain $\mathcal{X}$. If $f \in \mathcal{W}^{m_2}(\mathcal{X}, L_2)$, then by definition (equation 8) we know that $|f|_{m_2} \leq L_2$. Using equation 32, we have that:

$$|f|_{m_1,2} \leq K(m_2, \mathcal{X}) L_2^{\frac{m_1}{m_2}} \, \|f\|_2^{\frac{m_2-m_1}{m_2}} \tag{33}$$

To ensure that the two Sobolev balls are nested, L_1 should be larger than the right-hand side of the above inequality. The $\mathcal{L}_2$ norm of f is upper bounded by $\|f\|_2 \leq \gamma_0$. Plugging it in equation 33 incurs an lower bound for L_1:

$$L_1 \geq K(m_1, m_2, \gamma_0, \mathcal{X}) L_2^{\frac{m_1}{m_2}} = K(m_2, \mathcal{X}) \gamma_0^{\frac{m_2-m_1}{m_2}} L_2^{\frac{m_1}{m_2}}.$$

Having established $\mathcal{W}^{m_2}(\mathcal{X}, L_2) \subset \mathcal{W}^{m_1}(\mathcal{X}, L_1)$, we start with the formal proof of the adaptivity lower bound.

Function Construction Part I. This part is adapted from the regression lower bounds in Tsybakov (2004, Section 2.6). Let M be a positive integer parameter, which is the number of hypothesis functions we need. The value of M remains to be determined later in the proof. In the following, we shall assume $M \geq 2$ and eventually prove that this assumption holds. Further, define bandwidth $h = \frac{1}{2M}$. Let $\Delta > 0$ be a parameter that represents the maximum of the M hypothesis functions in $\mathcal{W}^{m_1,2}(\mathcal{X}, L_1)$. The value of Δ remains to be determined later in the proof same as M.

Partition the 1-dimensional domain $\mathcal{X} = [0, 1]$ into $M + 1$ bins: $H_{0 \ldots M}$, such that $\cup_{s=0 \ldots M} H_s = \mathcal{X}$. Define the bins and their middle points $\bar{x}_{0,\ldots M}$ as follows.

$$\begin{aligned} H_s &= \left[\frac{s-1}{2M}, \frac{s}{2M}\right], \ \bar{x}_s = \frac{s - \frac{1}{2}}{2M}, \text{ for } s = 1 \ldots M, \\ H_0 &= \left[\frac{1}{2}, 1\right], \ \bar{x}_0 = \frac{3}{4}. \end{aligned}$$

We use the bump function as a base function, then we shift the base function to construct the hypothesis functions. The bump function is defined as follows. It has compact support on $(-1, 1)$. Function $K_0(\cdot)$ is infinitely differentiable with continuous derivatives (Tsybakov, 2004, (2.34)).

$$K_0(x) = \exp\left(\frac{-1}{1 - x^2}\right) \mathbb{I}(|x| < 1). \tag{34}$$

Next, define $M + 1$ functions as follows, each one has support inside one of the $M + 1$ bins.

$$f_s = a h^{m_1 - \frac{1}{2}} K\left(\frac{x - \bar{x}_s}{h}\right), \ \ s = 1 \ldots M, \tag{35}$$

$$f_0 = \tilde{a} h^{m_2 - \frac{1}{2}} \tilde{K}\left(\frac{x - \bar{x}_0}{h}\right), \tag{36}$$

where

$$K(u) = K_0(bu), \tag{37}$$

$$\tilde{K}(u) = K_0(\tilde{b}u). \tag{38}$$

$a, b, \tilde{a}, \tilde{b}$ are non-negative parameters to be defined later. We require that $b \geq 2$ and $\tilde{b} \geq 4h$, so that the support of every function f_s is inside H_s, $\forall s \leq M$. Lemma 13 ensures that the requirements on $b, \tilde{b}$ hold, by posing constraints between Δ and M.

We introduce the following lemma to specify requirements on the variables $a, b, \tilde{a}, \tilde{b}$, with respect to Δ and L_1, L_2. This is to make sure that values of $a, b, \tilde{a}, \tilde{b}$ guarantee that $f_s \in \mathcal{W}^{m_1}(\mathcal{X}, L_1), \ \forall 1 \leq s \leq M$ and $f_0 \in \mathcal{W}^{m_2}(\mathcal{X}, L_2)$.

Lemma 12. *Let K_0^* to denote the maximum value of $K_0(\cdot)$, a constant less than 1. Let I_{m_1}, I_{m_2} denote the $\mathcal{L}_2$ norms of the m_1, m_2-th order derivatives of $K_0(\cdot)$, respectively. That is, $I_{m_1} = \int_{-1}^{1}\left[K_0^{(m_1)}(u)\right]^2 du$ and $I_{m_2} = \int_{-1}^{1}\left[K_0^{(m_2)}(u)\right]^2 du$. Then, if Δ is the maximum of f_s in $\mathcal{W}^{m_1,2}(\mathcal{X}, L_1)$, for all $s = 1 \dots M$ and $\Delta/2$ is the maximum of f_0 in $\mathcal{W}^{m_2,2}(\mathcal{X}, L_2)$, the function parameters $a, b, \tilde{a}, \tilde{b}$ satisfy the following:*

$$a = \Delta(2M)^{m_1-\frac{1}{2}}/K_0^* \tag{39}$$

$$\tilde{a} = \Delta(2M)^{m_2-\frac{1}{2}}/2K_0^* \tag{40}$$

$$b \leq \left(\frac{L_1^2 K_0^{*2}}{\Delta^2(2M)^{2m_1-1}I_{m_1}}\right)^{\frac{1}{2m_1-1}} \tag{41}$$

$$\tilde{b} \leq \left(\frac{4L_2^2 K_0^{*2}}{\Delta^2(2M)^{2m_2-1}I_{m_2}}\right)^{\frac{1}{2m_2-1}}, \tag{42}$$

Proof of Lemma 12. The constraints on $a, \tilde{a}$ follows trivially from the requirement that $f_s^* = \Delta$ for $s = 1 \dots M$, $f_0^* = \Delta/2$, and plugging in $h = 1/2M$.

The constraints on $b, \tilde{b}$ are to ensure that

$$\|f_s^{(m_1)}\|_2 \leq L_1, \quad s = 1 \dots M$$
$$\|f_0^{(m_2)}\|_2 \leq L_2$$

We first consider requirement for $\|f_s^{(m_1)}\|_2 \leq L_1, \quad s = 1 \dots M$. For $s \geq 1$,

$$\begin{aligned}
&\|f_s^{(m_1)}\|_2^2 \\
&= \int_0^1 \left[f^{(m_1)}(x)\right]^2 dx \\
&= \int_0^1 \left[a h^{m_1-\frac{1}{2}} \frac{\partial^{m_1}}{\partial x^{m_1}}\left(K(\frac{x-\bar{x}_s}{h})\right)\right]^2 dx \\
&= a^2 h^{2m_1-1} \int_0^1 \left[\frac{\partial^{m_1}}{\partial x^{m_1}}\left(K_0\left(\frac{b}{h}(x-\bar{x}_s)\right)\right)\right]^2 dx \\
&= a^2 h^{2m_1-1} \int_0^1 \left[(\frac{b}{h})^{m_1} K_0^{(m_1)}\left(\frac{b}{h}(x-\bar{x}_s)\right)\right]^2 dx \\
&\overset{u=\frac{b}{h}(x-\bar{x}_s)}{=} a^2 h^{-1} b^{2m_1} \int_{\frac{b}{h}(-\bar{x}_s)}^{\frac{b}{h}(1-\bar{x}_s)} \left[K_0^{(m_1)}(u)\right]^2 \frac{h}{b} du \\
&= a^2 b^{2m_1-1} \int_{-1}^{1} \left[K_0^{(m_1)}(u)\right]^2 du = a^2 b^{2m_1-1} I_{m_1}.
\end{aligned}$$

The second to last step follows because the bump function K_0 has compact support on $(-1, 1)$ and the upper and lower limits of the integral satisfy:

$$\frac{b}{h}(1-\bar{x}_s) = b(\frac{1}{h} - s + \frac{1}{2}) > 1,$$
$$\frac{b}{h}(-\bar{x}_s) = -b(s - \frac{1}{2}) \leq -1.$$

Therefore, for $\|f_s^{(m_1)}\|_2^2 \leq L_1^2$ to hold, we need $a^2 b^{2m_1-1} I_{m_1} \leq L_1^2$. This leads to

$$b \leq \left(\frac{L_1^2}{a^2 I_{m_1}}\right)^{\frac{1}{2m_1-1}} \tag{43}$$

$$= \left(\frac{L_1^2 (K_0^*)^2}{\Delta^2(2M)^{2m_1-1}I_{m_1}}\right)^{\frac{1}{2m_1-1}}. \tag{44}$$

Similarly, for $s = 0$, we have the following.

$$\begin{aligned}
&\|f_s^{(m_2)}\|_2^2 \\
&= \int_0^1 \left[f^{(m_2)}(x)\right]^2 dx \\
&= \int_0^1 \left[\tilde{a}h^{m_2-\frac{1}{2}} \frac{\partial^{m_2}}{\partial x^{m_2}} \left(\tilde{K}(\frac{x-\bar{x}_0}{h})\right)\right]^2 dx \\
&= \int_0^1 \tilde{a}^2 h^{2m_2-1} \left[\frac{\partial^{m_2}}{\partial x^{m_2}} \left(K_0(\frac{\tilde{b}(x-\bar{x}_0}{h}))\right)\right]^2 dx \\
&= \tilde{a}^2 h^{2m_2-1} \int_0^1 \left[\left(\frac{\tilde{b}}{h}\right)^{m_2} K_0^{(m_2)}(\frac{\tilde{b}}{h}(x-\bar{x}_0))\right]^2 dx \\
&\overset{u=\tilde{b}(x-\bar{x_0})/h}{=} \tilde{a}^2 \tilde{b}^{2m_2-1} \int_{-\frac{3\tilde{b}}{4h}}^{\frac{\tilde{b}}{h}(1-\frac{3}{4})} \left[K_0^{(m_2)}(u)\right]^2 du \\
&= \tilde{a}^2 \tilde{b}^{2m_2-1} \int_{-1}^1 \left[K_0^{(m_2)}(u)\right]^2 du \\
&= \tilde{a}^2 \tilde{b}^{2m_2-1} I_{m_2}.
\end{aligned}$$

Note that in the third last equation, the integral upper and lower limit satisfy:

$$\frac{\tilde{b}}{h}(1-\frac{3}{4}) > 1, \quad -\frac{3\tilde{b}}{4h} < -1.$$

For the above $\|f_s^{(m_2)}\|_2^2$ to be less or equal to L_2^2, we need:

$$\tilde{b} \leq \left(\frac{L_2^2}{\tilde{a}^2 I_{m_2}}\right)^{\frac{1}{2m_2-1}} = \left(\frac{4L_2^2 (K_0^*)^2}{\Delta^2 (2M)^{2m_2-1} I_{m_2}}\right)^{\frac{1}{2m_2-1}} \tag{45}$$

□

Combining Lemma 12 with what we required of the function parameters: $b \geq 2$ and $\tilde{b} \geq 4h$, we then need the following requirements for the parameter Δ. Intuitively, the following lemma says that the functions cannot be too "wavy", so that they stay within the corresponding balls in Sobolev spaces.

Lemma 13. *For $b \geq 2, \tilde{b} \geq 4h$ to hold, Δ needs to satisfy the following constraints with respect to M and the smoothness constants L_1, L_2.*

$$\Delta/L_1 \leq \frac{K_0^*}{2^{2m_1-1} M^{m_1-\frac{1}{2}} \sqrt{I_{m_1}}}, \tag{46}$$

$$\Delta/L_2 \leq \frac{K_0^*}{2^{2m_2-2} \sqrt{I_{m_2}}}. \tag{47}$$

Proof of Lemma 13. First, consider function f_s when $s \geq 1$. Using the conclusions in Lemma 12 we need the following,

$$\frac{L_1^2 (K_0^*)^2}{\Delta^2 (2M)^{2m_1-1} I_{m_1}} \geq b^{2m_1-1} \geq 2^{2m_1-1}.$$

What directly follows is the constraint on Δ:

$$\Delta^2 \leq \frac{L_1^2 (K_0^*)^2}{2^{4m_1-2} M^{2m_1-1} I_{m_1}}. \tag{48}$$

Similarly, for f_0, we need

$$\frac{L_2^2(K_0^*)^2}{\Delta^2(2M)^{2m_2-1}I_{m_2}} \geq \tilde{b}^{2m_2-1} \geq (4h)^{2m_2-1} = 2^{2m_2-1}M^{1-2m_2}.$$

This leads to second constraint on Δ:

$$\Delta^2 \leq \frac{L_2^2(K_0^*)^2}{2^{4m_2-4}I_{m_2}}. \tag{49}$$

□

Function Construction Part II. We have defined $f_0 \dots f_M$ in Part I, and identified the constraints between the floating parameters M and Δ, with respect to given parameters m_1, m_2, L_1, L_2 and known constants K_0^*, I_{m_1}, I_{m_2}. In this second part, we define $M+1$ bandit problems by defining their reward functions $\phi_s,\ s = 0 \dots M$ in the following way:

$$\phi_0 = f_0, \tag{50}$$

$$\phi_s = f_s + f_0,\ \forall 1 \leq s \leq M. \tag{51}$$

It is obvious that the reward functions satisfy the following conditions. The conditions below are the Sobolev version. They are necessary for the latter half of this proof. Similar conditions were required in Locatelli and Carpentier (2018); Hadiji (2019), see below for details.

1. The function ϕ_0 has peak value $\Delta/2$ and functions $\phi_s, 1 \leq s \leq M$ all have peak value Δ.
2. The function $\phi_0 \in \mathcal{W}^{m_2,2}(\mathcal{X}, L_2)$ and functions $\phi_s \in \mathcal{W}^{m_1,2}(\mathcal{X}, L_1), 1 \leq s \leq M$.
3. For $s \geq 1$, $\phi_s(x) = \phi_0(x)$ for $x \notin H_s$. Also, $\phi_s^* - \phi_s(x) \geq \frac{\Delta}{2}$ when $x \notin H_s$. Here $\phi_s^* = \max_{x \in \mathcal{X}} \phi_s(x)$.

RKHS Version of the Proof. We have now defined $M+1$ hypothesis functions in two balls in two different Sobolev spaces. By (i)the norm equivalency between Sobolev seminorm (Lemma 2) and the RKHS norm; and (ii) the relationships between B_1, L_1 and B_2, L_2 in equation 31, the reward functions also satisfy the following conditions. The conditions below are the RKHS version.

1. The function ϕ_0 has peak value $\Delta/2$ and functions $\phi_s, 1 \leq s \leq M$ all have peak value Δ.
2. $\phi_0 \in \mathcal{H}_{k_{m_2}}(\mathcal{X}, B_2)$, $\phi_s \in \mathcal{H}_{k_{m_1}}(\mathcal{X}, B_1)$, for $1 \leq s \leq M$.
3. $\forall s \geq 1$, $\phi_s(x) = \phi_0(x)$ when $x \notin H_s$. Also, $\phi_s^* - \phi_s(x) \geq \frac{\Delta}{2}$ when $x \notin H_s$.

Lower Bounding Cumulative Regret (Proof Sketch). This part shows the cumulative regret of an algorithm on functions $\phi_1 \dots \phi_M$ is lower bounded by a rate that depends reversely on $\tilde{R}$, if this algorithm has a regret upper bound of $\tilde{R}$ on reward function ϕ_0. The proof in the following directly follows from Hadiji (2019) and relies on Pinsker's inequality. We write down a proof sketch here for completeness, readers interested in the full version can refer to Hadiji (2019, Section F). We use their notations in this part unless otherwise specified. Those include $N_{H_s}(T)$ which is the number of times an algorithm selects an action in bin H_s; $\mathbb{P}_s^T(\cdot)$ which is the probability distribution of trajectory $\{x_t, y_t\}_{t=1\dots T}$, when the reward function in the bandit setting is defined by ϕ_s, for $0 \leq s \leq M$. Similarly, $\mathbb{E}_s[\cdot]$ is the expectation with respect to probability $\mathbb{P}_s$.

By definitions of the reward functions, when the underlying function is ϕ_s for some $s \geq 1$, the cumulative regret is lower bounded by

$$R_{T,s} \geq \frac{\Delta}{2}(T - \mathbb{E}_s[N_{H_s}(T)]) \tag{52}$$

For $s = 0$, the regret is lower bounded by

$$R_{T,0} \geq \frac{\Delta}{2} \sum_{s'=1}^{M} \mathbb{E}_0[N_{H_{s'}}(T)]. \tag{53}$$

Pinsker's inequality is used to establish a relationship between the two lower bounds defined above. The equation 54 is a core step of the proof.

$$\frac{1}{T}\mathbb{E}_s[N_{H_s}(T)] - \frac{1}{T}\mathbb{E}_0[N_{H_s}(T)] \le \sqrt{\frac{1}{2}D_{\mathrm{KL}}(\mathbb{P}_0^T, \mathbb{P}_s^T)}. \tag{54}$$

Calculation of KL distance $D_{\mathrm{KL}}(\cdot,\cdot)$ relies on condition 3 of $\phi_{0\ldots M}$, as well as the assumption that the noise is $1/4$-subgaussian. The result is that the KL distance is bounded by the following.

$$D_{\mathrm{KL}}(\mathbb{P}_0^T, \mathbb{P}_s^T) = 2\mathbb{E}_0[N_{H_s}(T)]\Delta^2. \tag{55}$$

With the above, a key intermediate result is reiterated below.

$$\frac{1}{M}\sum_{s=1}^{M} R_{T,s} \ge \frac{T}{2}\Delta\left(1 - \frac{1}{M} - \sqrt{\frac{\Delta \cdot R_{T,0}}{M}}\right). \tag{56}$$

Recall that our Theorem 11 assumes that $\sup_{f\in\mathcal{W}^{m_2,2}(\mathcal{X},L_2)} R_T \le \tilde{R}$, and since $\phi_0 \in \mathcal{W}^{m_2,2}(\mathcal{X}, L_2)$, it follows directly that $R_{T,0} \le \tilde{R}$. Therefore, the above inequality becomes

$$\begin{aligned}\frac{1}{M}\sum_{s=1}^{M} R_{T,s} &\ge \frac{T}{2}\Delta\left(1 - \frac{1}{M} - \sqrt{\frac{\Delta \cdot R_{T,0}}{M}}\right)\\ &\ge \frac{T}{2}\Delta\left(\frac{1}{2} - \sqrt{\frac{\Delta\tilde{R}}{M}}\right).\end{aligned}$$

In the last inequality, $M \ge 2$ is used. This assumption is *not* violated, as shown later.

Choosing the Appropriate value for Δ. Following the above lower bound, we need to choose a value for Δ that (i) does not violate any of the requirements (Lemma 13) and (ii) maximizes/tightens the lower bound. To do so, the value of Δ should satisfy:

1. $\sqrt{\frac{\Delta\tilde{R}}{M}} \le \frac{1}{4}$, where $\frac{1}{4}$ is a constant less than $\frac{1}{2}$ (chosen in an arbitrary manner).
2. $\Delta/L_1 \le \frac{(K_0^*)}{2^{2m_1-1}M^{m_1-\frac{1}{2}}I_{m_1}^{\frac{1}{2}}}$. Note that this condition satisfies only half of the requirements in Lemma 13. We later show that the other condition in Lemma 13 is also satisfied with the selected Δ.

When maximizing Δ, we first set $\Delta/L_1 \approx \frac{(K_0^*)}{2^{2m_1-1}M^{m_1-\frac{1}{2}}I_{m_1}^{\frac{1}{2}}}$ to achieve the optimal trade-off between M and Δ. That is, we set

$$M = \left\lfloor \left(\frac{L_1 K_0^*}{2^{2m_1-1}I_{m_1}^{\frac{1}{2}}\Delta}\right)^{\frac{1}{m_1-\frac{1}{2}}} \right\rfloor, \tag{57}$$

since M needs to be an integer. By simplifying the constant term:

$$C(m_1) \triangleq \left(\frac{K_0^*}{2^{2m_1-1}I_{m_1}^{\frac{1}{2}}}\right), \tag{58}$$

we get a simpler expression of M:

$$M = \left\lfloor C(m_1)\, L_1^{\frac{2}{2m_1-1}}\Delta^{\frac{-2}{2m_1-1}} \right\rfloor. \tag{59}$$

If $\Delta\tilde{R}/\left(C(m_1)L_1^{\frac{2}{2m_1-1}}\Delta^{\frac{-2}{2m_1-1}}\right) \le \frac{1}{32}$, the condition $\sqrt{\frac{\Delta\tilde{R}}{m}} \le \frac{1}{4}$ would be satisfied, using the fact that $\frac{x}{2} \le \lfloor x \rfloor, \forall x > 2$.

Shuffling some terms, the requirement $\Delta\tilde{R}/\left(C(m_1)L_1^{\frac{2}{2m_1-1}}\Delta^{\frac{-2}{2m_1-1}}\right) \leq \frac{1}{32}$ becomes:

$$\Delta \leq \frac{1}{32}C(m_1)L^{\frac{2}{2m_1-1}}\Delta^{\frac{-2}{2m_1-1}}B^{-1}$$
$$\Delta^{\frac{2m_1+1}{2m_1-1}} \leq \frac{C(m_1)}{32}L_1^{\frac{2}{2m_1-1}}\tilde{R}^{-1}$$
$$\Delta \leq \left(\frac{C(m_1)}{32}\right)^{\frac{m_1-\frac{1}{2}}{m_1+\frac{1}{2}}} L_1^{\frac{1}{m_1+\frac{1}{2}}}\tilde{R}^{-\frac{m_1-\frac{1}{2}}{m_1+\frac{1}{2}}}.$$

To maximize Δ, we thereby choose

$$\Delta = \left(\frac{C(m_1)}{32}\right)^{\frac{m_1-\frac{1}{2}}{m_1+\frac{1}{2}}} L_1^{\frac{1}{m_1+\frac{1}{2}}} M^{-\frac{m_1-\frac{1}{2}}{m_1+\frac{1}{2}}}. \tag{60}$$

This leads to the final lower bound:

$$\begin{aligned}&\frac{1}{M}\sum_{s=1}^{M}R_{T,s}\\ &\geq \frac{T}{2}\Delta\left(\frac{1}{2}-\sqrt{\frac{\Delta\tilde{R}}{M}}\right) \geq \frac{T\Delta}{8}\\ &= \frac{1}{8}\left(\frac{C(m_1)}{32}\right)^{\frac{m_1-1/2}{m_1+1/2}} TL^{\frac{1}{m_1+1/2}}\tilde{R}^{-\frac{m_1-1/2}{m_1+1/2}}.\end{aligned} \tag{61}$$

Verify Assumptions. Last but not least, we have to make sure that the assumptions made throughout the proof are satisfied, by our choice of Δ in equation 60 and M in equation 59.

1. $M \geq 2$. By the definition of M in equation 59, we need to ensure that $C(m_1)\,L_1^{\frac{2}{2m_1-1}}\Delta^{\frac{-2}{2m_1-1}} \geq 2+1=3$. Further, plugging in equation 60, this becomes the following requirement of L_1:

$$L_1 \geq \frac{3^{m_1+\frac{1}{2}}}{32}C(m_1)^{-m_1+\frac{1}{2}}\tilde{R}^{-1}. \tag{62}$$

2. $\Delta/L_2 \leq \frac{K_0^*}{2^{2m_2-2}\sqrt{I_{m_2}}}$. This is the second requirement in Lemma 13 that has not yet been verified to hold. For this condition to hold, the following constraint on L_2 should be met.

$$L_2 \geq C'(m_1,m_2)L_1^{\frac{1}{m_1+1/2}}\tilde{R}^{-\frac{m_1-1/2}{m_1+1/2}}, \tag{63}$$

where,

$$C'(m_1,m_2) = 2^{2m_2-2}\left(\frac{C(m_1)}{32}\right)^{\frac{m_1-1/2}{m_1+1/2}}\frac{\sqrt{I_{m_2}}}{K_0^*} \tag{64}$$

is a constant (independent of T) that depends on m_1, m_2. In other words, to make sure that the requirements in Lemma 13 are met, we need in the assumptions the following constraint.

$$L_1 \leq C'(m_1,m_2)^{-(m_1+\frac{1}{2})}L_2^{m_1+\frac{1}{2}}\tilde{R}^{m_1-\frac{1}{2}}. \tag{65}$$

We have proved Theorem 11 (Sobolev version).

The constraints on B_1 and B_2 in Theorem 3 are derived from the constraints on L_1, L_2 in Theorem 11 and setting B_1, B_2 as instructed in equation 31. Then the proof of Theorem 3 is also completed. $\square$

B.2 Proof of Corollary 5

When $d = 1$, Matérn kernel with regularity parameter ν has Fourier decay rate of $\nu + \frac{1}{2}$ (Definition 4). The algorithm considered in Corollary 5 thus satisfies the regret upper bound on an RKHS induced by a kernel with decay rate $m_2 = \nu_2 + \frac{1}{2}$ which is $\tilde{R} = \tilde{O}(T^{\frac{m_2+\frac{1}{2}}{2m_2}})$. Let m_1 be an integer larger than m_2. Applying Theorem 3, the lower bound on RKHS of a kernel with Fourier decay rate m_1 is $\Omega(\tilde{R}^{-\frac{m_1-\frac{1}{2}}{m_1+\frac{1}{2}}} T)$. For simplicity, we omit the dependence on B (and constant factors) and focus only on the dependence on T. Plugging in the rate of $\tilde{R}$, the lower bound then becomes $\Omega(T^{\frac{m_1 m_2 + \frac{3}{2} m_2 - \frac{1}{2} m_1 + \frac{1}{4}}{2 m_1 m_2 + m_2}})$. Set $m_1 = \nu_1 + \frac{1}{2}$ as the Fourier decay rate of $k_{\text{Matérn},\nu_1}$ in Corollary 5. Then, we get the lower bound by substituting $m_2 = \nu_2 + \frac{1}{2}$ and $m_1 = \nu_1 + \frac{1}{2}$, which is $\Omega(T^{\frac{\nu_1\nu_2 + 2\nu_2 + 1}{(\nu_1+1)(2\nu_2+1)}})$.

B.3 Proof of Theorem 6

UCB-Meta (Liu et al., 2021) achieves minimax regret rate in dependence on T (except log factors) in Hölder spaces with Hölder exponent $\alpha > 1$. For $0 < \alpha \leq 1$, it reduces to the minimax optimal continuum-armed bandit algorithm from Auer et al. (2007). For simplicity, we consider UCB-Meta as the general algorithm for continuum-armed bandits in Hölder spaces. To prove that it is also minimax optimal over RKHS of certain Matérn kernels, we establish the following embedding of RKHS of Matérn kernels to Hölder spaces, via (i) norm equivalency between RKHS of a Matérn-ν kernel and Sobolev space with order m and (ii) Sobolev embedding theorem that specifies the embedding of Sobolev space with order m to Hölder space with exponent α. Note that Singh (2021) have shown that the minimax bandit algorithm over a Besov or Sobolev space is the same as one that is minimax over the smallest Hölder space that the Besov or Sobolev space embeds onto, although not explicitly for RKHS. For completeness, we still include the following proof. We first state the Sobolev embedding theorem (Adams and Fournier, 2003, Theorem 5.4).

Theorem 14 (Sobolev embedding theorem (Adams and Fournier, 2003))**.** *Let m be a non-negative integer. Suppose that the dimension $d < p \cdot m$ and $\alpha = m - \frac{d}{p}$. Let Ω be a finite domain with Lipschitz boundary. Then, the Sobolev space $\mathcal{W}^{m,p}(\Omega)$ is embedded onto Hölder space with exponent α:*

$$\mathcal{W}^{m,p}(\Omega) \subset \Sigma^{\alpha}(\Omega). \tag{66}$$

For our problem setting, we set $p = 2$ and $d = 1$. The domain $\mathcal{X} = [0, 1]$ satisfies the Lipschitz boundary condition. Therefore, $\mathcal{W}^{m}(\mathcal{X}) \subset \Sigma^{\alpha}(\mathcal{X})$ where $\alpha = m - \frac{1}{2}$. Combining Sobolev embedding theorem with the norm equivalency between Sobolev space and RKHS (Lemma 1), we have the following result.

Corollary 15. *Suppose that $k_s : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is a positive-definite translation-invariant kernel, whose Fourier transformation decays polynomially with rate s, $s > d/2, s \in \mathbb{N}$. Then, the RKHS $\mathcal{H}_{k_s}(\mathcal{X})$ is embedded onto Hölder space $\Sigma^{\alpha}(\mathcal{X})$ with exponent $\alpha = s - \frac{d}{2}$:*

$$\mathcal{H}_{k_s}(\mathcal{X}) \subset \Sigma^{s-\frac{d}{2}}(\mathcal{X}). \tag{67}$$

The above relationship is also studied in the earlier work of Shekhar and Javidi (2020, Appendix B.1). Note that Matérn kernels with regularity parameter ν have a Fourier decay rate of $s = \nu + \frac{d}{2}$. Hence, $\mathcal{H}_{k_{\text{Matérn},\nu}}(\mathcal{X}) \subset \Sigma^{\alpha}(\mathcal{X})$, for $\alpha = \nu$. Therefore, since UCB-Meta achieves on $\Sigma^{\alpha}(\mathcal{X})$ the regret rate of $\tilde{O}(T^{\frac{\alpha+1}{2\alpha+1}})$ (Liu et al., 2021, Equation (19)), it achieves the same rate $\tilde{O}(T^{\frac{\nu+1}{2\nu+1}})$ on the subset $\mathcal{H}_{k_{\text{Matérn},\nu}}(\mathcal{X})$. Here, we omit the dependence on B, the RKHS norm bound. A function $f \in \mathcal{H}_{k_{\text{Matérn},\nu}}(\mathcal{X}, B)$ also has a finite Hölder norm $\|f\|_{\Sigma^{\alpha=\nu}}$. The norm $\|f\|_{\Sigma^{\nu}}$, by definition, poses an upper bound on L (using the notation from Liu et al. (2021, Definition 1), the Hölder-continuity coefficient of the l-th order derivative of f, where l is the largest integer strictly less than α. By Theorem 4 from Liu et al. (2021), we can see that L affects the regret only through a multiplicative term and not through the exponents of T. Therefore, we omit the dependence on B and write the regret rate of UCB-Meta as $\tilde{O}(T^{\frac{\nu+1}{2\nu+1}})$.

B.4 Proof of Theorem 7

Recall that Theorem 5.3 in Pacchiano et al. (2020b) provides general regret bounds for CORRAL. The proof of our Theorem 7 is an adaptation to the proof of Theorem 5.3 in Pacchiano et al. (2020b). We use the same notations as Pacchiano et al. (2020b) unless otherwise specified. M is the number of base algorithms (also aligning with the statement in Theorem 7). δ is the probability of failure. $U : \mathbb{R} \times [0, 1] \to \mathbb{R}^+$ is the cumulative regret function (for a base algorithm), such that $U(t,\delta)$ is the high-probability and anytime regret bound of a base algorithm. ρ is the maximum of reciprocals of the probability that the base algorithm is chosen by the master algorithm over all time steps. η is the learning rate of the master algorithm whose value is determined later in the proof.

In Section 5.1.1, we discussed briefly SupKernelUCB (Valko et al., 2013) versus GP-UCB Srinivas et al. (2009). Despite the convenient implementation and good empirical performance of GP-UCB, SupKernelUCB matches the non-adaptive lower bound in the dependence on T except log factors under the RKHS assumption and thus is minimax optimal while GP-UCB is not. UCB-Meta (Liu et al., 2021) as shown in Theorem 6 is also minimax optimal in the dependence on T except log factors for the Matérn RKHS setting. For this subsection, however, we use SupKernelUCB as base algorithms, since the regret bound of SupKernelUCB has an explicit dependence on B, while for UCB-Meta the dependence on B would rely on an implicit constant (see proof of Theorem 6 in Appendix B.3). We set $d = 1$ as specified in Section 5.

Given B and ν of a Matérn-ν kernel, the regret bound of SupKernelUCB is $\tilde{O}(B^{\frac{1}{2}}T^{\frac{\nu+1}{\nu+2}})$ in the RKHS of the Matérn kernel (Valko et al., 2013, Theorem 1). Note that the original SupKernelUCB (i) is for finite action set and (ii) takes T as input and therefore does not have any time regret guarantees. As mentioned in Section 5.1.1, Cai and Scarlett (2021) argue that the aforementioned problem (i) could be extended to the continuum-armed setting by a discretization argument with an extra $O(d(log(T)))$ term in the regret. The problem (ii) can be theoretically circumvented by the doubling procedure (Auer et al., 1995). Doubling converts an algorithm with (cumulative) regret bound for fixed T to one with anytime regret bound, suffering only up to constant factors in the regret. [16] Therefore, for theoretical interest, we treat SupKernelUCB as the minimax optimal base algorithm with anytime regret upper bound $\tilde{O}(B^{\frac{1}{2}}T^{\frac{\nu+1}{\nu+2}}), \forall T$.

We acknowledge that this is for theoretical convenience only and it remains an important open problem (Vakili et al., 2021b) to improve the regret bound of the practical GP-UCB algorithm under RKHS assumptions.

We plug in $U(T,\delta) = \tilde{O}(B^{\frac{1}{2}}T^{\frac{\nu+1}{2\nu+1}})$ for the base algorithms for CORRAL. Following the proof of Pacchiano et al. (2020b, Theorem 5.3), we have the following. Note that this upper bound holds with respect to any base algorithm with anytime high-probability regret $U(t,\delta)$. Therefore, we plug in the regret of the best base algorithm, which is $U(t,\delta) = \tilde{O}({B^*}^{\frac{1}{2}}t^{\frac{\nu^*+1}{2\nu^*+1}})$ because ν^*, B^* belong in the set of candidate values $\boldsymbol{u}$.

$$\begin{aligned}
R_T &\le O(\frac{M\ln(T)}{\eta} + T\eta) - \mathbb{E}\left[\frac{\rho}{40\eta\ln(T)} - \rho U(T/\rho,\delta)\log(T)\right] + \delta T + 8\sqrt{MT\log(\frac{4TM}{\delta})} \\
&\le \tilde{O}(\frac{M}{\eta} + T\eta + \delta T + \sqrt{MT}) - \mathbb{E}\left[\tilde{O}(\frac{\rho}{\eta} - \rho\sqrt{B^*}T^{\frac{\nu^*+1}{2\nu^*+1}}\rho^{-\frac{\nu^*+1}{2\nu^*+1}})\right] \\
&\overset{\text{set } \delta=\frac{1}{T}}{=} \tilde{O}(\frac{M}{\eta} + T\eta + \sqrt{MT}) - \mathbb{E}\left[\tilde{O}(\frac{\rho}{\eta} - \sqrt{B^*}T^{\frac{\nu^*+1}{2\nu^*+1}}\rho^{\frac{\nu^*}{2\nu^*+1}})\right]
\end{aligned}$$

Maximizing the above equation over ρ results in $\rho \propto \eta^{\frac{2\nu^*+1}{\nu^*+1}}{B^*}^{\frac{\nu^*+\frac{1}{2}}{\nu^*+1}}T$. If we plug this value for ρ in the above equation, then the regret is bounded by:

$$\begin{aligned}
R_T &= \tilde{O}(\frac{M}{\eta} + T\eta + \sqrt{MT}) - \tilde{O}(\eta^{\frac{\nu^*}{\nu^*+1}}{B^*}^{\frac{\nu^*+\frac{1}{2}}{\nu^*+1}}T - \eta^{\frac{\nu^*}{\nu^*+1}}{B^*}^{\frac{2\nu^*+1}{2\nu^*+2}}T) \\
&\le \tilde{O}(\frac{M}{\eta} + T\eta + \sqrt{MT} + \eta^{\frac{\nu^*}{\nu^*+1}}{B^*}^{\frac{2\nu^*+1}{2\nu^*+2}}T)
\end{aligned}$$

For the problem of adapting to kernel regularity (represented by ν^* when the kernel is a Matérn kernel), since CORRAL does not have access to ν^* (and B^*), we choose η with respect to the user-specified parameter $\tilde{\nu}$: $\eta = T^{-\frac{\tilde{\nu}+1}{2\tilde{\nu}+1}}$. Plugging this choice of η back in the above equation, we have:

$$R_T \le \tilde{O}(MT^{\frac{\tilde{\nu}+1}{2\tilde{\nu}+1}} + {B^*}^{\frac{2\nu^*+1}{2\nu^*+2}}T^{\frac{\tilde{\nu}\nu^*+2\tilde{\nu}+1}{(2\tilde{\nu}+1)(\nu^*+1)}}).$$

Absorbing the dependence on M and B in $\tilde{O}$, we then have the regret rate in equation 19.

B.5 Proof of Theorem 10

The proof follows from the general form of regret upper bound of RBBE (Theorem 5.1 from Pacchiano et al. (2020a)). The regret bound in Theorem 5.1 in Pacchiano et al. (2020a) is expressed with the "play ratio" $\sum_{i\in\mathcal{B}}\frac{n_i(t_i)}{n_*(t_i)}$, where $\mathcal{B}$ denotes

[16] The doubling procedure is also used in other works that use CORRAL to adapt to unknown parameters of the function space, for example Liu et al. (2021) which studied adaptivity to the Hölder exponent.

the set of misspecified base algorithms, t_i denotes the last round before base algorithm i is eliminated, and $n_i(t)$ denotes the number of times i is selected until time step $t \leq T$. In the following part, we use Lemma A.3 in Pacchiano et al. (2020a) to calculate the play ratio, then plug it in Theorem 5.1 of Pacchiano et al. (2020a) to get the final regret bound. For reasons why the more straightforward result (Theorem 5.4 in Pacchiano et al. (2020a)) is not used, see the end of this subsection for an explanation.

In the following, each base algorithm i has the following candidate pseudo regret bound (equation (7) in Pacchiano et al. (2020a)):

$$R_i(t) \leq C\theta_i T^{\beta_i}, \tag{68}$$

where $C \geq 1$ is some term independent of T or i, and $\theta_i \geq 1$ is some parameter dependent on i. For minimax optimal kernelised bandit algorithms instantiated with ν_i (parameter of the Matérn kernel), $\beta_i = \frac{\nu_i+d}{2\nu_i+d}$. We write down the general regret bound of RBBE here for completeness (Theorem 5.1 (Pacchiano et al., 2020a)). Below, $*$ denotes any well-specified learner, that is, a leaner whose actual (pseudo) regret Reg_i is upper bounded by its candidate (which means if well-specified) regret bound $R_i(T)$.

$$\begin{aligned}
R_T \leq & \sum_{i=1}^{M} R_*(n_*(t_i)) + \sum_{i\in\mathcal{B}} \frac{n_i(t_i)}{n_*(t_i)} R_*\left(n_*(t_i)\right) + 2M + 2c\sum_{i\in\mathcal{B}}\sqrt{n_i(t_i)\ln(\frac{M\ln(T)}{\delta})} \\
& + 2c\sum_{i\in\mathcal{B}}\sqrt{\frac{n_i(t_i)}{n_*(t_i)}}\sqrt{n_i(t_i)\ln(\frac{M\ln(T)}{\delta})}
\end{aligned}$$

We refer to the five terms in the above summation above as $\#1 \ldots \#5$.

The terms $\#1 + \#3$ can be bounded the same way as in the proof of Theorem 5.4 in Pacchiano et al. (2020a):

$$\sum_{i=1}^{M} R_*(n_*(t_i)) + 2M \leq MR_*(T) + 2M \leq \tilde{O}(M\theta_* T^{\beta_*}).$$

The term $\#4$ is bounded also following the proof in Pacchiano et al. (2020a):

$$\begin{aligned}
2c\sum_{i\in\mathcal{B}}\sqrt{n_i(t_i)\ln(\frac{M\ln(T)}{\delta})} &\leq 2c\sqrt{|\mathcal{B}|\ln\frac{M\ln(T)}{\delta}\sum_{i\in\mathcal{B}} n_i(t_i)} \\
&\leq 2c\sqrt{|\mathcal{B}|T\ln\frac{M\ln(T)}{\delta}}
\end{aligned}$$

Bounding the term $\#1$ and $\#5$, however, needs changes to the proof of Theorem 5.4 (Pacchiano et al., 2020a), since the play ratio is involved. Lemma A.3 in Pacchiano et al. (2020a) states that for two base learners i, j,

$$\frac{n_i(t)}{n_j(t)} \leq \max\left\{\left(2\frac{\theta_j}{\theta_i}\right)^{\frac{1}{\beta_i}} (n_j(t))^{\frac{\beta_j}{\beta_i}-1}, 2\right\}. \tag{69}$$

Therefore, the play ratio between a misspecified base learner i and a well-specified leaner $*$ can be bounded by:

$$\begin{aligned}
\frac{n_i(t)}{n_*(t)} &\leq 2 + \left(2\frac{\theta_*}{\theta_i}\right)^{\frac{1}{\beta_i}} n_*(t)^{\frac{\beta_*}{\beta_i}-1} \\
&\leq 2 + 4C_2B_* n_*(t)^{\frac{\beta_*}{\beta_i}-1} \\
&\leq 2 + 4C_2B_* n_*(t)^{2\beta_*-1}.
\end{aligned}$$

The first inequality above is simply plugging $j = *$ (representing a well-specified learner), and using that $\max\{x, y\} \leq x + y$. For the second inequality, recall that the minimax optimal SupKernelUCB algorithm has a regret rate (if the kernel parameter ν and RKHS norm bound B are known) of $\tilde{O}(\sqrt{B\gamma_T T}) = \tilde{O}(\sqrt{B}T^{\frac{\nu+d}{2\nu+d}})$. The $\tilde{O}$ notation hides polynomial terms that are dependent on $\log(T), d$. Therefore, the parameter θ_i in equation 68 that depends on the index of the base algorithm i is $\theta_i \propto \sqrt{B_i}$. Given the assumption that $\theta_i \geq 1$, $\frac{\theta_*}{\theta_i} \leq C_1\sqrt{B_*}$ for some constant C_1. Since $\beta_i \geq \frac{1}{2}$,

$\left(2\frac{\theta_*}{\theta_i}\right)^{\frac{1}{\beta_i}} \leq 4C_2B_*$ for some constant C_2. Also in the last two inequalities, we used $\beta_i \geq \frac{1}{2}$, that is, every base algorithm used in Theorem 10 have at least $\tilde{O}(T^{\frac{1}{2}})$ regret. Therefore, we have the following bound on the sum of play ratio:

$$\sum_{i\in\mathcal{B}} \frac{n_i(t)}{n_*(t)} \leq 2|\mathcal{B}| + 4C_2B_*|\mathcal{B}|(n_*(t))^{(2\beta_*-1)} \tag{70}$$

$$\leq 2|\mathcal{B}| + 4C_2B_*|\mathcal{B}|T^{(2\beta_*-1)} = 2|\mathcal{B}|(1+2C_2B_*T^{(2\beta_*-1)}) \tag{71}$$

We can plug equation equation 71 to bound $\#5$ as follows.

$$\begin{aligned}
2c\sum_{i\in\mathcal{B}}\sqrt{\frac{n_i(t_i)}{n_*(t_i)}}\sqrt{n_i(t_i)\ln(\frac{M\ln(T)}{\delta})} &\leq 2c\sqrt{\sum_{i\in\mathcal{B}}\frac{n_i(t_i)}{n_*(t_i)}\sum_{i\in\mathcal{B}}n_i(t_i)\ln\frac{M\ln(T)}{\delta}} \\
&\leq 2c\sqrt{\sum_{i\in\mathcal{B}}\frac{n_i(t)}{n_*(t)}T\ln\frac{M\ln(T)}{\delta}} \\
&\leq 2c\sqrt{2|\mathcal{B}|(1+2C_2B_*T^{(2\beta_*-1)})T\ln\frac{M\ln(T)}{\delta}} \\
&= \tilde{O}(|\mathcal{B}|^{\frac{1}{2}}{B_*}^{\frac{1}{2}}T^{\beta_*})
\end{aligned}$$

Similarly, the upper bound of term $\#2$ relies on equation 71 as well.

$$\begin{aligned}
\sum_{i\in\mathcal{B}}\frac{n_i(t_i)}{n_*(t_i)}R_*\left(n_*(t_i)\right) &\leq C\sum_{i\in\mathcal{B}}\frac{n_i(t)}{n_*(t)}\theta_*n_*(t_i) \\
&\leq C\sum_{i\in\mathcal{B}}\frac{n_i(t_i)}{(n_*(t_i))^{1-\beta_*}}\theta_* \\
&\leq C\left(\sum_{i\in\mathcal{B}}\frac{n_i(t_i)}{n_*(t_i)}\right)^{(1-\beta_*)}\theta_*(n_i(t_i))^{\beta_*} \\
&\leq C\theta_*\left(2|\mathcal{B}|(1+2C_2B_*T^{(2\beta_*-1)})\right)^{(1-\beta_*)}T^{\beta_*} \\
&= \tilde{O}(\theta_*|\mathcal{B}|^{(1-\beta_*)}{B_*}^{1-\beta^*}T^{(2\beta_*-1)(1-\beta_*)+\beta_*}) \\
&= \tilde{O}(\theta_*|\mathcal{B}|^{(1-\beta_*)}{B_*}^{1-\beta^*}T^{4\beta_*+2\beta_*^2-1})
\end{aligned}$$

Now that the asymptotic rates of the five terms are derived, we can see that term $\#2$ dominates in the dependence of T and $\#5$ dominates dependence on $|\mathcal{B}|$, B_*, and hence, the regret of RBBE can be bounded as follows.

$$R_T \leq \tilde{O}(\theta_*|\mathcal{B}|^{\frac{1}{2}}{B_*}^{\frac{1}{2}}T^{4\beta_*+2\beta_*^2-1}) \tag{72}$$

$$= \tilde{O}(\theta_*|\mathcal{B}|^{\frac{1}{2}}{B_*}^{\frac{1}{2}}T^{\frac{2\nu_*^2+4\nu^*+1}{(2\nu^*+1)^2}}) \tag{73}$$

$$= \tilde{O}(\theta_*M^{\frac{1}{2}}{B_*}^{\frac{1}{2}}T^{\frac{2\nu_*^2+4\nu^*+1}{(2\nu^*+1)^2}}) \tag{74}$$

Finally, the reason for not using the straightforward results in Theorem 5.4 of Pacchiano et al. (2020a) is as follows. In adaptation to the kernel regularity parameter ν, the candidate regret bounds of base algorithms do *not* have the same exponent of T. The candidate regret bounds having the same rates of T is a requirement for the more straightforward results, hence, those results are not directly applicable to our setting.