On Information Gain and Regret Bounds in Gaussian Process Bandits

Sattar Vakili, Prowler.io, Cambridge, UK
sattar@prowler.io
Kia Khezeli, Center for Individualized Medicine, Mayo Clinic, Rochester, MN, USA
khezeli.kia@mayo.edu
Victor Picheny, Prowler.io, Cambridge, UK
victor@prowler.io

Abstract

Consider the sequential optimization of an unknown, expensive to evaluate and possibly non-convex objective function \( f \) from noisy observations which can be considered as a continuum-armed bandit problem. Bayesian optimization algorithms based on Gaussian Process (GP) models are shown to perform favorably in this setting. In particular, upper bounds are proven on the regret performance of two popular algorithms — GP-UCB and GP-TS — under both Bayesian (when \( f \) is a sample from a GP) and frequentist (when \( f \) lives in a reproducing kernel Hilbert space) settings. The regret bounds crucially depend on a quantity referred to as the maximal information gain \( \gamma_T \) between \( T (\in \mathbb{N}) \) observations and the underlying GP (surrogate) model. In this paper, we build on the spectral properties of positive definite kernels to prove novel bounds on \( \gamma_T \). In comparison to the existing works which rely on specific kernels (such as Matérn and SE) to provide explicit bounds on \( \gamma_T \) and regret, we provide general results in terms of the decay rate of the eigenvalues of the kernel. Specialising our results for common kernels leads to significant improvements over the existing bounds on \( \gamma_T \) and regret. For the Matérn and SE kernels, where the lower bounds on regret are known, our results reduce the gap between the upper and lower bounds from a polynomial in \( T \) factor, in the existing work, to a logarithmic one, under the Bayesian setting. Furthermore, since our bounds on \( \gamma_T \) are independent of the optimisation algorithm, they impact the regret bounds under various other settings where \( \gamma_T \) is essential.

Keywords— Information gain, effective dimension, regret bounds, Bayesian optimization, GP-UCB, GP-TS, continuum-armed bandits.

1 Introduction

Sequential optimization is one of the fastest growing areas of machine learning. Bayesian optimization building on Gaussian Process (GP) models has proven a powerful tool for addressing the exploration-exploitation trade-off in the sequential optimization of non-convex objective functions with bandit feedback. This class of problems finds application in a variety of scientific and industrial settings that have the nature of experimental design. Examples include: scientific experiments to gain insights into physical and social phenomena; industrial production; clinical trials and drug discovery; software and web design; sensor networks to monitor ecological systems; and hyper-parameter tuning in machine learning models (see e.g. Shahriari et al., 2016, for a review).
There have been significant recent advances in the analysis of GP-based Bayesian optimization algorithms providing performance guarantees in terms of regret. That is defined as the cumulative loss in the value of the objective function \( f \) at a sequence of observation points \( \{x_t\}_{t=1}^{T} \), \( T \in \mathbb{N} \), in comparison to its value at a global maximum \( x^* \in \text{argmax}_{x \in \mathcal{X}} f(x) \) over the search space \( \mathcal{X} \subseteq \mathbb{R}^d \) (see (1)). In their seminal paper, Srinivas et al. (2010) established performance guarantees for GP-UCB, an optimistic optimization algorithm which sequentially selects \( x_t \) that maximize an upper confidence bound score over the search space. They considered a fully Bayesian setting where \( f \) is assumed to be a sample from a GP with a known kernel, as well as, a frequentist setting (referred to as agnostic in Srinivas et al. (2010)) where \( f \) is assumed to live in a reproducing kernel Hilbert space (RKHS) with a known kernel. They showed an \( \tilde{O}(\sqrt{T}) \) and an \( \tilde{O}(\gamma_T \sqrt{T}) \) regret bound for GP-UCB under the Bayesian and frequentist settings, respectively, where \( \gamma_T \) is a measure of the hardness of the optimization task referred to as the maximal information gain (see §3.3). The scaling of \( \gamma_T \) with \( T \) is sublinear and depends on the GP kernel. Since the pioneering result of Srinivas et al., there has been several works on improving the bounds toward their optimal value. Chowdhury and Gopalan (2017) improved the regret bounds under the frequentist setting by multiplicative logarithmic in \( T \) factors. Furthermore, they showed that \( \tilde{O}(\gamma_T \sqrt{T}) \) regret bounds, under the frequentist setting, also hold for GP-TS, a Bayesian optimization algorithm based on Thompson Sampling which sequentially draws \( x_t \) from the posterior distribution of \( x^* \). Under the Bayesian setting, Kandasamy et al. (2018) built on ideas from Russo and Van Roy (2014, 2016) to show that GP-TS achieves the same order of regret as GP-UCB.

The regret bounds mentioned above become complete only when \( \gamma_T \) is properly bounded which proves challenging. Srinivas et al. (2010) showed that \( \gamma_T = \tilde{O}(T^{\frac{d+1}{d(d+1)}}) \) for the Matérn-\( \nu \) kernel (see (10)) and \( \gamma_T = \mathcal{O}(\log^{d+1}(T)) \) for the Squared Exponential (SE) kernel. Plugging these bounds on \( \gamma_T \) into the regret bounds mentioned above results in explicit upper bounds in terms of \( T \) which are in the form of \( \tilde{O}(T^{\frac{d+1}{d(d+1)}}) \) and \( \tilde{O}(T^{\frac{d+1}{d(d+1)}}) \) for the Matérn-\( \nu \) kernel, under the Bayesian and frequentist settings, respectively; and in the form of \( \mathcal{O}(T^{\frac{d+1}{d(d+1)}}) \) and \( \mathcal{O}(T^{\frac{d+1}{d(d+1)}}) \) for the SE kernel, under the Bayesian and frequentist settings, respectively.

Finding the optimal regret bounds is a long standing open question. In the case of Matérn and SE kernels, there exist lower bounds on the performance of Bayesian optimization algorithms which facilitate the assessment of the upper bounds. Specifically, Scarlett et al. (2017) proved an \( \Omega(T^{\frac{d+1}{d+1}}) \) lower bound for Bayesian optimization algorithms with Matérn-\( \nu \) kernel. A comparison between this lower bound and the existing upper bounds unfortunately shows a drastic gap which can be as large as \( \mathcal{O}(\sqrt{T}) \), under both settings, with particular configurations of parameters \( \nu \) and \( d^2 \). In the following section, we discuss our contributions, one of which is reducing this polynomial gap in \( T \) to a logarithmic one.

### 1.1 Contribution

Our contribution is to establish novel bounds on \( \gamma_T \) which directly translate to new regret bounds for Bayesian optimization algorithms. Crucial to our analysis, we use Mercer’s theorem to represent the GP kernel in terms of its eigenvalue-eigenfeature decomposition — an inner product in the corresponding reproducing kernel Hilbert space (RKHS) — which is infinite dimensional for typical kernels. To overcome the difficulty of working in infinite dimensional spaces, we use a projection on a finite \( D \) dimensional space that allows us to bound the information gain in terms of \( D \) and the spectral properties of the GP kernel.

---

1. We use the notations \( \mathcal{O} \) and \( \Omega \) to denote the standard mathematical orders and the notation \( \tilde{O} \) to suppress the logarithmic factors.

2. E.g. consider a case where \( \nu \) and \( d \) grow large; \( \nu \) grows faster than \( d \) and slower than \( d^2 \). In this case, the lower bound and the upper bounds become arbitrarily close to \( \Omega(\sqrt{T}) \) and \( \mathcal{O}(T) \), respectively. Therefore, the worst case gap between them is in \( \mathcal{O}(\sqrt{T}) \).
For a GP kernel with decreasing eigenvalues \( \{\lambda_m\}_{m=1}^{\infty} \), we consider two cases of polynomial, \( \lambda_m = O(m^{-\beta_p}) \), \( \beta_p > 1 \), and exponential, \( \lambda_m = O(exp(-m^{\beta_e})) \), \( \beta_e > 0 \), decays. We prove \( O\left(T^{\frac{\beta_p}{\beta_p - 1}} \log^{\frac{1}{\beta_p}}(T)\right) \) and \( O\left(\log^{1+\frac{1}{\beta_e}}(T)\right) \) upper bounds on \( \gamma_T \) under these two cases, respectively. The application of our bounds on \( \gamma_T \) to the regret bounds results in new upper bounds based on \( \beta_p \) and \( \beta_e \) which are summarized in Table 1. In comparison to the existing works which rely on specific kernels (e.g. Matérn-\( \nu \), SE) for explicit regret bounds, our results provide general explicit regret bounds, as long as the conditions on the decay rate of the eigenvalues of the GP kernel (referred to as eigendecay for brevity) are satisfied.

| Decay rate of \( \lambda_m \) | Bound on \( \gamma_T \) | Regret Bound (GP-UCB) |
|-----------------------------|-------------------|---------------------|
| Polynomial \( \lambda_m = O(m^{-\beta_p}) \) | \( O\left(T^{\frac{\beta_p}{\beta_p - 1}} \log^{\frac{1}{\beta_p}}(T)\right) \) | Bayesian: \( O\left(T^{\frac{\beta_p}{\beta_p - 1}}, \log^{\frac{1}{\beta_p}}(T)\right) \) |
| Exponential \( \lambda_m = O(exp(-m^{\beta_e})) \) | \( O\left(\log^{1+\frac{1}{\beta_e}}(T)\right) \) | Bayesian: \( O\left(T^{\frac{1}{\beta_e}}, \log^{1+\frac{1}{\beta_e}}(T)\right) \) |

Table 1: The bounds on the information gain and the regret of GP-UCB under Bayesian and Frequentist settings and conditions on the eigendecay of the GP kernel. The regret bounds for GP-TS, under all cases, are the same as the respective ones for GP-UCB up to a multiplicative \( O(\sqrt{\log(T)}) \) factor.

As an instance of polynomially decaying eigenvalues, our results apply to the Matérn-\( \nu \) kernel (see (10)) showing \( O(T^{\frac{\nu+d}{\nu+d-\nu\nu}}) \) and \( O(T^{\frac{\nu+d}{\nu+d-\nu\nu}}) \) regret bounds under the Bayesian and frequentist settings, respectively (which are respectively improvements over the regret bounds reported in Srinivas et al. (2010) and Chowdhury and Gopalan (2017)). Our regret bound is tight under the Bayesian setting, closing the gap with the \( \Omega(T^{\frac{\nu+d}{\nu+d-\nu\nu}}) \) lower bound reported in Scarlett et al. (2017) and indicating the optimality of our bound on the effective dimension (both up to an \( O(\log(T)) \) factor). Under the frequentist setting, our regret bounds improve, polynomially in \( T \), upon the existing ones. However, the question of optimal regret bounds under this setting remains open (see §5). As an instance of exponentially decaying eigenvalues, our results apply to the Squared Exponential (SE) kernel. A summary of the results is given in Table 2.

| Kernel | Bound on \( \gamma_T \) | Lower Bound | Upper Bound (GP-UCB) |
|--------|-----------------|-------------|---------------------|
| Matérn-\( \nu \) | \( O\left(T^{\frac{\nu+d}{\nu+d-\nu\nu}}\right) \) | \( \Omega(T^{\frac{\nu+d}{\nu+d-\nu\nu}}) \) | Bayesian: \( O\left(T^{\frac{\nu+d}{\nu+d-\nu\nu}}\right) \) |
| | | | Frequentist: \( O\left(T^{\frac{\nu+d}{\nu+d-\nu\nu}}\right) \) |
| SE | \( O\left(\log^{d+1}(T)\right) \) | \( \Omega\left(T^{\frac{1}{2}}, \log^{\frac{d}{2}}(T)\right) \) | Bayesian: \( O\left(T^{\frac{1}{2}}, \log^{d+1}(T)\right) \) |
| | | | Frequentist: \( O\left(T^{\frac{1}{2}}, \log^{d+1}(T)\right) \) |

Table 2: The bounds on the information gain and the regret of GP-UCB under Bayesian and Frequentist settings with Matérn-\( \nu \) and SE kernels (established in this paper). The regret bounds for GP-TS, under all cases, are the same as the respective ones for GP-UCB up to a multiplicative \( O(\sqrt{\log(T)}) \) factor. The lower bounds on the second column of the table were reported in Scarlett et al. (2017). Under the Bayesian setting, the gap between the upper and lower bounds is reduced to a \( \log(T) \) factor. Under the frequentist setting with the SE kernel, the gap is reduced to a \( \log(T) \) factor. Under the frequentist setting with the Matérn kernel, the gap remains polynomial; although, significantly reduced with our results.

While we focus on the classic sequential optimization problem in this paper, it is worth noting that the bounds on \( \gamma_T \) are also essential for numerous variants of the problem such as the ones under the settings with contextual information, safety constraints and multi-fidelity evaluations (see §1.2 for a list of references). Our bounds on \( \gamma_T \) directly apply and improve the regret bounds depending on \( \gamma_T \) under these various settings.
The interest in the bounds on \( \gamma_T \) goes beyond the regret bounds. For example, the concentration inequalities of the GP models under the frequentist setting (cf. Theorem 2 of Chowdhury and Gopalan (2017)) depend on \( \gamma_T \). Our results improve such concentration inequalities.

A closely related quantity to the information gain is the so-called effective dimension \( D_T \) of the problem that satisfies \( D_T = O(\gamma_T) \) (see §2.5 for the details). Calandriello et al. (2019) introduced a variation of GP-UCB which improves its computational cost. The improved computational cost depends on \( D_T \). Our bounds on \( \gamma_T \) (consequently on \( D_T \)) improve such bounds on the algorithmic properties of GP-based methods.

1.2 Other Related Work

There has been an increasing interest in Bayesian optimization based on GP models in recent years. Performance guarantees for GP-UCB and GP-TS are studied under various settings including contextual information (Krause and Ong, 2011), high dimensional spaces (Djolonga et al., 2013; Mutny and Krause, 2018), safety constraints (Berkenkamp et al., 2016; Sui et al., 2018), parallelization (Kandasamy et al., 2018), meta-learning (Wang et al., 2018), multi-fidelity evaluations (Kandasamy et al., 2019), ordinal models (Picheny et al., 2019), and corruption tolerance (Bogunovic et al., 2020), to name a few. Javidi and Shekhar (2018) introduced an adaptive discretization of the search space improving the computational complexity of a GP-UCB based algorithm. Sparse approximation of GP posteriors are shown to preserve the regret orders while significantly improving the computational complexity of both GP-UCB (Mutny and Krause, 2018; Calandriello et al., 2019) and GP-TS (Vakili et al., 2020). Most of the existing works report their regret bounds in terms of \( \gamma_T \) (or equivalently, in terms of the bounds on \( D_T \)). Our results directly apply to all the works mentioned above and improve their regret bounds shall our bounds on \( \gamma_T \) replace the existing ones.

Our analytical approach and conditions on the eigendecay of the GP kernel bear similarity to Chatterji et al. (2019) which studied the problem of online learning with kernel losses. The problems and their analysis however have substantial differences. They considered a more challenging adversarial setting for the objective function. However, they restricted the objective function to the subspace of one dimensional functions in the RKHS which is very limiting for our purposes (one of the main challenges in our analysis is the infinite-dimensionality of the GP model in the RKHS). The algorithmic designs, based on exponential weights, under the adversarial setting are also significantly different from GP-UCB and GP-TS, especially, in the sense that their analysis does not rely on the information gain.

Both GP-UCB and GP-TS are rooted in the classic multi-armed bandit literature (see Auer et al., 2002; Russo and Van Roy, 2016; Slivkins, 2019; Zhao, 2019, and references therein). Our work strengthens the link between linear (Dani et al., 2008; Rusmevichientong and Tsitsiklis, 2010; Abbasi-Yadkori et al., 2011; Agrawal and Goyal, 2013; Abeille and Lazaric, 2017) and kernelized (GP-based) (Srinivas et al., 2010; Chowdhury and Gopalan, 2017) models for sequential optimization with bandit feedback, as we build our analysis based on a finite-dimensional projection that is equivalent to linear bandits.

The rest of the paper is organized as follows. The problem formulation, the preliminaries on GPs, GP-UCB and GP-TS, and the background on the connection between the regret bounds and the information gain are presented in §2. The analysis of the bounds on \( \gamma_T \) is provided in §3. The explicit regret bounds (in terms of \( T \)) for GP-UCB and GP-TS are given in §4. We conclude with discussing the open problem of the optimal regret bounds under the frequentist setting in §5.

2 Problem Formulation And Preliminaries

In this section, we provide background information on sequential optimization, GPs, and the connection between the information gain and the regret bounds for Bayesian optimization algorithms.
We use the following notations throughout the paper. For a square matrix $M \in \mathbb{R}^{n \times n}$, the notations $\det(M)$ and $\text{tr}(M)$ denote the determinant and the trace of $M$, respectively. The notation $M^\top$ denotes the transpose of an arbitrary matrix $M$. For a positive definite matrix $P$, $\log \det(P)$ denotes $\log(\det(P))$. The identity matrix of dimension $n$ is denoted by $I_n$. For a vector $z \in \mathbb{R}^n$, the notation $\|z\|_2$ denotes the 2-norm of $z$.

2.1 The Sequential Optimization Problem

Consider the sequential optimization of a fixed and unknown objective function $f$ over a compact set $\mathcal{X} \subset \mathbb{R}^d$. A learning algorithm $\pi$ sequentially selects an observation point $x_t \in \mathcal{X}$ at each discrete time instance $t = 1, 2, \ldots$, and receives the corresponding real-valued reward $y_t = f(x_t) + \epsilon_t$, where $\epsilon_t$ is the observation noise. Specifically, $\pi = \{\pi_t\}_{t=1}^{\infty}$ is a sequence of mappings $\pi_t : \mathcal{H}_{t-1} \rightarrow \mathcal{X}$ from the history of observations to a new observation point; $\mathcal{H}_t = \{X_t, y_t\}$, $X_t = [x_1, x_2, \ldots, x_t]^\top$, $y_t = [y_1, y_2, \ldots, y_t]^\top$, $x_s \in \mathcal{X}$, $y_s \in \mathbb{R}$, for all $s \geq 1$. The regularity assumptions on $f$ and $\epsilon_t$ are specified in §2.4.

The goal is to minimize regret defined as the expected cumulative loss compared to the maximum attainable objective, over a time horizon $T$. Specifically,

$$R(T; \pi) = \mathbb{E} \left[ \sum_{t=1}^{T} (f(x^\ast) - f(x_t)) \right],$$

(1)

where $x^\ast = \arg\max_{x \in \mathcal{X}} f(x)$ is a global maximum of $f$ and the expectation is taken with respect to the possible randomness in $X_t$ that is determined by $\pi$ and $\{\epsilon_t\}_{t=1}^{T}$. To keep the notation uncluttered, we do not include the dependency on $\pi$ in the notation of $X_t$, as it shall be clear from the context throughout the paper.

2.2 Gaussian Processes

The learning algorithms considered here build on GP (surrogate) models. A GP is a random process $\{\tilde{f}(x)\}_{x \in \mathcal{X}}$, whose finite subsets each follows a multivariate Gaussian distribution (Rasmussen and Williams). The distribution of a GP can be specified by its mean function $\mu(x) = \mathbb{E}[\tilde{f}(x)]$ and a positive definite kernel (or covariance function) $k(x, x') = \mathbb{E}[\tilde{f}(x) - \mu(x))(\tilde{f}(x') - \mu(x'))]$. Without loss of generality, it is typically assumed that $\forall x \in \mathcal{X}, \mu(x) = 0$ for prior GP distributions.

Conditioning GPs on available observations provides us with powerful non-parametric Bayesian (surrogate) models over the space of functions. In particular, conditioned on $\mathcal{H}_t$, the posterior of $\tilde{f}$ is a GP with mean function $\mu_t(x) = \mathbb{E}[\tilde{f}(x)|\mathcal{H}_t]$ and kernel function $k_t(x, x') = \mathbb{E}[(\tilde{f}(x) - \mu_t(x))(\tilde{f}(x') - \mu_t(x'))|\mathcal{H}_t]$ specified as follows:

$$\mu_t(x) = k_{X_t,x}^\top (K_{X_t,X_t} + \tau I)^{-1} y_t,$$

$$k_t(x, x') = k(x, x') - k_{X_t,x}^\top (K_{X_t,X_t} + \tau I)^{-1} k_{X_t,x},$$

where $k_{X_t,x} = [k(x_1, x), k(x_2, x), \ldots, k(x_t, x)]^\top$ and $K_{X_t,X_t}$ is the $t \times t$ positive definite covariance matrix $[k(x_i, x_j)]_{i,j=1}^{t}$. The posterior variance of $\tilde{f}(x)$ is given by $\sigma_t^2(x) = k_t(x, x)$.

2.3 Bayesian Optimization Algorithms (GP-UCB and GP-TS)

GP-UCB relies on an optimistic upper confidence bound score to select the observation points. Specifically, at each time $t$, $x_t$ is selected as

$$x_t = \arg\max_{x \in \mathcal{X}} \mu_{t-1}(x) + \alpha_t \sigma_{t-1}(x),$$

where $\mu_{t-1}$ and $\sigma_{t-1}$ are the posterior mean and the standard deviation based on previous observations defined in §2.2, and $\alpha_t$ is a user-specified scaling parameter.
GP-TS selects the observation points by posterior sampling. Specifically, at each time $t$, a sample $\hat{f}_t(x)$ is drawn from a GP with mean $\mu_{t-1}$ and kernel $\alpha_t^2 k_{t-1}$ where $\mu_{t-1}$ and $k_{t-1}$ are the posterior mean and the posterior kernel based on previous observations defined in §2.2, and $\alpha_t$ is a user-specified scaling parameter. Then, $x_t$ is selected as

$$x_t = \arg\max_{x \in \mathcal{X}} \hat{f}_t(x).$$

The scaling parameters $\alpha_t$ are designed to ensure sufficient exploration of the search space and increase with $t$ ($\alpha_t > \alpha_{t'}$ when $t > t'$). See, e.g., Srinivas et al. (2010); Chowdhury and Gopalan (2017) for the specifications of $\alpha_t$.

### 2.4 Regularity Assumptions

The regret performance of the learning algorithms is analysed under two different settings, referred to as Bayesian and frequentist.

Under the Bayesian setting, $f$ is assumed to be a sample from a prior GP with kernel $k$. The observation noise $\{\epsilon_t\}_{t=1}^T$ are assumed to be i.i.d. zero mean Gaussian random variables with variance $\tau$.

Under the frequentist setting, $f$ is assumed to live in the RKHS corresponding to $k$. In particular, $\|f\|_{H_k} \leq B$, for some $B > 0$, where $\|\cdot\|_{H_k}$ denote the RKHS norm (see §3.1 for the definition of the RKHS norm). The observation noise are assumed to be i.i.d. sub-Gaussian random variables. Specifically, it is assumed that $\forall h \in \mathbb{R}, \forall t \in \mathbb{N}, E[\epsilon_t e^{h\epsilon_t}] \leq \exp(\frac{h^2 R^2}{2})$, for some $R > 0$. The sub-Gaussian assumption implies that $E[\epsilon_t] = 0$, for all $t$.

### 2.5 The Information Gain and The Upper Bounds on Regret

The regret analysis of GP-UCB and GP-TS consists of two main components. One is a bound on the information gain, and the other is a concentration inequality on random processes. The information gain is treated the same under both Bayesian and frequentist settings. Concentration inequalities that are utilized under each setting are however different. We first discuss the regret under the Bayesian setting, and then point out the difference with the frequentist setting.

Srinivas et al. (2010) showed that, under the Bayesian setting on a finite search space,

$$R(T; \text{GP-UCB}) = O \left( \alpha_T \left( T \sum_{t=1}^{T} \sigma_{t-1}^2(x) \right)^{\frac{1}{2}} \right)$$

where $\alpha_t = (2 \log(|\mathcal{X}|^3 \pi^2 / 6))$ is the scaling parameter of GP-UCB.

The same order of the regret bound was then extended to the case of a general compact $\mathcal{X}$ under the following technical assumption, which ensures a dense set approximating the GP sample can efficiently be constructed within $\mathcal{X}$.

**Assumption 1.** For some $a, b > 0$, for all $j = 1, 2, \ldots, d$,

$$\Pr[\sup_{x \in \mathcal{X}} |\partial \hat{f} / \partial x_j| > L] \leq a e^{-(L/b)^2}.$$ 

In particular, when Assumption 1 is satisfied, under a Bayesian setting on a compact search space $\mathcal{X}$,

$$R(T; \text{GP-UCB}) = O(\alpha_T \sqrt{D_T T})$$

where $\alpha_T = O(\sqrt{\log(T)})$ and $D_T$ denotes the cumulative variance at the observation points, i.e.,

$$D_T = \sum_{t=1}^{T} \sigma_{t-1}^2(x_t).$$ (2)
Sometimes referred to as the effective dimension of the problem (cf. Calandriello et al. (2019)), \( D_T \) can be interpreted as the total uncertainty (measured by variance) experienced by a learning algorithm.

The standard approach for bounding \( D_T \) is to use the information gain that refers to the mutual information \( I(y_t; \hat{f}) \) (Cover, 1999) between \( y_t \) and \( \hat{f} \). Using the closed form expression of mutual information between two multivariate Gaussian distributions, we know

\[
I(y_t; \hat{f}) = \frac{1}{2} \log \det(I_t + \frac{1}{\tau} K_{X_t,X_t}).
\]

Using Jensen inequality, Srinivas et al. (2010) proved that

\[
D_T \leq c_1 I(y_T; \hat{f})
\]

where \( c_1 = 2/\log(1+1/\tau) \) is an absolute constant.

It is noteworthy that a reverse inequality can also be proven; \( I(y_T; \hat{f}) \leq c_2 D_T \) where \( c_2 = 1/(2\tau) \), which implies that the scaling of \( D_T \) and \( I(y_T; \hat{f}) \) with \( T \) is of the same order.

Most existing work proceed by defining a kernel-specific and \( X_T \)-independent maximal information gain,

\[
\gamma_T = \sup_{X_T \subseteq \mathcal{X}} I(y_T; \hat{f}),
\]

and giving the regret bounds in terms of \( \gamma_T \):

\[
R(T; \text{GP-UCB}) = \mathcal{O}(\alpha_T \sqrt{\gamma_T T}).
\]

For specific kernels (Matérn and SE), Srinivas et al. (2010) proved upper bounds on \( \gamma_T \) which are commonly used to provide explicit regret bounds.

Under the frequentist setting, similar regret bounds with similar analysis follow; albeit, with a much larger \( \alpha_T = B + R \sqrt{2(\gamma_{t-1} + \log(t) + 1)} \), scaling with \( \mathcal{O}(\sqrt{\gamma_t}) \) and resulting in

\[
R(T; \text{GP-UCB}) = \mathcal{O}(\gamma_T \sqrt{T}).
\]

The difference between the Bayesian and frequentist settings is due to the difference in the concentration inequalities which hold for the GP (surrogate) models under the two settings. In particular, under the Bayesian setting, for a fixed \( x \in \mathcal{X} \), it directly follows from the standard bounds on the CDF of normal distributions that, with probability at least \( 1 - \delta \),

\[
|\hat{f}(x) - \mu_t(x)| \leq u_t(\delta) \sigma_t(x)
\]

where \( u_t(\delta) = \sqrt{2 \log(\frac{1}{\delta})} \). Chowdhury and Gopalan (2017) proved a similar concentration inequality under the frequentist setting, however, with a significantly larger \( u_t(\delta) = B + R \sqrt{2(\gamma_{t-1} + 1 + \log(1/\delta))} \) which scales with \( \mathcal{O}(\sqrt{\gamma_t}) \) and manifests in the multiplicative \( \mathcal{O}(\sqrt{\gamma_T}) \) difference between the regret bounds under the Bayesian and frequentist settings.

The contribution of this paper is to derive novel bounds on \( \gamma_T \) (consequently, also, on \( D_T \)) which immediately translate to the improved regret bounds under various settings.

### 3 Upper Bounds On The Information Gain

In this section, we present our bounds on the information gain. Our results are achieved by a finite dimensional projection of the GP model in the RKHS corresponding to \( k \). We start with laying out the details of the RKHS and the finite dimensional projection of the GP model. We then present the bounds on \( \gamma_T \).
3.1 RKHS and Mercer’s Theorem

Consider a positive definite kernel $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+$. A Hilbert space $H_k$ of functions on $\mathcal{X}$ equipped with an inner product $\langle \cdot, \cdot \rangle_{H_k}$ is called an RKHS with reproducing kernel $k$ if the following are satisfied. For all $x \in \mathcal{X}$, $k(\cdot, x) \in H_k$, and for all $x \in \mathcal{X}$ and $f \in H_k$, $\langle f, k(\cdot, x) \rangle_{H_k} = f(x)$ (reproducing property).

An RKHS is completely specified with its kernel function and vice-versa. The inner product induces the RKHS norm $\|f\|_{H_k}^2 = \langle f, f \rangle_{H_k}$ that can be interpreted as a measure of the complexity of $f$.

Mercer’s theorem provides an alternative representation for GP kernels as an inner product of infinite dimensional feature maps (e.g. see Theorem 4.1 in Kanagawa et al.).

**Theorem 1** (Mercer’s Theorem). Let $k$ be a continuous kernel. There exists $\{ (\lambda_m, \phi_m) \}_{m=1}^\infty$ such that $\lambda_m \in \mathbb{R}^+$, $\phi_m \in H_k$, for $m \geq 1$, and

$$ k(x, x') = \sum_{m=1}^\infty \lambda_m \phi_m(x) \phi_m(x'). $$

The $\{ \lambda_m \}_{m=1}^\infty$ and the $\{ \phi_m \}_{m=1}^\infty$ are referred to as the eigenvalues and the eigenfeatures (or eigen-functions) of $k$, respectively. Throughout the paper, we assume that $\{ \lambda_m \}_{m=1}^\infty$ are in a decreasing order: $\lambda_1 \geq \lambda_2 \geq \ldots$.

As a result of Mercer’s theorem, we can express a GP sample $\hat{f}$ in terms of a weight vector in the feature space of $k$

$$ \hat{f}(\cdot) = \sum_{m=1}^\infty W_m \phi_m^\bot(\cdot), $$

(6)

where the weights $W_m$ are i.i.d. random variables with standard normal $\mathcal{N}(0, 1)$ distribution. It is straightforward to check that $\hat{f}$ given in (6) is a zero mean GP with kernel $k$. We refer to this representation as the feature space representation in contrast to the function space representation presented in §2.2.

The RKHS can also be represented in terms of $\{ (\lambda_m, \phi_m) \}_{m=1}^\infty$ using Mercer’s representation theorem (e.g. see Theorem 4.2 in Kanagawa et al. (2018)).

**Theorem 2** (Mercer’s Representation Theorem). For a continuous kernel $k$, let $\{ (\lambda_m, \phi_m) \}_{m=1}^\infty$ be as in Theorem 1. Then, the RKHS of $k$ is given by

$$ H_k = \left\{ f(\cdot) = \sum_{m=1}^\infty w_m \lambda_m^\bot \phi_m(\cdot) : ||f||_{H_k} \triangleq \sum_{m=1}^\infty w_m^2 < \infty \right\}. $$

Mercer’s representation theorem provides an explicit definition for the RKHS norm. It also indicates that $\{ \lambda_m^\bot \phi_m \}_{m=1}^\infty$ form an orthonormal basis for $H_k$.

3.2 Projection On A Finite Dimensional Space

The feature space representation of typical GP kernels is infinite dimensional. To overcome the difficulty of working in infinite dimensional spaces, we use a projection $P_D$ on a $D$ dimensional RKHS consisting of the first $D$ features (corresponding to the $D$ largest eigenvalues of the kernel). Specifically, consider the $D$-dimensional feature space $\phi_D(\cdot) = [\phi_1(\cdot), \phi_2(\cdot), \ldots, \phi_D(\cdot)]^\top$, the $D$-dimensional column vector $W_D = [W_1, W_2, \ldots, W_D]^\top$ and the diagonal matrix $\Lambda_D = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_D)$ with $[\lambda_1, \lambda_2, \ldots, \lambda_D]$ as the diagonal elements. The projection of $\hat{f}$ on the $D$-dimensional space is given by

$$ P_D[\hat{f}(\cdot)] = W_D^\top \Lambda_D^\bot \phi_D(\cdot). $$
Notice that $P_D[\hat{f}]$ is a zero mean GP with kernel $k_P(x, x') = \sum_{m=1}^{D} \lambda_m \phi_m(x)\phi_m(x')$. We used the subscript $P$ to signify the space resulted from the projection. In addition, let $P_D^\perp[\hat{f}] = \hat{f} - P_D[\hat{f}]$ be the orthogonal part of $\hat{f}$ with respect to the projection. Notice that $P_D^\perp[\hat{f}]$ is also a GP with kernel $k_O(x, x') = k(x, x') - k_P(x, x')$. We used the subscript $O$ to signify the orthogonal part.

Let $\delta_D$ be an upper bound on the tail mass of eigenvalues of $k$. In particular, assume $\psi = \sup_{x \in \mathcal{X}, i \in \mathbb{N}} |\phi_i(x)|$ exists. We define

$$
\delta_D = \sum_{m=D+1}^{\infty} \lambda_m \psi^2.
$$

Since $\lambda_m$ decrease as $m$ grows, $\delta_D$ becomes arbitrarily small when $D$ is large enough. We have the following uniform bound on $k_O(x, x')$:

$$
\forall x, x' \in \mathcal{X}, \quad k_O(x, x') \leq \delta_D.
$$

### 3.3 Analysis of the Information Gain

We establish a novel upper bound on $\gamma_T$ in the following theorem.

**Theorem 3 (Bounding $\gamma_T$).** Consider a GP with a continuous kernel $k$ satisfying $\forall x, x' \in \mathcal{X}$, $k(x, x') \leq \tilde{k}$, for some $\tilde{k} > 0$. For $D \in \mathbb{N}$, let $\delta_D$ be as defined in (7). The following upper bound on $\gamma_T$, defined in (3), is satisfied for all $D \in \mathbb{N}$.

$$
\gamma_T \leq \frac{c_1}{2} D \log \left( 1 + \frac{\tilde{k}T}{\tau D} \right) + \frac{c_1}{2} \delta_D T.
$$

where $c_1 = 2/\log(1 + 1/\tau)$.

The expression can be simplified as

$$
\gamma_T = O(D \log(T) + \delta_D T).
$$

In contrast to the existing results, Theorem 3 provides an upper bound in terms of the spectral properties of the GP kernel through $\delta_D$ that is applicable to all kernels based on their eigendecay. Specializing this bound for common kernels (e.g., Matérn and SE) significantly improves the upper bounds on $\gamma_T$ (and consequently the upper bounds on regret) compared to the existing ones.

**Proof Sketch.** Recall $I(y_i; \hat{f}) = \frac{1}{2} \log \det(I_t + \frac{1}{2} K_{X_t, X_t})$. The problem is thus bounding the log-det of the covariance matrix $I_t + K_{X_t, X_t}$ for an arbitrary sequence $X_t$ of observation points. To achieve this, we use the $D$-dimensional projection in the RKHS. Recall $k = k_P + k_O$. Let us use the notations $K_P, X, x_i = [k_P(x_i, x_{j})]_{j=1}^{T}$ and $K_O, X, x_i = [k_O(x_i, x_{j})]_{j=1}^{T}$ to denote the corresponding covariance matrices. We show that $\log \det(I_t + K_{X_t, X_t})$ is bounded in terms of $\log \det(I_t + K_P, X_t, X_t)$ and a residual term depending on $K_O, X, x_i$. The finite dimensionality of the RKHS of $k_P$ allows us to use Weinstein–Aronszajn identity and the Gram matrix $G_t$ in the feature space of $k_P$ to bound $\log \det(I_t + K_P, X_t, X_t)$ in terms of $\log \det(I_D + G_t)$. Elementary calculation can be used to establish a bound on the log-det of a positive definite matrix in terms of its trace. Utilizing this result, we bound $\log \det(I_D + G_t)$ by $O(D \log(T))$. We then use the bound on the log-det of a positive definite matrix in terms of its trace, again, to bound the residual term depending on $K_O, X, x_i$ by $\delta_D T$, taking advantage of the property that $k_O(x, x') \leq \delta_D$. A detailed proof is given in Appendix A.

In comparison, Srinivas et al. (2010), in their analysis of the information gain, first showed that $I(y_i; \hat{f}) = \log \det(I_t + K_{X_t, X_t})$ is a submodular function in $X_t$. While finding the observation sequence that maximizes $I(y_T; \hat{f})$ is NP-hard (Ko et al., 1995), Srinivas et al. (2010) used the properties of submodular functions to show that $\gamma_T$ is within a constant factor of $\log \det(I_T + K_{\hat{X}_T, X_T})$ where $\hat{X}_T$ is a sequence of observation points that is selected, in a greedy fashion, to maximize $D_T$ (see (2)). Then, in order to bound $\log \det(I_T + K_{\hat{X}_T, X_T})$, they used the proximity of the eigen-
values of $K_{\tilde{X}_t, \tilde{X}_t}$ and those of the kernel $k$. In contrast, we directly work with the eigenvalues of $k$. The key idea in our analysis is the finite dimensional projection in the RKHS which allows us to bound the information gain without having to handle the complexities of the greedy observation sequence and the eigenvalues of its covariance matrix. In a related work to the approach of Srinivas et al. (2010), Seeger et al. (2008) proved bounds on $E[I(y_t, \hat{f})]$ where the expectation is taken with respect to a prior distribution on $X_t$. Those bounds are not applicable to the sequential optimization problem due to the difference in the design of $X_t$.

**Remark 1.** In order to find an explicit bound on $\gamma_T$, we increase $D$ such that $D \log(T)$ and $T \delta_D$ on the right hand side of (9) become of the same order. For such sufficiently large $D$, we have $\gamma_T = O(D \log(T))$, consequently, $D_T = O(D \log(T))$, which explains the use of the term effective dimension. That is to say the behavior of the kernel (with regards to the cumulative variance which is of the same order as the information gain) becomes similar to that of a finite $D$-dimensional kernel (up to a $\log(T)$ factor).

### 3.4 Conditions on the Eigendecay of the GP kernel

We now discuss the implications of Theorem 3, under conditions on the eigendecay of $k$. In particular, we define the following characteristic eigendecay profiles (which are similar to the ones in Chatterji et al. (2019)).

**Definition 1 (Polynomial and Exponential Eigendecay).** Consider the eigenvalues $\{\lambda_m\}_{m=1}^\infty$ of $k$ as given in Theorem 1 in a decreasing order.

1. For some $C_p > 0$, $\beta_p > 1$, $k$ is said to have a $(C_p, \beta_p)$ polynomial eigendecay, if for all $m \in \mathbb{N}$, we have $\lambda_m \leq C_p m^{-\beta_p}$.

2. For some $C_{e,1}, C_{e,2}, \beta_e > 0$, $k$ is said to have a $(C_{e,1}, C_{e,2}, \beta_e)$ exponential eigendecay, if for all $m \in \mathbb{N}$, we have $\lambda_m \leq C_{e,1} \exp(-C_{e,2} m^{\beta_e})$.

The following corollary is a consequence of Theorem 3.

**Corollary 1.** Consider $\gamma_T$ defined in (3).

If $k$ has a $(C_p, \beta_p)$ polynomial eigendecay, we have

$$\gamma_T \leq \frac{c_1}{2} \left((C_p \psi^2 T) \frac{1}{\beta_p} \log^{1-\frac{1}{\beta_p}} (1 + \frac{kT}{\tau}) + 1\right).$$

The expression can be simplified as $\gamma_T = O \left( T \frac{1}{\beta_p} \log^{1-\frac{1}{\beta_p}} (T) \right)$.

If $k$ has a $(C_{e,1}, C_{e,2}, \beta_e)$ exponential eigendecay, we have

$$\gamma_T \leq \left( \left( \frac{2}{C_{e,2}} (\log(T) + C_{\beta_e}) \right)^{\frac{1}{\beta_e}} + 1 \right) \log(1 + \frac{kT}{\tau}),$$

where $C_{\beta_e} = \log \left( \frac{C_{e,2}}{C_{e,2}^2} \right)$ if $\beta_e = 1$ and $C_{\beta_e} = \log \left( 2C_{e,1} \psi^2 + \left( \frac{1}{\beta_e} - 1 \right) \log \left( \frac{C_{e,2}}{C_{e,2}^2} \left( \frac{1}{\beta_e} - 1 \right) \right) \right)$, otherwise.

The expression can be simplified as $\gamma_T = O \left( \log^{1+\frac{1}{\beta_e}} (T) \right)$

Corollary 1 provides general bounds on $\gamma_T$ as long as the polynomial and exponential conditions on the eigendecay of $k$ are satisfied. A detailed proof is provided in Appendix B.

### 4 The Regret Bounds

In this section, we utilize the upper bounds on $\gamma_T$ established in Theorem 3 and Corollary 1 to derive regret bounds for GP-UCB and GP-TS.
Theorem 4 (Regret Bounds). Consider the sequential optimization problem given in §2.1 and GP-UCB with properly tuned $\alpha_t$. Assume $k$ has a $(C_p, \beta_p)$ polynomial eigendecay. Under the Bayesian setting, when Assumption 1 is satisfied:

$$R(T; \text{GP-UCB}) = O\left(T^{\frac{\beta_p+1}{2\beta_p}} \log^{1+\frac{1}{\gamma P}}(T)\right).$$

Under the frequentist setting:

$$R(T; \text{GP-UCB}) = O\left(T^{\frac{\beta_p+2}{2\beta_p}} \log^{1+\frac{1}{\gamma P}}(T)\right).$$

Assume $k$ has $(C_{e,1}, C_{e,2}, \beta_e)$ exponential eigendecay. Under the Bayesian setting, when Assumption 1 is satisfied:

$$R(T; \text{GP-UCB}) = O\left(T^{\frac{1}{2}} \log^{1+\frac{1}{\gamma P}}(T)\right).$$

Under the frequentist setting:

$$R(T; \text{GP-UCB}) = O\left(T^{\frac{1}{2}} \log^{1+\frac{1}{\gamma P}}(T)\right).$$

In comparison to the existing results (which either are implicit due to dependency on $\gamma_T$, or are given for particular kernels, e.g. Matérn and SE), our general approach in bounding $\gamma_T$, allows us to provide explicit regret bounds (in terms of $T$) as long as the conditions on the eigendecay of $k$ are satisfied.

Remark 2. The regret bounds for GP-TS are the same as the respective ones for GP-UCB up to a multiplicative $O(\sqrt{\log(T)})$ factor under all the cases above. For example, under the frequentist setting, if $k$ has exponential eigendecay $R(T; \text{GP-TS}) = O\left(T^{\frac{1}{2}} \log^{1+\frac{1}{\gamma P}}(T)\right)$ (Chowdhury and Gopalan, 2017; Kandasamy et al., 2018). The extra multiplicative $O(\sqrt{\log(T)})$ factor in the regret bounds of GP-TS in comparison to GP-UCB is due to a union bound step used because of posterior sampling (Chowdhury and Gopalan, 2017).

Matérn and squared exponential (SE) are perhaps the most popular kernels in practice for Bayesian optimization (cf. Snoek et al.; Shahriari et al.),

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

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

where $r = ||x - x'||_2$ is the Euclidean distance between $x$ and $x'$, $l > 0$ is referred to as lengthscale, $\nu > 0$ is referred to as the smoothness parameter, and $B_\nu$ is the modified Bessel function of the second kind. Variation over parameter $\nu$ creates a rich class of kernels. The SE kernel can also be interpreted as a special case of Matérn when $\nu \to \infty$.

It is known that, in the case of a Matérn kernel with smoothness parameter $\nu > \frac{1}{2}$, $\lambda_j = O(j^{-\frac{\nu}{2\nu-1}})$ (Sanit and Schaback, 2016); and, in the case of SE kernel, $\lambda_j = O(\exp(-j^2))$ (Belkin, 2018). Also, see Riutort-Mayol1 et al. (2020) which gave closed form expression of their eigenvalue-eigenfeature pairs on hypercubes. We now formally give the regret bounds with Matérn and SE kernels as a direct result of Theorem 4.

Corollary 2. Consider the sequential optimization problem (§2.1) and GP-UCB and GP-TS with properly tuned $\alpha_t$. Assume $k$ is a Matérn kernel with smoothness parameter $\nu > 2$. Under the

3See Srinivas et al. (2010); Chowdhury and Gopalan (2017) for the specification of $\alpha_t$, which scales with $(O(\sqrt{\log(T)})$ and $O(\sqrt{\gamma T})$ under the Bayesian and frequentist settings, respectively.
Bayesian setting:

\[ R(T; \text{GP-UCB}) = \mathcal{O}\left(T^{\frac{\nu+1}{2\nu}} \log(T)\right). \]

Under the frequentist setting:

\[ R(T; \text{GP-UCB}) = \mathcal{O}\left(T^{\frac{\nu+1}{2\nu}} \log(T)\right). \]

Assume \( k \) is a SE kernel. Under the Bayesian setting:

\[ R(T; \text{GP-UCB}) = \mathcal{O}\left(T^{\frac{1}{2}}(\log(T))^{\frac{1}{2d}+1}\right). \]

Under the frequentist setting:

\[ R(T; \text{GP-UCB}) = \mathcal{O}\left(T^{\frac{1}{2}}(\log(T))^{d+1}\right). \]

The assumption \( \nu > 2 \) ensures that Assumption 1 used in Theorem 4 is satisfied (Srinivas et al., 2010).

In comparison to the existing work, our results significantly improve the regret bounds for the Matérn kernel, reducing the worst case gap with the \( \Omega(T^{\frac{\nu+1}{2\nu}}) \) lower bounds (Scarlett et al., 2017) to \( \mathcal{O}(\log(T)) \) (from \( \mathcal{O}(\sqrt{T}) \) in the existing work), under the Bayesian setting.

5 Conclusion

We introduced a new and general approach to bounding the information gain in Bayesian optimization problems. Using Mercer’s theorem, we decoupled the GP kernel into a finite \( D \)-dimensional projection in the corresponding RKHS and its orthogonal part. In the finite dimensional space, \( \gamma_T \) is bounded by \( D \log(T) \). Accounting for the effect of the orthogonal element using the eigendecay of the GP kernel completes the upper bound. We provided explicit bounds in terms of \( T \) on \( \gamma_T \) and the regret of GP-UCB and GP-TS under conditions on the eigendecay of the kernel which directly apply to common kernels such as Matérn and SE and show significant improvements over the state of the art.

Our results establish the first tight regret bounds (up to a \( \log(T) \) factor) with the Matérn kernel under the Bayesian setting which shows our bound on \( \gamma_T \) is tight (up to a \( \log(T) \) factor). Under the frequentist setting, although our regret bounds improve the existing ones, the question of optimal regret bounds remains open.

We, similar to Scarlett et al. (2017), conjecture that \( \tilde{O}(\sqrt{T}) \) regret bounds are provable under the frequentist setting. If true, the regret bounds under the frequentist setting will also become optimal (up to a \( \log(T) \) factor), with our bounds on \( \gamma_T \). The difference between the two settings is in the concentration inequalities for the GP models used under each one (cf. Theorem 2 of Chowdhury and Gopalan (2017)). The proof of stronger concentration inequalities applicable to the elements of RKHS is challenging and requires a separate investigation. We consider that as an interesting open problem.

References

Y. Abbasi-Yadkori, D. Pál, and C. Szepesvári. Improved algorithms for linear stochastic bandits. In Advances in Neural Information Processing Systems, pages 2312–2320, 2011.

M. Abeille and A. Lazaric. Linear thompson sampling revisited. In Proceedings of Machine Learning Research, volume 54, pages 176–184, 2017.
S. Agrawal and N. Goyal. Thompson sampling for contextual bandits with linear payoffs. In *International Conference on Machine Learning*, pages 127–135, 2013.

P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time analysis of the multiarmed bandit problem. *Mach. Learn.*, 47(2–3):235–256, May 2002. ISSN 0885-6125. doi: 10.1023/A:1013689704352. URL https://doi.org/10.1023/A:1013689704352.

M. Belkin. Approximation beats concentration? an approximation view on inference with smooth radial kernels. In *Conference On Learning Theory, COLT 2018, Stockholm, Sweden, 6-9 July 2018*, volume 75 of *Proceedings of Machine Learning Research*, pages 1348–1361. PMLR, 2018. URL http://proceedings.mlr.press/v75/belkin18a.html.

F. Berkenkamp, A. Krause, and A. P. Schoellig. Bayesian optimization with safety constraints: safe and automatic parameter tuning in robotics. *arXiv preprint arXiv:1602.04450*, 2016.

I. Bogunovic, A. Krause, and J. Scarlett. Corruption-tolerant gaussian process bandit optimization. *arXiv preprint arXiv:2003.01971*, 2020.

D. Calandriello, L. Carratino, A. Lazaric, M. Valko, and L. Rosasco. Gaussian process optimization with adaptive sketching: Scalable and no regret. In *Proceedings of the Thirty-Second Conference on Learning Theory*, volume 99 of *Proceedings of Machine Learning Research*, Phoenix, USA, 25–28 Jun 2019. PMLR.

N. Chatterji, A. Pacchiano, and P. Bartlett. Online learning with kernel losses. In *Proceedings of Machine Learning Research*, volume 97, pages 971–980, Long Beach, California, USA, 09–15 Jun 2019. PMLR. URL http://proceedings.mlr.press/v97/chatterji19a.html.

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

T. M. Cover. *Elements of information theory*. John Wiley & Sons, 1999.

V. Dani, T. P. Hayes, and S. M. Kakade. Stochastic linear optimization under bandit feedback. In *COLT*, pages 355–366, 2008.

J. Djolonga, A. Krause, and V. Cevher. High-dimensional gaussian process bandits. In *Advances in Neural Information Processing Systems 26*, pages 1025–1033. Curran Associates, Inc., 2013. URL http://papers.nips.cc/paper/5152-high-dimensional-gaussian-process-bandits.pdf.

Y. Fang, K. A. Loparo, and X. Feng. Inequalities for the trace of matrix product. *IEEE Transactions on Automatic Control*, 39(12):2489–2490, 1994.

T. Javidi and S. Shekhar. Gaussian process bandits with adaptive discretization. *Electron. J. Statist.*, 12(2):3829–3874, 2018.

M. Kanagawa, P. Hennig, D. Sejdinovic, and B. K. Sriperumbudur. Gaussian processes and kernel methods: A review on connections and equivalences. *Available at Arxiv*, 2018.

K. Kandasamy, A. Krishnamurthy, J. Schneider, and B. Póczos. Parallelised bayesian optimisation via thompson sampling. In *International Conference on Artificial Intelligence and Statistics*, pages 133–142, 2018.

K. Kandasamy, G. Dasarathy, J. Oliva, J. Schneider, and B. Poczos. Multi-fidelity gaussian process bandit optimisation. *Journal of Artificial Intelligence Research*, 66:151–196, 2019.

C.-W. Ko, J. Lee, and M. Queyranne. An exact algorithm for maximum entropy sampling. *Operations Research*, 43(4):684–691, 1995.
A. Krause and C. S. Ong. Contextual gaussian process bandit optimization. In Advances in Neural Information Processing Systems 24, pages 2447–2455. Curran Associates, Inc., 2011. URL http://papers.nips.cc/paper/4487-contextual-gaussian-process-bandit-optimization.pdf.

M. Mutny and A. Krause. Efficient high dimensional bayesian optimization with additivity and quadrature fourier features. In Advances in Neural Information Processing Systems 31, pages 9005–9016. Curran Associates, Inc., 2018.

V. Picheny, S. Vakili, and A. Artemev. Ordinal bayesian optimisation. arXiv preprint arXiv:1912.02493, 2019.

C. E. Rasmussen and C. K. Williams. Gaussian Processes for Machine Learning. MIT Press, 2006.

G. Riutort-Mayol1, P.-C. Burkner, M. R. Andersen, A. Solin, and A. Vehtari. Practical hilbert space approximate bayesian gaussian processes for probabilistic programming. Available at Arxiv., 2020.

P. Rusmevichientong and J. N. Tsitsiklis. Linearly parameterized bandits. Mathematics of Operations Research, 35(2):395–411, 2010.

D. Russo and B. Van Roy. Learning to optimize via posterior sampling. Mathematics of Operations Research, 39(4):1221–1243, 2014.

D. Russo and B. Van Roy. An information-theoretic analysis of thompson sampling. The Journal of Machine Learning Research, 17(1), 2016.

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

J. Scarlett, I. Bogunovic, and V. Cevher. Lower bounds on regret for noisy Gaussian process bandit optimization. In Proceedings of the 2017 Conference on Learning Theory, volume 65 of Proceedings of Machine Learning Research, pages 1723–1742, Amsterdam, Netherlands, 07–10 Jul 2017. PMLR. URL http://proceedings.mlr.press/v65/scarlett17a.html.

M. W. Seeger, S. M. Kakade, and D. P. Foster. Information consistency of nonparametric gaussian process methods. IEEE Transactions on Information Theory, 54(5):2376–2382, 2008.

B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas. Taking the human out of the loop: A review of bayesian optimization. Proceedings of the IEEE, 104(1):148–175, 2016.

A. Slivkins. Introduction to multi-armed bandits. arXiv preprint arXiv:1904.07272, 2019.

J. Snoek, H. Larochelle, and R. P. Adams. Practical bayesian optimization of machine learning algorithms. In Advances in Neural Information Processing Systems 25, pages 2951–2959. Curran Associates, Inc., 2012.

N. Srinivas, A. Krause, S. Kakade, and M. Seeger. Gaussian process optimization in the bandit setting: no regret and experimental design. In Proceedings of the 27th International Conference on International Conference on Machine Learning, pages 1015–1022. Omnipress, 2010.

Y. Sui, V. Zhuang, J. W. Burdick, and Y. Yue. Stagewise safe bayesian optimization with gaussian processes. arXiv preprint arXiv:1806.07555, 2018.

S. Vakili, V. Picheny, and A. Artemev. Scalable thompson sampling using sparse gaussian process models. Available at Arxiv., 2020.

Z. Wang, B. Kim, and L. P. Kaelbling. Regret bounds for meta bayesian optimization with an unknown gaussian process prior. In Advances in Neural Information Processing Systems, pages 10477–10488, 2018.
We bound \( I(y_t; \hat{f}) = \frac{1}{2} \log \det(I_t + K_{X_t,X_t}) \) for an arbitrary observation sequence \( X_t \). Recall \( k = k_P + k_O \) and the corresponding covariance matrices \( K_{P,X_t,X_t} = [k_P(x_i,x_j)]_{i,j=1}^T \) and \( K_{O,X_t,X_t} = [k_O(x_i,x_j)]_{i,j=1}^T \). We have

\[
I(y_t; \hat{f}) = \frac{1}{2} \log \det(I_t + \frac{1}{\tau}K_{P,X_t,X_t}) \]

\[
= \frac{1}{2} \log \det \left( I_t + \frac{1}{\tau}(K_{P,X_t,X_t} + K_{O,X_t,X_t}) \right) \]

\[
= \frac{1}{2} \log \det \left( (I_t + \frac{1}{\tau}K_{P,X_t,X_t})(I_t + (I_t + \frac{1}{\tau}K_{P,X_t,X_t})^{-1}K_{O,X_t,X_t}) \right) \]

\[
= \frac{1}{2} \log \det(I_t + \frac{1}{\tau}K_{P,X_t,X_t}) + \frac{1}{2} \log \det \left( I_t + (I_t + \frac{1}{\tau}K_{P,X_t,X_t})^{-1}K_{O,X_t,X_t} \right), \tag{11}
\]

where for the last line we used \( \det(AB) = \det(A) \det(B) \) which holds for all two square matrices of the same dimensions. The equation (11) decouples the \( \log \det \) of the covariance matrix corresponding to \( k \) into that of \( k_P \) and a residual term depending on \( k_O \). We now proceed to bounding the two terms on the right hand side of (11).

The first term on the right hand side of (11) corresponds to the covariance matrix of the \( D \)-dimensional \( P \). We can upper bound this term using a bound on the \( \log \det \) of the Gram matrix. Let us define \( \Phi_{t,D} = [\phi_D(x_1), \phi_D(x_2), \ldots, \phi_D(x_t)]^\top \), a \( t \times D \) matrix which stacks the feature vectors \( \phi_D(x_s) \), \( s = 1, \ldots, t \), at the observation points, as its rows. Notice that

\[
K_{P,X_t,X_t} = \Phi_{t,D} \Lambda_{D} \Phi_{t,D}^\top. \]

Consider the Gram matrix

\[
G_t = \Lambda_{D}^{\frac{1}{2}} \Phi_{t,D}^\top \Phi_{t,D} \Lambda_{D}^{\frac{1}{2}}.
\]

By Weinstein–Aronszajn identity (that is a special case of matrix determinant lemma)

\[
\det(I_D + \frac{1}{\tau}G_t) = \det(I_t + \frac{1}{\tau}K_{P,X_t,X_t}). \tag{12}
\]

We can prove the following lemma on the relation between the \( \log \det \) and the trace of a positive definite matrix.

**Lemma 1.** For all positive definite matrices \( P \in \mathbb{R}^{n \times n} \), we have

\[
\log \det(P) \leq n \log(tr(P)/n).
\]

See the proof at the end of the appendix.

---

Q. Zhao. Multi-armed bandits: Theory and applications to online learning in networks. *Synthesis Lectures on Communication Networks*, 12(1):1–165, 2019.
We next bound the trace of $I_D + \frac{1}{\tau} G_t$. Notice that, for all $x \in \mathcal{X}$,

$$\|\phi_D(x)\Lambda_D^\frac{1}{2}\|_2^2 = \sum_{m=1}^D \lambda_m \phi_m^2(x) = k_p(x, x) \leq \bar{k}.$$ 

Thus,

$$\text{tr}(I_D + \frac{1}{\tau} G_t) = D + \frac{1}{\tau} \text{tr} \left( \sum_{s=1}^t \Lambda^\frac{1}{2} \phi_D(x_s) \phi_D^\top(x_s) \Lambda^\frac{1}{2} \right)$$

$$= D + \frac{1}{\tau} \sum_{s=1}^t \text{tr} \left( \Lambda^\frac{1}{2} \phi_D(x_s) \phi_D^\top(x_s) \Lambda^\frac{1}{2} \right)$$

$$= D + \frac{1}{\tau} \sum_{s=1}^t \text{tr} \left( \Lambda^\frac{1}{2} \phi_D^\top(x_s) \phi_D(x_s) \Lambda^\frac{1}{2} \right)$$

$$= D + \frac{1}{\tau} \sum_{s=1}^t \|\phi_D(x_s)\Lambda^\frac{1}{2}\|_2^2$$

$$\leq D + \frac{tk}{\tau}. $$

For the first line we expanded the Gram matrix, the second line holds by distributivity of trace over sum, and the third line is a result of $\text{tr}(AA^\top) = \text{tr}(A^\top A)$ which holds for any matrix $A$.

Using Lemma 1 and (12), we have

$$\log \det (I_t + \frac{1}{\tau} K_{P,X_t,X_t}) = \log \det (I_D + \frac{1}{\tau} G_t) \leq D \log \left( \frac{\text{tr}(I_D + \frac{1}{\tau} G_t)}{D} \right) = D \log(1 + \frac{tk}{\tau D}).$$

(13)

To upper bound the second term on the right hand side of (11), we use $k_O(x, x') \leq \delta_D$. Notice that $(I_t + \frac{1}{\tau} K_{P,X_t,X_t})^{-1}$ is a positive definite matrix whose largest eigenvalue is upper bounded by 1. For two positive definite matrices $A, B$ with the same dimensions, we have $\text{tr}(AB) \leq \lambda_A \text{tr}(B)$ where $\lambda_A$ is the largest eigenvalue of $A$ (cf. Fang et al. (1994)). Thus

$$\text{tr} \left( I_t + \frac{1}{\tau} K_{P,X_t,X_t} \right)^{-1} K_{O,X_t,X_t} \leq \text{tr}(K_{O,X_t,X_t}).$$

Since $\forall x, x' \in \mathcal{X}, k_O(x, x') \leq \delta_D$, we have $\text{tr}(K_{O,X_t,X_t}) \leq \delta_D$. Therefore,

$$\text{tr} \left( I_t + \frac{1}{\tau} K_{P,X_t,X_t} \right)^{-1} K_{O,X_t,X_t} \leq t (1 + \delta_D).$$

Using Lemma 1, we have

$$\log \det \left( I_t + \frac{1}{\tau} K_{P,X_t,X_t} \right)^{-1} K_{O,X_t,X_t} \leq t \log \left( \frac{t(1 + \delta_D)}{t} \right) = t \log(1 + \delta_D) \leq t \delta_M,$$

(14)

where for the last line we used $\log(1 + z) \leq z$ which holds for all $z \in \mathbb{R}$. 

16
Putting (11), (13) and (14) together, we arrive at the following bound on the information gain.

\[
I(y_t; \hat{f}) \leq \frac{1}{2} D \log(1 + \frac{\bar{k} t}{\tau D}) + \frac{1}{2} t \delta_D,
\]

which holds for any arbitrary sequence \(X_t \subseteq X\). Thus

\[
\gamma_T = \sup_{X_t \subseteq X} I(y_t; \hat{f}) \leq \frac{1}{2} D \log(1 + \frac{\bar{k} t}{\tau D}) + \frac{1}{2} t \delta_D.
\]

**Proof of Lemma 1.** Let \(\{\kappa_m > 0\}_{m=1}^n\) denote the eigenvalues of \(P\). Using the inequality of arithmetic and geometric means

\[
\prod_{m=1}^n \kappa_m \leq \left(\frac{1}{n} \sum_{m=1}^n \kappa_m\right)^n.
\]

Thus,

\[
\log \det(P) = \log \left(\prod_{m=1}^n \kappa_m\right) \leq \log \left(\frac{1}{n} \sum_{m=1}^n \kappa_m\right)^n = \log \left(\frac{\text{tr}(P)}{n}\right)^n = n \log \left(\frac{\text{tr}(P)}{n}\right).
\]

\(\square\)

**B (Proof of Corollary 1)**

Under the \((C_p, \beta_p)\) polynomial eigendecay condition, the following bound on \(\delta_D\) is straightforwardly derived from the decay rate of \(\lambda_m\).

\[
\delta_D = \sum_{m=D+1}^\infty \lambda_m \psi^2 \leq \sum_{m=D+1}^\infty C_p m^{-\beta_p} \psi^2 \leq \int_{z=D}^{\infty} C_p z^{-\beta_p} \psi^2 dz = C_p D^{1-\beta_p} \psi^2.
\]

We select \(D = \lceil (C_p \psi^2 T) \frac{\beta}{\beta} \log^{-\beta} (1 + \frac{\bar{k} T}{\tau}) \rceil\) which results in the lowest growth rate for \(D_T\). Theorem 3 implies

\[
D_T \leq \frac{C_1}{2} \left( (C_p \psi^2 T) \frac{\beta}{\beta} \log^{1-\beta} (1 + \frac{\bar{k} T}{\tau}) + 1 \right).
\]
Under the \((C_{e,1}, C_{e,2}, \beta_e)\) exponential eigendecay condition,

\[
\delta_D = \sum_{m=D+1}^{\infty} \lambda_m \psi^2
\leq \sum_{m=D+1}^{\infty} C_{e,1} \exp(-C_{e,2} m^{\beta_e}) \psi^2
\leq \int_{z=D}^{\infty} C_{e,1} \exp(-C_{e,2} z^{\beta_e}) \psi^2 dz.
\]

Now, consider two different cases of \(\beta_e = 1\) and \(\beta_e \neq 1\). When \(\beta_e = 1\),

\[
\int_{z=D}^{\infty} \exp(-C_{e,2} z^{\beta_e}) dz = \int_{z=D}^{\infty} \exp(-C_{e,2} z^{1/2}) dz
= \frac{1}{C_{e,2}} \exp(-C_{e,2} D).
\]

When \(\beta_e \neq 1\), we have

\[
\int_{z=D}^{\infty} \exp(-C_{e,2} z^{\beta_e}) dz = \frac{1}{\beta_e} \int_{z=D^{1/\beta_e}}^{\infty} \left(\frac{2}{C_{e,2}} \left(\frac{1}{\beta_e} - 1\right)\right)^{1/\beta_e - 1} \exp(-\left(\frac{1}{\beta_e} - 1\right)) \exp(-C_{e,2} z^{1/2}) dz
\]

\[
\leq \frac{2}{C_{e,2} \beta_e} \left(\frac{2}{C_{e,2}} \left(\frac{1}{\beta_e} - 1\right)\right)^{1/\beta_e - 1} \exp(-\left(\frac{1}{\beta_e} - 1\right)) \exp(-C_{e,2} D^{1/\beta_e}).
\]

The first equality is obtained by a change of parameter. The inequality holds since

\[
\max_{z \in \mathbb{R}} z^{1/\beta_e - 1} \exp(-C_{e,2} z^{1/2}) = \left(\frac{2}{C_{e,2}} \left(\frac{1}{\beta_e} - 1\right)\right)^{1/\beta_e - 1} \exp(-\left(\frac{1}{\beta_e} - 1\right))
\]

which can be verified using the standard method of equating the derivative of the left hand side to zero.

When \(\beta_e = 1\), we select

\[
D = \left\lceil \frac{1}{C_{e,2}} \log\left(\frac{C_{e,1} \psi^2 T}{C_{e,2}}\right)\right\rceil.
\]

When \(\beta_e \neq 1\), we select

\[
D = \left\lceil \left(\frac{2}{C_{e,2}} \left(\log(T) + \log\left(\frac{2C_{e,1} \psi^2}{\beta_e C_{e,2}}\right) + \left(\frac{1}{\beta_e} - 1\right) \left(\log\left(\frac{2}{C_{e,2}} \left(\frac{1}{\beta_e} - 1\right)\right) - 1\right)\right)\right)^{\frac{1}{\beta_e}} + 1\right\rceil.
\]

Theorem 3 implies

\[
D_T \leq \left(\frac{2}{C_{e,2}} \left(\log(T) + C_{\beta_e}\right)^{\frac{1}{\beta_e}} + 1\right) \log(1 + \frac{kT}{r}),
\]

\(C_{\beta_e} = \log\left(\frac{C_{e,1} \psi^2}{C_{e,2}}\right)\) when \(\beta_e = 1\), and \(C_{\beta_e} = \log\left(\frac{2C_{e,1} \psi^2}{\beta_e C_{e,2}}\right) + \left(\frac{1}{\beta_e} - 1\right) \left(\log\left(\frac{2}{C_{e,2}} \left(\frac{1}{\beta_e} - 1\right)\right) - 1\right)\), otherwise.