Improved Regret Bounds for Gaussian Process Upper Confidence Bound in Bayesian Optimization

Shogo Iwazaki

LY Corporation

1 Introduction

2 Preliminaries

3 Improved Regret Bounds for GP-UCB

4 Possible Future Directions and Conclusion

Bayesian Optimization

- A sequential decision-making framework for optimizing a **black-box objective function**
  - Possibly expensive-to-evaluate and non-convex

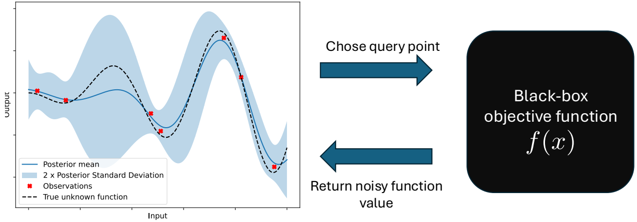


Figure: Illustrative image of Bayesian optimization

Bayesian Optimization

- A sequential decision-making framework for optimizing a **black-box objective function**
  - Possibly expensive-to-evaluate and non-convex

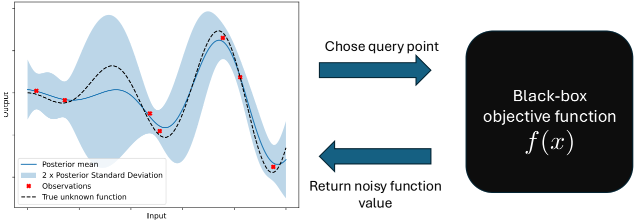


Figure: Illustrative image of Bayesian optimization

Goal: Design efficient adaptive query algorithms based on Gaussian process (GP) prediction.

Gaussian Upper Confidence Bound (GP-UCB)

Gaussian process upper confidence bound (GP-UCB) [Srinivas et al., 2010]

- An algorithm which combines GP prediction with optimism principle
- At round t, GP-UCB defines the query point $\boldsymbol{x}_t$ as the maximizer of the GP-based UCB score:

$$\boldsymbol{x}_t \in \operatorname*{argmax}_{\boldsymbol{x} \in \mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \sqrt{\beta_t}\, \sigma_{t-1}(\boldsymbol{x})$$

  - μ_{t-1} and σ_{t-1} are the GP posterior mean and standard deviation
  - β_t: confidence width parameter

Gaussian process upper confidence bound (GP-UCB) [Srinivas et al., 2010]

- An algorithm which combines GP prediction with optimism principle
- At round t, GP-UCB defines the query point $\boldsymbol{x}_t$ as the maximizer of the GP-based UCB score:

$$\boldsymbol{x}_t \in \operatorname*{argmax}_{\boldsymbol{x} \in \mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \sqrt{\beta_t}\, \sigma_{t-1}(\boldsymbol{x})$$

  - μ_{t-1} and σ_{t-1} are the GP posterior mean and standard deviation
  - β_t: confidence width parameter
- GP-UCB plays an important role in the advancement of Bayesian optimization, since:
  1. algorithm construction is simple and easy to extend to other advanced settings
  2. sophisticated theory was provided by Srinivas et al. [2010]

Gaussian process upper confidence bound (GP-UCB) [Srinivas et al., 2010]

- An algorithm which combines GP prediction with optimism principle
- At round t, GP-UCB defines the query point $\boldsymbol{x}_t$ as the maximizer of the GP-based UCB score:

$$\boldsymbol{x}_t \in \operatorname*{argmax}_{\boldsymbol{x} \in \mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \sqrt{\beta_t}\, \sigma_{t-1}(\boldsymbol{x})$$

  - μ_{t-1} and σ_{t-1} are the GP posterior mean and standard deviation
  - β_t: confidence width parameter
- GP-UCB plays an important role in the advancement of Bayesian optimization, since:
  1. algorithm construction is simple and easy to extend to other advanced settings
  2. sophisticated theory was provided by Srinivas et al. [2010]
- The existence of algorithms with theoretically stronger regret guarantees than GP-UCB has already been shown [Scarlett, 2018]

Gaussian process upper confidence bound (GP-UCB) [Srinivas et al., 2010]

- An algorithm which combines GP prediction with optimism principle
- At round t, GP-UCB defines the query point $\boldsymbol{x}_t$ as the maximizer of the GP-based UCB score:

$$\boldsymbol{x}_t \in \operatorname*{argmax}_{\boldsymbol{x} \in \mathcal{X}} \mu_{t-1}(\boldsymbol{x}) + \sqrt{\beta_t}\, \sigma_{t-1}(\boldsymbol{x})$$

  - μ_{t-1} and σ_{t-1} are the GP posterior mean and standard deviation
  - β_t: confidence width parameter
- GP-UCB plays an important role in the advancement of Bayesian optimization, since:
  1. algorithm construction is simple and easy to extend to other advanced settings
  2. sophisticated theory was provided by Srinivas et al. [2010]
- The existence of algorithms with theoretically stronger regret guarantees than GP-UCB has already been shown [Scarlett, 2018]

Question: Can we obtain improved regret bounds for GP-UCB?

Summary of Results

- **Improved high-probability regret bounds for the original GP-UCB algorithm**

- **Improved high-probability regret bounds for the original GP-UCB algorithm**

	Regret (Squared exponential kernel)	Regret (ν-Matérn kernel)
GP-UCB [Srinivas et al., 2010]	$O(\sqrt{T \ln^{d+2} T})$	$\widetilde{O}(T^{\frac{\nu+d}{2\nu+d}})$
GP-UCB (ours, improved analysis)	$O(\sqrt{T \ln^{2} T})$	$\widetilde{O}(\sqrt{T})$

  - **Strictly improve** the existing bounds of Srinivas et al. [2010]
  - **Comparable** to the state-of-the-art $\widetilde{O}(\sqrt{T})$ regret of Scarlett [2018] (up to polylog factors), while keeping the simple and popular GP-UCB rule

- **Improved high-probability regret bounds for the original GP-UCB algorithm**

	Regret (Squared exponential kernel)	Regret (ν-Matérn kernel)
GP-UCB [Srinivas et al., 2010]	$O(\sqrt{T \ln^{d+2} T})$	$\widetilde{O}(T^{\frac{\nu+d}{2\nu+d}})$
GP-UCB (ours, improved analysis)	$O(\sqrt{T \ln^{2} T})$	$\widetilde{O}(\sqrt{T})$

  - **Strictly improve** the existing bounds of Srinivas et al. [2010]
  - **Comparable** to the state-of-the-art $\widetilde{O}(\sqrt{T})$ regret of Scarlett [2018] (up to polylog factors), while keeping the simple and popular GP-UCB rule
- **Key technical idea:**
  - Refine the existing **worst-case** treatment of information gain
    - Study **algorithm- and sample-path dependent** behavior of information gain
  - Beyond GP-UCB in the standard setting, we expect that this analysis will motivate a reconsideration of other information gain-based guarantees

Next Section

1 Introduction

2 Preliminaries

3 Improved Regret Bounds for GP-UCB

4 Possible Future Directions and Conclusion

Notation

- $f : \mathcal{X} \to \mathbb{R}$: Objective function on $\mathcal{X} := [0, r]^d$
- $\boldsymbol{x}_t$: Learner's query point at step t (chosen based on GP-UCB)
- $y_t := f(\boldsymbol{x}_t) + \epsilon_t$: Learner's observed function value
- $R_T := \sum_{t=1}^{T} f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t)$: Cumulative regret up to step T
  - $\boldsymbol{x}^* \in \arg\max_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x})$: Maximizer of the function

Notation

- $f : \mathcal{X} \to \mathbb{R}$: Objective function on $\mathcal{X} := [0, r]^d$
- $\boldsymbol{x}_t$: Learner's query point at step t (chosen based on GP-UCB)
- $y_t := f(\boldsymbol{x}_t) + \epsilon_t$: Learner's observed function value
- $R_T := \sum_{t=1}^{T} f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t)$: Cumulative regret up to step T
    - $\boldsymbol{x}^* \in \arg\max_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x})$: Maximizer of the function

Assumpsions

- This paper studies **Bayesian setting**

Notation

- $f : \mathcal{X} \to \mathbb{R}$: Objective function on $\mathcal{X} := [0, r]^d$
- $\boldsymbol{x}_t$: Learner's query point at step t (chosen based on GP-UCB)
- $y_t := f(\boldsymbol{x}_t) + \epsilon_t$: Learner's observed function value
- $R_T := \sum_{t=1}^{T} f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t)$: Cumulative regret up to step T
  - $\boldsymbol{x}^* \in \arg\max_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x})$: Maximizer of the function

Assumpsions

- This paper studies **Bayesian setting**
  - **Assumption 1**: $f \sim \mathcal{GP}(0, k)$ with known positive definite kernel $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$
  - **Assumption 2**: $(\epsilon_t)_{t \in \mathbb{N}_+}$ are independent and identically distributed with $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$

Notation

- $f : \mathcal{X} \to \mathbb{R}$: Objective function on $\mathcal{X} := [0, r]^d$
- $\boldsymbol{x}_t$: Learner's query point at step t (chosen based on GP-UCB)
- $y_t := f(\boldsymbol{x}_t) + \epsilon_t$: Learner's observed function value
- $R_T := \sum_{t=1}^{T} f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t)$: Cumulative regret up to step T
  - $\boldsymbol{x}^* \in \arg\max_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x})$: Maximizer of the function

Assumpsions

- This paper studies **Bayesian setting**
  - **Assumption 1**: $f \sim \mathcal{GP}(0, k)$ with known positive definite kernel $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$
  - **Assumption 2**: $(\epsilon_t)_{t \in \mathbb{N}_+}$ are independent and identically distributed with $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$
- Focuses on the **squared exponential kernel** k_{SE} and ν-**Matérn kernel** $k_{\mathrm{Matérn}}$:

$$k_{\mathrm{SE}}(\boldsymbol{x}, \widetilde{\boldsymbol{x}}) = \exp\left(-\frac{\|\boldsymbol{x} - \widetilde{\boldsymbol{x}}\|_2^2}{2\ell^2}\right), \tag{1}$$

$$k_{\mathrm{Matérn}}(\boldsymbol{x}, \widetilde{\boldsymbol{x}}) = \frac{2^{1-\nu}}{\Gamma(\nu)}\left(\frac{\sqrt{2\nu}\,\|\boldsymbol{x} - \widetilde{\boldsymbol{x}}\|_2}{\ell}\right)^{\nu} K_\nu\left(\frac{\sqrt{2\nu}\,\|\boldsymbol{x} - \widetilde{\boldsymbol{x}}\|_2}{\ell}\right) \tag{2}$$

  - K_ν: Modified Bessel function of the second kind, ℓ: Length-scale parameter

Classical information-gain based analysis [Srinivas et al., 2010]

- Let us define $I(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)$ as the following quantity, called **information gain**:

$$I(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T) = \frac{1}{2} \log \det\left(\boldsymbol{I}_T + \sigma^{-2} \mathbf{K}_T\right).$$

  - $\mathbf{K}_T := [k(\boldsymbol{x}_i, \boldsymbol{x}_j)]_{1 \le i,j \le T}$: Gram matrix for inputs $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T$

Classical information-gain based analysis [Srinivas et al., 2010]

- Let us define $I(\boldsymbol{x}_1, \dots, \boldsymbol{x}_T)$ as the following quantity, called **information gain**:

$$I(\boldsymbol{x}_1, \dots, \boldsymbol{x}_T) = \frac{1}{2} \log \det \left(\boldsymbol{I}_T + \sigma^{-2} \mathbf{K}_T \right).$$

  - $\mathbf{K}_T \coloneqq [k(\boldsymbol{x}_i, \boldsymbol{x}_j)]_{1 \le i,j \le T}$: Gram matrix for inputs $\boldsymbol{x}_1, \dots, \boldsymbol{x}_T$

- Due to the UCB selection rule and the confidence bound of GP,

$$R_T \underset{\text{w.h.p.}}{\lesssim} 2\beta_T \sum_{t=1}^{T} \sigma_{t-1}(\boldsymbol{x}_t) \tag{3}$$

- [Srinivas et al., 2010] show the connection between R.H.S. and information gain:

$$R_T \underset{\text{w.h.p.}}{\lesssim} \widetilde{O}\left(\sqrt{T I(\boldsymbol{x}_1, \dots, \boldsymbol{x}_T)} \right) \le \widetilde{O}\left(\sqrt{T \gamma_T(\mathcal{X})} \right)$$

  - $\gamma_T(\mathcal{X}) \coloneqq \sup_{\boldsymbol{x}_1, \dots, \boldsymbol{x}_T \in \mathcal{X}} I(\boldsymbol{x}_1, \dots, \boldsymbol{x}_T)$ is the **worst-case (maximum) information gain**

Classical information-gain based analysis [Srinivas et al., 2010]

- Let us define $I(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)$ as the following quantity, called **information gain**:

$$I(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T) = \frac{1}{2} \log \det \Big(\boldsymbol{I}_T + \sigma^{-2} \mathbf{K}_T \Big).$$

  - $\mathbf{K}_T := [k(\boldsymbol{x}_i, \boldsymbol{x}_j)]_{1 \le i,j \le T}$: Gram matrix for inputs $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T$
- Due to the UCB selection rule and the confidence bound of GP,

$$R_T \underset{\text{w.h.p.}}{\lesssim} 2\beta_T \sum_{t=1}^{T} \sigma_{t-1}(\boldsymbol{x}_t) \tag{3}$$

- [Srinivas et al., 2010] show the connection between R.H.S. and information gain:

$$R_T \underset{\text{w.h.p.}}{\lesssim} \widetilde{O}\left(\sqrt{T I(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)} \right) \le \widetilde{O}\left(\sqrt{T \gamma_T(\mathcal{X})} \right)$$

  - $\gamma_T(\mathcal{X}) := \sup_{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T \in \mathcal{X}} I(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)$ is the **worst-case (maximum) information gain**

Resulting bounds

- For k_{SE}: $\gamma_T(\mathcal{X}) = O(\ln^{d+1} T) \Rightarrow \ R_T = O\big(\sqrt{T \ln^{d+2} T}\big)$
- For $k_{\text{Matérn}}$: $\gamma_T(\mathcal{X}) = \widetilde{O}(T^{\frac{d}{2\nu+d}}) \Rightarrow \ R_T = \widetilde{O}\Big(T^{\frac{\nu+d}{2\nu+d}}\Big)$

Next Section

1 Introduction

2 Preliminaries

3 Improved Regret Bounds for GP-UCB

4 Possible Future Directions and Conclusion

Improved Regret Bounds for GP-UCB (Informal)

Fix any $\delta \in (0,1)$, $\delta_{\mathrm{GP}} \in (0,1)$. Under the Bayesian setting, GP-UCB satisfies the following regret bounds with probability at least $1 - \delta - \delta_{\mathrm{GP}}$:

$$R_T = \begin{cases} O(\sqrt{T \ln^2 T}) & \text{if} \quad k = k_{\mathrm{SE}} \\ \widetilde{O}(\sqrt{T}) & \text{if} \quad k = k_{\text{Matérn}} \quad \text{(with sufficiently large } \nu\text{)} \end{cases} \tag{4}$$

Improved Regret Bounds for GP-UCB (Informal)

Fix any $\delta \in (0,1)$, $\delta_{\mathrm{GP}} \in (0,1)$. Under the Bayesian setting, GP-UCB satisfies the following regret bounds with probability at least $1-\delta-\delta_{\mathrm{GP}}$:

$$R_T = \begin{cases} O(\sqrt{T \ln^2 T}) & \text{if} \quad k = k_{\mathrm{SE}} \\ \widetilde{O}(\sqrt{T}) & \text{if} \quad k = k_{\text{Matérn}} \quad \text{(with sufficiently large } \nu\text{)} \end{cases} \tag{4}$$

- The above result improves upon the existing upper bounds in [Srinivas et al., 2010]:

$$\text{(Regret upper bound in [Srinivas et al., 2010])} \quad R_T = \begin{cases} O(\sqrt{T \ln^{d+2} T}) & \text{if} \quad k = k_{\mathrm{SE}} \\ \widetilde{O}(T^{\frac{\nu+d}{2\nu+d}}) & \text{if} \quad k = k_{\text{Matérn}} \end{cases} \tag{5}$$

Remarks:

- We do not change the algorithm itself to derive above results
  - We adopt completely the same confidence width parameters $(\beta_t)_{t \in \mathbb{N}_+}$ as in [Srinivas et al., 2010]
- We focus only on the scaling of T
  - The hidden constants depend on other parameters (such as d, ℓ, δ, δ_{GP}, and ν)

High-Level Idea behind New Analysis

Idea: Study the **algorithm- and sample-path-dependent behavior** of the information gain

Idea: Study the **algorithm- and sample-path-dependent behavior** of the information gain

(Existing Analysis [Srinivas et al., 2010]) $R_T \underset{\text{w.h.p.}}{\leq} \widetilde{O}\left(\sqrt{TI(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)}\right) \leq \widetilde{O}\left(\sqrt{T\gamma_T(\mathcal{X})}\right)$

- $I(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)$: Information gain under inputs $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T$
- $\gamma_T(\mathcal{X}) := \sup_{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T \in \mathcal{X}} I(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)$: Worst-case (maximum) information gain

Key observation: $I(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)$ can be **significantly smaller** than $\gamma_T(\mathcal{X})$ under GP-UCB

- Intuitively, GP-UCB eventually concentrates its queries around $\boldsymbol{x}^*$ to minimize regret
- But information gain from such inputs is small since they consists mostly of similar points

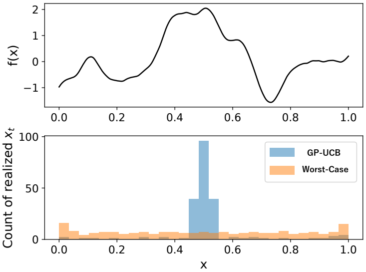
Idea: Study the **algorithm- and sample-path-dependent behavior** of the information gain

$$\text{(Existing Analysis [Srinivas et al., 2010])} \quad R_T \underset{\text{w.h.p.}}{\leq} \tilde{O}\left(\sqrt{T I(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)}\right) \leq \tilde{O}\left(\sqrt{T \gamma_T(\mathcal{X})}\right)$$

- $I(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)$: Information gain under inputs $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T$
- $\gamma_T(\mathcal{X}) := \sup_{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T \in \mathcal{X}} I(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)$: Worst-case (maximum) information gain

Key observation: $I(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)$ can be **significantly smaller** than $\gamma_T(\mathcal{X})$ under GP-UCB

- Intuitively, GP-UCB eventually concentrates its queries around $\boldsymbol{x}^*$ to minimize regret
- But information gain from such inputs is small since they consists mostly of similar points



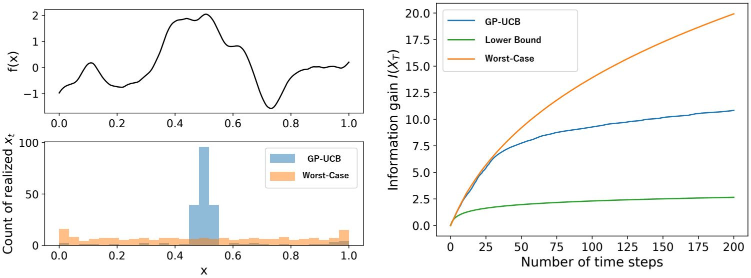
Idea: Study the **algorithm- and sample-path-dependent behavior** of the information gain

(Existing Analysis [Srinivas et al., 2010]) $R_T \underset{\text{w.h.p.}}{\leq} \widetilde{O}\left(\sqrt{TI(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)}\right) \leq \widetilde{O}\left(\sqrt{T\gamma_T(\mathcal{X})}\right)$

- $I(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)$: Information gain under inputs $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T$
- $\gamma_T(\mathcal{X}) := \sup_{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T \in \mathcal{X}} I(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)$: Worst-case (maximum) information gain

Key observation: $I(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)$ can be **significantly smaller** than $\gamma_T(\mathcal{X})$ under GP-UCB

- Intuitively, GP-UCB eventually concentrates its queries around $\boldsymbol{x}^*$ to minimize regret
- But information gain from such inputs is small since they consists mostly of similar points



Proof Overview 1: Regret Decomposition and Lenient Regret

Step 1: Decompose regret by a fixed threshold ε

Step 1: Decompose regret by a fixed threshold ε

$$
\begin{aligned}
R_T &= \sum_{t=1}^{T} f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t) \\
&= \underbrace{\sum_{t: f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t) > \varepsilon} f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t)}_{\text{Lenient regret [Cai et al., 2021], polylog}(T)} + \underbrace{\sum_{t: f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t) \leq \varepsilon} f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t)}_{?} \\
&:= R_T^{(1)} + R_T^{(2)}
\end{aligned}
$$

Proof Overview 1: Regret Decomposition and Lenient Regret

Step 1: Decompose regret by a fixed threshold ε

$$
\begin{aligned}
R_T &= \sum_{t=1}^{T} f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t) \\
&= \underbrace{\sum_{t: f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t) > \varepsilon} f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t)}_{\text{Lenient regret [Cai et al., 2021], polylog}(T)} + \underbrace{\sum_{t: f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t) \le \varepsilon} f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t)}_{?} \\
&:= R_T^{(1)} + R_T^{(2)}
\end{aligned}
$$

- For both k_{SE} and $k_{\mathrm{Matérn}}$, the lenient regret of GP-UCB is only $\mathrm{polylog}(T)$ when ε is fixed [Cai et al., 2021]
- Dominant term is $R_T^{(2)}$
  - Aim to show a good bound of $R_T^{(2)}$ by using **non-worst-case behavior** of information gain

Proof Overview 2-1: Bounding $R_T^{(2)}$

Introduce **two key facts** for bounding $R_T^{(2)}$ from the known properties of the GP-sample path

[1] This is only valid for sufficiently smooth stationary kernel such as $k = k_{\mathrm{SE}}$ or $k = k_{\mathrm{Matérn}}$ with $\nu > 2$.

Introduce **two key facts** for bounding $R_T^{(2)}$ from the known properties of the GP-sample path

Fact 1: locally quadratic behavior near x^*

- Near $\boldsymbol{x}^*$, f behaves **quadratically** with high probability [e.g., De Freitas et al., 2012][1]:
- With probability at least $1 - \delta_{\mathrm{GP}}$, we have

$$f(\boldsymbol{x}^*) - f(\boldsymbol{x}) \leq c\|\boldsymbol{x} - \boldsymbol{x}^*\|_2^2 \quad \text{for } \boldsymbol{x} \in \mathcal{B}_2(\rho; \boldsymbol{x}^*)$$

  - $\mathcal{B}_2(\rho; \boldsymbol{x}^*) \coloneqq \{\boldsymbol{x} \in \mathcal{X} \mid \|\boldsymbol{x} - \boldsymbol{x}^*\|_2 \leq \rho\}$
  - $c, \rho > 0$: constant depending on δ_{GP}

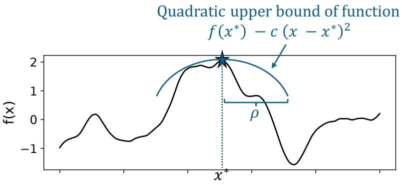


Figure: Illustrative image in 1-dimension

[1] This is only valid for sufficiently smooth stationary kernel such as $k = k_{\mathrm{SE}}$ or $k = k_{\mathrm{Matérn}}$ with $\nu > 2$.

Proof Overview 2-1: Bounding $R_T^{(2)}$

Introduce **two key facts** for bounding $R_T^{(2)}$ from the known properties of the GP-sample path

Fact 1: locally quadratic behavior near x^*

- Near $\boldsymbol{x}^*$, f behaves **quadratically** with high probability [e.g., De Freitas et al., 2012][1]:
- With probability at least $1 - \delta_{\mathrm{GP}}$, we have

$$f(\boldsymbol{x}^*) - f(\boldsymbol{x}) \le c\|\boldsymbol{x} - \boldsymbol{x}^*\|_2^2 \quad \text{for } \boldsymbol{x} \in \mathcal{B}_2(\rho; \boldsymbol{x}^*)$$

  - $\mathcal{B}_2(\rho; \boldsymbol{x}^*) := \{\boldsymbol{x} \in \mathcal{X} \mid \|\boldsymbol{x} - \boldsymbol{x}^*\|_2 \le \rho\}$
  - $c, \rho > 0$: constant depending on δ_{GP}

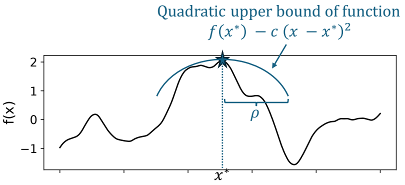


Figure: Illustrative image in 1-dimension

Fact 2: upper bound of the input radius

- If the **sub-optimality gap is sufficiently small**, upper bound of the radius of corresponding inputs is obtained [Scarlett, 2018]
  - Given the threshold $\eta > 0$ of sub-optimality gap, we can obtain

$$\{\boldsymbol{x}_t | f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t) \le \eta\} \subset \mathcal{B}_2(\sqrt{c^{-1}\eta}; \boldsymbol{x}^*) \quad (6)$$

[1] This is only valid for sufficiently smooth stationary kernel such as $k = k_{\mathrm{SE}}$ or $k = k_{\mathrm{Matérn}}$ with $\nu > 2$.

Proof Overview 2-2: Bounding $R_T^{(2)}$

Quantify the upper bound of $R_T^{(2)}$ by using aforementioned facts

- Define a sequence of thresholds $(\eta_i)_{i\geq 1}$:

$$\eta_i = \widetilde{\Theta}\left(\sqrt{\frac{\beta_T \gamma_{T_i}(\mathcal{X})}{T_i}}\right), \quad T_i = T/2^{i-1}$$

- Define the time indices where regret is η_i:

$$\mathcal{T}(\eta_i) := \{t \in [T] \mid f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t) \geq \eta_i\}$$

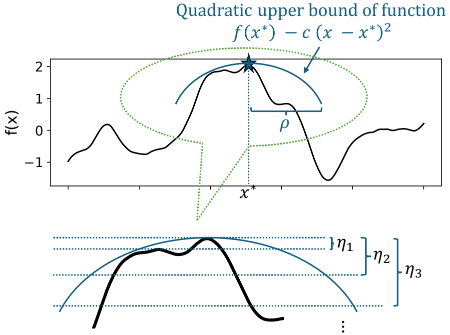
Quantify the upper bound of $R_T^{(2)}$ by using aforementioned facts

- Define a sequence of thresholds $(\eta_i)_{i \geq 1}$:

$$\eta_i = \widetilde{\Theta}\left(\sqrt{\frac{\beta_T \gamma_{T_i}(\mathcal{X})}{T_i}}\right), \quad T_i = T/2^{i-1}$$

- Define the time indices where regret is η_i:

$$\mathcal{T}(\eta_i) := \{t \in [T] \mid f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t) \geq \eta_i\}$$



Proof Overview 2-2: Bounding $R_T^{(2)}$

Quantify the upper bound of $R_T^{(2)}$ by using aforementioned facts

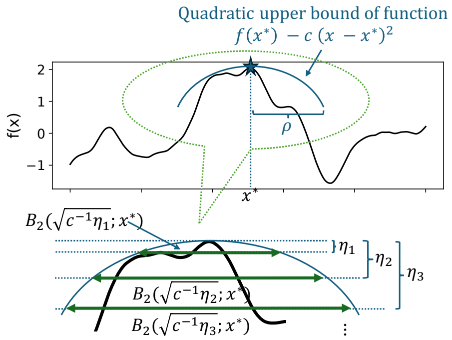
- Define a sequence of thresholds $(\eta_i)_{i\geq 1}$:

$$\eta_i = \widetilde{\Theta}\left(\sqrt{\frac{\beta_T \gamma_{T_i}(\mathcal{X})}{T_i}}\right), \quad T_i = T/2^{i-1}$$

- Define the time indices where regret is η_i:

$$\mathcal{T}(\eta_i) := \{t \in [T] \mid f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t) \geq \eta_i\}$$

- By a straightforward extension of the analysis in [Srinivas et al., 2010], we can derive $|\mathcal{T}(\eta_{i-1})| \leq T_i$
- From the quadratic conditions, $\{\boldsymbol{x}_t | t \in \mathcal{T}^c(\eta_i)\} \subset \mathcal{B}_2(\sqrt{c^{-1}\eta_i}; \boldsymbol{x}^*)$

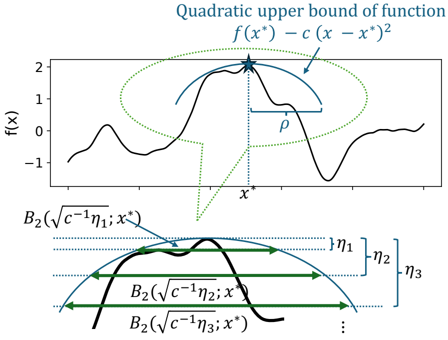


Proof Overview 2-2: Bounding $R_T^{(2)}$

Quantify the upper bound of $R_T^{(2)}$ by using aforementioned facts

- Define a sequence of thresholds $(\eta_i)_{i\geq 1}$:

$$\eta_i = \widetilde{\Theta}\left(\sqrt{\frac{\beta_T \gamma_{T_i}(\mathcal{X})}{T_i}}\right), \quad T_i = T/2^{i-1}$$



- Define the time indices where regret is η_i:

$$\mathcal{T}(\eta_i) := \{t \in [T] \mid f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t) \geq \eta_i\}$$

- By a straightforward extension of the analysis in [Srinivas et al., 2010], we can derive $|\mathcal{T}(\eta_{i-1})| \leq T_i$
- From the quadratic conditions, $\{\boldsymbol{x}_t | t \in \mathcal{T}^c(\eta_i)\} \subset \mathcal{B}_2(\sqrt{c^{-1}\eta_i}; \boldsymbol{x}^*)$
- By combining above properties with information gain bound, we have

$$R_T^{(2)} = \sum_i \underbrace{\sum_{\mathcal{T}^c(\varepsilon)\cap\mathcal{T}^c(\eta_i)\cap\mathcal{T}(\eta_{i-1})} f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t)}_{\leq O\left(\sqrt{\beta_T T_i \gamma_{T_i}(\mathcal{B}_2(\sqrt{c^{-1}\eta_i}; \boldsymbol{x}^*))}\right)} \underset{\text{w.h.p.}}{\leq} O\left(\sqrt{\beta_T T \max_i \gamma_{T_i}(\mathcal{B}_2(\sqrt{c^{-1}\eta_i}; \boldsymbol{x}^*))}\right)$$

Proof Overview 2-3: Bounding $R_T^{(2)}$

$$\text{(Derived bound)} \quad R_T^{(2)} \leq_{\text{w.h.p.}} O\left(\sqrt{\beta_T T \max_i \gamma_{T_i}(\mathcal{B}_2(\sqrt{c^{-1}\eta_i}; \boldsymbol{x}^*))}\right) \tag{7}$$

- Note that η_i is decreasing quantity as T_i increases
- Namely, the input radius **shrinks as time step of information gain increases**
  $\rightarrow \max_i \gamma_{T_i}(\mathcal{B}_2(\sqrt{c\eta_i}; \boldsymbol{x}^*))$ **becomes strictly smaller than** $\gamma_T(\mathcal{X})$ **due to this shrinking**
- Specifically,

$$\max_i \gamma_{T_i}(\mathcal{B}_2(\sqrt{c^{-1}\eta_i}; \boldsymbol{x}^*)) \leq \begin{cases} O(\ln T) & \text{if} \quad k = k_{\text{SE}}, \\ \text{polylog}(T) & \text{if} \quad k = k_{\text{Matérn}} \text{ (with sufficiently large } \nu) \end{cases}$$

- Therefore, we obtain the desired result:

$$\begin{aligned} R_T^{(2)} &\leq O\left(\sqrt{\beta_T T \max_i \gamma_{T_i}(\mathcal{B}_2(\sqrt{c^{-1}\eta_i}; \boldsymbol{x}^*))}\right) \\ &\leq \begin{cases} O(\sqrt{T \ln^2 T}) & \text{if} \quad k = k_{\text{SE}}, \\ \widetilde{O}(\sqrt{T}) & \text{if} \quad k = k_{\text{Matérn}} \text{(with sufficiently large } \nu) \end{cases} \end{aligned}$$

Next Section

1 Introduction

2 Preliminaries

3 Improved Regret Bounds for GP-UCB

4 Possible Future Directions and Conclusion

Possible Future Directions: Extension to other Settings

Beyond the basic Bayesian optimization setup

- The high-level idea of our analysis may be applicable for other settings
- The key aspects of our analysis are the following three-fold:
  1. Behavior of the objective function under the Bayesian assumption
  2. Behavior of the information gain
  3. Behavior of the algorithm

Beyond the basic Bayesian optimization setup

- The high-level idea of our analysis may be applicable for other settings
- The key aspects of our analysis are the following three-fold:
  1. Behavior of the objective function under the Bayesian assumption
  2. Behavior of the information gain
  3. Behavior of the algorithm
- It will be intriguing to analyze settings where one of the above three aspects differs from the standard setting. For example:

Possible Future Directions: Extension to other Settings

Beyond the basic Bayesian optimization setup

- The high-level idea of our analysis may be applicable for other settings
- The key aspects of our analysis are the following three-fold:
  1. Behavior of the objective function under the Bayesian assumption
  2. Behavior of the information gain
  3. Behavior of the algorithm
- It will be intriguing to analyze settings where one of the above three aspects differs from the standard setting. For example:
  - **Settings where the objective function differs from the standard setting:**
    - Multi-objective optimization [e.g., Zuluaga et al., 2016], Robust optimization [e.g., Bogunovic et al., 2018], Composite objective function [e.g., Li and Scarlett, 2024], etc.

Possible Future Directions: Extension to other Settings

Beyond the basic Bayesian optimization setup

- The high-level idea of our analysis may be applicable for other settings
- The key aspects of our analysis are the following three-fold:
    1. Behavior of the objective function under the Bayesian assumption
    2. Behavior of the information gain
    3. Behavior of the algorithm
- It will be intriguing to analyze settings where one of the above three aspects differs from the standard setting. For example:
    - **Settings where the objective function differs from the standard setting:**
        - Multi-objective optimization [e.g., Zuluaga et al., 2016], Robust optimization [e.g., Bogunovic et al., 2018], Composite objective function [e.g., Li and Scarlett, 2024], etc.
    - **Settings where the information gain behaves differently from the standard analysis:**
        - Multi-task setting [e.g., Chowdhury and Gopalan, 2021], Noise-free setting [e.g., Bull, 2011], etc.

Beyond the basic Bayesian optimization setup

- The high-level idea of our analysis may be applicable for other settings
- The key aspects of our analysis are the following three-fold:
  1. Behavior of the objective function under the Bayesian assumption
  2. Behavior of the information gain
  3. Behavior of the algorithm
- It will be intriguing to analyze settings where one of the above three aspects differs from the standard setting. For example:
  - **Settings where the objective function differs from the standard setting:**
    - Multi-objective optimization [e.g., Zuluaga et al., 2016], Robust optimization [e.g., Bogunovic et al., 2018], Composite objective function [e.g., Li and Scarlett, 2024], etc.
  - **Settings where the information gain behaves differently from the standard analysis:**
    - Multi-task setting [e.g., Chowdhury and Gopalan, 2021], Noise-free setting [e.g., Bull, 2011], etc.
  - **Algorithm beyond GP-UCB:**
    - Thompson sampling [Russo and Van Roy, 2014a], information directed sampling [Russo and Van Roy, 2014b]
    - But these algorithms mainly focus on expected regret
      $\rightarrow$ Although we believe that high-level idea of our analysis is beneficial, it is not directly applicable so far

Possible Future Directions: Extension to Instance Dependence Analysis

- Can we improve the **worst-case regret under the frequentist setting** by using non-worst-case information gain?
  - The setting whose f is not random and fixed in some reproducing kernel Hilbert space

- Can we improve the **worst-case regret under the frequentist setting** by using non-worst-case information gain?
  - The setting whose f is not random and fixed in some reproducing kernel Hilbert space

$\rightarrow$ We conjecture that it is impossible

  - We can construct worst-case functions such that $\boldsymbol{x}_t$ does not concentrate around $\boldsymbol{x}^*$
    - Such as worst-case bump functions [Bull, 2011; Scarlett et al., 2017]

- Can we improve the **worst-case regret under the frequentist setting** by using non-worst-case information gain?
  - The setting whose f is not random and fixed in some reproducing kernel Hilbert space

$\rightarrow$ We conjecture that it is impossible

  - We can construct worst-case functions such that $\boldsymbol{x}_t$ does not concentrate around $\boldsymbol{x}^*$
    - Such as worst-case bump functions [Bull, 2011; Scarlett et al., 2017]

- However, even under frequentist setting, an instance dependent analysis of regret [Shekhar and Javidi, 2022] is a promising direction
  - We expect that it is possible to use our analysis with some instance dependent property of f
  - One promising way is to leverage the growth condition in [Shekhar and Javidi, 2022]
    - growth condition: $\exists b, \underline{c} \|\boldsymbol{x}^* - \boldsymbol{x}\|^b \leq f(\boldsymbol{x}^*) - f(\boldsymbol{x}) \leq \overline{c} \|\boldsymbol{x}^* - \boldsymbol{x}\|^b$ for some neighborhood of $\boldsymbol{x}^*$
    - Interpreted as the generalization of the quadratic condition

Conclusion

- We provide improved regret upper bounds for GP-UCB under Bayesian setting
- The key idea is to consider non-worst-case information gain by:
  1. Specific sample-path behavior of GP
  2. Considering GP-UCB's realized input sequence by relating it to sample-path
- We believe that the proposed analysis will motivate refinements of many existing theories based on [Srinivas et al., 2010]

I. Bogunovic, J. Scarlett, S. Jegelka, and V. Cevher. Adversarially robust optimization with Gaussian processes. In *Proc. Neural Information Processing Systems (NeurIPS)*, 2018.

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

X. Cai, S. Gomes, and J. Scarlett. Lenient regret and good-action identification in Gaussian process bandits. In *International Conference on Machine Learning*, pages 1183–1192. PMLR, 2021.

S. R. Chowdhury and A. Gopalan. No-regret algorithms for multi-task bayesian optimization. In *International Conference on Artificial Intelligence and Statistics*, pages 1873–1881. PMLR, 2021.

N. De Freitas, A. J. Smola, and M. Zoghi. Exponential regret bounds for Gaussian process bandits with deterministic observations. In *Proceedings of the 29th International Conference on International Conference on Machine Learning*, page 955–962. Omnipress, 2012.

Z. Li and J. Scarlett. Regret bounds for noise-free cascaded kernelized bandits. *Trans. Mach. Learn. Res.*, 2024.

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

D. Russo and B. Van Roy. Learning to optimize via information-directed sampling. *Advances in neural information processing systems*, 27, 2014b.

J. Scarlett. Tight regret bounds for Bayesian optimization in one dimension. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 4500–4508. PMLR, 2018.

J. Scarlett, I. Bogunovic, and V. Cevher. Lower bounds on regret for noisy Gaussian process bandit optimization. In *Proc. Conference on Learning Theory (COLT)*, 2017.

Reference II

S. Shekhar and T. Javidi. Instance dependent regret analysis of kernelized bandits. In *International Conference on Machine Learning*, pages 19747–19772. PMLR, 2022.

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

M. Zuluaga, A. Krause, and M. Püschel. e-pal: An active learning approach to the multi-objective optimization problem. *Journal of Machine Learning Research*, 17(104):1–32, 2016.