On Kernelized Multi-armed Bandits

Sayak Ray Chowdhury    Aditya Gopalan

Department of Electrical Communication Engineering
Indian Institute of Science

ECE Student Seminar Series
August 2, 2017
Overview

Problem Formulation

Algorithms

Regret Bounds

Numerical Results

Proof Outline

Conclusion
Problem Statement

Sequentially Maximize $f : D \rightarrow \mathbb{R}$

- $f$ unknown, $D \subset \mathbb{R}^d$
Problem Statement

Sequentially Maximize $f : D \rightarrow \mathbb{R}$

- $f$ unknown, $D \subset \mathbb{R}^d$
- $x^* = \arg\max_{x \in D} f(x)$

Performance Metric
- Regret $r_t = f(x^*) - f(x_t)$
- Goal: Minimize cumulative regret $\sum_{t=1}^{T} r_t$
Problem Statement

Sequentially Maximize $f : D \rightarrow \mathbb{R}$

- $f$ unknown, $D \subset \mathbb{R}^d$
- $x^* = \arg\max_{x \in D} f(x)$
- At each round $t$:
  - Learner chooses $x_t \in D$ based on past
  - Observes noisy reward $y_t = f(x_t) + \varepsilon_t$
Problem Statement

Sequentially Maximize $f : D \rightarrow \mathbb{R}$

- $f$ unknown, $D \subset \mathbb{R}^d$
- $x^* = \arg\max_{x \in D} f(x)$
- At each round $t$:
  - Learner chooses $x_t \in D$ based on past
  - Observes noisy reward $y_t = f(x_t) + \varepsilon_t$

Performance Metric

- Regret $r_t = f(x^*) - f(x_t)$
- Goal: Minimize cumulative regret $\sum_{t=1}^{T} r_t$
Assumptions

- Noise $\varepsilon_t$ is $R$-sub-Gaussian
Assumptions

- Noise $\varepsilon_t$ is $\mathcal{R}$-sub-Gaussian
- $f$ lies in **RKHS** of functions: $D \rightarrow \mathbb{R}$
- Positive semi-definite kernel function $k : D \times D \rightarrow \mathbb{R}$ (known)
- Reproducing property: $f(x) = \langle f, k(x, \cdot) \rangle_k$
- Induces smoothness: $|f(x) - f(y)| \leq \|f\|_k \|k(x, \cdot) - k(y, \cdot)\|_k$
Assumptions

- Noise $\varepsilon_t$ is $R$-sub-Gaussian

- $f$ lies in **RKHS** of functions: $D \rightarrow \mathbb{R}$

- Positive semi-definite kernel function $k : D \times D \rightarrow \mathbb{R}$ (known)

- Reproducing property: $f(x) = \langle f, k(x, \cdot) \rangle_k$

- Induces smoothness: $|f(x) - f(y)| \leq \|f\|_k \|k(x, \cdot) - k(y, \cdot)\|_k$

- $D$ is compact, $\|f\|_k \leq B$ known
Assumptions

- Noise $\varepsilon_t$ is $\mathcal{R}$-sub-Gaussian

- $f$ lies in RKHS of functions: $D \to \mathcal{R}$

- Positive semi-definite kernel function $k : D \times D \to \mathcal{R}$ (known)

- Reproducing property: $f(x) = \langle f , k(x, \cdot) \rangle_k$

- Induces smoothness: $|f(x) - f(y)| \leq \|f\|_k \|k(x, \cdot) - k(y, \cdot)\|_k$

- $D$ is compact, $\|f\|_k \leq B$ known

- Bounded variance: $k(x, x) \leq 1$, for all $x \in D$
Example Kernels

- **Squared Exponential** kernel: \( k(x, y) = \exp \left( \frac{-\|x-y\|^2}{2l^2} \right) \)

- **Matérn** kernel: \( k(x, y) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left( \frac{\|x-y\| \sqrt{2\nu}}{l} \right)^\nu B_\nu \left( \frac{\|x-y\| \sqrt{2\nu}}{l} \right) \)

- **Stationary** kernels: \( k(x, y) \equiv k(x - y) \)
Example Kernels

- **Squared Exponential** kernel: \( k(x, y) = \exp\left(\frac{-\|x-y\|^2}{2l^2}\right) \)

- **Matérn** kernel: 
  \[
  k(x, y) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\|x-y\|_{2}\sqrt{2\nu}}{l}\right)^{\nu} B_{\nu}\left(\frac{\|x-y\|_{2}\sqrt{2\nu}}{l}\right)
  \]

- **Stationary** kernels: 
  \( k(x, y) \equiv k(x - y) \)

- **Linear** Kernel:
  - \( k(x, y) = x^T y \)
  - \( f(x) = \theta^T x, \theta \in \mathbb{R}^d \) unknown parameter
Example Kernels

- **Squared Exponential** kernel: \( k(x, y) = \exp\left( -\frac{\|x-y\|^2}{2l^2} \right) \)

- **Matérn** kernel: \( k(x, y) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left( \frac{\|x-y\| \sqrt{2\nu}}{l} \right)^\nu B_\nu \left( \frac{\|x-y\| \sqrt{2\nu}}{l} \right) \)

- **Stationary** kernels: \( k(x, y) \equiv k(x - y) \)

- **Linear Kernel:**
  - \( k(x, y) = x^T y \)
  - \( f(x) = \theta^T x, \theta \in \mathbb{R}^d \) unknown parameter

- Reduces to parametric **linear bandit** problem (Dani et al., COLT 2008, Abbasi-Yadkori et al., NIPS 2011, ...)
Algorithm Design Philosophy: Gaussian Processes

Assume:

- Gaussian Process Prior of $f$: $GP(0, \nu^2 k(x, y))$
- Noise $\varepsilon_t \sim \mathcal{N}(0, \lambda \nu^2)$
Algorithm Design Philosophy: Gaussian Processes

Assume:

- Gaussian Process Prior of $f$: $GP(0, \nu^2 k(x, y))$
- Noise $\varepsilon_t \sim \mathcal{N}(0, \lambda \nu^2)$
- After $t$ rounds, reward vector $y_{1:t} \sim \mathcal{N}(0, \nu^2(K_t + \lambda I))$
Algorithm Design Philosophy: Gaussian Processes

Assume:

- Gaussian Process Prior of $f$: $GP(0, \nu^2 k(x, y))$
- Noise $\varepsilon_t \sim \mathcal{N}(0, \lambda \nu^2)$
- After $t$ rounds, reward vector $y_{1:t} \sim \mathcal{N}(0, \nu^2(K_t + \lambda I))$

Posterior of $f$ after $t$ rounds: $GP(\mu_t(x), \nu^2 k_t(x, y))$

\[
\mu_t(x) = k_t(x)^T (K_t + \lambda I)^{-1} y_{1:t}
\]
\[
k_t(x, y) = k(x, y) - k_t(x)^T (K_t + \lambda I)^{-1} k_t(y)
\]
Algorithm 1: Improved GP-UCB (IGP-UCB)

**Key Idea:** Play the arm with highest UCB

At each round $t$, play:

$$x_t = \arg\max_{x \in D} \mu_t(x) + \beta_t \sigma_t(x)$$

$\beta_t$ trades off between exploration and exploitation.

Reduced width ($\beta_t$) of confidence interval compared to GP-UCB (Srinivas et al., ICML 2010).
Algorithm 1: Improved GP-UCB (IGP-UCB)

**Key Idea:** Play the arm with highest UCB

At each round $t$, play:

$$x_t = \arg\max_{x \in D} \mu_t(x) + \beta_t \sigma_t(x)$$
Algorithm 1: Improved GP-UCB (IGP-UCB)

Key Idea: Play the arm with highest UCB

At each round $t$, play:

$$x_t = \arg\max_{x \in D} \mu_t(x) + \beta_t \sigma_t(x)$$

- $\beta_t$ trades off b/w exploration and exploitation
- Reduced width ($\beta_t$) of confidence interval compared to GP-UCB (Srinivas et al., ICML 2010)
Algorithm 2: Gaussian Process Thompson Sampling (GP-TS)

**Key Idea:** Sample a random function and play its maximizer

![Diagram showing the key idea of sampling a random function and playing its maximizer.](image.png)
Algorithm 2: Gaussian Process Thompson Sampling (GP-TS)

**Key Idea:** Sample a random function and play its maximizer

At each round $t$:
- Sample $f_t$ from posterior of $f$
Algorithm 2: Gaussian Process Thompson Sampling (GP-TS)

**Key Idea:** Sample a *random* function and play its maximizer

At each round $t$:
- Sample $f_t$ from posterior of $f$
- Play $x_t = \arg\max_{x \in D_t} f_t(x)$

$D_t \subset D$: suitably chosen *Discretization* sets
Regret Bound for IGP-UCB

Result 1

Regret of IGP-UCB is $O\left(\sqrt{T}(B\sqrt{\gamma_T} + \gamma_T)\right)$ whp with the choice of confidence width $\beta_t \approx B + \sqrt{\gamma_t}$ for all $t$.
Result 1

Regret of IGP-UCB is $O\left(\sqrt{T}(B\sqrt{\gamma_T} + \gamma_T)\right)$ whp with the choice of confidence width $\beta_t \approx B + \sqrt{\gamma_t}$ for all $t$

- $\gamma_T$ is **Maximum Information Gain** after $T$ rounds:

  $$\gamma_T = \max_{A \subset D: |A| = T} I(y_A; f_A)$$

- **Mutual Information** b/w function values and rewards at set $A$
- **Reduction in uncertainty** about $f$ after observing rewards
- **SE kernel**: $\gamma_T = O((\ln T)^{d+1}) \rightarrow \text{sublinear}$ regret
Result 1

Regret of IGP-UCB is \( O\left(\sqrt{T}(B\sqrt{\gamma_T} + \gamma_T)\right) \) whp with the choice of confidence width \( \beta_t \approx B + \sqrt{\gamma_t} \) for all \( t \)

\[ \gamma_T = \max_{A \subset D : |A| = T} I(y_A; f_A) \]

- Maximum Information Gain after \( T \) rounds:
- Mutual Information b/w function values and rewards at set \( A \)
- Reduction in uncertainty about \( f \) after observing rewards
- SE kernel: \( \gamma_T = O((\ln T)^{d+1}) \) → sublinear regret

Regret of GP-UCB is \( O\left(\sqrt{T}(B\sqrt{\gamma_T} + \gamma_T \ln^{3/2} T)\right) \) whp and so we improve by \( O(\ln^{3/2} T) \)!
Regret Bound for GP-TS

Result 2

- Regret of GP-TS is $O\left(\sqrt{Td\ln(BdT)}(B\sqrt{\gamma_T} + \gamma_T)\right)$ whp
- First frequentist regret guarantee of TS in the non-parametric setting of infinite action spaces
Regret Bound for GP-TS

Result 2

- Regret of GP-TS is $O\left(\sqrt{Td \ln(BdT)}(B \sqrt{\gamma_T} + \gamma_T)\right)$ whp
- First frequentist regret guarantee of TS in the non-parametric setting of infinite action spaces

$\sqrt{d \ln(BdT)} \leftarrow$ Consequence of Discretization
Result 2

- Regret of GP-TS is $O\left(\sqrt{Td \ln(BdT)}(B \sqrt{\gamma_T} + \gamma_T)\right)$ whp
- First frequentist regret guarantee of TS in the non-parametric setting of infinite action spaces

\[ \sqrt{d \ln(BdT)} \leftrightarrow \text{Consequence of Discretization} \]

**Open Question:** Can the logarithmic dependency be removed?
Recovering Regret Bounds for Linear Bandits

**Linear Kernel**

- \( k(x, y) = x^T y \)
- \( f(x) = \theta^T x, \theta \in \mathbb{R}^d \) unknown parameter
- **Maximum Information Gain**: \( \gamma_T = O(d \ln T) \)
- Regret of IGP-UCB is \( \tilde{O}(d \sqrt{T}) \) and GP-TS is \( \tilde{O}(d^{3/2} \sqrt{T}) \)
Linear Kernel

- \( k(x, y) = x^T y \)
- \( f(x) = \theta^T x, \theta \in \mathbb{R}^d \) unknown parameter

- **Maximum Information Gain**: \( \gamma_T = O(d \ln T) \)
- Regret of IGP-UCB is \( \tilde{O}(d \sqrt{T}) \) and GP-TS is \( \tilde{O}(d^{3/2} \sqrt{T}) \)

- **Exactly** recovers regrets of OFUL (Abbasi-Yadkori et al., NIPS 2011) and Linear TS (Agrawal and Goyal, ICML 2013)
Recovering Regret Bounds for Linear Bandits

**Linear Kernel**

- \( k(x, y) = x^T y \)
- \( f(x) = \theta^T x, \theta \in \mathbb{R}^d \) unknown parameter

**Maximum Information Gain:** \( \gamma_T = O(d \ln T) \)

- Regret of IGP-UCB is \( \tilde{O}(d \sqrt{T}) \) and GP-TS is \( \tilde{O}(d^{3/2} \sqrt{T}) \)

- **Exactly** recovers regrets of OFUL (Abbasi-Yadkori et al., NIPS 2011) and Linear TS (Agrawal and Goyal, ICML 2013)

- **Lower Bound:** \( \Omega(d \sqrt{T}) \) (Dani et al., COLT 2008)
Numerical Results

Algorithms Compared:

1. **GP-Expected Improvement** (Močkus, 1975)
2. **GP-Probabilistic Improvement** (Kushner, 1964)
3. **GP-UCB** (Srinivas et al., 2010)
4. **IGP-UCB** (this work)
5. **GP-TS** (this work)
Numerical Results

$f$ sampled from RKHS
(Squared Exponential kernel)

\begin{itemize}
\item IGP-UCB improves over GP-UCB,
\item GP-TS fares reasonably well,
\item IGP-UCB performs similar to GP-UCB,
\item GP-TS performs the best,
\end{itemize}
Numerical Results

$f$ sampled from RKHS
(Squared Exponential kernel)

- IGP-UCB improves over GP-UCB 😊😊
- GP-TS fares reasonably well 😊
Numerical Results

- $f$ sampled from RKHS (Squared Exponential kernel)
- Temperature Sensor Data (Intel Berkeley Research lab)

- IGP-UCB improves over GP-UCB 😊😊
- GP-TS fares reasonably well 😊
Numerical Results

\( f \) sampled from RKHS
(Squared Exponential kernel)

Temperature Sensor Data
(Intel Berkeley Research lab)

- IGP-UCB improves over GP-UCB 😊😊
- GP-TS fares reasonably well 😊
- IGP-UCB performs similar to GP-UCB ✓
- GP-TS performs the best 😊
Key Technique: New Concentration Inequality

Setup:

- Feature map \( \varphi : D \rightarrow \text{RKHS} \)

- \( S_t = \sum_{s=1}^{t} \varepsilon_s \varphi(x_s) \leftarrow \text{RKHS-valued Martingale} \)

- \( V_t = I + \sum_{s=1}^{t} \varphi(x_s)\varphi(x_s)^T \leftarrow \text{possibly of infinite dimension} \)

Result 3: Self-Normalized CI for RKHS-valued Martingales

- For all \( t \):
  \[ \|S_t\|_2^2 V_t^{-1} \leq 2R^2 \ln(\sqrt{\det(K_t + I)} \delta) \]
  with probability at least \( 1 - \delta \) if \( K_t \) is positive-definite

- Generalizes finite-dimensional Inequality for vector-valued Martingales (Abbasi-Yadkori et al., NIPS 2011)

- Curse of Dimensionality → Mixing over Gaussian Processes
Key Technique: New Concentration Inequality

Setup:

- Feature map $\varphi : D \to \text{RKHS}$
- $S_t = \sum_{s=1}^{t} \varepsilon_s \varphi(x_s) \leftarrow \text{RKHS-valued Martingale}$
- $V_t = I + \sum_{s=1}^{t} \varphi(x_s)\varphi(x_s)^T \leftarrow$ possibly of infinite dimension

Result 3: Self-Normalized CI for RKHS-valued Martingales

- For all $t$: $\|S_t\|_{V_t^{-1}}^2 \leq 2R^2 \ln(\frac{\sqrt{\det(K_t+I)}}{\delta})$ with probability at least $1 - \delta$ if $K_t$ is positive-definite

- Generalizes finite-dimensional Inequality for vector-valued Martingales (Abbasi-Yadkori et al., NIPS 2011)

- Curse of Dimensionality $\rightarrow$ Mixing over Gaussian Processes
For Non-parametric Bandits:

- **Improved** existing UCB based algorithm
- **Introduced** new Thompson Sampling based algorithm
- **Developed** new self-normalized concentration inequality for RKHS-valued martingales

Agrawal, Shipra and Goyal, Navin. **Analysis of thompson sampling for the multi-armed bandit problem.** *In COLT, 2012.*

Srinivas, Niranjan, Krause, Andreas, Kakade, Sham M, and Seeger, Matthias. **Gaussian process optimization in the bandit setting: No regret and experimental design.** *In Proceedings of the 27th International Conference on Machine Learning, 2010*
Lemma: Concentration of Posterior Distribution

For all $t$ and for all $x \in D$:

$$\mu_t(x) - \beta_t \sigma_t(x) \leq f(x) \leq \mu_t(x) + \beta_t \sigma_t(x) \quad \text{whp}$$
Lemma: Concentration of Posterior Distribution

For all $t$ and for all $x \in D$:

$$\mu_t(x) - \beta_t \sigma_t(x) \leq f(x) \leq \mu_t(x) + \beta_t \sigma_t(x) \quad \text{whp}$$
Lemma: Concentration of Posterior Distribution

For all $t$ and for all $x \in D$:

$$
\mu_t(x) - \beta_t \sigma_t(x) \leq f(x) \leq \mu_t(x) + \beta_t \sigma_t(x) \quad \text{whp}
$$

“At every round, the unknown original function lies within properly constructed confidence intervals with shrinking width”
Proof Sketch: Regret bound for IGP-UCB

\[ \mu_t(x) - \beta_t \sigma_t(x) \leq f(x) \leq \mu_t(x) + \beta_t \sigma_t(x), \quad \beta_t \approx B + \sqrt{\gamma_t} \]
Proof Sketch: Regret bound for IGP-UCB

\[ \mu_t(x) - \beta_t \sigma_t(x) \leq f(x) \leq \mu_t(x) + \beta_t \sigma_t(x), \quad \beta_t \approx B + \sqrt{\gamma_t} \]

Regret at round \( t \):

\[ r_t = f(x^*) - f(x_t) \]
Proof Sketch: Regret bound for IGP-UCB

\[ \mu_t(x) - \beta_t \sigma_t(x) \leq f(x) \leq \mu_t(x) + \beta_t \sigma_t(x), \quad \beta_t \approx B + \sqrt{\gamma_t} \]

\[ f(x^*) \leq \mu_t(x^*) + \beta_t \sigma_t(x^*) \]

Regret at round \( t \):

\[ r_t = f(x^*) - f(x_t) \]
Proof Sketch: Regret bound for IGP-UCB

\[ \mu_t(x) - \beta_t \sigma_t(x) \leq f(x) \leq \mu_t(x) + \beta_t \sigma_t(x), \quad \beta_t \approx B + \sqrt{\gamma_t} \]

- \( f(x^*) \leq \mu_t(x^*) + \beta_t \sigma_t(x^*) \)

Regret at round \( t \):

\[
\begin{align*}
    r_t &= f(x^*) - f(x_t) \\
    &\leq \mu_t(x^*) + \beta_t \sigma_t(x^*) - f(x_t)
\end{align*}
\]
Proof Sketch: Regret bound for IGP-UCB

\[
\mu_t(x) - \beta_t \sigma_t(x) \leq f(x) \leq \mu_t(x) + \beta_t \sigma_t(x), \quad \beta_t \approx B + \sqrt{\gamma_t}
\]

\[\Rightarrow f(x^*) \leq \mu_t(x^*) + \beta_t \sigma_t(x^*)\]

Regret at round \(t\):

\[
r_t = f(x^*) - f(x_t)
\leq \mu_t(x^*) + \beta_t \sigma_t(x^*) - f(x_t)
\]

\[
x_t = \arg\max_{x \in D} \mu_t(x) + \beta_t \sigma_t(x)
\]

\[
\Downarrow
\]

\[
\mu_t(x^*) + \beta_t \sigma_t(x^*) \leq \mu_t(x_t) + \beta_t \sigma_t(x_t)
\]
Proof Sketch: Regret bound for IGP-UCB

\[ \mu_t(x) - \beta_t \sigma_t(x) \leq f(x) \leq \mu_t(x) + \beta_t \sigma_t(x), \quad \beta_t \approx B + \sqrt{\gamma_t} \]

- \( f(x^*) \leq \mu_t(x^*) + \beta_t \sigma_t(x^*) \)

Regret at round \( t \):

\[
\begin{align*}
    r_t &= f(x^*) - f(x_t) \\
    &\leq \mu_t(x^*) + \beta_t \sigma_t(x^*) - f(x_t) \\
    &\leq \mu_t(x_t) + \beta_t \sigma_t(x_t) - f(x_t)
\end{align*}
\]

\[ x_t = \arg\max_{x \in D} \mu_t(x) + \beta_t \sigma_t(x) \]

\[ \Downarrow \]

\[ \mu_t(x^*) + \beta_t \sigma_t(x^*) \leq \mu_t(x_t) + \beta_t \sigma_t(x_t) \]
Proof Sketch: Regret bound for IGP-UCB

\[ \mu_t(x) - \beta_t \sigma_t(x) \leq f(x) \leq \mu_t(x) + \beta_t \sigma_t(x), \quad \beta_t \approx B + \sqrt{\gamma t} \]

\[ f(x^*) \leq \mu_t(x^*) + \beta_t \sigma_t(x^*) \]

Regret at round \( t \):

\[ r_t = f(x^*) - f(x_t) \]
\[ \leq \mu_t(x^*) + \beta_t \sigma_t(x^*) - f(x_t) \]
\[ \leq \mu_t(x_t) + \beta_t \sigma_t(x_t) - f(x_t) \]
\[ = \mu_t(x_t) - f(x_t) + \beta_t \sigma_t(x_t) \]

\[ x_t = \arg\max_{x \in D} \mu_t(x) + \beta_t \sigma_t(x) \]
\[ \downarrow \]
\[ \mu_t(x^*) + \beta_t \sigma_t(x^*) \leq \mu_t(x_t) + \beta_t \sigma_t(x_t) \]
Proof Sketch: Regret bound for IGP-UCB

\[ \mu_t(x) - \beta_t \sigma_t(x) \leq f(x) \leq \mu_t(x) + \beta_t \sigma_t(x), \quad \beta_t \approx B + \sqrt{\gamma_t} \]

\[ x_t = \arg\max_{x \in D} \mu_t(x) + \beta_t \sigma_t(x) \]

\[ \mu_t(x^*) + \beta_t \sigma_t(x^*) \leq \mu_t(x_t) + \beta_t \sigma_t(x_t) \]

Regret at round \( t \):
\[ r_t = f(x^*) - f(x_t) \]
\[ \leq \mu_t(x^*) + \beta_t \sigma_t(x^*) - f(x_t) \]
\[ \leq \mu_t(x_t) + \beta_t \sigma_t(x_t) - f(x_t) \]
\[ = \mu_t(x_t) - f(x_t) + \beta_t \sigma_t(x_t) \]
Proof Sketch: Regret bound for IGP-UCB

\[ \mu_t(x) - \beta_t \sigma_t(x) \leq f(x) \leq \mu_t(x) + \beta_t \sigma_t(x), \quad \beta_t \approx B + \sqrt{\gamma_t} \]

- \( f(x^*) \leq \mu_t(x^*) + \beta_t \sigma_t(x^*) \)
- \( \mu_t(x_t) - f(x_t) \leq \beta_t \sigma_t(x_t) \)

\[ x_t = \arg\max_{x \in D} \mu_t(x) + \beta_t \sigma_t(x) \]

\[ \mu_t(x^*) + \beta_t \sigma_t(x^*) \leq \mu_t(x_t) + \beta_t \sigma_t(x_t) \]

Regret at round \( t \):

\[ r_t = f(x^*) - f(x_t) \]
\[ \leq \mu_t(x^*) + \beta_t \sigma_t(x^*) - f(x_t) \]
\[ \leq \mu_t(x_t) + \beta_t \sigma_t(x_t) - f(x_t) \]
\[ = \mu_t(x_t) - f(x_t) + \beta_t \sigma_t(x_t) \]
\[ \leq \beta_t \sigma_t(x_t) + \beta_t \sigma_t(x_t) \]
Proof Sketch: Regret bound for IGP-UCB

\[ \mu_t(x) - \beta_t \sigma_t(x) \leq f(x) \leq \mu_t(x) + \beta_t \sigma_t(x), \quad \beta_t \approx B + \sqrt{\gamma_t} \]

1. \[ f(x^*) \leq \mu_t(x^*) + \beta_t \sigma_t(x^*) \]
2. \[ \mu_t(x_t) - f(x_t) \leq \beta_t \sigma_t(x_t) \]

Regret at round \( t \):

\[ r_t = f(x^*) - f(x_t) \]
\[ \leq \mu_t(x^*) + \beta_t \sigma_t(x^*) - f(x_t) \]
\[ \leq \mu_t(x_t) + \beta_t \sigma_t(x_t) - f(x_t) \]
\[ = \mu_t(x_t) - f(x_t) + \beta_t \sigma_t(x_t) \]
\[ \leq \beta_t \sigma_t(x_t) + \beta_t \sigma_t(x_t) \]
\[ = 2\beta_t \sigma_t(x_t), \]

\[ x_t = \text{argmax}_{x \in D} \mu_t(x) + \beta_t \sigma_t(x) \]

\[ \downarrow \]

\[ \mu_t(x^*) + \beta_t \sigma_t(x^*) \leq \mu_t(x_t) + \beta_t \sigma_t(x_t) \]
Proof Sketch: Regret bound for IGP-UCB

\[ \mu_t(x) - \beta_t \sigma_t(x) \leq f(x) \leq \mu_t(x) + \beta_t \sigma_t(x), \quad \beta_t \approx B + \sqrt{\gamma_t} \]

\[ \mu_t(x^*) \leq \mu_t(x^*) + \beta_t \sigma_t(x^*) \]

\[ \mu_t(x_t) - f(x_t) \leq \beta_t \sigma_t(x_t) \]

\[ x_t = \arg\max_{x \in D} \mu_t(x) + \beta_t \sigma_t(x) \]

\[ \downarrow \]

\[ \mu_t(x^*) + \beta_t \sigma_t(x^*) \leq \mu_t(x_t) + \beta_t \sigma_t(x_t) \]

**Regret at round** \( t \):

\[
 r_t = f(x^*) - f(x_t) \\
 \leq \mu_t(x^*) + \beta_t \sigma_t(x^*) - f(x_t) \\
 \leq \mu_t(x_t) + \beta_t \sigma_t(x_t) - f(x_t) \\
 = \mu_t(x_t) - f(x_t) + \beta_t \sigma_t(x_t) \\
 \leq \beta_t \sigma_t(x_t) + \beta_t \sigma_t(x_t) \\
 = 2\beta_t \sigma_t(x_t),
\]

**Cumulative Regret:**

\[
 R_T = \sum_{t=1}^{T} r_t \leq \sum_{t=1}^{T} 2\beta_t \sigma_t(x_t) \leq 2\beta_T \sum_{t=1}^{T} \sigma_t(x_t)
\]

\[
 \mu_t(x^*) \leq f(x^*) \leq \mu_t(x_t) + \beta_t \sigma_t(x_t)
\]
Proof Sketch: Regret bound for IGP-UCB

\[ \sum_{t=1}^{T} \sigma_t(x_t) \leq \sqrt{T \sum_{t=1}^{T} \sigma_t^2(x_t)} \leq \sqrt{2T \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t))} \]
Proof Sketch: Regret bound for IGP-UCB

\[
\sum_{t=1}^{T} \sigma_t(x_t) \leq \sqrt{T \sum_{t=1}^{T} \sigma_t^2(x_t)} \leq \sqrt{2T \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t))}
\]

▶ **Mutual Information** b/w function values \(f_1:T\) and observed rewards \(y_1:T\) after \(T\) rounds is \(I(y_1:T; f_1:T) = \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t))\)
Proof Sketch: Regret bound for IGP-UCB

\[ \sum_{t=1}^{T} \sigma_t(x_t) \leq \sqrt{T \sum_{t=1}^{T} \sigma^2_t(x_t)} \leq \sqrt{2T \sum_{t=1}^{T} \ln(1 + \sigma^2_t(x_t))} \]

- **Mutual Information** b/w function values \( f_1:T \) and observed rewards \( y_1:T \) after \( T \) rounds is \( I(y_1:T; f_1:T) = \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma^2_t(x_t)) \)

- **Maximum Information Gain** \( \gamma_T = \max_{A \subset D:|A|=T} I(y_A; f_A) \)
Proof Sketch: Regret bound for IGP-UCB

\[
\sum_{t=1}^{T} \sigma_t(x_t) \leq \sqrt{T \sum_{t=1}^{T} \sigma_t^2(x_t)} \leq \sqrt{2T \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t))}
\]

- **Mutual Information** b/w function values \(f_{1:T}\) and observed rewards \(y_{1:T}\) after \(T\) rounds is \(I(y_{1:T}; f_{1:T}) = \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t))\)

- **Maximum Information Gain** \(\gamma_T = \max_{A \subset D: |A| = T} I(y_A; f_A)\)

\[
\frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t)) \leq \gamma_T
\]
Proof Sketch: Regret bound for IGP-UCB

\[ \sum_{t=1}^{T} \sigma_t(x_t) \leq \sqrt{T \sum_{t=1}^{T} \sigma_t^2(x_t)} \leq \sqrt{2T \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t))} \]

- **Mutual Information** between function values \( f_1: T \) and observed rewards \( y_1: T \) after \( T \) rounds is
  \[ I(y_1: T; f_1: T) = \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t)) \]

- **Maximum Information Gain** \( \gamma_T = \max_{A \subset D: |A| = T} I(y_A; f_A) \)

\[ \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t)) \leq \gamma_T \implies \sum_{t=1}^{T} \sigma_t(x_t) = O(\sqrt{T \gamma_T}) \]
Proof Sketch: Regret bound for IGP-UCB

\[ \sum_{t=1}^{T} \sigma_t(x_t) \leq \sqrt{T \sum_{t=1}^{T} \sigma_t^2(x_t)} \leq \sqrt{2T \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t))} \]

- **Mutual Information** b/w function values \( f_1:T \) and observed rewards \( y_1:T \) after \( T \) rounds is \( I(y_1:T; f_1:T) = \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t)) \)

- **Maximum Information Gain** \( \gamma_T = \max_{A \subset D: |A| = T} I(y_A; f_A) \)

- \( \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t)) \leq \gamma_T \implies \sum_{t=1}^{T} \sigma_t(x_t) = O(\sqrt{T \gamma_T}) \)

\[ R_T \leq 2 \beta_T \sum_{t=1}^{T} \sigma_t(x_t), \quad \beta_T \approx B + \sqrt{\gamma_T} \]
Proof Sketch: Regret bound for IGP-UCB

\[ \sum_{t=1}^{T} \sigma_t(x_t) \leq \sqrt{T} \sum_{t=1}^{T} \sigma_t^2(x_t) \leq \sqrt{2T \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t))} \]

- **Mutual Information** b/w function values \( f_1:T \) and observed rewards \( y_1:T \) after \( T \) rounds is \( I(y_1:T; f_1:T) = \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t)) \)

- **Maximum Information Gain** \( \gamma_T = \max_{A \subset D: |A|=T} I(y_A; f_A) \)

- \( \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t)) \leq \gamma_T \Rightarrow \sum_{t=1}^{T} \sigma_t(x_t) = O(\sqrt{T \gamma_T}) \)

\[ R_T \leq 2\beta_T \sum_{t=1}^{T} \sigma_t(x_t), \quad \beta_T \approx B + \sqrt{\gamma_T} \]

**Cumulative Regret** \( R_T = O\left(\sqrt{T}(B\sqrt{\gamma_T} + \gamma_T)\right) \)
Show: \( I(y_{1:T}; f_{1:T}) = \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t)) \)

Entropy of Gaussian: \( H(\mathcal{N}(\mu, \Sigma)) = \frac{1}{2} \ln(\det(2\pi e\Sigma)) \)
Show: $I(y_1:T; f_1:T) = \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t))$

Entropy of Gaussian: $H(\mathcal{N}(\mu, \Sigma)) = \frac{1}{2} \ln(\det(2\pi e \Sigma))$

▶ Posterior of $f(x) \sim \mathcal{N}(\mu_t(x), \nu^2 \sigma_t^2(x))$
▶ Reward $y_t = f(x_t) + \varepsilon_t$, Noise $\varepsilon_t \sim \mathcal{N}(0, \nu^2)$
Show: \[ I(y_1:T; f_1:T) = \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma^2_t(x_t)) \]

Entropy of Gaussian: \[ H(\mathcal{N}(\mu, \Sigma)) = \frac{1}{2} \ln(\det(2\pi e\Sigma)) \]

- Posterior of \( f(x) \sim \mathcal{N}(\mu_t(x), \nu^2\sigma^2_t(x)) \)
- Reward \( y_t = f(x_t) + \varepsilon_t \), Noise \( \varepsilon_t \sim \mathcal{N}(0, \nu^2) \)
- Posterior of reward \( y_t \sim \mathcal{N}(\mu_t(x_t), \nu^2(1 + \sigma^2_t(x_t))) \)
Show: \[ I(y_{1:T}; f_{1:T}) = \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t)) \]

Entropy of Gaussian: \[ H\left(\mathcal{N}(\mu, \Sigma)\right) = \frac{1}{2} \ln(\det(2\pi e \Sigma)) \]

- Posterior of \( f(x) \sim \mathcal{N}(\mu_t(x), \nu^2\sigma_t^2(x)) \)
- Reward \( y_t = f(x_t) + \varepsilon_t \), Noise \( \varepsilon_t \sim \mathcal{N}(0, \nu^2) \)
- Posterior of reward \( y_t \sim \mathcal{N}\left(\mu_t(x_t), \nu^2(1 + \sigma_t^2(x_t))\right) \)

\[ H(y_{1:T}) = \sum_{t=1}^{T} H(y_t \mid y_{1:t-1}) = \frac{T}{2} \ln(2\pi e \nu^2) + \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t)) \]
Show: $I(y_{1:T}; f_{1:T}) = \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t))$

Entropy of Gaussian: $H(\mathcal{N}(\mu, \Sigma)) = \frac{1}{2} \ln(\det(2\pi e \Sigma))$

- Posterior of $f(x) \sim \mathcal{N}(\mu_t(x), v^2 \sigma_t^2(x))$
- Reward $y_t = f(x_t) + \varepsilon_t$, Noise $\varepsilon_t \sim \mathcal{N}(0, v^2)$
- Posterior of reward $y_t \sim \mathcal{N}(\mu_t(x_t), v^2(1 + \sigma_t^2(x_t)))$
- $H(y_{1:T}) = \sum_{t=1}^{T} H(y_t | y_{1:t-1}) = \frac{T}{2} \ln(2\pi ev^2) + \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t))$
- Reward vector $y_{1:T} = f_{1:T} + \varepsilon_{1:T}$, Noise vector $\varepsilon_{1:T} \sim \mathcal{N}(0, v^2 I)$
Show: \[ I(y_{1:T}; f_{1:T}) = \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t)) \]

**Entropy of Gaussian:** \[ H(\mathcal{N}(\mu, \Sigma)) = \frac{1}{2} \ln(\det(2\pi e\Sigma)) \]

- Posterior of \( f(x) \sim \mathcal{N}(\mu_t(x), v^2\sigma_t^2(x)) \)
- Reward \( y_t = f(x_t) + \varepsilon_t \), Noise \( \varepsilon_t \sim \mathcal{N}(0, v^2) \)
- Posterior of reward \( y_t \sim \mathcal{N}(\mu_t(x_t), v^2(1 + \sigma_t^2(x_t))) \)

\[ H(y_{1:T}) = \sum_{t=1}^{T} H(y_t \mid y_{1:t-1}) = \frac{T}{2} \ln(2\pi ev^2) + \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t)) \]

- Reward vector \( y_{1:T} = f_{1:T} + \varepsilon_{1:T} \), Noise vector \( \varepsilon_{1:T} \sim \mathcal{N}(0, v^2 I) \)
- \( H(y_{1:T} \mid f_{1:T}) = H(\varepsilon_{1:T}) = \frac{T}{2} \ln(2\pi ev^2) \)
Show:  \( I(y_{1:T}; f_{1:T}) = \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t)) \)

**Entropy of Gaussian:**  \( H(\mathcal{N}(\mu, \Sigma)) = \frac{1}{2} \ln(\det(2\pi e\Sigma)) \)

- Posterior of \( f(x) \sim \mathcal{N}(\mu_t(x), \nu^2\sigma_t^2(x)) \)
- Reward \( y_t = f(x_t) + \varepsilon_t \), Noise \( \varepsilon_t \sim \mathcal{N}(0, \nu^2) \)
- Posterior of reward \( y_t \sim \mathcal{N}(\mu_t(x_t), \nu^2(1 + \sigma_t^2(x_t))) \)
  \[
  H(y_{1:T}) = \sum_{t=1}^{T} H(y_t \mid y_{1:t-1}) = \frac{T}{2} \ln(2\pi e\nu^2) + \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t))
  \]
- Reward vector \( y_{1:T} = f_{1:T} + \varepsilon_{1:T} \), Noise vector \( \varepsilon_{1:T} \sim \mathcal{N}(0, \nu^2I) \)
  \[
  H(y_{1:T} \mid f_{1:T}) = H(\varepsilon_{1:T}) = \frac{T}{2} \ln(2\pi e\nu^2)
  \]

\[
I(y_{1:T}; f_{1:T}) = H(y_{1:T}) - H(y_{1:T} \mid f_{1:T}) = \frac{1}{2} \sum_{t=1}^{T} \ln(1 + \sigma_t^2(x_t))
\]
Informal Sketch of Martingale Concentration Result

\[ \mathbb{P} \left[ \left\| S_t \right\| \leq \sqrt{2R^2 \ln \left( \frac{\sqrt{\det(K_t + I)}}{\delta} \right)} \right] \geq 1 - \delta \]
Informal Sketch of Martingale Concentration Result

\[ P \left[ \left\| S_t \right\|_{V_t^{-1}}^2 \leq 2R^2 \ln \left( \frac{\sqrt{\det(K_t+I)}}{\delta} \right) \right] \geq 1 - \delta \]

\[ S_t = \sum_{s=1}^{t} \varepsilon_s \varphi(x_s), \quad V_t = I + \sum_{s=1}^{t} \varphi(x_s)\varphi(x_s)^T, \quad K_t(i, j) = \varphi(x_i)^T \varphi(x_j) \]
Informal Sketch of Martingale Concentration Result

\[ \mathbb{P} \left[ \left\| S_t \right\|_2^2 V_t^{-1} \leq 2R^2 \ln \left( \frac{\sqrt{\det(K_t + I)}}{\delta} \right) \right] \geq 1 - \delta \]

- \( S_t = \sum_{s=1}^{t} \varepsilon_s \varphi(x_s), \ V_t = I + \sum_{s=1}^{t} \varphi(x_s)\varphi(x_s)^T, \ K_t(i,j) = \varphi(x_i)^T \varphi(x_j) \)
- Define \( \Phi_t := [\varphi(x_1) \cdots \varphi(x_t)]^T \)
Informal Sketch of Martingale Concentration Result

\[
\mathbb{P}\left[ \| S_t \|_V^{-1}^2 \leq 2R^2 \ln\left( \frac{\sqrt{\det(K_t+I)}}{\delta} \right) \right] \geq 1 - \delta
\]

- \( S_t = \sum_{s=1}^{t} \varepsilon_s \varphi(x_s), \ V_t = I + \sum_{s=1}^{t} \varphi(x_s)\varphi(x_s)^T, \ K_t(i,j) = \varphi(x_i)^T\varphi(x_j) \)
- Define \( \Phi_t := [\varphi(x_1) \cdots \varphi(x_t)]^T \)
- \( S_t = \Phi_t^T \varepsilon_{1:t}, \ V_t = I + \Phi_t^T \Phi_t \) and \( K_t = \Phi_t \Phi_t^T \)
Informal Sketch of Martingale Concentration Result

\[ \mathbb{P} \left[ \left\| S_t \right\|_{V_t^{-1}}^2 \leq 2R^2 \ln \left( \frac{\sqrt{\det(K_t+I)}}{\delta} \right) \right] \geq 1 - \delta \]

- \( S_t = \sum_{s=1}^{t} \varepsilon_s \varphi(x_s), \quad V_t = I + \sum_{s=1}^{t} \varphi(x_s)\varphi(x_s)^T, \quad K_t(i,j) = \varphi(x_i)^T \varphi(x_j) \)

- Define \( \Phi_t := [\varphi(x_1) \cdots \varphi(x_t)]^T \)

- \( S_t = \Phi_t^T \varepsilon_{1:t}, \quad V_t = I + \Phi_t^T \Phi_t \) and \( K_t = \Phi_t \Phi_t^T \)

- Hence
  \[ \left\| S_t \right\|_{V_t^{-1}}^2 = S_t^T V_t^{-1} S_t \]
  \[ = \varepsilon_{1:t}^T \Phi_t \left( I + \Phi_t^T \Phi_t \right)^{-1} \Phi_t^T \varepsilon_{1:t} \]
  \[ = \varepsilon_{1:t}^T \Phi_t \Phi_t^T \left( \Phi_t \Phi_t^T + I \right)^{-1} \varepsilon_{1:t} \]
  \[ = \varepsilon_{1:t}^T K_t \left( K_t + I \right)^{-1} \varepsilon_{1:t} \]
  \[ = \varepsilon_{1:t}^T \left( K_t^{-1} + I \right)^{-1} \varepsilon_{1:t} = \left\| \varepsilon_{1:t} \right\|^2 \left( K_t^{-1} + I \right)^{-1} \]
Show: \( \mathbb{P} \left[ \| \varepsilon_{1:t} \|^2 (K^{-1}_t + I)^{-1} \leq 2 \ln \left( \frac{\sqrt{\det(K_t + I)}}{\delta} \right) \right] \geq 1 - \delta \)

- For any function \( g : D \to \mathbb{R} \), define \( M_t^g := \exp(\varepsilon_{1:t}^T g_{1:t} - \frac{1}{2} \| g_{1:t} \|^2) \)
- \( M_t^g \) is a super-martingale with \( \mathbb{E} [M_t^g] \leq 1 \)
Show: \( \mathbb{P} \left[ \frac{1}{n} \leq 2 \ln \left( \frac{1}{\delta} \right) \right] \geq 1 - \delta \)

- For any function \( g : D \to \mathbb{R} \), define \( M_t^g := \exp(\varepsilon_{1:t}^T g_{1:t} - \frac{1}{2} \| g_{1:t} \|^2) \)
- \( M_t^g \) is a super-martingale with \( \mathbb{E} [M_t^g] \leq 1 \)
- **Method of Mixtures:** Construct a mixture martingale \( M_t \) by mixing \( M_t^g \) over \( g \) drawn from an independent Gaussian Process \( GP(0, k) \)
Show: \[ \mathbb{P} \left[ \| \varepsilon_{1:t} \|^2 (K_t^{-1} + I)^{-1} \leq 2 \ln \left( \frac{\sqrt{\det(K_t + I)}}{\delta} \right) \right] \geq 1 - \delta \]

- For any function \( g : D \to \mathbb{R} \), define \( M_t^g := \exp(\varepsilon_{1:t}^T g_{1:t} - \frac{1}{2} \| g_{1:t} \|^2) \)
- \( M_t^g \) is a super-martingale with \( \mathbb{E}[M_t^g] \leq 1 \)
- Method of Mixtures: Construct a mixture martingale \( M_t \) by mixing \( M_t^g \) over \( g \) drawn from an independent Gaussian Process \( GP(0, k) \)
- \( M_t = \int_{\mathbb{R}^D} \exp \left( \varepsilon_{1:t}^T g_{1:t} - \frac{1}{2} \| g_{1:t} \|^2 \right) d\mu(g) \),
- \( \mu \) is the GP-measure over function space \( \mathbb{R}^D \equiv \{ g : D \to \mathbb{R} \} \)
Show: \[ P \left[ \left\| \varepsilon_{1:t} \right\|^2 (K_t^{-1} + I)^{-1} \leq \frac{2 \ln \left( \frac{\sqrt{\det(K_t + I)}}{\delta} \right)}{} \right] \geq 1 - \delta \]

- For any function \( g : D \rightarrow \mathbb{R} \), define \( M^g_t := \exp(\varepsilon_{1:t}^T g_{1:t} - \frac{1}{2} \left\| g_{1:t} \right\|^2) \)
- \( M^g_t \) is a super-martingale with \( \mathbb{E} [M^g_t] \leq 1 \)
- **Method of Mixtures**: Construct a mixture martingale \( M_t \) by mixing \( M^g_t \) over \( g \) drawn from an independent **Gaussian Process** \( \text{GP}(0, k) \)
  
  \[ M_t = \int_{\mathbb{R}^D} \exp \left( \varepsilon_{1:t}^T g_{1:t} - \frac{1}{2} \left\| g_{1:t} \right\|^2 \right) d\mu(g), \]

  \( \mu \) is the GP-measure over function space \( \mathbb{R}^D \equiv \{ g : D \rightarrow \mathbb{R} \} \)

  **Change of measure**: Essentially induces a mixture distribution \( \mathcal{N}(0, K_t) \) over desired **finite** dimension \( t \)

  \[ M_t = \int_{\mathbb{R}^t} \exp \left( \varepsilon_{1:t}^T \lambda - \frac{1}{2} \left\| \lambda \right\|^2 \right) h(\lambda) d\lambda, \text{ where } h \text{ is pdf of } \mathcal{N}(0, K_t) \]
Show: \[ P \left[ \| \varepsilon_{1:t} \|^2_{(K_t^{-1} + I)^{-1}} \leq 2 \ln \left( \frac{\sqrt{\text{det}(K_t + I)}}{\delta} \right) \right] \geq 1 - \delta \]

- For any function \( g : D \to \mathbb{R} \), define \( M^g_t := \exp(\varepsilon_{1:t}^T g_{1:t} - \frac{1}{2} \| g_{1:t} \|^2) \)
- \( M^g_t \) is a super-martingale with \( \mathbb{E} [M^g_t] \leq 1 \)
- Method of Mixtures: Construct a mixture martingale \( M_t \) by mixing \( M^g_t \) over \( g \) drawn from an independent Gaussian Process \( GP(0, k) \)
  - \( M_t = \int_{\mathbb{R}^D} \exp \left( \varepsilon_{1:t}^T g_{1:t} - \frac{1}{2} \| g_{1:t} \|^2 \right) \, d\mu(g) \),
  - \( \mu \) is the GP-measure over function space \( \mathbb{R}^D \equiv \{ g : D \to \mathbb{R} \} \)
- Change of measure: Essentially induces a mixture distribution \( \mathcal{N}(0, K_t) \) over desired finite dimension \( t \)
  - \( M_t = \int_{\mathbb{R}^t} \exp \left( \varepsilon_{1:t}^T \lambda - \frac{1}{2} \| \lambda \|^2 \right) h(\lambda) \, d\lambda \), where \( h \) is pdf of \( \mathcal{N}(0, K_t) \)
  - \( M_t = \frac{1}{\sqrt{\text{det}(K_t + I)}} \exp \left( \frac{1}{2} \| \varepsilon_{1:t} \|^2 (K_t^{-1} + I)^{-1} \right) \)
- Result follows from \( \mathbb{E} [M_t] \leq 1 \) and Markov Inequality
Possible Extensions

- Kernel function not known to the learner

- Time varying functions from RKHS
Thank You