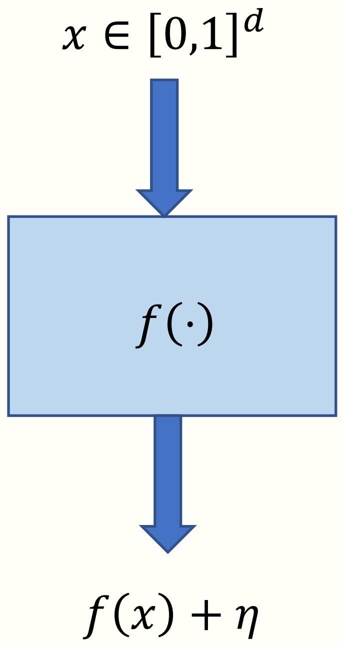
Instance Dependent Regret Analysis of Kernelized Bandits

ICML 2022

Shubhanshu Shekhar (CMU)
Tara Javidi (UCSD)

Problem Setup

- Unknown objective $f: [0,1]^d \to \mathbb{R}$
- f can be accessed via zeroth-order-oracle
- Assumptions:
  - $f \in$ RKHS of a kernel K, with norm $\leq M$
  - The noise η is i.i.d. $N(0, \sigma^2)$ with a known σ^2



Objective: Given a budget n, design an adaptive querying strategy ($\mathcal{A}$) to choose points $x_1, x_2, \dots, x_n$ such that the cumulative regret $\mathcal{R}_n(f, \mathcal{A})$ is small, where

$$\mathcal{R}_n(f, \mathcal{A}) = nf(x^*) - \sum_{i=1}^{n} f(x_t)$$

Prior Results: Minimax Framework

- Scarlett et al. (2017) showed that for Matern kernels with smoothness ν

$$\min_{\mathcal{A}} \max_{f \in \mathcal{H}_{\mathcal{K}}(M)} \mathbb{E}\left[\mathcal{R}_n(f, \mathcal{A})\right] = \Omega\left(n^{(\nu+d)/(2\nu+d)}\right)$$

- Several algorithms with theoretical guarantees.

Algorithm	**Regret bounds**
GP-UCB, GP-TS	sublinear for $\nu > d/2$
π-GP-UCB, LP-GP-UCB	sublinear for all ν values
SupKernelUCB, GP-ThreDS, RIPS	minimax near-optimal

Our Contributions

- We identify two gaps in existing results:
    1. Lower bounds provide limited information about the performance of 'good' algorithms on typical, non-adversarial problem instances
    2. Upper bounds do not adapt to easier problem instances

- Two main results:
    1. An instance dependent lower bound for algorithms that achieve $\mathcal{O}(n^{a_0})$ regret uniformly over $\mathcal{H}_{\mathcal{K}_\nu}(M)$
    2. A new algorithm ($\mathcal{A}_1$) that can adapt to some structure in problems.

Instance-Dependent Lower Bound

- (A1) $\mathcal{A}$ is a_0-consistent: $\mathbb{E}[\mathcal{R}_n(f,\mathcal{A})] = o(n^a)$ for $a > a_0$, uniformly over all $f \in \mathcal{H}_{\mathcal{K}_\nu}(M)$
- (A2) Local growth property: $f \approx (x - x^*)^b$ near its optimizer.

Theorem 1: If $(f, \mathcal{A})$ satisfy (A1-A2), then we have

$$\mathbb{E}[\mathcal{R}_n(f,\mathcal{A})] = \Omega(n^\alpha), \quad \text{for} \quad \alpha < (1 - a_0)\left(1 + \frac{d}{\nu}\left(1 - \frac{\nu}{b}\right)\right).$$

The lower bound increases with:

- Large a_0: stronger uniform regret assumption
- Large b: ‘flatter’ near-optimal regions

Instance-Dependent Upper Bound

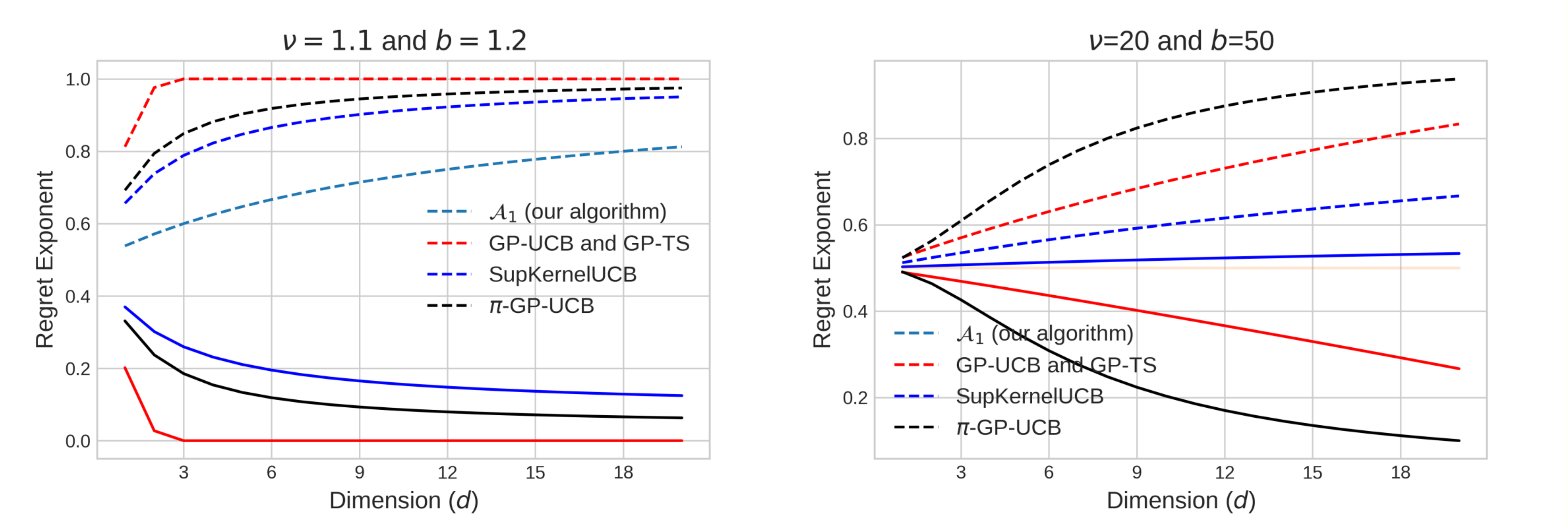
Theorem 2: We design an algorithm, $\mathcal{A}_1$, satisfying

$$\mathbb{E}[\mathcal{R}_n(f, \mathcal{A}_1)] = \mathcal{O}(n^{\alpha}),$$

where $\alpha = \min\left\{\frac{d+\nu}{d+2\nu}, \frac{d(1-\xi/b)^{+}+\xi}{d(1-\xi/b)^{+}+2\xi}\right\}$ and $\xi = \min\{1, \nu\}$.

- $\mathcal{A}_1$ achieves the minimax near-optimal worst-case regret.
- The regret of $\mathcal{A}_1$ is strictly better than the minimax rate when $b > 0, d \geq 1$ and $0 < \nu < \frac{1}{1-1/b}$.

Comparison



Future Work: Close the gap between the upper and lower bounds.

THANKS

Link to the paper:
https://arxiv.org/abs/2203.06297