Quantum Kernelized Bandits

Yasunari Hikima*, Kazunori Murao*, Sho Takemori*, Yuhei Umeda*
(Fujitsu Limited; * Alphabetical order)

Introduction and Problem Setup

We consider a quantum kernelized bandit problem (quantum bayesian optimization), where the player aims to maximize a noisy black-box non-linear reward function with a limited number of queries.

1. Select an action x_s based on the observation history
2. Input quantum circuit U containing $\mathcal{O}_{x_s}$

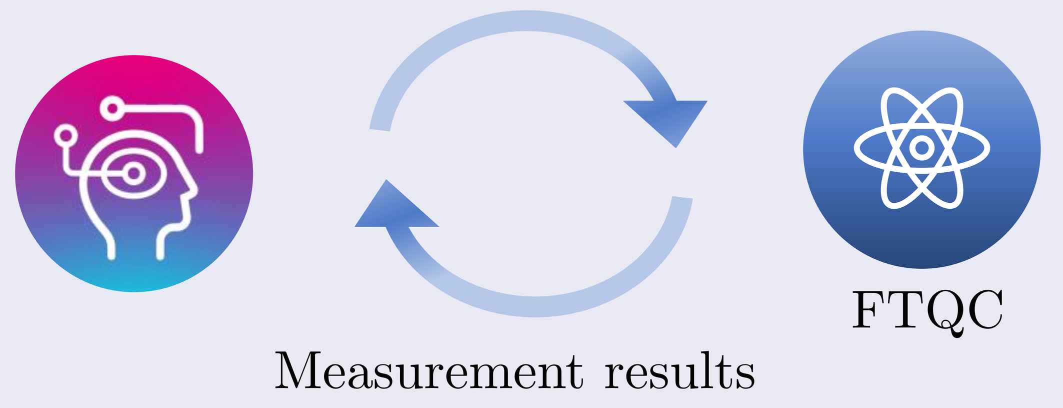

Figure: A conceptual diagram of quantum bandits

- Unlike the classical bandits, the player has a chance to access the quantum circuit (*quantum reward oracle*) that encodes the reward distribution, and this setting could perform much better than the classical setting.

■ **Problem Setup**:

For $1 \leq t \leq T$, the player selects an action $x_s \in \mathcal{X}$ and incurs an instantaneous regret $\mu(x^*) - \mu(x_t)$, where $\mu : \mathcal{X} \to [0, 1]$ is the mean reward function and $x^* \in \mathrm{argmin}_{x \in \mathcal{X}} \mu(x)$ is the optimal action in hindsight.

■ **Comparison**:

Reference	Kernel	Setting	Regret bound
Vakili et al. (2021)	β_p poly. eigen-decay	Classical	$O\left(T^{\frac{\beta_p+1}{2\beta_p}} \log^{1-\frac{1}{2\beta_p}}(T)\right)$
	β_e exp. eigen-decay	Classical	$O\left(T^{\frac{1}{2}} \log^{1+\frac{1}{2\beta_e}}(T)\right)$
Wan et al. (2023)	d-dim. linear	Quantum	$O\left(d^2 \log^{5/2} \log(T)\right)$
This work	β_p poly. eigen-decay	Quantum	$\widetilde{O}\left(T^{\frac{3}{1+\beta_p}} \log\left(\frac{1}{\delta}\right)\right)$
	β_e exp. eigen-decay	Quantum	$\widetilde{O}\left(\log^{3(1+\beta_e^{-1})/2}(T) \log\left(\frac{1}{\delta}\right)\right)$

Preliminaries on quantum computation and Quantum Monte Calro method

- A quantum computing relies on manipulating a *quantum state*, and one can extract partial information of the state by performing a *measurement*.
- Quantum Monte Calro Method (Montanaro, 2015), $\mathrm{QMC}(\mathcal{O}(y), \epsilon, \delta)$, provides a quadratic speedup compared to the classical estimator.
- Formally, the quantum algorithm $\mathrm{QMC}(\mathcal{O}(y), \epsilon, \delta)$ outputs an estimator of the expectation $\mathbb{E}[y]$ satisfying following both properties:
  (i) $P(|\widehat{y} - \mathbb{E}[y]| \geq \epsilon) \leq \delta$.
  (ii) $\mathrm{QMC}(\mathcal{O}(y), \epsilon, \delta)$ queries the unitary operator $\mathcal{O}(y)$ or its adjoint $\mathcal{O}(y)^\dagger$ at most $\frac{C}{\epsilon} \log\left(\frac{1}{\delta}\right)$ times.

Mercer Kernels and Eigenvalue Decay

■ **Mercer's Theorem**:

- Let $I \subseteq \mathbb{Z}_{\geq 1}$. We have

$$k(x, x') = \sum_{i \in I} \lambda_i \psi_i(x) \psi_i(x'), \qquad x, x' \in \mathcal{X},$$

  where $\{\lambda_i\}_{i \in I}$ and $\{\psi_i\}_{i \in I}$ be the eigenvalues of the Mercer operator and an orthogonal basis of the space of square-integrable functions on a measurable space $\mathcal{X}$, respectively.

■ **Eigen-decay**: We say a kernel has a

- (C_p, β_p)-polynomial eigen-decay, if for all $n \in I$, we have $\lambda_n \leq C_p n^{-\beta_p}$.
- $(C_{e,1}, C_{e,2}, \beta_e)$-exponential eigen-decay, if for all $n \in I$, we have $\lambda_n \leq C_{e,1} \exp(-C_{e,2} n^{\beta_e})$.

Algorithm: QMCKernelUCB

- For each stage $s = 1, 2, \ldots$, our algorithm, QMCKernelUCB, plays an action $x_s \in \mathcal{X}$, and calls $\mathrm{QMC}(\mathcal{O}_{x_s}, \eta\epsilon_s, \delta/M)$ with the error tolerance $\eta\epsilon_s$.

Algorithm 1 QMCKernelUCB

```
Input: fail probability δ ∈ (0, 1), the total number of rounds T, an upper bound of the total
  number of stages M, and a tradeoff parameter η > 0.
for each stage s = 1, 2, ... (terminate when we have used T queries to O_x, O_x†) do
    x_s ← argmax_{x∈X} μ̃_{s−1}(x) + β_{s−1} σ̃_{s−1}(x).
    ε_s ← σ̃_{s−1}(x_s).
    Run QMC(O_{x_s}, ηε_s, δ/M) obtain an output y_s of QMC.
    for the next (C/(ηε_s)) log(M/δ) rounds do
        play action x_s and the player incurs regret μ(x*) − μ(x_s).
    end for
end for
```

Theoretical Results: Regret Analysis

***Theorem* 1** (Upper Bounds of Algorithm 1)**.** *Assume $T > 1$.*

1. *Suppose that the kernel k has a β_p polynomial eigen-decay. Then, with probability at least $1 - \delta$, the cumulative regret of Algorithm 1 with $M = c_p \eta^{-2}$ and $\eta = T^{-\frac{1}{1+\beta_p}}$ is bounded as*

$$R(T) = O\left(T^{\frac{3}{1+\beta_p}} \log^{3(1-\beta_p^{-1})/2}(T) \log(T/\delta)\right).$$

2. *Suppose that the kernel k has a β_e exponential eigen-decay. Then with probability at least $1 - \delta$, the cumulative regret of Algorithm 1 with $M = c_e \log^{1+1/\beta_e}(T)$ and $\eta = 1$ is bounded as*

$$R(T) = O\left(\log^{3(1+\beta_e^{-1})/2}(T) \log\left(\frac{\log T}{\delta}\right)\right).$$

Comparison to Quantum Bayesian Optimization (Dai et al., 2023)

Compared to Dai et al. (2023), this work has the following advantages:

1. **Theoretical analysis without the unbiasedness assumption of the QMC estimator**: Dai et al. (2023) assumes that the QMC method is an unbiased estimator, however, this assumption is unlikely to hold. We provide a mathematically rigorous analysis without compromising the theoretical results.
2. **Improved regret bounds**: Even under the unbiasedness assumption, in β_p-polynomial eigen-decay case, our regret bound $\widetilde{O}(T^{\frac{3}{1+\beta_p}} \log(1/\delta))$ is better than that $\widetilde{O}(T^{\frac{3}{\beta_p}} \log(1/\delta))$ of Q-GP-UCB (Dai et al., 2023).
3. **A tradeoff parameter η in QMCKernelUCB**: Due to the tradeoff parameter η, our QMCKernelUCB can leverage the total number of stages and regret incurred in each stage.

Experiments

- We have conducted an ablation study regarding the tradeoff parameter η, since our algorithm with $\eta = 1$ is identical to that of Dai et al. (2023)
- We consider a simple synthetic environment as follows: $\mathcal{X} = [0, 1]^d$ with $d = 1$, k: Matèrn-ν kernel with $\nu = 1.5$, $\mathcal{O}_x$: a quantum circuit representing $\mathrm{Bern}(\mu(x))$, $\mu(x) = k(x, x_0)$ with $x_0 \in \mathcal{X}$.

■ **Results**: Our regret analysis indicates that by taking η as an appropriate small value, our algorithm achieves better performance.

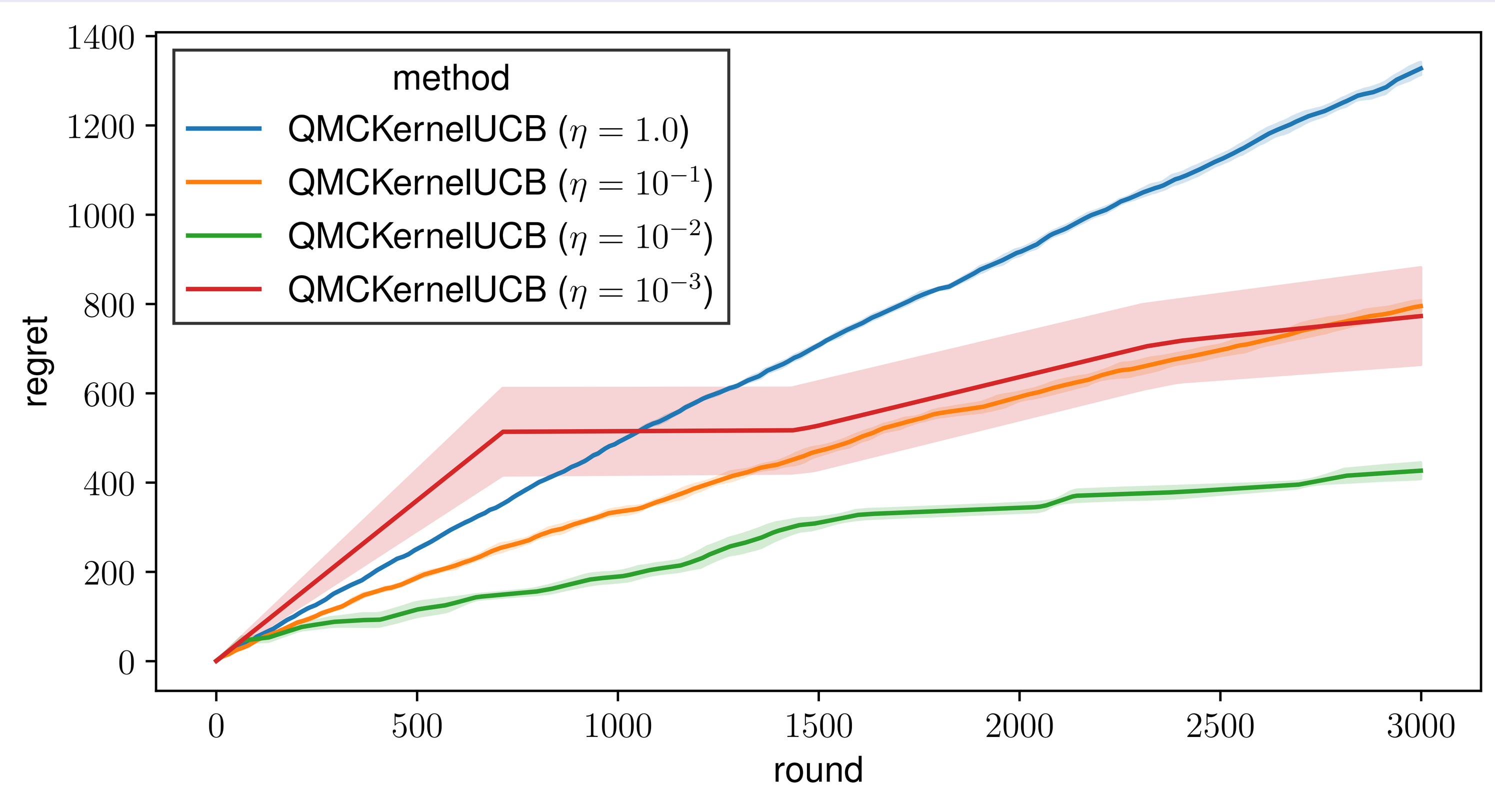

Figure: Ablation study of the parameter η.

- In the following figure, for better emprical performance, we introduce an exploration parameter v to the UCB $\widetilde{\mu}_{s-1}(x) + v\beta_{s-1}\widetilde{\sigma}_{s-1}(x)$.

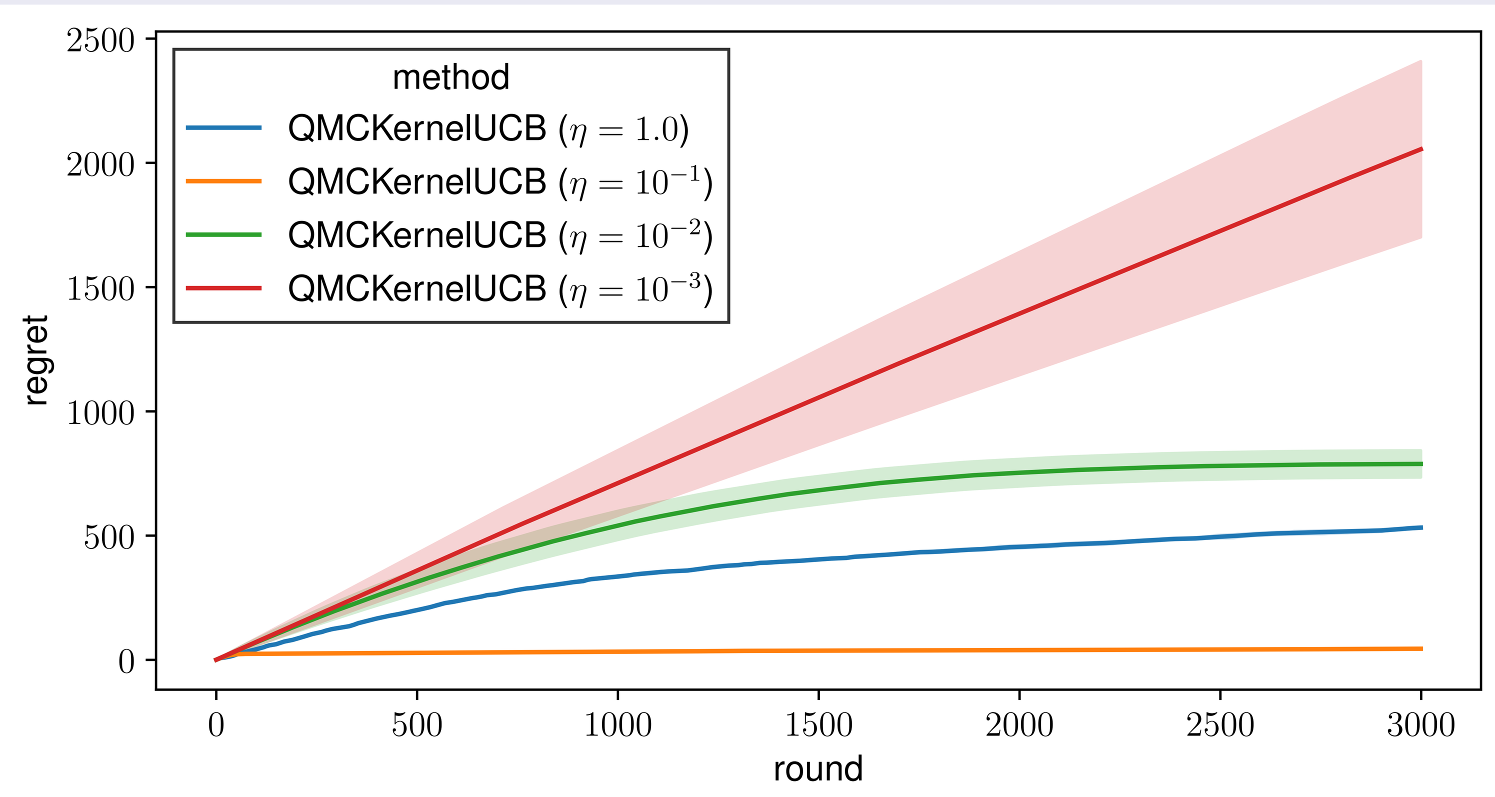

Figure: Ablation study of the parameter η with the exploration parameter $v = 0.1$

References

[1] Dai, Z., Lau, G. K. R., Verma, A., Shu, Y., Low, B. K. H., and Jaillet, P. (2023). Quantum Bayesian Optimization. In *Thirty-seventh Conference on Neural Information Processing Systems*.

[2] Montanaro, A. (2015). Quantum speedup of monte carlo methods. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 471(2181):20150301.

[3] Vakili, S., Khezeli, K., and Picheny, V. (2021). On information gain and regret bounds in Gaussian process bandits. In *International Conference on Artificial Intelligence and Statistics*, pages 82–90. PMLR.

[4] Wan, Z., Zhang, Z., Li, T., Zhang, J., and Sun, X. (2023). Quantum multi-armed bandits and stochastic linear bandits enjoy logarithmic regrets. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 10087–10094.