

# Accelerated Adaptive Decision Making for Autonomous Agents: Optimization and Coverage

Doctoral Consortium

Zakaria Shams Siam

University at Albany, State University of New York  
Albany, NY, United States  
zsiam@albany.edu

## ABSTRACT

We study adaptive decision making under uncertainty for autonomous agents that interact with expensive black-box evaluators under strict evaluation budgets. These settings arise naturally in agentic scientific discovery where an agent must sequentially choose what to test next while balancing exploration and exploitation and respecting safety or feasibility constraints in the objective space. A key theme of my dissertation is acceleration: characterizing how faster estimation primitives can fundamentally improve sample efficiency, and designing algorithms that exploit such acceleration in principled ways. As initial results, (i) for sequential decision making in non-linear bandits, we propose Q-NLB-UCB, a new input dimension-free quantum non-linear bandit algorithm with  $O(\text{poly log } T)$  regret, (ii) for adaptive constrained objective-space coverage, we further formulate the new multi-objective coverage (MOC) problem where our goal is to identify a small set of representative samples whose predicted outcomes broadly cover the feasible multi-objective space. To solve this MOC problem, we propose a novel search algorithm, MOC-CAS, which employs an upper confidence bound-based acquisition function to select optimistic samples guided by Gaussian process posterior predictions. Empirically, MOC-CAS achieves superior performances over competitive baselines on large-scale protein-target datasets. Going forward, we will develop acceleration-aware theory for optimistic exploration under explicit cost models, study multi-agent extensions for parallel experimentation, and extend to graph-structured action spaces using uncertainty-aware learned representations.

## KEYWORDS

Adaptive decision making; Bandit optimization; Constrained objective-space coverage; Autonomous scientific discovery

### ACM Reference Format:

Zakaria Shams Siam. 2026. Accelerated Adaptive Decision Making for Autonomous Agents: Optimization and Coverage: Doctoral Consortium. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026)*, Paphos, Cyprus, May 25 – 29, 2026, IFAAMAS, 3 pages. <https://doi.org/10.65109/ARNT2761>



This work is licensed under a Creative Commons Attribution International 4.0 License.

*Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026)*, C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems ([www.ifaamas.org](http://www.ifaamas.org)). <https://doi.org/10.65109/ARNT2761>

## 1 INTRODUCTION

Autonomous agents are increasingly utilized to support scientific discovery loops in which an agent repeatedly selects an experiment or candidate to evaluate, observes noisy outcomes, and updates its beliefs to decide what to experiment next [3, 4, 6, 8, 18]. In several real-world domains such as drug discovery [8] and materials design [4], evaluations are expensive and budgets are limited. Consequently, an agent must make adaptive decisions under uncertainty that carefully trade off exploration and exploitation while respecting feasibility and safety constraints in the objective space.

This paper focuses on sample-efficient sequential decision making for such agents. A central theme is acceleration: beyond algorithmic design alone, contemporary compute primitives (e.g., quantum subroutines) can improve what is possible under limited evaluation budgets. This raises a fundamental question: how can we design algorithms for autonomous agents that enable accelerated adaptive decision making for scientific discoveries?

## 2 DISSERTATION VISION AND RESEARCH QUESTIONS

The long-term goal of my dissertation is to develop a principled framework for accelerated adaptive decision making under uncertainty for autonomous agents. Concretely, we study (1) accelerated black-box optimization, where the agent seeks to maximize an unknown expensive black-box function under bandit feedback and (2) objective-space coverage, where the agent seeks to efficiently characterize the feasible multi-objective outcome space by selecting a small representative and safe set of candidates. This dissertation revolves around three major research questions:

- (i) For expensive non-linear bandit problems in high dimensions, how can an agent optimize without suffering from the curse of dimensionality?
- (ii) How can an agent efficiently cover the feasible multi-objective outcome space under constraints?
- (iii) How do acceleration mechanisms with the aid of quantum computing change achievable sample efficiency, and how can agents exploit them in a unified way across different decision objectives?

## 3 PROGRESS TO DATE

### 3.1 Quantum Computing-Accelerated Bandit Optimization

First, we studied non-linear bandit optimization where the learner maximizes a black-box function with zeroth order function oracle. Unfortunately, an  $\Omega(\sqrt{T})$  cumulative regret lower bound [14]

cannot be further improved in classical (non-quantum) setting. It implies that given total  $T$  rounds, no classical algorithm can incur cumulative regret less than  $\Omega(\sqrt{T})$  asymptotically. Existing works have showed that with the aid of quantum computing [13], it is possible to break the classical  $\Omega(\sqrt{T})$  regret lower bound and achieve the new  $O(\text{poly log } T)$  upper bound [1, 5, 17], where  $T$  denotes the number of function evaluations. However, they usually assume that the objective function sits within the reproducing kernel Hilbert space and their algorithms suffer from the curse of dimensionality. To address these limitations, we developed the new *Quantum Non-Linear Bandit with Upper Confidence Bound* (Q-NLB-UCB) algorithm [15] which enjoys an *input dimension-free*  $O(\text{poly log } T)$  upper bound, making it applicable for high-dimensional tasks. At the heart of our algorithm design are quantum Monte Carlo mean estimator [12], parametric function approximation technique [9], and a new quantum non-linear regression oracle, which can be of independent interests in more quantum machine learning problems. Different from existing works, Q-NLB-UCB does not suffer from the curse of dimensionality. We prove the first  $O(d_w^2 \log^{\frac{3}{2}}(T) \log(d_w \log T))$  regret bound with  $d_w$  being parameter complexity, which is also faster than the classical lower bound  $\Omega(\sqrt{T})$  but *independent* to input dimension  $d_x$  [15].

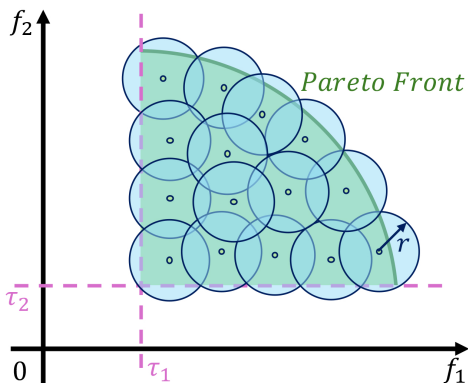


Figure 1: An example of the MOC problem where  $f_1, f_2$  are two objective functions and  $\tau_1, \tau_2$  are two thresholds defining the feasible region shown in light green. 15 samples are selected as representative feasible samples, each with a coverage ball whose radius is  $r$ .

### 3.2 Constrained Multi-Objective Outcome-Space Coverage

In parallel, we investigated settings where the agent’s goal is not to optimize a single objective, but to identify a small set of candidates whose outcomes *broadly cover* the feasible multi-objective space under constraints. This representative set can be evaluated much faster than the whole feasible set, thus significantly accelerating the scientific discovery process. This is a new problem formulation that we introduce, along with a principled algorithm to solve it. Existing works cannot be directly applied as they either focus on sample space coverage or multi-objective optimization that targets the

Pareto front. However, chemically diverse samples often yield identical objective profiles, and safety constraints are usually defined on the objectives. We therefore formulated the new multi-objective coverage (MOC) problem [16], explicitly targeting fast coverage and diversity of the feasible output region under per-objective thresholds. See Figure 1 for an example with two objective functions. This is distinct from the constraint active search (CAS) framed in the sample space [11] and also different from other frontier-centric objectives [2, 7, 19]. To solve the problem, we propose the *multi-objective coverage via constraint active search* (MOC-CAS) algorithm [16], which is a novel constraint active search algorithm that evaluates each candidate based on the optimistic estimate of the new feasible volume it is expected to cover within the feasible region. It further employs a tie-breaking mechanism that promotes diversity by encouraging dispersion in the predicted objective values. Compared to the competitive baselines, our MOC-CAS empirically achieves superior performances across large-scale protein-target datasets for SARS-CoV-2 and cancer [10], each assessed on five objectives derived from SMILES-based features [16].

## 4 RESEARCH PLAN AND FUTURE DIRECTIONS

My dissertation pursues acceleration-aware adaptive decision making under uncertainty for autonomous agents interacting with expensive oracles. My personal view is that progress depends on explicitly separating and optimizing the costs of real evaluations versus estimation or inference, since end-to-end runtime, not only sample efficiency, is the dominant bottleneck in discovery pipelines. We organize our future work around the following research questions (RQs):

**RQ 1: When and how does acceleration change what an agent can learn or decide under a fixed evaluation budget?** We will develop theory and algorithms that model the cost structure of decision making (expensive evaluations + potentially accelerated estimation or inference) and derive principled conditions under which acceleration provably improves achievable regret or sample complexity.

**RQ 2: How can multiple agents coordinate to achieve constraint-aware objective-space exploration efficiently?** We will extend our constraint-aware objective-space exploration to batch and multi-agent settings that reflect parallel experimentation, designing centralized or distributed coordination rules that balance optimism, feasibility, and diversity under shared budgets and limited communication.

**RQ 3: How can optimistic exploration scale to graph-structured decision spaces?** We will study graph-structured bandit problems (e.g., molecule/protein graphs) by integrating uncertainty-aware learned representations (e.g., graph neural network-based models) with optimism-based decision rules and efficient search over large discrete, graph-structured candidate spaces.

Ultimately, these advances aim to make closed-loop discovery agents more practical by reducing redundant evaluations and increasing the number of safe, diverse high-quality candidates identified per unit time and budget, thus supporting self-driving laboratory workflows in drug and materials discovery.

REFERENCES

[1] Zhongxiang Dai, Gregory Kang Ruey Lau, Arun Verma, Yao Shu, Bryan Kian Hsiang Low, and Patrick Jaillet. 2023. Quantum bayesian optimization. *NeurIPS* 36 (2023).

[2] Michael TM Emmerich, André H Deutz, and Jan Willem Klinkenberg. 2011. Hypervolume-based expected improvement: Monotonicity properties and exact computation. In *2011 IEEE congress of evolutionary computation (CEC)*. IEEE, 2147–2154.

[3] Peter I Frazier. 2018. A tutorial on Bayesian optimization. *arXiv preprint arXiv:1807.02811* (2018).

[4] Peter I Frazier and Jialei Wang. 2016. Bayesian optimization for materials design. *Information science for materials discovery and design* (2016), 45–75.

[5] Yasunari Hikima, Kazunori Murao, Sho Takemori, and Yuhei Umeda. 2024. Quantum Kernelized Bandits. In *UAI*.

[6] Kirthevasan Kandasamy, Willie Neiswanger, Jeff Schneider, Barnabas Poczos, and Eric P Xing. 2018. Neural architecture search with bayesian optimisation and optimal transport. *NIPS* 31 (2018).

[7] Joshua Knowles. 2006. ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE transactions on evolutionary computation* 10, 1 (2006), 50–66.

[8] Ksenia Korovina, Sailun Xu, Kirthevasan Kandasamy, Willie Neiswanger, Barnabas Poczos, Jeff Schneider, and Eric Xing. 2020. Chembo: Bayesian optimization of small organic molecules with synthesizable recommendations. In *AISTATS*.

[9] Chong Liu and Yu-Xiang Wang. 2023. Global optimization with parametric function approximation. In *ICML*.

[10] Xuefeng Liu, Songhao Jiang, Archit Vasan, Alexander Brace, Ozan Gokdemir, Thomas Brettin, Fangfang Xia, Ian Foster, and Rick Stevens. 2023. DRUGIM-PROVER: Utilizing reinforcement learning for multi-objective alignment in drug optimization. In *NeurIPS 2023 Workshop on New Frontiers of AI for Drug Discovery and Development*, Vol. 1. 8.

[11] Gustavo Malkomes, Bolong Cheng, Eric H Lee, and Mike Mccourt. 2021. Beyond the pareto efficient frontier: Constraint active search for multiobjective experimental design. In *ICML*.

[12] Ashley Montanaro. 2015. Quantum speedup of Monte Carlo methods. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 471, 2181 (2015), 1–20.

[13] Michael A Nielsen and Isaac L Chuang. 2010. *Quantum computation and quantum information*. Cambridge university press.

[14] Jonathan Scarlett, Ilija Bogunovic, and Volkan Cevher. 2017. Lower bounds on regret for noisy gaussian process bandit optimization. In *COLT*.

[15] Zakaria Shams Siam, Chaowen Guan, and Chong Liu. 2026. Quantum Non-Linear Bandit Optimization. In *Proceedings of the 40th AAAI Conference on Artificial Intelligence (AAAI)*.

[16] Zakaria Shams Siam, Xuefeng Liu, and Chong Liu. 2026. Multi-Objective Coverage via Constraint Active Search. In *Proceedings of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*.

[17] Zongqi Wan, Zhijie Zhang, Tongyang Li, Jialin Zhang, and Xiaoming Sun. 2023. Quantum multi-armed bandits and stochastic linear bandits enjoy logarithmic regrets. In *AAAI*.

[18] Jian Wu, Saul Toscano-Palmerin, Peter I Frazier, and Andrew Gordon Wilson. 2020. Practical multi-fidelity Bayesian optimization for hyperparameter tuning. In *UAI*.

[19] Kaifeng Yang, Michael Emmerich, André Deutz, and Thomas Bäck. 2019. Efficient computation of expected hypervolume improvement using box decomposition algorithms. *Journal of Global Optimization* 75, 1 (2019), 3–34.