Quantum Speedups for Sampling and Non-convex Optimization with Stochastic Oracles

Guneykan Ozgul, Xiantao Li, Mehrdad Mahdavi, and Chunhao Wang

We present quantum speedups for sampling from probability distributions of the form $\pi \propto e^{-f}$, where $f : \mathbb{R}^d \mapsto \mathbb{R}$. We consider two oracle models: (i) a stochastic gradient oracle, where f is in finite sum form, i.e., $f(x) = \frac{1}{n}\sum_{i=1}^{n} f_i(x)$ and individual component gradients are accessible, (ii) a stochastic zeroth-order oracle, where only noisy evaluations of f are available.

Our main contribution is a general framework for quantumly accelerating classical stochastic sampling algorithms, such as Langevin Monte Carlo (LMC) and Hamiltonian Monte Carlo (HMC), by replacing stochastic gradient computations with variance-controlled quantum mean and gradient estimation subroutines. In contrast to prior quantum sampling approaches based on quantum walks, our methods do not require reversibility, or exact gradient access, and preserve the structure of the underlying (possibly nonreversible) Markov chain.

In the finite-sum gradient oracle model, we integrate unbiased quantum mean estimation with classical variance-reduction techniques, including stochastic variance-reduced gradients (SVRG) and control variates (CV). By jointly optimizing the target variance of quantum estimators and the frequency of full-gradient recomputation, we obtain provable improvements in gradient query complexity over the best known classical samplers. These results apply both to strongly log-concave and to non-logconcave distributions satisfying a log-Sobolev inequality, with convergence guarantees in Wasserstein distance and Kullback–Leibler divergence.

In the stochastic zeroth-order model, we develop new quantum gradient estimation procedures that are robust to noisy and potentially unbounded function evaluations. These estimators lead to improved evaluation complexity for quantum-accelerated LMC and HMC under standard smoothness assumptions.

Finally, we show that faster quantum sampling yields quantum speedups for optimization, including nonsmooth and approximately convex objectives. This recovers known quantum advantages for finite-sum optimization and establishes new improvements in the zeroth-order stochastic setting.

1 Introduction

Given a potential function $f : \mathbb{R}^d \to \mathbb{R}$, we consider the problem of sampling from a probability distribution $\pi \in L^1(\mathbb{R}^d, d\mathbf{x})$ of the form

$$\pi(\mathbf{x}) = \frac{e^{-f(\mathbf{x})}}{\int e^{-f(\mathbf{x})}\mathrm{d}\mathbf{x}}. \tag{1}$$

This distribution is called the *Gibbs-Boltzmann distribution*, and our goal is to efficiently sample approximately from π.

Efficient sampling from Gibbs-Boltzmann distributions is a fundamental problem in many scientific and engineering disciplines, becoming increasingly important as modern applications deal with high-dimensional data and complex probabilistic models [11, 3, 7, 9]. However, even approximately sampling from this distribution is NP-hard in general and efficient sampling is only possible when f has certain local and global properties. In many applications with practical interest, the function value f or its gradients ∇f are not easily accessible. For example, in chemistry or molecular physics, f represents the energy of the system which is estimated by a Monte Carlo simulation.

Consequently, the overall sampling cost increases, since one typically needs very accurate estimates of f or ∇f to ensure that the output distribution is close to the target distribution. Motivated by these challenges, we consider the following *stochastic oracle* settings:

- The *stochastic gradient oracle*. In this setting we assume $f(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} f_i(\mathbf{x})$, and we assume access to the gradients of each component, i.e., $\nabla f_i(x)$ for any $i \in [n]$. This is a relevant problem, especially in statistics and machine learning problems, where access to the function f is provided only through sample data [11].

- The *stochastic zeroth-order oracle*. In this setting, we assume that we have access only to noisy function values $f(\mathbf{x};\xi)$, where ξ is a random variable that characterizes the noise. This setting is more relevant in contexts where the function value f is replaced by a statistical estimate, perhaps obtained through a numerical simulation [9]. Since this is essentially a noisy evaluation oracle, throughout the paper we use "zeroth-order oracle" and "evaluation oracle" interchangeably.

Given the challenging nature of the problem, it is natural to ask whether quantum computers can provide any speedup for the sampling task. The state-of-the-art classical algorithms are in fact Markov chains and this is a domain where quantum algorithms have been studied extensively [10, 12, 1]. Childs et al. initiated the study of sampling from log-concave distributions essentially by assuming that f is a strongly convex function and the gradient of f can be accessed up to very small error [4]. Their construction uses quantum walk framework of [10] and fixed point amplitude amplification similar to [12]. While the authors manage to improve the state-of-the-art classical algorithms in terms of certain parameters, their speedup is not applicable in stochastic setting. The authors also considered zeroth-order setting separately, however in this setting they could not use quantum walk framework. Therefore, they only obtained the same complexity as the classical algorithms (See Table 1 in [4]).

This work is later extended by Ozgul et al. in [8] to the stochastic non-convex setting. The setting in [8] allows stochastic gradients instead of exact gradients. However, the caveat in both these works is that the classical Markov chains do not satisfy a crucial assumption in quantum walk frameworks which is known as *reversibility* or *detailed-balanced*. A Markov chain with transition kernel P on Ω is said to be reversible with respect to π if for any $x, y \in \Omega$,

$$\pi(x)P(x,y) = \pi(y)P(y,x) \tag{2}$$

Due to this challenge, both works try to make the Markov chains reversible. While Childs et al. use Metropolis-Hastings adjustment, Ozgul et al. use very small step sizes so that the Markov chain stays close to a reversible Markov chain. However, it is well known that nonreversible Markov chains can be faster than their reversible counterparts, and the separation can be quadratic. Therefore, the improvements in these works can be sub-optimal even in the deterministic setting. We also note that, in the stochastic setting, even if the Markov chain is reversible, inexact gradients cause the chain to be nonreversible and unfortunately Metropolis–Hastings type corrections are typically infeasible because they require evaluating f exactly. Although there are some recent progress in accelerating non-reversible Markov chains by the authors in [5], this technique is similar to the technique used in [8] and requires a reversible Markov chain to bound the runtime.

Moreover, existing quantum results in this line primarily guarantee convergence in total variation distance. Obtaining quantum speedups with guarantees in Wasserstein distance or Kullback–Leibler divergence, metrics that are standard in classical sampling theory, remains largely open. Motivated by these challenges, we ask the following question:

Question 1. *Can quantum computers improve the oracle complexity of sampling algorithms such that*

1. *the algorithm can use stochastic gradient and/or zeroth-order oracles,*
2. *it can be based on fast (potentially nonreversible) Markov chains rather than a reversibilized surrogate, and*
3. *it provides guarantees in Wasserstein distance and Kullback–Leibler divergence?*

We answer Question 1 affirmatively by introducing a framework that replaces stochastic gradient computation with variance-controlled quantum subroutines, yielding provable oracle-complexity improvements while keeping the underlying (possibly nonreversible) Markov chain intact. Conceptually, our contribution is not a new sampler, but a principled way to quantum-accelerate existing stochastic samplers by exploiting variance structure (via quantum mean/gradient estimation) rather than relying on spectral-gap-based quantum-walk techniques. However, a naïve drop-in replacement of classical stochastic gradients with quantum mean/gradient estimation does not automatically yield an end-to-end speedup. The reason is that state-of-the-art finite-sum samplers already use variance reduction (e.g., SVRG and control variates) and periodically pay for full-gradient computations to control variance; if these full-gradient steps are still frequent, they can dominate the overall cost and dominate any quantum advantage. Thus, realizing a genuine improvement requires jointly tuning (i) how accurately the stochastic gradient is estimated and (ii) how often full gradients are recomputed. In the stochastic zeroth-order setting standard quantum gradient-estimation routines typically assume high-precision function access, which is incompatible with noisy evaluations. Another motivation for this work is to investigate whether quantum speedups for certain optimization problems can be obtained via our sampling algorithms. Sampling and optimization are tightly connected: sampling from a Gibbs distribution with potential βf at large inverse temperature β concentrates around (approximate) minimizers of f. Although using sampling as a generic optimization method is not always optimal, it can be effective in regimes where gradient information is noisy or the objective is nonsmooth/nonconvex. To this end, we ask the following question:

Question 2. *Can we use the quantum sampling algorithm to recover the existing quantum speedups for certain optimization problems and/or identify new class of optimization problems that can be accelerated through sampling?*

We also give an affirmative answer to this question by designing optimization algorithms based on quantum sampling algorithms. In particular, we recover the quantum speedup obtained in [13] for finite-sum optimization problems. In the zeroth-order setting, we improve some of the parameter dependencies for nonsmooth optimization problems considered in [6, 2]. While most of the resulting improvements are sub-quadratic, the point of our work is not focusing on the degree of the speedup for these particular problems. More broadly, we highlight that viewing optimization through the lens of sampling can be a useful organizing principle for identifying future quantum speedups.

References

[1] Simon Apers and Alain Sarlette. Quantum fast-forwarding: Markov chains and graph property testing. *Quantum Info. Comput.*, 19(3–4):181–213, March 2019.

[2] Shouvanik Chakrabarti, Dylan Herman, Jacob Watkins, Enrico Fontana, Brandon Augustino, Junhyung Lyle Kim, and Marco Pistoia. On speedups for convex optimization via quantum dynamics, 2025.

[3] David Chandler. Introduction to modern statistical. *Mechanics. Oxford University Press, Oxford, UK*, 5(449):11, 1987.

[4] Andrew M. Childs, Tongyang Li, Jin-Peng Liu, Chunhao Wang, and Ruizhe Zhang. Quantum algorithms for sampling log-concave distributions and estimating normalizing constants. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 23205–23217. Curran Associates, Inc., 2022.

[5] Baptiste Claudon, Jean-Philip Piquemal, and Pierre Monmarché. Quantum speedup for non-reversible markov chains, 2025.

[6] Tongyang Li and Ruizhe Zhang. Quantum speedups of optimizing approximately convex functions with applications to logarithmic regret stochastic convex bandits. In *Proceedings of the 36th International Conference on Neural Information Processing Systems*, NIPS '22, Red Hook, NY, USA, 2024. Curran Associates Inc.

[7] László Lovász and Santosh Vempala. Simulated annealing in convex bodies and an o*(n4) volume algorithm. *Journal of Computer and System Sciences*, 72(2):392–417, 2006. JCSS FOCS 2003 Special Issue.

[8] Guneykan Ozgul, Xiantao Li, Mehrdad Mahdavi, and Chunhao Wang. Stochastic quantum sampling for non-logconcave distributions and estimating partition functions. In Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp, editors, *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 38953–38982. PMLR, 21–27 Jul 2024.

[9] Abhishek Roy, Lingqing Shen, Krishnakumar Balasubramanian, and Saeed Ghadimi. Stochastic zeroth-order discretizations of langevin diffusions for bayesian inference, 2021.

[10] M. Szegedy. Quantum speed-up of markov chain based algorithms. In *45th Annual IEEE Symposium on Foundations of Computer Science*, pages 32–41, 2004.

[11] Max Welling and Yee Whye Teh. Bayesian learning via stochastic gradient langevin dynamics. In *Proceedings of the 28th International Conference on International Conference on Machine Learning*, ICML'11, page 681–688, Madison, WI, USA, 2011. Omnipress.

[12] Pawel Wocjan and Anura Abeyesinghe. Speedup via quantum sampling. *Phys. Rev. A*, 78:042336, Oct 2008.

[13] Yexin Zhang, Chenyi Zhang, Cong Fang, Liwei Wang, and Tongyang Li. Quantum algorithms and lower bounds for finite-sum optimization. In Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp, editors, *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 60244–60270. PMLR, 21–27 Jul 2024.