



Title	Hybrid quantum-classical Markov chain Monte Carlo algorithms for near-term quantum computers
Author(s)	中野, 裕一郎
Citation	大阪大学, 2026, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/105534
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

**Hybrid quantum-classical
Markov chain Monte Carlo algorithms
for near-term quantum computers**

YUICHIRO NAKANO

MARCH, 2026

**Hybrid quantum-classical
Markov chain Monte Carlo algorithms
for near-term quantum computers**

A dissertation submitted to
THE GRADUATE SCHOOL OF ENGINEERING SCIENCE
THE UNIVERSITY OF OSAKA
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY IN ENGINEERING

BY

YUICHIRO NAKANO

MARCH, 2026

Abstract

Quantum computing is expected to offer exponential or polynomial speedups over classical computing for many socially important problems, including integer factorization, combinatorial optimization, and quantum chemistry. This promise has fueled strong expectations for quantum computers as next-generation computing infrastructure. Most algorithms with provable quantum advantage require large-scale quantum error correction, making fault-tolerant quantum computers (FTQC) essential. However, currently available devices are noisy intermediate-scale quantum (NISQ) devices, and further breakthroughs, such as advances in large-scale integration, are required to realize practical implementations of quantum error correction.

Given these constraints, extensive research has focused on NISQ-friendly algorithms that can extract useful performance from small, noisy quantum circuits over the past 10 years. A central idea is hybrid quantum-classical computation, where quantum circuits are combined with classical processing. A representative approach is variational quantum algorithms (VQAs), which iteratively evaluate parameterized quantum circuits and update parameters using classical optimization, typically by estimating expectation values of physical quantities. Although VQAs have been proposed for many applications including combinatorial optimization and quantum chemistry, they face major obstacles, including high measurement costs for expectation estimation and optimization difficulties such as barren plateaus. As a result, clear practical quantum advantage has not yet been established. Against this background, sampling-based NISQ algorithms that directly exploit samples obtained from quantum computers have recently attracted attention as an approach distinct from VQAs.

This thesis investigates hybrid quantum-classical Markov chain Monte Carlo (MCMC), one of the sampling-based NISQ algorithms, as an alternative route to quantum speedup on NISQ devices. In this framework, the output distribution of a quantum circuit is used as the proposal distribution within MCMC. Sampling from the quantum circuit produces new transitions for the Markov chain, and the accept-reject step is performed classically based on the acceptance probability between transitions. Unlike VQAs, the method requires neither variational optimization nor expectation-value measurements, which can substantially reduce the number of quantum circuit executions. In particular, this thesis focuses on sampling from the Boltzmann distribution of Ising models, a task that underpins many problems in combinatorial optimization and machine learning, and experimentally demonstrates that the proposed approach can outperform existing classical MCMC techniques. Previous studies have applied hybrid quantum-classical MCMC to Boltzmann distribution sampling, but their implementations face challenges such as restrictions on the quantum circuits that can be used and the lack of systematic methods for optimizing circuit

parameters. Addressing these issues is crucial for realizing practical quantum speedup in the NISQ era. To advance the development and applicability of hybrid quantum-classical MCMC, this thesis presents three core studies.

First, we propose Quantum Alternating Operator Ansatz Monte Carlo (QAOA-MC), a hybrid quantum-classical MCMC algorithm that uses a fixed-depth parameterized quantum circuit based on QAOA as the proposal distribution to accelerate convergence while remaining compatible with NISQ devices. To systematically improve mixing, we introduce an optimization strategy based on the relationship between the spectral gap and the acceptance rate, showing that appropriate parameter restriction and acceptance rate minimization can enhance convergence. Numerical experiments on Boltzmann sampling of spin glass models demonstrate a near-quadratic speedup over uniform proposals and validate efficient magnetization estimation for systems of up to 15 spins.

Second, we develop a hybrid quantum-classical MCMC framework that integrates quantum circuit sampling with a generative neural sampler to overcome symmetry constraints imposed in existing methods. By approximating the output distribution of an arbitrary quantum circuit with a neural network surrogate, the proposed approach enables flexible circuit design without requiring explicit evaluation of proposal probabilities. Applied to Boltzmann sampling of fully connected spin glass models, the method achieves a polynomial speedup (over cubic) at low temperatures and enables efficient, unbiased sampling using shallow QAOA circuits, even without variational parameter optimization.

Finally, we investigate the fair sampling properties of hybrid quantum-classical MCMC and demonstrate its applicability to solution-sampling tasks in constraint satisfaction problems. Through analyses of Ising model toy systems, we show that combining quantum proposals with MCMC updates restores fair sampling. Extending the study to random k -SAT problems, we find that the proposed method achieves fairness comparable to state-of-the-art classical algorithms for 2-SAT and remains effective for 3-SAT, where such classical methods are no longer applicable.

Through these three studies, this thesis demonstrates the systematic development of hybrid quantum-classical MCMC algorithms as an effective framework for extracting the advantages of quantum computation even under the constraints of NISQ devices. In particular, by realizing quantum proposals through a combination of neural networks and quantum circuits without variational optimization, and by verifying the practically important property of fair sampling, this work substantially broadens the applicability of quantum heuristics. These results constitute an important step toward the realization of practical sampling algorithms that exploit near-term quantum devices, and they are expected to stimulate further developments across a wide range of fields, including statistical physics, machine learning, and combinatorial optimization.

Contents

1	Introduction	1
1.1	Dawn of Quantum Computation: Establishing the Theoretical Foundations . . .	1
1.2	NISQ Era: Hardware Development and Applications	2
1.3	Beyond NISQ Era	3
1.4	Overview	4
1.5	Outline	5
2	Preliminaries	9
2.1	Notations	9
2.2	Model	9
2.2.1	Ising Model	9
2.2.2	Boltzmann Distribution	11
2.2.3	Applications of Ising Model	11
2.3	Markov Chain Monte Carlo (MCMC)	12
2.3.1	Markov Chain	12
2.3.2	Metropolis–Hastings Algorithm	13
2.3.3	Proposal Distribution	14
2.3.4	Convergence	15
2.4	Foundations of Quantum Computation	17
2.4.1	Quantum State	17
2.4.2	Quantum Gate	18
2.4.3	Quantum Measurement	18
2.5	Quantum Approximate Optimization Algorithm (QAOA)	20
2.5.1	Algorithm	20
2.5.2	Challenges	21
2.6	Hybrid Quantum-Classical MCMC	22
2.7	Summary and Discussion	23
3	Optimization of Quantum Proposals Inspired by QAOA	27
3.1	Introduction	27
3.2	QAOA-MC	29
3.2.1	Parametrized Quantum Proposal	29
3.2.2	Optimization of Quantum Proposal	30
3.3	Numerical Experiments	31

3.3.1	Spectral Gap	33
3.3.2	Optimization with MCMC Estimator of AR	35
3.3.3	Magnetization Estimation	36
3.4	Conclusion	38
4	Neural-Network-Assisted Monte Carlo Sampling Trained by QAOA	41
4.1	Introduction	41
4.2	Generative Neural Sampler	43
4.3	QAOA-NMC	45
4.3.1	Algorithm	45
4.3.2	Boltzmann Sampling with Neural Proposal Trained by QAOA	46
4.4	Numerical Experiments	48
4.4.1	Spectral Gap	48
4.4.2	Magnetization Estimation	50
4.5	Conclusion	53
5	Fair Sampling of Ground-State Configurations	55
5.1	Introduction	55
5.2	Degenerate Ground-State Sampling for Small Systems	57
5.3	Random k -SAT Sampling	59
5.3.1	Problem Setting	59
5.3.2	Numerical Experiments	60
5.3.3	Fair Sampling	62
5.3.4	Model Counting	65
5.4	Conclusion	68
6	Conclusion	69
	Appendix A Spectral Gap For Different Temperatures	73
	Bibliography	75
	Acknowledgments	89
	List of Achievements	91

Chapter 1

Introduction

1.1 Dawn of Quantum Computation: Establishing the Theoretical Foundations

The history of quantum computation originated from fundamental concerns regarding the limitations of classical computers in simulating physical systems. In the early 1980s, Richard Feynman pointed out that an efficient simulation of nature, which obeys the laws of quantum mechanics, would require a computing device that itself operates according to quantum-mechanical principles [1]. In response to this insight, David Deutsch formalized the concept of a universal quantum Turing machine in 1985 [2]. He demonstrated that computation exploiting quantum superposition and interference can violate the extended Church–Turing thesis [3]. Through this work, quantum computers came to be defined not merely as tools for physics but as a new framework within computer science. As this theoretical foundation was established, two algorithms introduced in the mid-1990s ignited widespread interest in research on quantum computation. The first was the integer factorization algorithm proposed by Peter Shor in 1994, commonly known as Shor’s algorithm [4]. This algorithm showed that integer factorization, which underlies modern public-key cryptographic systems, can be solved in polynomial time on a quantum computer. As a result, it provided a clear example of an exponential computational advantage over classical algorithms. In addition, in 1996, Lov Grover proposed an algorithm that accelerates the search in an unstructured database [5]. These developments demonstrated that quantum computers can efficiently solve certain problems with specific mathematical structures that are intractable for classical computers within practical time scales, thereby attracting considerable attention as a potential next-generation computing paradigm.

Despite the demonstrated usefulness of quantum algorithms, there was widespread skepticism regarding their practical feasibility. A common concern was that quantum states are extremely fragile and that interactions with the environment or errors during computation would rapidly destroy the encoded information. This pessimistic view was fundamentally challenged by the concept of quantum error correction proposed by Shor in 1995 [6]. He showed that by encoding information redundantly using quantum entanglement, it is possible to detect and correct errors occurring on individual qubits without directly measuring and disturbing the stored quantum information. Subsequently, the threshold theorem was established, proving that arbitrarily long

quantum computations are possible provided that the error rate remains below a certain threshold [7, 8]. This result transformed quantum computation from a purely theoretical model into an engineering goal that is, in principle, achievable. Supported by the theory of quantum error correction, the development of quantum computers has since proceeded with the realization of fault-tolerant quantum computation (FTQC) as its ultimate objective. From that point onward, while research progressed on the theoretical foundations of quantum computation and on algorithms whose advantage over classical computation is theoretically guaranteed [9, 10], hardware development remained largely limited to experimental demonstrations of basic qubits and quantum gates [11–13]. As a result, research on quantum algorithms continued to be conducted primarily within a theoretical framework. This situation persisted until a major shift occurred in the 2010s.

1.2 NISQ Era: Hardware Development and Applications

The history of device development from the early to mid-2010s marked a transition of quantum computation from a theoretical concept to an object of practical engineering implementation. In recent years, the realization of high-fidelity quantum gates below the error threshold required for surface codes has driven rapid progress in the integration of quantum devices across various physical platforms, most notably superconducting circuits [14]. In 2018, John Preskill introduced the concept of noisy intermediate-scale quantum (NISQ) devices, referring to medium-scale quantum processors that lack full error-correction capabilities but may still outperform classical computers for certain tasks [15]. This definition provided a clear and widely shared research direction for the field, focusing on how imperfect qubits can be utilized effectively. Hardware development in the NISQ era has steadily improved both the number and quality of qubits, with superconducting platforms leading this progress. For example, IBM and Google have demonstrated increasing scalability with devices such as IBM’s 127-qubit Eagle processor [16] and Google’s 105-qubit Willow processor [17]. On the other hand, from the mid-2020s onward, significant advances have also been observed in platforms beyond superconducting circuits. Neutral atom systems [18] and trapped-ion systems [19, 20] have demonstrated high connectivity and gate fidelity, offering complementary strengths and expanding the range of viable architectures for commercial applications.

As quantum hardware reached a sufficient scale, a major milestone of the NISQ era was the experimental demonstration of “quantum supremacy”. In 2019, Google reported that its 53-qubit Sycamore processor performed a specific task known as random circuit sampling [21–23] significantly faster than the most powerful classical supercomputers available at that time [24]. This result demonstrated that quantum processors can execute computations that are difficult to simulate classically, providing concrete evidence that quantum computation can surpass classical computational capabilities for specific tasks. Subsequent experiments using both photonic and superconducting platforms reported increasingly sophisticated results [25–27], establishing quantum computational supremacy as a general milestone that does not depend on a single hardware implementation. While these achievements do not directly translate into immediate practical advantages, they have convincingly demonstrated the potential computational power of quantum devices and have served as a strong driving force for further development.

In parallel with demonstrations of computational capability on real hardware, active efforts have been made to develop applications that can deliver practical benefits despite the presence of noise. Rather than relying on ideal algorithms such as Shor's algorithm, which require an impractically large number of coherent operations, research has focused on hybrid quantum-classical approaches in which quantum devices are used to generate quantum states and classical computers are employed for optimization. This approach is called variational quantum algorithms (VQAs) [28]. Representative examples include the variational quantum eigensolver (VQE) [29], which targets applications in materials science and drug discovery, and the quantum approximate optimization algorithm (QAOA) [30], which aims to support complex decision-making problems in logistics and finance. These methods have opened pathways toward real-world applications in quantum chemistry and combinatorial optimization. From the present perspective, NISQ algorithms are best regarded as quantum heuristics, as they do not offer provable computational advantages. Nevertheless, when combined with quantum error mitigation techniques that statistically suppress the effects of noise, they have reached a stage where partial practical advantages can be demonstrated. As such, the NISQ paradigm plays a crucial role in bridging the gap between current noisy devices and the future realization of FTQC.

1.3 Beyond NISQ Era

At present, quantum computing is entering a transitional phase from the NISQ regime toward FTQC. Recent studies have successfully encoded a few logical qubits into more than one hundred physical qubits on platforms such as superconducting circuits [17] and neutral atom systems [31, 32]. Despite this progress, several major challenges must still be overcome before realizing a clear computational advantage with FTQC. In particular, the implementation of universal quantum computation requires T -gate operations, which cannot be directly realized and therefore rely on methods based on magic state distillation [33]. These approaches require a very large number of physical qubits. For example, the number of physical qubits required to factor standard public-key cryptographic schemes using fault-tolerant quantum algorithms is estimated to reach several million [34]. Given that current hardware is limited to devices with at most a few hundred qubits, further breakthroughs in large-scale integration remain essential for the realization of FTQC [35, 36].

At the same time, VQAs, which attracted significant attention during the NISQ era, face serious technical challenges. One of the most critical issues is the barren plateaus phenomenon [37], in which the gradients required for optimization decay exponentially as the number of qubits increases. This phenomenon has been shown both theoretically and experimentally to make parameter optimization using classical computers increasingly difficult for large-scale systems. In parallel, classical computing has continued to advance through the development of fast quantum simulation techniques based on graphics processing units and increasingly sophisticated heuristic algorithms [38]. As a result, demonstrating a clear quantum advantage using straightforward variational approaches has proven to be more challenging than initially anticipated.

Against this backdrop, new classes of NISQ algorithms that focus on sampling rather than optimization have been proposed [39, 40]. These approaches aim to extract useful information

directly from complex probability distributions generated by quantum devices, instead of relying primarily on iterative optimization loops. In this framework, noisy quantum hardware is used to tackle sampling tasks that are difficult for classical computation by drawing samples from quantum states prepared with suitably designed quantum circuits. One representative example is the quantum-selected configuration interaction (QSCI) method [41]. In this approach, approximate ground states generated by variational quantum circuits are sampled to construct effective Hamiltonians for configuration interaction calculations in quantum chemistry, enabling more accurate energy estimates than conventional VQAs [42]. Another important example is hybrid quantum-classical Markov chain Monte Carlo (MCMC) algorithms [43] for sampling from Boltzmann distributions, which play a central role in condensed matter physics. In these methods, samples generated by quantum circuits are used as proposal distributions to drive state transitions within the MCMC framework. These sampling-based approaches not only offer promising pathways for quantum computers to contribute to practical problem solving prior to the realization of full fault tolerance, but may also remain effective as quantum heuristics even in the FTQC era.

1.4 Overview

This thesis explores the potential advantages of quantum computation for practically relevant problem settings in both the NISQ era and the post-NISQ era by investigating a new class of sampling-based NISQ algorithms as an alternative to VQAs. To this end, we focus on hybrid quantum-classical MCMC as a novel approach for exploring the possibility of quantum speedup on NISQ devices. Hybrid quantum-classical MCMC is a framework in which the output distribution of a quantum circuit is used as a proposal distribution within Markov chain Monte Carlo, one of the most widely used probabilistic sampling methods. Unlike VQAs, this approach does not require variational optimization or repeated measurements of expectation values of objective functions. As a result, it can substantially reduce the number of quantum circuit executions required. Previous work [43] has primarily considered sampling from Boltzmann distributions of Ising models, which play a central role in combinatorial optimization and machine learning, and has experimentally demonstrated that the proposed methods can outperform existing classical MCMC techniques in certain settings. At the same time, several important challenges remain, including the lack of a systematic method for setting quantum circuit parameters [44,45] and the constraints imposed by circuit architectures that are implementable on NISQ devices. Addressing these challenges is therefore an essential research problem for realizing practical quantum speedup in the NISQ era.

The thesis consists of three studies. Two of them aim to address limitations identified in prior work, while the remaining study explores use cases enabled by the improved methods proposed in this thesis.

The first study addresses the challenge of systematically designing quantum circuits that can accelerate convergence in hybrid quantum-classical MCMC. To this end, we propose Quantum Alternating Operator Ansatz Monte Carlo (QAOA-MC), which employs a fixed-depth circuit inspired by QAOA as the proposal distribution in MCMC, thereby suppressing circuit depth growth irrespective of parameter choices and maintaining compatibility with NISQ devices. We establish a concrete optimization strategy by analyzing the relationship between the spectral gap

and the acceptance rate of the proposal distribution, and show that, after appropriately restricting parameter ranges and reducing the number of circuit parameters, minimizing the acceptance rate can lead to faster mixing. Numerical experiments on Boltzmann sampling of spin glass models demonstrate that QAOA-MC achieves a near-quadratic speedup in convergence compared with uniform proposals, and stable estimation of magnetization is confirmed for a 15-spin system.

The second study aims to overcome the structural limitations of existing hybrid quantum-classical MCMC methods that rely on symmetry constraints in quantum circuits to avoid explicit evaluation of proposal probabilities. We introduce a hybrid framework that combines quantum circuit sampling with a generative neural sampler, which serves as a classical surrogate model that approximates the output distribution of a quantum circuit in a tractable form. This approach removes the need for restrictive circuit structures and enables the use of arbitrary quantum circuits, including shallow QAOA circuits, as proposal distributions. Applied to Boltzmann sampling of fully connected spin glass models, the method yields a spectral gap that remains superior to that of classical proposals as the system size increases, with improvements of up to two orders of magnitude at low temperatures. Furthermore, we show that the method enables unbiased magnetization estimation in regimes where single spin-flip updates fail, and that using well-chosen fixed circuit parameters eliminates the need for costly variational optimization. This study demonstrates that integrating neural density estimation significantly enhances the flexibility and practicality of hybrid quantum-classical MCMC on NISQ devices.

The third study investigates the fair-sampling capability of hybrid quantum-classical MCMC, which is essential for applications where multiple solutions exist and uniform sampling over the solution space is required. We first analyze small Ising models and show that introducing MCMC steps on top of quantum dynamics can restore fair sampling that is otherwise lost in purely quantum evolution. We then consider random k -SAT problems as a more practically relevant setting. For 2-SAT instances, combining quantum-generated proposals with standard single spin-flip updates achieves a level of fairness comparable to that of parallel tempering with isoenergetic cluster moves. For 3-SAT, where such advanced classical methods are no longer applicable, the hybrid quantum-classical MCMC still attains fair sampling. In addition, for solution counting, the required number of transitions is comparable to that of WalkSAT, a representative classical heuristic solver. These results show that hybrid quantum-classical MCMC can overcome a long-standing weakness of quantum heuristics and broaden their applicability to solution-sampling tasks.

Through these studies, we seek to advance a new class of quantum algorithms applicable to the NISQ and post-NISQ eras, and to clarify how quantum computation can contribute to solving real-world problems.

1.5 Outline

An outline of each chapter is organized as follows.

Chapter 2: Preliminaries

This chapter introduces several key topics that form the background of this thesis. We first describe the Ising model and the Boltzmann distribution, which define the problem setting of this work, as well as Markov chain Monte Carlo (MCMC) as a representative sampling technique. We then review relevant NISQ algorithms, with particular emphasis on variational quantum algorithms and the quantum approximate optimization algorithm (QAOA). Finally, we introduce the hybrid quantum-classical MCMC algorithm, which constitutes the central topic of this thesis, and review its advantages and remaining challenges.

Chapter 3: Optimization of Quantum Proposals Inspired by QAOA

This chapter presents the first study on improving hybrid quantum-classical MCMC. We propose a method for optimizing quantum circuits in order to accelerate the convergence of MCMC. In existing studies, the relationship between quantum circuit parameters and MCMC convergence had not been clearly understood. By introducing a new objective function that directly reflects MCMC convergence properties, we demonstrate an effective acceleration of hybrid quantum-classical MCMC. This chapter is based on joint work conducted with H. Hakoshima, K. Mitarai, and K. Fujii [[Achievements 1](#)].

Chapter 4: Neural-Network-Assisted Monte Carlo Sampling Trained by QAOA

This chapter presents the second study on improving hybrid quantum-classical MCMC. We propose a new hybrid framework that combines quantum circuits with neural-network-based distribution estimation. To address the limitations imposed by the restricted set of quantum circuit architectures available on NISQ devices, we introduce a classical neural network that approximates the output distribution of a quantum circuit. This approach alleviates circuit-related constraints while preserving the benefits of quantum-generated proposals. This chapter is based on joint work conducted with K. N. Okada and K. Fujii [[Achievements 2](#)].

Chapter 5: Fair Sampling of Ground-State Configurations

This chapter explores use cases for hybrid quantum-classical MCMC. In particular, we analyze the fair sampling properties of hybrid quantum-classical MCMC and investigate their applications. Many important combinatorial optimization problems admit multiple solutions, and for such problems it is often desirable to sample solutions uniformly, a property commonly referred to as fair sampling. Through the analysis of toy models, we demonstrate that hybrid quantum-classical MCMC can achieve fair sampling. Furthermore, using the practically important random k -SAT problem as a benchmark, we apply the proposed methods and show that they achieve fair sampling and convergence efficiency comparable to state-of-the-art classical MCMC approaches. This chapter is based on joint work conducted with K. Fujii [[Achievements 3](#)].

Chapter 6: Conclusion

This chapter summarizes the contributions of this thesis and discusses possible future directions. The results presented in this work indicate that quantum algorithms available on NISQ devices may exhibit advantages for practically relevant problems under appropriate conditions. These findings represent a meaningful step toward the societal application of near-term quantum computers.

Chapter 2

Preliminaries

In this chapter, we provide background knowledge relevant to the research presented in this thesis. First, in Section 2.2, we introduce the Ising model and its Boltzmann distribution, which constitute the core problem setting throughout this thesis. While this model and its distribution have a wide range of important applications beyond statistical physics, their computational cost is extremely large, and approximate calculations are often performed via distribution sampling. In Section 2.3, we describe Markov chain Monte Carlo (MCMC) algorithms, which are widely used for sampling from the Boltzmann distribution. We review the fundamentals of MCMC as well as advanced techniques and associated challenges. Next, after introducing the basics of quantum computation in Section 2.4, we present NISQ algorithms developed for problems related to the Ising model. In Section 2.5, we review variational quantum algorithms, focusing in particular on QAOA for combinatorial optimization problems that can be mapped to the Ising model. Finally, in Section 2.6, we review hybrid quantum-classical MCMC, which constitutes the central topic of this thesis and is applicable to sampling from the Boltzmann distribution of the Ising model.

2.1 Notations

This section summarizes the notations used throughout this thesis. Unbolded x represents a single number, boldface $\mathbf{x}$ represents a vector, and the capital letter A represents a matrix. $\mathbf{x}^{(t)}$ indicates the state at time t represented by the vector $\mathbf{x}$. Other notations are shown in Table 2.1. Detailed definitions are provided in the following sections.

2.2 Model

2.2.1 Ising Model

The Ising model is one of the simplest and most fundamental models in statistical mechanics for describing phase-transition phenomena [46, 47]. The model consists of a collection of n -spin variables s_i placed on lattice sites i , denoted by $\mathbf{s} = \{s_1, \dots, s_i, \dots, s_n\}$. Each spin takes one of two discrete values, $s_i \in \{+1, -1\}$. The Hamiltonian H_P , which determines the energy of the

Table 2.1: Notations used in this thesis.

Notation	Description
$\mathbb{R}$	The set of real numbers.
$\mathbb{C}$	The set of complex numbers.
$\mathbf{x}^{\otimes n}$	The n -fold tensor product.
$ \psi\rangle$	A state vector. Also known as a ket.
$\langle\psi $	A vector dual to $ \psi\rangle$. Also known as a bra.
A^*	The complex conjugate of the A matrix.
$A^\top$	The transpose of the A matrix.
$A^\dagger$	The Hermitian conjugate or adjoint of the A matrix, $A^\dagger = (A^\top)^*$.
$\ A\ _F$	The Frobenius norm of the A matrix, $\ A\ _F = \{\text{tr}(A^\dagger A)\}^{1/2}$.
$\mathcal{N}(\mu, \sigma^2)$	A normal distribution with mean μ and standard deviation σ .
$\text{Uniform}(a, b)$	A uniform distribution on the interval $[a, b]$.

system, is generally written as

$$H_P(\mathbf{s}) = - \sum_{q=1}^k \sum_{1 \leq i_1 < i_2 < \dots < i_q \leq n} J_{i_1 i_2 \dots i_q}^{(q)} s_{i_1} s_{i_2} \dots s_{i_q}. \quad (2.2.1)$$

Here, $J_{i_1 \dots i_q}^{(q)}$ is the interaction coefficient among the q spins $s_{i_1}, \dots, s_{i_q}$. In many cases, one focuses on the case $k = 2$. In this case, the Hamiltonian is given by

$$H_P(\mathbf{s}) = - \sum_{\langle i, j \rangle} J_{ij} s_i s_j - \sum_i h_i s_i, \quad (2.2.2)$$

where J_{ij} is the interaction coefficient between neighboring spins $\langle i, j \rangle$. When $J_{ij} > 0$, the system is called ferromagnetic (the energy is lowered when spins align), whereas when $J_{ij} < 0$, it is called antiferromagnetic (the energy is lowered when spins alternate). The parameter h_i denotes an externally applied magnetic field. The Ising model has long played an important role in the theory and analysis of phase transitions. In particular, the existence of exact solutions in low-dimensional systems has been indispensable for validating theoretical frameworks. In the one-dimensional Ising model, no phase transition occurs at finite temperature because the effect of entropy outweighs the interactions. In contrast, in systems of two or more dimensions, spontaneous symmetry breaking occurs below a critical temperature, and long-range order is formed in which spins align in a single direction. The divergent behavior of physical quantities near the critical point (critical exponents) depends only on the dimension and symmetry of the system, and not on the microscopic details of the interactions. In addition, spin-glass models with random interactions $J_{ij} \sim \mathcal{N}(0, 1)$ were proposed to explain magnetic ordering in alloys containing magnetic impurities in paramagnetic metals [48–50]. Such models have complex energy landscapes with many local minima and constitute an important class of models for understanding systems with random interactions.

2.2.2 Boltzmann Distribution

The properties of a system at thermal equilibrium are described by the Boltzmann distribution (also known as the Gibbs distribution). This distribution characterizes the probability that a system in thermodynamic equilibrium occupies a particular energy state. Let T be the temperature of the system. The probability $\mu(s)$ that the system occurs in that state is given by

$$\mu(s) = \frac{e^{-\beta H(s)}}{Z(\beta)}. \quad (2.2.3)$$

Here, $\beta = 1/k_B T$ is the inverse temperature (with k_B being the Boltzmann constant). $Z(\beta) = \sum_s e^{-\beta H(s)}$ is called the partition function, and it is the constant required to normalize the total probability over all states to unity. The partition function is not merely a normalization constant; it is also a key function for deriving thermodynamic quantities such as the free energy and entropy. However, when the state space S grows exponentially, an exact computation of Z becomes computationally extremely difficult.

2.2.3 Applications of Ising Model

The Ising model and its Boltzmann distribution extend far beyond their original role in the study of magnetism and now serve as a standard prototype in the analysis of complex systems.

The spin glass model is important not only in statistical physics but also from the perspectives of information processing and machine learning. The Hopfield network [51], known as an associative memory model, treats binary neurons as spin variables, so that the stable states of the network correspond to low-energy states of an Ising model. Learning is equivalent to designing the couplings such that desired memory patterns become local minima of the energy, and the ability to recover correct patterns from noisy inputs can be interpreted as associative memory. On the other hand, limitations in memory capacity and the proliferation of spurious local minima give rise to difficulties similar to those encountered in spin glass systems, and are closely related to the problem of local optima in machine learning. Furthermore, by introducing stochastic updates, the framework can be extended to sampling based on the Boltzmann distribution, leading to energy-based models such as Boltzmann machines [52].

One of the most notable applications of the Ising model is to the solution of combinatorial optimization problems. Many difficult real-world problems, such as the traveling salesman problem, scheduling, and portfolio optimization, can be reformulated as the problem of searching for the ground state of the Ising model (a spin configuration that minimizes the energy). This is obtained by formulating the objective function of a combinatorial optimization problem in the form of a quadratic unconstrained binary optimization (QUBO) problem using binary variables $x \in \{0, 1\}$ [53, 54]. It can then be converted into an Ising Hamiltonian through the variable transformation: $s_i = 1 - 2x_i$. On the other hand, finding the ground state of a general Ising model is classified as NP-hard in computational complexity theory [55], and it is therefore necessary to develop efficient search algorithms.

There are two approaches to searching for the ground state of the Ising model. One is search based on energy minimization. A classical method in this category is simulated annealing (SA) [56], which aims to reach the ground state by gradually lowering the temperature while

exploiting thermal fluctuations. The other approach is sampling from the Boltzmann distribution at zero temperature (or at very low temperature). As the temperature approaches absolute zero ($\beta \rightarrow \infty$), the Boltzmann distribution concentrates on energy-minimizing configurations:

$$\lim_{\beta \rightarrow \infty} \mu(s) = \begin{cases} 1/g & (\text{if } H(s) = E_{\min}) \\ 0 & (\text{otherwise}) \end{cases}, \quad (2.2.4)$$

where g is the degeneracy of the ground state. Therefore, sampling from the Boltzmann distribution under an appropriately low-temperature parameter can be regarded as equivalent to the ground-state search problem for the Ising model. This is similar to SA in that the temperature is gradually decreased in the Boltzmann distribution; however, here one directly samples from the low-temperature Boltzmann distribution using a sampling algorithm and discovers the ground state [57].

2.3 Markov Chain Monte Carlo (MCMC)

Markov chain Monte Carlo (MCMC) algorithms [58, 59] provide a powerful framework that is widely used for sampling from complex probability distributions, including the Boltzmann distribution. The central idea is to construct a Markov chain whose stationary distribution is the desired distribution (the target distribution). In this section, we summarize Markov chains as the theoretical foundation of MCMC and explain the Metropolis–Hastings algorithm as a general procedure for constructing a Markov chain with a given target distribution as its stationary distribution. We also discuss proposal distributions, which are an important factor governing the performance of MCMC, and review associated challenges and several practical techniques. Finally, we introduce several metrics for evaluating the performance of MCMC.

2.3.1 Markov Chain

A Markov chain is a stochastic process with the property that the probability distribution of the next state depends only on the current state and does not depend on the previous history. This property is called the Markov property. In what follows, we focus on Markov chains on discrete state spaces. A sequence of random variables $\{\mathbf{X}^{(t)}\}_{t=0,1,2,\dots}$ on a discrete state space S is said to be a Markov chain if, for any states $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{t+1} \in S$, the following holds:

$$P(\mathbf{X}^{(t+1)} = \mathbf{x}_{t+1} \mid \mathbf{X}^{(t)} = \mathbf{x}_t, \mathbf{X}^{(t-1)} = \mathbf{x}_{t-1}, \dots, \mathbf{X}^{(0)} = \mathbf{x}_0) = P(\mathbf{X}^{(t+1)} = \mathbf{x}_{t+1} \mid \mathbf{X}^{(t)} = \mathbf{x}_t). \quad (2.3.1)$$

The matrix P with entries $P(\mathbf{y}|\mathbf{x}) = P(\mathbf{X}^{(t+1)} = \mathbf{y} \mid \mathbf{X}^{(t)} = \mathbf{x})$, which are transition probabilities that do not depend on time t , is called the transition probability matrix. Its entries satisfy $P(\mathbf{y} \mid \mathbf{x}) \geq 0$ and the row-sum condition $\sum_{\mathbf{y} \in S} P(\mathbf{y} \mid \mathbf{x}) = 1$.

When a Markov chain is used for sampling or optimization, it is important whether the distribution converges, as time evolves, to a particular distribution. In particular, a Markov chain that satisfies the following two properties converges to a unique stationary distribution π independently of the initial state.

Irreducibility For any two states $x, y \in S$, it is possible to move between them in a finite number of transitions.

Aperiodicity There is no restriction that forces the chain to return to the same state only at specific step intervals (periods).

A Markov chain with these properties is called *ergodic*, and its stationary distribution π is defined as the solution of the following equation:

$$\sum_{x \in S} \pi(x)P(y | x) = \pi(y), \quad \forall y \in S. \quad (2.3.2)$$

A stronger sufficient condition that guarantees π to be a stationary distribution is the detailed balance condition. It requires that, for any pair of states (x, y) , the probability flow in one direction equals that in the reverse direction:

$$\pi(x)P(y | x) = \pi(y)P(x | y), \quad \forall x, y \in S. \quad (2.3.3)$$

A Markov chain that satisfies this condition is called *reversible*. Since summing both sides of the above equation over x yields the definition of the stationary distribution, designing the transition probabilities $P(y | x)$ so as to satisfy detailed balance enables sampling from the desired distribution π .

2.3.2 Metropolis–Hastings Algorithm

As discussed above, if one can construct a reversible Markov chain whose stationary distribution is the desired distribution $\pi(x)$, then samples generated from the chain can be regarded as being drawn from $\pi(x)$. The Metropolis–Hastings (MH) algorithm [59] is one of the most general MCMC algorithms for drawing samples from a complex distribution $\pi(x)$ (the target distribution) from which direct sampling is difficult.

In the MH algorithm, one introduces a proposal distribution $Q(y | x)$ to propose a candidate next state y from the current state x . By deciding probabilistically whether to accept the proposed state y , the algorithm enforces the detailed balance condition. The procedure is as follows.

1. **Initialization:** Choose an initial state x_0 arbitrarily.
2. **Proposal:** Given the current state x_t , generate a candidate y from the proposal distribution $Q(y | x_t)$.
3. **Acceptance probability:** Compute the acceptance probability $A(y | x_t)$ defined by

$$A(y | x_t) = \min \left(1, \frac{\pi(y)Q(x_t | y)}{\pi(x_t)Q(y | x_t)} \right). \quad (2.3.4)$$

4. **Update:** Draw a uniform random number u from $[0, 1]$. If $u \leq A(y | x_t)$, accept the proposal and set $x_{t+1} = y$. Otherwise, reject the proposal and set $x_{t+1} = x_t$ (remain at the current state).

In particular, when the proposal distribution satisfies $Q(\mathbf{y} | \mathbf{x}) = Q(\mathbf{x} | \mathbf{y})$, the ratio of proposal probabilities becomes unity, and the acceptance probability $A(\mathbf{y} | \mathbf{x})$ reduces to a simpler form. This is called the Metropolis algorithm [58].

That the MH algorithm (or the Metropolis algorithm) correctly has the target distribution $\pi(\mathbf{x})$ as its stationary distribution follows from the fact that the detailed balance condition holds. The transition probability from $\mathbf{x}$ to $\mathbf{y}$ (for $\mathbf{x} \neq \mathbf{y}$) is given by the product of the proposal probability and the acceptance probability,

$$P(\mathbf{y} | \mathbf{x}) = \begin{cases} Q(\mathbf{y} | \mathbf{x})A(\mathbf{y} | \mathbf{x}) & \text{if } \mathbf{y} \neq \mathbf{x}, \\ 1 - \sum_{\mathbf{x}' \neq \mathbf{x}} Q(\mathbf{x}' | \mathbf{x})A(\mathbf{x}' | \mathbf{x}) & \text{if } \mathbf{y} = \mathbf{x}, \end{cases} \quad (2.3.5)$$

Substituting the definition of the acceptance probability yields

$$\begin{aligned} \pi(\mathbf{x})P(\mathbf{y} | \mathbf{x}) &= \pi(\mathbf{x})Q(\mathbf{y} | \mathbf{x}) \min \left(1, \frac{\pi(\mathbf{y})Q(\mathbf{x} | \mathbf{y})}{\pi(\mathbf{x})Q(\mathbf{y} | \mathbf{x})} \right) \\ &= \min (\pi(\mathbf{x})Q(\mathbf{y} | \mathbf{x}), \pi(\mathbf{y})Q(\mathbf{x} | \mathbf{y})) . \end{aligned} \quad (2.3.6)$$

Since this expression is symmetric in $\mathbf{x}$ and $\mathbf{y}$, it follows Eq. (2.3.3).

This leads to two major advantages of the MH algorithm. First, it does not require computing the normalization constant. Since the acceptance probability depends on the ratio $\pi(\mathbf{y})/\pi(\mathbf{x})$, this algorithm can be applied even when the normalization constant of the distribution (which often involves intractable integrals or sums) is unknown, as long as the relative values of the probability density can be computed. Second, it offers substantial flexibility in algorithm design. The proposal distribution $Q(\mathbf{y} | \mathbf{x})$ can be chosen freely as long as the resulting transition probabilities do not violate ergodicity, which is a very mild condition.

2.3.3 Proposal Distribution

MCMC, namely the MH algorithm, is a highly powerful tool for sampling from probability distributions due to its computational convenience and broad applicability. However, the price of this flexibility is that one must prepare an “appropriate” proposal distribution for the given target distribution. This becomes a serious issue particularly for high-dimensional and multimodal distributions such as the Boltzmann distribution of the Ising model. In what follows, we discuss proposal distributions for the Boltzmann distribution of the Ising model.

Single Spin-Flip (SSF) Update

The simplest proposal distribution for spin systems randomly selects a single spin from the current state and flips it. This is called the single spin-flip (SSF) update, and it has existed since the original studies of MCMC [58]. It is straightforward to show that this proposal is irreducible and aperiodic. Moreover, because the proposal distribution is symmetric, one can use the Metropolis algorithm. This update rule can be applied independently of the model.

However, a serious drawback is the performance degradation near a phase transition and in spin-glass systems. In the Boltzmann distribution, larger β makes low-energy states increasingly

dominant, and the distribution becomes sharply concentrated near the minimum energy. Since SSF induces only local transitions at Hamming distance one, sampling multiple modes of a low-temperature distribution requires transitions that pass through many high-energy states. As a result, proposed moves often involve a large energy increase, characterized by the energy difference $\Delta H = H_P(\mathbf{y}) - H_P(\mathbf{x})$ between the proposed state $\mathbf{y}$ and the current state $\mathbf{x}$. This leads to a very large rejection probability, $A(\mathbf{y} | \mathbf{x}) = \min(1, e^{-\beta\Delta H})$, and the dynamics concentrate on a single mode. As a consequence, the Markov chain generated by SSF effectively loses ergodicity and often suffers from severe slowing down.

Global Update

One approach to address the limitations of SSF is to use a proposal distribution that induces nonlocal transitions. Such methods are typically realized by generating some form of cluster and flipping all variables within the cluster, and this class of proposal distributions is referred to as cluster updates. For example, for the ferromagnetic Ising model, the Swendsen-Wang [60] and Wolff [61] algorithms have been proposed. There are also methods combined with techniques such as parallel tempering [62, 63], and for spin glasses, isoenergetic cluster update (PT-ICM) algorithms [64, 65] have been proposed. While these methods can achieve efficient sampling for specific models, satisfying detailed balance is not self-evident, and careful model-dependent design is required. In recent years, approaches have also been proposed in which neural networks are used to generate proposal distributions adapted to a given model [66].

2.3.4 Convergence

Finally, we describe how to evaluate the performance of MCMC. The reliability of MCMC depends on whether the generated sample sequence has sufficiently converged to the target distribution π and whether the chain explores the state space efficiently (good mixing). In theory, the time required for a Markov chain to be regarded as having become “sufficiently close” to the target distribution π is called the mixing time. First we define the total variation distance between two probability distributions μ and ν :

$$d_{\text{TV}} = \|\mu - \nu\|_{\text{TV}} = \max_{\mathbf{x} \in S} |\mu(\mathbf{x}) - \nu(\mathbf{x})| \quad (2.3.7)$$

Using the total variation distance d_{TV} , the distance between the t -step transition probability $P^{(t)}(\cdot | \mathbf{x})$ starting from the initial state $\mathbf{x}$ and the target distribution π is defined as

$$d(t) = \sup_{\mathbf{x} \in S} \|P^{(t)}(\cdot | \mathbf{x}) - \pi\|_{\text{TV}}. \quad (2.3.8)$$

For a tolerance $\epsilon > 0$, the mixing time $t_{\text{mix}}(\epsilon)$ is defined by

$$t_{\text{mix}}(\epsilon) = \min\{t : d(t) \leq \epsilon\}. \quad (2.3.9)$$

In practical computations, early samples prior to reaching the mixing time do not correctly reflect the target distribution and are therefore discarded as a burn-in period. A chain with a short mixing time (small t_{mix}) is said to be “rapidly mixing,” and this provides an indicator of

algorithmic efficiency. However, evaluation via the mixing time depends on the initial state and requires computation over the entire state space, which is computationally very expensive. For this reason, in this thesis we evaluate the performance of MCMC using the following metrics.

Spectral Gap

As an indicator of convergence speed that is closely related to the mixing time, one can consider the spectral gap of the transition probability matrix [67, 68]. The convergence rate of MCMC is governed by the eigenvalue spectrum of the transition matrix P . In particular, the spectral gap, defined as the difference between the largest and the second-largest eigenvalues, is one of the most important theoretical quantities for quantifying convergence efficiency. For a reversible Markov chain on a finite state space S , the transition matrix P has real eigenvalues λ_i , which can be ordered as

$$1 = \lambda_1 > \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_{|S|} \geq -1. \quad (2.3.10)$$

Here, the largest eigenvalue $\lambda_1 = 1$ corresponds to the stationary distribution π , and if the chain is irreducible and aperiodic, then the absolute values of all other eigenvalues are less than 1. Letting $\lambda_{\max} = \max_{i \geq 2} |\lambda_i|$ denote the eigenvalue with the second-largest absolute value, the (absolute) spectral gap is defined as

$$\delta = 1 - \lambda_{\max}. \quad (2.3.11)$$

The spectral gap is generally inversely proportional to the mixing time through $t_{\text{mix}} = O(\delta^{-1})$. It is known that the mixing time $t_{\text{mix}}(\epsilon)$ is bounded above and below in terms of the spectral gap as

$$(\delta^{-1} - 1) \ln \left(\frac{1}{2\epsilon} \right) \leq t_{\text{mix}}(\epsilon) \leq \delta^{-1} \ln \left(\frac{1}{\epsilon \pi_{\min}} \right), \quad (2.3.12)$$

where $\pi_{\min} = \min_{x \in S} \pi(x)$ [69]. This inequality implies that a larger spectral gap leads to a shorter mixing time and faster convergence. Since the spectral gap does not depend on the initial state and, in the case of the Boltzmann distribution, depends only on the proposal distribution and the temperature, we use the spectral gap to evaluate the performance of proposal distributions in this study. On the other hand, because computing eigenvalues of a $2^n \times 2^n$ transition matrix is required, different metrics are used for large-scale systems.

Autocorrelation Function

A sample sequence $\{\mathbf{X}^{(t)}\}$ obtained by MCMC is not statistically independent because, by definition, each state depends on the immediately preceding state. The autocorrelation function (ACF) quantifies this dependence between samples. The autocorrelation at lag τ , denoted by $c(\tau)$, is defined as

$$c(\tau) = \frac{\text{Cov}(\mathbf{X}^{(t)}, \mathbf{X}^{(t+\tau)})}{\text{Var}(\mathbf{X}^{(t)})}. \quad (2.3.13)$$

Here, Cov denotes the covariance and Var denotes the variance. For a well-converged chain, $c(\tau)$ decays rapidly to zero as τ increases. Conversely, if autocorrelation persists for a long range of lags, it suggests that the chain remains confined to a narrow region of the state space and that exploration is insufficient.

2.4 Foundations of Quantum Computation

In this section, we summarize basic concepts in quantum computation. As fundamental elements that constitute the theoretical framework of quantum computation, we describe the notions of quantum states, quantum gates, and quantum measurement. We define a computational procedure for a specific problem built from these elements as a quantum algorithm. This is a universal model of computation and does not depend on the hardware implementation used to realize it (for example, superconducting circuits or ion traps).

2.4.1 Quantum State

In classical information science, information is represented using a two-state system called a bit, taking one of two distinct states. A qubit extends this concept to the quantum world and is the smallest unit of information in quantum computing. While a classical bit can take only one of the states 0 or 1, a qubit can take a complex linear combination of these states, namely a superposition state. The state of a qubit is represented in a two-dimensional complex Hilbert space. Using the two computational basis states

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad (2.4.1)$$

an arbitrary single-qubit state can be expressed as a superposition of $|0\rangle$ and $|1\rangle$ as

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \quad (2.4.2)$$

where $\alpha, \beta \in \mathbb{C}$ are probability amplitudes and must satisfy $|\alpha|^2 + |\beta|^2 = 1$. A system consisting of n qubits can be represented as the tensor product of the states of the individual qubits. Denoting the state of the i -th qubit by $|\psi\rangle_i$, the state of the entire system is written as

$$|\psi\rangle = |\psi\rangle_1 \otimes |\psi\rangle_2 \otimes \cdots \otimes |\psi\rangle_n. \quad (2.4.3)$$

An n -qubit system lives in a 2^n -dimensional complex Hilbert space and can be expressed as a superposition of 2^n basis vectors:

$$|\psi\rangle = \sum_{i=1}^{2^n} c_i |x_i\rangle, \quad (2.4.4)$$

where $c_i \in \mathbb{C}$ and $\sum_{i=1}^{2^n} |c_i|^2 = 1$. Hereafter, unless otherwise specified, we consider the computational basis of an n -qubit system, labeled by n -bit strings $x_i = (x_1, \dots, x_n)$ with $i \in [0, 2^n - 1]$, and denote the corresponding basis states by $|x_i\rangle$.

2.4.2 Quantum Gate

In classical computation, operations on classical bits are described as a set of gates. Similarly, operations on quantum states are described by quantum gates. Since the time evolution of a closed quantum system follows a unitary transformation, a quantum gate is mathematically defined as a unitary matrix U . This is a complex square matrix that satisfies the unitary condition

$$U^\dagger U = U U^\dagger = I, \quad (2.4.5)$$

and the total probability of a quantum state is preserved under unitary transformations. A quantum gate acting on a single qubit is represented by a unitary matrix $U \in \mathbb{C}^{2 \times 2}$.

Representative single-qubit gates include the following four operators, called the Pauli operators:

$$I \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad X \equiv \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad Y \equiv \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad Z \equiv \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \quad (2.4.6)$$

Another important single-qubit gate is the Hadamard gate:

$$H \equiv \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad (2.4.7)$$

which can generate the superposition state $(|0\rangle + |1\rangle)/\sqrt{2}$ from the computational basis state $|0\rangle$.

Operations on an n -qubit system can likewise be represented by a unitary matrix $U \in \mathbb{C}^{2^n \times 2^n}$. By using multi-qubit gates such as the Controlled-NOT (CNOT) gate, which induces interactions between two qubits, one can generate quantum entanglement. In particular, it is known that any unitary transformation can be approximated to arbitrary accuracy by an appropriate combination of single-qubit gates and CNOT gates, and such a set of gates is called a universal gate set [70]. A computational process consisting of multiple gate operations is called a quantum circuit, by analogy with circuits in classical computation. An example of a quantum circuit is shown in Fig. 2.1.

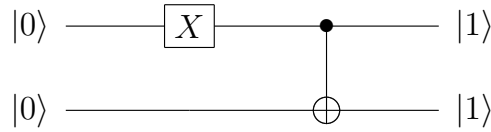


Figure 2.1: An example of quantum circuit diagram.

2.4.3 Quantum Measurement

At the final stage of quantum computation, measurement is performed to extract classical information from a quantum state. Measurement in quantum mechanics is inherently probabilistic,

unlike deterministic classical computation. Measurements in quantum systems and their associated outcome probabilities are given by the Born rule. For a quantum system in the state $|\psi\rangle$, when a measurement of a physical observable A is performed, the measurement outcome is randomly obtained as one of the eigenvalues of A . The probability of observing a particular eigenvalue $a \in \mathbb{R}$ is given by

$$\Pr[A = a \mid |\psi\rangle] = \langle \psi | M_a | \psi \rangle, \quad (2.4.8)$$

where M_a denotes the projector onto the eigenspace corresponding to the eigenvalue a and satisfies $\sum_a M_a^\dagger M_a = I$. Such an operation is called a projective measurement.

The simplest example is the measurement of a quantum state in the computational basis. This corresponds to a measurement defined by the operators $M_i = |x_i\rangle \langle x_i|$. When measuring the state $|\psi\rangle$, the probability of observing the bit string x_i is given by

$$\Pr[x_i] = \langle \psi | M_i | \psi \rangle = |c_i|^2. \quad (2.4.9)$$

Thus, measurement outcomes follow a probability distribution determined by the overlap between the quantum state and the measurement operators. Therefore, the goal in designing a quantum algorithm is to use a quantum circuit to amplify the probability amplitude of the states corresponding to desired solutions and then perform measurement.

Next, we consider the estimation of the expectation value of an observable A . When the observable A is a Hermitian operator with spectral decomposition $A = \sum_m a_m M_m$, The random variable obtained by measuring A takes the eigenvalues a_m with probability distribution $p(m)$. The expectation value of A in a quantum state ρ is therefore expressed as

$$\langle A \rangle = \sum_m a_m p(m) = \sum_m a_m \text{Tr}(M_m \rho) = \text{Tr}(A \rho). \quad (2.4.10)$$

In experiments or quantum circuit executions, the same state ρ is prepared N times and measured repeatedly. Using the eigenvalues $\{a_{m_i}\}_{i=1}^N$ corresponding to the observed outcomes $\{m_i\}_{i=1}^N$, the expectation value is estimated by the sample mean

$$\hat{A} = \frac{1}{N} \sum_{i=1}^N a_{m_i}. \quad (2.4.11)$$

This estimator satisfies $\mathbb{E}[\hat{A}] = \langle A \rangle$ and is therefore an unbiased estimator. Assuming that the measurement outcomes are independent and identically distributed, its variance scales as $\text{Var}(\hat{A}) = \text{Var}(a)/N$, and thus decreases on the order of $1/N$. In particular, when the spectrum of the observable is bounded, the central limit theorem implies that $\hat{A}$ approaches a normal distribution for large N , and the typical statistical error scales as $O(1/\sqrt{N})$. Consequently, in order to estimate the expectation value with a desired precision ϵ , one generally requires on the order of $N = O(1/\epsilon^2)$ measurement shots. In summary, while the Born rule determines the probability distribution of measurement outcomes and expectation values are defined as averages with respect to this distribution, practical estimation is inherently a statistical inference problem based on a finite number of measurements. As a result, there exists an intrinsic trade-off between the number of measurement shots and the achievable estimation accuracy.

2.5 Quantum Approximate Optimization Algorithm (QAOA)

In this section, we provide a comprehensive overview of variational quantum algorithms and introduce the Quantum Approximate Optimization Algorithm (QAOA) as a variant that is closely related to the Ising model. We also present several advanced approaches related to QAOA.

2.5.1 Algorithm

Variational quantum algorithms (VQAs) were developed as an approach for achieving quantum advantage on NISQ devices [24]. The core idea is a hybrid quantum-classical approach that restricts the role of quantum computation to subroutines where an advantage is expected. Specifically, a parameterized quantum circuit is executed on a quantum computer, while the optimization of its parameters is delegated to a classical optimizer. This approach allows the circuit depth to be kept shallow and thereby mitigates the impact of hardware noise.

A VQA begins by defining a cost function C that encodes the solution to the problem. As in classical machine learning, the cost function maps trainable parameters to a real number. In general, the cost function can be written as

$$C(\boldsymbol{\theta}) = f(\{|\psi_i\rangle\}, \{O_i\}, U(\boldsymbol{\theta})), \quad (2.5.1)$$

where f is a function, $U(\boldsymbol{\theta})$ is a parameterized unitary operator, and is also called a variational quantum circuit or an ansatz. The set $\{|\psi_i\rangle\}$ denotes input states, and $\{O_i\}$ denotes a set of observables. The task is formulated as the minimization (or maximization) problem

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} C(\boldsymbol{\theta}), \quad (2.5.2)$$

and the ansatz is trained through a quantum-classical hybrid loop. More specifically, the quantum computer is used to estimate $C(\boldsymbol{\theta})$ (or its gradient), and $\boldsymbol{\theta}$ is trained using a classical optimization method. The algorithm has a loop structure that repeats the following three steps:

1. **Preparation of a quantum state:** Starting from the initial state $|0\rangle^{\otimes n}$, a variational quantum circuit is applied to generate a trial state:

$$|\psi(\boldsymbol{\theta})\rangle = U(\boldsymbol{\theta})|0\rangle^{\otimes n}. \quad (2.5.3)$$

2. **Measurement and estimation of expectation values:** For the trial state $|\psi(\boldsymbol{\theta})\rangle$, the Hamiltonian H is measured. Typically, H is expanded as a linear combination of Pauli operators (X, Y, Z):

$$H = \sum_i c_i P_i \quad (P_i \in \{I, X, Y, Z\}^{\otimes n}). \quad (2.5.4)$$

For each term P_i , repeated measurements are performed, and the expectation value of $C(\boldsymbol{\theta})$ is computed by taking the sample average.

3. **Classical optimization:** The parameters $\boldsymbol{\theta}$ are updated using an optimization algorithm on a classical computer.

The loop is continued until the change in $C(\theta)$ becomes smaller than a threshold or until a predefined number of iterations is reached.

The quantum approximate optimization algorithm (QAOA) is a type of VQA for solving combinatorial optimization problems, proposed by Farhi *et al.* [30]. It aims to find the ground state of a problem mapped to an Ising model, using the energy as the objective function. The QAOA ansatz is inspired by quantum annealing (QA) [71, 72], which is a method for searching for the ground state of the problem Hamiltonian

$$H_P = - \sum_{\langle i,j \rangle} J_{ij} Z_i Z_j - \sum_i h_i Z_i, \quad (2.5.5)$$

by exploiting quantum fluctuations. Z_i denotes the Pauli-Z matrix acting on site i . As a mixer Hamiltonian with a trivial ground state, one prepares

$$H_M = - \sum_i X_i, \quad (2.5.6)$$

and the total Hamiltonian $H(t)$ is slowly varied from H_M to H_P as a function of time:

$$H(t) = A(t)H_M + B(t)H_P, \quad (2.5.7)$$

where the time $t \in [0, T_a]$ and the annealing schedules $A(t)$ and $B(t)$ are monotonic real-valued functions that satisfy $A(0) = 1, B(0) = 0$, and $A(T_a) = 0, B(T_a) = 1$. By the adiabatic theorem, if the system changes sufficiently slowly, a system initialized in the ground state of H_M remains in the instantaneous ground state of the Hamiltonian during the evolution [73, 74]. As a result, the ground state of H_P , namely the optimal solution, is obtained at the end of the evolution. QAOA can be interpreted as a time-discretized approach to the adiabatic evolution in QA. The time-evolution operator in QA is approximated, in a form analogous to a Suzuki–Trotter decomposition, by alternating applications of time evolution under H_P and time evolution under H_M (2.5.6). Concretely, for a quantum circuit with p layers, the state $|\psi(\gamma, \beta)\rangle$ is generated by the following unitary transformation:

$$|\psi(\gamma, \beta)\rangle = U(\gamma, \beta)|+\rangle^{\otimes n}, \quad U(\gamma, \beta) = \prod_{j=1}^p \left(e^{-i\beta_j H_M} e^{-i\gamma_j H_P} \right). \quad (2.5.8)$$

Here, $|+\rangle^{\otimes n}$ is the ground state of H_M (2.5.6), and $\gamma = (\gamma_1, \dots, \gamma_p)$ and $\beta = (\beta_1, \dots, \beta_p)$ are variational parameters. Rather than performing an adiabatic evolution over physical time as in QA, QAOA searches for an approximate solution by tuning the parameters (γ, β) using a classical optimization method so as to minimize the expectation value $\langle \psi | H_P | \psi \rangle$. It is known that, in the limit $p \rightarrow \infty$, QAOA has sufficient expressive power to reproduce adiabatic quantum computation [73].

2.5.2 Challenges

VQAs attracted substantial attention as a leading candidate algorithm for the NISQ era. However, several studies have pointed out that they face serious technical challenges. One of the major

bottlenecks in VQAs is the barren plateau (BP) phenomenon. For certain classes of quantum circuits, the associated cost function has partial derivatives whose magnitudes vanish exponentially with the system size on average [37]. This implies that the landscape becomes essentially flat, so parameter updates based on gradients require an exponentially large measurement cost. In particular, the increase in the number of measurements during the optimization process can eliminate the potential quantum computational advantage of VQAs, making this issue crucial.

Ansatz Variants

One approach to overcoming BPs is to design an appropriate ansatz tailored to the problem at hand. Highly expressive circuits such as a hardware-efficient ansatz [75] must optimize over a vast cost landscape as a trade-off, and thus they are more likely to encounter difficulties such as BPs [76]. Therefore, by using an ansatz specialized for a given problem, one can restrict the search space.

For QAOA, the original ansatz was introduced as a time discretization of QA, as expressed in Eq. (2.5.8). A more generalized framework is known as the Quantum Alternating Operator Ansatz [77]. This framework alternates time evolutions under two Hamiltonians, H_P and H_M . While H_M drives global transitions in the state space, it does not necessarily have to be a transverse-field Hamiltonian (2.5.6), unlike in the original proposal. For example, for constrained problems, there exist variants of H_M designed to induce transitions within a restricted search space [78–81].

Alternatively, restricting the parameter space of the ansatz can also limit the search space. A p -layer QAOA has only $2p$ parameters, which grow only linearly with the circuit depth p , and there are further methods to reduce the number of degrees of freedom. For instance, by assuming a linear ramp as in QA, the parameters can be expressed using only four variables [82–84].

Parameter Initialization

If the ansatz parameters are initialized at random, the algorithm may start from a point far from the solution, near a local minimum, or in a region where BP exists. Therefore, selecting suitable initial parameters can increase the likelihood that the optimization succeeds.

In particular, for QAOA, it is known that for problem instances sharing the same structure, there can exist parameters (so-called “fixed angles”) that commonly yield a good cost function [85–88]. By using such parameters as the initial values for optimization, one can expect a shorter optimization process and a reduced risk of being trapped in local minima.

2.6 Hybrid Quantum-Classical MCMC

In this section, we introduce hybrid quantum-classical MCMC algorithms as a sampling-based NISQ approach for drawing samples from the Boltzmann distribution of the Ising model, which constitutes a central topic of this thesis. In this approach, the output distribution of a quantum circuit is used as the proposal distribution in MCMC. The key idea is to exploit sampling from a quantum circuit without explicitly computing its output distribution, which typically requires

exponential time as the system size increases. In contrast, VQAs, which represent a conventional approach on NISQ devices, extract the output of a quantum circuit as an expectation value of an objective function. To suppress the statistical error of the expectation value to a required level, a large number of measurements and steps are needed, and the impact of noise is substantial on current NISQ devices. By comparison, hybrid quantum-classical MCMC directly uses samples from a quantum circuit as proposed transitions, and therefore can reduce the number of circuit measurements relative to VQAs. This is a highly desirable property on NISQ devices, where circuit execution is costly and hardware noise is significant.

The original study is referred to as quantum-enhanced MCMC (Qe-MCMC) and was proposed by Layden *et al.* [43]. In this method, a quantum circuit U is applied to a spin state s , and a new spin state s' is obtained by measuring the output state $U|s\rangle$ in the computational basis. This sequence of operations defines the proposal distribution:

$$Q(s' | s) = |\langle s' | U | s \rangle|^2. \quad (2.6.1)$$

To avoid computing the output distribution explicitly, a symmetry condition $U = U^\top$ is imposed on the quantum circuit. This requires the proposal distribution to be symmetric, and the acceptance probability reduces to that of the Metropolis algorithm:

$$Q(s' | s) = |\langle x' | U | x \rangle|^2 = |\langle x | U | x' \rangle|^2 = Q(s | s'). \quad (2.6.2)$$

In Qe-MCMC, time evolution under a time-independent quantum Hamiltonian associated with H_P is used to generate proposals. The quantum circuit U that drives the proposal is given by

$$U(u, t) = e^{-iH(u)t}, \quad (2.6.3)$$

$$H(u) = (1 - u)\alpha H_P + uH_M \quad (2.6.4)$$

and a time t , where

$$\alpha = \frac{\|H_M\|_F}{\|H_P\|_F} \quad (2.6.5)$$

is the normalization factor for H_M and $u \in [0, 1]$ is a parameter that controls the relative weights of both H_M and H_P .

The algorithm flow is shown in Algorithm 1. In the original work [43], the tunable parameters (u, t) are not optimized; instead, they are randomly chosen at each proposal step. It was shown that parameters selected from an appropriate range yield sufficient performance. On the other hand, a systematic method for parameter optimization has not been established, leaving room for further improvement [44, 45].

2.7 Summary and Discussion

In this chapter, as background to research on hybrid quantum-classical MCMC schemes, we reviewed Boltzmann distribution sampling of the Ising model and MCMC algorithms, as well

Algorithm 1 Quantum-enhanced MCMC (Qe-MCMC) [43]**Input:** Initial spin configuration s_0 , Target distribution π , MCMC iterations M **Output:** Samples $\{s^{(m)}\}_{m=1}^M$

```

1: function QUANTUMJUMP( $s, U, \pi$ )
2:   Generate  $|\psi\rangle = U|s\rangle$  on quantum device
3:    $s' =$  result of measuring  $|\psi\rangle$  in computational basis
4:    $A \leftarrow \min \left\{ 1, \frac{\pi(s')}{\pi(s)} \right\}$ 
5:   if  $A \geq \text{Uniform}(0, 1)$  then
6:     return  $s'$ 
7:   else
8:     return  $s$ 
9:   end if
10: end function

11: Initialize  $s^{(0)} \leftarrow s_0$ 
12:  $m \leftarrow 0$ 
13: for  $m < M$  do
14:    $u \sim \text{Uniform}(0.25, 0.6)$ 
15:    $t \sim \text{Uniform}(2, 20)$ 
16:    $s^{(m+1)} \leftarrow \text{QUANTUMJUMP}(s^{(m)}, U(u, t), \pi)$ 
17:    $m \leftarrow m + 1$ 
18: end for

```

as related approaches developed in the NISQ era, including QAOA and Qe-MCMC. The Ising model and its Boltzmann distribution sampling constitute an important class of models with broad applications extending beyond the original analysis of magnetism to combinatorial optimization, machine learning, and related fields. However, both ground state search and distribution estimation for such models are computationally extremely demanding tasks. MCMC is a representative approximate algorithm for distribution estimation, in which sampling is achieved by constructing a Markov chain through repeated steps consisting of proposal generation and acceptance or rejection. The proposal distribution is a crucial factor that determines the convergence properties of MCMC, yet it has long been known that convergence deteriorates in complex models such as spin glasses or in systems at low temperatures. In recent years, alongside advances in quantum hardware, numerous attempts have been made to identify quantum advantages for tasks related to the Ising model. A representative example is QAOA for ground state search. QAOA is a type of VQA designed to be executable on NISQ hardware. Nevertheless, it suffers from issues such as circuit optimization complexity and high measurement costs, and to date no clear advantage over existing classical methods has been demonstrated on real hardware. Motivated by this situation, hybrid quantum-classical MCMC schemes have been proposed as approaches better suited to NISQ hardware. In these schemes, the output distribution of a quantum circuit is utilized as a proposal distribution, while the role of quantum computation is restricted solely to sampling.

This design mitigates the disadvantages inherent to variational quantum algorithms and yields features that are more favorable for execution on NISQ hardware.

As a supplementary remark, quantum algorithms for Boltzmann distribution sampling have been studied for many years, and several representative algorithms exist for which computational advantages have been theoretically demonstrated [89–97]. These approaches construct quantum Markov chains for quantum Hamiltonians, and in some cases also for classical Hamiltonians, in order to perform sampling or estimate physical observables. However, all of these methods assume a FTQC architecture as their operating environment. Consequently, they cannot be executed in the NISQ era that is the focus of this thesis.

Based on the above considerations, hybrid quantum-classical MCMC has the potential to become a promising approach for solving academically and industrially important problems using real hardware in the NISQ and post-NISQ eras. However, as mentioned at the end of Section 2.6, the original work (Qe-MCMC) suffers from several challenges. In particular, the two challenges addressed in this thesis are summarized below.

1. Lack of circuit optimization methods: Hybrid quantum-classical MCMC is fundamentally different from VQAs and does not possess a systematic optimization mechanism comparable to that of VQAs. Although this algorithm also requires parameter tuning to accelerate MCMC convergence, an effective method for doing so is nontrivial. In Qe-MCMC, heuristic parameters obtained by pretesting the quantum circuit $U(u, t)$ are employed.

2. Restrictions on quantum circuits: Qe-MCMC introduces symmetric circuits satisfying $U = U^\top$ as a means to avoid explicit computation of the quantum proposal distribution Q . However, this requirement implies that the class of circuits that can be used as proposals is significantly restricted.

These two challenges constitute major obstacles to applying hybrid quantum-classical MCMC schemes to a wide range of practical problems and undermine their potential effectiveness.

In the subsequent chapters, we present solutions to these two challenges. In Chapter 3, we address the first challenge by proposing a new method for optimizing quantum circuits used as proposal distributions. We define an objective function that can be estimated within the MCMC framework and design an algorithm based on a newly parameterized quantum circuit together with an associated optimization loop. In Chapter 4, we address the second challenge by introducing a classical surrogate model that approximates the output distribution of a quantum circuit, instead of directly using the quantum circuit as the proposal distribution. In this study, we focus on Boltzmann distribution sampling of the Ising model, employ QAOA circuits as proposal distributions, and use a neural network for distribution estimation, specifically a masked autoencoder for distribution estimation (MADE), as the classical surrogate model.

Chapter 3

Optimization of Quantum Proposals Inspired by QAOA

This chapter presents our first study on improving hybrid quantum-classical MCMC. We propose a method for optimizing quantum circuits in order to accelerate the convergence of MCMC. In the original work, the relationship between quantum circuit parameters and MCMC convergence had not been clearly understood. By introducing a new objective function that directly reflects MCMC convergence properties, we demonstrate an effective acceleration of hybrid quantum-classical MCMC. This chapter is based on [Y. Nakano, *et al.*, Phys. Rev. Res. 6, 033105 (2024), [Achievement 1](#)] with modifications to fit the context.

3.1 Introduction

The number of qubits in current quantum computers is limited. They hence lack the capacity to implement quantum error correction, rendering them vulnerable to noise. These emerging systems are known as Noisy Intermediate-Scale Quantum (NISQ) devices [15]. These devices have successfully demonstrated the superiority of quantum computers over classical computers in practice: “quantum supremacy” [24, 26]. For practical applications, Variational Quantum Algorithms (VQAs) [28] have emerged as a promising approach to utilize NISQ devices. VQAs run parameterized quantum circuits (known as variational quantum circuits) on NISQ devices and optimize the parameters using an objective function that is expressed by the expectation value of an observable with respect to the output distribution. This approach keeps the quantum circuit depth shallow because the optimization is performed on classical computers. Some algorithms based on the VQA framework have been proposed for quantum chemistry [29], combinatorial optimization [30], and machine learning [98, 99]. Unfortunately, existing VQAs have not yet demonstrated a quantum advantage over the state-of-the-art classical approaches for solving those problems. A possible weakness of these algorithms is their use of expectation values of operators. This requires us to run quantum circuits many times to suppress statistical errors. Furthermore, since optimization using the expectation value is performed iteratively, the total runtime can be prohibitively large [100–102].

In contrast, algorithms that utilize each sampling output from quantum circuits may be more suitable for making use of NISQ devices. For example, Random Circuit Sampling (RCS) used to demonstrate quantum supremacy in NISQ devices is the task of sampling from the output distribution of a random quantum circuit [103] and is shown to be classically hard under a plausible conjecture [22]. Sampling from Instantaneous Quantum Polynomial (IQP) circuits [23] and random linear optical circuits [21] are other famous examples whose classical hardness is strongly believed. These examples motivate us to develop algorithms that fully exploit each sampling outcome from NISQ devices.

The quantum-enhanced Markov chain Monte Carlo (Qe-MCMC) method [43] is one of these algorithms, which uses samples from a quantum circuit as the proposal distribution in the Metropolis–Hastings method [59]. Markov chain Monte Carlo (MCMC) [58, 59] is a powerful technique for sampling from computationally difficult distributions such as the Boltzmann distribution and has many applications in statistical physics [104], combinatorial optimization [56], and machine learning [105]. The Metropolis–Hastings method, one of the MCMC methods, consists of two steps: the generation of a sample by the proposal distribution and the acceptance or rejection step for this sample. Since the proposal distribution determines the efficiency of the algorithm, using the proposal distribution generated by a quantum computer, including distributions that are difficult to simulate on classical computers, can improve the convergence speed of MCMC over existing ones. The Qe-MCMC employs a distribution defined by a classically intractable quantum state as its proposal. The circuit proposed in Ref. [43] is expressed as the time evolution governed by a time-independent Hamiltonian. However, when running it on quantum computers, the time evolution has to be decomposed using the Suzuki–Trotter expansion [106], which increases the circuit depth when the evolution time is long. In addition, the quantum circuits and their parameters are selected heuristically, and the strategy to construct a quantum circuit that improves the convergence speed of MCMC remains unclear.

In this chapter, we propose a new MCMC method called Quantum Alternating Operator Ansatz Monte Carlo (QAOA-MC). This algorithm uses a fixed-depth parameterized quantum circuit in the form of the so-called Quantum Alternating Operator Ansatz (QAOA) [77] as the proposal distribution. We thereby aim to suppress the increase in circuit depth regardless of the choice of parameters. Furthermore, we construct a systematic strategy to optimize the circuit to improve the convergence speed by examining the relationship between the spectral gap and the acceptance rate (AR) of the proposal distribution. More precisely, we find that MCMC can be accelerated by minimizing AR after properly limiting the parameter range and reducing the number of circuit parameters. Through numerical experiments, we evaluate the performance of QAOA-MC for the Boltzmann distribution of a spin glass model, which is one of the most challenging systems to simulate due to its complex energy landscape and slow relaxation dynamics. As a result, we show that QAOA-MC achieves a near-quadratic speed-up in the convergence speed compared to the proposal using the uniform distribution. Additionally, we demonstrate QAOA-MC through the estimation of the average magnetization in a spin glass consisting of 15 spins. These results suggest an acceleration of MCMC using NISQ devices and contribute to promoting the use of current NISQ devices.

The rest of this chapter is organized as follows. Section 3.2 outlines our scheme: QAOA-MC. There, we propose utilizing a parameterized quantum circuit for MCMC proposals and provide

guidance on optimizing the circuit. In Section 3.3, we describe the details of the numerical experiments and their results. Finally, a conclusion and future perspectives are presented in Section 3.4.

3.2 QAOA-MC

We propose to use a parameterized quantum circuit with a fixed depth as the proposal distribution of MCMC, and to optimize this circuit based on the MCMC acceptance rate. We call this method Quantum Alternating Operator Ansatz Monte Carlo (QAOA-MC). An overall view of this algorithm is shown in Fig. 3.1.

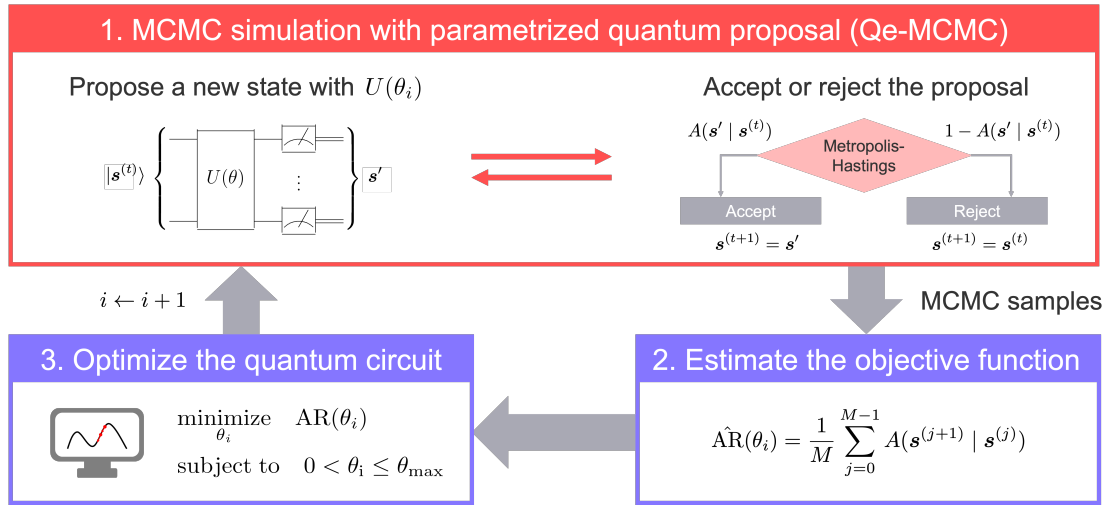


Figure 3.1: Schematic illustration of the QAOA-MC algorithm.

3.2.1 Parametrized Quantum Proposal

We apply the Quantum Alternating Operator Ansatz (QAOA) [77] to the structure of the circuit that generates MCMC proposals. Concretely, this circuit is defined as follows:

$$U = V(\gamma, \beta)^\top V(\gamma, \beta), \quad (3.2.1)$$

where $V(\gamma, \beta)$ is the p -layer QAOA ansatz (2.5.8). It is shown in Fig. 3.2. The quantum circuit U in Eq. (3.2.1) has a similar structure to Eq. (2.6.3). However, unlike the circuit implementation in Eq. (2.6.3), this circuit is more NISQ-friendly because the circuit depth is fixed. The initial state $|s\rangle$ is expected to be updated globally through H_M and H_P is responsible for proposing a transition respecting the energy landscape of the system. The generated probability distribution $Q(s'|s)$ includes those that are classically difficult to simulate and may realize acceleration of the convergence compared to existing proposal distributions.

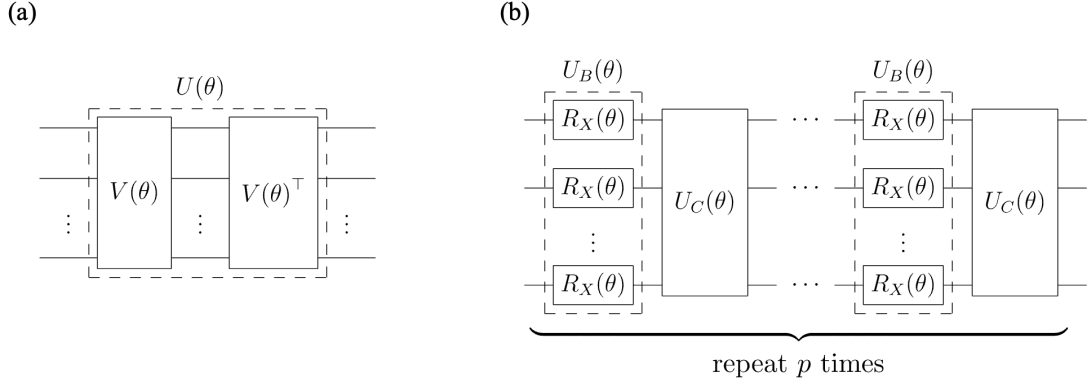


Figure 3.2: Parameterized quantum circuit used in the MCMC proposal. This structure of the circuit is inspired by QAOA, but is designed to satisfy the symmetry constraint that Qe-MCMC requires regarding arbitrary parameters, which is slightly different from the original QAOA.

3.2.2 Optimization of Quantum Proposal

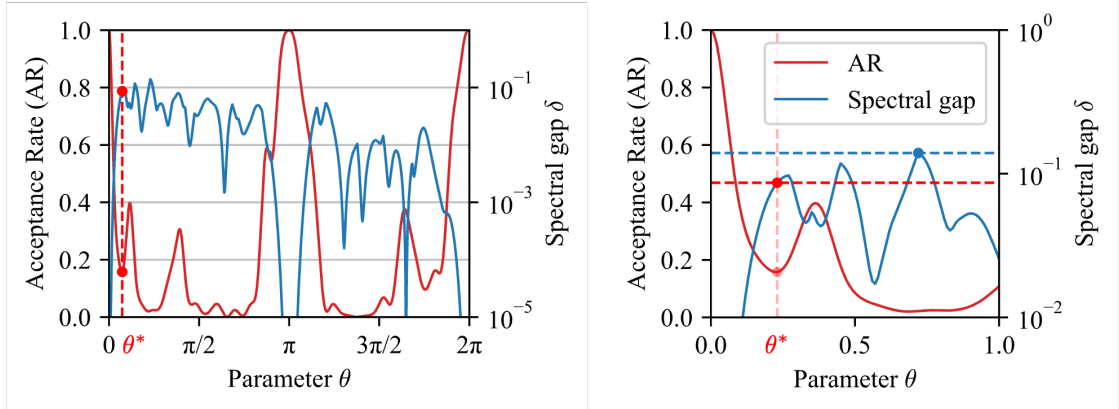


Figure 3.3: The relationship between the circuit parameter θ , AR and the spectral gap δ in a typical instance. The right figure is an enlarged image of the area around the smallest parameter θ^* among those taking AR minima in the left figure. In many cases, θ^* gives a large spectral gap. In these figures, we use a fully connected Ising model (2.2.2) with $n = 5$. Coefficients $\{J_{ij}\}$ and $\{h_i\}$ are randomly generated from the normal distribution $\mathcal{N}(0, 1)$.

Next, we explain how to optimize the proposal distribution generated by the proposed circuit (3.2.1) to achieve faster convergence. One might think that we can use the spectral gap δ (2.3.11) which directly determines the convergence speed as an objective function to maximize. However, computing δ requires solving for the eigenvalues of a transition probability matrix P with a size of $2^n \times 2^n$ for a system of size n and is not feasible. The objective function must be a quantity that reflects the convergence speed of MCMC and is easily computable. We find that the MCMC acceptance rate can be used as the objective function after some numerical

experiments. The acceptance rate (AR) [107] is defined as:

$$\text{AR} = \sum_{s, s' \in S} \pi(s) Q(s'|s) A(s'|s). \quad (3.2.2)$$

This formula includes $\pi(s)$, making it difficult to be calculated directly. However, it can be efficiently estimated by performing MCMC on $\pi(s)$. We estimate AR using M MCMC samples as follows:

$$\hat{\text{AR}} = \frac{1}{M} \sum_{j=0}^{M-1} A(s^{(j+1)} | s^{(j)}), \quad (3.2.3)$$

where $s^{(j)}$ represents the state at the j -th step of the MCMC.

After experimenting with the Boltzmann distribution for various Ising models, we have discovered a relationship between AR and the spectral gap δ . Figure 3.3 illustrates the relationship between δ and AR in a typical Ising model instance. In this experiment, we use a single-parameter circuit $U(\theta)$ which is defined by setting the parameters in Eq. (3.2.1) as

$$\theta = \beta_1 = \dots = \beta_p = \gamma_1 = \dots = \gamma_p.$$

Figure 3.3 shows that although there is usually no correlation between AR and δ , a correlation exists for small θ , where δ increases as AR decreases. It continues until the AR reaches a local minimum, which is often a local maximum value of δ . Based on these observations, we optimize $U(\theta)$ by searching for a small θ that achieves the locally minimal AR.

There are several reasons for AR minimization in QAOA-MC optimization. Firstly, the optimization process initiates near $\theta = 0$. At this point, for the Metropolis-Hastings method, the proposal distribution is maximized with $\text{AR} = 1$, yet it corresponds to the “delta-function” proposal which proposes the same state as before, rendering it the most inefficient proposal [108]. Consequently, in most instances, directing the search toward minimizing AR tends to significantly enhance convergence speed. Given these observations, it is reasonable to optimize the system with the aim of moderately reducing the AR value. The reason why AR local minima near the origin lead to effective global convergence speed remains an open question. While we expect that these considerations will remain valid when multiple parameters are involved, such an optimization becomes increasingly challenging as the dimensionality of the parameter space grows. Therefore, in this study, to maintain simplicity, the single parameterization is employed. It is important to note that our observation is given for the classical Ising model, and its applicability to other models remains still open.

The full procedure of QAOA-MC is summarized in Algorithm 2, and a schematic illustration of the overall process is shown in Fig. 3.1. We perform a search for the local minimum value of AR by restricting the search range to $\theta \in (0, \theta_{\max}]$. $\theta_{\max}$ is set as a hyperparameter in our optimization method.

3.3 Numerical Experiments

In this section, we evaluate the performance of QAOA-MC through the Boltzmann distribution associated with the fully connected spin glass model. This model is defined within the framework

Algorithm 2 QAOA-MC

Input: Initial spin configuration s_0 , Target distribution π , MCMC iterations M, M' **Output:** Samples $\{s^{(m)}\}_{m=1}^M$

```

1: function ARESTIMATOR( $U, \pi, M'$ )
2:   Initialize  $s^{(0)}$ 
3:    $m, \text{AR} \leftarrow 0$ 
4:   for  $m < M'$  do
5:      $s^{(m+1)} \leftarrow \text{QUANTUMJUMP}(s^{(m)}, U, \pi)$ 
6:      $A \leftarrow \min \left\{ 1, \frac{\pi(s^{(m+1)})}{\pi(s^{(m)})} \right\}$ 
7:      $\text{AR} \leftarrow \text{AR} + A/M'$ 
8:      $m \leftarrow m + 1$ 
9:   end for
10:  return AR
11: end function

12: Step 1: Optimization of parameter by minimizing AR
13:  $\theta \leftarrow 0.01$ 
14: while  $\theta$  not converged do
15:    $\hat{\text{AR}}(\theta) \leftarrow \text{ARESTIMATOR}(U(\theta), \pi, M')$ 
16:   Optimize  $\theta$  using  $\hat{\text{AR}}(\theta)$ 
17: end while
18:  $\theta^* \leftarrow \theta$ 

19: Step 2: MCMC simulation
20: Initialize  $s^{(0)} \leftarrow s_0$ 
21:  $m \leftarrow 0$ 
22: for  $m < M$  do
23:    $s^{(m+1)} \leftarrow \text{QUANTUMJUMP}(s^{(m)}, U(\theta^*), \pi)$ 
24:    $m \leftarrow m + 1$ 
25: end for

```

of the Ising model (2.2.2) and is given as follows:

$$H_P(\mathbf{s}) = - \sum_{1 \leq j < k}^n J_{jk} s_j s_k - \sum_{j=1}^n h_j s_j, \quad (3.3.1)$$

in which the coefficients $\{J_{jk}\}$ and $\{h_j\}$ are independently drawn at random from a normal distribution $\mathcal{N}(0, 1)$. In what follows, we present numerical experiments for this model and provide an analysis of the results.

3.3.1 Spectral Gap

First of all, to investigate the performance of QAOA-MC, we analyze the absolute spectral gap δ of the Boltzmann distribution $\mu(\mathbf{s})$ for fully connected spin glass instances of various sizes n . The temperature of the Boltzmann distribution is set to $\beta = 10.0$. We prepare 500 random spin glass instances (3.3.1) and calculate δ for each $\mu(\mathbf{s})$. The average convergence speed $\langle \delta \rangle$ for a model size of n is obtained from these 500 δ values. This is done for each $3 \leq n \leq 10$ to investigate the relationship between n and $\langle \delta \rangle$. In this numerical experiment, we compare the QAOA-MC proposal to three proposal distributions: SSF (local) update, uniform update, and “random circuit”. The random circuit corresponds to a distribution defined by Eq. (3.2.1) with a randomly chosen parameter $\theta \in [0, 2\pi]$ to verify the improvement of convergence speed through optimization. AR is calculated exactly by Eq. (3.2.2).

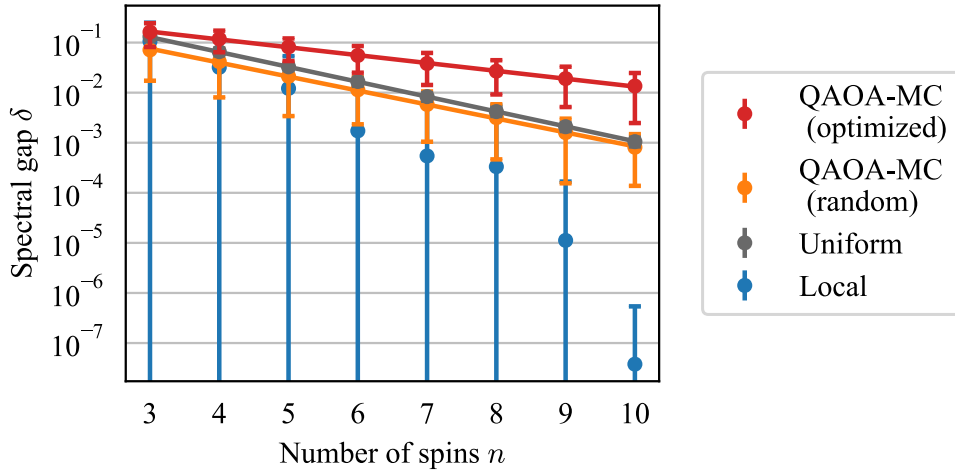


Figure 3.4: Numerical simulation results for the absolute spectral gap δ of the Boltzmann distribution at $\beta = 10.0$ for fully-connected Ising model instances. Relationship between model size n and average convergence rate $\langle \delta \rangle$. “QAOA-MC (random)” uses the randomly chosen parameter $\theta \in [0, 2\pi]$ in the parameterized quantum circuit (Fig. 3.2).

Figure 3.4 shows the relationship between n and $\langle \delta \rangle$ obtained from the numerical experiment. The points represent $\langle \delta \rangle$ computed using 500 random instances at each value of n . The error

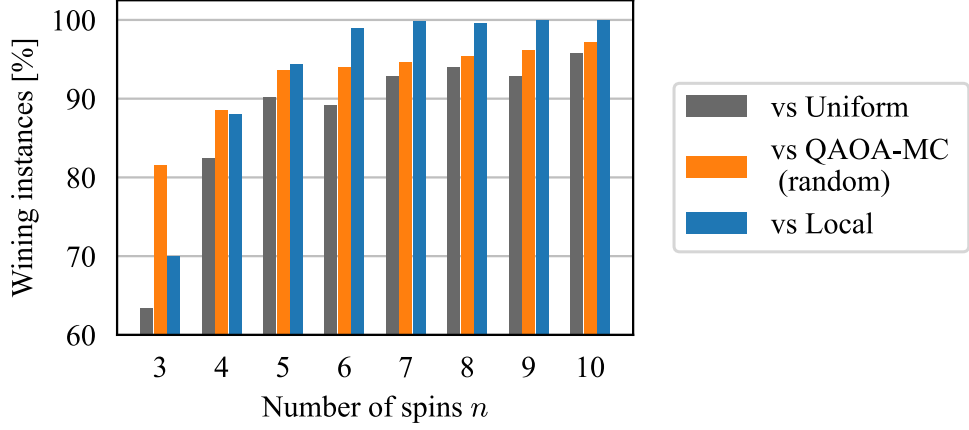


Figure 3.5: Percentage of cases where QAOA-MC surpasses other methods in convergence speed δ .

bars represent the standard deviations computed over 500 δ values. Although $\langle \delta \rangle$ decreases as n increases for all methods, QAOA-MC shows a slower rate of decrease compared to the others and is superior in terms of $\langle \delta \rangle$. We fit $\langle \delta \rangle$ by 2^{-kn} with a parameter k and show the result as straight lines in Fig. 3.4. The approximation curves fit the data well except for the local update. The fitting is calculated using the least squares method. Table 3.1 displays the scaling factor k for these curves. Uncertainties in Table 3.1 are obtained from the covariance matrices obtained in the fitting process. QAOA-MC has a scaling factor k approximately 1/1.89 times that of the

Table 3.1: The value of the scaling factor k obtained by fitting $\langle \delta \rangle$ with 2^{-kn} .

Proposal	k	Ratio to Uniform
QAOA-MC(optimized)	0.521(6)	1.89(2)
QAOA-MC(random)	0.926(9)	1.06(1)
Uniform	0.9869(11)	1

uniform update, which represents an approximately quadratic acceleration in terms of $\langle \delta \rangle$. On the other hand, the results of the “random circuit” are almost identical to those of the uniform update, suggesting that this acceleration is due to the optimization of the circuit.

Although QAOA-MC optimizes a parameter based on the observation that a local minimum of AR often gives a local maximum of δ , it does not always hold for all instances. To see the effect of this imperfect assumption, we show the percentage of instances for which QAOA-MC surpasses the δ of other methods for each size n in Fig. 3.5. The percentage of instances in which QAOA-MC is dominant increases with increasing n . For $n \geq 7$, QAOA-MC outperforms the others in more than 90% of the 500 instances, making it the best-performing proposal in this experiment. This result indicates that our optimization method, which searches for a local minimum of AR for small values of θ , works for many instances.

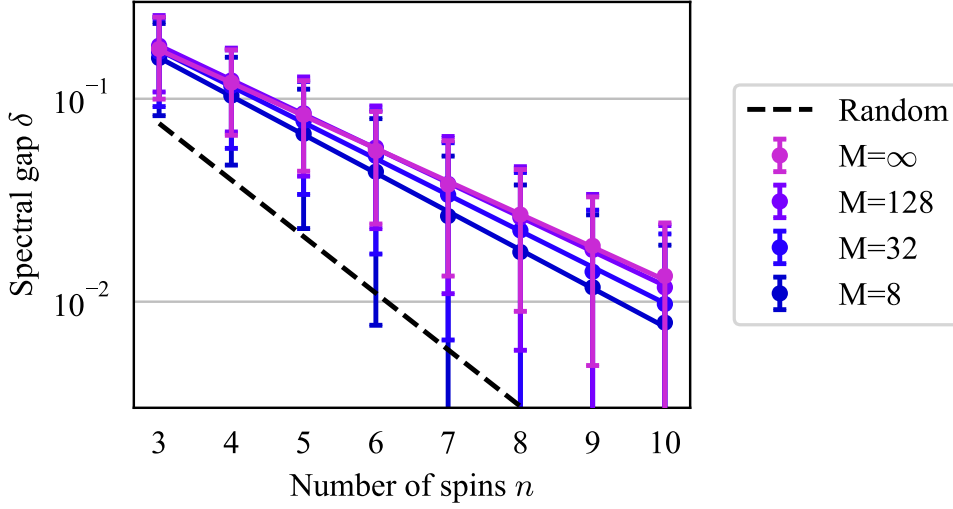


Figure 3.6: Relationship between model size n and average convergence rate $\langle \delta \rangle$ as a function of the number of MCMC samples M used for AR estimation. The dotted line represents the “QAOA-MC (random)” result in Fig. 3.4.

3.3.2 Optimization with MCMC Estimator of AR

We now examine the impact of MCMC estimation of AR on the performance of QAOA-MC. Since this algorithm must use an MCMC estimate to obtain AR in practice, the objective function contains statistical errors that could adversely affect the convergence performance. We analyze the relationship between the number of samples used in AR estimation and the performance. The numerical experiments performed here are under the same setup as discussed in Section 3.3.1 unless otherwise stated. AR is estimated from M samples, which are obtained through MCMC as described by Eq. (3.2.3). We set M to 8, 32, 128, and ∞ (where AR is calculated directly from the target distribution via Eq. (3.2.2)) and optimize θ . Note that in the AR estimation, the MCMC simulations start from a random initial configuration, and no post-processing techniques, such as burn-in, are employed. Then, using the optimized θ , we calculate the absolute spectral gap δ for the same instances used in Section 3.3.1. When using MCMC estimators, the L-BFGS-B method cannot be used as the optimization method because the objective function contains statistical errors. Here, we employ the bisection method for optimization, taking advantage of the fact that we only have a single parameter θ . In the numerical experiments, we used Brent’s method [109], which is implemented by Scipy. The circuit and the hyperparameter settings are the same as in Section 3.3.1.

Figure 3.6 displays the relationship between M and resulting $\langle \delta \rangle$. Table 3.2 shows the scaling factor k for the approximate curves obtained by the same fitting in Fig. 3.4. As M becomes smaller, the standard deviation of $\langle \delta \rangle$ increases and the scaling factor k deteriorates at the same time. This is because decreasing M results in a less accurate AR estimate. This then leads to poor optimization, making the result close to the outcomes of “random”. On the

Table 3.2: The value of the scaling factor k of the approximate curve $\langle \delta \rangle \approx 2^{-kn}$ in QAOA-MC, as M varies.

Number of samples M	k	Ratio to Uniform
∞	0.538(6)	1.834(19)
128	0.5615(26)	1.758(9)
32	0.592(6)	1.668(17)
8	0.629(6)	1.568(15)

other hand, if M is large enough, Brent's method can be used to achieve a performance that is nearly the same as that attained by the L-BFGS-B method. Note that the size of a sufficient M in QAOA-MC is much less than the number of measurements used for a single evaluation of an expectation value in VQAs. Additionally, once the parameters are optimized, only a single-shot measurement from the optimized circuit is required for each step of the MCMC.

3.3.3 Magnetization Estimation

Finally, we conduct numerical experiments with QAOA-MC by estimating a physical quantity, the average magnetization $\langle m \rangle$ of the Ising model. The average magnetization, defined in the context of the Boltzmann distribution $\mu(\mathbf{x})$, is expressed as:

$$\langle m \rangle = \sum_{\mathbf{s}} \mu(\mathbf{s}) m(\mathbf{s}), \quad (3.3.2)$$

where $m(\mathbf{s}) = \frac{1}{n} \sum_{j=1}^n s_j$ denotes the magnetization of a state $\mathbf{s}$. In this numerical experiment, we estimate the average magnetization of an $n = 15$ fully connected spin glass instance with respect to the Boltzmann distribution at $\beta = 1.0$. The Boltzmann probabilities for the ground, first, second, and third excited states (E_0, E_1, E_2 and E_3) are approximately 45.0%, 39.3%, 8.6%, and 4.8% respectively, resulting in a multimodal distribution with large probabilities corresponding to several low-energy configurations. Notably, this instance features several energy minima, each separated by considerable Hamming distances. As a result, MCMC for such an instance faces significant challenges in transitioning frequently between these distant energy minima, a task that proves difficult for traditional classical proposal distributions.

We simulate not only an ideal quantum computer but also examine the effects of noise on the convergence performance of MCMC. Various implementation methods for quantum computers exist, and noise models can vary significantly among them.

For our simulations, for simplicity, we employ one of the most commonly employed noise models, namely depolarizing error, and observe how noise impacts the MCMC convergence performance. For instance, a single-qubit gate is followed by a single-qubit depolarizing error:

$$\mathcal{D}(\rho) = (1 - p_e)\rho + \frac{p_e}{3}(X\rho X + Y\rho Y + Z\rho Z). \quad (3.3.3)$$

Here, p_e denotes the gate error probability, with Pauli operators X, Y , and Z acting on a quantum state ρ with equal probability. The two-qubit gate error is also similarly introduced by using the

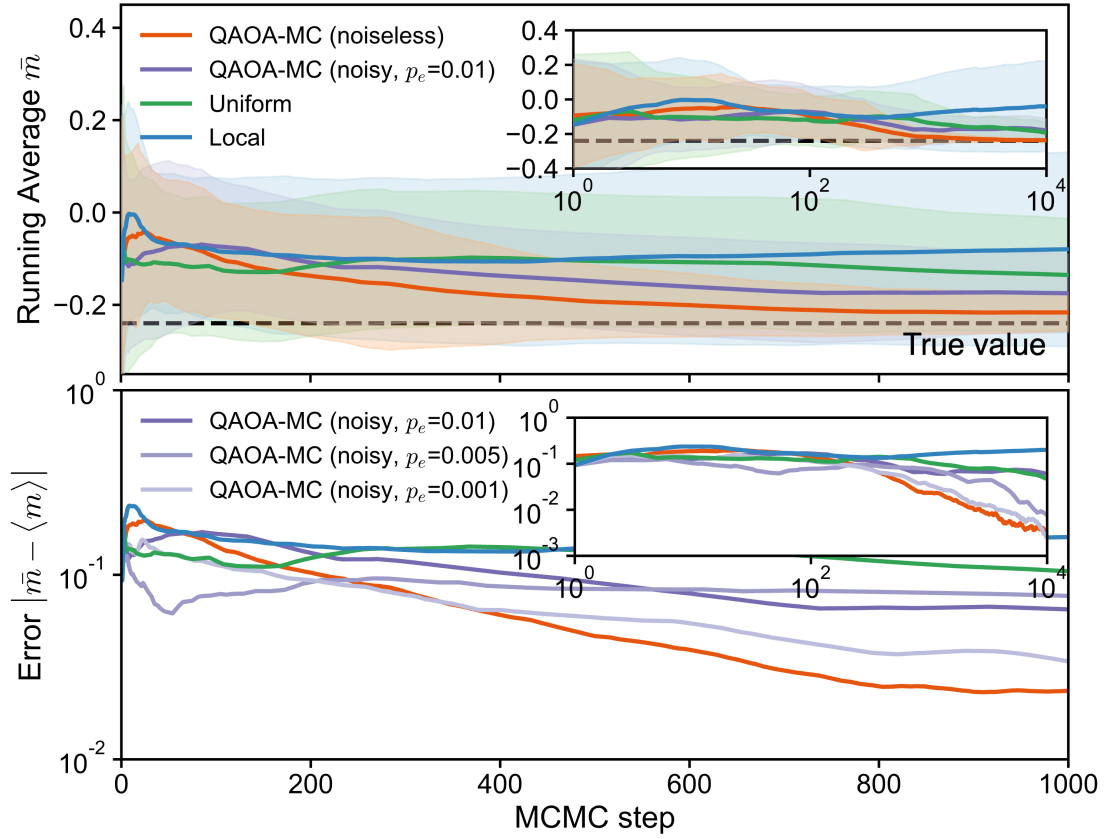


Figure 3.7: Estimation of the average magnetization $\langle m \rangle$ using MCMC. **(Top)** The relationship between the number of MCMC steps and the estimated value $\bar{m}$. **(Bottom)** The absolute difference between the mean of the estimated value $\bar{m}$ and the true value $\langle m \rangle$.

two-qubit depolarizing noise with error probability p_e . While actual devices encounter various other types of noise, this study simplifies the analysis by not delving into additional details. In this experiment, We set the error probability to $p_e = 1.0 \times 10^{-2}, 5.0 \times 10^{-3}, 1.0 \times 10^{-3}$, which is feasible with the current quantum computers.

In the experiment, optimization is conducted as previously described, utilizing AR estimated from MCMC samples. We employ Brent's method for this optimization. The number of samples used for AR estimation is set to $M = 10^3$. All optimization processes are computed through noise-free simulations, while noise simulations are conducted for the MCMC using the optimized parameter θ^* . The MCMC simulations begin from a spin configuration randomly selected according to a uniform distribution, and each of the 10 simulations runs for 10^4 steps, with each simulation starting from a new initial configuration.

Figure 3.7 presents the MCMC estimation results for the average magnetization $\langle m \rangle$. The solid lines depict the average of running averages $\bar{m}$, derived from 10 independent Markov chains at each step. The shaded bands around these lines indicate their standard deviations, illustrating the variability among the chains. The dotted line in Fig. 3.7 represents the true value of $\langle m \rangle$, which has been calculated from the target distribution $\mu(s)$. It is clear from the QAOA-MC results that the running average $\bar{m}$ converges more rapidly to the true value compared to other methods, under both noisy and noise-free conditions. Additionally, the standard deviation of $\bar{m}$ within this algorithm remains low and stable throughout the simulations. In the presence of noise, observations indicate that the convergence speed decreases with the increasing error probability p_e . Simultaneously, it is also observed that when p_e is sufficiently small, the advantage over classical proposals, such as the uniform proposal, still remains. In practical devices, there may be additional factors not accounted for in this experiment that could adversely affect the algorithm's performance. Notably, the presence of biased noise could disrupt the symmetry of the circuit. Nevertheless, this issue can be mitigated through the use of a noise-averaging technique known as Pauli Twirling [43, 110–112]. While this experiment does not consider such factors, and the implementation of twirling could potentially further impact performance, it is anticipated that this method would still substantially improve upon traditional classical proposal distributions.

3.4 Conclusion

In this chapter, we proposed QAOA-MC, which uses samples from quantum circuits in the form of QAOA as the proposal for MCMC. Quantum computation is used only for proposing transitions, and the other parts of the algorithm are executed on classical computers, which makes the algorithm feasible on current NISQ devices. We introduced the use of a QAOA-type circuit to realize the algorithm with shallow circuits. Furthermore, we showed that the convergence speed of MCMC can be improved by finding a local minimum of the AR. As shown in the numerical experiments, QAOA-MC confirmed an approximately quadratic speed-up in the absolute spectral gap for the Boltzmann distribution in a spin glass, when compared to the uniform distribution.

Some future directions are in order. First, the circuit in Fig. 3.2 originally has multiple parameters that could be further tuned to achieve a better proposal distribution. However, the optimization of multiple parameters using AR did not work in our trials. Building more advanced optimization methods remains a possible future work. Additionally, the results of the numerical

experiments in this study are based on the assumption of a quantum computer operating under an ideal or simplified noise model. It is unclear whether the acceleration can be achieved on real NISQ devices.

Chapter 4

Neural-Network-Assisted Monte Carlo Sampling Trained by QAOA

This chapter presents our second study on improving hybrid quantum-classical MCMC. We propose a new hybrid framework that combines quantum circuits with neural-network-based distribution estimation. To address the limitations imposed by the restricted set of quantum circuit architectures available on NISQ devices, we introduce a classical neural network that approximates the output distribution of a quantum circuit. This approach alleviates circuit-related constraints while preserving the benefits of quantum-generated proposals. This chapter is based on [Y. Nakano, K. N. Okada, and K. Fujii, PRX Quantum (2026) (in press), [Achievement 2](#)] with modifications to fit the context.

4.1 Introduction

Quantum computers are widely anticipated as the next generation of computing devices, as they are capable of solving certain problems faster than known algorithms on classical computers [4, 5, 10, 113]. However, current Noisy Intermediate-Scale Quantum (NISQ) devices [15] lack the capability to implement full quantum error correction, which makes it difficult to execute these algorithms with theoretically proven quantum speedups. Still, these devices possess the potential to outperform classical computation, as experimentally demonstrated by the achievement of “quantum computational supremacy” [24–26], and active efforts have been made to leverage this capability for practical tasks. Variational Quantum Algorithms (VQAs) [28] are the leading strategies to obtain a quantum advantage on NISQ devices. These methods heuristically find solutions by optimizing objective functions based on results from parameterized shallow quantum circuits and then updating parameters using classical optimization methods. For example, the Quantum Approximate Optimization Algorithm (QAOA) [30] is one of the most promising VQAs for solving combinatorial optimization problems by searching for the ground state of an Ising model that encodes the problem. While VQAs have achieved significant success in the NISQ era [29, 30, 98, 99], most methods that employ expectation values of Hamiltonian as the objective function remain impractical for real-world applications. This is primarily due to the

large number of measurements required as circuit size and noise increase [101], as well as the existence of barren plateaus in the optimization [37].

The fundamental source of quantum advantage lies in its ability to generate states that are difficult to produce classically, by operating in a Hilbert space that grows exponentially with system size, and to sample classical bitstrings by measuring them. Indeed, random circuit sampling [103], which is used to demonstrate quantum supremacy [22, 24], relies on these facts. On the other hand, if one is only interested in evaluating expectation values, then even for large-scale quantum circuits, classical simulation can be performed efficiently as long as the spread of correlations remains limited [114, 115]. This observation suggests a promising usage of NISQ devices as a sampler to generate certain probability distributions, instead of an estimator of expectation values. With this in mind, a hybrid quantum-classical algorithm known as Quantum-Selected Configuration Interaction has recently been proposed [41] and large-scale demonstrations using quantum hardware have already been conducted [42].

Another approach in this direction is hybrid quantum-classical MCMC [43], which uses a quantum circuit as the proposal distribution in Markov chain Monte Carlo (MCMC) [58, 59]. Sampling from the quantum circuit produces new transitions for the Markov chain, and the accept-reject step is performed classically based on the acceptance probability between transitions. The key idea is to take advantage of quantum circuit sampling without explicitly calculating its output distribution, which typically requires exponential time as system size increases. Existing works implement the idea by selecting a symmetric circuit $U = U^\top$, which can avoid the calculation of the output distribution entirely. However, this complicates the understanding of how the circuit influences the acceptance probability, making it difficult to control the convergence rate [44, 45, 116].

In this chapter, we investigate a hybrid quantum-classical MCMC framework that integrates quantum circuit sampling with a generative neural sampler (GNS) [117–119] to overcome existing limitations in these approaches. Unlike existing methods that impose symmetry constraints on the quantum circuit to avoid calculating output distributions, this approach uses a GNS as a neural surrogate model that approximates the output distribution in a computationally tractable form. This eliminates the need for restrictive circuit structures and enables the use of arbitrary quantum circuits as proposal distributions. Consequently, this study expands the applicability of the hybrid quantum-classical MCMC to a broader class of target distributions while retaining compatibility with NISQ devices. By enabling flexible and efficient quantum sampling, this framework paves the way for practical applications in domains such as statistical physics, machine learning, and combinatorial optimization.

We focus on Boltzmann sampling for the Ising model using a proposal distribution generated by a QAOA circuit, which does not satisfy the symmetry constraint and thus lies outside the scope of existing methods. It has been shown that the output distributions of shallow QAOA circuits can approximate low-temperature Boltzmann distributions that include contributions from several excited states [120–122]. A GNS trained on such QAOA outputs can induce transitions weighted toward low-temperature Boltzmann distributions, where long convergence times are typically a bottleneck, thereby enabling more efficient MCMC simulations. Moreover, shallow QAOA circuits are well suited for implementation on NISQ devices, while sampling from them remains intractable for classical computers [123]. We demonstrate the performance of the

proposed method on the Boltzmann distribution of fully-connected spin glasses. The spectral gap of the proposed method remains superior to that of the classical proposals as the system size increases. The proposed method achieves a polynomial speedup (over cubic) compared with the uniform distribution, particularly at low temperatures ($\beta \geq 5.0$). In magnetization estimation, we show that the method constructs a proposal distribution that approximates the Boltzmann distribution at low temperatures, allowing transitions in complex energy landscapes where the single spin-flip (SSF) update struggles and ensuring unbiased sampling. In particular, the use of well-chosen initial parameters, known as “fixed angles” [85–88], can obviate the need for parameter optimization, allowing the method to be deployed without costly variational procedures. Through numerical experiments, the proposed method with fixed-angle QAOA also surpasses the classical proposals. This feature may facilitate practical applications that leverage the computational power of NISQ devices with a realistic number of circuit executions.

The rest of this chapter is organized as follows. In Section 4.2, we introduce a generative neural sampler (GNS). Section 4.3 provides a hybrid quantum-classical MCMC algorithm integrating QAOA with GNS, which is named QAOA-NMC. Section 4.4 presents numerical experiments demonstrating that the proposed sampler improves convergence compared to existing methods and offers several additional advantages. Finally, Section 4.5 concludes this chapter and discusses future directions.

4.2 Generative Neural Sampler

Neural networks can infer an underlying unknown probability distribution from a given dataset, a task known as distribution estimation that has been studied for decades. Generative neural samplers (GNSs) employ this capability by sampling from the probability distribution learned by these neural networks. One of the earliest examples in this field is the Boltzmann machine [124]. In particular, the development of restricted Boltzmann machines (RBMs) [125] and gradient-based methods such as contrastive divergence [126] has made it feasible to avoid enumerating exponentially large state spaces, thereby reducing computational costs. However, distributions whose normalization constants are hard to compute, such as the Boltzmann distribution, remain challenging for RBMs. More recently, Neural Autoregressive Distribution Estimators (NADE) [117, 118] have proven effective for this task. These models offer greater ease of use and higher computational efficiency compared to conventional approaches.

In this study, we focus on the Masked Autoencoder Distribution Estimator (MADE) [119], which is a variant of NADE. MADE is a distribution estimator that uses an autoencoder. An autoencoder [127] is a feedforward network designed to produce an output vector $\mathbf{y}$ that closely matches the input vector $\mathbf{x}$. The main objective is to capture the underlying structure of the distribution from which the input data are drawn, thereby learning representations that can efficiently generate new samples from the same distribution. An autoencoder generally consists of an input layer, one or more hidden layers, and an output layer. A network with a single hidden layer can be written as

$$\mathbf{h} = g(\mathbf{W}\mathbf{x} + \mathbf{b}) \quad (4.2.1)$$

$$\mathbf{y} = \text{sigm}(\mathbf{V}\mathbf{h} + \mathbf{c}), \quad (4.2.2)$$

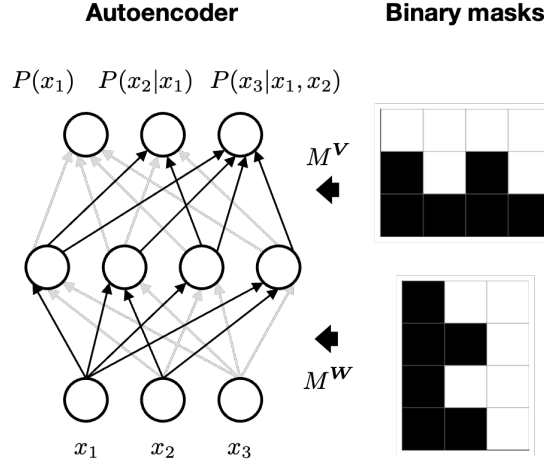


Figure 4.1: An example of a MADE network with a single hidden layer. By applying a binary mask (shown on the right) that restricts certain connections, the model obtains autoregressive properties absent in a standard autoencoder. The gray paths in MADE represent connections that are removed by the mask.

where $\mathbf{W}$ and $\mathbf{V}$ are weight matrices, $\mathbf{b}$ and $\mathbf{c}$ are bias terms, and g denotes the activation function. The sigmoid function $\text{sigm}(a) = 1/(1 + \exp(-a))$ is applied at the output layer.

MADE augments this autoencoder network with an autoregressive property to facilitate efficient distribution estimation. Any joint distribution can be factorized into a product of conditional probabilities:

$$p(\mathbf{x}) = \prod_{d=1}^D p(x_{o_d} | \mathbf{x}_{o_{<d}}) \quad (4.2.3)$$

where o_d specifies the ordering of the variables as a permutation of $\{1, \dots, D\}$. Autoregression imposes dependencies among the output-layer variables, allowing the joint distribution to be computed as the product of the network outputs. To implement this, masks M^W and M^V are introduced to constrain the network connections. Consequently, a single-hidden-layer MADE can be expressed as

$$\mathbf{h} = g \left((M^W \cdot \mathbf{W})\mathbf{x} + \mathbf{b} \right) \quad (4.2.4)$$

$$\mathbf{y} = p(x_{o_d} = 1 | \mathbf{x}_{o_{<d}}) = \text{sigm} \left((M^V \cdot \mathbf{V})\mathbf{h} + \mathbf{c} \right) \quad (4.2.5)$$

Figure 4.1 illustrates a case of a MADE network with a single hidden layer and the masks defined by Eqs. (4.2.4), (4.2.5). The network is trained by minimizing the cross-entropy loss:

$$l(\mathbf{x}) = \sum_{d=1}^D \{-x_d \log y_d - (1 - x_d) \log(1 - y_d)\} \quad (4.2.6)$$

which is equivalent to the negative log-likelihood of $p(\mathbf{x})$. In practice, the loss function can be optimized via stochastic gradient descent, such as Adam [128].

Recently, a strategy that integrates machine learning into MCMC has been proposed in the field of statistical physics, involving global updates through GNS that approximate the target distribution [129–131]. These neural proposals enable model-specific global updates by training a GNS to learn an auxiliary distribution that closely approximates the target. A major challenge of using the GNS as a proposal distribution lies in learning the target distribution. There are two primary strategies for this task: maximum likelihood estimation [130] and learning based on variational energy minimization [129, 132]. In maximum likelihood estimation, the model is trained on samples that closely approximate the target distribution, adjusting parameters to maximize the likelihood of the data. This approach depends heavily on the quality of the dataset, and it faces the difficulty of preparing samples that are sufficiently close to the target distribution when the latter is hard to sample from. Ref. [133] demonstrated that although a GNS trained with high-temperature SSF samples can be gradually refined through “sequential annealing” [130] to approximate the low-temperature Boltzmann distribution, the results reveal that it is not possible to obtain samples of sufficient quality in the low-temperature regime. In contrast, learning based on variational energy minimization does not require training samples. However, this method often suffers from mode collapse, where only a subset of the multiple peaks of the target distribution is learned. While this issue may not be severe for models where only a few dominant modes exist (e.g., a two-dimensional Edwards-Anderson model), it becomes a critical limitation for distributions with highly complex energy landscapes, such as spin glasses [133, 134].

4.3 QAOA-NMC

We present a hybrid quantum-classical MCMC algorithm that integrates GNS with quantum circuits. In this approach, GNS, which approximates the output distribution of the quantum circuit, serves as the proposal distribution for MCMC. Specifically, we propose using QAOA for the Boltzmann sampling of spin glasses.

4.3.1 Algorithm

The algorithm based on QAOA operates as follows. First, QAOA is executed, and samples are drawn from its final state. Although QAOA parameters typically require variational optimization, this computational cost can be significantly reduced or even avoided entirely by using the initial parameters introduced later in this study. The resulting binary samples are then used as training data for GNS, allowing it to learn the QAOA output distribution. Finally, an MCMC simulation is performed using GNS, which approximates the QAOA output distribution, as the proposal distribution. The GNS proposal is prepared and trained separately for each model instance. The full procedure of the proposed method is summarized in Algorithm 3, and a schematic illustration of the overall process is shown in Fig. 4.2.

The transition defined by GNS is independent of the current state, meaning that $Q(s'|s) = p(s')$. If $p(s)$ exactly matches $\pi(s)$, the acceptance probability becomes 1. Therefore, when GNS closely approximates the target distribution, a high acceptance rate can be expected.

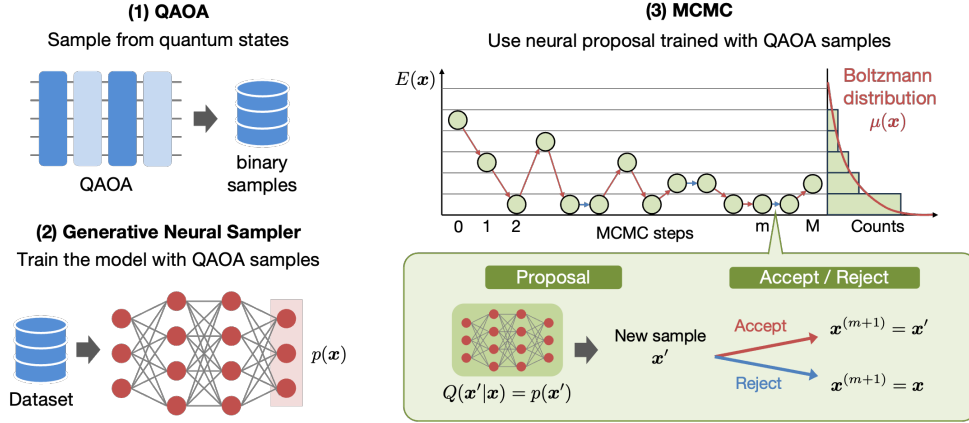


Figure 4.2: A schematic diagram of the proposed method in this study. The method comprises three main components: QAOA, GNS, and MCMC, and proceeds as follows: **(i)** Run QAOA and sample the output states of the quantum circuit. **(ii)** Use the sampled binary sequences as training data to train a generative neural network that reproduces the QAOA output distribution. **(iii)** Perform MCMC using the trained neural network as the proposal distribution.

In this algorithm, MADE is adopted as the GNS. MADE utilizes autoregressive properties to reduce connections, resulting in polynomial complexity of $O(n^2)$ for both evaluating and sampling from the distribution. This structure enables efficient computation. Furthermore, since the transition is independent of the current state, GNS proposals can be generated separately from the MCMC simulation. This property suggests that proposals may be precomputed, potentially reducing the overall computational cost.

4.3.2 Boltzmann Sampling with Neural Proposal Trained by QAOA

The main significance of combining quantum circuits with generative neural samplers is that, within hybrid quantum-classical MCMC, any quantum circuit can serve as a proposal distribution, provided its output distribution satisfies ergodicity. The trade-off of using quantum dynamics for MCMC proposals is the need to access output state probabilities, which are generally difficult to compute. Previous studies avoided computing output distributions by exploiting symmetric dynamics, but this approach restricted the set of applicable quantum circuits. Our method enables both the use of quantum dynamics and the tractable computation of proposal distributions by approximating the distribution generated by quantum dynamics with a surrogate model based on GNS. Earlier work proposed quantum annealing-like Hamiltonian dynamics [43, 135] for sampling from the Boltzmann distribution of the Ising model. In contrast, this study demonstrates that QAOA circuits without symmetry can also be handled. Moreover, this method may extend the benefits of quantum proposals to a broader class of distribution sampling problems beyond Boltzmann distributions. Nonetheless, identifying suitable quantum circuits for other distributions and examining the applicability of this method remain open directions for future research.

Algorithm 3 QAOA-NMC

Input: Initial spin configuration s_0 , Target distribution π , MCMC iterations M , MADE $MADE()$, QAOA circuit $U(\theta)$, Initial parameters θ_0 , Dataset size N .

Output: Samples $\{s^{(m)}\}_{m=1}^M$.

1: **Step 1: Sampling from QAOA circuit**

2: Set θ_0 in $U(\theta)$

3: Dataset $\mathcal{D} \leftarrow$ Sampling $U(\theta)$ in the computational basis

4: **Step 2: Training MADE on QAOA samples**

5: Train $MADE()$ on $\mathcal{D}$

6: **Step 3: MCMC simulation with trained MADE proposal**

7: Initialize $s^{(0)} \leftarrow s_0$

8: $m \leftarrow 0$

9: **for** $m < M$ **do**

10: $s', p(s') \leftarrow MADE()$

11: $A \leftarrow \min \left\{ 1, \frac{\pi(s')}{\pi(s^{(m)})} \frac{p(s^{(m)})}{p(s')} \right\}$

12: **if** $A \geq \text{Uniform}(0, 1)$ **then**

13: $s^{(m+1)} \leftarrow s'$

14: **else**

15: $s^{(m+1)} \leftarrow s^{(m)}$

16: **end if**

17: $m \leftarrow m + 1$

18: **end for**

There are several advantages to using QAOA samples as training data for the neural proposal. The QAOA output distribution is known to approximate a classical Boltzmann distribution at a low temperature, which can be especially beneficial at low temperatures compared to samples generated by traditional MCMC methods such as the SSF update. In practice, however, the optimized QAOA output distribution may deviate from the true Boltzmann distribution of the model. This deviation arises because the optimal QAOA solution corresponds to a zero-temperature thermal state, which may fail to capture peaks associated with certain low-energy excited states that are present in a finite-temperature Boltzmann distribution. Nevertheless, the method often performs well in practical applications. Indeed, studies have shown that under short-time evolution with QAOA-like Hamiltonians, wavefunction collapse upon measurement tends to converge with high probability to one of the lowest-energy configurations [43, 122, 136]. This result suggests that the output distribution of a shallow QAOA circuit is well-suited for proposing low-energy configurations, which remains a central challenge in Boltzmann sampling.

Another important feature of QAOA is the availability of heuristic initial parameters [85–88]. Utilizing such parameters enables rapid sample generation via QAOA without the computational overhead of optimization. This feature aligns well with the capabilities of NISQ devices, where

expectation value estimation can be challenging, but direct sampling remains feasible. As a result, this approach is especially promising for near-term quantum devices.

While this discussion focuses on QAOA implemented on quantum computers, related algorithms using quantum annealers are also possible [137], since QAOA is originally regarded as a discretized form of quantum annealing. Quantum annealers are capable of solving Ising models involving thousands of qubits [138], although hardware constraints limit the range of graph structures that can be directly implemented [139]. In contrast, universal quantum computers may support broader applications beyond the Ising model and may not be subject to such graph limitations, depending on the architecture. Moreover, VQAs are generally resilient to noise, and several noise mitigation techniques have been developed. Through these strategies, the proposed method may achieve quantum speedup on real NISQ devices for useful tasks such as Boltzmann sampling.

4.4 Numerical Experiments

In this section, we evaluate the performance of the proposed method by sampling from the Boltzmann distribution of spin glass systems. This model is a fully connected spin glass model without an external field and is given as follows:

$$H_P(s) = - \sum_{1 \leq j < k}^n J_{jk} s_j s_k, \quad (4.4.1)$$

where the coefficients $\{J_{jk}\}$ are randomly drawn from a normal distribution $\mathcal{N}(0, 1)$. For all experiments, we assume that computations are carried out on an ideal, noise-free quantum computer.

4.4.1 Spectral Gap

We begin by analyzing the average performance of the proposed method for small-scale systems. We calculate the spectral gap of the Markov chains generated by the proposed method and compare these results with those from various alternative proposal distributions.

In this experiment, we compute the spectral gap for 100 randomly generated instances (4.4.1) for each system size n , ranging from 3 to 12 spins, in order to examine how the performance of the proposed method depends on system size. MADE is trained separately for each instance, using the dataset generated by QAOA on that instance. We consider two types of QAOA circuits: one that uses fixed angles derived from the Sherrington-Kirkpatrick model [140] followed by optimization, and one that applies the same angles without any optimization. For comparison, we evaluate the proposed method against a quantum algorithm known as QAOA-MC. We also compare the proposed method with classical proposals, including the SSF update and a uniform proposal distribution. While more advanced techniques such as replica exchange methods are available for spin glasses, this study limits its scope to a range of simpler proposal distributions. Importantly, the proposed method could be extended to include replica exchange or other advanced MCMC algorithms for direct comparisons with state-of-the-art methods. However, such extensions are beyond the scope of this study.

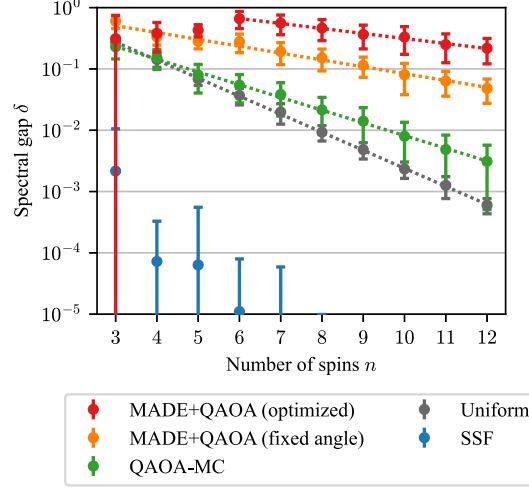


Figure 4.3: System size dependence of the spectral gap at $\beta = 10.0$. The data points represent the average spectral gap $\langle \delta \rangle$ calculated over 100 randomly generated instances for each system size n . The dotted lines indicate the scaling obtained through least-squares fitting.

Table 4.1: The exponential fit parameters in Fig. 4.3. Uncertainties are from the covariance matrices obtained in the fitting process.

Proposal type	Fitting $\langle \delta \rangle = 2^{-kn}$
MADE+QAOA (optimized)	0.272(8)
MADE+QAOA (fixed angle)	0.371(21)
QAOA-MC	0.689(12)
Uniform	0.977(5)
SSF	—
Enhancement (vs Uniform)	3.60(11)

Figure 4.3 shows how the spectral gap varies with n for each proposal distribution at $\beta = 10.0$. The average spectral gap $\langle \delta \rangle$ tends to decrease as n increases. However, in the case of optimized QAOA, the trend can reverse for small n : performance may actually degrade as n decreases. This occurs because, in the small- n regime, QAOA optimization becomes highly effective, concentrating nearly all probability on the optimal solution. The resulting output distribution thus approximates a zero-temperature Boltzmann distribution, which deviates from the desired proposal distribution.

At $\beta = 10.0$, the scaling of the spectral gap with n is well fitted by $\langle \delta \rangle = 2^{-kn}$, except in the case of SSF. For the optimized QAOA results, the fitting is applied only in the region where $\langle \delta \rangle$ decreases with increasing n , for the reasons discussed earlier. Table 4.1 reports the fitted values of k , from which it follows that the proposed method achieves a polynomial speedup (approximately cubic) compared with the classical uniform proposal. In this low-temperature regime, the proposed method consistently outperforms all of the other methods. Classical proposal distributions struggle to navigate the complex energy landscape, especially near local minima. By contrast, the proposed approach attains high convergence efficiency thanks to a neural network sampler trained on samples generated by QAOA, which effectively captures low-energy configurations of the spin glass. Although QAOA-MC also employs QAOA-inspired circuits to facilitate fast transitions between low-energy states, the proposed method still achieves better performance overall. Unlike QAOA-MC, which is constrained by available quantum circuits, this method, using the GNS representation, provides transitions that are better suited for low-temperature Boltzmann sampling through the QAOA output distribution. On the other hand, proposals based on the QAOA output distribution are not well suited for cases where the target distribution assigns appreciable probability to high-energy configurations, such as at high temperatures. In the high-temperature regime ($\beta \leq 1$), the scaling in fact worsens compared with classical SSF and uniform proposals. (See Appendix A for simulations and scaling results at other temperatures.) Nonetheless, it should be emphasized that in these temperature regimes, simple classical proposals are sufficiently efficient, and therefore fall outside the scope of this study.

Notably, the fixed-angle QAOA also yields superior outcomes compared to other methods aside from the optimized QAOA. This finding is particularly significant because the main bottleneck in VQAs generally lies in optimizing the parameterized quantum circuits. The possibility of achieving quantum advantage without extensive optimization could thus offer substantial benefits. Indeed, the hybrid quantum-classical MCMC algorithms typically face optimization and tuning challenges. For instance, quantum-enhanced MCMC methods still require careful circuit parameter adjustments [44]. Such difficulties underscore why the proposed method may represent a more practical quantum solution for current NISQ devices.

4.4.2 Magnetization Estimation

Next, we assess the performance of the proposed method for larger systems by estimating magnetization. Defined as the average spin value in the model, $m(s) = \frac{1}{n} \sum_{j=1}^n s_j$, magnetization quantifies the net polarization of spins. In this experiment, we estimate the average magnetization $\bar{m} = \sum_s \mu(s)m(s)$ from samples generated by MCMC simulations of a 25-site spin glass (4.4.1)

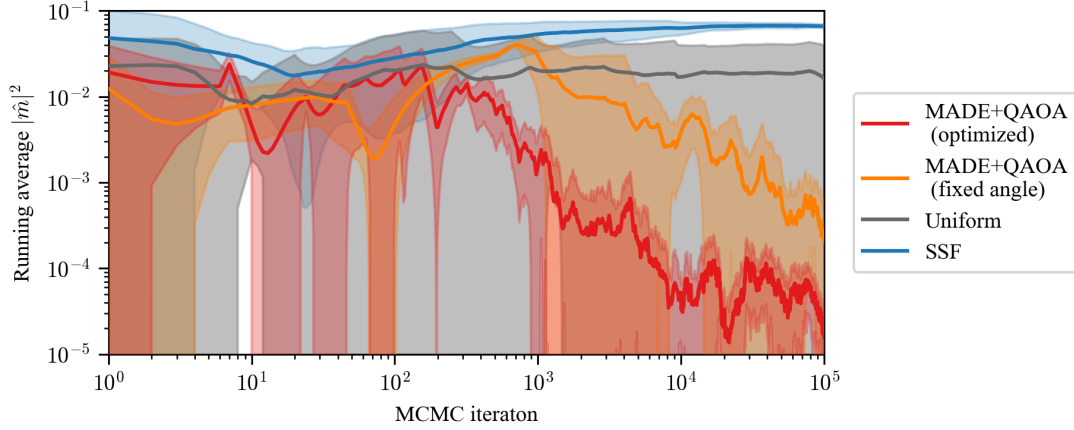


Figure 4.4: MCMC estimation results for the squared estimator of the average magnetization ($\overline{m} = 0$). The solid lines represent the average of running averages, computed from ten independent Markov chains at each step. The shaded bands around the lines indicate the standard deviations, illustrating variability among the chains.

at $\beta = 5.0$, corresponding to the Boltzmann distribution. We run ten separate MCMC simulations, each starting from a different initial configuration and running for 10^5 steps.

In Figure 4.5, we plot the evolution of the squared estimator of the average magnetization, $\hat{m}^2 = \left(\frac{1}{M} \sum_{j=0}^{M-1} m(s^{(j)}) \right)^2$, as computed from the MCMC simulations. Because the SSF update only allows local moves, it captures only a limited portion of the energy landscape, causing the magnetization estimate to converge to a different value. By contrast, the proposed method steadily approaches the true magnetization. Although a uniform proposal permits global transitions, it does so indiscriminately across all states. In a low-temperature Boltzmann distribution with sharp energy minima, this uniform search appears to hinder convergence. In comparison, the proposed method rapidly converges and efficiently estimates the magnetization by prioritizing proposals in low-energy regions of the energy landscape. Consequently, it focuses on the most significant configurations for accurate sampling and estimation, showcasing a clear advantage in exploring low-energy states.

Figure 4.5 presents the magnetization histogram for samples obtained from a single MCMC chain, all starting from the same initial configuration. Owing to the $\mathbb{Z}_2$ -symmetry of the spin glass defined in Eq. (4.4.1), the system features symmetric magnetization peaks corresponding to pairs of configurations with identical energies. While the SSF update tends to sample unevenly by favoring one side of these peaks, the proposed method successfully samples both peaks in a balanced manner. This highlights its capacity to respect the inherent symmetry and to explore the overall configuration space more thoroughly.

We examine the autocorrelation of the magnetization. The autocorrelation function measures the degree of statistical correlation among MCMC samples. Although samples drawn from an ideal probability distribution would be uncorrelated, MCMC samples inherently exhibit finite correlations due to the Markov chain nature of the process. When these correlations diminish

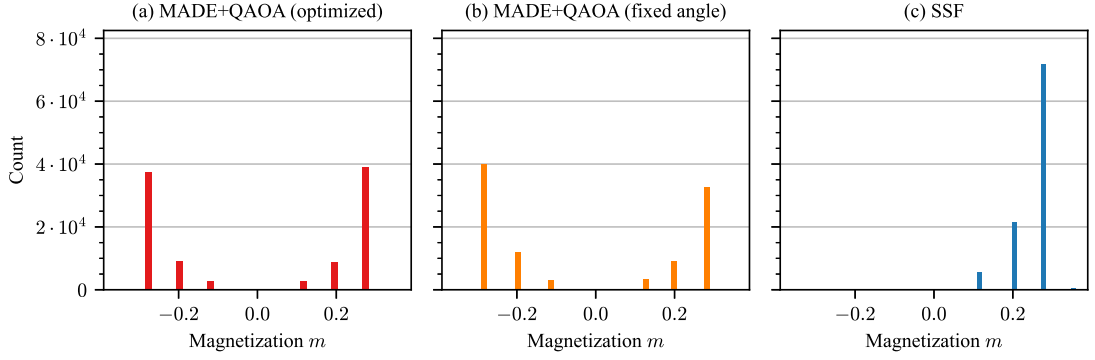


Figure 4.5: Magnetization histogram for samples obtained from a single MCMC chain.

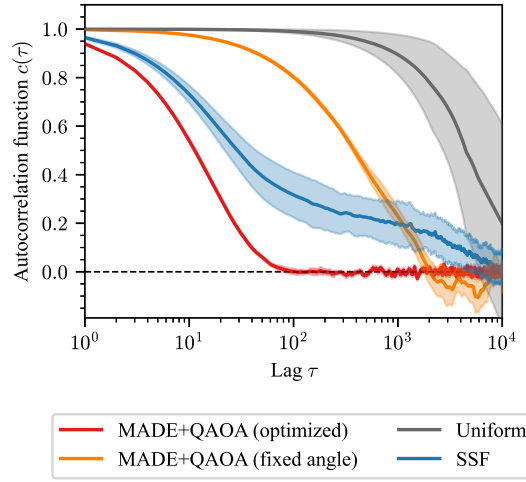


Figure 4.6: Evolution of the autocorrelation function of magnetization.

sufficiently, we regard the MCMC process as converged, making autocorrelation a key metric for gauging convergence efficiency. The autocorrelation function for magnetization is defined as

$$c(\tau) = \frac{\langle m(t + \tau)m(t) \rangle - \langle m(t) \rangle^2}{\langle m(t)m(t) \rangle - \langle m(t) \rangle^2}, \quad (4.4.2)$$

Here, $m(t)$ denotes the magnetization of sample $s^{(t)}$, and $\langle \cdot \rangle$ indicates an average over the entire Markov chain. To mitigate biases introduced by initial conditions, we discard the first 10^4 samples of each chain before computing autocorrelations. Figure 4.6 shows how the magnetization's autocorrelation evolves over time. The proposal distribution derived from optimized QAOA exhibits a notably faster decay in autocorrelation than other methods. By contrast, the fixed-angle QAOA, unlike the trend observed in the spectral gap results, performs on par with or worse than the SSF update. Nevertheless, it is important to emphasize that the SSF update fails

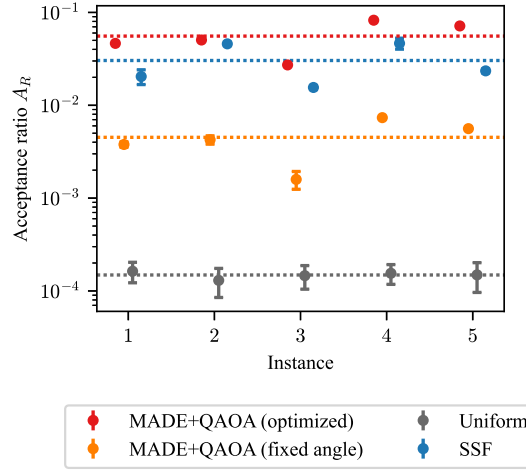


Figure 4.7: Acceptance rate for five random instances of 25-site spin glass. The dotted line shows the average over the five instances. Instance 1 is the same as that used in Section 4.4.2.

to converge to the correct magnetization, as evidenced by the mean magnetization estimates and magnetization histograms, while the proposed method reliably converges to the true value. These observations suggest that the proposed method yields an efficient proposal distribution for complex systems such as spin glasses. Unlike the SSF update, it achieves unbiased sampling while maintaining ergodicity and successfully capturing the Boltzmann distribution.

Finally, to test whether similar results hold for other instances, we calculated the acceptance rate. The acceptance rate is defined as the fraction of accepted proposals out of the total M steps. Figure 4.7 presents the acceptance rates for each proposal distribution across five random instances of a 25-site spin glass. Instance 1 is the same as that used in the previous experiments. The MCMC settings for the five instances are identical to those described earlier. Among the proposals that realize global transitions, the highest acceptance rate is achieved by optimized QAOA, followed by fixed-angle QAOA and then the uniform distribution, indicating that proposals based on the QAOA output distribution facilitate transitions to different configurations. The acceptance rate of SSF is comparable to that of optimized QAOA; however, as shown by the autocorrelation results, SSF introduces biased sampling, so its high acceptance rate does not necessarily correspond to efficient sampling.

4.5 Conclusion

In this study, we present a method to accelerate MCMC sampling for spin glasses by combining QAOA with a generative network sampler. While Boltzmann distributions of spin glasses have traditionally been studied in statistical physics, they are also central to binary combinatorial optimization and machine learning. However, sampling from these distributions remains a well-known challenge, particularly at low temperatures. The proposed approach leverages samples

from a QAOA circuit that approximates the Boltzmann distribution and uses them to train MADE, producing a proposal distribution closely aligned with the target. We benchmarked the proposed method on fully connected spin glasses and conducted several computational analyses. Spectral gap results show that the proposed method maintains strong performance as the system size increases. In magnetization estimation via MCMC, we demonstrated that, unlike the SSF update, the proposed method enables unbiased sampling and converges more rapidly than classical proposals. This advantage is maintained even when using unoptimized QAOA samples. This finding highlights the potential of the proposed approach to exploit current quantum devices without the extensive measurements that are typically required in VQAs.

One of the directions for future research is to extend beyond the Ising model to more general systems. This extension may require exploring alternative neural network architectures in place of MADE. Recent studies on neural network samplers for MCMC have indicated that MADE-based proposals may become inefficient for models that involve substantial computational complexity [133]. As a result, adopting more advanced architectures such as Transformers, which have shown strong performance in modern machine learning [141, 142], or customized architectures tailored to specific physical systems [143], could offer significant advantages.

Chapter 5

Fair Sampling of Ground-State Configurations

In this chapter, we explore use cases for hybrid quantum-classical MCMC. In particular, we analyze the fair sampling properties of hybrid quantum-classical MCMC and investigate their applications. Many important combinatorial optimization problems admit multiple solutions, and for such problems it is often desirable to sample solutions uniformly, a property commonly referred to as fair sampling. Through the analysis of toy models, we demonstrate that hybrid quantum-classical MCMC can achieve fair sampling. Furthermore, using the practically important random k -SAT problem as a benchmark, we apply the proposed methods and show that they achieve fair sampling and convergence efficiency comparable to state-of-the-art classical MCMC approaches. This chapter is based on [Y. Nakano and K. Fujii, arXiv:2512.14552 (2025), [Achievement 3](#)] with modifications to fit the context.

5.1 Introduction

Combinatorial optimization problems play an indispensable role across a wide range of fields, including logistics, finance, drug discovery, and machine learning [144]. However, many practically important problems are classified as NP-hard in computational complexity theory, meaning that their computational cost typically grows exponentially with problem size, which makes the computation of exact optimal solutions extremely challenging. Consequently, a great deal of effort has long been devoted to the development of approximation algorithms and heuristics that can provide high-quality approximate solutions within practical time scales [145]. In recent years, rapid advances in quantum hardware have spurred interest in new optimization heuristics that exploit the principles of quantum mechanics. Among these, quantum annealing (QA) [72, 73] and the quantum approximate optimization algorithm (QAOA) [30] are regarded as two of the most widely studied and promising approaches. In both methods, the cost function of an optimization problem is encoded into a classical Ising Hamiltonian H_P , which is combined with a driver Hamiltonian H_M (such as a transverse field) that induces quantum superposition and transitions between configurations, thereby enabling exploration of the solution space. QA

aims to reach the ground state of H_P (the optimal solution) through adiabatic time evolution implemented on a quantum annealer, whereas QAOA pursues the same goal using parameterized quantum circuits on universal quantum computers. These quantum heuristics have been applied to a variety of optimization tasks, and numerical evidence has suggested potential advantages in time-to-solution scaling over classical algorithms under specific conditions [146, 147], positioning them as promising candidates for next-generation computational paradigms.

On the other hand, in several important applications of optimization, it is not sufficient to find a single optimal solution. Instead, one must exhaustively explore the entire set of optimal solutions (ground states). Representative examples include model counting problems (#SAT) [148], counting solutions of knapsack instances (#Knapsack) [149], and SAT-based membership filters [150, 151]. For such tasks, the ability to draw solutions uniformly at random from the solution space, that is, *fair sampling*, is of critical importance. However, the standard transverse-field driver (2.5.6) widely used in QA and QAOA is known to induce biased transition probabilities due to quantum interference effects, leading to an exponential suppression of the occurrence probabilities of certain ground states [152–155]. To mitigate this bias and achieve perfectly fair sampling, the use of higher-order interaction drivers has been proposed in QA [152], while in QAOA, the introduction of Grover-mixers has been suggested [78–81]. Nonetheless, implementing these more complex drivers typically entails substantial overhead in terms of hardware connectivity and circuit depth, rendering them impractical on current noisy intermediate-scale quantum (NISQ) devices [154].

Against this backdrop, there has been growing interest in leveraging quantum optimization heuristics not only as solvers for optimization problems but also as building blocks for sampling from probability distributions. A particularly promising line of research in this direction is hybrid quantum-classical Markov chain Monte Carlo (MCMC) [43]. In this approach, samples drawn from a quantum circuit are used as a proposal distribution in the transition steps of MCMC methods, such as the Metropolis–Hastings algorithm. Previous work has demonstrated that transitions driven by a time-independent QA Hamiltonian can be effective for sampling from low-temperature Boltzmann distributions of Ising models [43]. In addition, methods have been proposed that employ autoregressive neural-network samplers trained to approximate the output distribution of QAOA circuits, enabling adaptive sampling [137, 156]. These hybrid schemes aim to enhance sampling efficiency by combining the local exploration capabilities of classical MCMC with the global exploration abilities provided by QA and QAOA. This naturally leads to a central open question: *Can hybrid quantum-classical MCMC with a standard transverse-field driver achieve fair sampling while retaining the practical advantages of easy hardware implementation?* If the sampling bias can indeed be mitigated within the MCMC framework using such a simple driver, without resorting to more elaborate mixers, it would constitute a highly practical and significant sampling algorithm for near-term quantum devices.

In this chapter, we investigate the fair-sampling capability of hybrid quantum-classical MCMC. Specifically, we explore its fair-sampling performance and applications through two problem settings based on Ising models. The first is an analysis of small-scale Ising models, including toy models studied in previous work. There we show that introducing MCMC steps on top of the quantum dynamics can restore fair sampling. Furthermore, as a more practically relevant problem setting, we consider sampling solutions of random k -SAT instances. For $k = 2$,

we find that combining proposals generated by quantum dynamics with a standard single spin-flip (SSF) update achieves a level of fairness comparable to that of PT-ICM [65], which is known to realize fair sampling. For $k = 3$, where sophisticated classical sampling algorithms such as PT-ICM are no longer applicable, the hybrid quantum-classical MCMC still attains fair sampling. In addition, for solution counting, we show that it requires a number of transitions comparable to that of WalkSAT [157], a representative classical heuristic solver. These results point to new possibilities for applying quantum heuristics to solution-sampling tasks, which have traditionally been regarded as a weak point of such methods.

The rest of this chapter is organized as follows. Section 5.2 examines the fair-sampling capability using representative small-scale instances that have been employed in previous work. Section 5.3 investigates the sampling of solutions for random k -SAT problems, comparing our approach with classical algorithms and discussing applications to counting problems. Finally, we present our conclusions and discuss future prospects in Section 5.4.

5.2 Degenerate Ground-State Sampling for Small Systems

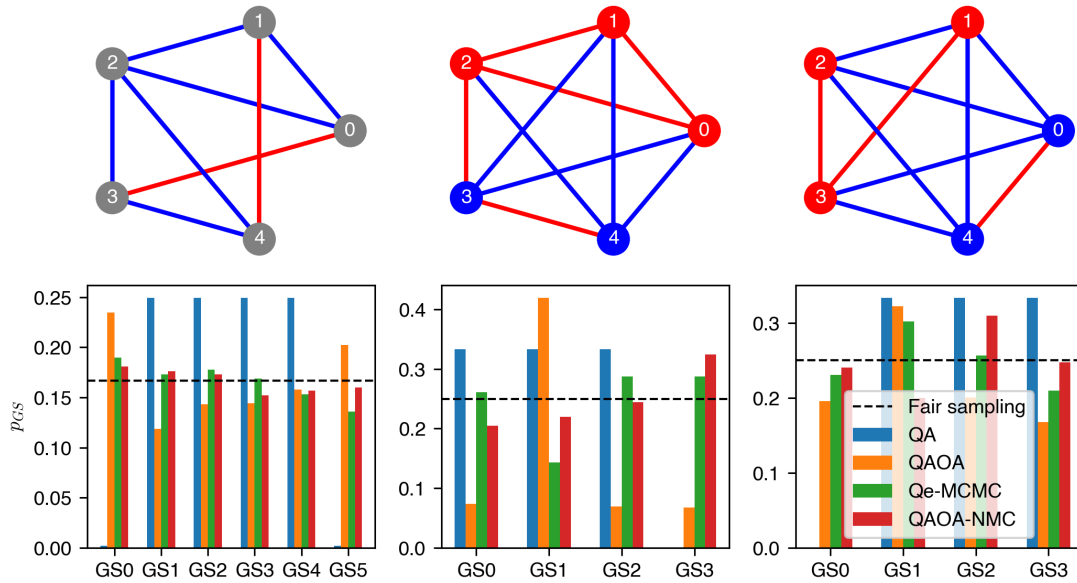


Figure 5.1: Three five-site Ising models with degenerate ground states and their corresponding sampling probabilities. **(Top)** Model specifications. Node colors represent the local fields h_i , and edge colors represent the couplings J_{ij} : red indicates coefficients $+1$, blue indicates -1 , and gray indicates 0 . **(Bottom)** Sampling probabilities over the degenerate ground states. The black dashed line indicates perfectly fair, that is, uniform, sampling over the ground-state manifold.

To investigate the sampling properties of quantum heuristics for the Ising model, we first perform numerical simulations on small-scale models with degenerate ground states. The models

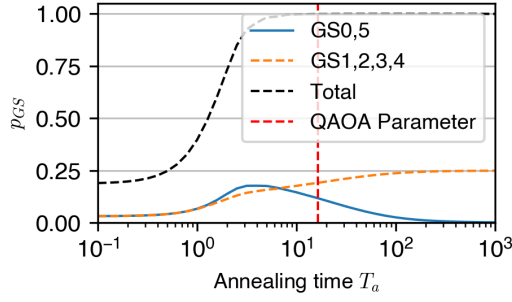


Figure 5.2: Dependence of the final QA state on the annealing time T_a for the instance with sixfold degeneracy shown in Fig. 5.1. The red line indicates the total effective evolution time, $T_{\text{QAOA}} = \sum_{i=1}^P (\beta_i + \gamma_i)$ obtained by interpreting the optimized QAOA parameters as effective evolution times.

considered are $k = 2$ Ising models (2.2.2) with five sites. The interaction coefficients take only two values, $J_{ij} \in \{+1, -1\}$, which results in a degenerate ground-state manifold.

We sample degenerate ground states using four methods: direct sampling of the final state via QA and QAOA, and two hybrid quantum-classical MCMC approaches, Qe-MCMC and QAOA-NMC. For all methods, a transverse-field driver Hamiltonian is employed as the mixer Hamiltonian. The specific settings are as follows: For QA, we calculate the final state by integrating the Schrödinger equation using the QuTiP library with an annealing time $T_a = 10^3$. For QAOA, we employ a circuit of depth $p = 5$, optimize parameters via the BFGS algorithm [158], and perform 10^3 measurements. QAOA-NMC uses these samples as training data, with all hyperparameters and settings adopted from Ref. [156]. Qe-MCMC is executed with hyperparameters based on Ref. [44]. For both MCMC approaches, 10^3 samples are drawn from the Boltzmann distribution with $\beta = 10$. Regarding implementation, Qulacs [159] is used for quantum circuit simulations of QAOA and Qe-MCMC, while PyTorch [160] is used for the training and inference of MADE.

Figure 5.1 shows the sampling results for the three models previously examined in Ref. [154, 155]. As confirmed in earlier studies, QA suppresses the occurrence probabilities of certain ground states, even for sufficiently large T_a . Interestingly, in the case of short-time annealing ($T_a \lesssim 10$), the sampling probabilities of all ground states remain fair in the final state (Fig. 5.2). This suggests that, in principle, fair sampling might be achieved by appropriately optimizing the annealing time. However, determining such an optimal value of T_a in advance is highly nontrivial.

By contrast, hybrid quantum-classical MCMC methods show no such suppression, retaining sampling probabilities for all degenerate ground states. This is primarily because both algorithms employ QA Hamiltonian dynamics with short evolution times for transitions, thereby mitigating the driver-induced suppression of degenerate ground states. In fact, the red dotted line in Fig. 5.2 denotes $T_{\text{QAOA}} = \sum_{i=1}^P (\beta_i + \gamma_i)$, the effective annealing time derived from optimized QAOA parameters, which corresponds to the temporal region where annealing maintains relatively fair sampling. Furthermore, the Metropolis–Hastings post-processing applied to the quantum

proposals also contributes significantly to restoring fair sampling. QAOA-NMC, in particular, shows reduced bias compared with the raw QAOA output distribution on which it is trained. This improvement arises because the low-temperature Boltzmann distribution imposes acceptance probabilities that effectively act as a filter, allowing only ground states and a subset of low-energy excited states to be accepted. These findings indicate that combining quantum heuristics with MCMC is a promising strategy for achieving fair sampling.

5.3 Random k -SAT Sampling

In this section, we perform a numerical analysis of the fair sampling and applications of hybrid quantum-classical MCMC on random k -SAT instances.

5.3.1 Problem Setting

The Boolean Satisfiability Problem (SAT) is one of the most fundamental problems in computer science. It asks whether there exists an assignment of variables that evaluates a given logical formula to **True**. In particular, the case where each constraint consists of a logical disjunction (OR) of k literals is referred to as k -SAT.

For n binary variables $x_1, \dots, x_n$, let a literal l be either the variable itself x_i or its negation $\neg x_i$. An instance of k -SAT is described in the conjunctive normal form (CNF) as a logical conjunction (AND) of “clauses” C_m , each consisting of the disjunction of k literals:

$$F(\mathbf{x}) = C_1 \wedge C_2 \wedge \dots \wedge C_M = \bigwedge_{m=1}^M \left(\bigvee_{j=1}^k l_{m,j} \right). \quad (5.3.1)$$

Here, M denotes the total number of clauses, and $l_{m,j}$ represents the j -th literal included in the m -th clause. The objective of the problem is to find a configuration of variables $\mathbf{x} \in \{0, 1\}^n$ such that $F(\mathbf{x}) = 1$ (**True**).

The k -SAT problem addressed in this study is a typical class of Constraint Satisfaction Problems (CSP). From a physical perspective, the SAT problem can be formulated as an energy minimization problem by treating each clause C_m as a local interaction energy term. We define a Hamiltonian H_m for each clause, which assigns an energy of 0 if the logical formula is satisfied **True** and a positive energy penalty $H_m > 0$ if it is not satisfied **False**. The total Hamiltonian of the system, H_{SAT} , is then constructed as:

$$H_{\text{SAT}} = \sum_{m=1}^M H_m(\mathbf{x}) \quad (5.3.2)$$

In k -SAT, the sole condition under which a clause C_m remains unsatisfied is when all k literals in the clause evaluate to **False**. Therefore, by constructing a penalty function that increases the energy only under this condition and expanding it using Ising variables $s_i \in \{+1, -1\}$ (or $x_i \in \{0, 1\}$), the problem can be mapped to the Ising model or QUBO formulation introduced in Section 2.2.

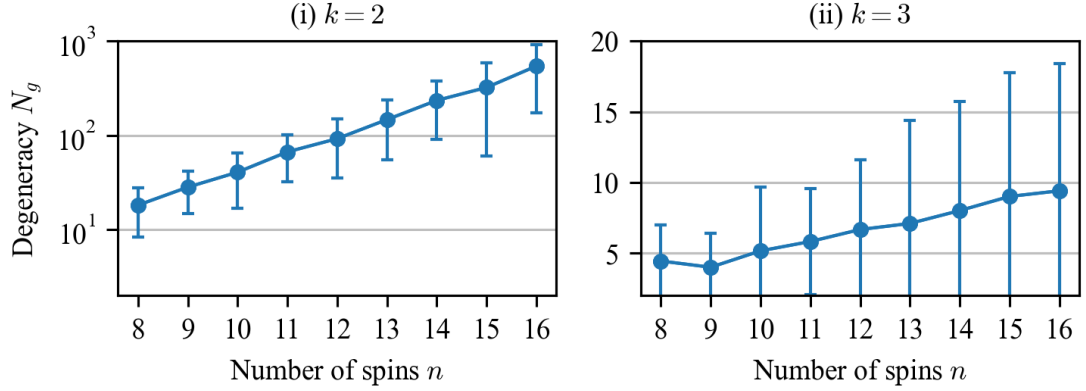


Figure 5.3: Number of solutions, or degeneracy, N_g , of the generated random 2-SAT and 3-SAT instances as a function of the problem size n .

For example, the Hamiltonian for a clause $C = x_i \vee x_j \vee x_k$ in 3-SAT (where variables are binary) can be constructed as follows. Since this clause evaluates to **False** only when $(x_i, x_j, x_k) = (0, 0, 0)$, the penalty term is given by:

$$H_m = (1 - x_i)(1 - x_j)(1 - x_k). \quad (5.3.3)$$

This term evaluates to 1 only for the state $(0, 0, 0)$ and 0 otherwise. By expanding this expression and converting the variables to spin variables via $s_i = 1 - 2x_i$, an Ising Hamiltonian containing up to three-body interactions is obtained.

Furthermore, it is known that the computational hardness of k -SAT changes dramatically depending on the constraint density $\alpha = M/n$, defined as the ratio of the number of clauses M to the number of variables n . A phase transition regarding the existence of a solution (SAT/UNSAT transition) occurs near a specific critical value α_c . This region is of particular importance as it represents the most difficult regime even for classical heuristics.

5.3.2 Numerical Experiments

Instance Set

In this study, we utilize Random k -SAT problems with $k = 2$ and 3. Specifically, to address problems with a constraint density near the critical point α_c [146], where quantum heuristics are expected to be particularly effective, and to ensure the existence of multiple SAT solutions, we generate the instance set as follows. First, we generate a Random k -SAT instance consisting of $M = \lfloor \alpha_c n \rfloor + 1$ clauses, randomly selected from the set of all possible $\binom{n}{k} 2^k$ clauses. Next, we use a classical solver to filter out instances that have one or zero SAT solution. Through this procedure, we prepare an instance set consisting of 100 random instances for each problem size n , ranging from $8 \leq n \leq 16$. Figure 5.3 shows the number of degenerate ground states, N_g , for each problem size n .

Method

In this experiment, we perform the sampling of SAT solutions for the generated instance set using a hybrid quantum-classical MCMC. We adopt QAOA-NMC for this purpose. The rationale for this choice is twofold: QAOA-NMC using an optimized QAOA circuit is expected to demonstrate faster convergence than Qe-MCMC, and it facilitates a direct comparison with the QAOA output distribution. This approach allows us to clearly elucidate the correction effects resulting from the MCMC steps.

The detailed settings for QAOA-NMC are described below. For QAOA, we use a circuit with depth $p = 5$. To reduce the computational cost associated with optimization, we introduce linear parameters [82]. This involves replacing the $2p$ parameters of the QAOA with a linear schedule consisting of four parameters: $\beta_l = \beta_{\text{slope}} \frac{l}{p} + \beta_{\text{intcp}}$ and $\gamma_l = \gamma_{\text{slope}} \frac{l}{p} + \gamma_{\text{intcp}}$. Furthermore, since such linear schedules contain parameters (referred to as “fixed angles”) that universally yield good cost functions for instances, we prepare two types of QAOA circuits: one based on optimized parameters and another based on these fixed parameters derived from the schedule [83, 84]. For MADE, the size of the training dataset is set to 10^3 , and all hyperparameters are identical to those in Ref. [156]. Sampling of the ground state is performed via the Boltzmann distribution with $\beta = 10$. We executed 10^4 sampling steps for each of 10 random initial states.

Hybrid Algorithm

As shown in Fig. 5.3, the instance set exhibits a high degree of degeneracy. Particularly for $k = 2$, the system possesses a vast number of solutions; consequently, it is highly probable that some ground states are missing from the training data, and the learned MADE may fail to transition to all ground states. To correct these biases, we introduce a new hybrid method that combines the MADE update with a sweep consisting of n single spin-flip (SSF) updates. The computational cost of an SSF sweep is minimal, and the hybrid update remains of the same order as the MADE update. The introduction of the SSF sweep is intended to explore other ground states located in the vicinity of the state proposed by MADE. Since many instances in $k = 2$ form clusters with close Hamming distances, this approach is expected to play a role in complementing the sampling of ground states that were missing during the MADE training. Hereafter, we denote this hybrid variant combining QAOA-NMC and SSF sweeps as QAOA-assisted Hybrid neural MCMC (QAOA-HMC).

Classical Algorithms

As benchmarks for classical algorithms in this study, we adopt PT-ICM, a state-of-the-art sampling algorithm for 2-body Ising models, and WalkSAT, a representative method for Stochastic Local Search (SLS).

PT-ICM PT-ICM [65] is an algorithm that combines Parallel Tempering (PT) [62, 63] with Isoenergetic Cluster Moves (ICM) [64]. PT simulates multiple replicas with different temperature parameters in parallel and periodically exchanges their configurations. High-temperature replicas perform global exploration, while low-temperature replicas perform local optimization, enabling

the system to efficiently overcome energy barriers. The ICM identifies clusters of interacting spins and performs non-local updates by flipping them collectively. This update is designed to preserve the total energy of the system (or cause minimal change), facilitating transitions to distant valleys in the state space by “tunneling” through regions with high energy barriers. The overall process is described by the procedure below. Due to the synergistic effect of PT and ICM, PT-ICM achieves extremely fast relaxation and efficient sampling even in complex systems like spin glasses.

WalkSAT WalkSAT [157] is an incomplete solver that heuristically explores the solution space, unlike complete solvers such as the DPLL (Davis-Putnam-Logemann-Loveland) algorithm [161], which performs an exhaustive search. It is known for its ability to discover satisfying assignments extremely quickly, especially for random k -SAT instances. The search process of WalkSAT begins with a random initial assignment and proceeds by repeating the following steps:

1. Randomly select one unsatisfied clause under the current variable assignment.
2. Determine which variable within that clause to flip. At this point, one of the following operations is performed based on a probability p (noise parameter):
 - Random Move (probability p): Randomly select a variable in the clause and flip its value. This promotes escape from local optima (local minima).
 - Greedy Move (probability $1 - p$): Select the variable that, when flipped, results in the greatest reduction (or minimal increase) in the total number of unsatisfied clauses.

By balancing this greedy optimization and random walk, WalkSAT efficiently explores the complex energy landscape, aiming to reach a global optimum (satisfying assignment). In this study, we utilize an algorithm called WalkSATlm [162], which is modified with a more complex cost function. Although WalkSAT is originally designed to find a single SAT solution, it can be utilized to enumerate SAT solutions with a simple modification. This involves adding a new clause to the problem’s CNF that prohibits the already discovered solution, thereby searching for undiscovered solutions. By repeating this process until the CNF becomes UNSAT, all SAT solutions can be enumerated.

5.3.3 Fair Sampling

First, to assess the fair sampling capability, we quantify the sampling bias using the ratio of the maximum to minimum probabilities of the ground states, denoted as $p_{\max}/p_{\min}$. For QAOA and MADE, this ratio is derived based on the directly computed output distributions; for MCMC, it is calculated from the histogram constructed from all sampled ground states. It should be noted that for MCMC, there are cases where the sampler fails to visit all degenerate ground states. In such instances, the ratio $p_{\max}/p_{\min}$ cannot be defined, and thus these results are excluded from the analysis.

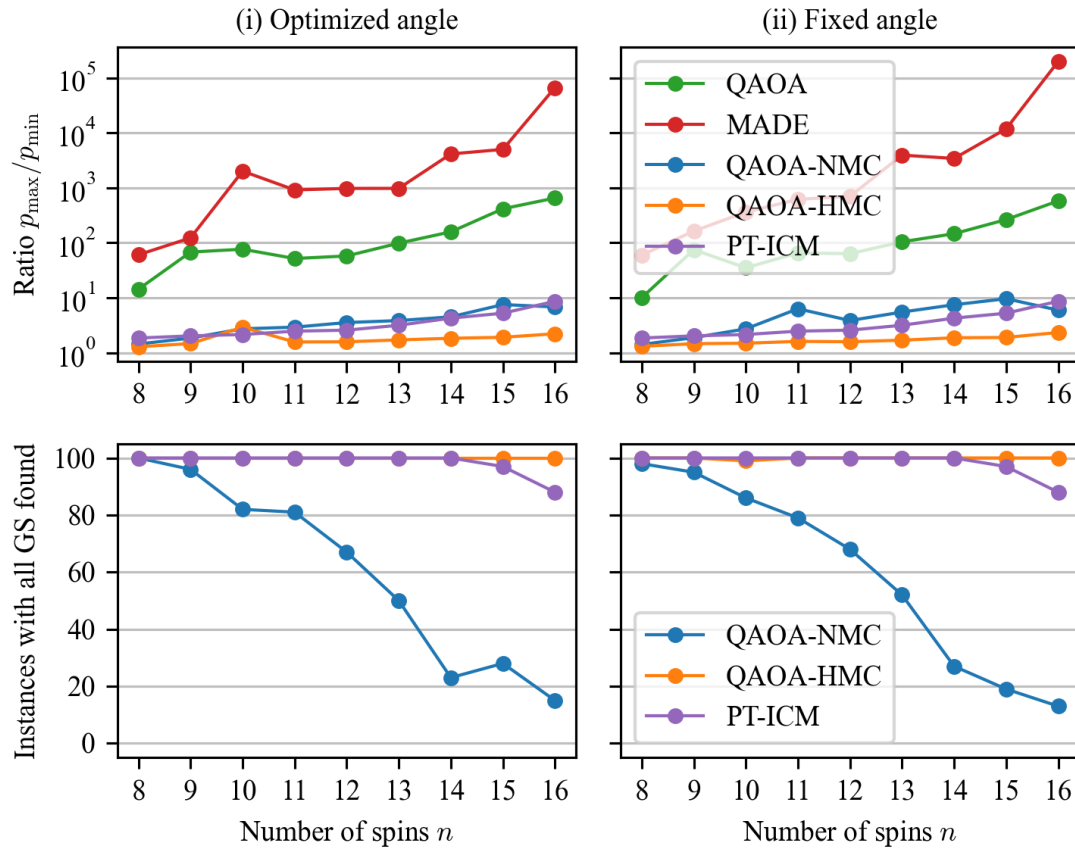


Figure 5.4: The ratio $p_{\max}/p_{\min}$ and the number of instances where all ground states were successfully discovered as a function of the problem size n for the $k = 2$ case.

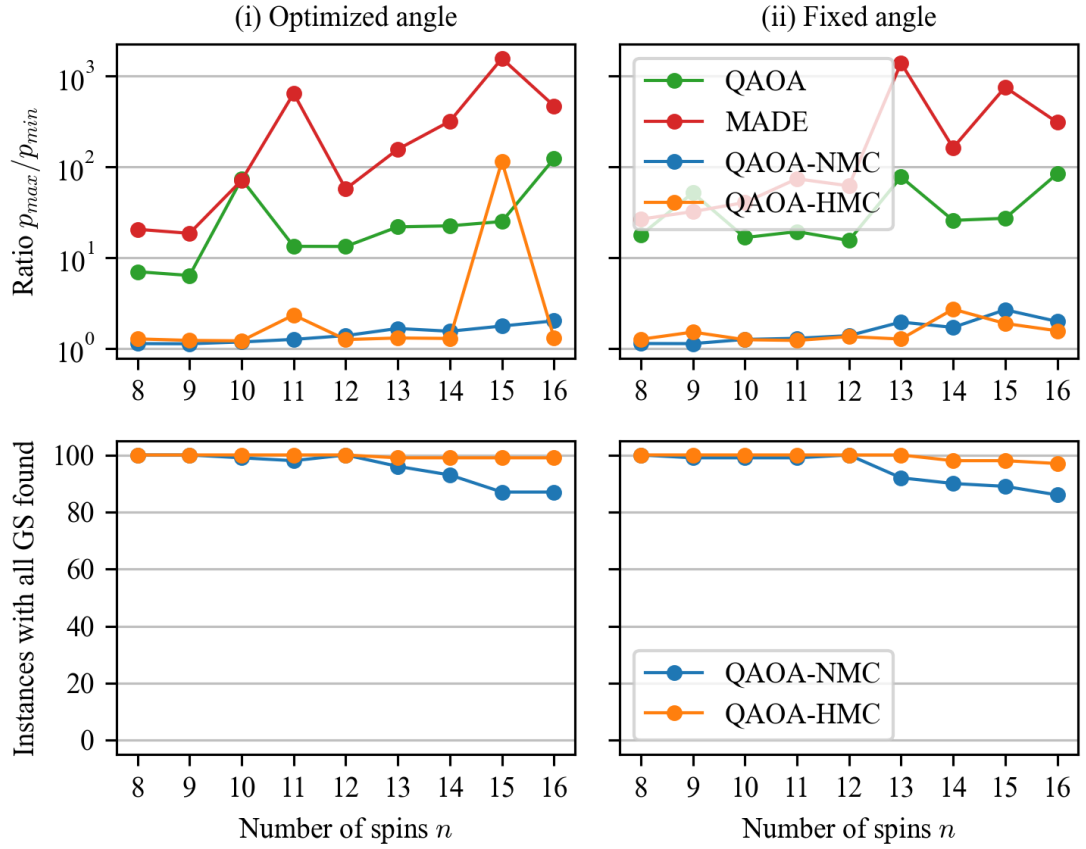


Figure 5.5: The ratio $p_{\max}/p_{\min}$ and the number of instances where all ground states were successfully discovered as a function of the problem size n for the $k = 3$ case.

Random 2-SAT

Figure 5.4 illustrates the ratio $p_{\max}/p_{\min}$ and the number of instances where all ground states were successfully discovered as a function of the problem size n for the $k = 2$ case. As discussed in the analysis of small systems, optimized QAOA exhibits significant bias. Consequently, the MADE trained on these outputs becomes an even more biased sampler. Therefore, QAOA-NMC often fails to sample all degenerate ground states. In contrast, QAOA-HMC successfully discovered all ground states for all instances. Furthermore, in terms of the ratio $p_{\max}/p_{\min}$, it demonstrates fairness comparable to, or even surpassing, that of PT-ICM, which is known for achieving fair sampling. Additionally, QAOA yielded nearly identical results regardless of whether optimized parameters or fixed angles were used. This is because the linear schedule employed in this study exhibits a strong concentration of optimal parameters [82], and the fixed angles, derived as the medians of these values, are very close to the optimal parameters. While this does not directly affect fair sampling performance, the use of fixed angles eliminates the need for the computationally expensive optimization of QAOA, making it a practically significant property.

Random 3-SAT

Figure 5.5 presents the ratio $p_{\max}/p_{\min}$ and the number of instances where all ground states were discovered as a function of problem size n for $k = 3$. While the trend is fundamentally similar to that of $k = 2$, the overall degree of bias is reduced. This is attributed to the smaller total number of solutions, N_g , compared to the $k = 2$ case. Consequently, even QAOA-NMC succeeded in exploring all ground states for many instances. However, in the case of $k = 3$, there were instances where even QAOA-HMC failed. All such instances possess configurations that are weakly correlated with other ground states; thus, they cannot benefit from the local neighborhood search provided by the SSF sweep.

5.3.4 Model Counting

Next, as an application of the fair sampling property, we perform the enumeration of SAT solutions. This is a problem known as Model Counting (#SAT) [148]. Here, we compare the counting performance between MCMC sampling and WalkSATlm. We define a state transition as a single step and evaluate the total number of steps required to discover all solutions. For QAOA-NMC, QAOA-HMC and WalkSATlm, we calculate the average number of steps over 10 trials, whereas for PT-ICM, we count the steps in a single trial. As in the previous section, instances where not all degenerate ground states were discovered within a limited number of steps are excluded from the analysis.

Random 2-SAT

Figure 5.6 displays the number of steps required to enumerate all ground states and the number of instances where MCMC demonstrated superiority over WalkSATlm, as a function of problem size n for $k = 2$. QAOA-NMC fails to surpass the performance of PT-ICM in almost all instances and is also inferior to WalkSATlm in many cases. This is because, as mentioned in the previous section, QAOA-NMC fails to sample all ground states for many instances. Consequently, the

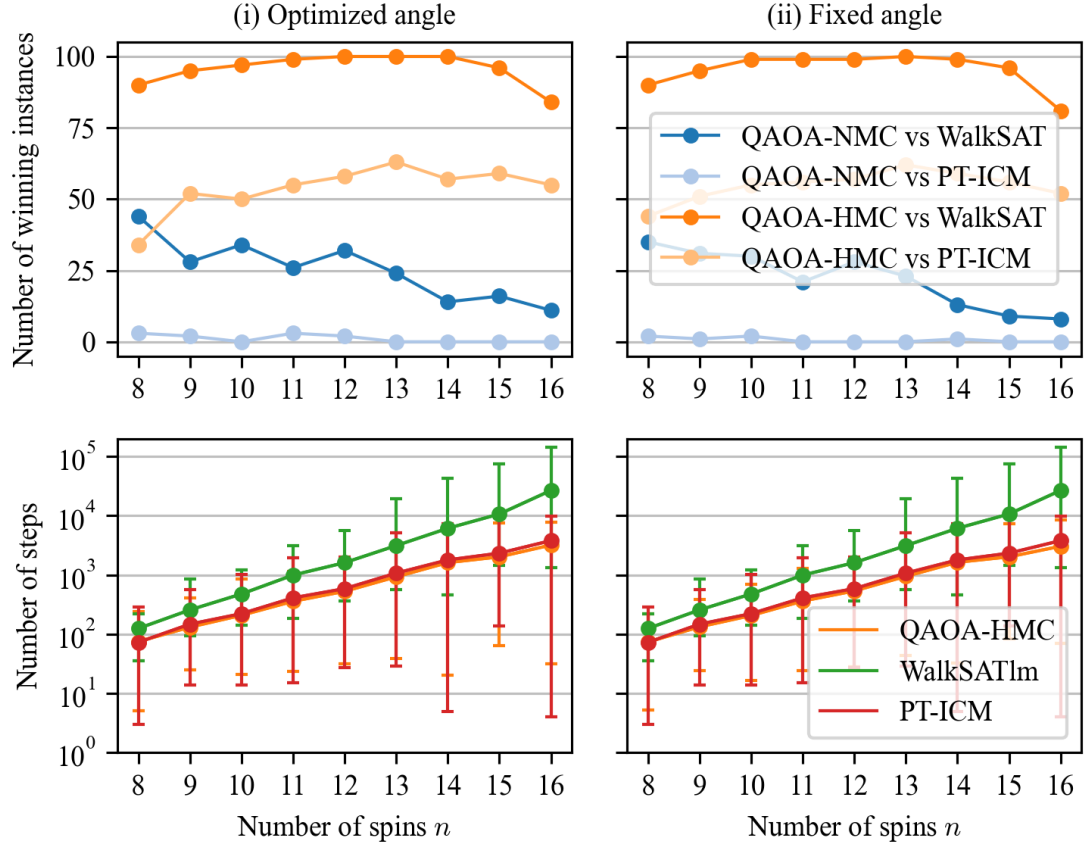


Figure 5.6: The number of steps required to enumerate all ground states and the number of instances where MCMC demonstrated superiority over WalkSATlm and PT-ICM, as a function of problem size n for $k = 2$.

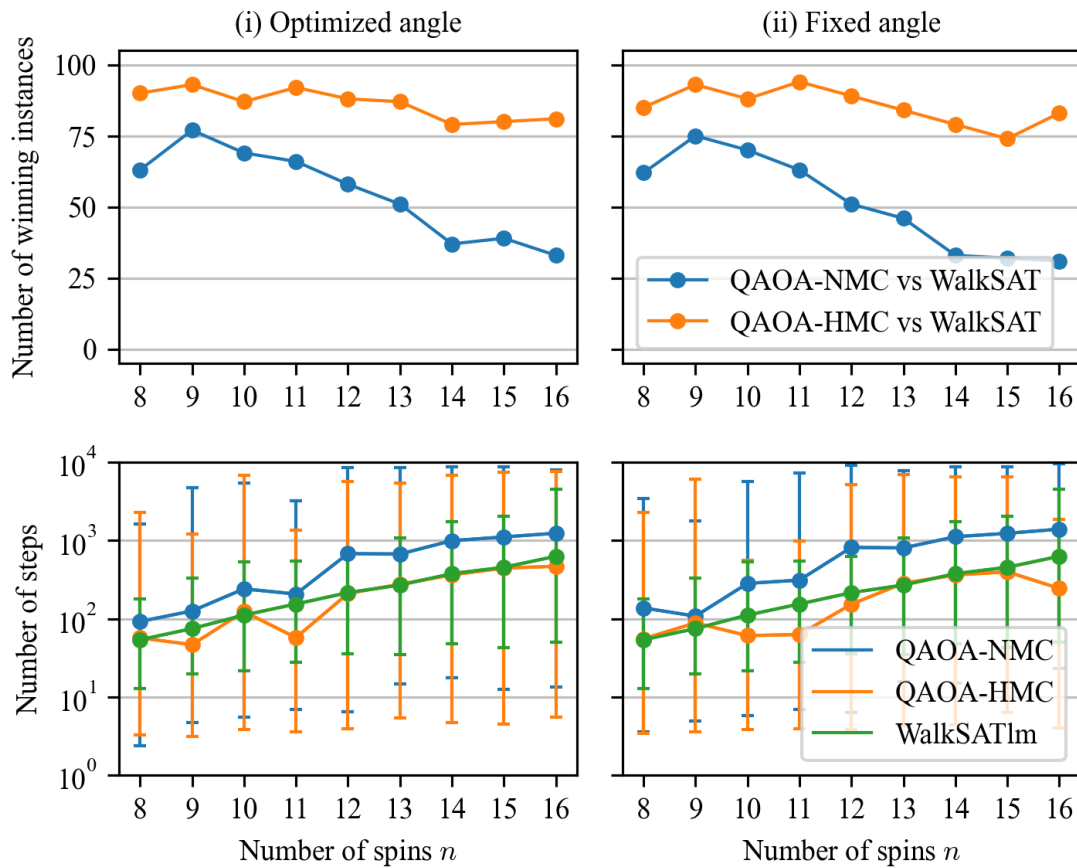


Figure 5.7: The number of steps required to enumerate all ground states and the number of instances where MCMC demonstrated superiority over WalkSATlm, as a function of problem size n for $k = 3$.

step count data for QAOA-NMC is omitted from Fig. 5.6. In contrast, QAOA-HMC outperforms WalkSATlm in most cases and achieves performance nearly equivalent to that of PT-ICM.

Random 3-SAT

Figure 5.7 displays the number of steps required to enumerate all ground states and the number of instances where MCMC was superior, as a function of problem size n for $k = 3$. Unlike the $k = 2$ case, QAOA-NMC possesses a certain degree of competitiveness against WalkSATlm. Although QAOA-HMC still outperforms WalkSATlm in many instances, the scaling of the average step count is comparable to that of WalkSATlm. As discussed in the previous section, the efficacy of the SSF sweep diminishes in this regime; consequently, the bias of QAOA-HMC increases, leading to significant performance variance across instances.

5.4 Conclusion

In this chapter, we investigated the fair sampling property of hybrid quantum-classical MCMC algorithms for problems with degenerate ground states.

Through numerical experiments on small Ising models, we showed that although quantum annealing and QAOA with transverse-field drivers exhibit strong sampling bias, the introduction of Metropolis–Hastings post-processing substantially restores uniformity over degenerate ground states. This result highlights that sampling bias arising from quantum dynamics does not necessarily preclude fair sampling when combined with classical MCMC updates that enforce detailed balance.

We further examined random k -SAT problems as practically relevant benchmarks with highly degenerate solution spaces. For random 2-SAT, we found that a hybrid MCMC combining QAOA-trained neural proposals with single spin-flip updates achieves a level of fairness comparable to that of PT-ICM, a well-established classical fair sampler. For random 3-SAT, where PT-ICM is no longer applicable, the hybrid approach still yields approximately uniform sampling, although instance-dependent limitations remain, particularly for solutions that are weakly connected in configuration space.

As an application, we studied model counting and observed that the number of transitions required to enumerate all solutions using hybrid MCMC is comparable to that of WalkSAT, a representative stochastic local search algorithm. This suggests that hybrid MCMC can serve as an alternative sampling-based approach for enumeration tasks, rather than a replacement for existing classical solvers.

Chapter 6

Conclusion

This thesis elucidates, from algorithmic and empirical perspectives, the effectiveness of hybrid quantum-classical MCMC algorithms that use samples generated by quantum circuits as proposal distributions and guarantee detailed balance through classical Metropolis–Hastings (MH) updates. The central claim is that, rather than employing quantum heuristics as the final sampler, one can restrict their role to generating transition proposals and justify the resulting procedure within classical MCMC, including accept-reject decisions. In this way, sampling performance, including convergence speed and fairness, can be improved in a form that is implementable even on the NISQ devices.

In Chapter 3, we proposed Quantum Alternating Operator Ansatz Monte Carlo (QAOA-MC), a hybrid quantum-classical MCMC algorithm that uses a fixed-depth QAOA circuit as a proposal distribution to accelerate convergence while remaining compatible with NISQ devices. By adopting a QAOA-type circuit, we increased practical feasibility with shallow circuits, and we showed that local optimization with respect to the acceptance rate can contribute to improved mixing. In numerical experiments for the Boltzmann distribution of spin glasses, we observed an approximately quadratic improvement in the absolute spectral gap compared with a uniform proposal, providing concrete evidence that quantum circuit based proposals can accelerate MCMC convergence.

In Chapter 4, we presented a method to accelerate MCMC sampling by training an autoregressive generative model (MADE) using QAOA samples, thereby constructing a proposal distribution consistent with the target Boltzmann distribution of spin glasses. Benchmarks on fully connected spin glasses showed that the performance of the proposed proposal mechanism remains strong as the system size increases. In estimating physical observables such as magnetization, we demonstrated that the method achieves faster convergence than the single spin-flip (SSF) update while preserving unbiased sampling. Moreover, the advantage persists even when the QAOA samples are not rigorously optimized. This indicates the potential for practical use of samples from quantum devices without relying on variational optimization or extensive measurement overhead.

In Chapter 5, we focused on fair sampling, which is crucial for problems with degenerate ground states. Using small Ising models, we showed that while quantum annealing and a transverse-field Hamiltonian mixer can exhibit strong sampling bias, a hybrid MCMC equipped

with MH post-processing substantially restores uniformity over the degenerate solution set. We further considered random k -SAT as a practically relevant benchmark with highly degenerate solution spaces. For random 2-SAT, we found that a hybrid MCMC combining a QAOA-trained proposal with SSF updates achieves a level of fairness comparable to PT-ICM, a well-established classical fair sampler. For random 3-SAT, we confirmed that the hybrid approach still yields approximately uniform sampling. As an application, we examined model counting, namely solution enumeration, and found that the number of transitions required to obtain all solutions is comparable to that of WalkSAT. This suggests that hybrid MCMC can function not as a simple replacement for existing classical solvers, but as an alternative sampling-based approach for enumeration and estimation tasks.

In summary, through (i) optimization strategies for accelerating convergence, (ii) quantum proposals that combine quantum samples with generative neural samplers, and (iii) verification of fair sampling properties together with applications to solution sampling tasks in SAT problems, this thesis demonstrates that hybrid quantum-classical MCMC algorithms provide an effective framework for extracting the advantages of quantum computation even under the constraints of NISQ devices. In the NISQ era, quantum algorithms have primarily developed around VQAs, but it has also become clear that obtaining a realistic computational advantage is difficult due to issues such as variational optimization and measurement costs. Against this background, sampling-based NISQ algorithms have attracted attention, leading to the emergence of hybrid quantum-classical MCMC. This thesis advances this approach and elevates it to a more practical level. In particular, QAOA-NMC based on fixed-angle QAOA inherits insights from VQAs while achieving accelerated MCMC without variational optimization, thereby opening a path toward obtaining computational advantage on currently available NISQ devices. Furthermore, applications to combinatorial optimization problems that exploit fair sampling properties demonstrate that NISQ algorithms can leverage quantum advantages for practically important tasks such as solution sampling. These results represent a major step toward the realization of practical sampling algorithms using near-term quantum devices and are expected to become an important milestone for achieving quantum advantage with NISQ devices across a wide range of fields, including statistical physics, machine learning, and combinatorial optimization.

Finally, we discuss directions for future research. A key unresolved issue for hybrid quantum-classical MCMC is whether the advantages observed in this thesis and in prior work persist for larger systems and on real hardware. Existing evaluations, including those in this thesis, have been limited to scaling analyses over small systems for which spectral gaps can be computed. In contrast, MCMC in statistical physics is routinely applied to large systems with hundreds to thousands of variables. If it can be confirmed that hybrid quantum-classical MCMC also outperforms classical MCMC in these regimes, it would constitute quantum advantage in a substantive sense. Another open question concerns robustness on actual hardware. The original Qe-MCMC [43] study demonstrated, using an IBM superconducting processor, that an advantage over classical proposal distributions can be maintained for a 10-qubit system. However, it remains unclear whether Qe-MCMC can overcome the difficulty of preserving symmetry under noise and still maintain performance for larger systems.

The QAOA-NMC proposed in Chapter 4, which combines QAOA with machine learning based distribution estimation, may address these issues. QAOA has already been executed on

systems exceeding one hundred qubits [84, 163, 164], and with fixed angles it should be possible to run QAOA-NMC with realistic circuit execution and measurement costs even at these scales. Regarding scalability, although the problem sizes that can be handled directly by current devices remain limited, it may be possible to apply the method to problems larger than the hardware size by employing divide and conquer strategies that partition the problem into subsystems and incorporate quantum subroutines [165]. QAOA-NMC can be naturally integrated with such structures, potentially enabling practical performance improvements using NISQ devices even for real world problems with hundreds to thousands of variables. While analyzing these directions is left to future work, the results of this thesis indicate the possibility that algorithms based on currently available NISQ devices can exhibit quantum advantage for problems of practical importance, and we believe that this represents an important step toward near-term societal applications of quantum computing.

Appendix A

Spectral Gap For Different Temperatures

Table A.1: The exponential fit parameters in Fig. A.1. The fitting compares the quantum proposal’s enhancement over the uniform proposal with the best scaling (highlighted in bold) among the data for which fitting was feasible. Uncertainties are from the covariance matrices obtained in the fitting process.

Proposal type	$\beta = 0.1$	$\beta = 1.0$	$\beta = 2.0$	$\beta = 5.0$	$\beta = 10.0$
MADE+QAOA (optimized)	1.73(8)	0.99(7)	–	0.23615	0.2728
MADE+QAOA (fixed angle)	1.430(34)	0.732(14)	0.27518	0.355(21)	0.371(21)
QAOA-MC	0.000(5)	0.272(8)	0.459(11)	0.640(12)	0.689(12)
Uniform	0.0954(8)	0.632(5)	0.829(6)	0.948(6)	0.977(5)
SSF	–	0.374(35)	0.60(6)	1.21(8)	–
Enhancement (vs Uniform)	None	None	3.02(20)	4.02(25)	3.60(11)

Figure A.1 shows how the spectral gap depends on system size across different temperatures, including the data at $\beta = 10.0$ previously shown in Fig. 4.3. The $\langle \delta \rangle = 2^{-kn}$ fitting is performed only on valid data; therefore, for some data the scaling of $\langle \delta \rangle$ could not be obtained. Table A.1 reports the fitted values of k . At high temperatures ($\beta \leq 1$), classical proposals are sufficiently efficient and outperform the proposed method. It is important to note that simulations in this regime are classically tractable, and the proposed method is specifically designed to utilize approximate Boltzmann distributions generated by QAOA at low temperatures. Therefore, it is not intended for high-temperature settings. In the low-temperature regime ($\beta > 1$), the optimized QAOA circuit yields the best performance. Interestingly, around $\beta = 1$, the fixed-angle QAOA performs better than its optimized counterpart. The proposed method consistently maintains an advantage over classical proposals throughout the low-temperature regime. This result suggests that even without parameter optimization, quantum circuits can serve as effective generators of Boltzmann samples at finite temperatures, which is a highly promising finding.

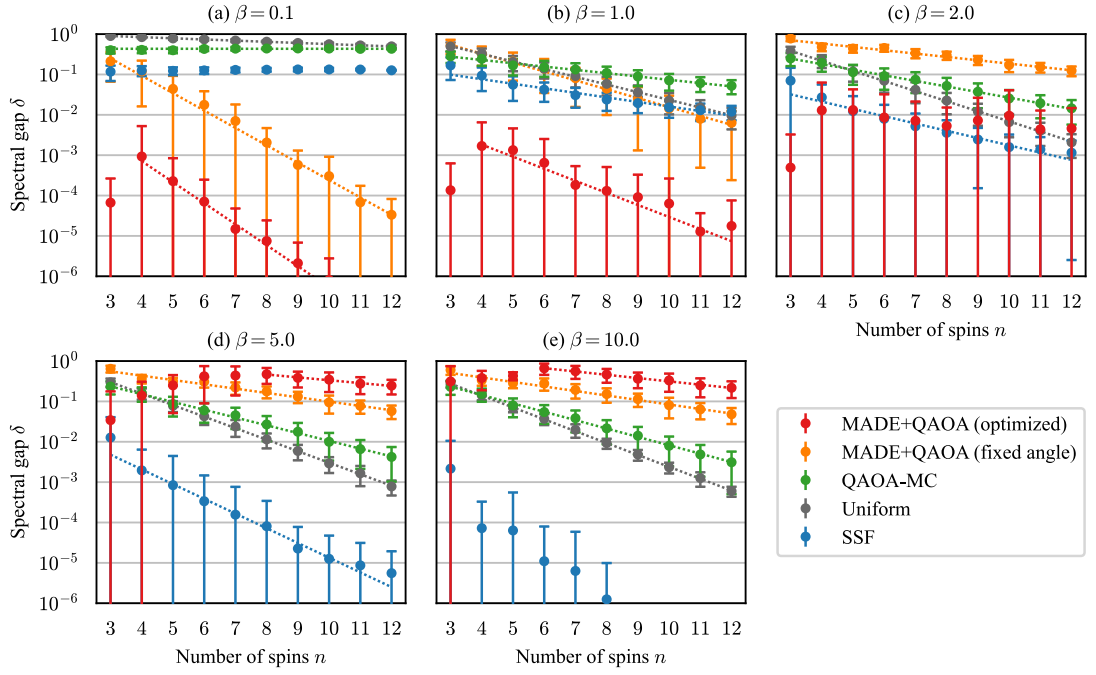


Figure A.1: System size dependence of the spectral gap at different temperatures. The dotted lines indicate the scaling obtained through least-squares fitting. The resulting values of fit parameter k are given in Table A.1.

Bibliography

- [1] Feynman, R. P. Simulating physics with computers. *International Journal of Theoretical Physics* **21**, 467–488 (1982). URL <https://doi.org/10.1007/BF02650179>.
- [2] Deutsch, D. & Penrose, R. Quantum theory, the Church-Turing principle and the universal quantum computer. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* **400**, 97–117 (1985). URL <https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1985.0070>.
- [3] Bernstein, E. & Vazirani, U. Quantum Complexity Theory. *SIAM Journal on Computing* **26**, 1411–1473 (1997). URL <https://doi.org/10.1137/S0097539796300921>.
- [4] Shor, P. W. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings of the 35th Annual Symposium on Foundations of Computer Science, SFCS '94*, 124–134 (IEEE Computer Society, USA, 1994). URL <https://doi.org/10.1109/SFCS.1994.365700>.
- [5] Grover, L. K. Quantum mechanics helps in searching for a needle in a haystack. *Phys. Rev. Lett.* **79**, 325 (1997). URL <https://doi.org/10.1103/PhysRevLett.79.325>.
- [6] Shor, P. W. Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A* **52**, R2493–R2496 (1995). URL <https://link.aps.org/doi/10.1103/PhysRevA.52.R2493>.
- [7] Aharonov, D. & Ben-Or, M. Fault-Tolerant Quantum Computation with Constant Error Rate. *SIAM Journal on Computing* **38**, 1207–1282 (2008). URL <https://doi.org/10.1137/S0097539799359385>.
- [8] Kitaev, A. Y. Quantum Error Correction with Imperfect Gates (1997). URL https://ideas.repec.org/h/spr/sprchp/978-1-4615-5923-8_19.html.
- [9] Lloyd, S. Universal Quantum Simulators. *Science* **273**, 1073–1078 (1996). URL <https://www.science.org/doi/abs/10.1126/science.273.5278.1073>.
- [10] Harrow, A. W., Hassidim, A. & Lloyd, S. Quantum Algorithm for Linear Systems of Equations. *Phys. Rev. Lett.* **103**, 150502 (2009). URL <https://link.aps.org/doi/10.1103/PhysRevLett.103.150502>.

- [11] Nakamura, Y., Pashkin, Y. A. & Tsai, J. S. Coherent control of macroscopic quantum states in a single-Cooper-pair box. *Nature* **398**, 786–788 (1999). URL <https://doi.org/10.1038/19718>.
- [12] Cirac, J. I. & Zoller, P. Quantum Computations with Cold Trapped Ions. *Phys. Rev. Lett.* **74**, 4091–4094 (1995). URL <https://link.aps.org/doi/10.1103/PhysRevLett.74.4091>.
- [13] H.-J. Briegel and T. Calarco and D. Jaksch and J. I. Cirac and P. Zoller. Quantum computing with neutral atoms. *Journal of Modern Optics* **47**, 415–451 (2000). URL <https://www.tandfonline.com/doi/abs/10.1080/09500340008244052>.
- [14] Barends, R. *et al.* Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature* **508**, 500–503 (2014). URL <https://doi.org/10.1038/nature13171>.
- [15] Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018). URL <https://doi.org/10.22331/q-2018-08-06-79>.
- [16] Ball, P. First quantum computer to pack 100 qubits enters crowded race. *Nature* **599**, 542 (2021). URL <https://doi.org/10.1038/d41586-021-03476-5>.
- [17] Acharya, R. *et al.* Quantum error correction below the surface code threshold. *Nature* **638**, 920–926 (2025). URL <https://doi.org/10.1038/s41586-024-08449-y>.
- [18] Wurtz, J. *et al.* Aquila: QuEra’s 256-qubit neutral-atom quantum computer (2023). URL <https://arxiv.org/abs/2306.11727>.
- [19] Chen, J.-S. *et al.* Benchmarking a trapped-ion quantum computer with 30 qubits. *Quantum* **8**, 1516 (2024). URL <https://doi.org/10.22331/q-2024-11-07-1516>.
- [20] Ransford, A. *et al.* Helios: A 98-qubit trapped-ion quantum computer (2025). URL <https://arxiv.org/abs/2511.05465>.
- [21] Aaronson, S. & Arkhipov, A. The computational complexity of linear optics. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, 333–342 (2011). URL <https://doi.org/10.1145/1993636.1993682>.
- [22] Aaronson, S. & Chen, L. Complexity-theoretic foundations of quantum supremacy experiments (2016). URL <https://doi.org/10.48550/arXiv.1612.05903>.
- [23] Bremner, M. J., Montanaro, A. & Shepherd, D. J. Achieving quantum supremacy with sparse and noisy commuting quantum computations. *Quantum* **1**, 8 (2017). URL <https://doi.org/10.22331/q-2017-04-25-8>.
- [24] Arute, F. *et al.* Quantum supremacy using a programmable superconducting processor. *Nature* **574**, 505–510 (2019). URL <https://doi.org/10.1038/s41586-019-1666-5>.

- [25] Han-Sen, Z. *et al.* Quantum computational advantage using photons. *Science* **370**, 1460–1463 (2020). URL <https://www.science.org/doi/abs/10.1126/science.abe8770>.
- [26] Madsen, L. S. *et al.* Quantum computational advantage with a programmable photonic processor. *Nature* **606**, 75–81 (2022). URL <https://doi.org/10.1038/s41586-022-04725-x>.
- [27] Morvan, A. *et al.* Phase transitions in random circuit sampling. *Nature* **634**, 328–333 (2024). URL <https://doi.org/10.1038/s41586-024-07998-6>.
- [28] Cerezo, M. *et al.* Variational quantum algorithms. *Nat. Rev. Phys.* **3**, 625–644 (2021). URL <https://doi.org/10.1038/s42254-021-00348-9>.
- [29] Peruzzo, A. *et al.* A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.* **5**, 4213 (2014). URL <https://doi.org/10.1038/ncomms5213>.
- [30] Farhi, E., Goldstone, J. & Gutmann, S. A quantum approximate optimization algorithm (2014). URL <https://doi.org/10.48550/arXiv.1411.4028>.
- [31] Logical quantum processor based on reconfigurable atom arrays. *Nature* **626**, 58–65 (2024). URL <https://doi.org/10.1038/s41586-023-06927-3>.
- [32] Reichardt, B. W. *et al.* Fault-tolerant quantum computation with a neutral atom processor (2025). URL <https://arxiv.org/abs/2411.11822>.
- [33] Knill, E. Fault-Tolerant Postselected Quantum Computation: Schemes (2004). URL <https://arxiv.org/abs/quant-ph/0402171>.
- [34] Gidney, C. How to factor 2048 bit RSA integers with less than a million noisy qubits (2025). URL <https://arxiv.org/abs/2505.15917>.
- [35] Gidney, C., Shutty, N. & Jones, C. Magic state cultivation: growing T states as cheap as CNOT gates (2024). URL <https://arxiv.org/abs/2409.17595>.
- [36] Itogawa, T., Takada, Y., Hirano, Y. & Fujii, K. Efficient Magic State Distillation by Zero-Level Distillation. *PRX Quantum* **6**, 020356 (2025). URL <https://link.aps.org/doi/10.1103/thxx-njr6>.
- [37] McClean, J. R., Boixo, S., Smelyanskiy, V. N., Babbush, R. & Neven, H. Barren plateaus in quantum neural network training landscapes. *Nat. Commun.* **9**, 4812 (2018). URL <https://doi.org/10.1038/s41467-018-07090-4>.
- [38] Tomislav, B., Johnnie, G. & Garnet, K.-L. C. Fast and converged classical simulations of evidence for the utility of quantum computing before fault tolerance. *Science Advances* **10**, eadk4321 (2024). URL <https://www.science.org/doi/abs/10.1126/sciadv.adk4321>.

- [39] Mansky, M. B. *et al.* Sampling problems on a Quantum Computer. In *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 01, 485–495 (2023). URL <https://doi.org/10.1109/QCE57702.2023.00062>.
- [40] Jiang, T. *et al.* Walking through Hilbert Space with Quantum Computers. *Chemical Reviews* **125**, 4569–4602 (2025). URL <https://doi.org/10.1021/acs.chemrev.4c00508>.
- [41] Kanno, K. *et al.* Quantum-Selected Configuration Interaction: classical diagonalization of Hamiltonians in subspaces selected by quantum computers (2023). URL <https://arxiv.org/abs/2302.11320>.
- [42] Robledo-Moreno, J. *et al.* Chemistry Beyond Exact Solutions on a Quantum-Centric Supercomputer (2024). URL <https://arxiv.org/abs/2405.05068>.
- [43] Layden, D. *et al.* Quantum-enhanced Markov chain Monte Carlo. *Nature* **619**, 282–287 (2023). URL <https://doi.org/10.1038/s41586-023-06095-4>.
- [44] Christmann, J., Ivashkov, P., Chiurco, M. & Mazzola, G. From quantum-enhanced to quantum-inspired Monte Carlo. *Phys. Rev. A* **111**, 042615 (2025). URL <https://link.aps.org/doi/10.1103/PhysRevA.111.042615>.
- [45] Orfi, A. & Sels, D. Bounding the speedup of the quantum-enhanced Markov-chain Monte Carlo algorithm. *Phys. Rev. A* **110**, 052414 (2024). URL <https://link.aps.org/doi/10.1103/PhysRevA.110.052414>.
- [46] Ising, E. Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik* **31**, 253–258 (1925). URL <https://doi.org/10.1007/BF02980577>.
- [47] BRUSH, S. G. History of the Lenz-Ising Model. *Rev. Mod. Phys.* **39**, 883–893 (1967). URL <https://link.aps.org/doi/10.1103/RevModPhys.39.883>.
- [48] Sherrington, D. & Kirkpatrick, S. Solvable Model of a Spin-Glass. *Phys. Rev. Lett.* **35**, 1792–1796 (1975). URL <https://link.aps.org/doi/10.1103/PhysRevLett.35.1792>.
- [49] Parisi, G. *Field Theory, Disorder and Simulations* (WORLD SCIENTIFIC, 1992). URL <https://www.worldscientific.com/doi/abs/10.1142/1655>.
- [50] Binder, K. & Young, A. P. Spin glasses: Experimental facts, theoretical concepts, and open questions. *Rev. Mod. Phys.* **58**, 801–976 (1986). URL <https://link.aps.org/doi/10.1103/RevModPhys.58.801>.
- [51] Hopfield, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences* **79**, 2554–2558 (1982). URL <https://www.pnas.org/doi/abs/10.1073/pnas.79.8.2554>.

- [52] Ackley, D. H., Hinton, G. E. & Sejnowski, T. J. A learning algorithm for Boltzmann machines. *Cogn. Sci.* **9**, 147–169 (1985). URL <https://www.sciencedirect.com/science/article/abs/pii/S0364021385800124>.
- [53] Kochenberger, G. *et al.* The unconstrained binary quadratic programming problem: a survey. *Journal of Combinatorial Optimization* **28**, 58–81 (2014). URL <https://doi.org/10.1007/s10878-014-9734-0>.
- [54] Lucas, A. Ising formulations of many NP problems. *Front. Phys.* **2**, 5 (2014). URL <https://www.frontiersin.org/journals/physics/articles/10.3389/fphy.2014.00005>.
- [55] Barahona, F. On the computational complexity of Ising spin glass models. *Journal of Physics A: Mathematical and General* **15**, 3241 (1982). URL <https://doi.org/10.1088/0305-4470/15/10/028>.
- [56] Kirkpatrick, S., Gelatt, C. D. & Vecchi, M. P. Optimization by Simulated Annealing. *Science* **220**, 671–680 (1983). URL <https://www.science.org/doi/abs/10.1126/science.220.4598.671>.
- [57] Wang, W., Machta, J. & Katzgraber, H. G. Comparing Monte Carlo methods for finding ground states of Ising spin glasses: Population annealing, simulated annealing, and parallel tempering. *Phys. Rev. E* **92**, 013303 (2015). URL <https://link.aps.org/doi/10.1103/PhysRevE.92.013303>.
- [58] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21**, 1087–1092 (1953). URL <https://doi.org/10.1063/1.1699114>.
- [59] Hastings, W. K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**, 97–109 (1970). URL <https://doi.org/10.1093/biomet/57.1.97>.
- [60] Swendsen, R. H. & Wang, J.-S. Nonuniversal critical dynamics in Monte Carlo simulations. *Phys. Rev. Lett.* **58**, 86 (1987). URL <https://doi.org/10.1103/PhysRevLett.58.86>.
- [61] Wolff, U. Collective Monte Carlo updating for spin systems. *Phys. Rev. Lett.* **62**, 361 (1989). URL <https://doi.org/10.1103/PhysRevLett.62.361>.
- [62] Swendsen, R. H. & Wang, J.-S. Replica Monte Carlo Simulation of Spin-Glasses. *Phys. Rev. Lett.* **57**, 2607–2609 (1986). URL <https://link.aps.org/doi/10.1103/PhysRevLett.57.2607>.
- [63] Hukushima, K. & Nemoto, K. Exchange Monte Carlo Method and Application to Spin Glass Simulations. *J. Phys. Soc. Jpn.* **65**, 1604–1608 (1996). URL <https://doi.org/10.1143/JPSJ.65.1604>.

- [64] Houdayer, J. A cluster Monte Carlo algorithm for 2-dimensional spin glasses. *Eur. Phys. J. B* **22**, 479–484 (2001). URL <https://doi.org/10.1007/PL00011151>.
- [65] Zhu, Z., Ochoa, A. J. & Katzgraber, H. G. Efficient cluster algorithm for spin glasses in any space dimension. *Phys. Rev. Lett.* **115**, 077201 (2015). URL <https://doi.org/10.1103/PhysRevLett.115.077201>.
- [66] Liu, J., Qi, Y., Meng, Z. Y. & Fu, L. Self-learning Monte Carlo method. *Phys. Rev. B* **95**, 041101 (2017). URL <https://link.aps.org/doi/10.1103/PhysRevB.95.041101>.
- [67] Peskun, P. H. Optimum Monte-carlo sampling using Markov chains. *Biometrika* **60**, 607–612 (1973). URL <https://doi.org/10.1093/biomet/60.3.607>.
- [68] Frigessi, A., Hwang, C.-R. & Younes, L. Optimal spectral structure of reversible stochastic matrices, Monte Carlo methods and the simulation of Markov random fields. *Ann. Appl. Probab.* **2**, 610–628 (1992). URL <https://doi.org/10.1214/aoap/1177005652>.
- [69] Levin, D. A. & Peres, Y. *Markov chains and mixing times* (American Mathematical Soc., 2017), 2nd revised edn. URL <https://doi.org/10.1090/mbk/058>.
- [70] Dawson, C. M. & Nielsen, M. A. The Solovay-Kitaev algorithm. *Quantum Info. Comput.* **6**, 81–95 (2006). URL <https://arxiv.org/abs/quant-ph/0505030>.
- [71] Quantum annealing: A new method for minimizing multidimensional functions. *Chemical Physics Letters* **219**, 343–348 (1994). URL <https://www.sciencedirect.com/science/article/pii/0009261494001170>.
- [72] Kadowaki, T. & Nishimori, H. Quantum annealing in the transverse Ising model. *Phys. Rev. E* **58**, 5355 (1998). URL <https://doi.org/10.1103/PhysRevE.58.5355>.
- [73] Farhi, E., Goldstone, J., Gutmann, S. & Sipser, M. Quantum Computation by Adiabatic Evolution (2000). URL <https://arxiv.org/abs/quant-ph/0001106>.
- [74] Aharonov, D. *et al.* Adiabatic Quantum Computation Is Equivalent to Standard Quantum Computation. *SIAM Review* **50**, 755–787 (2008). URL <https://doi.org/10.1137/080734479>.
- [75] Kandala, A. *et al.* Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature* **549**, 242–246 (2017). URL <https://doi.org/10.1038/nature23879>.
- [76] Cerezo, M., Sone, A., Volkoff, T., Cincio, L. & Coles, P. J. Cost function dependent barren plateaus in shallow parametrized quantum circuits. *Nature Communications* **12**, 1791 (2021). URL <https://doi.org/10.1038/s41467-021-21728-w>.
- [77] Hadfield, S. *et al.* From the Quantum Approximate Optimization Algorithm to a Quantum Alternating Operator Ansatz. *Algorithms* **12** (2019). URL <https://www.mdpi.com/1999-4893/12/2/34>.

- [78] Sundar, B., Paredes, R., Damanik, D. T., Dueñas-Osorio, L. & Hazzard, K. R. A. A quantum algorithm to count weighted ground states of classical spin Hamiltonians (2019). URL <https://arxiv.org/abs/1908.01745>.
- [79] Bärtschi, A. & Eidenbenz, S. Grover Mixers for QAOA: Shifting Complexity from Mixer Design to State Preparation. In *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*, 72–82 (2020). URL <https://doi.org/10.1109/QCE49297.2020.00020>.
- [80] Zhu, Y. *et al.* Multi-round QAOA and advanced mixers on a trapped-ion quantum computer. *Quantum Science and Technology* **8**, 015007 (2022). URL <https://doi.org/10.1088/2058-9565/ac91ef>.
- [81] Zhang, Z. *et al.* Grover-QAOA for 3-SAT: quadratic speedup, fair-sampling, and parameter clustering. *Quantum Science and Technology* **10**, 015022 (2024). URL <https://doi.org/10.1088/2058-9565/ad895c>.
- [82] Sakai, R., Matsuyama, H., Tam, W.-H., Yamashiro, Y. & Fujii, K. Linearly simplified QAOA parameters and transferability (2024). URL <https://arxiv.org/abs/2405.00655>.
- [83] Ichikawa, K. *et al.* Optimal elemental configuration search in crystal using quantum approximate optimization algorithm (2025). URL <https://arxiv.org/abs/2503.09356>.
- [84] Montañez-Barrera, J. A. & Michielsen, K. Toward a linear-ramp QAOA protocol: evidence of a scaling advantage in solving some combinatorial optimization problems. *npj Quantum Information* **11**, 131 (2025). URL <https://doi.org/10.1038/s41534-025-01082-1>.
- [85] Brandao, F. G. S. L., Broughton, M., Farhi, E., Gutmann, S. & Neven, H. For Fixed Control Parameters the Quantum Approximate Optimization Algorithm’s Objective Function Value Concentrates for Typical Instances (2018). URL <https://arxiv.org/abs/1812.04170>.
- [86] Galda, A., Liu, X., Lykov, D., Alexeev, Y. & Safro, I. Transferability of optimal QAOA parameters between random graphs. In *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)*, 171–180 (Broomfield, CO, USA, 2021). URL <https://doi.org/10.1109/QCE52317.2021.00034>.
- [87] Wurtz, J. & Lykov, D. Fixed-angle conjectures for the quantum approximate optimization algorithm on regular MaxCut graphs. *Phys. Rev. A* **104**, 052419 (2021). URL <https://link.aps.org/doi/10.1103/PhysRevA.104.052419>.
- [88] Shaydulin, R., Lotshaw, P. C., Larson, J., Ostrowski, J. & Humble, T. S. Parameter Transfer for Quantum Approximate Optimization of Weighted MaxCut. *ACM Trans. Quantum Comput.* **4** (2023). URL <https://doi.org/10.1145/3584706>.

- [89] Terhal, B. M. & DiVincenzo, D. P. Problem of equilibration and the computation of correlation functions on a quantum computer. *Phys. Rev. A* **61**, 022301 (2000). URL <https://link.aps.org/doi/10.1103/PhysRevA.61.022301>.
- [90] Szegedy, M. Quantum speed-up of Markov chain based algorithms. In *45th Annual IEEE symposium on foundations of computer science*, 32–41 (IEEE, 2004). URL <https://doi.org/10.1109/FOCS.2004.53>.
- [91] Temme, K., Osborne, T. J., Vollbrecht, K. G., Poulin, D. & Verstraete, F. Quantum Metropolis sampling. *Nature* **471**, 87–90 (2011). URL <https://doi.org/10.1038/nature09770>.
- [92] Yung, M.-H. & Aspuru-Guzik, A. A quantum–quantum Metropolis algorithm. *Proceedings of the National Academy of Sciences* **109**, 754–759 (2012). URL <https://www.pnas.org/doi/abs/10.1073/pnas.1111758109>.
- [93] Wocjan, P. & Temme, K. Szegedy Walk Unitaries for Quantum Maps. *Communications in Mathematical Physics* **402**, 3201–3231 (2023). URL <https://doi.org/10.1007/s00220-023-04797-4>.
- [94] Chen, C.-F., Kastoryano, M. J., Brandão, F. G. S. L. & Gilyén, A. Quantum Thermal State Preparation (2023). URL <https://arxiv.org/abs/2303.18224>.
- [95] Bergamaschi, T., Chen, C.-F. & Liu, Y. Quantum Computational Advantage with Constant-Temperature Gibbs Sampling. In *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, 1063–1085 (2024). URL <https://doi.org/10.1109/FOCS61266.2024.00071>.
- [96] Rouzé, C., França, D. S. & Álvaro M. Alhambra. Optimal quantum algorithm for Gibbs state preparation (2024). URL <https://arxiv.org/abs/2411.04885>.
- [97] Chen, C.-F., Kastoryano, M., Brandão, F. G. S. L. & Gilyén, A. Efficient quantum thermal simulation. *Nature* **646**, 561–566 (2025). URL <https://doi.org/10.1038/s41586-025-09583-x>.
- [98] Mitarai, K., Negoro, M., Kitagawa, M. & Fujii, K. Quantum circuit learning. *Phys. Rev. A* **98**, 032309 (2018). URL <https://doi.org/10.1103/PhysRevA.98.032309>.
- [99] Farhi, E. & Neven, H. Classification with quantum neural networks on near term processors (2018). URL <https://doi.org/10.48550/arXiv.1802.06002>.
- [100] Kübler, J. M., Arrasmith, A., Cincio, L. & Coles, P. J. An adaptive optimizer for measurement-frugal variational algorithms. *Quantum* **4**, 263 (2020). URL <https://doi.org/10.22331/q-2020-05-11-263>.
- [101] Gonthier, J. F. *et al.* Measurements as a roadblock to near-term practical quantum advantage in chemistry: Resource analysis. *Phys. Rev. Res.* **4**, 033154 (2022). URL <https://link.aps.org/doi/10.1103/PhysRevResearch.4.033154>.

- [102] Ito, K. Latency-aware adaptive shot allocation for run-time efficient variational quantum algorithms (2023). URL <https://doi.org/10.48550/arXiv.2302.04422>.
- [103] Hangleiter, D. & Eisert, J. Computational advantage of quantum random sampling. *Rev. Mod. Phys.* **95**, 035001 (2023). URL <https://doi.org/10.1103/RevModPhys.95.035001>.
- [104] Landau, D. & Binder, K. *A guide to Monte Carlo simulations in statistical physics* (Cambridge university press, 2021). URL <https://doi.org/10.1017/CB09781139696463>.
- [105] Andrieu, C., De Freitas, N., Doucet, A. & Jordan, M. I. An introduction to MCMC for machine learning. *Mach. Learn.* **50**, 5–43 (2003). URL <https://doi.org/10.1023/A:1020281327116>.
- [106] Suzuki, M. Generalized Trotter’s formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems. *Commun. Math. Phys.* **51**, 183–190 (1976). URL <https://doi.org/10.1007/BF01609348>.
- [107] Gelman, A., Gilks, W. R. & Roberts, G. O. Weak convergence and optimal scaling of random walk Metropolis algorithms. *Ann. Appl. Probab.* **7**, 110–120 (1997). URL <https://doi.org/10.1214/aoap/1034625254>.
- [108] Neklyudov, K., Egorov, E., Shvechikov, P. & Vetrov, D. Metropolis-Hastings view on variational inference and adversarial training (2018). URL <https://doi.org/10.48550/arXiv.1810.07151>.
- [109] Brent, R. P. *Algorithms for minimization without derivatives* (Courier Corporation, 2013).
- [110] Dür, W., Hein, M., Cirac, J. I. & Briegel, H.-J. Standard forms of noisy quantum operations via depolarization. *Phys. Rev. A* **72**, 052326 (2005). URL <https://link.aps.org/doi/10.1103/PhysRevA.72.052326>.
- [111] Magesan, E., Gambetta, J. M. & Emerson, J. Scalable and Robust Randomized Benchmarking of Quantum Processes. *Phys. Rev. Lett.* **106**, 180504 (2011). URL <https://link.aps.org/doi/10.1103/PhysRevLett.106.180504>.
- [112] Geller, M. R. & Zhou, Z. Efficient error models for fault-tolerant architectures and the Pauli twirling approximation. *Phys. Rev. A* **88**, 012314 (2013). URL <https://link.aps.org/doi/10.1103/PhysRevA.88.012314>.
- [113] Abrams, D. S. & Lloyd, S. Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors. *Phys. Rev. Lett.* **83**, 5162 (1999). URL <https://doi.org/10.1103/PhysRevLett.83.5162>.
- [114] Kim, Y. *et al.* Evidence for the utility of quantum computing before fault tolerance. *Nature* **618**, 500–505 (2023). URL <https://doi.org/10.1038/s41586-023-06096-3>.

- [115] Kechedzhi, K. *et al.* Effective quantum volume, fidelity and computational cost of noisy quantum processing experiments. *Future Gener. Comput. Syst.* **153**, 431–441 (2024). URL <https://www.sciencedirect.com/science/article/pii/S0167739X23004569>.
- [116] Nakano, Y., Hakoshima, H., Mitarai, K. & Fujii, K. Markov-chain Monte Carlo method enhanced by a quantum alternating operator ansatz. *Phys. Rev. Res.* **6**, 033105 (2024). URL <https://link.aps.org/doi/10.1103/PhysRevResearch.6.033105>.
- [117] Larochelle, H. & Murray, I. The Neural Autoregressive Distribution Estimator. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, vol. 15, 29–37 (PMLR, Fort Lauderdale, FL, USA, 2011). URL <https://proceedings.mlr.press/v15/larochelle11a.html>.
- [118] Uria, B., Côté, M.-A., Gregor, K., Murray, I. & Larochelle, H. Neural Autoregressive Distribution Estimation. *J. Mach. Learn. Res.* **17**, 1–37 (2016). URL <http://jmlr.org/papers/v17/16-272.html>.
- [119] Germain, M., Gregor, K., Murray, I. & Larochelle, H. MADE: Masked Autoencoder for Distribution Estimation. In *Proceedings of the 32nd International Conference on Machine Learning*, vol. 37, 881–889 (PMLR, Lille, France, 2015). URL <https://proceedings.mlr.press/v37/germain15.html>.
- [120] Lotshaw, P. C. *et al.* Approximate Boltzmann distributions in quantum approximate optimization. *Phys. Rev. A* **108**, 042411 (2023). URL <https://link.aps.org/doi/10.1103/PhysRevA.108.042411>.
- [121] Díez-Valle, P., Porras, D. & García-Ripoll, J. J. Quantum Approximate Optimization Algorithm Pseudo-Boltzmann States. *Phys. Rev. Lett.* **130**, 050601 (2023). URL <https://link.aps.org/doi/10.1103/PhysRevLett.130.050601>.
- [122] Leontica, S. & Amaro, D. Exploring the neighborhood of 1-layer QAOA with instantaneous quantum polynomial circuits. *Phys. Rev. Res.* **6**, 013071 (2024). URL <https://link.aps.org/doi/10.1103/PhysRevResearch.6.013071>.
- [123] Farhi, E. & Harrow, A. W. Quantum Supremacy through the Quantum Approximate Optimization Algorithm (2019). URL <https://arxiv.org/abs/1602.07674>.
- [124] Ackley, D. H., Hinton, G. E. & Sejnowski, T. J. A learning algorithm for boltzmann machines. *Cognit. Sci.* **9**, 147–169 (1985). URL <https://www.sciencedirect.com/science/article/pii/S0364021385800124>.
- [125] Smolensky, P. Information Processing in Dynamical Systems: Foundations of Harmony Theory. In *Parallel Distributed Processing, Volume 1: Explorations in the Microstructure of Cognition: Foundations* (The MIT Press, 1986). URL <https://doi.org/10.7551/mitpress/5236.003.0009>.

- [126] Hinton, G. E. Training Products of Experts by Minimizing Contrastive Divergence. *Neural Comput.* **14**, 1771–1800 (2002). URL <https://doi.org/10.1162/089976602760128018>.
- [127] Hinton, G. E. & Salakhutdinov, R. R. Reducing the Dimensionality of Data with Neural Networks. *Science* **313**, 504–507 (2006). URL <https://www.science.org/doi/abs/10.1126/science.1127647>.
- [128] Kingma, D. P. & Ba, J. Adam: A Method for Stochastic Optimization (2017). URL <https://arxiv.org/abs/1412.6980>.
- [129] Nicoli, K. A. *et al.* Asymptotically unbiased estimation of physical observables with neural samplers. *Phys. Rev. E* **101**, 023304 (2020). URL <https://link.aps.org/doi/10.1103/PhysRevE.101.023304>.
- [130] McNaughton, B., Milošević, M. V., Perali, A. & Pilati, S. Boosting Monte Carlo simulations of spin glasses using autoregressive neural networks. *Phys. Rev. E* **101**, 053312 (2020). URL <https://link.aps.org/doi/10.1103/PhysRevE.101.053312>.
- [131] Wu, D., Rossi, R. & Carleo, G. Unbiased Monte Carlo cluster updates with autoregressive neural networks. *Phys. Rev. Res.* **3**, L042024 (2021). URL <https://link.aps.org/doi/10.1103/PhysRevResearch.3.L042024>.
- [132] Hibat-Allah, M., Inack, E. M., Wiersema, R., Melko, R. G. & Carrasquilla, J. Variational neural annealing. *Nat. Mach. Intell.* **3**, 952–961 (2021). URL <https://doi.org/10.1038/s42256-021-00401-3>.
- [133] Ciarella, S., Trinquier, J., Weigt, M. & Zamponi, F. Machine-learning-assisted Monte Carlo fails at sampling computationally hard problems. *Mach. Learn.: Sci. Technol.* **4**, 010501 (2023). URL <https://dx.doi.org/10.1088/2632-2153/acbe91>.
- [134] Inack, E. M., Morawetz, S. & Melko, R. G. Neural annealing and visualization of autoregressive neural networks in the Newman-Moore model (2022). URL <https://arxiv.org/abs/2204.11272>.
- [135] Arai, S. & Kadowaki, T. Quantum annealing enhanced Markov-Chain Monte Carlo. *Sci. Rep.* **15**, 21427 (2025). URL <https://doi.org/10.1038/s41598-025-07293-y>.
- [136] Mazzola, G. Sampling, rates, and reaction currents through reverse stochastic quantization on quantum computers. *Phys. Rev. A* **104**, 022431 (2021). URL <https://doi.org/10.1103/PhysRevA.104.022431>.
- [137] Scriva, G., Costa, E., McNaughton, B. & Pilati, S. Accelerating equilibrium spin-glass simulations using quantum annealers via generative deep learning. *SciPost Phys.* **15**, 018 (2023). URL <https://scipost.org/10.21468/SciPostPhys.15.1.018>.
- [138] King, A. D. *et al.* Quantum critical dynamics in a 5,000-qubit programmable spin glass. *Nature* **617**, 61–66 (2023). URL <https://doi.org/10.1038/s41586-023-05867-2>.

- [139] Könz, M. S., Lechner, W., Katzgraber, H. G. & Troyer, M. Embedding Overhead Scaling of Optimization Problems in Quantum Annealing. *PRX Quantum* **2**, 040322 (2021). URL <https://link.aps.org/doi/10.1103/PRXQuantum.2.040322>.
- [140] Basso, J., Farhi, E., Marwaha, K., Villalonga, B. & Zhou, L. The Quantum Approximate Optimization Algorithm at High Depth for MaxCut on Large-Girth Regular Graphs and the Sherrington-Kirkpatrick Model. In *17th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2022)*, vol. 232, 7:1–7:21 (Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2022). URL <https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.TQC.2022.7>.
- [141] Vaswani, A. *et al.* Attention is All you Need. In *Advances in Neural Information Processing Systems*, vol. 30 (Curran Associates, Inc., 2017). URL https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.
- [142] Nagai, Y. & Tomiya, A. Self-learning Monte Carlo with equivariant transformer. *J. Phys. Soc. Jpn.* **93**, 114007 (2024). URL <https://doi.org/10.7566/JPSJ.93.114007>.
- [143] Biazzo, I. The autoregressive neural network architecture of the Boltzmann distribution of pairwise interacting spins systems. *Commun. Phys.* **6**, 296 (2023). URL <https://doi.org/10.1038/s42005-023-01416-5>.
- [144] Korte, B. & Vygen, J. *Approximation Algorithms*, 423–469 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2018). URL https://doi.org/10.1007/978-3-662-56039-6_16.
- [145] Williamson, D. P. & Shmoys, D. B. *The Design of Approximation Algorithms* (Cambridge University Press, 2011).
- [146] Boulebnane, S. & Montanaro, A. Solving Boolean Satisfiability Problems With The Quantum Approximate Optimization Algorithm. *PRX Quantum* **5**, 030348 (2024). URL <https://link.aps.org/doi/10.1103/PRXQuantum.5.030348>.
- [147] Munoz-Bauza, H. & Lidar, D. Scaling Advantage in Approximate Optimization with Quantum Annealing. *Phys. Rev. Lett.* **134**, 160601 (2025). URL <https://link.aps.org/doi/10.1103/PhysRevLett.134.160601>.
- [148] Valiant, L. The complexity of computing the permanent. *Theoretical Computer Science* **8**, 189–201 (1979). URL <https://www.sciencedirect.com/science/article/pii/0304397579900446>.
- [149] Gopalan, P. *et al.* An FPTAS for #Knapsack and Related Counting Problems. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, 817–826 (2011). URL <https://doi.org/10.1109/FOCS.2011.32>.

- [150] Weaver, S. A., Ray, K. J., Marek, V. W., Mayer, A. J. & Walker, A. K. Satisfiability-based set membership filters. *journal on Satisfiability, Boolean Modeling and Computation* **8**, 129–148 (2012). URL <https://doi.org/10.3233/SAT190095>.
- [151] Fang, C., Zhu, Z. & Katzgraber, H. G. NAE-SAT-based probabilistic membership filters (2018). URL <https://arxiv.org/abs/1801.06232>.
- [152] Matsuda, Y., Nishimori, H. & Katzgraber, H. G. Quantum annealing for problems with ground-state degeneracy. *journal of Physics: Conference Series* **143**, 012003 (2009). URL <https://doi.org/10.1088/1742-6596/143/1/012003>.
- [153] Mandrà, S., Zhu, Z. & Katzgraber, H. G. Exponentially Biased Ground-State Sampling of Quantum Annealing Machines with Transverse-Field Driving Hamiltonians. *Phys. Rev. Lett.* **118**, 070502 (2017). URL <https://link.aps.org/doi/10.1103/PhysRevLett.118.070502>.
- [154] Könz, M. S., Mazzola, G., Ochoa, A. J., Katzgraber, H. G. & Troyer, M. Uncertain fate of fair sampling in quantum annealing. *Phys. Rev. A* **100**, 030303 (2019). URL <https://link.aps.org/doi/10.1103/PhysRevA.100.030303>.
- [155] Pelofske, E. Biased degenerate ground-state sampling of small Ising models with converged quantum approximate optimization algorithm. *Phys. Rev. E* **111**, 054103 (2025). URL <https://link.aps.org/doi/10.1103/PhysRevE.111.054103>.
- [156] Nakano, Y., Okada, K. N. & Fujii, K. Neural-network-assisted Monte Carlo sampling trained by Quantum Approximate Optimization Algorithm (2025). URL <https://arxiv.org/abs/2506.01335>.
- [157] Wei, W., Erenrich, J. & Selman, B. Towards efficient sampling: exploiting random walk strategies. In *Proceedings of the 19th National Conference on Artificial Intelligence, AAAI'04*, 670–676 (AAAI Press, 2004). URL <https://dl.acm.org/doi/10.5555/1597148.1597256>.
- [158] Byrd, R. H., Lu, P., Nocedal, J. & Zhu, C. A limited memory algorithm for bound constrained optimization. *SIAM J. Sci. Comput.* **16**, 1190–1208 (1995). URL <https://doi.org/10.1137/0916069>.
- [159] Suzuki, Y. *et al.* Qulacs: a fast and versatile quantum circuit simulator for research purpose. *Quantum* **5**, 559 (2021). URL <https://doi.org/10.22331/q-2021-10-06-559>.
- [160] Paszke, A. *et al.* PyTorch: An Imperative Style, High-Performance Deep Learning Library. In *Advances in Neural Information Processing Systems*, vol. 32 (Curran Associates, Inc., 2019). URL https://proceedings.neurips.cc/paper_files/paper/2019/file/bdbca288fee7f92f2bfa9f7012727740-Paper.pdf.
- [161] Davis, M., Logemann, G. & Loveland, D. A machine program for theorem-proving. *Commun. ACM* **5**, 394–397 (1962). URL <https://doi.org/10.1145/368273.368557>.

-
- [162] Cai, S., Luo, C. & Su, K. Improving WalkSAT By Effective Tie-Breaking and Efficient Implementation. *The Computer journal* **58**, 2864–2875 (2014). URL <https://doi.org/10.1093/comjnl/bxu135>.
- [163] Pelofske, E., Bärtschi, A. & Eidenbenz, S. Quantum Annealing vs. QAOA: 127 Qubit Higher-Order Ising Problems on NISQ Computers. In Bhatele, A., Hammond, J., Baboulin, M. & Kruse, C. (eds.) *High Performance Computing*, 240–258 (Springer Nature Switzerland, Cham, 2023). URL https://doi.org/10.1007/978-3-031-32041-5_13.
- [164] Sachdeva, N. *et al.* Quantum optimization using a 127-qubit gate-model IBM quantum computer can outperform quantum annealers for nontrivial binary optimization problems (2024). URL <https://arxiv.org/abs/2406.01743>.
- [165] Ferguson, S. & Wallden, P. Quantum-enhanced Markov chain Monte Carlo for systems larger than a quantum computer. *Phys. Rev. Res.* **7**, 013231 (2025). URL <https://link.aps.org/doi/10.1103/PhysRevResearch.7.013231>.

Acknowledgments

I would like to express my sincere gratitude to all those who supported and guided me in completing this thesis.

First and foremost, I am deeply grateful to my supervisor, Prof. Keisuke Fujii, for your continuous guidance from the beginning of my research to the completion of this thesis. In addition to your careful supervision, You provided invaluable advice on how to conduct research, formulate problems, and interpret results. Through many discussions with you, I also learned the attitude and mindset required of a researcher. Although I started my Ph.D studies with little background in quantum computing, I was able to learn a great deal and complete this thesis thanks entirely to your support. I sincerely thank you.

I would also like to thank Prof. Takashi Yamamoto, Prof. Takeo Minamikawa, and Prof. Kosuke Mitarai for serving on my thesis committee and for providing valuable comments during the review process. Your feedback was greatly helpful in organizing the content of this thesis and clarifying it. I am also grateful to my collaborators, Prof. Hideaki Hakoshima and Prof. Ken N. Okada, for your substantial support through theoretical advice and fruitful discussions. I would further like to thank Prof. Yasuyuki Kato for helpful discussions related to this work.

I would like to thank all the members of Fujii Lab. for supporting me through daily discussions and by maintaining a productive research environment.

This work was supported by Grant-in-Aid for JSPS Fellows Grant No. 24KJ1606.

Finally, I would like to thank my family for their long-standing support and understanding. Their encouragement has been a great source of strength throughout my research.

List of Achievements

Publications included in this thesis

1. **Y. Nakano**, H. Hakoshima, K. Mitarai, and K. Fujii, “Markov-chain Monte Carlo method enhanced by a quantum alternating operator ansatz”, *Phys. Rev. Research* **6**, 033105 (2024).
2. **Y. Nakano**, K. N. Okada, and K. Fujii, “Neural-network-assisted Monte Carlo sampling trained by Quantum Approximate Optimization Algorithm”, *PRX Quantum* (2026) (in press).
3. **Y. Nakano**, and K. Fujii, “Fair sampling of ground-state configurations using hybrid quantum-classical MCMC algorithms”, *arXiv:2512.14552* (2025).

Publications not included in this thesis

4. R. Rahimi, A. SaiToh, A. Modiri, **Y. Nakano**, K. N. Okada, S. Tsukano, B. Zhang, K. Fujii, M. Kitagawa, and A. Sawant, “Quantum Computing for Radiation Therapy Optimization”, *Medical Physics*, **53**, 1 (2026).
5. D. Arai, K. N. Okada, **Y. Nakano**, K. Mitarai, and K. Fujii, “Scalable circuit depth reduction in feedback-based quantum optimization with a quadratic approximation” *Phys. Rev. Research* **7**, 013035 (2025).
6. T. Yoshioka, K. Sasada, **Y. Nakano** and K. Fujii, “Electric Power Demand Portfolio Optimization by Fermionic QAOA with Self-Consistent Local Field Modulation,” *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)* (2024).
7. T. Matsumoto, S. Sato, S. Akei, **Y. Nakano**, S. Iba, J. Ishihara, K. Miyamoto, N. Yokoshi, T. Omatsu, and K. Morita, “Coherent transfer of the higher-order polarization state of photons to the spin structure state of electrons in a semiconductor”, *Optica Quantum* **2**, 245–253 (2024).
8. T. Ichikawa, H. Hakoshima, K. Inui, K. Ito, R. Matsuda, K. Mitarai, K. Miyamoto, W. Mizukami, K. Mizuta, T. Mori, **Y. Nakano**, A. Nakayama, K. N. Okada, T. Sugimoto, S. Takahira, N. Takemori, S. Tsukano, H. Ueda, R. Watanabe, Y. Yoshida, and K. Fujii, “Current numbers of qubits and their uses”, *Nat Rev Phys* **6**, 345–347 (2024).

9. T. Yoshioka, K. Sasada, **Y. Nakano**, and K. Fujii, “Experimental Demonstration of Fermionic QAOA with One-Dimensional Cyclic Driver Hamiltonian”, [2023 IEEE International Conference on Quantum Computing and Engineering \(QCE\) \(2023\)](#).
10. T. Yoshioka, K. Sasada, **Y. Nakano**, and K. Fujii, “Fermionic Quantum Approximate Optimization Algorithm”, [Phys. Rev. Research 5, 023071 \(2023\)](#).

Oral presentations

10. **Y. Nakano**, K. N. Okada, and K. Fujii, “Neural-network-assisted Monte Carlo sampling with Quantum Approximate Optimization Algorithm” HPC198&QS14 (March, 2025).
11. **Y. Nakano**, “QAOA-MC: Markov chain Monte Carlo enhanced by Quantum Alternating Operator Ansatz”, 49th Q-LEAP Quantum AI Seminar (July, 2023).
12. **Y. Nakano**, H. Hakoshima, K. Mitarai, and K. Fujii, “Acceleration of Markov Chain Monte Carlo Method with Variational Quantum Circuit”, QS8 (March, 2023).

Poster presentations

13. **Y. Nakano**, K. N. Okada, and K. Fujii, “Neural-network-assisted Monte Carlo sampling trained by Quantum Approximate Optimization Algorithm”, AQIS2025 (August, 2025).
14. **Y. Nakano**, K. N. Okada, and K. Fujii, “Neural-network-assisted Monte Carlo sampling trained by Quantum Approximate Optimization Algorithm”, Quantum Innovation 2025 (July, 2025).
15. **Y. Nakano**, H. Hakoshima, K. Mitarai, and K. Fujii, “QAOA-MC: Markov chain Monte Carlo enhanced by Quantum Alternating Operator Ansatz”, AQIS2023 (August, 2023).

Awards

- Quantum Innovation 2025 Poster Presentation Awards for Young Researchers, Quantum Innovation 2025 (July, 2025).