Near-Term Quantum Algorithms for Classical Sampling

by

Alev Orfi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Physics (Quantum Information)

Waterloo, Ontario, Canada, 2023

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In the current era of noisy intermediate-scale quantum devices, quantum sampling algorithms have been of great interest as they permit errors in their execution while maintaining their advantage over classical counterparts [26]. However, the sampling problems considered often do not possess immediate practical relevance. This thesis explores two quantum algorithms for applicable classical sampling problems that can be implemented on today's quantum devices. Specifically, we are considering algorithms to sample from a Boltzmann distribution of a classical Hamiltonian. This sampling task is of significant importance in the fields of statistical physics, machine learning, and optimization.

The first such algorithm adiabatically prepares a quantum state which encodes the desired Boltzmann distribution [44]. Projectively measuring this state then produces uncorrelated samples from the desired distribution. The state preparation time scaling of this algorithm can be related to the properties of quantum phase transitions, giving physical insights into the mechanism of speedups found. Numerical investigations of the algorithmic performance on the Ising chain are reproduced, showing a quadratic improvement over a classical Markov chain Monte Carlo (MCMC) method. On the same model, counter-diabatic driving protocols are explored with the limitation of local driving terms. It is shown numerically this restriction of local driving terms leads to unfavourable scaling of the state preparation time.

Next, the quantum-enhanced Markov Chain Monte Carlo algorithm is explored [23]. This hybrid algorithm creates a Markov chain over the classical configuration space, where new configurations are proposed through a projectively measured quantum evolution. This algorithm has guaranteed convergence, independent of the quality of the evolution, making it an algorithm suited for near-term implementation. The performance of this algorithm on the Sherrington-Kirkpatrick model is numerically reproduced, showing faster mixing time than classical MCMC in the low-temperature limit. Bottlenecks of this chain are then explored for the Ising chain, giving an analytic bound on performance showing the algorithmic advantage found for small systems numerically persists for larger system sizes. Finally, this algorithm is tested numerically on the maximum independent set problem, which is native to an array of Rydberg atoms and has been experimentally realized on current quantum devices [12]. Our findings did not indicate any advantage of the quantum-enhanced MCMC algorithm over classical algorithms for the limited number of numerically accessible system sizes.

Acknowledgements

First, I would like to thank my co-supervisors Pooya Ronagh and Ray Laflamme, for supporting this work and providing valuable guidance. I am grateful for the opportunities they have provided throughout my Master's. I have learned tremendously through Ray's group meetings as well as his weekly group socials.

I would also like to express my gratitude to the other members of my committee, Roger Melko and Juan Carrasquilla. They have both been incredibly approachable sources of knowledge and advice. Attending their group meetings soon became a highlight of my week. I am particularly appreciative of the remarkably welcoming community of the Perimeter Institute Quantum Intelligence Lab (PIQuIL), which has shaped my development as a researcher.

I am very fortunate to have had the opportunity to spend the summer of 2022 at Flatiron Insitute working with Dries Sels. I cannot express how thankful I am for Dries' continual commitment to this project. Without our weekly meetings, the research presented in this thesis would not have been possible. He has been an endless source of stimulating scientific discussions as well as invaluable advice and encouragement.

During my time in Waterloo, I have been lucky to have been surrounded by such enthusiastic and impressive fellow students. I have grown significantly through discussions and I have made unforgettable friendships. A special thank you to everyone who gave me feedback on my thesis. Schuyler, Roeland, Esha, Ben, and Sebastian your comments were invaluable and changed how I viewed my own work.

Thank you to my family for their unconditional support. To my parents, who have put up with me coming home every weekend to steal food. And to Izzy, who is still convinced all I do is sit on a couch and think. Above all, Simon whose love and support holds immeasurable importance.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The task of sampling from a classical distribution is ubiquitous in statistical physics and has critical applications in other fields such as machine learning and optimization. A sampling problem is characterized by a probability distribution $D_{\{\lambda\}} : S \to \mathbb{R}$ which acts on a space S and may be parameterized by a set of parameters $\{\lambda\}$. The task is to output samples $y \in S$ according the distribution $D_{\{\lambda\}}$. Sampling algorithms play a crucial role in extracting insights into problems where exhaustive analysis or analytical solutions are infeasible. For problems with high-dimensional parameter spaces, sampling algorithms are often the only methods to acquire an understanding of the underlying distribution.

A quantum state $|\psi\rangle$ when measured in a defined basis $\{|n\rangle\}$ encodes a sampling problem. This measurement gives outcome n with probability $p(n) = |\langle n|\psi\rangle|^2$ according to Born's rule. This process generates probabilistic samples from a distribution determined by the state preparation method, for example, a quantum circuit. With this procedure, it is possible to construct distributions which can be efficiently sampled by a quantum computer but are difficult to sample classically [26]. Recent experimental demonstrations have shown instances where current quantum devices exhibit superior performance compared to classical devices for carefully designed sampling problems [3, 47, 48, 49, 27]. Although these experiments represent an important milestone in the field of quantum computing, the sampling problems considered do not have applications outside of the field.

This thesis will focus on a sampling task characterized by a specific distribution; the Boltzmann distribution of a classical Hamiltonian at a fixed temperature. This Hamiltonian will be defined on the finite space characterized by N spin variables $\sigma_i = \pm 1$. A spin configuration $\vec{\sigma}$, a vector of length N, is the assignment of a value to each spin variable. The state space is then spanned by the set of 2^N possible spin configurations, $S = \{\vec{\sigma}\}$. At

the core of equilibrium statistical mechanics is the notion of a partition function, encompassing all the crucial details pertaining to a system. The form of the partition function for a system with Hamiltonian $H(\vec{\sigma})$ is

$$\mathcal{Z} = \sum_{\{\vec{\sigma}\}} e^{-\beta H(\vec{\sigma})}, \tag{1.1}$$

where β is the inverse temperature. The above sum is over all possible states of the system, thus calculating $\mathcal{Z}$ is typically intractable. The partition function is essential in defining the Boltzmann distribution, which gives the probability that the system will be in a certain state. Specifically, the probability the system will be in the state $\vec{\sigma}$ at a given temperature $T = 1/\beta$ is,

$$\pi(\vec{\sigma}) = \frac{1}{\mathcal{Z}} e^{-\beta H(\vec{\sigma})}. \tag{1.2}$$

An algorithm which can sample efficiently from this distribution can be used to estimate numerical quantities which are hard to compute deterministically. A common application of these algorithms is to approximately compute thermodynamic quantities. For example, calculating expectation values of the form,

$$\langle \mathcal{O} \rangle_\pi = \frac{1}{\mathcal{Z}} \sum_{\{\vec{\sigma}\}} \mathcal{O}(\vec{\sigma}) e^{-\beta H(\vec{\sigma})}. \tag{1.3}$$

is typically infeasible. A sampling algorithms generates, for example M samples $\{\vec{\sigma}_1, \vec{\sigma}_2, ..., \vec{\sigma}_M\}$ according to π. The above observable can then be approximated with these generated samples,

$$\langle \mathcal{O} \rangle \sim \frac{1}{M} \sum_{i=1}^{M} \mathcal{O}(\vec{\sigma}_i) = \bar{\mathcal{O}}_M. \tag{1.4}$$

For many physical systems, this estimation method is the only method to gain insight into their thermodynamic properties and phase transitions [34, 36].

This class of sampling problems is not only essential in statistical physics but is a common subroutine in machine learning and combinatorial optimization [20]. For example, this sampling task is the primary training bottleneck of energy-based models. These are very successful probabilistic machine-learning networks which recently achieved state-of-the-art generative modelling performance [16, 41]. Due to the importance of this sampling, there is an active field of research proposing and investigating quantum algorithms for this task [17, 2, 24, 29, 40, 42, 46]. Several distinct approaches exist to explore this problem, many taking inspiration from their classical counterparts. For example, quantum walks

have been shown to give quadratically faster mixing time compared to classical Markov Chain Monte Carlo in a variety of special cases [33]. Additionally, quantum simulated annealing algorithms have also shown an algorithmic speedup over their classical analogues [46, 17, 24]. A subset of these algorithms will be briefly described in Chapter 2.

Current quantum processors are of limited size, have short coherence times, and can only perform error-prone gates. These physical devices are not yet able to achieve the algorithmic advantages found in the above algorithms. The current quantum processors, dubbed noisy intermediate-scale quantum (NISQ) devices [32], require quantum algorithms which are resistant to the inherent errors of the physical system [19]. The focus of this thesis is quantum algorithms for Boltzmann sampling which do not require fault-tolerant quantum computers but can be implemented on today's devices.

- Chapter 2 outlines the common classical and quantum sampling algorithms. The main focus is Markov chain Monte Carlo methods. Conditions for convergence, mixing and autocorrelation times, common chains, and bottleneck ratios are reviewed. Quantum walks and more sophisticated quantum algorithms for sampling are discussed.

- In Chapter 3 an adiabatic state preparation algorithm for sampling is examined. Known results for sampling the Ising chain at low temperatures with this algorithm are reproduced. The applicability of this algorithm to other more complicated sampling problems will be discussed. This algorithm will then be coupled with counter-diabatic driving protocols. The counter-diabatic drives will be approximated with local Hamiltonians in order to maintain the possibility of near-term implementation.

- Chapter 4 focuses on the quantum-enhanced Markov Chain Monte Carlo algorithm, examining its performance in various scenarios. We begin by numerically replicating its performance on the Sherrington-Kirkpatrick model. The mixing time of this method is then bounded analytically for the Ising chain by analyzing the bottleneck of the Markov chain. Numerically this algorithm is then tested on the maximum independent set (MIS) problem. This is motivated by the physical implementation of the MIS on Rydberg atom devices.

Chapter 2

Sampling Algorithms Background

2.1 Markov Chain Monte Carlo

In the context of algorithms considered in this thesis we are interested in a finite, discrete state space S. A Markov chain on this space is a process such that when at some $x \in S$ the next state, $y \in S$, is chosen with a fixed probability dependant only on the current state, x. We can specify this process with a transition matrix P which describes the probability of moving between any two states in S. Formally, this property states that a sequence of random variables $(X_0, X_1, ...)$ is a Markov chain on S if for every $t > 0$,

$$\mathbb{P}(X_{t+1} = y | X_0, X_1, ..., X_t = x) = P(x, y). \tag{2.1}$$

Where $P(x, y) \geq 0$ describes the probability to move from x to y. The transition matrix $P = (P(x, y))_{x,y \in S}$ then fully describes this process and will be essential in analyzing these algorithms. As the x-th row of P describes the probability for the state x to move all other possible states, the row describes a probability distribution. As a result, P is a stochastic matrix,

$$\sum_{y \in S} P(x, y) = 1 \quad \forall x \in S. \tag{2.2}$$

Let $(X_0, X_1, ...)$ be a Markov chain on S, and let μ_t be the vector form of the distribution of X_t over the state space S. An element of this vector μ_t is defined to be,

$$\mu_t(x) = \mathbb{P}\{X_t = x\} \tag{2.3}$$

for some $x \in S$. This distribution μ_t can be written in terms of previous distributions of the chain,

$$\mu_t = \mu_{t-1}P = \mu_0 P^t. \tag{2.4}$$

Ideally, we want this Markov chain to converge to a stationary distribution π, that is $\mu_t \to \pi$ as $t \to \infty$. In matrix notation, this is equivalently,

$$\pi = \pi P. \tag{2.5}$$

Markov chain Monte Carlo (MCMC) is the method of using a Markov chain to create samples from a desired probability distribution. If a chain can be constructed such that its stationary distribution is π, then after some large enough number of Markov chain steps, the chain will produce samples from a distribution sufficiently close to π. The rate of convergence of MCMC algorithms to their stationary distribution is an important aspect of their algorithmic complexity.

2.1.1 Convergence Conditions

The total variation (TV) distance between two probability distributions μ and ν is the max distance between the probability assigned by both distributions to a single event A [25],

$$||\mu - \nu||_{TV} = \max_{A \subseteq S} |\mu(A) - \nu(A)|. \tag{2.6}$$

Using this, we can then define a distance between a Markov chain at step t and the distribution π,

$$d(t) = \max_{x \in S} ||P^t(x, \cdot) - \pi||_{TV}. \tag{2.7}$$

Before considering the convergence of the chain, it is necessary to show the existence of π. In particular, if P is a transition matrix of an irreducible Markov chain then there exists a unique stationary distribution [25]. An irreducible chain has a non-zero probability of moving between any states after some t number of steps, $P^t(x, y) > 0 \ \forall x, y \in S$. If this were not the case, the chain would not be able to access all of the state space. To show that the chain converges to this stationary state, we also need the chain to be aperiodic [25]. Let,

$$T(x) = \{t \geq 1 : P^t(x, x) > 0\} \tag{2.8}$$

then the period of a state $x \in S$ is the greatest common divisor of $T(x)$. Here $T(x)$ is the set of t where there is a non-zero probability for the chain to return to x, and its greatest common divisor gives the period of any cyclic behaviour of this state's movement.

A chain is then aperiodic if the period of every state is 1. This is equivalent to saying that for every starting state, the chain has a non-zero probability of staying in its current position. Periodic chains return to the same states after a fixed number of steps and do not converge to a steady state distribution. A chain P with these two properties, irreducible and aperiodic, can be shown to converge to the stationary distribution of Eq. 2.5[25][5].

Another important property of a chain is reversibility. Specifically, a chain is reversible if it satisfies the detailed balance condition,

$$\pi(x)P(x,y) = \pi(y)P(y,x) \quad \forall x, y \in S. \tag{2.9}$$

If a distribution π satisfies Eq. 2.9 then it is stationary for P since P is stochastic,

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

This condition is not necessary for π to be a stationary distribution of P, however, it is a sufficient condition. As will see in Chapter 4, when given a desired steady-state distribution, it is common to use this condition to engineer a stochastic process that is guaranteed to converge.

2.1.2 Markov Chain Mixing

Assume we have an aperiodic, irreducible chain P that satisfies the detailed balance condition. If λ is an eigenvalue of P then, since P is stochastic $|\lambda| \leq 1$, as a result of the Perron-Frobenius theorem [25]. Due to the detailed balance condition, all eigenvalues are real and can be therefore ordered as,

$$\lambda_1 = 1 > \lambda_2 \geq \lambda_3 \geq ... \geq -1.$$

An important note is that since P is irreducible, there is a unique steady state distribution, and the eigenspace associated with the eigenvalue 1 is one-dimensional [25]. Specifically, from Eq. 2.5, the stationary state is the left eigenvector of this eigenvalue.

The mixing time of a Markov chain is defined as the time required for the total variational distance to the stationary state π to be ϵ-small,

$$t_{\text{mix}}(\epsilon) = \inf\{t \geq 0 : d(t) \leq \epsilon\}. \tag{2.10}$$

The distribution of the chain at time t, μ_t, is the repeated application of the transition matrix P. As π is associated with the eigenvalue 1 of P, the spectral gap $\delta = 1 - |\lambda_2|$

gives information about the rate of coverage. More precisely, for $t \to \infty$, $d(t)$ decays exponentially with the rate of decay governed by this spectral gap δ. This structure can then be used to bound $t_{\text{mix}}(\epsilon)$. Specifically, given a reversible, irreducible Markov chain on S with transition matrix P, then

$$(\delta^{-1} - 1) \ln\left(\frac{1}{2\epsilon}\right) \leq t_{\text{mix}}(\epsilon) \leq \delta^{-1} \ln\left(\frac{1}{\epsilon \pi_{\min}}\right) \tag{2.11}$$

where $\pi_{\min} = \min_{x \in S} \pi(x)$. The proof of this theorem can be found in Chapter 12 of Ref. [25]. Due to this bound, the spectral gap is commonly used in performance analyses of this type of Markov chain.

2.1.3 Common Chain Constructions

Let $S = \{\vec{\sigma}\}$ be the configuration space of a system of N discrete spin variables $\sigma_i = \pm 1$. Consider a classical Hamiltonian $H_c(\vec{\sigma})$ acting on these spins, we are interested in sampling from the Boltzmann distribution,

$$\pi(\vec{\sigma}) = \frac{1}{Z} e^{-\beta H_c(\vec{\sigma})}, \tag{2.12}$$

where $Z = \sum_{\{\vec{\sigma}\}} e^{-\beta H_c(\vec{\sigma})}$ is the partition function and β the inverse temperature.

In the previous sections, given some chain P we examined conditions needed for it to converge to a stationary distribution π. Here given a distribution, π, we want to construct a P that converges but can also be simulated efficiently. We need to design a stochastic process which produces the needed P. Often this is achieved by creating chains that are composed of two steps, a proposal step and an acceptance step. If the chain is at some configuration $X_t = \vec{\sigma}$, a new configuration $\vec{\sigma}'$ is proposed with probability $Q(\vec{\sigma}'|\vec{\sigma})$. Then this new configuration is accepted as the new state of the chain, $X_{t+1} = \vec{\sigma}'$, with probability $A(\vec{\sigma}'|\vec{\sigma})$. We can then construct the transition matrix P according to these steps.

$$P(\vec{\sigma}, \vec{\sigma}') = \begin{cases} Q(\vec{\sigma}'|\vec{\sigma}) A(\vec{\sigma}'|\vec{\sigma}), & \text{if } \vec{\sigma}' \neq \vec{\sigma} \\ 1 - \sum_{\{\vec{\sigma}_i \neq \vec{\sigma}'\}} Q(\vec{\sigma}'|\vec{\sigma}_i) A(\vec{\sigma}'|\vec{\sigma}_i), & \text{if } \vec{\sigma}' = \vec{\sigma} \end{cases} \tag{2.13}$$

The diagonal elements of this matrix are explicitly constructed so that P is stochastic and represents the probabilistic movement of the chain.

We will explore different proposal and acceptance strategies and investigate the produced chain's convergence rate for different problems. For both the proposal and acceptance steps, it is important to consider strategies that can be computed efficiently. Below

we will consider two such acceptance strategies which, although they involve π, only require a ratio so that the intractable partition function of the Boltzmann distribution cancels. The first such acceptance strategy, known as Metropolis-Hastings, accepts a new configuration with the following probability,

$$A(\vec{\sigma}'|\vec{\sigma})_{MH} = \min\left(1, \frac{\pi(\vec{\sigma}')}{\pi(\vec{\sigma})}\frac{Q(\vec{\sigma}|\vec{\sigma}')}{Q(\vec{\sigma}'|\vec{\sigma})}\right) = \min\left(1, e^{-\beta(H_c(\vec{\sigma}')-H_c(\vec{\sigma})}\frac{Q(\vec{\sigma}|\vec{\sigma}')}{Q(\vec{\sigma}'|\vec{\sigma})}\right). \tag{2.14}$$

If the new configuration is lower in energy, it will be accepted by the chain with unit probability. Another common method is to accept new configuration with the following Glauber probability,

$$A(\vec{\sigma}'|\vec{\sigma})_G = \left[1 + \left(\frac{\pi(\vec{\sigma}')}{\pi(\vec{\sigma})}\frac{Q(\vec{\sigma}|\vec{\sigma}')}{Q(\vec{\sigma}'|\vec{\sigma})}\right)^{-1}\right]^{-1} = \left[1 + \left(e^{-\beta(H_c(\vec{\sigma}')-H_c(\vec{\sigma}))}\frac{Q(\vec{\sigma}|\vec{\sigma}')}{Q(\vec{\sigma}'|\vec{\sigma})}\right)^{-1}\right]^{-1}. \tag{2.15}$$

Unlike Metropolis-Hastings, if the new configuration has lower energy, it is only accepted with 50% probability. This leads to a "lazy" chain, which is often implemented to avoid periodic behaviour [25]. Often the chosen proposal strategy is symmetric, that is $Q(\vec{\sigma}'|\vec{\sigma}) = Q(\vec{\sigma}|\vec{\sigma}')$, and both acceptance probability further reduce. Both of these chains are designed to satisfy the detailed balance condition [22].

2.1.4 Autocorrelation Time and Critical Slowing Down

So far we have discussed one time measure, the mixing time $t_{\text{mix}}(\epsilon)$. However, this measure is often not tractable to calculate for large systems as it requires diagonalizing the mixing matrix P. When using MCMC-generated samples to estimate some observable B, as in Eq. 1.4, correlations between samples lead to additional error in the estimate as the samples are not statistically independent. The autocorrelation function of the observable B is defined as [5],

$$A_B(x) = \frac{\langle B(t+x)B(t)\rangle - \langle B\rangle^2}{\langle B^2\rangle - \langle B\rangle^2}. \tag{2.16}$$

For samples far apart in the chain (i.e., large x) this function goes to zero. Typically, this decay is asymptotically exponential [22],

$$A_B(t) \sim e^{-t/\tau_B}, \tag{2.17}$$

which defines another time measure – the autocorrelation time τ_B [34]. This time is important in calculating estimator errors and is fundamentally related to the mixing time of the chain. Like the mixing time, it is upper bounded by the inverse spectral gap of the chain [15], validating the use of only the spectral gap as a measure of chain performance.

The autocorrelation time is also used to characterize an important effect when using MCMC methods to estimate observables near a phase transition. Near a critical point, there are large spatial correlations in the systems which correspond to time correlations in the chain. This leads to an effect known as critical slowing down, where near a phase transition, a Markov chain will require a longer time to mix. This is typically quantified with the autocorrelation time as, unlike the mixing time, it is tractable to compute. This critical behaviour of the chain can be expressed as a power law,

$$\tau_B \propto \xi^z$$

where ξ is the correlation length and z is the dynamical critical exponent. The value of z depends on the specific dynamics of the chain. For instance, near a critical point where there are large fluctuations, a local update strategy will have difficulty sampling relevant configurations and thus a larger z as compared to a cluster method. Throughout the rest of this thesis, we will only be considering the mixing time of chains. Mixing time is the maximum autocorrelation time over all possible observables for reversible chains [5]. Since mixing time upper bounds autocorrelation time, it is a justifiable measure of MCMC performance [25].

2.1.5 Cheeger Inequality

There are multiple types of bounds that can be put on the mixing time of a Markov chain. We have seen already how for a reversible, irreducible chain the spectral gap δ of P is an important metric. However, finding δ quickly becomes intractable since P is a $2^N \times 2^N$ dense matrix for the spin problems considered in this thesis. Instead, another avenue of analysis is to consider the bottlenecks of the chain – geometric features that restrict the flow of the chain and ultimately determine the mixing time.

There is a natural graph associated with a chain, where the vertices are elements of S, and edge weights are the equilibrium flow through that edge. Assuming a reversible chain, the equilibrium flow between vertex x and y is given by $E(x, y) = \pi(x)P(x, y)$. Given $A, B \subset S$, the flow from A to B is

$$E(A, B) = \sum_{x \in A, y \in B} E(x, y). \tag{2.18}$$

We are interested in determining where the flow is restricted in the chain, so it is useful to define a measure of the amount of flow out of a set relative to the size of the set. Specifically, this is the bottleneck ratio of a set A [25],

$$\Phi(A) = \frac{E(A, A^c)}{\pi(A)}, \tag{2.19}$$

where $\pi(A) = \sum_{x \in A} \pi(x)$. The bottleneck ratio of the whole chain is then,

$$\Phi_\star := \min_{A:\pi(A)\leq 1/2} \Phi(A), \tag{2.20}$$

which is a minimization over all subsets to determine where the restriction of movement occurs. As mentioned, this bottleneck is related to the mixing time of the chain. It bounds the spectral gap on both sides through Cheeger's inequality,

$$\frac{\Phi_\star^2}{2} \leq \delta \leq 2\Phi_\star. \tag{2.21}$$

The proof of this inequality can be found in Chapter 13 of [25]. This gives another avenue for understanding the mixing time, however the minimization in Eq. 2.20 is often difficult. For some problems, $\Phi_\star$ can be determined, and it is often used to prove the slow mixing of chains.

2.2 Quantum Sampling Algorithms

2.2.1 Quantum Walks

The first quantum algorithm we will discuss is the discrete-time random walk, as it is a quantum analog of the classical MCMC methods considered above [9]. A general framework to convert classical Markov chains into the quantum setting was presented by Szegedy [42]. Originally formulated for search problems, this process was later combined with the quantum adiabatic algorithm [40] in order to produce samples from a desired distribution.

At each time step, a classical MCMC method must occupy one possible state of the state space S, with transitions governed by P. Quantum walks instead transition the state x into a superposition of all possible y determined by P,

$$|x\rangle \to \sum_{y \in S} \sqrt{P(x, y)}\, |y\rangle\,. \tag{2.22}$$

This process is not unitary, requiring this algorithm to instead act on an enlarged Hilbert space $\mathcal{H} = \mathbb{C}^d \otimes \mathbb{C}^d = \{|x\rangle |y\rangle : x, y \in S\}$ [42]. In this space, this transition unitary W is defined as,

$$W |x\rangle \otimes |0\rangle = |w_x\rangle \otimes |x\rangle = \sum_{y \in S} \sqrt{P(x,y)} |y\rangle \otimes |x\rangle .$$

The quantum walk is performed through successive applications of the unitary U_W, which is defined as,

$$U_W = (2\Pi_0 - 1)W^\dagger \Lambda W.$$

Here Π_0 is the projector on the subspace spanned by $S_0 = \{|x\rangle \otimes |0\rangle\} \forall x \in S$ and Λ is the swap operator. The stationary state of this walk is the following superposition of the classical steady-state distribution π [42],

$$|\pi\rangle = \sum_x \sqrt{\pi_x} |x\rangle . \tag{2.23}$$

Projectively measuring this state gives an uncorrelation sample of the wanted classical distribution. What is advantageous about the quantum walk is that its spectral gap is quadratically larger than the classical chain on which it is based, $\delta_Q \geq \sqrt{\delta_C}$ [9].

To understand the properties of U_W it is useful to consider the operator X, which is obtained by projecting $W^\dagger \Lambda W$ onto S_0.

$$\begin{aligned} X &= \Pi_0 W^\dagger \Lambda W \Pi_0 \\ &= \sum_{x,y} \sqrt{P_{xy} P_{yx}} |y\rangle\langle x| \otimes |0\rangle\langle 0| \\ &= \sum_{x,y} \sqrt{\frac{\pi(x)}{\pi(y)}} P_{yx} |y\rangle\langle x| \otimes |0\rangle\langle 0| \end{aligned} \tag{2.24}$$

Where we have made use of the detailed balance condition of the classical chain. On the subspace S_0 we have that X and P are similar and thus share eigenvalues. Define the eigenstates of X and P as,

$$X |\tilde{e}_k\rangle = \lambda_k |\tilde{e}_k\rangle . \tag{2.25}$$

Extending these states to the larger space as $|e_k\rangle = |\tilde{e}_k\rangle \otimes |0\rangle$, we can consider U_W without the projection acting on this state,

$$W^\dagger \Lambda W |e_k\rangle = \lambda_k |e_k\rangle - \beta_k \left|e_k^\perp\right\rangle . \tag{2.26}$$

Here $\left|e_k^\perp\right\rangle$ orthogonal to S_0 and thus also to all $|e_k'\rangle$ with $k' \neq k$. Additionally, since $W^\dagger \Lambda W$ is unitary the $\left|e_k^\perp\right\rangle$ are also mutually orthogonal. We have now that $W^\dagger \Lambda W$ is

block diagonal in $\{|e_k\rangle, |e_k^\perp\rangle\}$ and again as it is unitary we can write $\beta_k = \sqrt{1-|\lambda_k|^2}$. Overall we have,

$$\begin{aligned} U_W |e_k\rangle &= \lambda_k |e_k\rangle + \sqrt{1-|\lambda_k|^2}\, |e_k^\perp\rangle \\ U_W |e_k^\perp\rangle &= \sqrt{1-|\lambda_k|^2}\, |e_k\rangle - \lambda_k |e_k^\perp\rangle \end{aligned} \tag{2.27}$$

Thus the eigenvalues of U_W are $e^{\pm i\theta_k}$ with $\cos(\theta_k) = \lambda_k$. Due to the construction of U_W, it can be shown that the state that corresponds to the eigenvalue $\theta_k = 0$ is the stationary distribution on S_0 [9],

$$|\pi\rangle \otimes |0\rangle . \tag{2.28}$$

Then quantum phase estimation, a common quantum subroutine which is used to measure the eigenvalues of a unitary is used on U_W [9]. We want this measurement of the eigenvalues to be accurate enough to resolve the eigenvalue $\theta = 0$. The spectral gap of U_W is $\delta_Q = \arccos(\lambda) \sim \sqrt{\delta_C}$, we need at least this precision, which leads to a runtime of $O(1/\delta_Q)$. In all this algorithm requires $1/\sqrt{\delta_C}$ steps of the quantum walker, or equivalently $1/\sqrt{\delta_C}$ applications of U_W. These types of algorithms are typically formulated with the assumption of oracle formation of the classical transition matrix. There have been further improvements to this algorithm that avoid the costly implementation of the oracle [24].

2.2.2 Quantum Simulated Annealing

One of the state-of-the-art quantum algorithms for approximating partition functions uses quantum simulated annealing [2]. As in their classical counterparts, the annealing process is characterized by a sequence of inverse temperatures β. The system will begin with the goal of approximating a partition function at high temperature and then, following this defined sequence, the temperature will be decreased until the desired distribution is reached [39].

Consider an initial state for the above quantum walk algorithm, $|\psi\rangle \otimes |0\rangle$. The probability of a measurement outcome of $\theta_k = 0$ is then $|\langle\psi|\pi\rangle|^2$, thus it is important to choose an initial state with large overlap with the steady-state distribution. In some cases, an initial state with a significant overlap can be efficiently prepared, but if this is not the case, adiabatic state preparation can be employed. This process adds additional runtime depending on the problem considered [24]. Another method is to consider a sequence of slow varying Markov chains [46]. These sequences are specifically chosen such that the stationary distributions of adjacent Markov chains are sufficiently close. The system can then stay in the $\theta_k = 0$ space for each of these quantum walks. By preparing intermediate

states through amplitude amplification, a speedup can be maintained [46], with further improvements by implementing an annealing schedule [39]. Adaptive sequences of quantum walks were later explored, allowing changes to the sequences throughout the algorithm [17, 2, 29]. These protocols often adapt optimized annealing schedules from their classical simulated annealing counterparts [2].

2.3 Computational Complexity

Classical Markov chains have a mixing time $O(\delta^{-1}\log(1/\min_x \pi(x)))$ coming from the bound of Eq. 2.11. In the most general context, quantum Markov chains have a mixing time $O(1/\sqrt{\delta \min_x \pi(x)})$ [17]. Due to this dependence on $\min_x \pi(x)$, there is no general quantum speedup for Monte Chain Monte Carlo methods, but speedups can be observed on certain problems.

In these quantum algorithms, the goal is to create a quantum state of the form,

$$|\pi\rangle = \sum_x \sqrt{\pi_x}\,|x\rangle\,. \tag{2.29}$$

Measuring this state leads to an uncorrelated sample from the distribution π. The state of Eq. 2.29 is known as a qsample. Preparing a qsample $|\pi\rangle$ is generally difficult, even if sampling from π is efficient [30]. For example, consider having access to an algorithm that can efficiently produce a state x with probability $\pi(x)$. One could make a quantum circuit which prepares the state $|\tilde{\pi}\rangle = \sum_x \sqrt{\pi(x)}\,|x\rangle\,|\phi(x)\rangle$ [30]. Here the extra register is required as the quantum circuit must be unitary. However, there is no efficient way to remove the information in the second register, making $|\pi\rangle$ unobtainable.

Preparing qsamples without information about the underlying distribution is part of a complexity class known as Statistical Zero Knowledge (SZK). This is a class of decision problems that can be verified by a statistical zero-knowledge proof [43]. SZK contains problems such as graph isomorphism. It is not believed that the problems in this class can be solved in bounded-error quantum polynomial (BQP) time, specifically BQP $\subseteq$ SZK [30]. If there were a generic method to prepare qsamples more efficiently than the classical chains, then it would imply SZK $\subseteq$ BQP.

The quantum and classical mixing time scaling have different dependencies on $\min_x \pi(x)$, which further restricts the possible quantum performance. There is not expected to be a generic method to prepare qsamples even polynomially more slowly than the mixing time of classical MCMC methods [17]. However, there are problems in which a quantum speedup

has been proven. In these cases, specific structures of the problems have been exploited, and the methods cannot be generalized. This will be the case throughout this thesis. We are considering algorithms which use information about the classical Hamiltonian of the desired Boltzmann distribution. Preparing a qsample with this information is no longer in the SZK class, allowing the possibility of a quantum speedup.

Chapter 3

Adiabatic Quantum Sampling

In contrast to the quantum algorithms considered in Chapter 2, the remainder of this thesis will focus on algorithms that do not require large universal quantum computers but are well-suited for current devices. The first such algorithm [44, 45], adiabatically prepares a quantum state which encodes the Boltzmann distribution of a classical problem. Projectively measuring this state then produces an independent sample from this distribution. Properties of the phase transitions in the space that is adiabatically traversed give physical insights into the computational speedup found.

3.1 Algorithm Overview

3.1.1 Parent Hamiltonian Definition

This algorithm [44] is designed to produce unbiased samples from the Boltzmann distribution, Eq. 2.12, of a classical Hamiltonian H_c. This distribution can be encoded in the following qsample, as in Eq. 2.29,

$$|\psi(\beta)\rangle = \frac{1}{\sqrt{Z}} \sum_{\vec{\sigma}} e^{-\beta H_c(\vec{\sigma})/2} |\vec{\sigma}\rangle \,. \tag{3.1}$$

Projectively measuring this state in the computational basis generates a sample of the Boltzmann distribution. The time complexity of the algorithm is related to the preparation process of this state. This is done adiabatically, starting from a state that is easy to prepare and slowly varying parameters until the desired state is reached. The state space which

is traversed adiabatically is defined by a parent Hamiltonian whose ground state is the qsample [44].

Consider a reversible Markov chain P whose stationary distribution π is the Boltzmann distribution of the classical Hamiltonian H_c. The parent Hamiltonian is defined with this transition matrix P as

$$H_q(\beta) = N(\mathbb{I} - e^{-\beta H_c/2} P e^{\beta H_c/2}). \tag{3.2}$$

The second term of the Hamiltonian is a similarity transformation on P, thus it conserves its eigenvectors and eigenvectors. The eigenspace corresponding to the $\lambda_1 = 1$ eigenvalue of P is one-dimensional since the chain has a unique steady state distribution. As a result, $H_q(\beta)$ has a unique ground state with zero energy. It can be easily shown that the ground state of this Hamiltonian is exactly the qsample of Eq. 3.1. Additionally, the spectral gap of the Markov chain is the energy difference between the first excited state and the ground state. The factor of N ensures that the spectrum of H_q is extensive to match physical Hamiltonians. The relationship between the classical mixing matrix and the spectra of the quantum Hamiltonian allows for algorithmic speed-ups depending on the specifics of the quantum phase transitions in the adiabatic space.

3.1.2 Adiabatic State Preparation

Starting at an easy-to-prepare ground state of H_q we can adiabatically modify parameters of H_q to obtain the qsample of Eq. 3.1 at a desired β. In order to maintain the adiabatic condition, it is required that the probability of moving out of the ground state is small as the Hamiltonian parameters are varied. Let E_n be the energy of the instantaneous eigenstates $|n\rangle$, with $|0\rangle$ the ground state. The adiabatic condition is be maintained by requiring that for a chosen small ε,

$$\sum_{n>0} \frac{1}{(E_n - E_0)^2} \left| \langle n| \frac{d}{dt} |0\rangle \right|^2 = \varepsilon^2, \tag{3.3}$$

where $\frac{d}{dt}|0\rangle$ is the time dependence of the ground state. This condition ensures that the Hamiltonian parameters are varied slowly when the gap is small, taking into account the probability of moving into a particular excited state $|n\rangle$. Eq. 3.3 established a limit on the transitions between the ground state and any excited state. This is necessary to ensure the adiabatic condition is maintained at a phase transition.

We will consider adiabatic paths parameterized by $s \in [0, 1]$. These paths will be in the parameter space of the Hamiltonian H_q. Labelling these Hamiltonian parameters $\{\lambda_\mu\}$

a path is specified through $\lambda_\mu(s)$'s dependence on s. Due to this parametrization of the path, the time dependence of the ground state is,

$$\frac{d}{dt}\left|0\right\rangle = \frac{ds}{dt}\sum_i \frac{\partial\left|0\right\rangle}{\partial\lambda_i}\frac{d\lambda_i}{ds} = \frac{ds}{dt}\sum_i \left|\partial_\mu 0\right\rangle \frac{d\lambda_i}{ds}. \tag{3.4}$$

The adiabatic condition of Eq. 3.3 restricts the rate of change in terms of s,

$$\frac{dt}{ds} = \frac{1}{\varepsilon}\sqrt{\sum_{\mu,\nu} g_{\mu\nu}(s)\frac{d\lambda_\mu}{ds}\frac{d\lambda_\nu}{ds}} \tag{3.5}$$

where the adiabatic metric $g_{\mu,\nu}$ is defined as [21],

$$g_{\mu\nu} = \sum_{n>0}\frac{1}{(E_n - E_0)^2}\left\langle\partial_\mu 0\right|\left.n\right\rangle\left\langle n\right|\left.\partial_\nu 0\right\rangle. \tag{3.6}$$

Finally, the adiabatic state preparation time can be found by integrating Eq. 3.5,

$$t_{\text{tot}} = \frac{1}{\varepsilon}\int_0^1 ds\sqrt{\sum_{\mu,\nu} g_{\mu,\nu}(s)\frac{d\lambda_\mu}{ds}\frac{d\lambda_\nu}{ds}} = \frac{l}{\varepsilon}. \tag{3.7}$$

This time is dependent on the ε considered. The value of ε controls how much movement out of the ground state is allowed. With a larger ε, the adiabatic path can be traversed faster at the cost of an inaccurate final state. The integral in Eq. 3.7, label as l, is the adiabatic path length,

$$l = \int_0^1 ds\sqrt{\sum_{\mu,\nu} g_{\mu,\nu}(s)\frac{d\lambda_\mu}{ds}\frac{d\lambda_\nu}{ds}}. \tag{3.8}$$

Adiabatic state preparation of the ground state of H_q can be achieved by varying only one parameter. However, considering an extended parameter space allows for adiabatic paths, which result in a speed up as compared to classical mixing times [45, 44].

3.2 Ising Chain

3.2.1 Parent Hamiltonian

The first model considered is the ferromagnetic Ising chain with periodic boundary conditions,

$$H_c = -\sum_{i=1}^{N}\sigma_i^z\sigma_{i+1}^z. \tag{3.9}$$

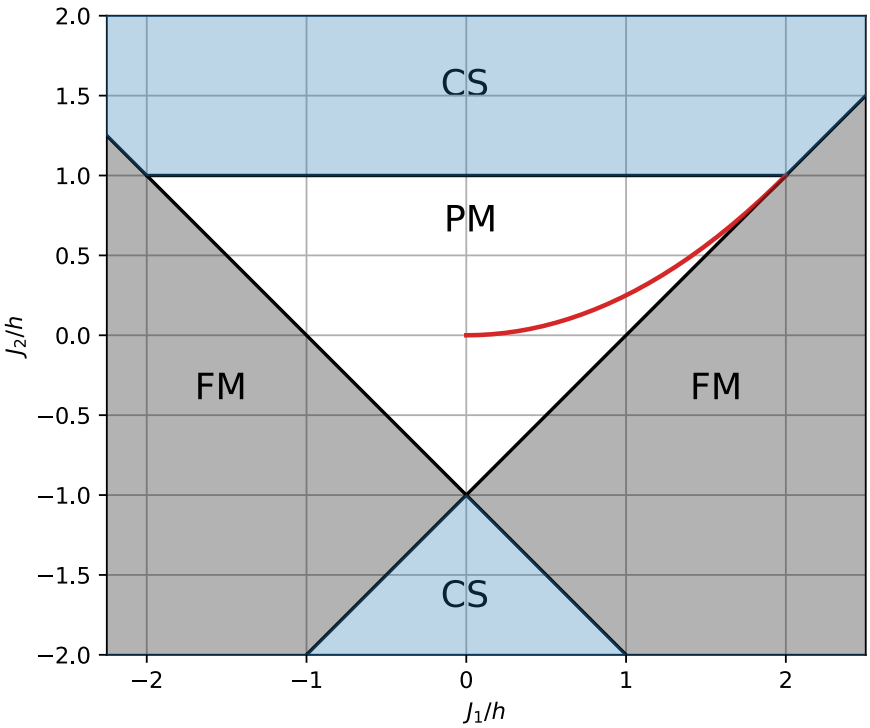

Figure 3.1: Quantum phase diagram of the parent Hamiltonian H_q corresponding the Ising chain. Three different phases are shown; a paramagnetic phase (PM), a ferromagnetic phase (FM), and a cluster-state-like phase (CS). The ground state on the red curve corresponds to the qsample of Eq. 2.29 at some β. The $\beta = 0$ point lies at the origin, and $\beta = \infty$ at $J_1/h = 2, J_2/h = 1$, at a tricritial point. Reproduced from [44].

Here I will reproduce the results in [44] which shows a quadratic improvement in adiabatic time scaling as compared to classical Glauber Markov chain Monte Carlo. The underlying Markov chain is defined on the space of spin configurations. At each step of the chain, a new configuration is proposed by flipping one spin from the previous chain configuration. This new configuration is then accepted according to the Glauber acceptance probability of Eq. 2.15. This update rule was chosen as the transition matrix of this chain can be written compactly with local terms. Consider a proposed configuration with spin i flipped,

the probability $P(\vec{\sigma}'|\vec{\sigma})$ is then,

$$Q(\vec{\sigma}'|\vec{\sigma})A(\vec{\sigma}'|\vec{\sigma})_G = \frac{1}{2N\cosh\big(\beta(\sigma^z_{i-1}+\sigma^z_{i-1})\big)} e^{-\beta\sigma^z_i(\sigma^z_{i-1}+\sigma^z_{i-1})} := p_i. \tag{3.10}$$

The transition matrix is thus,

$$P = \sum_i p_i \sigma^x_i + \left(\mathbb{I} - \sum_i p_i\right). \tag{3.11}$$

After some simplification, which can be found in Appendix B, the parent Hamiltonian of this chain is,

$$H_q(\beta) = \frac{N}{2}\mathbb{I} - \sum_i \left[J_1(\beta)\sigma^z_i\sigma^z_{i+1} + h(\beta)\sigma^x_i - J_2(\beta)\sigma^z_{i-1}\sigma^x_i\sigma^z_{i+1}\right] \tag{3.12}$$

with $h(\beta) = \frac{1}{4}(1 + 1/\cosh(2\beta))$, $J_1(\beta) = \frac{1}{2}\tanh(2\beta)$, and $J_2(\beta) = \frac{1}{4}(1 - 1/\cosh(2\beta))$. This Hamiltonian can be solved exactly through a Jordan-Wigner transformation where the spin degrees of freedom are mapped to free fermions, the details of which can also be found in Appendix B. This Hamiltonian is diagonalized in the momentum space of these fermions, where energy excitations at a given k are given by,

$$\varepsilon_k = \sqrt{(h + J_1(\beta)\cos(k) + J_2(\beta)\cos(2k))^2 + (J_1(\beta)\sin(k) + J_2(\beta)\sin(2k))^2}. \tag{3.13}$$

The ground state is at half filling, so a quantum phase transition occurs whenever $\varepsilon_k = 0$ for some k. For any finite β there is no phase transition, and the system is in a paramagnetic phase. Since the system is gapped, the Markov chain mixes rapidly away from large β. This spectral gap is calculated in Appendix B and for low temperature decreases as $\sim N^{-2}$. This can additionally be related to the dynamic critical exponent z at the $\beta = \infty$ quantum phase transition. The characteristic energy scale of the problem Δ, for example, the gap between the first excited state and the ground state, decreases to the zth power of the correlation length scale [35],

$$\Delta \sim \xi^{-z}. \tag{3.14}$$

For finite-sized systems $\xi \sim N$, and Eq. 3.13 gives a form for Δ in terms of $k \sim 1/N$. At the $\beta = \infty$ phase transition $\varepsilon_k = 0$ at multiple values of k and the dispersion relation is quadratic as seen in Fig. 3.2. This quadratic dependence on k shows this quantum phase transition has $z = 2$.

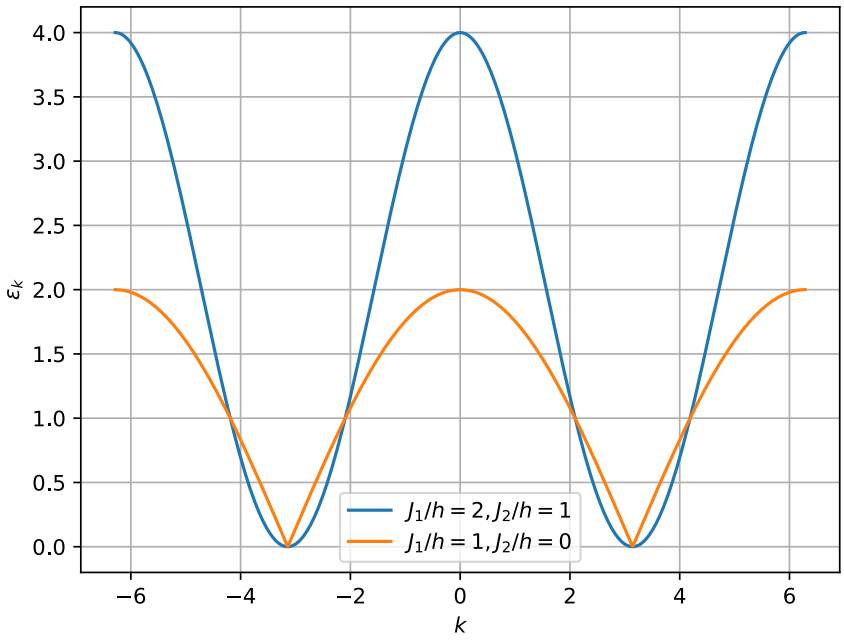

Figure 3.2: Ising chain dispersion relations at the tricritical point (shown in blue) and at the paramagnet to ferromagnetic transition (shown in orange). The quadratic and linear behaviour near the zeros shows the dynamic critical exponent value of $z = 2$ and $z = 1$, respectively.

We can also consider an extended parameter space where h, J_1, and J_2 can take on any value, not restricted by their β dependence. In this extended space, there are numerous phase transitions, for example, along the line $h + J_1 + J_2 = 0$ where $\varepsilon_{k=0} = 0$. When $h - J_1 + J_2 = 0$ then $\varepsilon_{k=\pi} = 0$. Additionally, when $h - J_2 = 0$ with $|J_1| < 2|h|$ then $\varepsilon_k = 0$ for $k = \pi - \cos^{-1}(J_1/(2J_2))$. This leads to the phase diagram seen in Fig. 3.1. The phases are identified through line cuts through parameters space leading to Hamiltonians with well-known phases [45]. There are three distinct phases, a paramagnet phase, a ferromagnetic phase, and a cluster-state-like phase. The red curve of Fig. 3.1 shows the path in parameter space corresponding to $h(\beta)$, $J_1(\beta)$, and $J_2(\beta)$ for some β. The $\beta = 0$ point lies at the origin, and $\beta = \infty$ at $J_1/h = 2, J_2/h = 1$. Only along this path does the ground state of the Hamiltonian correspond to the qsample state of Eq. 3.1. At all phase transitions besides the tricritical point at $J_1/h = 2, J_2/h = 1$, the dispersion relation is

linear, as seen in Fig. 3.2. Thus all other phase transitions have a dynamic critical exponent of $z = 1$.

3.2.2 Adiabatic Paths

Following [44], we will compare four different adiabatic paths through the extended parameter space of H_q. All methods will begin in the ground state of H_q at $J_1/h = J_2 = 0$, which corresponds to the $\beta = 0$ encoded distribution state. This is the equally mixed state, which is assumed to be easy to prepare. In order for the final state to describe a Boltzmann distribution of the classical Hamiltonian, the adiabatic path must return to the $h(\beta)$, $J_1(\beta)$, $J_2(\beta)$ curve for some β. Since this problem is only difficult at low temperatures the paths we will compare all end at $\beta = \infty$, the tricritical point of H_q.

The four different paths are shown in Fig. 3.3, with path ii) corresponding to the path where $h(\beta)$, $J_1(\beta)$, and $J_2(\beta)$ depend on β. These paths are parameterized by s, which is zero at the starting point, at one at the final point. The explicit parametrization of these paths is listed in Table. 3.1.

Path	J_1/h	J_2/h
i)	$2s$	s
ii)	$2s$	s^2
iii)	$3(1-s)^2s+7.5(1-s)s^2+2s^3$	$1.5(1-s)s^2+s^3$
iv)	$6(1-s)^2s+9(1-s)s^2+2s^3$	$-3(1-s)^2s+4.5(1-s)s^2+s^3$

Table 3.1: Ising-chain adiabatic paths parametrization in terms of $s \in [0, 1]$

To compare the adiabatic time along these different paths, we will fix a target fidelity of the final state. This is the overlap of the adiabatically evolved state and the wanted encoded distribution state,

$$\mathcal{F} = |\langle \psi(\beta)| \phi(t_{\text{tot}})\rangle|^2. \tag{3.15}$$

The final state $\phi(t_{\text{tot}})$ is found by numerically integrating the Schrodinger equations under the constraints of Eq. 3.3. The details of this calculation can be found in Appendix B. For each path an infidelity cutoff of $1 - \mathcal{F} < 10^{-3}$ is used to determine ε. The infidelity dependence on ε can be seen in Fig. 3.4 for the four paths considered.

Once an ε is fixed for each path, the state preparation time t_a is the shortest time needed to obtain $1-\mathcal{F} < 10^{-3}$. Since the adiabatic metric can be calculated, the adiabatic

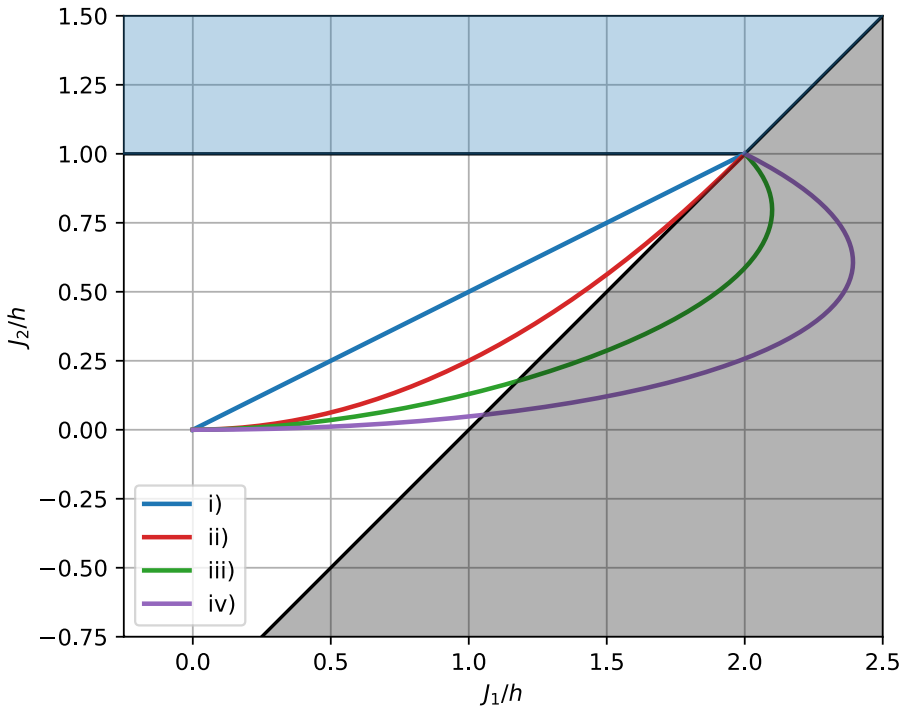

Figure 3.3: Four adiabatic paths in the parameter space of H_q. All methods begin at $H_q(\beta = 0)$, $(J_1/h = J_2/h = 0)$, and end at $H_q(\beta = \infty)$, $(J_1/h = 2, J_2/h = 1)$, in order to prepare the zero temperature qsample.

path length l from Eq. 3.8 can be found. With the values of ε found for each path, the state preparation time is given by

$$t_a = \frac{l}{\varepsilon}. \tag{3.16}$$

Fig. 3.5 shows the scaling with the system size of the adiabatic preparation time for each of the paths considered. The dependency on the system sizes matches the results of [45]. As the found ϵ is roughly independent of system size, this same scaling is found when considering the adiabatic path length l. Path ii), which corresponds to the Glauber mixing matrix, shows scaling with $\sim n^2$. This matches the analytic scaling of the spectral gap of the mixing matrix at low temperature found in Appendix B. The scaling of the paths is

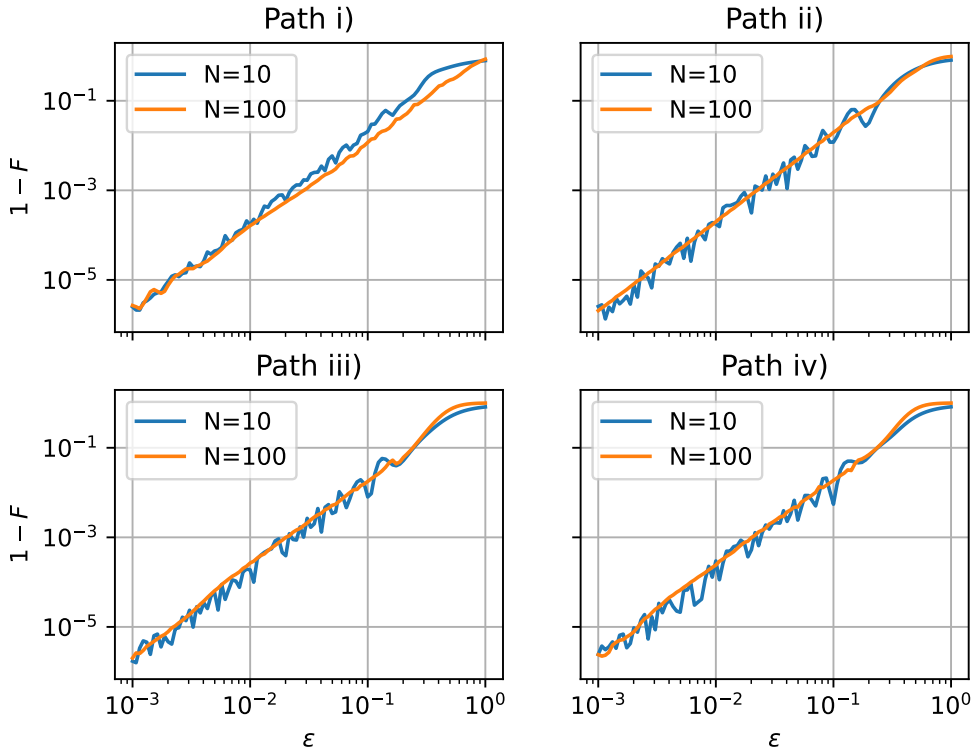

Figure 3.4: Infidelity of final prepared state after adiabatic evolution specified by ε for all four paths. Two system sizes are shown $N = 10$ and $N = 100$.

dominated by their behaviour around critical points.

In Fig. 3.6, the integrand of l is plotted for paths ii) and iii). Path iii) has an increase of this integrand over the phase transition from the paramagnet phase into the ferromagnetic phase and then as the path approaches the tricritical point. However, its behaviour around the tricritical point is different from the classical path ii). Paths which first cross over to the ferromagnetic phase before approaching the critical point show quadratic improvement in Fig. 3.5. This can be understood by considering how domain walls are removed from the paramagnetic state for the different paths. Along path ii) this process occurs while approaching the tricritical point. In paths iii) and iv), these domain walls are removed when crossing to the ferromagnetic phase. These phase transitions have different characteristics, most importantly different dynamic critical exponents z. By definition, Eq. 3.14, z describes how the energy gap scales with the correlation length ξ. For finite systems, at

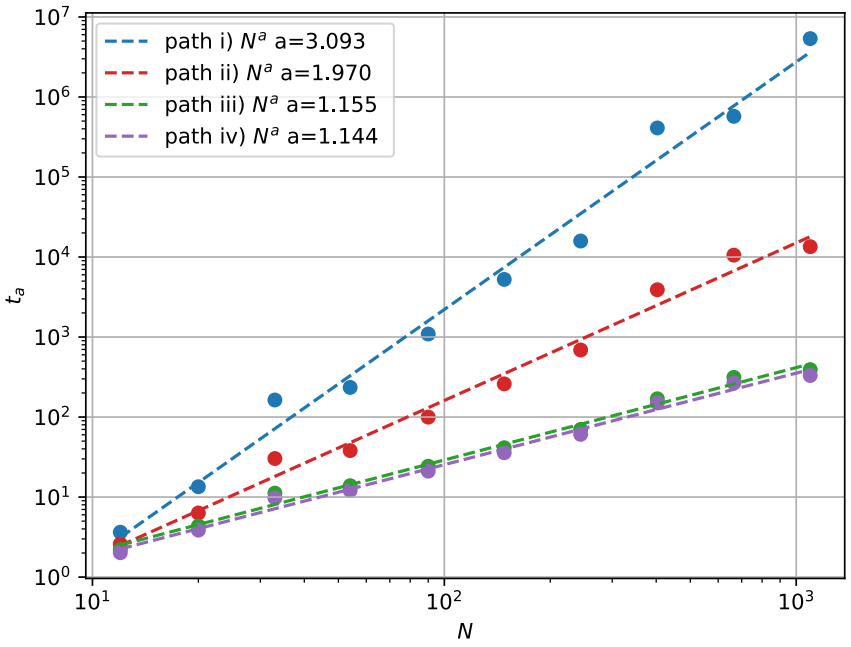

Figure 3.5: Adiabatic state preparation time scaling of the four paths considered. All adiabatic schedules are chosen to ensure $1 - \mathcal{F} < 10^{-3}$. The scaling is linearly fit to showcase the expected quadratic speedup between path iii) and iv) as compared to the classical scaling of path ii)

criticality, $\xi \sim N$ therefore, z describes how the energy gap scales with the system size. Thus, z also indicates the scaling of the slowdown needed to cross the phase transition adiabatically. Since the paramagnet to ferromagnet transition has $z = 1$ the domain walls can be removed with much better scaling in path iii) and iv) than in path ii) which is dependent on the $z = 2$ tricritical behaviour. The scaling of path iii) is also dependent on the behaviour of the phase transition out of the paramagnetic phase. However, unlike the other phase transitions the spectral gap can close exactly for finite-size systems during the transition from the paramagnetic to cluster-state-like phase [45]. This effect accounts for the additional slowdown observed for path iii).

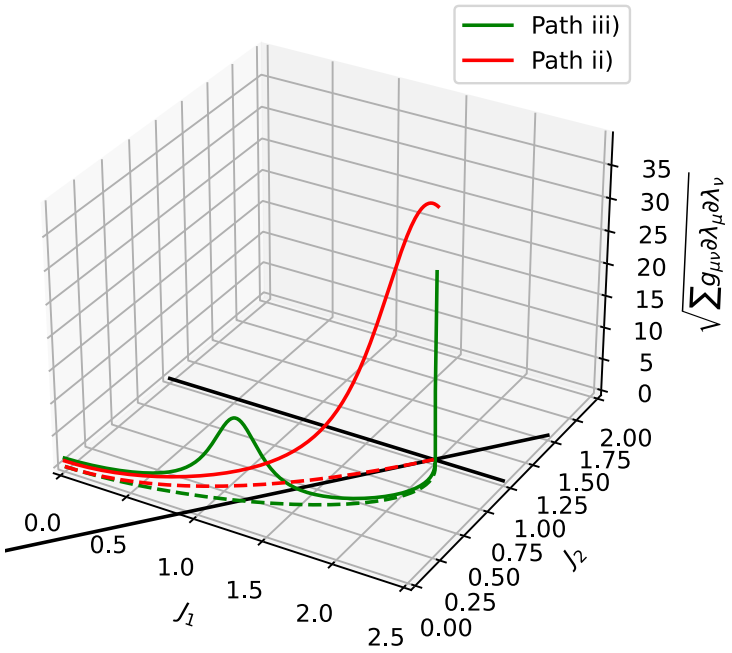

Figure 3.6: Adiabatic path length integrand plotted along path ii) and iii) for $N = 50$ showing different behaviour approaching the tricritical point.

3.3 Discussion

The results of [44], reproduced in Section. 3.2.2, show a quadratic improvement in adiabatic state preparation time using information of the phase transition of the parent Hamiltonian. The Ising chain example is quite unique as the parent Hamiltonian can be solved exactly. While this example is an interesting illustration of the connection between the algorithmic performance and the properties of the phase transitions, it is somewhat contrived. Ising chain sampling only becomes difficult at zero temperature and in the limit of large system sizes as only here are there long-range correlations. Away from this limit, classical local strategies mix quickly as there are no long-range correlations. In the zero temperature limit, the task becomes sampling from the ground state manifold, where optimization algorithms can be utilized. Additionally, for this problem, the parent Hamiltonian has a

three-body term, which is not typically experimentally achievable.

These problems are addressed in [44] by considering a variety of other demonstrative examples. For example, two weighted independent set problems were explored; a chain and star graph sampling. The chain graph, which also has slow mixing classically only at zero temperature, showed a similar quadratic speedup as the Ising chain attributed to fast domain wall propagation. The star graph mixes slowly classically at finite temperatures. The parent Hamiltonian of the star graph instead exhibits a first-order quantum phase transition, so the quadratic speed up found through this algorithm is a result of a larger tunnelling rate, not fast domain wall propagation [45]. A similar first-order transition was found when considering the unstructured search problem. All examples considered showed a speedup with an origin depending on the nature of the phase transition.

The original motivation for investigating this adiabatic algorithm was to explore its applicability for sampling from energy-based machine learning models [41]. However, gaining insight into the phases transitions of parent Hamiltonians derived from more complicated classical Hamiltonians is difficult. For example, even extending this algorithm to a higher dimensional Ising model leads to a parent Hamiltonian whose phases can only be investigated through numerical methods such as quantum Monte Carlo. Additionally, it is only on the one-parameter line of $H_q(\beta)$ where H_q is guaranteed to be sign-problem free. Knowledge of the phase transitions of the parent Hamiltonian is necessary for this algorithm, limiting its application.

3.4 Counterdiabatic Driving

The above algorithm's performance is determined by the rate at which paths can be traversed while maintaining the adiabatic condition. Counterdiabatic driving protocols, sometimes known as shortcuts to adiabaticity, drive the system in order to suppress transitions out of the ground state [21]. Here we explore this protocol on the one-parameter line of $H_q(\beta)$ as an alternative to the extended parameter space paths. Unlike the above algorithm, this process does not require knowledge of the phase transitions of H_q. The needed driving terms to counteract these diabatic transitions exactly are often nonlocal. Here we will consider local approximations to the counterdiabatic Hamiltonian for the qsample state preparation problem of the above section. The locality constraint is needed to run this algorithm on near-term devices.

3.4.1 Counterdiabatic Driving Background

Consider a general setting of evolution under the Hamiltonian $H_0(\lambda(t))$, parametrized by some parameter λ. We could equivalently consider a set of parameters $\{\lambda\}$. Instantaneously diagonalizing this Hamiltonian given the eigenstates $|m(\lambda)\rangle$. If λ is varied slowly, maintaining the adiabatic condition, the new instantaneous eigenstates are related to $|m(\lambda)\rangle$ through the adiabatic theorem. A specific unitary transformation can be made between these adiabatically connected states. Gauge potentials are the generators of continuous unitary transformations. The adiabatic gauge potential, $\mathcal{A}_\lambda$, is defined to be the gauge potential associated with the specific unitary made between these adiabatically connected eigenstates.

Consider a moving observer in the instantaneous eigenbasis of H_0. A transformation into this moving frame is given by $|\psi\rangle = U(\lambda)|\tilde{\psi}\rangle$. For this observer, the Hamiltonian is no longer H_0 but it gains an additional term associated with the moving frame of reference. The term that gets added is exactly this adiabatic gauge potential in this moving frame,

$$\begin{aligned} H_0^{\text{eff}} &= U^\dagger H_0 U - i\dot{\lambda} U^\dagger \partial_\lambda U, \\ &= U^\dagger H_0 U - \dot{\lambda} U^\dagger \mathcal{A}_\lambda U, \end{aligned} \tag{3.17}$$

where the gauge potential in the lab frame is,

$$\mathcal{A}_\lambda = i\partial_\lambda. \tag{3.18}$$

The form of Eq. 3.17 is found by solving the Schrodinger equation in the moving frame. Since we have defined the unitary U to be a transformation into the instantaneous eigenbasis of H_0, the term $U^\dagger H_0 U$ is diagonal. Thus it is the term involving the adiabatic gauge potential which causes non-adiabatic transitions. With this in mind, an alternative Hamiltonian can be constructed which cancels out this extra term in the moving frame. This is exactly the counterdiabatic Hamiltonian,

$$H_{\text{CD}}(t) = H_0 + \dot{\lambda}\mathcal{A}_\lambda. \tag{3.19}$$

The addition of the driving term allows arbitrary fast state preparation protocols. If the system is initialized in the ground state, the driving term counteracts any diabatic transitions that occur if the parameter λ are varied non-adiabatically, and the system will remain in the ground state. It can be shown that the adiabatic gauge potential defined for the Hamiltion H_0 satisfies the following equation [21],

$$[i\partial_\lambda H_0 - [\mathcal{A}_\lambda, H_0], H_0] = 0. \tag{3.20}$$

Although solving this equation does not require diagonalizing the Hamiltonian it is often difficult to solve and in some cases does not yield a solution [21]. It can be shown that this condition can be reformulated as a minimization problem, allowing approximate gauge potentials to be found variationally [37]. Consider the operator G_λ defined as,

$$G_\lambda(\mathcal{X}) = \partial_\lambda H_k + i\left[\mathcal{X}, H_k\right]. \tag{3.21}$$

Then we have,

$$G_\lambda(\mathcal{A}_\lambda) = -\sum_n \frac{\partial E_n}{\partial \lambda} |n(\lambda)\rangle\langle n(\lambda)| := -M_\lambda. \tag{3.22}$$

Finding $\mathcal{A}_\lambda$ is equivalent to minimizing the operator distance between $G_\lambda(\mathcal{X})$ and $-M_\lambda$ with respect to $\mathcal{X}$. Using the Frobenius norm, it can be shown that minimizing this distance is equivalent to minimizing the norm of $G_\lambda(\mathcal{X})$ [21]. In all, we can now consider an action associated with the gauge potential,

$$S = \mathrm{Tr}\left[G_\lambda^2(\mathcal{X})\right]. \tag{3.23}$$

With this problem now mapped to a minimization of this action, we can consider a variational form of $\mathcal{A}_\lambda$ under constraints such as locality [37].

3.4.2 Ising Chain

With this formalism, we can now return to the task of preparing the encoded distribution state for the Ising chain. Since we can solve the parent Hamiltonian of Eq. 3.12 through a Jordan-Wigner transformation, we can find the counterdiabatic Hamiltonian exactly [11]. The Jordan-Wigner solution, derived in Appendix B, leads to the following representation of H_q in terms of momentum space fermionic operators,

$$H_q = -2\sum_k \begin{pmatrix} a_k^\dagger & a_{-k} \end{pmatrix} \begin{pmatrix} A_k & -iB_k \\ iB_k & -A_k \end{pmatrix} \begin{pmatrix} a_k \\ a_{-k}^\dagger \end{pmatrix} = 2\sum_k \psi_k^\dagger H_k \psi_k, \tag{3.24}$$

where $A_k = -h - J_1\cos(k) - J_2\cos(2k)$ and $B_k = J_1\sin(k) + J_2\sin(2k)$. We can consider each of these k blocks separately since they are non-interacting, and the problem reduces to a set of uncoupled two-level systems,

$$H_k = -2\begin{pmatrix} A_k & -iB_k \\ iB_k & -A_k \end{pmatrix} = -2A_k\sigma_k^z - 2B_k\sigma_k^y. \tag{3.25}$$

Consider the following general form of $\mathcal{A}(k)$ in this space,

$$\mathcal{A}(k) = \frac{1}{2}\left(\alpha_x(k)\sigma_k^x + \alpha_y(k)\sigma_k^y + \alpha_z(k)\sigma_k^z\right). \tag{3.26}$$

The exact form of the adiabatic gauge potential can be found by minimizing the action, Eq. 3.23. Let us first find the commutator of this $\mathcal{A}(k)$ with H_k.

$$i\left[\mathcal{A}(k), H_k\right] = 2\left(\alpha_y A_k - \alpha_z B_k\right)\sigma_k^x - 2\alpha_x A_k \sigma_k^y + 2\alpha_x B_k \sigma_k^z \tag{3.27}$$

Recall the paths considered are parameterized by $s \in [0,1]$. To minimize the norm of $G_s(\mathcal{A})$ we require the additional term,

$$\partial_s H_k = -2\partial_s A_k \sigma_k^z - 2\partial_s B_k \sigma_k^y. \tag{3.28}$$

This leads to the following action,

$$\frac{S_s}{4} = \left(\alpha_y A_k - \alpha_z B_k\right)^2 + \left(\partial_s B_k + \alpha_x A_k\right)^2 + \left(\alpha_x B_k - \partial_s A_k\right)^2 \tag{3.29}$$

Minimizing with respect to α_x, α_y, and α_z gives the following form of the adiabatic gauge potential for a k block,

$$\mathcal{A}(k) = -\frac{A_k\partial_s B_k - B_k\partial_s A_k}{2\varepsilon_k^2}\sigma_k^x \tag{3.30}$$

where recall $\varepsilon_k^2 = A_k^2 + B_k^2$. When we diagonalized this Hamiltonian in Appendix B we used a Bogoliubov transformation which was a rotation in the σ_k^y, σ_k^z plane. Here the adiabatic gauge potential is the σ_k^x direction, as it is a generator of the diagonalizing transformation. The adiabatic gauge potential for the entire system is then,

$$\mathcal{A}_s = -\frac{1}{2}\sum_k \frac{A_k\partial_s B_k - B_k\partial_s A_k}{\varepsilon_k^2}\psi_k^\dagger \sigma_k^x \psi_k. \tag{3.31}$$

If we evolved the system under the counter diabatic Hamiltonian of Eq. 3.19 using this adiabatic gauge potential, we would be able to stay in the ground state of the system no matter how quickly the Hamiltonian parameters are varied. For the problem at hand, this would give perfect fidelity at $s = 1$ no matter the value of ds/dt. Although this $\mathcal{A}_s$ is compact in the momentum space representation when we transfer back into the original spin variables we get something highly non-local, and therefore not experimentally feasible.

$$\psi_k^\dagger \sigma_k^x \psi_k = a_k^\dagger a_{-k}^\dagger + a_{-k}a_k = \frac{i}{N}\sum_{j,l}\sin(lk)\left(a_j^\dagger a_{j+l}^\dagger + a_{j+l}a_j\right) \tag{3.32}$$

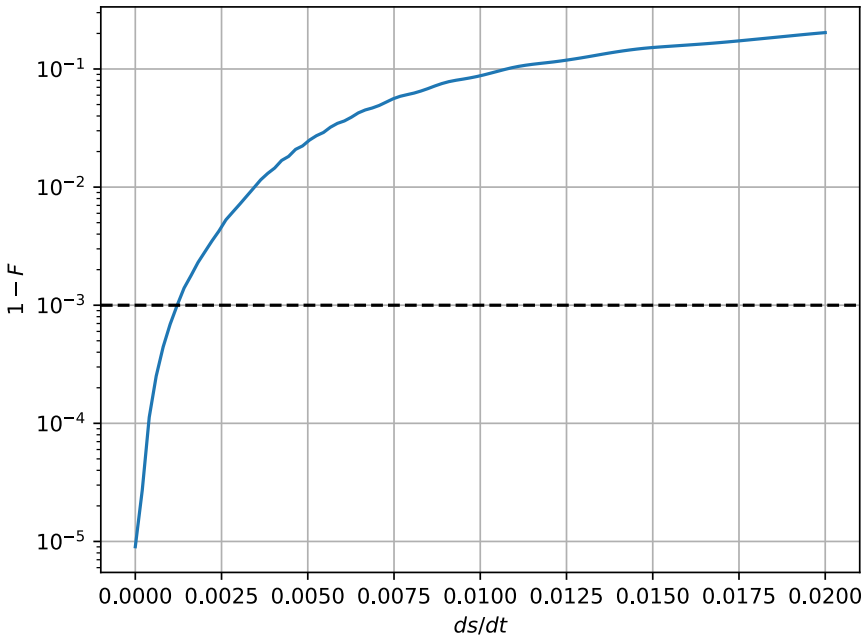

Figure 3.7: Infidelity with ds/dt of approximated counter adiabatic evolution for $N = 10$ with maximum string length $M = 3$.

In the real-space spin variables the adiabatic gauge potential is,

$$\mathcal{A}_s = \sum_{l=1}^{N} \alpha_l O_l, \tag{3.33}$$

where

$$\alpha_l = -\frac{1}{4N} \sum_k \frac{\sin(lk)\left(A_k \partial_s B_k - B_k \partial_s A_k\right)}{\varepsilon_k^2} \tag{3.34}$$

and the operators O_l are Pauli strings of length l,

$$O_l = 2i \sum_j \left(a_j^\dagger a_{j+l}^\dagger + a_{j+l} a_j \right) = \sum_j i \left(\sigma_j^y \sigma_{j+1}^x \sigma_{j+2}^x \dots \sigma_{j+l-1}^x \sigma_{j+l}^y - \sigma_j^z \sigma_{j+1}^x \sigma_{j+2}^x \dots \sigma_{j+l-1}^x \sigma_{j+l}^z \right). \tag{3.35}$$

We will consider the parent Hamiltonian along path ii) whose ground state corresponds

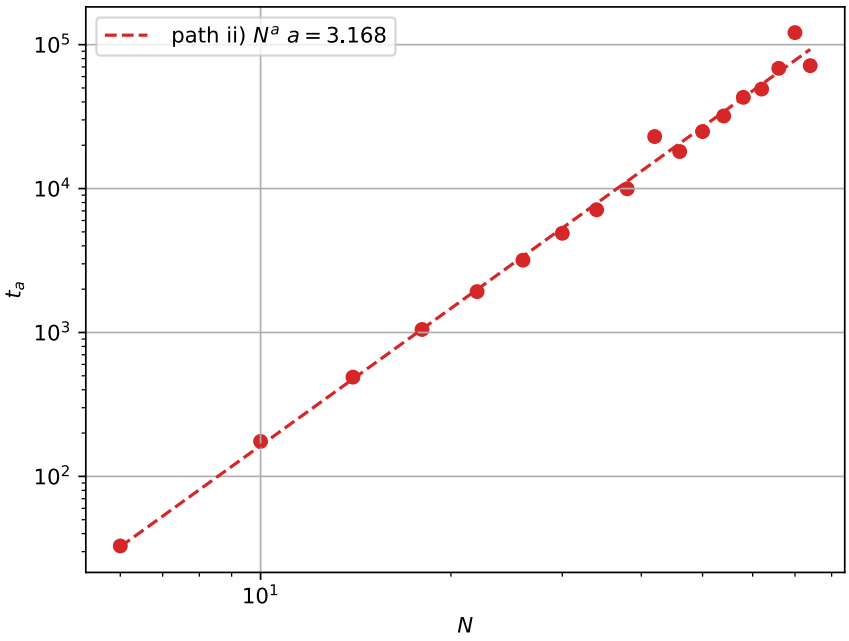

Figure 3.8: Counter diabatic state preparation time with approximated Hamiltonian with a fixed string length of $M = 3$.

to the encoded distributed state at some β. This path's parametrization is listed in Table. 3.1, where $s = \tanh(\beta)$. For this path we can make some simplifications,

$$\begin{aligned} A_k \partial_s B_k - B_k \partial_s A_k &= \left(-1 - 2s\cos(k) - s^2\cos(2k)\right)\left(2\sin(k) + 2s\sin(2k)\right) \\ &\quad - \left(2s\sin(k) + s^2\sin(2k)\right)\left(-2\cos(k) - 2s\cos(2k)\right) \\ &= -2\sin(k)\left(2s\cos(k) + s^2 + 1\right), \end{aligned} \tag{3.36}$$

$$\varepsilon_k^2 = A_k^2 + B_k^2 = \left(2s\cos(k) + s^2 + 1\right)^2. \tag{3.37}$$

For path ii) we then have the following coefficients on the string operators,

$$\alpha_l = -\frac{1}{2N}\sum_k \frac{\sin(k)\sin(lk)}{2s\cos(k) + s^2 + 1}. \tag{3.38}$$

Away from the critical point, these α_l decay exponentially with l, but at the critical point all the strings are relevant. These strings are difficult to implement experimentally, so we

can consider an approximation to this gauge potential and therefore the counter-diabatic Hamiltonian by taking only spin strings up to length M,

$$\mathcal{A}'_s = \sum_{l=1}^{M} \alpha_l O_l. \tag{3.39}$$

Return to the k space representation to see the effect of this truncation,

$$\mathcal{A}'_s = 4 \sum_k \left(\sum_{l=1}^{M} \alpha_l \sin(lk) \right) \psi_k^\dagger \sigma_k^x \psi_k. \tag{3.40}$$

Let us now consider evolving under the counter diabatic Hamiltonian of Eq. 3.19 with this approximation to the adiabatic gauge potential. As it is an approximation, some diabatic transitions will occur which limits the rate at which the paths can be traversed.

Following the procedure from the previous section, the fidelity of the final state Eq. B.34 is found as a function of allowed excitations. In this case, instead of limiting transitions with ε, they are limited by the rate of change of s. Fig. 3.7 shows the infidelity as a function of ds/dt for $N = 10$ with a fixed string length of $M = 3$. For each system size considered, the value of ds/dt is found to give an infidelity of 10^{-3}. With this rate of change, the adiabatic state preparation time can be found. The scaling of this time with system size is shown in Fig. 3.8. The nature of this scaling needs to be further investigated, but it shows unfavourable dependence on the system size as compared to the classical MCMC to sample at zero temperature.

This method of counterdiabatic driving gives a protocol to prepare the qsample state instantaneously but at the cost of Pauli strings with a length the size of the system. These types of interactions are difficult experimentally therefore a local approximation to the driving terms was made. Unfortunately, the adiabatic state preparation time of this local approximation scales as $\sim N^3$ as compared to the $\sim N^2$ scaling of the classical method. There are many possible directions for further research regarding this protocol. The nature of this scaling needed to be further investigated to determine if this scaling disadvantage would be present in other systems. This method was tested only on the Ising chain, where the adiabatic gauge potential can be calculated exactly. However, this protocol can be adapted to other more complicated systems as the gauge potential can be found variationally through minimizing the action of Eq. 3.23 [37].

Chapter 4

Quantum Enhanced Markov Chain Monte Carlo

The next algorithm considered, a quantum enchanced MCMC method [23], is also well suited for current quantum devices. This hybrid method constructs a Markov chain on the classical configuration space but uses a quantum evolution to propose new configurations. A typical method, such as Metropolis-Hastings, accepts or rejects this proposed configuration. This construction maintains detailed balance, ensuring the convergence to the desired distribution. Unlike the quantum algorithms of the previous chapters, this algorithm does not produce qsamples, but classical samples. This algorithm has shorter mixing times than standard classical methods for spin glass problems in the low-temperature regime, seen numerically and experimentally on small system sizes [23]. These results will be numerically reproduced here, and analytic bounds are found for simpler systems to investigate the nature of this speedup.

4.1 Algorithm Overview

The only quantum step of the quantum-enhanced Markov chain Monte Carlo algorithm [23] is to propose new spin configurations. Specifically, the current classical state of the MCMC is prepared as a computational basis state $|\vec{\sigma}\rangle$, evolved unitarily, and then measured in the computational basis. This procedure gives the following proposal probability,

$$Q(\vec{\sigma}'|\vec{\sigma}) = |\langle\vec{\sigma}'|U|\vec{\sigma}\rangle|^2.$$

The focus will be on the evolution $U = e^{-iHt}$ with the following time-independent Hamiltonian,

$$H = (1-\gamma)\alpha H_c + \gamma H_{\text{mix}}, \tag{4.1}$$

with

$$\alpha = \frac{\|H_{\text{mix}}\|_F}{\|H_c\|_F} \qquad H_{\text{mix}} = \sum_i \sigma_i^x. \tag{4.2}$$

Here α maintains a common energy scale between the two terms of the Hamiltonian.

4.1.1 Convergence

Recall in Section 2.1.1 a sufficient condition for π to be a stationary state of P is if the chain satisfies the detailed balance condition. This quantum algorithm satisfies the detailed balance condition if the proposal probability is symmetric,

$$Q(\vec{\sigma}'|\vec{\sigma}) = |\langle\vec{\sigma}'|U|\vec{\sigma}\rangle|^2 = Q(\vec{\sigma}|\vec{\sigma}'). \tag{4.3}$$

This is generally not the case for any arbitrary unitary U. However, it is satisfied in this algorithm as the Hamiltonian of Eq. 4.1 is symmetric for any H_c. There are many other choices of U that would maintain this condition. Recall the detailed balance condition requires $\pi(x)P(x,y) = \pi(y)P(y,x)\ \forall x, y \in S$. The case when $x = y$ is trivial, so assume $x \neq y$. Since the proposal probability is symmetric, this condition becomes $\pi(x)A(y|x) = \pi(y)A(x|y)$.

$$\begin{aligned}
\pi(x)A(y|x) &= \pi(x)\min\left(1, \frac{\pi(y)}{\pi(x)}\frac{Q(x|y)}{Q(y|x)}\right) \\
&= \pi(x)\min\left(1, \frac{\pi(y)}{\pi(x)}\right) \\
&= \pi(y)\min\left(1, \frac{\pi(x)}{\pi(y)}\right)
\end{aligned} \tag{4.4}$$

Where the last line can be seen by considering the two cases of $\pi(x) > \pi(y)$ and $\pi(x) < \pi(y)$. In order to prove the convergence to this steady state π, the chain must be aperiodic and irreducible. Both of these conditions are satisfied if $Q(\vec{\sigma}'|\vec{\sigma}) > 0\ \forall\vec{\sigma}', \vec{\sigma}$ [23]. However, the proposal probability is not greater than zero for all choices of Hamiltonian parameters in Eq. 4.1. For example, if $\gamma = 0$ the system is initialized in an eigenstate of H and thus remains in that state after the evolution. In this case, the MCMC would never propose new states and thus not mix. In later sections, the performance of the algorithm will be

investigated for different Hamiltonian parameters. We will see that many Hamiltonian parameters lead to short mixing times, justifying the procedure of [23] of choosing them at random at each step of the chain. This strategy additionally ensures that the chain remains irreducible.

Errors in the physical implementation of this evolution may break the symmetry of the proposal probably, a condition needed for convergence. It is only errors that break this symmetry which bias the MCMC sampling. Other errors may lead to longer mixing times, but the algorithm will still converge to the wanted distribution. In the physical implementation of this algorithm in [23], the quantum evolution is approximated through a Suzuki-Trotter expansion. This process induces errors which break the needed symmetry and are therefore randomized through a process of SPAM twirling [23]. This method restores the needed symmetry. Error effects on this algorithm are not studied in detail in this thesis but left as a possible extension. However, with this SPAM twirling addition when this algorithm is implemented on near-term error-prone devices, convergence to the wanted distribution is still achieved.

4.1.2 Hamiltonian Monte Carlo

Before exploring the performance of this algorithm, we can gain some intuition for why it may lead to a performance increase by relating it to Hamiltonian Monte Carlo (HMC) a successful classical algorithm for sampling continuous distributions[6]. Consider a possible target distribution $\pi(x)$ with $x \in \mathbb{R}^d$. HMC considers a different distribution, f, from which to sample in order to produce samples of $\pi(x)$ with a higher acceptance rate. These samples are proposed by running Hamiltonian dynamics for some time τ. Specifically, the Hamiltonian considered consists of a potential determined by the wanted distribution and the addition of a canonical momentum $p \in \mathbb{R}^d$,

$$H(x,p) = V(x) + \frac{1}{2}p^2, \tag{4.5}$$

where $V(x) = -\log(\pi(x))$. The Boltzmann distribution of this Hamiltonian at $\beta = 1$,

$$f(x,p) \propto e^{-H(x,p)}, \tag{4.6}$$

will be the distribution HMC samples from. From the definition of the Hamiltonian $H(x,p)$ it can be seen that the marginal distribution of the position variable is the desired $\pi(x)$. Starting at some (x, p_0) a new sample is generated by HMC with the following procedure.

1. Draw a random momentum vector, p, typically chosen to be a multivariate Gaussian.

2. Evolve the system under Hamiltonian dynamics starting at (x, p) for some time τ, specifically following
$$\frac{dx_i}{dt} = \frac{\partial H}{\partial p_i}, \qquad \frac{dp_i}{dt} = \frac{\partial H}{\partial x_i}. \tag{4.7}$$
The resulting point after this evolution is (x', p').

3. Accept or reject the new sample using the Metropolis-Hastings acceptance probability,
$$A((x', p')|(x, p)) = \max\left(1, \frac{e^{-H(x',p')}}{e^{-H(x,p)}}\right). \tag{4.8}$$

What is important about this process is that energy is conserved throughout the Hamiltonian dynamics so new samples are always accepted with unit probability. However, this evolution when done numerically introduces slight errors. This is why the Metropolis-Hastings acceptance step is then added to account for these added errors and ensure convergence. With some choice of τ depending on the problem, this algorithm allows one to generate samples far away in state space that are almost guaranteed to be accepted, leading to fast mixing times [6].

There are many parallels between HMC and the quantum-enhanced MCMC algorithm. In the quantum setting, we still have continuous Hamiltonian dynamics which conserve energy, but no longer have commutation of the canonical variables. Additionally, our target distribution is now over discrete variables. These complications mean in this setting we lose the guaranteed acceptance after evolution. However, it is still possible that it is beneficial to consider proposal through quantum evolution as we may obtain samples which are more likely to be accepted. We will see that this is the case for some problems at low temperatures. The quantum-enhanced MCMC algorithm is able to move between samples $\vec{\sigma} \to \vec{\sigma'}$ which are far in Hamming distance but are close is energy, therefore they are likely to be accepted.

4.2 Sherrington-Kirkpatrick Model

To numerically verify the results found in [23], the first model considered is a fully connected spin glass,
$$H_c = -\sum_{j>i=1} J_{ij}\sigma_j^z\sigma_i^z + \sum_i h_i\sigma_i^z, \tag{4.9}$$

where the parameters $\{J_{ij}\}_{j>i=1}^{N}$ and $\{h_i\}_{i=1}^{N}$ are drawn from a normal distribution with a mean of zero and a standard deviation of one. This model is a slightly modified version of the well-known Sherrington-Kirkpatrick (SK) model [38], as the normal distribution has a standard deviation of one, instead of the typical value of $1/\sqrt{N}$. This choice is made to compare the scaling of models with different connectivity more easily. Additionally, this Hamiltonian has random local fields $\{h_i\}_{i=1}^{N}$, which are not present in the first presentation of the SK model. This additional term further complicates the problem; a fully-connected N spin problem with local fields is equivalent to a $N+1$ fully-connected model without local fields. These models often have many local minima that are far in Hamming distance, making them difficult to sample from. At zero temperature, sampling from a Boltzmann distribution produces samples from the ground state manifold of H_c. For spin glass models, finding the ground state is NP-hard [4].

4.2.1 Numerical Results

The mixing time of this quantum algorithm can be probed by numerically finding the spectral gap δ of the associated transition matrix P. The configuration space grows exponentially with the system size, therefore this type of numerical investigation is limited to small system sizes.

As in [23], I will compare this quantum algorithm to two different classical proposal strategies. For the Ising model, other updates lead to faster mixing times, such as cluster algorithms [36]. However, these types of updates are not advantageous for more complex spin glasses problems. For these harder problems, typically annealing or parallel tempering methods are employed. Here we will compare constant temperature updates, with the possibility of expanding to these varying temperature methods.

The first is a uniform strategy which proposes a new configuration at random. In this case, we have $Q(\vec{\sigma}'|\vec{\sigma}) = 1/2^N$ for all $\vec{\sigma}', \vec{\sigma}$. These updates allow movement to configurations far in Hamming distance from the previous configuration. The second strategy, the local strategy, proposes a new configuration by flipping one of the N spins. In this case, $Q(\vec{\sigma}'|\vec{\sigma}) = 1/N$, for $\vec{\sigma}'$ and $\vec{\sigma}$ separated by a single spin flip, and zero otherwise. Due to the local structure of the Hamiltonian, this new configuration will be close in energy to the previous.

For all algorithms, quantum and classical, the mixing time of the chain depends on the system size N and the temperature $T = 1/\beta$. We will use the same acceptance probability in all cases, the Metropolis-Hastings probability of Eq. 2.14. Due to the symmetric

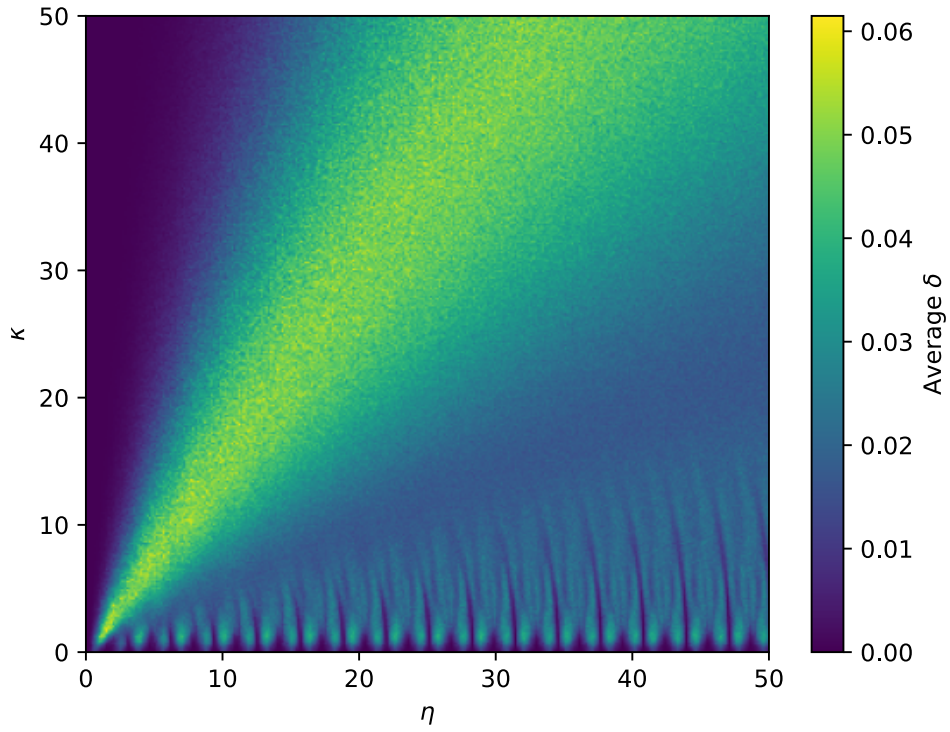

Figure 4.1: Spectral gap δ of quantum enhanced MCMC for $N = 7$ averaged over 100 instances of the SK model at $\beta = 5$.

proposals, this probability further simplifies to,

$$A(\vec{\sigma}'|\vec{\sigma})_{MH} = \min\left(1, e^{-\beta(H_c(\vec{\sigma}')-H_c(\vec{\sigma}))}\right). \tag{4.10}$$

Glauber acceptance probability was simulated for all methods but led to insignificant performance differences. The proposal strategy of this quantum algorithm involves evolving under the following Hamiltonian,

$$\begin{aligned} tH &= t\left((1-\gamma)\alpha H_c + \gamma H_{\text{mix}}\right) \\ &= \kappa\alpha\left(-\sum_{j>i=1} J_{ij}\sigma_j^z\sigma_i^z + \sum_i h_i\sigma_i^z\right) + \eta\sum_{i=1}^{N}\sigma_i^x \end{aligned} \tag{4.11}$$

Here I have redefined the evolution parameters γ and t in order to simplify the numerical analysis.

The quantum proposal depends non-trivially on κ and η. This dependence can be seen in Fig. 4.1. In this figure, the spectral gap at low-temperature averaged over 100 $N = 7$ Hamiltonian instances is shown as a function of κ and η. There is a substantial section of this parameter space which leads to a large gap and thus fast mixing of the chain. We can understand the limits of this diagram through perturbation theory.

When η is small we can approximate, to first order in η, an eigenstate of the Hamiltonian of Eq. 4.15 as,

$$\begin{aligned} &\left|n^0\right\rangle + \eta \left|n^1\right\rangle \\ &\left|n^0\right\rangle + \eta \sum_{i=1}^{N} \frac{1}{E_n^0 - E_{n_i}^0} \left|n_i^0\right\rangle \end{aligned} \tag{4.12}$$

Here $|n^0\rangle$ are the eigenstates of H_c, the computational basis states, and $|n_i^0\rangle$ is a configuration which differs from $|n^0\rangle$ by a single spin flip. There are N states $|n_i^0\rangle$, for each $|n^0\rangle$. The evolution on a computation basis state is then given by,

$$e^{-iH} \left|n^0\right\rangle = e^{-iE_n^0} \left|n^0\right\rangle + \eta \sum_{i=1}^{N} \left(e^{-iE_n^0} - e^{-iE_{n_i}^0}\right) \frac{1}{E_n^0 - E_{n_i}^0} \left|n_i^0\right\rangle . \tag{4.13}$$

As we then measure in the computational basis, the resulting proposal reduces to classical local strategy. In this limit, spectral gap scaling with the system size, which will discuss in more detail in a later section, matches that of the local strategy.

In the limit of small α, Fig. 4.1 shows a periodic behaviour with η. We can consider the evolution at $\alpha = 0$, which rotates the spins depending on the value of η. The spectral gap returns to zero at each factor of $\eta = \pi$ as this rotates all spins back to their original position. Additionally, at each factor of $\eta = \pi/2$ the spectral gap is also zero. At these values, the resulting state after evolution is the computational basis with all spins flipped compared to the starting state. This makes the chain periodic as it continues to flip between these two configurations. Periodic chains do not have unique stationary distribution, so there are multiple eigenstates with eigenvalues of 1, leading to a spectral gap of zero. At factors of $\eta = \pi/4$ this evolution leaves the system in an equal superposition state over all configurations. Through measurement in the computation basis, there is equal probability in proposing all configurations. At these values, the quantum algorithm reduces to the classical uniform MCMC. This is again verified numerically as the spectral gap scaling

with system size matches the uniform method. At $\alpha = 0$ the value of η that leads to the largest spectral gap is slightly larger than $\eta = \pi/4$ where there is still overlap with many basis states, but the ones with closer energy are slightly favoured. This optimal value of η is dependent on the specific Hamiltonian considered. As seen in Fig. 4.1, this periodic nature continues on average for small κ.

Away from these limits, the structure of Fig. 4.1 is harder to determine. We will see when considering simpler models that analyzing the bottlenecks of this chain will give further insight. However, for this model it is not known why there is a concentrated region of large average δ and if it persists at larger system sizes not accessible numerically.

Temperature Dependence

At low temperatures, the mixing time of the MCMC methods increases. This can be seen as the Metropolis-Hastings acceptance probability of Eq. 2.14 exponentially decreases with β. The chain accepts fewer proposed configurations, causing it to stay at one location in the configuration space and mix slowly. This temperature dependence can be seen in Fig. 4.2. Depending on the energy scale of the Hamiltonian, this plot also shows how the Boltzmann distribution changes the task of sampling. For example, below around $T = 0.5$, sampling from the Boltzmann distribution becomes finding the ground state manifold of the Hamiltonian. At high temperatures, all states become equally weighted, and sampling becomes much easier leading to much larger spectral gaps.

Unlike the other strategies, the local proposal method does not plateau at low temperatures but behaves erratically. When sampling from the ground state manifold, the strategy needs to be able to go between configurations potentially far in Hamming distance, a task that is very difficult for a local proposal strategy. The uniform strategy has an equal probability of proposing all configurations, leading to a plateaued spectral gap. Additionally, the gap of the local strategy decreases at high temperatures. This is because the chain is close to periodic, and an eigenvalue of its transition matrix is approaching -1. This known effect can be easily fixed with a lazy chain strategy, which keeps the chain at the same place with some probability, for example by using Glauber acceptance probability [5].

The quantum strategy depends on the choice κ and η as seen in Fig. 4.1. In order to get an indication of performance when taking random values of $\kappa, \eta \in [0, 50]$ at each step of the MCMC, the geometric mean over this range is plotted in Fig. 4.2 labelled as the random quantum proposal. The error bars on this value show a 95% confidence interval of this geometric mean. Additionally, the best quantum proposal shows the largest gap found in $\kappa, \eta \in [0, 50]$. Finally, the spectral gap was averaged over 100 Hamiltonian instances for

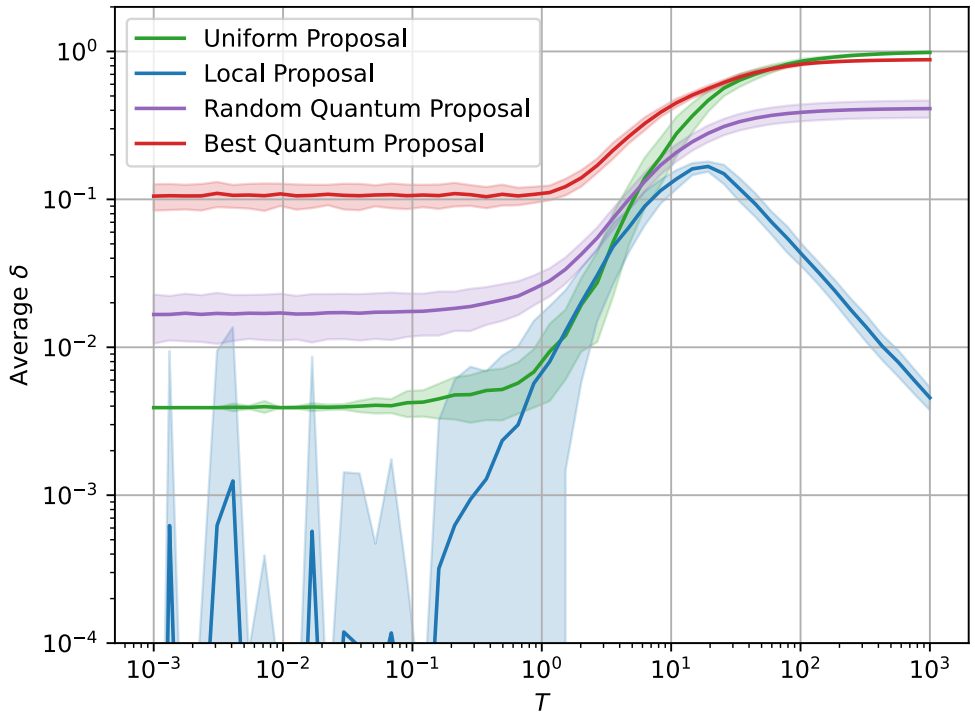

Figure 4.2: Spectral gap δ of classical and quantum strategies for $N = 8$ averaged over 100 instances of the SK model. The random quantum proposal is the geometric mean over $\kappa, \eta \in [0, 50]$ with error bars indicating a 95% confidence interval of this mean. The best quantum proposal shows the largest gap over the same range of κ and η.

all methods. At low temperatures, the quantum-enhanced MCMC method shows a larger spectral gap than the classical methods, showing a faster mixing time of the chain and matching the results of [23].

Low-Temperature System Size Scaling

The quantum proposal strategy is only beneficial compared to the classical method considered in the low-temperature limit. In Fig. 4.3, the spectral gap is shown at $\beta = 5$ for increasing system sizes. Here we numerically verify the results seen in [23], that this algorithm gives an average case polynomial enhancement compared to a uniform proposal

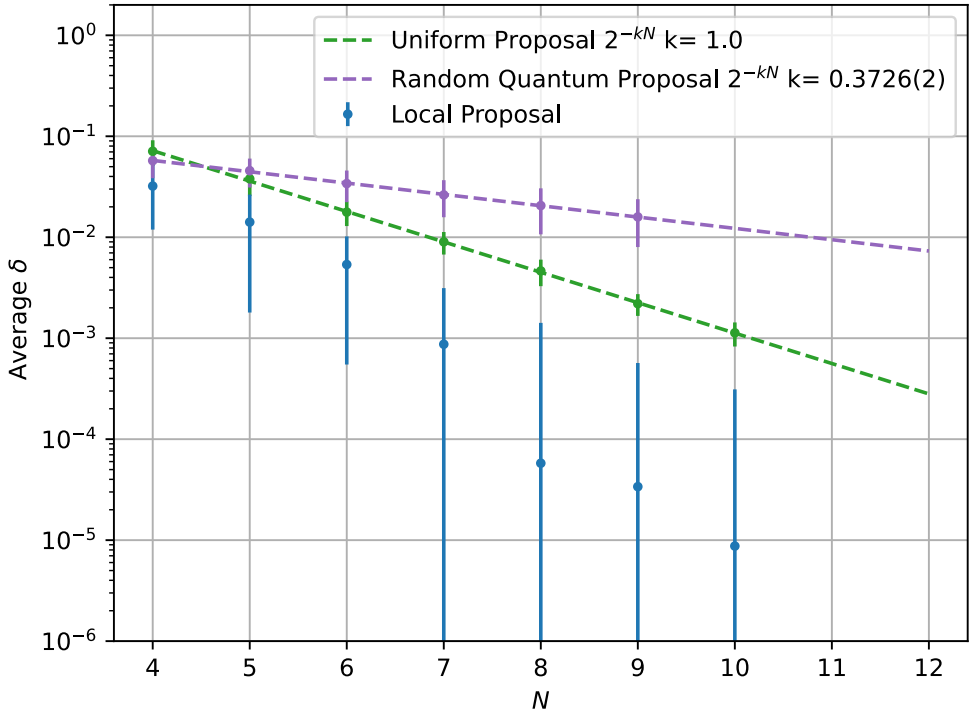

Figure 4.3: Average spectral gap scaling with system size at $\beta = 5$ for 100 instances of the SK model.

strategy for the SK model. The exact improvement depends on the temperature considered, but for all temperatures less than around one, this polynomial scaling improvement persists.

Due to the complexity of the system, it is difficult to bound this mixing time and get an understanding of the source of this improvement. Most critically, it is difficult to say if this improvement persists for larger system sizes which are not accessible numerically. Additionally, Fig. 4.1 shows areas with small κ and η, which still lead to a large spectral gap. This indicates that this improvement exists even with short-time evolution, which can be simulated classically. However, that effect may not persist for larger system sizes. Further understanding of the gap dependence on the Hamiltonian parameter is necessary to describe the generic behaviour of this algorithm. In comparing only the mixing time we are also not accounting for the extra time it takes to evolve quantum mechanically. This

is again related to the strategy used to choose κ and η, and requires further understanding of Fig. 4.1.

4.3 Ising Chain

Although the focus of [23] is on more complex models, a useful simplification is to consider an Ising chain. The quantum evolution associated with this system can be solved analytically, allowing theoretical exploration beyond the system sizes accessible numerically. Consider the N spin ferromagnetic Ising chain with periodic boundary conditions,

$$H_c = -\sum_{i=1}^{N} \sigma_i^z \sigma_{i+1}^z. \tag{4.14}$$

The choice to take periodic boundary conditions is to simplify the analytic scaling arguments but can be easily adapted to other boundary conditions. Using a Peierls argument, it can be shown that this model has no long-range order and hence no phase transition for $T > 0$ [14]. At $T = 0$ the system is ferromagnetically ordered. Due to the Z_2 symmetry of the system, this ground state is doubly degenerate. The proposal strategy of this quantum algorithm involves evolving under the transverse field Ising model,

$$tH = -\kappa \sum_{i=1}^{N} \sigma_i^z \sigma_{i+1}^z + \eta \sum_{i=1}^{N} \sigma_i^x, \tag{4.15}$$

as α defined in Eq. 4.2 equals one for this system.

4.3.1 Numerical Results

Although this quantum evolution can be solved in the thermodynamic limit, this only gives individual elements of transition matrix P. Calculating the spectral gap of P is still intractable for large systems, limiting the numerical exploration. Fig. 4.4 shows the spectral gap dependence on κ and η for the $N = 9$ Ising chain at $\beta = 5$. As in the SK model, the area with a larger gap is concentrated in the lower half of this plot. However, there is more intricate periodic behaviour.

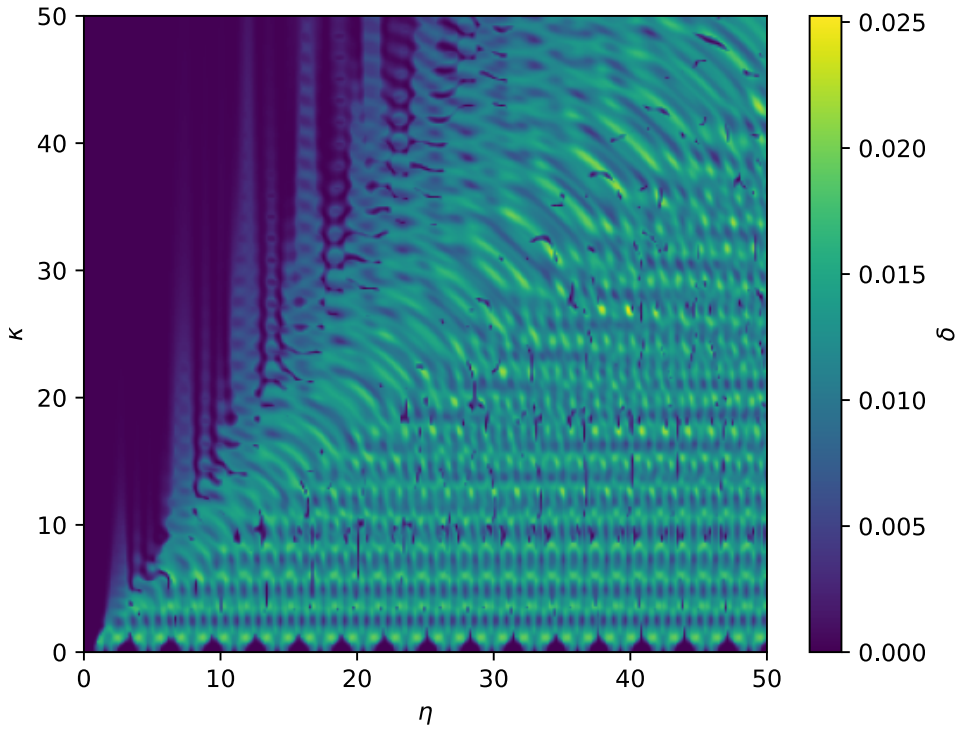

Figure 4.4: Spectral gap δ of quantum enhanced MCMC for $N = 9$ Ising chain at $\beta = 5$.

Temperature Dependence

The temperature dependence of these methods for $N = 8$ are shown in Fig. 4.5, with very similar behaviour as in the SK model. The low-temperature sampling task for this problem involves sampling from the two ferromagnetic states. The failure of the local proposal strategy is evident as in order to transition between these two configurations, a defect needs to be created. This initial step of flipping one spin when the others are aligned is energetically unfavourable. More specifically, it is exponentially (in β) suppressed by the Metropolis-Hastings probability, leading to the continuing decrease of the spectral gap with temperature. As in the SK model, the quantum algorithm has a larger spectral gap in this low-temperature regime.

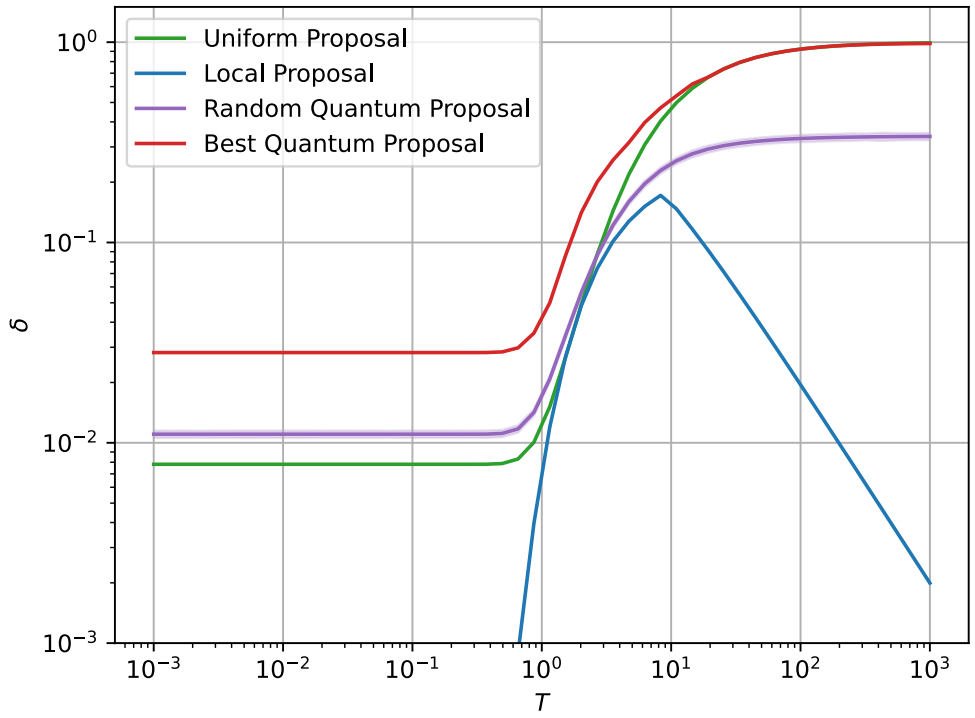

Figure 4.5: Spectral gap δ of classical and quantum strategies for $N = 8$. The random quantum proposal is the geometric mean over $\kappa, \eta \in [0, 50]$ with error bars indicating a 95% confidence interval of this mean. The best quantum proposal shows the largest gap over the same range of κ and η.

Low-Temperature System Size Scaling

Fig. 4.6 show the low-temperature scaling of the quantum and classical methods. This model shows a similar polynomial enhancement for the quantum method as compared to the uniform MCMC. A local proposal strategy is a typical choice for this model due to its polynomial scaling with the system size, which can be seen in this figure. However, as mentioned in the previous section it struggles to move between the Z_2 subspaces, thus exponentially suppressing its spectral gap. In practice, this is not an issue as the chain can be restarted and run multiple times in order to sample from both subspaces. This technique is not captured when considering the algorithm performance based on the spectral gap.

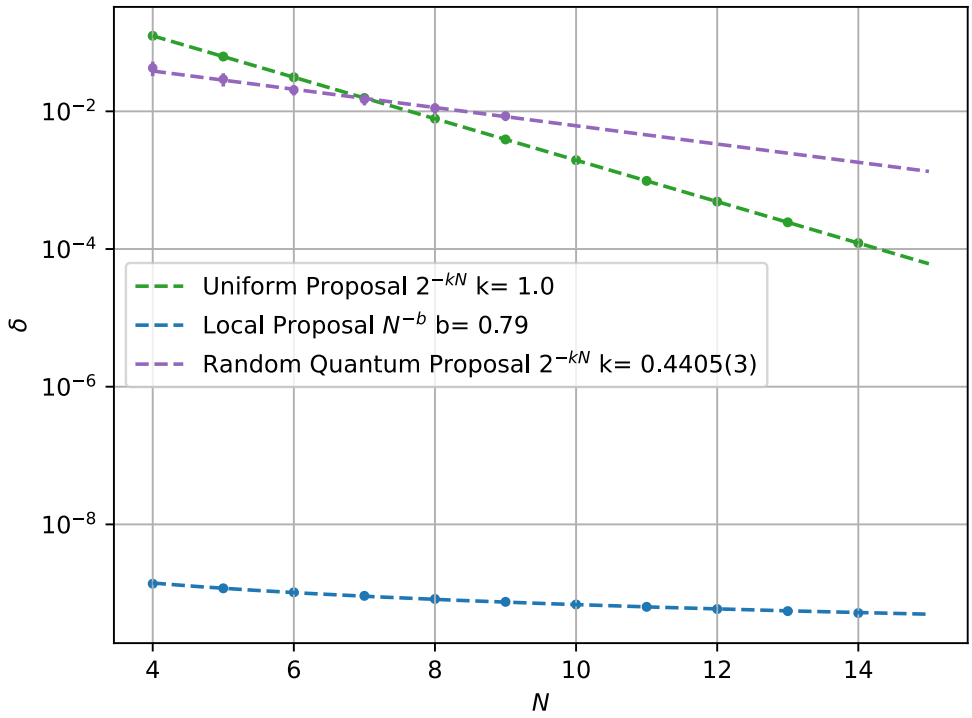

Figure 4.6: Spectral gap scaling at $\beta = 5$. The quantum strategy is averaged over $\kappa, \eta \in [0, 50]$.

We can solve exactly for the gap when we take a local proposal strategy with Glauber acceptance probability. Discussion of this scaling can be found in Appendix B.3. Although that analysis is done with a different acceptance probability (Glauber), both local proposal methods lead to polynomial scaling. The other strategies seen in Fig. 4.6 add additional complications, and their scaling is explored through Cheeger bounds in the next section.

4.3.2 Mixing Time Bound

We can gain further insight into the scaling of this algorithm through a bound of the mixing time. Recall that the Cheeger inequality, Eq. 2.21, provides a lower and upper bound on δ. This inequality requires the bottleneck ratio of the chain, Φ_*, which is defined as a

minimization over subsets A of the state space, such that $\pi(A) \leq 1/2$. This minimization is typically intractable, however, here we are considering the low-temperature limit that allows for some simplifying approximations.

Since the Markov chain updates depend on the energy of the configuration, it is useful to organize the space into subsets of configurations with the same energies. As we have periodic boundary conditions the possible energy values are,

$$E(k) = -N + 4k \tag{4.16}$$

with $k = \{0, 1, .., N/2\}$ for N even and $k = \{0, 1, .., (N-1)/2\}$ for N odd. Let S_k be the subset such that $\forall x \in S_k \; H_c(x) = E(k)$. The ground state subset, S_0, has a cardinally of 2, containing the all up and all down set. More generally,

$$|S_k| = 2\binom{N}{2k}. \tag{4.17}$$

Consider an arbitrary subset A containing elements from $\{S_0, S_1, S_2, S_3, ...\}$. Specifically A contains c_0 elements of S_0, c_1 elements of S_1,... with $0 \leq c_i \leq |S_i|$. Label the states of S_k as $\{x_k^1, x_k^2, ..\}$, with x_k a generic state in this subset. For all the above methods, a Metropolis-Hastings update probability is used. The bottleneck ratio of a generic set A can be expanded by using the problem's structure. Recall we have,

$$\pi(A) = \sum_{i=0} c_i \pi(x_i). \tag{4.18}$$

$$\begin{aligned}
\Phi(A) = \frac{E(A, A^c)}{\pi(A)} &= \frac{1}{\pi(A)} \left(\sum_{x \in A, y \in A^c} \pi(x) A(y|x) Q(y|x) \right) \\
&= \frac{1}{\pi(A)} \Bigg[c_0 \pi(x_0) \Bigg(\sum_{y \in A^c} Q(y|x_0) \min(1, e^{-\beta(H_c(y) - E_0)}) \Bigg) \\
&\quad + c_1 \pi(x_1) \Bigg(\sum_{y \in A^c} Q(y|x_1) \min\left(1, e^{-\beta(H_c(y) - E_1}\right) \Bigg) + ... \Bigg] \\
&= \frac{1}{\pi(A)} \Bigg[c_0 \pi(x_0) \Big[Q(x_0|x_0)(|S_0| - c_0) + Q(x_1|x_0)(|S_1| - c_1) e^{-4\beta} + Q(x_2|x_0)(|S_2| - c_2) e^{-8\beta} + ... \Big] \\
&\quad + c_1 \pi(x_1) \Big[Q(x_0|x_1)(|S_0| - c_0) + Q(x_1|x_1)(|S_1| - c_1) + Q(x_2|x_1)(|S_2| - c_2) e^{-4\beta} + ... \Big] + ... \Bigg]
\end{aligned} \tag{4.19}$$

In the low-temperature limit, many of these terms are exponentially suppressed. Additionally, only the lowest energy states contribute to the distribution π. To find the bottleneck of the entire chain Eq. 4.19 is minimized over any possible subset A.

Uniform Proposal

The bottleneck ratio further reduces for the classical strategy with a uniform proposal probability $Q(y|x) = 1/2^N \forall x, y$.

$$\begin{aligned}
\Phi(A) &= \frac{1}{2^N \pi(A)} \left(\sum_{x \in A, y \in A^c} \pi(x) A(y|x) \right) \\
&= \frac{1}{2^N \pi(A)} \Bigg[c_0 \pi(x_0) \Big[(|S_0| - c_0) + (|S_1| - c_1) e^{-4\beta} + (|S_2| - c_2) e^{-8\beta} + ... \Big] \\
&\quad + c_1 \pi(x_1) \Big[(|S_0| - c_0) + (|S_1| - c_1) + (|S_2| - c_2) e^{-4\beta} + ... \Big] + ... \Bigg] \\
&= \frac{1}{2^N \sum_{i=0} c_i \pi(x_i)} \Bigg[(|S_0| - c_0) \sum_{i=0} c_i \pi(x_i) + (|S_1| - c_1) \sum_{i=1} c_i \pi(x_i) + ... \\
&\quad + \sum_{j=0} c_j \pi(x_j) \sum_{i=1} (|S_{i+j}| - c_{i+j}) e^{-2i\beta} \Bigg]
\end{aligned} \tag{4.20}$$

In the low-temperature limit,

$$\begin{aligned}
\Phi(A) &\sim \frac{1}{2^N \sum_{i=0} c_i \pi(x_i)} \left((|S_0| - c_0) \sum_{i=0} c_i \pi(x_i) + (|S_1| - c_1) \sum_{i=1} c_i \pi(x_i) + ... \right) \\
&= \frac{1}{2^N} \left((|S_0| - c_0) + (|S_1| - c_1) \frac{\sum_{i=1} c_i \pi(x_i)}{\sum_{i=0} c_i \pi(x_i)} + ... \right) \\
&= \frac{1}{2^N} \left(\sum_{j=0} (|S_j| - c_j) \frac{\sum_{i=j} c_i \pi(x_i)}{\sum_{i=0} c_i \pi(x_i)} \right) \\
&= \frac{1}{2^N} \left(\sum_{j=0} (|S_j| - c_j) \frac{\sum_{i=j} c_i e^{-4i\beta}}{c_0 + \sum_{i=1} c_i e^{-4i\beta}} \right) \\
&\sim \frac{1}{2^N} (|S_0| - c_0)
\end{aligned} \tag{4.21}$$

As we have the constraint that $\pi(A) \leq 1/2$, this is minimal if we take A to contain one of our ground states. Thus, $\Phi_\star = \frac{1}{2^N}$, which matches the numerical minimization and shows the scaling seen in Fig. 4.6.

Quantum Proposal

Like the spectral gap δ, the bottleneck ratio of the quantum algorithm depends on κ and η. More specifically, the set A, which minimizes $\Phi(A)$ changes across κ and η. This minimization can be done numerically for small system sizes. As seen in Fig. 4.7, this minimization recovers the specific structure seen in the spectral gap. These patterns emerge by changing between two specific sets A. One, which I will label as A_1, contains just one of the ground states, $A_1 = \{x_0^1\}$. The other does not contain either of the ground states, $A_2^c = S_0$. As higher energy states are further suppressed in Eq. 4.19, it is a valid approximation to take $A_2 = S_1$ when considering large β.

When β is large, the constraint of $\pi(A) \leq 1/2$ limits our choice of A. We have a degenerate ground state, $\pi(x_0) \sim 1/2$. From the form of Eq. 4.19, the two sets for A described above emerge as the leading order contributions. Since we can solve this evolution exactly, we can calculate the bottleneck ratio in both these cases.

We will first investigate the bottleneck ratio with the set $A_1 = \{x_0^1\}$. Again in the low-temperature limit, transitions from x_0^1 into higher energy states are exponentially suppressed due to the Metropolis-Hastings acceptance probability.

$$\begin{aligned}\Phi(A_1) &= \frac{1}{\pi(x_0^1)} \left(\sum_{y \neq x_0^1} \pi(x_0^1) A(y|x_0^1) |\left\langle x_0^1 \right| U \left| y \right\rangle|^2 \right) \\ &= \sum_{y \neq x_0^1} A(x_0^1, y) |\left\langle x_0^1 \right| U \left| y \right\rangle|^2 \\ &\sim |\left\langle x_0^1 \right| U \left| x_0^2 \right\rangle|^2\end{aligned} \tag{4.22}$$

To calculate this overlap, I will make use of the exact solution of the transverse Ising model derived in Appendix A. As $|x_0^1\rangle$ and $|x_0^2\rangle$ denote the two ground states of the classical Ising Hamiltonian, they are also ground states of the transverse field Ising model (TFIM) of Eq. 4.15 when $\kappa = 1$ and $\eta = 0$. Due to the Z(2) symmetry of the problem, the TFIM Hamiltonian can be split into two fermionic parity sectors labelled as $p = 0$ and $p = 1$. The classical states $|x_0^1\rangle$ and $|x_0^2\rangle$ can then be written in terms of the lowest energy states

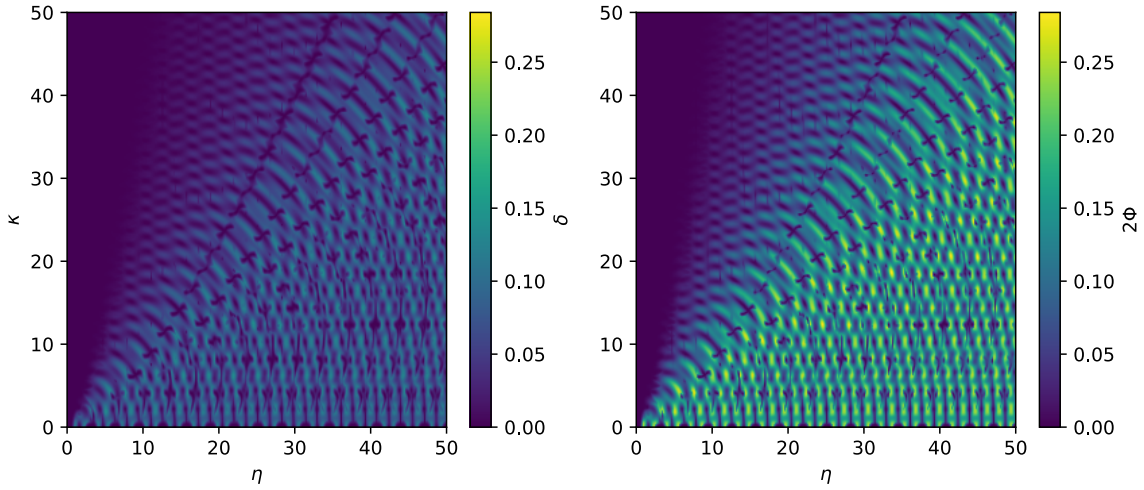

Figure 4.7: Left: Spectral gap of the quantum algorithm for $N = 4$ at $\beta = 5$. Right: Numerically minimized bottleneck for $N = 4$ at $\beta = 5$. The structure seen here comes from transitions between the sets A_1 and A_2.

of the $p = 0$ and $p = 1$ sectors, labelled $|gs_0\rangle$ and $|gs_1\rangle$, respectively.

$$\begin{aligned} \left|x_0^1\right\rangle &= \frac{1}{\sqrt{2}}\left(|gs_0\rangle + |gs_1\rangle\right) \\ \left|x_0^2\right\rangle &= \frac{1}{\sqrt{2}}\left(|gs_0\rangle - |gs_1\rangle\right) \end{aligned} \tag{4.23}$$

Expanding the overlap of Eq. 4.22 we have,

$$\begin{aligned} |\left\langle x_0^1\right| U \left|x_0^2\right\rangle|^2 &= \frac{1}{4}\left(\langle gs_0| U |gs_0\rangle - \langle gs_1| U |gs_1\rangle\right)\left(\langle gs_0| U^\dagger |gs_0\rangle - \langle gs_1| U^\dagger |gs_1\rangle\right) \\ &= \frac{1}{4}\Big[|\langle gs_0| U |gs_0\rangle|^2 + |\langle gs_1| U |gs_1\rangle|^2 \\ &\quad - \langle gs_0| U |gs_0\rangle \langle gs_1| U^\dagger |gs_1\rangle - \langle gs_1| U |gs_1\rangle \langle gs_0| U^\dagger |gs_0\rangle\Big]. \end{aligned} \tag{4.24}$$

Here we require the evolution of the classical ground state under the TFIM. Since these two models coincide when $\eta = 0$, we can write these classical states in terms of the momentum space fermions needed to diagonalize the TFIM. In this fermionic basis, the overlaps

present in the bottleneck ratio can be calculated exactly. The derivation is presented in Appendix C, with the final forms listed below.

$$|\langle gs_0|U|gs_0\rangle|^2 = \prod_{K_0^+}\left[1-\frac{\eta^2\sin^2(k)}{\varepsilon_k^2}\sin^2(2\varepsilon_k)\right] \tag{4.25}$$

$$|\langle gs_1|U|gs_1\rangle|^2 = \prod_{K_1^+}\left[1-\frac{\eta^2\sin^2(k)}{\varepsilon_k^2}\sin^2(2\varepsilon_k)\right] \tag{4.26}$$

$$\begin{aligned}&\langle gs_0|U|gs_0\rangle\langle gs_1|U^\dagger|gs_1\rangle\\ &= e^{-2i\alpha(N/2-1)}\prod_{\{K_0\}}\left[\cos(2\varepsilon_k)-\frac{i(\eta\cos(k)-\alpha)}{\varepsilon_k}\sin(2\varepsilon_k)\right]\prod_{\{K_1\}}\left[\cos(2\varepsilon_k)+\frac{i(\eta\cos(k)-\alpha)}{\varepsilon_k}\sin(2\varepsilon_k)\right]\end{aligned} \tag{4.27}$$

$$\begin{aligned}&\langle gs_1|U|gs_1\rangle\langle gs_0|U^\dagger|gs_0\rangle\\ &= e^{2i\alpha(N/2-1)}\prod_{\{K_0\}}\left[\cos(2\varepsilon_k)+\frac{i(\eta\cos(k)-\alpha)}{\varepsilon_k}\sin(2\varepsilon_k)\right]\prod_{\{K_1\}}\left[\cos(2\varepsilon_k)-\frac{i(\eta\cos(k)-\alpha)}{\varepsilon_k}\sin(2\varepsilon_k)\right]\end{aligned} \tag{4.28}$$

This analytic form can investigate much larger systems for the bottleneck ratio for $A_1 = \{x_0^1\}$. The other subset we need to consider is $A_2 = S_1$.

$$\begin{aligned}\Phi(A_2) &= \frac{1}{|S_1|\pi(x_1)}\left(\sum_{x\in S_1, y\in A^c}\pi(x_1)A(x,y)|\langle y|U|x\rangle|^2\right)\\ &= \frac{1}{|S_1|}\sum_{x\in S_1,y\in A^c}A(x,y)|\langle y|U|x\rangle|^2\\ &= \frac{1}{|S_1|}\sum_{x\in S_1,y\in S_0}A(x,y)|\langle y|U|x\rangle|^2+\frac{1}{|S_1|}\sum_{x\in S_1,y\in S_2}A(x,y)|\langle y|U|x\rangle|^2+..\\ &\sim \frac{1}{|S_1|}\sum_{x\in S_1,y\in S_0}|\langle y|U|x\rangle|^2\end{aligned} \tag{4.29}$$

Recall we have $|S_1| = N(N-1)$ from Eq. 2.21. This can also be seen as we need to create

2 domain walls, which can be placed at N sites.

$$\begin{aligned}\sum_{x\in S_1, y\in S_0} |\langle y|U|x\rangle|^2 &= \sum_{x\in S_1} \left\langle x_0^1\right| U|x\rangle\langle x|U^\dagger \left|x_0^1\right\rangle + \left\langle x_0^2\right| U|x\rangle\langle x|U^\dagger \left|x_0^2\right\rangle \\ &= \sum_{x\in S_1} \langle gs_0|U|x\rangle\langle x|U^\dagger|gs_0\rangle + \langle gs_1|U|x\rangle\langle x|U^\dagger|gs_1\rangle\end{aligned} \tag{4.30}$$

Compared to the previous calculation, these overlap terms also contain states in S_1. These excited states can also be divided into the $p = 0$ and $p = 1$ sectors, each containing $N(N-1)/2$ states. Since U perseveres the fermionic parity, there is no method to move from one parity sector to another under this evolution. This allows these two sectors to be considered separately.

$$\sum_{x\in S_1, y\in S_0} |\langle y|U|x\rangle|^2 = \sum_{x\in S_1^{p=0}} \langle gs_0|U|x\rangle\langle x|U^\dagger|gs_0\rangle + \sum_{x\in S_1^{p=1}} \langle gs_1|U|x\rangle\langle x|U^\dagger|gs_1\rangle \tag{4.31}$$

These terms are derived in detail in Appendix C, leading to the following bottleneck ratio,

$$\begin{aligned}\Phi(A_2) = \frac{1}{N(N-1)}\Bigg[&|\langle gs_0|U|gs_0\rangle|^2 \sum_{\{K_0\}} \left(\frac{\eta^2\sin^2(2\varepsilon_k)\sin^2(k)}{\varepsilon_k^2 - \eta^2\sin^2(2\varepsilon_k)\sin^2(k)}\right) \\ &+ |\langle gs_1|U|gs_0\rangle|^2 \sum_{\{K_1\}} \left(\frac{\eta^2\sin^2(2\varepsilon_k)\sin^2(k)}{\varepsilon_k^2 - \eta^2\sin^2(2\varepsilon_k)\sin^2(k)}\right)\Bigg].\end{aligned} \tag{4.32}$$

With the analytic form of the bottleneck ratio for A_1 and A_2, we can explore the system size dependence of this bound. Although the ratio of the set A_1 was shown to be relevant numerically for small systems, for larger N, the ratio associated with set A_2 is smaller for all η and κ. This can be seen as both ratios contain the Loschmidt echo term $|\langle gs_0|U|gs_0\rangle|^2$, but $\Phi(A_2)$ has additional polynomially suppression. We can rewrite this bottleneck ratio in a form to make the system size dependence evident.

$$\begin{aligned}|\langle gs|U|gs\rangle|^2 &= \prod_{\{K_0\}}\left[1 - \frac{\eta^2\sin^2(k)}{\varepsilon_k^2}\sin^2(2\varepsilon_k)\right] \\ &= \exp\left[-N\left(-\frac{1}{N}\sum_{\{K_0\}}\ln\left(1 - \frac{\eta^2\sin^2(k)}{\varepsilon_k^2}\sin^2(2\varepsilon_k)\right)\right)\right] \\ &= e^{-N\lambda(\kappa,\eta)}\end{aligned} \tag{4.33}$$

In the continuum limit, $\{K_0\} = \{K_1\}$ and we have for both sectors,

$$\begin{aligned}\lambda(\kappa,\eta) &= -\lim_{N\to\infty}\frac{1}{N}\sum_k \ln\left(1-\frac{\eta^2\sin^2(k)}{\varepsilon_k^2}\sin^2(2\varepsilon_k)\right)\\ &= -\int_0^\pi \frac{dk}{2\pi}\ln\left(1-\frac{\eta^2\sin^2(k)}{\varepsilon_k^2}\sin^2(2\varepsilon_k)\right).\end{aligned} \tag{4.34}$$

This gives the final form of $\Phi(A_2)$,

$$\begin{aligned}\Phi(A_2) &= \frac{1}{N(N-1)}\Bigg[|\langle gs_0|\,U\,|gs_0\rangle|^2\sum_{\{K_0\}}\left(\frac{\eta^2\sin^2(2\varepsilon_k)\sin^2(k)}{\varepsilon_k^2-\eta^2\sin^2(2\varepsilon_k)\sin^2(k)}\right)\\ &\quad + |\langle gs_1|\,U\,|gs_0\rangle|^2\sum_{\{K_1\}}\left(\frac{\eta^2\sin^2(2\varepsilon_k)\sin^2(k)}{\varepsilon_k^2-\eta^2\sin^2(2\varepsilon_k)\sin^2(k)}\right)\Bigg]\\ &= \frac{2}{N(N-1)}e^{-N\lambda(\kappa,\eta)}N\gamma(\kappa,\eta)\\ &= \frac{2e^{-N\lambda(\kappa,\eta)}\gamma(\kappa,\eta)}{N-1}\end{aligned} \tag{4.35}$$

Where $\gamma(\kappa,\eta)$ is defined as,

$$\begin{aligned}\gamma &= \lim_{N\to\infty}\frac{1}{N}\sum_k\frac{\eta^2\sin^2(2\varepsilon_k)\sin^2(k)}{\varepsilon_k^2-\eta^2\sin^2(2\varepsilon_k)\sin^2(k)}\\ &= \int_0^\pi\frac{dk}{2\pi}\frac{\eta^2\sin^2(2\varepsilon_k)\sin^2(k)}{\varepsilon_k^2-\eta^2\sin^2(2\varepsilon_k)\sin^2(k)}.\end{aligned} \tag{4.36}$$

The behaviour of the bottleneck ratio scaling is determined by the exponential dependence on N, weighted by the value of λ. The value of λ is only small in the limit of small η, where polynomial scaling is obtained. This matches the numerical results as in this limit the quantum-enhanced MCMC reduces to the local classical strategy, which scales polynomially. Outside of this limit, the bottleneck ratio scales exponentially with the system size. The value of λ depends on κ and η as seen in Fig. 4.8. In the region $\eta > \kappa$, λ is approximately constant. This can also be seen from Eq. 4.34 as the enveloping function η^2/ε_k^2 becomes constant in this region.

Taking this fixed value of λ the Cheeger inequality bounds are shown in Fig. 4.9 along with the fit of the random quantum proposal from Fig. 4.6. The numerical fit found for small system sizes remains in the bounds given by this bottleneck analysis. Although

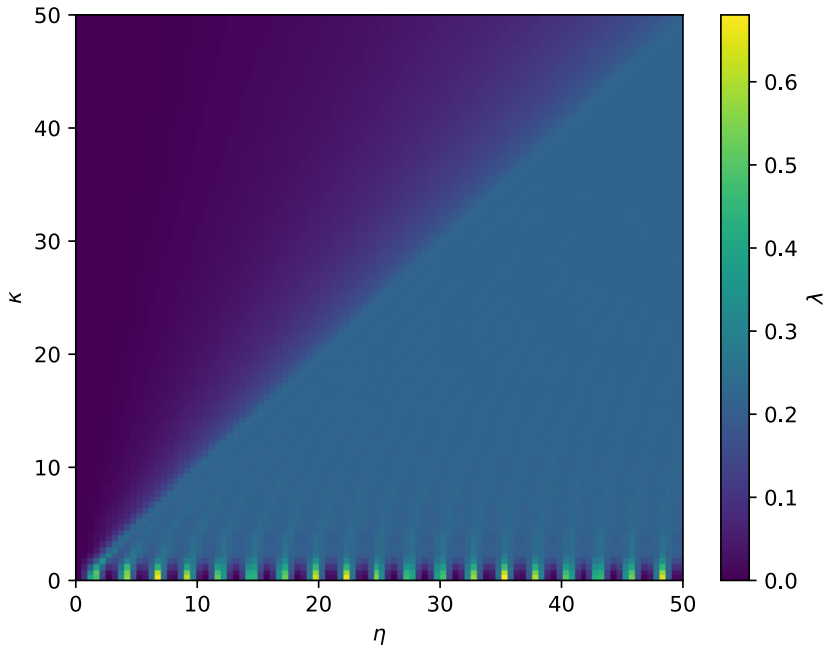

Figure 4.8: Exponential scaling factor λ of the chain's bottleneck ratio as a function of κ and η.

displayed here, the lower bound is not precise due to approximations made in the minimization over sets A. In the limit of $\beta \to \infty$ these approximations become exact. These results show that for a large η and κ the scaling found numerically for small systems falls within the analytic bounds of the Cheeger inequality at large system sizes. Additionally, the scaling of the uniform proposal is shown in Fig. 4.9. The lower Cheeger bound shows favourable scaling of the quantum strategy over this classical uniform method.

The bottleneck ratio of this chain gave insight into the dependence of κ and η seen in Fig. 4.4. Additionally, the Cheeger inequality shows the quantum improvement seen in the Ising chain persists at larger systems beyond what is numerically accessible. We were able to find the bottleneck ratio for this model because the quantum evolution is exactly solvable. As the transverse field Ising model is integrable, it does not obey the eigenstate thermalization hypothesis. The limit of large κ and η is the long time limit of the evolution when the system is expected to have thermalized. The specifics of this model which allowed the bottleneck ratio to be found, may also be causing the scaling speedup shown above.

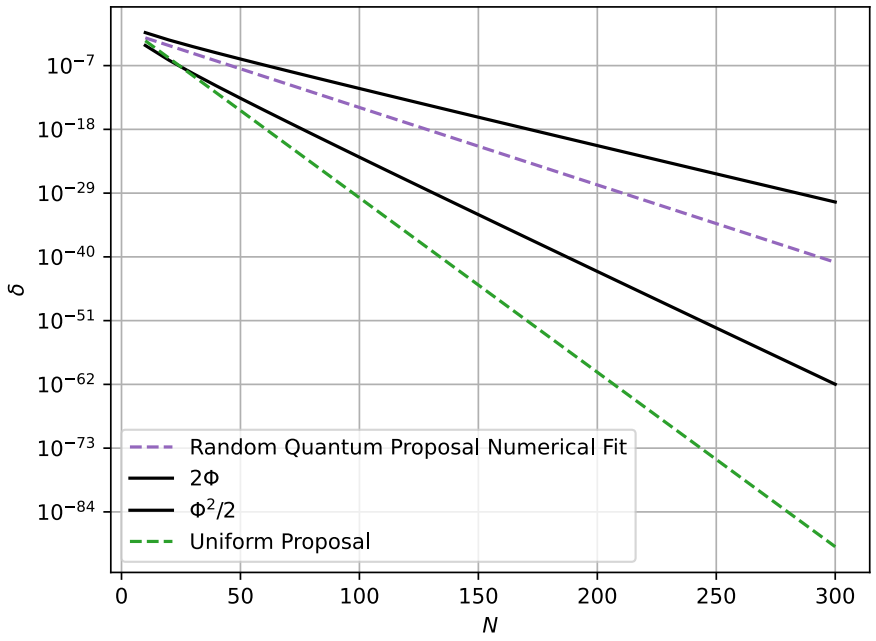

Figure 4.9: Cheeger bounds for quantum proposal strategy at large κ and η. The scaling fit from the small system size numerical results of the random quantum proposal falls within the Cheeger bounds. Additionally, the classical uniform proposal scaling is shown.

The thermalization of this and other models needs to be further investigated to see if the large system scaling found here can be generalized to other systems.

4.4 Maximum Independent Set Problem

4.4.1 Motivation

This algorithm was tested experimentally on a superconducting quantum device for random chain-connected SK model instances with near-degenerate local minimum [23]. The unitary $U = e^{-iHt}$ was approximately implemented using a second-order Suzuki-Trotter expansion. With this approximation and experimental errors, a mixing time speed-up was still observed compared to a classical uniform MCMC method [23]. However, due to device limitations,

this algorithm was only explored for systems with $N \leq 10$.

Fortunately, other candidate quantum devices, such as neutral atom arrays, are better suited for quantum simulation and have been scaled to sizes of hundreds of atoms [18]. Neutral atom arrays offer a highly coherent and controllable platform for implementing quantum many-body Hamiltonians. This algorithm only requires evolution under a specific Hamiltonian, thus avoiding the need for gate-based computation. Recently, it has been shown the Maximum Independent Set (MIS) problem can be encoded onto arrays of Rydberg atoms, a type of neutral atom-based quantum device [31]. The many-body Hamiltonian governing such Rydberg systems contains terms needed for the evolution performed in this algorithm, making it a promising candidate device for quantum-enhanced MCMC.

4.4.2 Rydberg Atom Arrays

Rydberg atom arrays consist of neutral atoms that are trapped and arranged using optical tweezers. These atoms, typically rubidium, are driven between their ground state and a Rydberg state, characterized by a large principal quantum number. Two atoms in the Rydberg state experience a dipole-dipole interaction. If the two atoms are at positions $\vec{x}_j$ and $\vec{x}_k$ respectively, typically the atoms are experimentally prepared so that the leading-order behaviour is a $1/|\vec{x}_j - \vec{x}_k|^6$ van der Waals interaction [8]. This interaction penalizes atoms of close proximity to both be excited into the Rydberg state, an effect known as the Rydberg blockade. Atomic excitations through homogeneous laser pulses are specified with a time-dependent Rabi frequency $\Omega(t)$ and detuning $\Delta(t)$. In combination with the van der Waals interaction, these arrays are described with the many-body Hamiltonian [13],

$$H(t) = \frac{\Omega(t)}{2} \sum_{i=1}^{N} \sigma_i^x - \Delta(t) \sum_j \hat{n}_j + \sum_{i<j} V_{ij} \hat{n}_i \hat{n}_j. \tag{4.37}$$

This Hamiltonian acts on the Rydberg state occupation basis, where each of the N atoms are described by the two-dimensional Hilbert space spanned by the atom's ground state $|g\rangle$ and Rydberg state $|r\rangle$. The Rydberg occupation operator is defined as $\hat{n}_i = |r_i\rangle\langle r_i|$, and here $\sigma_i^x = |g_i\rangle\langle r_i| + |r_i\rangle\langle g_i|$. The interaction term, which penalizes atoms both in Rydberg state, has the following power-law decay,

$$V_{ij} = V_0/|\vec{x}_j - \vec{x}_k|^6. \tag{4.38}$$

A distance scale R_b, known as the blockade radius, determines the radius at which the Rydberg interaction becomes strong. Specifically, this radius defines where the interaction

energy is comparable to the coupling determined by Ω,

$$\frac{V_0}{R_b^6} = \Omega. \tag{4.39}$$

If any atoms are closer than this radius, it is energetically unfavourable for them to both be in the Rydberg state. In an atom array created with optical tweezers, the effective blockade radius can be modified by programming the lattice spacing. This strongly-interacting Hamiltonian has been experimentally realized with a high degree of control for systems with hundreds of atoms [13].

Unit Disk Maximum Independent Set Problem

There is a hardware-efficient encoding of an NP-hard optimization problem on these arrays which requires the experimental ability to control the position of the atoms as well as the capability to control the atomic state [31]. Specifically, a maximum independent set problem can be encoded. On a graph G with vertices V and edges E an independent set is a collection of vertices in which no pair is connected by an edge. Finding the largest subset is a problem known as Maximum Independent Set (MIS). We will focus on the unit disk MIS problem, where the graph is embedded in a two-dimensional plane where two vertices are connected only if they are separated by a distance smaller than a unit radius. Even on this restricted graph type, this problem is NP-hard [10].

The Rydberg blockade mechanism that restricts two atoms in close proximity from becoming excited naturally enforces this independent set constraint. With a positive detuning Δ, the ground state of this Hamilton then finds the maximum independent set. It has been shown that the Rydberg Hamiltonian Eq. 4.37, with long-range interaction tails, can encode this NP-hard problem [31]. Additionally, Rydberg atom arrays have been used to optimize this MIS problem through variational algorithms [12]. Motivated by this physical device which can also implement the needed σ^x term, we will numerically explore this quantum algorithm on the unit disk maximum independent set problem. Ignoring additional effects from the interaction tails of the Rydberg interaction, we will consider the following Hamiltonian,

$$H_c^{\text{MIS}} = -\sum_{i \in V} \hat{n}_i + \sum_{(i,j) \in E} \nu \hat{n}_i \hat{n}_j \tag{4.40}$$

where $\nu = 100$. In order to exactly calculate δ we are limited to small systems simulations. These problems are embedded in a two-dimensional grid so two parameters are used the specify the problem. The first is edge length L and the second is a random dropout

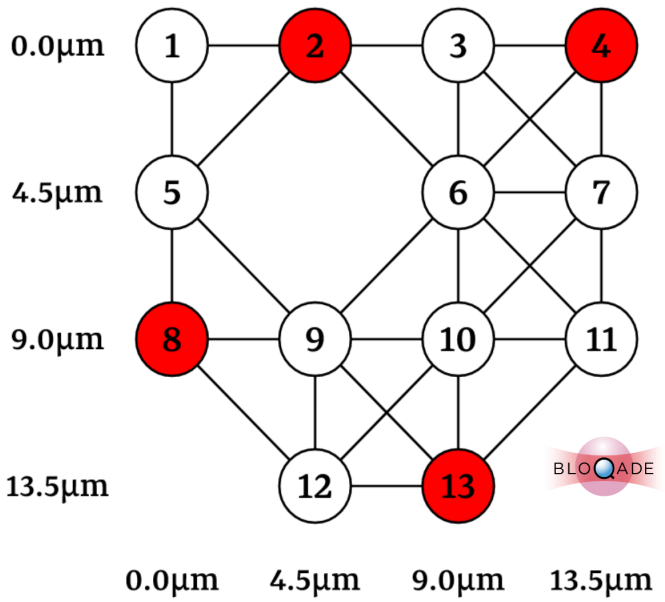

Figure 4.10: Unit disk maximum independent set solution on a $L = 4$ lattice with random dropout probability $D = 0.2$ [1].

probability D. For example, the lattice seen in Fig. 4.10 has $L = 4$ and $D = 0.2$. The maximum independent set is highlighted in red. Configurations were generated and solved using the Bloqade package [1].

4.4.3 Numerical Results

Using spin variable $\sigma_i = \pm 1$, the current configuration set is all spin with value one. At low temperatures, the MCMC will sample from sets which minimize Eq. 4.40, the maximum independent sets of the graph. We will compare the quantum-enhanced MCMC method to the uniform and local classical MCMC methods described above. However, we will additionally compare a classical update optimized for the MIS problem, the exchange MCMC (A fixed temperature variant of the Rydberg simulated annealing algorithm of [12]). Due to the independent set constraint Eq. 4.40, having two neighbouring spins part of the set is unfavourable. The exchange MCMC makes use of this structure and proposes new configurations primarily through a spin exchange. Spin exchanges remove one site from the set and add one of its neighbours to the set. This movement is not penalized by

the independent set constraint like random site addition is. Let free sites be the set of spins which have no neighbours in the set. The exchange MCMC generates new configurations through the following process.

1. Select a spin at random from the set of free sites and the sites in the set.
2. If the selected site is free, add it to the set.
3. If the selected site is in the set, randomly pick one of its neighbours. Spin exchange with this neighbour with probability 1/8. Remove the spin from the set with the remaining probability.

The spin exchange occurs with probability 1/8 to ensure similar behaviour for graphs with different connectivity. This chain does not satisfy the detailed balance conditions and is not fully mixing at high temperatures. However, for low-temperature, where the system is already constrained to independent sets, this chain mixes and has more favourable scaling than the other classical alternatives.

We will also compare a quantum alternative to the exchange MCMC. With the motivation of emulating the spin exchange advantage for this problem, we will consider the following term of the quantum evolution under Eq. 4.1,

$$H_{\text{mix}} = \sum_{i,j \in E} \frac{1}{2}(\sigma_i^+ \sigma_j^- + \sigma_i^- \sigma_j^+). \tag{4.41}$$

The quantum enhanced MCMC method is easy to adapt to variants optimized for the problem at hand. We will use the same steps as the classical algorithm, replacing the spin exchange with the quantum evolution.

1. Select a spin at random from the set of free sites and the sites in the set.
2. If the selected site is free, add it to the set.
3. If the selected site is in the set, with probability (# of neighbours)/8 evolve under the Hamiltonian of Eq. 4.1, with H_{mix} of Eq. 4.41. With the remaining probability, remove the spin from the set.

In the limit of small η this algorithm reduces to the above exchange MCMC. The dependence of the average δ for both these quantum strategies can be seen in Fig. 4.11 for $N = 7$ in the low-temperature limit. Unlike both the SK and Ising model, there is a concentration of large spectral gaps when η is small in Fig. 4.11. The reasoning for this concentration is unknown and needs to be further investigated.

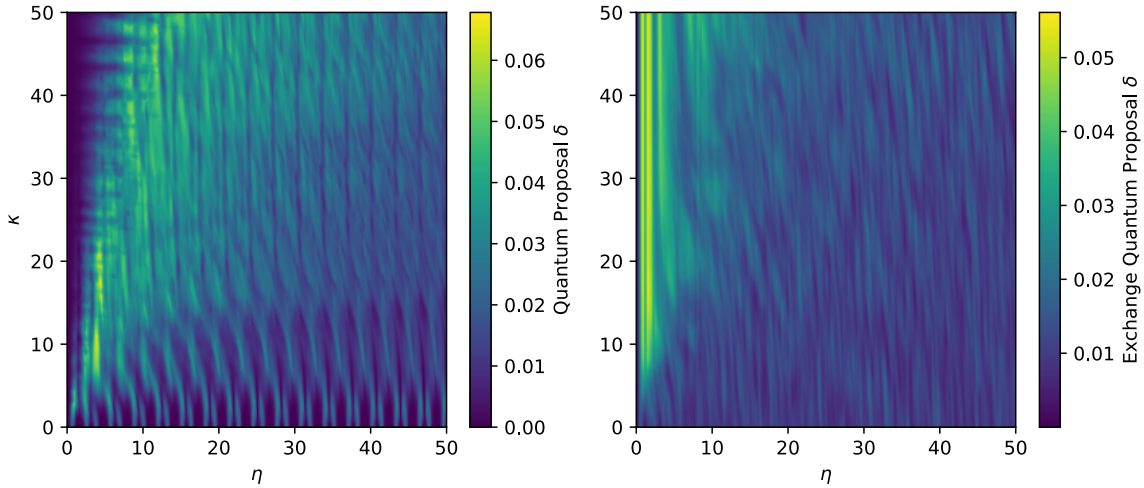

Figure 4.11: Spectral gap δ of quantum enhanced MCMC for $N = 7$ ($L = 3$ and $D = 0.2$) averaged over 100 instances of MIS at $\beta = 10$. Both the original quantum-enhanced MCMC proposal and the exchange proposal are shown.

Temperature Dependence

The average spectral gap of these algorithms are compared as a function of temperature in Fig. 4.12. In the uniform proposal method, the three energy regimes of this problem are evident. For low temperatures, below around $T = 0.1$, the chain samples from maximum independent sets. At slightly higher temperatures, independent sets are sampled. Finally, at high temperatures, non-independent sets are sampled as there is enough energy to overcome the large interaction term in the Hamiltonian. These three plateaus are also seen in the quantum proposal as that method is fully mixing at all temperatures. Both the classical and quantum exchange methods cannot add sites that break the independent set constraint and thus do not mix at large temperatures. Not evident in this figure, but at those temperatures, the exchange methods converge to an incorrect distribution. In the low-temperature limit, for this system size, the quantum strategies have longer mixing times as compared to the uniform proposal.

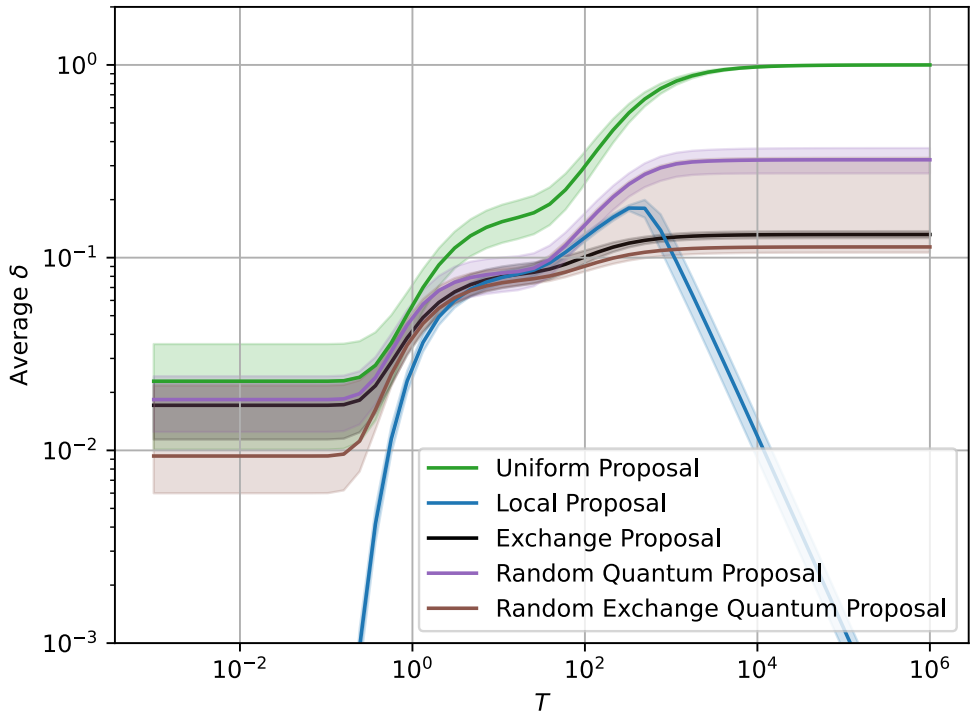

Figure 4.12: Spectral gap δ of classical and quantum strategies for $N = 8$ averaged over 100 graphs with $L = 4$ and $D = 0.5$.

Low-Temperautre System Size Scaling

Due to the graph structure of this problem, the system sizes are varied by considering different amounts of random dropouts. As a result, for some system sizes the graph is significantly more sparse, leading to varying difficulty in the task of finding the maximum independent set. Due to this added complication, we will consider scaling for graphs of comparable difficulty. Two measures of difficulty will be considered, the first being the degeneracy of the ground state manifold, $D_{|\mathrm{MIS}|}$. The second will be the graph's hardness parameter [12]. This is a measure of how hard it is for a random walker on the set of size $|\mathrm{MIS}| - 1$ to find a global minimum. It depends on the degeneracy of this set, $D_{|\mathrm{MIS}|-1}$, as

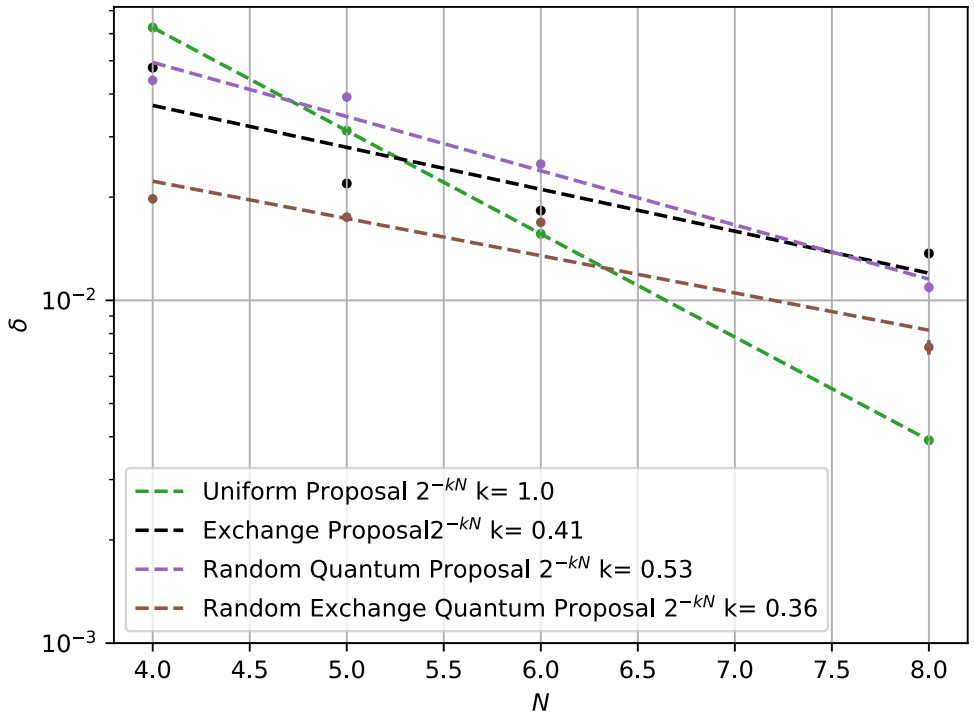

Figure 4.13: Mixing time scaling with system size of classical and quantum methods for graphs with fixed hardness parameter and degeneracy at $\beta = 10$.

well as the size and degeneracy of the optimal set. Specifically, it is defined as

$$\mathcal{HP} = \frac{D_{|\mathrm{MIS}|-1}}{|\mathrm{MIS}|D_{|\mathrm{MIS}|}}. \tag{4.42}$$

In Fig. 4.13 the $\beta = 10$ scaling is shown for fixed degeneracy and hardness parameters. Specifically, graphs with only one MIS set and a hardness parameter of 2 were used. Unfortunately, configurations with these specifications do not exist for all system sizes. This limitation makes it hard to determine the scaling characteristics of both the classical exchange proposal and the quantum strategies. However, from this data, there does not seem to be an advantage of the quantum algorithm over the exchange classical method. A possible extension to consider larger system sizes is to simulate only the subspace of the mixing matrix acting on the independent set configurations. As we are in the low-temperature

limit, transitions to non-independent set configurations are very unlikely. Simulating only this subspace would allow for larger system sizes to be considered in the scaling performance of the quantum algorithm. More scaling data is necessary for conclusions of the performance of this algorithm on this problem. However, this avenue was not yet pursued due to a lack of understanding of Fig. 4.11, instead, simpler models where more insight could be gained were considered.

4.5 Gaussian Distributed States

When exploring the bottleneck ratio of the Ising model, it was realized that the requirement of a unitary proposal might restrict the performance of the chain. The optimal unitary can be found for a model with Gaussian distributed states, giving insight into the best bound under this restriction of a unitary proposal. Assume we have a large system size such that the density of states of the classical Hamiltonian is a continuous function of energy, E. Specifically, it will be assumed to be Gaussian,

$$\rho(E) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{E^2}{2\sigma^2}}, \tag{4.43}$$

where σ is the standard deviation. The Boltzmann distribution π of this system can be found exactly as the assumption of a Gaussian density of states allows the partition function Z to be solved.

$$\begin{aligned} 1 &= \frac{1}{Z}\int_{-\infty}^{\infty} dE \rho(E) e^{-\beta E} \\ &= \frac{1}{Z}\int_{-\infty}^{\infty} dE \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{E^2}{2\sigma^2} - \beta E} \\ &= \frac{1}{Z}\int_{-\infty}^{\infty} dE \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(E+\beta\sigma^2)^2 + \frac{\beta^2\sigma^2}{2}} \\ &\rightarrow Z = e^{\beta^2\sigma^2/2} \end{aligned} \tag{4.44}$$

We will make use of the Cheeger inequality, Eq. 2.21, in order to gain insight of the scaling of the quantum-enhanced MCMC on this system. Recall we need the bottle ratio of the chain which is a minimization over subsets A,

$$\Phi_\star := \min_{A:\pi(A)\leq 1/2} \frac{\sum_{x\in A, y\in A^c} \pi(x)P(x,y)}{\pi(A)}, \tag{4.45}$$

In order to maximize the denominator of the bottleneck,

$$\pi(A) = \frac{1}{Z}\int_{E(A)} dE\rho(E)e^{-\beta E} = \int_{E(A)} dE\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(E+\beta\sigma^2)^2} \sim \frac{1}{2} \tag{4.46}$$

Additionally, in order to minimize the numerator, A should contain the largest energy states so they get suppressed by $\pi(x)$. Consider a cutoff E^* and let A be all states with $E > E^*$. To satisfy Eq. 4.46, E^* is the peak of the Gaussian,

$$E^* = -\beta\sigma^2. \tag{4.47}$$

Moving from this set A to A^c lowers the energy. Using Metropolis-Hastings acceptance probability, this new configuration is always acceptance. This simplifies the form of the bottleneck ratio,

$$\Phi(A, A^c) = \frac{\sum_{x\in A, y\in A^c} \pi(x) \langle y| U |x\rangle\langle x| U^\dagger |y\rangle}{\sum_{x\in A}\pi(x)}. \tag{4.48}$$

4.5.1 Optimal Unitary

Our goal is to see, given the optimal unitary, U, how the quantum method performs. Given the fixed A found above, we will consider the optimal unitary, the one which maximizes the bottleneck ratio. To maximize the numerator, the largest number of states from A should be moved to A^c. Due to the constraint of unitary evolution, at most $|A^c|$ states of A can be transferred into A^c. For classical methods, all states can be transferred into A^c, so the restriction of U limits the quantum strategy. However, in the numerator of the bottleneck ratio the terms are additionally weighted by $\pi(x)$. For high-energy states, when sampling from the low-temperature Boltzmann distribution, $\pi(x)$ for these states is exponentially small. The quantum algorithm can only transfer a fixed number of states, but many states barely contribute due to this exponential suppression. The optimal unitary U transfers the $|A^c|$ smallest energy states of A. Define E' to be the cutoff where all states in A with $E^* < E < E'$ will get transferred into A^c. The size of A^c is,

$$|A^c| = \int_{-\infty}^{E^*} dE\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{E^2}{2\sigma^2}} = \frac{1}{2}\left(1 - \text{erf}\left(\frac{\beta\sigma}{\sqrt{2}}\right)\right). \tag{4.49}$$

With $|A^c|$ we can now solve for E'.

$$\begin{aligned} \int_{-\infty}^{E^*} dE\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{E^2}{2\sigma^2}} &= \int_{E^*}^{E'} dE\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{E^2}{2\sigma^2}} \\ \frac{1}{2}\left(1 - \text{erf}\left(\frac{\beta\sigma}{\sqrt{2}}\right)\right) &= \frac{1}{2}\left(\text{erf}\left(\frac{E'}{\sqrt{2}\sigma}\right) + \text{erf}\left(\frac{\beta\sigma}{\sqrt{2}}\right)\right) \end{aligned} \tag{4.50}$$

$$E' = \sqrt{2}\sigma \text{erf}^{-1}\left(1 - 2\text{erf}\left(\frac{\beta\sigma}{\sqrt{2}}\right)\right) \tag{4.51}$$

The conductance for the optimal quantum strategy is then,

$$\begin{aligned}
\Phi(A, A^c) &= 2\int_{E^*}^{E'} dE \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(E+\beta\sigma^2)^2} \\
&= \text{erf}\left(\frac{\beta\sigma^2 + E'}{\sqrt{2}\sigma}\right) - \text{erf}\left(\frac{\beta\sigma^2 + E^*}{\sqrt{2}\sigma}\right) \\
&= \text{erf}\left(\frac{\beta\sigma^2 + E'}{\sqrt{2}\sigma}\right) \\
&= \text{erf}\left(\frac{\beta\sigma}{\sqrt{2}} + \text{erf}^{-1}\left(1 - 2\text{erf}\left(\frac{\beta\sigma}{\sqrt{2}}\right)\right)\right) \leq \frac{1}{\beta\sigma}.
\end{aligned}$$

Typically many-body systems have extensive bandwidth thus $\sigma^2 \sim N$ [35]. For this optimized unitary, we then have,

$$\Phi(A, A^c) \leq \frac{1}{\beta\sqrt{N}}. \tag{4.52}$$

This upper bound on the bottleneck ratio does not have exponential dependence on the system size. The situation outlined here is not practical as there is no general strategy which implements this optimal unitary. It does demonstrate that requiring a unitary proposal method does not immediately give exponentially increasing mixing time, thus not eliminating a possibility for a large system size quantum speedup.

4.6 Discussion

In this chapter, we saw the application of the quantum-enhanced MCMC algorithm on numerous problems. First, the SK model was explored numerically, confirming the low-temperature fast mixing as compared to uniform classical MCMC chain as seen in [23]. Next, the Ising chain was considered both numerically and analytically. We again observe a low-temperature advantage of the quantum algorithm as compared to a uniform classical strategy numerically. The bottleneck ratio of the quantum chain was then explored which not only gave insights into which Hamiltonian parameters lead to a large spectral gap, but allowed us to consider the behaviour of the chain for system sizes not accessible

numerically. In the limit of large κ and η the Cheeger inequality showed the quantum algorithm for low-temperature sampling maintains a mixing time advantage over the uniform classical method. Next, the algorithm was tested on the experimentally relevant MIS problem. Numerically we did not find any advantage of the quantum-enhanced MCMC algorithm over classical algorithms specified for this problem. The quantum algorithm was additionally extended to a version also optimized for this problem but still showed no advantage. Finally, Gaussian distributed states were explored in order to see if there was any limitation of requiring unitary proposal methods. Our analysis suggests that for this problem, this restriction does not lead to exponential scaling of the mixing time.

As we have seen from the above examples, the quantum-enhanced MCMC only ever shows an improvement over the classical methods in the low-temperature limit. In this limit, classical sampling methods have been shown to be substantially improved with annealing and parallel tempering methods [36, 7]. These methods are not compared here, and in order to claim any advantage of this quantum algorithm over its classical counterparts this comparison is necessary. Due to the Markov chain construction of the quantum-enhanced MCMC, this quantum method can also be extended with these same annealing and parallel tempering methods. However, it is unknown if this would lead to any algorithmic improvement over classical methods. At a fixed temperature, the quantum method is able to transition between states far in Hamming distance that are still close in energy. This is difficult for classical methods at fixed temperatures, however with the addition of annealing or parallel tempering, they no longer have this limitation. Therefore it is plausible that allowing variable temperature strategies may result in no quantum improvement. Additionally, often there are classical MCMC methods that are optimized for the specific problem at hand. We have seen an example of one such example when considering the MIS problem. In that case we were unable to show any numerical advantage of the quantum-enhanced MCMC algorithm over the classical strategy. This potentially limits the use case of this quantum algorithm to only difficult problems where uniform classical methods are the only option. More optimistically, as this quantum algorithm creates a classical chain, it can be easily adapted to the specific problem. This was showcased in the quantum exchange algorithm for the MIS problem.

There are still many open questions surrounding this quantum algorithm and possible directions for further research. From the investigation presented in this research, it is still unknown if this quantum algorithm can be simulated classically. For the examples considered, it seems large spectral gaps can be obtained for short-time quantum evolution, indicating there is a possibility of classical simulation. Another current path of investigation is to consider the long time limit of evolution where the quantum system is assumed to be thermalized. Hopefully, these explorations will give insight into the large system behaviour

of non-integrable systems, a previously inaccessible limit.

Chapter 5

Conclusion

Near-term quantum algorithms for sampling often focus on sampling tasks that are difficult classically but lack immediate practical relevance. In this thesis, we consider near-term quantum algorithms to sample from a Boltzmann distribution of a classical Hamiltonian. This sampling problem is ubiquitous in statistical physics in order to estimate observables of intractable systems. However, this task is also relevant much more broadly in fields such as machine learning and optimization.

The first quantum algorithm considered, adiabatically prepares a qsample which encodes the Boltzmann distribution [44]. The adiabatic state preparation time of this algorithm is related to the properties of the phase transitions traversed during the adiabatic schedule. Specifically, a quantum parent Hamiltonian such that its ground state is the wanted qsample. By considering adiabatic paths in an extended parameter space of this parent Hamiltonian, this algorithm can give a scaling improvement of the state preparation time as compared to the mixing time of the classical MCMC method.

We reproduce the quadratic scaling improvement found for a one-dimensional Ising model. The applicability of this algorithm to other more complicated classical problems is limited as it requires knowledge of the phase transitions of the quantum parent Hamiltonian. Instead, we consider counterdiabatic driving protocols on the adiabatic qsample preparation problem presented. Counterdiabatic driving protocols add additional driving terms to the Hamiltonian to suppress transitions out of the ground state. If these terms are found and implemented exactly, the qsample can be prepared instantaneously. The driving terms are often nonlocal and not practical for today's quantum devices. We therefore considered local approximations and numerically investigated the state preparation time of the qsample. This protocol was tested on the one-dimensional Ising model, where

the counterdiabatic terms can be calculated exactly. When the driving terms are locally restricted, the state preparation time scales unfavourably as compared to the classical MCMC. Further investigation is needed to determine the nature of this scaling and the applicability of this method to other problems.

In Chapter 4, another near-term quantum algorithm to sample from the Boltzmann distribution is investigated; the quantum-enhanced Markov chain Monte Carlo [23]. This algorithm does not produce qsamples. Instead, it is a Markov chain on the space of classical configurations. New configurations are proposed by encoding the current configuration as a computational basis state, evolving under a quantum Hamiltonian, and then protectively measuring to obtain a new classical configuration. This new configuration is then accepted as the new state of the MCMC according to the Metropolis-Hastings acceptance probability. This chain converges to the desired Boltzmann distribution, even if the quantum evolution is inaccurate.

First, the mixing time of this MCMC is investigated through numerical calculation of the spectral gap of the transition matrix on the Sherrington-Kirkpatrick model. We showed faster mixing times than standard classical MCMC methods at low temperatures for the numerically accessible system sizes, verifying the results of [23]. However, for this system, it is difficult to gain an understanding of the spectral gap's dependence on the parameters of the quantum evolution. For the accessible system sizes, the system maintains fast mixing even for short quantum evolutions, indicating the possibility of classical simulation of this evolution. Additionally, it is unknown if this speedup seen numerically is affected by finite-size effects and if it persists to large system sizes. The dependence of the spectral gap on the parameters of the quantum evolution is important for understanding the nature of the speed-up found and, unfortunately, very difficult to investigate for this system.

Next, a simplified model was considered, the one-dimensional Ising model. The spectral gap was found numerically, showing the same scaling advantage over uniform classical methods as in the SK model. The quantum evolution in the quantum-enhanced MCMC for this problem can be solved exactly for any system size. However, calculating the spectral gap is still limited to small systems accessible numerically. Instead, the bottlenecks of the quantum-enhanced MCMC were found in the limit of long quantum evolution. The Cheeger inequality can then be used to bound the mixing time for any system size. The scaling found numerically for small system sizes remains in the found bounds of the Cheeger inequality at large system sizes. Importantly, the lower bound on the spectral gap maintains an advantage over the uniform classical method for all system sizes. This shows that for this problem, the quantum-enhanced MCMC method with a long evolution time has a faster mixing time than the uniform classical MCMC. Analyzing the bottlenecks of the chain not only gave an analytical bond on the spectral gap but also gave insight into the spectral

gap's dependence on the parameters of the quantum evolution.

This algorithm has been implemented on superconducting devices, requiring an approximation of the quantum evolution [23]. Therefore, the maximum independent set problem was considered, as the quantum evolution of this algorithm can be implemented in Rydberg atom arrays without approximation. Here we compared to a classical MCMC method optimized for this problem and also considered an equivalently modified quantum method. The number of numerically accessible system sizes with the same problem difficulty is limited, and from the available data, there is no advantage of the quantum algorithm over its classical counterparts. Further research is required to understand the spectral gap dependence on the evolution parameters, as this dependence is significantly different than the other examples considered.

Finally, another simplified model is considered, a system with a gaussian distributed density of states. When investigating the bottlenecks of the Ising model, it was realized that the requirement of a proposal governed by a unitary might limit the chain's performance. Therefore this simplified model was considered as the optimal unitary can easily be found. This unitary maximizes the bottleneck ratio, and luckily does not lead to an upper bound that is exponentially small in the system size. These results show, at least for this model, that the requirement of a unitary proposal does not immediately discount this algorithm.

In all, this quantum-enhanced MCMC method was tested on four different classical problems and showed varying performance. There are still many avenues of investigation necessary in order to understand both the nature of the speedup found and for which problems this algorithm is beneficial. It is still unknown if a classical "quantum-inspired" version of this algorithm would maintain the same mixing time advantages found. Numerically, short quantum evolutions seem to still lead to large spectral gaps. If these evolutions can be approximated classically and the speed-up maintained, there is potential for a classical version of this algorithm.

References

[1] Bloqade.jl: Package for the quantum computation and quantum simulation based on the neutral-atom architecture. https://github.com/queracomputing/bloqade.jl, 2023.

[2] Srinivasan Arunachalam, Vojtech Havlicek, Giacomo Nannicini, Kristan Temme, and Pawel Wocjan. Simpler (classical) and faster (quantum) algorithms for Gibbs partition functions. *Quantum*, 6:789, 2022.

[3] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando GSL Brandao, David A Buell, et al. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, 2019.

[4] Francisco Barahona. On the computational complexity of Ising spin glass models. *Journal of Physics A: Mathematical and General*, 15(10):3241, 1982.

[5] Nathanaël Berestycki. Mixing times of Markov chains: Techniques and examples. *Alea-Latin American Journal of Probability and Mathematical Statistics*, 2016.

[6] Michael Betancourt. A conceptual introduction to Hamiltonian Monte Carlo. *arXiv preprint arXiv:1701.02434*, 2017.

[7] Elmar Bittner, Andreas Nußbaumer, and Wolfhard Janke. Make life simple: Unleash the full power of the parallel tempering algorithm. *Physical review letters*, 101(13):130603, 2008.

[8] Antoine Browaeys and Thierry Lahaye. Many-body physics with individually controlled Rydberg atoms. *Nature Physics*, 16(2):132–142, 2020.

[9] Andrew M Childs. Lecture notes on quantum algorithms. *Lecture notes at University of Maryland*, 2017.

[10] Brent N Clark, Charles J Colbourn, and David S Johnson. Unit disk graphs. *Discrete mathematics*, 86(1-3):165–177, 1990.

[11] Adolfo del Campo, Marek M Rams, and Wojciech H Zurek. Assisted finite-rate adiabatic passage across a quantum critical point: exact solution for the quantum ising model. *Physical review letters*, 109(11):115703, 2012.

[12] Sepehr Ebadi, Alexander Keesling, Madelyn Cain, Tout T Wang, Harry Levine, Dolev Bluvstein, Giulia Semeghini, Ahmed Omran, J-G Liu, Rhine Samajdar, et al. Quantum optimization of maximum independent set using Rydberg atom arrays. *Science*, 376(6598):1209–1215, 2022.

[13] Sepehr Ebadi, Tout T Wang, Harry Levine, Alexander Keesling, Giulia Semeghini, Ahmed Omran, Dolev Bluvstein, Rhine Samajdar, Hannes Pichler, Wen Wei Ho, et al. Quantum phases of matter on a 256-atom programmable quantum simulator. *Nature*, 595(7866):227–232, 2021.

[14] Nigel Goldenfeld. *Lectures on phase transitions and the renormalization group*. CRC Press, 2018.

[15] Jonathan Goodman. Lecture notes on Monte Carlo methods fall semester. Courant Institute of Mathematical Sciences, NYU, 2005.

[16] Will Grathwohl, Kuan-Chieh Wang, Jörn-Henrik Jacobsen, David Duvenaud, Mohammad Norouzi, and Kevin Swersky. Your classifier is secretly an energy based model and you should treat it like one. *arXiv preprint arXiv:1912.03263*, 2019.

[17] Aram W Harrow and Annie Y Wei. Adaptive quantum simulated annealing for Bayesian inference and estimating partition functions. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 193–212. SIAM, 2020.

[18] Loïc Henriet, Lucas Beguin, Adrien Signoles, Thierry Lahaye, Antoine Browaeys, Georges-Olivier Reymond, and Christophe Jurczak. Quantum computing with neutral atoms. *Quantum*, 4:327, 2020.

[19] He-Liang Huang, Xiao-Yue Xu, Chu Guo, Guojing Tian, Shi-Jie Wei, Xiaoming Sun, Wan-Su Bao, and Gui-Lu Long. Near-term quantum computing techniques: Variational quantum algorithms, error mitigation, circuit compilation, benchmarking and classical simulation. *Science China Physics, Mechanics & Astronomy*, 66(5):250302, 2023.

[20] Scott Kirkpatrick, C Daniel Gelatt Jr, and Mario P Vecchi. Optimization by simulated annealing. *science*, 220(4598):671–680, 1983.

[21] Michael Kolodrubetz, Dries Sels, Pankaj Mehta, and Anatoli Polkovnikov. Geometry and non-adiabatic response in quantum and classical systems. *Physics Reports*, 697:1–87, 2017.

[22] David Landau and Kurt Binder. *A guide to Monte Carlo simulations in statistical physics*. Cambridge university press, 2021.

[23] David Layden, Guglielmo Mazzola, Ryan V Mishmash, Mario Motta, Pawel Wocjan, Jin-Sung Kim, and Sarah Sheldon. Quantum-enhanced Markov chain Monte Carlo. *arXiv preprint arXiv:2203.12497*, 2022.

[24] Jessica Lemieux, Bettina Heim, David Poulin, Krysta Svore, and Matthias Troyer. Efficient quantum walk circuits for Metropolis-Hastings algorithm. *Quantum*, 4:287, 2020.

[25] David A Levin and Yuval Peres. *Markov chains and mixing times*, volume 107. American Mathematical Soc., 2017.

[26] Austin P Lund, Michael J Bremner, and Timothy C Ralph. Quantum sampling problems, boson sampling and quantum supremacy. *npj Quantum Information*, 3(1):15, 2017.

[27] Lars S Madsen, Fabian Laudenbach, Mohsen Falamarzi Askarani, Fabien Rortais, Trevor Vincent, Jacob FF Bulmer, Filippo M Miatto, Leonhard Neuhaus, Lukas G Helt, Matthew J Collins, et al. Quantum computational advantage with a programmable photonic processor. *Nature*, 606(7912):75–81, 2022.

[28] Glen Bigan Mbeng, Angelo Russomanno, and Giuseppe E Santoro. The quantum Ising chain for beginners. *arXiv preprint arXiv:2009.09208*, 114, 2020.

[29] Ashley Montanaro. Quantum speedup of Monte Carlo methods. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 471(2181):20150301, 2015.

[30] Davide Orsucci, Hans J Briegel, and Vedran Dunjko. Faster quantum mixing for slowly evolving sequences of Markov chains. *Quantum*, 2:105, 2018.

[31] Hannes Pichler, Sheng-Tao Wang, Leo Zhou, Soonwon Choi, and Mikhail D Lukin. Quantum optimization for maximum independent set using Rydberg atom arrays. *arXiv preprint arXiv:1808.10816*, 2018.

[32] John Preskill. Quantum computing in the NISQ era and beyond. *Quantum*, 2:79, 2018.

[33] Peter C Richter. Quantum speedup of classical mixing processes. *Physical Review A*, 76(4):042306, 2007.

[34] Christian P Robert, George Casella, and George Casella. *Monte Carlo statistical methods*, volume 2. Springer, 1999.

[35] Subir Sachdev. Quantum phase transitions. *Physics world*, 12(4):33, 1999.

[36] Anders W Sandvik. Computational studies of quantum spin systems. In *AIP Conference Proceedings*, volume 1297, pages 135–338. American Institute of Physics, 2010.

[37] Dries Sels and Anatoli Polkovnikov. Minimizing irreversible losses in quantum systems by local counterdiabatic driving. *Proceedings of the National Academy of Sciences*, 114(20):E3909–E3916, 2017.

[38] David Sherrington and Scott Kirkpatrick. Solvable model of a spin-glass. *Physical review letters*, 35(26):1792, 1975.

[39] Rolando Somma, Sergio Boixo, and Howard Barnum. Quantum simulated annealing. *arXiv preprint arXiv:0712.1008*, 2007.

[40] Rolando D Somma, Sergio Boixo, Howard Barnum, and Emanuel Knill. Quantum simulations of classical annealing processes. *Physical review letters*, 101(13):130504, 2008.

[41] Yang Song and Diederik P Kingma. How to train your energy-based models. *arXiv preprint arXiv:2101.03288*, 2021.

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

[43] John Watrous. Quantum computational complexity. *arXiv preprint arXiv:0804.3401*, 2008.

[44] Dominik S Wild, Dries Sels, Hannes Pichler, Cristian Zanoci, and Mikhail D Lukin. Quantum sampling algorithms for near-term devices. *Physical Review Letters*, 127(10):100504, 2021.

[45] Dominik S Wild, Dries Sels, Hannes Pichler, Cristian Zanoci, and Mikhail D Lukin. Quantum sampling algorithms, phase transitions, and computational complexity. *Physical Review A*, 104(3):032602, 2021.

[46] Pawel Wocjan and Anura Abeyesinghe. Speedup via quantum sampling. *Physical Review A*, 78(4):042336, 2008.

[47] Yulin Wu, Wan-Su Bao, Sirui Cao, Fusheng Chen, Ming-Cheng Chen, Xiawei Chen, Tung-Hsun Chung, Hui Deng, Yajie Du, Daojin Fan, et al. Strong quantum computational advantage using a superconducting quantum processor. *Physical review letters*, 127(18):180501, 2021.

[48] Han-Sen Zhong, Yu-Hao Deng, Jian Qin, Hui Wang, Ming-Cheng Chen, Li-Chao Peng, Yi-Han Luo, Dian Wu, Si-Qiu Gong, Hao Su, et al. Phase-programmable gaussian boson sampling using stimulated squeezed light. *Physical review letters*, 127(18):180502, 2021.

[49] Han-Sen Zhong, Hui Wang, Yu-Hao Deng, Ming-Cheng Chen, Li-Chao Peng, Yi-Han Luo, Jian Qin, Dian Wu, Xing Ding, Yi Hu, et al. Quantum computational advantage using photons. *Science*, 370(6523):1460–1463, 2020.

APPENDICES

Appendix A

Jordan-Wigner Transformation

The exact spectrum of a number of models throughout this thesis can be found using a Jordan-Wigner transformation [35]. This is a mapping between spin-1/2 degrees of freedom and spinless fermions which can hop between sites with single orbitals. Let $a^\dagger$ create one of these spinless fermions on the site i. Then we can relate the fermionic Hilbert space to the spin-1/2 Hilbert space through the operator mapping,

$$\sigma_i^x = 2a_i^\dagger a_i - 1. \tag{A.1}$$

This mapping is sometimes instead defined in terms of σ_i^z, however a rotation of the spin variables simplifies the analysis for the models considered. There is an added complication as fermionic operators on different sites anticommute. These additional minus signs can be handled in one-dimension where the spins are naturally ordered, through a non-local string operator,

$$\begin{aligned} \sigma_i^+ &= \frac{1}{2}(\sigma_i^z + i\sigma_i^y) = e^{i\pi \sum_{j=1}^{i-1} a_j^\dagger a_j} a_i, \\ \sigma_i^- &= \frac{1}{2}(\sigma_i^z - i\sigma_i^y) = e^{i\pi \sum_{j=1}^{i-1} a_j^\dagger a_j} a_i^\dagger. \end{aligned} \tag{A.2}$$

Which gives the following form of σ_i^z,

$$\sigma_i^z = e^{i\pi \sum_{j=1}^{i-1} a_j^\dagger a_j}(a_i + a_i^\dagger) = \prod_{j<i}(1 - 2a_j^\dagger a_j)(a_i + a_i^\dagger). \tag{A.3}$$

We will see that in this fermionic representation, the Hamiltonian of these models can be exactly diagonalized.

A.1 Transverse Field Ising Model

Consider the one-dimensional transverse field Ising model with periodic boundary conditions,

$$H = -\kappa \sum_{i=1}^{N} \sigma_i^z \sigma_{i+1}^z + \eta \sum_{i=1}^{N} \sigma_i^x. \tag{A.4}$$

We will assume that we have an even number of sites, which isn't a large restriction but makes the calculations easier. When transferring into fermionic operators we need keep track of the boundary term,

$$\begin{aligned}
\sigma_N^z \sigma_1^z &= \prod_{j<N} (1 - 2a_j^\dagger a_j)(a_N + a_N^\dagger)(a_1 + a_1^\dagger) \\
&= \prod_{j<N} (1 - 2a_j^\dagger a_j)(1 - 2a_N^\dagger a_N)(1 - 2a_N^\dagger a_N)(a_N + a_N^\dagger)(a_1 + a_1^\dagger) \\
&= \prod_{j=1}^{N} (1 - 2a_j^\dagger a_j)(1 - 2a_N^\dagger a_N)(a_N + a_N^\dagger)(a_1 + a_1^\dagger) \\
&= e^{i\pi \hat{N}} (a_N - a_N^\dagger)(a_1 + a_1^\dagger).
\end{aligned} \tag{A.5}$$

The system's Hamiltonian is then,

$$H = -\kappa \left(\sum_{i}^{N-1} (a_i^\dagger - a_i)(a_{i+1} + a_{i+1}^\dagger) + e^{i\pi N} (a_N^\dagger - a_N)(a_1 + a_1^\dagger) \right) + \eta \sum_{i}^{N-1} (2a_i^\dagger a_i - 1). \tag{A.6}$$

This Hamiltonian does not conserve the total number of fermions $\hat{N}$, but the fermionic parity $p = \frac{1}{2}(1 - e^{i\pi \hat{N}})$ is conserved. This symmetry is the $Z(2)$ symmetry evident in the spin problem. This allows us to consider two subspaces independently, one associated with $p = 0$ and one with $p = 1$. We can then write our Hamiltonian of each subspace as,

$$H_p = -\kappa \sum_{i}^{N} (a_i^\dagger - a_i)(a_{i+1} + a_{i+1}^\dagger) + \eta \sum_{i}^{N} (2a_i^\dagger a_i - 1) \tag{A.7}$$

with the boundary conditions for our fermionic operators defined as,

$$a_{N+1} = (-1)^{p+1} a_1. \tag{A.8}$$

This added complication is a direct consequence of the periodic boundary conditions of our original problem. In the open boundary conditions case, these two fermionic Hamiltonians

coincide as the boundary term is the same for both (zero). To further simplify this problem, define the following momentum space operators

$$a_k = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} e^{-ikj} a_j \quad \rightarrow \quad a_j = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} e^{ikj} a_k. \tag{A.9}$$

The terms in the Hamiltonian then become,

$$\begin{aligned} \sum_i \sigma_i^x &= \sum_i (2a_i^\dagger a_i - 1) \\ &= \sum_k (2a_k^\dagger a_k - 1) \\ &= \sum_k (a_k^\dagger a_k - a_{-k} a_{-k}^\dagger) \end{aligned} \tag{A.10}$$

$$\begin{aligned} \sum_i \sigma_i^z \sigma_{i+1}^z &= \sum_i (a_i^\dagger - a_i)(a_{i+1}^\dagger + a_{i+1}) \\ &= \sum_i -a_i a_{i+1} + a_i^\dagger a_{i+1}^\dagger - a_i a_{i+1}^\dagger + a_i^\dagger a_{i+1} \\ &= \sum_k -e^{ik} a_{-k} a_k + e^{ik} a_k^\dagger a_{-k}^\dagger - e^{ik} a_{-k} a_{-k}^\dagger + e^{ik} a_k^\dagger a_k \\ &= \sum_k \Big[-\frac{1}{2} e^{ik} a_{-k} a_k + \frac{1}{2} e^{-ik} a_{-k} a_k - \frac{1}{2} e^{-ik} \\ &+ \frac{1}{2} e^{ik} a_k^\dagger a_{-k}^\dagger - \frac{1}{2} e^{-ik} a_k^\dagger a_{-k}^\dagger + \frac{1}{2} e^{-ik} \\ &- \frac{1}{2} e^{ik} a_{-k} a_{-k}^\dagger + \frac{1}{2} e^{-ik} a_k^\dagger a_k - \frac{1}{2} e^{-ik} \\ &+ \frac{1}{2} e^{ik} a_k^\dagger a_k - \frac{1}{2} e^{-ik} a_{-k} a_{-k}^\dagger + \frac{1}{2} e^{-ik} \Big] \\ &= \sum_k \cos(k) \left(a_k^\dagger a_k - a_{-k} a_{-k}^\dagger \right) - i \sin(k) \left(a_{-k} a_k - a_k^\dagger a_{-k}^\dagger \right) \end{aligned} \tag{A.11}$$

To account for the different boundary conditions in each parity sector, the set of k values is then [28],

$$K_p = \left\{ k = \pm \frac{2\pi}{N} \times \begin{cases} (l - 1/2) \quad \text{with } l = 1, 2, ..., N/2 & p = 0 \\ l \quad \text{with } l = 1, 2, ..., N/2 - 1 & p = 1 \end{cases} \right. \tag{A.12}$$

The Hamiltonian for the two parity sectors can then be written as,

$$H_p = \sum_{K_p} \left[(\eta - \kappa \cos(k)) \left(a_k^\dagger a_k - a_{-k} a_{-k}^\dagger \right) + i\kappa \sin(k) \left(a_{-k} a_k - a_k^\dagger a_{-k}^\dagger \right) \right] \tag{A.13}$$

For the $p = 1$ sector we can separate the $k = 0$ and $k = \pi$ terms,

$$\begin{aligned} H_{k=0,\pi} &= (\eta - \kappa) \left(a_0^\dagger a_0 - a_0 a_0^\dagger \right) + (\eta + \kappa) \left(a_\pi^\dagger a_\pi - a_{-\pi} a_{-\pi}^\dagger \right) \\ &= (\eta - \kappa) \left(2 a_0^\dagger a_0 - 1 \right) + (\eta + \kappa) \left(2 a_\pi^\dagger a_\pi - 1 \right) \\ &= (\eta - \kappa) \left(2\hat{n}_0 - 1 \right) + (\eta + \kappa) \left(2\hat{n}_\pi - 1 \right) \\ &= 2\kappa(\hat{n}_\pi - \hat{n}_0) + 2\eta(\hat{n}_0 + \hat{n}_\pi - 1) \end{aligned} \tag{A.14}$$

The remaining k values in K_p can now be organized in pairs $(k, -k)$. For the $p = 0$ sector this pairing can be done without separating out any terms. Define K_p^+ to be the set of positive k values of Eq. A.12, we can then write the Hamiltionian as,

$$H_0 = \sum_{K_0^+} H_k \qquad H_1 = \sum_{K_1^+} H_k + H_{k=0,\pi}, \tag{A.15}$$

with

$$\begin{aligned} H_k &= 2 \left(\eta - \kappa \cos(k) \right) \left(a_k^\dagger a_k - a_{-k} a_{-k}^\dagger \right) + 2i\kappa \sin(k) \left(a_{-k} a_k - a_k^\dagger a_{-k}^\dagger \right) \\ &= 2 \begin{pmatrix} a_k^\dagger & a_{-k} \end{pmatrix} \begin{pmatrix} \eta - \kappa \cos(k) & -i\kappa \sin(k) \\ i\kappa \sin(k) & -\eta + \kappa \cos(k) \end{pmatrix} \begin{pmatrix} a_k \\ a_{-k}^\dagger \end{pmatrix} \\ &= 2 \begin{pmatrix} a_k^\dagger & a_{-k} \end{pmatrix} \mathbf{H}_k \begin{pmatrix} a_k \\ a_{-k}^\dagger \end{pmatrix} \end{aligned} \tag{A.16}$$

The task of exactly solving this system now only requires diagonalizing this 2×2 matrix $\mathbf{H}_k$ which can be done through a Bogoliubov transformation. This transformation maps current operators onto a new set of fermionic operators through a unitary transform,

$$\gamma_k = u_k a_k - i v_k a_{-k}^\dagger \tag{A.17}$$

Where $u_k = \cos(\theta_k/2)$ and $v_k = \sin(\theta_k/2)$ is chosen to diagonalize H_k under this unitary transformation. Specifically, take

$$\begin{aligned} \cos(\theta_k) &= \frac{\eta - \kappa \cos(k)}{\sqrt{(\eta - \kappa \cos(k))^2 + (\kappa \sin(k))^2}} \\ \sin(\theta_k) &= \frac{\kappa \sin(k)}{\sqrt{(\eta - \kappa \cos(k))^2 + (\kappa \sin(k))^2}} \end{aligned} \tag{A.18}$$

Which by construction diagonalizes $\mathbf{H}_k$,

$$\begin{aligned} A^\dagger \mathbf{H}_k A &= \begin{pmatrix} \cos(\theta_k/2) & i\sin(\theta_k/2) \\ i\sin(\theta_k/2) & \cos(\theta_k/2) \end{pmatrix} \begin{pmatrix} \eta - \kappa\cos(k) & -i\kappa\sin(k) \\ i\kappa\sin(k) & -\eta + \kappa\cos(k) \end{pmatrix} \begin{pmatrix} \cos(\theta_k/2) & -i\sin(\theta_k/2) \\ -i\sin(\theta_k/2) & \cos(\theta_k/2) \end{pmatrix} \\ &= \sqrt{(\eta - \kappa\cos(k))^2 + (\kappa\sin(k))^2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned} \tag{A.19}$$

Putting this together, we have

$$\begin{aligned} H_k &= 2 \begin{pmatrix} a_k^\dagger & a_{-k} \end{pmatrix} \mathbf{H}_k \begin{pmatrix} a_k \\ a_{-k}^\dagger \end{pmatrix} \\ &= 2 \begin{pmatrix} a_k^\dagger & a_{-k} \end{pmatrix} AA^\dagger \mathbf{H}_k AA^\dagger \begin{pmatrix} a_k \\ a_{-k}^\dagger \end{pmatrix} \\ &= 2\varepsilon_k(\gamma_k^\dagger \gamma_k - 1/2) \end{aligned} \tag{A.20}$$

Where ε_k is the dispersion relation,

$$\varepsilon_k = 2\sqrt{(\eta - \kappa\cos(k))^2 + (\kappa\sin(k))^2}. \tag{A.21}$$

Since $\varepsilon_k \geq 0$ the ground state is the state which annihilates γ_k for all k, known as the Bogoliuviov vacuum. It is typically to now think of this as a two-band problem, $\pm\varepsilon_k$, where higher energy states are created by moving fermions from the lower to the upper band. From this construction, it is evident that there is a phase transition when $\kappa = \eta$ as $\varepsilon_0 = 0$ and the energy bands coincide.

This band structure exists for each parity sector, however the $p = 1$ has the additional $H_{k=0,\pi}$ term. Due to this term, the global ground state is in the $p = 0$ sector for finite system sizes. In the ferromagnetic phase, the energy difference of the ground state in both sectors closes exponentially in the system size. This is what gives the ferromagnetic degeneracy in the thermodynamic limit.

Appendix B

Ising-Chain Parent Hamiltonian

B.1 Parent Hamiltonian Derivation

Recall the form of the parent Hamiltonian,

$$H_q(\beta) = N(\mathbb{I} - e^{-\beta H_c/2} P e^{\beta H_c/2}). \tag{B.1}$$

For the one-dimensional Ising model with periodic boundary conditions, if we take local spin flip updates and Glauber acceptance probability, the transition matrix takes the form,

$$P = \sum_i p_i \sigma_i^x + \left(\mathbb{I} - \sum_i p_i \right) \tag{B.2}$$

with

$$p_i = \frac{1}{2N \cosh\big(\beta(\sigma_{i-1}^z + \sigma_{i-1}^z)\big)} e^{-\beta \sigma_i^z (\sigma_{i-1}^z + \sigma_{i-1}^z)}. \tag{B.3}$$

The parent Hamiltonian then becomes,

$$\begin{aligned} H_q &= N\left(\mathbb{I} - e^{-\beta H_c/2}\sum_i p_i\sigma_i^x e^{\beta H_c/2} - e^{-\beta H_c/2} I e^{\beta H_c/2} + e^{-\beta H_c/2}\sum_i p_i e^{\beta H_c/2}\right) \\ &= N\left(\sum_i p_i - e^{-\beta H_c/2}\sum_i p_i\sigma_i^x e^{\beta H_c/2}\right) \\ &= N\left(\sum_i p_i - \sum_i \frac{1}{2n\cosh\big(\beta(\sigma_{i-1}^z + \sigma_{i-1}^z)\big)}\sigma_i^x\right) \\ &= \frac{1}{2}\sum_i \frac{1}{\cosh\big(\beta(\sigma_{i-1}^z + \sigma_{i+1}^z)\big)}\left(e^{-\beta\sigma_i^z(\sigma_{i-1}^z+\sigma_{i+1}^z)} - \sigma_i^x\right) \end{aligned} \tag{B.4}$$

We can simplify this form further using the following expansions.

$$\begin{aligned} \frac{1}{\cosh\big(\beta(\sigma_{i-1}^z + \sigma_{i+1}^z)\big)} &= \frac{1}{2}\left[\frac{1}{\cosh(2\beta)}(I + \sigma_{i-1}^z\sigma_{i+1}^z) + (I - \sigma_{i-1}^z\sigma_{i+1}^z)\right] \\ &= \frac{1}{2}\left[\left(\frac{1}{\cosh(2\beta)} + 1\right)I + \left(\frac{1}{\cosh(2\beta)} - 1\right)\sigma_{i-1}^z\sigma_{i+1}^z\right] \end{aligned} \tag{B.5}$$

$$\begin{aligned} e^{-\beta\sigma_i^z(\sigma_{i-1}^z+\sigma_{i+1}^z)} &= e^{-\beta\sigma_i^z\sigma_{i-1}^z}e^{-\beta\sigma_i^z\sigma_{i+1}^z} \\ &= \cosh^2(-\beta)I + \sinh^2(-\beta)\sigma_{i-1}^z\sigma_{i+1}^z + \frac{1}{2}\sinh(-2\beta)(\sigma_i^z\sigma_{i-1}^z + \sigma_i^z\sigma_{i+1}^z) \end{aligned} \tag{B.6}$$

Putting these results together we obtain,

$$
\begin{aligned}
H_q &= \frac{1}{2}\sum_i \frac{1}{\cosh\big(\beta(\sigma^z_{i-1}+\sigma^z_{i+1})\big)}\left[e^{-\beta\sigma^z_i(\sigma^z_{i-1}+\sigma^z_{i+1})}-\sigma^x_i\right] \\
&= \frac{1}{4}\sum_i \left[\left(\frac{1}{\cosh(2\beta)}+1\right)\mathbb{I}+\left(\frac{1}{\cosh(2\beta)}-1\right)\sigma^z_{i-1}\sigma^z_{i+1}\right]\Big[\cosh^2(-\beta)I+\sinh^2(-\beta)\sigma^z_{i-1}\sigma^z_{i+1} \\
&\quad +\frac{1}{2}\sinh(-2\beta)(\sigma^z_i\sigma^z_{i-1}+\sigma^z_i\sigma^z_{i+1})-\sigma^x_i\Big] \\
&= \frac{1}{4}\sum_i \Bigg[2\mathbb{I}+\frac{1}{2}\sinh(-2\beta)\left(\frac{1}{\cosh(2\beta)}+1\right)(\sigma^z_i\sigma^z_{i-1}+\sigma^z_i\sigma^z_{i+1}) \\
&\quad +\frac{1}{2}\sinh(-2\beta)\left(\frac{1}{\cosh(2\beta)}-1\right)(\sigma^z_{i+1}\sigma^z_i+\sigma^z_{i-1}\sigma^z_i) \\
&\quad -\left(\frac{1}{\cosh(2\beta)}+1\right)\sigma^x_i+\left(1-\frac{1}{\cosh(2\beta)}\right)\sigma^z_{i-1}\sigma^x_i\sigma^z_{i+1}\Bigg] \\
&= \frac{N}{2}\mathbb{I}-\frac{1}{4}\sum_i\left[2\tanh(2\beta)\sigma^z_i\sigma^z_{i+1}+\left(\frac{1}{\cosh(2\beta)}+1\right)\sigma^x_i-\left(1-\frac{1}{\cosh(2\beta)}\right)\sigma^z_{i-1}\sigma^x_i\sigma^z_{i+1}\right] \\
&= \frac{N}{2}\mathbb{I}-\sum_i\left(J_1(\beta)\sigma^z_i\sigma^z_{i+1}+h(\beta)\sigma^x_i-J_2(\beta)\sigma^z_{i-1}\sigma^x_i\sigma^z_{i+1}\right)
\end{aligned}
\tag{B.7}
$$

With $h(\beta)=\frac{1}{4}(1+1/\cosh(2\beta))$, $J_1(\beta)=\frac{1}{2}\tanh(2\beta)$, and $J_2(\beta)=\frac{1}{4}(1-1/\cosh(2\beta))$.

B.2 Exact Solution of the Parent Hamiltonian

The above Hamiltonian, Eq. B.7, has terms only of σ^x and σ^z and thus can be diagonalized with a Jordan-Wigner transformation. Following the same procedure outlined for the transverse field Ising model, see Appendix A, we can rewrite H_q as,

$$
\begin{aligned}
H_q &= \frac{N}{2}\mathbb{I} + \sum_{K_p} \Big[\left(-h - J_1\cos(k) - J_2\cos(2k)\right)\left(a_k^\dagger a_k - a_{-k}a_{-k}^\dagger\right) \\
&\quad + i\left(J_1\sin(k) + J_2\sin(2k)\right)\left(a_{-k}a_k - a_k^\dagger a_{-k}^\dagger\right)\Big] \\
&= \frac{N}{2}\mathbb{I} + \sum_{K_p}\Big[A_k\left(a_k^\dagger a_k - a_{-k}a_{-k}^\dagger\right) + iB_k\left(a_{-k}a_k - a_k^\dagger a_{-k}^\dagger\right)\Big]
\end{aligned}
\tag{B.8}
$$

with $A_k = -h - J_1\cos(k) - J_2\cos(2k)$ and $B_k = J_1\sin(k) + J_2\sin(2k)$. As above we can separate the $k = 0, \pi$ terms from the $p = 1$ sector,

$$
\begin{aligned}
H_{k=0,\pi} &= (h - J_1)\left(a_0^\dagger a_0 - a_0 a_0^\dagger\right) + (h + J_1)\left(a_\pi^\dagger a_\pi - a_{-\pi}a_{-\pi}^\dagger\right) \\
&= 2J_1(\hat{n}_\pi - \hat{n}_0) - 2h(\hat{n}_0 + \hat{n}_\pi - 1).
\end{aligned}
\tag{B.9}
$$

We get the form of the parent Hamiltonian in the two sectors,

$$
H_0 = \frac{N}{2}\mathbb{I} + \sum_{K_0^+} H_k \qquad H_1 = \frac{N}{2}\mathbb{I} + \sum_{K_1^+} H_k + H_{k=0,\pi},
\tag{B.10}
$$

with

$$
\begin{aligned}
H_k &= 2A_k\left(a_k^\dagger a_k - a_{-k}a_{-k}^\dagger\right) + 2iB_k\left(a_{-k}a_k - a_k^\dagger a_{-k}^\dagger\right) \\
&= 2\begin{pmatrix} a_k^\dagger & a_{-k}\end{pmatrix}\begin{pmatrix} A_k & -iB_k \\ iB_k & -A_k\end{pmatrix}\begin{pmatrix} a_k \\ a_{-k}^\dagger\end{pmatrix} \\
&= 2\begin{pmatrix} a_k^\dagger & a_{-k}\end{pmatrix}\mathbf{H}_k\begin{pmatrix} a_k \\ a_{-k}^\dagger\end{pmatrix}
\end{aligned}
\tag{B.11}
$$

Diagonalizing, in the same way, using a Bogoliubov transformation,

$$
H_k = 2\varepsilon_k(\gamma_k^\dagger\gamma_k - 1/2)
\tag{B.12}
$$

with

$$
\varepsilon_k = \sqrt{A_k^2 + B_k^2}.
\tag{B.13}
$$

B.3 Spectral Gap Scaling

With the exact solution to the parent Hamiltonian, we can solve for the spectral gap of the mixing matrix and analyze its scaling. The ground state of H_q is in the $p = 0$ sector with $-\varepsilon_k$ for each k. The energy is then,

$$\begin{aligned} E_0^{p=0} &= \frac{N}{2} - 2\sum_{K_0^+} \sqrt{A_k^2 + B_k^2} \\ &= \frac{N}{2} - 2\sum_{K_0^+} \sqrt{\left(-h - J_1\cos(k) - J_2\cos(2k)\right)^2 + \left(J_1\sin(k) + J_2\sin(2k)\right)^2} \end{aligned} \tag{B.14}$$

For the $p = 1$ to get an odd number of fermions, we need an odd contribution from $k = 0, \pi$. The configuration that gives the lowest energy has $\hat{n_0} = 1$ and $\hat{n_0} = 0$ which gives energy $-2J_1$ (as $J_1 > 0$). So then we have,

$$\begin{aligned} E_0^{p=1} &= \frac{N}{2} - 2\sum_{K_1^+} \sqrt{A_k^2 + B_k^2} - 2J_1 \\ &= \frac{N}{2} - 2\sum_{K_1^+} \sqrt{\left(-h - J_1\cos(k) - J_2\cos(2k)\right)^2 + \left(J_1\sin(k) + J_2\sin(2k)\right)^2} - 2J_1 \end{aligned} \tag{B.15}$$

These two energies' differences match the numerical gap from diagonalizing the mixing matrix (divided by N since this Hamiltonian has an extra N compared to the mixing matrix). As $\beta \to \infty$ and $N \to \infty$ the gap closes in the $p = 0$ sector, but for these size simulations, the gap between the two sectors is much smaller. We can simplify the above result by making use of the periodicity of $-2\varepsilon_k$ which has the property that $-2\varepsilon_a - 2\varepsilon_{\pi-a} = -1$. As $\{K_0^+\}$ is an even number of points equally spaced between zero and π. We can

pair up these points to give,

$$\begin{aligned} E_0^{p=0} &= \frac{N}{2} - 2\sum_{K_0^+} \varepsilon_k \\ &= \frac{N}{2} - 2\sum_{K_0^+/\{k_1,k_{N/2}\}} \varepsilon_k - 2(\varepsilon_{k_1} + \varepsilon_{k_{N/2}}) \\ &= \frac{N}{2} - 2\sum_{K_0^+/\{k_1,k_{N/2}\}} \varepsilon_k - 1 \\ &= \frac{N}{2} - \frac{N}{2} = 0 \end{aligned} \tag{B.16}$$

For the $p = 1$ sector, a similar pairing can be made as the middle point falls on -1,

$$\begin{aligned} E_0^{p=1} &= \frac{N}{2} - 2\sum_{K_1^+} \varepsilon_k - 2J_1 \\ &= \frac{N}{2} - \frac{N}{2} + 1 - 2J_1 \\ &= 1 - \tanh(2\beta) \end{aligned} \tag{B.17}$$

Then our gap of the mixing matrix is (recall the $1/N$ is need to cancel the N in the definition of H_q)

$$\delta = \frac{1}{N}(E_0^{p=1} - E_0^{p=0}) = \frac{1}{N}(1 - \tanh(2\beta)) \tag{B.18}$$

Consider the gap δ_1 between the ground state and first excited state in the $p = 0$ sector. From the form of ε_k, the k to flip to the upper band to get the first excited state is the largest k in K_0^+. The energy of this first excited state is then,

$$\begin{aligned} E_1^{p=0} &= E_0^{p=0} + 4\varepsilon_{k_{N/2}} \\ &= 4\varepsilon_{\pi-\pi/N} \end{aligned} \tag{B.19}$$

In the large β limit, then we have $h \sim \frac{1}{4}$, $J_1 \sim \frac{1}{2}$, $J_2 \sim \frac{1}{4}$ and $\varepsilon_k = \cos^2\left(\frac{k}{2}\right)$. In this limit,

$$E_1^{p=0} = 4\sin^2\left(\frac{\pi}{2N}\right). \tag{B.20}$$

The gap δ_1 is then,

$$\begin{aligned}\delta_1 &= \frac{1}{N}(E_1^{p=0} - E_0^{p=0}) \\ &= \frac{4}{N}\sin^2\left(\frac{\pi}{2N}\right) \\ &= \frac{2-2\cos(\pi/N)}{N} \\ &\sim \frac{\pi^2}{N^3} + O(1/N^5)\end{aligned} \tag{B.21}$$

This matches δ_1 found by diagonalizing. Critically this scaling with N shows that for large systems, this gap will determine the mixing time of the chain.

B.4 Adiabatic Evolution Fidelity Calculation

Above the fermionic Hamiltonian was diagonalized with a Bogoliubov transformation. Following [45], it is convenient to map it to non-interacting spins instead.

$$\begin{aligned}\tau_k^x &= a_k^\dagger a_{-k}^\dagger + a_{-k}a_k \\ \tau_k^y &= -i\left(a_k^\dagger a_{-k}^\dagger - a_{-k}a_k\right) \\ \tau_k^z &= a^\dagger a_k - a_{-k}a_{-k}^\dagger\end{aligned} \tag{B.22}$$

These operators are defined for $0 < k < \pi$, and commute for different k values. The special cases of $k = 0, \pi$ are not relevant in the $p = 0$ sector. Since we start the adiabatic evolution in this sector and the evolution maintains parity, we only need to consider the $p = 0$ sector.

$$H_q^{p=0} = \sum_{0<k<\pi} 2\varepsilon_k\left(\cos(\theta_k)\tau_k^z + \sin(\theta_k)\tau_k^y\right) \tag{B.23}$$

Where ε_k is the same as in the Bogoliubov transformation,

$$\varepsilon_k = \sqrt{A_k^2 + B_k^2}, \tag{B.24}$$

with $A_k = -h - J_1\cos(k) - J_2\cos(2k)$ and $B_k = J_1\sin(k) + J_2\sin(2k)$. The angles θ_k are defined as,

$$\begin{aligned}\varepsilon_k\cos\theta_k &= A_k, \\ \varepsilon_k\sin\theta_k &= B_k.\end{aligned} \tag{B.25}$$

As we saw, the ground state of the fermionic model is at half-filling, we will denote this ground state $|0\rangle$. In this spin notation this ground state can be written as,

$$|0\rangle = \prod_{0<k<\pi} e^{i\theta_k \tau_k^x/2} |\text{vac}\rangle . \tag{B.26}$$

Here $|\text{vac}\rangle$ is the vacuum state of the fermions a. The excited states, with energy $4\varepsilon_k + E_0$, can be constructed by flipping a spin at momentum k,

$$|k\rangle = \tau_k^x |0\rangle . \tag{B.27}$$

B.4.1 Adiabatic Metric Derivation

With this formulation in place, we can now evolve the system on different adiabatic paths. All paths are parameterized in terms of $s \in [0, 1]$ giving $J_1(s)$, $J_2(s)$, and $h = 1$. We require this evolution to be adiabatic, which gives the constrain of Eq. 3.5, restated here,

$$\frac{dt}{ds} = \frac{1}{\epsilon}\sqrt{\sum_{\mu,\nu} g_{\mu\nu}(s)\frac{d\lambda_\mu}{ds}\frac{d\lambda_\nu}{ds}}. \tag{B.28}$$

In this parametrization, we have $\lambda_1 = J_1$ and $\lambda_2 = J_2$. Recall the definition of the adiabatic metric,

$$g_{\mu\nu} = \sum_{n>0} \frac{1}{E_n^2} \langle \partial_\mu 0 | n \rangle \langle n | \partial_\nu 0 \rangle . \tag{B.29}$$

In the definition of the ground state, the only term with λ dependence is θ_k, leading to the following derivatives,

$$\partial_\mu |0\rangle = \frac{i}{2} \sum_{0<k<\pi} \partial_\mu \theta_k |k\rangle . \tag{B.30}$$

This gives the following simplification of the adiabatic metric.

$$\begin{aligned} g_{\mu\nu} &= \sum_{n>0} \frac{1}{E_n^2} \langle \partial_\mu 0 | n \rangle \langle n | \partial_\nu 0 \rangle \\ &= \sum_k \frac{1}{64\varepsilon_k^2} (\partial_\mu \theta_k)(\partial_\nu \theta_k) \end{aligned} \tag{B.31}$$

From the definition of θ_k we have,

$$\begin{aligned}\frac{\partial\theta_k}{\partial\lambda_i} &= \frac{1}{B_k}\left(\frac{\partial\varepsilon_k}{\partial\lambda_i}\frac{A_k}{\varepsilon_k} - \frac{\partial A_k}{\partial\lambda_i}\right), \\ &= \frac{1}{\epsilon_k^2}\left(A_k\frac{\partial B_k}{\partial\lambda_i} - B_k\frac{\partial A_k}{\partial\lambda_i}\right).\end{aligned} \tag{B.32}$$

With this form, we can now calculate the metric for $\lambda_1 = J_1$ and $\lambda_2 = J_2$,

$$g = \sum_{0<k<\pi}\frac{1}{64E_k^6}\sin^2(k)\begin{pmatrix}(J_2-h)^2 & (h-J_2)(2h\cos(k)+J_1)\\ (h-J_2)(2h\cos(k)+J_1) & (2h\cos(k)+J_1)^2\end{pmatrix} \tag{B.33}$$

For some parameterized path, like those listed in Table. 3.1, we now how a explicit form of dt/ds,

$$\begin{aligned}\frac{dt}{ds} &= \frac{1}{\epsilon}\sqrt{\sum_{\mu,\nu} g_{\mu\nu}(s)\frac{d\lambda_\mu}{ds}\frac{d\lambda_\nu}{ds}}, \\ &= \frac{1}{\epsilon}\sqrt{g_{11}(s)\left(\frac{dJ_1(s)}{ds}\right)^2 + g_{12}(s)\frac{dJ_1(s)}{ds}\frac{dJ_2(s)}{ds} + +g_{21}(s)\frac{dJ_2(s)}{ds}\frac{dJ_1(s)}{ds} + g_{22}(s)\left(\frac{dJ_2}{ds}\right)^2}.\end{aligned}$$

B.4.2 Adiabatic Evolution

In the above section, we derived dt/ds, which describes the rate of change of parameters required for the adiabatic condition to hold. In order to compare the time needed to traverse a path, we will consider the fidelity after evolution,

$$\mathcal{F} = |\langle\psi(\beta)|\,\phi(t_{\text{tot}})\rangle|^2. \tag{B.34}$$

In order to obtain $\phi(t_{\text{tot}})$, we will numerically integrate the time-dependent Schrodinger equation. Since τ_k commute on different sites, we can do this for each τ_k individually. Let H_k be the section Eq. B.23 acting on the spin with momentum k,

$$\begin{aligned}H_k &= 2\varepsilon_k\left(\cos(\theta_k)\tau_k^z + \sin(\theta_k)\tau_k^y\right) \\ &= 2\varepsilon_k\begin{bmatrix}\cos(\theta_k) & -i\sin(\theta_k)\\ i\sin(\theta_k) & -\cos(\theta_k)\end{bmatrix}\end{aligned} \tag{B.35}$$

This Hamiltonian has eigenstates,

$$\left|\chi_k^-(s)\right\rangle = \frac{1}{\sqrt{2(\cos(\theta_k)+1)}}\begin{bmatrix} i\sin(\theta_k) \\ \cos(\theta_k)+1 \end{bmatrix}, \quad \left|\chi_k^+(s)\right\rangle = \frac{1}{\sqrt{2(\cos(\theta_k)+1)}}\begin{bmatrix} \cos(\theta_k)+1 \\ i\sin(\theta_k) \end{bmatrix}, \tag{B.36}$$

with eigenvalues,

$$\lambda_- = -2\varepsilon_k, \quad \lambda_+ = 2\varepsilon_k. \tag{B.37}$$

Consider a general state parameterized in terms of s,

$$|\psi_k(s)\rangle = c_k(s)\left|\chi_k^-(s)\right\rangle + d_k(s)\left|\chi_k^+(s)\right\rangle.$$

Evolving this state in time gives,

$$\begin{aligned} i\frac{d}{dt}|\psi(t)\rangle &= H(t)|\psi(t)\rangle, \\ i\frac{d}{ds}\frac{ds}{dt}|\psi(s)\rangle &= H(s)|\psi(s)\rangle, \\ i\frac{d}{ds}\left(c_k(s)\left|\chi_k^-(s)\right\rangle + d_k(s)\left|\chi_k^+(s)\right\rangle\right) &= \frac{dt}{ds}\left(-2\varepsilon_k c_k(s)\left|\chi_k^-(s)\right\rangle + 2\varepsilon_k d_k(s)\left|\chi_k^+(s)\right\rangle\right). \end{aligned} \tag{B.38}$$

Expanding the left hand side gives,

$$\begin{aligned} &i\frac{d}{ds}\left(c_k(s)\left|\chi_k^-(s)\right\rangle + d_k(s)\left|\chi_k^+(s)\right\rangle\right) = \\ &i\left(\frac{dc_k(s)}{ds}\left|\chi_k^-(s)\right\rangle + c_k(s)\frac{d}{ds}\left|\chi_k^-(s)\right\rangle + \frac{dd_k(s)}{ds}\left|\chi_k^+(s)\right\rangle + d_k(s)\frac{d}{ds}\left|\chi_k^+(s)\right\rangle\right) \end{aligned} \tag{B.39}$$

Where from the form of the eigenstates,

$$\begin{aligned} \frac{d}{ds}\left|\chi_k^-(s)\right\rangle &= \frac{d}{ds}\frac{1}{\sqrt{2(\cos(\theta_k)+1)}}\begin{bmatrix} i\sin(\theta_k) \\ \cos(\theta_k)+1 \end{bmatrix} \\ &= \frac{d\theta}{ds}\frac{1}{\sqrt{2(\cos(\theta_k)+1)}}\begin{bmatrix} \frac{i}{2}(\cos(\theta_k)+1) \\ -\frac{1}{2}\sin(\theta_k) \end{bmatrix} \\ &= -\frac{i}{2}\frac{d\theta}{ds}\left|\chi_k^+(s)\right\rangle \end{aligned} \tag{B.40}$$

$$\frac{d}{ds}\left|\chi_k^+(s)\right\rangle = \frac{i}{2}\frac{d\theta}{ds}\left|\chi_k^-(s)\right\rangle \tag{B.41}$$

Combining these results together, we get the form of the Schrodinger equation presented in [45],

$$i\frac{d}{ds}\begin{bmatrix} c_k \\ d_k \end{bmatrix} = \begin{bmatrix} -2\varepsilon_k(s)\frac{dt}{ds} & -\frac{i}{2}\frac{d\theta_k}{ds} \\ \frac{i}{2}\frac{d\theta_k}{ds} & 2\varepsilon_k(s)\frac{dt}{ds} \end{bmatrix}\begin{bmatrix} c_k \\ d_k \end{bmatrix}.$$

With the result of the previous section as well as,

$$\frac{d\theta_k}{ds} = \frac{1}{E_k^2}\left(A_k\frac{dB_k}{ds} - B_k\frac{dA_k}{ds}\right), \tag{B.42}$$

$$\frac{dA_k}{ds} = -\cos(k)\frac{dJ_1(s)}{ds} - \cos(2k)\frac{dJ_2(s)}{ds}, \tag{B.43}$$

$$\frac{dB_k}{ds} = \sin(k)\frac{dJ_1(s)}{ds} + \sin(2k)\frac{dJ_2(s)}{ds}, \tag{B.44}$$

all terms in this Schrodinger equation are found and it can be numerically integrated. Since the adiabatic evolution begins in the ground state, the initial conditions of this integration are $c_k(0) = 1$ and $d_k(0) = 0$. The final fidelity is then the overlap with the ground state,

$$\mathcal{F} = \prod_{0<k<\pi} |c_k(1)|^2. \tag{B.45}$$

Appendix C

Bottleneck Ratio Derivations

C.1 Ising-Chain Ground State Overlap

Recall in Chapter 4, to calculate the bottleneck ratio of the set $A_1 = \{x_0^1\}$ we required the probability to move between the two classical ground states, $|\langle x_0^1| U |x_0^2\rangle|^2$. These states were then written in terms of the lowest energy states of the $p = 0$ and $p = 1$ sectors, $|gs_0\rangle$ and $|gs_1\rangle$. We need to find the four terms of Eq. 4.24, repeated here,

$$|\langle x_0^1| U |x_0^2\rangle|^2 = \frac{1}{4}\Big[|\langle gs_0| U |gs_0\rangle|^2 + |\langle gs_1| U |gs_1\rangle|^2 - \langle gs_0| U |gs_0\rangle \langle gs_1| U^\dagger |gs_1\rangle - \langle gs_1| U |gs_1\rangle \langle gs_0| U^\dagger |gs_0\rangle\Big]. \tag{C.1}$$

The free fermion solution to the transverse field Ising model, seen in Appendix A, transforms the full Hamiltonian into a sum over 2×2 blocks $\mathbf{H}_k$, labelled by the allowed k values. This process of writing the spin variables in terms of Fourier-transformed fermionic operators is the same for any values of κ and η. It is only the diagonalization of these 2×2 blocks that depend on the specific parameter values. We can then write this evolution $U = e^{-Ht}$ in this same basis, labelled as $\mathbf{U}_k$.

C.1.1 $p = 0$ Sector

We can write the classical ground state in the $p = 0$ sector as the TFIM ground state with $\kappa = 1$ and $\eta = 0$. I will label this state as $|gs_0\rangle = \prod_{K_0^+} |\emptyset\rangle_k^c$. Recall the matrix $\mathbf{H}_k$

was diagonalized with the transformation unitary $A(\theta_k^c)$ to the Bogoliubov fermions. Our goal is to evolve our classical ground state under different κ and η parameters. We will do this by writing $|\emptyset\rangle_k^c$ in terms of the Bogoliubov fermions specified for these κ and η. In this basis, our U will be diagonal and we can easily calculate this overlap. Let's begin by undoing the transformation $A(\theta_k^c)$ which was used to get $|\emptyset\rangle_k^c$.

$$\begin{aligned} A(\theta_k^c)A(\theta_k^c)^\dagger |\emptyset\rangle_k^c &= A(\theta_k^c) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta_k^c/2) & -i\sin(\theta_k^c/2) \\ -i\sin(\theta_k^c/2) & \cos(\theta_k^c/2) \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} -i\sin(\theta_k^c/2) \\ \cos(\theta_k^c/2) \end{pmatrix} \end{aligned} \tag{C.2}$$

Now transform this state into the new Bogoliubov fermions, γ_k, where θ_k is defined with κ and η.

$$\begin{aligned} A(\theta_k)^\dagger \begin{pmatrix} -i\sin(\theta_k^c/2) \\ \cos(\theta_k^c/2) \end{pmatrix} &= \begin{pmatrix} \cos(\theta_k/2) & i\sin(\theta_k/2) \\ i\sin(\theta_k/2) & \cos(\theta_k/2) \end{pmatrix} \begin{pmatrix} -i\sin(\theta_k^c/2) \\ \cos(\theta_k^c/2) \end{pmatrix} \\ &= \begin{pmatrix} -i\sin(\theta_k^c/2)\cos(\theta_k/2) + i\cos(\theta_k^c/2)\sin(\theta_k/2) \\ \sin(\theta_k^c/2)\sin(\theta_k/2) + \cos(\theta_k^c/2)\cos(\theta_k/2) \end{pmatrix} \\ &= \begin{pmatrix} -i\sin((\theta_k^c - \theta_k)/2) \\ \cos((\theta_k^c - \theta_k)/2) \end{pmatrix} \end{aligned} \tag{C.3}$$

Now we are now in the diagonal basis of our desired evolution.

$$\begin{aligned} \langle\emptyset|_k^c U_k |\emptyset\rangle_k^c &= \begin{pmatrix} i\sin((\theta_k^c - \theta_k)/2) & \cos((\theta_k^c - \theta_k)/2) \end{pmatrix} \begin{pmatrix} e^{-i\varepsilon_k} & 0 \\ 0 & e^{i\varepsilon_k} \end{pmatrix} \begin{pmatrix} -i\sin((\theta_k^c - \theta_k)/2) \\ \cos((\theta_k^c - \theta_k)/2) \end{pmatrix} \\ &= \sin^2((\theta_k^c - \theta_k)/2)e^{-i\varepsilon_k} + \cos^2((\theta_k^2 - \theta_k)/2)e^{i\varepsilon_k} \end{aligned} \tag{C.4}$$

We can simplify by solving for θ_k^c where $\kappa = 1$ and $\eta = 0$,

$$\begin{aligned} \cos(\theta_k^c) &= \frac{\eta - \kappa\cos(k)}{\sqrt{(\eta - \kappa\cos(k))^2 + (\kappa\sin(k))^2}} \\ &= -\cos(k) \\ &= \cos(-k + \pi) \end{aligned} \tag{C.5}$$

$$\begin{aligned}\sin(\theta_k^c) &= \frac{\kappa \sin(k)}{\sqrt{(\eta - \kappa\cos(k))^2 + (\kappa \sin(k))^2}} \\ &= \sin(k) \\ &= \sin(-k+\pi)\end{aligned} \tag{C.6}$$

Take $\theta_k^c = -k + \pi$ then we have $\cos(\theta_k^c/2) = \sin(k/2)$ and $\sin(\theta_k^c/2) = \cos(k/2)$.

$$\begin{aligned}\langle gs_0 | U | gs_0 \rangle &= \prod_{K_0^+} \left[\sin^2((-k+\pi-\theta_k/2)e^{-2i\varepsilon_k} + \cos^2((-k+\pi)-\theta_k/2)e^{2i\varepsilon_k}\right] \\ &= \prod_{K_0^+} \left[\cos^2((k+\theta_k)/2)e^{-2i\varepsilon_k} + \sin^2((k+\theta_k)/2)e^{2i\varepsilon_k}\right] \\ &= \prod_{K_0^+} \left[\cos(2\varepsilon_k) - \frac{i(\eta\cos(k)-\kappa)}{\epsilon_k}\sin(2\varepsilon_k)\right]\end{aligned} \tag{C.7}$$

C.1.2 $p = 1$ Sector

Things are quite similar in the $p = 1$ sector except we need to worry about $k = 0, \pi$ terms. Recall the TFIM Hamiltonian in this sector can be written as $H_1 = \sum_{K_1^+} H_k + H_{k=0,\pi}$, where

$$H_{k=0,\pi} = 2\kappa(\hat{n}_\pi - \hat{n}_0) + 2\eta(\hat{n}_0 + \hat{n}_\pi - 1) \tag{C.8}$$

Any of the other k contributes an even number of fermions, therefore we need an odd contribution from $k = 0, \pi$. The configuration that gives the lowest energy has $\hat{n_0} = 1$ and $\hat{n_0} = 0$ which gives energy -2κ. Following the same procedure as above, the classical ground state in this sector is,

$$|gs_1\rangle = a^\dagger_{k=0} \prod_{K_1^+} \begin{pmatrix} -i\sin((\theta_k^c - \theta_k)/2) \\ \cos((\theta_k^c - \theta_k)/2)\end{pmatrix} \tag{C.9}$$

This extra $k = 0$ fermion adds an extra term to the evolution,

$$\begin{aligned}\langle gs_1| U |gs_1\rangle &= e^{2i\kappa} \prod_{K_1^+} \left[\sin^2((-k+\pi-\theta_k)/2)e^{-2i\varepsilon_k} + \cos^2((-k+\pi)-\theta_k)/2)e^{2i\varepsilon_k}\right] \\ &= e^{2i\kappa} \prod_{K_1^+} \left[\cos^2((k+\theta_k)/2)e^{-2i\varepsilon_k} + \sin^2((k+\theta_k)/2)e^{2i\varepsilon_k}\right] \\ &= e^{2i\kappa} \prod_{K_1^+} \left[\cos(2\varepsilon_k) - \frac{i(\eta\cos(k)-\kappa)}{\varepsilon_k}\sin(2\varepsilon_k)\right]\end{aligned} \tag{C.10}$$

With these results from the two sectors, all terms of Eq. 4.24 can be found.

C.2 Ising-Chain Excited State Overlap

The other subset considered was $A_2 = S_1$. As seen in Eq. 4.31 this calculation required the following overlaps.

$$\sum_{x\in S_1, y\in S_0} |\langle y| U |x\rangle|^2 = \sum_{x\in S_1^{p=0}} \langle gs_0| U |x\rangle \langle x| U^\dagger |gs_0\rangle + \sum_{x\in S_1^{p=1}} \langle gs_1| U |x\rangle \langle x| U^\dagger |gs_1\rangle \tag{C.11}$$

For both sectors, the $N(N-1)/2$ first excited states are found by exciting a single fermion at momentum k from the bottom band to the top band. We will follow the same method as above, where we first consider the classical excited state in a single k block, $|x\rangle_k$.

$$\begin{aligned}A(\theta_k^c)A(\theta_k^c)^\dagger |x\rangle_k &= A(\theta_k^c)\begin{pmatrix}1\\0\end{pmatrix} \\ &= \begin{pmatrix}\cos(\theta_k^c/2)\\ -i\sin(\theta_k^c/2)\end{pmatrix}\end{aligned} \tag{C.12}$$

As above, transform this state into the new Bogoliubov fermions, γ_k, where θ_k is defined with κ and η.

$$\begin{aligned} A(\theta_k)^\dagger \begin{pmatrix} \cos(\theta_k^c/2) \\ -i\sin(\theta_k^c/2) \end{pmatrix} &= \begin{pmatrix} \cos(\theta_k/2) & i\sin(\theta_k/2) \\ i\sin(\theta_k/2) & \cos(\theta_k/2) \end{pmatrix} \begin{pmatrix} \cos(\theta_k^c/2) \\ -i\sin(\theta_k^c/2) \end{pmatrix} \\ &= \begin{pmatrix} -\cos(\theta_k^c/2)\cos(\theta_k/2) + \sin(\theta_k^c/2)\sin(\theta_k/2) \\ i\cos(\theta_k^c/2)\sin(\theta_k/2) - i\sin(\theta_k^c/2)\cos(\theta_k/2) \end{pmatrix} \\ &= \begin{pmatrix} \cos((\theta_k^c - \theta_k)/2) \\ -i\sin((\theta_k^c - \theta_k)/2) \end{pmatrix} \end{aligned} \tag{C.13}$$

We need to overlap between the classical ground state evolved under U with this classical excited state derived above. Say we are considering the excited state where the fermion at site k_1 is excited, labelled as $|\text{es}\rangle_{k_1}$,

$$|\text{es}\rangle_{k_1} = |x\rangle_{k_1} \prod_{k \neq k_1} |\emptyset\rangle_k \,. \tag{C.14}$$

Let us first consider the overlap in this k_1 block,

$$\begin{aligned} \langle x|_{k_1} U_{k_1} |\emptyset\rangle_{k_1} &= i\left(e^{2i\varepsilon_{k_1}} - e^{-2i\varepsilon_{k_1}}\right) \sin\big((\theta_{k_1}^c - \theta_{k_1})/2\big) \cos\big((\theta_{k_1}^c - \theta_{k_1})/2\big) \\ &= -\sin(2\varepsilon_{k_1}) \sin\big(\theta_{k_1}^c - \theta_{k_1}\big) \\ &= -\sin(2\varepsilon_{k_1}) \left[\sin(k_1)\frac{\eta - \kappa\cos(k_1)}{\varepsilon_{k_1}} + \cos(k_1)\frac{\kappa\sin(k_1)}{\varepsilon_{k_1}}\right] \\ &= -\frac{\eta\sin(2\varepsilon_{k_1})\sin(k_1)}{\varepsilon_{k_1}} \end{aligned} \tag{C.15}$$

The overlap for the entire state is then,

$$\begin{aligned} \langle \text{es}|_{k_1} U |gs_0\rangle &= \prod_{k \neq k_1} \langle\emptyset|_k \langle x|_{k_1} \prod_q U_q \prod_k |\emptyset\rangle_k \\ &= \prod_{k \neq k_1} \langle\emptyset|_k U_k |\emptyset\rangle_k \langle x|_{k_1} U_{k_1} |\emptyset\rangle_{k_1} \\ &= \prod_{k \neq k_1} \left[\cos^2((k+\theta_k)/2)e^{-2i\varepsilon_k} + \sin^2((k+\theta_k)/2)e^{2i\varepsilon_k}\right] \left(-\frac{\eta\sin(2\varepsilon_{k_1})\sin(k_1)}{\varepsilon_{k_1}}\right) \end{aligned} \tag{C.16}$$

Luckily this can be further simplified.

$$
\begin{aligned}
|\left\langle \mathrm{es}\right|_{k_1} U \left|gs_0\right\rangle|^2 &= \prod_{k\neq k_1}\left[1-\frac{\eta^2\sin^2(k)}{\varepsilon_k^2}\sin^2(2\varepsilon_k)\right]\left(\frac{\eta^2\sin^2(2\varepsilon_{k_1})\sin^2(k_1)}{\varepsilon_{k_1}^2}\right)\\
&= \prod_{k}\left[1-\frac{\eta^2\sin^2(k)}{\varepsilon_k^2}\sin^2(2\varepsilon_k)\right]\left[1-\frac{\eta^2\sin^2(k_1)}{\varepsilon_{k_1}^2}\sin^2(2\varepsilon_{k_1})\right]^{-1}\left(\frac{\eta^2\sin^2(2\varepsilon_{k_1})\sin^2(k_1)}{\varepsilon_{k_1}^2}\right)\\
&= \prod_{k}\left[1-\frac{\eta^2\sin^2(k)}{\varepsilon_k^2}\sin^2(2\varepsilon_k)\right]\left(\frac{\eta^2\sin^2(2\varepsilon_{k_1})\sin^2(k_1)}{\varepsilon_{k_1}^2-\eta^2\sin^2(2\varepsilon_{k_1})\sin^2(k_1)}\right)\\
&= |\left\langle gs_0\right| U \left|gs_0\right\rangle|^2\left(\frac{\eta^2\sin^2(2\varepsilon_{k_1})\sin^2(k_1)}{\varepsilon_{k_1}^2-\eta^2\sin^2(2\varepsilon_{k_1})\sin^2(k_1)}\right)
\end{aligned}
\tag{C.17}
$$