Case Studies in Quantum Adiabatic Optimization

by

David Gosset

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of
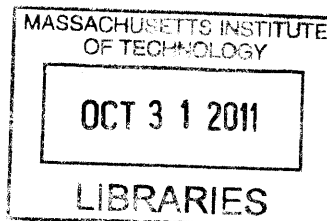
Doctor of Philosophy
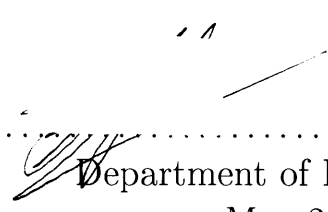
at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2011
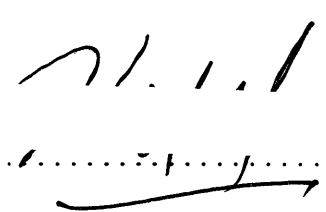
Author . . . .
Department of Physics
May 20, 2011

Certified by. . . . .
Edward Farhi
Cecil and Ida Green Professor of Physics
Thesis Supervisor

Accepted by . . . .
Krishna Rajagopal
Professor of Physics
Associate Head for Education, Physics

Case Studies in Quantum Adiabatic Optimization

by

David Gosset

Submitted to the Department of Physics
on May 20, 2011, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Quantum adiabatic optimization is a quantum algorithm for solving classical optimization problems (E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser. *Quantum computation by adiabatic evolution*, 2000. arXiv:quant-ph/0001106). The solution to an optimization problem is encoded in the ground state of a "problem Hamiltonian" H_P which acts on the Hilbert space of n spin $1/2$ particles and is diagonal in the Pauli z basis. To produce this ground state, one first initializes the quantum system in the ground state of a different Hamiltonian and then adiabatically changes the Hamiltonian into H_P. Farhi et al suggest the interpolating Hamiltonian

$$H(s) = (1-s) \sum_{i=1}^{n} \left(\frac{1 - \sigma_x^i}{2} \right) + s H_P \tag{1}$$

where the parameter s is slowly changed as a function of time between 0 and 1. The running time of this algorithm is related to the minimum spectral gap of $H(s)$ for $s \in [0,1]$.

We study such transverse field spin Hamiltonians using both analytic and numerical techniques. Our approach is example-based, that is, we study some specific choices for the problem Hamiltonian H_P which illustrate the breadth of phenomena which can occur. We present

I A random ensemble of 3SAT instances which this algorithm does not solve efficiently. For these instances $H(s)$ has a small eigenvalue gap at a value $s^\star$ which approaches 1 as $n \to \infty$.

II Theorems concerning the interpolating Hamiltonian when H_P is "scrambled" by conjugating with a random permutation matrix.

III Results pertaining to phase transitions that occur as a function of the transverse field.

IV A new quantum monte carlo method which can be used to compute ground state properties of such quantum systems.

We discuss the implications of our results for the performance of quantum adiabatic optimization algorithms.

Thesis Supervisor: Edward Farhi
Title: Cecil and Ida Green Professor of Physics

Acknowledgments

Contents

List of Figures

11

Chapter 1

Introduction

The subject of this thesis is the *quantum adiabatic algorithm*[21], which is a quantum algorithm for solving optimization problems.

Are there optimization problems which the quantum adiabatic algorithm can solve faster than the best classical algorithm? Answering this question is the objective of research on quantum adiabatic algorithms. It is a very difficult open question.

In this thesis we study the performance of the quantum adiabatic algorithm, focusing on a set of examples where we are able to make progress. The goal is to understand where to look for optimization problems on which the adiabatic algorithm performs well. Along the way we will develop technical tools which can be used to study transverse field quantum spin systems. We hope that the reader with some knowledge of classical and quantum computation (having perhaps already studied the standard text [52]) will find this thesis accessible.

We now describe the quantum adiabatic algorithm [21].

The Quantum Adiabatic Algorithm

The optimization problems to be solved have the following form. Given a cost function $f(z)$ which assigns a real number to each n-bit string, the goal is to find the bit string z_0 which minimizes f. Assume for simplicity that $f(z) > f(z_0)$ for $z \neq z_0$.

A first step in describing the quantum adiabatic algorithm is to rephrase the optimization problem to be solved. Consider the Hilbert space of n spin $\frac{1}{2}$ particles. Write each Pauli z-basis state of the n spin system as a bit string, where 0 is spin up and 1 is spin down. Then define the "Problem Hamiltonian"[21]

$$H_P = \sum_z f(z)|z\rangle\langle z| \tag{1.1}$$

which has the property that the ground state of this Hamiltonian is the basis state $|z_0\rangle$ corresponding to the bit string we are looking for.

Now let's suppose that we actually have a quantum system of n spin $\frac{1}{2}$ particles, and that we're able to control the Hamiltonian of the system. A simple version of the quantum adiabatic algorithm works as follows.

We assume that we can prepare a product state of the n spins

$$|\psi_B\rangle = \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right)^{\otimes n}.$$

Start by setting the Hamiltonian of the system to be

$$H_B = \sum_{i=1}^{n}\left(\frac{1 - \sigma_x^i}{2}\right) \tag{1.2}$$

which we call the beginning Hamiltonian[21], and initialize the state of the n spins in the state $|\psi_B\rangle$ which is the ground state of this Hamiltonian. Then slowly change the Hamiltonian of the system over a time T so that, at the end of the evolution, the Hamiltonian has been changed to H_P.

We now review the quantum adiabatic theorem (see for example [45]) which applies to very slowly changing Hamiltonians. Consider a Hamiltonian $H(s)$ which depends on a parameter $s \in [0,1]$ and which has a unique ground state at each value of s. The quantum adiabatic theorem can be applied under some fairly general conditions on the Hamiltonian $H(s)$. Suppose that we initialize a quantum system in the ground state $|\psi(0)\rangle$ of the Hamiltonian $H(0)$ and then change the Hamiltonian $H(s)$ as a function of time along the path $s(t) = \frac{t}{T}$. The quantum system evolves according to the Schrödinger equation

$$i\frac{d}{dt}|\psi(t)\rangle = H(s(t))|\psi(t)\rangle \tag{1.3}$$

(setting $\hbar = 1$). The quantum adiabatic theorem says that, for any $s \in [0,1]$, by taking T large enough the state $|\psi(sT)\rangle$ can be made arbitrarily close to the instantaneous ground state of $H(s)$.

In the quantum adiabatic algorithm we take the interpolating Hamiltonian

$$H(s) = (1 - s)H_B + sH_P \tag{1.4}$$

where $s(t) = \frac{t}{T}$. So if we choose T large then at the end of the time evolution the quantum system will be very close to the ground state of H_P. After the adiabatic evolution, the quantum system is measured in the Pauli z basis. Since the quantum system is very close to the state $|z_0\rangle$ this measurement will likely result in the output bit string z_0 which solves the optimization problem of interest.

From the adiabatic theorem we know that the quantum adiabatic algorithm prepares the state $|z_0\rangle$ in the limit $T \to \infty$. How big does T have to be to get a probability of finding the state $|z_0\rangle$ at the final measurement which is larger than some fixed threshold probability p? (A constant success probability is good enough because it can be amplified by running the algorithm independently a fixed number of times). The finite running time behaviour of quantum adiabatic evolution is governed by the adiabatic approximation which tells us how large to make T in order to get a constant success probability. A precise version of the adiabatic approximation is

given in reference [32].

A simpler question is that of computational efficiency. An efficient algorithm has a running time which grows as a polynomial function of the input size (in this case n), whereas an inefficient algorithm has a running time which is superpolynomial in n. Computational efficiency is a crude but useful measure of whether or not an algorithm is scalable in practice. For which problems is the quantum adiabatic algorithm efficient? It turns out that to answer this question the relevant property of the interpolating Hamiltonian 1.4 is the minimum spectral gap

$$g_{\text{min}} = \min_{s \in [0,1]} \left(E_1(s) - E_0(s) \right)$$

where $E_1(s)$ is the first excited state energy and $E_0(s)$ is the ground state energy. Let us assume that the norm of the Hamiltonian $H(s)$ is upper bounded for all s by a polynomial function of n. Then the adiabatic approximation tells us that the quantum adiabatic algorithm is efficient when g_{min} is asymptotically larger than an inverse polynomial in n [21]. On the other hand if g_{min} is exponentially small as a function of n then the adiabatic algorithm will be inefficient.

We now comment briefly on our motivation for studying quantum adiabatic optimization. As we have already mentioned, the primary goal is to look for problems on which the quantum adiabatic algorithm outperforms classical algorithms. Other quantum algorithms such as Shor's algorithm [59] for factoring integers and Grover's algorithm [26] for unstructured database search are remarkable because the running time of these algorithms is asymptotically faster than the running time of the best classical algorithm to date (in the case of factoring integers) or the best possible classical algorithm (in the case of unstructured database search) for the same task. These quantum algorithms were developed in the "circuit model" of quantum computation–a standard theoretical framework for quantum algorithms. Quantum adiabatic optimization is a different approach (although quantum adiabatic optimization with a local Hamiltonian can be simulated efficiently by a traditional quantum computer) and we do not yet know of any quantum speedups using this method.

A second reason to study quantum adiabatic optimization is to better understand the computational limitations which are imposed by the specific form of the interpolating Hamiltonians used. Transverse field spin Hamiltonians have a special property called stoquasticity [10] which means that all off diagonal matrix elements are nonpositive. As we will see in section 1.2, this special property can be exploited by various classical algorithms which simulate the quantum system of interest. Whether or not quantum adiabatic evolution with a stoquastic local Hamiltonian can be *efficiently* simulated on a classical computer is an open question (no such efficient simulation technique exists to date). Studying quantum adiabatic optimization by example may give some insight into the similarities and differences between quantum adiabatic optimization and classical optimization methods. We return to this point in section 1.4.

The rest of this Chapter is organized as follows. In section 1.1 we review some problems where it has been possible to determine the running time of the adiabatic

algorithm. In our discussion we will preview some of the examples which are treated in subsequent Chapters. In section 1.2 we introduce the numerical and analytic techniques that we will use in this thesis. Finally, in section 1.3 we give an outline of the rest of the thesis. In section 1.4 we conclude with an overview of complexity theoretic considerations which are relevant to the quantum adiabatic algorithm.

1.1 Performance of Quantum Adiabatic Algorithms

In this section we review progress towards understanding the running time of quantum adiabatic algorithms.

The first examples we review in section 1.1.1 all have the property that the function to be minimized has very little structure. In these cases it has been proven that the quantum adiabatic algorithm cannot succeed in subexponential time [17, 16].

We then discuss the case[15, 21, 19] where the function to be minimized only depends on the Hamming weight of its argument (an n bit string). In section 1.1.2 we review the strategy introduced in references [19, 21, 15] for analyzing the quantum adiabatic algorithm applied to such problems.

In section 1.1.3 we review results that pertain to quantum adiabatic algorithms for clause based decision problems such as 3SAT, Exact Cover, and XORSAT.

We conclude in section 1.1.4 with a discussion of the relationship between phase transitions and quantum adiabatic algorithms.

1.1.1 Grover and Scrambled Problem Hamiltonians

Grover's algorithm [27] is a quantum algorithm (originally defined in the circuit model of quantum computation) which solves the following problem. The goal is to find a special bit string z_0 out of the 2^n possible n-bit strings. In order to determine the bit string z_0 we are allowed to make queries to an "oracle" which takes as input a given bit string y and outputs "yes" if $y = z_0$ and "no" otherwise. This oracle is given to the user as a "black box" unitary transformation.

A different version of Grover's problem where the oracle is given as a Hamiltonian was defined in [17]. Here we are given the problem Hamiltonian

$$H_P = 1 - |z_0\rangle\langle z_0| \tag{1.5}$$

as a "black box" which we are allowed to apply to our system, and the goal is to find z_0. The quantum adiabatic algorithm with interpolating Hamiltonian

$$H(s) = (1 - s) \sum_{i=1}^{n} \left(\frac{1 - \sigma_x^i}{2} \right) + s \left(1 - |z_0\rangle\langle z_0| \right)$$

was studied in [21] where it was shown that for this problem the minimum gap is asymptotically $\sim 2 \cdot 2^{-\frac{n}{2}}$ and the running time of the adiabatic algorithm is exponential in n. (This followed previous work [17] which demonstrated upper and lower
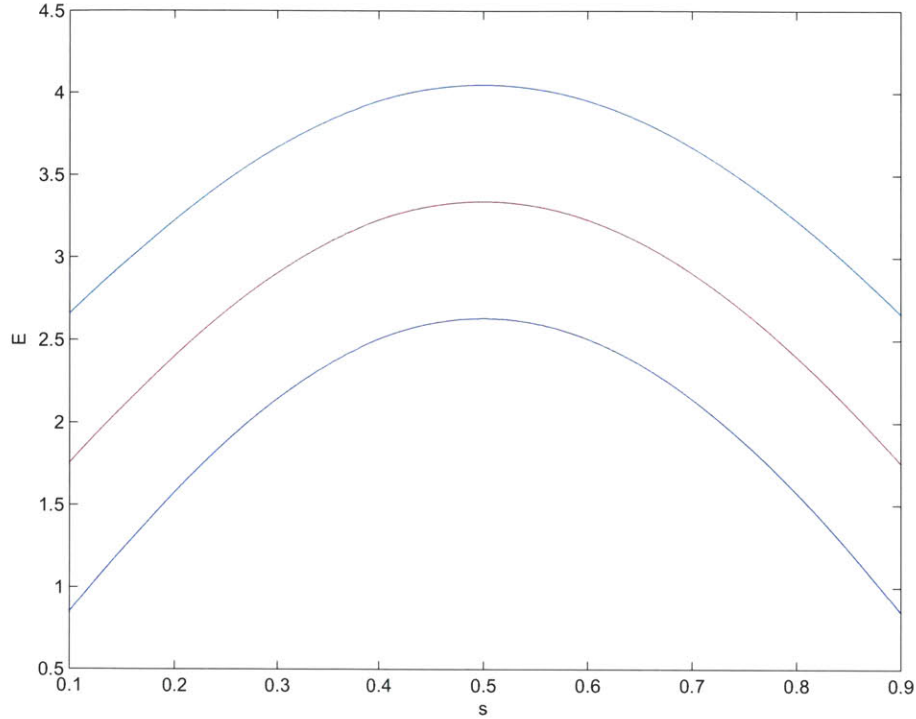
Figure 1-1: (From [13]) Lowest 25 levels for the unscrambled Hamiltonian $H_{\mathbb{I}}$ at $n = 18$.

bounds for solving Grover's problem within a more general "continuous time" model of quantum computation).

An extension of these results was given in [16] where it was shown that quantum adiabatic optimization cannot achieve success in subexponential time when the beginning Hamiltonian H_B is a rank 1 projector (as opposed to the Grover case where the problem Hamiltonian is rank 1).

Reference [16] also discussed quantum adiabatic algorithms for a class of problem Hamiltonians H_P which are "scrambled". Given a function $h(z)$ defined over the set of n bit strings and a permutation π which permutes the set of all 2^n n-bit strings, define

$$h_\pi(z) = h(\pi^{-1}(z)).$$

Each permutation π corresponds to a particular scrambled function h_π, so there are $2^n!$ possible scrambled functions.

In [16] it was proven that the quantum adiabatic algorithm cannot efficiently find the bit string which minimizes h_π for a randomly chosen π (with high probability over the choice of π) [16]. The theorem proven in [16] applies to a more general class of algorithms. It is (in a slightly adapted form)

17
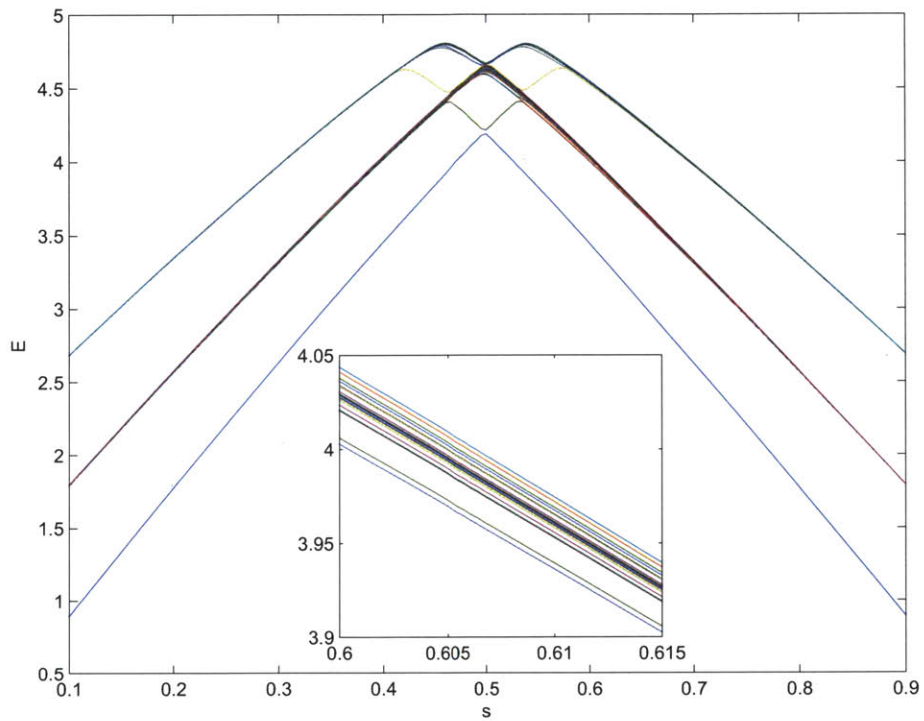
Figure 1-2: (From [13]) Lowest 25 energy levels for an instance of the scrambled Hamming weight problem $H_\pi(s)$ with a random permutation at $n = 18$. The inset shows a magnified view of levels 2 through 19 near $s = 0.6$.

Scrambled Theorem (modified from [16]). *Let $h(z)$ be a cost function, with $h(0) = 0$ and $h(1), h(2), ..., h(N-1)$ all positive. Let π be a permutation on N elements, and $H_D(t)$ be an arbitrary π-independent Hamiltonian. Consider the Hamiltonian*

$$\tilde{H}_\pi(t) = H_D(t) + c(t) \left(\sum_{z=0}^{N-1} h(\pi^{-1}(z))|z\rangle\langle z| \right),$$

where $|c(t)| \leq 1$ for all t. Let $|\psi_\pi(T)\rangle$ be the state obtained by Schrodinger evolution governed by $\tilde{H}_\pi(t)$ for time T, with a π-independent starting state. The bit string which minimizes the scrambled cost function is $\pi(0)$ and algorithmic success corresponds with measuring the basis vector $|\pi(0)\rangle$ at the end of the Schrodinger time evolution. Suppose that the success probability $|\langle\psi_\pi(T)|\pi(0)\rangle|^2 \geq \frac{1}{2}$, for a set of $\epsilon N!$ permutations. Then

$$T \geq \frac{\epsilon^2\sqrt{N}}{64h^\star} \ for \ N \geq \frac{256}{\epsilon}.$$

where

$$h^\star = \left(\frac{\sum_z h(z)^2}{N-1} \right)^{\frac{1}{2}}.$$

The above simplified version of the theorem bounds the time required to obtain success probability at least $\frac{1}{2}$ –the original theorem from [16] gives a similar bound for any choice of success probability.

The authors of [16] discussed the lessons learned from this result for quantum adiabatic algorithm design. A scrambled function h_π leads to a problem Hamiltonian

$$H_{P,\pi} = \sum_z h_\pi(z)|z\rangle\langle z|$$

and the interpolating Hamiltonian

$$H_\pi(s) = (1-s) \sum_{i=1}^{n} \left(\frac{1 - \sigma_x^i}{2} \right) + sH_{P,\pi}. \tag{1.6}$$

In the quantum adiabatic algorithm $s = s(t)$ is changed as a function of time. The Scrambled Theorem can be applied with the identifications $c(t) = s(t)$ and

$$H_D(t) = (1 - s(t)) \sum_{i=1}^{n} \left(\frac{1 - \sigma_x^i}{2} \right). \tag{1.7}$$

An example considered in [16] is where $h(z)$ computes the Hamming weight (number of ones) of the bit string z. In that case the unscrambled Hamiltonian $H_{P,\mathbb{I}}$ (where $\mathbb{I}$ is the identity permutation) is

$$H_{P,\mathbb{I}} = \sum_{i=1}^{n} \left(\frac{1 - \sigma_z^i}{2} \right)$$

19

and the Hamiltonian $H_{\mathbb{I}}(s)$ can be written as a sum of n single spin terms with no interactions. It was pointed out in [16] that the Hamiltonian $H_\pi(s)$ for most permutations π does not have the "bit structure" which is present in the Hamiltonian $H_\pi(s)$. The results of [16] show that the quantum adiabatic algorithm will not succeed on the scrambled Hamming weight problem when the permutation π is chosen randomly. This is in contrast to the unscrambled case where the minimum gap of $H_{\mathbb{I}}(s)$ is independent of n and the quantum adiabatic algorithm is efficient. It is argued in [16] that in this example the lack of bit structure is responsible for the failure of the quantum adiabatic algorithm.

It was also noted [16] that in the scrambled Hamming weight case the ground state energy of $H_\pi(s)$ has an interesting generic behaviour for randomly chosen permutations π. The spectrum of $H_{P,\pi}$ does not depend on π. But the spectrum of the Hamiltonian $H_\pi(s)$ does. In Figure 1-1 we show the spectrum of the decoupled Hamiltonian $H_{\mathbb{I}}(s)$ and in Figure 1-2 the spectrum of the Hamiltonian $H_\pi(s)$ for a random choice of the permutation π (these plots are reproduced from [13]). It was pointed out in [16] that as n increases the ground state energy of $H_\pi(s)$ appears to approach two straight lines which meet at $s = \frac{1}{2}$ (when π is chosen randomly). We can see this limiting behaviour already in Figure 1-2.

In Chapter 4 we will return to the subject of scrambled Hamiltonians where we will prove convergence results for the ground state energy and minimum eigenvalue gap for such models. Applying our results to the scrambled Hamming weight problem proves that the ground state energy per spin approaches two straight lines in the limit $n \to \infty$ and that the minimum eigenvalue gap g_{min} for this problem is exponentially small.

1.1.2 Cost Functions which depend only on the Hamming Weight

Quantum adiabatic algorithms for symmetric cost functions $f(z)$, that is, cost functions which depend only on the Hamming weight of their input, have been analyzed in references [15, 21, 19]. The Hamming weight of a bit string z is defined to be the number of ones, and is therefore always an integer between 0 and n. Obviously there is an efficient classical algorithm which finds a bit string z_0 for which $f(z)$ is minimal: for each $j \in \{0, 1, ...n\}$ choose a bit string z_j with Hamming weight j and compute $f(z_j)$. The minimum of f is then achieved by one of these $\{z_j\}$. It is not hoped that a quantum adiabatic algorithm achieves a speedup over a classical algorithm for any problem like this. The study of quantum adiabatic algorithms for symmetric cost functions is useful because they can be analyzed easily (as we discuss below), and because the lessons learned may help us understand quantum adiabatic algorithms for more difficult computational problems. We now review the techniques that have been used to study quantum adiabatic algorithms applied to such problems.

Write

$$H_P = \sum_z f(z)|z\rangle\langle z| \qquad (1.8)$$

where $f(z)$ depends only on the Hamming weight of z. If we add to this problem Hamiltonian 1.8 a beginning Hamiltonian

$$H_B = C \sum_{i=1}^{n} \left(\frac{1 - \sigma_x^i}{2} \right)$$

(where C is a constant that may depend on n) then the full Hamiltonian

$$H(s) = (1 - s) H_B + s H_p$$

is a function of the total spin operators

$$S_x = \frac{1}{2} \sum_{i=1}^{n} \sigma_x^i$$

and

$$S_z = \frac{1}{2} \sum_{i=1}^{n} \sigma_z^i.$$

The operators S_z and S_x both commute with $S^2 = S_x^2 + S_z^2 + S_y^2$. It is easy to verify that the Hamiltonian $H(s)$ commutes with S^2 since the problem Hamiltonian is a function of S_z and the beginning Hamiltonian is a function of S_x. Because of this symmetry, the ground state of the system remains in the $n + 1$ dimensional symmetric subspace throughout the quantum adiabatic evolution and so the only relevant eigenvalue gap is the gap within this subspace. This gap can be computed using numerical diagonalization up to very large values of n. This strategy has been used to convincingly obtain the asymptotic scaling of the gap with n for a number of different examples[21, 19, 15].

A second approach[15] for analyzing the quantum adiabatic algorithm applied to a symmetric cost function can be applied when (for large n) the function to be minimized is of the form

$$f(z) = n^r V(\frac{W(z)}{n}) + O(n^{r-1}) \tag{1.9}$$

(where $W(z)$ is the hamming weight function and $r \in 0, 1, 2, ...$) so that in the limit $n \to \infty$ the dependence of f on the Hamming weight is captured by the "intensive" function $V(u)$ with $u \in [0, 1]$. This function $V(u)$ is an energy landscape for the problem at $s = 1$ in the limit $n \to \infty$. In reference [15] it is shown that one can extend this quantity to an "effective potential" $V(u, s)$ whose minimum characterizes the ground state at any value of s in the limit $n \to \infty$ (however when $s \neq 1$ the quantity u no longer represents the rescaled Hamming weight of a bit string). The behaviour of this potential as a function of s can be used to infer whether or not an exponentially small gap is present for finite size systems[15].

In [19] an algorithmic strategy called "path change" was introduced and applied to symmetric cost functions. For some problem Hamiltonians the linear interpolation

1.4 between the initial Hamiltonian and the problem Hamiltonian may not be a good choice. Path change is the idea of randomly choosing interpolating Hamiltonians $H_R(s)$ which end at the given problem Hamiltonian $H_R(1) = H_P$ and running the quantum adiabatic algorithm for each random choice in the hopes that one of the runs will be successful. In [19] this path change strategy is applied to a symmetrized problem Hamiltonian for which the quantum adiabatic algorithm with the Hamiltonian 1.4 does not succeed. In that case it is shown [19] that path change can be used to give an efficient quantum adiabatic algorithm for this symmetrized problem. Farhi et al. suggest that path change may be a useful strategy to use with the quantum adiabatic algorithm applied to problems which are not symmetrized. This was the main purpose of introducing path change, since (as we have already discussed) symmetrized problems are not interesting from a computational point of view. In Chapter 3 we will see an example of path change applied to a non-symmetric problem.

1.1.3 Clause Based Decision Problems

In this section we review results pertaining to the performance of the quantum adiabatic algorithm on clause based decision problems such as 3SAT and Exact Cover and XORSAT. Unlike in the Grover and Scrambled cases, these problems have clause structure which an algorithm can try to exploit. And unlike symmetrized cost functions, there is presumably no simple limiting "energy landscape" for these problems as $n \to \infty$. We begin by giving some examples of these classical computational problems.

Examples of Clause Based Decision Problems

An instance of 3 Satisfiability (3SAT) on n bits is defined as follows. You are given a set of m *clauses*. A clause is a specific type of constraint on the n bits which is given as a subset $\{k_1, k_2, k_3\}$ of the n bits and a 3-bit string s. Each $k_i \in \{1, ..., n\}$ for $i \in \{1, 2, 3\}$ and $s \in \{000, 001, 010, ...111\}$. A clause is said to be *satisfied* by an n-bit string z if $z_{k_1} z_{k_2} z_{k_3} \neq s$, where z_i is the ith bit of z. Given the set of m clauses which specifies an instance of 3SAT, the question is to determine whether or not there exists an n-bit string z which satisfies all of the clauses. 3SAT is a decision problem, meaning the answer is either yes (if the instance is satisfiable) or no (otherwise).

Exact Cover is another example of a clause based decision problem. Like 3SAT, an instance of exact cover is specified by a set of m clauses and the question is to determine if there is an n-bit string z which satisfies all of them. Each clause in this case is represented by a subset k_1, k_2, k_3 of 3 bits. The clause is satisfied by an n-bit string if

$$z_{k_1} z_{k_2} z_{k_3} \in \{100, 010, 001\}.$$

Note that each exact cover clause can be written as a set of 5 3SAT clauses which penalize the bit strings $000, 011, 101, 110, 111$ on the given subset of bits.

A third example of a decision problem is called 3XORSAT. Here a clause on bits k_1, k_2, k_3 is a constraint of the form

$$z_{k_1} \oplus z_{k_2} \oplus z_{k_3} = b \tag{1.10}$$

where $b \in \{0, 1\}$ and the symbol $\oplus$ indicates the XOR function. The decision problem is to determine whether a set of m given constraints of the form 1.10 can be simultaneously satisfied.

How hard are the computational problems that we have just described? We briefly mention some ideas from complexity theory which help to answer this question. The complexity class NP is the class of decision problems where every "yes" instance has a proof that can be checked efficiently (on a classical computer). In 3SAT a proof consists of the n-bit string z which satisfies all of the clauses. Given this "winning" bit string, we can check that indeed all of the clauses are satisfied by it. For this reason 3SAT is contained in NP. It turns out that any instance of a decision problem in NP can be mapped into a corresponding instance of 3SAT (without expanding the number of input bits by more than a polynomial factor) [38]. This is the statement that 3SAT is NP-hard. A problem such as 3SAT which is both NP hard and in NP is said to be NP complete. Both 3SAT and Exact Cover are NP complete. Most computer scientists believe that there is no subexponential time algorithm which will solve NP complete problems. On the other hand 3XORSAT can be solved in polynomial time on a classical computer. This is the statement that 3XORSAT is in the complexity class P which consists of the set of decision problems that have polynomial time classical algorithms. To see why, note that each clause in 3XORSAT is a linear constraint over the field $\mathbb{F}_2^n$. Determining whether or not a set of linear equations have a solution can be solved in polynomial time using linear algebra.

Sometimes we will consider instances of 3SAT or other decision problems which are drawn at random from an ensemble. Properties of such random instances have been studied using methods from statistical mechanics (see [49] for a review). The canonical random ensemble of 3SAT instances is defined as follows. To generate a random instance on n bits, choose each clause uniformly at random but fix the ratio of clauses to bits $\alpha = \frac{m}{n}$. If α is sufficiently small then the instance will be satisfiable with high probability. Similarly if α is sufficiently large then a random instance drawn from this ensemble will be unsatisfiable with high probability. It is believed that there is a "phase transition" at a critical value α_c which separates a satisfiable phase from an unsatisfiable phase in the thermodynamic limit $n \to \infty$. Random instances with α close to α_c are believed to be computationally difficult for classical algorithms[49]. Similar phase transition phenomena are believed to occur in random ensembles of other clause based decision problems.

Quantum Adiabatic Algorithms for Decision Problems

A clause based decision problem can be turned into an optimization problem that can be solved by the quantum adiabatic algorithm. The strategy is to search for a bit string z which satisfies all of the clauses, using (for example) the number of violated clauses as a cost function which must be minimized. If the algorithm runs for a time T and then outputs the satisfying bit string z when the given instance is satisfiable, then we don't need to worry about what happens when the instance is not satisfiable. We can simply run the algorithm for time T , check if the output z satisfies all the clauses, and conclude that the instance is satisfiable if that is the case. That is why,

for decision problems, we can restrict to the set of satisfiable instances when studying the runtime of the quantum adiabatic algorithm.

Now let's turn to the quantum adiabatic algorithm for 3SAT from [21]. Take each clause (specified by the subset k_1, k_2, k_3 and the 3 bits $s = s_1 s_2 s_3$) and form the clause Hamiltonian

$$h_c = \left(\frac{1 + (-1)^{s_1}\sigma_z^{k_1}}{2}\right)\left(\frac{1 + (-1)^{s_2}\sigma_z^{k_2}}{2}\right)\left(\frac{1 + (-1)^{s_3}\sigma_z^{k_3}}{2}\right).$$

Summing the clause Hamiltonians associated with each of the m clauses we get the problem Hamiltonian for the given instance of 3SAT

$$H_P = \sum_{c=1}^{m} h_c \tag{1.11}$$

This Hamiltonian is diagonal in the Pauli z basis, and a bit string $|z\rangle$ has energy equal to the number of unsatisfied (violated) clauses. The form of equation 1.11 makes clear the *locality* of the problem Hamiltonian–it is a sum of terms h_c where each term h_c involves a bounded number of spins (in this case 3). If there is a zero energy state $|z_0\rangle$ then the corresponding bit string z_0 is a satisfying assignment for the instance. If we restrict to instances with a unique satisfying assignment then the eigenvalue gap of the Hamiltonian

$$H(s) = (1 - s)\sum_{i=1}^{n}\left(\frac{1 - \sigma_x^i}{2}\right) + sH_P$$

at $s = 1$ is at least 1 and the minimum gap $g_{\min}$ determines the running time of the quantum adiabatic algorithm.

Of course one can form problem Hamiltonians for other clause based problems such as Exact Cover or 3XORSAT in a similar manner.

Worst Case Performance for 3SAT

It is known that there are families of 3SAT instances for which the quantum adiabatic algorithm has running time which is exponential in the system size[61]. We now review the example from [61]. Start with the 3SAT clause which penalizes the bit string 100. This clause is associated with the 3 qubit Hamiltonian

$$h_{ijk}(\sigma_z^i, \sigma_z^j, \sigma_z^k) = \left(\frac{1 - \sigma_z^i}{2}\right)\left(\frac{1 + \sigma_z^j}{2}\right)\left(\frac{1 + \sigma_z^k}{2}\right)$$

Form a symmetrized instance of 3SAT by including this clause for each of the n^3 triplets of qubits

$$H_p' = \sum_{i,j,k=1}^{n} h_{ijk}(\sigma_z^i, \sigma_z^j, \sigma_z^k).$$

We have used the prime superscript in the above to indicate that this is an intermediate step in the construction of the problem Hamiltonian–we will modify H_P' to con-

struct the final problem Hamiltonian H_P. Since H'_P is symmetric under permutations, it can be written as a function of the Hamming weight operator $W = \sum_{i=1}^{n} \left(\frac{1-\sigma_z^i}{2} \right)$. We get

$$
\begin{aligned}
H'_P &= (n - W)\,W^2 \\
&= n^3 \left(1 - \frac{W}{n} \right) \left(\frac{W}{n} \right)^2
\end{aligned}
$$

so the rescaled cost function $g(u) = (1 - u)\,u^2$ (from equation 1.9). This cost function has two satisfying assignments, which correspond to the bit strings 000...0 and 111...1. It is possible to use the effective potential analysis of [15] to show that the quantum adiabatic algorithm with Hamiltonian

$$
H(s) = (1 - s) \sum_{i=1}^{n} \left(\frac{1 - \sigma_x^i}{2} \right) + sH'_P
$$

will end up in the solution 000...0 corresponding to the minimum at $u = 0$ as $s \to 1$. To construct the problem Hamiltonian H_P, add one additional clause to the Hamiltonian H'_P which removes the degeneracy and makes 111...1 the global minimum. The quantum adiabatic algorithm requires exponential time to find the winning bit string 111...1 in this example [61].

This example demonstrates that quantum adiabatic optimization does not succeed in the worst case for 3SAT (with the beginning Hamiltonian 1.2 and the problem Hamiltonian which counts the number of violated clauses). On the other hand the performance of the quantum adiabatic algorithm for *typical* instances of 3SAT (or other combinatorial optimization problems) drawn from a random ensemble has been more difficult to assess.

Performance on Random Instances of 3SAT and Exact Cover

In this section we will discuss results about the performance of the quantum adiabatic algorithm on random instances of clause based optimization problems. Since different studies have occasionally used different random ensembles of instances for the same problem, we will try to be clear about which random ensemble was used. The reader should be aware that not all the random instances which are discussed here are believed to be hard to solve classically.

The first numerical studies of quantum adiabatic algorithms for random instances of 3SAT [31] and Exact Cover [20] seemed to suggest that the quantum adiabatic algorithm might efficiently solve typical instances of these problems.

The authors of [20] generated random instances of Exact Cover with a unique satisfying assignment and simulated (on a classical computer) the time evolution of the quantum adiabatic algorithm applied to such instances. The random ensemble of instances which they studied was constructed by adding one clause at a time until the number of solutions became ≤ 1. If the final number of solutions after adding the last clause was zero then they started over, and if it was 1 then the instance was

kept. With this technique they were able to study the performance of the adiabatic algorithm up to $n = 20$. For each n between 10 and 20 they generated 75 random instances and simulated the adiabatic algorithm for each instance. They computed the median time to reach a constant (chosen to be $1/8$) success probability and found that this median time was fitted well to a quadratic function of n.

Another numerical study was undertaken by Hogg in [31]. Hogg studied the performance of the quantum adiabatic algorithm on random instances of 3SAT near the satisfiability phase transition. These instances are believed to be very difficult for classical algorithms. Hogg found a polynomial scaling of the running time with n for instances with n as large as 24. Hogg noted that, although the minimum gap determines the asymptotic scaling of the quantum adiabatic algorithm, for $n \leq 20$ he found using numerical diagonalization that the median of the minimum gap for the instances he considered did not vary too much with n. This led him to add the caveat "Hence, unlike for large n, the cost is not dominated by the minimum gap size and so the values of Fig. 1 may not reflect asymptotic scaling."[31] (referring to his Figure 1 which appeared to show polynomial scaling for the quantum adiabatic algorithm on the random instances he considered).

The reason that Farhi et al and Hogg were limited to studying relatively small systems is that they both used exact numerical techniques which require an exponential amount of computer memory. The dimension of the Hilbert space for an n spin quantum system is 2^n and computation involving a vector this large becomes difficult as n increases beyond 20 or so. A technique which has been developed to avoid this problem is called Quantum Monte Carlo. For now we will just mention that this method typically has significantly smaller space requirements, and often can be used at much higher bit number than techniques based on matrix or vector manipulation. Young et al. [63] were the first to apply Quantum Monte Carlo to the study of the quantum adiabatic algorithm. They studied an ensemble of random instances of Exact Cover with a unique satisfying assignment which are near the satisfiability transition for this problem (the ensemble studied by Young et al differed from the ensemble studied in [20]). They computed the median minimum gap over a set of random instances at each value of n studied, for n as large as 256. They found that this median minimum gap decreased only polynomially with n which supported the notion that the quantum adiabatic algorithm could efficiently solve typical instances drawn from this ensemble.

Despite this seemingly balmy weather for the quantum adiabatic algorithm, a new batch of papers [2, 3, 4, 5, 18] arrived which "cast clouds"[3] over the landscape. In these references a potentially dangerous problem for quantum adiabatic algorithms on random optimization problems was discussed. The central idea of all of these works is that there are certain types of instances of clause based optimization problems which may lead to failure of the quantum adiabatic algorithm. These problematic instances share some properties of the "worst case" example from [61] which we reviewed in the previous section.

We study the following random ensemble of problematic 3SAT instances (from [18]) in Chapter 3 of this thesis. Suppose that we have an instance of 3SAT with exactly 2 satisfying assignments, corresponding to a problem Hamiltonian H'_P. It will
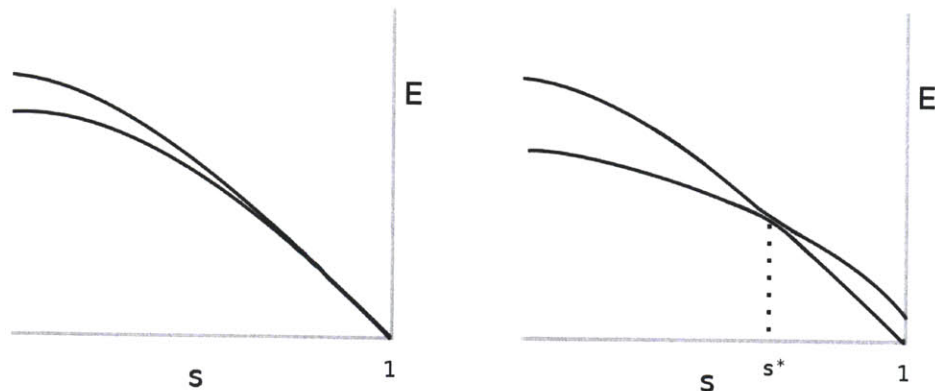
Figure 1-3: The two lowest energy levels near $s = 1$ before (left) and after (right) adding the final penalty clause. A "perturbative cross" is created by adding this single clause which breaks the degeneracy between two local minima.

be important that these two satisfying assignments differ by a number of bit flips which is proportional to n. In [18] we constructed a set of random instances where one of the satisfying assignnments is 000...0 and the other is 111...1. The Hamiltonian

$$H'(s) = (1 - s) \sum_{i=1}^{n} \left(\frac{1 - \sigma_x^i}{2} \right) + s H'_P$$

has 2 energy levels which are close together as $s \to 1$. The two lowest levels are illustrated in Figure 1-3 (left). Suppose that for s near 1 the lower energy level corresponds to the bit string 000...0. The final step in constructing the problematic instance is to add a single clause which penalizes the lower level 000...0 but not 111...1. This "drags" the lower energy level above the upper one and creates the picture in Figure 1-3 (right). At $s^\star$ there is an avoided cross and a very small gap which causes the adiabatic algorithm to fail. It turns out that the location of the avoided cross $s^\star$ approaches 1 as $n \to \infty$ in this scenario, and one can use low order perturbation theory to compute its value. This type of avoided cross has been called a "perturbative cross". The nature of the ground state changes abruptly at $s^\star$. For $s < s^\star$ most of the amplitude will be concentrated on bit strings with low Hamming weight whereas for $s > s^\star$ there is high amplitude only on states with Hamming weights close to n. It is argued in [4] that problematic instances similar to this will occur with nonneglible probability in random ensembles of optimization problems and generically lead to the failure of the quantum adiabatic algorithm.

Young et al again numerically studied random instances of Exact Cover using Quantum Monte Carlo[64]. The instances considered were chosen to have a unique satisfying assignment and ratio of clauses to bits near the phase transition. (There is a minor difference between the problem Hamiltonian H_P which was used here for each instance and that which was used in [63]–see these references for details.) This time, instead of looking at the median minimum gap, they tried to look for instances where there appeared to be a sharp change in the ground state at some value of the

27

interpolating parameter $s^\star$. In order to diagnose such a sharp change, they computed (using quantum monte carlo) the "order parameter"

$$q = \frac{1}{n} \sum_{i=1}^{n} \langle \sigma_z^i \rangle^2$$

at different values of s and observed that in some instances there is a value of s at which this order parameter changes abruptly. They found that the fraction of instances (out of 50 random instances considered at each value of n) where such an abrupt change occurs increases with n and may approach 1 for large n [64]. A recent paper of Neuhaus et. al has also reported abrupt changes in an order parameter for random instances of 3SAT [50]. We don't know if the sharp changes in the ground state observed by Young et al [64] (or the one observed by Neuhaus et al [50]) are associated with "perturbative crosses".

Performance on Random Regular Instances of XORSAT

In [21], Farhi et al showed that the quantum adiabatic algorithm can efficiently solve satisfiable instances of 2XORSAT on a ring. In this example the problem Hamiltonian to be minimized has nearest neighbor constraints on a line with periodic boundary conditions

$$H_P = \sum_{i=1}^{n} \left(\frac{1 - \sigma_z^i \sigma_z^{i+1} J_{i,i+1}}{2} \right) \tag{1.12}$$

where each $J_{i,i+1} \in \{\pm 1\}$ and $\sigma_z^{n+1} = \sigma_z^1$. This can be viewed as an instance of 2XORSAT since each term in the sum corresponds to a linear constraint on 2 variables of the form

$$z_i \oplus z_{i+1} = -J_{i,i+1}.$$

The cost function 1.12 computes the number of violated clauses. We call this 2-regular 2XORSAT since the instance is defined on a 2 regular graph (a ring). For any choice of the couplings which is not frustrated-that is, when the final constraint on bits n and 1 is consistent with all other constraints , the Hamiltonian 1.4 with the above problem Hamiltonian is exactly solvable and the spectrum does not depend on the (consistent) couplings $\{J_{i,i+1}\}$ [21]. Every satisfiable instance has the same spectrum (and consequently the same eigenvalue gap) as the transverse field ising model described by the Hamiltonian 1.12 with each $J_{i,i+1} = 1$. For the adiabatic algorithm with beginning Hamiltonian

$$H_B = \sum_{i=1}^{n} \left(\frac{1 - \sigma_x^i}{2} \right)$$

and problem Hamiltonian 1.12 it is possible to exactly compute the minimum gap as a function of n and one obtains that it scales as an inverse polynomial in n [21].

Another ensemble of XORSAT on regular graphs was studied in [37]. The problem

28

Hamiltonians considered in [37] are of the form[1]

$$H_P = \sum_c \left(\frac{1 - \sigma_z^{i_1,c} \sigma_z^{i_2,c} \sigma_z^{i_3,c} J_c}{2} \right)$$

each clause c is associated with 3 bits $i_{1,c}, i_{2,c}, i_{3,c}$ and a coupling $J_c \in \{\pm 1\}$. This problem Hamiltonian corresponds to an instance of 3XORSAT. The ensemble of instances was further restricted to those which are *3-regular* (meaning that each bit is involved in exactly 3 clauses). Using the quantum cavity method (which we will discuss further in section 1.2.4 and in Chapter 5) and Quantum Monte Carlo, Jörg et al [37] have shown that there is a first order phase transition in this model at a critical value $s \approx \frac{1}{2}$. At this critical value properties of the ground state change abruptly; for example the transverse magnetization changes discontinuously.

Jörg et al also studied the ensemble of random instances of 3 regular 3 XORSAT which are restricted to have a unique satisfying assignment. A random choice of the couplings $\{J_c\}$ and 3 regular hypergraph will produce an instance with a unique satisfying assignment with nonzero probability [37]. For instances with a unique satisfying assignment, we can always apply a sequence of bit flip operators which map the unique solution to the bit string 000...0. Conjugating the problem Hamiltonian with this unitary transformation has the effect of replacing each J_c with the value $+1$. So when restricted to instances with a unique satisfying assignment, the choice of the couplings does not affect the spectrum (as in the case of 2 regular 2XORSAT [21] reviewed in the introduction). The 3 regular hypergraph which specifies the instance is therefore the only factor which determines the spectrum in this model. Using numerical diagonalization, the authors of [37] demonstrated that the quantum adiabatic algorithm does not succeed in polynomial time for random instances of this problem due to an exponentially small gap which occurs at $s \approx \frac{1}{2}$. This agrees with the quantum cavity method and Quantum Monte Carlo results which apply to the random ensemble with no restrictions on the number of satisfying assignments.

In chapter 5 of this thesis we will demonstrate that the phase transition observed by Jörg et al. [37] in random instances of 3-Regular 3XORSAT with a unique satisfying assignment occurs at exactly $s = \frac{1}{2}$. In particular we show that there is a duality transformation which maps the problem Hamiltonian into the beginning Hamiltonian, and the beginning Hamiltonian into a problem Hamiltonian for a different instance (which occurs with equal weight in the ensemble). This demonstrates that the energy per spin, averaged over all instances, is symmetric about $s = \frac{1}{2}$. This implies that, if there is a unique phase transition as a function of s, it can only occur at $s = \frac{1}{2}$.

Note that although linear algebra solves 3XORSAT in polynomial time, heuristic algorithms such as the quantum adiabatic algorithm (which do not use linear algebra) presumably do not benefit from this fact. This is also true in the classical case–in reference [28] it was found numerically that the classical algorithm WALKSAT performs worse on random instances of 3-Regular 3XORSAT than on random instances

---

[1]We have rescaled the problem Hamiltonian and added an irrelevant constant in order to present the results of reference [37] in a consistent way with our notation.

of some NP complete problems.

1.1.4 Quantum Adiabatic Algorithms and Phase Transitions

Since we know that computing the scaling of the minimum eigenvalue gap can be difficult it would be nice to have an alternative way to study the running time of the quantum adiabatic algorithm. In some of the examples discussed above, the ground state energy per spin of the interpolating Hamiltonian $H(s)$ converges to a universal curve in the limit $n \to \infty$. Can we learn anything about the performance of the quantum adiabatic algorithm (for large but finite n) from this thermodynamic limit?

To begin to address this question let's look at simple illustrative example: the scrambled Hamming weight problem which was discussed in section 1.1.1. We will prove in Chapter 4 of this thesis that in this problem the energy per spin e for a randomly chosen scrambled instance converges to the curve

$$e(s) = \begin{cases} \frac{s}{2} & ,0 \leq s \leq \frac{1}{2} \\ \left(\frac{1-s}{2}\right) & ,\frac{1}{2} < s \leq 1 \end{cases}.$$

in the thermodynamic limit. In Figure 1-2 we can already see the two straight lines behaviour with a small example consisting of only 18 spins. This limiting curve has the property that its derivative at $s = \frac{1}{2}$ is discontinuous. This discontinuity is interpreted as a first order phase transition. The minimum gap at $s = \frac{1}{2}$ closes exponentially quickly as $n \to \infty$.

Another example where a first order phase transition occurs is the 3-regular 3XOR-SAT model studied in [37] using the quantum cavity method and Quantum Monte Carlo. There it was shown that the phase transition is first order–the first derivative of the ground state energy per spin has a discontinuity (in the limit $n \to \infty$) which was found numerically to be near $s = \frac{1}{2}$. This result applies to the ensemble of random instances without a restriction on the number of satisfying assignments. The instances with a unique satisfying occur with finite probability in this ensemble and were studied using numerical diagonalization in [37]. For finite system sizes the gap is exponentially small as a function of n near $s = \frac{1}{2}$ [37].

Second order phase transitions are also possible. In the (unfrustrated) 2XORSAT on a ring example [21] there is a limiting curve for the ground state energy per spin. Recall that for this problem each choice of couplings has the same spectrum so there is only one curve for each n. The limiting curve can be computed from the exact solution of the model which is reviewed in [21]. We have plotted the energy per spin e for this model in the thermodynamic limit in Figure 1-4. In this model the first derivative of the ground state energy per spin is continuous but its second derivative blows up at the transition point $s = \frac{1}{2}$. Here a second order phase transition is associated with a gap which scales as an inverse polynomial in n.

It has been suggested [58] that *first order* phase transitions in the $n \to \infty$ limit are typically associated with eigenvalue gaps which decrease exponentially quickly with n. This is consistent with all of the examples that we know of. When a first order phase transition occurs in a given model we expect that the quantum adiabatic algorithm
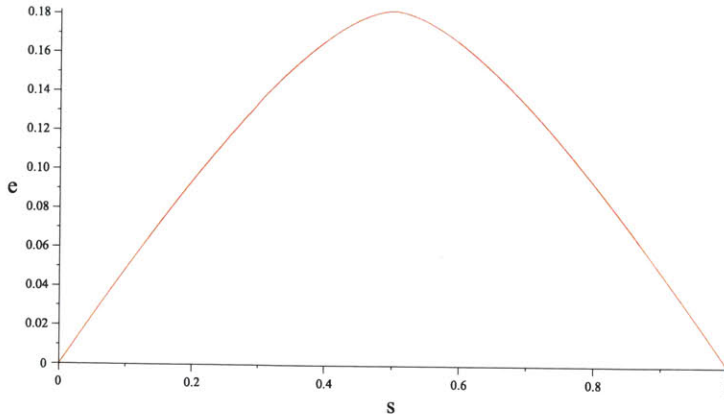
Figure 1-4: The energy per spin for (satisfiable) 2XORSAT on a ring.

does not succeed in polynomial time. However when a second order phase transition is present then we cannot make conclusions about the performance of the quantum adiabatic algorithm without further investigation of the minimum eigenvalue gap.

Perturbative crosses (which we discuss in more detail in Chapter 3) are a mechanism which leads to small gaps that is seemingly different from the phase transitions discussed above, although they have been called first order phase transitions by some authors. Is there a connection between perturbative crosses and phase transitions? We know that as $n \to \infty$ the perturbative crosses are expected to move to $s = 1$ but the effect of perturbative crosses on thermodynamic quantities is not clear.

Finally, we note that the crude distinction between "first order" (signalled by a discontinuous first derivative of the average energy per spin) and "second order" (indicating a continuous first derivative at the critical point) is not sufficient to capture the different behaviours of physical systems at quantum critical points. For example it is known in the context of quantum spin glasses that a wider range of phenomena is present (see for example[12] and [24, 23]).

1.2 Tools for Stoquastic Hamiltonians

Quantum spin Hamiltonians are difficult to study numerically because the size of the Hilbert space scales exponentially with the system size. Numerical methods which store a vector in this Hilbert space (such as numerical diagonalization) are limited to small system sizes for this reason. On the other hand there is no general analytic method that can be used to study quantum spin systems. Hamiltonians where analytic progress is possible seem to be few and far between.

In this thesis we will primarily use three technical tools to study transverse field spin Hamiltonians of the form 1.4 with H_B given by equation 1.2. The first is a variational lower bound on the ground state energy. The other tools we use are numerical methods called Quantum Monte Carlo[55, 60] and the quantum cavity method[43, 40]. All of these tools exploit a property of the Hamiltonian 1.4 called

stoquasticity [10].

A Hamiltonian H with real matrix elements is said to be stoquastic [10] with respect to a given basis $\{|z\rangle\}$ if all of its off diagonal matrix elements in this basis are nonpositive. For example the Hamiltonian in equation 1.4 has off diagonal matrix elements which are either 0 or $-\frac{1}{2}(1-s)$ in the Pauli z basis.

Many studies [10, 9, 44, 34, 11] have used complexity theory to demonstrate that computational problems involving Hamiltonians are easier when restricted to the class of stoquastic Hamiltonians. We will indeed see that stoquasticity makes our life easier in the examples considered in this thesis.

In this section we discuss the methods for stoquastic Hamiltonians that are used in subsequent Chapters of this thesis. The mathematical tools in sections 1.2.1 and 1.2.2 (including the variational lower bound on the ground state energy) consists of immediate applications of the Perron Frobenius theorem and its corollaries [46]. Quantum Monte Carlo and the quantum cavity method make use of a special feature of the thermal path integral for stoquastic Hamiltonians. We introduce this special property in section 1.2.3 and in section 1.2.4 we discuss briefly how these numerical techniques can be used to simulate properties of the Hamiltonian 1.4 on a classical computer.

The material discussed below is certainly not meant to be comprehensive–see references [10, 9, 44, 34, 11] for much more on stoquasticity. Portions of this section have appeared previously in [18, 13].

1.2.1   No Nodes

For any stoquastic Hamiltonian there is always a ground state with nonnegative amplitudes in the given basis. To see this, write

$$H = H_0 + V$$

where H_0 is diagonal in the basis $\{|z\rangle\}$ and V is off diagonal and satisfies $\langle z|V|z'\rangle \leq 0$ when $z \neq z'$. Consider a state

$$|\psi\rangle = \sum_z c_z |z\rangle$$

with all real amplitudes c_z. The energy of this state is given by

$$\langle\psi|H|\psi\rangle = \sum_z c_z^2 \langle z|H_0|z\rangle - \sum_z \sum_y c_z c_y \langle z|V|y\rangle.$$

Looking at this formula we see that the energy of $|\psi\rangle$ cannot be greater than the energy of the state

$$|\tilde{\psi}\rangle = \sum_z |c_z|\, |z\rangle$$

where we have made all of the coefficients nonnegative. So in particular, any stoquastic Hamiltonian H has a ground state with nonnegative coefficients in the z basis.

Note that if there is a ground state with nonnegative coefficients in the given

32

basis then no excited state can have coefficients which are all strictly positive in this basis (since eigenstates with different eigenvalues are orthogonal). This proves the following lemma

Lemma 1. *Let H be a Hermitian operator, stoquastic in the $|z\rangle$ basis, that is*

$$\langle z|H|z'\rangle \leq 0$$

for $z \neq z'$. If $H|\psi\rangle = E|\psi\rangle$ and $\langle z|\psi\rangle > 0$ for all z, then $E = E_g$, the ground state energy of H.

This statement can be strengthened for stoquastic Hamiltonians which have a property called *irreducibility*. A stoquastic Hamiltonian matrix H is reducible if there exists a permutation matrix P such that

$$PHP^T$$

is block diagonal (in the given basis $\{|z\rangle\}$). Otherwise H is irreducible. The Perron-Frobenius theorem [46] tells us that

If H is stoquastic and irreducible in the basis $\{|z\rangle\}$ then:

1. The ground state of H is unique.

2. The ground state $|\psi\rangle$ can be taken to have strictly positive coefficients $\langle z|\psi\rangle > 0$.

1.2.2 A Variational Lower Bound For the Ground State Energy

For any normalized state $|\psi\rangle$, and any Hamiltonian H, the quantity $\langle\psi|H|\psi\rangle$ is an upper bound on the ground state energy. It is less well known that, for stoquastic Hamiltonians, one can also obtain a variational lower bound on the ground state energy of H. Theorem 1 (given below) is an example of such a bound. (It is an elementary application of the Collatz Wielandt min-max formula [46] of nonnegative matrices). A short proof of this theorem from [13] is reprinted here.

Theorem 1. *Let H be a Hermitian operator, stoquastic in the $|z\rangle$ basis, that is $\langle z|H|z'\rangle \leq 0$ for $z \neq z'$. Let E_g be its lowest eigenvalue. Then*

$$E_g \geq \min_z \frac{\langle z|H|\phi\rangle}{\langle z|\phi\rangle}$$

for any state $|\phi\rangle$ such that $\langle z|\phi\rangle > 0$ for all z.

Proof. Let $\hat{H}$ be

$$\hat{H} = H - \sum_z |z\rangle\langle z| \frac{\langle z|H|\phi\rangle}{\langle z|\phi\rangle}.$$

Note that $|\phi\rangle$ is an eigenvector of $\hat{H}$ with eigenvalue 0. The ground state energy of $\hat{H}$ is 0 by Lemma 1.

For any normalized state $|\psi\rangle$ we have

$$\langle\psi|H|\psi\rangle - \langle\psi|\hat{H}|\psi\rangle = \sum_z |\langle z|\psi\rangle|^2 \left[\frac{\langle z|H|\phi\rangle}{\langle z|\phi\rangle}\right]$$

$$\geq \min_z \left[\frac{\langle z|H|\phi\rangle}{\langle z|\phi\rangle}\right].$$

Choosing $|\psi\rangle$ to be a ground state of H we have

$$E_g - \langle\psi|\hat{H}|\psi\rangle \geq \min_z \left[\frac{\langle z|H|\phi\rangle}{\langle z|\phi\rangle}\right].$$

But $\langle\psi|\hat{H}|\psi\rangle \geq 0$ since $\hat{H}$ has ground state energy 0, so

$$E_g \geq \min_z \frac{\langle z|H|\phi\rangle}{\langle z|\phi\rangle}.$$

$\square$

Note that this variational lower bound is tight (i.e it is possible to achieve equality) for irreducible stoquastic Hamiltonians.

In Chapter 4 of this thesis we will use this lower bound as well as the standard variational upper bound on the ground state energy to characterize the ground state energy for scrambled Hamiltonians. For these problems we are able to solve for the ground state energy per spin $\frac{1}{n}E_g$ by showing that we can get variational upper and lower bounds for this quantity which agree in the limit $n \to \infty$.

1.2.3 The Continuous Imaginary Time Path Integral and a Probability Distribution over Paths

Specialized numerical techniques such as the quantum cavity method [43, 40] and path integral Quantum Monte Carlo [55, 60] rely on the fact that the thermal path integral for a stoquastic Hamiltonian has a special form. Here we review the thermal path integral (in the continuous imaginary time formulation [22, 55]) and discuss how stoquasticity is useful in this context.

Any stoquastic Hamiltonian can be written as

$$H = H_0 + \lambda V \tag{1.13}$$

where H_0 is diagonal in the basis $\{|z\rangle\}$, V is purely off diagonal in this basis, and $\lambda \geq 0$. The parameter λ can always be set it to 1 by rescaling V. However, sometimes we will be interested in computing properties of the Hamiltonian 1.13 as a function of λ with fixed H_0 and V.

The continuous time thermal path integral for the Hamiltonian $H = H_0 + \lambda V$ can be obtained using the Dyson series as

$$Tr\left[e^{-\beta H}\right] = Tr\left[e^{-\beta H_0}\sum_{m=0}^{\infty}(-\lambda)^m \int_{t_m=0}^{\beta}dt_m\int_{t_{m-1}=0}^{t_m}dt_{m-1}...\int_{t1_1=0}^{t_2}dt_1 V_I(t_m)...V_I(t_1)\right]$$

Here we use the notation

$$V_I(t) = e^{tH_0}Ve^{-tH_0}. \tag{1.14}$$

We can then insert complete sets of states and take the trace to obtain the path integral

$$Tr\left[e^{-\beta H}\right] =$$

$$\sum_{m=0}^{\infty}\sum_{\{z_1,...,z_m\}}\left[\langle z_1| - \lambda V|z_m\rangle\langle z_m| - \lambda V|z_{m-1}\rangle...\langle z_2| - \lambda V|z_1\rangle\right.$$

$$\left.\int_{t_m=0}^{\beta}dt_m\int_{t_{m-1}=0}^{t_m}dt_{m-1}...\int_{t_1=0}^{t_2}dt_1 e^{-\int_{t=0}^{\beta}\mathcal{H}_0(z(t))dt}\right]. \tag{1.15}$$

In this formula we have used the notation $\mathcal{H}_0(z(t)) = \langle z(t)|H_0|z(t)\rangle$, where the function $z(t)$ is defined by

$$z(t) = \begin{cases} z_1, & 0 \leq t < t_1 \\ z_2, & t_1 \leq t < t_2 \\ \vdots \\ z_m, & t_{m-1} \leq t < t_m \\ z_1, & t_m \leq t \leq \beta. \end{cases}$$

So in particular

$$\int_{t=0}^{\beta}\mathcal{H}_0(z(t))dt = \langle z_1|H_0|z_1\rangle(t_1+\beta-t_m)+\langle z_2|H_0|z_2\rangle(t_2-t_1)+...\langle z_m|H_0|z_m\rangle(t_m-t_{m-1}).$$

We view the function $z(t)$ as a path in imaginary time which begins at $t = 0$ and ends at $t = \beta$. Then equation 1.15 is a sum over paths, where every path P is assigned a weight according to $\tilde{\rho}$

$$\tilde{\rho}(P) = \lambda^m\langle z_1| - V|z_m\rangle\langle z_m| - V|z_{m-1}\rangle...\langle z_2| - V|z_1\rangle dt_1...dt_m e^{-\int_{t=0}^{\beta}\mathcal{H}_0(z(t))dt}. \tag{1.16}$$

Note that because our Hamiltonian is stoquastic, $\langle z|V|z'\rangle \leq 0$ and the weight function $\tilde{\rho}(P)$ from equation 1.16 is nonnegative. Summed over all paths it gives the partition function $Z(\beta) = Tr[e^{-\beta H}]$. So we can define a probability distribution ρ

over paths by

$$\rho(P) = \frac{\tilde{\rho}(P)}{Z(\beta)}. \qquad (1.17)$$

When the Hamiltonian is not stoquastic one can define a normalized weight function but it will take both positive and negative values. (In Quantum Monte Carlo parlance, the fact that ρ can be negative for non stoquastic Hamiltonians is called the "sign problem".)

Physical properties of the quantum system at inverse temperature β can be expressed as expectation values with respect to the probability distribution ρ. (If β is sufficiently large then these expectation values will agree with the corresponding expectation values in the ground state.) It is straightforward to show that for any operator D which is diagonal in the $|z\rangle$ basis, the thermal expectation value can be written as

$$\frac{Tr[De^{-\beta H}]}{Tr[e^{-\beta H}]} = \langle \frac{1}{\beta} \int_0^\beta \mathcal{D}(z(t))dt \rangle_\rho \qquad (1.18)$$

where $\mathcal{D}(z(t)) = \langle z(t)|D|z(t)\rangle$. The notation $\langle \rangle_\rho$ means average with respect to the classical probability distribution ρ over paths.

This immediately allows us to express quantities such as the diagonal part of the Hamiltonian

$$\frac{Tr[H_0 e^{-\beta H}]}{Tr[e^{-\beta H}]} = \langle \frac{1}{\beta} \int_0^\beta \mathcal{H}_0(z(t))dt \rangle_\rho. \qquad (1.19)$$

The thermal expectation value of the full Hamiltonian $H_0 + \lambda V$ can be written using equation 1.19 as well as the expression

$$\frac{Tr[\lambda V e^{-\beta H}]}{Tr[e^{-\beta H}]} = -\langle \frac{m}{\beta} \rangle_\rho \qquad (1.20)$$

where m is the number of transitions in the path. (The derivation of equations 1.18 and 1.20 are given in Appendix A) .

1.2.4 Quantum Monte Carlo and The Quantum Cavity Method

We have seen that by generating paths from the distribution ρ and then computing averages with respect to ρ, we can evaluate the thermal expectation value of the energy at a given inverse temperature β. Generating paths according to the probability distribution ρ is not easy! Continuous imaginary time Quantum Monte Carlo methods (such as the method described in [40]) work by defining a Markov Chain which converges to the limiting distribution ρ over paths. By running many iterations of this Markov Chain one obtains a path which is approximately distributed according to ρ. Then one can use the estimators from equations 1.18 and 1.20 to compute thermal expectations values for the quantum system.

The quantum cavity method[43, 40] is another probabilistic simulation technique

which makes use of the distribution ρ over paths. It is the quantum analog of the classical cavity method which is used to study the Gibbs distribution in statistical physics. In the quantum case the Gibbs distribution is replaced with the distribution ρ over paths. We discuss the quantum cavity method in greater detail in appendix E and we use this method in Chapter 5.

In this thesis we make use of the numerical schemes for Quantum Monte Carlo and quantum cavity simulations which are due to Krzakala et al. [40]. We now briefly mention a modification of their scheme which we have developed and which we use in our simulations. Their Quantum Monte Carlo method uses as a subroutine an algorithm for sampling from the path integral for a single spin. Using this subroutine they are able to sample from the distribution ρ corresponding to the path integral of n spins. They also describe an implementation of the quantum cavity method which uses this single spin sampling algorithm. In appendix B we describe an alternative algorithm for sampling from the single spin path integral. In this thesis we will use the methods of Krzakala et al for the Quantum Monte Carlo simulations in Chapter 3 and the quantum cavity method in Chapter 5. The only difference is that we use our algorithm for sampling from the single spin path integral in place of the subroutine described in [40].

1.3   The Next Four Chapters

We are now ready to describe where this thesis fits in.

A Quantum Monte Carlo Method at Fixed Energy

We first take a short detour in Chapter 2 where we present a novel Quantum Monte Carlo method which is very different from the continuous imaginary time scheme discussed above. As we have already mentioned, Quantum Monte Carlo methods are often the best method for studying stoquastic Hamiltonians with large system sizes. Although this Chapter may seem tangential to our theme (the quantum adiabatic algorithm), the theory behind this method motivates a variational ansatz which is used in Chapter 4 to prove results about scrambled Hamiltonians.

Each subsequent Chapter relates in some way to the story we recounted in section 1.1 about the performance of quantum adiabatic algorithms.

Perturbative Crosses

In Chapter 3 we return to the subject of perturbative crosses. We introduce an ensemble of random instances of 3SAT which have these crosses and for which the quantum adiabatic algorithm will fail. We argue that, by randomly changing the beginning Hamiltonian in a prescribed way, the failure of the quantum adiabatic algorithm on these instances can be overcome. This follows the "path change" strategy originally outlined in [19] (although the way that we randomly change the path differs from the proposals in that paper). We confirm our results using Quantum Monte Carlo simulations on systems as large as $n = 150$ spins.

Scrambled Hamiltonians

Chapter 4 is concerned with the scrambled Hamiltonians first discussed in reference [16]. Farhi et al showed using information theoretic techniques that a quantum algorithm cannot succeed on typical instances of a scrambled problem[16]. We use this information-theoretic result to prove that in fact the gap is exponentially small with high probability over the choice of scrambled instance. We also use variational methods to solve for the ground state energy per spin in the limit $n \to \infty$. Our proof works for a large class of scrambled Hamiltonians including the scrambled Hamming weight example discussed above.

Phase Transitions in 3XORSAT and Max-Cut

Finally in Chapter 5 we discuss the quantum phase transitions which occur in two different ensembles of problems. The first is random 3-Regular 3XORSAT with a unique satisfying assignment. The results of Jörg et al suggest a first order phase transition in the ground state at $s \approx \frac{1}{2}$ for this model. We present a duality transform which shows that the average energy per spin for this problem is symmetric about $s = \frac{1}{2}$. This locates the phase transition (if it exists and is unique) at $s = \frac{1}{2}$.

The second problem we discuss is called Max-Cut which is associated with an antiferromagnetic 2 local Hamiltonian on a graph. We consider the ensemble of random 3 regular instances which we study using the quantum cavity method. Using this method we see a second order phase transition at inverse temperature $\beta = 4$.

1.4   Outlook

Where does this leave us?

We have seen examples where the quantum adiabatic algorithm fails and some examples of easy problems where it succeeds. In this thesis we will not answer the question of whether or not quantum adiabatic optimization is more or less powerful than classical computation. We hope our example based approach sheds some light on which problems are hard (and which might be easy) for quantum adiabatic optimization. A completely different approach is to study adiabatic algorithms using complexity theory. We conclude this Chapter by taking a step back and reviewing some proven results about broad classes of adiabatic algorithms.

Variations on the quantum adiabatic algorithm have been considered in the literature. Consider the following extension where $H(s)$ is allowed to be any local Hamiltonian and in particular is not restricted to the transverse field spin Hamiltonians of the form 1.4. Initialize a quantum system of n spins in a product state which is the ground state of a local Hamiltonian $H(0)$. Adiabatically change the Hamiltonian along some prescribed path $H(s)$ until $s = 1$ at which point you have prepared the ground state of $H(1)$. Then make local measurements on the final state. In this model one can solve a wide range of problems—not just optimization problems which are specified by a classical cost function. In fact, it has been shown that this model

of adiabatic quantum computation is as powerful as a full blown quantum computer [1].

On the other hand Bravyi and Terhal have shown that quantum adiabatic evolution in the ground state of a local, stoquastic, and *frustration free* Hamiltonian $H(s)$ can be simulated efficiently with a probabilistic classical computer [11]. (A frustration free local Hamiltonian is of the form $H = \sum_i H_i$ where each local term $H_i \geq 0$ and the ground state $|\psi\rangle$ satisfies $H_i|\psi\rangle = 0$ for each i.)

As pointed out by Bravyi and Terhal[11] the computational power of adiabatic evolution in the ground state of a stoquastic local Hamiltonian is unknown. The quantum adiabatic optimization algorithms discussed in this thesis are a subset of this class of quantum algorithms. We have seen in the previous section how stoquasticity allows us to use an expanded set of methods for studying quantum systems. But we should keep in mind that stoquasticity may be a double edged sword–it may also restrict the computational power of the quantum adiabatic algorithm. Does there exist an efficient technique for classically simulating stoquastic adiabatic evolution? If so then we cannot hope for a superpolynomial speedup through quantum adiabatic optimization. Such a classical simulation technique (if it exists) would likely be practically useful for condensed matter physicists studying quantum systems which arise in other contexts.

Since no efficient classical simulation algorithm for stoquastic adiabatic evolution exists to date, it makes sense to look for examples where quantum adiabatic optimization is efficient but classical optimization algorithms are not. We adopt this optimistic approach. The question of whether or not an exponential speedup can be obtained with quantum adiabatic optimization remains open.

Chapter 2

A Quantum Monte Carlo Method at Fixed Energy

In this Chapter we describe a new Quantum Monte Carlo method that allows one to numerically investigate ground state properties of a quantum system. Most of this Chapter is excerpted from [14] (which is joint work with Edward Farhi, Jeffrey Goldstone and Harvey Meyer).

2.1 Introduction

A virtue of Quantum Monte Carlo is that one is not required to manipulate vectors in the Hilbert space corresponding to the quantum system. The dimension of this Hilbert space typically grows exponentially with the physical size of the system. Instead, Quantum Monte Carlo methods map the problem of approximating the ground state energy (or some other observable) onto the problem of evaluating an expectation value with respect to a probability distribution $q(X)$ over a set of configurations C (so $X \in C$). In order to evaluate this expectation value, one can use a classical Markov chain Monte Carlo algorithm to sample configurations from the distribution q. Markov chain Monte Carlo works by defining a Markov chain on the space of configurations C. This Markov chain can be described by an update rule which tells you how to generate a new configuration of the chain from the current one. The Markov chain is constructed so that the limiting distribution is $q(X)$. One then applies some large number N_0 of iterations of the Markov chain to some initial configuration X_0. If N_0 is sufficiently large then after these iterations, the distribution of subsequent configurations will be arbitrarily close to q.

There are two key components of a quantum monte carlo scheme. The first component is the map between the quantum problem of interest and a classical sampling problem. One has to first show that, in theory, there is an ensemble of configurations $q(X)$ over a state space C such that certain ensemble averages with respect to q give properties of the quantum system. The second component is practical. One must show that, in practice, there is a way to calculate these estimators by defining a Markov Chain over the state space C which converges to the limiting distribution

41

q.

In section 1.2.3 we described the "in theory" component of continuous imaginary time path integral Quantum Monte Carlo[55]. In that method, the probability distribution ρ over paths is the object of interest and we have seen that by computing ensemble averages with respect to ρ we can evaluate properties of the quantum system at an inverse temperature β, such as the expectation value of the Hamiltonian H.

In this Chapter we present a new Quantum Monte Carlo method. Like most Quantum Monte Carlo methods, ours is applicable only to stoquastic Hamiltonians. In keeping with our earlier notation we write

$$H(\lambda) = H_0 + \lambda V \tag{2.1}$$

where H_0 is diagonal in the basis $\{|z\rangle\}$, V is purely off diagonal in this basis, and $\lambda \geq 0$.

Suppose we are interested in computing the ground state energy curve $E_g(\lambda)$ for the Hamiltonian (2.1) at a set of points $\lambda_0, \lambda_1, \lambda_2, ..., \lambda_R$. In order to compute $E_g(\lambda)$ by standard Quantum Monte Carlo methods such as the path integral method, one first fixes a value λ_0 as well as a large inverse temperature β. The Quantum Monte Carlo procedure then outputs the thermal expectation value of $H(\lambda_0)$ at inverse temperature β, which, if β is large, will approximate $E_g(\lambda_0)$. One then repeats this procedure for some list of values $\lambda_1, \lambda_2, ..., \lambda_R$. The Quantum Monte Carlo we describe in section 2.3 can also be used to compute points on the curve $E_g(\lambda)$. We will see that the curve $E_g(\lambda)$ for $\lambda > 0$ is in fact invertible. In order to use the new Quantum Monte Carlo method, one fixes an energy E and computes the value of the inverse function $\lambda(E)$. One can then repeat this procedure at different values of E to obtain a picture of the curve.

To motivate our new method and to get a general idea of how it works, consider the function

$$G(E, \lambda) = Tr\left[\left(\frac{-\lambda}{H(\lambda) - E}\right)V\right]. \tag{2.2}$$

Assuming that $E < 0$ and $\lambda > 0$ are chosen so that the Taylor series expansion converges, we can write

$$
\begin{aligned}
G(E, \lambda) &= Tr\left[\left(\frac{-\lambda}{1 + \frac{\lambda}{H_0 - E}V}\right)\frac{1}{H_0 - E}V\right] \\
&= \sum_{m=1}^{\infty} Tr\left[\left(\frac{-\lambda}{H_0 - E}V\right)^m\right].
\end{aligned}
\tag{2.3}
$$

It is clear from the expression in equation 2.2 that the function $G(E, \lambda)$ blows up when $E \to E_g(\lambda)$, where $E_g(\lambda)$ is the ground state energy of $H(\lambda)$. Equivalently we can say that at a fixed value of E the blow up occurs as $\lambda \to \lambda(E)$, where $E_g(\lambda(E)) = E$. At this value of λ the Taylor series expansion must diverge. In fact, this divergence occurs because as m becomes large, terms in the series approach 1 for large m (here

42

we have made some assumptions about the Hamiltonian which we discuss in the next section) so

$$Tr\left[\left(\frac{-\lambda(E)}{H_0 - E}V\right)^m\right] \approx 1 \, .$$

For the remainder of this section we assume that m is large enough to make $\approx$ close to $=$. By inserting complete sets of states in the basis that diagonalizes H_0 we can express the LHS as a sum over paths

$$(\lambda(E))^m \sum_{\{z_1,\ldots,z_m\}} \langle z_1| - V|z_m\rangle\langle z_m| - V|z_{m-1}\rangle\ldots\langle z_2| - V|z_1\rangle \prod_{i=1}^{m} \frac{1}{E_i - E} \approx 1$$

where $E_i = \langle z_i|H_0|z_i\rangle$. Now taking the log and differentiating with respect to E, we obtain

$$
\begin{aligned}
-\frac{1}{\lambda(E)}\frac{d\lambda(E)}{dE} &\approx \sum_{\{z_1,\ldots,z_m\}} f\left(\{z_1,\ldots,z_m\}\right)\left(\frac{1}{m}\sum_{i=1}^{m}\frac{1}{E_i - E}\right) \\
&= \langle\frac{1}{m}\sum_{i=1}^{m}\frac{1}{E_i - E}\rangle_f \, .
\end{aligned}
\tag{2.4}
$$

Here the expectation value is taken with respect to the measure f on paths defined by

$$f\left(\{z_1,\ldots,z_m\}\right) = \frac{1}{F}\langle z_1| - V|z_m\rangle\langle z_m| - V|z_{m-1}\rangle\ldots\langle z_2| - V|z_1\rangle \prod_{i=1}^{m}\frac{1}{E_i - E}$$

where F is a normalizing constant.

Sampling from the distribution f will also allow us to compute $-\frac{1}{\lambda(E)}\frac{d\lambda(E)}{dE}$ from equation 2.4 as well as other properties of the ground state. We will show how to sample with respect to the distribution f in a way that makes numerical work possible.

The rest of this Chapter is organized as follows. In section 2.2 we describe the Hamiltonians for which our method will work. In section 2.3 we fully describe the "in theory" component of our scheme which consists of the ensemble f over paths and estimators for ground state properties.

In section 2.4 we turn to the practical issue of sampling. We restrict to the case of transverse field spin Hamiltonians where the system is composed of n qubits,

$$V = -\sum_{i=1}^{n}\sigma_x^i,$$

and H_0 is diagonal in the Pauli z basis. We describe how one can sample configurations from the appropriate distribution for our new Monte Carlo method and we demonstrate numerically that our method works on a small (16 spin) instance of Exact Cover.

Finally, in section 2.5.1 we describe two connections between our quantum monte

carlo method and other topics considered in this thesis. We first demonstrate a connection between the $\beta \to \infty$ limit of the path integral distribution ρ and the $m \to \infty$ limit of the ensemble that appears in our method. We obtain a new estimator for the ground state energy in the standard path integral approach which is based on this connection. We then describe a variational ansatz for the ground state of a stoquastic Hamiltonian which we will return to in Chapter 4, where we use this ansatz in the lower bound from theorem 1 of Chapter 1.

2.2 The Hamiltonian

We now discuss the properties of the Hamiltonian $H(\lambda)$ which we require for our method to work. We require:

1. The ground state of $H(\lambda)$ is not degenerate for any value of $\lambda \in (-\infty, \infty)$.

2. The smallest eigenvalue of H_0 is zero. Note that this condition can be fulfilled without loss of generality by adding a constant term to the Hamiltonian. Writing $|z_0\rangle \in \{|z\rangle\}$ for the unique state with $H_0|z_0\rangle = 0$, we also require that $V|z_0\rangle \neq 0$. (This implies that $|z_0\rangle$ is not an eigenvector of V since $\langle z_0|V|z_0\rangle = 0$.)

We write $|\psi_g(\lambda)\rangle$ and $E_g(\lambda)$ for the ground state eigenvector and ground state energy of $H(\lambda)$, and write $|\psi_i(\lambda)\rangle$ and $E_i(\lambda)$ for the excited state eigenvectors and energies with $i \in \{1, ...N\}$. From second order perturbation theory in λ, we have that

$$\frac{d^2 E_g}{d\lambda^2} = -2 \sum_{i=1}^{N} \frac{|\langle \psi_i(\lambda)|V|\psi_g(\lambda)\rangle|^2}{E_i - E_g}$$
$$< 0. \tag{2.5}$$

Note that the inequality is strict for all λ. This holds because if it were not strict then $|\psi_g(\lambda_0)\rangle$ would be an eigenstate of V for some λ_0, which implies that $|\psi_g(\lambda)\rangle$ must be an eigenstate of V for all λ which is forbidden since $V|z_0\rangle \neq 0$.

Using the fact that

$$\frac{dE_g}{d\lambda} = \langle \psi_g(\lambda)|V|\psi_g(\lambda)\rangle \tag{2.6}$$

we show that

$$\frac{dE_g}{d\lambda} = \begin{cases} > 0 \text{ , for } \lambda < 0 \\ = 0 \text{ , for } \lambda = 0 \\ < 0 \text{ , for } \lambda > 0 . \end{cases}$$

In order to obtain the inequalities, we use the variational principle. When $\lambda > 0$, the ground state energy must be less than zero, since $|z_0\rangle$ has zero expectation value for H (and $|z_0\rangle$ is not an eigenvector of $H(\lambda)$). This, together with the fact that H_0 is positive semidefinite, implies that $\langle \psi_g|V|\psi_g\rangle < 0$. The analogous result for $\lambda < 0$ is

44
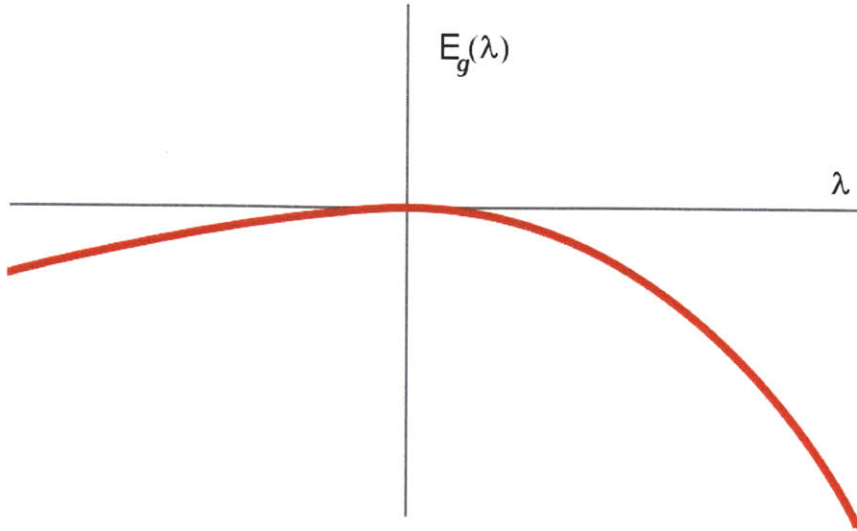
Figure 2-1: $E_g(\lambda)$ for the Hamiltonians we consider. As $\lambda \to \pm\infty$ we have $E_g \to -\infty$.

obtained in the same way. These inequalities give a qualitative picture of the curve $E_g(\lambda)$. Starting from $E_g(0) = 0$, the curve slopes downwards as it goes out from $\lambda=0$, and approaches $-\infty$ on both sides of the origin for sufficiently large $|\lambda|$. Note that this implies that for each $E < 0$ there is one positive and one negative value of λ (call them $\lambda(E)$ and $\lambda_-(E)$ respectively) such that $E_g(\lambda(E)) = E$ and $E_g(\lambda_-(E)) = E$. Furthermore, we show in section C that it is always the case that

$$\lambda(E) \leq |\lambda_-(E)|. \tag{2.7}$$

We refer to the case where the inequality is strict as the generic case. We illustrate the qualitative features of the curve $E_g(\lambda)$ (for the generic case) in Figure 2.2.

In the nongeneric case where equality holds at some particular value of E, then in fact equality holds at every value of E and the curve $E_g(\lambda)$ is symmetric about $\lambda = 0$.

2.3   Description of the method

As motivated in the Introduction, we now define an ensemble where the configurations are sequences $\{z_1, ..., z_m\}$, and we show how properties of the ground state can be computed in this ensemble. We refer to the sequences $\{z_1, ..., z_m\}$ as paths.

To begin, we fix $E < 0$ and a large integer m as parameters. As in the previous section, we take $\lambda(E)$ to be the positive value of λ such that $H(\lambda)$ has ground state energy E, with corresponding eigenvector $|\psi_g(\lambda(E))\rangle$. We now describe how our method allows us to approximate $\lambda(E)$ and other properties of the ground state.

Recall from the Introduction that the probability distribution f over paths is

defined by

$$f(z_1, ..., z_m) \equiv \frac{1}{F(E,m)} \langle z_1| - V|z_m\rangle ... \langle z_2| - V|z_1\rangle \prod_{i=1}^{m} \frac{1}{E_i - E} \qquad (2.8)$$

where $E_i = \langle z_i|H_0|z_i\rangle$ and [1]

$$F(E,m) \equiv \sum_{\{z_1,...,z_m\}} \langle z_1| - V|z_m\rangle ... \langle z_2| - V|z_1\rangle \prod_{i=1}^{m} \frac{1}{E_i - E} \qquad (2.9)$$

$$= Tr\left[(\frac{-1}{H_0 - E}V)^m\right]. \qquad (2.10)$$

As examples, we now define two quantities $\frac{\bar{\beta}(E,m)}{m}$ and $\overline{\lambda^2}(E,m)$ as ensemble averages with respect to the distribution f on paths

$$\bar{\beta}(E,m) \equiv \sum_{\{z_1,...,z_m\}} f(\{z_1, ..., z_m\})\beta_{est}(\{z_1, ..., z_m\}) = \langle \beta_{est}\rangle_f$$

$$\overline{\lambda^2}(E,m) \equiv \sum_{\{z_1,...,z_m\}} f(\{z_1, ..., z_m\})\lambda^2_{est}(\{z_1, ..., z_m\}) = \langle \lambda^2_{est}\rangle_f \qquad (2.11)$$

where we have defined the estimators (hence the subscript)

$$\beta_{est}(\{z_1, ..., z_m\}) \equiv \sum_{i=1}^{m} \frac{1}{E_i - E} \qquad (2.12)$$

$$\lambda^2_{est}(\{z_1, ..., z_m\}) \equiv \frac{1}{m}\sum_{i=1}^{m} \delta_{z_{i+2}z_i}(E_{i+1} - E)(E_i - E)\frac{1}{\langle z_i|V^2|z_i\rangle} \qquad (2.13)$$

with $z_{m+1} = z_1$, $z_{m+2} = z_2$. Our reason for using the symbol β_{est} will become clear in section 2.5.1 where we will discuss its interpretation as an inverse temperature. We show in section C that the ensemble averages $\frac{\bar{\beta}(E,m)}{m}$ and $\overline{\lambda^2}(E,m)$ correspond to properties of the quantum ground state in the limit $m \rightarrow \infty$

$$\lim_{m\rightarrow\infty} \frac{\bar{\beta}(E,m)}{m} = -\frac{1}{\lambda(E)}\frac{d\lambda(E)}{dE} \qquad (2.14)$$

$$= -\frac{1}{\lambda(E)}\frac{1}{\langle \psi_g(E)|V|\psi_g(E)\rangle} \qquad (2.15)$$

$$\lim_{m\rightarrow\infty} \overline{\lambda^2}(E,m) = (\lambda(E))^2. \qquad (2.16)$$

Equation 2.15 follows from 2.6. One can also derive expressions for higher derivatives of $\log(\lambda(E))$ as averages with respect to f.

---

[1]The nongeneric case where equality holds in equation 2.7 can arise when there exists a unitary transformation U such that $U^\dagger VU = -V$ and $U^\dagger H_0 U = H_0$. In this case it is seen from equation 2.10 that $F(E,m) = 0$ when m is odd. In the nongeneric case m must always be taken to be even.

We propose to use the measure f as the basis for Monte Carlo simulations. In particular, our Monte Carlo algorithm begins by choosing $E < 0$ and a large integer m and then samples sequences of bit strings from the distribution f. One can then use the estimators from equations 2.11 and 2.13 to evaluate the quantities $\lambda(E)$ and $\langle \psi_g(\lambda(E)) | V | \psi_g(\lambda(E)) \rangle$, using equations 2.15 and 2.16 for the limiting behaviour of these estimators. We do not construct a general method for sampling from f, instead we leave it to the reader to construct such a method for the particular choice of V at hand. We do note that it is essential that whatever method is used conserves the total number m of transitions in the path. In section 2.4 we give an example of this for a generic spin system with $V = -\sum_{i=1}^{n} \sigma_x^i$.

2.4 Sampling Paths for Transverse Field Spin Hamiltonians

In this section we explicitly construct a method to compute averages with respect to the distribution f in the case where the Hilbert space is that of n spin $\frac{1}{2}$ particles, and $V = -\sum_{i=1}^{n} \sigma_x^i$. The Hamiltonian H_0 is an arbitrary diagonal matrix in the Pauli z basis for n spins. Our algorithm can be used to compute the average of any function of the path which is invariant under cyclic permutations of the path. For this choice of V, the spectrum of $H(\lambda)$ is symmetric about $\lambda = 0$ (so this corresponds to the nongeneric case where equality holds in equation 2.7 for all $E < 0$). With these choices, the paths which have nonzero weight (with respect to f) are periodic paths of m bit strings of length n where each string z differs from the previous one by a bit flip. We must take m to be even since each bit must flip an even number of times so that the path is periodic (and so the total number of bit flips m in the path must be even). In order to sample from these paths according to f, we construct a Markov Chain which has f as its limiting distribution (actually, our Markov Chain converges to the correct distribution over equivalence classes of paths which are only defined up to cyclic permutation–this is why we restrict ourselves to estimating quantities which are cyclically invariant). Note that with our choice of V we can write (see equation 2.8)

$$f(\{z_1, ..., z_m\}) = \frac{1}{F(E, m)} \prod_{i=1}^{m} \frac{1}{E_i - E} \, .$$

Our Markov chain is defined by the following update rule which describes how the configuration is changed at each step

1. Choose an integer $i \in \{1, ..., m\}$ uniformly at random.

2. Consider the bit strings z_{i-1}, z_i, z_{i+1} in the current path (where $z_{m+1} = z_1$). Suppose that z_{i-1} and z_i differ in bit $q_1 \in \{1, ..., n\}$, which we write as $z_i = z_{i-1} \oplus \hat{e}_{q_1}$. Also write $q_2 \in \{1, ..., n\}$ for the bit in which z_i and z_{i+1} differ, so $z_{i+1} = z_i \oplus \hat{e}_{q_2}$.

3. If $q_1 \neq q_2$ then propose to change the bit string z_i to the new value $\tilde{z}_i = z_{i-1} \oplus \hat{e}_{q_2}$.
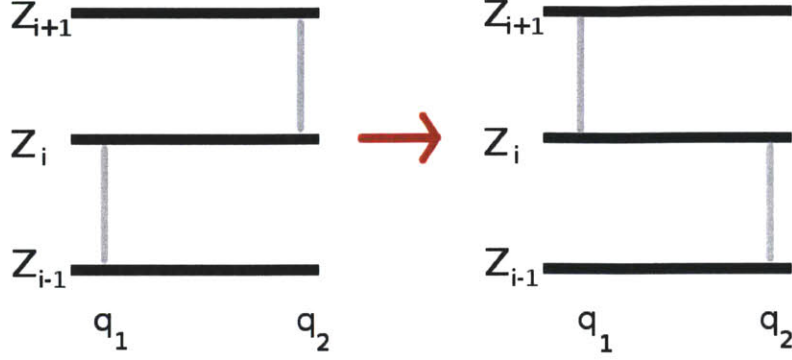
Figure 2-2: Monte Carlo update where the order of 2 flips in the path is interchanged.

Accept this proposal with probability

$$P_{accept} = \min\left\{1, \frac{E_i - E}{\tilde{E}_i - E}\right\}$$

where $\tilde{E}_i = \langle \tilde{z}_i | H_0 | \tilde{z}_i \rangle$. This Monte Carlo move has the effect of interchanging 2 consecutive flips in the path (see Figure 2-2).

4. If $q_1 = q_2$, then choose a new bit $q_{new} \in \{1, ..., n\}$ from the probability distribution

$$P(q_{new} = q) = \frac{1}{W} \frac{1}{E'_q - E} \tag{2.17}$$

where $E'_q = \langle z_{i-1} \oplus \hat{e}_q | H_0 | z_{i-1} \oplus \hat{e}_q \rangle$, and $W = \sum_{j=1}^{n} \frac{1}{E'_j - E}$. Then (with probability 1) change z_i to the new value $z_{i-1} \oplus \hat{e}_{q_{new}}$. This Monte Carlo move replaces a pair of consecutive flips which occur in the same bit with 2 new flips in a possibly different bit (see Figure 2-3).

We show in appendix C that this algorithm can be used to estimate any quantity which is invariant under cyclic permutations of the path (note that all estimators we have discussed have this property).

Numerical Simulation with a Particular Choice of H_0

We have numerically tested our Monte Carlo algorithm using a C++ computer program. In this section we show numerical data at 16 bits where we are able to compare results with exact numerical diagonalization. We studied the Hamiltonian with V as in the previous section, and H_0 corresponding to a problem Hamiltonian for the
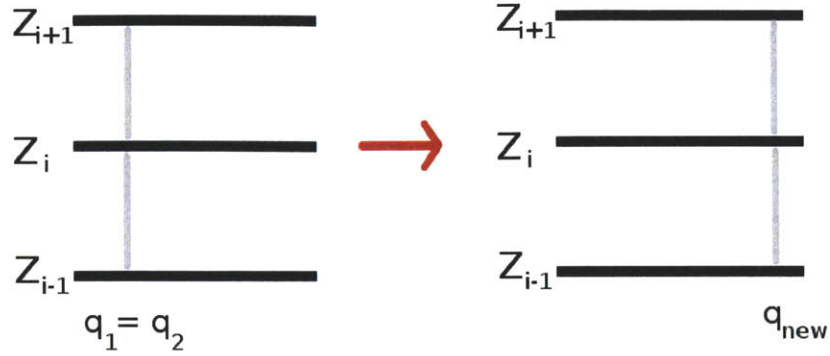
Figure 2-3: Monte Carlo update where 2 adjacent flips in the path which occur in the same bit are replaced by flips in a different bit.

combinatorial optimization problem Exact Cover [20]

$$H(\lambda) = H_0 - \lambda \sum_{i=1}^{n} \sigma_x^i$$

where

$$H_0 = \sum_{c=1}^{N_c} \left(\frac{1 - \sigma_z^{i_1(c)} - \sigma_z^{i_2(c)} - \sigma_z^{i_3(c)}}{2} \right)^2.$$

Here H_0 is a sum over N_c clauses.

We generated an instance of Exact Cover on 16 bits with a unique satisfying assignment through a random procedure. Figures 2-4 and 2-5 show the values of $\lambda(E)$ and $-\lambda(E)dE(\lambda)/d\lambda$ computed using equation 2.11 for 200 values of E, with $m = 1000$. Statistical errors were computed using Ulli Wolff's error analysis program [62]. This data set was taken by running 10^8 Monte Carlo updates on each of two processors of a dual core laptop computer for each value of E. The two processors ran simultaneously and the total time taken for all the data was under 5 hours. We also include the curves for these quantities obtained by exact diagonalization, which are in good agreement with the Monte Carlo data. Since it is hard to see the error bars in Figures 2-4 and 2-5, we have plotted the errors separately in Figures 2-6 and 2-7.

49
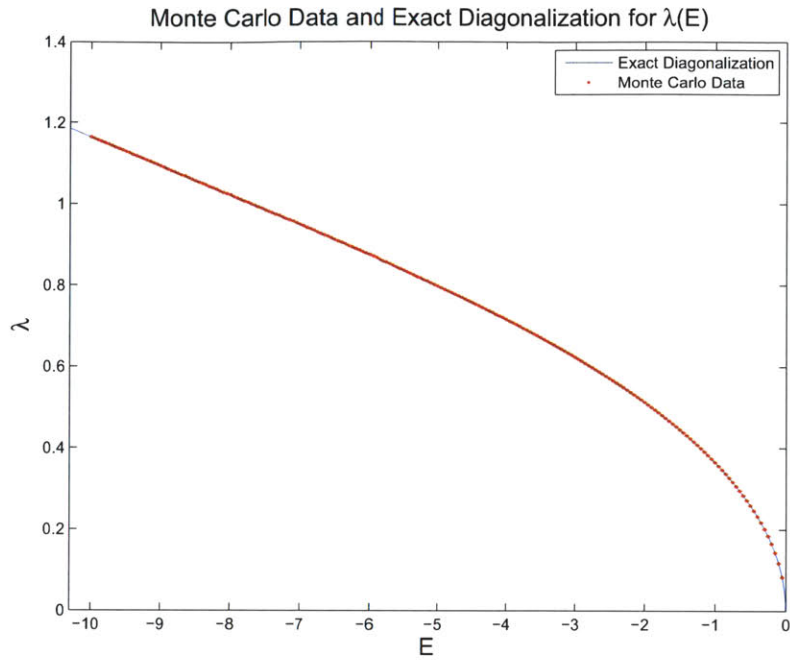
Figure 2-4: $\lambda(E)$ computed using Monte Carlo data and exact diagonalization for a 16 spin Hamiltonian. Statistical error bars are included for the Monte Carlo results, but they are barely visible. We have also plotted the errors separately in Figure 2-6.
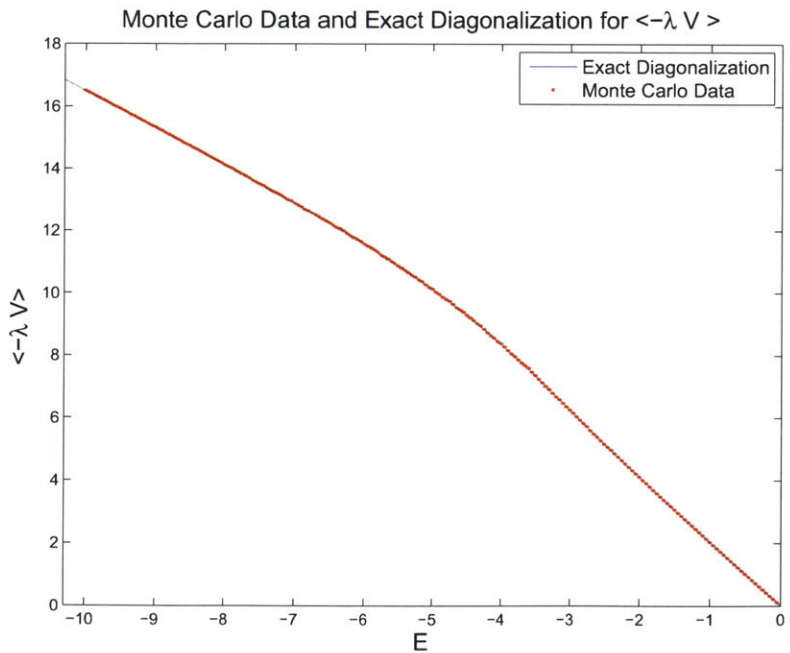
Figure 2-5: $-\lambda(E)\langle\psi_g(\lambda(E))|V|\psi_g(\lambda(E))\rangle$ computed using Monte Carlo data and exact diagonalization for a 16 spin Hamiltonian. Statistical error bars are included for the Monte Carlo results, but they are barely visible. Errors are also plotted separately in Figure 2-7.
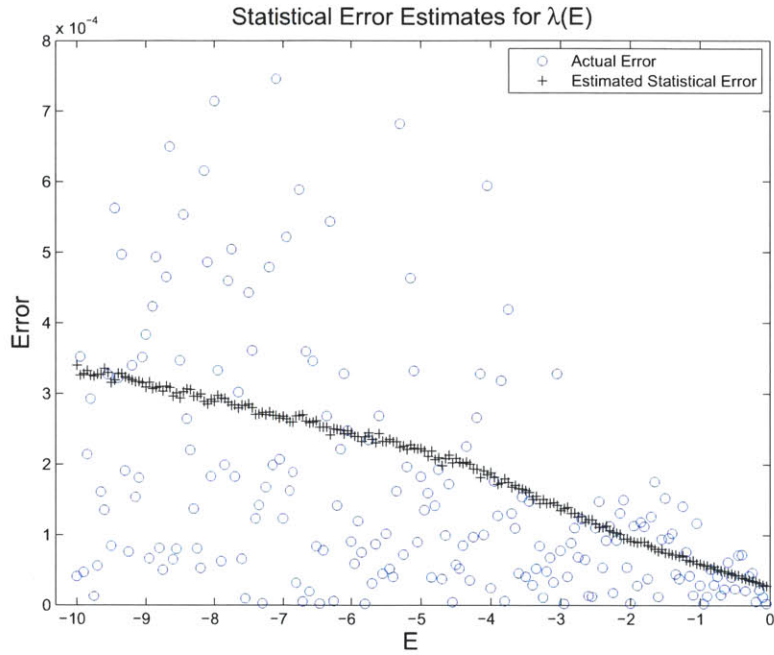
Figure 2-6: $\lambda(E)$ from Figure 2-4. The black crosses show the estimated statistical error. The blue circles show the magnitude of the difference between the Monte Carlo estimates and the result of exact numerical diagonalization.
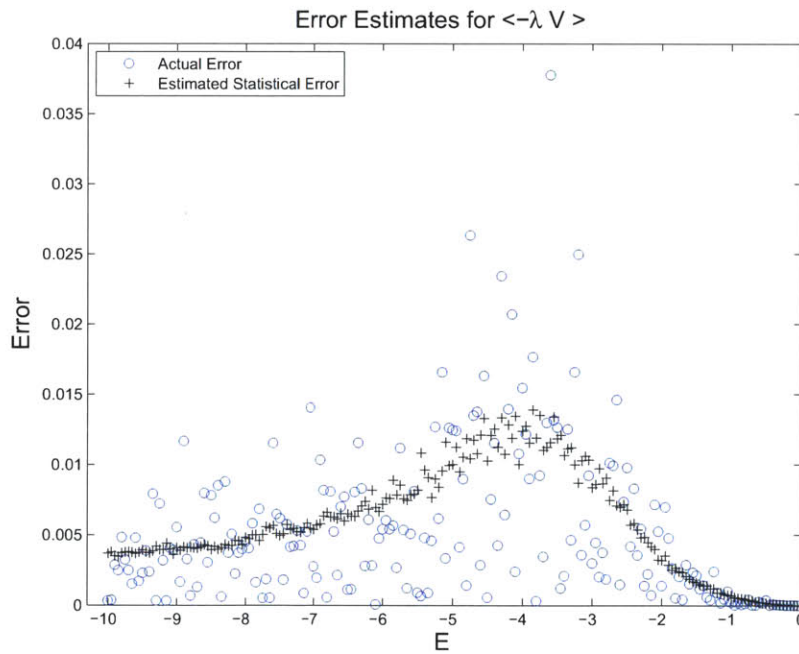
Figure 2-7: $-\lambda(E)\langle\psi_g(\lambda(E))|V|\psi_g(\lambda(E))\rangle$ from Figure 2-5. The black crosses show the estimated statistical error. The blue circles show the magnitude of the difference between the Monte Carlo estimates and the result of exact numerical diagonalization.

51

2.5 Two Observations

In this section we discuss two results which relate our new monte carlo method to other topics discussed in this thesis. Firstly, in section 2.5.1 we show how to obtain a new estimator for the ground state energy in the path integral quantum monte carlo method discussed in Chapter 1. We will see that the form of this estimator establishes a connection between the two quantum monte carlo methods.

In section 2.5.2 we give a preview of a variational ansatz used in Chapter 4 to solve for the ground state energy of scrambled Hamiltonians. We will see how this variational ansatz for the ground state is inspired by our Quantum Monte Carlo method.

2.5.1 A New Estimator for the Ground State Energy for Path Integral Quantum Monte Carlo

We now derive a novel estimator for the energy in the path integral ensemble described by the distribution ρ from equation 1.17 over paths in imaginary time. Our estimator, which is useful in the limit $\beta \to \infty$, is given by

$$\langle H \rangle = \langle E^\star \rangle_\rho + O\left(\frac{1}{\beta}\right)$$

where $E^\star$ is a function of the path defined to be the smallest value of E which satisfies the equation

$$\beta = \sum_{i=1}^{m+1} \frac{1}{E_i - E} \, . \tag{2.18}$$

In this expression the E_i are the energies of the states $|z_i\rangle$ visited along the path, with $E_{m+1} = E_1$. Note that the above equation is almost identical to equation 2.12 (this justifies our choice of notation β_{est}). We obtain this formula by a similar method to that used in reference [6] to obtain an alternate estimator for the energy $\langle H \rangle$. Our formula, however, is valid when some or all of the $\{E_i\}$ are the same, and therefore resolves a serious difficulty encountered in reference [6]. This may be of use in large β Monte Carlo simulations, as an alternative to the standard estimator for $\langle H \rangle$. We note in particular that this estimator does not involve the times $\{t_1, ..., t_m\}$ in the path.

Equation 2.18 can be derived by considering the Laplace transform of the partition function (with $s > -E_g$)

$$\int_0^\infty e^{-\beta s} Z(\beta) d\beta \;=\; Tr\left[\frac{1}{H + s}\right] \, .$$

We can also express this as

$$\int_0^\infty e^{-\beta s} Z(\beta) d\beta =$$

$$\int_0^\infty d\beta e^{-\beta s} \sum_{m=0}^\infty \left[\lambda^m \sum_{\{z_1,\ldots,z_m\}} \langle z_1| - V|z_m\rangle \langle z_m| - V|z_{m-1}\rangle \ldots \langle z_2| - V|z_1\rangle \right.$$

$$\left. \int_0^\infty du_{m+1} \ldots \int_0^\infty du_1 e^{-\sum_{i=1}^{m+1} E_i u_i} \delta(\beta - \sum_{i=1}^{m+1} u_i) \right]$$

$$= \sum_{m=0}^\infty \lambda^m \sum_{\{z_1,\ldots,z_m\}} \langle z_1| - V|z_m\rangle \ldots \langle z_2| - V|z_1\rangle \prod_{i=1}^{m+1} \frac{1}{E_i + s} \,.$$

Performing the inverse Laplace transform gives

$$Z(\beta) = \sum_{m=0}^\infty \lambda^m \sum_{\{z_1,\ldots,z_m\}} \langle z_1| - V|z_m\rangle \ldots \langle z_2| - V|z_1\rangle \frac{1}{2\pi i} \int_C ds e^{\beta s} \left(\prod_{i=1}^{m+1} \frac{1}{E_i + s} \right).$$

C is a contour in the complex plane which encircles the poles of the integrand, which are located at $\{-E_i\}$. The expectation value of the energy can then be expressed as

$$\frac{1}{Z(\beta)} Tr[He^{-\beta H}] \quad = - \quad \frac{1}{Z} \frac{dZ}{d\beta}$$

$$= \quad \left\langle \frac{\frac{1}{2\pi i} \int_C (-s) e^{\beta s} \left(\prod_{i=1}^{m+1} \frac{1}{E_i+s} \right) ds}{\frac{1}{2\pi i} \int_C e^{\beta s} \left(\prod_{i=1}^{m+1} \frac{1}{E_i+s} \right) ds} \right\rangle_\rho . \qquad (2.19)$$

The complex function $h(s) = e^{\beta s} \left(\prod_{i=1}^{m+1} \frac{1}{E_i+s} \right)$ will in general have multiple saddle points along the real axis. We can solve for the locations of these saddle points by writing

$$h(s) = e^{\beta g(s)}$$

where

$$g(s) = s - \frac{1}{\beta} \sum_{i=1}^{m+1} \log(E_i + s).$$

Saddle points occur at real values $s^\star$ where $\frac{dg}{ds}(s^\star) = 0$, which says that

$$\beta = \sum_{i=1}^{m+1} \frac{1}{E_i + s^\star} \,.$$

In the integrals in equation 2.19, we can choose the contour of integration along the curve of steepest descent through that saddle point $s^\star$ which is largest (it is possible

53

to show that one can deform the contour to follow this curve without changing the encircled poles).

Performing the integrals in equation 2.19 and letting $E^\star = -s^\star$ we obtain

$$\frac{1}{Z(\beta)} Tr[He^{-\beta H}] = \langle E^\star \rangle + O\left(\frac{1}{\beta}\right)$$

where $E^\star$ is the smallest solution to equation 2.18.

2.5.2 A Variational Ansatz for the Ground State

The central object in the Quantum Monte Carlo method at fixed energy is the operator

$$A(E) = \left(\frac{-\lambda(E)}{H_0 - E} V\right).$$

We have seen that, for large m, $A(E)^m$ projects onto the ground state of $H(\lambda(E))$. If $|\phi\rangle$ has nonzero overlap with the ground state, and if $E = E_g(\lambda)$, then the states

$$A(E)^m |\phi\rangle \tag{2.20}$$

become closer to the ground state of $H(\lambda)$ as m increases.

In Chapter 4 we prove upper and lower bounds on the ground state energy for scrambled Hamiltonians. We use the variational lower bound discussed in Chapter 1. We are able to prove a lower bound on the energy per spin which is exact as $n \to \infty$ using an ansatz of the form 2.20 with $m = 2$. Note that this procedure for characterising the ground energy is a bit backwards: in order to prove a bound on the energy we first guess the value of E to plug into the ansatz 2.20!

2.6 Conclusions

In this Chapter we outlined an approach to Quantum Monte Carlo simulations in which properties of the ground state of a quantum system are computed at a fixed value of the ground state energy. We confirmed the validity of our method by performing a numerical simulation of a system consisting of 16 spins. Our approach involves a path integral which does not include any jump time variables, as in the stochastic series expansion [56, 57], which is another Quantum Monte Carlo method.

We have mentioned two results which are inspired by our numerical method. By a saddle point calculation, we obtained an estimator for the ground state energy which is valid in continuous imaginary time Quantum Monte Carlo simulations but which does not involve the imaginary time variables. We also introduced a variational ansatz for the ground state of a stoquastic Hamiltonian which will be useful to us in Chapter 4.

Chapter 3

Perturbative Crosses

In this Chapter we discuss a random ensemble of 3SAT instances which exhibit the "perturbative crosses" that we mentioned in Chapter 1. Different examples with perturbative crosses have been considered by other authors in references [2, 4, 5, 18] . The material in this Chapter is for the most part excerpted from the paper [18] which is joint work with Edward Farhi, Jeffrey Goldstone, Sam Gutmann, Harvey Meyer, and Peter Shor.

In section 3.1 we construct a set of random instances of 3SAT for which the adiabatic interpolating Hamiltonian

$$H(s) = (1 - s) \sum_{i=1}^{n} \left(\frac{1 - \sigma_x^i}{2} \right) + sH_P$$

has a very small gap due to a perturbative cross. We then show in section 3.2 that by randomly choosing a new beginning Hamiltonian $\tilde{H}_B$ and using the interpolating Hamiltonian

$$\tilde{H}(s) = (1 - s) \, \tilde{H}_B + sH_P$$

the perturbative cross is removed with probability bounded away from zero by a constant. We confirm our results by Quantum Monte Carlo simulations in section 3.3.

3.1   Problematic Instances

We now describe the method we use to generate n bit random instances of 3SAT that lead to quantum adiabatic Hamiltonians with small minimum gaps. To do this we first generate an instance with exactly two satisfying assignments given by the bit strings 111...1 and 000...0 . Each clause c of the 3SAT instance specifies a subset of 3 bits $i_1(c), i_2(c), i_3(c) \in \{1, .., n\}$ and a particular assignment $w_1(c)w_2(c)w_3(c)$ to those three bits which is disallowed. In order to only generate instances which are consistent with the bit strings 111...1 and 000...0, we only use clauses for which

$$w_1(c)w_2(c)w_3(c) \in \{100, 010, 001, 110, 101, 011\} .$$

We add such clauses one at a time uniformly at random and stop as soon as these two bit strings are the only bit strings which satisfy all of the clauses that have been added. (In practice to check whether or not this is the case we use a classical 3SAT solver.) We write m for the total number of clauses in the instance. We note that the number of clauses m obtained using this procedure scales like $n \log n$. We need this many clauses in order to ensure that each bit is involved in some clause. On the other hand, when the number of clauses is $5n \log n$ the probability of additional satisfying assignments goes to zero as $n \to \infty$.

We now consider the problem Hamiltonian H'_P corresponding to this instance, which we define to be

$$H'_P = \sum_{c=1}^{m} \left(\frac{1 + (-1)^{w_1(c)} \sigma_z^{i_1(c)}}{2} \right) \left(\frac{1 + (-1)^{w_2(c)} \sigma_z^{i_2(c)}}{2} \right) \left(\frac{1 + (-1)^{w_3(c)} \sigma_z^{i_3(c)}}{2} \right).$$

Each term in this sum is 1 if $i_1(c)i_2(c)i_3(c)$ violates the clause and 0 otherwise. For the beginning Hamiltonian we choose

$$H_B = \sum_{i=1}^{n} \left(\frac{1 - \sigma_x^i}{2} \right). \tag{3.1}$$

The two lowest eigenvalues of the Hamiltonian $H'(s) = (1-s)H_B + sH'_P$ will then both approach 0 as $s \to 1$, as in Figure 1-3 (left). The ground state for values of s which are sufficiently close to 1 will approach either $|000...0\rangle$ or $|111...1\rangle$ as $s \to 1$. Suppose for values of s close to 1 the state that approaches $|000...0\rangle$ has lowest energy. Then we add an extra term h_0 which acts on bits $1, 2$ and 3 and which penalizes this state but not $|111...1\rangle$

$$h_0 = \frac{1}{2} \left(\frac{1 + \sigma_z^1}{2} \right) \left(\frac{1 + \sigma_z^2}{2} \right) \left(\frac{1 + \sigma_z^3}{2} \right)$$

and pick

$$H_P = H'_P + h_0.$$

(Note that this extra term has a multiplicative factor of $\frac{1}{2}$, to avoid any degeneracy of the first excited state at $s = 1$.) In the case where the lowest energy state near $s = 1$ approaches $|111...1\rangle$ we instead add

$$h_1 = \frac{1}{2} \left(\frac{1 - \sigma_z^1}{2} \right) \left(\frac{1 - \sigma_z^2}{2} \right) \left(\frac{1 - \sigma_z^3}{2} \right).$$

The Hamiltonian $H(s) = (1-s)H_B + sH_P$ is then expected to have a small gap near $s = 1$ as depicted in Figure 1-3 (right). We expect the location s^* of the minimum gap to approach $s = 1$ as $n \to \infty$.

56

Location of the Minimum Gap

In order to determine the dependence of the location $s^\star$ of the avoided crossing on the number of spins n and the number of clauses m, we consider the perturbative corrections to the energies of the states $|000...0\rangle$ and $|111...1\rangle$ around $s = 1$. We will show that the low order terms in the perturbation series reliably predict a crossing at[1] $s^\star = 1 - \Theta(\frac{1}{n^{1/4}} \left(\frac{m}{n}\right)^{\frac{3}{4}})$.

With the Hamiltonian constructed in the previous section, one of these states has energy 0 at $s = 1$ (call this the lower state $|z_L\rangle$) and the other state has energy $\frac{1}{2}$ at $s = 1$ (we call this the upper state $|z_U\rangle$). We write our Hamiltonian as

$$H(s) = (1-s)\frac{n}{2} + s \left[-\left(\frac{1-s}{s}\right) \sum_{i=1}^{n} \frac{\sigma_x^i}{2} + H_P \right] \tag{3.2}$$

and we then consider the term $-\left(\frac{1-s}{s}\right) \sum_{i=1}^{n} \frac{\sigma_x^i}{2}$ as a perturbation to H_P expanding around $s = 1$.

To understand why we trust perturbation theory to predict the location of the near crossing, consider a system which is composed of two disconnected sectors, A and B, so the corresponding Hamiltonian is of the form

$$\begin{pmatrix} H_A(s) & 0 \\ 0 & H_B(s) \end{pmatrix} .$$

In this situation the generic rule that levels do not cross[42] does not apply and we can easily imagine that the two lowest levels look like what we show in Figure 3-1, where the levels actually cross at $s^\star$.

Imagine that low order perturbation theory around $s = 1$ can be used to get good approximations to the ground state energies of $H_A(s)$ and $H_B(s)$ for s near $s^\star$, even for s somewhat to the left of $s^\star$. Then it is possible to accurately predict $s^\star$. Our situation is very close to this. We can think of A as consisting of the states close to z_L in Hamming weight, and B as states close to z_U in Hamming weight. Similarly, we view H_A and H_B as the restrictions of H to these sectors. Note that it takes n powers of the perturbation to connect $|000...0\rangle$ and $|111...1\rangle$ and this is why we view A and B as essentially disconnected.

Although Figure 3-1 looks like Figure 1-3 (right), it is Figure 1-3 (right) that depicts the actual situation, where the two levels avoid crossing. This true near cross means that the perturbation series in the actual theory will diverge very close to $s^\star$. However this divergence will only be seen at high order, in fact at an order which is proportional to n. The low order terms of the perturbation series in Figure 3-1 are the same as the low order terms of the perturbation series in Figure 1-3 (right), so we can trust low order perturbation theory to locate $s^\star$.

(We argue below that as a function of the number of bits, n, $s^\star$ goes to 1 as n goes to infinity. This implies that the radius of convergence of the perturbation theory

---

[1]The notation $f(n) = \Theta(g(n))$ means that, for n sufficiently large, $bg(n) \leq f(n) \leq cg(n)$ for some constants b and c.
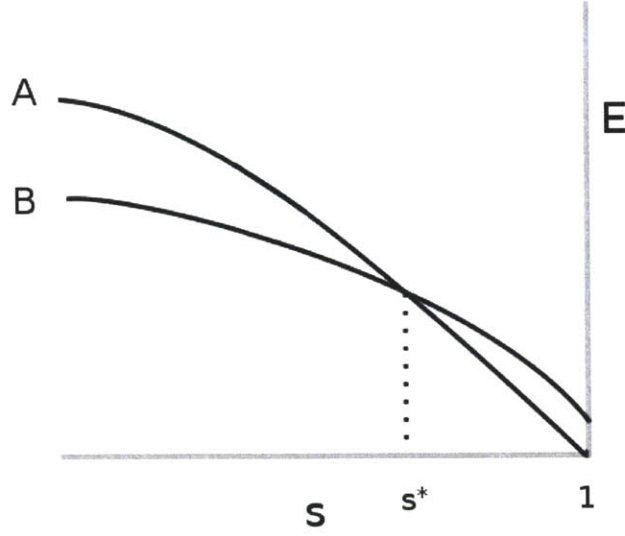
Figure 3-1: A true energy level crossing can arise from two disconnected sectors.

for the full $H(s)$, expanded about $s = 1$, goes to 0 as n goes to infinity. This fact has no bearing on our argument that low order perturbation theory can be used to accurately predict $s^\star$.)

For small values of the parameter $\frac{1-s}{s}$, the energies of the two states under consideration can be expanded as

$$E_L(s) = (1-s)\frac{n}{2} + s\left[0 + \left(\frac{1-s}{s}\right)^2 e_L^{(2)} + \left(\frac{1-s}{s}\right)^4 e_L^{(4)} + \ldots\right] \quad (3.3)$$

$$E_U(s) = (1-s)\frac{n}{2} + s\left[\frac{1}{2} + \left(\frac{1-s}{s}\right)^2 e_U^{(2)} + \left(\frac{1-s}{s}\right)^4 e_U^{(4)} + \ldots\right]. \quad (3.4)$$

It is easy to see that each expansion (inside the square brackets) only contains even powers. Note that $e_L^{(2)}$ is guaranteed to be negative as is always the case for 2nd order perturbation theory of the ground state. In addition, since we added a term to penalize the state which had smaller energy near $s = 1$ before adding the clause, we expect that $e_U^{(2)} < e_L^{(2)}$. We are interested in the behaviour of the difference $E_L(s) - E_U(s)$ for randomly generated instances (as generated using the prescription of the previous section) as a function of the number of spins n in the limit $n \to \infty$. This requires us to further investigate the behaviour of each coefficient $e_L^{(k)}$ and $e_U^{(k)}$ as a function of n and m.

In fact to locate $s^\star$ we need only to go to second order in perturbation theory. From equation 3.2 we view H_P as the unperturbed Hamiltonian, and $-\frac{1}{2}\sum_{i=1}^{n}\sigma_x^i$ as the perturbation, with $\frac{1-s}{s}$ as the expansion parameter. Now $\sigma_x^i|z\rangle = |z \oplus \hat{e}_i\rangle$ so at

second order we get

$$e_L^{(2)} = \frac{1}{4} \sum_{i=1}^{n} \frac{|\langle z_L \oplus \hat{e}_i | \sigma_x^i | z_L \rangle|^2}{\langle z_L | H_P | z_L \rangle - \langle z_L \oplus \hat{e}_i | H_P | z_L \oplus \hat{e}_i \rangle}$$

$$= -\frac{1}{4} \sum_{i=1}^{n} \frac{1}{\langle z_L \oplus \hat{e}_i | H_P | z_L \oplus \hat{e}_i \rangle} .$$

Similarly,

$$e_U^{(2)} = \frac{1}{4} \sum_{i=1}^{n} \frac{1}{\frac{1}{2} - \langle z_U \oplus \hat{e}_i | H_P | z_U \oplus \hat{e}_i \rangle}$$

where the $\frac{1}{2}$ in the denominator is $\langle z_U | H_P | z_U \rangle$.

The expected energy penalty incurred when flipping a bit of either z_U or z_L is of order $\frac{m}{n}$ since each bit is typically involved in $\Theta(\frac{m}{n})$ clauses. So the coefficients $e_L^{(2)}$ and $e_U^{(2)}$ are of order $n\left(\frac{n}{m}\right)$ since the energy denominators involved are $\Theta(\frac{m}{n})$. We now show that their difference is of order $\sqrt{n}\left(\frac{n}{m}\right)^{\frac{3}{2}}$. Write

$$e_U^{(2)} - e_L^{(2)} = \sum_{i=1}^{n} d_i$$

where for each i,

$$d_i = \frac{1}{4} \left(\frac{1}{\frac{1}{2} - \langle z_U \oplus \hat{e}_i | H_P | z_U \oplus \hat{e}_i \rangle} + \frac{1}{\langle z_L \oplus \hat{e}_i | H_P | z_L \oplus \hat{e}_i \rangle} \right). \qquad (3.5)$$

Recall that $H_P = H_P' + h$ where h is the penalty term from the final clause which acts only on the first 3 bits. Therefore, for $i = 4, 5, ..., n$

$$d_i = \frac{1}{4} \left(-\frac{1}{\langle z_U \oplus \hat{e}_i | H_P' | z_U \oplus \hat{e}_i \rangle} + \frac{1}{\langle z_L \oplus \hat{e}_i | H_P' | z_L \oplus \hat{e}_i \rangle} \right) .$$

Our procedure for generating instances is symmetric between the strings 000...0 and 111...1 so averaging over instances it is clear that the mean of d_i for $i = 4, 5..., n$ is 0. Thus we expect $\sum_{i=4}^{n} d_i$ to be (approximately) Gaussian with mean 0 and standard deviation proportional to $\sqrt{n}\sigma(d)$, where $\sigma(d)$ is the standard deviation of each d_i for $i \in \{4, 5, ..., n\}$. To compute $\sigma(d)$ we note that

$$d_i = \frac{1}{4} \left(\frac{\langle z_U \oplus \hat{e}_i | H_P' | z_U \oplus \hat{e}_i \rangle - \langle z_L \oplus \hat{e}_i | H_P' | z_L \oplus \hat{e}_i \rangle}{\langle z_L \oplus \hat{e}_i | H_P' | z_L \oplus \hat{e}_i \rangle \langle z_U \oplus \hat{e}_i | H_P' | z_U \oplus \hat{e}_i \rangle} \right) .$$

Again using the symmetry between all zeros and all ones, we conclude that the numerator is of order $\sqrt{\frac{m}{n}}$ and the denominator is of order $\left(\frac{m}{n}\right)^2$. Hence we expect $\sigma(d)$ to be $\Theta\left(\left(\frac{n}{m}\right)^{\frac{3}{2}} \right)$. So $e_U^{(2)} - e_L^{(2)}$ is of order $\sqrt{n}\left(\frac{n}{m}\right)^{\frac{3}{2}}$. We will now locate $s^\star$ using second order perturbation theory and afterwards argue that higher orders do not change the

result. Returning to equations 3.3 and 3.4, equating the two energies at second order we have

$$0 = s^\star \left[\frac{1}{2} + \left(\frac{1 - s^\star}{s^\star} \right)^2 \left(e_U^{(2)} - e_L^{(2)} \right) \right]$$

so

$$s^\star = 1 - \Theta\left(\frac{1}{n^{1/4}} \left(\frac{m}{n} \right)^{\frac{3}{4}} \right). \tag{3.6}$$

The 4th order correction to the energy of the lower state is given by

$$e_L^{(4)} = \langle z_L | V \left(\frac{\phi_L}{H_P} \right)^2 V | z_L \rangle \langle z_L | V \frac{\phi_L}{H_P} V | z_L \rangle - \langle z_L | V \frac{\phi_L}{H_P} V \frac{\phi_L}{H_P} V \frac{\phi_L}{H_P} V | z_L \rangle$$

where

$$V = -\frac{1}{2} \sum_{i=1}^{n} \sigma_x^i$$

and $\phi_L = 1 - |z_L\rangle\langle z_L|$. Writing $\mathcal{H}_P(z) = \langle z | H_P | z \rangle$, this can be expressed as

$$
\begin{aligned}
e_L^{(4)} &= \frac{1}{16} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{1}{\left(\mathcal{H}_P(z_L \oplus \hat{e}_i) \right)^2 \mathcal{H}_P(z_L \oplus \hat{e}_j)} \\
&\quad - \frac{1}{16} \sum_{i \neq j} \frac{1}{\mathcal{H}_P(z_L \oplus \hat{e}_i) \mathcal{H}_P(z_L \oplus \hat{e}_j) \mathcal{H}_P(z_L \oplus \hat{e}_i \oplus \hat{e}_j)} \\
&\quad - \frac{1}{16} \sum_{i \neq j} \frac{1}{\left(\mathcal{H}_P(z_L \oplus \hat{e}_i) \right)^2 \mathcal{H}_P(z_L \oplus \hat{e}_i \oplus \hat{e}_j)} \\
&= \sum_{i=1}^{n} \frac{1}{16} \frac{1}{\left(\mathcal{H}_P(z_L \oplus \hat{e}_i) \right)^3} \\
&\quad + \sum_{i \neq j} \frac{1}{16} \frac{\mathcal{H}_P(z_L \oplus \hat{e}_i \oplus \hat{e}_j) - \mathcal{H}_P(z_L \oplus \hat{e}_i) - \mathcal{H}_P(z_L \oplus \hat{e}_j)}{\left(\mathcal{H}_P(z_L \oplus \hat{e}_i) \right)^2 \mathcal{H}_P(z_L \oplus \hat{e}_j) \left(\mathcal{H}_P(z_L \oplus \hat{e}_i \oplus \hat{e}_j) \right)}.
\end{aligned}
$$

Now consider the terms in this expression corresponding to indices i, j for which $i \neq j$ and the bits i and j do not appear in a clause together. Under these conditions we have

$$\mathcal{H}_P(z_L \oplus \hat{e}_i \oplus \hat{e}_j) = \mathcal{H}_P(z_L \oplus \hat{e}_i) + \mathcal{H}_P(z_L \oplus \hat{e}_j).$$

So we can write

$$
\begin{aligned}
e_L^{(4)} &= \sum_{i=1}^{n} \frac{1}{16} \frac{1}{\left(\mathcal{H}_P(z_L \oplus \hat{e}_i) \right)^3} \tag{3.7} \\
&\quad + \sum_{\substack{i \neq j \\ \text{clausemates}}} \frac{1}{16} \frac{\mathcal{H}_P(z_L \oplus \hat{e}_i \oplus \hat{e}_j) - \mathcal{H}_P(z_L \oplus \hat{e}_i) - \mathcal{H}_P(z_L \oplus \hat{e}_j)}{\left(\mathcal{H}_P(z_L \oplus \hat{e}_i) \right)^2 \mathcal{H}_P(z_L \oplus \hat{e}_j) \left(\mathcal{H}_P(z_L \oplus \hat{e}_i \oplus \hat{e}_j) \right)}.
\end{aligned}
$$

Here the subscript "clausemates" indicates that we sum only over pairs of indices which appear together in at least one clause of the 3SAT instance corresponding to

H_P. For $e_U^{(4)}$, since the unperturbed energy is $\frac{1}{2}$ we obtain

$$e_U^{(4)} = \tag{3.8}$$

$$\sum_{i=1}^{n} \frac{1}{16} \frac{1}{\left(\mathcal{H}_P(z_U \oplus \hat{e}_i) - \frac{1}{2}\right)^3} \tag{3.9}$$

$$+ \sum_{\substack{i \neq j \\ \text{clausemates}}} \frac{\left(\mathcal{H}_P(z_U \oplus \hat{e}_i \oplus \hat{e}_j) - \frac{1}{2}\right) - \left(\mathcal{H}_P(z_U \oplus \hat{e}_i) - \frac{1}{2}\right) - \left(\mathcal{H}_P(z_U \oplus \hat{e}_j) - \frac{1}{2}\right)}{16 \left(\mathcal{H}_P(z_U \oplus \hat{e}_i) - \frac{1}{2}\right)^2 \left(\mathcal{H}_P(z_U \oplus \hat{e}_j) - \frac{1}{2}\right) \left(\mathcal{H}_P(z_U \oplus \hat{e}_i \oplus \hat{e}_j) - \frac{1}{2}\right)}$$

Let's look at the first sum in each of equations 3.7 and 3.9 where i goes from 1 to n. Each term as i goes from 4 to n is of order $\left(\frac{n}{m}\right)^3$ and so the difference of the two sums is $\Theta(\sqrt{n} \left(\frac{n}{m}\right)^3)$. The second sums (those which are restricted to clausemates) contain of order m terms. Each denominator is of order $\left(\frac{m}{n}\right)^4$ and the numerators are $\Theta(1)$ since the only contribution to the numerator is from clauses in H_P which involve bits i and j together. Separating out the terms where i and/or j are $1, 2$ or 3, we conclude that the contribution to $e_U^{(4)} - e_L^{(4)}$ from the clausemate terms is $\Theta(\sqrt{n} \left(\frac{n}{m}\right)^{\frac{7}{2}})$. For our instance generation m grows like $n \log n$ so this clausemate contribution is asymptotically dominated by the first term which scales like $\Theta(\sqrt{n} \left(\frac{n}{m}\right)^3)$. So the fourth order contribution to the difference of energies $E_U(s) - E_L(s)$ is $\Theta(s \left[\sqrt{n} \left(\frac{n}{m}\right)^3 \left(\frac{1-s}{s}\right)^4\right])$. At $s^\star$ which we determined at second order to be $1 - \Theta(\frac{1}{n^{1/4}} \left(\frac{m}{n}\right)^{\frac{3}{4}})$, the fourth order contribution to the energy difference is $\Theta\left(\frac{1}{\sqrt{n}}\right)$. The fourth order corrections can therefore be neglected in determining the location of $s^\star$. Sixth order and higher contributions to the difference are even smaller.

3.2 Fixing the Problem by Path Change

In our instance generation we manufactured a small gap by penalizing the planted assignment corresponding to the energy eigenstate with the smallest energy near $s = 1$. Since the slopes of the two curves in Figure 1-3(left) are the same at $s = 1$, the second derivatives determine which eigenvalue is smaller near $s = 1$. After penalization we have $e_U^{(2)} < e_L^{(2)}$ which is consistent with the near crossing in Figure 1-3(right). Suppose instead that we penalize the assignment corresponding to larger energy in Figure 1-3(left). Then we expect the situation depicted in Figure 3-2 where no level crossing is induced.

We imagine that the instances that we manufacture with a small gap as in Figure 1-3(right) are a model for what might be encountered in running the quantum adiabatic algorithm on some instance a quantum computer is actually trying to solve. There is a strategy for overcoming this problem. The idea is to produce Figure 3-2, with reasonable probability, by randomly modifying the adiabatic path which ends at H_P, of course making no use of the properties of the particular instance. For this purpose one could use any random ensemble of paths $H(s)$ such that $H(1) = H_P$ and the
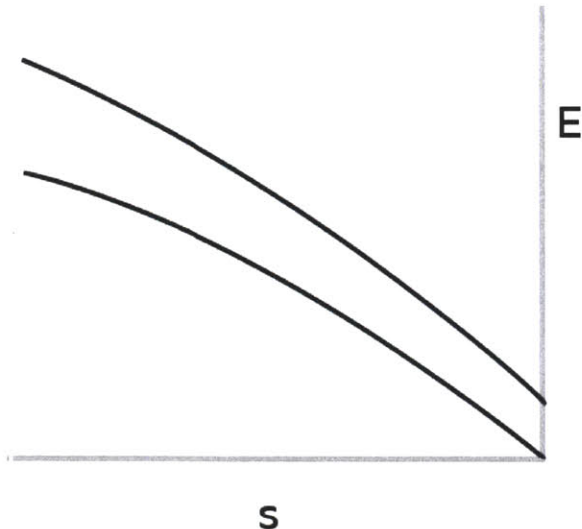
Figure 3-2: $s = 1$. Here the second derivative of the upper curve B is greater than the second derivative of the lower curve A.

ground state of $H(0)$ is simple to prepare. However in this paper we only consider randomly changing the beginning Hamiltonian. We have made this choice so that we are able to use our Quantum Monte Carlo method to numerically verify our arguments. Like other Quantum Monte Carlo methods, the method we use does not work when the Hamiltonian has nonzero off-diagonal matrix elements with positive sign.

Starting with an instance where $H(s) = (1 - s)H_B + sH_P$ has a tiny gap due to the problem discussed above, we now consider a different adiabatic path $\tilde{H}(s) = (1-s)\tilde{H}_B + sH_P$ obtained by keeping the same problem Hamiltonian H_P but choosing a different beginning Hamiltonian $\tilde{H}_B$ in a random fashion which we will prescribe below. We argue that the small gap near $s = 1$ is then removed with substantial probability, so that by repeating this procedure a constant number of times it is possible to find an adiabatic path without a small gap near $s = 1$.

The way that we choose a random Hamiltonian $\tilde{H}_B$ is to first draw n random variables c_i for $i = 1, 2, ..., n$, where each c_i is chosen to be $\frac{1}{2}$ or $\frac{3}{2}$ with equal probability. We then take

$$\tilde{H}_B = \sum_{i=1}^{n} \frac{c_i(1 - \sigma_x^i)}{2} . \tag{3.10}$$

We write $\tilde{e}_U^{(2)}$ and $\tilde{e}_L^{(2)}$ for the analogous quantities to $e_L^{(2)}$ and $e_U^{(2)}$ for the new Hamiltonian $\tilde{H}(s) = (1 - s)\tilde{H}_B + sH_P$. The point is that by randomizing $\tilde{H}_B$ in the way we prescribe, there is a substantial probability that one will obtain $\tilde{e}_U^{(2)} - \tilde{e}_L^{(2)} > 0$,

and in that case one expects no avoided crossing near $s = 1$. Write

$$\tilde{e}_U^{(2)} - \tilde{e}_L^{(2)} = \sum_{i=1}^{n} c_i^2 d_i$$

where the $\{d_i\}$ are fixed by the instance (and are defined in equation 3.5). Since we have fixed the problem Hamiltonian H_P, the only random variables appearing in the above equation are the c_i. We have $\overline{c_i^2} = \frac{5}{4}$ so the mean value of $\tilde{e}_U^{(2)} - \tilde{e}_L^{(2)}$ is then

$$\overline{\tilde{e}_U^{(2)} - \tilde{e}_L^{(2)}} = \frac{5}{4}(e_U^{(2)} - e_L^{(2)}) < 0.$$

But more importantly

$$\overline{\tilde{e}_U^{(2)} - \tilde{e}_L^{(2)}} = \Theta(\sqrt{n}\left(\frac{n}{m}\right)^{\frac{3}{2}}).$$

The variance of this difference is

$$\begin{aligned}
\mathrm{Var}\left(\tilde{e}_U^{(2)} - \tilde{e}_L^{(2)}\right) &= \sum_{i=1}^{n} d_i^2 \mathrm{Var}(c_i^2) \\
&= \sum_{i=1}^{n} d_i^2 \cdot 1,
\end{aligned}$$

which is $\Theta(n\left(\frac{n}{m}\right)^3)$. For a fixed instance with a corresponding fixed set $\{d_i\}$ the random variable $\sum_i c_i^2 d_i$ is approximately Gaussian and from its mean and variance we see that the probability that $\tilde{e}_U^{(2)} - \tilde{e}_L^{(2)}$ is positive and in fact greater than $a\sqrt{n}\left(\frac{n}{m}\right)^{\frac{3}{2}}$, for $a > 0$, is bounded away from 0 independent of n. This means that there is a good chance that randomizing H_B turns the situation depicted in Figure 1-3 (right) into the situation depicted in Figure 3-2.

In the case of two planted satisfying assignments with one penalized to produce a small gap when the beginning Hamiltonian is H_B of equation 3.1, we have shown that a random choice for the beginning Hamiltonian $\tilde{H}_B$ of equation 3.10 can with substantial probability remove the small gap. This gives further weight to the idea that when running the quantum adiabatic algorithm on a single instance of some optimization problem, the programmer should run the quantum adiabatic algorithm repeatedly with different paths ending at H_P [19].

3.3   Quantum Monte Carlo Simulations

We have seen in section 1.2.3 that sampling with respect to the probability distribution ρ over paths (from equation 1.17) allows one to compute thermal properties of a quantum system.

We have developed a quantum monte carlo simulator which samples from ρ using a modified version of the heat bath algorithm of Krzakala et al [41]. Our method differs from the scheme outlined in reference [41] in that we use our novel single spin

sampling algorithm which is given in Appendix B .

As with other Quantum Monte Carlo methods, this algorithm is a Markov Chain Monte Carlo method. In order to sample from the distribution ρ, one defines a Markov Chain over the state space consisting of all paths, where the limiting distribution of the chain is ρ. To obtain samples from ρ one starts in an initial path $z_{init}(t)$ and then applies some number N_{equil} of iterations of the Markov Chain. If N_{equil} is sufficiently large then the distribution of the paths found by further iteration will be close to ρ. We use these subsequent paths to compute averages with respect to ρ.

The figures which are referred to in this section are located in section 3.5.

3.3.1 Equilibration of the Quantum Monte Carlo and Identification of Level Crossings

Our discussion in Chapter 1 demonstrates how one can estimate expectation values of various quantities in the ground state of a quantum system. The method is "continuous time" so there is no discretization error and for fixed β quantities can be arbitrarily well approximated with enough statistics. We use the Quantum Monte Carlo method in a nonstandard way in order to be able to study the lowest two eigenstates of our Hamiltonian (as opposed to just the ground state).

As we have just noted, the standard procedure for generating configurations is to equilibrate to the distribution ρ from some initial path $z_{init}(t)$ (we call this the seeded path) by applying the Monte Carlo update N_{equil} times. In order to ensure that one has equilibrated after N_{equil} Monte Carlo updates, one could for example do simulations with two or more different initial paths (seeded paths) and check that the values of Monte Carlo observables appear to converge to the same seed independent values.

For each instance we consider, we run two different Monte Carlo simulations that are seeded with two different paths $z_{init}^0(t)$ and $z_{init}^1(t)$. These seeds are paths with no flips in them, corresponding to the two states 000...0 and 111...1,

$$
\begin{aligned}
z_{init}^0(t) &= 000...0 \text{ for all } t \in [0,\beta] \\
z_{init}^1(t) &= 111...1 \text{ for all } t \in [0,\beta] \,.
\end{aligned}
$$

With each seed, we run the Monte Carlo simulation for some total number N_{total} of Monte Carlo sweeps, taking data every k sweeps. Here a single sweep is defined to be n iterations of our Monte Carlo update rule as described in the appendix (where n is the number of spins). We then remove the first N_{equil} Monte Carlo samples, and use the remaining Monte Carlo samples to estimate the thermal averages of the energy and the Hamming weight operator

$$
\begin{aligned}
\langle H \rangle &= \frac{Tr[He^{-\beta H}]}{Tr[e^{-\beta H}]} \\
\langle W \rangle &= \frac{Tr[We^{-\beta H}]}{Tr[e^{-\beta H}]}
\end{aligned}
$$

64

(where $W = \sum_{i=1}^{n} \left(\frac{1-\sigma_z^i}{2} \right)$ is the Hamming weight operator) using equations 1.19, 1.20, and 1.18.

In order to convince the reader that the Monte Carlo algorithm works correctly, and that we can obtain information about the first excited state, we show data at 16 bits where exact numerical diagonalization is possible. In Figures 3-3 and 3-4 we show the result of this procedure for a 3SAT instance with 16 bits, with problem Hamiltonian H'_P (before adding the penalty term h, so that the levels are degenerate at $s = 1$). This instance has 122 clauses. The inverse temperature is $\beta = 150$. The total number of Monte Carlo sweeps at each value of s is $N_{total} = 200000$, with data taken after every fifth sweep (this gave us 40000 data samples). We use $N_{equil} = 2500$ samples solely for equilibration at each value of s. Note that the Monte Carlo simulations with the two different seeds (corresponding to the circles and crosses in the Figures) only agree for values of s less than roughly 0.4. We interpret this to mean that at these values of s the Quantum Monte Carlo has equilibrated to the proper limiting distribution ρ regardless of the seed. As s increases past 0.4, the two simulations abruptly begin to give different results. In this case the simulation seeded with $z_{init}^1(t)$ finds a metastable equilibrium of the Markov Chain (i.e not the limiting distribution ρ), which corresponds in this case to the first excited state of the Hamiltonian. For s above 0.4 we can see from the comparison with exact diagonalization that the two differently seeded values allow us to compute the energies (Figure 3-3) and Hamming weights (Figure 3-4) of the lowest two energy levels of our Hamiltonian.

What is happening here is that for large s, the quantum system can be thought of as consisting of two disconnected sectors. One sector consists of states in the z basis with low Hamming weight and the other with Hamming weight near n. The Quantum Monte Carlo "equilibrates" in each sector depending on the initial seed as can be seen by the smooth data in each sector. (Note that data is taken independently at each value of s.) The lower of the cross and circle at each value of s in Figure 3-3 is the ground state energy. Of course if the classical algorithm ran for long enough then the circles in Figure 3-3 would lie on the crosses for every value of s.

In Figures 3-5 and 3-6 we show data taken for the same instance after the penalty clause is added that removes the degeneracy at $s = 1$. In this case the two levels have an avoided crossing as expected and which can be seen by exact numerical diagonalization in Figure 3-5. However in the Monte Carlo simulation with two seeds there are two essentially disconnected sectors and the two levels being tracked do cross. When we see this behaviour in the Monte Carlo simulation, we interpret it as compelling evidence that there is a tiny gap in the actual system, which occurs where the curves cross. In Figure 3-6 first look at the curves coming from the exact numerical diagonalization. The Hamming weight of the ground state decreases and then abruptly rises at the location of the minimum gap. The Hamming weight of the first excited state also undergoes an abrupt change at the location of the minimum gap. Not surprisingly the exact diagonalization clearly shows the behaviour we expect with the manufactured tiny gap. Now look at the Monte Carlo data which is interpreted by first looking at Figure 3-5. In Figure 3-5 the true ground state is always the lower

of the circles and the crosses. For s below $s^\star \approx 0.423$ it is the crosses and for s larger than $s^\star$ it is the circles. Accordingly in Figure 3-6 the Hamming weight of the ground state is tracked by the crosses for s below $s^\star$ and by the circles for s above $s^\star$. We can therefore conclude, based only on the Monte Carlo data, that the Hamming weight of the ground state changes abruptly. At higher bit number we do not have exact numerical diagonalization available but we can still use the Quantum Monte Carlo to extract key information.

3.3.2 Randomizing the Beginning Hamiltonian

For the 16 spin Hamiltonian depicted above, we now consider randomizing the beginning Hamiltonian H_B as described in section 3.2. In order to minimize the number of times we run the Monte Carlo algorithm (this will be more of an issue at high bit number where simulations are very time consuming), we generated many different sets of coefficients $\{c_i\}$ and calculated the differences $\tilde{e}_U^{(2)} - \tilde{e}_L^{(2)}$ with fixed problem Hamiltonian H_P' and beginning Hamiltonians

$$\tilde{H}_B = \sum_{i=1}^{n} c_i \left(\frac{1 - \sigma_x^i}{2} \right) .$$

According to our discussion in section 3.2, we expect the avoided crossing near $s = 1$ to be removed for choices of coefficients c_i such that the difference $\tilde{e}_U^{(2)} - \tilde{e}_L^{(2)} > 0$. We made a histogram of the $\tilde{e}_U^{(2)} - \tilde{e}_L^{(2)}$ (shown in Figure 3-7) and randomly chose three sets of coefficients $\{c_i\}$ such that $\tilde{e}_U^{(2)} - \tilde{e}_L^{(2)} > \frac{1}{2}$ for each. Our analysis predicts that in these cases the crossing will be removed. As expected, each of these three sets of coefficients resulted in an adiabatic Hamiltonian with the small gap removed, although in one there was another small gap at a smaller value of s. We plot the Monte Carlo data for one of these sets in Figures 3-8 and 3-9.

In section 3.1 we argued that we could manufacture a small gap in a 3SAT instance as sketched in Figure 1-3(right). Our 16 bit instance is a concrete example of this and the tiny gap can be seen in Figure 3-5. In section 3.2 we argued that by randomizing the beginning Hamiltonian we should be able to produce Figure 3-2. This is what we see concretely at 16 bits in Figure 3-8. We now tell the same story at 150 bits using only Monte Carlo data since exact numerical diagonalization is not possible.

3.3.3 Data for an Instance of 3SAT with 150 bits

In addition to validating our method at 16 bits, we studied 3SAT instances with $25, 75$ and 150 bits using our Quantum Monte Carlo simulator. The data from these simulations for the most part supported our arguments. In this section we present Monte Carlo data taken for a double plant instance of 3SAT with 150 spins and 1783 clauses. In this case the inverse temperature $\beta = 300$, and the total number of Monte Carlo sweeps at each value of s is $N_{total} = 100000$, with data taken every fifth sweep (giving 20000 data samples). The first 2500 data samples at each value of s

are removed for equilibration. We ran our simulations on a 648 processor SiCortex computer cluster in embarrassingly parallel fashion. We used a different processor for each value of s and each value of the seed. Data taken at the lower values of s took the longest to accumulate, in some cases more than 10 days on a single processor for a single data point.

Figures 3-10 and 3-11 show the energy and the Hamming weight before the penalty clause is added, running with two different seeds. In Figure 3-10 for large values of s, the crosses are always below the circles and track the ground state which ends at 000...0 as can be seen in Figure 3-11. We have also plotted the second order perturbation theory energies for these two levels, expanding around $s = 1$. Note the good agreement between the second order perturbation theory and the Monte Carlo data all the way down to $s = 0.35$.

Since the lower curve corresponds to 000...0, we penalize this assignment by the addition of a single clause to form H_P attempting to manufacture a near crossing. In Figure 3-12, the circle data is below the cross data for s near 1 but the two curves cross near $s = .49$. This is seen clearly in Figure 3-13 where the energy difference is plotted. From the Monte Carlo data we conclude that $H(s)$ has a tiny gap. The location of the avoided crossing is well predicted by second order perturbation theory as can be seen in Figure 3-13.

The Monte Carlo data at 150 bits shows that we can make an instance of 3SAT with a tiny gap following the procedure outlined in section 3.1. We now use the Monte Carlo simulator to show that a randomly chosen $\tilde{H}_B$ can alter the Hamiltonian $H(s)$ so that this small gap becomes large. For the instance at hand we first compute the difference $\tilde{e}_U^{(2)} - \tilde{e}_L^{(2)}$ for 100000 randomly chosen sets of coefficients. The histogram of these differences is plotted in Figure 3-15. After doing this we randomly selected (from this set) two sets of coefficients such that $\tilde{e}_U^{(2)} - \tilde{e}_L^{(2)} > \frac{1}{2}$. For both of these sets of coefficients we saw that the crossing at $s \approx 0.49$ was no longer present, although in one case there appeared to be a new crossing at a much lower value of s. Figures 3-16 and 3-17 show the Monte Carlo data for the choice of $\tilde{H}_B$ which does not appear to have any crossing. At 150 bits we see compelling evidence that the story outlined in sections 3.1 and 3.2 is true.

Is $s^\star$ near 1?

We argued that $s^\star$ should go to 1 for large enough n. However our 16 bit example has $s^\star \approx 0.42$ and the 150 bit example has $s^\star \approx 0.49$. Recall from equation 3.6 that $s^\star = 1 - \Theta(\frac{1}{n^{1/4}} \left(\frac{m}{n}\right)^{\frac{3}{4}})$. At 16 bits with $m = 122$ we have $\frac{1}{n^{1/4}} \left(\frac{m}{n}\right)^{\frac{3}{4}} = 2.29$ and at 150 bits with $m = 1783$ we have $\frac{1}{n^{1/4}} \left(\frac{m}{n}\right)^{\frac{3}{4}} = 1.83$. Although asymptotically m is of order $n \log n$, for these values of n we are not yet in the regime where $\frac{1}{n^{1/4}} \left(\frac{m}{n}\right)^{\frac{3}{4}} \ll 1$.

Even though $s^\star$ is not near 1 for our instance at 150 bits, we see from Figure 3-13 that second order perturbation theory can be used to predict the location of $s^\star$. This is because the fourth order contribution to the energy difference is quite small. So already at 150 bits we can predict the presence or absence of an avoided crossing using second order perturbation theory.

67

3.4 Conclusions

We have introduced a new Quantum Monte Carlo technique to analyze the performance of quantum adiabatic algorithms for instances of satisfiability. Using seeded configurations (in the Monte Carlo simulation) corresponding to disparate low lying states, our technique exposed the presence or absence of an exponentially small gap without ever actually computing the gap.

We used this method to numerically investigate a set of random instances of 3SAT which were designed to expose a weakness of the adiabatic algorithm. We confirmed that this weakness can be overcome for our set of instances by using path change. Our numerical work and the analysis in this Chapter pertains to instances of 3SAT with 2 planted satisfying assignments. Reference [18] also discusses the scenario where an instance has k satisfying assignments and then all but one are penalized with a small number of clauses.

In light of our results we believe that path change must be considered an integral part of the quantum adiabatic algorithm. For any given instance, the algorithm should be run with many different randomly selected paths which end at the problem Hamiltonian. As long as the algorithm succeeds on at least a polynomially small fraction of the trials, it can be used to solve decision problems.

The random ensemble of instances considered in this Chapter were engineered to have perturbative crosses between their two lowest energy levels. It is argued in references [2, 4] that such crosses are a generic feature which causes the adiabatic algorithm to fail for other clause based decision problems. The question of how perturbative crosses manifest themselves in other random ensembles of clause based optimization problems is subtle since there may be many low lying excited states which are disparate in Hamming weight. We do not have a complete picture of the performance of the quantum adiabatic algorithm on such problems.

3.5 Numerical Results
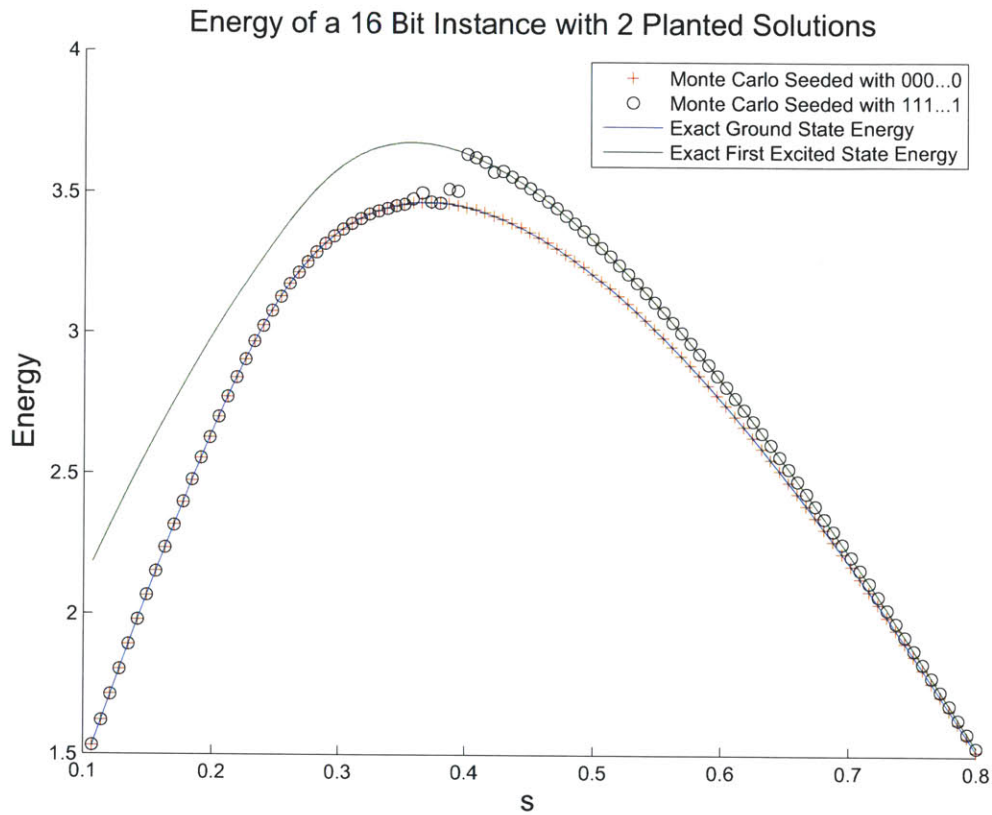
Figure 3-3: The discontinuity in the circle data that occurs near $s = 0.4$ is a Monte Carlo effect that we understand. As can be seen from the exact numerical diagonalization there is no true discontinuity in either the ground state energy or the first excited state energy. For s greater than 0.4 the Monte Carlo simulation is in a metastable equilibrium that corresponds to the first excited state. The true ground state energy at each s is always the lower of the circle and the cross at that value of s.
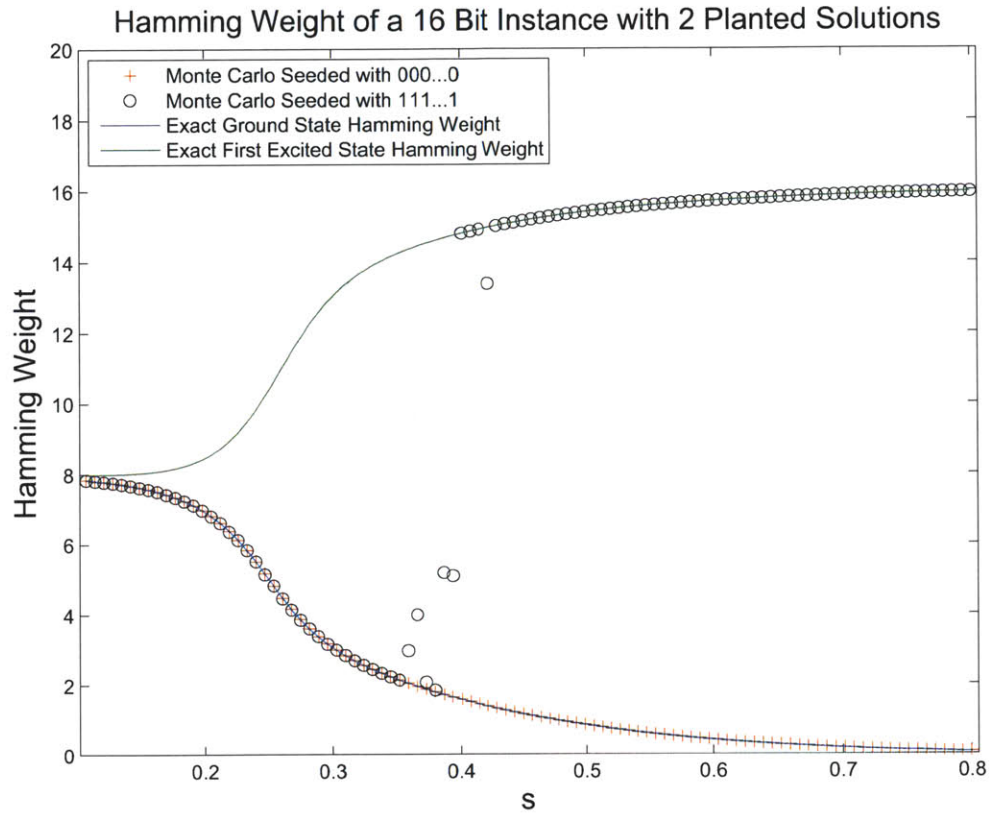
Figure 3-4: Together with Figure 3-3, we see that the discontinuity in the circles appears in data for both the energy and the Hamming weight. This is not indicative of a phase transition in the physical system (as evidenced by the smooth curves computed by exact numerical diagonalization), but is purely a result of the way in which we use the Monte Carlo method.

Energy of a 16 Bit Instance After Adding the Penalty Clause
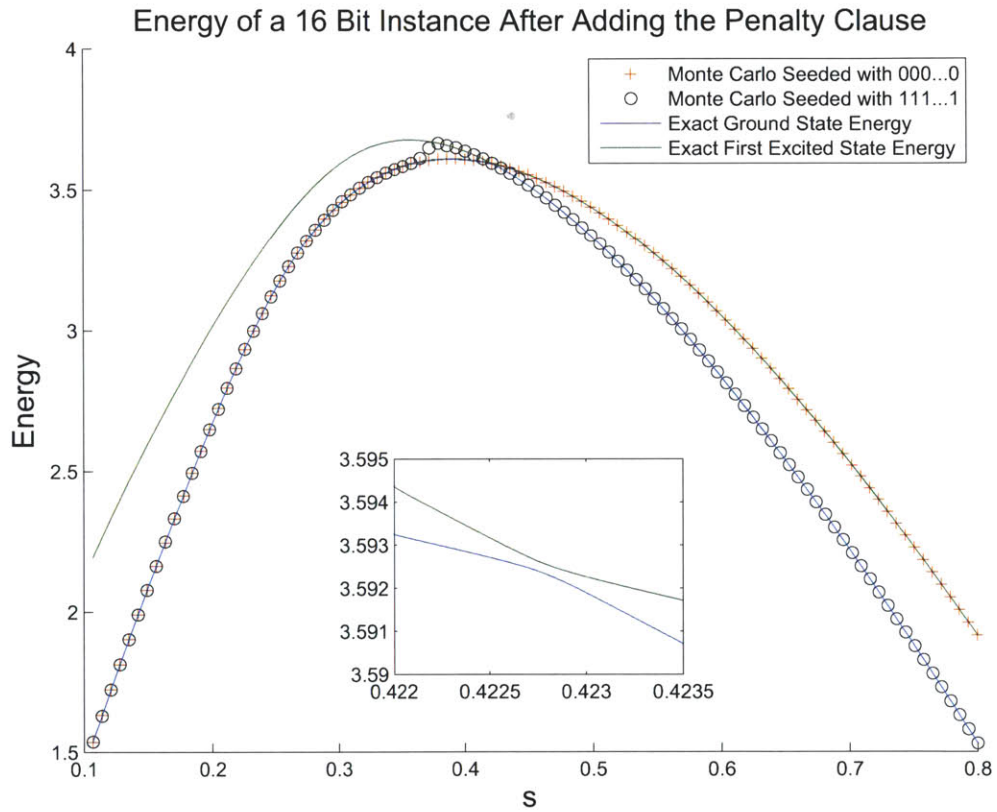
Figure 3-5: After adding the penalty clause we see that the energy levels have an avoided crossing at $s \approx 0.423$. The inset shows exact numerical diagonalization near the avoided crossing where we can resolve a tiny gap. The ground state energy is well approximated by the lower of the circle and cross at each value of s.
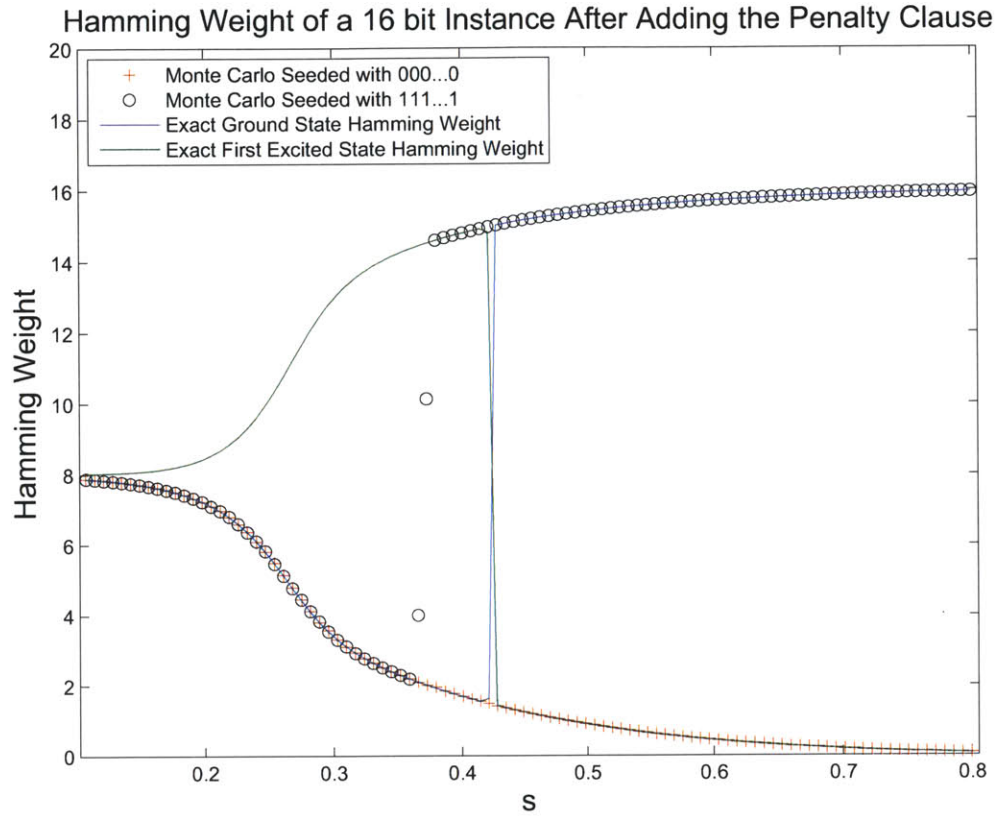
Figure 3-6: In this Figure there is a phase transition which occurs near $s \approx 0.423$. We see from the exact numerical diagonalization that the Hamming weights of the first excited state and ground state undergo abrupt transitions at the point where there is a tiny avoided crossing in Figure 3-5. There is also a jump in the Monte Carlo data plotted with circles that occurs before the avoided crossing: this is a Monte Carlo effect as discussed earlier and has no physical significance. If we look at the Monte Carlo data in Figure 3-5 we conclude that below $s \approx 0.423$ the crosses represent the ground state and after this point the circles represent the ground state. From the current Figure, along with Figure 3-5, we conclude that the Hamming weight of the ground state jumps abruptly at $s^\star \approx 0.423$.
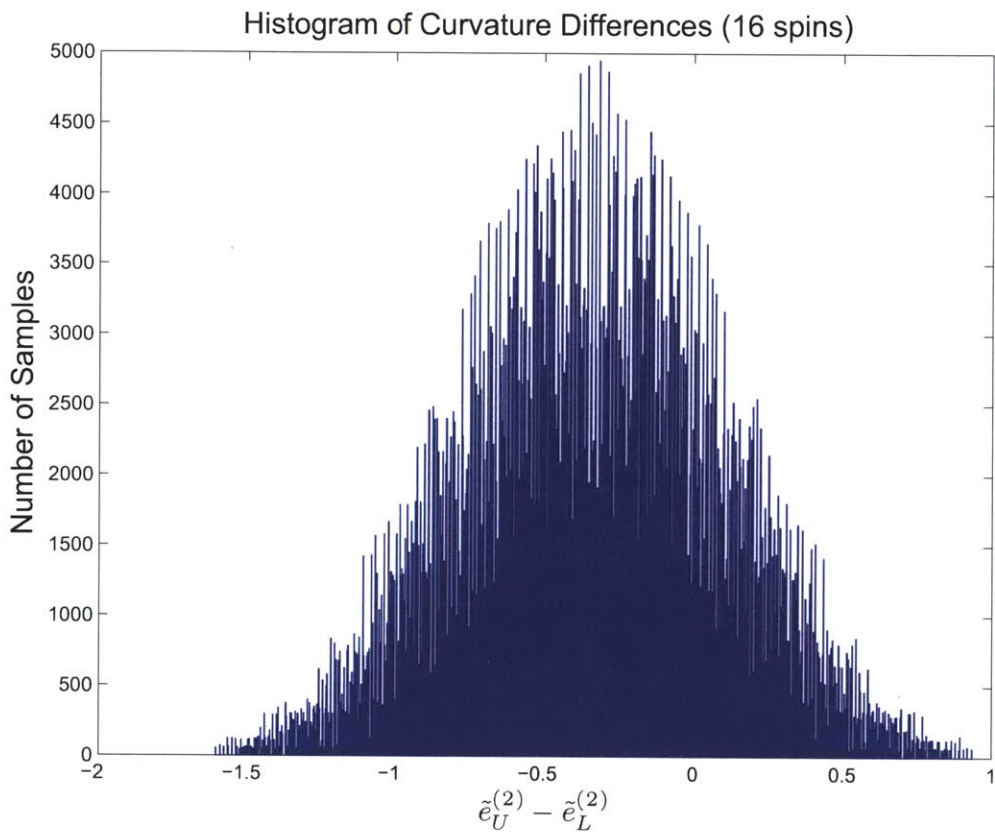
Figure 3-7: $\tilde{e}_U^{(2)} - \tilde{e}_L^{(2)}$ for 1 million different choices of coefficients c_i shows a substantial tail for which $\tilde{e}_U^{(2)} - \tilde{e}_L^{(2)} > 0$. These sets of coefficients correspond to beginning Hamiltonians $\tilde{H}_B$ for which we expect the small gap in Figure 3-5 to be removed.
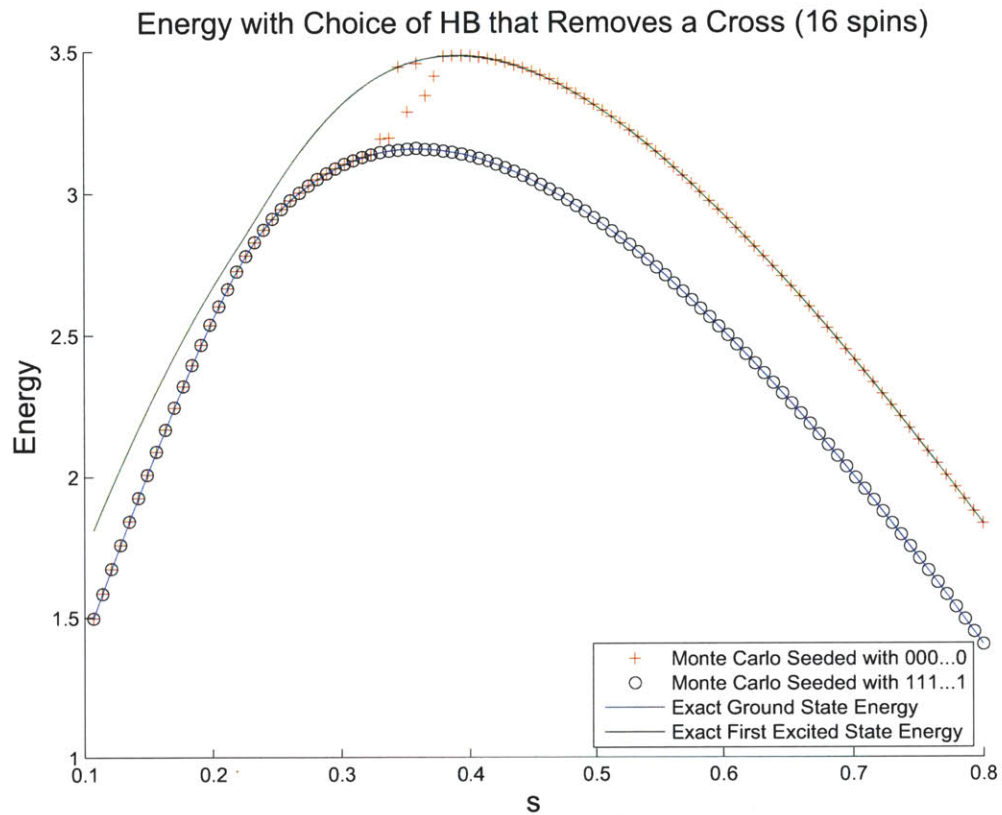
Figure 3-8: A random beginning Hamiltonian removes the crossing seen in Figure 3-5. The problem Hamiltonian is the same as that in Figure 3-5. The circles are always below (or equal to) the crosses for all values of s. This means that the circles track the ground state for all s and we see no sign of a small gap in the Monte Carlo data. This is consistent with the displayed exact numerical diagonalization.
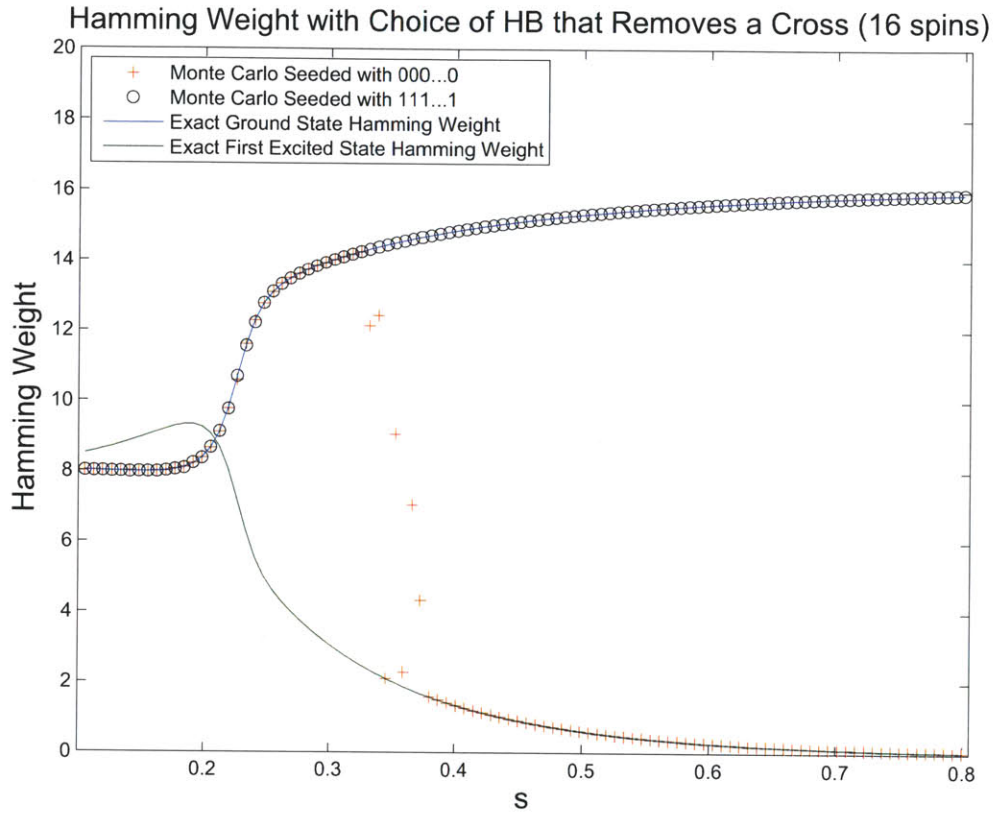
Figure 3-9: From Figure 3-8 we see that the ground state of the Hamiltonian corresponds to the circles for all values of s and the Hamming weight of the circles here goes smoothly to the Hamming weight of the unique satisfying assignment. (The jump in the Monte Carlo data corresponding to the crosses is due to the Monte Carlo effect discussed earlier.)
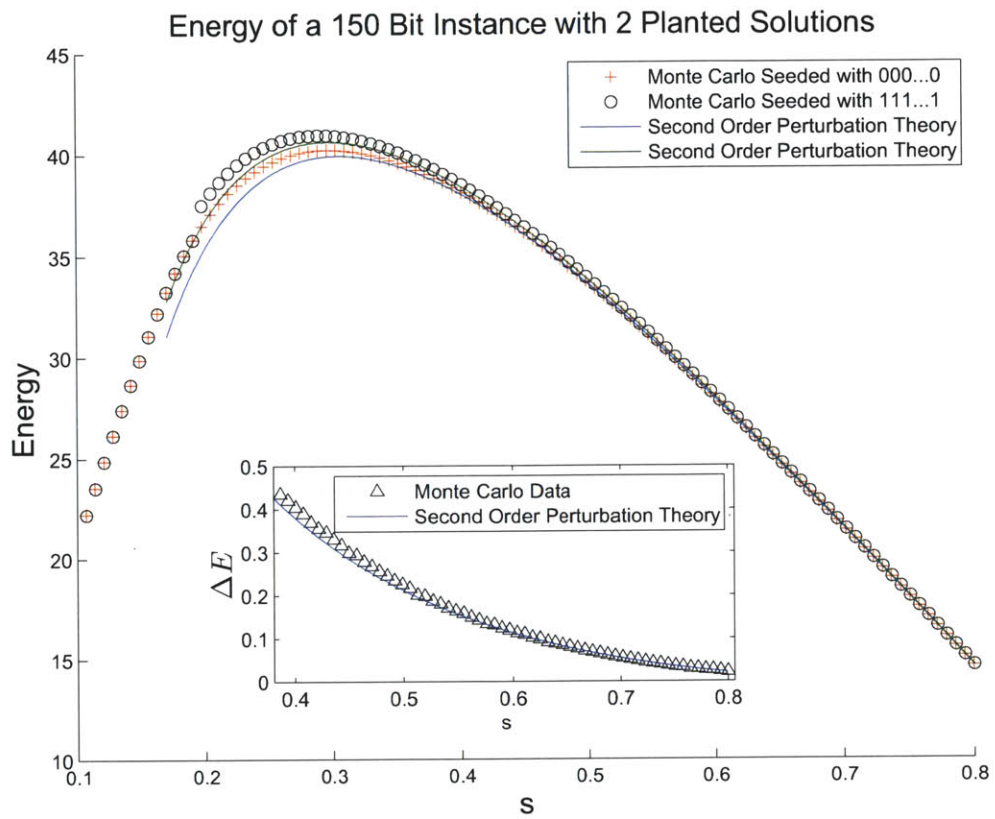
Figure 3-10: The crosses, which represent the Monte Carlo data seeded with 111...1, are below (or equal to) the circle data for all values of s. (This is seen more clearly in the inset which shows the positive difference between the circle and cross values). We conclude that the crosses track the ground state energy which is smoothly varying. The jump in the circle data is a Monte Carlo effect and for s above 0.2 the circles track the first excited state.
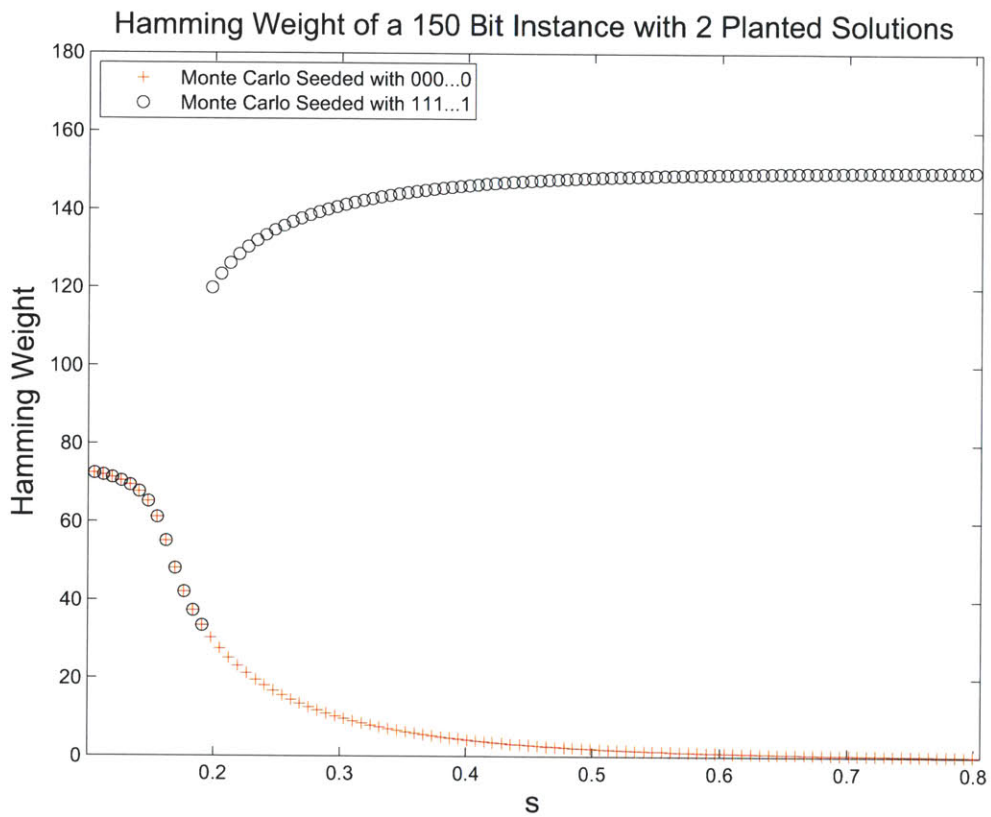
Figure 3-11: Comparing with Figure 3-10 we see that the Hamming weight for the first excited state and the ground state are continous functions of s. We only obtain data for the first excited state for values of s larger than $s \approx 0.2$.
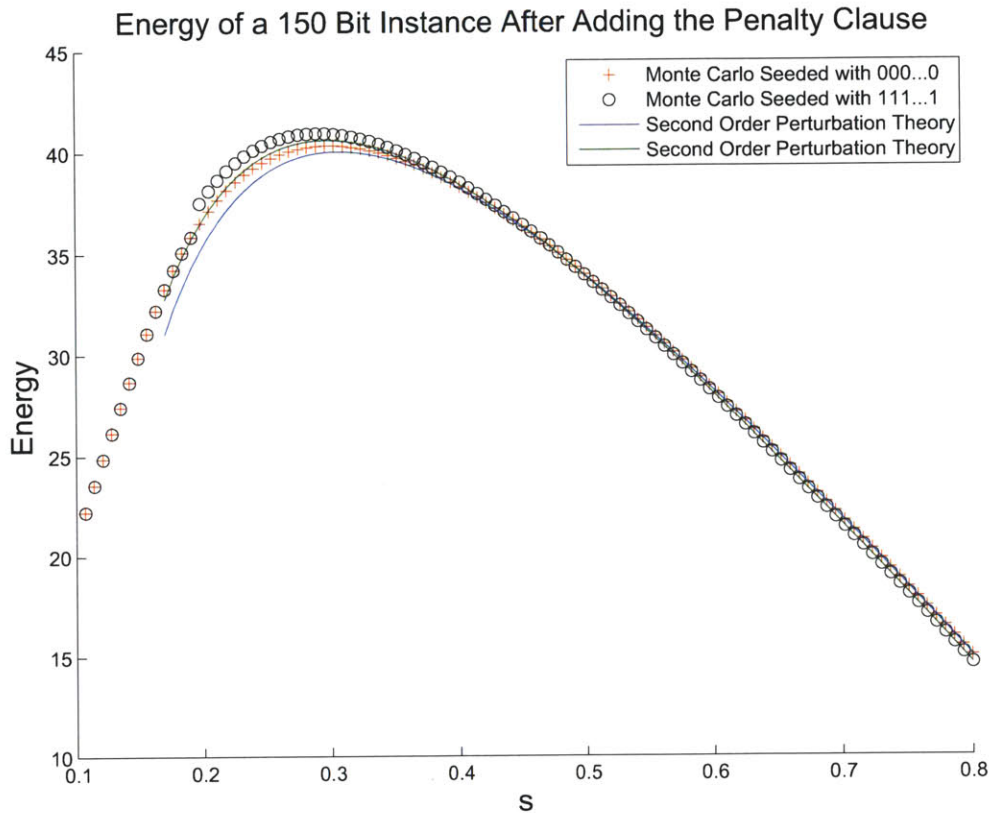
Figure 3-12: Adding the penalty clause makes the cross data go above the circle data at $s \approx 0.49$. This is shown in more detail in Figure 3-13 where we plot the energy difference between the first two levels as a function of s. We interpret this to mean that the Hamiltonian $H(s)$ has a tiny gap at $s^\star \approx 0.49$.
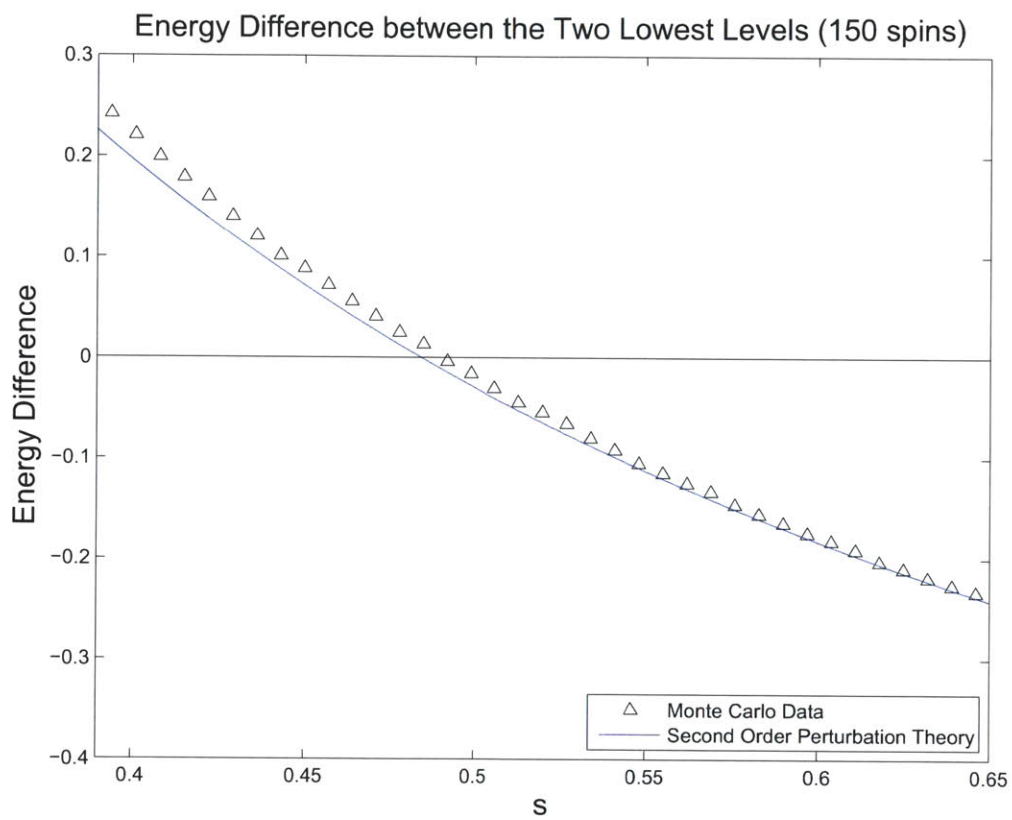
Figure 3-13: The energy difference, circles minus crosses, from Figure 3-12 near the value of s where the difference is 0. Note that second order perturbation theory does quite well in predicting where the difference goes through zero.
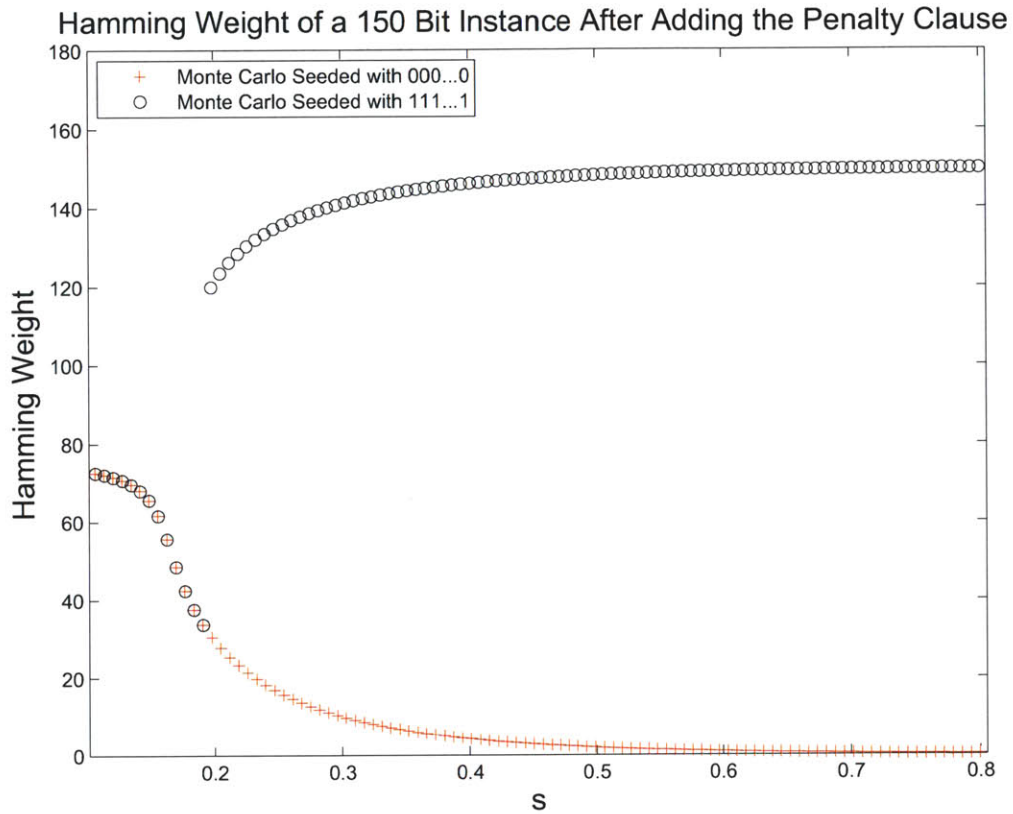
Figure 3-14: Looking at Figure 3-12 we see that the ground state is represented by the crosses to the left of $s \approx 0.49$ and is represented by the circles after this value of s. Tracking the Hamming weight of the ground state, we conclude that it changes abruptly at $s \approx 0.49$.
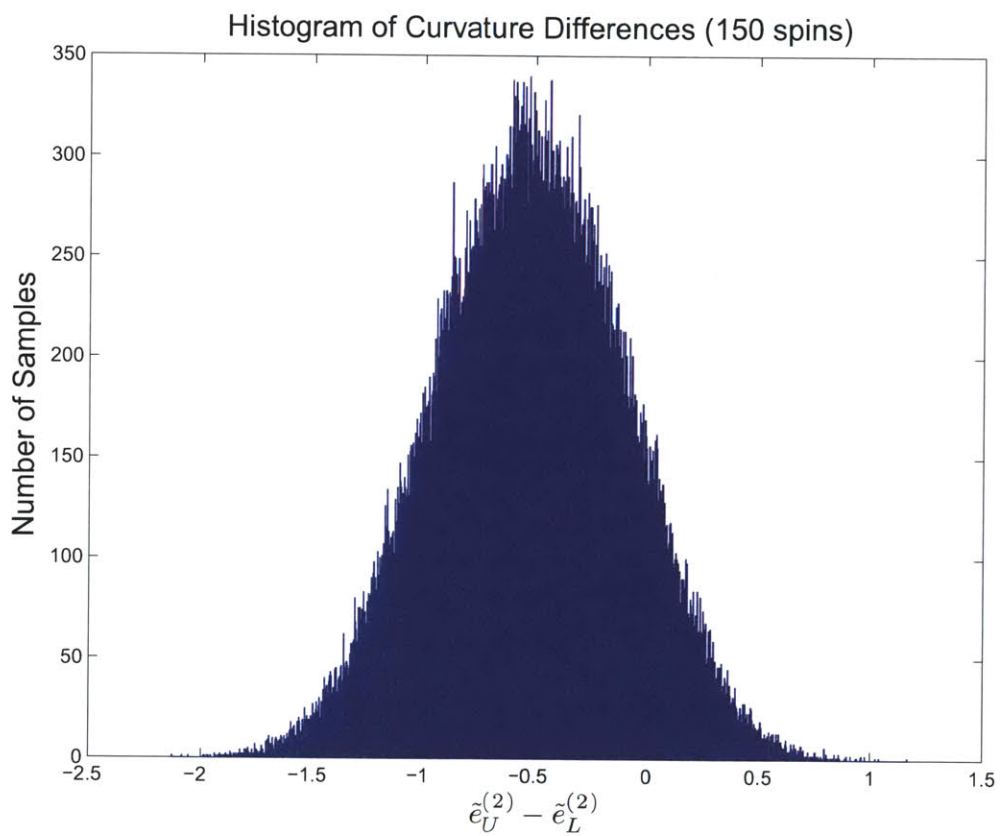
Figure 3-15: $\tilde{e}_U^{(2)} - \tilde{e}_L^{(2)}$ for 100000 choices of coefficients c_i for our 150 spin instance. Note that a good fraction have $\tilde{e}_U^{(2)} - \tilde{e}_L^{(2)} > 0$.

Energy with Choice of HB that Removes a Cross (150 spins)
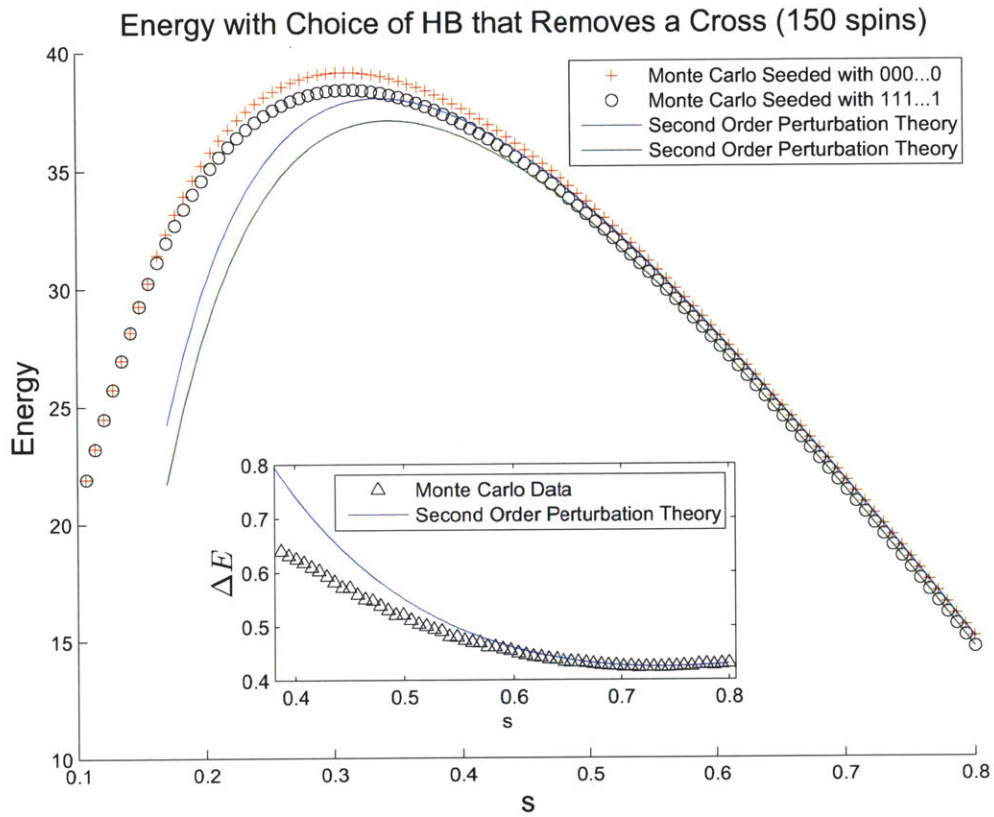
Figure 3-16: A random choice of coefficients such that $\tilde{e}_U^{(2)} - \tilde{e}_L^{(2)} > \frac{1}{2}$ gives rise to an $H(s)$ where there is no longer an avoided crossing. The circles here correspond to the ground state for all s since the cross data is always above (or equal to) the circle data for all s. This can be seen in the inset where we have plotted the energy difference, crosses minus circles. The crosses have a Monte Carlo discontinuity near $s \approx 0.2$, after which they correspond to the first excited state.
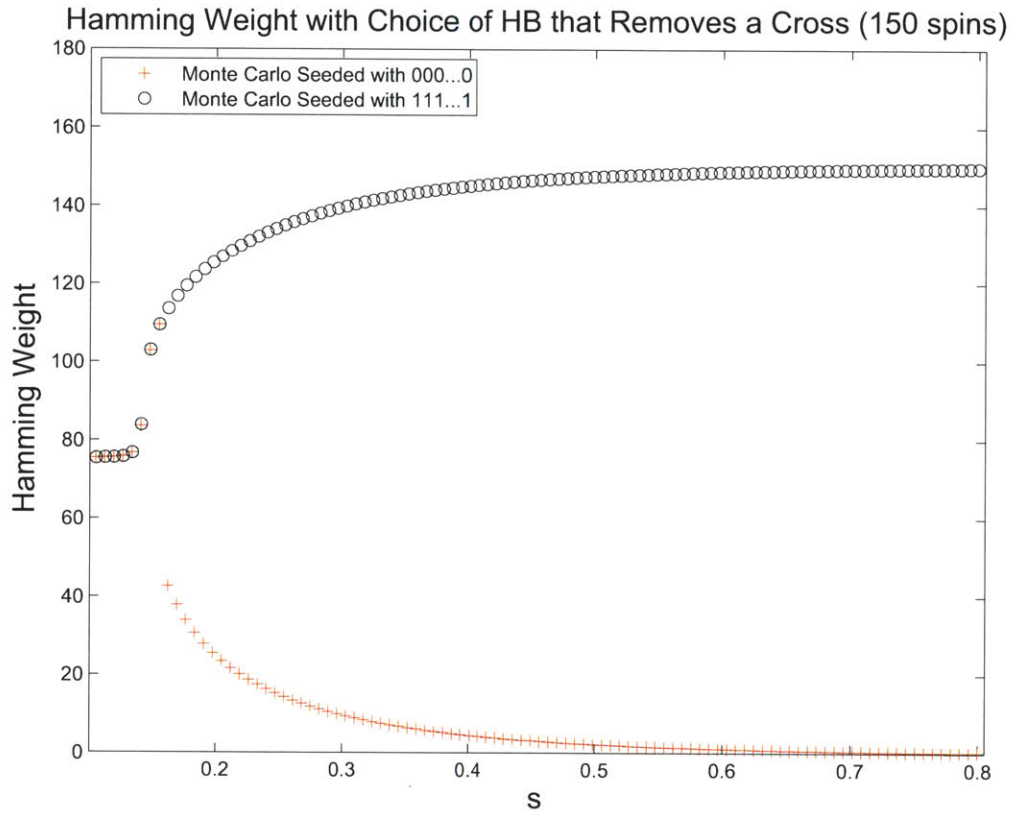
Figure 3-17: Looking at Figure 3-16 we see that the ground state corresponds to the circles for all values of s so we see here that the Hamming weight of the ground state goes smoothly to its final value as s is increased. We take this as further evidence that this choice of $\tilde{H}_B$ would correspond to success for the quantum adiabatic algorithm for this instance.

Chapter 4

Scrambled Hamiltonians and First Order Phase Transitions

In this Chapter we prove results concerning the scrambled Hamiltonians introduced in [16] and reviewed briefly in Chapter 1. We prove a convergence theorem for the ground state energy per spin and we show that the minimum gap is exponentially small. The results of this section are based on reference [13] (which is joint work with Edward Farhi, Sam Gutmann, Jeffrey Goldstone, and Peter Shor) and some of this Chapter is excerpted from that paper.

4.1   Scrambled Hamiltonians

We begin with some conventions and definitions. As in previous Chapters, Let H_P be a Hamiltonian which acts on the Hilbert space of n spin $\frac{1}{2}$ particles, and which is diagonal in the Pauli z basis, with matrix elements $H_P(z) = \langle z|H_P|z \rangle$. We will assume that $n > 1$. We assume that H_P satisfies the following three basic properties:

1. There is at least one n-bit string z_0 such that

$$H_P(z_0) = 0.$$

2. For any n-bit string z,
$$0 \le H_P(z) \le n.$$

3. There is at least one bit string z with $H_P(z) > 0$.

Note that properties 1 and 2 can always be achieved by adding a constant and then rescaling by a multiplicative factor. Property 3 only requires that H_p is not the zero matrix.

We also define

$$\alpha \; = \; \frac{1}{n}\left(\frac{1}{2^n}\sum_z H_P(z)\right) \tag{4.1}$$

$$\delta \; = \; \left[\frac{1}{2^n}\sum_z (H_P(z) - n\alpha)^2\right]^{\frac{1}{2}} \tag{4.2}$$

so αn is the average energy $H_P(z)$ of a randomly chosen bit string z and δ is the standard deviation of the random variable $H_P(z)$ (note that $\alpha > 0$ for any H_P satisfying the properties above).

For any permutation matrix π on 2^n elements, write

$$H_\pi(s) = \left(\frac{1-s}{2}\right)\sum_{i=1}^n \left(\frac{1-\sigma_x^i}{2}\right) + sH_{P,\pi} \tag{4.3}$$

where

$$H_{P,\pi} \; = \; \pi H_P \pi^{-1}$$
$$= \; \sum_{z=0}^{2^n-1} H_P(\pi^{-1}(z))|z\rangle\langle z|$$

The theorems we will prove in this Chapter will apply to the Hamiltonian $H_\pi(s)$ with π chosen uniformly at random from the $2^n!$ permutations of n-bit strings.

4.2   Results

Define

$$e_\alpha(s) = \begin{cases} \alpha s & , \; 0 \le s \le s^\star \\ \left(\frac{1-s}{2}\right) & , \; s^\star < s \le 1 \end{cases}$$

where

$$s^\star = \frac{1}{2\alpha + 1}.$$

Write $E_\pi(s)$ for the ground state energy of the Hamiltonian $H_\pi(s)$. The following theorem proves that $\frac{E_\pi(s)}{n}$ converges to $e_\alpha(s)$, under a condition on the standard deviation δ of the spectrum of H_P.

86

Theorem 2. *Suppose*

$$\frac{\delta}{n} = \frac{1}{n}\left[\frac{1}{2^n}\sum_z \left(H_P(z) - n\alpha\right)^2\right]^{\frac{1}{2}} \le \frac{c}{n^p}$$

where $c \ge 1$ and $p > 0$. Let π be a random permutation.
Then

$$\frac{E_\pi(s)}{n} \le e_\alpha(s) \ \text{ for } 0 \le s \le 1$$

and

$$\Pr\left[\frac{E_\pi(s)}{n} \ge e_\alpha(s)\left(1 - \frac{14c}{\alpha}\frac{1}{n^k}\right) \ \text{ for } 0 \le s \le 1\right] \ge 1 - e^{-n(1 - \ln 2)}$$

where

$$k = \begin{cases} \frac{p}{3} & , \ \text{if } 0 < p < \frac{1}{2} \\ \frac{1}{6} & , \ \text{if } \frac{1}{2} \le p \end{cases}$$

For each π let

$$g_\pi = \min_{0 \le s \le 1} \gamma_\pi(s).$$

where $\gamma_\pi(s)$ is the eigenvalue gap between the two lowest energy levels of $H_\pi(s)$. Our second theorem says that g_π is (with high probability) exponentially small as a function of n.

Theorem 3. *Let π be a random permutation. Suppose the ground state of H_P is unique, with an eigenvalue gap g_P. Then*

$$Pr\left[g_\pi \le n^2 2^{-\frac{n}{6}}\left(1 + \frac{1}{g_P^2}\right)^{\frac{1}{3}}\right] \to 1 \ as \ n \to \infty.$$

The proofs of theorems 2 and 3 are found in section 4.5.

4.3 Variational Ansatzes for the Ground State of $H_\pi(s)$

The proof of theorem 2 uses variational upper and lower bounds on the ground state energy. The variational upper bound we use is the standard elementary upper bound.

We also use Theorem 1 from Chapter 1 which is a lower bound on the ground state energy for stoquastic Hamiltonians.

In this section we discuss the states which are used as variational ansatzes in these bounds, as they give insight into the ground state of $H_\pi(s)$.

The upper bound

$$E_\pi(s) \leq ne_\alpha(s) \quad 0{\leq}s{\leq}1$$

is easy to prove. It can be obtained by using the state $|\psi(s)\rangle$ defined by

$$|\psi(s)\rangle = \begin{cases} \frac{1}{\sqrt{2^n}} \sum_z |z\rangle, & \text{for } 0 \leq s \leq s^\star \\ |w\rangle & , \text{for } s^\star < s \leq 1 \end{cases}$$

where $|w\rangle$ is a zero energy state for $H_{P,\pi}$(note that at least one such w exists by our conditions on the Hamiltonian H_P). This state satisfies

$$\langle\psi(s)|H_\pi(s)|\psi(s)\rangle = ne_\alpha(s)$$

which furnishes the desired bound.

Unfortunately the state $|\psi(s)\rangle$ does not give an acceptable lower bound on the ground state energy when used as a trial wavefunction in Theorem 1. The state $|\psi(s)\rangle$ does not even satisfy the requirements of the theorem for $s^\star < s \leq 1$ (since it has zero amplitude on all but one of the computational basis states). For $0 \leq s \leq s^\star$ we can try using the theorem applied to this state, which gives the bound

$$\begin{aligned} E_\pi(s) &\geq \min_z \frac{\langle z|H|\psi(s)\rangle}{\langle z|\psi(s)\rangle} \quad \text{for } 0 \leq s \leq s^\star \\ &= \min_z \left(sH_{P,\pi}(z) \right) \quad \text{for } 0 \leq s \leq s^\star \\ &= 0 \text{ for } 0 \leq s \leq s^\star. \end{aligned}$$

A better variational ansatz will be needed to prove theorem 2. The ansatz which we use is inspired by the novel Quantum Monte Carlo method introduced in Chapter 2. Recall that in Chapter 2 we encountered the operator

$$A(E) = \left(\frac{-\lambda V}{H_0 - E} \right)$$

which has the property that, if the parameter E is taken to be the ground state energy of the stoquastic Hamiltonian $H = H_0 + \lambda V$, then $A(E)^r$ projects onto the ground state for sufficiently large r[1]. For the problem at hand, we believe that we know the ground state energy per spin. In particular, we are aiming to show that $E_\pi(s) \approx ne_\alpha(s)$. This suggests that we try a variational ansatz

$$A(ne_\alpha(s))^r|\phi\rangle$$

---

[1]When $V = -\sum_i \sigma_x^i$ this operator also projects onto one other state–but we ignore this detail for the purposes of this motivational discussion

where r is chosen to be "large enough", and $|\phi\rangle$ is a suitably chosen state which we believe is already close to the ground state. By choosing $|\phi\rangle = \frac{1}{\sqrt{2^n}} \sum_z |z\rangle$ and $r = 2$ we are able to obtain a good lower bound for $0 \leq s < s^\star$. The bound we get is given in the following lemma

Lemma 2. *Fix $s \in [0, s^\star)$, where $s^\star = \frac{1}{2\alpha+1}$. If π is chosen uniformly at random, then, with probability at least $1 - e^{-n(1-\ln 2)}$*

$$\frac{E_\pi(s)}{n} \geq \alpha s - s \left(\frac{1 - (2\alpha - 1)s}{1 - (2\alpha + 1)s}\right)^2 \left[\frac{1 + \delta}{n} + \frac{1}{\sqrt{n}}\right]$$

This lemma, combined with the variational upper bound discussed above, shows that, for each $0 \leq s < s^\star$, the ground state energy $\frac{E_\pi(s)}{m}$ differs from the curve $e_\alpha(s)$ by terms which vanish in the large n limit if $\frac{\delta}{n} \to 0$ as $n \to \infty$.

This lemma does not prove a lower bound for $s \geq s^\star$, which we will need in order to prove theorem 2. In order to lower bound the ground state energy for all $s \in [0, 1]$ we use the concavity property of the ground state energy –we don't ever need to obtain a good variational ansatz for $s^\star \leq s \leq 1$. To see how this works, let's suppose that

$$\frac{1}{n} E_\pi(s^\star) = e_\alpha(s^\star) [1 - \epsilon] \tag{4.4}$$

for some $0 < \epsilon < 1$. It is easy to show, using the linearity of the Hamiltonian $H_\pi(s)$, that for any $s_1 < s_2$,

$$E_\pi(s) \geq \left(\frac{s - s_1}{s_2 - s_1}\right) E_\pi(s_2) + \left(\frac{s_2 - s}{s_2 - s_1}\right) E_\pi(s_1) \quad \text{for } s_1 \leq s \leq s_2.$$

This equation simply says that the true ground state energy curve lies above the straight line connecting any two points on the curve. Applying this to the two points $s_1 = 0$ and $s_2 = s^\star$ and then the two points $s_1 = s^\star$ and $s_2 = 1$ we get that $E_\pi(s)$ lies above $e_\alpha(s)(1 - \epsilon)$ for all $s \in [0, 1]$.

In summary, our proof technique is as follows. We first obtain the upper and lower bounds for $s < s^\star$ using the variational ansatzes discussed above. Using these results we get a bound on $E_\pi(s^\star)$ of the form 4.4 where $\epsilon \to 0$ as $n \to \infty$. We then use this fact to bound the curve for all s. The full proof of theorem 2 is given in section 4.5.

4.4 Discussion and Conclusions

Let's see what theorems 2 and 3 say about the scrambled Hamming weight problem which was introduced in [16] and reviewed in Chapter 1. Recall the (unscrambled)

problem Hamiltonian is

$$H_P = \sum_{i=1}^{n} \left(\frac{1 - \sigma_z^i}{2} \right).$$

For this problem Hamiltonian the mean energy is $\alpha = \frac{1}{2}$ (from equation 4.1) and the standard deviation δ (from equation 4.2) is the standard deviation of a Binomial $(n, \frac{1}{2})$ random variable, so $\delta = \frac{\sqrt{n}}{2}$. So in theorem 2 we can set $c = 1$ and $p = \frac{1}{2}$ to get that

$$\Pr\left[\frac{E_\pi(s)}{n} \geq e_{\frac{1}{2}}(s) \left(1 - \frac{28}{n^{\frac{1}{6}}} \right) \text{ for } 0 \leq s \leq 1 \right] \geq 1 - e^{-n(1 - \ln 2)}$$

which shows (combined with the upper bound from the theorem) that, for large n, with high probability the energy per spin will be close to the two straight lines $e_{\frac{1}{2}}(s)$. This confirms the limiting behaviour which was hypothesized in reference [16] and which is seen in Figure 1-2. A better rate of convergence for this problem is proven in reference [13] using the same variational method but with a different ansatz for the lower bound.

Since the gap g_P at $s = 1$ is equal to 1 for this problem Hamiltonian, theorem 3 says that the minimum gap g_π is exponentially small. It does not tell us where this minimum gap occurs as a function of s. From Figure 1-2 in Chapter 1 we can see that it is at $s = \frac{1}{2}$.

If we believe our variational ansatzes are close to the true ground state then for $s < \frac{1}{2}$, the ground state is close to the ground state at $s = 0$,

$$|x = 0\rangle = \frac{1}{\sqrt{2^n}} \sum_{z=0}^{2^n - 1} |z\rangle$$

which in the z basis is completely delocalized. And for $s > \frac{1}{2}$ the ground state is close to the ground state at $s = 1$ which is

$$|z = \pi(0)\rangle$$

corresponding to one z string, that is, a fully localized state.

The small gap at $s = \frac{1}{2}$ for the scrambled Hamming weight problem is associated with a "delocalized to localized" first order phase transition. The recent work on perturbative crosses[2, 3, 18, 39, 5] suggests that a "localized to localized" ground state transition can also lead to an exponentially small gap. Unlike the situation in this model, these ground state transitions occur for $s \to 1$ as $n \to \infty$. There are two key features of a model that exhibits perturbative crosses. First, for any string a single bit flip changes the cost function by $O(1)$. The other feature is that there are very disparate bit strings with low cost. The scrambled models that we study here have the second feature but not the first. In the case of the scrambled Hamming weight problem it is shown in [13] that perturbative crosses are not present.

In what ways are scrambled problem Hamiltonians different from random clause based decision problems such as 3SAT? There is the issue of perturbative crosses which

we have just mentioned. We also do not know if random instances of satisfiability have a "delocalized to localized" first order phase transition like the fully unstructured example shown in Figure 1-2 or have no "delocalized to localized" phase transition like the bit structured example shown in Figure 1-1. We hope that future work is able to better characterize the ways in which clause structure helps or hinders the quantum adiabatic algorithm for random optimization problems.

4.4.1 Connection to Previous work on Random Energy Models

We note that the scrambled Hamiltonians we consider in this Chapter are similar to the quantum random energy models which are Hamiltonians of the form

$$H(s) = (1 - s) \sum_{i=1}^{n} \left(\frac{1 - \sigma_x^i}{2} \right) + s \sum_{z=0}^{2^n - 1} E(z) |z\rangle \langle z|$$

where the "on site energies" $E(z)$ are independent and identically distributed random variables. Some of what we have proven is already known for random energy models[53, 54, 36]. For example, in [36, 35], Jörg et. al. use perturbation theory to show the existence of a first order phase transition in a random energy model and also discuss the connection to localization. Presilla and Ostilli have studied quantum random energy models in [54, 53] and have shown the two straight lines phenomenon but they seem to have made some assumptions in their proof. For example they "assume that the probability for the ground-state energy to assume a given value becomes infinitely peaked around its average E0"[53] .

In contrast, we have used variational methods which prove upper and lower bounds on the ground state energy. To prove that the gap is exponentially small in n, we use the information theoretic result[16] that no efficient quantum algorithm exists for locating the minimum of a scrambled cost function.

4.5 Proofs

Proof of Lemma 2

The main technical tools which are used in the proof of lemma 2 are the variational lower bound from Theorem 1 as well as Hoeffding's Inequality, which we reproduce here for completeness. Note that in the standard version of Hoeffding's inequality (Theorem 1 from [30]) the random variables X_i are taken to be independent and identically distributed. The version below, in which the random variables are drawn without replacement from a population, is proven in section 6 of [30].

Hoeffding's Theorem (from [30]). *Let $X_1, X_2, ..., X_n$ be randomly drawn without replacement from a population of N elements $\{y_i : 0 \leq i \leq N\}$ (here $N > n$). Suppose*

$$0 \leq y_i \leq 1 \quad \text{for all } 0 \leq i \leq N$$

and

$$\mu = \frac{1}{N} \sum_{i=1}^{N} y_i.$$

Then

$$Pr\left[\frac{1}{n} \sum_{j=1}^{n} X_j - \mu \geq t\right] \leq e^{-2nt^2}.$$

To prove lemma 2 we use the variational lower bound from theorem 1, using the variational ansatz $|\phi_\pi\rangle$ defined by

$$\langle z|\phi_\pi\rangle = \frac{1}{sH_{P,\pi}(z) + \left(\frac{1-s}{2}\right)n - \alpha ns}\left[\sum_{j=1}^{n}\left(\frac{1}{sH_{P,\pi}(z \oplus \hat{e}_j) + \left(\frac{1-s}{2}\right)n - \alpha ns}\right)\right].$$

We can use this to get a variational lower bound when $0 \leq s < s^\star$ because for this range of s the ansatz satisfies $\langle z|\phi_\pi\rangle > 0$ for all basis states $|z\rangle$.

Note that, up to normalization,

$$|\phi_\pi\rangle = A(ne_\alpha(s))^2 \left[\frac{1}{\sqrt{2^n}} \sum_z |z\rangle\right]$$

where

$$A(E) = \frac{1}{sH_{P,\pi} + \left(\frac{1-s}{2}\right)n - E}\left(\sum_{i=1}^{n} \frac{\sigma_x^i}{2}(1-s)\right).$$

Since $H_{P,\pi}(z) \leq n$, for any $0 \leq s < s^\star$,

$$
\begin{aligned}
|\langle z|\phi_\pi\rangle| &\geq \frac{1}{sn + \left(\frac{1-s}{2}\right)n - \alpha ns}\left[n \cdot \left(\frac{1}{sn + \left(\frac{1-s}{2}\right)n - \alpha ns}\right)\right] \\
&= \frac{1}{n}\left(\frac{1}{\frac{1}{2} - (\alpha - \frac{1}{2})s}\right)^2 \\
&= \frac{1}{n}\left(\frac{2}{1 - (2\alpha - 1)s}\right)^2.
\end{aligned}
\tag{4.5}
$$

We also bound

$$|\langle z|(H_\pi(s) - \alpha ns)|\phi_\pi\rangle|$$

$$= \left| \sum_{j=1}^{n} \left(\frac{1}{sH_{P,\pi}(z \oplus \hat{e}_j) + \left(\frac{1-s}{2}\right)n - \alpha ns} \right) - \left(\frac{1-s}{2}\right) \sum_{k=1}^{n} \langle z \oplus \hat{e}_k|\phi_\pi\rangle \right|$$

$$= \left| \sum_{j=1}^{n} \sum_{k=1}^{n} \left(\frac{1}{sH_{P,\pi}(z \oplus \hat{e}_j) + \left(\frac{1-s}{2}\right)n - \alpha ns} \right) \right.$$

$$\left. \left(\frac{1}{n} - \left(\frac{1-s}{2}\right) \frac{1}{sH_{P,\pi}(z \oplus \hat{e}_j \oplus \hat{e}_k) + \left(\frac{1-s}{2}\right)n - \alpha ns} \right) \right|$$

$$\leq \left[\left(\frac{1}{\left(\frac{1-s}{2}\right)n - \alpha ns} \right) \right.$$

$$\left. \sum_{j=1}^{n} \sum_{k=1}^{n} \left| \frac{1}{n} - \left(\frac{1-s}{2}\right) \frac{1}{sH_{P,\pi}(z \oplus \hat{e}_j \oplus \hat{e}_k) + \left(\frac{1-s}{2}\right)n - \alpha ns} \right| \right] \qquad (4.6)$$

where the last line follows since $H_{P,\pi}(z) \geq 0$. We can simplify the above expression by noting that (for $0 \leq s < s^\star$)

$$\left| \frac{1}{n} - \left(\frac{1-s}{2}\right) \frac{1}{sH_{P,\pi}(z \oplus \hat{e}_j \oplus \hat{e}_k) + \left(\frac{1-s}{2}\right)n - \alpha ns} \right|$$

$$= \left(\frac{s}{n}\right) \frac{1}{sH_{P,\pi}(z \oplus \hat{e}_j \oplus \hat{e}_k) + \left(\frac{1-s}{2}\right)n - \alpha ns} |H_{P,\pi}(z \oplus \hat{e}_j \oplus \hat{e}_k) - \alpha n|$$

$$\leq \left(\frac{s}{n^2}\right) \left(\frac{2}{1 - (2\alpha + 1)s}\right) |H_{P,\pi}(z \oplus \hat{e}_j \oplus \hat{e}_k) - \alpha n|.$$

Plugging this expression into equation 4.6 and using equation 4.5 we get

$$\left| \frac{\langle z|(H_\pi(s) - \alpha ns)|\phi_\pi\rangle}{\langle z|\phi_\pi\rangle} \right| \leq \left(\frac{s}{n^2}\right) \left(\frac{1 - (2\alpha - 1)s}{1 - (2\alpha + 1)s}\right)^2 \sum_{j=1}^{n} \sum_{k=1}^{n} |H_{P,\pi}(z \oplus \hat{e}_j \oplus \hat{e}_k) - \alpha n|.$$

We separate out the terms with $j = k$ and bound them using $n \cdot |H_{P,\pi}(z) - \alpha n| \leq n^2$,

which follows since $0 < \alpha \leq 1$ and $0 \leq H_{P,\pi}(z) \leq n$. Then

$$\left| \frac{\langle z | (H_\pi(s) - \alpha n s) | \phi_\pi \rangle}{\langle z | \phi_\pi \rangle} \right|$$

$$\leq \left(\frac{1 - (2\alpha - 1)s}{1 - (2\alpha + 1)s} \right)^2 \left[s + 2 \left(\frac{s}{n^2} \right) \sum_{j<k} |H_{P,\pi}(z \oplus \hat{e}_j \oplus \hat{e}_k) - \alpha n| \right]$$

$$= \left(\frac{1 - (2\alpha - 1)s}{1 - (2\alpha + 1)s} \right)^2 \left[s + 2 \left(\frac{s}{n^2} \right) \sum_{j<k} \mathbb{E} \left(|H_{P,\pi}(z \oplus \hat{e}_j \oplus \hat{e}_k) - \alpha n| \right) + \right.$$

$$\left. 2 \left(\frac{s}{n^2} \right) \sum_{j<k} \left(|H_{P,\pi}(z \oplus \hat{e}_j \oplus \hat{e}_k) - \alpha n| - \mathbb{E} \left(|H_{P,\pi}(z \oplus \hat{e}_j \oplus \hat{e}_k) - \alpha n| \right) \right) \right]$$

Now for any random variable X, $Var(|X|) = \mathbb{E}(|X|^2) - [\mathbb{E}(|X|)]^2 \geq 0$. Applying this to the case at hand we get

$$\mathbb{E} \left(|H_{P,\pi}(z \oplus \hat{e}_j \oplus \hat{e}_k) - \alpha n| \right) \leq \left[\mathbb{E} \left(|H_{P,\pi}(z \oplus \hat{e}_j \oplus \hat{e}_k) - \alpha n|^2 \right) \right]^{\frac{1}{2}}$$
$$= \delta$$

So

$$\left| \frac{\langle z | (H_\pi(s) - \alpha n s) | \phi_\pi \rangle}{\langle z | \phi_\pi \rangle} \right|$$

$$\leq \left(\frac{1 - (2\alpha - 1)s}{1 - (2\alpha + 1)s} \right)^2 \left[s + 2 \left(\frac{s}{n^2} \right) \left(\frac{n^2 - n}{2} \right) \delta \right.$$

$$\left. + 2 \left(\frac{s}{n^2} \right) \sum_{j<k} \left(|H_{P,\pi}(z \oplus \hat{e}_j \oplus \hat{e}_k) - \alpha n| - \mathbb{E} \left(|H_{P,\pi}(z \oplus \hat{e}_j \oplus \hat{e}_k) - \alpha n| \right) \right) \right].$$

$$\leq \left(\frac{1 - (2\alpha - 1)s}{1 - (2\alpha + 1)s} \right)^2 \left[s \left(1 + \delta \right) + s(n - 1) \left(\frac{2}{n(n-1)} \right) \sum_{j<k} \left(Y_{jk}^z - \mathbb{E} \left(Y_{jk}^z \right) \right) \right].$$

where

$$Y_{jk}^z = \left| \frac{1}{n} H_{P,\pi}(z \oplus \hat{e}_j \oplus \hat{e}_k) - \alpha \right|.$$

Note that

$$0 \leq Y_{jk}^z \leq 1.$$

The random variables Y_{jk}^z for fixed z and $j < k$ are not independent. Their joint probability distribution (over the choice of π) is the distribution governing a sample of $\frac{n(n-1)}{2}$ elements drawn from the population of 2^n values $\{ |\frac{1}{n} H_P(x) - \alpha|, x = 1, ..., 2^n \}$, without replacement. We can therefore use Hoeffding's inequality to bound the probability that the sum $\frac{2}{n(n-1)} \sum_{j<k} Y_{jk}^z$ differs substantially from its mean value. Ho-

effding's inequality says

$$\Pr\left[\frac{2}{n(n-1)}\sum_{j<k}\left(Y_{jk}-\mathbb{E}\left[Y_{jk}^z\right]\right)\geq t\right]\leq e^{-2\left(\frac{n^2-n}{2}\right)t^2}.$$

Now take $t=\frac{1}{\sqrt{n-1}}$ to get

$$\Pr\left[\frac{2}{n(n-1)}\sum_{j<k}\left(Y_{jk}-\mathbb{E}\left[Y_{jk}^z\right]\right)\geq\frac{1}{\sqrt{n-1}}\right]\leq e^{-n}.$$

So , fixing $z\in\{1,...,2^n\}$ and $0\leq s<s^\star$, with probability at least $1-e^{-n}$ we have

$$\left|\frac{\langle z|(H_\pi(s)-\alpha ns)|\phi_\pi\rangle}{\langle z|\phi_\pi\rangle}\right|\leq\left(\frac{1-(2\alpha-1)s}{1-(2\alpha+1)s}\right)^2\left[s(1+\delta)+s(n-1)\frac{1}{\sqrt{n-1}}\right]$$
$$\leq\left(\frac{1-(2\alpha-1)s}{1-(2\alpha+1)s}\right)^2\left[s\left(1+\delta+\sqrt{n}\right)\right].$$

We now apply a union bound over the z to bound the minimum of the LHS over all z

$$\Pr\left[\left|\min_z\frac{\langle z|(H_\pi(s)-\alpha ns)|\phi_\pi\rangle}{\langle z|\phi_\pi\rangle}\right|\leq\left(\frac{1-(2\alpha-1)s}{1-(2\alpha+1)s}\right)^2\left[s\left(1+\delta+\sqrt{n}\right)\right]\right]$$
$$\geq\ \Pr\left[\left|\frac{\langle z|(H_\pi(s)-\alpha ns)|\phi_\pi\rangle}{\langle z|\phi_\pi\rangle}\right|\leq\right.$$
$$\left.\left(\frac{1-(2\alpha-1)s}{1-(2\alpha+1)s}\right)^2\left[s\left(1+\delta+\sqrt{n}\right)\right]\text{ for all }z\in\{1,...,2^n\}\right]$$
$$=\ 1-\Pr\left[\left|\frac{\langle z|(H_\pi(s)-\alpha ns)|\phi_\pi\rangle}{\langle z|\phi_\pi\rangle}\right|>\right.$$
$$\left.\left(\frac{1-(2\alpha-1)s}{1-(2\alpha+1)s}\right)^2\left[s\left(1+\delta+\sqrt{n}\right)\right]\text{ for at least one }z\in\{1,...,2^n\}\right]$$
$$\geq\ 1-\sum_z\Pr\left[\left|\frac{\langle z|(H_\pi(s)-\alpha ns)|\phi_\pi\rangle}{\langle z|\phi_\pi\rangle}\right|>\left(\frac{1-(2\alpha-1)s}{1-(2\alpha+1)s}\right)^2\left[s\left(1+\delta+\sqrt{n}\right)\right]\right]$$
$$\geq\ 1-2^n e^{-n}. \tag{4.7}$$

Now the variational lower bound says

$$E_\pi(s)-\alpha ns\geq\min_z\frac{\langle z|(H_\pi(s)-\alpha ns)|\phi_\pi\rangle}{\langle z|\phi_\pi\rangle},$$

So

$$E_\pi(s)\geq\alpha ns-\left|\min_z\frac{\langle z|(H_\pi(s)-\alpha ns)|\phi_\pi\rangle}{\langle z|\phi_\pi\rangle}\right|.$$

Plugging in the bound in inequality 4.7 shows that, for any $0 \leq s < s^\star$,

$$\Pr\left[\frac{E_\pi(s)}{n} \geq \alpha s - \left(\frac{1-(2\alpha-1)s}{1-(2\alpha+1)s}\right)^2 \left[s\left(\frac{1+\delta}{n} + \frac{1}{\sqrt{n}}\right)\right]\right] \geq 1 - e^{-n(1-\ln 2)}.$$

Proof of Theorem 2

To prove the upper bound, we use the standard variational method. First, consider the state

$$|x = 0\rangle = \frac{1}{\sqrt{2^n}} \sum_z |z\rangle.$$

This state satisfies

$$\begin{aligned}
\langle x = 0|H_\pi(s)|x = 0\rangle &= s\langle x = 0|H_{P.\pi}|x = 0\rangle \\
&= \frac{s}{2^n} \sum_z H_{P,\pi}(z) \\
&= \alpha n s,
\end{aligned}$$

and so $\frac{E_\pi(s)}{n} \leq \alpha s$ for all $0 \leq s \leq 1$. Recall that by assumption our problem Hamiltonian $H_{P,\pi}$ has at least one state $|w\rangle = |\pi(z_0)\rangle$ which satisfies $H_{P,\pi}|w\rangle = 0$. This state gives

$$\langle w|H_{P,\pi}|w\rangle = \left(\frac{1-s}{2}\right)n,$$

and so $\frac{E_\pi(s)}{n} \leq \frac{1-s}{2}$ for all $0 \leq s \leq 1$. Combining these two upper bounds on the ground state energy gives the desired bound $E_\pi(s) \leq n e_\alpha(s)$ for all $0 \leq s \leq 1$.

We now prove the lower bound. Let $s_0 = \frac{1}{2\alpha+1}(1-\epsilon)$ where $\epsilon \leq 1$. Lemma 2 says that, with probability at least $1 - e^{-n(1-\ln 2)}$,

$$\begin{aligned}
\frac{E_\pi(s_0)}{n} &\geq \alpha s_0 - s_0 \left(\frac{1-(2\alpha-1)s_0}{1-(2\alpha+1)s_0}\right)^2 \left[\frac{1}{n} + \frac{\delta}{n} + \frac{1}{\sqrt{n}}\right] \\
&= \alpha s_0 - s_0 \left(\frac{1 - \frac{2\alpha-1}{2\alpha+1}(1-\epsilon)}{\epsilon}\right)^2 \left[\frac{1}{n} + \frac{\delta}{n} + \frac{1}{\sqrt{n}}\right] \\
&\geq \alpha s_0 - \frac{1}{2\alpha+1}\frac{4}{\epsilon^2}\left[\frac{\delta}{n} + \frac{2}{\sqrt{n}}\right] \\
&= \frac{\alpha}{2\alpha+1}\left[1 - \epsilon - \frac{4}{\alpha\epsilon^2}\left(\frac{\delta}{n} + \frac{2}{\sqrt{n}}\right)\right] \\
&\geq \frac{\alpha}{2\alpha+1}\left[1 - \epsilon - \frac{4}{\alpha\epsilon^2}\left(\frac{c}{n^p} + \frac{2}{\sqrt{n}}\right)\right] \\
&\geq \begin{cases} \frac{\alpha}{2\alpha+1}\left[1 - \epsilon - \frac{4}{\alpha\epsilon^2}\left(\frac{3c}{n^p}\right)\right] & \text{, if } 0 < p < \frac{1}{2} \\ \frac{\alpha}{2\alpha+1}\left[1 - \epsilon - \frac{4}{\alpha\epsilon^2}\left(\frac{3c}{n^{\frac{1}{2}}}\right)\right] & \text{, if } p \geq \frac{1}{2}. \end{cases}
\end{aligned}$$

96

Letting $\epsilon = \frac{1}{n^k}$ with

$$k = \begin{cases} \frac{p}{3} & , \text{if } 0 < p < \frac{1}{2} \\ \frac{1}{6} & , \text{if } \frac{1}{2} \leq p \end{cases}$$

we get (with probability at least $1 - e^{-n(1-\ln 2)}$)

$$\frac{E_\pi(s_0)}{n} \geq \frac{\alpha}{2\alpha + 1}\left[1 - \frac{1}{n^k} - \frac{12c}{\alpha}\frac{1}{n^k}\right]$$

$$\geq \frac{\alpha}{2\alpha + 1}\left[1 - \frac{13c}{\alpha}\frac{1}{n^k}\right]$$

Now, for any $0 \leq s_1 < s_2 \leq 1$ we can write

$$H_\pi(s) = \left(\frac{s - s_1}{s_2 - s_1}\right)H_\pi(s_2) + \left(\frac{s_2 - s}{s_2 - s_1}\right)H_\pi(s_1).$$

Taking the expectation value of both sides in the ground state of $H_\pi(s)$ gives

$$E_\pi(s) \geq \left(\frac{s - s_1}{s_2 - s_1}\right)E_\pi(s_2) + \left(\frac{s_2 - s}{s_2 - s_1}\right)E_\pi(s_1) \quad \text{for } s_1 \leq s \leq s_2. \qquad (4.8)$$

Plugging in $s_1 = s_0, s_2 = 1$ and $s = s^\star$ gives , with probability at least $1 - e^{-n(1-\ln 2)}$,

$$\begin{aligned}
\frac{1}{n}E_\pi(s^\star) &\geq \frac{1}{n}\left(\frac{1 - s^\star}{1 - s_0}\right)E_\pi(s_0) \\
&\geq \left(\frac{1 - \frac{1}{2\alpha+1}}{1 - \frac{1}{2\alpha+1}\left(1 - \frac{1}{n^k}\right)}\right)\frac{\alpha}{2\alpha + 1}\left[1 - \frac{13c}{\alpha}\frac{1}{n^k}\right] \\
&= \left(\frac{2\alpha}{2\alpha + \frac{1}{n^k}}\right)\frac{\alpha}{2\alpha + 1}\left[1 - \frac{13c}{\alpha}\frac{1}{n^k}\right] \\
&\geq \left(1 - \frac{1}{2\alpha n^k}\right)\frac{\alpha}{2\alpha + 1}\left[1 - \frac{13c}{\alpha}\frac{1}{n^k}\right] \\
&\geq \frac{\alpha}{2\alpha + 1}\left[1 - \frac{1}{2\alpha n^k} - \frac{13c}{\alpha}\frac{1}{n^k}\right] \\
&\geq \frac{\alpha}{2\alpha + 1}\left[1 - \frac{14c}{\alpha}\frac{1}{n^k}\right] \qquad (4.9)
\end{aligned}$$

since $c \geq 1$. Now use equation 4.8 with $s_1 = 0$ and $s_2 = s^\star$ to get, for all $s \in [0, s^\star)$,

$$\frac{E_\pi(s)}{n} \geq \frac{s}{s^\star}\frac{E_\pi(s^\star)}{n}.$$

Similarly, using equation 4.8 with $s_1 = s^\star$ and $s_2 = 1$ gives, for $s \in (s^\star, 1]$

$$\frac{E_\pi(s)}{n} \geq \frac{1 - s}{1 - s^\star}\frac{E_\pi(s^\star)}{n}.$$

If inequality 4.9 holds (which occurs with probability at least $1 - e^{-n(1-\ln 2)}$), then using the above two inequalities (as well as 4.9) gives

$$\frac{E_\pi(s)}{n} \geq e_\alpha(s) \left[1 - \frac{1}{n^k} \left(\frac{14c}{\alpha} \right) \right] \text{ for } 0 \leq s \leq 1.$$

This completes the proof.

Proof of Theorem 3

We will use the adiabatic theorem (in the form given in Theorem 3 of [32] with $m = 1$), which we reproduce below with the text of the hypotheses quoted from [33]. We will also use the Scrambled theorem (modified from reference [16]) reviewed in section 1.1.1.

Adiabatic Theorem (from [32] and [33]). *Let $H(s)$ be a finite-dimensional twice differentiable Hamiltonian on $0 \leq s \leq 1$ with a nondegenerate ground state $|\phi(s)\rangle$ separated by an energy gap $\gamma(s)$. Let $|\psi(t)\rangle$ be the state obtained by Schrödinger time evolution with Hamiltonian $H\left(\frac{t}{T}\right)$ starting with state $|\phi(0)\rangle$ at $t = 0$. Then*

$$\sqrt{1 - |\langle \psi(T)|\phi(1)\rangle|^2}$$
$$\leq \frac{1}{T} \left[\frac{1}{\gamma(0)^2} \left\| \frac{dH}{ds} \right\|_{s=0} + \frac{1}{\gamma(1)^2} \left\| \frac{dH}{ds} \right\|_{s=1} + \int_0^1 ds \left(\frac{7}{\gamma^3} \left\| \frac{dH}{ds} \right\|^2 + \frac{1}{\gamma^2} \left\| \frac{d^2H}{ds^2} \right\| \right) \right].$$

We now return to the quantum adiabatic Hamiltonian $H_\pi(s)$ given by (4.3). Recall that our problem Hamiltonian H_P has one zero energy ground state $|z_0\rangle$ with the first excited state energy separated by an energy gap g_P. We have

$$\gamma_\pi(0) = 1$$

and

$$\left\| \frac{dH}{ds} \right\| = \left\| -\sum_{i=1}^n \left(\frac{1 - \sigma_x^i}{2} \right) + H_{P,\pi} \right\|$$
$$\leq 2n.$$

Here $N = 2^n$. The ground state at $s = 1$ is $|\phi(1)\rangle = |\pi(z_0)\rangle$. The state $|\psi_\pi(T)\rangle$ is obtained by evolving with the Hamiltonian $H_\pi\left(\frac{t}{T}\right)$ starting from the state $|x = 0\rangle$.

98

Plugging into the Adiabatic Theorem we get (for any π such that $g_\pi > 0$),

$$\sqrt{\left(1 - |\langle\psi_\pi(T)|\pi(z_0)\rangle|^2\right)} \leq \frac{1}{T}\left[2n + \frac{2n}{g_P^2} + \int_0^1 \frac{28n^2}{\gamma_\pi^3}ds\right]$$

$$\leq \frac{1}{T}\left[\frac{4n^2}{g_\pi^3}\left(1 + \frac{1}{g_P^2}\right) + \frac{28n^2}{g_\pi^3}\right]$$

$$\leq \frac{32n^2}{Tg_\pi^3}\left(1 + \frac{1}{g_P^2}\right) \qquad (4.10)$$

where we used the fact that $g_\pi \leq \gamma_\pi(0) = 1$.

Fix $0 < \epsilon < 1$. Let R be the set of permutations π for which

$$g_\pi > \left(\frac{32n^3}{\left[\frac{\epsilon^2}{128}\sqrt{N}\right]}\left(1 + \frac{1}{g_P^2}\right)\right)^{\frac{1}{3}}. \qquad (4.11)$$

Here $N = 2^n$. Then plugging into (4.10) we get

$$\sqrt{\left(1 - |\langle\psi_\pi(T)|\pi(z_0)\rangle|^2\right)} \leq \frac{\left(\frac{\epsilon^2}{128n}\sqrt{N}\right)}{T} \text{ for all } \pi \in R.$$

Now choose

$$T = \frac{\sqrt{2}\epsilon^2}{128n}\sqrt{N} \qquad (4.12)$$

which guarantees that $|\langle\psi_\pi(T)|\pi(z_0)\rangle|^2 \geq \frac{1}{2}$ for all $\pi \in R$. Assume (to get a contradiction) that the size of R is at least $\epsilon N!$. Now apply the Scrambled Theorem (modified from reference [16]) in the form given in section 1.1.1 to obtain (for $N \geq \frac{256}{\epsilon}$)

$$T \geq \frac{\epsilon^2}{64n}\sqrt{N} \qquad (4.13)$$

where we have used the fact that in our case

$$h^\star = \left(\frac{1}{N-1}\sum_{z \neq z_0} H_P(z)^2\right)^{\frac{1}{2}}$$

$$\leq \left(\frac{1}{N-1}\sum_{z \neq z_0} n^2\right)^{\frac{1}{2}}$$

$$= n.$$

But (4.12) contradicts inequality 4.13. Therefore R cannot contain $\epsilon N!$ permutations

when N is sufficiently large. In other words, for $N \geq \frac{256}{\epsilon}$

$$\Pr\left[g_\pi \leq \left(\frac{32n^3}{\left[\frac{\epsilon^2}{128}\sqrt{N}\right]}\left(1 + \frac{1}{g_P^2}\right)\right)^{\frac{1}{3}}\right] \geq 1 - \epsilon.$$

For example choosing $\epsilon = \frac{64}{n\sqrt{n}}$ we get Theorem 3 in the stated form.

Chapter 5

Phase Transitions in 3XORSAT and Max-Cut

This Chapter is based on joint work with Edward Farhi, Itay Hen, Florent Krzakala, Peter Shor, Peter Young, and Francesco Zamponi.

5.1 Introduction

Here we present two results concerning phase transitions and the quantum adiabatic algorithm. For consistency with other studies, we will use the notation

$$H(\lambda) = H_P - \lambda \sum_{i=1}^{n} \sigma_x^i. \tag{5.1}$$

in this Chapter.

Recall from Chapter 1 that the authors of [37] considered the quantum adiabatic algorithm applied to an ensemble of random instances of 3-XORSAT. To conform to the notation of [37], we write the cost function for an instance of this problem as

$$H_P = \sum_c \sigma_z^{i_{1,c}} \sigma_z^{i_{2,c}} \sigma_z^{i_{3,c}} J_c \tag{5.2}$$

where each clause $c \in \{1, ..., n\}$ is associated with 3 bits $i_{1,c}, i_{2,c}, i_{3,c}$ and a coupling $J_c \in \{\pm 1\}$. Jörg et al [37] studied the ensemble of random instances where each bit is involved in exactly 3 clauses and which have a unique satisfying assignment. In this case the number of clauses is equal to the number of bits. Such instances also have the property that the cost function 5.2 can be mapped unitarily into the form

$$H_P = - \sum_c \sigma_z^{i_{1,c}} \sigma_z^{i_{2,c}} \sigma_z^{i_{3,c}} \tag{5.3}$$

by a product of bit flip operators which map the unique satisfying assignment to the all zeros bit string 000...0. Using numerical diagonalization, the authors of [37] gave evidence that there is a first order quantum phase transition which occurs at

$\lambda_c \approx 1$ in the ground state. (This was supported by their Quantum Monte Carlo and quantum cavity results which apply to the random ensemble with no restrictions on the number of satisfying assignments.) Their results also demonstrate that the minimum gap is exponentially small as a function of n at the transition point. In this Chapter we demonstrate using a duality transformation that the critical value of the transverse field is in fact at $\lambda_c = 1$ (if a unique phase transition is present as a function of λ). The duality that we construct is analogous to the well known duality in the transverse field Ising model.

The second cost function we discuss in this Chapter is

$$H_P = \sum_{(i,j) \in G} \sigma_z^i \sigma_z^j \qquad (5.4)$$

where the graph G of interactions is 3 regular. The ground state of this Hamiltonian encodes the solution to a computational problem called 3 regular Max-Cut which is NP hard in the worst case. Given a graph G, we consider labelings of the vertices with 2 different labels $+1$ and -1. The Max-Cut of the graph is defined to be the maximum number of edges which connect vertices labeled $+1$ to vertices labeled -1, where the maximum is taken over all such labelings. The Max-Cut of the graph G can be obtained from the ground state energy of the Hamiltonian 5.4 (the energy is equal to the total number of edges of the graph minus twice the Max-Cut). It is NP hard to approximate the Max-Cut of 3-regular graphs to within a multiplicative factor 0.997 [7] although there is a classical algorithm which achieves an approximation ratio of at least 0.9326 [29].

We study the Hamiltonian 5.1 where H_P is given by equation 5.4 for a random regular graph G. We use the quantum cavity method[43, 40] within an approximation that allows us to directly investigate the limit $n \to \infty$ at finite inverse temperature β. Working at $\beta = 4$ we show that there is a second order phase transition as a function of the transverse field.

The Hamiltonian 5.4 on a random 3 regular graph will in general be frustrated–the ground state does not separately minimize each of the terms in the Hamiltonian. As a result the problem Max-Cut cannot be solved by linear algebra (as was the case with XORSAT).

There is no "gauge transformation" equivalence between the Hamiltonian 5.4 and the Hamiltonian

$$H_P = \sum_{(i,j) \in G} J_{ij} \sigma_z^i \sigma_j^j \qquad (5.5)$$

where the $J_{ij} \in \{\pm 1\}$. However we do expect the solution to the cavity equations to be similar for these two models since the cavity equations exploit the fact that a random graph looks like a tree, and on a tree such a gauge transformation does exist (see reference [68] for a discussion of this point in the case where there is no transverse field present).

Laumann et al. have used the quantum cavity method to study the spin glass Hamiltonian 5.5 where each J_{ij} is chosen to be $+1$ or -1 with equal probability. Their method is very similar to ours, although the numerics performed in [43] have

some systematic errors which our calculations avoid. (The method used in [43] is a discrete imaginary time formulation of the quantum cavity method which has finite Trotter error, whereas our calculation works in continuous imaginary time[40] where this source of error is absent. Our calculation also does not use the approximation used in [43] where the "effective action" of a path in imaginary time is truncated at second order in a cluster expansion.)

5.2   3-Regular 3XORSAT

In this section we demonstrate a duality mapping for the ensemble of random instances of 3 Regular 3XORSAT with a unique satisfying assignment. This duality mapping explains the critical value $\lambda_c = 1$ of the quantum phase transition in this model [37].

Consider the Hamiltonian

$$H(\lambda) = -\sum_{c=1}^{n} \sigma_z^{i_1,c} \sigma_z^{i_2,c} \sigma_z^{i_3,c} - \lambda \sum_{i=1}^{n} \sigma_x^i \qquad (5.6)$$

corresponding to an instance of 3 regular 3XORSAT with a unique satisfying assignment. The 3 regular hypergraph specifying the instance can be represented by a matrix M where

$$M_{ij} = \begin{cases} 1 & \text{, if bit j is in clause i} \\ 0 & \text{, otherwise.} \end{cases}$$

and where M has 3 ones in each row and 3 ones in each column. The fact that there is a unique satisfying assignment 000...0 is equivalent to the statement that the matrix M is invertible over $\mathbb{F}_2^n$. To see this, consider the equation (with addition mod 2)

$$M\vec{v} = \vec{0}$$

This equation has the unique solution $\vec{v} = \vec{0}$ if and only if there is a unique satisfying assignment for the given instance. This is also the criterion for the matrix M to be invertible.

Write $F(\lambda, M, \beta)$ for the free energy of the Hamiltonian 5.6 at inverse temperature β. Let

$$f(\lambda, \beta) = \lim_{n \to \infty} \frac{1}{n} \mathbb{E}[F(\lambda, M, \beta)] \qquad (5.7)$$

where the expectation value is taken over all $n \times n$ invertible matrices M with 3 ones in each row and in each column. (We have implicitly assumed that this limiting curve exists). We prove in the following section that

$$F(\lambda, M, \beta) = \lambda F(\frac{1}{\lambda}, M^T, \beta). \qquad (5.8)$$

Note that if M is invertible then so is its transpose and therefore M^T corresponds to another instance of 3 regular 3XORSAT with a unique satisfying assignment. Using

103

this fact in equation 5.7 we get

$$\frac{1}{\sqrt{\lambda}} f(\lambda, \beta) = \sqrt{\lambda} f(\frac{1}{\lambda}, \beta).$$

which says that the function $\frac{1}{\sqrt{\lambda}} f(\lambda, \beta)$ is invariant under the transformation $\lambda \to \frac{1}{\lambda}$. Suppose there is a phase transition associated with a discontinuity in some derivative of $f(\lambda, \beta)$ when $\lambda = \lambda_c \in (0, \infty)$. Then there must also be be a phase transition at $\lambda = \frac{1}{\lambda_c}$. If there is only one transition point then it must occur at $\lambda_c = 1$.

Applied to the adiabatic interpolating Hamiltonian

$$H(s) = (1-s) \sum_{i=1}^{n} \left(\frac{1 - \sigma_x^i}{2} \right) + s \sum_{c=1}^{n} \left(\frac{1 - \sigma_z^{i_1,c} \sigma_z^{i_2,c} \sigma_z^{i_3,c}}{2} \right)$$

the duality that we construct in the next section shows that the spectrum of $H(s)$ is the same as the spectrum of $H_{\mathrm{DUAL}}(1 - s)$ where H_{DUAL} is obtained by replacing the problem Hamiltonian hypergraph by its dual. The ground state energy per spin (averaged over all 3 regular instances with a unique satisfying assignment) is symmetric about $s = \frac{1}{2}$ and the first order phase transition observed in [40] occurs at $s = \frac{1}{2}$.

5.2.1 Duality Mapping

We now prove equation 5.8 by constructing a duality transformation. For each $c = 1, ..., n$ define the operator

$$X_c = \sigma_z^{i_1,c} \sigma_z^{i_2,c} \sigma_z^{i_3,c}. \tag{5.9}$$

We also define, for each clause c, a bit string $\vec{y}^c$

$$\vec{y}^c = M^{-1} \hat{e}_c.$$

Here $\hat{e}_c$ is the unit vector with components $(\hat{e}_c)_i = \delta_{ic}$. Note that $\vec{y}^c$ is the unique bit string which violates clause c and satisfies all other clauses. Such a bit string is guaranteed to exist since M is invertible. Let y_i^c denote the ith bit of the string $\vec{y}^c$. Define, for each $c = 1, ..., n$,

$$Z_c = \prod_{i=1}^{n} \left[\sigma_x^i \right]^{y_i^c}. \tag{5.10}$$

Note that

$$\{Z_c, X_c\} = 0$$

and

$$[Z_c, X_{c'}] = 0 \text{ for } c \neq c'.$$

For each bit $i = 1, ..., n$ let $c_1(i), c_2(i), c_3(i)$ be the clauses which bit i participates in. Then

$$\sigma_x^i = Z_{c_1(i)} Z_{c_2(i)} Z_{c_3(i)}. \tag{5.11}$$

104

This follows from the fact that

$$M\hat{e}_i = \hat{e}_{c_1(i)} + \hat{e}_{c_2(i)} + \hat{e}_{c_3(i)}$$

and so

$$
\begin{aligned}
\hat{e}_i &= M^{-1}\left(\hat{e}_{c_1(i)} + \hat{e}_{c_2(i)} + \hat{e}_{c_3(i)}\right) \\
&= \vec{y}^{c_1(i)} + \vec{y}^{c_2(i)} + \vec{y}^{c_3(i)}.
\end{aligned}
$$

The above equation and the definition 5.10 show equation 5.11. Now using 5.11 and 5.9 write

$$
\begin{aligned}
H(\lambda) &= -\sum_{c=1}^{n} \sigma_z^{i_1,c}\sigma_z^{i_2,c}\sigma_z^{i_3,c} - \lambda\sum_{i=1}^{n}\sigma_x^i \\
&= -\sum_{c=1}^{n} X_c - \lambda\sum_{i=1}^{n} Z_{c_1(i)}Z_{c_2(i)}Z_{c_3(i)}. \\
&= \lambda\left[-\frac{1}{\lambda}\sum_{c=1}^{n} X_c - \sum_{i=1}^{n} Z_{c_1(i)}Z_{c_2(i)}Z_{c_3(i)}\right]
\end{aligned}
\tag{5.12}
$$

The X and Z operators satisfy the same commutation relations as the operators σ_x and σ_z. Comparing 5.12 with equation 5.6 we see that the spectrum of $H(\lambda)$ is the same as the spectrum of $\lambda H(\frac{1}{\lambda})$. This proves equation 5.8.

5.3   3 Regular Max-Cut

In this section we discuss the application of the quantum cavity method[43, 40] to the ensemble of Hamiltonians

$$H_G(\lambda) = \sum_{(i,j)\in G} \sigma_z^i\sigma_j^j - \lambda\sum_{i=1}^{n}\sigma_x^i \tag{5.13}$$

where G is chosen to be a random 3 regular graph. Recall from the introduction that the classical cost function in the above Hamiltonian corresponds to an instance of the computational problem 3 Regular Max-Cut.

The quantum cavity method was introduced in [43] and the continuous imaginary time approach that we use was initiated in [40]. Quantum Cavity methods have now been used to study a number of problems including the ferromagnet on the Bethe lattice[40], the spin glass on the Bethe Lattice[43], 3 regular 3-XORSAT [37] (as we have discussed in the previous section), and the quantum Biroli-Mézard model [25]. There are different levels of approximation within the quantum cavity method. The technical term for the approximation we have used is the 1 step replica symmetry breaking (1RSB) ansatz with Parisi parameter $m = 0$. This level of approximation is more powerful than the replica symmetric ansatz which has been used successfully on

the ferromagnet[40] but less powerful than the full 1RSB calculation which has been used to study the 3 regular 3-XORSAT problem[37].

5.3.1 Numerical Results

We have numerically solved the 1RSB quantum cavity equations (in continuous imaginary time) with Parisi parameter $m = 0$ at an inverse temperature $\beta = 4$. A description of the equations that we solve as well as the algorithm that we use to solve them is given in appendix E. The method we use is due to Krzakala et al [40].

We compare our cavity method results (at $\beta = 4$) to ground state quantities (obtained by numerical diagonalization) for 50 random instances of Max-Cut at each size $n = 16$ and $n = 24$. These diagonalization results were obtained by A.P Young[65]. The instances used for the numerical diagonalization were restricted to those which have exactly 2 minimal energy states (note that this is the smallest number possible since the problem is symmetric under flipping all the spins) and for which the energy density of the classical problem at $\lambda = 0$ has the value -1.25 corresponding to a Max-Cut of $\frac{11}{8}n$. The restriction to instances with a fixed Max-Cut was done to tighten the spread in the energy curves over the random ensemble.

In Figure 5-1 we have plotted the energy per spin from the cavity method numerics and the numerical diagonalization. The inset in Figure 5-1 shows a blow up of the data near the classical problem $\lambda = 0$. It is hard to discern the existence of a phase transition from this plot as we do not see evidence of a discontinuity in the first derivative of the energy per spin. The critical behaviour is easier to see using an order parameter such as the Edwards-Anderson order parameter which is defined through equation E.7 in the appendix. This order parameter captures a property of the distribution ρ over paths which arises from the path integral expansion of the partition function (see the appendix for details) and is not expressed as a simple physical observable. Figure 5-2 shows the Edwards-Anderson order parameter which shows a second order (continuous) phase transition at a critical value of $\lambda_c \approx 1.8$. A different order parameter is computed in the numerical diagonalization. The order parameter q_P defined by

$$q_P^2 = \left(\frac{1}{n^2 - n} \right) \sum_{i<j} \langle \sigma_z^i \sigma_z^j \rangle$$

was computed by numerical diagonalization in the ground state for the random sets of instances at $n = 16$ and $n = 24$. The x magnetization, computed with the cavity method, shows a kink signaling the second order phase transition as can be seen in Figure 5-3.

We can compare our results to studies which have been performed for the spin glass where the Hamiltonian is given by equation 5.5 with couplings $J_{ij} \in \{\pm 1\}$ with equal probability. This problem has been studied both in the classical[8, 48] and quantum[43] cases. In the classical case $\lambda = 0$ Boettcher has performed numerics on random regular graphs of increasing sizes up to 4096 vertices and has computed the average ground state energy density at each number of vertices. By extrapolating his results to the limit $n \to \infty$ he obtains the limiting value $-1.2716(1)[8]$. This is

consistent with the 1RSB result from [48] which is -1.2717. Note that the limiting value of the energy density which we computed for our problem as $\lambda \to 0$ (seen in the inset of Figure 5-1) is not far off from these values. In [43], Laumann et al show that there is a second order transition line for the spin glass as a function of λ and temperature T. Because they use a discrete time formulation of the cavity method their simulations begin to have substantial systematic Trotter errors at low temperature. These errors are associated with dividing up the interval $[0, \beta]$ into a finite number N_t of slices. The correct answer would correspond to the limit $N_t \to \infty$ where imaginary time is treated as a continuous variable. They nevertheless estimate the critical value of the field $\lambda_c(T)$ by fitting their results which are computed at $N_t = 6, 7, 8, 9, 10, 11$ and a range of temperatures as low as $T \sim 0.15[43]$. At $\beta = 4$ their estimate of the critical value of the transverse field is $\lambda_c \approx 1.85[43]$. This value is close to what we obtain for Max-Cut at $\beta = 4$ as can be seen in the inset of Figures 5-2 and 5-3.
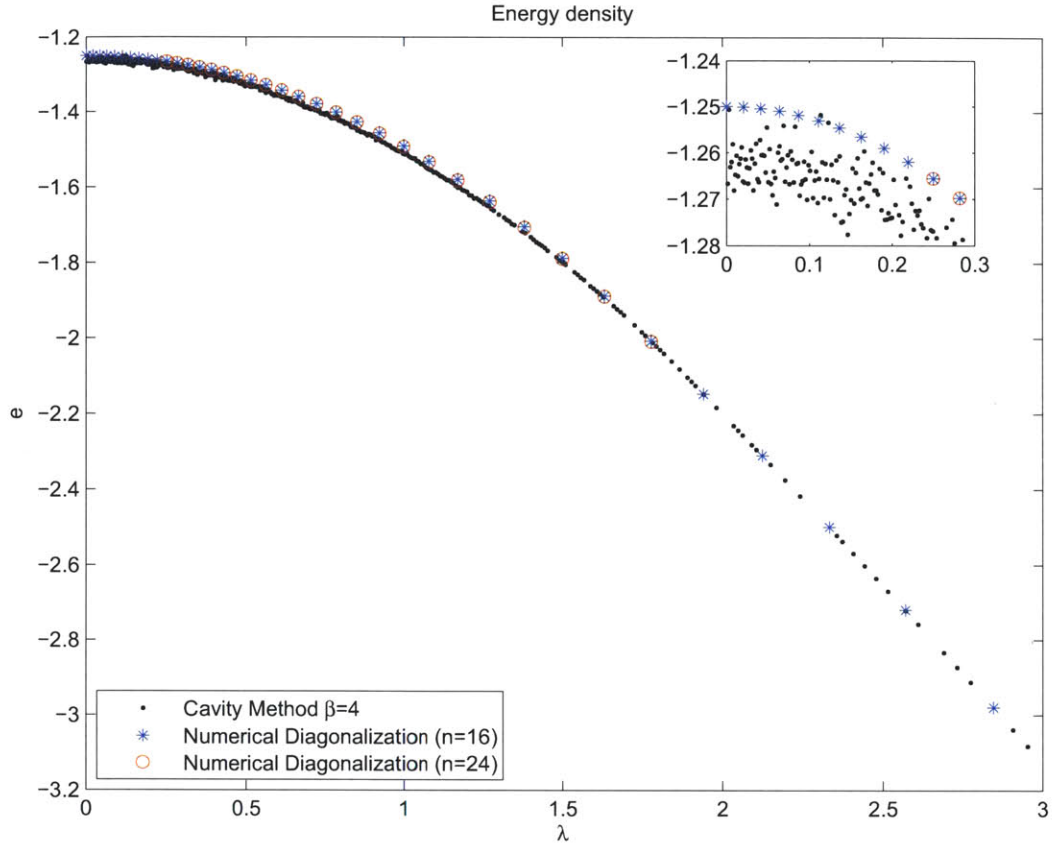
Figure 5-1: The black dots show the energy per spin for 3 Regular Max-Cut at $\beta = 4$ computed in the 1RSB cavity method with Parisi parameter $m = 0$ as a function of transverse field λ. The inset shows the behaviour near $\lambda = 0$. The blue circles show numerical diagonalization results for the ground state averaged over 50 random instances of Max-Cut on 16 bits which are restricted to the subset which have exactly two degenerate lowest energy states with energy per spin equal to -1.25. The red asterisks show numerical diagonalization results for the ground state averaged over 50 random instances of Max-Cut on 24 bits with exactly two degenerate lowest energy states with energy per spin equal to -1.25.
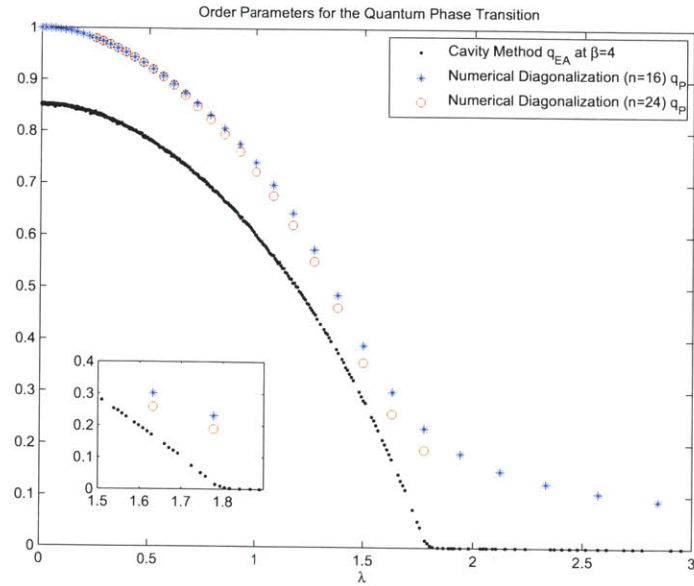
Figure 5-2: The black dots show the Edwards Anderson order parameter computed with the 1RSB cavity method with $m = 0$ at $\beta = 4$. The blue dots show the order parameter q_P in the ground state, averaged over 50 random instances of Max-Cut on 16 bits which are restricted to the subset which have exactly two degenerate lowest energy states with energy per spin equal to -1.25. The red dots show the order parameter q_P in the ground state, averaged over 50 random instances of Max-Cut on 24 bits which are restricted to the subset which have exactly two degenerate lowest energy states with energy per spin equal to -1.25.
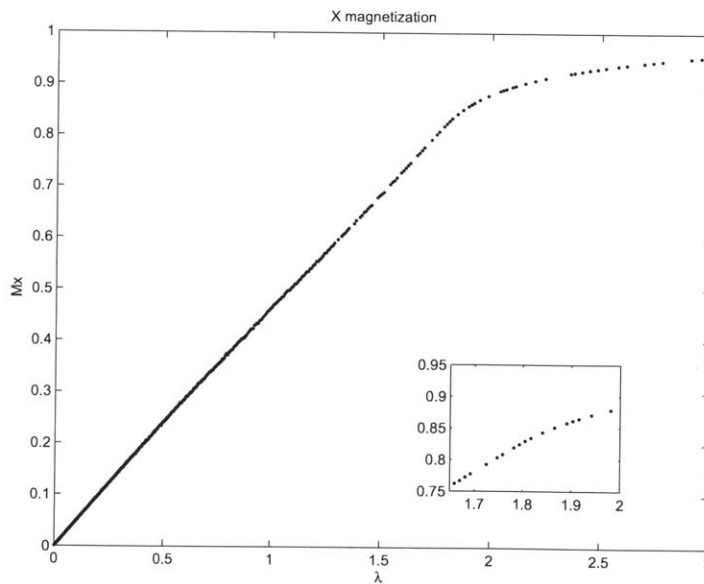
Figure 5-3: The magnetization along the transverse direction computed with the 1RSB cavity method with $m = 0$ at $\beta = 4$. The inset shows the region around the phase transition.

Details of the Cavity Method Simulations

Here we describe the numerical simulations in more detail, referring to some technical terminology from Appendix E.

Our simulation was run on a Sicortex computer cluster in an embarrassingly parallel fashion. For the data in Figures 5-1, 5-3 and 5-2 we ran two independent simulations at each value of λ. 59 sweeps of the cavity iteration from each data set were used for equilibration, and then the subsequent 151 sweeps from each of these two independent simulations were used to compute observables. We have confirmed our numerical results with those obtained from an independent implementation of the continuous time cavity method[67].

We used population sizes (see appendix E) $N_D = 200$ and $N_R = 15000$. We found numerically that there is a systematic error associated with taking N_R to be too small and that this error increases as β is increased. We believe that $N_R = 15000$ is large enough to make this error small for our simulation at $\beta = 4$. In appendix E.3 we support this by showing that the results obtained from numerics at $N_R = 10000$ differ only very slightly from those obtained with $N_R = 15000$ at this value of β. The simulations shown at $N_R = 10000$ had one data run for each value of λ , with 59 sweeps used for equilibration and data used from the subsequent 351 sweeps.

For unknown reasons our computer code sometimes (primarily at higher values of the transverse field λ and larger values of N_R) did not output the data file. This computer bug did not seem to compromise the results when the output was produced (we checked this by comparing with results from the independent implementation[67]). In Figures 5-1,5-3, and 5-2 we have only plotted data for values of λ where *both* independent simulations at $N_R = 15000$ outputted data files. This is responsible for the gaps in the data that can be seen on the plots.

5.4   Discussion

Using the quantum cavity method we have given evidence that there is a second order quantum phase transition at temperature $T = 0.25$ for random instances of 3 regular Max-Cut in a transverse field. This suggests that there may be a second order phase transition at zero temperature (i.e in the ground state),in contrast to the zero temperature first order phase transitions in random 3 regular 3-XORSAT [37] and for scrambled cost functions (from the previous Chapter and [16]) which are associated with exponentially small gaps and failure of the quantum adiabatic algorithm.

What does this have to do with the quantum adiabatic algorithm? The fact that we did not find a first order phase transition is good news. Further investigation is needed to determine the performance of the quantum adiabatic algorithm. Our collaborators Itay Hen and Peter Young are currently using Quantum Monte Carlo simulations to study the eigenvalue gap for this problem. Note that for the random ensemble which we studied using the quantum cavity method the number of satisfying assignments is unrestricted. As a result the minimum eigenvalue gap of the Hamiltonian $H(\lambda)$ can be zero due to degeneracy at $\lambda = 0$. For the moment let us

put that issue aside and discuss whether or not polynomial time quantum adiabatic evolution succeeds in moving past the critical point.

It is possible that the eigenvalue gap at the critical point is only polynomially small as a function of n and the quantum adiabatic algorithm succeeds in finding a solution in polynomial time. On the other hand the quantum adiabatic algorithm will not succeed if the gap is exponentially small at the critical value of the field. This would provide an interesting example where a continuous phase transition is associated with an exponentially small gap (we are aware of one example where this occurs [23] but this example has many strange properties). A third possibility is that the gap at the critical point λ_c is only polynomially small in n but that effects near zero transverse field (such as perturbative crosses[2, 4, 5, 18]) cause the adiabatic algorithm to fail for this problem. It is a topic for future research to distinguish which (if any) of these scenarios holds for this model. The same questions can also be asked about the spin glass model studied by Laumann et al [43].

Appendix A

Estimators in the Continuous Imaginary Time Ensemble of Paths

In this appendix we derive equations 1.18 and 1.20. These derivations have appeared previously in reference [14].

Estimator for a Diagonal Operator D

To derive the estimator for D given by equation 1.18 we write

$$\frac{Tr[De^{-\beta H}]}{Tr[e^{-\beta H}]} =$$

$$\frac{1}{Z(\beta)} Tr \left[D \sum_{m=0}^{\infty} (-\lambda)^m e^{-\beta D} \int_0^{\beta} dt_m \int_0^{t_m} dt_{m-1} ... \int_0^{t2} dt_1 V_I(t_m) V_I(t_{m-1}) ... V_I(t_1) \right]$$

$$(A.1)$$

where $V_I(t)$ is given by equation 1.14. (The $m = 0$ term in the above sum is $\frac{1}{Z(\beta)} Tr \left[e^{-\beta D} \right]$.) Inserting complete sets of states in the basis $\{|z\rangle\}$ which diagonalizes D we obtain

$$\begin{aligned}
\frac{Tr[De^{-\beta H}]}{Tr[e^{-\beta H}]} &= \frac{1}{Z(\beta)} \sum_{m=0}^{\infty} \left[(-\lambda)^m \sum_{\{z_1,...,z_m\}} \langle z_1|D|z_1\rangle \langle z_1|V|z_m\rangle \langle z_m|V|z_{m-1}\rangle ... \langle z_2|V|z_1\rangle \right. \\
&\quad \left. \int_0^{\beta} dt_m \int_0^{t_m} dt_{m-1} ... \int_0^{t2} dt_1 e^{-(E_1 t_1 + E_2(t_2-t_1) + ... + E_1(\beta-t_m))} \right] \\
&= \langle D(z(t=0)) \rangle_{\rho}.
\end{aligned}$$

where the expectation value is with respect to the measure ρ, and $D(z(t = 0)) = \langle z(0)|D|z(0)\rangle$. Noting that the measure ρ is invariant under a translation of the path by a time $x \in [0, \beta]$ (this corresponds to the transformation t_i goes to $(t_i + x) \mod \beta$ for $i \in 1, ..., m$ followed by a reordering of the labels i to maintain time ordering), we

have that
$$\langle D(z(t=0))\rangle_\rho = \langle D(z(t=x))\rangle_\rho \text{ , for all } x \in [0,\beta].$$
We obtain the stated estimator for $\langle D \rangle$ by averaging over all $x \in [0,\beta]$
$$\langle D \rangle \equiv \frac{Tr[De^{-\beta H}]}{Tr[e^{-\beta H}]} = \langle \frac{1}{\beta} \int_0^\beta D(z(x))dx \rangle_{\rho.} .$$

Estimator for $\langle \lambda V \rangle$

As in the previous section, we begin by expanding the operator $e^{-\beta H}$

$$\frac{Tr[\lambda V e^{-\beta H}]}{Tr[e^{-\beta H}]}$$
$$= \frac{1}{Z(\beta)}Tr\left[\lambda V \sum_{m=0}^{\infty}(-\lambda)^m e^{-\beta H_0} \right.$$
$$\left. \int_0^\beta dt_m \int_0^{t_m} dt_{m-1}... \int_0^{t2} dt_1 V_I(t_m)V_I(t_{m-1})...V_I(t_1) \right] .$$

For $m = 0, 1, 2...$ we have

$$\int_0^\beta dt_{m+1} \int_0^{t_{m+1}} dt_m ... \int_0^{t2} dt_1 V_I(t_{m+1})V_I(t_m)...V_I(t_1)\delta(t_1)$$
$$= \int_0^\beta dt_{m+1} \int_0^{t_{m+1}} dt_m ... \int_0^{t3} dt_2 V_I(t_{m+1})V_I(t_m)...V_I(t_2)V$$
$$= \int_0^\beta dt_m \int_0^{t_m} dt_{m-1} ... \int_0^{t2} dt_1 V_I(t_m)V_I(t_{m-1})...V_I(t_1)V .$$

Using this expression we obtain

$$\frac{Tr[\lambda V e^{-\beta H}]}{Tr[e^{-\beta H}]} = \frac{1}{Z(\beta)}Tr\left[(-1)\sum_{m=0}^{\infty}(-\lambda)^{m+1}e^{-\beta H_0}\right.$$
$$\left. \int_0^\beta dt_{m+1} \int_0^{t_{m+1}} dt_m ... \int_0^{t2} dt_1 V_I(t_{m+1})V_I(t_m)...V_I(t_1)\delta(t_1) \right]$$
$$= \frac{1}{Z(\beta)}Tr\left[(-1)\sum_{m=1}^{\infty}(-\lambda)^m e^{-\beta H_0}\right.$$
$$\left. \int_0^\beta dt_m \int_0^{t_m} dt_{m-1} ... \int_0^{t2} dt_1 V_I(t_m)V_I(t_{m-1})...V_I(t_1)\delta(t_1) \right]$$
$$= -\langle (1-\delta_{m,0})\delta(t_1) \rangle_\rho$$
$$= -\langle (1-\delta_{m,0})\sum_{l=1}^{m}\delta(t_l) \rangle_\rho \quad \text{(since only } t_1 \text{ can ever be 0).}$$

114

In the last two lines of the above m appears inside an expectation value $\langle ... \rangle_\rho$. In this context m is considered to be a function of the path. Now we can use the fact that the measure ρ over paths is invariant under translations of the path in imaginary time to write

$$\langle \lambda V \rangle \equiv \frac{Tr[\lambda V e^{-\beta H}]}{Tr[e^{-\beta H}]} \quad = \quad -\langle (1 - \delta_{m,0}) \frac{1}{\beta} \int_0^\beta \sum_{l=1}^m \delta(t_l - t) dt \rangle_\rho$$

$$= \quad -\langle \frac{m}{\beta} \rangle_\rho.$$

So $\langle \lambda V \rangle$ is $-\frac{1}{\beta}$ times the average number of jumps in a path of length β.

Appendix B

Algorithm for Sampling from the Single Spin Path Integral

In this appendix we present an algorithm which samples from the normalized probability distribution over single spin paths $\sigma(t)$ for $t \in [0, \alpha]$ with fixed boundary conditions. This algorithm was first presented in reference [18].

The Hamiltonian is

$$H = h\sigma_z - c\sigma_x$$

and we consider the path integral with boundary conditions b_1, b_2 :

$$\langle \sigma_z = b_2 | e^{-\alpha H} | \sigma_z = b_1 \rangle = \sum_{w=0}^{\infty} c^w \int_0^\alpha dt_w \int_0^{t_w} dt_{w-1} \ldots \int_0^{t_2} dt_1 e^{-h \int_0^\alpha \sigma(t)dt}$$

where

$$\sigma(t) = \begin{cases} b_1 & 0 \leq t < t_1 \\ -b_1 & t_1 \leq t < t_2 \\ \vdots \\ b_2 & t_w \leq t < \alpha \end{cases}$$

For each of the four possible sets of boundary conditions we can define a probability distribution over paths $\sigma(t)$ (which in the expression below are restricted to those which satisfy the given boundary conditions)

$$\rho_{b_1,b_2}(\sigma(t)) = \frac{c^w e^{-h \int_0^\alpha \sigma(t)dt} dt_1 \ldots dt_w}{\langle b_2 | e^{-\alpha H} | b_1 \rangle}.$$

It will be useful for us to make the change of variables from the times $(t_1, \ldots, t_w)$ to the waiting times $(u_1, \ldots, u_w)$ defined by

$$
\begin{aligned}
u_1 &= t_1 \\
u_j &= t_j - t_{j-1} \quad j \geq 2
\end{aligned}
$$

Then the weight assigned to each path is

$$\rho_{b_1,b_2}(\sigma(t)) = \frac{c^w e^{-h[b_1 u_1 - b_1 u_2 + b_1 u_3 - \ldots + b_2\left(\lambda - \sum_{k=1}^{w} u_k\right)]} du_1 \ldots du_w}{\langle b_2 | e^{-\alpha H} | b_1 \rangle} \tag{B.1}$$

We are now ready to describe our algorithm which, given boundary conditions b_1, b_2, samples from the distribution ρ_{b_1,b_2}.

Path Integral Sampling Algorithm

1. Start at t=0 in state $B_1 = b_1$ defined by the boundary conditions. Set i=1.

2. Draw the waiting time u_i until the next flip from the exponential distribution with parameter $[\sqrt{h^2 + c^2} + B_i h]$, which is described by the probability distribution function

$$f(u_i) = [\sqrt{h^2 + c^2} + B_i h] e^{-u_i [\sqrt{h^2 + c^2} + B_i h]}$$

   If $\sum_{j=1}^{i} u_j > \alpha$ then go to step 3. Otherwise define $B_{i+1} = -B_i$ and set $i \to i + 1$ and repeat step 2.

3. Take the path you have generated (which will in general be longer than α), and look at the segment $[0, \alpha]$. If this path satisfies the boundary condition at $t = \alpha$ then take this to be the generated path. Otherwise, throw away the path and repeat from step (1).

We now show that this algorithm generates paths from the distribution (B.1). Before conditioning on the boundary conditions being satisfied, the probability of generating a sequence of waiting times in $(u_1, u_1 + du_1), (u_2, u_2 + du_2), \ldots (u_w, u_w + du_w)$ followed by any waiting time u_{w+1} such that $u_{w+1} > \alpha - \sum_{j=1}^{w} u_i$ is given by

$$f(u_1)f(u_2)\ldots f(u_w) du_1 du_2 \ldots du_w \mathrm{Prob}\left(u_{w+1} > \alpha - \sum_{j=1}^{w} u_i\right)$$

$$= f(u_1)f(u_2)\ldots f(u_w) du_1 du_2 \ldots du_w e^{-(\alpha - \sum_{j=1}^{w} u_i)[\sqrt{h^2 + c^2} + B_{w+1} h]}$$

$$= \left[\left(\prod_{i=1}^{w}[\sqrt{h^2 + c^2} + B_i h]\right)\right.$$

$$\left. e^{-\sum_{i=1}^{w} u_i \sqrt{h^2 + c^2}} e^{-\sum_{i=1}^{w} u_i B_i h} du_1 \ldots du_w e^{-(\alpha - \sum_{j=1}^{w} u_j)[\sqrt{h^2 + c^2} + S B_{w+1} h]}\right]$$

$$= \begin{cases} c^w e^{-\alpha\sqrt{c^2 + h^2}} e^{-\int_{t=0}^{\alpha} \sigma(t) h \, dt} du_1 \ldots du_w, & \text{if } w \text{ is even} \\ c^w \left[\sqrt{1 + \left(\frac{h}{c}\right)^2} + B_1 \left(\frac{h}{c}\right)\right] e^{-\alpha\sqrt{c^2 + h^2}} e^{-\int_{t=0}^{\alpha} \sigma(t) h \, dt} du_1 \ldots du_w, & \text{if } w \text{ is odd} . \end{cases}$$

In the last line we have used the difference of squares formula to simplify consecutive

terms: $[\sqrt{c^2 + h^2} + h][\sqrt{c^2 + h^2} - h] = c^2$. When we condition on the given boundary conditions (whatever they may be), this generates the correct distribution over paths. (To see this note that two different paths with the same boundary conditions b_1, b_2 are generated by the above procedure with the correct ratio of probabilities.)

Appendix C

Analysis and Derivation of Estimators at Fixed Energy

Our analysis of the new Monte Carlo method of chapter 2 is based on properties of the operator

$$A(E) = \left(\frac{-\lambda(E)}{H_0 - E} V \right)$$

as was seen from the Introduction. We first derive some properties of this operator and then we derive equations (2.14) and (2.15). The material in this appendix has appeared previously in [14].

We take $E < 0$, and $\lambda(E)$ is defined to be the positive value of λ such that the ground state of $H(\lambda)$ has energy E. We will show the following properties of this operator:

1. All eigenvalues of $A(E)$ are real and ≤ 1 in absolute value.

2. $|\psi_g(\lambda(E))\rangle$ is an eigenvector of $A(E)$ with eigenvalue $+1$. It may also be the case that $|\psi_g(-\lambda(E))\rangle$ is an eigenvector of $A(E)$ with eigenvalue -1. There are no other eigenvectors of $A(E)$ with eigenvalues ± 1.

The second property says that the subspace of states spanned by eigenvectors of $A(E)$ with ± 1 eigenvalues is either 1 or 2 dimensional. We will see that the case where it is two dimensional occurs only when $H(-\lambda(E))$ has ground state energy E.

To show that eigenvalues of $A(E)$ are real whenever $E < 0$, note that the spectrum of $A(E)$ is the same as the spectrum of the operator

$$\sqrt{H_0 - E} A(E) \frac{1}{\sqrt{H_0 - E}}$$

which is Hermitian.

Now suppose (to reach a contradiction) that $|r\rangle$ is an eigenvector of $A(E)$ with eigenvalue $R > 1$. Then

$$\left[H_0 + \lambda(E) \frac{1}{R} V \right] |r\rangle = E |r\rangle$$

which says that there exists an eigenvector of $H(\frac{\lambda(E)}{R})$ with eigenvalue E. In section 2.2 we proved that the ground state energy of $H(\lambda)$ is strictly decreasing for positive λ, which means that no eigenvector of $H(\frac{\lambda(E)}{R})$ can have energy smaller than or equal to E. This is a contradiction and so all positive eigenvalues of $A(E)$ are ≤ 1.

Now, since we have proven that all eigenvalues of $A(E)$ are ≤ 1, if it is the case that some negative eigenvalue of $A(E)$ is < -1 then the eigenvalue W of $A(E)$ which is largest in magnitude has negative sign. If this were true, then the limit

$$\lim_{k \to \infty} \frac{1}{|W|^{2k+1}} Tr \left[(A(E))^{2k+1} \right]$$

would be equal to a negative constant. But this cannot be the case since all matrix elements of $A(E)$ are positive or zero. So we have shown property 1.

We now proceed to show property 2 which was stated earlier in the appendix, and in the process we prove the inequality 2.7 from section 2.2. The real nonzero eigenvalues of $A(E)$ can be related to eigenvalues of $H(\lambda)$ for some value of λ. In particular, suppose that ω is a real nonzero eigenvalue of $A(E)$ with eigenvector $|\omega\rangle$. Then

$$(H_0 - E)\,\omega|\omega\rangle \;=\; -\lambda(E)V|\omega\rangle$$

so

$$H \left(\frac{\lambda(E)}{\omega} \right) |\omega\rangle = E|\omega\rangle \,.$$

Let us write the eigenvalues of $A(E)$ as

$$1 = a_1(E) > a_2(E) \geq ... \geq a_{2^n}(E) \geq -1 \,.$$

(which follows by property 1.) Then the values of λ at which $H(\lambda)$ has an eigenvalue with energy E are

$$\frac{\lambda(E)}{a_j(E)}$$

for $j \in \{1, ..., 2^n\}$. Write $\lambda_-(E)$ for the negative value of lambda such that $H(\lambda_-(E))$ has ground state energy E. The values of λ which are smallest in magnitude (and hence closest to the axis $\lambda = 0$) are $\frac{\lambda(E)}{a_1(E)} = \lambda(E) > 0$ and $\frac{\lambda(E)}{a_{2^n}(E)} = \lambda_-(E) < 0$ and these correspond to the ground state at energy E, with $|\lambda_-(E)| \geq \lambda(E)$. All other values are greater than $\lambda(E)$ in magnitude. This proves inequality 2.7, and also shows property 2 described above.

Derivation of the Estimators β_{est} and λ^2_{est}

Having derived properties 1. and 2. of the operator $A(E)$ in the previous section, we now proceed to use these properties to prove equations 2.15 and 2.16.

Our treatment below applies to both the generic case where inequality 2.7 is strict

as well as the nongeneric case where equality holds as long as in the latter case m is always even. In either case we have

$$Tr\left[(A(E))^m\right] = (\lambda(E))^m F(E, m).$$ (C.1)

(Recall the definition of $F(E, m)$ from equation 2.9.) We can write this as

$$(\lambda(E))^m F(E, m) = \sum_{i=1}^{2^n} (a_i(E))^m.$$

Taking the log and differentiating both sides gives

$$\frac{d}{dE}\left[\log\left(\lambda(E)^m\right) + \log\left(F(E, m)\right)\right] = \frac{1}{\sum_{i=1}^{2^n} (a_i(E))^m} \sum_{j=1}^{2^n} m\,(a_j(E))^{m-1}\frac{da_j}{dE}.$$ (C.2)

Now take the limit as $m \to \infty$. Note that in the nongeneric case the fact that m is taken to be even ensures that the denominator of equation C.2 does not vanish in the limit of large (even) m. The limit of the RHS is zero for any fixed value of E. This is because in the large m limit the only terms which contribute to the sum in the numerator are those corresponding to values of j for which $|a_j(E)| = 1$. For these values of $a_j(E)$ (note j is either 1 or 2^n for these values) it is always the case that $\frac{da_j}{dE} = 0$ so these terms contribute zero to the sum.

So we have shown that

$$\lim_{m \to \infty}\left(m\frac{1}{\lambda(E)}\frac{d\lambda}{dE} + \frac{1}{F(E, m)}\frac{dF(E, m)}{dE}\right) = 0.$$ (C.3)

Hence

$$\lim_{m \to \infty}\frac{1}{m}\frac{1}{F(E, m)}\frac{dF(E, m)}{dE} = -\frac{1}{\lambda(E)}\frac{d\lambda}{dE}.$$

The left hand side of this equation can be rewritten as an ensemble average, which results in

$$\lim_{m \to \infty}\frac{1}{m}\frac{1}{F(E, m)}\frac{dF(E, m)}{dE} = \lim_{m \to \infty}\langle\frac{1}{m}\sum_{i=1}^{m}\frac{1}{E_i - E}\rangle_f = -\frac{1}{\lambda(E)}\frac{d\lambda}{dE}.$$ (C.4)

So for fixed m sufficiently large, one can estimate the quantity $-\frac{1}{\lambda(E)}\frac{d\lambda}{dE}$ using the ensemble average $\frac{\bar{\beta}(E, m)}{m}$. This proves equation 2.15.

We now derive an estimator for the quantity $\lambda(E)$ itself as an ensemble average. For this purpose we make use of the fact that

$$\lim_{m \to \infty}\frac{F(E, m - 2)}{F(E, m)} = \lambda(E)^2.$$

Expanding the numerator and denominator as sums over paths, we obtain

$$\frac{F(E, m-2)}{F(E, m)} = \frac{\sum_{\{z_1,\ldots,z_{m-2}\}} \langle z_1| - V|z_{m-2}\rangle \ldots \langle z_2| - V|z_1\rangle \prod_{i=1}^{m-2} \frac{1}{E_i - E}}{\sum_{\{z_1,\ldots,z_m\}} \langle z_1| - V|z_m\rangle \ldots \langle z_2| - V|z_1\rangle \prod_{i=1}^{m} \frac{1}{E_i - E}} . \qquad (C.5)$$

Now rewrite the numerator as

$$F(E, m-2)$$
$$= \sum_{\{z_1,\ldots,z_m\}} \bigg(\langle z_1| - V|z_m\rangle\langle z_m| - V|z_{m-1}\rangle\langle z_{m-1}| - V|z_{m-2}\rangle \ldots \langle z_2| - V|z_1\rangle$$
$$\prod_{i=1}^{m} \frac{1}{E_i - E} \left[\delta_{z_1 z_{m-1}} (E_m - E)(E_1 - E) \frac{1}{\langle z_1|V^2|z_1\rangle} \right] \bigg)$$

where $\delta_{z_1 z_{m-1}}$ is the Kronecker delta. Since our distribution $f(z_1, \ldots, z_m)$ is invariant under cyclic permutations of the bit strings $\{z_i\}$, we can write

$$F(E, m-2)$$
$$= \sum_{\{z_1,\ldots,z_m\}} \bigg(\langle z_1| - V|z_m\rangle\langle z_m| - V|z_{m-1}\rangle\langle z_{m-1}| - V|z_{m-2}\rangle \ldots \langle z_2| - V|z_1\rangle$$
$$\prod_{j=1}^{m} \frac{1}{E_j - E} \left[\frac{1}{m} \sum_{i=1}^{m} \delta_{z_{i+2} z_i} (E_{i+1} - E)(E_i - E) \frac{1}{\langle z_i|V^2|z_i\rangle} \right] \bigg)$$

where $z_{m+1} = z_1$ and $z_{m+2} = z_2$. Inserting this formula into equation C.5 gives the final expression for $\lambda(E)^2$ as an ensemble average

$$\lim_{m\to\infty} \left\langle \frac{1}{m} \sum_{i=1}^{m} \delta_{z_{i+2} z_i} (E_{i+1} - E)(E_i - E) \frac{1}{\langle z_i|V^2|z_i\rangle} \right\rangle_f = \lambda(E)^2 . \qquad (C.6)$$

This proves equation 2.16.

Appendix D

Convergence of The Markov Chain

We show in this section that the Markov Chain defined in section 2.4 can be used to estimate any quantity which is invariant under cyclic permutations of the path. This appendix has appeared previously in reference [14]. In order to streamline the proof, it will be useful to define a different Markov Chain over paths which has the update rule (for some fixed $0 < p < 1$)

1. With probability p do 1 update of the Markov Chain defined in section 2.4.

2. With probability $1 - p$ apply a random cyclic permutation to the path by letting $\{z_1, z_2, ..., z_m\} \to \{z_j, z_{j+1}, ..., z_m, z_1, ...z_{j-1}\}$ for uniformly random $j \in \{1, ..., m\}$.

In the next two sections we show that the above Markov Chain has limiting distribution f (defined in equation 2.8) for any choice of the parameter $0 < p < 1$. This Markov Chain with fixed $0 < p < 1$ induces a random walk on equivalence classes of paths where an equivalence class is the set of all paths related to a given path by cyclic permutation. In this equivalence class random walk, step 2 does nothing. So the limiting distribution over equivalence classes is the same whether or not step 2 is performed. If one only estimates quantities which are invariant under cyclic permutations (note that the estimators we have discussed have this property) then one can use the algorithm with $p = 1$ (so step 2 is never performed).

To show that the above Markov Chain converges to the limiting distribution f over paths, it is sufficient to verify that the update rule constructed above satisfies the following two conditions [51]:

- *Ergodicity*: Given any two paths A and B, it is possible to reach path B by starting in path A and applying the Markov chain update rule a finite number of times.

- *Detailed Balance:* For any two paths A and B,

$$f(A)P(A \to B) = f(B)P(B \to A) \tag{D.1}$$

where $P(X \to Y)$ is the probability of transitioning to the path Y given that you start in path X and apply one step of the Markov chain.

We now show that this Markov Chain satisfies these conditions.

Ergodicity

In order to show ergodicity of the Markov Chain defined above, we first note that a path can be specified either by a list of bit strings $\{z_1, ..., z_m\}$, or by one bit string z_{start} followed by a list of bits in which flips occur $\{b_1, ..., b_m\}$, with each $b_r \in \{1, ..., n\}$.

We now show that by applying the Monte Carlo update rules illustrated in figures 2-2 and 2-3, it is possible to transform an arbitrary path $A \longleftrightarrow \{z_{start}, \{b_1, ..., b_m\}\}$ into another arbitrary path $B \longleftrightarrow \{y_{start}, \{c_1, ..., c_m\}\}$ where y_{start} and z_{start} differ by an even number of bit flips. This is sufficient to show ergodicity because any path B can be cyclically permuted into a path which starts in a state $\tilde{y}_{start}$ that differs from z_{start} by an even number of flips (and our Markov chain includes moves which cyclically permute the path). We assume here that $m \geq 4$, since we are interested in the limit of large m anyways. In order to transform path A into path B, we give the following prescription:

1. First transform z_{start} into y_{start}. To do this, note that one can move any two flips b_i and b_j so that they are just before and just after z_{start} (i.e $\tilde{b}_1 = b_j$ and $\tilde{b}_m = b_i$), by applying the flip interchange rule illustrated in figure 2-2. So long as you do not interchange the first and last flip in the list, the bit string z_{start} will remain unchanged. Then one can flip both of these bits in the bit string z_{start} by interchanging the two flips $\tilde{b}_1$ and $\tilde{b}_m$. This describes how to flip any two bits in z_{start}, assuming that flips in these bits occur somewhere in the path. Now suppose that you wish to flip 2 bits in z_{start} but one or both of the bits does not occur in the current list of flips in the path. In that case you must first take some pair of flips which occur in some other bit q, and then move them until they are adjacent using the flip interchange rule (without ever moving them past z_{start}). Once they are adjacent, you can replace them with a pair of flips in another bit using the flip replacement rule illustrated in figure 2-3 . Assuming $m \geq 4$, there will always be two pairs of flips in the path which can be replaced by flips in the two bits that you desire to change in z_{start}.

2. After z_{start} has been transformed into y_{start}, one must then make the list of flips equal to $\{c_1, ..., c_m\}$. This can be done by interchanging flips and replacing pairs of flips as described above. Since this can always be achieved without interchanging the first and last flip in the path, the bit string y_{start} will remain unchanged by this procedure.

Detailed Balance

Here we demonstrate that the Markov Chain defined above satisfies the detailed balance condition from equation D.1. To show this, fix two paths A and B and consider the probability of transitioning between them in one step of the Monte Carlo update rule. This probability is zero except in the following cases

1. $A = B$. In this case detailed balance is trivially satisfied.

2. A cyclic permutation of the bit strings (which is not the identity) maps the path A into the path B. In this case $f(A) = f(B)$ and the probability of transitioning from A to B in one move of the Markov Chain is also equal to the probability of the reverse transition from B to A (this is because for every cyclic permutation which maps A to B the inverse permutation is also cyclic and maps B to A). So detailed balance is satisfied.

3. A and B are the same path except for at one location. In other words, A can be described by the sequence $\{z_1^A, ..., z_m^A\}$ and B can be described by the sequence $\{z_1^B, ..., z_m^B\}$ where the corresponding bit strings are all the same except for one pair z_i^A and z_i^B. Write q_1 for the bit in which z_i^A and z_{i-1}^A differ, and q_2 for the bit in which z_i^A and z_{i+1}^A differ. We have to consider two cases depending on whether or not $q_1 = q_2$:

- Case 1: $q_1 \neq q_2$. Detailed balance follows in this case since the transition probabilities follow the Metropolis Monte Carlo rule

$$\frac{P(A \to B)}{P(B \to A)} = \min\left\{1, \frac{E_i^A - E}{E_i^B - E}\right\} \frac{1}{\min\left\{1, \frac{E_i^B - E}{E_i^A - E}\right\}}$$

$$\frac{f(A)}{f(B)} = \frac{E_i^B - E}{E_i^A - E} = \frac{P(B \to A)}{P(A \to B)}.$$

- Case 2: $q_1 = q_2$. In this case, from equation 2.17 we have

$$\frac{P(A \to B)}{P(B \to A)} = \frac{E_i^A - E}{E_i^B - E}$$

and

$$\frac{f(A)}{f(B)} = \frac{E_i^B - E}{E_i^A - E} = \frac{P(B \to A)}{P(A \to B)}.$$

Appendix E

The Quantum Cavity Method

In this section we review the continuous time quantum cavity method[40, 43]. The purpose of this section is to motivate and then describe the equations which we have solved numerically in our study of random 3 regular Max-Cut. We first derive the cavity equations for a Hamiltonian with two local interactions on a finite tree. We then briefly mention the approximations which are used to investigate the infinite size limit for homogeneous Hamiltonians defined on random regular graphs (homogeneity means that the interaction is the same on each edge of the graph). Finally we describe the numerical procedure which was used in chapter 5 to solve the quantum cavity equations for Max-Cut in the 1RSB ansatz with Parisi parameter $m = 0$.

E.1 The Quantum Cavity Method For a Two local Transverse Field Spin Hamiltonian on a Tree

We now review the quantum cavity equations in the continuous imaginary time formulation [40].

We will consider transverse field spin Hamiltonians of the form

$$H = H_0 - \lambda \sum_{i=1}^{n} \sigma_x^i$$

where H_0 is diagonal in the Pauli basis. We will further restrict H_0 to be two local, that is

$$H_0 = \sum_{(i,j) \in T} H_{ij}$$

where H_{ij} only acts nontrivially on spins i and j and the graph of interactions T is a tree.

We begin with the distribution ρ over paths which arises from the continuous imaginary time path integral (recall from equation 1.17):

$$\rho(P) = \frac{1}{Z(\beta)} \Gamma^r dt_r dt_{r-1} ... dt_1 e^{-\int_0^\beta \langle P(t)|H_0|P(t)\rangle dt}.$$

Here we have changed notation slightly from chapter 1—we use $P(t)$ for the path in imaginary time and r for the number of flips (these were called $z(t)$ and m in chapter 1).

A path P can be specified by a number of flips r a sequence of bit strings $\{z_1, z_2, z_3, ..., z_{r+1} = z_1\}$ where z_{i+1} differs from z_i by a single bit flip, and times $\{t_1, t_2, ..., t_r\}$ at which transitions occur. A path P of n spins can also be specified as a collection of n one-spin paths $P^{(i)}$ for $i \in \{1, ..., n\}$ where $P^{(i)}$ is specified by $r(i)$ (the number of transitions in the path of the ith spin), a single bit $b(i) \in \{0,1\}$ which is the value taken by the spin at time $t = 0$, and a list of transition times $\{t_1^{(i)}, t_2^{(i)}, ..., t_{r(i)}^{(i)}\}$.

Then we can also write [1]

$$\rho(P) = \frac{1}{Z(\beta)} \left[\prod_{i=1}^{n} \lambda^{r(i)} dt_1^{(i)} dt_2^{(i)} ... dt_{r(i)}^{(i)} \right] e^{-\int_0^\beta \langle P(t)|H_0|P(t)\rangle dt}.$$

The quantum cavity equations allow one to determine $\mu_{i \to j}(P^{(i)})$, the marginal distribution of the path of spin i when the interaction H_{ij} between spins i and j is removed from H. This marginal distribution is defined through

$$\mu_{i \to j}(P^{(i)}) = \frac{1}{N_{i \to j}} \sum_{P_k : k \neq i} \rho(P) e^{\int_0^\beta \langle P(t)|H_{ij}|P(t)\rangle dt}.$$

where $N_{i \to j}$ is a normalizing factor.

We now derive the quantum cavity equations which are a closed set of equations for the cavity distributions $\{\mu_{i \to j}\}$. We begin by writing

$$\mu_{i \to j}(P^{(i)}) = \frac{1}{N_{i \to j}} \sum_{P_k : k \neq i} \rho(P) e^{\int_0^\beta \langle P(t)|H_{ij}|P(t)\rangle dt}$$

$$= \frac{1}{N_{i \to j}} \left(\lambda^{r(i)} dt_1^{(i)} dt_2^{(i)} ... dt_{r(i)}^{(i)} \right) \sum_{P^{(k)} : k \neq i} \left(\left[\prod_{k \neq i} \lambda^{r(k)} dt_1^{(k)} dt_2^{(k)} ... dt_{r(k)}^{(k)} \right] \right.$$

$$\left. \cdot e^{-\int_0^\beta \langle P(t)|H_0 - H_{ij}|P(t)\rangle dt} \right),$$

where $N_{i \to j}$ is a normalizing constant. Now let $T^{i \to j}$ be the component of G which is connected to spin i after removing the edge between i and j (see figure E-1). Then

---

[1] Note that the time ordering which is enforced in equation 1.15 simply says that we don't "double count" paths. This constraint is also enforced by making a change of variables to an ordered set of times for each spin.
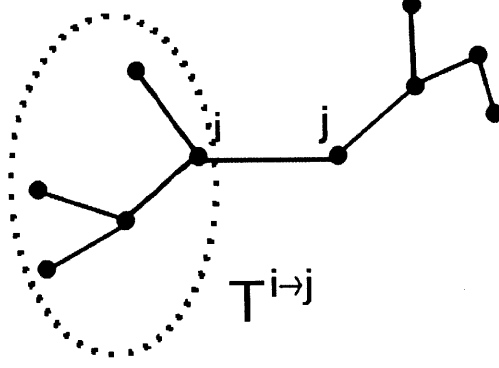
Figure E-1: The subtree $T^{i \to j}$.

$$\mu_{i \to j}(P^{(i)})$$

$$= C_{i \to j} \left(\lambda^{r^{(i)}} dt_1^{(i)} dt_2^{(i)} ... dt_{r(i)}^{(i)} \right) \sum_{P^{(k)}: k \in T^{i \to j}, k \neq i} \left(\left[\prod_{k \in T^{i \to j}, k \neq i} \lambda^{r^{(k)}} dt_1^{(k)} dt_2^{(k)} ... dt_{r(k)}^{(k)} \right] \right.$$

$$\left. \cdot e^{- \int_0^\beta \langle P(t) | \sum_{(l,m) \in T^{i \to j}} H_{lm} | P(t) \rangle dt} \right). \qquad (E.1)$$

where $C_{i \to j}$ is a constant. From this we can finally obtain the quantum cavity equations

$$\mu_{i \to j}(P^{(i)})$$

$$= \frac{1}{z_{i \to j}} \left(\lambda^{r^{(i)}} dt_1^{(i)} dt_2^{(i)} ... dt_{r(i)}^{(i)} \right) \sum_{\substack{P^{(k)} : \\ k \in \partial i \backslash j}} \left[\prod_{k \in \partial i \backslash j} \mu_{k \to i}(P^{(k)}) \, e^{- \int_0^\beta \langle P(t) | H_{ik} | P(t) \rangle dt} \right] \quad (E.2)$$

where $z_{i \to j}$ is the normalizing constant.

Now we show how to compute quantities such as the magnetization using the cavity distributions $\{\mu_{i \to j}\}$. Firstly, writing $\mu_i \left(P^{(i)} \right)$ for the marginal distribution of $P^{(i)}$ with respect to $\rho(P)$ we have

$$\mu_i \left(P^{(i)} \right) = \frac{1}{Z(\beta)} \sum_{P^{(k)}: k \neq i} \rho(P)$$

$$= \frac{1}{z_i} \sum_{P^{(k)}: k \in \partial i} \left[\prod_{k \in \partial i} \mu_{k \to i} \left(P^{(k)} \right) e^{- \int_0^\beta \langle P(t) | H_{ki} | P(t) \rangle dt} \right], \qquad (E.3)$$

with normalization z_i. (In the last line we have used equation E.1.) From this marginal

131

distribution μ_i one can compute, for example

$$\langle \sigma_x^i \rangle = \frac{\langle r \rangle_{\mu_i}}{\lambda \beta}, \tag{E.4}$$

where in the numerator one computes the average number of transitions r in $P^{(i)}$ with respect to the distribution μ_i. Similarly we have

$$\langle \sigma_z^i \rangle = \langle \frac{1}{\beta} \int_0^\beta P^{(i)}(t) dt \rangle_{\mu_i}.$$

We can also compute quantities which involve two neighboring spins. The joint distribution $\mu_{ij}(P^{(i)}, P^{(j)})$ of $P^{(i)}$ and $P^{(j)}$ is given by

$$
\begin{aligned}
\mu_{ij}(P^{(i)}, P^{(j)}) &= \frac{1}{Z(\beta)} \sum_{P^{(k)}: k \neq i, j} \rho(P) \\
&= \frac{1}{z_{ij}} \mu_{i \to j}(P^{(i)}) \mu_{j \to i}(P^{(j)}) e^{-\int_0^\beta \langle P(t) | H_{ij} | P(t) \rangle dt}, \tag{E.5}
\end{aligned}
$$

with normalization z_{ij}. These joint distributions (for each edge ij on the tree) can be used to compute the expectation value of H_0. By combining this with the estimators E.4 one can estimate the thermal average of the energy.

E.2 1RSB Quantum Cavity Equations with Parisi Parameter $m = 0$

We now specialize to the problem at hand–namely, Max-Cut on a random 3-regular graph. We briefly discuss the approximations which can be used in the cavity method to compute thermodynamic properties in the limit $n \to \infty$. We then go on to explain how we solve the 1RSB cavity equations with Parisi parameter $m = 0$.

The *replica symmetric* is the assumption that in the thermodynamic limit all the cavity distributions are the same ($\mu_{i \to j} = \mu$ for all directed edges $i \to j$). Roughly speaking this assumes that a random regular graph is modeled by an "infinite tree" which is obtained by assuming translation invariance for the recursion E.2. For a 3 regular antiferromagnet, this gives (from equation E.2)

$$\mu(P^{(0)}) = \frac{1}{Z} \left(\lambda^r dt_1 dt_2 ... dt_r \right) \sum_{P^{(1)}, P^{(2)}} \left[\mu\left(P^{(1)}\right) \mu\left(P^{(2)}\right) e^{-\int_0^\beta \langle P(t) | [\sigma_0 \sigma_1 + \sigma_0 \sigma_2] | P(t) \rangle dt} \right].$$

One can then attempt to solve for a distribution μ over paths which satisfies this recursion.

The 1RSB ansatz is described in [66] and [47] at the classical level and in the quantum case in reference [25]. The 1RSB ansatz makes the assumption that in the thermodynamic limit the distribution ρ for a random regular graph is a weighted

convex combination of distributions κ which have very little overlap (their support is on nonoverlapping sets of paths). In the 1RSB cavity method the Parisi parameter $m \in [0, 1]$ is used to assign the "states" κ different weights in the distribution. By choosing $m \in [0, 1]$ appropriately one obtains the correct weighting corresponding to the distribution ρ. The choice $m = 0$ corresponds to the assumption that each of these distributions κ is weighted evenly in the distribution ρ. In the 1RSB case with $m = 0$ we solve for a distribution $Q(\mu)$ over marginal distributions μ which has the property

- If μ_{i_1} and μ_{i_2} are drawn independently from Q, then $\tilde{\mu}$ defined by

$$
\tilde{\mu}(P^{(0)}) \tag{E.6}
$$
$$
= \frac{1}{Z} \left(\lambda^r dt_1 dt_2 ... dt_r \right) \sum_{P^{(1)}, P^{(2)}} \left[\mu_{i_1} \left(P^{(1)} \right) \mu_{i_2} \left(P^{(2)} \right) e^{- \int_0^\beta \langle P(t) | [\sigma_0 \sigma_1 + \sigma_0 \sigma_2] | P(t) \rangle dt} \right].
$$

is also distributed according to Q.

Numerical Solving for Q

We now describe how we solve for the distribution Q which has the property just described. "Population dynamics" methods such as the one we use have been used previously in the quantum case, but it has its origins in the classical method outlined in [48]. Our method uses the continuous imaginary time formulation of the quantum cavity method[40] but uses the single spin path integral sampling technique described in reference[18] which is also described in appendix B.

Since we cannot represent an arbitrary distribution $Q(\mu)$ in a finite amount of computer memory, we represent the distribution Q by a number N_D of representatives: that is, marginal distributions $\mu_1, \mu_2, ..., \mu_{N_D}$ which are each assigned an equal weight in the distribution. Our procedure is to repeat the following sweep iteratively until the fixed point Q is reached:

Distribution Sweep

For k from 1 to N_D do:

1. Choose i_1 and i_2 independently and at random from $1, ..., N_D$.

2. Generate $\tilde{\mu}$ according to equation E.6 using the Cavity Marginal Generating Subroutine (below) with inputs μ_{i_1} and μ_{i_2} and set $\mu_k = \tilde{\mu}$.

End do

The marginal generating subroutine takes as input two distributions μ_{i_1} and μ_{i_2} and outputs a new distribution $\tilde{\mu}$ according to equation E.6. Each cavity distribution

μ is stored as a list of N_R representative paths $P^{(1)}, P^{(2)}, ..., P^{(N_R)}$ which are given weights in the distribution $w^{(1)}, w^{(2)}, ..., w^{(N_R)}$ (with $\sum_i w^{(i)} = 1$).

Cavity Marginal Generating Subroutine

For k from 1 to N_R do

1. Choose a path $P^{(j_1)}$ according to μ_{i_1} and a path $P^{(j_2)}$ according to μ_{i_2}.

2. Generate a path $\tilde{P}$ according to the distribution

$$\frac{1}{A(P^{(j_1)}, P^{(j_2)})} \left(\lambda^{\tilde{m}} dt_1 dt_2 ... dt_{\tilde{m}} \right) e^{-\int_0^\beta \langle P(t) | [\tilde{\sigma}\sigma_1 + \tilde{\sigma}\sigma_2] | P(t) \rangle dt},$$

where $A(P^{(j_1)}, P^{(j_2)})$ is a normalizing constant. This is achieved by using the method outlined in reference [40] (with our procedure for sampling from the single spin path integral which was given in appendix B).

3. Set $\tilde{P}^{(k)} = \tilde{P}$ and let $\tilde{w}^{(k)} = A(P^{(j_1)}, P^{(j_2)})$.

End do

The new marginal distribution $\tilde{\mu}$ is represented by the paths $\tilde{P}^{(1)}, ..., \tilde{P}^{(N_R)}$ with weights

$$\frac{\tilde{w}^{(1)}}{\tilde{Z}}, \frac{\tilde{w}^{(2)}}{\tilde{Z}}, ..., \frac{\tilde{w}^{(N_R)}}{\tilde{Z}},$$

where $\tilde{Z} = \sum_k \tilde{w}^{(k)}$.

Estimators

The distribution $Q(\mu)$ is a probability distribution over marginal cavity distributions μ. Each cavity distribution is associated with a state κ and we have assumed that the distribution ρ is approximated by the uniform distibution ρ over these states. Recall that in order to compute single site properties such as the expected values of σ_z or σ_x we need to obtain the true marginal distribution from the cavity distribution (see equation E.3 on a tree). To do this we follow the procedure

---

True Marginal Generating Procedure

- Choose i_1, i_2 and i_3 independently and at random from $1, ..., N_D$.

- Generate $\tilde{\mu}$ as follows using the elements μ_{i_1}, μ_{i_2} and μ_{i_3} of the current population:

For k from 1 to N_R do

1. Sample paths $P^{(j_1)}$ according to μ_{i_1}, $P^{(j_2)}$ according to μ_{i_2} and $P^{(j_3)}$ according to μ_{i_3}

2. Generate a path $\tilde{P}$ according to the distribution

$$\frac{1}{A(P^{(j_1)}, P^{(j_2)}, P^{(j_3)})} \left(\lambda^{\tilde{r}} dt_1 dt_2 ... dt_{\tilde{r}} \right) e^{-\int_0^\beta \langle P(t) | [\tilde{\sigma}(\sigma_1 + \sigma_2 + \sigma_3)] | P(t) \rangle dt},$$

where $A(P^{(j_1)}, P^{(j_2)}, P^{(j_3)})$ is a normalizing constant. This is achieved by using the method outlined in reference [40] (with our procedure for sampling from the single spin path integral which was given in appendix B).

3. Set $\tilde{P}^{(k)} = \tilde{P}$ and let $\tilde{w}^{(k)} = A(P^{(j_1)}, P^{(j_2)}, P^{(j_3)})$.

End do

The true marginal distribution $\tilde{\mu}$ (which since $m = 0$ is ideally sampled from the states uniformly) is represented by the paths $\tilde{P}^{(1)}, ..., \tilde{P}^{(N_R)}$ with weights

$$\frac{\tilde{w}^{(1)}}{\tilde{Z}}, \frac{\tilde{w}^{(2)}}{\tilde{Z}}, ..., \frac{\tilde{w}^{(N_R)}}{\tilde{Z}},$$

where $\tilde{Z} = \sum_k \tilde{w}^{(k)}$.

Using this procedure we obtain a true marginal $\tilde{\mu}$ which is uniformly sampled from the set of thermodynamic states κ of the system. We can then compute, for example

$$\langle \sigma_x^i \rangle = \frac{\mathbb{E}\left[\langle r \rangle_{\tilde{\mu}} \right]}{\lambda \beta}$$

$$\langle \sigma_z^i \rangle = \mathbb{E}\left[\langle \frac{1}{\beta} \int_0^\beta P^{(i)}(t) dt \rangle_{\tilde{\mu}} \right]$$

where the outer expectation values $\mathbb{E}[\cdot]$ are averaged with respect to the distribution over $\tilde{\mu}$ which is obtained by repeating the true marginal generating procedure many times. We can also use the true marginal generating procedure to compute the

Edwards Anderson order parameter

$$q_{EA} = \mathbb{E}\left[\left(\langle \frac{1}{\beta} \int_0^\beta P^{(i)}(t)dt \rangle_{\tilde{\mu}}\right)^2\right]. \tag{E.7}$$

Note that this quantity does not represent a physical average with respect to the Gibbs distribution; we first compute the magnetization for a given state $\tilde{\mu}$ and then square it before averaging over the states $\tilde{\mu}$. It should now be clear that we can also compute link quantities in a similar manner. This is done by first sampling cavity distributions according to Q and then using equation E.5. From this we can compute the internal energy and other two site expectation values.

E.3   Data with $N_R = 10000$

In figures E-2, E-3 and E-4 we compare the cavity method numerics taken with $N_R = 10000$ and $N_R = 15000$ path representatives (see the previous section for the definition of N_R). In both cases the number of distributions $N_D = 200$. The data with $N_R = 15000$ is the same data which is plotted in figures 5-1,5-2 and5-3.
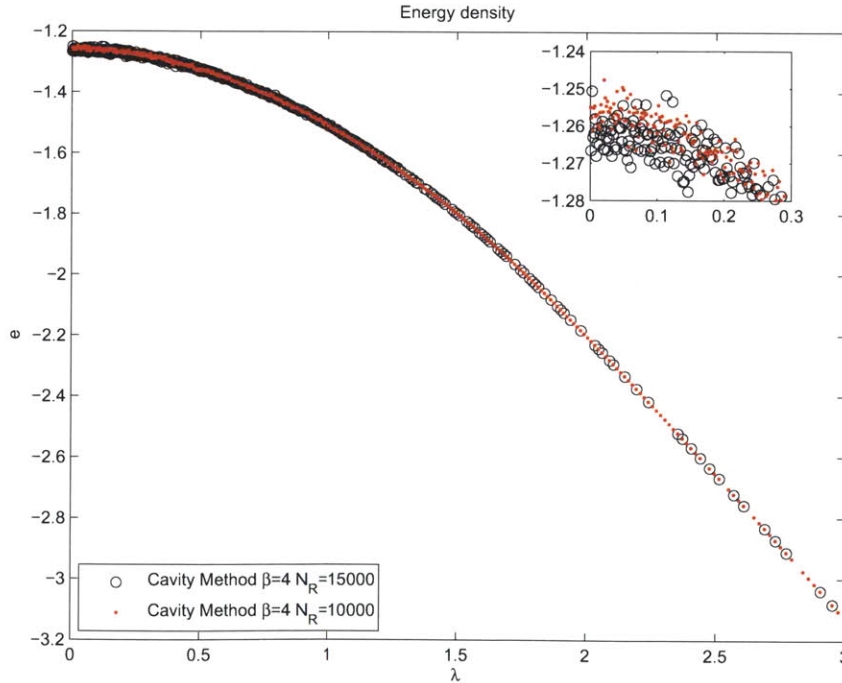

Figure E-2: The energy density computed with the cavity method at $\beta = 4$ and two different values of N_R. Note that the value at $\lambda = 0$ differs very slightly between the two data runs.
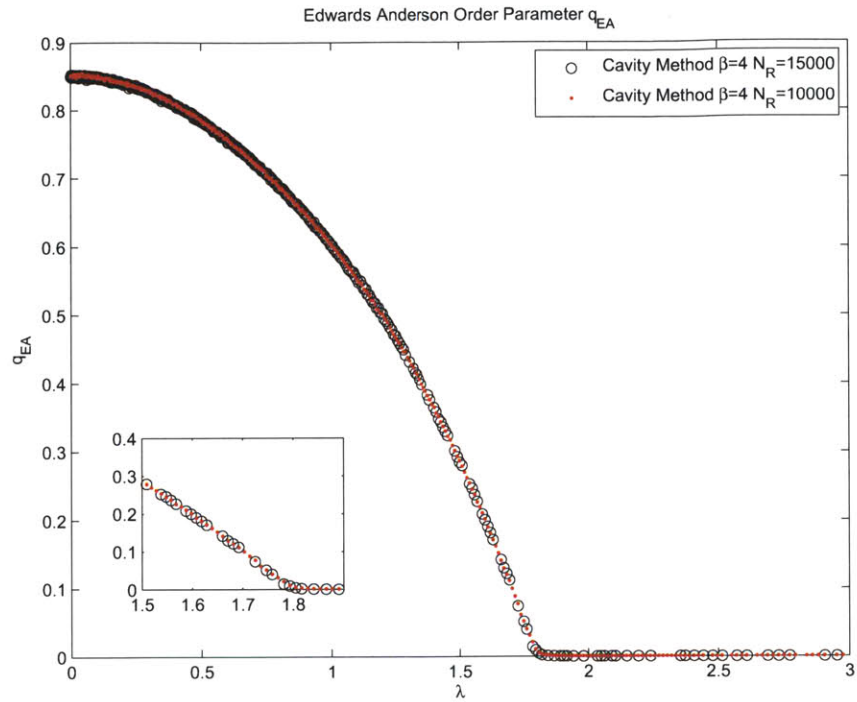
137

Figure E-3: The Edwards Anderson order parameter computed with the cavity method at $\beta = 4$ and two different values of N_R.
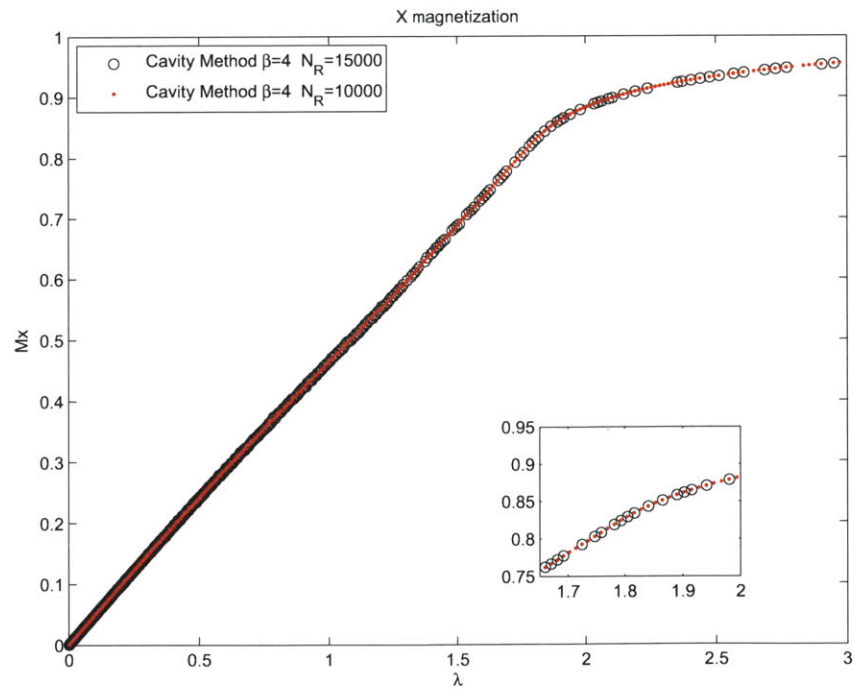

Figure E-4: The x magnetization computed with the cavity method at $\beta = 4$ and two different values of N_R.

Bibliography

[1] Dorit Aharonov, Wim van Dam, Julia Kempe, Zeph Landau, Seth Lloyd, and Oded Regev. Adiabatic quantum computation is equivalent to standard quantum computation. *SIAM JOURNAL OF COMPUTING*, 37:166, 2007. arXiv:quant-ph/0405098.

[2] B. Altshuler, H. Krovi, and J. Roland. Adiabatic quantum optimization fails for random instances of NP-complete problems. August 2009. arXiv:0908.2782.

[3] B. Altshuler, H. Krovi, and J. Roland. Anderson localization casts clouds over adiabatic quantum optimization. December 2009. arXiv:0912.0746.

[4] B. Altshuler, H. Krovi, and J. Roland. Anderson localization makes adiabatic quantum optimization fail. *Proceedings of the National Academy of Science*, 107:12446–12450, June 2010.

[5] M. H. S. Amin and V. Choi. First-order quantum phase transition in adiabatic quantum computation. *Physical Review A*, 80(6):062326–+, December 2009. arXiv:0904.1387.

[6] B.B. Beard. Extending monte carlo samples. *Nuclear Physics B - Proceedings Supplements*, 119:1012 – 1014, 2003. Proceedings of the XXth International Symposium on Lattice Field Theory.

[7] Piotr Berman and Marek Karpinski. On some tighter inapproximability results (extended abstract). In *Proceedings of the 26th International Colloquium on Automata, Languages and Programming*, ICAL '99, pages 200–209, London, UK, 1999. Springer-Verlag.

[8] Boettcher, S. Numerical results for spin glass ground states on bethe lattices: Gaussian bonds. *Eur. Phys. J. B*, 74(3):363–371, 2010.

[9] S. Bravyi, A. J. Bessen, and B. M. Terhal. Merlin-Arthur Games and Stoquastic Complexity. November 2006. arXiv:quant-ph/0611021.

[10] S. Bravyi, D. P. DiVincenzo, R. I. Oliveira, and B. M. Terhal. The Complexity of Stoquastic Local Hamiltonian Problems. June 2006. arXiv:quant-ph/0606140.

[11] Sergey Bravyi and Barbara M. Terhal. Complexity of stoquastic frustration-free hamiltonians. *SIAM J. Comput.*, 39(4):1462–1485, 2009. arXiv:0806.1746.

[12] Arnab Das and Bikas K. Chakrabarti (Eds.). *Quantum Annealing and Related Optimization Methods.* Springer, 2005.

[13] E. Farhi, J. Goldstone, D. Gosset, S. Gutmann, and P. Shor. Unstructured Randomness, Small Gaps and Localization. September 2010. arXiv:1010.0009.

[14] E. Farhi, J. Goldstone, D. Gosset, and H. B. Meyer. A Quantum Monte Carlo Method at Fixed Energy. *Computer Physics Communications*, 2011. In press. arXiv:0912.4271.

[15] E. Farhi, J. Goldstone, and S. Gutmann. Quantum Adiabatic Evolution Algorithms versus Simulated Annealing. January 2002. arXiv:quant-ph/0201031.

[16] E. Farhi, J. Goldstone, S. Gutmann, and D. Nagaj. How to Make the Quantum Adiabatic Algorithm Fail. *International Journal of Quantum Information*, 6(3):503–516, 2008. arXiv:quant-ph/0512159.

[17] E. Farhi and S. Gutmann. An Analog Analogue of a Digital Quantum Computation. December 1996. arXiv:quant-ph/9612026.

[18] Edward Farhi, Jeffrey Goldstone, David Gosset, Sam Gutmann, Harvey B. Meyer, and Peter Shor. Quantum adiabatic algorithms, small gaps, and different paths. *Quantum Information and Computation*, 11(3 & 4):0181–0214, 2011. arXiv:0909.4766.

[19] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. Quantum adiabatic evolution algorithms with different paths, 2002. arXiv:quant-ph/0208135.

[20] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, Joshua Lapan, Andrew Lundgren, and Daniel Preda. A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem. *Science*, 292:472–475, 2001. arXiv:quant-ph/0104129.

[21] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, and Michael Sipser. Quantum computation by adiabatic evolution, 2000. arXiv:quant-ph/0001106.

[22] Edward Farhi and Sam Gutmann. The functional integral constructed directly from the hamiltonian. *Annals of Physics*, 213(1):182 – 203, 1992.

[23] Daniel S. Fisher. Critical behavior of random transverse-field ising spin chains. *Phys. Rev. B*, 51(10):6411–6461, Mar 1995.

[24] Daniel S. Fisher and A. P. Young. Distributions of gaps and end-to-end correlations in random transverse-field ising spin chains. *Phys. Rev. B*, 58(14):9131–9141, Oct 1998.

[25] L. Foini, G. Semerjian, and F. Zamponi. Quantum Biroli-Mézard model: Glass transition and superfluidity in a quantum lattice glass model. *Physical Review B*, 83(9):094513–+, March 2011. arXiv:1011.6320.

[26] Lov K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, STOC '96, pages 212–219, New York, NY, USA, 1996. ACM.

[27] Lov K. Grover. A fast quantum mechanical algorithm for database search. In *STOC*, pages 212–219, 1996.

[28] M. Guidetti and A. P. Young. February 2011. arXiv:1102.5152.

[29] Eran Halperin, Dror Livnat, and Uri Zwick. Max cut in cubic graphs. *Journal of Algorithms*, 53(2):169 – 185, 2004.

[30] Wassily Hoeffding. Probability Inequalities for Sums of Bounded Random Variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[31] T. Hogg. Adiabatic quantum computing for random satisfiability problems. *Physical Review A*, 67(2):022314–+, February 2003. arXiv:quant-ph/0206059.

[32] S. Jansen, M.-B. Ruskai, and R. Seiler. Bounds for the adiabatic approximation with applications to quantum computation. *Journal of Mathematical Physics*, 48(10):102111–+, October 2007. arXiv:quant-ph/0603175.

[33] S. P. Jordan. *Quantum Computation Beyond the Circuit Model*. PhD thesis, MIT, 2008. arXiv:0809.2307.

[34] S. P. Jordan, D. Gosset, and P. J. Love. Quantum-Merlin-Arthur-complete problems for stoquastic Hamiltonians and Markov matrices. *Physical Review A*, 81(3):032331–+, March 2010. arXiv:0905.4755.

[35] T. Jörg, F. Krzakala, J. Kurchan, and A. C. Maggs. Simple Glass Models and Their Quantum Annealing. *Physical Review Letters*, 101(14):147204–+, October 2008. arXiv:0806.4144.

[36] T. Jörg, F. Krzakala, J. Kurchan, and A. C. Maggs. Quantum Annealing of Hard Problems. *Progress of Theoretical Physics Supplement*, 184:290–303, 2010. arXiv:0910.5644.

[37] T. Jörg, F. Krzakala, G. Semerjian, and F. Zamponi. First-Order Transitions and the Performance of Quantum Algorithms in Random Optimization Problems. *Physical Review Letters*, 104(20):207206–+, May 2010. arXiv:0911.3438.

[38] R. M. Karp. Reducibility Among Combinatorial Problems. In R. E. Miller and J. W. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103. Plenum Press, 1972.

[39] S. Knysh and V. Smelyanskiy. On the relevance of avoided crossings away from quantum critical point to the complexity of quantum adiabatic algorithm. May 2010. arXiv:1005.3011.

[40] F. Krzakala, A. Rosso, G. Semerjian, and F. Zamponi. Path-integral representation for quantum spin models: Application to the quantum cavity method and Monte Carlo simulations. *Physical Review B*, 78(13):134428–+, October 2008. arXiv:0807.2553.

[41] Florent Krzakala, Alberto Rosso, Guilhem Semerjian, and Francesco Zamponi. On the path integral representation for quantum spin models and its application to the quantum cavity method and to monte carlo simulations. *Physical Review B*, 78:134428, 2008. arXiv:0807.2553.

[42] L. D. Landau and E. M. Lifshitz. *Quantum Mechanics: Non-relativistic theory*, volume 3 of *Course of Theoretical Physics*. Pergamon Press, Oxford; New York, third edition, 1989, c1977. page 304.

[43] C. Laumann, A. Scardicchio, and S. L. Sondhi. Cavity method for quantum spin glasses on the Bethe lattice. *Physical Review B*, 78(13):134424–+, October 2008. arXiv:0706.4391.

[44] Y.-K. Liu. The Local Consistency Problem for Stoquastic and 1-D Quantum Systems. *ArXiv e-prints*, December 2007. arXiv:0712.1388.

[45] Albert Messiah. *Quantum Mechanics*. Dover Publications, 1999.

[46] Carl D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM, 2001.

[47] Marc Mezard and Andrea Montanari. *Information, Physics, and Computation*. Oxford University Press, Inc., New York, NY, USA, 2009.

[48] Marc Mezard and Giorgio Parisi. The cavity method at zero temperature. *Journal of Statistical Physics*, 111:1–34, 2003. arXiv:cond-mat/0207121.

[49] Remi Monasson, Riccardo Zecchina, Scott Kirkpatrick, Bart Selman, and Lidror Troyansky. Determining computational complexity from characteristic 'phase transitions'. *Nature*, 400(6740):133–137, July 1999.

[50] T. Neuhaus, M. Peschina, K. Michielsen, and H. de Raedt. Classical and quantum annealing in the median of three-satisfiability. *Physical Review A*, 83(1):012309–+, January 2011. arXiv:1103.3656.

[51] M. Newman and G. Barkema. *Monte Carlo Methods in Statistical Physics*. Clarendon Press, 1999.

[52] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2004.

[53] M. Ostilli and C. Presilla. The exact ground state for a class of matrix Hamiltonian models: quantum phase transition and universality in the thermodynamic limit. *Journal of Statistical Mechanics: Theory and Experiment*, 11:12–+, November 2006. arXiv:cond-mat/0610738.

[54] C. Presilla and M. Ostilli. Phase transition and annealing in quantum random energy models. February 2010. arXiv:1002.4409.

[55] N. V. Prokof'ev, B. V. Svistunov, and I. S. Tupitsyn. Exact quantum monte carlo process for the statistics of discrete systems. *ZH.EKS.TEOR.FIZ.*, 64:853, 1996. arXiv:cond-mat/9612091.

[56] A W Sandvik. A generalization of handscomb's quantum monte carlo scheme-application to the 1d hubbard model. *Journal of Physics A: Mathematical and General*, 25(13):3667–3682, 1992.

[57] A. W. Sandvik, R. R. P. Singh, and D. K. Campbell. Quantum monte carlo in the interaction representation — application to a spin-peierls model. *Physical Review B*, 56:14510, 1997. arXiv:9706046.

[58] R. Schützhold and G. Schaller. Adiabatic quantum algorithms as quantum phase transitions: First versus second order. *Physical Review A*, 74(6):060304–+, December 2006. arXiv:quant-ph/0608017.

[59] P.W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. *Foundations of Computer Science, Annual IEEE Symposium on*, 0:124–134, 1994.

[60] Masuo Suzuki. Relationship between d-dimensional quantal spin systems and $(d + 1)$-dimensional ising systems. *Progress of Theoretical Physics*, 56(5):1454–1469, 1976.

[61] Wim van Dam and Umesh Vazirani. Limits on quantum adiabatic optimization (unpublished manuscript).

[62] Ulli Wolff. Monte carlo errors with less errors. *Comput.Phys.Commun.* **156** *(2004) 143-153; Erratum-ibid.* **176** *(2007) 383.*

[63] A. P. Young, S. Knysh, and V. N. Smelyanskiy. Size Dependence of the Minimum Excitation Gap in the Quantum Adiabatic Algorithm. *Physical Review Letters*, 101(17):170503–+, October 2008. arXiv:0803.3971.

[64] A. P. Young, S. Knysh, and V. N. Smelyanskiy. First-Order Phase Transition in the Quantum Adiabatic Algorithm. *Physical Review Letters*, 104(2):020502–+, January 2010. arXiv:0910.1378.

[65] A.P Young. Private communication, 2011.

[66] F. Zamponi. Mean field theory of spin glasses. *ArXiv e-prints*, August 2010.

[67] F. Zamponi. Private communication, 2011.

[68] L. Zdeborová and S. Boettcher. A conjecture on the maximum cut and bisection width in random regular graphs. *Journal of Statistical Mechanics: Theory and Experiment*, 2:20–+, February 2010.