國立臺灣大學理學院物理學系

碩士論文

Department of Physics
College of Science
National Taiwan University

Master Thesis

矩陣積態在一維量子自旋系統的應用

Matrix Product State in One Dimensional Quantum Spin Systems

蘇育正

Yu-Cheng Su

指導教授：高英哲 博士

Advisor: Ying-Jer Kao, Ph.D.

中華民國 98 年 6 月

June, 2009

誌謝

研究所兩年，說長不長，說短不短，在這過程中所發生的點點滴滴，也非三言兩語便能道盡。這段時間的經歷讓我體認到一件相當重要的事情，一篇論文絕非僅僅一人便能完成；研究過程中遭遇的種種困難與挫折，有人可以討論與分擔是相當幸運的一件事。

首先我要謝謝我的指導老師高英哲教授，感謝他的指導，讓我瞭解如何從事研究以及從事研究時應有的嚴謹態度；與老師討論也獲益良多，遇到瓶頸時，與老師討論常能使問題迎刃而解。另外也要謝謝 Boston University 的 Anders Sandvik 教授。在他訪問台大期間，熱心地與我們討論以及給予我們建議，幫助我們克服瓶頸。

另外我要謝謝 group 裡的成員。謝謝昱均、俊谷、承瑋在我剛開始研究生生活時給予我的幫助，以及謝謝信智、揚智、雅琳、柯達的討論和幫忙。也要謝謝許多老師、B92 的同學們以及佳寧，在我遇到挫折時爲我打氣，並且給我建議。

我也要謝謝我打工的補習班的聯立老師與主任，還有一起工作的同事們：孝武、裕翔、澍寰、玉女和念慈姐。如果不是你們給予我經濟上的幫助與工作上的支援甚至是分享人生經驗，我想我很難走到今天這一步。

最後，我要謝謝我的家人，因爲你們的支持與鼓勵，我才能擁有今天的成就，謝謝你們。

摘要

在一維量子自旋系統中，矩陣積態可作爲變量數值模擬的試驗波函數。在此研究中，我們展示了兩種建構矩陣積態的方法，這些方法源自於密度矩陣重整群與量子資訊理論。我們發展了兩種在一維量子系統中矩陣積態的演算法，分別爲隨機最佳化的量子蒙地卡羅變量模擬 (Variational quantum Monte Carlo simulations with stochastic optimization)與時間演化間隔消除法 (Time-evolving block decimation)。我們推廣了隨機最佳化的方法至開放邊界 (open boundary condition)並且探討了伊辛模型加入橫向磁場與海森堡模型。另外，我們處理了無限長的伊辛模型加入橫向磁場，我們的結果顯示量子糾纏 (quantum entanglement)與量子相變息息相關。

關鍵字：矩陣積態、量子蒙地卡羅、隨機最佳化、時間演化間隔消除法、量子糾纏

Abstract

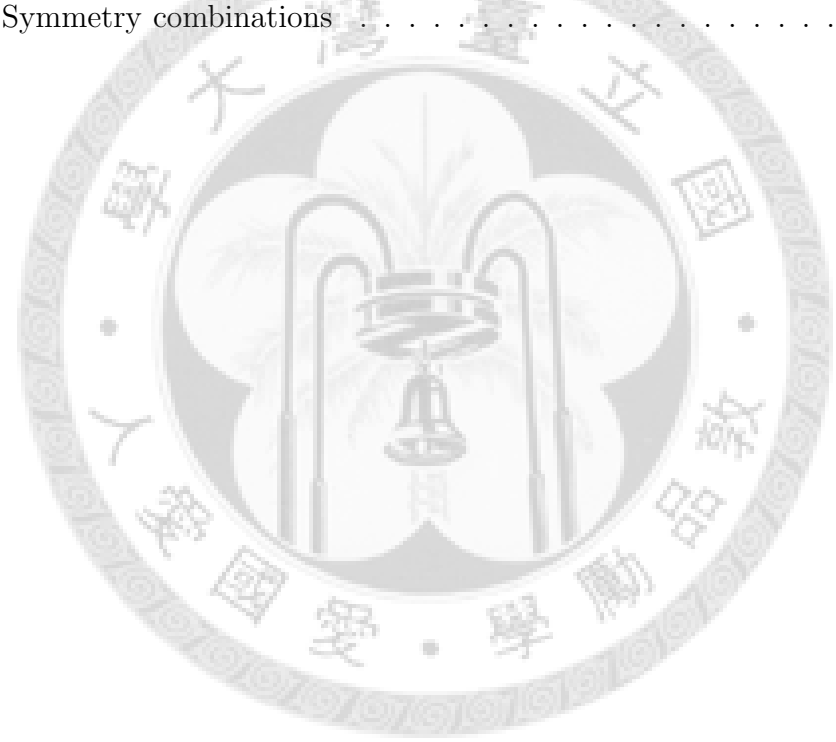
In one-dimensional quantum spin systems, the matrix product states (MPS) can be used as a trail wave function for variational numerical simulations. In this thesis, we investigate the construction of MPS which is related to the density matrix renormalization group (DMRG) and the Quantum information theory (QIT). We develop two algorithms, variational quantum Monte Carlo (QMC) simulations with stochastic optimization [1] and time-evolving block decimation (TEBD) [2, 3], in one dimensional systems. We generalize QMC with stochastic optimization to the open boundary condition and study the transverse Ising model and Heisenerg model. We also applied the infinite TEBD algorithm [4] to the infinite transverse Ising model and demonstrate that entanglement is a key ingredient in the quantum phase transition.

Keywords: matrix product state, quantum Monte Carlo, stochastic optimization, TEBD, iTEBD, entanglement

Contents

List of Figures

List of Tables



Chapter 1

Introduction

1.1 Overview

The study of low-dimensional strongly correlated quantum systems has been one of the intriguing topics in theoretical condensed matter physics. However, because of the complexity of the systems, few analytically exact solutions are available; on the other hand, numerical methods are able to extract useful properties from these systems. The density matrix renormalization group (DMRG) invented by White [5, 6] has provided accurate descriptions of the ground state in one dimensional strongly correlated Hamiltonians [7]. DMRG is developed to extend Wilson's numerical renormalization group [8] to general quantum lattice systems. The failure of Wilson's RG is identified as the inability to take the boundary conditions into account. DMRG first only apply to systems with the open boundary condition and later it is extended to the periodic boundary condition from the quantum information perspective [9, 10].

The success of DMRG is related to the matrix product states (MPS) [11] . MPS represents the wave function of the system in the DMRG method, and according to quantum information perspective [2], MPS can properly account for entanglement in one dimensional strongly correlated quantum systems. On the other hand, DMRG in dimensions higher than one is insufficient to properly account for entanglement of the system [12]. A generalization of MPS, projected enatngled-pair states (PEPS) [13], is proposed to deal with the simulation in two or higher spatial dimensions. MPS and PEPS can be used as trial wave functions [1, 13, 14, 15, 16, 17] in variational methods. In Ref [1], MPS is combined with quantum Monte Carlo (QMC) method to sample the physical states instead of summing over all states in order to reduce the computaional cost.

Over the last few years, various algorithms for simulating the time evolution of one dimensional quantum systems have been developed [18]. One of the algorithms called time-evolving block decimation (TEBD) [2, 3] uses MPS and emphasizes the connection between the computaioanl cost of a simulation and the amount of entaglement in the system. The scheme is later recast into the language of DMRG [19, 20]. TEBD can be applied efficiently to the quantum system which has a small amount of entanglement. A generalization of TEBD, infinite TEBD [4], is able to simulate one dimensional quantum lattice problems in the thermodynamic limit. The method simulates the infinite system directly without resorting to extrapolation.

1.2 Lattice models

1.2.1 Transverse Ising model

In this thesis, we study the one dimensional quantum transverse Ising model for the finite system and the infnite system. The transverse Ising model is the simplest quantum spin model to exhibit a quantum phase transition [21]. The hamiltonian is given by

$$H = -(\sum_i \sigma_i^z \sigma_{i+1}^z + h\sigma_i^x), \tag{1.1}$$

where σ^x and σ^z are the Pauli matrices, and sum over i is over all sites. For $h < 1$, the ground state of this system has long-range Ising order in the z-direction. For $h > 1$, the ground state becomes disorder in the z-direction. A quantum phase transition occurs when the system goes from $h < 1$ to $h > 1$. The critical point is at $h = 1$. The transverse Ising model provides a clear evidence that the lattice is most entangled at the critical point.

1.2.2 Heisenberg model

We also investigate the antiferro Heisenberg model with the next nearest neighbor interaction. The Hamiltonian is given by

$$H = \sum_i J_1 \mathbf{S}_i \cdot \mathbf{S}_{i+1} + J_2 \mathbf{S}_i \cdot \mathbf{S}_{i+2}, \tag{1.2}$$

where $\mathbf{S} = \frac{\hbar}{2}\sigma$ and J_1, $J_2 > 0$. The model is invariant under a global $SU(2)$ rotation, so the total spin should be conserved. For $J_2 = 0$, the model can be solved exactly by Bethe ansatz [22]. When $J_2 > 0$, it can be solved by taking the ground state as a superposition of the nearest neighbor valence

bond states [23]. For $J_1, J_2 > 0$, there is a competition between the nearest neighbor interaction and the next nearest neighbor interaction which is called *frustration*. This causes the "sign problems" [24, 25] in many QMC methods which originates from the minus sign in the wave function when interchanging two fermions. An illustration of frustration is shown in Fig. 1.1.

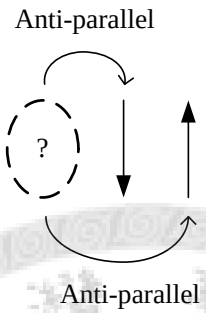


Figure 1.1: An illustration of frustration. The spins tend to lie anti-parallel to minimize the energy and form a competition between the nearest neighbor interaction and the next nearest neighbor interaction.

The thesis is organized as follows. In Chapter 2, we review the construction of MPS from two perspectives and discuss the connection between them. In Chapter 3, we review the Monte Carlo method and the stochatic optimization. We generalize this QMC method to the open boundary condition and study the transverse Ising model and the Heisenberg model. In chapter 4, we review the TEBD and iTEBD method in detail for imaginary time evolution. The infinite transverse Ising model is studied and the results imply a connection between entangelment and quantum phase transition. In chapter 5, we summarize the results and make a conclusion.

Chapter 2

Matrix Product States

In this chapter, we present the formulation of the matrix product states (MPS) from the density matrix renormalization group (DMRG) aspect and from the quantum information theory (QIT) [26] aspect. We will also discuss the notion of entanglement and use it to justify the use of MPS as a trial wave function.

2.1 MPS from the DMRG point of view

The development of the density matrix renormalization group method [5, 6] has enabled us to analyze and understand one-dimensional quantum many body systems with high accuracy . We shall describe the main idea of DMRG without going into too much details. First of all, we start from a smaller system which is small enough that we can diagonalize its Hamiltonian. This system is labeled as the *superblock* and it is divided into the *system block* and the *enviroment block*. The goal is to find a set of states of the system block which can optimally represent the superblock. We construct the reduced density matrix for the system block and diagonalize it, keeping only a number of states as basis states (e. g., D states) by dropping off the states with smaller eigenvalues in the reduced density matrix, as they are less likely to be accessed. The Hamiltonian of the system block are transformed to these basis states. Then we add a single spin to the system block and use the transformed Hamiltonian together with the added spin to construct the enviroment block; thus, the new superblock can be formed. Repeat this spin-adding procedure recursively until the system reach the desired size. In practice, the eigenvalues of the density matrix decrease rapidly so that the truncation errors are small. The name *density matrix renormalization group* reflects the fact that we keep those most relevant states in the density matrix.

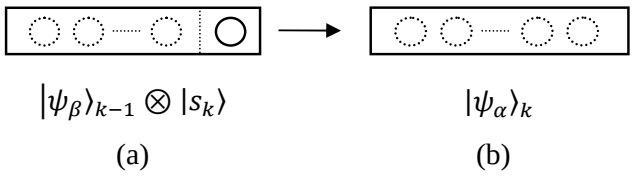


Figure 2.1: (a) Add kth new spin $|s_k\rangle$ to the system block basis states $|\psi_\beta\rangle_{k-1}$ containing $k-1$ sites. (b) Form the new super block and construct the reduced density matrix for the new system block keeping only D states to form a new set of basis.

The theoretical foundation of the success of DMRG is pointed out by Östlund and Rommer [11]. Their work shows that the DMRG construction is closely related to position-dependent matrix product state. At each recursion step, DMRG is a particularly effective way to generate a $D \times D$ projection operator $A^k_{\alpha,\beta}(s_k)$, which projects these states $|\psi_\beta\rangle_{k-1} \otimes |s_k\rangle$ to a larger system with a set of new basis states $|\psi_\alpha\rangle_k$. That is (Fig. 2.1(a) to Fig. 2.1(b))

$$|\psi_\alpha\rangle_k = \sum_{\beta=1}^{D} \sum_{s_k} A^k_{\alpha\beta}(s_k) |\psi_\beta\rangle_{k-1} \otimes |s_k\rangle. \tag{2.1}$$

For a state $|\Psi\rangle$ in an one-diensional system containing n spins with periodic boundary conditions($s_{n+1} = s_1$), we can construct a matrix product state using Eq. (2.1) recursively as shown in Fig. 2.2. $|\Psi\rangle$ is transformed into spin basis

$$|\Psi\rangle = \sum_{s_1, s_2, \ldots, s_n} tr(A^1(s_1) A^2(s_2) \ldots A^n(s_n)) |s_1 s_2 \ldots s_n\rangle, \tag{2.2}$$

where $tr(\cdots)$ is the matrix trace. This state contains $2nD^2$ parameters ranther than 2^n ones. Furthermore, if the system has translational symmetry,

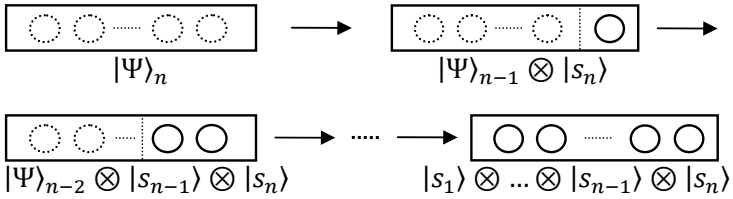


Figure 2.2: The recursive steps to transform the state into spin basis

the matrices A become independent of position. The wave function is now written

$$|\Psi\rangle = \sum_{s_1,s_2,\ldots,s_n} tr(A(s_1)A(s_2)\ldots A(s_n))|s_1 s_2 \ldots s_n\rangle. \tag{2.3}$$

This is the matrix product state we are going to use as a trial wave function to study the ground state of quantum spin systems, but before that, we would like to discuss more about the matrix product states and give the criteria for the validity of MPS as a trial wave function for such approximation.

2.2 MPS from the QIT point of view

To describe the state of n interacting quantum systems, it requires $\mathcal{O}(exp(n))$ parameters. We can see immediately that simulating such systems is a non-polynomial problem, and this makes the problem intractable for nowadays computer resources. Therefore, finding an effective approximation method to simulate quantum interacting systems is valuable to our understanding of quantum many body systems.

Recently, Quantum Information Theory (QIT) has provided a different point of view to describe condensed matter systems. A theory of entanglement has been established and has offered new insights in dealing with quantum many body systems. It is possible to reduce the number of parameters needed to describe a slightly entangled system from a non-polynomial order to a polynomial order.

In this section, we will first introduce two useful theorems in linear algebra which we are going to use to develop the theory, and then the concept of entanglement will be introduced. Finally, we construct a variant form of matrix product states and compare it with the form discussed in the previous section .

2.2.1 Tools in Linear Algebra

Linear algebra has played an important role in the formulation of quantum mechanics. The rich properties of linear algebra has helped us express the theory in a concise language, and this is especially true in the field of Quantum Information Theory. We present two important tools, singular value decomposition and Schmidt decomposition, which are used in QIT but not the standard materials in quantum mechanics textbooks. No attempt is made at mathematical rigor. These tools are suitable for the study of composite quantum systems, which are the key structures in quantum many body systems.

Singular value decomposition. Let A be a $m \times n$ matrix (suppose $n > m$). Then there exists an $m \times m$ unitary matrix U, an $n \times n$ unitary matrix V and an $m \times n$ diagonal matrix D such that

$$A = UDV^T = U \underbrace{\begin{pmatrix} d_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & d_2 & 0 & 0 & \dots & 0 \\ \vdots & & \ddots & & & \vdots \\ 0 & \dots & 0 & d_m & \dots & 0 \end{pmatrix}}_{n} V^T,$$

where the diagonal elements of D are called singular values of A, and $d_1 \geq d_2 \geq \ldots \geq d_m \geq 0$. V^T is the transpose of V.

The singular value decomposition is a useful technique to factorize a general matrix into two unitary operators and a diagonal operator. Dealing with these operators is easier than dealing with a general matrix. Here is a numerical example

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}. \tag{2.4}$$

Schmidt decomposition. A pure state $|\Psi_{AB}\rangle$ of a composite system , AB, can be decomposed into two orthonormal states $|i_A\rangle$ for system A, and $|i_B\rangle$ for system B such that

$$|\Psi_{AB}\rangle = \sum_i \lambda_i |i_A\rangle |i_B\rangle.$$

where $\lambda_i \geq 0$ and $\sum_i \lambda_i^2 = 1$ in order to normalize the state$|\Psi\rangle$. The number of λ_i is called Schmidt rank. The vectors $|i_A\rangle$ and $|i_B\rangle$ of the Schmidt decomposition are determined up to a phase, while Schmidt rank and λ_i are uniquely determined.

The Schmidt decomposition can be proved from the singular value decomposition. We will make use of this fact to perform our simulation in section 4.3. The proof is outlined and a numerical example is given below.

Let $|j_A\rangle$ and $|k_B\rangle$ be any orthornormal basis for systems A and B. Then $|\Psi_{AB}\rangle$ can be written

$$|\Psi_{AB}\rangle = \sum_{j,k} A_{jk} |j_A\rangle |k_B\rangle,$$

and by the singular value decomposition of A_{jk}

$$|\Psi_{AB}\rangle = \sum_{j,k} U_{ji} D_{ii} V_{ik}^T |j_A\rangle |k_B\rangle.$$

Defining $|i_A\rangle = \sum_j U_{ji}|j_A\rangle$, $|i_B\rangle = \sum_k V_{ki}|k_B\rangle$ and $D_{ii} = \lambda_i$ we get

$$|\Psi_{AB}\rangle = \sum_i \lambda_i |i_A\rangle|i_B\rangle, \tag{2.5}$$

where $\lambda_1 \geq \lambda_2 \geq \ldots \geq 0$. This completes our proof.

The numerical example is given according to Eq. (2.4)

$$\begin{aligned}|\Psi_{AB}\rangle &= |0_A\rangle|0_B\rangle + 2|0_A\rangle|1_B\rangle + 2|1_A\rangle|0_B\rangle + |1_A\rangle|1_B\rangle \\ &= 3[\ \frac{1}{\sqrt{2}}(|0_A\rangle + |1_A\rangle)\ \frac{1}{\sqrt{2}}(|0_B\rangle + |1_B\rangle)\ ] \\ &\quad + [\ \frac{1}{\sqrt{2}}(|0_A\rangle - |1_A\rangle)\ \frac{1}{\sqrt{2}}(-|0_B\rangle + |1_B\rangle)\ ].\end{aligned}$$

This suggests that we can perform the Schmidt decomposition through the singular value decomposition, and this approach will be used in our simulation to find λ_i in section 4.3.

2.2.2 Entanglement

For a state $|\Psi_{AB}\rangle$, it is entangled if it cannot be decomposed into a product state $|\Psi_A\rangle|\Psi_B\rangle$. An example below illustrates the notion of entanglement.

$$\begin{aligned}|\Psi_{AB}\rangle &= |0_A\rangle|1_B\rangle + |1_A\rangle|1_B\rangle \\ &= (|0_A\rangle + |1_A\rangle)|1_B\rangle = |\Psi_A\rangle|\Psi_B\rangle \qquad \text{(product state)} \\ |\Psi_{AB}\rangle &= |0_A\rangle|0_B\rangle + |1_A\rangle|1_B\rangle \neq |\Psi_A\rangle|\Psi_B\rangle \qquad \text{(entangled state)}\end{aligned}$$

We can clearly see that a state is entangled if and only if the Schmidt rank is greater than one, so the entanglement between system A and B can be characterized by the Schmidt rank. The larger the Schmidt rank, the more entangled the state, and vice versa. This is the key idea in justifying the approximation of reducing the number of parameters from the one that is exponential in the systme size n to the one that is linear in n.

2.2.3 A variant form of MPS

Let us now construct a variant form of MPS [2]. Consider a pure state $|\Psi\rangle$ for an n body system, we first perform the Schmidt decomposition to $|\Psi\rangle$ cutting the state at bond k between kth site and $(k+1)$th site as shown in Fig. 2.3

$$|\Psi\rangle = \sum_{\alpha_k=1}^{\chi_k} \lambda_{\alpha_k} |\psi_{\alpha_k}^{[1\cdots k]}\rangle |\psi_{\alpha_k}^{[k+1\cdots n]}\rangle \tag{2.6}$$

where χ_k is the Schmidt rank at bond k.

After cutting the state $|\Psi\rangle$ at bond k, we further cut the state $|\psi_{\alpha_k}^{[k+1\cdots n]}\rangle$ at bond $k+1$

$$|\psi_{\alpha_k}^{[k+1\cdots n]}\rangle = \sum_{\alpha_{k+1}=1}^{\chi_{k+1}} \lambda_{\alpha_{k+1}} |\psi_{\alpha_{k+1}}^{[k+1]}\rangle |\psi_{\alpha_{k+1}}^{[k+2\cdots n]}\rangle, \tag{2.7}$$

and expand the state $|\psi_{\alpha_k\alpha_{k+1}}^{[k+1]}\rangle$ in terms of the spin basis

$$|\psi_{\alpha_{k+1}}^{[k+1]}\rangle = \sum_{s_{k+1}} \Gamma_{\alpha_k\alpha_{k+1}}^{[k+1]s_{k+1}} |s_{k+1}\rangle. \tag{2.8}$$

Let us substitute Eq. (2.7) in Eq. (2.6)

$$|\psi_{\alpha_k}^{[k+1\cdots n]}\rangle = \sum_{\alpha_{k+1}=1}^{\chi_{k+1}} \sum_{s_{k+1}} \Gamma_{\alpha_k\alpha_{k+1}}^{[k+1]s_{k+1}} \lambda_{\alpha_{k+1}} |s_{k+1}\rangle |\psi_{\alpha_{k+1}}^{[k+2\cdots n]}\rangle, \tag{2.9}$$

and in a similar way, we can get

$$|\psi_{\alpha_k}^{[1\cdots k]}\rangle = \sum_{\alpha_{k-1}=1}^{\chi^{k-1}} \sum_{s_k} \lambda_{\alpha_{k-1}} \Gamma_{\alpha_{k-1}\alpha_k}^{[k]s_k} |\psi_{\alpha_{k-1}}^{[1\cdots k-1]}\rangle |s_k\rangle, \tag{2.10}$$

and therefore, Eq. (2.6) becomes

$$\begin{aligned} |\Psi\rangle = & \sum_{\alpha_{k-1}=1}^{\chi_{k-1}} \sum_{\alpha_k=1}^{\chi_k} \sum_{\alpha_{k+1}=1}^{\chi^{k+1}} \sum_{s_k,s_{k+1}} \\ & \lambda_{\alpha_{k-1}} \Gamma_{\alpha_{k-1}\alpha_k}^{[k]s_k} \lambda_{\alpha_k} \Gamma_{\alpha_k\alpha_{k+1}}^{[k+1]s_{k+1}} \lambda_{\alpha_{k+1}} |\psi_{\alpha_{k-1}}^{[1\cdots k-1]}\rangle |s_k\rangle |s_{k+1}\rangle |\psi_{\alpha_{k+1}}^{[k+2\cdots n]}\rangle. \end{aligned} \tag{2.11}$$

For a state $|\Psi\rangle$ with periodic boundary condition, repeat the above procedure recursively, and the state $|\Psi\rangle$ is decomposed into

$$|\Psi\rangle = \sum_{\alpha_1,\cdots,\alpha_n} \sum_{s_1,\cdots,s_n} \Gamma_{\alpha_n\alpha_1}^{[1]s_1} \lambda_{\alpha_1} \Gamma_{\alpha_1\alpha_2}^{[2]s_2} \lambda_{\alpha_2} \ldots \lambda_{\alpha_{n-1}} \Gamma_{\alpha_{n-1}\alpha_n}^{[n]s_n} \lambda_{\alpha_n} |s_1 s_2 \ldots s_n\rangle$$

where each α ranges from one to the corresponding Schmidt rank.

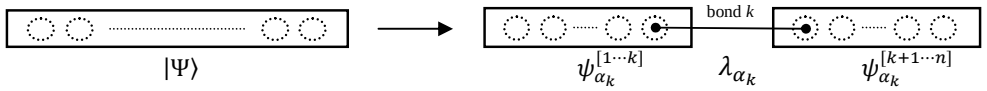


Figure 2.3: The Schmidt decomposition of $|\Psi\rangle$.

So far the expression is exact. The Schmidt rank χ_α for each bond grows exponentially with n. There are ground states in one-dimensional many body systems which are *slightly entangled* so that the values of λ_α decrease rapidly with α [2]. Hence, we can obtain a good approximation of the state $|\Psi\rangle$ by truncating λ_α at some α. We define $\chi_{max} \equiv max\{\chi_1, \chi_2, \ldots, \chi_n\}$, and for *slightly entangled* states, χ_{max} is truncated to some number D. In practice, D is determined through trial and error.

Thus, the state $|\Psi\rangle$ is now written

$$|\Psi\rangle = \sum_{\alpha_1,\cdots,\alpha_n=1}^{D} \sum_{s_1,\cdots,s_n} \Gamma^{[1]s_1}_{\alpha_n\alpha_1}\lambda_{\alpha_1}\Gamma^{[2]s_2}_{\alpha_1\alpha_2}\lambda_{\alpha_2}\ldots\lambda_{\alpha_{n-1}}\Gamma^{[n]s_n}_{\alpha_{n-1}\alpha_n}\lambda_{\alpha_n}|s_1 s_2 \ldots s_n\rangle, \tag{2.12}$$

where each α runs from one to D, or, to write it in a more compact form

$$|\Psi\rangle = \sum_{s_1,\cdots,s_n} tr(\Gamma^{[1]s_1}\lambda^{[1]}\Gamma^{[2]s_2}\lambda^{[2]}\ldots\lambda^{[n-1]}\Gamma^{[n]s_n}\lambda^{[n]})|s_1 s_2 \ldots s_n\rangle, \tag{2.13}$$

where $\Gamma^{[k]}$ is a $D \times D$ matrix and $\lambda^{[k]}$ is a D-dimensional vector. This state contains $n(2D^2 + D)$ parameters rather than 2^n ones.

To compare Eq. (2.2) and Eq. (2.13), we first note the similarity between Eq. (2.1) and Eq. (2.9). After the truncation step, Eq. (2.9) is written

$$|\psi^{[k+1\cdots n]}_{\alpha_k}\rangle = \sum_{\alpha_{k+1}=1}^{D} \sum_{s_{k+1}} \Gamma^{[k+1]s_{k+1}}_{\alpha_k\alpha_{k+1}}\lambda_{\alpha_{k+1}}|s_{k+1}\rangle|\psi^{[k+2\cdots n]}_{\alpha_{k+1}}\rangle.$$

Compare with Eq. (2.1)

$$|\psi_\alpha\rangle_k = \sum_{\beta=1}^{D}\sum_{s_k} A^k_{\alpha\beta}(s_k)|\psi_\beta\rangle_{k-1} \otimes |s_k\rangle.$$

We can immediately see that $A^k_{\alpha\beta}(s_k) \equiv \Gamma^{[k]s_k}_{\alpha\beta}\lambda_\beta$. Or, from Eq. (2.10), $A^k_{\alpha\beta}(s_k) \equiv \lambda_\alpha\Gamma^{[k]s_k}_{\alpha\beta}$. It depends on at which end the spin is added. Now the D states kept in DMRG iterating steps has a clear physical meaning. The number D is characterized by the amount of entanglement of the state; that is, the intrinsic computational cost is quantify by the entanglement. Also, the representation of the state $|\Psi\rangle$ in Eq. (2.2) and (2.13) are equivalent.

Before ending this chapter, we would like to present a form of MPS for a system with open boundary condition. The wave function is written

$$|\Psi\rangle = \sum_{s_1,s_2,\ldots,s_n} V^1(s_1)A^2(s_2)A^3(s_3)\ldots A^{n-1}(s_{n-1})V^n(s_n)|s_1 s_2 \ldots s_n\rangle \tag{2.14}$$

where $V^1(s_1)$ and $V^n(s_n)$ are D-dimensional vectors. This form can be derived from the above argument with a slight change in Eq. (2.12). We approximate this representation to simulate systems with open boundary conditions. More discussion about this form is in the next chapter.

Chapter 3

Variational quantum Monte Carlo simulation with stochastic optimization

When an analytical solution is hard to obtain, variational methods is a useful approach to getting an approximate solution. For a system described by the Hamiltonian H, we set up a form of the trial wave function $|\Psi\rangle$ with several parameters and implement the fact that

$$\frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} \geq E_{ground\ state}. \tag{3.1}$$

By adjusting the parameters, we lower the energy as much as we can and obtain the approximate ground state wave function for the system H.

In this chapter, MPS is used as a trial wave function. The matrix elements are regarded as parameters and are adjusted to approach the ground state. We will first discuss the Monte Carlo simulation, and then describe the optimization scheme. Finally, the above method is used to deal with one-dimensional quantum spin models.

3.1 Introduction to Monte Carlo simulation

Generally speaking, methods which involve making use of random numbers can be called Monte Carlo(MC) simulation. Monte Carlo simulation is an important method in condensed matter physics. Because of the complexity of many body systems, especially in strongly correlated systems where perturbation is unmanageable, many problems cannot be solved analytically. In this situation, Monte Carlo simulation is an effective method to extract statistical properties from the systems.

3.1.1 The Metropolis algorithm

The goal of the Monte Carlo simulation is to calculate the expectation value $\langle Q \rangle$, which is an ensemble average of some observable quantity Q,

$$\langle Q \rangle = \frac{\sum_{c_i} W_{c_i} Q_{c_i}}{\sum_{c_i} W_{c_i}}, \tag{3.2}$$

where Q_{c_i} is the value of the observable quantity in the ith configuration c_i and W_{c_i} is the weight of the ith configuration c_i. So the probability P_{c_i} for the ith configuration to occur is $P_{c_i} = \frac{W_{c_i}}{\sum_{c_i} W_{c_i}}$. To calculate $\langle Q \rangle$, we average the quantity Q_{c_i} over all states, $\langle Q \rangle = \sum_i P_{c_i} Q_{c_i}$; however, in reality, it is impossible to prepare all the configurations for large systems in our simulation. On the other hand, picking out the states randomly from the system is usually a poor method. In most cases, it only samples out a very small fraction of the states. The remedy of this problem is that instead of calculating the ensemble average, we calculate, through a *Markov process*, the average of the observable quantity by generating a chain of states which obey the probability of occurrence P_{c_i}.

A Markov process is a mechanism to generate a new state c_ν from a state c_μ through the transition probability $t(c_\mu \to c_\nu)$. The transition probability for a Markov process should depend only on the properties of the current configurations c_ν and c_μ ,and the transition probability should not vary over time.

To generate a chain of states which obey the probability of occurrence P_{c_i}, we first write down the master equation

$$\frac{dP_{c_\mu}}{dt} = \sum_{c_\nu} [P_{c_\nu} t(c_\nu \to c_\mu) - P_{c_\mu} t(c_\mu \to c_\nu)]. \tag{3.3}$$

For an equilibrium system, $\frac{dP_{c_\mu}}{dt} = 0$; furthermore, we enforce the *detail balance* condition such that

$$P_{c_\nu} t(c_\nu \to c_\mu) = P_{c_\mu} t(c_\mu \to c_\nu). \tag{3.4}$$

It is now clear that if the generated configurations are distributed according to P_{c_i}, the transition probability $t(c_\nu \to c_\mu)$ should satisfy

$$\frac{W_{c_\mu}}{W_{c_\nu}} = \frac{P_{c_\mu}}{P_{c_\nu}} = \frac{t(c_\nu \to c_\mu)}{t(c_\mu \to c_\nu)}. \tag{3.5}$$

Now the problem comes to "how do we manipulate the algorithm to satisfy the above condition?". Even if we can suggest many Markov processes

to generate a new state, we may not find one with the right transition probability. The solution to this problem is the concept called an *acceptance ratio*. We decompose the transition probability into two parts:

$$t(c_\mu \to c_\nu) = g(c_\mu \to c_\nu) a(c_\mu \to c_\nu), \tag{3.6}$$

where $g(c_\mu \to c_\nu)$ is the selection probability, and $a(c_\mu \to c_\nu)$ is the acceptance ratio (or called the "acceptance probability"). The selection probability is the probability that generates a new state c_ν from an old state c_μ. The acceptance ratio says that when we start from a state c_μ and our algorithm generates a new state c_ν, we should accept the change to the new state according to the acceptance ratio. So, no matter what Markov processes we choose to generate a new state which determine the selection probability, we can always satisfy the condition Eq. (3.2) as long as we adjust the acceptance ratio.

The most famous and widely used algorithm is called the *Metropolis algorithm* [27]. In the Metropolis algorithm the selection probability $g(c_\mu \to c_\nu)$ are all chosen to be equal. Therefore, Eq. (3.5) is now written

$$\frac{W_{c_\mu}}{W_{c_\nu}} = \frac{t(c_\nu \to c_\mu)}{t(c_\mu \to c_\nu)} = \frac{g(c_\nu \to c_\mu) a(c_\nu \to c_\mu)}{g(c_\mu \to c_\nu) a(c_\mu \to c_\nu)} = \frac{a(c_\nu \to c_\mu)}{a(c_\mu \to c_\nu)}. \tag{3.7}$$

The acceptance ratio is determined according to the weights. We define

$$a(c_\nu \to c_\mu) = min\{1, \frac{W_{c_\mu}}{W_{c_\nu}}\}. \tag{3.8}$$

In this chapter, we deal with one-dimensional $S = \frac{1}{2}$ quantum spin system, using the single-spin-flip method to generate a new configuration; in other words, we flip a spin at a time and calculate the acceptance ratio to decide whether to accept the flip or not. The Metropolis algorithm goes as follows.

1. Generate a spin configuration c_1 randomly as a starting configuration.
2. Flip a spin in the configuration c_i to generate a new one c_{i+1} .
3. Calculate the acceptance ratio

$$a(c_i \to c_{i+1}) = min\{1, \frac{W_{c_{i+1}}}{W_{c_i}}\}. \tag{3.9}$$

   Compare the acceptance ratio $a(c_i \to c_{i+1})$ with a random number $r \in [0, 1)$, if $a(c_i \to c_{i+1}) > r$, the configuration is update to c_{i+1}, otherwise it is rejected.

4. Go back to 2.

It is a standard algorithm in the Monte Carlo simulation.

3.1.2 Measuring observable quantities

As mentioned in the previous section, to measure the expectation value $\langle Q \rangle$, we have to run over all the configurations if we naively use Eq. (3.2). However, in the Metropolis algorithm, a chain of states is generated according to the occurrence probability of the states. That is, the number of a configuration generated is proportional to its weight in the system. So, we measure the observable quantity Q when the configuration is updated to a new one, and suppose the measurement performed M times, then the expectation value $\langle Q \rangle$ is now written

$$\langle Q \rangle = \frac{\sum_{i=1}^{M} Q_i}{M}. \tag{3.10}$$

In our simulation scheme, we flip a spin at a time to generate a new configuration. The difference between the old one and the new one is small. In other words, if we measure the observable quantity $\langle Q \rangle$ in each update, the measurement may be correlated. Then it is not a true Markov process. To solve this problem, we do not measure the observable quantity $\langle Q \rangle$ in each update. Instead, we perform the measurement after one Monte Carlo step. We define one Monte Carlo step as attempting to flip spins n times. The number n is the length of the spin chain.

In order to minimize the noise caused by the randomness in the Monte Carlo simulation, we use the blocking method called the *binning procedure*. First, we perform m Monte Carlo steps and perform the measurement after each Monte Carlo step. Second, we divide the measurements into m/b groups. Each group contains b measurement results, and the cumulative results are taken average over b. This step is called making a *bin*. Finally, the observable quantities are obtained by taking the bin average over m/b. The binning procedure is illustrated in Fig. 3.1.

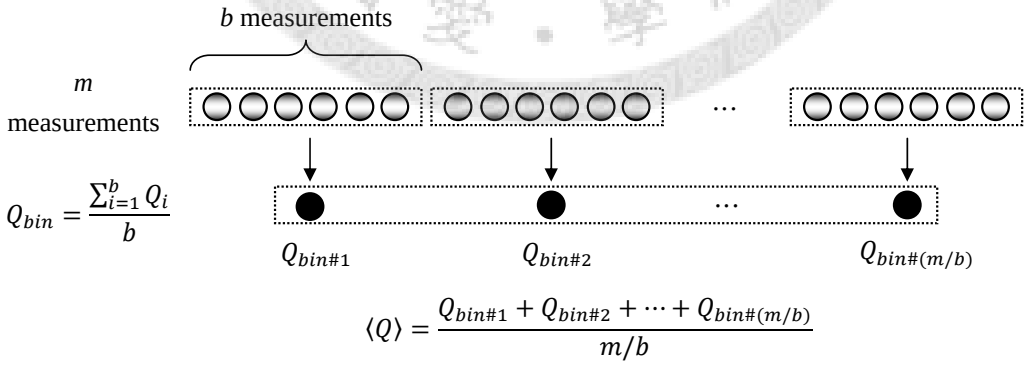


Figure 3.1: A sketch for the binning procedure.

3.2 Stochastic optimization

To obtain an approximate ground state wave function, our goal is to find the matrix elements that minimize the expectation value of the energy $E = \langle H \rangle$ by the derivatives of the energy with respect to the matrix elements. We illustrate the optimization scheme through a $S = \frac{1}{2}$ translationally invariant periodic spin system with n spins.

The first step is finding an appropriate form of E for Monte Carlo sampling. The wave function of a spin chain can be written

$$|\Psi\rangle = \sum_{s_1,s_2,\ldots,s_n} tr(A(s_1)A(s_2)\ldots A(s_n))|s_1 s_2 \ldots s_n\rangle = \sum_S F(S)|S\rangle, \quad (3.11)$$

where the spins $s_i = \pm 1$ are the eigenvalues of σ_i^z, $|S\rangle = |s_1, s_2, \ldots, s_n\rangle$, and $F(S)$ is the coefficient for state $|S\rangle$; $A(\pm 1)$ are two $D \times D$ general matrices, and they are taken to be real. The expectation value of energy can be expressed as

$$E = \frac{\sum_{S,S'} F(S)F(S')\langle S|H|S'\rangle}{\sum_S F^2(S)} = \frac{\sum_S F^2(S)E(S)}{\sum_S F^2(S)} = \langle E(S)\rangle, \quad (3.12)$$

where $E(S) = \sum_{S'} \frac{F(S')}{F(S)}\langle S|H|S'\rangle$. Identifying $F^2(S)$ as the weight of the configuration, we can use Metropolis algorithm, with the transition probability $t(S \to S') = min\{1, \frac{F^2(S')}{F^2(S)}\}$, to generate successive spin configurations by flipping a spin at a time. Through measuring $E(S)$ in the successive generated spin configurations, we then average over these measurements using the binning procedure to get $E = \langle E(S)\rangle$.

Second, the derivatives of the energy with respect to the matrix elements are also required. Taking derivatives directly, we have

$$\begin{aligned}\frac{\partial E}{\partial a_{ij}^s} &= \frac{\partial}{\partial a_{ij}^s}[\frac{\sum_{S,S'} F(S)F(S')\langle S|H|S'\rangle}{\sum_S F^2(S)}] \\ &= 2\langle \frac{1}{F(S)}\frac{\partial F(S)}{\partial a_{ij}^s}E(S)\rangle - 2\langle \frac{1}{F(S)}\frac{\partial F(S)}{\partial a_{ij}^s}\rangle\langle E(S)\rangle \end{aligned} \quad (3.13)$$

Calculating $\frac{\partial F(S)}{\partial a_{ij}^s}$ seems to be a formidable task; however, if we write out

those indices of the matrices, we can get

$$\begin{aligned}\frac{\partial F(S)}{\partial a^s_{ij}} &= \frac{\partial}{\partial a^s_{ij}} tr(A(s_1)A(s_2)\dots A(s_n)) \\ &= \sum_{k=1}^{n} \delta_{s,s_k} \frac{\partial}{\partial a^s_{ij}} [a^{s_1}_{\alpha_1\alpha_2} a^{s_2}_{\alpha_2\alpha_3} \cdots a^{s_{k-1}}_{\alpha_{k-1}\alpha_k} a^{s_k}_{\alpha_k\alpha_{k+1}} a^{s_{k+1}}_{\alpha_{k+1}\alpha_{k+2}} \cdots a^{s_{n-1}}_{\alpha_{n-1}\alpha_n} a^{s_n}_{\alpha_n\alpha_1}] \\ &= \sum_{k=1}^{n} \delta_{s,s_k} a^{s_{k+1}}_{j\alpha_{k+2}} a^{s_{k+2}}_{\alpha_{k+2}\alpha_{k+3}} \cdots a^{s_{n-1}}_{\alpha_{n-1}\alpha_n} a^{s_n}_{\alpha_n\alpha_1} a^{s_1}_{\alpha_1\alpha_2} a^{s_2}_{\alpha_2\alpha_3} \cdots a^{s_{k-1}}_{\alpha_{k-1}i} \\ &= \sum_{k=1}^{n} \delta_{s,s_k} [A(s_{k+1})A(s_{k+2}) \cdots A(s_n)A(s_1) \cdots A(s_{k-2})A(s_{k-1})]_{ji}.\end{aligned} \tag{3.14}$$

Define the matrices

$$M(k) = A(s_{k+1})A(s_{k+2}) \cdots A(s_n)A(s_1) \cdots A(s_{k-2})A(s_{k-1}), \tag{3.15}$$

the derivatives of the weight is rewritten

$$\frac{\partial F(S)}{\partial a^s_{ij}} = \sum_{k=1}^{n} \delta_{s,s_k} M_{ji}(k). \tag{3.16}$$

In order to minimize the energy, we have to calaulate $F(S)$, $F(S')$ and $M(k)$ during the update and the measurement. These terms involves many matrix multiplication and are time consuming. Here we implement a scheme to reduce the computational effort of calculating these terms for the update and the measurement. Noticing the form of $M(k)$, we introduce two more matrices $L(k) = A(s_k)A(s_{k+1}) \cdots A(s_n)$ and $R(k) = A(s_1)A(s_2) \cdots A(s_k)$ such that $M(k) = L(k+1)R(k-1)$. Also, we define$L(N+1) = R(0) = I$. The scheme goes as follows.

1. Calculate and store $L(2), L(3), \dots, L(n)$ based on a random initail configuration or the configuration from the previous run.

2. Start from s_1, we flip the spins sequentially to generate successive spin configurations. Each flip is based on Metropolis probability, $t(S \to S') = min\{1, \frac{F^2(S')}{F^2(S)}\}$. Using the cyclic property of the trace, we calculate $F(S')$ by $F(S'_k) = tr(A(-s_k)M(k))$.

3. After each attempt to flip the spin s_k, $L(k+1)$ is no longer needed, so we store $R(k) = R(k-1)A(s_k)$ in its place for future use.

4. After a sweep of spin updates, we start measurement from the spin s_n. Now the matrices $R(k)$ are all generated and stroed. Traverse the spin chain from $k = n$ to 1 for measuring observable quantities, and the matrices $L(k)$ are generated and stored.

5. Go back to 2 until the energy is converged.

After b measurements, we obtain a bin measurement of energy and the derivatives. We update the matrix elements after a bin is made. The bin measurement are used to update the matrix elements a^s_{ij}:

$$a^s_{ij} \to a^s_{ij} - q(p) \cdot r^s_{ij} \cdot sign(\frac{\partial E}{\partial a^s_{ij}}), \tag{3.17}$$

where $r^s_{ij} \in [0, 1)$ is random and $q(p)$ is the maximum change which decreases as a function of some number p. Each parameter is changed according to the sign of the derivative, but they are changed independently with a random but well bounded number. After each update of all the matrix elements, the matrices are normalized so that the largest element $|a^s_{ij}|$=1. In our simulation, we use the form $q(p) = q_0 p^{-\alpha}$, with $q_0 = 0.05 - 0.1$ and $\alpha = 0.7 - 0.8$.

As mentioned in section 3.1.2, the real output of the expextation value is the average over bins. Let p be the number of outputs; that is, $q(p)$ decreases with output numbers. Furthermore, when the energy approaches the minimum, the derivatives will become smaller, and the noise from randomness may affect the measurement of the derivatives. So, in order to minimize the noise, we use the form $b = b_0 p$ to increase measurements in a bin with $b_0 = 100 - 200$ and $20 - 40$ bins are averaged to make an output. Also, if we perform the above procedure to obtain optimized matrices, then, by using the optimized matrices as initial values of the matrices, we restart the above optimization but with a smaller q_0. Repeating this process produces better convergence.

A typical optimization process is shown in Fig. 3.2. Each point is an output with 20 bins and $b_0 = 100$. We notice that the energy decreased exponentially, and the variance of the energy was also reduced quickly during the optimization process. When using the matrices obtained in the previous run as the initial matrices, and with a smaller q_0, the energy reached a lower value.

The computaional cost of the above optimization scheme scales as ND^3, while DMRG scales as ND^6. Besides, it can generalized to the open boundary condition which scales as ND^2. The generalization is straightforward. We discuss the scheme in the next section.

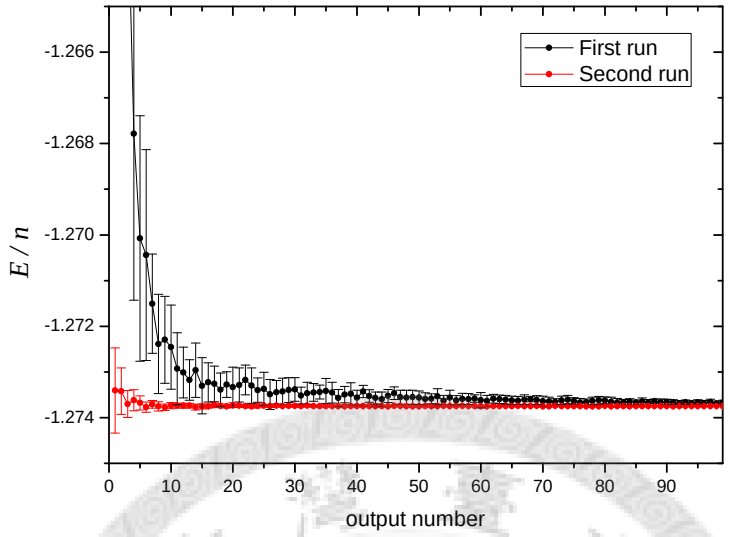


Figure 3.2: The periodic trnasverse Ising spin chain with 32 spins and $D = 6$ at $h = 1$. The first run starts with random matrices and $q_0 = 0.1$. The second run starts with the optimized matrices obtained in the first run and $q_0 = 0.01$. In the second run, the difference between the optimized energy and the exact energy is less than 10^{-4}.

3.3 Approximation forms for the open boundary condition

We present a form of MPS for open boundary condition at the end of section 2.2.3, and in this section, we suggest an approximate form to reduce the computational cost. For the open boundary condition, the vectors $V(s)$ and the matrices $A(s)$ are dependent of position;

$$|\Psi\rangle = \sum_{s_1,s_2,\ldots,s_n} V^1(s_1)A^2(s_2)A^3(s_3)\ldots A^{n-1}(s_{n-1})V^n(s_n)|s_1 s_2 \ldots s_n\rangle.$$

Update the vectors and the matrices one by one would be time consuming; also, if a long chain is simulated, it will require a lot of computer memory and become unmanagable. We approximate the above form by assuming $V(S)$ and $A(S)$ are independent of position. The wave function for the open boundary condition is now written

$$|\Psi\rangle = \sum_{s_1,s_2,\ldots,s_n} V(s_1)A(s_2)A(s_3)\ldots A(s_{n-1})V(s_n)|s_1 s_2 \ldots s_n\rangle = \sum_S F(S)|S\rangle.$$

The reason behind this approximation is that the boundary effect is expected to decrease quickly for the inner spins, so we still assume translational invariance for the matrices $A(s)$.

The optimization scheme can be adopted with some minor changes. For the open boundary condition, in order to update the elements v_i^s and a_{ij}^s, we need to compute

$$\frac{\partial E}{\partial v_i^s} = 2\langle \frac{1}{F(S)} \frac{\partial F(S)}{\partial v_i^s} E(S) \rangle - 2\langle \frac{1}{F(S)} \frac{\partial F(S)}{\partial v_i^s} \rangle \langle E(S) \rangle$$

and

$$\frac{\partial E}{\partial a_{ij}^s} = 2\langle \frac{1}{F(S)} \frac{\partial F(S)}{\partial a_{ij}^s} E(S) \rangle - 2\langle \frac{1}{F(S)} \frac{\partial F(S)}{\partial a_{ij}^s} \rangle \langle E(S) \rangle.$$

To calculate the derivatives of $F(S)$, we introduce the vectors $L_v(k) = V(s_1)A(s_2)\dots A(s_k)$ and $R_v(k) = A(s_k)\dots A(s_{n-1})V(s_n)$. Then the derivatives can be expressed as

$$\frac{\partial F(S)}{\partial v_i^s} = \delta_{s,s_n}[L_v(n-1)]_i + \delta_{s,s_1}[R_v(2)]_i$$

and

$$\frac{\partial F(S)}{\partial a_{ij}^s} = \sum_{k=2}^{n-1} \delta_{s,s_k}[L_v(k-1)]_i [R_v(k+1)]_j.$$

Now we can incorporate the above form into the optimization scheme in the previous section. One thing to notice is that we now calculate and store $R_v(2), R_v(3), \dots, R_v(n-1)$ first, and $L_v(k)$ is generated later in the optimization process.

We further explore the the approximate form of $F(S)$. The approximation is based on the assumption that we can ignore the boundary effect for the inner spins. However, we only assume different form at both ends of the spin chian. So in order to absorb the boundary effect, we insert matrices between $V(s)$ and $A(s)$ such that

$$F_B(S) = V(s_1)B(s_1)A(s_2)\dots A(s_{n-1})B(s_n)V(s_n)$$

or

$$F_C(S) = V(s_1)C(1)A(s_2)\dots A(s_{n-1})C(n)V(s_n)$$

where $B(s)$ depends on the spins at each ends; $C(1)$ and $C(n)$ are two matrices which are independent of spins direction. In Fig. 3.3, we investigate the D dependence of the the relative error in the energy for different forms. We observe that after adding the matrices $B(s)$, the optimized energy indeed converged to a lower value. It should be noticed that the number of

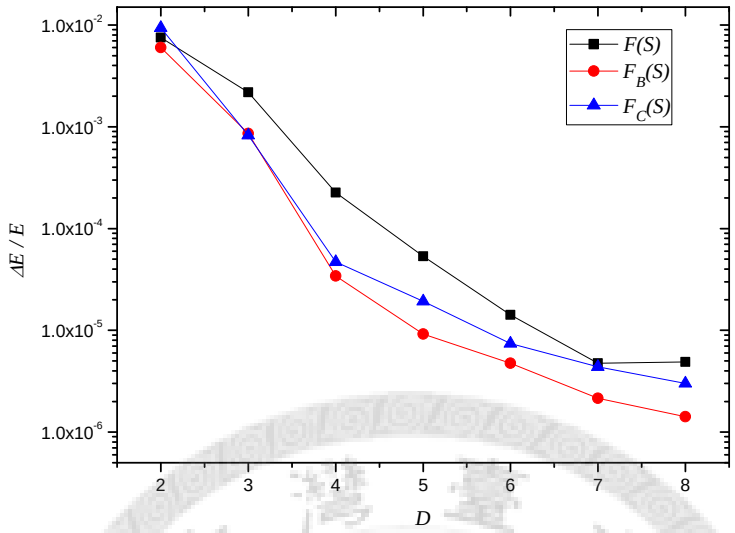


Figure 3.3: Comparison of the D dependence for the different coefficeint forms. The system is the open critical transverse Ising chain with 32 spins at $h = 1$.

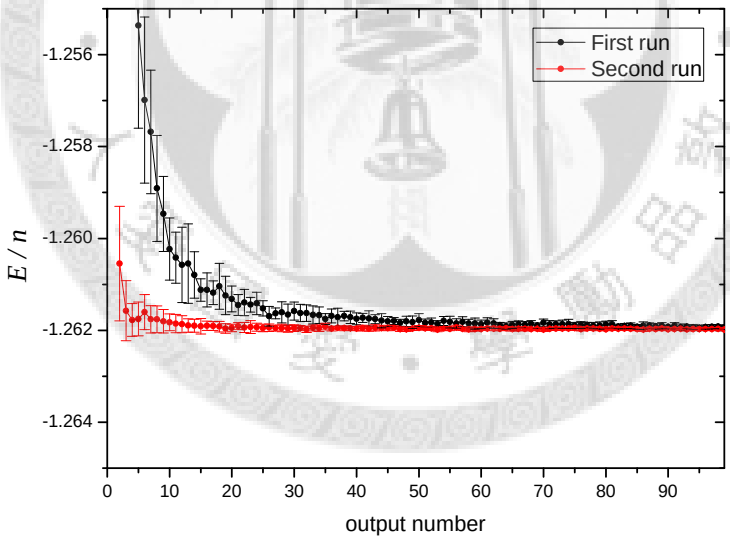


Figure 3.4: The open transverse Ising spin chain with 32 spins and $D = 6$ at $h = 1$. The first run starts with random matrices and $q_0 = 0.1$. The second run starts with the optimized matrices obtained in the first run and $q_0 = 0.01$. In the second run, the difference between the optimized energy and the exact energy is less than 10^{-4}.

the added parameters are the same for $F_B(S)$ and $F_C(S)$, but $F_B(S)$ has a better convergence. From now on, we use $F_B(S)$ as the coefficient of the trial wave function with the open boundary condition.

Fig. 3.4 shows a typical convergence process for open boundary condition. As with the periodic boundary condition, the energy also decreased exponetially, and the variance was reduced rapidly. Furthermore, when the vectors and the matrices obtained in the first run are used in the second run, the optimized enegy also reached a lower value.

3.4 Methods of measurements

In the previous sections, we measure the energy by Monte Carlo sampling because the measurement process can be incorporated into the optimization process; however, the MPS state features a powerful property which is able to sum over all states. We first discussed the measurements by Monte Carlo sampling and then the method to sum over all states.

3.4.1 Measurements by Monte Carlo sampling

Suppose we want to measure the expetation value of the operator $\hat{O}$. The expectation value can be expressed as

$$\langle \hat{O} \rangle = \frac{\sum_{S,S'} F(S)F(S')\langle S|\hat{O}|S'\rangle}{\sum_S F^2(S)} \tag{3.18}$$

$$= \frac{\sum_S F^2(S) \sum_{s'} \frac{F(S')}{F(S)} \langle S|\hat{O}|S'\rangle}{\sum_S F^2(S)}. \tag{3.19}$$

The Monte Carlo method can be applied to find the expectation value by taking $F^2(S)$ as the weight and use the Metropolis algorithm to sample the states. The advantage of the Monte Carlo method is that it avoids summing over all states which costs lots of computational resources, and sometimes it is even impossible to sum over all states; however, the price to pay is that we have to devise an efficient algorithm to sample the states. In the one dimensional case, it is suffice to use the Metropolis algorithm.

3.4.2 Measurements by summing over all states

We first demonstrate the method by calculating the normalized factor $\langle \Psi|\Psi\rangle$ for a periodic system. The generalization to the open boundary condition is

Table 3.1: Results for the critical transverse Ising model. PBC stands for the periodic boundary condition and OBC for the open boundary condition. The statitical errors in the last displayed digits are indicated; the statistical errors are less than 10^{-5}.

	n	D	E/n (MC)	E/n (ALL)	E/n (exact)
PBC	16	6	-1.27528(6)	-1.275285	-1.275287
	32	12	-1.27374(5)	-1.273744	-1.273751
	50	16	-1.27344(5)	-1.273445	-1.273449
OBC	16	6	-1.25101(9)	-1.251018	-1.251024
	32	12	-1.26200(6)	-1.262005	-1.262009
	50	16	-1.26602(0)	-1.266019	-1.266023

straight forward. Expand the MPS form directly,

$$\begin{aligned}
\langle\Psi|\Psi\rangle &= \sum_{S,S'} F(S)F(S')\langle S|S'\rangle \\
&= \sum_{\{s_i\}} tr(A(s_1)A(s_2)\dots A(s_n))tr(A(s_1)A(s_2)\dots A(s_n)) \\
&= \sum_{\{s_i\}} a^{s_1}_{\alpha_1\alpha_2} a^{s_1}_{\alpha'_1\alpha'_2} a^{s_2}_{\alpha_2\alpha_3} a^{s_2}_{\alpha'_2\alpha'_3} \cdots a^{s_n}_{\alpha_n\alpha_1} a^{s_n}_{\alpha'_n\alpha'_1} \\
&= \sum_{\{s_i\}} tr([A(s_1)\otimes A(s_1)][A(s_2)\otimes A(s_2)]\cdots[A(s_n)\otimes A(s_n)]).
\end{aligned}$$

We define the $D^2\times D^2$ tensor $T^{(k)} = \sum_{s_k} A(s_k)\otimes A(s_k)$ and the normalized factor is rewritten,

$$\langle\Psi|\Psi\rangle = tr(T^{(1)}T^{(2)}\cdots T^{(n)}). \tag{3.20}$$

Instead of conrtacting $D\times D$ matrices for all possible configurations, we contract $D^2\times D^2$ matrices for just one time.

A key notion of the MPS is that it allows *local updates*; that is, we can update some matrix $A(s_k)$ by acting the operator on it without changing other matrices. Suppose the operator $\hat{O}$ acts on the kth spin. We define $\tilde{A}(s_k) = \sum_{s'_k} \hat{O}_{s_k s'_k} A(s'_k)$ and the tensor $\tilde{T}^{(k)} = \sum_{s_k} A(s_k)\otimes\tilde{A}(s_k)$. So the expectation value of the operator $\hat{O}$ can be calculated

$$\langle\Psi|\hat{O}|\Psi\rangle = tr(T^{(1)}T^{(2)}\cdots\tilde{T}^{(k)}\cdots T^{(n)}). \tag{3.21}$$

In Table 3.1, we compare the energy obtained from the Monte Carlo method and the summing-over-all-states method. The optimized matrices

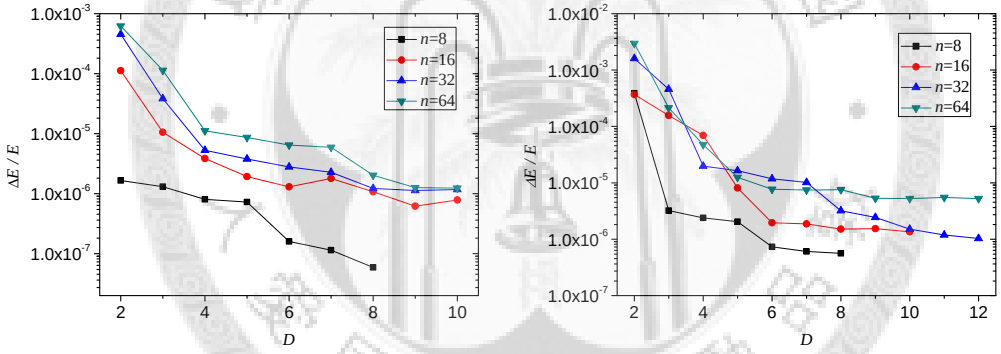
from the Monte Carlo method is used to sum over all states. The excellent agreement between these two methods indicates that the sampling of the states is faithful.

3.5 Studies of the transverse Ising model

In this section, we study the tansverse Ising model by the stochastic optimization and measure the physical quantity from the optimized MPS by summing over all states.

3.5.1 The D dependence

In Fig. 3.5, we investigated the D dependence for different sizes of the spin chain. The realtive energy error derceased exponetially with D; however, as D grows larger, the decreasing rate becomes smaller. This phenomenon results from the stochastic optimization where we update each matrix elements independently regardless of their mutual denpendence.



(a) Periodic boundary condition (b) Open boundary condition

Figure 3.5: Relative error in the energy as a function of D at $h = 1$.

3.5.2 Ground state energy, magnetization and correlation functions

The ground state structure of the transverse Ising model changes as the transvese field h is varied. Fig. 3.6 shows the energy per bond as a function of h and Fig. 3.7 the magnetization for the periodic boundary condition.

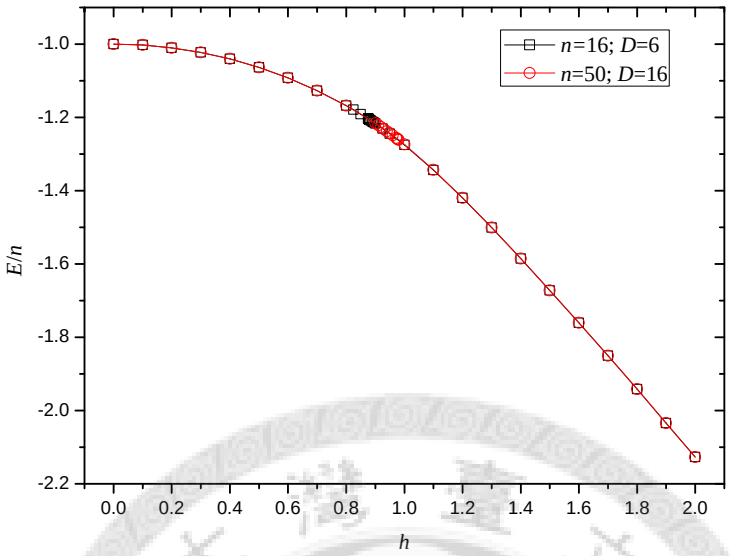


Figure 3.6: Energy per bond as a function of h for the periodic boundary condition.

These figures show that when $h \to 0$, the ground state is a product of spins pointing in the the same z direction,

$$|\Psi_g\rangle = \prod_{i=1}^{n} |\uparrow\rangle_i. \tag{3.22}$$

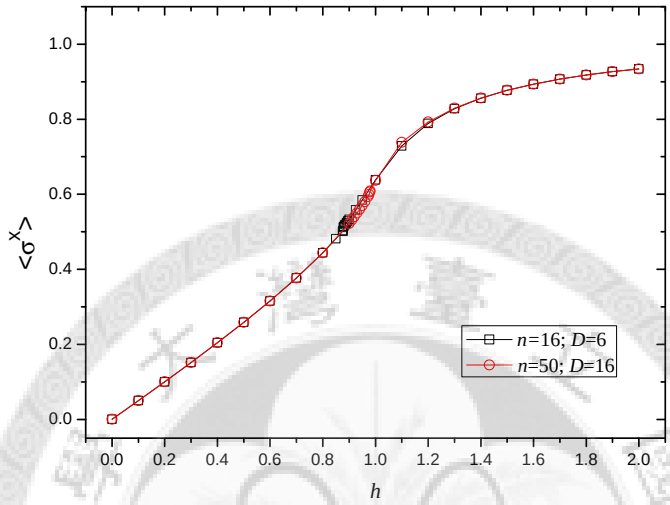
When h is much greater than one, the ground state approaches a product of spins pointing in the positive x direction,

$$|\Psi_g\rangle = \prod_{i=1}^{n} |\rightarrow\rangle_i. \tag{3.23}$$

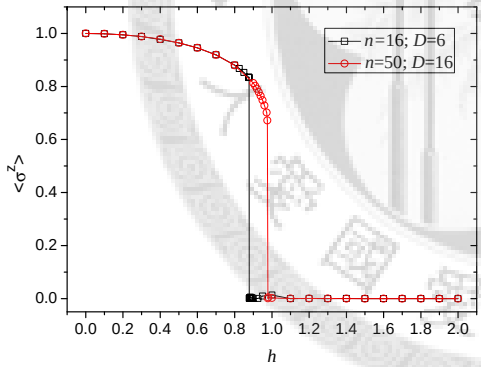
where $|\rightarrow\rangle_i = (|\uparrow\rangle_i + |\downarrow\rangle_i)/\sqrt{2}$ and $|\leftarrow\rangle_i = (|\uparrow\rangle_i - |\downarrow\rangle_i)/\sqrt{2}$ which are eigenstates of σ_x.

Using the stochastic optimization, a symmetry breaking state is found as shown in Fig. 3.7(b), but for finite size n, there is no symmetry breaking in the ground state. The ground state has spin inversion symmetry. That the magnetization dropped quickly to zero indicates that the symmetry under the global phase flip appeared. From the squared magnetization,

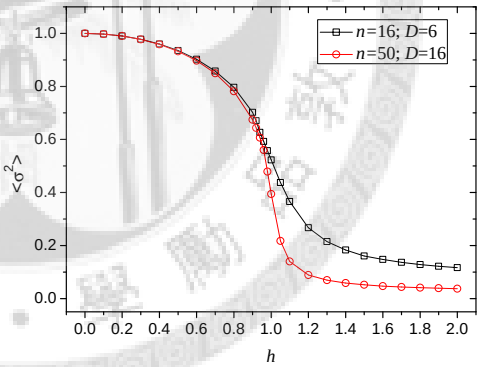
$$\langle \sigma^2 \rangle = (\frac{1}{n} \sum_{i=1}^{n} \sigma_i^z)^2, \tag{3.24}$$



(a) Magnetization $\langle \sigma^x \rangle$



(b) Magnetization $\langle \sigma^z \rangle$

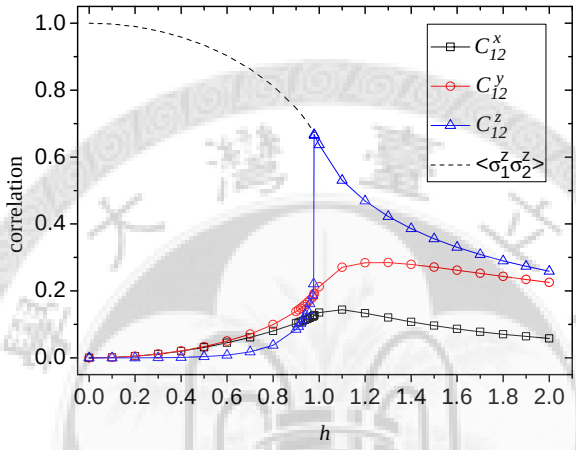


(c) Squared magnetization $\langle \sigma^2 \rangle$

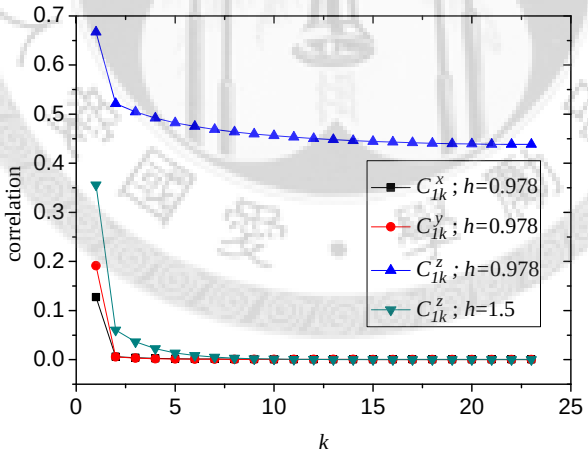
Figure 3.7: Magnetization as a function of h for the periodic boundary condition.

shown in Fig. 3.7(c), $\langle\sigma^2\rangle$ is not zero after the drop. Therefore, the results suggested that the state fluctuated between the spin inversion symmetry states.

The short range correlation function $C^i_{12} = \langle\sigma^i_1\sigma^i_2\rangle - \langle\sigma^i_1\rangle\langle\sigma^i_2\rangle$ where $i = x$, y, or z is shown in Fig. 3.8(a). The short range correlation also has a structural change near $h = 1$. The long range correlation $C^i_{1k} = \langle\sigma^i_1\sigma^i_k\rangle - \langle\sigma^i_1\rangle\langle\sigma^i_k\rangle$ near $h = 1$ is shown in Fig. 3.8(b). Near the critical point, we observe a strong long range correlation in the z-direction.



(a) Short range correlation C^i_{12}



(b) Long range correlation C^i_{1k}

Figure 3.8: Correlation for the periodic boundary condition. The system is the transverse Ising model with $n = 50$ and $D = 16$.

The open boundary system has a similar behavior with the periodic system; however, because of the dangling bonds at both ends, the spins near the edges behave differently. Fig. 3.9 shows the edge effect of the magnetization $\langle \sigma_i^x \rangle$ as a function of position.

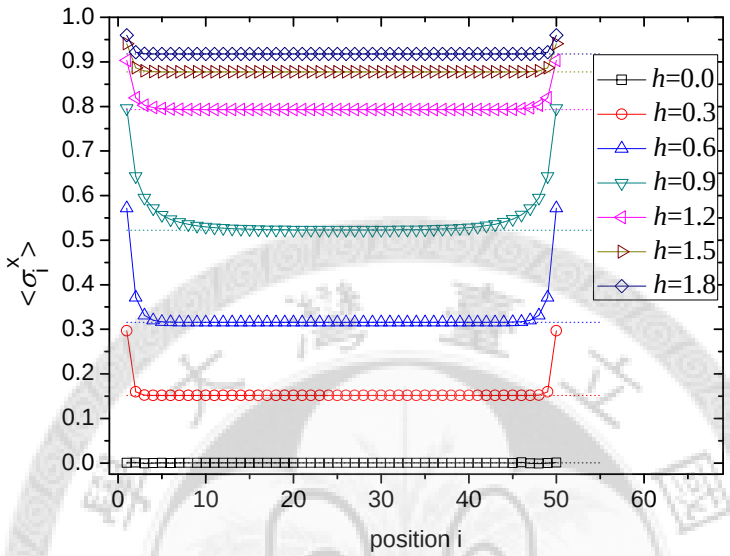


Figure 3.9: Magnetization $\langle \sigma_i^x \rangle$ as a function of position i for the open boundary condition with $n = 50$ and $D = 16$. The dotted line is the corresponding magnetization for the periodic boundary condition.

3.6 Applications to the Heisenberg model

The stochastic optimization can be applied to the $S = \frac{1}{2}$ antiferro Heisenberg chain with the next-nearest-neighbor interaction. The ground state of the antiferro Heisenberg chain can be divided into two sublattices as shown in Fig. 3.10; thus, for the periodic boundary condition, we use the trial wave function

$$|\Psi\rangle = \sum_{s_1,s_2,\ldots,s_n} tr(A(s_1)B(s_2)A(s_3)\ldots A(s_{n-1})B(s_n))|s_1 s_2 \ldots s_n\rangle,$$

where $A(s)$ and $B(s)$ are $D \times D$ matrices with $s = \pm 1$; for the open boundary condition,

$$|\Psi\rangle = \sum_{s_1,s_2,\ldots,s_n} V_A(s_1)C(s_1)B(s_2)A(s_3)\ldots A(s_{n-1})C(s_n)V_B(s_n)|s_1 s_2 \ldots s_n\rangle,$$

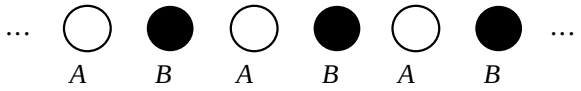


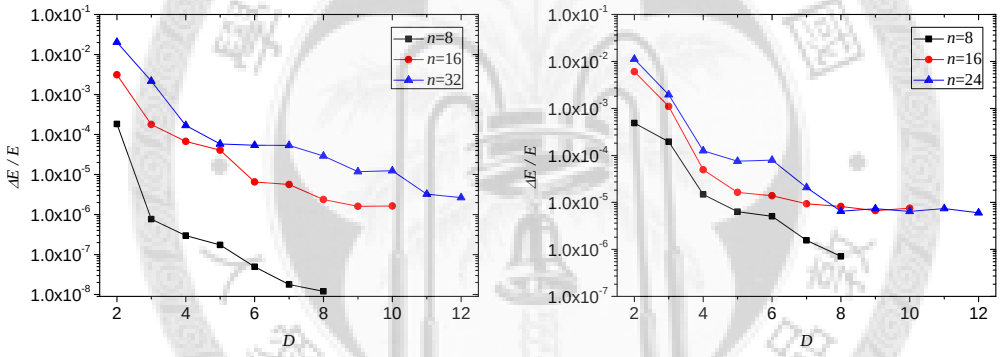
Figure 3.10: Two sublattices of the antiferro Heisenberg chain.

where $V_A(s)$, $V_B(s)$ are D dimensional vector and $C(s)$ are $D \times D$ matrices.

In this section, we add symmetry terms in the coefficients and explore the results for the $S = \frac{1}{2}$ antiferro Heisenberg chain with the next-nearest-neighbor interaction.

3.6.1 The D dependence

Fig. 3.11 shows the D dependence for different sizes of the antiferro Heisenberg chain. The relative error decreases exponentially with D as in the transverse Ising case.



(a) Periodic boundary condition (b) Open boundary condition

Figure 3.11: Relative error in the energy as a function of D.

3.6.2 Exploiting symmetry

The ground state of the antiferro Heisenberg chain has the following symmetries

$$\hat{Z}|s_1 s_2 \dots s_n\rangle = z|-s_1 - s_2 \dots - s_n\rangle \qquad \text{(spin inversion)}$$

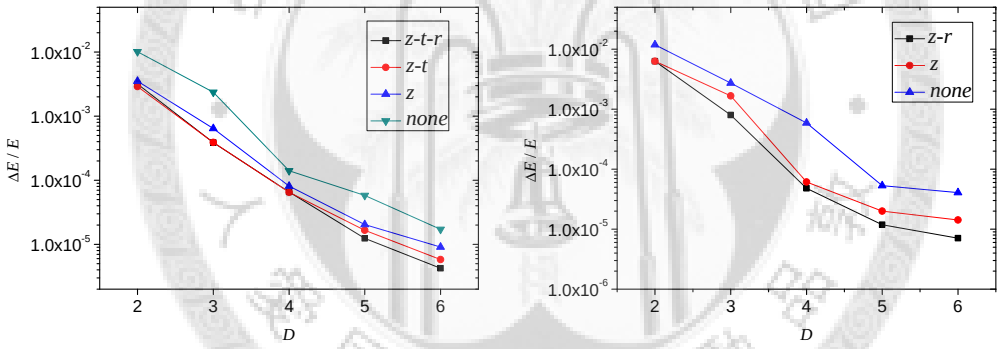
$$\hat{T}|s_1 s_2 \dots s_n\rangle = t|s_n s_1 s_2 \dots s_{n-1}\rangle \qquad \text{(translation)}$$

$$\hat{R}|s_1 s_2 \dots s_n\rangle = r|s_n s_{n-1} \dots s_1\rangle \qquad \text{(reflection)}$$

where z, t and r equals to $\pm$ 1. For the ground state, $z = t = r = 1$. These symmetries can be incorporated into the form of the trial wave function enforcing the states with these symmetries to have the same weight in the Monte Carlo simulation. It is written

$$|\Psi\rangle = \sum_S F(S) + zF(\hat{Z}S) + tF(\hat{T}S) + rF(\hat{R}S) + zrF(\hat{Z}\hat{R}S) + ztF(\hat{Z}\hat{T}S) + rtF(\hat{R}\hat{T}S) + zrtF(\hat{Z}\hat{R}\hat{T}S)|S\rangle.$$

In Fig. 3.12, we investigate the convergence behavior as the symmetry terms are added. Adding symmetry terms helps to recover the symmetries, we observe that the spin inversion term improves the energy most. Among these symmetries, updating the configuration to the spin inversion state requires a global spin flip which is hard to obtain. As mentioned in section 3.5.2, there is no symmetry breaking in the ground state with the finite size n. When the spin inversion term is added, the measured magnetization approaches zero; that is, it helps to recover the spin inversion symmetry and thus makes the optimized state more close to the truly ground state.



(a) Periodic boundary condition (b) Open boundary condition

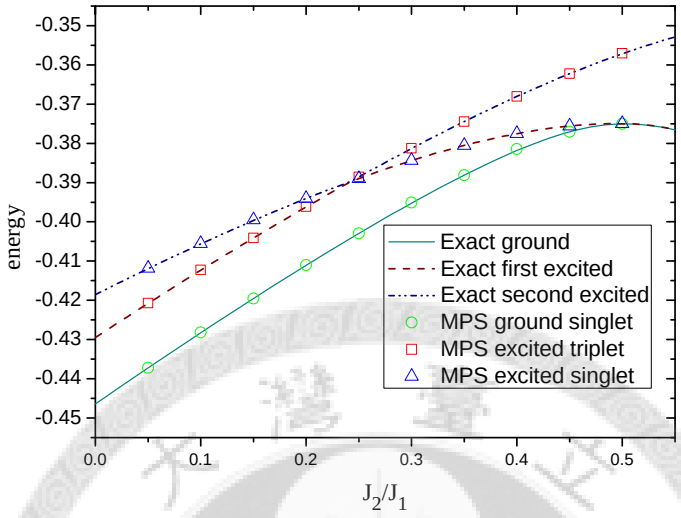
Figure 3.12: The convergence behavior as the symmetry terms are added. The system size is $n = 16$ and $D = 6$

By combining different symmetries, the method is able to find the ground state and the first few excited states of the antiferro Heisenberg model with the next-nearest-neighbor interaction which is frustrated [28]. The method is free from sign problems because the transition probability from states to states involves with $F^2(S)$ which is definitely positive. We extend the method to the open boundary condition and the results are summarized in Table 3.2 and in Fig. 3.13. As shown in Fig. 3.13, we can clearly see that the peak of the energy level is caused by the level crossing of the states with different

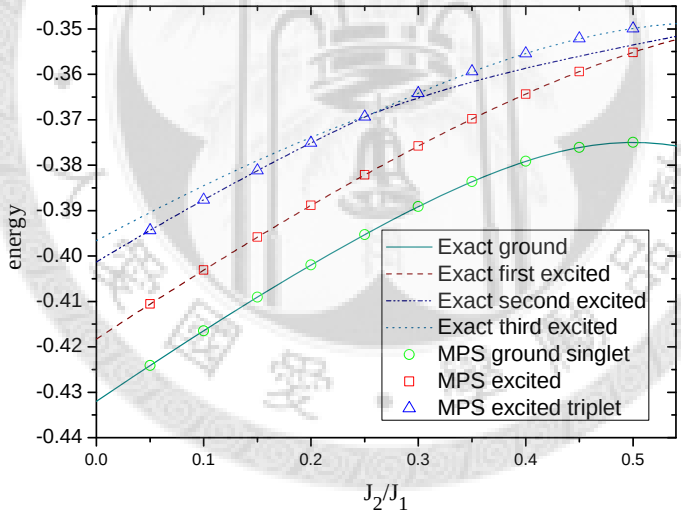
symmetries. In Fig. 3.13(a), at $J_2/J_1 = 0.5$, the energy per site of the ground state and the singlet excited state both are $-\frac{3}{8}$ regardless of the site number. It is called the Majumdar-Ghosh point [23]. The ground state and the singlet excited state are both doubly degenerate dimer products of singlet pairs on neighboring sites. On the other hand, in Fig. 3.13(b), only the ground state becomes dimer products of singlet pairs at $J_2/J_1 = 0.5$ and the energy is also $-\frac{3}{8}$ regardless of the site number.

Table 3.2: Symmetry combinations

PBC	$n = 4l$ (z, t, r)	$n = 4l + 2$ (z, t, r)
Ground singlet	$(+1, +1, +1)$	$(-1, -1, -1)$
Excited singlet	$(+1, -1, +1)$	$(-1, +1, -1)$
Excited triplet	$(-1, -1, -1)$	$(+1, +1, +1)$
OBC	$n = 4l$ (z, r)	$n = 4l + 2$ (z, r)
Ground singlet	$(+1, +1)$	$(-1, -1)$
Excited	$(-1, +1)$	$(+1, -1)$
Excited triplet	$(-1, -1)$	$(+1, +1)$



(a) Periodic boundary condition



(b) Open boundary condition

Figure 3.13: Optimized energy for the antiferro Heisenberg model with the next-nearest-neighbor interaction. The system size is $n = 16$ and $D = 6$; the relative errors in the energy are less than 10^{-4}.

Chapter 4

Imaginary time evolution with TEBD

In this chpter, the variant form of the MPS is used to simulate the imaginary time evolution. We first review the imaginary time evolution and the normalization conditions of MPS. Then the TEBD algorithm and the generalization of it, the infinte time-evolving block decimation (iTEBD), are discussed. Using these methods, we discuss the feasibility of the update scheme for the finite translationally invariant periodic systems and the ground state properties of the infinite transverse Ising model. The results show that entanglement indeed plays an important role in quantum phase transition.

4.1 Imaginary time evolution

The imaginary time evolution is a method to obtain the ground state of a system. Consider a time independent Hamiltonian H, the time dependent Schrödinger equation is

$$i\hbar\frac{\partial|\Psi(t)\rangle}{\partial t} = H|\Psi(t)\rangle.$$

It can be solved by the seperation of variables. The solution is

$$|\Psi(t)\rangle = \sum_i c_i e^{-iE_i t/\hbar}|\psi_i\rangle, \tag{4.1}$$

where c_i are complex numbers, E_i are the eigenvalues of H and $|\psi_i\rangle$ are the eigenstates of H. Set $\tau = \frac{it}{\hbar}$, Eq. (4.1) becomes

$$|\Psi(\tau)\rangle = \sum_i c_i e^{-\tau E_i}|\psi_i\rangle. \tag{4.2}$$

Because the ground state energy is the lowest, as $\tau \to \infty$, only the ground state $|\psi_0\rangle$ remains. In other words, the imaginary time evolution operator $e^{-\tau H}$ can be viewed as a projection operator which projects out the ground state. It can re-expressed as

$$|\psi_0\rangle = \frac{e^{-\tau H}|\Psi\rangle}{\|e^{-\tau H}|\Psi\rangle\|} \qquad \tau \to \infty. \tag{4.3}$$

For the transverse Ising model with n spins,

$$H = \sum_{i=1}^{n} \sigma_i^z \sigma_{i+1}^z + h\sigma_i^x = \sum_{i=1}^{n} H_i^{(z)} + H_i^{(x)}. \tag{4.4}$$

The evolution operator $e^{-\tau H}$ is a $2^n \times 2^n$ matrix which is hard to handle on a computer if n is large; however, we can appoximate the evolution operator by the second order Suzuki-Trotter decomposition [29],

$$e^{\delta(F+G)} = e^{\delta F/2} e^{\delta G} e^{\delta F/2} + O(\delta^3),$$

where F and G do not commute. Since $\sigma_i^z \sigma_{i+1}^z$ does not commute with the operator σ_i^x, we decompose the state into local operators

$$e^{-\tau H} = (e^{-\delta H})^{\frac{\tau}{\delta}} \approx ([\prod_{i=1}^{n} e^{-\delta H_i^{(x)}/2}][\prod_{i=1}^{n} e^{-\delta H_i^{(z)}}][\prod_{i=1}^{n} e^{-\delta H_i^{(x)}/2}])^{\frac{\tau}{\delta}}. \tag{4.5}$$

The operators $e^{-\delta H_i^{(z)}}$ and $e^{-\delta H_i^{(x)}/2}$ can be transformed exactly into the matrix form by the similarity transformation $M(e^{-\delta H}) = U e^{-\delta H_{diag}} U^\dagger$ where H_{diag} is the diagonalized H and U is formed by the eigenvectors of H. The way to decompose the Hamiltonian is not unique. See Ref. [3] for another way to decompose the Hamiltonain.

4.2 Normalization conditions for MPS

Recall that in section 2.2.3, we have the following equation:

$$|\Psi\rangle = \sum_{\alpha_k=1}^{\chi_k} \lambda_{\alpha_k} |\psi_{\alpha_k}^{[1\cdots k]}\rangle |\psi_{\alpha_k}^{[k+1\cdots n]}\rangle, \tag{2.6}$$

$$|\psi_{\alpha_k}^{[k+1\cdots n]}\rangle = \sum_{\alpha_{k+1}=1}^{\chi_{k+1}} \sum_{s_{k+1}} \Gamma_{\alpha_k \alpha_{k+1}}^{[k+1]s_{k+1}} \lambda_{\alpha_{k+1}} |s_{k+1}\rangle |\psi_{\alpha_{k+1}}^{[k+2\cdots n]}\rangle, \tag{2.9}$$

$$|\psi_{\alpha_k}^{[1\cdots k]}\rangle = \sum_{\alpha_{k-1}=1}^{\chi^{k-1}} \sum_{s_k} \lambda_{\alpha_{k-1}} \Gamma_{\alpha_{k-1}\alpha_k}^{[k]s_k} |\psi_{\alpha_{k-1}}^{[1\cdots k-1]}\rangle |s_k\rangle. \tag{2.10}$$

Requiring $|\psi_{\alpha_k}^{[k+1\cdots n]}\rangle$ and $|\psi_{\alpha_k}^{[1\cdots k]}\rangle$ to be orthonormal states, the vectors λ and the matrices Γ obey the following normalization conditios,

$$\langle\Psi|\Psi\rangle = \sum_{\alpha_k=1}^{\chi_k} \lambda_{\alpha_k}^2 = 1, \tag{4.6}$$

and from Eq. (2.9) and Eq. (2.10) we have

$$\langle\psi_{\alpha_k'}^{[k+1\cdots n]}|\psi_{\alpha_k}^{[k+1\cdots n]}\rangle = \sum_{\alpha_{k+1}=1}^{\chi_{k+1}} \sum_{s_{k+1}} \Gamma_{\alpha_k'\alpha_{k+1}}^{[k+1]s_{k+1}*} \lambda_{\alpha_{k+1}} \Gamma_{\alpha_k\alpha_{k+1}}^{[k+1]s_{k+1}} \lambda_{\alpha_{k+1}} = \delta_{\alpha_k'\alpha_k}, \tag{4.7}$$

and

$$\langle\psi_{\alpha_k'}^{[1\cdots k]}|\psi_{\alpha_k}^{[1\cdots k]}\rangle = \sum_{\alpha_{k-1}=1}^{\chi^{k-1}} \sum_{s_k} \lambda_{\alpha_{k-1}} \Gamma_{\alpha_{k-1}\alpha_k'}^{[k]s_k} \lambda_{\alpha_{k-1}} \Gamma_{\alpha_{k-1}\alpha_k}^{[k]s_k} = \delta_{\alpha_k'\alpha_k}. \tag{4.8}$$

The normalization conditions are important when applying the imaginary time evolution operator $e^{-\delta H}$ to the state $|\Psi\rangle$ because $e^{-\delta H}$ is not an unitary operator; that is, after $e^{-\delta H}$ acts on a normalized state $|\Psi\rangle$, we have $|\Psi'\rangle = e^{-\delta H}|\Psi\rangle$, and $\langle\Psi'|\Psi'\rangle$ no longer equals to 1. This causes the convergence process to be unstable during the numerical simulation, so it is important to normalize the state properly after applying a non-unitary operator. On the other hand, normalization by multiplying the inverse of $\langle\Psi'|\Psi'\rangle$ is not enough to keep Γ and λ in the normalization conditions. In the next section, we discuss a method for normalization which can solve this problem.

In the previous chapter, the normalzation conditions are not emphasized because we reach the ground state by varying the parameters in the function $\langle\Psi|H|\Psi\rangle/\langle\Psi|\Psi\rangle$. The numerical simulation is stable regardless of the normalization conditions. The wave function can be normalized by multiplying the inverse of $\langle\Psi|\Psi\rangle$ after the optimized matrices are obtained.

4.3 Updating the matrices

As menotioned in section 3.4.2, a useful feature of the MPS is that it allows local update. For a one-local operator $\hat{O}^{(1)}$ acting on the kth spin, the matrix $\Gamma^{[k]s_k}$ is updated to $\tilde{\Gamma}^{[k]s_k}$ according to

$$\tilde{\Gamma}_{\alpha_k\alpha_{k+1}}^{[k]s_k} = \sum_{s_k'} \hat{O}_{s_k s_k'}^{(1)} \Gamma_{\alpha_k\alpha_{k+1}}^{[k]s_k'}. \tag{4.9}$$

Updating a two-local operator $\hat{O}^{(2)}$ acting on the kth and the $k+1$th spin involves updating the matrices $\Gamma^{[k]}$, $\Gamma^{[k+1]}$ and the λ on the bond k. Recall that in section 2.2.1, the Schmidt decomposition of a state $|\Psi\rangle$ can be found by performing the singular value decomposition to the matrix that describes the relation of the two subsystems of the state $|\Psi\rangle$. From Eq. (2.11), $\sum_{\alpha_k} \lambda_{\alpha_{k-1}} \Gamma^{[k]s_k}_{\alpha_{k-1}\alpha_k} \lambda_{\alpha_k} \Gamma^{[k+1]s_{k+1}}_{\alpha_k\alpha_{k+1}} \lambda_{\alpha_{k+1}}$ describes the relation of the two subsystems $|\psi^{[1\cdots k-1]}_{\alpha_{k-1}}\rangle|s_k\rangle$ and $|s_{k+1}\rangle|\psi^{[k+2\cdots n]}_{\alpha_{k+1}}\rangle$ jointed by λ_{α_k} on the bond k. We define a tensor

$$\Theta^{s_k s_{k+1}}_{\alpha_{k-1}\alpha_{k+1}} = \sum_{\alpha_k=1}^{D} \lambda_{\alpha_{k-1}} \Gamma^{[k]s_k}_{\alpha_{k-1}\alpha_k} \lambda_{\alpha_k} \Gamma^{[k+1]s_{k+1}}_{\alpha_k\alpha_{k+1}} \lambda_{\alpha_{k+1}}, \tag{4.10}$$

and the staet $|\Psi\rangle$ is rewritten as

$$|\Psi\rangle = \sum_{\alpha_{k-1}=1}^{D} \sum_{\alpha_k=1}^{D} \sum_{\alpha_{k+1}=1}^{D} \sum_{s_k,s_{k+1}} \Theta^{s_k s_{k+1}}_{\alpha_{k-1}\alpha_{k+1}} |\psi^{[1\cdots k-1]}_{\alpha_{k-1}} s_k\rangle |s_{k+1}\psi^{[k+2\cdots n]}_{\alpha_{k+1}}\rangle. \tag{4.11}$$

When the operator $\hat{O}^{(2)}$ acts on the state $|\Psi\rangle$, the tensor Θ changes as

$$\tilde{\Theta}^{s_k s_{k+1}}_{\alpha_{k-1}\alpha_{k+1}} = \sum_{s'_k s'_{k+1}} \hat{O}^{(2)s_k s_{k+1}}_{s'_k s'_{k+1}} \Theta^{s'_k s'_{k+1}}_{\alpha_{k-1}\alpha_{k+1}}. \tag{4.12}$$

To find the updated Γ and λ, we recast the tensor $\tilde{\Theta}$ into a $2D \times 2D$ matrix

$$M_{(s_k\alpha_{k-1})(s_{k+1}\alpha_{k+1})} = \begin{pmatrix} \tilde{\Theta}^{\uparrow\uparrow}_{\alpha_{k-1}\alpha_{k+1}} & \tilde{\Theta}^{\uparrow\downarrow}_{\alpha_{k-1}\alpha_{k+1}} \\ \tilde{\Theta}^{\downarrow\uparrow}_{\alpha_{k-1}\alpha_{k+1}} & \tilde{\Theta}^{\downarrow\downarrow}_{\alpha_{k-1}\alpha_{k+1}} \end{pmatrix}, \tag{4.13}$$

Using the singular value decomposition, the matrix M is decomposed into two unitary matrices and one diagonal matrix. We truncate the the diagonal entries of the diagoal matrix by keeping only the D largest diagonal entries, and the D largest diagonal entries are identified as the updated λ; that is,

$$M_{(s_k\alpha_{k-1})(s_{k+1}\alpha_{k+1})} = \underbrace{\begin{pmatrix} U^{(\uparrow)} & \cdots \\ U^{(\downarrow)} & \cdots \end{pmatrix}}_{2D} \begin{pmatrix} \tilde{\lambda}^{[k]} & 0 \\ 0 & \ddots \end{pmatrix} \begin{pmatrix} V^{(\uparrow)} & V^{(\downarrow)} \\ \cdots & \cdots \end{pmatrix} \Big\} 2D, \tag{4.14}$$

where $\tilde{\lambda}^{[k]}$ is the updated vector of λ on the bond k and the dotted parts are all discarded. Comparing Eq. (4.14) with Eq. (4.13), we obtain the updated matrices $\tilde{\Gamma}$ by

$$\tilde{\Gamma}^{[k]s_k}_{\alpha_{k-1}\alpha_k} = (\lambda^{[k-1]}_{\alpha_{k-1}})^{-1} U^{s_k}_{\alpha_{k-1}\alpha_k}, \qquad \tilde{\Gamma}^{[k+1]s_{k+1}}_{\alpha_k\alpha_{k+1}} = V^{s_{k+1}}_{\alpha_k\alpha_{k+1}} (\lambda^{[k+1]}_{\alpha_{k+1}})^{-1}. \tag{4.15}$$

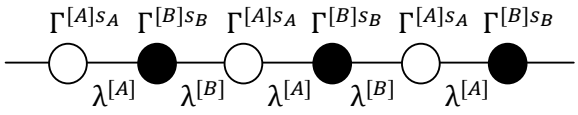


Figure 4.1: Translationally invariance translating by two.

There is another way to find the updated matrices and vectors, see Ref. [2].

For a translationally invariant system, the matrices Γ and the vectors λ should be site independent; however, the above update method breaks the translational invarinace translating by one site, but it still preserves the translational invariance translating by two sites. So we define matrices $\Gamma^{[A]s_A}$ on odd sites, $\Gamma^{[B]s_B}$ on even sites and two vectors $\lambda^{[A]}$ on odd bonds, $\lambda^{[B]}$ on even bonds. It is shown in figrue 4.1. In this circumstance, when performing the imaginary time evolution, instead of applying Eq. (4.5), we can just apply the following evolution operators

$$[e^{-\delta H_A^{(x)}/2} e^{-\delta H_B^{(x)}/2} e^{-\delta H_{BA}^{(z)}} e^{-\delta H_{AB}^{(z)}} e^{-\delta H_A^{(x)}/2} e^{-\delta H_B^{(x)}/2}]^{\frac{\tau}{\delta}}. \tag{4.16}$$

It is because the matrices are identical on even sites or on odd sites, and the terms in each square bracket are commute in Eq. (4.5). Updating a set of local matrices effectively updates other matrices in paralell. A sketch for this procedure is shown in Fig. 4.2. Although we decompose the matrices into A and B two groups, the numerical results show that we still obtain the translational invariance translating by one site after many updates.

We now introduce a method to keep Γ and λ in the normalization conditions [30]. We describe the method for the translationally invariant system.

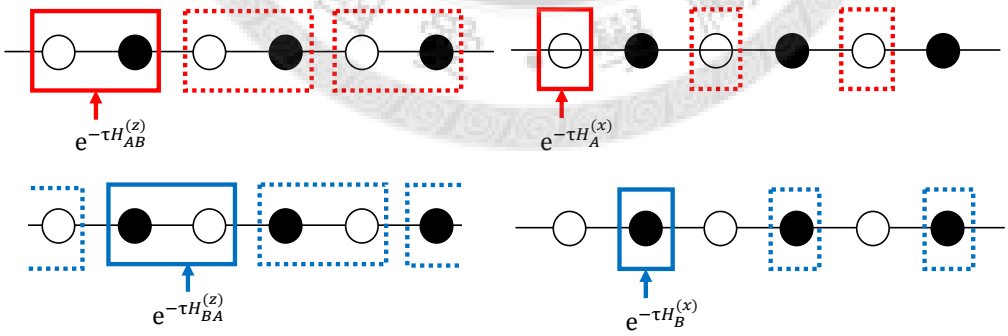


Figure 4.2: Update the matrices in parallel. The dashed square are updated simutaneously.

First, we form the matrix $M^{AB}_{(s_A\alpha)(s_B\gamma)} = \sum_\beta \lambda^{[B]}_\alpha \Gamma^{[A]s_A}_{\alpha\beta} \lambda^{[A]}_\beta \Gamma^{[B]s_B}_{\beta\gamma} \lambda^{[B]}_\gamma$. We perform the singular value decomposition to M^{AB} and obtain a new vector $\tilde{\lambda}^{[A]}$ and the new matrices $\tilde{\Gamma}^{[A]}$, $\tilde{\Gamma}^{[B]}$. The Schmidt rank does not increase in this case because it dose not involve any interaction between the two sites. We rescale $\tilde{\lambda}^{[A]}$ to satistfy $\sum_{\alpha=1}^{D} \tilde{\lambda}^{[A]2}_\alpha = 1$. Using $\tilde{\Gamma}^{[A]}$, $\tilde{\Gamma}^{[B]}$ and $\tilde{\lambda}^{[A]}$, we form the matrix M^{BA} and again obtain the new vector $\tilde{\lambda}^{[B]}$ and the new matrices. Repeating the above procedure over and over decreases the error in the normalization of the martices Γ. We perform this nomalization procedure after evolving the imaginary time δ each time.

4.4 The form of the wave function

For the transverse Ising model, the ground states of the finite periodic system and the infinite system both have the translational invairance, so the above update mehtod is suitable for studying these systems. In this section, we discuss the wave function form and the measurement method for these systems.

4.4.1 Infinite system

To construct the wave fuction form of the infinite system, we notice the form in Eq. (2.11)

$$|\Psi\rangle = \sum_{\alpha_{k-1}=1}^{\chi_{k-1}} \sum_{\alpha_k=1}^{\chi_k} \sum_{\alpha_{k+1}=1}^{\chi^{k+1}} \sum_{s_k,s_{k+1}} \lambda_{\alpha_{k-1}} \Gamma^{[k]s_k}_{\alpha_{k-1}\alpha_k} \lambda_{\alpha_k} \Gamma^{[k+1]s_{k+1}}_{\alpha_k\alpha_{k+1}} \lambda_{\alpha_{k+1}} |\psi^{[1\cdots k-1]}_{\alpha_{k-1}}\rangle |s_k\rangle |s_{k+1}\rangle |\psi^{[k+2\cdots n]}_{\alpha_{k+1}}\rangle. \quad (2.9)$$

This decomposition of the state $|\Psi\rangle$ does not contain any information about the boundary condition. For an infinite system, the decomposition can be visualized as in Fig. 4.3. $\lambda_{\alpha_{k-1}} \Gamma^{[k]s_k}_{\alpha_{k-1}\alpha_k}$ stores the information of the state $|\psi^{[-\infty\cdots k-1]}_{\alpha_{k-1}}\rangle |s_k\rangle$ and $\Gamma^{[k+1]s_{k+1}}_{\alpha_k\alpha_{k+1}} \lambda_{\alpha_{k+1}}$ stores the information of the state $|s_{k+1}\rangle |\psi^{[k+2\cdots\infty]}_{\alpha_{k+1}}\rangle$. The Schmidt coefficients λ_{α_k} describes how these two states are jointed together. Thus, the decomposition is actually an effective form for the infinite systems.

For the infinite system with translational invariance, we use the following form,

$$|\Psi\rangle = \sum_{\alpha,\beta,\gamma=1}^{D} \sum_{s_A,s_B} \lambda_\alpha \Gamma^{[A]s_A}_{\alpha\beta} \lambda_\beta \Gamma^{[B]s_B}_{\beta\gamma} \lambda_\gamma |\psi^{[-\infty\cdots B]}_\alpha\rangle |s_A\rangle |s_B\rangle |\psi^{[A\cdots\infty]}_\gamma\rangle. \quad (4.17)$$

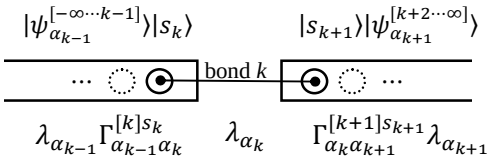


Figure 4.3: Visualization of the decomposition.

To measure the physical quantities, the site on which the operator acts is transformed to the spin basis. The expectation value of the operator $\hat{O}^{(1)}$ acting on a single site is

$$\langle\Psi|\hat{O}^{(1)}|\Psi\rangle = \sum_{\alpha,\beta=1}^{D}\sum_{s_A,s'_A}[\lambda_\alpha^{[B]}(\Gamma_{\alpha\beta}^{[A]s_A})^*\lambda_\beta^{[A]}]\hat{O}^{(1)}_{s_A,s'_A}[\lambda_\alpha^{[B]}\Gamma_{\alpha\beta}^{[A]s'_A}\lambda_\beta^{[A]}], \tag{4.18}$$

and the expectation value of a two-local operator $\hat{O}^{(2)}$ acting on the site and the nearest site is written

$$\langle\Psi|\hat{O}^{(2)}|\Psi\rangle = \sum_{\alpha,\beta,\beta',\gamma=1}^{D}\sum_{s_A,s_B,s'_A,s'_B}[\lambda_\alpha^{[B]}(\Gamma_{\alpha\beta}^{[A]s_A})^*\lambda_\beta^{[A]}(\Gamma_{\beta\gamma}^{[B]s_B})^*\lambda_\gamma^{[B]}]$$
$$\hat{O}^{(2)}_{s_A,s_B,s'_A,s'_B}[\lambda_\alpha^{[B]}\Gamma_{\alpha\beta'}^{[A]s'_A}\lambda_{\beta'}^{[A]}\Gamma_{\beta'\gamma}^{[B]s'_B}\lambda_\gamma^{[B]}]. \tag{4.19}$$

We have tried using the QMC with stochastic optimization to optimize the above wave fuction to see whether it would converge to the infinite case or not. It turns out that it does not converge to a fixed value. It can be explained by the fact that the decomposition contains no boundary information ,and thus there is too much degrees of freedom for variational methods.

4.4.2 Finite system

For the n spins periodic system with translational invariance, the following form seems to be a good candidate:

$$|\Psi\rangle = \sum_{s_1,\cdots,s_n} tr(\Gamma^{[A]s_1}\lambda^{[A]}\Gamma^{[B]s_2}\lambda^{[B]}\ldots\lambda^{[A]}\Gamma^{[B]s_n}\lambda^{[B]})|s_1s_2\ldots s_n\rangle. \tag{4.20}$$

To measure the physical quantity, we group the matrices and vecotrs into

$$A(s_k) \equiv \Gamma^{[A]s_k}\lambda^{[A]} \qquad B(s_k) \equiv \Gamma^{[B]s_k}\lambda^{[B]} \tag{4.21}$$

where $A(s_k)$ and $B(s_k)$ are $D\times D$ matrices. As mentioned in section 3.4.2, we form the $D^2\times D^2$ tensors $T^{[A](k)} = \sum_{s_k} A(s_k)\otimes A(s_k)$ and $T^{[B](k)} =$

$\sum_{s_k} B(s_k) \otimes B(s_k)$; also, when the operator $\tilde{O}$ acts on the kth site, we have $\tilde{T}^{[A](k)} = \sum_{s_k} A(s_k) \otimes \tilde{A}(s_k)$ or $\tilde{T}^{[B](k)} = \sum_{s_k} B(s_k) \otimes \tilde{B}(s_k)$. Because the matrices and vectors are all kept in the normalization condition, $\langle \Psi | \Psi \rangle = 1$, the expectation value is written

$$\langle \hat{O} \rangle = tr(T^{[A](1)} T^{[B](2)} \cdots \tilde{T}^{[m](k)} \cdots T^{[A](n-1)} T^{[B](n)}), \tag{4.22}$$

where $m = A$ or B depending on the odd site or the even site. Furthermore, the cyclic property of the trace can be used to simplify the calculation.

However, in this case, using the above wave function is only an approximation because there is no information about the periodic boundary during the process of update. Take the simplest case for example, consider a two spins system, the schmidt rank is two. On the other hand, the above update scheme gives a schmidt rank D regardless of the lattice size. To fully account for the boundary effect, one should use matrices and vectors which depend on position. However, in this scheme, we need to update the matrices and vectors site by site and a lot of computing time and memory is requried. The QMC algorithm developed in Chapter 3 avoids this porblem and there is no explicit truncation during the update. The boundary effect incorporates into the wave fuction automatically.

4.5 Imaginary time evolution algorithm

1. Decompose the Hamiltonian into two non-commute groups F and G. The operators in each group commutes.

2. Perform the Suzuki-Trotter decomposition to the imaginary time evolution operator $e^{-\tau H}$.

$$e^{-\tau H} = [e^{-\delta(F+G)}]^{\frac{\tau}{\delta}} \approx [e^{-\delta F/2} e^{-\delta G} e^{-\delta F/2}]^{\frac{\tau}{\delta}}$$

   where $\delta > 0$ and $\delta/\tau \ll 1$.

3. Apply the evolution operator evolving imaginary time δ [1] to the state.

$$|\Psi_{t_{imag}+\delta}\rangle \approx e^{-\delta F/2} e^{-\delta G} e^{-\delta F/2} |\Psi_{t_{imag}}\rangle.$$

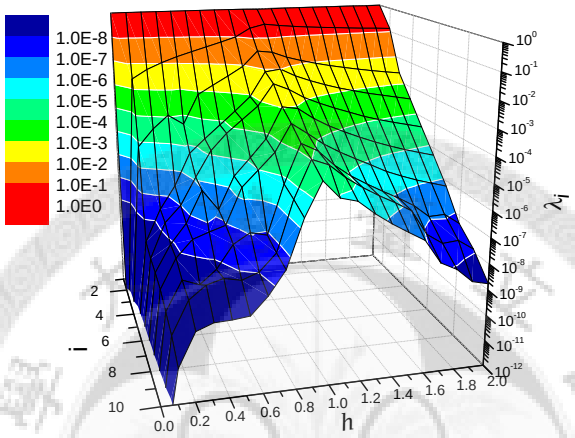
4. Perform the nomalization procedure.

5. Go back to 3 until it reaches time τ or until it reaches a optimal state.

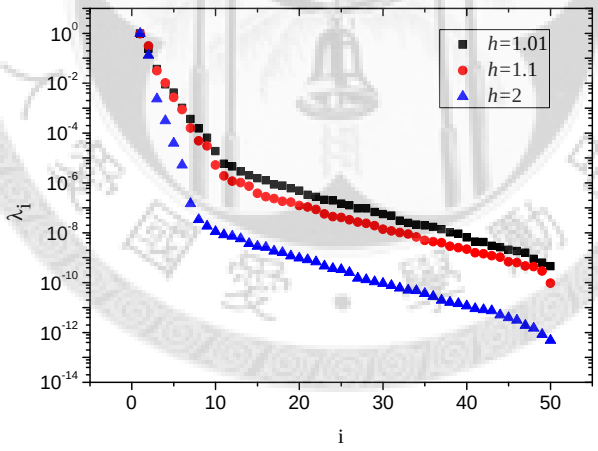
[1] We use progressively decreasing values of $\delta \in \{0.1, 0.01, 0.001, ...\}$ to reach a better convergence.

4.6 Infinite transeverse Ising model

In this section, we investigate the infinite transverse Ising model by iTEBD. The infinite transverse Ising model can be solved exactly via femionization [31]. Comparing the simulation results with the exact solutions, it shows that the imaginary time evolution with MPS provides accurate approximation for the ground state.



(a) As a fuction of h



(b) Typical ditrbutuin

Figure 4.4: The distribution of the singular values λ_i for different h.

Fig. 4.4(a) shows the distribution of the singular values λ_i as a function of h. Recall that λ_i can be viewed as a measure of entanglement. We can compare it with the single site entanglement which is obtained from the

von Neumann entropy S calculated from the single site density matrix ρ_1, $S = -tr(\rho_1 log \rho_1)$. The single site entanglement can be viewed as a measure of how entangled the lattice is [32]. In Fig. 4.5 we plot the exact single site entanglement, and we also plot the exact nearest site correlation [31], $C^z_{12} = \langle \sigma^z_1 \sigma^z_2 \rangle - \langle \sigma^z_1 \rangle \langle \sigma^z_2 \rangle$. The two show similar behavoir. The decreasing rate of the distribution of λ_i shows a qualitative agreement with the single site entanglement. For a fixed h, we observe that the singular values decrease exponentially and then after some value i, the decreasing rate becomes slower. Furthermore, as h approaches to 1, the decreasing rate of the singular values also becomes slower, indicating that the system has a maximum entanglement at the critical point [12, 32].

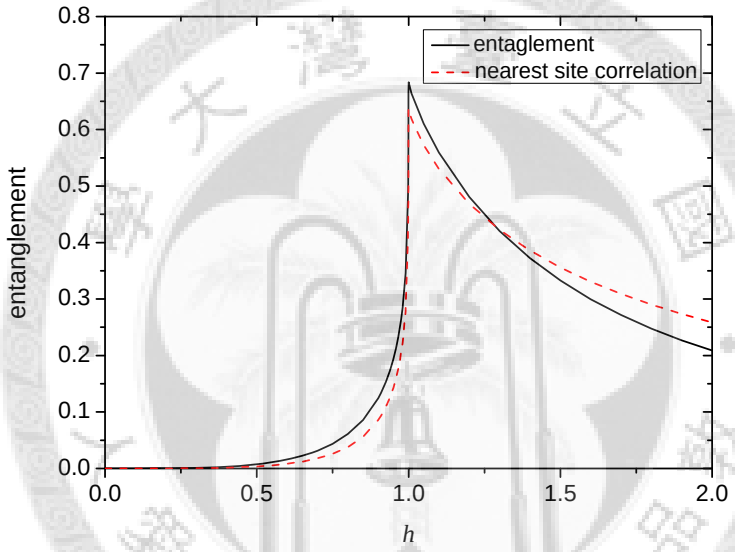


Figure 4.5: The single site entanglement and the nearest site correlation.

Fig. 4.6 shows a typical convergence process of the energy. We compare the ground state energy with the exact solutions. It is shown in Fig. 4.7. With $D = 1$, the wave function is approximated as a product state which fails to account for the entanglement. This result corresponds to the molecular field theory [33] which is a mean field theory. With $D > 1$, as D grows larger, the ground state energy obtained by the imaginary time evolution converges to the exact energy. There are two main errors. One is the error from Suzuki-Trotter decomposition, and one is the error from the truncation steps in the updating process. Comparing Fig. 4.4(a) with Fig. 4.7, the cuvrve of the relative error of energy has a similar shape with the decreasing rate of λ_i as a function of h. This indcates that the main error comes from the truncation

error.

In Fig. 4.8, we plot the magnetization and compare them with the exact solution. The $D = 1$ case corresponds to the mean field results which approximate the state as a product state, and the dotted line is obtained from the molecular field theory. In the case of $D = 10$, the magnetization $\langle \sigma^x \rangle$ shows a good agreement with the exact solution. Although the magnetization $\langle \sigma^z \rangle$ cannot be obtained via fermionization because of the spin inversion invariance of the hamiltoian, it can be obtained from the large$-k$ limit of the correlation function $\langle \sigma_1^z \sigma_k^z \rangle$ [31]. Fig. 4.9 shows that modest values of D can produce good approximations of the ground state, and as D becomes larger, the drop of the magnetization approaches the critical value $h = 1$. Comparing Fig. 4.5 with Fig. 4.8, the magnetization of the product state approximation begins to deviate as soon as the entanglement begins to grow. Also, the critical point in the product state approximation is not at $h = 1$ but $h = 2$. These results show that entanglement is essential to the quantum phase transition. The product state aprroximation fails to account for the correct characteristics of the system when it is near the critical point.

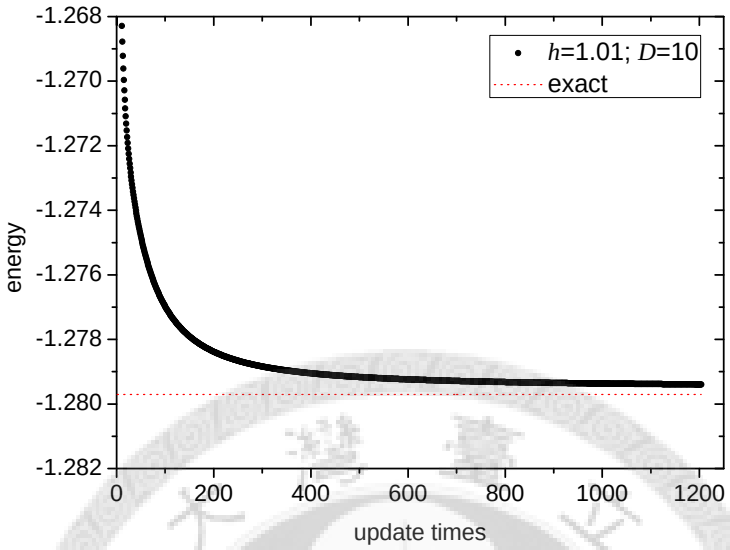


Figure 4.6: Energy convergence process.

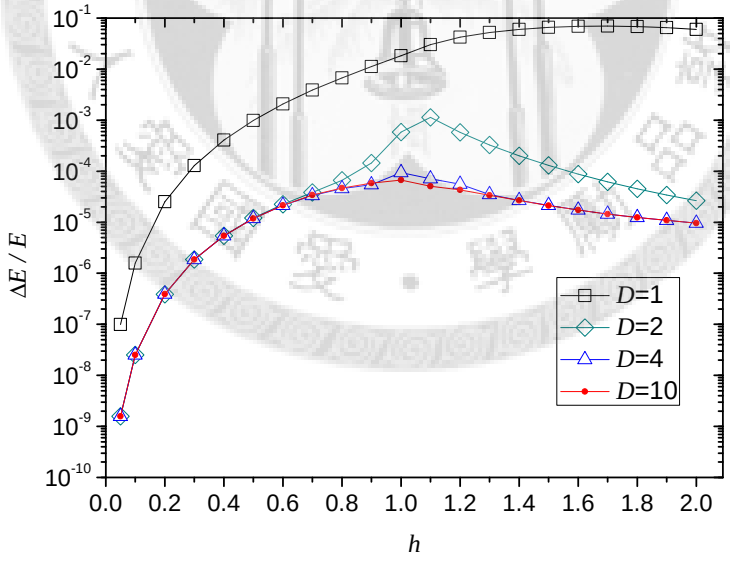


Figure 4.7: Energy convergence as a function of h for different D.

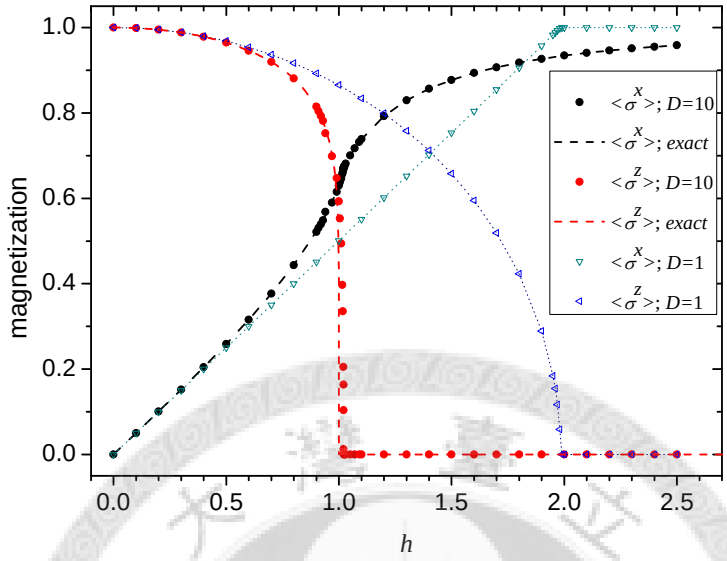


Figure 4.8: Magnetization as a function of h. The dotted lines are the molecular field results.

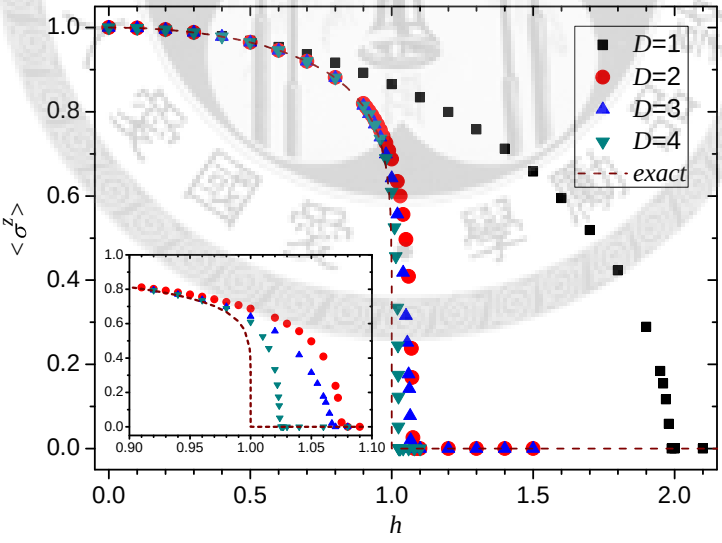


Figure 4.9: Magnetization as a function of h for different D.

Chapter 5

Conclusions

In this thesis, we have reviewed the construction of MPS from two perspectives: desity matrix renormalization group (DMRG) and quantum information theory (QIT). The connection between these two perspectives are shown leading to a clear picture of the physical meaning, which relates to entanglement, about the effectiveness of turncation steps in DMRG. Next, We develope two algorithms in detail within MPS ansatz which can deal with one dimensional quantum spin systems. First, we develope quantum Monte Carlo simulation with stochastic optimization. The notion behind this method is easy to understand. By rewriting the equation of expectation values, the sampling of states becomes classical-like. We show the convergence behavior of this method with the periodic and open boundary conditions. The results show excellent agreement with exact solutions even at critical point. We apply this method to a detailed study of the transverse Ising model. In addition, this method is free from sign-problem. By combining different symmetries, we calculate the gound state and the first few excited states of the Heisenberg model with the next nearest enighbor interaction, and the results give a clear evidence that the peak of the energy levels comes from the crossing of different symmetry states. Second, we develop the time-evolving block decimation algorithm (TEBD) and gives a detailed procedure for the imaginary tme evolution. We discuss the feasibility to apply TEBD to the periodic boundary condition. We also apply the infinite TEBD to the infinite transverse Ising model. The distribution of the Schmidt coefficients as a function of h is shown. Comparing with the convergence behavior of the energy for different D, we identify the main error source in the imaginary time evolution to be the trucation error. Finally, comparing the single site entanglement, the nearest site correlation and the phase diagram, these results imply that the entanglement can be used as an indicator to the quantum phase transition.

Bibliography

[1] A. W. Sandvik and G. Vidal, Phys. Rev. Lett. **99**, 220602 (2007).

[2] G. Vidal, Phys. Rev. Lett. **91**, 147902 (2003).

[3] G. Vidal, Phys. Rev. Lett. **93**, 040502 (2004).

[4] G. Vidal, Phys. Rev. Lett. **98**, 070201 (2007).

[5] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).

[6] S. R. White, Phys. Rev. B **48**, 10345 (1993).

[7] U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005).

[8] K. G. Wilson, Rev. Mod. Phys. **47**, 773 (1975).

[9] F. Verstraete, D. Porras, and J. I. Cirac, Phys. Rev. Lett. **93**, 227205 (2004).

[10] P. Pippan, S. R. White, and H. G. Evertz, arXiv:0801.1947v2.

[11] S. Östlund and S. Rommer, Phys. Rev. Lett. **75**, 3537 (1995).

[12] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev, Phys. Rev. Lett. **90**, 227902 (2003).

[13] F. Verstraete and J. I. Cirac, arXiv:cond-mat/0407066.

[14] A. Isacsson and Olav F. Syljuåsen Phys. Rev. E **74**, 026701 (2006).

[15] A. W. Sandvik, Phys. Rev. Lett. **101**, 140603 (2008).

[16] Z.-C. Gu, M. Levin, and X.-G. Wen, Phys. Rev. B **78**, 205116 (2008).

[17] L. Wang, Y.-J. Kao, and A. W. Sandvik, arXiv:0901.0214v1.

[18] U. Schollwöck, J. Phys. Soc. Jpn. 74 (2005) Supplement pp. 246-255.

[19] S. R. White and A. E. Feiguin, Phys. Rev. Lett. **93**, 076401 (2004).

[20] A. J. Daley *et al.*, J. Stat. Mech.: Theor. Exp. (2004) P04005.

[21] S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, Cambridge, 1999).

[22] H. A. Bethe, Z. Physik **71**, 205 (1931).

[23] C. K. Majumdar and D. K. Ghosh, J. Math. Phys. **10**, 1388 (1969); C. K. Majumdar and D. K. Ghosh, ibid. **10**, 1399 (1969).

[24] E. Y. Loh Jr. *et. al.,* Phys. Rev. B **41**, 9301 (1990).

[25] P. Henelius and A. W. Sandvik, Phys. Rev. B **62**, 1102 2000).

[26] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press,Cambridge, 2000).

[27] N. Metropolis *et. al.,* J. Chem. Phys. **21**, 1087 (1953).

[28] L. Wang, Y. J. Kao and A. W. Sandvik, unpublished.

[29] M. Suzuki, Phys. Lett. A 146, 6 (1990) 319-323; J. Math Phys. 32, 2 (1991) 400-407.

[30] D. Nagaj, E. Farhi, J. Goldstone, P. Shor, and I. Sylvester, Phys. Rev. B **77**, 214431 (2008).

[31] P. Pfeuty, Ann. Phys. **57**, 79 (1970).

[32] T. J. Osborne and M. A. Nielsen Phys. Rev. A **66**, 032110 (2002).

[33] R. Brout, K. A. Muller and H. Thomas, *Solid State Commun.* **4** (1966), 507.