NUMERICAL METHODS OF CHARACTERIZING SYMMETRY PROTECTED TOPOLOGICAL STATES IN ONE DIMENSION

BY

YE ZHUANG

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Physics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2018

Urbana, Illinois

Doctoral Committee:

Professor Bryce Gadway, Chair
Professor Taylor L. Hughes, Director of Research
Professor Nadya Mason
Professor Bryan K. Clark

Abstract

In this dissertation, we use numerical methods to study one dimensional symmetry protected topological (SPT) phases. We focus on the density matrix renormalization group (DMRG) methods and explore the machine learning methods. We investigated different SPT phases in the context of interactions and disorders. The application of machine learning methods reveals new insights into the topological phases. We begin by studying the Z_3 parafermionic chain, the simplest generalization of the Kitaev p-wave wire. The quantum entanglement diagnostics we performed allow us to determine phase boundaries, and the nature of the phase transitions. An intervening incommensurate phase is found between the topological and trivial phases. We locate and characterize a putative tricritical point in the phase diagram where the three above mentioned phases meet at a single point. The phase diagram is predicted to contain a Lifshitz type transition which we confirm using entanglement measures. As another generalization of the Kitaev p-wave wire, we study the interacting inversion symmetric superconductor. We introduce interaction and inversion symmetry and preserve its original time-reversal, particle-hole and chiral symmetry. The symmetries indicates a Z_2 classification. We study the quantum entanglement, teleportation and fractional Josephson effects of this system. The ground state of the topological phase is a condensation of four electrons instead of cooper-pairs. While there is a nonzero teleportation for cooper-pairs, the teleportation of one electron is suppressed. The inversion symmetry restricts the edge modes of the system to be cooper-pairs other than two uncorrelated electrons. It is also proved by the 2π periodicity in the fractional Josephson effects. At last we apply machine learning methods for classification of SPT phases when strong disorder is present. The entanglement spectrum is used as features to train the random forest model. We do the training using the data generated from a small fraction in the parameter space. The model can give high accuracy predictions to other regions in the phase space. It is even able to make correct predictions to system in a different symmetry class. A detailed analysis of the model indicates that it is able to capture the degeneracy in the entanglement spectrum.

To my family.

Acknowledgments

First and foremost, I would like to thank my academic advisor, Taylor L Hughes. His continuous support, insightful guidance have made my PhD study a great experience. He is always patient to answer my questions and can explain physical concepts with clear pictures. Whenever I came up with ideas, he is always generous in offering physical insight and giving me directions.

The Department of Physics is a great place for study. I am grateful to Lance Cooper for his counsel about graduate life here. I thank Taylor, Bryan Clark, Nadya Mason, and Bryce Gadway for taking the time to read my thesis and evaluate me in my final exam.

I would like to thank my collaborators and friends Norm Tubman and Luiz Santos. I appreciate their decent work and fruitful conversations, which have resulted in much of the work presented in this dissertation. I would like to thank Hitesh Changlani particularly for helping me resolve the technical problems and revise my writing. I would like to thank Victor Chua and Pouyan Ghaemi for patiently teaching me and sharing their ideas with me. I am also grateful to members in Taylor's group, especially Mayukh Khan, Ian Mondragon Shem, Wladimir Benalcazar, Matthew Lapa, Mao Lin and Tianhe Li. I appreciate the great times they have brought me.

I am thankful to my classmates Bo Han, Pak on Chan, Yizhi Fang, Xiongjie Yu, Tianci Zhou, and Wathid Assawasunthonnet for pursuing the degree together and the mutual encouragement and help throughout these years.

I would also like to express my sincere gratitude to my parents, for their constant and unconditional support. Last, but certainly not least, I am grateful to Yue. I feel very fortunate that I met him during my graduate study. I am grateful to him for his love, his encouragement, his support, for all that he has done for me. Thank you. I love you.

Finally, thanks to the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory and the Taub campus cluster at UIUC/NCSA for providing the computing resources.

Table of Contents

Chapter 1

Introduction

1.1 Topological phases

Characterizing the phases of a material is one of the most important problems in condensed matter physics. In the classical Landau theories, phases are characterized by local order parameters associated with the symmetries of the system [1]. Ordered states have non-zero local order parameters, indicating broken symmetries in those phases. However, in quantum systems there exists phases which cannot be described by Landau's symmetry breaking theory. These phases are denoted as topological phases [2, 3].

Gapped ground states are in the same phase if they can be connected adiabatically. The adiabatic evolution is realized by local unitary (LU) transformations, which acts a finite number of local operators on the ground state [4]. Two gapped ground states are considered in the same phase if they can be connected under the LU transformation without closing the gap. There exists states which can be mapped into states with no entanglement under finite number of LU transformations. These are topologically trivial states with short range entanglement, since the LU transformation only changes the entanglement locally. The remaining states must have long range entanglement, which are denoted as topologically ordered phases phases [4]. Some examples of topologically ordered states include the fractional quantum Hall states [5, 6], spin liquids [7], etc. There are usually fractional excitations in the bulk of these states with fractional charges and fractional statistics described by anyon models [8].

We require the LU transformations to preserve the symmetry, then new phases emerge: the symmetry protected topological/trivial (SPT) phases and the symmetry enriched topological (SET) phases [9]. The SPT phases are those that can only by mapped to the trivial states by symmetry-breaking LU transformations. Topological insulators and topological superconductors are SPT phases [10, 11]. They are classified by a ten-fold periodic table when interactions are not present [12, 13]. Similarly, the SET phases are emergent phases when symmetries are applied to the original topological phases instead of the trivial phases [9].

The most well known topological phase is the quantum Hall states [14, 15, 16]. It is an SPT phase protected by $U(1)$ symmetry (charge conservation). The Hall conductance is proportional to the first Chern

number, an integer valued topological invariant [15, 17]. Quantum Hall states are bulk insulators with conducting edges, which are general features of SPT phases (gapped bulk and gapless edges). Since topological phase transition is always accompanied by a bulk gap closing, the boundaries between vacuum and SPT phases must be gapless. The discovery of topological insulators opens a the door for searching new topological phases [18, 19]. Two dimensional topological insulators, also known as quantum spin Hall (QSH) states, can be understood as two copies of quantum Hall states stacked together. The magnetic fields of the two quantum Hall states are in opposite directions, making the whole system preserving time-reversal symmetry. The edge of QSH consists of two counter propagating currents with opposite spin polarizations, which are called helical edge states. Such kind of systems are realized by strong spin-orbital coupling. The Rashba effect induces effective magnetic field on the two spin bands. In real materials like HgTe, the strong spin-orbital coupling is caused by band inversion[19].

At first, only three symmetries are considered: time reversal, charge conjugation, and chiral symmetry. Recently, people realized the crystal point group symmetries can also have significant impact on the SPT phases. They are denoted as topological crystalline insulators [20, 21, 22]. Three-dimensional topological crystalline insulators have surface states protected by time reversal and discrete rotational symmetry. A new Z_2 topological invariant is defined to characterize their band structures [20].

1.2 Numerical methods for strongly correlated many-body systems

Though many interesting physics phenomena emerge from strongly correlated electron systems, they are hard to solve both theoretically and numerically. In most cases, we are only interested in the zero or low temperature properties. Therefore, finding the ground state wave function becomes the main target. Two types of numerical algorithms have been developed to tackle these questions. One start directly from the Hamiltonian and the other from trial wave function.

To solve a lattice model, the most straight forward way is to find eigen-vectors of the Hamiltonian matrix. This is what exact diagonalization does. Although this method gives us exact solutions, the systems we are able to solve are small due to the curse of dimensionality, i.e. exponential increase of the Hilbert space dimension. Many approximations and variations are developed to solve the Hamiltonian for large system size. There are mainly two approaches: approximate the original Hamiltonian with other solvable problems, or update trial wave functions.

Dynamical mean-field theory (DMFT) maps a many-body quantum problem to a local one embedded

in an effective medium [23]. The medium is solved from self-consistent conditions. The new system is an impurity model which is well studied and can be solved. The only approximation comes by setting the self-energy to be a local quantity, i.e. independent of momentum. This approximation becomes exact when the lattice has an infinite coordination. DMFT gained great success in solving Hubbard model in three dimensional spaces, but it is less accurate for low dimensional system.

Other numerical methods dependent more on wave functions: quantum Monte Carlo (QMC) and tensor network algorithms. There are many types and variations in QMC. All of them share the same key idea which is using Monte Carlo for performing high dimensional integrals. Variational Monte Carlo uses the variational method to find the trial wave function with the lowest energy [24]. Diffusion Monte Carlo (DMC) is the most common high-accuracy method for electrons. It effectively simulates the evolution of a trial wave function in imaginary time. However, DMC suffers from the fermion sign problem. One can circumvent the sign problem by fixing the node (FN) of the ground-state wavefunction to that of the trial wavefunction [25]. In practice, the accuracy of FN-DMC is limited by the quality of the node of the trial wavefunction. The path integral Monte Carlo solves systems at finite temperature [26]. Usually it behaves better for high temperatures.

The tensor network algorithms are also trial wave function methods. Among them the density matrix renormalization group has achieved great success in simulating one dimensional systems [27]. It is based on the matrix product states (MPS), which can efficiently represent states with low entanglement. The energy is lowered by iterative optimization of individual matrices. We will talk about this method in detail in the next chapter. Other variational ansatz such as the projected entangled-pair states (PEPS) [28, 29] and the mlti-scale entanglement renormalization ansatz (MERA) [30] are used in two dimensions. Recently, neural networks are used as trial wave functions [31]. they are shown to be equivalent with the tensor network states [32].

1.3 Organization of this thesis

In this thesis, we only focus on the SPT phases. The entanglement properties, such as entanglement entropy and entanglement spectrum, are widely used to detect these SPT phases [33, 34, 35, 36]. However, whether the entanglement properties are good metrics for SPT phases in general is still an open question [37]. In this paper, we will address this issue for one dimensional systems numerically.

The paper is arranged as follows. In the second chapter, we present the basic concepts of the entanglement and the numerical methods we used to calculate it. Free fermions have the special property that their

Hamiltonian only has quadratic terms. This makes it possible to exactly diagonalize the Hamiltonian and calculate desired properties. We illustrate the method with an example at the end of this chapter. The DMRG method is used for general one dimensional systems with short range entanglement. The entanglement entropy and spectrum can be extracted directly from the algorithm. This is the main technique we will use in chapter three and four. At last, the one dimensional p-wave superconductor is introduced. Models studied in the next two chapters are generalizations of it.

In the third chapter we calculate the phase diagram of the Z_3 parafermions based on the entanglement entropy diagnostic. And the ground state of the inversion symmetric topological superconductor is studied in chapter four. In chapter five, we introduce some of the most popular machine learning methods. An implementation of machine learning on the classic Ising ferromagnet is talked about. Then we apply the methods to classify the disordered SPT phase.

At last, we summarize this thesis and discuss the outlook of the work in chapter six.

Chapter 2

Background knowledge

2.1 Quantum entanglement

Entanglement is one of the most distinguishing property of a quantum system compared with a classical one. Here we give an intuitive picture of quantum entanglement. Consider a system of two $\frac{1}{2}$-spins with the single spin basis as $\{|\uparrow\rangle, |\downarrow\rangle\}$. For classical spins, this two-spin system can be in only one of the four states: $\uparrow\uparrow, \uparrow\downarrow, \downarrow\uparrow, \downarrow\downarrow$. Their corresponding states in quantum mechanics, such as $|\uparrow\downarrow\rangle$, are product states, that can be written as a product. For example, the state $|\uparrow\downarrow\uparrow\downarrow\rangle$ is a product state for a system with four spins. Moreover, the state $\frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle + |\uparrow\downarrow\rangle)$ is also a product state, because we can rewrite it as $|\uparrow\rangle \otimes \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$. Redefine the basis of the right spin as $|\pm\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle \pm |\downarrow\rangle)$. Then the state becomes $|\uparrow +\rangle$.

Product states have no entanglement. Changing or measuring one part has no effects on other parts. Product states behave the same as classical states. For example, acting an operator on the left spin of state $|\uparrow\downarrow\rangle$ will not change the state of the right spin. On the other hand, quantum states with entanglement can never have classical counterparts. Similarly, an entangled state can never be written in a product form. For example, the Bell state (or the EPR state) $\frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle)$ is the simplest entangled state [38]. If we measure one spin in this system, we get spin up or down with even probability. Now we measure the two spins together. For a classical system, we must get all four combinations with the same probability. However, this quantum state always gives both spin up or both spin down. Such behaviors make an entangled quantum state differ significantly from a classical state.

2.1.1 Entanglement entropy

How do we measure the entanglement? We continue to use the two-spin system as an example. Consider a general state $\alpha|\uparrow\uparrow\rangle + \beta|\downarrow\downarrow\rangle$ with $|\alpha|^2 + |\beta|^2 = 1$. When either α or β is zero, the state becomes a product state with zero entanglement. Intuitively, we know the state is maximally entangled when $|\alpha| = |\beta| = \frac{1}{\sqrt{2}}$. This is because both α and β are far away from zero when they are equal.

For those familiar with information theory, note that entanglement here is similar to information. Larger

entanglement indicates more connections between the two parts, therefore more information can be transferred. Consider the two spins at each end of a channel. This channel transforms $|\uparrow\rangle$ to $|\uparrow\rangle$, and $|\downarrow\rangle$ to $|\downarrow\rangle$, just like a channel transforming classical binary information. The source generates state $|\uparrow\rangle$ with probability $|\alpha|^2$ and state $|\downarrow\rangle$ with probability $|\beta|^2$. From Shannon's information theory for noiseless channels [39], we get the entropy for the channel (or entanglement for the quantum state) as $-|\alpha|^2 \log |\alpha|^2 - |\beta|^2 \log |\beta|^2$.

In general, we define the measurement of entanglement as the following. Given a system, we can always bipartite the full Hilbert space in real space into two parts $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. Here $\mathcal{H}_A$ and $\mathcal{H}_B$ should both be Hilbert spaces and be independent of each other. Then given a normalized state $|\Psi\rangle = \sum_{\alpha\beta} \Psi_{\alpha\beta} |\alpha\rangle|\beta\rangle$, a naive definition of entanglement would be $S' = -\sum_{\alpha\beta} |\Psi_{\alpha\beta}|^2$. However, this quantity is not well defined. We can always write the state in other basis by unitary transformations. Although the transformed state is essentially the same as the original one, the entanglement S' calculated this way will change. Recall that the entropy formula is true only for orthogonal states at the two ends of the channel. This corresponds to the Schmidt decomposition of the state

$$|\Psi\rangle = \sum_k \lambda_k |\alpha_k\rangle|\beta_k\rangle, \tag{2.1}$$

with $\{|\alpha_k\rangle\}$, $\{|\beta_k\rangle\}$ being the new orthogonal basis in the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively. The λ_k are non-negative real numbers and satisfy $\sum_k \lambda_k^2 = 1$. The Schmidt decomposition is the same as the singular value decomposition (SVD) of the coefficient matrix $\Psi_{\alpha\beta}$ with λ_k as singular values. Therefore, the decomposition always exists and is unique if there are no degeneracies. The *entanglement entropy* is then defined as

$$S = -\sum_k \lambda_k^2 \log \lambda_k^2, \tag{2.2}$$

This entanglement entropy is the minimum of S' from the naive definition.

From another perspective, we define entanglement from the density matrix. Define the reduced density matrix (RDM) of part A by tracing over all the degrees of freedom in region B of the density matrix:

$$\hat{\rho}_A = \mathrm{Tr}_B(\hat{\rho}) = \mathrm{Tr}_B(|\Psi\rangle\langle\Psi|). \tag{2.3}$$

Similarly, we can define the RDM of part B. It can be easily shown that $\hat{\rho}_A$ and $\hat{\rho}_B$ have the same eigen values λ_k^2. Define Renyi entropy as

$$S^{(n)} = \frac{1}{1-n} \log \mathrm{Tr}(\hat{\rho}_A^n) = \frac{\log(\sum_k \lambda_k^{2n})}{1-n}. \tag{2.4}$$

When $n \to 1$, it becomes the widely used entanglement entropy, also known as the von Neumann entropy:

$$S = -\mathrm{Tr}(\rho_A \log \rho_A), \tag{2.5}$$

which is just the entropy we defined in Eq.2.2.

Entanglement entropy has been widely used to detect topological phases in one and two dimensional systems [33, 34, 36]. In a gapped one dimensional system, the ground state entanglement entropy increases with the the block size l (the size of region A), and saturates when l reaches the correlation length [40]. Furthermore, if there is topological ground-state degeneracy we would expect an entanglement of order $\sim \log D$ where D is the degeneracy [36]. In numerical calculations, to eliminate the most harmful finite-size effects we usually take the central-cut entanglement entropy, i.e. the entanglement entropy calculated by cutting the chain in half.

In addition to entanglement entropy, the entanglement spectrum provides more information about topological phases [35]. Define the entanglement Hamiltonian $\mathcal{H}_s$ from the RDM

$$\rho_A = \frac{1}{\mathcal{Z}} e^{-\mathcal{H}_s}, \tag{2.6}$$

where $\mathcal{Z}$ is the normalization factor which ensures $\mathrm{Tr}(\rho_A) = 1$. Then the *entanglement spectrum* is the energy spectrum of the entanglement Hamiltonian $\mathcal{H}_s$. It has values $-\log \lambda_k^2$. For a general two dimensional topological quantum states, the entanglement Hamiltonian resembles the physical Hamiltonian of the one dimensional edge. The entanglement spectrum is related with the spectrum of the physical edge spectrum [35, 41].

2.1.2 Central charge

For critical one dimensional systems (gapless), it is known that the entanglement entropy grows logarithmically with l, and the scaling is characterized by the central charge. More specifically, for critical lattice systems with open boundary conditions, the entanglement entropy has the form [40]:

$$S = \frac{c}{6} \ln \left(\frac{2L}{\pi} \sin \frac{\pi l}{L} \right) + S_0 \tag{2.7}$$

where c is the *central charge*, and S_0 contains non-universal sub-leading corrections. For periodic boundary conditions,

$$S = \frac{c}{3} \ln \left(\frac{L}{\pi} \sin \frac{\pi l}{L} \right) + S_0. \tag{2.8}$$

Once we know the central charge we will have an important piece of information about the phase transition/critical phase, and can then appeal to previously known analytic results.

2.2 Methods for free fermions

To numerically solve a quantum system, we usually represent the Hamiltonian by a matrix in the predefined basis. As the system size increases, the dimension of this matrix increases exponentially, making it difficult to perform numerical calculations. However, free fermions have the advantage that there are only quadratic terms in the Hamiltonian. We do not need to deal with the full Hamiltonian, but only single-particle Hamiltonian, whose size is linear with the system size. In this section, we introduce numerical methods developed particularly for free fermions.

2.2.1 Wick's theorem

We first introduce Wick's theorem in the language of second quantization. We use the notation for free fermions with $\hat{c}$ as annihilation operator and $\hat{c}^\dagger$ as creation operator. Consider the vacuum state $|0\rangle$, then we have the following properties

$$\{\hat{c}_j, \hat{c}_k^\dagger\} = \delta_{ij}, \quad \hat{c}_j|0\rangle = 0, \quad \langle 0|\hat{c}_j^\dagger = 0. \tag{2.9}$$

For bosons, we replace anti-commutators with commutators. The fermion annihilation operator can be any operator that annihilates a given state. We only work on fermion creation and annihilation operators with the vacuum state. Other cases can be easily generalized.

Define *normal ordering* of an operator $\hat{O}$ as $:O:$ such that annihilation operators are all arranged to the right of all the creation operators. For example, set $\hat{O} = \hat{c}_1^\dagger \hat{c}_2 \hat{c}_3^\dagger \hat{c}_4$, then $:\hat{O} := -\hat{c}_1^\dagger \hat{c}_3^\dagger \hat{c}_2 \hat{c}_4$. The sign is from exchanging of fermionic operators.

Define *contraction* of two operators as $\widehat{\hat{O}_1 \hat{O}_2} = \hat{O}_1\hat{O}_2 - :\hat{O}_1\hat{O}_2:$. For $\hat{O}$ containing only one creation or annihilation operator, the contraction is either zero (when $\hat{O}_1\hat{O}_2$ is already normal ordered) or the commutator which is also a number:

$$\begin{aligned}
\overbracket{c_j \hat{c}_k} &= \hat{c}_j\hat{c}_k - :\hat{c}_j\hat{c}_k := 0 \\
\overbracket{c_j^\dagger \hat{c}_k} &= \hat{c}_j^\dagger\hat{c}_k - :\hat{c}_j^\dagger\hat{c}_k := 0 \\
\overbracket{c_j^\dagger \hat{c}_k^\dagger} &= \hat{c}_j^\dagger\hat{c}_k^\dagger - :\hat{c}_j^\dagger\hat{c}_k^\dagger := 0 \\
\overbracket{c_j \hat{c}_k^\dagger} &= \hat{c}_j\hat{c}_k^\dagger - :\hat{c}_j\hat{c}_k^\dagger := \delta_{jk}
\end{aligned}$$

Wick's theorem states that an operator $\hat{O}$ composed by $\hat{c}$ and $\hat{c}^\dagger$ can be rewritten as a sum of following terms: the normal ordered product of $\hat{O}$, the normal ordered product of $\hat{O}$ after all single contractions, all double contractions, etc., until no more contractions can be done.

$$\begin{aligned}\hat{O} = \hat{b}_1\hat{b}_2\cdots\hat{b}_n =& :\hat{b}_1\hat{b}_2\cdots\hat{b}_n: \\ &+\sum_{i<j} :\hat{b}_1\cdots\overbrace{\hat{b}_i\cdots\hat{b}_j}\cdots\hat{b}_n: \\ &+\sum_{i<j,k<l} :\hat{b}_1\cdots\overbrace{\hat{b}_i\cdots\overbrace{\hat{b}_k\cdots\hat{b}_j}\cdots\hat{b}_l}\cdots\hat{b}_n: \\ &+\ldots\end{aligned} \tag{2.10}$$

where $\hat{b}_j$ represents $\hat{c}_j$ or $\hat{c}_j^\dagger$. Now consider the expectation value of $\hat{O}$. Since the expectation value of any normal ordered operator is zero, The expectation value of operator $\hat{O}$ is non zero only when the number of creation and annihilation operators are the same. The result is just the final term containing only two-point functions.

We will not prove the theorem here, but only give an example.

$$\begin{aligned}\hat{c}_i\hat{c}_j^\dagger\hat{c}_k\hat{c}_l^\dagger &= :\hat{c}_i\hat{c}_j^\dagger\hat{c}_k\hat{c}_l^\dagger: + :\overbrace{\hat{c}_i\hat{c}_j^\dagger}\hat{c}_k\hat{c}_l^\dagger: + :\overbrace{\hat{c}_i\hat{c}_j^\dagger\hat{c}_k\hat{c}_l^\dagger}: + :\hat{c}_i\hat{c}_j^\dagger\overbrace{\hat{c}_k\hat{c}_l^\dagger}: + :\overbrace{\hat{c}_i\hat{c}_j^\dagger}\overbrace{\hat{c}_k\hat{c}_l^\dagger}: \\ &= -\hat{c}_2^\dagger\hat{c}_4^\dagger\hat{c}_1\hat{c}_3 - \delta_{ij}c_l^\dagger\hat{c}_k + \delta_{il}c_j^\dagger\hat{c}_k - \delta_{kl}c_j^\dagger\hat{c}_i + \delta_{ij}\delta_{kl}.\end{aligned}$$

The terms that are zero after contraction are not written down to save space. It can be easily checked that the last line is equal to the original operator. When we calculate the expectation value of this operator in vacuum, only the late term remains. Therefore, any n-point function can be written with two-point functions.

2.2.2 Reduced density matrices and entanglement entropy

Suppose the free fermion Hamiltonian has the form

$$\hat{H} = \sum_{j,k}\hat{c}_j^\dagger H_{jk}\hat{c}_k = \sum_j \epsilon_j \hat{f}_j^\dagger\hat{f}_j \tag{2.11}$$

with $\hat{f}_j^\dagger = \sum_k U_{kj}\hat{c}_k^\dagger$. Written in the matrix form, define column vector $\hat{c} = (\hat{c}_1, \hat{c}_2, \ldots)^T$ (similar for $\hat{f}$), hermitian matrix $\mathcal{H}$ with elements H_{jk}, and diagonal matrix ϵ with diagonal elements ϵ_j in ascending order. Then we have relations

$$U^\dagger\mathcal{H}U = \epsilon, \quad \hat{f}^\dagger = \hat{c}^\dagger U \tag{2.12}$$

The ground state with filling n is $|G\rangle = \prod_{j=1}^{n} f_j^\dagger |0\rangle$, and the corresponding energy is $E = \sum_{j=1}^{n} \epsilon_j$. The single particle correlation function $\rho_{ij} = \langle \hat{c}_i^\dagger \hat{c}_j \rangle = U_{ik}^* \langle \hat{f}_k^\dagger \hat{f}_l \rangle U_{jl}$. Since $\langle \hat{f}_k^\dagger \hat{f}_l \rangle = \delta_{k,l}$ for $k, l \leq n$, and $\langle \hat{f}_k^\dagger \hat{f}_l \rangle = 0$ for $k, l > n$, we get $\rho_{ij} = \sum_{k \leq n} U_{ik}^* U_{jk}$.

We separate the system into part A and part B. Define the $\tilde{U}$ matrix as the first n columns of the U matrix and write it into two parts as

$$\tilde{U} = \begin{pmatrix} U_A \\ U_B \end{pmatrix}$$

The correlation function ρ can be written in the block matrix form:

$$\rho = \tilde{U}\tilde{U}^\dagger = \begin{pmatrix} U_A U_A^\dagger & U_A U_B^\dagger \\ U_B U_A^\dagger & U_B U_B^\dagger \end{pmatrix}$$

Note that the right-singular vectors of matrix $U_B U_A^\dagger$ are just the eigenvectors of the matrix $M = U_A U_B^\dagger U_B U_A^\dagger$. Since $U_A^\dagger U_A + U_B^\dagger U_B = \tilde{U}^\dagger \tilde{U} = I$, we get $M = U_A U_A^\dagger - U_A U_A^\dagger U_A U_A^\dagger$, which has the same eigenvectors as the matrix $U_A U_A^\dagger$.

Construct a transformation matrix

$$T = \begin{pmatrix} T_A & 0 \\ 0 & T_B \end{pmatrix}.$$

T_A and T_B are the unitary transformation matrices that diagonalize $U_A U_A^\dagger$ and $U_B U_B^\dagger$. Then we have the following relations

$$\begin{aligned} T_A^\dagger \, U_A U_A^\dagger \, T_A &= D \\ T_B^\dagger \, U_B U_B^\dagger \, T_B &= I - D \\ T_B^\dagger \, U_B U_A^\dagger \, T_A &= S \\ S^2 &= D(I - D). \end{aligned}$$

Here D and S are diagonal matrices.

Therefore, ρ is block diagonal after the transformation:

$$T^\dagger \rho T = \begin{pmatrix} d_1 & & & \sqrt{d_1(1-d_1)} & & \\ & d_2 & & & \sqrt{d_2(1-d_2)} & \\ & & \ddots & & & \ddots \\ \sqrt{d_1(1-d_1)} & & & 1-d_1 & & \\ & \sqrt{d_2(1-d_2)} & & & 1-d_2 & \\ & & \ddots & & & \ddots \end{pmatrix}.$$

If part A is smaller than part B, the last diagonal terms should be 0 or 1. The number of the zeros and ones depends on the filling of the system. As a side note, if we diagonalize the whole density matrix, we can recover $\tilde{U}$ as eigenvectors with eigenvalues one. Other orbitals correspond to zero eigenvalues.

The ground state can be written as [42]

$$|G\rangle = \prod_{j=1}^{n} \hat{f}_j^\dagger |0\rangle = \prod_l \left(\sqrt{d_l}\hat{a}_l^\dagger + \sqrt{1-d_l}\hat{b}_l^\dagger \right) |0\rangle. \tag{2.13}$$

The $\hat{a}$ and $\hat{b}$ operators are defined within part A and B respectively.

$$\begin{aligned} \hat{a}_l^\dagger &= \sum_{j\in A} T_{jl}\hat{c}_j^\dagger \quad \text{for} \quad l \in A \\ \hat{b}_l^\dagger &= \sum_{j\in B} T_{jl}\hat{c}_j^\dagger \quad \text{for} \quad l \in B \end{aligned} \tag{2.14}$$

Operators $\hat{a}^\dagger$ and $\hat{b}^\dagger$ satisfy fermion anti-commutation relations, once we realize $T^\dagger T = I$.

The proof for Eq. 2.13 is not trivial. Here we provide a brief proof in three steps. First, we substitute $\hat{f}_j^\dagger$ with $\hat{a}_j^\dagger$ and $\hat{b}_j^\dagger$. Only $j \le n$ is considered. After simplification, we get $\hat{f}_j^\dagger = \sum_k \hat{a}_k^\dagger (T_A^\dagger U_A)_{kj} + \hat{b}_k^\dagger (T_B^\dagger U_B)_{kj}$. Second, we show that the single particle states for a Slater determinant ground state is not unique. They can be related by unitary transformations. Mathematically, we define new fermion operators $\hat{\eta}^\dagger = \hat{f}^\dagger Q$, where Q is a n-by-n unitary matrix. Then we get $\prod_{j=1}^n \hat{\eta}_j^\dagger |0\rangle = e^{i\theta} \prod_{j=1}^n \hat{f}_j^\dagger |0\rangle$. Expand the transformation explicitly,

$$\begin{aligned} & \textstyle\prod_{j=1}^n \hat{\eta}_j^\dagger |0\rangle \\ = \; & \textstyle\prod_{j=1}^n \left(\sum_{k=1}^n Q_{kj} \hat{f}_k^\dagger \right) |0\rangle \\ = \; & \textstyle\sum_{\{k_1,\ldots,k_n\}} \prod_{j=1}^n Q_{k_j j} \hat{f}_{k_j}^\dagger |0\rangle \\ = \; & \mathrm{Det}(Q) \textstyle\prod_{j=1}^n \hat{f}_j^\dagger |0\rangle. \end{aligned}$$

Here, $\sum_{\{k_1,\ldots,k_n\}}$ means sum over all possible permutations. The determinant of a unitary matrix is just a phase. Therefore, we get the freedom to write the ground state as

$$\hat{f}_j^\dagger = \sum_k \hat{a}_k^\dagger (T_A^\dagger U_A Q_A)_{kj} + \hat{b}_k^\dagger (T_B^\dagger U_B Q_B)_{kj}.$$

Third, we show that we can find unitary matrices Q_A and Q_B, so that

$$\begin{aligned} (T_A^\dagger U_A Q_A)_{kj} &= \sqrt{d_k}\delta_{kj}, \\ (T_B^\dagger U_B Q_B)_{kj} &= \sqrt{1-d_k}\delta_{kj}. \end{aligned}$$

Then we recover Eq. 2.13. We notice that the above equations are just SVD for U_A and U_B. Thus, Q_A and Q_B are unitary. From $U_A^\dagger U_A + U_B^\dagger U_B = I$, it is not too difficult to show that $Q_A = Q_B$.

A few comments on Eq. 2.13. First, this equation is exact for free fermions and can be easily generalized for free bosons. We do not use any approximations during the calculations. Second, the formula indicates that a particle in state $\hat{f}_l^\dagger|0\rangle$ has probability d_l in part A and probability $1-d_l$ in part B. The states with d_l closer to 0.5 is of more interest, because they contribute most to the entanglement of the two parts. Third, Eq. 2.13 can be expanded and written as the Schmidt decomposition of the ground state. Then we can get the singular values for the full Hamiltonian as

$$\sqrt{d_1^{n_1}(1-d_1)^{1-n_1}}\sqrt{d_2^{n_2}(1-d_2)^{1-n_2}}\ldots, \tag{2.15}$$

where $n_j = \{0,1\}$. One configuration of $\{n_1, n_2, \dots\}$ gives one value. The entanglement entropy can be calculated by

$$\begin{aligned}
S &= -\sum_{\{n_1,n_2,\dots\}} \left(\prod_l d_l^{n_l}(1-d_l)^{1-n_l}\right) \log\left(\prod_l d_l^{n_l}(1-d_l)^{1-n_l}\right) \\
&= -\left(\prod_l \sum_{n_l=0}^{1}\right)\left(\prod_l d_l^{n_l}(1-d_l)^{1-n_l}\right)\sum_l \left[n_l \log d_l + (1-n_l)\log(1-d_l)\right] \\
&= -\sum_l \sum_{n_l=0}^{1} d_l^{n_l}(1-d_l)^{1-n_l}\left[n_l \log d_l + (1-n_l)\log(1-d_l)\right] \\
&= -\sum_l \left[d_l \log d_l + (1-d_l)\log(1-d_l)\right].
\end{aligned} \tag{2.16}$$

Here we used $\sum_{n_l} d_l^{n_l}(1-d_l)^{1-n_l} = 1$. This formula is calculated from single particle Hamiltonian, different from Eq. 2.2 which is calculated from states of the full Hamiltonian.

2.3 Density matrix renormalization group

In this section, we first describe the concepts of the matrix product state and matrix product operators. They are building blocks for the density matrix renormalization group (DMRG) method [43]. DMRG is introduced in the third part.

2.3.1 Matrix product states

Consider a quantum system with N unites and each unite has d degree of freedom. This system can be electrons on a lattice, an array of qubits, etc. To describe the state of such system, we need d^N parameters.

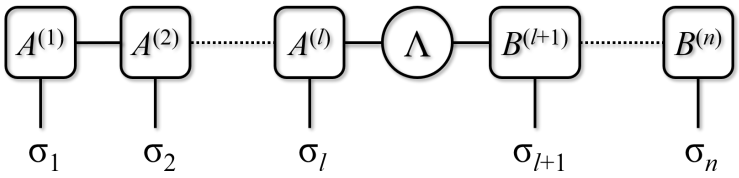


Figure 2.1: A graph representation of the MPS. Squares are tensors with their legs as indices. The circle represents a diagonal matrix. A line connecting two tensors means contracting of the corresponding indices.

The number of parameters increase exponentially with the size of the quantum system. This curse of dimensionality makes it impossible for us to represent a large quantum system exactly.

On the other hand, there is a group of special states that do not need so many parameters. Product states can be efficiently represented by Kronecker product of each unit. In principle, the number of parameters scales linearly with the system size. However, such states are too simple to exhibit interesting physics, because all unites are independent of each other.

The matrix product state (MPS) are just in between of product states and general states. They have the feature of products states that the number of parameters increases linearly, while they also have the advantage of general states of being complex enough. In general, a quantum state can be written as the matrix product form

$$|\Psi\rangle = \sum_{\{\sigma_j\}} \Psi_{\sigma_1,\sigma_2\cdots\sigma_n}|\sigma_1\sigma_2\cdots\sigma_n\rangle = \sum_{\{\sigma_j\}} A^{(1)}_{\sigma_1}A^{(2)}_{\sigma_2}\cdots A^{(n)}_{\sigma_n}|\sigma_1\sigma_2\cdots\sigma_n\rangle. \tag{2.17}$$

Here, σ_j represents local degree of freedom of unit j. $A^{(j)}_{\sigma_j}$ is a matrix for degree of freedom σ_j in unit j.

Think of $\Psi_{\sigma_1,\sigma_2\cdots\sigma_n}$ as a tensor. What we are doing is approximating this tensor by a sum of products of matrices. When no restrictions are applied on the matrices, the approximation is exact. Then we need as many parameters as the original tensor, because the dimension of matrix $A^{(j)}_{\sigma_j}$ increases dramatically as j gets further away from the edge. We restrict that the matrices should have dimensions no larger than D, the *bond dimension*. The total number of parameters is at most $D^2d\cdot N$, which is linear in N. The product states are special cases of the MPS with $A^{(j)}_{\sigma_j}$ as 1-by-1 matrices. For example, a product state of all spins up has $A^{(j)}_{\uparrow} = 1$ and $A^{(j)}_{\downarrow} = 0$ for all j.

Note that the MPS in Eq.2.17 is not unique. We can always insert the identity $I = U^\dagger U$ between $A^{(j)}_{\sigma_j}$ and $A^{(j+1)}_{\sigma_{j+1}}$ with any unitary matrix U. To get a unique representation, we apply the canonical convention

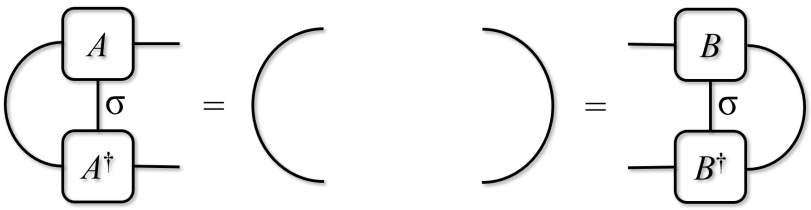


Figure 2.2: The canonical condition expressed in graph. The identity matrix is represented by a single line.

and write the MPS as

$$|\Psi\rangle = \sum_{\{\sigma_j\}} A^{(1)}_{\sigma_1} A^{(2)}_{\sigma_2} \cdots A^{(l)}_{\sigma_l} \Lambda B^{(l+1)}_{\sigma_{l+1}} \cdots B^{(n)}_{\sigma_n} |\sigma_1 \sigma_2 \cdots \sigma_n\rangle. \tag{2.18}$$

The $A^{(j)}_{\sigma_j}$ and $B^{(j)}_{\sigma_j}$ are matrices satisfying the canonical conditions

$$\sum_{\sigma} A^{\dagger}_{\sigma} A_{\sigma} = I, \qquad \sum_{\sigma} B_{\sigma} B^{\dagger}_{\sigma} = I. \tag{2.19}$$

Λ is a diagonal matrix. The canonical condition ensures that the state is normalized. It can be easily checked by applying the canonical conditions from either end of the chain.

We can get a graphical representation of the MPS as shown in Fig.2.1. We use squares to represent tensors with each leg as an index. The circle represents a diagonal matrix. The line connecting different tensors means contracting indices. A graph with dangling legs is still a tensor, while a graph with no external legs is just a number. The canonical conditions can also be represented similarly as shown in Fig.2.2. The unitary transformation mirror reflects the squares. A single line represents the identity matrix.

Any state can be written as a MPS by applying the SVD on the state recursively. For a given $\Psi_{\sigma_1,\sigma_2\cdots\sigma_n}$, we can treat it as a matrix with left index as σ_1 and right index as $(\sigma_2 \cdots \sigma_n)$. Then we decompose the matrix as

$$\Psi_{\sigma_1,(\sigma_2\cdots\sigma_n)} = A^{(1)}_{\sigma_1,k_1} \Lambda_{k_1,k_1} \tilde{\Psi}^{(1)}_{k_1,(\sigma_2\cdots\sigma_n)} = A^{(1)}_{\sigma_1,k_1} \Psi^{(1)}_{k_1,(\sigma_2\cdots\sigma_n)}. \tag{2.20}$$

Here, we write all the indices explicitly. Absorb the matrix Λ into $\tilde{\Psi}^{(1)}$ to get $\Psi^{(1)}$. We can do the SVD again on the new $\Psi^{(1)}$ with left index (k, σ_2) and right index $(\sigma_3 \cdots \sigma_n)$.

$$\Psi^{(1)}_{(k_1,\sigma_2)(\sigma_3\cdots\sigma_n)} = A^{(2)}_{(k_1,\sigma_2),k_2} \Lambda_{k_2,k_2} \tilde{\Psi}^{(2)}_{k_2,(\sigma_2\cdots\sigma_n)} = A^{(2)}_{(k_1,\sigma_2),k_2} \Psi^{(2)}_{k_2,(\sigma_3\cdots\sigma_n)}. \tag{2.21}$$

Keep doing it and rewrite $A^{(l)}_{(k_{l-1},\sigma_l),k_l}$ as matrices $(A^{(l)}_{\sigma_l})_{k_{l-1},k_l}$. Similarly, we can get matrices $B^{(l)}_{\sigma_l}$ by doing SVD from the right index.

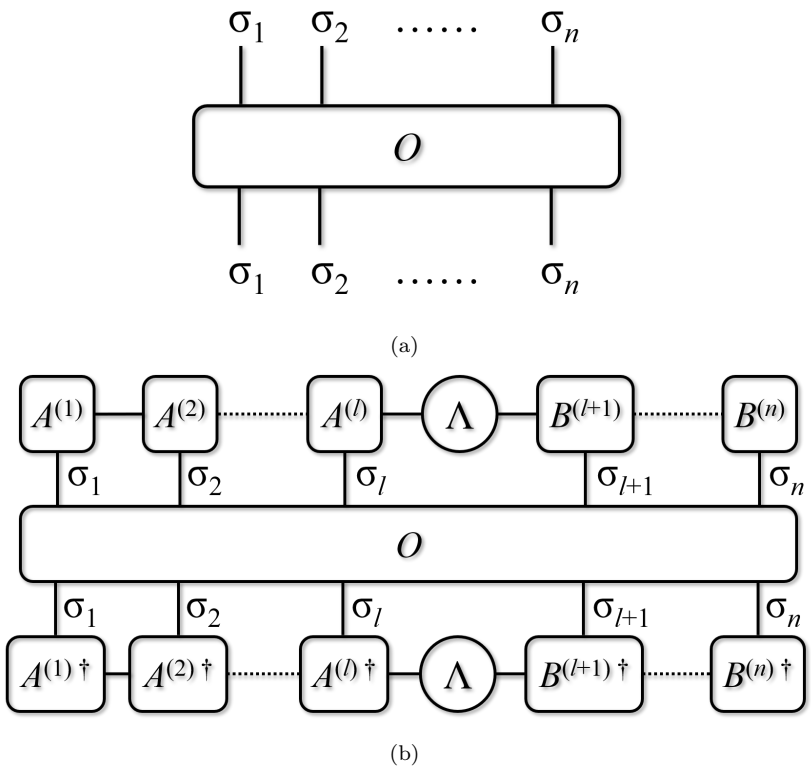


Figure 2.3: (a) A graph representation of an operator tensor. (b) The graph for the expectation value of an operator with given MPS.

We can get the MPS representation for state $|\Psi\rangle$ by applying the SVD recursively. However, it is impractical to keep all the matrices exactly. Suppose there are d local degrees of freedom on each site. Then the dimensions of matrices $A^{(1)}_{\sigma_1}$, $A^{(2)}_{\sigma_2}$, $\cdots$ are $1 \times d$, $d \times d^2$, $\cdots$, which grow exponentially. This is also true for the $B^{(j)}_{\sigma_j}$ matrices. To reduce the storage space, only the components that give the most contributions need to be kept. In practice, we take the singular vectors corresponding to the largest D singular values. Then all the matrices are at most of dimension $D \times D$. For gapped states with SRE, the error introduced by this truncation decreases exponentially as we increase the bond dimension D.

2.3.2 Matrix product operators

In physics, we are interested in physically measurable observables. From the wave function we can get expectation values of some operators. In general, the quantum operators can also be written in a form similar as the MPS. An operator in matrix form can be represented as a graph in Fig.2.3(a). The legs are indices of the tensor representing physical degree of freedom. Expectation value of this operator can be calculated by contracting all indices with the MPS (Fig. 2.3(b)).

In practice, we cannot express operator O as a matrix for the same reason as we do not explicitly write a wave function as a vector. The dimension of matrix O poses a severe restriction on the size of system we can solve. Similar to wave functions, we can also generate matrix product representations for operators. Then the expectation value of operator O is expressed by graph in Fig. 2.4(a). In this figure we omitted all the superscripts for clarity. The original big tensor O with $2n$ legs is decomposed by a production of n smaller tensors.

Unlike wavefunctions, the matrix product operators (MPO) are usually exact and straightforward to construct. Suppose we want to calculate a two-point correlation function $\langle \hat{c}_j^\dagger \hat{c}_{j+1} \rangle$. Since this operator only acts on sites j and $j+1$, we only need to apply operations on these two sites as shown in Fig. 2.4(b). Superscripts and subscripts are omitted because there is no ambiguity here. Now we explain how we get this graph in detail. The full matrix of the operator can be written as $I_1 \bigotimes \cdots \bigotimes I_{j-1} \bigotimes c_j^\dagger \bigotimes c_{j+1} \bigotimes I_{j+2} \bigotimes \cdots \bigotimes I_n$. Kronecker product of two matrix A_{ij} and B_{kl} can be written as tensor product of tensor A_{ij1} and tensor B_{1kl}. Both of them gives a new tensor with index $(ijkl)$. Here we explicitly show the index. Adding one index with dimension one does not change values of the tensor. Therefore, the Kronecker product is essentially a tensor product. We now obtain the MPO for operator $c_j^\dagger c_{j+1}$. This MPO can be further simplified. Contraction of index with dimension one can just be removed, so no horizontal legs connecting operators. Identity matrices are the same as lines. Thus, tensors in MPS are connected with their hermitian counterparts directly, except when $c^\dagger$ or c is present, which is the graph in Fig. 2.4(b).

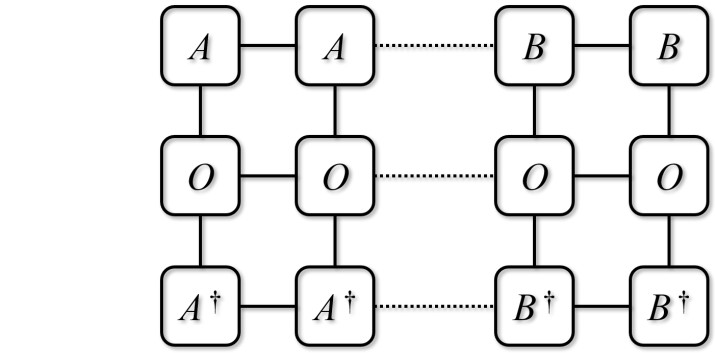
Theoretically, we can calculate any expectation values as long as we can write an operator as a product of local operators or as their sums. However, for Hamiltonian operators, it is not an efficient way to do the calculation. There are no general ways to construct MPO for Hamiltonian operators, but the idea is to create matrices for each site that can recover the original formula. We will elaborate the idea using an example.

Now we want construct the MPO for the Hamiltonian of free fermions in an open chain. The Hamiltonian of the system is

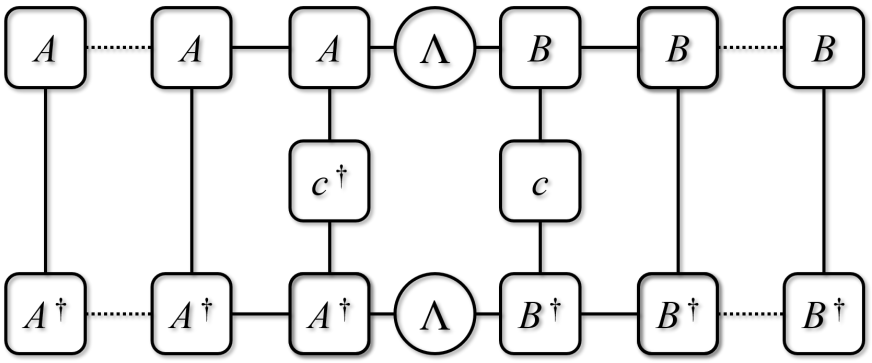
$$\hat{H} = -t \sum_{j=1}^{n-1} \left(\hat{c}_j^\dagger \hat{c}_{j+1} + h.c. \right). \tag{2.22}$$

We can rewrite the Hamiltonian as

$$\hat{H} = \prod_{j=1}^{n} O_j, \tag{2.23}$$



(a)



(b)

Figure 2.4: A graph for the expectation value of (a) a general MPO and (b) operator $c_j^\dagger c_{j+1}$. Superscripts are omitted for simplicity. Tensors can be different even though they are represented by the same symbol.

where

$$O_1 = \begin{pmatrix} 0 & \hat{c}_1 & \hat{c}_1^\dagger & 1 \end{pmatrix}, \quad O_{j\neq 1,n} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ t\hat{c}_j^\dagger & 0 & 0 & 0 \\ -t\hat{c}_j & 0 & 0 & 0 \\ 0 & \hat{c}_j & \hat{c}_j^\dagger & 1 \end{pmatrix}, \quad O_n = \begin{pmatrix} 1 \\ t\hat{c}_n^\dagger \\ -t\hat{c}_n \\ 0 \end{pmatrix} \tag{2.24}$$

It can be easily checked that we can recover the original Hamiltonian by substituting all the O_j. Note that O_j are four-tensors in the bulk (they are three-tensors at the edges). In the graph representation, the two indices of matrices O_j are represented by the horizontal legs in Fig.2.4(a). The vertical legs are physical degrees of freedom hidden in operators $\hat{c}_j$ and $\hat{c}_j$.

2.3.3 DMRG algorithm

The DMRG algorithm is ideal for solving one dimensional local Hamiltonians [27]. The key idea is updating the ground state site by site using effective local Hamiltonians. This method manifests itself under the language of MPS and MPO [43].

Starting from a random MPS, based on our previous discussion, it is not too difficult to calculate the energy of this state. Our goal is to minimize the energy by varying the parameters in the MPS, i.e. matrix elements. The general optimization methods do not work here because there are too many parameters and the derivative is not easy to calculate. We use an iterative method that updates one or two sites every time with other sites fixed. At each step, we can analytically get the minimum value and corresponding parameters. However, it is not guaranteed for this method to converge to the global minimum. In practice, we can reach reasonably low energies for systems with short range entanglement.

We talk about the algorithm in detail in the following.

Update two sites

Suppose we want to update two sites at l and $l+1$ of an MPS in the form of eq.(2.18). The system can be thought of as composed by four parts: the left environment, site l, site $l+1$ and the right environment. The structure is shown in Fig.2.5. The basis for a single site l is $|\sigma_l\rangle$. We assume the basis for the left and right environments are $|\alpha\rangle$ and $|\beta\rangle$ respectively. Then any state can be described by

$$|\Psi\rangle = \sum_{\alpha,\sigma_l,\sigma_{l+1},\beta} \Psi_{\alpha\sigma_l\sigma_{l+1}\beta}|\alpha\rangle|\sigma_l\rangle|\sigma_{l+1}\rangle|\beta\rangle \tag{2.25}$$

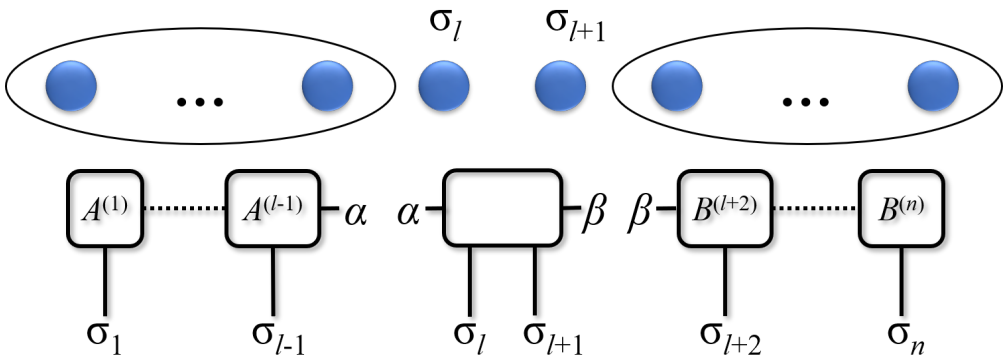


Figure 2.5: The set up for the DMRG algorithm. The system is separated as the left and right environment, and two sites in between. The corresponding MPS representations are shown in the below.

Since operators are expressed in the basis of $|\sigma_1 \dots \sigma_n\rangle$, we need to do a transform. In fact, the procedure of generating the MPS automatically creates the basis $|\alpha\rangle$ and $|\beta\rangle$ and the transformation matrix:

$$|\alpha\rangle = A^{(1)}_{\sigma_1} \cdots A^{(l-1)}_{\sigma_{l-1}} |\sigma_1 \cdots \sigma_{l-1}\rangle, \quad |\beta\rangle = B^{(l+2)}_{\sigma_{l+2}} \cdots B^{(n)}_{\sigma_n} |\sigma_{l+2} \cdots \sigma_n\rangle. \tag{2.26}$$

As represented by graphs in Fig. 2.5, the product of tensors $A^{(1)}_{\sigma_1} \cdots A^{(l-1)}_{\sigma_{l-1}}$ is a tensor with indices $(\sigma_1 \dots \sigma_{l-1}\alpha)$. It can be reformed as a transformation matrix by treating $(\sigma_1 \dots \sigma_{l-1})$ as one big index. From the canonical constraints (Eq. 2.19), the basis $|\alpha\rangle$ is orthonormal. Similar for basis $|\beta\rangle$.

Project the Hamiltonian of the system to the basis $|\alpha\rangle|\sigma_l\rangle|\sigma_{l+1}\rangle|\beta\rangle$. The projected Hamiltonian is an effective Hamiltonian for the two sites in the background environment of the other sites. This basis has dimension D^2d^2. Depending on the system and the accuracy required, the bond dimension D can range from tens to thousands. The dimension of the Hamiltonian matrix can be millions-by-millions. Such big matrix can be solved numerically using algorithms like Lanczos or Jacobi-Davidson large sparse matrix solvers.

Now we get the ground state wave function in the form of Eq. 2.25. The coefficients of the wave function $\Psi_{\alpha\sigma_l\sigma_{l+1}\beta}$ can be understood as a tensor, and represented by a graph (shown in Fig. 2.5). From the graph representation, it is clear that to get a complete ground state wave function for the full Hamiltonian, the only piece missing is $A^{(l)}_{\sigma_l} \Gamma B^{(l+1)}_{\sigma_{l+1}}$. The ground state solved from the effective Hamiltonian fits the place, but is not in the correct form. Treat $\Psi_{\alpha\sigma_l\sigma_{l+1}\beta}$ as a matrix with indices $(\alpha\sigma_l)$ and $(\sigma_{l+1}\beta)$. From the SVD:

$$\Psi_{\alpha\sigma_l\sigma_{l+1}\beta} = \sum_k U_{\alpha\sigma_l k} \Lambda_{kk} V_{k\sigma_{l+1}\beta}, \tag{2.27}$$

We get $\left(A^{(l)}_{\sigma_l}\right)_{\alpha k} = U_{\alpha\sigma_l k}$ and $\left(B^{(l+1)}_{\sigma_{l+1}}\right)_{k\beta} = V_{k\sigma_{l+1}\beta}$.

In practice, we can update any number of sites in one step. The generation to multiple sites is straightforward.

Iteration process

There are basically two types of iteration processes: sweeping in finite systems and growing in infinite systems. In finite systems, after the update of sites l and $l+1$, we move right and update the sites $l+1$ and $l+2$. We keep moving until we hit the edge and then go left. Usually, the ground state energy converges to constant after a few sweeps. In infinite systems, we insert two sites in the middle of the chain in every step. Unlike sweeping, growing changes the size of the system. The ground state energy per site converges to a constant. The finite process can start from any random MPS. Many people also like to generate the initial MPS from the infinite process, and then apply the sweeping.

During the iteration, the transformation matrix should be calculated. This includes contracting a chain of tensors. Notice that most contractions repeat themselves in the iterations. We can store the results from previous steps to speed up the program.

Although the DMRG is suitable for both fermionic and bosonic systems, attention should be paid to the fermionic ones. Suppose we want to calculate value $\langle\beta'\alpha'|\hat{O}_A\hat{O}_B|\alpha\beta\rangle$. Operators $\hat{O}_A$ and $\hat{O}_B$ only act on states $|\alpha\rangle$s and $|\beta\rangle$s respectively. This means for fermionic $\hat{O}_A$ and $|\beta\rangle$, we have $\hat{O}_A|\beta\rangle = -|\beta\rangle\hat{O}_A$ (This can be understood by thinking of $|\beta\rangle$ as $\hat{b}^\dagger|0\rangle$). If either of the two is bosonic, the minus sign is replace by a plus sign. Therefore, we can write

$$\langle\beta'\alpha'|\hat{O}_A\hat{O}_B|\alpha\beta\rangle = c\langle\alpha|\hat{O}_A|\alpha\rangle\langle\beta|\hat{O}_B|\beta\rangle, \tag{2.28}$$

where $c = \pm$ is a sign to be determined. If both $\hat{O}_A$ and $\hat{O}_B$ are bosonic, $c = 1$. Otherwise, we assume there are n_α fermions in state $|\alpha\rangle$. Then odd/even $n_\alpha - n_{\alpha'}$ indicates fermionic/bosonic $\hat{O}_A$. Since $\langle\beta'|$ need to switch with $\hat{O}_A$, we get the parity as $n_\alpha(n_{\beta'} - n_\beta)$. Therefore, $c = -1$ only when $n_\alpha(n_{\beta'} - n_\beta)$ is odd.

The main error of the DMRG comes from the truncation process that cuts matrices in the MPS to desired bond dimension D. Recall that the SVD for a general matrix is $M = U\Lambda V^\dagger$ with diagonals of Λ sorted in descending order $\lambda_1 \geq \lambda_2 \geq \ldots$. We do the cut by keeping only the D largest singular values and the corresponding left and right eigen-vectors to form a new matrix M'. The error measured by Frobenius norm between the two matrices is $\epsilon = ||M - M'||_F = \sqrt{\sum_{j=D+1}\lambda_j^2}$. In most cases the wave functions are short range entangled. The singular values decay exponentially. Therefore the error decrease exponentially with

D, making a small D (∼hundreds) sufficient for general models.

Details of the DMRG algorithm can be found in Ref. [44] and more details and discussions in Ref. [43]. We provide a pseudo code in the following. We use L as the length of the chain and H_j as the MPO of the Hamiltonian on site j. For now we assume L is even.

Algorithm 1 DMRG algorithm

procedure SETUP
  $l \leftarrow 1$.
  $r \leftarrow L$.
  Construct the Hamiltonian matrix H_{two} of a two site system by contracting H_l and H_r.
  Solve H_{two} and find the ground state wave function $\Psi_{\sigma_l\sigma_r}$.
  Write $\Psi_{\sigma_l\sigma_r}$ in form of the MPS using SVD: $\Psi_{\sigma_l\sigma_r}=A_{\sigma_l\alpha}\Lambda_{\alpha\beta}B_{\beta\sigma_r}$.
  Define the left and right environments: $E_l \leftarrow AH_lA^\dagger, \quad E_r \leftarrow BH_rB^\dagger$.
procedure INFINITE DMRG(growing)
  repeat
    $l \leftarrow l+1$.
    $r \leftarrow r-1$.
    $E_{\text{left}} \leftarrow E_{l-1}$
    $E_{\text{right}} \leftarrow E_{r+1}$
    Construct the Hamiltonian matrix H by contracting E_{left}, H_l, H_r, and E_{right}.
    Solve H and find the ground state wave function $\Psi_{\alpha\sigma_l\sigma_r\beta}$.
    Write $\Psi_{\alpha\sigma_l\sigma_r\beta}$ in form of the MPS using SVD: $\Psi_{\alpha\sigma_l\sigma_r\beta}=A_{\alpha\sigma_l\alpha'}\Lambda_{\alpha'\beta'}B_{\beta'\sigma_r\beta}$.
    Define the left and right environments: $E_l \leftarrow AE_{\text{left}}A^\dagger, \quad E_r \leftarrow BE_{\text{right}}B^\dagger$.
  until $l+1 == r$
procedure FINITE DMRG(sweeping)
  toRight $\leftarrow$ *true*
  repeat
    if *toRight* **then**
      $l \leftarrow l+1$.
      $r \leftarrow r+1$.
    else
      $l \leftarrow l-1$.
      $r \leftarrow r-1$.
    $E_{\text{left}} \leftarrow E_{l-1}$
    $E_{\text{right}} \leftarrow E_{r+1}$
    Construct the Hamiltonian matrix H by contracting E_{left}, H_l, H_r, and E_{right}.
    Solve H and find the ground state wave function $\Psi_{\alpha\sigma_l\sigma_r\beta}$.
    Write $\Psi_{\alpha\sigma_l\sigma_r\beta}$ in form of the MPS using SVD: $\Psi_{\alpha\sigma_l\sigma_r\beta}=A_{\alpha\sigma_l\alpha'}\Lambda_{\alpha'\beta'}B_{\beta'\sigma_r\beta}$.
    Update the left and right environments: $E_l \leftarrow AE_{\text{left}}A^\dagger, \quad E_r \leftarrow BE_{\text{right}}B^\dagger$.
    if $r+1 == L$ **then**
      toRight $\leftarrow$ *false*
    else if $l-1 == 1$ **then**
      toRight $\leftarrow$ *true*
  until the ground state energy converges
Calculate any desired property using the ground state wave function.

A few comments on the algorithm. In general, A, B, E, and H are tensors. The equation $\Psi = A\Lambda B$ is actually contraction over non-physical indices. The matrix productions $AHA^\dagger$ and $BHB^\dagger$ are tensor contractions over the physical indices. Since contraction of tensors is much clearer when using graphs, we

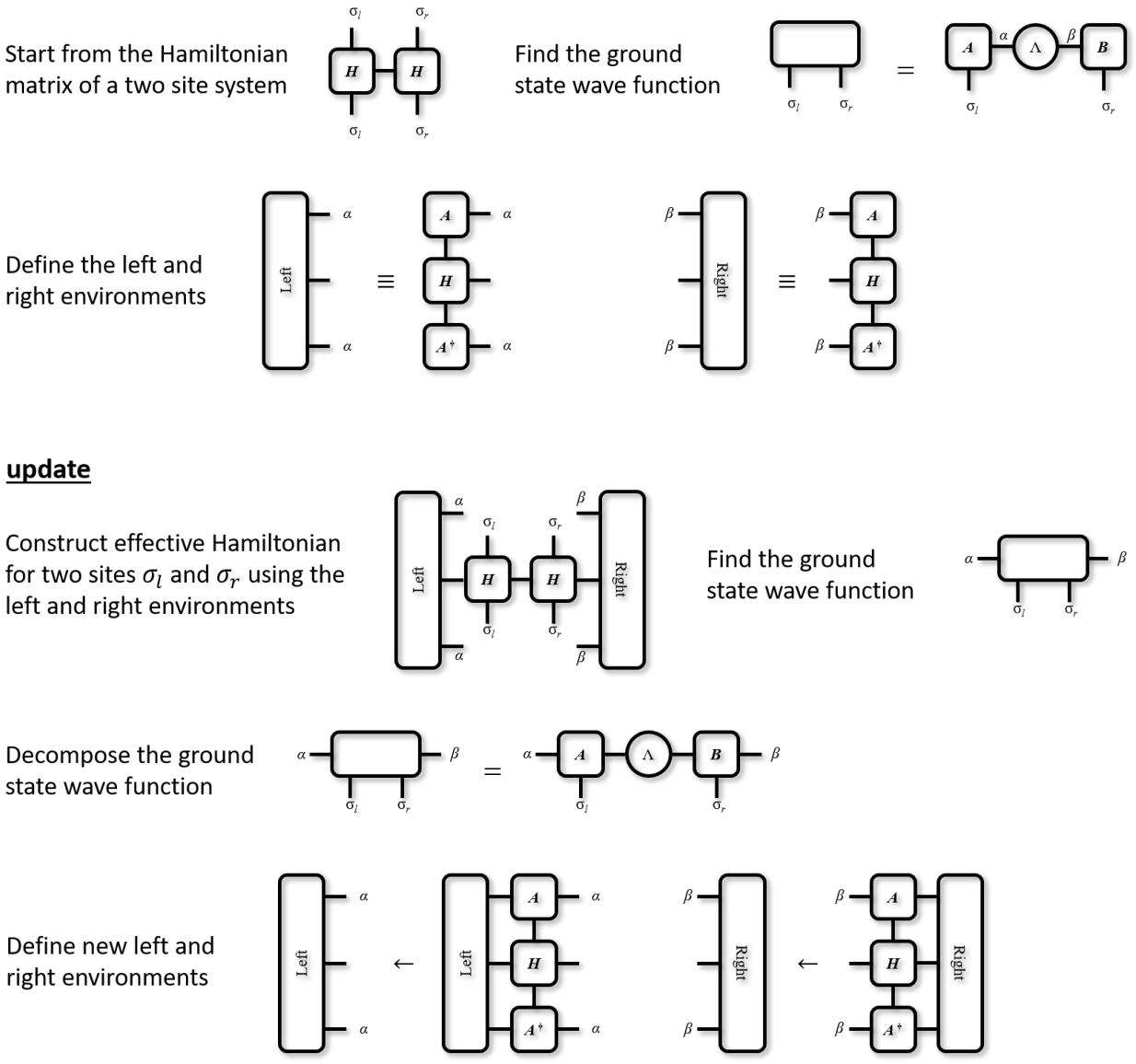


Figure 2.6: A graphical representation of the DMRG algorithm. The setup corresponds to setps from 4 to 7. The update corresponds to steps from 14 to 17 and steps from 30 to 33.

show the graph representations of the key steps in Fig. 2.6. The setup part corresponds to setps from 4 to 7. The update part appears in both infinite (steps 14 to 17) and finite (steps 30 to 33) DMRG.

Although we indicated to generate the matrix H, it is usually not explicitly done in practice. One reason for not doing so is the high dimension of this matrix. With the bond dimension set as one thousand, the matrix will have 10^{12} elements. It is not efficient to store such a big matrix; and it takes too much resources to diagonalize it exactly. Algorithms designed to solve large matrices do not require a dense matrix as input. Only a function that takes a vector and returns a vector is needed. This function calculates the action of the matrix on the input vector. With proper contraction order, such function can be efficiently implemented.

2.4 An example: 1D p-wave superconductor

In this section, we solve a well-known lattice model: the spin-polarized (or spinless) p-wave superconductor (pSC) [45]. It is the simplest model for a topological superconductor. We will first solve the spectrum of the system in the bulk analytically, and calculate the winding number as an indicator for topological phase. This model can also be solved numerically using the technique for free fermions. Phase diagrams and some entanglement properties will be discussed. We will show the deep connection between the p-wave superconductor, the Kitaev chain and the transverse-field Ising model. Some physical quantities manifest themselves in one model other than in others.

The Hamiltonian of the p-wave superconductor is

$$\hat{H}_{\mathrm{p-SC}} = \sum_j \left(-t\hat{c}_j^\dagger \hat{c}_{j+1} + \Delta \hat{c}_j \hat{c}_{j+1} + h.c. \right) + 2\mu \sum_j \left(\hat{c}_j^\dagger \hat{c}_j - \frac{1}{2} \right). \tag{2.29}$$

t is electron hopping parameter and μ is chemical potential. Δ is a p-wave pairing coefficient. This is the simplest pairing we can have because the s-wave paring potential is not possible for spinless fermions.

2.4.1 Bogoliubov-de-Gennes Hamiltonian

For chains with periodic boundary conditions or infinitely long, we can apply Fourier transformation and get the single-particle Hamiltonian in the momentum space.

$$\hat{H}_{\mathrm{p-SC}} = \sum_k \left(-2t\cos(k)\hat{c}_k^\dagger \hat{c}_k - i\Delta \sin(k)\hat{c}_k \hat{c}_{-k} - i\Delta^* \sin(k)\hat{c}_k^\dagger \hat{c}_{-k}^\dagger \right) + 2\mu \sum_k \left(\hat{c}_k^\dagger \hat{c}_k - \frac{1}{2} \right). \tag{2.30}$$

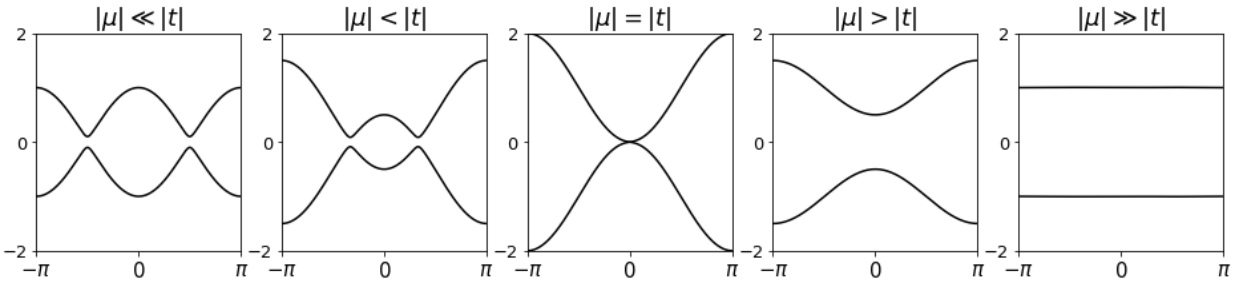


Figure 2.7: The band structures for the p-wave superconductor at different parameters. From left to right, μ increases from zero and t decreases to zero. The phase transition happens when $\mu = t$. A small Δ is kept unchanged. It opens a gap in the topological phases on the left.

k takes all values in the 1st Brillouin zone. For small p, keep only quadratic terms, we get

$$\hat{H}_{\mathrm{p-SC}} = \sum_k \epsilon_k \hat{c}_k^\dagger \hat{c}_k - \sum_k \left(i\Delta k \hat{c}_k \hat{c}_{-k} + i\Delta^* k \hat{c}_k^\dagger \hat{c}_{-k}^\dagger \right) \tag{2.31}$$

where $\epsilon_k = -tk^2 - 2t + 2\mu$. We ignored the constant term. The first term is just the kinetic energy for free fermions. The second term gives the momentum dependent pairings.

Unlike free fermions, here the Hamiltonian is not diagonalized in the momentum space. We use the Nambu representation. Define spinor as $\Psi_k \equiv (c_k \; c_{-k}^\dagger)^T$. Omit constant terms, we can write the Hamiltonian in Eq. 2.30 as $H = \sum_k \Psi_p^\dagger \mathcal{H}_{\mathrm{BdG}}(k) \Psi_k$. The Bogoliubov-de-Gennes (BdG) Hamiltonian is then

$$\mathcal{H}_{\mathrm{BdG}}(k) = \begin{pmatrix} \mu - t\cos(k) & i\Delta^* \sin(k) \\ -i\Delta \sin(k) & -\mu + t\cos(k) \end{pmatrix} = -\Delta^* \sin(k)\sigma_y + (\mu - t\cos(k))\sigma_z.$$

σ_y and σ_z are Pauli matrices. This is a two band system with energies

$$E = \pm\sqrt{(t\cos(k) - \mu)^2 + |\Delta|^2 \sin(k)^2} \tag{2.32}$$

The energy spectrum is gapless only when $|\mu| = |t|$. This critical point separates two physical regimes: the topological ($|\mu| < |t|$) and trivial ($|\mu| > |t|$) phases. In the topological phase, the system is metallic if there is no pairing term. It has fermion surface at two points $k = \pm \arccos(\mu/t)$. These are crossing points of the free fermion band and the free hole band. The pairing term opens a gap at these points and destroys the fermion surface. In the trivial phase, the system is a gapped insulator even when there is no pairing term. In Fig. 2.7, we plot the band structures at five different parameter sets. The pairing coefficient Δ is always set small compared with the other two. From left to right, the first one is the limiting case when $\mu = 0$,

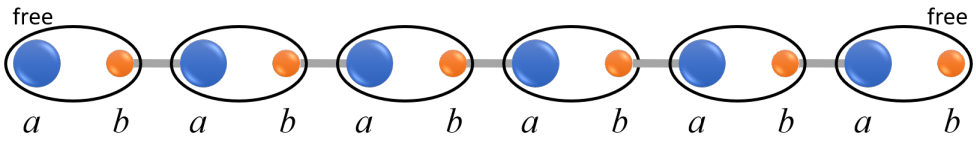


Figure 2.8: Limiting case when there are free Majorana modes at the ends of the chain. One black circle indicates one complex fermion. The big blue and small red dots are Majorana fermions.

while the last one is $t = 0$.

The topological invariant is given by the Berry phase.

$$\nu = \frac{1}{2\pi}\int_{-\pi}^{\pi} A(k)\mathrm{d}k, \quad A(k) = i\sum_{\alpha\in\mathrm{occ}}\langle u_\alpha(k)|\partial_k|u_\alpha(k)\rangle. \tag{2.33}$$

$A(k)$ is the Berry connection. The integral calculates the winding number, which is always an integer. Write the Bloch Hamiltonian as $\mathcal{H} = v_y\sigma_y + v_z\sigma_z$. Define point $V = (v_y, v_z)$ on the yz-plane. Intuitively, the winding number is the number of times that point V winds around the origin as k changes from $-\pi$ to π.

2.4.2 Majorana fermions

Define Majorana fermions as

$$\hat{a}_j = \hat{c}_j + \hat{c}_j^\dagger, \qquad \hat{b}_j = \frac{\hat{c}_j - \hat{c}_j^\dagger}{i}. \tag{2.34}$$

They satisfy conditions

$$\hat{a}_j = \hat{a}_j^\dagger, \quad \hat{b}_j = \hat{b}_j^\dagger, \quad \{\hat{a}_j, \hat{a}_k\} = 2\delta_{jk}, \quad \{\hat{b}_j, \hat{b}_k\} = 2\delta_{jk}, \quad \{\hat{a}_j, \hat{b}_k\} = 0. \tag{2.35}$$

We can express the p-wave superconductor Hamiltonian using Majorana fermions. With $\hat{c}_j = \frac{1}{2}(\hat{a}_j + i\hat{b}_j)$ and $\hat{c}_j^\dagger = \frac{1}{2}(\hat{a}_j - i\hat{b}_j)$, we write Eq. 2.29 as:

$$\hat{H}_{\mathrm{Majorana}} = \sum_j \left[\frac{-t+\Delta}{2}\, i\hat{a}_j\hat{b}_{j+1} + \frac{t+\Delta}{2}\, i\hat{b}_j\hat{a}_{j+1} + \mu\, i\hat{a}_j\hat{b}_j\right]. \tag{2.36}$$

Since we can always absorb the phase of Δ in to the definition of fermion operators, we assume Δ is a real number. When $t = \Delta$, the model becomes the Kitaev Chain [45].

In the limiting case when $t = \Delta$ and $\mu = 0$, the Majorana Hamiltonian with open boundary conditions does not contain operators $\hat{a}_1$ and $\hat{b}_L$. They correspond to the zero modes of the system. We plot such a chain in Fig. 2.8. The decomposition of complex fermion as Majorana fermions is represented as two dots

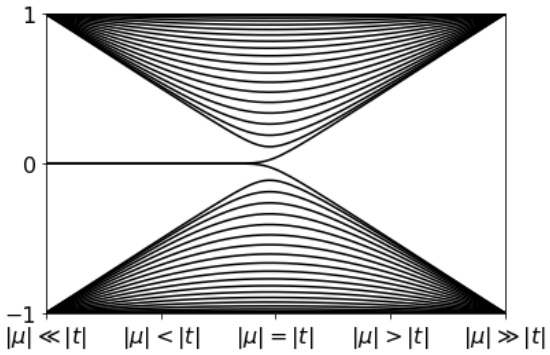


Figure 2.9: Energy bands at different parameters. There are two zero modes in the topological phase ($|\mu| < |t|$). They split at the phase transition point $|\mu| = |t|$.

within one circle. Majorana fermions from adjacent complex fermions are connected, while independent of each other if they are from the same complex fermion. Therefore, the edge Majorana fermions have no partners to interact with and are completely free.

To show the edge zero modes, we solve the Hamiltonian numerically with open boundary conditions. We set $\Delta = t$ and $t + \mu = 1$ for convenience. We used a chain of 20 sites. The energy spectrum is show in Fig. 2.9. Different from Fig. 2.7, the x-axis is parameters of the Hamiltonian instead of the momentum. We notice two degenerate zero-energy states appear only when $|\mu| < |t|$. The corresponding wave functions are localized at the edges, and decays exponentially into the bulk.

2.4.3 Jordan-Wigner transformation

Spin models and fermion models can have deep connections. They can be mapped to each other exactly under the Jordan-Wigner transformation.

The Jordan-Winger transformation is defined as

$$\hat{c}_j = \left(\prod_{k=1}^{j-1} \hat{\sigma}_k^z \right) \frac{\hat{\sigma}_j^x - i\hat{\sigma}_j^y}{2}. \tag{2.37}$$

Here $\hat{\sigma}$ are Pauli matrices representing operators for half spin. Then we have relations

$$\begin{aligned} \hat{c}_j^\dagger \hat{c}_j &= \frac{1 + \hat{\sigma}_j^z}{2} \\ \hat{c}_j^\dagger \hat{c}_{j+1} + h.c. &= -\frac{1}{2}(\hat{\sigma}_j^x \hat{\sigma}_{j+1}^x + \hat{\sigma}_j^y \hat{\sigma}_{j+1}^y) \\ \hat{c}_j \hat{c}_{j+1} + h.c. &= \frac{1}{2}(\hat{\sigma}_j^x \hat{\sigma}_{j+1}^x - \hat{\sigma}_j^y \hat{\sigma}_{j+1}^y) \end{aligned} \tag{2.38}$$

The p-wave superconductor Hamiltonian (Eq. 2.29) after Jordan-Wigner transformation is thus

$$\hat{H}_{\mathrm{XY}} = \sum_j \left[\frac{t+\Delta}{2} \hat{\sigma}_j^x \hat{\sigma}_{j+1}^x + \frac{t-\Delta}{2} \hat{\sigma}_j^y \hat{\sigma}_{j+1}^y + \mu \hat{\sigma}_j^z \right]. \tag{2.39}$$

This is the XY chain with magnetic field in the z direction. Again we set $t = \Delta$. The Hamiltonian reduces to

$$\hat{H}_{\mathrm{Ising}} = t \sum_j \hat{\sigma}_j^x \hat{\sigma}_{j+1}^x + \mu \sum_j \hat{\sigma}_j^z, \tag{2.40}$$

which is the transverse field Ising model with $-t$ as ferro magnetic pairing and μ as the external magnetic field.

This system has $\mathbb{Z}_2$ symmetry with the symmetry operator $\hat{P}_Z = \prod_{j=1}^L \hat{\sigma}_j^z$. It is easy to check that $\hat{P}_Z^2 = 1$, and $[H_{\mathrm{Ising}}, P_Z] = 0$. Note that $\hat{\sigma}_j^z = 2\hat{c}_j^\dagger \hat{c}_j - 1$. The $\mathbb{Z}_2$ symmetry of the spin chain then becomes the fermionic parity with $\hat{P}_F = \prod_{j=1}^L \left(2\hat{c}_j^\dagger \hat{c}_j - 1\right)$.

2.4.4 Entanglement analysis

Entanglement properties are widely used to detect topological phases. In this part we use the numerical methods introduced before to analyze the p-wave superconductor. It serves as an exercise for the techniques and gives a better understanding of the models.

The energy spectrum and the entanglement energy spectrum are the same for the two models under the Jordan-Wigner transformation. Mathematically, the transformation acts as a non-local unitary transformation between the two Hamiltonian, thus keeping the eigenvalues unchanged. Similarly, the entanglement Hamiltonian defined from the ground state can be related by the same unitary transformation. Therefore, the entanglement spectra are the same for the two systems.

In fig. 2.10, we show the low energy spectrum of both the energy and the entanglement energy of the p-wave superconductor (pSC). They are obtained using DMRG with system size $L = 200$ and bond dimension $m = 100$. Open boundary condition is used. The spectra of the transverse field Ising model (TFIM) are not shown because they overlap exactly with those of the pSC. This is consistent with our arguments above. Both the energy and the entanglement energy have double degeneracy in the topological phase. However, the origins of the degeneracy are different. The degeneracy in the ordered phase of the TFIM comes from spontaneous symmetry breaking. It can be understood from the limiting case when the transverse field is zero. Then the ground states are the perfect ferromagnetic with two possible directions of magnetization. In the pSC, the ground state degeneracy comes from the edge zero modes which are Majorana fermions. In the trivial phases there are some accidental degeneracies in the excited states, but the ground levels are always

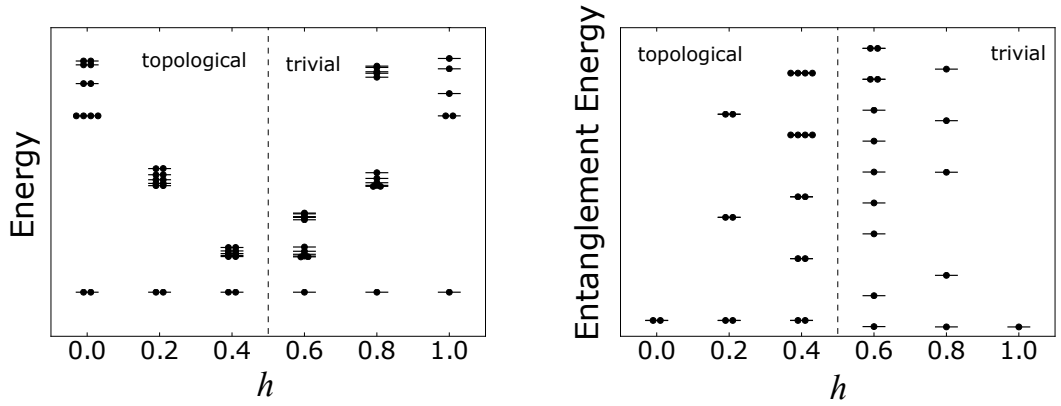


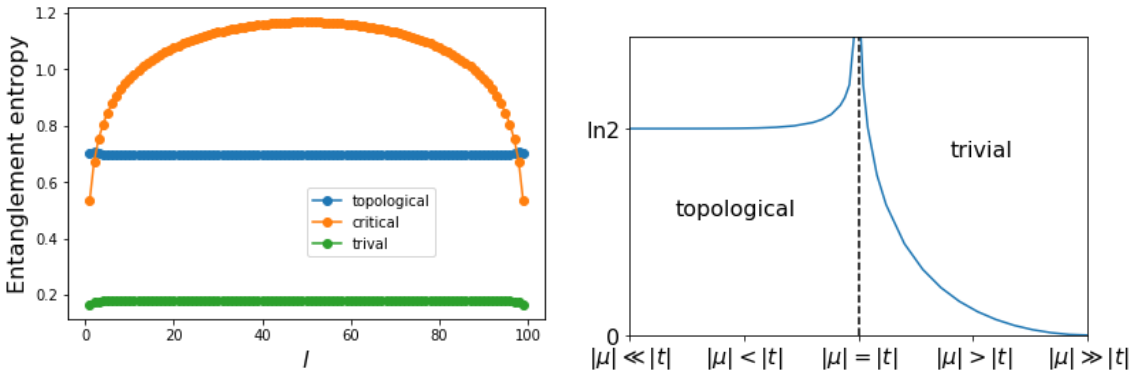
Figure 2.10: The (a) energy spectrum and (b) entanglement energy spectrum of the p-wave superconductor. We set $J = 1 - h$. The spectra of the transverse field Ising mode are exactly the same. Double degeneracy is seen for all levels in the topological phases for both the energy and the entanglement energy.

non-degenerate. The boundary conditions have different effects on the degeneracies of the two models. If we switch to periodic boundary conditions, the degeneracy in pSC energy spectrum disappears because the boundary term that connects the two zero modes lifts the degeneracy. On the other hand, the TFIM is not affected.

Now we focus on the entanglement entropy of the pSC. We still use the open boundary condition. Recall that entanglement entropy is calculated for one bi-partition of the system. For a chain with L sites, we can cut the chain at $L - 1$ different positions. We plot the entanglement entropy at all possible cuts in Fig. 2.11(a) for topological phase, trivial phase, and at the transition point which is also known as a critical point. We set $L = 100$. The upper line is the typical curve for the entanglement entropy in the topological phase ($t = \Delta = 2$ and $\mu = 1$). When the cut is close to the edge, one of the subsystem is smaller than the correlation length. The entanglement entropy grows as the cut gets into the bulk. For the pSC, it saturates at $\ln 2$. The trivial phase (we chose $t = \Delta = 0.5$ and $\mu = 1$) always has a small entanglement entropy as shown by the lower line. The middle curve is the entanglement entropy at the critical point. It satisfies Eq. 2.7.

The central cut entanglement entropy of the system as parameter changes is shown in Fig. 2.11(b). We set $t = \Delta = 1 - \mu$ for convenience. In the bulk of the topological phase, the entanglement entropy is always $\ln 2$. The value diverges near the transition point, and decreases after entering the trivial phase. Note that the entanglement entropy becomes zero only in the atomic limit.

The central charge at the critical point can be extracted from the entanglement entropy S with Eq. 2.7. We calculated the entanglement entropy at different cuts for open and periodic boundary conditions with $L =$



(a) (b)

Figure 2.11: (a) Entanglement entropy at different cuts l for topological (upper line), critical (middle), and trivial (lower line) phases. In the topological phase, the entanglement entropy is around $\ln 2$. At the critical point, the entanglement entropy increases with the subsystem size. The trivial phase always has a low entanglement entropy. (b) Central-cut entanglement entropy at different parameter sets. It is a constant $\ln 2$ in the topological phase until reaches the transition point, and then decreases gradually in the trivial phase.

100. Recall that S is linear with $\ln(\sin(\pi l/L))$. We plot S verses x in Fig. 2.12. We set $x = \frac{1}{6}\ln\left(\frac{2L}{\pi}\sin\frac{\pi l}{L}\right)$ for open boundary conditions and $x = \frac{1}{3}\ln\left(\frac{L}{\pi}\sin\frac{\pi l}{L}\right)$ for open boundary conditions. A clear linear relation can be seen for both boundary conditions. A fitted line for each of them is also shown in the figure. The central charge calculated from the fitting is printed in the legend. Although we use the same system size, the central charge from the periodic boundary condition is closer to the true value $c = 1$. The small deviation from the true value comes from the finite size effects near the edges. We take values only in the middle of the chain (from $\frac{1}{4}L$ to $\frac{3}{4}L$), and do the fitting again. Now the central charge is 1.016 and 1.000018 for open and periodic boundary conditions respectively. Compare with the results using all the data, we see a lot of improvements especially for the periodic boundary condition.

2.4.5 Finite size effects

In this part, we discuss finite size effects in numerical simulations. Recall that there are two edge zero modes in the topological phase of an open chain. In general, the eigenvectors with the same eigenvalues are not uniquely determined. Linear combinations of the zero-energy states are also eigenstates of the Hamiltonian with zero energy. When the length of the chain is comparable or smaller than the correlation length, the two edge modes can interact with each other and lift the degeneracy. Therefore, we do not have the freedom to choose the ground state we want to work with for small systems.

We plot the central-cut entanglement entropy (blue dots) for different system sizes and the corresponding energy gap (red line) between the two zero modes in Fig. 2.13(a). We set $t = \Delta = 0.7$ and $\mu = 0.3$ while

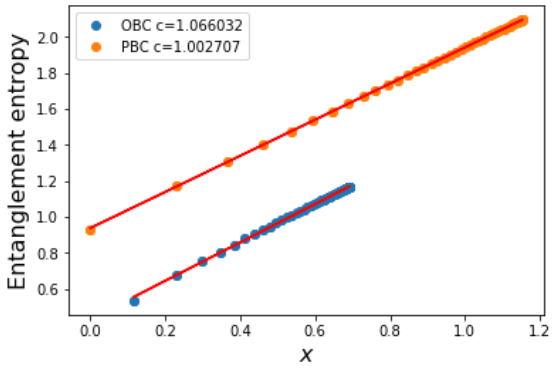


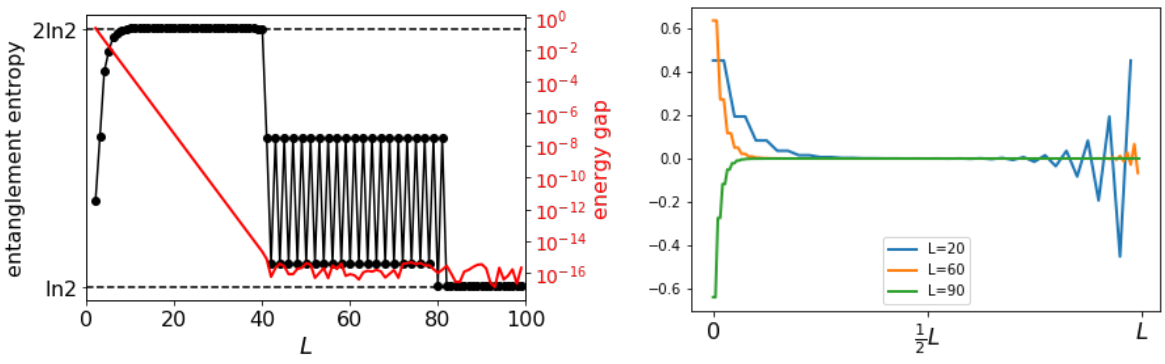
Figure 2.12: Entanglement entropy at the transition point for different cut position. We use both the open boundary condition (OBC, lower line) and the periodic boundary condition (PBC, upper line). x is $\frac{1}{6}\ln\left(\frac{2L}{\pi}\sin\frac{\pi l}{L}\right)$ for OBC and $\frac{1}{3}\ln\left(\frac{L}{\pi}\sin\frac{\pi l}{L}\right)$ for PBC. The fitted straight lines are also plotted.

varying the system size L. When L is small ($L < 10$), the entanglement entropy grows quickly with the system size, then saturates at $2\ln 2$. The $2\ln 2$ can be understood as two parts. The cut in the middle of the chain contributes one $\ln 2$. Besides, the zero modes at the two ends of the chain are entangled to lower the energy. The cut of this entangled state gives an additional $\ln 2$. This can be verified by plots of the energy gap and wave functions. The energy difference between the two zero modes are plotted in Fig. 2.13(a) as a red line. The values are plotted in a log scale. With small system size L, the gap reduces exponentially with L until it reaches machine accuracy. The entanglement entropy remains as $2\ln 2$ until the gap closes exactly. The typical wave function of the zero mode that is filled is plotted in Fig. 2.13(b). The blue line with non-zero values at both ends is the amplitude of the zero mode at different sites for $L = 20$. It is a combination of the two edge modes.

When L is large ($L > 80$), the entanglement entropy drops to $\ln 2$. There is no contributions from the entangled state as in the short chain case. This is consistent with our argument, because the two edge modes in this case are too far away to interact with each other. The eigen problem solver just picks one edge state. As shown in Fig. 2.13(b) by the green line, the wave function only has non-zero values at one end of the chain.

When L is neither small nor large ($40 < L < 80$), the system is in an intermediate state with the two edge modes partially entangled. The entanglement entropy oscillates as the size of the chain changes between even and odd numbers. The zero mode has non-zero values at both ends of the chain (orange line in Fig 2.13(b)), but the amplitudes at the right side is significantly smaller than those at the left side.

We Found similar effects in other one dimensional models with edge zero modes, such as the Su-Schrieffer-Heeger (SSH) chain. Same as the pSC, the zero modes of a short SSH chain are located at two edges. The



(a) (b)

Figure 2.13: (a) The central-cut entanglement entropy for different system size. We set $t = \Delta = 0.7$ and $h = 0.3$. The corresponding energy difference between the two zero modes is plotted as a red line. The scales are on the right. (b) The wave functions of the filled zero mode at different system sizes. All of them are edge modes. The blue line with non-zero values at both ends is for $L = 20$. The orange line is calculated when $L = 60$. It has high amplitude on the left but low on the right. The state with $L = 90$ only has one peak on the left as shown by the green line.

only one difference is that the wave function oscillates at both ends for the SSH chain, while the wave function decrease monotonically at one end and oscillates at the other end for the pSC.

There are several ways to eliminated such effects. The most straight forward way is to use larger system sizes, but it sometimes consumes too much computation resources. We can apply periodic boundary conditions, or cut the middle part out of an open chain. In this case, the two edge modes are in the same area. How do they interact each other will not affect the entanglement calculation at the cut. For central-cut entanglement entropy of an open chain, we can add a small symmetry breaking term to split the degenerate states into the corresponding symmetry sectors.

Chapter 3

Phase diagram of the Z_3 parafermionic chain with chiral interactions

In this chapter, we characterize the phase diagram of a parafermionic chain using numerical methods.[1] Parafermions are exotic quasiparticles with non-Abelian fractional statistics that can be realized and stabilized in 1-dimensional models that are generalizations of the Kitaev p-wave wire [46]. We study the simplest generalization, i.e. the Z_3 parafermionic chain. Using a Jordan-Wigner transform we focus on the equivalent three-state chiral clock model [47], and study its rich phase diagram using the density matrix renormalization group (DMRG) technique [27, 43] . We perform our analyses using quantum entanglement diagnostics which allow us to determine phase boundaries, and the nature of the phase transitions. In particular, we study the transition between the topological and trivial phases, as well as to an intervening incommensurate phase which appears in a wide region of the phase diagram [48, 49]. The phase diagram is predicted to contain a Lifshitz type transition [50] which we confirm using entanglement measures. We also attempt to locate and characterize a putative tricritical point in the phase diagram where the three above mentioned phases meet at a single point.

This chapter is arranged as follows. We first talk about the background of this work in the first section, and then discuss the details the entanglement criteria used to map out the phase diagram in the next section. The DMRG algorithm gives immediate access to the entanglement entropy (EE), and therefore the central charge, at critical points/regions in the phase diagram [40]. The third section introduces the model we will work on and its properties. The general features of the phase diagram and regions of the topological phase (where para-fermion boundary modes may exist) are in section four. We also discuss the nature of the phase transitions out of the topological phase, and the numerical difficulties encountered when locating the transition point. Section five contains our observations pertaining to a critical incommensurate phase, and the possibility of a tricritical point in the phase diagram at the intersection of the topological, trivial, and incommensurate phases. We also find a region of the phase diagram which exhibits the critical entanglement features of a Lifshitz transition. Finally, we conclude by discussing future directions and possible relevance

[1] The material presented in this chapter was published in: Ye Zhuang, Hitesh J. Changlani, Norm M. Tubman and Taylor L. Hughes, Phys. Rev. B 92 (3), 035154 (2015); Some of the figures and their captions in this chapter are reprinted from these publications with minor modifications.

to experiments looking for para-fermions in the last section.

3.1 Background

There has been concerted effort to engineer systems with stable Majorana bound states, and other anyonic quasiparticles, for use in the topological quantum computation architecture [51, 52, 53, 54, 55, 56, 57]. For example, there has been recent progress in attempts to isolate Majorana bound states in quantum nanowires [55, 58, 59, 60] and in superconductor surfaces implanted with a line of magnetic impurities [61]. These quasi-1D systems effectively realize a version of the Kitaev p-wave wire model [45], and are predicted to have a gapped topological phase which supports characteristic Majorana bound states at the ends of the wire.

While the boundary modes in these heterostructure systems are non-Abelian anyons, they are unfortunately known to beinsufficient for universal quantum computation. A possible remedy for this problem has been to look for more exotic non-Abelian excitations. For example, Fendley has recently suggested exploring one-dimensional Z_N para-fermionic models which support topological phases with more computationally efficient non-Abelian anyon bound states [46]. Still, the Z_N non-Abelian anyons are not able to perform universal quantum computation, however they can be leveraged to create a 2D phase with Fibonaccci anyons, which are universal [62]. These promising features have spurred wide spread interest in these models, and has led to many analytical and numerical studies, including several experimental proposals for realizing these topological phases [63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87].

In this work, we continue along these lines of research by exploring the rich phase diagram of the Z_3 para-fermionic chain; though for ease of calculation we actually study the Jordan-Wigner transformed para-fermionic chain [47], including chiral interactions. The resulting model is the three state *chiral* clock model. This model re-surfaced in this context in Ref. [46] as a candidate for exhibiting non-Abelian bound states beyond Majorana fermions. It was shown analytically that para-fermionic boundary zero modes can exist in this model when spatial-parity and time-reversal symmetries are broken via chiral interactions [46]. This was verified numerically in Ref. [88], which confirms that chiral interactions can help to stabilize the boundary zero modes, although the zero modes themselves are more fragile than one might initially expect.

Here we are interested in studying the full phase diagram of the chiral clock model as a function of two chiral-interaction phase-parameters (θ, ϕ), as well as the relative strength of the nearest neighbor coupling (J) to the local Zeeman field (f). Using entanglement techniques, we have been able to locate the phase boundaries that separate the topological phase from the trivial gapped phase, and a critical incommensurate

phase, the latter of which has no analog in the Kitaev p-wave wire model. We have conclusively identified the region in which there is a topological phase, and have explored the nature of the quantum phase transitions in and out of the three adjoining phases. In addition, by studying oscillatory properties of the system in, or near, the incommensurate phase, we establish the approximate location of a putative tricritical point[48, 49], and further support the entanglement signatures that were recently proposed for identifying Lifshitz transitions[50].

3.2 Z_3 chiral clock model and para-fermions

3.2.1 The Hamiltonian

For our study we use the 1D 3-state (Z_3) chiral clock model [48, 89, 46, 90]. The Hamiltonian for the 3-state chiral clock model is:

$$H_3 = -f \sum_{j=1}^{L} \tau_j^\dagger e^{-i\phi} - J \sum_{j=1}^{L-1} \sigma_j^\dagger \sigma_{j+1} e^{-i\theta} + \text{h.c.} \tag{3.1}$$

following the notation in previous work [46], where f, J, θ and ϕ are scalar parameters, and σ_i and τ_i are local three state spin operators on site i. The spin operators have the properties $\tau^3 = \sigma^3 = I$, $\sigma\tau = \omega\, \tau\sigma$, where $\omega = e^{2\pi i/3}$. Specifically, we use the matrix representation

$$\tau = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix}, \quad \sigma = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

The chiral clock model is related to the para-fermionic chain through a Jordan-Wigner transformation [47, 46], similar to the well-known, analogous case that the Kitaev p-wave wire is related to the transverse-field Ising model via the same type of transformation. The parafermion operators are defined as

$$\chi_j = \left(\prod_{k=1}^{j-1} \tau_k \right) \sigma_j, \tag{3.2}$$

$$\psi_j = \left(\prod_{k=1}^{j-1} \tau_k \right) \sigma_j \tau_j, \tag{3.3}$$

at site j. The corresponding para-fermionic Hamiltonian is

$$H_3 = -f \sum_{j=1}^{L} \psi_j^\dagger \chi_j e^{-i\phi} - J\omega^2 \sum_{j=1}^{L-1} \psi_j^\dagger \chi_{j+1} e^{-i\theta} + \text{h.c.} \tag{3.4}$$

3.2.2 Symmetry analysis

The chiral clock model has a global Z_3 symmetry that can be represented with $\chi \equiv \prod_{j=1}^{L} \tau_j^\dagger \equiv e^{\frac{2\pi i}{3}\mathcal{Z}}$. Here $\mathcal{Z}$ is the generator of the symmetry, and has three different eigenvalues 0,1,2. In addition, when all of the coefficients of the Hamiltonian are real, i.e. when the system is a Z_3-ferromagnet or Z_3-antiferromagnet Hamiltonian, then the Hamiltonian is invariant under time-reversal, charge-conjugation, and parity symmetries. This can be easily seen from the following definitions of these symmetries. Charge conjugation C acts on the spin operators via $C\sigma_j C = \omega^2 \sigma_j^\dagger$, $C\tau_j C = \tau_j^\dagger$, $C^2 = 1$. As an aside, note that charge conjugation, together with the Z_3 symmetry, forms the S_3 permutation symmetry, i.e. the symmetry obeyed when the 3-state clock model is restricted to the 3-state Potts model. Time reversal T acts on the spin operators via $T\sigma_j T = \sigma_j$, $T\tau_j T = \tau_j^\dagger$, $T^2 = 1$, and complex conjugates any scalar coefficients. Spatial parity P acts on the spin operators via $P\sigma_j P = \sigma_{-j}$, $P\tau_j P = \tau_{-j}$, $P^2 = 1$. Finally, we note two things: (i) due to the symmetry of the Hamiltonian with respect to ϕ and θ, we only need to consider the region of the phase diagram where ϕ and θ each range from 0 to $\frac{\pi}{3}$, and (ii) for $f = J$, the system is self-dual along the line $\phi = \theta$.

We first address the property (i) in detail: the Hamiltonian in Eq. 3.1 has the same properties when either of the two phases θ, ϕ are shifted by multiples of $\frac{2}{3}\pi$. To show this we can see that the transformation

$$\theta' \to \theta + \frac{2n\pi}{3} \quad \phi' \to \phi + \frac{2m\pi}{3} \tag{3.5}$$

changes the Hamiltonian to:

$$H_3 = -f\omega^{-n} \sum_{j=1}^{L} \tau_j^\dagger e^{-i\phi} - J\omega^{-m} \sum_{j=1}^{L-1} \sigma_j^\dagger \sigma_{j+1} e^{-i\theta} + \text{h.c.} \tag{3.6}$$

Then we can redefine the operators:

$$\tau' = \omega^{-n}\tau \quad \sigma'_{2j} = \omega^{-m}\sigma_{2j} \quad \sigma'_{2j+1} = \sigma_{2j+1}. \tag{3.7}$$

This new set of operators preserves the properties $\tau^3 = \sigma^3 = I$, $\sigma\tau = \omega\, \tau\sigma$, where $\omega = e^{2\pi i/3}$. After this redefinition we end up with a Hamiltonian with the same form as the original.

Additionally, the transformation that flips the signs of the two phases at the same time, i.e.,

$$\theta' \to -\theta \quad \phi' \to -\phi \tag{3.8}$$

changes the Hamiltonian to,

$$H_3 = -f \sum_{j=1}^{L} \tau_j e^{-i\phi} - J \sum_{j=1}^{L-1} \sigma_j \sigma_{j+1}^{\dagger} e^{-i\theta} + \text{h.c.} \tag{3.9}$$

Here we can redefine the operators as

$$\tau' = \tau^{\dagger} \quad \sigma' = \sigma^{\dagger} \tag{3.10}$$

to recover the form of the original Hamiltonian.

We can also just flip the sign of just one of the phases, say $\phi' \to -\phi$, and then the redefinition:

$$\tau'_j = \tau_{-j} \quad \sigma'_j = \sigma_{-j} \tag{3.11}$$

leaves the Hamiltonian unchanged. If instead we flipped the sign of θ, will need a transformation that involves both Eqs. 3.10 and 3.11.

Now, we can consider the property (ii) duality transformation:

$$\mu_{j+\frac{1}{2}} = \prod_{k=1}^{j} \tau_k, \quad \nu_{j+\frac{1}{2}} = \sigma_j^{\dagger} \sigma_{j+1}. \tag{3.12}$$

These dual operators satisfy $\mu^3 = 1$, $\nu^3 = 1$, and $\mu\nu = w\, \nu\mu$. The dual Hamiltonian is then

$$H_3^{\text{dual}} = -J \sum_{j=1}^{L-1} \nu_{j+\frac{1}{2}}^{\dagger} e^{-i\theta} - f \sum_{j=1}^{L} \mu_{j-\frac{1}{2}}^{\dagger} \mu_{j+\frac{1}{2}} e^{-i\phi} + \text{h.c.} \tag{3.13}$$

Comparing with the original Hamiltonian, the dual Hamiltonian returns to the original form if we exchange θ and ϕ, and at the same time J and f.

3.2.3 Analytical results

There are many previously known results about this model (Eq.3.1), beginning with the original proposals of Ostlund [48] and Huse [89]. For example, the corresponding two-dimensional classical Hamiltonian for $\phi = 0$ was studied in Ref. [48], and the one-dimensional quantum Hamiltonian was studied in Ref. [49] for the restricted case $\phi = \theta$. One of the most important early results is that Eq. 3.1 has a second order quantum phase transition at $f = J$ when $\theta = \phi = 0$. At this point the model realizes the full S_3 permutation symmetry (instead of just Z_3), and the critical point is described by the critical conformal field theory for the 3-state Potts model, which has central charge 4/5 [91]. In addition, the line $f\cos(3\phi) = J\cos(3\theta)$ [92]

is known to be integrable and $\phi = \theta = \frac{\pi}{6}$ is super integrable [93, 94]. Despite this, the knowledge of the location of some important critical points and their associated properties is an open question.

Generically, it is known that the phase diagram is divided up into two gapped regions, one of which is identified with small values of f (compared with J), and the other with large values of f. These regions are separated by continuous quantum phase transitions that we will identify and discuss further below. Using a more modern terminology, the gapped phase for small f is a symmetry broken phase of the 3-state clock model and it exactly corresponds to the "topological" phase in the Jordan-Wigner transformed para-fermionic chain. The gapped phase for large f is a disordered phase of the 3-state clock model, and maps onto the "trivial" phase of the para-fermionic chain. This gives another example of a case where the degeneracy associated to symmetry breaking is mapped to topological degeneracy via the Jordan-Wigner transformation [95, 96]. Hence, in either representation this phase has a three-fold ground-state degeneracy, which can be detected by measuring the ground state EE. On the other hand, the trivial phase is equivalent to the spin disordered phase, which does not have a generic ground-state degeneracy. The parameter f is thus an important tuning parameter for the phase diagram, and analogous to the external transverse field in the Ising model.

While we expect these general features to pervade the phase diagram, the phase space for generic θ and ϕ is largely unexplored. Additionally, it is known that the combination of the Z_3 symmetry and the chiral nature of the interactions, gives rise to interesting behavior that cannot be found in the Majorana/Ising case. For example, this model supports a so-called "incommensurate phase" which is not present in the transverse-field Ising model with chiral interactions [48].

This motivates the main objective of our article, which is to characterize the phases and the nature of the phase transitions over the entire phase space. We will show that there are two types of phase transitions that occur to destabilize the topological phase, and there is a large region of critical incommensurate phase that separates the topological from the trivial phase over a wide range of parameters.

3.3 Phase diagrams

3.3.1 Phases and transitions

We primarily use the spatial entanglement entropy (EE) to characterize the phase diagram. This measure has been widely used to detect topological order in 2D [33, 34], and has been applied more recently to 1D topological phases [36]. To arrive at our results for the phase diagram (and to obtain reasonable estimates of the phase boundaries in the thermodynamic limit), we simulated Hamiltonians using open-boundary

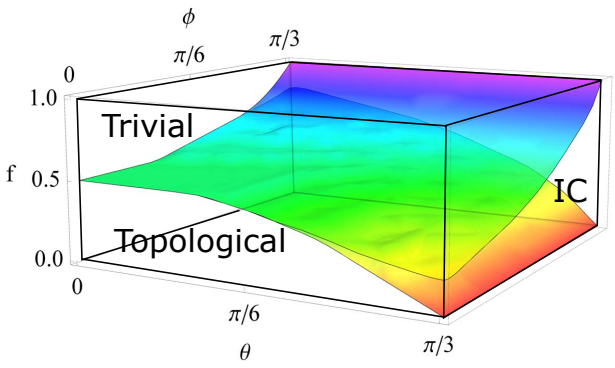


Figure 3.1: Three-dimensional phase diagram of the chiral 3-state clock model in terms of f, θ and ϕ with $J = 1 - f$. For details of the Hamiltonian see Eq. 3.1. The topological, trivial, and incommensurate (IC) phases are indicated. The coloring is a function of the value of f at the critical surface separating the phases. The dashed line that connects points (0, 0, 0.5) and ($\pi/3$, $\pi/3$, 0.5) is the self-dual line.

DMRG with 100 sites, and a bond dimension $m = 100$. We find this to be sufficient for the phases with low entanglement entropy. For the critical phases, additional checks were performed with bond dimension $m = 200$. For establishing characteristics of other phases, for example, the region of critical incommensurate phase, larger lengths of 400 sites were also tested.

First, we present the full three-parameter phase diagram (f,θ,ϕ) over the reduced domain in Fig. 3.1, where we have set $J = 1 - f$. The basic topology of the phase structure is clear. We find three distinct phases as mentioned above. The phase corresponding to largest f values is generically the trivial phase, and the phase corresponding to the smallest f values is generically the topological phase. They share a common/direct phase boundary between them when θ and ϕ are small. For large θ or ϕ, an intermediate incommensurate phase appears between the two.

We show the central-cut EE in Fig.3.2(a),(b),(c) for several 2D cross-sections of the 3D phase diagram. These plots help to identify the gapped phases and the topology of the phase boundaries. To more clearly identify the nature of the critical regions/boundaries we also calculate the central charge via the scaling relation. It is interesting to see that the observed locations of the phase boundaries for cross sections $\phi = 0$ and $\theta = \phi$ are broadly consistent with earlier works [48, 49], and that the topological phase itself is stable over a large part of the phase diagram [2].

We indicate several special points on these cross sections: Point A in Fig. 3.2(a) and Fig. 3.2(c) is the transition point of the three-state Potts model associated with $c = 4/5$ [91], and Point B and C are putative tri-critical points. We indicate approximate locations of the phase boundaries with solid, dashed, or dot-dashed lines, depending on the nature of the phase transition, as indicated in the figure caption. Finite size

[2]By this we mean that the system remains gapped and the topological ground-state degeneracy is robust. We do not mean that the edge zero-modes remain exact over the entire phase range. See Ref. [88] for discussion on this distinction.

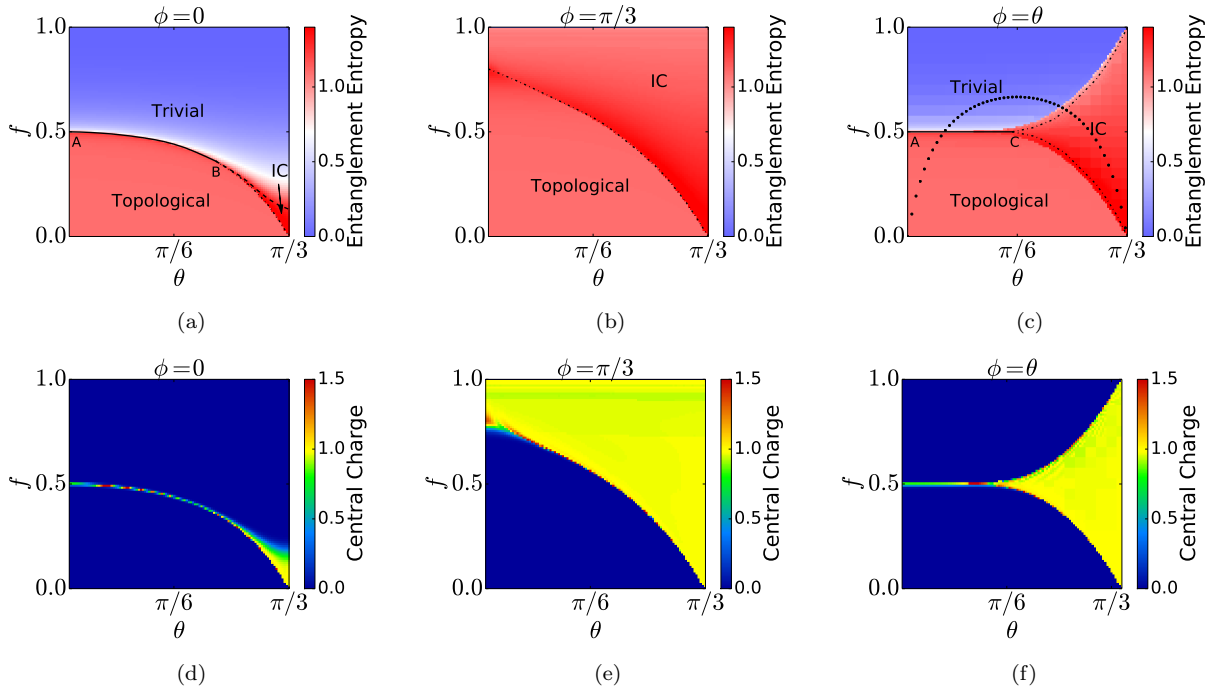


Figure 3.2: Three cross-sections corresponding to (a) $\phi = 0$ (b) $\phi = \pi/3$ and (c) $\phi = \theta$ of the three dimensional phase diagram, and all for $L = 100$. Topological, trivial, and incommensurate (IC) phases are identified by the central-cut entanglement entropy (color coded). For (a) and (b) a 2D grid in increments of 0.01 was used to resolve fine features of the transitions. (c) was mapped out on a 2D grid in increments of 0.05. Point A is the transition point of the 3-state Potts model, i.e. the chiral clock model for ($\theta = \phi = 0$). Points B and C are Lifshitz points and are associated with putative tricritical behavior. The solid lines, dashed lines, and dotted-dashed lines indicate direct topological-trivial ($c = 4/5$) type, Kosterlitz-Thouless type, and Pokrovskii-Talapov type [97] transitions respectively. The thick circularly-dotted line represents an upper bound on the region where exact parafermionic zero modes can exist [46]. Panels (d), (e) and (f) show the corresponding central charges for cross sections (a),(b),(c) respectively. The IC phase is associated with central charge $c = 1$ (yellow) whereas the critical regions close to point A have $c = 4/5$ (green).

effects were checked around specific points along the critical lines by running system sizes of L=100 to 400 on a finer grid. The locations of these lines did not change significantly in comparison to the resolution of our grid, except in certain regions which are discussed in further detail in later sections.

From the central-cut EE we see that the trivial phase is characterized by a small EE, while the topological phase has a nearly uniform EE of $\approx \ln 3$ indicating a three-fold degeneracy of the ground state. The change of EE is abrupt between the two phases as can clearly been seen in Fig. 3.2(a) and Fig. 3.2(c) for $\theta \lesssim \pi/4$ and $\theta \lesssim \pi/6$ respectively. We also verified that this transition is accompanied by a divergence in the second order derivative of the ground state energy (not shown).

The third phase in the phase diagram is the incommensurate phase. This is a critical phase in which the correlation functions generically behave as $A(r)e^{(2\pi i/3)Qr}$, where A decays algebraically and Q is irrational. The oscillatory properties of the correlation functions also manifest themselves in oscillatory behavior seen in energy gaps, which we address later. Although there is not an extremely sharp distinction between the central-cut EE for the topological and incommensurate phases, the EE scaling with system size is markedly different. The former has an EE that quickly saturates to a constant value of $\ln 3$ with sub-system size, while the latter has EE that diverges logarithmically with sub-system size. By fitting our data to Eq. 2.7, we establish that the incommensurate phase is critical and its central charge is $c = 1$ over the entire phase.

3.3.2 Extract central charge near critical points

While constructing the detailed phase diagram cross sections, we found that while it was easy to approximate the locations of the phase boundaries, we often encountered difficulties in precisely nailing down the central charge of the corresponding critical points. As an example, we note the appearance of a few points with (apparently) high central charge, indicated by red color, on the direct topological-trivial phase boundary in Fig. 3.2(d). While in some cases there may be real physics associated to this behavior, A primary source for these spurious effects is fitting to a region of the phase diagram that is just slightly off-criticality. We show that the central charge is very sensitive to the precise location of the critical point, and can easily give $O(1)$ errors even when only slightly tuned away from criticality, and even with reasonably large-size calculations.

When performing the fit to EE data obtained from a finite size system, and for a point in parameter space that is close to (but not at) a critical point, it is often difficult to obtain a reasonable estimate of the central charge. One possible explanation is that, when the system size is smaller than the correlation length, the fit to Eq. 2.7 may appear to be good, but the central charge obtained from the fit may not match the actual central charge of the nearby critical point. This is not unique to our model, and we were also able to observe this effect for free Dirac fermions with a tunable mass term as shown below. Eventually, if the

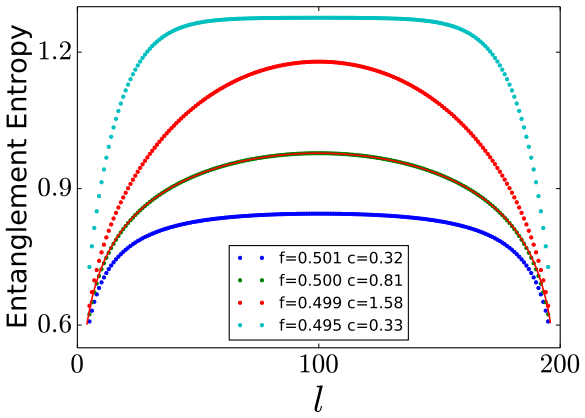


Figure 3.3: The EE as a function of the subsystem size l at $\phi = \theta = 0$ and several different f close to or at the critical point ($f = 0.50$). From the highest curve to the lowest one, the corresponding f is 0.495, 0.499, 0.500, and 0.501. The central charges obtained from the fitting are shown in the legend. For $f = 0.495$ and $f = 0.501$, a plateau in the EE is seen indicating a gapped phase. For $f = 0.499$, an apparent critical phase is seen which is attributed to an artifact of finite size effects.

system is tuned off criticality, and when the system size is larger than the correlation length, the EE will saturate and hence reveal the gapped phase.

To provide an example of such behavior, we refer to known analytic results that the central charge should be 4/5 at ($f = J = 0.5$, $\phi = \theta = 0$), and zero for all other f at $\phi = \theta = 0$. In Fig. 3.3, we show that at the critical point $f = J = 0.5$, the central charge is $c = 0.81 \pm 0.01$, close to the analytical result. However, when we are slightly away from this point, say $f = 0.499$, the system still appears critical with an (apparent) central charge of $c = 1.58$, much larger than the expected value of 0.80. On going slightly further away, $f = 0.495$, a plateau in the EE profile is seen consistent with our expectation of a gapped phase. Thus, the fitting procedure produces misleading results in the neighborhood of the critical point, and can make it difficult to determine the central charge for critical points in which the position of the point is not known to extremely high accuracy.

To further confirm our discussion above we performed similar calculations for 1D gapless Dirac fermions using exact diagonalization. We use the 2-band free-fermion lattice Dirac model as the test model

$$\begin{aligned} H = \quad & -\sum_n \left(i c^\dagger_{n+1,\uparrow} c_{n,\downarrow} + i c^\dagger_{n+1,\downarrow} c_{n,\uparrow} + h.c. \right) \\ & -\sum_n \left(c^\dagger_{n+1,\uparrow} c_{n,\uparrow} - c^\dagger_{n+1,\downarrow} c_{n,\downarrow} + h.c. \right) \\ & +(2-m) \sum_n \left(c^\dagger_{n,\uparrow} c_{n,\uparrow} - c^\dagger_{n,\downarrow} c_{n,\downarrow} \right) \end{aligned}$$

This model is gapless at $k = 0$ if m is zero, and the critical point should have a central charge of 1. If m is tuned away from zero the system exhibits an energy gap of the size 2m. For our entanglement calculations

the system was filled to half filling, such that when it is gapless, the filling hits exactly at the Dirac node, and if it is gapped, the filling includes all the states in the lower band. In this model the correlation length is controlled by the scale 1/m (with units it would be $\hbar v_F/m$ but $\hbar$ and v_F are effectively unity for our model).

To compare closely with our DMRG results we fit the central charge of this model using entanglement scaling with open boundary conditions. When gapless, we find the central charge to 2 or 3 digits of accuracy. For example, we find c=1.006 when the chain is of length 400. In addition to calculating the scaling law over the entire chain we can improve the fit by taking symmetric cuts around the center of the chain which reduces the edge effects. We get slightly improved accuracy for ranges such as 120-280, i.e., c=1.004. If we increase system size to L=500 and fit over 120-380 we find c=1.003.

	40-360	120-280	40-80
L=300	1.0198	1.0202	1.0199
L=400	1.022	1.024	1.021
L=500	1.0243	1.0273	1.0219.

Table 3.1: The central charge obtained by fitting from different region of the system and different system size L. The mass gap is set to be m=1/10000

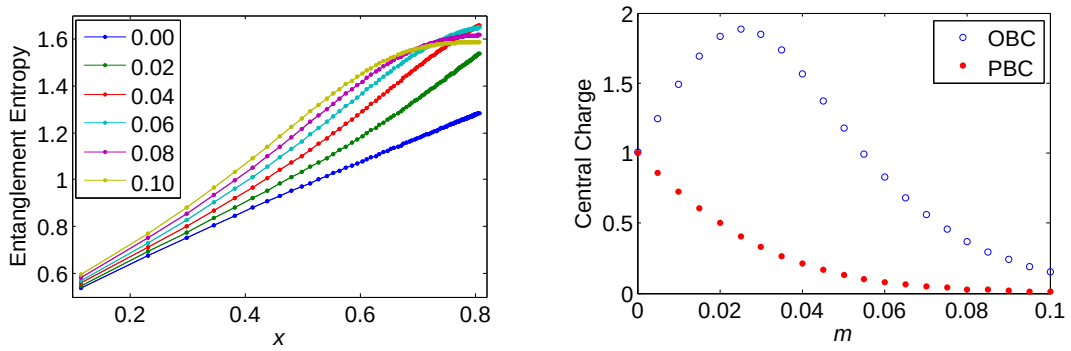
Now let us perturb the system slightly away from the critical point. For this test we turn on a gap size of m=1/10000 as a start. As an estimate, this should give a correlation length of ξ =10000 sites. For system size 400, if we fit from 40-360, we find c=1.022; if we fit from 120-280 we find c=1.024. If we try to fit a different range, e.g., 40-80 we find c=1.021. Either way, the result is already 1% different than the gapless case even for this tiny gap (compared with the bandwidth). Next we repeated the same 3 fits for L=300 and we find c= 1.0198, 1.0202, 1.0199. And then for L=500 and find c=1.0243,1.0273,1.0219. These results are summarized in Table. 3.1. We observe that the fits get worse when we increase the system size, and when we fit over the region restricted mostly to lie over the center. The latter result may be expected since the scaling function varies most slowly over the center. The fact that the fits get worse as we increase system size is most likely just an indicator that there is a finite correlation length and that the critical scaling form will eventually break down. For additional tests we also fit the central charge for larger (but still very small) mass gaps with $m = 1/1000$ and $m = 1/100$ in Table. 3.2 and Table. 3.3 respectively.

	40-360	120-280	40-80
L=300	1.1374	1.1633	1.1230
L=400	1.1699	1.2082	1.1372
L=500	1.2014	1.2499	1.1473

Table 3.2: The central charge obtained by fitting from different region of the system and different system size L. The mass gap is set to be m=1/1000

	40-360	120-280	40-80
L=300	1.8811	1.7267	1.9246
L=400	1.7563	1.4033	1.9438
L=500	1.5701	1.0878	1.9385
L=600	1.3874	0.84507	1.9273

Table 3.3: The central charge obtained by fitting from different region of the system and different system size L. The mass gap is set to be m=1/100

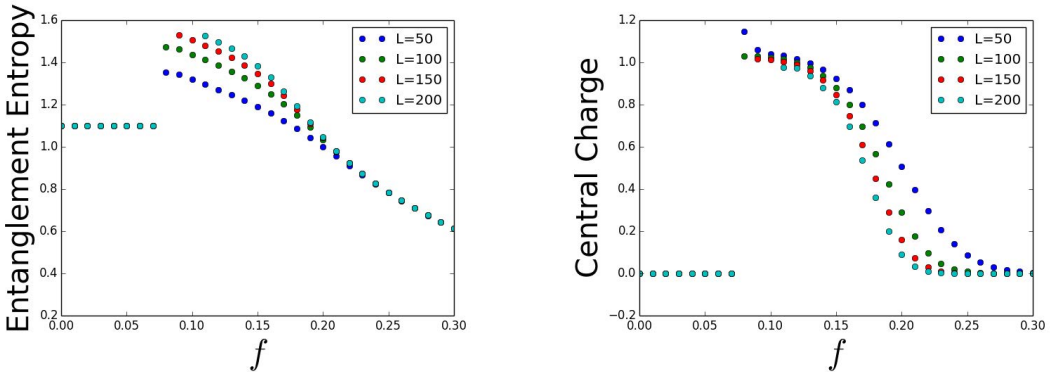


(a) (b)

Figure 3.4: (a) The transformed curve of entanglement entropy. Here x is defined as $\ln(\frac{L}{\pi}\sin\frac{\pi l}{L})$, where l is the block size and the system size $L = 200$. For small x the curves are ordered by increasing the mass gap (lowest mass is the lowest curve). (b) The central charge obtained by fitting the entanglement entropy from site 40-160 for a 200 sites chain. The blue open circles are for open boundary conditions and the red dots are for periodic boundary conditions.

We see that when we are tuned near, but not at, the critical point the best fits in the gapped case seem to come from smaller system sizes, and over ranges which do not include the flat middle portion of the scaling range nor the far tails of the scaling range. The unfortunate thing is that once we are a bit further away from the critical point this optimized fitting pattern no longer works. In this case none of the fitting regimes we used give accurate results because the system begins to reveal its gapped nature. We do find something close to $c = 1$ when $m = 1/100$ and $L = 500$ (Table. 3.3), but this seems accidental since we tested it for L=600 and got a worse results. From this data we would claim that for the Dirac model when the central charge differs by 20% from its expected value then we are too far away from the critical point to do any fitting and should claim that it is not critical. In fact for a system size of 500 and mass gap of $m = 1/1000$ the fitted values are closer to 6/5 instead of 1 and could easily lead to misidentification of critical points in models where their location is not known exactly.

As a possible diagnostic we plot the entanglement entropy as a function of $x = \ln(\frac{L}{\pi}\sin\frac{\pi l}{L})$, where l is the sub-block size. The slope of the entanglement entropy vs x should be interpreted as $c/6$. The only feature that could be used as a diagnostic is that if the transformed curve has a decrease in slope then we are definitely too far away from a critical point to fit properly as can be seen in Fig. 3.4(a). The final two



(a) (b)

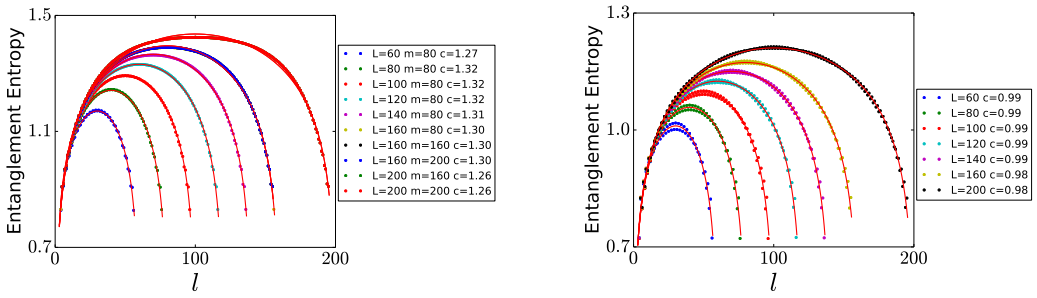
Figure 3.5: Panel (a) shows the entanglement entropy (for the central cut), as a function of f for $\theta = 1.00$. The EE increases for larger sizes of the system for f from 0.07 to 0.17, indicating a critical phase at this region. (b) shows the corresponding central charge calculated for various system sizes. The change of the central charge becomes sharper for larger systems.

curves have clear decreases in slope as we move far away from criticality. Note that all these artificially high central charges only occur when we use open boundary conditions. As can be seen in Fig. 3.4(b), the central charge first goes up then drops for open boundary conditions when we tune the system away from criticality. However, it decreases monotonically for periodic boundary conditions.

3.3.3 Kosterlitz-Thouless transition

Although most phase boundaries were easily identified, there are three regions where difficulties arise: (i) the trivial-incommensurate phase transition at $\phi = 0$ and large θ (lower-right corner of Figs. 3.2(a) and 3.2(d)), (ii) the topological-incommensurate phase transition at $\phi = \pi/3$ and small θ (upper-left corner of Figs. 3.2(b) and 3.2(e)), and (iii) the Lifshitz transition area for $f = 0.5$ and $\phi = \theta \sim \pi/6$ as seen in Figs. 3.2(c) and 3.2(f). Regions (i) and (ii) are related by duality, and the explanation of the numerical difficulties in these regions may have a common origin. To explain, we recall that the trivial-incommensurate phase transition at $\phi = 0$ and large θ, i.e. region (i), is of the Kosterlitz-Thouless type [48]. Hence, the correlation length decays as $\exp(c(T - T_{KT})^{-1/2})$ away from the transition point [98, 99], and this results in a long correlation length (compared to our system size $L = 100$) for this region of the phase diagram. The duality indicates that region (ii) may also be near a Kosterlitz-Thouless phase transition point. Thus, we attribute the issues with these regions as likely artifacts due to finite size effects. We elaborate further on this in this part. The remaining region (iii) requires more discussion, to which we turn in the next part.

We first discuss the features seen in Figs. 3.2(a) and 3.2(d), i.e., the cross section for $\phi = 0$. For small



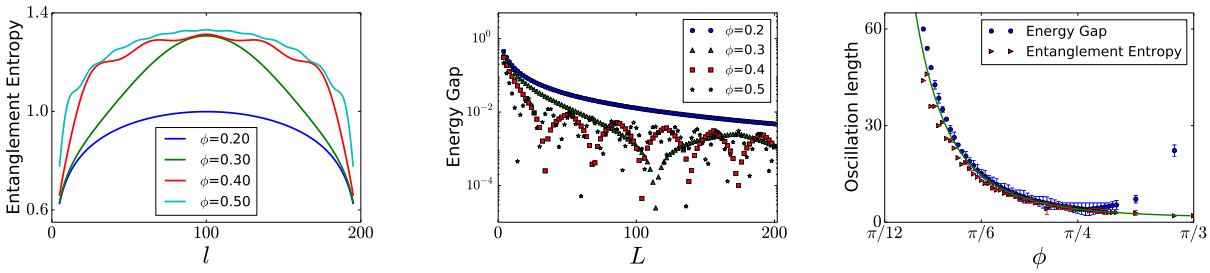
(a) (b)

Figure 3.6: The profile of the entanglement entropy as s function of block size l at $\phi = \pi/3$, $\theta = 0$ and (a) $f = 0.8$ (b) $f = 0.9$ for different system size. The continuous lines are the fit to the DMRG data.

f and large θ, the phase transition between the topological and trivial phase is indirect: it is mediated by the incommensurate phase. To establish the fact that the incommensurate region is of non-zero extent, we performed finite size analyses on both the entanglement entropy and central charge as is shown in Figs 3.5(a) and 3.5(b) This extent is found to be from $f \approx 0.07$ to $f \approx 0.15$. We find that the central charge of the trivial-incommensurate transition is consistent with that of the Kosterlitz-Thouless (KT) type, i.e., $c = 1$ [48].

Because of the duality in the Hamiltonian (Eq. 3.1), the phase diagram is symmetric with respect to the line $f = J = 0.5$, $\phi = \theta$. Thus, the above mentioned phase transition is dual to the incommensurate-topological phase transition, for large ϕ and small θ. That is to say, the region with the smooth change of the central charge in the lower-right corner of Fig. 3.2(d) is dual to the (red) region in the upper-left corner of Fig. 3.2(e). This region, being near the KT phase transition point is also plagued by finite size errors: the correlation length is long compared with the system size ($L = 100$).

To test this assertion, we studied the (apparently) large central charge that was calculated near the critical region, as is shown in Fig. 3.6. For example, as is shown in Fig. 3.6(a), the point $\phi = \pi/3$, $\theta = 0$, and $f = 0.8$ appears to be critical, but for larger system sizes is shown to be gapped. We base this conclusion on the appearance of a saturation plateau in the profile of the EE scaling as a function of subsystem size. As a comparative check, we went deeper into the critical regime (i.e. $f = 0.9$). As can be seen in Fig. 3.6(b) and as is expected, we found no such plateau in the EE.



(a) (b) (c)

Figure 3.7: (Color online) Properties of the critical line at $f = J = 1/2$ for various values of $\theta = \phi$ (a) Profile of the EE as a function of block size shows Lifshitz oscillations. We predict the oscillation length for $\phi < 0.2$ to be larger than our system size $L = 200$. (b) Energy gap between the ground and first excited state, which displays similar oscillatory behavior on varying the system size. (c) Characteristic oscillation lengths in the EE and energy gap, which are nearly identical for a large range of ϕ. The (green) line is the fit with $\zeta = \phi^{-3.75} + 1.16$.

3.4 Lifshitz behavior

3.4.1 Lifshitz transition in chiral clock model

Let us now focus on the cross-section in Figs. 3.2(c), 3.2(f), which corresponds to $\phi = \theta$. Since the system is self-dual on the line $f = J$, the trivial-topological phase boundary should just be the line $f = J = 0.5$, a fact verified in our numerical calculations when $\theta = \phi$ are small. On top of the phase diagram we also plot the function $f = [2\sin(3\phi)][1 + 2\sin(3\phi)]^{-1}$ (in a thick circular dotted line), which represents an upper bound on the region in which exact parafermionic zero modes are expected to exist as proven in Ref. [46]. The region of the phase diagram above this curve are guaranteed to not have exact parafermionic zero modes, despite still being in the topological phase with the topological ground state degeneracy. Along the critical line $f = J = 0.5$, $c = 4/5$ at the ferromagnetic point ($\phi = \theta = 0$), and $c = 1$ at the antiferromagnetic point ($\phi = \theta = \pi/3$) [91]. It is *a priori* unclear how the central charge transitions from $c = 4/5$ to $c = 1$, i.e., is it an abrupt jump at some transition point, or does it change incrementally in stages, or perhaps something else entirely? Only a few studies address this question directly: among them is the work of Howes et al. [49] who used fermion analyses and series expansions to conjecture that a tricritical point connecting the ordered (topological), disordered (trivial), and incommensurate phases exists at *exactly* $\phi = \theta = \pi/6$. McCoy et al. [93, 94] studied the super integrable line $\phi = \theta = \pi/6$ and suggested a modified picture with the incommensurate phase stretched all the way down to the point $\phi = \theta = 0$ and $f = J = 0.5$. Our results seem to support the latter picture, as we will further develop below.

To address the questions posed above, we studied the critical line $f = J$ carefully. We observed (see Fig 3.7(a)) that before we reach the putative tricritcal (Lifshitz) point at $\phi = \theta = \pi/6$, the EE starts

to show oscillatory behavior [3] . The frequency of the oscillations increases as we approach the Lifshitz point from small $\phi = \theta$, and when further increasing $\phi = \theta$ its amplitude dies out after the system clearly enters the incommensurate phase. Conventionally, a Lifshitz transition point of this nature corresponds to a continuously varying oscillation length, and in this case it is the length scale associated with the incommensurate order. Interestingly, the shapes of the EE oscillation curves match those observed recently in 1D free, and interacting, fermion systems near Lifshitz points where the Fermi surface is augmented by additional Fermi points [50]. Thus, our result adds to the evidence of Ref. [50] that these types of EE oscillations are a fingerprint of the Lifshitz-type phase transition. As an aside, we mention that the Lifshitz oscillations are only present in the EE when one uses *open* boundary conditions. One can easily check this by calculating the EE for free fermions as a function of next-nearest neighbor hopping[50], but with periodic boundary conditions.

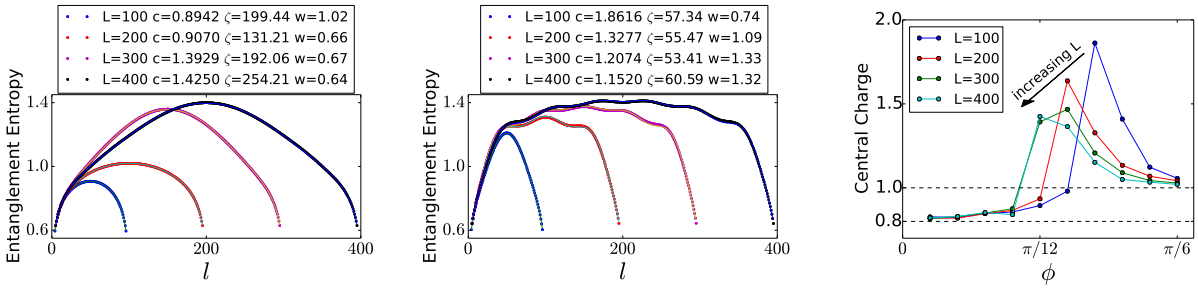
To quantitatively study the nature of this critical regime we want to investigate the variation of the central charge. However, in the presence of oscillations in the EE, we must modify Eq. 2.7 if we wish to extract the central charge. Empirically, the observed oscillations appear to have a similar form to those in the work Ref. [100], and we propose a phenomenological scaling form which can fit the EE with oscillations:

$$S(l)_{\text{cor}} = \frac{c}{6} \ln \left(\frac{L}{\pi} \sin \frac{\pi l}{L} \right) + S_0 + \frac{\cos(2\pi l/\zeta + p)}{(L/2 - |L/2 - l|)^w},$$

where the first two terms are the same as in Eq. 2.7, and the third term incorporates oscillations and a symmetrized damping function. The parameter ζ is the oscillation length and p is a phase factor. These parameters, along with the exponent w, are free-parameters determined by fitting. Some representative fits are shown in Figs. 3.8(a) and 3.8(b), which clearly capture the sub-leading oscillations accurately.

The results of calculating the central charge from this procedure are shown as a function of ϕ in Fig. 3.8(c). One can see that there is still an unaccounted for effect that leads to a peak in the central charge at a system-size dependent ϕ value. More careful inspection reveals that the peak is located at a ϕ^* that corresponds to an oscillation length $\zeta \approx L/2$. Thus, as seen in the figure, the peak location ϕ^*, occurs at values closer and closer to $\phi = \theta = 0$ when system size is increased, and all other parameters remain fixed. Our observations indicate that the central charge converges to $c \approx 1$ when $\phi \geq \phi^*$, and $c \approx 4/5$ for $\phi < \phi^*$. This strongly suggests that the transition from $c = 4/5$ to $c = 1$ along the line $f = J = 0.5$ is an abrupt one that occurs at $\phi = \theta = \phi^*$. From our numerical data it appears that $\phi^* \to 0$ as $L \to \infty$. Hence, our data supports a scenario where there is an immediate onset of oscillations as one tunes away from $\phi = \theta = 0$ in the thermodynamic

[3] Note, in Fig. 3.7(a) the EE curves in the incommensurate phase are not shown because they overlap with the curve at $\phi = \theta = 0.5$.



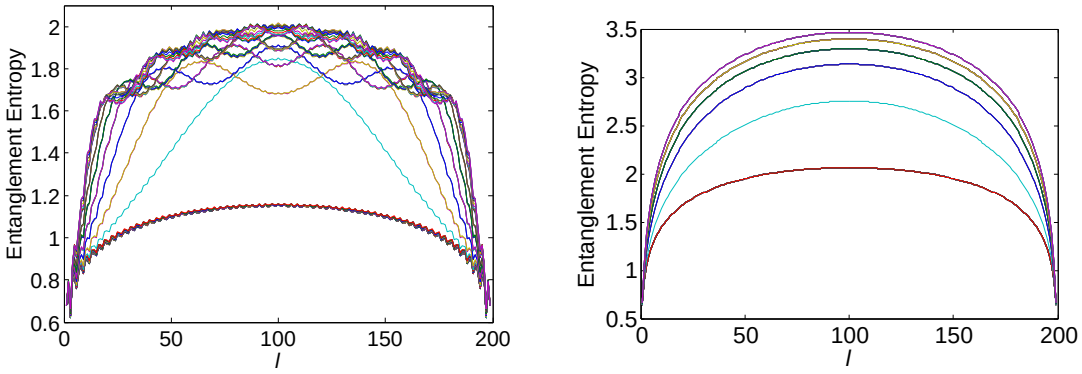
(a) (b) (c)

Figure 3.8: Panels (a) and (b) show the profile of the entanglement entropy (as a function of block size) for various values of system size at $\phi = \theta = 0.25$ and $\phi = \theta = 0.35$ respectively. Panel (c) shows the central charge obtained by fitting the entanglement entropy with the corrected formula along the line $\phi = \theta$ and $f = J = 0.5$. The two dashed lines are at $c = 0.8$ and $c = 1$. The arrow indicates the trend of the peak when L is increased.

limit.

We corroborate this by observing that oscillations are not seen in the EE if the oscillation length itself exceeds the system size L. For example, for $L = 200$, the oscillations are not explicitly visible for $\phi \lesssim \pi/12$, however upon increasing the system size, with all other parameters fixed, the oscillations appear over a larger region of ϕ, as is shown in Fig. 3.8(a). As ϕ is decreased the oscillation length increases, and thus we must use larger and larger systems to observe the oscillations. Thus, we believe that this is evidence that, in the thermodynamic limit, the oscillations are a feature for all $\theta = \phi$ except $\theta = \phi = 0$. An alternate scenario, which we can not rule out completely based on this numerical data, is that the incommensurate phase persists to small but non-zero values of $\theta = \phi$. Thus, a conservative estimate of the location of the tricritical point is $0 \leq (\theta = \phi) < 0.25$, which is well below the previously conjectured location at $\theta = \phi = \pi/6$. We aim to shed further light on this transition through larger scale simulations in future work.

Finally, we note that matching oscillations are observed in the splitting of the lowest two energy states (Fig. 3.7(b)), as a function of system size. We can extract the characteristic length scale ζ of the oscillations from both the EE (for a given system length), and the energy gap (as a function of system length). Our results are shown in Fig. 3.7(c) where a clear correlation between the two is observed for $\phi = \theta < \pi/4$. The solid (green) line in Fig. 3.7(c) is the fit of the oscillation length for $\phi = \theta < \pi/4$ to the function $\zeta = \phi^{-3.75} + 1.16$. When $\phi = \theta = 0$, the oscillation length appears to diverge, indicating that no such oscillations survive in the non-chiral 3-state Potts model limit. Attempts to relax the fit with $\zeta = (\phi - \phi^*)^{-\eta} + const$ (i.e. with a possibly non-zero ϕ^*) gave $\phi^* \sim 0.09$ indicating that the conjectured tricritical point may be in close proximity to $\phi^* = 0$.



(a) (b)

Figure 3.9: Entanglement entropy as a function of block size for different next nearest hopping t with (a) open and (b) periodic boundary condition. t is chosen from 0.97 to 1.03 with a step of 0.001. The lower part corresponds to $t \leq 1$. The cyan peak appears in the middle of panel (a) is of $t = 1.001$. Panel (b) shows the entanglement entropy increases as we increase t. Clearly, there are some steps for such increase.

3.4.2 Lifshitz transition in 1D free fermions

For comparison with our discussion of the Lifshitz transition in the chiral clock model we consider a version with 1D free fermions hopping on a chain with nearest neighbor and next nearest neighbor hopping. As the n.n.n. hopping is increased additional Fermi-points can enter the spectrum and eventually hit the chemical potential which leads to a Lifshitz transition of the Fermi-surface topology. As our model we consider free fermions with next nearest neighbor hopping.

$$H = -\sum_{n} \left[c^{\dagger}_{n+1} c_n + t c^{\dagger}_{n+2} c_n + h.c. \right] \tag{3.14}$$

Here, t is the parameter for the next nearest hopping. The energy spectrum of this model is $E = -2\cos(k) - 2t\cos(2k)$. When t increases from zero, the topology of the Fermi surface at zero energy changes from two points to four points at $t = 1$, which is the Lifshitz transition.

We calculate the entanglement entropy of this model with open boundaries and the periodic boundaries at half filling. The results are shown in Fig. 3.9(a) and Fig. 3.9(b) and one can immediately recognize the pattern of oscillations that we saw earlier for the chiral clock model. One interesting thing to notice is that the oscillations go away when we use periodic boundary conditions. This model, and the related entanglement properties, are carefully studied in Ref. [50]. For periodic boundary conditions the curves gradually increase from the scaling form with $c = 1$ to a scaling form of $c = 2$ which is the result expected for two sets of left and right movers at the Fermi-level.

3.5 Conclusions

In summary, we have mapped out the three dimensional phase diagram of the Z_3 *chiral* clock model using the density matrix renormalization group method. Using the entanglement entropy (of the half-chain) as a diagnostic, we have been able to locate the phase boundaries of the various topological-trivial-incommensurate phase transitions. Quantitatively, we have also been able to see the variation of the central charge along the various critical surfaces that divide these phases. Another outcome of this study is the identification of the Lifshitz transition using the entanglement entropy, along with an estimate of the location of the putative tricritical point. We discussed several competing qualitative scenarios for the cross section of the phase diagram in which the tricritical point has been predicted to exist. Our data suggests that the tricritical point (along $f = J = 1/2$) is not at $\phi = \theta = \pi/6$: rather we find it to be shifted to a much smaller value in the range $0 \leq \theta = \phi < 0.25$.

Finally, our results must be viewed in a broader context as providing further confirmation of the stability of the parafermionic topological phase to chiral interactions, over a wide range of parameters. We expect a further study of this and related models to elucidate the conditions under which these phases can be practically realized.

Chapter 4

Inversion symmetric topological superconductor with interactions

It is well known that interaction can reduce the Z classification of topological superconductors in one-dimensional BDI class to Z_8. If inversion symmetry is conserved, the classification can be further reduced to Z_2. The ground state of the topological phase in this interacting inversion symmetric superconductor is a condensation of four electrons instead of cooper-pairs. In this paper, we study the teleportation and fractional Josephson effects of this system. While there is a nonzero teleportation for cooper-pairs, the teleportation of one electron is suppressed. To restore the one electron teleportation, inversion symmetry must be broken. The inversion symmetry restricts the edge modes of the system to be cooper-pairs other than two uncorrelated electrons. It is also proved in the fractional Josephson effects. The 2π periodicity in the ground states indicating the tunneling of cooper-pairs.

4.1 Introduction

Topological states are of great interest due to it's potential application to fault tolerance quantum computation [101, 51, 56]. A full classification of topological phases for free fermions is obtained in all dimensions [12, 13]. The systems are classified into ten classes according to the time-reversal, charge-conjugation and chiral symmetry. Although only three symmetries are considered in the ten-fold classification, SPT phases can also be protected by other symmetries such as spatial symmetries [102, 103, 104, 105].

When interaction is considered, the classification can either be reduced or enhanced. For example, the classification of free fermions in class BDI in one dimension is reduced from Z to Z_8 [106]. The Z classification can be understood as n parallel Kitaev chains protected by time-reversal symmetry. No quadratic terms that preserve time-reversal symmetry can be added at the edge to gap out the edge state. However, a quartic interaction term that couples eight Majoranas can adiabatically connect eight parallel Kitaev chains with a topological trivial state [106].

On the other hand, if systems in class BDI have inversion symmetry, they are always trivial. Inversion symmetry changes ν to $-\nu$. Preserving inversion symmetry requires $\nu = -\nu$, which makes $\nu = 0$. However,

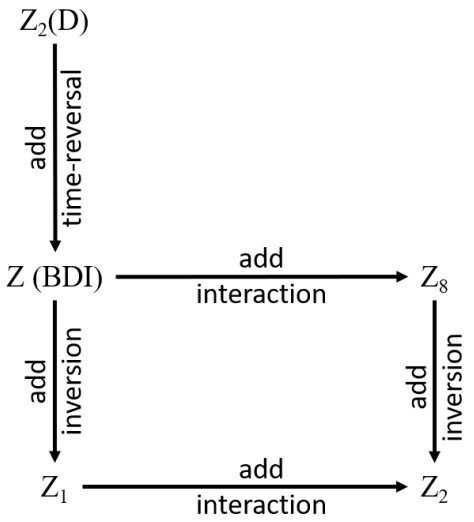


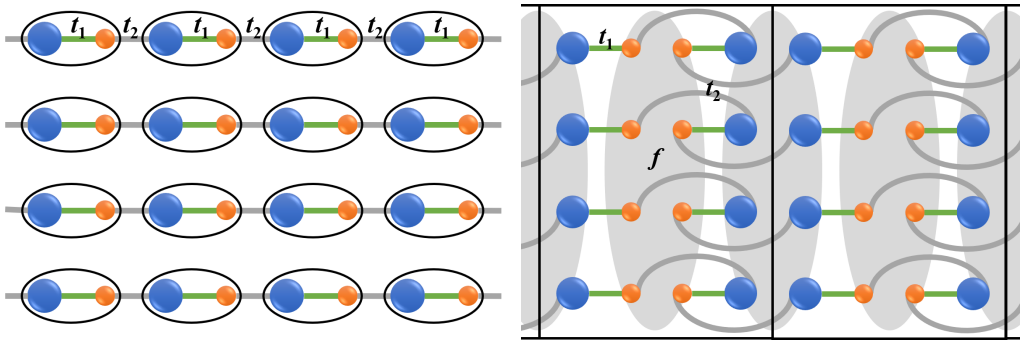
Figure 4.1: The schematic relationship between different systems and their classifications.

introducing interactions can enhance the classification to Z_2 [107]. This can be seen as follows. If interactions are allowed, the classification is reduced to Z_8 [106]. The topological phase with $\nu = 4$ can remain itself under inversion symmetry $\nu = 4 \to \nu = -4 \pmod 8 = 4$. A schematic relationship between these states is shown in Fig. 4.1.

It is now natural to ask, what is the difference between the time-reversal invariant SPT phase and the interaction enhanced SPT phase. In both cases, there are Majorana fermions at the ends of the chain. In the first case, the Majoranas are of different types while in the inversion symmetric case, the Majoranas are of the same type. It is interesting to know if we can distinguish these two types of Majoranas by some mesurable effects. In this paper, we will discuss the effect of inversion symmetry and interaction on this system through entanglement entropy and espectrum, teleportation and charge polarization. A phase transition between topological phases with and without inversion symmetry is discovered. The teleportation properties and fractional Josephson effects of the systems are also discussed in detail.

4.2 Construct the Hamiltonian

To construct the Hamiltonian, we first consider a system with four parallel Kitaev chains which does not have inversion symmetry (Fig. 4.2(a)). We denote t_1 and t_2 as intra-cell and inter-cell pairing terms for the Majorana fermions. The system is in a topological phase when $t_1 < t_2$. To impose inversion symmetry to this system, we enlarge the unit cell to include two sites for each chain (eight complex fermions or sixteen Majorana fermions in one unit cell). Then exchange the position of the two Majorana fermions in one chain



(a) (b)

Figure 4.2: (a) Four parallel Kitaev chains. We plot four unit cells with four fermions in one unit cell. The fermions are represented by the elliptical circles with two Majorana fermions inside each circle. The blue (big) dots represent Majorana fermions of type a and the red (small) dots are Majorana fermions of type b. The bars indicate the pairing of Majorana fermions with alternating strength t_1 and t_2. (b) The inversion symmetric 4 electron superconductor. We plot two unit cells. The oval shadow that connects eight Majorana fermions represents the intra-cell nearest-neighbor interaction f_1 between them.

for the right half of the unit cell. The pairing term t_2 connecting two neighboring complex fermions should be set to zero for a inversion symmetric system. However, turning off t_2 also makes the system transform from the topological phase to a trivial phase. Then we add interactions that couple the four chains (the ovals in Fig. 4.2(b)), so that it can transform back to a topological phase. We will talk about the phase diagram and the transition between the two topological phases in detail in next sections.

For convenience, we redefine complex fermions by combining Majorana fermions in the same column in Fig. 4.2(b) and consider the lattice with only four Majorana fermions (one column) on each lattice site. The two complex fermions combined from the upper/lower two Majorana fermions are labeled by spin up/down. Note that the complex fermion defined this way do satisfy the spin properties at the edge [107].

We use the eight-Majorana Fidkowski-Kitaev terms as the interaction in our system [106]. Written as complex fermions, the interaction includes terms between two sites j and k

$$W_{j,k} = -4(\hat{n}_{j\uparrow} + \hat{n}_{j\downarrow} - 1)(\hat{n}_{k\uparrow} + \hat{n}_{k\downarrow} - 1) + 8(\hat{c}_{j\uparrow}\hat{c}_{j\downarrow}\hat{c}_{k\uparrow}\hat{c}_{k\downarrow} + \hat{c}^{\dagger}_{j\uparrow}\hat{c}^{\dagger}_{j\downarrow}\hat{c}^{\dagger}_{k\uparrow}\hat{c}^{\dagger}_{k\downarrow}), \tag{4.1}$$

together with the on-site interaction

$$V_j = -(2\hat{n}_{j\uparrow} - 1)(2\hat{n}_{j\downarrow} - 1). \tag{4.2}$$

The full Hamiltonian we consider is then

$$\begin{aligned} H =& \sum_{j=1}^{N} f_0 V_j + \sum_{j=1}^{N-1} \left(f_1 W_{4j,4j+1} + t_2 P_{4j-1,4j+1} \right) \\ &+ \sum_{j=1}^{N} \left(f_1 W_{4j-2,4j-1} + t_1 P_{4j-3,4j-2} + t_1 P_{4j-1,4j} + t_2 P_{4j-2,4j} \right). \end{aligned} \tag{4.3}$$

Here N is the number of unit cells. $P_{j,k}$ includes the pairing terms between sites j and k

$$P_{j,k} = i\hat{c}^{\dagger}_{j,\uparrow}\hat{c}_{k,\uparrow} + i\hat{c}^{\dagger}_{j,\downarrow}\hat{c}_{k,\downarrow} + \text{h.c.} \tag{4.4}$$

4.2.1 Symmetries of the system

Inversion - The inversion symmetry $\mathcal{I}$ reverses the lattice. It takes the operator on site j to site $N-j+1$. All the terms in the Hamiltonian (eq. 4.3) preserves the inversion symmetry except for the next-nearest-neighbor hopping terms t_2. The t_2 hopping terms are always on odd/even sites for inter-cell/intra-cell hopping. After inversion the hopping terms on even and odd sites switches.

Fermion parity - We define the fermion parity for spin up and spin down fermions separately as $\mathcal{P}_{\uparrow}$, $\mathcal{P}_{\downarrow}$. It can be easily seen from eq. 4.3 that no term breaks either of the two fermion parities.

Time-reversal - The time reversal $\mathcal{T}$ preserves the physical fermion operators. For our redefined complex fermions, we have $\mathcal{T}\hat{c}_{j,\sigma}\mathcal{T}^{-1} = \pm\hat{c}^{\dagger}_{j,\sigma}$, and $\mathcal{T}\hat{c}^{\dagger}_{j,\sigma}\mathcal{T}^{-1} = \pm\hat{c}_{j,\sigma}$, since we pair the same type of Majorana fermions. The sign depends on the site i. Within one unit cell of four sites, we take "+" for site 1, 4 and take "−" for site 2 and 3. Note that if we define the states on sites 2 and 3 by acting on the vacuum state $|0\rangle$ with operators $i\hat{c}^{\dagger}_{j,\sigma}$, there will be no difference between different sites when acting $\mathcal{T}$ on the states. Then $\mathcal{T}$ is the same as charge conjugation $\mathcal{C}$.

Above are all the symmetries we consider in our system. To implement the symmetries in the numerical calculation, we cannot use the basis of occupation numbers because they respect non of those symmetries. For a two-site system, we list the basis we use in Table 4.1 (not normalized) and label them with the eigenvalues of the symmetries $\mathcal{P}_{\uparrow}$, $\mathcal{P}_{\downarrow}$ and $\mathcal{T}$.

4.2.2 Limiting cases

When interactions are turned off, the system is just four decoupled parallel Kitaev chains. A phase transition happens at $t_1 = t_2$. The topological phase is when $t_1 < t_2$. This can be understood by considering the limiting case $t_1 = 0$. The system has four free Majoranas at each end of the chain. Note that although the free

basis states	$\mathcal{P}_\uparrow$	$\mathcal{P}_\downarrow$	$\mathcal{T}$
$\|0\rangle\|0\rangle + \|\uparrow\downarrow\rangle\|\uparrow\downarrow\rangle$	+	+	+
$\|\uparrow\rangle\|\uparrow\rangle + \|\downarrow\rangle\|\downarrow\rangle$			
$\|0\rangle\|0\rangle - \|\uparrow\downarrow\rangle\|\uparrow\downarrow\rangle$	+	+	−
$\|\uparrow\rangle\|\uparrow\rangle - \|\downarrow\rangle\|\downarrow\rangle$			
$\|0\rangle\|\downarrow\rangle - \|\uparrow\downarrow\rangle\|\uparrow\rangle$	+	−	+
$\|\downarrow\rangle\|0\rangle - \|\uparrow\rangle\|\uparrow\downarrow\rangle$			
$\|0\rangle\|\downarrow\rangle + \|\uparrow\downarrow\rangle\|\uparrow\rangle$	+	−	−
$\|\downarrow\rangle\|0\rangle + \|\uparrow\rangle\|\uparrow\downarrow\rangle$			
$\|\uparrow\rangle\|0\rangle + \|\downarrow\rangle\|\uparrow\downarrow\rangle$	−	+	+
$\|0\rangle\|\uparrow\rangle + \|\uparrow\downarrow\rangle\|\downarrow\rangle$			
$\|\uparrow\rangle\|0\rangle - \|\downarrow\rangle\|\uparrow\downarrow\rangle$	−	+	−
$\|0\rangle\|\uparrow\rangle - \|\uparrow\downarrow\rangle\|\downarrow\rangle$			
$\|\uparrow\rangle\|\downarrow\rangle - \|\downarrow\rangle\|\uparrow\rangle$	−	−	+
$\|0\rangle\|\uparrow\downarrow\rangle - \|\uparrow\downarrow\rangle\|0\rangle$			
$\|\uparrow\rangle\|\downarrow\rangle + \|\downarrow\rangle\|\uparrow\rangle$	−	−	−
$\|0\rangle\|\uparrow\downarrow\rangle + \|\uparrow\downarrow\rangle\|0\rangle$			

Table 4.1: Basis states and the corresponding symmetry values for a two-site system.

Majoranas are in the edge unit cells, only on the left side they are on the edge site. When $t_2 = 0$, all the unit cells are decoupled, indicating a topological trivial phase when $t_1 > t_2$.

When the hopping terms are turned off, the system is discussed by Lapa et. al. [107]. The model can be mapped to an anti-ferromagnetic Heisenberg spin-1/2 chain when $f_0 = \infty$, which fixes the fermion parity to be even on each site. Though the bulk has time-reversal symmetry $T^2 = 1$, the edges have $T^2 = -1$. In this case, the free Majoranas are all on the edge sites.

4.3 Numerical results

From the limiting case analysis, we know that there are two topological phases; one is interacting and preserves inversion symmetry (II-phase) but the other is non-interacting and breaks inversion symmetry (NN-phase). There are a few questions that can hardly be answered analytically: 1) Is the II-phase stable with respect to general hopping terms such as t_1? 2) Is the NN-phase stable with respect to interactions? 3) are the two topological phases the same? If not, is there a phase transition between the two? Are there any physical observables that can be measured to distinguish the two?

In the following, we will address all these questions by solving the Hamiltonian (eq. 4.3) numerically. We adopt the density matrix renormalization group (DMRG) method for all the calculations.

4.3.1 Phase diagram

To characterize the different phases, we use the entanglement entropy and the entanglement spectrum at the central cut (cEE and cES) [35, 36]. They can be extracted naturally by implementing the DMRG calculation. The entanglement entropy can be naturally extracted from the DMRG calculation. To find the phase diagram, we use open boundary conditions and set the number of sites $L = 160$ ($N = 40$ unit cells) and bond-dimension $m = 40$.

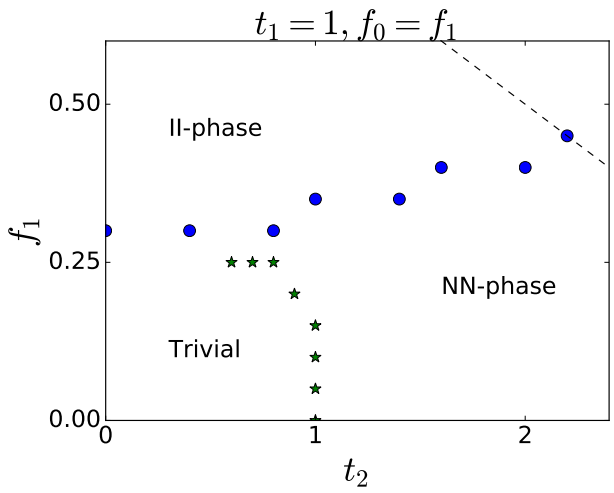
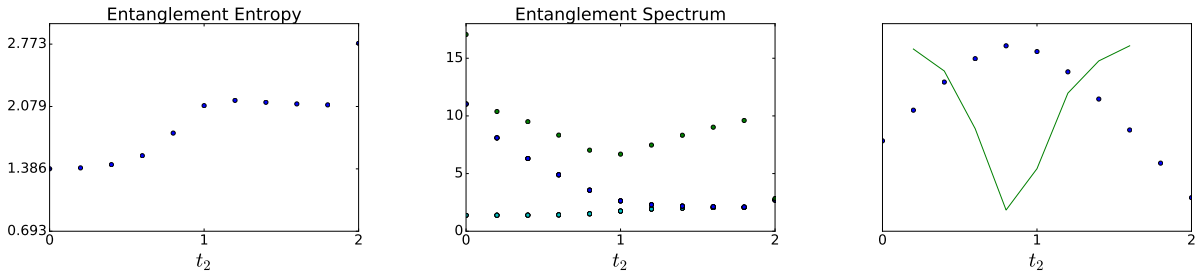


Figure 4.3: The phase diagram. We set $t_1 = 1$ and $f_0 = f_1$. The blue dots are obtained from the cES and the green stars are from cEE. The dashed line is $t_2 + 4f_1 = 4$.

In Fig. 5.12, we show the phase diagram obtained from the entanglement properties of the system. The lower left region is the trivial phase with zero cEE. The rest parts of the phase diagram have cEE $S \sim \ln 4$. The value of S deviates near the phase boundaries. We separate the region of zero and non-zero S by the green stars.

To find the phase boundary of the two topological phases, we have to look at the cEE and cES in detail. In Fig. 4.4, we plot the change of cEE, cES and the ground state energy by varying t_2 along the line $t_2 + 4f_1 = 4$. As can be seen in Fig. 4.4(a), the cEE diverges at $t_2 \sim 2.4$, indicating a phase transition at that point. The divergence in cEE corresponds to the split of the cES (Fig. 4.4(b)). In the II-phase, the cES is four-fold degenerate for all the entanglement energies. In the NN-phase, the lowest entanglement energy is singly degenerate, while there are large degeneracies in the higher energy levels. We indicate this kind of phase boundary by blue dots in the phase diagram in Fig. 5.12.

In addition, we plot the ground state energy in Fig. 4.4(c). The solid line is the second order derivative of the energy respect to f_1. From the peak of the derivative at point $f_1 = 0.4$, we reach the conclusion that the phase transition between the two topological phase is of second order.



(a) (b) (c)

Figure 4.4: (a) The entanglement entropy, (b) the entanglement spectrum and (c) the ground state energy along the line $t_2 + 4f_1 = 4t_1$. The solid line in (c) is the second order derivative of the energy. The discontinuous near $t_2 \sim 2.4$ in all these properties indicates there is a phase transition between the two topological phases.

4.3.2 teleportation

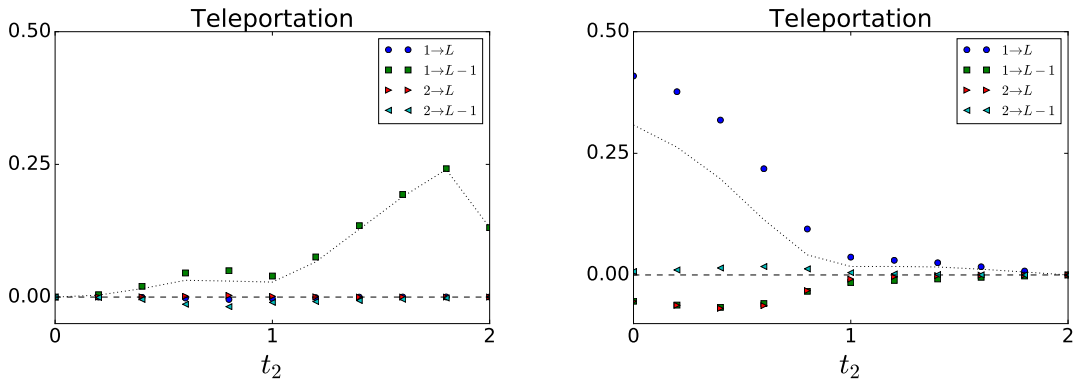
From the phase diagram, we can say with confident that both the topological phase are stable. However, even though the two phases are separated by a second order phase transition, it is still not clear whether they are truly different or not. In the following, we will show that the two topological phases are actually different. They have different teleportation properties.

The teleportation is measured from the static correlation function $C_{1,L} = \langle c^\dagger_{1\uparrow} c^\dagger_{1\downarrow} c_{L\uparrow} c_{L\downarrow} \rangle$. To get some intuition on this property, we first consider the case when the Majorana fermions at the edges are free. Then we have four free Majorana fermions at site 1 and the other four at site L. Since the Majorana fermions at one edge are of the same type, the only on-site term that preserves time-reversal and couples the four is the on-site interaction term V_j. This term locks the on-site fermion parity, so the teleportation of one fermion is always suppressed. This is because one fermion teleportation changes the on-site fermion parity of both edges, making the teleported state orthogonal to the original one.

On the other hand, the teleportation of cooper pairs is allowed when the fermion parity at the edges are even. Consider a two-site system consisting only the edge sites. The only states that have non-zero cooper-pair teleporataion are $\frac{1}{\sqrt{2}}(|0\rangle_1 |\uparrow\downarrow\rangle_L \pm |\uparrow\downarrow\rangle_1 |0\rangle_L)$. They have teleportation value $C_{1,L} = \mp 0.5$.

When the hopping terms are turned on, the local fermion parity is not fixed, so the single fermion teleportation is possible. However, from the numerical calculation, all types of the single fermion teleportation is always suppressed for all the parameter values.

We investigate the cooper-pair teleportation along the line $t_2 + 4f_1 = 4$ in the phase diagram. In Fig. 4.5, we plot the teleportation valuse between sites 1 and L, sites 1 and $L-1$, sites 2 and L, and sites 2 and $L-1$. The teleportation values are calculated using the ground state in the symmetry sector



(a) (b)

Figure 4.5: The teleportation value between the edge sites for ground state of symmetry (a) $\{\mathcal{P}_\uparrow, \mathcal{P}_\downarrow, \mathcal{T}\} = \{+,+,-\}$ and (b) $\{\mathcal{P}_\uparrow, \mathcal{P}_\downarrow, \mathcal{T}\} = \{-,-,+\}$

$\{\mathcal{P}_\uparrow, \mathcal{P}_\downarrow, \mathcal{T}\} = \{+,+,-\}$ (Fig. 4.5(a)) and $\{\mathcal{P}_\uparrow, \mathcal{P}_\downarrow, \mathcal{T}\} = \{-,-,+\}$ (Fig. 4.5(b)). For the symmetry sectors with opposite time-reversal value, the teleportatioin values are of opposite signs.

Note that all the teleportation values drops to zero at the transition point $t_2 \sim 2.4$. This point is consistent with the transition point obtained from the entanglement properties of the system.

4.3.3 Review of the periodicity of the Josephson effect for the Majorana/Kitaev chain

Consider first a normal Kitaev chain with a single species of complex fermion $\psi_j = \frac{1}{2}(a_j + ib_j)$ located at site j. Suppose the Hamiltonian contains $2e$ pairing terms of the form $\Delta\psi_j\psi_{j+1}$+h.c., where the superconducting order parameter $\Delta = \Delta_0 e^{-i\theta}$ with Δ_0 real and positive. Here we will slightly formalize Kitaev's explanation for why the ground state of the system in the topological phase is only periodic in θ with period 4π, and not 2π.

The phase θ can be removed from the Hamiltonian by performing a unitary transformation on the Hamiltonian which will add a compensating phase to the fermions ψ_j. To do this we need to find an operator U such that $U\psi_j U^\dagger = e^{i\frac{\theta}{2}}\psi_j$. Since the two types of Majorana's are given in terms of ψ_j by the formulas

$$a_j = \psi_j + \psi_j^\dagger \tag{4.5}$$

$$b_j = -i(\psi_j - \psi_j^\dagger)\ , \tag{4.6}$$

we are looking for a transformation which acts as

$$Ua_jU^\dagger = \cos(\tfrac{\theta}{2})a_j - \sin(\tfrac{\theta}{2})b_j \tag{4.7a}$$

$$Ub_jU^\dagger = \sin(\tfrac{\theta}{2})a_j + \cos(\tfrac{\theta}{2})b_j \ . \tag{4.7b}$$

The Hamiltonian H' with the phase θ removed is then given by $H' = UHU^\dagger$.

The operator which can do this transformation for the Majorana's a_j, b_j is $U_j = e^{-i\frac{\theta}{4}(ia_jb_j)}$, where $ia_jb_j = 2\psi_j^\dagger\psi_j - 1$ is just the local fermion parity operator for site j. Therefore the operator U for the whole chain is given by

$$\begin{aligned} U &= \prod_{j=1}^{N} e^{-i\frac{\theta}{4}(ia_jb_j)} \\ &= e^{-i\frac{\theta}{4}\sum_{j=1}^{N}(ia_jb_j)} \\ &= e^{i\frac{N\theta}{4}} e^{-i\frac{\theta}{2}F} \ , \end{aligned} \tag{4.8}$$

where $F = \sum_{j=1}^{N} \psi_j^\dagger\psi_j$ is the fermion number operator, not to be confused with N which is the length of the chain.

Now we can see that when $\theta \to \theta + 2\pi$, the Hamiltonian returns to itself, but at $\theta = 2\pi$ the operator U becomes

$$U(\theta = 2\pi) = e^{i\frac{N\pi}{2}}(-1)^F \ , \tag{4.9}$$

so it is proportional to the total fermion parity operator $(-1)^F$. In the ground state subspace of a chain in the topological phase (consisting of two states), U acts as the matrix $U(\theta = 2\pi) = e^{i\frac{N\pi}{2}}\sigma^z$, which is not the identity. This explains why the system is only periodic in θ with period 4π, and not 2π.

4.3.4 Josephson effect for the Fidkowski-Kitaev chain

Although the FK features $4e$ pairing as well as $2e$, for now let's continue to define the phase as in the $2e$ case, so we still consider unitary transformations that act on the Majorana's as in Eq. (4.7) above. The FK chain has eight complex fermions $\psi_{j,J} = \frac{1}{2}(a_{j,J} + ib_{j,J})$ in each unit cell, so the correct transformation that rotates the phases of all fermions is now

$$\begin{aligned} U &= \prod_{j=1}^{N}\prod_{J=1}^{8} e^{-i\frac{\theta}{4}(ia_{j,J}b_{j,J})} \\ &= e^{i2N\theta} e^{-i\frac{\theta}{2}F} \ , \end{aligned} \tag{4.10}$$

where now $F = \sum_{j=1}^{N} \sum_{J=1}^{8} \psi_{j,J}^{\dagger} \psi_{j,J}$ is the fermion number operator. For $\theta = 2\pi$ we have $U(\theta = 2\pi) = (-1)^F$.

For the FK chain with open boundary conditions, we can add a time-reversal symmetric quartic perturbation for the four Majorana's at each end of the chain, as long as the perturbations on both ends are identical in order to preserve the inversion symmetry. In this case the ground state subspace of the chain consists of four states, and all four states have the same fermion parity (the parity of these four states is odd or even depending on the sign of the quartic perturbation added at the ends of the chain). So we find that $U(\theta = 2\pi) = \pm \mathbb{I}_4$ (where $\mathbb{I}_4$ is the 4×4 identity matrix) in the ground state subspace of the FK chain. This evidence suggests that in terms of the $2e$ pairing phase θ, an ordinary Kitaev chain has a 4π periodic Josephson effect, while the FK chain has a 2π periodic Josephson effect.

4.3.5 Suppression of 4π Josephson effect for non-interacting TSC's due to interactions

In this section we consider the Josephson effect (JE) for a set of four ordinary Kitaev/Majorana chains. We consider two half-infinite chains with different superconducting phases which meet at the origin $n = 0$. At the origin we consider single fermion tunneling between the two chains (which leads to the 4π JE), but we also turn on quartic interactions at the ends of each chain.

The Majorana chain in the region $n \geq 0$ (the right half of the space) consists of the four flavors of complex fermion $\psi_{n,J}^r$, where $J = 1, 2, 3, 4$ and the superscript r stands for "right". The $2e$ superconducting phase for this chain is θ_r, so the complex fermions are written in terms of Majorana fermions as

$$\psi_{n,J}^r = \frac{e^{i\frac{\theta_r}{2}}}{2}(a_{n,J}^r + ib_{n,J}^r) \ . \tag{4.11}$$

The Majorana chain in the region $n \leq 0$ (the left half) consists of the complex fermions $\psi_{n,J}^l$. In this region the superconducting phase is θ_l and the complex fermions are written in terms of Majorana fermions as

$$\psi_{n,J}^l = \frac{e^{i\frac{\theta_l}{2}}}{2}(a_{n,J}^l + ib_{n,J}^l) \ . \tag{4.12}$$

Now we assume that each of the chains is in the topological superconducting phase, so that at the origin $n = 0$ we have the four unpaired Majorana modes $a_{0,J}^r$ coming from the right chain, and the four unpaired Majorana modes $b_{0,J}^l$ coming from the left chain. We assume single fermion (number-conserving) tunneling

between the two chains at the origin, of the form

$$\begin{aligned} V &= \sum_{I,J} t_{IJ}(\psi^{r,\dagger}_{0,I}\psi^{l}_{0,J} + \text{h.c}) \\ &\approx i\sum_{I,J} \tilde{t}_{IJ} a^{r}_{0,I} b^{l}_{0,J} \ , \end{aligned} \tag{4.13}$$

where $\tilde{t}_{IJ} = t_{IJ}\cos\left(\frac{\theta_r - \theta_l}{2}\right)$ and we have approximated this term by ignoring all Majorana's which are paired in the bulk of the superconductors. From now on I will drop the "0" index (the position index) on the Majorana's $a^r_{0,J}$ and $b^l_{0,J}$, since these are the only Majorana's that we will be dealing with. The tunneling term between the chains takes the form

$$V = i\sum_{I,J} \tilde{t}_{IJ} a^r_I b^l_J \ . \tag{4.14}$$

We now study the tunneling term V when strong interactions are turned on for the unpaired Majorana's at the junction. So we consider a Hamiltonian for the junction of the form

$$H = H_0 + V \ , \tag{4.15}$$

where

$$H_0 = -u(P^l + P^r) \tag{4.16}$$

with $P^l = b^l_1 b^l_2 b^l_3 b^l_4$ and $P^l = a^r_1 a^r_2 a^r_3 a^r_4$. So H_0 is the interaction which fixes the local fermion parity. We assume that H_0 is the dominant interaction and treat V as a perturbation.

We now investigate the effects of V in the subspace of the total Hilbert space at the junction consisting of the states $|\psi\rangle$ with even local fermion parity, i.e., $P^l|\psi\rangle = P^r|\psi\rangle = |\psi\rangle$. To do this we perform a Schrieffer-Wolf transformation on the Hamiltonian H, following the formalism in arXiv:1105.0675. First define the projector $\mathcal{P}_0$ onto the subspace of even local fermion parity:

$$\mathcal{P}_0 = \left(\frac{P^r + 1}{2}\right)\left(\frac{P^l + 1}{2}\right) \ . \tag{4.17}$$

We also define $\mathcal{Q}_0 = \mathbb{I} - \mathcal{P}_0$. Then we split the perturbation into diagonal and off-diagonal pieces $V = V_d + V_{od}$, where

$$\begin{aligned} V_d &= \mathcal{P}_0 V \mathcal{P}_0 + \mathcal{Q}_0 V \mathcal{Q}_0 \\ &= V + 2\mathcal{P}_0 V \mathcal{P}_0 - \{\mathcal{P}_0, V\} \ , \end{aligned} \tag{4.18}$$

and

$$\begin{aligned} V_{od} &= \mathcal{P}_0 V \mathcal{Q}_0 + \mathcal{Q}_0 V \mathcal{P}_0 \\ &= -2\mathcal{P}_0 V \mathcal{P}_0 + \{\mathcal{P}_0, V\} \ . \end{aligned} \tag{4.19}$$

In our case one can show that $\mathcal{P}_0 V \mathcal{P}_0 = 0$ (it must be the case because V is a hopping term and so it negates the local parity when it acts).

Now we perform a unitary transformation to obtain a new Hamiltonian $H' = UHU^\dagger$ which has the off-diagonal piece of the perturbation, V_{od}, removed. We write $U = e^S$ where $S^\dagger = -S$. Then expanding out $e^S H e^{-S}$ shows that to first order we must choose S such that

$$[H_0, S] = V_{od} \ . \tag{4.20}$$

The corrected Hamiltonian H' to lowest order is then

$$H' = H_0 + V_d + [S, V] \ . \tag{4.21}$$

Finally we project H' into the space of states with even local fermion parity to see how the perturbation acts in that sector (which is the low energy sector), i.e., the final Hamiltonian operating in the low-energy sector is

$$H'' = \mathcal{P}_0 H' \mathcal{P}_0 \ . \tag{4.22}$$

This is equivalent to just taking the matrix elements $\langle\psi|H'|\phi\rangle$ of H' in states $|\psi\rangle, |\phi\rangle$ which are in the sector of even local fermion parity $P^r = P^l = 1$.

In our case the solution to Eq. (4.20) takes the form

$$S = \alpha V P^r + \beta P^l V \ , \tag{4.23}$$

where α and β must satisfy $\alpha - \beta = \frac{1}{4u}$. Using the formula $[AB, C] = A\{B, C\} - \{A, C\}B$ and the fact that $\{V, P^r\} = \{V, P^l\} = 0$, we find that the correction to the Hamiltonian is

$$[S, V] = -2\alpha V^2 P^r + 2\beta P^l V^2 \ . \tag{4.24}$$

We can compute

$$V^2 = \sum_{I,J,K,L} \tilde{t}_{IJ}\tilde{t}_{KL} a_I^r a_K^r b_J^l b_L^l \ . \tag{4.25}$$

Then we can see that for $K \neq I$ and $L \neq J$, multiplication of this by P^r or P^l will still give a quartic interaction. Now suppose there is a term in which $K = I$. Naively this gives a quadratic term of the form

$$\sum_{I,J,L} \tilde{t}_{IJ}\tilde{t}_{IL} b_J^l b_L^l \ , \tag{4.26}$$

however, this term sums to zero since $\tilde{t}_{IJ}\tilde{t}_{IL}$ is even under $J \leftrightarrow L$ while $b_J^l b_L^l$ is odd under $J \leftrightarrow L$. A similar argument shows that any term with $J = L$ also vanishes. Finally, any term with $I = K$ and $J = L$ is just a constant.

The main point is that since $\tilde{t}_{IJ} \propto \cos\left(\frac{\theta_r - \theta_l}{2}\right)$, the induced interaction in the low-energy sector is proportional to $\cos^2\left(\frac{\theta_r - \theta_l}{2}\right) = \frac{1}{2} + \frac{1}{2}\cos(\theta_r - \theta_l)$ which means that the periodicity of the Josephson effect is reduced back down to 2π in the low energy sector. The locking of the local fermion parity by the strong interaction H_0 has suppressed single fermion tunneling, and only Cooper pair tunneling survives to contribute to the Josephson effect.

Finally, there is one peculiar property of the induced interaction term. Since we only need $\alpha - \beta = \frac{1}{4u}$ to satisfy Eq. (4.20), there are many possible choices. For example we could choose $\alpha = \frac{1}{4u}$, or $\alpha = -\beta = \frac{1}{8u}$. The second choice seems to be the most symmetric, and it leads to an induced interaction of the form

$$[S, V] = -\sum_{I,J,K,L} \frac{\tilde{t}_{IJ}\tilde{t}_{KL}}{4u} \left[(a_I^r a_K^r b_J^l b_L^l)P^r + P^l(a_I^r a_K^r b_J^l b_L^l)\right] \ . \tag{4.27}$$

In any case, we always get a quartic interaction with coefficient $\propto \frac{t^2}{u}$. The resolution of this puzzle is that after projection into the space of even local parity $P^r = P^l = 1$ all choices of α and β give the same interaction.

4.4 Conclusions

In summary, we have constructed a one dimensional topological superconductor with interactions. The time-reversal, particle-hole, chiral, and inversion symmetries are preserved. The ground state of the topological phase in this interacting inversion symmetric superconductor is a condensation of four electrons instead of cooper-pairs. The teleportation and fractional Josephson effects of this system are calculated. While there is a nonzero teleportation for cooper-pairs, the teleportation of one electron is suppressed by the inversion

symmetry. This is because the edge modes of the system must be cooper-pairs other than two uncorrelated electrons with the presence of inversion symmetry. It is also verified by the fractional Josephson effects. The 2π periodicity in the ground states indicating the tunneling of cooper-pairs.

Chapter 5

Machine learning symmetry protected topological phases

We apply random forest as a machine learning model to classify topological phases when strong disorder is present. We show that using entanglement spectrum as features for the training, the model gives high accuracy of classification. This model can be applied to other regions in the phase space and even to other systems. A detailed analysis of the model indicates that it captures the degeneracy in the entanglement spectrum.

5.1 Introduction

One of the most important problems in condensed matter physics is phase classification. In the classical Landau theories, phases are characterized by local order parameters associated with symmetries of the system [1, 108]. This unified criteria breaks down when topological order was discovered in quantum many-body systems [2, 3]. Many methods have been developed to classify topological phases, such as string order parameters, entanglement properties, topological indexes and etc. [109, 12, 110, 36]. Although a classification table was later proposed for free fermion systems in different symmetry classes [12, 13], explicit confirmations are still needed at strong disorder for each symmetry class. Many successful attempts have been made with symmetry class A, AIII, and BDI using entanglement properties, level statistics analysis, and real-space topological index [111, 112, 113, 114, 115].

With the development of computer technologies, machine learning now provides a new framework for solving problems in physics. Promising developments have been achieved in applying machine learning techniques to condensed matter physics. Supervised learnings have been used directly in characterizing phases in both classical spin systems [31, 116, 117, 118, 119] and quantum many-body systems [120, 121, 122, 123, 124, 125, 126]. Specifically, neural networks are the most widely used model to identity phases especially topological phases. Chen insulators and fractional Chen insulators can be classified by feeding quantum loop topography into the neural network [123]. On the other hand, entanglement spectrum was used as features for training to locate phase transition points [124]. Neural networks are powerful models that

have universal approximation capabilities [127, 128]. Their black-box attribute makes it hard to interpret the trained models, therefore not so insightful in physics applications. Besides, the large number of hyper-parameters in a neural network makes it difficult to train.

In this paper, we use random forest (RF) as our machine learning model instead of neural networks to detect topological phases with strong disorder. Random forest is an ensemble method that is capable of representing complicated functions with much fewer parameters compared with neural networks, but with better interpretation abilities [129, 130]. It is a collection of decision trees, which can be understood as piece wise constant functions in the feature space. An individual decision tree cannot make good predictions because in general predictions of decision trees have large variance. Averaging over decision trees reduces variance, making random forest a popular method in machine learning community. One major advantage of random forest is that it has few hyper-parameters. Besides, It is immune to problems such as over-fitting, collinearity, etc.

We propose the use of the entanglement spectrum (ES) as our input data. The ES has been widely used to characterize topological phases. A robust degeneracy in entanglement spectrum is an indicator for topological phase in general [35, 36]. We calculated the ES of a disordered chiral Hamiltonian in symmetry class AIII. The RF model trained by the data generated from a small fraction of phase diagram can be generalized to the full phase space with high prediction accuracy. Further more, the trained RF model shows high prediction ability in Kitaev chain [45], which is in symmetry class BDI. A detailed analysis reveals that the RF model is capturing the degeneracy in the ES.

5.2 Machine learning methods

In general, a machine learning problem can be described mathematically as follows. Use X as features of the training data and Y as responses. X and Y are matrices that have the same number rows. Each row of X is an observation, and each column is a feature. In many cases Y is just a column vector, i.e. one value for each observation, but we can also have Y as matrices. For example, in a classification problem the response of one observation is a vector. Each element of this vector indicates the probability of being in the corresponding class. The goal of machine learning is to find a function $f(X)$ that can best represent Y. This is the *supervised learning*. For the *unsupervised learning*, no response Y is provided. People use supervised learning for regression and classification, but unsupervised learning can only be used to do classification.

The training process is to find the best f from a family of functions. It is crucial to choose an appropriate function family. First, we introduce two concepts: training error and test error. We usually use the loss

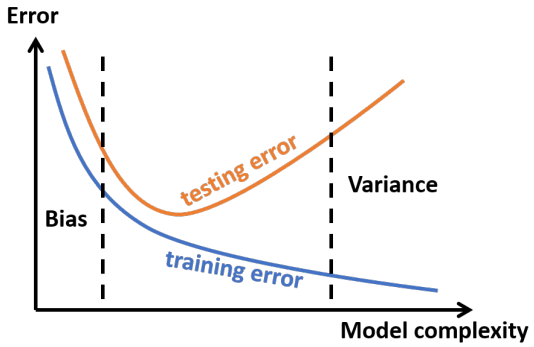


Figure 5.1: Bias variance trade-off. When the model is simple, we are in the high bias region (left) with high training and testing error. As the model complexity increases, both error decrease. If we make the model more complex, we reach the high variance region (right) with low training error but high testing error.

function $L(Y, f(X))$ as a measurement of the error. Divide the data set into two parts: training and testing. Only the training data is used for the training process, while testing data is completely hidden during the training. After the training, the training and testing error are calculated for the training and testing data respectively.

If the functions are simple, such as linear functions, we are posing a lot of restrictions to the model. In this case, we are biased and the model usually has high training and testing error. If the functions are too complex, the model tends to over-fit. This corresponds to the high variance region. In this area the training error usually is small, because we find the We show the relations between model complexity and error in Fig. 5.1. This is known as the bias-variance trade-off in the machine learning community. A good model should have low error for both training and testing data.

5.2.1 Linear models

Simple linear regression

We start from the simplest machine learning method: linear regression. Suppose the response Y is linearly related with the features X. Define function

$$f(X) = \beta_0 + X\beta. \tag{5.1}$$

Here the intercept β_0 is a number. The coefficients β is a column vector. We want to find the β_0 and β that minimize the loss function

$$L(Y, f(X)) = ||Y - f(X)||^2, \tag{5.2}$$

where $||\cdot||$ is the $L-2$ norm. This is also known as least square.

Although linear regression is the simplest model, it can still over fit when there are a lot of features. We can reduce the complexity of the linear regression by adding regularization terms. The most popular regularization terms are $L-2$ and $L-1$ norms, corresponding to ridge regression and LASSO. For the ridge regression, the $L-2$ norm of the coefficients are added to the loss function Eq. 5.2, posing restrictions on large amplitudes.

$$L_{\text{ridge}} = ||Y - (\beta_0 + X\beta)||^2 + \lambda||\beta||^2. \tag{5.3}$$

Here, λ is a positive number. It is a hyper-parameter to be tuned. Larger λ means higher regularization and simpler model.

Logistic regression

We can use linear models for classification problems. However, directly applying linear regression with Y as indicators makes terrible predictions. This is due to the fact that the Y values are discrete while linear regression makes continuous predictions. What we need to do is to use a step-like function to do the transformation. The most commonly used function is the sigmoid function $\sigma(x) = 1/(1+e^{-x})$.

For a binary classification problem, suppose the probability of being in one class is

$$f(X) = \frac{1}{1+e^{-(\beta_0+X\beta)}}. \tag{5.4}$$

We used the same notation as for the linear regression. Then Y is either one or zero, because we know for sure whether the observation is in this class or not. We do not use the same loss function Eq. 5.2, as it is not convex in this case. Instead we use the cross entropy as the loss function.

$$L(Y, f(X)) = -\frac{1}{n}\sum_{j=1}^{n}\left[y_j \log(f(x_j)) + (1-y_j)\log(1-f(x_j))\right], \tag{5.5}$$

where x_j and y_j are rows of X and Y respectively. Same as linear regression, we can add regularization terms to the loss function to poss additional constraints on the coefficients.

For a multi-classification problem with K classes, we use the softmax function as the probability for class k:

$$P(y=k|X) = \frac{e^{b_k+XW_k}}{\sum_{j=1}^{K} e^{b_j+XW_j}}. \tag{5.6}$$

Here, the number b_j is the intercept for class j and the column vector W_k is the coefficient. The corresponding

loss function is then

$$L(Y, f(X)) = -\frac{1}{n}\sum_{j=1}^{n}\sum_{k} p_k \log[P(y = k|X)]. \tag{5.7}$$

We used p_k as the true probability distribution. It takes value one if the observation is in class k, otherwise zero. When $K = 2$, it reduces to the binary classification problem.

5.2.2 Neural networks

Neural networks can be understood as generalizations of linear and logistic regressions. The input information is processed step by step within each layer until the last one. The features generated in the last layer are put into a linear or logistic regression to give the final result. Different neural networks have their unique processing schemes. We will introduce two neural networks in this part: the feed forward and the convolutional neural networks. We will not go into the details too much as it is time consuming to develop a neural network from scratch. Besides, there are a lot of well developed packages ready for us to use, such as Keras, Tensorflow, Caffee, Pytorch, and etc.

Feed forward neural networks

The feed forward neural network is the most basic neural network. A neural network is often represented by a figure with each neuron as a feature. As shown in Fig. 5.2, a feed-forward neural network is composed by a bunch of layers. Each layer contains several features or neurons represented by circles in the figure. Layers are calculated in order from left to right. The left most layer is the input layer and the right most layer is the output layer. All the other layers in between are the hidden layers. Neurons within the same layer are not connected. Mathematically, we denote each layer as X_0, X_1, ..., X_n. For a single observation, they are row vectors with each element as one neuron. In general, for an input data with n observations, they are matrices with n rows and columns as features. The input layer $X_0 = X$ is known; and the output layer $X_n = \hat{Y}$ is the value our model predicts. Two neighboring layers are connected by

$$X_{i+1} = f(X_i) = \sigma(b_i + X_i W_i), \tag{5.8}$$

where $\sigma(x) = 1/(1 + e^{-x})$ is the sigmoid function as before.

All the non-linearity comes from the sigmoid function. We can use other functions instead of the sigmoid function, such as the hyperbolic tangent and the rectified linear unit (ReLU). ReLU is defined as the positive part of the argument $\text{ReLU}(x) = \max(0, x)$. All these functions are called activation functions. Among them the most widely used is the ReLU. Although it is not smooth at zero, it over performances others in most

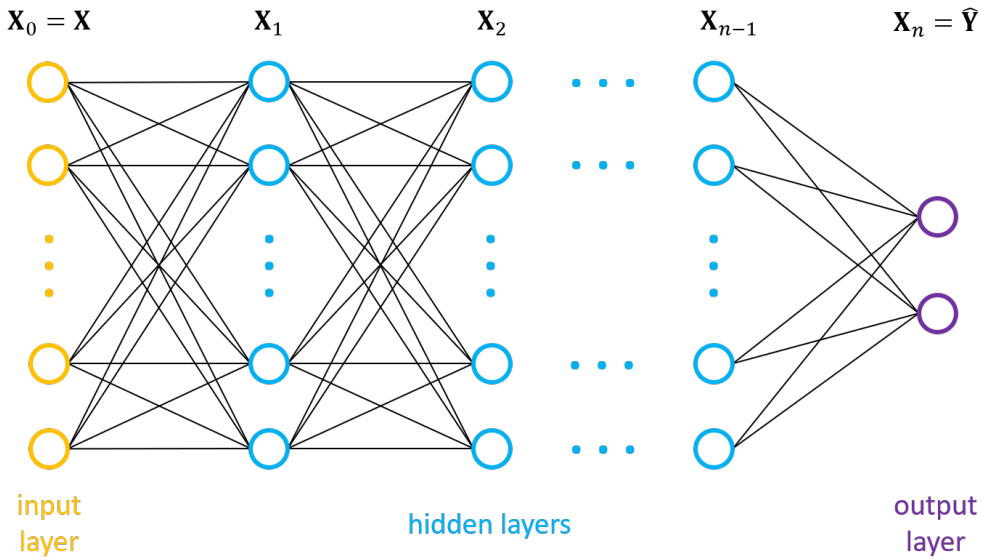


Figure 5.2: A feed forward neural network with $n+1$ layers. Each circle is a neuron, representing one feature. The first layer (yellow) is the input layer and the last one (purple) is the output layer. Other layers (blue) are hidden layers.

problems.

The choice of loss functions is the same as linear cases. We use mean square error for regressions and cross entropy for classifications in general. A regularization term like $\lambda \sum_i ||W_i||^2$ can be added to the loss function to prevent over fitting as well. Moreover, the technique of drop-off is used more often as a regularization method. We mask neurons in each layer with some probability α, i.e. independently set values of neurons as zero with probability α. This prevents neurons to collaborate with other neurons.

In linear regression the coefficients can be solved directly from the analytic formula. In logistic regression the loss function is convex, so we can always use the gradient descent to find the best parameters. However, the feed forward neural networks have no analytic solutions nor convex loss functions. Many numerical methods as generalizations of the gradient descent have been developed to find the minimum of the loss function. Among them, Adam is the most popular. It is a kind of stochastic gradient descent combined with the momentum and self-adjusting learning rate. More details can be found in the paper proposing Adam [131].

Convolutional neural networks

The convolutional neural network (CNN) is known for its ability of image processing. One distinguishing feature of the CNN is it takes into account the spatial structure of data. In addition to the ordinary feed-forward layers, the CNN has convolution layers, rectifier layers and pooling layers. We list the mathematical

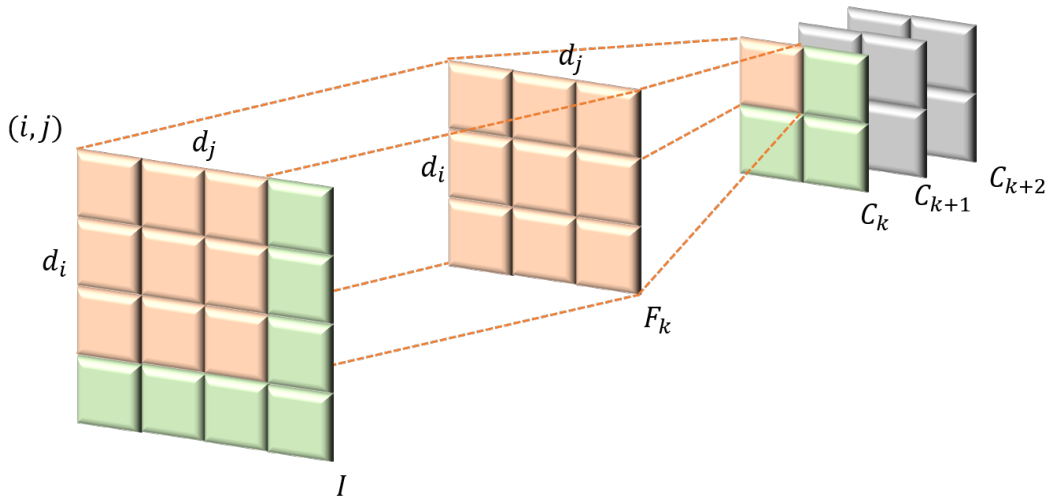


Figure 5.3: Convolution of input image I with filter image F_k. The output value is one element of image C_k. With multiple filters, we can get different output images.

operations for these layers:

$$\begin{aligned} \text{convolution} \quad & C_{ijk} = \sum_{d_i,d_j,q} I_{i+d_i,j+d_j,q} F_{d_i,d_j,q,k} \\ \text{rectifier} \quad & R = \max\{C, 0\} \\ \text{pooling} \quad & O_{ij} = \max\{R_{i:i+s_i,j:j+s_j}\}. \end{aligned}$$

Here, I is input image with q channels. F is filter images with size (d_i, d_j). q is index for input channels and k for output channels.

As shown in Fig. 5.2.2, the convolution layer takes one image I as an input, convolves with filter images F_k, and outputs a stack of filtered images C_k. If the input images have multiple channels, the convolved results are summed up for all input channels. The rectifier layer sets negative values to zero, as images always have non-negative values at each pixel. This is just the ReLU activation function. The pooling layer takes the maximum value among an window of the filtered image stack. One pixel of an image is considered as a feature. While convolution in general creates layers with more features, pooling reduces the feature number. Convolution and pooling together extract information from the image and transform them as abstract numbers, so that the following layers can use them to make predictions.

5.2.3 Random forest

Random forest is an ensemble method. This means the prediction is given as the averaged results of the individual regressors or classifiers. Random forest averages over randomized decision trees. We first introduce decision trees.

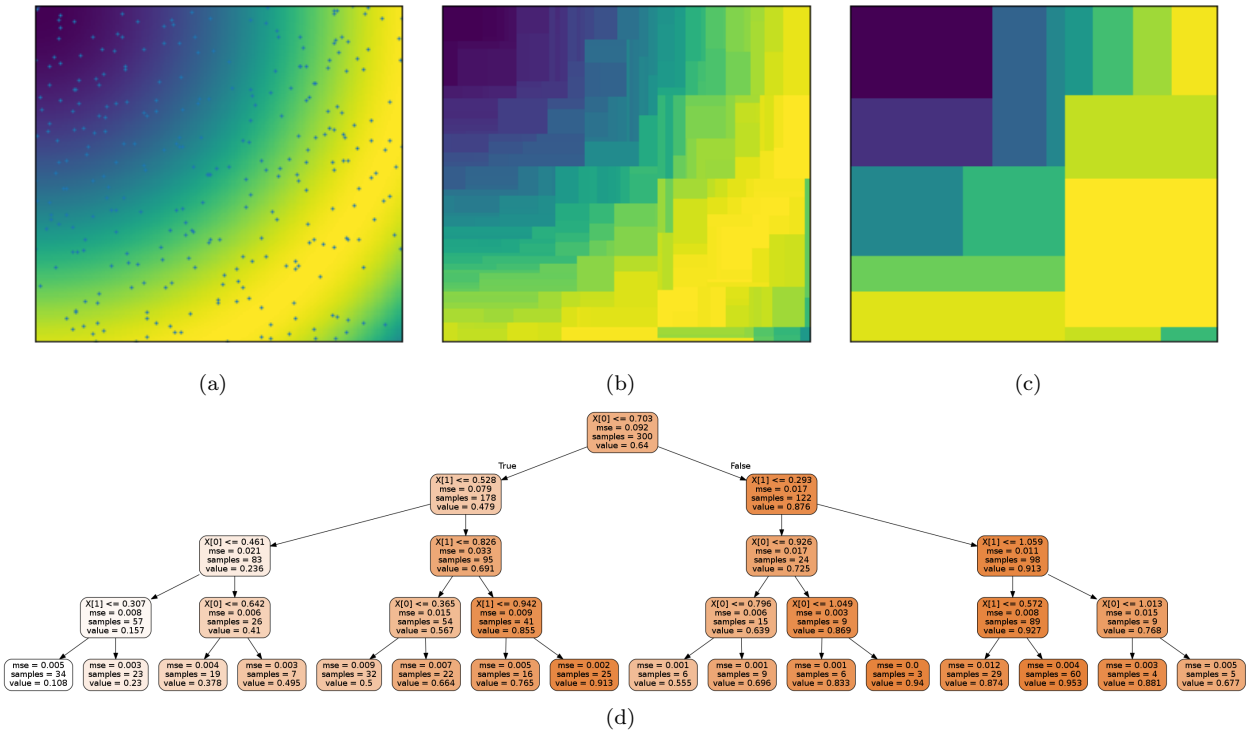


Figure 5.4: Decision tree prediction of a function plotted in panel (a). Dark color means higher values. The small dots are the data samples used to train the model. Panels (b) and (c) are predictions by a fully grown tree and a tree with maximum dept as 4 respectively. The tree with depth 4 is shown in panel (d).

Decision trees

Decision Tree is a non-parametric supervised learning method. It can be used for both regression and classification. It makes predictions by learning simple decision rules inferred from the data. Shown in Fig. 5.4(d) is a typical decision tree. It is simple to understand and to interpret. Start from the root node, anyone can follow the rule to do predictions.

We give an example of decision tree by fitting it with a continuous function. The function is plotted in Fig. 5.4(a). It takes a point on a two dimensional plane and returns a value. We randomly selected 300 data points as training data. They are shown as small dots in the figure. The root node of the tree contains all the data. A tree grows with the node split into two. The split is just cutting the parameter space in half. The position of the cut is chosen so that it maximally reduces the error. The nodes without child nodes are called leaf nodes. A tree stops growing when all its leaf nodes are pure. For a regression problem, this means there is one sample in the leaf. For a classification problem, this means all the samples in the leaf belong to the same class. We show the prediction results of a fully grown tree in Fig. 5.4(b). The maximal depth of this tree is 19. As can be seen in the figure, the parameter space is divided into small rectangles.

The predicted values within the same rectangle of the parameter space are the same. A bigger tree has finer rectangles, thus gives smaller training error. As a comparison, We plot the prediction of a smaller tree in Fig. 5.4(c). The corresponding tree is shown in Fig. 5.4(d). The color of the tree indicates the amplitude of the predicted value. In general, a tree divides the parameter space with n-orthotope when there are n features. It can be understood as finding the best piece wise constant functions.

Although the decision tree method has many advantages such as simple interpretation, it usually does not perform well on test data. Decision trees are prone to be over-complex and over-fit the data. Moreover, they are unstable because a completely different tree might be generated even with small variations in the training data.

Random forest

Random forest is an ensemble of decision trees. By averaging over trees, the over-fitting and unstable problems can be mitigated. The averaging process is also known as bagging. Think of each tree as one random variable T_b with mean f_b and variance σ_b^2. The random forest is then represented by the random variable $T = \frac{1}{B}\sum_{b=1}^{B} T_b$. If all the trees have the same mean and variance and are not correlated, T has mean f_b and variance $\frac{1}{B}\sigma_b^2$. Therefore, bagging can improve the variance of the model. Decision trees are usually unbiased but with high variance. Thus, they benefit from bagging.

Random forest differs in only one way from bagging. Only a subset of features are used at each node split when growing a tree. This reduces correlations between the trees. If some features are strong predictors, these features will always be selected making the trees correlated.

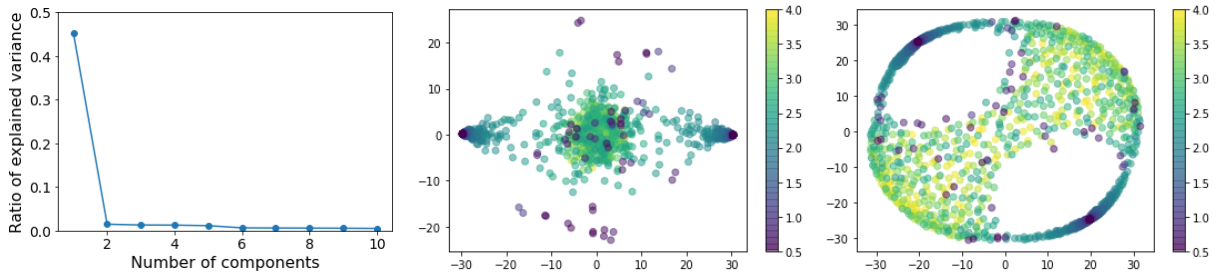
5.3 Ising ferromagnet

We first consider the simplest model: classical Ising ferromagnet on a square lattice with nearest-neighbor interaction. The Hamiltonian is

$$H = -\sum_{i,j}(S_{i,j}S_{i+1,j} + S_{i,j}S_{i,j+1}), \tag{5.9}$$

with $S_{ij} = \pm 1$. At low temperature, the system is in the ferromagnetic phase with all spins parallel. At high temperature, the thermal fluctuation overrides the ferromagnetic interaction, making the spins fluctuate randomly. Theoretically, the phase transition happens at temperature $T = 2/\ln(1+\sqrt{2})$.

We generate spin configurations from classical Monte Carlo. The system size is chosen as (30,30). We take periodic boundary condition, because there are large finite effects from open boundaries. We take



(a) (b) (c)

Figure 5.5: Analysis of Ising spin configuration data. Panel (a) shows the percentage of sample variances of the first ten principal components. We plot the data points in 2D using (b) PCA and (c) MDS. The color indicates temperature.

spin configurations after thermalization as input data. Thus, the data has 900 features. The samples are labeled as ferromagnetic or paramagnetic based on the temperature used. We generate spin configurations at temperatures from 0.5 to 4 with step 0.05. The number of samples should be more than the number of features for linear models. Otherwise, the model will definitely over fits the data due to the small sample size. Therefore, we need to keep at least 13 configurations at each temperature. We choose to take 200.

5.3.1 Data visualization

Before we build any model, we would like to know what the data looks like. We take 20 configurations at each temperature for visualization purpose. Our data has 900 features, so the data points are in a 900-dimensional space. Since the data elements can only be one or minus one, all data points are at corners of a high dimensional cube. It is hard to think about objects in a space of more than three dimensions. For simplicity consider a three dimensional cube with corners at $(\pm 1, \pm 1, \pm 1)$. The ferromagnetic spin configurations have all spins parallel. They occupy the two corners that are the most further away from each other. The remaining corners in between are paramagnetic.

In the 900-dimensional space, the situation is similar. The ferromagnetic phase has spin configurations that locate near two corners. We want to project the data points onto a two dimensional plane. The plane is chosen so that the data spreads the most on this plane. This is basically what principal component analysis (PCA) does. It finds a direction that maximize the variance if all data are projected on it. The direction is called the first principal component (1st PC). Then among directions perpendicular to the 1st PC, the direction with maximal variance is found. We can keep doing it until reaching the dimension of the space.

We calculate the variance of data along all the PC, and plot their percentage in Fig. 5.5(a). As can be seen from the figure, the 1st PC explained almost half of the total variance. Therefore, the data can be well separated along a line. We plot 2D projection of the data in Fig. 5.5(b). The light color indicates

high temperature, thus the paramagnetic phase. As can be seen, data generated from low temperature concentrates on two points ($\pm$30,0), while high temperature configurations lay between them. The 1st PC is basically sum of all spins. There are some dark points scattered in the middle region. Those are configurations with domain walls, as the sum of spins is close to zero.

In addition to the PCA, we can visualize high-dimensional data with muli-dimensional scaling (MDS). This method project data to low dimensional space with the pair distance preserved. In general, we find the projections that minimize the difference between the distance matrix of projected and original data points. We plot the result in Fig. 5.5(c). Same as the PCA projection, the configurations with domain walls scatter among the paramagnetic ones.

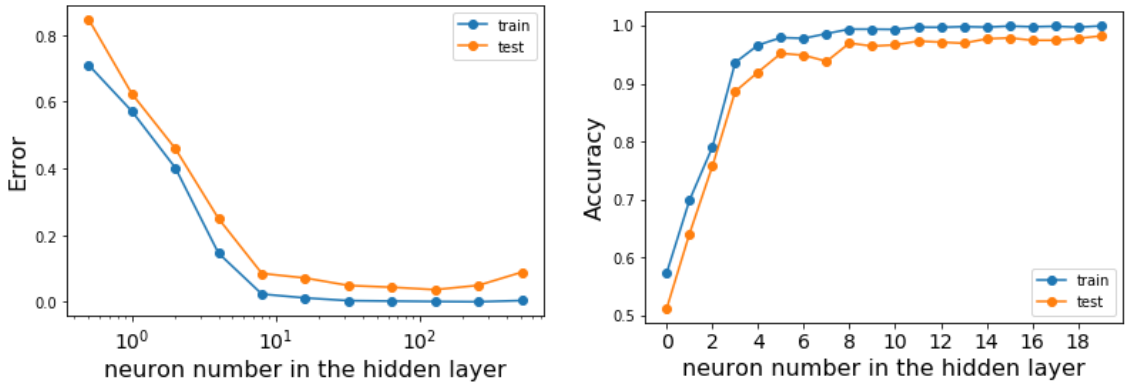
5.3.2 Supervised learning

Before we go to any sophisticated models, we always use a linear model to set benchmarks. For this binary classification problem, we apply the logistic regression with $L-2$ norm regularization. We expect the linear model to be sufficient for this problem, because we can tell the phase from the total magnetization, which is just the mean spin value. Unfortunately, the logistic regression can only give random guesses even for the training data. The reason is that the ferromagnetic phase is divided by the paramagnetic phase into two disconnected regions in the parameter space (shown in Fig. 5.5(b)). We cannot use one plane to separate the two phases in the parameter space. We set benchmarks with predictions by magnetization M. We predict the configuration to be in the magnetic phase if M is larger than 0.5. With this criteria, we get prediction accuracy as 0.945.

Feed-forward neural networks

We apply neural networks to do the classification. The model is trained using a feed-forward neural network with one hidden layer. The number of neurons in the hidden layer is a hyper-parameter that can be tuned. We vary this number from zero to 512. When it is zero, the model reduces to the logistic regression. Increasing the number of neurons makes the model more complex. We used ReLU as our activation function. The optimization algorithm is chosen as Adam.

Use the cross entropy as error. We plot the error for models with different number of neurons in Fig. 5.6(a). The data set is split into training and testing set with ratio 2:1. We use log scales for the x-axis. This figure has the similar structure with Fig. 5.1. Fewer neurons indicates simpler model. Both training and testing errors are large when the neuron number is small. While training error keeps decreasing with more neurons, the testing error goes up with larger neuron numbers, indicating over-fitting. The model has



(a) (b)

Figure 5.6: Model performance of the feed-forward neural network with one hidden layer. We vary the neuron number. The error and accuracy for both the training and testing data are shown in panel (a) and (b) respectively. Training data performances slightly better than testing data with smaller error and higher accuracy.

low training and testing errors when the number of neurons is between 8 to 128.

Define accuracy as the percentage of correct predictions. In Fig. 5.6(b), we plot the training and testing accuracy verses the neuron number. Both of them get closer to one for larger neuron numbers. This is consistent with the decreasing error in Fig. 5.6(a).

We take the model with 8 neurons in the hidden layer. A simple calculation yields the total number of parameters of this model is $(1+900)\times 8+(1+8)\times 2 = 7,226$. It is close to $8,174$, the number of samples of the training data. The training accuracy is 0.99 and testing accuracy is 0.96. The errors are 0.037 and 0.11 for training and testing data respectively. Training takes only a few seconds. We plot the predictions for the test data in Fig. 5.7(a). The blue dots are predictions for individual spin configurations. And the orange line is their average for the corresponding temperature. The vertical black dashed line indicates the theoretical transition temperature, while the horizontal black dashed line is a guide of the eye for 0.5 probability.

Note that most wrong predictions appear near the transition point. However, some configurations at very low temperature are predicted as paramagnetic with high probability. These configurations contain domain walls that separate the space into regions with parallel spins within each region.

Convolutional neural networks

The input data are spin configurations of a 2D lattice. they are images which can be best trained using convolutional neural networks. We use the same training and testing data. The convolutional neural network has one convolution cell with 32 filter channels, followed by max-pooling with 2×2 windows. The parameter number of this model is comparable with the feed-forward neural network we used before.

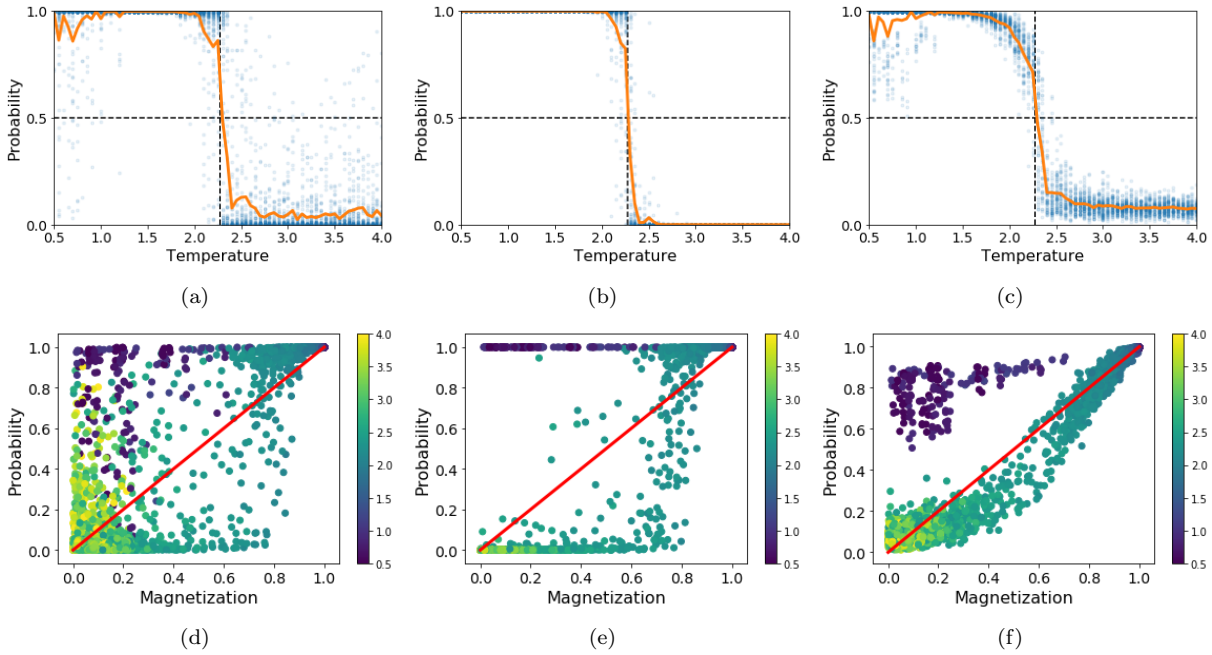


Figure 5.7: Predictions of (a) feed-forward neural networks with one hidden layer, (b) convolutional neural networks, and (c) random forest. The blue dots are predicted probabilities of being ferromagnetic for individual spin configurations. The orange line is the averaged probabilities. The vertical black dashed line indicates the theoretical transition temperature. Comparison between the prediction and magnetization is shown in panels (d) (e) and (f) for feed-forward neural networks, convolutional neural networks, and random forest respectively. The color indicates temperature used to generate the spin configuration. The red line from (0,0) to (1,1) is drawn to guide the eye.

The training accuracy is 0.997 and testing accuracy is 0.991. The errors are 0.0886 and 0.024 for training and testing data respectively. Training takes approximately one minutes. The convolutional neural network behaves better than the feed-forward neural network. The predictions on testing data is plotted in Fig. 5.7(b). Compared with Fig. 5.7(a), the convolutional network makes correct predictions for all spin configurations deep in the phase. It also has better performances near the transition temperature.

Random forest

Next, we use random forest as our model. We choose 100 trees with maximum depth as 10. The training takes only a few seconds. The training accuracy is 0.996 and testing accuracy is 0.988. Compared with the neural networks, random forest can make high accuracy predictions with short training time. The prediction results are shown in Fig. 5.7(c).

Conclusion

As a summary for the four models. The logistic regression cannot make reasonable predictions because the underlying relation is not linear. All the other three models can make predictions with accuracy of more than 0.95.

A comparison of random forest and neural network models is shown in Fig. 5.7. The convolutional neural network gives the most accurate predictions. As can be seen from Fig. 5.7(b), the model only gives uncertain predictions near phase boundaries. With the blue dots as predicted probability of being in the ferromagnetic phase, and the orange line as their average, we only see deviation of blue dots from the orange line near transition temperature. On the other hand, predictions of the other two models varies for configurations from the same temperature.

Since the predicted probability is a number from zero to one and the absolute value of magnetization also falls in this range, we plot their relationship in Fig. 5.7. The random forest model gives the highest correlation between the magnetization and the probability. Moreover, it makes correct predictions for those configurations trapped in local minimum. Those configurations are represented by dark purple dots in Fig. 5.7(f). Although they have low magnetization, the model can predict them as in the ferromagnetic phase. The predictions from neural networks do not show significant relationship with the magnetization.

5.4 Disordered chiral chain

We start with the disordered chiral chain in Ref. [114]. It is a one-dimensional topological insulator of the AIII symmetry class. Defined on a one-dimensional chain with two sites A and B in one unit cell, the Hamiltonian is written as

$$H = \sum_n \left[\frac{t_n}{2} c_n^\dagger \left(\sigma_x + i\sigma_y \right) c_{n+1} + h.c \right] + \sum_n m_n c_n^\dagger \sigma_y c_n. \tag{5.10}$$

Here $c_n^\dagger = (c_{n,A}^\dagger, c_{n,B}^\dagger)$ are fermion creation operators in unit cell n. We put disorder to both the hopping and mass terms, i.e. $t_n = 1 + W_1\omega_1$, and $m_n = m + W_2\omega_2$, where ω_1 and ω_2 are random variables generated from box distribution on $[-0.5, 0.5]$.

5.4.1 Analytic properties of the Hamiltonian

In the clean limit, the system has translational symmetry. We can write the Bloch Hamiltonian

$$\mathcal{H}(k) = t\cos k\sigma_x + (t\sin k + m)\sigma_y \tag{5.11}$$

The chiral symmetry operator $\mathcal{C} = \sigma_z$ anti-commutes with $\mathcal{H}(k)$: $\mathcal{C}\mathcal{H}(k)\mathcal{C}^{-1} = -\mathcal{H}(k)$. Winding number ν can be calculated as the topological invariant for the AIII class, which is represented by the $\mathbb{Z}$ classification [109]. ν can be calculated for an infinite chain as [132]

$$\nu = \frac{1}{2\pi i}\int_0^{2\pi}\frac{\partial_k q(k)}{q(k)}\mathrm{d}k. \tag{5.12}$$

Here, $q(k) = te^{ik} - im$ is the off-diagonal term of $\mathcal{H}(k)$. Write the Bloch Hamiltonian in the form of $\mathcal{H}(k) = d_x(k)\sigma_x + d_y(k)\sigma_y$ with $d_x^2(k) + d_y^2(k) = 1$. Then the winding number ν calculates the number of times point (d_x, d_y) goes around the origin as k goes through the whole Brillouin Zone. When $|m| < |t|$, the system is in symmetry protected topological (SPT) phase with winding number $\nu = 1$. Otherwise, $\nu = 0$. On a finite lattice of length L, the integral becomes sum. We get the discrete version of Eq. 5.12

$$\nu = \frac{1}{L}\sum_{k=0}^{L-1}\frac{te^{ik/2\pi}}{te^{ik/2\pi} - im}. \tag{5.13}$$

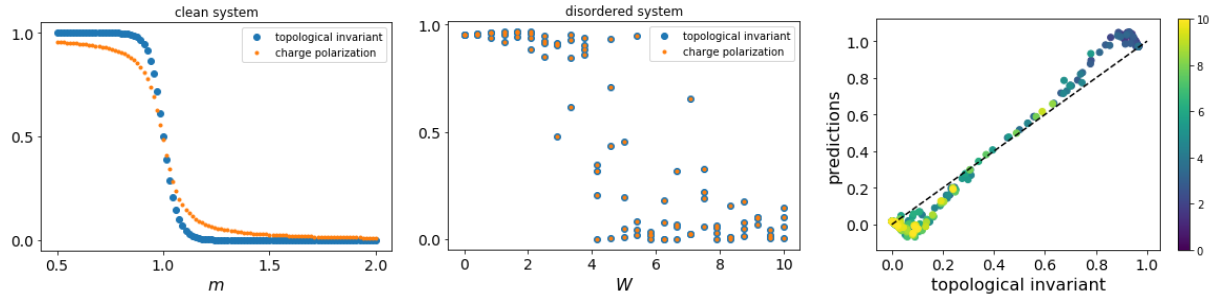
When disorder is turned on, in the limit that $W_2 \gg t$, the system is completely dimerized within individual unit cells despite the value of m. This gives the same product states as in the atomic limit, which are topologically trivial. Therefore, there must be a phase transition when W_2 is gradually increasing with $|m| < |t|$. This topological phase transition point is consistent with the divergence in localization length and can be calculated from formula [114].

$$\frac{|2t + W_1|^{t/W_1+1/2}|2m - W_2|^{m/W_2-1/2}}{|2t - W_1|^{t/W_1-1/2}|2m + W_2|^{m/W_2+1/2}} = 1. \tag{5.14}$$

We set $t = 1$. This equation defines a surface that divides the three-dimensional parameter space into two parts. The topological invariant can be define in real space by Ref [114]

$$\nu = -\mathrm{Tr}\{Q_\mp[X, Q_\pm]\}. \tag{5.15}$$

Here X is position operator. $Q_\pm$ and $Q_\mp$ are projects of homotopically equivalent flatband Hamiltonian Q.



(a) (b) (c)

Figure 5.8: Topological invariant (big blue dots) and charge polarization (orange small dots) are calculated for the chiral chain with $L = 30$. The clean and disordered cases are shown in panel (a) and (b) respectively. Panel (c) shows the predicted topological invariant verses the calculated value for disordered data. The color bar is W, indicating the strength of disorder. The dashed line connects points (0,0) and (1,1).

We have relations $Q = Q_{\pm} + Q_{\mp}$ and $Q_{\pm} = Q_{\mp}^{-1}$.

5.4.2 Machine learning topological invariant

Predict disordered systems with models trained by clean systems

Note that the real space topological invariant (Eq. 5.15) is the same as the charge polarization of an open chain. The topological invariant defined in Eq. 5.12 is for an infinite or periodic chain. In a clean system, we can calculate both invariants. The two do not match each other near the transition point due to finite size effect from the edges.

We calculate the topological invariant for a chain with $L = 30$. Both the clean limit and the disordered chain are considered. In Fig. 5.8(a), we plot the topological invariant calculated by Eq. 5.12 as blue dots. Even though the system size is small, the topological invariant has a sharp decrease at the transition point $m = 1$. The small orange dots are the charge polarization calculated with the same parameter but in an open chain. As can be seen, the two values deviates a lot near the phase boundary. With larger system size L, the deviation will be smaller.

In Fig. 5.8(b), the blue dots are the topological invariant calculated by Eq. 5.15. We set $m = 0.5$ and $W = 2W_1 = W_2$. The charge polarization is shown as smaller orange dots. It is clear that the charge polarization overlaps exactly with the topological invariant calculated in real space.

We apply machine learning methods to predict topological invariant with the existence of disorder for small systems. We use charge density as our features, because they are directly related with the charge polarization. The training data is the obtained from the clean systems; and we test on disordered systems with the topological invariant calculated in real space. A linear regression with $L - 2$ norm regularization

is used. We take 100 data samples by varying m from 0.5 to 2 for a $L = 30$ clean chain as training data. We only used a small data set for training because we want to deal with small systems. We used the same data set as in Fig. 5.8(a). The testing data is chosen as samples plotted in Fig. 5.8(b). There are 100 data samples in total. The prediction results are shown in Fig. 5.8(c). If the predictions are the same as topological invariant calculated in real space, the points should fall onto the dashed back line. It is a line from point (0,0) to point(1,1). The color represents the disorder strength W. Lighter color means higher disorder. From the figure, we see that predictions are consistent with the calculated values.

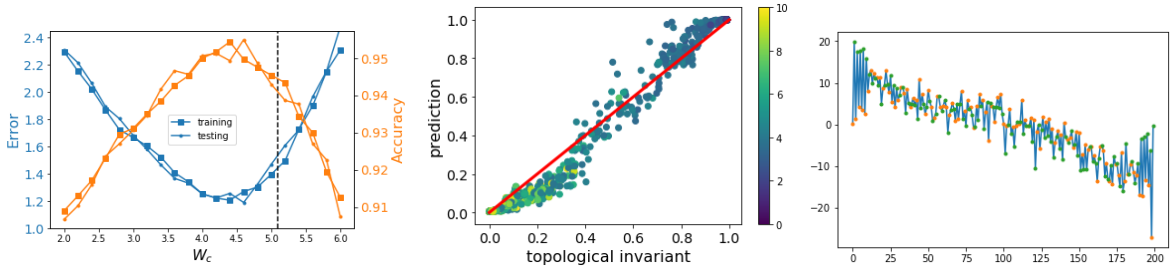
Linear models well predict topological invariant based on charge densities. The model trained with clean systems can be generalized to make predictions on systems with disorder. The predictions are close to the calculated values. They get closer to each other with increasing L. Note that we have avoided to choose L to be $4n$ with n as integers. The reason is there are divergences near phase boundaries when $L = 4n$. It is a finite size numerical effect. When L gets larger, the divergence goes away. This divergence messes up the linearity of the data, making the linear model behave strangely.

Trial-and-test method to locate phase boundary

Without any knowledge of the phase, we can use a trial-and-test method to tell determine the phase boundary [124, 125]. The idea is simple. We assume the system is characterized by one parameter p. Assume further that there are two phases separated by p_c. We do not know the value of p_c, but we can assume $p_c = p_i$ for several p_i values and do training and testing. If p_i is not the true critical point, the testing error would be large. We only need to try a few p_i values and find the value with the lowest error or highest accuracy. With this method, we can find the phase boundary of a system along parameter p.

We work with the disordered system of size $L = 100$ with open boundary conditions. Same as previous part, we set $m = 0.5$ and $W = 2W_1 = W_2$. 10,000 charge density configurations are generated for W from 0 to 10. We randomly choose 70% as training data and the remaining as testing data. We apply a simple logistic regression with $L - 2$ regularization. The error is calculated by cross-entropy and accuracy is the percentage of correct predictions. We vary the trial critical point from 2 to 8. The training and testing error and accuracy is shown in Fig. 5.9(a). We get the best performance when $W_c = 4$, which is close to the theoretical value.

Using the best model, i.e. model trained by setting $W_c = 4$, we compare the predicted probability with the topological invariant in Fig. 5.9(b). The red line is drawn as a guide of the eye. As can be seen from the figure, all data points falls close to the line, indicating that the topological invariant can best characterize the phases.



(a) (b) (c)

Figure 5.9: (a) Error (left axis) and accuracy (right axis) for training and testing data for different trial critical point values W_c. The best prediction happens at $W_c = 4$ with lowest testing error and highest testing accuracy. (b) Predicted probability of being in the topological phase verses topological invariant. The red line is from point (0,0) to point (1,1). Color indicates W, the strength of disorder. Lighter color means stronger disorder. (c) Coefficients of the logistic regression model.

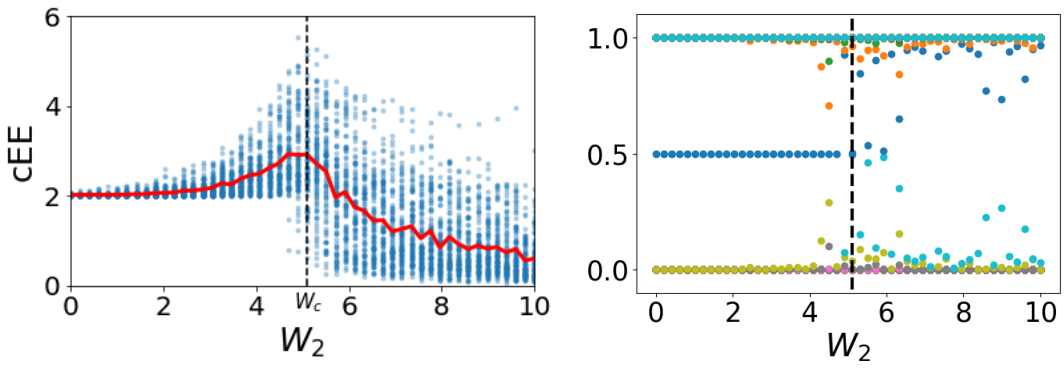
We plot the coefficients of the model in Fig. 5.9(c). Since we have two sites in one unit cell, there are $2L = 200$ coefficients, one for each site. The orange dots are values for sites of type A and green dots for type B sites. The charge densities on A sites do not change for different parameters. Therefore their coefficients are set to be zero. It is interesting to see that the coefficients for B sites increase linearly (approximately) with the site position. With a proper rescale, it resembles the calculation of charge density.

Note that the model finds the best coefficients from a statistical point of view. It is impossible to recover topological invariant from the training. However, we can gain some intuition or inspiration from the machine learning model.

5.4.3 Machine learning topological phase from entanglement spectrum

Entanglement entropy has been widely used to detect 1D topological phases [36]. Set $t = 1$, $m = 0.5$ and a small disorder strength for hopping $W_1 = 1$. We increase W_2 and the calculated central-cut entanglement entropy (cEE) using periodic boundary condition on a chain of length $L = 400$. We calculate 100 disorder configurations at each W_2 and present the raw data in Fig. 5.10(a) as blue dots. The values of entanglement entropy spread out when disorder gets large. We indicate the theoretical transition point with a dashed line. Notice that in the topological phase, the distribution of entanglement entropy at fixed W_2 never goes below 2. The upper bond of entanglement entropy increases with W_2 until transition point and then decreases to a fixed value. We average the entanglement entropy of the 100 configurations and plot in Fig. 5.10(a) as a red solid line. The averaged entanglement entropy bumps near the theoretical transition point. We can hardly get the transition point even with the averaged entanglement entropy.

Since entanglement entropy cannot be used to characterize topological phases when disorder is present,



(a) (b)

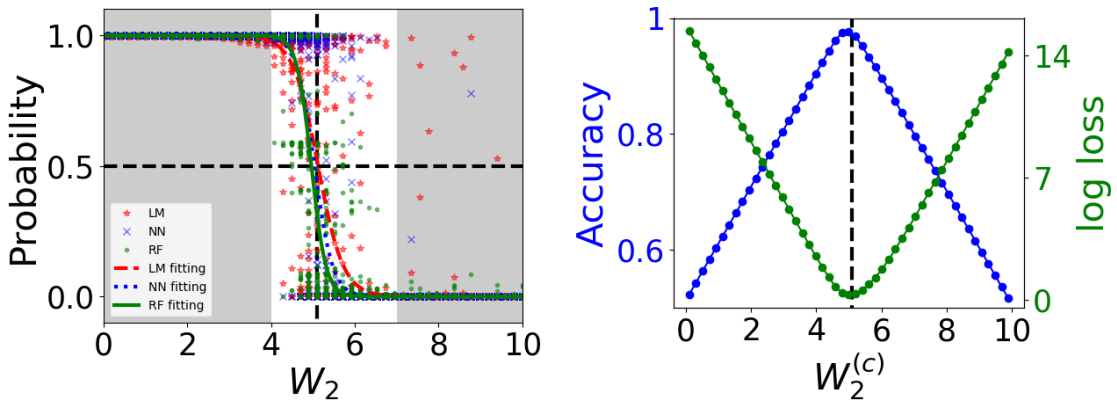
Figure 5.10: (a) Central-cut entanglement entropy and (b) Single particle entanglement spectrum of the disordered chiral chain. The vertical black dash line is the analytical transition point. The red line in panel (a) is the average value. Double degeneracy of the spectrum at 0.5 on the left hand side indicates SPT phase. There may be accidental degeneracies in the trivial phase due to disorder.

we apply machine learning techniques to ameliorate it. We calculated single particle ES [133] of the lattice model as the input data. Periodic boundary conditions are used on a chain of length $L = 400$ with t set as one. For simplicity, we first focus on a line in the 3D phase space $\{m, W_1, W_2\}$ with $m = 0.5$ and $W_1 = 1$. We plot the ES of one disorder configuration at each value of W_2 in Fig. 5.10(b). The black vertical line indicates the theoretical transition point calculated from Eq. 5.14. We can clearly see double degeneracy at 0.5 on the left part, which is a signature for SPT phases. In the region of strong disorder, there are no such degeneracies in general. However, there may be accidental degeneracies induced by disorder.

5000 training samples were generated with W_2 ranging from 0 to 4 and from 7 to 10. We intentionally skipped the area near the phase transition point, in hope that the RF model can locate it only with knowledge deep in the phase. Test data was generated separately over the whole range of W_2 from 0 to 10. Using the same training and testing data, we fit three models: linear model (LM), neural network (NN), and random forest (RF). The first two models were fitted as comparisons with the RF model. The python package sklearn [134] was used for training and predicting.

We show the prediction results of the three models in Fig. 5.11(a). The red stars, blue crosses, and green dots are predicted probabilities of being in the topological phase from LM, NN, and RF model respectively. We can see that most states are correctly classified for all the three models. The linear model has several misclassified states in the region of strong disorder, while NN and RF models only have wrong classifications near the phase boundary. We fit the predicted probability with function

$$f(x) = 1/(1 + e^{b+wx}) \tag{5.16}$$



(a) (b)

Figure 5.11: (a) Predicted probability of being in the topological phase by the three models: linear model (red star and dashed line), neural network (blue cross and dotted line), and random forest (green dot and solid line). The lines are fitted by Eq. 5.16. (b) Accuracy and log loss of prediction by random forest with different transition points $W_c^{(c)}$. The point with the highest accuracy (lowest log loss) is the true transition point.

The fitted lines are shown in the figure as a guide of eye. The black vertical dashed line indicates the true transition point; and the horizontal one is 0.5 probability. We chose the cutoff value as 0.5, i.e. when predicted probability is larger than 0.5, we say the state is in the topological phase. Otherwise, it is in the trivial phase.

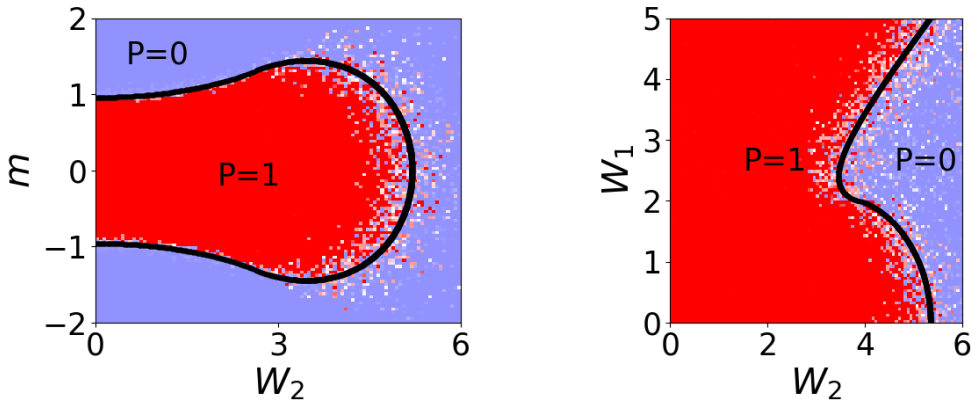
Define accuracy as the percentage of correctly predicted samples. We get the accuracy of the LM, NN, and RF model as 0.966, 0.974, and 0.977 respectively. Linear model has the lowest accuracy among the three, due to its simple linear assumption. NN and RF model behavior similarly.

To locate the phase transition point, we can find the crossing point of the 0.5 probability line with the probability fitting line. On the other hand, we can also use a confusion scheme [124]. We set the true transition point is at $W_2^{(c)}$ and calculate the prediction accuracy. If $W_2^{(c)}$ is the true transition point, we get high accuracy. Otherwise the accuracy is low. We plot the accuracy at different $W_2^{(c)}$ in Fig. 5.11(b). The $\wedge$ shape of the plot suggests the true transition point at $W_2 \approx 5$, which is consistent with the analytic result indicated by the vertical dashed line.

For completeness, we make the same plot for error. We measure the error of fitting by log loss or cross entropy

$$H(p, q) = \mathrm{E}_p[-\log q] \tag{5.17}$$

where, p is the true probability distribution and q is the predicted probability distribution. The $\vee$ shape of the plot is consistent with what we get from accuracy.



(a) (b)

Figure 5.12: The predicted phase diagram of (a) $W_1 = 1.0$ and (b) $m = 0.5$. The black solid lines are theoretical phase boundaries. P is the predicted probability of being in the topological phase.

It is interesting to see that the trained model can locate the transition point with respectively high accuracy. Even though the model is not given the information near the phase boundary, it can make correct predictions. Further, we apply the trained model on other regions of the phase space.

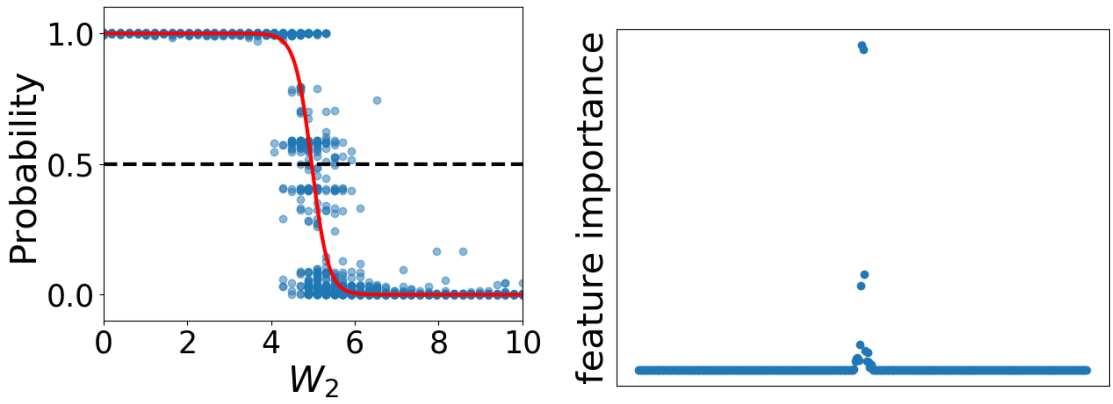
We take two cross sections of the three-dimensional phase diagram. One with $W_1 = 1$ and the other with $m = 0.5$. The phase diagrams are plotted in Fig 5.12. Color indicates the predicted probability of being in topological phase. Theoretical phase boundaries are plotted as solid black lines. As can be seen, the model makes predictions with high confidence deep in the phase. When disorder is small, predicted phase boundaries match perfectly with theoretical ones, while there are some deviations near the phase boundaries at large disorder.

Since the ES is a general feature for topological phases, we should be able to extend it to other topological systems. This is indeed true when we apply the RF model to the Kitaev chain. The Hamiltonian of the Kitaev chain is [45]

$$H = \sum_n [t_n \; ib_n a_{n+1} + m_n \; ia_n b_n], \tag{5.18}$$

where a_n and b_n are Majorana fermions. We add disorder to the parameters $t_n = 1 + W_1\omega_1$ and $m_n = m + W_2\omega_2$. Note that same as the t_n and m_n in the chiral model of Eq. 5.10, t_n and m_n here can also be interpreted as inter-cell and intra-cell hopping terms.

The prediction results are shown in Fig. 5.13(a). We added terms that break time reversal and chiral symmetry: $\sum_n ia_n a_{n+1} + ib_n b_{n+1}$. Similar with Fig. 5.11(a), the dots representing predicted probabilities of being in the topological phase. The red line is fitted using Eq. 5.16. There is clearly a phase transition near $W_2 = 8$.



(a) (b)

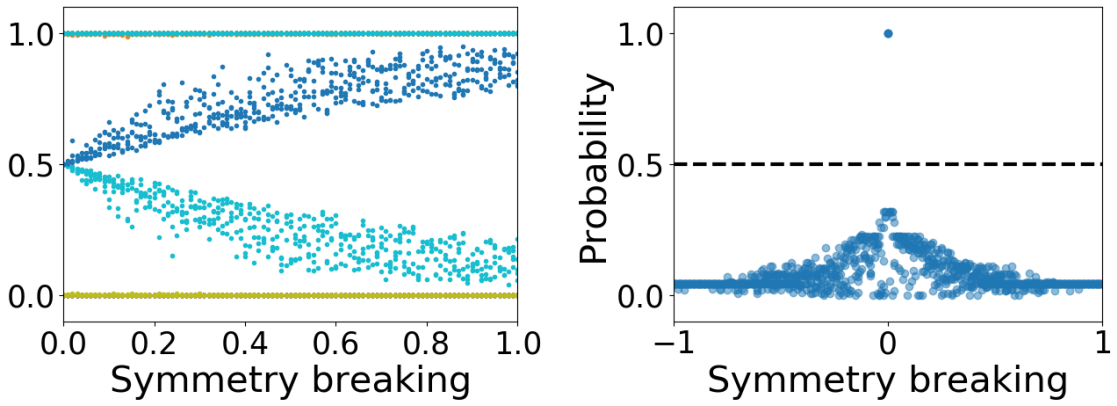
Figure 5.13: (a) The predicted probability of being in the topological phase of Kitaev model. (b) Feature importance from random forest model. The high value in the middle indicates high importance of mid-gap states in entanglement spectrum.

To see what random forest does exactly, we plot the feature importance of the model (Fig. 5.13(b)). Feature importance measures the number of splits in a tree that includes the feature [135]. High feature importance means the feature is more likely to be used to seperate classes. As shown by the figure, the middle values of the ES shows the highest influence on predictions. It suggests that the RF model focuses on the degeneracy of the ES to do classification.

To illustrate this point, we add a small σ_z term that breaks chiral symmetry. We fix other parameters as $m = 0.5$ and $W_1 = W_2 = 1$, so that the system is in topological phase when symmetry is not broken. Since the topological phase is protected by chiral symmetry, turning on a σ_z term immediately breaks down the topological phase and the degeneracy in entanglement spectrum is lifted.

As the symmetry breaking term becomes stronger, the degenerate states in the entanglement spectrum get further away from each other (shown in Fig. 5.14(a)). The predictions made by the RF model are shown in Fig. 5.14(b). The blue dots are the raw prediction probabilities. When there is no symmetry breaking term, i.e. chiral symmetry is preserved, the model predicts the states as topological. As long as the symmetry breaking strength is non-zero, the predicted probability immediately drops below 0.5. This probability goes down gradually as the symmetry breaking strength increase.

We check whether the prediction is related with entanglement gap. We use the entanglement degeneracy to make predictions of the phase. If the gap is smaller than 0.001, we say the state has probability one as being in the topological phase. Otherwise, we say the state is in the trivial phase. The predicted results give accuracy 0.977, which is close to the accuracy of random forest classifier. We plot the predictions from degeneracy in Fig. 5.15(a) by orange dots. Since we can only predict one or zero, we take average of the



(a) (b)

Figure 5.14: (a) Entanglement spectrum with symmetry breaking term. (b) The probability of being in the topological phase with symmetry breaking term added to the system. All configurations are in the trivial phase except when symmetry is preserved.

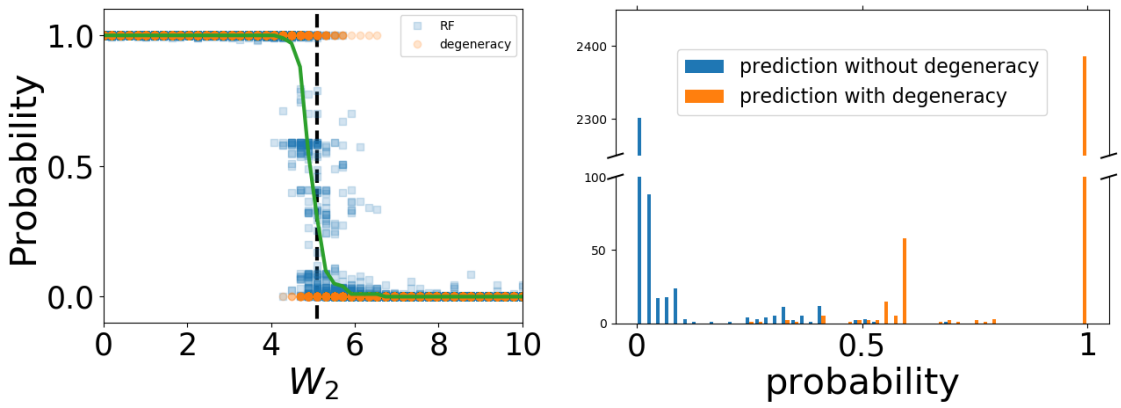
predictions as the probability (green line). The predictions of random forest is show in blue square for comparison. Clearly, the two predictions matches perfectly.

We plot the distribution of random forest predictions in Fig. 5.15(b) when degeneracy exists or not. Note that the y-axis is cut in the middle. When there is degeneracy, the random forest classifier predict with probability one in most cases. Similarly, the predictions by random forest classifier are mostly zero when there are no degeneracies. In addition to entanglement gaps, random forest makes predictions based on other features as well. As can be seen from the distribution plot, there are small/big prediction probabilities when there are/not degeneracies.

Random forest can extract most information from features when feature number is not too large. In this case, entanglement spectrum may not be able to capture all the information with disorder. This is the main obstacle to improve the prediction accuracy. On the other hand, the states near phase boundaries may be in the wrong side of phase, making accuracy even lower.

5.5 Summary

In summary, we used machine learning techniques to classify phases. Both traditional methods, such as logistic regression and random forest, and deep learning methods, such as neural networks are used as our models. We applied logistic regression, random forest and neural networks to classify classical Ising ferromagnets and topological phases. Compared with linear model, random forest gives better predictions. On the other hand, it preserves the easy interpretation ability of linear model. Neural network gives similar



(a) (b)

Figure 5.15: (a) Predictions by degeneracy in entanglement spectrum (blue dots) compared with random forest classifier (orange square). The green line is average of predictions by degeneracy. Theoretical phase transition point is indicated by a vertical dashed line. (b) Distribution of predictions by random forest classifier. When entanglement spectrum has (no) double degeneracy, most predictions give probability one as being in the topological (trivial) phase.

accuracy as the random forest, but requires more time for training. The behavior of neural networks can be improved by fine tuning hyper parameters.

With entanglement spectrum as features, we trained a random forest model. Because of the robustness of entanglement spectrum, the model trained on a small training data can be well generalized to test data in a larger phase space, and even to other models. A closer look at the RF model indicates that the model is capturing the degeneracy of entanglement spectrum in the middle.

Chapter 6

Conclusions and outlook

In this thesis, we studied one dimensional symmetry protected topological (SPT) phases. The density matrix renormalization group (DMRG) method is applied to solve the chiral Z_3 parafermionic chain and the interacting inversion symmetric topological superconductor. We determined the phase diagrams of the two systems based on quantum entanglement analysis. We discovered three phases in the Z_3 parafermionic chain: a topological phase, a trivial phase and an intervening incommensurate phase. The tricritical point where the three phases meet at a single point is located by analyzing the Lifshitz type transition. The topological superconductor is generalized from the Kitaev p-wave wire. The calculation of teleportation and the fractional Josephson effects indicates that the ground state of the topological phase is a condensation of four electrons instead of cooper-pairs. Moreover, the inversion symmetry restricts the edge modes of the system to be cooper-pairs other than two uncorrelated electrons. We then applied machine learning methods to find universal classifiers for SPT phases. After comparing three machine learning methods, we chose random forest as our model The trained model can give high accuracy predictions to other regions in the phase space, even other systems with different symmetries.

Many efforts have been made on applying machine learning techniques to physics problems, but there are still many things can be done. We used entanglement spectra as features to train the model. They are properties calculated from the ground states. The performance of the model depends highly on the type of features we choose. We still need to use our prior knowledge about the system to choose proper features. It is interesting to see if it is possible to train a model with ground state wavefunctions. For a classical model there is no big obstacle, because many work has been done based on spin configurations. However, a quantum state wavefunction cannot be represented exactly in general. The best approximation for now belongs to tensor networks. For one dimensional systems, it would thus be interesting to build models that can extract information automatically from matrix product states and then make predictions.

Another potentially interesting problem to consider is to extend the ability of the unsupervised learning. Most physics models people study with machine learning have analytic solutions, because we need to label the phases before any training can be done. With unsupervised learning we do not need those knowledge,

so we can work on unsolved problems. There have already been quite a few works on this direction, but still with the solved models. It will be exciting if unsupervised learning can detect phases that are note previously known.

References

[1] L. Landau, "Zur theorie der phasenumwandlungen ii," *Phys. Z. Sowjetunion*, vol. 11, pp. 26–35, 1937.

[2] X. G. Wen, "Vacuum degeneracy of chiral spin states in compactified space," *Phys. Rev. B*, vol. 40, pp. 7387–7390, Oct 1989.

[3] X.-G. Wen, "Topological orders in rigid states," *International Journal of Modern Physics B*, vol. 4, no. 02, pp. 239–271, 1990.

[4] X. Chen, Z.-C. Gu, and X.-G. Wen, "Local unitary transformation, long-range quantum entanglement, wave function renormalization, and topological order," *Phys. Rev. B*, vol. 82, p. 155138, Oct 2010.

[5] R. B. Laughlin, "Anomalous quantum hall effect: An incompressible quantum fluid with fractionally charged excitations," *Phys. Rev. Lett.*, vol. 50, pp. 1395–1398, May 1983.

[6] N. Read and S. Sachdev, "Large-n expansion for frustrated quantum antiferromagnets," *Phys. Rev. Lett.*, vol. 66, pp. 1773–1776, Apr 1991.

[7] X. G. Wen, F. Wilczek, and A. Zee, "Chiral spin states and superconductivity," *Phys. Rev. B*, vol. 39, pp. 11413–11423, Jun 1989.

[8] A. Kitaev, "Anyons in an exactly solved model and beyond," *Annals of Physics*, vol. 321, no. 1, pp. 2 – 111, 2006. January Special Issue.

[9] X. Chen, Z.-C. Gu, Z.-X. Liu, and X.-G. Wen, "Symmetry protected topological orders and the group cohomology of their symmetry group," *Phys. Rev. B*, vol. 87, p. 155114, Apr 2013.

[10] M. Z. Hasan and C. L. Kane, "Colloquium: Topological insulators," *Rev. Mod. Phys.*, vol. 82, pp. 3045–3067, Nov 2010.

[11] X.-L. Qi and S.-C. Zhang, "Topological insulators and superconductors," *Rev. Mod. Phys.*, vol. 83, pp. 1057–1110, Oct 2011.

[12] A. P. Schnyder, S. Ryu, A. Furusaki, and A. W. W. Ludwig, "Classification of topological insulators and superconductors in three spatial dimensions," *Phys. Rev. B*, vol. 78, p. 195125, Nov 2008.

[13] S. Ryu, A. P. Schnyder, A. Furusaki, and A. W. Ludwig, "Topological insulators and superconductors: tenfold way and dimensional hierarchy," *New Journal of Physics*, vol. 12, no. 6, p. 065010, 2010.

[14] R. B. Laughlin, "Quantized hall conductivity in two dimensions," *Phys. Rev. B*, vol. 23, pp. 5632–5633, May 1981.

[15] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, "Quantized hall conductance in a two-dimensional periodic potential," *Phys. Rev. Lett.*, vol. 49, pp. 405–408, Aug 1982.

[16] F. D. M. Haldane, "Model for a quantum hall effect without landau levels: Condensed-matter realization of the "parity anomaly"," *Phys. Rev. Lett.*, vol. 61, pp. 2015–2018, Oct 1988.

[17] Y. Hatsugai, "Chern number and edge states in the integer quantum hall effect," *Phys. Rev. Lett.*, vol. 71, pp. 3697–3700, Nov 1993.

[18] C. L. Kane and E. J. Mele, "Quantum spin hall effect in graphene," *Phys. Rev. Lett.*, vol. 95, p. 226801, Nov 2005.

[19] B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, "Quantum spin hall effect and topological phase transition in hgte quantum wells," *Science*, vol. 314, no. 5806, pp. 1757–1761, 2006.

[20] L. Fu, "Topological crystalline insulators," *Phys. Rev. Lett.*, vol. 106, p. 106802, Mar 2011.

[21] T. H. Hsieh, H. Lin, J. Liu, W. Duan, A. Bansil, and L. Fu, "Topological crystalline insulators in the snte material class," *Nature Communications*, vol. 3, pp. 982 EP –, Jul 2012. Article.

[22] Y. Tanaka, Z. Ren, T. Sato, K. Nakayama, S. Souma, T. Takahashi, K. Segawa, and Y. Ando, "Experimental realization of a topological crystalline insulator in snte," *Nature Physics*, vol. 8, pp. 800 EP –, Sep 2012.

[23] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, "Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions," *Rev. Mod. Phys.*, vol. 68, pp. 13–125, Jan 1996.

[24] D. Ceperley, G. V. Chester, and M. H. Kalos, "Monte carlo simulation of a many-fermion study," *Phys. Rev. B*, vol. 16, pp. 3081–3099, Oct 1977.

[25] J. B. Anderson, "Quantum chemistry by random walk. h 2 p, h+ 3 d 3 h 1 a 1, h2 $3\sigma+$ u, h4 $1\sigma+$ g, be 1 s," *The Journal of Chemical Physics*, vol. 65, no. 10, pp. 4121–4127, 1976.

[26] J. Barker, "A quantum-statistical monte carlo method; path integrals with boundary conditions," *The Journal of Chemical Physics*, vol. 70, no. 6, pp. 2914–2918, 1979.

[27] S. R. White, "Density matrix formulation for quantum renormalization groups," *Phys. Rev. Lett.*, vol. 69, pp. 2863–2866, Nov 1992.

[28] G. Vidal, "Classical simulation of infinite-size quantum lattice systems in one spatial dimension," *Phys. Rev. Lett.*, vol. 98, p. 070201, Feb 2007.

[29] J. Jordan, R. Orús, G. Vidal, F. Verstraete, and J. I. Cirac, "Classical simulation of infinite-size quantum lattice systems in two spatial dimensions," *Phys. Rev. Lett.*, vol. 101, p. 250602, Dec 2008.

[30] G. Vidal, "Class of quantum many-body states that can be efficiently simulated," *Phys. Rev. Lett.*, vol. 101, p. 110501, Sep 2008.

[31] G. Torlai and R. G. Melko, "Learning thermodynamics with boltzmann machines," *Phys. Rev. B*, vol. 94, p. 165134, Oct 2016.

[32] J. Chen, S. Cheng, H. Xie, L. Wang, and T. Xiang, "Equivalence of restricted boltzmann machines and tensor network states," *Phys. Rev. B*, vol. 97, p. 085104, Feb 2018.

[33] M. Levin and X.-G. Wen, "Detecting topological order in a ground state wave function," *Phys. Rev. Lett.*, vol. 96, p. 110405, Mar 2006.

[34] A. Kitaev and J. Preskill, "Topological entanglement entropy," *Phys. Rev. Lett.*, vol. 96, p. 110404, Mar 2006.

[35] H. Li and F. D. M. Haldane, "Entanglement spectrum as a generalization of entanglement entropy: Identification of topological order in non-abelian fractional quantum hall effect states," *Phys. Rev. Lett.*, vol. 101, p. 010504, Jul 2008.

[36] F. Pollmann, A. M. Turner, E. Berg, and M. Oshikawa, "Entanglement spectrum of a topological phase in one dimension," *Phys. Rev. B*, vol. 81, p. 064439, Feb 2010.

[37] A. Chandran, V. Khemani, and S. L. Sondhi, "How universal is the entanglement spectrum?," *Phys. Rev. Lett.*, vol. 113, p. 060501, Aug 2014.

[38] J. S. Bell, "Einstein-podolsky-rosen experiments," in *John S Bell on the Foundations of Quantum Mechanics*, pp. 74–83, World Scientific, 2001.

[39] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE mobile computing and communications review*, vol. 5, no. 1, pp. 3–55, 2001.

[40] P. Calabrese and J. Cardy, "Entanglement entropy and quantum field theory," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2004, no. 06, p. P06002, 2004.

[41] X.-L. Qi, H. Katsura, and A. W. W. Ludwig, "General relationship between the entanglement spectrum and the edge state spectrum of topological quantum states," *Phys. Rev. Lett.*, vol. 108, p. 196402, May 2012.

[42] I. Peschel, "Special review: Entanglement in solvable many-particle models," *Brazilian Journal of Physics*, vol. 42, pp. 267–291, Aug 2012.

[43] U. Schollwck, "The density-matrix renormalization group in the age of matrix product states," *Annals of Physics*, vol. 326, no. 1, pp. 96 – 192, 2011. January 2011 Special Issue.

[44] J. A. Kjäll, M. P. Zaletel, R. S. Mong, J. H. Bardarson, and F. Pollmann, "Phase diagram of the anisotropic spin-2 xxz model: Infinite-system density matrix renormalization group study," *Physical Review B*, vol. 87, no. 23, p. 235106, 2013.

[45] A. Y. Kitaev, "Unpaired majorana fermions in quantum wires," *Physics-Uspekhi*, vol. 44, no. 10S, p. 131, 2001.

[46] P. Fendley, "Parafermionic edge zero modes in zn-invariant spin chains," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2012, no. 11, p. P11020, 2012.

[47] E. Fradkin and L. P. Kadanoff, "Disorder variables and para-fermions in two-dimensional statistical mechanics," *Nuclear Physics B*, vol. 170, no. 1, pp. 1 – 15, 1980.

[48] S. Ostlund, "Incommensurate and commensurate phases in asymmetric clock models," *Phys. Rev. B*, vol. 24, pp. 398–405, Jul 1981.

[49] S. Howes, L. P. Kadanoff, and M. D. Nijs, "Quantum model for commensurate-incommensurate transitions," *Nuclear Physics B*, vol. 215, no. 2, pp. 169 – 208, 1983.

[50] M. Rodney, H. F. Song, S.-S. Lee, K. Le Hur, and E. S. Sørensen, "Scaling of entanglement entropy across lifshitz transitions," *Phys. Rev. B*, vol. 87, p. 115132, Mar 2013.

[51] C. Nayak, S. H. Simon, A. Stern, M. Freedman, and S. D. Sarma, "Non-abelian anyons and topological quantum computation," *Reviews of Modern Physics*, vol. 80, no. 3, p. 1083, 2008.

[52] L. Fu and C. L. Kane, "Superconducting proximity effect and majorana fermions at the surface of a topological insulator," *Physical review letters*, vol. 100, no. 9, p. 096407, 2008.

[53] L. Fu and C. L. Kane, "Probing neutral majorana fermion edge modes with charge transport," *Physical review letters*, vol. 102, no. 21, p. 216403, 2009.

[54] R. M. Lutchyn, J. D. Sau, and S. D. Sarma, "Majorana fermions and a topological phase transition in semiconductor-superconductor heterostructures," *Phys. Rev. Lett.*, vol. 105, no. 7, p. 077001, 2010.

[55] Y. Oreg, G. Refael, and F. von Oppen, "Helical liquids and majorana bound states in quantum wires," *Phys. Rev. Lett.*, vol. 105, no. 17, p. 177002, 2010.

[56] J. Alicea, Y. Oreg, G. Refael, F. von Oppen, and M. P. Fisher, "Non-abelian statistics and topological quantum information processing in 1d wire networks," *Nature Physics*, vol. 7, no. 5, pp. 412–417, 2011.

[57] J. Alicea, "New directions in the pursuit of majorana fermions in solid state systems," *Reports on Progress in Physics*, vol. 75, no. 7, p. 076501, 2012.

[58] V. Mourik, K. Zuo, S. M. Frolov, S. R. Plissard, E. P. A. M. Bakkers, and L. P. Kouwenhoven, "Signatures of majorana fermions in hybrid superconductor-semiconductor nanowire devices," *Science*, vol. 336, no. 6084, pp. 1003–1007, 2012.

[59] L. P. Rokhinson, X. Liu, and J. K. Furdyna, "The fractional ac josephson effect in a semiconductor-superconductor nanowire as a signature of majorana particles," *Nature Physics*, vol. 8, no. 11, pp. 795–799, 2012.

[60] A. Finck, D. Van Harlingen, P. Mohseni, K. Jung, and X. Li, "Anomalous modulation of a zero-bias peak in a hybrid nanowire-superconductor device," *Phys. Rev. Lett.*, vol. 110, no. 12, p. 126406, 2013.

[61] S. Nadj-Perge, I. K. Drozdov, J. Li, H. Chen, S. Jeon, J. Seo, A. H. MacDonald, B. A. Bernevig, and A. Yazdani, "Observation of majorana fermions in ferromagnetic atomic chains on a superconductor," *Science*, vol. 346, no. 6209, pp. 602–607, 2014.

[62] R. S. K. Mong, D. J. Clarke, J. Alicea, N. H. Lindner, P. Fendley, C. Nayak, Y. Oreg, A. Stern, E. Berg, K. Shtengel, and M. P. A. Fisher, "Universal topological quantum computation from a superconductor-abelian quantum hall heterostructure," *Phys. Rev. X*, vol. 4, p. 011036, Mar 2014.

[63] G. Ortiz, E. Cobanera, and Z. Nussinov, "Dualities and the phase diagram of the p-clock model," *Nuclear Physics B*, vol. 854, no. 3, pp. 780 – 814, 2012.

[64] N. H. Lindner, E. Berg, G. Refael, and A. Stern, "Fractionalizing majorana fermions: Non-abelian statistics on the edges of abelian quantum hall states," *Phys. Rev. X*, vol. 2, p. 041002, Oct 2012.

[65] M. Cheng, "Superconducting proximity effect on the edge of fractional topological insulators," *Phys. Rev. B*, vol. 86, p. 195126, Nov 2012.

[66] D. J. Clarke, J. Alicea, and K. Shtengel, "Exotic non-abelian anyons from conventional fractional quantum hall states," *Nature communications*, vol. 4, p. 1348, 2013.

[67] A. Vaezi, "Fractional topological superconductor with fractionalized majorana fermions," *Phys. Rev. B*, vol. 87, p. 035132, Jan 2013.

[68] A. Rapp, P. Schmitteckert, G. Takacs, and G. Zarand, "Asymptotic scattering and duality in the one-dimensional three-state quantum potts model on a lattice," *New Journal of Physics*, vol. 15, no. 1, p. 013058, 2013.

[69] M. B. Hastings, C. Nayak, and Z. Wang, "Metaplectic anyons, majorana zero modes, and their computational power," *Phys. Rev. B*, vol. 87, p. 165421, Apr 2013.

[70] M. Burrello, B. van Heck, and E. Cobanera, "Topological phases in two-dimensional arrays of parafermionic zero modes," *Phys. Rev. B*, vol. 87, p. 195422, May 2013.

[71] R. Bondesan and T. Quella, "Topological and symmetry broken phases of zn parafermions in one dimension," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2013, no. 10, p. P10024, 2013.

[72] M. Barkeshli, C.-M. Jian, and X.-L. Qi, "Theory of defects in abelian topological states," *Phys. Rev. B*, vol. 88, p. 235103, Dec 2013.

[73] M. Barkeshli, Y. Oreg, and X.-L. Qi, "Experimental proposal to detect topological ground state degeneracy," *arXiv preprint arXiv:1401.3750*, 2014.

[74] E. Cobanera and G. Ortiz, "Fock parafermions and self-dual representations of the braid group," *Phys. Rev. A*, vol. 89, p. 012328, Jan 2014.

[75] J. C. Y. Teo and C. L. Kane, "From luttinger liquid to non-abelian quantum hall states," *Phys. Rev. B*, vol. 89, p. 085101, Feb 2014.

[76] Y. Oreg, E. Sela, and A. Stern, "Fractional helical liquids in quantum wires," *Phys. Rev. B*, vol. 89, p. 115402, Mar 2014.

[77] C. P. Orth, R. P. Tiwari, T. Meng, and T. L. Schmidt, "Non-abelian parafermions in time-reversal invariant interacting helical systems," *arXiv preprint arXiv:1405.4353*, 2014.

[78] J. Klinovaja and D. Loss, "Parafermions in an interacting nanowire bundle," *Phys. Rev. Lett.*, vol. 112, p. 246403, Jun 2014.

[79] F. Zhang and C. Kane, "Time-reversal-invariant Z_4 fractional josephson effect," *Phys. Rev. Lett.*, vol. 113, p. 036401, Jul 2014.

[80] W. Li, S. Yang, H.-H. Tu, and M. Cheng, "Criticality in translation-invariant parafermion chains," *Phys. Rev. B*, vol. 91, p. 115133, Mar 2015.

[81] A. Tsvelik, "Z $_n$ parafermion zero modes without fractional quantum hall effect," *arXiv preprint arXiv:1407.4002*, 2014.

[82] J. Klinovaja and D. Loss, "Time-reversal invariant parafermions in interacting rashba nanowires," *Phys. Rev. B*, vol. 90, p. 045118, Jul 2014.

[83] J. Klinovaja, A. Yacoby, and D. Loss, "Kramers pairs of majorana fermions and parafermions in fractional topological insulators," *Phys. Rev. B*, vol. 90, p. 155447, Oct 2014.

[84] R. S. Mong, D. J. Clarke, J. Alicea, N. H. Lindner, and P. Fendley, "Parafermionic conformal field theory on the lattice," *Journal of Physics A: Mathematical and Theoretical*, vol. 47, no. 45, p. 452001, 2014.

[85] J. Alicea and A. Stern, "Designer non-Abelian anyon platforms: from Majorana to Fibonacci," *ArXiv e-prints*, Oct. 2014.

[86] A. Milsted, E. Cobanera, M. Burrello, and G. Ortiz, "Commensurate and incommensurate states of topological quantum matter," *Phys. Rev. B*, vol. 90, p. 195101, Nov 2014.

[87] M. Dalmonte, J. Carrasquilla, L. Taddia, E. Ercolessi, and M. Rigol, "Gap scaling at Berezinskii-Kosterlitz-Thouless quantum critical points in one-dimensional Hubbard and Heisenberg models," *ArXiv e-prints*, Dec. 2014.

[88] A. S. Jermyn, R. S. K. Mong, J. Alicea, and P. Fendley, "Stability of zero modes in parafermion chains," *Phys. Rev. B*, vol. 90, p. 165106, Oct 2014.

[89] D. A. Huse, "Simple three-state model with infinitely many phases," *Phys. Rev. B*, vol. 24, pp. 5180–5194, Nov 1981.

[90] H.-C. Jiang, Z. Wang, and L. Balents, "Identifying topological order by entanglement entropy," *Nature Physics*, vol. 8, no. 12, pp. 902–905, 2012.

[91] P. H. Ginsparg, "Applied conformal field theory," *arXiv preprint hep-th/9108028*, vol. 63, 1988.

[92] H. Au-Yang, B. M. McCoy, J. H. Perk, S. Tang, and M.-L. Yan, "Commuting transfer matrices in the chiral potts models: Solutions of star-triangle equations with genus¿1," *Physics Letters A*, vol. 123, no. 5, pp. 219 – 223, 1987.

[93] G. Albertini, B. M. McCoy, and J. H. Perk, "Commensurate-incommensurate transition in the ground state of the superintegrable chiral potts model," *Physics Letters A*, vol. 135, no. 3, pp. 159 – 166, 1989.

[94] B. M. McCoy and S. shyr Roan, "Excitation spectrum and phase structure of the chiral potts model," *Physics Letters A*, vol. 150, no. 89, pp. 347 – 354, 1990.

[95] H.-D. Chen and J. Hu, "Exact mapping between classical and topological orders in two-dimensional spin systems," *Phys. Rev. B*, vol. 76, p. 193101, Nov 2007.

[96] M. Greiter, V. Schnells, and R. Thomale, "The 1d ising model and the topological phase of the kitaev chain," *Annals of Physics*, vol. 351, no. 0, pp. 1026 – 1033, 2014.

[97] V. L. Pokrovsky and A. L. Talapov, "Ground state, spectrum, and phase diagram of two-dimensional incommensurate crystals," *Phys. Rev. Lett.*, vol. 42, pp. 65–67, Jan 1979.

[98] J. M. Kosterlitz and D. J. Thouless, "Ordering, metastability and phase transitions in two-dimensional systems," *Journal of Physics C: Solid State Physics*, vol. 6, no. 7, p. 1181, 1973.

[99] J. Kosterlitz, "The critical properties of the two-dimensional xy model," *Journal of Physics C: Solid State Physics*, vol. 7, no. 6, p. 1046, 1974.

[100] B. Swingle, J. McMinis, and N. M. Tubman, "Oscillating terms in the renyi entropy of fermi gases and liquids," *Phys. Rev. B*, vol. 87, p. 235112, Jun 2013.

[101] A. Y. Kitaev, "Fault-tolerant quantum computation by anyons," *Annals of Physics*, vol. 303, no. 1, pp. 2–30, 2003.

[102] L. Fu and C. L. Kane, "Topological insulators with inversion symmetry," *Phys. Rev. B*, vol. 76, p. 045302, Jul 2007.

[103] A. M. Turner, Y. Zhang, and A. Vishwanath, "Entanglement and inversion symmetry in topological insulators," *Phys. Rev. B*, vol. 82, p. 241102, Dec 2010.

[104] T. L. Hughes, E. Prodan, and B. A. Bernevig, "Inversion-symmetric topological insulators," *Phys. Rev. B*, vol. 83, p. 245132, Jun 2011.

[105] Y.-M. Lu and D.-H. Lee, "Inversion symmetry protected topological insulators and superconductors," *arXiv preprint arXiv:1403.5558*, 2014.

[106] L. Fidkowski and A. Kitaev, "Effects of interactions on the topological classification of free fermion systems," *Phys. Rev. B*, vol. 81, p. 134509, Apr 2010.

[107] M. F. Lapa, J. C. Teo, and T. L. Hughes, "Interaction-enabled topological crystalline phases," *Physical Review B*, vol. 93, no. 11, p. 115131, 2016.

[108] V. L. Ginzburg and L. D. Landau, "On the theory of superconductivity," *Zh. eksp. teor. Fiz*, vol. 20, no. 1064-1082, p. 35, 1950.

[109] X.-L. Qi, T. L. Hughes, and S.-C. Zhang, "Topological field theory of time-reversal invariant insulators," *Phys. Rev. B*, vol. 78, p. 195424, Nov 2008.

[110] A. Kitaev, "Periodic table for topological insulators and superconductors," in *AIP Conference Proceedings*, vol. 1134, pp. 22–30, AIP, 2009.

[111] E. Prodan, T. L. Hughes, and B. A. Bernevig, "Entanglement spectrum of a disordered topological chern insulator," *Phys. Rev. Lett.*, vol. 105, p. 115501, Sep 2010.

[112] J. Song, C. Fine, and E. Prodan, "Effect of strong disorder on three-dimensional chiral topological insulators: Phase diagrams, maps of the bulk invariant, and existence of topological extended bulk states," *Phys. Rev. B*, vol. 90, p. 184201, Nov 2014.

[113] I. Mondragon-Shem and T. L. Hughes, "Signatures of metal-insulator and topological phase transitions in the entanglement of one-dimensional disordered fermions," *Phys. Rev. B*, vol. 90, p. 104204, Sep 2014.

[114] I. Mondragon-Shem, T. L. Hughes, J. Song, and E. Prodan, "Topological criticality in the chiral-symmetric aiii class at strong disorder," *Phys. Rev. Lett.*, vol. 113, p. 046802, Jul 2014.

[115] J. Song and E. Prodan, "Aiii and bdi topological systems at strong disorder," *Phys. Rev. B*, vol. 89, p. 224203, Jun 2014.

[116] L. Wang, "Discovering phase transitions with unsupervised learning," *Phys. Rev. B*, vol. 94, p. 195105, Nov 2016.

[117] J. Carrasquilla and R. G. Melko, "Machine learning phases of matter," *Nat Phys*, vol. 13, pp. 431–434, May 2017. Letter.

[118] S. J. Wetzel, "Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders," *Phys. Rev. E*, vol. 96, p. 022140, Aug 2017.

[119] W. Hu, R. R. P. Singh, and R. T. Scalettar, "Discovering phases, phase transitions, and crossovers through unsupervised machine learning: A critical examination," *Phys. Rev. E*, vol. 95, p. 062122, Jun 2017.

[120] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst, "Machine learning quantum phases of matter beyond the fermion sign problem," *Scientific Reports*, vol. 7, no. 1, p. 8823, 2017.

[121] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, "Machine learning phases of strongly correlated fermions," *Phys. Rev. X*, vol. 7, p. 031038, Aug 2017.

[122] L. Huang and L. Wang, "Accelerated monte carlo simulations with restricted boltzmann machines," *Phys. Rev. B*, vol. 95, p. 035105, Jan 2017.

[123] Y. Zhang and E.-A. Kim, "Quantum loop topography for machine learning," *Phys. Rev. Lett.*, vol. 118, p. 216401, May 2017.

[124] E. P. L. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, "Learning phase transitions by confusion," *Nat Phys*, vol. 13, pp. 435–439, May 2017. Letter.

[125] Y.-H. Liu and E. P. L. van Nieuwenburg, "Discriminative cooperative networks for detecting phase transitions," *Phys. Rev. Lett.*, vol. 120, p. 176401, Apr 2018.

[126] P. Zhang, H. Shen, and H. Zhai, "Machine learning topological invariants with neural networks," *Phys. Rev. Lett.*, vol. 120, p. 066401, Feb 2018.

[127] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359 – 366, 1989.

[128] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural Networks*, vol. 4, no. 2, pp. 251 – 257, 1991.

[129] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5–32, Oct 2001.

[130] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning*, vol. 1. Springer series in statistics New York, 2001.

[131] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[132] A. P. Schnyder, S. Ryu, and A. W. W. Ludwig, "Lattice model of a three-dimensional topological singlet superconductor with time-reversal symmetry," *Phys. Rev. Lett.*, vol. 102, p. 196804, May 2009.

[133] I. Peschel, "Calculation of reduced density matrices from correlation functions," *Journal of Physics A: Mathematical and General*, vol. 36, no. 14, p. L205, 2003.

[134] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[135] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*. The Wadsworth statistics/probability series, Monterey, CA: Wadsworth and Brooks/Cole Advanced Books and Software, 1984.