NQS Study of Deconfined Quantum Criticality from a U(1) Quantum spin Liquid to a Spinon Bose-Einstein Condensate

Erfan Keshavarz

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Physics

Eastern Mediterranean University
Spring 2025
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Ali Hakan Ulusoy
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science in Physics.

Prof. Dr. İzzet Sakallı
Chair, Department of Physics

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Master of Science in Physics.

Prof. Dr. S. Habib Mazharimousavi
Supervisor

Examining Committee

1. Prof. Dr. Özay GÜRTUĞ
2. Prof. Dr. Mustafa Halilsoy
3. Prof. Dr. S. Habib Mazharimousavi
4. Asst. Prof. Dr. Mustafa Riza

ABSTRACT

Artificial intelligence has significantly impacted modern scientific research, with quantum physics standing at the forefront of these developments. In this thesis, we employ Neural Quantum States (NQS), a variational framework based on neural network ansatze, for ab initio simulations of quantum many-body systems.

Building upon this foundation, we address key challenges in the application of NQS, particularly their efficiency and expressivity in frustrated quantum magnets. Frustrated systems, in which competing interactions hinder classical ordering, are essential for studying emergent quantum phases.

As a case study, we focus on quantum spin ice, a highly frustrated system known to host exotic excitations and a potential U(1) quantum spin liquid phase. By applying NQS to this system, we assess the accuracy, scalability, and limitations of this method in comparison to established approaches, such as variational Monte Carlo and tensor-network-based techniques.

Our results highlight the capability of NQS to capture complex correlations in frustrated magnets, thereby advancing their role in the simulation of strongly correlated quantum matter.

Keywords: Quantum many-body system, Neural quantum state, Frustrated Magnets, Quantum spin ice, U(1) quantum spin liquid.

ÖZ

Yapay zeka, modern bilimsel araştırmalar üzerinde önemli bir etki yaratmış olup, kuantum fiziği bu gelişmelerin ön saflarında yer almaktadır. Bu tezde, nöral ağ ansätzelerine dayalı bir varyasyonel çerçeve olan Nöral Kuantum Durumları (NQS), kuantum çok parçacıklı sistemlerin ab initio simülasyonları için kullanılmaktadır.

Bu temelin üzerine, özellikle frustrasyonlu kuantum mıknatıslar söz konusu olduğunda, NQS'nin verimliliği ve ifade gücü ile ilgili temel zorluklar ele alınmaktadır. Rekabet eden etkileşimlerin klasik düzeni engellediği frustrasyonlu sistemler, ortaya çıkan kuantum fazlarını incelemek için kritik öneme sahiptir.

Vaka çalışması olarak, egzotik uyarılmaların gözlendiği ve potansiyel bir U(1) kuantum spin sıvısı fazı barındırmasıyla bilinen kuantum spin buzu sistemine odaklanıyoruz. NQS'yi bu sistem üzerinde uygulayarak, yöntemin doğruluğu, ölçeklenebilirliği ve sınırlamaları mevcut yöntemler olan varyasyonel Monte Carlo ve tensör ağ tabanlı teknikler ile karşılaştırılmaktadır.

Elde edilen sonuçlar, NQS'nin frustrasyonlu mıknatıslar içindeki karmaşık korelasyonları yakalama yeteneğini ortaya koymakta ve böylece güçlü korelasyonlu kuantum maddelerin simülasyonunda rolünü ileriye taşımaktadır.

Anahtar Kelimeler: Kuantum çok-cisimli sistem, Sinirsel kuantum durumu, Frustrasyonlu manyetikler, Kuantum spin buzu, U(1) kuantum spin sıvısı.

Dedication

I dedicate this thesis to my family and Maryam, whose unwavering love, support, and sacrifices have been the foundation of my academic and personal journey. Their constant encouragement has given me the strength to overcome challenges and the determination to pursue my goals with perseverance. I am deeply grateful for the values they instilled in me, the patience they have shown throughout the years, and their continuous belief in my potential. Without their guidance and support, the completion of this work would not have been possible.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my supervisor, Prof. Dr. S. Habib Mazharimousavi, for their invaluable guidance, continuous support, and encouragement throughout the course of this research. Their expertise and insights have been essential in shaping this work. I am also grateful to the members of my thesis committee for their constructive feedback and helpful suggestions.

I wish to acknowledge the physics department for providing the resources and support necessary for the completion of this thesis.

Finally, I extend my heartfelt appreciation to Asst. Prof. Dr. A. M. Dawid Lekowska and Asst. Prof. Dr. Filippo Vicentini for their invaluable guidance and instruction, from which I have learned a great deal. Their support and insights have significantly enriched my understanding and contributed to the success of this research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI	Artificial Intelegence
ANN	Artificial Neural Network
ARNN	Autoregressive Neural Network
DL	Deep Learning
DQC	Deconfined Quantum Criticality
ED	Exact Diagonalization
FFN	Feed Forward Neural Network
FM	Ferromagnetism
G-CNNs	Group Convolution Neural Network
MCMC	Markove chain Monte Carlo
MF	Mean Field
ML	Machine Learning
MPS	Matrix Product state
NQS	Neural Quantum State
QSI	Quantum Spin Ice
QSL	Quantum Spin Liquid
RBM	Restricted Boltzmann Machine

Chapter 1

INTRODUCTION

The field of machine learning has profoundly transformed scientific research, with physics being no exception. Numerous studies have underscored the pivotal role of machine learning in advancing modern scientific inquiry [ref]. In the realm of quantum science, in particular, machine learning has emerged as an increasingly powerful tool. For instance, in ab initio simulations of quantum many-body systems, researchers often encounter substantial challenges: analytical solutions are typically unattainable, rendering approximation methods essential for investigating these complex systems [1–7].

Despite the development of numerous computational methods, including tensor networks, quantum Monte Carlo methods, and matrix product states (MPS) [8–15], simulating frustrated quantum systems remains challenging. In these systems, competing interactions prevent classical ordering, giving rise to exotic quantum phases. A key open question is whether neural-network-based variational methods, such as NQS, can accurately and efficiently capture the ground-state properties, correlations, and emergent excitations of such frustrated systems.

In computational physics, a diverse array of techniques has been developed to address quantum many-body problems, including tensor networks, quantum Monte Carlo methods, and matrix product states (MPS), among others [8–15]. More recently, with the rapid advancement of machine learning, Carleo and Troyer introduced a novel

approach for addressing quantum many-body systems using a restricted Boltzmann machine, which they termed the Neural Quantum State (NQS) [16]. Since this pioneering work, extensive research has been devoted to exploring the capabilities and limitations of NQS, including outstanding questions regarding its effectiveness in studying frustrated systems [17–22]. This thesis aims to investigate the potential of NQS for studying frustrated systems.

The thesis employs a combination of variational and computational approaches to study frustrated quantum systems. In Chapter 2, we introduce the fundamentals of neural networks, including architecture and backpropagation, and present the NQS method as a variational approach for solving ab initio quantum many-body problems. We also discuss Markov Chain Monte Carlo (MCMC) sampling and stochastic reconfiguration via imaginary time evolution as tools for efficiently optimizing variational parameters. Chapter 3 extends the study to frustrated pyrochlore spin-ice systems, incorporating quantum fluctuations to model quantum spin ice capable of hosting a U(1) quantum spin liquid phase [23–27]. Chapter 4 focuses on the XXZ Hamiltonian on a pyrochlore lattice with 16 spin-½ particles, where the ground state is first obtained using exact diagonalization as a benchmark. The system is then explored using various variational ansätze, including mean-field, Jastrow, and NQS with different neural network architectures. Physical observables, such as the nearest-neighbor correlator, are computed to evaluate the ability of NQS to capture the phase transition from a U(1) quantum spin liquid to XY long-range order at the critical coupling ratio.

The application of NQS demonstrates that neural-network-based variational methods can successfully capture highly entangled and correlated ground states in frustrated

systems. Comparisons with exact diagonalization and other variational methods reveal that NQS can achieve high accuracy and scalability. Analysis of nearest-neighbor correlators shows that NQS is able to detect phase transitions and reproduce key physical features of quantum spin ice, confirming its effectiveness for studying exotic frustrated quantum phases.

Lastly, in Chapter 5, we summarize and discuss the main findings of this thesis, highlighting their significance and potential directions for future research.

Chapter 2

NEURAL NETWORK QUANTUM STATE

In this section, we first outline the fundamental concepts of neural networks, highlighting their underlying principles and the key factors behind their success. Particular attention is given to widely used network architectures that have proven especially effective in quantum physics research (2.1). We then present a detailed overview of the backpropagation algorithm, discussing both its mechanics and its importance. Finally, we examine the structure of automatic differentiation, which is essential for performing efficient gradient computations during neural network training (2.2).

Building on these foundational concepts, the discussion proceeds with a summary of the main branches of machine learning that are relevant to quantum physics, along with a critical examination of the contemporary challenges faced in this interdisciplinary field. Particular attention is devoted to a recent methodological innovation the neural network quantum state (NQS) with an in-depth exploration of its historical development and theoretical foundations (2.4).

In this context, the Markov Chain Monte Carlo (MCMC) sampling method, which plays a central role in the implementation of NQS, is examined in detail. In addition, this section discusses stochastic reconfiguration, an advanced optimization technique developed to overcome the challenges associated with NQS (2.5).

2.1 Introduction to Neural Networks

Artificial Neural Networks (ANNs) are machine learning models composed of multiple layers of interconnected nodes, commonly referred to as neurons. Each neuron receives input in the form of a weighted sum of the input vector and a bias term, after which an activation function is applied to this linear combination to generate a scalar output. Inspired by the structure of the human brain, ANNs are widely used for data processing tasks. As noted above, an individual neuron takes a vector as input and produces a scalar output, as illustrated below.

$$a_i = \zeta(Z), \tag{2.1}$$

$$Z = \sum_j W_{ij} x_j + b_i. \tag{2.2}$$

ζ denotes the activation function (a nonlinear function). Common examples include the hyperbolic tangent, sigmoid, and rectified linear unit (ReLU).

A neural network generally consists of three types of layers: an input layer, one or more hidden layers, and an output layer. The overall design of these layers is referred to as the network architecture. For example, when each neuron is connected to all neurons in the preceding layer, the structure is called a fully connected network. Beyond this, several specialized architectures have been developed, including Convolutional Neural Networks, Autoencoders, Variational Autoencoders, Autoregressive Neural Networks, Recurrent Neural Networks, Transformer Networks, and Graph Neural Networks, among others. A network is described as deep when it contains multiple hidden layers, and its study falls under the field of deep learning. Importantly, Kolmogorov and Arnold demonstrated that any function can be represented through a polynomial number of compositions of nonlinear

one-dimensional functions, as illustrated below

$$f(x) = \sum_{i=0}^{2m} \xi_i \left(\sum_{j=1}^{m} \zeta_{ij}(x_j) \right). \tag{2.3}$$

This shows similarity with a neural network with two hidden layers. Furthermore, nonlinear functions are essential for representing complex functions. George Cybenko stated that A feedforward neural network with at least one hidden layer and a nonlinear activation function can approximate any continuous function. This theory leads to universal approximation theory and is furthermore generalized by Kurt Hornik. Neural networks are trained using gradient-based optimization methods, which work to minimize the chosen loss function. The method of computing the gradient is backpropagation, which is introduced in the subsequent section.

2.2 Backpropagation

In order to train a Neural network, one should use gradient-based methods, which require computing derivatives of the loss function. Backpropagation is an algorithm based on automatic differentiation, designed to efficiently and systematically compute gradients. It uses the chain rule in the reverse direction of the network's computation flow. To aid in understanding, an illustration of backpropagation applied to a simple feedforward network is presented below, noting that the same principle extends to other neural network architectures. The activation of the neurons in the layer l of an FFN is given by

$$a^l = \zeta^l(W^l a^{l-1} + b^l). \tag{2.4}$$

a^{l-1} represents the vector that holds the activations from the preceding layer. W^l represents the weight matrix connecting layer $l-1$ to layer l, b^l is the bias term for layer l, and ζ^l corresponds to the activation function in layer l. A neural network with

l layers and $l-1$ hidden layers is given by

$$NN(x) = a^l(x) = \zeta^l(W^l a^{l-1} + b^l), \tag{2.5}$$

For a training dataset $D = (x_i, y_i)_{i=1}^n$, the loss function can be expressed as

$$\mathscr{L} = \frac{1}{n}\sum_{i=1}^{n} l(NN(x_i), y_i). \tag{2.6}$$

where l represents the deviation between the prediction $NN(x_i)$ and the desired output y_i. The primary objective is to compute the derivatives of the loss function with respect to all network parameters (weights and biases) in order to minimize the loss and optimize these parameters for improved performance on unseen data. Accordingly, the derivative of the loss function with respect to its weights is given by

$$\frac{\partial \mathscr{L}}{\partial w} = \frac{1}{n}\sum_{i=1}^{n} \frac{\partial l(NN(x_i), y_i)}{\partial w}, \tag{2.7}$$

Where w is a single weight of the NN

$$\frac{\partial l(NN(x_i), y_i)}{\partial w} = \frac{\partial l(NN(x_i), y_i)}{\partial NN} . \frac{\partial NN(x_i)}{\partial w} = \frac{\partial l(a^L(x_i), y_i)}{\partial a^L} \frac{\partial a^L(x_i)}{\partial w}. \tag{2.8}$$

The first term can be calculated for a specified loss function. For instance, in the case of MSE loss

$$l_{MSE}(NN(x_i), y_i) = ||NN(x_i) - y_i||^2, \tag{2.9}$$

ergo,

$$\frac{\partial l_{MSE}(NN(x_i), y_i)}{\partial NN} = 2(NN(x_i) - y_i), \tag{2.10}$$

The central derivative of interest is

$$\frac{\partial NN(x)}{\partial w} = \frac{\partial a^L}{\partial w}. \tag{2.11}$$

A suggestive way of computation is through the chain rule. Since a forward neural network processes information layer-by-layer, the weight in the layer l influences the loss through the next layer $l+1$. Starting from the last layer l, using the chain rule,

$$\frac{\partial a^L}{\partial w} = \frac{\partial a^L}{\partial Z^L}\frac{\partial Z^L}{\partial w} = \frac{\partial a^L}{\partial z^L}(\frac{\partial W^L}{\partial w}a^{L-1} + W^L\frac{\partial a^{L-1}}{\partial w}), \tag{2.12}$$

where, $\frac{\partial a^L}{\partial Z^L} = J^L_\xi$ denotes the Jacobian of ζ^L, whose entries are the derivatives of the activation functions, as shown below

$$(J^L_\zeta)_{ij} = \frac{\partial \zeta^L_i}{\partial z_j}, \tag{2.13}$$

If w denotes a weight in layer L, the result can be expressed as follows

$$\frac{\partial a^L}{\partial w} = J^L_\zeta e^L_i a^{L-1}_j, \tag{2.14}$$

e^L_i is the activation vector of the layer L where all neuron activations are zero except for the i-th neuron, which has an activation of one. Otherwise, all entries remain zero

$$\frac{\partial a^L}{\partial w} = J^L_\zeta W^L\frac{\partial a^{L-1}}{\partial w}, \tag{2.15}$$

subsequently,

$$\frac{\partial a^{L-1}}{\partial w} = \frac{\partial a^{L-1}}{\partial Z^{L-1}}\frac{\partial Z^{L-1}}{\partial w} = J^{L-1}_\zeta(\frac{\partial W^{L-1}}{\partial w}a^{L-2} + W^{L-1}\frac{\partial a^{L-2}}{\partial w}), \tag{2.16}$$

If w denotes a weight in layer $L-1$, the result can be expressed as follows

$$\frac{\partial a^{L-1}}{\partial w} = J^{L-1}_\zeta e^{L-1}_i a^{L-2}_j, \tag{2.17}$$

Otherwise,

$$\frac{\partial a^{L-1}}{\partial w} = J_{\zeta}^{L-1} W^{L-1} \frac{\partial a^{L-2}}{\partial w}, \tag{2.18}$$

Since the computation of the above derivative is recursive, the following result is expected

$$\frac{\partial a^{L}}{\partial w} = J_{\zeta}^{L} W^{L} J_{\zeta}^{L-1} W^{L-1} ... J_{\zeta}^{L'+1} W^{L'+1} \frac{\partial a^{L'}}{\partial w}, \tag{2.19}$$

As observed from the above equation, w does not correspond to the weight of layer L through L'.

Otherwise,

$$\frac{\partial a^{L}}{\partial w} = J_{\zeta}^{L} W^{L} J_{\zeta}^{L-1} W^{L-1} ... J_{\zeta}^{L'+1} W^{L'+1} J_{\zeta}^{L'} e_i^{L'} a_j^{L'-1}, \tag{2.20}$$

To compute the backpropagation method, the following deviation from the output layer is needed

$$\Delta^{L} = \frac{\partial l(a^{L}(x_i), y_i)}{\partial a^{L}} \odot J_{\zeta}^{L}, \tag{2.21}$$

For the case of MSE as a loss function, deviation leads to

$$\Delta^{L} = 2(a^{L}(x_i), y_i) \odot J_{\zeta}^{L}, \tag{2.22}$$

Substituting equations (2.22) and (2.14) into equation (2.8) yields the following expression

$$\frac{\partial \mathscr{L}}{\partial w} = \Delta^{L} e_i^{L} a_j^{L-1}, \tag{2.23}$$

Furthermore, by taking the summation over all the components

$$\frac{\partial l(a^L(x_i), y_i)}{\partial w_{ij}^L} = \sum_k (\Delta^L e_i^L a_j^{L-1})_k. \tag{2.24}$$

After computing the derivatives with respect to all weights of the neural network, we move one layer back. The connection between layers is illustrated in Equations (2.15), (2.19), and (2.20). This backward passing is the reason for calling it the backpropagation algorithm. This process continues until it reaches the first layer. In the subsequent Section, a review of machine learning in quantum physics is given.

2.3 Machine Learning for Quantum Physics

The advancement in the field of artificial intelligence and machine learning has revolutionized the technological world. Additionally, it has significantly influenced the realm of physics, especially in the fields of particle physics, condensed matter physics, and astronomy. When it comes to quantum physics, we can typically divide the subject into four main branches, which are given below:

- Ab-initio Quantum Simulation: This involves modeling quantum systems based on the fundamental principles of physics with the aid of machine learning methods, without depending on experimental data. Machine learning assists in simulating intricate systems like quantum many-body systems and extracting interpretable information from our models.
- Analysis and Detection for Quantum Data: In experimental physics, quantum systems generate extensive datasets, which can be leveraged through machine learning for applications like pattern recognition, signal detection within our quantum systems, and gaining insights from the noise present in our experiments.

- Optimal Control for Quantum Devices:In this area, we can utilize machine learning techniques to create and refine control protocols that improve the performance of quantum hardware, like qubits in quantum computers, enhancing both the precision and stability of these devices.
- Optimization of Quantum Algorithms: Similar to the earlier branch, we can utilize machine learning to enhance quantum hardware. In this context, machine learning can be applied to refine and optimize quantum algorithms, enabling more efficient solutions to computational problems.

  In our research, we primarily concentrate on the first aspect, which is the ab-initio quantum simulation of a quantum many-body system. We commence this section by defining what constitutes a quantum many-body system and outlining the key pathway for ab-initio simulation, as well as the challenges we face during this process.

2.3.1 Quantum Many Body System

A quantum many-body system consists of an exponentially large number of interacting particles, such as spins, electrons, and atoms. These interactions give rise to complex phenomena, including correlations and entanglement, among many other exotic effects. The initial step in simulating such systems is to solve the Schrödinger equation to determine the ground state, excited states, and their corresponding energies. It has been observed that for many real-world quantum scenarios involving a few interacting particles, analytical solutions to the Schrödinger equation do not exist. Consequently, we rely on approximation methods, particularly for many-body systems, which is the focus of our discussion. Approximation techniques such as mean field theory, tensor networks, matrix product states (MPS), and variational methods are crucial for understanding these systems.

As our work is exclusively centered on spin systems, we will now solely concentrate on these systems. From quantum physics, we understand that all information about our system is contained in a complex function called the wave function. Investigating single spin-1/2 systems can be advantageous for enhancing comprehension for educational purposes. The wave function on the computational basis Z is expressed as follows

$$|\Psi\rangle = C_{\uparrow}|\uparrow\rangle + C_{\downarrow}|\downarrow\rangle . \tag{2.25}$$

where $C_{\uparrow}$ and $C_{\downarrow}$ represent the probability amplitudes for the system to align in either of the computational basis, and the system must also adhere to the Born probability amplitude, which is

$$|C_{\uparrow}|^2 + |C_{\downarrow}|^2 = 1. \tag{2.26}$$

Moreover, to calculate each probability amplitude, we simply need to find the dot product of the state with each computational basis, as demonstrated below

$$C_{\uparrow} = \langle\uparrow|\Psi\rangle , \tag{2.27}$$

$$C_{\downarrow} = \langle\downarrow|\Psi\rangle . \tag{2.28}$$

Let's now transition to a quantum many-body system composed of N spin-$1/2$ particles. The value of N can be on the order of Avogadro's number, which is approximately 10^{23}. As a result, the total count of probability amplitudes is 2^N, which matches the dimension of the Hilbert space. In a manner comparable to the previous example, the wavefunction in this system can be represented as

$$
\begin{aligned}
|\Psi\rangle &= C_{\uparrow\uparrow\ldots\uparrow}|\uparrow\uparrow\ldots\uparrow\rangle + C_{\uparrow\uparrow\ldots\downarrow}|\uparrow\uparrow\ldots\downarrow\rangle + \cdots + C_{\downarrow\downarrow\ldots\downarrow}|\downarrow\downarrow\ldots\downarrow\rangle \\
&= \sum_{\phi_1,\phi_2,\ldots,\phi_N} C_{\phi_1,\phi_2,\ldots,\phi_N}\ |\phi_1\rangle \otimes |\phi_2\rangle \otimes \cdots \otimes |\phi_N\rangle\,.
\end{aligned}
\tag{2.29}
$$

By defining $|\sigma\rangle = |\phi_1\rangle \otimes |\phi_2\rangle \otimes \cdots \otimes |\phi_N\rangle$ and setting $C_{\phi_1,\phi_2,\ldots,\phi_N} = \psi(\sigma)$, equation (2.29) transforms into

$$
|\Psi\rangle = \sum_{\sigma=1}^{2^N} \psi(\sigma)\,|\sigma\rangle\,. \tag{2.30}
$$

The initial significant challenge in performing computations on quantum many-body systems stems from the exponential size of the Hilbert space, which contains 2^n distinct configurations. To represent a random state, we are required to store 2^n complex numbers in our computational device. The memory required to store a complex number using double precision is 16 bytes, with 8 bytes allocated for the real part and another 8 bytes for the imaginary part. The table below details the memory requirements for storing the wavefunction for different numbers of spins

Spin	Hilbert size	Memory Cost
20	10^6	16 MB
30	10^9	16 GB
40	10^{12}	16 TB
60	10^{15}	16 EB

Table 2.1: the memory requirement for storing the wave function associated with their Hilbert space size

Based on the information in Table (2.1), storing the wave function for 60 spins requires approximately 16 exabytes of memory, which is 500 times more than the most powerful supercomputer currently available. This problem is referred to as memory complexity. Additionally, to calculate the expectation value of the

Hamiltonian, another challenge emerges known as computational complexity, since the Hamiltonian matrix is now of the order of $2^N \times 2^N$, as shown below

$$\langle \hat{H} \rangle = \sum_{\sigma,\eta} \psi^*(\sigma) H_{\sigma,\eta} \psi(\eta). \tag{2.31}$$

As indicated by equation (2.31), if it were possible to store all the probability amplitudes to calculate the expectation value, the runtime would be affected by the enormous size of the Hilbert space, which is referred to as computational complexity. In simple terms, a quantum many-body system can be likened to a coin with two sides: one representing memory complexity and the other representing runtime complexity. In the following section, a method will be proposed to address memory complexity.

2.3.2 Variational Ansätze

The Hilbert space of a quantum many-body system has an exponentially large size; however, a physically relevant state is typically concentrated at a corner of the Hilbert space, particularly for Hamiltonians with local interactions, such as systems with nearest-neighbor interactions. Consequently, the central concept of the variational method is to discover an efficient representation of these physically meaningful states within our Hilbert space. As a result, rather than storing the complete wavefunction, we store a parameter vector $\vec{\theta}$ that lives within the variational space, which is a submanifold of the real Hilbert space, and these parameters are polynomial in dimension. These parameters establish a function within the variational space that can calculate the amplitude for any given configuration instantly, effectively acting as a smart and compact representation of the complete wavefunction. The variational wave function is represented below

$$|\Psi_{\vec{\theta}}\rangle = \sum_{\sigma=1}^{2^N} \psi_{\vec{\theta}}(\sigma) \, |\sigma\rangle \, . \tag{2.32}$$

It is worth noting that the parameterized wave function will not be stored; rather, one would compute the probability amplitude for a specific configuration on demand. As a result, the objective will be to identify the optimal parametrization that represents the quantum state of a system based on a specified Hamiltonian.

Given that the issue of memory complexity has been addressed, it is time to explore the expectation value of an operator like the Hamiltonian. Consequently, one can calculate the expectation value in the following manner

$$\frac{\langle \Psi_{\vec{\theta}} | H | \Psi_{\vec{\theta}} \rangle}{\langle \Psi_{\vec{\theta}} | \Psi_{\vec{\theta}} \rangle} = \sum_{\sigma,\eta} \frac{\langle \Psi_{\vec{\theta}} | \sigma \rangle \langle \sigma | H | \eta \rangle \langle \eta | \Psi_{\vec{\theta}} \rangle}{\sum_{\sigma} \langle \Psi_{\vec{\theta}} | \sigma \rangle \langle \sigma | \Psi_{\vec{\theta}} \rangle}. \tag{2.33}$$

As indicated by equation (2.33), while $\langle \Psi_{\vec{\theta}} | \sigma \rangle$ can be determined using variational Ansätze, the issue of runtime has not been resolved yet. To summarize this section, we have mentioned that variational Ansätze have been introduced to address memory complexity. There are various examples of variational Ansätze, including mean field Ansätze, matrix product Ansätze, Jastrow Ansätze, and many others. In the following section, we will concentrate on a particular variational Ansatz referred to as the neural quantum state.

2.4 Neural Network Quantum State

In their pioneering study in 2017, Giuseppe Carleo and Matthias Troyer proposed that variational Ansätze could be viewed as the exponential output of a neural network

$$\Psi_{\vec{\theta}} = \exp\left(NN_{\vec{\theta}}(\sigma)\right). \tag{2.34}$$

where $\left(NN_{\vec{\theta}}\right)$ represents the output generated by the neural network for a specific configuration on demand. As previously highlighted in section (2.1), neural networks serve as universal function approximators. Thus, the concept is that when a particular

configuration is input into the neural network, it produces the associated probability amplitude as the output. The weights and biases of the neural network will be stored as parameters. Additionally, these parameters are optimized by minimizing a loss function, which we will introduce in the following section; this loss function is essentially energy as a function of these parameters, similar to the approach used in variational methods in quantum physics.

To address the issue of runtime complexity, it is proposed that rather than summing across the entire Hilbert space, one should only sum over a carefully chosen set of configurations within the Hilbert space. This approach effectively involves utilizing Monte Carlo sampling to identify the most significant configurations. Consequently, evaluating the energy transforms into an energy estimator or local estimator based on specific configurations. As a result, one can achieve the following outcome

$$\frac{\langle \Psi_{\vec{\theta}} | H | \Psi_{\vec{\theta}} \rangle}{\langle \Psi_{\vec{\theta}} | \Psi_{\vec{\theta}} \rangle} = \sum_{\sigma} \frac{\langle \Psi_{\vec{\theta}} | \sigma \rangle \langle \sigma | H | \Psi_{\vec{\theta}} \rangle}{\langle \Psi_{\vec{\theta}} | \Psi_{\vec{\theta}} \rangle}, \tag{2.35}$$

By multiplying equation (2.35) by $\frac{\langle \sigma | \Psi_{\vec{\theta}} \rangle}{\langle \sigma | \Psi_{\vec{\theta}} \rangle}$, the following outcome can be achieved

$$\frac{\langle \Psi_{\vec{\theta}} | H | \Psi_{\vec{\theta}} \rangle}{\langle \Psi_{\vec{\theta}} | \Psi_{\vec{\theta}} \rangle}, = \sum_{\sigma} \frac{|\langle \Psi_{\vec{\theta}} | \sigma \rangle|^2}{\langle \Psi_{\vec{\theta}} | \Psi_{\vec{\theta}} \rangle} \frac{\langle \sigma | H | \Psi_{\vec{\theta}} \rangle}{\langle \sigma | \Psi_{\vec{\theta}} \rangle}. \tag{2.36}$$

It is important to mention that in equation (2.36), there is a possibility for $\langle \sigma | \Psi_{\vec{\theta}} \rangle$ to equal zero. This situation can lead to catastrophic outcomes; however, this problem can be easily resolved by considering that the summation in equation (2.36) can be divided into two parts: one summation over configurations where the probability amplitude is zero, which consequently results in that term being zero, and another summation over non-zero probability amplitudes. Thus, equation (2.36) transforms into

$$\frac{\langle \Psi_{\vec{\theta}} | H | \Psi_{\vec{\theta}} \rangle}{\langle \Psi_{\vec{\theta}} | \Psi_{\vec{\theta}} \rangle}, = \sum_{\sigma | \psi_{\vec{\theta}}(\sigma) \neq 0} \frac{|\langle \Psi_{\vec{\theta}} | \sigma \rangle|^2}{\langle \Psi_{\vec{\theta}} | \Psi_{\vec{\theta}} \rangle} \frac{\langle \sigma | H | \Psi_{\vec{\theta}} \rangle}{\langle \sigma | \Psi_{\vec{\theta}} \rangle}. \tag{2.37}$$

It is clear that the initial term signifies the Born probability amplitude, while the subsequent term refers to the local energy or local estimator, as illustrated below.

$$P_\theta(\sigma) = \frac{|\langle \Psi_{\vec{\theta}} | \sigma \rangle|^2}{\langle \Psi_{\vec{\theta}} | \Psi_{\vec{\theta}} \rangle}, \tag{2.38}$$

Additionally,

$$\sum_\sigma P_\theta(\sigma) = 1. \tag{2.39}$$

and

$$H_\theta^{loc}(\sigma) = \frac{\langle \sigma | H | \Psi_{\vec{\theta}} \rangle}{\langle \sigma | \Psi_{\vec{\theta}} \rangle}. \tag{2.40}$$

As a result, equation (2.13) can be rewritten as follows

$$\frac{\langle \Psi_{\vec{\theta}} | H | \Psi_{\vec{\theta}} \rangle}{\langle \Psi_{\vec{\theta}} | \Psi_{\vec{\theta}} \rangle} = \sum_\sigma P_\theta(\sigma) H_\theta^{loc}(\sigma). \tag{2.41}$$

The presented equation corresponds to the expectation value of a local estimator computed with respect to the probability distribution; therefore, it can be formally expressed as

$$\sum_\sigma P_\theta(\sigma) H_\theta^{\mathrm{loc}}(\sigma) = \mathbb{E}_{\sigma \sim P_\theta(\sigma)} \left[H_\theta^{\mathrm{loc}}(\sigma) \right]. \tag{2.42}$$

To grasp the concept better, one might imagine a classroom where the objective is to calculate the average age of the students. Instead of asking each person individually, one can simply select a few representatives and calculate the average from that sample. In this quantum system, one can draw configurations from the probability distribution based on the Born amplitude, then compute a local estimator for those configurations, which effectively indicates how much each configuration contributes to the energy. As

a result, it is possible to obtain

$$\mathbb{E}_{\sigma\sim P_\theta(\sigma)}\left[H_\theta^{\mathrm{loc}}(\sigma)\right] \approx \frac{1}{N_s}\sum_{i=1}^{N_s} H_\theta^{\mathrm{loc}}(\sigma_i). \tag{2.43}$$

began with a summation over an exponentially large number of entries and ultimately reduced it to a summation involving only a few entries; this approach is effective and polynomial if the local estimator is also polynomial in time.

$$\frac{\langle\sigma|H|\Psi_{\vec{\theta}}\rangle}{\langle\sigma|\Psi_{\vec{\theta}}\rangle} = \sum_\eta \langle\sigma|H|\eta\rangle \frac{\langle\eta|\Psi_\theta\rangle}{\langle\sigma|\Psi_\theta\rangle}. \tag{2.44}$$

In the equation mentioned above, there are no issues related to memory or runtime complexity for $\psi_\theta(\sigma)$ or $\psi_\theta(\eta)$, as they are merely the outputs produced by the neural network. but how about the $\langle\sigma|H|\eta\rangle$, Quantum many-body problems The Hamiltonian acts on a large Hilbert space, but local terms like spin spin interaction only affect a small number of configurations, resulting in a log-sparse Hamiltonian. On the other hand, the number of non-zero entries grows logarithmically with the size of the Hilbert space. For example, if the Hamiltonian has a dimension of 2^N, there will only be about N non-zero entries per row or column. These are what we regard as connected entries, which in a sense means what are the configurations the H links to it in such a manner that $H_{\sigma\eta}$ is not zero, and in general they are polynomially many as long as Hamiltonian have local interaction. Now, one may question how to produce samples selected from a distribution, assuming the Neural Network is completely arbitrary.

2.5 Markov Chain Monte Carlo Sampling

In this section, a sampling method known as Markov Chain Monte Carlo sampling with a Metropolis-Hastings acceptance rule is introduced, generating several samples drawn from a probability distribution. This acceptance rule constructsa Markovian stochastic process that fulfills the detailed balance condition with respect to the target

probability distribution, as given below

$$P(\sigma^t)T(\sigma^{t+1}|\sigma^t) = P(\sigma^{t+1})T(\sigma^t|\sigma^{t+1}) \tag{2.45}$$

This essentially tells that the probability of taking one random configuration from the probability distribution and transforming it to the next configuration should be equal to the probability of having that next configuration and transforming it back. Therefore, the detailed condition rule can be written as a ratio, which is shown below

$$\frac{T(\sigma^{t+1}|\sigma^t)}{T(\sigma^t|\sigma^{t+1})} = \frac{P(\sigma^{t+1})}{P(\sigma^t)}. \tag{2.46}$$

Furthermore, the primary task is to devise a systematic method for constructing this chain of configurations, which means that a random configuration at time t can generate the one at time t+1. Therefore, by defining the transition function as below, which is the product of two functions, the purposing configuration function as g and the accepting configuration function as A

$$T(\sigma^{t+1}|\sigma^t) = g(\sigma^{t+1}|\sigma^t)A(\sigma^{t+1}|\sigma^t), \tag{2.47}$$

by putting equation (2.47) in (2.46) the following result can be obtained

$$\frac{A(\sigma^{t+1}|\sigma^t)}{A(\sigma^t|\sigma^{t+1})} = \frac{P(\sigma^{t+1})}{P(\sigma^t)}\frac{g(\sigma^{t+1}|\sigma^t)}{g(\sigma^t|\sigma^{t+1})}. \tag{2.48}$$

There are several options for defining the acceptance function, which already exist in the literature, one of which can be the following function

$$A(\sigma^{t+1}|\sigma^t) = \min\left[1, \frac{P(\sigma^{t+1})}{P(\sigma^t)} \cdot \frac{g(\sigma^{t+1}|\sigma^t)}{g(\sigma^t|\sigma^{t+1})}\right]. \tag{2.49}$$

For the case of a purposing function, it can be as simple as flipping one spin, but some

systems are different. For instance, for a highly correlated system, just flipping one spin can give a correlated sample, so we need to do it many times for decorrelation. Also, if the Hilbert space is constrained, then the transition should satisfy this constraint.

In order to summarize this section, the initial step was to have a completely random configuration given according to a probability distribution. Furthermore, by using g, a new configuration is proposed. By using the acceptance rule, we can accept it with the probability of A by doing these steps many times, a chain of configurations is generated according to the probability distribution. A key point here is that in the previous section, the normalization constant had no analytical expression. Now, in here, one can easily see that these normalization constants are canceled out, and that's why the MCMC method is working well for the system.

2.6 Stochastic Reconfiguration

In this section, we start by finding the ground state energy and introduce a method called stochastic reconfiguration through imaginary time evolution for optimization purposes.

2.6.1 Probing the ground energy landscape

Normal in any ML task to find the best parameter; a loss function is always needed to minimize it for finding the best optimized parameter. For the case of the variational method, which is discussed here, the suggested loss function is the variational energy defined below

$$\mathscr{L}(\theta) = E(\theta) = \frac{\langle \Psi_\theta | H | \Psi_\theta \rangle}{\langle \Psi_\theta | \Psi_\theta \rangle}. \tag{2.50}$$

Whereas the variational principle in quantum physics states that variational energy is

always equal to or greater than the ground state energy, which needs to be minimized, by using a gradient-based optimization algorithm selected to update the model parameters, as demonstrated below

$$\theta_{i+1} = \theta_i - \eta \nabla E_{\theta_i}. \tag{2.51}$$

This method simply starts with initializing the Neural network with some random parameters and looking at how the state overlaps with all the excited states of the Hamiltonian, which essentially shows completely random overlap with different eigenstates, then starting by computing the gradient to increase the overlap with the ground state and diminish the overlap with higher energy states. There are several issues to be addressed in the optimization problem, as one of them is what happens if, during the optimization, one gets stuck at a local minimum. This issue can be addressed by stating that high-dimensional functions with so many parameters are highly unlikely to have local minima. Another problem comes from the saddle point, where it is a massive issue. For a better grasp of this concept, starting from the definition of the gradient of energy can be helpful. θ can be either complex or real, where in this case considering complex θ will result in becoming holomorphic with respect to *theta*, which indicates two important notes, as illustrated below.

- $\partial_{\theta^*} \Psi_\theta(\sigma) = 0$
- Having a complex parameter indicates that to minimize the loss function, one needs to follow the conjugate gradient, not the gradient itself

One can simply get the definition of the conjugate gradient as follows

$$\partial_{\theta^*} E(\theta) = \mathbb{E}_{\sigma \sim P_\theta(\sigma)} \left[\partial_{\theta^*} \log \psi_\theta(\sigma) \left(H_\theta^{\mathrm{loc}}(\sigma) - \mathbb{E}_{\sigma \sim P_\theta(\sigma)} \left[H_\theta^{\mathrm{loc}}(\sigma) \right] \right) \right]. \tag{2.52}$$

The important feature in this covariant form shown above is that the moment one reaches an eigenstate, it indicates that the gradient is zero and, therefore, the optimization will stop. In a simple word, the whole idea is to compress the wave function down into some variational neural network and then optimize those parameters to minimize the energy and hence reach the ground-state energy.

As mentioned before, to compute the estimator efficiently, one should use MCMC sampling; therefore, the conjugate gradient becomes

$$\nabla_\theta = \frac{1}{N_s} \sum_{k=1}^{N_s} \partial_{\theta^*} \log(\psi_\theta(\sigma)) \Delta H^{\mathrm{loc}}(\sigma_k), \tag{2.53}$$

where

$$\Delta H^{\mathrm{loc}}(\sigma_k) = H_\theta^{\mathrm{loc}}(\sigma) - \frac{1}{N_s} \sum_{k=1}^{N_s} H_\theta^{\mathrm{loc}}(\sigma_k). \tag{2.54}$$

From this point on, the learning process is straightforward: First, initialize the system with complex random parameters. Then, at each step, sample several configurations according to Born's probability distribution using the MCMC method with the Metropolis–Hastings acceptance rule. Next, determine the local estimator, which provides an estimate of the expectation value of the Hamiltonian, and compute the conjugate gradient from the above equation. The final step is to apply gradient-based optimization to update the parameters. This procedure is repeated until convergence to a minimum in the energy landscape is achieved. These algorithms can cause some issues, for instance, what will happen if the gradient becomes zero at a local minimum? In the subsequent section, an imaginary time evolution representation is introduced to enhance the optimization challenges.

2.6.2 Stochastic reconfiguration from Imaginary Time Evolution Dynamic

The energy is quadratic with respect to the wave function, which implies that there are no local minima. This suggests that, if it were possible to take the gradient with respect to the wave function (which, unfortunately, is not likely), an algorithm following gradient descent could perform the optimization efficiently. Afterwards, by simply mapping the result back to the variational space, the problem would be resolved. This approach is called stochastic reconfiguration, or, in machine learning, natural gradient descent. In theory, this method guarantees convergence. Stochastic reconfiguration can be derived from the imaginary-time evolution of the variational ansatz

$$|\Psi_{\theta(\tau)}\rangle = e^{-H\tau}\,|\Psi_{\theta(0)}\rangle\,. \tag{2.55}$$

This shows that the parameters are time-dependent, and $|\Psi_{\theta(0)}\rangle$ is merely a random NQS initial state. By considering $\tau << 1$, the following result can be calculated

$$|\Psi_{\theta(\tau)}\rangle \approx (\mathbf{1} - \tau H)\,|\Psi_{\theta(0)}\rangle + \mathscr{O}(\tau^2). \tag{2.56}$$

The main goal is to remap the dynamics from the Hilbert space, which is unbearably large, to a differential equation in the variational space. This is achieved by finding the update rule for a single time step as follows

$$|\Psi_{\theta+\dot{\theta}\tau}\rangle \approx |\Psi_\theta\rangle + \tau\dot{\theta}.\partial_\theta\,|\Psi_\theta\rangle + \mathscr{O}(\tau^2). \tag{2.57}$$

Note that this time evolution step changes the norm. Therefore, the centered tangent vectors are

$$|\Psi_t\rangle = |\partial_{\theta_j}\Psi_\theta\rangle - \frac{\langle\Psi_\theta|\partial_{\theta_j}\Psi_\theta\rangle}{\langle\Psi_\theta|\Psi_\theta\rangle}\,|\Psi_\theta\rangle\,. \tag{2.58}$$

This ensure that $|\Psi_t\rangle$ is orthogonal to $|\Psi_\theta\rangle$. Therefore equation (2.55) reads

$$|\Psi_{\theta+\dot{\theta}\tau}\rangle \approx |\Psi_\theta\rangle + \tau\dot{\theta}_j |\Psi_j\rangle . \tag{2.59}$$

The following objective is to determine the distance between the two states given in (2.54) and (2.57), by the choice of $\dot{\theta}$ in such a way that this distance is minimized, therefore,

$$\mathscr{L}(\dot{\theta}) = \mathscr{D}(e^{-H\tau}|\Psi_\theta\rangle, |\Psi_{\theta+\dot{\theta}\tau}\rangle). \tag{2.60}$$

Here $\mathscr{D}$ is the distance between two corresponding states in the Hilbert space. For instance, to compute the distance of a vector in a Euclidean space, one can easily find the following

$$\mathscr{D}(|x+\varepsilon\rangle, |x\rangle) = \varepsilon^\dagger \mathbf{1}\varepsilon. \tag{2.61}$$

The metric tensor is the identity. Moreover, for the case of quantum mechanical distance

$$\mathscr{D}(|\Psi_{\theta+\delta\theta}\rangle, |\Psi_\theta\rangle) = \delta\theta^\dagger \mathbf{S}\delta\theta. \tag{2.62}$$

Where $\mathbf{S}$ is the metric tensor and is not the identity, since the distance here is not the Euclidean distance that is embedded in the Euclidean manifold of parameters, this quantum mechanical distance is the Fubini-Study distance, which is just a physically motivated distance and $\mathbf{S}$ known as the quantum geometric tensor given by the Fubini-Study metric tensor of the manifold on which a variational state is defined. Lastly by minimizing (2.60) one can get

$$\mathbf{S}\dot{\theta} = -\mathbf{F}, \tag{2.63}$$

Where,

$$\mathbf{S_{ij}} = \frac{\langle \partial_{\theta_i} \Psi_\theta | \partial_{\theta_j} \Psi_\theta \rangle}{\langle \Psi_\theta | \Psi_\theta \rangle} - \frac{\langle \partial_{\theta_i} \Psi_\theta | \Psi_\theta \rangle \langle \Psi_\theta | \partial_{\theta_j} \Psi_\theta \rangle}{\langle \Psi_\theta | \Psi_\theta \rangle^2}, \tag{2.64}$$

$$\mathbf{F_j} = \frac{\langle \Psi_\theta | H | \partial_{\theta_j} \Psi_\theta \rangle}{\langle \Psi_\theta | \Psi_\theta \rangle} - E_\theta \frac{\langle \Psi_\theta | \partial_{\theta_j} \Psi_\theta \rangle}{\langle \Psi_\theta | \Psi_\theta \rangle}. \tag{2.65}$$

By using the same method for calculating the local estimator and using the Monte Carlo Markov chain method similarly to compute **S** and **F**, one gets the following

$$\mathbf{F_j} \approx \frac{1}{N_s} \sum_{k=1}^{N_s} \partial_{\theta_j} \log(\psi_\theta(\sigma)) \Delta H^{\text{loc}}(\sigma_k), \tag{2.66}$$

$$\Delta H^{\text{loc}}(\sigma_k) = H_\theta^{\text{loc}}(\sigma) - \frac{1}{N_s} \sum_{k=1}^{N_s} H_\theta^{\text{loc}}(\sigma_k), \tag{2.67}$$

$$\begin{aligned} \mathbf{S_{ij}} \approx & \frac{1}{N_s} \sum_{k=1}^{N_s} \left(\partial_{\theta_i} \log(\psi_\theta(\sigma_k))\right)^* \left(\partial_{\theta_j} \log(\psi_\theta(\sigma_k))\right) \\ & - \frac{1}{N_s} \sum_{k=1}^{N_s} \left(\partial_{\theta_i} \log(\psi_\theta(\sigma_k))\right)^* \frac{1}{N_s} \sum_{k=1}^{N_s} \left(\partial_{\theta_j} \log(\psi_\theta(\sigma_k))\right). \end{aligned} \tag{2.68}$$

Furthermore, by inverting the **S** matrix, equation 2.39 can be solved; however, another issue arises: the dimension of the **S** matrix is N_p^2, and the computational cost of matrix inversion scales cubically with its size. Consequently, this calculation becomes computationally expensive. Recently, Markus Heyl and Ao Chen proposed a method that eliminates the need to directly invert the quantum geometric tensor. Instead, they exploit specific properties of this tensor to construct a significantly smaller matrix from a set of sampled square matrices, enabling more efficient computations.

Chapter 3

QUANTUM SPIN ICE

At low temperatures, most magnetic materials adopt an ordered state. However, the geometry of the lattice can strongly influence this behavior. In certain frustrated geometries, such as kagome and pyrochlore lattices, this ordering can be suppressed, giving rise to exotic phases. In this section, we focus on spin ice materials like $Ho_2Ti_2O_7$, examining their Hamiltonian and emergent magnetic monopole excitations. By introducing quantum fluctuations into spin ice systems, we arrive at a minimal model on the pyrochlore lattice that describes quantum spin ice—a special type of U(1) quantum spin liquid characterized by photon-like excitations, fractionalized particles, and long-range quantum entanglement. Finally, we discuss these models in the context of potential candidate materials that may host such phases.

3.1 Spin Ice

Spin ice materials are considered by putting magnetic moments on the sites of a pyrochlore lattice 3.1 with corner-sharing tetrahedra, where as these moments point along the local Ising axis, the Hamiltonian of such materials can be written as follows

$$H = \frac{J}{3}\sum_{<ij>} S_i S_j + Da^3 \sum_{<ij>} [\frac{\hat{e}_i.\hat{e}_j}{|r_{ij}|^3} - \frac{3(\hat{e}_i.\hat{r}_{ij})(\hat{e}_j.\hat{r}_{ij})}{|r_{ij}|^5}]S_i S_j, \tag{3.1}$$

Here,

$$D = \frac{\mu_0 \mu^2}{4\pi a^3} = 1.41 \text{ K}. \tag{3.2}$$

is the dipolar interaction coupling constant, where r_{ij} denotes the distance between

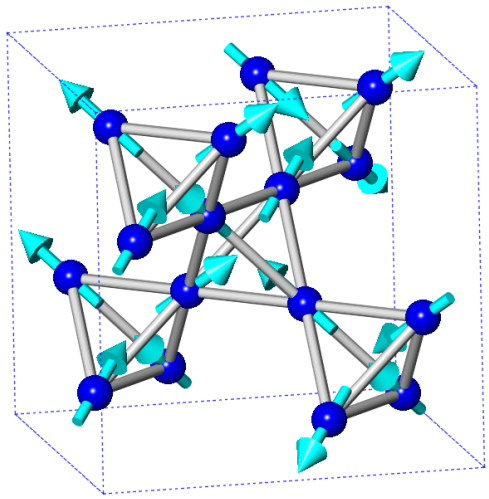

Figure 3.1: Half-integer spin particles on a pyrochlore lattice following the ice rule

spins on the pyrochlore lattice and a is the nearest-neighbor spacing.

Spin ice is a special type of magnet that, even at low temperatures, does not exhibit conventional ordering. These materials allow for many possible spin configurations that obey the ice rule—meaning two spins point in and two point out at each vertex—to minimize the system's ground-state energy. This constraint leads to a residual entropy similar to that of water ice.

If the ice rule is locally violated by flipping a spin, an excited state is created, resulting in a defect that behaves like a magnetic monopole. These monopoles arise due to long-range interactions between spins. To visualize this, each spin can be represented as a dumbbell with positive and negative magnetic charges at its ends, aligned along the bonds of the underlying diamond lattice. Flipping a spin corresponds to changing the arrangement of these charges. The magnitude of the charges is chosen such that the dumbbell accurately represents the original spin.

The energy of the system is determined by the interactions between the magnetic

charges at the ends of the dumbbells. Flipping a single spin creates two adjacent magnetic charges, and by successive flips, these monopoles can move apart without violating the ice rule. The interaction between monopoles follows the magnetic Coulomb law. Separating them requires a finite amount of energy, making these monopoles the elementary excitations of the system. The energy of a configuration of dipoles is thus governed by the magnetic Coulomb law, as given below

$$V_{ab} = \begin{cases} \frac{\mu_0}{4\pi} \frac{Q_\alpha Q_\beta}{r_{\alpha\beta}}, & \alpha \neq \beta \\ \\ \frac{1}{2}\mu_0 \frac{Q^2}{\alpha}, & \alpha = \beta \end{cases} \tag{3.3}$$

Q_α is the total magnetic charge at site α, while $r_{\alpha\beta}$ is the distance between the two sites α and β. In the classical spin ice materials, several experiments studied the monopole excitation, such as neutron scattering, magnetization relaxation, partially supporting the monopole behavior. The most successful result comes from the field-driven transition, which can be interpreted by the crystallization of positive and negative monopoles. In the subsequent section, by adding a fluctuation term, the subject moves towards quantum spin ice, and a new excitation will appear.

3.2 Quantum Spin Ice

Quantum spin ice is a type of U(1) quantum spin liquid found in pyrochlore magnets. The system remains disordered even at very low temperatures and hosts fractionalized excitations that behave like emergent particles. These excitations are gapless, meaning that no energy is required to create them, and their dispersion is linear, resembling that of light waves.

The excitation particles possess two transverse polarizations, analogous to photons, despite being spin excitations. In conventional photon theory, these polarizations arise

from gauge invariance in a system of massless spin-1 particles. In quantum spin ice, however, the spin dynamics generate an emergent gauge field, giving rise to photon-like excitations. In contrast, in a typical magnetically ordered state at low temperatures, magnons also have two polarizations, but one direction is fixed by broken symmetry. In quantum spin ice, the two polarizations arise from gauge invariance rather than symmetry breaking.

Quantum spin ice is particularly intriguing because it exhibits exotic behavior analogous to quantum electrodynamics within a condensed-matter setting.

To construct the Hamiltonian of the system, we begin with the classical Ising model for nearest-neighbor interactions, which describes the classical spin ice regime, as illustrated below

$$H_{CSI} = J_{||} \sum_{<ij>} S_i^z S_j^z. \tag{3.4}$$

There are many different configurations that obey the local ice rules at the ground state; therefore, the system is highly degenerate. The ice rule for each tetrahedron can be formulated as follows

$$\sum_a S_a . \hat{z}_a = 0. \tag{3.5}$$

This ice rule suggests a divergence-free spin field, as shown below

$$\nabla \cdot \vec{B} = 0. \tag{3.6}$$

Any vector field can be decomposed into two components: one with zero divergence and one with zero curl. In spin ice, the zero-divergence condition originates from the

ice rule, which means that only the zero-curl component contributes to thermodynamic effects. Furthermore, the divergence-free condition implies an emergent gauge invariance, analogous to that in electromagnetism.

At low but finite temperatures, some spins can flip, locally violating the ice rule and breaking the divergence-free condition, which gives rise to emergent magnetic monopoles. When these monopoles are sparse, spin ice behaves like a dilute plasma of magnetic charges.

By introducing a small additional interaction that allows spin flips—known as the transverse nearest-neighbor term added to the Ising term—quantum fluctuations are incorporated while largely preserving the ice rule. This transverse term generates quantum dynamics within the ice-rule manifold. Since the coupling favors the Ising exchange term, the resulting Hamiltonian corresponds to the XXZ model on the pyrochlore lattice.

$$H_{QSI} = J_{||} \sum_{<ij>} S_i^z S_j^z - J_{\pm} \sum_{<ij>} (S_i^+ S_j^- + S_i^- S_j^+). \tag{3.7}$$

The classical spin ice ground states form a highly degenerate manifold, constituting a classical spin liquid. A small transverse exchange coupling acts as a perturbation, and, using degenerate perturbation theory, the zeroth-order states are mapped onto the full ice-rule manifold. The lowest-order terms that preserve the ice rule are ring-exchange interactions on hexagonal plaquettes. Consequently, the resulting effective low-energy Hamiltonian captures the quantum fluctuations within the ice-rule manifold, as shown below

$$H_{\text{ring}} \approx \frac{J_{\pm}^3}{J_{\parallel}^2} \sum_{h \in \hexagon} S_{h,1}^+ S_{h,2}^- S_{h,3}^+ S_{h,4}^- S_{h,5}^+ S_{h,6}^- + \text{h.c.}. \tag{3.8}$$

This effective low-energy Hamiltonian is invariant under local $U(1)$ gauge transformations. As a result, the corresponding phase is referred to as a U(1) quantum spin liquid, due to the emergent U(1) gauge symmetry.

The ring-exchange terms dominate the low-energy Hamiltonian. To better understand their effects, it is useful to map the ring-exchange model onto a quantum dimer model on the diamond lattice. In this mapping, spin ice states correspond to loop coverings of the diamond lattice, with each dimer representing a spin along a link and dimers connected end-to-end. In this representation, the ice rule becomes the constraint that each diamond lattice site is connected to exactly two dimers, and the ring-exchange Hamiltonian captures the quantum dynamics of these dimers. Under this mapping, the ring-exchange term becomes

$$H_{\text{dimer}} = -V \sum_{\hexagon} (|\, \hexagon\rangle\langle \hexagon\, | + \text{h.c.}) + \mu \sum_{\hexagon} (|\, \hexagon\rangle\langle \hexagon\, | + |\, \hexagon\rangle\langle \hexagon\, |)\,. \tag{3.9}$$

Here, V denotes the amplitude for the resonance of dimers around a hexagonal loop, and μ represents the potential energy associated with flippable hexagons. The operators $|\, \hexagon\rangle\langle \hexagon\, |$ corresponds to the two possible dimer configurations on a hexagon. Each term in the Hamiltonian flips a loop of dimers around a hexagon while maintaining the ice-rule constraint. Overall, this Hamiltonian describes the quantum dynamics of the dimer configurations. At the RK point ($V = \mu$), the ground state wavefunction can be written exactly as a superposition of all loop coverings. The phases around this point are as follows, for $\mu > V$, the ground state forms a long-range ordered crystalline dimer phase, for $\mu < 0$, it transitions into the squiggle state, and for $\mu < V$, the system resides in a $U(1)$ quantum liquid phase.

Chapter 4

NQS STUDY OF XXZ MODEL ON PYROCHLORE LATTICE

There are several challenges and open questions related to Neural Quantum States (NQS), including NQS capacity, long-time dynamics, open quantum systems, frustrated systems, simulation of quantum circuits, quantum state tomography, and more. In this section, we apply NQS to a minimal model of quantum spin ice to explore its capability in simulating such a frustrated system.

4.1 Constructing the XXZ Model on the Pyrochlore Lattice

To initialize the simulation, the Hilbert space is constructed from 16 spin-1/2 particles arranged on a pyrochlore lattice, which consists of corner-sharing tetrahedra with a spin at each site. The Hamiltonian of the XXZ model is given in equation (3.7). In the following section, exact diagonalization is employed to determine the ground state, serving as a benchmark for comparison.

4.2 Exact Diagonalization as a Benchmark

Exact diagonalization [28] is a numerical technique in many-body quantum physics that involves directly diagonalizing the Hamiltonian matrix, as illustrated below

$$\langle \Psi | H | \Psi \rangle . \tag{4.1}$$

Then one computes the eigenvalues and eigenstates of the following matrix

$$\hat{H} \, | \Psi_n \rangle = E_n \, | \Psi_n \rangle . \tag{4.2}$$

For the XXZ Hamiltonian defined on a pyrochlore lattice with 16 spin-$\frac{1}{2}$ particles, the system can be studied through exact diagonalization. Figure 4.1 displays the corresponding ground state.

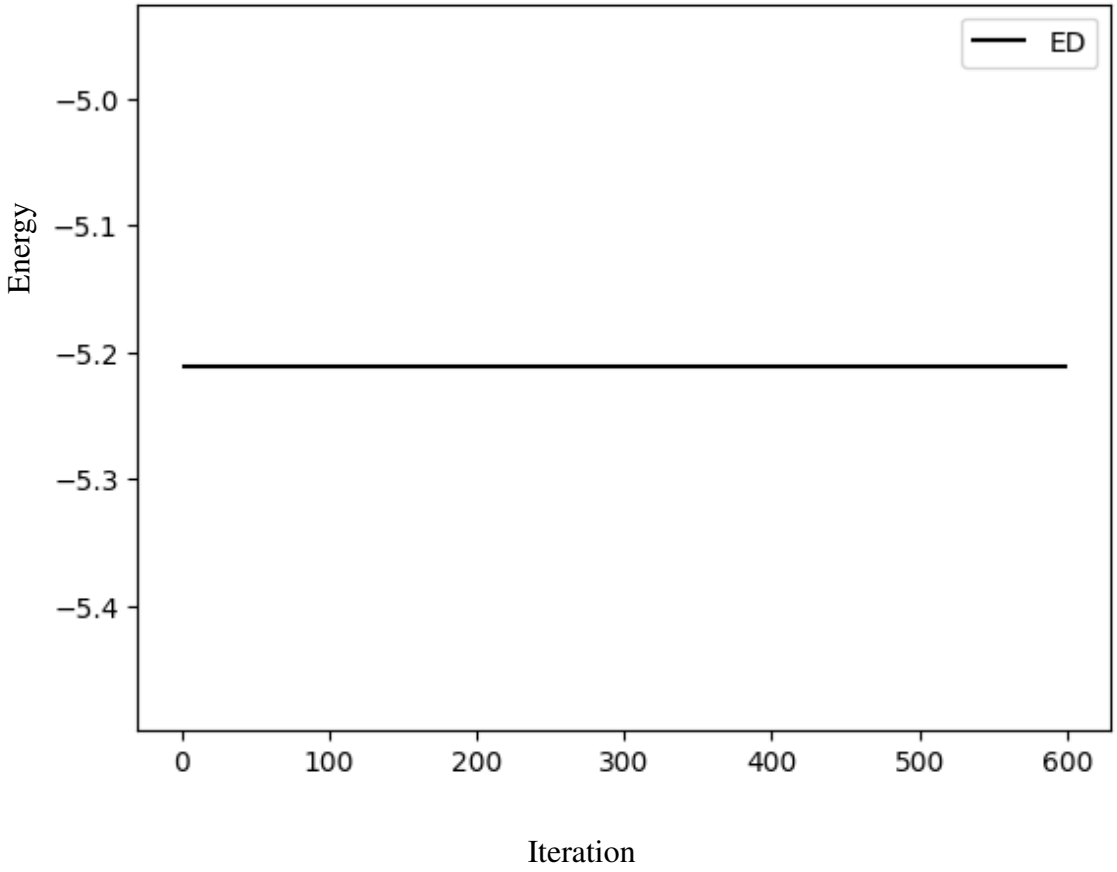

Figure 4.1: Ground state energy of the XXZ Hamiltonian on the pyrochlore lattice via exact diagonalization

4.3 Mean field Ansatze

A mean-field [29] ansatz is a variational approach used to approximate the ground state of a specific Hamiltonian within a given Hilbert space. Its general form is presented below

$$\langle \sigma_1^z, \sigma_2^z, ..., \sigma_N^z | \Psi_{mf} \rangle = \Pi_{i=1}^{N} \Phi(\sigma_i^z). \tag{4.3}$$

Instead of trying to solve the full interacting system, which is usually intractable, the

mean-field ansatz approximates each particle as interacting with an effective average field produced by all the others. In this framework, the variational parameters are the single-spin wavefunctions, which can be chosen to be normalized as shown below

$$|\Phi(\uparrow)|^2 + |\Phi(\downarrow)|^2 = 1. \tag{4.4}$$

Using the definition $\Phi(\sigma^z) = \sqrt{P(\sigma^z)}e^{i\Phi(\sigma^z)}$, for simplicity, we set the phase to zero and express the single-spin probabilities in the sigmoid form as shown below.

$$P(\sigma_z;\lambda) = \frac{1}{1+\exp(-\lambda\sigma_z)}. \tag{4.5}$$

As a result, the variational parameter is identified as λ. By optimizing this parameter, the ground-state energy can be obtained, as illustrated in Figure (4.2).

As depicted in Figure 4.2, the mean-field ansatz for the ground-state energy yields a catastrophic result. However, this outcome is not surprising, since the very definition of the mean-field ansatz relies on approximating interactions through an effective average field. Consequently, it fails to capture the correlations present in the system. In the next section, we turn to another ansatz, namely the Jastrow ansatz, to examine whether it can provide improved results and reduce the gap between the variational approximation and the exact diagonalization results.

4.4 Jastrow Ansatze

In the mean-field ansatz, the correlations between particles could not be properly captured, as indicated by the results. Since quantum spin ice materials are highly correlated systems, an alternative ansatz is required to account for these correlations. In this section, the Jastrow ansatz [30] is introduced, which consists of two parameters and incorporates entanglement between nearest and next-to-nearest neighbors. With

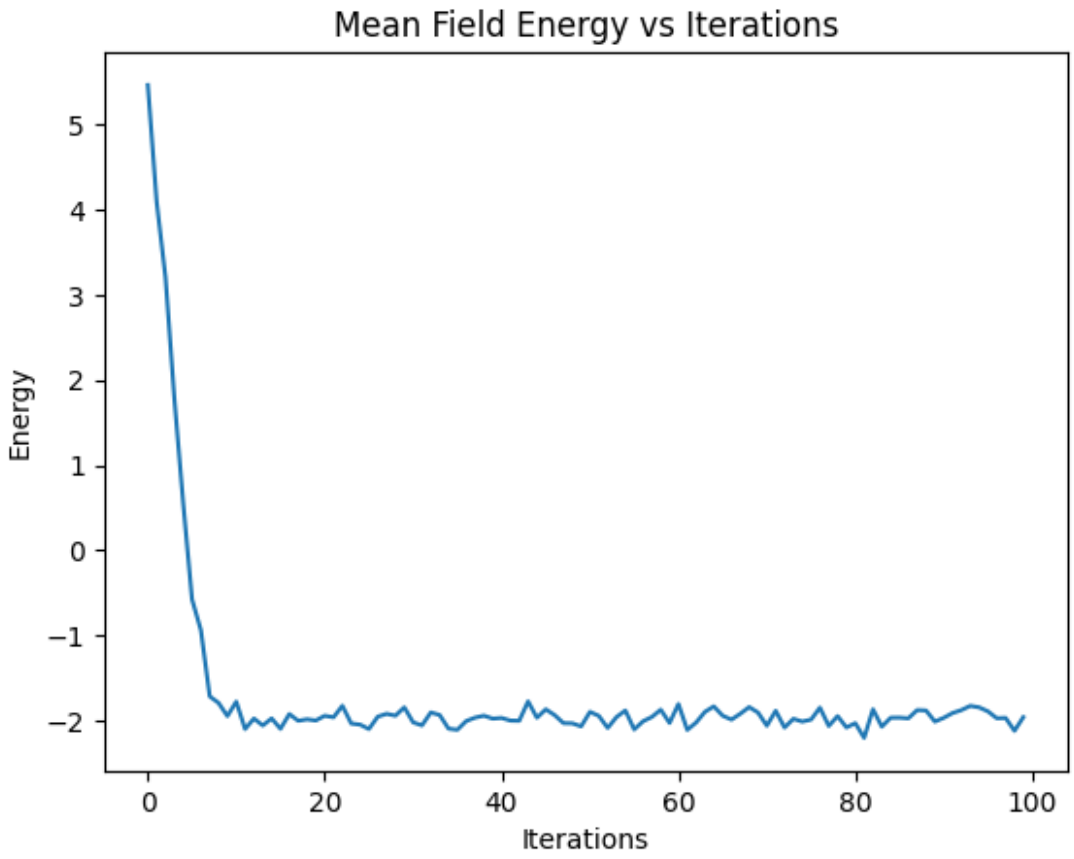

Figure 4.2: Ground-state energy of the pyrochlore-lattice XXZ Hamiltonian computed with mean-field variational ansatze

this ansatz, the ground state energy is obtained as shown in Figure (4.3)

$$\langle \sigma_1^z, \sigma_2^z, ..., \sigma_N^z | \Psi_{jas} \rangle = \exp \sum_i (J_1 \sigma_i^z \sigma_{i+1}^z + J_2 \sigma_i^z \sigma_{i+2}^z). \tag{4.6}$$

where J_1 and J_2 are the parameters that need to be optimizedized.

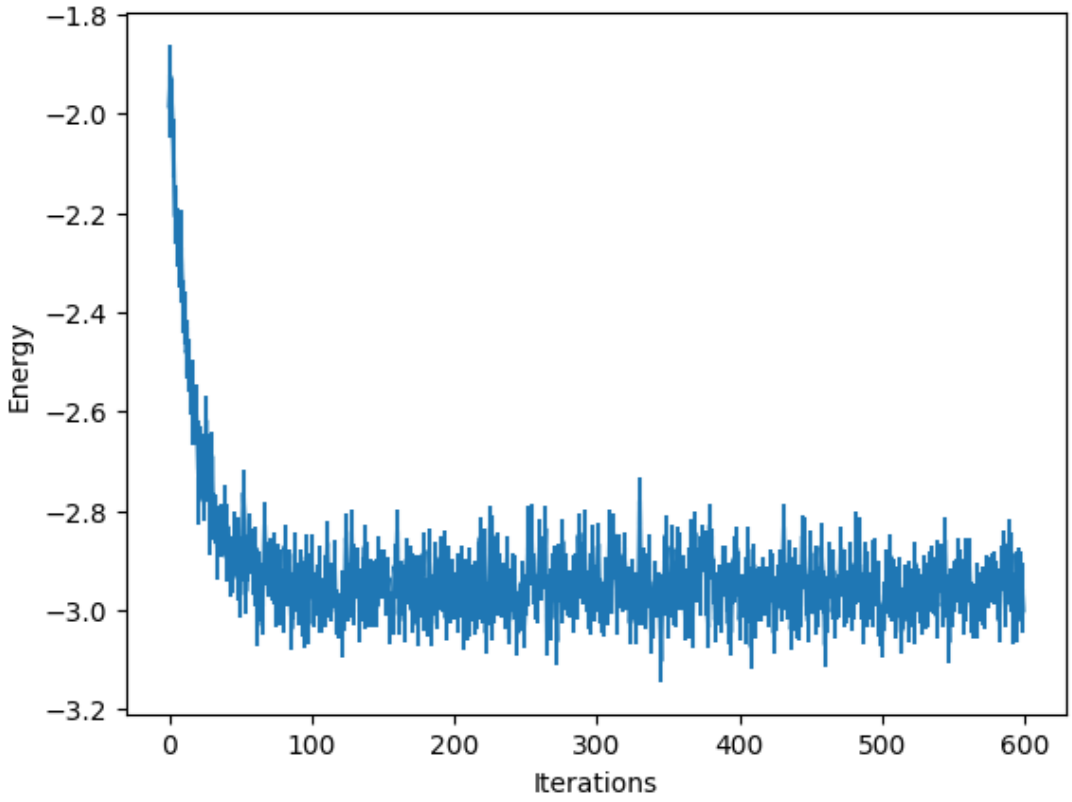

Figure 4.3: XXZ Hamiltonian ground-state energy on the pyrochlore lattice via Jastrow-based variational ansatze

4.5 RBM and symmetrical RBM Ansatze

The Restricted Boltzmann Machine (RBM) [31] was the first neural network applied in the framework of Neural Quantum States (NQS). An RBM is a simple neural network consisting of only two layers: a visible layer with N neurons and a hidden layer with M neurons. Every visible neuron is connected to every hidden neuron, while no connections exist within the same layer. The variational ansatz of the RBM is given below

$$\langle \sigma_1^z, \sigma_2^z, ..., \sigma_N^z | \Psi_{RBM} \rangle = \sum_h e^{b_v^\dagger s + b_h^\dagger h + h^\dagger W s}. \tag{4.7}$$

Whereas v and h indicate the visible and hidden neurons, the parameters $\theta = \{b_v, b_h, W\}$ represent the biases of the visible and hidden neurons, as well as the weight matrix connecting these neurons, and s denotes the input configuration. The RBM ansatze can be reconstructed as follows in order to become tractable

$$\langle \sigma_1^z, \sigma_2^z, ..., \sigma_N^z | \Psi_{RBM} \rangle = e^{b_v^\dagger s} \Pi_{i=1}^M 2\cosh(b_{h,i} + W_i \cdot s). \tag{4.8}$$

Furthermore, by applying this ansatz to the XXZ model on the pyrochlore lattice and optimizing the parameters, the ground-state energy is obtained, as shown in Figure (4.4). Note that in this simulation, there are two ground states: one obtained using the plain RBM and another using the RBM with the pyrochlore lattice symmetry imposed.

The gap between the ground-state energies and the exact result remains significant. In the subsequent section, by introducing a new architecture known as group convolutional neural networks (G-CNNs), we explore how the results can be further improved.

4.6 G-CNNs Ansatze

Convolutional Neural Networks (CNNs) [32] are a type of neural network that works a bit differently from standard feed-forward networks. Instead of connecting every neuron to all neurons in the next layer, each neuron only looks at a small, local region of the input, called a filter. This design not only reduces the number of parameters the network has to learn but also makes it great at picking up local patterns. CNNs achieve this by sliding these filters across an image or lattice, scanning for features.

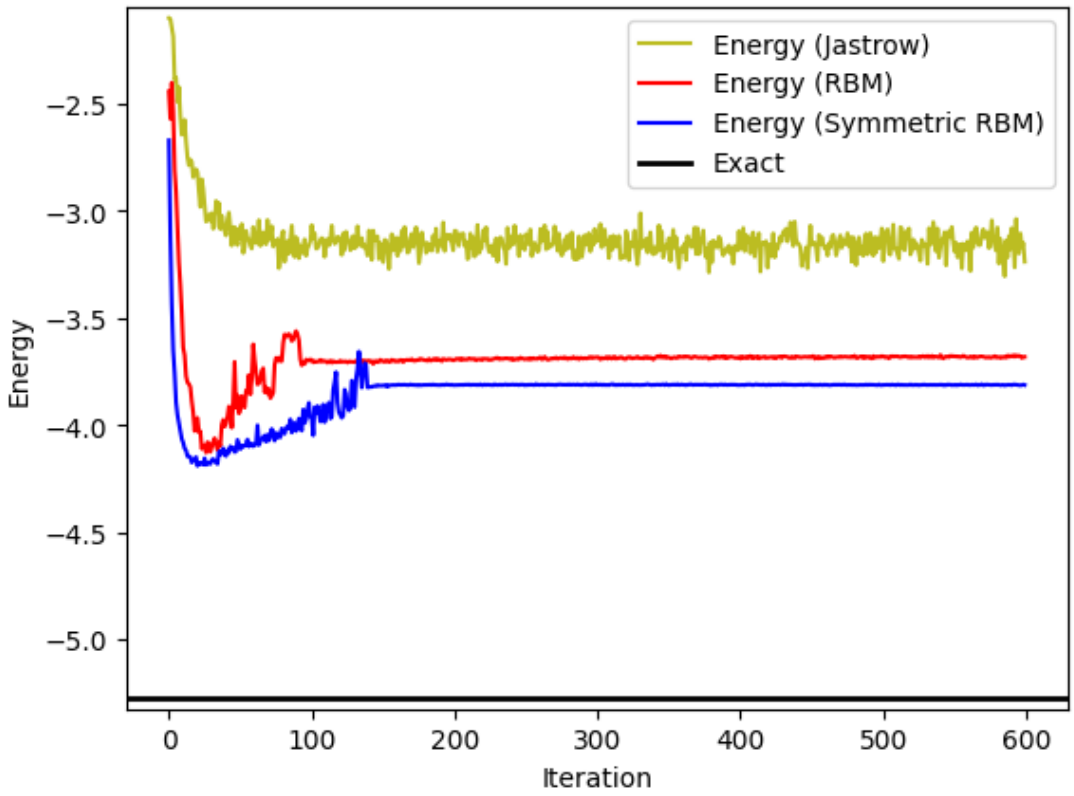

Figure 4.4: Ground-state energy of the pyrochlore-lattice XXZ Hamiltonian computed with RBM variational ansatze

Taking this idea further, Group Convolutional Neural Networks (G-CNNs) let the network take advantage of symmetries in the data like rotations, reflections, and translations, even when the order of these operations matters. By doing so, G-CNNs can understand patterns in a more flexible and powerful way, often leading to much better results. By using this architecture, the ground state energy estimation is shown in figure 4.5

4.7 FNN and symmetrical FNN Ansatze

A feed-forward neural network [33] was already introduced in Chapter 2.1, where two types of networks were constructed: one without symmetry and another with imposed lattice symmetry. The corresponding ground state is shown in Figure 4.6.

As can be clearly seen from the figure, the symmetrical model provides an excellent

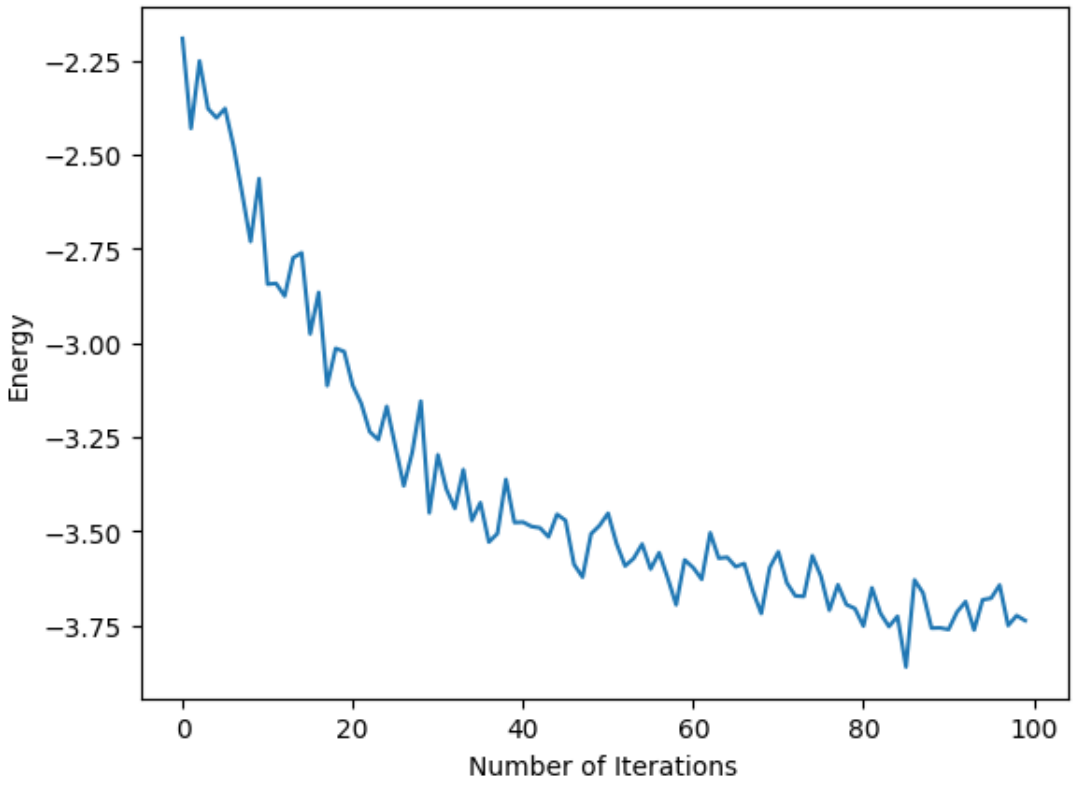

Figure 4.5: Ground-state energy of the XXZ Hamiltonian on the pyrochlore lattice obtained using G-CNN variational ansatze

result for the ground state. Now that the ground state has been determined, the next section aims to examine whether the NQS approach can capture the quantum phase transition.

4.8 Quantum Phase Transition

Since we now have a reliable variational state that provides a good estimate of the ground state, we can proceed to compute other observables from this neural quantum state representation. One of the most important observables is the nearest-neighbor zz and xy spin correlations, as shown below

$$C_{zz} = \langle \Psi_{gs} | \sigma_i^z \sigma_{i+1}^z | \Psi_{gs} \rangle \,, \tag{4.9}$$

$$C_{XY} = \langle \Psi_{gs} | \sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y | \Psi_{gs} \rangle \,. \tag{4.10}$$

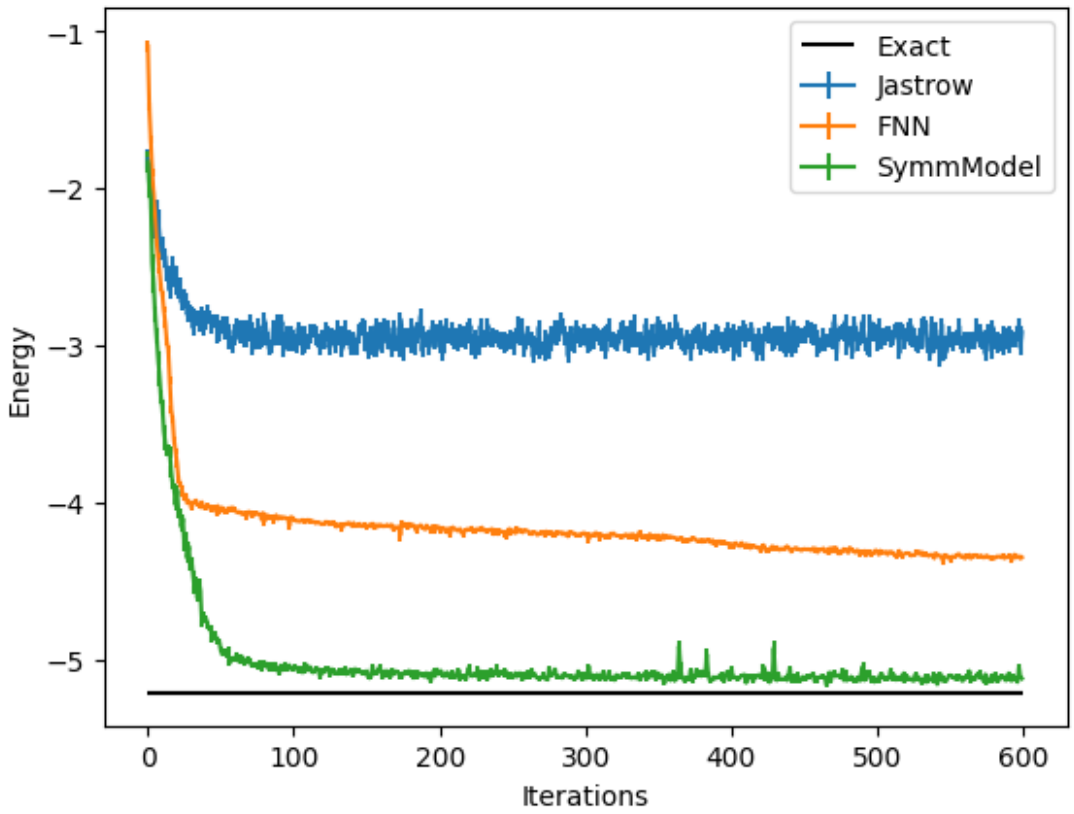

Figure 4.6: Ground-state energy of the XXZ Hamiltonian on the pyrochlore lattice obtained using FFN variational ansatze

These quantities measure the correlations between spin operators on neighboring sites in the quantum spin system. Such observables are crucial for understanding entanglement patterns, symmetry breaking, and quantum phase transitions. For example, if $C_{zz} < 0$, the system exhibits antiferromagnetic order; if $C_{zz} > 0$, it indicates ferromagnetic order; and if $|C_{zz}| \ll 1$, The system behaves as a paramagnet. In the case of the XXZ model on the pyrochlore lattice, the nearest-neighbor correlations for various coupling ratios are presented in Figure (4.7).

As shown in the figure 4.7, three distinct phases can be observed as the coupling ratio $J_{\pm}/J_{zz}$ is varied. When the transverse coupling is close to zero, the system behaves like a classical spin ice with an antiferromagnetic phase. The XY correlator is nearly zero, indicating that the spins are mostly constrained along the local z-axis, satisfying

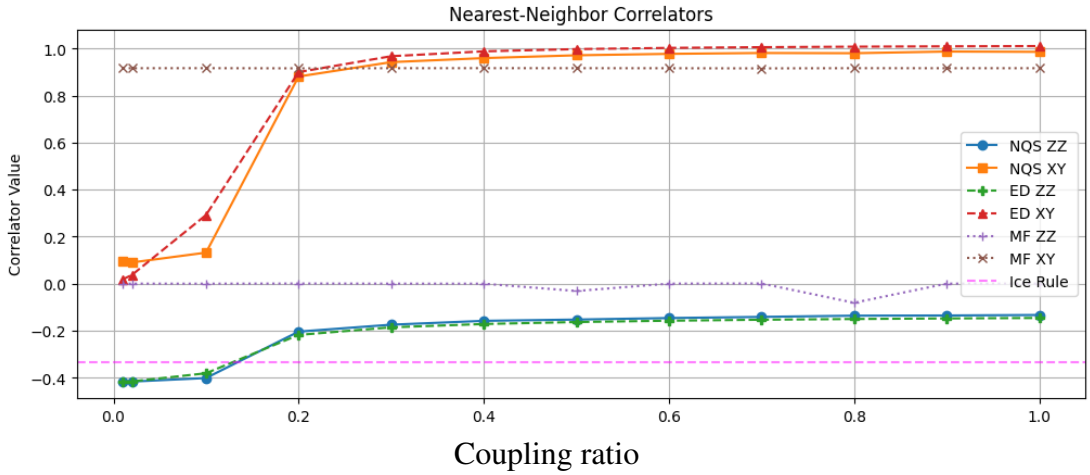

Figure 4.7: Nearest-neighbor correlations across various coupling ratios

the 2-in–2-out ice rule.

As $J_{\pm}$ increases, quantum fluctuations become significant. This is reflected in the growth of the XY correlations, while the ZZ correlations remain roughly unchanged, indicating that the fluctuations are still small. In this regime, the Hamiltonian can be described by degenerate perturbation theory, leading to an exchange ring term where spins fluctuate along closed loops in the XY plane, while still satisfying the ice rule. The classical order is thus melted by quantum fluctuations, giving rise to a U(1) quantum spin liquid with emergent gauge fields, photon-like excitations, and gapped magnetic monopoles (fractionalized quasiparticles). This transition does not involve any symmetry breaking.

When the XY term dominates, spins tend to align within the XY plane, forming a ferromagnetic state and violating the ice rule. Consequently, the spin liquid melts, and the U(1) symmetry is spontaneously broken. Both transitions are quantum phase transitions, occurring at zero temperature.

Furthermore, the U(1) quantum spin liquid (QSL) phase is characterized by an emergent U(1) gauge field and fractionalized spinons (magnetic monopoles), which are gapped excitations requiring energy to violate the ice rule and create them. This phase also supports gapless photon-like excitations and exhibits topological order. In quantum spin ice, spinons are typically treated as deconfined bosonic quasiparticles carrying magnetic charge. They are gapped and dilute in the QSL, behaving effectively as hard-core bosons.

As the transverse coupling increases, the system undergoes a quantum phase transition to an XY-ferromagnetic phase. At this transition, the spinon gap closes, and the previously gapped spinons become gapless. These gapless spinons then condense, forming a spinon Bose–Einstein condensate. This condensation breaks the U(1) symmetry, leading to long-range ferromagnetic order. Consequently, the emergent gauge field is Higgsed, deconfinement is lost, and spinon fractionalization disappears.

An interesting question arises when considering the connection between the quantum spin liquid (QSL) and XY-ferromagnetic (XY-FM) phases. In the QSL phase, the system hosts fractionalized spinons that are deconfined and gapped, whereas the XY-FM phase emerges when these spinons condense, giving rise to spontaneously broken symmetry and long-range magnetic order. How, then, can these two seemingly very different phases, one fractionalized, the other ordered, be directly connected? In the QSL phase, spinons remain gapped and deconfined, while the emergent photon-like excitations are gapless, a reflection of the underlying U(1) gauge structure constrained by the ice rule. At the quantum critical point, the spinon gap closes, producing gapless spinons that act as strong sources of gauge fluctuations while remaining deconfined. Once the system enters the spinon Bose–Einstein condensate

phase, the spinons become confined, and the system develops long-range magnetic order. This transition from QSL to XY-FM cannot be explained by conventional Landau theory. Remarkably, the deconfinement at criticality allows for a direct transition between a fractionalized and an ordered phase. This phenomenon is a candidate for deconfined quantum criticality (DQC), which is best described within a non-Landau–Ginzburg–Wilson framework.

Chapter 5

CONCLUSION

In this thesis, we begin with an introduction to neural networks and the backpropagation algorithm. Building on this foundation, we explore the application of machine learning to ab initio simulations of many-body quantum systems, inspired by the pioneering work of Giuseppe Carleo and Matthias Troyer, who first introduced the concept of Neural Quantum States (NQS). This emerging field, still in its early stages, raises several open questions, one of which concerns the study of frustrated quantum systems through the NQS framework.

Our investigation focuses on a particularly important frustrated magnet: quantum spin ice. This system can be understood as a quantum extension of classical spin ice, where the inclusion of transverse fluctuations in the Hamiltonian leads to the minimal XXZ model defined on the pyrochlore lattice—a lattice composed of corner-sharing tetrahedra. To establish a benchmark, we first apply exact diagonalization to a 16-site spin-$\frac{1}{2}$ cluster, successfully obtaining the ground-state energy.

We then employ a variety of variational ansätze to examine how well different approaches capture the complex entanglement and correlations of this highly nontrivial quantum system. Among the tested methods, a feed-forward neural network with lattice symmetries explicitly imposed provides the most accurate approximation of the ground state.

To further probe the nature of the system, we calculate spectra of nearest-neighbor correlators, both longitudinal and transverse. These results reveal the presence of three distinct phases: the classical spin ice, the U(1) quantum spin liquid, and the XY ferromagnetic ordered state. Notably, the transition between the quantum spin liquid and the XY ferromagnetic phase cannot be explained within the conventional Landau–Ginzburg–Wilson framework. This observation strongly suggests that the transition may be a candidate for deconfined quantum criticality (DQC), an exotic non-Landau paradigm of quantum phase transitions

REFERENCES

[1] Y. Xu, X. Liu, X. Cao, C. Huang, E. Liu, S. Qian, X. Liu, Y. Wu, F. Dong, C.-W. Qiu *et al.*, "Artificial intelligence: A powerful paradigm for scientific research," *The Innovation*, vol. 2, no. 4, 2021.

[2] S. Bianchini, M. Müller, and P. Pelletier, "Artificial intelligence in science: An emerging general method of invention," *Research Policy*, vol. 51, no. 10, p. 104604, 2022.

[3] J. Willard, X. Jia, S. Xu, M. Steinbach, and V. Kumar, "Integrating physics-based modeling with machine learning: A survey," *arXiv preprint arXiv:2003.04919*, vol. 1, no. 1, pp. 1–34, 2020.

[4] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, "Machine learning and the physical sciences," *Reviews of Modern Physics*, vol. 91, no. 4, p. 045002, 2019.

[5] D. Guest, K. Cranmer, and D. Whiteson, "Deep learning and its application to lhc physics," *Annual Review of Nuclear and Particle Science*, vol. 68, no. 1, pp. 161–181, 2018.

[6] G. Karagiorgi, G. Kasieczka, S. Kravitz, B. Nachman, and D. Shih, "Machine learning in the search for new fundamental physics," *Nature Reviews Physics*, vol. 4, no. 6, pp. 399–412, 2022.

[7] A. Tanaka, A. Tomiya, and K. Hashimoto, *Deep learning and physics*. Springer, 2021, vol. 1.

[8] M. C. Bañuls, "Tensor network algorithms: A route map," *Annual Review of Condensed Matter Physics*, vol. 14, no. 1, pp. 173–191, 2023.

[9] P. Silvi, F. Tschirsich, M. Gerster, J. Jünemann, D. Jaschke, M. Rizzi, and S. Montangero, "The tensor networks anthology: Simulation techniques for many-body quantum lattice systems," *SciPost Physics Lecture Notes*, p. 008, 2019.

[10] J. I. Cirac and F. Verstraete, "Renormalization and tensor product states in spin chains and lattices," *Journal of physics a: mathematical and theoretical*, vol. 42, no. 50, p. 504004, 2009.

[11] S. Paeckel, T. Köhler, A. Swoboda, S. R. Manmana, U. Schollwöck, and C. Hubig, "Time-evolution methods for matrix-product states," *Annals of Physics*, vol. 411, p. 167998, 2019.

[12] A. Nüßeler, I. Dhand, S. F. Huelga, and M. B. Plenio, "Efficient construction of matrix-product representations of many-body gaussian states," *Physical Review A*, vol. 104, no. 1, p. 012415, 2021.

[13] J.-W. Li and X. Waintal, "Matrix product states and first quantization," *arXiv preprint arXiv:2404.07105*, 2024.

[14] A. W. Sandvik and G. Vidal, "Variational quantum monte carlo simulations with tensor-network states," *Physical review letters*, vol. 99, no. 22, p. 220602, 2007.

[15] W. Huggins, C. D. Freeman, M. Stoudenmire, N. M. Tubman, and K. B. Whaley, "Monte carlo tensor network renormalization," *arXiv preprint arXiv:1710.03757*, 2017.

[16] G. Carleo and M. Troyer, "Solving the quantum many-body problem with artificial neural networks," *Science*, vol. 355, no. 6325, pp. 602–606, 2017.

[17] I. Glasser, N. Pancotti, M. August, I. D. Rodriguez, and J. I. Cirac, "Neural-network quantum states, string-bond states, and chiral topological states," *Physical Review X*, vol. 8, no. 1, p. 011006, 2018.

[18] K. Ch'Ng, J. Carrasquilla, R. G. Melko, and E. Khatami, "Machine learning phases of strongly correlated fermions," *Physical Review X*, vol. 7, no. 3, p. 031038, 2017.

[19] K. Choo, T. Neupert, and G. Carleo, "Two-dimensional frustrated j 1-j 2 model studied with neural network quantum states," *Physical Review B*, vol. 100, no. 12, p. 125124, 2019.

[20] Y. Nomura, A. S. Darmawan, Y. Yamaji, and M. Imada, "Restricted boltzmann machine learning for solving strongly correlated quantum systems," *Physical Review B*, vol. 96, no. 20, p. 205152, 2017.

[21] M. Y. Pei and S. R. Clark, "Neural-network quantum states for spin-1 systems: spin-basis and parameterization effects on compactness of representations," *Entropy*, vol. 23, no. 7, p. 879, 2021.

[22] T. Westerhout, N. Astrakhantsev, K. S. Tikhonov, M. I. Katsnelson, and A. A. Bagrov, "Generalization properties of neural network approximations to frustrated magnet ground states," *Nature communications*, vol. 11, no. 1, p. 1593, 2020.

[23] S. T. Bramwell and M. J. Gingras, "Spin ice state in frustrated magnetic pyrochlore materials," *Science*, vol. 294, no. 5546, pp. 1495–1501, 2001.

[24] C. Castelnovo, R. Moessner, and S. L. Sondhi, "Spin ice, fractionalization, and topological order," *Annu. Rev. Condens. Matter Phys.*, vol. 3, no. 1, pp. 35–55, 2012.

[25] M. J. Gingras and P. A. McClarty, "Quantum spin ice: a search for gapless quantum spin liquids in pyrochlore magnets," *Reports on Progress in Physics*, vol. 77, no. 5, p. 056501, 2014.

[26] C. Broholm, R. J. Cava, S. Kivelson, D. Nocera, M. Norman, and T. Senthil, "Quantum spin liquids," *Science*, vol. 367, no. 6475, p. eaay0668, 2020.

[27] L. Savary and L. Balents, "Coulombic quantum liquids in spin-1/2 pyrochlores," *Physical review letters*, vol. 108, no. 3, p. 037202, 2012.

[28] A. Wietek, L. Staszewski, M. Ulaga, P. L. Ebert, H. Karlsson, S. Sarkar, H. Shackleton, A. Sinha, and R. D. Soares, "Xdiag: Exact diagonalization for quantum many-body systems," *arXiv preprint arXiv:2505.02901*, 2025.

[29] H. Bruus and K. Flensberg, *Many-body quantum theory in condensed matter physics: an introduction*. Oxford university press, 2004.

[30] J. Yang and A. del Campo, "Time-dependent jastrow ansatz: Exact quantum dynamics, shortcuts to adiabaticity, and quantum quenches in strongly correlated many-body systems," *Physical Review A*, vol. 111, no. 5, p. 053315, 2025.

[31] A. Fischer and C. Igel, "An introduction to restricted boltzmann machines," in *Iberoamerican congress on pattern recognition*. Springer, 2012, pp. 14–36.

[32] M. Najibi, M. Rastegari, and L. S. Davis, "G-cnn: an iterative grid based object detector," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2369–2377.

[33] G. Bebis and M. Georgiopoulos, "Feed-forward neural networks," *Ieee Potentials*, vol. 13, no. 4, pp. 27–31, 2002.