Tensor network descriptions of quantum entanglement in path integrals, thermalisation and machine learning

Andrew Hallam

Supervisors:
Prof. Andrew G. Green
Dr. Andrew J. A. James

A thesis presented for the degree of
Doctor of Philosophy

Department of Physics and Astronomy
University College London
Disclaimer

I, Andrew Hallam confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature: ..........................
Abstract

One of the major ways in which quantum mechanics differs from classical mechanics is the existence of special quantum correlations - entanglement. Typical quantum states are highly entangled, making them complex and inefficient to represent. Physically interesting states are unusual, they are only weakly entangled. By restricting ourselves to weak entanglement, efficient representations of quantum states can be found.

A tensor network is constructed by taking objects called tensors that encode spatially local information and gluing them together to create a large network that describes a complex quantum state. The manner in which the tensors are connected defines the entanglement structure of the quantum state. Tensors networks are therefore a natural framework for describing physical behaviour of complex quantum systems.

In this thesis we utilize tensor networks to solve a number of interesting problems. Firstly, we study a Feynman path integral written over tensor network states. As a sum over classical trajectories, a Feynman path integral can struggle to capture entanglement. Combining the path integral with tensor networks overcomes this. We consider the effect of quadratic fluctuations on the tensor network path integral and calculate corrections to observables numerically and analytically.

We also study the time evolution of complex quantum systems. By projecting quantum dynamics onto a classical phase space defined using tensor networks, we relate thermal behaviour of quantum systems to classical chaos. In doing so we demonstrate a relationship between entanglement growth and chaos. By studying the dynamics of
coupled quantum chains we also gain insight into how quantum correlations spread over
time.

As noted, tensor networks are remarkably efficient. In the final section of this thesis
we use tensor networks to create compressed machine learning algorithms. Their effi-
ciency means that tensor networks can use 50 times fewer parameters with no significant
decrease in performance.
Impact statement

Tensor networks are a key technique for understanding complex quantum systems. An area of major interest in recent years has been the study of complex quantum systems far from equilibrium. For example, the relaxation of an isolated quantum system to a steady state is still poorly understood. One puzzling aspect of this relaxation is its profound difference from the classical equivalent. In this thesis we have bridged the gap between classical and quantum systems by using tensor networks to bring classical concepts such as deterministic chaos into a quantum setting. This work makes it possible to apply the entire framework of classical chaos to study complex quantum systems.

Another contribution of this work has been to compress neural networks using tensor networks techniques. Many technology companies use neural networks for tasks such as image recognition or natural language processing. Unfortunately, the millions or billions of parameters needed to define a neural network limits their applicability on memory limited devices - such as mobile phones. New compression techniques could therefore be incredibly useful.

Combining tensor networks with machine learning has a second interesting consequence - machine learning algorithms can be deployed on quantum computers. Tensor networks can naturally describe quantum circuits, they are therefore the perfect language for applying machine learning to quantum computers. Some believe quantum machine learning will be an useful application of a near-term quantum computer is it may be more robust against noise and decoherence.
# Contents

1 Overview 9

2 An introduction to tensor networks 12
   2.1 Entanglement and area laws 12
   2.2 Tensor networks and graphical notation 13
   2.3 Properties of matrix product states 15
   2.4 Manifold and tangent-space methods for matrix product states 19
   2.5 Tree-tensor networks 20

3 Feynman path integral over matrix product states I: General theory 23
   3.1 Feynman path integral over product states 24
   3.2 Constructing the path integral 25
   3.3 Feynman path integral over matrix product states 26
   3.4 Fluctuations around a product state saddle point 28
   3.5 Fluctuations around a matrix product saddle point 31
   3.6 Calculating $M_k(\mathcal{H}, A_0)$ 34
      3.6.1 Calculating $\epsilon_k(\mathcal{H}, A_0)$ 34
      3.6.2 Calculating $\Delta_k(\mathcal{H}, A_0)$ 36
   3.7 Diagonalising the fluctuation around an MPS saddle point 37
   3.8 Properties of fluctuations around an MPS saddle point 38
Chapter 1

Overview

By the late 1920s it was clear to most physicists that the world of atoms could not be explained by classical mechanics. Instead, a strange new approach called quantum theory was developed. Quantum theory differs from classical mechanics in many ways. Particles are described as complex wavefunctions rather than point particles with well defined position and momentum. Measuring the observable quantities of the wavefunction results in inherently probabilistic outcomes. While some embraced quantum theory as the new foundation of physics others rejected the idea that a probabilistic theory could be fundamental and suggested that it must emerge from some more intuitive classical theory.

Despite contributing significantly to its development, Albert Einstein was one of those who became sceptical about the new theory. In 1935 along with his collaborators Boris Podolsky and Nathan Rosen he drew attention to a paradox which he believed fundamentally undermined quantum theory. Einstein pointed out if two quantum particles shared correlations then a measurement on one particle appears to have an instantaneous effect on the other, regardless of the distance between them. This “spooky action at a distance” was taken as evidence by Einstein that quantum theory must emerge from a more fundamental classical theory.
In the 1960s John Stewart Bell demonstrated that no local, classical theory could have the special quantum correlations - *quantum entanglement* - necessary to reproduce this phenomenon. Subsequent experiments have confirmed that nature cannot be described by local, classical theory. Rather than undermining quantum mechanics, Einstein’s thought experiment resulted in a devastating blow to classical mechanics.

Quantum entanglement is not just conceptually interesting, it has a profound impact on the behaviour of many-body quantum systems. There are many more possible configurations of quantum particles than there are of their classical equivalents. A linear increase in the number of quantum particles involved in a system results in an exponential increase in the number of possible quantum states. A consequence of this exponential growth is that exact simulation a system of even fifty particles is impossible using any classical approach.

However, in many ways this exponential growth is an illusion. While in principal there are exponentially many possible quantum states, the number which are relevant physically is significantly smaller. A randomly chosen quantum state is highly entangled. If we take any section of a random quantum state and measure its entanglement with the rest of the system we find that it grows with the overall size of the patch. This is called an entanglement *volume law*. Fortunately, physical states have an unusual entanglement structure. For example, the low energy states of Hamiltonians with local interactions and a finite gap between the ground state and first excited state obey an an entanglement *area law*. If a section of these systems is considered, the entanglement instead grows with the size of its boundary. As a result, rather than requiring exponentially many numbers, only polynomially many are necessary. Physically relevant states live in a small region of weakly entangled states rather than the vast expanse of highly entangled states. In fact, the vast majority of possible states take exponentially long to reach through time evolution for any local Hamiltonian.

*Tensor networks* are descriptions of quantum wavefunctions constructed by taking
objects called *tensors* that describe spatially local information in a system and gluing them together to create a large network that can capture the most important information about the system. The tensor network is chosen to compliment the entanglement structure of the wavefunction. Tensor networks are therefore the natural framework for capturing the set of physically relevant quantum states.

Tensor networks are a new language for studying many-body quantum systems. They have been essential in the characterization of new, exotic phases of matter, they have revolutionised our ability to study dynamics of one-dimensional systems, and there are suggestions they may explain how spacetime geometry can emerge from entanglement in quantum gravity \([8,9]\). There is a second important reason to be interested in tensor networks. Algorithms built upon tensor networks are some of the most effective ways of finding the ground state properties, or the time evolution of quantum systems \([10]\).

In this thesis, we build upon both of these points. We combine tensor networks with other methods to extend their applicability. For example, combining them with the Feynman path integral or the truncated spectrum approach to better understand quantum systems. Additionally we combine them with neural networks to improve machine learning algorithms. We also take advantage of the fact that tensor networks represent physically relevant states. This allows us to study phenomenon such as how quantum systems approach thermal equilibrium.

In chapter 2 we give a brief overview of tensor networks and specifically *matrix product states*. We then move on to discuss how matrix product states can be combined with the Feynman path integral in chapter 3. In chapter 4 we apply the matrix product state path integral to the transverse field Ising model. In chapter 5 we try to unify classical and quantum ideas of how systems approach thermal equilibrium using matrix-product states. In chapter 6 we study the dynamics of a system created from coupling two quantum chains. Finally, we study how tensor networks can be used to compress neural networks in chapter 7.
Chapter 2

An introduction to tensor networks

2.1 Entanglement and area laws

Before describing tensor networks, we must first understand entanglement and how it is quantified. Take a quantum wavefunction and divide it into two sections $A$ and $B$

$$|\psi\rangle = \sum_{i,j} c_{i,j} |i\rangle_A \otimes |j\rangle_B,$$  

(2.1)

where $|i\rangle_A$ is an orthonormal basis for $A$ and $|j\rangle_B$ is an orthonormal basis for $B$. $|\psi\rangle$ can be rewritten in a useful form called a Schmidt decomposition. The Schmidt decomposition can be achieved by performing a singular value decomposition (SVD) on $c_{ij}$. An SVD decomposes $c = U \Lambda V$ where $U$ and $V$ are unitary matrices and $\Lambda$ is a diagonal matrices of real numbers $\{\lambda_\alpha\}$ called singular values. These unitary matrices simply rotate the basis vectors on subsystem $A$ and $B$. $|\psi\rangle$ can now be written as

$$|\psi\rangle = \sum_{\alpha=1}^{\chi} \lambda_\alpha |\alpha\rangle_A \otimes |\alpha\rangle_B.$$  

(2.2)

where $|\alpha\rangle_A = \sum_i U_{\alpha,i} |i\rangle_A$, $|\alpha\rangle_B = \sum_j V_{\alpha,j} |j\rangle_B$. The Schmidt decomposition of $|\psi\rangle$ is useful precisely because it is convenient form for measuring the entanglement between $A$
and $B$. The dimension of the Schmidt decomposition is $\chi = \min[\dim(A), \dim(B)]$. We calculate the reduced density matrix of subsystem $A$ by tracing out subsystem $B$. In the Schmidt basis, the reduced density matrix of subsystem $A$ is diagonal:

$$\rho_A = \sum_{\alpha, \alpha', \beta} \lambda_\alpha \lambda_{\alpha'} \bra{\beta}_B \bra{\alpha}_A (\alpha' \otimes \alpha}_B \bra{\alpha}_B \bra{\beta}_B = \sum_{\alpha=1}^{\chi} |\lambda_\alpha|^2 |\alpha}_A \bra{\alpha}_A$$

(2.3)

Here $|\beta}_B$ is an arbitrary orthonormal basis for $B$. The Von Neumann entropy - the canonical measure of entanglement between two parts of a pure state - is defined as the Shannon entropy of the eigenspectrum of the reduced density matrix:

$$S_A = -\text{tr}(\rho_A \log \rho_A) = -\sum_{\alpha=1}^{\chi} |\lambda_\alpha|^2 \log |\lambda_\alpha|^2.$$  

(2.4)

If there is no entanglement between $A$ and $B$ then the Von Neumann entropy would be 0. The Von Neumann entropy can reach a maximum value of $\log \chi$ when $\lambda_\alpha = \frac{1}{\sqrt{\chi}}$.

In many-body quantum physics, it is often interesting to study how $S_A$ changes as the size of $A$ increases for a specific system. In a $d$ dimensional system, $S_A$ scaling like $S_A \propto c^d$ is an entanglement volume law and $S_A \propto c^{d-1}$ is an entanglement area law. In one dimension, an entanglement area law therefore means as the size of $A$ increases the entanglement remains roughly constant, a rather useful feature for a system to possess.

2.2 Tensor networks and graphical notation

For our purposes a tensor is merely a multidimensional array. Scalars can be thought of as 0-dimensional arrays, a single number $\alpha$. Vectors are 1-dimensional arrays with several numbers $v_i$. Matrices are again an extension of this, two dimensional arrays with members $M_{ij}$. Tensors take this notion and generalise it to arbitrary $N$-dimensional arrays $T_{i_1, i_2, \ldots, i_N}$.

Several tensors may be combine to form new tensors through a process of tensor
contraction, for example

\[ T_{i,k,p} = \sum_{j,l,m,p} A_{i,j,k,l} B_{j,m,n,p} C_{l,n}. \]  

(2.5)

Matrix multiplication is a one example of a common tensor contraction. We call many tensors contracted together a tensor network.

A convenient graphical representation can be used to represent tensor contraction. In this representation tensors are nodes in a graph with tensor indices being summed over represented as lines connecting the tensors being contracted. Indices not being summed over are represented as free legs. A demonstration of this graphical notation applied to equation 2.5 can be seen in figure 2.1. The indices being summed over \( j, i, m \) and \( p \) are connecting the tensors being contracted whereas \( i, k \) and \( p \) are free legs, being uncontracted.
To understand the application of tensor network methods to many-body physics, first notice an arbitrary state of $N$ $d$-dimensional particles can be written as:

$$|\psi\rangle = \sum_{\sigma_1, \sigma_2, \ldots, \sigma_N} C_{\sigma_1, \sigma_2, \ldots, \sigma_N} |\sigma_1\rangle \otimes |\sigma_2\rangle \otimes \cdots \otimes |\sigma_N\rangle.$$ 

(2.6)

The array of $d^N$ complex coefficients, $C_{\sigma_1, \sigma_2, \ldots, \sigma_N}$ is an $N$-dimensional tensor. By decomposing $C_{\sigma_1, \sigma_2, \ldots, \sigma_N}$ and writing it as a tensor network constructed from a set of smaller tensors, a more efficient representation of the state can be found.

### 2.3 Properties of matrix product states

![Diagram of matrix product state](image)

Figure 2.2: By parametrising $C_{\sigma_1, \sigma_2, \ldots, \sigma_N}$ as $A^{\sigma_1} A^{\sigma_2} \cdots A^{\sigma_N}$ any quantum wavefunction can be written as a matrix product state.

Matrix product states (MPS) are the most well-known and successful tensor network for one dimensional systems. To construct an MPS a rank-3 tensor is associated with each site of a chain, $A^{\sigma_\alpha}_{i_\alpha j_\alpha}$. One of these indices is a $d$-dimensional index representing the local physical degrees of freedom on site $\alpha$, $\sigma_\alpha$. $i_\alpha$ and $j_\alpha$ are $\chi_\alpha$ and $\chi'_\alpha$ dimensional respectively. These auxillary indices are tensor contracted in the network. $\chi_\alpha$ is often called the bond dimension of the MPS.

A wavefunction can be constructed from $N$ tensors by contracting all of the auxillary
indices together:

\[ |\psi\rangle = \sum_{\sigma_1, \sigma_2, \ldots, \sigma_N} \text{tr}(A^{\sigma_1}A^{\sigma_2} \ldots A^{\sigma_N}) |\sigma_1\rangle \otimes |\sigma_2\rangle \otimes \ldots \otimes |\sigma_N\rangle \]  \hspace{1cm} (2.7)

here we use periodic boundary conditions by contracting the first and last tensors together.

Entanglement of an MPS: In principal every wavefunction can be written as a matrix product state. If \( C_{\sigma_1, \sigma_2, \ldots, \sigma_N} \) is reshaped as a matrix \( C_{(\sigma_1, \ldots, \sigma_{\alpha}), (\sigma_{\alpha+1}, \ldots, \sigma_N)} \) an SVD can be performed. If this process is performed for all \( \alpha \) and the first \( \chi_{\alpha} \) singular values are kept then an MPS is obtained. If the system is divided between site \( \alpha \) and \( \alpha + 1 \) the entanglement between the two halves of the system is therefore \( S_E(\alpha) \leq \log_d \chi_{\alpha} \). Matrix product states naturally satisfy the entanglement area law in one dimension.

Gauge freedom and canonical form: From equation 2.7 it is clear that for a given state \( |\psi\rangle \) there is not a unique choice of tensors \( \{A_{ij}^{\sigma}\} \). In fact, any transformation of the form

\[ A^{\sigma} \rightarrow X_{\alpha} A^{\sigma} (X_{\alpha+1})^{-1} \]  \hspace{1cm} (2.8)

does not alter the state described, provided all \( X_{\alpha} \) are invertible. There is a particularly useful gauge choice that will make calculations more convenient. We write an MPS as a rank-4 tensor \( U^{\sigma, \gamma}_{ij} \) contracted with a reference state \( A^{\sigma}_{ij} = \sum_{\gamma} U^{\sigma, \gamma}_{ij} \phi^{\gamma} \). For simplicity we assume the reference state is of the form \( \phi = (1, 0, \ldots, 0) \).

A convenient result of this representation is that we can group together the indices \( V = U_{(\sigma,i),(\gamma,j)} \) such that they form a unitary and therefore \( V^\dagger V = VV^\dagger = I \). This is convenient because we now find that:

\[ \sum_{\sigma} \bar{A}^{\sigma} A^{\sigma} = I_x \]  \hspace{1cm} (2.9)

An MPS represented in this manner is called left-orthogonal. By instead grouping \( V = }
Figure 2.3: The conditions on $A$ in left-canonical form. When contracted to the left the identity is obtained. When contracted to the right with the environment tensor $\Gamma(\alpha - 1)$ $\Gamma(\alpha)$ is obtained.

$U_{(\sigma,j),(\gamma,i)}$ an equivalent, right-orthogonal form can be achieved. In left-orthogonal form, we also find the following relationship:

$$\sum_\sigma A^{\sigma \alpha} \Gamma_\alpha \bar{A}^{\sigma \alpha} = \Gamma_{\alpha - 1} \quad (2.10)$$

where $\Gamma_\alpha$ is the right environment. Performing a singular value decomposition on $\Gamma_\alpha$ is can be written as $\Gamma_\alpha = U_\alpha \tilde{\Gamma}_\alpha U_\alpha^\dagger$ where $\tilde{\Gamma}_\alpha$ is diagonal. A final gauge transformation can be made to transform to left-canonical form,

$$A^{\sigma \alpha} \rightarrow \tilde{A}^{\sigma \alpha} = U_{\alpha - 1} A^{\sigma \alpha} U_\alpha^\dagger \quad (2.11)$$

$$\sum_\sigma \tilde{A}^{\sigma \alpha} \tilde{A}^{\sigma \alpha} = I_\chi \quad (2.12)$$

$$\sum_\sigma \tilde{A}^{\sigma \alpha} \tilde{\Gamma}_\alpha \tilde{A}^{\sigma \alpha} = \tilde{\Gamma}_{\alpha - 1} \quad (2.13)$$

The left-canonical gauge conditions in tensor graphical notation can be seen in figure 2.3.
**Expectation values**: Using the canonical form of an MPS expectation values can be calculated efficiently. The expectation value of a single site operator $M$ can be calculated using the method depicted in figure 2.4. To the left of the observable the left canonical gauge condition described in equation 2.9 can be used. To the right of the observable the condition described in equation 2.10 may be used. Using these two conditions calculating the expectation value is reduced to the contraction of four tensors.

![Figure 2.4](image.png)

Figure 2.4: The expectation value of local observables can be efficiently calculated with matrix product states.

The expectation value of two operators separated by $n$ sites can also be calculated conveniently using matrix product states. To do so, we introduce the transfer matrix of $A$ defined as

$$E(A) = \sum_\sigma \bar{A}^\sigma \otimes \bar{\bar{A}}^\sigma.$$  \hspace{1cm} (2.14)

Rather than contracting the $n - 1$ $A$ tensors between the two operators, it is more convenient to calculate the transfer matrix to the $(n - 1)$-th power and then insert this between the two operators. An example can be seen in figure 2.5.
In canonical form the largest eigenvalue of the transfer matrix is scaled to be exactly 1. The left and right eigenvector of the transfer matrix associated with this eigenvalue are the left and right environment of the matrix product state, $\mathbb{2}_\chi$ and $\Gamma$ respectively.

![Diagram](image)

Figure 2.5: The expectation value of two single site observables can also be efficiently calculated with matrix product states using the transfer matrix, $E) = \sum_{\sigma} A^\sigma \otimes \bar{A}^\sigma$.

2.4 Manifold and tangent-space methods for matrix product states

Matrix product states have another key feature that we depend upon. If we consider matrix-product states of a fixed bond dimension, they form a special complex Riemannian manifold embedded in Hilbert space called a Kahler manifold \[11\]. This manifold is also a symplectic manifold, possessing a symplectic form. As a consequence of MPS forming a symplectic manifold, they can be thought of as forming a classical phase-space. We can define all typical properties of a classical phase-space including a Poisson bracket. Thinking of MPS as forming a classical phase-space will be central to our investigation of thermalisation in chapter 5.
To utilize the manifold properties of MPS, we need to define an MPS that points tangentially along the manifold at each point. This can be achieved using the unitary representation of the MPS which we have previously introduced. $A_{i,j}^{\sigma}$ can be created by operating onto a reference state with a unitary. By applying the unitary to any orthogonal reference state an MPS tangential to $A_{i,j}^{\sigma}$ will be produced. If the reference state is $\phi^\gamma = \delta_{\gamma,1}$ then the set of states tangential to $A$ can be parametrised as

$$dA_{i,j}^{\sigma}(x) = U_{i,k}^{\sigma,\gamma \neq 1} x_{k,l}^{\gamma \neq 1} \Gamma_{l,j}^{-\frac{1}{2}}$$

(2.15)

where $\Gamma$ is the right environment as previously introduced and $x$ is a general $(d - 1)\chi$ by $\chi$ matrix. This parametrisation was originally introduced by Haegeman et al. [12].

A useful property of $dA$ is that it is orthogonal to $A$ in the following sense:

$$\sum_{\sigma} dA^{\sigma\alpha} A_{\sigma} = 0.$$  

(2.16)

### 2.5 Tree-tensor networks

Matrix product states are not the only useful class of tensor network for one-dimensional systems. Another natural construction is a network with a tree-like shape. These networks emerge naturally from a discussion of entanglement renormalization [13].

Calculating the ground state properties of a Hamiltonian exactly is not always efficient. One solution is to simplify the system, systematically removing short range degrees of freedom while preserving important features. This is called renormalization [14,15]. It can be implemented in real space using an isometry which maps from a lattice $\mathcal{L}$ to a smaller lattice, $\mathcal{L}'$. Mapping from an $N$ site chain to an $\frac{N}{2}$ site chain can be achieved using a rank-3 tensor $\omega_{i,j}^{\sigma\alpha}$ that maps sites $i$ and $j$ to site $\alpha$. This process can be repeated to map from an $\frac{N}{2}$ site chain to an $\frac{N}{4}$ site chain and so on until only a single site remains after $O(\log_2 N)$ layers. The set of tensors that perform this mapping
can be used to define a tree tensor network. An example of a tree tensor network for an eight site chain can be seen in figure 2.6 a).

The tree tensor network’s hierarchical structure is designed to capture short range correlations between neighbouring sites at the bottom of the network with long range correlations only being captured at the top of the network. Unfortunately, tree tensor networks have a fundamental flaw. While some of the correlations between neighbouring sites will be captured at the lowest level, others will not. For example in figure 2.6 a) the correlations between site 3 and 4 will be captured by an isometry at the first layer of the network but those between site 4 and 5 will not be captured until the final layer.

The solution is to modify the network with a new set of tensors which can capture correlations at the appropriate level of the network. Between each isometry tensor $\omega$, a rank-4 tensor $u_{\sigma \alpha, \sigma \beta}$ called a disentangler is introduced. With these additional elements the tree tensor network becomes the multiscale entanglement renormalization ansatz.
(MERA) [16]. An example of a MERA for 8 sites can be seen in figure 2.6 b).

The MERA has been shown to be the appropriate network for describing quantum systems as they undergo phase transitions [13, 17–20]. At these points the spectral gap of the Hamiltonian disappears. For gapless one dimension, the entanglement of a length $l$ subsystem scaling logarithmically, $S_l \propto \log(l)$. Correlation functions also decay algebraically with distance rather than exponentially. MERA’s tree structure means that it naturally possess both of these properties.
Chapter 3

Feynman path integral over matrix product states I: General theory

Path integrals are a powerful tool for studying many-body quantum systems. However, they also have some weaknesses. One of these is that entanglement cannot be captured at the saddle point of a product state path integral. In this chapter a Feynman path integral over matrix product states is introduced to overcome these issues. After considering how to analyse quadratic fluctuations around a product state saddle point we move onto discuss how to calculate quadratic fluctuations around a matrix product state path integral. We then discuss how this can be used to calculate corrections to mean-field value of observables, including the ground state energy. Finally, we discuss how an improved approximation to the ground state can be calculated. Applications of these techniques can be found in chapter 4.

The introductory discussion of the MPS path integral is based upon the work of Green et al. [21]. The explicit construction of the quadratic expansion are presented here for the first time, as are the explicit form for the zero-point fluctuation corrections to the energy and other observables. This chapter is based upon a paper that is in preparation.
3.1 Feynman path integral over product states

Quantum mechanics can be developed in two equivalent ways. The first of these is canonical quantization in which observables are associated with linear operators acting on a complex vector space. This approach seems to ignore many quantities that play a central role in classical mechanics. In particular, Dirac noted that the action - defined in terms of a Lagrangian as \( S = \int L dt \) does not appear to be relevant in canonical quantization [22].

In 1948 Feynman developed the second approach to quantum theory in which the action plays a central role: the Feynman path integral [23]. Feynman considered the probability amplitude of a particle starting at a position \( q_i \) at \( t = 0 \) and finishing at \( q_f \) at \( t = T \): \( G(q_f, q_i; T) = \langle q_f | e^{-iHT/\hbar} | q_i \rangle \). He realised that \( G(q_f, q_i; T) \) could be calculated as a sum over paths

\[
G(q_f, q_i; T) = A \sum_{\text{all paths}} \exp \left[ i \frac{S[q(t)]}{\hbar} \right] \tag{3.1}
\]

where \( A \) is a normalisation factor and \( S[q(t)] \) is the classical action of each path. This formulation of quantum mechanics has a number of appealing features. One of them is that the relationship between classical and quantum dynamics is much clearer. Paths with rapidly fluctuating phase factors \( \exp \left[ i \frac{S[q(t)]}{\hbar} \right] \) will average to zero. The sum in equation 3.1 is therefore dominated by trajectories that minimise the action, these are the classical trajectories given by the Euler-Lagrange equations.

One of the drawbacks of the path integral is related to this. Being a sum over classical trajectories, it can be inconvenient to capture entanglement and the behaviour of systems that depends upon entanglement. We try to overcome this shortcoming by writing a Feynman path integral over trajectories that are already entangled, matrix product states [21].

One of the advantages of the matrix product state field integral is that the true ground state of the system can in principal be found from a perturbative expansion.
around the correct MPS saddle. After defining the matrix product state path integral in section 3.3 we outline how to consider quadratic fluctuations around an MPS saddle point in sections 3.5 and 3.7. We then consider applications of this in chapter 4.

3.2 Constructing the path integral

We will demonstrate the path integral approach by calculating the partition function. The partition function is the central object of study in quantum statistical mechanics:

$$Z = \text{Tr}(e^{-\beta \hat{H}}) = \int dq_i \langle q_i | e^{-\beta \hat{H}} | q_i \rangle$$  \hspace{1cm} (3.2)

From the partition function we can - in principal - calculate all the quantities we need to understand a system. To write the partition function as a path integral we divide $e^{-\beta \hat{H}}$ into $N$ timeslices:

$$e^{-\beta \hat{H}} = (e^{-\frac{\beta}{N} \hat{H}})^N = e^{-\frac{\beta}{N} \hat{H}} I e^{-\frac{\beta}{N} \hat{H}} I ... I e^{-\frac{\beta}{N} \hat{H}}$$  \hspace{1cm} (3.3)

where we have inserted the identity between each timeslice. We can then replace the identity here with the resolution over the identity over a certain set of states. We will use coherent states. A coherent state we mean a state $|\alpha\rangle$ defined as

$$|\alpha\rangle = e^{\alpha \hat{a}^\dagger} |0\rangle,$$  \hspace{1cm} (3.4)

where $\hat{a}^\dagger$ is a bosonic creation operator and $|0\rangle$ is the vacuum of the creation operator. They are eigenvalues of the bosonic annihilation operator $\hat{a} |\alpha\rangle = \alpha |\alpha\rangle$ and the overlap between two coherent states is $\langle \alpha |\alpha\rangle = e^{\alpha^\ast \alpha}$. The resolution of the identity over these states takes the form

$$I = \int \frac{d\alpha d\bar{\alpha}}{2\pi i} e^{-\alpha \bar{\alpha}} |\alpha\rangle \langle \alpha|.$$  \hspace{1cm} (3.5)
Their nonzero overlap demonstrates that coherent states form an overcomplete basis. Inserting this resolution of the identity into the path integral we find the partition function takes the following form:

$$Z = \int \left( \prod_j \frac{d\alpha_j d\bar{\alpha}_j}{2\pi i} \right) e^{-\sum_j \alpha_j \bar{\alpha}_j} \prod_j \langle \alpha_{j+1} | e^{-\frac{\beta}{N} \hat{H}} | \alpha_j \rangle.$$  \hspace{1cm} (3.6)

In the large $N$ limit, we can write

$$\langle \alpha_{j+1} | e^{-\frac{\beta}{N} H(a, a^\dagger)} | \alpha_j \rangle \approx \langle \alpha_j | (1 - \frac{\beta}{N} H(\alpha_j, \bar{\alpha}_{j+1})) e^{a_j \bar{\alpha}_{j+1}} (1 - \frac{\beta}{N} H(\alpha_j, \bar{\alpha}_{j+1})) \rangle.$$  \hspace{1cm} (3.7)

Inserting this back into equation (3.6) we find

$$Z = \int \left( \prod_j \frac{d\alpha_j d\bar{\alpha}_j}{2\pi i} \right) e^{\sum_j \alpha_j \bar{\alpha}_j + \alpha_j \bar{\alpha}_{j+1}} \prod_j \left( 1 - \frac{\beta}{N} H(\alpha_j, \bar{\alpha}_{j+1}) \right).$$  \hspace{1cm} (3.8)

Finally, the partition function can be extracted by taking the continuum limit of this, $N \to \infty$. In this limit $1 - \frac{\beta}{N} H(\alpha_j, \bar{\alpha}_{j+1})$ can be exponentiated to find

$$Z = \int D\alpha D\bar{\alpha} e^{-\int_0^\beta d\tau \{ a \partial_\tau - \bar{\alpha} \partial_\tau - a - H(a, \bar{\alpha}) \}}.$$  \hspace{1cm} (3.9)

### 3.3 Feynman path integral over matrix product states

We now take a new approach to constructing the Feynman path integral. Instead of using coherent states, we write the path integral over the matrix product states introduced in chapter 2. This was first introduced in [21].

An MPS path integral can be constructed precisely as in the previous section but
with a resolution of the identity now written over matrix product states:

$$\mathbb{I} = \int dA \langle \psi(A) \rangle \langle \psi(A) \rangle.$$  \hfill (3.10)

Inserting this resolution of the identity we have a path integral over matrix-product states:

$$Z = \text{Tr}(e^{-\beta \mathcal{H}}) = \int DA(\tau) \exp \left[ - \int_0^\beta \langle \psi(A) \rangle \frac{d}{d\tau} |\psi(A)\rangle - \langle \psi(A) | \mathcal{H} | \psi(A) \rangle \right], \hfill (3.11)$$

It is not entirely clear what the measure $DA$ being integrated means in this case. We can take advantage of the construction of a matrix product state as a unitary matrix operating on a reference state in section 2.4. Using this construction the Haar measure can be chosen \cite{24}. The Haar measure is an unbiased over the set of unitary matrices.

The second feature we want to guarantee is the *locality* of the field integral. Calculating expectation values using matrix product states involves performing a contraction along the chain. At first glance the MPS field integral is therefore not local. Using the left canonical construction in chapter 2 the expectation of a local operator will not depend on tensors to the left. To the right the non-local contributions can be described by the *environmental tensor*

$$\Gamma_{i-1} = \sum_\alpha A_\alpha^i \Gamma_i \bar{A}_\alpha^i. \hfill (3.12)$$

Locality can therefore be ensured by modifying the field integral:

$$Z = \int DA(\tau) D\Gamma(\tau) \delta \left[ \Gamma_{i-1} - \sum_\alpha A_\alpha^i \Gamma_i \bar{A}_\alpha^i \right] \exp \left[ - \int_0^\beta \langle \psi(\bar{A}) | \dot{\psi}(A) \rangle - \langle \psi(\bar{A}) | \mathcal{H} | \psi(A) \rangle \right] \hfill (3.13)$$

Having made these modifications this path integral can now be treated using standard techniques.
**Adiabatic continuity:** In what circumstances will a higher bond dimension field theory be useful? Often the true ground state of the system can be obtained by a perturbative dressing of a low bond order saddle. This means the saddle is *adiabatically connected* to its ground state.

In some cases the true ground state of a system cannot be adiabatically connected to a product state saddle. For example, the ground state may be require tunnelling between different product state saddle point configurations. This requires non-perturbative *instantons*. Matrix product states are already a sum over product states, they can directly capture physics at the saddle point that requires resummations of instantons in a product state field integral.

A simple example of this exists even in the case of two spins. Suppose the ground state of the system was an entangled state, such as \( \alpha |\uparrow_1,\uparrow_2\rangle + \beta |\downarrow_1,\downarrow_2\rangle \). A product state path integral could not capture this at the saddle point.

**Euler-Lagrange equations:** What are the Euler-Lagrange equations for the MPS field integral defined in equation (3.13)? Requiring the action \( S[A,\bar{A}] \) is stationary with respect to a variation \( \bar{A} \rightarrow \bar{A} + d\bar{A} \) results in the following Euler-Lagrange equations

\[
\langle \partial_{\bar{A}} \psi | \partial_{A} \psi \rangle \hat{A}^i = -\langle \partial_{\bar{A}} \psi | \hat{H} | \psi \rangle.
\]  

(3.14)

This is the *time-dependent variational principal*, originally introduced by Dirac [25]. It was first derived for matrix product states by Haegeman et al. [12].

### 3.4 Fluctuations around a product state saddle point

Before studying fluctuations around a quadratic MPS field theory, we will consider the example of the \( \frac{1}{2} \) Heisenberg antiferromagnet

\[
\mathcal{H} = J \sum_i \sigma_i^z \sigma_{i+1}^z.
\]

(3.15)
where $\sigma = \{\sigma_x, \sigma_y, \sigma_z\}$ are the Pauli matrices. We study the Heisenberg antiferromagnet using a product state path integral defined in terms of spin coherent states $|L_i\rangle$ defined in terms of an $O(3)$ $L_i$ that points in the direction in which the spin on site $i$ is polarized.

In terms of spin coherent states the expectation value of $H$ is

$$H[L] = J \sum_i L_i L_{i+1}$$

(3.16)

We introduce spin-wave modes by writing $L_i$ in terms of fluctuations $m_i$ around a saddle-point value $l_i$,

$$L_i = l_i \left(1 - \frac{1}{2}|m|^2 \right) + m_i$$

(3.17)

where $m_i, l_i = 0$. For the antiferromagnet, we expand around the Néel state, $l_i = -l_{i+1}$.

If $\hat{\theta}_i$ and $\hat{\phi}_i$ are the unit vectors in the azimuthal and longitudinal directions at $l_i$ then $m_i$ can be written $m_i = m_1 \hat{\theta}_i + m_2 \hat{\phi}_i$. The spin-waves around $l_i$ can be written in terms of Bose coherent state variables by performing a change of variables $z_i = m_1 + im_2$.

Having performed this change of variables, equation (3.16) becomes

$$H = \sum_i H[l] - \frac{J}{2} \sum_k \left(\begin{array}{c} z_k^* \\ z_{-k}^* \end{array}\right)^T \left(\begin{array}{cc} 1 & \cos(k) \\ \cos(k) & 1 \end{array}\right) \left(\begin{array}{c} z_k \\ z_{-k} \end{array}\right)$$

(3.18)

This coherent state Hamiltonian can be used to define an equivalent second quantized Hamiltonian. We associated $z_k$ and $z_k^*$ with creation and annihilation operators $\hat{b}_k$ and $\hat{b}_k^\dagger$ respectively:

$$\mathcal{H} = \sum_i H[l] - \frac{J}{2} \sum_k \left(\begin{array}{c} \hat{b}_k^\dagger \\ \hat{b}_{-k}^\dagger \end{array}\right)^T \left(\begin{array}{cc} 1 & \cos(k) \\ \cos(k) & 1 \end{array}\right) \left(\begin{array}{c} \hat{b}_k \\ \hat{b}_{-k} \end{array}\right) + E'_0$$

(3.19)

A constant factor $E'_0$ must be introduced due to the ordering ambiguity when transforming from coherent state variables to second quantized operators. We address this
ambiguity below.

The Hamiltonian in equation 3.19 can now be diagonalised using a Bogoliubov transformation. We spend some time discussing this in detail as it will be important in sections 3.5, 3.7 and 3.8. The Bogoliubov transformation is a linear combination of creation and annihilation operators chosen to satisfy two criteria, firstly it must diagonalise the Hamiltonian and secondly, it must preserve the commutation relations of the bosonic operators. It takes the form:

\[
\hat{\beta}_k = u_k \hat{b}_k - v_k \hat{b}^\dagger_{-k}.
\] (3.20)

The restriction that it must preserve the bosonic commutation relations requires \(|u_k|^2 - |v_k|^2 = 1\). This restriction means the useful parametrisation \(u_k = \cosh \theta_k\) and \(v_k = \sinh \theta_k\) may be used. It is important to note that the transformation applied to the bosonic operators is not unitary, rather it is an example of a symplectic transformation.

The Hamiltonian in equation 3.19 can be diagonalised by requiring

\[
\tanh 2\theta_k = -\cos(k).
\] (3.21)

The Hamiltonian can now be written as

\[
\mathcal{H} = \sum_i H[i] + \sum_k \omega_k \left( \hat{\beta}_k^\dagger \hat{\beta}_k + \frac{1}{2} \right) + E'_0
\] (3.22)

where \(\omega_k = \sqrt{1 - \cos^2 k}\). We can estimate the change to the ground-state energy as a result of the spin-wave fluctuations. However, first we must address the \(E'_0\) term. Is it possible to determine what value this takes?

One approach would be to work directly with operators. For spin-waves this would mean performing a Holstein-Primakoff expansion of the Pauli operators in order to determine \(E'_0\) [26]. Fortunately, a second approach exists that does not depend upon working in terms of operators. The transformation \(\hat{b}_k \rightarrow \hat{\beta}_k = u_k \hat{b}_k - v_k \hat{b}^\dagger_{-k}\) can be expressed as
a transformation between coherent state variables $z_k \rightarrow \eta_k = u_k z_k - v_k z_{-k}^\dagger$. Blasone et al. have shown how a path integral written in terms of $z_k$ coherent state variables can consistently be written in terms of $\eta_k$.

In this framework $E'_0$ can be shown to take the form

$$E'_0 = \frac{\partial^2 H(\eta, \bar{\eta})}{\partial \eta \partial \bar{\eta}} \left[ \frac{\partial \eta}{\partial \bar{z}} \frac{\partial \bar{\eta}}{\partial z} + \frac{\partial \eta}{\partial z} \frac{\partial \bar{\eta}}{\partial \bar{z}} \right] = \frac{\partial^2 H(z, \bar{z})}{\partial z \partial \bar{z}}$$

(3.23)

Using this result, the zero-point energy contribution to the ground-state energy of the Heisenberg antiferromagnet due to spin-waves is

$$\Delta E_0 = \frac{1}{2} \sum_k (\omega_k - 1).$$

(3.24)

Since this correction is negative, it is clear that the Néel state is not the true ground state of a system. A new estimate for the ground state can be made, the vacuum of the $\beta_k$ bosons. This is the Néel state dressed with spin wave fluctuations:

$$|\Psi_0\rangle = \exp \left[ \frac{1}{2} \sum_k \tanh \theta_k b_k^\dagger b_{-k}^\dagger \right] |\psi_0\rangle.$$

(3.25)

It is important to note that this correction to the Néel state differs from significantly from finding a higher bond dimension approximation to a ground state using a matrix product state algorithm. The dressed state is in principle an infinite dimensional matrix product state. This is a useful demonstration of the difference between using a higher dimensional matrix product state saddle point compared to calculating fluctuation corrections to a matrix product state of a fixed bond dimension.

### 3.5 Fluctuations around a matrix product saddle point

We now turn to calculating fluctuations around a matrix-product saddle point. In the previous section the spin coherent state on site $i$, $L_i$ was written as variations around

\[31\]
a saddle point value $\mathbf{l}_i$. For the antiferromagnetic Heisenberg model we chose the Néel state as our saddle point. The Néel state is the best product state approximate to the ground state of the antiferromagnet but it can be improved by dressing it with spin-wave excitations. In this chapter a matrix-product state on site $i$ will be written in terms of variations around an MPS saddle $A_0^\sigma$. The saddle point MPS $A_0^\sigma$ could be determined through an analytical approach or numerically by using algorithms such the imaginary time time-dependent variational principal defined in equation 3.14. $A_0^\sigma$ is the best mean-field approximation to the ground state of a Hamiltonian on the variational manifold of bond dimension $\chi$ MPS. We improve $A_0^\sigma$ by dressing it with quadratic fluctuations. A graphical depiction of this can be seen in figure 3.2: the new ground state is a sum over states in the vicinity of $A_0^\sigma$ on the manifold of $\chi$-dimensional MPS.

Variations will be parametrised in terms of the tangent vector introduced in equation
\[ A = A_0 + V x_i e^{i \theta} - \frac{1}{2} A_0 r^{-1/2} x_i r^{-1/2} x_i^\dagger r^{-1/2} + -\frac{1}{2} \]

Figure 3.2: A graphical representation of the parametrisation we use to expand to quadratic order in fluctuations around an MPS saddle \( A_0 \).

In close analogy with equation 3.17, \( A^{\sigma_i} \) is written as:

\[ A^{\sigma_i}(x_i) = A_0^{\sigma_i}\left[1 - \frac{1}{2} \Gamma^{-\frac{1}{2}} x_i^\dagger x_i \Gamma^{-\frac{1}{2}}\right] + V x_i \Gamma^{-\frac{1}{2}}. \tag{3.26} \]

Here \( \Gamma \) is the right environment as defined in equation 2.10. To consistently expand to quadratic order around an MPS saddle it is necessary to include the normalisation term \( \Gamma^{-\frac{1}{2}} x_i^\dagger x_i \Gamma^{-\frac{1}{2}} \). A representation of this state in tensor graphical notation can be seen in figure 3.2.

The expectation of a Hamiltonian with respect to a matrix product state parametrised as in equation 3.26 is

\[ \langle \psi(A^{\sigma_i}) | H | \psi(A^{\sigma_i}) \rangle = \langle \psi(A_0) | H | \psi(A_0) \rangle + \sum_k \begin{pmatrix} x_k^* \\ x_{-k} \end{pmatrix}^T \begin{pmatrix} \epsilon_k(H, A_0) & \Delta_k(H, A_0) \\ \Delta_k^T(H, A_0) & \epsilon_k^T(H, A_0) \end{pmatrix} \begin{pmatrix} x_k \\ x_{-k}^* \end{pmatrix} \]

\[ \begin{array}{c} \text{M}_k(H, A_0) \end{array} \tag{3.27} \]

where \( \epsilon_k(H, A_0) = \langle \partial_{A_0^\sigma_k} \psi | H | \partial_{A_0^\sigma_k} \psi \rangle \) and \( \Delta_k(H, A_0) = \langle \partial_{A_0^\sigma_{-k}} \partial_{A_0^\sigma_k} \psi | H | \psi \rangle \). Due to the fact that \( \epsilon_k(H, A_0) = \epsilon_k^T(H, A_0) \) and \( \Delta_k(H, A_0) = \Delta_k^T(H, A_0) \) we find \( M_k(H, A_0) \) is a Hermitian matrix. \( \epsilon_k(H, A_0) \) is the projected Hamiltonian originally introduced by Haegeman et al. [28].

We can similarly expand the Berry phase term of the action. Noting that \( \dot{A}_0 = 0 \) at
the saddle point, the Berry phase is

$$\sum_k \text{tr}[x_k \dot{x}_k]$$ (3.28)

This reduces to the $CP_1$ case when $\chi = 1$, $\sum_k x_k \dot{x}_k$.

3.6 Calculating $M_k(\mathcal{H}, A_0)$

We will now give an overview of how to calculate $M_k(\mathcal{H}, A_0)$. In this section we will explain some of the important techniques required to make this calculation tractable. Full details of all of the tensor diagrams required to calculate $M_k(\mathcal{H}, A_0)$ can be found in appendix F.

3.6.1 Calculating $\epsilon_k(\mathcal{H}, A_0)$

Determining $\epsilon_k(\mathcal{H}, A_0)$ involves contracting a large number of tensor diagrams. These all take the following form: they are a sum over diagrams with a tangent MPS $dA$ at site $a$, a Hamiltonian at site $c$ and $c + 1$ and a second tangent MPS $\bar{d}A$ at site $b$. A graphical representation of this can be seen in figure 3.3.
To calculate $\epsilon_k(H, A_0)$, we note that $|\partial A^\sigma_k \psi\rangle$ can be parametrised as

$$|\partial A^\sigma_k \psi\rangle = \sum_n e^{ikn} \sum_{\{\sigma_i\}} \text{tr} \left[ ... A^{\sigma_{n-1}}_0 d A^{\sigma_n} A^{\sigma_{n+1}}_0 ... \right] |\sigma\rangle,$$

where $dA^\sigma_n = V^{\sigma,\gamma\neq 1} x \Gamma^{-\frac{1}{2}}$ as in equation 2.15. The calculation of $\epsilon_k(H, A_0)$ is simplified by using properties of $dA$. In chapter 2 we noted that $dA$ is orthogonal to $A$:

$$\sum_{\sigma} dA^\sigma A^\sigma = 0. \tag{3.30}$$

As a consequence, all diagrams in which $a < c$ or $b < c$ are automatically zero. $\epsilon_k(H, A_0)$ still involves terms with $dA_a$ or $dA_b$ far to the right of the Hamiltonian. These will be simplified by summing powers of the transfer matrix $E(A_0) = \sum_{\sigma} A^\sigma_0 \otimes A^\sigma_0$. In chapter 2 we noted that the transfer matrix’s largest eigenvalue is $\lambda = 1$, with left and right eigenvectors, $I_\chi$ and $\Gamma$ respectively. $E^n$ can be decomposed as

$$E^n = P_{\lambda=1} + Q(QEQ)^nQ, \tag{3.31}$$

where $P_{\lambda=1}$ projects onto the eigenvalue $\lambda = 1$ subspace of $E$ and $Q = I - P_{\lambda=1}$ is its compliment. As $I_\chi$ is the $\lambda = 1$ left eigenvalue, the contribution of $P_{\lambda=1}$ vanishes when calculating $\epsilon_k(H, A_0)$ due to equation 3.30. As $Q(QEQ)^nQ$ now has a spectral radius smaller than one and may be summed, its contribution can be calculated:

$$\sum_{n=0}^\infty Q(e^{ikQEQ})^nQ = Q(I - e^{ikQEQ})^{-1}Q = T_k. \tag{3.32}$$

Using this result, the diagrams in figure 3.3 can all be determined. As an example, tensor diagrams in which $b = c$ and $a > c + 1$ can be seen in figure 3.4. The form of all of the non-zero diagrams from figure 3.3 can be found in appendix F, section F.1.

The quadratic order normalisation terms in $A^\sigma$, $A_N = -\frac{1}{2} A^\sigma_0 \Gamma^{-\frac{1}{2}} x_i^\dagger x_i \Gamma^{-\frac{1}{2}}$ also con-
Figure 3.4: An example of one of the tensor diagrams which contribute to $\epsilon_k(\mathcal{H}, A_0)$. These are tensor diagrams with either $A_N$ at a site $a$ or $\bar{A}_N$ at a site $b$. They are calculated using the tools introduced above. It is important to note that these diagrams are not momentum-dependent. The contributions to $\epsilon_k(\mathcal{H}, A_0)$ due to $A_N$ can be found in section F.2.

3.6.2 Calculating $\Delta_k(\mathcal{H}, A_0)$

To calculate $\Delta_k(\mathcal{H}, A_0) = \langle \partial_{A_0^a} \partial_{A_0^{a'}} \psi | \mathcal{H} | \psi \rangle$ we parametrise $|\partial_{A_0^a} \partial_{A_0^{a'}} \psi \rangle$ as

$$|\partial_{A_0^a} \partial_{A_0^{a'}} \psi \rangle = \sum_{n=-\infty}^{\infty} e^{-ik_1 A_0^{a} A_0^{a'}} \sum_{\sigma_1} \langle . . A_{0_0} \sigma_{n+1} A_{0_0} \sigma_{n+1} . . A_{0_0} \sigma_{m+1} A_{0_0} \sigma_{m+1} . . \rangle |\sigma\rangle$$

(3.33)

$\Delta_k(\mathcal{H}, A_0)$ can be calculated by summing diagrams with the following form: $dA$ at site $a$, a Hamiltonian at $c$ and $c+1$ and another $dA$ at site $b$. A graphical representation can be seen in figure 3.5. As with $\epsilon_k(\mathcal{H}, A_0)$, this calculation can be simplified by using the orthogonal condition for tangent vectors, $\sum_{a} dA_{a} = 0$. Once again, all diagrams with $a < c$ or $b < c$ are zero automatically due to this orthogonality. The complete set of tensor diagrams which determine $\Delta_k(\mathcal{H}, A_0)$ can be found in section F.3.
3.7 Diagonalising the fluctuation around an MPS saddle point

We now explain how to calculate the correction to the zero-point energy for a matrix product state field theory. Similarly to the antiferromagnet example in section 3.4, we need to diagonalise the quadratic fluctuations defined by $M_k(\mathcal{H}, A_0)$. For the sake of clarity we write $M_k(\mathcal{H}, A_0)$ as $M$ in this section.

As in the product state case, the appropriate matrix for diagonalising $M$ must preserve the commutation relations of the bosonic operators $\hat{b}_k$ associated with the MPS path integral variables $x_k$. If we group together the creation and annihilation operators like $B = \left(\hat{b}_k, \hat{b}_k^\dagger\right)$ we find that the commutation relations take the form

$$[B_i, B_j^\dagger] = \Omega_{ij} \quad \text{(3.34)}$$

where

$$\Omega = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad \text{(3.35)}$$

If the bosonic creation and annihilation operators transform like $A = \left(\hat{a}_k, \hat{a}_k^\dagger\right) = SB$ then to preserve the commutation relations $S$ must satisfy

$$S^\dagger \Omega S = \Omega. \quad \text{(3.36)}$$

Figure 3.5: A graphical representation of the tensor diagrams required to calculate $\Delta_k(\mathcal{H}, A_0)$. 

\[ \sum_{a=-\infty}^{a=b-1} \sum_{b=a+1}^{\infty} \sum_{c=-\infty}^{\infty} e^{ika} e^{-ik'b} \]
We now construct an $S$ which satisfies equation \[3.36\] and diagonalises $M$:

\[S^\dagger MS = \text{diag}(\omega_1, ..., \omega_N, \omega_1, ..., \omega_N) = \omega. \tag{3.37}\]

Conveniently, $\omega$ can be calculated without constructing $S$. The equations of motion for $(x_k, x_k^*)$ can be found using equations \[3.27\] and \[3.28\]. $\omega$ can then be calculated using the following relation:

\[\text{eig}(\Omega M) = \Omega \omega. \tag{3.38}\]

Having calculated $\omega$, $S$ can be written without loss of generality as

\[S = M^{-\frac{1}{2}}U\omega^{\frac{1}{2}}. \tag{3.39}\]

Written in this form, $S$ already satisfies equation \[3.37\] but $U$ must be fixed to satisfy equation \[3.36\]. Substituting equation \[3.39\] into equation \[3.36\] we find:

\[U^\dagger(M^\dagger)^{-\frac{1}{2}}\Omega M^{-\frac{1}{2}}U = \omega^{-\frac{1}{2}}\Omega \omega^{\frac{1}{2}}. \tag{3.40}\]

By noticing that the right hand side of this equation is diagonal we see that $U$ is the unitary that diagonalises $(M^\dagger)^{-\frac{1}{2}}\Omega M^{-\frac{1}{2}}$.

### 3.8 Properties of fluctuations around an MPS saddle point

**Corrections to observables:** Using this construction, the ground state energy can be calculated as in section \[3.4\]

\[E_{GS} = E_0 + \Delta E = \langle \psi(A_0)|H|\psi(A_0)\rangle + \frac{1}{2} \sum_k [\text{tr}(\omega_k) - \text{tr}(M_k(H, A_0))]. \tag{3.41}\]
While we do not construct the state explicitly, this is the energy of a new approximation to the ground state. This ground state is the vacuum of the Bogoliubov transformed bosonic operators which diagonalise $M_k(\mathcal{H}, A_0)$. Corrections to other observables can be calculated by expressing them in terms of the Bogoliubov transformed observables and noting that the new ground state is the vacuum for these bosons.

\[
O_{GS} = \langle \psi(A_0)|O|\psi(A_0)\rangle + \frac{1}{2} \sum_k [\text{tr}(S_k^\dagger M_k(O, A_0)S_k - M_k(O, A_0))] \quad (3.42)
\]

**A new ground state:** While $A_0$ is the MPS that minimises the mean-field energy, it may not minimise $E_{GS}$ as defined in equation 3.41. Taking into account fluctuations $A_0$ may be shifted $A_0 \rightarrow \tilde{A}_0$. The new state, $\tilde{A}_0$ can be calculating through an imaginary time evolution algorithm:

\[
\dot{A}^{\sigma_i} = -\partial_{A^{\sigma_i}} E_{GS}. \quad (3.43)
\]

**Existence of Bogoliubov transformation:** Under what circumstances can fluctuations around a saddle be consistently calculated? It is required that $M_k(\mathcal{H}, A_0)$ is positive-definite for a consistent Bogoliubov transformation to exist. If $M_k(\mathcal{H}, A_0)$ isn’t positive definite then $M_k(\mathcal{H}, A_0)^{-\frac{1}{2}}$ ceases to be Hermitian and no Bogoliubov transformation can be found. As an example consider the Hamiltonian

\[
\hat{\mathcal{H}} = \sum_k \begin{pmatrix} \hat{b}_{k}^\dagger \\ \hat{b}_{-k} \end{pmatrix}^T \begin{pmatrix} \epsilon & \Delta \\ \Delta & \epsilon \end{pmatrix} \begin{pmatrix} \hat{b}_{k} \\ \hat{b}_{-k}^\dagger \end{pmatrix}. \quad (3.44)
\]

A Bogoliubov transformation $\hat{\beta}_k = \cosh \theta_k \hat{b}_k - \sinh \theta_k \hat{b}_k^\dagger$ can diagonalise the Hamiltonian by requiring $\tanh 2\theta_k = \frac{\Delta}{\epsilon}$. If $\Delta > \epsilon$ then no Bogoliubov transformation can be found.

This happens in situations in which the expansion is carried out around an incorrect saddle point. Suppose we use imaginary time-evolution TDVP to find a uniform MPS approximation to the ground state. In this case the $k = 0$ fluctuation Hamiltonian is
guaranteed to be positive definite, unfortunately there is no such guarantee at $k \neq 0$. An obvious example is the case of the Heisenberg anti-ferromagnet in which a purely uniform state is not a reasonable saddle point to expand around.
Chapter 4

Feynman path integral over matrix product states II: Applications

We now apply the MPS path integral techniques introduced in chapter 3. The majority of this chapter focuses on the MPS path integral applied to the transverse-field Ising model. The bilinear-biquadratic $S = 1$ antiferromagnetic Heisenberg model is also studied. In section 4.2.1 the transverse-field Ising model is studied using an analytical $\chi = 3$ MPS ansatz. The ansatz significantly outperforms product states at mean field, predicting a critical point of $h_c \approx 1.709$ compared to $h_c = 2$ for product states. In section 4.2.2 spin-wave fluctuations around this MPS ansatz are studied. Finally in section 4.3 general MPS fluctuations are considered. The regime of validity for this approach is examined using the bilinear-biquadratic $S = 1$ antiferromagnetic Heisenberg model. Fluctuation corrections are then calculated for this model and the transverse-field Ising model.

The MPS ansatz used to study the transverse field Ising model was introduced by Green et al. [29]. The calculations of corrections to observables due to zero-point fluctuations are presented here for the first time. The MPS saddle points calculated numerically were found using TDVP code written by James Morley. This work is based upon a paper that is in preparation.
4.1 Models studied

Having introduced the machinery for studying fluctuations around an MPS saddle point in chapter 4, we apply these methods in this chapter. In the majority of this chapter we study the transverse-field Ising model, an exactly solvable model suitable for benchmarking the techniques previously introduced. We also apply the techniques to the bilinear-biquadratic, spin-1 Heisenberg antiferromagnet.

4.1.1 The transverse-field Ising model

The transverse-field Ising model takes the following form:

\[ \mathcal{H} = - \sum_i \sigma_i^x \sigma_{i+1}^x + h \sigma_i^x. \] (4.1)

The transverse-field Ising model is interesting due to the competition between its two terms. In the limit \( h \to \infty \), the model has a unique groundstate \( |\psi_0\rangle = \bigotimes_i \frac{1}{\sqrt{2}} (|\uparrow\rangle + |\downarrow\rangle). \) In the opposite limit \( h \to 0 \), there are two possible groundstates, \( |\psi_0\rangle = \bigotimes_i |\uparrow\rangle \) or
$|\psi_0\rangle = \bigotimes_i |\downarrow\rangle$. The $\sigma_z \rightarrow -\sigma_z$ symmetry of the Hamiltonian has been *spontaneously broken*. It is not possible to continuously vary a unique, symmetric groundstate in the $h \rightarrow \infty$ limit into two, symmetry breaking groundstates when $h = 0$. Therefore between these two extremes there must be a *quantum phase transition*.

In the ferromagnetically ordered phase, there is a non-zero magnetisation, $M = \langle \sigma^z \rangle$. As the system approaches the critical point this magnetisation reduces until it vanishes at $h_c$. The excitations in this phase are domain walls. When $h$ is larger than $h_c$ the ground state is no longer magnetically ordered. In this phase the excitations are spin flips. The tranverse-field Ising model can be mapped to free fermions using a Jordon-Wigner transformation. The critical point of the model can be calculated to be at $h_c = 1$ exactly. A phase diagram of the transverse-field Ising model can be seen in figure 4.1a.

### 4.1.2 The bilinear-biquadratic, spin-1 Heisenberg antiferromagnet

We also consider fluctuation corrections to the bilinear, biquadratic, spin-1 Heisenberg antiferromagnet:

$$H_{BB} = J \sum_i \cos \theta (S_i \cdot S_{i+1}) + \sin \theta (S_i \cdot S_{i+1})^2. \tag{4.2}$$

This model has a number of interesting phases and phase transitions as $\theta$ is varied, as can be seen in figure 4.1b. We study the model in the *Haldane phase*, which exists for $\theta \in (-\frac{\pi}{4}, \frac{\pi}{4})$. In particular, we consider fluctuation corrections to a matrix product state saddle point in the vicinity of the AKLT point, $\tan \theta = \frac{1}{3}$. At this point the ground state of the system can be exactly represented by an bond dimension $\chi = 2$ matrix product state $[30][31]$. 

43
4.2 Studying the transverse-field Ising model with an MPS ansatz

4.2.1 The saddle point of the transverse-field Ising model.

Figure 4.2: In an MPS ansatz was introduced which can be used to study spin $\frac{1}{2}$ chains. It is defined in terms of $l_i$, an $O(3)$ vector which describes local spin order, $-l_i$ which is an $O(3)$ vector pointing in the opposite direction and $n_i$ an $SU(2)$ spinor which captures nearest neighbour entanglement.

We now begin to study the transverse-field Ising model using the MPS field integral. It is possible to work directly with the field integral as defined in equation 3.13 - and we consider this in section 4.3 - but we first use restricted MPS ansatz introduced by A. G. Green et al. [29]. This ansatz takes the following form:

$$A_i^\dagger = \begin{pmatrix} 0 & n_{i-1}^1 \langle \uparrow | l_i \rangle & n_{i-1}^2 \langle \uparrow | -l_i \rangle \\ \langle \uparrow | l_i \rangle & 0 & 0 \\ \langle \uparrow | -l_i \rangle & 0 & 0 \end{pmatrix}, \quad (4.3)$$
\[ A_i^+ = \begin{pmatrix} 0 & n_{i-1}^1 \langle \downarrow | l_i \rangle & n_{i-1}^2 \langle \downarrow | -l_i \rangle \\ \langle \downarrow | l_i \rangle & 0 & 0 \\ \langle \downarrow | -l_i \rangle & 0 & 0 \end{pmatrix} \] \tag{4.4}

This MPS ansatz has two variables associated with each site of the chain. An \(O(3)\) vector \(l_i\) which describes the direction of the local spin on site and an \(SU(2)\) spinor \(n_i\) describes the entanglement between site \(i - 1\) and site \(i\). \(n_i\) can be mapped to an \(O(3)\) vector, \(n_i = n_i^\dagger \sigma n_i\). When \(n_z = 1\) there is no entanglement between neighbouring sites. When \(n_z = 0\) the ansatz describes a singlet decoration of the one-dimensional chain. An illustration of how to interpret this ansatz can be seen in figure 4.2.

This MPS is in left-canonical form. Its right-environment tensor takes the form

\[ \Gamma_i = \begin{pmatrix} \lambda_i & 0 & 0 \\ 0 & (1 - \lambda_i)|n_i^1| & 0 \\ 0 & 0 & (1 - \lambda_i)|n_i^2| \end{pmatrix}, \tag{4.5} \]

where \(\lambda_i = \lambda\) on even sites and \(\lambda_i = 1 - \lambda\) on odd sites. This variable chooses the weighting of the odd and even singlet tilings of the chain.

To begin, we consider the mean field case of the ansatz, \(l_i = 1, n_i = n, \lambda_i = \frac{1}{2}\). The Hamiltonian takes the form

\[ H_{mf} = hx - J(zz) = \frac{1}{2} \left[ 2hnz^x + J(1 + n_z^2)l_z^2 + Jnz^x l_z^2 \right] , \tag{4.6} \]

where \(l_z^2 = 1 - l_z^2\). The behaviour of \(l_z\) and \(n_x\) as a function of transverse field can be seen in figure 4.3a and fig 4.3b. The local spin degree of freedom drives the transition from the ordered to disordered phase. \(n_x\) reaches a maxima at the critical point, indicating the uniform MPS reaches its maximal entanglement at the critical point, as one would expect. A comparison of the magnetisation can be seen in figure 4.4. The uniform MPS ansatz predicts the critical point appears at \(h_c = \frac{1}{2} \sqrt{\frac{1}{2} (11 + 3\sqrt{17})} \approx 1.709\), a
$l_z$ is the parameter of our chosen uniform MPS ansatz which determines the local spin in the $z$ direction. It approaches zero as the transverse-field is increased and reaches zero at the critical point.

$n_x$ is the parameter which controls the entanglement between neighbouring sites in the MPS ansatz defined in equation 4.3. When $n_x = 1$ the ansatz is maximally entangled. Here $n_x$ is shown as a function of the transverse-field $h$ for the transverse-field Ising model, reaching a maximum at $h_c$.

(a) $l_z$ is the parameter of our chosen uniform MPS ansatz which determines the local spin in the $z$ direction. It approaches zero as the transverse-field is increased and reaches zero at the critical point.

(b) $n_x$ is the parameter which control the entanglement between neighbouring sites in the MPS ansatz defined in equation 4.3. When $n_x = 1$ the ansatz is maximally entangled. Here $n_x$ is shown as a function of the transverse-field $h$ for the transverse-field Ising model, reaching a maximum at $h_c$.

Significant improvement over the product state ansatz which predicts it appears at $h_c = 2$. Fluctuations have the potential to further improve upon this.

4.2.2 Spin-wave fluctuations around an MPS ansatz saddle

Obtaining the zero-point energy.

Having considered the potential of using an MPS ansatz at mean field, we now move onto discussing fluctuations. We begin this by restricting the set of fluctuations we consider and consider spin-wave fluctuations as they were introduced for the Heisenberg antiferromagnet in section 3.4.

The spin-wave fluctuations are described by taking a uniform matrix product state approximation for the ground-state $|\psi(A)\rangle$ and applying a local unitary rotation to each site:

$$|\psi(A)\rangle \rightarrow \prod_i e^{i\sigma_i \cdot m_i} |\psi(A)\rangle$$

with $\langle \sigma \rangle \cdot m_i = 0$. Linking this to the discussion in section 3.5 we note by expanding the unitary to quadratic order we find $dA_{\alpha,\beta}^{\gamma} = i(\sigma \cdot m_\eta)_{\gamma_{\alpha},\gamma_{\beta}} A_{\alpha,\beta}^{\gamma}$. The requirement
Figure 4.4: The magnetisation $\langle \sigma_z \rangle$ is the order parameter of the transverse-field Ising model. Our chosen uniform MPS ansatz predicts the transition is at $h_c \approx 1.709$, significantly closer to the exact result of $h_c = 1$ than the mean-field result $h_c = 2$. The exact result for the magnetization is calculated using the formula found in [32].

that $\langle \sigma \rangle \cdot \mathbf{m}_i = 0$ ensures that this is a tangent vector. To calculate the fluctuation Hamiltonian we utilize the following result:

$$\sigma \rightarrow e^{-i\sigma \cdot \mathbf{m}} e^{i\sigma \cdot \mathbf{m}} = \sigma - \mathbf{m} \times \sigma + \frac{1}{2} (\mathbf{m}(\sigma \cdot \mathbf{m}) - \sigma(\mathbf{m} \cdot \mathbf{m})) + O(\mathbf{m}^3). \quad (4.8)$$

We can applying this to the transverse-field Ising model evaluated at an arbitrary matrix-product state. The quadratic order fluctuations can be written in terms of the expectation value of one or two point correlators evaluated at the MPS saddle.

Proceeding as in section 3.4, $\mathbf{m}_i$ can be written in terms of azimuthal and longitudinal angles, $\mathbf{m}_i = m_1 \hat{\theta}_i + m_2 \hat{\phi}_i$. $\mathbf{m}_i$ can be transformed into coherent state variables through a change of variables $z_i = m_1 + im_2$. By once again identifying these variables with bosonic operators, the Hamiltonian can be written as:

$$\mathcal{H} = N \hbar \omega_f + \frac{1}{2} \sum_q \begin{pmatrix} \hat{b}_q^\dagger \\ \hat{b}_{-q} \end{pmatrix}^T \begin{pmatrix} M_1 + M_2 & M_1 - M_2 \\ M_1 - M_2 & M_1 + M_2 \end{pmatrix} \begin{pmatrix} \hat{b}_q^\dagger \\ \hat{b}_{-q} \end{pmatrix}. \quad (4.9)$$

where $M_1 = \frac{1}{2} \hbar \langle \sigma_z \rangle + 2 \frac{\langle \sigma_x \rangle^2}{\sigma_z^2} \langle \sigma_x \rangle \langle \sigma_z \rangle + 2 \frac{\langle \sigma_z \rangle^2}{\sigma_z^2} \langle \sigma_x \sigma_z \rangle - 2 \frac{\langle \sigma_y \rangle^2}{\sigma_x^2} \langle \sigma_y \sigma_y \rangle \cos(q)$, $M_2 = \cdots$
\[ \frac{1}{2}(h(x) + 2\langle \sigma_z\sigma_z \rangle - 2\langle \sigma_x\sigma_x \rangle \cos(q)) \] and \[ \sigma^2 = \langle \sigma_x \rangle^2 + \langle \sigma_x \rangle^2. \] Details of this calculation can be found in Appendix A. After diagonalising it with a Bogoliubov transformation, the Hamiltonian takes the form

\[ \mathcal{H} = NH_{mf} + \frac{1}{2} \sum_q \left( \hat{a}_q^\dagger \left[ \begin{array}{c} 2\sqrt{M_1 M_2} \\ 0 \\ 2\sqrt{M_1 M_2} \end{array} \right] \hat{a}_q \right). \] (4.10)

Using equation 3.41, the new estimate for the ground state energy is:

\[ E_{GS} = NH_{mf} + \sqrt{M_1 M_2} - \frac{1}{2}(M_1 + M_2). \] (4.11)

Fluctuation corrections to observables.

For the mean-field ground state calculated in section 4.2.1 we can calculate the correction to observables due to these spin-wave fluctuations. Using the approach to corrections to observables found in 3.8, the corrections to single-site observables is:

\[ \Delta \langle \sigma_\alpha \rangle = -\frac{\langle \sigma_\alpha \rangle}{2} \sum_k \left[ \frac{M_1(k) + M_2(k)}{2\sqrt{M_1(k)M_2(k)}} - 1 \right]. \] (4.12)

Two-site observables, \[ \langle \sigma_x, \sigma_x \rangle, \langle \sigma_y, \sigma_y \rangle \] and \[ \langle \sigma_z, \sigma_z \rangle \] take the following form:

\[ \Delta \langle \sigma_x, \sigma_x \rangle = \sum_k \frac{1}{2} \left[ -(xx)(1 - \frac{z^2}{\sigma^2}) - \frac{zx}{\sigma}(xz) + \frac{x^2}{\sigma^2}(yy) \cos k \right] \left[ \frac{M_1(k)}{\sqrt{M_1(k)M_2(k)}} - 1 \right] \]
\[ + \left[ -(xx) + (zz) \cos k \right] \left[ \frac{M_2(k)}{\sqrt{M_1(k)M_2(k)}} - 1 \right], \] (4.13)

\[ \Delta \langle \sigma_y, \sigma_y \rangle = \frac{1}{2} \left[ -(yy) + \frac{\cos k}{\sigma^2}(z^2 + 2zx(xz) + 2x^2) \right] \left[ \frac{M_1(k)}{\sqrt{M_1(k)M_2(k)}} - 1 \right]. \] (4.14)
(a) Correction to $\langle \sigma_x \rangle$.

(b) Correction to $\langle \sigma_z \rangle$.

(c) Correction to $\langle \sigma_x \sigma_x \rangle$.

Figure 4.5: The correction to the observables for the transverse-field Ising model as a result of spin-wave fluctuations around a uniform MPS ansatz of the form defined in equation 4.3. Results are shown as a function of transverse field for the mean field ansatz, the ansatz with fluctuation corrections and the exact result.

and

$$\Delta \langle \sigma_x \sigma_z \rangle = \frac{1}{2} \left[ -(zz)(1 - \frac{x^2}{\sigma^2}) - \frac{zx}{\sigma}(xz) + \frac{z^2}{\sigma^2}(yy) \cos k \right] \left[ \frac{M_1(k)}{\sqrt{M_1(k)M_2(k)}} - 1 \right] \right]$$

(4.15)

Here $\alpha = \psi(A_0)|\sigma_\alpha^0|\psi(A_0)$, $(\alpha\beta) = \langle \psi(A_0)|\sigma_\alpha \sigma_\beta|\psi(A_0) \rangle$ and $\sigma^2 = (z^2 + x^2)$. In figure 4.5a the expectation of $\langle \sigma_x \rangle$ as a function of transverse field can be seen for the mean-field ansatz and the ansatz modified by spin-wave fluctuations can be seen. The exact result for $\langle \sigma_x \rangle$ is also plotted for a comparison. Unfortunately the magnetisation $\langle \sigma_x \rangle$ is not changed as significantly, as can be seen in figure 4.5b. In figure 4.5c we show the
effect of the spin wave corrections for $\langle \sigma_z \sigma_z \rangle$. In this case we also see a fairly significant improvement over the mean field results.

We finish by noting that domain walls are the low-energy excitations in the ordered phase of the transverse-field Ising model. Domain walls are topologically non-trivial excitations; all spins to the left of the domain wall are polarized in the opposite direction to those to the right. Spin flip excitations - which are relevant in the disordered phase - are purely local. Spin-wave fluctuations as described by equation 4.7 cannot capture topologically nontrivial excitations. Therefore it is unsurprising that the correction to $\langle \sigma_z \sigma_z \rangle$ due to spin-waves is more significant in the paramagnetic phase.
4.3 General fluctuations around an MPS saddle

(a) The eigenvalues of $M_k(H_{BB}(\theta), A_0)$ at $\theta = 0$ and $\chi = 2$. The Bogoliubov transformation cannot be constructed due to negative eigenvalues around $k = \pi$.

(b) The eigenvalues of $M_k(H_{BB}(\theta), A_0)$ at $\theta = \tan(1/3)$ and $\chi = 2$. This is the special AKLT point where the ground state can be represented exactly using a $\chi = 2$ MPS.

Figure 4.6: The eigenvalues of $M_k(H, A_0)$ as a function of momentum for the bilinear, biquadratic, $S = 1$ antiferromagnetic Heisenberg model at $\chi = 2$.

Having considered spin-wave corrections to an MPS ansatz, we now consider general fluctuations around a numerically calculated saddle point. This is done using the techniques introduced in sections 3.5, 3.7 and 3.8. The saddle points used are translationally invariant states found using the time-dependent variational principle. All results in this section are generated for bond dimension $\chi = 2$ MPS.

The spin-wave fluctuations considered in section 4.2.2 were introduced as local unitary rotations. They can therefore be thought of as a $\chi = 1$ product state restriction of the general fluctuations. As noted, the $\chi = 1$ spin-waves could not directly describe topologically nontrivial excitations such as domain walls. This is still true of the generalised fluctuations considered in this section.

4.3.1 The bilinear-biquadratic antiferromagnetic Heisenberg model

Before discussing the transverse-field Ising model, we examine the spin-1 bilinear-biquadratic antiferromagnetic Heisenberg model. This Hamiltonian elucidates a number of interest-
(a) The correction to the energy of the bilinear-biquadratic antiferromagnetic Heisenberg model as function of $\theta$ at $D = 2$. There is no correction to the energy at $\theta = 0.321$ (dashed line) due to the fact the AKLT model can be represented exactly at $\chi = 2$.

(b) The correction to $\langle \sigma_z \sigma_z \rangle$ for the bilinear-biquadratic antiferromagnetic Heisenberg model as function of $\theta$ at $D = 2$. The results are compared to a ground state with $\chi = 100$ found using time-evolving block decimation.

Figure 4.7

In section 3.8 we mentioned that fluctuation corrections to an MPS saddle could not always be consistently calculated. This can happen if the translationally invariant saddle point is a poor approximation to the ground state. For example, a ferromagnetic state is a bad approximation to the ground state of the Heisenberg antiferromagnet. In these cases a Bogoliubov transformation cannot be found due to the negative eigenvalues of $M_k(\mathcal{H}, A_0)$.

In figure 4.6a the eigenvalues of $M_k(\mathcal{H}_{BB}, A_0)$ as a function of momentum can be seen at $\theta = 0$. At $\theta = 0$ the model reduces to the $S = 1$ Heisenberg antiferromagnet. The negative eigenvalues around $k = \pi$ demonstrate that the translationally invariant ground state is a poor choice at the antiferromagnetic point - fluctuation corrections cannot be calculated. Fortunately, as $\theta$ increases the negative eigenvalues disappear. In figure 4.6b the eigenvalues of $M_k(\mathcal{H}_{BB}, A_0)$ can be seen at $\theta = \arctan(\frac{1}{3}) \approx 0.321$, they are positive for all $k$.

We now turn to fluctuation corrections to the $\chi = 2$ saddle in the region around the AKLT point. Since the AKLT model has an exact $\chi = 2$ ground state, there should be
no corrections to the ground state due to fluctuations. In figure 4.7a the correction to the ground state energy $\Delta E$ is plotted a function of $\theta$. As we would expect, we find a finite correction to the energy at every value of $\theta$ except the AKLT point.

Similar behaviour is observed for other observables. In figure 4.7b the fluctuation corrections to $\langle \sigma_z \sigma_z \rangle$ are shown. The corrected result is a significant improvement over the $\chi = 2$ ground state which is constant as $\theta$ is varied.

4.3.2 The transverse-field Ising model

We now move on to a discussion of corrections to the transverse-field Ising model. In figure 4.8a the correction to the magnetisation for a $\chi = 2$ matrix product state ground state as a function of transverse-field can be seen. Unfortunately, the fluctuations do not significantly correct the saddle point. As noted in the MPS ansatz case, this may be due to the fact that domain walls are the pertinent excitations in the ordered phase.

The correction to other observables is more significant. In figure 4.8b the corrections to the $\langle \sigma_y \sigma_y \rangle$ can be seen. The fluctuations move the correlator significantly closer to the exact result.

Figure 4.8
4.4 Discussion

In chapter 3 the methods for calculating fluctuation corrections to an MPS saddle were introduced. As we have seen in chapter 4, these techniques can be applied to improve numerically and analytically calculated MPS ansatz.

In chapter 3 we suggested shifting the saddle point $A_0 \rightarrow \tilde{A}_0$ so as to minimise the new estimate for the ground state energy $E_{GS}$. This has the potential to significantly improve the approximation to the ground state compared to $A_0$. Unfortunately, there are some technical issues that need to be resolved. As we have noted, the Bogoliubov transformation can become undefined when expanding around an inappropriately selected MPS. This can be an issue when shifting $A_0$ to optimise $E_{GS}$, if the imaginary time evolution algorithm overshoots $\tilde{A}_0$ then the algorithm becomes unstable.

There are a number of possible extensions to this work. Any technique that can be applied to a product state path integral can - in principle - be applied to an MPS path integral. Perhaps the simplest idea is to continue to expand around an MPS saddle - to cubic or quartic order. When determining the ground-state of a system, it may prove more convenient to expand around a low bond dimension MPS saddle rather than continue to increase the bond dimension to produce a more accurate saddle point. It would also be interesting to consider the renormalization of an MPS path integral. It is unclear how such a process would relate to other tensor network renormalization schemes [33–35].
Chapter 5

The Lyapunov spectrum of quantum thermalization

The cause of thermalising behaviour in classical and quantum systems differs significantly. In classical systems dynamical chaos leads to ergodic behaviour on a classical phase space. In quantum systems the eigenstate thermalisation hypothesis suggests thermalisation occurs due to dephasing between eigenstates of a system. In this chapter we attempt to bridge the gap between these pictures using matrix product states. Matrix product states of a given bond dimension form a classical phase space. We study the chaotic behaviour of quantum dynamics projected onto this phase space. We also study the chaotic behaviour of dynamics projected onto a manifold of density matrices. In the wavefunction case we observe a relationship between the chaotic behaviour and the growth of entanglement entropy. In the density matrix case we observe significant differences between integrable and nonintegrable systems - suggesting the projected dynamics is encoding important information about the system.

The wavefunction TDVP code was originally written by James Morley. I wrote the code to calculate the Lyapunov spectrum code and also the thermofield double code. This chapter is based upon work available online at [36].
5.1 Introduction

The extra information required to specify a pure quantum state compared to that required for a classical or thermal state underpins many of the apparent paradoxes of quantum mechanics [37]. These may be profoundly philosophical, such as when attempting to apply quantum mechanics to the whole universe, e.g. the black hole information paradox, and the very long scale entanglement implied by the origin of microwave background anisotropy in zero-point fluctuations [38]. Whilst there are fewer philosophical difficulties in the description of finite quantum systems, there are practical consequences. We focus upon one, namely the thermalisation of local observables in the quantum evolution of a closed system.

Accurate numerical description of a quantum system evolving from a weakly entangled initial state requires an exponentially growing number of parameters. The eigenstate thermalisation hypothesis implies that, beyond a certain point in time, an accurate representation of this dynamics should require a reducing number of parameters. The eigenstate thermalisation hypothesis [39–41] has made great strides in demonstrating how thermal correlations present in local observations of eigenstates are revealed through a process of de-phasing due to entanglement with regions of the system not directly under observation. The ultimate consequence is that late-time, local observations are characterised by just the energy density. Evidently, the reduction of parameters required to capture the late-time dynamics [42,43] is related to the emergence of classicality in local observations of the closed quantum system. Here we demonstrate a new way to analyse quantum thermalisation that extends the connection between classical and quantum thermalisation. The central idea is to project the quantum dynamics onto an effective semi-classical, Hamiltonian dynamics on variational manifolds [12,44].

\footnote{The increase and then reduction of parameters required to accurately describe a thermalising quantum system is akin to the Page curve [42,43] for the entanglement entropy of partitions of a system. The Page curve often appears in the context of the black hole information paradox and indeed, its appearance here is for very similar reasons; the difference being that the horizon for observations is imposed by hand in our case and does not evaporate.}
in these classical systems occurs via dynamical chaos \[45\-47\], which we characterise by extracting the full Lyapunov spectrum.

We apply this reasoning to a translationally invariant spin chain, a system over which we have analytical and numerical control using matrix product state methods introduced in chapter 2 and 3.3. Previously we considered MPS representations of the wavefunction, we will now also consider MPS representations of the density matrix. In both cases we follow the dynamics using the time-dependent variational principle. Amongst our key results, in the case of wavefunction MPS we find a zero-parameter fit between the Lyapunov spectrum and the time-dependence of entanglement. In the thermofield MPS near the centre of the spectrum, we recover a semi-circular distribution of Lyapunov exponents for thermalising systems, as found previously in the case of gravitational systems \[48\-49\], and a Gaussian distribution for integrable systems.

By bringing the study of many-body quantum chaos into contact with that of classical chaos, our approach opens up the full range of techniques available in the latter. For example, it allows the potential to examine how the classical KAM theorem for deformations from integrable behaviour and periodic orbits in classically chaotic systems may manifest in quantum systems \[50\-54\]. It also suggests natural possibilities for efficient descriptions of late-time dynamics. This complementary perspective brings the study of quantum chaos full circle, recapitulating the characterisation of few particle quantum chaos through its projection to classically chaotic systems.

In Section 5.2 we begin by reviewing the role of classical dynamical chaos in enabling the ergodicity and thermalisation of classical closed system. This introduces some of the ideas, methods, and nomenclature of classical chaos that we will later apply to projected quantum dynamics. We then turn to quantum dynamics and give brief expositions of eigenstate thermalisation, the important role played by quantum chaos and the conventional characterisation of the latter through the eigen spectrum. In Section 5.3 we discuss the projection of pure quantum dynamics to a variational manifold, the condi-
tions under which this captures the dynamics of a restricted set of observations, and how this projected dynamics reduces to an effective (semi-)classical dynamics. Section 5.4 gives some of the technical detail (expanded upon in appendices) of how to extract the Lyapunov spectrum of projected quantum dynamics. Section 5.5 summarises our numerical results and the relationship of the Lyapunov spectrum to other measures of quantum chaos. Finally, we discuss the broader implications of our results.

5.2 Classical and Quantum Thermalization

5.2.1 Classical Thermalization

*Ergodicity and Chaos:* Thermalisation in closed classical systems occurs due to dynamical chaos. Every dynamical mode of the system is chaotic, revealed on timescales given by the inverse of its corresponding Lyapunov exponent. On the longest timescales, evolution leads to ergodic exploration of states in phase space with a given energy (or other conserved quantities). This picture of classical thermalisation requires an explicit ensemble or time-averaging to obtain thermal averages, a point that we will return to later. On these timescales, only conserved quantities can be used to locally distinguish states of the system - small differences in states with the same values of conserved quantities are eventually randomised by the chaotic evolution and so cannot be used to characterise the state.

On shorter timescales, dynamical modes can be divided into two classes; those that have revealed their chaotic behaviour, and those that have not. We will refer to these as chaotic and residual regular modes (a classification that is determined by a choice of timescale). Residual regular modes can be used to discriminate between states of the system on a given timescale. The chaotic modes effectively form part of a thermal bath, and time or ensemble averaging will draw uniformly from the possible amplitudes of deformations in chaotic directions. On increasing timescales, the number of residual
regular modes decreases monotonically, until ultimately only conserved quantities remain as distinguishing features of classical states. It is important to note that this is not simply a matter of averaging out high-frequency modes. A high frequency mode with high quality factor can distinguish states on timescales longer than its period. Of course, it is plausible that frequencies and decay rates may be linked in some cases, but this is not necessarily so.

The Lyapunov Spectrum: After setting the scene in this way, we now give an overview of how the Lyapunov spectrum is extracted for a classical dynamical system. There are many excellent reviews of this subject \[45\–47\]. We confine ourselves to a brief outline in order to contextualise our analysis of the quantum system. Consider a dynamical system whose parameters are contained in a vector $X(t)$ that evolves according to

$$\frac{d}{dt} X(t) = F(X(t)).$$  \hspace{1cm} (5.1)

The Lyapunov spectrum is found by considering the evolution of the displacement between neighbouring trajectories $X(t)$ and $X(t) + dX(t)$, where $dX$ is initially small. Expanding Eq. (5.1) to leading order, we obtain the following equation for the evolution of the displacement between trajectories:

$$\frac{d}{dt} dX_i = \partial_j F_i(X) dX_j.$$  \hspace{1cm} (5.2)

The formal solution of this equation, $dX(t) = T(\exp[\int_0^t \partial F(t') dt'])dX(0)$, shows that instantaneously, $dX$ grows exponentially and decreases exponentially with $t$ in the eigendirections of $\partial_i F_j$. The exponents for this growth and decay are the instantaneous Lyapunov exponents and the Lyapunov exponents are the time average of them along a trajectory. These equations are manipulated in various ways to determine the exponents numerically \[45\] (See Appendix B). Conservation of phase space volume under Hamiltonian dynamics implies that the exponents (both instantaneous and averaged)
sum to zero. Moreover, time-reversal invariance demands that they come in positive and
negative pairs. As we shall see below, projection from unitary quantum dynamics to
classical dynamics on a variational manifold leads to additional constraints.

5.2.2 Quantum Thermalisation

There are various ways to express the eigenstate thermalisation hypothesis. Perhaps the
simplest is to state that the expectation of observables should typically have a smooth
dependence upon the energy of the state. If this is true for arbitrary states, then it
ought to be true for an eigenstate. The expectations of local operators are the same as
in a Gibbs state with the same energy density.

This locality is crucial. In the conventional view of quantum thermalisation, it allows
the major part of the system, over which the observable has no support, to act as a bath
for the parts of the system engaged directly in the observation. The remarkable conclu-
sion of the eigenstate thermalisation hypothesis [39–41] is that the information about
thermal averages of local operators is contained in the quantum eigenstates themselves.
Coherences in an initial superposition of eigenstates can obscure this fact. Time evolution
reveals the underlying thermal properties by a process of dephasing. For thermalisation
to occur, the part of the system within the observation window must be highly entangled
with the system beyond. This is reflected in the fact that states towards the top and
bottom of the spectrum — that in one-dimension are provably weakly entangled [55] and
suspected to be so in higher dimensions — obey the eigenstate thermalisation hypoth-
esis less well than those in the centre [41]. Even in the centre of the spectrum, special
eigenstates may exhibit non-thermalizing behaviour. [51–54]. In this picture, quantum
thermalisation depends upon rates of de-phasing, which in turn depend upon differences
in the frequencies of the quantum eigenstates of the full system. This is apparently very
different from the dependence of classical thermalisation upon the Lyapunov spectrum,
although there clearly some relation, since systems that display quantum chaos more
rapidly dephase spatial partitions. Our aim in the following is to further explore the links between classical and quantum thermalisation.

5.3 Projecting Quantum to Classical Dynamics

Noting the importance of the locality of observation permits an alternative way to frame eigenstate thermalisation that makes much closer contact with its classical counterpart. Central to this is recognising that observations on a spatial partition of a system can be captured accurately by projecting states to some manifold of variational approximates. Following the projected dynamics on this manifold using the time-dependent variational principle results in a classical Hamiltonian dynamics whose thermalisation is driven by its chaotic properties and characterised by the Lyapunov spectrum. We present two ways to achieve this mapping: approximating the wavefunction of the system using matrix product states (MPS), and approximating the thermofield double purification of the density matrix by matrix product states. The numerical implementation of these two protocols is very similar — indeed, we use the same code (with suitable modification) for both cases — but both their regime of applicability and the manner in which they encode the physics is rather different.

5.3.1 TDVP applied to the wavefunction

A variational parametrisation of a system’s wavefunction picks out a sub-manifold of Hilbert space. There are a number of ways that one might choose to project dynamics onto this manifold. The time-dependent variational principle does so by mapping an updated quantum state — which in general will lie outside of the manifold — onto the state on the manifold with which it has the highest fidelity. A remarkable property of this mapping is that the reduced equations are those of a classical Hamiltonian system. In particular, a quantity conserved by the exact dynamics will also be conserved by the
projected dynamics, provided that the symmetry transformation generating it can be captured on the manifold. This permits sensible results to be obtained even at very long times \[44\].

Consider a variational parametrization with a set of complex parameters \(\{X_i\}\). The time derivative of the wavefunction may be written
\[
\frac{\partial}{\partial t} |\psi\rangle \approx |\partial X_i \psi\rangle \dot{X}_i.
\]
It is tempting to substitute this into the Schrödinger equation, but the result is not correct since the action of the Hamiltonian on the state \(|\psi(X)\rangle\) will generally take the state out of the variational manifold. Contracting with a tangent vector \(\langle \partial \bar{X}_i \psi | \partial X_j \psi \rangle\) fixes this and permits us to write
\[
\langle \partial \bar{X}_i \psi | \partial X_j \psi \rangle \dot{X}_j = i \langle \partial \bar{X}_i \psi | \hat{H} | \psi \rangle.
\]
(5.3)

Using a particular basis for the tangent space, one may fix the Gramm matrix \(\langle \partial \bar{X}_i \psi | \partial X_j \psi \rangle = \delta_{ij}\) after which identifying positions and momenta \(q_i \equiv \sqrt{2} \text{Im}[M_{ij} X_j]\) and \(p_i \equiv \sqrt{2} \text{Re}[M_{ij} X_j]\) reduces Eq. (5.3) to Hamilton’s equation for a classical system\[^2\]. Even though the parameters \(\{X_i\}\) may quantify aspects of the entanglement structure of the wavefunction, they nevertheless provide a (semi-) classical description\[^3\]. The technical details of applying this to matrix product states was developed in a seminal work of Jutho Haegeman \textit{et al}. We summarise pertinent details in Appendix A. Once the quantum dynamics has been mapped to classical dynamics in this way, we may proceed to evaluation the Lyapunov spectrum associated with this dynamics.

5.3.2 TDVP applied to the thermofield double

As we discuss presently, the MPS ansatz applied in the usual way efficiently describes states near to the top and bottom of the spectrum. States near to the centre of the spectrum we require an alternative variational parametrization. We use an MPS parametrization
\[^2\]One may alternatively choose positions and momenta \(q_i \equiv \sqrt{2} \text{Im}[M_{ij} X_j]\) and \(p_i \equiv \sqrt{2} \text{Re}[M_{ij} X_j]\) with \(M_{ij}^{-2} = \langle \partial \bar{X}_i \psi | \partial X_j \psi \rangle\) in order to demonstrate the mapping to a classical system without this explicit gauge fixing.
\[^3\]This extends the notion of classical chaos considered in Ref. [56] to semi-classical properties present even without the strict limit of \(\hbar \to 0\).
tion of the thermofield double. The thermofield double \[57\] is a purification of the density matrix. In the eigenbasis of the density matrix \( \hat{\rho} = \sum_\alpha \gamma_\alpha |\alpha\rangle \langle \alpha | \), it may be written as

\[
|\psi\rangle = \sum_\alpha \sqrt{\gamma_\alpha} |\alpha\rangle \otimes |\alpha\rangle,
\]

where \( \gamma_\alpha \) are real positive weights that correspond to the Gibbs weights in thermal equilibrium, and \( \alpha \) labels the eigenstates, \( |\alpha\rangle \). Physical operators act on the first copy of the state only, so that expectations with the thermofield double are identical to those obtained from the density matrix:

\[
\langle \psi | \hat{\theta} | \psi \rangle = \text{Tr}(\hat{\rho} \hat{\theta}).
\]

The time-evolution of the thermofield double is determined by the Hamiltonian \( \mathcal{H} = \mathcal{H} \otimes 1 + 1 \otimes \mathcal{H} \), which acts symmetrically on the doubled space.

Having identified the thermofield double and the appropriate Hamiltonian, we are free to construct an MPS ansatz for it and to evolve using the time-dependent variational principle. The time-dependent variational principle projects to the variational manifold by optimising the fidelity of the thermofield double. This amounts to optimising over a certain set of observations — specifically the trace-norm of the square root of the updated density matrix with the square root of its variational approximation \[4\]. The bond order of the MPS for the thermofield double does not have a direct interpretation in terms of the entanglement structure of individual states. Moreover, although evolution under \( \mathcal{H} = \mathcal{H} \otimes 1 + 1 \otimes \mathcal{H} \) without approximation would preserve the purity of the state, projection to the variational manifold takes pure states into mixed states. This is consistent with optimising over a certain set of observations, but quite different from the wavefunction MPS which remains pure. Although TDVP has been applied to the density matrix before \[58\], as far as we are aware, this is the first time that it has been used to follow real time evolution of a matrix product ansatz for it (though see \[59\] for a related work). In order to obtain accurate results, we have made an important modification to the algorithm developed in \[12\] for MPS representations of the state. The MPS for the thermofield double can be written such that the symmetry between the two copies of the physical space is explicit. This is achieved for a bond order \( D = D^2 \) thermofield

\[\text{The square root guarantees that the fidelity is 1 for identical density matrices}\]
MPS by imposing the symmetry $A_{\sigma \delta I,J} = A_{\delta \sigma \bar{I},\bar{J}}$, where $I \equiv i \otimes i'$ and $\bar{I} \equiv i' \otimes i$ with the indices $i, i', j, j' \in \{1, 2, \ldots D\}$\footnote{As described in Appendix A, we employ an additional gauge fixing on the tangent manifold that imposes a constraint gauge equivalent to this. This reduces the number of variables and increases the accuracy considerably.} As described in Appendix A, we employ an additional gauge fixing on the tangent manifold that imposes a constraint gauge equivalent to this. The Infinite temperature state takes a particularly simple and instructive form when represented in terms of a thermofield MPS. At $\mathcal{D} = 1$ it is given by $A^{\sigma \delta} = \delta^{\sigma \delta}/\sqrt{2}$. At $\mathcal{D} = D^2 > 1$ there are many ways to represent the state. A class of symmetrical thermofield MPS states can be constructed from a unitary matrix $U \in SU(dD)$ as

$$A_{\sigma \delta I,J} = \frac{1}{\sqrt{d}} \sum_{\gamma=1}^{d} U_{(\sigma i),(\gamma j)} U_{(\delta i'),(\gamma j')}.$$  

This follows from noting that i. the infinite temperature state is the same for any Hamiltonian and ii. that it is invariant under evolution with the Hamiltonian. Eq. (5.4) follows by representing an arbitrary time evolution of $A^{\sigma \delta} = \delta^{\sigma \delta}/\sqrt{2}$ with a bond operator representation of the time-evolution operator using the unitary $U$. This manifold of equivalent representations of the infinite temperature state resolves an apparent contradiction: on the one hand a state at the middle of the spectrum of a given Hamiltonian is expected to evolve towards the infinite temperature state, whilst on the other hand the projected dynamics is classically Hamiltonian and so cannot evolve to a single point in phase space. It also holds the seed of how to compress the thermofield MPS representation of a thermalising system at late time\footnote{\emph{A. Hallam and A. G Green work in progress.}}.

5.3.3 Comparing Classical Projections

These two schemes for projecting quantum dynamics to classical Hamiltonian dynamics capture the physics in rather different ways and have different regimes of validity. The

\footnote{Note that a pure state with bond order $D$ wavefunction MPS tensor $A^\sigma_{i,j}$ can be represented as a thermofield MPS of bond order $D = \mathcal{D}$ and tensor $A^\sigma_{i,j} = A^\sigma_{i,j} A^\sigma_{i',j'}$.}
MPS approximation for a state is efficient near the top and the bottom of the spectrum. The bond order required to accurately describe a thermal state at temperature $T$ scales as a double exponential [60]. The thermofield MPS is efficient both at the edges and near to the centre of the spectrum. The latter is justified heuristically as follows: a thermofield MPS of bond order $D$ accurately describes observations up to a lengthscale $\sim \log_d D$. If this is longer than the thermal correlation length in the final state, the description will accurately capture the dynamics. This occurs near the centre of the spectrum, where the effective temperature is high and the correlation length is short.

These differences are also revealed in correlation lengths and the factorisation of averages such as $\langle \sigma_n^x \sigma_{n+N}^x \rangle$ for $N$ greater than the thermal correlation length. The wavefunction MPS at low bond order captures such properties in explicit time-averages. The instantaneous correlation length of the wavefunction MPS extracted from its transfer matrix [61] can be longer than the thermal correlation length, reflecting the long-distance entanglement of its constituent eigenstates. The thermofield MPS captures the thermal correlation length in a rather different way. Since it is a purification of the density matrix, the thermofield MPS is directly related to observations and already includes the effects of dephasing. In this case, the instantaneous correlation length deduced from the transfer matrix is equal to the thermal correlation length and long distance correlators factorise in instantaneous observations.

5.4 Lyapunov Spectrum of Projected Dynamics

In this section, we summarize how to extract Lyapunov spectra from projected quantum dynamics. The details of this are similar for our two projection schemes. For clarity, we will focus our discussion upon the wavefunction MPS, noting modifications necessary for the thermofield MPS as appropriate.
Figure 5.1: Convergence Plot for a Typical Thermalising System: The time-averaged Lyapunov exponents versus time are shown for an MPS representation of the wavefunction of a typical thermalising system. We consider an Ising model with anti-ferromagnetic coupling $J = 1$, transverse field $h_x = 0.5$ and longitudinal field $h_z = 1$. The dynamics are obtained by integrating Eq. (C.1) and the spectrum from averaging instantaneous exponents obtained from Eq. (5.6) both using bond order $D = 10$ (as a representative example).
5.4.1 Distance on the Variational Manifold

As a first step to deducing the Lyapunov spectrum, we must assign a distance measure on the variational manifold. This is done using the fidelity between states with two different coordinates $X$ and $X + dX$. As a simple example, consider a spin-1/2 coherent state given by

$$|\theta, \phi \rangle = e^{-i\phi/2} \cos \left( \frac{\theta}{2} \right) \uparrow + e^{i\phi/2} \sin \left( \frac{\theta}{2} \right) \downarrow.$$ 

The square of the distance between two such states $|\theta, \phi \rangle$ and $|\theta + d\theta, \phi + d\phi \rangle$ can be written, after expanding the fidelity between them to quadratic order, by

$$dS^2 = 1 - |\langle \theta, \phi | \theta + d\theta, \phi + d\phi \rangle|^2 = \frac{1}{4} (\sin^2 \theta d^2 \phi^2 + \phi d^2 \theta)$$

corresponding to the usual distance measure on the Bloch sphere. In the case of translationally invariant states, we must use the fidelity density rather than fidelity, since the fidelity between translationally invariant states described by an MPS tensor $A_{ij}^\sigma$ and one described by a tensor $A_{ij}^\sigma + dA_{ij}^\sigma$ scales as one over the total length of the system. We will parametrise a small deviation from $A_{ij}^\sigma$ as $A_{ij}^\sigma \rightarrow A_{ij}^\sigma + dA_{ij}^\sigma (X_{ij}^\sigma)$ where $dA(X_{ij}^\sigma)$ is parametrised as in equation 2.15 in chapter 2. The distance measure takes a particularly simple form in terms of $X$:

$$dS^2 = \sum_{\sigma ij} X_{ij}^\sigma X_{ji}^\sigma.$$ \hspace{1cm} (5.5)

This parametrisation is useful in determining the Lyapunov spectrum, the details of which we turn to next.

5.4.2 Linearised TDVP and the Lyapunov Spectrum

To extract the Lyapunov spectrum we must characterise the divergence between nearby trajectories. Consider two trajectories both in the vicinity of a point on the MPS manifold with tensor $A_{ij}^\sigma$. Let these trajectories have parametrisations in terms of $X_{ij}^\sigma(t)$
and \(X^\sigma_{ij}(t) + dX^\sigma_{ij}(t)\), respectively. Substituting each of these into the time-dependent variational principle Eq.(5.3) and subtracting, we obtain the following equation for the evolution of the difference between trajectories

\[
d\dot{X}^\sigma_{ij}(t) = i(\partial_{X^\sigma_{ij}} \partial_{X^\gamma_{kl}} \psi |\hat{H}|\psi) dX^\sigma_{ij}(t) + i(\partial_{X^\sigma_{ij}} \psi |\hat{H}|\partial_{X^\gamma_{kl}} \psi) dX^\gamma_{kl}(t).
\] (5.6)

With the minor modification of allowing complex parameters, this equation is clearly analogous to Eq.(5.2) used to extract the Lyapunov spectrum for classical trajectories. We can therefore define the Lyapunov exponent \(\lambda\) associated with this trajectory in the following manner:

\[
\lambda = \lim_{t \to \infty} \frac{1}{t} \log \frac{dX(t)}{dX(0)}.
\] (5.7)

The full Lyapunov spectrum can be found by creating a \((d-1)D^2\) dimensional parallelepiped which spans the tangent space of \(X^\sigma_{ij}(t)\), \(U(t) = \{dX^1(t), dX^2(t), \ldots\}\) and performing an equivalent time averaging,

\[
\sum_i \lambda_i = \lim_{t \to \infty} \frac{1}{t} \log \frac{\text{Vol}(U(t))}{\text{Vol}(U(0))}.
\] (5.8)

Similar structures have been used by Haegeman et al in order to construct the excitation ansatz [62], and form the zero-wavevector part of the kernel of a quadratic expansion of MPS path integral about its saddle-point [29]. Appendix B gives details of how this equation is evaluated. Extraction of the Lyapunov spectrum now proceeds as in the classical case, using Eq.(5.6) to find the instantaneous Lyapunov spectrum at each point along a trajectory given by Eq.(C.1) and averaging.

A final addition to this procedure — not usually used in calculating Lyapunov exponents for classical dynamical systems — is to parallel transport displacements between nearby trajectories along the variational manifold (see Appendix C). This enables us to
Figure 5.2: Lyapunov Spectrum for a wavefunction MPS representation of Ising model dynamics: a) Non-integrable case with $J = 1$, $h^x = 0.5$, $h^z = 1$. b) Integrable case with $J = 1$, $h^x = 0.5$, $h^z = 0$. c) Nearly Integrable case with $J = 1$, $h^x = 0.5$, $h^z = 0.1$ In all cases the spectrum is obtained for an MPS representation of the wavefunction at bond order $D = 20$.

satisfy some constraints of projected quantum dynamics to numerical precision. The Lyapunov spectra of classical Hamiltonian systems are constrained by time-reversal invariance to have all of the exponents in positive/negative pairs with the same modulus. This property is inherited by the spectrum of projected quantum dynamics. An additional important property follows from using fidelity to determine the measure on the variational manifold. Fidelity is not changed by unitary time evolution. As a result, Lyapunov exponents calculated for unitary evolution must be identically zero. Evolution under a purely local Hamiltonian provides a useful test case, since it does not change the entanglement structure of a quantum state and the time-dependent variational principle Eq.(C.1) reproduces the full Schrödinger equation under projection onto any manifold. The Lyapunov exponents in this case must be identically zero. Fig. 5.1 shows typical convergence plots for an MPS approximation to the wavefunction of a thermalising system. The corresponding Lyapunov spectrum is show in Fig. 5.2.
5.5 Numerical Results

In this section, we summarise the results of applying the above methods to the thermalisation of the Ising model with longitudinal and transverse fields:

\[ H = \sum_i \left[ J \sigma_i^z \sigma_{i+1}^z + h^z \sigma_i^z + h^x \sigma_i^x \right] . \] (5.9)

The properties of this model are well known; it is integrable when the longitudinal field \( h^z \) is zero and non-integrable otherwise. This allows us to investigate: i. integrable systems \( (J = O(1), h^x = O(1) \) and \( h^z = 0) \), ii. non-integrable/thermalising systems \( J = O(1), h^x = O(1) \) and \( h^z = O(1) \)), and iii. nearly integrable systems \( J = O(1), h^x = O(1) \) and \( h^z \ll h^x \)). We apply the machinery of the time-dependent variational principle to determine trajectories, and the linearised time-dependent variational principle to determine Lyapunov spectra. Reflecting their different encodings of the relevant physics and different regimes of validity, we separate our discussions of the wavefunction MPS and thermofield MPS.

5.5.1 Wavefunction MPS

We now consider the Lyapunov spectra evaluated from the wavefunction MPS starting from an initial product state \( |\psi(0)\rangle_i = (0.382 - 0.382i) |\uparrow\rangle_i + (-0.595 + 0.595i) |\downarrow\rangle_i \) near the bottom of the spectrum. Due to the ergodicity of the dynamics, all initial states of the same energy density result in the same Lyapunov spectrum. The Lyapunov spectrum for the non-integrable, integrable and nearly-integrable cases are shown in Fig. 5.2. All show a broad distribution of exponents, with no strong differences apparent between integrable and non-integrable cases.

\[
\text{spectrum vs bond order:} \quad \text{since the nonlinearities and chaos of our dynamics arise from projection to the variational manifold, the Lyapunov spectrum varies with bond order.}
\]
Figure 5.3: *Maximum Lyapunov exponent versus bond order.* The maximum Lyapunov exponent depends strongly upon the projection non-linearities at different bond orders, tending to zero in the limit $D \to \infty$. Here we show the largest exponent varying with bond order for Non-Integrable (circles), Integrable (crosses) and Nearly integrable (pluses) systems. The largest exponent decreases like $\lambda_{\text{max}}(D) = 0.320(D - 1)^{-0.219}$ for Non-Integrable systems, $\lambda_{\text{max}}(D) = 0.553(D - 1)^{-0.280}$ for Integrable systems and $\lambda_{\text{max}}(D) = 0.480(D - 1)^{-0.245}$ for Nearly-Integrable systems.

This situation is unlike the conventional use of matrix product methods, where increasing bond order give increasingly accurate results. The dependence of the maximum Lyapunov exponent, $\lambda_{\text{max}}$, with $D$ is shown in Fig. 5.3. This appears to show a monotonic decrease from $D = 2$ as $D \to \infty$. Note that in the translationally invariant case with spin $1/2$, the projected dynamics is not chaotic at $D = 1$ by the Poincaré-Bendixson theorem, since the phase space is two-dimensional. The following discussion demonstrates the consistency of these results with physical observations.

Maldacena et al. [63] have conjectured that the largest Lyapunov exponent of a quantum system has an upper bound related to its temperature $\lambda_{\text{max}} \leq 2\pi k_B T / \hbar$. The
Figure 5.4: *Kolmogorov-Sinai entropy versus bond order:* The Kolmogorov-Sinai entropy is related to entanglement growth at short times, it appears to be diverging with bond dimension. Here we show the KS entropy varying with bond order for Non-Integrable (circles), Integrable (crosses) and Nearly integrable (pluses) systems. The Non-Integrable KS entropy grows like $S_{\text{KS}}(D) = 0.518(D-1)^{1.296}$, the Integrable KS entropy grows like $S_{\text{KS}}(D) = 0.733(D-1)^{1.337}$ and the Nearly-Integrable KS entropy grows like $S_{\text{KS}}(D) = 0.638(D-1)^{1.386}$.

At low energies the exponent appears to increase as a power law before saturating at $E \approx 0.6$.

The dependence of the *Entanglement Entropy, $S_E$* upon time is shown in Figs. 5.6 and 5.7. For a given bond order, $S_E$ saturates. To a good approximation (that we can manipulate analytically) this saturation value corresponds to drawing the Schmidt coefficients $s_n$ from a distribution given by the modulus of the elements of a random $O(D)$ vector. The mean Schmidt coefficients then correspond to $s_n = n\sqrt{6}/\sqrt{D(1+D)(1+2D)}$, from
Figure 5.5: Maximum Lyapunov exponent versus energy density: It has previously been conjectured that $\lambda_{\text{max}} \leq 2\pi k_B T/\hbar$, here observe that $\lambda_{\text{max}}(D = 2)$ increases with energy density above the ground state but appears to saturated at $E \approx 0.6$.

which one may deduce a saturation entanglement at large bond order given by

$$S_E^{\text{Sat}}(D) = -\sum_{n=1}^{D} s_n^2 \log s_n^2 \approx \log \left( \frac{D e^{2/3}}{3} \right).$$  \hspace{1cm} (5.10)

With growing entanglement, the effective bond order of the quantum state (the bond order required for an accurate description) grows. We can use Eq. (5.10) to deduce the time-dependence of the bond order; $D$ attains a particular value at the point where $S_E(t)$ crosses the corresponding saturation value. A continuous approximation can be found by equating $S_E^{\text{Sat}}(D) = S_E(t)$, from which we obtain

$$D(t) = 3e^{-2/3}(e^{S_E(t)} - 1) + 2$$ \hspace{1cm} (5.11)
Figure 5.6: **Entanglement entropy across a bond compared to randomly distributed Schmidt coefficients:** At a given bond dimension the entanglement entropy will saturate after a short time. The saturation value for the entanglement entropy is in strong agreement with a random uniform distribution of Schmidt coefficients as discussed in the text.

which fits both the large bond-order limit implied by Eq. (5.10) and the initial bond order. This dependence of bond order upon time allow us to demonstrate the consistency of the Lyapunov spectrum and its variation with $D$ with the physically relevant dependence of the entanglement entropy upon time.

The *Kolmogorov Sinai* entropy is a measure of how quickly knowledge of a system’s initial state is lost in a chaotic system. It determines the rate of growth of the volume (of gyration) of a region of phase space and, following Pesin’s theorem [64], is given by the sum of the positive Lyapunov exponents. Fig. 5.4 shows the Kolmogorov Sinai entropy calculated from our Lyapunov spectra and its dependence upon bond order. The
Figure 5.7: Entanglement entropy and Kolmogorov-Sinai entropy: The gradient of the entanglement entropy is determined by the Kolmogorov-Sinai entropy. Here we demonstrate the short time behaviour of the entanglement entropy is determined by $S_{KS}(D = 2)$. By integrating Eq.(5.13) after substituting $D(t)$ from Eq.(5.11) and the fitted form of $S_{KS}(D)$ from Eq.(5.12) we find a zero-parameter fit between the Lyapunov spectrum and entanglement.
Figure 5.8: Lyapunov Spectrum for a thermofield MPS representation of Ising model dynamics: a) Non-integrable case with $J = 1$, $h^x = 0.5$, $h^z = 1$. b) Integrable case with $J = 1$, $h^x = 0.5$, $h^z = 0$. c) Nearly Integrable case with $J = 1$, $h^x = 0.5$, $h^z = 0.1$. In all cases the spectrum is obtained for a wavefunction MPS at bond order $D = 16$. The non-integrable case appears to fit a semicircle distribution with radius $r = 0.39$, the integrable case appears to be Gaussian with standard deviation $\sigma = 0.167$ and the nearly integrable case appears to be Gaussian with standard deviation $\sigma = 0.161$.

latter dependence is fitted in the non-integrable case with a polynomial approximation:

$$S_{KS}(D) = 0.518(D - 1)^{1.296}.$$ \hfill (5.12)

Studies of single particle quantum chaos have shown the relationship $\dot{S}_E(t = 0) = S_{KS}$, provided that starting wavefunction is as classical as possible \[65–67\]. Here we find — as indicated in Fig. 5.7 — that $\dot{S}_E(t = 0) = S_{KS}(D = 2)$. $D = 2$ corresponds to the most classical, non-trivial (recall that $D = 1$ has vanishing Lyapunov exponents) projected dynamics and is the many-body equivalent of the single particle result. We speculate the following extension of this result:

$$\dot{S}_E(t) = \frac{S_{KS}(D(t))}{(D(t) - 1)^2}. \hfill (5.13)$$

Our main justification for this is the very good, zero-parameter fit that it gives between our results for the entanglement and Lyapunov spectrum. Fig. 5.7 shows the time-integral of the right-hand side of Eq.(5.13) substituting $S_{KS}(D)$ from Eq.(5.12) and

\footnote{Note that the expansion is in $D - 1$ since the Lyapunov exponents and so $S_{KS}$ are zero at $D = 1$.}
\(D(t)\) from Eq.\((5.11)\), alongside the entanglement entropy.

Eq.\((5.13)\) can also be used to place bounds upon the Lyapunov spectrum. At long times we expect \(S_E(t) \sim t\) for thermalising systems. Assuming the validity of Eq.\((5.13)\) then, \(S_{KS}(D) = \alpha D^2\) at large \(t\) suggesting that the fit in Eq.\((5.12)\) must be modified at large \(D\). Moreover, if the exponents converge to a consistent distribution then this also implies that \(\lambda_{max}\) converges. For example, if the exponents approach a semicircular distribution then \(\pi \lambda_{max}^2(D)/4 = \alpha\) to be consistent with \(S_{KS}(D) = \alpha D^2\).

It is apparent from these observations that the Lyapunov spectrum extracted from mapping the quantum dynamics of the wavefunction to classical Hamiltonian dynamics is not unique. Although the dependence is rather slow with bond order, there is no sense in which spectra collected in this way show numerical convergence, for example with increasing bond order. A moments reflection about the way in which the wavefunction MPS captures the physics of thermalisation shows why. At low bond order, the dynamics is very non-linear and thermalisation occurs via chaotic classical dynamics. Thermal averages are recovered in temporal averages of the simulated dynamics. As bond order increases, the MPS ansatz make better and better approximation to the underlying eigenstates and ultimately, thermalisation is captured in the same way as the conventional picture of eigenstate thermalisation. Thermal averages are obtained in instantaneous measurements after an initial period of dephasing reveals the intrinsic properties of the underlying eigenstates. However, the Lyapunov spectrum does have physical meaning. We have demonstrated how the physical quantity, \(S_E(t)\), is related to the Lyapunov spectrum obtained on a variational manifolds.

### 5.5.2 Thermofield MPS

The above analysis allows us to relate the chaos of projected quantum dynamics near to the edge of the spectrum to the process of thermalisation. As discussed in Sec. \(5.3\) a matrix product state description of the wavefunction cannot work near to the centre
Figure 5.9: Maximum Lyapunov Exponent vs Thermofield MPS Bond Order for Non-Integrable System: The largest Lyapunov exponent for the Ising model with $J = 1$, $h_x = 0.5$, $h_z = 1.0$ obtained for an MPS representation of the Thermofield double. The exponent appears to be approaching zero like $\lambda_{max} = 1.09 D^{-0.373}$.

of the spectrum. In this subsection, we apply our analysis of the Lyapunov spectrum to an MPS of the thermofield double. We consider an initial pure state near to the middle of spectrum, $|\psi(0)\rangle_i = 0.448 \uparrow_i + 0.873 \downarrow_i$. The late time dynamics of this are similar to the infinite-temperature state.

The Lyapunov spectra for the thermofield MPS dynamics are shown in Fig. 5.8. There is a clear distinction between the non-integrable, and integrable and nearly-integrable cases. The former has a semi-circular distribution, whereas the latter are narrower and fit a Gaussian distribution (with long tails that have been cut off in Fig. 5.8). The semi-circular distribution in the non-integrable case suggests a connection to random matrix theory. Such a connection has previously been explored in the context of quantum
Figure 5.10: Kolmogorov-Sinai entropy vs Thermofield MPS Bond Order for Non-Integrable System: The Kolmogorov-Sinai entropy for the Ising model with $J = 1$, $h_x = 0.5$, $h_z = 1.0$ obtained for an MPS representation of the Thermofield double. The KS entropy appears to be diverging, growing like $S_{KS} = 1.427D^{1.58}$.

Fig. 5.9 shows the variation of the maximum Lyapunov exponent with bond order. The symmetry constraint that we impose upon the thermofield MPS tensor restrict the bond order to $D = 1, 4, 9, 16$ etc. and together with the rapid growth of the number of Lyapunov exponents as $2(d^2 - 1)D^2$ this leads to rather few points in the figure. Our numerics are fit by $1.09D^{-0.373}$, or $1.17e^{-0.0173D}$, but are also consistent with convergence $0.410 + 0.1740e^{-0.0116D}$. The latter might be expected since the thermofield double (being a purification of the density matrix) encodes a limited set of observations corresponding roughly to a window of size $\frac{1}{2}\log_2 D$. When this window is larger than the correlation length timescales of the dynamics are expected to converge to values

---

#NB: since the dimension of the local Hilbert space is $d^2$, dynamical chaos occurs at $D = 1$. 

79
Figure 5.11: Entanglement of the midspectrum state: The entanglement between sites of a Thermofield MPS state starting in the middle of the spectrum saturates at a value close to value obtained by averaging over random unitaries.
characteristic of the observable thermalisation process. The Kolmogorov Sinai entropy for the thermofield MPS is shown in Fig. 5.10. This is fit with $1.427 \mathcal{D}^{1.58}$ to high accuracy. This scaling is less than $\mathcal{D}^2$ (the volume of phase space) of a typical classical dynamical system. This is consistent with unitary dynamics as $\mathcal{D}$ tends to infinity. The thermofield entanglement is shown in Fig. 5.11. At short times the thermofield time evolution is exact and the entanglement is double the matrix-product entanglement. When the thermofield state becomes mixed the thermofield entanglement appears to be more closely related to operator entanglement $[68, 69]$. It is interesting to note that the saturation of this thermofield entanglement is near to the mean value obtained by averaging the thermofield entanglement of the infinite-temperature state given by Eq. (5.4) with a Haar measure over $U$. Finally, we note that unlike wavefunction MPS, we have not been able to determine an simple relationship between the Kolmogorov-Sinai entropy and the thermofield entanglement.

5.6 Discussion

The analysis presented above allows the thermalisation of local observables to be recast as a chaotic classical Hamiltonian dynamics in two different ways: using the time-dependent variational principle to evolve MPS representations of the wavefunction and of the thermofield double. This picture is complementary to the dephasing of eigenstates in the conventional picture of eigenstate thermalisation and brings the study of quantum chaos full circle. Early studies of quantum chaos focussed upon single particle quantum systems whose semi-classical limit is chaotic. The impact of this upon the level statistics provides a convenient way to discriminate between chaotic and non-chaotic behaviour that can be extended to many-body systems. Our approach returns to a semi-classical analysis for many-body systems. Albeit, the semi-classical dynamics that we study de-
scribes entanglement structure whose origin is quantum mechanical. We have applied this to the Ising model with a longitudinal and transverse field using the time-dependent variational principle applied to matrix product states.

This analysis has afforded several insights. An MPS description of the wavefunction suggests a new relationship between the Komogorov-Sinai entropy (and its dependence upon bond order) and the entanglement, Eq. (5.13). This relationship holds not just for the initial entanglement growth, but rather for the entire time-dependence of the entanglement. Using the thermofield MPS reveals a semi-circular distribution of Lyapunov exponents in the non-integrable case and Gaussian distribution in the integrable case. The former result has been anticipated in the context of gravitation [48, 49], where it was hypothesised that it may be universal.

There are several natural extensions of the present work. Similar Lyapunov spectra calculated for finite systems would enable comparison calculations of the out-of-time ordered correlator. The latter has become a canonical tool for studying quantum thermalisation [70–73]. When studying finite systems it may be more convenient to calculate Lyapunov exponents using a time-series approach [74–77]. This would involve extracting exponents from the evolution of observables, it is currently unclear if exponents can be accurately calculated in the quantum context using this approach.

**Effective Long-time Dynamics:** The exponential increase of data required to accurately describe the dynamics of a quantum system at early times and the ultimate decrease of this data at late times for thermalising quantum systems presents an acute difficulty for efficient numerical simulation. Whilst the mechanism of this decrease can be understood by dephasing, it is difficult to turn this insight into a practical scheme. The new perspective provided here might provide a route. The dynamical modes of a classically chaotic system divide naturally into those that have revealed their chaotic nature on a given timescale and those that have not. The latter behave as quasi-regular modes and the former as a chaotic bath for them. This division, suggests an appealing
way to describe late-time dynamics of the wavefunction MPS. On the longest timescales, the majority of modes form a bath, with energy density equal to that of the initial state. It ought to be possible to develop a Langevin description of this late-time dynamics. The cross over between early- and late-time dynamics being captured as the crossover from dominance of inertial dynamics to diffusive dynamics driven by the noise and dissipation. Such a description may be developed by adding noise and dissipation to the time-dependent variational principle, to derive an MPS Langevin equation. This picture provides a suggestive link to random circuit analyses of thermalisation [79–87].

However, our implementation of the time-dependent variational principle for the evolution of a thermofield MPS may obviate the need for such a Langevin description. Eigenstate thermalisation suggest that there should be an efficient description of both early and late time dynamics. If a single variational scheme can capture both limits — and if it is imbued with sufficient variables to surmount the information barrier in the middle of the dynamics — then it should be possible to obtain an accurate numerical description that runs from the earliest times to the latest times. The time-dependent variational principle applied to the thermofield MPS seems to satisfy these requirements. The remaining ingredient is to find a way of compressing the thermofield description at late times. As commented in Sec. 5.3.2 the multiple, equivalent descriptions of an infinite temperature state contains the essence of such a compression. Implementing this is a subject of ongoing investigation.

*Glimmers of a Quantum KAM Theorem:* Classical integrable systems show a remarkable robustness to perturbation. The KAM theorem shows that aspects of integrability remain through the presence of residual invariant tori (essentially periodic motions of action angle variables) when perturbations away from integrability are below some threshold. There has been speculation recently of whether such effects could be apparent in a quantum system [50]. It is *inevitable* that they are possible when quan-

---

10 Using the recently developed path integral over matrix product states [29, 78] from a Keldysh path integral, a Langevin equation can be constructed in the usual way.
tum dynamics is projected to (semi-)classical dynamics by observing on a finite window. This is a promising direction for future study, for example in the study of many body localisation.

**Thermalisation in Quantum Critical Systems:** Matters of thermalisation and chaotic dynamics come to a head in quantum critical systems. These are the most rapidly dissipating and de-phasing of quantum systems [88], and it is no coincidence that recent years has seen their mapping to black holes — through the AdS/CFT correspondence — themselves the most rapidly scrambling (classically chaotic) of objects [89]. The semicircle distribution of Lyapunov exponents that we have uncovered already makes links to works carried out in this context [48,49]. A direct application of MPS methods has limitations for the study of quantum criticality, however, because of the high levels of entanglement. It may be that other variational schemes such as MERA can do a better job, although in that case, dynamics are trickier. The view of quantum dynamics that we present should, then give an interesting complementary view of dynamical transitions observed after quenches and sweeps through quantum critical points.

To conclude, we have presented techniques that provide a bridge between thermalising quantum systems and classically chaotic Hamiltonian systems. Moreover, we have demonstrated how techniques developed in the latter may be applied fruitfully to the study of quantum thermalisation. We hope that this approach will provide a useful insights into other aspects of quantum thermalisation and chaos. As anticipated by many others, the intrinsic chaos of non-linear classical mechanics is the very property that permits the stability of a classical view of the underlying quantum world.
Chapter 6

Out-of-equilibrium dynamics in coupled quantum chains

Cold atom gas experiments make it possible to study out-of-equilibrium behaviour of many-body quantum systems more precisely than ever before. Despite much progress, there are significant questions about the dynamics of many-body systems at both short and late times. In this chapter we use the truncated spectrum approach and a number of analytical tools to study the behaviour of two one-dimensional Bose gases after they are suddenly coupled. We observe strong disagreements with existing results using a semiclassical approximation. Additionally, we consider the spreading of correlations throughout the system following a quench.

In this chapter TSA code developed by Andrew James and Robert Konik is used. All analytic calculations are my work. This is based upon work currently being prepared for publication.
6.1 Out of equilibrium dynamics

Since the creation of the first experimentally observed Bose-Einstein condensate in 1995 [90,91], the field of ultracold gases has developed into a key tool for the study of many-body quantum systems.

By trapping different atomic gases in optical lattices many toy Hamiltonians can be created. These experiments can accurately simulate Hamiltonians that can only be approximately realised in solid-state experiments. They can be created without complicating features such as lattice defects or impurities. Careful measurements can be made of these gases to extract complex properties of the system. For example, measures of the entanglement structure of the many-body wavefunction can be made using entanglement witnesses [92–94]. Additionally, real-time measurements of ultracold gases are now possible so nonequilibrium properties can also be studied [95].

These are topics of significant interest. The nonequilibrium properties of many-body quantum systems are still poorly understood. For example, we still do not fully understand the process by which a many-body quantum system, starting far from equilibrium, approaches its equilibrium state. A simple framework in which to study such a problem is a quantum quench [96]: a system is prepared in the ground state of an initial Hamiltonian before being evolving in time according to a final Hamiltonian. One can then ask questions such as, how do correlations spread through a system? How does the entanglement structure change? For certain systems there is a well-known picture due to Cardy and Calabrese in which correlations are spread by quasiparticles which move ballistically though the system [97]. If there is a maximum velocity of propagation in the system, then one expects to see a light cone effect in the equal time correlation functions. However, this picture breaks down in certain cases - for example when the quasiparticles form bound states [98,99].

Another unique feature of ultracold atoms is their isolation from an external envi-
ronment. This isolation raises interesting questions about late-time behaviour and in what sense a closed quantum system can thermalize. It is clear that not all systems relax to thermal states, instead many approach non-thermal steady states \cite{41}. For example, integrable systems relax to a \textit{generalized Gibbs ensemble}, characterised by many quasi-local conserved quantities \cite{100}. There is interest in probing not only these steady states but also the relaxation dynamics as they are approached \cite{101,103}. Theoretical techniques are needed to compliment the leap forwards that ultracold gas experiments represent.

In this section of the thesis we study the time evolution of one-dimensional bosonic chains which are suddenly coupled together. In this case we limit ourselves to coupling together two bosonic chains, although we hope this will illuminate some of the physics one should expect for many chains and some of the technical issues that may occur in simulating the many chain limit. In section \ref{sec:6.2} we introduce the description we use to describe the one-dimensional bosonic chains, the Luttinger liquid. In section \ref{sec:6.3} we describe the coupled chain system and its relation to the sine-Gordon model. In section \ref{sec:6.4} the techniques used to study the system will be explained. Finally in section \ref{sec:6.5} we describe our results and in section \ref{sec:6.6} we consider future directions for this work.

### 6.2 The Luttinger liquid

The Lieb-Liniger model is a widely studied one-dimensional bosonic model that can be accurately realized using cold atom experiments \cite{104}. It is a model of bosons that interact with a contact interaction of strength $c$:

\begin{equation}
H = \int_{0}^{R} dx \left[ \frac{1}{2m} \partial_{x} \psi^\dagger(x) \partial_{x} \psi(x) + c \psi^\dagger(x) \psi^\dagger(x) \psi(x) \psi(x) \right] \[ \psi(x, t), \psi^\dagger(y, t) \] = \delta(x - y). \tag{6.1}
\end{equation}
Figure 6.1: The relationship between the Luttinger parameter $K$ and the underlying bosonic theory. When $K \geq 1$ the Luttinger liquid represents the Lieb-Liniger model with increasing interaction strength $c$ as $K$ is decreased. For $K < 1$ the underlying theory has long range interactions.

While the Lieb-Liniger model is exactly solvable \cite{100,105}, we will not be working with it directly. Instead we consider a lower energy theory that can obtained from the Lieb-Liniger model through a process known as *bosonization* \cite{106,107}. The boson field $\psi(x)$ can be written in terms of the density ($\rho$) and the phase ($\theta$) as follows:

\begin{align}
\psi^\dagger(x) &= \sqrt{\rho}e^{-i\theta}, \\
\psi^\dagger(x)\psi(x) &= \rho = \left(\rho_0 - \frac{1}{\pi}\nabla\phi(x)\right) \sum_p \exp \left[2ip(\pi\rho_0x - \phi(x))\right], \\
[\rho(x), \theta(y)] &= i\delta(x - y), \\
\left[\frac{1}{\pi}\partial_x\phi(x), \theta(y)\right] &= -i\delta(x - y).
\end{align}
The last line can be integrated by parts to show that the momentum canonically conjugate to $\phi$ is

$$\Pi_\phi = \frac{1}{\pi} \partial_x \theta. \quad (6.6)$$

We can obtain the low energy theory - the *Luttinger liquid* - by expanding to quadratic order in terms of the $\phi$ and $\theta$ fields. We find that the Hamiltonian of the Luttinger liquid takes the following form:

$$H = \frac{v}{2\pi} \int_0^R dx \left[ K(\pi \Pi_\phi)^2 + \frac{1}{K}(\partial_x \phi)^2 \right] = \frac{v}{2\pi} \int_0^R dx \left[ \frac{1}{K}(\pi \Pi_\theta)^2 + K(\partial_x \theta)^2 \right]$$

$$= \frac{v}{2\pi} \int_0^R dx \left[ K(\partial_x \theta)^2 + \frac{1}{K}(\partial_x \phi)^2 \right]. \quad (6.7)$$

The Lagrangian can when written in terms of the $\theta$ field is:

$$L(\theta, \dot{\theta}, t) = \frac{K}{2\pi} \int_0^R dx \left[ \frac{1}{v}(\partial_t \theta)^2 - v(\partial_x \theta)^2 \right]. \quad (6.8)$$

The Luttinger liquid is defined in terms of two parameters, a velocity $v$ and the *Luttinger parameter* $K$. While we have introduced and motivated the Luttinger liquid in terms of the Lieb-Liniger model, it actually describes the low energy physics of a variety of different one-dimensional systems - both bosonic and fermionic.

When $K$ becomes large, fluctuations in the phase $\theta$, are strongly penalised leading it to become classical, while fluctuations of the displacement field are energetically favourable. Such behaviour describes a superfluid. $K \to \infty$ corresponds to the case of non-interacting bosons and $K = 1$ represents the $c \to \infty$ limit of the Lieb-Liniger model. Interestingly this point also corresponds to a system of free fermions.

$K = 1$ also represents the crossover from the phase field being fixed to the displacement field becoming increasingly classical. In this regime there is a tendency towards density wave formation. This captures the behaviour of bosons with long range interac-
The relationship between the Luttinger liquids and the underlying bosonic theory can be seen in figure 6.1.

6.2.1 Mode expansion

In order to better understand the Luttinger liquid, we will describe its behaviour in terms of a free boson on a ring, following the conventions in Di Francesco et al [109]. We will identify this boson with the phase field which possesses an angular character due to its exponential relation to the original interacting bosons. The action for the free bosonic field (with velocity $v=1$) is:

$$S = \frac{1}{8\pi} \int dx dt \partial_\mu \varphi(x,t) \partial^\mu \varphi(x,t).$$  \hfill (6.9)

We consider a general boundary condition for the free field:

$$\varphi(x + R, t) = \varphi(x, t) + 2\pi m/\beta,$$  \hfill (6.10)

where $R$ is the physical chain length and $\varphi(x, t)$ is angular in character, with its values confined to the circumference of a circle with radius $1/\beta$. With this condition, the field $\varphi(x, t)$ has the following Fourier expansion in terms of bosonic modes

$$\varphi(x, t) = \varphi_0 + \frac{4\pi \beta n}{R} x + \frac{2\pi m}{R \beta} t + \sum_{k \neq 0} \frac{1}{k} \left( a_k e^{2\pi ik(x-t)/R} + \bar{a}_{-k} e^{2\pi ik(x+t)/R} \right).$$  \hfill (6.11)

where $\varphi_0$ is the zero mode of the field, $n$ is an integer related to the canonical momentum conjugate to $\varphi(x, t)$ and $m$ is the winding number for the field $\varphi(x, t)$. Finally, $a_k$ and $\bar{a}_k$ are chiral (right or left moving) bosonic modes. The bosonic modes as defined above have unusual commutators:

$$[a_k, a_p] = k \delta_{k+p}, \quad [a_k, \bar{a}_p] = 0, \quad [\bar{a}_k, \bar{a}_p] = k \delta_{k+p}, \quad k, p \in \mathbb{Z} \neq 0.$$  \hfill (6.12)
so that $a_k, \tilde{a}_k$ are creation operators for $k < 0$ and annihilation operators for $k > 0$.

How does $\varphi(x,t)$ relate to the phase field $\theta(x,t)$? Setting $v = 1$ and comparing Lagrangian densities for $\theta(x)$ and $\varphi(x)$ - defined in Eqs. 6.8 and 6.9 - we find that

$$\frac{1}{2} \varphi(x) = \sqrt{K} \theta(x). \quad (6.13)$$

Then we use the boundary conditions, Eq. 6.10 to obtain

$$\theta(x + R) = \theta(x) + \frac{2\pi m}{2\sqrt{K} \beta} \quad (6.14)$$

Note that $\theta$ has a compactification radius of unity based on Eq. 6.2, so we must have

$$\beta = \frac{1}{2\sqrt{K}} \quad (6.15)$$

$$\theta(x) = \beta \varphi(x). \quad (6.16)$$

Hence we can use the free boson $\varphi(x,t)$ with $\beta$ to calculate quantities for the bosonized system with parameter $K$. We can now associate $\varphi_0$ with the global phase; $n$ with the number of bosons relative to the average density; and $m$ with the centre of mass momentum.

6.3 Coupled Luttinger liquids

6.3.1 Coupled Luttinger liquids and the sine-Gordon model.

We model interactions between two one-dimensional Bose gases by taking two Luttinger liquids and introducing a tunnelling term between them. In terms of the original Lieb-Liniger bosons, the interaction term will take the form:

$$H_{\text{int}} = -t_\perp \int_0^R dx \left( \psi_1^\dagger(x) \psi_2(x) + \text{H.c.} \right). \quad (6.17)$$
We will transform from fields defined on each chain, to symmetric/antisymmetric combinations, \( \varphi_{S/A} = \frac{\varphi_1 \pm \varphi_2}{\sqrt{2}} \). In terms of the symmetric/antisymmetric bosonic fields the Hamiltonian is

\[
H = H_{LL}(\varphi_S) + H_{LL}(\varphi_A) - 2t_\perp \int_0^R dx \cos(\sqrt{2}\beta \varphi_A(x)) \\
= H_{LL}(\varphi_S) + H_{s-G}(\varphi_A),
\]

(6.18)

where \( H_{LL}(\varphi_S) \) is a standard Luttinger liquid Hamiltonian and \( H_{s-G}(\varphi_A) \) is the sine-Gordon Hamiltonian [110]. The spectrum of the sine-Gordon model consists of soliton (and antisoliton) excitations called kinks. These kink excitations have a rest mass given by [111]

\[
M = \left( t_\perp \frac{\pi \Gamma(1 - \beta^2)}{\Gamma(\beta^2)} \right)^{\frac{1}{2-2\beta^2}} \frac{2\Gamma\left(\frac{\xi}{2}\right)}{\sqrt{\pi \Gamma\left(\frac{1}{2} + \frac{\xi}{2}\right)}},
\]

(6.19)

\[
\xi = \frac{\beta^2}{1 - \beta^2}.
\]

(6.20)

Moreover, for \( \beta < 1/\sqrt{2} \) there are also \( \lfloor \xi^{-1} \rfloor \) soliton - antisoliton bound states called breathers. The \( \ell \)th breather mass is given by

\[
M_\ell = 2M \sin\left(\frac{\ell \pi \xi}{2}\right).
\]

(6.21)

6.3.2 Quenching coupled Luttinger liquids.

We are interested in studying the dynamics of this system. Before \( t = 0 \) we prepare the two Luttinger liquids in their uncoupled ground states. At \( t = 0 \) we quench the system by suddenly changing the Hamiltonian, introducing a coupling between the two Luttinger liquids. This procedure takes the following form:
\[
\begin{align*}
t < 0 & \quad H = H_{1,\text{free}} + H_{2,\text{free}}, \\
t \geq 0 & \quad H = H_{1,\text{free}} + H_{2,\text{free}} - t_\perp \int_0^R dx \left( \psi_1^\dagger(x)\psi_2(x) + \text{H.c.} \right), \quad t_\perp \neq 0.
\end{align*}
\] (6.22)

Since the Hamiltonian of the symmetric field \(\varphi_S\) is independent of \(t_\perp\) only the antisymmetric field contributes in a non-trivial way to the dynamics.

It is important to note that this quench differs from those considered by Schmiedmayer et al. [102]. In these experiments a typical quench involves creating a single Luttinger liquid by confining the gas to one-dimension. Then at \(t = 0\) the gas is divided in two to create two separate Luttinger liquids. However the quench setup described here should contain similar sine-Gordon physics, and moreover it extends in a natural way to systems of more than two chains, a subject of future interest.

6.4 Methods

6.4.1 The truncated spectrum approach

We study the system using the truncated spectrum approach [112–114]. Suppose that we are considering a multi chain system of the following form:

\[
H_{2D} = \sum_i \left[ H_{1D,i} + J_\perp \int_0^R dx O_i^1(x)O_{i+1}^2(x) \right].
\] (6.23)

This is a quasi-two dimensional system constructed by taking exactly solvable one-dimensional chains (with periodic boundary conditions) and perturbing them by introducing nearest neighbour couplings between them.

Working with exactly solvable 1D Hamiltonians, \(H_{1D}\) means that we know their exact spectrum and matrix elements. This may appear to be a severe restriction but there are a large set of physically interesting Hamiltonians that are exactly solvable e.g.
Heisenberg spin chains, Hubbard models etc.

The 1D models in equation [6.23] have an infinitely large Hilbert space. By working with periodic chains of size $R$ the spectrum becomes discrete. We then introduce an energy cutoff $E_C$ and only keep eigenstates of the individual 1D systems with an energy less than $E_C$. After this truncation we can in principle form the finite dimensional Hamiltonian for the coupled systems and exactly diagonalize it. The dynamics of the system can also be calculated by forming the time evolution operator in the restricted low energy sector and time evolving an initial state.

The truncated spectrum approach suggests that for relevant (in an RG sense) inter-chain interactions, the low energy sector of a perturbed integrable system is primarily described as a mixture of low energy eigenstates of the unperturbed system. Therefore, provided $E_C$ is sufficiently high, the correct dynamics can be captured even for a finite Hilbert space. Relevant interactions flow to zero in the UV limit, it is therefore reasonable to think the perturbation is less significant at higher energies.

A cylindrical system can be studied as if it were a one-dimensional chain. In this case each site of the chain is the 1D system described by $H_{1D}$ with an energy cutoff $E_C$. In what follows we restrict ourselves to the case of two coupled chains. The aim is to use this simpler model as a guide before a future investigation of a many-chain system. We note that the time evolution of a 2D Bose system on a square lattice with strong nearest neighbour repulsion has been studied using sophisticated MPO techniques in Ref. [115].

6.4.2 Semiclassical limit of sine-Gordon model.

We also analytically study the system by making a semiclassical approximation. We expand the cosine term which leads to interactions between the two chains to quadratic order [116],

$$2t_\perp \int_0^R dx \cos \left( \sqrt{2} \beta \varphi_A(x) \right) \to t_\perp \int_0^R dx \left( \sqrt{2} \beta \varphi_A(x) \right)^2. \quad (6.24)$$
The new two-chain Hamiltonian can be diagonalised using a Bogoliubov transformation and its dynamics can be calculated exactly. This follows the work of Foini and Giamarchi [116]. Details can be found in Appendix E.

6.5 Results

(a) Comparison of the time evolution of the first mode occupation calculated for the full cosine potential (using TSA) versus the semiclassical result.

(b) Comparison of the time evolution of the second mode occupation calculated for the full cosine potential (using TSA) versus the semiclassical result.

Figure 6.2: Mode occupation of coupled Luttinger liquids as a function of time comparing the semiclassical calculation to the full interaction. These results are generated for $K = 25 (\beta = 0.1)$, $R = 100m$, $v = 1m/s$ and $t_\perp = 0.1Hz$.

Disagreement between TSA and semiclassics: In our work we have observed major discrepancies in the dynamics between the semiclassical Hamiltonian and the full sine-Gordon Hamiltonian. To demonstrate this, we calculate the time-evolution of the mode occupation of each of the bosonic modes $n_{i,k} = \langle \hat{a}_{i,k}^\dagger \hat{a}_{i,k} \rangle$, the full calculation can be found in Appendix E. Using unitary perturbation theory [117], we find that the evolution of the bosonic modes for the full cosine interaction at short times is

$$\langle n_{k,i}(t) \rangle \approx \frac{1}{k^2} 4R^2 \beta^3 \left( \frac{2\pi}{R} \right)^{4\beta^2} t^2 t_\perp^2.$$

In comparison, expanding the cosine to quadratic order we find that the mode occupation
Figure 6.3: Density plots of $C[x,t] - C[R/2,t]$ for coupled Luttinger liquids as a function of time using a semiclassical approximation. These results are generated for $K = 25$ ($\beta = 0.1$), $R = 200\mu m$, $v = 1600\mu m/ms$, $t_\perp = 20Hz$. (a) is a plot for the initial state $|0,0\rangle$, (b) is a plot for the initial squeezed state. The largest values of the correlator are cut off in the plot to improve contrast.

at short times is

$$\langle n_{k,i}(t) \rangle_{SC} \approx \frac{1}{k^2} 8R^2 \beta^4 t^2_\perp.$$  \hspace{1cm} (6.26)

Comparing these two calculations it is clear that these two results differ

$$\langle n_{k,i}(t) \rangle_{SC} = \left(\frac{2\pi}{R}\right)^4 \beta^2 \frac{1}{2} \langle n_{k,i}(t) \rangle.$$  \hspace{1cm} (6.27)

Note, even at small $\beta$ these results do not agree except at $t = 0$. This relation (at short times) is borne out by numerical evaluation of the dynamics for the full (cosine) model versus the semiclassical result. Examples are shown in Figs. 6.2a and 6.2b. Moreover, this is merely the difference at first order. Numerically it is clear that the difference grows more significant as $t$ becomes large.
Figure 6.4: Plots of $C[x,t]$ at different timeslices for coupled Luttinger liquids as a function of time using a semiclassical approximation. These results are generated for $K = 25$ ($\beta = 0.1$), $R = 200 \mu m$, $v = 1600 \mu m/ms$, $t_\perp = 20 Hz$. (a) is a plot for the initial state $|0, 0\rangle$, (b) is a plot for the initial squeezed state.

**Lightcone of correlations:** We now calculate

\[ C[x,t] = \langle e^{i(\theta_A(x,t) - \theta_A(0,t))} \rangle \]
\[ = e^{\frac{1}{2}[(\theta_A(x,t) - \theta_A(0,t))^2]} , \]

performing the calculation for the semiclassical Hamiltonian using two different initial states. Note that the second inequality is only true for the semiclassical Hamiltonian.

While we expect the behaviour of this correlator to be significantly different for the full Hamiltonian, it will be a useful comparison for future work.

The first of the initial states is each of the chains prepared in their uncoupled ground state $|0\rangle_1 \otimes |0\rangle_2$. From equation 6.29 it is clear that $C[x,t]$ can written in terms of \( \langle \theta_A(x,t)\theta_A(y,t) \rangle \). In Appendix E we calculate this for the $|0\rangle_1 \otimes |0\rangle_2$ and find it is:

\[ \langle \theta_A(x,t)\theta_A(y,t) \rangle = \frac{1}{R^2} \theta_0^2 + \beta^2 \sum_{q \in Z \geq 0} \frac{4}{|q|} \cos \left( \frac{2\pi i q (x - y)}{R} \right) e^{2\phi_q} (\cosh 2\phi_q + \cos 2\omega_q t \sinh 2\phi_q) . \]

(6.30)

where $\omega_p = \sqrt{p^2 + 16\pi \beta^2 t_\perp}$, and $\phi_p = \frac{-t_\perp}{t_\perp + Kvp}$ is the angle which defines the Bogoliubov bosons diagonalising the Hamiltonian. The sum over $q$ is a sum over bosonic modes as
Our results are generated for $K = 25$ ($\beta = 0.1$), $R = 200\mu m$, $v = 1600\mu m/ms$, $t_\perp = 20Hz$. (a) is a plot for the initial state $|0,0\rangle$, (b) is a plot for the initial squeezed state.

We compare results for this initial state to the behaviour of the system initially prepared in a squeezed state

$$|\psi_0\rangle = \exp \left[ \sum_p W_p b_p^\dagger b_{-p}^\dagger \right] |0,0\rangle .$$

A similar expression for $\langle \theta_A(x,t)\theta_A(y,t) \rangle$ in this case can be found in appendix E. Density plots of $C[x,t] - C[R/2,t]$ for the two initial states can be seen in Fig 6.3a and Fig 6.3b. A lightcone is significantly easier to see when the initial state is the squeezed state. This can also be clearly seen by looking at timeslices of this density plot. As can be seen in figures Fig 6.4a and Fig 6.4b, there is a bump in the correlation function which moves out with time in both cases, however in the case of the squeezed state this bump stands out much more clearly from the background behaviour.

The lightcone can be more easily seen for the vacuum state by considering the deriva-
Figure 6.6: Density plots of $\frac{\partial C[x,t]}{\partial x}$ for coupled Luttinger liquids as a function of time using a semiclassical approximation for various $q_{\text{max}}$. These results are generated for $K = 25$ ($\beta = 0.1$), $R = 200\mu m$, $v = 1600\mu m/ms$, $t_\perp = 20Hz$.

Unfortunately, detecting a lightcone for the full cosine interaction is quite numerically
intensive. A clear lightcone can only be detected in the semiclassical case after dozens of different bosonic modes in equation 6.30. An example can be seen in figure 6.6; a lightcone has only clearly emerged by $q_{\text{max}} = 50$. To capture such a high number of bosonic modes using the truncated spectrum approach, these modes could involve states with a very high energy and therefore a very high cutoff is necessary. Reproducing even the $q_{\text{max}} = 20$ results involves hundreds of states in the TSA.

6.6 Discussion

In this chapter we studied the dynamics of two coupled Luttinger liquids using the truncated spectrum approach and a semiclassical approximation. The semiclassical approximation appears to give a relatively poor account of the full dynamics of the system, as can be seen by the discrepancy between the mode occupations in figure 6.2. However using the semiclassical model one can make progress analytically and compute equal time correlation functions. These suggest, at least qualitatively, that a lightcone should be visible in the correlations. An interesting question is then whether this behaviour will persist for the full cosine version of the model, which admits breather bound states that might be expected to suppress the spread of correlations. Unfortunately, a high momentum and energy cutoff is necessary to observe a lightcone of correlations for the full cosine interaction using the truncated spectrum approach. Such high cutoffs are beyond the current reach of the code used, but a more sophisticated version that does not explicitly diagonalize the truncated (but still very large) two chain Hamiltonian is in development.

Going forwards, these techniques could similarly be applied to study correlations in multi-chain systems by combining them with matrix-product state algorithms. Such methods have already been applied to coupled Ising chains [118].
Chapter 7

Machine learning using tensor networks

In this chapter we apply tensor network techniques to compress neural networks. Neural networks have proven to be incredibly effective at tasks such as image recognition. However, this comes at the cost of needing millions or billions of parameters to perform effectively. By replacing the largest layers of a neural network with layers inspired by the multiscale entanglement renormalisation ansatz these layers can be compressed by a factor of 14,000 with a minor drop in accuracy.

These results were generated using machine learning code written equally with Edward Grant. The original idea was due to Andrew G. Green. Vid Stojevic and Simone Severini contributed significantly. This chapter is based upon work which can be found in Ref [119]

7.1 A brief overview of machine learning

Since the internet became widely used in the late 1990s the amount of data being shared between users has exploded. There is significant interest in interpreting the large volume of videos, images and text being shared. For example, email software is designed to automatically filter spam emails, smartphone cameras automatically identify faces or search engines suggest the optimal results given a particular search.
The appropriate tool to detect meaningful patterns in data is *machine learning*. While machine learning algorithms have existed for decades, the development of powerful GPUs in 2000s suddenly made these algorithms useful for a characterising large datasets. *Neural networks* are some of the most widely used machine learning algorithms. Inspired by a biological brain, neural networks are a network of nodes sorted into layers. The output of a node in one layer is a non-linear function of its inputs. This output is then fed into the next layer of the network. A neural network can be used to perform tasks such as image classification: given an image sort it into one of a number of different classes. An image is fed into the first layer of a neural network and the class which it falls into is the output of the final layer. An example of this network can be seen in figure 7.1.

Unfortunately, a neural network has significant drawbacks for image classification. A typical image may be 256 × 256 pixels. Reshaping the image into a vector and

![Figure 7.1: An example of a neural network used for a classification task. An image is fed into the network and the network places it in to a particular class. The network is optimised by varying the weight between the nodes of the network.](image-url)
inputting it into the network would require an intractable number of parameters. The solution is to use a *convolutional neural network*. A convolutional neural network uses a series of convolutional mask and coarse graining steps to reduce the image down to its most important features. These can then be fed into a conventional neural network. Convolutional neural networks are now state-of-the-art for image classification and are widely used in both academia and industry. Typically convolutional neural networks still require millions or billions of parameters. Previous work has demonstrated that CNNs can be highly compressed without significant reduction in network performance, suggesting significant redundancy [120], [121], [122].

In this chapter we use tensor network methods - which are effective at compressing quantum states - to compress convolutional neural networks. We parametrise layers of the convolutional neural networks in terms of the multi-scale entanglement renormalization ansatz (MERA) and use the compressed network to classify the CIFAR-10 and CIFAR-100 image datasets.

In section 7.2 we give a brief overview of the structure of a convolutional neural network. In section 7.3 we explore the relationship between neural networks, physics and tensor networks. In section 7.4 we explain how tensor networks can be used to compress CNNs. In section 7.5 we fully explain the structure of the CNN used and in section 7.6 we state some of the results found for the compressed neural networks.

### 7.2 Convolutional neural networks

A convolutional neural network is constructed using four main operations.

1. The first of these is the *convolution operator*. Convolutions are filters which can be applied to an image to create a new image, with the filter emphasising certain details.

   As an example, suppose we apply a convolution to a greyscale image where each
Figure 7.2: Convolutions can be used to filter images and detect important features. Here a convolution matrix which detects edges is applied to a greyscale image. Image is courtesy of Jonathan Chng on Unsplash.

A pixel of the new image is defined by placing this convolution matrix over that pixel in the original image, taking the element-wise product and summing the result. The convolution defined in equation (7.1) can be used to detect edges in an image as can be seen in figure 7.2.

2. The second operation used is an elementwise nonlinear function. After a convolutional layer of a CNN a nonlinear function is applied. Commonly chosen functions
include \( f(x) = \tanh(x) \) and ReLU,

\[
f_{ReLU}(x) = \max(0, x). \tag{7.2}
\]

Without the use of nonlinear functions all layers of a neural network can be combined, the network effectively becomes one layer deep.

3. The third operation used in CNNs is a **pooling layer**. Pooling layers are a type of coarse graining used to reduce the dimension of the data. A pooling layer applied to an image maps an \( n \times n \) block of pixels and maps them to a single pixel in a smaller image. **Max pooling** - taking the maximum pixel value of the \( n \times n \) block - is the most commonly used method in CNNs.

4. The final element of a CNN are the **fully connected layers**. These are layers in which every nodes is connected to every node in the next layer, as can be seen in figure 7.1. After several fully connected layers there is a final output layer which classifies the image. In this work we only replace the fully connected layers of the CNN with tensor network inspired layers.

All of these operations can be applied to form a convolutional neural network for tasks such as image classification, an example CNN can be seen in figure 7.3. The *weights* in the fully connected layers and the values in the convolutional filters are chosen so that the CNN does the best job of classifying images. This can be achieved by training of a initial set of images. If \( C(x_i) \) is the correct class of an image \( x_i \) and \( C'(x_i) \) is the class as predicted by a CNN, the parameters in the CNN are varied to minimise the *loss function*

\[
E = \sum_{i=1}^{N} ||C(x_i) - C'(x_i)||^2. \tag{7.3}
\]

This is optimization is often achieved using a gradient descent algorithm called *back-*
7.3 Connecting machine learning to physics

*Machine learning and renormalization:* There is a close resemble between machine learning algorithms and renormalisation in physics. Both involve throwing away irrelevant details of data while preserving the most important features. There have been a number of works looking at the relationship between renormalization and deep learning. Ref. [123] argue that the effectiveness of deep neural networks should be thought of in terms of renormalization and ref. [124] demonstrate an exact mapping between the variational renormalization group and restricted Boltzmann machines. *Machine learning and tensor networks:* In recent years a second significant connections between machine learning and physics has been made. In 2017 Levine *et al.* used a tensor network con-
struction of a neural network to argue that quantum entanglement was a useful measure of the appropriate network structure for a neural network \[125\].

There have been other significant applications of tensor networks to machine learning. Matrix product operators - in this context called tensor trains - have previously been deployed to compress neural networks \[120,121\].

7.4 Tensor Factorization of Fully Connected Layers

Figure 7.4: Schematic diagrams of various tensor factorizations of linear layers. a) a general linear layer, b) its tensor train factorization. The squares represent smaller tensors. Connections represent contractions as indicated in Eq.(1). c) Tree network factorization. d) MERA factorization. placeholder graphic

In this chapter we replace the fully connected layers of a CNN seen in figure 7.3 tensor network layers. We consider two tensor network compressions. The first of these is a matrix-product state inspired compression, already applied in CNNs under the name tensor trains. \[120,121\]. The matrix product state factorization of a linear layer will take
the form seen in figure 7.4 (b).

In chapter 2 tree tensor networks were also introduced. Tree networks were built from isometries, rank-3 tensors which reduced the size of the system by half after each layer. The structure is intended to capture short-range correlations on the lowest layer of the system with long range correlation being captured in the highest layers. A tree tensor network inspired layer uses tree structure although the rank-3 tensors are replaced with rank-4 tensors, the extra index going directly into the next layer. An example can be seen in figure 7.4 (c).

Of course, as we noted in 2, tree tensor networks have a major issue. Not all short range correlations are captured at the lowest layers of the network, consider inputs 4 and 5 in 7.4 (c). Unlike other neighbouring inputs, they are not connected until the final layer of the tree. The MERA (126) factorization was introduced in order to solve this problem. MERA adds additional rank-4 disentanglers which overlap with neighbouring tensors of the tree tensor network. The consequence of the disentanglers is to cause all correlations on the same length scale to be treated similarly.

7.5 Experiments & network structure

We have considered the performance of a neural network with the two penultimate fully connected layers of the model replaced with MERA layers, similar to the study of compression of fully connected layers using matrix product states. We have quantified the performance of the MERA layer through comparisons with two other classes of networks: fully connected layers with varying numbers of nodes and matrix product layers with varying internal dimension. The three types of network are otherwise identical.

The networks consisted of three sets of two convolutional layers each followed by max pooling layers defined with $3 \times 3$ matrices and stride 2. The matrices defining the
convolutional layers were $3 \times 3$. The final convolutional layer was followed by two more layers, these were either fully connected, MERA layers or MPS-layers depending upon the network. The final layer had 10 or 100 nodes corresponding to the image classes in CIFAR-10 and CIFAR-100. The initial values of all the parameters in the network was chosen by sampling from a Gaussian distribution with with standard deviation equal to $\frac{1}{\sqrt{n_{in}}}$, where $n_{in}$ was the number of inputs [127].

In this report we considered networks with two varieties of fully-connected layers. The first of these networks had a $4096 \times 4096$ fully connected layer followed by one which was $4096 \times 64$; this network was used as a benchmark against which the other models could be compared. The second network instead had a $4096 \times n$ fully connected layer followed by a $n \times 64$ layer where $n = 5$ for the CIFAR-10 network and $n = 10$ for the CIFAR-100 network. We trained these networks to compare the MERA and tensor train layers to a fully connected model with a comparable number of parameters, in order to evaluate how detrimental naive compression is to accuracy.

A schematic of the two MERA layers can be found in Figure 7.3. The input to the first MERA layer was reshaped into a rank-12 tensor with each index being dimension 2, as described in Section 2. The MERA layer was then constructed from a set of rank-4 tensors using the method described in Section 2. The first MERA layer works as follows: It contains a column of 6 rank-4 tree elements, followed by 3 tree elements and finally a single tree element. 5 disentanglers are placed before the first column of tree elements and 2 more disentanglers are placed before the second column of tree elements. The second MERA layer has an identical structure to the first MERA layer, one of the outputs of the first set of tree elements is fixed. As a result the output of the second MERA layer is 64 nodes.

Finally, a network with its fully connected layers replaced with a matrix product decomposition was trained in order to provide a comparison with the MERA layers. The matrix product layers were constructed using matrix product states with bond
Figure 7.5: A schematic of the MERA layers of the model. The small rectangles represent linear elements to factorize a general linear layer. White rectangles represent disentanglers. Red rectangles represent tree elements. Solid black lines connecting nodes represent tensor contraction and dashed lines with arrow heads represent the nonlinearities being applied. Dashed lines ending in a circle represent fixed outputs.

dimension $\chi = 3$. In the second tensor train layer, half of the output indices were fixed to match the second MERA layer.

We tested performance on the CIFAR-10 and CIFAR-100 datasets. We used 45,000 images for training, 5,000 for validation and 10,000 for testing. Training data was augmented by randomly flipping and translating the input images by up to 4 pixels in order to make the network more robust to minor errors in the images.

7.6 Experimental results

In Table 1 we compare the different models described in section 3 trained on the CIFAR-10 dataset. The compression rate stated is with respect to the number of parameters used in a fully-connected state of the art benchmark model, FC-1.

When comparing the MERA network to the fully connected model, FC-1 we see a considerable drop in the number of parameters required with only a modest drop in the accuracy of the network. MERA compresses the fully connected layers by a factor of 14,000 with a drop in the accuracy of only 0.4%. We do not attempt to compress
the convolutional layers in this work so in the MERA network the vast majority of the parameters are used in the convolutional layers which are identical to the fully connected model.

How significant is the MERA network structure we have chosen to the results obtained? To test this we compare the MERA results obtained to the fully connected model with many fewer parameters in the fully connected layers, FC-2. Despite having around 20 times more parameters in the fully connected layer than the MERA model, the MERA model significantly outperforms FC-2, with a 1.2% drop in the accuracy of FC-2 compared to MERA.

The MERA network also compares favourably to a tensor train network. In this case, the two networks have a comparable number of parameters but the MERA appears to achieve a higher accuracy than the tensor train network in this case.

Results for the CIFAR-100 model can be seen in Table 2. While none of the networks are as accurate as the benchmark case, the MERA network continues to outperform the tensor train and ablated fully connected network. However, the reduction in accuracy compared to the fully connected network is larger than for the CIFAR-10 dataset.

In addition to the degree of compression achieved by these networks, we also address the time to optimize. There is evidently a degree of compromise required here: the number of multiplications required to apply a MERA layer scales with the input size $N$ and bond order $D$ as $N \log_2 D$. The equivalent scaling for a matrix product and fully connected layer are $ND^2$ and $N^2$, respectively. This is reflected in the times taken to optimize these networks. Note however, that MERA can accommodate correlations at all scales of its input even at low bond order, whereas tensor trains require a bond order that scales exponentially with the length scale of correlation (128). MERA is, therefore, expected to scale better for very large data sets than either matrix product or fully connected layers.
Table 7.1: The CIFAR-10 experimental results for the different models. FC1 was the state of the art fully-connected model and FC2 was the fully-connected model with severely reduced number of parameters in the fully-connected layers. MERA are the result for the MERA inspired network. Finally MPS is the matrix product model with the bond dimension $\chi = 3$.

<table>
<thead>
<tr>
<th>Network</th>
<th>Parameters (FC Layer)</th>
<th>Parameters (Total)</th>
<th>Compression (FC Layer)</th>
<th>Compression (Total)</th>
<th>Accuracy</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>FC-1</td>
<td>17,040,000</td>
<td>17,336,640</td>
<td>1</td>
<td>1</td>
<td>88.9</td>
<td>0.2</td>
</tr>
<tr>
<td>FC-2</td>
<td>21,440</td>
<td>318,080</td>
<td>795</td>
<td>54.5</td>
<td>86.5</td>
<td>0.8</td>
</tr>
<tr>
<td>MERA</td>
<td>1192</td>
<td>297,832</td>
<td>14,295</td>
<td>58.21</td>
<td>88.5</td>
<td>0.1</td>
</tr>
<tr>
<td>MPS</td>
<td>1312</td>
<td>297,952</td>
<td>12,987</td>
<td>58.19</td>
<td>87.9</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Table 7.2: The CIFAR-100 experimental results for the different models. FC1 was the fully-connected model and FC2 was the fully-connected model with severely reduced number of parameters in the fully-connected layers. MERA are the result for the MERA inspired network. Finally MPS is the matrix product model with the bond dimension being 3.

<table>
<thead>
<tr>
<th>Network</th>
<th>Parameters (FC Layer)</th>
<th>Parameters (Total)</th>
<th>Compression (FC Layer)</th>
<th>Compression (Total)</th>
<th>Accuracy</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>FC-1</td>
<td>17,045,760</td>
<td>17,342,400</td>
<td>1</td>
<td>1</td>
<td>61.8</td>
<td>0.7</td>
</tr>
<tr>
<td>FC-2</td>
<td>48,000</td>
<td>344,640</td>
<td>355</td>
<td>50.3</td>
<td>53.4</td>
<td>0.6</td>
</tr>
<tr>
<td>MERA</td>
<td>6952</td>
<td>303,592</td>
<td>2451</td>
<td>57.12</td>
<td>58.4</td>
<td>0.6</td>
</tr>
<tr>
<td>MPS</td>
<td>7072</td>
<td>303,712</td>
<td>2410</td>
<td>57.10</td>
<td>57.9</td>
<td>0.6</td>
</tr>
</tbody>
</table>

7.7 Discussion

We have shown that replacing the fully connected layers of a deep neural network with layers based upon the multi-scale entanglement renormalization ansatz can generate significant efficiency gains with only small reduction in accuracy. When applied to the CIFAR-10 data we found the fully connected layers can be replaced with MERA layers with 14,000 times less parameters with a reduction in the accuracy of less than 1%. The model significantly outperformed compact fully connected layers with $70 - 100$
times as many parameters. Moreover, it outperformed a similar replacement of the fully connected layers with matrix product states, both in terms of accuracy for a given compression and compression for a given accuracy. While the MERA layer resulted in a larger accuracy drop in the CIFAR-100 case, it still outperformed a comparable tensor train network.

An added advantage — not explored here — is that a factorized layer can potentially handle much larger input data sets, thus enabling entirely new types of computation. Correlations across these large inputs can be handled much more efficiently by MERA than by tensor trains. Moreover, a compressed network may provide a convenient way to avoid over-fitting of large data sets. The compression achieved by networks with these factorized layers comes at a cost. They can take longer to train than networks containing the large fully connected layers due to the number of tensor contractions required to apply the factorized layer.

Our results suggest several immediate directions for future inquiry. Firstly, there are some questions about how to improve the existing model. For example, before the MERA layer is used the input is reshaped into a rank-12 tensor. There isn’t a well defined method for how to perform this reshaping optimally and some experimentation is necessary. The best way to initialize the MERA layers is also still an open question.

The results presented here are a promising first step for using MERA in a more fundamental way. Since MERA can be viewed as a coarse graining procedure (as explained in Section 2), and image data is often well represented in a hierarchical manner, one possibility would be to simply train a two-dimensional MERA directly on an image dataset, with no reference to a neural network. In [129] a similar idea was explored with matrix product states being trained directly on MNIST. An alternative possibility would be the replacement of just the convolutional layers of the network with a two-dimensional MERA. Both of these approaches would be closer in spirit to the fundamental ideas about the relationships between quantum physics and machine learning.
proposed in \cite{123} and \cite{124}.

Additionally, there has been some work using entanglement measures to explore how correlations are distributed in deep neural networks, and then utilizing these in order to optimize the design of networks (\cite{130}, \cite{125}). It would be intriguing to explore such ideas using MERA, for example by using the concrete MERA model explored in this paper, or one of the more ambitious possibilities mentioned above.

We end by noting two facts: any variational approximation to a quantum wavefunction can be used to construct a replacement for linear layers of a network. There are many examples and each may have its sphere of useful application. Moreover, quantum computers of the type being developed currently by several groups are precisely described by a type of tensor network (a finite-depth circuit - and one that may very soon be too large to manipulate classically) and could be used as direct replacement for linear layers in a hybrid quantum/classical neural computation scheme.
Chapter 8

Conclusion

Tensor networks are a unique approach for studying complex quantum systems. In this thesis we have introduced and expanded upon a number of techniques based upon tensor networks.

In chapter 3 we gave an overview of the application of tensor network techniques to the Feynman path integral. Path integrals poorly capture quantum entanglement. Writing a path integral as a sum over tensor network states overcomes this issue. The saddle point of a tensor network path integral is described by the time-dependent variational principle. In chapter 3 we introduced the techniques needed to calculate quadratic order fluctuations around a tensor network saddle point. We then calculated corrections to observables using these methods in chapter 4. There are a number of possible extensions of this work.

Firstly, the only tensor network for which a path integral was constructed was matrix-product states. In principle a path integral can be constructed for any set of tensor networks for which a resolution of the identity can be found. This includes hierarchical networks such as the multiscale entanglement renormalization ansatz or continuous ansatz such as continuous matrix product states. A second avenue is the calculation of further corrections to matrix product state path integrals. This could include nonper-
turbative corrections such as instantons.

In chapter 5 we considered an interesting application of tensor network techniques to out-of-equilibrium dynamics. The approach to thermal equilibrium was explained through the chaotic dynamics of a system on manifold of matrix product states. We also observed a relationship between the chaotic dynamics and the growth of entanglement entropy. Using the path integral techniques of chapter 3 it may be possible to rigorously derive this relationship in general.

We also studied out-of-equilibrium dynamics in chapter 6. In this case we studied the dynamics of one dimensional Bose gases that were suddenly coupled together, a scenario that can be realised in cold atomic gas experiments. We demonstrated the inadequacy of semiclassical methods for studying these systems compared to the truncated spectrum approach. These techniques could also be applied to many coupled bosonic chains.

Finally, we considered the application of tensor networks to machine learning algorithms. Tensor networks have been incredibly successful at compressing exponentially large quantum wavefunctions down to a manageable size, in chapter 7 we saw that they can similarly compress large neural networks. We found that networks compressed using tensor networks required fifty times fewer parameters without a significant drop in performance.

In summary, while tensor networks are already an incredibly powerful method for understanding quantum systems they are still limited in a number of ways. In this thesis we have combine tensor networks with a number of different methods, greatly increasing their applicability.
Appendix A

Spin-waves around an MPS saddle point

Here we calculate spin-wave fluctuations around an MPS saddle point $A_0$ for the transverse-field Ising model. We’ll use the following notation throughout this appendix:

$$\langle \psi(A_0) | \sigma_j^\alpha | \psi(A_0) \rangle = \alpha,$$

$$\langle \psi(A_0) | \sigma_j^\alpha \sigma_j^{\beta+1} | \psi(A_0) \rangle = (\alpha \beta).$$

We now consider a spin wave expansion about the mean field solution. We perform this spin-wave expansion in the follows manner:

$$\sigma \rightarrow \sigma - m \times \sigma + \frac{1}{2} (m(\sigma \cdot m) - \sigma (m \cdot m))$$

Consider the term $\langle A | \sigma_j^\alpha | A \rangle$,

$$\langle A | \sigma_j^\alpha | A \rangle \rightarrow h \langle A | \sigma_x - \epsilon_{xab} m^a_i \sigma_b + \frac{1}{2} m^a_i (\sigma \cdot m^i) - \frac{1}{2} \sigma_x (m^i \cdot m^i) | A \rangle$$

$$= x[1 - \frac{1}{2} |m^i|^2] - h\epsilon_{xab} m^a_i b.$$
Where we have used the fact that $m \cdot (\sigma) = 0$. Now consider the term $\langle A | \sigma_z^j \sigma_z^{j+1} | A \rangle$,

$$
\langle A | \sigma_z^j \sigma_z^{j+1} | A \rangle \rightarrow -J \langle A | (\sigma_z - \epsilon z a b m_a^j \sigma_b + \frac{1}{2} m_z^j (\sigma \cdot m^j) - \frac{1}{2} \sigma_z (m^j \cdot m^j))
$$

$$=
\langle (zz) (1 - \frac{1}{2} m^j \cdot m^j) \rangle - \epsilon z a b m_a^j \sigma_b - \epsilon z c d m_c^j (dz) + \frac{1}{2} m_z^j (z e) m^j + \frac{1}{2} m_z^j (z e) m^j + O(m^3)
$$

$$=
\langle (zz) (1 - \frac{1}{2} m^j \cdot m^j) \rangle - \epsilon z a b m_a^j \sigma_b - \epsilon z c d m_c^j (dz) + m_x^j m_x^{j+1} (yy) + m_y^j m_y^{j+1} (xx) + \frac{1}{2} m_z^j m_z^j (zx) + \frac{1}{2} m_z^j m_z^j (zz) + \frac{1}{2} m_z^j m_z^{j+1} (zx) + \frac{1}{2} m_z^j m_z^{j+1} (zz)
$$

$$=
\langle (zz) (1 - \frac{1}{2} m^j \cdot m^j) \rangle - \epsilon z a b m_a^j \sigma_b - \epsilon z c d m_c^j (dz) + m_x^j m_x^{j+1} (yy) + m_y^j m_y^{j+1} (xx) + \frac{1}{2} m_z^j m_z^j (zx) + \frac{1}{2} m_z^j m_z^j (zz) + \frac{1}{2} m_z^j m_z^{j+1} (zx) + \frac{1}{2} m_z^j m_z^{j+1} (zz)
$$

$$=
\langle (zz) (1 - \frac{1}{2} m^j \cdot m^j) \rangle - \epsilon z a b m_a^j \sigma_b - \epsilon z c d m_c^j (dz) + \frac{1}{2} m_z^j m_z^j (zx) + \frac{1}{2} m_z^j m_z^j (zz)
$$

$$=
\langle (zz) (1 - \frac{1}{2} m^j \cdot m^j) \rangle - \epsilon z a b m_a^j \sigma_b - \epsilon z c d m_c^j (dz) + \frac{1}{2} m_z^j m_z^j (zx) + \frac{1}{2} m_z^j m_z^j (zz)
$$

$$-
\epsilon z a b m_a^j \sigma_b - \epsilon z c d m_c^j (dz) + \frac{1}{2} m_z^j m_z^j (zx) + \frac{1}{2} m_z^j m_z^j (zz)
$$

$$- \frac{1}{2} m_z^j m_z^j (zx) + \frac{1}{2} m_z^j m_z^j (zz)
$$

$$= (A.5)$$
Where we use the fact that \( y = 0 \). We define \( \mathbf{m} = m_1 \hat{\theta} + m_2 \hat{\phi} \) and therefore \( m_x = \hat{z} m_1 \), \( m_y = m_2 \) and \( m_z = \frac{\hat{z}}{\hat{y}} m_1 \) where \( \sigma^2 = \langle z^2 + x^2 \rangle \). The Hamiltonian takes the form:

\[
H = \sum_j -h_x[1 - \frac{1}{2}(m_1^j)^2 - \frac{1}{2}(m_2^j)^2]
- (zz)(1 - \frac{z^2}{2\sigma^2}(m_1^j)^2 - \frac{1}{2}(m_2^j)^2 - \frac{z^2}{2\sigma^2}(m_1^{j+1})^2 - \frac{1}{2}(m_2^{j+1})^2)
- (zz)(1 - \frac{z^2}{2\sigma^2}(m_1^{j-1})^2 - \frac{1}{2}(m_2^{j-1})^2 - \frac{z^2}{2\sigma^2}(m_1^j)^2 - \frac{1}{2}(m_2^j)^2)
- J\frac{z^2}{\sigma^2}(m_1^j m_1^{j+1} + m_1^{j-1} m_1^j)(yy) - (m_2^j m_2^{j+1} + m_2^{j-1} m_2^j)(xx)
+ \frac{zx}{2\sigma^2}m_1^j m_1^{j+1}(zx) + \frac{zx}{2\sigma^2}m_1^{j-1} m_1^j(zx)
+ \frac{zx}{2\sigma^2}m_1^{j-1} m_1^{j+1}(zx) + J\frac{zx}{2\sigma^2}m_1^j m_1^{j+1}(zx)
\]

(A.6)

The Fourier series for \( \mathbf{m}^j \) is:

\[
\mathbf{m}^j = \sqrt{\frac{1}{N}} \sum_q \mathbf{m}^q e^{iqj}
\]

(A.7)

Using this Fourier series, the Hamiltonian becomes:

\[
H = NH_{mf} + \sum_q (m_1^q m_1^{-q} + \frac{1}{2}h_x + \frac{z^2}{2\sigma^2}(zz) + \frac{zx}{\sigma^2}(zx) - \frac{z^2}{\sigma^2}(yy) \cos(q))
+ (m_2^q m_2^{-q} + \frac{1}{2}h_x + (zz) - (xx) \cos(q))
= NH_{mf} + \sum_q \begin{pmatrix} m_1^q \\ m_2^q \end{pmatrix}^T \begin{pmatrix} M_1 & 0 \\ 0 & M_2 \end{pmatrix} \begin{pmatrix} m_1^{-q} \\ m_2^{-q} \end{pmatrix}
\]

(A.8)

Where \( M_1 = \frac{1}{2}(h(\sigma_x) + 2\frac{\langle \sigma_x \rangle}{\sigma_x} \langle \sigma_x \sigma_x \rangle + 2\frac{\langle \sigma_x \rangle}{\sigma_x} \langle \sigma_x \rangle - 2\frac{\langle \sigma_x \rangle^2}{\sigma_x^2} \langle \sigma_x \sigma_y \rangle \cos(q)) \), \( M_2 = \frac{1}{2}(h(\sigma_x) + 2\langle \sigma_x \sigma_x \rangle - 2\langle \sigma_x \sigma_x \rangle \cos(q)) \) and \( \sigma^2 = \langle \sigma_x \rangle^2 + \langle \sigma_x \rangle^2 \). We transform the spin
waves to Bose coherent state variables as follows:

$$
\begin{pmatrix}
z_i \\
z_i^* 
\end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \end{pmatrix}.
$$

(A.9)

We then identify $z_i$ with the bosonic operator $b_i$, and $z_i^*$ with $b_i^\dagger$. Expressing the Hamiltonian in terms of these variables, we obtain a second quantized Bose Hamiltonian. This Hamiltonian is:

$$
\mathcal{H} = NHmf + \frac{1}{2} \sum_q \begin{pmatrix} b_q^\dagger \\ b_{-q} \end{pmatrix}^T \begin{pmatrix} M_1 + M_2 & M_1 - M_2 \\ M_1 - M_2 & M_1 + M_2 \end{pmatrix} \begin{pmatrix} b_q \\ b_{-q}^\dagger \end{pmatrix}
$$

(A.10)

We diagonalize this with a Bogoliubov transformation:

$$
b_q = u_q a_k - v_q a_{-q}^\dagger,
$$

(A.11)

where $u_q^2 - v_q^2 = 1$. For the Hamiltonian to be diagonal

$$
u_q^2 = \frac{1}{2} + \frac{M_1 + M_2}{4\sqrt{M_1M_2}}
$$

(A.12)

and

$$
v_q^2 = \frac{1}{2} + \frac{M_1 + M_2}{4\sqrt{M_1M_2}}.
$$

(A.13)

The Hamiltonian is therefore:

$$
\mathcal{H} = NHmf + \frac{1}{2} \sum_q \begin{pmatrix} a_q^\dagger \\ a_{-q} \end{pmatrix}^T \begin{pmatrix} 2\sqrt{M_1M_2} & 0 \\ 0 & 2\sqrt{M_1M_2} \end{pmatrix} \begin{pmatrix} a_q \\ a_{-q}^\dagger \end{pmatrix}
$$

(A.14)

$$
= NHmf + \sqrt{M_1M_2}(a_q^\dagger a_q + a_{-q}a_{-q}^\dagger) = NHmf + 2\sqrt{M_1M_2}(a_q^\dagger a_q + \frac{1}{2})
$$
Using this result we find that the quantum zero-point energy is:

\[ E_{ZP} = \sqrt{M_1 M_2} - \frac{1}{2} (M_1 + M_2) = -\frac{1}{2} (\sqrt{M_1} - \sqrt{M_2})^2 \]  

(A.15)
Appendix B

Extracting Lyapunov Exponents

Here we provide some additional details of how to extract Lyapunov spectra from linearised equations of motion, Eq. (5.2), describing the evolution of the displacement between neighbouring trajectories \( X(t) \) and \( X(t) + dX(t) \). The asymptotic rate at which these two trajectories diverge (or converge) is characterized with a Lyapunov exponent. If the solution for this equation is \( dX(t) = Y(X,t)dX(0) \) then the Lyapunov exponent associated with these trajectories is

\[
\lambda = \lim_{t \to \infty} \frac{1}{t} \log \left| \frac{dX(t)}{dX(0)} \right| = \lim_{t \to \infty} \frac{1}{t} \log (Y(X,t)dX(0)). \tag{B.1}
\]

For almost all trajectories \( X(t) \) and almost all tangent vectors \( dX(t) \) the limit in Eq. (B.1) converges to the largest Lyapunov exponent of the system \([131,132]\).

Using a similar approach it is possible to calculate the entire Lyapunov spectrum. Instead of a single trajectory, consider a \( d \)-dimensional parallelepiped defined by \( d \) vectors tangent to the manifold at point \( X(t), U(t) = \{dX^1(t), dX^2(t), ..., dX^d(t)\} \). The volume of the parallelepiped will evolve over time in a manner determined by the \( d \) Lyapunov exponents

\[
\sum_{i=1}^{d} \lambda_i = \lim_{t \to \infty} \frac{1}{t} \log (\text{Vol}^d(Y(X,t)U(0))). \tag{B.2}
\]
Unfortunately, the Lyapunov spectrum cannot be easily extracted using this method. As \( t \to \infty \) the different tangent vectors comprising the parallelepiped all begin to point in the direction of the largest Lyapunov exponent. Many methods have been introduced to circumvent this issue. We use an algorithm introduced by Bennetin et al. \[46\].

An orthonormal basis for the tangent space \( \mathbf{V}(t) = \{d\hat{\mathbf{X}}^1(t), d\hat{\mathbf{X}}^2(t), ..., d\hat{\mathbf{X}}^d(t)\} \) is defined and then evolved for a short time:

\[
\mathbf{U}(t + \delta t) = Y(X, t)\mathbf{V}(t). \tag{B.3}
\]

This evolution rotates and changes the length of each of the unit vectors in \( \mathbf{V}(t) \). By performing a QR decomposition on \( \mathbf{U}(t + \delta t) \) we can separate these two effects:

\[
\mathbf{U}(t + \delta t) = \mathbf{Q}(t + \delta t)\mathbf{R}(t + \delta t). \quad \mathbf{V}(t + \delta t) \equiv \mathbf{Q}(t + \delta t)
\]

is a new orthonormal basis for the tangent space, obtained by rotating the basis vectors from the previous time step. Since \( \det[\mathbf{U}(t + \delta t)] = \prod_i R_{ii} \) the diagonal elements of \( \mathbf{R}(t + \delta t) \) capture the extent to which the volume of the parallelepiped at the previous time step has changed.

This process is repeated iteratively to obtain a sequence of matrices \( \mathbf{R}(t) \) from which we may extract the Lyapunov spectrum using

\[
\lambda_i = \lim_{N \to \infty} \frac{1}{N\delta t} \sum_{n=1}^{N} \log |R_{ii}(n\delta t)|. \tag{B.4}
\]
Appendix C

Implementing the TDVP

Here we provide details of the time-dependent variational principle used to generate our numerical results. Our implementation closely follows that of Haegeman et al. [12] and we refer to the original papers for further details. Here we give a brief summary noting in particular aspects that require modification for the thermofield MPS.

C.1 Matrix product state TDVP

A variational wavefunction $|\psi(A)\rangle$ defined by a matrix product state $A^\sigma_{ij}$ evolves on the manifold of matrix product states according to Eq.(5.3) with the appropriate identification of variables and indices: $X \rightarrow A$, $i \rightarrow I \equiv \{i,j,\sigma\}$ giving

$$\langle \partial \bar{A}^I|\partial A^J|\psi\rangle = \frac{i}{\hbar} \langle \partial \bar{A}^I|\hat{H}|\psi\rangle.$$  \hspace{1cm} (C.1)

Determining the time evolution of $|\psi(A)\rangle$ from equation Eq.(C.1) requires inversion of the Gram matrix $\langle \partial \bar{A}^I|\partial A^J|\psi\rangle$. In the case of matrix product states this is a $dD^2 \times dD^2$ matrix, however not all of the $dD^2$ tangent vectors are linearly independent so the Gram matrix cannot be inverted. As noted in Ref. [12], this can be resolved by imposing a gauge fixing condition on the states $|\partial A^J|\psi(A)\rangle$ parameterizing the tangent space. We
follow Ref. [12] and use the left tangent gauge fixing condition, \( \sum_{\sigma=1}^{d} A_{ij}^{\sigma \dagger} l_{jk} d A_{kl}^{\sigma} = 0 \), where \( l \) is the left environment, \( i.e. \) the result of contracting the MPS state with its conjugate on every site to the left of a given site and \( dA \) is an update to the MPS tensor such that \( A \rightarrow A + dA \). This gauge condition can be achieved by constructing \( L_{i,(\sigma_j)} = [A^{\sigma l}]_{ij} \) and calculating its null vectors, \([V_L]_{(i\sigma),j}\). If the null space is reshaped to \([V_L]_{ij}^{\sigma}\) then a \( dA \) that satisfies the tangent gauge condition can be written as

\[
dA^\sigma(x) = l^{-\frac{1}{2}} V_L^\sigma X r^{-\frac{1}{2}}, \tag{C.2}
\]

where \( r \) is the right environment. Using this parameterizing the Gram matrix becomes diagonal and the time evolution of the state can determined by evaluating Eq.\((5.3)\) to find the \((d^2 - 1)D \times D\) matrix \( X \).

### C.2 Inverse-free algorithm

While this algorithm is sufficient to determine the time evolution of a matrix product state at fixed bond dimension it has two flaws. Firstly, it necessarily involves inverting Schmidt coefficients and therefore encounters issues when a state has small Schmidt values. Secondly, there is no easy way to increase the bond dimension of the matrix product state as we may need to if we start from say a product state initial state. Both of these issues can be solved by using an inverse-free version of the TDVP algorithm [133]. Here we provide minor modifications to this algorithm required to study real-time evolution rather than imaginary-time evolution as studied in [133].

An inverse-free algorithm uses \( A \) in both left and right canonical forms, \( A_L \) and \( A_R \) respectively. For \( A_L \) the dominant left eigenvector of the transfer matrix is \( l = I \) and dominant right eigenvector is \( r = cc^\dagger \). For \( A_R \) the dominant left eigenvector of the transfer matrix is \( l = c^\dagger c \) and dominant right eigenvector is \( r = I \). The algorithm has three key steps:
i. $A_R$ and $c$ can be calculated from $A_L$ in an inverse-free method by iterating

$$[c_{(i+1)}, A_{R(i+1)}] = RQ(A_Lc_i) \quad (C.3)$$

until $c_{i+1} \approx c_i$ where $RQ(M)$ is an RQ decomposition.

ii. An inverse-free update of $A_L(t)$ is found by solving

$$\min_{\tilde{A}_L} |\tilde{A}_Lc(t + \delta t) - A_C(t + \delta t)| \quad (C.4)$$

where we have defined $A_C = A_Lc = cA_R$, with

$$A_C(t + \delta t) = A_L(t)c(t) + \delta t \, d(A_Lc)/dt$$

and $c(t + \delta t) = c(t) + \delta t \, dc/dt$. The time derivative of $A_L$ is obtained from Eq.(C.1) and that of $c$ from

$$\left(1 - \sum_{\sigma=1}^d A^\sigma_L \otimes \bar{A}^\sigma_R\right) \frac{dc}{dt} = \sum_{\sigma=1}^d A^\sigma_L \frac{dc}{dt} A^\sigma_R. \quad (C.5)$$

Eq.(C.5) follows from writing $\frac{dc}{dt} = \frac{d}{dt} (\sum_{\sigma=1}^d A^\sigma_L c \bar{A}^\sigma_R)$ and using the right gauge fixing condition on $A_R$ to impose \(\sum_{\sigma=1}^d A^\sigma_L c_{ij} dA^\sigma_{R,kl} = 0\). iv. Eq.(C.4) can be solved performing qr decompositions on $c(t + \delta t)$ and $A_C(t + \delta t)_{(\sigma),j}$, $c(t + \delta t) = qr$ and $A_C(t + \delta t)_{(\sigma),j} = QR$. We find $r = R$ so $A_L(t + \delta t)_{(\sigma),j} = Qq^j$ and $A_L(t + \delta t)_{ij}$ can be found by reshaping this matrix.

### C.3 Thermofield Double

The main modifications that we require to the standard MPS machinery is in its application to the thermofield double and its time evolution. We parametrize the thermofield double state $|\psi(\mathcal{H})\rangle$ by an expanded matrix product state $M^\delta_{ij}$ with a doubled physical index representing the two copies of the system. The thermofield double state is evolved using the expanded Hamiltonian, $\mathcal{H} = \mathcal{H} \otimes 1 + 1 \otimes \mathcal{H}$. The time-
dependent variational principle Eq.(5.3) is modified accordingly with $A \rightarrow A_A$, $H \rightarrow H_H$ and $\{i, j, \sigma\} \rightarrow \{I, J, \sigma, \delta\}$.

The thermofield double is evidently symmetrical between the two copies of the physical space; observations made on either copy will yield the same result. However, this is not necessarily reflected in an explicit symmetry of the tensor $A_I^\sigma_j^\delta$. We therefore impose the symmetry $A_I^\sigma_j^\delta = A_{\tilde{I}}^{\delta\sigma}_{\tilde{J}}$ on our state, where $I \equiv i \otimes i'$ and $\tilde{I} \equiv i' \otimes i$, explicitly using an additional tangent space gauge fixing. This prevents a key source of errors: thermofield doubles form a subspace of the doubled Hilbert space. Strictly the dynamics under $H_H$ remains in this subspace, but numerical errors can take the dynamics away from it.

For calculating our tangent state we find it more convenient to work in the a slightly different gauge in which the symmetry condition is $A_I^\sigma_j^\delta = M_{IK} A_{KL}^\delta_j^\sigma$, where

$$M = \begin{pmatrix} \frac{I}{2} D(D+1) & 0 \\ 0 & -\frac{I}{2} D(D-1) \end{pmatrix}$$

(C.6)

The tangent gauge fixing is then achieved as follows: We first calculate $V_{L,I,j}^\sigma_j^\delta$ using the method described in section C.1. Symmetric ($\frac{1}{2} V_{L,(I,J)}^\sigma_j^\delta + M_{LL'} \frac{1}{2} V_{L',(I,J)}^\delta_j^\sigma$) and anti-symmetric ($\frac{1}{2} V_{L,(I,J)}^\sigma_j^\delta - M_{LL'} \frac{1}{2} V_{L',(I,J)}^\delta_j^\sigma$) parts of $V$ contribute separately to $dA$ with corresponding symmetric and antisymmetric parts of the matrices $X$. The symmetrised and anti-symmetrised spaces are each smaller than the doubled space. A complete orthonormal basis for $V$ keeping the first and a complete orthogonal basis for each is obtained by keeping the first $(D^2 - 1)D(D-1)/2 - D$ or $(D^2 - 1)D(D-1)/2 + D$ (where $D = D^2$) columns of the $Q$ from a QR decomposition of the symmetrised or anti-symmetrised $V$ respectively. Full details of the implementation of this algorithm will be communicated elsewhere. This constraint also requires the modification of step iv. in the inverse-free algorithm. $A_L(t + \delta t)$ is calculated using QR decompositions on $c(t + \delta t)$ and $A_C(t + \delta t)$.
but the symmetry constraint requires $Q$ to be modified. A new $Q$ is obtained by performing a QR decomposition on the symmetrised $\frac{1}{2}Q_{((\sigma\delta)I),(\sigma\delta)'I')} + \frac{1}{2}MQ_{((\delta\sigma)\tilde{I}),(\delta\sigma)'\tilde{I}')}M^\dagger$ and keeping the first $D$ columns.
Appendix D

Extraction of Lyapunov Spectrum

In Appendix A we discussed how to calculate the Lyapunov spectrum of a trajectory in a dynamical system using vectors in its tangent space, in Appendix B we explained how time evolution of a quantum state can be determined using the time dependent variational principle, we will now explain how to extract the Lyapunov spectrum of a quantum system using these methods. As described in Section 5.4.2 we are interested in the evolution of the difference of two trajectories, i.e. the tangent vectors to the variational manifold whose equation of motion is given by linearizing Eq.(C.1) using the parametrization in terms of $X$ given by Eq.(C.2):

$$d\dot{X}_a(t) = i\langle \partial_{X_a} \partial_{\bar{X}_b} \psi | \hat{H} | \psi \rangle dX_b(t) + i\langle \partial_{X_a} \psi | \hat{H} | \partial_{\bar{X}_b} \psi \rangle d\bar{X}_b(t).$$

(D.1)

Our notation indicates a reshaping of the $(d^2 - 1)D \times D$ matrix $X$ into a complex $(d^2 - 1)D^2$ vector. In the case of thermofield double states, $X$ is complex $(d(d+1)/2 - 1)D^2$-dimensional vector. The right hand side of Eq.(D.1) contains two parts: $F_1 = \langle \partial_{X_a} \psi | \hat{H} | \partial_{X_b} \psi \rangle$ is manifestly Hermitian and generates unitary rotations of the tangent vectors. $F_2 = \langle \partial_{X_a} \partial_{\bar{X}_b} \psi | \hat{H} | \psi \rangle$ is not Hermitian. Instead it is a symmetric matrix
\(F_2 = F_2^T\) and is responsible for the non-unitary evolution of tangent vectors.

The Lyapunov spectrum is calculated by measuring the extent to which a tangent vector \(dX(t)\) has changed in magnitude between a time \(t\) and \(t + \delta t\). Eq.\((D.1)\) describes how the components \(dX_a\) transform but does not account for the transformation of the tangent space basis. This may be captured by allowing for parallel transport along the trajectory. Taking into account the parallel transport, Eq.\((D.1)\) can now be written as

\[
d\dot{X}(t) = \tilde{F}_1 dX(t) + \tilde{F}_2 d\bar{X}(t),
\]

where \(F_1\) and \(F_2\) have been modified as follows:

\[
F_1 \rightarrow \tilde{F}_1 = \langle \partial_{\bar{X}_b} \psi | \hat{H} | \partial X_a \psi \rangle - \Gamma_{a,b}^c \dot{X}_c
\]

\[
F_2 \rightarrow \tilde{F}_2 = \langle \partial_{\bar{X}_a} \partial_{\bar{X}_b} \psi | \hat{H} | \psi \rangle - \bar{\Gamma}_{a,b}^c \dot{X}_c,
\]

with \(\bar{\Gamma}_{a,b,c} = \langle \partial_{\bar{X}_b} \partial_{\bar{X}_c} \psi | \partial X_a \psi \rangle\) and \(\Gamma_{a,b,c} = \langle \partial_{\bar{X}_b} \psi | \partial X_c \partial X_a \psi \rangle\) the Christoffel symbols for the variational manifold. With this modification we can calculate the Lyapunov spectrum. We separate the real and imaginary components of \(dX = dX^R + idX^I\), \(\tilde{F}_1 = \tilde{F}_1^R + i\tilde{F}_1^I\) and \(\tilde{F}_2 = \tilde{F}_2^R + i\tilde{F}_2^I\). The real vector space is \(2(d-1)D^2\) dimensional for matrix product states and \(\frac{(d(d+1)\times2-2)}{2}D^2\) for thermofield double states. Eq.\((D.1)\) can be rewritten as:

\[
\begin{pmatrix}
  dX^R \\
  dX^I
\end{pmatrix} =
\begin{pmatrix}
  \tilde{F}_1^I + \tilde{F}_2^I & \tilde{F}_1^R - \tilde{F}_2^R \\
  -\tilde{F}_1^R - \tilde{F}_2^R & \tilde{F}_1^I - \tilde{F}_2^I
\end{pmatrix}
\begin{pmatrix}
  dX^R \\
  dX^I
\end{pmatrix}
\]

If \(\tilde{F}_2 = 0\) the Hermitian property of \(\tilde{F}_1\) would result in a totally antisymmetric matrix in Eq.\((D.5)\), generating purely orthogonal rotations on the tangent vectors. \(\tilde{F}_2\) is responsible for the changing magnitude of a tangent vector upon moving along a trajectory, and therefore for the generation of a non-zero Lyapunov spectrum. One important example
in which $F_2 = 0$ is local Hamiltonians $H = \sum_i h_i$. In this case, the parallel transport term cancels with $F_2$, guaranteeing that the Lyapunov spectrum is zero for every possible state.

Having accounted for these details, the Lyapunov spectrum of the system can be calculated using Eq. (D.5) and the methods in Appendix A.
Appendix E

Dynamics of coupled bosonic chains

In this appendix we will calculate observables for a system of two coupled Luttinger liquids. We will begin by calculating mode occupations using unitary perturbation theory and a semiclassical approximation in section E.1. We will then calculate the correlation function $C[x, t]$ defined in equation 6.28.

E.1 Mode occupations

E.1.1 Unitary perturbation theory

We will begin by calculating the mode occupation following the method described in Ref. [134]. This requires our initial state to be an eigenstate of the original (unperturbed) Hamiltonian, $H_0$, and the operator we wish to measure must commute with all constants of motion of $H_0$, hence our choice of $n_{k,i}$ as the quantity to calculate. At lowest order we have [134]

$$\langle n_{k,i}(t) \rangle = \langle n_{k,i}(t) \rangle_{t=0} + 4 \int_{-\infty}^{\infty} d\omega J(\omega) \frac{\sin^2 \omega t/2}{\omega^2}, \quad (E.1)$$

$$J(\omega) = \sum_\ell |\langle I | H_{\text{int}} | \ell \rangle|^2 \langle \ell | n_{k,i} | \ell \rangle \delta(\omega - (E_\ell - E_I)). \quad (E.2)$$
Here $\mathcal{I}$ is the quench initial state, which is a tensor product of the two chains in their ground states,

$$|\mathcal{I}\rangle = |0\rangle_1 \otimes |0\rangle_2$$  \hspace{1cm} (E.3)

while $\ell$ is a more general product state of chain eigenstates. This state takes the form

$$|\ell\rangle = |u\rangle_1 \otimes |v\rangle_2$$

where $|u\rangle_i$ and $|v\rangle_i$ are (non-vacuum) eigenstates of the $i$th chain Hamiltonian. In a system of two coupled chains, labelled by $i$ and $\bar{i}$, we have

$$\langle n_{k,i}(t) \rangle = 4t_\perp^2 R^2 \sum_{u,v} \left\{ |\langle 0|_i \varphi_i^1(0)|u\rangle_i|^2 |\langle 0|_i \varphi_i(0)|v\rangle_i|^2 \left[ \langle v|_i n_{k,i}|v\rangle_i + \langle u|_i n_{k,i}|u\rangle_i \right] \right\}$$

$$\times \frac{\sin^2 \left( t(E_u + E_v - 2E_0)/2 \right)}{(E_u + E_v - 2E_0)^2}, \hspace{1cm} i = 1, 2.$$

(E.4)

The small time expansion of this quantity is

$$\langle n_{k,i}(t) \rangle = t^2 t_\perp^2 R^2 \sum_{u,v} \left\{ |\langle 0|_i \varphi_i^1(0)|u\rangle_i|^2 |\langle 0|_i \varphi_i(0)|v\rangle_i|^2 \left[ \langle v|_i n_{k,i}|v\rangle_i + \langle u|_i n_{k,i}|u\rangle_i \right] \right\}.$$

(E.5)

Let us evaluate this expression for small $\beta$. At lowest order in $\beta$, and remembering the constraint that total momentum vanishes, the states $u, v$ (with appropriate normaliza-
tion) that contribute are:

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>k\rangle^{-1} a_{-k} \bar{a}_{-k}</td>
<td>{-1}\rangle$</td>
<td>$</td>
</tr>
<tr>
<td>$</td>
<td>{-1}\rangle$</td>
<td>$</td>
<td>k\rangle^{-1} a_{-k} \bar{a}_{-k}</td>
</tr>
<tr>
<td>$</td>
<td>k\rangle^{-1/2} a_{-k}</td>
<td>{-1}\rangle$</td>
<td>$</td>
</tr>
<tr>
<td>$</td>
<td>k\rangle^{-1/2} \bar{a}_{-k}</td>
<td>{-1}\rangle$</td>
<td>$</td>
</tr>
</tbody>
</table>

Therefore our approximation to the mode occupation is

$$\langle n_{k,i}(t) \rangle \approx \frac{1}{k^2} 4R^2 \beta^4 \left(\frac{2\pi}{R}\right)^{4\beta^2} t^2 t^2. \quad (E.7)$$

An extra factor of $\beta^2$ suppresses the involvement of other states.

E.1.2 Semiclassical calculation

We now wish to calculate the mode occupation within the semiclassical approximation.

The semiclassical Hamiltonian is diagonalised by the Bogoliubov transformation

$$b_p^\dagger = c_\phi \gamma_p^\dagger + s_\phi \gamma_{-p}, \quad (E.8)$$

with $c_\phi = \cosh \phi_p$, $s_\phi = \sinh \phi_p$ and

$$\phi_p = \frac{-t_{\perp}}{t_{\perp} + Kvp^2}. \quad (E.9)$$
Firstly we consider

$$\langle b_p^+(t)b_p(t) \rangle = \langle I | e^{iH_{hc}t}b_p^+e^{-iH_{hc}t}e^{iH_{hc}t}b_p e^{-iH_{hc}t} | I \rangle$$

$$= \langle I | e^{iH_{hc}t}(c_\phi^\dagger \gamma_p^\dagger + s_\phi \gamma_{-p}) e^{-iH_{hc}t}e^{iH_{hc}t}(c_\phi \gamma_p^\dagger + s_\phi \gamma_{-p}) e^{-iH_{hc}t} | I \rangle$$

$$= \langle I | e^{iH_{hc}t}(c_\phi^\dagger \gamma_p^\dagger + s_\phi \gamma_{-p}) e^{-iH_{hc}t}e^{iH_{hc}t}(c_\phi \gamma_p^\dagger + s_\phi \gamma_{-p}) e^{-iH_{hc}t} | I \rangle$$

$$= \langle I | (e^{i\omega_p^t c_\phi \gamma_p^\dagger + e^{-i\omega_p^t s_\phi \gamma_{-p}})(e^{-i\omega_p^t c_\phi \gamma_p^\dagger + e^{i\omega_p^t s_\phi \gamma_{-p}}}) | I \rangle, \quad (E.10)$$

where $\omega_p = \sqrt{p^2 + 16\pi\beta^2 t_\perp}$. Performing the reverse Bogoliubov transform, $\gamma_p^\dagger = c_\phi b_k^\dagger + s_\phi b_{-k}$ yields

$$\langle b_p^+(t)b_p(t) \rangle = \langle I | (2 - e^{2i\omega_p^t} - e^{-2i\omega_p^t}) (\sinh \phi_p \cosh \phi_p)^2 b_{-p} b_{-p}^\dagger | I \rangle$$

$$= 2(1 - \cos 2\omega_p^t) (\sinh \phi_p \cosh \phi_p)^2$$

$$= 4 \sin^2 \omega_p^t \sinh \phi_p \cosh \phi_p$$

$$= \sin^2 \omega_p^t \sinh^2 2\phi_p. \quad (E.11)$$

We now use the following expression for the transformation parameter

$$\tanh 2\phi_p = \frac{1}{1 + \frac{|p|^2}{8\pi\beta^2 t_\perp}}, \quad (E.12)$$

from which we find

$$\sinh^2 2\phi_p = \left( \frac{8\pi\beta^2 t_\perp}{|p|^2} \right)^2 \frac{1}{1 + \frac{16\pi\beta^2 t_\perp}{|p|^2}}. \quad (E.13)$$

Therefore

$$\langle b_p^+(t)b_p(t) \rangle = \sin^2 \omega_p^t \left( \frac{8\pi\beta^2 t_\perp}{|p|^2} \right)^2 \frac{1}{1 + \frac{16\pi\beta^2 t_\perp}{|p|^2}}, \quad (E.14)$$

135
and the occupation of mode $k$ (with $2\pi k/R = p$) on one of the chains is

$$
\langle n_{k,1}(t) \rangle = \frac{1}{2} \sin^2 \omega_p t \left( \frac{8\pi \beta^2 t_{\perp}}{|p|^2} \right)^2 \frac{1}{1 + \frac{16\pi \beta^2 t_{\perp}}{|p|^2}}.
$$

(E.15)

Expanding for short times we find

$$
\langle n_{k,1}(t) \rangle \approx \frac{1}{k^2} 8R^2 \beta^4 t^2 t_{\perp}^2, \quad t \ll \omega_p^{-1}.
$$

(E.16)

### E.2 Correlation function

We will now calculate the correlation function

$$
C[x,t] = \langle e^{i(\theta_A(x,t) - \theta_A(0,t))} \rangle
$$

(E.17)

$$
= e^{-\frac{1}{2} \langle (\theta_A(x,t) - \theta_A(0,t))^2 \rangle}
$$

(E.18)

in the semiclassical Hamiltonian using two different initial states. We will determine this correlation function by calculating

$$
\langle \theta_A(x,t) \theta_A(y,t) \rangle
$$

(E.19)

The mode expansion of $\theta_A$ can be calculated using equation [6.11]

$$
\theta_A = \frac{1}{\sqrt{R}} \theta_0 + i\beta \sum_{q \in \mathbb{Z}_{\neq 0}} \sqrt{\frac{2}{|q|}} e^{2\pi i q x/R} (b_{-q}^\dagger - b_q).
$$

(E.20)
E.2.1 Ground state of uncoupled chains

We will initially calculate \( \langle \theta_A(x,t)\theta_A(y,t) \rangle \) for a system of initially uncoupled chains. \( \theta_A(x) \) can be written in terms of the Bogoliubov bosons as

\[
\theta_A(x) = \frac{1}{\sqrt{R}} \theta_0 + i\beta \sum_{q \in \mathbb{Z}, q \neq 0} \sqrt{\frac{2}{|q|}} e^{2\pi i q x/R} e^{\phi_p (\gamma^\dagger_q - \gamma_q)}. \tag{E.21}
\]

In the Heisenberg picture we find that for the semiclassical Hamiltonian \( \theta_A(x,t) \) is

\[
\theta_A(x,t) = \frac{1}{\sqrt{R}} \theta_0 - i\beta \sum_{q \in \mathbb{Z}, q \neq 0} \sqrt{\frac{2}{|q|}} e^{2\pi i q x/R} e^{\phi_p (e^{i\omega_q t} \gamma^\dagger_{-q} + e^{-i\omega_q t} \gamma_q)}. \tag{E.22}
\]

Noting that this initial state for the system is the vacuum for the \( b_p \) bosons, we rewrite \( \theta_A(x,t) \) in terms of \( b_p \) and \( b_p^\dagger \).

\[
\theta_A(x,t) = \frac{1}{\sqrt{R}} \theta_0 - i\beta \sum_{q \in \mathbb{Z}, q \neq 0} \sqrt{\frac{2}{|q|}} e^{2\pi i q x/R} e^{\phi_p}
\times (e^{i\omega_q t} (\cosh \phi_q b_{-q}^\dagger - \sinh \phi_q b_q) - e^{-i\omega_q t} (\sinh \phi_q b_q^\dagger - \cosh \phi_q b_{-q})). \tag{E.23}
\]

The correlator \( \langle \theta_A(x,t)\theta_A(y,t) \rangle \) for the initial state \( |\psi(0)\rangle = |0,0\rangle \) can therefore be written as:

\[
\langle \theta_A(x,t)\theta_A(y,t) \rangle = \frac{1}{R} \theta_0^2 + \beta^2 \sum_{q \in \mathbb{Z}, q \neq 0} \frac{4}{|q|} \cos \left( \frac{2\pi q (x-y)}{R} \right) e^{2\phi_p} (\cosh 2\phi_q + \cos 2\omega_q t \sinh 2\phi_q). \tag{E.25}
\]

E.2.2 Squeezed state

We will now repeat this calculation for initial squeezed state:

\[
|\psi(0)\rangle = \exp \left[ \sum_q W_q b_q^\dagger b_{-q}^\dagger \right] |0,0\rangle \tag{E.26}
\]
where

\[ W_p = \frac{1}{2} \frac{\pi \rho}{\pi \rho + |q| K}. \]  

(E.27)

This initial state is annihilated by a new boson operator,

\[ \alpha_q = \frac{1}{\sqrt{1 - (2 W_q)^2}} b_q - \frac{2 W_q}{\sqrt{1 - (2 W_q)^2}} b_{-q}^\dagger = A_q b_q - B_q b_{-q}^\dagger. \]  

(E.28)

With this initial state the \( \langle \theta_A(x,t) \theta_A(y,t) \rangle \) correlator becomes:

\[
\langle \theta_A(x,t) \theta_A(y,t) \rangle = \frac{1}{R} \theta_0^2 + \beta^2 \sum_{q \in \mathbb{Z} \geq 0} \frac{4}{|q|} \cos \left( \frac{2 \pi i q (x - y)}{R} \right) e^{2 \phi_q} \cos (2 \omega_q t - 2 A_q B_q \cosh 2 \phi_q) \]  

(E.29)

\[
\times ((A_q^2 + B_q^2) \cosh 2 \phi_q - 2 A_q B_q \sinh 2 \phi_q + \cos 2 \omega_q t ((A_q^2 + B_q^2) \sinh 2 \phi_q - 2 A_q B_q \cosh 2 \phi_q)). \]  

(E.30)

This reduces to equation \( \text{E.25} \) when \( W_q \to 0 \) which causes \( A_q \to 1 \) and \( B_q \to 0 \). Using these different initial states results in large differences between for \( C[x, t] \).
Appendix F

Calculating $M_k(\mathcal{H}, A_0)$

Here we note all of the tensor diagrams required for the calculation of $M_k(\mathcal{H}, A_0)$, defined as

$$
\sum_k \left( \begin{array}{c} x_k^* \\ x_{-k}^* \end{array} \right) ^T \left( \begin{array}{cc} \epsilon_k(\mathcal{H}, A_0) & \Delta_k(\mathcal{H}, A_0) \\ \Delta_k(\mathcal{H}, A_0)^T & \epsilon_k^T(\mathcal{H}, A_0) \end{array} \right) \left( \begin{array}{c} x_k \\ x_{-k}^* \end{array} \right)
$$

where $\epsilon_k(\mathcal{H}, A_0) = \langle \partial^{A^\sigma_k} \psi | \mathcal{H} | \partial^{A^\sigma_k} \psi \rangle$ and $\Delta_k(\mathcal{H}, A_0) = \langle \partial^{A^\sigma_k} \partial^{A^\sigma_{-k}} \psi | \mathcal{H} | \psi \rangle$.

As we have noted, the orthogonality of the tangent states $dA$ to $A_0$ simplify these calculations significantly:

$$
\sum_{\sigma} dA^\sigma A_0^\sigma = 0.
$$

Additionally, the tensor diagrams involve summing powers of the transfer matrix as defined in equation 2.14. Due to the orthogonality condition, $E^n$ can be projected onto its $\lambda < 1$ subspace, $E^n \rightarrow Q(QE)^nQ$. This can now be summed

$$
\sum_{n=0}^{\infty} Q(e^{ik}QE)^n Q = Q(I - e^{ik}QE)^{-1} Q = T_k.
$$
F.1 Standard terms of $\epsilon_k(\mathcal{H}, A_0)$

As previously discussed, the standard terms of $\epsilon_k(\mathcal{H}, A_0)$ are a sum over diagrams with a tangent MPS $dA$ at site $a$, a Hamiltonian at site $c$ and $c+1$ and a second tangent MPS $\overline{dA}$ at site $b$. Using the simplifications discussed in that section, these take the following form:
F.2 Normalization terms of $\epsilon_k(H, A_0)$

The quadratic order terms in $A^\sigma_i$, $A_N = -\frac{1}{2} A_0^\sigma \Gamma^{-\frac{1}{2}} x_i \Gamma^{-\frac{1}{2}}$ also contribute to $\epsilon_k(H, A_0)$. These are tensor diagrams with either $A_N$ at a site $a$ or $\bar{A}_N$ at a site $b$:
F.3 Tensor diagrams of terms of $\Delta_k(\mathcal{H}, A_0)$

We now state all the diagrams which contribute to $\Delta_k(\mathcal{H}, A_0)$. These take the following form: $\bar{d}A$ at site $a$, a Hamiltonian at $c$ and $c + 1$ and another $\bar{d}A$ at site $b$:
\[ e^{-ik} H T_0 T_\pi \]
Bibliography


[53] Henning Labuhn, Daniel Barredo, Sylvain Ravets, Sylvain De Léséleuc, Tommaso Macrì, Thierry Lahaye, and Antoine Browaeys. Tunable two-dimensional arrays of


155


