Novel Ways of Solving the Single Particle Wavefunction in Quantum Mechanics

Saraswat Bhattacharyya
Undergraduate Programme
Indian Institute of Science, Bangalore

Supervisor
Prof. Ambarish Ghosh, Prof. Vijay Shenoy

In fulfillment of the requirements for the degree of

Bachelor of Science (Research) in Physics

December 1, 2019
Acknowledgements

I would like to thank the Undergraduate Programme, IISc for giving me the opportunity to pursue my interests in Physics. IISc provided me with a platform to interact with and learn from the best researchers in India and abroad. I would also like to thank the Kishore Vaigyanik Protsahan Yojana (KVPY) fellowship for funding my time in IISc, and motivating me to do regular research projects in the summer. I would like to thank Prof. Ambarish Ghosh for his expert guidance on every step of my project. I would also like to thank him for all the advice he has given me outside of Physics, and for all the times he has made time for me, even at odd hours, so I could resolve all my doubts. I would also like to thank Prof. Vijay Shenoy for agreeing to work with me on a topic he had never worked on before. I am grateful for the way he helped us frame the pertinent questions- and the way he explained basic things in a lucid and patient manner. I would like to thank him for resolving all my doubts about the project, for giving me the unique opportunity to help set up a computational cluster, and for hopefully teaching me better how to think like a physicist. I would like to thank Arjun Dey for working along with me on many aspects of applying machine learning techniques to quantum mechanics, and helping me out whenever I ran into any issues. I am grateful to Prof. Chandan Dasgupta and Prof. Sumilan Banerjee for giving me excellent advice about life and physics, and for being available whenever I needed them. I would like to thank all my friends who made my time at IISc special. I thank them for stimulating discussions about Physics and other topics, and for their constant companionship and support during all times good and bad. Finally, this thesis would not be possible without the constant moral support from my family. They provided me the opportunities and experience to make me who I am, and motivated me to continue working every time I felt I had hit a dead end.
Abstract

The quantum wavefunction $\psi$ is one of the most fundamental properties of a quantum system. For a single particle, knowing the wavefunction is sufficient to describe all the observable properties of the system. This also makes the wavefunction a challenging quantity to calculate and describe. In this thesis, we discuss two new methods to calculate the wavefunction of a single particle system.

The first of these methods, the symplectic method developed here, enables one to calculate the wavefunction and energy for an arbitrary potential in three dimensions with axial symmetry. I have used this technique to tackle the problem of an electron inside Liquid Helium, which forms a cavity around itself (called an "electron bubble"). I have reproduced studies of an electron bubble in a potential well with infinite walls. The technique may be then used to generalize to finite walls with electron tunnelling, or to account for an additional effective potential due to other electrons in an electron bubbles.

The second method discussed here uses a machine learning framework called "restricted Boltzmann machine" (RBM) to represent a ground-state wavefunction. I study the physical significance of the structure of this RBM. This helps us gain insight in a completely new method of looking at the single particle wavefunction, and also develops deeper ideas about what sort of information can be used to characterize a quantum wavefunction.
# Contents

1 Introduction 1
   1.1 The Electron Bubble Problem . . . . . . . . . . . . . . . . . . . 1
   1.2 Machine Learning Quantum Mechanics . . . . . . . . . . . . . 2
   1.3 Overview of the Thesis . . . . . . . . . . . . . . . . . . . . . . . 3

2 Symplectic Method for an Axially Symmetric 3d Potential 5
   2.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5
   2.2 Calculations . . . . . . . . . . . . . . . . . . . . . . . . . . . . 7
   2.3 The $\partial_\phi \psi \neq 0$ Case . . . . . . . . . . . . . . . . . 11
   2.4 The Numerical Implementation . . . . . . . . . . . . . . . . . . 12

3 Symplectic Results 13
   3.1 Simple Harmonic Oscillator . . . . . . . . . . . . . . . . . . . . 13
   3.2 Infinite Spherical Potential Well . . . . . . . . . . . . . . . . . 14
   3.3 Finite Spherical Potential Well . . . . . . . . . . . . . . . . . . 16
   3.4 Infinite Well of Arbitrary Geometry . . . . . . . . . . . . . . . 17
   3.5 Future Scope . . . . . . . . . . . . . . . . . . . . . . . . . . . . 21

4 Quantum Mechanics using Restricted Boltzmann Machines 22
   4.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . 23
   4.2 Restricted Boltzmann Machines . . . . . . . . . . . . . . . . . . 24
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3</td>
<td>The Variational Method</td>
<td>27</td>
</tr>
<tr>
<td>4.4</td>
<td>Representing a single particle</td>
<td>28</td>
</tr>
<tr>
<td>5</td>
<td>Applying the RBM Formalism</td>
<td>31</td>
</tr>
<tr>
<td>5.1</td>
<td>Fitting the RBM trial function</td>
<td>31</td>
</tr>
<tr>
<td>5.2</td>
<td>Transverse Field Ising model</td>
<td>34</td>
</tr>
<tr>
<td>5.3</td>
<td>Tight Binding Model</td>
<td>38</td>
</tr>
<tr>
<td>5.3.1</td>
<td>The &quot;Weights&quot; W</td>
<td>43</td>
</tr>
<tr>
<td>5.3.2</td>
<td>The &quot;Biases&quot; A and B</td>
<td>45</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Adding on-site disorder</td>
<td>49</td>
</tr>
<tr>
<td>5.4</td>
<td>Discussion</td>
<td>51</td>
</tr>
<tr>
<td>6</td>
<td>Conclusions</td>
<td>53</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>57</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 The Electron Bubble Problem

When an electron is injected inside liquid Helium, it opens a cavity about itself free of Helium atoms. This cavity is called an electron bubble [1][2][3]. The reason for the cavity may be thought of as follows: when the electron is far away from the Helium atom, it sees the atom as a net charge-neutral entity and gets attracted by the polarization it induces. When the electron gets too close, however, the outer electrons repel the electron strongly. Thus, it settles to equilibrium at a certain distance from the neighbouring Helium atoms with a cavity.

When the electron bubble is illuminated by light, the electron is excited from the lowest spherically symmetric 1s state to the asymmetric 1p state. The electron wavefunction creates the charge density responsible for maintaining the shape of the bubble, and hence a change in the wavefunction must change the shape of the bubble. Thus, we are in an interesting scenario where the wavefunction depends on the boundary conditions, but the boundary conditions also depend on the wavefunction. One of the interesting questions is to ask what will be the final shape of the bubble when it is excited from its initially spherical shape.
1.2 Machine Learning Quantum Mechanics

If one could solve the energy eigenvalue and wavefunction for any shape efficiently, one can do a minimization over the set of shapes to find the optimal ground state. This was solved in papers from the Maris group [4][1], [2], where the electron is in a potential well of a certain shape with infinite walls. Under this approximation, the calculation could only be done up to a liquid pressure of 5 bars, since the bubble acquired a peanut shape and started pinching at the waist until it split into two different bubbles. Experimentally, however, the bubbles can be probed up to 25 bars. It is possible that some assumption breaks down for Humphrey’s calculations, and it is possible that quantum tunneling may be the culprit, as it would have a strong contribution near the limits of Humphrey’s calculations where the waist of the bubble gets a sharp concave pinch. It may also be the case that the bubble above 5 bars pressure absorbs light and becomes hydrodynamically unstable, in which case the calculations will be valid. Humphrey’s method also relies on the fact that the wavefunction must vanish outside the bubble, so it cannot generalize to cases like quantum tunneling, where it decays exponentially outside the bubble instead of vanishing. We find an alternate way of solving the Schrödinger equation, which matches with Humphrey’s results for infinite potential wells and also generalizes to cases with finite potential wells.

1.2 Machine Learning Quantum Mechanics

Doing quantum many body calculations is difficult because the size of the Hilbert space explodes exponentially with the number of particles. The two ways to deal with this are with finite probabilistic sampling (Quantum Monte Carlo), and restricting the wavefunction to lie in a smaller subspaces by physical approximations (Tensor Networks). The first one suffers from numerical issues like the sign problem, whereas for the tensor networks, the method does not generalize
1.3 Overview of the Thesis

well to higher dimensions and large-scale entanglement. A variational technique using highly expressive neural network trial functions may provide a new method of solving quantum many body problems in regimes which the above-mentioned techniques cannot access. A neural network is basically a parametric model that can represent a function mapping a multidimensional input space to a multidimensional output space. $f(x^{(1)}_{in}, x^{(2)}_{in}, ...) = (y^{(1)}_{out}, y^{(2)}_{out}, ...)$. This particular model is special because it can approximate any ”well-behaved” function with a finite number of neurons (the Universal Approximation Theorem [5]), and also because there exist efficient methods (like backpropagation and gradient descent) which make it easy to find the optimal parameter values to represent any given function. Thus, it makes sense to use such a description to represent a variational quantum trial function. This was done by Carleo and Troyer [6], who used a restricted Boltzmann machine (RBM) as a trial variational wavefunction and found that ”the method performed at least as well as state-of-the-art approaches.” The question we are interested in are: what is the physical meaning of the RBM as a trial wavefunction i.e. what do the parameters represent physically? This leads us deeper to ask what information is actually required to describe a system, and these questions are explored in the second part of this thesis.

1.3 Overview of the Thesis

- In Chapter 2, we develop the symplectic method for solving the Schrodinger equation in three dimensions, where the potential has axial symmetry.

- In Chapter 3, we verify the accuracy of the symplectic method by running it on standard quantum mechanical problems with analytical solutions. We also reproduce Humphrey’s results for the infinite wall bubble case, and explain why this method will also give reliable results for the finite wall
problem.

• In Chapter 4, we switch gears and look at other ways of solving for the wavefunction of a system. We study the restricted Boltzmann machine framework for describing the wavefunction, and try to understand why it works and what the parameters represent.

• In Chapter 5, we apply the RBM formalism to different physical Hamiltonians. We analyze the resulting solutions and try to estimate what the RBM parameters physically mean, and how many parameters can characterize a given system.
Chapter 2

Symplectic Method for an Axially Symmetric 3d Potential

Summary

Symplectic schemes have been developed to numerically solve the Schrodinger equation in two dimensions [7][8][9][10]. We generalize the scheme to axially symmetric three dimensional problems. We show that such a symplectic scheme can be developed in 3d polar coordinates with trivial φ dependence, and discuss how to account for φ dependence in the wavefunction when the potential is axially symmetric. We use the SciPy wrapper for ARPACK[11] to solve the resulting eigenvalue problem in an efficient way.

2.1 Introduction

One of the first methods we learnt to solve the Schrodinger equation in 1 dimension is called the shooting method [12]. This method relies on fixing the wavefunction for the first two points (since we need two numbers ψ(x = 0) and ψ′(x = 0)
2.1 Introduction

to find an unique solution to the second order Schrodinger equation), and then
finding what the resultant wavefunction at all the remaining points would be by
extrapolating using the Finite Difference method. This search would be repeated
over all energies (since the extrapolation depends on the energy), and the energy
eigenvalue would have the wavefunction at the other end of the system go to zero
(i.e. obey the required boundary condition).

If we think on the wavefunction at every (grid)point as a vector,

$$\Psi = \begin{bmatrix} \psi(x = 0) & \psi(x = \Delta x) & \psi(x = 2\Delta x) & \ldots & \psi(x = N\Delta x) \end{bmatrix}^T$$

then we can write the Schrodinger equation as a matrix equation in $\Psi$ as the
differentiation operators just become matrices. If we define $\Phi$ to be

$$\Phi = \begin{bmatrix} \psi(x = \Delta x) & \psi(x = 2\Delta x) & \ldots & \psi(x = (N - 1)\Delta x) \end{bmatrix}^T$$

i.e. endpoints excluded, we can in fact implement the boundary conditions requiring $\psi(0) = \psi(N) = 0$, because the matrices representing the differential operators enforce continuity.

Now, if we want to generalize this method to higher dimensions, we need
a method to "wrap" the higher dimensions into a vector so we can utilise this
matrix form again. This was done for two dimensions by Liu et.al. [7], using
what they called the symplectic scheme. Under this generalization, they used a
$N_x \times N_y$ vector for the wavefunction defined as:

$$\Phi = \begin{bmatrix} \psi(x = -N\Delta x, -N\Delta y) & \psi(x = -N\Delta x, (-N + 1)\Delta y) & \ldots \\
\psi(x = -N\Delta x, N\Delta y) & \psi(x = (-N + 1)\Delta x, -N\Delta y) & \ldots \end{bmatrix}^T$$

and they calculated how the wrapping was to be done appropriately. Similar
methods have been used to solve other variants of the Schrodinger equation [8][9][10] and also the Gross-Pitaevski equation[13]. In our case, we generalize the same calculations to 3 dimensions with axial symmetry.

2.2 Calculations

In spherical polar coordinates, the Hamiltonian is given by (where the solution is axially symmetric with m = 0):

\[ \mathcal{H}\psi = -\frac{1}{2r^2} \frac{d}{dr} (r^2 \frac{d}{dr} \psi) - \frac{1}{2r^2 \sin \theta} \frac{d}{d\theta} (\sin \theta \frac{d}{d\theta} \psi) + V(r, \theta) \psi(r, \theta) = E\psi(r, \theta) \quad (2.1) \]

Using the central finite difference algorithm for defining theta, we get:

\[ \frac{d}{d\theta} (\sin \theta \frac{d\psi}{d\theta}) = \frac{1}{\Delta \theta} \left( \sin(\theta_{j+\frac{1}{2}}) \frac{d\psi}{d\theta}(r, \theta_{j+\frac{1}{2}}) - \sin(\theta_{j-\frac{1}{2}}) \frac{d\psi}{d\theta}(r, \theta_{j-\frac{1}{2}}) \right) \quad (2.2) \]

\[ \frac{1}{r^2 \sin \theta} \frac{d}{d\theta} (\sin \theta \frac{d\psi}{d\theta}(r, \theta_j)) = \frac{1}{r^2 \Delta \theta^2 \sin(\theta_j)} \times \left[ \sin(\theta_{j+\frac{1}{2}}) \psi(r, \theta_{j+\frac{1}{2}}) + \sin(\theta_{j-\frac{1}{2}}) \psi(r, \theta_{j-\frac{1}{2}}) \right. \]

\[ \left. \sin(\theta_{j+\frac{1}{2}}) \psi(r, \theta_{j+\frac{1}{2}}) - (\sin(\theta_{j+\frac{1}{2}}) + \sin(\theta_{j-\frac{1}{2}})) \psi(r, \theta_j) \right] \quad (2.3) \]

Define the coefficients of \( \psi(r, \theta_{j-\frac{1}{2}}), \psi(r, \theta_{j+\frac{1}{2}}), \psi(r, \theta) \) as \( A_-, A_+, B \) respectively, each dependent on the chosen \( j \) index. We now have:

\[ -\frac{1}{2r^2} \frac{d}{dr} (r^2 \frac{d}{dr} \psi) - \frac{1}{2r^2} (A_-^j \psi(r, \theta_{j-\frac{1}{2}}) + A_+^j \psi(r, \theta_{j+\frac{1}{2}}) - B^j \psi(r, \theta_j)) + V(r, \theta) \psi(r, \theta) = E\psi(r, \theta) \quad (2.4) \]
2.2 Calculations

where

\[
A_\pm^j = \frac{\sin(\theta_j \pm \frac{1}{2})}{\sin(\theta_j) \Delta \theta^2}
\] (2.5)

\[
B^j = \frac{\sin(\theta_{j+\frac{1}{2}}) + \sin(\theta_{j-\frac{1}{2}})}{\sin(\theta_j) \Delta \theta^2}
\] (2.6)

We can simplify by defining \( \phi = \psi r \). Then the new Hamiltonian is:

\[
-\frac{1}{2} \frac{d^2 \phi}{dr^2}(r, \theta_j) - \frac{1}{2r^2}\left(A_\pm^j \phi(r, \theta_j - \frac{1}{2}) + A_\pm^j \phi(r, \theta_j + \frac{1}{2}) - B^j \phi(r, \theta_j)\right) + V(r, \theta_j) \phi(r, \theta_j) = E \phi(r, \theta_j)
\] (2.7)

This can be written as:

\[
\frac{d^2 \phi}{dr^2}(r, \theta_j) = \frac{A_\pm^j \phi(r, \theta_j - \frac{1}{2})}{2r^2} + \frac{A_\pm^j \phi(r, \theta_j + \frac{1}{2})}{2r^2} - 2[E - C(r, \theta_j)] \phi(r, \theta_j)
\] (2.8)

where

\[
C(r, \theta_j) = V(r, \theta_j) + \frac{B^j}{2r^2}
\] (2.9)

We want the wavefunction to be zero at \( R = R_{max} \), so it is physical. Thus, we have one of our boundary conditions:

\[
\phi(r_N, \theta) = 0 \quad \forall \theta
\] (2.10)

Also, since \( \phi = \psi r \), we must have:

\[
\phi(0, \theta) = 0 \quad \forall \theta
\] (2.11)

We also must remember that as we have chosen \( \theta \) between \([0, \pi]\), that the boundary conditions are:

\[
\theta_{-1} = \theta_0 \quad \& \quad \theta_N = \theta_{N-1}
\] (2.12)
This brings us to our desired form and now, defining $x = [\phi(r, \theta_0) \ \phi(r, \theta_1) \ \ldots \ \phi(r, \theta_{N-1})]^T$, and $y = \dot{x}$, we can write:

$$\dot{y} = -S \dot{x} \quad \text{(2.13)}$$

$$\dot{x} = y \quad \text{(2.14)}$$

where $S$ is given by

$$
\begin{pmatrix}
2(E - C(\theta_0, r)) + \frac{A_-(\theta_0)}{r^2} & \frac{A_+(-\theta_0)}{r^2} & 0 & 0 & \ldots & 0 \\
\frac{A_-(\theta_1)}{r^2} & 2(E - C(\theta_1, r)) & \frac{A_+(-\theta_1)}{r^2} & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \ldots & \ldots & 0 & \frac{A_-(\theta_N)}{r^2} & 2(E - C(\theta_N, r)) + \frac{A_+(-\theta_N)}{r^2}
\end{pmatrix}
$$

Applying the first order symplectic method for Eqn. (2.13), (2.14), we find:

$$y^{n+1} = y^n - (\Delta r) S^{n+\frac{1}{r^2}} x_n \quad \text{(2.15)}$$

$$x^{n+1} = x^n + (\Delta r) y_{n+1} \quad \text{(2.16)}$$

Eliminating $y_N$, we get:

$$x^{n-1} + [(\Delta r)^2 S^{n+\frac{1}{r^2}} - 2I] x^n + x^{n+1} = 0 \quad \text{(2.17)}$$

Using the boundary conditions $x_0 = x_M = 0$, and doing some simple algebraic
manipulations, we find:

\[
A \begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_{M-1}
\end{bmatrix} + 2(\Delta r)^2 E \begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_{M-1}
\end{bmatrix} = 0
\]  

(2.18)

where

\[
A = \begin{bmatrix}
  (\Delta r)^2 U^{1+0.5} - 2I & I_{N \times N} & 0 & 0 & \cdots \\
  I_{N \times N} & (\Delta r)^2 U^{2+0.5} - 2I & I_{N \times N} & 0 & \cdots \\
  \vdots & \vdots & \vdots & \vdots & \ddots \\
  0 & \cdots & 0 & I_{N \times N} & (\Delta r)^2 U^{M-0.5} - 2I
\end{bmatrix}
\]

(2.19)

and the U matrix has the form:

\[
U_{j, j^2}^{n+\frac{1}{2}} = -2 \left[ V(r_{n+\frac{1}{2}}, \theta_j) + \frac{B^i}{2r_{n+\frac{1}{2}}^2} \right] = -2 \left[ V(r_{n+\frac{1}{2}}, \theta_j) + \frac{\sin(\theta_j + \frac{1}{2}) + \sin(\theta_j - \frac{1}{2})}{2r_{n+\frac{1}{2}}^2 (\Delta \theta)^2 \sin(\theta_j)} \right]
\]

(2.20)

\[
U_{j, j^{\pm1}}^{n+\frac{1}{2}} = A_{\pm} = \frac{\sin(\theta_j \pm \frac{1}{2})}{r_{n+\frac{1}{2}}^2 (\Delta \theta)^2 \sin(\theta_j)}
\]

(2.21)

where it is understood that the terms \(U_{-1,0}\) and \(U_{N,N+1}\) are added to the terms in
2.3 The $\partial_\phi \psi \neq 0$ Case

$U_{0,0}$ and $U_{N,N}$ respectively, where periodic boundary conditions are encountered.

Going back to Eqn. (2.18), finding the eigenvalues of $A$ denoted by $\lambda_k$ allows us to calculate the energy levels of a system using the relation:

$$E_k = -\frac{\lambda_k}{2(\Delta r)^2}$$  \hspace{1cm} (2.22)

The eigenvectors denote the wavefunctions in the $\phi = 0$ plane.

2.3 The $\partial_\phi \psi \neq 0$ Case

We have assumed all along that the wavefunction has no axial variation i.e. $\partial_\phi \psi = 0$. If this were not so, we would have to add an extra term

$$-\frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \psi$$

to the left hand side of Equation 2.1. However, since there is no other $\phi$ dependence, we can simply do separation of variables and solve the $\phi$ equation independently. This gives us

$$\Psi(r, \theta, \phi) = \psi(r, \theta) e^{im\phi}$$  \hspace{1cm} (2.23)

The non-trivial $\phi$ dependence adds a term

$$-\frac{m^2}{r^2 \sin^2 \theta} \psi$$

to the left hand side of Equation 2.7. Since this term only depends on $\phi(r, \theta_j)$, it can be simply absorbed into $B^j$, or equivalently we can redefine

$$C(r, \theta_j) = V(r, \theta_j) + \frac{B^j}{2r^2} \frac{m^2}{r^2 \sin^2 \theta_j}$$  \hspace{1cm} (2.24)
The rest of the calculation proceeds as before. Thus, we can simply generalize this method for the \( m \neq 0 \) case.

2.4 The Numerical Implementation

Since we only want the extreme eigenvalues of a large but sparse matrix, ARPACK (ARnoldi PACKage) subroutines [11] allow us to find these results very quickly and efficiently. This package, used via a scipy wrapper, uses the implicitly restarted Arnoldi Method (IRAM) [14][15]

We have also optimized the creation of this large matrix by noting that only the V term in \( U_{j,j} \) is dependent on the external potential, thus one can create the rest of the A matrix for a given grid configuration and save it to disk, and then add a diagonal matrix for each potential well configuration.
Chapter 3

Symplectic Results

Summary

In this chapter, we describe the various calculations done using the symplectic method developed in the previous chapter. We compare the results with the analytical results for a Simple Harmonic Oscillator, Infinite Spherical Potential Well, Finite Spherical Potential Well with tunneling. We also compare with the numerical results for infinite potential waves of arbitrary bubble shapes obtained by Humphrey et.al. (2003). We discuss how the method can generalize to other problems.

3.1 Simple Harmonic Oscillator

We solve the problem with different values of M, which shows the effect that improving the fineness of the radial grid increases the accuracy. Since the problem is symmetric in $\theta$, the change in N should not cause any appreciable changes to the result.

This can be solved by the separation of variables [16], and the allowed energies
3.2 Infinite Spherical Potential Well

are:

\[ E = (n_x + n_y + n_z + 3/2) \hbar \omega \]  

(3.1)

We test the problem with \( \hbar = 1 \) and \( \omega = 1 \), and thus the expected energy levels are \( \frac{3}{2}, \frac{5}{2}, \frac{7}{2}, \ldots \). The results are given in Table 3.1.

<table>
<thead>
<tr>
<th>Level</th>
<th>Exact</th>
<th>M=41 N=50</th>
<th>M=81 N=50</th>
<th>M = 161 N = 50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ground State</td>
<td>1.5000</td>
<td>1.5831</td>
<td>1.5420</td>
<td>1.5210</td>
</tr>
<tr>
<td>1st Excited</td>
<td>2.5000</td>
<td>2.4963</td>
<td>2.4985</td>
<td>2.4993</td>
</tr>
<tr>
<td>2nd Excited</td>
<td>3.5000</td>
<td>3.4940</td>
<td>3.4966</td>
<td>3.4985</td>
</tr>
<tr>
<td>3rd Excited</td>
<td>4.5000</td>
<td>4.4900</td>
<td>4.4953</td>
<td>4.4984</td>
</tr>
<tr>
<td>4th Excited</td>
<td>5.5000</td>
<td>5.4824</td>
<td>5.4935</td>
<td>5.4966</td>
</tr>
</tbody>
</table>

Table 3.1: Simple Harmonic Oscillator Results

We can see that the accuracy increases with increase in M. The eigenvectors corresponding to the eigenvalues give us the wavefunctions, and some of them are plotted in Figure 3.1.

3.2 Infinite Spherical Potential Well

In this case, too, the solution proceeds by separation of variables [16]. Defining

\[ \psi(r, \theta, \phi) = R_{nl}(r)Y_{lm}(\theta, \phi) \]

and using the standard properties of the spherical harmonics, we find that the radial function \( R_{nl}(r) \) satisfies

\[ \frac{d^2 R_{nl}}{dr^2} + \frac{2}{r} \frac{d R_{nl}}{dr} + \left( \frac{k^2}{r^2} - \frac{l(l+1)}{r^2} \right) R_{nl} = 0 \]  

(3.2)
3.2 Infinite Spherical Potential Well

Defining $z = kr$, the equation transforms to

$$
\frac{d^2 R_{nl}}{dz^2} + \frac{2}{z} \frac{dR_{nl}}{dz} + \left(1 - \frac{l(l+1)}{z^2}\right) R_{nl} = 0
$$

(3.3)

The solutions to this equation are the spherical Bessel functions. Since the wavefunction must not blow up at $z = 0$, the only acceptable solution is:

$$
j_l(z) = z^l \left(-\frac{1}{z} \frac{d}{dz}\right)^l \left(\frac{\sin z}{z}\right)
$$

(3.4)

Our second boundary condition is that $R_{nl}(a) = 0$, where $a$ is the radius of the well. Thus, we want $ka = z_{nl}$, where $z_{nl}$ denotes the n-th zero of $j_l(z)$. These values are tabulated in Table 3.2

<table>
<thead>
<tr>
<th></th>
<th>n=1</th>
<th>n=2</th>
<th>n=3</th>
<th>n=4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$l = 0$</td>
<td>3.142</td>
<td>6.283</td>
<td>9.425</td>
<td>12.566</td>
</tr>
<tr>
<td>$l = 1$</td>
<td>4.493</td>
<td>7.725</td>
<td>10.904</td>
<td>14.066</td>
</tr>
<tr>
<td>$l = 2$</td>
<td>5.763</td>
<td>9.095</td>
<td>12.323</td>
<td>15.515</td>
</tr>
</tbody>
</table>

Table 3.2: Zeros of the Spherical Bessel Function

The acceptable energy values are

$$
E = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2 z_{nl}^2}{2m a^2}
$$

(3.5)

In natural units which I have used, $\hbar = 1$, $m = 1$. Also, for convenience, we choose a well of radius equal to the Bohr radius, $a_0 = 0.529 \times 10^{-10} m$, so the energy numbers are in Hartree units. For reference, 1 Hartree = 27.2 eV.
Proceeding similarly as before, we find that for $0 < E < V_0$, the radial function satisfies Equation 3.2:

$$\frac{d^2 R_{nl}}{dz^2} + \frac{2}{z} \frac{dR_{nl}}{dz} - \frac{l(l+1)}{z^2} R_{nl} + \left(\frac{q^2}{\kappa^2}\right) R_{nl} = 0$$

(3.6)

where $q = \sqrt{2mE}$ inside the well and $\kappa = \sqrt{2m(V_0 - E)}$ outside the well. Using our boundary conditions that the solution is well behaved at $r = 0$ and decays to zero as $r \to \infty$, we have the following wavefunctions:

$$R(r) = A j_l(qr) \quad 0 < r < a \quad (3.7)$$

$$R(r) = B h_l^{(1)}(ikr) \quad r > a \quad (3.8)$$

Enforcing continuity and differentiability at the boundary $r = a$ gives us the constraint equation [16]:

$$q \frac{d \log j_l}{dz} \bigg|_{qa} = i\kappa \frac{d \log h_l^{(1)}}{dz} \bigg|_{i\kappa a}$$

(3.9)

This equation can be solved numerically and/or graphically for different values
3.4 Infinite Well of Arbitrary Geometry

of l to get the values of q, and hence E. This was done using mathematica, and the analytical results are compared with the symplectic results in Table 3.4.

<table>
<thead>
<tr>
<th>Level (n, l)</th>
<th>Exact Energy</th>
<th>M=81 N=50</th>
<th>Error Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 0</td>
<td>4.31</td>
<td>4.37</td>
<td>1.39</td>
</tr>
<tr>
<td>1, 1</td>
<td>8.79</td>
<td>8.77</td>
<td>0.23</td>
</tr>
<tr>
<td>1, 2</td>
<td>14.45</td>
<td>14.41</td>
<td>0.27</td>
</tr>
<tr>
<td>2, 0</td>
<td>17.15</td>
<td>17.43</td>
<td>1.63</td>
</tr>
<tr>
<td>1, 3</td>
<td>21.22</td>
<td>21.14</td>
<td>0.38</td>
</tr>
<tr>
<td>2, 1</td>
<td>25.87</td>
<td>25.81</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Table 3.4: Finite Spherical Well with \( R = a_0 \) and \( V = 100 \)

The wavefunction cross-section along the symmetry axis is plotted for some of the levels in Figure 3.2.

3.4 Infinite Well of Arbitrary Geometry

To solve the electron bubble problem described in Section 1.1, it is necessary to calculate the electronic energy in an infinite potential well but with an arbitrary geometry in an efficient way. This problem was tackled by Humphrey et.al. [1], [2] using the following method. First, parametrize the shape by Legendre polynomials. Since the shape has to be axially symmetric, we only need the even ones. Thus:

\[
R(\theta) = \sum_{l=0,2,4,...} a_L P_L(cos \theta) \tag{3.10}
\]

Now, the wavefunction inside the well must again satisfy the radial equation (Equation 3.2). Thus, we can write the radial function as a superposition of
3.4 Infinite Well of Arbitrary Geometry

Figure 3.1: Wavefunctions for the Simple Harmonic Oscillator

(a) Ground State  
(b) First Excited State  
(c) Second Excited State  
(d) Third Excited State  
(e) Fourth Excited State

Figure 3.2: Wavefunction Cross-sections for the Finite Potential Well

(a) (1,0) state  
(b) (1, 1) state  
(c) (2, 0) state  
(d) (2, 1) State
3.4 Infinite Well of Arbitrary Geometry

Bessel functions. This gives the wavefunction the final form:

$$\psi(r, \theta, \phi) = \sum_l A_l P_l^m(\cos \theta) e^{im\phi} j_l(kr)$$  \hspace{1cm} (3.11)

Given an arbitrary shape, one can randomly pick points on the surface of the bubble and minimize $|\psi|^2$ there. If one chooses the correct number of points, this leads to a linear system in terms of $A$ and $B$, and then it is exactly solvable. The energy $E_{el}$ was found by doing a linear search over all energies and minimizing $|\psi|^2$ for each value, and finding the value of energy which minimizes the function overall.

I have compared the results from this method to the results obtained from the symplectic method. For my calculations, all the shape coefficients have units of angstroms, and all energy units are in Hartree. The comparison for the 1p results (the first excited state) is presented in Table 3.5. The wavefunctions for the 1p and 1d levels are given in Figure 3.3.
3.4 Infinite Well of Arbitrary Geometry

(a) 1p state

(b) 1d state

Figure 3.3: Wavefunctions for the Peanut-Shaped Electron Bubble

<table>
<thead>
<tr>
<th>$a_0$</th>
<th>$a_2$</th>
<th>$a_4$</th>
<th>$a_6$</th>
<th>Humphrey</th>
<th>M=81 N=50</th>
<th>Error Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>29.75</td>
<td>6.11</td>
<td>-0.23</td>
<td>0.01</td>
<td>2.753</td>
<td>2.724</td>
<td>1.063</td>
</tr>
<tr>
<td>26.38</td>
<td>7.16</td>
<td>-0.37</td>
<td>0.03</td>
<td>3.359</td>
<td>3.314</td>
<td>1.352</td>
</tr>
<tr>
<td>23.80</td>
<td>7.99</td>
<td>-0.54</td>
<td>0.05</td>
<td>3.972</td>
<td>3.907</td>
<td>1.648</td>
</tr>
<tr>
<td>22.35</td>
<td>8.46</td>
<td>-0.67</td>
<td>0.07</td>
<td>4.397</td>
<td>4.324</td>
<td>1.666</td>
</tr>
<tr>
<td>21.33</td>
<td>8.75</td>
<td>-0.77</td>
<td>0.12</td>
<td>4.742</td>
<td>4.671</td>
<td>1.490</td>
</tr>
<tr>
<td>20.92</td>
<td>8.87</td>
<td>-0.82</td>
<td>0.14</td>
<td>4.890</td>
<td>4.820</td>
<td>1.437</td>
</tr>
<tr>
<td>19.91</td>
<td>9.14</td>
<td>-0.96</td>
<td>0.19</td>
<td>5.294</td>
<td>5.211</td>
<td>1.572</td>
</tr>
<tr>
<td>19.13</td>
<td>9.35</td>
<td>-1.08</td>
<td>0.25</td>
<td>5.640</td>
<td>5.535</td>
<td>1.866</td>
</tr>
<tr>
<td>18.32</td>
<td>9.55</td>
<td>-1.21</td>
<td>0.32</td>
<td>6.044</td>
<td>5.932</td>
<td>1.849</td>
</tr>
<tr>
<td>17.89</td>
<td>9.67</td>
<td>-1.31</td>
<td>0.37</td>
<td>6.271</td>
<td>6.162</td>
<td>1.737</td>
</tr>
<tr>
<td>17.50</td>
<td>9.78</td>
<td>-1.41</td>
<td>0.42</td>
<td>6.485</td>
<td>6.373</td>
<td>1.728</td>
</tr>
<tr>
<td>17.15</td>
<td>9.88</td>
<td>-1.49</td>
<td>0.47</td>
<td>6.687</td>
<td>6.576</td>
<td>1.667</td>
</tr>
<tr>
<td>16.83</td>
<td>9.98</td>
<td>-1.58</td>
<td>0.52</td>
<td>6.876</td>
<td>6.761</td>
<td>1.678</td>
</tr>
<tr>
<td>16.28</td>
<td>10.16</td>
<td>-1.75</td>
<td>0.63</td>
<td>7.056</td>
<td>7.098</td>
<td>0.593</td>
</tr>
</tbody>
</table>

Table 3.5: 1p levels for Bubbles with Infinite Walls
3.5 Future Scope

We have seen that the symplectic method works for the infinite spherical potential well, and also for the spherical potential well with finite walls. We have also verified the energy results from our method to Humphrey’s for infinite potential wells of arbitrary cross-section to reasonable accuracy. We are as of yet unable to find the shape of the bubble which minimizes the overall energy - the shape coefficients used in Table 3.5 were used from results from the Maris group [2].

One possible application of the symplectic method is to now apply it to find the energy levels and wavefunctions for bubbles with arbitrary cross-sections but a finite potential well. Some of the bubble shapes found by Humphrey et.al. have sharp concave corners which may lead to significant tunnelling and drastically affect the final result.

Another possible application of the method may be to solve multi-electron bubbles - where there are many electrons in the same bubble (potential well) which can interact with each other also. In this case, we can model the interaction between electrons as an effective classical potential using the Hartree approximation. We can use this potential to solve for the wavefunctions and energy levels of a single electron in the system, and use iterative algorithms to try and obtain a self-consistent solution.
Chapter 4

Quantum Mechanics using
Restricted Boltzmann Machines

Summary

In this chapter, we discuss the various problems plaguing modern methods to solve the quantum many-body wavefunction. We explain the restricted Boltzmann machine scheme as an alternative method to solve these problems. Then, we try to gain insight into the physical significance of the Restricted Boltzmann machine (RBM) parameters. We argue that for a single particle in a Hilbert space of dimension N, one can describe almost any state as an RBM with a single visible node taking N allowed values and $M \geq \frac{2N-3}{4}$ binary valued hidden nodes. We then study the physical significance of the parameters for the transverse field Ising model.
4.1 Introduction

The quantum wavefunction characterizing a many-particle state is a very complex object. For a Hilbert space with L sites/particles and N states per site, the total number of basis vectors is $N^L$. Thus, we need an exponential number of complex numbers to specify a given state in the Hilbert space - which makes exploring the entire state space clearly impossible even for moderate N and L.

Two typical strategies are used to overcome this problem. The first one is called quantum Monte Carlo (QMC). In this case, the method samples only a finite number of configurations using a probabilistic framework, and hopes that the configurations are statistically representative of the full Hilbert space. The second strategy is called the tensor network (TN) methods, where the idea is that we are only interested in states of the Hilbert space with weak correlations (entanglement), and hence we can choose a basis where we can efficiently compress the wavefunction to have fewer free parameters. If the L particles were completely independent, we could describe each particle with N complex numbers and we would need only $N \times L$ complex numbers to describe the entire system, a massive improvement. Even for small entanglement, structures like Matrix Product States (MPS) reduces the complexity similarly. However, current tensor networks approaches also fail to compress efficiently in higher dimensions, and cannot handle systems with moderate levels of entanglement.

Neural networks tackle this same problem through the lens of dimensional reduction and feature extraction. Effectively, a neural network tries to learn the best ”basis” to represent the quantum state in. Thus, it is one of the ”compression” approaches which falls in the same family as the tensor networks.
4.2 Restricted Boltzmann Machines

A restricted Boltzmann Machine (RBM) [17] is an undirected graphical model which can be used to learn a given probability distribution. It is a recurrent, energy based, stochastic and generative model.

An energy based model attaches a number, defined as the "Energy", to a particular configuration of the model. The energy functional needs to be trained to have the desired properties. Define the energy of a configuration $x$ of the network as $E(x)$. The probability distribution the network represents is given by:

$$p(x) = \frac{e^{-E(x)}}{Z}$$

(4.1)

where $Z$ is a normalizing factor, known as the partition function, and is given by:

$$Z = \sum_x e^{-E(x)}$$

(4.2)

The reason for choosing this particular form (Equation 4.1) can be seen from defining an energy function and then looking for the stationary distribution - which will turn out to have this particular form.

Given our training data, we can find the best energy-based model to represent the training data by maximizing the likelihood of the training data. Since the logarithm is an increasing function, we can maximize the log-likelihood, which is equivalent to minimizing the negative log-likelihood (NLL) function, denoted by $\mathcal{L}$, through gradient descent:

$$\mathcal{L}(\theta, \mathcal{D}) = \frac{1}{N} \sum_{x^{(i)} \in \mathcal{D}} \log p(x^{(i)})$$

(4.3)
The loss function is given by the negative of the NLL function:

\[ \ell(\theta, \mathcal{D}) = -\mathcal{L}(\theta, \mathcal{D}) \]  

(4.4)

and the best function can be trained by using the stochastic gradient: \(-\frac{\partial \log p(x^{(i)})}{\partial \theta}\), where \(x^{(i)}\) denote training data and \(\theta\) denote the model parameters.

The ability of the energy-based model to represent a given probability distribution can be enhanced by adding extra degrees of freedom to the energy functional, called hidden nodes. These cannot be observed, and any probability distribution observed over the visible nodes must trace over these nodes. Thus:

\[ P(x) = \sum_h P(x, h) = \sum_h \frac{e^{-E(x,h)}}{Z} \]  

(4.5)

and we can also define the free energy \(\mathcal{F}\) as:

\[ \mathcal{F}(x) = -\log \sum_h e^{-E(x,h)} \]  

(4.6)

Figure 4.1: Restricted Boltzmann machine

Given this background, let us move on to Restricted Boltzmann machines. A typical RBM is denoted in Figure 4.1. \(v\) denotes the visible layers, indexed by \(I\), and \(h\) denotes the hidden layers, indexed by \(J\). The visible layers have a bias \(a_I\), whereas the hidden layers have a bias \(b_J\). The hidden and visible layers are
connected by the weights $W_{IJ}$. The energy is now defined by

$$E(v, h) = - \sum_I a_I v_I - \sum_J b_J h_J - \sum_{I,J} h_J W_{IJ} v_I \quad (4.7)$$

Because of the structure of the RBM, visible and hidden units are conditionally independent if we know the value of the other set. Hence:

$$p(h|v) = \prod_i p(h_i|v)$$

$$p(v|h) = \prod_j p(v_j|h) \quad (4.8)$$

Given an input vector $v$, the probability of the hidden node $h_J$ being activated is:

$$p(h_J = 1|v) = \frac{1}{1 + e^{-(b_J + \sum_i v_i W_{ij})}} = \text{sig} \left( b_J + \sum_i v_i w_{ij} \right) \quad (4.9)$$

where

$$\text{sig}(x) = \frac{1}{1 + e^{-x}} \quad (4.10)$$

If we assume binary hidden nodes, $h_J \in \{-1, 1\}$, then we find:

$$P(v; W) = \exp(\sum_I a_I v_I) \prod_J 2 \cosh(b_J + \sum_I W_{IJ} v_I) \quad (4.11)$$

The Boltzmann machine is very powerful because there are easy ways to train this system by either gradient descent or the contrastive divergence method. Also, the RBM is also an universal approximator [5], meaning it can approximate any "reasonably-behaved" function with sufficient number of hidden units - this leads to very high expressive power of the RBM.
4.3 The Variational Method

The variational principle works on the claim that given a Hamiltonian $H$, choose any arbitrary normalized wavefunction $\psi$, and

$$E_{gs} \leq \langle \psi | H | \psi \rangle = \langle H \rangle$$  \hspace{1cm} (4.12)

The proof goes as follows [18]: Let us denote the true eigenvectors of the Hamiltonian as $\psi_n$ with energy $E_n$. Any arbitrary state $\psi$ can be expanded in a linear combination:

$$\psi = \sum_n c_n \psi_n$$  \hspace{1cm} (4.13)

where $\sum_n |c_n|^2 = 1$ since $\psi$ is normalized. Now,

$$\langle H \rangle = \left\langle \sum_m c_m \psi_m | H \sum_n c_n \psi_n \right\rangle = \sum_m \sum_n c_m^* E_n c_n \langle \psi_m | \psi_n \rangle = \sum_n E_n |c_n|^2$$  \hspace{1cm} (4.14)

Now, since the ground state has the least energy, $E_{gs} \leq E_n$. Thus,

$$\langle H \rangle \geq E_{gs} \sum_n |c_n|^2 = E_{gs}$$  \hspace{1cm} (4.15)

The importance of the variational principle is that any function one can dream of gives an upper bound for the ground state energy of the system. Thus, a carefully designed trial wavefunction can give one a very close estimate of the energy of the system. Also, the closer the trial energy is to the true energy, the closer the variational wavefunction is to the true wavefunction [18].

The basic strategy of the variational approach is:

- Guess a good trial function with parameters $\theta$
- Find the expectation value of the energy in the state as a function of $\theta$
4.4 Representing a single particle

- Find $\theta$ that minimizes the energy overall

The hard part is to find the good trial function. RBMs take care of this problem by providing a highly expressive, easily trainable (minimizable) parametric functional form that can approximate the wavefunction and energy quite accurately. Let us study now how this works on a single particle problem.

4.4 Representing a single particle

Let us denote the wavefunction of the ising spins to be

$$|\Psi\rangle = \sum_{\sigma_I} \Psi(\{\sigma_I\}) |\sigma_I\rangle$$  \hspace{1cm} (4.16)

We want to denote the quantum amplitude of a many particle state as the partition function of a classical spin system. We add the many body coordinates in the visible layer, and the classical spins in the hidden layer.

Continuing in standard machine learning parlance, we define an "energy":

$$E(\{\sigma_I\}, \{h_J\}) = \sum_I a_I \sigma_I + \sum_J b_J h_J + \sum_{IJ} W_{IJ} \sigma_I h_J$$  \hspace{1cm} (4.17)

The claim is that:

$$\Psi(\{\sigma_J\}) \sim \text{tr}_h e^{-E(\{\sigma_I\}, \{h_J\})}$$

$$\Psi(\{\sigma_J\}) \sim e^{-\sum_I a_I \sigma_I} \prod_J \left(1 + \exp[-(b_J + \sum_I W_{IJ} \sigma_I)]\right)$$  \hspace{1cm} (4.18)

where we have summed over $h_J \in \{0, 1\}$. This is our trial wavefunction.

For a single particle in an N dimensional Hilbert space spanned by orthonor-
4.4 Representing a single particle

<table>
<thead>
<tr>
<th>(N)</th>
<th>(\psi_0)</th>
<th>(\psi_1)</th>
<th>(\psi_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>((1 + e^{-b}))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>((1 + e^{-b}))</td>
<td>(e^{-a}(1 + e^{-b+W}))</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>((1 + e^{-b}))</td>
<td>(e^{-a}(1 + e^{-b+W}))</td>
<td>(e^{-2a}(1 + e^{-b+2W}))</td>
</tr>
</tbody>
</table>

Table 4.1: Wavefunction Expressions for the first few \(N\) basis vectors \(|I\rangle, I = 0, 1, ..., N - 1\), a generic Hamiltonian is

\[
\mathcal{H} = \sum_{I_1, I_2} H_{I_1, I_2} |I_1\rangle \langle I_2|
\] (4.19)

while a generic single particle state is given by

\[
|\psi\rangle = \sum_I \psi_I |I\rangle
\] (4.20)

We require \(|\psi\rangle\) to be a non-zero vector (unnormalized). This state is described by \(2N\) real numbers. We can remove an overall phase, which is equivalent to choosing \(\psi_0\) positive. The wavefunction is given by:

\[
\psi_I \sim e^{-Ia} \prod_j \left(1 + e^{-(b_j + IW_j)}\right)
\] (4.21)

This has \(4M + 2\) real parameters. For an arbitrary \(N\), we need that we have sufficient real parameters i.e.

\[
4M + 2 \geq 2N - 1
\] (4.22)

\[
M \geq \frac{2N - 3}{4}
\] (4.23)

We can explicitly tabulate the values for a few \(N\) for \(M=1\), as shown in Table 4.1.

For \(N=1\), the solution obviously exists. For \(N=2\), \(a\) can store the phase of
4.4 Representing a single particle

\( \psi_1 \) and \( W \) it’s magnitude. For \( N=3 \), we have two complex numbers \( a, W \) to be determined from two complex numbers \( \psi_1, \psi_2 \) - one would expect an unique solution. Then, we would need a second hidden node starting from \( N=4 \), which is in agreement with Equation 4.23.
Chapter 5

Applying the RBM Formalism

Summary

In this chapter, we apply the RBM variational trial function to the Transverse Field Ising (TFI) model, the tight binding (TB) model, and the Anderson tight-binding model in 1 dimension. We find out the simplest RBM which can represent the ground state wavefunction in each case, and try to develop physical intuition about what each parameter represents.

5.1 Fitting the RBM trial function

For this section only, let the set of all RBM parameters be denoted by $W$, and each individual parameter be denoted as $W_k$. To recap, we have defined a wavefunction $\psi$ as

$$\psi(S; W) = e^{-\sum_i a_i \sigma_i} \prod_j (cosh(b_j + \sum_i W_{ij} \sigma_i))$$

and we have defined

$$W = \{\{a\}, \{b\}, \{W\}\}$$
The variational method demands that we choose $W$ so as to minimize the expectation value of the Hamiltonian for this wavefunction: i.e. the function we want to minimize is:

$$E(W) = \langle \Psi_M | \mathcal{H} | \Psi_M \rangle / \langle \Psi_M | \Psi_M \rangle$$

(5.3)

where the denominator takes care of the fact that $\psi$ is not normalized. The optimal solution, given by $\nabla E(W^\ast) = 0$, can be found if we could minimize the energy functional using something like gradient descent.

The only problem with this approach is that we can calculate $\psi(S) = \langle S | \psi \rangle$, but calculating the entire wavefunction will take an addition over an exponential number of spin configurations.

The way around this is to find an expected value of the gradient instead by importance sampling the gradient of $E$ with respect to the parameter by Monte-Carlo. This is done using the following steps:

- For a given spin configuration $S$, calculate the gradient of energy explicitly.
- Now choose a new configuration $S'$. Choose that state with the probability $\min\left(1, \frac{|\psi(S' | \psi) |^2}{|\psi(S | \psi) |^2} \right)$. Otherwise stay on the same state $S$.
- Repeat steps 1 and 2. After a desired number of cycles, the average value of the gradient of energy is the expectation value according to importance sampling Monte-Carlo.

It can be shown [19] [6] that the expectation value for gradient for the energy function is given by:

$$F_k = \langle E_{\text{loc}} (O^\ast_k) \rangle - \langle E_{\text{loc}} \rangle \langle O^\ast_k \rangle$$

(5.4)
5.1 Fitting the RBM trial function

where

\[ O_k(S) = \frac{1}{\Psi_M(S)} \partial_{w_k} \Psi_M(S) \]  
(5.5)

\[ E_{\text{loc}}(S) = \frac{\langle S|H|\Psi_M \rangle}{\Psi_M(\delta)} \]  
(5.6)

and all expectation values are assumed to be importance sampled. Thus, using stochastic gradient descent, we would find:

\[ W_k(p + 1) = W_k(p) - \gamma(p) F_k(p) \]  
(5.7)

where \( k \) denotes the parameter index, \( p \) denotes the iteration, and \( \gamma \) is a scaling parameter that decides how large the step is going to be. \( \gamma \) depends on \( p \) because it is usually chosen to get smaller with larger iterations, so the solution can get closer and closer to the true minimum.

However, this simple gradient descent can be improved by accounting for the correlations between each variable. This was done by Sorella et.al [20], who found that a better algorithm would use the update rule:

\[ W(p + 1) = W(p) - \gamma(p) S^{-1}(p) F(p) \]  
(5.8)

where

\[ S_{kk'}(p) = \langle O_{k'}^* O_k \rangle - \langle O_{k'}^* \rangle \langle O_k \rangle \]  
(5.9)

is the hermitian covariance matrix. This method is called the stochastic reconfiguration (SR) method and has been used to train the RBM. The SR method has been shown [20] to be an effective imaginary-time evolution in the variational subspace.

An open source framework called "NetKet" to study quantum mechanics us-
ing machine learning has been developed by Carleo et.al [21]. We have used this package to generate the RBM representation and train it via stochastic reconfiguration.

5.2 Transverse Field Ising model

Having developed the RBM formalism, let us apply it to the transverse field Ising model (TFI) and see what we get. For simplicity, let us consider only two sites. The Hamiltonian is given by:

$$H_{TFI} = -h \sum_i \sigma_i^x - J \sum_{\langle i,j \rangle} \sigma_i^z \sigma_j^z$$  \hspace{1cm} (5.10)

Choosing a basis to be

$$\mathcal{H} = \text{span}(|\uparrow\rangle, |\downarrow\rangle) \otimes \{|\uparrow\rangle, |\downarrow\rangle\})$$  \hspace{1cm} (5.11)

we find that

$$\sigma_1^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \text{Id}_{2\times2}$$  \hspace{1cm} (5.12)

$$\sigma_2^x = \text{Id}_{2\times2} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$  \hspace{1cm} (5.13)
5.2 Transverse Field Ising model

\[
\sigma_1^z \sigma_2^z = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\] (5.14)

This gives the overall Hamiltonian in this basis,

\[
H = \begin{pmatrix}
-J & -h & -h & 0 \\
-h & J & 0 & -h \\
-h & 0 & J & -h \\
0 & -h & -h & -J
\end{pmatrix}
\] (5.15)

Let the final wavefunction be denoted as:

\[
|\psi\rangle = c_1 |\uparrow\uparrow\rangle + c_2 |\uparrow\downarrow\rangle + c_3 |\downarrow\uparrow\rangle + c_4 |\downarrow\downarrow\rangle
\] (5.16)

From the symmetry of the Hamiltonian, \(c_1 = c_4\) and \(c_2 = c_3\). We want to set up an RBM with 2 visible nodes to denote the 2 spins. We have M binary hidden nodes, and we want to find the parameters such that inputting a certain configuration in the visible nodes makes the RBM evaluate to the coefficient of the same configuration. Let us write out the expressions explicitly for a single hidden node, using the notation \(\psi(S) = \langle S|\psi\rangle\), where \(S\) is a spin configuration.
of the visible nodes:

\[
\psi(\uparrow\uparrow) = c_1 = Ae^{a_1 + a_2} \cosh(w_1 + w_2 + b) \tag{5.17}
\]

\[
\psi(\uparrow\downarrow) = c_2 = Ae^{a_1 - a_2} \cosh(w_1 - w_2 + b) \tag{5.18}
\]

\[
\psi(\downarrow\uparrow) = c_3 = Ae^{-a_1 + a_2} \cosh(-w_1 + w_2 + b) \tag{5.19}
\]

\[
\psi(\downarrow\downarrow) = c_4 = Ae^{-a_1 - a_2} \cosh(-w_1 - w_2 + b) \tag{5.20}
\]

Multiplying the first equation with the fourth, and the second with the third gives us:

\[
c_1 c_4 = A^2 \cosh(w_1 + w_2 + b) \cosh(-w_1 - w_2 + b) \tag{5.21}
\]

\[
c_2 c_3 = A^2 \cosh(-w_1 + w_2 + b) \cosh(w_1 - w_2 + b) \tag{5.22}
\]

Using hyperbolic sum-angle formulae, we find:

\[
\frac{c_2 c_3}{c_1 c_4} = \frac{\cosh(2(w_1 - w_2)) + \cosh(2b)}{\cosh(2(w_1 + w_2)) + \cosh(2b)} = p \tag{5.23}
\]

Let us try to find a solution with \(a_1 = a_2 = b = 0\). The requirement \(c_1 = c_4\) and \(c_2 = c_3\) is immediately satisfied. We want the coefficients to be real. Since \(\cosh(x)\) is real only if \(x\) is purely real or purely imaginary, we must have

\[
W_1 = \pm W_2^* \tag{5.24}
\]

Now, we can use our physical intuition to see that for \(h = 0\), we expect both the spins to be pointing in the same direction. For small \(h\), thinking perturbatively, we should expect \(c_1\) to be a large number close to 1, and \(c_2\) to be a small number close to 0. If \(W_1 = -W_2^*\), we have \(c_1 \sim \cos(x) < 1\) and \(c_2 \sim \cosh(x) > 1\). Thus,
from physical considerations, we must choose

\[ W_1 = W_2^* = x + iy \quad (5.25) \]

where \( x \) and \( y \) are real numbers.

With this notation, we can simplify Equation 5.23 to:

\[
\begin{align*}
\frac{1 + \cosh(4y)}{1 + \cosh(4x)} &= p \\
\frac{1 + \cos(4y)}{1 + \cosh(4x)} &= p
\end{align*}
\]

(5.26) (5.27)

Now, \( \cosh x \geq 1 \) and \( -1 \leq \cos(y) \leq 1 \), thus we find:

\[ 0 \leq p \leq 1 \quad (5.28) \]

This is in-line with our physical intuition that the ground state wavefunction can range from both being aligned with each other:

\[ \psi_{\text{grnd}} = \frac{1}{\sqrt{2}} |\uparrow\uparrow\rangle + \frac{1}{\sqrt{2}} |\downarrow\downarrow\rangle \]

corresponding to \( p=0 \), upto the fully mixed state:

\[ \psi_{\text{grnd}} = \frac{1}{2} \left( |\uparrow\uparrow\rangle + |\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle + |\downarrow\downarrow\rangle \right) \]

corresponding to \( p = 1 \). \( p \) defines the ratio of \( c_1 \) and \( c_2 \), and thus fully specifies the wavefunction. Thus, just 1 real number \( p \) is required to define the system.

We can represent this system with an RBM without any biases. This has one complex parameter, \( W \) - or equivalently, the two real parameters \( x \) and \( y \), related by Equation 5.27.
5.3 Tight Binding Model

Equation 5.27 represents a closed ellipse-like in the X-Y plane with the p value basically defining the ratio of the semimajor to the semiminor axis. The scale, of course, is set by the normalization condition.

If we let $p = 0$, we must have $\cos(4y) = -1$ or equivalently

$$y = \pm \frac{\pi}{4} \pm \frac{2\pi m}{4} \quad m \in \mathbb{Z}$$

This is verified in Figure 5.1b where we get two horizontal lines at precisely these fixed values of $y$ and all allowed $x$ values.

For the more general $p \neq 0$, proceeding from Equation 5.27, we get:

$$\cos^2(2y) = 2p(1 + \cosh 4x)$$

$$y = \arccos \left( \sqrt{\frac{2p(1 + \cosh 4x)}{2}} \right)$$

This is shown in Figure 5.1a. Moreover, given the manifold of weights obtained, it was also possible to reproduce $p$ accurately by fitting the curve (Eqn 5.31). This is very interesting because every point on the complex curve $W$ uniquely defines the variable $p$, and hence can identify the defining parameter of the ground state. In other words, the weights capture enough information to obtain the ground state uniquely - but they represent it as an unusual "weight" quantity, whose physical meaning we do not know yet. This leads us to our next question - what physical quantities do these $W$ physically represent?

5.3 Tight Binding Model

To try and answer this question, let us turn to the single spinless particle in a tight-binding model with $N$ sites, and periodic boundary conditions. The Hamil-
5.3 Tight Binding Model

(a) Manifold of $W$: $p \neq 0$

Manifold of $W$ for different values of $p$ (Analytical & Numerical)

(b) Dependence of $W$ Manifold on $P$

Figure 5.1: Figures show the manifold of allowed $W$ values obtained for different values of $p$ - for the transverse field Ising model

39
Tight Binding Model

The Hamiltonian is given by:

\[ H = -t \sum_{\langle i,j \rangle} \left( c_i^\dagger c_j + h.c. \right) \]  \hspace{1cm} (5.32)

where \( \langle i,j \rangle \) means that \( i \) and \( j \) are nearest neighbours. We can turn this into a single particle with \( N \) spin values allowed per site as follows:

We expect the resultant solution to have the form:

\[ |\psi\rangle = \sum_{i=0}^{N-1} \gamma_i |i\rangle \]  \hspace{1cm} (5.33)

Thus, we can represent this in a RBM by using a single visible node which is allowed to take on an integer value between 0 and \( (N-1) \). Plugging into the visible node each site number should make the RBM evaluate to the corresponding coefficient \( \gamma_i \) in the given site. The term \( (c_i^\dagger c_{i+1}) \) correspond to destroying a spin in the \( (i+1) \)th position and creating one in the \( i \)-th position - thus transferring the particle from site \( (i+1) \) to site \( i \). For the single hidden node, it defines an interaction term of strength \( -t \) from the spin state \( i \) to the spin state \( i-1 \). Thus, if we include the Hamiltonian of the system as an on-site spin operator between spin levels, we must have

\[ H_{i,i-1} = -t \]  \hspace{1cm} (5.34)

Similarly, the other term gives a contribution to \( (i,i+1) \) element of the spin operator. If we allow \( t \) to be complex, hermiticity requires that this position will have a matrix element \( t^* \). This we find:

\[ H_{i,i+1} = -t^* \]  \hspace{1cm} (5.35)

Since we only allow nearest neighbour interactions, and no on-site energy, all the other terms must be zero. Keeping in mind periodic boundary conditions (PBC),
we can now describe the tight binding Hamiltonian as a spin operator acting on
a single spin, defined by:

\[
H = \begin{pmatrix}
0 & -t^* & 0 & 0 & \ldots & 0 & -t \\
-t & 0 & -t^* & 0 & 0 & \ldots & 0 \\
0 & -t & 0 & -t^* & 0 & \ldots & 0 \\
& \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
-t^* & 0 & 0 & \ldots & 0 & -t & 0
\end{pmatrix}
\] (5.36)

With this matrix, we can now set up a stochastic gradient descent algorithm
to minimize the energy \( E = \langle \psi | H | \psi \rangle \) as described by Carleo and Troyer [6].
Thus, we can solve for the optimal RBM parameters to describe this system.

Before proceeding to the RBM solution, let us note that this model has an
analytical solution. To do this, we first fourier transform to the momentum space
basis:

\[
|k\rangle = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{inka} |n\rangle 
\] (5.37)

\[
c_k^\dagger = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{inka} c_n^\dagger 
\] (5.38)

\[
c_k = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{-inka} c_n 
\] (5.39)
5.3 Tight Binding Model

In this basis, the Hamiltonian becomes diagonalized to:

\[ H = -2t \sum_k \cos(ka) \langle k \rangle \langle k | \]  \hspace{1cm} (5.40)

Periodic boundary conditions demand \( \cos(k(a + Na)) = \cos(ka) \), which force

\[ \cos(kNa) = 1 \]  \hspace{1cm} (5.41)

Choosing \( a = 1 \) for convenience, and choosing \( k \) symmetric about 0, we find the allowed values of \( k \):

\[ k \in \left[ -\pi, -\pi + \frac{2\pi}{N}, -\pi + \frac{4\pi}{N}, ... , \pi - \frac{2\pi}{N}, \pi \right] \]  \hspace{1cm} (5.42)

, and the ground state with \( k = 0 \) has energy \(-2t\). We can see that this is an eigenvalue of the Hamiltonian matrix (Equation 5.36), and that the normalized eigenvector is

\[ \psi_{\text{ground}} = \frac{1}{\sqrt{N}}[1 1 1 ... 1]^T \]  \hspace{1cm} (5.43)

Even if \( t \) is complex, the eigenvector remains the same but the eigenvalue becomes \(-2 \text{Re}(t)\).

Looking at the RBM variational form (Equation 4.11), we see that the eigenvalues (Equation 5.37) can be fit trivially by setting

\[ a = ik \quad b = 0 \quad W = 0 \]  \hspace{1cm} (5.44)

Let us see if the RBM has any other way of representing the TB eigenstates using the weights \( W \).
5.3 Tight Binding Model

5.3.1 The "Weights" W

To do this, we train the machine from multiple random initial points keeping \( a = b = 0 \) fixed. We look at the resulting distribution of W’s for different values of N, and the results are shown in Figure 5.2. Looking at Fig. 5.2 (a), (b), we find that the W’s form an interesting pattern. There are solutions clumped near W=0, but also solutions where the imaginary part of W is a fixed number and the real part can vary slightly around 0.

The W=0 solutions are trivial solutions for the tight-binding ground state. The wavefunction for each site,

\[
\psi(\sigma) = A \cosh(w\sigma)
\]

(5.45)
gives an uniform distribution A for each site independent of location. This is the \( k=0 \) state, and the normalization condition

\[
\sum_{i=1}^{N} A^2 = 1
\]

(5.46)
implies \( A = \frac{1}{\sqrt{N}} \). This fact can be seen in Figure 5.2 (c), where W=0 gives \( \psi = \frac{1}{\sqrt{20}} = 0.2236 \), with a very small cosh residual of the order of \( 10^{-5} \). The energy of the system also turns out to be correct to the order of \( 10^{-6} \).

Also very interesting is the solution with \( \text{Im}(W) \) to be a fixed non-zero value. It turns out that this value is \( 2\pi/N \) (from Figure 5.3), which is also the value of k for the first exited state. Also, the solution for the wavefunction turns out to be the first excited state, the cosine wave over the entire lattice, as seen in Figure 5.2 (d). For large N (N=50, Figure 5.2b), we even find the higher levels spaced at \( \text{Im}W = 2 \times \frac{2\pi}{N} \). Thus, the imaginary part of W captures the only relevant parameter of the problem, k, and hence is discretized in steps of \( 2\pi/N \).
5.3 Tight Binding Model

Going back to Equation 5.45, we can write the wavefunction (for \( w = x + iy \)) as:

\[
\cosh(w) = \frac{e^{iy\sigma}e^{x\sigma} - e^{-iy\sigma}e^{-x\sigma}}{2} \tag{5.47}
\]

It can be seen that the imaginary part of \( W \) plays the correct role of the momentum, but this is a superposition of two different bloch states with opposite signs of momentum. Since the energy dispersion spectra is symmetric about \( k = 0 \), this is also an eigenstate of the energy with the correct energy value, but it is not a bloch wavefunction.
5.3 Tight Binding Model

5.3.2 The ”Biases” A and B

Given that W alone is sufficient to describe the tight-binding model, which has only one parameter k, what happens if we add the other parameters, the visible and hidden. In effect, we want to look at how descriptive the RBM is and how the additional degrees of freedom can conspire together to give a consistent solution.

First, let’s study the effect of the visible bias $a$ by fixing the hidden bias $b$ to be zero. We plot a pairplot of W and A in Figure 5.4, which basically brings out all the correlations between the variables. The diagonal terms show a histogram of the distribution of the variable.

Qualitatively, we see that the $A = W = 0$ solution is present in this case as well. This is evidenced by the cluster of points near $W=0$ and $A=0$, seen in subplots (2,1) and (4, 3) respectively. Also notice that that the W have lost their discreteness property. This is because both Im($A$) and Im($W$) can represent the momentum of the state, and hence there is a lot of freedom in the choice of parameters. Subplot (3, 1) is symmetric about the $Re(W)=0$ axis, as we expect. This is because the cosh function is even, and the A parameter outside cannot distinguish between a positive and a negative W. When the real part of

![Figure 5.3: Imaginary part of W varies as $\frac{2\pi}{N}$](image)
Figure 5.4: Pair plot between weight $W$ and visible bias $A$, $N=20$
Figure 5.5: Pair plot between weight $W$ and hidden bias $B$, $N=20$
W increases, the \( \cosh \) function starts to blow up exponentially (due to one of its terms), and the \( A \) function must proportionately kill of the exponential. Thus, the lower branch of the \( \text{Re}(W) - \text{Re}(A) \) graph has a slope of \(-1\). However, the \( \text{Re}(A) \) branch where \( \text{Re}(W) = 0 \) is unphysical, and the solutions found are spurious local minima of the RBM trial function. The wavefunctions and energies evaluated from these points do not match with the theoretical results. The \( \text{Im}(W) - \text{Im}(A) \) graph has two crisscrossing straight lines of slope \( \pm 1 \). This is because the \( \cosh \) can be broken into two terms as in Equation 5.47, and the \( \text{Im}(W) \) part cancels the \( \text{Im}(A) \) part of one of the terms in the \( \cosh \). The other term remains as an error term of small magnitude. Thus, since the bias \( A \) overall has something to do with momentum, it reduces the strict restrictions on the weights \( W \), and instead compensates for the momentum that the \( W \) term does not provide.

We can also study the effect of the hidden bias \( b \) although in this case, it is not very interesting. The pair-plot is given in figure 5.5. Looking at subplots (4,2), (4, 3) and (4, 4), we find that the \( W = b = 0 \) solution is again present. Observing that the scale of \( B \) is an order of magnitude larger than the scale of \( W \), it is clear that in this case, the large value of \( B \) makes

\[
\psi(\sigma) = \cosh(W\sigma + B) \approx \cosh(B) = \text{constant}
\]

(5.48)

Since the solution to our equation is indeed very trivial, the hidden bias does not provide us any interesting insight for this problem. Of course, the solutions with \( B = 0 \), and \( W \) corresponding to the results in the previous section also pop up, as is evidenced by the peaks at the ends given by subplot (2,2) at \( \text{Im}(W) = 2\pi/N \).
5.3.3 Adding on-site disorder

Now let us consider a tight-binding model where every site contains an uncorrelated disordered on-site energy term $\epsilon_i$. Let us define:

$$\epsilon_i = u \times \epsilon$$  \hspace{1cm} (5.49)

where $\epsilon$ is a measure of the disorder strength and $u$ is a random number picked from a uniform distribution between -1 and 1.

$$u \in Unif[-1, 1]$$

Since the on-site term is diagonal in our particle representation, our Hamiltonian becomes:

$$H = \begin{pmatrix}
\epsilon_0 & -t^* & 0 & 0 & \ldots & 0 & -t \\
-t & \epsilon_1 & -t^* & 0 & \ldots & 0 \\
0 & -t & \epsilon_2 & -t^* & \ldots & 0 \\
\vdots & & & & & & \\
-t^* & 0 & 0 & \ldots & 0 & -t & \epsilon_{N-1}
\end{pmatrix}$$  \hspace{1cm} (5.50)

We solve this for different $\epsilon$ by choosing a single disorder realization $u_i, i = 0, 1, 2, \ldots, N - 1$ and then scaling epsilon. Under this model, we found that the RBM with larger number of hidden nodes performed worse than the one with fewer hidden nodes, especially for larger disorder strength $\epsilon$. This is because the number of degrees of freedom of the model increase significantly with the number of hidden nodes. This makes many more spurious local energy minima.
available, and the stochastic gradient descent can converge to any one of them. To avoid this issue, I started the RBM with trivial $\epsilon = 0$ solution, and then slowly varied $\epsilon$ in small steps and trained the RBM from the previous state (i.e. convergence point for smaller $\epsilon$). Since a small change in $\epsilon$ gives a small change in the wavefunction, the RBM can track that easily and hence avoid fitting spurious minima in a complex high-dimensional space.

Having taken care of this issue, the results obtained are given in Figure 5.6. $\alpha$ denotes the number of hidden nodes. We can think about the disorder strength as a knob to slowly increase the number of independent degrees of freedom of the system. For a system without disorder, there is only one real parameter, the momentum, controlling the system. For a system with $\epsilon \gg t$, each site acts as a separate system so we have N independent degrees of freedom. Thus, increasing $\epsilon$ slowly increases the degrees of freedom of the system. More importantly, disorder breaks the translational symmetry inherent in the problem, and makes the energy levels of the system less smooth and more erratic.

There are two key features to notice from Figure 5.6. The first thing is that for very low $\epsilon$ values, the $\alpha = 1$ does a very good job but it is quite off for larger $\epsilon$ values. The larger $\alpha$ each performs better than its lower values (except for probably a numerical error for $\alpha = 4$, $\sigma = 0.01$). Since we have 20 sites, the increase in disorder strength significantly increases the number of free parameters in the system, and we can clearly see that more hidden nodes are needed to represent the correct wavefunction more accurately. In fact, for large enough disorder strengths, we can see that the number of hidden nodes required to describe the system is roughly equal to the number of different sites, which in turn is equal to the number of complex numbers needed to describe the system.

The second feature to note is that for $\sigma > 0.5$, the solutions all seem to perform quite poorly. In fact, the higher $\alpha$ values ($\alpha = 8, 16$) seem to fare worse
5.4 Discussion

Figure 5.6: Solving the RBM trial function for the 1d Anderson model than the lower alpha values $\alpha = 1, 4$. This clearly illustrates the other problem of finding spurious minima. For higher disorder strengths, the energy landscape stops being smooth and has lots of local minima. These can be mistaken by the RBM for global minima instead. Thus, the RBMs are finding spurious energy minima instead. In the case of higher $\alpha$, the more degrees of freedom of the representation can represent/learn a larger number of local minima, and hence the energy found by the higher $\alpha$ machines is significantly worse than the lower alpha machines, and their estimated error bars (bootstrapping) significantly larger.

5.4 Discussion

We have seen that in the case of single particle physics, the RBM description does not always find the true ground state. It can often converge to a higher excited state, which is also a local minimum of the energy function. For RBMs with too many free parameters compared to the problem it is describing, the
problem tends to converge to spurious minima of the RBM system. Thus, it is important to maintain a fine balance between the number of degrees of freedom of the system, and the degrees of freedom of the RBM trying to describe it.

We have seen that for the tight-binding model, an RBM with only $W$ values gives the exact solution, whereas adding the biases $a$ and $b$ give it a little extra freedom - this gives new representations of the same solution, and we start finding spurious minima as well. Similarly, for disordered systems with broken translational invariances, the stronger the disorder, the worse the RBM seems to perform.

The RBM description often underwhelms for single particle physics, but it is exceedingly powerful for many-particle physics problems. This may be because the many degrees of freedom of a quantum many-body problem often conspire together to create a unique ground state. Moreover, larger degrees of freedom does not always create more complexity, they can simplify problems also - smoothening dispersion relations, making discrete variables into continuous variables and overall making it much simpler for a gradient descent to find the optimal solution. The examples studied above provides a cautionary tale of the problems that can arise when applying the RBM framework to a problem without properly balancing the number of degrees of freedom first.
Chapter 6

Conclusions

In this thesis, we have investigated two different ways to solve the Schrödinger equation for a single particle system. The symplectic method we have developed has been able to reproduce energy results for arbitrary bubble geometries obtained by the Maris group [2]. This method can also generalize to include tunnelling and effective potentials to calculate further corrections to the electron bubble energy and wavefunction. I intend to answer the question of whether tunneling can prevent the collapse of an electron bubble in Liquid Helium above 5 bars pressure with this method, and try to extend the analysis to multi-electron bubbles.

We have also looked at using Restricted Boltzmann machines as a trial function for quantum systems. We have found the physical interpretation of weights as representing momenta for a energy dispersion symmetric about $k = 0$. We have seen how visible and hidden biases can affect these interpretations. We have explored the limitations of the RBM representation in representing single particle states, and have come to the conclusion that the number of degrees of freedom of the RBM and of the wavefunction we are trying to describe must be closely matched to get a reasonable solution. We have gained a fleeting glimpse at how the RBM description fails when the simplification from translational symmetry is broken by disorder.
One way forward is to try to understand the meaning of the different weights and biases for other simple Hamiltonians. This would give us a further insight into both the ground state wavefunctions, and also what the parameters with which an RBM characterizes the state signify physically. Another future direction might be to add a second particle to the tight-binding model, and try to estimate the strength of the interaction needed between them to smoothen the results enough for an RBM to consistently find accurate solutions. Then, we can repeat this for more particles and more interactions. This will better quantify the sort of energy landscape required for an RBM to find an accurate solution.
References


56


