Spring 2020-21, Senior Project II – Final Report

Constructing the Phase Diagram of the 1D Bose-Hubbard Model with an Artificial Neural Network

Kadir Çeven

May 24, 2021

Abstract

This project aims to calculate the phase diagram of the 1D Bose-Hubbard model using the suggested method for quantum many-body problems in *Carleo et al.* [5], based on neural network quantum states. For this case, our artificial neural network is a restricted Boltzmann machine (RBM). Before focusing on the RBM, we discuss the Bose-Hubbard model. Then, we introduce the RBM and its ansatz. To optimize the RBM ansatz, we use a variant of the stochastic gradient descent called the Root Mean Square Propagation (RMSProp) [7] with the Metropolis-Hastings algorithm and automatic differentiation [5]. As a result, we obtain the phase diagram of the 1D Bose-Hubbard model, which is coherent with the results of DMRG [1].

Contents

1 Bose-Hubbard Model

The Bose-Hubbard model describes spinless bosons' system on an optical lattice, which can be considered repulsive interacting atoms at low temperatures. The Hamiltonian of the model can be written as follows

$$\mathcal{H} = -t \sum_{\langle i,j \rangle} \left(\hat{b}_i^\dagger \hat{b}_j + \hat{b}_j^\dagger \hat{b}_i \right) + \frac{U}{2} \sum_i \hat{n}_i (\hat{n}_i - 1) - \mu \sum_i \hat{n}_i \tag{1}$$

where t is the hopping strength, U is the term of on-site interaction strength, and μ is the chemical potential that determines the number of particles in the ground state. $\hat{b}_i^\dagger$, $\hat{b}_i$, and $\hat{n}_i$ are the creation, annihilation, and number operators of bosons for the i-th lattice site, respectively. The term of $\langle i, j \rangle$ means the nearest-neighbor sites.

By studying the model's limits, we can show that the system can exist in either the Mott insulator or superfluid phases.

1.1 Mott Insulator Phase ($t \ll U$)

Since we can set $t = 0$ for this case, the Hamiltonian becomes

$$\begin{aligned} \mathcal{H}_{\mathrm{MI}} &= \frac{U}{2} \sum_i \hat{n}_i (\hat{n}_i - 1) - \mu \sum_i \hat{n}_i \\ &= \sum_i \left(\frac{U}{2} \hat{n}_i (\hat{n}_i - 1) - \mu \hat{n}_i \right) \\ &= \sum_i \hat{h}_i \end{aligned} \tag{2}$$

where $\hat{h}_i$ is the local Hamiltonian. It is clear to see that the eigenstate of $\hat{h}_i$ is the Fock state so that

$$\begin{aligned} \hat{h}_i |n_i\rangle &= \left(\frac{U}{2} \hat{n}_i (\hat{n}_i - 1) - \mu \hat{n}_i \right) |n_i\rangle \\ &= \left(\frac{U}{2} n_i (n_i - 1) - \mu n_i \right) |n_i\rangle \\ &= E_i |n_i\rangle \end{aligned} \tag{3}$$

where E_i is the eigenvalue of $\hat{h}_i$.

To find the local occupation number $n_0(U, \mu)$ of the ground state (the perfect Mott insulator state), we first can say that

$$\begin{aligned} E_{n_0} &> E_{n_0 - 1} \\ \frac{U}{2} n_0 (n_0 - 1) - \mu n_0 &> \frac{U}{2} (n_0 - 1)(n_0 - 2) - \mu (n_0 - 1) \end{aligned} \tag{4}$$

which results in

$$\frac{\mu}{U} > n_0 - 1 \tag{5}$$

where $E_{n_0} = \min_n E_n$ is the ground state energy.

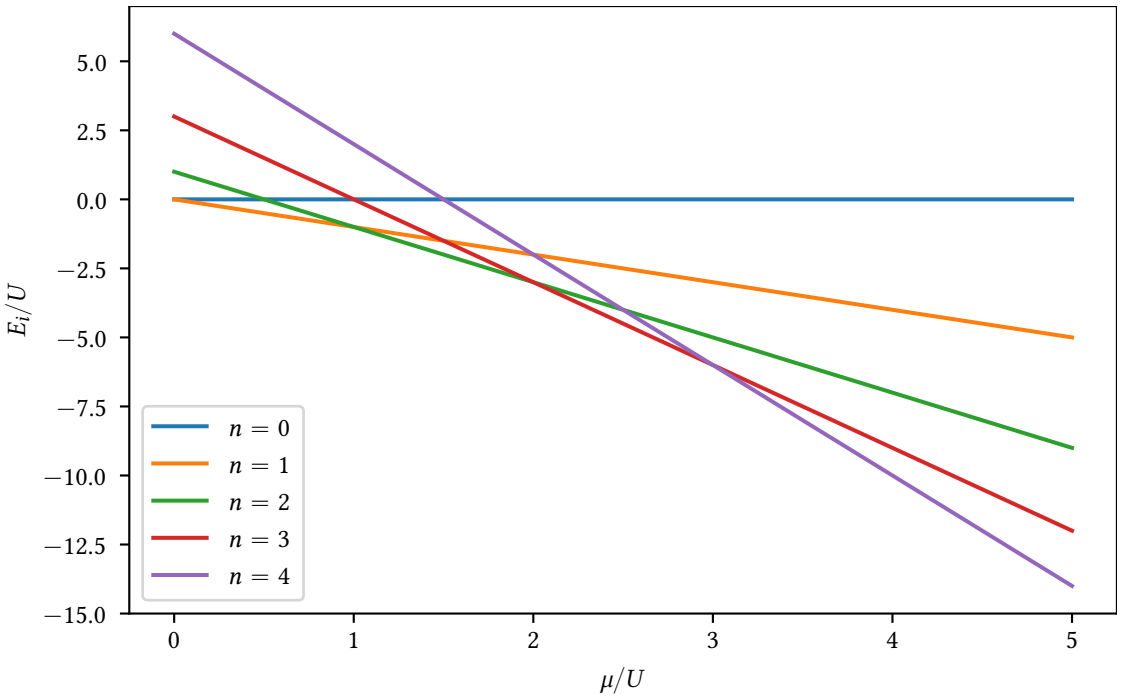


Figure 1: E_i/U vs. μ/U for the Bose-Hubbard model with $t = 0$.

Similarly, we can write the following

$$\begin{gathered} E_{n_0+1} > E_{n_0} \\ \frac{U}{2}(n_0+1)n_0 - \mu(n_0+1) > \frac{U}{2}n_0(n_0-1) - \mu n_0 \end{gathered} \tag{6}$$

which results in

$$n_0 > \frac{\mu}{U}\,. \tag{7}$$

The ground state of the system ($t = 0$) with n_0 bosonic particles exists when

$$n_0 > \frac{\mu}{U} > n_0 - 1\,. \tag{8}$$

Generally, the local occupation number $n_0(U,\mu)$ of the ground state can be written as follows

$$n_0(U,\mu) = \begin{cases} 0 & \text{if } 0 > \frac{\mu}{U} \text{ and } n = 0 \\ n & \text{if } n > \frac{\mu}{U} > n-1 \text{ and } n \in \mathbb{Z}^+ \end{cases}\,. \tag{9}$$

Therefore, this ground state has the following exact solution for the ideal case ($t = 0$) [3]

$$|\Psi_{\mathrm{MI}}\rangle = \prod_{i=1}^{N_s} \frac{1}{\sqrt{n_0(U,\mu)}} \left(\hat{b}_i^\dagger\right)^{n_0(U,\mu)} |0\rangle \tag{10}$$

where N_s is the number of lattice sites, and $|0\rangle$ is the vacuum state, which indicates that each lattice site has the same number of particles – $n_0(U,\mu)$.

1.2 Superfluid Phase ($t \gg U$)

In the case of $t \gg U$, the Hamiltonian reduces to

$$\mathcal{H}_{\mathrm{SF}} = -t \sum_{\langle i,j \rangle} \left(\hat{b}_i^\dagger \hat{b}_j + \hat{b}_j^\dagger \hat{b}_i \right) - \mu \sum_i \hat{n}_i \tag{11}$$

which can be diagonalized only in the Fourier space by finding the Fourier transform of the creation and annihilation operators as follows [2]

$$\hat{b}_i^\dagger = \frac{1}{\sqrt{N_s}} \sum_{\mathbf{k}} \exp\left(-i(\mathbf{k} \cdot \mathbf{r}_i)\right) \hat{b}_{\mathbf{k}}^\dagger \tag{12}$$

$$\hat{b}_i = \frac{1}{\sqrt{N_s}} \sum_{\mathbf{k}} \exp\left(i(\mathbf{k} \cdot \mathbf{r}_i)\right) \hat{b}_{\mathbf{k}} \tag{13}$$

where N is the number of particles in the system. After inserting these transformed operators into the transformed Hamiltonian, we get [2]

$$\mathcal{H}_{\mathrm{SF}} = \sum_{\mathbf{k}} - \left(\bar{\epsilon}_{\mathbf{k}} + \mu\right) \hat{b}_{\mathbf{k}}^\dagger \hat{b}_{\mathbf{k}} \tag{14}$$

where $\bar{\epsilon}_{\mathbf{k}} = 2t \sum_{m=1}^{d} \cos(k_m a)$ and d is the number of dimensions of the lattice.

In this limit, the ground state is a Bose-Einstein condensate in which all bosonic particles are in the state where $\mathbf{k} = 0$ [4], which can be written as

$$|\Psi_{\mathrm{SF}}\rangle = \frac{1}{\sqrt{N!}} \left(\hat{b}_{\mathbf{k}=0}\right)^N |0\rangle = \frac{1}{\sqrt{N!}} \left(\frac{1}{\sqrt{N_s}} \sum_{i=1}^{N_s} \hat{b}_i^\dagger \right)^N |0\rangle \,. \tag{15}$$

1.3 Mean-Field Theory

We can trivially study the 1D Bose-Hubbard model within the mean-field theory.

To do so, we define a 1D lattice. Then, the Hamiltonian for the m-th lattice site becomes

$$\mathcal{H}_{\mathrm{m}} = -t \left(\hat{b}_m^\dagger \hat{b}_{m-1} + \hat{b}_{m-1}^\dagger \hat{b}_m + \hat{b}_m^\dagger \hat{b}_{m+1} + \hat{b}_{m+1}^\dagger \hat{b}_m \right) + \frac{U}{2} \hat{n}_m \left(\hat{n}_m - 1\right) - \mu \hat{n}_m \,. \tag{16}$$

We can simplify this Hamiltonian by setting that

$$\alpha \equiv \hat{b}_{m-1}^\dagger \equiv \hat{b}_{m-1} \equiv \hat{b}_{m+1}^\dagger \equiv \hat{b}_{m+1} \tag{17}$$

which is defined as the mean-field parameter so that the mean-field Hamiltonian $\mathcal{H}_{\mathrm{m}}^{\mathrm{MF}}$ turns out to be

$$\mathcal{H}_{\mathrm{m}}^{\mathrm{MF}} = -2\alpha t \left(\hat{b}_m^\dagger + \hat{b}_m \right) + \frac{U}{2} \hat{n}_m \left(\hat{n}_m - 1\right) - \mu \hat{n}_m \,. \tag{18}$$

To numerically calculate and plot the phase diagram for the 1D mean-field Bose-Hubbard Hamiltonian $\mathcal{H}_{\mathrm{m}}^{\mathrm{MF}}$, we first need to write the creation and annihilation operators as matrices. For bosonic Fock states, these operators can be represented as the following matrices

$$\hat{b}_m^\dagger = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \sqrt{1} & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \sqrt{3} & \ddots & \vdots & 0 & 0 \\ \vdots & \vdots & \ddots & \sqrt{4} & \ddots & 0 & 0 \\ 0 & 0 & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \sqrt{N} & 0 \end{bmatrix} \tag{19}$$

$$\hat{b}_m = \begin{bmatrix} 0 & \sqrt{1} & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \sqrt{3} & \ddots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \sqrt{4} & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 & \sqrt{N} \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}. \tag{20}$$

Secondly, we need to optimize the mean-field parameter α. Ideally, α must be equal to $\langle \hat{b} \rangle$. Thus, with a self-consistency procedure, the optimum value of the mean-field parameter α_{opt} can be found as in Algorithm 1.

Algorithm 1: The Self-Consistency Procedure to Find α_{opt}

$\alpha_{\text{opt}}^{\text{pre}} \leftarrow 0$
$\alpha_{\text{opt}} \leftarrow 1$
while not $|\alpha_{\text{opt}} - \alpha_{\text{opt}}^{\text{pre}}| < \epsilon$ **do**
  find the normalized ground state $|\Psi_{\text{GS}}\rangle$ for α_{opt} via the NumPy's function of `linalg.eigh`
  $\alpha_{\text{opt}}^{\text{pre}} \leftarrow \alpha_{\text{opt}}$
  $\alpha_{\text{opt}} \leftarrow \hat{b}\Psi_{\text{GS}}$
end while

After making the mean-field Hamiltonian $\hat{H}_{\text{m}}^{\text{MF}}$ dimensionless as follows

$$\frac{\mathcal{H}_{\text{m}}^{\text{MF}}}{U} = -2\alpha\frac{t}{U}\left(\hat{b}_m^\dagger + \hat{b}_m\right) + \frac{1}{2}\hat{n}_m\left(\hat{n}_m - 1\right) - \frac{\mu}{U}\hat{n}_m, \tag{21}$$

we can construct the mean-field phase diagram for the 1D Bose-Hubbard model as in Figure 2.

2 Neural Network Quantum States

Carleo et al. [5] has shown that the quantum many-body problems such as the Heisenberg model and Ising model can be solved by using an artificial neural network (ANN), so-called a restricted Boltzmann machine (RBM) with a reinforcement-learning scheme. Using this method, *McBrian et al.* [6] and *Vargas-Calderón et al.* [8] found the ground state of the 1D Bose-Hubbard model and constructed its phase diagram.

2.1 Restricted Boltzmann Machine Ansatz for Wavefunction

In our case, we consider the Fock basis; therefore, the ground state of the 1D Bose-Hubbard model can be written as follows [8]

$$|\Psi_{\text{GS}}\rangle = \sum_{\mathcal{S}} \Psi(\mathcal{S})\,|n(\mathcal{S})\rangle \tag{22}$$

where $\mathcal{S} = (n_1, \cdots, n_{N_s})$ is the occupation number configuration, n_i is the bosonic occupation number of the i-th lattice site, and $|n(\mathcal{S})\rangle = |n_1, \cdots, n_{N_s}\rangle$ is the Fock state concerning the

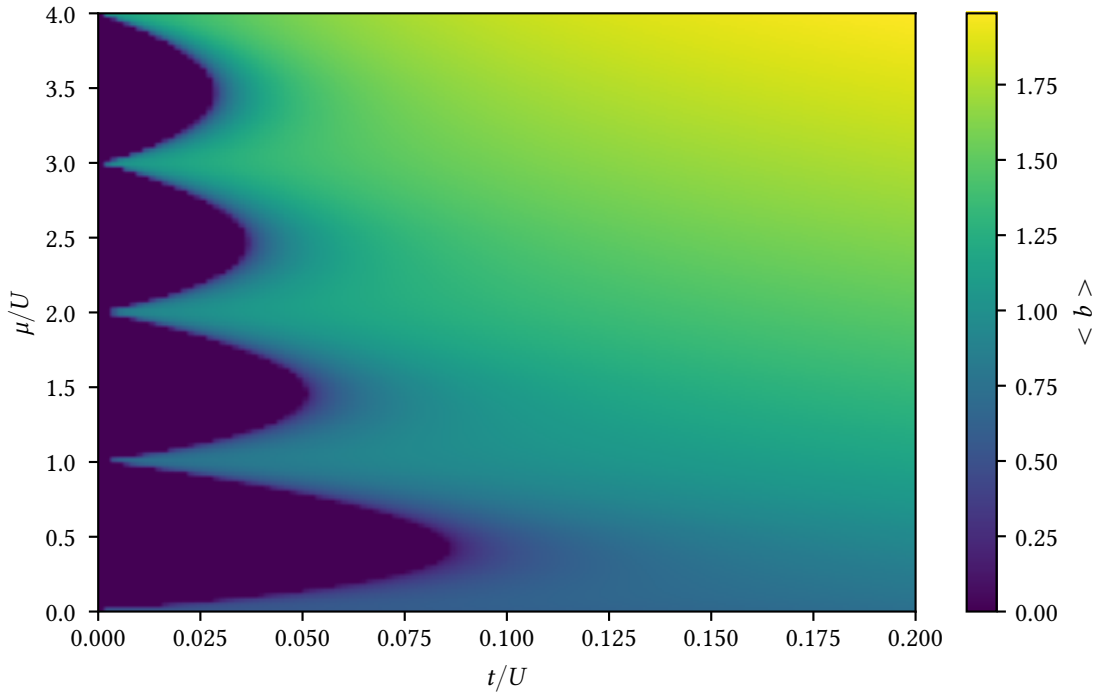


Figure 2: Mean-field phase diagram of the 1D Bose-Hubbard model.

configuration $\mathcal{S}$. To implement $\mathcal{S}$ into our RBM, we subdivide $\mathcal{S}$ into another configuration called the one-hot encoding-like configuration $\mathcal{S}_{\text{OHEL}}$. This configuration has no difference from $\mathcal{S}$ rather than having extra neurons $n_{i,j}$ (or v_k) to represent each occupation number. These extra neurons are limited by $n_{\max}$ – the maximum number of particles in each lattice site.

Figure 3: Restricted Boltzmann machine used in our case.

The wavefunction of $\Psi(\mathcal{S})$ can be approximated as the RBM ansatz as follows [5]

$$\Psi(\mathcal{S}) \approx \Psi(\mathcal{S}_{\text{OHEL}}; \mathcal{W}) = \sum_{\{h_k=\pm 1\}} \exp\left(\sum_{j=1}^{N_s \times n_{\max}} \alpha_j v_j + \sum_{i=1}^{M} \beta_i h_k + \sum_{i=1}^{M} \sum_{j=1}^{N_s \times n_{\max}} W_{ij} h_k v_j \right) \tag{23}$$

where $\mathcal{W} = \{\alpha_j, \beta_i, W_{ij}\}$ is the set of the parameters of the RBM, h_k is the hidden variable, α_j and β_i are the local weights for the visible and hidden layers, W_{ij} is the interaction weight matrix, and M is the number of hidden neurons.

By definition, the wavefunction of $\Psi(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})$ can also be written as

$$\Psi(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W}) = \exp\left(\sum_{j=1}^{N_s\times n_{\max}} \alpha_j v_j\right) \prod_{i=1}^{M} 2\cosh\left[\theta_i(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})\right] \tag{24}$$

where $\theta_i(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})$ is the effective angle

$$\theta_i(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W}) = \beta_i + \sum_{j=1}^{N_s\times n_{\max}} W_{ij} v_j\,. \tag{25}$$

2.2 Ground State Energy

To calculate the ground state energy $E(\mathcal{W})$, we use the variational Monte Carlo (VMC) as follows

$$\begin{aligned}
E(\mathcal{W}) = E = \frac{\Psi_{\mathrm{GS}}\mathcal{H}\Psi_{\mathrm{GS}}}{\langle\Psi_{\mathrm{GS}}|\Psi_{\mathrm{GS}}\rangle} &= \frac{\int \Psi^*(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})\,\mathcal{H}\,\Psi(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})\,\mathrm{d}\mathcal{S}_{\mathrm{OHEL}}}{\int |\Psi(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})|^2\,\mathrm{d}\mathcal{S}_{\mathrm{OHEL}}} \\
&= \frac{\int \Psi^*(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})\,\frac{\Psi(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})}{\Psi(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})}\,\mathcal{H}\,\Psi(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})\,\mathrm{d}\mathcal{S}_{\mathrm{OHEL}}}{\int |\Psi(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})|^2\,\mathrm{d}\mathcal{S}_{\mathrm{OHEL}}} \\
&= \frac{\int |\Psi(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})|^2\,\frac{\mathcal{H}\,\Psi(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})}{\Psi(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})}\,\mathrm{d}\mathcal{S}_{\mathrm{OHEL}}}{\int |\Psi(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})|^2\,\mathrm{d}\mathcal{S}_{\mathrm{OHEL}}} \\
&= \frac{\int |\Psi(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})|^2\,E_{\mathrm{loc}}(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})\,\mathrm{d}\mathcal{S}_{\mathrm{OHEL}}}{\int |\Psi(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})|^2\,\mathrm{d}\mathcal{S}_{\mathrm{OHEL}}} \\
&= \int \rho(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})\,E_{\mathrm{loc}}(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})\,\mathrm{d}\mathcal{S}_{\mathrm{OHEL}}
\end{aligned} \tag{26}$$

where $\mathcal{H}$ is the 1D Bose-Hubbard Hamiltonian, $E_{\mathrm{loc}}(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W}) = \frac{\mathcal{H}\,\Psi(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})}{\Psi(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})}$ is the so-called local energy, and $\rho(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W}) = \frac{|\Psi(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})|^2}{\int |\Psi(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})|^2\,\mathrm{d}\mathcal{S}_{\mathrm{OHEL}}}$ is the probability density.

By sampling $|\Psi(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})|^2$, we can approximate $E(\mathcal{W})$ statistically as follows

$$E(\mathcal{W}) \approx E_{\mathrm{locMH}} = \frac{1}{P}\sum_{i=1}^{P} E_{\mathrm{loc}}(\mathcal{S}^{(i)};\mathcal{W}) \tag{27}$$

where $\cdots_{\mathrm{MH}}$ is the notation of the average over the sampling of $|\Psi(\mathcal{S}_{\mathrm{OHEL}};\mathcal{W})|^2$ with the Metropolis-Hastings algorithm, P is the number of samples, and $\mathcal{S}^{(i)}$ is the one-hot encoding-like configuration of the i-th sample.

2.3 Metropolis-Hastings Algorithm

To carry out this sampling, we use the Metropolis-Hastings algorithm as in *Carleo et al.* [5]. It is a Markov Chain Monte Carlo algorithm to draw random samples from a probability distribution iteratively. Direct sampling is laborious so that the formed Markov chain is expected to converge to the chosen probability distribution without depending on the first sample.

We generate a Markov chain of the one-hot encoding-like configurations $\mathcal{S}^{(1)} \to \mathcal{S}^{(2)} \to \cdots \to \mathcal{S}^{(P)}$ as explained in Algorithm 2.

Algorithm 2: The Metropolis-Hastings Algorithm for the 1D Bose-Hubbard Model

randomly initialize $\mathcal{S}^{(1)}$
for $i \in \{2, 3, \cdots, P\}$ **do**
  $\mathcal{S}' \leftarrow \mathcal{S}^{(i-1)}$
  generate a random integer $m \in [1, N_s]$
  $n_m \leftarrow$ the occupation number of the m-th lattice site in $\mathcal{S}^{(i-1)}$
  generate a random integer $n' \in [0, n_{\max}] \setminus \{n_m\}$
  set the occupation number of the m-th lattice site in $\mathcal{S}'$ to n'
  $R \leftarrow U(0, 1)$
  if $R < \frac{\Psi(\mathcal{S}';\mathcal{W})}{\Psi(\mathcal{S}^{(i-1)};\mathcal{W})}^2$ **then**
    $\mathcal{S}^{(i)} \leftarrow \mathcal{S}'$
  else
    $\mathcal{S}^{(i)} \leftarrow \mathcal{S}^{(i-1)}$
  end if
end for

2.4 Optimization of RBM Ansatz

Since the parameters of RBM are randomly initialized, we need to optimize our RBM ansatz to obtain the ground state wavefunction. It is also known that the ground state is the one that has the lowest energy in a given system, so the cost function that will be minimized is the expectation value of energy.

Carleo et al. [5] state that this optimization can be achieved with stochastic gradient descent. For each iteration k, we first do a sampling of $|\Psi(\mathcal{S}; \mathcal{W}_k)|^2$ with the Metropolis-Hastings algorithm and calculate the local energy for each sample in the Markov chain. Next, we determine $\mathcal{W}_{k+1}$ by calculating the variational derivatives $\mathcal{O}_{k,p}(\mathcal{S}; \mathcal{W}_k)$ with respect to the p-th parameter of $\mathcal{W}_k$ of $\Psi(\mathcal{S}; \mathcal{W}_{k,p})$

$$\mathcal{O}_{k,p}(\mathcal{S}; \mathcal{W}_k) = \frac{\partial}{\partial \mathcal{W}_{k,p}} \ln\left[\Psi(\mathcal{S}; \mathcal{W}_k)\right] = \frac{1}{\Psi(\mathcal{S}; \mathcal{W}_k)} \frac{\partial \Psi(\mathcal{S}; \mathcal{W}_k)}{\partial \mathcal{W}_{k,p}} . \tag{28}$$

For the network parameters, the derivatives are

$$\mathcal{O}_{k,\alpha_j}(\mathcal{S}; \mathcal{W}_k) = \frac{1}{\Psi(\mathcal{S}; \mathcal{W}_k)} \frac{\partial \Psi(\mathcal{S}; \mathcal{W}_k)}{\partial \alpha_j} = v_j \tag{29}$$

$$\mathcal{O}_{k,\beta_i}(\mathcal{S}; \mathcal{W}_k) = \frac{1}{\Psi(\mathcal{S}; \mathcal{W}_k)} \frac{\partial \Psi(\mathcal{S}; \mathcal{W}_k)}{\partial \beta_i} = \tanh\left[\theta_i(\mathcal{S}; \mathcal{W}_k)\right] \tag{30}$$

$$\mathcal{O}_{k,W_{ij}}(\mathcal{S}; \mathcal{W}_k) = \frac{1}{\Psi(\mathcal{S}; \mathcal{W}_k)} \frac{\partial \Psi(\mathcal{S}; \mathcal{W}_k)}{\partial W_{ij}} = v_j \tanh\left[\theta_i(\mathcal{S}; \mathcal{W}_k)\right] . \tag{31}$$

Thus, the gradient of the energy with respect to the p-th parameter can be approximated

as follows [5]

$$\begin{aligned}\frac{\partial E(\mathcal{W}_k)}{\partial \mathcal{W}_{k,p}} \approx G_p(\mathcal{W}_k) &= E_{\text{loc}}\, \mathcal{O}^*_{k,p_{\text{MH}}} - E_{\text{locMH}}\, \mathcal{O}^*_{k,p_{\text{MH}}} \\ &= \frac{1}{P}\sum_{i=1}^{P} E_{\text{loc}}(\mathcal{S}^{(i)};\mathcal{W}_k)\, \mathcal{O}^*_{k,p}(\mathcal{S}^{(i)};\mathcal{W}_k) - E(\mathcal{W}_k)\frac{1}{P}\sum_{i=1}^{P} \mathcal{O}^*_{k,p}(\mathcal{S}^{(i)};\mathcal{W}_k) \\ &= \frac{1}{P}\sum_{i=1}^{P} \left[E_{\text{loc}}(\mathcal{S}^{(i)};\mathcal{W}_k) - E(\mathcal{W}_k)\right] \mathcal{O}^*_{k,p}(\mathcal{S}^{(i)};\mathcal{W}_k)\end{aligned} \tag{32}$$

so $\mathcal{W}_{k+1,p}$ turns out to be

$$\mathcal{W}_{k+1,p} = \mathcal{W}_{k,p} - \eta\, G_p(\mathcal{W}_k) \tag{33}$$

with a fixed learning rate η.

To optimize the RBM ansatz efficiently, we can apply the Root Mean Square Propagation (RMSProp) [7] to obtain an adaptive learning rate $\eta(k,p)$

$$\eta(k,p) = \frac{\eta}{\sqrt{s_p(\mathcal{W}_k)} + \epsilon} \tag{34}$$

where ϵ is a small cutoff value, and

$$s_p(\mathcal{W}_k) = \gamma s_p(\mathcal{W}_{k-1}) + (1-\gamma) G_p(\mathcal{W}_{k-1})^2 \tag{35}$$

with the exponential decay rate γ so that

$$\mathcal{W}_{k+1,p} = \mathcal{W}_{k,p} - \eta(k,p)\, G_p(\mathcal{W}_k)\,. \tag{36}$$

2.5 Results

We choose $N_s = 4$, $n_{\max} = 4$, $M = 6$, $P = 1500$, $\eta = 0.01$, $\gamma = 0.9$, and $\epsilon = 10^{-8}$.

2.5.1 Ground State Energy

As seen in Figure 4, the variational ground state energy with the RBM converges to its exact value after a while.

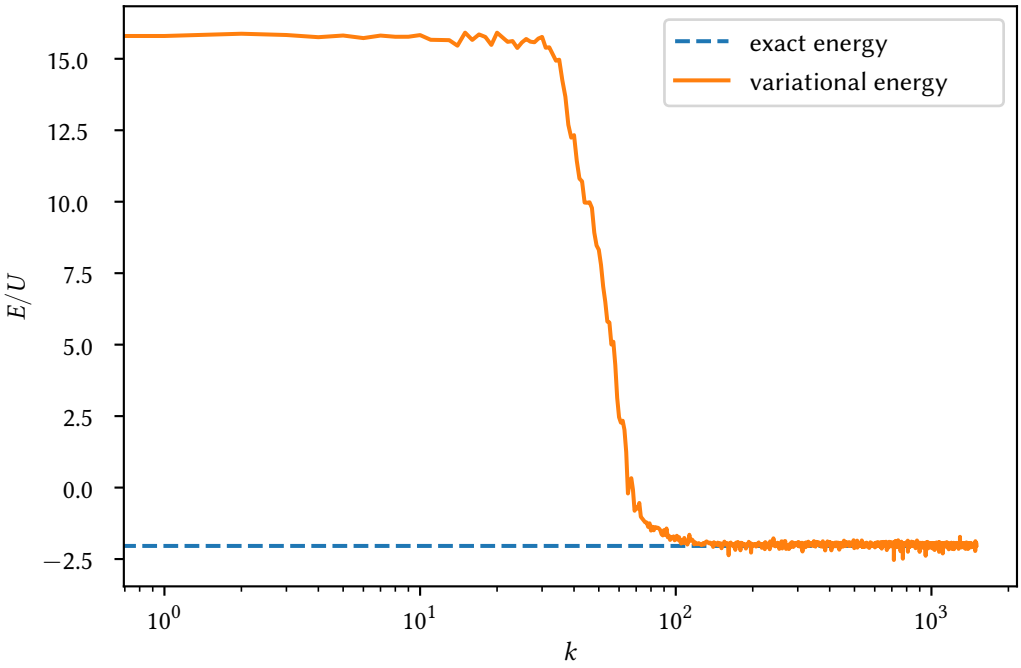
2.5.2 Phase Diagram

To plot the phase diagram, we need to choose an order parameter to distinguish the phase transitions. The variance of the on-site number operator $\text{Var}(\hat{n}_i)$ is appropriate for this because it is expected to be zero for the Mott insulator phase and greater than zero for the superfluid phase.

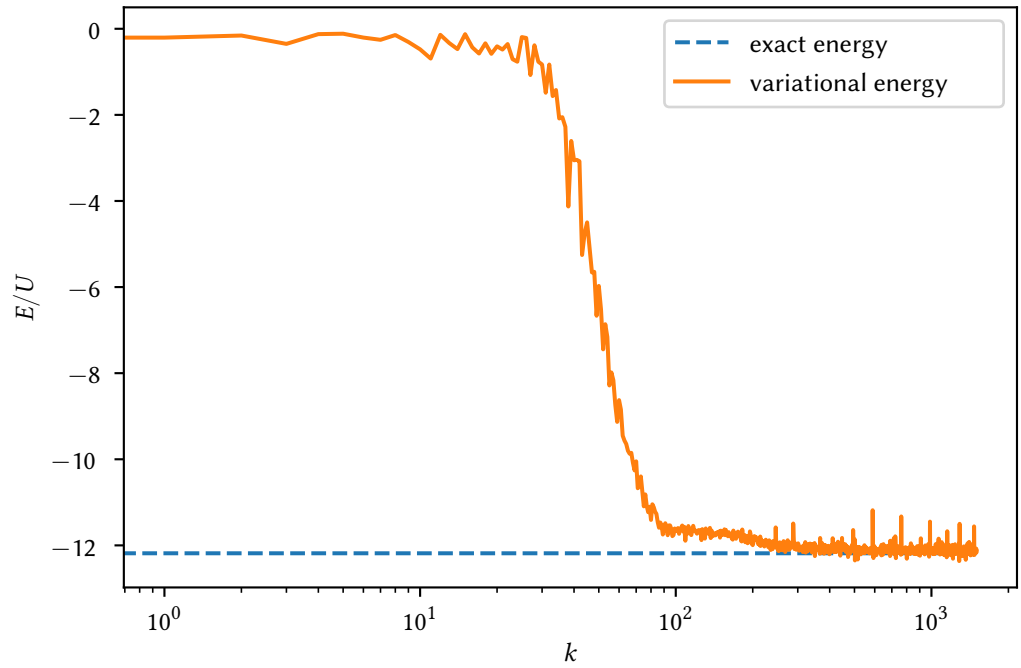
Once again, we calculate this value by using the samples produced by the Metropolis-Hastings algorithm as follows

$$\text{Var}(\hat{n}_i) \approx \frac{\sum_{j=1}^{P} n_i^2(\mathcal{S}^{(j)})\, \Psi(\mathcal{S}^{(j)};\mathcal{W})^2 - \left[\sum_{j=1}^{P} n_i(\mathcal{S}^{(j)})\, \Psi(\mathcal{S}^{(j)};\mathcal{W})^2\right]^2}{\sum_{j=1}^{P} \Psi(\mathcal{S}^{(j)};\mathcal{W})^2}\,. \tag{37}$$

We plot a grid of the averages of $\text{Var}(\hat{n}_1)$ resulted in the last 300 sampling steps by using 1500 samples for each iteration. The size of the grid is 33×33. The produced plot is in Figure 5.



(a) $t/U = 0.05$ & $\mu/U = 0.5$



(b) $t/U = 0.35$ & $\mu/U = 1.5$

Figure 4: Variational energy of the 1D Bose-Hubbard model as a function of the iteration k.

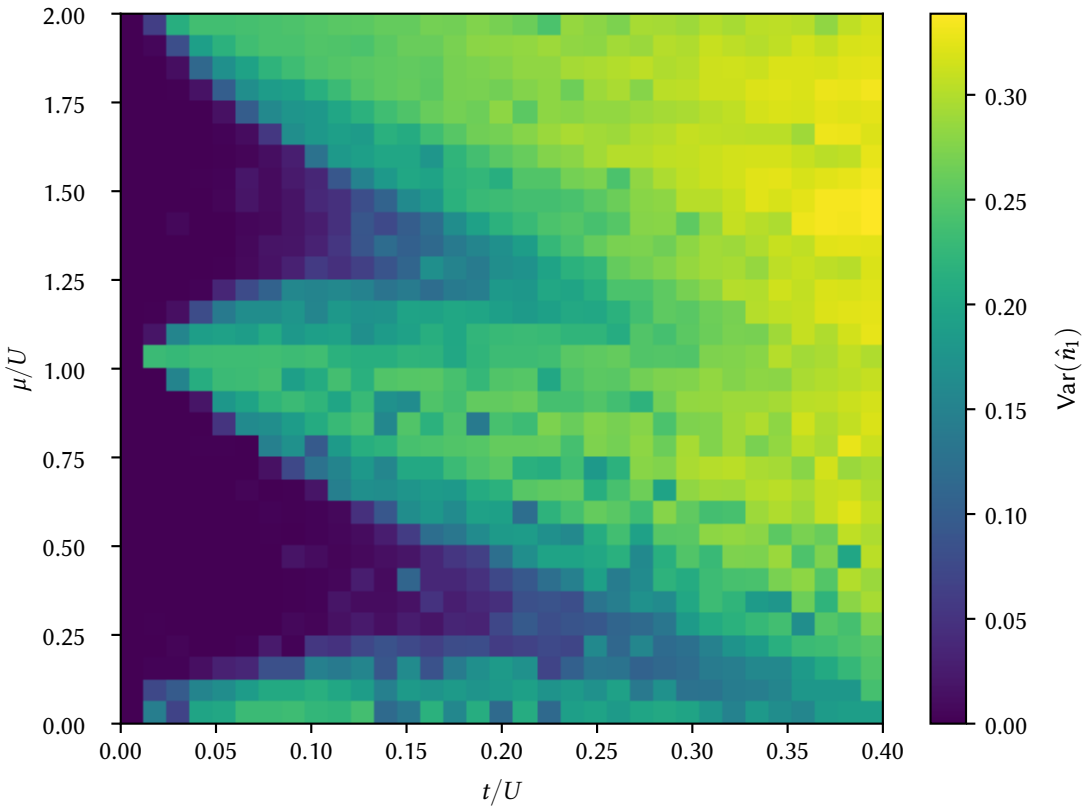


Figure 5: Phase diagram of the 1D Bose-Hubbard model, calculated by the given RBM in Section 2.5.

3 Conclusions

In this project, we successfully tested whether a restricted Boltzmann machine can be trained to determine the ground state wavefunction of the 1D Bose-Hubbard model. Even though some noises appear in the phase diagram (see Figure 5) due to the low-resolution grid and small sample size, we can calculate it, which has better results than the mean-field theory. Thus, it is more consistent with the numerical data from DMRG [1].

Firstly, as we say in Section 2.1, we prefer to use the one-hot encoding-like configuration rather than the typical occupation number configuration since we have more fluctuations in the variational energy convergence with the typical one. Those fluctuations may occur because of the small number of neurons in the visible layer of the RBM. Even if we try to increase the number of neurons in the hidden layer, they keep appearing.

Secondly, the Metropolis-Hastings algorithm provides us a sample from the Fock space instead of the whole Fock space to carry out all calculations without dealing with many Fock states. In our case, it is not so valid due to time restrictions; however, when N_s and $n_{\max}$ increase, the sampling turns out to be more significant.

Finally, optimizing the RBM ansatz is one of the crucial points in this whole process. Thanks to the automatic differentiation, we can efficiently compute the gradients. Besides, the RMSProp allows us to update the learning rate so that the ground state energy converges faster. Instead of the RMSProp, *Carleo et al.* [5] use the stochastic reconfiguration method.

For future work, one can determine the phase boundaries in Figure 5 differently, such as applying the neural network quantum states to the canonical ensemble of the 1D Bose-Hubbard model.

References

[1] T. D. Kühner and H. Monien, Physical Review B **58**, R14741 (1998).

[2] D. van Oosten, P. van der Straten, and H. T. C. Stoof, Phys. Rev. A **63**, 053601 (2001).

[3] F. Gerbier, A. Widera, S. Fölling, O. Mandel, T. Gericke, and I. Bloch, Phys. Rev. Lett. **95**, 050404 (2005).

[4] I. Bloch, J. Dalibard, and W. Zwerger, Rev. Mod. Phys. **80**, 885 (2008).

[5] G. Carleo and M. Troyer, Science **355**, 602 (2017).

[6] K. McBrian, G. Carleo, and E. Khatami, Journal of Physics: Conference Series **1290**, 012005 (2019).

[7] C.-Y. Park and M. J. Kastoryano, Phys. Rev. Research **2**, 023232 (2020).

[8] V. Vargas-Calderón, H. Vinck-Posada, and F. A. González, Journal of the Physical Society of Japan **89**, 094002 (2020).