Ground state phase diagram of the one-dimensional Bose-Hubbard model from restricted Boltzmann machines

Kristopher McBrien
Department of Physics and Astronomy, San Jose State University, San Jose, CA 95192, USA

Giuseppe Carleo
Center for Computational Quantum Physics, Flatiron Institute, New York, NY 10010, USA

Ehsan Khatami
Department of Physics and Astronomy, San Jose State University, San Jose, CA 95192, USA

Abstract. Motivated by recent advances in the representation of ground state wavefunctions of quantum many-body systems using restricted Boltzmann machines as variational ansatz, we utilize an open-source platform for constructing such ansatz called NetKet to explore the extent of applicability of restricted Boltzmann machines to bosonic lattice models. Within NetKet, we design and train these machines for the one-dimensional Bose-Hubbard model through a Monte Carlo sampling of the Fock space. We vary parameters such as the strength of the onsite repulsion, the chemical potential, the system size and the maximum site occupancy and use converged equations of state to identify phase boundaries between the Mott insulating and superfluid phases. We compare the average density and the energy to results from exact diagonalization and map out the ground state phase diagram, which agrees qualitatively with previous finding obtained through conventional means.

1. Introduction
In recent years, artificial neural networks have been employed to study the Bose-Hubbard model [1, 2], which describes the many-body systems of interacting bosonic particles on lattices [3]. In Ref. [1], the author approaches the problem using feedforward artificial neural networks where the inputs are Fock states (see below) and the output layer, consisting of two nodes, provides the Re and Im parts of the log of the wavefunction. The networks are trained following conventional supervised learning algorithms for feedforward networks with the cost function taken to be the ground state energy. Ref. [2] expands on this idea and examines the role multiple hidden layers or convolutional layers play in the accuracy of the output and the efficiency of the algorithm.

Restricted Boltzmann machines (RBMs) on the other hand have emerged as useful tools among artificial neural networks for providing a superior ansatz for representing quantum wave functions [4, 5, 6, 7, 8], or mimicking thermodynamics of classical and quantum many-body systems [9, 10, 11]. For example, in Ref. [4], it was shown that RBMs lead to lower
variational ground state energy for classical and quantum magnetic models than the state-of-the-art ansatz. This in turn motivated several other studies in which the RBMs were used to represent topological [5, 6] or chiral [8] states, draw equivalences between RBMs and tensor network states [7], and to accelerate Monte Carlo simulations [12], to name a few.

Here, we use RBMs to represent ground state wavefunctions of the one-dimensional Bose-Hubbard model and study the extent to which its equilibrium properties can be captured in this architecture for a range of model parameters. It has been shown that multi-valued RBMs, such as the one we have used here, can represent a wide class of many-body entangled states efficiently [13], and to the best of our knowledge, our results are the first to demonstrate that. We perform the training using Monte Carlo sampling with the same goal of minimizing the ground state energy as in Refs. [1, 2]. We employ NetKet [14], an open-source platform for solving quantum many-body problems using artificial intelligence. We show that RBMs can be trained to represent the ground state of the bosonic system with a high degree of accuracy. We find that the energy and density both in the Mott insulating and superfluid phases rapidly converge to final values during the training. By calculating equations of state for different values of the interaction strength in the grand canonical ensemble, we map out the phase diagram of the model and show that it agrees well with one obtained in the past directly from quantum Monte Carlo simulations.

2. The Bose-Hubbard Model

The Bose-Hubbard Hamiltonian is expressed as

$$H = -t \sum_{\langle ij \rangle} b_i^\dagger b_j + \frac{U}{2} \sum_i n_i(n_i - 1) - \mu \sum_i n_i,$$

(1)

where $b_i^\dagger$ and $b_j$ are the bosonic creation and annihilation operators, respectively, $\langle ij \rangle$ indicates that $i$ and $j$ are nearest neighbors, $t$ is the corresponding hopping amplitude, which we set to unity to serve as the unit of energy throughout the paper, $U$ is the strength of the onsite repulsion, $n_i = b_i^\dagger b_i$ is the site occupation operator, and $\mu$ is the chemical potential. In this study, we consider the periodic one-dimensional (1D) geometry.

For large enough interactions, the ground state of the system is the interaction-driven Mott insulating phase while it remains a superfluid for smaller interactions. This was shown in a seminal paper by Fisher et al. [3] where the universality class of the model was explored and a mean-field solution was presented. Later, the model received much attention in the 1990’s, mostly confirming numerically the transition in different dimensions [15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25]. The Mott insulator (MI) to superfluid (SF) transition was also observed in experiments with cold atoms in optical lattices [26].

3. Restricted Boltzmann Machine

3.1. The Architecture

The RBM is a two-layer fully-connected artificial neural network. The architecture is shown in Fig. 1. It has one visible layer, and one hidden layer. Each layer consists of several nodes with an external bias parameter associated with them ($b_i$ for the visible and $c_i$ for the hidden layer). Each node in a layer is connected to all the nodes in the other layer and a weight parameter ($W_{ij}$) is associated to every connection.

The RBM lends itself well to representing ground state wavefunctions of the Bose-Hubbard model because a probability distribution of possible Fock states of the system can then be calculated directly from the RBM parameters after training. This is similar to representing thermodynamics of a many-body system using an RBM where a probability distribution of microstates can be extracted [10].
Figure 1. Sample architecture of the RBM used in this study. We choose visible layers of different sizes for \( L = 4, 5, 6, \) and 7, which also represent the system sizes and twice as many nodes in the hidden layer as in the input layer. Both the input nodes and hidden nodes take binary values; \( \sigma_i \) and \( h_i \), respectively. Every site occupation \( n_j = \{0, 1, ..., N_{\text{max}}\} \) is converted to \( N_{\text{max}} + 1 \) binary digits, all 0’s except the \( n_j \)th digit, which is 1. Therefore, the number of input nodes is \( (N_{\text{max}} + 1) \times L \). Each node has a bias value (shown as \( b_i \) or \( c_i \)) and nodes are fully connected between the two layers with weights \( w_{ij} \) associated with each connection. The weights and biases are determined through a stochastic optimization process that involves local Metropolis moves between Fock states where the expectation value of the Bose-Hubbard Hamiltonian is minimized.

Within NetKet, we define a wave function ansatz for the \( q = 0 \) momentum sector using a RBM with spin-1/2 hidden nodes and permutation symmetry. The input layer can be visualized as a two-dimensional \( L \times (N_{\text{max}} + 1) \) array of binary numbers, where each row represents a site, and a 1 in column number \( n \) indicates that there are \( n \) bosons at that site. This so-called “one-hot” encoding of site occupations is not unique. For example, an alternative would be to use \( L \) non-binary input nodes whose values can vary from 0 to \( N_{\text{max}} \) instead. One advantage of the one-hot representation over the latter is that the number of network parameters is much larger and therefore, the training will presumably be more refined. We choose \( L \) and \( N_{\text{max}} \) values as large as 7 and 5, respectively, and a hidden layer with twice as many nodes as in the visible layer.

3.2. The Training

We write the ground state wave function as a linear combination of the Fock state basis

\[
|\psi\rangle = \sum_n \psi_n |n\rangle ,
\]

where \( |n\rangle = |n_1, n_2, ..., n_L\rangle \) are the particle number Fock states with site occupancies \( n_i \in [0, N_{\text{max}}] \). The goal is to create a RBM that can stochastically represent \( |\psi\rangle \) by producing the correct probability, \( |\psi_n|^2 \) for every input Fock state \( |n\rangle \) through the input layer. For the Bose-Hubbard model, the coefficients \( \psi_n \) are known to be real and positive and are written for the RBM as

\[
\psi_n = \sum_{\{h\}} e^{\sum_i b_i \sigma_i + \sum_j c_j h_j + \sum_{ij} W_{ij} \sigma_i h_j}.
\]

where \( \sigma_i \) and \( h_i \) denote the input node and hidden node binary values, respectively.

The hidden degrees of freedom can be integrated out exactly for the simple architecture.
shown in Fig. 1,

\[
\sum_{\{h\}} e^{\sum_i b_i \sigma_i + \sum_j c_j h_j + \sum_{ij} W_{ij} \sigma_i h_j} = e^{\sum_i b_i \sigma_i + \sum_j e^{c_j h_j + \sum_i W_{ij} \sigma_i h_j}} = e^{\sum_i b_i \sigma_i} \prod_j e^{c_j h_j + \sum_i W_{ij} \sigma_i h_j} = e^{\sum_i b_i \sigma_i} \prod_j e^{\log(1 + e^{c_j h_j + \sum_i W_{ij} \sigma_i h_j})},
\]

which results in

\[
\psi_n = e^{\sum_i b_i \sigma_i + \sum_j \log(1 + e^{c_j h_j + \sum_i W_{ij} \sigma_i h_j})}.
\]

The goal in the training of the RBM is to optimize its weights and biases in order to arrive at a probability distribution for \(\psi_n\) so that \(|\psi\rangle\) represents the ground state. This is accomplished through the minimization of the expectation value of the Hamiltonian,

\[
E = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle}.
\]

NetKet adopts a Metropolis algorithm for importance sampling of \(|n\rangle\)'s during which \(E\) is minimized also using stochastic techniques. Since we are working in the grand canonical ensemble, we choose local moves in which a site is picked at random and its occupation is proposed to change to a value in the range \([0, N_{\text{max}}]\). The move is accepted with the probability \(\min[1, |\psi_{n'}|^2/|\psi_n|^2]\), where \(n'\) and \(n\) are the new and old states, respectively. This guarantees that after the training is completed, the probability distribution is \(P_n = |\psi_{n'}|^2/\sum_m |\psi_m|^2 = |\psi_n|^2/\langle \psi | \psi \rangle\). The minimization of energy is done using the Stochastic Reconfiguration algorithm [27] with AdaMax optimizer in NetKet and the default parameters. Our sampler performs 1000 sweeps before sampling the probability distribution and updating the RBM parameters. More details about the method and the training parameters can be found in Ref. [14].

4. Results

In Figs. 2 and 3, we show the average energy, \(E\), and the average density, \(\rho\), as the training of our RBM progresses for two sets of values for the interaction strength and the chemical potential. The system with \(\mu/U = 0.56\) (blue lines) is a Mott insulator with a density of one particle per site and the system with \(\mu/U = 1.00\) (orange lines) is a superfluid for the system size we have considered (\(L = 6\)). Here, \(N_{\text{max}} = 5\) is much larger than the average density. We observe that the energies and densities quickly converge to their final values regardless of whether the system is in the MI or SF phase. However, we find that the average density has larger fluctuations over the training steps inside the SF phase while it is pinned to unity when the system is in the first Mott lobe.

For comparison, we also show in Figs. 2 and 3 results from exact diagonalization as horizontal dashed and dotted lines. We find that the error in both the energy and the density is larger in the superfluid phase. The fluctuations in the density seem to extend to very long training times.
Figure 2. Training a RBM using stochastic reconfiguration aims to find the ground state of the system. This is done by minimizing the expectation value of the energy, which is plotted above as a function of training step. This convergence is an indication that the system has entered the ground state of the Hamiltonian. Here, $L = 6$ and $N_{\text{max}} = 5$.

Figure 3. Same as in Fig. 2 but for the evolution of the average density over training step.

and point to the limitations of the method and the fact that more complex networks or better optimization may be needed to capture the physics in the superfluid phase.

Taking long-time values of the converged density as we vary the chemical potential at a fixed interaction strength, we obtain the equation of state. Figure 4 shows this property for $U/t = 15$ and several different combinations of $L$ and $N_{\text{max}}$. We find the expected behavior \cite{3,15} where the density is pinned at integer numbers of bosons per site as we cross multiple SF and MI phase boundaries. Two density plateaus corresponding to the first and second Mott lobes are clearly visible in this plot. Increasing the system size from 4 to 7 seems to lead to smoother curves, signaling better convergence during the training, but no appreciable change in the onsets of the Mott regions. Moreover, the results do not change significantly by increasing $N_{\text{max}}$ from 4 to
Figure 4. Equation of state for a fixed interaction strength of $U = 15$, different system sizes and different limits for the boson site occupation numbers. The density, $\rho$, does not show significant variations by changing the system size or the maximum occupation numbers as long as $N_{\text{max}} > \rho$.

5. These observations do not mean that our results are not suffering from finite size effects. Much larger clusters have been shown to give different results in conventional treatments of the model [15, 28].

By training different RBMs as we sweep the chemical potential at increasing values of the interaction strength, we are able to map out the ground state phase diagram of the model in the interaction-chemical potential space. The findings are summarized in Fig. 5. We extend the results up to the second Mott lobe. They qualitatively agree with those obtained using Monte Carlo algorithms and much larger system sizes (see e.g. Ref. [28]). As with other techniques, the accurate locating of the phase boundary near the tip of the Mott lobes are the most challenging due to the Kosterlitz-Thouless nature of the transition [29].

5. Summary
Using the open-source package NetKet, recently developed for the study of quantum many-body systems through machine learning algorithms, we design and train RBMs to represent the ground state wavefunction of the 1D Bose-Hubbard model on small chains. We demonstrate the degree of applicability of the RBM ansatz in various regions of the parameter space as MI or SF phases set in. We show that the optimization of free parameters of the RBMs during training in order to minimize the ground state energy leads to physical properties, including expectation values of the Hamiltonian and a phase diagram, that agree with known exact results from conventional treatments of the model.

6. Acknowledgements
KM acknowledges support from Undergraduate Research Grants at San Jose State University. EK acknowledges support from the National Science Foundation (NSF) under Grant No. DMR-1609560.
Figure 5. Ground state phase diagram of the Bose-Hubbard model for $L = 6$ and $N_{max} = 5$ in the $t/U - \mu/U$ space from the RBM. The first two Mott lobes for $\rho = 1$ and $\rho = 2$ are clearly captured in the strong-coupling region where $t/U \ll 1$. We use a tolerance of $\pm 0.05$ for the density around integer values to determine the departure from the MI phases. Dashed lines are exact results for $L = 128$ from Ref. \[28\].

7. References

[1] Saito H 2017 *Journal of the Physical Society of Japan* **86** 093001 URL https://doi.org/10.7566/JPSJ.86.093001

[2] Saito H and Kato M 2018 *Journal of the Physical Society of Japan* **87** 014001 URL https://doi.org/10.7566/JPSJ.87.014001

[3] Fisher M P A, Weichman P B, Grinstein G and Fisher D S 1989 *Phys. Rev. B* **40**(1) 546–570 URL https://link.aps.org/doi/10.1103/PhysRevB.40.546

[4] Carleo G and Troyer M 2017 *Science* **355** 602–606 ISSN 0036-8075 URL http://science.sciencemag.org/content/355/6325/602

[5] Deng D L, Li X and Das Sarma S 2017 *Phys. Rev. B* **96**(19) 195145 URL https://link.aps.org/doi/10.1103/PhysRevB.96.195145

[6] Torlai G and Melko R G 2017 *Phys. Rev. Lett.* **119**(3) 030501 URL https://link.aps.org/doi/10.1103/PhysRevLett.119.030501

[7] Chen J, Cheng S, Xie H, Wang L and Xiang T 2018 *Phys. Rev. B* **97**(8) 085104 URL https://link.aps.org/doi/10.1103/PhysRevB.97.085104

[8] Huang Y and Moore J E 2017 (Preprint arXiv:1701.06246) URL https://arxiv.org/pdf/1701.06246.pdf

[9] Torlai G and Melko R G 2016 *Phys. Rev. B* **94**(16) 165134 URL https://link.aps.org/doi/10.1103/PhysRevB.94.165134

[10] Torlai G, Mazzola G, Carrasquilla J, Troyer M, Melko R and Carleo G 2018 *Nature Physics* **14** 447–450 ISSN 1745-2481 URL https://doi.org/10.1038/s41567-018-0048-5

[11] Nomura Y, Darmawan A S, Yamaji Y and Imada M 2017 *Phys. Rev. B* **96**(20) 205152 URL https://link.aps.org/doi/10.1103/PhysRevB.96.205152

[12] Huang L and Wang L 2017 *Phys. Rev. B* **95**(3) 035105 URL https://link.aps.org/doi/10.1103/PhysRevB.95.035105

[13] Lu S, Gao X and Duan L M 2018 (Preprint arXiv:1810.02352) URL https://arxiv.org/pdf/1810.02352v2.pdf

[14] Source code and tutorials by Giuseppe Carleo are available at https://www.netket.org/

[15] Batrouni G G, Scalettar R T and Zimanyi G T 1990 *Phys. Rev. Lett.* **65**(14) 1765–1768 URL https://doi.org/10.1103/PhysRevLett.65.1765.
[16] Scalettar R T, Batrouni G G and Zimanyi G T 1991 Phys. Rev. Lett. 66(24) 3144–3147 URL https://link.aps.org/10.1103/PhysRevLett.66.3144
[17] Niyaz P, Scalettar R T, Fong C Y and Batrouni G G 1991 Phys. Rev. B 44(13) 7143–7146 URL https://link.aps.org/10.1103/PhysRevB.44.7143
[18] Rokhsar D S and Kotliar B G 1991 Phys. Rev. B 44(18) 10328–10332 URL https://link.aps.org/10.1103/PhysRevB.44.10328
[19] Krauth W, Trivedi N and Ceperley D 1991 Phys. Rev. Lett. 67(17) 2307–2310 URL https://link.aps.org/10.1103/PhysRevLett.67.2307
[20] Krauth W and Trivedi N 1991 EPL (Europhysics Letters) 14 627 URL http://stacks.iop.org/0295-5075/i=14/a=003
[21] Krauth W, Caffarel M and Bouchaud J P 1992 Phys. Rev. B 45(6) 3137–3140 URL https://link.aps.org/10.1103/PhysRevB.45.3137
[22] Kampf A P and Zimanyi G T 1993 Phys. Rev. B 47(1) 279–286 URL https://link.aps.org/10.1103/PhysRevB.47.279
[23] Bruder C, Fazio R and Schö'n G 1993 Phys. Rev. B 47(1) 342–347 URL https://link.aps.org/10.1103/PhysRevB.47.342
[24] van Otterlo A, Wagenblast K H, Baltin R, Bruder C, Fazio R and Schö'n G 1995 Phys. Rev. B 52(22) 16176–16186 URL https://link.aps.org/10.1103/PhysRevB.52.16176
[25] Prokof’ev N, Svistunov B and Tupitsyn I 1998 Physics Letters A 238 253 – 257 ISSN 0375-9601 URL http://www.sciencedirect.com/science/article/pii/S0375960197009572
[26] Jaksch D, Bruder C, Cirac J I, Gardiner C W and Zoller P 1998 Phys. Rev. Lett. 81(15) 3108–3111
[27] Becca F and Sorella S 2017 Quantum Monte Carlo Approaches for Correlated Systems (Cambridge University Press)
[28] Ejima S, Fehske H, Gebhard F, zu Münster K, Knap M, Arrigoni E and von der Linden W 2012 Phys. Rev. A 85(5) 053644 URL https://link.aps.org/10.1103/PhysRevA.85.053644
[29] Carrasquilla J, Manmana S R and Rigol M 2013 Phys. Rev. A 87(4) 043606 URL https://link.aps.org/10.1103/PhysRevA.87.043606