Machine Learning of Many Body Localization

Philipp Krüger
(Dated: July 18, 2020)

Exact diagonalization was used to find the reduced density matrices for different block sizes of the lowest energy eigenstate of the Heisenberg Model with an additional random field in z-direction at low and high disorder strength. The resulting dataset representing extended and localized phases was used to train a neural network. Afterwards, the trained network was applied on intermediate disorder strengths to deduct the critical disorder strength for a phase transition. The phase transition occurred for all system sizes at $W_c = J$.

I. INTRODUCTION

Firstly, the physical model is introduced. Secondly, the concept of exact diagonalization is briefly presented. As we use the reduced density matrices as the feature for the neural network, we state briefly their derivation.

A. Physical model

1. Hamiltonian of the Heisenberg model

The hamiltonian of the Heisenberg model is shown in equation 1. In the course of further analysis, we choose $J = 1$ and sample h from a uniform distribution such that $h_i \in [-W, W]$.

$$H = \underbrace{J \sum_i \vec{S}_i \cdot \vec{S}_{i+1}}_{\text{Exchange Energy}} - \underbrace{\sum_i h_i S_i^z}_{\text{Random Field}} \tag{1}$$

2. Expectations for the ground state

The expectation for the ground state is dependent on the ratio of the coupling and the local random field.

For $\frac{W}{J} \ll 1$, we expect an extended, delocalized phase, since the exchange energy dominates over the small external field. Therefore, the system can relax to thermal equilibrium serving as its own heat bath in the limit of large system size $L \to \infty$. Here, the reduced density operator of a finite subsystem converges to the equilibrium thermal distribution for $L \to \infty$.[1]

For $\frac{W}{J} \gg 1$, we can expect a localized phase, since the h_i factors dominate over the exchange energy. The resulting states are expected to be product states of spins "up" or "down", as the external field points in z-direction. Also an infinite system cannot equilibrate itself. The local configurations are set by the initial conditions at all times and are adiabatically connected to the trivial state.[1]

B. Exact diagonalization

Exact diagonalization (ED) is a numerical technique we can use to solve the Schrödinger Equation $H|\psi\rangle = E|\psi\rangle$ for the eigenvalues E and eigenvectors $|\psi\rangle$. This only works of the Hamiltonian H represents a discrete and finite system. Most quantum many-particle problems lead to a sparse matrix representation of the Hamiltonian, where only a very small fraction of the matrix elements is non-zero.[2] An efficient method to find ground states is the Lanczos algorithm.[3] At first, the algorithm was numerically unstable. This issue was overcome in 1970 by Ojalvo and Newman.[4]

C. Reduced Density Matrix

The usefulness of reduced density matrices has already been shown by White in 1992 with ground states of Heisenberg chains [5]. In our case we use areal density matrices as features for the neural network to predict the critical disorder strength of a phase change from delocalized to localized. The reduced density matrix is defined in equation 3. Physically, the reduced density matrix ρ_A, provides correct measurement statistics for subsystem A.

$$\rho_{AB} = |\psi_A\rangle\langle\psi_A| \otimes |\psi_B\rangle\langle\psi_B| \tag{2}$$

$$\rho_A = \mathrm{Tr}_B(\rho_{AB}) = |\psi_A\rangle\langle\psi_A| \,\mathrm{Tr}\left(|\psi_B\rangle\langle\psi_B|\right) \tag{3}$$

The reduced density matrix was also used by Zhang in 2019 to learn the localization transition in disordered quantum Ising spin chains. Here, the motivation was to reduce the dimension and filter out redundant information. However, it proved to be inferior in comparison to the full density matrix in the analysis. [6]

D. Neural Networks and Convolutional Neural Networks

Nice, since we can label our dataset with parameters, where we know the outcome. Image classification is a can be solved efficiently with machine learning.

II. COMPUTATIONAL METHODS

The strategy for implementation was as follows:

1. Generate Hamiltonian from random disorder strength and system size and calculate lowest eigenstate near Energy $E = 0$.

2. Generate density matrix from the eigenstate and the respective reduced density matrices for defined block sizes n.

3. Set up machine learning model per n, L that takes density matrices of different W as an input and predicts whether the state represents an extended or a localized phase.

4. Make predictions for different system sizes L and block sizes n and plot the predicitons over W. Then extract W_c from the data by using a fit function.

Critical decisions and specifications for each steps are listed below.

1. *Eigenvalue solver*

scipy.sparse.linalg.eigsh(A, k=6, sigma=0, which='SM',)
sigma = 0:which='SM': smallest magnitude.
Eigenvalue solver, stability convergence

2. *Density matrix reducer*

Density matrix reducer

```
v = np.arange(8)
dm =v*np.identity(8)
print("density matrix:\n",dm)
qutip_dm = qutip.Qobj(dm, dims=[[2,2,2]]*2, type='oper')

reduced_dm_via_qutip = qutip_dm.ptrace([0]).full()
print("Summation over first lattice site:\n",reduced_dm_via_qutip)
reduced_dm_via_qutip = qutip_dm.ptrace([1]).full()
print("Summation over second lattice site:\n",reduced_dm_via_qutip)
reduced_dm_via_qutip = qutip_dm.ptrace([2]).full()
print("Summation over third lattice site:\n",reduced_dm_via_qutip)
```

```
density matrix:
 [[0. 0. 0. 0. 0. 0. 0. 0.]
 [0. 1. 0. 0. 0. 0. 0. 0.]
 [0. 0. 2. 0. 0. 0. 0. 0.]
 [0. 0. 0. 3. 0. 0. 0. 0.]
 [0. 0. 0. 0. 4. 0. 0. 0.]
 [0. 0. 0. 0. 0. 5. 0. 0.]
 [0. 0. 0. 0. 0. 0. 6. 0.]
 [0. 0. 0. 0. 0. 0. 0. 7.]]
Summation over first lattice site:
 [[ 6.+0.j  0.+0.j]
 [ 0.+0.j 22.+0.j]]
Summation over second lattice site:
 [[10.+0.j  0.+0.j]
 [ 0.+0.j 18.+0.j]]
Summation over third lattice site:
 [[12.+0.j  0.+0.j]
 [ 0.+0.j 16.+0.j]]
```

Figure 1: Proof of concept for partial trace calculation.

3. *Machine Learning Models and Error Metrics*

A neural network was used. As the sample size was a number of layers was selected to avoid overfitting. Dropout for regularization.

Cite Low sample size techniques:

Prominent optimzer Adam was used which has these and that strengths. For a binary classification problem, it is common to use as a loss binary crossentropy:

Loss: BinaryCrossentropy from `https://keras.io/api/losses/probabilistic_losses/#binary_crossentropy-function`: Computes the cross-entropy loss between true labels and predicted labels.

Reason: Use this cross-entropy loss when there are only two label classes (assumed to be 0 and 1). For each example, there should be a single floating-point value per prediction.

4. *W_c finder*

Fit with step like function: Candidates: logistic, heaviside

Choose logisitc nicer, because differentiable and transition is not abrupt?

III. RESULTS

A. Generation of density matrix training set

Chosen parameters: $L \in \{10, 11, 12\}$. Repetitions: 500. Measure of variation in the test set??

Plots: What is computationally realizable in 1h concerning time? The training set was sufficiently large enough

We only need M Eigenstates

This is how corresponding density matrices look like

This will be our parameter space for n, L

B. Prediction of extended vs localized phase

Training and validation scores:

C. W_c analysis

Now we generate testing set with $W_{max} \in [0, 4]$. We suspect W_c to be at 1 ??We fit a logistic curve and extract W_c as a parameter.

These are our W_c depending on n, L.

IV. CONCLUSION

W_c depends on n, L (yes/no).

W_c prediction coincides with the expectation (yes/no)

W_c is dependent on these and that effects =¿ scaling analysis? (yes/no)

Citations are numerical[**?**], some more citations [**? ? ? ? ?**].

[1] A. Pal and D. A. Huse, Many-body localization phase transition, Phys. Rev. B **82**, 174411 (2010).

[2] A. Weiße and H. Fehske, Exact diagonalization techniques, in *Computational Many-Particle Physics*, edited by H. Fehske, R. Schneider, and A. Weiße (Springer Berlin Heidelberg, Berlin, Heidelberg, 2008) pp. 529–544.

[3] C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, Journal of Research of the National Bureau of Standards **45**, 255 (1950).

[4] I. U. Ojalvo and M. Newman, Vibration modes of large structures by an automatic matrix-reductionmethod, AIAA Journal **8**, 1234 (1970).

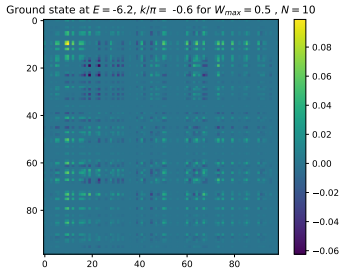
[5] S. R. White, Density matrix formulation for quantum renormalization groups, Physical Review Letters **69**, 2863 (1992).

[6] W. Zhang, L. Wang, and Z. Wang, Interpretable machine learning study of the many-body localization transition in disordered quantum ising spin chains, Physical Review B **99**, 10.1103/physrevb.99.054208 (2019).

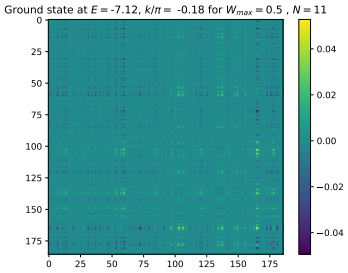
Appendix A: Code listing

Please copy your code in the appendix.

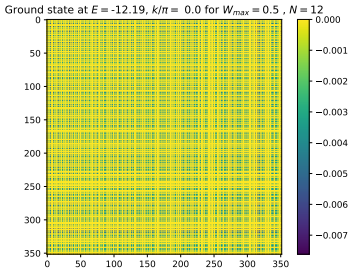
```
"""

Description

"""

import numpy as np

code
```



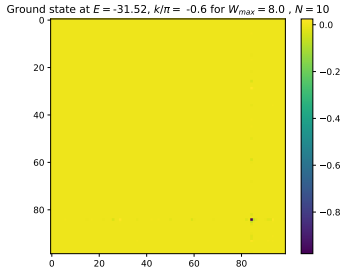
a) Ergodic phase L = 10.



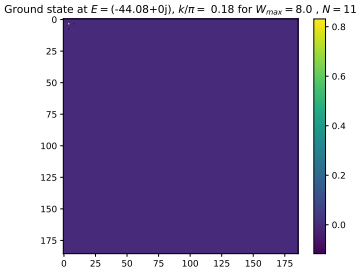
b) Ergodic phase L = 11.



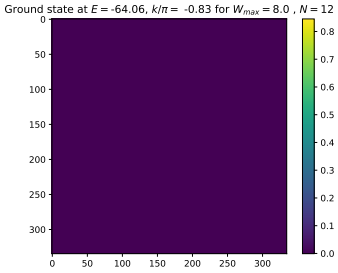
c) Ergodic phase L = 12.



d) Localized phase L = 10.

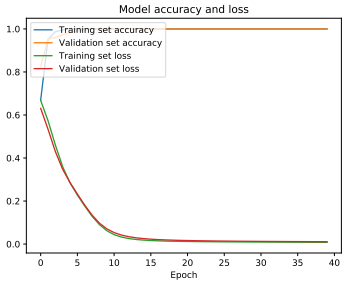


e) Localized phase L = 11.

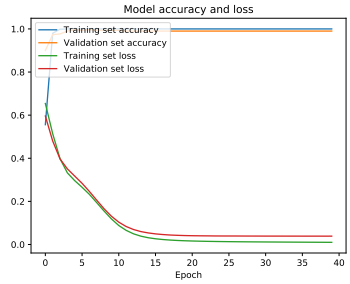


f) Localized phase L = 12.

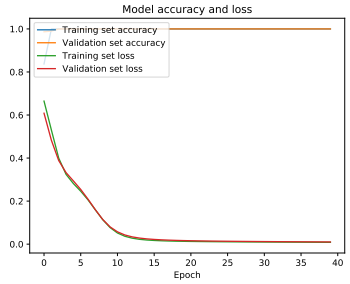
Figure 2: Real part of the density matrix of an ergodic/localized phase for different system sizes L.



a) $N = 10$

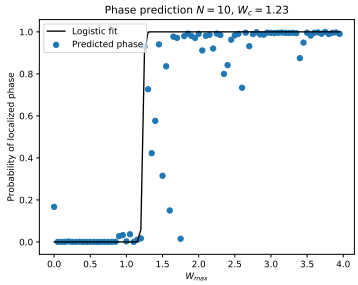


b) $N = 11$

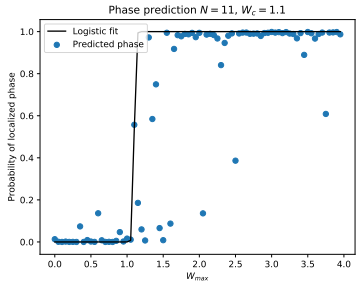


c) $N = 12$

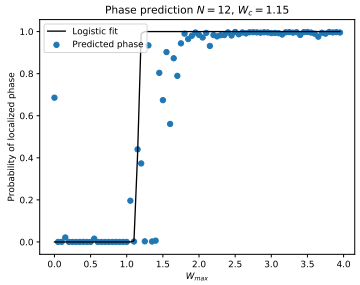
Figure 3: Accuracy and loss of neuronal network plotted over training epochs.



a) $N = 10$



b) $N = 11$



c) $N = 12$

Figure 4: Phase prediction with localized and ergodic phase defined as 1, 0.

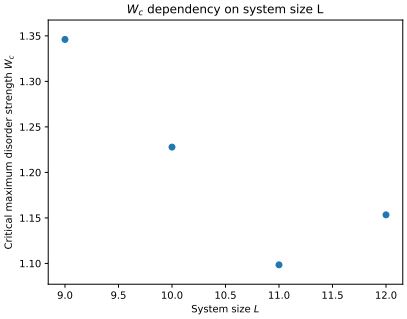


Figure 5

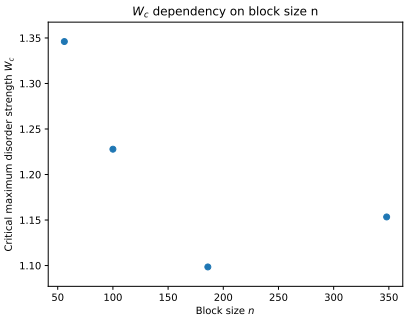


Figure 6