Machine Learning of Many Body Localization

Philipp Krüger
(Dated: July 14, 2020)

Exact diagonalization was used to find the reduced density matrices of the lowest energy eigenstate of the Heisenberg Model with an additional 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 occured for all system sizes at $W_c = J$.

I. INTRODUCTION

Review Literature on task
Why is the topic interesting? =¿ Finding Wc?

A. Hamiltonian of the Heisenberg model

What is it used for? With $J = 1$, $h_i \in [-W, W]$

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

Outcome expectation:
In the ergodic phase (delocalized phase) $h < h_c$, the many-body eigenstates are thermal,20–23 so the isolated quantum system can relax to thermal equilibrium under the dynamics due to its Hamiltonian. In the thermodynamic limit $L \to \infty$, the system thus successfully serves as its own heat bath in the ergodic phase. In a thermal eigenstate, the reduced density operator of a finite subsystem converges to the equilibrium thermal distribution for $L \to \infty$. Thus the entanglement entropy between a finite subsystem and the remainder of the system is, for $L \to \infty$, the thermal equilibrium entropy of the subsystem. At nonzero temperature, this entanglement entropy is extensive, proportional to the number of degrees of freedom in the subsystem.

In the many-body localized phase $h > h_c$, on the other hand, the many-body eigenstates are not thermal:2 the "eigenstate-thermalization hypothesis"20–23 is false in the localized phase. Thus in the localized phase, the isolated quantum system does not relax to thermal equilibrium under the dynamics of its Hamiltonian. The infinite system fails to be a heat bath that can equilibrate itself. It is a "glass" whose local configurations at all times are set by the initial conditions. Here the eigenstates do not have extensive entanglement, making them accessible to density-matrixrenormalization-group–type numerical techniques.5 A limit of the localized phase that is simple is J= 0 with $h > 0$. Here the spins do not interact, all that happens dynamically is local Larmor precession of the spins about their localrandom fields. No transport of energy or spin happens and the many-body eigenstates are simply product states with each spin either "up" or "down." `https://doi.org/10.1103/PhysRevB.82.174411`

non-thermalising phase, in which the system violates the Eigenstate Thermalisation hypothesis `https://arxiv.org/pdf/1610.03042.pdf`

J¡g: (large W) localized phase (adiabatically connected to trivial state), J¿g extended phase (small W) (ordered??)

Is scaling important?

B. Exact Diagonalization

Introduce concepts: Exact Diagonalization,

C. Areal (reduced??) Density Matrix

areal Density Matrix: `http://www.thphys.nuim.ie/staff/jvala/Lecture_9.pdf`

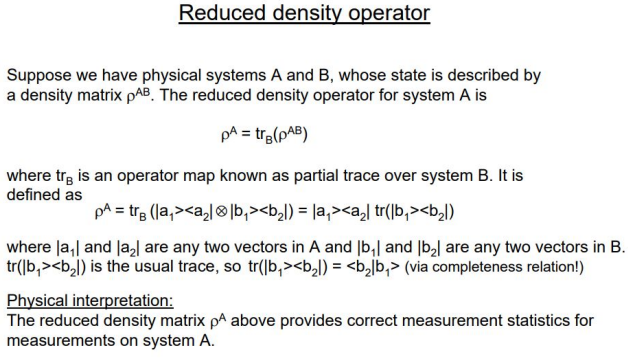


FIG. 1: Example of a figure [7].

vs
direct density matrix: (DOI: 10.1103/PhysRevB.99.054208 says Instead of dividing the system into two subsystems A and B to calculate the reduced density matrix of an eigenstate rho a and using the entanglement spectrum as the training data set 34,35, we directly feed the probability density of the eigenstate psi i computed in the spin basis to the machines as the training data set. The reason for doing so is that, although by preprocessing the training data can reduce the dimension and filter out redundant information,

useful information contained in the wavefunction of the entire system can also be lost.)

Conclusion: =¿ areal

D. Neural Networks and Convolutional Neural Networks

Neural Network, CNN

II. MATERIALS AND METHODS

A. Code Protocol

1. Function: Generate random disorder strength using a uniform distribution
2. Function: Generate Hamiltonian from disorder strength and system size
3. Define and understand phase transition from extended phase to localized phase
4. Function: Picks a number M of lowest eigenstates near Energy E = 0
5. Function: Generate density matrix for an eigenstate
6. Function: Visualize density matrices
7. Function: Set up machine learning model that takes density matrices of different W as an input, and predicts whether the state represents an extended or a localized phase.
8. Function: Make predictions for different system sizes L and block sizes n.
9. Function: Plot the predicitons over W.
10. Function: Extract W_c from the data.

Explain Flow with figure
Fig. 2

B. Machine Learning Models and Error Metrics

Two different approaches:
regression like: Optimizer RMSprop, Loss MAE
classification focused: Optimizer Adam
Loss: BinaryCrossentropy from `https://keras.io/api/losses/probabilistic_losses/#binary_crossentropy-function`: Computes the cross-entropy loss between true labels and predicted labels.

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.

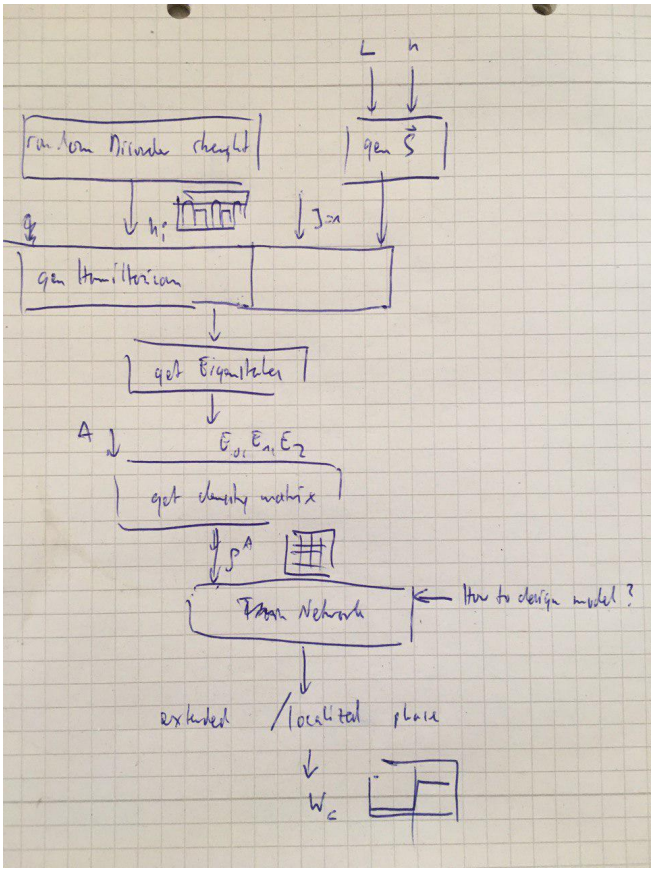


FIG. 2: Example of a figure [7].

Optimizer: Stochastic Gradient Descent, Learning rate? Formula? Principle?

Explain metrics and errors and why they are used. Which ml models are used and why?

Use drop out to reduce overfitting. Low sample size techniques:

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[1], some more citations [2–6].

[1] A. Einstein, Yu. Podolsky, and N. Rosen (EPR), Phys. Rev. **47**, 777 (1935).

[2] R. P. Feynman, Phys. Rev. **94**, 262 (1954).

[3] N. D. Birell and P. C. W. Davies, *Quantum Fields in Curved Space* (Cambridge University Press, 1982).

[4] J. G. P. Berman and J. F. M. Izrailev, Stability of nonlinear modes, Physica D **88**, 445 (1983).

[5] E. Witten, (2001), hep-th/0106109.

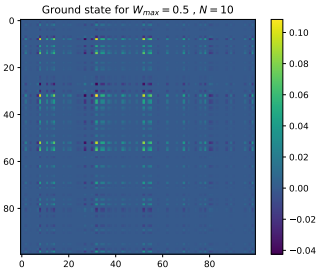
[6] E. B. Davies and L. Parns, Trapped modes in acoustic waveguides, Q. J. Mech. Appl. Math. **51**, 477 (1988).

[7] R. Orus, A practical introduction to tensor networks: Matrix product states and projected entangled pair states, Annals of Physics **349**, 117 (2013), 1306.2164.

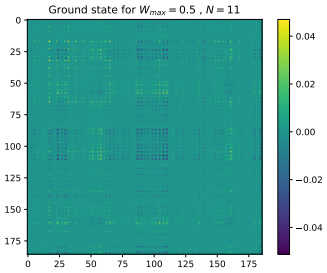
Appendix A: Code listing

Please copy your code in the appendix.

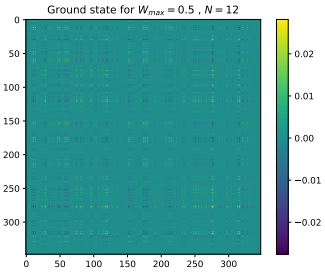
```
"""

Description

"""

import numpy as np

code
```



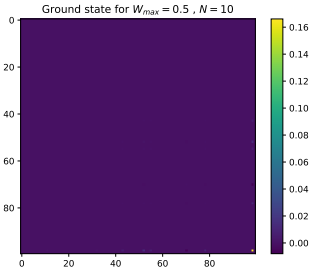
a) Ergodic phase L = 10.



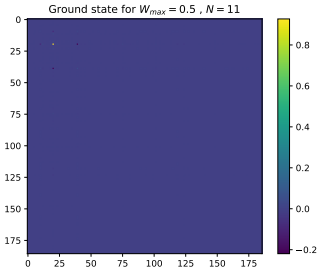
b) Ergodic phase L = 11.



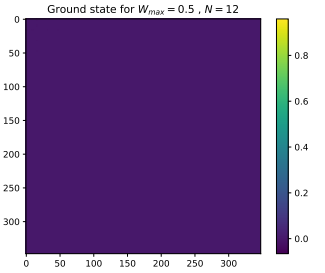
c) Ergodic phase L = 12.



d) Localized phase L = 10.

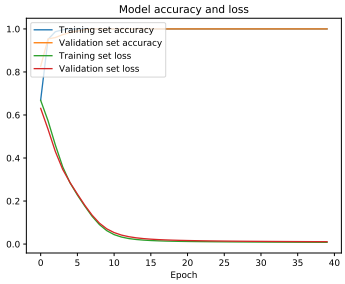


e) Localized phase L = 11.

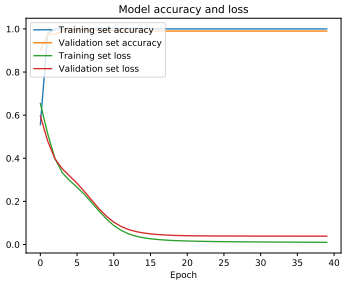


f) Localized phase L = 12.

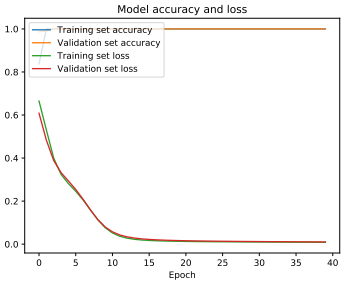
FIG. 3: Real part of the density matrix of an ergodic/localized phase for different system sizes L.



a) $N = 10$

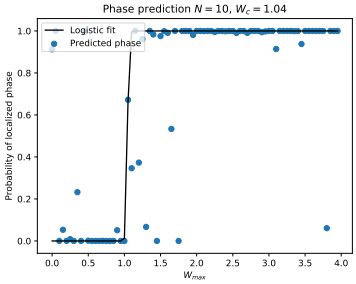


b) $N = 11$

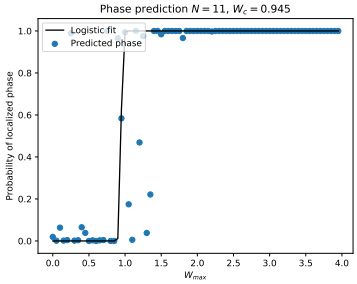


c) $N = 12$

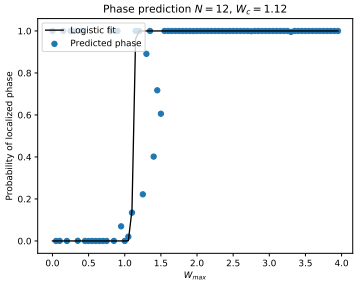
FIG. 4: Accuracy and loss of neuronal network plotted over training epochs.



a) $N = 10$



b) $N = 11$



c) $N = 12$

FIG. 5: Phase prediction with localized and ergodic phase defined as 1, 0.

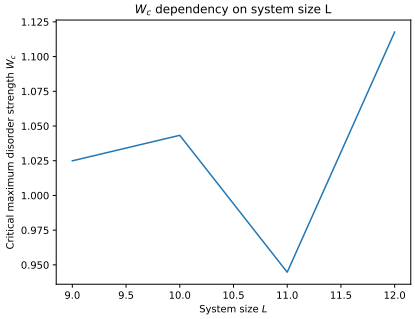

W_c dependency on system size L
Critical maximum disorder strength W_c
1.125
1.100
1.075
1.050
1.025
1.000
0.975
0.950
9.0
9.5
10.0
10.5
11.0
11.5
12.0
System size L

FIG. 6

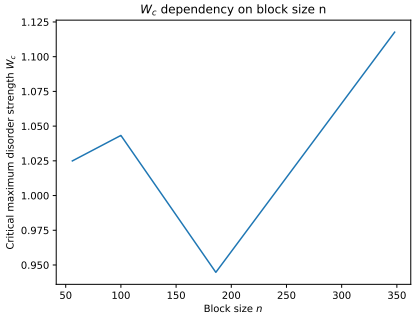

W_c dependency on block size n
Critical maximum disorder strength W_c
1.125
1.100
1.075
1.050
1.025
1.000
0.975
0.950
50
100
150
200
250
300
350
Block size n

FIG. 7