A universal neural network for learning phases

In collaboration with D.-R. Tan, J.-H. Peng, and Y.-H. Hseng

F.-J. Jiang

Physics Department, National Taiwan Normal University, Taipei, Taiwan

July, 2021

Outlines

- Introduction.
- Supervised Neural Network (NN): Multi-layer Perceptron (MLP).
- Training and prediction strategies.
- Applications to the phase transitions of various models: three-dimensional (3D) classical $O(3)$ model, 3D 5-state ferromagnetic Potts model, 3D dimerized quantum spin Heisenberg model, and 2D classical XY model.
- Discussions and Conclusions

Introduction

- Machine learning (ML) is a research field of information and computer science.
- Recently, the ideas and methods from ML, particularly the neural network (NN) have been applied to the studies of many-body problems in Physics.
- Restricted Boltzmann machine $\Rightarrow$ speed up Monte Carlo simulations.
- Supervised (and unsupervised) NNs $\Rightarrow$ Ground state properties and critical points of various models.
- These applications are very successful.
- Great potential applications of ML in physical systems.
- In this talk, we will focus on employing methods of NN in studying phase transitions of several three-dimensional (3D) and two-dimensional (2D) models.

Employed Neural Network: Multi-layer Perceptron (MLP)

- Multi-layer Perceptron (MLP):
  - Input layer (training objects x_0, assign a label $\hat{y}_0$ for each x_0).
  - One hidden layer.
  - Output layer (output y_0 of x_0).
  - The hidden layer consists of 512 (or 1024) independent nodes.
  - Each of the nodes in the hidden layer is connected to each of the elements in its consecutive layers (input and output layers) $\Rightarrow$ weights.
  - The activation functions applied between the relevant layers: ReLU and softmax.
  - The loss function which is considered to determine the difference between the real output (y_0) and desired output ($\hat{y}_0$): categorical cross entropy.
  - The optimizer which is used to adjust weights so that real output will match the expected labels eventually $\Rightarrow$ get the desired weights: Adam

Conventional procedure of applying (supervised) NN to study the phase transition of a many-body system

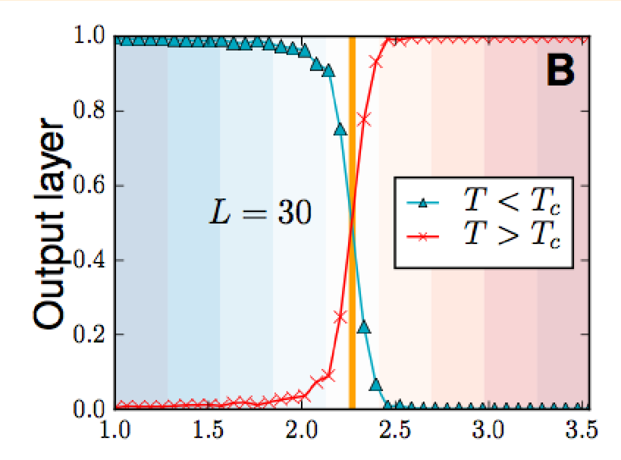
- Training stage:
  - Assign a label (in the output layer) to each element in the training set (input layer) $\Rightarrow$ optmization $\Rightarrow$ desired weights.
- Validation stage: to make sure that the trained-NN produces the outcomes correctly (using part of the training set).
- Testing stage: use the obtained weights (trained-NN) to calculate the desired results.

Conventional NN procedure for studying Ising model

- Several temperatures below T_c are picked.
- Generate (a) few thousands configurations for each of the chosen temperatures.
- Assign the label $(1, 0)$ for every generated configuration described above.
- Similar steps apply to the other side of the phase transition ($T > T_c$), but this time the assigned label is $(0, 1)$.
- With this set up of training, one expects that in the testing stage, the temperature T corresponding to the output vector $(0.5, 0.5)$ is the temperature where the phase transition takes place.

Critical temperature of 2D Ising model

Carrasquilla and Melko (2017)



Disadvantages of the standard training procedure

- One needs to generate few to several thousand configurations for the training.
- A separate training is required for every studied model and for each used system size.
- Such a procedure is time comsuming.

Disadvantages of the standard prediction (testing) stage

- Few to several thousand real configurations are used for the prediction.
- It takes huge amount of storage space if system of larger sizes are considered.

Training set considered in our study

- In our study, the training set consists of only **TWO** configurations.
- Each of the configurations is a **1D lattice made up of 120 sites**.
- One of the two configurations takes the integer 1 as the values for all of its elements.
- The sites of the other configuration have the integer 0 as their values.
- The related labels are $(0, 1)$ and $(1, 0)$.
- With our training strategy, a factor of several hundred to a few thousand in efficiency is obtained (The time needed for the training in our study is much less than the standard approaches).

Prediction procedure in our study

- Instead of using real configurations, we use bulk quantities to construct the needed configurations (120 sites 1D lattices) for the prediction.
- Bulk quantities that saturate to distinct values at the high and the low temperature regions can be employed in our calculations.
- The detailed procedures are in arXiv:2103.10846 (Which contains a second method of conducting the prediction not described here).
- With such a prediction procedure, only few permille fraction of the storage space needed for a typical NN calculation is required in our investigation.
- Binder ratios Q_1 and Q_2.

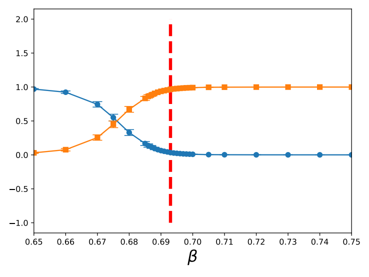
The method of determining the critical temperatures of the studied models

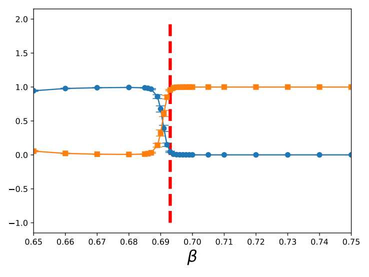
- Crossing of the two components of the output vectors.
- The minimum of the magnitude of the output vectors.

- *Only one NN is trained, and this only one NN is applied to calculate the critical points of all the considered models.*

Results of the 3D classical $O(3)$ model

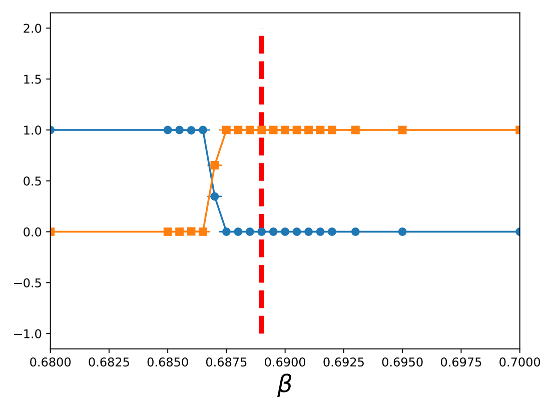
Q_1, 1D training, $L = 12$(left) and $L = 48$(right).





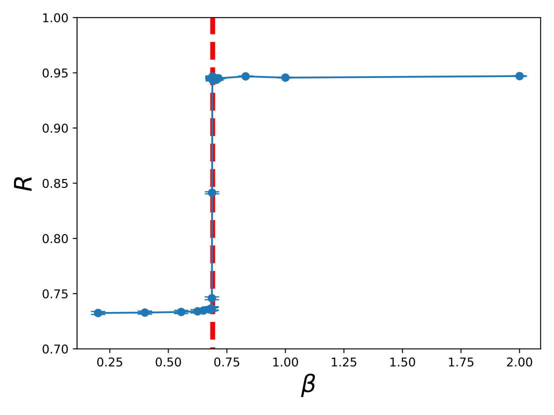
Result of the 3D 5-state ferromagnetic Potts model 1

Q_1, 1D training, $L = 20$



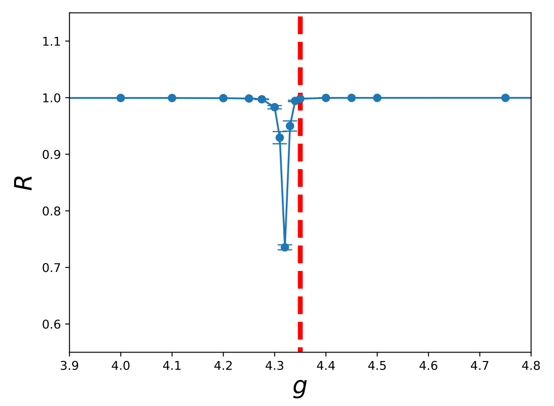
Result of the 3D 5-state ferromagnetic Potts model 2

Configuration, 1D training, $L = 20$



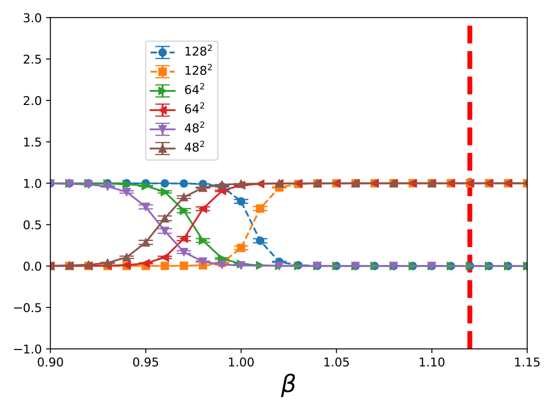
Result of a 3D dimerized quantum spin Heisenberg model

Q_2, 2D training.



Results of a 2D classical *XY* model

Q_1, 1D training of 200 sites.



Discussions and Conclusions

- An extremely efficient and universal NN is built. The infrastructure of our NN is much implier than the conventional ones typical used in the literature when studying phase transitions is concerned.
- The training is conducted on a 1D lattice of 120 sites.
- Bulk quantities are employed to construct the needed configurations (120 sites 1D lattices) for the prediction.
- The (only one) NN successfully calculates the critical points of several 3D and 2D models including the 3D $O(3)$ model, the 3D 5-state Potts model, a 3D dimerized quantum spin Heisenberg model, and the 2D classical XY model.
- The built NN can be applied to many (other) models as well when studying the associated phase transitions (critical points) is concerned.
- Compared to any known schemes in the literature, our NN is several hundred to a few thousand times more efficient in both the computation and storage.

Acknowledgement

- The NN results are obtained using keras and tensorflow.
- Partial support from MOST of Taiwan is acknowledged.

References

- Juan Carrasquilla, Roger G. Melko, Nature Physics **13**, 431434 (2017).
- https://keras.io
- https://www.tensorflow.org
- C.-D. Li, D.-R. Tan, and F.-J. Jiang, Ann. Phys. **391** 312 (2018).
- D.-R. Tan *et al.*, 2020 New J. Phys. 22 063016.
- D.-R. Tan and F.-J. Jiang, Phys. Rev. B **102**, 224434 (2020).
- D.-R. Tan, J.-H. Peng, Y.-H. Hseng, and F.-J. Jiang, arXiv:2103.10846.