A NEURAL NETWORKS APPROACH TO THE THERMODYNAMIC AVERAGES CALCULATION OF THE EDWARDS-ANDERSON SPIN GLASS MODEL

D.Yu. Kapitan, P.D. Andriushchenko, A.E. Rybin, A.O. Korol, E.V. Vasiliev, M.A. Padalko, A.V. Perzhu, R.A. Volotovskiy, A.G. Makarov, Yu.A. Shevchenko, K.S. Soldatov, V.Yu. Kapitan, K.V. Nefedev

kapitan.diu@dvfu.ru

Far Eastern Federal University
Vladivostok, Russian Federation

Introduction

One of the most popular approaches to studying systems with complex interactions in recent decades has been the Monte Carlo method. At the same time, this approach is evolving and improving due to modern advances in computing technologies and new computational approaches, such as machine learning and neural networks. The application of machine learning to statistical physics began relatively recently but is rapidly evolving. For example, in [1,2,3] approaches to calculate the critical temperature from Monte Carlo-generated spin configurations were shown. In [4] a similar approach is shown which allows recognizing different phases of spin systems depending on parameters such as DMI and external magnetic field H_z using a neural network. In this research, we will focus on the study of the ability of different neural network architectures to predict the main thermodynamic characteristics of the Edwards-Anderson spin glass models.

Spin glass model

The model is a square lattice of interacting Ising spins, characterized by random distribution of interactions between spins. Hamiltonian of system :

$$E = \sum_{\langle i,j \rangle} J_{ij} S_i S_j$$

where S_i, S_j – spins of the lattice, <i,j> denotes the summation over pairs of interacting spins in a system with size N.

Models and Data

For each considered configuration of the spin glass, 60-temperatures were calculated from 0.1 to 6 in steps of 0.1. Simulations were performed using a parallel replica-exchange Monte Carlo (MC) method. To train neural networks, it was necessary to prepare data sets for training, validation and testing of neural network models.

In total, two datasets each were calculated for two system sizes. The datasets contained $2N$ values of all bonds J , one value of temperature T , and 2 output values of average energy $\langle E \rangle$ and average magnetization $\langle M \rangle$, that is a total of $2N + 3$ values. For the $6x6$ spin glass, the small dataset consisted of 834 configurations (with dimension $50040x75$, since each spin glass configuration was calculated at 60 different temperature values) and the large one consisted of 41405 configurations ($2484300x75$). For the $10x10$ spin glass, the small dataset consisted of 10302 configurations ($618120x203$) and the large of 43596 configurations ($2615760x203$). All datasets were divided into train, validation and test data in proportion to 0.8:0.15:0.05.

In this paper, we compared the performance of fully-connected neural networks(FC) with the custom architectures which we proposed. We created two NN architectures with two levels of spin lattice abstraction CC1 and CC2.

The first architecture CC1 proposes to consider the first hidden layer h_1 as virtual connections, and the second layer h_2 as virtual spins. In such a network, all neurons of layer h_1, except the temperature neuron, are connected to the corresponding neurons of layer h_2 in the same way as connections in a square lattice are connected to spins. Layer h_2 and h_3 are fully-connected.

The second architecture CC2 offers the same approach as in CC1, except for the connections between layers h_2 and h_3. In CC2 it was proposed to consider layer h_3 also as virtual spins, and to connect layer h_2 with h_3 similarly to neighboring spins in a square lattice.

Models and Data

To control the overfitting of neural networks during training, the loss function L was calculated on the validation dataset, which did not participate in the training. The graph of the loss function L as a function of the number of training epochs is shown in Fig 1.

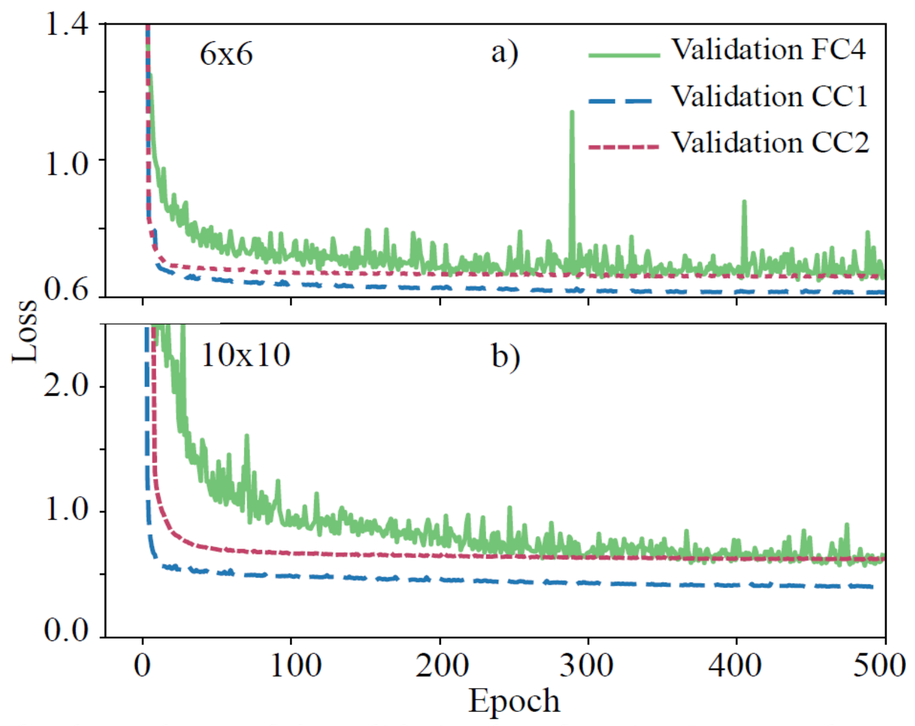

Fig. 1 - The dependence of the validation loss function L on epoch number for the fully-connected FC4 network and the proposed CC1 and CC2 architectures trained on a big dataset for the models $6x6$ (a), $10x10$ (b) .

Results

Examples of comparison of prediction quality of different neural network approaches are shown in Fig. 2.

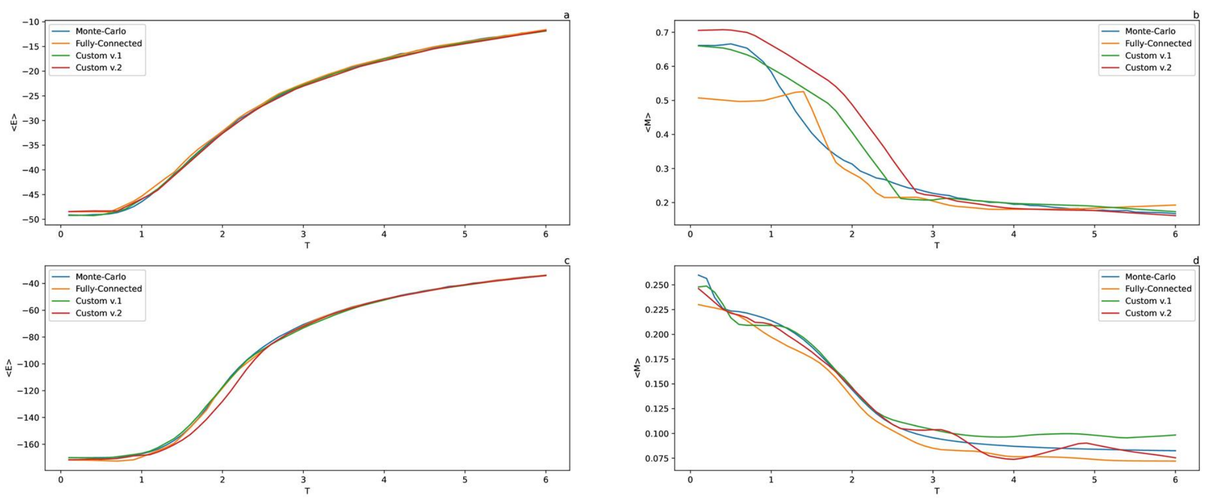

Fig. 2. - Comparison of results of neural networks with different architecture: (a) Dependence of energy values on temperature for model N= $6x6$ (b) Dependence of magnetization values on temperature for model N= $6x6$ (c) Dependence of energy values on temperature for model N= $10x10$ (d) Dependence of magnetization values on temperature for model N= $10x10$.

Conclusion

In this work we have demonstrated the possibility of using neural networks to calculate the basic thermodynamic characteristics of the Edwards-Anderson spin glass models. We also compared the accuracy results of neural network implementations of different architectures. On average, the custom neural network architecture showed more accurate results compared to the full-link architecture.

This work was supported by RSF (project No. 21-72-00157).

[1]. Carrasquilla, J., & Melko, R. G. (2017). Machine learning phases of matter. Nature Physics, 13(5), 431-434. [2]. Shiina, K., Mori, H., Okabe, Y., & Lee, H. K. (2020). Machine-learning studies on spin models. Scientific reports, 10(1), 1-6 [3]. Korol, A. O., & Kapitan, V. Y. E. (2021). Neural Network for Prediction of Curie Temperature of Two-Dimensional Ising Model. Dal'nevostochnyi Matematicheskii Zhurnal, 21(1), 51-60.[4]. Kapitan, V., Vasiliev, E., Perzhu, A., Kapitan, D., Rybin, A., Korol, A., ... & Shevchenko, Y. (2021). Numerical simulation of magnetic skyrmions on flat lattices. AIP Advances, 11(1), 015041.