EXPLORING THERMODYNAMIC BEHAVIOR OF SPIN GLASSES WITH MACHINE LEARNING

Vitalii Kapitan
Department of Statistics and Data Science, National University of Singapore
Singapore
kapitanv@nus.edu.sg

Dmitrii Kapitan
Department of Theoretical Physics and Intelligent Technologies, Far Eastern Federal University
Vladivostok, Russia
kapitan.diu@dvfu.ru

Petr Andriushchenko
Heidelberg Institute of Global Health, Heidelberg University Hospital, Heidelberg University
Heidelberg, Germany
petr.andriushchenko@uni-heidelberg.de

ABSTRACT

In this paper, we consider the regression problem of predicting thermodynamic quantities – specifically the average energy $\langle E \rangle$ – as a function of temperature T for spin glasses on a square lattice. The spin glass is represented as a weighted graph, where exchange interactions define the edge weights. We investigate how the spatial distribution of these interactions relates to $\langle E \rangle$, leveraging several machine learning approaches that we specifically developed for this task. While $\langle E \rangle$ is used to demonstrate the approach, our framework is general and can be applicable to the prediction of other thermodynamic characteristics.

The nature of the low-temperature phase in frustrated spin systems, especially in nearest-neighbor models remains an open question even after many years of active research. The analysis of the partition function, which contains information about all states of the system, is crucial for accurately determining the features of interacting spin systems in thermodynamic equilibrium. However, due to the exponentially large number of states and the complexity of specifying the generating function, it is frequently impossible to compute the partition function exactly. Long relaxation times, rough energy landscapes, and macroscopic degeneracy of the ground states in frustrated systems further complicate this problem Altieri & Baity-Jesi (2024).

A spin glass is a set of interacting magnetic moments, originating from spins, in which the magnetic interactions are randomly distributed in sign. Two important characteristics that distinguish spin glasses from other lattice models are disorder brought on by the freezing of spins at low temperatures and frustration, where competing magnetic interactions prohibit all interactions in the system from being satisfied simultaneously. No simple symmetric configuration of the set of spins corresponds to an equilibrium state Vincent (2024). These problems belong to the class of nondeterministic polynomial problems (NP-class) Lucas (2014). For these reasons, the development of efficient algorithms is one of the relevant goals in the theory of frustrated magnetism for studying the low-temperature phase of spin glass models.

Our research presents a broader approach to this fundamental challenge: we demonstrate how machine learning techniques can be applied to predict thermodynamic properties of a spin glass on a square lattice in the frame of two-dimensional Edwards - Anderson (EA) model with periodic boundary conditions (PBC) on an Ising square lattice $N = L \times L$, where N is a total number of spins and L is a linear size of the system Edwards & Anderson (1975). Specifically, we focus on modeling how the mean energy $\langle E \rangle$ varies with temperature T, though our method can be extended to other physical observables of the system.

In order to do this, we consider the spin glass model as a weighted graph,where the graph architecture corresponds to the lattice and the values of the edges indicate the values of the exchange interactions. Thus, we are looking for a functional relationship between the spatial distribution of the exchange integrals $J_k = f_J(x_k, y_k)$ and $\langle E \rangle$. Here, x_k and y_k represent the bond coordinates for bond k, J_k is the bond value, and f_J is the function of the spatial distribution of spin glass bond values.

For a lattice of size $N = L \times L$, there are 2^{2N} possible configurations of the exchange integrals. These configurations range from the fully antiferromagnetic to the fully ferromagnetic state. In the first case, all exchange interactions $J_k = -1, \forall k$ leading to $\sum_{k=1}^{2N} J_k = -2N$. For the second case, all interactions are $J_k = 1, \forall k$ and $\sum_{k=1}^{2N} J_k = 2N$.

We study $N = 6 \times 6$ and 10×10 systems of spins. Using the complete enumeration method, we calculated datasets with all possible distributions of $\sum_{k=1}^{2N} J_k$ for each system size. Since we used several types of neural networks, described in detail below, we had to slightly modify the input data format to fit the input layer of the neural networks – for example, for the fully connected neural networks, each bond configuration was fed to the input as a one-dimensional structure, while for convolutional neural networks we fed the bonds by dividing them into two channels – horizontal and vertical (by analogy with the classical example of dividing an RGB image into three channels). Also, we utilized data proportions of $0.8 : 0.15 : 0.05$ for training, validation, and testing, respectively.

To capture a relation between J_k and $\langle E \rangle$, a model can be trained to recognize patterns in the arrangement of exchange interaction values on the lattice and their impact on macroscopic parameters. In particular, the spin glass configuration defined by the set $\{J_1, J_2, J_3, \ldots, J_{2N}\}$ and temperature T should be provided to the network, that then can be trained to predict the average energy $\langle E \rangle$.

We use several different types of machine learning algorithms to conduct a comparative analysis of their predictive ability to solve this class of problems:

1. Fully connected (FC) network architectures with different numbers and sizes of hidden layers. We used fully connected neural networks as a baseline solution. After training NN's, we selected a pair of architectures with the lowest RMSE on the test sample and applied these architectures further to evaluate the effectiveness of our proposed approaches.

2. Custom Connected (CC) network architectures. To reduce the prediction error, we investigated DNNs whose architecture would convey information about the spatial arrangement of bonds on a square lattice. We proposed to replace fully connected hidden layers with layers in which neurons would be connected similarly to spins on a square lattice. The resulting architectures proposes to consider the first hidden layer h_1 as virtual bonds, and the layers h_2 and h_3 as virtual spins. The differences in the architectures are in how we connect the "spins" between layers, we use both fully connected connections and those that mimic the bonds between spins in a lattice..

3. Convolutional Neural Networks (CNNs) are well-suited for predicting the average energy levels in spin glass models because of their capacity to learn from the complex patterns observed in the bond configurations. Our method takes advantage of the intrinsic structure of spin glass systems, whose bond interactions have different spatial orientations, by using different CNN channels for vertical and horizontal interactions. We experimented with various of configurations, by varying different layers, such as upscaling, amount of FC layers. In addition, we used a gradient scaling approach, the Landscape Modification (LM) method Choi (2024), to optimise the CNN. The LM method enhances optimisation by transforming the objective function $g(x)$ into a modified form $\hat{g}(x)$ managed by parameters a and threshold c. This can help gradient-based optimizers like Adam to improve the avoidance of local minima and saddle points, leading to faster convergence to a global or high quality local optimum. In our case, the scaling procedure is as follows: the gradient ∇g_t is scaled using a transformation function f of the running loss rl_t and the parameter c_t on a step t:

$$\nabla \hat{g}_t = \frac{\nabla g_t}{af((rl_t - c_t)_+) + 1}, \tag{1}$$

where c_t was set close to the minimum value of the loss function of the default NN model.

The results of the work of neural networks were scored by root mean squared error (RMSE) of the average energy $\langle E \rangle$. The resulting RMSE values depending on NN architecture and system size are presented in Table 1.

	RMSE							
	FC1	FC2	CC1	CC2	CNN1	CNN2	CNN1LM	CNN2LM
6×6	0.05552	0.05568	0.04126	0.04909	0.01182	0.00249	0.00338	0.00437
10×10	0.03766	0.038168	0.026071	0.030173	0.00178	0.00187	0.00309	0.00172

Table 1: The root mean squared error (RMSE) metric for the predicted $\langle E \rangle (T)$ using the proposed architectures.

A comparative analysis of different neural network architectures for predicting the average energy $\langle \mathrm{E} \rangle$ of spin glass systems shows that convolutional neural networks (CNNs) significantly outperform fully connected (FC) and custom connected (CC) networks. The superior performance of CNNs is due to their ability to effectively capture spatial patterns in connection configurations, a critical aspect of spin glass models. Although CC networks outperform FC networks by incorporating spatial connectivity, they still underperform CNNs. Model performance consistently improves with increasing system size, likely due to a richer data representation that enables better learning. Applying the landscape modification (LM) method yields mixed results, demonstrating its potential to improve optimization in certain cases, especially for larger systems. In this study, we tested two hyperparameter selection strategies for LM: with respect to 6×6 and 10×10 systems, which may explain the differences in the obtained results. This outcome could be attributed to suboptimal choices of the hyperparameters f, a, or c, as identifying the ideal values can be challenging. We plan to study and improve it in further research.

We demonstrates that several novel architectures of Deep neural networks that we developed, made it feasible to determine the system's global features with a high degree of accuracy based on the microarchitecture (a certain distribution of bonds values) of a specific spin glass configuration.

In the future, we plan to extend our study to significantly larger spin systems. This presents a considerable challenge, as direct DOS calculations become infeasible for such systems. Therefore, it is crucial to train the model to extrapolate the identified dependencies from small lattices to systems with a larger number of particles.

We also plan to use a Reinforcement Learning approach using the Deep Q-Network (DQN) algorithm, which uses a Graph Neural Network (GNN) as an agent. GNNs have several advantages over other approaches when studying spin glasses, such as being able to capture the spatial structure of the lattice as well as the interactions between adjacent spins. This may allow researchers to model more complex spin glass models.

Author Contributions

All authors contributed equally to this work.

Acknowledgments

V. Kapitan performed the optimization using the LM method and contributed to the manuscript with financial support provided by the project "MAPLE" with grant number 22-5715-P0001 under the National University of Singapore Faculty of Science Ministry of Education Tier 1 grant Data for Science and Science for Data collaborative scheme.

References

A. Altieri and M. Baity-Jesi. An introduction to the theory of spin glasses. In Tapash Chakraborty (ed.), *Encyclopedia of Condensed Matter Physics (2nd Ed.)*, pp. 361. Academic Press, 2024.

M. Choi. Improved metropolis–hastings algorithms via landscape modification with applications to simulated annealing and the curie–weiss model. *Adv. in Applied Probability*, 56(2):587, 2024.

S.F. Edwards and P.W. Anderson. Theory of spin glasses. *J. of Phys. F: Metal Phys.*, 5(5):965, 1975.

A. Lucas. Ising formulations of many np problems. *Frontiers in physics*, 2:5, 2014.

Eric Vincent. Spin glass experiments. In Tapash Chakraborty (ed.), *Encyclopedia of Condensed Matter Physics (2nd Ed.)*, pp. 371. Academic Press, 2024.