UMYU Scientifica, Vol. 1 NO. 1, September 2022, Pp 336 – 341
https://doi.org/10.56919/usci.1122.043
ISSN: 2955 – 1145 (print); 2955 – 1153 (online)
A periodical of the Faculty of Natural and Applied Sciences, UMYU, Katsina

ORIGINAL RESEARCH ARTICLE

Sampling of thermodynamic systems with neural network surrogates

Yahaya Ibrahim
Department of Physics, Umaru Musa Yarádua University, Katsina, Nigeria

ARTICLE HISTORY
Received August 30, 2022
Accepted September 27, 2022
Published September 30, 2022

KEYWORDS
Neural networks, surrogate modeling, Ising model, statistical sampling, Monte Carlo simulation, machine learning

ABSTRACT
Traditional sampling methods such as the Monte Carlo method are computationally expensive and not feasible for studying large and complex systems. These methods are essential for developing new materials, optimizing chemical reactions, and understanding biological processes. However, simulating thermodynamic systems for physically relevant system sizes is computationally challenging. This is partly due to the exponential growth of the configuration space with the system size. With the current Monte Carlo methods, studying the same system for different investigation of its properties means repeating the expensive computation multiple times. In this article, I showed that thermodynamic systems can be sampled using a surrogate neural network model thereby avoiding the computationally expensive proposal Monte Carlo methods for subsequent investigations. To demonstrate the method, I trained a feed-forward neural network surrogate for the Boltzmann distribution of the Ising model. This approach would potentially help accelerate Monte Carlo simulations towards understanding the physics of novel materials and some biological processes.

INTRODUCTION

Efficient sampling of thermodynamic systems is a key area of research in many fields, as it provides insights into the behavior of complex systems that are difficult to observe experimentally (Hylton, 2020; Nicoli *et al*., 2021a, 2021b; Wang *et al*., 2019). Sampling methods are essential for designing new materials, optimizing chemical reactions, and increasingly in understanding biological processes (Husic *et al*., 2020; Lee *et al*., 2021; Sultan *et al*., 2018). Monte Carlo simulation is a powerful tool for modeling complex systems and analyzing the associated risks (Barzegar *et al*., 2018; Blöte & Deng, 2002). However, Monte Carlo simulations can be computationally expensive and time-consuming, especially for large-scale problems (Anderson *et al*., 2018; Shapiro, 2011). To speed up the process, many techniques were developed over the years.

One of the widely employed method for speeding up Monte Carlo simulations is Importance sampling (Bayrakci *et al*., 2010; Elvira *et al*., 2015). This method is variance reduction technique that involves sampling from a distribution that is better suited to the problem being solved. The idea is to sample the configurations more frequently from the regions of the distribution that contribute the most to the solution, reducing the overall variance of the simulation.

A recent trend is the application of machine learning models to classify phases of matter (Broecker *et al*., 2017) and solve partial differential equations modeling physical phenomena (Cai *et al*., 2021; Shukla *et al*., 2021). The latter is known as Physics-Informed Neural Networks. The machine learning approach have also been successfully to learning probability distributions in statistics (Kosmatopoulos & Christodoulou, 1994; MacKay, 1995). The question then arise that: can we learn statistical physics probability distributions to speed up simulations and inference of physical investigations?

In this article, I provide evidence that indeed the Boltzmann equilibrium can be learned by a neural

Correspondence: Ibrahim, Y. Department of Department of Physics, Umaru Musa Yarádua University, Katsina, Nigeria. ✉ yahaya.ibrahim@gmail.com
How to cite: Yahaya, Ibrahim. (2022) Sampling of thermodynamic systems with neural network surrogates. UMYU Scientifica, 1(1), 336 – 341. https://doi.org/10.56919/usci.1122.043

network efficiently. I demonstrate application of the neural network surrogate technique to approximation of the canonical 2D Ising probability distribution. The paper is organized as follows: in the next section, I present the methodology used to conduct this study (including model definitions), thereafter discuss the results and then conclude.

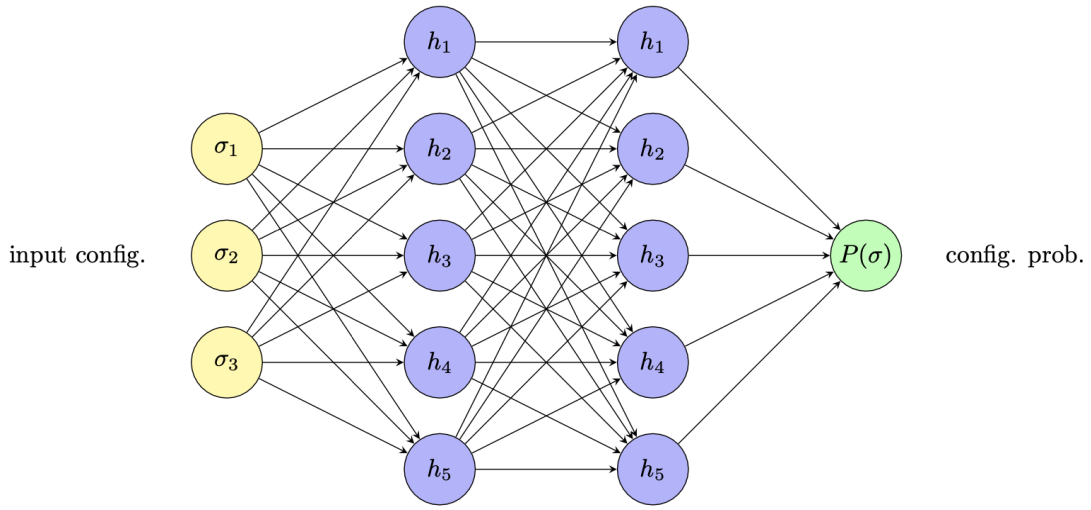

Figure 1: Feed-forward neural network as a surrogate for the Boltzmann distribution.

MATERIALS AND METHODS

The methodology for training a neural network surrogate using exact computation of probability distribution data involves two main steps. First, configuration data is generated from the Boltzmann distribution for a system with 16 spins (see a sketch for 3 spins in Figure 1). This data represents a set of configurations of the system at different energy levels, with a higher probability of sampling configurations with lower energy levels.

Next, the generated data is used to train the neural network surrogate. The neural network is implemented with two hidden layers and a sigmoid output layer using the Tensorflow 2 framework. The sigmoid output layer is chosen to model the probability distribution of the system, as it outputs a value between 0 and 1, which can be interpreted as the probability of sampling a particular configuration.

During training, the neural network is adjusted to minimize the difference between its predicted probability distribution and the exact computation of the probability distribution from the Boltzmann distribution. This is done using a loss function, which measures the difference between the predicted probability distribution and the exact computation of the probability distribution. The training is performed using stochastic gradient descent, which iteratively adjusts the weights and biases of the neural network to improve its predictions.

Once the neural network has been trained, it can be used as a surrogate to efficiently sample from the probability distribution of the system. This involves generating new configurations by sampling from the neural network's learned probability distribution, which is much faster than performing exact computations of the probability distribution. The neural network surrogate can be used to explore the phase space of the system and obtain thermodynamic properties of interest.

The model

We consider the Ising model. The energy of the Ising model arranged in a lattice is given by the Boltzmann distribution

$$E(\sigma) = -\langle\sigma|W|\sigma\rangle - \langle h||\sigma\rangle \qquad (1)$$

where $|\sigma\rangle = (\sigma_1, \cdots, \sigma_N)^T$ is the spin vector and we have used the Dirac's bra-ket notation. W and $|h\rangle$ are the spin-spin (coupling) interaction matrix and external field bias respectively. The probability of finding the system in a configuration σ is given by the Boltzmann distribution

$$P(\sigma) = \frac{1}{Z} e^{-\beta E(\sigma)}, \qquad (2)$$

where Z is the partition function. This is a general formulation which allows us to compute the energy for both 1D and 2D easily using linear algebra packages. For example, for four spin system, the corresponding 1D coupling matrix is

$$W^{(1D)} = \begin{pmatrix} 0 & w_{12} & 0 & 0 \\ w_{21} & 0 & w_{23} & 0 \\ 0 & w_{32} & 0 & w_{34} \\ 0 & 0 & w_{43} & 0 \end{pmatrix}, \quad (3)$$

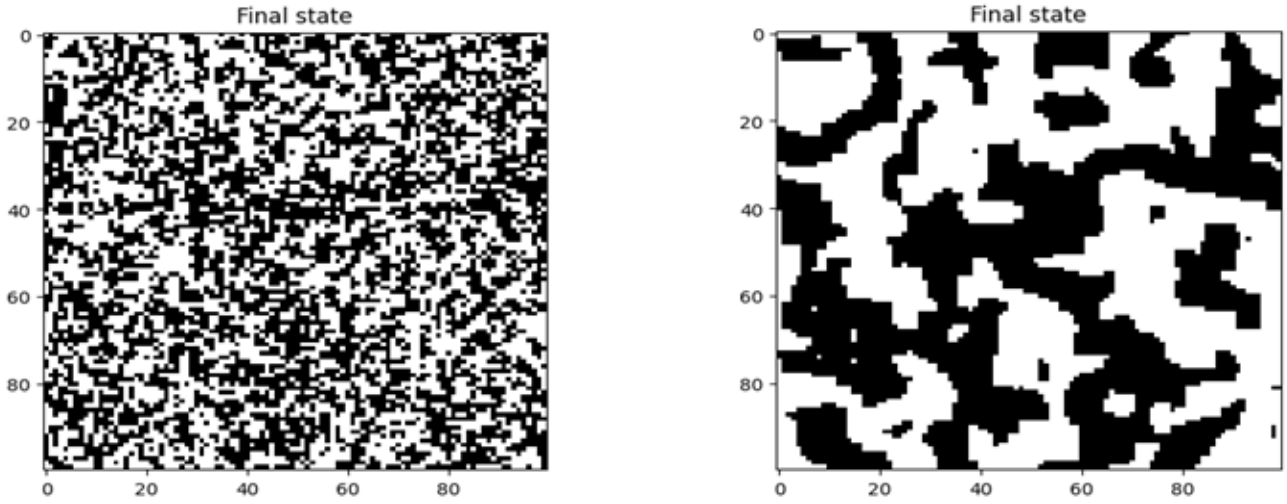

Figure 2: 2D Final state of a simulated Ising Model with temperature T=0.2 (in units where k_B=1) and weights fluctuations of . (Left figure) external field (right figure)

while that of the 2D system allows for long range interaction is

$$W^{(2D)} = \begin{pmatrix} 0 & w_{12} & w_{13} & 0 \\ w_{21} & 0 & 0 & w_{24} \\ w_{31} & 0 & 0 & w_{34} \\ 0 & w_{42} & w_{43} & 0 \end{pmatrix}. \quad (4)$$

Of course, the choice of $W^{(2D)}$ is not unique. Another equally valid coupling for this system is

$$W^{(2D)} = \begin{pmatrix} 0 & w_{12} & 0 & w_{14} \\ w_{21} & 0 & w_{23} & 0 \\ 0 & w_{32} & 0 & w_{34} \\ w_{41} & 0 & w_{43} & 0 \end{pmatrix}. \quad (5)$$

This implies we can study Ising model of different dimensions using the same formulation we have developed in this section. Two advantages can be drawn from this: first, it makes easier intuitively understanding the effects of long range interactions between the spins and the influence of the topology of the lattice on the statistics. Second, the configuration energies and hence the probabilities can be computed relatively efficiently using linear algebra software packages such as Python numpy or Jax. Some final state configuration of 2D spin lattice can be found in Figure 2.

Neural network training

Training the neural network seek to minimize the mean squared error loss function

$$L = \frac{1}{M}\sum_{\sigma}\left(P(\sigma) - P_{nn}(\sigma)\right)^2, \quad (6)$$

where $P(\sigma)$ is the exact probability, $P_{nn}(\sigma)$ is the neural network and M is the number of training samples. The neural network specification of the trained model can be found in Table 1 below.

RESULTS AND DISCUSSION

After the generating the training data from the Boltzmann distribution, the trained neural network surrogate was used to predict the probability distribution which is compared with the ground truth values in Figure 3. It is noteworthy that we can generate the probability distribution for $1D$ and $2D$ Ising systems in a straightforward manner using optimized linear algebra (or machine learning) libraries (see eqns. 3, 4 and 5).

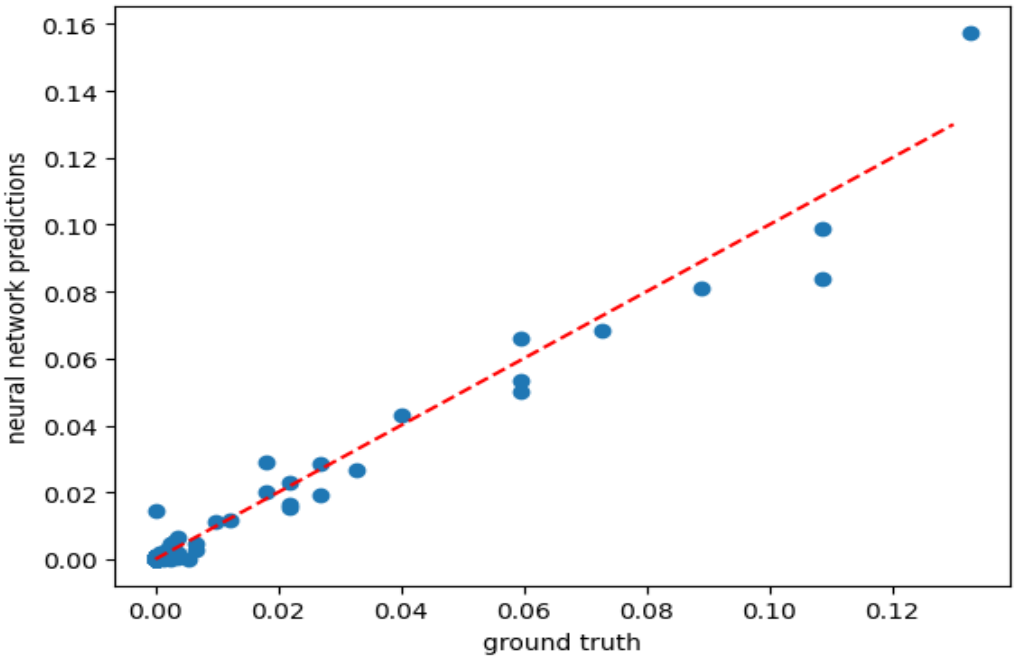

Figure 3: Neural network prediction against the exact Boltzmann distribution after 20 epochs with two hidden layers of 25 nodes each. The network has a Mean Absolute Error (MAE) of 5.166e-6.

Table 1: Table of parameters specification of the neural network.

Layer (type)	Output Shape	# of Parameters
dense1 (Dense)	(None,25)	425
dense2 (Dense)	(None,25)	650
dense3 (Dense)	(None,25)	26

Related works of applying neural networks to statistical physics tends to focus on finding more accurate force fields (Botu *et al.*, 2017; Gkeka *et al.*, 2020; Poltavsky & Tkatchenko, 2021; Unke *et al.*, 2021), classification of phases of matter (Broecker *et al.*, 2017), physics-informed solutions of partial differential equations (Cai *et al.*, 2021; Laubscher, 2021) and relatively weak attention has been given to the potential of the neural networks to speedup the ubiquitous Monte Carlo simulations (McNaughton *et al.*, 2020; Wu *et al.*, 2021). This work complements these efforts by demonstrating the first step in the direction of making the process of sampling from complex physics distributions efficient, focusing on building computationally cheap sampling distribution as a neural network surrogate.

CONCLUSION

I have studied an alternative approach to sampling of thermodynamic systems probability distribution functions. Ising model was used to demonstrate our proposed method of using neural network surrogate. This article main contribution is the demonstration that the thermodynamic probability distribution can be learned by a neural network and subsequent sampling of the distribution can be inexpensively sampled using the surrogate neural network. A particular drawback of this method is that system size is a fixed parameter of the training and therefore later sampling of the distribution can be restricting to similar system sizes without the need for re-training. Future work would focus on applying this method to simulating some of the challenging problems in materials research and living matter physics.

REFERENCES

Anderson, D. F., Higham, D. J., & Sun, Y. (2018). Computational Complexity Analysis for Monte Carlo Approximations of Classically Scaled Population Processes. *Multiscale Modeling & Simulation*, *16*(3), 1206–1226. [Crossref]

Barzegar, A., Pattison, C., Wang, W., & Katzgraber, H. G. (2018). Optimization of population annealing Monte Carlo for large-scale spin-glass simulations. *Physical Review E*, *98*(5), 053308. [Crossref]

Bayrakci, A. A., Demir, A., & Tasiran, S. (2010). Fast Monte Carlo Estimation of Timing Yield With Importance Sampling and Transistor-Level Circuit Simulation. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, *29*(9), 1328–1341. [Crossref]

Blöte, H. W. J., & Deng, Y. (2002). Cluster Monte Carlo simulation of the transverse Ising model. *Physical Review E*, *66*(6), 066110. [Crossref]

Botu, V., Batra, R., Chapman, J., & Ramprasad, R. (2017). Machine Learning Force Fields: Construction, Validation, and Outlook. *The Journal of Physical Chemistry C*, *121*(1), 511–522. [Crossref]

Broecker, P., Carrasquilla, J., Melko, R. G., & Trebst, S. (2017). Machine learning quantum phases of matter beyond the fermion sign problem. *Scientific Reports*, *7*(1), 8823. [Crossref]

Cai, S., Wang, Z., Wang, S., Perdikaris, P., & Karniadakis, G. E. (2021). Physics-Informed Neural Networks for Heat Transfer Problems. *Journal of Heat Transfer*, *143*(6), 060801. [Crossref]

Elvira, V., Martino, L., Luengo, D., & Bugallo, M. F. (2015). Efficient Multiple Importance Sampling Estimators. *IEEE Signal Processing Letters*, *22*(10), 1757–1761. [Crossref]

Gkeka, P., Stoltz, G., Barati Farimani, A., Belkacemi, Z., Ceriotti, M., Chodera, J. D., Dinner, A. R., Ferguson, A. L., Maillet, J.-B., Minoux, H., Peter, C., Pietrucci, F., Silveira, A., Tkatchenko, A., Trstanova, Z., Wiewiora, R., & Lelièvre, T. (2020). Machine Learning Force Fields and Coarse-Grained Variables in Molecular Dynamics: Application to Materials and Biological Systems. *Journal of Chemical Theory and Computation*, *16*(8), 4757–4775. [Crossref]

Husic, B. E., Charron, N. E., Lemm, D., Wang, J., Pérez, A., Majewski, M., Krämer, A., Chen, Y., Olsson, S., de Fabritiis, G., Noé, F., & Clementi, C. (2020). Coarse graining molecular dynamics with graph neural networks. *The Journal of Chemical Physics*, *153*(19), 194101. [Crossref]

Hylton, T. (2020). Thermodynamic Neural Network. *Entropy*, *22*(3), 256. [Crossref]

Kosmatopoulos, E. B., & Christodoulou, M. A. (1994). The Boltzmann g-RHONN: A learning machine for estimating unknown probability distributions. *Neural Networks*, *7*(2), 271–278. [Crossref]

Laubscher, R. (2021). Simulation of multi-species flow and heat transfer using physics-informed neural networks. *Physics of Fluids*, *33*(8), 087101. [Crossref]

Lee, E. M. Y., Ludwig, T., Yu, B., Singh, A. R., Gygi, F., Nørskov, J. K., & de Pablo, J. J. (2021). Neural Network Sampling of the Free Energy Landscape for Nitrogen Dissociation on Ruthenium. *The Journal of Physical Chemistry Letters*, *12*(11), 2954–2962. [Crossref]

MacKay, D. J. C. (1995). Developments in Probabilistic Modelling with Neural Networks—Ensemble Learning. In B. Kappen & S. Gielen (Eds.), *Neural Networks: Artificial Intelligence and Industrial Applications* (pp. 191–198). Springer London. [Crossref]

McNaughton, B., Milošević, M. V., Perali, A., & Pilati, S. (2020). Boosting Monte Carlo simulations of spin glasses using autoregressive neural networks. *Physical Review E*, *101*(5), 053312. [Crossref]

Nicoli, K. A., Anders, C., Funcke, L., Hartung, T., Jansen, K., Kessel, P., Nakajima, S., & Stornati, P. (2021a). Machine learning of thermodynamic observables in the presence of mode collapse. *ArXiv Preprint ArXiv:2111.11303*.

Nicoli, K. A., Anders, C. J., Funcke, L., Hartung, T., Jansen, K., Kessel, P., Nakajima, S., & Stornati, P. (2021b). Estimation of

Thermodynamic Observables in Lattice Field Theories with Deep Generative Models. *Physical Review Letters*, *126*(3), 032001. [Crossref]

Poltavsky, I., & Tkatchenko, A. (2021). Machine Learning Force Fields: Recent Advances and Remaining Challenges. *The Journal of Physical Chemistry Letters*, *12*(28), 6551–6564. [Crossref]

Shapiro, A. (2011). Computational Complexity of Stochastic Programming: Monte Carlo Sampling Approach. *Proceedings of the International Congress of Mathematicians 2010 (ICM 2010)*, 2979–2995. [Crossref]

Shukla, K., Jagtap, A. D., & Karniadakis, G. E. (2021). Parallel physics-informed neural networks via domain decomposition. *Journal of Computational Physics*, *447*, 110683. [Crossref]

Sultan, M. M., Wayment-Steele, H. K., & Pande, V. S. (2018). Transferable Neural Networks for Enhanced Sampling of Protein Dynamics. *Journal of Chemical Theory and Computation*, *14*(4), 1887–1894. [Crossref]

Unke, O. T., Chmiela, S., Sauceda, H. E., Gastegger, M., Poltavsky, I., Schütt, K. T., Tkatchenko, A., & Müller, K.-R. (2021). Machine Learning Force Fields. *Chemical Reviews*, *121*(16), 10142–10186. [Crossref]

Wang, Y., Ribeiro, J. M. L., & Tiwary, P. (2019). Past–future information bottleneck for sampling molecular reaction coordinate simultaneously with thermodynamics and kinetics. *Nature Communications*, *10*(1), 3573. [Crossref]

Wu, D., Rossi, R., & Carleo, G. (2021). Unbiased Monte Carlo cluster updates with autoregressive neural networks. *Physical Review Research*, *3*(4), L042024. [Crossref]