Bayesian Neural Networks for Boltzmann-Gibbs Distribution Generative Modelling

LIST CEA, LPTMS CNRS, UPS

Autumn 2023/2026

1 Context

Artificial neural networks correspond to highly parameterized differentiable non-linear functions that can, for instance, help defining probability density functions. The resulting parameterized distributions are known as deep generative models named after their ability to learn how to "generate" random instances. Several types of generative models have been developed corresponding to different combinations of parametric distributions and deep neural networks. Among the great diversity of existing models, we can mention mixture models, energy based models such as restricted Boltzmann machines, flow-based generative models such as normalizing flows, generative adversarial networks, autoregressive models, variational autoencoders, diffusion models and discrete indexed flows [1].

In the context of statistical physics and computational chemistry, deep generative models have been used to approximate the Boltzmann-Gibbs distribution [2] and perform tasks such as speeding up physical computations [3] or compute observables of physical systems [4, 5]. The parameters of these models are optimized either by density estimation and maximum likelihood estimation, i.e. training on data that is distributed according to the target distribution, or by variational inference where the target energy function is used to maximize the so-called Evidence Lower Bound (ELBO).

Both optimization targets seek at minimizing a loss function which can be interpreted as the minimization of a variational free energy [6] using the Bogoliubov inequality well known within the mean field approximation in statistical physics. Critically, deep generative models do not require a factorized partition function and can thus approximate distributions that are unreachable by traditional mean field approaches [7]. On the contrary, the bias introduced by the variational modelling is typically expected to vanish for a sufficiently large number of parameters. The challenge of these methods rather lies in the correct estimation of the parameters that should minimize the loss function and that cannot be computed analytically.

In practice, this minimization is carried out numerically with the help of automatic differentiation. Its outcome is a single set of parameters that is ob-

tained through a stochastic gradient descent. It is well known that the set of parameters obtained following this procedure does generally not correspond to the global minimum of the loss function. Despite the indisputable potential of such methods, neural networks are limited by the large number of free parameters to optimize that is accompanied by an increase of the training data required for good performance [8]. Moreover, one major weakness of deep models is the difficulty to estimate the uncertainty associated with their predictions that should be representative of the error introduced by replacing the target distribution with a sub-optimal variational one.

On the other hand, Bayesian inference is well known to be able to describe the epistemic uncertainty about parameters conditional on a collection of observed data. This conditional distribution over statistical parameters is known as the posterior distribution. It can be estimated using Variational Bayesian methods or by Bayesian Nonparametric Learning for instance. Nevertheless, the gold standard approach is still to use Markov Chain Monte Carlo methods to sample a Neural Network posterior distribution and using it, for instance, to compute the predictive posterior distribution. It is a notoriously difficult task due to the high-dimensionality and the numerous local minima of the likelihood function. Finding optimized transition kernels that can mitigate such an issue remains an open research topic, e.g. [9]. Bayesian inference applied to neural networks parameters is referred to as Bayesian Neural Networks.

2 PhD proposal

The overall goal of the PhD candidate will be to study the impact of Bayesian neural networks for generative modelling of a Boltzmann distribution. A first research axis will consist in developing methods able to sample neural posteriors using Markovian diffusion processes such as Brownian motion, Langevin diffusion in addition to HMC. Such methods will then be used to estimate the bias introduced by replacing a Bayesian neural network approximation in a physical numerical computation, by applying it for example on a quantum field theory or a spin system.

Secondly, Bayesian neural networks will be used to capture the epistemic uncertainty of the parameters of a surrogate variational distribution i.e. a deep generative model. Lastly, the PhD student will study the impact of a model combining both the physical aleatoric and Bayesian epistemic uncertainties for physical simulation, in particular on Monte Carlo simulations, with application in statistical physics and material science.

3 Profile

The candidate should have a strong theoretical and numerical interest and a special care about software development. A Master's degree in probability theory & machine learning and/or theoretical physics & statistical physics is required.

Programming experience in python is a distinct advantage. Fluency in French is not required.

4 Supervision & Application

Supervisors - **Eiji Kawasaki**, Eric Barat CEA LIST
Director - Alberto Rosso LPTMS CNRS

Please email your resume, cover letter and any letters of recommendation to eiji.kawasaki@cea.fr

References

[1] Elouan Argouarc'h, François Desbouvries, Eric Barat, Eiji Kawasaki, and Thomas Dautremer. Discretely Indexed Flows, April 2022. Number: arXiv:2204.01361 arXiv:2204.01361 [cs, stat].

[2] Frank Noé, Simon Olsson, Jonas Köhler, and Hao Wu. Boltzmann generators: Sampling equilibrium states of many-body systems with deep learning. *Science*, 365(6457):eaaw1147, September 2019.

[3] Li Huang and Lei Wang. Accelerated Monte Carlo simulations with restricted Boltzmann machines. *Physical Review B*, 95(3):035105, January 2017.

[4] Dian Wu, Lei Wang, and Pan Zhang. Solving Statistical Mechanics Using Variational Autoregressive Networks. *Physical Review Letters*, 122(8):080602, February 2019.

[5] Maciej J. Karcz, Luca Messina, Eiji Kawasaki, Serenah Rajaonson, Didier Bathellier, and Emeric Bourasseau. Semi-supervised generative approach to point-defect formation in chemically disordered compounds: application to uranium-plutonium mixed oxides, November 2022. Number: arXiv:2211.12086 arXiv:2211.12086 [cond-mat].

[6] Diederik P. Kingma. Variational inference & deep learning: A new synthesis. 2017.

[7] George T. Cantwell. Approximate sampling and estimation of partition functions using neural networks, September 2022. Number: arXiv:2209.10423 arXiv:2209.10423 [cs, stat].

[8] Javad Komijani and Marina K. Marinkovic. Generative models for scalar field theories: how to deal with poor scaling?, January 2023. Number: arXiv:2301.01504 arXiv:2301.01504 [hep-lat].

[9] Eiji Kawasaki and Markus Holzmann. Data Subsampling for Bayesian Neural Networks, October 2022. Number: arXiv:2210.09141 arXiv:2210.09141 [cs, stat].