KIM REVIEW
https://kimreview.org

Volume 2, Article 03, 2024
DOI:10.25950/bfa99422

Boltzmann Generators and the New Frontier of Computational Sampling in Many-Body Systems

Commentary by

Alessandro Coretti [*]
Faculty of Physics, University of Vienna, Austria
Sebastian Falkner [*]
Faculty of Physics, University of Vienna, Austria
Jan Weinreich [*]
Faculty of Physics, University of Vienna, Austria
Christoph Dellago
Faculty of Physics, University of Vienna, Austria
O. Anatole von Lilienfeld[†]
Vector Institute; Departments of Chemistry and Materials Science and Engineering, University of Toronto, Canada

on

Boltzmann generators: Sampling equilibrium states of many-body systems with deep learning
Frank Noé, *et al.*, Science, 365:6457 (2019)

Received November 29, 2023
Published April 22, 2024

Statement of Significance

The paper by Noé *et al.* [1] introduced the concept of Boltzmann Generators (BGs), a deep generative model that can produce unbiased independent samples of many-body systems. They can generate equilibrium configurations from different metastable states, compute relative stabilities between different structures of proteins or other organic molecules, and discover new states. In this commentary, we motivate the necessity for a new generation of sampling methods beyond molecular dynamics, explain the methodology, and give our perspective on the future role of BGs.

Background

Before delving into the discussion of the paper by Noé *et al.* [1], it is essential to first outline the main challenge it seeks to address. Within the domain of numerical atomistic simulations, two significant issues frequently dominate computational complexity: the first is the computational "curse" of solving the electronic Schrödinger equation, prohibiting chemically accurate *first principles* investigations of large molecules. The second is the so-called sampling problem: Even when using predictive machine learned potentials, i.e. data-driven and cost-effective approximations of the electronic potential, or more conventional force fields, it is impossible to reach the timescales necessary for many chemical and biological processes. While machine learned energies [2], or forces [3–6] recover even highly accurate quantum labels orders of magnitude faster than numerical solutions

[*] These authors contributed equally
[†] anatole.vonlilienfeld@utoronto.ca

of approximate variants of Schrödinger's equation, they can still be substantially slower than traditional force fields [7–10]. Furthermore, the sampling problem of *uncorrelated* physical configurations within statistical mechanics ensembles remains. This latter problem is closely intertwined with computing free energies that govern the phase diagrams of condensed matter. To achieve that, sufficient coverage of uncorrelated configurations must be guaranteed. However, the challenge is that directly simulating trajectories of the many atoms that make up materials and molecules in order to integrate Newton's equations of motion, and to compute their essential properties for most relevant time scales, is computationally prohibitive, exceeding even the capabilities of supercomputers.

Using generative deep neural networks, Noé *et al.* tackled the sampling problem from a novel direction in 2019. From their first appearance, the use of generative neural networks has been tempting in statistical mechanics, due to their ability to produce independent samples from a given distribution. Provided sufficient training data, this could effectively overcome some of the most challenging problems of standard sampling algorithms, commonly used in statistical mechanics, in particular correlations between subsequently sampled states. Generative models have originally been developed in the realm of image/text/audio generation, where a large set of examples is available and no analytical form of the target distribution exists. Within the realm of the atomistic sciences, they were introduced for the purpose of molecular materials design already one year earlier in 2018 [11]. The physical sampling problem, however, is profoundly different since the exact target distribution is known (up to a proportionality constant) and since it is crucial to sample the target distribution exactly, to avoid any bias in the result of the numerical simulation. The contribution of Noé *et al.* was also contextualized in the same issue within a perspective by Tuckerman [12].

Normalizing flows are a particular class of deep generative models well suited to accommodate these different premises. First, they can be trained by exploiting the analytical likelihood of the target space to sample from. Second, the architecture of the network allows to analytically compute the likelihood of a generated sample. This allows for the generation of a fully unbiased distribution in target space. These two features make normalizing flows a very promising tool for tackling the sampling problem of physical configurations. Incidentally, related work was published two weeks earlier in *Phys Rev D* on flow-based generative models for Markov chain Monte Carlo in lattice field theory [13]. Corresponding successful applications have been reported more recently, [14–17] and also for the calculation of free energies [18, 19]. In their paper, Noé *et al.* have adapted the architecture of normalizing flows to the physical sampling problem, using the energy of the target system as the likelihood for training the model. They then introduce Boltzmann generators (BGs), which are normalizing flows aimed at solving the sampling problem of statistical mechanics.

Summary of the article

The architecture of BGs is built on normalizing flows: The idea is to train a deep neural network to approximate a transformation from a "latent" space z to a target space x such that

$$x = F_{\mathrm{zx}}(z) \tag{1}$$

where z and x are samples from the spaces z and x with distributions $\mu_{\mathrm{z}}(z)$ and $\mu_{\mathrm{x}}(x)$, respectively. The distribution of the latent space, which is sometimes called "prior", is usually chosen to be very simple (Gaussian or even uniform). In this way, once the network is trained, it is possible to easily get samples from z and to transform them, via F_{zx}, to samples distributed according to μ_{x}. Choosing the target distribution $\mu_{\mathrm{x}}(x) = e^{-\beta U(x)}$, the BG samples the NVT ensemble of the system defined by the potential energy $U(x)$ at inverse temperature, $\beta = (k_B T)^{-1}$.

The neural network representing the transformation F is built to be invertible. In this way, one can also compute $z = F_{\mathrm{xz}}(x)$ where $F_{\mathrm{xz}} = F_{\mathrm{zx}}^{-1}$. The invertibility of the transformation F guarantees that, given the likelihood of a sample in latent space $\mu_{\mathrm{z}}(z)$, it is possible to exactly compute the corresponding likelihood in the target space after the transformation as $p_{\mathrm{x}}(x) = \mu_{\mathrm{z}}(z) \det |J_{\mathrm{zx}}(z)|^{-1}$ where J_{zx} is the Jacobian matrix associated to F_{zx}. The same is true for the inverse transformation which yields $p_{\mathrm{z}}(z) = \mu_{\mathrm{x}}(x) \det |J_{\mathrm{xz}}(x)|^{-1}$. It is important to highlight the difference between the distributions μ and p. The former corresponds to the *true* distribution of samples in space, while the latter is the distribution of samples that is generated by the network. The two will be different in general.

The goal of making the target distributions μ_{x} and p_{x} as close as possible provides a natural way of training the network. The Kullback-Liebler divergence as a “distance” in distribution space provides the loss function

$$J_{\mathrm{KL}} = \mathbb{E}_{\mathrm{z}}\big[\beta U(F_{\mathrm{zx}}(z)) - \log(\det J_{\mathrm{zx}}(z))\big] \tag{2}$$

where $\mathbb{E}_{\mathrm{z}}$ is the mean value over a batch of samples from z. This loss function corresponds to the free-energy difference between the prior and the target distributions. Minimizing Eq. (2) corresponds to a) minimizing the internal energy (first term) and therefore training the network to sample low-energy configurations, and b) maximizing the entropy of the target distribution at the given temperature (second term) and therefore avoiding mode-collapse in the lowest energy configuration. The invertibility of the network also allows for training in the other direction. This proves to be particularly useful to initialize the flow at the beginning of the training process. Given a subset of initial configurations from the target space sampled by means of standard algorithms such as Molecular Dynamics (MD) or Monte Carlo (MC), one can maximize the likelihood in the distribution of latent space (from here on assumed to be normal) of samples transformed via $F_{\mathrm{xz}}(x)$. This yields the loss function

$$J_{\mathrm{ML}} = \mathbb{E}_{\mathrm{x}}\left[\frac{1}{2}\|F_{\mathrm{xz}}(x)\|^2 - \log(\det J_{\mathrm{xz}}(x))\right] \tag{3}$$

This second loss function is crucial when multiple minima are present in the potential energy surface of the target system. In this case, the entropic term in Eq. (2) alone is not sufficient to avoid the collapse of the generated configurations around a single minimum. Eqs. (2) and (3) represent the two main terms of the total loss function used for training and they are in practice often combined in a single loss function (sometimes with different weights). Also additional, system-specific terms can be added at convenience. In particular, the authors discuss a Reaction Coordinate (RC) loss that can be optionally included to force the system to generate configurations closer to energy barriers.

Finally, knowledge of the likelihood of a transformed sample allows for the removal of any residual bias in the generated distribution. For example, a weight $\omega(x) = \mu_{\mathrm{x}}(x)/p_{\mathrm{x}}(x)$ can be assigned to every generated configuration x and statistical mechanics estimators can be computed as

$$A = \frac{\Sigma_i \omega(x_i) a(x_i)}{\Sigma_i \omega(x_i)} \tag{4}$$

where A is an unbiased estimator of an observable $\langle A \rangle$.

Given this basic architecture, the authors also provide an algorithm to explore the target space while training the BG. Starting from a pool of known physical configurations, they perform a MC simulation in latent space and progressively add more configurations to the initial pool. Training the flow with the newly added configurations as the exploration proceeds, guarantees that the mapping to the latent space is always consistent with the currently explored physical space. The acceptance criterion of the MC scheme ensures that the generated samples approach the correct Boltzmann distribution.

The authors show applications of the proposed methods to a wide range of cases, from proof-of-principle calculations for systems with two degrees of freedom to complex biochemical systems such as bovine pancreatic trypsin inhibitor (BPTI) protein, passing through a two dimensional example of a condensed matter system.

Commentary and critical analysis

From the dawn of molecular simulations in the 1950s, the introduction of BGs represents one of the few attempts at a paradigm shift in the calculation of statistical mechanics observables via numerical experiments. Most of the efforts in the development of methods for generating physical configurations are focused on the improvement of the Markov chain upon which the methods are based. Enhanced sampling algorithms, replica exchange methods, and multi-scale techniques (just to name a few) allowed tremendous advances in the field, but none of these method drastically changed the way in which sampling is carried out, i.e. via sequential updates of configurations. Boltzmann generators tackle the sampling problem from a different angle, making the most out of the increase in computational power provided by the latest advances in deep learning.

Due to this change of perspective, assessing the performances of BGs with respect to standard sampling algorithms is not straightforward. The generation of configurations via BGs is astonishingly fast. The time needed to transform samples from latent space is many orders of magnitude lower than the time needed for producing the same number of independent configurations using standard Markov-chain-based methods. This comes at the cost of a painful and time-consuming training process, which involves fine-tuning different hyperparameters and a large number of energy evaluations. Even if it is true that the training process is a one-time procedure, it is also true that, for the time being, there is only a limited transferability from system to system. In practice, it has been observed that the performance gain obtained using BGs tends to be strongly dependent on the system under investigation. The training process is complicated by complex energy landscapes and a large number of degrees of freedom as this task usually requires bigger networks and larger training sets. On the other hand, strong correlations between sampled configurations and the presence of high energy barriers and basins of attraction in configuration space can become problematic for standard methods and time needed to decorrelate or to "jump" out of such basins can quickly amortize for the training time.

Some limitations in the applicability of the methods also arise from the relatively new architecture of normalizing flows underpinning BGs. The network's invertibility prevents a significant reduction in the number of degrees of freedom that the transformation must act upon. The treatment of large systems is therefore hampered by the inherently global nature of the transformation which limits its expressiveness. This can be particularly limiting for biochemical and biological systems in which solvation effects are important and for which good models of implicit solvents are not available. A high number of degrees of freedom is also expected to have an impact on the difficulty of the training process and on the reweighting algorithm. The latter, in particular, to be effective requires a degree of superposition between the target and generated distributions which scales exponentially with the number of degrees of freedom.

Our experience [20, 21] has also shown that increasing the "power" of the network, i.e. the ability to faithfully reproduce a given complex distribution, is not so straightforward as in other neural network applications. Oftentimes, the mantra *bigger is better* does not apply to normalizing flows, and increasing the network size has only a mild effect or indeed no effect on the quality of the training. What seems to be effective in increasing the efficiency of the training is a different and somewhat smarter representation of the input. A very good example is represented by the mixed-coordinate transformation the authors used to treat the BPTI protein. The designing of smart transformation layers (usually non-learnable layers that are placed between the raw input, e.g. the $3N$-dimensional array containing the particle coordinates, and the normalizing flow), can make a difference in the training process.

While BGs, for the time being, are still too "young" to provide a comparable alternative to standard sampling methods, in particular for big and complex systems, we see the enormous potential that the introduction of BGs represents for the molecular simulation community. On the other hand, the use of BGs in conjunction with other standard sampling algorithms, in the spirit of the exploration method proposed by the authors, can already drastically improve the efficiency of molecular simulations today. One clear example is the use of BGs to propose smart Monte Carlo moves [22, 23].

Further progress in the field of generative models is expected to mitigate many of the remaining issues of the method and eventually make it the new standard for generating equilibrium configurations of statistical mechanics systems.

Potential new directions

Many research directions stemmed, directly or indirectly, from the publication of BGs in [1] and from the efforts of the community to improve on the baseline provided by this paper. Some of these efforts are discussed and referenced below.

Soon after the publication of [1], the same group explored different techniques that could improve on previous results. For example, in Ref. 24 they designed a BG which could be trained to sample exactly many thermodynamics states with a single training. This was achieved by parameterizing the generated distribution through the temperature of the target ensemble. To this end, they employed a different architecture with respect to the original paper: they augmented the physical space by doubling the degrees of freedom involved in the transformation and they designed the architecture of the network to be volume preserving, (i.e. $\det|J| = 1$).

In a different paper [25] they introduced stochastic dynamics between the deterministic blocks of the normalizing flow. They show how to overcome, in this way, possible topological constraints in the target density and increase the expressiveness of the network in reproducing complex target distributions. Moreover, they show how the combined optimization of the network parameters and of the stochastic sampler also improves the efficiency of the latter. This approach has been pursued also by other groups in different fields of statistical mechanics [26, 27].

The problem of topological constraints is tackled by a different perspective in a later paper [28], where the condition of the smoothness of the flow is also preserved and the gradient computation is also addressed for these cases. In the same paper, the possibility of using forces in the target space $\mathcal{F}(x)$ to improve training is also introduced. The authors show how the addition in the loss function of a Force Matching (FM) term

$$J_{\mathrm{FM}} = \mathbb{E}_{\mathrm{x}}\left[\|\mathcal{F}(x) - \nabla_x \log p_{\mathrm{x}}(x)\|^2\right] \tag{5}$$

can drastically improve the training efficiency while compensating for the additional burden of computing force on data in target space.

In a more recent paper [29], the power of continuous normalizing flows [30, 31] has been combined with the simplified training process arising from flow matching [32, 33] and with the possibility of encoding some of the symmetries of the system in the transformation itself via the design of equivariant flows [34, 35]. Rotational symmetries are also discussed using quaternions in Ref. 36.

The training process can be facilitated by the introduction of physical information in the prior distribution. Wirnsberger *et al.* [37] generated configurations for a solid using normalizing flows, relying on a prior distribution that consisted in Gaussians placed at the lattice sites of the crystalline phase of the target system. In this way, they managed to generate very accurate configurations with relatively little effort and without the need for reweighting. Moreover, they managed to achieve this result with no reference structures from the target system. Coretti *et al.* [21] followed a similar line of research for liquid systems, generating liquid configurations using simulations with a simpler potential energy at a higher temperature as the prior distribution.

By integrating physical priors about the intrinsic distribution of internal coordinates, new approaches could capitalize on a) the inherently Gaussian characteristics of bonded interactions derived from the harmonic potential of bonds, and b) the specific distributions of non-bonded atom pairs. Embedding such physical priors to intelligently parameterize the learning model could reduce it to just a handful of parameters for each type of interaction. For example, in Ref. 38, a model was trained to predict the average geometry of small molecules at finite temperatures using a single machine learning model for each bond pair.

The use of BGs is also promising in relation to the study of rare events. Falkner *et al.* [20] produced a version of BGs that can be conditioned to generate configurations biased along a reaction coordinate in target space. They show how this is very powerful either for computing free energy or for producing shooting points for transition path algorithms.

Statistical ensembles other than NVT have also been investigated by different research groups: Wirnsberger *et al.* [39] and van Leeuwen *et al.* [40], almost at the same time, came out with an algorithm for the generation of configurations in the isobaric-isothermal ensemble whose main features are based on BGs.

Finally, a word on research directions that have not yet been addressed but where normalizing flows could make a positive impact: In situations where an analytical expression for the nonequilibrium distribution is present, the use of BGs to sample nonequilibrium configurations could be of interest to many applications. For instance, Quantum Monte Carlo or Path Integrals could also be among the applications that could benefit from machines which generate one-shot configurations out of a given distribution [41].

Acknowledgment

O.A.v.L. has received support as the Ed Clark Chair of Advanced Materials and as a Canada CIFAR AI Chair. A.C., S.F. and C.D. acknowledge support of the Austrian Science Fund (FWF) through the SFB TACO, grant number F 81-N.

References

[1] Noé, F., Olsson, S., Köhler, J. and Wu, H., Boltzmann generators: Sampling equilibrium states of many-body systems with deep learning, *Science*, 365:eaaw1147 (2019).

[2] Huang, B. and von Lilienfeld, O. A., Ab initio machine learning in chemical compound space, *Chemical Reviews*, 121:10001 (2021).

[3] Bartók, A., Csányi, G., Gaussian approximation potentials: A brief tutorial introduction, *International Journal of Quantum Chemistry*, 115:1051 (2015).

[4] Behler, J. Four Generations of High-Dimensional Neural Network Potentials, *Chemical Reviews*, 121:10037–10072 (2021).

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

[6] Käser, S., Vazquez-Salazar, L. I., Meuwly, M., Töpfer, K. Neural network potentials for chemistry: concepts, applications and prospects, *Digital Discovery*, 2:28–58 (2023).

[7] Wang, J., Wang, W., Kollman, P. A., Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations, *J. Mol. Graph. Model.*, 25:247–260 (2006).

[8] Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. Development and testing of a general amber force field, *J. Comp. Chem.*, 25:1157–1174 (2004).

[9] Qiu, Y. et al. Development and Benchmarking of Open Force Field v1.0.0—the Parsley Small-Molecule Force Field, *Journal of Chemical Theory and Computation*, 17:6262–6280 (2021).

[10] Bjelkmar, P., Larsson, P., Cuendet, M. A., Hess, B., Lindahl, E. Implementation of the CHARMM Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models, *Journal of Chemical Theory and Computation*, 6:459–466 (2010).

[11] Gómez-Bombarelli, R., Wei, J. N., Duvenaud, D., Hernández-Lobato, J. M., Sánchez-Lengeling, B., Sheberla, D., Aguilera-Iparraguirre, J., Hirzel, T. D., Adams, R. P., Aspuru-Guzik, A. Automatic chemical design using a data-driven continuous representation of molecules, *ACS central science*, 4:268–276 (2018).

[12] Tuckerman, M. E. Machine learning transforms how microstates are sampled, *Science*, 365:982–983 (2019).

[13] Albergo, M. S., Kanwar, G., Shanahan, P. E. Flow-based generative models for Markov chain Monte Carlo in lattice field theory, *Physical Review D*, 100:034515 (2019).

[14] Nicoli, K. A., Anders, C. J., Funcke, L., Hartung, T., Jansen, K., Kessel, P., Nakajima, S., Stornati, P. Estimation of thermodynamic observables in lattice field theories with deep generative models, *Physical review letters*, 126:032001 (2021).

[15] Nicoli, K. A., Nakajima, S., Strodthoff, N., Samek, W., Müller, K.-R., Kessel, P. Asymptotically unbiased estimation of physical observables with neural samplers, *Physical Review E*, 101:023304 (2020).

[16] Singha, A., Chakrabarti, D., Arora, V. Conditional normalizing flow for Markov chain Monte Carlo sampling in the critical region of lattice field theory, *Physical Review D*, 107:014512 (2023).

[17] Gabrié, M., Rotskoff, G. M., Vanden-Eijnden, E. Adaptive Monte Carlo augmented with normalizing flows, *Proceedings of the National Academy of Sciences*, 119:e2109420119 (2022).

[18] Ahmad, R., Cai, W. Free energy calculation of crystalline solids using normalizing flows, *Modelling and Simulation in Materials Science and Engineering*, 30:065007 (2022).

[19] Wirnsberger, P., Ballard, A. J., Papamakarios, G., Abercrombie, S., Racaniére, S., Pritzel, A., Jimenez Rezende, D., Blundell, C. Targeted free energy estimation via learned mappings, *The Journal of Chemical Physics*, 153 (2020).

[20] Falkner, S., Coretti, A., Romano, S., Geissler, P. L., Dellago, C. Conditioning Boltzmann generators for rare event sampling, *Machine Learning: Science and Technology*, 4:035050 (2023).

[21] Coretti, A., Falkner, S., Geissler, P., Dellago, C. Learning mappings between equilibrium states of liquid systems using normalizing flows, arXiv preprint, arXiv:2208.10420 (2022).

[22] Sbailò, L., Dibak, M., Noé, F. Neural mode jump monte carlo, *The Journal of Chemical Physics*, 154 (2021).

[23] Invernizzi, M., Krämer, A., Clementi, C., Noé, F. Skipping the replica exchange ladder with normalizing flows, *The Journal of Physical Chemistry Letters*, 13:11643–11649 (2022).

[24] Dibak, M., Klein, L., Krämer, A., Noé, F. Temperature steerable flows and Boltzmann generators, *Phys. Rev. Res.*, 4:L042005 (2022).

[25] Wu, H., Köhler, J., Noé, F. Stochastic normalizing flows, *Advances in Neural Information Processing Systems*, 33:5933–5944 (2020).

[26] Caselle, M., Cellini, E., Nada, A., Panero, M. Stochastic normalizing flows for lattice field theory, arXiv preprint, arXiv:2210.03139 (2022).

[27] Caselle, M., Cellini, E., Nada, A., Panero, M. Stochastic normalizing flows as non-equilibrium transformations, *Journal of High Energy Physics*, 1–31 (2022).

[28] Köhler, J., Krämer, A., Noé, F. Smooth normalizing flows, *Advances in Neural Information Processing Systems*, 34:2796–2809 (2021).

[29] Klein, L., Krämer, A., Noé, F. Equivariant flow matching, arXiv preprint, arXiv:2306.15030 (2023).

[30] Chen, R. T., Rubanova, Y., Bettencourt, J., Duvenaud, D. K., Neural ordinary differential equations, *Advances in neural information processing systems*, 31 (2018).

[31] Grathwohl, W., Chen, R. T., Bettencourt, J., Duvenaud, D., Scalable reversible generative models with free-form continuous dynamics, *International Conference on Learning Representations* (2019) p 7.

[32] Lipman, Y., Chen, R. T., Ben-Hamu, H., Nickel, M., Le, M., Flow matching for generative modeling, arXiv preprint, arXiv:2210.02747 (2022).

[33] Tong, A., Malkin, N., Huguet, G., Zhang, Y., Rector-Brooks, J., Fatras, K., Wolf, G., Bengio, Y., Conditional flow matching: Simulation-free dynamic optimal transport, arXiv preprint arXiv:2302.00482 (2023).

[34] Rezende, D. J., Racanière, S., Higgins, I., Toth, P. Equivariant hamiltonian flows, arXiv preprint, arXiv:1909.13739 (2019).

[35] Köhler, J., Klein, L., Noé, F., Equivariant flows: exact likelihood generative learning for symmetric densities, *International conference on machine learning* (2020) pp 5361–5370.

[36] Köhler, J., Invernizzi, M., De Haan, P., Noé, F. Rigid body flows for sampling molecular crystal structures, arXiv preprint, arXiv:2301.11355 (2023).

[37] Wirnsberger, P., Papamakarios, G., Ibarz, B., Racanière, S., Ballard, A. J., Pritzel, A., Blundell, C. Normalizing flows for atomic solids, *Machine Learning: Science and Technology*, 3:025009 (2022).

[38] Weinreich, J., Lemm, D., von Rudorff, G. F., von Lilienfeld, O. A. Ab initio machine learning of phase space averages, *The Journal of Chemical Physics*, 157:024303 (2022).

[39] Wirnsberger, P., Ibarz, B., Papamakarios, G. Estimating Gibbs free energies via isobaric-isothermal flows, *Machine Learning: Science and Technology*, 4:035039 (2023).

[40] van Leeuwen, S. and Ortíz, A. P. d. A. and Dijkstra, M. A, Boltzmann generator for the isobaric-isothermal ensemble, arXiv preprint, arXiv:2305.08483 (2023).

[41] Tuckerman, M. E. Statistical Mechanics: Theory and Molecular Simulation, *Oxford University Press* (2010), p 712.