UNIVERSITY OF OXFORD

DEPARTMENT OF COMPUTER SCIENCE

Deep Potential Moelcular Dynamics: a scalable model with the accuracy of quantum mechanics - A summary

Coursework 1 of Topics in Computational Biology

submitted by

Jonas Bernhard Wildberger

Student Number: 1335851

January 31, 2021

Summary

A substantial drawback of Molecular Dynamics simulations is the trade-off between accuracy and efficiency. *Ab initio* molecular dynamics (AIMD) yields the most accurate results but is limited to space- and time scales too small for most applications. On the other hand, the deployment of empirical force fields enables the study of larger scales but is approximative in nature and therefore lacking the accuracy of AIMD.

In their paper [ZHW+18] L. Zhang et al. present the Deep Potential Molecular Dynamics (DeePMD) method for molecular simulations addressing this trade-off. This method leverages Machine Learning (ML) to learn the functional dependence of the potential on the atomic configurations and consolidate the benefits of the two approaches. Contrary to former ML approaches, DeePMD does not make use of auxiliary quantities like symmetry functions or the Coulomb matrix in order to preserve the symmetry of the system under translations, rotations and permutations. Instead, it makes use of the extensive character of the potential energy to assign each atom a local reference frame contributing to one summand E_i of the potential energy $E = \sum_i E_i$. The scheme to find E_i can be split up into two steps.

1. In this first step, the inputs for the neural network (NN) are generated. This is done by setting up a coordinate system for each atom i with the atom as its center. Then, for each neighbor atom j within a preset radius R_c, the reciprocal distances[1] are calculated and combined to the ith input vector of the NN in step 2. The idea behind this representation is the *Embedded Atom Concept* stating that the energy E_i only depends on the distance between to neighboring atoms.

2. The vectors calculated in step 1 are now input in a fully connected NN consisting of 5 hidden layers with 240, 120, 60, 30, 10 units respectively from the innermost to the outermost layer. It uses the hyperbolic tangent as an activation function for all hidden layers and the identity activation for the output of the NN from the last hidden layer. The weights and biases are adapted by the Adam method [KB17] with an exponentially decaying learning rate, which is a variant of stochastic gradient descent, applied to the parameterised loss functions:

$$\mathcal{L}(p_\varepsilon, p_f, p_\xi) = p_\varepsilon \Delta \varepsilon^2 + \frac{p_f}{3N} \sum_i |\Delta \vec{F}_i|^2 + \frac{p_\xi}{9} \|\Delta \Xi\|^2$$

   Δ indicates the difference between the prediction and the training data, N is the number of atoms, ε is the energy per atom, $\vec{F}_i$ is the force on atom i and ξ is the virial tensor $\Xi = -\frac{1}{2} \sum_i \vec{R}_i \otimes \vec{F}_i$. The parameters p_ε and p_ξ are steadily increased during training whereas p_f is decreased. Including the forces and the virial tensor in the loss function serves as a regularization of the network to prevent overfitting and to reduce the training time. The denominators in the fractions normalise the respective contributions based on

---

[1]The reciprocal distances can also be supplemented by additional angular information

the number of terms in each part.

The main advantage of the approach by L. Zhang et al. is its accuracy without sacrificing scalability. In the case of water the DeePMD method achieves similar performance to the original AIMD simulations and for molecules the results were slightly better than the GDML benchmark. Since the physical quantities are all sums of local contributions, the computational costs increase only linearly with the number of atoms and the calculations can be easily parallelised. This allows for accurate MD simulations on larger timescales and with more atoms. Additionally, the model benefits from not requiring any auxiliary quantities as opposed to other machine learning approaches like [CTS+17] or [BP07], which simplifies the scheme.

One drawback of this local decomposition is the emergence of discontinuities in the outputs from the restriction to balls of a fixed radius. Furthermore, the model's ability to represent long-range interactions is in question. These contribute only implicitly to the energy from interactions within the cutoff distance. S. Yue et al. [YMA+21] investigate the impact this shortcoming and conclude that it strongly affects for example vapor phases properties, such as densities and second virial coefficients. Recent efforts to include long-range interactions comprise among others a decomposition of energy and forces into electrostatic and non-electrostatic contributions [YMA+21]. However, this approach results in 2x higher computational costs. Thus, the inclusion of long-range interactions without sacrificing efficiency is still topic of ongoing research.

REFERENCES

[BP07]   Jörg Behler and Michele Parrinello. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.*, 98:146401, Apr 2007.

[CTS⁺17]   Stefan Chmiela, Alexandre Tkatchenko, Huziel E. Sauceda, Igor Poltavsky, Kristof T. Schütt, and Klaus-Robert Müller. Machine learning of accurate energy-conserving molecular force fields. *Science Advances*, 3(5):e1603015, May 2017.

[KB17]   Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2017.

[YMA⁺21]   Shuwen Yue, Maria Muniz, Marcos Andrade, Linfeng Zhang, Roberto Car, and Athanassios Panagiotopoulos. When do short-range atomistic machine-learning models fall short? *The Journal of Chemical Physics*, 154:034111, 01 2021.

[ZHW⁺18]   Linfeng Zhang, Jiequn Han, Han Wang, Roberto Car, and Weinan E. Deep potential molecular dynamics: A scalable model with the accuracy of quantum mechanics. *Phys. Rev. Lett.*, 120:143001, Apr 2018.