JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

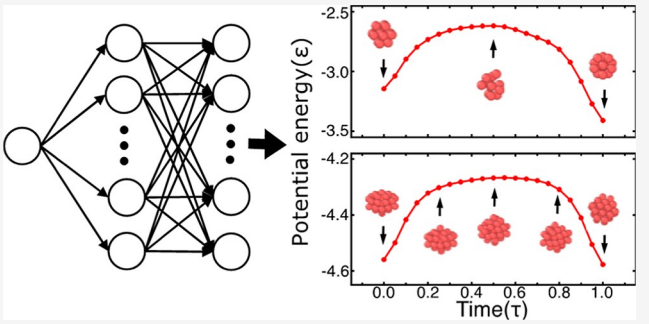
Neural Network for Principle of Least Action

Beibei Wang,* Shane Jackson, Aiichiro Nakano, Ken-ichi Nomura, Priya Vashishta, Rajiv Kalia, and Mark Stevens

Cite This: J. Chem. Inf. Model. 2022, 62, 3346−3351 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The principle of least action is the cornerstone of classical mechanics, theory of relativity, quantum mechanics, and thermodynamics. Here, we describe how a neural network (NN) learns to find the trajectory for a Lennard-Jones (LJ) system that maintains balance in minimizing the Onsager−Machlup (OM) action and maintaining the energy conservation. The phase-space trajectory thus calculated is in excellent agreement with the corresponding results from the "ground-truth" molecular dynamics (MD) simulation. Furthermore, we show that the NN can easily find structural transformation pathways for LJ clusters, for example, the basin-hopping transformation of an LJ_{38} from an incomplete Mackay icosahedron to a truncated face-centered cubic octahedron. Unlike MD, the NN computes atomic trajectories over the entire temporal domain in one fell swoop, and the NN time step is a factor of 20 larger than the MD time step. The NN approach to OM action is quite general and can be adapted to model morphometrics in a variety of applications.



1. INTRODUCTION

The principle of least action is a foundational law of physics. It pervades classical mechanics,[1] theory of relativity,[2] quantum mechanics,[3] and thermodynamics.[4] The action is defined as

$$A = \int_{t_0}^{t_1} L(q(t),\ \dot{q}(t))\ \mathrm{d}t \tag{1}$$

where L is the Lagrangian in classical mechanics and $q(t)$ and $\dot{q}(t)$ are the system trajectory and its time derivative, respectively. Equation 1 embodies a boundary-value problem in which A is minimized with respect to $q(t)$, subject to the constraints imposed by the initial and final configurations at time t_0 and t_1, respectively.[5] The differential equations arising from the principle of least action are the Euler−Lagrange equations, which constitute an initial-value problem. These equations are at the core of molecular dynamics (MD), the preeminent simulation approach in physics, chemistry, materials science, and biology.[6] The essential input to MD is interatomic potential energy from which forces are calculated and the equations of motion are integrated over discretized time with a finite-difference scheme.[7] The output is phase-space trajectories $\{q(t),\ \dot{q}(t)\}$ from which structural, mechanical, thermodynamic, and dynamical properties of the system can be computed and compared with experimental measurements.

In recent years, there has been a surge of interest in applying machine learning (ML) tools in MD tasks. Atomic force fields have been developed using neural networks trained by data from quantum-mechanical calculations.[8,9] Predictive models based on kernel ridge regression, support vector machine, random forest, and other techniques have been employed to predict material properties such as band gaps,[10] elastic constants,[11] dielectric constants, and thermoelectric properties.[12−14] For example, Bayesian optimization methods have been used to discover optimal layered materials for targeted properties such as electronic band structure and thermal-transport coefficients.[15] On the other hand, original ML models have been built to learn the basic laws of physics for simple systems such as a mass-spring, a double pendulum, and even the case of light refraction.[16−18]

In this paper, we explore the application of ML in a different context, that is, how a NN learns the principle of least action and provides atomic trajectories by minimizing the Onsager−Machlup (OM) action[19]

$$S_{\mathrm{OM}} = \int_0^T \left[\sum_{i=0}^{N} m_i \ddot{\vec{q}}_i(t) + \frac{\partial V(Q(t))^2}{\partial \vec{q}_i} \right] \mathrm{d}t \tag{2}$$

where m_i is the mass and $\ddot{\vec{q}}_i(t)$ is the second-order time derivative of the position vector $\vec{q}_i(t)$ of the ith particle. In eq 2,

Received: April 29, 2022
Published: July 5, 2022

$Q(t)$ collectively represents the coordinates of all the particles in the system at time t and $V(Q(t))$ is the potential energy of the system. The genesis of the OM action was Onsager's seminal papers in which he developed reciprocal relations for irreversible transport process and proposed that the probability of paths of a diffusion process is exponentially small with the exponent proportional to the time integral of dissipation function of the history.[20−22] Later on, Onsager and Machlup showed that minimization of eq 2 would produce the most probable trajectory.[20] Action formulation is a popular optimization approach to atomistic boundary-value problems since a global minimum is guaranteed.[23] A straightforward minimization of the discretized OM action involves second-order derivatives which are computationally expensive to calculate for large systems. Passerone and Parrinello add the energy-conservation constraint to the OM action and minimize the resulting function by making the following transformation

$$\boldsymbol{q}_j = \boldsymbol{q}_A + \frac{j\Delta}{\tau}(\boldsymbol{q}_B - \boldsymbol{q}_A) + \sum_{n=1}^{P} \boldsymbol{a}_n \sin\left(\pi n \frac{j\Delta}{\tau}\right) \tag{3}$$

which automatically satisfies the boundary condition. The minimization is performed with respect to Fourier coefficients $\boldsymbol{a}_n$ to get atomic trajectories. It has been pointed out that these solutions are strongly dependent upon the initial values of $\boldsymbol{a}_n$.[24]

Here, we demonstrate how a NN optimizes the OM action (eq 2) for a system of Lennard-Jones (LJ) atoms[25] and produces atomic trajectories in one fell swoop over the entire time domain between t_0 and t_1. The atomic trajectories generated by the NN are in excellent agreement with the "ground-truth" MD trajectories even though the time step taken by the NN is 20 times larger than the MD time step. We also demonstrate that the NN can easily find transformation pathways and energy barriers between different structures of LJ clusters. The NN-based approach to OM action can be easily implemented in any statistical ensemble, and it is straightforward to include constraints, invariances, and conservation laws in the NN loss function.

2. METHODS

Figure 1 shows our NN architecture. It consists of an input node, a hidden layer with n units, and an output layer with

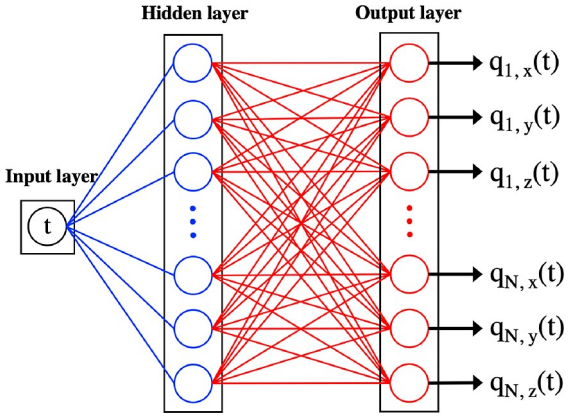


Figure 1. NN consists of an input layer, a hidden layer (blue), and an output layer (red), which gives the atomic coordinates whose first and second derivatives provide velocity and acceleration, respectively. The blue and red edges represent that the weights and each node has a bias.

$d\times N$ units, where N is the number of atoms and d is the dimensionality of the system. The network outputs the trajectories of all the atoms in the system, that is, the cartesian components $\{q_{i,x}, q_{i,y}, \text{and } q_{i,z}\}$ from $i = 1$ to N. The system configuration for all spatial degrees of freedom is calculated by feeding time into the input layer and letting it propagate through the hidden layer to the output layer. More details of the model output can be found in the Supporting Information.

We evaluate the OM action numerically by discretizing the time integral in eq 3 into N_t grid points. The time increment is $\Delta t = \tau/N_t$, where τ is the time unit computed from the chosen LJ interaction for liquid Argon.[26] The loss function L for the NN includes not only the OM action but also constraints for boundary conditions and energy conservation

$$\begin{aligned} L(\{w,\ b\}, \lambda_1, \lambda_2, \lambda_3) &= S_{\mathrm{OM}}^{\mathrm{dis}} + \lambda_1(Q(0) - Q_0)^2 + \lambda_2(Q(T) - Q_{\mathrm{T}})^2 \\ &\quad + \lambda_3 \frac{1}{N_{\mathrm{T}}} \sum_{t=0}^{T} [E(Q(t), \dot{Q}(t)) - E_0]^2 \end{aligned} \tag{4}$$

$$S_{\mathrm{OM}}^{\mathrm{dis}} = \sum_{t=0}^{T} \left[\ddot{Q}(t) + \frac{\mathrm{d}V(Q(t))}{\mathrm{d}Q}\right]^2 \Delta t \tag{5}$$

$$\begin{aligned} E(Q(t), \dot{Q}(t)) &= \sum_{i=1}^{N} \frac{1}{2} m_i \dot{\vec{q}}_i^{\,2} \\ &\quad + \frac{1}{2} \sum_{i=1}^{N} \sum_{j\neq i}^{N} 4\varepsilon \left(\frac{\sigma^{12}}{|\vec{q}_i - \vec{q}_j|^{12}} - \frac{\sigma^{6}}{|\vec{q}_i - \vec{q}_j|^{6}} \right) \end{aligned} \tag{6}$$

where $\{w, b\}$ represents all the weights and biases in the NN. For simplicity, we use $Q(t)$ to denote the model output $Q(\{w, b\}, t)$. Q_0 and Q_{T} respectively represent initial and final coordinates of all the atoms in the system. λ's are hyperparameters, which are chosen using random search schema for balancing the magnitude of individual terms (details provided in Supporting Information). Note that the last term in eq 4 is the energy-conservation constraint, where E_0 is the total energy of the system. In eq 6, $E(Q(t),\dot{Q}(t))$ is the total energy, and ε and σ are the parameters for the LJ potential.

We minimize the loss function in eq 4 using the Nestrov-and-Adam (NADAM) optimizer.[27] The gradient of the loss function is calculated numerically using Google's JAX library.[28] After training the network, we compute atomic trajectories and compare them with the "ground-truth" MD simulations performed using the same set of initial conditions. We have trained the network to predict trajectories of liquid Argon systems consisting of up to 500 atoms in three dimensions. We have also used the NN to find transition pathways and energy barriers between different structures of LJ clusters. More details about the NN model and the atomic systems we study are given in the Supporting Information.

3. RESULTS

The NN learns to compute atomic trajectories of two- and three-dimensional LJ systems from the OM least action principle. Here, we present results for a 3D system containing 500 atoms, which was trained on a NN with 125 neurons in the hidden layer. Results in this section are generated on N_t = 25 time increments, which correspond to a time step (∼40 fs)

https://doi.org/10.1021/acs.jcim.2c00515
J. Chem. Inf. Model. 2022, 62, 3346−3351

that is 20 times the MD timestep (2 fs). It is worth noting that the NN is designed to approach the most probable phase-space trajectory for a system given initial and final configurations, rather than outperform MD at the computing speed of atomic trajectories.

Figure 2 shows a comparison of NN and MD trajectories for four randomly chosen particles in the 500-atom LJ system over

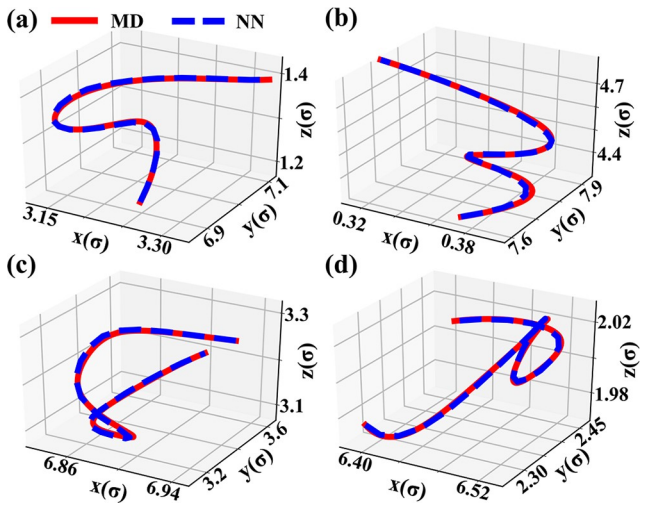


Figure 2. Visualization of atomic trajectories in the 500-atom system. NN (solid blue) and MD (solid red) trajectories show that the NN can reproduce non-trivial paths over a time period of 1 ps.

a time period of 1 ps. These trajectories cover a time scale beyond the straight-line ballistic motion and possess sharp changes in direction. It is remarkable that the NN is able to reproduce the dynamics at the individual particle level. From this level of matching for particle trajectories, it is not surprising that the NN results for all other quantities are in excellent agreement with the "ground truth" MD results.

Figure 3 shows the root mean-square deviations (RMSDs) between the NN and MD trajectories over the entire time domain as a function of epochs for 256- and 500-atom LJ systems. The RMSDs for positions Q are indeed very small, resulting in excellent matching of the NN and MD trajectories. The RMSDs between the NN and MD trajectories are negligible near the initial and final configurations because of boundary-condition constraints in the loss function. Even in the middle of the time domain, the RMSD does not exceed 10^{-4}. The difference between the potential energies computed by the NN and MD is therefore negligible. The behavior of the RMSDs for momenta P as a function of epoch is similar to the RMSDs of trajectories, but the deviations are larger for the following reasons: first, referring to eqs 4−6, the time derivatives are described with fewer parameters than the positions as the biases outside the activation function are killed by the derivative; second, all quantities in the loss function depend explicitly on Q and only two depend on P. Since we have a finite number of parameters, they can only approximate the function up to some error. These errors magnify as we take higher derivatives.

In Figure 4, we compare the results from the NN and MD for some standard simulation quantities. For example, the

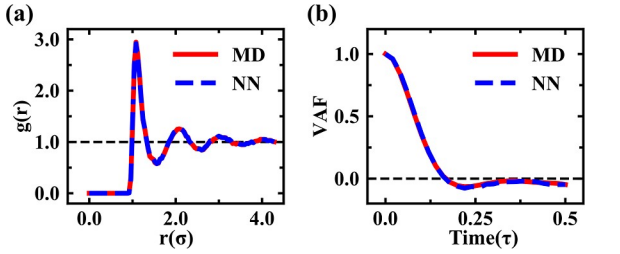


Figure 4. Comparison between the NN and MD simulation results. (a) Radial distribution functions and (b) velocity autocorrelation functions from MD (red) and NN (blue) dynamics.

structure of the liquid is characterized by the radial distribution function[29] $g(r)$ shown in Figure 4a. Slightly below $r = 1$, $g(r) = 0$ because the particles cannot overlap and the NN trajectory maintains this feature. The peaks in $g(r)$ corresponding to the nearest and second nearest neighbors indicate that our model is able to obtain the correct liquid structure. To characterize the particle dynamics, we calculated the velocity autocorrelation function[30] (VAF) shown in Figure 4b. The match between NN and MD is good, and the correlation time t_c, when VAF = 0 for the first time, is about 0.16 in the unit of $\tau = 2$ ps. Beyond t_c the velocities and, in general, the particle dynamics are not correlated with the initial dynamics of the particles. This quantity is important in calculating time averages, as a time series of data with points separated by at least t_c are not correlated and contribute independently to the average quantity.

We have also used our NN to find transition pathways between different structures of LJ clusters. The number of

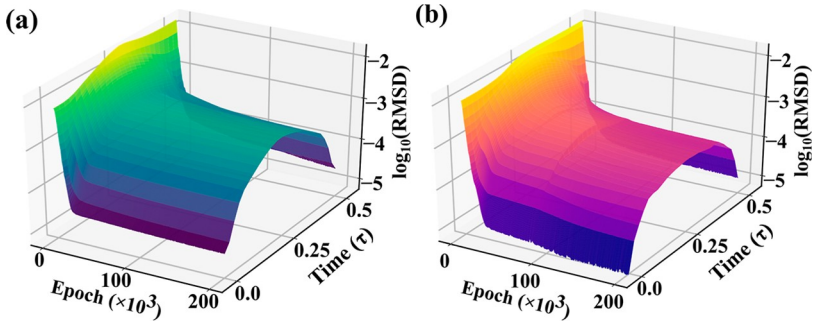


Figure 3. Root-mean-square deviation (RMSD) between MD and NN trajectories (a) Shows the deviation between MD and NN trajectories for a 256-atom LJ system. (b) figure shows how the RMSD between NN and MD trajectories for a 500-particle system decreases with the number of epochs and changes with time. The boundary-condition constraints in the loss function ensure that the RMSDs are much smaller near the initial and final configurations than in the middle of the time domain. Here, $\tau = 2$ ps.

https://doi.org/10.1021/acs.jcim.2c00515
J. Chem. Inf. Model. 2022, 62, 3346−3351

conformations of a cluster can be very large,[31] and the potential energy surface (PES) of a cluster has a large number of local minima in which it can get trapped, making it very difficult to identify the global minimum.[32] Finding the global minimum of a cluster or protein is considered an NP-hard problem;[33] that is, the global minimum cannot be found in polynomial time.

There have been numerous computational studies of LJ clusters consisting of tens to hundreds of atoms. Global minima of almost all the clusters up to LJ_{150} atoms have been found. Most of them have Mackay icosahedral structures as their global minima.[34] However, there are a few exceptions, namely, LJ_{38}, LJ_{75-77}, and $LJ_{102-104}$.[35] The global minimum of LJ_{38} is a face centered cubic (FCC) truncated octahedron, and the global minima of LJ_{75-77} and $LJ_{102-104}$ are Marks dodecahedra.[35]

Figure 5 shows the structural transition pathway of an LJ_{38} from a local to the global minimum on the PES. For

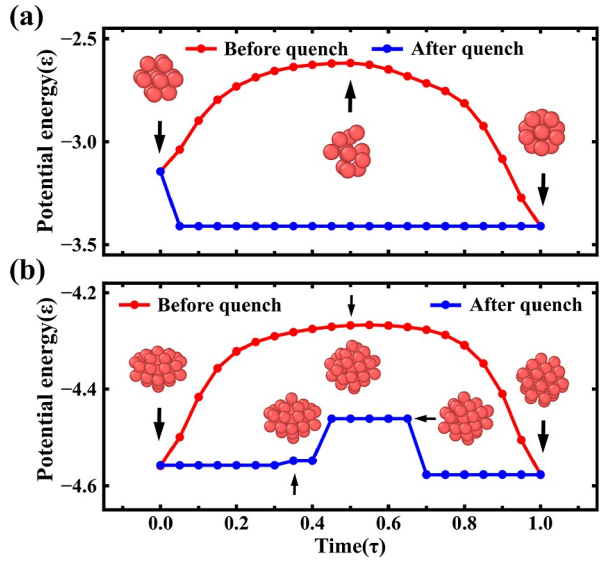


Figure 5. Panels (a) and (b) show potential energies per atom for LJ_{13} and LJ_{38} clusters, respectively, as a function of time. The intermediate structures of the clusters are also shown in panels (a) and (b) red curves are the NN potential energies and the blue curves are the PES obtained by quenching NN configurations. LJ_{13} quickly quenches into a Mackay icosahedra global minimum. The LJ_{38} structure indicates that the cluster remains an incomplete Mackay icosahedron for nearly 0.8 ps. In the next 0.6 ps, the cluster crosses an energy barrier and transforms into an FCC truncated octahedron, which is the global minimum. The LJ_{38} structures just before, during, and after the transition are shown in (b).

comparison, we have also trained our NN to find the transition of an LJ_{13} to the Mackay icosahedron global minimum. The structural change in LJ_{13} from a truncated FCC to an icosahedron can be easily spotted within a thousand MD steps. The NN finds this transition within a few time steps as shown in Figure 5a. In the case of an LJ_{38}, we used the NN to find the transition pathway between the second lowest and the global minimum, that is, from an incomplete Mackay icosahedron to an FCC truncated octahedron. This transition is hard to find by MD simulation, whereas the NN rapidly finds not only the transition pathway but also the potential energy barrier and intermediate conformations between the Mackay icosahedron and the FCC truncated octahedron (see Figure 5b).

To quantify the difference between the transition pathways of LJ_{13} and LJ_{38}, we compute the Euclidean distance of the initial and final configurations Q_i and Q_f. The distance between the two structures is the minimal arc length traveled by a cluster as it transitions from Q_i to Q_f in a time period T. The minimal distance is defined by a transformation $\bar{Q}(t)$, which is a 3N-dimensional vector parameterized by time t. The infinitesimal distance traveled during the transformation is $\sqrt{\left(\frac{d\bar{Q}(t)}{dt}\right)^2}$, and the minimal distance between Q_i and Q_f is given by[36]

$$d(Q_i, Q_f) = \int_0^T \sqrt{\left(\frac{d\bar{Q}(t)^*}{dt}\right)^2}\, dt \tag{7}$$

where $\bar{Q}(t)^*$ is the minimal path determined by the NN. We find that the Euclidean distances between the initial and final configurations of LJ_{13} and LJ_{38} are 0.9 σ and 10.4 σ, respectively. It has been suggested that the global minimum of an LJ_{13} can be located much more easily than that of an LJ_{38} because the former has a single funnel energy landscape and the latter has double funnels leading to the global minimum.[37]

4. CONCLUSIONS

In this work, we have shown that a NN can correctly generate atomic trajectories of a LJ system by minimizing the OM action. The NN phase-space trajectories, potential and kinetic energies, radial distribution function, and velocity autocorrelation function are in excellent agreement with the corresponding results from the "ground-truth" MD simulation. We have also demonstrated that the NN can easily find transition pathways and energy barriers between different configurations of typical LJ clusters. Despite the limitation of our method in computing speed, the NN approach to principle of least action has some advantages over the traditional MD method: the NN solves boundary condition problems while the MD cannot, and the NN provides the entire trajectory in one fell swoop with a time step much larger than that of the MD, while maintaining energy conservation over the entire time domain. Our NN approach is generally applicable to systems with complex interatomic interactions other than LJ potential. The encouraging results indicate the possibility of making the NN method a candidate for modeling experimental data. For example, Hills et al. have developed an algorithm based on the least action principle to predict the dynamics of physical systems using observed data.[38] Our approach combined with theirs can model observational data and predict the dynamics beyond the experimental measurements of systems more than LJ clusters.

DATA AVAILABILITY

The code and data that support the findings of this study are provided in the Supporting Information; they are also available from the corresponding author upon request.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.2c00515.

https://doi.org/10.1021/acs.jcim.2c00515
J. Chem. Inf. Model. 2022, 62, 3346−3351

Details of the LJ interaction for the system, model pretraining protocol, scaling curves for the time performance and RMSD with respect to the system size, comparisons between MD and NN trajectories with correlation functions for systems having different number of atoms, and transition pathway generated by the NN model for a surface-hopping event in a small 2D LJ cluster (PDF)

AUTHOR INFORMATION

Corresponding Author

Beibei Wang − *Collaboratory for Advanced Computing and Simulations, University of Southern California, Los Angeles, California 90089, United States;* orcid.org/0000-0002-9581-2698; Email: beibeiw@usc.edu

Authors

Shane Jackson − *Collaboratory for Advanced Computing and Simulations, University of Southern California, Los Angeles, California 90089, United States;* orcid.org/0000-0002-2663-8235

Aiichiro Nakano − *Collaboratory for Advanced Computing and Simulations, University of Southern California, Los Angeles, California 90089, United States;* orcid.org/0000-0003-3228-3896

Ken-ichi Nomura − *Collaboratory for Advanced Computing and Simulations, University of Southern California, Los Angeles, California 90089, United States*

Priya Vashishta − *Collaboratory for Advanced Computing and Simulations, University of Southern California, Los Angeles, California 90089, United States;* orcid.org/0000-0003-4683-429X

Rajiv Kalia − *Collaboratory for Advanced Computing and Simulations, University of Southern California, Los Angeles, California 90089, United States*

Mark Stevens − *Center for Integrated Nanotechnologies, Sandia National Laboratory, Albuquerque, New Mexico 87185, United States;* orcid.org/0000-0002-2290-0224

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.2c00515

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation, Future Manufacturing Program, Award 2036359. The coauthor K.N. was partially supported by an NSF grant, Cyber Training on Materials Genome Innovation for Computational Software (CyberMAGICS), under award number 2118061.

REFERENCES

(1) Landau, L. D.; Lifshitz, E. M. *Course of Theoretical Physics Volume 1: Mechanics*, 3rd ed.; Butterworth-Heinemann: Oxford, 1982.

(2) Landau, L. D.; Lifshitz, E. M. *Course of Theoretical Physics Volume 2: The Classical Theory of Fields*, 4th ed.; Butterworth-Heinemann: Oxford, 1987.

(3) Feynman, R. P. The Principle of Least Action in Quantum Mechanics. *Feynman's Thesis—A New Approach to Quantum Theory*; World Scientific, 2005; pp 1−69.

(4) García-Morales, V.; Pellicer, J.; Manzanares, J. A. Thermodynamics Based on the Principle of Least Abbreviated Action: Entropy Production in a Network of Coupled Oscillators. *Ann. Phys.* **2008**, *323*, 1844−1858.

(5) Goldstein, H. *Classical Mechanics*, 3rd ed.; Addison-Wesley: Boston, 2002.

(6) Rahman, A. Correlations in the Motion of Atoms in Liquid Argon. *Phys. Rev.* **1964**, *136*, A405−A411.

(7) Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. A Computer Simulation Method for the Calculation of Equilibrium Constants for the Formation of Physical Clusters of Molecules: Application to Small Water Clusters. *J. Chem. Phys.* **1982**, *76*, 637−649.

(8) Behler, J. Perspective: Machine Learning Potentials for Atomistic Simulations. *J. Chem. Phys.* **2016**, *145*, 170901.

(9) Fukushima, S.; Ushijima, E.; Kumazoe, H.; Koura, A.; Shimojo, F.; Shimamura, K.; Misawa, M.; Kalia, R. K.; Nakano, A.; Vashishta, P. Thermodynamic Integration by Neural Network Potentials Based on First-Principles Dynamic Calculations. *Phys. Rev. B* **2019**, *100*, 214108.

(10) Rajan, A. C.; Mishra, A.; Satsangi, S.; Vaish, R.; Mizuseki, H.; Lee, K.-R.; Singh, A. K. Machine-Learning-Assisted Accurate Band Gap Predictions of Functionalized Mxene. *Chem. Mater.* **2018**, *30*, 4031−4038.

(11) Wang, J.; Yang, X.; Zeng, Z.; Zhang, X.; Zhao, X.; Wang, Z. New Methods for Prediction of Elastic Constants Based on Density Functional Theory Combined with Machine Learning. *Comput. Mater. Sci.* **2017**, *138*, 135−148.

(12) Iwasaki, Y.; Takeuchi, I.; Stanev, V.; Kusne, A. G.; Ishida, M.; Kirihara, A.; Ihara, K.; Sawada, R.; Terashima, K.; Someya, H.; Uchida, K.-i.; Saitoh, E.; Yorozu, S. Machine-Learning Guided Discovery of a New Thermoelectric Material. *Sci. Rep.* **2019**, *9*, 2751.

(13) Wang, T.; Zhang, C.; Snoussi, H.; Zhang, G. Machine Learning Approaches for Thermoelectric Materials Research. *Adv. Funct. Mater.* **2020**, *30*, 1906041.

(14) Ramprasad, R.; Batra, R.; Pilania, G.; Mannodi-Kanakkithodi, A.; Kim, C. Machine Learning in Materials Informatics: Recent Applications and Prospects. *npj Comput. Mater.* **2017**, *3*, 54.

(15) Bassman, L.; Rajak, P.; Kalia, R. K.; Nakano, A.; Sha, F.; Sun, J.; Singh, D. J.; Aykol, M.; Huck, P.; Persson, K.; Vashishta, P. Active Learning for Accelerated Design of Layered Materials. *npj Comput. Mater.* **2018**, *4*, 74.

(16) Greydanus, S.; Dzamba, M.; Yosinski, J. Hamiltonian Neural Networks. *Adv. Neural Inf. Process. Syst.* **2019**, *32*, 1−11.

(17) Cranmer, M.; Greydanus, S.; Hoyer, S.; Battaglia, P.; Spergel, D.; Ho, S. Lagrangian Neural Networks. **2020**, Vol. *1*−*9*, arXiv:2003.04630.

(18) Jin, Z.; Lin, J. Y.-Y.; Li, S.-F. Learning Principle of Least Action with Reinforcement Learning. *CEUR Workshop Proc.* **2020**, *2964*, 1−5.

(19) Elber, R.; Meller, J.; Olender, R. Stochastic Path Approach to Compute Atomically Detailed Trajectories: Application to the Folding of C Peptide. *J. Phys. Chem. B* **1999**, *103*, 899−911.

(20) Onsager, L.; MacHlup, S. Fluctuations and Irreversible Processes. *Phys. Rev.* **1953**, *91*, 1505−1512.

(21) Onsager, L. Reciprocal Relations in Irreversible Processes. I. *Phys. Rev.* **1931**, *37*, 405−426.

(22) Onsager, L. Reciprocal Relations in Irreversible Processes. II. *Phys. Rev.* **1931**, *38*, 2265−2279.

(23) Yasue, K. The Role of the Onsager−Machlup Lagrangian in the Theory of Stationary Diffusion Process. *J. Math. Phys.* **1979**, *20*, 1861−1864.

(24) Passerone, D.; Parrinello, M. Action-Derived Molecular Dynamics in the Study of Rare Events. *Phys. Rev. Lett.* **2001**, *87*, 108302.

(25) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*, 2nd ed.; Oxford University Press: Oxford, 2017.

(26) White, J. A. Lennard-Jones as a Model for Argon and Test of Extended Renormalization Group Calculations. *J. Chem. Phys.* **1999**, *111*, 9352−9356.

(27) Dozat, T. *Incorporating Nesterov Momentum into Adam*; ICLR Workshop, 2016, pp 2013−2016.
(28) Frostig, R.; Johnson, M. J.; Leary, C. *Compiling Machine Learning Programs via High-Level Tracing*; SysML, 2018.
(29) Kirkwood, J. G.; Boggs, E. M. The Radial Distribution Function in Liquids. *J. Chem. Phys.* **1942**, *10*, 394−402.
(30) Alder, B. J.; Wainwright, T. E. Decay of the Velocity Autocorrelation Function. *Phys. Rev. A* **1970**, *1*, 18−21.
(31) Doye, J. P. K.; Wales, D. J. On Potential Energy Surfaces and Relaxation to the Global Minimum. *J. Chem. Phys.* **1996**, *105*, 8428−8445.
(32) Mandelshtam, V. A.; Frantsuzov, P. A. Multiple Structural Transformations in Lennard-Jones Clusters: Generic versus Size-Specific Behavior. *J. Chem. Phys.* **2006**, *124*, 204511.
(33) Wille, L. T.; Vennik, J. Computational Complexity of the Ground-State Determination of Atomic Clusters. *J. Phys. A: Math. Gen.* **1985**, *18*, L419−L422.
(34) Northby, J. A. Structure and Binding of Lennard-Jones Clusters: 13≤N≤147. *J. Chem. Phys.* **1987**, *87*, 6166−6177.
(35) Wales, D. J.; Doye, J. P. K. Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms. *J. Phys. Chem. A* **1997**, *101*, 5111−5116.
(36) Gelfand, I. M.; Fomin, S. V. *Calculus of Variations*; Courier Corporation: New York, 2000.
(37) Doye, J. P. K.; Miller, M. A.; Wales, D. J. The Double-Funnel Energy Landscape of the 38-Atom Lennard-Jones Cluster. *J. Chem. Phys.* **1999**, *110*, 6896−6906.
(38) Hills, D. J. A.; Grütter, A. M.; Hudson, J. J. An Algorithm for Discovering Lagrangians Automatically from Data. *PeerJ Comput. Sci.* **2015**, *1*, No. e31.

https://doi.org/10.1021/acs.jcim.2c00515
J. Chem. Inf. Model. 2022, 62, 3346−3351