

Accurate nuclear quantum statistics on machine-learned classical effective potentials

Cite as: *J. Chem. Phys.* **161**, 134102 (2024); doi: [10.1063/5.0226764](https://doi.org/10.1063/5.0226764)

Submitted: 3 July 2024 • Accepted: 13 September 2024 •

Published Online: 1 October 2024



View Online



Export Citation



CrossMark

Iryna Zaporozhets,^{1,2,3}  Félix Musil,¹  Venkat Kapil,^{4,5,6,a)}  and Cecilia Clementi^{1,3,2,b)} 

AFFILIATIONS

¹Department of Physics, Freie Universität Berlin, Arnimallee 12, 14195 Berlin, Germany

²Department of Chemistry, Rice University, Houston, Texas 77005, USA

³Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005, USA

⁴Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom

⁵Department of Physics and Astronomy, University College, London WC1E 6BT, United Kingdom

⁶Thomas Young Centre and London Centre for Nanotechnology, London WC1E 6BT, United Kingdom

Note: This paper is part of the JCP Special Topic on Molecular Dynamics, Methods and Applications 60 Years After Rahman.

^{a)}v.kapil@ucl.ac.uk

^{b)}Author to whom correspondence should be addressed: cecilia.clementi@fu-berlin.de

ABSTRACT

The contribution of nuclear quantum effects (NQE) to the properties of various hydrogen-bound systems, including biomolecules, is increasingly recognized. Despite the development of many acceleration techniques, the computational overhead of incorporating NQEs in complex systems is sizable, particularly at low temperatures. In this work, we leverage deep learning and multiscale coarse-graining techniques to mitigate the computational burden of path integral molecular dynamics (PIMD). In particular, we employ a machine-learned potential to accurately represent corrections to classical potentials, thereby significantly reducing the computational cost of simulating NQEs. We validate our approach using four distinct systems: Morse potential, Zundel cation, single water molecule, and bulk water. Our framework allows us to accurately compute position-dependent static properties, as demonstrated by the excellent agreement obtained between the machine-learned potential and computationally intensive PIMD calculations, even in the presence of strong NQEs. This approach opens the way to the development of transferable machine-learned potentials capable of accurately reproducing NQEs in a wide range of molecular systems.

© 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0226764>

I. INTRODUCTION

Molecular dynamics (MD) simulations provide insights into phenomena at atomic resolution and are an invaluable tool in chemistry, biophysics, and material science. Recent advances in force field,^{1–3} algorithms,⁴ and hardware development⁵ allow for increasingly higher accuracy and a broader scope of application of molecular dynamics simulations.

Traditional MD simulations rely on the classical evolution of nuclear positions, neglecting an explicit treatment of so-called nuclear quantum effects (NQEs) such as zero-point fluctuations, tunneling phenomena, and isotope effects. This simplification has a historical justification as the parameters of classical empirical force field simulations are fit to reproduce (quantum) experimental data. Hence, NQEs are effectively included in the classical potential

energy surface, at least partially, albeit with limited transferability. Meanwhile, bottom-up *ab initio* MD methods or machine learning interatomic potentials (MLIPs), based on the Born–Oppenheimer approximation,⁶ only account for the quantization of electrons. Hence, these simulations require an explicit treatment of quantum nuclear motion on the Born–Oppenheimer potential energy surface (PES) in theory. Particularly when accurate electronic structure methods are used to estimate the Born–Oppenheimer PES, the neglect of quantum nuclear motion may be the largest source of error for obtaining quantitative agreement with experiments.^{7–10} Unfortunately, NQEs are typically neglected due to their significant computational overhead or because they are assumed to have small effects. However, even at ambient temperatures, NQEs significantly affect the properties of hydrogen-bounded systems, e.g., static and dynamic properties of water,^{11–23} free energies of hydrogen bonding

and proton transfer,^{24–29} stability of molecular crystals,^{30,31} strength of intermolecular interactions,^{9,32} and enzyme activity.^{19,33}

The standard condensed-phase simulation technique for incorporating NQEs in equilibrium properties is Feynman's imaginary time path integral (PI) formalism.³⁴ In this framework, a system's quantum statistical mechanics is mapped onto the classical statistical mechanics of a ring polymer, which consists of multiple replicas of the system connected by harmonic interactions that depend on temperature and mass. The effective classical PI partition function of the ring polymer system can be sampled using MD or Monte Carlo techniques, yielding the PIMD³⁵ and PI Monte Carlo³⁶ methods. In theory, PI methods incorporate quantum statistics exactly with classical sampling, in the limit of the number of replicas going to infinity.³⁷ However, typically, the required number of replicas to converge structural properties, and thus the increased computational cost compared to classical MD, rises with increasing physical frequencies or lowering temperatures.³⁸ As a result, most molecular systems require 16–32 replicas at room temperature, which increases proportionally with lowering temperature, such as over 10^4 replicas to obtain converged energies for molecular systems at cryogenic temperatures.³⁹

Over the past decades, several systematically improvable reduced-cost PI approaches have been proposed. These include high-order factorization of the Boltzmann operator,^{40–45} perturbative expansion of the Boltzmann operator,^{46,47} multiple time stepping,⁴⁸ ring polymer contraction,⁴⁹ and colored-noise thermostats.^{50–52} Despite their notable success toward mainstreaming the use of PI methods,³⁸ these approaches do not yet present a “fix-all” solution. Issues still exist, including, but not limited to, lack of universal applicability of multiple time stepping and ring polymer contraction, tedious implementation of high-order path integral techniques, and poor ergodicity⁵³ or zero-point energy leakage⁵⁰ due to the coupling between physical and fictitious PI collective modes. Finally, these approaches do not eliminate the diverging cost of path integral simulations at low temperatures.

An appealing direction that avoids the computational overhead of PI methods is to avoid modeling the entire imaginary-time path. For instance, the Wigner–Kirkwood effective potential^{54,55} expands the partition functions in powers of the reduced Planck's constant and retains computationally tractable terms. Similarly, an effective quantum potential was suggested by Feynman and Hibbs^{34,56} by integrating out all degrees of freedom except the centroid of the ring polymer, i.e., by coarse-graining the ring polymer to the centroid. Ever since, the idea of ring polymer coarse-graining has been explored for over 30 years^{57–65} and has recently gained increasingly more attention thanks to the advent of MLIPs.^{66–68}

A number of coarse-grained effective potentials can be developed by identifying (a) a suitable mapping from the fine-grained ring polymer to the coarse-grained classical system and (b) the functional form of the interactions in the coarse-grained system. For instance, the so-called CG-PI theory, developed by Voth and co-workers,^{69–71} proposes to map the imaginary time path to a two-replica coarse-grained system. This mapping gives access to both diagonal and nondiagonal elements of the thermal density matrix. However, the corresponding formalism is challenging to implement for realistic molecular systems in the regime of strong NQEs.⁷¹ Another common mapping involves coarse-graining of the ring polymer to the centroid.^{65–68} Among others, we have

recently shown that⁶⁶ the resulting effective potential can be used in the same manner as a classical force field. It is well known that coarse-graining the ring polymer to its centroid can still give rigorous access to equilibrium properties, in terms of the so-called centroid-constrained imaginary time correlation function.⁵⁷ However, estimating equilibrium properties, such as the average of a non-linear position-dependent operator, in this manner is usually complicated for generic systems because ensemble averages of the operators evaluated on the centroid coordinates need to be integrated in the high-dimensional configurational space with the centroid-constrained imaginary time correlation function.⁵⁷ On the other hand, the centroid-based description is most effective for an approximate treatment of dynamical properties of complex systems in terms of time correlation functions estimated on the centroid trajectory.⁵⁸ A prominent feature of our work is the path-integral coarse-grained simulation (PIGS) method, which uses many-body universal approximators⁷² to fit the coarse-grained potential of mean force. The PIGS approach utilizes machine-learned potentials, capable of approximating arbitrary complex functions, to learn the full many-body coarse-grained potentials as demonstrated recently on biomolecular simulations.^{73–76}

Here, instead of dynamical properties, we focus on characterizing static position-dependent properties, including NQEs, and we develop an effective machine-learned potential to compute them accurately and at the cost of classical MD. Our approach builds upon the PIGS approach that uses multiscale coarse-graining⁷² and our work on the use of deep learning architectures for classical coarse-grained force-fields⁷⁷ but maps the ring polymer onto a single replica instead of the centroid. This coarse-grained mapping is distinct from earlier choices and allows quantum thermodynamic expectation values of position-dependent observables to be estimated rigorously as simple time averages akin to classical MD. We demonstrate that this approach yields accurate quantum nuclear statistics of bulk and isolated systems. The remaining sections of this paper are organized as follows: In Sec. II, we briefly summarize the imaginary-time PI theory and the application of thermodynamic coarse-graining to the problem. Section III discusses the numeric implementation of the approach for four model systems: one-dimensional Morse potential, single H₂O molecule, Zundel cation, and bulk water. The results are presented and discussed in Sec. IV. Section V provides concluding remarks.

II. THEORY AND METHODS

A. The path-integral approach

We consider a Hamiltonian H of N distinguishable nuclei, living in three-dimensional Cartesian space, with masses $\{m_1, \dots, m_N\} \equiv \{m_i\}_i$, and position vector $\mathbf{q} \equiv \{\mathbf{q}_i\}_i$, interacting with the potential $U(\mathbf{q})$. The equivalent PI partition function comprising P replicas can be expressed as

$$Z = \text{tr} \left[e^{-\beta \hat{H}} \right] = \lim_{P \rightarrow \infty} Z_P, \quad (1)$$

where

$$Z_P \propto \int d\mathbf{Q} e^{-\beta U^{\text{PI}}(\mathbf{Q})}, \\ = \int \prod_{j=1}^P d\mathbf{q}^{(j)} e^{-\frac{\beta}{P} \left[\sum_{j=1}^P \left[\sum_{i=1}^N \frac{1}{2} \frac{m_i p_i^2}{\beta^2 \hbar^2} (\mathbf{q}_i^{(j+1)} - \mathbf{q}_i^{(j)})^2 + U(\mathbf{q}^{(j)}) \right] \right]}. \quad (2)$$

Here, $\mathbf{q}^{(j)}$ represents the position vector of the j th replica of the system in the ring polymer with cyclic boundary conditions $\mathbf{q}^{(j)} \equiv \mathbf{q}^{(j+P)}$. In addition, $\mathbf{Q} \equiv \{\mathbf{q}^{(j)}\}_j$ is a shorthand notation for the set of positions of the P replicas of the ring polymer, and $U^{\text{PI}}(\mathbf{Q})$ is a shorthand notation for the total potential experienced by the ring polymer. In the limit $P \rightarrow \infty$, the ring polymer statistics is an unbiased estimator of quantum statistics of the target system.

The quantum thermodynamic average of a position-dependent operator $O(\hat{\mathbf{q}})$ is typically estimated as the ensemble average of the position-dependent function averaged over all the replicas,

$$\langle O \rangle = \left\langle \frac{1}{P} \sum_{j=1}^P O(\mathbf{q}^{(j)}) \right\rangle_{\text{PI}}, \quad (3)$$

where $\langle \square \rangle_{\text{PI}}$ corresponds to the ensemble average defined in Eq. (2). Estimating Eq. (3) is computationally demanding due to the need for simulating P replicas of the system. Typically, P should be larger than $\beta \hbar \omega_{\text{max}}$ with ω_{max} being the largest physical frequency of the system.⁴⁹ Notably, for a given system (with fixed ω_{max}), the number of replicas scales inversely with T , implying a steeply rising computational cost as the temperature goes to zero.

B. Path-integral coarse-graining for quantum statistics

Our goal is to reduce the computational cost of simulating ring polymer statistics by integrating additional degrees of freedom, such as (linear combinations of) the replicas of the system, in a way that retains the accuracy of Eq. (2). In the most extreme and computationally beneficial scenario, we can integrate the $P - 1$ replicas of the physical system to an effective potential felt by a single replica, reducing the dimensionality of the ring polymer to the physical system. In doing so, we reduce the cost of estimating thermodynamic properties to that of classical MD.

The problem can be reformulated using the bottom-up coarse-graining technique widely used in biomolecular simulations. We wish to coarse-grain the ring polymer system in the high-dimensional configurational space $\mathbf{Q} \in \mathbb{R}^{3NP}$ described by a potential $U^{\text{PI}}(\mathbf{Q})$ into an effective thermodynamically consistent potential $U^{\text{eff}}(\mathbf{q})$ with $\mathbf{q} \in \mathbb{R}^{3N}$,

$$U^{\text{eff}}(\mathbf{q}) = -\beta^{-1} \ln \left[\frac{\int d\mathbf{Q} e^{-\beta U^{\text{PI}}(\mathbf{Q})} \delta(\mathbf{q} - \xi(\mathbf{Q}))}{\int d\mathbf{Q} e^{-\beta U^{\text{PI}}(\mathbf{Q})}} \right], \quad (4)$$

where δ is a Dirac delta function and ξ is a linear operator that maps coordinates from $\mathbb{R}^{3NP} \rightarrow \mathbb{R}^{3N}$. The effective potential of mean force in Eq. (4) can be obtained using the force-matching variational approach.⁷² We first represent the effective potential as a function, $\tilde{U}^{\text{eff}}(\mathbf{q}, \boldsymbol{\theta})$, parameterized by $\boldsymbol{\theta}$, and then minimize the force-matching loss⁷³ with respect to $\boldsymbol{\theta}$,

$$\mathcal{L} = \|\xi_f(\mathbf{F}) + \nabla_{\mathbf{q}} \tilde{U}^{\text{eff}}(\mathbf{q}; \boldsymbol{\theta})\|. \quad (5)$$

Here, $\mathbf{F} \equiv -\nabla_{\mathbf{Q}} U^{\text{PI}}(\mathbf{Q})$ is the force associated with the (high-dimensional) ring polymer systems and ξ_f is a linear operator that maps the forces from $\mathbb{R}^{3NP} \rightarrow \mathbb{R}^{3N}$, i.e., the high-dimensional ring polymer space into the low-dimensional space associated with the physical system.

C. Single replica coarse-graining

While several choices of the mapping operators ξ and ξ_f are possible,⁷⁷ not all choices are ideal for estimating generic position-dependent operators with the full accuracy of Eq. (3). To make a physically meaningful and computationally favorable choice, we note that all replicas of the ring polymer are equivalent thanks to the invariance of the trace in Eq. (2) to cyclic shifts. Hence, a thermodynamic expectation value can be estimated by averaging the position-dependent function estimated on *any* replica j' ,

$$\langle O \rangle = \left\langle O(\mathbf{q}^{(j')}) \right\rangle_{\beta}. \quad (6)$$

Equation (6) suggests that the statistics of a single replica are sufficient to estimate the equilibrium averages of position-dependent observables. From the perspective of developing an effective potential, it is sufficient to integrate all but one (arbitrary) replica. Hence, in this work, we coarse-grain the ring polymer to a single replica using as a configurational map the following transformation:

$$\xi(\mathbf{Q}) = \mathbf{q}^{(j')}. \quad (7)$$

Given this configurational map, a thermodynamically consistent force map ξ_f can be constructed with the matrix elements given by the following relationship:⁷⁷

$$\xi_f^{ji} = \begin{cases} 1 & \text{for atom } i \text{ identified with replica } j, \\ 0 & \text{for atom } i \text{ identified with replica } k \neq j, \\ c_{ji} \in \mathbb{R} & \text{for all other atoms,} \end{cases} \quad (8)$$

where c_{ji} are arbitrary. For a ring polymer, setting $c_{ji} = 0$ leads to the following primitive estimator of force projected onto a coarse-grained space:

$$\xi_f(\mathbf{F}) = -\frac{\nabla U(\mathbf{q}^{(j')})}{P} - \frac{mP}{\beta^2 \hbar^2} (2\mathbf{q}^{(j')} - \mathbf{q}^{(j'+1)} - \mathbf{q}^{(j'-1)}). \quad (9)$$

However, the numerical efficiency of the force-matching minimization can be further improved by optimizing c_{ji} to minimize the variance of the mapped force.⁷⁷ With the configurational mapping operator given by Eq. (7), the expectation values of a position-dependent operator can be estimated as a simple ensemble average,

$$\langle O \rangle = \langle O(\mathbf{q}) \rangle_{\beta}, \quad (10)$$

akin to standard MD, albeit in the classical ensemble of the effective potential $U^{\text{eff}}(\mathbf{q})$.

The choice of the mapping operator in Eq. (7) is distinct from earlier choices made in the context of coarse-graining of imaginary-time PIs. The most popular coarse-graining approach maps the ring polymer positions to its centroid. While this approach is useful for simulating dynamical properties, it does not give correct equilibrium averages of generic position-dependent operators when estimated as an ensemble average similar to Eq. (10). Similarly, the PICG theory introduced in Refs. 69–71 coarse-grains the ring polymer to two replicas, which are related to the positions at which the off-diagonal elements of the Boltzmann operator are estimated. This approach

gives access to generic position- and momentum-dependent operators; however, it requires derivations of bespoke (complicated) operators. Meanwhile, our mapping reduces the ring polymer to the position at which the trace of the Boltzmann operator is estimated, with $U^{\text{eff}}(\mathbf{q}) = -\beta^{-1} \log(\langle \mathbf{q} | e^{-\beta \hat{H}} | \mathbf{q} \rangle)$, modulo an additive constant, giving access to position-dependent operators as simple ensemble averages.

D. Workflow using the PIGS method

To develop the effective potentials that incorporate NQEs, we utilize the PIGS approach, outlined in Fig. 1. First, we generate training data via the PIMD simulation. The resulting ring polymer trajectories are then projected onto a coarse-grained space using a configurational and force map discussed above. Notice that the remaining replica j' in Eq. (7) is selected arbitrarily, and for a ring polymer with P replicas, Eq. (7) defines P equivalent projections. Thus, each ring polymer configuration yields P data points for training.

The effective potential is represented in the following form:

$$U^{\text{eff}}(\mathbf{q}, \boldsymbol{\theta}) = \alpha U(\mathbf{q}) + U^{\text{ML}}(\mathbf{q}, \boldsymbol{\theta}), \quad (11)$$

where $U^{\text{ML}}(\mathbf{q}, \boldsymbol{\theta})$ is a learnable correction. The hyperparameter α determines the contribution of the physical potential to the total energy. A common approach is to set $\alpha = 1$, so the ML network learns an effective NQE correction to the classical potential.^{66,68} This choice is effective at high temperatures when NQEs are small. However, as temperature decreases, prior classical distribution quickly becomes more localized, and the difference between the classical and target quantum potential increases, making the correction harder to learn. However, in the harmonic limit, it is possible to make classical distribution more “quantum-like” by changing the simulation temperature. For this reason, to simplify the learning and enhance the data efficiency of the training and the model stability, we consider α to be temperature dependent: $\alpha = T/T_0$, where $T < T_0$ denotes the target temperature and T_0 is a hyperparameter. If not specified otherwise, we set $T_0 = 600$ K. The learnable model parameters are then optimized with standard ML approaches. The resulting effective potential can then be used for calculations as a standard MD force field.

E. Applicability of the models to different thermodynamic conditions

The thermodynamically consistent CG model approximates the potential of mean force, which depends on the thermodynamic state, e.g., system temperature. Consequently, the model is strictly applicable only under the same conditions it was trained, and targeting the coarse-grained model to different conditions requires a different

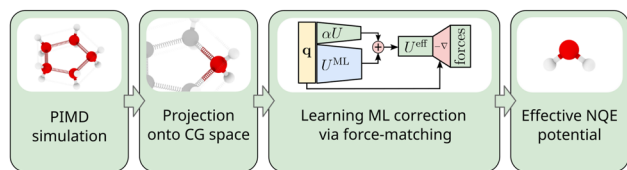


FIG. 1. Procedure to train an effective NQE potential.

set of training data. However, in the case of ring polymer coarse-graining, we notice two cases where the applicability of a trained model can be extended, or training data generated for one condition can be reused for a different one. In the first scenario, the molecule predominantly populates the ground vibrational state, and the population of the excited states is negligible. In this case, the general equation for the quantum PMF [see Eq. (18)] simplifies to

$$U^{\text{eff}} = -\frac{1}{\beta} \ln |\Psi_0|^2, \quad (12)$$

where Ψ_0 denotes the ground-state nuclear wave function.

Consequently, the effective CG potentials at different temperatures in the ground vibrational state are related to each other as

$$U^{\text{eff}}(\beta_2) = \frac{\beta_1}{\beta_2} U^{\text{eff}}(\beta_1). \quad (13)$$

This relationship implies that if a model is trained at a temperature that is low enough to predominately populate the ground state, the same model can be used to run simulations and compute the quantum statistics at any lower temperature by rescaling U^{eff} and the corresponding forces according to Eq. (13).

This simple relation breaks down when the population of the excited vibrational states is non-negligible. However, a dataset generated with an expensive PIMD simulation at one temperature can be used to generate synthetic datasets corresponding to different temperatures and isotope compositions. This data recycling process involves two steps. First, we reweight the configuration to the target temperature and isotope composition, efficiently using the coordinate rescaling approaches from Refs. 78 and 79. Within the coordinate scaling approach, given the initial coordinates \mathbf{q} , inverse temperature β_1 , and mass m_1 , to model a system at inverse temperature β_2 and mass m_2 , we rescale the coordinates as

$$\mathbf{q}'^{(j)} = \bar{\mathbf{q}} - \sqrt{\frac{m_2 \beta_1}{m_1 \beta_2}} (\bar{\mathbf{q}} - \mathbf{q}^{(j)}), \quad (14)$$

where $\bar{\mathbf{q}} = \frac{1}{P} \sum_{j=1}^P \mathbf{q}^{(j)}$ represents the centroid position. Then, the probability $\rho(\mathbf{q})$ of a configuration in the ensemble characterized by m_2 and β_2 is given by (see the [supplementary material](#) for the details)

$$\rho(\mathbf{q}'; \beta_2; m_2) = \rho(\mathbf{q}; \beta_1; m_1) \exp \left(\sum_{j=1}^P (\beta/P_1 \mathbf{q}^{(j)} - \beta/P_2 \mathbf{q}'^{(j)}) \right). \quad (15)$$

We used Eq. (15) to reweight the original dataset.

Second, we recompute the forces originating from the harmonic couplings between adjacent replicas, \mathbf{F}' , for the target temperature and isotope compositions as

$$\mathbf{F}'(\mathbf{q}; m_2; \beta_2) = \frac{\beta_1^2 m_2}{\beta_2^2 m_1} \mathbf{F}'(\mathbf{q}; m_1; \beta_1). \quad (16)$$

The resulting dataset is then used for model training, as described above.

III. COMPUTATIONAL DETAILS

A. 1D Morse potential

1. Model and training procedure

We first demonstrate the approach on a system for which the exact effective potential can be estimated numerically. We study a particle with mass μ experiencing a one-dimensional Morse potential,

$$U^{\text{Morse}}(q) = D_e(1 - e^{-a(q-q_0)})^2, \quad (17)$$

where q is the particle position. The parameters q_0 , a , and D_e , which determine the equilibrium position, shape, and depth of the potential, respectively, are chosen to reproduce the potential of the O–H bond, and μ is set to the reduced mass of the O–H bond (see the [supplementary material](#) for the numeric values).

The effective potential corresponding to the mapping in Eq. (7) can be estimated numerically by solving the Schrödinger equation for the bound states of a Morse potential⁸⁰ as

$$U^{\text{eff}}(q) = -\beta^{-1} \ln \frac{\sum_i e^{-\beta E_i} |\Psi_i(q)|^2}{\sum_i e^{-\beta E_i}}, \quad (18)$$

where E_i and $\Psi_i(q)$ are the energy eigenvalues and wave functions associated with the i th eigenstate, respectively. To estimate the effective potential using the PIGS approach, we optimize a coarse-grained potential in Eq. (11) with $\alpha = P^{-1}$,

$$U^{\text{eff}}(q, \theta) = \frac{U^{\text{Morse}}(q)}{P} + U^{\text{ML}}(q, \theta), \quad (19)$$

where $U^{\text{ML}} = \sum_{i=1}^M \theta_i \phi_i(q)$ is a linear combination of radial basis functions $\phi_i(q)$ (see the [supplementary material](#) for details)⁸¹ and M is the total number of basis functions (we set $M = 10$). Hence, instead of learning the full effective potential, we fit a correction to the classical potential that captures the quantum effects learned from the full-ring polymer simulations. The adjustable parameters θ are obtained by minimizing Eq. (5) using ridge regression, with the L_2 regularization parameter set to 0.1 and fivefold cross-validation.

2. Training data generation

The training data were generated from PIMD simulations using the *i*-PI code⁸² and a custom Python-based force provider implementing the one-dimensional Morse potential. Three simulations were performed in an NVT ensemble for 100, 300, and 600 K, using 64 replicas and a PILE-L thermostat with a time constant of 100 fs.⁸³ The integration was performed with a time step of 0.5 fs, and positions and forces acting on each bead were sampled every 20 steps. The first 2.5 ps of the simulation was discarded, and the positions and forces sampled in the subsequent 500 ps were used to train the model.

B. Molecular systems: Water molecule, Zundel cation, and liquid water

1. Model and training procedure

We use Eq. (11) to model the effective potential of the single water molecule, the Zundel cation, and bulk water. We represent U^{ML} with the MACE graph neural network architecture.⁸⁴ One of

the hallmarks of the MACE architecture is the use of many-body features to represent atomic local environments, improving the model accuracy and data efficiency.

The models were trained with the AdamW optimizer, a learning rate of 0.001, and a weight decay of 0.001. The MACE neural network cutoff was set to 6.0 Å, with 15 radial basis functions, 6 polynomial, and 3 radial channels. The correlation order for Zundel cation and bulk water was set to 3, which corresponds to four-body features. For a single water molecule, three-body features were used. In all cases, 80% of the data available were used for training and the remaining 20% were used for validation. The training was performed for 60 epochs for the H₂O molecule and the Zundel cation and for 15 epochs for liquid H₂O. For each system, ten different models were trained, each using different parameter initialization and train-validation splitting.

2. Training data generation

The training data were generated from PIMD simulations using the *i*-PI code using the Partridge and Schwenke potential for a single water molecule,⁸⁵ the HBB potential⁸⁶ for the Zundel cation, and the MB-pol for bulk water.⁸⁷ The simulation parameters are summarized in Table S1. For each system, all the data generated during the production phase of PIMD simulation were used as a training dataset. Since all the replicas are equivalent, each frame of the trajectory yields P data points for model training. We use a force map optimized to minimize the average magnitude of the mapped forces, subject to the constraints of the valid force map. This force optimization was performed according to the procedure introduced in Ref. 77, with an L_2 regularization parameter set to 0.05.

3. Model evaluation

For each of the trained models, we performed molecular dynamics simulations, using as a force field a combination of classical potential and a trained correction as defined in Eq. (11). The simulations were performed in the NVT ensemble, with a Langevin integrator and a friction thermostat with a time constant set to 100 ps⁻¹. The time step and periodic boundary conditions were the same as in the training simulations. All the frames were used for further analysis. For each of the systems, we performed additional simulations with only the classical potential as a control to compare the classical and quantum statistics.

IV. RESULTS AND DISCUSSION

A. 1D Morse potential

First, we study a particle in a one-dimensional Morse potential, a model for a single O–H bond. The Schrödinger equation for this system can be solved analytically,⁸⁰ thus making Morse potential an ideal test system. It also exhibits significant nuclear quantum effects, as shown in Fig. 2(a), where the classical Morse potential [Eq. (17)] is compared with the corresponding quantum PMFs [Eq. (18)] at different temperatures. The classical approximation leads to the over-localization of the particle, and, as expected, the quantum effects are more prominent at low temperatures. However, for this system, even at a temperature of 600 K, there are significant deviations between the classical and quantum pictures.

The developed approach allows us to get a quantitative agreement with the reference, especially in the high-probability regions

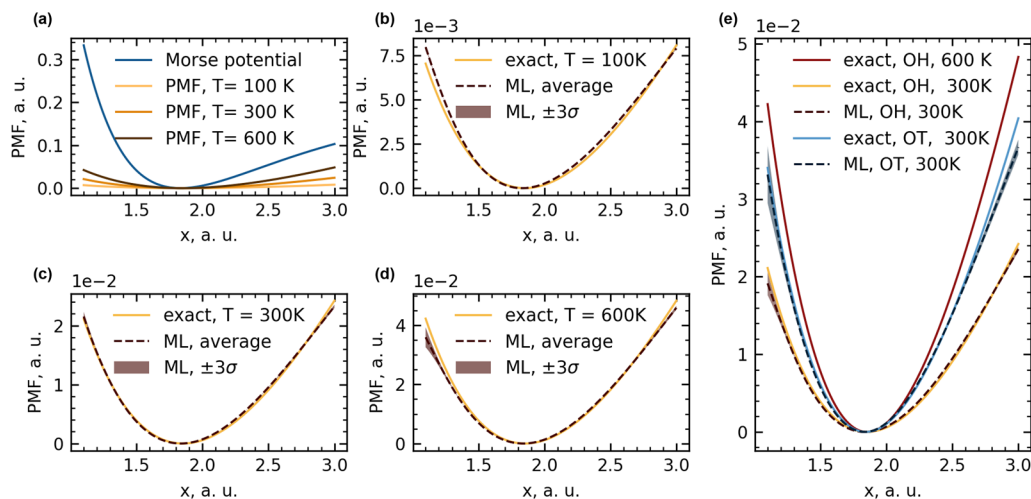


FIG. 2. Exact and machine-learned PMF for the Morse potential. (a) Classical Morse potential and quantum PMFs at different temperatures. Even at 600 K, the system exhibits significant quantum effects. (b)–(d) Comparison of exact quantum PMF and machine-learned PMF at 100 K (b), 300 K (c), and 600 K (d). (e) Results for models trained at 300 K for OH and OT from reweighted simulations performed at 600 K. The average PMF over five models is reported as a dashed line, and ± 3 standard deviations are reported as shaded areas.

that are well-sampled in the training dataset [Figs. 2(b)–2(d)]. Further away from the minimum energy, the model's prediction slightly deviates from the ground truth. This is to be expected, as these regions of the phase space are not as well represented in the training data as the high-probability regions.

We test the reweighting scheme proposed in Sec. II E by generating datasets corresponding to different temperatures (300 K) and mass of the particle that is either protium (corresponding to an O–H bond) or increased to tritium (corresponding to an O–T bond) from the 600 K O–H dataset previously used. The resulting reweighted datasets were used to train new models. As shown in Fig. 2(e), the models trained with the reweighted dataset are in excellent agreement with the analytical results for the corresponding temperature and isotope composition. Such data recycling allows for a further decrease in the computational cost of generating training data, although it becomes less effective for more complex systems. Nevertheless, we expect that this approach may be particularly useful for generating large datasets for training transferable models.

B. Water molecule

The next test case we considered is a single water molecule in vacuum. The quantum statistics of this three-body system can be exhaustively described in terms of the length of two O–H bonds and the H–O–H bond angle. The comparison of the probability distributions for the O–H distances and H–O–H bond angles is shown in Fig. 3, with additional details provided in Fig. S4. As was the case for the Morse potential, the classical distribution is significantly more localized than the quantum counterpart. The ML model consistently samples the correct quantum distribution: the results obtained from ten independently trained models produce similar results with low variance.

For all trained models, we run the MD simulation for 1 ns, which is four times longer than the PIMD simulation used to generate the training data. This allows us to obtain slowly converging properties by performing short PIMD simulations and using those to train a model that can be run for a much longer time.

We further investigated the impact of PIMD simulation length on model performance by training additional models using 2.5, 12.5, 25, and 125 ps of post-equilibration PIMD data, compared

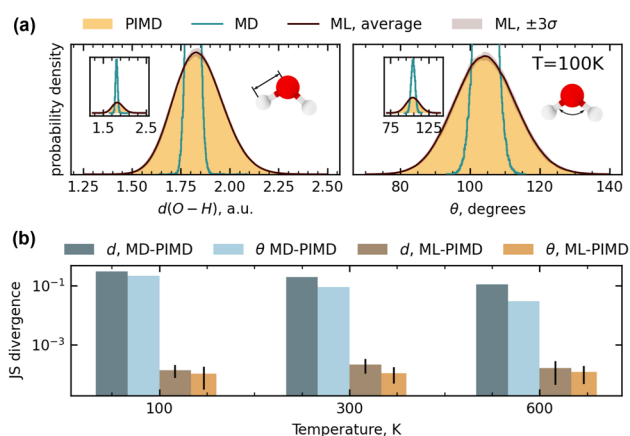


FIG. 3. (a) Comparison of the probability distributions of the O–H bond length (left) and the H–O–H angle (right) for classical MD ($P = 1$), PIMD ($P = 64$), and ML models at 100 K. The insets show the same distributions at a different scale. (b) Jensen–Shannon (JS) divergence between collective variable distributions obtained with PIMD and either classical potential (shades of blue) or with our machine-learned model (shades of brown). For each case, JS divergence for the O–H bond length distribution and the H–O–H angle distribution at three different temperatures are shown.

to the primary model trained with 250 ps. All models generated stable 1 ns trajectories. Notably, even with just 5% of the original dataset, the models showed on average good agreement with the reference PIMD data [Figs. S1 and S2]. As the dataset size increased, the Jensen–Shannon divergence decreased and eventually plateaued by 250 ps (Fig. S3).

Even at 600 K, the water molecule predominantly populates the ground vibrational state, with a negligible population of the excited states. Thus, as discussed in Sec. II E, the model trained at 600 K can be used to simulate the system at any lower temperatures. To demonstrate this point further, we report in Fig. S5 the results obtained with a model trained on PIMD data generated at 600 K and then simulated at lower temperatures, with forces rescaled according to Eq. (13).

C. Zundel cation

A more challenging case is the Zundel cation, which consists of two water molecules sharing a hydrogen-bonded proton H^* . This system has been extensively studied both theoretically^{88,89} and experimentally.^{90–94} The Zundel cation exhibits large conformational fluctuations, including anharmonic and floppy motion associated with the transfer of the hydrogen-bounded proton. Following Ref. 89, we characterize these fluctuations by means of the O–O distance $d(O-O)$, the proton transfer coordinate $\delta = d(O_1H^*) - d(O_2H^*)$, the dihedral angle ϕ that measures the rotation of the H_3O^+ fragments around the O– H^* –O axis, and the $d(O-plane)$ that measures the non-planarity of the H_3O^+ fragment. The dihedral

angle ϕ is defined by points $X_1-O_1-O_2-X_2$, where X_1 and X_2 represent a midpoint between the hydrogen atoms that are not involved in the hydrogen bond formation and bound to O_1 and O_2 , correspondingly. The non-planarity of H_3O^+ is characterized by a distance between an O atom and the plane, defined by the three closest H atoms (see the lower panel of Fig. 4).

At 300 K, the proton transfer coordinate and the rotational motion are almost classical, and only a small correction is needed to account for NQEs. The O–O distance and the O–plane distance require larger corrections, yet the NQEs are relatively small. At 100 K, however, significant quantum effects appear in all the collective variables, including changes in the equilibrium geometry. Nevertheless, our methods recover the accurate quantum distribution both at 300 K and at 100 K. Despite the diversity of conformational fluctuations, the model not only reproduces the target probability distributions but also leads to remarkably stable simulations: all the models were used to perform 1 ns simulation and did not exhibit any instabilities, often plaguing ML force fields.^{95,96}

D. Liquid water

Finally, we applied our approach to model liquid water to reproduce the NQEs manifesting in both bonded and nonbonded interactions. Since NQEs are highly local, the former are more affected than the latter. The quality of the machine-learned model is evaluated in Fig. 5 by computing radial distribution functions (RDFs). At 300 K, by computing the first peak in the O–H [Fig. 5(b)] and H–H [Fig. 5(c)] RDFs predicted by classical MD ($P = 1$) are too narrow and significantly deviate from the correct quantum PIMD results.

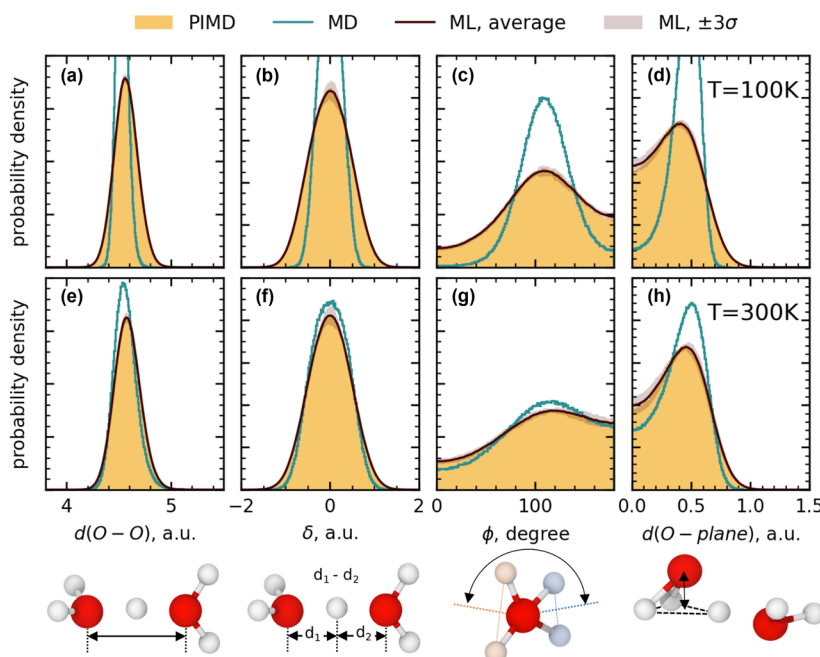


FIG. 4. Probability distribution for the Zundel cation, sampled with the PIMD simulation, the MD simulation, and the proposed ML model. The top row [panels (a)–(d)] represents the model trained and simulated at 100 K, and the bottom row [panels (e)–(h)] shows the results trained at 300 K. Under the plots, a visual representation of the corresponding collective variable is shown.

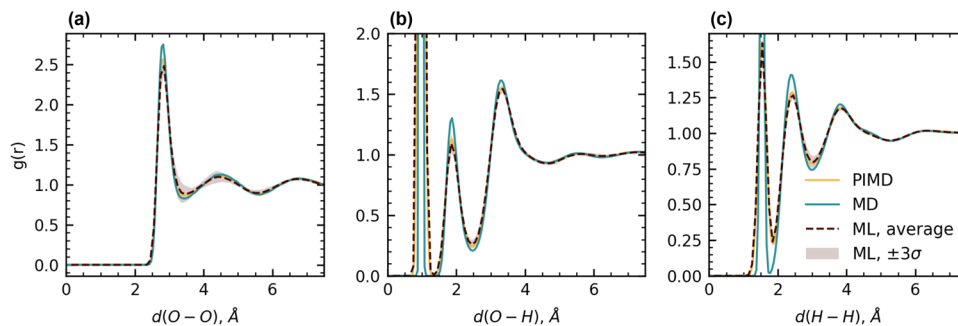


FIG. 5. Radial distribution functions $g(r)^{OO}$, $g(r)^{OH}$, and $g(r)^{HH}$ for bulk H_2O at 300 K. The teal line represents the RDF obtained with a classical MD simulation ($P = 1$), the orange line represents the PIMD result, and the brown dashed line represents the results obtained with the ML model. The result reported for the ML model is the mean over ten independent models with different train–test splits and model initialization.

The deviations between the classical and quantum RDFs rapidly decrease with the distance.

Our neural-network-based potential is able to accurately predict the correct quantum RDFs, including the contributions that arise from both bonded and nonbonded interactions (Fig. 5), in contrast to another similar study.⁶⁸ There is also a remarkable agreement between independently trained models. The high accuracy and low uncertainty of our ML model prediction show that it reliably captures the subtle differences between classical and quantum RDFs.

Compared to the simulations for the ML models of a single water molecule and of the Zundel cation, the simulations for bulk water exhibit a decreased stability. We note that some simulations of our ML models, trained on 35 ps of PIMD simulations, exhibit numerical instabilities. However, the ML models allow, on average, at least 38 ps of simulation time before the onset of any instability. Simulations can be reinitialized and restarted to improve statistics. The numerical instability of the model can be potentially cured by increasing the size of the training dataset. Alternatively, one can modify the learning problem by simplifying the ML correction, for example, by introducing physics-informed energy terms, as it is done in the learning of classical ML coarse-grained force-fields.^{73–76} Given that the stability of the simulation strongly depends on the parameter α [see Eq. (11)], this approach may be especially promising.

V. CONCLUSION

We developed an approach to coarse-grain the ring-polymer dynamics describing NQEs, and we obtain results thermodynamically consistent with PIMD simulations. This approach allows us to accurately compute position-dependent NQEs at the cost of classical MD. We have demonstrated this method on four distinct systems where NQEs are prevalent—the Morse potential, a single water molecule, the Zundel cation, and bulk water—and reported results in excellent agreement with (much more expensive) reference PIMD simulations.

The proposed method builds an effective NQE correction potential by combining the rigorous framework of multiscale

coarse-graining⁷² with the MACE architecture,⁸⁴ an expressive state-of-the-art graph neural network potential. We have shown that these effective potentials accurately reproduce the quantum statistics of nuclei, both near the classical limit and in the regime of strong NQE. Moreover, our method eliminates the need to simulate multiple replicas. Once parameterized, the model seamlessly integrates into existing classical force fields and does not require specialized algorithms other than ones already used for classical MD simulations. However, this approach is limited to reproducing quantum thermodynamics, so dynamical and momentum-dependent properties should be approximated using a previously proposed method.⁶⁶

For the results presented here, different sets of training data were used to train models for different molecular species at different temperatures. However, if the temperature associated with the training data is low enough to significantly populate only the quantum ground state, the trained model can be used to also simulate any lower temperature, as we have demonstrated in the case of the H_2O molecule. The proposed method is designed such that the forces originating from the coupling between the adjacent replicas have a non-zero contribution to the force projected onto the coarse-grained space. As the dependence of these forces on the atomic mass and temperature is known [Eq. (9)], the training dataset obtained for one isotope composition and temperature can be, in principle, reweighted to construct a dataset for different isotope compositions and temperatures. Thus, a single PIMD simulation can provide data for training several models, as we have shown in the case of the Morse potential.

The rapid proliferation of machine-learned force fields parameterized using accurate quantum chemical calculations to account for the electronic effects means that the missing NQEs are now becoming a dominant source of error. At the same time, NQEs are crucial to describe biomolecular processes accurately, for instance, in the case of proton transfer reactions. While the results presented here are model-specific, building upon the recent development of machine-learned force-fields (either coarse-grained^{73–76} or atomistic^{81,84,97–100}), we believe that this study represents a significant step forward toward the design of transferable potentials that explicitly incorporate also NQEs but retain the computational cost of classical MD.

SUPPLEMENTARY MATERIAL

Section I of [supplementary material](#) details the Morse potential (Sec. I A) and ML model used for a particle in Morse potential (Sec. I B). Section II contains the simulation parameters used to generate the training data. Section III provides the derivation of coordinate rescaling. Section IV presents extended results for a water molecule.

ACKNOWLEDGMENTS

We thank Gregory Voth and members of the Clementi's group at FU for insightful discussions and comments on the manuscript. We acknowledge funding from the Deutsche Forschungsgemeinschaft (DFG) (Grant Nos. SFB/TRR 186, Project A12; SFB 1114, Projects B03, B08, and A04; and SFB 1078, Project C7), the National Science Foundation (Grant No. PHY-2019745), and the Einstein Foundation Berlin (Project No. 0420815101), and we also acknowledge the computing time provided on the supercomputer Lise at NHR@ZIB as part of the NHR infrastructure (Project No. beb00040), the HPC Service of FUB-IT, Freie Universität Berlin, the HPC Service of the Physics department, Freie Universität Berlin, and the Swiss National Supercomputing Centre (under Project Nos. s1209 and s1288). V.K. acknowledges the support from the Ernest Oppenheimer Early Career Fellowship and the Sydney Harvey Junior Research Fellowship, Churchill College, University of Cambridge.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Iryna Zaporozhets: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Validation (equal); Visualization (equal); Writing – original draft (equal). **Félix Musil:** Conceptualization (equal); Data curation (supporting); Formal analysis (supporting); Methodology (equal); Software (equal); Validation (supporting); Visualization (supporting); Writing – original draft (supporting); Writing – review & editing (equal). **Venkat Kapil:** Conceptualization (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Supervision (equal); Validation (equal); Writing – review & editing (equal). **Cecilia Clementi:** Conceptualization (lead); Funding acquisition (lead); Investigation (supporting); Methodology (equal); Project administration (lead); Resources (lead); Supervision (lead); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

All the training data, code, and scripts that support the findings of this study are openly available in the online repository: `cg_nuclear_quantum_statistics`.

REFERENCES

- S. Chmiela, V. Vassilev-Galindo, O. T. Unke, A. Kabylda, H. E. Sauceda, A. Tkatchenko, and K.-R. Müller, *Sci. Adv.* **9**, eadf0873 (2023).
- I. Batatia, P. Benner, Y. Chiang, A. M. Elena, D. P. Kovács, J. Riebesell, X. R. Advincula, M. Asta, W. J. Baldwin, N. Bernstein, A. Bhowmik, S. M. Blau, V. Cărare, J. P. Darby, S. De, F. D. Pia, V. L. Deringer, R. Elijošius, Z. El-Machachi, E. Fako, A. C. Ferrari, A. Genreith-Schriever, J. George, R. E. A. Goodall, C. P. Grey, S. Han, W. Handley, H. H. Heenen, K. Hermansson, C. Holm, J. Jaafar, S. Hofmann, K. S. Jakob, H. Jung, V. Kapil, A. D. Kaplan, N. Karimitari, N. Kroupa, J. Kullgren, M. C. Kuner, D. Kuryla, G. Liepuoniute, J. T. Margraf, I.-B. Magdău, A. Michaelides, J. H. Moore, A. A. Naik, S. P. Niblett, S. W. Norwood, N. O'Neill, C. Ortner, K. A. Persson, K. Reuter, A. S. Rosen, L. L. Schaaf, C. Schran, E. Sivonxay, T. K. Stenczel, V. Svahn, C. Sutton, C. van der Oord, E. Varga-Umbrich, T. Vegge, M. Vondrák, Y. Wang, W. C. Witt, F. Zills, and G. Csányi, *arXiv:2401.00096* [physics.chem-ph] (2023).
- D. van der Spoel, *Curr. Opin. Struct. Biol.* **67**, 18 (2021).
- J. Jung, C. Kobayashi, K. Kasahara, C. Tan, A. Kuroda, K. Minami, S. Ishiduki, T. Nishiki, H. Inoue, Y. Ishikawa, M. Feig, and Y. Sugita, *J. Comput. Chem.* **42**, 231 (2021).
- D. E. Shaw, P. J. Adams, A. Azaria, J. A. Bank, B. Batson, A. Bell, M. Bergdorf, J. Bhatt, J. A. Butts, T. Correia, R. M. Dirks, R. O. Dror, M. P. Eastwood, B. Edwards, A. Even, P. Feldmann, M. Fenn, C. H. Fenton, A. Forte, J. Gagliardo, G. Gill, M. Gorlatova, B. Greskamp, J. Grossman, J. Gullingsrud, A. Harper, W. Hasenplough, M. Heily, B. C. Heshmat, J. Hunt, D. J. Ierardi, L. Iserovich, B. L. Jackson, N. P. Johnson, M. M. Kirk, J. L. Klepeis, J. S. Kuskin, K. M. Mackenzie, R. J. Mader, R. McGowen, A. McLaughlin, M. A. Moraes, M. H. Nasr, L. J. Nociolo, L. O'Donnell, A. Parker, J. L. Peticolas, G. Pocina, C. Predescu, T. Quan, J. K. Salmon, C. Schwink, K. S. Shim, N. Siddique, J. Spengler, T. Szalay, R. Tabladillo, R. Tartler, A. G. Taube, M. Theobald, B. Towles, W. Vick, S. C. Wang, M. Wazlowski, M. J. Weingarten, J. M. Williams, and K. A. Yuh, in *Proceedings of International Conference for High Performance Computing, Networking, Storage and Analysis, SC '21* (Association for Computing Machinery, New York, 2021).
- M. Born and R. Oppenheimer, *Ann. Phys.* **389**, 457 (1927).
- G. R. Medders and F. Paesani, *J. Am. Chem. Soc.* **138**, 3912 (2016).
- O. Marsalek and T. E. Markland, *J. Phys. Chem. Lett.* **8**, 1545 (2017).
- L. Pereyaslavets, I. Kurnikov, G. Kamath, O. Butin, A. Illarionov, I. Leontyev, M. Olevanov, M. Levitt, R. D. Kornberg, and B. Fain, *Proc. Natl. Acad. Sci. U. S. A.* **115**, 8878 (2018).
- J. Daru, H. Forbert, J. Behler, and D. Marx, *Phys. Rev. Lett.* **129**, 226001 (2022).
- H. A. Stern and B. J. Berne, *J. Chem. Phys.* **115**, 7622 (2001).
- M. Shiga and W. Shinoda, *J. Chem. Phys.* **123**, 134502 (2005).
- F. Paesani, W. Zhang, D. A. Case, T. E. Cheatham III, and G. A. Voth, *J. Chem. Phys.* **125**, 184507 (2006).
- J. A. Morrone and R. Car, *Phys. Rev. Lett.* **101**, 017801 (2008).
- S. Habershon, T. E. Markland, and D. E. Manolopoulos, *J. Chem. Phys.* **131**, 024501 (2009).
- C. Vega, M. M. Conde, C. McBride, J. L. F. Abascal, E. G. Noya, R. Ramirez, and L. M. Sesé, *J. Chem. Phys.* **132**, 046101 (2010).
- B. Pamuk, J. M. Soler, R. Ramírez, C. P. Herrero, P. W. Stephens, P. B. Allen, and M.-V. Fernández-Serra, *Phys. Rev. Lett.* **108**, 193003 (2012).
- S. Fritsch, R. Potestio, D. Donadio, and K. Kremer, *J. Chem. Theory Comput.* **10**, 816 (2014).
- L. Wang, M. Ceriotti, and T. E. Markland, *J. Chem. Phys.* **141**, 104502 (2014).
- Y. Litman, D. Donadio, M. Ceriotti, and M. Rossi, *J. Chem. Phys.* **148**, 102320 (2017).
- S. Shepherd, J. Lan, D. M. Wilkins, and V. Kapil, *J. Phys. Chem. Lett.* **12**, 9108 (2021).
- A. Eltareb, G. E. Lopez, and N. Giovambattista, *J. Phys. Chem. B* **127**, 4633 (2023).
- V. Kapil, D. P. Kovacs, G. Csányi, and A. Michaelides, *Faraday Discuss.* **249**, 50 (2024).
- M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello, *Science* **275**, 817 (1997).
- B. Walker and A. Michaelides, *J. Chem. Phys.* **133**, 174306 (2010).

- ²⁶M. Ceriotti, J. Cuny, M. Parrinello, and D. E. Manolopoulos, *Proc. Natl. Acad. Sci. U. S. A.* **110**, 15591 (2013).
- ²⁷W. Fang, J. Chen, M. Rossi, Y. Feng, X.-Z. Li, and A. Michaelides, *J. Phys. Chem. Lett.* **7**, 2125 (2016).
- ²⁸V. Kapil, E. Engel, M. Rossi, and M. Ceriotti, *J. Chem. Theory Comput.* **15**, 5845 (2019).
- ²⁹V. Kapil and E. A. Engel, *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2111769119 (2022).
- ³⁰S. J. Buxton, D. Quigley, and S. Habershon, *J. Chem. Phys.* **151**, 144503 (2019).
- ³¹J. R. Štoček, O. Socha, I. Čásařová, T. Slanina, and M. Dračinský, *J. Am. Chem. Soc.* **144**, 7111 (2022).
- ³²H. E. Saucedo, V. Vassilev-Galindo, S. Chmiela, K.-R. Müller, and A. Tkatchenko, *Nat. Commun.* **12**, 442 (2021).
- ³³D. T. Major, A. Heroux, A. M. Orville, M. P. Valley, P. F. Fitzpatrick, and J. Gao, *Proc. Natl. Acad. Sci. U. S. A.* **106**, 20734 (2009).
- ³⁴R. Feynman and A. Hibbs, *Quantum Mechanics and Path Integrals, International Series in Pure and Applied Physics* (McGraw-Hill, 1965).
- ³⁵D. Chandler and P. G. Wolynes, *J. Chem. Phys.* **74**, 4078 (1981).
- ³⁶B. J. Berne and D. Thirumalai, *Annu. Rev. Phys. Chem.* **37**, 401 (1986).
- ³⁷M. E. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation* (Oxford University Press, 2023).
- ³⁸T. E. Markland and M. Ceriotti, *Nat. Rev. Chem.* **2**, 0109 (2018).
- ³⁹F. Uhl, D. Marx, and M. Ceriotti, *J. Chem. Phys.* **145**, 054101 (2016).
- ⁴⁰M. Takahashi and M. Imada, *J. Phys. Soc. Jpn.* **53**, 3765 (1984).
- ⁴¹S. A. Chin, *Phys. Lett. A* **226**, 344 (1997).
- ⁴²A. Pérez and M. E. Tuckerman, *J. Chem. Phys.* **135**, 064104 (2011).
- ⁴³V. Kapil, J. Behler, and M. Ceriotti, *J. Chem. Phys.* **145**, 234103 (2016).
- ⁴⁴V. Kapil, J. Wieme, S. Vandenbrande, A. Lamaire, V. Van Speybroeck, and M. Ceriotti, *J. Chem. Theory Comput.* **15**, 3237 (2019).
- ⁴⁵Y. Kamibayashi and S. Miura, *J. Chem. Phys.* **145**, 074114 (2016).
- ⁴⁶I. Poltavsky, V. Kapil, M. Ceriotti, K. S. Kim, and A. Tkatchenko, *J. Chem. Theory Comput.* **16**, 1128 (2020).
- ⁴⁷I. Poltavsky and A. Tkatchenko, *Chem. Sci.* **7**, 1368 (2016).
- ⁴⁸V. Kapil, J. VandeVondele, and M. Ceriotti, *J. Chem. Phys.* **144**, 054111 (2016).
- ⁴⁹T. E. Markland and D. E. Manolopoulos, *Chem. Phys. Lett.* **464**, 256 (2008).
- ⁵⁰M. Ceriotti, G. Bussi, and M. Parrinello, *Phys. Rev. Lett.* **103**, 030603 (2009).
- ⁵¹H. Dammak, Y. Chalopin, M. Laroche, M. Hayoun, and J.-J. Greffet, *Phys. Rev. Lett.* **103**, 190601 (2009).
- ⁵²M. Ceriotti and D. E. Manolopoulos, *Phys. Rev. Lett.* **109**, 100604 (2012).
- ⁵³R. Korol, N. Bou-Rabee, and T. F. Miller III, *J. Chem. Phys.* **151**, 124103 (2019).
- ⁵⁴E. Wigner, *Phys. Rev.* **40**, 749 (1932).
- ⁵⁵J. G. Kirkwood, *Phys. Rev.* **44**, 31 (1933).
- ⁵⁶R. P. Feynman, *Statistical Mechanics* (Benjamin, New York, 1972).
- ⁵⁷J. Cao and G. A. Voth, *J. Chem. Phys.* **100**, 5093 (1994).
- ⁵⁸J. Cao and G. A. Voth, *J. Chem. Phys.* **100**, 5106 (1994).
- ⁵⁹J. Cao and G. A. Voth, *J. Chem. Phys.* **101**, 6157 (1994).
- ⁶⁰J. Cao and G. A. Voth, *J. Chem. Phys.* **101**, 6168 (1994).
- ⁶¹J. Cao and G. A. Voth, *J. Chem. Phys.* **101**, 6184 (1994).
- ⁶²N. V. Blinov, P.-N. Roy, and G. A. Voth, *J. Chem. Phys.* **115**, 4484 (2001).
- ⁶³N. Blinov and P.-N. Roy, *J. Chem. Phys.* **115**, 7822 (2001).
- ⁶⁴P.-N. Roy and N. Blinov, *Isr. J. Chem.* **42**, 183 (2002).
- ⁶⁵T. D. Hone, S. Izvekov, and G. A. Voth, *J. Chem. Phys.* **122**, 054105 (2005).
- ⁶⁶F. Musil, I. Zaporozhets, F. Noé, C. Clementi, and V. Kapil, *J. Chem. Phys.* **157**, 181102 (2022).
- ⁶⁷T. D. Loose, P. G. Sahrman, and G. A. Voth, *J. Chem. Theory Comput.* **18**, 5856 (2022).
- ⁶⁸I. V. Kurnikov, L. Pereyaslavets, G. Kamath, S. N. Sakipov, E. Voronina, O. Butin, A. Illarionov, I. Leontyev, G. Nawrocki, M. Darkhovskiy, M. Olevanov, I. Ivahnenko, Y. Chen, C. B. Lock, M. Levitt, R. D. Kornberg, and B. Fain, *J. Chem. Theory Comput.* **20**, 1347 (2024).
- ⁶⁹W. H. Ryu, Y. Han, and G. A. Voth, *J. Chem. Phys.* **150**, 244103 (2019).
- ⁷⁰A. V. Sinititskiy and G. A. Voth, *J. Chem. Phys.* **143**, 094104 (2015).
- ⁷¹W. H. Ryu and G. A. Voth, *J. Phys. Chem. A* **126**, 6004 (2022).
- ⁷²W. G. Noid, J.-W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, and H. C. Andersen, *J. Chem. Phys.* **128**, 244114 (2008).
- ⁷³J. Wang, S. Olsson, C. Wehmeyer, A. Pérez, N. E. Charron, G. de Fabritiis, F. Noé, and C. Clementi, *ACS Cent. Sci.* **5**, 755 (2019).
- ⁷⁴B. E. Husic, N. E. Charron, D. Lemm, J. Wang, A. Pérez, M. Majewski, A. Krämer, Y. Chen, S. Olsson, G. de Fabritiis, F. Noé, and C. Clementi, *J. Chem. Phys.* **153**, 194101 (2020).
- ⁷⁵M. Majewski, A. Pérez, P. Thölke, S. Doerr, N. E. Charron, T. Giorgino, B. E. Husic, C. Clementi, F. Noé, and G. De Fabritiis, *Nat. Commun.* **14**, 5739 (2023).
- ⁷⁶N. E. Charron, F. Musil, A. Guljas, Y. Chen, K. Bonneau, A. S. Pasos-Trejo, J. Venturin, D. Gusew, I. Zaporozhets, A. Krämer, C. Templeton, A. Kelkar, A. E. P. Durumeric, S. Olsson, A. Pérez, M. Majewski, B. E. Husic, A. Patel, G. De Fabritiis, F. Noé, and C. Clementi, [arXiv:2310.18278](https://arxiv.org/abs/2310.18278) (2023).
- ⁷⁷A. Krämer, A. E. P. Durumeric, N. E. Charron, Y. Chen, C. Clementi, and F. Noé, *J. Phys. Chem. Lett.* **14**, 3970 (2023).
- ⁷⁸T. M. Yamamoto, *J. Chem. Phys.* **123**, 104101 (2005).
- ⁷⁹M. Ceriotti and T. E. Markland, *J. Chem. Phys.* **138**, 014112 (2013).
- ⁸⁰J. P. Dahl and M. Springborg, *J. Chem. Phys.* **88**, 4535 (1988).
- ⁸¹O. T. Unke and M. Meuwly, *J. Chem. Theory Comput.* **15**, 3678 (2019).
- ⁸²V. Kapil, M. Rossi, O. Marsalek, R. Petraglia, Y. Litman, T. Spura, B. Cheng, A. Cuzzocrea, R. H. Meißner, D. M. Wilkins, B. A. Helfrecht, P. Juda, S. P. Bienvenue, W. Fang, J. Kessler, I. Poltavsky, S. Vandenbrande, J. Wieme, C. Corminboeuf, T. D. Kuhne, D. E. Manolopoulos, T. E. Markland, J. O. Richardson, A. Tkatchenko, G. A. Tribello, V. Van Speybroeck, and M. Ceriotti, *Comput. Phys. Commun.* **236**, 214 (2019).
- ⁸³M. Ceriotti, M. Parrinello, T. E. Markland, and D. E. Manolopoulos, *J. Chem. Phys.* **133**, 124104 (2010).
- ⁸⁴I. Batatia, D. P. Kovacs, G. Simm, C. Ortner, and G. Csanyi, “MACE: Higher order equivariant message passing neural networks for fast and accurate force fields,” in *Advances in Neural Information Processing Systems*, edited by S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (Curran Associates, Inc., 2022), Vol. 35, pp. 11423–11436.
- ⁸⁵H. Partridge and D. W. Schwenke, *J. Chem. Phys.* **106**, 4618 (1997).
- ⁸⁶X. Huang, B. J. Braams, and J. M. Bowman, *J. Chem. Phys.* **122**, 044308 (2005).
- ⁸⁷V. Babin, C. Leforestier, and F. Paesani, *J. Chem. Theory Comput.* **9**, 5395 (2013).
- ⁸⁸M. Schröder, F. Gatti, D. Lauvergnat, H.-D. Meyer, and O. Vendrell, *Nat. Commun.* **13**, 6170 (2022).
- ⁸⁹K. Suzuki, M. Tachikawa, and M. Shiga, *J. Chem. Phys.* **138**, 184307 (2013).
- ⁹⁰J. M. Headrick, J. C. Bopp, and M. A. Johnson, *J. Chem. Phys.* **121**, 11523 (2004).
- ⁹¹E. G. Diken, J. M. Headrick, J. R. Roscioli, J. C. Bopp, M. A. Johnson, and A. B. McCoy, *J. Phys. Chem. A* **109**, 1487 (2005).
- ⁹²N. I. Hammer, E. G. Diken, J. R. Roscioli, M. A. Johnson, E. M. Myshakin, K. D. Jordan, A. B. McCoy, X. Huang, J. M. Bowman, and S. Carter, *J. Chem. Phys.* **122**, 244301 (2005).
- ⁹³L. R. McCunn, J. R. Roscioli, M. A. Johnson, and A. B. McCoy, *J. Phys. Chem. B* **112**, 321 (2008).
- ⁹⁴O. Vendrell, F. Gatti, and H. Meyer, *Angew. Chem.* **121**, 358–361 (2008).
- ⁹⁵X. Fu, Z. Wu, W. Wang, T. Xie, S. Ketten, R. Gomez-Bombarelli, and T. Jaakkola, *Transactions on Machine Learning Research* (OpenReview, 2023), <https://openreview.net/forum?id=A8pqQipwkt>.
- ⁹⁶S. Stocker, J. Gasteiger, F. Becker, S. Günnemann, and J. T. Margraf, *Mach. Learn. Sci. Technol.* **3**, 045010 (2022).
- ⁹⁷K. T. Schütt, H. E. Saucedo, P.-J. Kindermans, A. Tkatchenko, and K.-R. Müller, *J. Chem. Phys.* **148**, 241722 (2018).
- ⁹⁸O. T. Unke, S. Chmiela, M. Gastegger, K. T. Schütt, H. E. Saucedo, and K.-R. Müller, *Nat. Commun.* **12**, 7273 (2021).
- ⁹⁹O. T. Unke, S. Chmiela, H. E. Saucedo, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko, and K.-R. Müller, *Chem. Rev.* **121**, 10142 (2021).
- ¹⁰⁰S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt, and B. Kozinsky, *Nat. Commun.* **13**, 2453 (2022).