

Machine learning phase space quantum dynamics approaches

Cite as: J. Chem. Phys. **154**, 184104 (2021); <https://doi.org/10.1063/5.0046689>

Submitted: 05 February 2021 • Accepted: 16 April 2021 • Published Online: 11 May 2021

 Xinzijian Liu,  Linfeng Zhang and  Jian Liu

COLLECTIONS

Paper published as part of the special topic on [Quantum Dynamics with ab Initio Potentials](#)



[View Online](#)



[Export Citation](#)



[CrossMark](#)

ARTICLES YOU MAY BE INTERESTED IN

Machine learning meets chemical physics

The Journal of Chemical Physics **154**, 160401 (2021); <https://doi.org/10.1063/5.0051418>

When do short-range atomistic machine-learning models fall short?

The Journal of Chemical Physics **154**, 034111 (2021); <https://doi.org/10.1063/5.0031215>

Chemical physics software

The Journal of Chemical Physics **155**, 010401 (2021); <https://doi.org/10.1063/5.0059886>

[Submit Today!](#)

The Journal
of Chemical Physics

SPECIAL TOPIC: Low-Dimensional
Materials for Quantum Information Science

Machine learning phase space quantum dynamics approaches

Cite as: J. Chem. Phys. 154, 184104 (2021); doi: 10.1063/5.0046689

Submitted: 5 February 2021 • Accepted: 16 April 2021 •

Published Online: 11 May 2021



View Online



Export Citation



CrossMark

Xinzijian Liu,¹ Linfeng Zhang,² and Jian Liu^{1,a)}

AFFILIATIONS

¹ Beijing National Laboratory for Molecular Sciences, Institute of Theoretical and Computational Chemistry, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

² Beijing Institute of Big Data Research, Beijing 100871, China

Note: This paper is part of the JCP Special Topic on Quantum Dynamics with *Ab Initio* Potentials.

a) Author to whom correspondence should be addressed: jianliupku@pku.edu.cn. URL: <http://jianliugroup.pku.edu.cn>

ABSTRACT

Derived from phase space expressions of the quantum Liouville theorem, equilibrium continuity dynamics is a category of trajectory-based phase space dynamics methods, which satisfies the two critical fundamental criteria: conservation of the quantum Boltzmann distribution for the thermal equilibrium system and being exact for any thermal correlation functions (even of nonlinear operators) in the classical and harmonic limits. The effective force and effective mass matrix are important elements in the equations of motion of equilibrium continuity dynamics, where only the zeroth term of an exact series expansion of the phase space propagator is involved. We introduce a machine learning approach for fitting these elements in quantum phase space, leading to a much more efficient integration of the equations of motion. Proof-of-concept applications to realistic molecules demonstrate that machine learning phase space dynamics approaches are possible as well as competent in producing reasonably accurate results with a modest computation effort.

Published under license by AIP Publishing. <https://doi.org/10.1063/5.0046689>

I. INTRODUCTION

A great deal of theoretical effort has been focused on developing approaches for including nuclear quantum mechanical effects in trajectory-based molecular dynamics (MD) simulations^{1–6} for calculating quantum thermal correlation functions that encode most dynamical quantities of interest.^{7–9} For example, energy current correlation functions yield thermal conductivities,^{9,10} flux correlation functions produce reaction rates,^{11,12} force correlation functions are related to vibrational energy relaxation rate constants,¹³ and infrared(IR)/Raman vibrational spectra can be expressed in terms of collective dipole-moment/polarizability correlation functions.^{8,9,14–19} Quantum thermal correlation functions are of the form^{1–9}

$$C_{AB}(t) = \frac{1}{Z} \text{Tr} [\hat{A}^\beta(0) \hat{B}(t)] = \frac{1}{Z} \text{Tr} [\hat{A}^\beta e^{i\hat{H}t/\hbar} \hat{B} e^{-i\hat{H}t/\hbar}], \quad (1)$$

or equivalently,^{20–26}

$$C_{AB}(t) = \frac{1}{Z} \text{Tr} [e^{-i\hat{H}t/\hbar} \hat{A}^\beta e^{i\hat{H}t/\hbar} \hat{B}] = \frac{1}{Z} \text{Tr} [\hat{A}^\beta(t) \hat{B}(0)], \quad (2)$$

where $\hat{A}_{\text{std}}^\beta = e^{-\beta\hat{H}} \hat{A}$ for the standard version of the correlation function, $\hat{A}_{\text{sym}}^\beta = e^{-\beta\hat{H}/2} \hat{A} e^{-\beta\hat{H}/2}$ for the symmetrized version,¹¹ or $\hat{A}_{\text{Kubo}}^\beta = \frac{1}{\beta} \int_0^\beta d\lambda e^{-(\beta-\lambda)\hat{H}} \hat{A} e^{-\lambda\hat{H}}$ for the Kubo-transformed version.²⁷ These three versions are related to one another by the following identities between their Fourier transforms:

$$\frac{\beta\hbar\omega}{1 - e^{-\beta\hbar\omega}} I_{AB}^{\text{Kubo}}(\omega) = I_{AB}^{\text{std}}(\omega) = e^{\beta\hbar\omega/2} I_{AB}^{\text{sym}}(\omega), \quad (3)$$

where $I_{AB}(\omega) = \int_{-\infty}^{\infty} dt e^{-i\omega t} C_{AB}(t)$. Here, $Z = \text{Tr}[e^{-\beta\hat{H}}]$ ($\beta = 1/k_B T$) is the partition function, \hat{H} the (time-independent) Hamiltonian of the system with the total number of degrees of freedom $3N$, which we assume to be of standard Cartesian form

$$\hat{H} = \frac{1}{2} \hat{\mathbf{p}}^T \mathbf{M}^{-1} \hat{\mathbf{p}} + V(\hat{\mathbf{x}}), \quad (4)$$

where \mathbf{M} is the diagonal “mass matrix” with elements $\{m_j\}$ and $\hat{\mathbf{p}}$ and $\hat{\mathbf{x}}$ are the momentum and coordinate operators, respectively;

and \hat{A} and \hat{B} are operators relevant to the specific property of interest, which are often nonlinear functions of coordinate or momentum operators.

A consistent practical quantum dynamics method for studying quantum thermal correlation functions of complex (large) molecular systems should have the two important properties:

1. Treat both linear and nonlinear operators (i.e., linear and nonlinear functions of the position or momentum operators) equally well and recover exact correlation functions in the classical $\hbar \rightarrow 0$, high temperature $\beta \rightarrow 0$, and harmonic limits.^{22,23,28–34}
2. Conserve the mapping quantum canonical density distribution, which corresponds to the quantum commutation relation

$$[e^{-\beta\hat{H}}, e^{-i\hat{H}t/\hbar}] = 0, \quad (5)$$

or equivalently,

$$e^{-i\hat{H}t/\hbar} e^{-\beta\hat{H}} e^{i\hat{H}t/\hbar} \equiv \hat{\rho}(t) = \hat{\rho}(0). \quad (6)$$

In addition, such a trajectory-based method should not require any specific form of the potential energy surface (PES) and is then, in principle, feasible to be employed with *ab initio* calculations on the fly or *ab initio*-based force fields.

We have recently developed trajectory-based approaches in phase space formulations of quantum mechanics^{20–26} for estimating the quantum thermal correlation function. The starting point is from phase space expressions of the quantum Liouville theorem for the “general density operator” in quantum mechanics,^{6,20,22,23}

$$\frac{\partial \hat{A}^\beta(t)}{\partial t} = -\frac{1}{i\hbar} [\hat{A}^\beta(t), \hat{H}]. \quad (7)$$

When $\hat{A} = 1$, the “general density operator” becomes the density operator $\hat{A}^\beta(t) = \hat{\rho}^{\text{eq}}(t) = e^{-i\hat{H}t/\hbar} e^{-\beta\hat{H}} e^{i\hat{H}t/\hbar}$. Because solving quantum Liouville dynamics in phase space is often difficult, if not impossible, for general multi-dimensional systems, we have proposed equilibrium continuity dynamics (ECD), a category of trajectory-based dynamics for solving quantum Liouville dynamics approximately but reasonably well.²⁵ In the ECD category, equilibrium Hamiltonian dynamics (EHD) is a special kind in which an effective Hamiltonian can be defined to govern the equations of motion. Both ECD and EHD satisfy the two important properties. The ECD or EHD equations of motion for the phase space variables (\mathbf{x}, \mathbf{p}) often involve the effective force as well as effective mass matrix, which can be computed on the fly by path integral,²⁴ Feynman–Kleinert approximation,^{12,22,35–38} thermal Gaussian approximation,^{22,23,39–42} etc. It is worth mentioning that, after we first utilized the quantum Liouville theorem in Wigner phase space to develop trajectory-based approaches for satisfying the two important properties,^{20,23} more trajectory-based methods along this line (in addition to our progress) have been developed.^{43,44}

The purpose of this paper is to develop a machine learning approach for fitting the effective force and effective mass matrix in advance such that the equations of motion can be integrated much more efficiently. In the field of fitting a PES from first-principles quantum chemistry data, the Deep Potential scheme has been

verified to be competent in constructing accurate and robust PES models for a wide variety of systems in an end-to-end way.^{45–47} This was achieved by the smooth symmetry-preserving feature descriptor, as well as the adaptive data generating scheme Deep Potential Generator.⁴⁸ As a result, Deep Potential-based molecular dynamics (DeepMD) can preserve the accuracy of *ab initio* molecular dynamics while reducing the cost by several orders of magnitude. This paper shows that the successful strategies of Deep Potential scheme can be further extended in a nontrivial way to fitting the effective force and effective mass matrix in phase space quantum dynamics.

Section II first introduces ECD and shows that ECD involves the zeroth order term of an exact series expansion of the phase space propagator. After expressing the effective force and effective mass matrix by path integral, we then introduce the machine learning scheme for fitting these quantities necessary for ECD by further development of the strategies introduced in Deep Potential. Section III applies the machine learning ECD methods to vibrational spectra for hydrogen peroxide (H_2O_2) and its isotope molecules, as well as the rovibrational spectrum of the H_2O molecule. The former involve a type of challenging molecular systems that have large-amplitude motion as well as tunneling splitting effects. The latter tests the performance for describing the vibrational motions and free rotational motions that are coupled with one another. Conclusions and outlook follow in Sec. IV.

II. THEORY

A. The correlation function in the phase space formulation of quantum mechanics

The expression of Eq. (2) in the phase space formulation of quantum mechanics can be derived^{20–22} as

$$C_{AB}(t) = \frac{1}{Z} \frac{1}{(2\pi\hbar)^{3N}} \int d\mathbf{x} \int d\mathbf{p} A^\beta(\mathbf{x}, \mathbf{p}; t) \tilde{B}(\mathbf{x}, \mathbf{p}; 0). \quad (8)$$

Here, the “general” phase space distribution $A^\beta(\mathbf{x}, \mathbf{p}; t)$ and the dual function $\tilde{B}(\mathbf{x}, \mathbf{p}; 0)$ can be expressed in a more generic framework for the one-to-one correspondence mapping in quantum mechanics,⁴⁹ which we show below to reformulate the unified classification scheme for the expression of the quantum thermal correlation function presented in Ref. 21 as a generalization from Ref. 50,

$$A^\beta(\mathbf{x}, \mathbf{p}; t) = \text{Tr}[\hat{A}^\beta(t)\hat{K}(\mathbf{x}, \mathbf{p})] \quad (9)$$

and

$$\tilde{B}(\mathbf{x}, \mathbf{p}; t) = \text{Tr}[\hat{K}^{-1}(\mathbf{x}, \mathbf{p})\hat{B}(t)]. \quad (10)$$

Or reversely,

$$\begin{aligned} \hat{A}^\beta(t) &= \frac{1}{(2\pi\hbar)^{3N}} \int d\mathbf{x} \int d\mathbf{p} A^\beta(\mathbf{x}, \mathbf{p}; t) \hat{K}^{-1}(\mathbf{x}, \mathbf{p}), \\ \hat{B}(t) &= \frac{1}{(2\pi\hbar)^{3N}} \int d\mathbf{x} \int d\mathbf{p} \tilde{B}(\mathbf{x}, \mathbf{p}; t) \hat{K}(\mathbf{x}, \mathbf{p}). \end{aligned} \quad (11)$$

The mapping kernel and the corresponding inverse satisfy the normalization,

$$\text{Tr}[\hat{K}(\mathbf{x}, \mathbf{p})] = \text{Tr}[\hat{K}^{-1}(\mathbf{x}, \mathbf{p})] = 1 \quad (12)$$

and

$$\frac{1}{(2\pi\hbar)^{3N}} \int d\mathbf{x} \int d\mathbf{p} \hat{K}(\mathbf{x}, \mathbf{p}) = \frac{1}{(2\pi\hbar)^{3N}} \int d\mathbf{x} \int d\mathbf{p} \hat{K}^{-1}(\mathbf{x}, \mathbf{p}) = \hat{I}. \quad (13)$$

Here, \hat{I} is the identity operator. The kernel of mapping is

$$\hat{K}(\mathbf{x}, \mathbf{p}) = \left(\frac{\hbar}{2\pi}\right)^{3N} \int d\zeta \int d\boldsymbol{\eta} e^{i\zeta^T(\hat{\mathbf{x}}-\mathbf{x}) + i\boldsymbol{\eta}^T(\hat{\mathbf{p}}-\mathbf{p})} f(\zeta, \boldsymbol{\eta}) \quad (14)$$

and

$$\hat{K}^{-1}(\mathbf{x}, \mathbf{p}) = \left(\frac{\hbar}{2\pi}\right)^{3N} \int d\zeta \int d\boldsymbol{\eta} e^{-i\zeta^T(\hat{\mathbf{x}}-\mathbf{x}) - i\boldsymbol{\eta}^T(\hat{\mathbf{p}}-\mathbf{p})} \frac{1}{f(-\zeta, -\boldsymbol{\eta})}, \quad (15)$$

where $f(\zeta, \boldsymbol{\eta})$ is a scalar function that defines the corresponding quantum phase space.

For instance, the Wigner function^{51,52} has

$$f(\zeta, \boldsymbol{\eta}) = 1, \quad (16)$$

the “general density distribution” $A^\beta(\mathbf{x}, \mathbf{p}; t)$ in the Wigner phase space is

$$A_W^\beta(\mathbf{x}, \mathbf{p}; t) = \frac{1}{(2\pi\hbar)^{3N}} \int d\mathbf{y} \left\langle \mathbf{x} - \frac{\mathbf{y}}{2} \middle| \hat{A}^\beta(t) \middle| \mathbf{x} + \frac{\mathbf{y}}{2} \right\rangle e^{i\mathbf{y}^T \mathbf{p}/\hbar}, \quad (17)$$

and the corresponding dual function $\tilde{B}_W(\mathbf{x}, \mathbf{p}; t)$ becomes

$$\tilde{B}_W(\mathbf{x}, \mathbf{p}; t) = \int d\mathbf{y} \left\langle \mathbf{x} - \frac{\mathbf{y}}{2} \middle| \hat{B}(t) \middle| \mathbf{x} + \frac{\mathbf{y}}{2} \right\rangle e^{i\mathbf{y}^T \mathbf{p}/\hbar}. \quad (18)$$

Following Refs. 20, 22, and 23, one can show that the quantum Liouville equation [Eq. (7)] in the Wigner phase space representation is

$$\begin{aligned} \frac{\partial A_W^\beta(\mathbf{x}, \mathbf{p}; t)}{\partial t} &= \mathcal{L} A_W^\beta(\mathbf{x}, \mathbf{p}; t) \\ &= -\mathbf{p}^T \mathbf{M}^{-1} \frac{\partial A_W^\beta}{\partial \mathbf{x}} + V'(\mathbf{x}) \cdot \frac{\partial A_W^\beta}{\partial \mathbf{p}} \\ &\quad - \frac{\hbar^2}{24} V(\mathbf{x}) \left(\frac{\overleftarrow{\partial}}{\partial \mathbf{x}} \cdot \frac{\overrightarrow{\partial}}{\partial \mathbf{p}} \right)^3 A_W^\beta + \dots, \end{aligned} \quad (19)$$

or equivalently,

$$\begin{aligned} \frac{\partial A_W^\beta(\mathbf{x}, \mathbf{p}; t)}{\partial t} &= \mathcal{L} A_W^\beta(\mathbf{x}, \mathbf{p}; t) \\ &= \left[-\mathbf{p}^T \mathbf{M}^{-1} \frac{\partial}{\partial \mathbf{x}} + \frac{2}{\hbar} V(\mathbf{x}) \sin \left(\frac{\hbar}{2} \frac{\overleftarrow{\partial}}{\partial \mathbf{x}} \cdot \frac{\overrightarrow{\partial}}{\partial \mathbf{p}} \right) \right] A_W^\beta, \end{aligned} \quad (20)$$

which is a generalization of the Wigner–Moyal equation.^{51,53} The arrows above the gradient operators in Eq. (19) or Eq. (20) imply the directions in which they act.

More trivially, the Wigner–Moyal equation leads to

$$\begin{aligned} \frac{\partial \tilde{B}_W(\mathbf{x}, \mathbf{p}; t)}{\partial t} &= \mathcal{L}^* \tilde{B}_W(\mathbf{x}, \mathbf{p}; t) \\ &= \mathbf{p}^T \mathbf{M}^{-1} \frac{\partial \tilde{B}_W}{\partial \mathbf{x}} - V'(\mathbf{x}) \cdot \frac{\partial \tilde{B}_W}{\partial \mathbf{p}} \\ &\quad + \frac{\hbar^2}{24} V(\mathbf{x}) \left(\frac{\overleftarrow{\partial}}{\partial \mathbf{x}} \cdot \frac{\overrightarrow{\partial}}{\partial \mathbf{p}} \right)^3 \tilde{B}_W + \dots, \end{aligned} \quad (21)$$

or equivalently,

$$\begin{aligned} \frac{\partial \tilde{B}_W(\mathbf{x}, \mathbf{p}; t)}{\partial t} &= \mathcal{L}^* \tilde{B}_W(\mathbf{x}, \mathbf{p}; t) \\ &= \left[\mathbf{p}^T \mathbf{M}^{-1} \frac{\partial}{\partial \mathbf{x}} - \frac{2}{\hbar} V(\mathbf{x}) \sin \left(\frac{\hbar}{2} \frac{\overleftarrow{\partial}}{\partial \mathbf{x}} \cdot \frac{\overrightarrow{\partial}}{\partial \mathbf{p}} \right) \right] \tilde{B}_W. \end{aligned} \quad (22)$$

Here, \mathcal{L}^* and \mathcal{L} are adjoint operators. \mathcal{L} is anti-self-adjoint in Wigner phase space because

$$\mathcal{L}^* = -\mathcal{L}. \quad (23)$$

The propagators in phase space then have the relation

$$\exp(\mathcal{L}^* t) = \exp(-\mathcal{L} t). \quad (24)$$

The quantum thermal correlation function of the form Eq. (2) or Eq. (8) is then

$$C_{AB}(t) = \frac{1}{Z} \int d\mathbf{x} \int d\mathbf{p} (\exp(\mathcal{L} t) A^\beta(\mathbf{x}, \mathbf{p}; 0)) \tilde{B}(\mathbf{x}, \mathbf{p}; 0), \quad (25)$$

while that of Eq. (1) is

$$C_{AB}(t) = \frac{1}{Z} \int d\mathbf{x} \int d\mathbf{p} A^\beta(\mathbf{x}, \mathbf{p}; 0) \tilde{B}(\mathbf{x}, \mathbf{p}; t) \quad (26)$$

$$= \frac{1}{Z} \int d\mathbf{x} \int d\mathbf{p} A^\beta(\mathbf{x}, \mathbf{p}; 0) (\exp(\mathcal{L}^* t) \tilde{B}(\mathbf{x}, \mathbf{p}; 0)) \quad (27)$$

$$= \frac{1}{Z} \int d\mathbf{x} \int d\mathbf{p} A^\beta(\mathbf{x}, \mathbf{p}; 0) (\exp(-\mathcal{L} t) \tilde{B}(\mathbf{x}, \mathbf{p}; 0)). \quad (28)$$

For convenience, one can define the function $f_{A^\beta}^W(\mathbf{x}, \mathbf{p}; t)$ as the ratio of the “general density distribution” to the density distribution,

$$f_{A^\beta}^W(\mathbf{x}, \mathbf{p}; t) = \frac{A_W^\beta(\mathbf{x}, \mathbf{p}; t)}{\rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; t)} = \frac{\int d\mathbf{y} \langle \mathbf{x} - \frac{\mathbf{y}}{2} | \hat{A}^\beta(t) | \mathbf{x} + \frac{\mathbf{y}}{2} \rangle e^{i\mathbf{y}^T \mathbf{p}/\hbar}}{\int d\mathbf{y} \langle \mathbf{x} - \frac{\mathbf{y}}{2} | e^{-\beta \hat{H}} | \mathbf{x} + \frac{\mathbf{y}}{2} \rangle e^{i\mathbf{y}^T \mathbf{p}/\hbar}}. \quad (29)$$

Then, the quantum thermal correlation function Eq. (8) in the Wigner representation becomes

$$C_{AB}(t) = \frac{1}{Z} \int d\mathbf{x} \int d\mathbf{p} \rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; 0) f_{A^\beta}^W(\mathbf{x}, \mathbf{p}; t) \tilde{B}_W(\mathbf{x}, \mathbf{p}; 0). \quad (30)$$

Similarly, one can express the (Wigner) phase space representation of Eq. (1) as

$$C_{AB}(t) = \frac{1}{Z} \int d\mathbf{x} \int d\mathbf{p} \rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; 0) f_{A^\beta}^W(\mathbf{x}, \mathbf{p}; 0) \tilde{B}_W(\mathbf{x}, \mathbf{p}; t). \quad (31)$$

Note that the phase space formulation of Eq. (6) in the second criterion reads

$$\frac{\partial \rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; t)}{\partial t} = 0. \quad (32)$$

Here, $\rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; t)/Z$ is the density distribution function, not the probability distribution function, because the uncertainty principle in quantum mechanics prevents the definition of the probability at a given point in phase space.

B. Evaluation of the correlation function with trajectory-based dynamics in quantum phase space

1. The equilibrium continuity dynamics correlation function in quantum phase space

The ECD formulation of the quantum thermal correlation function as first proposed in Ref. 20 and later in Ref. 23 suggests

$$C_{AB}(t) \approx \frac{1}{Z} \int d\mathbf{x}_0 \int d\mathbf{p}_0 \rho_W^{\text{eq}}(\mathbf{x}_0, \mathbf{p}_0; 0) f_{A^\beta}^W(\mathbf{x}_0, \mathbf{p}_0; 0) \tilde{B}_W(\mathbf{x}_t, \mathbf{p}_t; 0), \quad (33)$$

where the equations of motion for trajectory $(\mathbf{x}_t, \mathbf{p}_t)$ are

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{M}^{-1} \mathbf{p}, \\ \dot{\mathbf{p}} = -\frac{\partial V_{\text{eff}}(\mathbf{x}, \mathbf{p})}{\partial \mathbf{x}}. \end{cases} \quad (34)$$

with the effective force $-\partial V_{\text{eff}}(\mathbf{x}, \mathbf{p})/\partial \mathbf{x}$ obtained from

$$\mathbf{p}^T \mathbf{M}^{-1} \frac{\partial \rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; 0)}{\partial \mathbf{x}} = \frac{\partial}{\partial \mathbf{p}} \cdot \left(\rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; 0) \frac{\partial V_{\text{eff}}(\mathbf{x}, \mathbf{p})}{\partial \mathbf{x}} \right) \quad (35)$$

in Wigner phase space. Note that in many cases, there exist various choices for the effective force satisfying Eq. (35). (Appendix A offers more comprehensive discussions.)

It is easy to verify that Eq. (34) with Eq. (35) leads to

$$\mathcal{L}_{\text{con}} \rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; 0) = \left[-\frac{\partial}{\partial \mathbf{x}} \cdot (\dot{\mathbf{x}}) - \frac{\partial}{\partial \mathbf{p}} \cdot (\dot{\mathbf{p}}) \right] \rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; 0) = 0, \quad (36)$$

which states the stationarity of the density distribution in the continuity equation,

$$\frac{\partial \rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; t)}{\partial t} = \mathcal{L}_{\text{con}} \rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; t) = 0. \quad (37)$$

Using the operator adjoint to \mathcal{L}_{con} ,

$$\mathcal{L}_{\text{con}}^* = \dot{\mathbf{x}} \cdot \frac{\partial}{\partial \mathbf{x}} + \dot{\mathbf{p}} \cdot \frac{\partial}{\partial \mathbf{p}}, \quad (38)$$

we recast Eq. (33) into

$$\begin{aligned} C_{AB}(t) &\approx \frac{1}{Z} \int d\mathbf{x}_0 \int d\mathbf{p}_0 \rho_W^{\text{eq}}(\mathbf{x}_0, \mathbf{p}_0; 0) f_{A^\beta}^W(\mathbf{x}_0, \mathbf{p}_0; 0) \\ &\times [\exp(\mathcal{L}_{\text{con}}^* t) \tilde{B}_W(\mathbf{x}_0, \mathbf{p}_0; 0)] \\ &= \frac{1}{Z} \int d\mathbf{x}_0 \int d\mathbf{p}_0 \rho_W^{\text{eq}}(\mathbf{x}_0, \mathbf{p}_0; 0) f_{A^\beta}^W(\mathbf{x}_0, \mathbf{p}_0; 0) \tilde{B}_W(\mathbf{x}_t, \mathbf{p}_t; 0). \end{aligned} \quad (39)$$

In comparison, the exact formulation of the quantum thermal correlation function Eq. (31) can be expressed as

$$\begin{aligned} C_{AB}(t) &= \frac{1}{Z} \int d\mathbf{x}_0 \int d\mathbf{p}_0 \rho_W^{\text{eq}}(\mathbf{x}_0, \mathbf{p}_0; 0) f_{A^\beta}^W(\mathbf{x}_0, \mathbf{p}_0; 0) \\ &\times [\exp(\mathcal{Y}_{\text{mod}} t) \tilde{B}_W(\mathbf{x}_t, \mathbf{p}_t; 0)], \end{aligned} \quad (40)$$

with the modification propagator $\exp(\mathcal{Y}_{\text{mod}} t)$ defined by

$$\exp(\mathcal{Y}_{\text{mod}} t) \exp(\mathcal{L}_{\text{con}}^* t) = \exp(\mathcal{L}^* t). \quad (41)$$

In a harmonic system, because only first derivatives exist in the Liouville operator \mathcal{L}^* in Eq. (21), it is evident that $\mathcal{L}_{\text{con}}^* = \mathcal{L}^*$ produces exact dynamics. That is, the modification propagator $\exp(\mathcal{Y}_{\text{mod}} t) = 1$ in the harmonic limit. Therefore, it is not unreasonable to make the simplest approximation

$$\exp(\mathcal{Y}_{\text{mod}} t) \approx 1 \quad (42)$$

for general systems, as done in ECD.

It is obvious that the zeroth order propagator $U_0(t) = \exp(\mathcal{L}_{\text{con}}^* t)$ instead of the exact propagator $U(t) = \exp(\mathcal{L}^* t) = \exp(\mathcal{Y}_{\text{mod}} t) \exp(\mathcal{L}_{\text{con}}^* t)$ is used in Eq. (33) or Eq. (39). The strategy of Ref. 54 can be implemented to obtain the exact propagator from the zeroth order one. Propose a correction operator

$$\mathcal{C}(t) = (\mathcal{L}^* - \mathcal{L}_{\text{con}}^*) U_0(t). \quad (43)$$

It is then straightforward to prove

$$U(t) - U_0(t) = \int_0^t dt' U(t-t') \mathcal{C}(t') \quad (44)$$

and that the exact propagator can be represented as

$$U(t) = \sum_{j=0}^{\infty} U_j(t), \quad (45)$$

with the recursion relation

$$U_{j+1}(t) = \int_0^t dt' U_j(t-t') \mathcal{C}(t'). \quad (46)$$

A rigorous formulation of a series of propagators, where ECD involves only the zeroth order propagator, is then constructed for general systems. Even though it is often computationally demanding to obtain the exact quantum thermal correlation function through Eq. (45), it is useful to employ low order correction terms to estimate the accuracy of Eq. (A44) for a specified system. That is, the influence of the first order correction $U_1(t)$ indicates the performance of ECD even when no exact results are available. Provided that ECD is reasonably good, low order correction terms of Eq. (45) can be constructed to improve over ECD in a systematic and rigorous way, as long as the computation effort is affordable. (See Appendix B for example.)

It is not difficult to verify that in Eq. (19), the second order term ($\sim \hbar^2$) and higher order terms of the RHS vanish in the classical ($\hbar \rightarrow 0$), high temperature ($\beta \rightarrow 0$), and harmonic limits. In these limits, \mathcal{L} becomes the same as \mathcal{L}_{con} . That is, the ECD formulation of Eq. (33) with the equations of motion [Eqs. (34) and (35)] is exact in the classical ($\hbar \rightarrow 0$), high temperature ($\beta \rightarrow 0$), and harmonic limits. In the three limits, ECD approaches the linearized semiclassical initial value representation (LSC-IVR) of Miller *et al.*^{12,13,38,55–58} Because ECD conserves the canonical distribution [i.e., Eq. (32) or

Eq. (37)], it offers a framework to develop trajectory-based dynamics methods that have the two important properties listed in Sec. I. In Ref. 23, it is shown that there exists EHD—a special type of ECD in which the effective Hamiltonian can be identified to generate the equations of motion after transforming the phase variables.

2. Local Gaussian approximation

The calculation of $\tilde{\rho}_W(\mathbf{x}, \mathbf{p}; t = 0)$ for operator \hat{B} in Eq. (18) is usually straightforward; in fact, \hat{B} is often a function only of coordinates or only of momenta, in which case its Wigner function is simply the classical function itself. Calculating $A^\beta(\mathbf{x}, \mathbf{p}; t = 0)$ for operator \hat{A}^β in Eq. (17), however, involves the Boltzmann operator with the total Hamiltonian of the complete system so that carrying out the multidimensional Fourier transform to obtain it is non-trivial. The quantum canonical distribution in the Wigner phase space, i.e., Wigner density distribution function or Wigner function of the Boltzmann operator, can often be approximated by the local Gaussian approximation (LGA) of the momentum distribution,^{12,59}

$$\begin{aligned}\rho_W^{\text{eq,LGA}}(\mathbf{x}, \mathbf{p}) &= \langle \mathbf{x} | e^{-\beta \hat{H}} | \mathbf{x} \rangle \left(\frac{\beta}{2\pi} \right)^{3N/2} \det(\mathbf{M}_{\text{therm}}(\mathbf{x}))^{-1/2} \\ &\quad \times \exp \left[-\frac{\beta}{2} \mathbf{p}^T \mathbf{M}_{\text{therm}}^{-1}(\mathbf{x}) \mathbf{p} \right] \\ &= \left(\frac{\beta}{2\pi} \right)^{3N/2} \det(\mathbf{M}_{\text{therm}}(\mathbf{x}))^{-1/2} \\ &\quad \times \exp \left[-\beta \tilde{V}_{\text{eff}}(\mathbf{x}) - \frac{\beta}{2} \mathbf{p}^T \mathbf{M}_{\text{therm}}^{-1}(\mathbf{x}) \mathbf{p} \right].\end{aligned}\quad (47)$$

Here, the effective potential is defined as

$$\tilde{V}_{\text{eff}}(\mathbf{x}) = -\frac{1}{\beta} \ln \langle \mathbf{x} | e^{-\beta \hat{H}} | \mathbf{x} \rangle. \quad (48)$$

When the thermal mass matrix $\mathbf{M}_{\text{therm}}$ is constant in Eq. (47), one choice for the effective force defined by Eq. (35) is

$$\begin{aligned}-\frac{\partial V_{\text{eff}}(\mathbf{x}, \mathbf{p})}{\partial \mathbf{x}} &= \frac{1}{\beta} \mathbf{M}_{\text{therm}} \mathbf{M}^{-1} \frac{\partial \ln \langle \mathbf{x} | e^{-\beta \hat{H}} | \mathbf{x} \rangle}{\partial \mathbf{x}} \\ &= -\mathbf{M}_{\text{therm}} \mathbf{M}^{-1} \frac{\partial \tilde{V}_{\text{eff}}}{\partial \mathbf{x}}.\end{aligned}\quad (49)$$

$\partial \ln \langle \mathbf{x} | e^{-\beta \hat{H}} | \mathbf{x} \rangle / \partial \mathbf{x}$ can be evaluated by path integral on the fly as done in path integral Liouville dynamics (PILD).^{24–26} PILD is a specific option in the category of EHD or ECD as discussed in Ref. 60.

When the thermal mass matrix $\mathbf{M}_{\text{therm}}$ is strongly position-dependent, a reasonably good approximation is to employ the covariance matrix of momenta²⁴ to evaluate the inverse of $\mathbf{M}_{\text{therm}}$ locally, i.e.,

$$\mathbf{M}_{\text{therm}}^{-1} = \frac{1}{\beta \hbar^2} \frac{\int d\mathbf{y} \langle \mathbf{x} - \frac{\mathbf{y}}{2} | e^{-\beta \hat{H}} | \mathbf{x} + \frac{\mathbf{y}}{2} \rangle \mathbf{y} \mathbf{y}^T}{\int d\mathbf{y} \langle \mathbf{x} - \frac{\mathbf{y}}{2} | e^{-\beta \hat{H}} | \mathbf{x} + \frac{\mathbf{y}}{2} \rangle}, \quad (50)$$

which can be evaluated by open-chain path integral molecular dynamics (PIMD) or Monte Carlo (PIMC). Some other options that are even much more numerically demanding employ open-chain

PIMD or PIMC for directly performing the Fourier transform to obtain the Wigner function of the Boltzmann operator (they have to deal with the sign problem).^{61–64} In this paper, we consider only Eq. (50) that is possibly useful for general large molecular systems.

On the assumption that the Wigner density function is approximated by Eq. (47), (at least) two kinds of ECD equations of motion can be achieved. One is

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{M}^{-1} \mathbf{p}, \\ (\dot{\mathbf{p}})_l &= \sum_i (\mathbf{M}_{\text{therm}})_{li} m_i^{-1} \left[-\frac{\partial \tilde{V}_{\text{eff}}}{\partial x_i} - \frac{1}{2} \mathbf{p}^T \frac{\partial \mathbf{M}_{\text{therm}}^{-1}}{\partial x_i} \mathbf{p} \right. \\ &\quad \left. - \frac{1}{2\beta} \frac{\partial}{\partial x_i} \log \det(\mathbf{M}_{\text{therm}}(\mathbf{x})) \right] + \frac{1}{\beta} \sum_k m_k^{-1} \frac{\partial (\mathbf{M}_{\text{therm}})_{lk}}{\partial x_k}.\end{aligned}\quad (51)$$

The dynamics produced by Eq. (51) is denoted ECD-1, which has a conserved quantity

$$\begin{aligned}H_{\text{eff}}^{\text{ECD-1}} &= \tilde{V}_{\text{eff}} + \frac{1}{2} \mathbf{p}^T \mathbf{M}_{\text{therm}}^{-1} \mathbf{p} + \frac{1}{2\beta} \log \det(\mathbf{M}_{\text{therm}}) \\ &\quad - \frac{1}{\beta} \int_0^t \sum_{i,j,k} \frac{p_k}{m_i} (\mathbf{M}_{\text{therm}}^{-1})_{kj} \frac{\partial (\mathbf{M}_{\text{therm}})_{ji}}{\partial x_i} dt'\end{aligned}\quad (52)$$

that is the explicit expression of Eq. (A45) or Eq. (A48) in Appendix A. The conserved quantity of Eq. (52), however, does not lead to any form of Hamiltonian dynamics. The other is

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{M}^{-1} \mathbf{p}, \\ (\dot{\mathbf{p}})_l &= \sum_i (\mathbf{M}_{\text{therm}})_{li} m_i^{-1} \left[-\frac{\partial \tilde{V}_{\text{eff}}}{\partial x_i} + \frac{1}{2} \mathbf{p}^T \frac{\partial \mathbf{M}_{\text{therm}}^{-1}}{\partial x_i} \mathbf{p} \right. \\ &\quad \left. + \frac{1}{2\beta} \frac{\partial}{\partial x_i} \log \det(\mathbf{M}_{\text{therm}}(\mathbf{x})) \right] \\ &\quad + \sum_{i,j,k} p_k m_k^{-1} \frac{\partial (\mathbf{M}_{\text{therm}})_{li}}{\partial x_k} (\mathbf{M}_{\text{therm}}^{-1})_{ij} p_j.\end{aligned}\quad (53)$$

The dynamics yielded by Eq. (53) is denoted ECD-2 or EHD, which has a conserved property,

$$H_{\text{eff}}^{\text{ECD-2}} = \tilde{V}_{\text{eff}} + \frac{1}{2} \mathbf{p}^T \mathbf{M}_{\text{therm}}^{-1} \mathbf{p} - \frac{1}{2\beta} \log \det(\mathbf{M}_{\text{therm}}). \quad (54)$$

Note that Eq. (53) of ECD-2 (or EHD) becomes Hamilton's equations of motion generated by effective Hamiltonian defined in Eq. (54), under the transformation of phase space variables,²³

$$\begin{cases} \mathbf{x}_{\text{eff}} = \mathbf{x}, \\ \mathbf{p}_{\text{eff}} = \mathbf{M} \mathbf{M}_{\text{therm}}^{-1} \mathbf{p}. \end{cases} \quad (55)$$

It is straightforward to verify that ECD with either option above still rigorously satisfies Eq. (36) and produces the exact correlation function (even of nonlinear operators) in the classical, high-temperature, and harmonic limits. Note that the thermal mass matrix $\mathbf{M}_{\text{therm}}$ defined by the quantum canonical distribution or Wigner density distribution $\rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p})$ is often different from the (diagonal) physical mass matrix \mathbf{M} except in the high temperature or classical limit.

3. Machine learning implementation

As demonstrated in Eqs. (51) and (53), the effective force in the coordinate space $-\partial\tilde{V}_{\text{eff}}/\partial\mathbf{x}$, thermal mass matrix $\mathbf{M}_{\text{therm}}$, and its derivative $\partial\mathbf{M}_{\text{therm}}/\partial\mathbf{x}$ are required for ECD. Estimation of such quantities by using PIMD/PIMC at each time step of ECD is expensive. Modeling them with machine learning should greatly improve the computation efficiency.

Although employing machine learning to fit the PES from electronic structure data is an old idea,^{65–72} efficiently fitting the important elements in the ECD equations of motion is not trivial. In this paper, we adopt and extend the Deep Potential scheme,⁴⁷ which represents the inter-atomic PES as a sum of atomic contributions. The atomic contribution, being a function of the coordinates of each atom and its near neighbors, is represented by deep neural networks that ensures the symmetry preserving property of the PES. To date, Deep Potential has been used to model various phenomena in physical chemistry^{73–78} and materials sciences.^{79–84} In particular, it was later extended to representing vectors like the positions of Wannier centers,⁷⁷ as well as tensor properties like the polarizability,⁷³ at the level of *ab initio* electronic structure theory.

Below, we introduce the strategy to develop machine learning ECD (ML-ECD). We will see that, although the Deep Potential construction can be directly used for the effective potential, the construction of the thermal mass matrix, being a $3N \times 3N$ symmetric matrix, requests entirely non-trivial extension.

The effective potential \tilde{V}_{eff} defined in Eq. (48) satisfies

1. Translational invariance: For any vector $\mathbf{t} \in \mathbb{R}^3$, define the translation vector $\tilde{\mathbf{t}} = (\mathbf{t}^T \dots \mathbf{t}^T)^T$, then $\tilde{V}_{\text{eff}}(\mathbf{x} + \tilde{\mathbf{t}}) = \tilde{V}_{\text{eff}}(\mathbf{x})$.
2. Rotational invariance: For any orthogonal matrix $\mathbf{R} \in \mathbb{R}^{3 \times 3}$ satisfying $\mathbf{R}^T \mathbf{R} = \mathbf{R} \mathbf{R}^T = \mathbf{1}_{3 \times 3}$, define the rotation matrix $\tilde{\mathbf{R}} = \text{diag}\{\mathbf{R}, \dots, \mathbf{R}\}$, then $\tilde{V}_{\text{eff}}(\tilde{\mathbf{R}}\mathbf{x}) = \tilde{V}_{\text{eff}}(\mathbf{x})$.
3. Permutational invariance: For any pair of atom i and atom j of the same type, i.e., $\alpha_i = \alpha_j$, define matrix $\mathbf{P} \in \mathbb{R}^{N \times N}$, whose elements are $P_{kl} = \delta_{kl} - (\delta_{ki} - \delta_{kj})(\delta_{li} - \delta_{lj})$. Define permutation matrix $\tilde{\mathbf{P}} = \mathbf{P} \otimes \mathbf{1}_{3 \times 3}$, then $\tilde{V}_{\text{eff}}(\tilde{\mathbf{P}}\mathbf{x}) = \tilde{V}_{\text{eff}}(\mathbf{x})$.

These requirements are met by the Deep Potential scheme.⁴⁷ In the construction of the Deep Potential descriptor, the feature of the environment of atom i is represented by

$$D_i^{lm} = \sum_{j,k \in \mathcal{N}_i} G_{\alpha_i \alpha_j}^l(s(r_{ij})) G_{\alpha_i \alpha_k}^m(s(r_{ik})) \mathbf{r}_{ij}^T \mathbf{r}_{ik}, \quad l = 1, \dots, M_1, \quad m = 1, \dots, M_2. \quad (56)$$

\mathcal{N}_i is the set of neighbors for atom i . The local embedding function \mathbf{G} is a neural network (NN) function, which takes one input and produces M_1 outputs and whose parameters are dependent on the types of atom i and atom j . In Eq. (56), the translational invariance is satisfied by the relative position (\mathbf{r}_{ij}) , the rotational invariance is met by the inner product $(\mathbf{r}_{ij}^T \mathbf{r}_{ik})$, and the permutational invariance is fulfilled by the summation over all neighbor atoms. The function $s(r)$ in Eq. (56) vanishes smoothly as r increases

$$s(r) = \begin{cases} \frac{1}{r}, & r < r_{cs}, \\ \frac{1}{r} \left[\frac{1}{2} \cos\left(\pi \frac{r - r_{cs}}{r_c - r_{cs}}\right) + \frac{1}{2} \right], & r_{cs} < r < r_c, \\ 0, & r > r_c, \end{cases} \quad (57)$$

where r_c and r_{cs} are the cutoff radius and smooth cutoff radius, respectively.

We express the feature of the environment of atom i [Eq. (56)] as a vector \mathbf{D}_i of length $M_1 M_2$ and then define the local energy of atom i as the output of an NN function f that employs \mathbf{D}_i as the input. Finally, the effective potential is constructed as the sum of all local energies,

$$\tilde{V}_{\text{eff}} = \sum_i \tilde{V}_i = \sum_i f(\mathbf{D}_i). \quad (58)$$

Different from the effective potential, the thermal mass matrix $\mathbf{M}_{\text{therm}}$ satisfies

1. Translational invariance: For any translation vector $\tilde{\mathbf{t}}$, $\mathbf{M}_{\text{therm}}(\mathbf{x} + \tilde{\mathbf{t}}) = \mathbf{M}_{\text{therm}}(\mathbf{x})$.
2. Rotational covariance: For any rotation matrix $\tilde{\mathbf{R}}$, $\mathbf{M}_{\text{therm}}(\tilde{\mathbf{R}}\mathbf{x}) = \tilde{\mathbf{R}}\mathbf{M}_{\text{therm}}(\mathbf{x})\tilde{\mathbf{R}}^T$.
3. Permutational covariance: For any permutation matrix $\tilde{\mathbf{P}}$, $\mathbf{M}_{\text{therm}}(\tilde{\mathbf{P}}\mathbf{x}) = \tilde{\mathbf{P}}\mathbf{M}_{\text{therm}}(\mathbf{x})\tilde{\mathbf{P}}^T$.

A model for constructing the thermal mass meeting all the requirements above is to express $\mathbf{M}_{\text{therm}}$ in the block form $\{\mathbf{M}_{ij}, i = 1, \dots, N; j = 1, \dots, N\}$. Similar to Eq. (56), we introduce the feature descriptor in the tensor form

$$\overleftrightarrow{D}_i^{lm} = \frac{1}{2} \sum_{j,k} G_{\alpha_i \alpha_j}^l(s(r_{ij})) G_{\alpha_i \alpha_k}^m(s(r_{ik})) (\mathbf{r}_{ij} \mathbf{r}_{ik}^T + \mathbf{r}_{ik} \mathbf{r}_{ij}^T), \quad l = 1, \dots, M_1, \quad m = 1, \dots, M_2, \quad (59)$$

which replaces the inner product in Eq. (56) with the outer product and outputs a symmetric 3×3 matrix for each l and m . It is straightforward to prove that the rotational covariance is satisfied by the outer product. Then, the tensor feature of the environment of atom i , $\overleftrightarrow{\mathbf{D}}_i$, can be designed as a $(M_1 M_2)$ -dimensional vector of 3×3 matrices. It is easy to see that the scalar feature [Eq. (56)] is simply the trace of the tensor feature [Eq. (59)], i.e., $\text{Tr } \overleftrightarrow{D}_i^{lm} = D_i^{lm}$.

Equation (56) or Eq. (59), however, does not use the information of the angle between \mathbf{r}_{ij} and \mathbf{r}_{ik} , which may have a considerable effect on the tensor descriptor [Eq. (59)]. More importantly, Eq. (56) or Eq. (59) fails to describe high symmetry points in molecular systems. (See Appendix C for more discussions.) A more general form of the Deep Potential scalar descriptor reads

$$D_i^l = \sum_{j,k} G_{\alpha_i \alpha_j \alpha_k}^l(s(r_{ij}), s(r_{ik}), \cos \theta_{ijk}) \mathbf{r}_{ij}^T \mathbf{r}_{ik}, \quad l = 1, \dots, F. \quad (60)$$

The local embedding function \mathbf{G} is an NN function, which takes three inputs $[s(r_{ij}), s(r_{ik}), \cos \theta_{ijk} \equiv \mathbf{r}_{ij}^T \mathbf{r}_{ik}/(r_{ij} r_{ik})]$ and yields F outputs and whose parameters are dependent on the types of atom i , atom j , and atom k .

The corresponding Deep Potential descriptor in the tensor form is

$$\overleftrightarrow{D}_i^l = \frac{1}{2} \sum_{j,k} G_{\alpha_i \alpha_j \alpha_k}^l(s(r_{ij}), s(r_{ik}), \cos \theta_{ijk}) (\mathbf{r}_{ij} \mathbf{r}_{ik}^T + \mathbf{r}_{ik} \mathbf{r}_{ij}^T), \quad l = 1, \dots, F, \quad (61)$$

which is a symmetric 3×3 matrix for each l . Thus, the tensor feature of the environment of atom i , $\overleftrightarrow{\mathbf{D}}_i$, can be viewed as an F -dimensional

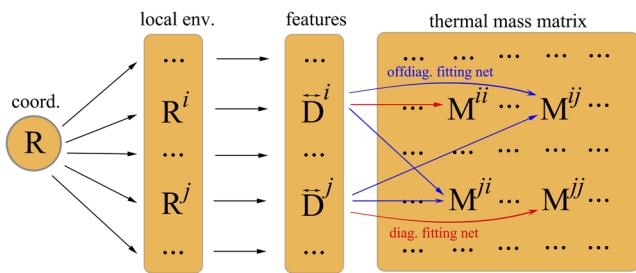


FIG. 1. Schematic representation of the neural networks for the thermal mass matrix.

vector of 3×3 matrices. Similarly, the scalar feature [Eq. (60)] is the trace of the tensor feature [Eq. (61)], i.e., $\text{Tr } \overset{\leftrightarrow}{D}_i^l = D_i^l$. In Appendix C, we illustrate that in some cases, Eq. (61) may be better to describe tensor properties than Eq. (59).

With the scalar and tensor forms of the feature descriptor designed by Eqs. (60) and (61), the diagonal blocks of the thermal mass (that are 3×3 matrices) can be constructed as

$$M_{ii} = \sum_{l=1}^F f^l(\mathbf{D}_i) \overset{\leftrightarrow}{D}_i^l \overset{\leftrightarrow}{D}_i^l, \quad (62)$$

where f is an NN function that takes $\mathbf{D}_i = \{D_i^l, l = 1, \dots, F\}$ as the input and outputs F values. The off-diagonal blocks can be designed by

$$M_{ij} = \sum_l g^l(\mathbf{D}_i s(r_{ij}), \mathbf{D}_j s(r_{ij}), s(r_{ij})) \overset{\leftrightarrow}{D}_i^l \overset{\leftrightarrow}{D}_j^l, \quad i \neq j, \quad (63)$$

where g is an NN function that takes $\{D_i^l s(r_{ij}), l = 1, \dots, F\}$, $\{D_j^l s(r_{ij}), l = 1, \dots, F\}$, and $s(r_{ij})$ as the input and generates F output values. In addition, $g(\mathbf{D}_1, \mathbf{D}_2, s)$ should be symmetric with respect to \mathbf{D}_1 and \mathbf{D}_2 such that $M_{ij} = M_{ji}^T$. Therefore, we can make $g(\mathbf{D}_1, \mathbf{D}_2, s) = \tilde{g}(\mathbf{D}_1, \mathbf{D}_2, s) + \tilde{g}(\mathbf{D}_2, \mathbf{D}_1, s)$, where \tilde{g} is a general NN function that employs $2F + 1$ inputs and yields F outputs. A schematic representation of the NNs for the thermal mass matrix is demonstrated in Fig. 1.

III. SIMULATION DETAILS AND RESULTS

Below, we implement ML-ECD to illustrate examples of realistic molecular systems.

A. Simulation details

$\text{H}_2^{16}\text{O}_2$ and its isotope molecules are often used as fundamental benchmark nonlinear and nonplanar systems. Hydrogen peroxide is one of the simplest molecules involving large-amplitude motion, e.g., the internal rotation (torsion) around the O–O bond. The molecule system presents a challenging test for both experimental and theoretical studies. The accurate six dimensional PES was developed by Koput, Carter, and Handy.¹⁰⁴ We use Tapenade⁸⁶ to generate the code for the (classical) force of the PES.

In applications of ML-ECD in this paper, PIMD is used for equilibrating the molecular system and for evaluating the effective

potential $\hat{V}_{\text{eff}}(\mathbf{x})$ as well as thermal mass matrix $\mathbf{M}_{\text{therm}}$ before ML-ECD is employed for real time dynamics. For each molecular system at 300 K, closed-chain PIMD with 256 PI beads is used to evaluate the effective force for 1000 typical samples. The time step is about 9–12 a.u. Data for each sample are averaged from three constrained PIMD trajectories of 1×10^6 a.u., which leads to a statistical error bar of the effective force $\sim 1\%$ of the standard deviation of the effective force. The covariance matrix is evaluated by open-chain PIMD with 128 beads for 1000 samples, each of which is averaged from three constrained open-chain PIMD trajectories of 6×10^7 a.u., which gives a statistical error bar of the covariance matrix, which is $\sim 1\%$ of the standard deviation of the covariance matrix. Both closed-chain and open-chain PIMD simulations employ the “middle” thermostat scheme,^{87–89} in which we use the Langevin equation^{85,90,91} with 0.017 a.u. as the friction coefficient for the first staging mode and $\omega_P = \sqrt{P}/\beta\hbar$ as the friction constants for other staging modes.^{87,88} For H_2O_2 at 100 K, closed-chain PIMD with 768 PI beads is used to evaluate the effective force, and the covariance matrix is evaluated by open-chain PIMD with 384 beads.

In the machine learning model of the effective potential $\hat{V}_{\text{eff}}(\mathbf{x})$, we employ two hidden layers (with 20 nodes for the first layer and 40 nodes for the second one) for the local embedding net. Take $M_1 = M_2 = 6$. Three hidden layers (with 20 nodes for each layer) are used for the fitting net. In the machine learning model of the thermal mass matrix $\mathbf{M}_{\text{therm}}$, we use two hidden layers (with 20 nodes for each layer) for the local embedding net. The number of encoded features is set to be $F = 20$. The fitting net for the thermal mass matrix involves three hidden layers (with 20 nodes for each layer). The thermal mass matrix is pre-processed before the training. A unitless matrix $\mathbf{M}^{-1/2} \mathbf{M}_{\text{therm}} \mathbf{M}^{-1/2} - \mathbf{I}$ is constructed such that the values of the matrix elements are closer to zero and better-scaled. In

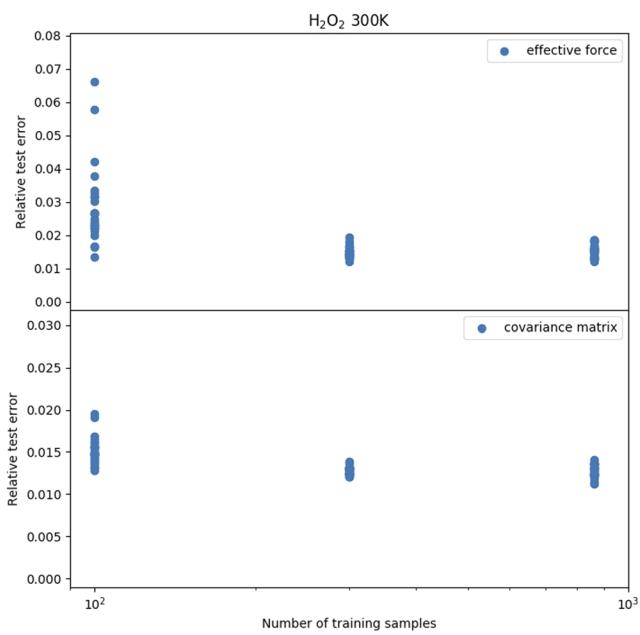


FIG. 2. Relative test error of the effective force or covariance matrix as a function of the number of training samples. The logarithmic scale is used in the x-axis.

the training process, the samples are randomly split into the training data, validation data, and test data with a ratio of 8:1:1. A converged model is obtained after training with 1000 epochs. For each converged model, the (largest) test error is below 2%, which is on the same scale with the statistical error of training data.

Figure 2 shows the convergence of relative test error (ratio of the test error to the standard deviation of test data) as the size of training set increases. H_2O_2 at 300 K is used as an example. For each training set, 24 models are trained, which are different only in initial parameters. As demonstrated in Fig. 2, irrespective of the number of samples (100, 300, or 864) used to train the models, the relative test error of the best model(s) is below 1.5%. While the deviation among models is relatively large when 100 samples are used, the performance with 300 training samples is similar to that with 864 training samples. When the size of the training set is 300 or more, the largest relative test error is expected to be smaller than 2%, which does not significantly decrease as the number of samples increases. In this paper, 1000 typical samples are employed, which are enough for the current type of machine learning models for the molecular systems under investigation. The time cost to integrate the ECD equations of motion per time step is reduced by 6–7 orders of magnitude in comparison to the brute-forcing way of calculating the effective force and thermal mass matrix [by employing the covariance matrix Eq. (50)] on the fly with PIMD.

In the real time dynamics simulation of $\text{H}_2^{16}\text{O}_2$ and its isotope molecules, 240 trajectories propagated to 131 072 a.u. (3.17 ps) are used to calculate the correlation function. The time step is 4 a.u. (0.096 fs). The translation and rotation are removed along the trajectory each 100 steps using the technique introduced in Ref. 92. The dipole moment μ and its time derivative $\dot{\mu}$ are obtained by the point charge model ($\mu = \sum_{i=1}^{N_{\text{atom}}} q_i \mathbf{x}_i$ and $\dot{\mu} = \sum_{i=1}^{N_{\text{atom}}} q_i \dot{\mathbf{x}}_i$, where q_i , \mathbf{x}_i , and $\dot{\mathbf{x}}_i$ represent the point charge, position, and velocity of the i -th atom, respectively). The Fourier transform of the Kubo-transformed collective dipole-derivative autocorrelation function^{14,25,26,93} is used for ML-ECD to calculate the vibrational IR spectrum.

We also apply ML-ECD to the rovibrational IR spectrum of the H_2^{16}O molecule at $T = 300$ K. The accurate PES from Ref. 94 and dipole moment surface (DMS) from Ref. 95 are used. Simulation details of PIMD and ML procedures are the same as those for H_2O_2 . In the real time dynamics simulation, while the time step size is 4 a.u. (0.096 fs) and each ECD trajectory is propagated up to 131 072 a.u. (3.17 ps), 240 such trajectories are used to calculate the dipole-derivative autocorrelation function, the Fourier transform of which leads to the rovibrational IR spectrum. The (translational) velocities of the center of mass of the H_2O molecule are removed along the trajectory each 100 steps. The accurate spectrum line data for 296 K of Ref. 96 were

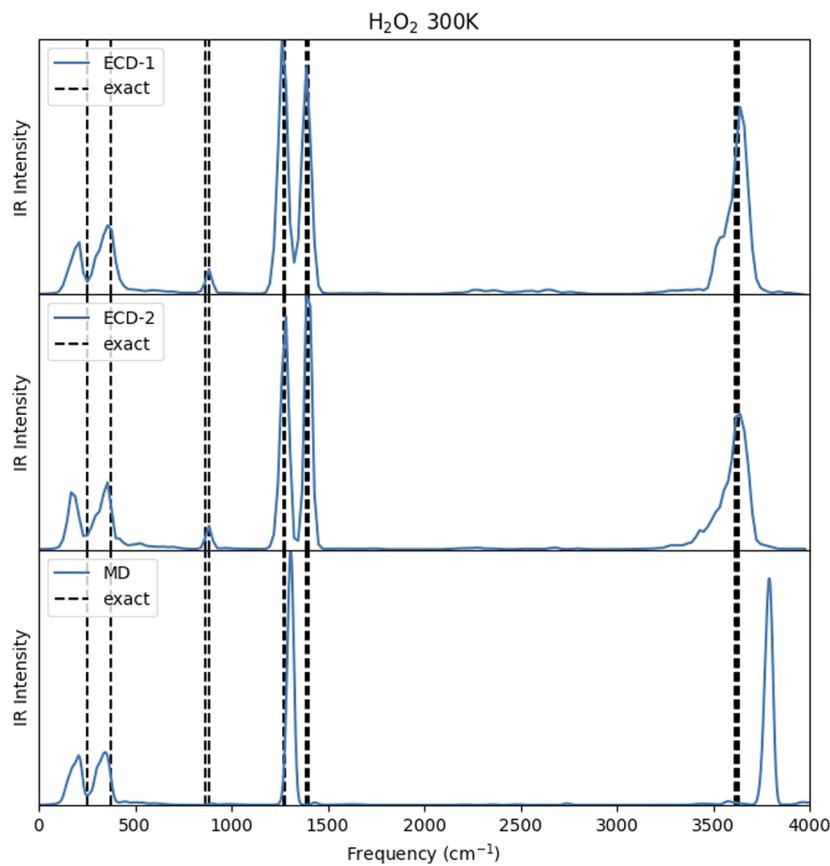


FIG. 3. The IR spectra for H_2O_2 at 300 K.

computed for the same PES⁹⁴ and DMS,⁹⁵ which are available on <http://www.tampa.phys.ucl.ac.uk/ftp/astrodata/water/BT2/> as well as included in the HITRAN database (<https://hitran.org/>). The exact results for 300 K are calculated with a modification of the Boltzmann weight from those benchmark line data for 296 K.

B. Results and discussions

We first study the vibrational IR spectrum of H₂O₂ at T = 300 K and that at T = 100 K (the results of which are demonstrated in Figs. 3 and 4). Spectra of H₂¹⁸O₂, HDO₂, and D₂O₂ are then simulated for T = 300 K for studying isotope effects (as presented in Figs. 5–7). ML-ECD, MD, and normal-mode analysis (NMA) are employed. NMA data are obtained by diagonalization of the mass-weighted Hessian matrix at the equilibrium configuration (i.e., the minimum of the PES). The simulation results are compared to the “exact” vibrational frequencies,⁹⁷ which were obtained using the fully symmetry-adapted finite-basis variational method described in Ref. 98.

As shown in Figs. 3–7 and Table I, ML-ECD methods are able to describe all peak positions in the vibrational spectrum reasonably well, overall faithfully capturing nuclear quantum effects when they are important and missed by MD and NMA. While MD and NMA

results are blue-shifted by ~170–190 cm⁻¹ for the high frequency O-H stretching mode(s) and by ~100 cm⁻¹ for the high frequency O-D stretching mode(s), ML-ECD results (obtained by such as ECD-1) are within the absolute error less than 20 cm⁻¹. ML-ECD produces a reasonably accurate peak position with a relatively small full width at half maximum such that two close peaks (the distance of which is no less than 40 cm⁻¹) are distinguished. For instance, ML-ECD methods produce two distinct peaks, the difference between which is about 70–130 cm⁻¹, in 1260–1405 cm⁻¹ for H₂O₂ (in Figs. 3 and 4) and H₂¹⁸O₂ (in Fig. 7), and in 945–1030 cm⁻¹ for D₂O₂ (in Fig. 6), as the “exact” results⁹⁷ suggest. In comparison, MD captures only one peak in the same region. The peak of the O-O stretching mode between 820 and 880 cm⁻¹ is noticeable (very small though) in the ML-ECD spectra but vanishes in the counterpart MD spectra, as demonstrated for H₂O₂ (Figs. 3 and 4), H₂¹⁸O₂ (Fig. 7), and D₂O₂ (Fig. 6).

The ML-ECD peak position (above 800 cm⁻¹) barely changes as the temperature is decreased from T = 300 K (Fig. 3) to T = 100 K (Fig. 4). This is in good agreement with the fact that the line position of the IR spectrum of the isolated molecule is independent of the change of temperature. Figures 3 and 5–7 indicate that isotopic effects are captured well by the ML-ECD results in comparison to those in the “exact” vibrational frequencies.⁹⁷ For this series of molecular systems, ECD-1 yields slightly more accurate results than ECD-2 does. The two variants of

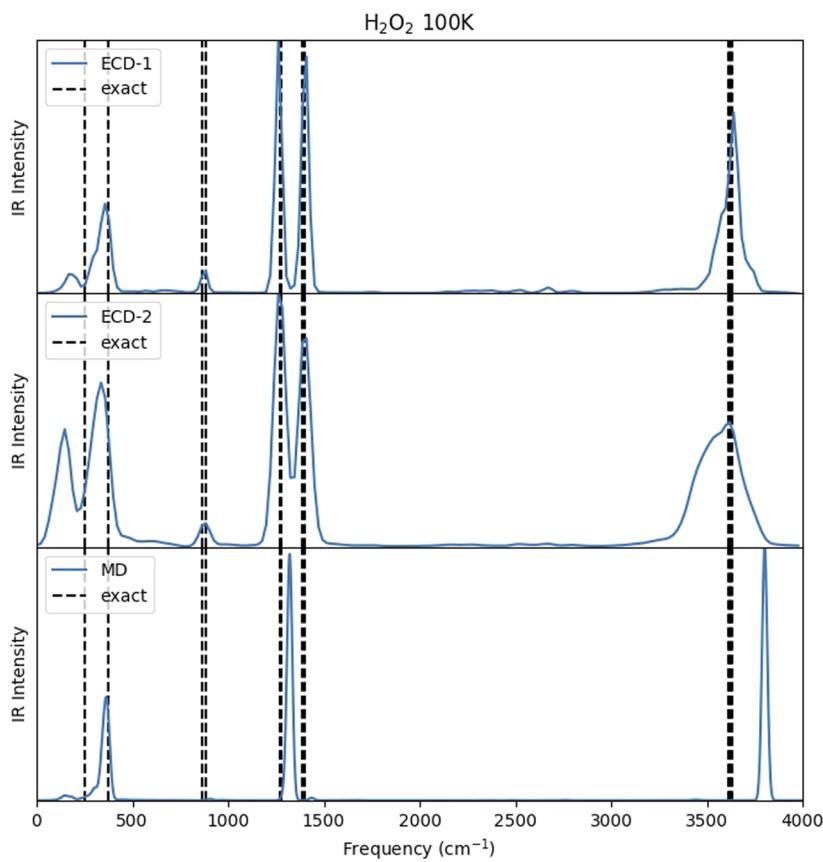


FIG. 4. The IR spectra for H₂O₂ at 100 K.

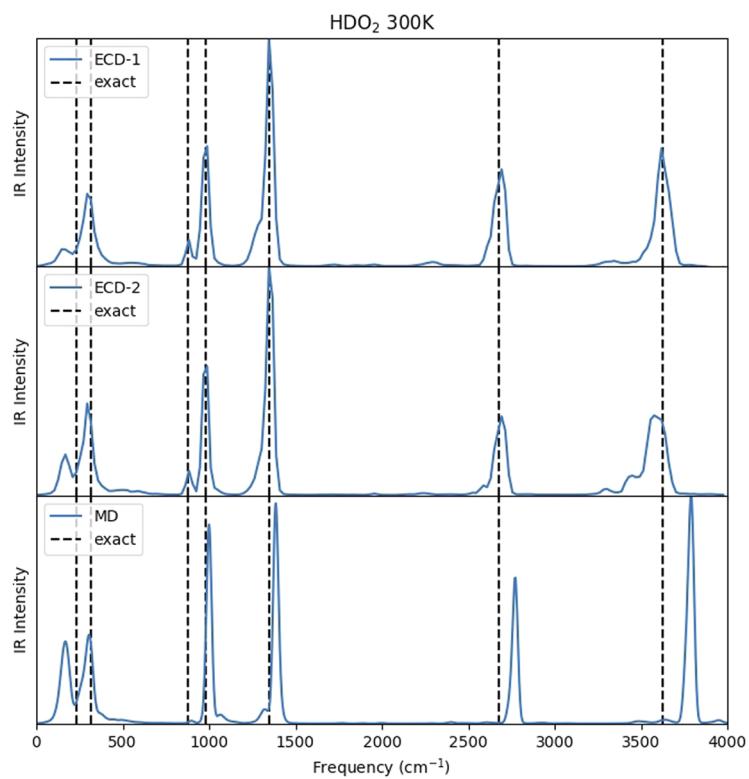


FIG. 5. The IR spectra for HDO_2 at 300 K.

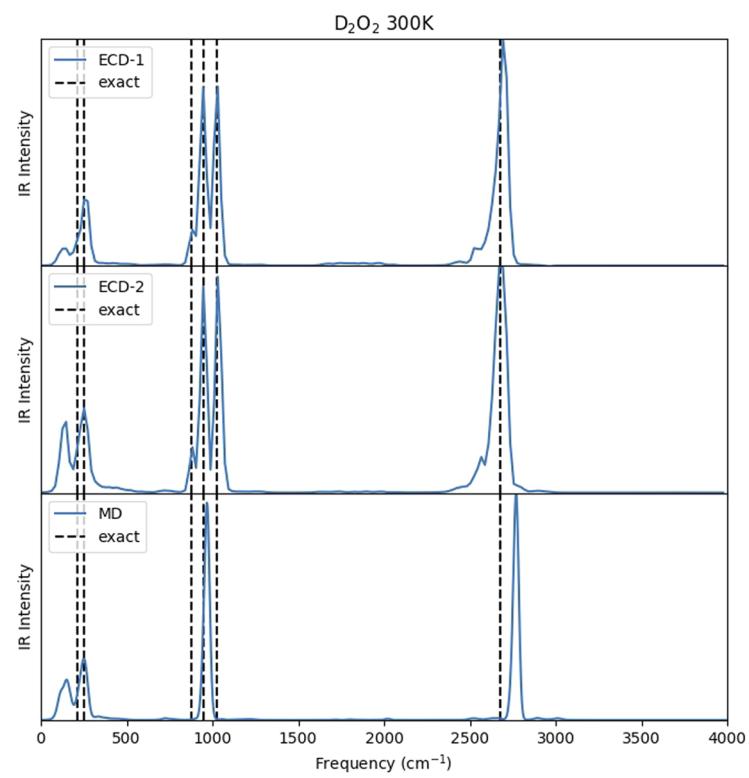
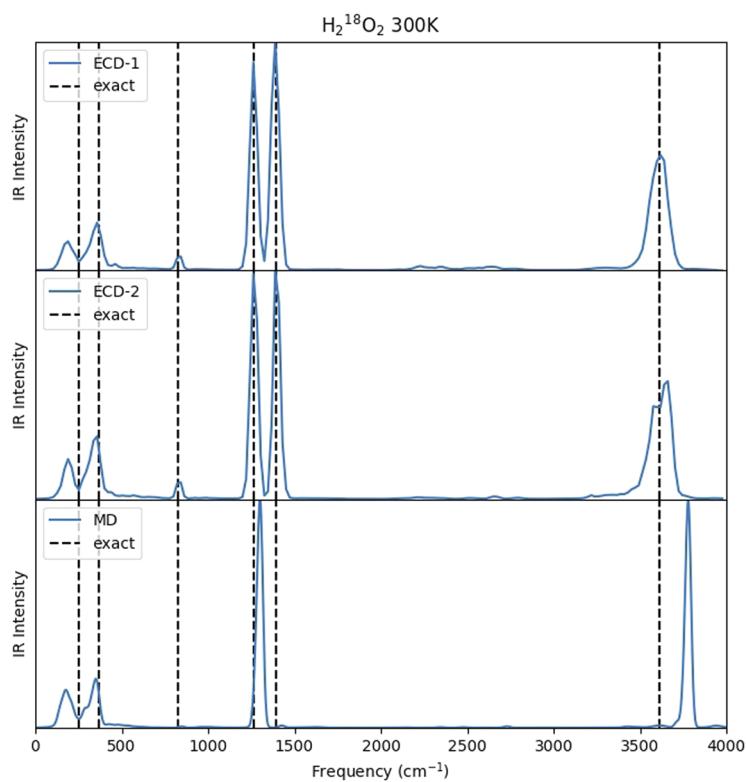


FIG. 6. The IR spectra for D_2O_2 at 300 K.

FIG. 7. The IR spectra for $\text{H}_2^{18}\text{O}_2$ at 300 K.

ECD involve different equations of motion although they fall into the same category. Therefore, they produce slightly different results. The numerical performance of ML-ECD (especially ECD-1) is, in general, robust for the IR spectrum as a function of temperature and isotopic substitution as shown in Figs. 3–7 and Table I.

We then investigate the rovibrational spectrum of the H_2^{16}O molecule at 300 K based on the accurate PES⁹⁴ and DMS⁹⁵ for this realistic molecule. In such a case, rotational and vibrational motions are coupled. As is well known, exact quantum data for the rovibrational spectrum of such a molecule consist of lines. Figure 8 demonstrates the comparison of MD/ML-ECD results to exact data. Because the frequency difference between the lines is so small and ECD is not able to distinguish between two peaks whose difference is less than 40 cm^{-1} , ML-ECD yields three bands: the rotational band below 500 cm^{-1} , the bending band between 1300 and 2000 cm^{-1} , and the stretching band between 3400 and 4000 cm^{-1} , which are in good agreement with the corresponding regimes where exact line data fall in. ECD-1 shows an overall better performance for the “envelopes” of the exact lines. In comparison to ECD-1 (as well as exact data), classical MD also produces three bands, the stretching band ($3400\text{--}4200\text{ cm}^{-1}$) of which is significantly blue-shifted by 180 cm^{-1} and the bending band does so by $30\text{--}40\text{ cm}^{-1}$. It is evident that ML-ECD is capable of describing the rovibrational motion and yielding reasonably accurate bands. Provided that ML-ECD depicts both vibrational and rovibrational spectra

practically well as shown in Figs. 3–8, ML-ECD (in which ECD-1 is recommended) is anticipated to be useful to study vibrational spectra in more complicated environment where both rotational and translational motions are hindered, e.g., in solution or at interfaces.

Despite the overall reasonably good performance of ML-ECD methods, we note that ECD methods fail to describe true quantum recurrence/rephasing effects^{21–26} in the quantum thermal correlation function. In the simulation of the one-dimensional pure quartic potential in Appendix B, despite that ECD works reasonably well for short times for correlation functions (even of nonlinear operators), it is incapable of capturing quantum recurrence/rephasing in longer time. As shown in Figs. 3–7 and Table I, while ML-ECD methods semi-quantitatively capture the large tunneling splitting of the torsion motion (in the fundamental with the lowest frequency), which is about $40\text{--}120\text{ cm}^{-1}$, they are incapable of describing the tunneling splitting of any other fundamental, whose value is smaller than 20 cm^{-1} . ML-ECD methods are not able to distinguish the symmetric and asymmetric O-H (or O-D) stretching modes that are near-degenerate in the peroxide molecule, e.g., the difference between the two modes is less than 3 cm^{-1} for D_2O_2 . Distinguishing the small distance in frequency between two peaks indicates accurate long time quantum dynamics. The envelope of the ECD correlation function does not demonstrate rephasing behaviors for true quantum coherence and deep tunneling splitting. Although long-time quantum recurrence or rephrasing effects (often shown in small bounded

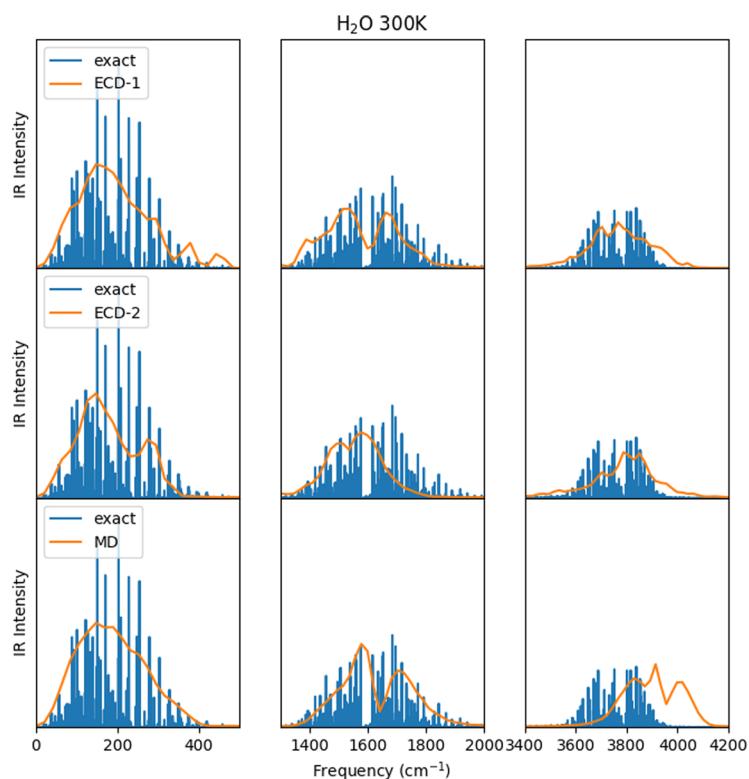
TABLE I. Peak positions of the H₂O₂ molecule at different temperatures and isotope molecules of H₂O₂ at 300 K. The spectra calculated from ML-ECD simulations are accurate to 21 cm⁻¹. PES is from Ref. 104 and “exact” results are from Ref. 97 (unit: cm⁻¹).

System	Peak positions (cm ⁻¹)				
	ECD-1	ECD-2	MD	NMA	Exact
H ₂ O ₂ 300 K	210/357	168/357	207/344	381.9	251.3/371.6
	883	883		910.8	865.4/884.9
	1262	1283	1307	1329.8	1267.8/1277.2
	1388	1388		1437.2	1387.0/1402.4
	3640	3640	3790	3807.7	3614.4/3631.4
				3809.2	3625.1/3622.2
H ₂ O ₂ 100 K	168/357	147/336	362	381.9	251.3/371.6
	883	883		910.8	865.4/884.9
	1262	1283	1323	1329.8	1267.8/1277.2
	1388	1388		1437.2	1387.0/1402.4
	3640	3619	3803	3807.7	3614.4/3631.4
				3809.2	3625.1/3622.2
HDO ₂ 300 K	147/294	168/294	168/305	334.7	234.2/313.9
	883	883		909.6	876.5
	988	988	999	1014.1	982.3
	1346	1346	1386	1388.0	1349.5
	2693	2693	2772	2774.9	2675.2
	3619	3577	3790	3808.4	3622.9
D ₂ O ₂ 300 K	147/252	147/252	149/249	279.4	214.9/254.2
	883	883		908.6	877.4
	946	946	967	983.8	948.2
	1031	1031		1052.7	1027.6
	2693	2672	2769	2772.5	2675.6
				2777.1	2676.3
H ₂ ¹⁸ O ₂ 300 K	189/357	189/357	176/349	380.3	250.8/369.7
	841	841		858.9	825.7
	1262	1262	1301	1322.7	1262.5
	1388	1409		1431.1	1391.5
	3619	3661	3779	3795.0	3611.0
				3796.1	3613.5

systems) are often expected to be quenched by coupling among the various degrees of freedom in condensed phase systems, ML-ECD itself is incompetent in suggesting where these effects do become important. Such problems related to the accuracy of long time quantum dynamics present challenges for ECD and other trajectory-based methods. As shown in Sec. II B, ECD involves only the zeroth order term of an exact series of expansion of the phase space propagator. It will be interesting as well as challenging to see whether the drawbacks of ML-ECD methods can be practically improved for realistic molecular systems in the future.

The $N \times N \times N$ tensor $\left\{ \frac{\partial(M_{\text{therm}})_{ij}}{\partial x_k} \right\}$ is requested in the current ECD equations of motion [Eq. (51) or Eq. (53)] proposed in

Sec. II B 2. Therefore, the numerical effort of each time step of ECD is on the order N^3 . In comparison, the computational cost of a time step in conventional molecular dynamics is from $O(N)$ to $O(N \log N)$, at most no greater than $O(N^2)$. For achieving more economic computation cost on the order of $O(N)$ or $O(N \log N)$, one can certainly employ only the diagonal elements in the current ECD equations of motion or utilize the strategy in Ref. 99, which partitions the nuclear vibrational space into subspaces, but the accuracy is anticipated to be decreased. In addition to these strategies, in the category of ECD, there exist some other types of methods (in the Wigner representation) that lead to a more economic scaling, as we will investigate in the future.

FIG. 8. The rovibrational IR spectra for H_2O at 300 K.

IV. CONCLUDING REMARKS

ECD is a category of phase space quantum dynamics methods in the spirit of the continuity equation of probability distribution, which maintains the stationarity of the quantum canonical distribution and works well for quantum thermal correlation functions even when nonlinear operators are involved. ECD is exact in the harmonic and classical limits and includes only the zeroth order propagator of a rigorous series expansion of the exact propagator. In this paper, we demonstrate that machine learning implementation of ECD can be established by nontrivial extension of the strategies in Deep Potential to fitting the effective force and effective mass matrix from their path integral evaluations prior to integration of the ECD equations of motion. Proof-of-concept applications of ML-ECD to vibrational IR spectra of H_2O_2 and the isotope molecules as well as the rovibrational spectrum of H_2O show that ML-ECD methods are reasonably accurate with modest numerical effort for complex realistic systems. In addition to Eqs. (51) and (53), there exist other possible choices for ECD. It will be interesting in the future to extend the machine learning strategies in this paper to develop even more efficient phase space quantum dynamics methods with reasonable accuracy for studying large molecular systems at surfaces or in the condensed phase.

Along with recent progress on machine learning methods for fitting dynamics-friendly potential energy surfaces from *ab initio* data,^{45–48,65–72} the strategies developed in this paper offer reduction in computation effort via machine learning for phase space quantum dynamics. A combination of advances in fitting the

ab initio PES as well as those in accelerating ECD equations of motion will shed light on more numerically favorable approaches on *ab initio* quantum dynamics in the phase space formulation.

SUPPLEMENTARY MATERIAL

See the [supplementary material](#) for consistency of ML-ECD standard and Kubo-transformed correlation functions.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (NSFC) under Grant No. 21961142017 and by the Ministry of Science and Technology of China (MOST) under Grant No. 2017YFA0204901. We acknowledge the High-Performance Computing Platform of Peking University, Beijing PARATERA Tech Co., Ltd., and the Guangzhou supercomputer center for providing computational resources. We thank Hui Li for useful discussions on the accurate rovibrational spectrum of the H_2O molecule in Ref. 96.

APPENDIX A: SOLVING QUANTUM LIOUVILLE DYNAMICS IN THE SPIRIT OF THE CONTINUITY EQUATION

The essential task for evaluating the quantum thermal correlation function is to numerically solve Eq. (19) or Eq. (21). This

is, however, inconvenient as well as far from trivial in the Eulerian viewpoint (involving partial derivatives at a fixed point in phase space) for multi-dimensional systems. It is often more useful to employ the Lagrangian viewpoint by introducing trajectory-based dynamics. It is well known that Bohm's formulation offers an exact interpretation of quantum mechanics by employing an ensemble of quantum trajectories, equivalent to solving the time-dependent Schrödinger equation. Quantum Liouville dynamics, in principle, can be solved in an exact fashion by an ensemble of (interacting) trajectories,

$$C_{AB}(t) = \frac{1}{Z} \int \mu(d\mathbf{x}_0 d\mathbf{p}_0) \rho_W^{\text{eq}}(\mathbf{x}_0, \mathbf{p}_0; 0) f_{A^\beta}^W(\mathbf{x}_0, \mathbf{p}_0; 0) g_B^W(\mathbf{x}_t, \mathbf{p}_t; t), \quad (\text{A1})$$

where $\mu(d\mathbf{x}_0 d\mathbf{p}_0)$ stands for the invariant measure on the trajectory space. Equation (A1) is open to any type of trajectory-based dynamics. $g_B^W(\mathbf{x}_t, \mathbf{p}_t; t)$ is often an explicitly time-dependent function along the trajectory.

1. Perspective from ordinary differential equations

The exact formulation for the quantum thermal correlation function [Eq. (31)] requests the solution to the partial differential equation (PDE) [Eq. (21)]. If Eq. (21) can be replaced by

$$\frac{\partial \tilde{B}_W(\mathbf{x}, \mathbf{p}; t)}{\partial t} = \mathcal{L}_{\text{con}}^* \tilde{B}_W(\mathbf{x}, \mathbf{p}; t) = \left[\dot{\mathbf{x}} \cdot \frac{\partial}{\partial \mathbf{x}} + \dot{\mathbf{p}} \cdot \frac{\partial}{\partial \mathbf{p}} \right] \tilde{B}_W(\mathbf{x}, \mathbf{p}; t), \quad (\text{A2})$$

where the expression for the dynamical function at the very beginning $\tilde{B}_W(\mathbf{x}, \mathbf{p}; 0)$ is obtained by Eq. (18) with $t = 0$, the explicit formulations of the two ordinary differential equations (ODEs) for $\dot{\mathbf{x}}$ and $\dot{\mathbf{p}}$ should then, in principle, ensure that the right-hand side (RHS) of Eq. (21) is *equal to* that of Eq. (A2).

Consider the Wigner function for operator $\hat{B} = B(\dot{\mathbf{x}})$ at time $t = 0$, i.e., $\tilde{B}_W(\mathbf{x}, \mathbf{p}; 0) = B(\mathbf{x})$. Its exact evolution at time $t = 0$ from Eq. (21) states that

$$\frac{\partial \tilde{B}_W(\mathbf{x}, \mathbf{p}; t)}{\partial t} \Big|_{t=0} = \mathbf{p}^T \mathbf{M}^{-1} \frac{\partial B(\mathbf{x})}{\partial \mathbf{x}}. \quad (\text{A3})$$

Correspondingly, Eq. (A2) produces

$$\frac{\partial \tilde{B}_W(\mathbf{x}, \mathbf{p}; t)}{\partial t} \Big|_{t=0} = \dot{\mathbf{x}} \cdot \frac{\partial B(\mathbf{x})}{\partial \mathbf{x}}. \quad (\text{A4})$$

A reasonable option for the ODE for $\dot{\mathbf{x}}$ satisfying Eq. (A3) then reads

$$\dot{\mathbf{x}} = \mathbf{M}^{-1} \mathbf{p}. \quad (\text{A5})$$

It yields the ODE for $\dot{\mathbf{p}}$ requested by

$$\dot{\mathbf{p}} \cdot \frac{\partial}{\partial \mathbf{p}} \tilde{B}_W(\mathbf{x}, \mathbf{p}; t) = \left[-\frac{2}{\hbar} V(\mathbf{x}) \sin\left(\frac{\hbar}{2} \overleftarrow{\frac{\partial}{\partial \mathbf{x}}} \cdot \overrightarrow{\frac{\partial}{\partial \mathbf{p}}}\right) \right] \tilde{B}_W(\mathbf{x}, \mathbf{p}; t) \quad (\text{A6})$$

such that Eq. (21) is equivalent to Eq. (A2). When the RHS of Eq. (A6) involves third and higher-order derivatives, the ODE for $\dot{\mathbf{p}}$ is often *ill-defined*. More generally, except for linear systems (such as harmonic potentials), it is impossible to design a pair of ODEs for $\{\dot{\mathbf{x}}, \dot{\mathbf{p}}\}$ in Eq. (A2) to make Eqs. (21) and (A2) equivalent. That is, Eq. (A2) is doomed to be an imperfect approximation to Eq. (21) for

general systems. There exist other options for the definition $\mathcal{L}_{\text{con}}^*$ in Eq. (A2). As long as only first derivatives of $\{\tilde{B}_W, \dot{\mathbf{x}}, \dot{\mathbf{p}}\}$ over (\mathbf{x}, \mathbf{p}) are included, the conclusion remains the same.

Provided that a pair of ODEs for $\{\dot{\mathbf{x}}, \dot{\mathbf{p}}\}$ are well-defined in the approximation Eq. (A2), it is trivial to obtain the solution to Eq. (A2),

$$\exp(\mathcal{L}_{\text{con}}^* t) \tilde{B}_W(\mathbf{x}, \mathbf{p}; 0) = \tilde{B}_W(\exp(\mathcal{L}_{\text{con}}^* t) \mathbf{x}, \exp(\mathcal{L}_{\text{con}}^* t) \mathbf{p}; 0) \quad (\text{A7})$$

$$= \tilde{B}_W(\mathbf{x}_t(\mathbf{x}, \mathbf{p}; t), \mathbf{p}_t(\mathbf{x}, \mathbf{p}; t); 0), \quad (\text{A8})$$

as long as there exists the uniqueness of the solution $\{\mathbf{x}_t(\mathbf{x}, \mathbf{p}; t), \mathbf{p}_t(\mathbf{x}, \mathbf{p}; t)\}$ to the two ODEs for $\{\dot{\mathbf{x}}, \dot{\mathbf{p}}\}$, which passes through any specified phase point at given time t .

On condition that Eq. (A7) holds, the approximation to the quantum thermal correlation function Eq. (31) reads

$$C_{AB}(t) \approx \frac{1}{Z} \int d\mathbf{x} \int d\mathbf{p} A_W^\beta(\mathbf{x}, \mathbf{p}; 0) \tilde{B}_W(\mathbf{x}_t, \mathbf{p}_t; 0). \quad (\text{A9})$$

The RHS of Eq. (A9) is straightforwardly evaluated by following the trajectory $(\mathbf{x}_t, \mathbf{p}_t)$, i.e., the solution to the pair of ODEs for $\{\dot{\mathbf{x}}, \dot{\mathbf{p}}\}$.

On the assumption that $\dot{\Omega} A_W^\beta(\Omega; 0) \tilde{B}_W(\Omega; 0)$ vanishes on the boundary of (Wigner) phase space, the volume integral

$$\begin{aligned} & \int_{\text{volume}} d\Omega \frac{\partial}{\partial \Omega} \cdot \left(\dot{\Omega} A_W^\beta(\Omega; 0) \tilde{B}_W(\Omega; 0) \right) \\ &= \oint_{\text{surface}} d\mathbf{S} \cdot \left(\dot{\Omega} A_W^\beta(\Omega; 0) \tilde{B}_W(\Omega; 0) \right) = 0. \end{aligned} \quad (\text{A10})$$

Here, \mathbf{S} is a unit vector outward normal to the boundary, Ω denotes $\{\mathbf{x}, \mathbf{p}\}$, and $\dot{\Omega}$ denotes $\{\dot{\mathbf{x}}, \dot{\mathbf{p}}\}$. Provided that Eq. (A10) holds, an equivalent expression for the RHS of Eq. (A9) is

$$C_{AB}(t) \approx \frac{1}{Z} \int d\mathbf{x} \int d\mathbf{p} \left[\exp(\mathcal{L}_{\text{con}} t) A_W^\beta(\mathbf{x}, \mathbf{p}; 0) \right] \tilde{B}_W(\mathbf{x}, \mathbf{p}; 0), \quad (\text{A11})$$

where \mathcal{L}_{con} is an operator adjoint to $\mathcal{L}_{\text{con}}^*$ of Eq. (A2), i.e.,

$$\begin{aligned} \frac{\partial A_W^\beta(\mathbf{x}, \mathbf{p}; t)}{\partial t} &= \mathcal{L}_{\text{con}} A_W^\beta(\mathbf{x}, \mathbf{p}; t) \\ &= \left[-\frac{\partial}{\partial \mathbf{x}} \cdot (\dot{\mathbf{x}}) - \frac{\partial}{\partial \mathbf{p}} \cdot (\dot{\mathbf{p}}) \right] A_W^\beta(\mathbf{x}, \mathbf{p}; t). \end{aligned} \quad (\text{A12})$$

Note that the adjoint operators \mathcal{L}_{con} and $\mathcal{L}_{\text{con}}^*$ are *not* anti-self-adjoint. The RHS of Eq. (A12) contains only first derivatives, while that of Eq. (19) often includes third and higher-order derivatives. Except for linear systems, regardless of the expression of the pair of well-defined ODEs for $\{\dot{\mathbf{x}}, \dot{\mathbf{p}}\}$, it is also unattainable to make Eq. (A12) equivalent to Eq. (19). This states that Eq. (A11) with (A12) is, in general, an approximation to the quantum thermal correlation function Eq. (8). A comparison between Eqs. (A12) and (19) sheds little light on any reasonable choice for the pair of ODEs for $\{\dot{\mathbf{x}}, \dot{\mathbf{p}}\}$. The exception is that, when operator $\hat{A} = \hat{1}$ is an identity operator in Eq. (A12), the analogy to the stationary condition of Eq. (32) suggests the left-hand side (LHS) of Eq. (A12) should be zero, i.e.,

$$\left[-\frac{\partial}{\partial \mathbf{x}} \cdot (\dot{\mathbf{x}}) - \frac{\partial}{\partial \mathbf{p}} \cdot (\dot{\mathbf{p}}) \right] \rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; 0) = 0. \quad (\text{A13})$$

We emphasize that Eq. (A13) is not equivalent to Eq. (32) whose intrinsic dynamics depends on Eq. (19) although the stationarity in Eq. (A13) also states

$$\rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; t) = \rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; 0). \quad (\text{A14})$$

When the ODE for $\dot{\mathbf{x}}$ is chosen to be Eq. (A5), Eq. (A13) yields

$$\mathbf{p}^T \mathbf{M}^{-1} \frac{\partial \rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; 0)}{\partial \mathbf{x}} = \frac{\partial}{\partial \mathbf{p}} \cdot (\rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; 0) \dot{\mathbf{p}}). \quad (\text{A15})$$

In many cases, various options are possible for the ODE for $\dot{\mathbf{p}}$ satisfying Eq. (A15).

2. A more general viewpoint from the continuity equation

a. Probability distribution function and the continuity equation

As long as the probability distribution function in phase space $\mathcal{P}(\Omega; t)$ can be defined, we have the conservation

$$\int d\Omega \mathcal{P}(\Omega; t) = 1 \quad \text{all } t. \quad (\text{A16})$$

$\mathcal{P}(\Omega; t)d\Omega$ represents the probability of finding the system state in the infinitesimal area $d\Omega$ around the phase point Ω at time t .

$\int d\Omega \mathcal{P}(\Omega; t)$ is then the probability in the volume. $\dot{\Omega}(\Omega; t)$ is the volume velocity (of the system) at a given phase point Ω at time t . Assume that no creation or annihilation occurs in the full space at any time. The probability flowing in unit time through an element dS of the surface bounding the volume is $\mathcal{P}(\Omega; t)\dot{\Omega}(\Omega; t) \cdot dS$. The magnitude of dS is equal to the area of the surface element and the direction of dS is directed normally outward from the volume that the entire surface encloses. Denote

$$\mathbf{j}(\Omega; t) = \mathcal{P}(\Omega; t)\dot{\Omega}(\Omega; t) \quad (\text{A17})$$

the probability current density. The total probability flowing out of the volume (through the boundary) in unit time is

$$\oint_{\text{surface}} dS \cdot \mathbf{j}(\Omega; t). \quad (\text{A18})$$

The decrease in the probability in the volume per unit time is

$$-\frac{\partial}{\partial t} \int_{\text{volume}} d\Omega \mathcal{P}(\Omega; t). \quad (\text{A19})$$

The conservation of probability Eq. (A16) states that these two quantities must be equal, i.e.,

$$\begin{aligned} \int_{\text{volume}} d\Omega \left[-\frac{\partial}{\partial t} \mathcal{P}(\Omega; t) \right] &= \oint_{\text{surface}} dS \cdot \mathbf{j}(\Omega; t) \\ &= \int_{\text{volume}} d\Omega \frac{\partial}{\partial \Omega} \cdot \mathbf{j}(\Omega; t). \end{aligned} \quad (\text{A20})$$

Since Eq. (A20) must hold for any volume, it yields the continuity equation

$$\frac{\partial}{\partial t} \mathcal{P}(\Omega; t) = -\frac{\partial}{\partial \Omega} \cdot \mathbf{j}(\Omega; t). \quad (\text{A21})$$

Substitution of Eq. (A17) into Eq. (A21) produces

$$\begin{aligned} \frac{\partial \mathcal{P}(\mathbf{x}, \mathbf{p}; t)}{\partial t} &= \left[-\frac{\partial}{\partial \mathbf{x}} \cdot (\dot{\mathbf{x}}) - \frac{\partial}{\partial \mathbf{p}} \cdot (\dot{\mathbf{p}}) \right] \mathcal{P}(\mathbf{x}, \mathbf{p}; t) \\ &= \mathcal{L}_{\text{con}} \mathcal{P}(\mathbf{x}, \mathbf{p}; t). \end{aligned} \quad (\text{A22})$$

The Kolmogorov operator \mathcal{L}_{con} in the evolution of the probability distribution Eq. (A22) contains only first derivatives, which is different from the Liouville operator in Eq. (19) that involves third or higher-order derivatives in addition. The continuity equation [Eq. (A21) or Eq. (A22)] always holds, independent of the expression of the correlation function Eq. (A1) and regardless of whether deterministic or stochastic equations of motion [for $\dot{\Omega}(\Omega; t)$] are involved.

In quantum mechanics, the probability distribution function is well-defined in the coordinate space. That is, $\mathcal{P}(\mathbf{x}; t)d\mathbf{x} = \langle \mathbf{x} | \hat{\rho}(t) | \mathbf{x} \rangle d\mathbf{x}$ is the probability of finding the system in the region $d\mathbf{x}$ around the point \mathbf{x} at time t . The probability current density is

$$\mathbf{j}(\mathbf{x}; t) = \left\langle \mathbf{x} \left| \frac{1}{2} \mathbf{M}^{-1} (\hat{\rho}(t) \hat{\mathbf{p}} + \hat{\mathbf{p}} \hat{\rho}(t)) \right| \mathbf{x} \right\rangle. \quad (\text{A23})$$

The conservation of probability

$$\int d\mathbf{x} \mathcal{P}(\mathbf{x}; t) = 1 \quad \text{all } t \quad (\text{A24})$$

yields the corresponding continuity equation

$$\frac{\partial \mathcal{P}(\mathbf{x}; t)}{\partial t} = -\frac{\partial}{\partial \mathbf{x}} \cdot \mathbf{j}(\mathbf{x}; t). \quad (\text{A25})$$

Due to the Heisenberg uncertainty principle, it is *impossible* to define a joint probability function at the phase point (\mathbf{x}, \mathbf{p}) in quantum mechanics. When the (Wigner) phase space is used, the subtlety is that the Wigner density distribution function $\rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; t)/Z$ is not a probability distribution function, or in other words, $\rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; t)d\mathbf{x}d\mathbf{p}/Z$ is *not* a probability. Even though the conservation

$$\frac{1}{Z} \int d\mathbf{x} \int d\mathbf{p} \rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; t) = 1 \quad \text{all } t \quad (\text{A26})$$

holds, it does *not* ensure the corresponding continuity equation

$$\frac{\partial \rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; t)}{\partial t} = \left[-\frac{\partial}{\partial \mathbf{x}} \cdot (\dot{\mathbf{x}}) - \frac{\partial}{\partial \mathbf{p}} \cdot (\dot{\mathbf{p}}) \right] \rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; t) \quad (\text{A27})$$

at any time t in physics.

As trajectory-based dynamics is employed in multi-dimensional systems, it is often necessary to define the probability distribution function for Monte Carlo as well as molecular dynamics techniques. When the Wigner density function $\rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; 0) \geq 0$, one can use it as the initial probability distribution, i.e., $\mathcal{P}(\mathbf{x}, \mathbf{p}; 0) = \rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; 0)/Z$, trajectory-based dynamics leads to the evolution of the probability distribution governed by Eq. (A22), which does *not* rely on the expression of the correlation function Eq. (A1) as well as the formulation of the trajectory-based dynamics. Even though

the initial probability is the same as the initial density distribution, the continuity equation that the probability evolution must obey never reproduces the (generalized) Wigner–Moyal equation [Eq. (19) or Eq. (20)] that governs the evolution of the (general) density distribution, except for linear systems.

When the Wigner density function contains both positive and negative values, one can divide the whole phase space into positive-valued and negative-valued regions. The Wigner density function on the boundary surface between a positive-valued region and a negative-valued one is zero. (Consequently, $A_W^\beta(\Omega; 0)$ vanishes on the boundary surface.) In a region where the Wigner density function $\rho_W^{\text{eq}}(\Omega; 0)$ is negative, one can define the initial probability distribution function $\mathcal{P}(\Omega; 0) = -\rho_W^{\text{eq}}(\Omega; 0)/Z$ in this region. It is easy to show that the overall probability in such a region $\int d\Omega \mathcal{P}(\Omega; t)$ is constant for all time, or equivalently, the continuity equation still holds in this region. One can then use the evolution of $-\mathcal{P}(\Omega; t)$ in the region as a starting point for studying the evolution of the Wigner density distribution in the same negative-valued region.

The conclusion also applies to the Wigner density distribution function as well as other types of density distribution functions for nonequilibrium systems. Except for linear systems (such as quadratic Hamiltonian systems), the continuity equation used for the density distribution function in phase space^{20,100} [Eq. (A27)] is not exact but an approximation. More accurately, Eq. (A27) involves the zeroth order propagator of a rigorous series expansion of the exact propagator, as shown in Sec. II B.

b. Equilibrium continuity dynamics

On condition that Eq. (A10) is satisfied, integration by parts yields the operator adjoint to the Kolmogorov operator \mathcal{L}_{con} of Eq. (A22), which is $\mathcal{L}_{\text{con}}^*$ in Eq. (A2). Consequently, the Kolmogorov operator \mathcal{L}_{con} is not anti-self-adjoint in phase space. When no trajectories cross with each other at any time, Eq. (A7) holds.

As the probability distribution evolves, the time-dependent ensemble average of any property with the dual function \tilde{B}_W is

$$\langle \tilde{B}_W(t) \rangle_{\text{ens}} = \int d\mathbf{x} \int d\mathbf{p} \mathcal{P}(\mathbf{x}, \mathbf{p}; t) \tilde{B}_W(\mathbf{x}, \mathbf{p}; 0). \quad (\text{A28})$$

It is equivalent to the average of the dynamical property along the ensemble of trajectories in phase space,

$$\langle \tilde{B}_W(t) \rangle_{\text{ens}} = \int d\mathbf{x} \int d\mathbf{p} \mathcal{P}(\mathbf{x}, \mathbf{p}; 0) \tilde{B}_W(\mathbf{x}_t(\mathbf{x}, \mathbf{p}; t), \mathbf{p}_t(\mathbf{x}, \mathbf{p}; t); 0). \quad (\text{A29})$$

When the equations of motion for the trajectory satisfy

$$\frac{\partial \mathcal{P}(\mathbf{x}, \mathbf{p}; t)}{\partial t} = 0, \quad (\text{A30})$$

the probability distribution

$$\mathcal{P}(\mathbf{x}, \mathbf{p}; t) = \mathcal{P}_{\text{eq}}(\mathbf{x}, \mathbf{p}) = \rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; 0)/Z \quad (\text{A31})$$

is stationary. Or equivalently,

$$\mathcal{L}_{\text{con}} \mathcal{P}_{\text{eq}}(\mathbf{x}, \mathbf{p}) = \left[-\frac{\partial}{\partial \mathbf{x}} \cdot (\dot{\mathbf{x}}) - \frac{\partial}{\partial \mathbf{p}} \cdot (\dot{\mathbf{p}}) \right] \mathcal{P}_{\text{eq}}(\mathbf{x}, \mathbf{p}) = 0. \quad (\text{A32})$$

It is then easy to verify

$$\mathcal{L}_{\text{con}} [\mathcal{P}(\mathbf{x}, \mathbf{p}; t) f_{A^\beta}^W(\mathbf{x}, \mathbf{p}; t)] = \mathcal{P}_{\text{eq}}(\mathbf{x}, \mathbf{p}) (-\mathcal{L}_{\text{con}}^* f_{A^\beta}^W(\mathbf{x}, \mathbf{p}; t)). \quad (\text{A33})$$

Define the correlation function along the ensemble of trajectories in phase space for Eq. (A30),

$$\langle f_{A^\beta}^W(0) \tilde{B}_W(t) \rangle_{\text{ens}} = \int d\mathbf{x} \int d\mathbf{p} \mathcal{P}_{\text{eq}}(\mathbf{x}, \mathbf{p}) f_{A^\beta}^W(\mathbf{x}, \mathbf{p}; 0) \times \tilde{B}_W(\mathbf{x}_t(\mathbf{x}, \mathbf{p}; t), \mathbf{p}_t(\mathbf{x}, \mathbf{p}; t); 0). \quad (\text{A34})$$

Equation (A34) can be recast into

$$\langle f_{A^\beta}^W(0) \tilde{B}_W(t) \rangle_{\text{ens}} = \int d\mathbf{x} \int d\mathbf{p} \exp(\mathcal{L}_{\text{con}} t) \times [\mathcal{P}_{\text{eq}}(\mathbf{x}, \mathbf{p}) f_{A^\beta}^W(\mathbf{x}, \mathbf{p}; 0)] \tilde{B}_W(\mathbf{x}, \mathbf{p}; 0). \quad (\text{A35})$$

By virtue of Eq. (A33), we obtain from Eq. (A35) the following equation:

$$\langle f_{A^\beta}^W(0) \tilde{B}_W(t) \rangle_{\text{ens}} = \int d\mathbf{x} \int d\mathbf{p} \mathcal{P}_{\text{eq}}(\mathbf{x}, \mathbf{p}) \times [\exp(-\mathcal{L}_{\text{con}}^* t) f_{A^\beta}^W(\mathbf{x}, \mathbf{p}; 0)] \tilde{B}_W(\mathbf{x}, \mathbf{p}; 0). \quad (\text{A36})$$

The RHS of Eq. (A36) is equivalent to

$$\langle f_{A^\beta}^W(-t) \tilde{B}_W(0) \rangle_{\text{ens}} = \int d\mathbf{x} \int d\mathbf{p} \mathcal{P}_{\text{eq}}(\mathbf{x}, \mathbf{p}) \times f_{A^\beta}^W(\mathbf{x}_{-t}(\mathbf{x}, \mathbf{p}; -t), \mathbf{p}_{-t}(\mathbf{x}, \mathbf{p}; -t); 0) \times \tilde{B}_W(\mathbf{x}, \mathbf{p}; 0) \quad (\text{A37})$$

after following Eq. (A7). Similarly, it is trivial to verify a more general relation

$$\langle f_{A^\beta}^W(0) \tilde{B}_W(t) \rangle_{\text{ens}} = \langle f_{A^\beta}^W(t') \tilde{B}_W(t' + t) \rangle_{\text{ens}}. \quad (\text{A38})$$

When $f_{A^\beta}^W = 1$ is an identity property, Eq. (A36) states that $\langle \tilde{B}_W(t) \rangle_{\text{ens}}$ is time-independent, i.e.,

$$\langle \tilde{B}_W(t) \rangle_{\text{ens}} = \langle \tilde{B}_W(0) \rangle_{\text{ens}}. \quad (\text{A39})$$

Not less importantly, substitution of $\tilde{B} = 1$ into Eqs. (A35) and (A37) proves that the ensemble average

$$\int d\mathbf{x} \int d\mathbf{p} \exp(\mathcal{L}_{\text{con}} t) [\mathcal{P}_{\text{eq}}(\mathbf{x}, \mathbf{p}) f_{A^\beta}^W(\mathbf{x}, \mathbf{p}; 0)] = \langle f_{A^\beta}^W(-t) \rangle_{\text{ens}} = \langle f_{A^\beta}^W(0) \rangle_{\text{ens}} \quad (\text{A40})$$

is invariant with time.

A general formulation for the equations of motion for $(\mathbf{x}_t, \mathbf{p}_t)$ reads

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{M}^{-1} \mathbf{P}_{\text{eff}}(\mathbf{x}, \mathbf{p}, t), \\ \dot{\mathbf{p}} = \mathbf{F}_{\text{eff}}(\mathbf{x}, \mathbf{p}, t), \end{cases} \quad (\text{A41})$$

where $\mathbf{P}_{\text{eff}}(\mathbf{x}, \mathbf{p}, t)$ and $\mathbf{F}_{\text{eff}}(\mathbf{x}, \mathbf{p}, t)$ are the effective momentum vector and effective force, respectively. Equations (A31) and (A32) request

$$\frac{\partial}{\partial \mathbf{x}} \cdot [\rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; 0) \mathbf{M}^{-1} \mathbf{P}_{\text{eff}}(\mathbf{x}, \mathbf{p}, t)] + \frac{\partial}{\partial \mathbf{p}} \cdot [\rho_W^{\text{eq}}(\mathbf{x}, \mathbf{p}; 0) \mathbf{F}_{\text{eff}}(\mathbf{x}, \mathbf{p}, t)] = 0. \quad (\text{A42})$$

Any deterministic or stochastic trajectory-based dynamics satisfying Eqs. (A41) and (A42) falls into the category of equilibrium continuity dynamics (ECD).

We then employ an ensemble of ECD trajectories to express the quantum thermal correlation function. The initial phase point $(\mathbf{x}_0, \mathbf{p}_0)$ can be any point in the Wigner phase space and is sampled from the initial probability distribution function $\mathcal{P}_{\text{eq}}(\mathbf{x}_0, \mathbf{p}_0)$. The formulation of the quantum thermal correlation function Eq. (A1) can be recast into

$$C_{AB}(t) = \int d\mathbf{x}_0 \int d\mathbf{p}_0 \mathcal{P}_{\text{eq}}(\mathbf{x}_0, \mathbf{p}_0) f_{A^\beta}^W(\mathbf{x}_0, \mathbf{p}_0; 0) \times [\exp(\mathcal{Y}_{\text{mod}} t) \tilde{B}_W(\mathbf{x}_t, \mathbf{p}_t; 0)], \quad (\text{A43})$$

with the modification propagator $\exp(\mathcal{Y}_{\text{mod}} t)$ defined by Eq. (41). Because we have the modification propagator $\exp(\mathcal{Y}_{\text{mod}} t) = 1$ in the harmonic limit, the simplest reasonable approximation Eq. (42) for general systems produces the approximate expression of the quantum thermal correlation function,

$$C_{AB}(t) \approx \int d\mathbf{x}_0 \int d\mathbf{p}_0 \mathcal{P}_{\text{eq}}(\mathbf{x}_0, \mathbf{p}_0) f_{A^\beta}^W(\mathbf{x}_0, \mathbf{p}_0; 0) \tilde{B}_W(\mathbf{x}_t, \mathbf{p}_t; 0) = \langle f_{A^\beta}^W(0) \tilde{B}_W(t) \rangle_{\text{ens}}. \quad (\text{A44})$$

Equations (43)–(46) represent a series expansion of $U(t) = \exp(\mathcal{L}^* t) = \exp(\mathcal{Y}_{\text{mod}} t) \exp(\mathcal{L}_{\text{cont}}^* t)$, where ECD includes only the zeroth order propagator. Other than simply postulating the continuity equation in quantum phase space or Eq. (A2), Appendix A 2 offers much more useful insight.

It is straightforward to show that the continuity equation [Eq. (A22)] leads to

$$\mathcal{P}_{\text{eq}}(\mathbf{x}_0, \mathbf{p}_0) d\mathbf{x}_0 d\mathbf{p}_0 = \mathcal{P}_{\text{eq}}(\mathbf{x}_t, \mathbf{p}_t) d\mathbf{x}_t d\mathbf{p}_t. \quad (\text{A45})$$

Thus, a time averaging can be employed

$$C_{AB}(t) = \frac{1}{T_0} \int_0^{T_0} dt' \langle \hat{A}(t') \hat{B}(t' + t) \rangle \quad (\text{A46})$$

$$\approx \left\langle \frac{1}{T_0} \int_0^{T_0} dt' f_{A^\beta}^W(\mathbf{x}_{t'}, \mathbf{p}_{t'}; 0) B_W(\mathbf{x}_{t'+t}, \mathbf{p}_{t'+t}; 0) \right\rangle_{\text{ens}}. \quad (\text{A47})$$

If the trajectory-based dynamics is ergodic for the system, the quantum thermal correlation function can (approximately) be estimated by several long time trajectories.

For the same reason discussed in Appendix A 2 a, it is easy to extend ECD to the case in which the equilibrium Wigner density function includes both positive and negative values. In each positive-valued or negative-valued region, most equations stay the same. For instance, the continuity equation in this case yields Eq. (A45) in each region such that a more compact form in the whole Wigner phase space is

$$\rho_W^{\text{eq}}(\mathbf{x}_0, \mathbf{p}_0) d\mathbf{x}_0 d\mathbf{p}_0 = \rho_W^{\text{eq}}(\mathbf{x}_t, \mathbf{p}_t) d\mathbf{x}_t d\mathbf{p}_t. \quad (\text{A48})$$

It is more convenient to employ *independent* rather than interacting ECD trajectories for the approximate formulation Eq. (A44) or Eq. (A47). In this paper, the deterministic equations of motion for ECD in Eq. (A41) are employed, as it is warranted to recover (deterministic) classical dynamics in the classical limit ($\hbar \rightarrow 0$). Equation (A44) evidently yields exact results at time $t = 0$. Consider the short-time limit described in Eq. (A3) for the Wigner function for the coordinate-dependent operator $\hat{B} = B(\hat{\mathbf{x}})$. Equation (A5) is requested in order to obtain the exact results in the short-time limit. The ECD equations of motion Eqs. (A41) and (A42) become Eq. (34) with the effective force obtained from Eq. (35). Evidently, various options for the effective force satisfying Eq. (35) are possible.

APPENDIX B: FIRST ORDER CORRECTION TO ECD DYNAMICS

Equation (45) with Eq. (46) forms an exact series expansion, where ECD involves only the zeroth order term. We investigate the influence of the first order correction of Eq. (45) on ECD dynamics. We use the quartic potential $U(x) = x^4/4$ with mass $M = 1$ as an example where the exact operator $\mathcal{L}^* = -\mathcal{L}$ is available,

$$\mathcal{L}^* B = \frac{p}{M} \frac{\partial B}{\partial x} - x^3 \frac{\partial B}{\partial p} + \frac{\hbar^2}{4} x \frac{\partial^3 B}{\partial p^3}. \quad (\text{B1})$$

Note that when one-dimensional systems are considered, the equations of motion of ECD-1 occur to be the same as those of ECD-2, which are

$$\begin{aligned} \dot{x} &= \frac{p}{M}, \\ \dot{p} &= -\frac{M_{\text{therm}}}{M} \frac{\partial \tilde{V}_{\text{eff}}}{\partial x} - \frac{p^2}{2M_{\text{therm}}M} \frac{\partial M_{\text{therm}}}{\partial x} - \frac{1}{2\beta M} \frac{\partial M_{\text{therm}}}{\partial x}. \end{aligned} \quad (\text{B2})$$

Exact data are obtained by the discrete variable representation (DVR) method of Ref. 101 with 100 grids distributed in $(-6, 6)$ in the coordinate space. For ECD, we use 100×100 grids in phase space, the range of position is $x \in (-3, 3)$, and the range of momentum is $p \in (-6, 6)$. LGA with the thermal mass defined in Eq. (50) is employed for the initial condition where the density matrix elements in Eq. (50) are obtained from the DVR discussed above. The Crank-Nicolson method¹⁰² with a time step 0.1 a.u. is employed to propagate the function $B(x, p; t)$ for ECD dynamics as well as its first-order correction. We study the position correlation function $\langle \hat{x}(0) \hat{x}(t) \rangle_{\text{std}}$, position square correlation function $\langle \hat{x}^2(0) \hat{x}^2(t) \rangle_{\text{std}}$, and Kubo-transformed momentum correlation function $\langle \hat{p}(0) \hat{p}(t) \rangle_{\text{Kubo}}$ at $\beta = 8$ and at $\beta = 1$. Results are demonstrated in Figs. 9 and 10. In this challenging example where no harmonic term exists in the potential function, ECD works reasonably good for quantum thermal correlation functions (of even nonlinear operators) for short times (e.g., in the first period of oscillation of the correlation function). It is evident that ECD fails to capture quantum recurrence/coherence in longer time, e.g., as shown in Panel (b) of Fig. 10. Encouragingly, the first order correction in Eq. (45) is able to systematically improve over ECD in Figs. 9 and 10. The improvement is more pronounced at

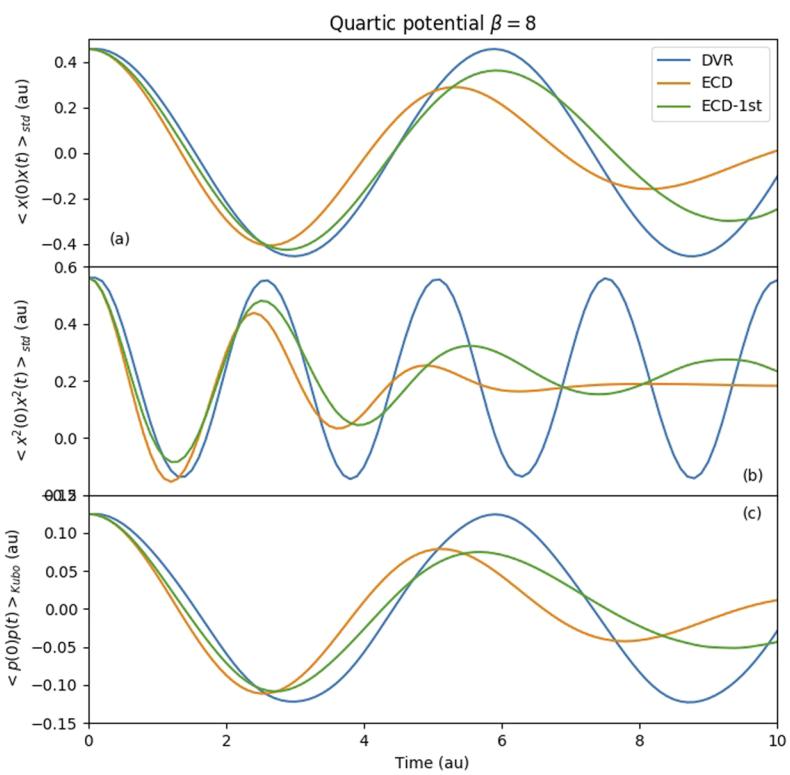


FIG. 9. The position autocorrelation function, position square autocorrelation function, and Kubo-transformed momentum autocorrelation function of 1D quartic potential at $\beta = 8$. “ECD-1st” represents the first-order correction to ECD in Eq. (45).

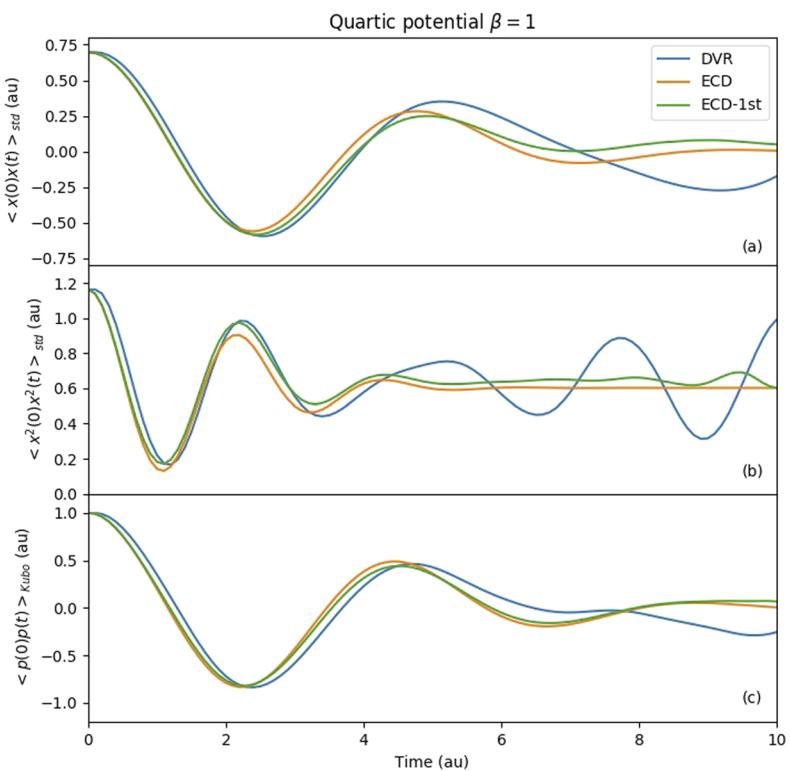


FIG. 10. The position autocorrelation function, position square autocorrelation function, and Kubo-transformed momentum autocorrelation function of 1D quartic potential at $\beta = 1$. “ECD-1st” represents the first-order correction to ECD in Eq. (45).

lower temperature, where quantum effects are often more significant. It is expected from the exact series expansion of Eq. (45) that higher order corrections will improve dynamics behaviors for longer time. It will be interesting as well as challenging in the future to investigate how to utilize the correction strategy for general large molecular systems.

APPENDIX C: A GENERAL FORMULATION FOR THE DESCRIPTOR

Consider the triatomic molecule H₂O. Denote the vector from the O atom to one H atom \mathbf{r}_1 and that from the O atom to the other H atom \mathbf{r}_2 . If we use Eq. (59) to describe the tensor feature of atoms, the feature of atom O reads

$$\overleftrightarrow{D}_O^{lm} = \frac{1}{2} \sum_{j,k=1}^2 G_{OH}^l(s(r_j)) G_{OH}^m(s(r_k)) (\mathbf{r}_j \mathbf{r}_k^T + \mathbf{r}_k \mathbf{r}_j^T), \quad (C1)$$

$$l = 1, \dots, M_1, \quad m = 1, \dots, M_2.$$

When $r_1 = r_2 = r$, the feature of atom O becomes

$$\overleftrightarrow{D}_O^{lm} = G_{OH}^l(s(r)) G_{OH}^m(s(r)) (\mathbf{r}_1 \mathbf{r}_1^T + \mathbf{r}_1 \mathbf{r}_2^T + \mathbf{r}_2 \mathbf{r}_1^T + \mathbf{r}_2 \mathbf{r}_2^T) \\ = G_{OH}^l(s(r)) G_{OH}^m(s(r)) (\mathbf{r}_1 + \mathbf{r}_2)(\mathbf{r}_1 + \mathbf{r}_2)^T, \quad (C2)$$

$$l = 1, \dots, M_1, \quad m = 1, \dots, M_2$$

whose rank is always 1, while the rank of the tensor property of interest may be larger than 1. Thus, it buries over-redundant symmetry that impairs the competence of description in the machine learning model and causes discontinuity.

In contrast, when Eq. (61) is employed to describe the tensor feature of atoms, the feature of atom O is

$$\overleftrightarrow{D}_O^l = \frac{1}{2} \sum_{j,k=1}^2 G_{OHH}^l(s(r_j), s(r_k), \cos \theta_{Ojk}) (\mathbf{r}_j \mathbf{r}_k^T + \mathbf{r}_k \mathbf{r}_j^T), \quad (C3)$$

$$l = 1, \dots, F.$$

Denote the HOH angle $\theta = \theta_{O12} = \theta_{O21} = \mathbf{r}_1^T \mathbf{r}_2 / (r_1 r_2)$. Since $\cos \theta_{O11} = \mathbf{r}_1^T \mathbf{r}_1 / (r_1 r_1) = \cos \theta_{O22} = \mathbf{r}_2^T \mathbf{r}_2 / (r_2 r_2) = 1$, when $r_1 = r_2 = r$, the feature of atom O will be

$$\overleftrightarrow{D}_O^l = G_{OHH}^l(s(r), s(r), 1) (\mathbf{r}_1 \mathbf{r}_1^T + \mathbf{r}_2 \mathbf{r}_2^T) \\ + G_{OHH}^l(s(r), s(r), \cos \theta) (\mathbf{r}_1 \mathbf{r}_2^T + \mathbf{r}_2 \mathbf{r}_1^T), \quad l = 1, \dots, F \quad (C4)$$

whose rank is not 1 for general cases. This offers a more powerful and useful description than Eq. (59) does.

It is well known that high symmetry points in realistic molecular systems often cause problems in fitting the PES (a scalar function), especially when the PES is used for quantum dynamics for scattering reactions and (vibrational) spectra. Such problems could be more severe for vector and tensor properties. The H₂O case described above is a simple example of many high symmetry points in realistic molecular systems. There exists no guarantee that Eqs. (56) and (59), or Eqs. (60) and (61), offer a complete representation. While Eqs. (56) and (59) of the Deep Potential scheme is rather naive, Eqs. (60) and (61) are more useful but will also fail when higher symmetry is encountered. It is expected that Eqs. (56)

and (59) or Eqs. (60) and (61) are not as accurate as the permutation invariant polynomial NN and fundamental invariant NN^{70,72} (which are much more computationally expensive) for fitting the *ab initio* PES for quantum reactive scattering calculations.¹⁰³ Despite that such problems may probably be alleviated in condensed phase systems, the influence of high symmetry points and other demanding characteristics can sometime become significant, especially when properties such as nonlinear response functions are studied. The electronic polarizability is an example. That the peak position of the high-frequency O–H stretching band of the Raman spectrum produced by DPMD in Ref. 73 is in good agreement with experiment simply suggests the failure of either the Deep Potential fitting scheme or the empirical functional of DFT used in the literature in describing the high-frequency vibrational motion. The peak of the high-frequency stretching band yielded by MD should always be blue-shifted from the exact value (i.e., the experiment value), as well explained in Refs. 14, 15, and 93. Any triumph claimed based on reproducing the high-frequency band of the experimental vibrational IR or Raman spectrum with MD simulations only indicates over-fitting of the functional or/and PES in the wrong direction.

It will probably be useful to develop a unified formulation of the Deep Potential scalar descriptor with the hierarchy of expansion,

$$\overleftrightarrow{D}_i^l = \sum_{n=2}^{N_i} \tilde{D}_i^{(n),l}, \quad (C5)$$

where N_i is the total number of neighbor atoms for atom i and the n -neighbor term for the environment of atom i is defined by

$$\tilde{D}_i^{(n),l} = \sum_{j_1, j_2, \dots, j_n} \tilde{G}_{\alpha_1, \alpha_2, \dots, \alpha_n}^{(n),l} (\{s(r_{ijk})\}, \{\theta_{ijk,jm}\}) \quad (C6)$$

$$\cdot \left(\sum_{\substack{\text{all pair combinations} \\ \text{in the } n \text{ neighbors}}} \mathbf{r}_{ijk}^T \mathbf{r}_{ijm} \right), \quad (C7)$$

with $r_{ijk} = |\mathbf{r}_{ijk}|$ the distance between atom i and atom j_k , and $\theta_{ijk,jm} \equiv \arccos(\mathbf{r}_{ijk}^T \mathbf{r}_{ijm} / (r_{ijk} r_{ijm}))$. The n -neighbour local embedding function $\tilde{G}^{(n)}$ is an NN function, which takes $n + \frac{n(n-1)}{2}$ inputs ($\{s(r_{ijk})\}$ and $\{\theta_{ijk,jm}\}$) and yields F outputs and whose parameters are dependent on the types of atom i , atom j_1 , atom j_2, \dots , and atom j_n .

Similarly, the hierarchy of expansion for the second-order tensor descriptor reads

$$\overleftrightarrow{D}_i^l = \sum_{n=2}^{N_i} \overleftrightarrow{D}_i^{(n),l}, \quad (C8)$$

with

$$\overleftrightarrow{D}_i^{(n),l} = \sum_{j_1, j_2, \dots, j_n} \tilde{G}_{\alpha_1, \alpha_2, \dots, \alpha_n}^{(n),l} (\{s(r_{ijk})\}, \{\theta_{ijk,jm}\}) \quad (C9)$$

$$\cdot \left(\sum_{\substack{\text{all pair combinations} \\ \text{in the } n \text{ neighbors}}} (\mathbf{r}_{ijm} \mathbf{r}_{ijk}^T + \mathbf{r}_{ijk} \mathbf{r}_{ijm}^T) \right). \quad (C10)$$

While the two-neighbor terms $\tilde{D}_i^{(2),l}$ for the scalar and $\overset{\leftrightarrow}{D}_i^{(2),l}$ for the second-order tensor can be equivalent to Eqs. (60) and (61), respectively, many-neighbor terms take care of higher symmetry points and many-body effects. Truncation in Eq. (C5) or Eq. (C8) often leads to higher numerical efficiency as long as convergence of error is accepted.

In addition to Eqs. (C5) and (C8), the Deep Potential descriptor for higher-order tensor features of atoms can be constructed in an analogous fashion.

DATA AVAILABILITY

The data that support the findings of this study are available within the article and its [supplementary material](#).

REFERENCES

- ¹ W. H. Miller, "Quantum dynamics of complex molecular systems," *Proc. Natl. Acad. Sci. U. S. A.* **102**, 6660–6664 (2005).
- ² W. H. Miller, "Including quantum effects in the dynamics of complex (i.e., large) molecular systems," *J. Chem. Phys.* **125**, 132305 (2006).
- ³ N. Makri, A. Nakayama, and N. J. Wright, "Forward-backward semiclassical simulation of dynamical properties in liquids," *J. Theor. Comput. Chem.* **03**, 391–417 (2004).
- ⁴ G. A. Voth, "Path integral centroid methods in quantum statistical mechanics and dynamics," *Adv. Chem. Phys.* **93**, 135 (1996).
- ⁵ S. Habershon, D. E. Manolopoulos, T. E. Markland, and T. F. Miller III, "Ring-polymer molecular dynamics: Quantum effects in chemical dynamics from classical trajectories in an extended phase space," *Annu. Rev. Phys. Chem.* **64**, 387–413 (2013).
- ⁶ J. Liu, "Recent advances in the linearized semiclassical initial value representation/classical Wigner model for the thermal correlation function," *Int. J. Quantum Chem.* **115**, 657–670 (2015).
- ⁷ B. J. Berne and G. D. Harp, "On the calculation of time correlation functions," *Adv. Chem. Phys.* **17**, 63–227 (1970).
- ⁸ R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, New York, 2001).
- ⁹ D. A. McQuarrie, *Statistical Mechanics* (University Science Books, Sausalito, CA, 2000).
- ¹⁰ J. Liu, B. J. Alder, and W. H. Miller, "A semiclassical study of the thermal conductivity of low temperature liquids," *J. Chem. Phys.* **135**, 114105 (2011).
- ¹¹ W. H. Miller, S. D. Schwartz, and J. W. Tromp, "Quantum-mechanical rate constants for bimolecular reactions," *J. Chem. Phys.* **79**, 4889–4898 (1983).
- ¹² J. Liu and W. H. Miller, "A simple model for the treatment of imaginary frequencies in chemical reaction rates and molecular liquids," *J. Chem. Phys.* **131**, 074113 (2009).
- ¹³ Q. Shi and E. Geva, "A relationship between semiclassical and centroid correlation functions," *J. Chem. Phys.* **118**, 8173–8184 (2003).
- ¹⁴ J. Liu, W. H. Miller, G. S. Fanourgakis, S. S. Xantheas, S. Imoto, and S. Saito, "Insights in quantum dynamical effects in the infrared spectroscopy of liquid water from a semiclassical study with an *ab initio*-based flexible and polarizable force field," *J. Chem. Phys.* **135**, 244503 (2011).
- ¹⁵ X. Liu and J. Liu, "Critical role of quantum dynamical effects in the Raman spectroscopy of liquid water," *Mol. Phys.* **116**, 755–779 (2018).
- ¹⁶ F. Paesani, S. S. Xantheas, and G. A. Voth, "Infrared spectroscopy and hydrogen-bond dynamics of liquid water from centroid molecular dynamics with an *ab initio*-based force field," *J. Phys. Chem. B* **113**, 13118–13130 (2009).
- ¹⁷ A. Witt, S. D. Ivanov, M. Shiga, H. Forbert, and D. Marx, "On the applicability of centroid and ring polymer path integral molecular dynamics for vibrational spectroscopy," *J. Chem. Phys.* **130**, 194510 (2009).
- ¹⁸ M. Rossi, M. Ceriotti, and D. E. Manolopoulos, "How to remove the spurious resonances from ring polymer molecular dynamics," *J. Chem. Phys.* **140**, 234116 (2014).
- ¹⁹ G. Trenins, M. J. Willatt, and S. C. Althorpe, "Path-integral dynamics of water using curvilinear centroids," *J. Chem. Phys.* **151**, 054109 (2019).
- ²⁰ J. Liu and W. H. Miller, "Real time correlation function in a single phase space integral beyond the linearized semiclassical initial value representation," *J. Chem. Phys.* **126**, 234110 (2007).
- ²¹ J. Liu and W. H. Miller, "An approach for generating trajectory-based dynamics which conserves the canonical distribution in the phase space formulation of quantum mechanics. I. Theories," *J. Chem. Phys.* **134**, 104101 (2011).
- ²² J. Liu and W. H. Miller, "An approach for generating trajectory-based dynamics which conserves the canonical distribution in the phase space formulation of quantum mechanics. II. Thermal correlation functions," *J. Chem. Phys.* **134**, 104102 (2011).
- ²³ J. Liu, "Two more approaches for generating trajectory-based dynamics which conserves the canonical distribution in the phase space formulation of quantum mechanics," *J. Chem. Phys.* **134**, 194110 (2011).
- ²⁴ J. Liu, "Path integral Liouville dynamics for thermal equilibrium systems," *J. Chem. Phys.* **140**, 224107 (2014).
- ²⁵ J. Liu and Z. Zhang, "Path integral Liouville dynamics: Applications to infrared spectra of OH, water, ammonia, and methane," *J. Chem. Phys.* **144**, 034307 (2016).
- ²⁶ Z. Zhang, Z. Chen, and J. Liu, "Path integral Liouville dynamics simulations of vibrational spectra of formaldehyde and hydrogen peroxide," *Chin. J. Chem. Phys.* **33**, 613–622 (2020).
- ²⁷ R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II: Nonequilibrium Statistical Mechanics*, 2nd ed. (Springer-Verlag, Heidelberg, 1991), Vol. 2.
- ²⁸ A. Horikoshi and K. Kinugawa, "Effective potential analytic continuation approach for real time quantum correlation functions involving nonlinear operators," *J. Chem. Phys.* **122**, 174104 (2005).
- ²⁹ Q. Shi and E. Geva, "On the calculation of vibrational energy relaxation rate constants from centroid molecular dynamics simulations," *J. Chem. Phys.* **119**, 9030–9046 (2003).
- ³⁰ D. R. Reichman, P.-N. Roy, S. Jang, and G. A. Voth, "A Feynman path centroid dynamics approach for the computation of time correlation functions involving nonlinear operators," *J. Chem. Phys.* **113**, 919–929 (2000).
- ³¹ T. D. Hone and G. A. Voth, "A centroid molecular dynamics study of liquid *para*-hydrogen and *ortho*-deuterium," *J. Chem. Phys.* **121**, 6412–6422 (2004).
- ³² F. Paesani and G. A. Voth, "Nonlinear quantum time correlation functions from centroid molecular dynamics and the maximum entropy method," *J. Chem. Phys.* **129**, 194113 (2008).
- ³³ S. Habershon, B. J. Braams, and D. E. Manolopoulos, "Quantum mechanical correlation functions, maximum entropy analytic continuation, and ring polymer molecular dynamics," *J. Chem. Phys.* **127**, 174108 (2007).
- ³⁴ I. R. Craig and D. E. Manolopoulos, "Inelastic neutron scattering from liquid *para*-hydrogen by ring polymer molecular dynamics," *Chem. Phys.* **322**, 236–246 (2006).
- ³⁵ R. Giachetti and V. Tognetti, "Variational approach to quantum statistical mechanics of nonlinear systems with application to sine-Gordon chains," *Phys. Rev. Lett.* **55**, 912–915 (1985).
- ³⁶ R. P. Feynman and H. Kleinert, "Effective classical partition-functions," *Phys. Rev. A* **34**, 5080–5084 (1986).
- ³⁷ A. Cuccoli, R. Giachetti, V. Tognetti, R. Vaia, and P. Verrucchi, "The effective potential and effective Hamiltonian in quantum statistical mechanics," *J. Phys.: Condens. Matter* **7**, 7891–7938 (1995).
- ³⁸ J. A. Poulsen, G. Nyman, and P. J. Rossky, "Practical evaluation of condensed phase quantum correlation functions: A Feynman-Kleinert variational linearized path integral method," *J. Chem. Phys.* **119**, 12179–12193 (2003).
- ³⁹ J. Liu and W. H. Miller, "Using the thermal Gaussian approximation for the Boltzmann operator in semiclassical initial value time correlation functions," *J. Chem. Phys.* **125**, 224104 (2006).
- ⁴⁰ J. Shao and E. Pollak, "A new time evolving Gaussian series representation of the imaginary time propagator," *J. Chem. Phys.* **125**, 133502 (2006).

- ⁴¹P. A. Frantsuzov and V. A. Mandelshtam, "Quantum statistical mechanics with Gaussians: Equilibrium properties of van der Waals clusters," *J. Chem. Phys.* **121**, 9247–9256 (2004).
- ⁴²I. Georgescu and V. A. Mandelshtam, "Molecular dynamics with quantum fluctuations," *Phys. Rev. B* **82**, 094305 (2010).
- ⁴³K. K. G. Smith, J. A. Poulsen, G. Nyman, and P. J. Rossky, "A new class of ensemble conserving algorithms for approximate quantum dynamics: Theoretical formulation and model problems," *J. Chem. Phys.* **142**, 244112 (2015).
- ⁴⁴T. J. H. Hele, M. J. Willatt, A. Muolo, and S. C. Althorpe, "Boltzmann-conserving classical dynamics in quantum time-correlation functions: Matsubara dynamics," *J. Chem. Phys.* **142**, 134103 (2015).
- ⁴⁵J. Han, L. Zhang, R. Car, and W. E, "Deep potential: A general representation of a many-body potential energy surface," *Commun. Comput. Phys.* **23**, 629–639 (2018).
- ⁴⁶L. Zhang, J. Han, H. Wang, R. Car, and W. E, "Deep potential molecular dynamics: A scalable model with the accuracy of quantum mechanics," *Phys. Rev. Lett.* **120**, 143001 (2018).
- ⁴⁷L. Zhang, J. Han, H. Wang, W. Saidi, R. Car, and W. E, "End-to-end symmetry preserving inter-atomic potential energy model for finite and extended systems," in *Advances in Neural Information Processing Systems 31*, edited by S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (Curran Associates, Inc., 2018), pp. 4441–4451.
- ⁴⁸L. Zhang, D. Lin, H. Wang, R. Car, and W. E, "Active learning of uniformly accurate interatomic potentials for materials simulation," *Phys. Rev. Mater.* **3**, 023804 (2019).
- ⁴⁹X. He, Z. Gong, B. Wu, and J. Liu, "Negative zero-point-energy parameter in the Meyer-Miller mapping model for nonadiabatic dynamics," *J. Phys. Chem. Lett.* **12**, 2496–2501 (2021).
- ⁵⁰L. Cohen, "Generalized phase-space distribution functions," *J. Math. Phys.* **7**, 781–786 (1966).
- ⁵¹E. Wigner, "On the quantum correction for thermodynamic equilibrium," *Phys. Rev.* **40**, 749 (1932).
- ⁵²E. P. Wigner, *Perspectives in Quantum Theory* (MIT Press, Cambridge, 1971).
- ⁵³J. E. Moyal, "Quantum mechanics as a statistical theory," *Proc. Cambridge Philos. Soc.* **45**, 99–124 (1949).
- ⁵⁴J. Moix, E. Pollak, and J. Shao, "Generalized Liouville time-dependent perturbation theory," *Phys. Rev. A* **80**, 052103 (2009).
- ⁵⁵H. Wang, X. Sun, and W. H. Miller, "Semiclassical approximations for the calculation of thermal rate constants for chemical reactions in complex molecular systems," *J. Chem. Phys.* **108**, 9726–9736 (1998).
- ⁵⁶X. Sun, H. Wang, and W. H. Miller, "Semiclassical theory of electronically nonadiabatic dynamics: Results of a linearized approximation to the initial value representation," *J. Chem. Phys.* **109**, 7064–7074 (1998).
- ⁵⁷R. Hernandez and G. A. Voth, "Quantum time correlation functions and classical coherence," *Chem. Phys.* **233**, 243–255 (1998).
- ⁵⁸E. Pollak and J.-L. Liao, "A new quantum transition state theory," *J. Chem. Phys.* **108**, 2733–2743 (1998).
- ⁵⁹Q. Shi and E. Geva, "Semiclassical theory of vibrational energy relaxation in the condensed phase," *J. Phys. Chem. A* **107**, 9059–9069 (2003).
- ⁶⁰J. Liu, D. Li, and X. Liu, "Further study of path integral Liouville dynamics," *Sci. Sin.: Chim.* **46**, 27–37 (2016).
- ⁶¹A. Bose and N. Makri, "Coherent state-based path integral methodology for computing the Wigner phase space distribution," *J. Phys. Chem. A* **123**, 4284 (2019).
- ⁶²S. Bonella, M. Monteferante, C. Pierleoni, and G. Ciccotti, "Path integral based calculations of symmetrized time correlation functions. I," *J. Chem. Phys.* **133**, 164104 (2010).
- ⁶³S. Bonella and G. Ciccotti, "Approximating time-dependent quantum statistical properties," *Entropy* **16**, 86–109 (2013).
- ⁶⁴S. K.-M. Svensson, J. A. Poulsen, and G. Nyman, "Classical Wigner model based on a Feynman path integral open polymer," *J. Chem. Phys.* **152**, 094111 (2020).
- ⁶⁵T. B. Blank, S. D. Brown, A. W. Calhoun, and D. J. Doren, "Neural network models of potential energy surfaces," *J. Chem. Theory Comput.* **103**, 4129–4137 (1995).
- ⁶⁶D. F. R. Brown, M. N. Gibbs, and D. C. Clary, "Combining *ab initio* computations, neural networks, and diffusion Monte Carlo: An efficient method to treat weakly bound molecules," *J. Chem. Phys.* **105**, 7597–7604 (1996).
- ⁶⁷C. M. Handley and P. L. A. Popelier, "Potential energy surfaces fitted by artificial neural networks," *J. Phys. Chem. A* **114**, 3371–3383 (2010).
- ⁶⁸J. Behler, "Neural network potential-energy surfaces in chemistry: A tool for large-scale simulations," *Phys. Chem. Chem. Phys.* **13**, 17930–17955 (2011).
- ⁶⁹J. Chen, X. Xu, X. Xu, and D. H. Zhang, "Communication: An accurate global potential energy surface for the OH + CO → H + CO₂ reaction using neural networks," *J. Chem. Phys.* **138**, 221104 (2013).
- ⁷⁰B. Jiang and H. Guo, "Permutation invariant polynomial neural network approach to fitting potential energy surfaces. III. Molecule-surface interactions," *J. Chem. Phys.* **139**, 054112 (2013).
- ⁷¹S. Manzhos, R. Dawes, and T. Carrington, "Neural network-based approaches for building high dimensional and quantum dynamics-friendly potential energy surfaces," *Int. J. Quantum Chem.* **115**, 1012–1020 (2015).
- ⁷²K. Shao, J. Chen, Z. Zhao, and D. H. Zhang, "Communication: Fitting potential energy surfaces with fundamental invariant neural network," *J. Chem. Phys.* **145**, 071101 (2016).
- ⁷³G. M. Sommers, M. F. Calegari Andrade, L. Zhang, H. Wang, and R. Car, "Raman spectrum and polarizability of liquid water from deep neural networks," *Phys. Chem. Chem. Phys.* **22**, 10592–10602 (2020).
- ⁷⁴M. F. C. Andrade, H. Ko, L. Zhang, R. Car, and A. Selloni, "Free energy of proton transfer at the water-TiO₂ interface from *ab initio* deep potential molecular dynamics," *Chem. Sci.* **11**, 2335–2341 (2020).
- ⁷⁵J. Zeng, L. Cao, M. Xu, T. Zhu, and J. Z. Zhang, "Complex reaction processes in combustion unraveled by neural network-based molecular dynamics simulation," *Nat. Commun.* **11**, 5713 (2020).
- ⁷⁶W.-K. Chen, X.-Y. Liu, W.-H. Fang, P. O. Dral, and G. Cui, "Deep learning for nonadiabatic excited-state dynamics," *J. Phys. Chem. Lett.* **9**, 6702–6708 (2018).
- ⁷⁷L. Zhang, M. Chen, X. Wu, H. Wang, W. E, and R. Car, "Deep neural network for the dielectric response of insulators," *Phys. Rev. B* **102**, 041121 (2020).
- ⁷⁸H.-Y. Ko, L. Zhang, B. Santra, H. Wang, W. E, R. A. DiStasio, Jr., and R. Car, "Isotope effects in liquid water via deep potential molecular dynamics," *Mol. Phys.* **117**, 3269–3281 (2019).
- ⁷⁹F.-Z. Dai, B. Wen, Y. Sun, H. Xiang, and Y. Zhou, "Theoretical prediction on thermal and mechanical properties of high entropy (Zr_{0.2}Hf_{0.2}Ti_{0.2}Nb_{0.2}Ta_{0.2})C by deep learning potential," *J. Mater. Sci. Technol.* **43**, 168–174 (2020).
- ⁸⁰A. Marcolongo, T. Binniger, F. Zipoli, and T. Laino, "Simulating diffusion properties of solid-state electrolytes via a neural network potential: Performance and training scheme," *ChemSystemsChem* **2**, e1900031 (2019).
- ⁸¹H. Wang, X. Guo, L. Zhang, H. Wang, and J. Xue, "Deep learning inter-atomic potential model for accurate irradiation damage simulations," *Appl. Phys. Lett.* **114**, 244101 (2019).
- ⁸²Q. Liu, D. Lu, and M. Chen, "Structure and dynamics of warm dense aluminum: A molecular dynamics study with density functional theory and deep potential," *J. Phys.: Condens. Matter* **32**, 144002 (2020).
- ⁸³L. Bourgeois, Y. Zhang, Z. Zhang, Y. Chen, and N. V. Medhekar, "Transforming solid-state precipitates via excess vacancies," *Nat. Commun.* **11**, 1248 (2020).
- ⁸⁴H. Niu, L. Bonati, P. M. Piaggi, and M. Parrinello, "Ab initio phase diagram and nucleation of gallium," *Nat. Commun.* **11**, 2654 (2020).
- ⁸⁵B. Leimkuhler and C. Matthews, "Rational construction of stochastic numerical methods for molecular sampling," *Appl. Math. Res. Express* **2013**(1), 34–56.
- ⁸⁶L. Hascoet and V. Pascual, "The Tapenade automatic differentiation tool: Principles, model, and specification," *ACM Trans. Math. Software* **39**, 1–43 (2013).
- ⁸⁷J. Liu, D. Li, and X. Liu, "A simple and accurate algorithm for path integral molecular dynamics with the Langevin thermostat," *J. Chem. Phys.* **145**, 024103 (2016).
- ⁸⁸Z. Zhang, X. Liu, Z. Chen, H. Zheng, K. Yan, and J. Liu, "A unified thermostat scheme for efficient configurational sampling for classical/quantum canonical ensembles via molecular dynamics," *J. Chem. Phys.* **147**, 034109 (2017).
- ⁸⁹Z. Zhang, X. Liu, K. Yan, M. E. Tuckerman, and J. Liu, "A unified efficient thermostat scheme for the canonical ensemble with holonomic or isokinetic constraints via molecular dynamics," *J. Phys. Chem. A* **123**, 6056–6079 (2019).

- ⁹⁰N. Grönbeck-Jensen and O. Farago, “A simple and effective Verlet-type algorithm for simulating Langevin dynamics,” *Mol. Phys.* **111**, 983–991 (2013).
- ⁹¹D. Li, X. Han, Y. Chai, C. Wang, Z. Zhang, Z. Chen, J. Liu, and J. Shao, “Stationary state distribution and efficiency analysis of the Langevin equation via real or virtual dynamics,” *J. Chem. Phys.* **147**, 184104 (2017).
- ⁹²K. P. Padma and D. Marx, “Understanding hydrogen scrambling and infrared spectrum of bare CH_5^+ based on *ab initio* simulations,” *Phys. Chem. Chem. Phys.* **8**, 573–586 (2006).
- ⁹³J. Liu, W. H. Miller, F. Paesani, W. Zhang, and D. A. Case, “Quantum dynamical effects in liquid water: A semiclassical study on the diffusion and the infrared absorption spectrum,” *J. Chem. Phys.* **131**, 164509 (2009).
- ⁹⁴S. V. Shirin, O. L. Polyansky, N. F. Zobov, P. Barletta, and J. Tennyson, “Spectroscopically determined potential energy surface of H_2^{16}O up to 25 000 cm^{-1} ,” *J. Chem. Phys.* **118**, 2124 (2003).
- ⁹⁵D. W. Schwenke and H. Partridge, “Convergence testing of the analytic representation of an *ab initio* dipole moment function for water: Improved fitting yields improved intensities,” *J. Chem. Phys.* **113**, 6592–6597 (2000).
- ⁹⁶R. J. Barber, J. Tennyson, G. J. Harris, and R. N. Tolchenov, “A high-accuracy computed water line list,” *Mon. Not. R. Astron. Soc.* **368**, 1087–1094 (2006).
- ⁹⁷J. Koput, S. Carter, and N. C. Handy, “*Ab initio* prediction of the vibrational-rotational energy levels of hydrogen peroxide and its isotopomers,” *J. Chem. Phys.* **115**, 8345–8350 (2001).
- ⁹⁸M. J. Bramley and N. C. Handy, “Efficient calculation of rovibrational eigenstates of sequentially bonded four-atom molecules,” *J. Chem. Phys.* **98**, 1378–1397 (1993).
- ⁹⁹M. Gandolfi, A. Rognoni, C. Aieta, R. Conte, and M. Ceotto, “Machine learning for vibrational spectroscopy via divide-and-conquer semiclassical initial value representation molecular dynamics with application to *N*-methylacetamide,” *J. Chem. Phys.* **153**, 204104 (2020).
- ¹⁰⁰A. Donoso and C. C. Martens, “Quantum tunneling using entangled classical trajectories,” *Phys. Rev. Lett.* **87**, 223202 (2001).
- ¹⁰¹D. T. Colbert and W. H. Miller, “A novel discrete variable representation for quantum mechanical reactive scattering via the S-matrix Kohn method,” *J. Chem. Phys.* **96**, 1982–1991 (1992).
- ¹⁰²J. Crank and P. Nicolson, “A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type,” *Adv. Comput. Math.* **6**, 207–226 (1996).
- ¹⁰³D. H. Zhang and H. Guo, “Recent advances in quantum dynamics of bimolecular reactions,” *Annu. Rev. Phys. Chem.* **67**, 135 (2016).
- ¹⁰⁴J. Koput, S. Carter, and N. C. Handy, “Potential energy surface and vibrational-rotational energy levels of hydrogen peroxide,” *J. Phys. Chem. A* **102**, 6325 (1998).