Efficiently Calculating Anharmonic Frequencies of Molecular Vibration by Molecular Dynamics Trajectory Analysis

Shaoqing Wang*

Shenyang National Laboratory for Materials Science, Institute of Metal Research, Chinese Academy of Sciences, Shenyang 110016, China

ABSTRACT: Two efficient methods, the Eckart frame algorithm and the multiorder derivative algorithm, for vibrational frequency calculation directly based on the raw data of atomic trajectory from the state-of-the-art first-principles molecular dynamics simulation are presented. The Eckart frame approach is robust to retrieve the full set of anharmonic fundamental frequencies of any molecule from the atomic trajectory for a sufficiently long molecular dynamics simulation at a temperature close to 0 K. In addition to the fundamental vibrational frequencies, the multiorder derivative approach is universal for the calculations of vibrational frequencies based on the molecular dynamics result in a wide range of temperatures. The accuracy, efficiency, and applicability of these two methods are demonstrated through several successful examples in calculating the anharmonic fundamental vibrational frequencies of methane, ethylene, water, and cyclobutadiene.

INTRODUCTION

The fundamental frequencies of molecular vibration, often simplified as fundamentals, are exclusively dependent on the geometrical configuration and elemental composition of the molecule, and therefore, are considered as the molecular fingerprint. Information on fundamental frequencies is important to identify molecular species and to determine the molecular structure. The accurate calculation of fundamental frequencies is essential for a reasonable interpretation of the experimental observation of the molecular vibrational spectrum. The methodological development of the vibrational frequency calculation was initiated in the early part of the 20th century for explaining the experimental vibrational spectra of several small molecules. The traditional frequency calculations are typically performed by the force field method with the force field parameters either fitted by experimental data or from ab initio calculations. Because the principle of this realization is dependent on the knowledge of molecular normal vibrational modes, it is often referred to as the normal mode analysis (NMA) method. At present, it has become a routine task to perform fundamental frequency calculations through the NMA method based on the second derivatives (SDs) of the adiabatic potential energy at the local energy minimum evaluated by ab initio calculations. In this way, the calculated frequencies are the harmonic frequencies. The algorithms for NMA and harmonic frequency calculations have been integrated in many of the commercial/non-commercial electronic structure software packages such as GAMESS, GAUSSIAN, HyperChem, MOLPRO, and others. However, the anharmonicity of the potential energy surface, the vibrational mode–mode coupling, and the Coriolis coupling are not taken into consideration in this kind of calculation. The accuracy of such calculations is poor, and subsequent scaling of the calculated frequencies should be applied. The scaling factor is usually a single-parameter empirical value between 0.89 and 0.99, and therefore, leads to a certain degree of uncertainty.

Molecular vibration is the periodic motion of the constituent atoms around their mean positions in a molecule. Each atom oscillates around an equilibrium point in the molecular vibration. The interatomic bonds are non-rigid, may be stretched or compressed, thus have a variable bond-length. The bond angles that are formed between adjacent bonds may also change to cause bond bending, dihedral or torsion motions. This leads to the synergic motions of atoms in the molecule like a spring network, i.e. the intra-molecular atomic vibration at a finite temperature. The geometrical symmetry of a molecule can be described by a specific point group in the group theory, e.g. C₄ᵥ for water, T₄ for methane, and D₉h for benzene. According to the NMA theory, molecular vibration is described by a set of normal vibrational modes. Each of these normal modes belongs to an irreducible representation of a symmetry species in the point group. Each normal mode involves the associated motions of all atoms in the molecule. Based on the harmonic vibration theory, the atomic motion between different vibrational modes must be independent of each other and the potential surface must be a parabola. However, the potential surface is always slightly anharmonic in any real molecular system. This anharmonicity causes the coupling between different vibrational modes and results in anharmonic molecular vibration.

Received: December 1, 2018
Accepted: May 13, 2019
Published: May 28, 2019

DOI: 10.1021/acsomega.8b03364
ACS Omega 2019, 4, 9271–9283
Several first-principles theoretical methods to calculate high resolution vibrational spectra by taking the anharmonic effect into consideration were proposed since the late 20th century, a few of which can sometimes even be compared with the experimental accuracy.16 These methods are basically rooted from the vibrational self-consistent field (VSCF) approach and vibrational perturbation theory (VPT), including the configuration interaction VSCF,27 correlation-corrected VSCF,28 vibrational multi-configuration self-consistent field,29 size-extensive VSCF method,30 vibrational second-order Moller–Plesset perturbation theory (VMP2 or VPT2),31–34 vibrational coupled cluster theory,35,36 vibrational multi-reference coupled cluster theory,37 etc. These high-accuracy methods are only limited to small molecules except VSCF and VPT2. A fully automated code for an anharmonic vibration study with VPT by the second-order Moller–Plesset perturbation had been developed.34 The VPT2 method had demonstrated its accuracy for anharmonic vibrational frequency calculations for molecular systems consisting of dozens of atoms.38–40 A VSCF approximation with the second-order perturbation correction (VSCF-PT2) algorithm28 had been recently applied to anharmonic vibration calculations for a few large biomolecules with several hundred atoms.41–43 Nevertheless, both VPT2 and VSCF-PT2 suffer the problem of the resonance effect43–45 and several terms in the formalism of the VSCF algorithm show nonphysical size dependence.30 A more general computational method to solve vibrational Schrödinger equations for larger molecules and solids remains to be developed.

All information on molecular vibration is encrypted in the trajectory of the atomic motion evolved over time. Unfortunately, it is not feasible to directly record the atomic trajectory by any means of experiment due to the micro-amplitude and rapid displacement of the atoms in the molecular vibration. Molecular vibration energy is quantized. Molecules can change their vibrational state and interact with the outside world through absorbing or releasing energy that matches their vibrational frequencies. Based on this principle, experimental scientists try to use a variety of ways to measure the amount of energy exchanged by molecular vibration with the outside world to probe the behavior of the molecular vibration. Various spectroscopic techniques were invented owing to these efforts, such as infrared (IR) spectroscopy, ultraviolet spectroscopy,40 Raman spectroscopy, photon fluorescence spectroscopy, cold ion spectroscopy,41 etc. Because it is hardly possible to conduct a comprehensive study on the molecular vibrational behavior only using a single spectroscopic device in most cases, it is always a tough task for experimental study of the vibrational properties even for many well-known molecules. On the other hand, there is no difficulty to record the atomic trajectory in theoretical simulations. The only foreseeable obstacle might just be the limitation on data storage of the computer. The simplest and most straightforward way to reasonably study the anharmonic behavior in the molecular vibration can be done by molecular dynamics (MD) simulations. MD was first developed by Fermi and co-workers in the mid 50s of the last century,46 and successfully applied to simulate the elastic collisions between hard spheres with an IBM 704 computer by Alder and Wainwright in 1960.49 MD is designed to simulate the physical movements of atoms and molecules. The trajectories of atoms and molecules are determined by numerically solving Newton’s equations of motion for the system. The forces between the constituent atoms are calculated either using the empirical interatomic potentials or by directly solving the Schrödinger equation.

Theoretically, the dynamic behavior of atoms in any physical system can be simulated by MD as long as the exact information on the interaction between atoms is known. An empirical interatomic potential function incorporated correctly with non-harmonic and polarization effects is required in the classical MD simulation for vibrational frequency calculations. As long as the potential function is accurate enough, detailed information on molecular vibration can be recorded in the atomic trajectory of a sufficiently long MD simulation. However, such a potential function is very difficult to find for most interested molecules. The first-principles MD (FPMD)48 or ab initio MD methods have emerged in the last three decades to study the dynamic problems of atomic and molecular systems in quantum mechanical precision. FPMD within the framework of density-functional theory (DFT)41,52 deals with the interatomic forces by accurate electronic structure calculation, the effects of charge redistribution, potential energy surface anharmonicity, and polarization can be naturally taken into account in such calculations. The state-of-the-art FPMD offers one of the most appropriate choices for molecular vibration simulations.

The earliest application of computer simulation to molecular vibrational spectra was attempted by Noid and co-workers for anharmonic two- and three-dimensional model systems in 1977.53 Adams et al. studied the vibration of icosahedral C60 by a simplified quantum-molecular-dynamics simulation.54 The thermal vibration of hydrogen on the Si(111) surface was simulated by Car–Parrinello MD (CPMD); the vibration frequencies of Si–H stretching and bending, and Si–Si–H angle bending modes were calculated by the Fourier transformation of the simulated atomic trajectories near the surface.55 Wathelet et al. tested the possibility to calculate the vibrational frequencies of diatomic molecules by the first-principles CPMD simulation.56 The IR absorption spectrum of water was calculated by Fourier transformations of the velocity autocorrelation function57 and the dipole moment autocorrelation function58,59 from MD simulations, respectively. The hydrogen molecule vibration in crystalline silane under high-pressure was simulated by ab initio canonical ensemble MD at 300 K.60 The phenomenon of vibrational spectral diffusion in liquid methanol was captured by the CPMD simulation.61 The Fourier transform of atom-velocity and bond-length autocorrelation functions from the CPMD simulation was carried out to study the molecular vibration of methane in clathrate hydrate.62 Analysis on the electronic excitation state from the simulated trajectory of the Born–Oppenheimer MD (BOMD) simulation was realized.63 On the whole, the reported MD studies are mostly focused on some specific aspects of molecular vibration. The calculation and analysis methods adopted are limited only to certain special purposes. The feasibility of accurately calculating the anharmonic frequencies of all vibrational modes of a molecule directly by the MD trajectory has not been verified. A universal, easy-to-use method for this purpose has yet to be developed.

The Fourier transform provides a primary theoretical tool for spectral analysis. A Fourier transform maps a time series g(t) into a frequency spectrum G(ν) using
\[ G(\nu) = \mathcal{F}(g(t)) = \int_{-\infty}^{\infty} g(t) e^{-2\pi i \nu t} dt \]  

The MD trajectory consists of discrete time series of atomic coordinates. The high-efficient fast Fourier transform (FFT) algorithm is usually used for such discrete Fourier transform. All information about the vibrational frequencies should be obtained in principle through the FFT of a sufficiently long MD trajectory in molecular vibrational simulations. However, the problem that complicates the matter is that the equilibrium positions of atoms in vibration are constantly changing because of the overall rotating, random tilting, and drifting motions of the molecule in a vacuum. The regular atomic vibrational signal and the irregular random signal are mixed together in the MD trajectory. This problem will cause serious noise in the spectrogram by the direct FFT of the raw MD trajectory. As a typical example, the influence of noise on the obtained spectrogram of methane is shown in Figure 1. Figure 1a presents the raw atomic trajectory for the \( y \) coordinate of a hydrogen atom in the molecule between 5 000 000 and 5 200 000 steps in the FPMD simulation. The inset of Figure 1a shows the enlarged trajectory of the small rectangular area to show that the faint vibrational signal is buried in the strong noises. The direct FFT of this raw trajectory is given in Figure 1b. It is seen that the extremely strong noise frequencies occupy the low-frequency end on the left side of the figure. Still, the fundamental frequencies of the normal vibration modes \( \nu_1 \) and \( \nu_3 \) in the vicinity of 3000 cm\(^{-1}\) are distinguishable, and the other two fundamental frequencies, \( \nu_2 \) and \( \nu_4 \), can also be identified in the inset enlarged three hundred times. The pure vibrational trajectory between 5 090 000 and 5 100 000 MD steps of this atom after noise reduction using the Eckart frame algorithm is shown in Figure 1c which is retrieved from the inset area of the raw MD trajectory in Figure 1a. Figure 1a shows that the maximum amplitude of the raw MD trajectory is about 0.6 Å. Figure 1c shows that the maximum amplitude of the pure vibrational signal is less than 0.002 Å. Therefore, the noise is a few hundred times stronger than the pure molecular vibration in the raw MD trajectory. Despite the serious noise, the information on molecular vibration is completely kept in the raw MD trajectory as shown in the inset of Figure 1a and in the Fourier transform spectrogram of Figure 1b. Evidently, the calculation of highly accurate molecular vibrational frequencies from the MD trajectory depends on powerful noise reduction and signal enhancement techniques.

Although the MD simulation is one of the most natural and direct methods to study molecular vibration theoretically, it is usually still rather difficult to directly retrieve the vibrational information from the raw data of the MD trajectory because of the serious noise interference. The great potential of MD simulation in molecular vibration research has not been fully demonstrated. Therefore, it is highly necessary to find efficient numerical algorithms for MD trajectory analysis. The purpose of the present study is to develop accurate, robust, and easy-to-use methods for retrieving molecular vibrational information based on the result of FPMD simulations.

\section*{Algorithms for Molecular Vibrational Frequency Calculation by MD Trajectory Analysis}

Two algorithms, the Eckart frame algorithm and the multiorder derivative (MOD) algorithm, will be developed for retrieving the vibrational information from the raw data of the MD trajectory through noise reduction and signal enhancement in the following two subsections.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{image1.png}
\caption{Spectrogram calculation based on the MD trajectory. (a) The raw MD trajectory of the \( y \)-coordinate of a hydrogen atom in methane by FPMD simulations. The trajectory in the small rectangular frame area is enlarged in the inset which shows that the faint vibrational signal is buried in strong noise. (b) FFT of the raw MD trajectory in (a). (c) The pure vibrational trajectory of the same atom after noise reduction using the Eckart frame algorithm.}
\end{figure}
that there are serious noises in the MD trajectory resulted from the rotation, the random drift, and tilting motions of the molecule as a whole. In order to eliminate these noises, the algorithm for vibrational frequency calculation using the Eckart molecule-fixed frame includes the following steps.

First of all, the random drift motion can be eliminated by the translational transformation of the coordinate system through always setting the mass center of the molecule as the origin of the coordinate system in each MD step. For a molecule with \( n \) atoms assigned by coordinates \( \vec{r}_i \) and masses \( m_i \) \((i = 1...n)\), its mass center is given by

\[
\vec{r}_c = \frac{\sum_i m_i \vec{r}_i}{\sum_i m_i}
\]  

(2)

Secondly, a molecule-fixed Eckart frame coordinate system is adopted to remove the effect of global tilting and rotation of the molecule. Eckart proposed a moving Cartesian coordinate system to separate the rotational and vibrational behaviors of polyatomic molecules to the maximum extent possible, i.e., the Eckart frame.\(^6\) The instantaneous orientation of the Eckart frame changes with the instantaneous coordinate vectors of the atoms in the molecule at the moment. All atoms look seem to be undergoing the vibration motion around their equilibrium positions in the Eckart frame. The condition of the Eckart frame is\(^{6,4-66}\)

\[
\sum_{i=1}^{n} m_i (\vec{r}_i^e \times \vec{r}_i^e) = 0
\]

(3)

Here, \( \vec{r}_i^e \) is the initial coordinate vector of the \( i \)th atom at equilibrium

\[
\vec{r}_i^e = \sum_{j=1}^{3} x_{ij} \vec{a}_j
\]

(4)

\( \vec{a}_1, \vec{a}_2, \) and \( \vec{a}_3 \) are the Cartesian basic vectors of the real space \( \mathbb{R}^3 \). The instantaneous coordinate vector \( \vec{r}_i(t) \) of the \( i \)th atom in real space \( \mathbb{R}^3 \) is

\[
\vec{r}_i(t) = \sum_{j=1}^{3} x_{ij}(t) \vec{a}_j
\]

(5)

The instantaneous coordinate vector \( \vec{r}_i^e(t) \) of the \( i \)th atom in the Eckart frame at time \( t \) is given by

\[
\vec{r}_i^e(t) = \sum_{j=1}^{3} x_{ij}^e(t) \vec{c}_j(t)
\]

(6)

\( \vec{c}_1(t), \vec{c}_2(t), \) and \( \vec{c}_3(t) \) are the three unitary Eckart vectors\(^{65}\) that define the instantaneous Eckart frame. The relationship between \( \vec{r}_i^e(t) \) and \( \vec{r}_i(t) \) is

\[
\begin{pmatrix}
    x_{11}^e \\
    x_{12}^e \\
    x_{13}^e
\end{pmatrix} = U
\begin{pmatrix}
    x_{11} \\
    x_{12} \\
    x_{13}
\end{pmatrix}
\]

(7)

For the sake of brevity, the time dependence of the atomic coordinate will be hidden hereafter. \( U \) is the direction-cosine matrix for the coordinate transformation from the \( \mathbb{R}^3 \) real space to Eckart space. It is expressed in quaternion \( Q = (q_1, q_2, q_3, q_4) \) algebra by\(^{67}\)

\[
U = \begin{pmatrix}
    q_1^2 + q_2^2 - q_3^2 - q_4^2 & 2(q_2q_3 + q_4q_1) & 2(q_3q_4 - q_2q_1) \\
    2(q_2q_3 - q_4q_1) & q_1^2 + q_2^2 - q_3^2 - q_4^2 & 2(q_1q_4 + q_3q_2) \\
    2(q_3q_4 + q_2q_1) & 2(q_1q_4 - q_3q_2) & q_1^2 + q_2^2 - q_3^2 - q_4^2
\end{pmatrix}
\]

(8)

with \( |Q|^2 = 1 \). Through minimizing the sum of mass-weighted squared deviations of the atomic coordinates of the displaced and the equilibrium configurations, it was deduced that \( Q \) is the eigenvector with the minimal eigenvalue of the following symmetrical matrix \( C \)

\[
C = \begin{pmatrix}
    c_{11} & c_{12} & c_{13} & c_{14} \\
    c_{12} & c_{22} & c_{23} & c_{24} \\
    c_{13} & c_{23} & c_{33} & c_{34} \\
    c_{14} & c_{24} & c_{34} & c_{44}
\end{pmatrix}
\]

(9)

The matrix elements of \( C \) are calculated using

\[
\begin{align*}
    c_{11} &= \sum_{i=1}^{n} m_i (\xi_{1i}^2 + \xi_{2i}^2 + \xi_{3i}^2), \\
    c_{12} &= \sum_{i=1}^{n} m_i (\xi_{1i} \xi_{2i} - \xi_{3i} \xi_{4i}), \\
    c_{13} &= \sum_{i=1}^{n} m_i (\xi_{1i} \xi_{3i} - \xi_{2i} \xi_{4i}), \\
    c_{14} &= \sum_{i=1}^{n} m_i (\xi_{1i} \xi_{4i} - \xi_{2i} \xi_{3i}), \\
    c_{22} &= \sum_{i=1}^{n} m_i (\xi_{2i}^2 + \xi_{3i}^2 + \xi_{4i}^2), \\
    c_{23} &= \sum_{i=1}^{n} m_i (\xi_{2i} \xi_{3i} - \xi_{4i} \xi_{1i}), \\
    c_{24} &= \sum_{i=1}^{n} m_i (\xi_{2i} \xi_{4i} - \xi_{3i} \xi_{1i}), \\
    c_{33} &= \sum_{i=1}^{n} m_i (\xi_{3i}^2 + \xi_{2i}^2 + \xi_{4i}^2), \\
    c_{34} &= \sum_{i=1}^{n} m_i (\xi_{3i} \xi_{4i} - \xi_{2i} \xi_{1i}), \\
    c_{44} &= \sum_{i=1}^{n} m_i (\xi_{4i}^2 + \xi_{2i}^2 + \xi_{3i}^2)
\end{align*}
\]

(10)

Here, the \( \xi \) and \( \zeta \) variables are defined as

\[
\xi_{ij} = x_{ij}^e + x_{ij},
\]

(11)

\[
\zeta_{ij} = x_{ij}^e - x_{ij}
\]

(12)

The physics behind this algorithm is that the Eckart frame rotates together with the global rotation of the molecule; thus, the atoms seem to engage only in the vibrational motion in the instantaneous Eckart coordinate system. In this way, the details of molecular vibration are completely kept in the time series \( \vec{r}_i^e(t) \) \((i = 1...n)\), and all information on the global molecule rotation is kept in the time series of the Eckart frame, \( \vec{c}_i(t) \) \((j = 1...3)\).

Finally, it is a simple task by the FFT of these atomic trajectories in the Eckart frame to calculate the vibrational frequencies of the molecule as soon as the time series \( \vec{r}_i^e(t) \) \((i = 1...n)\) is ready. The spectrograms for each of the three coordinate components of the \( i \)th atom are calculated using

\[
F_{11}^i(k) = \sum_{m=0}^{N-1} x_{ijm}^e(t)e^{-2\pi imk/N},
\]

(13)

where \( \Delta t \) is the timestep in the MD simulation, \( N \) represents the total MD steps, and \( i \) in the exponent is the imaginary unit. The spectrogram of the \( i \)th atom is given by
\[ F_l^2(\nu) = \frac{3}{\sum_{j=1}^{3} (F_j^0(\nu))^2} \]  

(14)

**Calculation of Molecular Vibrational Frequencies using the MOD Algorithm.** The Eckart frame algorithm presented in the former subsection enables one to effectively separate vibration from rotational motion through intensely compressing the rotational effect. Because the fundamental vibrational modes of a molecule are only related to the pure vibrational motion of atoms, the method is especially suitable for fundamental frequency calculations. However, the algorithm may fail at high temperature when the strict Eckart frame is hard to find due to the serious molecular geometry distortion. The Eckart frame algorithm is also not applicable to analyze the rovibrational spectrum caused by the interactive coupling between the rotational and vibrational motions. An alternate solution will be presented in this subsection.

Vibration is the periodic motion of an object around its equilibrium position. In addition to the periodic change of the atomic coordinate, the velocity, acceleration, and all the MODs of the atomic coordinate trajectory also periodically vary with the same frequency as the atomic coordinate in molecular vibration. Unlike the constant drift of the atomic equilibrium position in the MD simulation of an isolated molecule, the equilibrium points of velocity, acceleration, and any MODs of the atomic coordinate trajectory are always zero. Hence, the problem of the equilibrium point drift of the vibratory variable in molecular vibration can be solved effectively. For this reason, the FFT spectrum of the MOD of the MD trajectory should provide more accurate information on the vibrational frequency than the atomic coordinate trajectory. Mathematically, the MD trajectory is the superposition of the vibrational motions, non-vibrational motions, and random noises. The coordinates of all atoms in the molecule can be written as

\[ \bar{r}_i(t) = \bar{v}_i(t) + \bar{\mu}_i(t), \quad (i = 1 \ldots n) \]  

(15)

where \( \bar{r}_i(t) \) is the contribution of the vibrational motion to \( \bar{r}_i(t) \). Its three components can be written as

\[ v_i^j(t) = \sum_{j=1}^{M} A_i^j(i) \sin(2\pi \nu_j t + \varphi_j) + \omega_i^j(t) + \varphi_i^j(t), \]

\[ (i = 1 \ldots n, \ k = 1 \ldots 3) \]  

(16)

\( M \) is the total number of vibrational modes, \( \omega_i \) is the angular frequency of vibrational mode \( j \), and \( \varphi_j \) is the phase difference of the mode. A strong anharmonic vibration gives rise to frequency multiplication by overtones \( \omega_i^j(t) \), and combination bands \( \omega_i^j(t) \), upon the fundamental vibrational motions to the system. \( \bar{\mu}_i(t) \) includes all non-vibrational motions and noises. This item represents the random thermal fluctuation of the system, its lack of periodicity. Because NVE ensemble is adopted, no periodic noise source is involved in the present FPMD simulation. Thermal fluctuation leads to an irregular drift and tilt of the molecule in a short period of time. All of these aperiodic noises can be described using a polynomial expansion function of time \( t \)

\[ \mu_i^k(t) = c_0^k + c_1^k t + c_2^k t^2 + c_3^k t^3 + \ldots, \]

\[ (i = 1 \ldots n, \ k = 1 \ldots 3) \]  

(17)

Based on eqs 15–17, a noise reduction algorithm by the successive derivative of the atomic coordinate trajectory with respect to \( t \) is suggested

\[ \frac{dr^k(t)}{dt} = \sum_{j=1}^{M} 2\pi \nu_j A_i^j(i) \cos(2\pi \nu_j t + \varphi_j) + c_1^k + 2c_2^k t \]

\[ + 3c_3^k t^2 + \ldots \]  

(18)

\[ \frac{d^2r^k(t)}{dt^2} = -\sum_{j=1}^{M} (2\pi \nu_j)^2 A_i^j(i) \sin(2\pi \nu_j t + \varphi_j) + 2c_2^k \]

\[ + 6c_3^k t + \ldots \]  

(19)

eqs 18 and 19 show that the derivative operation will change \( \pi/2 \) in the phase angle and enhance the amplitude by \( 2\pi \nu_j \) to the periodic part of the derivative function. All information on the vibration frequencies will be retained. As for those non-vibrational signals and noises, the first derivative of \( \bar{r}_i(t) \) will remove the signal by transient random jumps of the molecule. The SD eliminates the effect of global molecule translation at a steady speed, and the third derivative removes the irregular accelerating and tilting motions, etc. It is seen that the net effect of derivative operations will lead to the significant enhancement of the vibrational signal/noise ratio in the derivative functions. Therefore, a much clean molecular vibrational spectrogram can be recorded through the Fourier transform of these MOD functions of the MD trajectory.

The numerical five-point central difference equations\(^{68}\) were used for MOD operation in this study

\[ \frac{dr^k(t)}{dt} \approx \frac{r^k(t + 2\Delta t) - 8r^k(t) + 8r^k(t - \Delta t) - r^k(t - 2\Delta t)}{12\Delta t} \]

\[ \frac{d^2r^k(t)}{dt^2} \approx \frac{-r^k(t - 2\Delta t) + 16r^k(t - \Delta t) - 30r^k(t) + 16r^k(t + \Delta t) - r^k(t + 2\Delta t)}{12(\Delta t)^2} \]

\[ \frac{d^3r^k(t)}{dt^3} \approx \frac{-r^k(t - 2\Delta t) + 2r^k(t - \Delta t) - 2r^k(t + \Delta t) + r^k(t + 2\Delta t)}{2(\Delta t)^3} \]

\[ \frac{d^4r^k(t)}{dt^4} \approx \frac{r^k(t - 2\Delta t) - 4r^k(t - \Delta t) + 6r^k(t) - 4r^k(t + \Delta t) + r^k(t + 2\Delta t)}{(\Delta t)^4} \]

(20)
The vibrational spectrogram is calculated using
\[ G(\gamma) = \left( \sum_{k=1}^{3} \left( \int \frac{d^3 r_i(t)}{dt^2} \right)^2 \right)^{1/2}, \]
\[ (k = 1 \ldots 3, \ i = 1 \ldots n, \ l = 0, 1, 2, 3, \ldots) \] (21)

By this algorithm, all information related to the vibrational frequencies will be preserved.

### DETAILS OF THE FPMD SIMULATION

All FPMD simulations in this study were carried out using the CP2K code\(^{69}\), which is an open-source software package for atomic and molecular simulations. The quickstep algorithm\(^{70}\) by Gaussian and plane wave formalism, using a dual basis of atom centered Gaussian orbitals and plane waves for regular grids, is implemented for solving DFT Kohn–Sham equations in the code. FPMD simulations were performed using the quickstep module of CP2K in the scheme of BOMD. The double-zeta valence polarized Gaussian basis sets optimized for the Goedecker–Teter–Hutter pseudopotentials\(^{51}\) and the Pade exchange–correlation functional in the local density approximation (LDA) were chosen for C and H elements in methane, cyclobutadiene, and ethylene calculations. The triple-zeta valence Gaussian basis sets with two sets of polarization functions optimized for GTH pseudopotentials and the Pade LDA exchange–correlation functional were used for H and O elements in H\(_2\)O molecule calculations. A 280 Ry cutoff for the finest grid level of plane waves was used for GTH pseudopotentials. Orbital transformation methods by the conjugate gradient and direct inversion in the iterative subspace were adopted for wave function optimization in molecular geometry optimization and MD simulation, respectively.\(^{50}\)

The supercell modeling method is adopted to build the structure model for the FPMD simulation; the molecule is placed in the center of a vacuum cubic box. The edges of these boxes are set to 20 Å for ethylene and cyclobutadiene, 15 Å for water, and 10 Å for methane, respectively. Both periodic and aperiodic boundary conditions were tested. The Kohn–Sham equation under aperiodic boundary conditions was solved using the wavelet algorithm. The decoupling technique\(^{68}\) was adopted to eliminate the effect of the periodic charge images in the evaluation of the Coulomb interaction for the infinitely replicated periodic system. The vibrational frequencies calculated using these two boundary conditions with the above model settings are almost the same; thus, the reliability of these models for simulating an isolated molecule is verified. Hereafter, all the reported results were simulated using the periodic boundary conditions. As a large displacement from atom’s equilibrium position will cause an overtone phenomenon, the MD simulation at extreme low temperature is preferable to calculate the fundamental frequencies of molecular vibration. The microcanonical ensemble (NVE ensemble), in which the total energy, particle number, and volume of the system remain unchanged during the FPMD simulation, was adopted throughout this study. It was suggested that a timestep shorter than 0.5 fs is essential for accurate spectrum calculations by the MD simulation.\(^{77}\) A variety of timesteps from 0.01 to 0.5 fs were tested for the system stability during the FPMD simulation with the NVE ensemble in this work. It is found that a timestep of 0.1 fs will keep the system energy stable up to a few million steps when the system temperature is close to 0 K, which will be enough for fundamental frequency calculation of small molecules. Before MD simulation, the ground-state equilibrium geometry of the molecule was obtained using the geometry optimization routine in the CP2K code. The convergence criterion of force is set to 1.0 × 10\(^{-4}\) hartree/Bohr. The optimized equilibrium geometries of the molecules in this work are listed in Table 1.

| molecule       | geometry | optimized | experimental | error (%) |
|----------------|----------|-----------|--------------|-----------|
| cyclobutadiene | r\(_{CC}\) | 1.347     | 1.380\(^{\text{a}}\) | -2.4      |
|                | r\(_{CC}\) | 1.571     | 1.560\(^{\text{a}}\) | 0.7       |
|                | r\(_{CH}\) | 1.094     |              |           |
|                | \(\alpha\) | 134.62    |              |           |
|                | \(\alpha\) | 135.38    |              |           |
| ethylene       | r\(_{CC}\) | 1.343     | 1.335\(^{\text{b}}\) | 0.6       |
|                | r\(_{CH}\) | 1.098     | 1.090\(^{\text{b}}\) | 0.7       |
|                | \(\alpha\) | 121.51    | 121.7\(^{\text{c}}\) | -0.2      |
| methane        | r\(_{CH}\) | 1.102     | 1.086\(^{\text{c}}\) | 1.5       |
|                | \(\alpha\) | 109.47    |              |           |
| water          | r\(_{OH}\) | 0.971     | 0.959\(^{\text{b}}\) | 1.3       |
|                | \(\alpha\) | 104.79    | 103.9\(^{\text{b}}\) | 0.9       |

It is seen that the optimized configurations agree well with the experimental reports. Then, the \(\text{NVE FPMD simulation was setup up from the optimized ground-state configuration by slightly displacing the atoms from their equilibrium positions according to the initial temperature. No atomic velocity or temperature rescale was performed throughout the simulation. This realization is essential for simulating the intrinsic vibrational behavior of a molecule under the conditions without outside interference. Starting from 0 K, the final stabilized temperatures of our simulated systems are 0.16, 0.17, 0.35, and 0.12 K for water, ethylene, cyclobutadiene, and methane, respectively, in the fundamental frequency calculations. The closer the optimized molecule configuration is to the ground state, the lower the stabilized system temperature. The molecular vibrational frequencies are calculated using the atomic coordinate trajectories after the system temperature was stabilized. The total time of the MD simulation for spectrum calculations is 0.29, 0.29, 0.24, and 1.80 ns for water, ethylene, cyclobutadiene, and methane, respectively, these correspond to 2.4–18.0 million MD steps. |

### VERIFICATION OF THE NUMERICAL ALGORITHMS FOR CALCULATING MOLECULAR VIBRATIONAL FREQUENCIES

The obtained spectrograms of water, methane, ethylene, and cyclobutadiene in this work are presented in Figures 2–5, respectively. The vibrational frequency calculations were conducted by applying eq 14 to an arbitrarily selected hydrogen atom in these molecules. The quantitative comparisons of the present calculated fundamental frequencies with the published experimental results are given in Tables 2–5. The fundamental frequencies in the “Eckart calc.” column are calculated using the Eckart frame algorithm, the data in the “Error” column gives the relative error between the present calculated and the experimental results. It is seen from Tables 2 and 3 that the calculated fundamental frequencies of water and methane match well with the experimental
measurements. Table 3 shows that the frequencies of $\nu_1$ and $\nu_3$, for the stretching vibrations along the C−H bond of methane, are slightly overestimated; meanwhile, $\nu_2$ and $\nu_4$ for the waging vibrations are underestimated, which means that the present DFT description somewhat enhanced the tension−compression strength of the C−H bond but weakened the bending strength between C−H bonds. The calculated fundamental frequencies of ethylene match closely with the infra-red spectrum analysis by Arnett and Crawford in Table 4. The vibration modes of $\omega_1$, $\omega_9$, and $\omega_{11}$ are typical stretching vibrations along the C−H bond. The calculated frequencies of these modes are all higher than those obtained in the experimental measurements, which is similar to the enhancement of C−H tension−compression strength in methane. The mode of $\omega_2$ is for the stretching vibration between the two carbon atoms in the middle of ethylene, and $\omega_{10}$ is for the waging vibration between the neighboring C−H bonds of the two carbons, respectively. The calculated frequencies of these two modes reveal that the tension−compression enhancing and bending weakening are also true for this molecule. The fundamental frequencies of cyclobutadiene by the SD of MD trajectories, using the same MD trajectories as for the Eckart frame algorithm, are presented in the "SD calc." column in Table 5 for a quantitative comparison with the results from the Eckart frame algorithm. It is seen that both the Eckart frame and MOD algorithms have the same

Figure 2. FFT spectrogram of water using the Eckart frame algorithm. The spectrogram was calculated for an arbitrarily selected hydrogen atom in the molecule.

Figure 3. FFT spectrogram of methane using the Eckart frame algorithm. The spectrogram was calculated for an arbitrarily selected hydrogen atom in the molecule.

Figure 4. FFT spectrogram of ethylene using the Eckart frame algorithm. The spectrogram was calculated for an arbitrarily selected hydrogen atom in the molecule.

Figure 5. FFT spectrogram of cyclobutadiene using the Eckart frame algorithm. The spectrogram was calculated for an arbitrarily selected hydrogen atom in the molecule.

Table 2. Calculated and Experimental Fundamental Frequencies of Water (Unit in cm$^{-1}$)

| band | $\omega_1$ | $\omega_2$ | $\omega_3$ | $\omega_4$ |
|------|--------|--------|--------|--------|
| expt | 3755.92 | 3657.06 | 1594.76 | 1533.68 |
| Eckart | 3819.3 | 3711.1 | 1590.6 | 1310.8 |
| error (%) | 1.7 | 1.5 | -2.1 |

Table 3. Calculated and Experimental Fundamental Frequencies of Methane (Unit in cm$^{-1}$)

| band | $\nu_1$ | $\nu_2$ | $\nu_3$ | $\nu_4$ |
|------|--------|--------|--------|--------|
| symmetry | $A_1$ | $E$ | $F_2$ | $F_2$ |
| Eckart calc. | 2969.8 | 1488.1 | 3097.7 | 1253.6 |
| exp. | 2917.6 | 1533.6 | 3019.7 | 1310.8 |
| error (%) | 1.8 | -3.0 | 2.6 | -4.4 |
Table 4. Calculated and Experimental Fundamental Frequencies of Ethylene (Unit in cm\textsuperscript{-1})

| band | symmetry | \(\omega_1\) | \(\omega_2\) | \(\omega_3\) | \(\omega_4\) | \(\omega_5\) | \(\omega_6\) |
|------|----------|--------------|--------------|--------------|--------------|--------------|--------------|
| Eckart calc. | \(A_{1g}\) | 3076.5 | 1661.6 | 1317.9 | 1017.4 | 3142.2 | 1162.0 |
| exp. | \(A_{1g}\) | 3019.3 | 1623.3 | 1342.4 | 1027.0 | 3272.3 | 1236.0 |
| error (%) | 1.9 | 2.4 | -1.8 | -0.9 | -4.0 | -6.0 |

The data in the SD calc. column are calculated using the second derivative of the MD trajectory.

Table 5. Calculated and Experimental Fundamental Frequencies of Cyclobutadiene (Unit in cm\textsuperscript{-1})

| no. | sym. | exp. | Eckart calc. | SD calc. | err. (%) | active | action |
|-----|------|------|--------------|----------|----------|--------|--------|
| 1   | \(b_{1u}\) | 531.8 | 505.7 | 505.3 | -4.8 | Raman | CH OPLA deform |
| 2   | \(b_{2u}\) | 569.5 | 556.5 | 556.9 | -2.1 | IR | CH OPLA |
| 3   | \(b_{3u}\) | 719.6 | 709.9 | 710.0 | -1.4 | IR | ring deform |
| 4   | \(b_{1g}\) | 723.3 | 827.8 | 827.6 | 14.5 | Raman | ring deform |
| 5   | \(a_{1}\) | 989.5 | 935.0 | 935.1 | -5.5 | Raman | CH deform |
| 6   | \(b_{1u}\) | 1028.6 | 1002.1 | 1002.2 | -2.5 | IR | CH deform |
| 7   | \(a_{g}\) | 1059.8 | 1093.7 | 1093.6 | 3.2 | Raman | CH deform |
| 8   | \(b_{1g}\) | 1121.9 | 1121.8 | 1121.8 | 3.5 | Raman | C–C stretch |
| 9   | \(b_{1u}\) | 1242.6 | 1198.8 | 1198.8 | -3.5 | IR | C–H deform |
| 10  | \(a_{g}\) | 1678.5 | 1611.0 | 1611.1 | -4.0 | Raman | C=C stretch |
| 11  | \(b_{1g}\) | 3093.8 | 3121.4 | 3121.3 | 0.9 | Raman | CH stretch |
| 12  | \(b_{2g}\) | 3105.5,5,6 | 3139.3 | 3139.3 | 1.0 | IR | CH stretch |
| 13  | \(b_{1u}\) | 3124.5,5,6 | 3159.7 | 3159.9 | 1.1 | IR | CH stretch |
| 14  | \(a_{g}\) | 3140.8 | 3172.6 | 3172.7 | 1.0 | Raman | CH stretch |

The data in the SD calc. column are calculated using the second derivative of the MD trajectory.

accuracy for the fundamental frequency calculation of molecular vibration, but the latter approach is simpler in concept and much easy to implement.

In general, it should be enough to obtain the full set of fundamental frequencies by performing FFT of the MD trajectory for any one of the atoms with the Eckart frame algorithm because of the overall synergic movement of all atoms in the molecule. However, exceptions may occur for some molecules with a high structural symmetry. For example, the carbon atom is immobile for the vibration modes with \(A_{1}\) and \(E\) symmetries, but all hydrogen atoms are active for any vibration mode in methane. The trajectory of the lightest atom in the molecule is assumed to carry the most detailed information on the molecular vibration. For some molecules with a high structural symmetry, it is seen that the multiorder derivation operation does not affect the enhancement of the vibrational signal in the MD trajectory.

Figure 6 presents an example to demonstrate the effect of the MOD noise-reduction algorithm for the vibrational spectrum calculation of ethylene. The MD simulation was initialized in the NVE ensemble at 10 K in this case. The system temperature was finally stabilized at 6.77 K. The thick black lines are the raw MD trajectory of the \(x\)-coordinate for a carbon atom in ethylene, and the thin blue lines show the first to fourth derivative curves of the raw MD trajectory in sequence from Figure 6a–d. The amplitudes of the curves for the first to fourth derivatives are enlarged 35.0, 116.0, 233.0, and 437.0 times, respectively, to facilitate direct visual comparison with the raw MD trajectory. The insets in these figures are the FFT spectrograms of the MODs of the raw trajectory. It is seen from Figure 6a that the vibrational signal is greatly enhanced against the noise by the first derivative operation, but the noise interference is still clearly seen in the first derivative curve. The influence of noise is also reflected in the left lower-end of the inset spectrogram near zero cm\textsuperscript{-1}. The noise interference is almost completely removed in the SD curve and in the corresponding spectrogram as shown in Figure 6b. By comparing the derivative curves and the inset spectrograms from Figure 6a–d, it is seen that the multiform derivation has a remarkable effect on the enhancement of the periodic signals, especially the high frequency signals. As is shown in eqs 18 and 19, the derivation operation does not cause any loss of vibrational frequency information other than changing the amplitude of the vibrational signal. The MOD algorithm has a special advantage in the spectral calculation of the weak high-frequency vibrations, for example, the polyad structures of methane in the region from several thousands to tens of thousands cm\textsuperscript{-1}. Because the initial system temperature of the MD simulation is 10 K in this example, the spectrograms shown in the insets of Figure 6 are actually the rovibrational spectra of ethylene, which are reflected from the broadening of the peaks around the fundamental frequencies.

DISCUSSION

Molecular vibration is intrinsically anharmonic because of the anharmonicity of the potential surface in real molecules. FPMD can simulate the vibrational process of real molecules without approximation. The simulation takes into account the anharmonic effect of molecular vibration in the most natural way. The respective scopes of application of the vibrational frequency calculation methods developed in this study, as well as the considerations in the application, are discussed below.

DOI: 10.1021/acs.omega.8b03364
ACS Omega 2019, 4, 9271–9283
Performance Comparison between the Eckart Frame and MOD Algorithms. The Eckart frame algorithm of the vibrational frequency calculation developed in this work enables the separation of the molecular vibration from molecular rotational and random motions to the maximum extent possible. This method is most suitable for calculating the anharmonic fundamental frequencies, which can be conveniently realized to any molecule with the atomic coordinate trajectory simulated by FPMD at a temperature close to 0 K. It is a robust method, and is very suitable for the construction of large scale molecular databases based on the idea of materials genome research.91−93 The MOD algorithm is inspired by the analysis that the MD atomic coordinate trajectory comprises periodic sinusoidal vibrational signs, aperiodic non-vibrational signals, and random noises. The derivative operation specializes in compressing aperiodic signals and intensifying the periodic signals. The higher the derivative order, the more significant the effect is. A visual comparison of these two algorithms is presented in Figure 7.

The FPMD simulations of an isolated methane molecule were performed for three stabilized temperatures \( T_s \) at 0.41, 9.70, and 198.02 K, respectively. It is seen from the left column spectrograms, obtained using the Eckart frame algorithm, in Figure 7a,c,e that there are very sharp peaks at the positions of the fundamental frequencies, but the rovibrational, combination bands, and overtone features are seriously distorted because of the suppression effect of the algorithm on the rotational motion. In contrast, all rovibrational, combination bands, and overtones are gradually resolved with the increase of temperature in the spectrograms calculated using the MOD algorithm as shown in the right column FFT spectrograms of Figure 7. It is seen that the fundamental frequencies can always be determined by the local strongest peaks in the recorded spectrogram at various simulated temperatures with the Eckart frame method. As the simulation temperatures rise, it becomes increasingly difficult to distinguish the fundamental frequencies in the obtained spectrogram using the MOD algorithm. However, the latter spectrogram contains more information on the vibrational property of the real molecule.

Correlation between Calculation Accuracy and the Simulation Time. For a dynamic system experiencing a sufficient long running time, almost all points in its phase space will be visited and all dynamic properties of the system will be included in system’s evolutionary trajectory according to the Ergodic theory in statistical physics. There are two reasons for using a sufficiently long-time trajectory in vibrational frequency calculations. Firstly, the resolution \( \Delta \nu \) (unit in cm\(^{-1}\)) of the FFT spectrogram is depended on the total MD steps \( N \), the MD timestep \( \Delta t \), and the light velocity \( c \) by

\[
\Delta \nu = \frac{1}{Nc\Delta t} \tag{22}
\]

It is seen that the longer the MD trajectory the higher the frequency resolution. For example, the MD trajectories of 330,000 and 3.3 million steps bring about 1.0 and 0.1 cm\(^{-1}\) frequency resolutions, respectively, when the MD timestep is 0.1 fs. Secondly, the MD simulation in this study starts with a small perturbation to the equilibrium molecular configuration. It is most likely that only a few vibrational modes are excited...
at the initial MD steps. The other modes can be gradually excited through modes coupling during the simulation. A sufficiently long-time trajectory is essential to include the information on all vibrational modes. At higher simulated temperatures, this process can be accelerated, but the spectrogram will become too complex to analyse.

Target Atom Selection in the Case of Macromolecules. In the examples of the former section, only an arbitrarily selected hydrogen atom is used for the FFT spectrogram calculations. This target atom choice method in the vibrational spectrogram calculation is feasible for small, single branched, and covalently bonded molecular systems. But, it might be failure for large organic molecules and biomolecular systems that are composed of multiple functional groups with many complex multi-branched, multi-stranded, and weak hydrogen bonding structures in which the interactions between different functional groups or branches are very weak. In such a case, the selection of a light atom from each of the functional groups or branches for molecular vibration analysis will show better results.

Intensity Comparison between the Calculated and Experimental Spectra. The geometrical configuration and elemental composition are uniquely correlated with a specific set of fundamental frequencies. However, the experimental spectrum intensity is closely related to the experimental methods and test conditions. For example, the intensities of the \(a_1\) and \(e_2\) vibration modes of benzene can be detected by Raman spectroscopy, but the intensities of these modes are zero in IR spectroscopy. On the contrary, the intensities of the three benzene \(e_1\) modes can only be observed by IR spectroscopy rather than the Raman spectroscopy. Temperature is also a critical factor of the experimental spectral intensity. Therefore, in addition to the atomic trajectory, additional information on polarizability or dipole moment is

Figure 7. Temperature dependence of the FFT spectrograms obtained for an arbitrarily selected hydrogen atom in methane. The spectrograms in the left (a,c,e) and right (b,d,f) columns are recorded using the Eckart frame algorithm and the SD algorithm, respectively. The features of rovibrational bands (d), combination bands, and overtones (f) are emerged successively as temperature increases in the spectrograms by the latter algorithm.
needed to reproduce the experimental intensities of IR and Raman spectra. Because the purpose of the present study is to develop general methods for the molecular vibrational frequency calculation only based on the MD trajectory, the calculated spectrogram intensity is not suitable for direct comparison with the relevant experimental results.

Information on fundamental vibrational frequencies is important for detection, identification, and characterization of organic and inorganic molecules. Molecular vibrational frequencies carry abundant and subtle information of the static and dynamic properties of the molecule. Theoretically, the successful molecular vibrational simulation can be used as an important criterion for verifying the accuracy and validity of various first-principles theories and methods. Such research works will surely push the development of the first-principles calculation to a new height.

**CONCLUSIONS**

In this paper, two algorithms for calculating molecular vibrational frequencies based on the raw data of the atomic coordinate trajectory by the MD simulation are presented. The Eckart frame approach enables the separation of the vibrational and rotational motions of atoms to the greatest extent. It is robust to calculate the full set of fundamental frequencies of any molecules by taking the anharmonic effect into consideration using the atomic trajectory of a sufficiently long time MD simulation near 0 K. The MOD approach fully takes advantage of the enhancement effect of derivative operation on periodic signals. The higher the order, the better the effect of the high-frequency enhancement is. This approach is generally applicable to the calculation of molecular vibrational frequencies based on the MD trajectory in a wide temperature range. Several successful examples are presented to demonstrate the applicability of these two algorithms and the potential of the state-of-the-art FPMD simulation in the study of molecular vibrational properties.

**AUTHOR INFORMATION**

Corresponding Author
*E-mail: sqwang@imr.ac.cn.

ORCID

Shaoqing Wang: 0000-0002-5102-901X

Notes

The author declares no competing financial interest.

**ACKNOWLEDGMENTS**

This work was supported by the CAS Frontier Science Research Project (no. QYZDJ-SSW-JSC015), the National Key R&D Program of China (no. 2016YFB0701302), and the National Natural Science Foundation of China (no. 51471164). The Special Program for Applied Research on Super Computation of the NSFC-Guangdong Joint Fund (the second phase) is also highly acknowledged. Some of the calculations in this study were done on the Tianhe-II high performance computer system in the National Supercomputer Center in Guangzhou, China.

**REFERENCES**

(1) Wilson, E. B. The normal modes and frequencies of vibration of the regular plane hexagon model of the benzene molecule. *Phys. Rev.* 1934, 45, 706–714.

(2) Wilson, E. B.; Decius, J. C.; Cross, P. C. *Molecular Vibrations*; Dover: New York, 1980.

(3) Hecht, K. T. The vibration-rotation energies of tetrahedral XY4 molecules. *J. Mol. Spectrosc.* 1960, 5, 355–389.

(4) Whitten, D. H. The force field, vibrational frequencies, normal coordinates, infra-red and Raman intensities for benzene. *Philos. Trans. R. Soc. London* 1955, 248, 131–154.

(5) Guo, H.; Karplus, M. Ab initio force field for the planar vibrations of benzene. *J. Chem. Phys.* 1988, 89, 4235–4245.

(6) Bercès, A.; Ziegler, T. The harmonic force field of benzene. A local density functional study. *J. Chem. Phys.* 1993, 98, 4793–4804.

(7) Christiansen, O. Vibrational structure theory: new vibrational wave function methods for calculation of anharmonic vibrational energies and vibrational contributions to molecular properties. *Phys. Chem. Chem. Phys.* 2007, 9, 2942–2953.

(8) Alecu, I. M.; Zheng, J.; Zhao, Y.; Truhlar, D. G. Computational thermochemistry: scale factor databases and scale factors for vibrational frequencies obtained from electronic model chemistries. *J. Chem. Theory Comput.* 2010, 6, 2872–2887.

(9) Ochotzki, D.; Neff, M.; Meier, P.; Pfeiffer, F.; Rauhut, G. Selected aspects concerning the efficient calculation of vibrational spectra beyond the harmonic approximation. *Croat. Chem. Acta* 2012, 85, 379–390.

(10) Carter, S.; Bowman, J. M.; Handy, N. C. Extensions and tests of “multimode”: a code to obtain accurate vibration/rotation energies of many-mode molecules. *Theor. Chem. Acc.* 1998, 100, 191–198.

(11) Yagi, K.; Taketsugu, T.; Hiroa, K.; Gordon, M. S. Direct vibrational self-consistent field method: applications to H2O and H2Co. *J. Chem. Phys.* 2000, 113, 1005–1017.

(12) Hermes, M. R.; Hirata, S. Stochastic many-body perturbation theory for anharmonic molecular vibrations. *J. Chem. Phys.* 2014, 141, 084105–084117.

(13) Fischer, S. A.; Ueltschi, T. W.; El-Khoury, P. Z.; Mifflin, A. L.; Hess, W. P.; Wang, H.-F.; Cramer, C. J.; Govind, N. Infrared and Raman spectroscopy from ab initio molecular dynamics and static normal mode analysis: the C-H region of DNSO as a case study. *J. Phys. Chem. B* 2016, 120, 1429–1436.

(14) Pulay, P.; Fogarasi, G.; Boggs, J. E. Force field, dipole moment derivatives, and vibronic constants of benzene from a combination of experimental and ab initio quantum chemical information. *J. Chem. Phys.* 1981, 74, 3999–4014.

(15) Andersson, M. P.; Uvdal, P. New scale factors for harmonic vibrational frequencies using the B3LYP density functional method with the triple-ζ basis set 6-311+G(d,p). *J. Phys. Chem. A* 2005, 109, 2937–2941.

(16) Gerber, R. B.; Chaban, G. M.; Brauer, B.; Miller, Y. First-principles calculations of anharmonic vibrational spectroscopy of large molecules. In *Theory and Applications of Computational Chemistry: The First Forty Years*; Dykstra, C., Ed.; Elsevier: B. V., 2005; pp 165–194.

(17) Blöino, J.; Biczysko, M.; Crescenzi, O.; Barone, V. Integrated computational approach to vibrationally resolved electron spectra: Anisole as a test case. *J. Chem. Phys.* 2008, 128, 244105–244115.

(18) Rauhut, G.; Pulay, P. Transferable scaling factors for density functional derived vibrational force fields. *J. Phys. Chem.* 1995, 99, 3093–3100.

(19) Roy, T. K.; Carrington, T., Jr.; Gerber, R. B. Approximate first-principles anharmonic calculations of polyatomic spectra using MP2 and B3LYP potentials: comparisons with experiment. *J. Phys. Chem. A* 2014, 118, 6730–6739.

(20) Roy, T. K.; Sharma, R.; Gerber, R. B. First-principles anharmonic quantum calculations for peptide spectroscopy; VSCF calculations and comparison experiments. *Phys. Chem. Chem. Phys.* 2016, 18, 1607–1614.

(21) Teixeira, F.; Melo, A.; Cordeiro, M. N. D. S.; Cordeiro, D. S. Calibration sets and the accuracy of vibrational scaling factors: A case study with the X3LYP hybrid functional. *J. Chem. Phys.* 2010, 133, 114109–114117.
Carnimeo, I.; Panek, P. Toward anharmonic computations of Chem. decapeptide hydrated by two waters: conformers determined by 121 ACS Omega field method. frequencies. Scaling factors and uncertainties for ab initio anharmonic vibrational 2011 vibrations. Phys. Rev. A 1997, 55, 124115.

Boyarkin, O. V.; Gerber, R. B. Conformational structures of a 124110. logical background and benchmark studies. J. Chem. Theory Comput. 2015, 11, 4082–4092.

Roy, T. K.; Kopysov, V.; Pereverzev, A.; Šebek, J.; Gerber, R. B.; Boyarkin, O. V. Intrinsic structure of pentapeptide Leu-enkephalin geometry optimization and validation by comparison of VSCF-PT2 calculations with cold ion spectroscopy. Phys. Chem. Chem. Phys. 2018, 20, 24894–24901.

Yagi, K.; Hirata, S.; Hiroa, K. Vibrational quasi-degenerated perturbation theory: applications to Fermi resonance in CO2, H2CO, and C2H2. Phys. Chem. Chem. Phys. 2008, 10, 1781–1788.

Pele, L.; Šebek; Potma, E. O.; Benny Gerber, R. Raman and IR spectra of butane: Anharmonic calculations and interpretation of room temperature spectra. Chem. Phys. Lett. 2011, 515, 7–12.

Zhao, H.; Bi, S.; Wang, Y.; Zhao, T. Characterization of the binding of paylean and DNA by fluorescence, UV spectroscopy and molecular docking techniques. Luminescence 2016, 31, 1013–1019.

Nagornova, N. S.; Rizzo, T. R.; Boyarkin, O. V. Interplay of intra- and intermolecular H-bonding in a progressively solvated macrocyclic peptide. Science 2012, 336, 320–323.

Fermi, E.; Pasta, J.; Ulam, S. Los Alamos Report LA-1940, 1955.

Alder, B. J.; Wainwright, T. E. Studies in molecular dynamics. I. General method. J. Chem. Phys. 1959, 31, 459–466.

Car, R.; de Angelis, F.; Giannozzi, P.; Marzari, N. First-principles molecular dynamics. In Handbook of Materials Modeling; Yip, S., Ed.; Springer: The Netherlands, 2005; pp 59–76.

Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. Phys. Rev. [Sect.] B 1964, 136, B864–B871.

Kohn, W.; Sham, L. J. Self-consistent equations including exchange and correlation effects. Phys. Rev. [Sect.] A 1965, 140, A1133–A1138.

Noid, D. W.; Kosezykowski, M. L.; Marcus, R. A. A spectral analysis method of obtaining molecular spectra from classical trajectories. J. Chem. Phys. 1977, 67, 404–408.

Adams, G. B.; Page, J. B.; Sankey, O. F.; Sinha, K.; Menendez, J.; Huffman, D. R. First-principles quantum-molecular-dynamics study of the vibrations of icosahehdral C60. Phys. Rev. B Condens. Matter Mater. Phys. 1991, 44, 4052–4055.

Gai, H.; Voth, G. A. First-principles molecular dynamics study of surface vibrations and vibrational mode coupling on the H/Si(111) 1x1 surface. J. Chem. Phys. 1994, 101, 1734–1737.

Wathelet, V.; Champagne, B.; Mosley, D. H.; André, J.-M.; Massida, S. Vibrational frequencies of diatomic molecules from Car and Parrinello molecular dynamics. Chem. Phys. Lett. 1997, 275, 506–512.

Martí, J. Analysis of the hydrogen bonding and vibrational spectra of supercritical model water by molecular dynamics simulations. J. Chem. Phys. 1999, 110, 6876–6886.

Praprotnik, M.; Janeičić, D. Molecular dynamics integration and molecular vibrational theory. III. The infrared spectrum of water. J. Chem. Phys. 2005, 122, 174103–174110.

Silvestrelli, P. L.; Bernasconi, M.; Parrinello, M. Ab initio infrared spectrum of liquid water. Chem. Phys. Lett. 1997, 277, 478–482.

Yin, W.-L.; Tse, J. S.; Itaka, T. Pressure-induced intermolecular interactions in crystalline silane-hydrogen. Phys. Rev. Lett. 2010, 105, 215501–215504.

Yadav, V. K.; Karmakar, A.; Choudhuri, J. R.; Chandra, A. A first principles molecular dynamics study of vibrational spectral diffusion and hydrogen bond dynamics in liquid methanol. Chem. Phys. 2012, 408, 36–42.

Hiratsuka, M.; Ohmura, R.; Sum, A. K.; Yasukou, K. Molecular vibrations of methane molecules in the structure 1 clathrate hydrate from ab initio molecular dynamics simulation. J. Chem. Phys. 2012, 136, 044508–44516.

Van Yperen-De Deyne, A.; De Meyer, T.; Pauwels, E.; Ghysels, A.; De Clerck, K.; Watoquier, M.; Van Speybroeck, V.; Hemelsoet, K. Exploring the vibrational fingerprint of the electronic excitation energy via molecular dynamics. J. Chem. Phys. 2014, 130, 134105–134114.

Eckart, C. Some studies concerning rotating axes and polyatomic molecules. Phys. Rev. 1935, 47, 552–558.
(65) Louck, J. D.; Galbraith, H. W. Eckart vectors, Eckart frames, and polyatomic molecules. Rev. Mod. Phys. 1976, 48, 69–106.

(66) Krasnoshechekov, S. V.; Isaeva, E. V.; Stepanov, N. F. Determination of the Eckart molecule-fixed frame by use of the apparatus of quaternion algebra. J. Chem. Phys. 2014, 140, 154104–154107.

(67) Kearse, S. K. An algorithm for the simultaneous superposition of a structural series. J. Comput. Chem. 1990, 11, 1187–1192.

(68) Mathews, J. H. Computer derivations of numerical differentiation formulae. Int. J. of Math. Education in Sci. and Tech. 2003, 34, 280–287.

(69) Hutter, J.; Iannuzzi, M.; Schöffmann, F.; VandeVondele, J. CP2K: atomistic simulations of condensed matter systems. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2014, 4, 15–25.

(70) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. Comput. Phys. Commun. 2005, 167, 103–128.

(71) Goedecker, S.; Teter, M.; Hutter, J. Separable dual-space Gaussian pseudopotentials. Phys. Rev. B: Condens. Matter Mater. Phys. 1996, 54, 1703–1710.

(72) Legrand, Y.-M.; van der Lee, A.; Barboiu, M. Single-crystal X-ray structure of 1,3-dimethylcyclobutadiene by confinement in a crystalline matrix. Science 2010, 329, 299–302.

(73) Kuchitsu, K. Calculation of the average structure of ethylene. J. Chem. Phys. 1966, 44, 906–911.

(74) Hollenstein, H.; Marquardt, R. R.; Quack, M.; Suhm, M. A. Dipole moment function and equilibrium structure of methane in an analytical anharmonic nine-dimensional potential surface related to experimental rotational constants and transition moments by quantum Monte Carlo calculations. J. Chem. Phys. 1994, 101, 3588–3602.

(75) Defrees, D. J.; Levi, B. A.; Pollack, S. K.; Hehre, W. J.; Binkley, J. S.; Pople, J. A. Effect of electron correlation on theoretical equilibrium geometries. J. Am. Chem. Soc. 1979, 101, 4085–4089.

(76) Martyna, G. J.; Tuckerman, M. E. A reciprocal space based method for treating long range interactions in ab initio and force-field-based calculations in clusters. J. Chem. Phys. 1999, 110, 2810–2821.

(77) Hornák, J.; Kapralova, P.; Bour, P. Simulations of vibrational spectra from classical trajectories: Calibration with ab initio force fields. J. Chem. Phys. 2007, 127, 084502–084509.

(78) Arnett, R. L.; Crawford, B. L. The vibrational frequencies of ethylene. J. Chem. Phys. 1950, 18, 118–126.

(79) Pugh, L. A.; Narahari Rao, K. Spectrum of water vapor in the 1.9 and 2.7 μ regions. J. Mol. Spectrosc. 1973, 47, 403–408.

(80) Camy-Peyret, C.; Flaud, J. M.; Maillard, J. P. 4ν-2 band of H2O-16. J. Phys., Lett. 1980, 41, 23–26.

(81) Thomas, M. A.; Welsh, H. L. The Raman spectrum of methane. Can. J. Phys. 1960, 38, 1291–1303.

(82) Gray, D. L.; Ribiette, A. G. Simultaneous analysis of the ν2 and ν4 bands of methane. Mol. Phys. 1976, 32, 1609–1625.

(83) Arnold, B. R.; Radziszewski, J. G.; Campion, A.; Perry, S. S.; Michl, J. Raman spectrum of matrix-isolated cyclobutadiene. Evidence for environmental hindrance to heavy-atom tunneling? J. Am. Chem. Soc. 1991, 113, 692–694.

(84) Orendt, A. H.; Arnold, B. R.; Radziszewski, J. G.; Facelli, J. C.; Malsch, K. D.; Strub, H.; Grant, D. M.; Michl, J. 13C NMR and polarized IR spectra of vicinal labeled[13C2] cyclobutadiene in an argon matrix: interconversion of valence tautomerism. J. Am. Chem. Soc. 1988, 110, 2648–2650.

(85) Arnold, B. R.; Michl, J. Ultraviolet and polarized infrared spectroscopy of matrix-isolated cyclobutadiene and its isotopomers. J. Phys. Chem. 1993, 97, 13348–13354.

(86) Maier, G. Tetrahedrane and cyclobutadiene. Angew. Chem., Int. Ed. Engl. 1988, 27, 309–332.

(87) Ulenikov, O. N.; Bekhtereva, E. S.; Albert, S.; Bauerecker, S.; Niederer, H. M.; Quack, M. Survey of the high resolution infrared spectrum of methane (12CH4 and 14CH4): partial vibrational assignment extended towards 12000 cm−1. J. Chem. Phys. 2014, 141, 234302–234333.

(88) Halonen, L. Internal coordinate Hamiltonian model for Fermi resonances and local modes in methane. J. Chem. Phys. 1997, 106, 831–845.

(89) Wang, X.-G.; Sibert, E. L., III A nine-dimensional perturbative treatment of the vibrations of methane and its isotopomers. J. Chem. Phys. 1999, 111, 4510–4522.

(90) Albert, S.; Bauerecker, S.; Boudon, V.; Brown, L. R.; Champion, J.-P.; Loëte, M.; Nitkin, A.; Quack, M. Global analysis of the high resolution infrared spectrum of methane 14CH4 in the region from 0 to 4800 cm−1. Chem. Phys. 2009, 356, 131–146.

(91) Ceder, G. Opportunities and challenges for first-principles materials design and applications to Li battery materials. MRS Bull. 2010, 35, 693–701.

(92) Wang, S.; Ye, H. First-principles calculation of crystalline materials genome: a preliminary study. Chin. Sci. Bull. 2014, 59, 1624–1634.

(93) Liu, Y.; Ferralis, N.; Bryndzia, L. T.; Grossman, J. C. Genome-inspired molecular identification in organic matter via Raman spectroscopy. Carbon 2016, 101, 361–367.