How Much Can Molecular Dynamics Improve on Harmonic Analysis for Vibrational Spectrum Predictions? Insights from Microcanonical Molecular Simulations of Highly Anharmonic Water Clusters

Yuzhe Zhang,[†] Yiwen Wang,[†] Xi Xu,[‡] Zehua Chen,[†] and Yang Yang*,[†]

[†] *Theoretical Chemistry Institute and Department of Chemistry, University of Wisconsin-Madison, 1101 University Avenue, Madison, Wisconsin 53706, USA*

[‡] *Center for Advanced Materials Research, Beijing Normal University, Zhuhai 519087, China*

E-mail: yyang222@wisc.edu

Abstract

Vibrational spectroscopy is widely used to gain insights into structural and dynamic properties of chemical, biological and material systems. Thus, an efficient and accurate method to simulate vibrational spectra is desired. In this paper, we propose a microcanonical molecular simulation scheme for efficient calculations of vibrational spectra of molecular clusters. Within the new scheme, we perform both harmonic analyses and molecular dynamics simulations and predict vibrational spectra of three challenging water clusters: the neutral water dimer (H_4O_2), the protonated water trimer ($H_7O_3^+$), and the protonated water tetramer ($H_9O_4^+$). We find that with the accurate description of quantum nuclear delocalization effects through the constrained nuclear-electronic

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 ORCID: https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. License: CC BY 4.0

orbital framework, including vibrational mode coupling effects through molecular dynamics simulations can additionally improve the vibrational spectrum calculations. In contrast, without the quantum nuclear delocalization picture, conventional *ab initio* molecular dynamics may even lead to less accurate results than harmonic analysis.

1 INTRODUCTION

Vibrational spectroscopy is a powerful tool for investigating physical properties and chemical processes in a variety of chemical, biological and material systems.[1–4] Theoretical predictions and simulations of vibrational spectra can provide valuable insights into both static and dynamic properties of a system, including structural information, interactions with the surrounding environment, reaction pathways and dynamic evolutions.

Many theoretical methods have been developed to compute vibrational spectra. Based on pre-calculated high quality potential energy surfaces (PESs), there are highly accurate methods such as vibrational self-consistent field/configuration interaction (VSCF/VCI),[5,6] multiconfiguration time-dependent Hartree (MCTDH),[7] and diffusion Monte Carlo (DMC).[8,9] These methods have been applied to a series of challenging molecular systems and provided accurate predictions of vibrational frequencies and successful identifications of subtle features in vibrational spectra.[10–13] In addition, there are path-integral based methods that can calculate vibrational spectra with reasonable accuracy and less computational cost. Examples often include centroid molecular dynamics (CMD),[14,15] ring polymer molecular dynamics (RPMD),[16] thermostatted RPMD (TRPMD)[17] and quasi-centroid molecular dynamics (QCMD).[18–20] Although CMD may suffer from curvature problems[21,22] and vibrational spectra calculated by RPMD may be contaminated by internal ring-polymer vibrational modes,[23,24] recently developed QCMD and TRPMD can mitigate these problems. Combined with accurate force fields or on-the-fly *ab initio* calculations, these path-integral based methods have been applied to gas phase and condensed phase systems and produced vibrational spectra that agree well with experimental results,[17,25,26] although in general the

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

direct combination with *ab initio* calculations remains expensive. A less expensive method is the vibrational second-order perturbation theory (VPT2).[27] VPT2 takes into account anharmonicities and mode coupling effects through a perturbative expansion around the local stationary geometry, and it has been very successful in predicting the vibrational frequencies of many molecular systems. However, VPT2 may face challenges in highly anharmonic systems such as shared proton systems.[28]

Despite the accuracy of the aforementioned methods, currently in the field the most widely used methods are harmonic analysis and molecular dynamics (MD) simulations based on either force fields or *ab initio* calculations. In harmonic analysis, the PES is assumed to be harmonic and vibrational frequencies are directly obtained from diagonalizing the mass-weighted Hessian matrix at a stationary geometry. In practical calculations, empirical scaling parameters are often applied to the harmonic frequencies to account for the lack of anharmonic and nuclear quantum effects, thereby reducing the deviation from experimental results. Despite the popularity of empirical scaling, the choice of scaling parameters is *ad hoc* and can therefore vary significantly depending on systems, vibrational modes, and underlying electronic structure methods. Compared to harmonic analysis, MD simulations, especially *ab initio* MD (AIMD) simulations, can often give better vibrational spectra with more detailed information. They can sample the anharmonic region of PESs as well as incorporate coupling effects between different vibrational modes through thermal motions.

A commonly used procedure to obtain vibrational spectra with AIMD is first to thermally equilibrate the system with a canonical ensemble (NVT) simulation, subsequently perform many microcanonical (NVE) simulations starting from uncorrelated frames picked from the equilibrated NVT trajectory, and perform a Fourier transform on obtained NVE autocorrelation functions before taking an average to obtain the spectrum.[29–31] Unfortunately, this procedure, which we will refer to as the NVT-NVE scheme hereafter, requires numerous NVE simulation runs to reach sufficient Boltzmann sampling and obtain converged spectra. Later, Kirchner and co-workers proposed the direct use of equilibrated NVT trajectories to

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

calculate vibrational spectra, which we will refer to as the direct *NVT* scheme hereafter.[32,33] This direct *NVT* scheme greatly reduced the simulation cost and worked well for bulk systems, although for gas phase small molecules with small couplings between a limited number of degrees of freedom, it often needs an impractically long time for modes to exchange energy and to reach equilibrium.[33] Furthermore, as has been shown in model systems, the direct *NVT* scheme may introduce artifacts caused by thermostats into vibrational spectra.[33]

Recently, our group developed a molecular dynamics framework based on constrained nuclear-electronic orbital density functional theory (CNEO-DFT)[34,35] to efficiently incorporate nuclear quantum effects, especially zero-point effects due to quantum delocalization, into molecular simulations.[36,37] In CNEO-DFT, both electrons and key nuclei are treated quantum mechanically, but the classical molecular geometry picture is retained using the expectation values of position operators for quantum nuclei together with the positions of classical nuclei. This treatment is justified by the physical intuition that quantum nuclei are still relatively localized in space in most chemical systems of interest. Therefore, in CNEO-DFT, constraints on the expectation values of quantum nuclear positions are imposed, which leads to effective PESs that inherently incorporate nuclear quantum effects. On these effective PESs, molecular dynamics, termed CNEO-MD,[37] can be performed, and our group has found that with essentially the same computational cost as conventional DFT-based AIMD (less than 15% more expensive as benchmarked, see Supporting Information of ref 28), CNEO-MD significantly outperforms conventional AIMD in describing vibrational modes with large hydrogen-motion character.[28,37]

However, like conventional AIMD, CNEO-MD may still suffer from the insufficient Boltzmann sampling problem either from the *NVT*-*NVE* or the direct *NVT* scheme. Additionally, in the past, CNEO-MD was mostly performed on simple small molecules, and it awaits further tests on more complicated systems with much stronger mode coupling effects. In this paper, we propose a new scheme for efficient calculations of vibrational spectra for gas phase clusters using only microcanonical molecular simulations. This new scheme, which we refer

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

to as the direct *NVE* scheme, is compared to both *NVT*-*NVE* and direct *NVT* schemes, and is found to be able to obtain converged spectra with much less simulation time. We further calculate the vibrational spectra of three highly challenging water clusters: the neutral water dimer (H_4O_2), the protonated water trimer ($H_7O_3{}^+$), and the protonated water tetramer ($H_9O_4{}^+$) and demonstrate the excellent performance of CNEO-MD in combination with the direct *NVE* scheme. We find that in these systems, CNEO-MD can further improve over CNEO-DFT harmonic analysis in vibrational spectra calculations, due to the incorporation of strong mode coupling effects on top of quantum nuclear delocalization effects. In contrast, conventional DFT-based AIMD may not necessarily perform better than DFT harmonic analysis because of the lack of quantum nuclear delocalization effects.

2 METHODS

2.1 Review of infrared spectra calculations with classical simulations

To introduce the direct *NVE* scheme, we first briefly review infrared (IR) spectra calculations with classical simulations. IR spectroscopy measures the frequency-dependent absorbance ($A(\omega)$) of IR waves by a sample. The measured absorbance is directly proportional to the sample's absorption cross-section ($\sigma(\omega)$) and thus the rate of energy loss from the radiation ($\dot{E}_{rad}(\omega)$) through

$$A(\omega) = \frac{\sigma(\omega) l}{V} = \frac{\dot{E}_{rad}(\omega) l}{S(\omega) V} \tag{1}$$

where l is the length and V is the volume of the sample under radiation, and $S(\omega)$ is the energy flux, or the time-averaged amplitude of the Poynting vector.

With a continuous and perturbative electromagnetic field $E(t) = E_0 \cos \omega t$, the radiation

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

energy loss rate can be computed using Fermi's golden rule:

$$\dot{E}_{rad}(\omega) = \big[R(\omega) - R(-\omega)\big]\hbar\omega = \frac{\omega E_0^2}{12\hbar}(1 - \mathrm{e}^{-\beta\hbar\omega}) \int_{-\infty}^{\infty} \mathrm{d}t \ \mathrm{e}^{-i\omega t} \langle \hat{\boldsymbol{\mu}}(0) \hat{\boldsymbol{\mu}}(t) \rangle_{\mathrm{qm}} \tag{2}$$

where $R(\omega)$ and $R(-\omega)$ are the excitation and de-excitation transition rates, $\beta = 1/k_B T$, with k_B the Boltzmann constant and T the temperature of the system, and $\langle \hat{\boldsymbol{\mu}}(0) \hat{\boldsymbol{\mu}}(t) \rangle_{\mathrm{qm}}$ is the quantum dipole autocorrelation function.[38–40] The energy flux can be calculated by classical electromagnetic theory, and in the end, the IR spectra can be obtained with the Fourier transform of the quantum dipole autocorrelation function

$$A(\omega) = \frac{\omega \mu_r l}{6\hbar n(\omega)\epsilon_0 c V}(1 - \mathrm{e}^{-\beta\hbar\omega}) \int_{-\infty}^{\infty} \mathrm{d}t \ \mathrm{e}^{-i\omega t} \langle \hat{\boldsymbol{\mu}}(0) \hat{\boldsymbol{\mu}}(t) \rangle_{\mathrm{qm}} \tag{3}$$

where $n = \sqrt{\epsilon_r \mu_r}$ is the frequency-dependent refractive index, $c = 1/\sqrt{\mu_0 \epsilon_0}$ is the speed of light in vacuum, and the relative permeability μ_r is often approximated to be unity.

In practice, because quantum correlation functions are usually difficult to calculate, classical autocorrelation functions obtained from classical simulations are more often used for spectrum computation. These two types of autocorrection functions are not rigorously equal, but within the harmonic oscillator model, if the quantum and classical oscillators both reach equilibrium at the temperature T within a canonical ensemble, the Fourier transforms of quantum and classical position autocorrelation functions only differ by a frequency-dependent prefactor,

$$\frac{\int_{-\infty}^{\infty} \mathrm{d}t \ \mathrm{e}^{-i\omega t} \langle \hat{x}(0) \hat{x}(t) \rangle_{\mathrm{qm}}}{\int_{-\infty}^{\infty} \mathrm{d}t \ \mathrm{e}^{-i\omega t} \langle x(0) x(t) \rangle_{\mathrm{cl}}} = \frac{\hbar\omega}{kT(1 - e^{-\beta\hbar\omega})}. \tag{4}$$

Although this prefactor is derived from the harmonic oscillator model, it is often used to approximately account for the differences between quantum and classical position and dipole autocorrelation functions in real systems. Thus, with this harmonic approximation, the IR

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

absorbance $A(\omega)$ can be expressed with classical dipole autocorrelation functions

$$A(\omega) = \frac{B(\omega)\omega^2}{kT} \int_{-\infty}^{\infty} \mathrm{d}t \ \mathrm{e}^{-i\omega t} \langle \boldsymbol{\mu}(0)\boldsymbol{\mu}(t) \rangle_{\mathrm{cl}} \tag{5}$$

where $B(\omega) = \frac{\mu_r l}{6n(\omega)\epsilon_0 cV}$. For further simplification, integration by parts can be performed, and the absorbance can be expressed with the autocorrelation function of the time derivative of the dipole vector

$$A(\omega) \equiv \frac{B(\omega)}{kT} \int_{-\infty}^{\infty} \mathrm{d}t \ \mathrm{e}^{-i\omega t} \langle \dot{\boldsymbol{\mu}}(0)\dot{\boldsymbol{\mu}}(t) \rangle_{\mathrm{cl}}. \tag{6}$$

This equation implies that in principle, in order to obtain the IR absorption spectra, one needs to generate enough sampling from the canonical ensemble. This is usually achieved by running an NVT trajectory at a given temperature, and then performing NVE simulations starting from each sampled configuration. The final absorption spectra can be obtained by taking the average of the Fourier transformed dipole derivative autocorrelation function, with a prefactor that is essentially frequency independent if the refractive index $n(\omega)$ does not significantly vary with frequency. This equation is the theoretical foundation for the calculations of vibrational spectra with the NVT-NVE scheme. The NVT-NVE scheme may sound easy to perform, but to achieve enough sampling from the canonical ensemble, the number of configurations needed as well as subsequent NVE trajectories is very large even in model systems.[33]

In order to reduce the computational cost, Kirchner and co-workers introduced the direct NVT scheme for spectrum calculations.[32,33] Instead of an NVT simulation followed by many NVE simulations, the direct NVT scheme only uses one equilibrated NVT trajectory to calculate the ensemble averaged dipole derivative autocorrelation function. In principle, this scheme is not theoretically rigorous because the thermostat will affect the evolution of the system and thus bring artifacts into the spectrum. However, in practical simulations and especially those in the condensed phase, if a weak thermostat is used, the artifacts are negligible.[32] Despite its success, a scheme that is both computationally inexpensive and

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 ORCID: https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. License: CC BY 4.0

theoretically more rigorous for gas phase vibrational spectrum simulations is highly desirable.

2.2 Direct *NVE* scheme for vibrational spectra calculations

In this paper, we propose the direct NVE scheme. Different from the previous two schemes, which are built on calculating the dipole derivative autocorrelation functions for a canonical ensemble equilibrated at a given temperature T, we provide the theoretical foundation of the direct NVE scheme by revisiting the autocorrelation function of the harmonic model within a microcanonical ensemble.

For a microcanonical ensemble of harmonic oscillators with energy E, the classical position autocorrelation function is

$$\langle x(0)x(t)\rangle_{\mathrm{cl}}^{NVE} = \frac{E}{m\omega^2}\cos\omega t \tag{7}$$

This is very similar to the classical position autocorrelation function within the canonical ensemble at temperature T, which is

$$\langle x(0)x(t)\rangle_{\mathrm{cl}}^{NVT} = \frac{k_B T}{m\omega^2}\cos\omega t. \tag{8}$$

These two equations indicate that within the harmonic oscillator model, the classical position autocorrelation functions of the microcanonical ensemble and the canonical ensemble only differ by a prefactor k_BT/E. In previous work, classical autocorrelation functions from a canonical ensemble were widely used to approximate quantum autocorrelation functions. A natural question then arises: can the autocorrelation functions from a microcanonical ensemble be used for this approximation? If so, we can easily calculate IR spectra with

$$A(\omega) = \frac{B(\omega)}{E}\int_{-\infty}^{\infty}\mathrm{d}t\ \mathrm{e}^{-i\omega t}\langle\dot{\boldsymbol{\mu}}(0)\dot{\boldsymbol{\mu}}(t)\rangle_{\mathrm{cl}}^{NVE}. \tag{9}$$

This equation seems trivially similar to equation 6, but there is a key difference: the autocorrelation function in equation 6 needs to be obtained from a canonical ensemble whereas

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 ORCID: https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. License: CC BY 4.0

here it needs to be obtained from a microcanonical ensemble. This difference is especially beneficial because for a given system, the phase space of a microcanonical ensemble with a fixed energy E is significantly smaller than that of a canonical ensemble with a fixed temperature T. As such, the sampling need for a microcanonical ensemble simulation is much smaller.

Based on equation 9, the most straightforward way to obtain IR spectra is to run an NVE MD trajectory and then perform a Fourier transform on the obtained dipole derivative autocorrelation function. In practice, to further accelerate the sampling and to reduce the wall time, we can run multiple NVE trajectories in parallel with the same energy but different initial phase space configurations and subsequently average the results from multiple trajectories. This is the direct NVE scheme we are introducing in this paper. Later, we will show that the direct NVE scheme performs better than both NVT-NVE and direct NVT schemes in obtaining converged spectra with minimal simulation time.

One should notice that in conventional NVT simulations, with the help of thermostats, all modes will equilibrate to the same temperature and there is no need for special treatment of each specific mode. However, in the direct NVE scheme, although in principle we only need to fix the total energy of the system and allow the energy to flow naturally between different modes during the simulation, in gas phase systems, particularly in small molecules with relatively weak mode couplings, the energy flow is slow, and it could take a long time to reach ergodicity and thus energy equipartition. This is the main drawback of the direct NVE scheme. To alleviate this problem, in practice we directly start the NVE simulations from configurations in which all modes have the same energy. This will artificially favor those configurations close to energy equipartition, but we will show later that the simulated spectra are reasonably accurate compared to experimental results in a series of highly challenging water cluster systems.

We also note that in conventional NVT simulations, the simulation temperature can affect the final spectra, including both peak positions and peak intensities. This is in con-

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

trast to quantum simulations, in which only peak intensities change with temperature. The temperature dependence of peak positions in classical simulations is a known artifact,[32,33] although temperature elevation is often used to account for anharmonicities and mode coupling effects, which are absent in harmonic analysis. Similarly, since the direct NVE scheme is also based on classical simulations, the final spectra will depend on the simulation energy E. A finite energy E will naturally lead to the incorporation of anharmonic and mode coupling effects, which we will show later to be important for accurate spectrum prediction in the challenging water cluster systems.

2.3 Brief review of CNEO-MD working equations

In this paper, We will combine the direct NVE scheme with both conventional DFT-based AIMD and CNEO-MD. The CNEO-MD framework was recently developed by our group to incorporate nuclear quantum effects, especially zero-point effects due to quantum delocalization, in practical molecular simulations.[37] Its working equations are[36,37,41]

$$m\frac{\mathrm{d}\langle \boldsymbol{x}\rangle}{\mathrm{d}t} = \langle \boldsymbol{p}\rangle \tag{10}$$

and

$$\frac{\mathrm{d}\langle \boldsymbol{p}\rangle}{\mathrm{d}t} \approx -\nabla_{\langle \boldsymbol{x}\rangle} V^{\mathrm{CNEO}}(\langle \boldsymbol{x}\rangle) \tag{11}$$

where $\langle \boldsymbol{x}\rangle$ and $\langle \boldsymbol{p}\rangle$ are expectation position and expectation momentum, respectively, and $V^{\mathrm{CNEO}}(\langle \boldsymbol{x}\rangle)$ is the CNEO-DFT energy surface. Note that CNEO-MD is essentially an AIMD method using the energies and gradients of CNEO-DFT computed on-the-fly at each MD step. Compared with conventional DFT-based AIMD, CNEO-MD only leads to a small increase in computational cost ($<15\,\%$ with recent implementation) but much more accurate vibrational spectra with quantum nuclear delocalization effects incorporated.[28,37,42]

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

3 COMPUTATIONAL DETAILS

The gas phase methanol molecule was used as an example to compare the performance of three different MD sampling schemes and demonstrate the strength of the direct *NVE* scheme. For the scheme comparison, we only performed CNEO-MD since the results from conventional AIMD will be highly similar. To further show the power of the new sampling scheme along with CNEO-MD, as well as the importance of vibrational mode coupling effects, we chose three highly challenging water clusters, including the neutral water dimer(H_4O_2), the protonated water trimer ($H_7{O_3}^+$), and the protonated water tetramer ($H_9{O_4}^+$) for study. It is known that the choice of electronic density functional could significantly influence the calculation of vibrational spectra,[37,43] and here we adopted the PBE0[44,45] functional for both CNEO-DFT and DFT calculations as it was previously found that PBE0 is the best functional for free X-H stretches as benchmarked against the CCSD(T) method[46,47] for a series of small organic molecules,[37] where X is a heavy atom such as carbon, nitrogen and oxygen. However, because hydrogen bonds are not present in the small organic molecules that were benchmarked before, here for water cluster systems, we additionally benchmarked several commonly used functionals against CCSD(T) on the O-H stretch modes in H_4O_2. We found that the meta-GGA functional ωB97MV[48] with VV10 dispersion correction[49] has the best performance and PBE0 is a close second, with less than $25\,\mathrm{cm}^{-1}$ deviations from the CCSD(T) reference (Table S1). Because of the relatively low computational cost and relatively high accuracy, we therefore continued using PBE0 in this paper. We used the def2-TZVP[50] electronic basis set for methanol calculations and def2-TZVPPD[50,51] for water clusters calculations. As for nuclei, we only treated protons quantum mechanically because they feature the most significant nuclear quantum effects, but we note that a full quantum treatment for all nuclei is possible.[35] The PB4D basis set[52] was used for protons. Because here protons are relatively localized with little overlap between each other, they were treated as distinguishable particles and no proton-proton correlation was included. We did not include electron-proton correlation (epc), either, because we have previously found

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

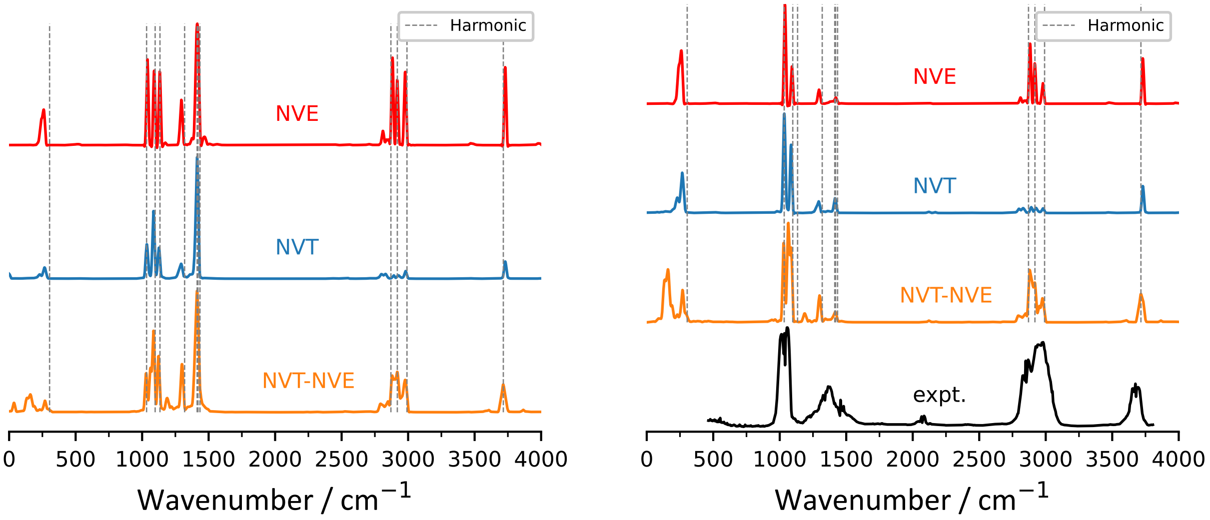
that without an epc functional,[53–55] CNEO-DFT and CNEO-MD are already sufficiently accurate for vibrational spectra calculations.[28,37,56] We leave the investigation of epc effects for future studies. Both DFT and CNEO-DFT harmonic analyses were performed using an in-house version[57] of PySCF,[58,59] and MD simulations were performed with Atomic Simulation Environment (ASE).[60] For the direct *NVE* scheme, unless otherwise specified, 50 2 ps simulations were performed for spectra computation. To initialize each trajectory, the starting geometry was the optimized geometry and a kinetic energy $E = k_B T$ was added to each mode with random positive or negative velocity directions.[61] For the direct *NVT* scheme, following ref 32, we employed the Nosé-Hoover chain thermostat[62–64] which has been implemented in an in-house version of ASE. First, a 10 ps simulation was performed for the system to reach equilibrium, then another 10 ps simulation was performed for the production run. It is known that the Nosé-Hoover chain thermostat may suffer from ergodicity issues in small or strongly harmonic systems,[65] therefore for the *NVT*-*NVE* scheme, we used the Andersen thermostat[66] for *NVT* sampling. A 20 ps *NVT* simulation was first performed, and then uncorrelated configurations from the later 15 ps were picked as initial configurations for subsequent 2 ps *NVE* simulations.

4 RESULTS AND DISCUSSION

4.1 Comparison of simulation schemes

Figure 1 presents the power and IR spectra of a single methanol molecule at 300 K or its equivalent energy by CNEO-MD with direct *NVE*, direct *NVT*, and *NVT*-*NVE* schemes. For the direct *NVE* scheme, we used five 2 ps trajectories to obtain converged spectra. As power spectra reflect the energy distribution among different vibrational modes, we can see from Figure 1(a) that energy is about equally distributed among different modes from the direct *NVE* scheme, which leads to good agreement between its IR spectrum and the experimental spectrum. For the direct *NVT* scheme, from the power spectrum, we can see that with the

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0



(a) Power spectra

(b) IR spectra

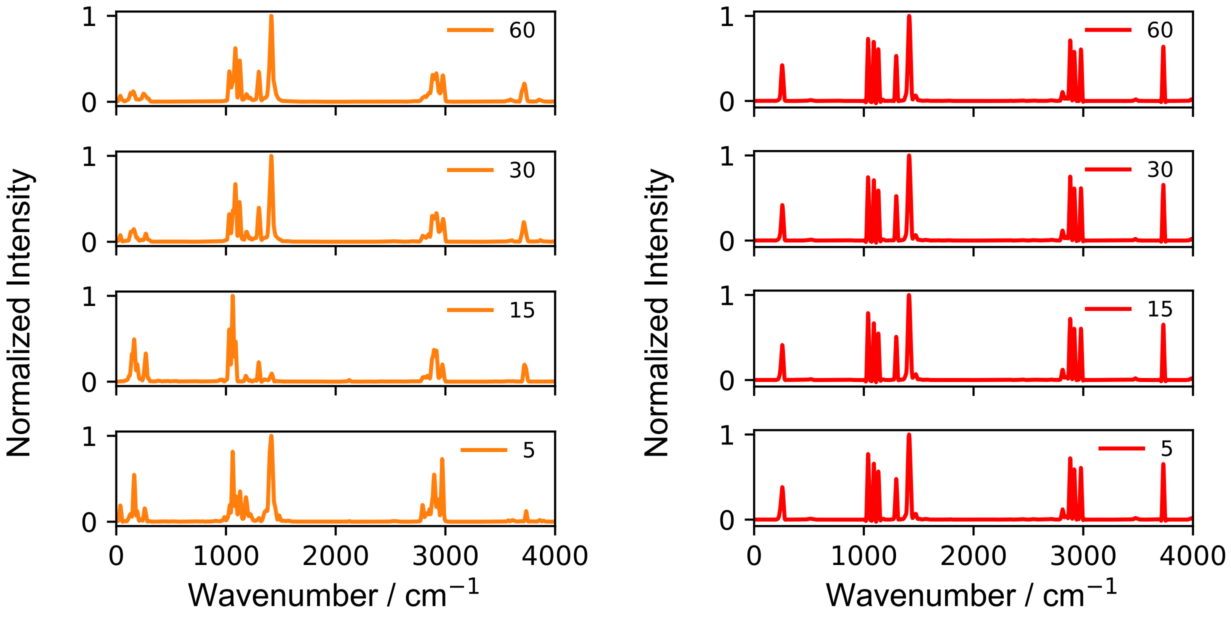
Figure 1: Power and IR spectra of a single methanol molecule by different MD schemes as indicated. Dashed vertical lines are harmonic analysis results from CNEO-DFT. The experimental IR spectrum of gas phase methanol is obtained from the National Institute of Standards and Technology (NIST) WebBook.

same amount of total simulation time as the direct NVE scheme, energy equipartition in the direct NVT scheme is far from being achieved. Specifically for this trajectory, energies of the O-H and C-H stretch modes above 2500 cm^{-1} are much lower than the remaining modes, which leads to weak IR intensities for the highly IR active C-H and O-H stretch modes in the experimental spectrum. This observation agrees well with previous literature that similar unequal energy distributions across different vibrational modes were observed for a single methanol molecule in conventional AIMD simulations.[32] As to the NVT-NVE scheme, here the spectra were obtained from an average of 30 2 ps NVE trajectories, whose initial configurations were sampled every 500 fs from an NVT trajectory. This choice of 500 fs sampling time gap is justified by the fast decay of autocorrelation functions of the NVT trajectory (Figure S1). From the power and IR spectra, it can be seen that the results from the NVT-NVE scheme are close to the direct NVE scheme and in good agreement with the experimental spectrum as well.

Note that here we only ran 5 trajectories for the direct NVE scheme but 30 trajectories

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

for the *NVT*-*NVE* scheme. This is because the spectra calculated from the direct *NVE* scheme need fewer trajectories to converge. Figure 2 shows the convergence test results for power spectra as a function of number of trajectories used in both schemes. We can see that at least 30 trajectories are needed for convergence within the *NVT*-*NVE* scheme. In contrast, only 5 trajectories are sufficient to achieve convergence within the direct *NVE* scheme for the methanol molecule. Although later in more complicated water clusters, more trajectories will be needed for spectrum convergence, the general trend is always that the direct *NVE* scheme needs significantly fewer trajectories than the *NVT*-*NVE* scheme to reach convergence. As was discussed before, this difference of convergence behavior is due to the much smaller phase space of a microcanonical ensemble compared to that of the corresponding canonical ensemble, and it makes the direct *NVE* scheme more efficient in vibrational spectrum simulations.



(a) *NVT*-*NVE*

(b) *NVE*

Figure 2: Methanol molecule power spectra as a function of number of trajectories averaged for *NVT*-*NVE* and direct *NVE* schemes. The direct *NVE* scheme needs significantly fewer trajectories than the *NVT*-*NVE* scheme to reach spectrum convergence.

Another difference between the direct *NVE* and the *NVT*-*NVE* schemes is that peaks are

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

generally broader in the *NVT*-*NVE* scheme. It is understandable that with the *NVT*-*NVE* scheme, there is a total energy distribution with more fluctuation, whereas the total energy is fixed in the direct *NVE* scheme. However, we emphasize that this does not mean that energy will not flow with the direct *NVE* scheme. Instead, through mode coupling effects, energy can exchange among different modes and also broaden their peaks. Therefore, both schemes are able to capture the mode coupling effects through finite temperature or energy, and it is simply the fact that the fluctuation is larger in the *NVT*-*NVE* scheme that makes the peaks broader. However, we cannot argue which peak width is more physical since peak broadening can be observed even in simple harmonic and Morse oscillator models, in which quantum references give sharp peaks and thus the broad peaks obtained from classical simulations are purely artifacts.

From the spectrum comparison and spectrum convergence results, we can conclude that both the direct *NVE* scheme and the *NVT*-*NVE* scheme can produce reliable spectra, but the direct *NVE* scheme is significantly more efficient in spectrum convergence with minimal simulation time. The direct *NVT* scheme may be as efficient as the direct *NVE* scheme, but it may suffer from difficulty in effectively equilibrating different vibrational modes in a single trajectory due to relative weak mode couplings between a limited number of degrees of freedom in gas phase clusters. However, we note that in condensed phase systems where modes couple more strongly and interact with the surroundings more frequently, reaching equipartition will be easily achieved and thus the direct *NVT* scheme will be more successful in that paradigm.[32,33]

4.2 H_4O_2

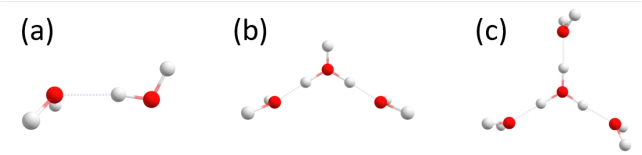


Figure 3: Schematic pictures of (a) H_4O_2, (b) $H_7O_3^+$ and (c) $H_9O_4^+$ water clusters.

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 ORCID: https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. License: CC BY 4.0

Table 1: Vibrational frequencies of O-H stretch modes in H_4O_2 (in cm^{-1}) from harmonic analysis and experiments.

Vibrational mode	expt.[a]	CCSD(T)[a]	DFT	CNEO-DFT
free O-H	3745	3941	3964	3778
	3735	3923	3944	3766
	3660	3836	3861	3679
bound O-H	3601	3753	3731	3494

[a] From ref. 67.

To further show the capability of the combination of the direct *NVE* scheme and CNEO-MD, we calculated the IR spectra of three water cluster systems. The structures of three water clusters to be studied are shown in Figure 3 (Structure files are given in the Supporting Information). To begin, the water dimer H_4O_2 is the simplest water cluster and its vibrational spectra have been studied experimentally and theoretically for decades.[68–75] In H_4O_2, the two water molecules are bound by the intermolecular hydrogen bond, and there are three free O-H stretch modes and one bound O-H stretch mode. The vibrational frequencies of these four O-H stretch modes obtained from CCSD(T), DFT, and CNEO-DFT harmonic analysis, along with experimental references, are presented in Table 1. Within the harmonic approximation, CCSD(T) overestimates all O-H stretch frequencies by over $150\,\mathrm{cm}^{-1}$ and DFT with PBE0 gives similar results. In contrast, although the harmonic approximation is still used, because of the incorporation of quantum nuclear delocalization effects, CNEO-DFT with PBE0 accurately describes the vibrational frequencies of the free O-H stretches, whose errors are within $30\,\mathrm{cm}^{-1}$ of the experimental references. Interestingly, for the bound O-H stretch, CNEO-DFT harmonic analysis gives a larger error and underestimates the frequency by 107 cm^{-1}, but it still slightly outperforms DFT, which overestimates the frequency by $130\,\mathrm{cm}^{-1}$.

The large error of the CNEO-DFT harmonic analysis for the bound O-H stretch seems to indicate a failure of CNEO-DFT in this particular case, but one should note a special feature of H_4O_2: its bound O-H stretch strongly couples with its low-frequency modes, making the system highly anharmonic.[76,77] Therefore, although CNEO-DFT harmonic analysis

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

incorporates quantum nuclear delocalization effects and thus some single mode anharmonicity, the absence of mode coupling effects beyond the harmonic picture could be the reason for its lower accuracy in describing the bound O-H stretch. Since MD simulations provide an efficient way of incorporating mode coupling effects through finite temperature thermal motions,[32] we performed CNEO-MD with the new direct *NVE* scheme to investigate the importance of these coupling effects. Note that the water dimer system is more complex than the methanol molecule and different trajectories tend to give significantly different spectra (Figure S3). Therefore, instead of using 5 trajectories as in the methanol case, we used 50 trajectories to obtain the converged IR spectrum, although we found that in fact 30 trajectories are already sufficient to reach good spectra convergence. (See Figure S4-S6 for convergence test results, for H_4O_2, ${H_7O_3}^+$, and ${H_9O_4}^+$, respectively).

Figure 4 shows the calculated IR spectra from CNEO-MD and AIMD with the total energy corresponding to 300 K. For comparison, we also show their harmonic analysis results with artificial peak broadening as well as the experimental peak position references. We can see that in spectra obtained from both CNEO-MD and AIMD, the bound O-H stretch blueshifts about $100\,\mathrm{cm}^{-1}$ compared with their corresponding harmonic results. This shift makes CNEO-MD accurately capture the frequency of the bound O-H stretch but makes the AIMD results worse in that the underlying DFT already largely overestimates the vibrational frequency from harmonic analysis without quantum nuclear delocalization effects incorporated. Because the free O-H stretches are barely coupled with other modes, MD simulations hardly change their peak positions compared to harmonic analyses. Therefore, with the incorporation of both quantum nuclear delocalization effects and mode coupling effects, CNEO-MD significantly improves over CNEO-DFT harmonic analysis and can excellently describe both free and bound O-H stretches in the water dimer. In contrast, due to the absence of quantum nuclear delocalization effects, DFT-based AIMD may not necessarily perform better than DFT harmonic analysis. Ironically, for the bound O-H stretch, due to the error cancellation from missing nuclear quantum effects and mode coupling effects, the

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

simple DFT harmonic analysis is much better than the more expensive DFT-based AIMD.

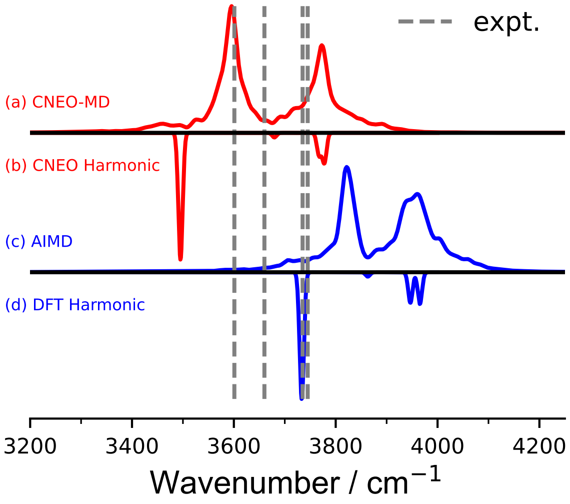


Figure 4: Calculated IR spectra of H_4O_2 by (a) CNEO-MD, (b) CNEO harmonic analysis, (c) AIMD and (d) DFT harmonic analysis. Harmonic spectra are obtained by adding 10 cm^{-1} Gaussian width to each transition. Experimental reference is from ref. 67.

As mentioned in Section 2.2, energy in classical *NVE* simulations will artificially affect computed vibrational spectra, and for H_4O_2, we indeed observed that the simulation energy influences how much the bound O-H stretch shifts. (See Figure S2 for temperature dependence test). Here we emphasize that the simulation temperatures/energies do not mean that experiments are carried out at this temperature, instead, they should be purely perceived as a parameter that can modulate the mode coupling strength. We found that the 300 K parameter works well for both the water dimer and the protonated water clusters studied here, and it also seems to work well for other hydrogen bonded systems.[78]

4.3 $H_7O_3^+$ AND $H_9O_4^+$

CNEO-MD was also applied to two significantly more challenging protonated water cluster systems, $H_7O_3^+$ and $H_9O_4^+$, whose vibrational spectra have been studied extensively with high-level and computationally intensive methods such as VSCF/VCI,[10–12] MCTDH[13] and DMC.[79] Relatively cheap AIMD and TRMPD have also been applied to these two systems,

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 ORCID: https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. License: CC BY 4.0

but the spectra of the former are blueshifted significantly and the latter suffers from artificial broadening.[80]

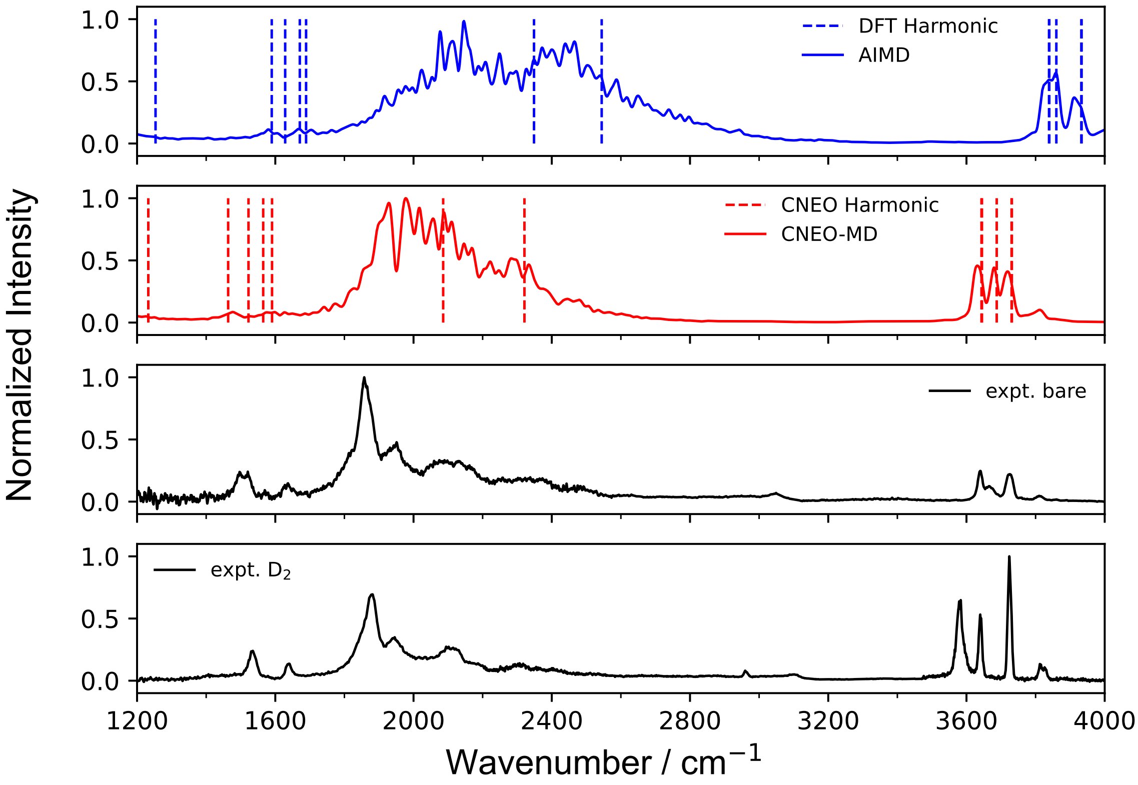


Figure 5: Overview of IR spectra of $H_7O_3^+$ obtained by the indicated methods. Experimental IR spectra are from ref. 10.

Figure 5 shows the calculated IR spectra of $H_7O_3^+$ from AIMD and CNEO-MD along with experimental references. In experimental spectra, the dominant feature is the broad progression of bands from 1800 to 2400 cm^{-1} with the maximum peak located at 1858 cm^{-1} for the bare cluster and 1878 cm^{-1} for the D_2-tagged cluster.[10] VSCF/VCI analysis assigned the most intense peak to the hydronium H_3O^+ asymmetric stretch with strong coupling to some low-frequency modes, and the broad tail extended to 2400 cm^{-1} is attributed to many overtones and combination bands.[10] CNEO harmonic analysis overestimates the asymmetric stretch of bound O-H of H_3O^+ by 200 cm^{-1} (2086 cm^{-1}), although it is already much better than the DFT harmonic frequency (2349 cm^{-1}). With the incorporation of mode

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 ORCID: https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. License: CC BY 4.0

coupling effects, CNEO-MD gives further improved results over CNEO harmonic analysis with an intense peak positioned around 1900 cm^{-1} and shows a progression of bands extended to 2400 cm^{-1}. Although the relative intensity contrast between the main peak and the band progression is not as sharp as that in the experimental spectra, this significant improvement again demonstrates that CNEO-MD overall successfully reproduces most features presented in experiment and captures the strong mode coupling effects between bound O-H stretches and other low-frequency modes. In contrast, conventional AIMD predicts an overly broadened progression from 1800 to 2800 cm^{-1} with its most intense peak appearing around 2100 cm^{-1}. For the free O-H stretches around 3600 cm^{-1} that have limited coupling with other modes, MD barely changes relative to the harmonic analysis results, and both CNEO harmonic analysis and CNEO-MD can accurately predict their vibrational frequencies. This observation is again consistent with the water dimer case as well as a series of molecular systems our group has previously studied.[37]

Figure 6 shows the calculated IR spectra of $H_9O_4^+$ from AIMD and CNEO-MD along with experimental references. For this cluster, there were controversies in the past about the existence of the linear-chain Zundel isomer in the observed gas phase IR spectrum,[31] but later joint experimental and theoretical efforts concluded that only the Eigen isomer exists in the experimental spectrum.[11] Therefore, we only performed CNEO-MD and conventional AIMD simulations on the Eigen isomer. The prominent feature of the experimental IR spectra for $H_9O_4^+$ is the band around 2650 cm^{-1}, which is assigned to the hydronium H_3O^+ asymmetric stretch coupled to low-frequency modes.[12] CNEO harmonic analysis predicts the frequency of this mode to be 2585 cm^{-1}, which is 65 cm^{-1} lower than experiment. With CNEO-MD, this peak blueshifts to 2600 cm^{-1}, slightly closer to the experimental results. In contrast, DFT harmonic analysis overestimates the frequency by about 200 cm^{-1} and AIMD makes the overestimation slightly worse. Despite the great success of CNEO-MD in reproducing the main bands of the experimental IR spectra, we point out a small feature that seems absent in CNEO-MD: a diffuse feature that lies about 380 cm^{-1} above the dominant

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 ORCID: https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. License: CC BY 4.0

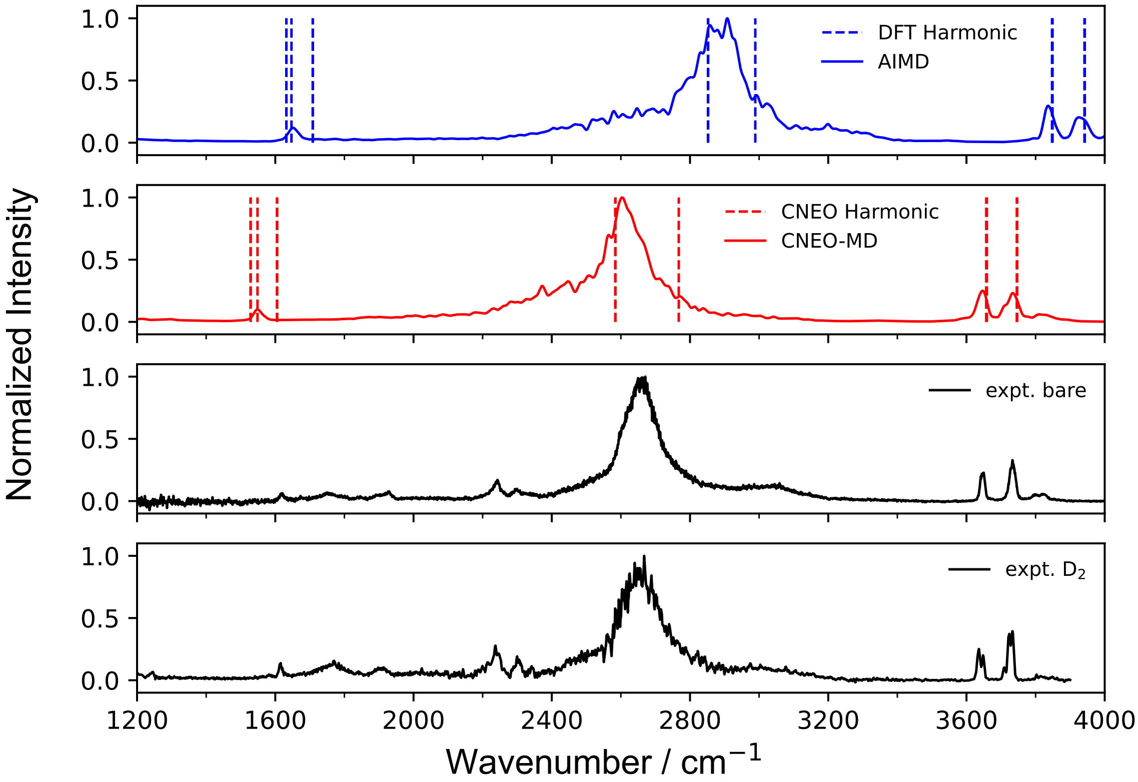


Figure 6: Overview of IR spectra of $H_9O_4^+$ obtained by the indicated methods. The experimental IR spectrum of bare cluster is from ref 12 and the experimental IR spectrum of D_2-tagged cluster is from ref 81.

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

peak in both experimental spectra.[12,81] With several high level methods, this feature has been attributed to various overtones and combination bands, especially the combination band involving the O-O stretch mode.[12,82] AIMD seems to have fortuitously captured some diffuse feature at around $3200\,\mathrm{cm}^{-1}$, but far from the experimental reference, and for CNEO-MD, this feature is not observed in the spectrum. The reason is that although CNEO-MD accounts for quantum nuclear delocalization effects, it is still based on classical MD, and it is known in the past that the accurate description of overtones, combination bands, and Fermi resonances have been particularly challenging for classical MD.[32] Therefore, these features should not be anticipated from either AIMD or CNEO-MD, although some Fermi resonance has been fortunately captured by CNEO-MD in the past.[37] The frequencies of free O-H stretches can be accurately described with both CNEO harmonic analysis and CNEO-MD, which is consistent with the results from all previous investigations.

Finally, we comment that despite their geometric similarity, the two protonated water clusters have very different vibrational features, especially in the dominant bands featuring bound O-H stretches. The $H_7O_3^+$ spectrum is relatively more complex, and CNEO harmonic analysis greatly overestimates its bound O-H stretch frequency, whereas the $H_9O_4^+$ spectrum features an intense band associated with the asymmetric bound O-H stretches, and CNEO harmonic analysis gives a reasonable prediction of the frequency. We found that CNEO-MD improves much over the poor CNEO harmonic results for $H_7O_3^+$ but only makes a small correction on top of the already good CNEO harmonic results for $H_9O_4^+$. Thus, we can see that CNEO-MD can correctly incorporate the coupling effects between different vibrational modes, whether they are strong or weak.

5 SUMMARY

In summary, we introduced the direct *NVE* scheme to efficiently compute vibrational spectra by directly running microcanonical ensemble simulations. We proved that under harmonic

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 ORCID: https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. License: CC BY 4.0

approximations, the new scheme can obtain essentially the same position and dipole autocorrelation function as conventional *NVT-NVE* and direct *NVT* schemes. With the gas phase methanol molecule as an example, we showed that the direct *NVE* scheme can reach spectrum convergence faster with less total simulation time and thus is much more efficient for vibrational spectrum calculations for gas phase clusters. In combination with the direct *NVE* scheme, we demonstrated again the power of CNEO-MD for accurate and efficient vibrational spectrum calculations. With three highly challenging anharmonic water clusters, H_4O_2, $H_7O_3^+$ and $H_9O_4^+$, as examples, we showed that CNEO-DFT harmonic analysis improves over DFT harmonic analysis with the incorporation of quantum nuclear delocalization effects, and CNEO-MD further improves the results with the incorporation of mode coupling effects. In contrast, DFT-based AIMD may not necessarily improve over DFT harmonic analysis due to the lack of quantum nuclear delocalization effects. Therefore, both quantum nuclear delocalization effects and mode coupling effects are major reasons for the excellent performance of CNEO-MD. As a cost-effective method based on classical simulations, CNEO-MD in combination with the direct *NVE* scheme holds significant promise as a robust tool for conducting vibrational spectrum simulations in more challenging systems.

Acknowledgement

The authors thank Nan Yang and Mark A. Johnson for providing experimental spectra of protonated water clusters studied in this paper, and Yunrui Qiu, Zhongyi Wan and Tanner Culpitt for helpful discussions. The authors are grateful for the funding support from the National Science Foundation under Grant No. 2238473 and from the University of Wisconsin via the Wisconsin Alumni Research Foundation. The computational resource support from Center for High Throughput Computing at the University of Wisconsin-Madison is acknowledged.[83]

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

Supporting Information Available

Detailed derivation of the direct *NVE* scheme; Structures of water clusters; benchmark of different electronic functionals on H_4O_2; temperature dependence of H_4O_2 spectra; spectrum convergence test

References

(1) Stuart, B. H. *Infrared spectroscopy: fundamentals and applications*; 2004.

(2) Barth, A. Infrared spectroscopy of proteins. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **2007**, *1767*, 1073–1101.

(3) Baker, M. J.; Trevisan, J.; Bassan, P.; Bhargava, R.; Butler, H. J.; Dorling, K. M.; Fielden, P. R.; Fogarty, S. W.; Fullwood, N. J.; Heys, K. A.; others Using Fourier transform IR spectroscopy to analyze biological materials. *Nature protocols* **2014**, *9*, 1771–1791.

(4) Baiz, C. R.; Błasiak, B.; Bredenbeck, J.; Cho, M.; Choi, J.-H.; Corcelli, S. A.; Dijkstra, A. G.; Feng, C.-J.; Garrett-Roe, S.; Ge, N.-H.; others Vibrational spectroscopic map, vibrational spectroscopy, and intermolecular interaction. *Chemical reviews* **2020**, *120*, 7152–7218.

(5) Bowman, J. M. The self-consistent-field approach to polyatomic vibrations. *Accounts of Chemical Research* **1986**, *19*, 202–208.

(6) Bowman, J. M.; Carter, S.; Huang, X. MULTIMODE: a code to calculate rovibrational energies of polyatomic molecules. *International Reviews in Physical Chemistry* **2003**, *22*, 533–549.

(7) Vendrell, O.; Gatti, F.; Meyer, H.-D. Full dimensional (15-dimensional) quantum-

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

dynamical simulation of the protonated water dimer. II. Infrared spectrum and vibrational dynamics. *The Journal of chemical physics* **2007**, *127*, 184303.

(8) Reynolds, P. J.; Ceperley, D. M.; Alder, B. J.; Lester Jr, W. A. Fixed-node quantum Monte Carlo for molecules. *The Journal of Chemical Physics* **1982**, *77*, 5593–5603.

(9) McCoy, A. B. Diffusion Monte Carlo approaches for investigating the structure and vibrational spectra of fluxional systems. *International Reviews in Physical Chemistry* **2006**, *25*, 77–107.

(10) Duong, C. H.; Gorlova, O.; Yang, N.; Kelleher, P. J.; Johnson, M. A.; McCoy, A. B.; Yu, Q.; Bowman, J. M. Disentangling the Complex Vibrational Spectrum of the Protonated Water Trimer, H+(H2O)3, with Two-Color IR-IR Photodissociation of the Bare Ion and Anharmonic VSCF/VCI Theory. *The Journal of Physical Chemistry Letters* **2017**, *8*, 3782–3789.

(11) Yu, Q.; Bowman, J. M. High-Level Quantum Calculations of the IR Spectra of the Eigen, Zundel, and Ring Isomers of H+(H2O)4 Find a Single Match to Experiment. *Journal of the American Chemical Society* **2017**, *139*, 10984–10987.

(12) Duong, C. H.; Yang, N.; Kelleher, P. J.; Johnson, M. A.; DiRisio, R. J.; McCoy, A. B.; Yu, Q.; Bowman, J. M.; Henderson, B. V.; Jordan, K. D. Tag-Free and Isotopomer-Selective Vibrational Spectroscopy of the Cryogenically Cooled H9O4+ Cation with Two-Color, IR–IR Double-Resonance Photoexcitation: Isolating the Spectral Signature of a Single OH Group in the Hydronium Ion Core. *The Journal of Physical Chemistry A* **2018**, *122*, 9275–9284.

(13) Schröder, M.; Gatti, F.; Lauvergnat, D.; Meyer, H.-D.; Vendrell, O. The coupling of the hydrated proton to its first solvation shell. *Nature Communications* **2022**, *13*, 6170.

(14) Cao, J.; Voth, G. A. The formulation of quantum statistical mechanics based on the

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

Feynman path centroid density. II. Dynamical properties. *The Journal of chemical physics* **1994**, *100*, 5106–5117.

(15) Jang, S.; Voth, G. A. A derivation of centroid molecular dynamics and other approximate time evolution methods for path integral centroid variables. *The Journal of chemical physics* **1999**, *111*, 2371–2384.

(16) Craig, I. R.; Manolopoulos, D. E. Quantum statistics and classical mechanics: Real time correlation functions from ring polymer molecular dynamics. *The Journal of chemical physics* **2004**, *121*, 3368–3373.

(17) Rossi, M.; Ceriotti, M.; Manolopoulos, D. E. How to remove the spurious resonances from ring polymer molecular dynamics. *The Journal of chemical physics* **2014**, *140*, 234116.

(18) Trenins, G.; Willatt, M. J.; Althorpe, S. C. Path-integral dynamics of water using curvilinear centroids. *The Journal of Chemical Physics* **2019**, *151*, 054109.

(19) Benson, R. L.; Trenins, G.; Althorpe, S. C. Which quantum statistics–classical dynamics method is best for water? *Faraday Discussions* **2020**, *221*, 350–366.

(20) Fletcher, T.; Zhu, A.; Lawrence, J. E.; Manolopoulos, D. E. Fast quasi-centroid molecular dynamics. *The Journal of Chemical Physics* **2021**, *155*, 231101.

(21) Witt, A.; Ivanov, S. D.; Shiga, M.; Forbert, H.; Marx, D. On the applicability of centroid and ring polymer path integral molecular dynamics for vibrational spectroscopy. *The Journal of chemical physics* **2009**, *130*, 194510.

(22) Ivanov, S. D.; Witt, A.; Shiga, M.; Marx, D. Communications: On artificial frequency shifts in infrared spectra obtained from centroid molecular dynamics: Quantum liquid water. *The Journal of chemical physics* **2010**, *132*, 031101.

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

(23) Shiga, M.; Nakayama, A. Ab initio path integral ring polymer molecular dynamics: Vibrational spectra of molecules. *Chemical Physics Letters* **2008**, *451*, 175–181.

(24) Habershon, S.; Fanourgakis, G. S.; Manolopoulos, D. E. Comparison of path integral molecular dynamics methods for the infrared absorption spectrum of liquid water. *The Journal of chemical physics* **2008**, *129*, 074501.

(25) Marsalek, O.; Markland, T. E. Quantum dynamics and spectroscopy of ab initio liquid water: The interplay of nuclear and electronic quantum effects. *The journal of physical chemistry letters* **2017**, *8*, 1545–1551.

(26) Lieberherr, A. Z.; Furniss, S. T.; Lawrence, J. E.; Manolopoulos, D. E. Vibrational strong coupling in liquid water from cavity molecular dynamics. *The Journal of Chemical Physics* **2023**, *158*, 234106.

(27) Barone, V. Anharmonic vibrational properties by a fully automated second-order perturbative approach. *The Journal of chemical physics* **2005**, *122*, 014108.

(28) Zhang, Y.; Xu, X.; Yang, N.; Chen, Z.; Yang, Y. Describing proton transfer modes in shared proton systems with constrained nuclear–electronic orbital methods. *The Journal of Chemical Physics* **2023**, *158*, 231101.

(29) Baer, M.; Marx, D.; Mathias, G. Assigning Predissociation Infrared Spectra of Microsolvated Hydronium Cations H3O+·(H2)n (n=0, 1, 2, 3) by Ab Initio Molecular Dynamics. *ChemPhysChem* **2011**, *12*, 1906–1915.

(30) Heine, N.; Fagiani, M. R.; Rossi, M.; Wende, T.; Berden, G.; Blum, V.; Asmis, K. R. Isomer-selective detection of hydrogen-bond vibrations in the protonated water hexamer. *Journal of the American Chemical Society* **2013**, *135*, 8266–8273.

(31) Wang, H.; Agmon, N. Reinvestigation of the Infrared Spectrum of the Gas-Phase Protonated Water Tetramer. *The Journal of Physical Chemistry A* **2017**, *121*, 3056–3070.

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

(32) Thomas, M.; Brehm, M.; Fligg, R.; Vöhringer, P.; Kirchner, B. Computing vibrational spectra from ab initio molecular dynamics. *Physical Chemistry Chemical Physics* **2013**, *15*, 6608–6622.

(33) Thomas, M. *Theoretical Modeling of Vibrational Spectra in the Liquid Phase*; Springer, 2016.

(34) Xu, X.; Yang, Y. Constrained nuclear-electronic orbital density functional theory: Energy surfaces with nuclear quantum effects. *The Journal of Chemical Physics* **2020**, *152*, 084107.

(35) Xu, X.; Yang, Y. Full-quantum descriptions of molecular systems from constrained nuclear–electronic orbital density functional theory. *The Journal of Chemical Physics* **2020**, *153*, 074106.

(36) Chen, Z.; Yang, Y. Incorporating nuclear quantum effects in molecular dynamics with a constrained minimized energy surface. *The Journal of Physical Chemistry Letters* **2023**, *14*, 279–286.

(37) Xu, X.; Chen, Z.; Yang, Y. Molecular dynamics with constrained nuclear electronic orbital density functional theory: Accurate vibrational spectra from efficient incorporation of nuclear quantum effects. *Journal of the American Chemical Society* **2022**, *144*, 4039–4046.

(38) Gordon, R. Molecular motion in infrared and Raman spectra. *The Journal of Chemical Physics* **1965**, *43*, 1307–1312.

(39) Berne, B. J.; Harp, G. On the calculation of time correlation functions. *Advances in chemical physics* **1970**, 63–227.

(40) Tuckerman, M. *Statistical Mechanics: Theory and Molecular Simulation*; OUP Oxford, 2010.

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

(41) Wang, Y.; Chen, Z.; Yang, Y. Calculating vibrational excited state absorptions with excited state constrained minimized energy surfaces. *The Journal of Physical Chemistry A* **2023**, *127*, 5491–5501.

(42) Xu, X. Constrained Nuclear-Electronic Orbital Density Functional Theory with a Dielectric Continuum Solvent Model. *The Journal of Physical Chemistry A* **2023**, *127*, 6329–6334.

(43) Taherivardanjani, S.; Elfgen, R.; Reckien, W.; Suarez, E.; Perlt, E.; Kirchner, B. Benchmarking the Computational Costs and Quality of Vibrational Spectra from Ab Initio Simulations. *Advanced Theory and Simulations* **2022**, *5*, 2100293.

(44) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Physical review letters* **1996**, *77*, 3865.

(45) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *The Journal of chemical physics* **1999**, *110*, 6158–6170.

(46) Bartlett, R. J. Many-body perturbation theory and coupled cluster theory for electron correlation in molecules. *Annual review of physical chemistry* **1981**, *32*, 359–401.

(47) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chemical Physics Letters* **1989**, *157*, 479–483.

(48) Mardirossian, N.; Head-Gordon, M. ωB97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *The Journal of chemical physics* **2016**, *144*, 214110.

(49) Vydrov, O. A.; Van Voorhis, T. Nonlocal van der Waals density functional: The simpler the better. *The Journal of chemical physics* **2010**, *133*, 244103.

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

(50) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Physical Chemistry Chemical Physics* **2005**, *7*, 3297–3305.

(51) Rappoport, D.; Furche, F. Property-optimized Gaussian basis sets for molecular response calculations. *The Journal of chemical physics* **2010**, *133*, 134105.

(52) Yu, Q.; Pavošević, F.; Hammes-Schiffer, S. Development of nuclear basis sets for multicomponent quantum chemistry methods. *The Journal of Chemical Physics* **2020**, *152*, 244123.

(53) Yang, Y.; Brorsen, K. R.; Culpitt, T.; Pak, M. V.; Hammes-Schiffer, S. Development of a practical multicomponent density functional for electron-proton correlation to produce accurate proton densities. *The Journal of Chemical Physics* **2017**, *147*, 114113.

(54) Brorsen, K. R.; Yang, Y.; Hammes-Schiffer, S. Multicomponent density functional theory: Impact of nuclear quantum effects on proton affinities and geometries. *The Journal of Physical Chemistry Letters* **2017**, *8*, 3488–3493.

(55) Tao, Z.; Yang, Y.; Hammes-Schiffer, S. Multicomponent density functional theory: Including the density gradient in the electron-proton correlation functional for hydrogen and deuterium. *The Journal of Chemical Physics* **2019**, *151*, 124102.

(56) Xu, X.; Yang, Y. Molecular vibrational frequencies from analytic Hessian of constrained nuclear–electronic orbital density functional theory. *The Journal of Chemical Physics* **2021**, *154*, 244110.

(57) Available at https://github.com/theorychemyang/pyscf.

(58) Sun, Q.; Berkelbach, T. C.; Blunt, N. S.; Booth, G. H.; Guo, S.; Li, Z.; Liu, J.; McClain, J. D.; Sayfutyarova, E. R.; Sharma, S.; Wouters, S.; Chan, G. K.-L. PySCF: the

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

Python-based simulations of chemistry framework. *WIREs Computational Molecular Science* **2018**, *8*, e1340.

(59) Sun, Q.; Zhang, X.; Banerjee, S.; Bao, P.; Barbry, M.; Blunt, N. S.; Bogdanov, N. A.; Booth, G. H.; Chen, J.; Cui, Z.-H.; Eriksen, J. J.; Gao, Y.; Guo, S.; Hermann, J.; Hermes, M. R.; Koh, K.; Koval, P.; Lehtola, S.; Li, Z.; Liu, J.; Mardirossian, N.; McClain, J. D.; Motta, M.; Mussard, B.; Pham, H. Q.; Pulkin, A.; Purwanto, W.; Robinson, P. J.; Ronca, E.; Sayfutyarova, E. R.; Scheurer, M.; Schurkus, H. F.; Smith, J. E. T.; Sun, C.; Sun, S.-N.; Upadhyay, S.; Wagner, L. K.; Wang, X.; White, A.; Whitfield, J. D.; Williamson, M. J.; Wouters, S.; Yang, J.; Yu, J. M.; Zhu, T.; Berkelbach, T. C.; Sharma, S.; Sokolov, A. Y.; Chan, G. K.-L. Recent developments in the PySCF program package. *The Journal of Chemical Physics* **2020**, *153*, 024109.

(60) Larsen, A. H.; Mortensen, J. J.; Blomqvist, J.; Castelli, I. E.; Christensen, R.; Dułak, M.; Friis, J.; Groves, M. N.; Hammer, B.; Hargus, C.; Hermes, E. D.; Jennings, P. C.; Jensen, P. B.; Kermode, J.; Kitchin, J. R.; Kolsbjerg, E. L.; Kubal, J.; Kaasbjerg, K.; Lysgaard, S.; Maronsson, J. B.; Maxson, T.; Olsen, T.; Pastewka, L.; Peterson, A.; Rostgaard, C.; Schiøtz, J.; Schütt, O.; Strange, M.; Thygesen, K. S.; Vegge, T.; Vilhelmsen, L.; Walter, M.; Zeng, Z.; Jacobsen, K. W. The atomic simulation environment—a Python library for working with atoms. *Journal of Physics: Condensed Matter* **2017**, *29*, 273002.

(61) West, D.; Estreicher, S. First-principles calculations of vibrational lifetimes and decay channels: Hydrogen-related modes in Si. *Physical review letters* **2006**, *96*, 115504.

(62) Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics* **1984**, *81*, 511–519.

(63) Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A* **1985**, *31*, 1695–1697.

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

(64) Martyna, G. J.; Klein, M. L.; Tuckerman, M. Nosé–Hoover chains: The canonical ensemble via continuous dynamics. *The Journal of chemical physics* **1992**, *97*, 2635–2643.

(65) Cooke, B.; Schmidler, S. C. Preserving the Boltzmann ensemble in replica-exchange molecular dynamics. *The Journal of chemical physics* **2008**, *129*, 164112.

(66) Andersen, H. C. Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of chemical physics* **1980**, *72*, 2384–2393.

(67) Johnson, R. Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database 101. 2002; `http://cccbdb.nist.gov/`.

(68) Huisken, F.; Kaloudis, M.; Kulcke, A. Infrared spectroscopy of small size-selected water clusters. *The Journal of Chemical Physics* **1996**, *104*, 17–25.

(69) Low, G. R.; Kjaergaard, H. G. Calculation of OH-stretching band intensities of the water dimer and trimer. *The Journal of chemical physics* **1999**, *110*, 9104–9115.

(70) Dunn, M. E.; Evans, T. M.; Kirschner, K. N.; Shields, G. C. Prediction of Accurate Anharmonic Experimental Vibrational Frequencies for Water Clusters, (H2O)n, n = 2-5. *The Journal of Physical Chemistry A* **2006**, *110*, 303–309.

(71) Kjaergaard, H. G.; Garden, A. L.; Chaban, G. M.; Gerber, R. B.; Matthews, D. A.; Stanton, J. F. Calculation of vibrational transition frequencies and intensities in water dimer: Comparison of different vibrational approaches. *The Journal of Physical Chemistry A* **2008**, *112*, 4324–4335.

(72) Kuyanov-Prozument, K.; Choi, M. Y.; Vilesov, A. F. Spectrum and infrared intensities of OH-stretching bands of water dimers. *The Journal of chemical physics* **2010**, *132*, 014304.

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

(73) Perlt, E.; Berger, S. A.; Kelterer, A.-M.; Kirchner, B. Anharmonicity of Vibrational Modes in Hydrogen Chloride–Water Mixtures. *Journal of Chemical Theory and Computation* **2019**, *15*, 2535–2547.

(74) Zhang, B.; Yu, Y.; Zhang, Z.; Zhang, Y.-Y.; Jiang, S.; Li, Q.; Yang, S.; Hu, H.-S.; Zhang, W.; Dai, D.; others Infrared spectroscopy of neutral water dimer based on a tunable vacuum ultraviolet free electron laser. *The Journal of Physical Chemistry Letters* **2020**, *11*, 851–855.

(75) Vogt, E.; Kjaergaard, H. G. Vibrational spectroscopy of the water dimer at jet-cooled and atmospheric temperatures. *Annual Review of Physical Chemistry* **2022**, *73*, 209–231.

(76) Mackeprang, K.; Kjaergaard, H. G.; Salmi, T.; Hänninen, V.; Halonen, L. The effect of large amplitude motions on the transition frequency redshift in hydrogen bonded complexes: A physical picture. *The Journal of Chemical Physics* **2014**, *140*, 184309.

(77) Mackeprang, K.; Hänninen, V.; Halonen, L.; Kjaergaard, H. G. The effect of large amplitude motions on the vibrational intensities in hydrogen bonded complexes. *The Journal of Chemical Physics* **2015**, *142*, 094304.

(78) Langford, J.; Zhang, Y.; Chen, Z.; Yang, Y. Where is the Hidden Intramolecular H-bonding Vibrational Signal in Proline? **2023**, 10.26434/chemrxiv–2023–tp95r.

(79) DiRisio, R. J.; Finney, J. M.; Dzugan, L. C.; Madison, L. R.; McCoy, A. B. Using Diffusion Monte Carlo Wave Functions to Analyze the Vibrational Spectra of H7O3+ and H9O4+. *The Journal of Physical Chemistry A* **2021**, *125*, 7185–7197.

(80) Yu, Q.; Bowman, J. M. Classical, Thermostated Ring Polymer, and Quantum VSCF/VCI Calculations of IR Spectra of H7O3+ and H9O4+ (Eigen) and Comparison with Experiment. *The Journal of Physical Chemistry A* **2019**, *123*, 1399–1409.

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

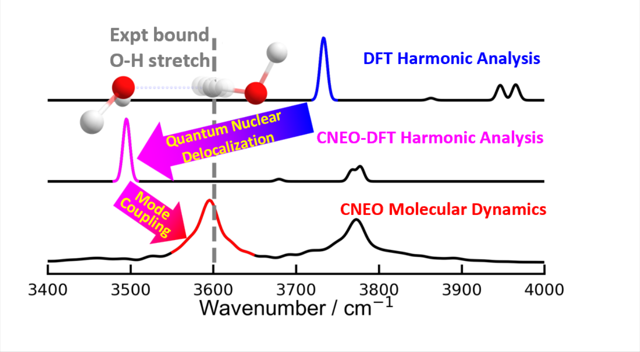
(81) Wolke, C. T.; Fournier, J. A.; Dzugan, L. C.; Fagiani, M. R.; Odbadrakh, T. T.; Knorke, H.; Jordan, K. D.; McCoy, A. B.; Asmis, K. R.; Johnson, M. A. Spectroscopic snapshots of the proton-transfer mechanism in water. *Science* **2016**, *354*, 1131–1135.

(82) Yagi, K.; Thomsen, B. Infrared spectra of protonated water clusters, H+ (H2O) 4, in Eigen and Zundel forms studied by vibrational quasi-degenerate perturbation theory. *The Journal of Physical Chemistry A* **2017**, *121*, 2386–2398.

(83) Center for High Throughput Computing Center for High Throughput Computing. 2006; `https://chtc.cs.wisc.edu/`.

https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

TOC Graphic



https://doi.org/10.26434/chemrxiv-2023-0lb60-v2 **ORCID:** https://orcid.org/0000-0001-8572-5155 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0