Efficient, "On-the-Fly", Born–Oppenheimer and Car–Parrinello-type Dynamics with Coupled Cluster Accuracy through Fragment Based Electronic Structure

Cody Haycraft, Junjie Li, and Srinivasan S. Iyengar*

Department of Chemistry and Department of Physics, Indiana University, 800 East Kirkwood Avenue, Bloomington, Indiana 47405, United States

Ⓢ Supporting Information

ABSTRACT: We recently developed two fragment based *ab initio* molecular dynamics methods, and in this publication we have demonstrated both approaches by constructing efficient classical trajectories in agreement with trajectories obtained from "on-the-fly" CCSD. The dynamics trajectories are obtained using both Born–Oppenheimer and extended Lagrangian (Car–Parrinello-style) options, and hence, here, for the first time, we present Car–Parrinello-like AIMD trajectories that are accurate to the CCSD level of post-Hartree–Fock theory. The specific extended Lagrangian implementation used here is a generalization to atom-centered density matrix propagation (ADMP) that provides post-Hartree–Fock accuracy, and hence

the new method is abbreviated as ADMP-pHF; whereas the Born–Oppenheimer version is called frag-BOMD. The fragmentation methodology is based on a set-theoretic, inclusion-exclusion principle based generalization of the well-known ONIOM method. Thus, the fragmentation scheme contains multiple overlapping "model" systems, and overcounting is compensated through the inclusion-exclusion principle. The energy functional thus obtained is used to construct Born–Oppenheimer forces (frag-BOMD) and is also embedded within an extended Lagrangian (ADMP-pHF). The dynamics is tested by computing structural and vibrational properties for protonated water clusters. The frag-BOMD trajectories yield structural and vibrational properties in excellent agreement with full CCSD-based "on-the-fly" BOMD trajectories, at a small fraction of the cost. The asymptotic (large system) computational scaling of both frag-BOMD and ADMP-pHF is inferred as $O(N^{3.5})$, for on-the-fly CCSD accuracy. The extended Lagrangian implementation, ADMP-pHF, also provides structural features in excellent agreement with full "on-the-fly" CCSD calculations, but the dynamical frequencies are slightly red-shifted. Furthermore, we study the behavior of ADMP-pHF as a function of the electronic inertia tensor and find a monotonic improvement in the red-shift as we reduce the electronic inertia. In all cases a uniform spectral scaling factor, that in our preliminary studies appears to be independent of system and independent of level of theory (same scaling factor for both MP2 and CCSD implementations ADMP-pHF and for ADMP DFT), improves on agreement between ADMP-pHF and full CCSD calculations. Hence, we believe both frag-BOMD and ADMP-pHF will find significant utility in modeling complex systems. The computational power of frag-BOMD and ADMP-pHF is demonstrated through preliminary studies on a much larger protonated 21-water cluster, for which AIMD trajectories with "on-the-fly" CCSD are not feasible.
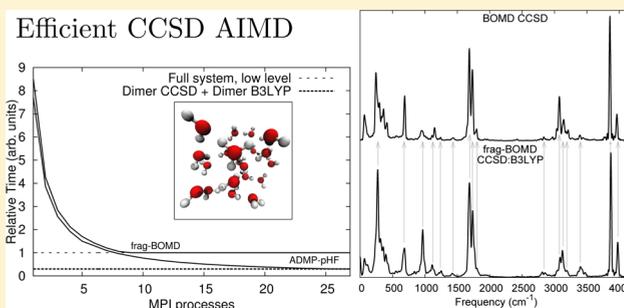
I. INTRODUCTION

Ab initio molecular dynamics (AIMD)[1,2] is a central tool in the study of reactive processes and is used, for example, to compute vibrational spectral properties beyond harmonic approximation and for biochemical studies when these methods are combined with hybrid (QM/MM and QM/QM) calculations. Here, the instantaneous electronic energy and forces determine nuclear motion, and hence AIMD is much more computationally intensive compared to force-field based, empirical molecular dynamics and electronic structure applications such as geometry optimization and frequency calculations. As a result, AIMD methods are restricted in their applications, and, apart from the case for a few small to medium sized systems where

one can afford post-Hartree–Fock studies, most are typically based on "on-the-fly" DFT.[3−7] Unfortunately, despite the large volume of literature on DFT based electronic structure, the method is still restricted in many ways,[8−10] and post-Hartree–Fock methods such as coupled cluster and MP2 are desirable but clearly beyond existing capabilities.

One alternative to this dilemma has recently emerged in the literature through the advent of fragment based electronic structure methods.[11−32] Details can be found in several review articles.[33−37] Essentially there are two related sets of ideas here:
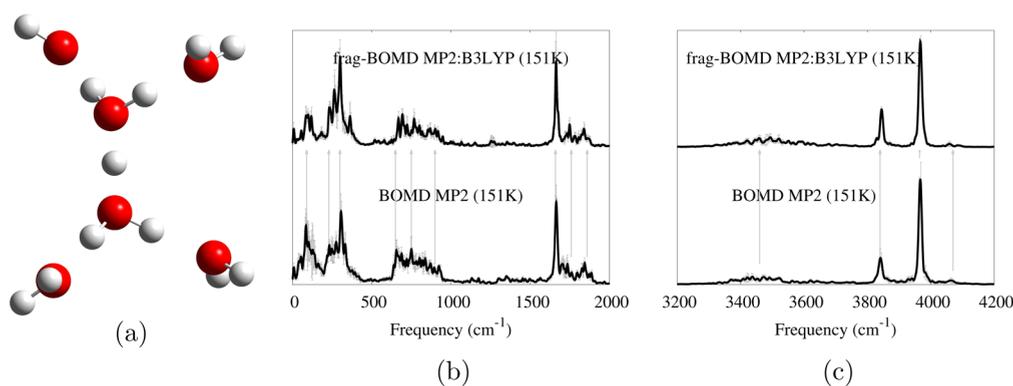
Figure 1. Vibrational density of states computed from dynamics trajectories at 150 K for $H_{13}O_6^+$. Two different spectral ranges are shown, and the agreement is highlighted through gray arrows. See ref 38.

In one case a system is subdivided into fragments, and the overall energy and properties are computed by assembling the results from individual fragment calculations. The second idea is based on many-body expansion, where the overall energy is written as a combination of one-body, two-body, and higher order interaction terms when necessary. In both cases, overlapping fragments have been shown to be treated in a consistent manner, and computational advantages obviously arise from the need to only compute fragments of the full system at higher levels of electronic structure theory. These methods have been used to compute a number of structural parameters that are in agreement with high level electronic structure treatment.

While there have been some efforts to combine these methods with dynamics,[29,30,39−41] in recent publications,[38,42] we have developed a method that utilizes fragment based electronic energy and gradients to compute the on-the-fly electronic energy and gradients required during AIMD. The approach has both Born–Oppenheimer dynamics[42] and extended Lagrangian[38] flavors and has also been shown to provide good quality potential surfaces.[42] It has been demonstrated through the study of small to medium sized protonated water clusters and polypeptide fragments with accuracy comparison of structural distribution functions and dynamical time-correlation functions with more expensive post-Hartree−Fock calculations, and the agreement is near perfect! (See Figure 1.) The approach is part of a C++ module of computer programs that work in parallel through a combination of MPI and OpenMP protocols that invoke electronic structure packages. While the current implementation invokes the Gaussian series of electronic structure programs, we are also developing a new platform that will allow the simultaneous utilization of multiple electronic structure codes within a single dynamics calculation.

In this publication we show that the approach can be used to compute accurate coupled cluster trajectories at very low cost. We compute ab initio molecular dynamics trajectories for the protonated water clusters, $H_9O_4^+$ and $H_{13}O_6^+$, and obtain vibrational properties for these systems. For $H_9O_4^+$ we also compute the AIMD trajectories with "on-the-fly" CCSD calculations of energy and forces, and these calculations are extremely demanding. The agreement with the fragment-based calculations is very good as will be seen later in this publications, and as expected the computational cost is negligible compared to the full calculation (which, in general, is prohibitive). Funneling from this agreement, we compute

fragment based CCSD calculations for the larger protonated water cluster where the full CCSD calculations are impossible to compute in a dynamical fashion. This system is chosen to demonstrate the power of potential application from this approach, but the general choice of systems, protonated water clusters, is particularly significant. These systems have proven a great challenge to theory in terms of obtaining vibrational characteristics.[43−50] This has been thought to be due to the following: (a) the anharmonicity of the proton stretch leads to nontrivial and unexpected coupling between the harmonic states that has led to discrepancies in the past between the measured and high-level computed spectra,[51−54] and (b), as a result of point (a), the shared proton tends to have a delocalized distribution function, that may require quantum nuclear treatment,[55−57] although *ab initio* molecular dynamics has been useful in modeling the behavior of such shared H/D hydrogen bonded systems.[54]

The paper is organized as follows: In Section II, accompanied by the Supporting Information, we present a brief overview of the theoretical formalism with details in refs 38 and 42. Results are provided in Section III. Specifically, the stability of the dynamics simulations is discussed in Section III; structural distribution functions computed from all dynamics trajectories are presented in Section IIIA, and vibrational density of states is provided in Section IIIB. Computational efficiency and implementational aspects are discussed in Section IIA and Section IV followed by conclusions in Section V.

II. FRAGMENT BASED *AB INITIO* MOLECULAR DYNAMICS: BORN-OPPENHEIMER AND EXTENDED LAGRANGIAN IMPLEMENTATIONS

Our fragment based AIMD procedure has been discussed in previous publications[38,42], and the key aspects are presented here. A system may be partitioned into overlapping regions as shown in Figure 2, and the overall molecular energy expression[42] is assembled in accordance with the set-theoretic principle of inclusion-exclusion[58] as

$$E^{PIE-ONIOM} \equiv E_{low}(0) + \sum_{i=1}^{n} \mathcal{S}(i) - \sum_{1<=i<j<=n} \mathcal{S}(i \cap j)$$
$$+ \sum_{1<=i<j<k<=n} \mathcal{S}(i \cap j \cap k) - \cdots + (-1)^{n-1} \sum \mathcal{S}(1 \cap \cdots \cap n)$$

$$(1)$$

That is, as illustrated in Figure 2, the energy is first computed at a lower level of theory, and this energy is represented in eq 1 as $E_{low}(0)$. The energy is then corrected for every possible

$$E^{\text{PIE-ONIOM}} = E^{\text{low}}_{A \cup B \cup C} + S_A + S_B + S_C$$
$$- S_{A \cap B} - S_{B \cap C} - S_{A \cap C} + S_{A \cap B \cap C}$$
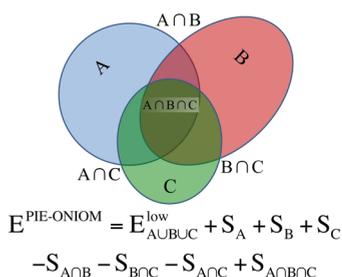
Figure 2. Illustration of the energy expression in eq 1.

molecular fragment where the extrapolation term, $\mathcal{S}(i)$, for two levels of theory is defined as

$$\mathcal{S}(i) = E_{\text{high}}(i) - E_{\text{low}}(i) \tag{2}$$

Thus, the fragment contributions to $E_{\text{low}}(0)$ are improved by adding in the difference between the energy at a higher level of theory, $E_{\text{high}}(i)$, and the lower level of theory, $E_{\text{low}}(i)$. This idea for extrapolation is based on the well-known ONIOM scheme[12] but is generalized using the set-theoretic inclusion-exclusion (PIE) principle. As a result, the expression in eq 1 is abbreviated as $E^{\text{PIE-ONIOM}}$. Similarly, the term $\mathcal{S}(i \cap j)$ extrapolates the energy for fragment $(i \cap j)$ that is obtained from the intersection of fragment numbers, i and j. As is evident, each fragment is treated at two levels of theory, and the entire system (the zeroth fragment) is only considered at the lowest level of theory. The approach may be generalized to more levels as outlined in ref 42, but this generalization is not considered here.

A powerful feature of our approach is that the identity of the overlapping fragments is evaluated using a bit-manipulation algorithm[42] that reconfigures the fragments during dynamics. Fragment identities are stored as bits in a large integer, and all possible sets of fragments are computed using a sequence of bit-shifting, bitwise symmetric ORs and two's complement operations,[42] and overlapping regions between fragments are computed using binary AND operations. The use of bitwise arithmetic improves efficiency tremendously in this otherwise exponential scaling bottleneck, but a large integer library is necessary for large systems. When recomputing the fragment definition is not required, as might be the case for several bonded systems, a user defined fragment topology is retained all through dynamics. Problems associated with dynamical fragmentation are not considered in this publication.[41,59,60]

Another important feature that is already present in eq 1 and needs emphasis is that $E_{\text{low}}(0)$, $E_{\text{low}}(i)$, and $E_{\text{high}}(i)$ can refer to two different levels of theories, two different basis sets for the same level of theory (thus treating on-the-fly basis set extrapolation), or two different numerical approximations for the same level of theory. The ideas presented below are independent of these choices, and such choices have been explored by the ONIOM community, which the fragment-based extrapolation method here is based on. Furthermore, more levels can also be included in the treatment[42] in an adaptive fashion, but these aspects are not considered in this publication. Here we are concerned with providing "on-the-fly" CCSD accuracy at low computational cost.

From the energy expression in eq 1 one can write equations of motion for the nuclei using gradients, and conservative, accurate Born–Oppenheimer dynamics based on this idea has been achieved in refs 38 and 42. However, for large systems the

computational expense may be dictated by the calculation of $E_{\text{low}}(0)$. Thus, in ref 38 the electronic degrees of freedom that influence $E_{\text{low}}(0)$ are directly propagated using an extended Lagrangian formalism for cases where $E_{\text{low}}(0)$ is obtained from a single particle formalism, such as Hartree–Fock or DFT. In this method, the system is retained in its ground electronic state through a Car–Parrinello-style dynamics of the single-particle electronic density matrix propagated in-step with the nuclei. As such, a fully converged SCF at each step is then avoided during the calculation of $E_{\text{low}}(0)$. The Hamiltonian that governs such a dynamics is given by

$$\mathcal{H} = \frac{1}{2}\text{Tr}[\mathbf{V^T M^{-1} V}] + \frac{1}{2}\text{Tr}\left[(\boldsymbol{\mu}_{\text{low},0}^{-1/4}\mathbf{W_{\text{low},0}}\boldsymbol{\mu}_{\text{low},0}^{-1/4})^2\right]$$
$$+ E^{\text{PIE-ONIOM}}(\mathbf{R}, \mathbf{P_{\text{low},0}}) + \text{Tr}[\boldsymbol{\Lambda_{\text{low},0}}(\mathbf{P_{\text{low},0}^2} - \mathbf{P_{\text{low},0}})] \tag{3}$$

and the associated Lagrangian obtained from a Legendre transform[61] is

$$\mathcal{L} = \frac{1}{2}\text{Tr}[\mathbf{V^T M V}] + \frac{1}{2}\text{Tr}\left[(\boldsymbol{\mu}_{\text{low},0}^{1/4}\mathbf{W_{\text{low},0}}\boldsymbol{\mu}_{\text{low},0}^{1/4})^2\right]$$
$$- E^{\text{PIE-ONIOM}}(\mathbf{R}, \mathbf{P_{\text{low},0}}) - \text{Tr}[\boldsymbol{\Lambda_{\text{low},0}}(\mathbf{P_{\text{low},0}^2} - \mathbf{P_{\text{low},0}})] \tag{4}$$

where the quantities, $\mathbf{R}$, $\mathbf{V}$, and $\mathbf{M}$ are the nuclear coordinates, velocities, and mass respectively, and the single particle density matrix, $\mathbf{P_{\text{low},0}}$, which determines $E_{\text{low}}(0)$ (the first term in eq 1), has fictitious inertia tensor $\boldsymbol{\mu_{\text{low},0}}$ and velocity $\mathbf{W_{\text{low},0}}$. The quantity $\boldsymbol{\Lambda_{\text{low},0}}$ is a Lagrangian multiplier that maintains the N-representability of $\mathbf{P_{\text{low},0}}$. The energy functional in eq 1 is thus augmented here through propagation of $\mathbf{P_{\text{low},0}}$ which leads to the additional dependence, $E_{\text{low}}(0) \equiv E_{\text{low}}(0, \mathbf{P_{\text{low},0}})$. Since we use the single-particle density matrix as a dynamical variable, represented using an atom-centered Gaussian basis set, this idea is along the lines of the atom-centered density matrix (ADMP)[62−65] formalism but with post-Hartree−Fock theory as afforded through the use of the composite energy functional in eq 1. As a result, the dynamical approach derived from eqs 3 and 4 is called ADMP-pHF. The corresponding Euler−Lagrange equations of motion are derived in ref 38 along with the demonstration of conservative, accurate, and efficient *ab initio* molecular dynamics. (Born−Oppenheimer dynamics trajectories directly constructed using the energy expressed in eq 1 are referred to as frag-BOMD.) The associated equations of motion in each case are integrated using the velocity Verlet algorithm.[66] The equations of motion, nuclear, and density matrix forces, including additional terms to nuclear forces arising in ADMP-pHF due to nonconvergence of $\mathbf{P_{\text{low},0}}$, are all described in ref 38 and are also summarized for completeness in the Supporting Information. The fictitious inertia tensor for valence electrons is fixed at a specific scalar value, μ_{valence}, and core orbitals are weighted using μ_{valence} and diagonal elements of the single-particle Fock matrix, $\mathbf{F_{\text{low},0}}$, which determines $E_{\text{low}}(0)$.[62,63] Specifically the elements of $\boldsymbol{\mu_{\text{low},0}}$ are represented using a diagonal mass-weighting matrix, $\mathbf{A_{\text{low},0}}$, with elements defined as

$$\mathbf{A}_{\text{low},0}^{i,i} = \mathbf{I}, \qquad\qquad \mathbf{F}_{\text{low},0}^{ii} > -2 \text{ au}$$
$$\mathbf{A}_{\text{low},0}^{i,i} = [2|\mathbf{F}_{\text{low},0}^{ii} + 2|^{1/2} + 1]^2, \quad \mathbf{F}_{\text{low},0}^{ii} < -2 \text{ au}$$
$$\mathbf{A}_{\text{low},0}^{i,j} = 0, \qquad\qquad i \neq j \tag{5}$$

and

$$\mu_{\text{low},0} \equiv \mu_{\text{valence}} \mathbf{A}_{\text{low},0} \tag{6}$$

where $\mathbf{F}_{\text{low},0}{}^{ii}$ are diagonal elements of the Fock matrix representing $E_{\text{low}}(0)$. This setup has been widely used for many single particle ADMP applications[42,51–54,64,67–76] and for ADMP-pHF in ref 38. Other options for the fictitious inertia tensor are possible but are not tested here. As seen from the Supporting Information, the ADMP-pHF nuclear forces have additional terms that depend on the commutator $[\mathbf{P}_{\text{low},0}, \mathbf{F}_{\text{low},0}]$, and these correct for the fact that $\mathbf{P}_{\text{low},0}$ is propagated as highlighted above, as opposed to being converged. Furthermore, based on the analysis presented in ref 65 for ADMP, we expected that the choice of fictitious inertia tensor $\mu_{\text{low},0} \propto [\mathbf{P}_{\text{low},0}, \mathbf{F}_{\text{low},0}]$.

While extended-Lagrangian methods have thus far been confined to single-particle formalisms such as Hartree−Fock or DFT, a powerful feature of ADMP-pHF is that it can include many-body, post-Hartree−Fock effects such as those in MP2 and CCSD. Hence ADMP-pHF is one recipe to generalize Car−Parrinello-style methods beyond the realm of DFT. In the next section we probe the utility of frag-BOMD and ADMP-pHF in computing accurate Coupled Cluster (CCSD) based post-Hartree−Fock AIMD trajectories for small and medium sized protonated water clusters. In Section III we compare results obtained from ADMP-pHF with (a) full BOMD constructed using "on-the-fly" CCSD for systems where such calculations can be done and (b) associated frag-BOMD trajectories. We also present an analysis of the reduced scaling afforded by both ADMP-pHF and frag-BOMD while obtaining results in reasonable agreement with BOMD CCSD; but before we undertake a full numerical comparison, we present below a brief discussion on formal scaling of our method and how additional features can be added based on existing efforts in the literature.

IIA. Formal Computational Scaling for frag-BOMD and ADMP-pHF. The power of the frag-BOMD and ADMP-pHF approaches is clear from the enormous reduction in computational effort. We first discuss the scaling in computational effort for the class of systems presented in this publication. We then present the general scaling principles for systems with *local electronic effects*. All calculations performed in this publication utilize hydrogen-bonded (neighboring) water-dimer fragments, independent of full system size. See Figure 3 where the primary fragments are explicitly noted. The identity of secondary fragments is obtained through the bit-manipulation algorithm discussed earlier and the detailed approach for fragmentation of protonated water clusters is listed in the Appendix. For a

(a)Eigen cation ($H_9O_4^+$)

(b)solvated Zundel cation
($H_{13}O_6^+$)
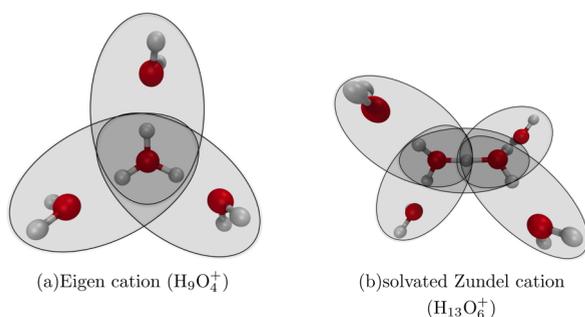
Figure 3. Two protonated water clusters are depicted here. The ellipses show primitive "dimer" fragments.

water cluster with N-water molecules all the dimer fragment CCSD and fragment DFT calculations scale as $O(4N[M^6 + M^{3.5}])$ where M is the size of the individual (dimer) fragments that undergo CCSD (and thus the M^6 contribution) and DFT (and thus the $M^{3.5}$ contribution) calculations, and the factor of $4N$ arises from the fact that each water molecule, on average, resides inside four different neighboring fragments. Calculation of energy and gradients for the monomer fragments arising from overlapping dimer fragments scale as $O(N[(M_m)^6 + (M_m)^{3.5}])$, where M_m represents the size of the overlapping monomer fragments, which in this case is approximately $M/2$. While CCSD formally scales as $O(N^6)$, note already that the CCSD scaling depends on fragment size and not system size, and for all calculation here the fragment size is a constant (water dimer). The full system low-level DFT calculation formally scales as $O(N^4)$, but for larger systems the two-electron integrals can be reused[77] which reduces scaling in practice to about $O(N^{3.5})$. Employing linear scaling methods for SCF[78–83] and SCF parallelism[78,84] will further affect the full system low-level calculations, but these developments are independent of our approach and can readily be included due to the independence of our fragment based AIMD approach on the actual electronic structure package employed to calculate energy and forces. Thus, the overall reduced asymptotic (large N) formal scaling for the CCSD:B3LYP frag-BOMD and ADMP-pHF approaches here is

$$O(4N[M^6 + M^{3.5}] + N[(M_m)^6 + (M_m)^{3.5}] + N^{3.5}) \approx O(N^{3.5})$$

to obtain CCSD accuracy. Thus, based on our results the approach here appears to be quite promising and may be able to provide CCSD accuracy for "on-the-fly" dynamics at DFT cost. (We have assumed above that the number of atoms is also proportional to the number of basis functions, which is generally the case when atom-centered Gaussian basis functions are used to describe the electronic structure.) Note that two of the water clusters treated here ($H_9O_4^+$ and $H_{13}O_6^+$) are small enough that the dimer fragments have comparable sizes. Hence the CCSD-level fragment calculations in these cases are more expensive as compared to the full-system low-level calculations. For larger water clusters, such as the protonated 21-water cluster described in Section IV, the dimer fragments are much smaller than the full system, and correspondingly the full-system DFT calculations do become the computational bottleneck as compared to the fragment high level calculations. The discussion above also assumes that the dimer fragment sufficiently describes the system. At least for the case of protonated water clusters, we find this to be true. Other systems may require additional layers of fragmentation as allowed by eq 1. For systems requiring an accurate description of nonlocal effects (such as π-conjugated systems), larger fragments may be needed, and the scaling ideas presented above will need to be modified accordingly.

Here we must note a few critical efforts in the literature. In ref 85, the author shows that the MP2 portion of gradients is a slowly varying quantity as a function of nuclear coordinates. This allows the author to construct a potentially powerful multiple time-step method.[86] We will consider similar extensions for our approach in the future. In ref 87 the authors present a mechanism where the steeper scaling of "on-the-fly" MP2 is alleviated through the use of GPUs for this portion of

the calculation. These developments are critical and complementary to our approach and can be merged with our approach with little effort. In this regard it is also critical to note the effort in ref 88 that allows AIMD on GPUs for single-particle electronic structure treatment.

III. RESULTS AND DISCUSSION

Two protonated water clusters, the Eigen cation ($H_9O_4^+$) and the solvated Zundel cation ($H_{13}O_6^+$), with significance in biological,[89−96] atmospheric,[70,76,97,98] and condensed phase chemistry are studied here. Of particular interest is the role these two structures have in bulk water and clusters,[51,69,99−108] where they exist simultaneously as competing moieties. These have been studied extensively by both experiments[43−50,109] and by other theories see refs 43, 51, 52, 55, 69, 71, 101, 102, 105, 106, 108, 110−120. Accurate treatment of hydrogen bonding is crucial to model structural and dynamical behavior, and, to that end, a "dimer" topology is employed, where primitive fragments are formed from neighboring water molecules (see Figure 1 and the Appendix for method of fragmentation). It is possible, during the course of a simulation, for the ($H_{13}O_6^+$) cluster to undergo structural rearrangement to a more "eigen-like" isomer, which would need to be accompanied by a corresponding change in the fragment topology definition. Such "topological hops" present challenges to the theory and will be addressed in future publications. For all simulations presented fragment topology is static.

A variety of simulations have been performed. The stability of dynamics and the accuracy (gauged by structural and vibrational properties) are judged for both of the aforementioned water clusters. These are investigated with (1) BOMD using pure CCSD for the Eigen cation only (abbreviated as BOMD CCSD). Pure CCSD for the solvated Zundel cation is computationally prohibitive. (2) BOMD using the two-layer PIE-ONIOM fragment-based electronic structure, with CCSD as the high level electronic structure and B3LYP as the low (abbreviated BOMD CCSD:B3LYP). (3) Extended Lagrangian dynamics is also performed at the CCSD:B3LYP level (abbreviated as ADMP-pHF CCSD:B3LYP). Initial geometries in each case were obtained from geometry optimization at the B3LYP level of theory with 6-31+g(d,p) Gaussian basis. In addition we have also studied the behavior of ADMP-pHF trajectories for multiple values of the fictitious inertia tensor in eqs 3 and 4. These are investigated with a series of simulations performed at the ADMP-pHF MP2:B3LYP level. Our computational implementation is achieved through a C++ program that works in parallel using MPI across nodes and OpenMP within each node. The electronic structure energies and nuclear gradients were obtained from the Gaussian series of electronic structure programs. It may be noted that in ref 38, we have computed MP2 trajectories and established that our fragmentation methodology provided good agreement with results from the full MP2 level of theory. (See Figure 1.) In addition, potential energy surfaces computed using the CCSD(T):B3LYP level of PIE-ONIOM fragmentation theory along normal mode coordinates for these water clusters were found to be in excellent agreement with the surface obtained from CCSD(T) calculations. (See Figure 1(b) in ref 38.)

All simulations are performed under constant total energy conditions, where the conserved Hamiltonian is given by eq 3 for ADMP-pHF simulations. For BOMD trajectories there is no constraint penalty or fictitious kinetic energies. A summary of the simulations performed can be found in Tables I and II.

Table I. Energy Conservation in the Dynamics Simulations (NVE Simulations)

system	hybrid method[a]	time[b]	av temp (K)[c]	$\Delta\mathcal{H}$[d]	$\mathcal{H}_{Drift}$[e]
$H_9O_4^+$	BOMD CCSD	2.3 ps	157.7 K ± 29.4 K	0.002	0.000
	BOMD CCSD:B3LYP	3.43 ps	159.4 K ± 29.3 K	0.021	0.017
	ADMP-pHF CCSD:B3LYP[f]	2.61 ps	146.3 K ± 27.0 K	0.010	0.012
$H_{13}O_6^+$	BOMD CCSD:B3LYP	3.00 ps	151.1 K ± 24.5 K	0.018	0.073
	ADMP-pHF CCSD:B3LYP[f]	3.92 ps	139.6 K ± 18.8 K	0.027	0.009

[a]The 6-31+g(d,p) basis is used for all simulations. [b]The time step is 0.2 fs for all simulations. For production calculations, it is known from previous studies that BOMD can allow time steps of the order of 0.4 fs steps,[64] and for larger $\mu_{valence}$ values, ADMP can also allow similarly large steps;[64] but these are not considered in the current publication. We only benchmark the possible accuracy allowed by these simulations. [c]The average and RMS temperatures are computed assuming an equipartition theorem where the total kinetic energy is $\frac{3}{2}(N-1)kT$. These are simply a measure of the amount of kinetic energy in the system that in turn affects the extent to which the potential surface is sampled in these classical nuclear simulations. The initial kinetic energy was chosen such that the corresponding temperature was 300 K. Within this prescription the nuclear velocities were chosen to be random. [d]RMS deviation of total energy in kcal/mol. The total energy reflects the Hamiltonian in eq 3. [e]Drift in total Hamiltonian, eq 3, is obtained as the difference between the average total energies for the first and last 100 fs of dynamics data (in kcal/mol). [f]$\mu_{valence}$ = 180 au. See eq 6. 180 au = 0.1 amu·bohr^2.

Table II. Energy Conservation for ADMP-pHF Simulations at the MP2:B3LYP Level of Theory

system[a]	$\mu_{valence}$[b]	time[c]	av temp (K)[d]	$\Delta\mathcal{H}$[e]	$\mathcal{H}_{Drift}$[f]
$H_9O_4^+$	90	4.7 ps	146.4 K ± 26.5 K	0.043	0.062
	180	8.6 ps	153.98 K ± 25.61 K	0.030	0.060
	270	6.0 ps	152.43 K ± 23.13 K	0.030	0.080
	360	10.8 ps	133.7 K ± 24.4 K	0.079	0.319[g]

[a]The 6-31+g(d,p) basis is used for all simulations. [b]As noted in eq 6, $\mu_{valence}$ scales the electronic inertia tensor, with all numbers in atomic units. 180 au = 0.1 amu·bohr^2. [c]The time step is chosen to be 0.2 fs for all simulations, except for $\mu_{valence}$ = 90 au, which uses a 0.05 fs time step. [d]As noted in Table I, the average and RMS temperature in Kelvin are computed assuming an equipartition theorem. [e]RMS deviation in total energy in kcal/mol. The total energy represents the Hamiltonian in eq 3. [f]Drift in total Hamiltonian, eq 3, is computed as a difference between the average total energies for the first and last 100 fs of dynamics data (in kcal/mol). [g]The larger $\mu_{valence}$ = 360 au simulation has a pronounced drift for long time-scale dynamics. This is on account of larger deviations from the Born−Oppenheimer surface. See the Supporting Information, where it is noted that $\mu_{low,0} \propto [\mathbf{F}_{low,0}, \mathbf{P}_{low,0}]$. While one needs to be careful in using such a large $\mu_{valence}$ for production calculations, it appears from our discussion of results later in the publication that these simulations do describe the system to reasonable accuracy and in agreement with the remaining simulations. It must be further noted that these simulations are included here to facilitate the discussion in Section IIIC.

Table I contains simulations primarily used in Sections IIIA and B to establish the accuracy of frag-BOMD and ADMP-pHF
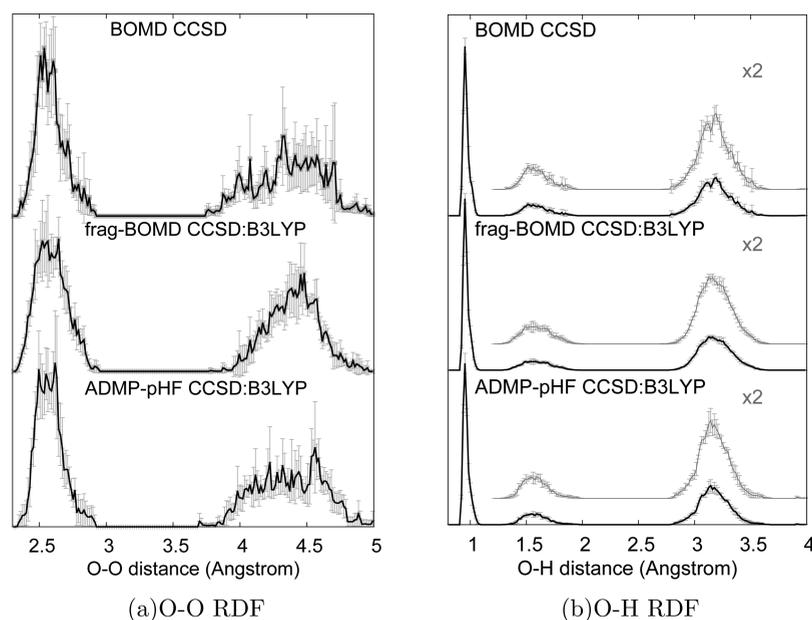
Figure 4. Radial distribution functions for oxygen−oxygen and oxygen−hydrogen distances for $H_9O_4^+$ trajectories. Shown in lighter colors in part (b) is a trace of the O−H radial distribution, magnified for ease of readability. The statistics are obtained by constructing averages from incremental 1 ps snapshots taken from the full trajectory. We compute the distributions using data from 0 to 1 ps, 0.25−1.25 ps, etc., for all trajectories. The distributions are then averaged, and standard deviations are presented as error bars.

simulations. The simulations in Table II are used in Section IIIC to study the effect of choice of $\mu_{valence}$ on accuracy.

For uniformity in comparison of results, most simulations are carried out using a time step of 0.2 fs. The only exception is the $\mu_{valence} = 90$ au, ADMP-pHF simulation which required a much smaller time step of 0.05 fs, and this choice is consistent with previous studies.[64] Production simulations using BOMD and those involving ADMP-pHF with larger $\mu_{valence}$ values can use larger time steps, but an analysis of time steps is not considered here. Our focus here is to determine the accuracy of these simulations in reproducing results in agreement with full CCSD BOMD calculations. All trajectories conserve the total energy to well under a kcal/mol, and we note that the drift in total energy is similarly small. The one exception is the relatively larger drift seen for the $\mu_{valence} = 360$ au ADMP-pHF simulation over an ≈ 10 ps trajectory length. This is on account of larger deviations from the Born−Oppenheimer surface. See the Supporting Information, where it is noted that $\mu_{low,0}$ and $[F_{low,0}, P_{low,0}]$ are directly related, and thus the choice of $\mu_{valence}$ presents a bound on deviations from the Born−Oppenheimer surface.[65] While one needs to be careful in using such a large $\mu_{valence}$ for production calculations as noted elsewhere,[121−123] it appears from our discussion of results later in the publication that these simulations do describe the system to reasonable accuracy and in agreement with the remaining simulations. Furthermore, this larger $\mu_{valence}$ value is now in the range of the inertia that is commonly used in Car−Parrinello molecular dynamics (CPMD) calculations,[121,122] where such drifts have been noted,[123] and, in fact, ref 123 was developed to overcome such drifts in CPMD. It must however be noted that these larger $\mu_{valence}$ simulations are only included in the current study to facilitate the discussion of $\mu_{valence}$-dependence of results in Section IIIC. These large values are not generally recommended for production simulations. In addition, as we have seen in our previous fragment-based AIMD publications the correlated AIMD trajectories (MP2 in refs 38 and 42) show 1

order of magnitude better $\Delta \mathcal{H}$ and $\mathcal{H}_{Drift}$ as compared to (a) DFT-based BOMD and ADMP trajectories and (b) frag-BOMD and ADMP-pHF trajectories. See, for example, Table 1 in ref 38 and Table 3 in ref 42, where it is seen that all DFT based BOMD and ADMP trajectories have $\Delta \mathcal{H}$ and $\mathcal{H}_{Drift}$ that are 1 order of magnitude larger as compared to BOMD MP2 trajectories. This is similarly the case for the fragmentation dynamics trajectories in refs 38 and 42 and also the case here, as noted in Table I for the frag-BOMD and ADMP-pHF results. Thus, $\Delta \mathcal{H}$ and $\mathcal{H}_{Drift}$ in frag-BOMD and ADMP-pHF are limited by the intrinsic error in the DFT forces due to the exchange-correlation quadrature grid.[124] While this can, in principle, be improved, by using ultrafine dense grids, we find that this is (a) neither necessary as we can already see from the accuracy of the spectra in refs 38 and 42 and those presented later here in this section, (b) nor is it desirable given the additional expense of ultrafine grids to be incurred at every time step.

The computational gain is discussed in detail in Section IV.

IIIA. Structural Features from Dynamics Simulations. We compute and compare the oxygen−oxygen and oxygen−hydrogen radial distribution functions (RDF) for each of the simulations listed above. The Eigen cation[125,126] comprises a single H_3O^+ cation solvated by three water molecules that are symmetrically arranged and act as hydrogen bond acceptors and for the case of the minimum energy geometry has the charge symmetrically delocalized[107] about the central oxygen atom. The solvated Zundel cation, on the other hand, features a centrally located proton shared between two equally spaced water molecules, the so-called Zundel cation,[127] that is then solvated by four other water molecules that again act as hydrogen bond acceptors. Figure 4(a) shows the Eigen O−O radial distribution function, whereas Figure 4(b) shows the Eigen O−H radial distribution function. The average distributions and the method of computing error bars are described in the figure caption. The broad peak extending
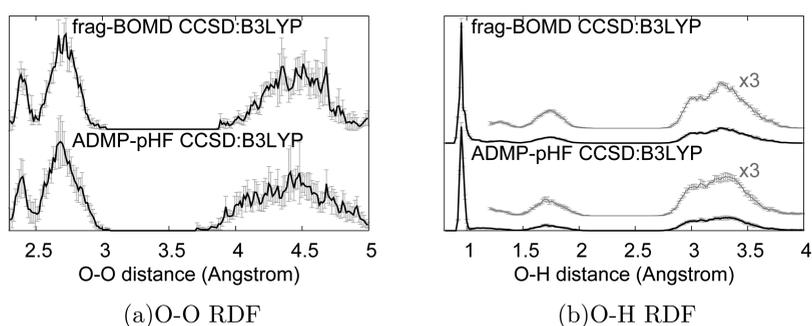
Figure 5. Radial distribution function for O−O and O−H distances from $H_{13}O_6^+$ trajectories. As in Figure 4, peaks at larger OH distances are magnified. Statistics and error bars are as in Figure 4.
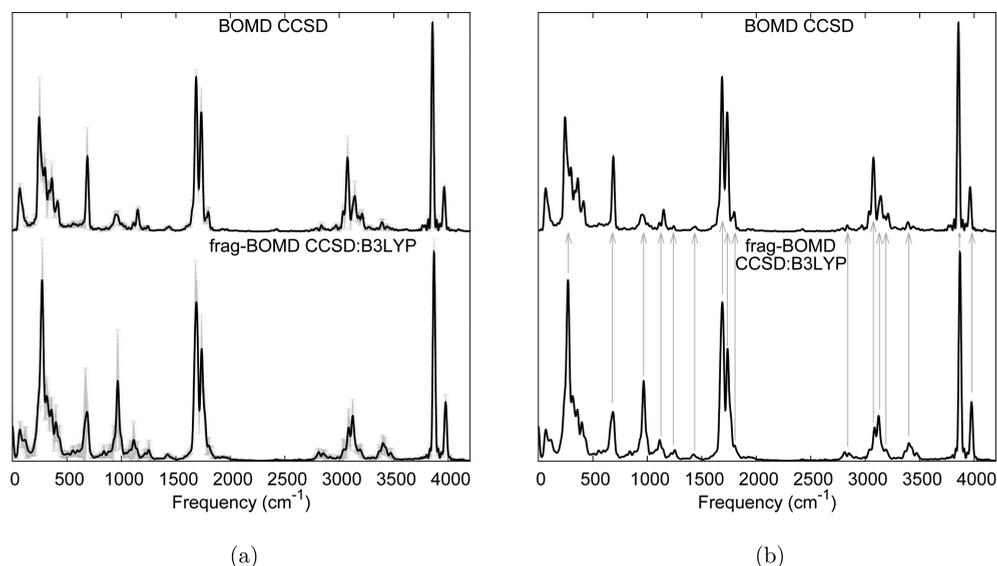
Figure 6. Vibrational density of states from full CCSD and frag-BOMD CCSD:B3LYP trajectories for $H_9O_4^+$ (at ≈150 K). The fragment-based calculation replicates the spectral activity predicted at the CCSD level of theory to very good accuracy. In part (a), the statistics is obtained by constructing averages from incremental 1 ps snapshots. We compute $I_V(\omega)$ using data from 0 to 1 ps, 0.25−1.25 ps, etc., for all trajectories. The results are averaged standard deviations presented by error bars. See the text for details. In part (b), the averaged spectra from (a) are reproduced, and the agreement in frequencies is emphasized through arrows to guide the eye.

between 2.4 and 2.9 Å describes the Eigen O−O inner solvation shell and is followed by a broader outer solvation shell spanning 3.7 Å−4.8 Å. The lower peak has two features, one wider compared to the other. The first feature, consistently represented in all three simulations, is at ≈2.5 Å and indicates the presence of an unsymmetric Zundel substructure sampled as part of the Eigen simulations. The second feature at ≈2.7 Å, which appears as a broad shoulder in all three cases, is the classical Eigen configuration. The distribution beyond 3.8 Å is the incomplete second solvation shell present in Eigen and is sampled only to a reasonable extent in all simulations; the frag-BOMD result appearing more structured in comparison, but this is likely to be due to the relatively short simulation time for this system. Figure 4(b) shows the Eigen cation O−H distribution. The first and most intense peak corresponds to the covalently bound hydrogens, and there is a shoulder on the right side of this intense peak which corresponds to anharmonic vibrations of the covalent OH bonds. The peak at ≈1.6 Å encompasses the region between 1.4 and 1.8 Å and corresponds to the internal hydrogen bonds between the central hydronium and the peripheral water molecules. At the shorter end of the distribution (≈1.5 Å), this peak along with the covalently bound distribution at ≈1 Å together depict an unsymmetric

Zundel sampled during the finite temperature Eigen simulations consistent with the corresponding OO distribution feature noted above. Critically, all of these characteristic features (including the second solvation peak at ≈3.2 Å in the OH distribution) are captured in our fragment-based trajectories and agree with those from the CCSD simulation.

For the solvated Zundel cation (Figure 5), the O−O distribution contains two types of nearest oxygen pairs: (1) a tight O−O pair caused by the strong Zundel-like hydrogen bond at <2.5 Å, and (2) the Eigen-like O−O peak centered around ≈2.7 Å. While full CCSD-based AIMD is intractable for a system as large as the solvated Zundel cation, both frag-BOMD and ADMP-pHF predict a broadening of the larger O−O peak, consistent with CCSD predictions made for the Eigen cation. In Figure 5(b), the O−H distribution exhibits a strong peak corresponding to the covalently bound OH, a feature expectedly shared with the Eigen cation simulation. Consistent with the O−H distribution in Figure 4(b), the solvated Zundel cation also exhibits two types of hydrogen bonds. The small bump centered at approximately 1.2 Å is seen only upon enhancement and is consistent with a hydrogen bond in a pure, unsolvated Zundel cation. At ≈1.6 Å, we note a second hydrogen bonded OH distribution, consistent with the one
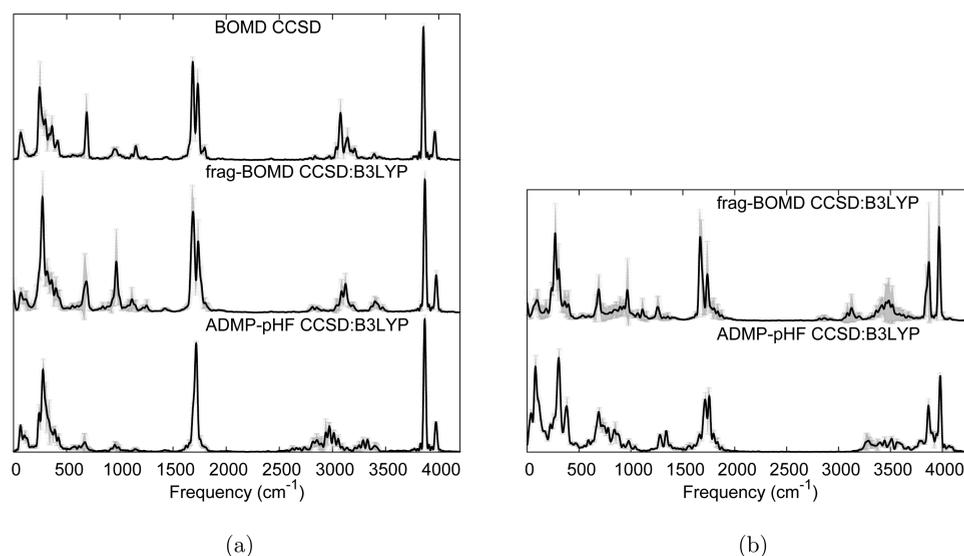
Figure 7. Comparison of vibrational density of states from CCSD BOMD, frag-BOMD CCSD:B3LYP, and ADMP-pHF CCSD:B3LYP. Part (a): Eigen cation, part (b): solvated Zundel cation. The ADMP-pHF frequencies are scaled by 1.021. Statistics and error bars are as in Figure 6.

exhibited for the Eigen cation. Thus, it appears that both solvated Zundel trajectories do exhibit Eigen-like substructures! Additional solvation shells are also seen in the 2.8 Å–3.6 Å range which specifically include two features, one peaked around 3 Å and a second at approximately 3.3 Å which correspond to second solvation shell peaks. All of these features are noted in both frag-BOMD and ADMP-pHF simulations.

IIIB. Vibrational Properties with CCSD Accuracy. In this section we benchmark the accuracy of our fragment-based dynamics by computing the vibrational density of states (VDOS) for all trajectory. This is computed using the Fourier transform of the velocity autocorrelation function[128]

$$
\begin{aligned}
I_V(\omega) &= \lim_{T \to \infty} \int_{t=0}^{t=T} dt\, \exp(-\imath \omega t) \langle \mathbf{V}(0) \cdot \mathbf{V}(t) \rangle \\
&= \tilde{\mathbf{V}}(\omega) \cdot \tilde{\mathbf{V}}(\omega) = |\tilde{\mathbf{V}}(\omega)|^2
\end{aligned}
\tag{7}
$$

where $\mathbf{V}(t)$ represents the 3N velocity vector at time t with Fourier transform $\tilde{\mathbf{V}}(\omega)$. We have used the convolution theorem[129] in presenting the second and third equalities in the equation above. Our previous studies on hydrogen bonded systems[51–54,70,71,119,130] have shown that vibrational properties obtained from AIMD are particularly sensitive to the internal cluster temperature, which dictates the extent to which the potential energy surface is sampled. It has been shown in previous publications[54,71,119] that results from simulations generally in the temperature range of ≈150 K are found to be in good agreement with argon-tagged action spectroscopy and those >300 K are in agreement with infrared multiphoton dissociation (IRMPD) for hydrogen bonded systems. Hence the choice of cluster temperature is significant, and initial conditions were chosen for all trajectories such that the average kinetic energy was in the 150 K range in all cases.

For each trajectory described in Table I we have computed $I_V(\omega)$. Shown in Figure 6 is a comparison between the spectral activity predicted by full CCSD and our frag-BOMD method for $H_9O_4^+$. In computing these spectra, we have used the following steps: (a) The convolution theorem[129] is used to simplify the Fourier transform of the autocorrelation function, as already noted in eq 7. This reduces the Fourier transform of the velocity autocorrelation function to simply the power

spectrum of velocities. This is also known as the Wiener-Khinchin theorem.[129] (b) The idea of introducing a 1 ps moving window that is explained in the caption of Figure 6 is related to two aspects in the theory of computing accurate Fourier transforms that (i) reduce "leakage"[129] and (ii) reduce variance in the Fourier transform. Specifically, whenever a finite sample of frequencies is constructed, there is always a leakage (or smudging phenomenon) where information from one frequency leaks to the neighboring frequency regions of the spectrum. The extent of the leakage domain is generally quite substantial and falls of as $1/\omega^2$. Data windowing reduces the leakage of information in the frequency domain, and the recommended form of data windowing is one that includes overlapping segments, as explained in the caption of Figure 6. The second reason for computing the Fourier transforms of these segments and subsequent averaging of these is to reduce variance in the Fourier transform as explained in ref 129.

We find excellent agreement between the spectra in Figure 6. The high frequency proton stretching motions are replicated in frag-BOMD in excellent agreement, and the lower frequency peaks are also in very good agreement. Error bars are explained in the figure caption and shown in Figure 6(a), whereas in Figure 6(b) the gray arrows are placed as a guide to demonstrate the level of agreement.

A well-known issue in ADMP and other extended Lagrangian methods is the dependence of vibrational frequencies on the fictitious electronic inertia parameters.[64,121,122,131–134] The fictitious electronic oscillations are much faster than nuclear oscillations but contribute to deviations from the BOMD forces, as indicated in the Supporting Information. These additional forces act as fluctuations to the nuclear gradients and may add coherently to produce a net phase shift in ADMP forces from BOMD forces,[131,132] which manifests as a red-shift in vibrational frequency. Hence we apply a scaling factor to the frequencies in all ADMP-pHF spectra shown here. Figure 7 compares spectra obtained from frag-BOMD and ADMP-pHF after the ADMP-pHF spectral frequencies are scaled by 1.021. This implies that a 2.1% blue-shift in the ADMP-pHF frequencies was necessary for the level of agreement with frag-BOMD and full CCSD BOMD results shown in Figure 7.

(a)O-O RDF   (b)O-H RDF
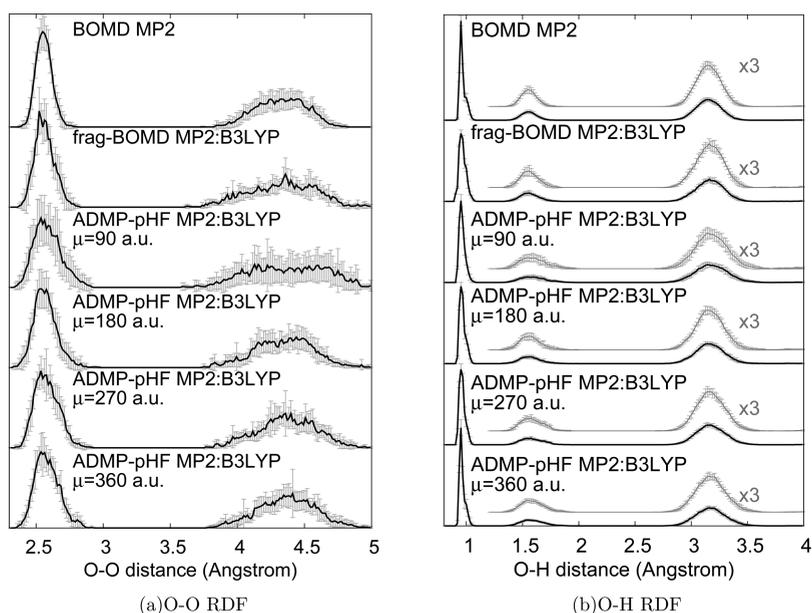
Figure 8. Radial distribution function for O−O and O−H distances at multiple values of μ_{valence}. The structural distributions are in reasonable agreement for all μ_{valence} values.

(a)Unscaled   (b)Scaled
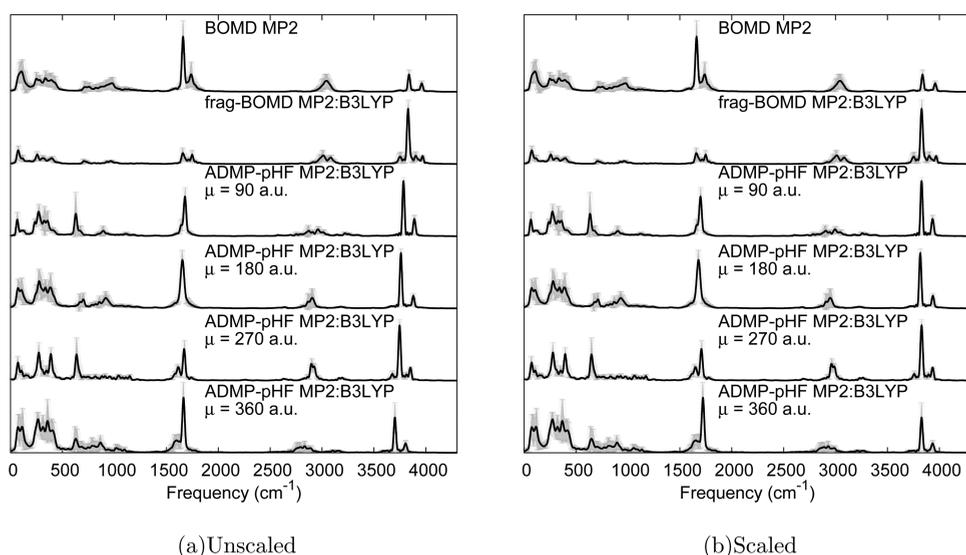
Figure 9. Vibrational density of states calculated from dynamical trajectories. ADMP-pHF simulations using multiple values of μ_{valence} are shown to illustrate the μ_{valence}-dependence of the red-shift in the higher frequency modes. (a) As the ficticious mass parameter is reduced, the red-shift in higher frequency modes reduces, and modes tend to converge to those predicted by frag-BOMD and full MP2. (b) Each simulation is independently scaled by a μ_{valence}-dependent uniform scaling factor as shown in Figure 11. Statistics and error bars are as in Figure 6.

Qualitatively, all of the significant spectral features are captured in ADMP-pHF as compared to frag-BOMD and BOMD with full CCSD, except for the peak at ≈3000 cm^{-1} for the Eigen case (Figure 7(a)). This region is the highly polarizable Eigen-cation proton stretch and likely leads to enormous coupling with electronic structure. To probe the level of ADMP-pHF accuracy in this region of the spectrum, in the next subsection we present the spectroscopic dependence on choice of μ_{valence}; but, before that, we also note that low frequency regions of the spectra in Figure 7 are in good agreement even before the scaling factor is applied, and the uniform frequency scaling factor of 1.021 applied to ADMP-pHF appears to be sufficient to correct the ADMP red-shift for the most part and provide accurate spectra in the higher frequencies. There still remains a

slight, but notable red-shift in the 2500−3500 cm^{-1} frequency window. Apart from that, both frag-BOMD and ADMP-pHF are in agreement with the full CCSD simulation.

IIIC. Effect of μ_{valence} on the Structural and Vibrational Properties from ADMP-pHF and Comparison of Vibrational Properties with DFT. To probe the effect of the choice of inertia tensor we compute $I_V(\omega)$ and radial distribution functions using the MP2:B3LYP level of theory for a series of simulations varying in μ_{valence}. The details for these trajectories are summarized in Table II, and our computed results are shown in Figures 8 and 9. The structural properties are reproduced with good accuracy for all values of μ_{valence} as seen in Figure 8. Higher frequency peaks in the vibrational density of states are red-shifted as expected, but reducing the ficticious
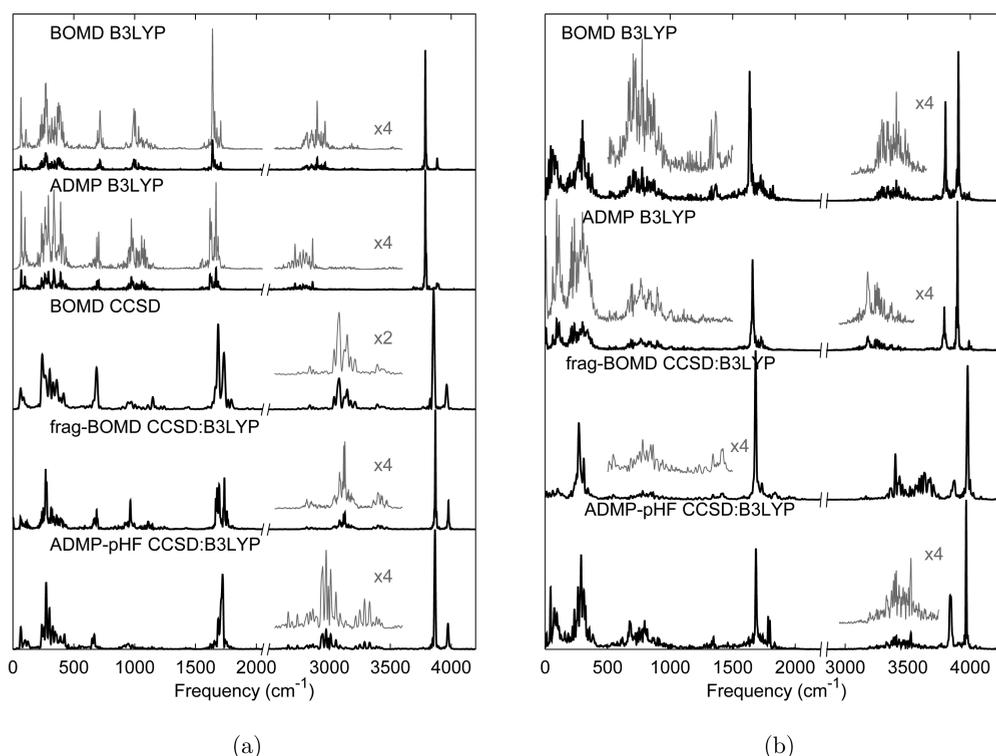
Figure 10. Vibrational density of states calculated from dynamical trajectories. The high-frequency, free−OH stretches are noticeably red-shifted in DFT simulations as compared to CCSD and fragment-based CCSD. (a) The Eigen cation. (b) The solvated Zundel cation. The ADMP and ADMP-pHF frequencies are scaled by the same 1.021 scaling factor used for all other $\mu_{valence}$ = 180 au simulations.

inertia reduces the red-shift. Furthermore, the same uniform scaling of 1.021 can be employed for both MP2 and CCSD when the fictitious inertia tensor scaling parameter is chosen to be 180 amu bohr². Thus, the scaling factors appear to be independent of theory based on this limited analysis and, as noted in Figure 7, system independent as well at least for the hydrogen bonded water clusters studied here. Future studies will further probe this dependence for larger water clusters and then for other hydrogen bonded systems, but below we also find our choice of scaling factors to be consistent for DFT.

DFT based BOMD and ADMP studies for Eigen and solvated Zundel were conducted in ref 38. While structural features were in agreement, the DFT vibrational peaks were red-shifted with respect to MP2. The corresponding comparison between DFT and CCSD is presented in Figure 10. Both ADMP and ADMP-pHF frequencies have been scaled by 1.021, consistent with our observations above. As is clear the free−OH stretch peaks on the far right are red-shifted in DFT, and the effect on the hydrogen bonded Eigen-like OH stretch peaks is even greater when one compares DFT and CCSD.

As the value of $\mu_{valence}$ is decreased, the scaling factor also gets smaller, and in Figure 11 we present the behavior of this scaling factor as a function of $\mu_{valence}$ and compare it to a quadratic fit that is expected from the discussion in ref 65. Specifically, in ref 65 it has been shown that the commutator $[\mathbf{P}_{low,0},\mathbf{F}_{low,0}] \propto \mu_{valence}$. The difference between ADMP and Born−Oppenheimer forces is also proportional to the commutator $[\mathbf{P}_{low,0},\mathbf{F}_{low,0}]$ (see the Supporting Information), and hence it follows that the difference in forces must be proportional to $\mu_{valence}$. Additionally, due to the quadratic dependence of the correlation function, $I_V(\omega)$ in eq 7, we expect a shift in the velocity spectrum that is quadratic in $\mu_{valence}$, and this aspect is consistent with Figure 11.
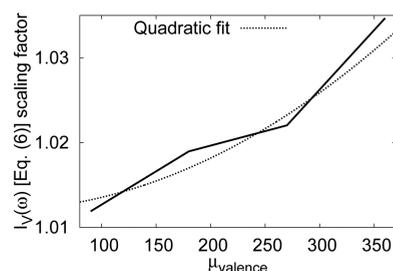
Figure 11. Scaling factors used in Figure 9 follow the expected quadratic dependence in $\mu_{valence}$.

It is clear that while structural features are well reproduced by ADMP-pHF, vibrational properties can be in acceptable agreement provided a scaling factor is applied.

IV. NUMERICAL EFFICIENCY IN COMPUTING CCSD BASED AIMD TRAJECTORIES FOR $(H_2O)_{21}H^+$

Our computational implementation for frag-BOMD and ADMP-pHF is achieved through a C++ program that is written in parallel and works using MPI across nodes and with OpenMP inside each node. For small water clusters such as the ones treated here, the computational bottleneck involves evaluating energies and forces for water dimers at the high level of theory (CCSD in this case), and we see little performance difference between ADMP-pHF and frag-BOMD. For larger systems, however, the low level calculation over the full system becomes a bottleneck. In Figure 12 we present the computational scaling of our implementation of fragment-based AIMD for the challenging protonated 21-water cluster. The protonated 21-mer system is interesting because it presents a long-standing challenge to theory.[135] The system, noted to be a

(a) Theoretical MPI Scaling

(b) Actual MPI Scaling
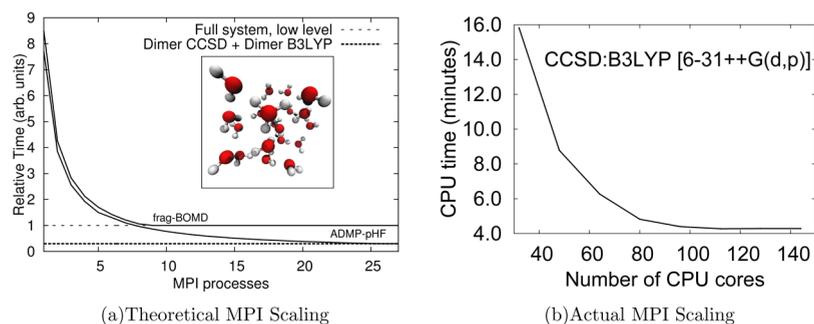
Figure 12. MPI scaling for the 21-water cluster: $(H_2O)_{21}H^+$. (a) frag-BOMD shows linear parallel speedup until bottle-necked by the full system calculated at the lower level of theory. ADMP-pHF calculations are bottle-necked by the cost of the largest primitive fragments. (b) The MPI scaling shown for a single frag-BOMD step.

"magic" cluster with pronounced stability,[43,136] has been studied experimentally using argon-tagged action spectroscopy[43−45,135] and theoretically by us[51,52] and by others.[43,116,117,135] While AIMD simulations using DFT and lower levels of theory have been able to qualitatively capture the essence of experimental findings,[43,51,52] the problem remains a great challenge since it is beyond the realm of dynamics based on electronic structure treatments at the post-Hartree−Fock level. While we reserve a full discussion of the vibrational spectrum obtained from frag-BOMD and ADMP-pHF to future publications, here we use this system to gauge the computational power of our methodology and its numerical implementation.

A typical structure found during a protonated $(H_2O)_{21}H^+$ dynamical simulation may require upward of 25 primitive fragments (water dimer units) and nearly 18 derivative fragments (water monomers formed from overlapping primitive fragments) to accurately describe the system. Computations involving these fragments are evenly divided over the ensemble of MPI processes and evaluated in parallel. A theoretical scaling curve is provided in Figure 12(a), and the real scaling for CCSD:B3LYP is provided in Figure 12(b). Here, frag-BOMD is expected to provide near-linear parallel speedup until the full system calculation at the lower level of theory becomes the primary bottleneck at roughly 10 MPI processes. ADMP-pHF alleviates the bottleneck and allows efficiently parallelization until the point at which the primitive dimer fragments become the bottleneck. As seen in Figure 12(b) our approach is capable of studying these complex systems at the CCSD level.

V. CONCLUSION

In a recent set of publications[38,42] we developed a method for dynamical treatment of molecular systems using fragment based electronic structure. The fragmentation method is based on ONIOM but includes overlapping model systems that are suitably treated using the set-theoretic inclusion-exclusion principle. Both Born−Oppenheimer and extended Lagrangian treatments are possible as shown in ref 38, and the latter is particularly interesting since the approach now allows *Car−Parrinello-style dynamics but with post-Hartree−Fock accuracy.* The particular extended Lagrangian employed here is a post-Hartree−Fock extension of the previously develop atom-centered density matrix propagation and is hence called ADMP-pHF. It is in this setting that the current publication makes a contribution as we discuss here a coupled-cluster implementation of "on-the-fly" dynamics both within the Car−Parrinello (or ADMP-pHF) and Born−Oppenheimer (frag-BOMD) paradigms.

The approach is rigorously tested for two protonated water clusters: $H_9O_4^+$ and $H_{13}O_6^+$, where in the case of the former full CCSD dynamics results are also presented. It is noted that the vibrational density of states and temperature dependent structure distribution functions computed from Born−Oppenheimer implementation of fragment based CCSD dynamics are in near perfect agreement with the full CCSD dynamics simulations. As a further demonstration of the computational power of the method, we have also benchmarked the computation times required to obtain dynamics in agreement with CCSD for the larger sized, protonated 21-water cluster, $(H_2O)_{21}H^+$, and this turns out to be an attractive option for large scale simulations in complex systems. The approach is promising because both frag-BOMD and ADMP-pHF may be able to provide CCSD accuracy for "on-the-fly" dynamics at DFT expense.

The corresponding ADMP-pHF simulations are faster than the Born−Oppenheimer, frag-BOMD, counterpart as seen in the protonated 21-water cluster, $(H_2O)_{21}H^+$ test case, and provide excellent agreement for the temperature governed structural distribution for $H_9O_4^+$ and $H_{13}O_6^+$. The vibrational properties, on the contrary, show an inertia tensor dependent red-shift, that is overcome here through introduction of a uniform scaling factor, which then improves the agreement between CCSD-based ADMP-pHF and the full Born−Oppenheimer based CCSD simulations. We have also studied the behavior of properties as a function of choice of inertia tensor and find that while the structural properties remain in good agreement all through, the vibrational properties show a monotonic improvement with decreasing inertia, converging toward the accurate full CCSD simulation results. Thus, we expect both frag-BOMD and ADMP-pHF to have significant impact in computing accurate *ab initio* dynamics trajectories for complex systems.

◼ APPENDIX

Fragmentation Algorithm for Protonated Water Clusters

We obtain fragments for a given molecular structure as follows:

1) Given the minimum oxygen−oxygen distance for a chosen oxygen atom, denoted as OO_i, a maximum of three additional water molecules are included inside the solvation sphere of the water molecule such that the respective oxygen−oxygen separations are within 10% of OO_i.

2) Bonded hydrogens are then included to form primitive fragments. Here, we use two well-known limits for the

protonated water cluster system: (a) the Zundel $H_5O_2^+$ system where the oxygen−oxygen distance is 2.5 Å, on average, and the shared proton is situated in the middle. (b) Second the Eigen, $H_9O_4^+$, system, which is most H_3O^+ like but fully solvated. To include the definition of delocalized charge as afforded by these two systems, the cutoff for the O−H distance is 1.4 Å in all our simulations.

3) All fragments that include the excess proton inside their envelope have a net charge of +1. The rest are neutral.

4) Finally, all fragments that do not contain electrons are excluded from the calculation. See for example the intervening proton in Figure 3(b). This fragment does not contain any electrons. Since the energies computed using electronic structure are interaction energies, a fragment containing just one proton has zero interaction energy. (This definition would need to be modified if molecular mechanics is used at the lower level, which is not the case in this publication.)

Primitive fragments generated by this algorithm overlap with adjacent fragments through an intervening hydronium ion or a water molecule. These intersections ensure all hydrogen bonds are equally treated. Two examples are given in Figure 3.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.6b01107.

> Equations of motion for nuclei and density matrices, derived from eq 4, nuclear and density matrix gradients for completeness, and further details on these in the cited references (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: iyengar@indiana.edu.

ORCID ⓘ

Srinivasan S. Iyengar: 0000-0001-6526-2907

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Field, C.; Bash, P. A.; Karplus, M. A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations. *J. Comput. Chem.* **1990**, *11*, 700.

(2) Car, R.; Parrinello, M. Unified Approach for Molecular Dynamics and Density-Functional Theory. *Phys. Rev. Lett.* **1985**, *55*, 2471.

(3) Koitz, R.; Hutter, J.; Iannuzzi, M. Formation and properties of a terpyridine-based 2D MOF on the surface of water. *2D Mater.* **2016**, *3*, 025026.

(4) Gerber, R. B.; Varner, M. E.; Hammerich, A. D.; Riikonen, S.; Murdachaew, G.; Shemesh, D.; Finlayson-Pitts, B. J. Computational Studies Of Atmospherically-Relevant Chemical Reactions In Water Clusters And On Liquid Water And Ice Surfaces. *Acc. Chem. Res.* **2015**, *48*, 399−406.

(5) Gerber, R. B.; Sebek, J. Dynamics Simulations of Atmospherically Relevant Molecular Reactions. *Int. Rev. Phys. Chem.* **2009**, *28*, 207−222.

(6) Hutter, J. Car-Parrinello Molecular Dynamics. *WIREs-Comp. Mol. Sci.* **2012**, *2*, 604−612.

(7) Kirchner, B.; di Dio, P. J.; Hutter, J. Real-World Predictions From Ab Initio Molecular Dynamics Simulations. In *Multiscale Molecular Methods in Applied Chemistry*; Kirchner, B., Vrabec, J., Eds.; Topics in Current Chemistry-Series; 2012; Vol. 307, DOI: 10.1007/128_2011_195.

(8) Cohen, A. J.; Mori-Sanchez, P.; Yang, W. T. Challenges for Density Functional Theory. *Chem. Rev.* **2012**, *112*, 289.

(9) Klimes, J.; Michaelides, A. Perspective: Advances and challenges in treating van der Waals dispersion forces in density functional theory. *J. Chem. Phys.* **2012**, *137*, 120901.

(10) Peverati, R.; Truhlar, D. Quest for a Universal Density Functional: The Accuracy of Density Functionals Across a Broad Spectrum of Databases in Chemistry and Physics. *Philos. Trans. R. Soc., A* **2014**, *372*, 20120476.

(11) Yang, W. Direct Calculation of Electron Density in Density-Functional Theory. *Phys. Rev. Lett.* **1991**, *66*, 1438.

(12) Maseras, F.; Morokuma, K. A New "Ab Initio + Molecular Mechanics" Geometry Optimization Scheme of Equilibrium Structures and Transition States. *J. Comput. Chem.* **1995**, *16*, 1170.

(13) Kerdcharoen, T.; Morokuma, K. ONIOM-XS: An Extension of the ONIOM Method for Molecular Simulation in Condensed Phase. *Chem. Phys. Lett.* **2002**, *355*, 257.

(14) Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. The Effective Fragment Potential Method: A QM-Based MM Approach to Modeling Environmental Effects in Chemistry. *J. Phys. Chem. A* **2001**, *105*, 293.

(15) Gordon, M.; Mullin, J.; Pruitt, S.; Roskop, L.; Slipchenko, L.; Boatz, J. Accurate Methods for Large Molecular Systems. *J. Phys. Chem. B* **2009**, *113*, 9646.

(16) Zhang, D. W.; Zhang, J. Z. H. Molecular Fractionation with Conjugate Caps for Full Quantum Mechanical Calculation of Protein-molecule Interaction Energy. *J. Chem. Phys.* **2003**, *119*, 3599.

(17) Dahlke, E. E.; Truhlar, D. G. Electrostatically Embedded Many Body Expansion for Large Systems, with Applications to Water Clusters. *J. Chem. Theory Comput.* **2007**, *3*, 46.

(18) Dahlke, E. E.; Truhlar, D. G. Electrostatically Embedded Many Body Expansion for Simulations. *J. Chem. Theory Comput.* **2008**, *4*, 1.

(19) Li, S.; Li, W.; Ma, J. Generalized Energy-Based Fragmentation Approach and Its Applications to Macromolecules and Molecular Aggregates. *Acc. Chem. Res.* **2014**, *47*, 2712.

(20) Wang, L.-W.; Zhao, Z.; Meza, J. Linear-Scaling Three-Dimensional Fragment method for Large-scale Electronic Structure Calculations. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2008**, *77*, 165113.

(21) Ganesh, V.; Dongare, R. K.; Balanarayan, P.; Gadre, S. R. Molecular Tailoring Approach for Geometry Optimization of Large Molecules: Energy Evaluation and Parallelization Strategies. *J. Chem. Phys.* **2006**, *125*, 104109.

(22) Guo, W.; Wu, A.; Xu, X. XO: An Extended ONIOM Method for Accurate and Efficient Geometry Optimization of Large Molecules. *Chem. Phys. Lett.* **2010**, *498*, 203−208.

(23) Mayhall, N. J.; Raghavachari, K. Molecules-In-Molecules: An Extrapolated Fragment-Based Approach for Accurate Calculations on Large Molecules and Materials. *J. Chem. Theory Comput.* **2011**, *7*, 1336.

(24) Mayhall, N. J.; Raghavachari, K. Many-Overlapping-Body (MOB) Expansion: A Generalized Many Body Expansion for Nondisjoint Monomers in Molecular Fragmentation Calculations of Covalent Molecules. *J. Chem. Theory Comput.* **2012**, *8*, 2669.

(25) Jacobson, L. D.; Herbert, J. M. An Efficient, Fragment-Based Electronic Structure Method for Molecular Systems: Self-Consistent Polarization with Perturbative Two-Body Exchange and Dispersion. *J. Chem. Phys.* **2011**, *134*, 094118.

(26) Richard, R. M.; Herbert, J. M. A Generalized Many-Body Expansion and a Unified View of Fragment-Based Methods in Electronic Structure Theory. *J. Chem. Phys.* **2012**, *137*, 064113.

(27) Hirata, S. Fast Electron-Correlation Methods for Molecular Crystals: an Application to the α, $\beta(1)$, and $\beta(2)$ Modifications of Solid Formic Acid. *J. Chem. Phys.* **2008**, *129*, 204104.

(28) Kamiya, M.; Hirata, S.; Valiev, M. Fast Electron-Correlation Methods for Molecular Crystals Without Basis Set Superposition Errors. *J. Chem. Phys.* **2008**, *128*, 074103.

(29) Brorsen, K. R.; Minezawa, N.; Xu, F.; Windus, T. L.; Gordon, M. S. Fragment Molecular Orbital Molecular Dynamics with the Fully Analytic Energy Gradient. *J. Chem. Theory Comput.* **2012**, *8*, 5008.

(30) Brorsen, K. R.; Zahariev, F.; Nakata, H.; Fedorov, D. G.; Gordon, M. S. Analytic Gradient for Density Functional Theory Based on the Fragment Molecular Orbital Method. *J. Chem. Theory Comput.* **2014**, *10*, 5297.

(31) Le, H.-A.; Tan, H.-J.; Ouyang, J. F.; Bettens, R. P. A. Combined Fragmentation Method: A Simple Method for Fragmentation of Large Molecules. *J. Chem. Theory Comput.* **2012**, *8*, 469.

(32) Han, J.; Mazack, M. J. M.; Zhang, P.; Truhlar, D. G.; Gao, J. Quantum Mechanical Force Field for Water with Explicit Electronic Polarization. *J. Chem. Phys.* **2013**, *139*, 054503.

(33) Chung, L. W.; Hirao, H.; Li, X.; Morokuma, K. The ONIOM Method: Its Foundation and Applications to Metalloenzymes and Photobiology. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012**, *2*, 327.

(34) Raghavachari, K.; Saha, A. Accurate Composite and Fragment-Based Quantum Chemical Models for Large Molecules. *Chem. Rev.* **2015**, *115*, 5643.

(35) Chung, L. W.; Sameera, W. M. C.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z.; Liu, F.; Li, H.-B.; Ding, L.; Morokuma, K. The ONIOM Method and Its Applications. *Chem. Rev.* **2015**, *115*, 5678.

(36) Collins, M. A.; Bettens, R. P. A. Energy-Based Molecular Fragmentation Methods. *Chem. Rev.* **2015**, *115*, 5607.

(37) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. Fragmentation Methods: A Route to Accurate Calculations on Large Systems. *Chem. Rev.* **2012**, *112*, 632.

(38) Li, J.; Haycraft, C.; Iyengar, S. S. Hybrid extended Lagrangian, post-Hartree-Fock Born-Oppenheimer ab initio molecular dynamics using fragment-based electronic structure. *J. Chem. Theory Comput.* **2016**, *12*, 2493.

(39) Liu, J.; Zhu, T.; Wang, X.; He, X.; Zhang, J. Z. H. Quantum Fragment Based ab Initio Molecular Dynamics for Proteins. *J. Chem. Theory Comput.* **2015**, *11*, 5897.

(40) Willow, S. Y.; Salim, M. A.; Kim, K. S.; Hirata, S. Ab initio molecular dynamics of liquid water using embedded fragment second-order many-body perturbation theory towards its accurate property prediction. *Sci. Rep.* **2015**, *5*, 14358.

(41) Lange, A. W.; Voth, G. A. Multi-State Approach to Chemical Reactivity in Fragment Based Quantum Chemistry Calculations. *J. Chem. Theory Comput.* **2013**, *9*, 4018.

(42) Li, J.; Iyengar, S. S. Ab initio Molecular Dynamics using Recursive, Spatially Separated, Overlapping Model Subsystems Mixed Within an ONIOM Based Fragmentation Energy Extrapolation Technique. *J. Chem. Theory Comput.* **2015**, *11*, 3978−3991.

(43) Shin, J.-W.; Hammer, N. I.; Diken, E. G.; Johnson, M. A.; Walters, R. S.; Jaeger, T. D.; Duncan, M. A.; Christie, R. A.; Jordan, K. D. Infrared Signature of Structures Associated with the $H^+(H_2O)_n$ (N = 6 to 27) Clusters. *Science* **2004**, *304*, 1137.

(44) Miyazaki, M.; Fujii, A.; Ebata, T.; Mikami, N. Infrared Spectroscopic Evidence for Protonated Water Clusters Forming Nanoscale Cages. *Science* **2004**, *304*, 1134.

(45) Douberly, G. E.; Ricks, A. M.; Duncan, M. A. Infrared Spectroscopy of Perdeuterated Protonated Water Clusters in the Vicinity of the Clathrate Cage. *J. Phys. Chem. A* **2009**, *113*, 8449.

(46) Headrick, J. M.; Diken, E. G.; Walters, R. S.; Hammer, N. I.; Christie, R. A.; Cui, J.; Myshakin, E. M.; Duncan, M. A.; Johnson, M. A.; Jordan, K. Spectral Signatures of Hydrated Proton Vibrations in Water Clusters. *Science* **2005**, *308*, 1765.

(47) Hammer, N. I.; Diken, E. G.; Roscioli, J. R.; Johnson, M. A.; Myshakin, E. M.; Jordan, K. D.; McCoy, A. B.; Huang, X.; Bowman, J. M.; Carter, S. The Vibrational Predissociation Spectra of the $H_5O_2^+ \cdot RG_n(RG = Ar,Ne)$ clusters: Correlation of the solvent perturbations in the free OH and shared proton transitions of the Zundel ion. *J. Chem. Phys.* **2005**, *122*, 244301.

(48) Diken, E. G.; Headrick, J. M.; Roscioli, J. R.; Bopp, J. C.; Johnson, M. A.; McCoy, A. B. Fundamental Excitations of the Shared Proton in the $H_3O_2^-$ and $H_5O_2^+$ Complexes. *J. Phys. Chem. A* **2005**, *109*, 1487.

(49) Asmis, K. R.; Pivonka, N. L.; Santambrogio, G.; Brümmer, M.; Kaposta, C.; Neumark, D. M.; Wöste, L. Gas-Phase Infrared Spectrum of the Protonated Water Dimer. *Science* **2003**, *299*, 1375.

(50) Fridgen, T. D.; MacAleese, P.; Maitre, L.; McMahon, T. B.; Boissel, P.; Lemaire, J. Infrared Spectra of Homogeneous and Heterogeneous Proton-Bound Dimers in the Gas Phase. *Phys. Chem. Chem. Phys.* **2005**, *7*, 2747.

(51) Iyengar, S. S.; Petersen, M. K.; Day, T. J. F.; Burnham, C. J.; Teige, V. E.; Voth, G. A. The Properties of Ion-Water Clusters. I. the Protonated 21-Water Cluster. *J. Chem. Phys.* **2005**, *123*, 084309.

(52) Iyengar, S. S. Further Analysis of the Dynamically Averaged Vibrational Spectrum for the "Magic" Protonated 21-Water Cluster. *J. Chem. Phys.* **2007**, *126*, 216101.

(53) Li, X.; Moore, D. T.; Iyengar, S. S. Insights from First Principles Molecular Dynamics Studies Towards Infra-Red Multiple-Photon and Single-Photon Action Spectroscopy: Case Study of the Proton-Bound Di-Methyl Ether Dimer. *J. Chem. Phys.* **2008**, *128*, 184308.

(54) Li, X.; Oomens, J.; Eyler, J. R.; Moore, D. T.; Iyengar, S. S. Isotope Dependent, Temperature Regulated, Energy Repartitioning in a Low-Barrier, Short-Strong Hydrogen Bonded Cluster. *J. Chem. Phys.* **2010**, *132*, 244301.

(55) Vendrell, O.; Gatti, F.; Meyer, H.-D. Dynamics and Infrared Spectroscopy of the Protonated Water Dimer. *Angew. Chem., Int. Ed.* **2007**, *46*, 6918.

(56) Vendrell, O.; Gatti, F.; Meyer, H.-D. Full Dimensional (15D) Quantum-Dynamical Simulation of the Protonated Water-Dimer II: Infrared Spectrum and Vibrational Dynamics Dynamics and Infrared Spectroscopy of the Protonated Water Dimer. *J. Chem. Phys.* **2007**, *127*, 184303.

(57) Vendrell, O.; Gatti, F.; Meyer, H.-D. Strong Isotope Effects in the Infrared Spectrum of the Zundel Cation. *Angew. Chem., Int. Ed.* **2009**, *48*, 352.

(58) Björklund, A.; Husfeldt, T.; Koivisto, M. Set Partitioning via Inclusion Exclusion. *SIAM J. Comput.* **2009**, *39*, 546.

(59) Komeiji, Y.; Mochizuki, Y.; Nakano, T. Three-Body Expansion and Generalized Dynamic Fragmentation Improve the Fragment Molecular Orbital-Based Molecular Dynamics (FMO-MD). *Chem. Phys. Lett.* **2010**, *484*, 380.

(60) Collins, M. A. Can Systematic Molecular Fragmentation Be Applied to Direct Ab Initio Molecular Dynamics? *J. Phys. Chem. A* **2016**, *120*, 9281−9291.

(61) Goldstein, H.; Poole, C.; Safko, J. *Classical Mechanics*; Addison Wesley: San Francisco, CA, 2002.

(62) Schlegel, H. B.; Millam, J. M.; Iyengar, S. S.; Voth, G. A.; Daniels, A. D.; Scuseria, G. E.; Frisch, M. J. Ab Initio Molecular Dynamics: Propagating the Density Matrix with Gaussian Orbitals. *J. Chem. Phys.* **2001**, *114*, 9758.

(63) Iyengar, S. S.; Schlegel, H. B.; Millam, J. M.; Voth, G. A.; Scuseria, G. E.; Frisch, M. J. Ab Initio Molecular Dynamics: Propagating the Density Matrix with Gaussian Orbitals. II. Generalizations Based on Mass-Weighting, Idempotency, Energy Conservation and Choice of Initial Conditions. *J. Chem. Phys.* **2001**, *115*, 10291.

(64) Schlegel, H. B.; Iyengar, S. S.; Li, X.; Millam, J. M.; Voth, G. A.; Scuseria, G. E.; Frisch, M. J. Ab Initio Molecular Dynamics: Propagating the Density Matrix with Gaussian Orbitals. III. Comparison with Born-Oppenheimer Dynamics. *J. Chem. Phys.* **2002**, *117*, 8694.

(65) Iyengar, S. S.; Schlegel, H. B.; Voth, G. A.; Millam, J. M.; Scuseria, G. E.; Frisch, M. J. Ab Initio Molecular Dynamics: Propagating the Density Matrix with Gaussian Orbitals. IV. Formal Analysis of the Deviations from Born-Oppenheimer Dynamics. *Isr. J. Chem.* **2002**, *42*, 191.

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

(67) Iyengar, S. S.; Frisch, M. J. Effect of Time-Dependent Basis Functions and Their Superposition Error on Atom-Centered Density Matrix Propagation (ADMP): Connections to Wavelet Theory of Multi-Resolution Analysis. *J. Chem. Phys.* **2004**, *121*, 5061.

(68) Rega, N.; Iyengar, S. S.; Voth, G. A.; Schlegel, H. B.; Vreven, T.; Frisch, M. J. Hybrid Ab-Initio/Empirical Molecular Dynamics: Combining the ONIOM Scheme with the Atom-Centered Density Matrix Propagation (ADMP) Approach. *J. Phys. Chem. B* **2004**, *108*, 4210.

(69) Iyengar, S. S.; Day, T. J. F.; Voth, G. A. On the Amphiphilic Behavior of the Hydrated Proton: An Ab Initio Molecular Dynamics Study. *Int. J. Mass Spectrom.* **2005**, *241*, 197.

(70) Iyengar, S. S. Dynamical Effects on Vibrational and Electronic Spectra of Hydroperoxyl Radical Water Clusters. *J. Chem. Phys.* **2005**, *123*, 084310.

(71) Li, X.; Teige, V. E.; Iyengar, S. S. Can the Four-Coordinated, Penta-Valent Oxygen in Hydroxide Water Clusters Be Detected Through Experimental Vibrational Spectroscopy? *J. Phys. Chem. A* **2007**, *111*, 4815.

(72) Phatak, P.; Sumner, I.; Iyengar, S. S. Gauging the Flexibility of the Active Site in Soybean Lipoxygenase-1 (SLO-1) Through an Atom-Centered Density Matrix Propagation (ADMP) Treatment That Facilitates the Sampling of Rare Events. *J. Phys. Chem. B* **2012**, *116*, 10145.

(73) Phatak, P.; Venderley, J.; Debrota, J.; Li, J.; Iyengar, S. S. Active Site Dynamical Effects That Facilitate the Hydrogen Transfer Process in Soybean Lipoxygenase-1 (SLO-1): Isotope Effects. *J. Phys. Chem. B* **2015**, *119*, 9532.

(74) Vimal, D.; Pacheco, A. B.; Iyengar, S. S.; Stevens, P. S. Experimental and Ab Initio Dynamical Investigations of the Kinetics and Intramolecular Energy Transfer Mechanisms for the OH + 1,3-Butadiene Reaction Between 263 and 423 K at Low Pressure. *J. Phys. Chem. A* **2008**, *112*, 7227.

(75) Pacheco, A. B.; Dietrick, S. M.; Stevens, P. S.; Iyengar, S. S. Pump-Probe" Atom-Centered Density Matrix Propagation Studies to Gauge Anharmonicity and Energy Repartitioning in Atmospheric Reactive Adducts: Case Study of the OH + Isoprene and OH + Butadiene Reaction Intermediates. *J. Phys. Chem. A* **2012**, *116*, 4108.

(76) Dietrick, S. M.; Pacheco, A. B.; Phatak, P.; Stevens, P. S.; Iyengar, S. S. The Influence of Water on Anharmonicity, Stability and Vibrational Energy Distribution of Hydrogen-Bonded Adducts in Atmospheric Reactions: Case Study of the OH + Isoprene Reaction Intermediate Using Ab-Initio Molecular Dynamics. *J. Phys. Chem. A* **2012**, *116*, 399.

(77) Schlegel, H. B.; Frisch, M. J. Computational Bottlenecks In Molecular-Orbital Calculations. *Theoretical and Computational Models for Organic Chemistry*; Formosinho, S. J., Csizmadia, I. G., Csizmadia, I. G., Eds.; Nato Advanced Science Institutes Series, Series C, Mathematical And Physical Sciences; 1991; Vol. *339*, pp 5−33, DOI: 10.1007/978-94-011-3584-9_2.

(78) Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. Efficient, approximate and parallel Hartree-Fock and hybrid DFT calculations. A 'chain-of-spheres' algorithm for the Hartree-Fock exchange. *Chem. Phys.* **2009**, *356*, 98−109.

(79) Kussmann, J.; Beer, M.; Ochsenfeld, C. Linear-scaling self-consistent field methods for large molecules. *Wiley Interdisciplinary Reviews-Computational Molecular Science* **2013**, *3*, 614−636.

(80) Scuseria, G. E. Linear Scaling Density Functional Calculations with Gaussian Orbitals. *J. Phys. Chem. A* **1999**, *103*, 4782.

(81) Goedecker, S. Linear Scaling Electronic Structure Methods. *Rev. Mod. Phys.* **1999**, *71*, 1085.

(82) White, C. A.; Head-Gordon, M. Derivation and Efficient Implementation of the Fast Multipole Method. *J. Chem. Phys.* **1994**, *101*, 6593.

(83) Strain, M. C.; Scuseria, G. E.; Frisch, M. J. *Science* **1996**, *271*, 51.

(84) Valiev, M.; Bylaska, E. J.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Van Dam, H. J. J.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T. L.; de Jong, W. NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations. *Comput. Phys. Commun.* **2010**, *181*, 1477−1489.

(85) Steele, R. P. Multiple-timestep ab initio molecular dynamics using an atomic basis set partitioning. *J. Phys. Chem. A* **2015**, *119*, 12119−12130.

(86) Margul, D. T.; Tuckerman, M. E. A Stochastic, Resonance-Free Multiple Time-Step Algorithm for Polarizable Models That Permits Very Large Time Steps. *J. Chem. Theory Comput.* **2016**, *12*, 2170−2180.

(87) Del Ben, M.; Hutter, J.; VandeVondele, J. Forces and stress in second order Møller-Plesset perturbation theory for condensed phase systems within the resolution-of-identity Gaussian and plane waves approach. *J. Chem. Phys.* **2015**, *143*, 102803.

(88) Ufimtsev, I. S.; Martinez, T. J. Quantum Chemistry on Graphical Processing Units. 3. Analytical Energy Gradients, Geometry Optimization, and First Principles Molecular Dynamics. *J. Chem. Theory Comput.* **2009**, *5*, 2619−2628.

(89) Pomes, R.; Roux, B. Structure and Dynamics of a Proton Wire: A Theoretical Study of H+ Translocation Along the Single-File Water Chain in the Gramicidin a Channel. *Biophys. J.* **1996**, *71*, 19.

(90) Pomes, R.; Roux, B. Theoretical Study of H^+ Translocation Along a Model Proton Wire. *J. Phys. Chem.* **1996**, *100*, 2519.

(91) Decornez, H.; Drukker, K.; Hammes-Schiffer, S. Solvation and Hydrogen-Bonding Effects on Proton Wires. *J. Phys. Chem. A* **1999**, *103*, 2891.

(92) Brewer, M. L.; Schmitt, U. W.; Voth, G. A. The Formation and Dynamics of Proton Wires in Channel Environments. *Biophys. J.* **2001**, *80*, 1691.

(93) Teeter, M. M. Water Structure of a Hydrophobic Protein at Atomic Resolution. *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81*, 6014.

(94) Neidle, S.; Berman, H. M.; Shieh, H. S. Highly Structured Water Network in Crystals of a Deoxydinucleoside-Drug Complex. *Nature* **1980**, *288*, 129.

(95) Lipscomb, L. A.; Peek, M. E.; Zhou, F. X.; Bertrand, J. A.; VanDerveer, D.; Williams, L. D. Water Ring Structure at Dna Interfaces - Hydration and Dynamics of Dna Anthracycline Complexes. *Biochemistry* **1994**, *33*, 3649.

(96) Tu, C.; Rowlett, R. S.; Tripp, B. C.; Ferry, J. G.; Silverman, D. N. Chemical Rescue of Proton Transfer in Catalysis by Carbonic Anhydrases in the Beta- and Gamma-Class. *Biochemistry* **2002**, *41*, 15429.

(97) McEwan, M. J.; Phillips, L. F. *Chemistry of the Atmosphere*; Eward Arnold: London, 1975.

(98) Wayne, R. P. *Chemistry of the Atmosphere*; Clarendon Press: Oxford, 1994.

(99) Agmon, N. Hydrogen Bonds, Water Rotation and Proton Mobility. *J. Chim. Phys.* **1996**, *93*, 1714.

(100) Agmon, N. The Grotthuss Mechanism. *Chem. Phys. Lett.* **1995**, *244*, 456.

(101) Cheng, H.-P. Water Clusters: Fascinating Hydrogen-Bonding Networks, Solvation Shell Structures, and Proton Motion. *J. Phys. Chem. A* **1998**, *102*, 6201.

(102) Wei, D.; Salahub, D. R. Hydrated Proton Clusters: Ab Initio Molecular Dynamics Simulation and Simulated Annealing. *J. Chem. Phys.* **1997**, *106*, 6086.

(103) Tuckerman, M. E.; Laasonen, K.; Sprik, M.; Parrinello, M. Ab Initio Molecular Dynamics Simulation of the Solvation and Transport of H_3O^+ and OH^-. *J. Phys. Chem.* **1995**, *99*, 5749.

(104) Marx, D.; Tuckerman, M. E.; Hutter, J.; Parrinello, M. The Nature of the Hydrated Excess Proton in Water. *Nature* **1999**, *397*, 601.

(105) Sadhukhan, S.; Munoz, D.; Adamo, C.; Scuseria, G. E. Predicting Proton Transfer Barriers with Density Functional Methods. *Chem. Phys. Lett.* **1999**, *306*, 83.

(106) Xie, Y.; Remington, R. B.; Schaefer, H. F., III The Protonated Water Dimer: Extensive Theoretical Studies of $H_5O_2^+$. *J. Chem. Phys.* **1994**, *101*, 4878.

(107) Swanson, J. M. J.; Simons, J. Role of Charge Transfer in the Structure and Dynamics of the Hydrated Proton. *J. Phys. Chem. B* **2009**, *113*, 5149.

(108) Hodges, M. P.; Wales, D. J. Global Minima of Protonated Water Clusters. *Chem. Phys. Lett.* **2000**, *324*, 279.

(109) Okumura, M.; Yeh, L. I.; Myers, J. D.; Lee, Y. T. Infrared Spectra of the Solvated Hydronium Ion: Vibrational Predissociation Spectroscopy of Mass-Selected H_3O^+ $(H_2O)_nH_2$. *J. Phys. Chem.* **1990**, *94*, 3416.

(110) Tsai, C. J.; Jordan, K. D. Theoretical Study of the $(H_2O)_6$ Cluster. *Chem. Phys. Lett.* **1993**, *213*, 181.

(111) Xantheas, S. S. *Ab Initio* Studies of Cyclic Water Clusters $(H_2O)_n$, N = 1−6. II. Analysis of Many body Interactions. *J. Chem. Phys.* **1994**, *100*, 7523.

(112) Xantheas, S. S. *Ab Initio* Studies of Cyclic Water Clusters $(H_2O)_n$, N = 1−6. III. Comparison of Density Functional with MP2 Results. *J. Chem. Phys.* **1995**, *102*, 4505.

(113) Kryachko, E. S. Ab Initio Studies of the Conformations of Water Hexamer: Modelling the Penta-Coordinated Hydrogen-Bonded Pattern in Liquid Water. *Chem. Phys. Lett.* **1999**, *314*, 353.

(114) Sadeghi, R. R.; Cheng, H.-P. The Dynamics of Proton Transfer in a Water Chain. *J. Chem. Phys.* **1999**, *111*, 2086.

(115) Cheng, H.-P.; Krause, J. L. The Dynamics of Proton Transfer in $H_5O_2^+$. *J. Chem. Phys.* **1997**, *107*, 8461.

(116) Wu, C.-C.; Lin, C.-K.; Chang, H.-C.; Jiang, J.-C.; Kuo, J.-L.; Klein, M. L. Protonated Clathrate Cages Enclosing Neutral Water Molecules: $H^+(H_2O)_{21}$ and $H^+(H_2O)_{28}$. *J. Chem. Phys.* **2005**, *122*, 074315.

(117) Yu, H. B.; Cui, Q. The Vibrational Spectra of Protonated Water Clusters: A Benchmark for Self-Consistent-Charge Density-Functional Tight Binding. *J. Chem. Phys.* **2007**, *127*, 234504.

(118) Petersen, M. K.; Iyengar, S. S.; Day, T. J. F.; Voth, G. A. The Hydrated Proton at Water Liquid/Vapour Interfaces. *J. Phys. Chem. B* **2004**, *108*, 14804.

(119) Dietrick, S. M.; Iyengar, S. S. Constructing Periodic Phase Space Orbits from Ab Initio Molecular Dynamics Trajectories to Analyze Vibrational Spectra: Case Study of the Zundel $(H_5O_2^+)$ Cation. *J. Chem. Theory Comput.* **2012**, *8*, 4876.

(120) Kaledin, M.; Kaledin, A. L.; Bowman, J. M. Vibrational Analysis of the $H_5O_2^+$ Infrared Spectrum Using Molecular and Driven Molecular Dynamics. *J. Phys. Chem. A* **2006**, *110*, 2933.

(121) Schwegler, E.; Grossman, J. C.; Gygi, F.; Galli, G. Towards an Assessment of the Accuracy of Density Functional Theory for First Principles Simulations of Water. II. *J. Chem. Phys.* **2004**, *121*, 5400.

(122) Grossman, J. C.; Schwegler, E.; Draeger, E. W.; Gygi, F.; Galli, G. Towards an Assessment of the Accuracy of Density Functional Theory for First Principles Simulations of Water. *J. Chem. Phys.* **2004**, *120*, 300.

(123) Niklasson, A. M. N.; Cawkwell, M. J. Generalized extended Lagrangian Born-Oppenheimer molecular dynamics. *J. Chem. Phys.* **2014**, *141*, 164123.

(124) Becke, A. D. A Multi-Center Numerical Integration Scheme for Polyatomic Molecules. *J. Chem. Phys.* **1988**, *88*, 2547.

(125) Eigen, M.; de Maeyer, L. Self-Dissociation and Protonic Charge Transport in Water and Ice. *Proc. R. Soc. London, Ser. A* **1958**, *247*, 505.

(126) Eigen, M. Proton Transfer, Acid-Base Catalysis, and Enzymatic Hydrolysis. Part I: Elementary Processes. *Angew. Chem., Int. Ed. Engl.* **1964**, *3*, 1.

(127) Zundel, G. In *The Hydrogen Bond-Recent Developments in Theory and Experiments II. Structure and Spectroscopy*; Schuster, P., Zundel, G., Sandorfy, C., Eds.; North-Holland: Amsterdam, 1976; p 683.

(128) Thomas, M.; Brehm, M.; Fligg, R.; Vohringer, P.; Kirchner, B. Computing Vibrational Spectra From Ab Initio Molecular Dynamics. *Phys. Chem. Chem. Phys.* **2013**, *15*, 6608.

(129) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in C*; Cambridge University Press: New York, 1992.

(130) Sumner, I.; Iyengar, S. S. Quantum Wavepacket *Ab Initio* Molecular Dynamics: An Approach for Computing Dynamically Averaged Vibrational Spectra Including Critical Nuclear Quantum Effects. *J. Phys. Chem. A* **2007**, *111*, 10313.

(131) Tangney, P.; Scandolo, S. How Well Do Car-Parrinello Simulations Reproduce the Born-Oppenheimer Surface? Theory and Examples. *J. Chem. Phys.* **2002**, *116*, 14.

(132) Herbert, J. M.; Head-Gordon, M. Curvy-Steps Approach to Constraint-Free Extended-Lagrangian Ab Initio Molecular Dynamics, Using Atom-Centered Basis Functions: Convergence Toward Born Oppenheimer Trajectories. *J. Chem. Phys.* **2004**, *121*, 11542.

(133) Iyengar, S. S.; Schlegel, H. B.; Scuseria, G. E.; Frisch, M. J.; Millam, J. M. Comment On: Curvy-Steps Approach to Constraint-Free Extended-Lagrangian Ab Initio Molecular Dynamics, Using Atom-Centered Basis Functions: Convergence Toward Born Oppenheimer Trajectories. *J. Chem. Phys.* **2005**, *123*, 027101.

(134) Herbert, J. M.; Head-Gordon, M. *J. Chem. Phys.* **2005**, *123*, 027102.

(135) Shin, J.-W.; Hammer, N. I.; Diken, E. G.; Johnson, M. A.; Walters, R. S.; Jaeger, T. D.; Duncan, M. A.; Christie, R. A.; Jordan, K. D. Snapshots of Proton Accommodation at a Microscopic Water Surface: Understanding the Vibrational Spectral Signatures of the Charge Defect in Cryogenically Cooled $H(H_2O)_n^+$=2−28 Clusters. *J. Phys. Chem. A* **2015**, *119*, 9425.

(136) Wei, S.; Shi, Z.; Castleman, A. W., Jr. Mixed Cluster Ions As a Structure Probe: Experimental Evidence for Clathrate Structure of $(H_2O)_{20}H^+$ and $(H_2O)_{21}H^+$. *J. Chem. Phys.* **1991**, *94*, 3268.