X-ray absorption of liquid water by advanced *ab initio* methods

Zhaoru Sun,[1] Mohan Chen,[1] Lixin Zheng,[1] Jianping Wang,[1] Biswajit Santra,[2] Huaze Shen,[3] Limei Xu,[3] Wei Kang,[4] Michael L. Klein,[1,5,6] and Xifan Wu[1,6,*]

[1]*Department of Physics, Temple University, Philadelphia, Pennsylvania 19122, USA*
[2]*Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA*
[3]*International Center for Quantum Materials, School of Physics, Peking University, Beijing 100871, China*
[4]*College of Engineering, Peking University, Beijing 100871, China*
[5]*Department of Chemistry, Temple University, Philadelphia, Pennsylvania 19122, USA*
[6]*Institute for Computational Molecular Science, Temple University, Philadelphia, Pennsylvania 19122, USA*

(Received 17 April 2017; revised manuscript received 6 July 2017; published 11 September 2017)

Oxygen K-edge x-ray absorption spectra of liquid water are computed based on configurations from advanced *ab initio* molecular dynamics simulations, as well as an electron excitation theory from the GW method. One the one hand, the molecular structures of liquid water are accurately predicted by including both van der Waals interactions and a hybrid functional (PBE0). On the other hand, the dynamic screening effects on electron excitation are approximately described by the recently developed enhanced static Coulomb-hole and screened-exchange approximation of W. Kang and M. S. Hybertsen [Phys. Rev. B **82**, 195108 (2010)]. The resulting spectra of liquid water are in better quantitative agreement with the experimental spectra due to the softened hydrogen bonds and the slightly broadened spectra originating from the better screening model.

DOI: 10.1103/PhysRevB.96.104202

I. INTRODUCTION

Water is arguably one of the most important materials on Earth and needs to be thoroughly understood [1]. However, the understanding of liquid water is itself a challenge in many aspects. Unlike other liquids, water shows a lot of anomalies such as the density maximum at 4 °C and the isobaric heat-capacity minimum at 35 °C, among many other things [2,3]. Understanding the microscopic structures and dynamics of liquid water, in particular its hydrogen-bond (HB) network, is the key to understanding these anomalies [4–7]. Recently, high-resolution oxygen K-edge core-level spectroscopy, such as x-ray absorption spectra (XAS) and x-ray Raman scattering, has emerged as a powerful experimental technique to probe the electronic structure and infer the molecular structure of water and ice, as well as aqueous solutions [6,8–22]. Excited from the oxygen 1s core level, the electron excitation probes the unoccupied electronic states, which are antibonding along the covalent OH bonds and particularly sensitive to the HBs. Therefore, the XAS technique serves as a local probe for the HB structures of liquid water and ice.

The experimental XAS of water have three distinct features as a function of increasing excitation energies: a preedge starts from the absorption threshold at 533 to 536 eV with a peak centered at 535 eV, a main edge spans from 537 to 539 eV, and a postedge exists from 539 eV and beyond [11,18,20–22]. Experimentally, the preedge feature is present in both liquid water and ice but is more intense in the former. The relative intensities between the main edge and postedge of liquid water and crystalline ice Ih are substantially different [6,8–11,20–24]. As one of the most qualitative differences, the intensity of the main edge is higher than that of the postedge in the XAS of liquid water, while the opposite trend is true for the spectra in the ice.

The unambiguous assignments of the XAS features to the underlying HB structures are prerequisites for the physical interpretation of the experimental spectra, which can be achieved by first-principles methods including both the modeling of the molecular structure of liquid water and the electron-hole excitation process. With snapshots of the represented molecular configurations from an equilibrated molecular dynamics trajectory, the XAS can be computed with the knowledge of electronic structures of the excited core hole. In this regard, various approximations [17,25,26] for excited electronic states have been proposed within the framework of density functional theory (DFT) [27,28]. In the seminal work of Prendergast and Galli [17], the proposed excited electron and core-hole approximation yielded XAS in close agreement with experimental measurements. More rigorously, the XAS of liquid water can be computed by solving the Bethe-Salpeter equation (BSE) describing the electron-hole interaction [15,29,30]. The BSE approach involves calculations of the self-energy operator and quasiparticles, which are, in general, computationally expensive for liquid water. An approximate way of solving the BSE with fewer computational resources was introduced by Chen *et al.* [9] based on a model electron-screening function in the static Coulomb-hole and screened-exchange (COHSEX) approximation [31], which was evaluated using the maximally localized Wannier functions as basis [32], reducing the computational cost significantly. In the above work, the molecular origins of the spectra features as well as the intensity difference between the main edge and postedge in water and ice have been validated [9].

Despite these recent developments, two uncertainties still remain in the theoretical XAS of liquid water obtained with the approximate solutions of the BSE. The first uncertainty comes from the drawback in modeling the liquid-water structure by employing the generalized gradient approximation (GGA) in the framework of DFT [33–43]. It is known that GGA predicts overstructured liquid water [40–43]. Thus, a significantly elevated temperature is often adopted to generate a softer HB

*Corresponding author: xifanwu@temple.edu

structure closer to the experimental measurement [41–43]. The lack of physics in accurately describing the water structure is due to the neglected van der Waals (vdW) interactions and the spurious self-interaction error [44] in the GGA functional. Specifically, by including the vdW interactions, the water population in the interstitial region between the first and second coordination shells of water molecules is increased to better match experiment [43]. Furthermore, by mitigating the self-interaction error through the hybrid functional, the directional HB strength between water molecules is weakened so it is closer to the experiment; as a result, the protons are less easily donated to neighboring water molecules [43]. However, the effects of this improved water structure on the theoretical XAS have not been elucidated. Second, in the series of computational works adopting the static COHSEX approximation for the electron-hole excitation of liquid water, a homogeneous electronic screening model was first adopted [9] and then extended by using the Hybertsen-Louie ansatz [45] to account for the inhomogeneous screening effects from the molecular environment [10]. However, some discrepancies still exist, and it is not yet clear to what extent the dynamic screening effect will affect the quasiparticle wave functions (QWs) and the computed XAS. For example, it was observed that the width of the theoretical XAS by static COHSEX is slightly narrower than the experimental data [10].

In an effort to address the above issues, we adopted a systematic way to study the XAS of liquid water at ambient conditions. Specifically, we used more advanced *ab initio* modeling of molecular structures and electronic excitations. We generated liquid-water trajectories from the *ab initio* molecular dynamics (AIMD) [46] simulations by employing a hierarchy of exchange-correlation (XC) functionals of Perdew, Burke, and Ernzerhof (PBE) [47], PBE with the vdW interactions in the form of Tkatchenko and Scheffler (PBE+vdW) [48], and the hybrid functional PBE0 [49,50] with vdW interactions (PBE0+vdW). By utilizing the enhanced static COHSEX method to treat the excitations [51], we find that the XAS computed from the snapshot generated by the PBE0+vdW functional agree well with the experiment among the three XC functionals studied. The vdW interactions soften the water structures by increasing the population of water molecules in the interstitial region, while the hybrid functional mitigates the self-interaction error and weakens the HB strength so it is closer to experiment [43]. Both structural corrections and improved excitation theory are crucial in giving rise to an overall improvement of the three edges of XAS. In particular, the postedge feature in the high-energy region of XAS is in slightly better agreement with experiment because of better screening modeling using the enhanced static COHSEX. In addition, we also compare a set of XAS computed from different excitation theories to the experiment in order to show the importance of self-consistently diagonalized QWs in capturing the qualitative features of XAS.

II. METHODS

We performed AIMD simulations to generate liquid-water trajectories using a modified version of the QUANTUM ESPRESSO package [52]. We simulated 128 water molecules in a periodic cubic cell with a cell length of 15.68 Å using the Car-Parrinello molecular dynamics (CPMD) [46] within the canonical (NVT) ensemble. We employed norm-conserving pseudopotentials in the form of Troullier and Martins [53] and set the kinetic-energy cutoff of the electronic wave functions as 71 Ry. We used a hierarchy of XC functionals, including PBE, PBE+vdW, and PBE0+vdW, as mentioned. The hybrid functional PBE0 with a mixing of 25% exact exchange was evaluated in a linear-scaling manner by taking advantage of maximally localized Wannier functions [32]. The ionic temperatures were controlled through Nosé-Hoover chain thermostats with a chain length of 4 for each ion [54–56]. All AIMD simulations were performed at 330 K, where an increase of 30 K has been found to mimic the nuclear quantum effect in structural quantities such as the oxygen-oxygen radial distribution function in DFT-based simulations of liquid water [57]. A time step of 4.0 a.u. and a fictitious electron mass of 300 a.u. were chosen.

We calculated the x-ray absorption cross section using Fermi's golden rule:

$$\sigma(\omega) = 4\pi^2\alpha_0\hbar\omega \sum_f |M_{if}|^2 \delta(\omega_{if} - \omega), \tag{1}$$

where α_0 is the fine-structure constant and $\hbar\omega$ is the absorbed photon energy matching the energy difference $\hbar\omega_{if} = E_f - E_i$. E_f and E_i are the eigenvalues of the final and initial states, respectively. M_{if} are the transition matrix elements between the initial state $|\phi_i\rangle$ and final state $|\phi_f\rangle$, which can be evaluated within the electric-dipole approximation as $M_{if} \sim \langle\phi_i|\, x\, |\phi_f\rangle$, averaging over the three Cartesian directions. We take the $1s$ atomic core wave function from DFT calculations as the initial state $|\phi_i\rangle$. For the final state $|\phi_f\rangle$, we apply a self-consistent diagonalization procedure within the enhanced static COHSEX approach [51], and the details are as follows.

The final state $|\phi_f\rangle$ is obtained by utilizing the enhanced COHSEX approach, which has been implemented within the framework of the CPMD [46] scheme. Specifically, the COHSEX method has been implemented in the CPMD module within the QUANTUM ESPRESSO package [52]. For each set of input wave functions $|\phi_f\rangle$, we fix the ion positions and damp the wave functions of the system in a self-consistent way, as explained below. Note that the formulas we describe here are suitable only for the excitation theory we adopt and are independent of the other CPMD simulations we performed with vdW and PBE0 functionals.

First, the Lagrangian from the Car-Parrinello approach is

$$\mathcal{L} = \frac{\mu}{2}\sum_i \langle\dot{\psi}_i|\dot{\psi}_i\rangle + \tfrac{1}{2}\sum_I M_I\dot{\mathbf{R}}_I^2 - E_{\rm tot}(\mathbf{R},\{\psi\}) + \lambda_{ij}(\langle\psi_i|\psi_j\rangle - \delta_{ij}), \tag{2}$$

where ν is a fictitious mass of electrons and ψ_i is the orbital of state i. M_I is the mass for atom I that is located at $\mathbf{R}_I$. $E_{\rm tot}(\mathbf{R},\{\psi\})$ is the total energy calculated from first-principles methods. The last part is the orthogonality constraint imposed on orbitals by the Lagrangian multiplier to be λ_{ij}. Note that the initial ion positions are fixed, and only wave functions are generated. Therefore, we need to damp the wave functions towards the ground state, which is realized via the equations

of motion of electrons in plane-wave basis set,

$$\mu\ddot{\psi}_i = -H(\mathbf{R},\{\psi\})\psi_i + \sum_j \lambda_{ij}\psi_j. \tag{3}$$

Here, $H(\mathbf{R},\{\psi\})$ is the Hamiltonian matrix of the system. Furthermore, the Hamiltonian part can be evaluated in real space as

$$H\psi(\mathbf{r}) = \left[-\frac{1}{2}\nabla^2(\mathbf{r}) + V_{\text{ext}}(\mathbf{r},\mathbf{R}) + V_H(\mathbf{r})\right]\psi(\mathbf{r}) \tag{4}$$

$$+ \int d\mathbf{r}'\Sigma(\mathbf{r},\mathbf{r}',E)\psi(\mathbf{r}'), \tag{5}$$

where $V_{\text{ext}}(\mathbf{r},\mathbf{R})$ is the external potential and $V_H(\mathbf{r})$ is the Hartree potential. In particular, $\Sigma(\mathbf{r},\mathbf{r}',E)$ is the self-energy operator that is nonlocal in real space and depends on the self-energy E. In the static COHSEX approximation, the self-energy operator can be approximated as

$$\Sigma^{\text{static}}_{\text{COHSEX}}(\mathbf{r},\mathbf{r}') = \Sigma^{\text{static}}_{\text{COH}}(\mathbf{r},\mathbf{r}') + \Sigma^{\text{static}}_{\text{SEX}}(\mathbf{r},\mathbf{r}'). \tag{6}$$

The first part is

$$\Sigma^{\text{static}}_{\text{COH}}(\mathbf{r},\mathbf{r}') = \tfrac{1}{2}\delta(\mathbf{r}-\mathbf{r}')W_p(\mathbf{r},\mathbf{r}';E=0) \tag{7}$$

$$= \tfrac{1}{2}\delta(\mathbf{r}-\mathbf{r}')(W-v), \tag{8}$$

where W_p is the Coulomb hole, W is the screened Coulomb interaction, and v is the bare Coulomb interaction. The Hybertsen-Louie ansatz [45] proposed that W generally follows the local charge density and has the form

$$W(\mathbf{r},\mathbf{r}';E=0) = \tfrac{1}{2}\{W[\mathbf{r}-\mathbf{r}';\rho(\mathbf{r}')] + W[\mathbf{r}'-\mathbf{r};\rho(\mathbf{r})]\}, \tag{9}$$

where W can be written as

$$W[\mathbf{r}'-\mathbf{r};\rho(\mathbf{r})] = \frac{1}{2\pi^3}\int \epsilon^{-1}[\mathbf{q};\rho(\mathbf{r})]v(\mathbf{q})e^{i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')}d\mathbf{q}. \tag{10}$$

Here, we use the Bechstedt model [58] for the dielectric function, 0

$$\epsilon[\mathbf{q},\rho(\mathbf{r})] = 1 + \left[(\epsilon_0-1)^{-1} + \alpha(q/q_{\text{TF}})^2 + q^4\Big/\left(\tfrac{4}{3}k_F^2q_{\text{TF}}^2\right)\right]^{-1}, \tag{11}$$

where q_{TF} is the Thomas-Fermi wave vector, k_{F} is the Fermi wave vector, ϵ_0 is taken from experiment, and α is fixed by matching the Bechstedt model to the q^2 dependence of the Penn model [59]. Next, we can transform W to

$$W[\mathbf{r}'-\mathbf{r};\rho(\mathbf{r})] = \frac{v(\mathbf{r}-\mathbf{r}')}{\epsilon_0} - \frac{1}{a(x_1-x_2)|\mathbf{r}'-\mathbf{r}|} \times \left(\frac{e^{i\sqrt{x_1}|\mathbf{r}'-\mathbf{r}|}}{x_1} - \frac{e^{i\sqrt{x_2}|\mathbf{r}'-\mathbf{r}|}}{x_2}\right). \tag{12}$$

Here, $x_{1,2} = (-b \pm \sqrt{b^2-4ac})/2a$, $a = (\frac{4}{3}k_F^2q_{\text{TF}}^2)^{-1}$, $b = \alpha/q_{\text{TF}}^2$, and $c = \epsilon_0/(\epsilon_0-1)$. The second part is

$$\Sigma^{\text{static}}_{\text{SEX}}(\mathbf{r}',\mathbf{r}) = -\sum_i^{occ}\psi_i(\mathbf{r})\psi_i^*(\mathbf{r}')W(\mathbf{r},\mathbf{r}';E=0). \tag{13}$$

Numerically, it has been shown that most of the error from using the static COHSEX approximation, when compared to the GW approximation, comes from the short-wavelength part of the assumed adiabatic accumulation of the Coulomb hole W_p, namely, the COH part, while the SEX term in the static COHSEX approximation yields values relatively close to those of the GW calculations [51]. Therefore, the enhanced static COHSEX was proposed to introduce a universal function f to approximately include the dynamics screening in the original static model in the COHSEX formula. Specifically, the enhanced static COH term can be evaluated as

$$\Sigma^{\text{new}}_{\text{COH}}(\mathbf{r},\mathbf{r}') = \frac{\delta(\mathbf{r}-\mathbf{r}')}{2}\int W_p(q;E=0)f(q/k_f)e^{-i\mathbf{q}\cdot\mathbf{r}}d\mathbf{q}, \tag{14}$$

where q is a plane wave and k_f is the Fermi vector. The scaling function is

$$f(x) = \frac{1+a_1x+a_2x^2+a_3x^3+a_4x^4+a_5x^5+a_6x^6}{1+b_1x+b_2x^2+b_3x^3+b_4x^4+b_5x^5+b_6x^6}, \tag{15}$$

where $a_1 = 1.9085$, $a_2 = -0.542572$, $a_3 = -2.45811$, $a_4 = 3.08067$, $a_5 = -1.806$, $a_6 = 0.410031$, $b_1 = 2.01317$, $b_2 = -1.55088$, $b_3 = 1.58466$, $b_4 = 0.368325$, $b_5 = -1.68927$, and $b_6 = 0.599225$.

Above, we described the procedures to compute XAS based on an excited water molecule in a snapshot. We then excited every water molecule in the 128-molecule supercell in order to sample the different local environments of the disordered liquid-water structure. In our case, we found converged XAS after randomly exciting 64 water molecules in the snapshot. We note that in a previous study, 10 individual and uncorrelated snapshots of 32 water molecules were chosen, and only small differences were observed between these snapshots [17]. Therefore, we consider one snapshot of a large cell to be sufficient to yield meaningful results and take a representative snapshot from each AIMD trajectory reflecting the equilibrated structure of liquid water to compute XAS.

Due to the different local environments in liquid water for each excited oxygen atom, we adopted the core-hole energy shift of each excitation by following Ref. [60], which is a standard approach to compute the core-level shifts. We used Gaussian broadening of 0.4 eV for all spectra, which was used in previously calculated XAS of liquid water [9,10]. The computed XAS were aligned to the onset of the preedge (535 eV) and then normalized to the same area of experimental data ranging from 533 to 546 eV.

To analyze the real-space locations of QWs in terms of the three edges in XAS of liquid water, we also define the one-dimensional density of the QW as

$$\rho_i(r) = \iint |\psi_i(r,\theta,\phi)^2|d\theta d\phi \tag{16}$$

along the radial direction, and the origin is taken to be the position of the excited oxygen with a core hole. In the equation, i represents the index of the conduction band with excitation energy ε_i. In order to present the different localizations of the QWs, the index i is chosen so that $\varepsilon_i \in [534.5\text{ eV}, 535.5\text{ eV}]$ for QWs representing the preedge

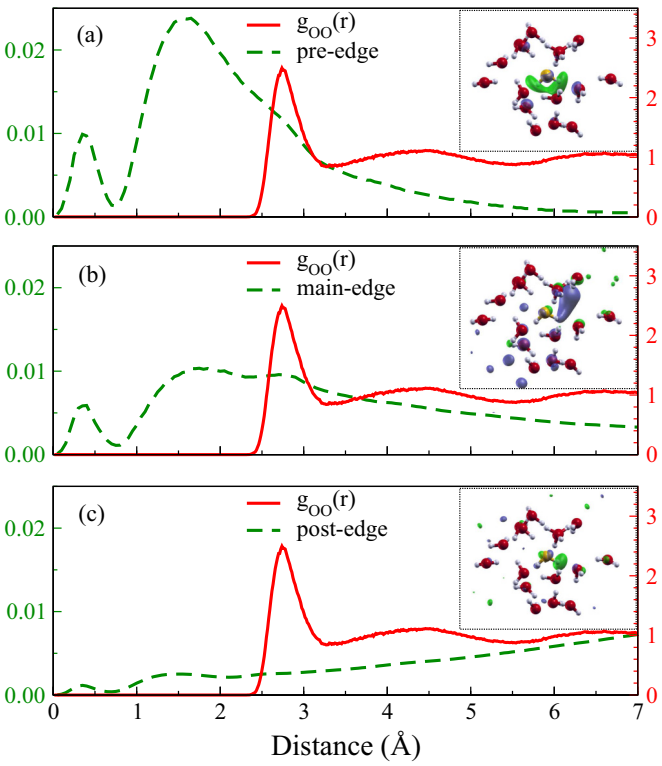

FIG. 1. Density distributions of the QW (green dashed line) of (a) the preedge, (b) main edge, and (c) postedge as a function of oxygen-oxygen distance computed from the snapshot of the PBE0+vdW trajectory using the enhanced static COHSEX method. The $g_{OO}(r)$ (red line) from the PBE0+vdW trajectory is shown for comparison. The insets show the representative QWs of the three edges around the excited water molecule. Water molecules residing within the second coordination shell of the excited oxygen are shown. Red, white, and yellow spheres represent oxygen, hydrogen, and oxygen atoms with a core hole, respectively. QWs with opposite signs are depicted in blue and green.

of XAS, $\varepsilon_i \in [537\ \text{eV}, 539\ \text{eV}]$ for QWs representing the mainedge of XAS, and $\varepsilon_i \in [540\ \text{eV}, 542\ \text{eV}]$ for QWs representing the postedge of XAS.

III. RESULTS AND DISCUSSION

A. HB structure probed by the preedge, main edge, and postedge of XAS

The QWs can be strongly perturbed by the local liquid-water structures. The preedge, main edge, and postedge of XAS are found to have distinguishable molecular signatures that relate to different spatial regions of the HB network [8,9,18]; the preedge has $4a_1$ character, while the main edge and postedge have b_2 character, both of which originate from the molecular excitations in the gas phase [8]. In order to quantitatively study the spatial regions in terms of different XAS edges, we present the density distributions of QWs as a function of oxygen-oxygen distance in Fig. 1. The oxygen-oxygen radial distribution function $g_{OO}(r)$ and an excited oxygen with QWs distributed within the HB network (insets) are also shown for comparison. Overall, Figs. 1(a), 1(b), and 1(c) show that the density distributions of QWs become more delocalized from the preedge to the main edge to the postedge.

The density distribution of QWs of the preedge illustrated in Fig. 1(a) has the highest peak at 1.7 Å and is mostly localized within 2.75 Å, the latter of which is the first peak position of $g_{OO}(r)$. The inset in Fig. 1(a) shows that the QW of the preedge resembles the first excited state of a water molecule in the gas phase with $4a_1$ symmetry. In this regard, our result is consistent with a previous assignment [9] of the preedge to a bound exciton state, where the electron orbital was found to be mostly localized within the first coordination shell. Therefore, the preedge features can be largely affected by the short-range structures, such as broken HBs and covalent bond strength, around the excited oxygens. With the weakened HBs described by the hybrid DFT functional (PBE0) and vdW interactions, that is, the short-range HB network, therefore, the computed preedge of XAS is expected to be improved in both energies and intensities.

As shown in Fig. 1(b), the density distribution of the QW of the main edge is more delocalized than that of the preedge. The inset shows that the QW of the main edge can be found not only on the excited molecule itself but also on its first- and second-shell neighbors. In addition, a clear b_2 character can be identified for a typical QW of the main edge. The above is consistent with the fact that the main edge was found to originate from the second excited state of a water molecule in the gas phase. By comparing the localization of the main-edge density of the QW to that of the preedge, it can be seen that the former is more localized between the first and second coordination shells of the liquid-water structure. In this regard, it is expected that the main-edge feature of XAS will be sensitive to the intermediate-range order of the liquid-water structure, i.e., water molecules in the interstitial region.

In contrast to the density distributions of QWs of the preedge and main edge, the QW of the postedge shown in Fig. 1(c) is the most delocalized one but still preserves b_2 character. The strong delocalization can be clearly seen by the increased density of the QW as a function of the distance away from the excited water molecule. Hence, the water molecules in long-range order are critical in determining the postedge features. Using a small simulation cell containing 32 water molecules fails to yield a postedge intensity of XAS in close agreement with experimental data [9]. The delocalization feature is consistent with the fact that the postedge is a resonant exciton state.

B. XAS calculated from PBE, PBE+vdW, and PBE0+vdW AIMD trajectories

Among the three levels of XC functionals investigated, the XAS computed from the snapshot obtained with PBE show the least agreement with the experimental spectra in Fig. 2(a). Four major discrepancies can be identified. First, the intensity of the computed preedge is underestimated compared to experiment. Second, the theoretical main edge is centered around 538.5 eV, showing a large blueshift (around 1 eV) compared to the experimental value at 537.5 eV. Third, a significantly overestimated postedge intensity leads to the fact that the main edge and post edge have almost the same intensities, which contradicts the experimental fact

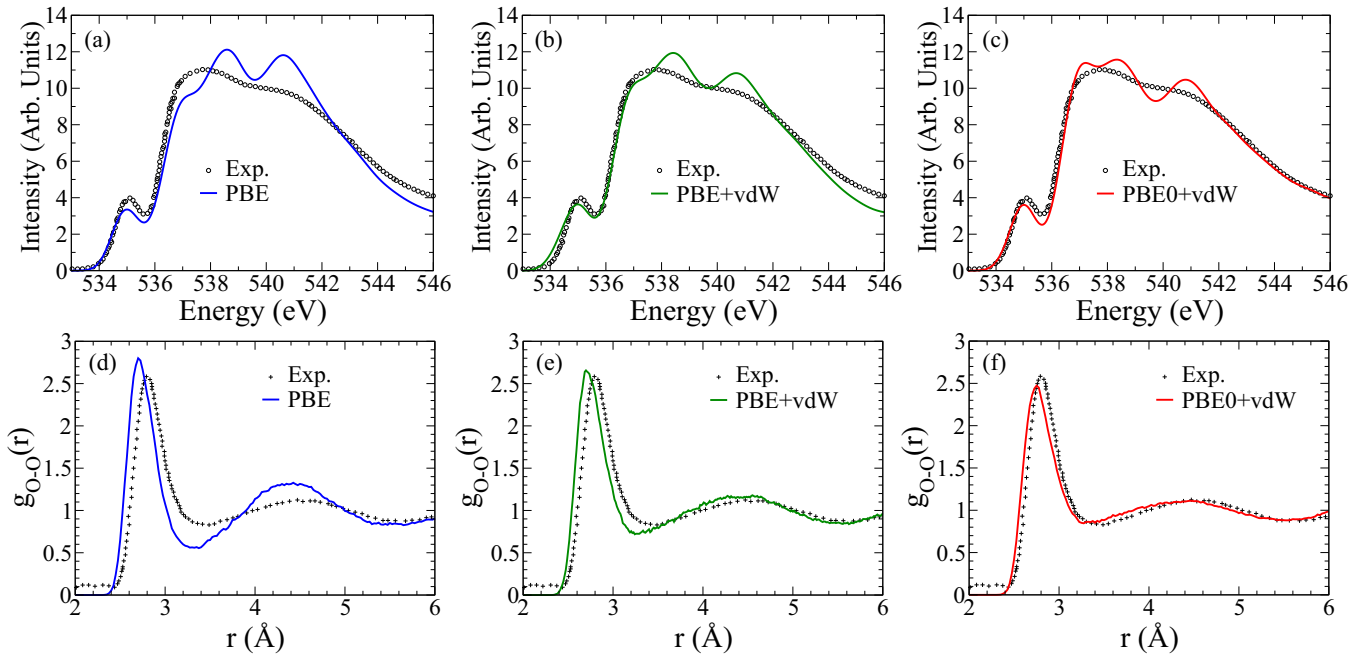

FIG. 2. Computed XAS and $g_{OO}(r)$ from three levels of exchange-correlation functionals used in the AIMD simulations, namely, PBE, PBE+vdW, and PBE0+vdW. A representative snapshot consisting of 128 water molecules from each equilibrated AIMD trajectory was used for spectra calculation. The enhanced static COHSEX method is adopted as the excitation theory. The experimental (Exp.) data of XAS [21] and $g_{OO}(r)$ [61] are also shown for comparison.

that the main edge is more prominent than the postedge in liquid water. Fourth, the width of the computed XAS is slightly narrower than the experimental XAS, especially in the high-energy region. Experimentally, the XAS of crystalline ice Ih with an intact HB network show a more prominent peak of the postedge than that of the main edge. Therefore, the discrepancies imply that the HB network of liquid water from the PBE AIMD trajectory is overstructured. Indeed, $g_{OO}(r)$ computed from the PBE AIMD trajectory significantly deviates from experimental measurement [61], as shown in Fig. 2(d). Specifically, the first and second peaks of $g_{OO}(r)$ from simulations are significantly overestimated, and the first minimum is largely underestimated. Consistently, the average number of HBs per water molecule is found to be 3.76 in the PBE trajectory according to the popular HB definition proposed by Luzar and Chandler [62], which is the highest among all the functionals studied herein. All the above indicate that the HB network of liquid water is overstructured from the PBE AIMD trajectory. Not surprisingly, the theoretical XAS predicted by an overstructured HB network from the PBE AIMD trajectory yield icelike spectra with a relatively more prominent postedge feature.

As shown in Fig. 2(b), the XAS computed from the PBE+vdW trajectory are largely improved with regard to the experimental spectra compared to those obtained from the PBE trajectory. The improvement can be seen by a higher preedge intensity, a shift of the main edge towards lower energy, and a lower postedge intensity. The better agreement is attributed to the improved description of the HB network by including the vdW interactions in the AIMD simulation. An explicit account of vdW forces strengthens the attractive interactions among water molecules, which significantly increase the population of water molecules in the interstitial region. The increased population of water molecules brings $g_{OO}(r)$ in closer agreement with experiment within the first and second coordination shells. Moreover, the increased population of water molecules in the interstitial region weakens the HBs among the water molecules in the first coordination shell, resulting in a reduced first peak in $g_{OO}(r)$. The average number of HBs per molecule is found to be 3.56, which is about 5% smaller than that of PBE. Hence, an excited oxygen atom experiences a more disordered environment by the surrounding water molecules. First of all, the preedge intensity is increased due to the breaking of more HBs. In order to verify this, we selected two excited water molecules, one with broken HBs and the other one with four intact HBs, and plotted their density distributions of QWs in Fig. 3(a). The QW of the preedge from the excited molecule with the broken HBs is more localized with enhanced p character due to the more disordered short-range molecular environment. Therefore, larger amplitudes of transition matrices M_{ij} are obtained according to Fermi's golden rule, as shown in Eq. (1). As a result, the preedge intensity is increased with more broken HBs. Second, the main-edge intensity is also enhanced by an increased population of water molecules in the interstitial region due to vdW interactions. In order to further explain how the water molecules in the interstitial region affect the main-edge features, we also selected two representative excited water molecules and calculated their density distributions of QWs; one of the excited water molecules has four intact HBs, and the other one has extra neighboring water molecules within the interstitial region in addition to the four intact HBs. Figure 3(b) suggests that the density of the main-edge QW of the excited molecule with extra neighboring water

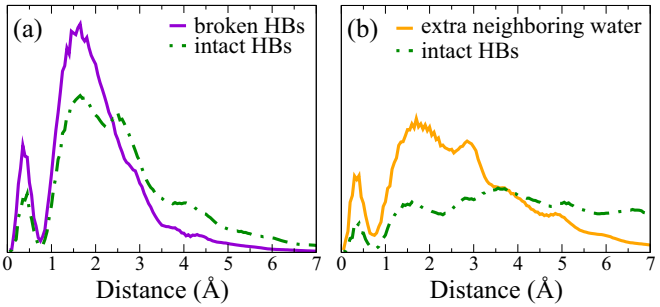

FIG. 3. Density distributions of QWs computed from different excited water molecules within a snapshot of 128 water molecules for the (a) preedge and (b) main-edge regions of XAS. The QWs are chosen based on different H-bonding environments and different electronic states. In (a), density distributions of QWs are shown for two water molecules with intact HBs (green dashed line) and broken HBs (purple solid line). Note the intact HBs refer to two accepting and two donating HBs, and prominent peaks are observed for water molecules with broken HBs, as shown in (a). In (b), density distributions of QWs are shown for two water molecules with intact HBs. In particular, the presence of an extra nonbonded neighboring water molecules within the interstitial region (orange solid line) results in more prominent peaks, as shown in (b).

molecules within the interstitial region is more localized than the other by the disordered molecular environment in the intermediate range. Like the discussion of the amplitudes of transition matrices and the QW localization associated with the preedge feature, we conclude that the increased population of water molecules within the interstitial region leads to larger amplitudes of transition matrices in the main edge. The above discussion explains the improved main-edge features, namely, the increased intensity and shift of the peak position (centering at 538.2 eV) towards the experimental value (centering at 537.5 eV). The postedge feature, whose QW is orthogonal to that of the mainedge, is therefore predicted to have a lower spectral intensity.

The theoretical XAS are further improved with respect to the experimental measurements by using the molecular configurations generated from the PBE0+vdW AIMD trajectory, as shown in Fig. 2(c). By mixing a fraction of exact exchange in the hybrid functional, PBE0 mitigates the self-interaction error and lowers the tendency of hydrogen atoms to be donated to neighboring water molecules. Under the influence of exact exchange, the covalent bonds of water molecules are strengthened, resulting in a shorter OH bond with the weakened directional HBs. Consequently, the average number of HBs is reduced to 3.48 in the PBE0+vdW AIMD trajectory, which is lower than the ones from the PBE and PBE+vdW AIMD trajectories. As a result, the water structure is further softened with a larger fraction of broken HBs, in agreement with the lower of the first peak of $g_{OO}(r)$, as illustrated in Fig. 2(f). Moreover, the width of the first peak in $g_{OO}(r)$ is broadened, and the position is slightly increased with respect to the experimental direction, in agreement with the weakened directional HB strength. The revised HB network in liquid water essentially affects the computed XAS in the following three ways. First, the mitigated self-interaction error included in the hybrid functional revises the short-range order of the HB network in such a way that the excited water molecules experience a more disordered molecular environment comprising its first coordination shell, which is sensitively probed by the XAS technique in the preedge region. Using the same argument as in the previous paragraph, the QW of the preedge with broken HBs is more localized with enhanced p character, resulting in the more prominent preedge intensity with more broken HBs. Therefore, the intensity of the preedge from the PBE0+vdW trajectory is slightly increased towards that of the experimental spectra, reflecting the softer liquid structure at short range. Second, as a direct consequence of the weakened directional HB strength, more nonbonded water molecules flow into the interstitial region, as shown by the increased first minimum with respect to the experimental measure in Fig. 2(f). As a result, the intermediate-range order of the HB network is revised as well, in which the excited water molecule is placed in a more disordered molecular environment beyond the first coordination shell. As in the previous discussion of the effect of vdW interactions on the main edge, the main-edge features of XAS, which are sensitive to the intermediate-range order of the HB network, are modified with increased intensity ranging from 537 to 538 eV. In addition, the peak position of the main edge is shifted to 537.9 eV, which is closer to the experimental peak at 537.5 eV. The changes in the main-edge features provide better agreement with experimental spectra. Third, the postedge becomes more delocalized with decreased intensity, shifting towards higher energies. The peak position of the postedge is 540.8 eV, which is slightly smaller than the experimental value of 541.0 eV. The more prominent feature of the main edge than that of the postedge is best captured by the PBE0+vdW trajectory. In particular, the width of XAS is broadened in the high-energy region (544 to 546 eV), matching experiments better.

C. Enhanced static COHSEX and self-consistently diagonalized QWs

Besides the accurate description of water structures gained by including the vdW interactions and hybrid functional (PBE0), proper treatment of excitations is also critical for obtaining accurate XAS. The approximation of the electron self-energy has been widely used in electron excitation problems but is computationally very expensive in simulating XAS of liquid water. In order to reduce the computational cost, excitation theories based on the COHSEX method with electron screening models were used [9,10]. Compared to the GW methods, the static COHSEX method suffers from the wavelength-dependent error from the COH part, which is more significant at short wavelength but negligible at long wavelength [51]. As a result, the high-energy region of XAS is affected but can be improved by using the enhanced static COHSEX method [51].

Figure 4 shows the XAS of liquid water computed using both static COHSEX and enhanced static COHSEX methods, and the water structure is chosen from the PBE0+vdW trajectory. By using the enhanced static COHSEX method, the XAS of liquid water are improved with respect to experiment when compared to the spectra calculated from the static COHSEX. For example, the intensity of the postedge region

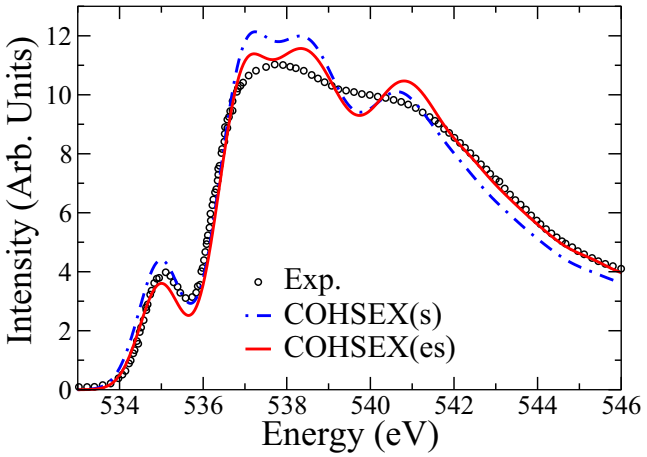

FIG. 4. Comparison of XAS computed between static COHSEX [COHSEX(s)], enhanced static COHSEX [COHSEX(es)], and the experimental data [21]. A representative snapshot from the PBE0+vdW AIMD trajectory containing 128 water molecules was used for all spectra calculations. Self-consistent calculations were carried out to update the input wave functions.

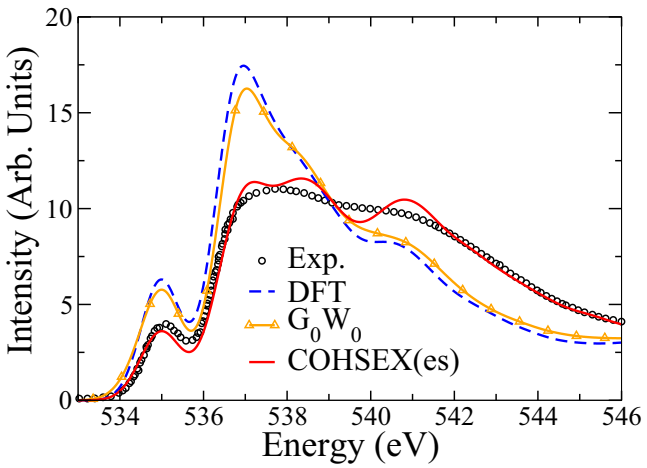

FIG. 5. Comparison of XAS computed from three different excitation schemes and the experiment [21]. A representative snapshot from the PBE0+vdW AIMD trajectory containing 128 water molecules was used for all spectra calculations. The first excitation scheme by GGA-level DFT (blue dashed line) with the FCH approximation is a one-step calculation without changing the input wave functions. Second, the G_0W_0 (orange line) scheme was also performed with one-step diagonalizaiton. Third, the enhanced static COHSEX [COHSEX(es); red solid line] was adopted, and a self-consistent calculation was carried out to update the input wave functions.

(>541 eV) increases while intensities of the preedge and main edge decrease, as obtained from the enhanced static COHSEX. This is due to the broadening effect introduced by considering the dynamic screening in the enhanced static COHSEX approach. The above broadening effect is more significant in the short-wavelength part, as can be seen by the corrections on the XAS in the relatively high excitation energy region. By normalizing the XAS of both experiment and theory to the same area, the increase in the high-energy spectra also leads to the decrease of the preedge and main-edge intensities.

In order to emphasize the importance of self-consistently diagonalized QWs, we further compare the XAS obtained from three different excitation treatments: (i) the full core-hole (FCH) approximation, where both the energy levels and wave functions are from unoccupied Kohn-Sham eigenstates, (ii) a perturbative treatment based on the G_0W_0 approximation, in which the quasiparticle energies are from the excitation theory but the QWs are still approximated by the Kohn-Sham eigenfunctions, and (ii) the current enhanced static COHSEX approach in which both the excitation energies and the QWs are generated by the self-consistently diagonalized self-energy operator. The resulting spectra yielded by these three different approaches are reported in Fig. 5. The G_0W_0 approximation has been widely used in calculating the band structures of water [63], aqueous solutions [64], and many other organic systems [65–70]. However, calculations of the XAS significantly depend on the QWs as they are explicitly involved in the evaluation of the transition matrix elements M_{ij} in Eq. (1). Not surprisingly, the use of the wave functions from the FCH approximation in DFT and G_0W_0 approach leads to very similar predicted XAS. As a result, XAS from both the FCH and G_0W_0 approaches yield a narrowed spectra width with spectral features significantly deviating from experiments, as shown in Fig. 5. In sharp contrast, XAS from a self-consistently diagonalized self-energy operator within the enhanced static COHSEX approach are in much better agreement with the experimental spectra.

IV. SUMMARY

We reported systematic modeling of the liquid-water XAS using advanced *ab initio* approaches. We found that the computed XAS agree better with the experiment spectra based on the liquid structure generated from AIMD when including the vdW interactions and hybrid functional (PBE0). The predicted XAS are further improved by the enhanced static COHSEX method that includes the approximate dynamic screening. Specifically, the features of the three edges (preedge, main edge, and postedge) of liquid water are improved. First, the short-range order in liquid water is improved by the inclusion of the vdW interactions and PBE0 functional. The vdW interactions increase the population of water molecules within the interstitial region and therefore weaken the short-range HBs. Under the influence of exact exchange by PBE0, the covalent OH bonds are shorter, which further weakens the directional HB strength. As a result, the preedge intensity is increased by the presence of more broken HBs in liquid water. An accurate description of the intermediate-range order of water, especially the water molecules in the interstitial region, is crucial to get a main-edge intensity and peak position close to experiment. The inclusion of both vdW interactions and the hybrid functional increases the population of water molecules in the interstitial region, resulting in main-edge features that agree well with the experimental spectra. Last, the overall spectra are improved by considering the dynamic screening effects in the enhanced static COHSEX. In conclusion, the PBE0+vdW AIMD trajectory with the

enhanced static COHSEX yields better theoretical XAS. QWs from the diagonalization of the GW self-energy instead of the DFT wave functions are crucial for obtaining spectra that quantitatively agree with experiments.

ACKNOWLEDGMENTS

The authors thank R. Car and G. Galli for helpful discussions. This work was mainly supported by the National Science Foundation (NSF), DMR, under Grant No. DMR-1552287 (design of the project). M.C. and M.L.K. were supported by U.S. Department of Energy SciDAC under Grant No. DE-SC0008726 (hybrid functional and van der Waals AIMD algorithms). This research used computational resources of the National Energy Research Scientific Computing Center (NERSC), a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

[1] P. Ball, Chem. Rev. **108**, 74 (2008).
[2] P. Gallo, K. Amann-Winkel, C. A. Angell, M. A. Anisimov, F. Caupin, C. Chakravarty, E. Lascaris, T. Loerting, A. Z. Panagiotopoulos, J. Russo *et al.*, Chem. Rev. **116**, 7463 (2016).
[3] Z. Sun, G. Sun, Y. Chen, and L. Xu, Sci. China Phys. Mech. Astron. **57**, 810 (2014).
[4] J. C. Palmer, F. Martelli, Y. Liu, R. Car, A. Z. Panagiotopoulos, and P. G. Debenedetti, Nature (London) **510**, 385 (2014).
[5] A. Nilsson and L. G. Pettersson, Chem. Phys. **389**, 1 (2011).
[6] M. Leetmaa, M. Ljungberg, A. Lyubartsev, A. Nilsson, and L. G. Pettersson, J. Electron Spectrosc. Relat. Phenom. **177**, 135 (2010).
[7] C. W. Swartz and X. Wu, Phys. Rev. Lett. **111**, 087801 (2013).
[8] P. Wernet, D. Nordlund, U. Bergmann, M. Cavalleri, M. Odelius, H. Ogasawara, L. Å. Näslund, T. K. Hirsch, L. Ojamäe, P. Glatzel *et al.*, Science **304**, 995 (2004).
[9] W. Chen, X. Wu, and R. Car, Phys. Rev. Lett. **105**, 017802 (2010).
[10] L. Kong, X. Wu, and R. Car, Phys. Rev. B **86**, 134203 (2012).
[11] S. T. John, D. M. Shaw, D. D. Klug, S. Patchkovskii, G. Vankó, G. Monaco, and M. Krisch, Phys. Rev. Lett. **100**, 095502 (2008).
[12] T. Head-Gordon and M. E. Johnson, Proc. Natl. Acad. Sci. USA **103**, 7973 (2006).
[13] J. D. Smith, C. D. Cappa, B. M. Messer, W. S. Drisdell, R. C. Cohen, and R. J. Saykally, J. Phys. Chem. B **110**, 20038 (2006).
[14] D. Nordlund, H. Ogasawara, K. Andersson, M. Tatarkhanov, M. Salmerón, L. Pettersson, and A. Nilsson, Phys. Rev. B **80**, 233404 (2009).
[15] J. Vinson, J. Kas, F. Vila, J. J. Rehr, and E. Shirley, Phys. Rev. B **85**, 045101 (2012).
[16] O. Fuchs, M. Zharnikov, L. Weinhardt, M. Blum, M. Weigand, Y. Zubavichus, M. Bär, F. Maier, J. D. Denlinger, C. Heske *et al.*, Phys. Rev. Lett. **100**, 027801 (2008).
[17] D. Prendergast and G. Galli, Phys. Rev. Lett. **96**, 215502 (2006).
[18] A. Nilsson, D. Nordlund, I. Waluyo, N. Huang, H. Ogasawara, S. Kaya, U. Bergmann, L.-Å. Näslund, H. Öström, P. Wernet *et al.*, J. Electron Spectrosc. Relat. Phenom. **177**, 99 (2010).
[19] T. Fransson, I. Zhovtobriukh, S. Coriani, K. T. Wikfeldt, P. Norman, and L. G. Pettersson, Phys. Chem. Chem. Phys. **18**, 566 (2016).
[20] J. A. Sellberg, S. Kaya, V. H. Segtnan, C. Chen, T. Tyliszczak, H. Ogasawara, D. Nordlund, L. G. Pettersson, and A. Nilsson, J. Chem. Phys. **141**, 034507 (2014).
[21] S. Schreck and P. Wernet, J. Chem. Phys. **145**, 104502 (2016).
[22] M. Nagasaka, H. Yuzawa, T. Horigome, and N. Kosugi, J. Electron Spectrosc. Relat. Phenom. (2017), doi: 10.1016/j.elspec.2017.05.004.
[23] T. Pylkkanen, V. M. Giordano, J.-C. Chervin, A. Sakko, M. Hakala, J. A. Soininen, K. Hamalainen, G. Monaco, and S. Huotari, J. Phys. Chem. B **114**, 3804 (2010).
[24] C. J. Sahle, C. Sternemann, C. Schmidt, S. Lehtola, S. Jahn, L. Simonelli, S. Huotari, M. Hakala, T. Pylkkänen, A. Nyrow *et al.*, Proc. Natl. Acad. Sci. USA **110**, 6301 (2013).
[25] B. Hetényi, F. De Angelis, P. Giannozzi, and R. Car, J. Chem. Phys. **120**, 8632 (2004).
[26] M. Cavalleri, M. Odelius, A. Nilsson, and L. G. Pettersson, J. Chem. Phys. **121**, 10065 (2004).
[27] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[28] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[29] J. Rehr, J. Soininen, and E. L. Shirley, Phys. Scr. **2005**, 207 (2005).
[30] J. Vinson, J. Rehr, J. Kas, and E. Shirley, Phys. Rev. B **83**, 115106 (2011).
[31] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).
[32] X. Wu, A. Selloni, and R. Car, Phys. Rev. B **79**, 085102 (2009).
[33] D. Asthagiri, L. R. Pratt, and J. Kress, Phys. Rev. E **68**, 041505 (2003).
[34] J. C. Grossman, E. Schwegler, E. W. Draeger, F. Gygi, and G. Galli, J. Chem. Phys. **120**, 300 (2004).
[35] E. Schwegler, J. C. Grossman, F. Gygi, and G. Galli, J. Chem. Phys. **121**, 5400 (2004).
[36] J. VandeVondele, F. Mohamed, M. Krack, J. Hutter, M. Sprik, and M. Parrinello, J. Chem. Phys. **122**, 014515 (2005).
[37] S. Yoo and S. S. Xantheas, J. Chem. Phys. **134**, 121105 (2011).
[38] A. Møgelhøj, A. K. Kelkkanen, K. T. Wikfeldt, J. Schiøtz, J. J. Mortensen, L. G. Pettersson, B. I. Lundqvist, K. W. Jacobsen, A. Nilsson, and J. K. Nørskov, J. Phys. Chem. B **115**, 14149 (2011).
[39] J. Wang, G. Román-Pérez, J. M. Soler, E. Artacho, and M.-V. Fernández-Serra, J. Chem. Phys. **134**, 024516 (2011).
[40] S. Yoo, X. C. Zeng, and S. S. Xantheas, J. Chem. Phys. **130**, 221102 (2009).
[41] C. Zhang, D. Donadio, F. Gygi, and G. Galli, J. Chem. Theory Comput. **7**, 1443 (2011).
[42] C. Zhang, J. Wu, G. Galli, and F. Gygi, J. Chem. Theory Comput. **7**, 3054 (2011).
[43] R. A. DiStasio, Jr., B. Santra, Z. Li, X. Wu, and R. Car, J. Chem. Phys. **141**, 084502 (2014).
[44] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
[45] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **37**, 2733 (1988).
[46] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

[47] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[48] A. Tkatchenko and M. Scheffler, Phys. Rev. Lett. **102**, 073005 (2009).
[49] J. P. Perdew, M. Ernzerhof, and K. Burke, J. Chem. Phys. **105**, 9982 (1996).
[50] C. Adamo and V. Barone, J. Chem. Phys. **110**, 6158 (1999).
[51] W. Kang and M. S. Hybertsen, Phys. Rev. B **82**, 195108 (2010).
[52] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo *et al.*, J. Phys.: Condens. Matter **21**, 395502 (2009).
[53] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).
[54] S. Nosé, J. Chem. Phys. **81**, 511 (1984).
[55] W. G. Hoover, Phys. Rev. A **31**, 1695 (1985).
[56] G. J. Martyna, M. L. Klein, and M. Tuckerman, J. Chem. Phys. **97**, 2635 (1992).
[57] J. A. Morrone and R. Car, Phys. Rev. Lett. **101**, 017801 (2008).
[58] F. Bechstedt, R. Del Sole, G. Cappellini, and L. Reining, Solid State Commun. **84**, 765 (1992).
[59] D. R. Penn, Phys. Rev. **128**, 2093 (1962).
[60] E. Pehlke and M. Scheffler, Phys. Rev. Lett. **71**, 2338 (1993).
[61] L. B. Skinner, C. Huang, D. Schlesinger, L. G. Pettersson, A. Nilsson, and C. J. Benmore, J. Chem. Phys. **138**, 074506 (2013).
[62] A. Luzar and D. Chandler, Phys. Rev. Lett. **76**, 928 (1996).
[63] T. A. Pham, C. Zhang, E. Schwegler, and G. Galli, Phys. Rev. B **89**, 060202 (2014).
[64] D. Opalka, T. A. Pham, M. Sprik, and G. Galli, J. Chem. Phys. **141**, 034501 (2014).
[65] X. Blase, C. Attaccalite, and V. Olevano, Phys. Rev. B **83**, 115103 (2011).
[66] N. Marom, F. Caruso, X. Ren, O. T. Hofmann, T. Körzdörfer, J. R. Chelikowsky, A. Rubio, M. Scheffler, and P. Rinke, Phys. Rev. B **86**, 245127 (2012).
[67] A. Droghetti, M. Cinchetti, and S. Sanvito, Phys. Rev. B **89**, 245137 (2014).
[68] T. Körzdörfer and N. Marom, Phys. Rev. B **86**, 041110 (2012).
[69] J. W. Knight, X. Wang, L. Gallandi, O. Dolgounitcheva, X. Ren, J. V. Ortiz, P. Rinke, T. Körzdörfer, and N. Marom, J. Chem. Theory Comput. **12**, 615 (2016).
[70] M. J. van Setten, F. Caruso, S. Sharifzadeh, X. Ren, M. Scheffler, F. Liu, J. Lischner, L. Lin, J. R. Deslippe, S. G. Louie *et al.*, J. Chem. Theory Comput. **11**, 5665 (2015).