I. INTRODUCTION

Recent studies using x-ray absorption and Raman spectroscopy have focused on the details of H-bonding environments around specific molecules in bulk water and at the water surface. Comparisons of the oxygen K-edge spectra in hexagonal ice, liquid, and gas phases have shown sharp bands in the gas phase which broaden and shift in the condensed phases. A pre-edge absorption feature below the main absorption edge is found in both ice and liquid, however, this feature is stronger in the liquid phase. A comparison of measurements using surface versus bulk sensitive absorption methods was also reported indicating that x-ray-absorption spectra are very different depending on the level of bulk penetration of the experimental method. Spectra taken by a surface sensitive technique are reminiscent of gas-phase spectra with sharp distinguishable bands, while the bulk sensitive method results in a broadened pre-edge and main edge as in Ref. 2. Interestingly there are differences between the pre-edge/main-edge ratios reported in Ref. 2 and Refs. 5 and 6 for water, which could be due to experimental limitations such as the difficulty in estimating the degree of bulk penetration and saturation effects in x-ray fluorescence measurements.

Calculations based on density-functional theory using clusters of water molecules generated via classical potentials enabled the interpretation of the growth of the pre-edge feature in water as a signature of a decrease in the extent of donor H bonding around the absorbing molecule. The growth in pre-edge in bulk water was attributed to water molecules with one donor hydrogen bond broken. In ice no pre-edge feature was found from calculation. The sharp structure at the surface was shown to be due to water molecules with no donor bonds at all.

In this study we calculate the complete near-edge x-ray-absorption fine structure (NEXAFS) of bulk water simulated by first-principles molecular dynamics (FPMD). Besides water we also calculate the spectra for the gas and hexagonal ice phases, and wherever possible we make comparisons with the recent experiments mentioned above. Furthermore, we compare two different methods of calculating the spectra, one in which valence relaxation is achieved by using a full core hole (FCH) to model the x-ray excitation and one in which a half core hole (HCH) is used. Our calculations result in good qualitative agreement with experiment with FCH generally giving an edge structure closer to experiment than HCH.

We find that a well defined pre-edge structure is present both in ice and in liquid water. In agreement with experiment, this feature is stronger in water than in ice, an effect that we attribute to a partial rupture of the H-bond network. In particular we confirm the result of Refs. 2 and 3 in that the growth in pre-edge is due mainly to the breaking of a donor H bond on the absorbing molecule, moreover, smaller but qualitative changes can be attributed to the breaking of an acceptor H bond. Our pre-edge to main-edge ratio is more in agreement with those reported in Refs. 5 and 6 than the ones reported in Refs. 2 and 3. Furthermore, we find that the fraction of broken H bonds in liquid water is 19% compared to hexagonal ice, where each molecule is tetrahedrally H bonded. This result is in reasonable agreement with the one obtained from comparing the sublimation and melting energies of water.

In the x-ray-absorption process a core electron is excited into a near-edge conduction state. The shape of the conduc-
atom the Bethe–Salpeter equation simplifies to a single-particle calculation. This approach has been recently applied with success to the calculation of the x-ray-absorption spectra of several crystals. These calculations used a model dielectric function to describe the screened electron-hole interaction and they are quite demanding computationally, since a calculation is performed for each individual excitation.

In this study we adopt approximate approaches to NEXAFS based on density-functional theory (DFT). The relaxation of the valence electrons due to the core hole is determined self-consistently. The electron excitation can be modeled by the transition state method introduced by Slater, where half an electronic occupation is promoted from a core state to an excited state. The excitation energy corresponds to the difference between the final and initial state eigenvalues, and includes relaxation effects up to second order. Alternatively one may calculate excitation energies as differences in total energy (DSCF) between the system in which a whole electron is promoted from a core state to an excited state and the system in the ground state. Straightforward implementation of either the Slater transition state theory or the DSCF approach requires doing a DFT calculation for each individual excited state.

A computationally less demanding alternative is to calculate the virtual orbitals of a system in the presence of a full core hole (FCH) or a half core hole (HCH) on the absorbing nucleus. In these calculations the excited electron is usually removed from the system. For delocalized final states the neglect of the dependence of the spectrum on the final occupied state is expected to be small. In the case of small test molecules and clusters HCH has provided good agreement with experiment, while FCH seems to be preferred for condensed phase systems. For water our calculations show that while the edge structure in the gas phase is reproduced for both HCH and FCH, FCH compares better with experiment in the condensed phases.

Our paper is organized as follows: In the following section we describe the methods we used in obtaining spectra (orbital reconstruction method). Since the Car–Parrinello method and in particular its variant that is combined with ultrasoft pseudopotentials that we use in this study have been described elsewhere we omit its description here. In Sec. III we provide our simulation details. In Sec. IV we present our results and analysis, and in Sec. V we conclude our paper.

II. SPECTRAL CALCULATION AND ORBITAL RECONSTRUCTION

The spectral transition rate between some initial state \( i \) and some final state \( f \) can be written within the dipole approximation as

\[
\Gamma = \frac{2\pi}{\hbar} |T_{i\rightarrow f}|^2 \delta(\epsilon_f - \epsilon_i),
\]

where \( \epsilon_f \) and \( \epsilon_i \) are energies associated with the final and initial states, respectively. In our calculations the initial and final states are represented by Kohn–Sham orbitals. The initial state is a 1s state obtained from a calculation on an isolated atom. The final states are the Kohn–Sham orbitals of the system in which a core hole is placed on an atom assumed to be absorbing. The excited electron is absent in the approximation we use. Thus in order to calculate the spectra the evaluation of transition matrix elements of the form

\[
T_{i\rightarrow f} = \langle \psi_i | \nabla | \psi_f \rangle,
\]

is needed where \( \psi_i \) indicates the 1s core state of an oxygen atom, and \( \psi_f \) denotes a final excited state. While the 1s core state is available from an atomic calculation, the final state \( \psi_f \) is not directly available from our electronic structure calculation, since we use pseudopotentials to represent atomic cores. When a pseudopotential is used the orbitals obtained are modified in the regions near atomic cores, precisely where the transition matrix in Eq. (2) is expected to be significant.

In order to circumvent this problem, we apply a reconstruction procedure around the absorbing atom. Our reconstruction procedure is based on the transformation due to Bloechl to be described below. The Bloecbl transformation is an exact transformation between an all-electron (AE) orbital and a corresponding pseudo (PS) orbital in terms of atomic AE and PS orbitals. For the full derivation see Ref. 26.

In the following we denote the AE (PS) orbitals of the extended system by \( |\Psi_n\rangle \) (\( |\tilde{\Psi}_n\rangle \)). The transformation that relates the AE and the PS orbitals is given by

\[
|\Psi_n\rangle = |\tilde{\Psi}_n\rangle + \sum_i (|\phi_i\rangle - |\tilde{\phi}_i\rangle) \langle \tilde{\phi}_i | \Psi_n \rangle R_i,
\]

where \( |\phi_i\rangle \) denote atomic AE partial waves obtained by integrating the radial Schrödinger equation, \( |\tilde{\phi}_i\rangle \) denote atomic PS partial waves chosen to reproduce the corresponding AE partial waves outside the core radius \( R_i \). The index \( i \) in Eq. (3) is a composite index denoting the atomic site \( R \), energy state \( n \), and angular momentum state \( l m \). The notation \( \langle \rangle | R \rangle \) stands for taking the scalar product within the core radius \( R \). The atomic PS partial waves are constructed to be nodeless and usually smoother than the atomic AE functions within the core radius. Furthermore, the atomic PS partial waves match continuously onto the atomic AE functions at the core radius.

The functions \( |\tilde{\phi}_i\rangle \) are projector functions to be determined. A condition on the projector functions \( \tilde{\phi}_i \) is arrived at by requiring that an atomic PS orbital transform into an atomic AE orbital. In general, this requirement is satisfied if

\[
\langle \tilde{\phi}_i | \phi_j \rangle_R = \delta_{ij}.
\]
III. SIMULATION DETAILS

We use first-principles molecular dynamics\(^9\) (FPMD) to optimize structures of molecules and ice, and to generate thermally distributed liquid water configurations. We use a pseudopotential (PP) plane-wave approach in our electronic structure calculation. Our calculation procedure consists of three parts: (i) generation of atomic configurations by FPMD; (ii) calculation of unoccupied orbitals in the presence of a half/full core hole; (iii) calculation of transition matrix elements by applying a reconstruction procedure described in the previous section. HCH/FCH are implemented by the generation of a PP for the absorbing oxygen atom with a half/full core electron missing. The missing half/full electron is replaced by an equally charged uniform background in our calculations using periodic boundary conditions. The 1s orbital of the oxygen (i.e., the initial state in the transition matrix element) is obtained from a calculation on the isolated oxygen atom.

Our simulations were performed using a FPMD code based on the Car–Parrinello method combined with Vanderbilt’s ultrasoft PPs.\(^{24,25}\) Throughout we used the PBE-GGA (Ref. 30) exchange-correlation functional. It has been demonstrated\(^{31,32}\) that in order to describe the structure of water a gradient-corrected exchange-correlation functional is needed in FPMD simulations. In particular the PBE-GGA has been shown to provide excellent agreement with experiment regarding structure.\(^{33}\) The plane-wave cutoff energy on the orbitals (density) was 25 (200) Ry.

For the calculation on ice\(^34\) we used a supercell including 96 molecules at a density of 0.961 g/cm\(^3\) with periodic boundary conditions. We generated a hexagonal lattice with proton disorder satisfying the Pauling ice rules and the requirement that the dipole moment of the cell is zero. In generating such a lattice, we followed the procedure of Kroes.\(^{35}\) The resulting atomistic structure was fully relaxed at zero temperature.

The water simulation was performed using a constant volume supercell including 64 molecules at a density of 1.00 g/cm\(^3\). The electronic mass was 700 amu. An equilibration run of 1.5 ps was performed at finite temperature, subsequently we obtained 4 ps of configurations from a microcanonical run, whose average temperature was 343 K. This temperature is considerably higher than the experimental one (see comparison in Fig. 1) due to an intrinsic limitation in the current approximations of density functionals. In an FPMD run at ambient temperatures, diffusion tends to be extremely slow. This limitation can be corrected by raising the temperature by 50 K (see also Refs. 36 and 37).

The spectrum was calculated by averaging over all 64 possible locations of the hole in one instantaneous configuration during the run. We checked that a different instantaneous configuration, well separated in time from the former (by 2 ps) gave essentially the same spectrum. The reasons why the spectra are insensitive to which instantaneous configuration is used is that all configurations in the microcanonical run are representative of the local H bonding environment around oxygens (see below), furthermore the features of the excited orbitals are strongly perturbed by the core hole around the absorbing atom.

Since extrinsic broadening is expected to dominate, the reported spectra are broadened using a Gaussian convolution of width 0.6 eV in the case of the ice and the liquid and 0.4 eV in the gas phase. In calculating the spectra for ice (water) we have used 300 (40) virtual orbitals relaxed in the presence of a half or full core hole. This is sufficient to describe the edge structure.

IV. RESULTS AND ANALYSIS

We have calculated the radial distribution functions and compared them to two recent high quality experiments by Hura et al.\(^{38}\) and Soper.\(^{39}\) In Fig. 1, we show the oxygen–oxygen radial distribution functions obtained from FPMD and the two experiments. The three curves are in excellent agreement, the positions and intensities of the calculated maxima and minima are close to the experimental ones. We attribute the small differences between theory and experiment to the simulation temperature which was higher than the experimental ones (ambient conditions), to the finite size of the simulation cell, to the neglect of quantum motion of the H nuclei, and to the intrinsic limitations of the generalized gradient approximation.\(^{36}\) Our calculated distribution function also compares well with the one obtained by Silvestrelli and Parrinello\(^{37,40}\) who used the gradient-corrected BLYP (Refs. 41 and 42) exchange-correlation functional.

The comparison of spectra between theory and experiment for the monomer is shown in Fig. 2. Calculations for the dimer are also presented (we are not aware of experimental results for the dimer). Our calculations give only the eigenvalues corresponding to the final virtual states, the initial-state eigenvalue (the oxygenic core state \(\varepsilon_{1s}\)) is not available. As a result absolute excitation energies are not available. Relative positions of peaks, however, correspond to the relative positions of our calculated final-state eigenvalues. Since in our calculations only relative excitation energies are known, throughout the paper we shift the excitation
energies (x axis) by an arbitrary constant for better visual comparison with experiment. For the monomer, we compare HCH and FCH. Essentially both methods reproduce the edge structure, while HCH is better in the high-energy region. It has to be kept in mind that it is difficult to obtain accurately the excitation energies over a wide spectral range by a method in which all excited states feel the same potential. In the case of the dimer we have also shifted the excitation energy axis of the spectra by the same amount, since we have made the assumption that the initial-state eigenvalue ($\epsilon_{1s}$) is equal for the two oxygen atoms. This is a reasonable assumption since the core states in water are not significantly perturbed by the valence electronic states. For the dimer, the spectrum of the acceptor molecule is not significantly perturbed compared to the monomer. In contrast, the donor molecule shows a first peak that is enhanced relative to the second one, and a more intense band at high energies. The difference between the acceptor and the donor indicates that the spectra depend strongly on the H-bonding environment experienced by the absorbing molecule.

The calculated and experimental spectra of ice and liquid water are shown in Figs. 3 and 4. The intensities of the experimental results between different phases are not comparable, therefore we can only compare intensity ratios of different bands between the experimental and theoretical spectra of a given phase. Moreover, as was mentioned above, the x axis of our calculated spectra were also shifted for better comparison with experiment, assuming as before that the $\epsilon_{1s}$ eigenvalues corresponding to different molecules are equal. FCH (HCH) results are presented in Fig. 3 (Fig. 4).

In order to assess the overall accuracy of our calculation we consider first the spectra in the ice phase. The calculated spectra for ice have been obtained using a single configuration in which the core hole is placed on a given molecule. For this calculation we used 300 virtual states. We separately checked that no significant changes occur when the core hole is placed on a different molecule (i.e., the effect of proton disorder is negligible). We then checked the effect of $k$-point sampling by calculating the spectrum at the Baldereschi point. As we can see in Fig. 3, the sharp rise of the edge seen in experiment is reproduced by FCH, which in addition gives well the three main features seen in experiment, i.e., pre-edge, shoulder, and broad peak, showing a minor dependence on $k$-point sampling, in particular in the edge region. This is not surprising for a molecular crystal modeled by a
TABLE I. Hydrogen bond criteria based on Ref. 44.

<table>
<thead>
<tr>
<th>Hydrogen Bond Criteria</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>O−···O distance</td>
<td>2.2–3.2 Å</td>
</tr>
<tr>
<td>H···O distance</td>
<td>1.2–2.2 Å</td>
</tr>
<tr>
<td>O−H···O angle</td>
<td>130–180°</td>
</tr>
</tbody>
</table>

relatively large periodic cell with 96 molecules. The most noticeable effect of better k-point sampling is to broaden the spectrum at higher energy, bringing it even closer to agreement with experiment. Brillouin-zone sampling should be even less important in liquid water because of the strong positional disorder. Therefore all the liquid water calculations have been done using the Γ point only. For the water calculation we used 40 (300) virtual states in the case of FCH (HCH), which is why the overall band in the case of FCH is narrow compared to our calculated ice spectrum, however, 40 states were enough to converge the edge structure. The main conclusions are the following: (i) a well defined pre-edge is present in both ice and water for both FCH and HCH, (ii) the growth in the pre-edge relative to the main edge in liquid compared to solid is well reproduced by both calculations. While the qualitative conclusions are similar for both FCH and HCH the FCH edge structure seems to be in better agreement with experiment. While the HCH pre-edge overall intensity seems to be in agreement it is more structured than the experimental result.

FCH calculations underestimate the relative intensity of the pre-edge to the main absorption edge in comparison to the experiment shown. This discrepancy is considerably less severe when comparing with the results of Ref. 6. The discrepancies are possibly due to the limitations of our theory, but there may also be limitations of the experimental techniques used in the cited studies. These possible limitations include the difficulty in estimating the degree to which x-ray-absorption techniques penetrate bulk water, but saturation effects measured by X-ray fluorescence could also be a contributing factor.

In order to rationalize the observed trends in the spectra, we study the environment of water molecules in terms of H bonding. In ice, each molecule is tetrahedrally coordinated, and is H bonded to all four neighboring molecules. Naturally, two of the four H bonds are donor bonds, and the other two are acceptor bonds. We denote a water molecule in such a H-bonding configuration (HBC) as 2A−2D. In water, due to thermal motion, a significant portion of H bonds are broken, leading to molecules in other HBC’s (2A−1D, 1A−2D, etc.). We have averaged the number of molecules in different HBC’s using the criteria in Ref. 44 (O−···O distance: 2.2–3.2 Å H···O distance: 1.2–2.2 Å O−H···O angle 130–180°, see Table I) and found that out of the 64 molecules 27±4, 12±2, 9±2, and 10±3 molecules are in HBC’s 2A−2D, 1A−2D, 2A−1D, 1A−1D, respectively. In the configuration for our spectral calculation the number of HBC’s was 24, 14, 11, and 10 in respective order. We also found significantly smaller fractions of other networks. Overall, there are 128 H bonds in a 64 molecule ice configuration with periodic boundary conditions. In water we found 104±4 H bonds on average, the configuration which was used to calculate the spectrum turned out to have 104.
ably more sensitive to the breaking of an acceptor bond and to local environment in general, since it is a smaller perturbation on the virtual states. While there may be experimentally detectable effects of the breaking of an acceptor bond, they are likely to be small in particular in comparison to the breaking of a donor bond.

V. CONCLUSION

In conclusion we have calculated absorption spectra using relaxed Kohn–Sham excitations with a full core hole (FCH) and half core hole (HCH) pseudopotential and have determined that better agreement in the condensed phases is obtained in the case of the former. We have shown by calculating average spectra for the most abundant HBC’s that the prominent features of the spectra depend mainly on the extent of donor bonding around the absorbing molecule. The breaking of donor bonds strengthens the “pre-edge” feature and decreases absorption in the high-energy range. The breaking of an acceptor bond seems to have very little effect on the spectrum, and according to our FCH calculations may only be noticeable if a donor bond is already broken. Our prediction of the fraction of broken H bonds in liquid water compared to ice (19%) is supported by the quantitative agreement of our calculated spectra with experiment, and by previous calculations.10

ACKNOWLEDGMENTS

We thank S. Myneni for sharing unpublished experimental results and for helpful discussions. We thank G. Hura and H. Stern for providing the experimental radial distribution functions. We thank E. R. Batista for providing the code to generate a disordered hexagonal ice lattice. We also thank Dr. N. Marzari, Dr. S. Scandolo, and Dr. M. Lazzeri for helpful discussions. We acknowledge support from an NSF ITR grant. The calculations were performed at the Materials Science Computing Center of the W. M. Keck Foundation.