Electronically Coarse-Grained Model for Water

A. Jones,[1,2] F. Cipcigan,[1] V. P. Sokhan,[2] J. Crain,[1,2] and G. J. Martyna[3]

[1]School of Physics and Astronomy, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JZ, United Kingdom
[2]National Physical Laboratory, Hampton Road, Teddington, Middlesex TW11 0LW, United Kingdom
[3]IBM T. J. Watson Research Center, Yorktown Heights, New York 10598, USA

(Received 12 December 2012; published 31 May 2013)

We introduce an electronically coarse-grained description of water representing all long range, many-body electronic responses via an embedded quantum oscillator. Leading-order response coefficients and gas phase electrostatic moments are exactly reproduced. Molecular dynamics, using electronic path integral sampling, shows that this framework is sufficient for a realistic liquid to emerge naturally with transferability extending further to nonambient state points and to the free water surface. The model allows the strength of many-body dispersion and polarization to be adjusted independently and these are found to have significant effects on the condensed phase.

DOI: 10.1103/PhysRevLett.110.227801 PACS numbers: 61.20.Ja, 31.15.xk, 61.25.Em

Since the pioneering work of Bernal and Fowler [1], there has been sustained interest in the molecular origin of water's anomalous physical properties [2]. Examples of these include an extremely large dipole enhancement in the condensed phase leading to a high dielectric constant, unusual pressure and temperature dependencies of the diffusion coefficient and the existence of a temperature of maximum density. At low temperatures, there exist two distinct densities of amorphous ice [3] and a proposed coexistence line between high- and low-density liquids terminating at an inaccessible (second) critical point [4].

Ultimately, these signature features of water arise from a unique combination of molecular properties such as water's high polarizability, directional hydrogen bonding sites and ubiquitous van der Waals (vdW) forces. Though underexplored and rarely emphasized for water, vdW interactions are believed to affect high-pressure polymorphism [5], wetting phenomena [6], surface melting [7], and stability of cluster geometries where many body effects are particularly relevant [8].

Developing a minimal model that captures these phenomena and exposes their molecular origins remains a grand challenge in atomistic simulation. There are many empirical water models that can be divided into fixed-charge and dipole-polarizable families (see, e.g., [9,10] and references therein for recent reviews). The former are not generally transferable to a wide variety of state points or heterogeneous aqueous systems (e.g., water interfaces). The latter, which include shell models and classical oscillators, offer improvements in transferability but neglect both many-body dispersion and higher-order polarizabilities which are particularly important for water due to the strongly nonuniform local Coulomb field.

First principles simulations of water based on the density functional theory (DFT) within the generalized gradient approximation [11] tend to produce an overstructured liquid with a significantly underestimated equilibrium density. The prevailing opinion is that the root of the problem is the failure of standard local functionals to capture the electronic correlations [12–15] which give rise to vdW forces. It now appears that the addition of empirical pairwise vdW dispersion corrections to exchange and correlation functionals are required to improve the equilibrium density [16]; empirical many-body dispersion corrections in the dipole limit have also been explored in this context [17].

Recently, a framework for the treatment of long-range responses for atoms and molecules has been developed [18–20]. The method is based on replacing the valence electrons of an atom by an embedded quantum Drude oscillator (QDO)—a charged quasiparticle tethered by a harmonic spring. This framework defines a coarse grained electronic structure capable of generating many-body polarization and dispersion to all orders (both in multipoles and particle number) [18–20]. Indeed, despite their inherent simplicity, QDO's have been shown to generate remarkably accurate descriptions of long-range response coefficients beyond the dipole approximation and the quantum mechanical nature of the QDO naturally provides screening at short range.

The present work is the first QDO description of a condensed phase system with a highly nontrivial phase diagram, water, for which both polarization and dispersion are important. Here, we do not set out to obtain a fully optimized parameterization through extensive fitting. We simply aim to input an accurate description of the long range electronic responses of the isolated monomer and observe the emergence of the liquid state physics to be understood in terms of molecular properties.

The QDO model of water is constructed as follows. First, the charge density of the water monomer is treated using a rigid 4-site representation: Three atoms positioned at the equilibrium geometry are supplemented by an "M site" along the $\angle$HOH bisector following the TIP4P empirical model. The charges and M-site position are fit to

TABLE I. The dipole moment μ, quadrupole moment θ, dipole-dipole polarizability α_1, quadrupole-quadrupole polarizability α_2, and dispersion coefficients C_6 and C_8. The corresponding values in condensed phase, are also shown, if determined. The enhancement of μ and θ_T can be seen graphically in Fig. 4. The molecular coordinate frame is defined with the molecule lying in the $x-y$ plane with x axis along the C_{2v} symmetry axis pointing from oxygen to hydrogens. Data in the column 'reference' are all from the Ref. [21]. From α_1, α_2, and C_6, the relations in the text yield QDO parameters $\omega = 0.6287E_h/\hbar, m = 0.3656m_e$, and $q = -1.1973e$.

Molecular property	Gas-phase reference	Gas-phase model	Condensed-phase model	Condensed-phase change
μ/D	1.855	1.855	2.61	+41%
θ_{xx}/DÅ	−0.13	−0.32	−0.64	
θ_{yy}	2.63	2.66	3.02	
θ_{zz}	−2.50	−2.34	−2.38	
$\theta_T \equiv (\theta_{yy} - \theta_{zz})/2$	2.51	2.50	2.70	+8%
α_{xx}/a_0^3	9.91	9.92	10.61	+7%
α_{yy}	10.31	9.92	9.82	−1%
α_{zz}	9.55	9.92	10.22	+3%
$\lvert\alpha_1\rvert$	9.92	9.92	10.22	+3%
$\lvert\alpha_2\rvert/a_0^5$	32.4	32.4	···	···
$C_6/E_h a_0^6$	46.4	46.4	···	···
$C_8/E_h a_0^8$	1140	1010	···	···
$C_9/E_h a_0^9$	115	103	···	···

reproduce exactly the monomer dipole moment and to give the best fit of the quadrupole moment components, θ_{ij}. Then, in order to capture the electronic responses, a quantum Drude particle of mass m and charge $-q$ is tethered to the M site with a spring constant $k = m\omega^2$, yielding a Gaussian charge density having width $\sigma^2 = \hbar/2m\omega$. The tether point carries a neutralizing charge $+q$, and is placed on the M site—a choice which minimizes the square error in the dipole-quadrupole polarizability tensor [21]. Since the polarizability of water is nearly isotropic (see Table I), a single QDO oscillator suffices. Third, short-range repulsion effects due to fermion exchange are absent from the model by construction. Therefore, empirical short-range pairwise repulsion terms between the O atoms, and short-range cutoffs to the electrostatic interactions, are added. These are not finely tuned here and will be explored fully in later work. The model parameters and form of the interaction potential are listed in the supplementary material [22].

Our baseline QDO model parameters, $\{m, \omega, q\}$ are chosen to reproduce the isotropic parts of the dipole, quadrupole polarizabilities α_1, α_2, and the C_6 dispersion coefficient (see Table I; note that some of these reference values are obtained from *ab initio* computation rather than measurement) by solving the relations $\alpha_1 = q^2/m\omega^2$, $\alpha_2 = (3/4)(\hbar/m\omega)\alpha_1$, $C_6 = (3/4)\hbar\omega\alpha_1^2$ [20]. The higher order C_8 and leading order many-body dispersion term C_9, though not explicitly fit are nonetheless in reasonable agreement with available reference values (see Table I).

To explore the effect of scaling the dispersion interaction while preserving all polarizabilities we introduce a scaling factor κ, which modifies the baseline parameters, namely, $\omega'(\kappa) = \omega\kappa$, $m'(\kappa) = m/\kappa$, and $q'(\kappa) = q\sqrt{\kappa}$. To scale the polarization at fixed dispersion we have a second scaling factor λ which modifies the parameters by $\omega'(\lambda) = \omega/\lambda^2$, $m'(\lambda) = m\lambda^2$, and $q'(\lambda) = q/\sqrt{\lambda}$. The case of $\kappa = \lambda = 1$ recovers the baseline model. We stress that κ has the effect of scaling *all many-body* dispersion effects linearly and λ scales *all* polarizabilities linearly.

We use the adiabatic path integral molecular dynamics of Refs. [18,23], which scales linearly in the number of particles, to study our electronically coarse-grained model at ambient conditions, $T = 300$ K, $\rho = 0.997$ g cm^{-3} for systems containing between $N = 72$ and 4000 water molecules. Partial radial distribution functions for the baseline and scaled models collected over 2 ns runs are shown in Fig. 1 against the TIP4P-pol2 model [24], known to be in excellent agreement with experimental diffraction data [25]. All three models produce a reasonable description of liquid water given that we have not optimized the short-range interactions.

We first discuss the baseline model which generates the well-known structural features of the liquid state: The O-O distribution has a broad feature near 4.4 Å, which is the hallmark of a tetrahedral hydrogen-bonded network. However, the peak is slightly amplified and as a result, the separation between first and second shells (minimum at 3.4 Å) is emphasized. The enthalpy of vaporization at ambient pressure (extrapolated to infinite system size), is $\Delta h_{vap} = 40 \pm 2$ kJ/mol, close to the experimental value of 43.91 kJ/mol. Overall, the baseline model produces a realistic but mildly overstructured liquid, with a pressure of 6.6(2) kbar at experimental density. Ambient pressure is recovered when the density is reduced by 15%. Similar observations have been made based on DFT simulations [5], where an overstructured liquid is produced, with a density that is too low by up to 20%. It is believed that neglect of dispersion is the primary cause of the discrepancy in that calculation. In the present model, which includes dispersion, short-range effects will also be important (currently O-H repulsion is neglected [21] but would tend to weaken the H-bonding structure—this is being explored currently). Since this Letter was accepted for publication, using *ab initio* computations at the level of coupled cluster CCSD(T) on the dimer as input, we have fit an improved water model which gives a good distribution function and ambient density without enhanced dispersion or depleted polarization, making only slight adjustments to the parameters of the repulsion potential. On the other hand, dispersion coefficients have not been measured experimentally for water, and there is some variation in reported values.

When dispersion is scaled according to the prescription given previously, the equilibrium density at ambient

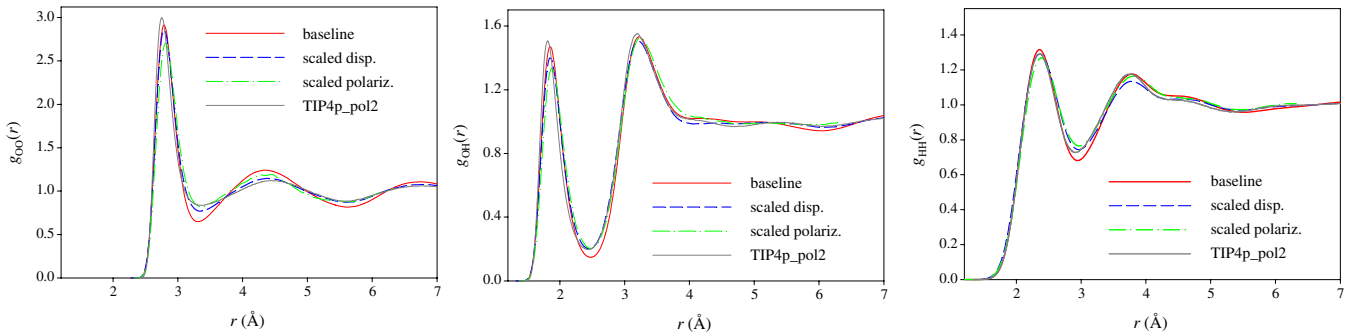

FIG. 1 (color online). Radial distribution functions (RDFs) for oxygen-oxygen (left), oxygen-hydrogen (middle), and hydrogen-hydrogen (right), obtained using the baseline QDO model for water (solid lines), with scaled dispersion (dashed lines), and scaled polarizability (dash-dotted lines). For comparison, RDFs obtained with TIP4P-pol2 model [24] (gray lines) are overlapped.

pressure increases with increasing dispersion at fixed polarization reaching its experimental value at $\kappa = 1.37$. As can be seen in Fig. 1, the increased dispersion leads to an overall weakening of the H-bond network bringing the correlations into better agreement with the benchmark data at the equilibrium density. The hydrogen-bonding structure is more fully explored in Fig. 2, exposing its similarity to the TIP4P-pol2 structure. However, the enthalpy of vaporization increases to $\Delta h_{\mathrm{vap}} = 46 \pm 2$ kJ/mol, which is 4% higher than the experimental value. Thus the enhanced dispersion at fixed polarization model has yielded intriguing results.

We find that the H-bond structure is also sensitive to changing polarization at fixed dispersion. Figures 1 and 2 show that scaling down the water polarizability to $\lambda = 0.9$, also weakens the H-bond network relative to the baseline model. The effect is therefore similar to that found for increasing dispersion though pressure and enthalpy of vaporization are less affected (4.2 kbar and 41.2 kJ/mol, respectively).

The three-dimensional view of the structure presented in Fig. 3 reveals the strong similarity in orientational tetrahedral correlations between QDO and both TIP4P and TIP4P-pol2 models. The isosurfaces for generalised gradient approximation-DFT based *ab initio* and several empirical models can be found elsewhere [26].

As the QDO model produces a coarse-grained electronic structure, we can study many-body effects on the induced moments in the liquid phase. The mean dipole moment in the liquid phase is 2.6 D for the baseline model, 2.5 D in the case of scaled dispersion at $\kappa = 1.37$ and 2.5 D for the scaled polarization at $\lambda = 0.9$. Since scaling the dispersion by κ does not change the polarizability, the difference in enhancement in that case must be due to disruption of the H-bond structure. This enhancement relative to the gas phase, by $\approx 40\%$, is close to proposed values for water [24] but is lower than the result of *ab initio* molecular dynamics, which gave ≈ 3.0 D [27], and than that predicted by a point-multipole model of polarizability [28].

We now contrast our Gaussian electronic structure model with the results of point polarizable models and the similar classical charge on spring (or classical Drude oscillator) models. It is known that point-polarizable models of water yield an exaggerated enhancement because the

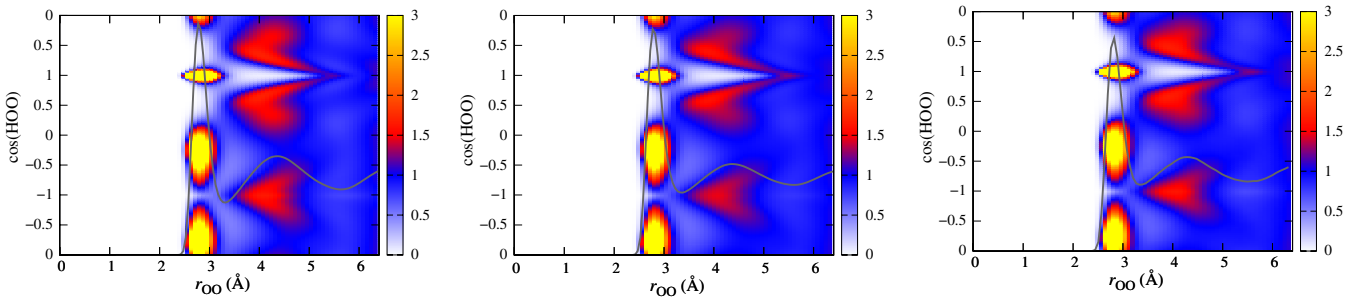

FIG. 2 (color online). Density maps of the H-bond angle H-O-O vs O-O distance distribution, obtained for the baseline QDO model (left), with scaled dispersion (middle), and scaled polarizability (right). The O-O RDF from Fig. 1 have been overlaid on each. Perfect tetrahedral structure (ice) would give peaks at cos(HOO) $= +1$ and $-1/3$ in the first shell, and at -1 and $+1/3$ in the second shell. Significant density is observed at each of these points. The liquid has additional density in the line (3.5,0.5)-(5.5,1), which implies water molecules have considerable flexibility to slide behind the H-bonded molecule in the first shell, unless they fall into the ''locked'' tetrahedral position at about (4.4, + 1/3) which is slightly further out and which results in the observed lower density of ice. The baseline model is more strongly structured at this point, confirming that the differences between the models shown in Fig. 1 are due to enhanced ice-like character.

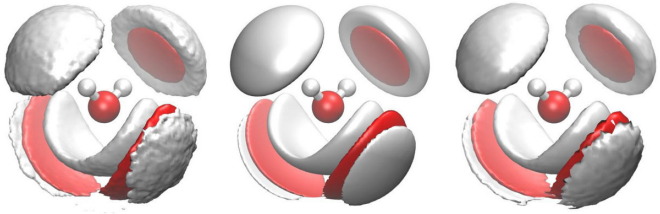

FIG. 3 (color online). Nearest neighbor distributions of oxygen (red) and hydrogen (gray), for the baseline QDO model (left), TIP4P (middle), and compared with published data [26] for TIP4P-pol2 (right), showing significant agreement between all three models at 300 K. The isosurfaces are drawn for oxygen and hydrogen densities at $\times 3.5$ and $\times 1.75$ the bulk density, respectively.

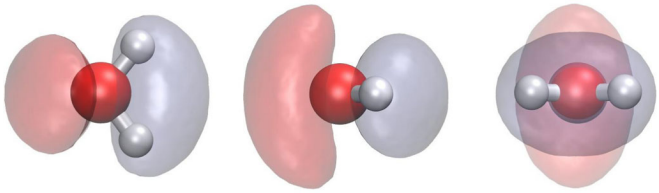

FIG. 4 (color online). The detailed average electronic polarization for the baseline QDO model water in the bulk (that is, far from surfaces and other inhomogeneities) at 300 K, relative to an unperturbed gas-phase or isolated monomer, represented by losses (blue) and gains (pink) of density, seen along the z axis (left), y axis (middle), and x axis (right). The observed pattern reflects the tetrahedral structure present in liquid water; electronic density is lost (blue) from around outgoing hydrogen bonds and gained (pink) around incoming hydrogen bonds, yielding a significant enhancement of the molecular dipole by 40% along the x axis, but also enhancement of the molecular quadrupole (see Table I). This distribution was generated by averaging over many molecular configurations in the liquid, sampling Drude density by projecting bead positions onto the molecular frame, binning into a three-dimensional array, and then subtracting the exactly known unperturbed density distribution of the isolated monomer, $\rho_0(x) = \sqrt{m\omega/\pi\hbar}\exp(-(m\omega/\hbar)x^2)$.

electrostatic field experienced by a polarizable point dipole in the center of the molecule is considerably larger than the same field averaged over the whole electronic distribution of the molecule [29]. The optimal Gaussian width for this averaging was determined empirically to be ≈ 0.7 Å [29]. The present model is in good agreement with this conclusion, having a Gaussian width of $\sigma = 0.78$ Å which arises in the QDO from the input leading order responses. The models SWM4-DP and -NDP [30], which are fit empirically to equation-of-state data, compensate for the exaggerated enhancement by fitting a smaller dipole polarizability than the gas-phase monomer [29]. Therefore, gas phase polarizability is inaccurate for these models. In contrast, for the QDO model, polarizability and dipole moment are accurate in both gas and liquid.

The QDO framework is also unique among empirical models (including all shell and classical oscillator models) in that it retains a highly nontrivial electronic structure which can be directly obtained from the Drude distribution. Polarization occurs by a physical redistribution of the coarse-grained electronic density. This is shown in Fig. 4 where regions of enhanced and depleted density (leading to altered condensed phase multipoles including the quadrupole moment, see Table I) are visible and reflect the tetrahedral environment of the bulk liquid.

Collective properties, such as the static dielectric permittivity, provide a stringent test of the accuracy of the structural and electronic properties. Using the expression for polarizable models [31], we obtain the static dielectric constant for trajectories acquired over ≈ 2 ns, the baseline model ($\kappa = \lambda = 1$) has a dielectric constant of $\epsilon = 60 \pm 3$, which is comparable to the value for the nonpolarizable TIP4P model but lower than the experimental value of 78. However, when dispersion is scaled to $\kappa = 1.37$, the value at which the experimental density occurs at zero pressure, we find the permittivity increases to $\epsilon = 79 \pm 2$ in excellent agreement with experiment. The origin of the enhancement of the dielectric constant is subtle and not intuitive since the dipole moment for the scaled dispersion case is marginally lower than that of the baseline parameters. However, the scaled system is also more weakly structured and may therefore exhibit larger dipolar fluctuations. This may also affect convergence behavior which we have not thoroughly explored.

As the QDO framework is based upon an accurate description of the isolated molecule, it should be transferable to interfaces and other heterogeneous environments where nonpolarizable models are questionable and the effects of dispersion are unexplored. As a first example, we consider here the free water surface where the H-bond network is abruptly truncated. In particular, we compute the surface tension γ in slab geometry using $\gamma = L_z(p_n - p_t)/2$, where p_n and p_t are the normal and tangential components of the macroscopic pressure tensor. For the scaled dispersion case we obtain $\gamma = 72.6 \pm 1$ mN/m, in excellent agreement with the experimental estimate of 71.73 mN/m. We also make an initial exploration of the liquid properties at nonambient state points along the liquid-gas coexistence curve to further examine transferability. In particular, we perform NPT ensemble simulations at temperatures $T = 300$ and 400 K. The system densities for $\kappa = 1.37$ are found to be 0.988 and 0.928 g cm^{-3} both within 1% of experimental values.

The principal result reported here is that an accurate description of the long-range electrostatic responses in the isolated water molecule (through which many-body dispersion and polarization effects are accounted for, to all orders, in a simple coarse-grained framework) produces a realistic liquid which under ambient conditions exhibits a well-developed hydrogen bond network and reasonable values for the heat of vaporization, enhanced dipole, and static permittivity. Transferability appears to extend further to the liquid-vapor interface where the surface tension is in

excellent agreement with experiment and to state points along the liquid-gas coexistence curve.

It has for some time been argued that the structure and properties of water emerge from an effective competition between directional hydrogen bonding favoring open, low density configurations, and more isotropic dispersion interactions which favor compact, higher density molecular arrangements [32,33]. This behavior emerges naturally in the QDO framework. The tetrahedral structure shows particular sensitivity to small changes in polarizability and the dipole enhancement is sensitive to changes in structure.

We emphasize that this is a first generation QDO model for water (indeed, for any molecular liquid) and that there is considerable scope for (mild) tuning of input molecular responses within physically reasonable limits, adjustment of repulsive and regularization terms and of alternative parameterization strategies. In the context of the latter, it would be of interest, for example, to determine what molecular properties are required to produce the more disordered liquid state inferred from neutron diffraction measurements [34] (that is, compared to the TIP4P-pol2 standard which we use here). The present work suggests that a full quantum treatment, albeit with Gaussian fluctuations, resolves many issues inherent in classical-limit polarizable models. The implication is that the framework introduced here may be an attractive foundation for a highly transferable model for water and other molecular liquids. Last, the technique is linear scale and easily implemented for parallel supercomputers allowing large complex systems to be examined.

This work was supported in part by the National Physical Laboratory Strategic Research Programme and IBM Research. We thank Yves A. Mantz for providing us with the TIP4P-pol2 data sets. Generous allocation of CPU time on BlueGene/Q at STFC Hartree Centre in Daresbury, UK, is gratefully acknowledged.

[1] J. D. Bernal and R. H. Fowler, J. Chem. Phys. **1**, 515 (1933).
[2] *Water: A Comprehensive Treatise*, edited by F. Franks (Plenum, New York, 1972), Vols. 1–7.
[3] O. Mishima, L. D. Calvert, and E. Whalley, Nature (London) **314**, 76 (1985).
[4] P. H. Poole, F. Sciortino, U. Essmann, and H. E. Stanley, Nature (London) **360**, 324 (1992).
[5] B. Santra, J. Klimeš, D. Alfè, A. Tkatchenko, B. Slater, A. Michaelides, R. Car, and M. Scheffler, Phys. Rev. Lett. **107**, 185701 (2011).
[6] I. Hamada, K. Lee, and Y. Morikawa, Phys. Rev. B **81**, 115452 (2010).
[7] J. Carrasco, B. Santra, J. Klimeš, and A. Michaelides, Phys. Rev. Lett. **106**, 026101 (2011).
[8] N. Goldman and R. Saykally, J. Chem. Phys. **120**, 4777 (2004).
[9] C. Vega and J. L. F. Abascal, Phys. Chem. Chem. Phys. **13**, 19 663 (2011).
[10] P. T. Kiss and A. Baranyai, J. Chem. Phys. **137**, 084506 (2012).
[11] K. Laasonen, M. Sprik, M. Parrinello, and R. Car, J. Chem. Phys. **99**, 9080 (1993).
[12] W. Kohn, Y. Meir, and D. E. Makarov, Phys. Rev. Lett. **80**, 4153 (1998).
[13] R. W. Williams and D. Malhotra, Chem. Phys. **327**, 54 (2006).
[14] M. J. Gillan, F. R. Manby, M. D. Towler, and D. Alfè, J. Chem. Phys. **136**, 244105 (2012).
[15] B. Santra, J. Klimeš, D. Alfè, A. Tkatchenko, B. Slater, A. Michaelides, R. Car, and M. Scheffler, Phys. Rev. Lett. **107**, 185701 (2011).
[16] J. Schmidt, J. VandeVondele, I.-F. W. Kuo, D. Sebastiani, J. I. Siepmann, J. Hutter, and C. J. Mundy, J. Phys. Chem. B **113**, 11 959 (2009).
[17] M. W. Cole, D. Velegol, H.-Y. Kim, and A. A. Lucas, Mol. Simul. **35**, 849 (2009).
[18] T. W. Whitfield and G. J. Martyna, Chem. Phys. Lett. **424**, 409 (2006).
[19] A. Jones, A. Thompson, J. Crain, M. H. Müser, and G. J. Martyna, Phys. Rev. B **79**, 144119 (2009).
[20] A. P. Jones, J. Crain, V. P. Sokhan, T. W. Whitfield, and G. J. Martyna, Phys. Rev. B **87**, 144103 (2013).
[21] C. Millot, J.-C. Soetens, M. Costa, M. Hodges, and A. Stone, J. Phys. Chem. A **102**, 754 (1998).
[22] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.110.227801 for the model parameters and form of the interaction potential.
[23] T. W. Whitfield and G. J. Martyna, J. Chem. Phys. **126**, 074104 (2007).
[24] B. Chen, J. Xing, and J. I. Siepmann, J. Phys. Chem. B **104**, 2391 (2000).
[25] G. L. Hura, D. Russo, R. M. Glaeser, T. Head-Gordon, M. Krack, and M. Parrinello, Phys. Chem. Chem. Phys. **5**, 1981 (2003).
[26] Y. A. Mantz, B. Chen, and G. J. Martyna, J. Phys. Chem. B **110**, 3540 (2006).
[27] P. L. Silvestrelli and M. Parrinello, Phys. Rev. Lett. **82**, 3308 (1999).
[28] E. R. Batista, S. S. Xantheas, and H. Jónsson, J. Chem. Phys. **109**, 4546 (1998).
[29] B. Schropp and P. Tavan, J. Phys. Chem. B **112**, 6233 (2008).
[30] G. Lamoureux, J. Alexander, D. MacKerell, and B. Roux, J. Chem. Phys. **119**, 5185 (2003).
[31] M. Neumann and O. Steinhauser, Chem. Phys. Lett. **106**, 563 (1984).
[32] I.-C. Lin, A. P. Seitsonen, M. D. Coutinho-Neto, I. Tavernelli, and U. Rothlisberger, J. Phys. Chem. B **113**, 1127 (2009).
[33] A. Møgelhøj, A. K. Kelkkanen, K. T. Wikfeldt, J. Schiøtz, J. J. Mortensen, L. G. M. Pettersson, B. I. Lundqvist, K. W. Jacobsen, A. Nilsson, and J. K. Nørskov, J. Phys. Chem. B **115**, 14 149 (2011).
[34] A. Soper, J. Phys. Condens. Matter **19**, 335206 (2007).