



PCCP

## Raman Spectrum and Polarizability of Liquid Water from Deep Neural Networks

Journal:	<i>Physical Chemistry Chemical Physics</i>
Manuscript ID	CP-ART-04-2020-001893.R1
Article Type:	Paper
Date Submitted by the Author:	28-Apr-2020
Complete List of Authors:	Sommers, Grace; Princeton University, Physics Calegari, Marcos; Princeton University, Chemistry Zhang, Linfeng; Princeton University, Program in Applied and Computational Mathematics Wang, Han; Institute of Applied Physics and Computational Mathematics, Laboratory of Computational Physics Car, Roberto; Princeton University, Chemistry; Princeton University, Physics; Princeton University, Program in Applied and Computational Mathematics

SCHOLARONE™  
Manuscripts

Cite this: DOI: 00.0000/xxxxxxxxxx

# Raman Spectrum and Polarizability of Liquid Water from Deep Neural Networks

Grace M. Sommers,<sup>a\*</sup> Marcos F. Calegari Andrade,<sup>b\*</sup> Linfeng Zhang<sup>c</sup>, Han Wang<sup>d</sup>, and Roberto Car<sup>a,b,c†</sup>

Received Date

Accepted Date

DOI: 00.0000/xxxxxxxxxx

We introduce a scheme based on machine learning and deep neural networks to model the environmental dependence of the electronic polarizability in insulating materials. Application to liquid water shows that training the network with a relatively small number of molecular configurations is sufficient to predict the polarizability of arbitrary liquid configurations in close agreement with *ab initio* density functional theory calculations. In combination with a neural network representation of the interatomic potential energy surface, the scheme allows us to calculate the Raman spectra along 2-nanosecond classical trajectories at different temperatures for H<sub>2</sub>O and D<sub>2</sub>O. The vast gains in efficiency provided by the machine learning approach enable longer trajectories and larger system sizes relative to *ab initio* methods, reducing the statistical error and improving the resolution of the low-frequency Raman spectra. Decomposing the spectra into intramolecular and intermolecular contributions elucidates the mechanisms behind the temperature dependence of the low-frequency and stretch modes.

## 1 Introduction

Raman scattering has been widely used to study rotational and vibrational spectra of gases and condensed-phase systems<sup>1</sup>. The Raman effect arises from the inelastic scattering of visible light with matter, in which incident radiation is shifted in the frequency domain due to vibrations and rotations of the scatterer<sup>2</sup>. Raman spectroscopy thus probes the same frequency region as infrared (IR) spectroscopy, but different selection rules apply for each technique, making Raman and IR complementary tools to investigate the rotational and vibrational signatures of condensed-phase systems.

The Raman spectrum of a system can be modeled from the polarizability time-correlation function obtained from molecular simulations. The sensitivity of the polarizability to the environmental dependence of the electronic structure demands consistent quantum mechanical approaches to model the potential and the polarizability surfaces. Traditional numerical simulations employ *ab initio* molecular dynamics (AIMD), in which electronic interactions are computed on the fly using density functional theory

(DFT) and the system polarization described through the modern theory of polarization<sup>3,4</sup>. This approach has been used to analyze Raman spectra of liquid water<sup>5</sup>, with polarizabilities evaluated at every AIMD step with Density Functional Perturbation Theory (DFPT)<sup>6</sup> or by calculating the displacement of the Wannier centers under a finite applied electric field<sup>7</sup>. However, while AIMD methods with the appropriate functional bring much-needed predictive accuracy, their computational intensity forbids their application on large systems size and time scales<sup>8</sup>. Empirical potentials provide a less expensive alternative, but are less robust and generalizable. One concern is that the optimal parameters for the potential energy surface do not reproduce the correct polarizability surface<sup>9</sup>; thus, successful polarizable models must take care to parameterize the polarizability surface. One such potential, POLI2VS, closely matches the observed IR spectrum, but the low-frequency and librational modes of the Raman spectrum are inconsistent with experiment<sup>10</sup>. Moreover, non-reactive potentials are unable to model systems with mutating chemical environments, such as acidic or alkaline aqueous solutions in which proton transfer occurs on a picosecond timescale<sup>11</sup>. An alternative to empirical potentials is MB-pol, a many-body potential including up to 3-body terms plus induction, parameterized with high-level quantum mechanical calculations on small molecular clusters. This is quite accurate for liquid water and more promising than empirical potentials, as it models the potential and polarization/polarizability surfaces within a self-consistent framework, but, being limited to molecular systems, this method cannot treat

<sup>a</sup> Department of Physics, Princeton University, Princeton, NJ 08544, USA.

<sup>b</sup> Department of Chemistry, Princeton University, Princeton, NJ 08544, USA.

<sup>c</sup> Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544, USA.

<sup>d</sup> Laboratory of Computational Physics, Institute of Applied Physics and Computational Mathematics, Huayuan Road 6, Beijing 100088, P. R. China.

† Email: rcar@princeton.edu

\* Equal contribution from both authors

dissociation<sup>12</sup>.

A second consideration in modeling the Raman spectrum is the treatment of nuclear quantum effects (NQEs). Recent studies have approached the quantum TCF in several ways: approximate path integral methods such as centroid molecular dynamics (CMD) and (thermostatted) ring polymer molecular dynamics (TRPMD)<sup>13,14</sup>; the local monomer approximation and other mixed quantum-classical methods, which treat exactly a small subset of vibrational modes<sup>15</sup>; and the linearized semiclassical initial value representation (LSC-IVR)<sup>16</sup>. However, each of these approximations has its drawbacks<sup>17,18</sup>, and it is difficult to distinguish effects due to approximate quantum dynamics from those due to the choice for the potential energy surface (PES). Our goal in this paper is to demonstrate the extent to which a *classical-nuclei* approach using a consistent neural-network-based PES and polarizability surface successfully reproduces experimental results.

Recently, machine learning methods have been used to express *ab initio* potential energy surfaces as a function of nuclear coordinates<sup>19–27</sup>. These methods preserve the accuracy of AIMD while improving on its efficiency. One implementation is the Deep Potential Molecular Dynamics (DPMD), a generalizable, accurate, and linearly scalable deep neural network (DNN)-based framework that can generate long trajectories in agreement with AIMD. Some of the authors have also recently used DNNs to learn the electric polarization as a sum of local atomic contributions<sup>28</sup>. This method allows the construction of *ab initio*-level polarization surfaces, which were used to compute the IR spectra of liquid and crystalline water under different pressure conditions. Alternative machine learning models based on kernel instead of DNN representations have been recently reported in the literature to predict tensorial properties extracted from *ab initio* calculations<sup>29–31</sup>. These models have been used to describe the polarizability and the Raman spectra of isolated molecules and dispersion bound molecular crystals<sup>31</sup>.

The strategy we adopt here is to use models based on DNNs to represent the quantities required by Raman spectra calculations: the interatomic potential energy and force, as well as the polarizability tensor. Upon training with *ab initio* data, DNN-based simulations reproduce AIMD results at orders of magnitude lower computational cost, enabling accurate simulations of large-scale systems at timescales prohibitively long for AIMD. We use our DNN-predicted interatomic potential and polarizabilities to compute the Raman spectrum of liquid water at different temperatures.

Due to the wide relevance of liquid water for science and technology, experimental studies over the past several decades have investigated Raman spectra of water at different thermodynamic conditions<sup>32–36</sup>. Although most experiments agree on the position of primary and secondary peaks, their interpretation is still not free of controversy. For instance, it is well understood that a decrease in water temperature redshifts and broadens the primary OH stretch peak, a fact attributed to larger stability of hydrogen bonds at lower temperatures. But interpretations diverge on whether and how specific features can be assigned to unique hydrogen bond configurations. Therefore, theoretical modeling

can point the way forward in identifying the physical processes at play.

## 2 Methods

In this section, we first introduce the theory of Raman spectra, relating the Raman line shape to the autocorrelator of the system polarizability. We then discuss the DNN representation of the Wannier centroid polarizabilities, before providing detail on the DFT calculations used to train the network.

### 2.1 Raman Line Shape from *Ab initio* Molecular Dynamics

The differential cross section of Raman scattering can be written in terms of the Fourier transform of the time autocorrelation function of the electronic polarizability of the system according to<sup>2</sup>:

$$\frac{d^2\sigma}{d\omega d\Omega}(\omega) = \left(\frac{2\pi}{\lambda_s}\right)^4 \frac{1}{2\pi} \int_{-\infty}^{\infty} dt e^{-i\omega t} \langle \hat{\epsilon}_s \cdot \boldsymbol{\alpha}(0) \cdot \hat{\epsilon}_i \hat{\epsilon}_s \cdot \boldsymbol{\alpha}(t) \cdot \hat{\epsilon}_i \rangle \quad (1)$$

Here  $\omega$  is the Raman frequency shift,  $\boldsymbol{\alpha}$  is the polarizability tensor of the sample,  $\hat{\epsilon}_s$  and  $\hat{\epsilon}_i$  are, respectively, the polarization directions of the scattered and incident light,  $\lambda_s$  is the wavelength of the scattered light. The angular brackets denote ensemble average, and the integration is over the time  $t$ . Factoring out the dependence on  $\lambda_s$ , experiments typically report reduced line shapes  $R(\omega)$  containing an arbitrary constant factor<sup>37</sup>:

$$R(\omega) = n_{BE}(\omega) I(\omega) \propto n_{BE}(\omega) \frac{1}{2\pi} \int_{-\infty}^{\infty} dt e^{-i\omega t} \langle \hat{\epsilon}_s \cdot \boldsymbol{\alpha}(0) \cdot \hat{\epsilon}_i \hat{\epsilon}_s \cdot \boldsymbol{\alpha}(t) \cdot \hat{\epsilon}_i \rangle \quad (2)$$

Here the Bose Einstein (BE) factor  $n_{BE}(\omega) = 1 - \exp(-\beta\hbar\omega)$  is introduced when studying the low-frequency features of the spectrum, which would otherwise be obscured by the Rayleigh line<sup>38</sup>. For fluid systems, it is convenient to decompose the polarizability tensor into a spherical part  $\bar{\alpha} = \frac{1}{3}\text{Tr}\boldsymbol{\alpha}$  and a traceless anisotropic tensor  $\boldsymbol{\beta} = \boldsymbol{\alpha} - \mathbb{1}\bar{\alpha}$ , yielding the isotropic and anisotropic components of the line shape:

$$R_{iso}(\omega) \propto n_{BE}(\omega) \int_{-\infty}^{\infty} dt e^{-i\omega t} \langle \bar{\alpha}(0) \bar{\alpha}(t) \rangle$$

$$R_{aniso}(\omega) \propto n_{BE}(\omega) \int_{-\infty}^{\infty} dt e^{-i\omega t} \frac{2}{15} \text{Tr}(\boldsymbol{\beta}(0) \boldsymbol{\beta}(t)) \quad (3)$$

In the above formulae, the polarizability depends on the nuclear coordinates, which are quantum mechanical operators but are treated here classically to compute the equilibrium time correlation functions via a molecular dynamics simulation. This amounts to neglecting nuclear quantum effects (NQEs) in the dynamics of the nuclei. In liquid water, NQEs are small but not negligible. Their influence on static equilibrium properties has been quantified in experiments and simulations based on Feynman path integrals<sup>39</sup>, but it is difficult to predict the effect on dynamic properties using statistical simulation methods. Leaving this issue aside, we approximate the environmental dependence of the polarizability with classical mechanics.

Nevertheless, the electronic polarizability itself can only be derived from quantum mechanics. Within the Born-Oppenheimer approximation, the electronic polarizability at time  $t$  measures the response of the instantaneous polarization ( $\boldsymbol{\mu}$ ) of the sample

at time  $t$  to an infinitesimally small uniform electric field while the nuclear positions are held fixed:

$$\boldsymbol{\alpha}(t) = \frac{\delta \boldsymbol{\mu}(t)}{\delta \mathbf{E}} \quad (4)$$

We adopt *ab initio* density functional theory (DFT) to describe the electronic ground state of the system and use the modern theory of polarization to compute  $\boldsymbol{\mu}$ . The derivative in Eq. 4 can be expressed analytically with density functional perturbation theory (DFPT)<sup>40</sup>, requiring the solution of the self-consistent response equations for the electrons, or it can be calculated numerically within the electric enthalpy framework<sup>41,42</sup> by applying small but finite electric fields  $\pm \boldsymbol{\varepsilon}$  to the sample:

$$\boldsymbol{\alpha} \approx \frac{\boldsymbol{\mu}(\boldsymbol{\varepsilon}) - \boldsymbol{\mu}(-\boldsymbol{\varepsilon})}{2\boldsymbol{\varepsilon}} \quad (5)$$

In practice, the two formulations are equivalent. We adopt here the one based on Eq. 5, which does not require a specialized DFPT code but only a DFT minimization code. The macroscopic polarization  $\boldsymbol{\mu}$  of a bulk periodic system is conveniently expressed, modulo a quantum, in terms of the position vectors of the nuclei ( $\mathbf{r}_i$ ) and the maximally localized Wannier centers ( $\mathbf{w}_k$ ),  $\boldsymbol{\mu} = e \sum_i Z_i \mathbf{r}_i - 2e \sum_k \mathbf{w}_k$ , where  $e$  is the unit electronic charge,  $Z_i$  are atomic numbers, and we have assumed a spin-saturated system<sup>43</sup>. The Wannier centers are obtained from a unitary transformation that minimizes the spatial spread in the occupied orbital subspace<sup>44</sup>. We use a valence-only pseudopotential approach so that the nuclear charges  $eZ_i$  are the charges of the ions consisting of the nuclei and the frozen core electrons, and the Wannier centers correspond to the valence electrons. Specializing to water, which contains oxygen ( $\mathbf{r}_{O_i}$ ) and hydrogen ( $\mathbf{r}_{H_m}$ ) ions, the polarization vector is:

$$\boldsymbol{\mu} = 6e \sum_i \mathbf{r}_{O_i} + e \sum_m \mathbf{r}_{H_m} - 2e \sum_k \mathbf{w}_k \quad (6)$$

In water, four Wannier centers can be uniquely associated to their nearest oxygen ion. These four Wannier centers remain close to the same oxygen during dynamical evolution, even when the water molecule to which the oxygen belongs dissociates leading to formation of hydronium and hydroxyl complexes. It is convenient to define a Wannier centroid by the average position of the four Wannier centers associated to oxygen  $i$ :

$$\mathbf{w}_i = \frac{1}{4} \sum_{l_i=1}^4 \mathbf{w}_{l_i} \quad (7)$$

Then, the electronic polarizability of the liquid water sample is just the sum of the centroid polarizabilities  $\boldsymbol{\alpha}_i$ :

$$\boldsymbol{\alpha}(t) = -8e \frac{\partial}{\partial \mathbf{E}} [\sum_i \mathbf{w}_i(t)] = \sum_i \boldsymbol{\alpha}_i(t) \quad (8)$$

In absence of molecular dissociation, the centroid polarizabilities can be viewed as effective molecular polarizabilities. These are useful for interpreting the spectrum, but only their sum is a physical observable. Its time correlation function yields the Raman line shape through Eq. 3. The polarizability  $\boldsymbol{\alpha}(t)$  is accessible on

the fly in AIMD trajectories, which thus provide a way to compute the Raman spectra using the same DFT approximation for the spectral calculations and for modeling the potential interactions that generate the atomic trajectories. Computational cost restricts these calculations to relatively short trajectories ( $\approx 100$  ps) and small simulation boxes ( $\approx 100$  molecules). The ensuing statistical errors limit the accuracy of Raman spectra calculations for water, particularly at low frequency where the Raman signal is very weak.

Recent progress with machine learning (ML) techniques applied to molecular simulation greatly alleviates these difficulties, as ML interatomic potentials can reproduce accurately AIMD trajectories at a cost that is several orders of magnitude lower than AIMD and scales linearly with system size. Our group has developed a versatile deep neural network (DNN) representation of the potential energy surface of multi-atomic systems called deep potential (DP) that has been used in several applications<sup>19,27</sup>. Recently, the DP representation was generalized to describe the environmental dependence of the polarization  $\boldsymbol{\mu}$ <sup>28</sup>. In the next section we discuss how this approach can be extended to the electronic polarizability  $\boldsymbol{\alpha}$ .

## 2.2 Deep Neural Network for the Environmental Dependence of the Polarizability Tensor

Let  $\mathcal{X}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_i, \dots, \mathbf{r}_N)$  be an extensive physical property, such as the potential energy  $U$ , the polarization  $\boldsymbol{\mu}$ , or the polarizability  $\boldsymbol{\alpha}$ , which depends on the atomic positions. We consider systems for which  $\mathcal{X}$  can be decomposed into a sum of local components  $\mathcal{X}_i$  that depend on the coordinates of all atoms inside a finite neighborhood  $\mathcal{N}_i$  of the atom at  $\mathbf{r}_i$ , i.e.:

$$\mathcal{X} = \sum_i \mathcal{X}_i = \sum_i \mathcal{X}_i(\{\mathbf{r}_j \in \mathcal{N}_i\}), \quad \mathcal{N}_i = \{j, r_{ij} < r_c\}, \quad (9)$$

where  $r_{ij} = \|\mathbf{r}_i - \mathbf{r}_j\|$  is the distance between  $i$  and  $j$ , and  $r_c$  is a predefined cut-off distance. There is no restriction, besides finite range, on the functional form of the environmental dependence in Eq. 9. In a condensed phase, only the global property  $\mathcal{X}$  is a meaningful observable while the local quantities  $\mathcal{X}_i$  are effective properties that depend on the adopted decomposition. As discussed in the previous section, in water the electronic contribution to the polarization and the electronic polarizability are conveniently given by sums of Wannier centroid contributions. Since each centroid is uniquely associated to an oxygen atom, only the oxygen neighborhoods enter the sum in Eq. 9. By contrast, all atomic neighborhoods, oxygen and hydrogen, are included in the sum when  $\mathcal{X}$  is the potential energy  $U$ .

The local quantity is an effective property whose precise value depends on the adopted decomposition. In the case of potential energy,  $\mathcal{X}_i$  denotes the atomic energy  $U_i$ , whose summation gives rise to the total potential energy  $U$  of the system in the DNN model. The analytical negative gradients of  $U$  with respect to atomic positions define the interatomic forces. In the case of polarizability, for the liquid water system we consider here,  $\mathcal{X}_i$  becomes the molecular polarizability  $\boldsymbol{\alpha}_i$ , where we only consider  $i$  to be oxygen.

The environmental dependence of  $\mathcal{X}_i$  can be accurately represented by a DNN-parametrized function  $\mathcal{X}_i^\gamma$ , where  $\gamma$  denotes the parameters of the DNN model. Due to the local dependence of  $\mathcal{X}_i^\gamma$  on the neighborhood of  $i$ , the DNN model is scalable by construction.  $\mathcal{X}_i^\gamma$  should also satisfy some additional criteria. First, it should depend continuously on the atomic positions and, in terms of efficiency, be orders of magnitude faster than *ab initio* models. Moreover, it is crucial to preserve the translational, rotational, and permutational symmetry of the quantity that is learned. The potential energy is a scalar quantity and should be invariant upon translation and rotation of the system and identical particle permutation. The polarization and the polarizability are vectorial and tensorial quantities, respectively, and should be translationally and permutationally invariant, but rotationally covariant. The above requirements are achieved by means of two DNNs, an *embedding* DNN and a *fitting* DNN. The embedding DNN maps the positions of the atoms belonging to the  $i$ th neighborhoods to features that are invariant under symmetry operations, while the fitting DNN maps these features in a way that is symmetry invariant for the potential energy, or covariant for the polarization and the polarizability. The number of hidden layers and outputs is refined in the training procedure.

In detail, we make a *local* frame transformation to the primed coordinates, which are defined relative to  $\mathbf{r}_i$ :

$$\mathbf{r}'_k \equiv \mathbf{r}_k - \mathbf{r}_i, \quad (10)$$

Translational symmetry is preserved by construction in the local frame. It is convenient to use generalized coordinates  $\mathbf{q}_k$ , which weight atoms according to their distance from site  $i$ ,  $r_{k'} \equiv (\mathbf{r}_{k'} \cdot \mathbf{r}_{k'})^{1/2}$ , and provide continuous evolution as atoms enter/exit the neighborhood. For this purpose we introduce a smooth weight function  $s(r')$  equal to  $1/r'_k$  at short distances and decaying smoothly to zero as  $r'_k$  approaches  $r_c$ , the radius of the neighborhood. The four-component vector  $\mathbf{q}_k = (q_k^1, q_k^2, q_k^3, q_k^4)$  is then given by  $(s(r'_k), s(r'_k)x'_k/r'_k, s(r'_k)y'_k/r'_k, s(r'_k)z'_k/r'_k)$ , in terms of the Cartesian components of  $\mathbf{r}'_k$ . We use the matrix  $\mathbf{Q} = (Q_{ki}) = (q_k^j)$ , which has  $N_i$  rows and 4 columns, to represent the set of generalized coordinates  $\{\mathbf{q}_k\}$  in a neighborhood.

The embedding DNN is the matrix  $\mathbf{E} = (E_{ik}) = E_i(s(r'_k))$  with  $M$  rows and  $N_i$  columns, whose elements are found by training, which maps each element in the set  $\{s(r'_k), r'_k \in \mathcal{N}_i\}$  onto  $M$  outputs. Multiplication of  $\mathbf{E}$  by  $\mathbf{Q}$  gives the matrix  $\mathbf{T} = \mathbf{E}\mathbf{Q}$  with  $M$  rows and 4 columns, whose generic element is:

$$T_{ij} = \sum_{k=1}^{N_i} E_i(s(r'_k)) Q_{kj}. \quad (11)$$

In Eq. 11, the permutationally invariant sum over the atoms is a smooth function of  $N_i$ . The last three columns of  $\mathbf{T}$  ( $j = 2, 3, 4$ ) transform covariantly under rotation because  $(Q_{k2}, Q_{k3}, Q_{k4})$  transforms like  $\mathbf{r}'_k$ . Let  $\mathbf{S}$  be the matrix formed by the first  $M'$  ( $< M$ ) rows of  $\mathbf{T}$ . Multiplication of  $\mathbf{T}$  by  $\mathbf{S}^T$ , the transpose of  $\mathbf{S}$ , gives the matrix  $\mathbf{D}$  of dimension  $M \times M'$ , called the *feature* matrix:

$$\mathbf{D} = \mathbf{T}\mathbf{S}^T, \quad (12)$$

The elements of  $\mathbf{D}$  are invariant under rotation and permutation.

$\mathbf{D}$  captures the local features of the neighboring pattern of  $i$  in a faithful and adaptive way. While a fitting network from  $\mathbf{D}$  to a scalar would properly define  $U_i$ , one needs different procedures for vectorial and tensorial quantities. The case of the polarization vector was discussed in Ref. 28. Here we specialize to the polarizability tensor. In this case, the output of the fitting network is an  $M \times M$  diagonal matrix  $\mathbf{F} = (F_{jk})$ , which is mapped onto  $\alpha_i$ , in a permutationally invariant and rotationally covariant way by right- and left-multiplications with the last three columns of  $\mathbf{T}$  and its transpose:

$$(\alpha_i)_{kl} = \sum_{j=1}^M T_{j,k+1} F_{jj}(\mathbf{D}) T_{j,l+1}.$$

Finally, the parameters  $\gamma$  are determined by training, i.e., an optimization process that minimizes a loss function, which is here the mean square difference between the DNN prediction and the training data. The Adam stochastic gradient descent method<sup>45</sup> is adopted for the optimization.

The DNN for the polarizability should be combined with a DNN for the PES to study the evolution of the polarizability along MD trajectories. For consistency, the two networks should be trained with electronic structure data at the same level of theory. In this paper, we use a DNN for the polarizability and a DP representation of the PES based on the same DFT data generated with the SCAN functional approximation. Since the *ab initio* electronic structure data are expensive, efficient learning strategies are crucial. To collect a minimal set of *ab initio* data for training, we use the iterative learning scheme Deep Potential Generator (DP-GEN) of Ref. 46. In this approach, an ensemble of DNN models, initially trained with the same limited set of *ab initio* data but different random initializations of the network parameters, are used to efficiently explore the configuration space. As done previously<sup>46,47</sup> we used 4 independent DNN models to represent the ensemble. A small subset of the visited configurations is selected with an error indicator, defined as the variance of the predictions within the ensemble DNNs. The protocol is repeated until all the explored configurations are described with satisfactory accuracy. The error indicator, here chosen to be the maximum standard deviation of atomic forces, exploits the highly non-linear dependence of the DNN models on the network parameters. As a consequence, different initializations of the parameters lead to different local minima in the landscape of the loss function, originating an ensemble of minimizing DNNs. In our experience, good DNN models constructed with the above procedure require significantly less *ab initio* data in the target thermodynamic range than learning approaches based on independent AIMD sampling data.

### 2.3 DFT Data and Deep Neural Networks for the Potential and Polarizability of Water

The ground state electronic structures of the equilibrium configurations of liquid water within DFT were calculated using the SCAN functional approximation<sup>48</sup>. SCAN predicts with sufficient accuracy the molecular structure of liquid water at equilibrium<sup>49</sup> as well as the dipole moment (SCAN:1.84 D, experi-

ment<sup>50</sup>: 1.86 D) and isotropic polarizability (SCAN: 1.41 Å<sup>3</sup>, experiment<sup>51</sup>: 1.47 Å<sup>3</sup>) of the water molecule in gas phase. We used the CP code of the Quantum ESPRESSO package<sup>52,53</sup> to compute the electronic ground state, the potential energy, and the forces on the atoms at selected molecular configurations. The same code was also used to compute the Wannier center coordinates and the electronic polarizability at fixed nuclear positions via the electric enthalpy method<sup>41,54</sup>. Norm-conserving pseudo-potentials of Troullier-Martins type<sup>55</sup> were used for both oxygen and hydrogen atoms, and the wavefunctions and charge density were plane-wave expanded with an energy cutoff of 110 and 440 Ry, respectively. Total energy was converged to 10<sup>-8</sup> Hartree or lower. The polarizability was estimated numerically with Eq. 5 using  $\epsilon = 0.001$  a.u., which falls well within the linear response regime. An average error of 0.0005 Å<sup>3</sup> for the DFT polarizabilities was estimated from a higher-order finite difference method based on a 5-point stencil. All Raman response calculations were performed at the experimental equilibrium density of water.

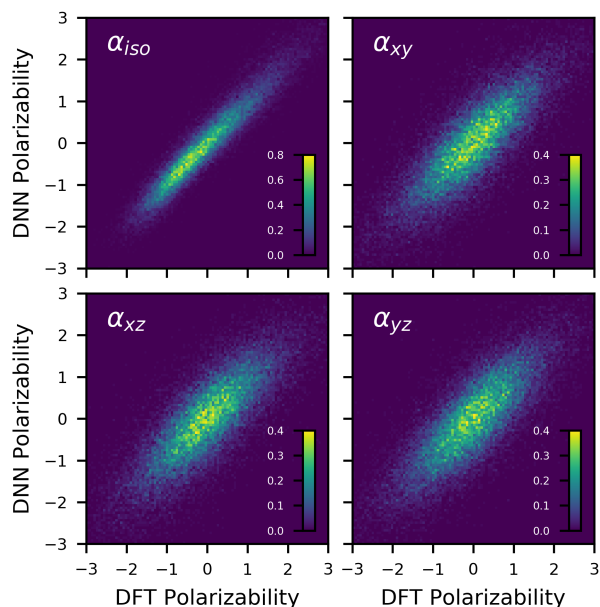


Fig. 1 Normalized probability distribution of predicted DNN and calculated DFT effective molecular polarizabilities for sample configurations of liquid water (see text). The polarizabilities are given in normalized units. Normalization consisted of subtraction of average DFT polarizability and division by the standard deviation of the DFT data. Isotropic polarizability and the off-diagonal components of the polarizability tensor are shown. Data compared here was not included in the training data of the deep neural network. The colorbar gives the scale used in each panel.

With the DP-GEN scheme we explored a pressure range of 1–10<sup>5</sup> bar with liquid water (192 atoms/cell), ice Ic (192 atoms/cell) and ice Ih (288 atoms/cell). The temperature range of the exploration ranged from 270–370 K for liquid water and 50–270 K for ices Ih and Ic. Simulations were sampled for at least 100 ps and atomic configurations with maximum standard deviation of atomic forces larger than 0.1 eV/Å were selected to retrain. No more than 400 atomic configurations in the exploration trajectory, separated on average by several ps, were selected per DP-GEN iteration. At

the end of the iterative training procedure, our training data contained a set of 2056, 2015 and 1708 configurations of liquid water, ice Ih and ice Ic, respectively. The referred training data set included only atomic forces and energy for each atomic configuration, the data needed to train the DNN potential energy surface (DP). The polarizability DNN, on the other hand, was trained only with effective molecular polarizabilities evaluated for the entire DP training set of liquid water. The resulting DP predicts water density ( $\rho$ ) and diffusion coefficient ( $D$ ) of  $1.07 \pm 0.02$  g/cm<sup>3</sup> and  $0.17 \pm 0.01$  Å<sup>2</sup>/ps, respectively, in close agreement with SCAN-AIMD ( $\rho = 1.05 \pm 0.03$  g/cm<sup>3</sup> and  $D = 0.19 \pm 0.02$  Å<sup>2</sup>/ps)<sup>49</sup>.

DP-based Molecular dynamics (DPMD) simulations of liquid water were carried out with 512 water molecules in a periodically repeated cubic cell of 24.9 Å size (H<sub>2</sub>O density of 1.0 g/cm<sup>3</sup>). The system was initially equilibrated at constant volume coupled to a single Nosé-Hoover thermostat<sup>56,57</sup> for 200 ps. The simulations proceeded at constant volume and energy for 2 ns, the only section of the simulation used to compute the spectra. The classical equations of motion were integrated with the velocity-Verlet algorithm with a time step of 0.25 and 0.5 fs for H<sub>2</sub>O and D<sub>2</sub>O, respectively. All simulations were performed with the LAMMPS<sup>58</sup> package interfaced with the DeepMD-Kit<sup>59</sup>. The DP-GEN package<sup>60</sup> was used to realize the iterative learning scheme.

## 2.4 Numerical Modeling of Raman Spectra

Given the polarizability tensor as a function of time, we numerically evaluate the classical TCF of the cell polarizability and take its discrete Fourier transform (Eq. 3). Furthermore, we also obtain the contribution of intermolecular coupling to the Raman spectra by decomposing the TCF of the cell polarizability into intramolecular and intermolecular terms.

The decomposition of the system polarizability into effective molecular polarizabilities enables us to distinguish spectral features due to autocorrelations (within the same effective molecule) from those due to intermolecular coupling, which weaken with increasing temperature. Analogously to Wan *et al.*<sup>5</sup>, in evaluating Eq. 3, we first calculated the intramolecular and intermolecular TCF. For the isotropic TCF, these are defined as:

$$C_{intra}(t) = \sum_{i=1}^N \langle \bar{\alpha}_i(0) \bar{\alpha}_i(t) \rangle_c \quad (13)$$

where  $N = 512$ , and

$$C_{inter}(t) = \sum_{i \neq j} \langle \bar{\alpha}_i(0) \bar{\alpha}_j(t) \rangle_c \quad (14)$$

with analogous definitions for the anisotropic TCF.

The decomposition of the cell polarizability into effective molecular polarizabilities also enables us to isolate sources of noise in the intermolecular spectrum. The sum in Eq. 14 was taken only over pairs  $\{i, j\}$  within the first two shells of neighboring molecules (6 Å), as determined by plotting the radial distribution function  $g_{OO}(r)$ . Including interactions from the third shell and beyond only adds to the zero-frequency component of the Fourier spectrum. The cutoffs were enforced on the coordinates of the oxygen atoms at each initial time used in the computation

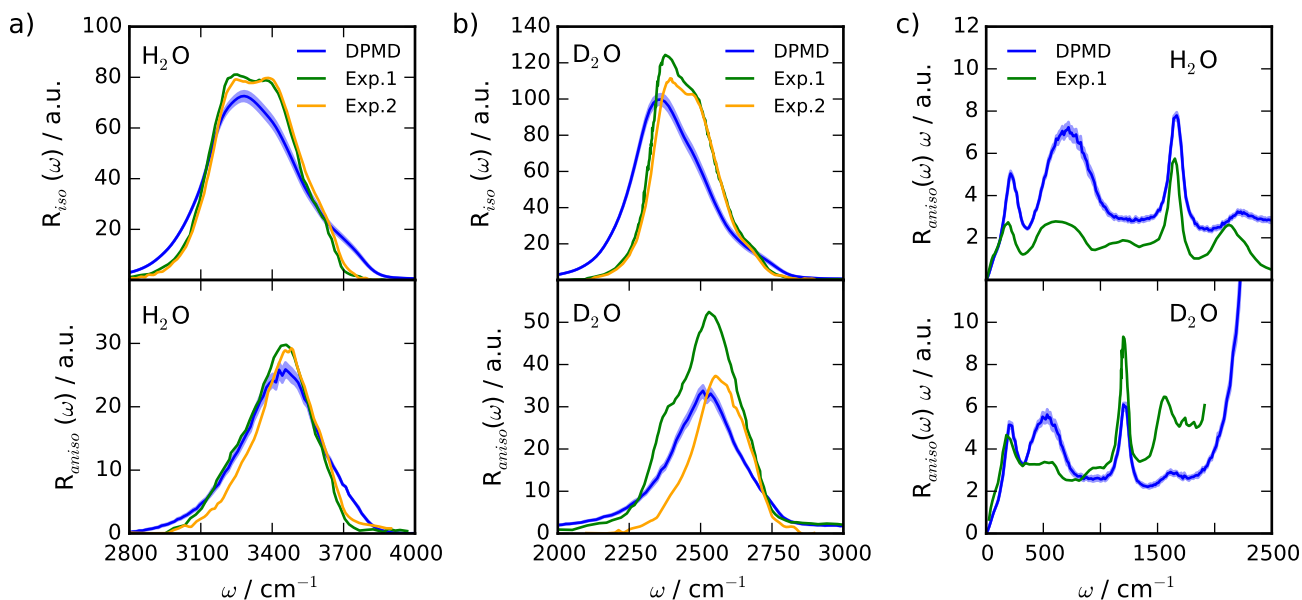


Fig. 2 a) Isotropic (top) and anisotropic (bottom) Raman frequency-reduced spectra of H<sub>2</sub>O at the OH stretching frequency region. b) Isotropic (top) and anisotropic (bottom) Raman spectra of D<sub>2</sub>O at the OD stretching frequency region. c) Low-frequency anisotropic Raman spectra of H<sub>2</sub>O (top) and D<sub>2</sub>O (bottom). Intensities in (c) were divided by 1000 for clarity. Experimental data obtained from Brooker *et al.*<sup>61</sup> (Exp.1, green lines) and Scherer *et al.*<sup>62</sup> (Exp.2, orange lines). Light blue regions indicate the standard deviation of DPMD-predicted spectra estimated with an ensemble of polarizability DNNs (see text). All intensities reported in arbitrary units (a.u.).

of the TCF.

Next, the spectrum was obtained by taking a discrete Fourier transform of the TCF and multiplying by the appropriate prefactors:

$$R(\omega_k) \propto n_{BE}(\omega_k) \sum_{t_m=0}^T e^{-i\omega_k t_m} C(t_m) \Delta t \quad (15)$$

where  $T = 2.5$  ps is the length of the TCF and the times  $t_m$  were discretized into intervals of  $\Delta t = 0.25$  fs for H<sub>2</sub>O and 0.5 fs for D<sub>2</sub>O. The frequencies  $\omega_k = 2\pi k/T$ , where  $k$  is an integer, run from 0 to the Nyquist frequency, discretized into units of  $13.33$  cm<sup>-1</sup>. This provides a bound on the resolution of the spectrum, so we report the locations of peaks to the nearest 10 cm<sup>-1</sup>.

### 3 Results and discussion

In this section, we present the isotropic and anisotropic Raman spectra computed from DPMD, compare to experiment, and discuss the temperature dependence of the low-frequency and stretch modes.

We first show in Fig. 1 the ability of our neural network model to predict effective molecular polarizabilities in excellent agreement with *ab initio* data. The units in Fig. 1 are relative to the standard deviation of DFT data. The molecular polarizabilities were obtained from a set of 416 liquid water atomic configurations not included in the training data. Our neural network model predicts similar polarizability distributions to DFT, with a better agreement for the isotropic polarizability than for the off-diagonal components of the polarizability tensor. By analyzing the data in Fig. 1, we find that 99.5 % of the DNN-predicted isotropic polarizabilities deviate by less than one standard deviation from the corresponding DFT isotropic polarizabilities. This percentage

decreases to 92.7, 91.9 and 92.3 % for the  $xy$ ,  $xz$  and  $yz$  components, respectively, of the polarizability tensor. It should be noted that the typical absolute values of the off-diagonal components of the polarizability tensor are approximately an order of magnitude smaller than its average diagonal component. To further estimate the uncertainty in the DNN-predicted Raman spectra of water we computed the Raman spectra of H<sub>2</sub>O and D<sub>2</sub>O from the polarizabilities predicted by 4 independently trained DNN models representative of a DNN polarizability ensemble. These models used the same training dataset and differed only in the random initialization of the network parameters, as we did to characterize the DNN ensemble when constructing the DP model of the PES. The small deviation between the four models, shown by the filled light blue regions in Fig. 2, confirms the consistency of our DNN model for the polarizability. Moreover, since Fig. 1 attests to the close agreement between the DNN and DFT polarizabilities, and the DPMD trajectories used in this paper closely approximate AIMD trajectories, as shown also in previous work<sup>19,47</sup>, this gives us confidence that the reported spectra accurately represent SCAN DFT spectra.

#### 3.1 Comparison with Experiment

Our key results are shown in Fig. 2, which compares the DPMD spectra to experimental data obtained from Brooker *et al.*<sup>61</sup> and Scherer *et al.*<sup>62</sup>. In order to properly compare the intensities from simulation and experiment, we set the integral of the DPMD isotropic spectrum of water between 2700 and 4000 cm<sup>-1</sup> equal to the integral of the same region in the experimental spectrum. The intensities reported for all the other spectra in Fig. 2 (except the experimental low-frequency spectra) are relative to the units de-



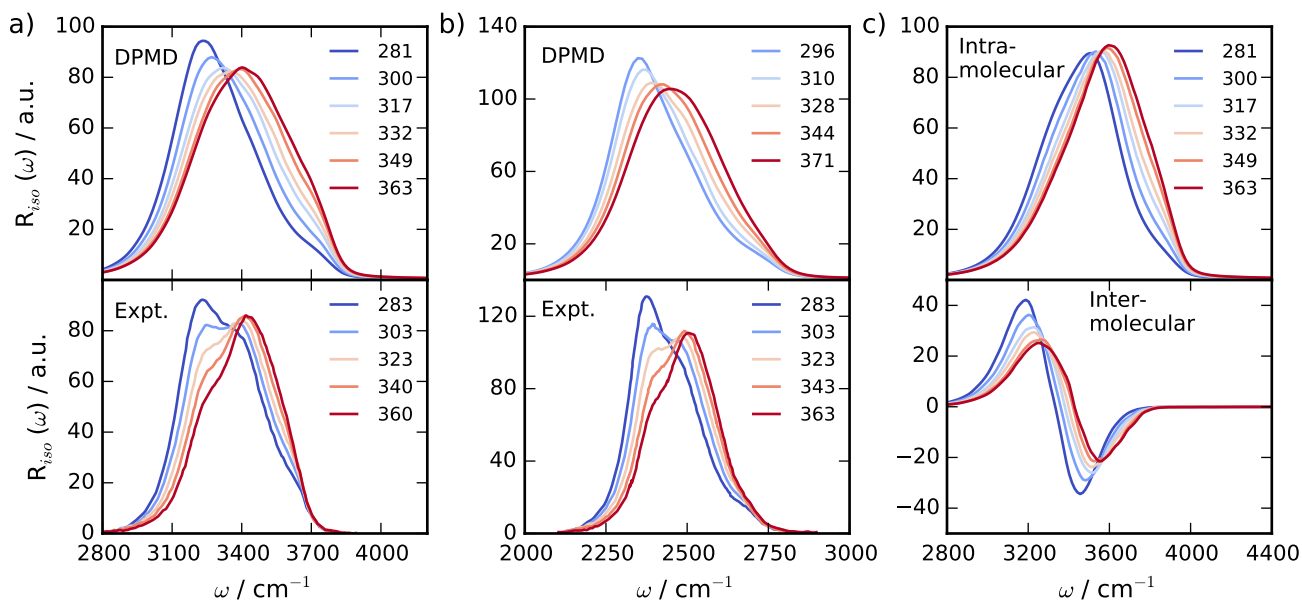


Fig. 3 a) Isotropic Raman spectrum of H<sub>2</sub>O at the OH stretching region as a function of temperature. b) Isotropic Raman spectrum of D<sub>2</sub>O at the OD stretching region as a function of temperature. c) Intramolecular (top) and intermolecular (bottom) contributions to the isotropic Raman spectrum of H<sub>2</sub>O at the OH stretching region as a function of temperature. Experimental data obtained from Scherer *et al.*<sup>62</sup>. Temperature reported in Kelvin units. All intensities reported in arbitrary units (a.u.).

fined in Fig. 2a. The experimental intensities of low-frequency region were estimated by comparing the amplified and unamplified intensities of the HOH (or DOD) bending peaks reported by Brooker *et al.*<sup>61</sup>. Magnification factors of 50× and 40× for the amplified low frequency spectra relative to their unamplified counterparts were crudely estimated for H<sub>2</sub>O and D<sub>2</sub>O, respectively, but this crude estimate has an error of at least 30%, and therefore the experimental intensities shown in Fig. 2c should be taken with reservation. To make a direct comparison, we plot the temperature-reduced spectra at the OH stretch region,  $R(\omega)$ , and the frequency-reduced spectra,  $R(\omega)\omega$ , in the low-frequency region as done in Brooker *et al.*<sup>61</sup>

For H<sub>2</sub>O, we successfully captured the peaks in the isotropic and anisotropic spectra, at 3250 cm<sup>-1</sup> and 3470 cm<sup>-1</sup> respectively. We are also able to reproduce the accurate location of the peaks in the D<sub>2</sub>O spectrum, at  $\approx$  2350 cm<sup>-1</sup> for the isotropic spectrum and  $\approx$  2530 cm<sup>-1</sup> for the anisotropic spectrum. The quoted wavenumbers of the peaks are defined as the positions, to the nearest 10 cm<sup>-1</sup>, of the maxima in  $R(\omega)$  in the interval 2800 to 4000 cm<sup>-1</sup> for H<sub>2</sub>O, and 2000 to 3000 cm<sup>-1</sup> for D<sub>2</sub>O. We also observe a shoulder at  $\approx$  3750 cm<sup>-1</sup> of the H<sub>2</sub>O spectrum, blueshifted relative to experiment, which originates from the OH stretch of transiently broken H-bonds. The main shortcoming of our results is the absence of a shoulder at 3400 cm<sup>-1</sup> in the H<sub>2</sub>O spectrum, previously assigned to the asymmetric stretch of water. With the exception of LSC-IVR<sup>16</sup>, this shoulder was likewise absent from the theoretical Raman spectra obtained with other methods<sup>5,14</sup>.

Another metric for comparison to experiment is the full width at half maximum (FWHM) of the peaks in the OH and OD stretch, which is sensitive to the choice of functional and the method of approximate quantum dynamics. Several theoretical studies have

predicted a stretch band that is either too broad (such as spectra computed using the PBE functional<sup>5,14</sup>) or too narrow (such as the H<sub>2</sub>O spectra computed using a semiclassical approach<sup>15</sup> and MB-pol<sup>13</sup>). We find a FWHM of 430 cm<sup>-1</sup> (270 cm<sup>-1</sup>) for the isotropic OH (OD) stretch, greater than the experimental width of 418 cm<sup>-1</sup> (236 cm<sup>-1</sup>) determined from the temperature-reduced spectrum in Brooker *et al.* In the anisotropic spectrum, we find widths of 370 cm<sup>-1</sup> (286 cm<sup>-1</sup>) for OH (OD), compared to 328 cm<sup>-1</sup> (306 cm<sup>-1</sup>) from Brooker's experiment and 296 cm<sup>-1</sup> (225 cm<sup>-1</sup>) from Scherer *et al.* Possible causes for this artificial broadness are the spurious electronic self-interaction present in the SCAN functional, the missing quantum fluctuations in our treatment of molecular dynamics, or a combination of both effects.

With DPMD, we are able to produce accurate results not only in the stretch band, but across the entire spectrum. We qualitatively reproduce an array of features present in the anisotropic spectra, with approximately correct intensities, over an order of magnitude less than the stretching band. These features include the network modes below 300 cm<sup>-1</sup>, examined in greater detail below; the libration mode from  $\approx$  400 cm<sup>-1</sup> to  $\approx$  800 cm<sup>-1</sup>; the bending peaks at 1670 cm<sup>-1</sup> and 1210 cm<sup>-1</sup> in H<sub>2</sub>O and D<sub>2</sub>O respectively; and the combination bands at 1640 cm<sup>-1</sup> in D<sub>2</sub>O and 2250 cm<sup>-1</sup> in H<sub>2</sub>O. The most significant discrepancy between the DPMD-predicted and experimental data is the librational peak, which extends approximately 100 cm<sup>-1</sup> higher in wavenumber compared to experiment. This indicates that the libration potential is stiffer than in experiment, which could result from overbonding. Moreover, our results do not capture the significant suppression of the libration peak in D<sub>2</sub>O compared to H<sub>2</sub>O seen in the experimental data. Previous simulations, both those using



the harmonic approximation<sup>5</sup> and those employing approximate quantum dynamics<sup>13,14</sup>, have also produced an incorrect intensity in this region.

### 3.2 Temperature Dependence of the OH/OD Stretch

A novel contribution of this paper is that, by running the DPMD simulations at six temperatures from approximately 280 to 370 K, we can observe the temperature dependence of OH/OD stretch. Further, we decompose the spectra into intermolecular and intramolecular contributions, which provides insights into the mechanisms driving the temperature dependence.

The theoretical spectra in Figs. 3a and 3b exhibit a blueshift in the isotropic stretch bands due to weakening hydrogen bonds as the temperature increases. Experimentally, this blueshift manifests as a change in the relative intensities of the symmetric ( $\approx 3200\text{ cm}^{-1}$ ) and asymmetric ( $\approx 3400\text{ cm}^{-1}$ ) OH stretch peaks. The shoulder is not visible in the DPMD-computed spectra, but we do accurately reproduce the magnitude of the blueshift. Qualitatively, the experimentally observed decrease in maximum intensity is captured for both H<sub>2</sub>O and D<sub>2</sub>O.

While several experimental studies have proposed Gaussian spectral decomposition of the stretch band as a mean for understanding the role of different hydrogen bond configurations, there is no consensus on the number of components<sup>33</sup>, or even whether the multistructure model is the correct starting point<sup>32</sup>. *Ab initio* modeling offers an alternative route to decomposing the spectrum: by splitting into the intramolecular and intermolecular contributions, we can understand how the relative intensity of couplings between and within molecules varies as the temperature increases. This is shown for the isotropic spectrum of H<sub>2</sub>O in Fig. 3c.

There are several features worth noting in Fig. 3c, which together are responsible for the temperature dependence of the total spectrum. First, the intermolecular spectrum exhibits a maximum in the range of 3190 to 3280  $\text{cm}^{-1}$ , followed by a minimum in the range of 3450 to 3570  $\text{cm}^{-1}$ , both of approximately equal magnitude. At a given temperature, both extrema tend to redshift the frequency at which the total spectrum has a maximum. Second, as the temperature increases, both extrema are blueshifted and reduced in intensity. The same behavior occurs in the intramolecular spectrum, but to a lesser extent: the peak is blueshifted from 3430 to 3520  $\text{cm}^{-1}$  with a slight decrease in intensity. Thus, we can attribute the overall blueshift in the spectrum not only to the separate blueshifting of the two components, but perhaps more importantly, to the reduced role of intermolecular coupling as the temperature rises. We also note that at all temperatures, the intramolecular contribution dominates.

The extent of intermolecular coupling in the Raman spectra of liquid water can be controlled experimentally through isotopic substitution. Raman spectra of dilute HOD in D<sub>2</sub>O, for instance, is able to probe the OH stretch almost uncoupled from the surrounding oscillators. The temperature dependence of both experiment and our DPMD simulations of HOD in D<sub>2</sub>O (Fig. 4) shows a weaker temperature dependence of the OH stretch peak relative to pure H<sub>2</sub>O, confirming the significance of the intermolecular

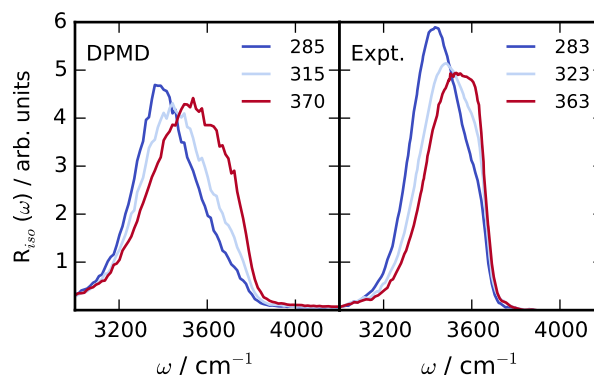


Fig. 4 DPMD-predicted and experimental isotropic Raman spectra of dilute HOD in D<sub>2</sub>O at the OH stretching frequency. Experimental data obtained from Scherer *et al.*<sup>62</sup>. Intensities of the experimental spectra were normalized such that the integral of the 283 K experimental spectrum was equal to the integral of the 285 K DPMD spectrum between 2700 and 4000  $\text{cm}^{-1}$ . Temperature reported in Kelvin units.

coupling to the temperature dependence of H<sub>2</sub>O Raman spectra. From Fig. 4 we also observe a good agreement between DPMD-predicted and the experimental spectra, although DPMD predicts broader peaks. The experimental spectra also contains a shoulder at  $\approx 3600\text{ cm}^{-1}$ , usually attributed to the oscillation of non-H-bonded OH groups. This shoulder is blueshifted in the DPMD-predicted spectrum at 370 K.

### 3.3 Temperature Dependence of the Low Frequency Spectrum

Turning now to the low-frequency regime, the anisotropic spectrum of H<sub>2</sub>O below 300  $\text{cm}^{-1}$  is shown in Fig. 5. At 300 K, we find the nominal 60  $\text{cm}^{-1}$  and 180  $\text{cm}^{-1}$  peaks whose presence has been firmly established in experiment<sup>35,63</sup>. Walrafen *et al.* further observe that as the temperature increases, the 180  $\text{cm}^{-1}$  peak decreases in intensity, disappearing entirely in the gas phase. This temperature dependence is evident, but slightly less pronounced in the DNN-predicted spectra relative to experiment.

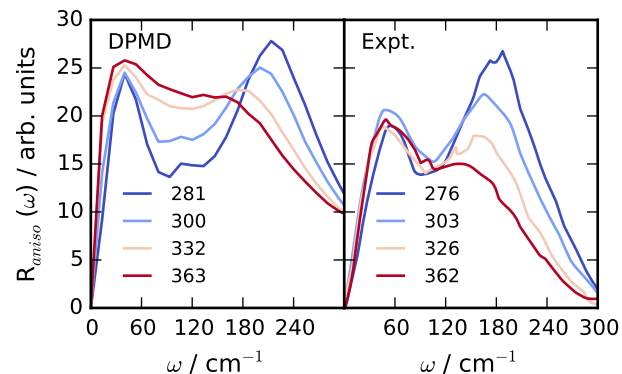


Fig. 5 Low frequency region of H<sub>2</sub>O anisotropic Raman spectrum as a function of temperature (K). Experimental spectra were multiplied by a constant factor in order to match maximum intensities of lowest temperature DPMD and experimental spectra. Experimental data taken from Walrafen *et al.*<sup>35</sup>.

While the two low-frequency peaks are experimentally attested, their origin is not entirely clear. Previous *ab initio* simulations of the IR spectrum of water and ice have assigned these peaks to H bond network modes: bending of H bonded oxygen atoms ( $\approx 70 \text{ cm}^{-1}$ ) and stretching of H bonded oxygen atoms ( $\approx 200 \text{ cm}^{-1}$ )<sup>64</sup>. Another school of thought assigns both peaks to restricted translational modes, modeling the  $180 \text{ cm}^{-1}$  mode as a harmonic oscillator which changes from underdamped to overdamped as the temperature increases<sup>65</sup>.

In an AIMD study of the low-frequency  $\text{D}_2\text{O}$  spectrum, Wan *et al.* attribute the  $60 \text{ cm}^{-1}$  peak to intramolecular dipole induced-dipole modes and the  $200 \text{ cm}^{-1}$  peak to intermolecular charge fluctuations<sup>5</sup>. The DNN-predicted spectra in Fig. 6, obtained from a more accurate functional, larger system size, and longer trajectories than Wan *et al.*, are consistent with these conclusions. At 300 K, the nominal  $60 \text{ cm}^{-1}$  peak is dominated by the maximum in the intramolecular spectrum at  $70 \text{ cm}^{-1}$ , with a much weaker, redshifted contribution from the intermolecular spectrum at  $30 \text{ cm}^{-1}$ . In addition, while the intramolecular spectrum exhibits a shoulder at  $190 \text{ cm}^{-1}$ , a more distinct peak appears in the intermolecular spectrum, at  $200 \text{ cm}^{-1}$ .

The division into intramolecular and intermolecular components also sheds some light on the temperature dependence of the peaks. As the temperature rises, the  $60 \text{ cm}^{-1}$  peak in the intramolecular spectrum is slightly redshifted, and the shoulder at  $210 \text{ cm}^{-1}$  redshifts and disappears. As in the OH stretch, the intermolecular contribution is far more sensitive to temperature; the nominal  $180 \text{ cm}^{-1}$  peak redshifts from  $210 \text{ cm}^{-1}$  at 281 K to  $160 \text{ cm}^{-1}$  at 363 K, dropping in intensity by a factor of 0.6. By accessing long time scales and controlling noise, DPMD enables us to accurately model this intermolecular contribution to the spectrum, which is crucial in explaining the temperature dependence at both low and high frequencies.

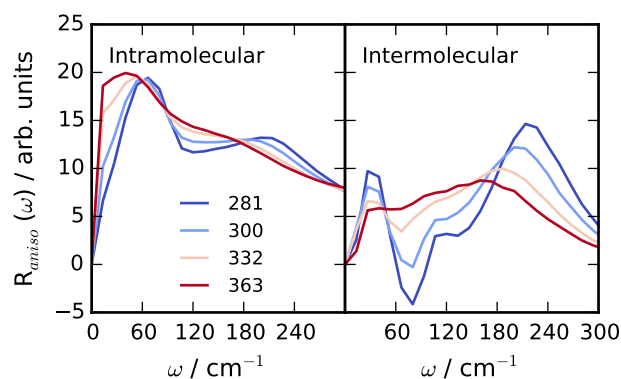


Fig. 6 Intramolecular (top) and intermolecular (bottom) contributions to the low frequency region of  $\text{H}_2\text{O}$  anisotropic Raman spectrum as a function of temperature (K).

## 4 Conclusion

In this paper, we have constructed a DNN representation of the physical properties of a molecular system, focusing specifically on the effective molecular polarizability. By doing so, we were able to model the temperature dependence of the Raman spec-

trum, a useful tool for examining the local structure of condensed phase systems. Liquid water offers a particularly interesting testing ground for the DNN framework, as its complex hydrogen bond networks lead to unique vibrational signatures which have thus far defied full theoretical explanation. The DPMD-predicted trends in the spectra for  $\text{H}_2\text{O}$  and  $\text{D}_2\text{O}$  at temperatures ranging from 280 to 360 K are in good agreement with experiment, although nuclear quantum effects must be included to complete the picture.

We emphasize, however, that while the Raman spectra presented here are for water, the DNN framework is fully generalizable to other systems of interest in chemical physics. Since DPMD greatly reduces the uncertainty due to limited statistics in AIMD, it offers a useful starting point for assessing the systematic errors of DFT functionals and approximate quantum corrections relative to experiment. Although the SCAN functional describes H-bonds and intermediate van der Waals forces significantly better than GGA functionals, it is still affected by self-interaction errors. Going forward, a study of the effect of using different functional approximations in the Raman spectra will be useful. Of particular interest is the effect of functional approximations, which reduce the self-interaction errors. It will also be worthwhile to study how quantum corrections affect the Raman spectra. The orders-of-magnitude efficiency gains of DPMD over AIMD should enable various semi-classical methods and even permit the analytic continuation of imaginary time data.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

This work was conducted within the Computational Chemical Center: Chemistry in Solution and at Interfaces funded by the DoE under Award DE-SC0019394. We used resources of the National Energy Research Scientific Computing Center (DoE Contract No. DE-AC02-05cH11231). We also acknowledge use of the TIGRESS High Performance Computer Center at Princeton University. GS was supported by the PACM Summer Fellowship through the Program in Applied and Computational Mathematics at Princeton University.

## Notes and references

- 1 P. J. Hendra and P. M. Stratton, *Chemical Reviews*, 1969, **69**, 325–344.
- 2 D. McQuarrie, *Statistical Mechanics*, University Science Books, 2000.
- 3 R. Resta, *Ferroelectrics*, 1992, **136**, 51–55.
- 4 R. D. King-Smith and D. Vanderbilt, *Physical Review B*, 1993, **47**, 1651–1654.
- 5 Q. Wan, L. Spanu, G. A. Galli and F. Gygi, *Journal of chemical theory and computation*, 2013, **9**, 4124–4130.
- 6 A. Putrino and M. Parrinello, *Physical Review Letters*, 2002, **88**, 176401.
- 7 G. Cassone, J. Spöner, S. Trusso and F. Saija, *Physical Chemistry Chemical Physics*, 2019, **21**, 21205–21212.

- 8 T. D. Kühne, *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2014, **4**, 391–406.
- 9 C. Vega, *Molecular Physics*, 2015, **113**, 1145–1163.
- 10 T. Hasegawa and Y. Tanimura, *The Journal of Physical Chemistry B*, 2011, **115**, 5545–5553.
- 11 M. Chen, L. Zheng, B. Santra, H. Y. Ko, R. A. Distasio, M. L. Klein, R. Car and X. Wu, *Nat. Chem.*, 2018, **10**, 413–419.
- 12 S. K. Reddy, S. C. Straight, P. Bajaj, C. Huy Pham, M. Riera, D. R. Moberg, M. A. Morales, C. Knight, A. W. Götz and F. Paesani, *The Journal of Chemical Physics*, 2016, **145**, 194504.
- 13 G. R. Medders and F. Paesani, *Journal of Chemical Theory and Computation*, 2015, **11**, 1145–1154.
- 14 O. Marsalek and T. E. Markland, *The Journal of Physical Chemistry Letters*, 2017, **8**, 1545–1551.
- 15 M. Yang and J. L. Skinner, *Phys. Chem. Chem. Phys.*, 2010, **12**, 982–991.
- 16 X. Liu and J. Liu, *Molecular Physics*, 2018, **116**, 755–779.
- 17 R. L. Benson, G. Trenins and S. C. Althorpe, *Faraday Discuss.*, 2020, **221**, 350–366.
- 18 M. Rossi, H. Liu, F. Paesani, J. Bowman and M. Ceriotti, *The Journal of Chemical Physics*, 2014, **141**, 181101.
- 19 L. Zhang, J. Han, H. Wang, R. Car and W. E, *Phys. Rev. Lett.*, 2018, **120**, 143001.
- 20 J. Behler and M. Parrinello, *Phys. Rev. Lett.*, 2007, **98**, 146401.
- 21 S. Chmiela, H. E. Sauceda, K. R. Müller and A. Tkatchenko, *Nature Communications*, 2018, **9**, 3887.
- 22 A. P. Bartók, M. C. Payne, R. Kondor and G. Csányi, *Physical Review Letters*, 2010, **104**, 136403.
- 23 M. Rupp, A. Tkatchenko, K.-R. Müller and O. A. VonLilienfeld, *Physical Review Letters*, 2012, **108**, 058301.
- 24 S. Chmiela, A. Tkatchenko, H. E. Sauceda, I. Poltavsky, K. T. Schütt and K.-R. Müller, *Science Advances*, 2017, **3**, e1603015.
- 25 K. Schütt, P.-J. Kindermans, H. E. S. Felix, S. Chmiela, A. Tkatchenko and K.-R. Müller, *Advances in Neural Information Processing Systems*, 2017, pp. 992–1002.
- 26 J. Han, L. Zhang, R. Car and W. E, *Communications in Computational Physics*, 2018, **23**, 629–639.
- 27 L. Zhang, J. Han, H. Wang, W. Saidi, R. Car and W. E, *Advances in Neural Information Processing Systems 31*, Curran Associates, Inc., 2018, pp. 4441–4451.
- 28 L. Zhang, M. Chen, X. Wu, H. Wang, W. E and R. Car, *arXiv preprint arXiv:1906.11434*, 2019.
- 29 A. Grisafi, D. M. Wilkins, M. J. Willatt and M. Ceriotti, in *Atomic-Scale Representation and Statistical Learning of Tensorial Properties*, ch. 1, pp. 1–21.
- 30 D. M. Wilkins, A. Grisafi, Y. Yang, K. U. Lao, R. A. DiStasio and M. Ceriotti, *Proceedings of the National Academy of Sciences of the United States of America*, 2019, **116**, 3401–3406.
- 31 N. Raimbault, A. Grisafi, M. Ceriotti and M. Rossi, *New Journal of Physics*, 2019, **21**, 105001.
- 32 J. D. Smith, C. D. Cappa, K. R. Wilson, R. C. Cohen, P. L. Geissler and R. J. Saykally, *Proceedings of the National Academy of Sciences*, 2005, **102**, 14171–14174.
- 33 Q. Hu, H. Zhao and S. Ouyang, *Phys. Chem. Chem. Phys.*, 2017, **19**, 21540–21547.
- 34 J. R. Scherer, M. K. Go and S. Kint, *The Journal of Physical Chemistry*, 1974, **78**, 1304–1313.
- 35 G. E. Walrafen, M. R. Fisher, M. S. Hokmabadi and W. H. Yang, *The Journal of Chemical Physics*, 1986, **85**, 6970–6982.
- 36 M. H. Brooker, G. Hancock, B. C. Rice and J. Shapter, *Journal of Raman Spectroscopy*, 1989, **20**, 683 – 694.
- 37 B. Schrader and D. Moore, *Pure Appl. Chem.*, 1997, **69**, 1451–1468.
- 38 O. F. Nielsen, *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.*, 1996, **93**, 57–99.
- 39 L. Wang, M. Ceriotti and T. E. Markland, *The Journal of Chemical Physics*, 2014, **141**, 104502.
- 40 S. Baroni, S. de Gironcoli and A. Dal Corso, *Reviews of Modern Physics*, 2001, **73**, 515–562.
- 41 P. Umari and A. Pasquarello, *Phys. Rev. Lett.*, 2002, **89**, 157602.
- 42 I. Souza, J. Iñiguez and D. Vanderbilt, *Physical review letters*, 2002, **89**, 117602.
- 43 N. Marzari and D. Vanderbilt, *Phys. Rev. B*, 1997, **56**, 12847–12865.
- 44 M. Sharma, Y. Wu and R. Car, *Int. J. Quantum Chem.*, 2003, **95**, 821–829.
- 45 D. P. Kingma and J. Ba, *arXiv preprint arXiv:1412.6980*, 2015, **9**, year.
- 46 L. Zhang, D.-Y. Lin, H. Wang, R. Car and W. E, *Physical Review Materials*, 2019, **3**, 023804.
- 47 M. F. Calegari Andrade, H.-Y. Ko, L. Zhang, R. Car and A. Selloni, *Chemical Science*, 2020, **11**, 2335–2341.
- 48 J. Sun, A. Ruzsinszky and J. P. Perdew, *Phys. Rev. Lett.*, 2015, **115**, 036402.
- 49 M. Chen, H.-Y. Ko, R. Remsing, M. F. Calegari Andrade, B. Santra, Z. Sun, A. Selloni, R. Car, M. Klein, J. Perdew and X. Wu, *Proceedings of the National Academy of Sciences of the United States of America*, 2017, **114**, 10846–10851.
- 50 S. A. Clough, Y. Beers, G. P. Klein and L. S. Rothman, *The Journal of Chemical Physics*, 1973, **2254**, 2254–2259.
- 51 W. F. Murphy, *The Journal of Chemical Physics*, 1977, **67**, 5877–5882.
- 52 P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari and R. M. Wentzcovitch, *J. Phys.: Condens. Matter*, 2009, **21**, 395502.
- 53 P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. Buongiorno Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. Dal Corso, S. De Gironcoli, P. Delugas, R. A. Distasio, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H. Y. Ko, A. Kokalj,

- E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H. V. Nguyen, A. Otero-De-La-Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu and S. Baroni, *J. Phys.: Condens. Matter*, 2017, **29**, 465901.
- 54 R. W. Nunes and D. Vanderbilt, *Phys. Rev. Lett.*, 1994, **73**, 712–715.
- 55 N. Troullier and J. L. Martins, *Phys. Rev. B*, 1991, **43**, 1993–2006.
- 56 S. Nosé, *Molecular Physics*, 1984, **52**, 255–268.
- 57 W. G. Hoover, *Physical Review A*, 1985, **31**, 1695–1697.
- 58 S. Plimpton, *Journal of Computational Physics*, 1995, **117**, 1–19.
- 59 H. Wang, L. Zhang, J. Han and W. E, *Computer Physics Communications*, 2018, **228**, 178–184.
- 60 Y. Zhang, H. Wang, W. Chen, J. Zeng, L. Zhang, H. Wang and W. E, *Computer Physics Communications*, 2020, 107206.
- 61 M. H. Brooker, G. Hancock, B. C. Rice and J. Shapter, *Journal of Raman Spectroscopy*, 1989, **20**, 683–694.
- 62 J. R. Scherer, M. K. Go and S. Kint, *The Journal of Physical Chemistry*, 2005, **78**, 1304–1313.
- 63 G. E. Walrafen, Y. C. Chu and G. J. Piermarini, *The Journal of Physical Chemistry*, 1996, **100**, 10363–10372.
- 64 W. Chen, M. Sharma, R. Resta, G. Galli and R. Car, *Physical Review B - Condensed Matter and Materials Physics*, 2008, **77**, 1–5.
- 65 K. Mizoguchi, Y. Hori and Y. Tominaga, *The Journal of Chemical Physics*, 1992, **97**, 1961–1968.