Exploring the Absorption Spectrum of Simulated Water from MHz to the Infrared - Supplementary Information

Shane Carlson[1], Florian N. Brünig[1], Philip Loche[1], Douwe Jan Bonthuis[2], and Roland R. Netz*[1]

[1]Fachbereich Physik, Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany
[2]Institute of Theoretical and Computational Physics, Graz University of Technology, 8010 Graz, Austria

June 22, 2020

S1 Compiled Experimental Data

We compiled experimental spectral data for water at room temperature from existing literature for comparison with our simulated spectra. The compiled experimental dataset consists of data from five different sources for liquid water at or near 300 K, covering different frequency ranges to overlap:

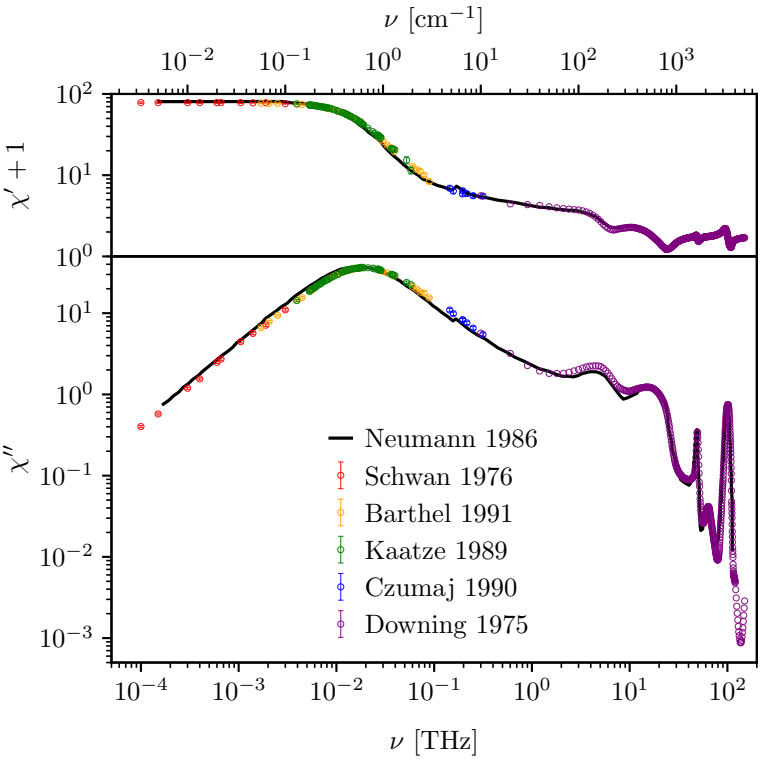

Figure S1: Experimental spectra for water at 300 ± 2 K, compared, with $\varepsilon'(\nu) = \chi'(\nu) + 1$ in the top panel, $\chi''(\nu)$ in the bottom. The data shown as colored circles were found in Refs.[1,2,3,4,5] and are the data that comprise the compiled dataset that we use for comparison to simulated spectra. The black curve shows another dataset shown only in this figure for comparison.[6] Error bars are shown for each of the compiled datasets except for the Downing dataset.

- Ref.[1] (Schwan 1976), at a reported temperature of 25°C, covers 0.1–3 GHz. It consists of datasets from two separate laboratories, spanning the ranges 0.1–0.7 GHz and 0.8–3 GHz. The highest- and lowest-frequency points of the first and second datasets respectively (at 0.7 and 0.8 GHz) are omitted, as they appear to represent outliers.
- Ref.[2] (Barthel 1991), at a reported temperature of 25°C, covers 1.7–89 GHz.

- Ref.[3] (Kaatze 1989), at a reported temperature of 25°C, covers 1.8–58 GHz. The first ten datapoints, covering 1.8–3.75 GHz are noisy, and as this range is also covered by Ref.[2], they are omitted.
- Ref.[4] (Czumaj 1990), at a reported temperature of 25°C, covers 57–315 GHz. The first five datapoints, covering 57-90 GHz are noisy, and as this range is also covered by Ref.[2], they are omitted.
- Ref.[5] (Downing 1975), at a reported temperature of 27°C, covers 0.3–150 THz. The quantities given are the indices of refraction and absorption, $n(\nu)$ and $k(\nu)$, from which we calculated the susceptibility via $\chi'(\nu) + 1 = n(\nu)^2 - k(\nu)^2$, and $\chi''(\nu) = 2n(\nu)k(\nu)$.

Thus, the compiled experimental dataset covers a frequency range of 100 MHz to 150 THz. Figure S1 shows the compiled data as colored circles compared with a dataset digitized from a figure in Ref.[6] (Neumann 1986) spanning from ∼0.2 GHz to ∼110 THz. In figures comparing experimental and simulated spectra in the main work, the data are plotted directly as a heavy black curve without errorbars. Note that Refs.[1,2,4] were found via a compilation in Ref.[7].

S2 Complex Electric Susceptibility from Fluctuation Dissipation Theorem

Assuming an isotropic medium perturbed by an applied electric field $\boldsymbol{E}$, whose Hamiltonian takes the form $H = H_0 - \boldsymbol{P} \cdot \boldsymbol{E}$, the linear response of the total system dipole moment $\boldsymbol{P}$ (not to be confused with the polarization density) is described by the time-dependent electric susceptibility $\chi(t)$ via

$$\boldsymbol{P}(t) = \int_{-\infty}^{t} \mathrm{d}t' \; V\varepsilon_0 \chi(t-t')\boldsymbol{E}(t') \,. \tag{S1}$$

where V is the system volume. Fourier transforming Eq. (S1) gives

$$\widetilde{\boldsymbol{P}}(\nu) = V\varepsilon_0 \chi(\nu)\widetilde{\boldsymbol{E}}(\nu) \,, \tag{S2}$$

where $\chi(\nu)$ is the positive-domain Fourier transform of $\chi(t)$, and is known as the complex electric susceptibility, generalized electric susceptibility, or frequency-dependent electric susceptibility. $\chi(\nu)$ is a dimensionless, complex quantity, denoted here as $\chi(\nu) = \chi'(\nu) - i\chi''(\nu)$, in order that $\chi''(\nu)$ be positive for positive ν. The fluctuation dissipation relation expresses the linear response function $V\varepsilon_0\chi(\nu)$ in terms of an equilibrium ensemble average

$$V\varepsilon_0\chi(\nu) = \frac{-1}{3k_BT} \int_0^{\infty} \mathrm{d}t \; e^{-2\pi i\nu t} \frac{\mathrm{d}}{\mathrm{d}t} \langle \boldsymbol{P}(0) \cdot \boldsymbol{P}(t) \rangle \,, \tag{S3}$$

where the arithmetic mean has been taken over the three spatial dimensions. Here we have taken the Fourier transform of the function $f(t)$ to be defined as

$$\tilde{f}(\nu) = \int_{-\infty}^{\infty} \mathrm{d}t \; e^{-2\pi i\nu t} f(t) \,, \tag{S4}$$

whose inverse Fourier-transform is

$$f(t) = \int_{-\infty}^{\infty} \mathrm{d}\nu \; e^{2\pi i\nu t} \tilde{f}(\nu) \,. \tag{S5}$$

S3 The Dissipative Part of the Susceptibility

Starting from Eq. (S3) and using that $\langle \boldsymbol{P}(0) \cdot \boldsymbol{P}(t) \rangle$ is real, the imaginary (dissipative) part of $\chi(\nu)$ is found to be

$$\chi''(\nu) = \frac{-1}{3Vk_BT\varepsilon_0} \int_0^{\infty} \mathrm{d}t \; \sin(2\pi\nu t) \; \frac{d}{dt} \langle \boldsymbol{P}(0) \cdot \boldsymbol{P}(t) \rangle \,. \tag{S6}$$

Because the autocorrelaton function is symmetric, the integrand is symmetric, thus

$$\begin{aligned}\chi''(\nu) &= \frac{-1}{3Vk_BT\varepsilon_0}\frac{1}{2}\int_{-\infty}^{\infty}\mathrm{d}t\ \sin(2\pi\nu t)\ \frac{d}{dt}\langle \boldsymbol{P}(0)\cdot\boldsymbol{P}(t)\rangle \\ &= \frac{1}{3Vk_BT\varepsilon_0}\frac{1}{2}\ \mathrm{Im}\int_{-\infty}^{\infty}\mathrm{d}t\ e^{-2\pi i\nu t}\frac{d}{dt}\langle \boldsymbol{P}(0)\cdot\boldsymbol{P}(t)\rangle \\ &= \frac{\pi}{3Vk_BT\varepsilon_0}\nu\ \mathrm{Re}\int_{-\infty}^{\infty}\mathrm{d}t\ e^{-2\pi i\nu t}\ \langle \boldsymbol{P}(0)\cdot\boldsymbol{P}(t)\rangle\,. \end{aligned} \tag{S7}$$

Using that the Fourier transform of a symmetric function is always real-valued gives for the dissipative part,

$$\chi''(\nu) = \frac{\pi}{3V\varepsilon_0 k_BT}\,\nu\int_{-\infty}^{\infty}\mathrm{d}t\ e^{-2\pi i\nu t}\ \langle \boldsymbol{P}(0)\cdot\boldsymbol{P}(t)\rangle\,. \tag{S8}$$

Application of the Wiener-Khinchin theorem (Eq. (S14) below) gives

$$\chi''(\nu) = \frac{\pi}{3L_tV\varepsilon_0 k_BT}\,\nu\ \left|\widetilde{\boldsymbol{P}}(\nu)\right|^2\,, \tag{S9}$$

where L_t is the length in time of $\boldsymbol{P}(t)$. Eq. (S9) is the Equation implemented in this work as it involves only a single Fourier transform of each Cartesian component of $\boldsymbol{P}(t)$ and otherwise simple array operations.

S4 Spectrum of Flexible Classical Water Model TIP4P/2005f

We also calculated spectra using other classical water models, including of a 5 ns trajectory with a 0.5 fs writeout frequency of TIP4P/2005f, a flexible 4-point water model. [8] The resulting spectrum, along with self and collective components, is shown in Figure S2. Here, smoothing of segment spectra was carried out by convolution with a Gaussian of $\sigma = 0.1$ THz and the resulting thinned spectrum was again smoothed by convolution with a Gaussian of $\sigma = 0.5$ THz. Below ∼40 THz, the spectrum is very similar to that for SPC/E, but shows peaks for both intramolecular modes at ∼50 and ∼100 THz. The OH-stretch mode at ∼100 THz lacks the collectivity of the DFT MD spectrum, which is mostly due to the interactions of lone-pair Wannier centers with donor hydrogens across HBs.

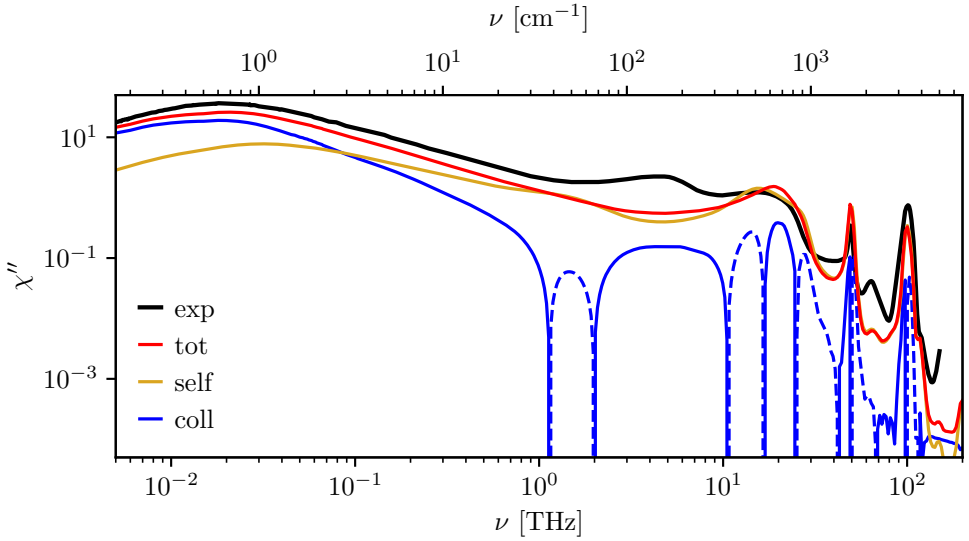

Figure S2: The TIP4P/2005f spectrum along with its self and collective components.

S5 Error Estimates

Spectra for all figures in this and the main work are calculated as the mean of ten smoothed spectra, from which the standard deviation is obtained. Error estimates are omitted from figures for clarity, excepting one

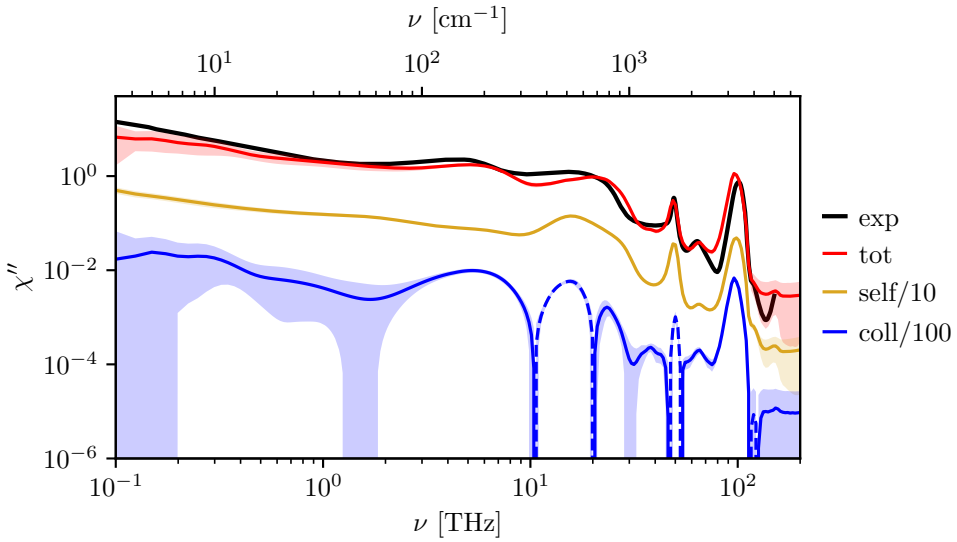

Figure S3: The DFT MD spectrum along with its self and collective components, shifted down by factors of 10 and 100 respectively, shown with their estimated standard deviations (shaded areas).

prototypical example, Figure S3, where they are shown for the DFT MD spectrum and its self and collective components along with their respective standard deviations, shown as shaded zones above and below the curves. Generally, as in Figure S3, standard deviations are small for self spectral components, and large for collective ones.

S6 Wiener-Khinchin Theorem

Assuming the available data of a signal $f(t)$ is limited to a finite time interval $[0,\, L_t]$, we formally define $f(t)$ as being zero outside this interval. We define the autocorrelation function of $f(t)$ as the mean over the interval I (of length $L_t - |t|$) where there is available data

$$C(t) = \frac{1}{L_t - |t|} \int_I \mathrm{d}t'\; f^*(t')\, f(t'+t)\,, \tag{S10}$$

where $f^*(t)$ is the complex conjugate of $f(t)$. If $f(t)$ is an observable in an equilibrium system, then by the ergodic theorem, $C(t)$ is the best estimate of the equilibrium ensemble average $\langle f^*(0) f(t)\rangle$. For $t \geq 0$, $I = [0,\, L_t - t]$, and for $t \leq 0$, $I = [|t|,\, L_t]$. The integrand $f^*(t')\, f(t'+t)$ is always zero for t' outside of I, so in both cases, the integration bounds can be extended arbitrarily to give a generally applicable expression,

$$C(t) = \frac{1}{L_t - |t|} \int_{-\infty}^{\infty} \mathrm{d}t'\; f^*(t')\, f(t'+t)\,. \tag{S11}$$

This step in the proof demands that in practice, the signal $f(t)$ be zero padded: that is, zeros of length L_t should be appended to the end of $f(t)$ before Fourier transforming. Substituting Eq. (S5) for $f^*(t')$ and $f(t'+t)$ gives

$$\begin{aligned}
C(t) &= \frac{1}{L_t - |t|} \int_{-\infty}^{\infty} \mathrm{d}t' \int_{-\infty}^{\infty} \mathrm{d}\nu\; e^{-2\pi i \nu t'}\, \tilde{f}^*(\nu) \int_{-\infty}^{\infty} \mathrm{d}\mu\; e^{2\pi i \mu (t'+t)}\, \tilde{f}(\mu)\,, \\
&= \frac{1}{L_t - |t|} \int_{-\infty}^{\infty} \mathrm{d}\nu\; \tilde{f}^*(\nu) \int_{-\infty}^{\infty} \mathrm{d}\mu\; e^{2\pi i \mu t}\, \tilde{f}(\mu) \int_{-\infty}^{\infty} \mathrm{d}t'\; e^{2\pi i t' (\mu - \nu)}\,, \\
&= \frac{1}{L_t - |t|} \int_{-\infty}^{\infty} \mathrm{d}\nu\; \tilde{f}^*(\nu) \int_{-\infty}^{\infty} \mathrm{d}\mu\; e^{2\pi i \mu t}\, \tilde{f}(\mu)\, \delta(\mu - \nu)\,, \\
&= \frac{1}{L_t - |t|} \int_{-\infty}^{\infty} \mathrm{d}\nu\; e^{2\pi i \nu t}\, \tilde{f}^*(\nu)\, \tilde{f}(\nu)\,.
\end{aligned} \tag{S12}$$

Rearranging and Fourier transforming both sides gives

$$\tilde{f}^*(\nu)\,\tilde{f}(\nu) = \int_{-\infty}^{\infty} \mathrm{d}t\; e^{-2\pi i\nu t}\,(L_t - |t|)\,C(t)\,. \tag{S13}$$

We assume correlations are large for small t and decay over time, so both terms of Eq. (S13) should be dominated by the small-$|t|$ regime. In this regime, in the limit of large L_t, $|t|/L_t \to 0$, so we may neglect the $|t|$ term. Additionally, for large L_t, $C(t) \to \langle f^*(0)f(t)\rangle$. Thus Eq. (S13) may be rewritten as the Wiener-Khinchin theorem

$$\int_{-\infty}^{\infty} \mathrm{d}t\; e^{-2\pi i\nu t}\,\langle f^*(0)f(t)\rangle = \frac{1}{L_t}\left|\tilde{f}(\nu)\right|^2 . \tag{S14}$$

S7 Acceptor Lone Pair and Donor Hydrogen Interactions

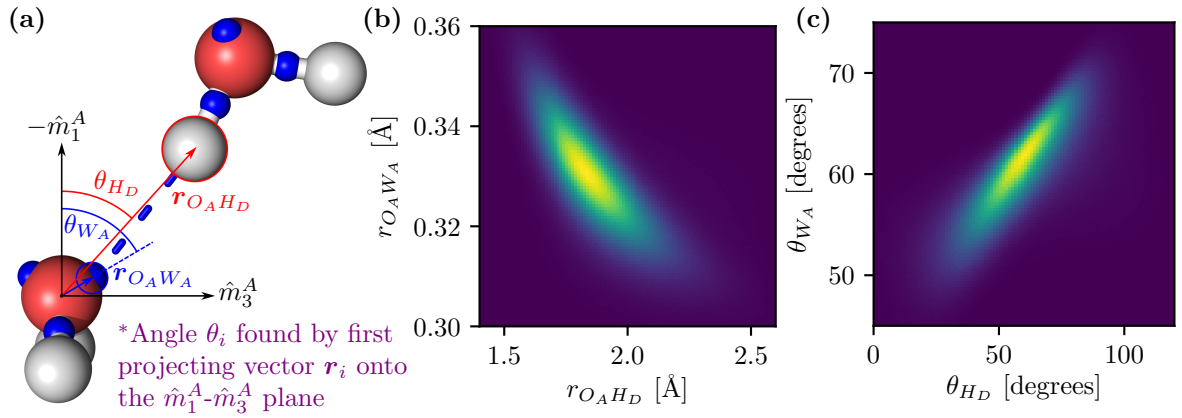

Figure S4: Positional dependence of acceptor Wannier centers on donor hydrogens. **(a)** Schematic illustrating vector and angle definitions. Wannier centers are represented as small blue spheres. Vectors $\boldsymbol{r}_{O_AH_D}$ and $\boldsymbol{r}_{O_AW_A}$ are the displacement vectors from acceptor oxygen (O_A) to donor hydrogen (H_D) and the acceptor Wannier center nearest to H_D (W_A). Angles θ_{H_D} and θ_{W_A} are between the acceptor's orientational axis ($-\hat{m}_1^A$) and the component of $\boldsymbol{r}_{O_AH_D}$ or $\boldsymbol{r}_{O_AW_A}$ in the $\hat{m}_1^A$-$\hat{m}_3^A$ plane. Thus θ_{H_D} and θ_{W_A} are independent of position along $\hat{m}_2^A$. **(b)** Bivariate histogram over distances $r_{O_AH_D}$ and $r_{O_AW_A}$, which reveals a pronounced negative dependence (yellow represents a high count, navy blue a low count). **(c)** Bivariate histogram over θ_{H_D} and θ_{W_A}, which reveals a pronounced positive linear dependence.

To aid in understanding the following, refer to the schematic Figure S4a, which shows a snapshot of a hydrogen-bonded dimer from our DFT MD simulation (small blue spheres indicate Wannier centers). We define two water molecules at a simulation timestep as hydrogen-bonded when they fulfill the geometrical Luzar criterion,[9] and denote the donor hydrogen as H_D and the acceptor oxygen as O_A. For our DFT MD trajectory, calculation of Wannier centers and assignment of each to the nearest oxygen consistently results in exactly four Wannier centers per water molecule, arranged around the oxygen in a roughly tetrahedral configuration, with two lying along the OH-bonds, and two on the back side of the oxygen where the lone pair electron density is high. For each hydrogen-bonded dimer in the DFT MD trajectory, we define the lone pair Wannier center of the acceptor molecule that is nearest the donor hydrogen as the "acceptor Wannier center", denoting it W_A.

We find that an acceptor Wannier center's position with respect to the acceptor oxygen $\boldsymbol{r}_{O_AW_A}$ tends to depend strongly on relative donor hydrogen position $\boldsymbol{r}_{O_AH_D}$, which we interpret as resulting from the Coulomb attraction between donor hydrogens and lone pair electrons. In this picture, lone pair electrons might be expected to stretch away from the parent oxygen as a donor hydrogen approaches. Indeed, Figure S4b agrees with this picture; it shows a bivariate joint histogram over the distances $r_{O_AH_D}$ and $r_{O_AW_A}$, which show a clear negative correlation. As $r_{O_AH_D}$ can be expected to fluctuate under molecular translations (HB stretching) and OH-stretching, this relationship has significant implications for the peaks at ∼5 and ∼100 THz.

We define the angle between the acceptor's orientational axis ($-\hat{m}_1^A$) and the components of $\boldsymbol{r}_{O_AH_D}$ or $\boldsymbol{r}_{O_AW_A}$ in the $\hat{m}_1^A$-$\hat{m}_3^A$ plane as θ_{H_D} and θ_{W_A} respectively. Figure S4c shows a bivariate histogram of θ_{H_D} and

θ_{W_A}, where there is a positive linear correlation. This indicates that lone pair electrons, attracted by donor hydrogens, track their movement about the acceptor oxygen. The angle θ_{H_D} can be expected to fluctuate under librations of the acceptor and/or donor molecules, and under bending of ∠HOH of the donor molecule, so this relationship has significant implications for the peaks at ∼20 and ∼50 THz.

S8 Further Studies of Librational Spectra

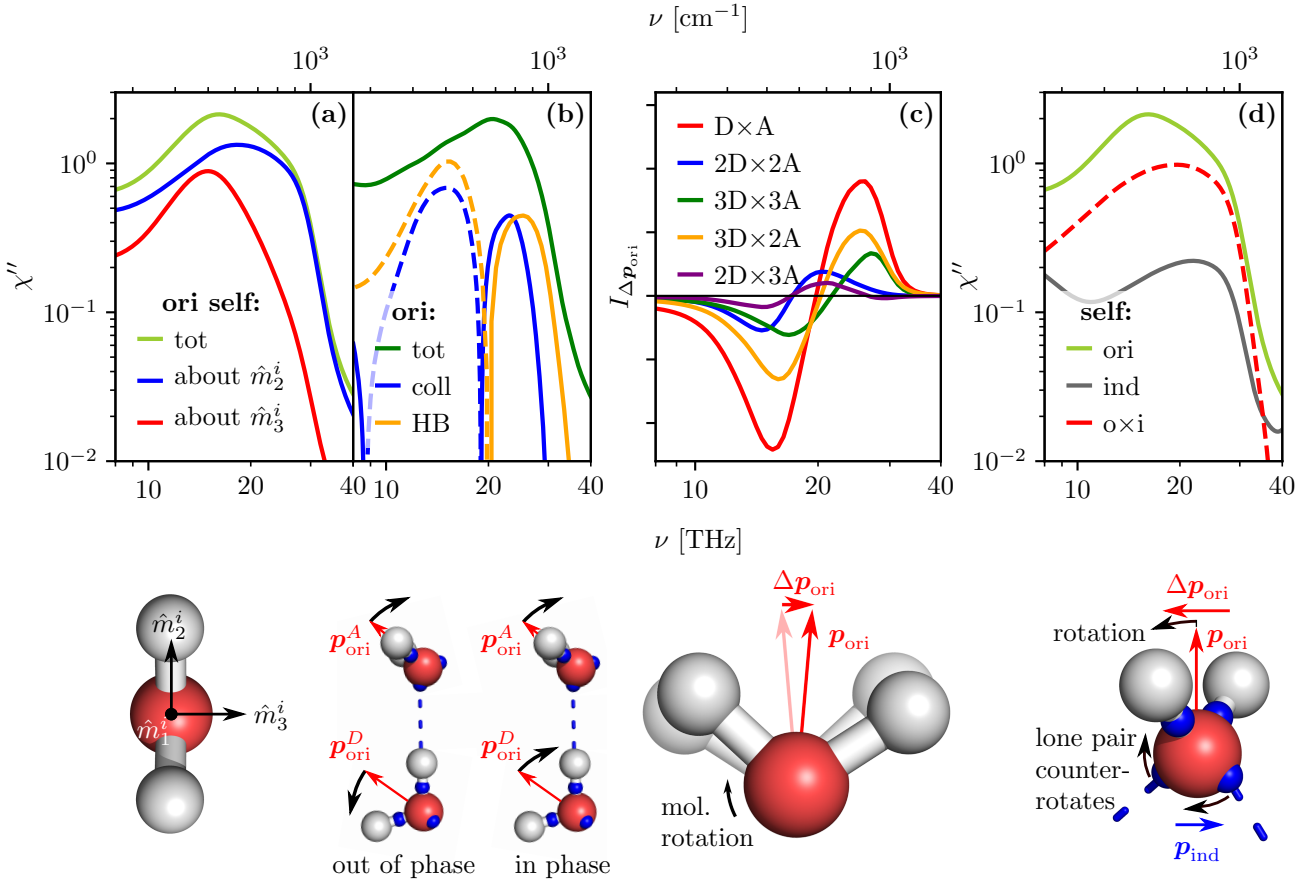

Figure S5: Components of the susceptibility and other power spectra in the libration regime. The schematic below each plot pertains to that plot. Small blue spheres in the schematics are Wannier center positions. **(a)** The orientational self component, along with the subcomponents due to librations about $\hat{m}_2$ and $\hat{m}_3$. **(b)** The orientational spectrum and two of its subcomponents: the collective component and the part of the collective component due only to cross correlations within hydrogen-bonded dimers, which is apparently the source of the in- and out-of-phase modes. **(c)** The spectral components due to cross correlations in the change in the orientational component $\Delta\boldsymbol{p}_{\text{ori}}$ over one simulation time step (see schematic) between hydrogen-bonded donors and acceptors. $\Delta\boldsymbol{p}_{\text{ori}}$ is decomposed into its $\hat{m}_2^i$ and $\hat{m}_3^i$ components for both the donor and the acceptor: the cross spectral components among these are shown in the plot. E.g. 2D×3A denotes the cross component due to $(\hat{m}_2^D \cdot \Delta\boldsymbol{p}_{\text{ori}}^D)\hat{m}_2^D \cdot (\hat{m}_3^A \cdot \Delta\boldsymbol{p}_{\text{ori}}^A)\hat{m}_3^A$. These correspond to librations about different molecular axes. It is clear from the plot that there are out-of-phase and in-phase modes at lower and higher frequencies for all combinations of libration axes. **(d)** Subcomponents of the self component: namely its orientational, induced, and orientational×induced cross components. The negative cross component implies that molecular dipole moments are induced opposite to librations, the mechanism shown in the schematic.

In a 1964 study of Raman spectra of water, normal mode analysis of intermolecular motions in a C_{2v}-symmetric tetrahedral water cluster predicted modes at roughly 13, 17 and 22 THz, due to librations about $\hat{m}_1$, $\hat{m}_3$, and $\hat{m}_2$ respectively.[10,11,12] These differences in frequency have been attributed to differences in the

moment of inertia of a water molecule about these three principle axes.[12] As the libration about the C_2-axis doesn't change the system dipole for the C_{2v}-symmetric system, only the 17 and 22 THz modes were predicted to be IR active. A 1995 study fit experimental IR absorption spectra in the libration regime with Gaussians, finding two modes at roughly 11.5 and 20 THz.[13] Figure S5a shows the spectral contributions due to librations about $\hat{m}_2$ and $\hat{m}_3$ from our DFT MD simulation. The librations about $\hat{m}_3$ peak sharply at ∼15 THz, and those about $\hat{m}_2$ peak at ∼18 THz with a shoulder at ∼25 THz. This is in relatively good agreement with the normal mode analysis in Ref.[10].

A 2008 study of IR spectra extracted from Carr-Parrinello simulations of water revealed features at roughly 11.5 and 18 THz in components of a self/collective decomposition, where the collective component was negative at 11.5 THz and positive at 18 THz.[14] These negative and positive modes appear in the collective components of DFT and forcefield MD as well at ∼15 and ∼25 THz respectively (see main text). Figure S5b plots the orientational component, its collective component, and finally a further subcomponent of this collective component: that due only to cross correlations between hydrogen-bonded molecules, which apparently accounts almost entirely for the collective behavior of the orientation there. Thus, we can rule out correlations among non-hydrogen-bonded molecules and conclude that it is in- and out-of-phase dimer modes that underlie these positive and negative collective features, as illustrated in the schematic below Figure S5b.

As the frequencies of these in- and out-of-phase dimer modes correspond roughly to those of the different libration directions, a naive explanation is that librations about $\hat{m}_2$ tend to correlate positively between hydrogen-bonded molecules, and those about $\hat{m}_3$, negatively. Figure S5c shows the power spectra (as defined in Eq. (3) of the main text) in lin-log, due to cross correlations between hydrogen-bonded molecules of $\Delta\boldsymbol{p}_{\mathrm{ori}}$, the *change* in the orientational component. Further, $\Delta\boldsymbol{p}_{\mathrm{ori}}$ is decomposed in each molecule into $\hat{m}_2$ and $\hat{m}_3$, (which correspond closely to librations about $\hat{m}_3$ and $\hat{m}_2$ respectively), among which cross power spectra are calculated and plotted. For example, the configuration shown in the schematic below Figure S5b is from correlations between $(\hat{m}_2^D \cdot \Delta\boldsymbol{p}_{\mathrm{ori}}^D)\hat{m}_2^D$ and $(\hat{m}_3^A \cdot \Delta\boldsymbol{p}_{\mathrm{ori}}^A)\hat{m}_3^A$, which is labeled 2D×3A in Figure S5c. It is clear from Figure S5c that for any combination of librational axes, there are negative and positive modes at lower and higher frequencies respectively, which rules out libration axes as an explanation of the frequency difference between the in-phase and out-of-phase libration modes. Our explanation is simply that the shared hydrogen bond is stretched more for in-phase librations, and therefore exerts a stronger restoring force, resulting in a higher frequency, as illustrated in Figures 4a and b in the main text.

S8.1 Lone-Pair Induced Molecular Dipole Moments Under Libration

Finally, Figure S5d shows all three components of the orientational/induced decomposition of the self component in the libration regime. The line shapes are broadly similar, though the cross component is negative, which indicates that in a librating molecule, a dipole moment is induced opposite to the libration. Figure S4c in Section S7 provides the explanation: under librations of an acceptor molecule, its lone-pair electrons track the relative motion of an attractive donor hydrogen, inducing a dipole moment, which is illustrated in the schematic below Figure S5d.

References

[1] Schwan, H. P.; Sheppard, R. J.; Grant, E. H. Complex permittivity of water at 25 °C. *J. Chem. Phys.* **1976**, *64*, 2257–2258.

[2] Barthel, J.; Bachhuber, K.; Buchner, R.; Hetzenauer, H.; Kleebauer, M. A Computer-controlled System of Transmission Lines for the Determination of the Complex Permittivity of Lossy Liquids between 8.5 and 90 GHz. *Ber. Bunsenges. Phys. Chem.* **1991**, *95*, 853–859.

[3] Kaatze, U. Complex permittivity of water as a function of frequency and temperature. *J. Chem. Eng. Data* **1989**, *34*, 371–374.

[4] Czumaj, Z. Absorption coefficient and refractive index measurements of water in the millimetre spectral range. *Mol. Phys.* **1990**, *69*, 787–790.

[5] Downing, H. D.; Williams, D. Optical constants of water in the infrared. *J. Geophys. Res.* **1975**, *80*, 1656–1661.

[6] Neumann, M. Dielectric relaxation in water. Computer simulations with the TIP4P potential. *J. Chem. Phys.* **1986**, *85*, 1567–1580.

[7] Ellison, W. Permittivity of Pure Water, at Standard Atmospheric Pressure, over the Frequency Range 0–25 THz and the Temperature Range 0–100 ^{o}C. *J. Phys. Chem. Ref. Data* **2007**, *36*, 1.

[8] González, M. A.; Abascal, J. L. F. A flexible model for water based on TIP4P/2005. *J. Chem. Phys.* **2011**, *135*, 224516.

[9] Luzar, A.; Chandler, D. Effect of Environment on Hydrogen Bond Dynamics in Liquid Water. *Phys. Rev. Lett.* **1996**, *76*, 928–931.

[10] Walrafen, G. E. Raman Spectral Studies of Water Structure. *J. Chem. Phys.* **1964**, *40*, 3249–3256.

[11] Walrafen, G. E. *The Physics and Physical Chemistry of Water*; Franks, Felix v. 1; Springer US, 1972; Vol. 1; Chapter 5, pp 151–214.

[12] Walrafen, G. E. Raman spectrum of water: transverse and longitudinal acoustic modes below ≈300 cm^{-1} and optic modes above ≈300 cm^{-1}. *J. Phys. Chem.* **1990**, *94*, 2237–2239.

[13] Zelsmann, H. R. Temperature dependence of the optical constants for liquid H2O and D2O in the far IR region. *J. Mol. Struct.* **1995**, *350*, 95 – 114.

[14] Chen, W.; Sharma, M.; Resta, R.; Galli, G.; Car, R. Role of dipolar correlations in the IR spectra of water and ice. *Phys. Rev. B* **2008**, *77*, 245114.