Neural Network-Based Sum-Frequency Generation Spectra of Pure and Acidified Water Interfaces with Air

Miguel de la Puente, Axel Gomez, Damien Laage

Neural Network-Based Sum Frequency Generation Spectra of Pure and Acidified Water Interfaces with Air

Miguel de la Puente, Axel Gomez, and Damien Laage*

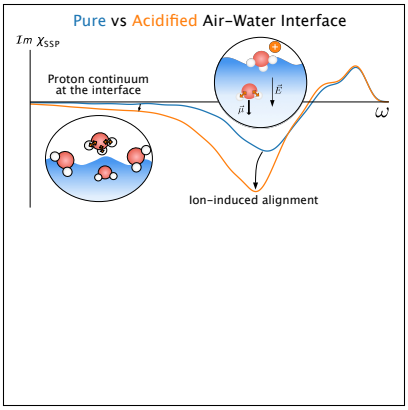
PASTEUR, Department of Chemistry, École Normale Supérieure, PSL University, Sorbonne Université, CNRS, 75005 Paris, France

E-mail: damien.laage@ens.psl.eu

Abstract

The affinity of hydronium ions (H_3O^+) for the air-water interface is a crucial question in environmental chemistry. While sum-frequency generation (SFG) spectroscopy has been instrumental to indicate the preference of H_3O^+ for the interface, key questions persist regarding the molecular origin of the SFG spectral changes in acidified water. Here we combine ns-long neural network (NN) reactive simulations of pure and acidified water slabs with NN predictions of molecular dipoles and polarizabilities to calculate SFG spectra of long reactive trajectories including proton transfer events. Our simulations show that H_3O^+ ions cause two distinct changes in phase-resolved SFG spectra: first, a low-frequency tail due to the vibrations of H_3O^+ and its first hydration shell, analogous to the bulk proton continuum, and second, an enhanced hydrogen-bonded band due to the ion-induced static field polarizing molecules in deeper layers. Our calculations confirm that changes in SFG spectra of acidic solutions are caused by hydronium ions preferentially residing at the interface.

TOC Graphic



Water interfaces exhibit a unique chemical reactivity that differs from both the bulk liquid and the gas phase.[1,2] Understanding this specific reactivity is critical in numerous fields, ranging from atmospheric chemistry[3] to microdroplet catalysis.[4] With acidity recognized as a key determinant of reactivity in aqueous phases, considerable experimental[5–15] and simulation[16–20] effort has been devoted to characterizing the acid-base properties of the air-water interface in recent decades.

Although no definitive answer to this question has been obtained yet, there is accumulating evidence of substantial surface activity of hydronium (H_3O^+) cations.[6–8,12–14,16,17,20] Among the experimental techniques which have supported this acid-enriched picture of the air-water interface, vibrational sum-frequency generation (SFG) spectroscopy has been instrumental because of its surface specificity. SFG measurements have revealed that the vibrational spectrum of the neat water surface changes markedly when the solution is acidified,[6–8,12–14] which has been suggested to arise from the presence of hydronium ions at the air-water interface.

However, interpreting SFG spectra can be complex and ambiguous, and a consensus on the molecular origin of these spectral changes has not yet been reached. A first difficulty arises from the nature of the SFG signal being collected. Spectral changes in acidic solutions have been found to be very different between experiments collecting the overall SFG intensity[6–8,11,12] and phase-resolved SFG measurements.[13,14] In addition, a non-resonant background contribution has also been proposed[11] to explain the observed changes instead of an interfacial effect. And finally, because these characterizations consider the difference between spectra, the need for a high signal-to-noise ratio[7] has also been emphasized.

A second question arises from the two types of effects that ions like hydronium can have on the SFG spectra. SFG measurements collect signal from all regions which are not centrosymmetric. At the interface between neat water and air, SFG thus selectively probes the outermost water molecular layer, where the isotropic character of the bulk liquid disappears.[21–23] However, when charged species like H_3O^+ are added at the interface, an

additional contribution to the SFG signal arises from deeper-lying water molecules, whose local environment is now no longer symmetric because of the static electric field generated by the surface charges. The effective resonant second order susceptibility $\chi_{eff}^{(2),res}$ measured in the SFG spectra becomes[21,24]

$$\chi_{eff}^{(2),res}(\omega) = \chi^{(2),res}(\omega) + \chi^{(3)}(\omega)\Delta\Phi \tag{1}$$

where $\chi^{(2),res}$ and $\chi^{(3)}$ are the intrinsic second-order susceptibility of the interface and the third-order susceptibility of bulk water, and $\Delta\Phi$ is the difference in electrostatic potential between the interface and the bulk liquid ($\chi_{eff}^{(2),res}$ includes an additional interference term,[25] which is ignored here because it is negligible for ionic strengths above 0.01 M, as will be considered in the following).

A key question is then to determine which term is responsible for the changes observed in SFG experiments on acidified interfaces, and whether the latter result from an acid enrichment of the air-water interface. Several explanations (sometimes in combination) have been proposed in the literature to explain the spectral changes: first the intrinsic vibrational response of hydronium ions lying at the interface[8,12,13] (*i.e.*, a $\chi^{(2),res}$ term), second a structural rearrangement of first layer water molecules induced by neighboring hydronium ions[7,8] (*i.e.*, another $\chi^{(2),res}$ contribution), third a long-range electrostatic effect which causes the water molecules further from the interface to partially align with the field produced by surface ions[7,8,14,26] (*i.e.*, a $\chi^{(3)}$ contribution), and finally a non-resonant background.[11] However, disentangling these different contributions has proven extremely challenging so far.

Molecular dynamics simulations appear as ideally suited to determine these different factors and to provide a molecular interpretation of the observed spectral changes at the surface of acidic aqueous solutions. In the case of neat water, simulations have already demonstrated that a quantitative description of the SFG spectrum can be obtained, at the air-water interface (see, *e.g.*, refs. 22,23,27–32) and at other aqueous interfaces including the

water-silica,[33,34] water-phospholipid[33,35] and water-graphene[36–38] interfaces.

However, SFG calculations of acidic aqueous solutions face two major difficulties due to the highly reactive character of hydronium ions, which exchange their excess proton with neighboring water molecules on a picosecond timescale.

The first one arises from the need to propagate trajectories long enough to obtain converged spectra. Pioneering simulations[39] of SFG spectra at the surface of acidified water have for example employed classical polarizable forcefields to describe the hydronium ion. While such non-reactive simulations have the advantage of their limited computational cost, they impose a fixed (*Eigen*-like) localized structure to the fluctuating excess proton which cannot be exchanged, and the impact on its vibrational properties at the air-water interface is not known. In contrast, ab initio molecular dynamics simulations are reactive and allow for proton exchanges,[13,26] but their substantial computational cost imposes important limitations on the trajectory length and thus on the convergence of the SFG spectra.

The second difficulty stems from the calculation of the molecular dipoles and polarizabilities that enter in the $\chi^{(2),res}$ susceptibility. In the non-reactive situation of neat water, several approaches have been successfully used to calculate the SFG spectrum: Morita and Hynes[22,23] related the water OH bond dipoles and polarizabilities to the molecular geometry, Auer and Skinner[27,28] designed an empirical map between the local electric field and the bond dipoles and polarizabilities, and Nagata and coworkers[29] used some features of this map to relate the dipole-polarizability correlation to the bond velocity autocorrelation function. However, a common feature of all these methods is that they require intact covalent bonds. Therefore they cannot describe the change in molecular dipole and polarizability when a hydronium ion is converted into a water molecule upon proton transfer. Another set of methods based on novel machine-learning approaches have recently reproduced the IR, Raman and SFG spectra of liquid water[40–43] but they have so far only been developed on non-reactive systems.

Here we solve both difficulties and address the challenge of simulating the SFG spectrum

of a reactive excess proton at the air-water interface by combining several neural network-based approaches. A first neural network is used to propagate a reactive molecular dynamics trajectory at the density-functional theory level. Two additional neural networks are constructed to obtain the dipoles and polarizabilities along this trajectory and a procedure is introduced to account for proton exchanges. We then show that our calculated SFG spectrum reproduces the features of the experimental spectra and we provide a molecular interpretation of the changes observed in acidic aqueous solutions, disentangling the different $\chi^{(2),res}$ and $\chi^{(3)}$ contributions.

We first introduce our simulation methodology. A series of 16 independent 1 ns-long reactive molecular dynamics simulations of a pure water slab and of a slab of water with one excess proton (at an effective concentration of approximately 0.3 mol/L, see SI) are propagated using a neural network potential (NNP). This NNP was constructed as part of our recent study[20] of water self-dissociation using the DeePMD-kit software[44,45] to reproduce the hybrid functional revPBE0-D3 level of theory. Since this network was trained to describe proton exchanges, our trajectory is reactive. We emphasize that this NNP allows simulating trajectories that are approximately one order of magnitude longer than in the most recent ab initio molecular dynamics studies, while retaining the same accuracy.

The resonant second order susceptibility $\chi^{(2),res}$ is expressed within the time-dependent representation of Morita and Hynes[22,23] as

$$\chi^{(2),res}_{pqr}(\omega) = \frac{i\omega Q(\omega)}{k_B T} \int_0^\infty dt\ e^{i\omega t} \left\langle \sum_{i,j} \alpha_{pq,j}(t)\mu_{r,i}(0) \right\rangle \tag{2}$$

where k_B is the Boltzmann constant, T is the temperature, $Q(\omega) = (\hbar\omega/k_B T)\,/\,[1 - \exp(-\hbar\omega/k_B T)]$ is the harmonic quantum correction factor,[46] $\mu_{r,i}$ is the r component of the dipole moment of molecule i and $\alpha_{pq,j}$ is the pq component of the polarizability tensor of molecule j.

We determine the molecular dipoles and polarizabilities using the maximally localized Wannier functions,[47] whose charge centers provide the locations of electronic pairs. The

modern theory of bulk polarization[48] provides a natural connection between the total dipole of the system and the vector sum of the Wannier centers and of the nuclei.[47] Wannier centers have already been used in combination with DFT-based molecular dynamics simulations to calculate the molecular dipole of water[49] and SFG spectra.[31,50] However, for SFG spectroscopy, the computational cost of determining the dipole and polarizability via DFT at every step along the trajectory is prohibitively expensive, limiting the accuracy of the calculated spectrum.[50]

We therefore draw on recent work[40,41] where a NN was trained to predict the location $\mathbf{w}$ of the vector sum of Wannier centers, *i.e.*, the Wannier centroid, relative to each atom. In the case of a water molecule, each oxygen atom is uniquely associated with four maximally localized Wannier centers, grouped into a Wannier centroid which thus represents the effective center of negative charge for the molecule. Since the centers of positive charge are simply the oxygen and hydrogen nuclear positions $\mathbf{r}_{O,H}$, the dipole of molecule i is then expressed as

$$\mu_i(t) = 6e\ \mathbf{r}_{O_i}(t) + e\sum_{j\in i}\mathbf{r}_{H_i^j}(t) - 8e\ \mathbf{w}_i(t) \tag{3}$$

where the sum runs over the H_i^j hydrogen atoms covalently bonded to the i-th oxygen atom O_i. The derivatives of $\mathbf{w}_i$ with respect to electric fields applied along each axis (computed via finite differences) provide the effective molecular polarizabilities $\alpha_{\mathbf{i}}$, and are learned by a separate NN. This method was successfully applied to the calculation of the bulk water IR[40] and Raman[41] spectra, and we extend it here to obtain the SFG spectra (we note that a related NN approach has also recently been used to calculate the IR spectrum of bulk water[51]). Details on the data sets and training procedures employed to obtain the dipole and polarizability NN are provided in the SI.

For the hydrated excess proton, an additional difficulty arises from the discontinuities caused by proton transfer events. Covalent bond breaking and making events abruptly change the identity of the oxygen atom carrying the excess proton and induce discontinuities

in the molecular dipoles of the proton donating and accepting molecules. While experimentally the excess proton charge is compensated by a counterion, we focus here on the effect of a single excess proton and our system carries a net charge (this approximation is justified by the absence of counterion effect on experimental SFG spectra in dilute conditions[14]). The dipole moment therefore depends on the chosen charge origin. In non-reactive situations one typically sets the origin on the center of excess charge, and this choice does not affect the spectrum[52] since its very slow motion does not contribute to the frequency range of interest. However, this is not the case here because of the fast proton transfer events. We solve this difficulty by drawing on ideas from the modern theory of polarization, which stressed that in periodic systems changes in polarization are well defined although the absolute polarization is not. We therefore add an *a posteriori* time-dependent analytic correction to the NN molecular dipoles to ensure that they are continuous, and that their derivatives are equal to their uncorrected values between proton transfers. This amounts to keeping the center of charge on the original charge carrier (see details in SI). We finally calculate the susceptibility from the time derivatives of the corrected dipoles and polarizabilities as

$$\chi^{(2),res}_{pqr}(\omega) = \frac{iQ(\omega)}{\omega k_B T} \int_0^\infty dt\ e^{i\omega t} \left\langle \sum_{i,j} \dot{\alpha}_{pq,j}(t) \dot{\mu}_{r,i}(0) \right\rangle \tag{4}$$

Our simulations treat all nuclear motions classically. Although nuclear quantum effects have a substantial impact on the vibrational spectrum of water, previous studies have shown that classical simulations of water yield the same spectral features as with quantum nuclei[53] but with a shift of vibrational bands. To facilitate the comparison with experimental results, our calculated spectra therefore include an additional frequency axis whose frequencies are scaled by a factor of 0.96 to approximate the impact of nuclear quantum effects in the high frequency region, as done in previous calculations.[29]

We now present our calculated phase-resolved SFG spectra at the interface of neat water and acidified water with air in Fig. 1. Pioneering SFG studies[7,8] and several subsequent

works[12,13] have measured the *intensity* of the SFG signal for the interfaces of pure and acidified water, but interpreting these changes is difficult because the intensity results from both the real and imaginary parts of the response function. In our study, we therefore focus on the imaginary part of the signal which can be more readily connected to a molecular interpretation. Regarding the frequency range to be explored, all recent phase-resolved measurements focus on frequencies larger than 3,000 cm^{-1} but because the "proton continuum" is known[54] to extend to lower frequencies in the bulk and because SFG intensity measurements[8,11] have reported a change upon acidification at low frequencies, we consider the 2,700-4,000 cm^{-1} frequency window.

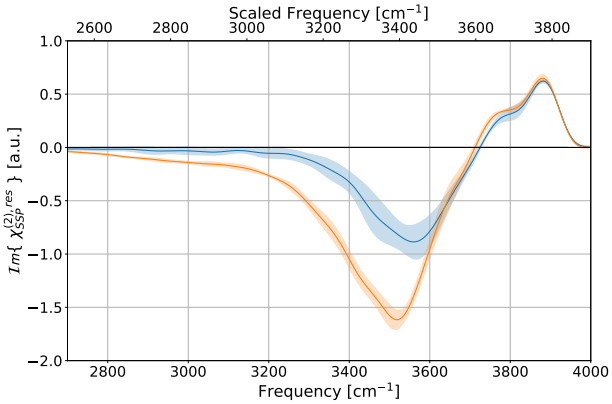


Figure 1: Imaginary part of the SFG response function with the SSP polarization combination for pure (blue) and acidified (orange) water slabs, normalized by the slab surface area. Colored shaded areas indicate the 95% confidence intervals.

Figure 1 shows that our calculations reproduce all the well-known features present in the air-water SFG spectrum, which supports our simulation methodology (to facilitate comparisons with experimental spectra, our following discussions on calculated spectra will employ the scaled frequency). These include the positive peak at high frequencies ($\approx$ 3,750 cm^{-1}) associated with dangling OH groups pointing towards the vapor phase, the hump close to 3,700 cm^{-1} coming from intra- and inter-molecular vibrational couplings[55] (see SI for a detailed analysis) and the negative band in the "bonded" region due to hydrogen-bonded OH groups pointing towards the liquid phase ($\approx$ 3,000$-$3,600 cm^{-1}). As shown in Fig. S10, our simulated spectra are in excellent agreement with experiments for the hydrogen-bonded

band, but slightly overestimate the dangling OH band width as in other previous simulations.[27,28,30,32,39,56] However, this in no way hinders our aim of determining the molecular factors behind the changes in SFG spectra upon acidification.

We now compare the SFG spectra at the air interface of neat and acidified water, focussing on three frequency regions. First, in the high frequency domain ($\approx$ 3,700$-$3,800 cm^{-1}) corresponding to the dangling OH groups, our spectra do not reveal any substantial change upon acidification. Early SFG intensity measurements[7,8] had reported a decrease in the amplitude of the free OH peak but its magnitude was found to greatly depend on the experimental signal-to-noise ratio; more importantly these measurements considered the intensity of the SFG signal, and subsequent studies have shown[14,26,39] that this intensity decrease is caused by a change in the real part of the response function. This is therefore consistent with the absence of any change in the imaginary part of our calculated signal. Second, in the bonded region (3,300$-$3,600 cm^{-1}), the amplitude of the negative band is enhanced in the acidified solution, in good agreement with previous measurements[13,14] (see Fig. S10) and simulations.[26] Finally, at low frequencies (below $\approx$ 3,000 cm^{-1}), our calculations reveal a negative tail in the acidified solution. This result is consistent with low-frequency measurements of the SFG intensity,[8] but to the best of our knowledge, subsequent phase-resolved measurements and simulations do not seem to have considered this frequency region, and recent SFG intensity measurements have attributed it to a non-resonant contribution.[11] Although the non-resonant contribution cannot be determined from simulations, the presence of this tail in our calculated *resonant* response function suggests instead that this spectral feature is associated (at least in part) to the interfacial presence of H_3O^+ and our analysis below will determine its molecular origin. We have complemented this spectral comparison with a structural analysis of the two interfaces in the SI.

In order to interpret these spectra and to determine the molecular origin of the spectral changes observed in acidic solutions, we decompose the simulated signals into their contributions from molecules in successive interfacial layers (see Fig. 2). For neat water, it is well

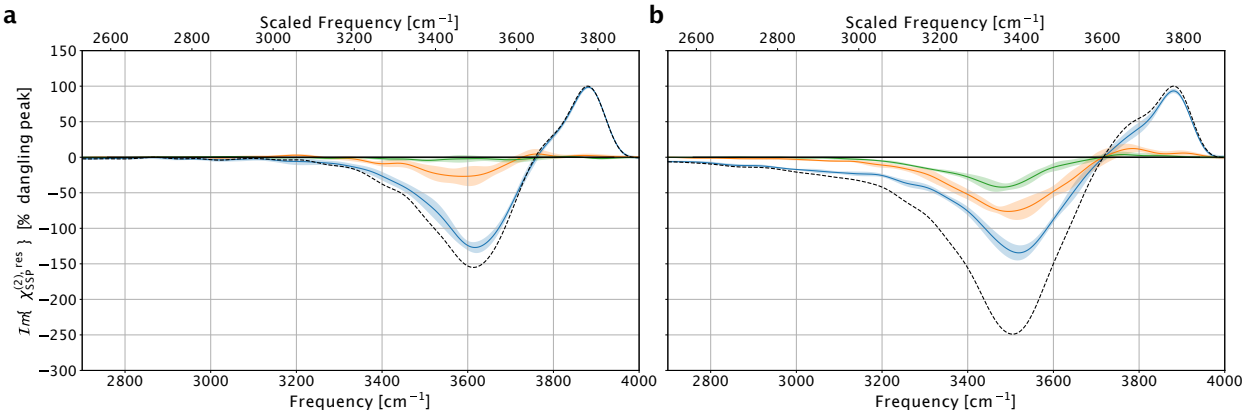


Figure 2: Imaginary part of the SSP SFG response function for a) pure and b) acidified water slabs decomposed into contributions from molecules lying in the first (blue), second (orange) and third (green) interfacial layers. Black dashed lines represent the total signals. Signals are normalized with respect to the maximal amplitude of the dangling OH peak. The first, second, and third layers are defined to include water molecules whose oxygen atom lies respectively above and less than 3 Å below the Gibbs dividing surface (GDS), between 3 Å and 6 Å below the GDS and between 6 Å and 9 Å below the GDS (with 9 Å below the GDS approximately matching the middle of our simulation slab).

known[22,57] that the SFG signal majoritarily comes from molecules lying in the first interfacial layer, with a minor contribution from molecules in deeper layers and a negligible signal below ≈ 5-6 Å, and Fig. 2a shows that our simulations support this picture. We first examine the negative hydrogen-bonded band. In the acidified solution, the contribution of the first layer is only marginally enhanced compared to that in pure water (less than 5% units of the dangling peak intensity used as a reference, see Fig. 2b). This observation is corroborated by our structural analysis, showing that within the first layer only the few water molecules in the hydronium cation's first hydration shell experience a downward reorientation, and this rearrangement remains limited (see Fig. S4). The increased amplitude of the total signal in the hydrogen-bonded band is shown in Fig. 2b to arise from the increased contributions of the second and third molecular layers, by ≈ -50 % compared to the pure water situation . Our structural analysis shows that this originates from a change in the average orientation of these water molecules (Fig. S5) due to the electric field generated by the cation lying at the interface (Fig. S8), leading to the previously suggested[14,26] $\chi^{(3)}$ effect.

We note that the depth over which this static field extends depends on the solution screening length, and thus on the ionic strength and on the nature of the counterion. Therefore, while our calculations are expected to provide a good description of the $\chi^{(3)}(\omega)$ lineshape, the amplitude of this contribution should be considered an estimation for two reasons. First, since this term arises from the static field induced by the surface charges, the simulated slab (half-) thickness should significantly exceed the Debye screening length, which is approximately 5 Å at the 0.3 mol/L concentration considered here; thus a thicker slab would be necessary. Second, counterions should be included, whereas our simulations focus solely on a single hydronium ion. Despite these limitations, the good agreement that is demonstrated between our simulated spectra and the experimental measurements (Fig. S10) suggests that our simulations capture the key effects.

We then turn to the low frequency tail present in the acidified interface signal. Figure 2b shows that it originates from molecules lying in the first interfacial layer, and not from a bulk contribution. This therefore rules out a solely electrostatic origin for the vibrational impact of excess protons at the air-water interface.

We now determine whether this signal is directly caused by the vibrations of H_3O^+ or by a rearrangement of water molecules in the topmost interfacial layer. To this end, we decompose the full SFG signal into contributions from three different kinds of molecules: the hydronium cation, its first hydration shell molecules, and the rest of the slab molecules. The result in Fig. 3a clearly shows that the low-frequency tail directly comes from the excess proton and its solvating water molecules, with no contribution from the rest of the slab. This broad red-shifted OH stretch band can thus be interpreted as part of the interfacial equivalent of the "proton continuum" characterized in the bulk via IR and Raman measurements.[8,54]

A previous study[13] has proposed that the vibrations of the proton are responsible for all the changes measured in the imaginary part of $\chi^{(2),res}$ upon acidification, extending up to frequencies corresponding to the hydrogen-bonded band, thus much larger than typically observed for the proton continuum in bulk acidified solutions (≈3,400-3,500 cm^{-1} at the

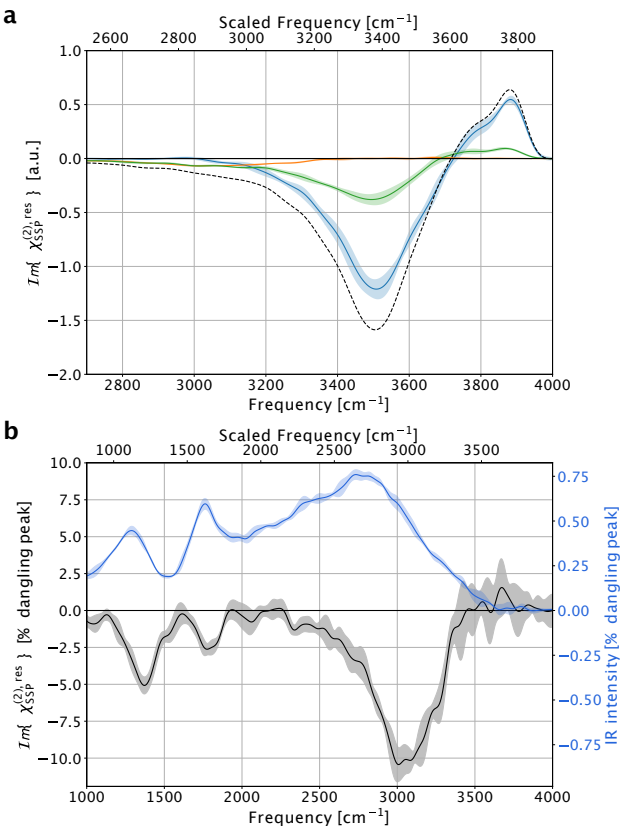


Figure 3: a) Imaginary part of the SSP SFG response function for the acidified water slab decomposed into contributions coming from the hydronium cation (orange), its first hydration shell (green), and the other water molecules in the system (blue). Signals are normalized by the slab surface area and black dashes represent the total signal. b) Contribution of the hydronium cation to the total SFG (black) and bulk IR (blue) signals normalized by the dangling OH peak amplitude.

interface vs ≈ 3,000 cm^{-1} in bulk). To understand how the hydronium vibrations change between the bulk and the air-water interface, we isolate the hydronium contribution to the imaginary part of $\chi^{(2),res}$ and compare it to the hydronium contribution to the bulk IR spectrum, obtained with the same NN-based methodology[40] from an acidified aqueous solution (one excess proton in 192 water molecules in bulk conditions, see SI for details).

The results in Fig. 3b show that the bulk H_3O^+ vibrational spectrum exhibits three well-known[54] features, namely the "proton transfer mode" (≈ 1,100−1,200 cm^{-1}), the bending peak (≈ 1,700 cm^{-1}) and the broad stretching band (≈ 2,000−3,000 cm^{-1}) known as the proton "continuum". The same three features are present in the SFG spectrum of H_3O^+ at

the interface, but with a negative amplitude due to the downward (ie. toward the liquid) orientation of the H_3O^+ molecular dipole, consistent with the literature[39] and our own structural analyses (Fig. S3). The relative contribution of H_3O^+ to the total SFG signal is more than an order of magnitude larger than in the IR signal, due to the larger effective concentration of protons at the interface compared to that in the bulk (see Fig. S1). But while the signs and amplitudes of the SFG and IR spectra allow only a qualitative comparison, contrasting the frequencies of the different H_3O^+ bands in the two environments is very instructive. When going from the bulk to the air-water interface, we observe a moderate blue-shift of the proton transfer (≈ 100 cm^{-1}) and stretching (≈ 200 cm^{-1} for the maximum amplitude) modes, and a narrowing of the latter band compared to the bulk, with a signal that rapidly vanishes above $\approx 3{,}300$ cm^{-1}. To understand the origin of this blueshift, we have analyzed the distribution of oxygen-oxygen distances (see Fig. S2) between the hydronium ion and its hydrogen-bond acceptors. Both the average oxygen-oxygen distance and its fluctuations are identical (within numerical errors) between the bulk and the interface. This strongly suggests that the observed blueshift of the H_3O^+ stretching mode at the interface does not originate from an elongation of its hydrogen bonds, but from the reduced polarity of the interfacial environment, which disfavors charge separation, increases the energetic cost of the OH bond elongation and thus the stretch frequency. Finally, the question of whether the excess proton at the air-water interface adopts the Eigen or Zundel structure has attracted some attention in the literature: our calculated SFG spectra for each form (Fig. S13) suggest that these two structures cannot be easily distinguished spectrally.

Our spectral analysis thus unambiguously shows that the increased intensity of the phase-resolved SFG signal observed at acidified air-water interfaces around (and below) $\approx 3{,}000$ cm^{-1} is directly due to the OH vibrations of the hydronium cation and its first solvation shell. An important implication of our results is therefore that this spectral region provides a clear signature of the presence of hydronium ions at the interface. In addition, because of the large relative contribution of the hydronium ion to the total SFG signal, the low-frequency proton

transfer mode which has been quite elusive in the bulk should be detectable experimentally with low-frequency phase-resolved SFG. This should also apply to protons at more complex aqueous interfaces, including, e.g. the phospholipid/water interface, and our results provide a molecular interpretation of recent experimental SFG spectra[58] on such systems, which have shown an enhanced intensity in the 2,500-2,800 cm^{-1} frequency range in the presence of acids.

Finally, while our simulations have focused on a single hydronium ion with no counterion, our calculated SFG spectra effectively capture the key changes observed upon acidification. Experiments show that counterions have a significant impact only at high concentration and for surface-active ions. Our results suggest that they can affect two spectral regions: first, the hydrogen-bonded band, by partially screening the hydronium-induced static field and changing the $\chi^{(3)}$ term amplitude, and second, the dangling OH peak by displacing some surface water molecules and reducing the peak amplitude. However, a more limited effect is expected in the low-frequency region below 3,000 cm^{-1} where the hydronium vibrational signature is present.

Our study therefore demonstrates that integrating state-of-the-art NN-based reactive MD simulations with NN determinations of molecular dipoles and polarizabilities offers a powerful method for calculating SFG spectra of reactive systems. Our application to the paradigm situation of an excess proton at the air-water interface yields an accurate, statistically converged characterization of the spectroscopic signature for this challenging system. Our results show that this elusive species brings two spectrally distinct contributions to the SFG vibrational spectrum: first, a static field $\chi^{(3)}$ effect[7,8,14,26] expanding the interfacial region probed by SFG, and second a unique broad low-frequency vibrational signature akin to that observed in bulk acidic aqueous solutions, characterized here for the first time. The strong agreement between our simulated spectra and experimental measurements, along with the identification of the molecular factors responsible for the spectroscopic features of acidified solutions, strongly supports the preferential adsorption of excess hydronium cations at

the air-water interface. Further extensions of this study will explore the mechanism of the impact of counterions and of acid concentration on the interfacial affinity of H_3O^+ and on the SFG spectra, as measured experimentally.[8,14] These issues will be addressed in a future work building on the methodology presented here.

Acknowledgement

The authors thank Dr Hilton B. de Aguiar (ENS, Paris) and Dr Simone Pezzotti (ENS, Paris) for stimulating discussions. This work has been supported through CDSN Ph.D. fellowships to M. P. and A. G. from the French Ministry of Higher Education and Research and by an HPC allocation from GENCI–IDRIS (Grant A0130707156).

Supporting Information Available

simulation details, structural analysis, NN architecture, training and validation, time-dependent correction of the molecular dipoles, calculation of the SFG response functions, additional SFG and IR spectra, electric field.

Data Availability Statement: the NNs used in this work are available from the authors upon reasonable request.

References

(1) Narayan, S.; Muldoon, J.; Finn, M. G.; Fokin, V. V.; Kolb, H. C.; Sharpless, K. B. "On water": unique reactivity of organic compounds in aqueous suspension. *Angew. Chem. Int. Ed.* **2005**, *44*, 3275–3279.

(2) Ruiz-Lopez, M. F.; Francisco, J. S.; Martins-Costa, M. T.; Anglada, J. M. Molecular reactions at aqueous interfaces. *Nat. Rev. Chem.* **2020**, *4*, 459–475.

(3) Finlayson-Pitts, B. J. Reactions at surfaces in the atmosphere: integration of experiments and theory as necessary (but not necessarily sufficient) for predicting the physical chemistry of aerosols. *Phys. Chem. Chem. Phys.* **2009**, *11*, 7760–7779.

(4) Yan, X.; Bain, R. M.; Cooks, R. G. Organic reactions in microdroplets: reaction acceleration revealed by mass spectrometry. *Angew. Chem. Int. Ed.* **2016**, *55*, 12960–12972.

(5) Radüge, C.; Pflumio, V.; Shen, Y. Surface vibrational spectroscopy of sulfuric acid-water mixtures at the liquid-vapor interface. *Chem. Phys. Lett.* **1997**, *274*, 140–144.

(6) Mucha, M.; Frigato, T.; Levering, L. M.; Allen, H. C.; Tobias, D. J.; Dang, L. X.; Jungwirth, P. Unified molecular picture of the surfaces of aqueous acid, base, and salt solutions. *J. Phys. Chem. B* **2005**, *109*, 7617–7623.

(7) Tarbuck, T. L.; Ota, S. T.; Richmond, G. L. Spectroscopic studies of solvated hydrogen and hydroxide ions at aqueous surfaces. *J. Am. Chem. Soc.* **2006**, *128*, 14519–14527.

(8) Levering, L. M.; Sierra-Hernández, M. R.; Allen, H. C. Observation of hydronium ions at the air-aqueous acid interface: vibrational spectroscopic studies of aqueous HCl, HBr, and HI. *J. Phys. Chem. C* **2007**, *111*, 8814–8826.

(9) Enami, S.; Stewart, L. A.; Hoffmann, M. R.; Colussi, A. J. Superacid chemistry on mildly acidic water. *J. Phys. Chem. Lett.* **2010**, *1*, 3488–3493.

(10) Yamaguchi, S.; Kundu, A.; Sen, P.; Tahara, T. Communication: Quantitative estimate of the water surface pH using heterodyne-detected electronic sum frequency generation. *J. Chem. Phys.* **2012**, *137*, 151101.

(11) Sengupta, S.; Moberg, D. R.; Paesani, F.; Tyrode, E. Neat water-vapor interface: proton continuum and the nonresonant background. *J. Phys. Chem. Lett.* **2018**, *9*, 6744–6749.

(12) Das, S.; Bonn, M.; Backus, E. H. G. The surface activity of the hydrated proton is

substantially higher than that of the hydroxide ion. *Angew. Chem. Int. Ed.* **2019**, *58*, 15636–15639.

(13) Das, S.; Imoto, S.; Sun, S.; Nagata, Y.; Backus, E. H. G.; Bonn, M. Nature of excess hydrated proton at the water-air interface. *J. Am. Chem. Soc.* **2020**, *142*, 945–952.

(14) Chiang, K.-Y.; Dalstein, L.; Wen, Y.-C. Affinity of hydrated protons at intrinsic water/vapor interface revealed by ion-induced water alignment. *J. Phys. Chem. Lett.* **2020**, *11*, 696–701.

(15) Colussi, A. J.; Enami, S.; Ishizuka, S. Hydronium ion acidity above and below the interface of aqueous microdroplets. *ACS Earth Space Chem.* **2021**, *5*, 2341–2346.

(16) Buch, V.; Milet, A.; Vácha, R.; Jungwirth, P.; Devlin, J. P. Water surface is acidic. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 7342–7347.

(17) Kumar, R.; Knight, C.; A. Voth, G. Exploring the behaviour of the hydrated excess proton at hydrophobic interfaces. *Faraday Discuss.* **2013**, *167*, 263–278.

(18) Wick, C. D. Comparing hydroxide and hydronium at the instantaneous air-water interface using polarizable multi-state empirical valence bond models. *Comput. Theor. Chem.* **2017**, *1116*, 64–72.

(19) de la Puente, M.; David, R.; Gomez, A.; Laage, D. Acids at the edge: why nitric and formic acid dissociations at air-water interfaces depend on depth and on interface specific area. *J. Am. Chem. Soc.* **2022**, *144*, 10524–10529.

(20) de la Puente, M.; Laage, D. How the acidity of water droplets and films is controlled by the air-water interface. *J. Am. Chem. Soc.* **2023**, *145*, 25186–25194.

(21) Eisenthal, K. B. Liquid interfaces probed by second-harmonic and sum-frequency spectroscopy. *Chem. Rev.* **1996**, *96*, 1343–1360.

(22) Morita, A.; Hynes, J. T. A theoretical analysis of the sum frequency generation spectrum of the water surface. II. time-dependent approach. *J. Phys. Chem. B* **2002**, *106*, 673–685.

(23) Morita, A. *Theory of Sum Frequency Generation Spectroscopy*; Springer: Singapore, 2018.

(24) Wen, Y.-C.; Zha, S.; Liu, X.; Yang, S.; Guo, P.; Shi, G.; Fang, H.; Shen, Y. R.; Tian, C. Unveiling microscopic structures of charged water interfaces by surface-specific vibrational spectroscopy. *Phys. Rev. Lett.* **2016**, *116*, 016101.

(25) Gonella, G.; Lütgebaucks, C.; de Beer, A. G. F.; Roke, S. Second harmonic and sum-frequency generation from aqueous interfaces is modulated by interference. *J. Phys. Chem. C* **2016**, *120*, 9165–9173.

(26) Pezzotti, S.; Gaigeot, M.-P. Spectroscopic BIL-SFG invariance hides the chaotropic effect of protons at the air-water interface. *Atmosphere* **2018**, *9*, 396.

(27) Auer, B. M.; Skinner, J. L. Vibrational sum-frequency spectroscopy of the liquid/vapor interface for dilute hod in D_2O. *J. Chem. Phys.* **2008**, *129*, 214705.

(28) Auer, B. M.; Skinner, J. L. Vibrational sum-frequency spectroscopy of the water liquid/vapor interface. *J. Phys. Chem. B* **2009**, *113*, 4125–4130.

(29) Ohto, T.; Usui, K.; Hasegawa, T.; Bonn, M.; Nagata, Y. Toward *ab initio* molecular dynamics modeling for sum-frequency generation spectra; an efficient algorithm based on surface-specific velocity-velocity correlation function. *J. Chem. Phys.* **2015**, *143*, 124702.

(30) Medders, G. R.; Paesani, F. Dissecting the molecular structure of the air/Water interface from quantum simulations of the sum-frequency generation spectrum. *J. Am. Chem. Soc.* **2016**, *138*, 3912–3919.

(31) Liang, C.; Jeon, J.; Cho, M. Ab initio modeling of the vibrational sum-frequency generation spectrum of interfacial water. *J. Phys. Chem. Lett.* **2019**, *10*, 1153–1158.

(32) Yamaguchi, S.; Takayama, T.; Goto, Y.; Otosu, T.; Yagasaki, T. Experimental and theoretical heterodyne-detected sum frequency generation spectroscopy of isotopically pure and diluted water surfaces. *J. Phys. Chem. Lett.* **2022**, 9649–9653.

(33) Nihonyanagi, S.; Yamaguchi, S.; Tahara, T. Ultrafast dynamics at water interfaces studied by vibrational sum frequency generation spectroscopy. *Chem. Rev.* **2017**, *117*, 10665–10693.

(34) Bañuelos, J.; Borguet, E.; Brown, G.; Cygan, R.; DeYoreo, J.; Dove, P.; Gaigeot, M.; Geiger, F.; Gibbs, J.; Grassian, V. et al. Oxide- and silicate-water interfaces and their roles in technology and the environment. *Chem. Rev.* **2023**, *123*, 6413–6544.

(35) Roy, S.; Gruenbaum, S.; Skinner, J. Theoretical vibrational sum-frequency generation spectroscopy of water near lipid and surfactant monolayer interfaces. *J. Chem. Phys.* **2014**, *141*, 18C502.

(36) Ohto, T.; Tada, H.; Nagata, Y. Structure and dynamics of water at water-graphene and water-hexagonal boron-nitride sheet interfaces revealed by ab initio sum-frequency generation spectroscopy. *Phys. Chem. Chem. Phys.* **2018**, *20*, 12979–12985.

(37) Zhang, Y.; de Aguiar, H. B.; Hynes, J. T.; Laage, D. Water structure, dynamics, and sum-frequency generation spectra at electrified graphene interfaces. *J. Phys. Chem. Lett.* **2020**, *11*, 624–631.

(38) Olivieri, J.-F.; Hynes, J. T.; Laage, D. Water dynamics and sum-frequency generation spectra at electrode/aqueous electrolyte interfaces. *Faraday Discuss.* **2024**, *249*, 289–302.

(39) Ishiyama, T.; Morita, A. Molecular dynamics analysis of interfacial structures and sum frequency generation spectra of aqueous hydrogen halide solutions. *J. Phys. Chem. A* **2007**, *111*, 9277–9285.

(40) Zhang, L.; Chen, M.; Wu, X.; Wang, H.; E, W.; Car, R. Deep neural network for the dielectric response of insulators. *Phys. Rev. B* **2020**, *102*, 041121.

(41) Sommers, G. M.; Calegari Andrade, M. F.; Zhang, L.; Wang, H.; Car, R. Raman spectrum and polarizability of liquid water from deep neural networks. *Phys. Chem. Chem. Phys.* **2020**, *22*, 10592–10602.

(42) Shepherd, S.; Lan, J.; Wilkins, D. M.; Kapil, V. Efficient quantum vibrational spectroscopy of water with high-order path integrals: from bulk to interfaces. *J. Phys. Chem. Lett.* **2021**, *12*, 9108–9114.

(43) Litman, Y.; Lan, J.; Nagata, Y.; Wilkins, D. M. Fully first-principles surface spectroscopy with machine learning. *J. Phys. Chem. Lett.* **2023**, *14*, 8175–8182.

(44) Wang, H.; Zhang, L.; Han, J.; E, W. DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics. *Comput. Phys. Commun.* **2018**, *228*, 178–184.

(45) Zeng, J.; Zhang, D.; Lu, D.; Mo, P.; Li, Z.; Chen, Y.; Rynik, M.; Huang, L.; Li, Z.; Shi, S. et al. DeePMD-kit v2: A software package for deep potential models. *J. Chem. Phys.* **2023**, *159*, 054801.

(46) Bader, J. S.; Berne, B. J. Quantum and classical relaxation rates from classical simulations. *J. Chem. Phys.* **1994**, *100*, 8359–8366.

(47) Marzari, N.; Mostofi, A. A.; Yates, J. R.; Souza, I.; Vanderbilt, D. Maximally localized wannier functions: theory and applications. *Rev. Mod. Phys.* **2012**, *84*, 1419–1475.

(48) King-Smith, R.; Vanderbilt, D. Theory of polarization of crystalline solids. *Phys. Rev. B* **1993**, *47*, 1651.

(49) Silvestrelli, P. L.; Parrinello, M. Water molecule dipole in the gas and in the liquid phase. *Phys. Rev. Lett.* **1999**, *82*, 3308.

(50) Sulpizi, M.; Salanne, M.; Sprik, M.; Gaigeot, M.-P. Vibrational sum frequency generation spectroscopy of the water liquid–Vapor interface from density functional theory-based molecular dynamics simulations. *J. Phys. Chem. Lett.* **2013**, *4*, 83–87.

(51) Schienbein, P. Spectroscopy from machine learning by accurately representing the atomic polar tensor. *J Chem Theory Comput* **2023**, *19*, 705–712.

(52) Yu, H.; Cui, Q. The vibrational spectra of protonated water clusters: a benchmark for self-consistent-charge density-functional tight binding. *J. Chem. Phys.* **2007**, *127*, 234504.

(53) Rossi, M.; Ceriotti, M.; Manolopoulos, D. E. How to remove the spurious resonances from ring polymer molecular dynamics. *J. Chem. Phys.* **2014**, *140*, 234116.

(54) Thämer, M.; De Marco, L.; Ramasesha, K.; Mandal, A.; Tokmakoff, A. Ultrafast 2D IR spectroscopy of the excess proton in liquid water. *Science* **2015**, *350*, 78–82.

(55) Auer, B. M.; Skinner, J. L. Vibrational sum-frequency spectroscopy of the water liquid/vapor interface. *J. Phys. Chem. B* **2009**, *113*, 4125–4130.

(56) Ohto, T.; Dodia, M.; Xu, J.; Imoto, S.; Tang, F.; Zysk, F.; Kühne, T. D.; Shigeta, Y.; Bonn, M.; Wu, X. et al. Accessing the accuracy of density functional theory through structure and dynamics of the water-air interface. *J. Phys. Chem. Lett.* **2019**, *10*, 4914–4919.

(57) Pieniazek, P. A.; Tainter, C. J.; Skinner, J. L. Interpretation of the water surface vibrational sum-frequency spectrum. *J. Chem. Phys.* **2011**, *135*, 044701.

(58) Biswas, B.; Singh, P. Signature of the surface hydrated proton and associated restructuring of water at model membrane interfaces: a vibrational sum frequency generation study. *Phys. Chem. Chem. Phys.* **2021**, *23*, 14764–14769.