Toward a simple molecular understanding of sum frequency generation at air-water interfaces

Joyce Noah-Vanhoucke,[1] Jared D. Smith,[2] and Phillip L. Geissler[1,2]

[1]*Department of Chemistry, University of California, Berkeley, California 94720*

[2]*Chemical Sciences Division Lawrence Berkeley National Lab, Berkeley, California 94720*

(Dated: January 13, 2009)

Abstract

Second-order vibrational spectroscopies successfully isolate signals from interfaces, but they report on intermolecular structure in a complicated and indirect way. Here we adapt a perspective on vibrational response developed for bulk spectroscopies to explore the microscopic fluctuations to which sum frequency generation (SFG), a popular surface-specific measurement, is most sensitive. We focus exclusively on inhomogeneous broadening of spectral susceptibilities for OH stretching of HOD as a dilute solute in D_2O. Exploiting a simple connection between vibrational frequency shifts and an electric field variable, we identify several functions of molecular orientation whose averages govern SFG. The frequency-dependence of these quantities is well captured by a pair of averages, involving alignment of OH and OD bonds with the surface normal at corresponding values of the electric field. The approximate form we obtain for SFG susceptibility highlights a dramatic sensitivity to the way a simulated liquid slab is partitioned for calculating second-order response.

I. INTRODUCTION

Molecular organization in heterogeneous environments, and in particular near interfaces, can differ substantially from that in homogeneous bulk materials. As an important example, the microscopic structure of many metals undergoes reconstruction at surfaces, with dramatic consequences for catalysis.[1, 2] The nature and extent of intermolecular rearrangement at liquid interfaces, while no less intriguing, are poorly understood. Of especial interest to biological, chemical and environmental sciences are interfaces involving water and salt solutions.[3–8]

Understanding molecular structure at solid-vapor interfaces has been greatly facilitated by surface-specific, high-resolution vibrational spectroscopy, which in those systems can resolve distinct microscopic arrangements. While it is unrealistic to expect well-separated spectral signatures of diverse intermolecular arrangements in a dense fluid, the adaptation of such tools to study liquid interfaces promises to provide substantial new insight into their chemistry and physics. Sum frequency generation spectroscopy (SFG), in particular, has emerged as an important tool for the investigation of the aqueous liquid-vapor interface at the molecular level. SFG spectroscopy is a second-order optical process, and therefore forbidden in bulk centrosymmetric media. Thus the SFG response, which is proportional to the square of the second-order susceptibility $\boldsymbol{\chi}$, results entirely from the interface where symmetry is necessarily broken. Unfortunately, this complex orientationally weighted vibrational spectrum does not admit a straightforward interpretation of the interfacial structure.

Despite the ambiguous link between the SFG spectrum and interfacial structure, a number of groups have made detailed assignments based on the SFG spectrum of pure water. For instance, Raymond *et al.* and Liu *et al.* have both decomposed the SFG spectrum into multiple hydrogen bonding configurations.[9, 10] Yet recent theoretical and experimental findings have indicated that such a multi-state interpretation is inappropriate even for the much simpler spectral response of bulk liquid water.[11–13] Computer simulations point to a more continuous view of intermolecular arrangements at the interface.[14, 15] Spectroscopic formalism developed by Morita and Hynes has enabled estimates of SFG response from such simulations.[16–21] But the complexity of these calculations, together with uncertainties in several important input parameters (such as nonresonant contributions to SFG susceptibility, transition dipoles and polarizabilities, and the strength of couplings between vibrational

modes) make it difficult to draw firm conclusions from comparison with experiment.

In this work we present a new approach for calculating resonant SFG susceptibility. Our purpose is not to generate more accurate predictions of spectroscopic response, though the methods we use have proved quantitatively accurate in the case of bulk Raman spectra.[11, 22] Instead, we aim to clarify how SFG reveals specific aspects of intermolecular structure, with a focus on hydrogen bonding arrangements at the air-water interface. In Sec. II we introduce a series of physically motivated approximations that yield from the formalism of Morita and Hynes a set of time-independent orientational averages governing SFG response. These simplifications are reminiscent of our earlier work on bulk vibrational spectroscopy, which emphasized a simple connection between shifts in hydroxyl stretching frequency and fluctuations of a specific component of the liquid's electric field. We further show that the frequency dependence of relevant orientational averages is well-characterized by a small number of readily interpreted quantities.

We explore the implications of our theoretical results using computer simulations of a molecular model for liquid water in coexistence with its vapor phase. Our numerical methods are described in detail in Sec. III, while practical issues of symmetry breaking that cannot be avoided in the computation of SFG signals are discussed in Sec. IV. We report simulation results in Sec. V a dilute solution of HOD in D_2O. These calculations illustrate the benefits of simple perspectives described in Sec. II. But they also highlight the crucial importance of resolving problematic symmetry-breaking issues, as well as the need for accurate computation of transition dipoles and polarizabilities that determine the strengths and signs of the contributions we have identified. In Sec. VI we conclude with a discussion of outstanding theoretical problems and the prospects for assigning structural motifs to features of measured SFG frequency dependence.

II. A SIMPLIFIED VIEW OF SFG SUSCEPTIBILITY

Our analysis of SFG focuses on an interface's susceptibility, $\boldsymbol{\chi}(\omega)$, which is primarily responsible for the IR frequency dependence of observed signals.[23] More specifically, we consider the imaginary part of the resonant contribution, $\mathrm{Im}\,\boldsymbol{\chi}^{(\mathrm{R})}$, whose connection to absorptive response offers the most straightforward link with intermolecular structure. In a significant advance, Shen and coworkers have recently demonstrated direct observation of

this quantity.[24] By contrast, conventional experiments measure only a squared SFG intensity that combines $\mathrm{Im}\,\boldsymbol{\chi}^{(\mathrm{R})}$ with its real counterpart, as well as a nonresonant contribution. These latter components obscure the frequency dependence of $\mathrm{Im}\,\boldsymbol{\chi}^{(\mathrm{R})}$ without providing additional useful information about variations in intermolecular structure across the spectrum: The real part of the resonant contribution can be obtained from $\mathrm{Im}\,\boldsymbol{\chi}^{(\mathrm{R})}$ via Kramers-Krönig relations, while the nonresonant contribution is nearly frequency independent.[9, 21, 25, 26]

In this paper we consider the most common choice of light polarizations, *ssp*, corresponding to a single element of the susceptibility tensor, $\chi^{(\mathrm{R})}_{xxz}$. We further restrict our attention to a dilute isotopic mixture of liquid water. (Generalizing our results for different polarizations should be straightforward.) As exploited in spectroscopic studies of bulk water, great simplification can be achieved by mixing H_2O with a large excess of D_2O, yielding HOD as the dominant solute species. In the case of dilute isotopic mixtures, stretching of an OH bond can be reasonably considered a normal mode of vibration, uncoupled to lower frequency OD stretching motions in its environment. Despite this simplification, few studies have been performed for such mixtures at the interface, and none have probed very dilute mixtures in detail. As will become apparent in our analysis, even intramolecular vibrational coupling can seriously threaten a simple microscopic interpretation of SFG.

Morita and Hynes have derived approximate formulas for the resonant susceptibility that may be evaluated by averaging functions of molecular orientation over an ensemble of thermal fluctuations.[16, 17] These results, based on the simple connection between SFG susceptibility and average molecular hyperpolarizability, $\boldsymbol{\chi} \propto \boldsymbol{\beta}$, provide a straightforward numerical prescription for estimating SFG signals using molecular simulations. But they do not reveal in a transparent way how SFG reports on specific aspects of molecular arrangements. Indeed, most applications arrive at a structural interpretation by dissecting the computed signal into contributions from various molecular populations. Our goal in this work is to simplify the theoretical basis of such calculations through a series of physically motivated approximations, in order to clarify the sensitivity of spectral features to interesting aspects of microscopic structure.

For our purposes, expressions of $\beta^{(\mathrm{R})}_{xxz}(\omega)$ either in terms of the dynamics of pertinent fluctuations at equilibrium[17] or more explicitly in terms of thermally distributed vibrational frequencies[16] would suffice. Existing theories pursuing the latter route require specification of a phenomenological dephasing rate γ and are therefore somewhat less fundamental. We

adopt the alternative, time domain approach of Ref. 17. Shortly, however, we will make an approximation amounting to the limit of infinitely slow dephasing, $\gamma \to 0$, at which point the two approaches become equivalent. Before doing so, we will cast the calculation of SFG susceptibility in a semiclassical form reminiscent of Kubo's line shape theory for simple absorptive processes.

Following the perturbative treatment of radiation-matter interactions due to Morita and Hynes, we begin by writing the hyperpolarizability as the one-sided Fourier transform of a correlation function involving the molecular dipole $\boldsymbol{\mu}$ at time zero and the molecular polarizability $\boldsymbol{\alpha}(t)$ at a later time t:

$$\beta^{(\mathrm{R})}_{xxz}(\omega) = \int_0^\infty dt\, e^{i\omega t}\phi(t) \tag{1}$$

$$\phi(t) = \langle \alpha_{xx}(t)\mu_z(0)\rangle\,. \tag{2}$$

Angular brackets denote unrestricted averaging over a canonical distribution of microscopic degrees of freedom, including both the OH vibrational coordinate R and the solvent configuration Γ. We regard $\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$ as quantum mechanical operators acting on R, rather than as classical variables as in Refs. 17, 18, 27 and 28.

The Hamiltonian governing dynamics of the vibrational wavefunction acquires explicit time dependence from interactions with the fluctuating solvent environment. We consider the solvent's influence as "pure dephasing," neglecting vibrational relaxation induced by coupling between the vibrational "system" and the solvent "bath." We instead focus on solvent-induced modulation of the energy gap $\hbar\omega_{10}(t)$ between ground $|0\rangle$ and first excited $|1\rangle$ vibrational states. Ignoring the very low probability of excitations to higher-lying vibrational states, and invoking the Condon approximation, we obtain

$$\phi(t) = \sum_{j,k,l} \alpha'_{jk}\mu'_l \left\langle j_x(t)k_x(t)l_z(0)\exp\left[-i\int_0^t d\tau\,\omega_{10}(\tau)\right]\right\rangle. \tag{3}$$

Here, j_x, k_x, and l_z are projections of the Cartesian unit vectors $\hat{\mathbf{j}}$, $\hat{\mathbf{k}}$, and $\hat{\mathbf{l}}$ of the molecular reference frame onto the x- and z-axes of the laboratory frame. Figure 1 depicts our choice of coordinate system, which includes the unit vector $\hat{\mathbf{u}}$ pointing along the OH bond as a principal direction. The sums of Eq. 3 thus run over vectors $\hat{\mathbf{u}}$, $\hat{\mathbf{v}}$ (also lying in the plane of the molecule), and $\hat{\mathbf{w}} \equiv \hat{\mathbf{u}} \times \hat{\mathbf{v}}$. The coefficients $\alpha'_{jk} = \hat{\mathbf{j}} \cdot \langle 0|(\partial\boldsymbol{\alpha}/\partial R)|1\rangle \cdot \hat{\mathbf{k}}$ and

$\mu'_l = \hat{\mathbf{l}} \cdot \langle 1|(\partial \boldsymbol{\mu}/\partial R)|0\rangle$ are the transition matrix elements of the polarizability and dipole derivatives with respect to the vibrational coordinate.

The correlation function $\phi(t)$ in Eq. 3 closely resembles those describing vibrational dephasing in the contexts of bulk Raman and infrared spectroscopies.[29, 30] Differing only in the appearance of molecular orientations relative to input light polarization, it could be evaluated using the extensive body of methods developed for computer simulations of those more straightforward probes.[11, 29, 31, 32] Independent of the work described in this paper, Skinner and coworkers have performed such a detailed calculation.[33]. Here we pursue a simplification of Eq. 3 that is similar in spirit to those of Refs. 11 and 22. Our primary approximations from this point are twofold: First, we imagine that spectral lineshapes are dominated by inhomogeneous broadening and only weakly manifest effects of motional narrowing. Second, we assert a linear relationship between instantaneous OH vibrational frequency and electrostatic forces experienced by the proton.

Neglect of dynamical broadening is strictly justified only when relaxation of a molecule's environment, as well as its own reorientation, proceeds slowly compared to the inverse width of the vibrational line shape. Although these time scales are not likely to be well separated in the case of liquid water, careful calculations by Skinner and coworkers show that the effects of motional narrowing are qualitatively insignificant.[33] Moreover, an analogous static approximation yields surprisingly accurate predictions for bulk Raman spectroscopy.[22] By dictating a correspondence between measured frequency and an oscillator's instantaneous environment, this assertion greatly facilitates a statistical interpretation of spectral features. Mathematically, it allows trivial evaluation of the time integral in Eq. 3, rendering $\phi(t)$ a Boltzmann average of purely sinusoidal functions of time,

$$\phi(t) = \sum_{j,k,l} \alpha'_{jk} \mu'_l \left\langle j_x k_x l_z e^{-i\omega_{10} t} \right\rangle . \tag{4}$$

Subsequent Laplace transformation yields for the imaginary part of $\beta^{(\mathrm{R})}_{xxz}(\omega)$ an average of $\delta(\omega - \omega_{10})$, *i.e.*, a probability distribution of vibrational frequencies, weighted by orientational statistics. Although this limit does not appear to have been explicitly considered in previous work on SFG, its usage is often implicit in studies that employ values of γ much smaller than the spectral line width[18, 26] and in discussions that associate spectral frequencies with molecular subensembles.[9, 10, 29, 34]

Finally, we introduce a simple relationship between ω_{10} and a component $\mathcal{E}$ of the liquid's

electric field, acting on the proton in the direction of the OH bond:

$$\omega_{10} = \omega_{10}^{(0)} + Q\mathcal{E}. \tag{5}$$

We and others have provided theoretical justification for this proportionality, both in the context of aqueous vibrational spectroscopy and more generally.[35–39] In Ref. 22 we reported values for the reference frequency $\omega_{10} = 3745\,\mathrm{cm}^{-1}$ and system-bath coupling $Q = 160.514\,\mathrm{cm}^{-1}\text{Å/V}$ that provide excellent correspondence between the probability distribution $P(\mathcal{E}) = \langle\delta(\mathcal{E} - \mathcal{E}(\Gamma))\rangle$ of electric field strength and the Raman spectrum of a dilute $\mathrm{HOD/D_2O}$ solution. (The variable Γ represents the solvent configuration.) The quantity Q is a property of a single water molecule, which, to first approximation, is independent of its environment. For the interfacial system under consideration, small changes in the magnitude of Q in the outer layer of the liquid would have little influence on our results. In the context of SFG, Eq. 5 allows us to isolate the spectral consequences of symmetry breaking at an interface. Combining Eqs. 2, 4, and 5, we arrive at our central theoretical result,

$$\mathrm{Im}\,\beta_{xxz}^{(\mathrm{R})}(\omega) = \frac{\pi}{Q} P(\mathcal{E}) \sum_{j,k,l} \alpha'_{jk}\mu'_l \langle j_x k_x l_z\rangle_{\mathcal{E}}\,, \tag{6}$$

where $P(\mathcal{E})$ is the distribution of electric field strengths for the entire sample, a distribution dominated by the bulk contributions, as discussed below. Aside from multiplicative constants, the hyperpolarizability within our approximations is just the product of $P(\mathcal{E})$ and a linear combination of orientational terms,

$$\langle j_x k_x l_z\rangle_{\mathcal{E}} = \frac{\langle j_x k_x l_z \delta(\mathcal{E} - \mathcal{E}(\Gamma)\rangle}{\langle\delta(\mathcal{E} - \mathcal{E}(\Gamma)\rangle}. \tag{7}$$

These conditional averages quantify the correlation between a molecule's hydrogen bonding environment, as characterized by $\mathcal{E}$, and its orientational bias due to the interface. They provide a concrete basis for rationalizing features of spectra computed according to Eq. 2 in terms of intermolecular structure.

Note that, in the thermodynamic limit, the electric field distribution $P(\mathcal{E})$ in Eq. 6 is not altered by the interface. Generally speaking, the physical influence of an interface is limited to a microscopic depth Δ_{int} below the surface.[52] Many studies have concluded that for liquid water near ambient conditions $\Delta_{\mathrm{int}} \sim 1\mathrm{nm}$ is not much larger than a molecular diameter.[7, 16] The thickness D of a typical experimental sample, on the other hand, is macroscopic. The fraction of molecules contributing to $P(\mathcal{E})$ whose statistics differ from

those of the bulk interior is thus extremely small. Mathematically, we could write $P(\mathcal{E})$ as the sum of its bulk counterpart and a perturbation of order $\epsilon = \Delta_{\text{int}}/D \ll 1$,

$$P(\mathcal{E}) = P_{\text{bulk}}(\mathcal{E}) + \epsilon g_{\text{int}}(\mathcal{E}). \tag{8}$$

The function $g_{\text{int}}(\mathcal{E})$, which quantifies changes in electric field statistics near the interface, is comparable in magnitude to $P_{\text{bulk}}(\mathcal{E})$. By contrast, the bulk contribution to orientational averages appearing in Eq. 6 vanish due to symmetry regardless of $\mathcal{E}$, so that $\langle j_x k_x l_z \rangle_{\mathcal{E}} \sim O(\epsilon)$. Replacing $P(\mathcal{E})$ by $P_{\text{bulk}}(\mathcal{E})$ in our result 6 for the hyperpolarizability therefore introduces an error of only $O(\epsilon)$, which may be neglected entirely for a macroscopic sample. As a matter of practice, interfacial systems represented in computer simulations are typically only a few nm thick, so that the ratio ϵ is not necessarily negligible. Nonetheless, for simulated aqueous systems, changes in $P(\mathcal{E})$ are not dramatic even in the first few monolayers of the liquid (as evidenced by computed surface Raman and IR spectra that differ little from their bulk counterparts).[17] One remains safe in considering only bulk electric field statistics for the purpose of evaluating $\text{Im}\,\beta^{(\text{R})}_{xxz}(\omega)$.

There are 18 choices of the set $\{j, k, l\}$ that yield distinct averages from Eq. 7. Several of these averages vanish identically due to symmetry: Because the interfacial system is statistically achiral, any microscopic configuration Γ can be reflected across the xz- or yz-planes to yield a mirror image microstate $\overline{\Gamma}$ with equal Boltzmann weight. In the case of yz-plane reflection, the x-component of OH and OD bond vectors changes sign. Linear combinations of these bond orientations determine basis vectors $\hat{u}$ and $\hat{v}$ of the molecular frame, which thus transform as $\overline{u}_x = -u_x$, $\overline{v}_x = -v_x$; all other components of $\hat{u}$ and $\hat{v}$ are unaffected by yz-reflection. As the vector product of the first two basis vectors, $\hat{\mathbf{w}}$ transforms differently: $\overline{w}_x = w_x$, $\overline{w}_y = -w_y$, and $\overline{w}_z = -w_z$.[53] As a result, products such as $u_x v_x w_z$ and $w_x u_x u_z$ yield zero upon averaging, regardless of electric field. Table 1 lists the ten distinct orientational averages that survive this elimination by symmetry.

The entries in Table 1 are ordered by the magnitude of the coefficients that multiply them in Eq. 6, as determined by electronic structure calculations for an isolated water molecule. Two of the averages that survive symmetry operations on the system as a whole are multiplied in Eq. 6 by coefficients that vanish due to internal symmetry of the molecule. For any typical configuration of the liquid, this latter symmetry will be broken due to interactions with surrounding molecules. High-level electronic structure calculations are

$\langle j_x k_x l_z \rangle_{\mathcal{E}}$	$\langle u_x^2 u_z \rangle_{\mathcal{E}}$	$\langle u_x^2 v_z \rangle_{\mathcal{E}}$	$\langle w_x^2 u_z \rangle_{\mathcal{E}}$	$\langle v_x^2 u_z \rangle_{\mathcal{E}}$	$\langle w_x^2 v_z \rangle_{\mathcal{E}}$	$\langle v_x^2 v_z \rangle_{\mathcal{E}}$	$\langle u_x v_x u_z \rangle_{\mathcal{E}}$	$\langle u_x v_x v_z \rangle_{\mathcal{E}}$	$\langle u_x w_x w_z \rangle_{\mathcal{E}}$	$\langle v_x w_x w_z \rangle_{\mathcal{E}}$
$\alpha'_{jk}\mu'_l/\alpha'_{uu}\mu'_u$	1	-0.37	0.23	0.21	0.09	-0.08	-0.02	0.007	0	0

TABLE I: Symmetry-allowed orientational averages and their coefficients in Eq. 6. Ab initio calculations were performed with GAUSSIAN 03 on an isolated water molecule employing DFT at the B3LYP/d-aug-cc-pVTZ level of theory. Polarizability and dipole derivatives were calculated numerically by displacing a single OH bond length by 0.01 Å, and calculating the polarizability and dipole moment at each position. This method uses a slightly different bond geometry, but is otherwise identical to that employed in Ref. 16. As expected the numerical results are only slightly different.

needed to establish whether $\langle u_x w_x w_z \rangle_{\mathcal{E}}$ and $\langle v_x w_x w_z \rangle_{\mathcal{E}}$ in fact remain negligible.

Only two of a molecule's three basis vectors $\hat{\mathbf{u}}$, $\hat{\mathbf{v}}$, and $\hat{\mathbf{w}}$ can be considered independent (since one can always be defined as a vector product of the other two). It may therefore not be surprising that the ten nonzero orientational averages carry a significant amount of redundant structural information. For example, we will show in Sec. V that $\langle w_x^2 u_z \rangle_{\mathcal{E}}$ is very nearly equal to $\langle v_x^2 u_z \rangle_{\mathcal{E}}$ at all values of $\mathcal{E}$. Fixing the angle $\theta = \cos^{-1}(u_z)$ between an OH bond and the surface normal evidently does little to bias a molecule's remaining orientational degrees of freedom. This observation highlights the weakness of molecular alignment at the air-water interface, a theme that will run through many of our numerical results. Because $w_x^2 + v_x^2 + u_x^2 = 1$ and $u_x^2 + u_y^2 + u_z^2 = 1$ due to geometry, and because $\langle u_x^2 u_z \rangle_{\mathcal{E}} = \langle u_y^2 u_z \rangle_{\mathcal{E}}$ due to symmetry, a near equivalence of $\langle w_x^2 u_z \rangle_{\mathcal{E}}$ and $\langle v_x^2 u_z \rangle_{\mathcal{E}}$ implies that

$$\left\langle v_x^2 u_z \right\rangle_{\mathcal{E}} \approx \left\langle w_x^2 u_z \right\rangle_{\mathcal{E}} \approx \frac{1}{4} \left(\langle u_z \rangle_{\mathcal{E}} + \left\langle u_z^3 \right\rangle_{\mathcal{E}} \right). \tag{9}$$

Noting that

$$\left\langle u_x^2 u_z \right\rangle_{\mathcal{E}} = \frac{1}{2} \left(\langle u_z \rangle_{\mathcal{E}} - \left\langle u_z^3 \right\rangle_{\mathcal{E}} \right) \tag{10}$$

simply as a result of symmetry, we see that three of the orientational averages in Table 1 are essentially determined by the simpler quantities $\langle u_z \rangle_{\mathcal{E}}$ and $\langle u_z^3 \rangle_{\mathcal{E}}$. Approximate relationships between SFG susceptibility and orientational averages of the form $\langle \cos^n \theta \rangle$ have been presented previously,[40, 41] but without the dependence on electric field that is the focus of our work. Not surprisingly, $\langle u_z \rangle_{\mathcal{E}}$ and $\langle u_z^3 \rangle_{\mathcal{E}}$ differ little in their qualitative dependence on electric field. All three contributions $\langle u_x^2 u_z \rangle_{\mathcal{E}}$, $\langle v_x^2 u_z \rangle_{\mathcal{E}}$, and $\langle w_x^2 u_z \rangle_{\mathcal{E}}$ can thus be understood

in terms of $\langle u_z \rangle_{\mathcal{E}}$, *i.e.*, the average projection of a hydrogen bond onto the outward surface normal given the electric field it experiences.

We will similarly show simulation results indicating that $\langle w_x^2 v_z \rangle_{\mathcal{E}} \approx \langle u_x^2 v_z \rangle_{\mathcal{E}}$ to a very good approximation. This analogous effect of weak alignment, together with the result of symmetry

$$\left\langle v_x^2 v_z \right\rangle_{\mathcal{E}} = \frac{1}{2} \left(\langle v_z \rangle_{\mathcal{E}} - \left\langle v_z^3 \right\rangle_{\mathcal{E}} \right), \tag{11}$$

carries similar implications for orientational averages involving v_z:

$$\left\langle u_x^2 v_z \right\rangle_{\mathcal{E}} \approx \left\langle w_x^2 v_z \right\rangle_{\mathcal{E}} \approx \frac{1}{4} \left(\langle v_z \rangle_{\mathcal{E}} + \left\langle v_z^3 \right\rangle_{\mathcal{E}} \right). \tag{12}$$

As before, the quantities $\langle v_z \rangle_{\mathcal{E}}$ and $\langle v_z^3 \rangle_{\mathcal{E}}$ depend on $\mathcal{E}$ in very similar ways, so that the contributions $\langle u_x^2 v_z \rangle_{\mathcal{E}}$, $\langle v_x^2 v_z \rangle_{\mathcal{E}}$, and $\langle w_x^2 v_z \rangle_{\mathcal{E}}$ can be understood through $\langle v_z \rangle_{\mathcal{E}}$.

Because $\hat{\mathbf{v}}$ lies in the plane including both the OH bond and the OD bond (directed along $\hat{\mathbf{u}}'$, see Fig. 1), it can be written as a linear combination of the two bond vectors, $\hat{\mathbf{v}} = \csc(\theta_{\mathrm{HOD}})\hat{\mathbf{u}}' - \cot(\theta_{\mathrm{HOD}})\hat{\mathbf{u}}$, where θ_{HOD} is the HOD bond angle (a fixed parameter in the models we will consider). This separation provides a very appealing interpretation for the average $\langle v_z \rangle_{\mathcal{E}}$ as a linear combination of $\langle u_z \rangle_{\mathcal{E}}$ and $\langle u_z' \rangle_{\mathcal{E}}$. The former quantity appeared already in Eqs. 9 and 10, and its meaning was explained above. The latter, $\langle u_z' \rangle_{\mathcal{E}}$, quantifies the average projection of the OD bond onto the surface normal, given the electric field modulating the OH vibrational frequency. As discussed in later sections, the two functions of $\mathcal{E}$ give complementary views on the relationship between hydrogen bonding and molecular orientation at the interface. Together, they thoroughly characterize the first six entries in Table 1, and we will see that their frequency dependencies dominate distinct ranges of $\mathcal{E}$.

As judged by electronic and vibrational properties of an isolated HOD molecule, contributions from the final four entries in Table 1 are suppressed by weak corresponding transition dipoles and polarizabilities. It is nonetheless notable that their electric field dependencies are also well characterized by $\langle u_z \rangle_{\mathcal{E}}$ and $\langle u_z' \rangle_{\mathcal{E}}$. Results supporting this conclusion will be presented in Sec. V.

Through a series of approximations guided both by physical considerations and by empirical observations, we have distilled the essence of a calculated SFG spectrum down to two functions of electric field, $\langle u_z \rangle_{\mathcal{E}}$ and $\langle u_z' \rangle_{\mathcal{E}}$. A linear combination of these orientational averages, and the closely related quantities $\langle u_z^3 \rangle_{\mathcal{E}}$ and $\langle v_z^3 \rangle_{\mathcal{E}}$, multiply the bulk electric field

distribution to determine nonlinear spectral response. In subsequent sections we will explore these functions in detail as a guide for understanding how SFG measurements reflect on microscopic structure at the air-water interface.

III. SIMULATION METHODS

We explore the statistics of orientation and electric field at liquid-vapor interfaces using computer simulations of a simple molecular model. For this purpose we have employed standard techniques of molecular dynamics and Monte Carlo (MC) importance sampling. All results reported here are obtained from systems including $N = 512$ molecules, interacting via the SPC/E potential[42] computed with Ewald sums, at ambient temperature and at fixed densities that enforce liquid-vapor coexistence. Since we will focus exclusively on static, classical properties, masses of the atoms are irrelevant. For spectroscopic purposes we view the system as one HOD solute amongst 511 D_2O solvent molecules. For each sampled configuration of the system we may regard any one of 2×512 atoms as the deuteron. Computed averages therefore converge about 1000 times more rapidly than they would in corresponding simulations that explicitly include a single HOD solute.

Periodic boundary conditions were imposed in all three principal directions of the laboratory frame, resulting in a liquid slab oriented perpendicular to the z-axis. With simulation cell dimensions $L_x = L_y = 20$Å in the x- and y-directions, the slab adopts a thickness of roughly 40Å. The size of the simulation cell in the z direction, $L_z = 100$Å, was thus sufficient to provide nearly 60Å of vapor (essentially vacuum on these scales at 298 K) between periodically replicated slabs. Equilibration was achieved by allowing a well-equilibrated configuration from the uniform liquid phase to expand in the z-direction, followed by extensive sampling (at least 200 ps or, in the case of MC simulations, 10^6 sweeps) before data was collected over trajectories of at least 2 ns (or 10^6 MC sweeps). We report vertical coordinates $\Delta z = z - z_{\text{Gibbs}}$ of H, O, and D atoms relative to the location z_{Gibbs} of the closest Gibbs dividing surface, *i.e.*, the value of z where solvent density falls to half its bulk value.

Molecular dynamics simulations were performed using the LAMMPS software package.[43] In these calculations rigid molecular geometries were maintained using the SHAKE algorithm.[44] Equilibration runs employed a Nosé-Hoover thermostat,[45] while production runs were propagated by integrating Newton's equations of motion using the

velocity Verlet algorithm.[46] In both cases $\Delta t = 1$ fs served as the fundamental time step of integration.

If realized in the laboratory, the liquid samples simulated in this and most other theoretical studies of air-water interfaces would in fact exhibit no second-order spectroscopic response at all. The problem is not a lack but an excess of interfaces. Because a liquid slab possesses two identical interfaces with antiparallel surface normal vectors, it does not in fact break symmetry in the way required for SFG. Computing second-order susceptibilities from such a system thus requires an artificial breaking of symmetry. The natural strategy for doing so is to place an imaginary dividing surface through the slab, parallel to the interfaces. An isolated interfacial region then comprises all molecules residing on one side of the dividing surface.[54] Considered independently from the second notional interfacial region, such a subsystem would yield a nonzero SFG signal. This procedure is so reasonable that many computational SFG studies do not report that it has been implemented, much less how it has been implemented.

There is no ambiguity in assigning a point particle to one side of a planar dividing surface. The internal structure of an HOD molecule, however, mandates that many molecules in a typical configuration of the liquid will straddle the dividing surface. To which interface they should be assigned is surprisingly a far from academic question. We will show in Sec. V that the convention one uses to divide an aqueous liquid slab into two separate interfaces can dramatically influence predictions for the frequency-dependent susceptibility. We explain how such sensitivity arises and describe its relationship to experiments in the following section.

IV. SYMMETRY BREAKING CONVENTIONS

Consider the average of an odd function $f(u_z)$ of molecular orientation (*e.g.*, $f = u_z$ or $f = u_z^3$), which vanishes in a centrosymmetric bulk medium, but can be nonzero at an interface. Here we will mimic the attenuation with depth of radiation incident on a sample by multiplying contributions to $\langle f(u_z)\rangle$ with a decaying weight $e^{z/L}$. (The liquid is taken to extend from $z = -\infty$ to $z \approx 0$.) This factor could also be viewed as a mathematical device to break symmetry of a simulated liquid slab in a smooth way. In either case the decay length L should be large compared to the scale Δ_{int} over which the interface substantially

influences orientational statistics, but small compared to the sample thickness D. Although the arguments we make here are unrelated to electric field fluctuations, we will condition the average on $\mathcal{E}$ for application to the quantities controlling SFG:

$$\langle f(u_z)\rangle_{\mathcal{E}} = \int_{-\infty}^{d} dz \int_{-1}^{1} du_z \, e^{z/L} p(u_z, z|\mathcal{E}) \, f(u_z). \tag{13}$$

The upper limit d for z-integration refers to an elevation above the interface at which the density is sufficiently low as to strongly suppress any contribution to $\langle f(u_z)\rangle_{\mathcal{E}}$ from the region $z > d$. For aqueous systems $d = O(\Delta_{\rm int}) \ll L$.

The joint probability $p(u_z, z|\mathcal{E})$ that an HOD molecule is located at vertical coordinate z and has orientation u_z, given electric field $\mathcal{E}$, requires precise specification of molecular location. Let us consider two extreme cases: The vertical coordinate z of an HOD molecule is assigned to be (1) that of its oxygen atom, $z^{(\rm O)}$, or (2) that of its hydrogen atom, $z^{(\rm H)}$. If an OH bond straddles the dividing surface distinguishing top and bottom interfacial regions of a liquid slab, say $z^{(\rm O)} > 0$ and $z^{(\rm H)} < 0$, then convention (1) would associate the molecule with the top region, while convention (2) would assign it to the bottom region. In this configuration $u_z < 0$ from the perspective of the top region, while $u_z > 0$ from the perspective of the bottom region, whose surface normal points opposite to $\hat{\mathbf{z}}$. (Recall that u_z denotes the projection of an OH bond vector onto the outward surface normal.) As a result, averages such as $\langle u_z\rangle_{\mathcal{E}}$ will be systematically biased to lower values for convention (1) and to higher values for convention (2). As we have presented them, these biases would appear to be an unphysical artifact of the sharp and arbitrary manner in which symmetry has been broken. They would also appear to be quite small in magnitude, and therefore unimportant to the qualitative nature of a predicted SFG signal. Both of these appearances are deceiving.

In an SFG measurement the contribution of a molecule at vertical coordinate z to the observed signal is of course not decided in a sharp way according to an artificial dividing surface. Rather, it is determined by the smooth attenuation of incident radiation as it is absorbed and scattered by the sample. Because the length scale L of this attenuation greatly exceeds the OH bond length $\ell \approx 1$Å, one might reasonably expect that the distinctions described above become unimportant. To evaluate this expectation, it is useful to resolve the statistics of orientation and electric field by vertical coordinate z. For an HOD molecule probability distributions involving z are only well-defined once we specify whether $z = z^{(\rm O)}$,

$z = z^{(\mathrm{H})}$, or otherwise. For the first convention introduced above we would define the joint distribution of u_z and z given $\mathcal{E}$ as

$$p^{(\mathrm{O})}(u_z, z|\mathcal{E}) = \frac{\left\langle \delta(u_z - u_z(\Gamma))\delta(z - z^{(\mathrm{O})}(\Gamma))\delta(\mathcal{E} - \mathcal{E}(\Gamma))\right\rangle}{\langle \delta(\mathcal{E} - \mathcal{E}(\Gamma))\rangle}, \tag{14}$$

while for the second we would define

$$p^{(\mathrm{H})}(u_z, z|\mathcal{E}) = \frac{\left\langle \delta(u_z - u_z(\Gamma))\delta(z - z^{(\mathrm{H})}(\Gamma))\delta(\mathcal{E} - \mathcal{E}(\Gamma))\right\rangle}{\langle \delta(\mathcal{E} - \mathcal{E}(\Gamma))\rangle}. \tag{15}$$

Since $z^{(\mathrm{H})} = z^{(\mathrm{O})} + \ell u_z$, the two distributions are related simply by

$$p^{(\mathrm{O})}(u_z, z|\mathcal{E}) = p^{(\mathrm{H})}(u_z, z + \ell u_z|\mathcal{E}). \tag{16}$$

To the extent that these probabilities vary slowly with vertical coordinate, averages taken over $p^{(\mathrm{O})}(u_z, z|\mathcal{E})$ and $p^{(\mathrm{H})}(u_z, z|\mathcal{E})$ differ in general by an amount proportional to the bond length ℓ. This difference is a matter of concern only because the range Δ_{int} of interfacial influence is not dramatically larger than ℓ.

Assuming such probabilities vary smoothly with z, we can exploit the connection 16 between different conventions to write

$$p^{(\mathrm{O})}(u_z, z|\mathcal{E}) \approx p^{(\mathrm{H})}(u_z, z|\mathcal{E}) + \ell u_z \left(\frac{\partial p^{(\mathrm{H})}(u_z, z|\mathcal{E})}{\partial z}\right)_{u_z}. \tag{17}$$

Averages of $f(u_z)$ calculated using the two conventions are then related by

$$\langle f(u_z)\rangle_{\mathcal{E}}^{(\mathrm{O})} = \langle f(u_z)\rangle_{\mathcal{E}}^{(\mathrm{H})} + \ell \int_{-\infty}^{d} dz \int du_z e^{-(d-z)/L} \left(\frac{\partial p^{(\mathrm{H})}(u_z, z|\mathcal{E})}{\partial z}\right)_{u_z} u_z f(u_z). \tag{18}$$

Integrating by parts, we obtain

$$\langle f(u_z)\rangle_{\mathcal{E}}^{(\mathrm{O})} = \langle f(u_z)\rangle_{\mathcal{E}}^{(\mathrm{H})} - \frac{\ell}{L} \int_{-\infty}^{d} dz \int du_z e^{-(d-z)/L} p^{(\mathrm{H})}(u_z, z|\mathcal{E}) u_z f(u_z) \tag{19}$$

where boundary terms have vanished by construction.

Following the arguments leading to Eq. 8, we can separate bulk and interfacial contributions to $p^{(\mathrm{H})}$:

$$p^{(\mathrm{H})}(u_z, z|\mathcal{E}) = p^{(\mathrm{H})}_{\mathrm{bulk}}(u_z, z|\mathcal{E}) + \epsilon g^{(\mathrm{H})}_{\mathrm{int}}(u_z, z|\mathcal{E}) \tag{20}$$

where $\epsilon = \Delta_{\mathrm{int}}/D \ll 1$. As with $P(\mathcal{E})$ in Eq. 6, the bulk contribution will dominate any average that does not strictly vanish away from the interface. Unlike $f(u_z)$, the even function $f(u_z)u_z$ will not average to zero in the bulk due to symmetry. The interfacial contribution to the second term in Eq. 19 can therefore be discarded. Recognizing that orientational averages in the bulk are independent of both z and $\mathcal{E}$, and noting that the depth distribution in bulk is uniform, $p^{(\mathrm{H})}_{\mathrm{bulk}}(z) = D^{-1}$, we obtain

$$\langle f(u_z)\rangle^{(\mathrm{O})}_{\mathcal{E}} = \langle f(u_z)\rangle^{(\mathrm{H})}_{\mathcal{E}} - \frac{\ell}{D}\langle f(u_z)u_z\rangle_{\mathrm{bulk}}. \tag{21}$$

The second term of Eq. 21 is by no means negligible, since $\langle f(u_z)\rangle^{(\mathrm{H})}_{\mathcal{E}} \sim O(\Delta_{\mathrm{int}}/D)$ and the interfacial region of a dense liquid spans only a handful of molecular layers.

When the differences between smoothly attenuated averages of a function $f(u_z)$ over the distributions 14 and 15 are applied to the specific case $\langle u_z\rangle_{\mathcal{E}}$ we obtain

$$\langle u_z\rangle^{(\mathrm{O})}_{\mathcal{E}} = \langle u_z\rangle^{(\mathrm{H})}_{\mathcal{E}} - \frac{\ell}{3D}. \tag{22}$$

As anticipated, the bias of convention (1) is negative relative to that of convention (2). The difference $\ell/3D$ is small in an absolute sense (recall that D is the thickness of the entire sample), but so is $\langle u_z\rangle_{\mathcal{E}} = O(\Delta_{\mathrm{int}}/D)$. Remarkably, this result does not depend on the distance L over which averaging is attenuated. The issue we encounter when using a sharp cutoff is recapitulated when using a smooth cutoff due to an accumulation of bias over the entire range of averaging. Similar consequences associating a molecular property with that of a single constituent atom have been described in the context of computing electrostatic potentials on molecular centers.[47]

Which convention is more physically realistic is not a simple question to answer. The central issue is where, and to what extent, the spectroscopic response of a molecule is localized. One expects the transition dipole contribution to be most sensitive to the position of the proton, whose motion largely constitutes the OH stretching mode. Polarizability, on the other hand, arises from electronic fluctuations that are distributed more broadly. Careful quantum chemistry calculations, as well as considerations of radiation-matter interaction that do not caricature HOD as a point dipole, may be needed to resolve this issue.

The convention of our computer simulations is to break symmetry by considering contributions only from protons situated above the slab's center of mass, $z^{(\mathrm{H})} > z^{(\mathrm{c.o.m.})}$. In the context of the above discussion, we have in effect attenuated averaging over a scale $L \approx D/2$. For the geometry of our system, $D \approx 40$Å, we thus expect $\langle u_z \rangle_{\mathcal{E}}^{(\mathrm{O})}$ and $\langle u_z \rangle_{\mathcal{E}}^{(\mathrm{H})}$ to differ by $\ell/3D \approx 0.01$. Simulation results exhibit an offset within a factor of two of this estimate.

V. SIMULATION RESULTS

In Sec. II we introduced two orientational averages which control SFG susceptibility in our approximations. Our discussion will concentrate on these quantities and their microscopic significance. Because the dependence on electric field $\mathcal{E}$ is of particular interest, it is useful to begin by considering the statistics of $\mathcal{E}$ alone, as a backdrop for understanding variations of $\langle u_z \rangle_{\mathcal{E}}$ and $\langle u'_z \rangle_{\mathcal{E}}$.

Thermal distributions of electric field, resolved for water molecules in 1Å-thick horizontal slices below the interface, are shown in Fig. 2. In order to compare fluctuations at different depths, we have scaled each distribution by its value at $\mathcal{E}_0$, the most likely value of $\mathcal{E}$ in the bulk. As the interface is approached, a distinct peak emerges at the right extreme of the distribution. We distinguish between this extreme and the rest of the distribution by the dividing point $\mathcal{E}^*$, as indicated in Fig. 2. We and others have shown for bulk systems a strong correlation between $\mathcal{E}$ and hydrogen bond geometry, with more positive values of $\mathcal{E}$ corresponding to weaker hydrogen bonds.[11, 22, 37, 38, 48] It is therefore natural to associate the growing weight for $\mathcal{E} > \mathcal{E}^*$ with a population of HOD molecules at the interface whose protons engage in very weak hydrogen bonds or lack a hydrogen bond acceptor altogether. Indeed, molecules corresponding to this population possess with high probability a "dangling" or "free" hydroxyl group.[55] Others have drawn similar conclusions.[6, 49, 50]

Perhaps more strikingly, electric field statistics within the hydrogen bonded population are essentially invariant with depth. Even molecules above the Gibbs dividing surface experience extremely negative fluctuations in $\mathcal{E}$ (say, $\mathcal{E} < -3$V/Å, corresponding to abnormally strong hydrogen bonds) with a probability relative to $P(\mathcal{E}_0)$ that is nearly identical to the bulk value. It is reasonable to conclude that the distribution of intact hydrogen bond geometries is very weakly perturbed by proximity to the interface. In support of this conclusion, we find that the distance r_{OH} between a proton and its hydrogen bond accepting oxygen

atom follows a distribution near the interface whose shape for $r_{\mathrm{OH}} < 2\text{Å}$ is hardly changed from the bulk.

Before focusing on the conceptually simple functions $\langle u_z \rangle_{\mathcal{E}}$ and $\langle u'_z \rangle_{\mathcal{E}}$, we present results for the full set of nonvanishing orientational averages listed in Table 1. These quantities, plotted in Fig. 3 as functions of $\mathcal{E}$, fall roughly into two classes: those that peak in the region $\mathcal{E} > \mathcal{E}^*$ and are very small in magnitude for $\mathcal{E} < \mathcal{E}^*$; and those that change sign near $\mathcal{E} \approx \mathcal{E}_0$ and then grow approximately linearly in magnitude as $\mathcal{E}$ decreases. Panels (b) and (c) further demonstrate these similarities by showing the accuracy of Eqs. 9 and 12. Here we have multiplied orientational averages by $P(\mathcal{E})$, as they appear in our prediction (6) for SFG susceptibility. This multiplication emphasizes small differences near the peak of $P(\mathcal{E})$, which decays rapidly in the region $\mathcal{E} > \mathcal{E}^*$.

For the remainder of this paper we will examine $\langle u_z \rangle_{\mathcal{E}}$ and $\langle u'_z \rangle_{\mathcal{E}}$ as prototypes for the frequency (or, equivalently, electric field) dependence of averages in Fig. 3. The first of these functions quantifies correlations that are the subject of many discussions on SFG: If OH bonds with vibrational frequency in a particular range are strongly aligned parallel (antiparallel) to the outward surface normal, $\langle u_z \rangle_{\mathcal{E}}$ will be large and positive (negative) at those frequencies. The second function carries more subtle structural information: Nonzero $\langle u'_z \rangle_{\mathcal{E}}$ at a certain frequency signifies correlations between the electric field acting on the OH bond and orientation of the OD bond vector. The geometry of a water molecule ensures that $\langle u_z \rangle_{\mathcal{E}}$ and $\langle u'_z \rangle_{\mathcal{E}}$ will oppose one another to some degree. Strong alignment of one bond vector in any direction implies partial anti-alignment of the other. Nonetheless, we will see that $\langle u'_z \rangle_{\mathcal{E}}$ reveals aspects of molecular arrangement that are not apparent from $\langle u_z \rangle_{\mathcal{E}}$.

Figure 4 shows $\langle u_z \rangle_{\mathcal{E}}$ and $\langle u'_z \rangle_{\mathcal{E}}$ as functions of electric field, as well as their products with the distribution $P(\mathcal{E})$. The very small magnitude of $\langle u_z \rangle_{\mathcal{E}}$ for $\mathcal{E} < \mathcal{E}^*$ indicates that correlations between OH vibrational frequency and OH bond orientation are significant only for free OH groups, which preferentially align parallel to the surface normal. Hydroxyl groups that do engage in hydrogen bonding exhibit very little orientational bias. Because $P(\mathcal{E})$ peaks in the corresponding frequency range, however, any residual bias is greatly amplified. Together, these effects conspire to make the convention for breaking symmetry of a liquid slab an important issue. We have shown that even a smooth cutoff can impart a frequency-independent bias depending on which atom of an HOD molecule is used to gauge its depth. Although this bias is weak, it is effectively the only source of preferential

alignment for hydrogen bonded molecules. We illustrate this situation in Fig. 4 by including results for both of the conventions we have described.

By contrast, $\langle u'_z \rangle_{\mathcal{E}}$ exhibits non-negligible frequency dependence over the whole range of $\mathcal{E}$. Net anti-alignment in the free OH region is expected from the arguments above. If a hydroxyl group is substantially aligned with the surface normal, geometry requires the OD bond to point on average opposite $\hat{\mathbf{z}}$. Variations over the range $\mathcal{E} < \mathcal{E}^*$ are more interesting. Following the minimum at $\mathcal{E} \approx 0$, $\langle u'_z \rangle_{\mathcal{E}}$ systematically increases with decreasing field. When the hydroxyl group engages in strong hydrogen bonding, the OD bond tends to align, weakly but noticeably, parallel to the surface normal. In other words, breaking one of a water molecule's hydrogen bonds at the interface tends to strengthen its other interactions. This interpretation is illustrated more directly in Fig. 5, which shows average alignment of u_z and u'_z as functions of hydrogen bond length (rather than electric field). Indeed, short OH$\cdots$O hydrogen bonds bias the OD bond vector to point upwards.

By considering only dilute isotopic mixtures, we have avoided complications due to coupling among different vibrational modes. In pure H_2O or D_2O a typical normal mode of vibration includes significant participation of more than one hydroxyl group, and perhaps more than one molecule. By mixing the character of different hydroxyl stretches on the same molecule, intramolecular coupling will have the basic effect of mixing contributions due to $\langle u_z \rangle_{\mathcal{E}}$ and to $\langle u'_z \rangle_{\mathcal{E}}$. If mixing is strong, it will not even be possible to decompose a computed SFG susceptibility into simply understood contributions. Orientational averages at all frequencies will be sensitive to alignment of both a molecule's hydroxyl groups. In this case one cannot unambiguously reason, *e.g.*, that a negative contribution to the computed susceptibility for $\mathcal{E} < \mathcal{E}^*$ indicates downward alignment of strong hydrogen bonds.

VI. CONCLUSIONS

We have presented a theoretical framework for understanding, and perhaps ultimately predicting, SFG response for the air-water interface. The same analysis can also be used to investigate the SFG response for dilute electrolyte solutions at the air interface, and such studies are currently underway. By design, our treatment is less elaborate in its details than previous calculations. With complications of bath dynamics, vibrational coupling and relaxation, and fluctuating electronic polarization removed, a relatively simple physical picture

emerges for the microscopic origins of interfacial hyperpolarizability. Specifically, we can identify a handful of specific measures of orientational bias that determine our approximation to resonant second-order response. Because the orienting effects of air-water interfaces are rather weak, the frequency dependence of all relevant orientational averages are well captured by a pair of functions involving average projections of OH and OD bond vectors onto the outward surface normal.

These simplifications, together with results of computer simulations, provide several insights that would have been difficult to extract from more detailed expressions for hyperpolarizability or from more elaborate molecular models. Our results reveal, for example, an alarming sensitivity to the convention one chooses for breaking symmetry of a simulated liquid slab. Obtaining accurate predictions of SFG from related approaches will require close attention to this accumulated effect of water molecules' internal structure. It is not clear that a spectroscopic formalism based on the dipole approximation is sufficient for this purpose, especially when normal modes involve motion of multiple protons (as is the case for pure H_2O and D_2O). Due to the ambiguity facing current methodologies, we recommend that questions such as, in which direction do hydrogen bonded OH groups point on average at the interface, would be better posed as, do hydrogen bonded OH groups at the interface adopt a noteworthy net orientation on average? An affirmative answer would require an insensitivity to symmetry-breaking convention. For our simulations of the SPC/E model in a liquid slab geometry, the answer according to this criterion is decidedly negative. These issues are unfortunately important not only for achieving quantitative SFG predictions. Weak, frequency-independent net orientation can alter the qualitative character of computed susceptibilities, due to the weighting of orientational averages by the bulk electric field distribution, which decays rapidly outside the frequency range corresponding to hydrogen bonded species.

We have stopped short of presenting full numerical results for SFG signals that can be measured in the laboratory. The primary reason for doing so is that the coefficients determining the linear combination of orientational averages in Eq. 6 have not been well characterized for water molecules in solution or in heterogeneous interfacial environments. Our principal result for neat solvent is that Eq. 6 is essentially a linear combination of the conceptually appealing functions $\langle u_z \rangle_{\mathcal{E}}$ and $\langle u'_z \rangle_{\mathcal{E}}$. The range of susceptibilities that could be obtained from the data plotted in Fig. 4 is not difficult to assess by visual inspection, an exercise

we leave to the interested reader. Ambiguity in methods of symmetry-breaking, significant nonresonant contributions to most experimental data, and the possible importance of non-Condon effects discourage us from following fitting procedures others have followed to infer these parameters.[56] We note, however, that the contributions to second-order susceptibility we have calculated could be combined with reasonable coefficients to roughly reproduce results reported from more involved approaches.

The theoretical perspective we have developed argues strongly against interpreting SFG data as one would a simple absorption spectrum. Associating spectral peaks with distinct intermolecular arrangements, and their intensities as populations, can in fact be misguided even for simple absorptive measurements. In the context of vibrational spectroscopy of bulk liquid water, we have shown for simulated systems that a pronounced shoulder in the Raman line shape signifies nothing more than the nonlinear relationship between frequency and hydrogen bond geometry.[11] For liquid-vapor interfaces as well, distributions of r_{OH} for hydrogen bonded species do not suggest a meaningful segregation into ice-like and liquid-like structures[51] as previously suggested.[50] By this measure only free OH species stand out as a discrete variation in intermolecular structure.

Despite a lack of discrete variety in the local geometry of hydrogen bonded molecules, it is clear that rich SFG signals could nonetheless be obtained from our numerical results. For example, a simple sum of the two contributions plotted in Fig. 4 would yield a susceptibility that changes sign twice as a function of frequency. A peak at very negative $\mathcal{E}$, like that observed in recent experiments,[24] would result. Its origin can be traced not to a distinct microscopic structure, but instead to a melange of free hydroxyl alignment, weak correlation between OH vibrational frequency and OD orientation, and weighting by the bulk electric field distribution. The location and height of this feature, and indeed its very existence, are sensitive to changing the coefficients of linear combination.

Acknowledgments

This work was supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division and the Chemical Sciences, Geosciences, and Biosciences Division under the Department of Energy Contract No. DE-AC02-05CH11231. Gaussian calculations were performed at the U.C. Berkeley College of

Chemistry Molecular Graphics and Computation Facility (NSF Grant No. CHE-0233882). J.D.S is supported by the Camille and Henry Dreyfus Foundation postdoctoral program in environmental chemistry.

---

[1] G. A. Somorjai, *Introduction to Surface Chemisty and Catalysis* (Wiley, 1994).

[2] G. A. Somorjai, Annu. Rev. Phys. Chem. **45**, 721 (1994).

[3] S. McLaughlin, Annu. Rev. Biophys. Biophys. Chem. **18**, 113 (1989).

[4] G. Spicer, E. Chapman, B. Finlayson-Pitts, R. Plastridge, J. Hubbe, J. Fast, and C. Berkowitz, Nature **394**, 353 (1998).

[5] E. Knipping, M. Lakin, K. Foster, P. Jungwirth, D. Tobias, R. Gerber, D. Dabdub, and B. Finlayson-Pitts, Science **288**, 301 (2000).

[6] K. B. Eisenthal, Chem. Rev. **96**, 1343 (1996).

[7] P. B. Petersen and R. J. Saykally, Annu. Rev. Chem. **57**, 333 (2006).

[8] T.-M. Chang and L. X. Dang, Chem. Rev. **106**, 1305 (2006).

[9] D. Liu, G. Ma, L. Levering, and H. Allen, J. Phys. Chem. B **108**, 2252 (2004).

[10] E. A. Raymond and G. L. Richmond, J. Phys. Chem. B **108**, 5051 (2004).

[11] J. D. Smith, C. D. Cappa, K. R. Wilson, R. C. Cohen, P. L. Geissler, and R. J. Saykally, Proc. Nat. Acad. Sci. USA **102**, 14171 (2005).

[12] J. D. Eaves, J. J. Loparo, C. J. Fecko, S. T. Roberts, A. Tokmakoff, and P. L. Geissler, Proc. Nat. Acad. Sci. USA **102**, 13019 (2005).

[13] P. L. Geissler, J. Am. Chem. Soc. **127**, 14930 (2005).

[14] R. M. Townsend and S. A. Rice, J. Chem. Phys. **94**, 2207 (1991).

[15] K. A. Motakabbir and M. L. Berkowitz, Chem. Phys. Lett. **176**, 61 (1991).

[16] A. Morita and J. T. Hynes, Chem. Phys. **258**, 371 (2000).

[17] A. Morita and J. T. Hynes, J. Phys. Chem. B **106**, 673 (2002).

[18] E. C. Brown, M. Mucha, P. Jungwirth, and D. J. Tobias, J. Phys. Chem. B **109**, 7934 (2005).

[19] V. Buch, T. Tarbuck, G. L. Richmond, H. Groenzin, I. Li, and M. J. Shultz, J. Chem. Phys. **127**, 204710 (2007).

[20] D. S. Walker, D. K. Hore, and G. L. Richmond, J. Phys. Chem. B **110**, 20451 (2006).

[21] T. Ishiyama and A. Morita, J. Phys. Chem. C **111**, 721 (2007).

[22] J. D. Smith, R. J. Saykally, and P. L. Geissler, J. Am. Chem. Soc. **129**, 13847 (2007).

[23] Y. R. Shen, *The Principles of Nonlinear Optics* (Wiley, 1984).

[24] N. Yi, V. Ostroverkhov, C. Tian, and Y. R. Shen, Phys. Rev. Lett. **100**, 096102 (2008).

[25] E. V. Alieva, L. A. Kuzik, and V. A. Yakovlev, Chem. Phys. Lett. **292**, 542 (1998).

[26] D. S. Walker and G. L. Richmond, J. Phys. Chem. C **111**, 8321 (2007).

[27] A. Perry, H. Ahlborn, B. Space, and P. B. Moore, J. Chem. Phys. **118**, 8411 (2003).

[28] A. Perry, C. Neipert, C. Ridley, B. Space, and P. B. Moore, Phys. Rev. E. **71**, 050601(R) (2005).

[29] S. A. Corcelli and J. L. Skinner, J. Phys. Chem. A **109**, 6154 (2005).

[30] J. R. Schmidt, S. A. Corcelli, and J. L. Skinner, J. Chem. Phys. **123**, 044513 (2005).

[31] W. B. Bosma, L. E. Fried, and S. Mukamel, J. Chem. Phys. **98**, 4413 (1993).

[32] E. Harder, J. D. Eaves, A. Tokmakoff, and B. J. Berne, Proc. Nat. Acad. Sci. USA **102**, 11611 (2005).

[33] J. L. Skinner (2008), personal communication.

[34] T. Steinel, J. B. Asbury, S. A. Corcelli, C. P. Lawrence, J. L. Skinner, and M. D. Fayer, Chem. Phys. Lett. **386**, 295 (2004).

[35] V. Buch and J. P. Devlin, J. Chem. Phys. p. 3437 (1999).

[36] J. B. Asbury, T. Steinel, K. Kwak, S. A. Corcelli, C. P. Lawrence, J. L. Skinner, and M. D. Fayer, J. Chem. Phys. **121**, 12431 (2004).

[37] S. A. Corcelli, C. P. Lawrence, and J. L. Skinner, J. Chem. Phys. **120**, 8107 (2004).

[38] J. D. Eaves, A. Tokmakoff, and P. L. Geissler, J. Phys. Chem. A **109**, 9424 (2005).

[39] H. Sumi and R. Marcus, J. Chem. Phys. **84**, 4894 (1986).

[40] J. Wang, Z. Paszti, M. A. Even, and Z. Chen, J. Am. Chem. Soc. **124**, 7016 (2002).

[41] Y. Rao, Y.-S. Tao, and H.-F. Wang, J. Chem. Phys. **119**, 5226 (2003).

[42] H. J. Berendsen, J. R. Grigera, and T. P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).

[43] S. Plimpton, J. Comp. Phys. **117**, 1 (1995), http://lammps.sandia.gov.

[44] J.-P. Ryckaert, G. Ciccotti, and H. J. Berendsen, J. Comput. Phys. **23**, 327 (1977).

[45] w. G. Hoover, Phys. Rev. A **31**, 1695 (1985).

[46] W. C. Swope and H. C. Andersen, J. Chem. Phys. **76**, 637 (1982).

[47] G. Hummer, L. R. Pratt, A. E. Garca, B. J. Berne, and S. W. Rick, J. Phys. Chem. B **101**, 3017 (1999).

[48] R. Rey and J. T. Hynes, J. Chem. Phys. **108**, 142 (1998).

[49] Q. Du, R. Superfine, E. Freysz, and Y. R. Shen, Phys. Rev. Lett. **70**, 2313 (1993).

[50] Y. R. Shen and V. Ostroverkhov, Chem. Rev. **106**, 1140 (2006).

[51] M. Sovago, R. K. Campen, G. W. Wurpel, M. Müller, H. J. Bakker, and M. Bonn, Phys. Rev. Lett. **100**, 17390 (2008).

[52] This separation of interfacial and bulk length scales rests on the assumption that systems of interest are well removed from critical points, as is certainly the case for liquid water at ambient conditions.

[53] Although we notionally generated the configuration $\overline{\Gamma}$ by reflection, we nonetheless define a molecular reference frame for each of its molecules as we would in the case of Γ. The resulting coordinate system is therefore right-handed (which would not be the case had we transformed $\hat{\mathbf{w}}$ in the same way as $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$). As we have described it, this symmetry argument holds only for achiral molecules such as HOD.

[54] Periodic boundary conditions prevent rigorous association of a molecule and all its images with one side of a dividing surface. However, our simulations include ample space, usually devoid of mass, between replicated liquid regions. We can thus unambiguously identify a liquid subsystem whose center of mass $z^{(\mathrm{c.o.m.})}$ in the z-direction is well-defined. When we segregate molecules using a dividing surface, we refer to their periodic images whose z-displacements from $z^{(\mathrm{c.o.m.})}$ are smaller than $L_z/2$.

[55] We judge a hydroxyl group to be dangling according to the distance r_{OH} between the proton and the nearest oxygen atom within a cone of half-width $\pi/6$ emanating from the proton in the direction of the OH bond vector. Values of r_{OH} much greater than the range of typical bulk fluctuations (say, two standard deviations) can safely be considered to lack a hydrogen bond acceptor.

[56] Eq. 6 permits a straightforward modification to include non-Condon effects through electric field dependence of the transition dipole.

VII. FIGURES

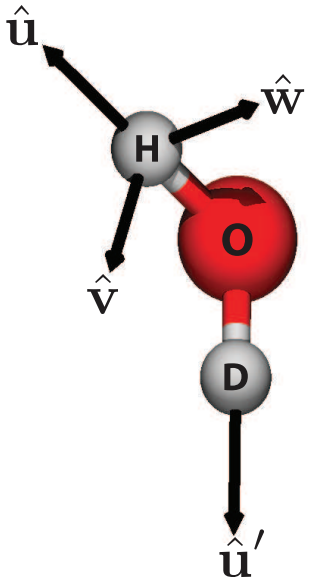

FIG. 1: Rectilinear coordinate system specifying a molecular reference frame for HOD. $\hat{\mathbf{u}}$ points along the OH bond. $\hat{\mathbf{v}}$ also lies in the plane of the molecule and can therefore be considered a linear combination of $\hat{\mathbf{u}}$ and a unit vector $\hat{\mathbf{u}}'$ pointing along the OD bond. Defined as the vector product of $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$, the third basis vector $\hat{\mathbf{w}}$ is perpendicular to the plane of the molecule.

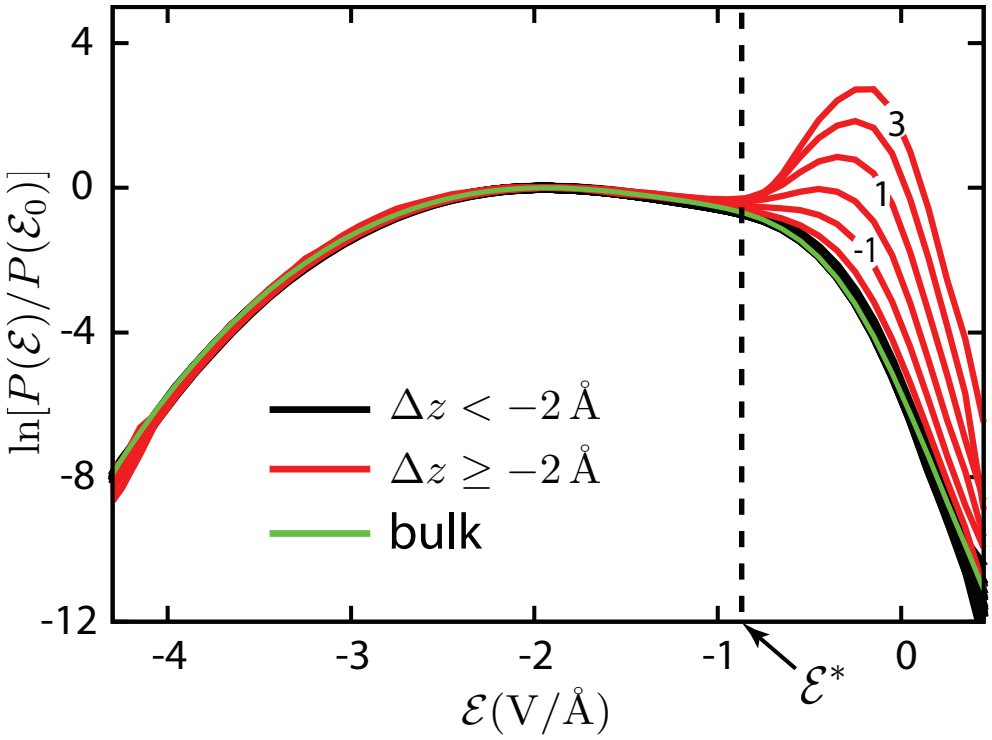

FIG. 2: Depth-resolved distributions of $\mathcal{E}$, the electric field on an H atom projected onto the corresponding OH bond vector, for dilute HOD in D_2O. Each curve reports molecular dynamics results for water molecules in a 1 Å-thick slice of the slab parallel to the interface. Black lines show results for slices more than 2Å below the Gibbs dividing surface. Red lines show results for the interfacial region, specifically for slices centered at $\Delta z = -2$Å, $\Delta z = -1$Å, $\Delta z = 0$, $\Delta z = 1$Å, $\Delta z = 2$Å, and $\Delta z = 3$Å. (A few of these lines are labeled by the corresponding value of Δz.) We plot the logarithm of each distribution, scaled by its value at $\mathcal{E}_0$, where the bulk distribution (green) is maximum. The dashed line at $\mathcal{E} = \mathcal{E}^*$ distinguishes a range where electric field statistics are unaltered by the interface from a range that shows dramatic change near the Gibbs dividing surface.

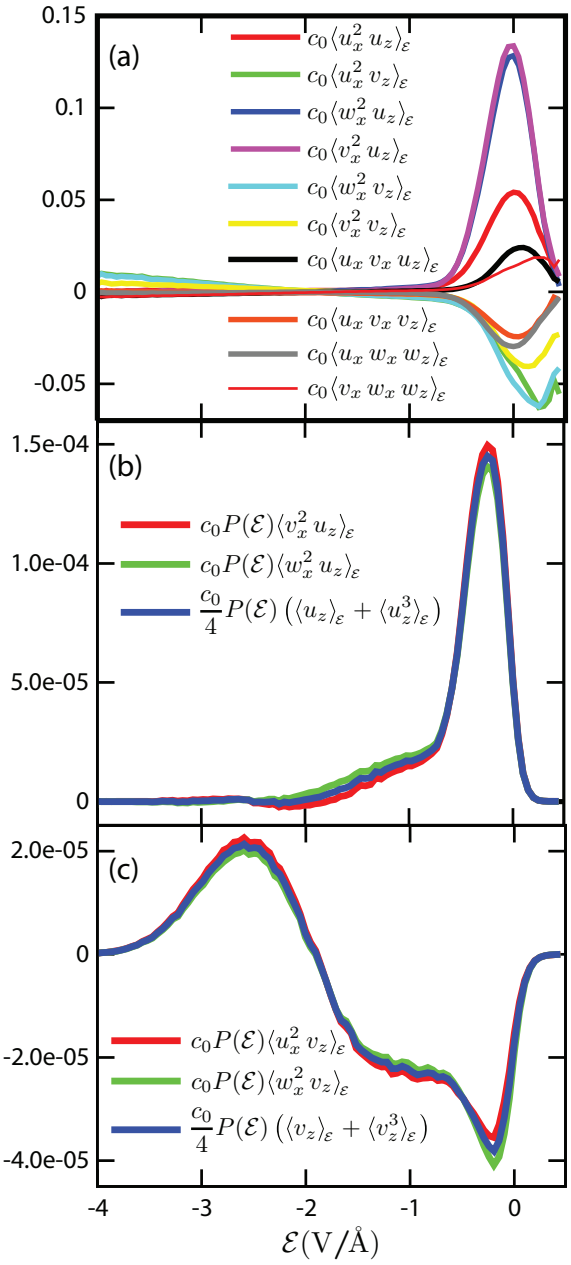

FIG. 3: Nonvanishing orientational averages contributing to the sum in Eq. 6, computed from Monte Carlo simulations of dilute HOD in D_2O. All ten averages are plotted in (a) as functions of the electric field variable $\mathcal{E}$. Each has been multiplied by the factor $c_0 \equiv N/(2L_xL_y)\,\text{Å}^{-2}$ in order to remove a trivial dependence on system size. Bottom two panels show selected averages, weighted by the electric field distribution $P(\mathcal{E})$, emphasizing the weakness of orientational bias at the interface. We show $c_0P(\mathcal{E})\langle v_x^2 u_z\rangle_\mathcal{E}$ and $c_0P(\mathcal{E})\langle w_x^2 u_z\rangle_\mathcal{E}$ in (b), along with the estimate $(c_0/2)P(\mathcal{E})(\langle u_z\rangle_\mathcal{E} + \langle u_z^3\rangle_\mathcal{E})$ from Eq. 9. Similarly, in (c) we show $c_0P(\mathcal{E})\langle u_x^2 v_z\rangle_\mathcal{E}$ and $c_0P(\mathcal{E})\langle w_x^2 v_z\rangle_\mathcal{E}$, along with the estimate $(c_0/2)P(\mathcal{E})(\langle v_z\rangle_\mathcal{E} + \langle v_z^3\rangle_\mathcal{E})$ from Eq. 12.

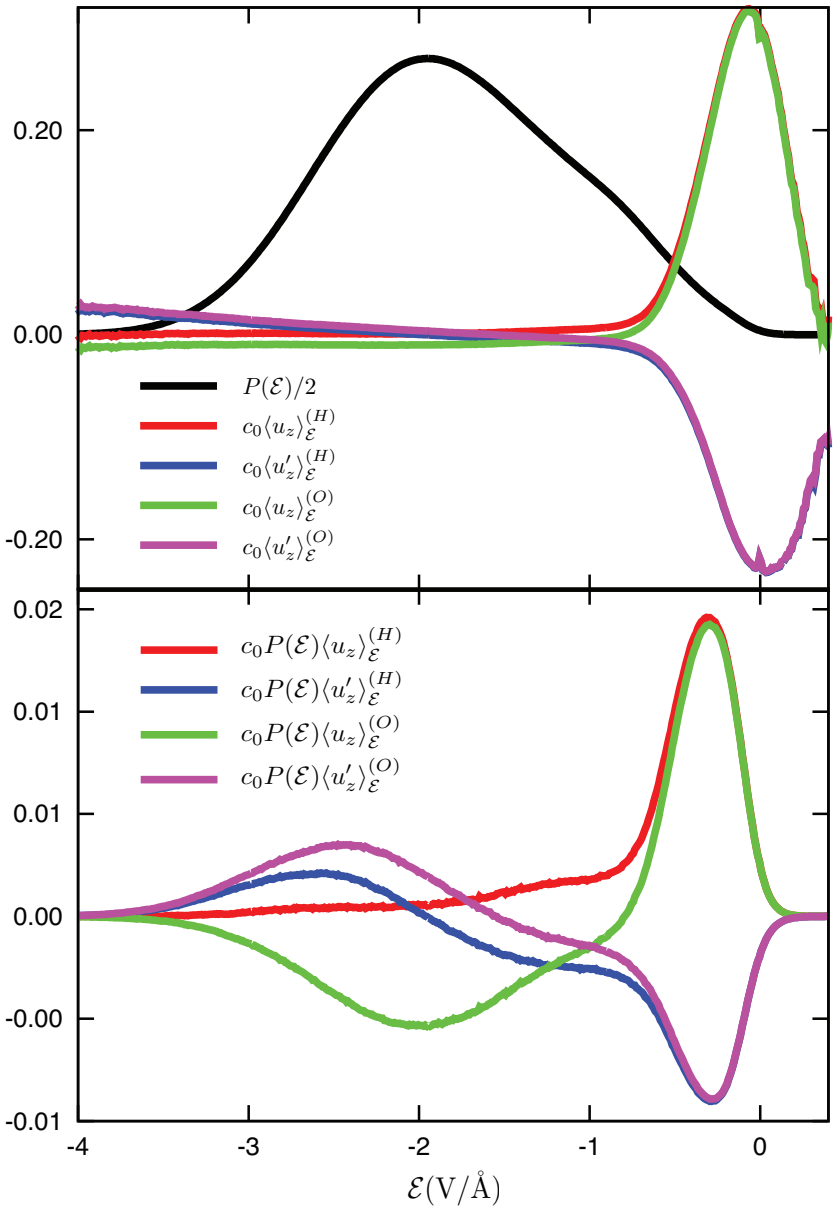

FIG. 4: Average projections of OH and OD bond vectors onto the surface normal $\hat{\mathbf{z}}$ for dilute HOD in D_2O. Panel (a) shows $c_0\langle u_z\rangle_{\mathcal{E}}$ and $c_0\langle u'_z\rangle_{\mathcal{E}}$ as functions of electric field. We include results for both of the symmetry-breaking conventions described in Sec. III: The superscripts (O) and (H) indicate that the liquid slab was divided into separate interfacial regions according to oxygen and hydrogen atom positions, respectively. As in Fig. 3 and in subsequent figures, the factor $c_0 = N/(2L_xL_y)\text{Å}^{-2}$ renders calculated averages as intensive quantities. The bulk electric field distribution, scaled by an arbitrary constant, is also shown for reference. Each of these orientational averages is weighted in (b) by the electric field distribution $P(\mathcal{E})$. Because $\langle u_z\rangle$ is small in magnitude near the peak of $P(\mathcal{E})$, minor differences due to choice of convention effect significant shifts in the weighted average.

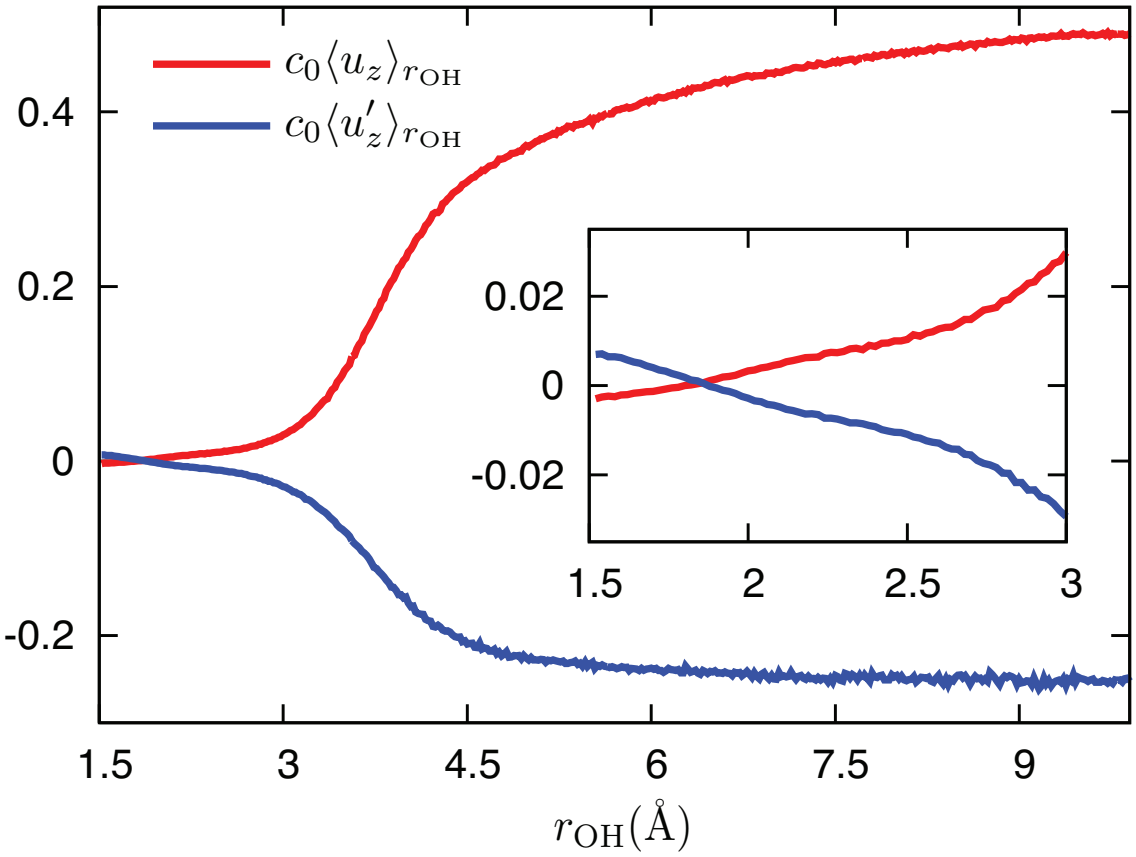

FIG. 5: Average projections of OH and OD bond vectors onto the surface normal, resolved by OH$\cdots$O distance. Black curve shows $c_0\langle u_z\delta(r_{\rm OH}-r_{\rm OH}(\Gamma))\rangle/\langle\delta(r_{\rm OH}-r_{\rm OH}(\Gamma))\rangle$ for dilute HOD in D_2O. Red curve shows $c_0\langle u'_z\delta(r_{\rm OH}-r_{\rm OH}(\Gamma))\rangle/\langle\delta(r_{\rm OH}-r_{\rm OH}(\Gamma))\rangle$ for the same system. The inset provides a magnified view of these functions for distances typical of intact hydrogen bonds.