

# *De Novo* Determination of Molecular Orientation at Interfaces through Multimode Polarization-Dependent Sum-Frequency Generation Spectroscopy

*Chun-Chieh Yu,<sup>1,#</sup> Sho Imoto,<sup>2,#</sup> Takakazu Seki,<sup>1</sup> Kuo-Yang Chiang,<sup>1</sup> Shumei Sun,<sup>3</sup> Mischa Bonn,<sup>1</sup> and Yuki Nagata<sup>1,\*</sup>*

1. Molecular Spectroscopy Department, Max Planck Institute for Polymer Research, Ackermannweg 10, Mainz 55128, Germany
2. Analysis Technology Center, Fujifilm R&D, 210 Nakanuma, Minamiashigara, Kanagawa, 250-0123, Japan
3. Department of Physics, Applied Optics Beijing Area Major Laboratory, Beijing Normal University, 100875 Beijing, China

## **Corresponding Author**

\* [nagata@mpip-mainz.mpg.de](mailto:nagata@mpip-mainz.mpg.de)

#These authors contributed to this work equally

## **Abstract**

Many essential processes occur at soft interfaces, from chemical reactions on aqueous aerosols in the atmosphere to biochemical recognition and binding at the surface of cell membranes. The spatial arrangement of molecules specifically at these interfaces is crucial for many of such processes. The accurate determination of the interfacial molecular orientation has been challenging due to the low number of molecules at interfaces and the ambiguity of their orientational distribution. Here, we combine phase- and polarization-resolved sum-frequency generation spectroscopy to obtain the molecular orientation at the interface. We extend an exponentially decaying orientational distribution to multiple dimensions, which, in conjunction with multiple SFG data sets obtained from the different vibrational modes, allows us to determine molecular orientation. We apply this new approach to formic acid molecules at the air-water interface. The inferred orientation of formic acid agrees very well with *ab initio* molecular dynamics data. The phase-resolved SFG multimode analysis scheme using the multi-dimensional orientational distribution thus provides a universal approach for obtaining the interfacial molecular orientation.

## I. INTRODUCTION

The surfaces of soft matter and liquid systems are ubiquitous in nature and technology. Typically, a strong connection exists between the molecular-level structure at the interface or surface and the macroscopic properties. For instance, for foams, the nature and density of surfactants determine surface tension and foam stability;<sup>1</sup> for liquid aerosols in the atmosphere, the aerosol surface composition determines the evolution of photochemical processes occurring on the surface.<sup>2</sup> Hence, knowledge of the composition and spatial arrangement of interfacial molecules is highly desirable.

Sum-frequency generation (SFG) spectroscopy is a surface-specific second-order optical technique that provides molecular responses from interfaces.<sup>3-6</sup> The SFG signal is generated by shining infrared (IR) and visible pulses onto the sample. The SFG signal is enhanced when the IR frequency is resonant with the molecular vibrations, providing molecular specificity. As such, SFG spectra can selectively probe the specific moiety of the molecules located at the interface. SFG signals recorded with different polarization combinations for the SFG, visible and infrared fields further contain the orientational information of interfacial molecular groups, accessible through their vibrations.<sup>7-17</sup> For example, Shen and Wang groups explored the orientation of the free O-H groups of the interfacial water molecules.<sup>8-11,18,19</sup> Tyrode, Johnson, and co-workers investigated the molecular orientation of several organic molecules at the air-water interface.<sup>20-23</sup> Such analyses have previously also been applied to study the interfacial conformations of proteins.<sup>24-28</sup>

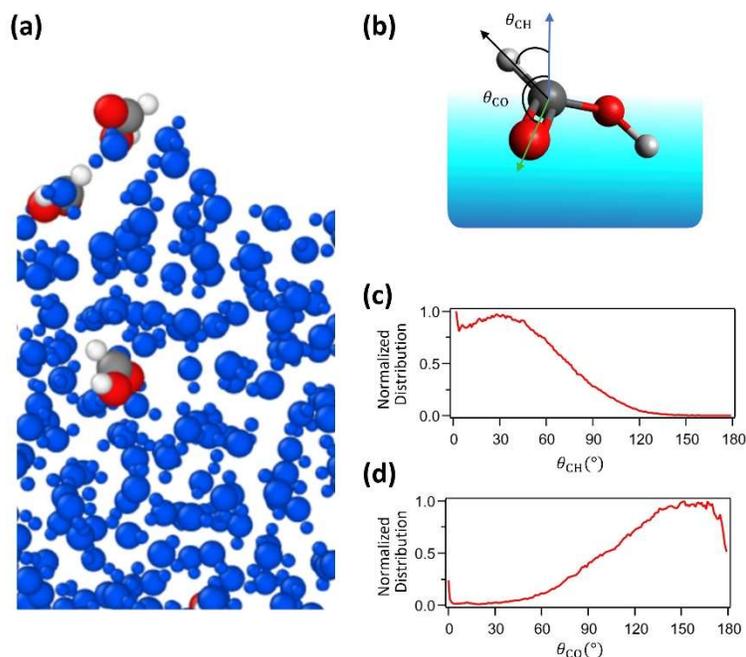
Despite these applications of the orientational analysis through SFG, it remains challenging to accurately predict the interfacial molecular orientation,<sup>29</sup> because extracting the molecular orientation from the SFG signals poses two major challenges. The first challenge arises from the difficulty in accurately determining the peak amplitudes in the absorptive ( $\text{Im}\chi^{(2)}$ ) spectra;

conventional SFG measurements provide the intensity ( $|\chi^{(2)}|^2$ ), which does not give unique access to  $\text{Im}\chi^{(2)}$  spectra. Thanks to the recent development of the heterodyne-detected SFG (HD-SFG) measurement,<sup>30-32</sup> one can unambiguously obtain the peak amplitude in the  $\text{Im}\chi^{(2)}$  spectra.<sup>12,14,33</sup> The second challenge originates from the orientational distribution. To extract information on the molecular orientation from the SFG peak amplitude, one must assume an orientational distribution function of the molecular moiety at the interface. The choice of the orientational distribution function critically affects the inferred molecular orientation.<sup>29,34,35</sup> Previous studies have used rectangular functions,<sup>19,28,33</sup> Gaussian-shaped functions<sup>8,18</sup> and the exponential decay functions<sup>35</sup> to describe the orientational distribution. For the free O-H groups sticking out from a water surface, a comparison of the simulation and experiment as indicated that the free O-H groups at the air-water<sup>35,36</sup> and air-ice interfaces<sup>37</sup> have exponential decay-shaped distributions.

Can such an exponential decay function be universally used for describing the orientational distributions of different interfacial molecules? Molecular dynamics (MD) simulations have shown that an exponential decay function seems not universally applicable.<sup>29,38-40</sup> For example, we carried out an *ab initio* MD (AIMD) simulation of formic acid molecules at the air-water interface (Fig. 1(a)) and calculated the distributions of  $\theta_{\text{CH}}$  and  $\theta_{\text{CO}}$ , where  $\theta_{\text{CH}}$  ( $\theta_{\text{CO}}$ ) represents the angle formed by the C-H (C=O) group and the surface normal (Fig. 1(b)). Figs. 1(c) and (d) show the simulated  $\theta_{\text{CH}}$  and  $\theta_{\text{CO}}$  distributions of interfacial formic acid molecules. These distributions deviate from an exponential decay-shape. Here, a question is what causes the deviation of the distribution function from an exponential decay-shape and how we can construct the orientational distribution functions.

Here, we hypothesized that the deviation of the distribution function from an exponential decay-shape arises from the fact that the C-H group and C=O group are not independent, because

the intramolecular H-C=O angle is fixed at  $\sim 120^\circ$ . Based on this hypothesis, we develop a methodology to couple the individual orientational information obtained from the multimode SFG probes via the multi-dimensional orientational distribution. We show how our multimode SFG coupling scheme can predict the molecular orientations by targeting a formic acid molecule located at the air-water interface. We benchmark the experimental data against data obtained from AIMD simulations.



**FIG. 1. Conformation of formic acid molecules at the air-water interface predicted by AIMD simulation.** (a) Snapshot of simulated air-formic acid solution interface. White, red, and grey spheres represent the hydrogen, oxygen, and carbon atoms of formic acid molecules, respectively, while water molecules are drawn as blue spheres. (b) Schematic of a formic acid molecule at the air-water interface. The blue arrow represents the surface normal. The black and green arrows represent the C→H and C→O vectors, respectively. The angle  $\theta_{CH}$  ( $\theta_{CO}$ ) is formed by the C→H (C→O) vector and the surface normal. (c,d) The orientational distributions of (c)  $\theta_{CH}$  and (d)  $\theta_{CO}$  obtained from the AIMD simulation.

## II. THEORY FOR ACCESSING MOLECULAR ORIENTATION FROM MULTI-MODE SFG

The following four steps are required to access the angle  $\theta_{\text{CH}}$  ( $\theta_{\text{CO}}$ ) from the SFG spectra: (i) Measuring the complex  $\chi^{(2)}$  spectra at the *ssp* and *sps* polarization combinations, where *abc* polarization represents the *a*-, *b*-, and *c*-polarizations of SFG, visible, and IR beams, respectively.

(ii) Obtaining  $\chi_{yzy}^{(2)}$  and  $\chi_{zyy}^{(2)}$  spectra from  $\chi_{\text{eff},ssp}^{(2)}$  and  $\chi_{\text{eff},sps}^{(2)}$  via;

$$\chi_{\text{eff},ssp}^{(2)}(\omega_{\text{IR}}; \omega_{\text{SFG}}, \omega_{\text{Vis}}) = L_{yy}(\omega_{\text{SFG}})L_{yy}(\omega_{\text{Vis}})L_{zz}(\omega_{\text{IR}})\sin\beta_{\text{IR}}\chi_{yzy}^{(2)}(\omega_{\text{IR}}), \quad (1)$$

$$\chi_{\text{eff},sps}^{(2)}(\omega_{\text{IR}}; \omega_{\text{SFG}}, \omega_{\text{Vis}}) = L_{yy}(\omega_{\text{SFG}})L_{zz}(\omega_{\text{Vis}})L_{yy}(\omega_{\text{IR}})\sin\beta_{\text{Vis}}\chi_{zyy}^{(2)}(\omega_{\text{IR}}), \quad (2)$$

where  $\beta$  is the incident angle of the light of frequency  $\omega$  with respect to the surface normal and  $L_{ii}$  ( $i = x, y, z$ ) is the Fresnel factor.<sup>41</sup> The *xz*-plane forms the incident beam plane, and the *z*-axis forms the macroscopic surface normal. (iii) Obtaining  $\langle \cos\theta \rangle / \langle \cos^3\theta \rangle$  from the peak amplitudes  $A_{yzy}$  and  $A_{yyz}$  in the  $\text{Im}\chi_{zyy}^{(2)}$  and  $\text{Im}\chi_{yzy}^{(2)}$  spectra via;<sup>19</sup>

$$\frac{A_{yzy}}{A_{yyz}} \propto \frac{(1-r)(\langle \cos\theta \rangle - \langle \cos^3\theta \rangle)}{(1+r)\langle \cos\theta \rangle - (1-r)\langle \cos^3\theta \rangle}, \quad (3)$$

where  $r$  represents the depolarization ratio, and the values of  $r$  were obtained from *ab initio* calculations (see Supplementary material). (iv) Obtaining any angle information, including average angle  $\langle \theta \rangle$  from  $\langle \cos\theta \rangle / \langle \cos^3\theta \rangle$  value by assuming the orientational distribution function,  $f(\theta)$ . Note that the use of the *sps* configuration has the advantage over the use of the *ppp* configuration in this step, in that the amplitude ratio of  $A_{yzy}/A_{yyz}$  is much less sensitive to the

poorly defined interfacial dielectric constant in the Fresnel factor correction (appearing in the next step) than the ratio of  $A_{yyz}/A_{zzz}$ .<sup>42</sup>

A key part of rendering the SFG data into the angular information is step (iv). When a single driving force generates a preferential interfacial orientation, the orientational distribution can often be well approximated by an exponentially decaying function.<sup>35-37</sup> For instance, for an interfacial water molecule with the free O-H group, the other half O-H group is usually hydrogen-bonded (H-bonded), which is the driving force for determining the orientation of the free O-H group.<sup>43,44</sup> In contrast, when a complex molecule such as a biomolecule can have several energetically stable conformations, it possesses multiple driving forces to govern the molecular orientation. As a result, its orientational distribution cannot be described by one exponential decay function; rather, the different driving forces will each drive towards a different exponential function. To couple multiple orientation distribution for different moieties and include the geometrical constraint for the intramolecular conformation, we propose to couple multiple vibrational mode SFG data by introducing the multi-dimensional orientational distribution function. For formic acid, we consider the two-dimensional orientational distribution,  $f(\theta_{CH}, \theta_{CO})$ .  $\theta_{CH}$  and  $\theta_{CO}$  cannot be determined independently, because they should satisfy the condition that the intramolecular H-C=O angle is  $\sim 120^\circ$ . By taking this H-C=O angle constraint into account within the free rotation model (see Supplementary material), the orientational distribution function can be given as;

$$f(\theta_{CH}, \theta_{CO}) = N_E \exp(-\theta_{CH}/\theta_{E,CH}) \exp(-\theta_{CO}/\theta_{E,CO}) g(\theta_{CH}, \theta_{CO}), \quad (4)$$

based on the exponential decay distribution function,<sup>35</sup> where  $N_E$  is the normalization factor,  $\theta_E$  is a parameter determining the steepness/width of the exponential decay function, and  $g(\theta_{CH}, \theta_{CO})$

represents the geometric constraint which  $\theta_{\text{CH}}$  and  $\theta_{\text{CO}}$  should satisfy, given that the intramolecular H-C=O angle is  $\sim 120^\circ$ .

### III. EXPERIMENTAL and SIMULATION PROCEDURES

#### A. HD-SFG Experiment

For probing the C=O stretch mode, a visible pulse centered at  $\sim 800$  nm and a tunable IR pulse were combined collinearly. They were first focused onto a  $20 \mu\text{m}$ -y-cut quartz plate to generate the local oscillator (LO) signal. Subsequently, they passed through a  $8$  mm CaF<sub>2</sub> plate for the phase modulation and were again focused onto the sample surface with their incident angles of  $45^\circ$ . The SFG signal from the sample interfered with the LO beam, generating the SFG interferogram. The SFG interferogram was decomposed into the frequency domain through a spectrometer (SpectraPro HRS-300, Princeton Instruments) and detected by a liquid N<sub>2</sub>-cooled CCD (PyLoN®, Princeton Instruments). To avoid the effect of water vapor absorption of IR, the optical path of SFG setup was purged with N<sub>2</sub>.

For probing the C-H stretch mode, we used a visible pulse centered at  $\sim 800$  nm and a tunable IR pulse with their incident angles of  $64^\circ$  and  $50^\circ$ , respectively. The IR and visible beam were firstly focused onto a  $200$  nm ZnO thin film deposited on a  $1$  mm CaF<sub>2</sub> window to generate the LO signal. A  $1.5$  mm fused silica plate was inserted in the optical path of the LO beam. The IR, visible, and LO beams were then focused onto the sample interface. The SFG signal from the sample interfered with the LO signal, generating the SFG interferogram. The SFG interferogram was decomposed into the frequency domain through a spectrometer (Shamrock 303i, Andor Technology), and subsequently detected by an EMCCD camera (Newton, Andor Technology). The details can be found in the material.

## B. MD Simulation

The force field MD and AIMD simulations were performed using the CP2K package.<sup>45</sup> The simulation cell contained 5 formic acid and 195 water molecules. The system is set into a rectangular cell of  $16.63 \text{ \AA} \times 16.63 \text{ \AA} \times 40.0 \text{ \AA}$ . We prepared all five *cis*-conformations of formic acid molecules as well as all five *trans*-conformations. We generated equilibrated 20 initial configurations for each conformation by using force field. These initial configurations were further equilibrated by AIMD for 10 ps. Subsequently, we ran 20 ps AIMD simulation. We then generated the total 400 ps AIMD trajectories at the revPBE level level together with the Grimmes' van der Waals correction.<sup>46</sup> We integrated the equation of motion in the NVT ensemble with a time step of 0.5 fs. The target temperature was set to 300 K. The details can be found in the Supplementary material.

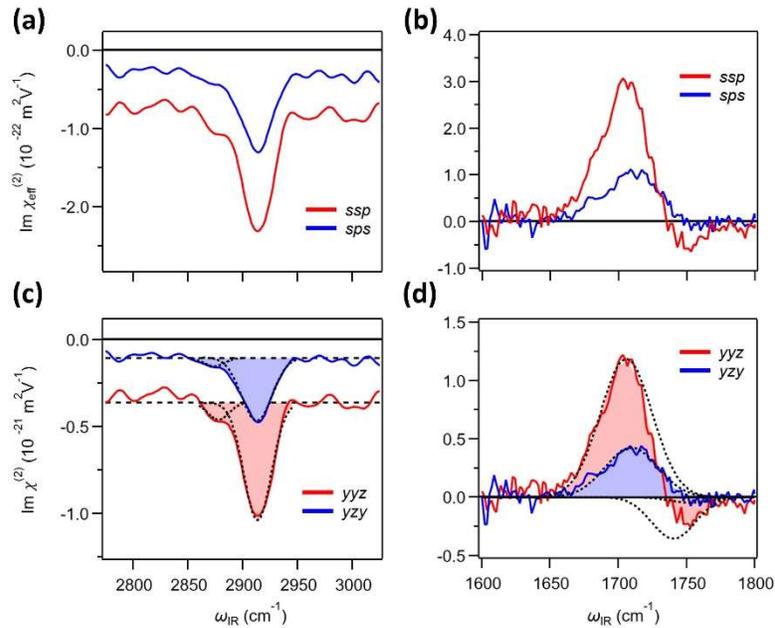
## III. RESULTS

### A. Polarization-Dependent Multi-Mode HD-SFG Spectra

We examined how accurately this multimode coupling scheme can predict the molecular orientation of formic acid molecules at the air-water interface. Fig. 2(a) shows the  $\text{Im}\chi_{\text{eff,ssp}}^{(2)}$  SFG response of the C-H stretch mode of a 2.5% molar fraction aqueous formic acid solution. The spectrum shows a  $2920 \text{ cm}^{-1}$  negative peak and a  $2880 \text{ cm}^{-1}$  negative shoulder peak. These peaks can be assigned to the C-H stretch mode of the *trans*- and *cis*-conformations of formic acid, respectively.<sup>20,38,47,48</sup> Thus, our data indicates that both *trans*- and *cis*-conformations are present at the air-water interface. The presence of the *trans*- and *cis*-conformations is corroborated by the

$\text{Im}\chi_{\text{eff},ssp}^{(2)}$  spectrum of the C=O stretch mode (Fig. 2(b)). The spectrum shows a 1710  $\text{cm}^{-1}$  positive feature and a 1750  $\text{cm}^{-1}$  negative feature which can be assigned to the *trans*- and *cis*-conformations of a formic acid molecule at the air-water interface, respectively.<sup>49</sup>

The signs of the peaks in the  $\text{Im}\chi_{\text{eff},ssp}^{(2)}$  spectra provide information on the absolute orientation of the C→H and C→O groups. The negative C-H peaks indicate that the C→H groups of both *trans*- and *cis*-conformations point *up* towards the air.<sup>33</sup> This can be explained by the inability of the hydrophobic C-H group to form a hydrogen bond with water. In the case of the C=O stretch mode, a positive (negative) sign indicates that the C→O group points *down* to the bulk water (*up* to the air).<sup>50</sup> The 1710  $\text{cm}^{-1}$  positive feature indicates that the C→O group of the *trans*-conformations points *down* at the air-water interface, while the 1750  $\text{cm}^{-1}$  negative feature indicates that the C→O group of the *cis*-conformation points *up*. The *trans*- and *cis*-conformations of the formic acid have different orientations at the air-water interface.



**FIG. 2. SFG spectra of formic acid at the air-water interface.**  $\text{Im}\chi_{\text{eff},ssp}^{(2)}$  and  $\text{Im}\chi_{\text{eff},sps}^{(2)}$  spectra in (a) the C-H stretch and (b) C=O stretch mode regions.  $\text{Im}\chi_{yyz}^{(2)}$  and  $\text{Im}\chi_{yzy}^{(2)}$  spectra in (c) the C-H stretch and (d) C=O stretch mode regions. The dotted line represents Gaussian lineshapes obtained from the fit to the spectra, while the filled area represents the sum of the two Gaussians.

To further understand the orientation of the molecules, we measured the  $\text{Im}\chi_{\text{eff},sps}^{(2)}$  spectra of the C-H and C=O stretch modes. The data are plotted in Figs. 2(a) and (b), respectively. Subsequently, we obtained  $\text{Im}\chi_{yyz}^{(2)}$  and  $\text{Im}\chi_{yzy}^{(2)}$  spectra using Eqs. (1) and (2), which are displayed in Figs. 2(c) and 2(d), respectively. The  $\text{Im}\chi_{yyz}^{(2)}$  and  $\text{Im}\chi_{yzy}^{(2)}$  features are very similar to the  $\text{Im}\chi_{\text{eff},ssp}^{(2)}$  and  $\text{Im}\chi_{\text{eff},sps}^{(2)}$  features, respectively. The comparison of the  $\text{Im}\chi_{yyz}^{(2)}$  and  $\text{Im}\chi_{yzy}^{(2)}$  indicates that the amplitude of the spectra at different polarization differs significantly.

## B. C=O and C-H Orientation of Interfacial Formic Acid Molecule

Here, we focus on the *trans*-conformation of the formic acid (1710  $\text{cm}^{-1}$  positive C=O stretch and 2920  $\text{cm}^{-1}$  negative C-H stretch features). From the  $\text{Im}\chi_{yyz}^{(2)}$  and  $\text{Im}\chi_{yzy}^{(2)}$  spectra, we obtained the values of  $A_{yzy}/A_{yyz}$  of  $0.60 \pm 0.01$  for the C-H stretch mode and of  $0.36 \pm 0.01$  for the C=O stretch, which provides the left side of Eq. (3). The right side of Eq. (3) can be obtained using  $f(\theta_{CH}, \theta_{CO})$ , where the ensemble average of  $B$  is given by;

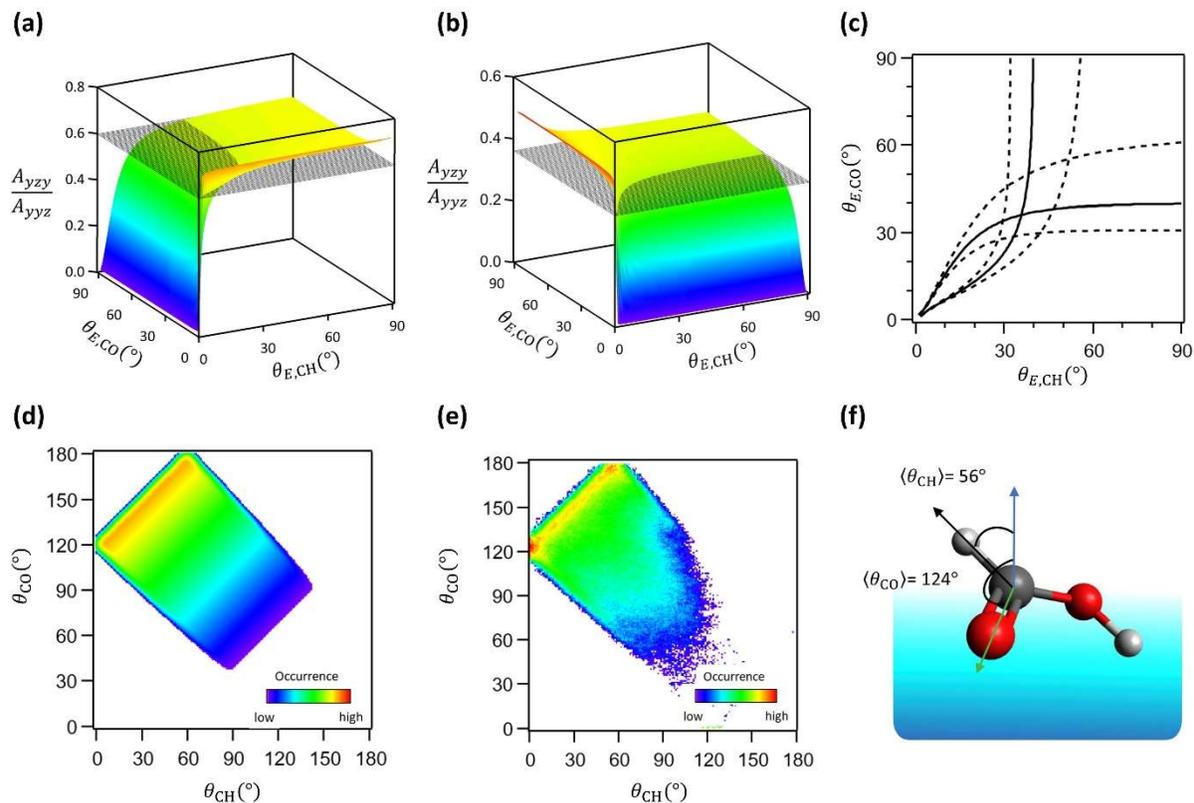
$$\langle B \rangle = \int_0^\pi d\theta_{CO} \int_0^\pi d\theta_{CH} B f(\theta_{CH}, \theta_{CO}) \sin\theta_{CH} \sin\theta_{CO}.$$

(5)

Here, we replaced  $\theta_{\text{CO}}$  by  $180^\circ - \theta_{\text{CO}}$  in Eq. (3) by knowing that the C→O group of the *trans*-conformation points *down* to the bulk water. Eq. (3) will determine the values of parameters,  $\theta_{\text{E,CH}}$  and  $\theta_{\text{E,CO}}$ .

Figs. 3(a) and (b) display the  $A_{yzy}/A_{yyz}$  values calculated for various  $\theta_{\text{E,CH}}$  and  $\theta_{\text{E,CO}}$  via the right side of Eq. (3) (rainbow curves) and the experimentally determined  $A_{yzy}/A_{yyz}$  (grey planes). The crossing lines of the rainbow curves and grey planes in Figs. 3(a) and (b) represent the condition that  $\theta_{\text{E,CH}}$  and  $\theta_{\text{E,CO}}$  should satisfy in the C-H and C=O stretch modes, respectively. By coupling these crossing curves, one can find a crossing point (Fig. 3(c)). As such, we obtained  $\theta_{\text{E,CH}} = 36 \pm 9^\circ$  and  $\theta_{\text{E,CO}} = 36 \pm 10^\circ$ .

Fig. 3(d) shows the two-dimensional orientational distribution  $f(\theta_{\text{CH}}, \theta_{\text{CO}})$  for  $\theta_{\text{E,CH}} = 36^\circ$  and  $\theta_{\text{E,CO}} = 36^\circ$ . The orientational distribution obtained from the experimental data agrees with that obtained from the AIMD simulation data (Fig. 3(e)). This shows that our multimode coupling scheme can accurately predict the orientation of the formic acid molecules. The obtained distribution functions provide  $\langle \theta_{\text{CH}} \rangle = 56 \pm 5^\circ$  and  $\langle \theta_{\text{CO}} \rangle = 124 \pm 5^\circ$ . The difference in the angle between  $\langle \theta_{\text{CH}} \rangle$  and  $\langle \theta_{\text{CO}} \rangle$  being much smaller than the H-C=O angle of  $120^\circ$  indicates that the H-C=O plane is strongly tilted away from the surface normal, tending toward being parallel to the surface. The summary of the *trans*-conformation of the interfacial formic acid molecule is shown in Fig. 3(f).

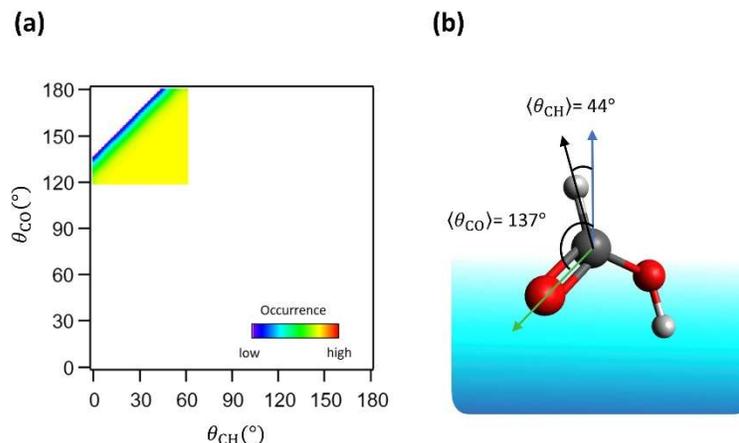


**FIG. 3. C→H and C→O angle analysis based on exponential decay-shaped function.**  $A_{yzy}/A_{yyz}$  vs.  $\theta_{E,CH}$  and  $\theta_{E,CO}$  for (a) the C-H stretch mode and (b) the C=O stretch mode. The rainbow 3D curves represent the numerical data based on Eq. (3), while the grey planes represent the experimental values. (c) The lines obtained from the crossing of rainbow 3D curves and grey planes in (a) and (b). The dotted lines represent the experimental error. (d) The 2D orientational distributions inferred from the crossing point of (c). (e) The 2D orientational distributions obtained from the AIMD simulation. (f) Schematic of the average orientation of a formic acid molecule at the air-water interface. The blue arrow represents the surface normal. The black and green arrows represent the C→H and C→O vectors, respectively.

Above, we have hypothesized that the H-C=O angle constraint is essential to estimate the orientation of the formic acid molecules and can be modeled based on the exponentially decay-shaped function. Here, we examine these assumptions in more detail. First, to explore the impact of the H-C=O angle constraint, we considered the distribution without the H-C=O angle constraint.

In this case, the 3D curve based on Eq. (3) does not cross the experimentally obtained  $A_{yzy}/A_{yyz}$  value (see Supplementary material). This result clearly shows that the intuitive constraint of intramolecular geometry is essential to obtain the orientational distribution function.

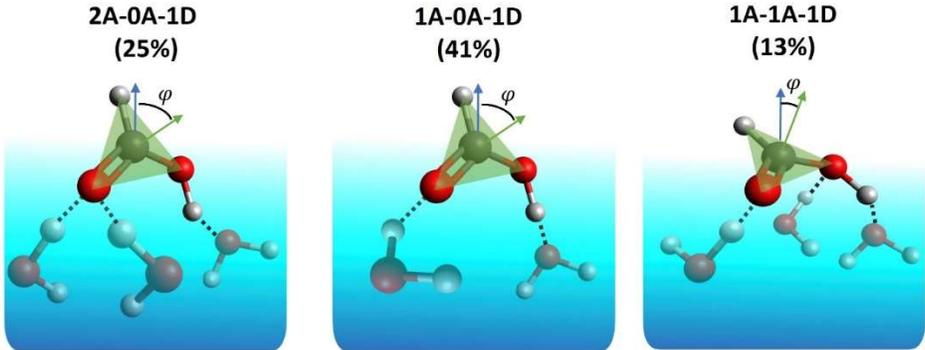
Second, to verify the appropriateness of the exponential decay-shaped distribution, rather than the rectangular function, we computed the distribution functions based on the rectangular function, including the H-C=O constraint.<sup>19</sup> The obtained distribution function based on the rectangular function is displayed in Fig. 4(a). It differs substantially from those obtained from the AIMD simulation (Fig. 3(e)). The rectangular distribution provides  $\langle\theta_{CH}\rangle = 44 \pm 3^\circ$  and  $\langle\theta_{CO}\rangle = 137 \pm 3^\circ$ , quite different from the  $\langle\theta_{CH}\rangle = 56^\circ$  and  $\langle\theta_{CO}\rangle = 124^\circ$  predicted using the exponential decay distribution. The smaller  $\langle\theta_{CH}\rangle$  and larger  $\langle\theta_{CO}\rangle$  predicted by the rectangular distribution function indicates that the H-C=O plane prediction is closer to the surface normal. As such, the predicted three-dimensional structure of formic acid with the rectangular distribution function differs that with the exponential decay function or simulated structure in the AIMD simulation.



**FIG. 4. C→H and C→O angle analysis based on rectangular function. a** The obtained 2D orientational distribution based on the rectangular function. **b** Schematic of the average orientation of a formic acid molecule at the air-water interface, assuming a rectangular distribution function. The blue arrow represents the surface normal. The black and green arrows represent the C→H and C→O vectors, respectively.

We estimate the interaction of the interfacial formic acid molecule with surrounding water from the conformation of the formic acid molecule. To do so, we calculated the angle ( $\varphi$ ) formed by the normal to the H-C=O plane and the surface normal from  $\langle \theta_{CH} \rangle$  and  $\langle \theta_{CO} \rangle$ . The exponential decay distribution provides  $\varphi = 40^\circ$ . This  $\varphi$ -value indicates that the H-C=O plane is not so perpendicular to the surface; rather, the H-C=O plane lies down to the surface. This conformation allows a formic acid molecule to form a hydrogen bond between the O(-H) atom and the H atom of the surrounding water molecule, in addition to the hydrogen bonds of O(=C) atom and H(-O) atom of formic acid molecule (see Fig. 5). In fact, the hydrogen bond analysis<sup>51</sup> using the AIMD trajectory indicates three major conformations of the hydrogen bond of the interfacial formic acid molecules: 2A-0A-1D conformation, i.e., the O(=C), O(-C), H(-O) atoms respectively accept two H-bonds, accept no H-bonds, and donate one H-bond with the surrounding water molecules, in the

same notation, 1A-0A-1D, and 1A-1A-1D conformations (Fig. 5). As such, the *trans*-conformation of the formic acid can have a variety of hydrogen bond conformation at the air-water interface.



**FIG. 5. Schematics of major hydrogen bond conformations of the interfacial formic acid molecule at the air-water interface.** The blue arrow represents the surface normal. The green arrows represent the normal of the H-C=O plane.  $\varphi$  represents the angle formed by the normal of the H-C=O plane and the surface normal. The number in the parenthesis represents the fraction of the conformation obtained in the AIMD simulation.

We comment on the *cis*-conformation of the formic acid. It is extremely challenging to obtain robust  $A_{yzy}/A_{yyz}$  values for both C-H and C=O stretch mode, because the signal of the *cis*-conformation is one order magnitude smaller than that of the *trans*-conformation. Thus, we could not obtain reliable values for  $\theta_{CH}$  and  $\theta_{CO}$  of the *cis*-conformation. On the other hand, the signs of the C=O stretch peaks for the *trans*- and *cis*-conformations in the  $\text{Im}\chi_{\text{eff},ssp}^{(2)}$  spectra are opposite, reflecting that both the orientations of the C→O groups are different between the *trans* and *cis*-conformations. This notion agrees with the AIMD data ( $\langle\theta_{CO}\rangle = 80^\circ$ ), showing that our SFG data and AIMD data are consistent not only quantitatively for the *trans*-conformation but also at least qualitatively for the *cis*-conformation of the formic acid molecule at the air-water interface.

## V. CONCLUSION

We have developed the framework for determining the orientation of a molecule at the interface by coupling the angle information obtained from the multimode polarization-dependent HD-SFG spectroscopy. The distribution function can be generated based on the exponential decay function together with the intramolecular geometry constraint. We applied this SFG multimode coupling technique to obtain the molecular orientation of *trans*-conformation of formic acid molecules at the air-water interface. We identified  $\langle\theta_{\text{CH}}\rangle = 56 \pm 5^\circ$  and  $\langle\theta_{\text{CO}}\rangle = 124 \pm 5^\circ$ , in good agreement with the AIMD data. Our analysis indicates that the intramolecular geometry constraint is essential to determine the molecular orientation. Furthermore, we showed that the rectangular function for the orientational distribution function is inappropriate. We expected that this multimode polarization-dependent SFG technique can be applied to the biomolecules by probing the different moieties of the amino acid group or by using the isotope labeling for the C=O group of the amide backbone.<sup>25,52</sup>

## ACKNOWLEDGMENTS

We are grateful for the financial support from the MaxWater Initiative of the Max Planck Society. We acknowledge the financial support from the DAAD (Deutscher Akademischer Austauschdienst) Project Based Personnel Exchange Program (#57526761). We especially thank Yair Litman and Xiaoqing Yu for fruitful discussions and support for the measurement.

## REFERENCES

- <sup>1</sup> D.Beneventi, B.Carre, andA.Gandini, Colloids Surfaces A Physicochem. Eng. Asp. **189**, 65 (2001).
- <sup>2</sup> C.George, M.Ammann, B.D'Anna, D.J.Donaldson, andS.A.Nizkorodov, Chem. Rev. **115**, 4218 (2015).
- <sup>3</sup> Q.Du, R.Superfine, E.Freysz, andY.R.Shen, Phys. Rev. Lett. **70**, 2313 (1993).
- <sup>4</sup> J.E.Laaser, D.R.Skoff, J.J.Ho, Y.Joo, A.L.Serrano, J.D.Steinkruger, P.Gopalan, S.H.Gellman, andM.T.Zanni, J. Am. Chem. Soc. **136**, 956 (2014).
- <sup>5</sup> L.Fu, J.Liu, andE.C.Y.Yan, J. Am. Chem. Soc. **133**, 8094 (2011).
- <sup>6</sup> L.Fu, G.Ma, andE.C.Y.Yan, J. Am. Chem. Soc. **132**, 5405 (2010).
- <sup>7</sup> J.Wang, C.Chen, S.M.Buck, andZ.Chen, J. Phys. Chem. B **105**, 12118 (2001).
- <sup>8</sup> R.-R.Feng, Y.Guo, andH.-F.Wang, J. Chem. Phys. **141**, 18C507 (2014).
- <sup>9</sup> E.Tyrode, C.M.Johnson, A.Kumpulainen, M.W.Rutland, andP.M.Claesson, J. Am. Chem. Soc. **127**, 16848 (2005).
- <sup>10</sup> X.Zhuang., P.B.Miranda, D.Kim, andY.R.Shen, Phys. Rev. B **59**, 12632 (1999).
- <sup>11</sup> Y.Rao, Y.S.Tao, andH.F.Wang, J. Chem. Phys. **119**, 5226 (2003).
- <sup>12</sup> J.Kirschner, A.H.A.Gomes, R.R.T.Marinho, O.Björneholm, H.Ågren, V.Carravetta, N.Ottosson, A.N.deBrito, andH.J.Bakker, Phys. Chem. Chem. Phys. **23**, 11568 (2021).
- <sup>13</sup> K.T.Nguyen, S.V.LeClair, S.Ye, andZ.Chen, J. Phys. Chem. B **113**, 12169 (2009).
- <sup>14</sup> Y.R.Farah andA.T.Krummel, J. Chem. Phys. **154**, 124702 (2021).
- <sup>15</sup> C.S.Santos andS.Baldelli, J. Phys. Chem. B **111**, 4715 (2007).
- <sup>16</sup> S.N.Wren, B.P.Gordon, N.A.Valley, L.E.McWilliams, andG.L.Richmond, J. Phys. Chem. A **119**, 6391 (2015).

- <sup>17</sup> S.Baldelli, J.Bao, W.Wu, and S.S.Pei, Chem. Phys. Lett. **516**, 171 (2011).
- <sup>18</sup> W.Gan, D.Wu, Z.Zhang, R.R.Feng, and H.F.Wang, J. Chem. Phys. **124**, 114705 (2006).
- <sup>19</sup> X.Weiz and Y.R.Shen, Phys. Rev. Lett. **86**, 4799 (2001).
- <sup>20</sup> C.M.Johnson, E.Tyrode, A.Kumpulainen, and C.Leygraf, J. Phys. Chem. C **113**, 13209 (2009).
- <sup>21</sup> E.Tyrode, C.M.Johnson, S.Baldelli, C.Leygraf, and M.W.Rutland, J. Phys. Chem. B **109**, 329 (2005).
- <sup>22</sup> C.M.Johnson, E.Tyrode, S.Baldelli, M.W.Rutland, and C.Leygraf, J. Phys. Chem. B **109**, 321 (2005).
- <sup>23</sup> E.Tyrode, C.M.Johnson, M.W.Rutland, and P.M.Claesson, J. Phys. Chem. C **111**, 11642 (2007).
- <sup>24</sup> K.T.Nguyen, J.T.King, and Z.Chen, J. Phys. Chem. B **114**, 8291 (2010).
- <sup>25</sup> W.Guo, T.Lu, Z.Gandhi, and Z.Chen, J. Phys. Chem. Lett. **12**, 10144 (2021).
- <sup>26</sup> S.Alamdari, S.J.Roeters, T.W.Golbek, L.Schmüser, T.Weidner, and J.Pfaendtner, Langmuir **36**, 11855 (2020).
- <sup>27</sup> T.Weidner, N.F.Breen, K.Li, G.P.Drobny, and D.G.Castner, Proc. Natl. Acad. Sci. U. S. A. **107**, 13288 (2010).
- <sup>28</sup> S.Ye, H.Li, F.Weiz, J.Jasensky, A.P.Boughton, P.Yang, and Z.Chen, J. Am. Chem. Soc. **134**, 6237 (2012).
- <sup>29</sup> K.Saito, Q.Peng, L.Qiao, L.Wang, T.Joutsuka, T.Ishiyama, S.Ye, and A.Morita, Phys. Chem. Chem. Phys. **19**, 8941 (2017).
- <sup>30</sup> S.Nihonyanagi, R.Kusaka, K.I.Inoue, A.Adhikari, S.Yamaguchi, and T.Tahara, J. Chem. Phys. **143**, 124707 (2015).
- <sup>31</sup> Shoichi Yamaguchi and Takuhiro Otsu, Phys. Chem. Chem. Phys. **23**, 18253 (2021).

- <sup>32</sup> S.Nihonyanagi, S.Yamaguchi, and T.Tahara, *J. Chem. Phys.* **130**, 204704 (2009).
- <sup>33</sup> N.Takeshita, M.Okuno, and T.A.Ishibashi, *Phys. Chem. Chem. Phys.* **19**, 2060 (2017).
- <sup>34</sup> F.Tang, T.Ohto, S.Sun, J.R.Rouxel, S.Imoto, E.H.G.Backus, S.Mukamel, M.Bonn, and Y.Nagata, *Chem. Rev.* **120**, 3633 (2020).
- <sup>35</sup> S.Sun, F.Tang, S.Imoto, D.R.Moberg, T.Ohto, F.Paesani, M.Bonn, E.H.G.Backus, and Y.Nagata, *Phys. Rev. Lett.* **121**, 246101 (2018).
- <sup>36</sup> F.Tang, T.Ohto, T.Hasegawa, W.J.Xie, L.Xu, M.Bonn, and Y.Nagata, *J. Chem. Theory Comput.* **14**, 357 (2018).
- <sup>37</sup> W.J.Smit, F.Tang, M.A.Sánchez, E.H.G.Backus, L.Xu, T.Hasegawa, M.Bonn, H.J.Bakker, and Y.Nagata, *Phys. Rev. Lett.* **119**, 133003 (2017).
- <sup>38</sup> X.Yu, T.Seki, C.-C.Yu, K.Zhong, S.Sun, M.Okuno, E.H.G.Backus, J.Hunger, M.Bonn, and Y.Nagata, *J. Phys. Chem. B* **125**, 10639 (2021).
- <sup>39</sup> X.Li, J.Liu, K.Lin, Y.Zhang, Y.Zhang, R.Zheng, Q.Shi, Y.Guo, and Z.Lu, *J. Phys. Chem. C* **123**, 12975 (2019).
- <sup>40</sup> G.Murdachaew, G.M.Nathanson, R.B.Gerber, and L.Halonen, *Phys. Chem. Chem. Phys.* **18**, 29756 (2016).
- <sup>41</sup> A.Morita, *Theory of Sum Frequency Generation Spectroscopy* (Springer Singapore, Singapore, 2018).
- <sup>42</sup> Z.Zhang, Y.Guo, Z.Lu, L.Velarde, and H.F.Wang, *J. Phys. Chem. C* **116**, 2976 (2012).
- <sup>43</sup> P.A.Pieniazek, C.J.Tainter, and J.L.Skinner, *J. Chem. Phys.* **135**, 044701 (2011).
- <sup>44</sup> T.Seki, S.Sun, K.Zhong, C.C.Yu, K.Machel, L.B.Dreier, E.H.G.Backus, M.Bonn, and Y.Nagata, *J. Phys. Chem. Lett.* **10**, 6936 (2019).

- <sup>45</sup> J.Hutter, M.Iannuzzi, F.Schiffmann, andJ.Vandevondele, Wiley Interdiscip. Rev. Comput. Mol. Sci. **4**, 15 (2014).
- <sup>46</sup> S.Grimme, J.Antony, S.Ehrlich, andH.Krieg, J. Chem. Phys. **132**, 154104 (2010).
- <sup>47</sup> K.Marushkevich, L.Khriachtchev, andM.Räsänen, J. Phys. Chem. A **111**, 2040 (2007).
- <sup>48</sup> E.M.S.Maçôas, J.Lundell, M.Petterson, L.Khriachtchev, R.Fausto, andM.Räsänen, J. Mol. Spectrosc. **219**, 70 (2003).
- <sup>49</sup> G.Giubertoni, O.O.Sofronov, andH.J.Bakker, J. Phys. Chem. Lett. **10**, 3217 (2019).
- <sup>50</sup> L.B.Dreier, M.Bonn, andE.H.G.Backus, J. Phys. Chem. B **123**, 1085 (2019).
- <sup>51</sup> A.Luzar andD.Chandler, Nature **379**, 55 (1996).
- <sup>52</sup> S.D.Moran andM.T.Zanni, J. Phys. Chem. Lett. **5**, 1984 (2014).