A Simplified Method for Theoretical Sum Frequency Generation Spectroscopy Calculation and Interpretation: the "pop model"

Wanlin Chen,[1,2] Dorian Louaas,[1] Flavio Siro Brigiano,[3] Simone Pezzotti,[4] and Marie-Pierre Gaigeot[1]
[1)] *Université Paris-Saclay, Univ Evry, CNRS, LAMBE UMR8587, 91025 Evry-Courcouronnes, France*
[2)] *Department of Physical Chemistry II, Ruhr University Bochum, D-44801 Bochum, Germany*
[3)] *Laboratoire de Chimie Théorique, Sorbonne Université, UMR 7616 CNRS, 4 Place Jussieu, 75005 Paris, France*
[4)] *Department of Physical Chemistry II, Ruhr University Bochum, D-44801 Bochum, Germany*
**Corresponding author: Marie-Pierre Gaigeot, mgaigeot@univ-evry.fr, Simone Pezzotti, simone.pezzotti@rub.de*

Sum Frequency Generation (SFG) spectroscopy is a powerful tool to probe molecular environments of otherwise nearly inaccessible buried interfaces. Theoretical spectroscopy is required to reveal the structure-spectroscopy relationship. Existing methods to compute theoretical spectra are restricted to the use of time-correlation functions evaluated from accurate atomistic molecular dynamics simulations, often at the ab-initio level. The interpretation of the computed spectra requires additional steps to deconvolve the spectroscopic contributions from local water and surface structural populations at the interface. The lack of a standard procedure to do this often hampers rationalization. To overcome these challenges, we rewrite the equations for spectra calculation into a sum of partial contributions from interfacial populations, weighted by their abundance at the interface. We show that SFG signatures from each population can be parameterized into a minimum dataset of reference partial spectra. Accurate spectra can then be predicted by just evaluating the statistics of interfacial populations, which can be done even with force field simulations as well as with analytic models. This approach broadens the range of simulation techniques from which theoretical spectra can be calculated, opening toward non-atomistic and Monte Carlo simulation approaches. Most notably, it allows constructing accurate theoretical spectra for interfacial conditions that can not even be simulated, as we demonstrate for the pH-dependent SFG spectra of silica/water interfaces.

INTRODUCTION

Sum Frequency Generation (SFG) spectroscopy[1,2] serves as an effective tool for interface characterization. As a non-linear vibrational spectroscopy technique, SFG depends on the second-order susceptibility $\chi^{(2)}(\omega)$, which is zero in centrosymmetric media (e.g. bulk of liquids and many solids). Thanks to its surface specificity, SFG spectroscopy has been widely utilized to investigate aqueous interfaces to unravel their structural, vibrational and dynamical properties at a molecular level[3–10].

Interpreting SFG spectra in terms of microscopic surface structures requires theoretical analyses. Extensive research has been dedicated for theoretical $\chi^{(2)}(\omega)$ calculation via molecular dynamics (MD) simulations, which are based on correlation functions between molecular dipole moment and polarizability tensor[11–15] or velocity-velocity correlation functions (VVCF)[16–22]. However, the interpretation of SFG spectra remains challenging due to the overload of information contained in the signals.

Taking the example of the 3000–4000 cm^{-1} spectral domain, both water OHs and surface OH groups (e.g. silanols on silica surface) contribute to the SFG signal. Dissecting these contributions requires specific deconvolution of the spectrum and the assignment of the different spectroscopic fingerprints into the structural populations of both water and surface. In Heterodyne-detected (HD)-SFG, the interpretation of the signal is furthermore complicated by the compensation of positive and negative peaks due to different orientations of the OH groups, which also requires specific theoretical deconvolution.

Moreover, the SFG signals arising from water have spectral contributions from different water layers, each one with distinct structural properties. An aqueous interface can be separated into three layers[18,19,23]: the Binding Interfacial Layer (BIL), i.e. the water layer directly in contact with the surface together with the top-surface contribution, the Diffuse Layer (DL), i.e. the subsequent water layer with bulk-like water reoriented by the surface charges[23], and liquid bulk water. Only BIL and DL are SFG active with broken centrosymmetry. By calculating separately the SFG signals arising from water and top-surface OHs in the BIL and DL, the total SFG spectrum can be deconvolved: [19]

$$\chi^{(2)}(\omega) = \chi^{(2)}_{BIL}(\omega) + \chi^{(2)}_{DL}(\omega) \quad (1)$$

where $\chi^{(2)}_{BIL}(\omega)$ and $\chi^{(2)}_{DL}(\omega)$ are spectral contributions from BIL and DL.

The physics behind $\chi^{(2)}_{BIL}(\omega)$ and $\chi^{(2)}_{DL}(\omega)$ are different. $\chi^{(2)}_{DL}(\omega)$ arises solely from a population of bulk-like water molecules reoriented by the electric field generated by the surface charges. $\chi^{(2)}_{DL}(\omega)$ is thus a universal two-band signal centered at 3200 cm^{-1} and 3400 cm^{-1}, which sign (in HD-SFG) and amplitude depend on the surface po-

tential[19,23]. $\chi^{(2)}_{BIL}(\omega)$ is composed of contributions from the surface functional groups (e.g. OH terminations) and the water molecules (denoted BIL-water) interacting with both the surface and the subsequent water layer (DL). The interactions between BIL-water and surface depend on the types and local distribution of the surface functional groups, the surface morphology, the topology, etc..., resulting in a variety of structural populations providing the surface-specific fingerprints in the SFG spectrum.

Because of these complexities, just computing theoretical SFG spectra is usually far from sufficient. Their interpretation indeed requires deconvolution steps to dissect BIL from DL signals, and to dissect the contributions from distinct BIL populations. This is based on the structural knowledge extracted from MD simulations. This is however limited to systems that can be simulated with good enough accuracy for the theoretical spectroscopy calculation, e.g. usually at the DFT-MD electronic level of representation for instance for aqueous oxide interfaces (where water-surface interactions are hard to describe with sufficient accuracy for spectroscopy by classical force fields). Many relevant interfacial conditions can however not even be realistically simulated, such as variations in pH conditions or electrolyte solutions with low ionic concentrations. It is thus highly challenging to accurately calculate and interpret the SFG spectra of interfaces in the wide range of conditions that they experience in nature and in common applications.

We hereby propose a simplified theoretical method denoted the "pop model", that constructs SFG spectra of aqueous interfaces as the sum of spectral contributions from identified populations located in the different interfacial SFG active layers. It consists in rewriting the standard equations to compute a theoretical spectrum into a sum of partial contributions arising from the different interfacial populations, weighted by the abundance of these populations. We will show that it is possible to parameterize the SFG signatures of a minimum dataset of populations from DFT-MD simulations, and predict with high accuracy the SFG spectrum of a given interface only from the knowledge of the statistics of the population at that interface. The statistics on the population can be obtained with classical force fields simulations, which enlarges the complexity of interfaces over which theoretical spectra can be simulated. More importantly, the "pop model" provides a way to calculate theoretical SFG spectra even in cases where simulations are not available, as long as the abundance of interfacial populations can be deduced from existing analytical models. We provide a proof of principle for this exciting perspective by predicting pH-dependent surface SFG spectra of silica/water interfaces as a function of pH in excellent quantitative agreement with experiments.

METHOD AND APPLICATIONS

In the following, we introduce the equations for the "pop model" in order of increasing complexity. We start from the DL, which can be described by a single water population, and then extend the approach to the BIL, where the signals of multiple water and surface populations overlap. Hereafter, we define these populations based on local (first shell) coordination and orientation of water and surface OH groups.

A. SFG Signal in the Diffuse Layer (DL)

As shown in refs.[19,23–25], the SFG $\chi^{(2)}_{DL}(\omega)$ response from water in the DL region of the interface can be rewritten in terms of the bulk liquid water third order susceptibility $\chi^{(3)}_{Bulk}(\omega)$ contribution, hence providing the generic two bands at 3200 and 3400 cm^{-1} of any DL-SFG spectrum, weighted in sign and magnitude by the potential difference $\Delta\phi_{DL}$ across the DL:

$$\chi^{(2)}_{DL}(\omega) = \chi^{(3)}_{Bulk}(\omega) \cdot \Delta\phi_{DL} \tag{2}$$

where $\chi^{(3)}_{Bulk}(\omega)$ is the third order susceptibility of liquid water and $\Delta\phi_{DL}$ is:

$$\Delta\phi_{DL} = \int_{z_a}^{\infty} \mathrm{d}z \cdot E_{\mathrm{DC}}(z) \mathrm{e}^{i\Delta k_z z} \tag{3}$$

with z_a the vertical (i.e. perpendicular to the surface) boundary between the BIL and DL and ∞ is the boundary between DL and bulk water, $E_{\mathrm{DC}}(z)$ the electrostatic field, and Δk_z a phase factor that takes into account interferences between the emitted light at different depths from the surface.

When theoretical DL-SFG spectra are calculated from MD simulations using the approach we have developed and validated in ref. 19, the knowledge of $\chi^{(3)}_{Bulk}(\omega)$ and $\Delta\phi_{DL}$ is not required anymore. $\chi^{(2)}_{DL}(\omega)$ is simply obtained from the Fourier transform of the dipole-polarizability correlation function (i.e. the standard time-dependent method introduced by Morita *et al.*,[11,12]) and by considering only the contribution of the DL-water molecules (identified based on a clear definition of the evolution of water density as a function of the vertical distance from the instantaneous surface[26], as detailed in ref. 18,19), so that:

$$\chi^{(2)}_{DL}(\omega) = \frac{i\omega}{k_B T} \sum_{i=1}^{N_{DL}} \left(\sum_{j=1}^{N_W} \int_0^{\infty} dt e^{i\omega t} \langle \alpha^j_{xx}(t) \mu^i_z(0) \rangle \right) \tag{4}$$

where all the possible cross-correlation terms between the N_{DL} water molecules located in the DL layer and all the N_W water molecules of the simulated system ($N_W = N_{BIL} + N_{DL} + N_{Bulk}$) are explicitly included.

The xxz component of the $\chi^{(2)}_{DL}(\omega)$ tensor is here considered, corresponding to *ssp* polarization of the SFG lasers. $\mu^i_z(0)$ and $\alpha^j_{xx}(t)$ respectively refer to the individual dipole and polarizability of the i-th and j-th water molecules, j and i are indices respectively running over the total number of water molecules in the simulation box (Nw) and over the number of water molecules in the DL (N_{DL}). Recent applications of DFT-MD for SFG calculations replace the polarizability-dipole correlation function in eq. 4 by a time correlation of velocities weighted by APT and Raman tensors (VVCF approach[17], often including self-correlation terms only). Hereafter, the derivation of the "pop model" is done from eq. 4.

The $\frac{i\omega}{k_B T}\sum_{j=1}^{N_W}\int_0^\infty dt\exp(i\omega t)\langle\alpha^j_{xx}(t)\mu^i_z(0)\rangle)$ term in the equation represents the contribution to the total SFG spectrum of the i-th water molecule located in the DL. This is labeled $\beta^i_{DL}(\omega)$ hereafter.

We can thus rewrite Eq. 4 as:

$$\chi^{(2)}_{DL}(\omega) = \sum_{i=1}^{N_{DL}} \beta^i_{DL}(\omega) \tag{5}$$

where $\beta^i_{DL}(\omega) = \beta^i_{DL,xxz}(\omega)$ for the *ssp* SFG polarization of interest in this work.

As demonstrated in refs.[19,23,24,27], the SFG activity of the DL is due to a pure second order effect, dominated by bulk-water reorientation induced by the static-field arising from the charged surface. Any difference between DL-SFG spectra of interfaces with different surface charges can thus be related to the difference in the average orientation of the DL-water molecules.

In order to explicitly show the dependence of the DL-SFG signal on the water orientation in Eq. 5, $\beta^i_{DL}(\omega)$ can be divided and multiplied by the orientation of each i-th water molecule, denoted by $cos\theta_i$. Specifically, $cos\theta_i$ is the cosine of the angle formed between the dipole direction of the i-th water molecule and the normal to the surface (defined from the liquid to the solid/vapor phase). Assuming that differences between the SFG-activity of the DL-water molecules only arise from their orientation and that $\beta^i_{DL}(\omega)$ linearly depends on $cos\theta_i$, so that $\frac{\beta^i_{DL,xxz}(\omega)}{cos\theta_i}$=$\beta^{eff}_{DL,xxz}(\omega)$ is a constant for each given ω, we obtain:

$$\chi^{(2)}_{DL}(\omega) = \sum_{i=1}^{N_{DL}} \frac{\beta^i_{DL}(\omega)}{cos\theta_i}\cdot cos\theta_i \tag{6}$$

$$= \beta^{eff}_{DL}(\omega)\cdot\sum_{i=1}^{N_{DL}} cos\theta_i \tag{7}$$

where $\frac{\beta^i_{DL}(\omega)}{cos\theta_i}$=$\beta^{eff}_{DL}(\omega)$ is defined as the effective hyperpolarizability of the *i-th* DL-water molecule, constant for all DL-water molecules for each given ω, i.e. a function of ω only.

This equation allows separating the orientation part of any DL-SFG spectrum (due to the surface charge), i.e. $\sum_{i=1}^{N_{DL}} cos\theta_i$, from the $\beta^{eff}_{DL}(\omega)$ term which is the contribution to the SFG spectrum of a DL water molecule per unitary orientation. This latter is identical for all DL-water molecules at the charged interface. It is important to stress here that Eq. 7 and Eq. 2 are equivalent under the assumption that orientation dominates the DL-SFG activity (which is indeed the usual assumption). The $\sum_{i=1}^{N_{DL}} cos\theta_i$ orientation term in Eq. 7 is hence the counterpart of $\Delta\phi_{DL}$ in Eq.2, at the origin of the sign and intensity of DL-SFG spectra, while $\beta^{eff}_{DL}(\omega)$ is the counterpart of $\chi^{(3)}_{bulk}(\omega)$, providing the two bands structure of any DL-SFG spectrum.

In order to prove that $\beta^{eff}_{DL}(\omega)$ is, as $\chi^{(3)}_{bulk}(\omega)$, a universal term at any aqueous interface, and therefore that Eq. 2 and Eq. 7 are equivalent formulations, we now consider the same DFT-MD simulations of charged aqueous interfaces as the ones used in refs. 19 and 28, which allowed us to obtain the theoretical $\chi^{(3)}_{bulk}(\omega)$ using Eq. 2. The same trajectories are used to calculate $\beta^{eff}_{DL}(\omega)$ from the theoretical $\chi^{(2)}_{DL}(\omega)$ via Eq. 7. The systems considered include both neat and electrolytic silica-water interfaces with different surface charge values and ions concentrations. The neat interfaces are: the fully protonated (0001)-α-Quartz-water (QW) interface, QW with 3% of surface SiOH sites being deprotonated, and QW with 12% of deprotonated sites. Three electrolytic interfaces are generated by adding different concentrations of KCl ion-pairs in the simulation box of the fully hydroxylated QW interface. The resulting excess concentrations $[K^+]_{BIL}$ of K^+ in the BIL, are 1.6 M, 4.2 M and 7.2 M for the investigated three systems. We refer to ref. 28 for more details on these DFT-MD simulations.

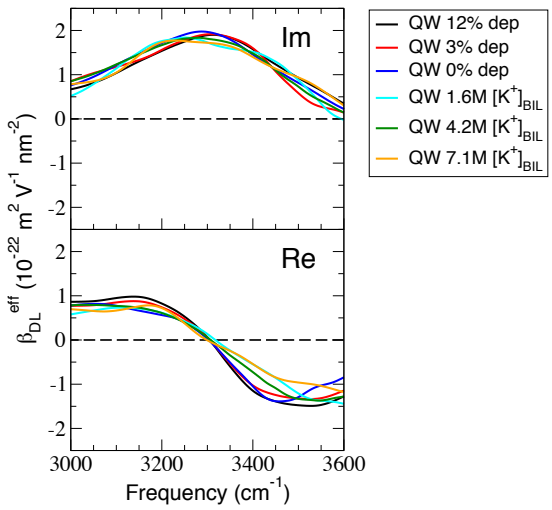

FIG. 1. $\beta^{eff}_{DL}(\omega)$ spectra (imaginary, Im and real, Re components) calculated using Eq.7 for the DFT-MD simulations already used to parametrize $\chi^3(\omega)$ in refs.[19,28]: Quartz-water interfaces with different degrees of surface deprotonation (0%, 3% and 12%) and ions (KCl) concentrations, reported as excess cation (K^+) concentration in the BIL (1.6, 4.2 and 7.1 M).

As shown in Fig. 1, similar $\beta^{eff}_{DL}(\omega)$ spectra are obtained for all considered systems. They have the same shape as the $\chi^{(3)}_{Bulk}(\omega)$ spectrum reported in refs.[19,28], with the same two positive bands at 3200 and 3400 cm^{-1}. With this universal $\beta^{eff}_{DL}(\omega)$ parameterized based on these DFT-MD simulations, we can now calculate any DL-SFG spectrum from any MD simulation (DFT-MD or classical MD) or even from Monte Carlo simulations by simply evaluating the orientation term $\sum_{i=1}^{N_{DL}} cos\theta_i$ (Eq. 7) for the water molecules located in the DL.

B. SFG Signal in the Binding Interfacial Layer (BIL)

Contrary to the DL, the water structural organization in the Binding Interfacial Layer (BIL) cannot be simply described by the orientation of the water molecules. Because of the actual balance between water-water and water-surface interactions at a given aqueous surface, several distinct water populations contribute to the BIL spectrum, each one characterized by specific local coordination and orientation. We hereby define such populations simply based on the water OH groups being either H-bonded to the surface, or to the water, or non H-bonded at all, together with their orientations with respect to the surface normal (oriented from the liquid to the surface as is the convention in SFG spectroscopy). These criteria are easily obtained as output of simulations, making the classification of BIL populations and their statistical analysis straightforward. The types and relative abundances of such BIL-populations are surface specific as they depend on their interaction strength with surface functional groups as well as on the local surface morphology, topology, hydrophilicity, charge state, etc. However, we will show in the following that a small number of spectroscopically relevant populations is sufficient to describe various families of aqueous interfaces (e.g. hydrophobic interfaces, interfaces between water and organic monolayers). As for the DL, we need to extract the $\beta^{eff}(\omega)$ of each defined population located in the BIL, which requires building a database for each family of surfaces. In this work, the SFG fingerprints of a series of BIL-populations have been calculated from our database of DFT-MD simulations on various aqueous interfaces. The BIL-SFG spectrum of each population has been obtained with the VVCF method[18–22,28].

As summarized in Fig. 2 and demonstrated hereafter, a limited set of 6 BIL-water populations were required as building blocks for the "pop model" to describe the SFG spectra of a large number of liquid/solid, liquid/vapor and liquid/organic aqueous interfaces. These interfaces include charged and neutral surfaces, ranging from pure hydrophobic to hydrophilic, with in between the interfaces having coexistent local hydrophobic and hydrophilic patches.

BIL-water OH groups are first categorized into (1) OH groups pointing towards the subsequent water layer (opposite to the conventional SFG normal vector, denoted "OH-down", with $\cos\theta < 0$ where θ is the angle between an OH group and the SFG normal vector, i.e. perpendicular to the surface, pointing from water to the surface, see fig. 2), resulting in a negative SFG band, and (2) OH groups pointing towards the surface (along the conventional SFG normal vector, denoted "OH-up" with $\cos\theta > 0$), leading to a positive SFG band. The spectroscopic fingerprint of "OH-down" groups (Fig. 2A, solid cyan line) is generalizable for any investigated aqueous interface as these groups are systematically HB-donors to bulk (or DL) water molecules with the same interaction strength, as demonstrated in ref. 22. This population is denoted by pop 1 in the following.

The "OH-up" fingerprints are system dependent due to the strength of interaction between BIL-water and the top-surface sites.

At hydrophobic interfaces, the "OH-up" groups are dangling, i.e. pointing towards the surface without forming H-bonds, resulting in a free-OH peak located at above 3600 cm^{-1} in SFG. Note that the free-OH peak is usually recognized as a molecular proof of hydrophobicity[22,29,30]. In the case of the air/water interface, the prototype hydrophobic interface, the dangling OH groups pointing towards the air is free, corresponding to a positive SFG peak at 3700 cm^{-1}, denoted pop 2 (Fig. 2B, solid red line).

At other hydrophobic interfaces, e.g. graphene, or hydrophilic interfaces with local hydrophobic patches, like the amorphous silica with low hydroxylation degree, the dangling OH groups point towards the hydrophobic functional groups (i.e. siloxanes) with weak interactions (e.g. van der Waals forces), leading to a red-shifted free OH peak (denoted quasi-free, Fig. 2B, dashed red line) compared to that of the air/water interface[22]. The representative fingerprint depicted in Fig. 2B with the dashed red line is parameterized from the SFG fingerprint arising from OH groups pointing to hydrophobic patches of the amorphous silica surface with a hydroxylation degree of 4.5 SiOH/nm^2.[20,22] Note that this quasi-free OH peak frequency shows very little variation between hydrophobic interfaces ($\pm$ 15 cm^{-1}) investigated in ref. 22, which makes the representative fingerprint semi-transferable to other hydrophobic interfaces. This quasi-free OH population is denoted pop 3.

In the case of hydrophilic interfaces, "OH-up" groups are H-bonded to surface hydrophilic functional groups, corresponding to a positive band in the H-bonded frequency range (below 3600 cm^{-1}). The band shape and frequency center of this peak is system specific, depending on the types of surface hydrophilic functional groups. Here, we present two examples of H-bonded OH-up fingerprints, one parameterized from the water OH groups forming donor H-bonds with surface SiOHs at an amorphous silica/water interface with a hydroxylation degree of 4.5 SiOH/nm^2[20,22] (denoted pop 4, Fig. 2C, dashed blue), and the other from water OH groups forming donor H-bonds with oxygen atoms of the organic self-assembled monolayer (SAM) composed of polyethy-

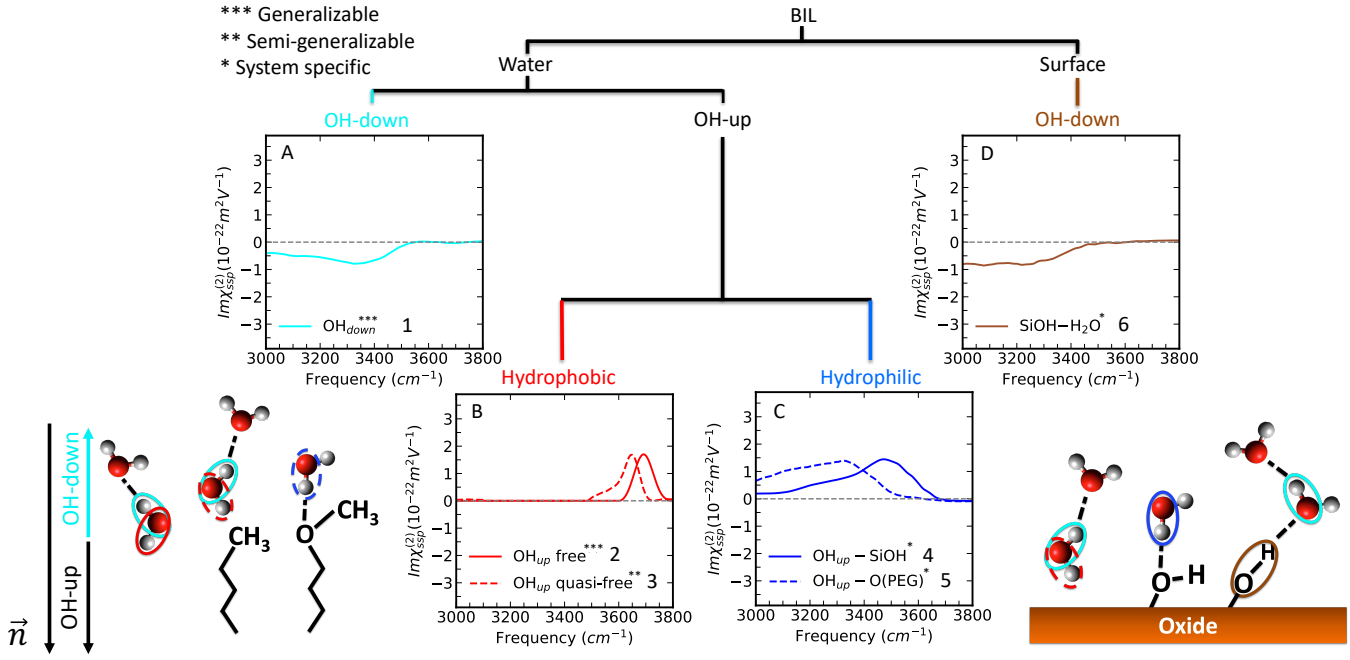

FIG. 2. Spectroscopic fingerprints $\beta^{eff}(\omega)$ of BIL-water OH groups and surface SiOH groups in the OH stretch frequency range, parameterized from various aqueous interfaces including the air/water, amorphous silica/water, self-assembled monolayer/water interfaces. Parameterization on DFT-MD trajectories and VVCF SFG calculations (see text, section B).

lene glycol (PEG, $CH_3(OCH_2CH_2)_7CH_2SiCl_3$), denoted pop 5, see Fig. 2C, solid blue. While being system dependent, these fingerprints can be utilized to reconstruct the SFG spectra of interfaces within the same family but with different parameters, e.g. with different degrees of hydroxylation and crystallinity of silica/water interfaces, with different ratios of PEG polymer chains involved in the SAM monolayer composition. These examples will be demonstrated in the following section of applications.

One last population is hereby defined for SFG signals in the OH stretch frequency range arising from the surface O-H functional groups. As for pop 4 and pop 5, this population is system dependent. Among the investigated interfaces in our database, the only SFG-active surface group in the OH stretch range arises from surface SiOHs[20,21,31], which is hence defined as pop 6 in the present work (Fig. 2D, solid brown line).

With the $\beta^{eff}(\omega)$ of different populations being parameterized from DFT-MD simulations of various aqueous interfaces through the VVCF-based SFG calculation method and shown in Fig. 2, we are now able to construct the SFG spectrum of any given interface within the families present in our current (small) database. It requires the knowledge of the statistics of each population given by MD simulations (either DFT-MD or standard classical MD). In analogy to Eq. 7 for the DL, the $\chi^{(2)}_{BIL}(\omega)$ is calculated by summing up the contributions arising from each of identified i-th population:

$$\chi^{(2)}_{BIL}(\omega) = \sum_{i=1}^{N_{pop}} \beta_i^{eff}(\omega) \cdot \left(\sum_{j=1}^{n_{pop_i}} cos\theta_{ij} \right) \qquad (8)$$

Note that unlike the DL where water orientation dominates the $\chi^{(2)}_{DL}(\omega)$ SFG activity, the orientation of the water OH groups is dictated by the balance between water-surface and water-water interactions. This balance is intrinsically taken into account in the definition of the populations and is ultimately accounted for in equation 8 with the summation over the cosine values.

C. Applications

In the following we present a series of DFT-MD and FF-MD simulations of aqueous interfaces for which the $\chi^{(2)}_{BIL}(\omega)$ SFG signal is calculated by using the simplified 'pop model' of $\beta_i^{eff}(\omega)$ fingerprints described in the previous section.

The protocole is the following. *Stage 1*- The DFT-MD/FF-MD trajectory is analyzed in terms of the structures adopted by the water molecules in the BIL and of the structures of the solid/polymer/air top-surface in order to extract the population(s) of the O-H groups regarding their orientational and H-Bonding properties. The O-H populations hence uncovered are attributed to one of the populations that have been described in the

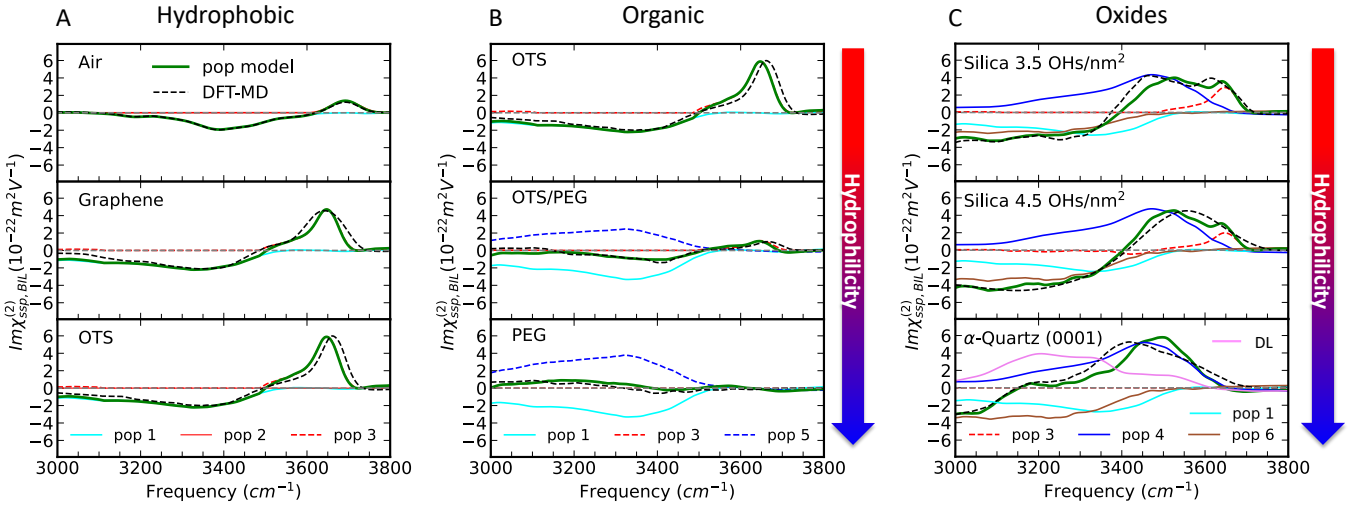

FIG. 3. BIL-SFG $\Im(\chi^{(2)}_{BIL}(\omega))$ spectra of various hydrophobic/hydrophilic interfaces calculated using "pop model" approach in solid green lines. (A) Hydrophobic interfaces. Top: the air/water interface calculated over a classical MD simulation of 20 ns (SPCE[32] water); Middle: graphene/water interface calculated over a DFT-MD simulation of 50 ps. Bottom: OTS/water interface calculated over a DFT-MD simulation of 400 ps. (B) Organic SAM/water interfaces SFG spectra calculated over DFT-MD simulations of 400 ps using the "pop model". Top: OTS/water interface; Middle: PEG/water interface; and Bottom: homogeneously (1:1) mixed OTS-PEG/water interface. (C) Silica/water interfaces with different degrees of crystallinity and hydroxylation. Top: amorphous silica/water interface with a hydroxylation degree of 3.5 OHs/nm^2; Middle, amorphous silica/water interface with a hydroxylation degree of 4.5 OHs/nm^2; Bottom: α-Quartz (0001)/water interface with added DL contribution (in pink) for a direct comparison with the DFT-MD calculated spectrum reported in ref. 21. The "pop model" SFG spectra are calculated over DFT-MD trajectories, respectively of 20 ps (top), 20 ps (middle) and 15 ps (bottom). These trajectories can be found in refs. 20,21. The spectral contributions from each population are depicted for each system using the same color code and line style as in Fig. 2. In all plots, the VVCF based DFT-MD calculated spectra are plotted in dashed black for comparison with the "pop model" spectra in green.

previous section and which $\beta_i^{eff}(\omega)$ spectral fingerprints have been presented in figure 2. This provides the knowledge of the density of each population found in the BIL (i.e. the ensemble of N_{pop} in eq. 8). *Stage 2*- The $\beta_i^{eff}(\omega)$ spectral fingerprint from the database is associated to each density population. *Stage 3*- The $\chi^{(2)}_{BIL}(\omega)$ SFG signal of the BIL is calculated through equation 8 of the 'pop model'.

In the following applications, when the 'pop model' is applied on a DFT-MD trajectory that served for the parameterization of a $\beta_i^{eff}(\omega)$ fingerprint, the validity of eq. 8 is tested in terms of the $\beta_i^{eff}(\omega)$ fingerprint per unitary orientation. When the 'pop model' is applied on a FF-MD trajectory that was not used in the parameterization of a $\beta_i^{eff}(\omega)$ fingerprint, the transferrability of a DFT-MD based $\beta_i^{eff}(\omega)$ to a classical FF-MD trajectory is tested together with the final accuracy of the FF-MD-pop model signal for the positions, intensities and shapes of the spectral bands.

1. Hydrophobic Interfaces

We start our demonstration of the "pop model" for SFG calculations by hydrophobic aqueous interfaces. The air/water interface has been extensively studied through SFG spectroscopy[4,30,33–35] during the last decades: the $\Im(\chi^{(2)}_{BIL}(\omega))$ of the air/water interface is composed of a positive peak located at ∼3680 cm^{-1} (in the free OH region) and a negative band centered at ∼3400 cm^{-1} (in the H-bonded region). Apart from the free OH peak, a two-dimensional collective H-bonded network (2DN)[18] composed of interconnected interfacial water molecules through H-Bonds being oriented parallel to the water surface was uncovered. The 2DN is another molecular hydrophobic descriptor, as shown in refs. 18,22,36, which is however SFG inactive in *ssp* or *ppp* polarization due to the orientation of the OH groups of water being perpendicular to the SFG normal vector. The theoretical $\Im(\chi^{(2)}_{BIL}(\omega))$ of the air/water interface calculated from DFT-MD simulation using the VVCF based method[17,18] was shown in good agreement with the experimental SFG spectrum.[18]

The deconvolution of the two SFG-band spectrum made in ref[18] showed that the O-H groups of the water in the 2DN that are pointing towards the air contribute to the $\sim$ 3700cm^{-1} free O-H positive band, while the O-H groups of the water in the 2DN that are pointing towards and making H-Bonds with the subsequent bulk water layer are responsible for the $\sim$ 3400cm^{-1} negative band.

A classical FF-MD simulation using the SPC/E force field[32] of the air-water interface has been accumulated with the same box size and same number of water molecules as in the DFT-MD simulation used for the SFG calculation in ref. 18, but larger time scale (20 ns), (see ref. 37 for simulation details). This trajectory is analyzed for the water O-H density populations in the BIL as described in the protocole above, finding without surprise the two building block populations identified in our 'pop model' as respectively OH_{down} population (pop 1, Fig. 2A for $\beta_{eff}(\omega)$) and OH_{up} free (pop 2, Fig. 2B for $\beta_{eff}(\omega)$). The SFG signal of this FF-MD air/water interface can then be reconstructed by using these two populations with our "pop model" of Eq. 8. The result is shown in Fig. 3A, upper panel. The excellent agreement between the FF-MD-pop model SFG spectrum calculated over a classical MD trajectory with the DFT-MD calculated spectrum using the VVCF method validates the "pop model". This shows that the "pop model" hence makes it possible to construct a SFG spectrum in excellent quality using classical MD simulations without a specialized force field and without calculating any kind of correlation functions.

The SFG spectrum of other hydrophobic interfaces, e.g. graphene/water, are similar to that of the air/water interface, as demonstrated in ref. 22. Taking the example of the graphene/water SFG spectrum calculated from DFT-MD (fig. 3A middle panel), one indeed recognizes a similar positive free OH peak at 3644 cm^{-1} and a negative H-bonded band at $\sim$ 3550 cm^{-1}, similar to the ones in the air/water SFG spectrum. The decomposition of these peaks in ref. 22 revealed the same two populations, OH_{down} population (pop 1, Fig. 2A) and OH_{up} free (pop 2, Fig. 2B), as the two sole contributors to the SFG spectrum. Hence, we use the same two-population model as for the air/water interface to construct the SFG spectrum of the graphene/water by extracting the density of each population from the DFT-MD simulation (see ref. 22 for simulation details).

However, the free OH peak, in this case, is slightly redshifted compared to that of the air/water interface due to the weak interactions between water and the surface. This is well captured by the OH_{up} quasi-free population (pop 3, Fig. 2B for $\beta_{eff}(\omega)$), while the negative band arises from the same OH_{down} (pop 1, Fig. 2A for $\beta_{eff}(\omega)$)). The SFG spectrum deduced from the "pop model" in Fig. 3A (middle panel) matches extremely well the one calculated from the DFT-MD simulation using the VVCF method in terms of frequency, intensity and band-shape.

2. Organic (SAMs) Aqueous Interfaces

To increase complexity in the application of the "pop model" to more complex and more dynamical aqueous interfaces, the SFG spectroscopy of a set of aqueous interfaces composed of self-assembled monolayers (SAMs) is now investigated. SAM/water interfaces have attracted considerable attention as a model platform to study hydrophobicity and hydrophilicity both theoretically[38–40] and experimentally[12,41–43]. In ref.[36], the water SFG spectrum of three different silane-based SAM/water interfaces have been measured and calculated from DFT-MD simulations for (1) pure octadecyltrichlorosilane (OTS, CH_3(–CH_2)$_{17}$–$SiCl_3$) monolayer (intrinsically hydrophobic), (2) pure polyethylene glycol (PEG, $CH_3(OCH_2CH_2)_7CH_2SiCl_3$) monolayer (intrinsically hydrophilic with oxygens at the interface), and (3) homogeneously mixed OTS-PEG monolayer (1:1 mixing ratio in the calculation).

We start with the hydrophobic OTS/water interface. The structural analysis of the BIL in the DFT-MD trajectory (see details in ref.[36]) reveals the same water populations as found in the other hydrophobic interfaces, i.e. the OH_{up} free (pop 2) or the quasi-free (pop 3), and the OH_{down} (pop 1). The water SFG spectrum of the OTS/water interface is reconstructed with the "pop model" over the same DFT-MD simulation as in ref. 36. The result is displayed in Fig. 3B, upper panel together with the DFT-MD spectrum calculated with the VVCF method. As for the graphene/water interface, the frequency shift of the OH_{up} with respect to that of the air/water interface is taken into account by the OH_{up} quasi-free (pop 3) fingerprint. The good agreement with the DFT-MD calculated SFG spectrum once again validates our choice of the two populations for hydrophobic interfaces and of eq. 8 for reconstructing the $\chi^{(2)}_{BIL}(\omega)$. The small discrepancy in the frequency of the OH_{up} between the DFT-MD calculated and the "pop model" constructed SFG spectrum (within 15 cm^{-1}) does not prevent us from correctly interpreting the spectrum.

Turning now to a hydrophilic interface, the $\Im(\chi^{(2)}_{BIL}(\omega))$ of the PEG/water interface calculated from DFT-MD surprisingly shows an almost zero signal (Fig. 3B, middle panel, dashed black). As shown in ref. 36, this is due to the compensation between the contributions of 2 populations: OH_{down} (pop 1) and the specific OH_{up}-O(PEG) (pop 5). The spectral contributions arising from each population are also depicted in Fig. 3B, middle panel, using the same color coding and line style as in Fig. 2, with pop 1 in solid cyan and pop 5 in dashed blue. The compensation of the two contributions results from: (1) the OH_{up}-O(PEG) and OH_{down} spectra having the same vibrational frequency, implying the same strength of the hydrogen bonds for water-water and water-O(PEG) and (2) having an equivalent density of population.

The homogeneously (1:1) mixed OTS-PEG/water interface has both OH_{up}-O(PEG) (pop 5) and OH_{up} quasi-free (pop 3) as OH-up populations, depending on whether the OH-up is in proximity and H-bonded to a PEG or pointing to an OTS and remaining dangling. Three populations (pop 1, 3, 5) thus contribute to the SFG spectrum. Once evaluating each corresponding population density from the DFT-MD simulation, the water SFG spectrum at the interface is obtained using Eq. 8. The

result in Fig. 3B, lower panel is in perfect agreement with the one calculated from the DFT-MD simulation using the VVCF method.

The resulting SFG spectrum has a tiny free OH peak corresponding to the few dangling OH groups pointing towards the OTS chains and a negative peak in the H-bonded region, whose intensity is reduced compared to other hydrophobic interfaces. This is due to the compensation of signals arising from up OHs H-bonded to PEGs and down OHs H-bonded to the subsequent water layer, as indicated by the separated population contributions from pop 1 (solid cyan) and pop 5 (dashed blue).

3. Aqueous Silica

For the systems demonstrated above, the SFG signals in the OH stretch region solely arise from the O-H stretch of water molecules. However, for some aqueous interfaces with the presence of surface hydroxylated functional groups (e.g. silica/water interface), the surface OHs also contribute to the SFG spectra, whose signals are often overlapped with contributions from water OHs in the H-bonded region.

This is the case with silica-water interfaces. The organization of interfacial water molecules and the corresponding SFG spectra varies from system to system as a function of degrees of hydroxylation and crystallinity of the silica surface[19–21,27,44,45]. We show in Fig. 3C that these variations are fully captured by the "pop model" and can be rationalized in changes in the relative abundance of hydrophilic and hydrophobic populations.

Though being macroscopically hydrophilic, local hydrophobic patches consisting of siloxane bridges (Si-O-Si) can coexist with hydrophilic functional groups (SiOH) at the silica surface (especially amorphous silica). Similar to the mixed OTS-PEG/water interface, there are three BIL-water populations that are responsible for the SFG spectrum of a set of silica/water interfaces proposed in Fig. 3C: OH_{down} (pop 1), OH_{up} quasi-free (pop 3) and OH_{up}-SiOH (pop 4). In addition to BIL-water populations, surface silanol groups perpendicular to the surface pointing out-of-plane to water (pop 6) also contribute to the SFG spectrum in the O-H stretch frequency range.

Three silica/water interfaces are presented here: two amorphous silica surfaces with a hydroxylation degree of 3.5 OHs/nm^2 and 4.5 OHs/nm^2, together with one crystalline surface of α-quartz (0001), all fully hydroxylated (corresponding to pH < 4 conditions). All interfaces have been analyzed for structures at the interface and SFG spectroscopy through the VVCF method by DFT-MD in refs. 20,21. Here with the knowledge of the individual populations, SFG spectra are calculated with eq. 8 with the "pop model". Results are in Fig. 3C, with each population contribution illustrated in the corresponding color code and line style as used in Fig. 2. The SFG spectra of amorphous silica/water interfaces of both 3.5 and 4.5 OHs/nm^2 are in excellent agreement with that from the VVCF DFT-MD method. The individual spectral contributions from each population manage to capture the decrease of ratio between the spectral intensities of OH_{up} quasi-free (pop 3) and OH_{up}-SiOH (pop 4), with the increase in the degree of hydroxylation of the surface (from 3.5 to 4.5 OHs/nm^2). This informs us that more BIL-water OHs are H-bonded to SiOHs and less are pointing to the siloxane bridge (hence being quasi-free), when the hydroxylation degree increases.

Unlike the neutral amorphous silica/water interfaces, the aqueous α-quartz (0001) surface is slightly charged despite under PZC condition, because of the alternatively distributed in-plane and out-of-plane silanols at the surface[21,46,47]. This results in a $\Im(\chi^{(2)}_{DL}(\omega))$ contribution to the total $\Im(\chi^{(2)}(\omega))$ compared to the other two amorphous silica/water interfaces, for which $\Im(\chi^{(2)}_{DL}(\omega))$ equals zero. The DL-water contribution calculated with the "pop model" from Eq. 7 is thus added to the $\Im(\chi^{(2)}_{BIL}(\omega))$ of Eq. 8 for direct comparison with the VVCF DFT-MD calculated spectrum. Note that the "pop model" confirms that SFG-DL equals zero for the amorphous surfaces. The "pop model" calculated SFG spectrum shown in Fig. 3 C-bottom panel has a blue-shifted positive peak by 100 cm^{-1} compared to the VVCF DFT-MD calculation. Since the OH_{up}-SiOH $\beta^{eff}(\omega)$ is parameterized from an amorphous silica/water interface with a hydroxylation degree of 4.5 OHs/nm^2, this frequency shift due to stronger H-bonds between BIL-water and SiOHs of the crystalline silica surface can not be captured by the parameterization done here. However, this does not impede us from interpreting the spectrum by attributing each spectral contribution to the corresponding population and capturing the population variation of the interface components under different conditions.

4. The "Pop Model" Applied to Surface Chemistry in the Phonon Range

As regards surfaces, only one population (pop 6) corresponds to the surface in terms of out-of-plane silanol groups pointing towards and H-bonded to water as an SFG-active population in the O-H stretch range. This population does not inform on the actual surface structure organization of silica. To unravel the surface structure and chemistry of silica in contact with water, we now go from the high O-H stretch range (3000 – 3800 cm^{-1}) to the lower phonon frequency range, corresponding to the Si-O stretch region (750 – 1100 cm^{-1}) and parameterize our pop model also on this vibrational range.

The molecular understanding of surface chemistry of silica surfaces in contact with water has long been a hard nut to crack due to difficulties in probing buried oxide surfaces[48]. In 1992, Eisenthal and coworkers[49] first employed the second harmonic generation (SHG) technique to monitor the silica/water interface under different pH

conditions to uncover surface chemistry. They discovered two pKa values associated with surface silanol deprotonation, giving rise to the historical bimodal titration behavior, which has been extensively discussed in the literature. A substantial amount of both experimental[50–52] and theoretical works[53–56] have been dedicated to investigating the microscopic origins of such bimodal behavior of silica surface acidities. As a result, the surface chemistry behind the bimodal behavior has been highly debated.

In order to fully uncover the surface chemistry of silica in contact with water, we have carried out in ref.[57] DFT-MD simulations in conjunction with calculation of SFG spectra in the Si-O stretch frequency range to elucidate the silica surface structure under different pH conditions. The simulations revealed an unexpected surface chemistry of the silica surface. In the pH range between pH∼2 (Point of Zero Charge for silica) to 6.7, the surface is fully hydroxylated. At $6.7 < \text{pH} < 11.6$, the acidic silanols start to be deprotonated. The simulations showed that each deprotonated oxygen attacks and covalently bonds to a neighboring Si-O(H), forming a five-coordinated and negatively charged Si species, denoted Si(5c)O^-. When pH rises above 11.6, the basic SiOH groups are now deprotonated to SiO^- (i.e. the normal four-coordinated and negatively charged Si(4c)O^-, denoted SiO^-) while Si(5c)O^- species become unstable, split up and form SiO^-.

With this picture, the surface coverages of each surface species/population, i.e. SiOH, Si(5c)O^- and SiO^-, as a function of pH is known through Henderson-Hasselbalch equations (Fig. 4A). The $\beta^{eff}(\omega)$ SFG fingerprints of each surface population are calculated using the same VVCF method as for water in the O-H stretch region in the previous applications. They are displayed in Fig. 4B. (all detailed in Section computational methodology).

The pH-dependent theoretical SFG spectra of the silica surface in the Si-O stretch region can thus be constructed using the "pop model", as for water O-H spectra before, without requiring DFT-MD simulations at a given pH. These DFT-MD (as well as FF-MD) would be impossible to obtain because of the limited box sizes that can not take into account varying pH and ionic strength conditions.

In Fig. 4C, we hence observe that at pH 2, where the silica surface is fully hydroxylated and covered by SiOH groups, the spectrum showed a single peak arising from Si-OH stretching vibrations at ∼920 cm^{-1}. As pH increases, the silica surface starts to be deprotonated, resulting in a decrease in the intensity of the Si-OH peak, accompanied by a redshift to ∼840 cm^{-1} before disappearing at ∼pH 10. This corresponds to the active zone of the Si(5c)O^- species in the spectroscopy. Above pH 10, a new mode at ∼1000 cm^{-1} emerged, which is due to the Si-O^- stretching mode only.

An experimental scheme enabling in situ vibrational spectroscopy of oxide surfaces in liquid water was developed by Liu *et al.* in ref.[57] and applied to the amorphous

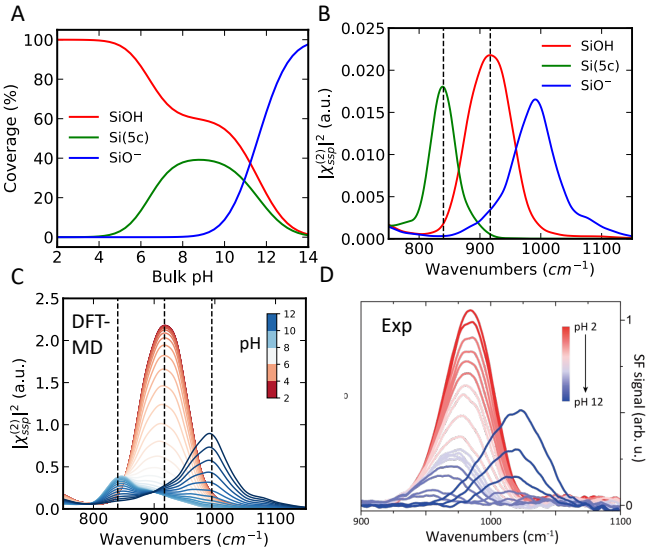

FIG. 4. Theoretical pH-dependent SFG calculation: (A) Surface coverages of each population as a function of pH deduced by Henderson-Hasselbalch equations based on the pKa values of surface silanols from metadynamics simulations. (B) SFG fingerprints in the Si-O stretch frequency range of each surface population calculated from DFT-MD simulations based on VVCF method. (C) Theoretical pH-dependent SFG spectra of amorphous silica surface in contact with water in the Si-O stretch frequency range constructed by "pop model" through combination of (A) and (B). (D) Experimental pH-dependent SFG spectra of amorphous silica surface in contact with water in the Si-O stretch frequency range probed by Liu *et al.* in ref.[57].

silica surface in water to record the pH-dependent SFG spectroscopy of silica in the Si-O stretch region (fig. 4D). The agreement between experimental pH-dependent SFG spectroscopy and theoretical SFG spectroscopy obtained through the "pop model" validates at the same time the "pop model" and the newly discovered Si(5c)O^- species. Note that the shifts in the absolute positions of the peaks between theory and experiment are due to the DFT functional. More details are in ref[57].

CONCLUSIONS

In conclusion, the "pop model" presents a simplified and efficient method for constructing and simultaneously interpreting SFG spectra of aqueous interfaces. Our approach has the potential to broaden the scope of theoretical SFG calculations, as well as simplify (and in the near future automatize with the help of machine learning techniques) the interpretation of experimental spectra, in various scientific and technological applications. It allows to calculate accurate spectra from any simulation type, including the cheapest force field-based molecular dynamics as well as Monte Carlo simulations. Strikingly, the "pop model" is even able to construct theoretical spec-

tra without simulations, directly from analytical models. Thanks to this, the "pop model" opens theoretical spectroscopy at various interfacial conditions that are out of reach of the current methods, such as pH and low electrolytic concentrations.

COMPUTATIONAL METHODOLOGY

The details of DFT-MD and classical MD simulations with the SPC/E force field for the air/water interface are presented in ref. 37. The DFT-MD simulation details of the graphene/water interface and the silica/water interface with a hydroxylation degree of 3.5OHs/nm^2 are presented in ref. 22. We refer to ref. 21 for the DFT-MD simulation details of silica/water interface with a hydroxylation degree of 4.5OHs/nm^2 and of $\alpha-$quartz (0001)/water interface. The simulation details for the SAM/water interfaces are presented in ref. 36. We refer to ref. 57 for all the simulation details related to the surface chemistry of aqueous silica and the phonon range spectroscopy, including details on the DFT-MD metadynamics that were performed for the pKa calculations of surface silanols.

All the VVCF-based DFT-MD theoretical SFG spectra are calculated following the approach introduced in ref. 17. For the water contribution to $\chi^{(2)}(\omega)$, individual molecular dipole moments and polarizability tensors are calculated with the model from ref. 17, supposing that only the O-H stretching motions contribute to the spectrum in the high frequency region ($> 3000\ cm^{-1}$). All details of derivation for the theoretical expression of a SFG spectrum can be found in refs. 19,21 with the parameterization of APT (Atomic Polar Tensor) and Raman tensors of water found in ref. 17. For the out-of-plane SiOH contribution to the SFG spectrum in the O-H stretch region, the parameterization of APT and Raman tensors is adopted from ref. 21. For SFG spectra calculated in the Si-O stretch region, the parameterization of APT and Raman tensors are set to unity since we are more interested in the influence on the SFG intensity as the result of a change in group population (as it happens when changing the pH) than in the actual change in the SFG activity of the group. For a more accurate description of the activity, the APT and Raman tensors must be rigorously evaluated from ab initio calculations as in ref. 17 for water.

For all structural analyses of H-bonds, the H-bond definition proposed by White and co-workers[58] has been adopted, with $O(-H)\cdots O \leq 3.2$ Å and the $O-H\cdots O$ angle between 140° and 220°. We have shown that any other definition from the literature leads to the same final conclusions.

ACKNOWLEDGMENTS

We acknowledge HPC resources from GENCI-France Grant 072484 (CINES/IDRIS/TGCC). We also acknowledge PRACE for awarding access to the Fenix infrastructures at GENCI-TGCC-France, which are partially funded by the European Union's Horizon 2020 research and innovation program through the ICEI project under grant agreement No. 800858. S.P. acknowledges financial support from European Research Council (ERC) Advanced Grant 695437 THz-Calorimetry and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC2033-390677874-RESOLV for funding this investigation.

AUTHOR DECLARATIONS

Conflict of interest

The authors have no conflicts to disclose.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

[1] J. Hunt, P. Guyot-Sionnest, and Y. Shen, *cpl*, 1987, **133**(3), 189–192.
[2] X. Zhu, H. Suhr, and Y. Shen, *prb*, 1987, **35**(6), 3047.
[3] Y. Shen, *Fundamentals of sum-frequency spectroscopy*, Cambridge University Press, 2016.
[4] Q. Du, R. Superfine, E. Freysz, and Y. Shen, *Phys. Rev. Lett.*, 1993, **70**(15), 2313.
[5] A. Eftekhari-Bafrooei and E. Borguet, *J. Am. Chem. Soc.*, 2010, **132**(11), 3756–3761.
[6] Z. Zhang, L. Piatkowski, H. J. Bakker, and M. Bonn, *Nat. Chem.*, 2011, **3**(11), 888–893.
[7] C.-S. Hsieh, M. Okuno, J. Hunger, E. H. Backus, Y. Nagata, and M. Bonn, *Angew. Chem. Int. Ed.*, 2014, **53**(31), 8146–8149.
[8] C. Tian and Y. Shen, *Surf. Sci. Rep.*, 2014, **69**(2-3), 105–131.
[9] A. M. Jubb, W. Hua, and H. C. Allen, *Annu. Rev. Phys. Chem.*, 2012, **63**, 107–130.
[10] L. Dalstein, E. Potapova, and E. Tyrode, *Phys. Chem. Chem. Phys.*, 2017, **19**(16), 10343–10349.
[11] A. Morita and J. T. Hynes, *J. Phys. Chem. B.*, 2002, **106**, 673–685.
[12] A. Morita and T. Ishiyama, *Phys.Chem.Chem.Phys.*, 2008, **10**, 5801–5816.
[13] Y. Ni and J. L. Skinner, *J. Chem. Phys.*, 2016, **145**(3), 031103.
[14] T. Ishiyama, T. Imamura, and A. Morita, *Chem. Rev.*, 2014, **114**(17), 8447–8470.
[15] T. Ishiyama and A. Morita, *Ann. Rev. Phys. Chem.*, 2017, **68**, 355–377.
[16] T. Ohto, K. Usui, T. Hasegawa, M. Bonn, and Y. Nagata, *jcp*, 2015, **143**(12), 124702.
[17] R. Khatib, E. H. Backus, M. Bonn, M.-J. Perez-Haro, M.-P. Gaigeot, and M. Sulpizi, *Sci. Reports*, 2016, **6**(1), 1–10.
[18] S. Pezzotti, D. R. Galimberti, and M.-P. Gaigeot, *J. Phys. Chem. Lett.*, 2017, **8**(13), 3133–3141.
[19] S. Pezzotti, D. R. Galimberti, Y. R. Shen, and M.-P. Gaigeot, *Phys. Chem. Chem. Phys.*, 2018, **20**, 5190–5199.

[20] J. D. Cyran, M. A. Donovan, D. Vollmer, F. S. Brigiano, S. Pezzotti, D. R. Galimberti, M.-P. Gaigeot, M. Bonn, and E. H. Backus, *Proc. Natl. Acad. Sci.*, 2019, **116**(5), 1520–1525.

[21] S. Pezzotti, D. R. Galimberti, and M.-P. Gaigeot, *Phys. Chem. Chem. Phys.*, 2019, **21**(40), 22188–22202.

[22] S. Pezzotti, A. Serva, F. Sebastiani, F. S. Brigiano, D. R. Galimberti, L. Potier, S. Alfarano, G. Schwaab, M. Havenith, and M.-P. Gaigeot, *J. Phys. Chem. Lett.*, 2021, **12**(15), 3827–3836.

[23] Y.-C. Wen, S. Zha, X. Liu, S. Yang, P. Guo, G. Shi, H. Fang, Y. R. Shen, and C. Tian, *Phys. Rev. Lett.*, 2016, **116**(1), 016101.

[24] T. Joutsuka, T. Hirano, M. Sprik, and A. Morita, *Phys. Chem. Chem. Phys.*, 2018, **20**, 3040–3053.

[25] G. Gonella, C. Lutgebaucks, A. G. F. de Beer, and S. Roke, *J. Phys. Chem. C.*, 2016, **120**, 9165–9173.

[26] A. P. Willard and D. Chandler, *J. Phys. Chem. B*, 2010, **114**(5), 1954–1958.

[27] S.-h. Urashima, A. Myalitsin, S. Nihonyanagi, and T. Tahara, *J. Phys. Chem. Lett.*, 2018, **9**(14), 4109–4114.

[28] S. Pezzotti, D. R. Galimberti, Y. R. Shen, and M.-P. Gaigeot, *Minerals*, 2018, **8**(7), 305.

[29] F. Tang, T. Ohto, S. Sun, J. R. Rouxel, S. Imoto, E. H. Backus, S. Mukamel, M. Bonn, and Y. Nagata, *Chem. Rev.*, 2020, **120**(8), 3633–3667.

[30] Q. Du, E. Freysz, and Y. R. Shen, *Science*, 1994, **264**(5160), 826–828.

[31] A. Tuladhar, S. Dewan, S. Pezzotti, F. S. Brigiano, F. Creazzo, M.-P. Gaigeot, and E. Borguet, *J. Am. Chem. Soc.*, 2020, **142**(15), 6991–7000.

[32] H. J. Berendsen, J. v. Postma, W. F. Van Gunsteren, A. DiNola, and J. R. Haak, *The Journal of chemical physics*, 1984, **81**(8), 3684–3690.

[33] C. Tainter, Y. Ni, L. a. Shi, and J. Skinner, *J. Phys. Chem. Lett.*, 2013, **4**(1), 12–17.

[34] Y. Ni and J. Skinner, *J. Chem. Phys.*, 2016, **145**(3), 031103.

[35] G. R. Medders and F. Paesani, *J. Am. Chem. Soc.*, 2016, **138**(11), 3912–3919.

[36] W. Chen, S. E. Sanders, B. Ozdamar, D. Louaas, F. S. Brigiano, S. Pezzotti, P. B. Petersen, and M.-P. Gaigeot, *J. Phys. Chem. Lett.*, 2023, **14**, 1301–1309.

[37] A. Serva, S. Pezzotti, S. Bougueroua, D. R. Galimberti, and M.-P. Gaigeot, *J. Mol. Struct.*, 2018, **1165**, 71–78.

[38] J. I. Monroe and M. S. Shell, *Proc. Natl. Acad. Sci.*, 2018, **115**(32), 8093–8098.

[39] B. C. Dallin and R. C. Van Lehn, *J. Phys. Chem. Lett.*, 2019, **10**(14), 3991–3997.

[40] J. Monroe, M. Barry, A. DeStefano, P. Aydogan Gokturk, S. Jiao, D. Robinson-Brown, T. Webber, E. J. Crumlin, S. Han, and M. S. Shell, *Annu. Rev. Chem. Biomol.*, 2020, **11**, 523–557.

[41] C. Tian and Y. Shen, *Proc. Natl. Acad. Sci.*, 2009, **106**(36), 15148–15153.

[42] S. Ye, S. Nihonyanagi, and K. Uosaki, *Phys. Chem. Chem. Phys.*, 2001, **3**(16), 3463–3469.

[43] S. E. Sanders and P. B. Petersen, *J. Chem. Phys.*, 2019, **150**(20), 204708.

[44] Q. Du, E. Freysz, and Y. R. Shen, *Phys. Rev. Lett.*, 1994, **72**, 238.

[45] V. Ostroverkhov, G. A. Waychunas, and Y. Shen, *prl*, 2005, **94**(4), 046102.

[46] M. Sulpizi, M.-P. Gaigeot, and M. Sprik, *J. Chem. Theory Comput.*, 2012, **8**(3), 1037–1047.

[47] M. Pfeiffer-Laplaud and M.-P. Gaigeot, *J. Phys. Chem. C*, 2016, **120**(26), 14034–14047.

[48] U. Diebold, S.-C. Li, and M. Schmid, *Annu. Rev. Phys. Chem.*, 2010, **61**, 129–148.

[49] S. Ong, X. Zhao, and K. B. Eisenthal, *cpl*, 1992, **191**(3-4), 327–335.

[50] V. Ostroverkhov, G. A. Waychunas, and Y. Shen, *prl*, 2005, **94**(4), 046102.

[51] J. D. Fisk, R. Batten, G. Jones, J. P. O'Reill, and A. M. Shaw, *jpcb*, 2005, **109**(30), 14475–14480.

[52] A. M. Darlington and J. M. Gibbs-Davis, *jpcc*, 2015, **119**(29), 16560–16567.

[53] M. Sulpizi, M.-P. Gaigeot, and M. Sprik, *J. Chem. Theory Comput.*, 2012, **8**(3), 1037–1047.

[54] M. Pfeiffer-Laplaud, D. Costa, F. Tielens, M.-P. Gaigeot, and M. Sulpizi, *jpcc*, 2015, **119**(49), 27354–27362.

[55] F. Tielens, M. Gierada, J. Handzlik, and M. Calatayud, *Catal*, 2020, **354**, 3–18.

[56] M. Gierada, F. De Proft, M. Sulpizi, and F. Tielens, *jpcc*, 2019, **123**(28), 17343–17352.

[57] X. Li, F. Siro Brigiano, S. Pezzotti, X. Liu, W. Chen, H. Chen, Y. Li, H. Li, Y. R. Shen, M.-P. Gaigeot, and W.-T. Liu, *under review in Nature.*, 2023.

[58] J. A. White, E. Schwegler, G. Galli, and F. Gygi, *J. Chem. Phys.*, 2000, **113**(11), 4668–4673.