






 Cite this: *Phys. Chem. Chem. Phys.*, 2022, 24, 19463

# First hyperpolarizability of water in bulk liquid phase: long-range electrostatic effects included via the second hyperpolarizability†

 Guillaume Le Breton, , Oriane Bonhomme, , Emmanuel Benichou  and Claire Loison \*

The molecular first hyperpolarizability  $\beta$  contributes to second-order optical non-linear signals collected from molecular liquids. For the Second Harmonic Generation (SHG) response, the first hyperpolarizability  $\beta(2\omega, \omega, \omega)$  often depends on the molecular electrostatic environment. This is especially true for water, due to its large second hyperpolarizability  $\gamma(2\omega, \omega, \omega, 0)$ . In this study we compute the electronic  $\gamma(2\omega, \omega, \omega, 0)$  and  $\beta(2\omega, \omega, \omega)$  for water molecules in liquid water using QM/MM calculations. The average value of  $\gamma(2\omega, \omega, \omega, 0)$  is smaller than the one for the gas phase, and its standard deviation among the molecules is relatively small. In addition, we demonstrate that the average bulk second hyperpolarizability  $\langle \gamma(2\omega, \omega, \omega, 0) \rangle$  can be used to describe the electrostatic effects of the distant neighborhood on the first hyperpolarizability  $\beta(2\omega, \omega, \omega)$ . In comparison with more complex schemes to take into account long-range effects, the approximation is simple, and does not require any modifications of the QM/MM implementation. The long-range correction can be added explicitly, using an average value of  $\gamma$  for water in the condensed phase. It can also be easily added implicitly in QM/MM calculations through an additional embedding electric field, without the explicit calculation of  $\gamma$ .

 Received 17th February 2022,  
 Accepted 30th June 2022

DOI: 10.1039/d2cp00803c

rsc.li/pccp

## 1 Introduction

Non-Linear Optical (NLO) technique(s) are increasingly used to probe structural properties of matter. In Second Harmonic Generation (SHG) or Hyper-Rayleigh Scattering (HRS), two photons with the same fundamental frequency interact with a nonlinear material that generates a new photon with twice the energy of the initial photons.<sup>1</sup> SHG-based technologies have been developed to investigate simple and complex liquids,<sup>2,3</sup> biomimetic systems,<sup>4–6</sup> or even biological materials.<sup>7,8</sup> One interesting property of non-resonant SHG is the sensitivity of the response to the electrostatic environment.<sup>9</sup> Historically, the prototypical application is the electric-field-induced second harmonic generation (EFISHG) of molecules in a gas phase.<sup>10</sup> During such experiments, a macroscopic electrostatic field  $\mathbf{E}^{\text{DC}}$  is applied and the SHG response of the system is described using second and third order susceptibility tensor:  $\chi^{(2)}(2\omega, \omega, \omega)$  and  $\chi^{(3)}(2\omega, \omega, \omega, 0)$ , respectively. For a mesoscopic unitary

volume, the induced total dipole moment at the second harmonic frequency  $\mathbf{P}^{2\omega}$  expressed as a sum of two terms:

$$\mathbf{P}^{2\omega} \propto \chi^{(2)}(2\omega, \omega, \omega): \mathbf{E}^\omega \mathbf{E}^\omega + \chi^{(3)}(2\omega, \omega, \omega, 0): \mathbf{E}^\omega \mathbf{E}^\omega \mathbf{E}^{\text{DC}}, \quad (1)$$

where  $\mathbf{E}^\omega$  is the exciting electromagnetic field at the fundamental frequency, and  $\mathbf{E}^{\text{DC}}$  is the external electrostatic field that potentially modifies (i) the orientation of dipoles, and (ii) the electronic-dependent molecular intrinsic SHG response. More recently, EFISHG has received renewed interest as a non-intrusive tool for probing electric fields in diluted media with spatial and temporal definition, using laser pulses.<sup>11</sup>

Eqn (1) was also applied to condensed phases,<sup>12,13</sup> and the EFISHG is an established technique to determine the first and the second hyperpolarizabilities of compounds in solutions.<sup>14</sup> Even more, studies on liquid/solid or liquid/air interfaces have reported the evolution of the Surface-SHG (S-SHG) response when the surface charge is modulated. Frameworks based on eqn (1) for the fluid nearby the surface have been commonly used in the S-SHG community to extract a surface potential or an effective surface charge.<sup>15–19</sup> But the interpretation of the different terms in the S-SHG intensity generated by aqueous solutions is still the subject of many recent works<sup>20–23</sup> requiring theoretical calculations at the molecular level.<sup>24</sup>

University of Lyon, Université Claude Bernard Lyon 1, CNRS, Institut Lumière Matière, F-69622, Villeurbanne, France. E-mail: [claire.loison@univ-lyon1.fr](mailto:claire.loison@univ-lyon1.fr)

† Electronic supplementary information (ESI) available: Methodological details; symmetry, averages and standard deviations of water second hyperpolarizability in the liquid phase; heterogeneity of the electrostatic field generated by the neighborhood; impact of second hyperpolarizability fluctuations on the  $\gamma$  long correction. See DOI: <https://doi.org/10.1039/d2cp00803c>



To provide an interpretation at the microscopic scale, eqn (1) can also be rewritten at the molecular level:

$$\begin{aligned} \mu^{2\omega} &= \frac{1}{2!} \beta^{(2)}(2\omega, \omega, \omega) : \mathbf{e}^\omega \mathbf{e}^\omega \\ &+ \frac{1}{2!} \gamma^{(3)}(2\omega, \omega, \omega, 0) : \mathbf{e}^\omega \mathbf{e}^\omega \mathbf{e}^{\text{DC}} \end{aligned} \quad (2)$$

where  $\mu^{2\omega}$  is the induced dipole moment of the molecule at the second harmonic frequency,  $\mathbf{e}^\omega$  the exciting electromagnetic field in the molecular frame,  $\beta^{(2)}(2\omega, \omega, \omega)$  and  $\gamma^{(3)}(2\omega, \omega, \omega, 0)$  are respectively the first and second molecular hyperpolarizabilities, noted  $\beta$  and  $\gamma$  later on. The second hyperpolarizability  $\gamma$  describes the variations of the first hyperpolarizability value under a static external electrostatic field  $\mathbf{e}^{\text{DC}}$ , present in the molecular frame. This approach permits to link the EFISHG macroscopic measurements for the gas phase to quantum calculations of the molecular first hyperpolarizability  $\beta$ .<sup>25–30</sup>

The application of field expansions such as eqn (2) for condensed phases is very useful for the interpretation of S-SHG of aqueous solutions, and the values of  $\beta$  and  $\gamma$  for water in liquid water become parameters of the equation. To obtain values of  $\beta$  and  $\gamma$  for water in liquid water, one can use Quantum Mechanical/Molecular Mechanic (QM/MM) calculations.<sup>9,24,26,31,32</sup> In such approaches, the electrostatic effect of the environment is included within the QM calculations of a molecular hyperpolarizability  $\beta^{\text{env}}$  for a given electrostatic environment, so that

$$\mu^{2\omega} = \frac{1}{2!} \beta^{\text{env}} : \mathbf{e}^\omega \mathbf{e}^\omega, \quad (3)$$

where  $\beta^{\text{env}}$  depends on the electric field exerted by the environment  $\mathbf{e}^{\text{env}}$  (T convention<sup>33</sup>). Using the values of  $\beta^{\text{env}}$ , one can also calculate  $\gamma^{\text{env}}$  from its definition as  $\gamma_{ijkl} = \frac{\delta\beta_{ijk}}{\delta e_l}$ , where  $\beta_{ijk}(e_l)$  is the  $ijk$ -component of the first hyperpolarizability of a molecule on which an extra electrostatic field  $e_l$  is applied along the  $l$  molecular direction. In the present work, we first provide the  $\beta^{\text{env}}$  and  $\gamma^{\text{env}}$  tensors of individual water molecules in the liquid phase using QM/MM approaches at an optical wavelength of 800 nm (typically used for experiments<sup>34,35</sup>). Relatively to previous works,<sup>9,26,32</sup> we detail all the components of both tensors, and discuss the distribution of  $\gamma$  among water molecules.

In a second step, possessing a database of  $\{\beta^{\text{env}}, \gamma^{\text{env}}\}$  for individual molecules in liquid water, we can additionally scrutinize the validity of the field perturbation approach represented by eqn (2). Similar field expansions for the calculation of solvent effects have been tested for observables such as nuclear magnetic shieldings,<sup>36</sup> or linear and non linear optical properties.<sup>37</sup> However, the precision of the field expansion for the non-linear optical properties were limited,<sup>37</sup> and more terms are needed in the perturbation expression.<sup>38</sup> Concerning the first hyperpolarizability of water, Liang *et al.*<sup>9</sup> have compared  $\beta^{\text{env}}$  of water within the liquid phase obtained by QM/MM calculations to the linear field expansion of eqn (2), where  $\mathbf{e}^{\text{DC}}$  was the electrostatic field generated by the

surrounding water molecules, noted  $\mathbf{e}^{\text{env}}$ , and  $\beta$  and  $\gamma$  were the gas-phase values: a quantitative difference emerged. It is therefore expected that the strong solvation of water molecules within liquid water cannot be simplified to eqn (2). Nevertheless, in this work, we show that the perturbation field Taylor expansion can be useful to predict long-range electrostatic effects on the hyperpolarizability of water molecules in their liquid phase.

In the following, Section 2 describes the numerical details. In Section 3, the individual values of  $\gamma^{\text{env}}$  and  $\beta^{\text{env}}$  are first reported. Then, we show that the electrostatic field generated by the environment, namely  $\mathbf{e}^{\text{env}}$ , is strongly heterogeneous in space. Based on this result, we propose to separate the effects of the electric field generated by the environment into a short- and a long-range part. Using this separation, we show that, for a water molecule in the bulk phase, the effect of the long-range electrostatic environment on  $\beta^{\text{env}}$  can be included using a linear correction proportional to  $\gamma$ .

## 2 Method

The calculations are based on a sequential QM/MM approach in two steps. First, we use a classical Molecular Dynamics (MD) simulation of bulk water to obtain typical structures of the liquid. Then, we compute the first and second hyperpolarizabilities of individual water molecules at the Density Functional Theory (DFT) level, within an electrostatic embedding framework. To investigate fluctuations due to the changing environment, statistics are performed over numerous configurations of the MD simulation.

### 2.1 Molecular dynamics

LAMMPS,<sup>39</sup> V.11.08.2017 is used to perform the MD simulation along with the rigid TIP4P/2005 water force field.<sup>40</sup> 15625 rigid TIP4P/2005 water molecules are placed in a simulation box (approx  $7.8 \times 7.8 \times 7.8 \text{ nm}^3$ ) to form a 3D-periodic bulk system. We have used an isothermal, isobaric ensemble (NPT) with Nose–Hoover thermostat at 300 K ( $\tau_T = 0.4 \text{ ps}$ ) and Nose–Hoover barostat at 1 atm ( $\tau_P = 2 \text{ ps}$ ).<sup>41</sup> 1 ns of equilibration is performed before the 1 ns production run, both with a time step of 2 fs. Both electrostatic and Lennard-Jones intermolecular interactions are computed using the long-range Particle-Particle-Particle-Mesh (PPPM) formalism.<sup>42,43</sup> Neighbor lists are updated every time step within a radius of 10 Å. This simulation leads to a density of about  $0.996 \text{ kg L}^{-1}$ . Such a large system was necessary to investigate environmental effects up to 40 Å.

### 2.2 First hyperpolarizability in the liquid phase

To compute the first hyperpolarizability of water molecules within the liquid phase, an explicit environment composed of point charges is used: the Polarizable Embedding model at the zero<sup>th</sup>-order (PE0),<sup>44,45</sup> implemented in the DALTON software,<sup>46</sup> release 2018.2 package. As in our previous recent work,<sup>32</sup> the QM calculations were carried out on individual water molecules. Point charges represent the surrounding water molecules,



and the same electrostatic description is used as for the MD (TIP4P/2005 model). The MD trajectories are used directly in the QM calculations without further optimization, in contra. We define the parameter  $R_c$  as the maximal distance upon which neighbors are included in the PE formalism. The electrostatic field generated by this environment is spatially-heterogeneous in the vicinity of the target molecule and creates a new potential for the target molecule electronic Hamiltonian.

The QM/MM approach could be refined by optimizing geometries,<sup>47</sup> or by adding dipole, quadrupole or polarizability of the environment,<sup>45,48–51</sup> but we shall see in the results that eqn (2) becomes relevant only for long-range effects, where the point-charges electrostatic field is dominant. We have therefore chosen to keep the environment description as simple and robust as possible. To optimize parts of the QM/MM routines, the home-made software FROG is used.‡

The whole procedure is composed of the following steps: (1) perform MD simulation using LAMMPS. (2) Read the MD trajectories using FROG to build the electrostatic environment for each water molecule defined into a sphere of radius  $R_c$  centered on the mass center of the target molecule. If a neighbor possesses at least one atom in the sphere, the whole molecule is considered in the environment. The output files for DALTON are created for 300 molecules for 8 selected MD configurations. (3) For each of the 2400 target molecules, calculate the electrostatic field, and its gradients, generated by the environment using FROG. (4) For each of the 2400 target molecules, calculate the value of  $\beta$  at the fundamental wavelength of 800 nm using DFT with the functional CAM-B3LYP<sup>52</sup> and the basis d-aug-cc-pVTZ,<sup>53</sup> using the quadratic response formalism as implemented in DALTON.<sup>54</sup> (5) Read all DALTON output files and analyze the data using FROG.

The largest inaccuracy of our approach is due to the QM method and the corresponding functional. For the first hyperpolarizability of water molecules, the values in vacuum obtained with our method (DFT/CAM-B3LYP) are typically overestimated by about 10%<sup>32</sup> in comparison to the golden standard for QM/MM calculation of small molecules (CCSD<sup>55–57</sup>). Both calculation schemes use the TIP4P/2005 rigid force field, the MM size  $R_c = 15 \text{ \AA}$  and the basis is d-aug-pVTZ. The basis set d-aug-cc-pVTZ<sup>53</sup> used is widespread across the community<sup>26,55–57</sup> and presents a good balance between accuracy and cost. In a previous work,<sup>32</sup> we have observed a smooth convergence of the first hyperpolarizability relative to the basis-set size for water in bulk water, and similar absolute values of the nuclear and electronic contributions to the electrostatic and polarization energies. Therefore, we estimate that spill-out effects are not dominant,<sup>50</sup> even if we use the large d-aug-cc-pVTZ basis set. We have not taken into account the nuclear degree of freedom (vibrational effects) since Beaujean *et al.*<sup>58</sup> have shown that this effect is small for the second hyperpolarizability of water, especially at small wavelengths.

‡ This code is deposited on Zenodo platform <https://doi.org/10.5281/zenodo.5998193>, and available on demand to the authors.

### 2.3 Second hyperpolarizability in the liquid phase

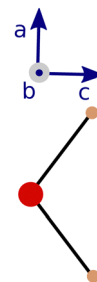
In this work, we are interested in the evolution of the first hyperpolarizability under a static and homogeneous electrostatic field, which is reflected by  $\gamma(2\omega, \omega, \omega, 0)$  noted here  $\gamma$ . For each individual molecule, the second hyperpolarizability  $\gamma$  can be computed from the first hyperpolarizability  $\beta$  using the Finite Field (FF) framework. For that, we use the PE formalism described above to compute  $\beta$  when an extra homogeneous electrostatic field is applied, and then compute  $\gamma$  according to

$$\gamma_{ijkl} = \frac{\delta\beta_{ijk}}{\delta e_l} \simeq \frac{\beta_{ijk}(e_l) - \beta_{ijk}(0)}{e_l}, \quad (4)$$

where  $\beta_{ijk}(e_l)$  is the first hyperpolarizability of a molecule on which an extra static electrostatic field  $e_l$  is applied along the  $l$  molecular direction. Numerically, the  $\gamma_{ijkl}$  is obtained for each molecule as the slope of a linear fit performed on seven calculations with  $e_l$  from 0 to  $1.5 \times 10^{-3}$  a.u. Noticeably, during the  $\gamma$  calculation, the target molecule is embedded in a total electrostatic field composed of a spatially-heterogeneous contribution from its PE environment plus the spatially-homogeneous  $e_l$ . More details regarding the FF procedure can be found in the ESI,† Section S1.1. The first and second hyperpolarizabilities are given in atomic units, our conventions are defined by eqn (2) and (4). They are expressed in the molecular frame  $\{a, b, c\}$  with the axis  $c$  along the dipole moment pointing from the Oxygen towards the Hydrogens, and the atoms within the  $\{a, c\}$  plane, see insert in Table 1. Noticeably, one could alternatively have used the response scheme to the third order, as implemented in Dalton, to obtain the second hyperpolarizability. For a given QM/MM conformation, the numerical cost of the response scheme is about ten times one of the finite field (FF) scheme, for similar results (see ESI,† Table S1). The comparison of the FF and the response

Table 1 Water second hyperpolarizability  $\gamma^{(3)}(2\omega, \omega, \omega, 0)$  components, either in the vacuum ( $\gamma^{\text{vac}}$ ) or in its liquid phase ( $\gamma^{\text{env}}$ ) for an exciting wavelength of 800 nm, in atomic units. For the bulk phase, the average value  $\langle \gamma^{\text{env}} \rangle$  and the standard deviation  $\sigma[\gamma^{\text{env}}]$  are reported. The tensors components are expressed in the molecular frame  $\{a, b, c\}$ , see inserted figure. Comparing with response calculations, the error on  $\gamma$  due to the finite field differentiation is estimated at about 20 a.u. (see ESI, Table S1)

$ijkl$	Vacuum	Liquid	
	$\gamma^{\text{vac}}$	$\langle \gamma^{\text{env}} \rangle$	$\sigma[\gamma^{\text{env}}]$
aaaa	1110	740	140
bbbb	4090	2940	650
cccc	2000	1400	230
aabb	930	640	200
bbaa	1060	710	150
baab	1010	680	220
abba	940	660	140
aacc	590	370	60
ccaa	610	380	80
caac	610	380	60
acca	590	370	80
bbcc	1120	770	170
cccb	1010	720	180
cbbc	1010	730	150
bccb	1060	750	200



schemes provides an estimation of the numerical error due to the FF procedure: about 5 to 20 a.u.

Frequency dispersion can be important for  $\gamma$ : for example, Beaujean *et al.*<sup>58</sup> have shown a variation of about 25% between an excitation wavelength of 800 nm and at the infinite wavelength limit for  $\gamma(3\omega, \omega, \omega, \omega)$ . Here, even if we are using a Finite Field approach, the frequency dispersion of  $\beta$  related to the fundamental frequency (here 800 nm) is correctly reproduced since we use the first hyperpolarizability obtained using the frequency dependent response scheme.

#### 2.4 Embedding environment size: $R_c$

One of the objectives of this study is to understand the impact of the electrostatic environment on the first hyperpolarizability, distinguishing the effects of close and far neighbors. Therefore, different values of  $R_c$  are used for  $\beta$  calculations. In the following, the notation  $\beta^{\text{PE}}(R_c)$  denotes the hyperpolarizability obtained by a QM/MM calculation at a given  $R_c$  and  $\mathbf{e}(R_c)$  the electric field generated by this PE environment. In this notation,  $R_c = 0$  corresponds to a gas phase calculation. According to previous works,<sup>32,57,59</sup> a large environment (typically between 10 and 20 Å) is necessary to obtain a good convergence of the  $\beta$  values. Here, for each molecule, a reference calculation is performed using the PE formalism with an extremely large radius of  $R_f = 40$  Å, and we consider that  $\beta^{\text{env}} = \beta_{\text{PE}}(40 \text{ Å})$ . Similarly, we consider that the electrostatic field  $\mathbf{e}^{\text{env}} = \mathbf{e}(40 \text{ Å})$ . Note that such a large radius still does not represent the complete MD liquid phase, described using periodic boundary conditions.

On the contrary, the variation of  $\gamma$  with the size of the environment is not investigated in details here. Therefore, the PE radius used to compute the second hyperpolarizability  $\gamma$  is set to 10 Å. Yet, we have verified that the  $\gamma$  does not evolve much using larger radius in agreement with Osted *et al.*<sup>59</sup> (data not shown) and have noted it  $\gamma^{\text{env}}$ .

#### 2.5 Statistical averaging

The first and second hyperpolarizability in the liquid phase are obtained from 2400 configurations of the PE environment. These configurations are extracted from 8 MD snapshots, with 300 water molecules randomly selected in each frame. These frames are separated by 100 ps to ensure time decorrelation.<sup>32</sup> We have verified that this number of configurations is sufficient to reach convergence; indeed, the numerical error due to the configuration number is about 1% (see ESI,† Fig. S2 and S3). This is smaller than the error due to the DFT Finite field calculation relative to CCSD calculations (see ESI,† Tables S2 and S3). The  $C_{2v}$  symmetry of the first and second hyperpolarizability obtained by QM/MM confirms this good convergence (see the Section 3 and ESI,† Section S2.1).

## 3 Results

In the following, Section 3.1 first presents the second hyperpolarizability tensor  $\gamma$  calculated using QM/MM approaches for single water molecules within the liquid bulk phase, and the

standard deviation of the calculated values. Then, Section 3.2.1 presents some difficulties encountered when applying eqn (2) for condensed phases. We attribute them to the properties of the electrostatic field generated by the environment,  $\mathbf{e}^{\text{env}}$ , described in 3.2.2. Based on these results, and on eqn (2), in Section 3.2.3, we propose a correction to  $\beta$  that permits to include long-range interactions. Finally, we show that this correction enhances the precision of standard QM/MM calculations.

### 3.1 QM/MM results: second hyperpolarizability of water in the bulk phase

Table 1 presents the values of selected  $\gamma$  components calculated either in the vacuum or in liquid phase – the other non-zero components of  $\gamma$  are presented in ESI,† Table S4.

For the vacuum phase, the molecular  $C_{2v}$  symmetry is fulfilled: the only non-vanishing components are the  $\gamma_{iiii}$  and  $\gamma_{ijij}$  for  $i$  and  $j$  in  $\{a, b, c\}$ , plus their permuted terms<sup>1</sup> (see ESI,† Table S4). In the liquid phase, this symmetry is not valid for individual values, but is recovered for the average  $\langle \gamma^{\text{env}} \rangle$ . Our results agree with the ones by Osted *et al.*<sup>59</sup> within about 15%. We impute the discrepancies to both the difference of methods (DFT vs. CCSD, different MD models) and the difference of excitation wavelengths (800 vs. 1080 nm). More detailed comparisons with literature are provided in ESI,† Section S1.1.

Noticeably, when transferred into the laboratory frame, the average second hyperpolarizability becomes centro-symmetric (see ESI,† Table S5), indicating that our sampling of molecular orientation in the liquid phase is sufficient. Interestingly, the average value of all  $C_{2v}$ -authorized  $\gamma^{\text{env}}$  components are positive. Moreover, in the liquid phase, the  $\gamma^{\text{env}}$  components are about 30% smaller than the ones in the vacuum – see ESI,† Fig. S4.

Beyond the average values, the environment induces fluctuations of  $\gamma^{\text{env}}$ , and of  $\beta^{\text{env}}$ . A typical and relevant case is illustrated by Fig. 1, displaying the joint probabilities of  $\{\beta_{ccc}, \gamma_{ccc}\}$  obtained for the 2400 water molecules within their environment. Projections on the axes display the distributions of  $\beta_{ccc}$  and  $\gamma_{ccc}$  separately. As already discussed in the literature,<sup>32,57,59</sup> the water first hyperpolarizability fluctuations in the liquid phase are strong, with standard deviations similar to the absolute values of the averages – for instance  $\langle \beta_{ccc}^{\text{env}} \rangle = 4.1$  a.u. and  $\sigma[\beta_{ccc}^{\text{env}}] = 3.1$  a.u. For individual molecules, some  $\beta$  components, null in average, can be larger than the one with a net average (see ESI,† Fig. S4).

In complement to the results published by Osted *et al.*,<sup>59</sup> we show that the second hyperpolarizability  $\gamma$  also fluctuates, but in a smaller extent than  $\beta$ : the standard deviations of the different components  $\sigma[\gamma^{\text{env}}]$  are about 20% of their average values  $\langle \gamma^{\text{env}} \rangle$  – see Fig. 1 for  $\gamma_{ccc}$ , and ESI,† Table S4.

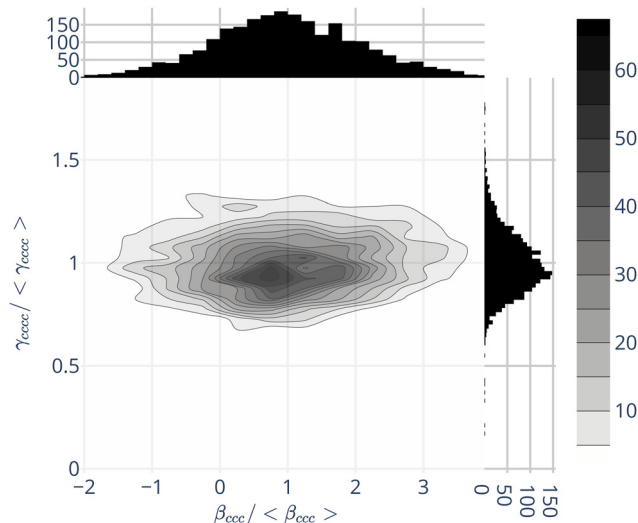
Fig. 1 moreover shows that the values of  $\beta$  and  $\gamma$  are not strongly correlated. Hence, in the following, we neglect the  $\gamma^{\text{env}}$  dispersion and attribute the same second hyperpolarizability  $\langle \gamma^{\text{env}} \rangle$  to all the water molecules.

### 3.2 Using $\gamma$ to calculate individual $\beta$ in condensed phases

After describing the second hyperpolarizability  $\gamma^{\text{env}}$  in the liquid phase, we will question if it can be used to predict the







**Fig. 1** Joint population of  $(\beta_{ccc}^{\text{env}}, \gamma_{cccc}^{\text{env}})$  obtained by QM/MM approaches for bulk liquid water. The independent ones are presented on the top and right for the  $\beta_{ccc}^{\text{env}}$  and  $\gamma_{cccc}^{\text{env}}$  respectively. The values are normalized with respect to their average ones:  $\langle \beta_{ccc}^{\text{env}} \rangle = 4.1$  a.u. and  $\langle \gamma_{cccc}^{\text{env}} \rangle = 1400$  a.u.

first hyperpolarizability of water molecules in the liquid phase using different approximations inspired by eqn (2). We shall compare the  $\beta^{\text{env}}$  obtained by QM/MM to linear expansions relative to a perturbing electric field.

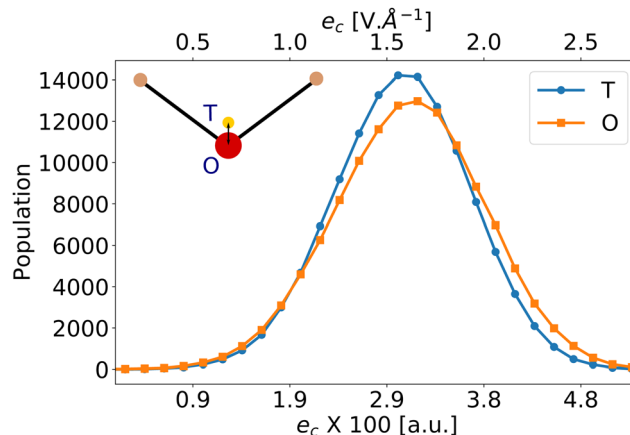
**3.2.1 Limitations of the gas-phase reference model.** Comparing eqn (2) and (3), one straightforward approach to include the effect of the environment on the first hyperpolarizability  $\beta$  is to use the second hyperpolarizability  $\gamma$  and the field created by the environment  $\mathbf{e}^{\text{env}}$ :

$$\beta^{\text{GR}} = \beta^{\text{vac}} + \gamma^0 \cdot \mathbf{e}^{\text{env}}, \quad (5)$$

where a linear perturbation is done relative to the Gas-phase Reference (GR), for which  $\mathbf{e}^{\text{env}} = 0$  and the reference first and second hyperpolarizability would be the one of the vacuum for instance. For pedagogical reasons, we will first highlight here the limitation of this gas-phase reference model, that was already mentioned in the literature.<sup>36,57</sup>

If the linear proportionality factor  $\gamma^0$  is known, and the same for all molecules, this approach would be very helpful: the liquid first hyperpolarizability depends only on  $\mathbf{e}^{\text{env}}$ . An MD simulation can provide  $\mathbf{e}^{\text{env}}$  for each molecule, and the liquid phase first hyperpolarizability is calculated using eqn (5), so that no expensive QM/MM approach is needed. Moreover, such an approximation is commonly used in the Surface-SHG community (for example ref. 18 and 19).

Liang *et al.*<sup>57</sup> have reported that eqn (5) using the vacuum value for the second hyperpolarizability  $\gamma^0 = \gamma^{\text{vac}}$  is a poor estimator for  $\beta^{\text{env}}$ . We have thus applied eqn (5) using the average bulk second hyperpolarizability tensor  $\gamma^0 = \langle \gamma^{\text{env}} \rangle$  for all the molecules (values from Table 1 and from ESI,<sup>†</sup> Table S4), and the electrostatic field generated by the environment  $\mathbf{e}^{\text{env}}$ . For liquid water, the distribution of the largest component of  $\mathbf{e}^{\text{env}}$ , along the water molecular axis  $c$  is represented on Fig. 2: it has typical values around  $3 \times 10^{-2}$  a.u., *i.e.*  $1.6 \text{ V } \text{\AA}^{-1}$ . The other



**Fig. 2** Distribution of the electrostatic field generated by the environment  $\mathbf{e}^{\text{env}}$  along the molecular  $c$ -axis. The electrostatic field is measured at 2 positions: at the oxygen atom position (O: orange squares), or at the negative charge one (T: blue dots) according to the TIP4P/2005 force field. The O and T positions are distant by  $0.15 \text{ \AA}$ .

components average to zero, with approximately the same standard deviations (data not shown).

The predicted average values  $\langle \beta^{\text{GR}} \rangle$  are compared to the QM/MM values  $\langle \beta^{\text{env}} \rangle$  in Table 2. First, the addition of the linear correction in eqn (5) allows recovering signs for  $\beta^{\text{GR}}$  components in agreement with  $\beta^{\text{env}}$ . However, the  $\beta^{\text{GR}}$  model largely over-estimates the reference values: the correction  $\gamma^0 \cdot \mathbf{e}^{\text{env}}$  should be smaller. This overshooting is even worse when the vacuum value  $\gamma^{\text{vac}}$  is used for  $\gamma^0$ , as Liang *et al.* had tested.<sup>9</sup> Thus, even if it contains some relevant physics, eqn (5) describing a correction that is proportional to  $\mathbf{e}^{\text{env}}$  is not workable: it predicts an average hyperpolarizability  $\langle \beta^{\text{GR}} \rangle$  one order of magnitude too large.

Moreover, the value of  $\mathbf{e}^{\text{env}}$  in eqn (5) is ill-defined because of its strong spatial heterogeneity. To highlight this fact, Fig. 2 presents the electrostatic field created by the environment along the molecular  $c$ -direction,  $e_c^{\text{env}}$ , calculated at 2 positions within the embedded molecule: either at the position of the negative charge of the TIP4P/2005 model, or at the position of the oxygen atom. While the 2 distributions look very similar, a difference on the average values of about  $10^{-3}$  a.u. appears. Given the large value of  $\gamma_{cccc}$  ( $\approx 1400$  a.u., see Table 1), the component  $\beta_{ccc}$  predicted by  $\beta^{\text{GR}}$  changes by about 40%, depending on where  $\mathbf{e}^{\text{env}}$  is calculated. Therefore, we attribute

**Table 2** Comparison between relevant QM/MM components of the first hyperpolarizability,  $\beta^{\text{env}}$ , the vacuum ones,  $\beta^{\text{vac}}$ , and  $\beta^{\text{GR}}$  predicted by eqn (5) with  $\gamma^0 = \langle \gamma^{\text{env}} \rangle$  from Table 1. All values are in atomic units and are calculated for fundamental wavelength of 800 nm. The values of  $\mathbf{e}^{\text{env}}$  obtained at the center-of-charge position have been used

$ijk$	$\langle \beta^{\text{vac}} \rangle$	$\langle \beta^{\text{env}} \rangle$	$\langle \beta^{\text{GR}} \rangle$
<i>ccc</i>	-15.3	4.1	28.5
<i>caa</i>	-12.5	-2.0	-0.6
<i>aca</i>	-12.4	-2.0	-0.8
<i>cbb</i>	-5.0	2.5	17.7
<i>bcb</i>	-7.4	2.2	16.9



the failures of eqn (5) to the strong intensity and large gradient of the embedding field  $\mathbf{e}^{\text{env}}$ , which are incompatible with a simple linear expansion. Such limitations of perturbation field expansions to model solvation effects had already been discussed in the literature. When studying the nuclear magnetic shielding tensors of the atoms of *N*-methylacetamide solvated by liquid water, Kjaer *et al.*<sup>36</sup> have shown that the field experienced by the nuclei is very inhomogeneous, so that considering one averaged field from all configurations and/or across all nuclei is not a good approximation.

**3.2.2 Separation into short and long-range neighbors.** To go beyond previous works,<sup>57,59</sup> we propose to rewrite the perturbation field expansion so that the perturbing electric field and electric field gradients are weak, and the linear regime is verified. The evolution of the electrostatic field  $\mathbf{e}(R_c)$  generated by the environment around the target molecule of the QM/MM calculations with the environment size  $R_c$  is presented on Fig. 3. The main contribution to this electrostatic field is created in the first and second solvation shells, at distances between 2 and 5 Å from the target molecule. This neighborhood region also creates the largest electrostatic field gradients, see ESI,† Fig. S6.

Given that the closest shells are responsible for both the intensity and the heterogeneity of  $\mathbf{e}^{\text{env}}$ , we divide the neighborhood into two parts, as illustrated on Fig. 4.

The direct area contains the closest neighbors of the target molecule (at the center), up to a distance  $R_c$ : they create a strong and heterogeneous electric field, noted  $\mathbf{e}(R_c)$ . The long-range area, due to neighbors beyond  $R_c$  from the target molecules, creates a less intense and more homogeneous electric field on the target molecule.

The total embedding electrostatic field  $\mathbf{e}^{\text{env}}$  is thus separated into a short-range contribution  $\mathbf{e}(R_c)$ , and a correction,  $\Delta\mathbf{e}(R_c)$  due to the long-range electrostatic interactions:

$$\mathbf{e}^{\text{env}} = \mathbf{e}(R_c) + \Delta\mathbf{e}(R_c). \quad (6)$$

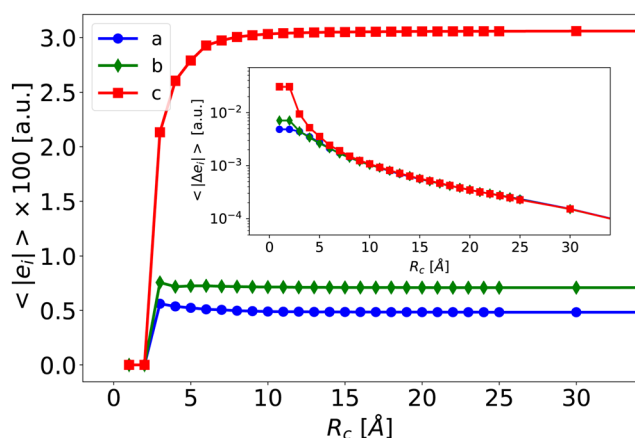


Fig. 3 Averaged norm of the electrostatic field generated by the electrostatic environment ( $\mathbf{e}(R_c)$ ) felt at the T-position, for increasing environment size  $R_c$ . In insert, the difference  $\Delta\mathbf{e}(R_c)$  relative to the value at  $R_c = 40$  Å, in logarithmic scale as a function of the size  $R_c$ .

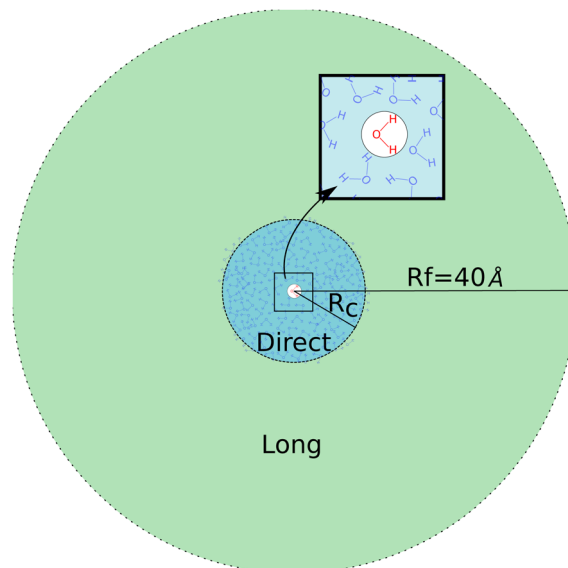


Fig. 4 Sketch of the electrostatic embedding procedure. The QM box is defined for only one molecule of water at the center. Direct area, until  $R_c$ : all the neighbors are included in a QM/MM calculation with discrete solvation procedure, such as the PE formalism. Long-range area, from  $R_c$  to  $R_f$ : the neighbors are included by the electrostatic field they generate at one specific point of the QM box.

where all the fields are evaluated at the position of the negative charge in the TIP4P/2005 model (T-position, see Fig. 2 inset). The evolution of  $\Delta\mathbf{e}(R_c)$  as a function of  $R_c$  is available in the insert of Fig. 3. At  $R_c = 10$  Å, the electrostatic field generated by the long-range neighbors is very small compared to the one created by close neighbors: the average correction  $\Delta\mathbf{e}(R_c = 10 \text{ Å})$  is close to  $10^{-3}$  a.u. The finite field calculations reported in Section 3.1 have shown that for such amplitudes of the perturbation, the dependence of the QM/MM  $\beta^{\text{env}}$  relative to an external, homogeneous field is linear (see ESI,† Fig. S1). Moreover, we have investigated whether the correction  $\Delta\mathbf{e}(R_c)$  is spatially homogeneous by measuring the correction to the electric field gradient  $\Delta\frac{\delta e}{\delta x}(R_c)$  (see ESI,† Fig. S1). Generally, the corrections to the gradients are small (typically  $10^{-4}$  a.u.) for neighbors further away than 10 Å. Since the correction  $\Delta\mathbf{e}(R_c)$  is more homogeneous than  $\mathbf{e}^{\text{env}}$ , it hardly depends on the point on which it is calculated. In the following, we use the value of  $\Delta\mathbf{e}(R_c)$  evaluated at the T-position.

**3.2.3 Long-range correction to  $\beta$  using  $\gamma$ .** The further the neighbor is from the target molecule, the better its contribution to  $\mathbf{e}^{\text{env}}$  can be modeled as a weak and homogeneous electrostatic field. A linear  $\gamma$  correction to  $\beta$  could be tested to describe the long-range electrostatic effects.

For the first hyperpolarizability, as for  $\mathbf{e}^{\text{env}}$ , we describe the impact of the environment in two different ways depending on the neighbors distance. The direct neighbors within the sphere of radius  $R_c$  are included explicitly in the QM/MM calculation through the PE approach. The long-range neighbors, beyond  $R_c$ , are included implicitly *via* the homogeneous field they produce on the target molecule,  $\Delta\mathbf{e}(R_c)$ , with a correction



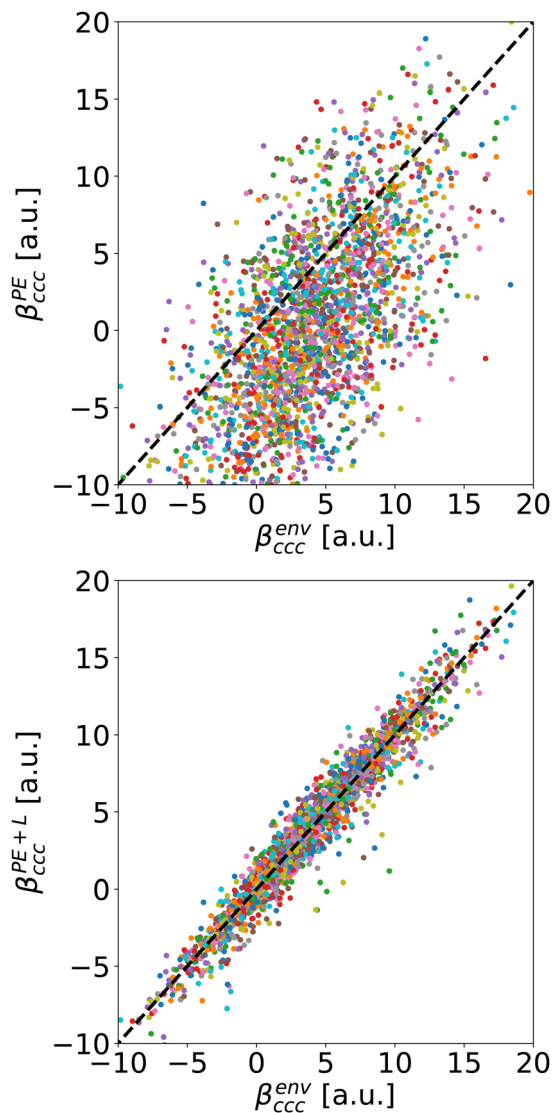


Fig. 5 Comparison of the reference individual molecular  $\beta_{ccc}^{env}$  component and the one obtained using PE (TOP) and PE + L (BOTTOM) approximation with  $R_c = 5 \text{ \AA}$ . The dashed line corresponds to the ideal case where a correction matches perfectly the expected value. Each dot represents a molecule.

proportional to  $\langle \gamma^{env} \rangle$ . This approximation can be viewed as a linear expansion relative to the value of first hyperpolarizability already perturbed by the direct neighbors,  $\beta^{PE}(R_c)$ :

$$\beta^{PE+L}(R_c) = \beta^{PE}(R_c) + \langle \gamma^{env} \rangle \cdot \Delta e(R_c). \quad (7)$$

In the following, we compare this approximation of  $\beta$ , noted  $\beta^{PE+L}(R_c)$  and the usual QM/MM calculation where only the PE scheme up to a distance  $R_c$  is used, noted  $\beta^{PE}(R_c)$ . To estimate the accuracy of the two approximations PE and PE + L as a function of  $R_c$ , we compare them to our reference values  $\beta^{env}$ .

As a typical example, we compare on Fig. 5 the  $\beta_{ccc}$  computed with  $R_c = 5 \text{ \AA}$  either using PE approximation (top) or the PE + L one (bottom) to the reference value  $\beta_{ccc}^{env}$ . The diagonal

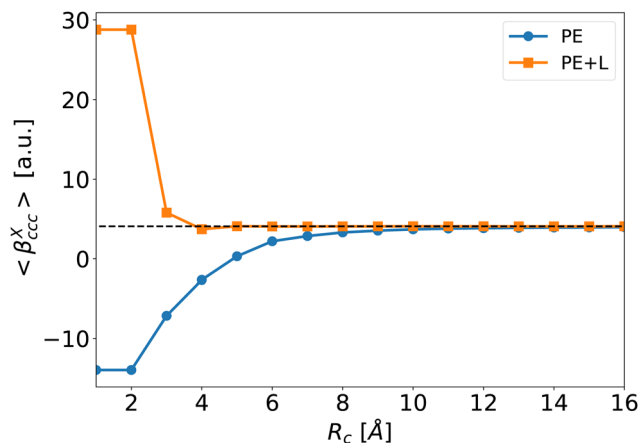


Fig. 6 Average of the  $\beta_{ccc}^X(R_c)$  component as a function of  $R_c$  for the PE (blue disks) and PE + L (orange squares, see eqn (5)) approximations. The dashed line is the reference value  $\beta^{env}$  obtained for  $R_f = 40 \text{ \AA}$ .

dashed lines represent the case where the approximation reproduces exactly the reference.

For the usual calculation ( $\beta^{PE}$ ), there is a systematic error leading to inaccuracy: at  $R_c = 5 \text{ \AA}$ , the average value of  $\beta_{ccc}^{PE}$  is smaller than the one of  $\beta_{ccc}^{env}$ . Moreover, the error fluctuates a lot, it can reach tens of a.u., depending on the molecule. On the contrary, the PE + L approach almost quantitatively reproduces the reference, even for  $R_c$  as small as  $5 \text{ \AA}$ .

To illustrate the effect of  $R_c$ , Fig. 6 depicts the evolution of  $\langle \beta_{PE} \rangle$  and  $\langle \beta_{PE+L} \rangle$  as a function of  $R_c$  for the *ccc* component (similar trends are obtained for other components). As expected, the inaccuracy of the PE and PE + L approaches decreases when  $R_c$  increases.

When the correction is added, the convergence is reached faster ( $R_c = 4 \text{ \AA}$ ) compared to the usual  $\beta^{PE}$  calculation ( $R_c \approx 12 \text{ \AA}$ ). The effect of the neighbors further than  $5 \text{ \AA}$  can thus be very efficiently included by a second hyperpolarizability corrective term. Beyond the description of the average first hyperpolarizability, Fig. 5 also illustrates that the PE + L approach improves the precision of the individual predictions. This can be quantified through the mean absolute error (MAE), averaged over all the  $N = 2400$  molecular configurations and all the 27 components, noted  $\Delta\beta_T$ :

$$\Delta\beta_T^X = \frac{1}{27N} \sum_{ijk} \sum_{n=1}^N \left| \beta_{ijk}^{env(n)} - \beta_{ijk}^{X(n)} \right| \quad (8)$$

where X stands for the approximation PE, or PE + L, and the exponent ( $n$ ) indicates that the values of the  $n$ -th molecule are used. Fig. 7 reports the evolution of  $\Delta\beta_T^X$  as a function of  $R_c$ .

The long-range correction provides very good results and reaches an error below 0.1 a.u. at  $R_c = 12 \text{ \AA}$ . To obtain the same degree of accuracy, one would need to include explicitly neighbors up to  $R_c \approx 30 \text{ \AA}$  in the usual PE scheme without long-range correction. For  $R_c$  larger than  $10 \text{ \AA}$ , the error obtained with the long range correction (PE + L) is about one order of magnitude lower than the one with the traditional PE approach. This shows that – once the strong short-range effects are explicitly



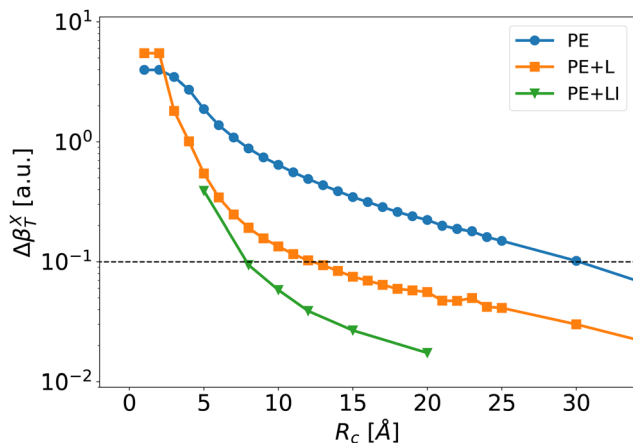


Fig. 7 Mean absolute error of the individual first hyperpolarizability tensor defined by eqn (8) as a function of  $R_c$  for the PE (blue disks), PE + L (orange squares) and PE + LI (green triangles) approximations. The dashed line corresponds to an error of 0.1 a.u. at which we define convergence.

included – it is possible to use the knowledge of  $\gamma$  to predict accurately the long-range electrostatic effects on  $\beta$ . Note that the long-range correction is proportional to the average value  $\langle\gamma^{\text{env}}\rangle$ , *i.e.* Eqn (9) neglects  $\gamma^{\text{env}}$  fluctuations. As discussed further in ESI,<sup>†</sup> Section S4, including the fluctuations of  $\gamma^{\text{env}}$  in eqn (9) only slightly improves the results. However, it implies to compute  $\gamma^{\text{env}}$  for every molecule, which greatly increases the numerical cost. Therefore, using an average  $\gamma^{\text{env}}$  is a very good compromise between computational cost and accuracy, at least for bulk liquid water.

**3.2.4 Implicit long-range correction.** The linear correction is validated here for water molecules in bulk liquid water. We consider this result as a first step towards more complex systems, anisotropic or charged ones, where electrostatic effects become even more important. In these new environments, the second hyperpolarizability of water might differ from the one of the bulk phase, and its fluctuations may become larger. With these next steps in mind, we show here that our explicit long-range correction can also be added to the system implicitly, during the QM calculation, *via* the addition of the homogeneous field  $\mathbf{e}(R_c)$  on the top of the heterogeneous MM embedding field. We note this approach PE + LI for Long-range Implicit correction, than can be noted

$$\beta^{\text{PE+LI}}(R_c) = \beta^{\text{PE}}(R_c; \mathbf{e}(R_c)). \quad (9)$$

Fig. 7 illustrates that the PE + LI values of the first hyperpolarizability are very close to the reference values. The PE + LI approach is even more precise than the PE + L because (1) the second hyperpolarizability fluctuations are neglected in PE + L – not in PE + LI, and (2) the second hyperpolarizability is calculated with finite field approach in PE + L – not in PE + LI. The PE + LI correction does not require the knowledge of the second hyperpolarizability. It can be seen as a simplification of more elaborated QM/MM developments taking into account long-range environment in non-periodic QM/MM methods.<sup>60–63</sup> Indeed, long-range effects can be approximated

by boundary potential models, by cutoff method coupled with switching and shifting functions, by the so-called Long-Range Electrostatic Correction, or by additional charges on the edge of the MM embedding, or in the MM embedding. Our simpler approach with a homogeneous electrostatic field is intended to provide the impact of the remote MM environment on the non-linear optical property of the QM system. It is limited to sequential QM/MM approaches where the trajectory is obtained independently. The advantages of the PE + LI are (1) its insight in the physical origin of the correction and its link with usual equations used in the experimentalist community, and (2) its simplicity and low computational cost. Finally, the value of  $\mathbf{e}(R_c)$  could include periodic boundary conditions, but it is beyond the present work.

## 4 Conclusions

In this work, we have computed the first and second hyperpolarizability of water in the liquid phase using state-of-the-art QM/MM calculations. With this set of values, several approaches were tested to take into account the electrostatic environment in the first hyperpolarizability.

The second hyperpolarizability is high in the bulk phase. The neighborhood induces a dispersion on the second hyperpolarizability  $\gamma(2\omega, \omega, \omega, 0)$  values, but less than for the first hyperpolarizability.

Concerning the first hyperpolarizability  $\beta$ , linear expansions are valid for a weak and spatially-homogeneous electrostatic field, while the embedding electrostatic field created by the liquid surroundings is not. To compute individual molecular hyperpolarizability in the liquid phase, the molecular structure of the nearest neighbors have to be described explicitly, and QM/MM models are very relevant.

But long-range effects, up to several nanometers, can become heavy to take into account using QM/MM approaches, since a very large number of neighbors have to be included. This problem is expected to become all the more pregnant when the environment is anisotropic, with charges or polarization. We have proposed to consider the effect of long-range electrostatic environment on the liquid first hyperpolarizability using the average second hyperpolarizability  $\gamma$ . This long-term correction increases the precision of the hyperpolarizability calculation, and speeds up the convergence of the average value relative to the QM/MM environment size  $R_c$ . For pure bulk water, our approximation makes sense for a long-range part defined beyond 5 to 10 Å, depending on the quantity required. From a numerical point of view, the correction can be powerful for charged environments because the size of the explicit embedding region can be reduced drastically.

This work has demonstrated the promising potentialities of using the second hyperpolarizability, explicitly or implicitly to predict first hyperpolarizabilities. The present study is applied to pure water bulk phase at 300 K. However, the correction developed here should be also relevant for many other system geometries or compositions, where the long-range





electrostatic effects can modify optical responses. If necessary, an extension including electric field gradients would also be straightforward using the Dalton package. Precision tests remain to be performed for more complex systems such as salted aqueous solutions, or charged interfaces.

## Author contributions

GLB: conceptualization, methodology, validation, writing. OB: supervision, validation, writing. EB: supervision, writing, review, editing. CL: supervision, validation, writing.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

We thank P.-F. Brevet for insightful discussions. We gratefully acknowledge support from the PSMN (Pôle Scientifique de Modélisation Numérique) of the ENS de Lyon for the computing resources.

## Notes and references

- P.-F. Brevet, Surface Second Harmonic Generation, 1997.
- M. C. Goh, J. M. Hicks, K. Kemnitz, G. R. Pinto, K. Bhattacharyya, K. B. Eisenthal and T. F. Heinz, *J. Phys. Chem.*, 1988, **92**, 5074–5075.
- Q. Wei, D. Zhou and H. Bian, *Phys. Chem. Chem. Phys.*, 2018, **20**, 11758–11767.
- M. N. Nasir, E. Benichou, C. Loison, I. Russier-Antoine, F. Besson and P.-F. Brevet, *Phys. Chem. Chem. Phys.*, 2013, **15**, 19919–19924.
- G. Licari, J. S. Beckwith, S. Soleimanpour, S. Matile and E. Vauthey, *Phys. Chem. Chem. Phys.*, 2018, **20**, 9328–9336.
- E. Donohue, S. Khorsand, G. Mercado, K. M. Varney, P. T. Wilder, W. Yu, A. D. MacKerell, P. Alexander, Q. N. Van, B. Moree, A. G. Stephen, D. J. Weber, J. Salafsky and F. McCormick, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 17290–17297.
- M. E. Didier, O. B. Tarun, P. Jourdain, P. Magistretti and S. Roke, *Nat. Commun.*, 2018, **9**, 1–7.
- S. Bancelin, C. Aimé, I. Gusachenko, L. Kowalczyk, G. Latour, T. Coradin and M. C. Schanne-Klein, *Nat. Commun.*, 2014, **5**, 1–8.
- C. Liang, G. Tocci, D. M. Wilkins, A. Grisafi, S. Roke and M. Ceriotti, *Phys. Rev. B*, 2017, **96**, 1–6.
- P. Kaatz, E. A. Donley and D. P. Shelton, *J. Chem. Phys.*, 1998, **108**, 849–856.
- J. Yang, E. V. Barnat, S.-k Im and D. B. Go, *J. Phys. D: Appl. Phys.*, 2022, **55**(22), 225203.
- B. Levine and C. Bethea, *J. Chem. Phys.*, 1976, **65**, 2429–2438.
- T. N. Ramos, S. Canuto and B. Champagne, *J. Chem. Inf. Model.*, 2020, **60**, 4817–4826.
- B. F. Levine and C. G. Bethea, *J. Chem. Phys.*, 1975, **63**, 2666–2682.
- E. C. Yan, Y. Liu and K. B. Eisenthal, *J. Phys. Chem. B*, 1998, **102**, 6331–6336.
- T. Joutsuka, T. Hirano, M. Sprik and A. Morita, *Phys. Chem. Chem. Phys.*, 2018, **20**, 3040–3053.
- S. Pezzotti, D. R. Galimberti, Y. R. Shen and M.-P. Gaigeot, *Phys. Chem. Chem. Phys.*, 2018, **20**, 5190–5199.
- L. Dalstein, K.-Y. Chiang and Y.-C. Wen, *J. Phys. Chem. Lett.*, 2019, **10**, 5200–5205.
- A. Marchioro, M. Bischoff, C. Lütgebaucks, D. Biriukov, M. Předota and S. Roke, *J. Phys. Chem. C*, 2019, **123**, 20393–20404.
- R. Hartkamp, A. L. Biance, L. Fu, J. F. Dufrêche, O. Bonhomme and L. Joly, *Curr. Opin. Colloid Interface Sci.*, 2018, **37**, 101–114.
- L. B. Dreier, C. Bernhard, G. Gonella, E. H. Backus and M. Bonn, *J. Phys. Chem. Lett.*, 2018, **9**, 5685–5691.
- E. Ma, P. E. Ohno, J. Kim, Y. Liu, E. H. Lozier, T. F. Miller, H.-F. Wang and F. M. Geiger, *J. Phys. Chem. Lett.*, 2021, **12**, 5649–5659.
- H. Chang, P. E. Ohno, Y. Liu, E. H. Lozier, N. Dalchand and F. M. Geiger, *J. Phys. Chem. B*, 2020, **124**, 641–649.
- Y. Foucaud, B. Siboulet, M. Duvail, A. Jonchere, O. Diat, R. Vuilleumier and J.-F. Dufrêche, *Chem. Sci.*, 2021, **12**, 15134–15142.
- G. Maroulis, *Chem. Phys. Lett.*, 1998, **289**, 403–411.
- J. Kongsted, A. Osted, K. V. Mikkelsen and O. Christiansen, *J. Chem. Phys.*, 2004, **120**, 3787–3798.
- M. Hidalgo Cardenuto and B. Champagne, *Phys. Chem. Chem. Phys.*, 2015, **17**, 23634–23642.
- T. Giovannini, M. Ambrosetti and C. Cappelli, *Theor. Chem. Acc.*, 2018, **137**(6), 74.
- P. Beaujean and B. Champagne, *J. Chem. Phys.*, 2019, **151**, 064303.
- T. N. Ramos, F. Castet and B. Champagne, *J. Phys. Chem. B*, 2021, **125**, 3386–3397.
- L. Jensen and P. T. Van Duijnen, *J. Chem. Phys.*, 2005, **123**(7), 074307.
- G. Le Breton, O. Bonhomme, P.-F. Brevet, E. Benichou and C. Loison, *Phys. Chem. Chem. Phys.*, 2021, **23**, 24932–24941.
- A. Willetts, J. E. Rice, D. M. Burland and D. P. Shelton, *J. Chem. Phys.*, 1992, **97**, 7590–7599.
- O. Bonhomme, L. Sanchez, E. Benichou and P. Brevet, *J. Phys. Chem. B*, 2021, **125**, 10876–10881.
- A. Pardon, O. Bonhomme, C. Gaillard, P.-F. Brevet and E. Benichou, *J. Mol. Liq.*, 2021, **322**, 114976.
- H. Kjaer, S. P. A. Sauer and J. Kongsted, *J. Chem. Phys.*, 2011, **134**, 44514.
- Y. Luo, P. Norman, I. Introduction and H. Agren, *J. Chem. Phys.*, 1998, **109**, 3589–3595.
- K. O. Sylvester-Hvid, K. V. Mikkelsen, P. Norman, D. Jonsson and H. Ågren, *J. Phys. Chem. A*, 2004, **108**, 8961–8965.
- S. Plimpton, *J. Comput. Phys.*, 1995, **117**, 1–19.
- J. L. Abascal and C. Vega, *J. Chem. Phys.*, 2005, **123**, 234505.
- G. J. Martyna, D. J. Tobias and M. L. Klein, *J. Chem. Phys.*, 1994, **101**, 4177–4189.



- 42 R. E. Isele-Holder, W. Mitchell and A. E. Ismail, *J. Chem. Phys.*, 2012, **137**.
- 43 R. E. Isele-Holder, W. Mitchell, J. R. Hammond, A. Kohlmeyer and A. E. Ismail, *J. Chem. Theory Comput.*, 2013, **9**, 5412–5420.
- 44 J. M. H. Olsen and J. Kongsted, *Advances in Quantum Chemistry*, Academic Press, 2011, vol. 61, pp. 107–143.
- 45 E. R. Kjellgren, J. M. Haugaard Olsen and J. Kongsted, *J. Chem. Theory Comput.*, 2018, **14**, 4309–4319.
- 46 K. Aidas, C. Angeli, K. L. Bak, V. Bakken, R. Bast, L. Boman, O. Christiansen, R. Cimraglia, S. Coriani and P. Dahle, *et al.*, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**, 269–284.
- 47 H. C. Georg and S. Canuto, *J. Phys. Chem. B*, 2012, **116**, 11247–11254.
- 48 N. H. List, H. J. A. Jensen and J. Kongsted, *Phys. Chem. Chem. Phys.*, 2016, **18**, 10070–10080.
- 49 M. T. Beerepoot, A. H. Steindal, N. H. List, J. Kongsted and J. M. H. Olsen, *J. Chem. Theory Comput.*, 2016, **12**, 1684–1695.
- 50 C. Steinmann, P. Reinholdt, M. S. Nørby, J. Kongsted and J. M. H. Olsen, *Int. J. Quantum Chem.*, 2019, **119**, e25717.
- 51 A. Marefat Khah, P. Reinholdt, J. M. H. Olsen, J. Kongsted and C. Hattig, *J. Chem. Theory Comput.*, 2020, **16**, 1373–1381.
- 52 T. Yanai, D. P. Tew and N. C. Handy, *Chem. Phys. Lett.*, 2008, **393**, 51–57.
- 53 R. A. Kendall, T. H. Dunning and R. J. Harrison, *J. Chem. Phys.*, 1992, **96**, 6796–6806.
- 54 P. Salek, O. Vahtras, T. Helgaker and H. Ågren, *J. Chem. Phys.*, 2002, **117**, 9630–9645.
- 55 P. Besalú-Sala, S. P. Sitkiewicz, P. Salvador, E. Matito and J. M. Luis, *Phys. Chem. Chem. Phys.*, 2020, **22**, 11871–11880.
- 56 P. Beaujean and B. Champagne, *Theor. Chem. Acc.*, 2018, **137**, 50.
- 57 C. Liang, G. Tocci, D. M. Wilkins, A. Grisafi, S. Roke and M. Ceriotti, *Phys. Rev. B*, 2017, **96**, 1–6.
- 58 P. Beaujean and B. Champagne, *J. Chem. Phys.*, 2019, **151**, 064303.
- 59 A. Osted, J. Kongsted, K. V. Mikkelsen, P. O. Åstrand and O. Christiansen, *J. Chem. Phys.*, 2006, **124**, 124503.
- 60 G. A. Cisneros, M. Karttunen, P. Ren and C. Sagui, *Chem. Rev.*, 2014, **114**, 779–814.
- 61 T. Vasilevskaya and W. Thiel, *J. Chem. Theory Comput.*, 2016, **12**, 3561–3570.
- 62 X. Pan, K. Nam, E. Epifanovsky, A. C. Simmonett, E. Rosta and Y. Shao, *J. Chem. Phys.*, 2021, **154**, 024115.
- 63 J. Nochebuena, S. Naseem-Khan and G. A. Cisneros, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2021, **11**, e1515.

