PAPER

A quantum-mechanics molecular-mechanics scheme for extended systems

IOP Publishing Journal of Physics: Condensed Matter
J. Phys.: Condens. Matter **28** (2016) 335201 (10pp) doi:10.1088/0953-8984/28/33/335201

A quantum-mechanics molecular-mechanics scheme for extended systems

Diego Hunt[1], Veronica M Sanchez[2] and Damián A Scherlis[1]

[1] Departamento de Química Inorgánica, Analítica y Química Física/INQUIMAE, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Pab. II, Buenos Aires (C1428EHA) Argentina
[2] Centro de Simulación Computacional Para Aplicaciones Tecnológicas, Polo Científico Tecnológico, CONICET, Godoy Cruz 2201, Buenos Aires, Argentina

E-mail: damian@qi.fcen.uba.ar

Received 5 April 2016, revised 25 May 2016
Accepted for publication 1 June 2016
Published 28 June 2016

Abstract

We introduce and discuss a hybrid quantum-mechanics molecular-mechanics (QM-MM) approach for Car–Parrinello DFT simulations with pseudopotentials and planewaves basis, designed for the treatment of periodic systems. In this implementation the MM atoms are considered as additional QM ions having fractional charges of either sign, which provides conceptual and computational simplicity by exploiting the machinery already existing in planewave codes to deal with electrostatics in periodic boundary conditions. With this strategy, both the QM and MM regions are contained in the same supercell, which determines the periodicity for the whole system. Thus, while this method is not meant to compete with non-periodic QM-MM schemes able to handle extremely large but finite MM regions, it is shown that for periodic systems of a few hundred atoms, our approach provides substantial savings in computational times by treating classically a fraction of the particles. The performance and accuracy of the method is assessed through the study of energetic, structural, and dynamical aspects of the water dimer and of the aqueous bulk phase. Finally, the QM-MM scheme is applied to the computation of the vibrational spectra of water layers adsorbed at the TiO_2 anatase (1 0 1) solid–liquid interface. This investigation suggests that the inclusion of a second monolayer of H_2O molecules is sufficient to induce on the first adsorbed layer, a vibrational dynamics similar to that taking place in the presence of an aqueous environment. The present QM-MM scheme appears as a very interesting tool to efficiently perform molecular dynamics simulations of complex condensed matter systems, from solutions to nanoconfined fluids to different kind of interfaces.

Keywords: QM-MM, molecular dynamics, DFT, intefaces

(Some figures may appear in colour only in the online journal)

Introduction

In the context of molecular simulations, hybrid quantum mechanics-molecular mechanics (QM-MM) schemes consider the system as a sum of two parts: solute (QM fragment) and solvent (MM fragment) [1–20]. The particles are assigned to one of these two groups according to their role: atoms directly involved in bonds breaking or forming, or in polarization or charge transfer effects, must be considered in the QM region, whereas those atoms not participating in these processes are included within the MM subsystem. These two groups are described at different levels with different Hamiltonians, but they interact with each other, generally self-consistently.

QM-MM schemes have been applied extensively and successfuly along the last couple of decades, to model finite chemical and biological systems. The impact of this methodology has been acknowledged through the Chemistry Nobel Prize of the year 2013, which was awarded to some of its founders for the development of multiscale modelling. One of the major successes of this approach was in the study of chemical

reactions inside the active site of proteins. In this type of simulations the solute described quantum-mechanically comprises the active site, while the rest of the protein plus hydration water molecules are treated classically [1, 2, 17, 21]. Leaving aside the applications in biochemistry, hybrid QM-MM methodologies in a non-periodic setting have also been employed in various other contexts, as for example proton transfer reactions in water clusters [5, 22], or in different kinds of materials which were modeled as finite structures: these works have addressed solid–liquid [9], metal-organic [16], and oxide interfaces [23].

On the other hand, the QM-MM methodology applied in periodic boundary conditions (PBC) for both the QM and MM parts, has been rarely reported in the literature. A few examples imposing periodicity to the MM region only to model pure phases and dilute solutions can be found for semiempirical or first-principles approaches [3, 15, 24]. Laino *et al* [10, 11] developed a QM-MM method in periodic boundary conditions based on Gaussian basis sets and multigrids to treat the long-range interactions, which was tested on the simulation of surface defects present at the α-quartz phase of silica. Other periodic QM-MM implementations with Gaussian basis sets have been proposed based on the reduction of the electron density to point charges, after which the classical Ewald summation can be applied [19, 25]. Such strategy has been implemented for both semiempirical [25] and *ab initio* [19] Hamiltonians. In this line, Golze and co-workers elaborated a method for the treatment of metallic interfaces, where the interactions between the quantum-mechanical adsorbate and the classical substrate are handled at the molecular mechanics level [20]. To the best of our knowledge only Yarne *et al* [6] developed a hybrid QM-MM methodology imposing PBC to the whole system in a pseudopotentials planewaves (PPW) code [26]. In this case electrons were confined to a smaller unit cell inside the supercell needed to describe the whole system, and periodicity was limited to 1 or 2D [6].

In the present article, we present a formulation for hybrid QM-MM calculations based on density functional theory (DFT) in a PPW framework. In particular, this scheme has been devised for the Car–Parrinello method as implemented in the Quantum Espresso code [27]. The goal is to have available a hybrid QM-MM methodology in PBC appropriate to describe condensed matter in complex environments, as solid–liquid or liquid–liquid interfaces, where the main interest or the 'chemistry' involves one of the two phases—the QM part—under the influence of the other—the MM part. This could be useful in solid-water interfaces, where the solid and any adsorbed species can be described quantum-mechanically, while the solution may be modeled using a classical force-field. We illustrate this kind of application through the study of titania in contact with an aqueous phase. The opposite representation, in which the solid constitutes the MM part, might also be appealing if the interest were in the properties of the other phase, as it could be the case of nanoconfined molecules or fluids. To the best of our knowledge, no other QM-MM model has been based on the present strategy, which we believe is a very interesting one for atomistic simulations of interfaces or nano-spaces with high accuracy at an affordable computational cost.

QM-MM method in the pseudopotential plane wave framework

Partitioning of the total energy

In the DFT-PPW approach used in the Quantum Espresso code, the QM energy can be cast as [26, 28]:

$$E_{\mathrm{QM}} = T_e[\rho] + E_{\mathrm{H}}[\rho] + E_{ii} + E_{\mathrm{PS}}^{\mathrm{loc}}[\rho] + E_{\mathrm{PS}}^{\mathrm{nl}} + E_{\mathrm{XC}}[\rho] \tag{1}$$

On the right hand side of the above equation, from left to right, there is the kinetic energy of the electrons, the Hartree energy, the ion–ion repulsion, the local and non-local contributions to the pseudopotential energies, and the exchange-correlation functional. Here $\rho(\mathbf{r})$ is the electron charge density.

In the context of QM-MM models, the Hamiltonian and the energy of the system are written as:

$$\hat{H}_{\mathrm{tot}} = \hat{H}_{\mathrm{QM}} + \hat{H}_{\mathrm{MM}} + \hat{H}_{\mathrm{QM-MM}} \tag{2}$$

$$E_{\mathrm{tot}} = E_{\mathrm{QM}} + E_{\mathrm{MM}} + E_{\mathrm{QM-MM}} \tag{3}$$

where $H_{\mathrm{QM-MM}}$ (and the related energy $E_{\mathrm{QM-MM}}$) is a coupling term describing the interaction between the two regions of the system. In the MM region, atoms are typically treated as point charges of charge Z_I interacting with each other through electrostatics, dispersive-repulsive and harmonic potentials, so that the molecular mechanics energy E_{MM} is the sum of three contributions [29]:

$$E_{\mathrm{MM}} = E_{\mathrm{ele}} + E_{\mathrm{LJ}} + E_{\mathrm{bond}} \tag{4}$$

where E_{ele}, E_{LJ} and E_{bond} denote the electrostatic, the Lennard-Jones, and the bonding energy respectively, the later of which models the intramolecular degrees of freedom. In turn, these terms are normally computed as:

$$E_{\mathrm{ele}} = \frac{1}{2}\sum_{I=1}\sum_{J=1,J\neq I}\frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \tag{5}$$

$$E_{\mathrm{LJ}} = \sum_I \sum_J 4\epsilon_{IJ}\left[\left(\frac{\sigma_{IJ}}{|\mathbf{R}_I - \mathbf{R}_J|}\right)^{12} - \left(\frac{\sigma_{IJ}}{|\mathbf{R}_I - \mathbf{R}_J|}\right)^{6}\right] \tag{6}$$

$$E_{\mathrm{bond}} = \sum_{\mathrm{bonds}}\frac{k_i}{2}(l_i - l_{i0})^2 + \sum_{\mathrm{angles}}\frac{a_i}{2}(\theta_i - \theta_{i0})^2 + \sum_{\mathrm{dihedrals}}\frac{v_n}{2}(1 + \cos(n\omega - \gamma)) \tag{7}$$

In the second of these three equations σ_{IJ} and ϵ_{IJ} are the Lennard-Jones radius and interaction energy between atoms I and J. In the last expression, k_i, a_i, and v_n, represent force constants for the harmonic potentials controlling bond lengths, angles and torsions, respectively. Within the MM region we will consider only water molecules, which internal degress of freedom are described through the O-H distances and H-O-H angles, and therefore the third term in the last equation will not be present.

The $E_{\mathrm{QM-MM}}$ contribution appearing in equation (3) can normally be explicitely written as the sum of an electrostatic and a non-electrostatic term. As it will be shown below, however, in the working formula implemented here the electrostatic

contribution to $E_{\mathrm{QM-MM}}$ can not be written separately, because it is intertwined with the total electrostatic energy.

The electrostatic energy in the PPW framework

In the PPW method, the electrostatic contribution comes from the sum of the second, third, and fourth terms on the right hand side of equation (1):

$$E_{\mathrm{es}}[\rho] = E_{\mathrm{H}}[\rho] + E_{\mathrm{PS}}^{\mathrm{loc}}[\rho] + E_{ii} = \frac{1}{2}\iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' + \sum_{s=1}^{N}\sum_{l=1}^{P_s}\int \rho(\mathbf{r})v_{\mathrm{PS}}^{\mathrm{loc},s}(|\mathbf{r}-\mathbf{R}_I|)\mathrm{d}\mathbf{r} + \frac{1}{2}\sum_{I=1}^{P}\sum_{J=1,J\neq I}^{P}\frac{Z_I Z_J}{|\mathbf{R}_I-\mathbf{R}_J|} \tag{8}$$

where s indicates the atomic species, N is the number of different atomic species, and $v_{\mathrm{PS}}^{\mathrm{loc},s}$ is the local part of the pseudopotential for each species. Z_I is the ionic charge of the nuclei (which amounts to the atomic number minus the valence electrons) and $\mathbf{R}_I$ their positions. P stands for the number of ions and P_s for the number of ions corresponding to the atomic species s. We adopt the same convention used in the computational code, in which the sign of the electronic charge is taken as positive and the ionic charge as negative.

Due to the long-range decay of electrostatic interactions, $E_{\mathrm{H}}[\rho]$, $E_{\mathrm{PS}}^{\mathrm{loc}}[\rho]$ and E_{ii} diverge if they are calculated separately. It turns out to be convenient to introduce a fictitious ionic charge, $\rho_\alpha(\mathbf{r})$, which is added to the Hartree energy term and substructed from the other two. In this way, a total neutral charge density is defined, $\rho_T(\mathbf{r}) = \rho_\alpha(\mathbf{r}) + \rho(\mathbf{r})$, and the electrostatic energy can be rewritten as:

$$\begin{aligned} E_{\mathrm{es}}[\rho] = {} & \frac{1}{2}\iint \frac{\rho_T(\mathbf{r})\rho_T(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \\ & + \int \rho(\mathbf{r})\left(\sum_{s=1}^{N}\sum_{l=1}^{P_s} v_{\mathrm{PS}}^{\mathrm{loc},s}(|\mathbf{r}-\mathbf{R}_I|)\mathrm{d}\mathbf{r} - \int \frac{\rho_\alpha(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\right)\mathrm{d}\mathbf{r} \\ & + \frac{1}{2}\left(\sum_{I=1}^{P}\sum_{J=1,J\neq I}^{P}\frac{Z_I Z_J}{|\mathbf{R}_I-\mathbf{R}_J|} - \iint \frac{\rho_\alpha(\mathbf{r})\rho_\alpha(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'\right) \end{aligned} \tag{9}$$

The Hartree and the local pseudopotential contributions can be expanded in Fourier space, and the ion-repulsions treated with the Ewald method. In particular, if the ionic charge is defined as a sum of Gaussian functions centered on every nuclei,

$$\rho_\alpha(\mathbf{r}) = -\frac{\eta^3}{\pi^{3/2}}\sum_{I=1}^{P} Z_I \mathrm{e}^{-2\eta^2|\mathbf{r}-\mathbf{R}_I|^2}, \tag{10}$$

after some manipulation the electrostatic energy can be expressed as [26]:

$$\begin{aligned} E_{\mathrm{es}}[\rho] = {} & \frac{\Omega}{2}\sum_{\mathbf{G}}\frac{4\pi}{G^2}\tilde{\rho}_T(\mathbf{G})\tilde{\rho}_T(-\mathbf{G}) \\ & + \Omega\sum_{\mathbf{G}}\left[\sum_{s=1}^{N} S_s(\mathbf{G})\tilde{v}_{\mathrm{PS}}^{\mathrm{loc},s}(\mathbf{G}) - \frac{4\pi}{G^2}\tilde{\rho}_\alpha(\mathbf{G})\right]\tilde{\rho}(-\mathbf{G}) \\ & + \frac{1}{2}\sum_{I=1}^{P}\sum_{J\neq 1}^{P} Z_I Z_J\left[\sum_{n=-n_{\max}}^{n_{\max}}\frac{erfc(|\mathbf{R}_I+nL-\mathbf{R}_J|\eta)}{|\mathbf{R}_I+nL-\mathbf{R}_J|}\right] - \frac{\eta}{\sqrt{\pi}}\sum_{I=1}^{P} Z_I^2. \end{aligned} \tag{11}$$

where Ω is the volume of the supercell, $1/\eta$ is a cutoff distance parameter, and $S_s(\mathbf{G})$ is an atomic structure factor for each species s,

$$S_s(\mathbf{G}) = \sum_{I=1}^{P_s}\mathrm{e}^{-\mathrm{i}\mathbf{G}\cdot\mathbf{R}_I^s}. \tag{12}$$

In the last couple of equations, $\tilde{\rho}(\mathbf{G})$, $\tilde{\rho}_\alpha(\mathbf{G})$, $\tilde{\rho}_T(\mathbf{G})$, and $\tilde{v}_{\mathrm{PS}}^{\mathrm{loc},s}(\mathbf{G})$ are the coefficients of the Fourier expansions of the corresponding real space functions, with $\mathbf{G}$ the reciprocal lattice vectors ($\tilde{f}(\mathbf{G}) = 1/\Omega\int_\Omega f(\mathbf{r})\mathrm{e}^{-\mathrm{i}\mathbf{G}\mathbf{r}}\mathrm{d}\mathbf{r}$, $f(\mathbf{r}) = \sum_{\mathbf{G}}\tilde{f}(\mathbf{G})\mathrm{e}^{\mathrm{i}\mathbf{G}\mathbf{r}}$). The complementary error function $erfc(x)$ arises from the point-charges interactions screened by the Gaussian functions, with n an index running over cells in real space. The short-ranged nature of this interaction ensures that the sum converges very fast: typically, only first neighbours need to be considered.

Total energy in the QM-MM implementation

One of the key points in our hybrid approach is to conceive the MM atoms in the same way as the pseudoions of the QM region within the PPW framework. There are basically two differences between MM and QM ions in this case: (i) the MM atoms do not include a non-local pseudopotential term, and (ii) the MM ions can have a partial charge, which can be either negative or positive, according to the charge parameter in the force field. Hence, in our implementation, the electrostatic energy is extended to include the MM atoms:

$$E_{\mathrm{es}}[\rho] = E_{\mathrm{H}}[\rho] + E_{\mathrm{PS}}^{\mathrm{loc}}[\rho] + E_{\mathrm{em}}[\rho] + E_{\mathrm{im}} \tag{13}$$

where $E_{\mathrm{em}}[\rho]$ and E_{im} represent, respectively, the interaction of the electron density with the classical charges, and the Coulomb interaction between all ions, both QM and MM.

$$E_{\mathrm{em}}[\rho] = \sum_{m=1}^{M}\sum_{l=1}^{P_m}\int \rho(\mathbf{r})v_{\mathrm{MM}}^{m}(|\mathbf{r}-\mathbf{R}_I|)\mathrm{d}\mathbf{r} \tag{14}$$

$$E_{\mathrm{im}} = \frac{1}{2}\sum_{I=1}^{T}\sum_{J=1,J\neq I}^{T}\frac{Z_I Z_J}{|\mathbf{R}_I-\mathbf{R}_J|} \tag{15}$$

Here M and P_m are, respectively, the number of classical species and the number of atoms for the m species. The function v_{MM}^m is the pseudopotential associated with the classical species m, to be defined below. $\mathbf{R}_I$ is the position of every atom, irrespective of being quantum or classical, and Z_I is its charge, that will be typically a non-integer number in the MM region. T denotes the total number of atoms in the system ($T = \sum_s^N P_s + \sum_m^M M_m$).

The pseudopotential associated with the classical atoms, v_{MM}^m, has to verify a few properties: has to be a smooth continuous function to be numerically tractable with Fast Fourier Transforms, has to decay as the inverse of the distance r at long ranges, and must avoid the divergence when $r \to 0$. We have adopted the functional form proposed by Laio *et al* [7] :

$$v_{\mathrm{MM}}^m(|\mathbf{r}-\mathbf{R}_I|) = v_{\mathrm{MM}}^m(r) = Z_m\frac{r_{\mathrm{cm}}^4 - r^4}{r_{\mathrm{cm}}^5 - r^5} \tag{16}$$

with m the classical atom species, Z_m its charge, and $r_{\rm cm}$ a cutoff radius appropriate for every species. This function approaches Z_m/r for $r \gg r_{\rm cm}$, and goes smoothly to $Z_m/r_{\rm cm}$ for $r = 0$. Even if the exact value of $v_{\rm MM}$ at short ranges is not critical, it has to be small enough not to become a trap for the electrons. In the case of plane-wave basis, sharp MM potentials of positive species may cause electronic charge localization on the classical atoms: this is called the *spill out* effect. The possibility of electron density flowing to the MM region can be minimized using a classical pseudopotential which varies softly and has a small magnitude at short distances. The function defined in equation (16) satisfies these conditions, providing at the same time an appropriate interaction between MM and QM atoms. The values of $r_{\rm cm}$ have to be parameterized for every classical species, and may also depend on the particular combination of DFT scheme and force-field. In practical terms, however, as mentioned in [7] or discussed below, the dependence of the QM-MM interactions on $r_{\rm cm}$ is relatively weak within a broad range, and therefore the exact values of the cutoff radii, while kept within prudential limits, do not appear to be a crucial feature.

With these modifications, the electrostatic energy amounts to the following final form:

$$\begin{aligned} E_{\rm es}[\rho] = {} & \frac{\Omega}{2}\sum_{\mathbf{G}}\frac{4\pi}{G^2}\tilde{\rho}_T(\mathbf{G})\tilde{\rho}_T(-\mathbf{G}) \\ & + \Omega\sum_{\mathbf{G}}\left[\sum_{s=1}^{N} S_s(\mathbf{G})\tilde{v}_{\rm PS}^{{\rm loc},s}(\mathbf{G}) + \sum_{m=1}^{M} S_m(\mathbf{G})\tilde{v}_{\rm MM}^{m}(\mathbf{G}) - \frac{4\pi}{G^2}\tilde{\rho}_\alpha(\mathbf{G})\right]\tilde{\rho}(-\mathbf{G}) \\ & + \frac{1}{2}\sum_{I=1}^{T}\sum_{J\neq I}^{T} Z_I Z_J\left[\sum_{n=-n_{\rm max}}^{n_{\rm max}}\frac{erfc(|\mathbf{R}_I + nL - \mathbf{R}_J|\eta)}{|\mathbf{R}_I + nL - \mathbf{R}_J|}\right] - \frac{\eta}{\sqrt{\pi}}\sum_{I=1}^{T} Z_I^2. \end{aligned} \tag{17}$$

This expression is identical to equation (11), aside from the term involving the structure factor $S_m(\mathbf{G})$ corresponding to the MM species, and from the fact that the sums in the last two terms run over MM and QM atoms. The charge ρ_α now includes the contribution of the MM ions, and therefore it can take either negative or positive values across space.

There is a non-electrostatic contribution to the QM-MM energy which is analogous to that between atoms in the MM region:

$$E_{\rm LJ,im} = \sum_{s=1}^{N}\sum_{I=1}^{P_s}\sum_{m=1}^{M}\sum_{J=1}^{P_m} 4\epsilon_{sm}\left[\left(\frac{\sigma_{sm}}{|\mathbf{R}_{s,I} - \mathbf{R}_{m,J}|}\right)^{12} - \left(\frac{\sigma_{sm}}{|\mathbf{R}_{s,I} - \mathbf{R}_{m,J}|}\right)^{6}\right] \tag{18}$$

where now ϵ_{sm} and σ_{sm} are the parameters for the Lennard-Jones interaction of a classical atom of species s with a quantum atom m. This energy prevents the MM charges of negative sign from collapsing on the positive QM nuclei. The MM subsystem in the present study involved H_2O molecules, which were described through the SPC flexible water model (SPC/Fw) proposed by Wu, Tepper and Voth [30]. The same set of parameters for σ and ϵ were used in both the MM-MM and the QM-MM non-electrostatic interactions, given respectively in equations (6) and (18).

At this point it must be noticed that most force-fields, including the SPC/Fw potential, do not consider any Coulomb interactions between atoms belonging to the same molecule. In the present formulation, however, the electrostatic energy in equation (17) arises from pairwise interactions of every ion with all the others, and those of intramolecular origin can not be easily individualized and excluded from the rest. A simple way to correct for this overcounting could be to separately compute the Coulomb interactions inside each molecule, and then substract it from the total energy. In this case, such a correction would be:

$$E_{\rm intra} = \sum_{i=1}^{n_{\rm H_2O}}\frac{Z_{\rm O}Z_{\rm H}}{|\mathbf{R}^i_{\rm O} - \mathbf{R}^i_{H1}|} + \frac{Z_{\rm O}Z_{\rm H}}{|\mathbf{R}^i_{\rm O} - \mathbf{R}^i_{\rm H2}|} + \frac{Z_{\rm H}Z_{\rm H}}{|\mathbf{R}^i_{H1} - \mathbf{R}^i_{\rm H2}|} \tag{19}$$

where $n_{\rm H_2O}$ is the number of water molecules in the MM region, and $\mathbf{R}^i_O$, $\mathbf{R}^i_{H1}$ and $\mathbf{R}^i_{\rm H2}$ are the positions of the three atoms belonging to molecule i.

Finally, combining all the contributions together, we compute the total energy as:

$$\begin{aligned} E_{\rm tot}[\rho] = {} & E_{\rm es}[\rho] + T_e[\rho] + E_{\rm XC}[\rho] + E^{\rm nl}_{\rm PS} + E_{\rm LJ,im} + E_{\rm LJ,MM} \\ & + E_{\rm bond} - E_{\rm intra} \end{aligned} \tag{20}$$

where $E_{\rm LJ,MM}$ considers the Lennard-Jones interactions within the MM region, and $E_{\rm bond}$ the intramolecular harmonic contributions between connected MM atoms (equation (7)).

Forces

The atomic forces can be calculated for the QM and for the MM atoms as the derivative of the total energy, equation (20), with respect to the ionic positions. This leads to analytical forces in all cases. For the QM atoms, there is no explicit dependence of $T_e[\rho]$ and $E_{\rm XC}[\rho]$ on R_I, and therefore only three terms survive in the derivative:

$$F^{\rm QM}_I = -\frac{{\rm d}E_{\rm tot}}{{\rm d}\mathbf{R}_I} = -\frac{\partial E_{\rm es}}{\partial \mathbf{R}_I} - \frac{\partial E_{\rm LJ,im}}{\partial \mathbf{R}_I} - \frac{\partial E^{\rm nl}_{\rm PS}}{\partial \mathbf{R}_I} \tag{21}$$

The former of these terms on the right hand side above can be developed as:

$$\begin{aligned} -\frac{\partial E_{\rm es}}{\partial \mathbf{R}_I} = {} & \frac{Z_I}{2}\sum_{J\neq I}^{T}\sum_{n=-n_{\rm max}}^{n_{\rm max}}(\mathbf{R}_I + nL - \mathbf{R}_J) \\ & \times\left[\frac{erfc(|\mathbf{R}_I + nL - \mathbf{R}_J|\eta)}{|\mathbf{R}_I + nL - \mathbf{R}_J|^3} + \frac{\eta e^{-\eta^2|\mathbf{R}_I + nL - \mathbf{R}_J|^2}}{|\mathbf{R}_I + nL - \mathbf{R}_J|}\right] \\ & + \Omega\sum_{G\neq 0} i\mathbf{G}e^{i\mathbf{G}\cdot\mathbf{R}_I}\left(\tilde{v}_{\rm PS}^{{\rm loc},s}(\mathbf{G}) + \tilde{v}^m_{\rm MM}(\mathbf{G}) + \frac{4\pi Z_s}{G^2\Omega}\right)\tilde{\rho}(-\mathbf{G}) \end{aligned} \tag{22}$$

On the other hand, the Lennard-Jones contribution to the force is simply:

$$-\frac{\partial E_{\rm LJ,im}}{\partial \mathbf{R}_I} = \sum_{m=1}^{M}\sum_{J=1}^{P_m} 4\epsilon_{sm}\left[\frac{12\sigma^{12}_{sm}}{|\mathbf{R}_{s,I} - \mathbf{R}_{m,J}|^{13}} - \frac{6\sigma^6_{sm}}{|\mathbf{R}_{s,I} - \mathbf{R}_{m,J}|^7}\right] \tag{23}$$

The contribution originating in the non local part of the pseudopotential energy is part of the standard QM implementation and will not be discussed in the present context.

The expression for the forces on the atoms belonging to the MM subsystem will have two terms in common with

equation (21), plus the pure MM contributions, whose derivatives are straightforward:

$$F_J^{\mathrm{MM}} = -\frac{\partial E_{\mathrm{es}}}{\partial \mathbf{R}_J} - \frac{\partial E_{\mathrm{LJ,im}}}{\partial \mathbf{R}_J} - \frac{\partial E_{\mathrm{LJ,MM}}}{\partial \mathbf{R}_J} - \frac{\partial E_{\mathrm{bond}}}{\partial \mathbf{R}_J} + \frac{\partial E_{\mathrm{intra}}}{\partial \mathbf{R}_J} \tag{24}$$

Assessment of the model: water dimer and aqueous liquid phase

The water dimer

The potential energy curve of a water dimer, with one molecule in the QM domain and the other in the MM region, was calculated as a first test. The geometry of the dimer was optimized for a series of oxygen–oxygen separations. It is important to note that there are two inequivalent configurations in which this curve can be obtained, depending on whether the MM molecule plays the role of donor or acceptor of the hydrogen bond. In the water dimer depicted in the inset of figure 1, the hydrogen bond donor is the molecule on the left and the acceptor is on the right. Therefore, two curves can be obtained.

Calculations were performed in a supercell of dimensions $19 \times 9.5 \times 9.5$ Å^3, to minimize the interactions of the water dimer with their periodic images. The cell dimension is longer along the x axis because this is the direction in which the potential energy is scanned. The PBE approach to the DFT exchange-correlation energy [31] in combination with ultrasoft pseudopotentials [32] were adopted to compute total energies and forces. The Kohn–Sham orbitals and charge density were expanded in planewaves up to a kinetic energy cutoff of 50 and 200 Ry, respectively. An electronic mass of 400 a.u. was used to propagate the wavefunctions according to the Car–Parrinello scheme. The r_{cm} values adopted for the QM-MM interactions were 0.21 Å and 0.27 Å for the hydrogen and oxygen atoms, respectively. These values provide the best agreement with respect to the classical and quantum-mechanical interaction energies and radial distribution functions (see below).

Before discussing the QM-MM results, we will examine the curves provided by the pure QM (DFT) and MM (SPC/Fw) methodologies. These are plotted in figure 1. The difference in the description is readily apparent: the QM curve is shifted to a weaker interaction and to a longer distance at the minimum, with respect to the MM curve. Reported experimental values, of 2.98 Å for the oxygen–oxygen separation [33, 34], and ranging from 3.6 to 5.2 kcal mol^{-1} for the magnitude of the interaction at the minimum energy geometry [35–38], are better reproduced by the quantum-mechanical model, which provides an optimized hydrogen-bond energy of 4.9 kcal mol^{-1} at 2.9 Å. Similar results have been obtained from previous DFT calculations based on both localized and extended basis functions [39–43]. On the other hand, the classical-mechanics water dimer interaction, of nearly 7.4 kcal mol^{-1}, is clearly above the available experimental data. The reason for this overestimation is that the SPC and SPC/Fw potentials, as most water force-fields, are parameterized to

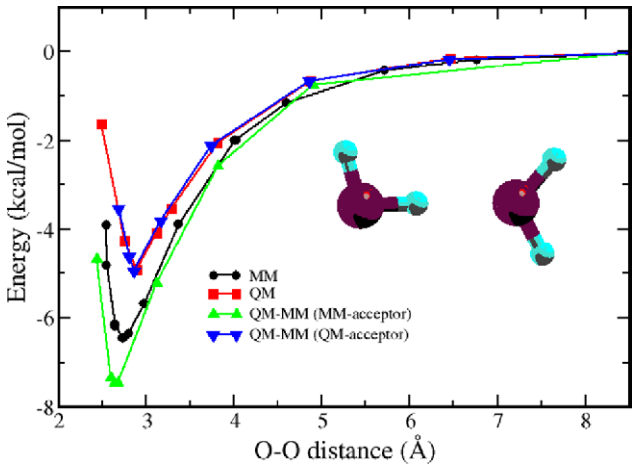

Figure 1. Interaction energy for a water dimer as a function of the O-O distance, according to DFT (QM), SPC/Fw (MM), and hybrid QM-MM calculations. In the dimer depicted in the inset, the molecule on the right plays the role of H-bond acceptor.

reflect the properties of the bulk liquid phase [30, 44], where the molecular dipole moments are significantly enhanced with respect to the isolated molecule (see next section). As a matter of fact, the minimum of the MM curve, at nearly 2.7 Å, is coincident with the first peak of the oxygen–oxygen radial distribution function (RDF) for liquid water at room temperature.

The QM-MM potential energy curves, presented in figure 1 together with the results corresponding to the pure SPC/Fw and DFT calculations, turn out to be quite interesting. As mentioned above, two configurations can be considered in this case, depending on the identity of the donor of the H-bond. Examination of the curves leads to the following observation: the QM-MM curve in which the acceptor is the MM molecule, roughly reproduces the SPC/Fw curve, whereas the QM-MM curve where this role is played by the QM molecule, is very close to the DFT results. In other words, the QM-MM curves are essentially reflecting the identity of the acceptor. This is a meaningful result, understandable when we recall that most of the charge density involved in the bond, and therefore the polarization effect, corresponds to the oxygen atom. The electron density associated with the H atom is much lower and localized, and there is not a major effect if this electron density is replaced by a bare pseudopotential.

Ideally, both QM-MM curves should be identical. In practice, however, a discrepancy is immanent to all QM-MM models, since these curves are tied to the QM and MM Hamiltonians, which necessarily provide different descriptions of the bond. For the calculations in the condensed phase, we expect that the difference between the two kinds of interactions (involving the QM water molecule as the donor or as the acceptor) will be substantially attenuated. The results in the next section, concerning the properties in the bulk, suggest that this is certainly the case.

Figure 2 presents the interaction energies at the minima of the curve, as a function of the hydrogen and oxygen r_{cm} parameters. A decrease in $r_{\mathrm{cm}}^{\mathrm{H}}$ strengthens the interaction when the MM molecule is the hydrogen-bond donor, as a consequence of an enhanced Coulombic attraction. A decrease in

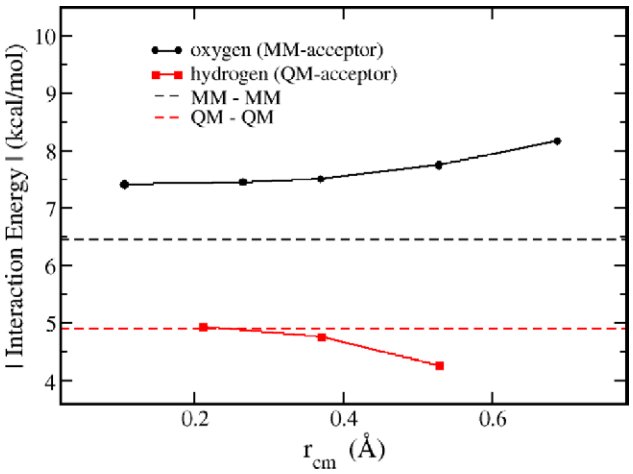

Figure 2. Interaction energy (absolute value) at the minima of the QM-MM curves, as a function of the r_{cm} parameters. The red curve shows the dependence on r^{H}_{cm} when the MM molecule is the hydrogen-bond donor, while the black curve depicts the effect of r^{O}_{cm} when the MM molecule is the acceptor. Dashed lines indicate the interaction energies in the full QM and MM models.

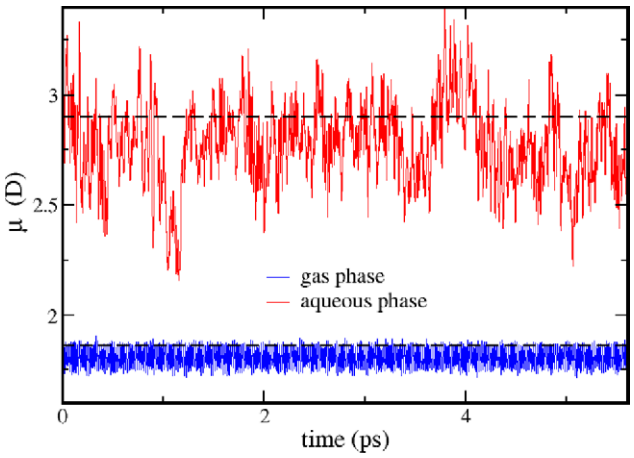

Figure 3. Time evolution of the dipole moment of a quantum-mechanical water molecule in the gas phase, and in an aqueous environment consisting of 63 classical molecules in a cubic cell. The dashed horizontal lines show the experimental values, from [46] and [49].

r^{O}_{cm} when the MM molecule is the acceptor, instead, has the opposite effect, as could be expected from the negative charge of the MM oxygen. The chosen cutoff radii minimize the difference between the interaction energies of the two possible configurations (QM-acceptor and QM-donor), maximizing the agreement between the QM-MM potential energy surfaces with either the full QM or full MM curves, according to the identity of the acceptor. The effect of the r_{cm} values on the interaction energies is minor over a broad range. In [7] larger values were adopted, but we have found that these deteriorate the quality of the radial distribution functions in the present scheme.

Bulk phase properties

Car–Parrinello molecular dynamics simulations were performed on a system of 64 water molecules, of which one was described quantum-mechanically and the rest classically. The simulations were conducted at 300 K using the Nosé–Hoover thermostat in a cubic box in periodic boundary conditions, with a density corresponding to 1 g cm^{-3}, and a time-length of 6 ps. A planewaves basis of 25 Ry for the Kohn–Sham states and 200 Ry for the charge density were employed, together with the PW91 exchange-correlation functional [45].

The average dipole moment for a water molecule in the gas phase obtained from the Car–Parrinello dynamics is 1.81 D, very close to the experimental value of 1.86 D [46, 47]. In the system of 64 H_2O molecules representing the bulk phase, the polarization exerted by the classical environment raises the dipole moment of the QM molecule to an average of 2.88 D. This number is in full agreement with reported estimates of 2.95 D from *ab initio* simulations [48], or 2.9 ± 0.6 D from the x-ray structure factor [49]. Figure 3 depicts the computed dipole moments of water as a function of time, both isolated and in the classical aqueous environment. The dashed lines represent the experimental values. Such a good accord demonstrates that the polarization effect is finely accomplished by the QM-MM scheme.

Figure 4 presents the vibrational frequencies of the water molecule in the gas and in the liquid phases, computed from the Fourier transform of the time correlation function of the atomic velocities $\mathbf{v}(t)$ [50],

$$I(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}t\ \mathrm{e}^{-\mathrm{i}\omega t}\frac{1}{N}\sum_{i=1}^{N}\langle \mathbf{v}_i(0)\mathbf{v}_i(t)\rangle \qquad (25)$$

(all spectral lineshapes presented here have been subject to a Gaussian smoothing). Depending on the system, on the vibrational mode, and on the electronic mass, ionic frequencies in Car–Parrinello dynamics may show redshifts of a few percent with respect to spectroscopic data [51]. In the present case, the positions of the peaks fall between 100 to 200 cm^{-1} below the experimental frequencies, consistently with previous Car–Parrinello simulations of H_2O [52, 53]. For the isolated molecule, it is possible to recognize two bands in figure 4, corresponding to the stretching and bending modes, centered at 3500 and 1480 cm^{-1} respectively. In the liquid state these bands become broader and noisier, with an additional set of peaks below 1000 cm^{-1} arising from librations. The first thing to note is that the stretching frequencies shift to lower wavenumbers in the liquid, whereas the bending experiences the opposite trend. This is the same behavior as observed from IR spectroscopy, where the stretching in the liquid is redshifted in about 300 cm^{-1}, and the bending mode is blueshifted in nearly 50 cm^{-1} [54]. Our simulations in the liquid give peaks which are spread and too much splitted to establish unambiguously the magnitudes of these shifts; however, considering the center of mass of the bands, it turns out that the shift of 300 cm^{-1} in the stretching frequency is pretty much reproduced by the QM-MM model, while the change in the bending frequency appears overestimated by a factor of two. These predictions for the spectral shifts in water are, from a quantitative point of view, of a quality comparable to that obtained from full quantum-mechanical Car–Parrinello simulations [52, 53].

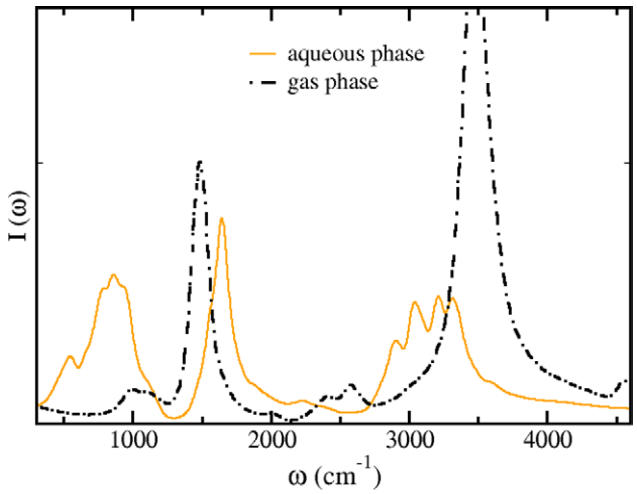

Figure 4. Simulated vibrational spectra of water in the gas and in the liquid phases. In the later case, the aqueous environment is represented by 63 classical molecules in a cubic cell. The computed frequency shifts are in qualitative agreement with IR spectroscopic data.

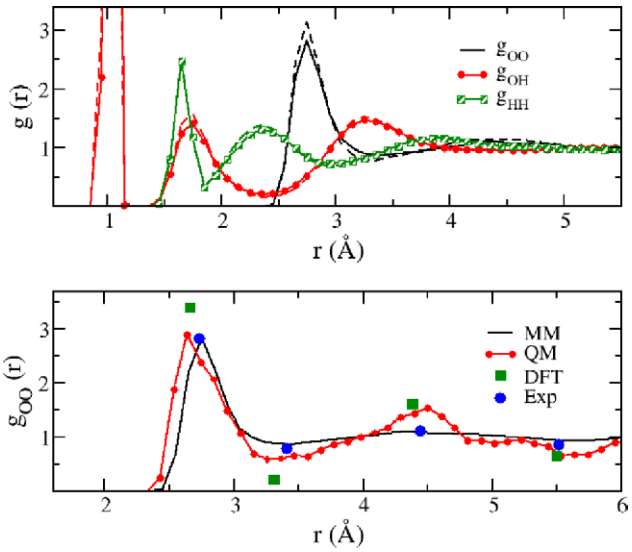

Figure 5. Upper panel: RDFs computed for all pairs of MM atoms in the QM-MM system of 64 water molecules. The dashed lines correspond to the curves obtained from classical molecular dynamics simulations with the SPC/Fw potential. Lower panel: RDFs for MM oxygen atoms (black) and for the QM oxygen atom (red), from the simulations of the 64 molecules system. The amount of data-points to construct the QM RDF is only 1/63 of that involved in the MM RDF, which explains the uneven, unconverged structure of the former. The symbols show the positions of the maxima and the minima in the RDF obtained from quantum-mechanical Car–Parrinello simulations (squares) and from x-ray diffraction experiments (circles), extracted from [43] and [58] respectively.

Radial distribution functions for the model of 64 water molecules are displayed in figure 5. Classical molecular dynamics simulations with the SPC/Fw potential in PBC were performed in the same system with the LAMMPS code [55]. The upper panel confirms that the electrostatic description of the MM atoms within the QM-MM approach, including the correction to the intramolecular Coulomb forces, reproduces the dynamics dictated by the SPC/Fw force-field. The subtle discrepancies between the classical and the QM-MM curves are attributable to differences in the length of the simulations, to the distinctive numerical implementation of the Ewald sums in every code, and possibly to the presence of the QM water molecule. The radial distribution function corresponding to the quantum-mechanical oxygen atom is presented in the lower panel of figure 5. In comparison to SPC/Fw, this curve is more structured, with maxima and minima appearing respectively above and below. Its shape does not seem to be converged, probably because insufficient sampling: note that in this case the RDF is built from a single water molecule out of 64. Various studies of liquid water using the Car–Parrinello method with GGA functionals and an electronic mass comparable to the one employed here, have found overstructured oxygen–oxygen RDFs at room temperature and pressure [43, 53, 56, 57], suggesting that in these conditions this approach represents bulk water in a glassy or supercooled state. Data from one of these works is depicted in figure 5. The radial distribution function of the QM oxygen atom in our QM-MM simulation turns out to be intermediate between those obtained from pure classical and pure quantum-mechanical Car–Parrinello molecular dynamics.

Considerations on computational efficiency

At variance with other QM-MM schemes in the PPW framework, in which the classical region does not have to be included in the simulation box, in the present treatment the MM atoms need to be contained in the supercell together with the QM atoms. Then, the amount of planewaves and the size of real space grids are the same or about the same as in a quantum-mechanical calculation with an equal number of total atoms, and therefore the QM-MM implementation does not involve any significant decrease in memory requirements. In spite of this, the reduction of the QM region cuts the quantity of Kohn–Sham states to be evolved in the Car–Parrinello dynamics, which may have a substantial impact on the overall computing time. As a matter of fact, the computational effort in the quantum-mechanical Car–Parrinello scheme for a given unit cell size is approximately proportional to the number of atoms, with a slope larger than 1, which tends to increase with system dimensions. As a consequence, the speedup achieved by the QM-MM approach is roughly linear with the replacement of QM by MM atoms. This behavior is reflected in figure 6, which illustrates the relative decrease in computation time with the increase of the number of atoms represented classically, for the cases of 32 and 64 water molecules. As expected, the gain in performance becomes more significant as the system grows bigger.

In *ab initio* modelling of solid interfaces in contact with a bulk liquid, the solvent fills a major fraction of the supercell, often representing between 1/2 to 2/3 of the total atoms. In the solid–liquid interface model for anatase (1 0 1) discussed in the next section, for instance, the number of classical atoms is 144, out of a total of 264. In this situation, the QM-MM calculation turns out to be 5 times faster than a full QM simulation. In the case of a quantum-mechanical water molecule

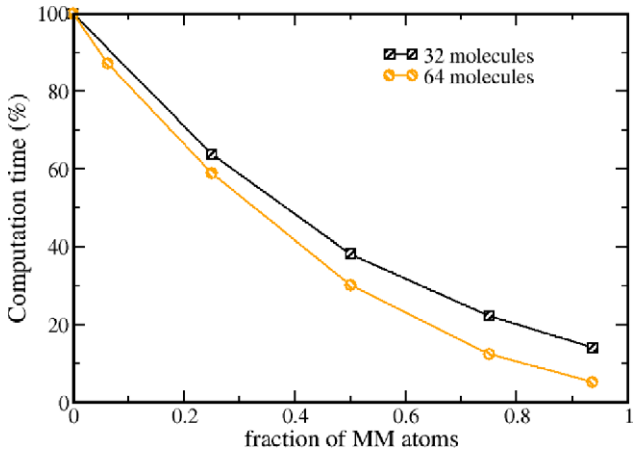

Figure 6. Dependence of the total computational time with the size of the classical domain, keeping constant the total (QM + MM) number of atoms. The benchmarks correspond to 100 Car–Parrinello steps parallelized on four processors for systems of 64 and 32 water molecules in a cubic unit cell. Tests on different number of processors, up to 16, provide identical trends.

surrounded by 63 classical ones, examined in the previous section, the acceleration goes above one order of magnitude.

Water vibrational frequencies at the solid–liquid interface of TiO_2 anatase

Many of the most relevant applications of TiO_2 implicate the solid–liquid interface, which entails a serious challenge to first-principles modelling, given the need for extensive simulations to achieve an appropriate configurational sampling of the fluid phase. Thus, *ab initio* molecular dynamics investigations considering the water liquid phase in contact with titania have been carried out only in a limited number of occasions, to address the rutile (1 1 0) [59–62] and the anatase (1 0 1) and (0 0 1) surfaces [63]. Instead, there is a large number of DFT studies which have examined stoichiometric and defective titania interfaces in the presence of just a few water monolayers, typically ranging from one to three [64–72]. In this section we illustrate the applicability of our hybrid quantum-mechanics molecular-mechanics scheme through the calculation of the water vibrational frequencies at the anatase (1 0 1) interface. We aim at determining to what extent the explicit inclusion of the liquid phase affects the dynamical properties of the first adsorbed layers. To this end, results from QM-MM molecular dynamics simulations representing the liquid phase, are compared with those coming from QM simulations incorporating just one or two H_2O monolayers.

The DFT parameters concerning planewave basis, pseudopotentials, and exchange correlation functional, were the same as employed to describe the water bulk phase in the previous section. The anatase (1 0 1) surface was represented using a 2×2 supercell, containing six layers of TiO_2 units, and Γ-point sampling. Cell dimensions were $7.56 \times 10.24 \times 22.67 \text{Å}^3$ for the QM calculations. In the QM-MM simulations the z-parameter was extended to 30.23 Å to accommodate the aqueous phase, consisting of 16 QM plus 48 MM water molecules. In particular, the first two water monolayers adjacent

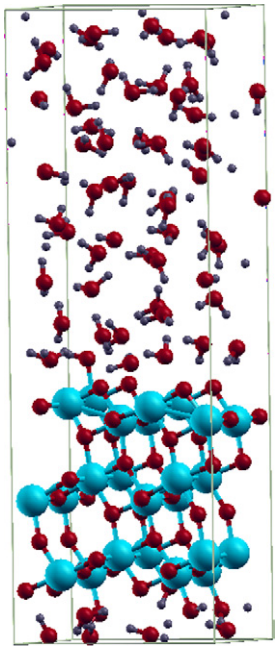

Figure 7. Slab model used in the QM-MM simulations of the TiO_2 anatase (1 0 1) surface in contact with a bulk water phase. The water molecules were represented classically, except those forming the first and second adsorbed layers.

to the solid surfaces were modelled quantum-mechanically, to get the corresponding vibrational frequencies and to avoid a direct interaction between SPC/Fw water and titania. The probability of molecule exchanges between water layers is very low, in particular in the first adsorbed monolayer, where residence times have been estimated to be in the order of several nanoseconds through molecular dynamics simulations [73, 74]. We have not observed any such exchanges in our simulations (in any case, if they were detected in a particular model interface, this could be resolved by increasing the size of the QM domain). Molecular dynamics simulations were performed at 300 K with the Nosé–Hoover thermostat, with sampling windows of 6 ps. Figure 7 displays the model structure employed in the QM-MM calculations of the solid–liquid interface.

Figure 8 shows the vibrational density of states corresponding to the water molecules directly adsorbed on the TiO_2 surface, computed through equation (25). The three panels compare three different coverages: monolayer (top), bilayer (center), and the liquid environment (bottom). For the single layer, the stretching mode, appearing at around $3500\,\text{cm}^{-1}$, is the strongest one, vaguely resembling the vibrational patterns of water in the gas phase. With the incorporation of a second layer, the intensities arising from librations and bending become larger than that associated with the stretching, which in turn moves slightly to the left. This trend is similarly observed in the presence of the bulk liquid, where the librations dominate the spectra. Hence, in general terms it is observed that the vibrational behavior of the H_2O molecules in the first layer, roughly shifts from gas-like to liquid-like as the degree of hydration is increased. The main features in our computed spectra are in line with those obtained from neutron scattering experiments on water confined in anatase

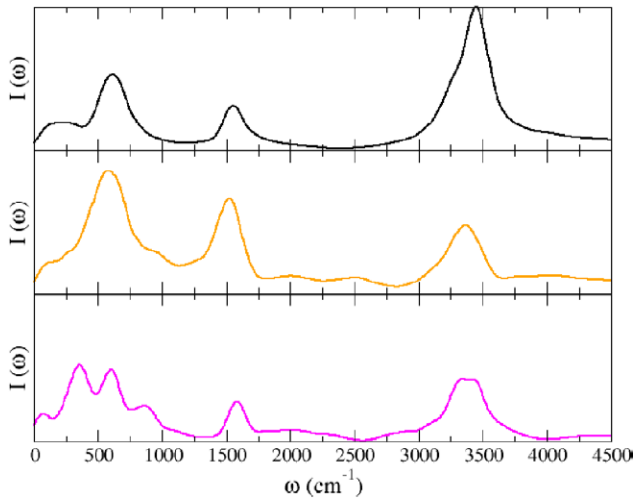

Figure 8. Simulated vibrational spectra corresponding to the first water monolayer adsorbed on the anatase (1 0 1) surface, at different coverages: monolayer (top panel), bilayer (middle panel), and bulk liquid (bottom panel).

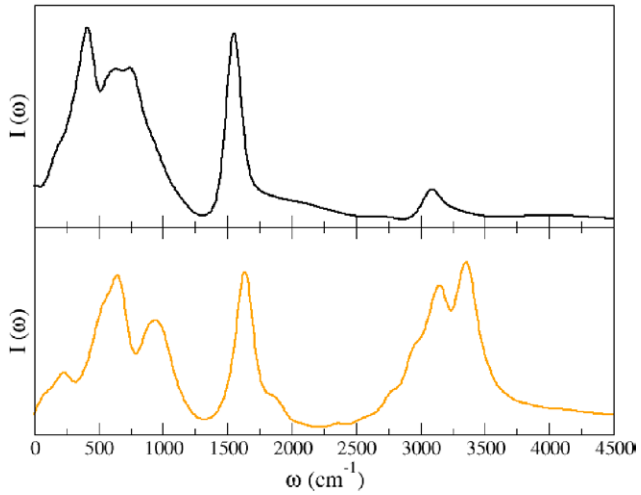

Figure 9. Simulated vibrational spectra corresponding to the second water monolayer adsorbed on the anatase (1 0 1) surface, for a bilayer (top panel) and the bulk liquid phase (bottom panel).

nanoparticles, where librations and bending are predominant, and the stretching absorption band is shifted to lower frequencies with respect to the gas phase [75].

The aqueous media has a major influence on the spectra of the second water layer. Figure 9 shows that in the absence of the liquid environment, the stretching of the water molecules in the second layer is very weak in comparison with librations and bending. The typical bulk-water features are recovered as the MM environment is included in the simulation.

The above results suggest that the internal degrees of freedom of a water molecule in a single layer on anatase (1 0 1) retain some of the character they have in the gas phase. Moreover, the sole inclusion of a second H_2O layer appears to be enough to recreate, at least to some extent, the dynamics in the presence of a bulk aqueous phase. This behavior can be tracked to the hydrogen bond network arising in each case. For a single adsorbed layer, the H_2O molecules are tightly bound to the five-coordinate Ti atoms, with one or both of their hydrogen atoms forming relatively weak hydrogen-bonds with the bridging oxygen sites on the surface. On the other hand, an inspection of the trajectories in the presence of a second row of solvent molecules reveals that, most of the time, the molecules in the first layer are involved in hydrogen-bonds with at least another H_2O molecule, either as a donor or as an acceptor. In the liquid phase the number of hydrogen bonds that can be formed, and the polarization effects, are even larger. Yet, the incorporation of a few water molecules beyond the first monolayer is sufficient to induce on the adsorbate, a dynamical behavior very close to the one corresponding to full hydration.

Summary

We have presented an approach to perform hybrid quantum-mechanics molecular-mechanics simulations with the Car–Parrinello method in the context of the pseudopotential-planewaves setting. At variance with other QM-MM implementations existing in planewaves codes, in the present approach the classical atoms are treated on the same footing as the quantum-mechanical ions, which naturally leads to periodic boundary conditions for the totality of the system. Thus, all QM and MM atoms need to be contained within the same real space grid determined by the simulation cell. As a consequence, the size of the MM region has an impact on the computational cost, and this method would not be convenient for extended systems where this region is extremely large, or exceeds by far the size of the QM part. It turns out that, for a given unit cell, the scalability is approximately linear with the substitution of QM by MM atoms. In typical calculations of solid–liquid interface models involving a few hundred atoms, in which the solvent represents more than one half of the system, speedup factors above five can be attained with the present scheme. We applied our implementation to the computation of the vibrational spectra of water adsorbed at the TiO_2 anatase (1 0 1) surface, at various coverages. It was found that the presence of a second monolayer of water molecules is enough to mimick the effect of an aqueous environment on the vibrational frequencies of the first adsorbed layer. This methodology seems particularly suited for molecular dynamics simulations in condensed matter systems including one or more fluid phases. Solutions, nanoconfined fluids, or solid–liquid and liquid–liquid interfaces, are all examples where this scheme could be extremely valuable.

Acknowledgments

We express our gratitude to Ivan Girotto and Paolo Gianozzi for precious help related to the structure and parallelization of the Quantum-Espresso code. We also thank Davide Ceresoli for useful discussions. This study has been partially supported by grants of the Agencia Nacional de Promocion Cientifica y Tecnologica de Argentina, PICT 2012-2292, and UBACYT 20020120100333BA. We acknowledge CSC-CONICET for granting the use of the TUPAC HPC cluster, which allowed to perform most of the computations included in this work.

References

[1] Monard G and Merz K M Jr 1999 *Acc. Chem. Res.* **32** 904
[2] Orozco M and Luque F J 2000 *Chem. Rev.* **100** 4187
[3] Gao J and Alhambra C 1997 *J. Chem. Phys.* **107** 1212
[4] Eichinger M, Tavan P, Hutter J and Parrinello M 1999 *J. Phys. Chem.* **110** 10452
[5] Elola M D, Estrin D A and Laria D 1999 *J. Phys. Chem.* A **103** 5105
[6] Yarne D A, Tuckerman M E and Martyna G J 2001 *J. Chem. Phys.* **115** 3531
[7] Laio A, VandeVondele J and Rothlisberger U 2002 *J. Chem. Phys.* **116** 6941
[8] Crespo A, Scherlis D A, Marti M A, Ordejón P, Roitberg A E and Estrin D A 2003 *J. Phys. Chem.* B **107** 13728
[9] Du M-H, Kolchin A and Cheng H P 2003 *J. Chem. Phys.* **119** 6418
[10] Laino T, Mohamed F, Laio A and Parrinello M 2005 *J. Chem. Theory Comput.* **1** 1176
[11] Laino T, Mohamed F, Laio A and Parrinello M 2006 *J. Chem. Theory Comput.* **2** 1370
[12] Sánchez V M, Crespo A, Gutkind J S and Turjanski A G 2006 *J. Phys. Chem.* B **110** 18052
[13] Maurer P, Laio A, Hugosson H, Colombo M and Rothlisberger U 2007 *J. Chem. Theory Comput.* **3** 628
[14] Rinaldo D, Philipp D M, Lippard S J and Friesner R A 2007 *J. Am. Chem. Soc.* **129** 3135–47
[15] Lebrero M C G and Estrin D A 2007 *J. Chem. Theory Comput.* **3** 1405
[16] Sushko M L, Sushko P V, Abarenkov I V and Shluger A L 2010 *J. Comput. Chem.* **31** 2955
[17] Bongards C and Gärtner W 2010 *Acc. Chem. Res.* **43** 485–95
[18] Kamerlin S C L, Haranczyk M and Warshel A 2009 *J. Phys. Chem.* B **113** 1253
[19] Holden Z C, Richard R M and Herbert J M 2013 *J. Phys. Chem.* **139** 244108
[20] Golze D, Iannuzzi M, Nguyen M T, Passerone D and Hutter J 2013 *J. Chem. Theory Comput.* **9** 5086
[21] Capece L, Estrin D A and Marti M A 2008 *Biochemistry* **47** 9416
[22] Lin H and Truhlar D G 2005 *J. Phys. Chem.* A **109** 3991–4004
[23] Ellis D and Warschkow O 2003 *Coord. Chem. Rev.* **238–9** 31
[24] Ishiyama T, Takahashi H and Morita A 2012 *J. Phys.: Condens. Matter* **24** 124107
[25] Nam K, Gao J and York D M 2005 *J. Chem. Theory Comput.* **1** 2
[26] Kohanoff J 2006 *Electronic Structure Calculations for Solids and Molecules—Theory and Computational Methods* (Cambridge: Cambridge University Press)
[27] Giannozzi P *et al* 2009 *J. Phys.: Condens. Matter* **21** 395502
[28] Marx D and Hutter J 2000 *Modern Methods and Algorithms of Quantum Chemistry* ed J Grotendorst (Jülich: John Von Neumann Insitute for Computing)
[29] Leach A 2001 *Molecular Modelling: Principles and Applications* 2nd edn (Harlow: Prentice Hall)
[30] Wu Y, Tepper H L and Voth G A 2006 *J. Phys. Chem.* **124** 024503
[31] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865
[32] Vanderbilt D 1990 *Phys. Rev.* B **41** 7892
[33] Dyke T R, Mack K M and Muenter J S 1977 *J. Chem. Phys.* **66** 498
[34] Odutola J A and Dyke T R 1980 *J. Chem. Phys.* **72** 5062
[35] Curtiss L A, Frurip D J and Blander M 1979 *J. Phys. Chem.* **71** 2703
[36] Gebbie H A, Burroughs W J, Chamberlain J, Harries J E and Jones R G 1969 *Nature* **221** 143
[37] Dianov-Klokov V I, Ivanov V M, Arefev V N and Sizov N I 1981 *J. Quant. Spectrosc. Radiat. Transf.* **25** 83
[38] Bondarenko G V and Gorbaty Y E 1991 *Mol. Phys.* **74** 639
[39] Feyereisen M W, Feller D and Dixon D A 1996 *J. Phys. Chem.* **100** 2993
[40] Famulari A, Raimondi M, Sironi M and Gianinetti E 1998 *Chem. Phys.* **232** 275
[41] Xu X and Goddard W A 2004 *J. Phys. Chem.* A **108** 2305
[42] Haynes P, Skylaris C K, Mostofi A A and Payne M C 2006 *Chem. Phys. Lett.* **422** 345
[43] Sit P H-L and Marzari N 2005 *J. Chem. Phys.* **122** 204510
[44] Berendsen H J C, Grigera J R and Straatsma T P 1987 *J. Phys. Chem.* **91** 6269
[45] Perdew J P and Wang Y 1992 *Phys. Rev.* B **45** 13244
[46] Clough S A, Beers Y, Klein G P and Rothman L S 1973 *J. Phys. Chem.* **59** 2254
[47] Dyke T R and Muenter J S 1973 *J. Phys. Chem.* **59** 3125
[48] Silvestrelli P L and Parrinello M 1999 *Phys. Rev. Lett.* **82** 3308
[49] Badyal Y S, Saboungi M-L, Price D L, Shastri S D, Haeffner D R and Soper A K 2000 *J. Chem. Phys.* **112** 9206
[50] McQuarrie D A 2000 *Statistical Mechanics* (Sausalito, CA: University Science Books)
[51] Tangney P and Scandolo S 2002 *J. Chem. Phys.* **116** 14
[52] Kuo I-F W, Mundy C J, McGrath M J and Siepmann J I 2006 *J. Chem. Theory Comput.* **2** 1274
[53] Guardia E, Skarmoutsos I and Masia M 2015 *J. Phys. Chem.* B **119** 8926
[54] Water vibrational spectra www1.lsbu.ac.uk/water/water_vibrational_spectrum.html (visited June 2016) and references therein
[55] Plimpton S 1995 *J. Comput. Phys.* **117** 1
[56] Grossman J C, Schwegler E, Draeger E W, Gygi F and Galli G 2004 *J. Chem. Phys.* **120** 300
[57] Schwegler E, Grossman J C, Gygi F and Galli G 2004 *J. Chem. Phys.* **121** 5400
[58] Sorenson J M, Hura G, Glaeser R M and Head-Gordon T 2000 *J. Chem. Phys.* **113** 9149
[59] Liu L-M, Zhang C, Thornton G and Michaelides A 2010 *Phys. Rev.* B **82** 161415
[60] Cheng J and Sprik M 2010 *J. Chem. Theory Comput.* **6** 880
[61] Cheng J and Sprik M 2010 *Phys. Rev.* B **82** 081406
[62] Cheng J and Sprik M 2014 *J. Phys.: Condens. Matter* **26** 244108
[63] Sumita M, Hu C and Tateyama Y 2010 *J. Phys. Chem.* C **114** 18529
[64] Zhang C and Lindan P J D 2003 *J. Chem. Phys.* **119** 9183
[65] Tilocca A and Selloni A 2004 *J. Phys. Chem.* B **108** 4743
[66] Tilocca A and Selloni A 2004 *Langmuir* **20** 8379
[67] Harris L A and Quong A A 2004 *Phys. Rev. Lett.* **93** 086105
[68] Lindan P J D and Zhang C 2005 *Phys. Rev.* B **72** 075439
[69] Mattioli G, Filippone F, Caminiti R and Bonapasta A A 2008 *J. Chem. Phys.* C **112** 13579
[70] Lindan P J D, Harrison N M and Gillan M J 2010 *Phys. Rev. Lett.* **80** 762
[71] Tilocca A and Selloni A 2012 *J. Phys. Chem.* C **116** 9114
[72] English N J 2013 *Chem. Phys. Lett.* **583** 125
[73] Solveyra E G, de la Llave E, Molinero V, Soler-Illia G J A A and Scherlis D A 2013 *J. Phys. Chem.* C **117** 3330
[74] Wei M-J, Zhou J, Lu X, Zhu Y, Liu W, Lu L and Zhang L 2011 *Fluid Phase Equilib.* **302** 316
[75] Levchenko A A, Kolesnikov A I, Ross N L, Boerio-Goates J, Woodfield B F, Li G and Navrotsky A 2007 *J. Phys. Chem.* A **111** 12584