Reciprocal Polarizable Embedding with a Transferable H_2O Potential Function I: Formulation & Tests on Dimer

Elvar Örn Jónsson,* Asmus Ougaard Dohn, and Hannes Jónsson

Science Institute and Faculty of Physical Sciences, VR-III, University of Iceland, Reykjavík, Iceland

E-mail: eojons@gmail.com

Abstract

The incorporation of mutual polarization in multiscale simulations where different regions of the system are treated at different level of theory is important in studies of, for example, electronic excitations and charge transfer processes. We present here an energy functional for describing a quantum mechanics / molecular mechanics (QM/MM) scheme that includes reciprocal polarization between the two subsystems. The inclusion of polarization alleviates shortcomings inherent in electrostatic embedding QM/MM models based on point-charge force fields. A density functional theory (DFT) description of the QM subsystem is coupled to a single center multipole expansion (SCME) description of H_2O molecules in the MM subsystem that includes anisotropic dipole and quadrupole polarizability as well as static multipoles up to and including the hexadecapole. The energy functional and the coupling scheme is general and can be extended to arbitrary order in terms of both the static and induced moments. Tests of the energy surface for the H_2O dimer show that the QM/MM results lie

in between the pure DFT and pure SCME values. The consistency of the many-body contributions to the energy and analytical forces is demonstrated for an H_2O pentamer.

1 Introduction

The basic principle in multiscale methods is a compromise between computational cost, accuracy, and transferability by combining two (or more) levels of theory for different parts of the simulated system. The more sophisticated model is used to describe the region of the system where the main part of the chemical/physical processes take place while the less sophisticated model describes the remainder of the system needed to include the effect of the surroundings. Often, the more sophisticated model involves calculating the electronic structure (quantum mechanics, QM), while the less sophisticated model is based on a (semi)empirical potential energy function (molecular mechanics, MM). This class of multiscale methods is usually referred to as quantum mechanics / molecular mechanics (QM/MM). After the pioneering work of Levitt, Warshel,[1] and Karplus, the methodology for combining electronic structure methods and potential energy functions in this way has been developed further in many ways,[2–21] and applied to a variety of problems in several scientific fields. While the initial formulation and some subsequent advances included the effect of polarization of the MM subsystem, this effect is usually not included in QM/MM simulations.

The present work seeks to remedy this by providing a general formulation for the coupling of a QM model with a single center multipole expansion (SCME) polarizable force field for molecules in the MM region. Here, the SCME describes H_2O molecules.[22,23] First, a brief outline of the relevant background leading up to this work is given. Then, the SCME model is reviewed in the context of multiscale schemes, followed by a formulation of the mutual polarization of the QM and MM regions including analytical expressions for the atomic forces. The third part describes the non-electrostatic/induction interactions at the QM/MM interface. An assessment of the accuracy of the energetics and forces is provided by studies

of the water dimer. In an accompanying article, more extensive tests including liquid water and H_2O clusters are presented.[24]

2 Background

The QM/MM approach has been used extensively in studies of biochemical systems[14,25–27] such as enzyme-based catalysis[28] and medicine[29] where the division into subsystems typically cuts covalent bonds. Other important application areas are photochemistry[30] and solvation dynamics[31–34] where subdivision can be made without cutting covalent bonds, e.g. between the first and second solvation shells, or between solute and solvent. Likewise, theoretical studies of heterogeneous electrocatalysis can benefit from a QM/MM approach to include electrolyte structure and solvation effects in a more complete manner than is typically done so far.[35–41]

Thanks to the continued development of exchange-correlation functionals in the past decades, Kohn–Sham density functional theory[42,43] (KS-DFT) has become an essential work-horse in modern day computational chemistry and solid-state physics,[44] and is thus also a common choice for multiscale methods.[14,45–50] While calculations of large, condensed phase systems are usually limited to the generalized gradient approximation (GGA)and thereby relatively low accuracy in many cases, the use of a QM/MM scheme opens the possibility of using more accurate functionals for just the QM region.

Most commonly applied MM force fields are based on fixed charges on atomic centers or charge sites.[51–54] The parametrization of these models is often carried out in such a way that a few thermally averaged bulk properties are reproduced by the model. Point charge models are convenient choices for multiscale simulations since the electrostatic interaction coupling to a QM system is then quite simple. The static partial charge values assigned to the MM atoms do not allow for polarization of the MM subsystem. Therefore, they have limited applicability beyond the physical conditions (such as pressure or temperature) and

types of systems from which their parameters are derived. The H_2O molecule is a notable and important example, having a dipole moment of 1.8 D in gas phase and 3.1 D in ice Ih.[55] The interaction between H_2O is therefore strongly dependent on the local environment. The electrical field around a water cluster and within ice obtained from a fixed, point charge model fitted to liquid water has been shown to be inconsistent with results obtained from electronic structure calculations.[56]

One way to improve the static charge models is to include explicit and self-consistent polarization by (1) allowing the charge values at atomic sites to vary, or (2) by introducing a molecular or atomic dipole as modeled by charges connected with springs (Drude oscillators), or (3) through a more rigorous framework based, for example, on a perturbative decomposition of quantum mechanical interactions[57] – resulting in the multipole moment expansion. The fragments within these models naturally adjust their representations of charge densities to the electrostatic environment, and hence have the potential to be more transferable to heterogeneous environments as well as be able to provide an accurate response to changes due to chemical reactions and charge transfer processes. The molecular moment tensors and polarizabilities as defined from the perturbative expansion are well defined quantities, and can be extracted from high level quantum chemistry calculations or, in some cases, from experiments.

Many models of polarizability exist,[58–60] and extensive work has been done to utilizing them in multiscale methodologies, to produce polarizable embedding (PE-) QM/MM methods.[61–92] Originally, these efforts mainly focused on excited electronic states and Hartree-Fock or configuration interaction methods,[61,62,93] where the quantum chemical wave-function and the polarizable sites are variationally relaxed in a mutual fashion. Further development extended the efforts to more sophisticated quantum chemistry wave-function methods such as coupled cluster,[70–75,89] or time-dependent density functional theory.[70,76–84,90] It is evident from the literature that PE-QM/MM is mostly scrutinized in the context of electronic excitations.[61,73,75–77,80–83,85,87,89,90,94] However, systems involving highly polar species,[62,76,81]

inter- and intramolecular charge transfer,[61,87,95] biological reaction centers,[73,74,83,85] as well as solvated ionic and small molecular systems have also been studied.[63,74,76,78,91–93] Here, we present a formulation and validation of a globally self-consistent polarizable multiscale method that couples an explicit electronic description to a transferable and polarizable H_2O potential of high accuracy.[22,23] The potential is based on a single center multipole expansion (SCME) model and includes static multipole moment tensors up to and including the hexadecapole. No point charges are included since the expansion is around a single center in each of the neutral molecules. Furthermore, induced dipoles and induced quadrupoles in response to both the electric field and the electric field gradient are included, whereas commonly only dipole-induction is included. In contrast to isotropic atomic polarization sites that have most frequently been used in PE-QM/MM methods, the present formulation includes molecular anisotropic polarizability, a well defined quantity. Along with the globally self-consistent scheme, the analytical derivatives with respect to the QM nuclear coordinates and MM sites are presented, which allows for seamless use of the method in structural relaxation and Born-Oppenheimer molecular dynamics.

In this article, a derivation is given of the self-consistent solution to the coupled QM and MM energy functional that describes the static electrostatic interactions and mutual electrostatic polarization. The coupling between the QM and MM parts is handled by the atomic simulation environment (ASE) python based software.[96,97] Electrostatic potentials and gradients thereof – describing the interaction of the charge density of the QM system and the multipole moment tensors of the MM system – are implemented in the real-space grid-based projector augmented wave[98] software GPAW.[99,100] The self-consistent solution of the induced moments of the MM system under the influence of the QM system is implemented in a modified version of the original SCME potential function.[22,23] In the following, we limit the derivation to the point moments provided in the SCME model and do not provide the specifics of the GPAW Hamiltonian and atomic forces. The Supplementary Information gives more technical and implementation details.

3 Single-Center Multipole Expansion

In the SCME model,[22,23] the energy due to interaction between the water molecules is split into two parts

$$E^{\mathrm{MM}} = E^{\mathrm{MM}}_{\mathrm{ele+ind}} + E^{\mathrm{MM}}_{\mathrm{NE}} \tag{1}$$

where the first term describes the electrostatic plus induction interaction while the second term represents non-electrostatic effects, including repulsive exchange interaction and attractive dispersion interaction.

In the SCME model, each water molecule, i, is ascribed a dipole (μ^i_α), quadrupole ($\theta^i_{\alpha\beta}$), octupole ($\Omega^i_{\alpha\beta\gamma}$) and hexadecapole ($\Phi^i_{\alpha\beta\gamma\delta}$) moment tensor, with origin at the center of mass (COM). Furthermore, the external field, $\mathbf{V}^i_\alpha$ (negative of the electric field), and field gradient, $\mathbf{V}^i_{\alpha\beta}$ – evaluated at the COM of i due to the presence of $j(\neq i)$ SCME water molecules – further induces the dipole, $\Delta\mu^i_\alpha$, and quadrupole, $\Delta\theta^i_{\alpha\beta}$, moment tensors.[57] The induction terms include dipole–dipole ($\alpha_{\alpha\beta}$), dipole–quadrupole, ($A_{\alpha\beta\gamma}$), and quadrupole–quadrupole ($C_{\alpha\beta\gamma\delta}$) polarizability such that [1]

$$\Delta\mu^i_\alpha = -\alpha_{\alpha\beta}\mathbf{V}^i_\beta - \frac{1}{3}A_{\alpha,\beta\gamma}\mathbf{V}^i_{\beta\gamma} \tag{2}$$

$$\Delta\theta^i_{\alpha\beta} = -A_{\gamma,\alpha\beta}\mathbf{V}^i_\gamma - C_{\gamma\delta,\alpha\beta}\mathbf{V}^i_{\gamma\delta} \tag{3}$$

where the external field is additive

$$\mathbf{V}^i_\alpha = \sum_{j\neq i}^{n_{\mathrm{MM}}} \mathbf{V}^{ij}_\alpha \tag{4}$$

[1]Throughout this work we make use of the Einstein summation convention. Greek indexes $(\alpha, \beta, \ldots \eta)$ represent the set of Cartesian components x, y and z; i.e. α is $\{x, y, z\}$, $\alpha_1 = x$, and $\alpha = \beta = \cdots = \eta$. Repeated Greek indexes are to be summed over; for example $-\alpha_{\alpha\beta}\mathbf{V}^i_\beta = \sum_{s=1}^{3} -\alpha_{\alpha\beta_s}\mathbf{V}^i_{\beta_s}$. Vectors (and matrices) between points in Cartesian space are naturally represented in terms of the indexes; for example $\mathbf{r}_\alpha = (\mathbf{r}_x, \mathbf{r}_y, \mathbf{r}_z)$. Similarly the gadient operator is a vector operator, $\nabla_\alpha = (\nabla_x, \nabla_y, \nabla_z)$ – and hence expands a scalar field into a three component vector field, a three component vector field into a three-by-three matrix field, etc.

and given by

$$\mathbf{V}^{ij}_{\alpha} = -\mathbf{T}^{ij}_{\alpha\beta}(\mu^{j}_{\beta} + \Delta\mu^{j}_{\beta}) + \frac{1}{3}\mathbf{T}^{ij}_{\alpha\beta\gamma}(\theta^{j}_{\beta\gamma} + \Delta\theta^{j}_{\beta\gamma}) - \frac{1}{15}\mathbf{T}^{ij}_{\alpha\beta\gamma\delta}\Omega^{j}_{\beta\gamma\delta} + \frac{1}{105}\mathbf{T}^{ij}_{\alpha\beta\gamma\delta\epsilon}\Phi^{j}_{\beta\gamma\delta\epsilon} \tag{5}$$

Here we have introduced Coulomb interaction tensors between COM pairs i and j at Cartesian coordinates $\mathbf{r}^i$ and $\mathbf{r}^j$ where the zeroth order tensor is

$$T^{ij} = \frac{1}{r} = \frac{1}{|\mathbf{r}^i - \mathbf{r}^j|} \tag{6}$$

and higher orders are given by the successive use of the gradient operator ∇_α

$$\nabla_\alpha T^{ij} = \mathbf{T}^{ij}_{\alpha},\ \nabla_\beta \mathbf{T}^{ij}_{\alpha} = \mathbf{T}^{ij}_{\alpha\beta},\ \nabla_\eta \mathbf{T}^{ij}_{\alpha\beta\ldots} = \mathbf{T}^{ij}_{\alpha\beta\ldots\eta} \tag{7}$$

where

$$\nabla_\alpha = \frac{\partial}{\partial \mathbf{r}_\alpha} = \left(\frac{\partial}{\partial \mathbf{r}_x}, \frac{\partial}{\partial \mathbf{r}_y}, \frac{\partial}{\partial \mathbf{r}_z}\right) \tag{8}$$

and $\mathbf{r}_\alpha = (\mathbf{r}^i - \mathbf{r}^j)_\alpha$. Similarly, the field gradient in equations (2) and (3) is given by $\mathbf{V}^{i}_{\alpha\beta} = \nabla_\beta \mathbf{V}^{i}_{\alpha}$.

This nonlinear scheme between equations (2)–(5) is solved iteratively since the induction of the dipole and quadrupole moment tensors results in a change in the external field and field gradient at the COM of the other $j(\neq i)$ molecules, and vice versa.

The relationship between the external field in equation (5) and induced moments in equations (2)-(3) is a general result for linearly induced moments where the internal energy change required to induce the moments depends bilinearly on the on-site potential. By

solving[57,101]

$$\frac{\partial E^{\mathrm{MM}}_{\mathrm{ele+ind}}}{\partial \Delta\mu^i_\alpha} = 0 \tag{9}$$

$$\frac{\partial E^{\mathrm{MM}}_{\mathrm{ele+ind}}}{\partial \Delta\theta^i_{\alpha\beta}} = 0, \tag{10}$$

the resulting electrostatic plus induction energy functional becomes

$$\begin{aligned} E^{\mathrm{MM}}_{\mathrm{ele+ind}} =& \frac{1}{2}\sum_i^{n_{\mathrm{MM}}} \left((\mu^i_\alpha + \Delta\mu^i_\alpha)\mathbf{V}^i_\alpha + \frac{1}{3}(\theta^i_{\alpha\beta} + \Delta\theta^i_{\alpha\beta})\mathbf{V}^i_{\alpha\beta} + \frac{1}{15}\Omega^i_{\alpha\beta\gamma}\mathbf{V}^i_{\alpha\beta\gamma} + \frac{1}{105}\Phi^i_{\alpha\beta\gamma\delta}\mathbf{V}^i_{\alpha\beta\gamma\delta} \right) \\ & - \frac{1}{2}\sum_i^{n_{\mathrm{MM}}} \left(\Delta\mu^i_\alpha \mathbf{V}^i_\alpha + \frac{1}{3}\Delta\theta^i_{\alpha\beta}\mathbf{V}^i_{\alpha\beta} \right) \\ =& E^{\mathrm{MM}*}_{\mathrm{ele+ind}} + \sum_i^{n_{\mathrm{MM}}} E^{\mathrm{MM},i}_{\mathrm{self}} \end{aligned} \tag{11}$$

where the first term on the right hand side describes the intermolecular electrostatic energy of the static and induced moments. The second term on the right hand side is a self-energy term that accounts for the energy required to distort a ground state charge density (whose multipole expansion results in the static moment tensors) to a polarized charge density (whose multipole expansion results in the static plus induced moment tensors) for each MM molecule in the system. A generalization of the interaction between linearly induced moments and resulting self-energy terms can be found elsewhere.[101]

It is customary to write the energy expression in equation (11) at self-consistency as

$$E^{\mathrm{MM}}_{\mathrm{ele+ind}} = \frac{1}{2}\sum_i^{n_{\mathrm{MM}}} \left(\mu^i_\alpha \mathbf{V}^i_\alpha + \frac{1}{3}\theta^i_{\alpha\beta}\mathbf{V}^i_{\alpha\beta} + \frac{1}{15}\Omega^i_{\alpha\beta\gamma}\mathbf{V}^i_{\alpha\beta\gamma} + \frac{1}{105}\Phi^i_{\alpha\beta\gamma\delta}\mathbf{V}^i_{\alpha\beta\gamma\delta} \right) \tag{12}$$

since there is a net cancellation between the self-energy and the intermolecular induced-static and induced-induced electrostatic interactions. However, they cannot be assumed to simply cancel out in a mutually polarizable QM/MM interface. A distinction is made made between the total electrostatic plus induction and the intermolecular (or interspecies) electrostatic

interaction plus induction with an asterisk, and the general form of equation (11) is used throughout.

4 Global Interaction

In order to arrive at a self-consistent model for the mutual polarization of the QM subsystem and MM subsystem, we start by considering the interaction of the total system in a global reference frame, chosen to be an orthorhombic box with a left handed Cartesian coordinate system with origin at a box corner. The reference frame encompasses both the MM and QM subsystems and has a total charge density $\rho(\mathbf{r})$

$$\rho(\mathbf{r}) = \sum_{i}^{n_{\text{sys}}} \rho_i(\mathbf{r}) = \rho_{\text{QM}}(\mathbf{r}) + \rho_{\text{MM1}}(\mathbf{r}) + \rho_{\text{MM2}}(\mathbf{r}) + \dots \tag{13}$$

where i runs over all n_{sys} charge densities in the system, and includes both the charge density of the QM subsystem and conceptual charge densities in the MM subsystem, one for each SCME molecule. The energy functional of the total system can then be written as

$$E^{\text{sys}}_{\text{tot}}[\rho(\mathbf{r})] = E^{\text{KS}}[\rho_{\text{QM}}(\mathbf{r})] + E^{\text{sys}}_{\text{int}}[\rho(\mathbf{r})] \tag{14}$$

where $E^{\text{KS}}[\rho_{\text{QM}}(\mathbf{r})]$ is the conventional KS–DFT energy functional of the QM charge density, and $E^{\text{sys}}_{\text{int}}[\rho(\mathbf{r})]$ is a global interaction energy functional of the total charge density.

The QM energy functional is

$$E^{\text{KS}}[\rho_{\text{QM}}(\mathbf{r})] = T[\{\psi_{\text{QM}}\}] + E_{\text{Coul}}[\rho_{\text{QM}}(\mathbf{r})] + E_{\text{xc}}[n_{\text{QM}}(\mathbf{r})] + \int \rho_{\text{QM}}(\mathbf{r}) V_{\text{ext}}(\mathbf{r}) d\mathbf{r}. \tag{15}$$

It contains a kinetic energy term of the non-interacting KS single particle states, $\{\psi_{\text{QM}}\}$, a Coulomb term of the charge density, an exchange–correlation term that depends on the QM electron density, $n_{\text{QM}}(\mathbf{r})$, and, for now, some arbitrary external potential term, $V_{\text{ext}}(\mathbf{r})$. A

corresponding Hamiltonian is defined as

$$H_{\text{tot}}^{\text{KS}} = H^{\text{KS}} + V_{\text{ext}} \tag{16}$$

The global interaction term that couples the different charge densities is furthermore split into two contributions

$$E_{\text{int}}^{\text{sys}}[\rho(\mathbf{r})] = E_{\text{ele+ind}}^{\text{sys}}[\rho(\mathbf{r})] + E_{\text{NE}}^{\text{sys}}$$

an electrostatic plus induction term that depends on the total charge density, and a non-electostatic term that does not depend on the charge density. It is assumed that the electrostatic plus induction energy functional splits into two contributions

$$E_{\text{ele+ind}}^{\text{sys}}[\rho(\mathbf{r})] = E_{\text{ele+ind}}^{\text{sys*}}[\rho(\mathbf{r})] + \sum_{i}^{n_{\text{MM}}} E_{\text{self}}^{\text{MM,i}}[\rho(\mathbf{r})] \tag{17}$$

along the same lines as the energy expression in equation (11) for the linearly induced moments of the MM system.

A total energy expression for the system is now written as

$$E_{\text{tot}}^{\text{sys}}[\rho(\mathbf{r})] = E^{\text{KS}}[\rho_{\text{QM}}(\mathbf{r})] + E_{\text{ele+ind}}^{\text{sys*}}[\rho(\mathbf{r})] + \sum_{i}^{n_{\text{MM}}} E_{\text{self}}^{\text{MM},i}[\rho(\mathbf{r})] + E_{\text{NE}}^{\text{sys}} \tag{18}$$

Thus, it is apparent that a self-consistent solution of the QM charge density simultaneously minimizing the KS-DFT energy functional and the total energy functional depends on the self-consistent solution of the MM charge densities and the corresponding self-energy, and vice versa.

4.1 Electrostatic Multipole Embedding Model

We at first consider the static case where all of the individual charge densities are decoupled (i.e. are in their ground state) and non-overlapping, and assume that the global intermolecular electrostatic plus induction can be separated into two terms

$$E^{\mathrm{sys}*}_{\mathrm{ele+ind}}[\rho(\mathbf{r})] = E^{\mathrm{sys}}_{\mathrm{ele}}[\rho(\mathbf{r})] + E^{\mathrm{sys}*}_{\mathrm{ind}}[\rho(\mathbf{r})] \tag{19}$$

In terms of the total charge density, the electrostatic interaction can be written as

$$E^{\mathrm{sys}}_{\mathrm{ele}}[\rho(\mathbf{r})] = \frac{1}{2}\sum_{i}^{n_{\mathrm{sys}}}\sum_{j\neq i}^{n_{\mathrm{sys}}}\int\int\frac{\rho_i(\mathbf{r})\rho_j(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' \tag{20}$$

The partial derivative of the electrostatic interaction leads to

$$\frac{\partial E^{\mathrm{sys}}_{\mathrm{ele}}[\rho(\mathbf{r})]}{\partial\rho_l(\mathbf{r}'')} = \sum_{j\neq l}^{n_{\mathrm{sys}}}\int\frac{\rho_j(\mathbf{r}')}{|\mathbf{r}''-\mathbf{r}'|}d\mathbf{r}' \tag{21}$$

expressing the electrostatic interaction potential $V^l(\mathbf{r}'')$ at some arbitrary coordinate $\mathbf{r}''$ in the global reference frame. We can choose this charge density to be the explicit QM charge density, or one of the conceptual charge densities of the MM subsystem and corresponding coordinate of the COM of a SCME molecule. Choosing the latter we can write the potential at site i due to all of the other $j(\neq i)$ charge densities as

$$V^i = \sum_{j\neq i}^{n_{\mathrm{sys}}} V^{ij} \tag{22}$$

$$V^{ij} = \int\frac{\rho_j(\mathbf{r})}{|\mathbf{r}^i-\mathbf{r}|}d\mathbf{r} \tag{23}$$

where the shorthand abbreviation $V^i(\mathbf{r}^i) = V^i$ is used. Isolating the term due to the QM charge density gives

$$V^i = V^{i\mathrm{QM}} + \sum_{j\neq i}^{n_{\mathrm{MM}}} V^{ij}. \tag{24}$$

A Taylor expansion of the potential at site i due to a charge density associated with MM molecule j in the V^{ij} term leads to[57]

$$V^{ij} = -\mathbf{T}^{ij}_{\alpha}\mu^{j}_{\alpha} + \frac{1}{3}\mathbf{T}^{ij}_{\alpha\beta}\theta^{j}_{\alpha\beta} - \frac{1}{15}\mathbf{T}^{ij}_{\alpha\beta\gamma}\Omega^{j}_{\alpha\beta\gamma} + \frac{1}{105}\mathbf{T}^{ij}_{\alpha\beta\gamma\delta}\Phi^{j}_{\alpha\beta\gamma\delta} \tag{25}$$

which is expressed in terms of interaction tensors and the multipole moment tensors associated with each MM molecule. This Taylor expansion in turn expresses the electrostatic potential acting on the QM charge density due to the MM subsystem

$$\begin{aligned}
\frac{\partial E^{\mathrm{sys}}_{\mathrm{ele}}[\rho(\mathbf{r}')]}{\partial \rho_{\mathrm{QM}}(\mathbf{r})} =& \frac{\partial}{\partial \rho_{\mathrm{QM}}(\mathbf{r})} \sum_{j}^{n_{\mathrm{MM}}} \int \rho_{\mathrm{QM}}(\mathbf{r}') V^{r'j} d\mathbf{r}' = \sum_{j}^{n_{\mathrm{MM}}} \int \delta(\mathbf{r}-\mathbf{r}') V^{r'j} d\mathbf{r}' \\
=& \sum_{j}^{n_{\mathrm{MM}}} \left(-\mathbf{T}^{rj}_{\alpha}\mu^{j}_{\alpha} + \frac{1}{3}\mathbf{T}^{rj}_{\alpha\beta}\theta^{j}_{\alpha\beta} - \frac{1}{15}\mathbf{T}^{rj}_{\alpha\beta\gamma}\Omega^{j}_{\alpha\beta\gamma} + \frac{1}{105}\mathbf{T}^{rj}_{\alpha\beta\gamma\delta}\Phi^{j}_{\alpha\beta\gamma\delta} \right) \\
=& V^{\mathrm{QM/MM}}_{\mathrm{ele}}(\mathbf{r})
\end{aligned} \tag{26}$$

where the zeroth order tensor is in this case $T^{rj} = 1/(|\mathbf{r} - \mathbf{r}^{j}|)$. Using all of the above, the electrostatic interaction of the system can be expressed as

$$\begin{aligned}
E^{\mathrm{sys}}_{\mathrm{ele}}[\rho(\mathbf{r})] =& \int \rho_{\mathrm{QM}}(\mathbf{r}) V^{\mathrm{QM/MM}}_{\mathrm{ele}}(\mathbf{r}) d\mathbf{r} \\
&+ \frac{1}{2} \sum_{i}^{n_{\mathrm{MM}}} \sum_{j\neq i}^{n_{\mathrm{MM}}} \left(\mu^{i}_{\alpha}\mathbf{V}^{ij}_{\alpha} + \frac{1}{3}\theta^{i}_{\alpha\beta}\mathbf{V}^{ij}_{\alpha\beta} + \frac{1}{15}\Omega^{i}_{\alpha\beta\gamma}\mathbf{V}^{ij}_{\alpha\beta\gamma} + \frac{1}{105}\Phi^{i}_{\alpha\beta\gamma\delta}\mathbf{V}^{ij}_{\alpha\beta\gamma\delta} \right) \\
=& E^{\mathrm{QM/MM}}_{\mathrm{ele}}[\rho_{\mathrm{QM}}(\mathbf{r})] + E^{\mathrm{MM}}_{\mathrm{ele}}.
\end{aligned} \tag{27}$$

The global electrostatic interaction splits into an explicit QM/MM electrostatic interaction – written in terms of the QM charge density and the Cartesian multipole moments of MM molecules – and an explicit MM electrostatic interaction.

At this stage the energy functional can be written as

$$E_{\mathrm{tot}}^{\mathrm{sys}}[\rho(\mathbf{r})] = E^{\mathrm{KS}}[\rho_{\mathrm{QM}}(\mathbf{r})] + E_{\mathrm{ele}}^{\mathrm{QM/MM}}[\rho_{\mathrm{QM}}(\mathbf{r})] + E_{\mathrm{ind}}^{\mathrm{sys*}}[\rho(\mathbf{r})] + E_{\mathrm{ele}}^{\mathrm{MM}} + \sum_{i}^{n_{\mathrm{MM}}} E_{\mathrm{self}}^{\mathrm{MM},i}[\rho(\mathbf{r})] + E_{\mathrm{NE}}^{\mathrm{sys}} \tag{28}$$

and what is left to be described is the global induction interaction, which includes the interspecies QM/MM and intermolecular MM/MM induction interactions and corresponding self-energy of the MM subsystem.

4.2 QM/MM Induction Model

The evaluation of the induction involves two self-consistent cycles that are coupled. One cycle is for the MM subsystem as described in Section 3. The second is for the quantum mechanical system where a conventional iterative self-consistent field (SCF) method is used to solve for the electron density that minimizes the KS–DFT energy functional, equation (15). These two cycles are analogous in the sense that they are both based on the SCF approach, hence cycle one is referred to as MM-SCF, and cycle two as QM-SCF. The idea of coupling two different SCF cycles is not new and was first introduced in pioneering work on PE-QM/MM.[61] The SCF solver for the combined system is referred to as QM/MM-SCF.

In order to couple the two SCF solvers the relevant potentials and gradients thereof are included in the self-consistent equations of the MM-SCF solver, such that the induced moments are polarized by the QM charge density. Similarly, for the QM-SCF, a self-consistent solution involves diagonalizing the KS–DFT Hamiltonian corresponding to equation (15)

$$(H^{\mathrm{KS}} + V_{\mathrm{ext}}^{\mathrm{QM/MM}})\,|\psi_m\rangle = \epsilon_m\,|\psi_m\rangle \tag{29}$$

where $V_{\mathrm{ext}}^{\mathrm{QM/MM}}$ includes both the electrostatic plus induction interaction as defined in this section. In analogy to the classical variational condition of equations (9) and (10), the

stationary point satisfies

$$\frac{\partial E^{\mathrm{KS}}[\rho_{\mathrm{QM}}(\mathbf{r})]}{\partial \langle\psi_m|} = 0. \tag{30}$$

The computational cost of reaching self-consistency for the MM-SCF is small compared to the QM-SCF, so before coupling the two subsystems the MM-SCF is iterated until it reaches the stationary point of the MM energy functional, equation (11). The MM moments are

$$\mu_\alpha^{i'} = \mu_\alpha^i + \Delta\mu_\alpha^i \tag{31}$$

$$\theta_{\alpha\beta}^{i'} = \theta_{\alpha\beta}^i + \Delta\theta_{\alpha\beta}^i. \tag{32}$$

At this stage the QM and MM subsystems are coupled, and from the point of view of the QM subsystem the MM point moments in equations (31) and (32) are fixed, hence the QM/MM electrostatic interaction energy is

$$E_{\mathrm{ele}}^{\mathrm{QM/MM}}[\rho_{\mathrm{QM}}(\mathbf{r})] = \int \rho_{\mathrm{QM}}(\mathbf{r}) \sum_j^{n_{\mathrm{MM}}} \left(-\mathbf{T}_\alpha^{rj}\mu_\alpha^{j'} + \frac{1}{3}\mathbf{T}_{\alpha\beta}^{rj}\theta_{\alpha\beta}^{j'} - \frac{1}{15}\mathbf{T}_{\alpha\beta\gamma}^{rj}\Omega_{\alpha\beta\gamma}^j + \frac{1}{105}\mathbf{T}_{\alpha\beta\gamma\delta}^{rj}\Phi_{\alpha\beta\gamma\delta}^j \right) d\mathbf{r}. \tag{33}$$

Similarly, from the point of view of the MM subsystem, there is a non-uniform field due to the presence of the QM charge density. This can be expressed as an additional potential at each MM site i as

$$V^{i\mathrm{QM}} = \int \rho_{\mathrm{QM}}(\mathbf{r}) T^{ir} d\mathbf{r} \tag{34}$$

with external field and field gradient

$$\mathbf{V}_\alpha^{i\mathrm{QM}} = \int \rho_{\mathrm{QM}}(\mathbf{r}) \left(-\frac{1}{r^3}\mathbf{r}_\alpha \right) d\mathbf{r} = \int \rho_{\mathrm{QM}}(\mathbf{r}) \mathbf{T}_\alpha^{ir} d\mathbf{r} \tag{35}$$

$$\mathbf{V}_{\alpha\beta}^{i\mathrm{QM}} = \int \rho_{\mathrm{QM}}(\mathbf{r}) \left(\frac{3}{r^5}\mathbf{r}_\alpha\mathbf{r}_\beta - \frac{\delta_{\alpha\beta}}{r^3} \right) d\mathbf{r} = \int \rho_{\mathrm{QM}}(\mathbf{r}) \mathbf{T}_{\alpha\beta}^{ir} d\mathbf{r}. \tag{36}$$

The instantaneous response of the MM sites to the additional potential has two effects. First,

the induced moments as a result of the QM external field and field gradient will be

$$\Delta\mu_{\alpha}^{i\mathrm{QM}} = -\alpha_{\alpha\beta}\mathbf{V}_{\beta}^{i\mathrm{QM}} - \frac{1}{3}A_{\alpha,\beta\gamma}\mathbf{V}_{\beta\gamma}^{i\mathrm{QM}} \tag{37}$$

$$\Delta\theta_{\alpha\beta}^{i\mathrm{QM}} = -A_{\gamma,\alpha\beta}\mathbf{V}_{\gamma}^{i\mathrm{QM}} - C_{\gamma\delta,\alpha\beta}\mathbf{V}_{\gamma\delta}^{i\mathrm{QM}} \tag{38}$$

so the interspecies electrostatic plus induction interaction energy of the QM/MM system becomes

$$E_{\mathrm{ele+ind}}^{\mathrm{QM/MM*}}[\rho_{\mathrm{QM}}(\mathbf{r})] = \int \rho_{\mathrm{QM}}(\mathbf{r}) \sum_{j}^{n_{\mathrm{MM}}} \Big(-\mathbf{T}_{\alpha}^{rj}(\mu_{\alpha}^{j'} + \Delta\mu_{\alpha}^{j\mathrm{QM}}) + \frac{1}{3}\mathbf{T}_{\alpha\beta}^{rj}(\theta_{\alpha\beta}^{j'} + \Delta\theta_{\alpha\beta}^{j\mathrm{QM}}) - \frac{1}{15}\mathbf{T}_{\alpha\beta\gamma}^{rj}\Omega_{\alpha\beta\gamma}^{j} + \frac{1}{105}\mathbf{T}_{\alpha\beta\gamma\delta}^{rj}\Phi_{\alpha\beta\gamma\delta}^{j}\Big) d\mathbf{r}. \tag{39}$$

Secondly, the self-energy of the MM sites also changes due to this instantaneous polarization in a way that depends bi-linearly on the on-site potential[101]

$$E_{\mathrm{self}}^{\mathrm{QM/MM}}[\rho_{\mathrm{QM}}(\mathbf{r})] = -\frac{1}{2}\sum_{j}^{n_{\mathrm{MM}}} \left(\Delta\mu_{\alpha}^{j\mathrm{QM}}\mathbf{V}_{\alpha}^{j\mathrm{QM}} + \frac{1}{3}\Delta\theta_{\alpha\beta}^{j\mathrm{QM}}\mathbf{V}_{\alpha\beta}^{j\mathrm{QM}} \right). \tag{40}$$

In this way, the total QM/MM electrostatic plus induction energy is defined as

$$E_{\mathrm{ele+ind}}^{\mathrm{QM/MM}}[\rho_{\mathrm{QM}}(\mathbf{r})] = E_{\mathrm{ele+ind}}^{\mathrm{QM/MM*}}[\rho_{\mathrm{QM}}(\mathbf{r})] + E_{\mathrm{self}}^{\mathrm{QM/MM}}[\rho_{\mathrm{QM}}(\mathbf{r})]. \tag{41}$$

The induced moments in equations (37) and (38) are now fixed, and by using the expression above, the form of the external potential that couples the MM subsystem to the QM

subsystem becomes

$$\begin{aligned} V_{\text{ext}}^{\text{QM/MM}}(\mathbf{r}) =& \frac{\partial E_{\text{ele+ind}}^{\text{QM/MM}}[\rho_{\text{QM}}(\mathbf{r})]}{\partial \rho_{\text{QM}}(\mathbf{r})} \\ =& \sum_{j}^{n_{\text{MM}}} \left(\mathbf{T}_{\alpha}^{rj} \left(\mu_{\alpha}^{j'} + \frac{1}{2} \Delta \mu_{\alpha}^{j\text{QM}} \right) + \frac{1}{3} \mathbf{T}_{\alpha\beta}^{rj} \left(\theta_{\alpha\beta}^{j'} + \frac{1}{2} \Delta \theta_{\alpha\beta}^{j\text{QM}} \right) \right. \\ & \left. + \frac{1}{15} \mathbf{T}_{\alpha\beta\gamma}^{rj} \Omega_{\alpha\beta\gamma}^{j} + \frac{1}{105} \mathbf{T}_{\alpha\beta\gamma\delta}^{rj} \Phi_{\alpha\beta\gamma\delta}^{j} \right) \\ =& V_{\text{ele+ind}}^{\text{QM/MM*}}(\mathbf{r}) + V_{\text{self}}^{\text{QM/MM}}(\mathbf{r}) \end{aligned} \tag{42}$$

where the explicit Coulomb interaction with the static and induced MM moments combines into a single term that includes the cost of inducing the moments – i.e. the self-energy. The QM/MM interface energy can then be written as

$$E_{\text{ele+ind}}^{\text{QM/MM}}[\rho_{\text{QM}}(\mathbf{r})] = \int \rho_{\text{QM}}(\mathbf{r}) V_{\text{ext}}^{\text{QM/MM}}(\mathbf{r}) d\mathbf{r}. \tag{43}$$

This external potential is a key result, since if the self-energy counterbalancing the explicit instantaneous polarization is taken into account, the QM and MM subsystems would over-polarize each other. This term is added to the KS-DFT Hamiltonian of equation (29) and a single iteration performed with the QM-SCF solver.

After a single QM-SCF iteration the QM charge density is perturbed due to the presence of the MM moments

$$\rho'_{\text{QM}}(\mathbf{r}) = \rho_{\text{QM}}(\mathbf{r}) + \Delta \rho_{\text{QM}}(\mathbf{r}). \tag{44}$$

This is the QM polarization analog of the polarization in the MM subsystem. In the next step, the induced dipoles and induced quadrupoles are evaluated again, but the potential at the start of the new MM-SCF cycle now includes the polarized QM charge density and the

optimized moments of equations (31) and (32) as an initial guess

$$V^i = \int \rho'_{\rm QM}(\mathbf{r}) T^{ir} d\mathbf{r} + \sum_j^{n_{\rm MM}} \left(-\mathbf{T}^{ij}_{\alpha}\mu^{j'}_{\alpha} + \frac{1}{3}\mathbf{T}^{ij}_{\alpha\beta}\theta^{j'}_{\alpha\beta} - \frac{1}{15}\mathbf{T}^{ij}_{\alpha\beta\gamma}\Omega_{\alpha\beta\gamma} + \frac{1}{105}\mathbf{T}^{ij}_{\alpha\beta\gamma\delta}\Phi^{j}_{\alpha\beta\gamma\delta} \right) \tag{45}$$

The induced moments are at first

$$\Delta\mu^i_{\alpha} = -\alpha_{\alpha\beta}\left(\mathbf{V}^{i\rm QM}_{\beta} + \sum_j^{n_{\rm MM}} \mathbf{V}^{ij}_{\beta} \right) - \frac{1}{3}A_{\alpha,\beta\gamma}\left(\mathbf{V}^{i\rm QM}_{\beta\gamma} + \sum_j^{n_{\rm MM}} \mathbf{V}^{ij}_{\beta\gamma} \right) \tag{46}$$

$$\Delta\theta^i_{\alpha\beta} = -A_{\gamma,\alpha\beta}\left(\mathbf{V}^{i\rm QM}_{\gamma} + \sum_j^{n_{\rm MM}} \mathbf{V}^{ij}_{\gamma} \right) - C_{\gamma\delta,\alpha\beta}\left(\mathbf{V}^{i\rm QM}_{\gamma\delta} + \sum_j^{n_{\rm MM}} \mathbf{V}^{ij}_{\gamma\delta} \right) \tag{47}$$

giving rise to new MM site-to-site external fields due to the induced moments by virtue of equation (5). Equations (46)–(47) and (5) are iterated to convergence. At MM-SCF self-consistency,

$$\mu^{i''}_{\alpha} = \Delta\mu^{i\rm QM}_{\alpha} + (\mu^{i'}_{\alpha}) \tag{48}$$

$$\theta^{i''}_{\alpha\beta} = \Delta\theta^{i\rm QM}_{\alpha\beta} + (\theta^{i'}_{\alpha\beta}) \tag{49}$$

and the resulting MM intermolecular electrostatic plus induction energy is

$$E^{\rm MM*}_{\rm ele+ind} = \frac{1}{2}\sum_i^{n_{\rm MM}}\sum_{j\neq i}^{n_{\rm MM}} \left(\mu^{i''}_{\alpha}\mathbf{V}^{ij}_{\alpha} + \frac{1}{3}\theta^{i''}_{\alpha\beta}\mathbf{V}^{ij}_{\alpha\beta} + \frac{1}{15}\Omega^{i}_{\alpha\beta\gamma}\mathbf{V}^{ij}_{\alpha\beta\gamma} + \frac{1}{105}\Phi^{i}_{\alpha\beta\gamma\delta}\mathbf{V}^{ij}_{\alpha\beta\gamma\delta} \right). \tag{50}$$

The potential in the equation above acts between the MM pairs only but it includes the effects of the QM charge density through the polarized moments as defined in equations (48) and (49). However, the MM self-energy at each site is given by the total external field and field gradient at that site

$$E^{\rm MM,i}_{\rm self} = -\frac{1}{2}\left(\Delta\mu^i_{\alpha}\mathbf{V}^i_{\alpha} + \frac{1}{3}\Delta\theta^i_{\alpha\beta}\mathbf{V}^i_{\alpha\beta} \right). \tag{51}$$

This concludes one QM/MM-SCF iteration. A new iteration is then initialized. The moments in the parenthesis of equations (48) and (49) enter equation (33), and the polarized QM charge density of equation (44) enters equation (34). This process is continued until both the QM charge density and the MM induced moments reach a desired convergence threshold. Figure 1 shows a schematic overview of the PE-QM/MM scheme.

At global self-consistency, the following identity holds

$$\int \rho'_{\mathrm{QM}}(\mathbf{r}) V_{\mathrm{ext}}^{\mathrm{QM/MM}}(\mathbf{r}) d\mathbf{r} = \sum_{i}^{n_{\mathrm{MM}}} \left(\mu_{\alpha}^{i''} \mathbf{V}_{\alpha}^{i\mathrm{QM}} + \frac{1}{3} \theta_{\alpha\beta}^{i''} \mathbf{V}_{\alpha\beta}^{i\mathrm{QM}} + \frac{1}{15} \Omega_{\alpha\beta\gamma}^{i} \mathbf{V}_{\alpha\beta\gamma}^{i\mathrm{QM}} + \frac{1}{105} \Phi_{\alpha\beta\gamma\delta}^{i} \mathbf{V}_{\alpha\beta\gamma\delta}^{i\mathrm{QM}} \right) - \frac{1}{2} \sum_{i}^{n_{\mathrm{MM}}} \left(\Delta\mu_{\alpha}^{i\mathrm{QM}} \mathbf{V}_{\alpha}^{i\mathrm{QM}} + \frac{1}{3} \Delta\theta_{\alpha\beta}^{i\mathrm{QM}} \mathbf{V}_{\alpha\beta}^{i\mathrm{QM}} \right) \tag{52}$$

and hence the energy functional of the combined system can be written as

$$E_{\mathrm{tot}}^{\mathrm{sys}} = E^{\mathrm{KS}}[\rho'_{\mathrm{QM}}(\mathbf{r})] + E_{\mathrm{ele+ind}}^{\mathrm{QM/MM}}[\rho'_{\mathrm{QM}}(\mathbf{r})] + E_{\mathrm{ele+ind}}^{\mathrm{MM*}} + \sum_{i}^{n_{\mathrm{MM}}} \left(E_{\mathrm{self}}^{\mathrm{MM},i} - E_{\mathrm{self}}^{\mathrm{MM},i\mathrm{QM}} \right) + E_{\mathrm{NE}}^{\mathrm{sys}} \tag{53}$$

where the fifth term on the right hand side is

$$E_{\mathrm{self}}^{\mathrm{MM,iQM}} = -\frac{1}{2} \left(\Delta\mu_{\alpha}^{i\mathrm{QM}} \mathbf{V}_{\alpha}^{i\mathrm{QM}} + \frac{1}{3} \Delta\theta_{\alpha\beta}^{i\mathrm{QM}} \mathbf{V}_{\alpha\beta}^{i\mathrm{QM}} \right). \tag{54}$$

It is subtracted from the total MM self-energy such as to not double count this term, as it is already present in the QM/MM electrostatic plus induction energy expression (equation (40)) as a result of the coupling scheme. Equations (52) is rigorously true at the QM/MM-SCF stationary point, where the conditions of equations (9)–(10) and (30) are all simultaneously valid. The energy functional expression in equation (53) is used for the derivation of the analytical derivatives with respect to both the QM nuclei positions and the COM of the MM sites. Figure 1 shows the resulting polarizable embedding QM/MM scheme, based on the energy functional in equation 53.

Note that the total energy functional expression can be simplified further at the QM/MM-

SCF stationary point by using equations (39), (52) and (54) to give

$$E_{\text{tot}}^{\text{sys}} = E^{\text{KS}}[\rho'_{\text{QM}}(\mathbf{r})] + E_{\text{ele+ind}}^{\text{QM/MM*}}[\rho'_{\text{QM}}(\mathbf{r})] + E_{\text{ele+ind}}^{\text{MM*}} + \sum_{i}^{n_{\text{MM}}} E_{\text{self}}^{\text{MM},i} + E_{\text{NE}}^{\text{sys}} \tag{55}$$

such that the interspecies QM/MM and intermolecular MM/MM electrostatic plus induction interactions are described by the second and third term on the right hand side, and the MM self-energy is expressed with a single term.

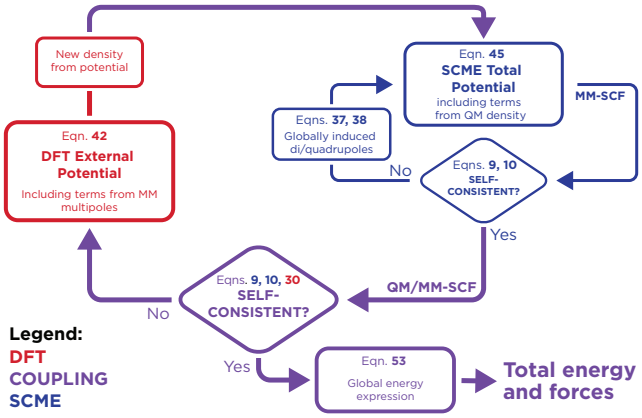

Figure 1: Schematic outlining the global self-consistent scheme where a MM-SCF (blue) and a QM-SCF (red) are coupled, resulting in the QM/MM-SCF cycle (purple). One iteration of the QM/MM-SCF involves: first, fully self-consistently calculating the induced MM moments – under the influence of the QM charge density (from a previous iteration) through an additional static potential at each classical site. Second, using the new updated numerical values of the MM induced moments to evaluate the external potential as derived from the QM/MM interface energy, equation (41), and iterate the QM-SCF once resulting in a new QM charge-density polarized by the MM moments. This concludes one QM/MM-SCF iteration.

4.3 Forces

Analytical expressions for the atomic forces are obtained from the negative gradient of equation (53) with respect to the position of the MM sites, or the position of the nuclei in the QM system. Note that expression (55) leads to incorrect expressions for the atomic forces, which will be adressed in the numerical tests presented later in this work. Consider first a

QM nucleus, indexed as z

$$\begin{aligned}\mathbf{F}_{\alpha}^{z} &= -\frac{dE_{\mathrm{tot}}^{\mathrm{sys}}}{d\mathbf{r}_{\alpha}^{z}} \\ &= -\frac{\partial E_{\mathrm{tot}}^{\mathrm{sys}}}{\partial \mathbf{r}_{\alpha}^{z}} - \frac{\partial E_{\mathrm{tot}}^{\mathrm{sys}}}{\partial \Delta\mu_{\beta}^{i}}\frac{\partial \Delta\mu_{\beta}^{i}}{\partial \mathbf{r}_{\alpha}^{z}} - \frac{\partial E_{\mathrm{tot}}^{\mathrm{sys}}}{\partial \Delta\theta_{\beta}^{i}}\frac{\partial \Delta\theta_{\beta}^{i}}{\partial \mathbf{r}_{\alpha}^{z}} - \frac{\partial E_{\mathrm{tot}}^{\mathrm{sys}}}{\partial \langle\psi_{m}|}\frac{\partial \langle\psi_{m}|}{\partial \mathbf{r}_{\alpha}^{z}}\end{aligned} \tag{56}$$

The atomic forces are evaluated at the stationary point of the total energy functional, where the following conditions apply

$$\frac{\partial E_{\mathrm{tot}}^{\mathrm{sys}}}{\partial \Delta\mu_{\alpha}^{i}} = 0,\ \frac{\partial E_{\mathrm{tot}}^{\mathrm{sys}}}{\partial \Delta\theta_{\alpha}^{i}} = 0,\ \frac{\partial E_{\mathrm{tot}}^{\mathrm{sys}}}{\partial \langle\psi_{m}|} = 0$$

Hence we only need to explicitly evaluate the first term on the right hand side of equation (56), which is

$$\begin{aligned}\mathbf{F}_{\alpha}^{z} = -\frac{\partial E_{\mathrm{tot}}^{\mathrm{sys}}}{\partial \mathbf{r}_{\alpha}^{z}} &= -\frac{\partial E^{\mathrm{KS}}}{\partial \mathbf{r}_{\alpha}^{z}} - \frac{\partial}{\partial \mathbf{r}_{\alpha}^{z}}\int \rho'_{\mathrm{QM}}(\mathbf{r}) V_{\mathrm{ext}}^{\mathrm{QM/MM}}(\mathbf{r}) d\mathbf{r} - \frac{\partial E_{\mathrm{NE}}^{\mathrm{sys}}}{\partial \mathbf{r}_{\alpha}^{z}} \\ &= \mathbf{F}_{\alpha}^{z,\mathrm{KS}} + \mathbf{F}_{\alpha}^{z,\mathrm{QM/MM}} + \mathbf{F}_{\alpha}^{z,\mathrm{NE}}\end{aligned} \tag{57}$$

where the total force splits up into three individual components corresponding the KS–DFT forces, the PE-QM/MM forces, and forces due to the non-electrostatic terms (see Section 4.4). The first two terms on the right hand side depend on the particular KS-DFT implementation, and for this implementation in GPAW the analytical forces due to the KS term are given elsewhere.[99,100] The contributions due to the QM/MM coupling are given explicitly in the Supporting Information.

To evaluate the atomic forces on a MM COM site i, a term corresponding to the first

term on the right hand side of equation (56) needs to be evaluated

$$\begin{aligned}\mathbf{F}_{\alpha}^{i} = -\frac{\partial E_{\mathrm{tot}}^{\mathrm{sys}}}{\partial \mathbf{r}_{\alpha}^{i}} =& -\frac{\partial E_{\mathrm{ele+ind}}^{\mathrm{MM}}}{\partial \mathbf{r}_{\alpha}^{i}} - \frac{\partial}{\partial \mathbf{r}_{\alpha}^{i}} \int \rho_{\mathrm{QM}}'(\mathbf{r}) V_{\mathrm{ext}}^{\mathrm{QM/MM}}(\mathbf{r}) d\mathbf{r} \\ & - \frac{\partial}{\partial \mathbf{r}_{\alpha}^{i}} \left(E_{\mathrm{self}}^{\mathrm{MM},i} - E_{\mathrm{self}}^{\mathrm{MM},i\mathrm{QM}} \right) - \frac{\partial E^{\mathrm{NE}}}{\partial \mathbf{r}_{\alpha}^{i}} \\ =& \mathbf{F}_{\alpha}^{i,\mathrm{MM}} + \mathbf{F}_{\alpha}^{i,\mathrm{QM/MM}} + \mathbf{F}_{\alpha}^{i,\mathrm{NE}} \end{aligned} \tag{58}$$

with three components corresponding to MM, QM/MM interface and non-electrostatics terms. Note that there is no term associated with the self-energy. The proof is presented in the Supporting Information.

Using the identity in equation (52), the second term on the right hand side is

$$\begin{aligned}\mathbf{F}_{\alpha}^{i,\mathrm{QM/MM}} =& -\frac{\partial}{\partial \mathbf{r}_{\alpha}^{i}} \sum_{j}^{n_{\mathrm{MM}}} \Bigg(\left(\mu_{\beta}^{j'} + \frac{1}{2} \Delta \mu_{\beta}^{j\mathrm{QM}} \right) \mathbf{V}_{\beta}^{j\mathrm{QM}} + \frac{1}{3} \left(\theta_{\beta\gamma}^{j'} + \frac{1}{2} \Delta \theta_{\beta\gamma}^{j\mathrm{QM}} \right) \mathbf{V}_{\beta\gamma}^{j\mathrm{QM}} \\ & + \frac{1}{15} \Omega_{\beta\gamma\delta}^{j} \mathbf{V}_{\beta\gamma\delta}^{j\mathrm{QM}} + \frac{1}{105} \Phi_{\beta\gamma\delta\epsilon}^{j} \mathbf{V}_{\beta\gamma\delta\epsilon}^{j\mathrm{QM}} \Bigg) \\ =& -\left(\mu_{\beta}^{i'} + \frac{1}{2} \Delta \mu_{\beta}^{i\mathrm{QM}} \right) \mathbf{V}_{\alpha\beta}^{i\mathrm{QM}} - \frac{1}{3} \left(\theta_{\beta\gamma}^{i'} + \frac{1}{2} \Delta \theta_{\beta\gamma}^{i\mathrm{QM}} \right) \mathbf{V}_{\alpha\beta\gamma}^{i\mathrm{QM}} \\ & - \frac{1}{15} \Omega_{\beta\gamma\delta}^{i} \mathbf{V}_{\alpha\beta\gamma\delta}^{i\mathrm{QM}} - \frac{1}{105} \Phi_{\beta\gamma\delta\epsilon}^{i} \mathbf{V}_{\alpha\beta\gamma\delta\epsilon}^{i\mathrm{QM}} \end{aligned} \tag{59}$$

where the third to fifth order gradients of the electrostatic potential due to the QM charge

density are given by

$$\mathbf{V}^{i\mathrm{QM}}_{\alpha\beta\gamma} = \int \rho_{\mathrm{QM}}(\mathbf{r}) \left(-\frac{15}{r^7}\mathbf{r}_\alpha\mathbf{r}_\beta\mathbf{r}_\gamma + \frac{3}{r^5}\left(\mathbf{r}_\alpha\delta_{\beta\gamma} + \mathbf{r}_\beta\delta_{\alpha\gamma} + \mathbf{r}_\gamma\delta_{\alpha\beta}\right)\right) d\mathbf{r} \tag{60}$$

$$\begin{aligned}\mathbf{V}^{i\mathrm{QM}}_{\alpha\beta\gamma\delta} = \int \rho_{\mathrm{QM}}(\mathbf{r}) \Bigg(& \frac{105}{r^9}\mathbf{r}_\alpha\mathbf{r}_\beta\mathbf{r}_\gamma\mathbf{r}_\delta - \frac{15}{r^7}\left\{\begin{array}{l} \mathbf{r}_\alpha\mathbf{r}_\beta\delta_{\gamma\delta} + \mathbf{r}_\alpha\mathbf{r}_\gamma\delta_{\beta\delta} + \mathbf{r}_\alpha\mathbf{r}_\delta\delta_{\beta\gamma} \\ + \mathbf{r}_\beta\mathbf{r}_\gamma\delta_{\alpha\delta} + \mathbf{r}_\beta\mathbf{r}_\delta\delta_{\alpha\gamma} + \mathbf{r}_\gamma\mathbf{r}_\delta\delta_{\alpha\beta}\end{array}\right\} \\ & + \frac{3}{r^5}\left(\delta_{\alpha\beta}\delta_{\gamma\delta} + \delta_{\alpha\gamma}\delta_{\beta\delta} + \delta_{\alpha\delta}\delta_{\beta\gamma}\right)\Bigg) d\mathbf{r}\end{aligned} \tag{61}$$

$$\begin{aligned}\mathbf{V}^{i\mathrm{QM}}_{\alpha\beta\gamma\delta\epsilon} = \int \rho_{\mathrm{QM}}(\mathbf{r}) \Bigg(& -\frac{945}{r^{11}}\mathbf{r}_\alpha\mathbf{r}_\beta\mathbf{r}_\gamma\mathbf{r}_\delta\mathbf{r}_\epsilon \\ & + \frac{105}{r^9}\left\{\begin{array}{l} \mathbf{r}_\alpha\mathbf{r}_\beta\mathbf{r}_\gamma\delta_{\delta\epsilon} + \mathbf{r}_\alpha\mathbf{r}_\beta\mathbf{r}_\delta\delta_{\gamma\epsilon} + \mathbf{r}_\alpha\mathbf{r}_\beta\mathbf{r}_\epsilon\delta_{\gamma\delta} + \mathbf{r}_\alpha\mathbf{r}_\gamma\mathbf{r}_\delta\delta_{\beta\epsilon} + \mathbf{r}_\alpha\mathbf{r}_\gamma\mathbf{r}_\epsilon\delta_{\beta\delta} \\ + \mathbf{r}_\alpha\mathbf{r}_\delta\mathbf{r}_\epsilon\delta_{\beta\gamma} + \mathbf{r}_\beta\mathbf{r}_\gamma\mathbf{r}_\delta\delta_{\alpha\epsilon} + \mathbf{r}_\beta\mathbf{r}_\gamma\mathbf{r}_\epsilon\delta_{\alpha\delta} + \mathbf{r}_\beta\mathbf{r}_\delta\mathbf{r}_\epsilon\delta_{\alpha\gamma} + \mathbf{r}_\gamma\mathbf{r}_\delta\mathbf{r}_\epsilon\delta_{\alpha\beta}\end{array}\right\} \\ & - \frac{15}{r^7}\left\{\begin{array}{l} \mathbf{r}_\alpha\delta_{\beta\gamma}\delta_{\delta\epsilon} + \mathbf{r}_\alpha\delta_{\beta\gamma}\delta_{\gamma\epsilon} + \mathbf{r}_\alpha\delta_{\beta\epsilon}\delta_{\gamma\delta} + \mathbf{r}_\beta\delta_{\alpha\gamma}\delta_{\delta\epsilon} + \mathbf{r}_\beta\delta_{\alpha\delta}\delta_{\gamma\epsilon} \\ + \mathbf{r}_\beta\delta_{\alpha\epsilon}\delta_{\gamma\delta} + \mathbf{r}_\gamma\delta_{\alpha\beta}\delta_{\gamma\epsilon} + \mathbf{r}_\gamma\delta_{\alpha\delta}\delta_{\beta\epsilon} + \mathbf{r}_\gamma\delta_{\alpha\epsilon}\delta_{\beta\delta} + \mathbf{r}_\delta\delta_{\alpha\beta}\delta_{\gamma\epsilon} \\ + \mathbf{r}_\delta\delta_{\alpha\gamma}\delta_{\beta\epsilon} + \mathbf{r}_\delta\delta_{\alpha\epsilon}\delta_{\beta\gamma} + \mathbf{r}_\epsilon\delta_{\alpha\beta}\delta_{\gamma\delta} + \mathbf{r}_\epsilon\delta_{\alpha\gamma}\delta_{\beta\delta} + \mathbf{r}_\epsilon\delta_{\alpha\delta}\delta_{\beta\gamma}\end{array}\right\}\Bigg) d\mathbf{r}\end{aligned} \tag{62}$$

The MM COM force terms due to both the MM-MM interactions and QM/MM interface interactions have two contributions; the first contribution is an explicit term corresponding to derivatives of the external field and gradients thereof (as shown above); and the second term is an implicit one depending on the rotation used to define the MM moment tensors in the global reference frame relative to the local frame for each MM COM site. The implicit terms (often referred to as the torque) are outside the scope of this work, as well as redistribution of the COM forces onto the atomic positions of each MM water molecule. This is described in the original SCME work.[22,23] A brief summary on the derivatives of commonly used local frames and associated rotation operations can be found elsewhere.[102]

4.4 Non-electrostatic Interactions

The non-electrostatic interactions in the QM/MM interface are split into two contributions, repulsion and dispersion

$$E_{\mathrm{NE}} = E_{\mathrm{rep}} + E_{\mathrm{disp}} \tag{63}$$

The dispersion and repulsion interactions between QM and MM molecules are taken to be the same as in SCME[22,23] with a physically justifiable change to the phenomenological repulsion function as discussed below. The interactions are centered on the oxygen atoms of the QM and MM molecules, so the distance r refers to the O-O distance, $r = |\mathbf{r}_{\mathrm{O}}^{i} - \mathbf{r}_{\mathrm{O}}^{j}|$, between molecular pairs ij. $n_{\mathrm{tot}} = n_{\mathrm{QM}} + n_{\mathrm{MM}}$, where n_{QM} refers to the number of QM oxygen in the combined system.

The dispersion energy is

$$E_{\mathrm{disp}} = -\sum_{i}^{n_{\mathrm{tot}}} \sum_{j<i}^{n_{\mathrm{tot}}} \left(\frac{C_6}{r^6} g_6(r) + \frac{C_8}{r^8} g_8(r) + \frac{C_{10}}{r^{10}} g_{10}(r) \right) \tag{64}$$

with isotropic coefficients up to tenth order. This series corresponds to second-order perturbation correlation between dipole–dipole, dipole–quadrupole, and quadrupole–quadrupole (or dipole–octupole) polarizabilities.[57] The coefficients C_6, C_8 and C_{10} are taken from Wormer and Hettema.[103] At short range, the interaction is smoothly switched off with a Tang-Toennies[104] type damping function

$$g_m(r) = 1 - e^{-\tau_d r} \sum_{k=0}^{m} \frac{(\tau_d r)^k}{k!} \tag{65}$$

to avoid divergence of the series expansion. The function depends on the damping parameter, τ_d, approximately corresponding to the inverse decay length of the charge density of an H_2O molecule.

A modified Born-Mayer potential is used for the exchange-repulsion. Before presenting the expression for the QM/MM interface, we recall the original expression in the SCME

potential function[22,23]

$$E_{\text{rep}} = \sum_{i<j}^{n_{\text{MM}}} A(1 + B(\xi_i) + B(\xi_j)) r_{ij}^{-b} e^{-cr_{ij}} \tag{66}$$

where the function B depends on the local density of molecules, ξ_i, and has the following values

$$B(\xi_i) = \begin{cases} 0 \text{ if } \xi_i \leq 1600 \\ \sum_{n=0}^{5} a_n \xi_i^n \text{ if } 1600 < \xi_i \leq 8000 \\ 0.0875 \text{ if } 8000 < \xi_i \end{cases} \tag{67}$$

while the molecular densities are evaluated as

$$\xi_i = C \sum_{j \neq i}^{n_{\text{MM}}} \frac{e^{-dr_{ij}}}{r_{ij}^3}. \tag{68}$$

Since the repulsion potential has its interaction centers in the molecular centers of mass, the resulting repulsion interaction is isotropic. The molecular density term in equation (68) provides a phenomenological way of taking into account the change in exchange repulsion due to the polarization where the mixing of excited electronic states changes the decay length of the electron density.[22,23] At the QM/MM interface, the MM-MM repulsion terms then need to be evaluated with molecular densities evaluated over the entire system, so (68) becomes

$$\xi_i = C \sum_{j \neq i}^{n_{\text{tot}}} \frac{e^{-dr_{ij}}}{r_{ij}^3} \tag{69}$$

For a QM molecule, there is no molecular density dependent term associated with it, since its charge density responds and adjusts to the environment by virtue of the self-consistent equations. In particular, the external potential at each MM site due to the QM charge density is modified by interaction tensor damping functions, which effectively introduce short-range repulsion. This effective repulsion becomes more pronounced if the MM site moves closer to

a QM molecule, or the electron density of the QM region polarizes (decay length changes) and overlaps more with an MM site. See the Supplementary Information for more details. So while the MM-MM repulsion terms are still evaluated from equation (66), the QM-MM terms are evaluated from

$$E_{\text{rep}} = \sum_{i}^{n_{\text{MM}}} \sum_{j}^{n_{\text{QM}}} A(1 + B(\xi_i)) r_{ij}^{-b} e^{-cr_{ij}}. \tag{70}$$

Note that the sum over the molecules in the MM subsystem still has the molecular density term associated with it.

5 Implementation and Computational Details

In order to implement the mutually polarizable coupling scheme efficiently, the QM-SCF cycle is modified in such a way as to simultaneously solve for the KS ground state charge density and the external potential. A procedure where the de-coupled ground state charge density is first obtained and then the two subsystems coupled is not efficient. Instead, the external potential at each QM-SCF step is updated. However, since the octu- and hexadecapoles are static these higher order terms only need to be evaluated once at the start of the cycle, while the first and second order terms corresponding to the induced dipole and induced quadrupole need to be evaluated at every iteration. We make use of the highly paralellizable grid-based nature of GPAW as well as permutation symmetry of the tensors to make the interface efficient. See the Supplementary Information for more details. The coupling scheme is presented schematically in figure 1.

The position of MM sites in the global reference frame will possibly place them within the QM grid space, and hence close to or on top of a grid-point. Potentials and gradients in terms of the interaction tensors will diverge resulting in what is commonly known as the polarization catastrophe – as coined by Thole in his seminal paper.[105] This is the polarization analog of the charge spill-out effect encountered in electrostatic embedding QM/MM models. In order

to avoid the catastrophe, we have included tensor damping functions based on Gaussians.[106] Effectively, the functions smear out the point moments and describe a screened interaction as the distributions start to overlap. See the Supplementary Information for more details.

The DFT calculations are carried out within the generalized gradient approximation, in particular the PBE[107] and BLYP[108,109] functionals. The calculations were carried out using a real space grid with mesh size of 0.15 Å. The dimer calculations were started from the optimized SCME dimer geometry and the interaction energy calculated by subtracting the DFT monomer energy from the total energy of the system (as the SCME molecules have by definition no ground state energy). Possible energetic irregularities arising from the coarse real-space grid were avoided by making sure that the QM molecule position relative to the grid did not change for any of the coordinates scanned.

A accurate coupling scheme does not introduce artifacts that affect the total energy and atomic forces.[13] The artifacts can be divided into two classes: (1) The energy of the coupled QM/MM system is smaller or larger than the energy obtained from using either the pure QM or pure MM model. (2) Different sub-divisions into QM and MM regions give substantially different energy.

6 H_2O Dimer

Studies of the water dimer interaction energy reveal how the hydrogen bonding is represented in the multiscale framework; the roles of the H-donating and the H-accepting water molecules in a hydrogen bond are different. This provides a good test case for assessing the second type of artifacts described above. Throughout this section we let 'xc/SCME' denote the case where the hydrogen donating molecule is in the QM region described either with the PBE or BLYP xc-functional, while the hydrogen accepting molecule is described with SCME, and vice versa for 'SCME/xc'.

Figure 2 shows various cuts through the potential energy surface of the H_2O dimer. The

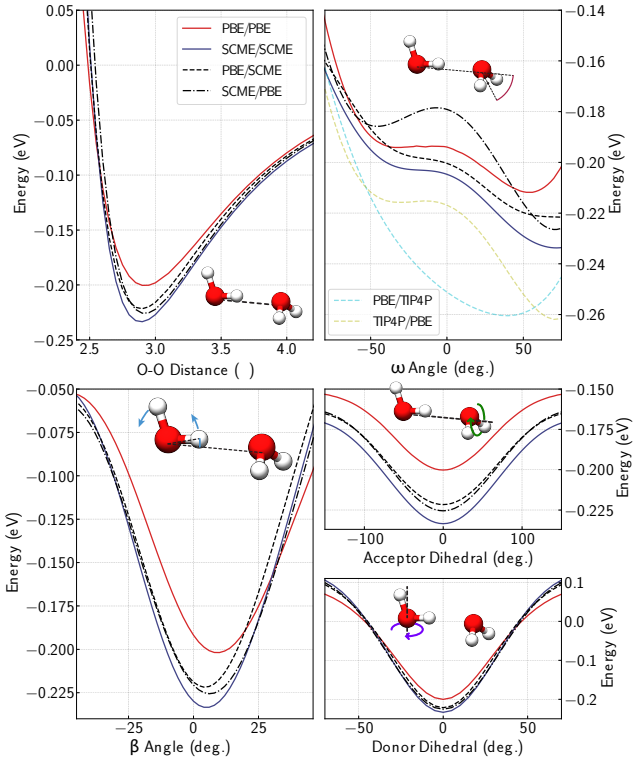

Figure 2: Scans through the potential energy surface of the H_2O dimer, calculated with PBE (labeled PBE/PBE), with SCME (labeled SCME/SCME) and the two possible QM/MM configurations (labeled PBE/SCME and SCME/PBE). The curves are generated by taking the relaxed SCME dimer geometry and scanning the variable indicated on the horizontal axes. Only the degree of freedom defining each scan is varied without structural relaxation.

top left graph shows the interaction energy as a function of distance between the two O-atoms. For almost all of the binding region, both PBE/SCME and SCME/PBE curves are within the energy range spanned by the pure PBE and pure SCME results. Furthermore, the two QM/MM results are similar, differing by only 3.6% with respect to the average at an O-O distance of 2.9 Å. The position of the minimum varies slightly. This is due to the importance of the electronic density associated with the lone pair on the H-acceptor molecule for defining the hydrogen bonding geometry. When this electron density is obtained from DFT, the resulting SCME/PBE curve minimum is closer to the pure PBE result. On the other hand, when the H-acceptor molecule is modeled with SCME, the PBE/SCME curve minimum is at a O-O distance closer to the pure SCME result. At short distances where the repulsion is strong, the QM/MM results are too repulsive. This is ascribed to shortcomings in our QM/MM adaptation of the of the repulsion model in the SCME potential function[58] and will be addressed in future work.

The top right graph of figure 2 shows the acceptor wagging-angle ω. The scan corresponds to a transference of the hydrogen bond from one lone pair on the O-atom of the H-acceptor molecule to another and should result in a shoulder on the curve. Both QM/MM configurations reproduce the behaviour of this subtle but important interaction, whereas QM/MM models based on simple point charge MM potentials miss this feature altogether when the H-acceptor molecule is in the MM region, because the two lone pairs are reduced to a single point of partial charge. The shoulder is somewhat exaggerated in the SCME/PBE curve, corresponding to the slightly deeper O-O distance well. This could be a result of the effective polarizability produced by GGA functionals being too high,[110] which results in a QM density producing a stronger field which thus again more strongly induces the MM moments. However, the two coupled curves still lie within the two single-model limits.

The results on the interaction energy of rotating the H-donor molecule around its oxygen and the normal spanned from there to the hydrogen atoms produces the β-angle plot, in the bottom left region of figure 2. Here, again, the PBE/SCME minimum is shifted towards the

pure SCME minimum and SCME/PBE towards PBE, in accordance with the analysis of the O-O distance cut. Lastly, the two bottom right plots of the figure show consistent results between the QM/MM results and the single-scale models for the two dihedral angles.

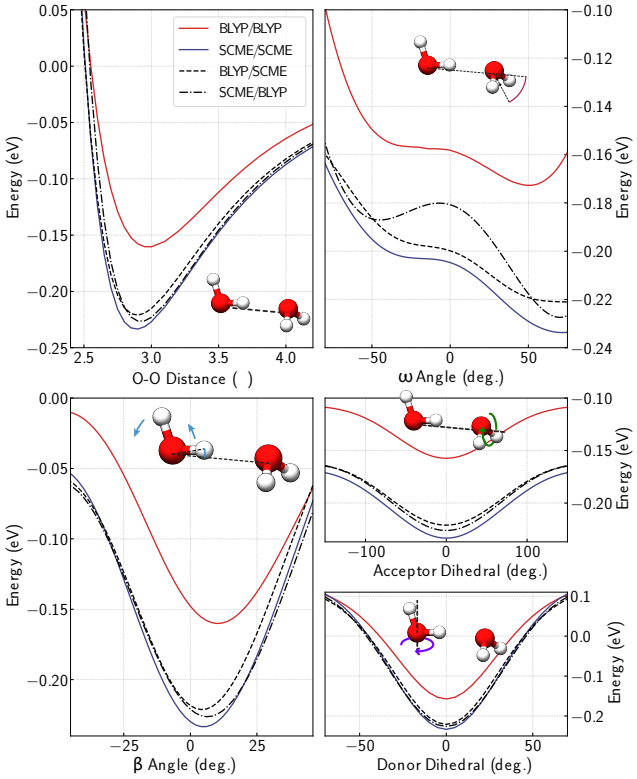

Figure 3: Potential energy curves for the water dimer, this time using BLYP.

Figure 3 shows the analogous results when the BLYP functional is used for the QM region instead of the PBE functional. In general, the difference between the two single-scale descriptions are then larger, but the QM/MM curves are similar to those obtained using the PBE functional, showing the same trend depending on with model is used for the lone pairs on the H-acceptor molecule.

It is well known that GGA functionals have several, but different, shortcomings when describing liquid water.[110] This difference can be traced to the different exchange-enhancement factor in the functionals. The two functionals compared here represent extrema in this regard – BLYP being strongly repulsive in the exchange-overlap region (i.e. the region where electron densities between two water molecules overlap), whereas PBE is weakly repulsive.

This results in clear differences in the binding energy curves of the H_2O dimer, for example. However, the two functionals give similar static and induced moments for an H_2O molecule, and as a result their electrostatic coupling to the MM multipole moment tensors are near identical for the dimer, further supporting the validity of the self-consistent electrostatic coupling scheme presented here. Another shortcoming of most commonly used GGA functionals is the lack of dispersion interaction.

6.1 Tests of forces

An essential part of atomic scale simulations is to have accurate evaluation of the forces acting on the atoms, for efficient structure relaxation, energy barrier calculations and dynamics simulations. In order to test and verify the analytical expressions for the forces presented in section 4.3, the forces acting on the molecules in the H_2O dimer are compared with numerical, finite difference calculations. Figure 4 shows the results. The graph shows the force acting

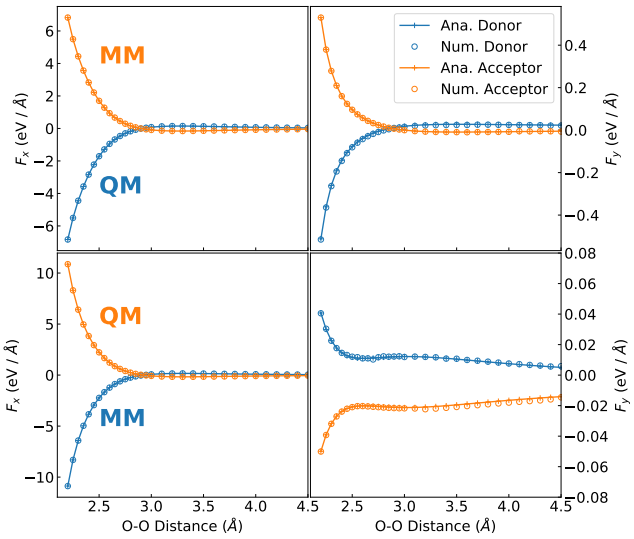

Figure 4: Test comparing the analytical forces on the center of mass to the numerical forces. Top row: the QM/MM configuration with the H-donor molecule being described with PBE. Bottom row: MM/QM, where the H-acceptor molecule is described with PBE.

on the center of mass of each one of the H_2O molecules in the dimer, since this is the most direct test within the SCME formalism where there is just a single expansion center per

water molecule. The dimer is placed in the x, y plane, so there is no force in the z direction. The force can subsequently be distributed on the atoms, or the molecules propagated in time using rigid-body dynamics since the torque on each molecule is also calculated.

Figure 5 shows how the analytical forces converge as convergence criteria for the various quantities, such as density, wavefunctions, total energy, dipole, and quadrupole are tightened by an order of magnitude. The difference computed with respect to the value obtained with

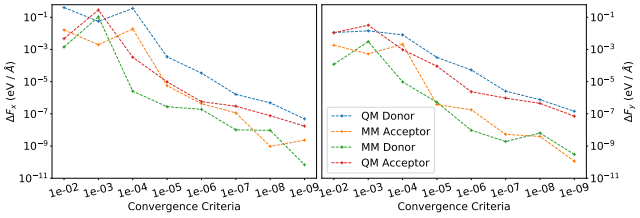

Figure 5: Differences in analytical QM/MM forces of the H_2O dimer in the SCME-optimized geometry. The dimer is oriented in the simulation cell in the same way as in figure 4.

the tightest tolerance, $\Delta F = F - F_{1e-10}$ is shown. As expected, the difference decreases systematically as the convergence criteria are tightened. The difference is largest for the QM molecule due to the coarseness of the real space grid used in the DFT calculation.

7 H_2O Pentamer

Due to the many-body nature of both the repulsion potential and the MM-SCF and QM/MM-SCF induction loops, it is important to test the analytical forces in systems containing more than a single molecule within each subsystem. An incorrect inclusion of the MM self-energy in the QM/MM coupling can lead to inconsistent forces. For example, in the QM/MM dimer the external field at the MM site is due to the QM charge density only, hence the self-energy terms $E_{\text{self}}^{\text{MM},i}$ and $E_{\text{self}}^{\text{MM},i\text{QM}}$ in equation (53) cancel each other out exactly. However, in a system containing more than one MM site such a cancellation will not occur. Furthermore, the total energy expressions in equations (53) and (55) would result in the same numerical forces, but different analytical forces, with the latter expression giving wrong force terms.

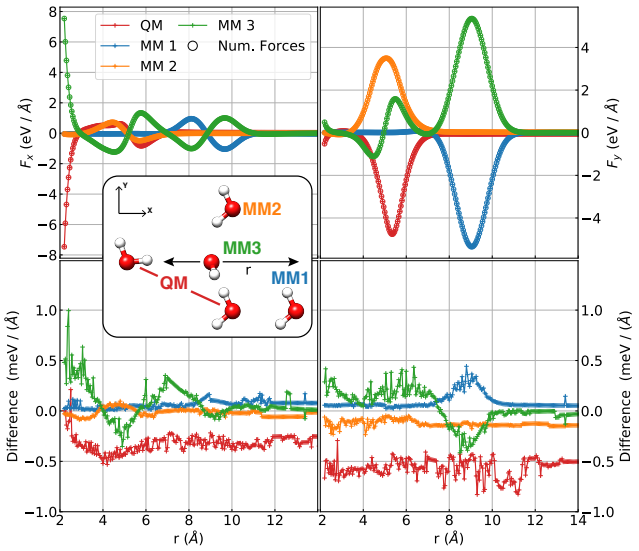

Figure 6: x- and y-components (left and column, respectively) of analytical (lines) and numerical (circles) forces of a test-case using an arficial configuration designed to best probe our coupling (inset). The bottom two plots show the difference in meV / Å between the analytical and numerical forces.

Figure 6 shows the results of a comparison between analytical and numerical forces in a pentamer H_2O cluster in a non-equilibrium geometry chosen to test all aspects of the $B(\xi_i)$ function in equation 70 (see inset in in figure 6). The standard convergence criteria of the DFT calculation is used, while the di- and quadrupole convergence criteria were set to 10^{-9} ea_0 and ea_0^2, respectively, in units of the elementary charge e, and the Bohr radius a_0. The lower two graphs show the difference between the analytical and numerical forces in the two relevant dimensions. The difference is overall largest for the QM subsystem, consistent with the tests on the dimer. For the SCME-molecules a small error becomes evident at extremely short distances due to the use of the Tang-Toennies damping function in the interaction tensors, which is not included in an entirely consistent way in the original SCME formulation.[23] However, this error manifests itself only at O-O distances below ≈2.4 Å, (r in the figure from 8-10 Å) and will not be relevant in, for example, simulations of liquid water at ambient conditions.

Conclusions

A reciprocal polarizable embedding QM/MM energy functional is presented as well as a minimization scheme that couples ground state electronic structure calculations based on KS-DFT and an accurate single center multipole expansion model of H_2O molecules. In the former case the QM/MM interface energy is expressed as an external potential such that the ground state charge density of the QM system is polarized by the MM region. Similarly, the self-consistent induced moments of the MM sites include the external field and field gradient due to the QM charge density. In this way both the QM-SCF or MM-SCF calculations are carried out in the usual way except that the relevant numerical values are passed between the QM and MM parts. The MM model of the H_2O molecules includes static molecular moments up to and including the hexadecapole, as well as induced dipole and quadrupole moments in response to the external field and field gradient. The energy and resulting analytical forces are in agreement with numerical tests for both the dimer and pentamer clusters, the consistency being systematically improved as convergence tolerances in the calculations are reduced.

The PE-QM/MM energy surface for the H_2O dimer is intermediate between the pure QM and pure MM energy surfaces. Furthermore, the two different GGA functionals used, PBE and BLYP, give nearly identical results since they give similar polarizability. The difference between the pure PBE and pure BLYP dimer results is mainly due to the different semi-local part and is most apparent when two water molecular densities start to overlap.[110] In future work, we will explore higher-level exchange-correlation functional approximations and self-interaction corrected energy functional description[111] of the QM region for improved accuracy, as well as an extension of the SCME potential function for flexible H_2O molecules.

Acknowledgement

The authors thank K. T. Wikfeldt, E. R. Batista and J. Öström for helpful discussions. Funding was provided by the Icelandic Research Fund, Grants 174082-051, 174244-051, and 141080-051. The calculations were carried out at the Icelandic High-Performance Computing Center. Furthermore, E.O.J acknowledges additional computer access from Delingr. A.O.D. acknowledges support from Villum Fonden.

References

(1) Warshel, A.; Levitt, M. Theoretical Studies of Enzymatic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249.

(2) Field, M. J.; Bash, P. A.; Karplus, M. A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem.* **1990**, *11*, 700–733.

(3) Bakowies, D.; Thiel, W. Hybrid Models for Combined Quantum Mechanical and Molecular Mechanical Approaches. *J. Phys. Chem.* **1996**, *100*, 10580–10594.

(4) Antes, I.; Thiel, W.; Gao, J. Hybrid Quantum Mechanical and Molecular Mechanical Methods. ACS Symposium Series. 1998; pp 50–65.

(5) Bentzien, J.; Florián, J.; Glennon, T. M.; Warshel, A. *Combined Quantum Mechanical and Molecular Mechanical Methods*; ACS Publications, 1998; Chapter 2, pp 16–34.

(6) Merz Jr, K. M. *Combined Quantum Mechanical and Molecular Mechanical Methods*; ACS Publications, 1998; Chapter 1, pp 2–15.

(7) Mordasini, T. Z.; Thiel, W. Computational Chemistry Column: Combined Quantum Mechanical and Molecular Mechanical Approaches. *CHIMIA* **1998**, *52*, 288–291.

(8) Woo, T.; Margl, P.; Deng, L.; Cavallo, L.; Ziegler, T. *Transition State Modeling for Catalysis*; ACS Publications, 1999; Chapter 14, pp 173–186.

(9) Monard, G.; Merz, K. M. Combined quantum mechanical/molecular mechanical methodologies applied to biomolecular systems. *Acc. Chem. Res.* **1999**, *32*, 904–911.

(10) Hillier, I. H. Chemical reactivity studied by hybrid QM/MM methods. *THEOCHEM* **1999**, *463*, 45–52.

(11) Morokuma, K. New challenges in quantum chemistry: quests for accurate calculations for large molecular systems. *Philos. Trans. A. Math. Phys. Eng. Sci.* **2002**, *360*, 1149–1164.

(12) Gao, J.; Truhlar, D. G. Quantum mechanical methods for enzyme kinetics. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467–505.

(13) Lin, H.; Truhlar, D. G. QM/MM: what have we learned, where are we, and where do we go from here? *Theor. Chem. Acc.* **2006**, *117*, 185.

(14) Senn, H. M.; Thiel, W. QM/MM methods for biomolecular systems. *Angew. Chem. Int. Ed. (English)* **2009**, *48*, 1198–1229.

(15) Kirchner, B.; Vrabec, J. Multiscale Molecular Methods in Applied Chemistry. *Top. Curr. Chem.* **2012**, *307*.

(16) van der Kamp, M. W.; Mulholland, A. J. Combined quantum mechanics/molecular mechanics (QM/MM) methods in computational enzymology. *Biochemistry* **2013**, *52*, 2708–2728.

(17) Brunk, E.; Rothlisberger, U. Mixed Quantum Mechanical/Molecular Mechanical Molecular Dynamics Simulations of Biological Systems in Ground and Electronically Excited States. *Chem. Rev.* **2015**, *115*, 6217–6263.

(18) Pezeshki, S.; Lin, H. Recent Advances in the Molecular Simulation of Chemical Reactions. *Mol. Sim.* **2015**, *41*, 168–189.

(19) Zheng, M.; Waller, M. P. Adaptive quantum mechanics/molecular mechanics methods. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2016**, *6*, 369–385.

(20) Sousa, S. F.; Ribeiro, A. J. M.; Neves, R. P. P.; Brs, N. F.; Cerqueira, N. M. F. S. A.; Fernandes, P. A.; Ramos, M. J. Application of quantum mechanics/molecular mechanics methods in the study of enzymatic reaction mechanisms. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2017**, *7*, e1281.

(21) Zhang, Y.-J.; Khorshidi, A.; Kastlunger, G.; Peterson, A. A. The potential for machine learning in hybrid QM/MM calculations. *J. Chem. Phys* **2018**, *148*, 241740.

(22) Batista, E. *Development of a New Water-Water Interaction Potential and Applications to Molecular Processes in Ice*; University of Washington, 1999.

(23) Wikfeldt, K. T.; Batista, E. R.; Vila, F. D.; Jónsson, H. A Transferable H_2O Interaction Potential Based on a Single Center Multipole Expansion: SCME. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16542.

(24) Dohn, A. O.; Jónsson, E. O.; Jónsson, H. Reciprocal Polarizable Embedding with a Transferable H_2O Potential Function II: Application to $(H_2O)_n$ Clusters & Liquid Water. *Submitted to JCTC simultaneously with the current manuscript* **2019**,

(25) Smirnov, I. V.; Golovin, A. V.; Chatziefthimiou, S. D.; Stepanova, A. V.; Peng, Y.; Zolotareva, O. I.; Belogurov, A. A.; Kurkova, I. N.; Ponomarenko, N. A.; Wilmanns, M.; Blackburn, G. M.; Gabibov, A. G.; Lerner, R. A. Robotic QM/MM-driven maturation of antibody combining sites. *Sci. Adv.* **2016**, *2*, e1501695.

(26) Barends, T. R. M.; Foucar, L.; Ardevol, A.; Nass, K.; Aquila, A.; Botha, S.; Doak, R. B.; Falahati, K.; Hartmann, E.; Hilpert, M.; Heinz, M.; Hoffmann, M. C.;

Köfinger, J.; Koglin, J. E.; Kovacsova, G.; Liang, M.; Milathianaki, D.; Lemke, H. T.; Reinstein, J.; Roome, C. M.; Shoeman, R. L.; Williams, G. J.; Burghardt, I.; Hummer, G.; Boutet, S.; Schlichting, I. Direct observation of ultrafast collective motions in CO myoglobin upon ligand dissociation. *Science* **2015**, *350*, 445–450.

(27) Senthilkumar, K.; Mujika, J. I.; Ranaghan, K. E.; Manby, F. R.; Mulholland, A. J.; Harvey, J. N. Analysis of polarization in QM/MM modelling of biologically relevant hydrogen bonds. *J. Royal Soc. Interface* **2008**, *5*, 207–216.

(28) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. M. Electrostatic basis for enzyme catalysis. *Chem. Rev.* **2006**, *106*, 3210–3235.

(29) Zheng, F.; Xue, L.; Hou, S.; Liu, J.; Zhan, M.; Yang, W.; Zhan, C.-G. A highly efficient cocaine-detoxifying enzyme obtained by computational design. *Nat. Commun.* **2014**, *5*, 3457.

(30) Knorr, J.; Sokkar, P.; Schott, S.; Costa, P.; Thiel, W.; Sander, W.; Sanchez-Garcia, E.; Nuernberger, P. Competitive solvent-molecule interactions govern primary processes of diphenylcarbene in solvent mixtures. *Nat. Commun.* **2016**, *7*, 12968.

(31) Pham, V.-T.; Penfold, T. J.; van der Veen, R. M.; Lima, F.; Nahhas, A. E.; Johnson, S. L.; Beaud, P.; Abela, R.; Bressler, C.; Tavernelli, I.; Milne, C. J.; Chergui, M. Probing the Transition from Hydrophilic to Hydrophobic Solvation with Atomic Scale Resolution. *J. Am. Chem. Soc.* **2011**, *133*, 12740–12748.

(32) Dohn, A. O.; Jónsson, E. O.; Kjær, K. S.; B. van Driel, T.; Nielsen, M. M.; Jacobsen, K. W.; Henriksen, N. E.; Møller, K. B. Direct Dynamics Studies of a Binuclear Metal Complex in Solution: The Interplay Between Vibrational Relaxation, Coherence, and Solvent Effects. *J. Phys. Chem. Letters* **2014**, *5*, 2414–2418.

(33) Dohn, A. O.; Kjær, K. S.; Harlang, T. B.; Canton, S. E.; Nielsen, M. M.; Møller, K. B.

Electron Transfer and Solvent-Mediated Electronic Localization in Molecular Photocatalysis. *Inorg. Chem.* **2016**, *55*, 10637–10644.

(34) Levi, G.; Pápai, M.; Henriksen, N. E.; Dohn, A. O.; Møller, K. B. Solution Structure and Ultrafast Vibrational Relaxation of the PtPOP Complex Revealed by ΔSCF-QM/MM Direct Dynamics Simulations. *J. Chem. Phys. C* **2018**, *122*, 7100–7119.

(35) Hörmann, N. G.; Jäckle, M.; Gossenberger, F.; Roman, T.; Forster-Tonigold, K.; Naderian, M.; Sakong, S.; Groß, A. Some challenges in the first-principles modeling of structures and processes in electrochemical energy storage and transfer. *J. Power Sources* **2015**, *275*, 531–538.

(36) Calculations of Product Selectivity in Electrochemical CO2 Reduction. *ACS Catal.* **2018**, *8*, 5240–5249.

(37) Cheng, T.; Xiao, H.; Goddard, W. A. Free-Energy Barriers and Reaction Mechanisms for the Electrochemical Reduction of CO on the Cu(100) Surface, Including Multiple Layers of Explicit Solvent at pH 0. *TJ. Phys. Chem. Lett.* **2015**, *6*, 4767–4773.

(38) Shi, C.; O'Grady, C. P.; Peterson, A.; Hansen, H.; Nørskov, J. K. Modeling CO2 reduction on Pt(111). *Phys. Chem. Chem. Phys.* **2013**, *15*, 7114–7122.

(39) Skúlason, E.; Jónsson, H. Atomic scale simulations of heterogeneous electrocatalysis: recent advances. *Adv.Phys. X* **2017**, *2*, 481–495.

(40) Nielsen, M.; Björketun, M. E.; Hansen, M. H.; Rossmeisl, J. Towards first principles modeling of electrochemical electrodeelectrolyte interfaces. *Surf. Sci* **2015**, *631*, 2–7.

(41) Vassilev, P.; van Santen, R. A.; Koper, M. T. M. Ab initio studies of a water layer at transition metal surfaces. *J. Chem. Phys* **2005**, *122*, 054701.

(42) Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **1964**, *136*, B864.

(43) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, 1133–1138.

(44) Becke, A. D. Perspective: Fifty years of density-functional theory in chemical physics. *J. Chem. Phys* **2014**, *140*, 18A301.

(45) Hu, H.; Yang, W. Free energies of chemical reactions in solution and in enzymes with ab initio quantum mechanics/molecular mechanics methods. *Annu. Rev. Phys. Chem.* **2008**, *59*, 573–601.

(46) Takahashi, H.; Hori, T.; Hashimoto, H.; Nitta, T. *J. Comp. Chem.* **2001**, *22*, 1252.

(47) Crespo, A.; Scherlis, D. A.; Martí, M. A.; Ordejón, P.; Roitberg, A. E.; Estrin, D. A. *J. Phys. Chem. B* **2003**, *107*, 13728.

(48) de Visser, S. P. Density functional theory (DFT) and combined quantum mechanical/molecular mechanics (QM/MM) studies on the oxygen activation step in nitric oxide synthase enzymes. *Biochem. Soc. Trans.* **2009**, *37*, 373–377.

(49) Dohn, A. O.; Jónsson, E. Ö.; Levi, G.; Mortensen, J. J.; Lopez-Acevedo, O.; Thygesen, K. S.; Jacobsen, K. W.; Ulstrup, J.; Henriksen, N. E.; Møller, K. B.; Jónsson, H. Grid-Based Projector Augmented Wave (GPAW) Implementation of Quantum Mechanics/Molecular Mechanics (QM/MM) Electrostatic Embedding and Application to a Solvated Diplatinum Complex. *J. Chem. Theory Comput.* **2017**, *13*, 6010–6022.

(50) Dohn, A. O.; Selli, D.; Fazio, G.; Ferraro, L.; Mortensen, J.; Civalleri, B.; Valentin, C. D. Interfacing CRYSTAL/AMBER to Optimize QM/MM Lennard–Jones Parameters for Water and to Study Solvation of TiO2 Nanoparticles. *Molecules* **2018**, *23*, 2958.

(51) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L.

Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(52) Jorgensen, W. L. Quantum and statistical mechanical studies of liquids. 10. Transferable intermolecular potential functions for water, alcohols, and ethers. Application to liquid water. *J. Am. Chem. Soc.* **1981**, *103*, 335–340.

(53) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an Improved Four-Site Water Model for Biomolecular Simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120*, 9665.

(54) Zielkiewicz, J. Structural properties of water: Comparison of the SPC, SPCE, TIP4P, and TIP5P models of water. *J. Chem. Phys* **2005**, *123*, 104501.

(55) Batista, E. R.; Xantheas, S. S.; Jónsson, H. Molecular Multipole Moments of Water Molecules in Ice Ih. *J. Chem. Phys.* **1998**, *109*, 4546.

(56) Batista, E. R.; Xantheas, S. S.; Jónsson, H. Electric fields in ice and near water clusters. *J. Chem. Phys.* **2000**, *112*, 3285.

(57) Stone, A. *The Theory of Intermolecular Forces*; The Theory of Intermolecular Forces; OUP Oxford, 2013.

(58) Cisneros, G. A.; Wikfeldt, K. T.; Ojamäe, L.; Lu, J.; Xu, Y.; Torabifard, H.; Bartók, A. P.; Csányi, G.; Molinero, V.; Paesani, F. Modeling Molecular Interactions in Water: From Pairwise to Many-Body Potential Energy Functions. *Chem. Rev.* **2016**, *116*, 7501–7528.

(59) Yu, H.; Van Gunsteren, W. F. Accounting for polarization in molecular simulation. *Comput. Phys. Commun.* **2005**, *172*, 69–85.

(60) Lopes, P. E.; Roux, B.; MacKerell, A. D. Molecular modeling and dynamics studies

with explicit inclusion of electronic polarizability: theory and applications. *Theor. Chem. Acc.* **2009**, *124*, 11–28.

(61) Thompson, M. A.; Schenter, G. K. Excited states of the bacteriochlorophyll b dimer of Rhodopseudomonas viridis: a QM/MM study of the photosynthetic reaction center that includes MM polarization. *J. Phys. Chem.* **1995**, *99*, 6374–6386.

(62) Thompson, M. A. QM/MMpol: A Consistent Model for Solute/Solvent Polarization. Application to the Aqueous Solvation and Spectroscopy of Formaldehyde, Acetaldehyde, and Acetone. *J. Phys. Chem.* **1996**, *100*, 14492–14507.

(63) Bryce, R. A.; Buesnel, R.; Hillier, I. H.; Burton, N. A. A solvation model using a hybrid quantum mechanical/molecular mechanical potential with fluctuating solvent charges. *Chemical Physics Letters* **1997**, *279*, 367 – 371.

(64) Lipparini, F.; Barone, V. Polarizable force fields and polarizable continuum model: a fluctuating charges/PCM approach. 1. theory and implementation. *J. Chem. Theory Comput.* **2011**, *7*, 3711–3724.

(65) Boulanger, E.; Thiel, W. Solvent Boundary Potentials for Hybrid QM/MM Computations Using Classical Drude Oscillators: A Fully Polarizable Model. *J. Chem. Theory Comput.* **2012**, *8*, 4527–4538.

(66) Lu, Z.; Zhang, Y. Interfacing ab initio quantum mechanical method with classical Drude osillator polarizable model for molecular dynamics simulation of chemical reactions. *J. Chem. Theory Comput.* **2008**, *4*, 1237–1248.

(67) Thellamurege, N. M.; Si, D.; Cui, F.; Zhu, H.; Lai, R.; Li, H. QuanPol: A full spectrum and seamless QM/MM program. *J. Comput. Chem.* **2013**, *34*, 2816–2833.

(68) Kratz, E. G.; Walker, A. R.; Lagardère, L.; Lipparini, F.; Piquemal, J.-P.; Andrés Cis-

neros, G. LICHEM: A QM/MM program for simulations with multipolar and polarizable force fields. *J. Comput. Chem.* **2016**, *37*, 1019–1029.

(69) Dziedzic, J.; Mao, Y.; Shao, Y.; Ponder, J.; Head-Gordon, T.; Head-Gordon, M.; Skylaris, C.-K. TINKTEP: A fully self-consistent, mutually polarizable QM/MM approach based on the AMOEBA force field. *J. Chem. Phys* **2016**, *145*, 124106.

(70) Gomes, A. S. P.; Jacob, C. R. Quantum-chemical embedding methods for treating local electronic excitations in complex chemical systems. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem* **2012**, *108*, 222–277.

(71) Söderhjelm, P.; Husberg, C.; Strambi, A.; Olivucci, M.; Ryde, U. Protein influence on electronic spectra modeled by multipoles and polarizabilities. *J. Chem. Theory Comput.* **2009**, *5*, 649–658.

(72) Sneskov, K.; Schwabe, T.; Christiansen, O.; Kongsted, J. Scrutinizing the effects of polarization in QM/MM excited state calculations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 18551–18560.

(73) Sneskov, K.; Schwabe, T.; Kongsted, J.; Christiansen, O. The polarizable embedding coupled cluster method. *J. Chem. Phys* **2011**, *134*, 03B608.

(74) Caprasecca, S.; Jurinovich, S.; Viani, L.; Curutchet, C.; Mennucci, B. Geometry optimization in polarizable QM/MM models: the induced dipole formulation. *J. Chem. Theory Comput.* **2014**, *10*, 1588–1598.

(75) Kongsted, J.; Osted, A.; Mikkelsen, K. V.; Christiansen, O. The QM/MM approach for wavefunctions, energies and response functions within self-consistent field and coupled cluster theories. *Mol. Phys.* **2002**, *100*, 1813–1828.

(76) Zeng, Q.; Liang, W. Analytic energy gradient of excited electronic state within

TDDFT/MMpol framework: Benchmark tests and parallel implementation. *J. Chem. Phys* **2015**, *143*, 134104.

(77) Loco, D.; Polack, É.; Caprasecca, S.; Lagardère, L.; Lipparini, F.; Piquemal, J.-P.; Mennucci, B. A QM/MM Approach Using the AMOEBA Polarizable Embedding: From Ground State Energies to Electronic Excitations. *J. Chem. Theor. Comput.* **2016**, *12*, 3654–3661.

(78) Loco, D.; Lagardère, L.; Caprasecca, S.; Lipparini, F.; Mennucci, B.; Piquemal, J.-P. Hybrid QM/MM molecular dynamics with AMOEBA polarizable embedding. *J. Chem. Theory Comput.* **2017**, *13*, 4025–4033.

(79) Jensen, L.; van Duijnen, P. T.; Snijders, J. G. A discrete solvent reaction field model for calculating molecular linear response properties in solution. *J. Chem. Phys* **2003**, *119*, 3800–3809.

(80) Steindal, A. H.; Ruud, K.; Frediani, L.; Aidas, K.; Kongsted, J. Excitation Energies in Solution: The Fully Polarizable QM/MM/PCM Method. *J. Phys. Chem. B* **2011**, *115*, 3027–3037.

(81) Nielsen, C. B.; Christiansen, O.; Mikkelsen, K. V.; Kongsted, J. Density functional self-consistent quantum mechanics/molecular mechanics theory for linear and nonlinear molecular properties: Applications to solvated water and formaldehyde. *J. Chem. Phys* **2007**, *126*, 154112.

(82) Olsen, J. M.; Aidas, K.; Kongsted, J. Excited states in solution through polarizable embedding. *J. Chem. Theory Comput.* **2010**, *6*, 3721–3734.

(83) Lipparini, F.; Cappelli, C.; Barone, V. Linear response theory and electronic transition energies for a fully polarizable QM/classical Hamiltonian. *J. Chem. Theory Comput.* **2012**, *8*, 4153–4165.

(84) Curutchet, C.; Muoz-Losa, A.; Monti, S.; Kongsted, J.; Scholes, G. D.; Mennucci, B. Electronic Energy Transfer in Condensed Phase Studied by a Polarizable QM/MM Model. *J. Chem. Theory Comput.* **2009**, *5*, 1838–1848.

(85) List, N. H.; Olsen, J. M. H.; Kongsted, J. Excited states in large molecular systems through polarizable embedding. *Phys. Chem. Chem. Phys.* **2016**, *18*, 20234–20250.

(86) Schwörer, M.; Breitenfeld, B.; Tröster, P.; Bauer, S.; Lorenzen, K.; Tavan, P.; Mathias, G. Coupling density functional theory to polarizable force fields for efficient and accurate Hamiltonian molecular dynamics simulations. *J. Chem. Phys* **2013**, *138*, 244103.

(87) Curutchet, C.; Muñoz-Losa, A.; Monti, S.; Kongsted, J.; Scholes, G. D.; Mennucci, B. Electronic Energy Transfer in Condensed Phase Studied by a Polarizable QM/MM Model. *J. Chem. Theory Comput.* **2009**, *5*, 1838–1848.

(88) Visscher, K.; Swope, W.; Geerke, D. A QM/MM Derived Polarizable Water Model for Molecular Simulation. *Molecules* **2018**, *23*, 3131.

(89) Hršak, D.; Olsen, J. M. H.; Kongsted, J. Polarizable Density Embedding Coupled Cluster Method. *J. Chem. Theory Comput.* **2018**, acs.jctc.7b01153.

(90) Menger, M. F. S. J.; Caprasecca, S.; Mennucci, B. Excited-State Gradients in Polarizable QM/MM Models: An Induced Dipole Formulation. *J. Chem. Theory Comput.* **2017**, *13*, 3778–3786.

(91) Mao, Y.; Shao, Y.; Dziedzic, J.; Skylaris, C.-K.; Head-Gordon, T.; Head-Gordon, M. Performance of the AMOEBA Water Model in the Vicinity of QM Solutes: A Diagnosis Using Energy Decomposition Analysis. *J. Chem. Theory Comput.* **2017**, *13*, 1963–1979.

(92) Dziedzic, J.; Head-Gordon, T.; Head-Gordon, M.; Skylaris, C.-K. Mutually polarizable QM/MM model with in situ optimized localized basis functions. *J. Chem. Phys* **2019**, *150*, 074103.

(93) Gao, J. Energy components of aqueous solution: insight from hybrid QM/MM simulations using a polarizable solvent model. *J. Comput. Chem.* **1997**, *18*, 1061–1071.

(94) Liu, H.; Jenkins, A. J.; Wildman, A.; Frisch, M. J.; Lipparini, F.; Mennucci, B.; Li, X. Time-Dependent Complete Active Space Embedded in a Polarizable Force Field. *J. Chem. Theory Comput.* **2019**, *15*, 1633–1641.

(95) Nakano, H.; Sato, H. Introducing the mean field approximation to CDFT/MMpol method: Statistically converged equilibrium and nonequilibrium free energy calculation for electron transfer reactions in condensed phases. *J. Chem. Phys* **2017**, *146*, 154101.

(96) Bahn, S. R.; Jacobsen, K. W. An object-oriented scripting interface to a legacy electronic structure code. *Comput. Sci. Eng.* **2002**, *4*, 55.

(97) Larsen, A. H.; Mortensen, J. J.; Blomqvist, J.; Castelli, I. E.; Christensen, R.; Dułak, M.; Friis, J.; Groves, M. N.; Hammer, B.; Hargus, C.; Hermes, E. D.; Jennings, P. C.; Jensen, P. B.; Kermode, J.; Kitchin, J. R.; Kolsbjerg, E. L.; Kubal, J.; Kaasbjerg, K.; Lysgaard, S.; Maronsson, J. B.; Maxson, T.; Olsen, T.; Pastewka, L.; Peterson, A.; Rostgaard, C.; Schiøtz, J.; Schtt, O.; Strange, M.; Thygesen, K. S.; Vegge, T.; Vilhelmsen, L.; Walter, M.; Zeng, Z.; Jacobsen, K. W. The atomic simulation environment—a Python library for working with atoms. *J. Phys. Condens. Matter* **2017**, *29*, 273002.

(98) Blöchl, P. E. Projector Augmented–Wave Method. *Phys. Rev. B* **1994**, *50*, 17953.

(99) Mortensen, J.; Hansen, L.; Jacobsen, K. W. Real-space grid implementation of the projector augmented wave method. *Phys. Rev. B* **2005**, *71*, 035109.

(100) Enkovaara, J.; Rostgaard, C.; Mortensen, J. J.; Chen, J.; Dulak, M.; Ferrighi, L.; Gavnholt, J.; Glinsvad, C.; Haikola, V.; Hansen, H. A.; Kristoffersen, H. H.; Kuisma, M.; Larsen, A. H.; Lehtovaara, L.; Ljungberg, M.; Lopez-Acevedo, O.; Moses, P. G.; Ojanen, J.; Olsen, T.; Petzold, V.; Romero, N. A.; Stausholm-Møller, J.; Strange, M.; Tritsaris, G. A.; Vanin, M.; Walter, M.; Hammer, B.; Häkkinen, H.; Madsen, G. K. H.; Nieminen, R. M.; Nørskov, J. K.; Puska, M.; Rantala, T. T.; Schiøtz, J.; Thygesen, K. S.; Jacobsen, K. W. Electronic structure calculations with GPAW: a real-space implementation of the projector augmented-wave method. *J. Phys. Condens. Matter* **2010**, *22*, 253202.

(101) Stone, A. J. The Induction Energy of a Assembly of Polarizable Molecules. *Chem. Phys. Lett.* **1989**, *155*, 102–110.

(102) Lipparini, F.; Lagardre, L.; Stamm, B.; Cancs, E.; Schnieders, M.; Ren, P.; Maday, Y.; Piquemal, J.-P. Scalable Evaluation of Polarization Energy and Associated Forces in Polarizable Molecular Dynamics: I. Toward Massively Parallel Direct Space Computations. *J. Chem. Theory Comput.* **2014**, *10*, 1638–1651.

(103) Wormer, P. E.; Hettema, H. Many-body perturbation theory of frequency-dependent polarizabilities and van der Waals coefficients: Application to H2O–H2O and Ar–NH3. *J. Chem. Phys* **1992**, *97*, 5592–5606.

(104) Tang, K.; Toennies, J. P. An improved simple model for the van der Waals potential based on universal damping functions for the dispersion coefficients. *J. Chem. Phys* **1984**, *80*, 3726–3741.

(105) Thole, B. Molecular polarizabilities calculated with a modified dipole interaction. *Chem. Phys.* **1981**, *59*, 341 – 350.

(106) Stone, A. J. Electrostatic damping functions and the penetration energy. *J. Phys. Chem. A* **2011**, *115*, 7017–7027.

(107) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(108) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098.

(109) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789.

(110) Gillan, M. J.; Alfè, D.; Michaelides, A. Perspective: How good is DFT for water? *J. Chem. Phys.* **2016**, *144*, 130901.

(111) Lehtola, S.; Jónsson, H. Variational, self-consistent implementation of the PerdewZunger self-interaction correction with complex optimal orbitals. *J. Chem. Theo. Comput.* **2014**, *10*, 5324.

Graphical TOC Entry