Coupling density functional theory to polarizable force fields for efficient and accurate Hamiltonian molecular dynamics simulations

Magnus Schwörer, Benedikt Breitenfeld, Philipp Tröster, Sebastian Bauer, Konstantin Lorenzen, Paul Tavan, and Gerald Mathias

Citation: The Journal of Chemical Physics **138**, 244103 (2013); doi: 10.1063/1.4811292
View online: http://dx.doi.org/10.1063/1.4811292
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/138/24?ver=pdfcov
Published by the AIP Publishing

Coupling density functional theory to polarizable force fields for efficient and accurate Hamiltonian molecular dynamics simulations

Magnus Schwörer, Benedikt Breitenfeld, Philipp Tröster, Sebastian Bauer, Konstantin Lorenzen, Paul Tavan, and Gerald Mathias[a)]
Lehrstuhl für BioMolekulare Optik, Ludwig–Maximilians Universität München, Oettingenstr. 67, 80538 München, Germany

(Received 5 April 2013; accepted 3 June 2013; published online 25 June 2013)

Hybrid molecular dynamics (MD) simulations, in which the forces acting on the atoms are calculated by grid-based density functional theory (DFT) for a solute molecule and by a polarizable molecular mechanics (PMM) force field for a large solvent environment composed of several 10^3–10^5 molecules, pose a challenge. A corresponding computational approach should guarantee energy conservation, exclude artificial distortions of the electron density at the interface between the DFT and PMM fragments, and should treat the long-range electrostatic interactions within the hybrid simulation system in a linearly scaling fashion. Here we describe a corresponding Hamiltonian DFT/(P)MM implementation, which accounts for inducible atomic dipoles of a PMM environment in a joint DFT/PMM self-consistency iteration. The long-range parts of the electrostatics are treated by hierarchically nested fast multipole expansions up to a maximum distance dictated by the minimum image convention of toroidal boundary conditions and, beyond that distance, by a reaction field approach such that the computation scales linearly with the number of PMM atoms. Short-range over-polarization artifacts are excluded by using Gaussian inducible dipoles throughout the system and Gaussian partial charges in the PMM region close to the DFT fragment. The Hamiltonian character, the stability, and efficiency of the implementation are investigated by hybrid DFT/PMM-MD simulations treating one molecule of the water dimer and of bulk water by DFT and the respective remainder by PMM. [http://dx.doi.org/10.1063/1.4811292]

I. INTRODUCTION

In a seminal paper devoted to the study of enzymatic reactions, Warshel and Levitt[1] introduced in 1976 a quantum-classical coupling scheme for a molecule, which is described by quantum mechanics (QM) and is embedded in a condensed phase environment modeled by a polarizable molecular mechanics (PMM) force field. In their abstract, these authors emphasized that the "solvation energy resulting from this polarization is considerable and must be included in any realistic calculation" of molecules in condensed phase.

However, as documented in a recent review on QM/MM methods for biomolecular systems,[2] this advice was subsequently ignored in most applications. Instead so-called standard MM force fields like AMBER,[3] CHARMM,[4] OPLS-AA,[5] or GROMOS[6] were generally applied to the MM part of hybrid simulation systems. These force fields model the electrostatic signatures of molecules or of molecular fragments by static partial charges localized at the atoms and, therefore, can account for the effects of electronic polarization only by the mean field approximation, which is highly questionable for inhomogeneous and non-isotropic biomolecular systems.[7] There are notable exceptions which combined a polarizable force field for the MM fragment with semi-empirical quantum chemistry for the QM fragment.[8–12] Combinations of higher-level QM treatments (density functional theory[13,14] (DFT) or *ab initio* quantum chemistry) with PMM force fields were either restricted to the energetics of static systems,[15–23] to small molecular clusters,[24–29] or describe the dynamics only in parts of the simulation system.[30,31] Other approaches augment DFT atoms with self-consistent polarization terms (SCP-DFT) to correct the deficiencies of the long-range electrostatics and dispersion description within certain exchange-correlation functionals.[32,33]

The development of hybrid methods combining grid-based DFT with non-polarizable MM force fields started with the work of Eichinger *et al.*,[34] which particularly aimed at accurate computations of vibrational spectra of molecules in condensed phase environments from hybrid MD simulations. Since then corresponding applications have demonstrated the power of this approach.[35–37] Subsequently, two further DFT/MM implementations[38,39] took up the challenge posed by the requirement to combine DFT treatments of a molecule in an efficient and accurate way with large scale MM environments. Here, Laio *et al.*[38] emphasized the need of a fully Hamiltonian description, which was violated by certain approximations applied by Eichinger *et al.*,[34] while Laino *et al.*[39] additionally provided a clever suggestion for the efficient computation of the electrostatic interaction between the DFT and MM fragments.

However, applications of the above DFT/MM setting to the computation of infrared (IR) spectra of biological chromophores like retinal in bacteriorhodopsin[40,41] or flavin in blue light sensing domains[42] through instantaneous normal

a)Electronic mail: gerald.mathias@physik.uni-muenchen.de

mode analyses[36,43] also revealed those limitations, which are due to the neglected polarizability of the MM protein environments surrounding the DFT chromophores. The corresponding errors in the computed vibrational spectra could be largely removed by iterative DFT/MM calculations of polarized force fields in the respective chromophore binding pockets, thus, uniquely proving that the neglected polarizabilities were the main cause of the earlier ill-descriptions.

A similar attempt to compute the vibrational spectra of phosphate ions in aqueous solution[44] showed that the DFT/MM calculations largely underestimate the solvatochromic shifts in the IR spectra. Here, these underestimates were erroneously attributed to the neglected polarizability of the solvating water, which had been modeled by Jorgensen's "transferable three point interaction potential" (TIP3P).[45] By contrast, recent "first principles" DFT-MD simulations of phosphates in small periodic water boxes have clearly shown[46] that the use of the TIP3P model entailed highly erroneous structures for the first solvation shell, which are mainly due to its simplified structure and to a lesser degree due to the neglected polarizability. Hence, it remains to be seen whether improved (and polarizable) MM water models combined with a DFT description of the phosphate solutes can predict the solvatochromic shifts in the phosphate IR spectra at a quality comparable to that of the very expensive "first principles" DFT-MD simulations.

To enable rapid and nevertheless accurate computations of solvatochromic effects in chromophore IR spectra, the construction of a new and efficient DFT/PMM implementation therefore seemed necessary.

In this paper, we address two issues. First, we thoroughly revise the DFT/MM suggestion made by Eichinger *et al.*[34] and develop an efficient, accurate, and fully Hamiltonian electrostatic DFT/MM coupling scheme whose computational effort scales logarithmically with the number of condensed phase atoms surrounding the DFT fragment. Preserving the thereby achieved levels of accuracy and efficiency, we next extend this scheme by including dynamic polarization effects through inducible atomic dipoles. We give analytical expressions for the calculation of the forces and, therefore, are able to employ the new DFT/PMM scheme for molecular dynamics simulations. For the implementation, the program packages of choice are the parallelized PMM-MD program IPHIGENIE[47] and the parallelized grid-based plane wave DFT program CPMD.[48]

II. THEORY

The Hamiltonian of a DFT/PMM hybrid system can be decomposed into the following four contributions:

$$H = H_{\mathrm{MM}} + H_{\mathrm{PMM}} + H_{\mathrm{DFT}} + H_{\mathrm{DFT/(P)MM}}. \tag{1}$$

Here, H_{MM} represents one of the standard MM force fields[3–6] including the kinetic energy of the atoms, and

$$H_{\mathrm{PMM}} = \frac{1}{2}\sum_{i,j\neq i} q_i \Phi(\mathbf{r}_i \,|\, \mathbf{p}_j, \mathbf{r}_j, \tilde{\sigma}_j) - \frac{1}{2}\sum_i \mathbf{p}_i \cdot \langle \mathbf{E}^{q,\mathbf{p}}(\mathbf{r}_i)\rangle_{\tilde{\sigma}_i} + \frac{1}{2}\sum_i \mathbf{p}_i^2/\alpha_i \tag{2}$$

accounts[1,49–51] for the energy contribution of polarizable Gaussian dipoles

$$\mathbf{p}_i^{\mathrm{G}}(\mathbf{r} \,|\, \mathbf{r}_i, \tilde{\sigma}_i) = \mathbf{p}_i \, g(\mathbf{r} \,|\, \mathbf{r}_i, \tilde{\sigma}_i)$$

of strengths $\mathbf{p}_i$ and widths $\tilde{\sigma}_i$, which are located at the atomic positions $\mathbf{r}_i$ and have the shape functions

$$g(\mathbf{r} \,|\, \mathbf{r}_i, \tilde{\sigma}_i) = \frac{1}{\left(2\pi\tilde{\sigma}_i^2\right)^{3/2}} \exp\left[-\frac{(\mathbf{r}-\mathbf{r}_i)^2}{2\tilde{\sigma}_i^2}\right]. \tag{3}$$

Note that in PMM force fields, the use of Gaussian dipoles yields an enhanced algorithmic stability,[52,53] if the widths $\tilde{\sigma}_i$ are chosen sufficiently large, i.e., typically $\tilde{\sigma}_i \approx 0.1$ nm.

The symbol $\Phi(\mathbf{r}_i \,|\, \mathbf{p}_j, \mathbf{r}_j, \tilde{\sigma}_j)$ in Eq. (2) denotes the electrostatic potential generated at the position $\mathbf{r}_i$ of an atom $i \neq j$ by a Gaussian dipole $\mathbf{p}_j^{\mathrm{G}}(\mathbf{r} \,|\, \mathbf{r}_j, \tilde{\sigma}_j)$. Furthermore, the bracket expression

$$\langle f(\mathbf{r}_i)\rangle_{\tilde{\sigma}_i} \equiv \int f(\mathbf{r}) g(\mathbf{r} \,|\, \mathbf{r}_i, \tilde{\sigma}_i) d\mathbf{r} \tag{4}$$

denotes the average of a function $f(\mathbf{r})$ over the volume occupied by $g(\mathbf{r} \,|\, \mathbf{r}_i, \tilde{\sigma}_i)$. If $\mathbf{E}(\mathbf{r} \,|\, q_j, \mathbf{r}_j) = -\nabla(q_j/|\mathbf{r} - \mathbf{r}_j|)$ is the field of a point charge q_j and $\mathbf{E}(\mathbf{r} \,|\, \mathbf{p}_j, \mathbf{r}_j, \tilde{\sigma}_j) = -\nabla\Phi(\mathbf{r} \,|\, \mathbf{p}_j, \mathbf{r}_j, \tilde{\sigma}_j)$ is the field of a Gaussian dipole at $\mathbf{r}_j$, then

$$\langle \mathbf{E}^{q,\mathbf{p}}(\mathbf{r}_i)\rangle_{\tilde{\sigma}_i} \equiv \sum_{j\neq i} \left\langle \mathbf{E}(\mathbf{r}_i \,|\, q_j, \mathbf{r}_j) + \mathbf{E}(\mathbf{r}_i \,|\, \mathbf{p}_j, \mathbf{r}_j, \tilde{\sigma}_j)\right\rangle_{\tilde{\sigma}_i} \tag{5}$$

is the field polarizing atom i. Assuming linear response, the dipole strengths $\mathbf{p}_i$ are calculated by

$$\mathbf{p}_i = \alpha_i \langle \mathbf{E}^{q,\mathbf{p}}(\mathbf{r}_i)\rangle_{\tilde{\sigma}_i} \tag{6}$$

from the scalar atomic polarizabilities α_i and from the polarizing fields (5) in a self-consistent field iteration (PMM-SCF).[54,55] The last term in Eq. (2) is the self-energy required to create the dipoles $\mathbf{p}_i$. If Eq. (6) is self-consistently fulfilled, this self-energy cancels the second term in Eq. (2) and the first term remains as the polarization contribution to the total energy.

In Eq. (1), H_{DFT} is the energy function of the isolated quantum system. The DFT/(P)MM interaction energy

$$H_{\mathrm{DFT/(P)MM}} = H^{\mathrm{vdW}}_{\mathrm{DFT/MM}} + H^{\mathrm{bonded}}_{\mathrm{DFT/(P)MM}} + H^{\mathrm{elec}}_{\mathrm{DFT/(P)MM}} \tag{7}$$

has a contribution from van der Waals interactions $H^{\mathrm{vdW}}_{\mathrm{DFT/MM}}$ between MM and DFT atoms, which is calculated with the applied MM force field. If chemical bonds between the DFT and (P)MM fragments exist, a term $H^{\mathrm{bonded}}_{\mathrm{DFT/(P)MM}}$ has to be included, for which several suggestions exist.[2,34]

We will, however, focus here on chemically non-bonded PMM and DFT fragments, for which the electrostatic interaction energy

$$H^{\mathrm{elec}}_{\mathrm{DFT/(P)MM}} = \int d\mathbf{r}\rho(\mathbf{r})\Phi_{\mathrm{ext}}(\mathbf{r}) \tag{8}$$

is given by the classical expression for the energy of the DFT fragment's charge density ρ in the external potential Φ_{ext} generated by the partial charges and induced Gaussian dipoles in the PMM fragment. The DFT charge density $\rho(\mathbf{r}) = \rho_{\mathrm{e}}(\mathbf{r}) + \rho_{\mathrm{c}}(\mathbf{r})$ comprises contributions $\rho_{\mathrm{e}}(\mathbf{r})$ of the valence

electrons and $\rho_c(\mathbf{r})$ of the nuclear cores. Correspondingly, the interaction Hamiltonian

$$H^{\text{elec}}_{\text{DFT/(P)MM}} = H_e + H_c \tag{9}$$

decomposes into energies

$$H_\kappa = \int d\mathbf{r}\, \rho_\kappa(\mathbf{r})\, \Phi_{\text{ext}}(\mathbf{r}), \qquad \kappa \in \{e, c\}, \tag{10}$$

associated to the electrons (e) and nuclear cores (c) of the DFT fragment.

In the computation of those atomic forces, which are caused by an external potential $\Phi_{\text{ext}}(\mathbf{r})$, the employed DFT program CPMD[48] treats the effective core charges as Gaussian distributions $q_\mu\, g(\mathbf{r}\,|\,\mathbf{r}_\mu, \sigma_\mu)$ centered with widths σ_μ around the positions $\mathbf{r}_\mu$ of the DFT atoms μ. Thus, with definition (4) of Gaussian averages one obtains from Eq. (10) for the core Hamiltonian:

$$H_c = \sum_\mu q_\mu \langle \Phi_{\text{ext}}(\mathbf{r}_\mu)\rangle_{\sigma_\mu}. \tag{11}$$

The Gaussian averages $\langle \Phi_{\text{ext}}(\mathbf{r}_\mu)\rangle_{\sigma_\mu}$ are readily calculated for external potentials generated by Gaussian charge or dipole distributions of widths σ_i, because one solely has to replace the σ_i by the widths $\sigma_{i\mu} = (\sigma_i^2 + \sigma_\mu^2)^{1/2}$ in the respective expressions for the potentials. Note that these expressions reduce for distances $r_{i\mu} \equiv |\mathbf{r}_{i\mu}| \equiv |\mathbf{r}_i - \mathbf{r}_\mu| \gg \sigma_{i\mu}$ to the potentials of point charges and dipoles.

As a result, the Gaussian approximation $q_\mu\, g(\mathbf{r}\,|\,\mathbf{r}_\mu, \sigma_\mu)$ of the nuclear pseudo-potentials enables a speedy evaluation of H_c. By contrast, the computation of H_e requires the evaluation of $\Phi_{\text{ext}}(\mathbf{r})$ at all N_γ points γ of the grid, on which ρ_e is represented in real space by CPMD.[48] Introducing the electronic grid charges

$$q_\gamma = (V_{\text{box}}/N_\gamma)\rho_e(\mathbf{r}_\gamma), \tag{12}$$

whose sum over all grid points γ in the DFT box volume V_{box} is the total charge of all valence electrons in the DFT fragment, one can numerically evaluate the integral from Eq. (10) for the electronic contribution to the DFT/(P)MM interaction Hamiltonian (9) by the expression

$$H_e = \sum_\gamma q_\gamma\, \Phi_{\text{ext}}(\mathbf{r}_\gamma). \tag{13}$$

This approximation is valid because our way of constructing Φ_{ext} guarantees, as will be explained in more detail below, that the external potential is smooth on the spatial scale defined by the DFT grid.

A. Computational issues

The forces required for MD simulations are obtained from the Hamiltonian (1) by taking negative gradients with respect to the atomic coordinates at every time step of the numerical integration of the classical equations of motion. In a DFT/MM setting, only the Kohn-Sham wave functions have to be determined in a SCF iteration (DFT-SCF) at every time step, whereas DFT/PMM simulations additionally require a PMM-SCF procedure, which has to be properly interfaced with the DFT-SCF calculations.

The computational effort required for the four components of the Hamiltonian (1) depends on the respective sizes of the DFT and PMM fragments, on the spatial resolution of the DFT grid, and on the choice of the SCF convergence criteria. Nevertheless, for a system composed of about 10^4 PMM solvent atoms and a rather small DFT solute molecule comprising about 10 DFT atoms, one may estimate that all four components of H will pose comparable numerical tasks.

Treating, for instance, a water molecule by DFT with a 70 Ry plane wave cutoff requires $N_\gamma \approx 10^6$ grid points. Within a brute force computational approach, the evaluation of Eq. (13), which is the most expensive contribution to $H_{\text{DFT/(P)MM}}$, would lead for the PMM fragment characterized above to about 10^{10} distance calculations. The associated effort would then definitely represent the computational bottleneck. For the reduction of this effort three different suggestions exist,[34,38,39] which all utilize multi-scale concepts. Here, we will adopt and extend the suggestion by Eichinger *et al.*[34] ensuring, however, the Hamiltonian character of the resulting dynamics (cf. Laio *et al.*[38]).

Accordingly, we will treat the DFT fragment as a component of the nested hierarchy into which a simulation system with toroidal boundary conditions[57] is decomposed, if the electrostatic interactions are calculated by the combination of the pth-order "structure adapted fast multipole method" with a reaction field approach (SAMM$_p$/RF) developed in Refs. 47 and 58–60. In pure MM- or PMM-MD simulations of large systems, this fast multipole method (FMM)[61–63] enables an efficient and accurate calculation of the electrostatics, which scales linearly with the number of atoms. Due to a balanced combination of mth order multipole moments with nth order local Taylor expansions, which is expressed by the equation $p = n + m$, the electrostatic forces calculated with SAMM$_p$ exactly obey Newton's reaction principle.[47,64] Furthermore, this choice additionally guarantees a minimal computational effort for a predefined level p of accuracy ($p = 4$ is the standard of the current implementation available in IPHIGENIE). A predecessor version called SAMM had been used in Eichinger's DFT/MM approach,[34] whose fully Hamiltonian DFT/(P)MM extension will be explained below.

B. DFT/PMM with SAMM$_p$

According to Eq. (13), the external potential Φ_{ext} is imported into the DFT Hamiltonian through its evaluation at the points γ of the DFT grid. For an efficient solution of this computational task, Eichinger *et al.*[34] applied and extended (cf. Figs. 4 and 5 in Ref. 34) the distance class scheme of SAMM.[58,59,65] The extended scheme partitions the PMM environment of each DFT atom μ into disjoint distance classes C_μ^l, $l = 0, \ldots, l_{\max}$. For periodic systems, a RF continuum starts beyond the outermost distance class ($l_{\max}$) at a distance dictated by the minimum image convention.[60] Figure 1 illustrates the three innermost distance classes C_μ^l, $l = 0, 1, 2$ for an atom μ of a DFT water molecule embedded in liquid PMM water.

Figure 1 shows one PMM water molecule in each of the three classes and indicates the distances d used for their

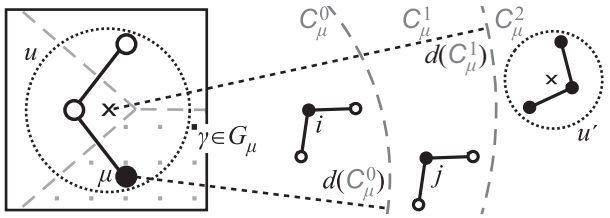

FIG. 1. Scheme of inner distance classes for the FMM evaluation of the DFT/PMM electrostatics: A water molecule (left) representing a structural unit u (dotted circle) of the SAMM$_p$ hierarchy is embedded in a rectangular DFT box, which is discretized by a grid (dots). Only those points $\gamma \in G_\mu$ are shown which belong to a selected DFT atom μ (large black dot) through the Voronoi tessellation of the box. This tessellation defines the index sets G_μ and is indicated by the dashed gray lines. Two dashed gray segments of circles [radii $d(C_\mu^0)$, $d(C_\mu^1)$] around μ and the reference point "×" of the structural unit u indicate the outer limits of the distance classes C_μ^0 and C_μ^1, respectively. Representative atoms $i \in C_\mu^0$ and $j \in C_\mu^1$ of PMM water molecules belonging to these two classes and of a structural unit $u' \in C_\mu^2$ are drawn as black dots.

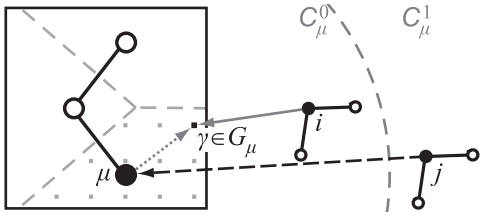

FIG. 2. Evaluation of Φ_{ext} at a grid point $\gamma \in G_\mu$: Contributions from Gaussian charges and induced dipoles of a PMM atom $i \in C_\mu^0$ are evaluated directly (solid gray arrow), whereas the contributions from more distant atoms, like the one indicated by the dashed black arrow for a PMM atom $j \in C_\mu^1$, are calculated by a Taylor expansion around the position of the DFT atom μ. The dotted gray arrow marks the connection $\mathbf{r}_{\gamma\mu}$ of the points μ and γ used in the Taylor expansion.

definition. Typical values are $d(C_\mu^0) \approx 6$ Å and $d(C_\mu^1) \approx 8$ Å. The electrostatics of the PMM atoms $i \in C_\mu^0$ is represented by Gaussian partial charges of widths σ_i and by Gaussian induced dipoles of widths $\tilde{\sigma}_i$, which all are typically smaller than 1 Å but much larger than the spacing of the DFT grid. The parameters $(\sigma_i, \tilde{\sigma}_i)$ steer the strengths of the near-field electrostatic interactions between the DFT and the PMM atoms. In the case of an aqueous PMM environment, for instance, the strength of the DFT/PMM hydrogen bonding interactions can be tuned by proper choices of these Gaussian widths.[38] Because they are about one order of magnitude smaller than the typical distances $r_{\mu j} > d(C_\mu^0)$ of atoms $j \in C_\mu^1$ from the given DFT atom μ, the Gaussian character of the PMM charges and dipoles can safely be neglected for the class C_μ^1 and beyond. In C_μ^0, the use of smoothed charge and dipole distributions is mandatory[34,38,39,56] to avoid artificial distortions of the DFT electron density $\rho_e(\mathbf{r}_\gamma)$ and to guarantee that Φ_{ext} is sufficiently smooth on the scale of the DFT grid spacing.

The Voronoi tessellation of the DFT box characterized by the gray dashed lines in Fig. 1 decomposes the whole DFT grid into disjoint subsets G_μ associated to the various DFT atoms μ. Correspondingly, the sum $\sum_\gamma$ in Eq. (13) can be expressed as the double sum

$$H_e = \sum_\mu \sum_{\gamma \in G_\mu} q_\gamma \Phi_{\text{ext}}(\mathbf{r}_\gamma), \tag{14}$$

which partitions H_e into a sum over contributions associated to the DFT atoms μ. The proximity of the grid points $\gamma \in G_\mu$ to the positions $\mathbf{r}_\mu$ of the DFT atoms can now be exploited for the rapid evaluation of Eq. (14) by taking advantage of the SAMM$_p$ algorithm.[47]

C. Efficient computation of Φ_{ext}

According to the SAMM scheme,[34] the electrostatic potential at points $\mathbf{r}_\gamma$ in the vicinity of a given atom μ is calculated as a sum

$$\Phi_{\text{ext}}(\mathbf{r}_\gamma) = \sum_{l=0}^{l_{\max}} \Phi\left(\mathbf{r}_\gamma \mid C_\mu^l\right), \qquad \gamma \in G_\mu \tag{15}$$

over contributions $\Phi(\mathbf{r}_\gamma \mid C_\mu^l)$ from sources located in the distance classes C_μ^l, $l = 0, \ldots, l_{\max}$, to which for periodic systems a reaction field contribution $\Phi(\mathbf{r}_\gamma \mid \text{RF})$ is added[60] (for notational simplicity it will be omitted in the subsequent discussion).

Figure 2 illustrates how the external potential $\Phi_{\text{ext}}(\mathbf{r}_\gamma)$ is calculated using SAMM$_p$ for two PMM water molecules belonging to the distance classes C_μ^0 and C_μ^1 of a DFT atom μ. Here, the solid gray arrow marks the computation of the potential generated by the electrostatic moments of the PMM atoms $i \in C_\mu^0$ through

$$\Phi\left(\mathbf{r}_\gamma \mid C_\mu^0\right) = \sum_{i \in C_\mu^0} [\Phi(\mathbf{r}_\gamma \mid q_i, \mathbf{r}_i, \sigma_i) + \Phi(\mathbf{r}_\gamma \mid \mathbf{p}_i, \mathbf{r}_i, \tilde{\sigma}_i)] \tag{16}$$

with the potentials

$$\Phi(\mathbf{r}_\gamma \mid q_i, \mathbf{r}_i, \sigma_i) = \frac{q_i \text{erf}[r_{\gamma i}/(\sqrt{2}\sigma_i)]}{r_{\gamma i}} \tag{17}$$

of Gaussian partial charges q_i and

$$\Phi(\mathbf{r}_\gamma \mid \mathbf{p}_i, \mathbf{r}_i, \tilde{\sigma}_i) = -\mathbf{p}_i \cdot \partial_{(1)} \Phi(\mathbf{r}_\gamma \mid q_i, \mathbf{r}_i, \tilde{\sigma}_i)/q_i \tag{18}$$

of Gaussian dipoles $\mathbf{p}_i^G$. In Eq. (18), the gradient is written as $\partial_{(1)}$.

For electrostatic PMM moments in all higher ($l \geq 1$) distance classes C_μ^l, the potentials

$$\Phi\left(\mathbf{r}_\gamma \mid C_\mu^l\right) = \sum_{n=0}^{p} \frac{1}{n!} \mathbf{r}_{\gamma\mu}^{(n)} \odot \mathbf{T}^{n,p}\left(\mathbf{r}_\mu \mid C_\mu^l\right) \tag{19}$$

are calculated through pth order Taylor expansions around the position $\mathbf{r}_\mu$ of the DFT atom μ. The symbol $\mathbf{r}_{\gamma\mu}^{(n)}$ is the n-fold outer product of $\mathbf{r}_{\gamma\mu}$ with itself. This vector $\mathbf{r}_{\gamma\mu}$ connects the reference point μ with the grid point $\gamma \in G_\mu$ (dotted gray arrow in Fig. 2). The symbol $\odot$ denotes the inner contraction product of two tensors (Ref. 47 thoroughly explains the employed tensorial notation). Finally, the class specific expansion coefficients

$$\mathbf{T}^{n,p}\left(\mathbf{r}_\mu \mid C_\mu^l\right) \equiv \partial_{(n)} \Phi^T\left(\mathbf{r} \mid C_\mu^l\right)\big|_{\mathbf{r}_\mu} \tag{20}$$

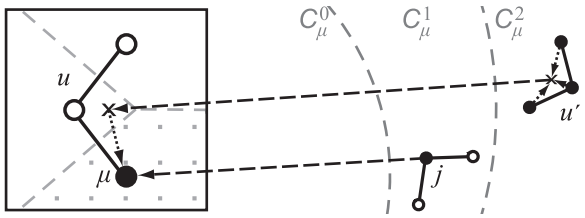

FIG. 3. Computation of Taylor expansion coefficients: The charges and induced dipoles of the PMM atoms $j \in C_\mu^1$ generate the coefficients $\mathbf{T}^{n,p}(\mathbf{r}_\mu \mid C_\mu^1)$ (lower dashed arrow). The PMM atoms in C_μ^2 are collected into structural units, whose electrostatic signatures are represented by multipole expansions.[47] For the PMM unit u', for instance, such an expansion is symbolized by three black dotted arrows pointing toward its reference point "×". The multipole potentials originating from u' are expanded into a Taylor series at the reference point of the DFT unit u (upper dashed arrow), from which the additional contributions $\mathbf{T}^{n,p}(\mathbf{r}_\mu \mid C_\mu^2)$ to the atom-centered expansion coefficients are inherited (dotted arrow) by a simple shifting procedure.[47]

are nth rank tensors generated by the nth order partial derivatives of the potentials $\Phi^{\mathrm{T}}(\mathbf{r} \mid C_\mu^l)$ at $\mathbf{r}_\mu$, which originate from point-like electrostatic moments (cf. the discussion of Fig. 1 above) occupying the distance class C_μ^l.

Here, the nature of the electrostatic moments, which generate the potential appearing in Eq. (20), differs for the distance classes at level $l = 1$ and at levels $l \geq 2$, respectively. As is schematically indicated by the lower dashed arrow in Figure 3, at $l = 1$ the point charges q_j and induced point dipoles $\mathbf{p}_j$ of the individual atoms j are considered to be the sources of the potential. The associated expansion coefficients

$$\mathbf{T}^{n,p}\left(\mathbf{r}_\mu \mid C_\mu^1\right) = \partial_{(n)} \sum_{j \in C_\mu^1} [q_j - (1 - \delta_{np})\, \mathbf{p}_j \cdot \partial_{(1)}] \frac{1}{r_{\mu j}} \tag{21}$$

are essentially given by the nth rank tensors[47] $\partial_{(n)}(1/r)$. Here, the prefactor $1 - \delta_{np}$, in which δ is the Kronecker symbol, ensures that the expansion is of comparable accuracy for the atomic charges and dipoles.

For distance class levels $l \geq 2$, by contrast, mth order multipole moments ($m = 0, \ldots, p$) of nested and increasingly larger charge and dipole distributions are considered as the sources of the potential in Eq. (20). Within the SAMM$_p$ algorithm[47] the Taylor expansion coefficients

$$\mathbf{T}^{n,p}\left(\mathbf{r}_\mu \mid C_\mu^l\right) = \partial_{(n)} \sum_{m=0}^{p-n} \Phi^m\left(\mathbf{r}_\mu \mid C_\mu^l\right), \qquad n = 0, \ldots, p, \quad l \geq 2 \tag{22}$$

are then the partial derivatives of the mth order multipole potentials $\Phi^m(\mathbf{r}_\mu \mid C_\mu^l)$. As documented in the Appendix, the SAMM$_p$ algorithm, which has been originally developed for distributions of static charges,[47] can meanwhile also account for induced dipoles.

Figure 3 indicates for three atoms $k \in C_\mu^2$ collected into a unit u' the computation of multipole moments by dotted arrows and the calculation of the Taylor coefficients (22) by a two-step process (upper dashed and left dotted arrows) as is common in FMM methods.[47]

D. Forces on the DFT atoms μ

Equations (16) and (19) specify the two basically different procedures by which the external potential is imported onto the DFT grid. This import enables the DFT program CPMD to compute a polarized electron density $\rho_{\mathrm{e}}(\mathbf{r}_\gamma)$. Using the updated grid charges q_γ [cf. Eq. (12)] and the Gaussian core charge q_μ centered at $\mathbf{r}_\mu$, CPMD then calculates the electrostatic interaction energies

$$H_\kappa = \sum_\mu \sum_{l=0}^{l_{\max}} H_\kappa\left(C_\mu^l\right), \qquad \kappa \in \{\mathrm{e}, \mathrm{c}\}, \tag{23}$$

because of Eqs. (11) and (13)–(15) as sums of electronic

$$H_{\mathrm{e}}\left(C_\mu^l\right) = \sum_{\gamma \in G_\mu} q_\gamma\, \Phi\left(\mathbf{r}_\gamma \mid C_\mu^l\right) \tag{24}$$

and nuclear

$$H_{\mathrm{c}}\left(C_\mu^l\right) = q_\mu \langle \Phi\left(\mathbf{r}_\mu \mid C_\mu^l\right) \rangle_{\sigma_\mu} \tag{25}$$

contributions associated to the DFT atoms μ and distance class levels l. According to Eq. (9) an update of the interaction Hamiltonian $H^{\mathrm{elec}}_{\mathrm{DFT/(P)MM}}$ has thus been determined. By taking gradients of $H_{\mathrm{DFT}} + H^{\mathrm{elec}}_{\mathrm{DFT/(P)MM}}$ with respect to the coordinates $\mathbf{r}_\mu$, CPMD can now compute the electrostatic forces on the DFT atoms μ.

The electrostatic Hellmann-Feynman reaction forces[66] exerted by the charges q_γ and q_μ on the PMM atoms i follow from the gradients of $H^{\mathrm{elec}}_{\mathrm{DFT/(P)MM}}$ with respect to the coordinates $\mathbf{r}_i$. However, in the current form, the contributions (24) and (25) to the electrostatic interaction energy do not immediately reveal how $H^{\mathrm{elec}}_{\mathrm{DFT/(P)MM}}$ depends on the PMM coordinates $\mathbf{r}_i$ and, therefore, how these forces should be calculated. To uncover this dependence, the electronic and nuclear interaction energies $H_\kappa(C_\mu^l)$ will now be separately analyzed for the distance classes $l = 0$, 1, and $l \geq 2$.

E. Reaction forces on the PMM atoms $i \in C_\mu^0$

The electronic near-field interaction Hamiltonian

$$H_{\mathrm{e}}\left(C_\mu^0\right) = \sum_{\gamma \in G_\mu} q_\gamma \sum_{i \in C_\mu^0} [\Phi(\mathbf{r}_\gamma \mid q_i, \mathbf{r}_i, \sigma_i) + \Phi(\mathbf{r}_\gamma \mid \mathbf{p}_i, \mathbf{r}_i, \tilde{\sigma}_i)] \tag{26}$$

is obtained by inserting the external potential (16) originating from class C_μ^0 into Eq. (24). It is the energy of the point-like grid charges q_γ in the potentials of the Gaussian charges q_i of widths σ_i and Gaussian induced dipoles $\mathbf{p}_i$ of widths $\tilde{\sigma}_i$ collected in C_μ^0. Equations (17) and (18), respectively, specify these potentials in terms of the connection vectors $\mathbf{r}_{\gamma i}$ pointing from PMM atom i to the grid point γ.

Replacing these vectors by their inverses $\mathbf{r}_{i\gamma} = -\mathbf{r}_{\gamma i}$ and interchanging in Eq. (26), the sums over γ and i lead to the

strictly equivalent expression

$$H_e(C_\mu^0) = \sum_{i\in C_\mu^0}\Bigg[q_i \sum_{\gamma\in G_\mu}\Phi(\mathbf{r}_i \mid q_\gamma, \mathbf{r}_\gamma, \sigma_i) - \mathbf{p}_i \cdot \sum_{\gamma\in G_\mu}\mathbf{E}(\mathbf{r}_i \mid q_\gamma, \mathbf{r}_\gamma, \tilde{\sigma}_i)\Bigg], \quad (27)$$

which is the energy of point charges q_i and dipoles $\mathbf{p}_i$ at positions $\mathbf{r}_i$ in the potentials $\Phi(\mathbf{r}_i \mid q_\gamma, \mathbf{r}_\gamma, \sigma_i)$ and fields $\mathbf{E}(\mathbf{r}_i \mid q_\gamma, \mathbf{r}_\gamma, \tilde{\sigma}_i) = -\partial_{(1)}\Phi(\mathbf{r}_i \mid q_\gamma, \mathbf{r}_\gamma, \tilde{\sigma}_i)$ of Gaussian grid charges q_γ of widths σ_i and $\tilde{\sigma}_i$, respectively. The negative gradients of $H_e(C_\mu^0)$ with respect to the positions $\mathbf{r}_i$ are then the electrostatic Hellmann-Feynman forces[66] exerted by the grid charges q_γ, $\gamma \in G_\mu$, on the PMM atoms $i \in C_\mu^0$. If one inverts the solid gray arrow in Fig. 2, the inverted arrow can serve to symbolize such a reverse action of one of the grid charges on a nearby PMM atom.

The core contribution

$$H_c(C_\mu^0) = \sum_{i\in C_\mu^0}[q_i\Phi(\mathbf{r}_i \mid q_\mu, \mathbf{r}_\mu, \sigma_{\mu i}) - \mathbf{p}_i \cdot \mathbf{E}(\mathbf{r}_i \mid q_\mu, \mathbf{r}_\mu, \tilde{\sigma}_{\mu i})] \quad (28)$$

is analogously obtained by inserting the potential $\Phi(\mathbf{r}_\mu \mid C_\mu^0)$ as defined by Eq. (16) into Eq. (25), by executing the Gaussian averages (4) through an increase of the Gaussian widths (as explained in connection with Eq. (11)), and by repeating the arguments, which lead from Eq. (26) to Eq. (27).

F. Reaction forces on the PMM atoms $j \in C_\mu^1$

Inserting the Taylor expansion (19) with the coefficients (21) into (24) yields the electronic interaction energy

$$H_e(C_\mu^1) = \sum_{\gamma\in G_\mu} q_\gamma \sum_{n=0}^{p}\frac{1}{n!}\mathbf{r}_{\gamma\mu}^{(n)} \odot \partial_{(n)} \sum_{j\in C_\mu^1}[q_j - (1-\delta_{np})\,\mathbf{p}_j \cdot \partial_{(1)}]\frac{1}{r_{\mu j}} \quad (29)$$

for DFT atom μ and class C_μ^1. Interchanging the sums over j and γ and employing the identity $\partial_{(n)}(1/r_{\mu j}) = (-1)^n\partial_{(n)}(1/r_{j\mu})$, one finds

$$H_e(C_\mu^1) = \sum_{j\in C_\mu^1}\sum_{n=0}^{p}[q_j + (1-\delta_{np})\,\mathbf{p}_j \cdot \partial_{(1)}]\Phi^n(\mathbf{r}_j \mid Q_\mu, \mathbf{r}_\mu) \quad (30)$$

with the potentials

$$\Phi^n(\mathbf{r}_j \mid Q_\mu, \mathbf{r}_\mu) = \frac{(-1)^n}{n!}\left(\partial_{(n)}\frac{1}{r_{j\mu}}\right) \odot \sum_{\gamma\in G_\mu} q_\gamma \mathbf{r}_{\gamma\mu}^{(n)} \quad (31)$$

generated at the positions $\mathbf{r}_j$ of the PMM atoms by the nth order multipole moments of the electronic grid charges $Q_\mu \equiv \{q_\gamma \mid \gamma \in G_\mu\}$ of the DFT atom μ. According to Lorenzen *et al.*,[47] the potentials (31) can be equivalently written as

$$\Phi^n(\mathbf{r}_j \mid Q_\mu, \mathbf{r}_\mu) = \frac{(-2)^n}{(2n)!}\left(\partial_{(n)}\frac{1}{r_{j\mu}}\right) \odot \mathbf{M}^n(\mathbf{r}_\mu \mid Q_\mu). \quad (32)$$

Here,

$$\mathbf{M}^n(\mathbf{r}_\mu \mid Q_\mu) = \sum_{\gamma\in G_\mu} q_\gamma r_{\mu\gamma}^{2n+1}\left(\partial_{(n)}\frac{1}{r_{\mu\gamma}}\right) \quad (33)$$

are the reduced totally symmetric multipole tensors, which have only $2n+1$ independent components, because they are traceless with respect to every pair of tensor components.[67,68]

A slight rearrangement of Eq. (30) and the introduction of the multipole fields

$$\mathbf{E}^n(\mathbf{r}_j \mid Q_\mu, \mathbf{r}_\mu) = -\partial_{(1)}\Phi^n(\mathbf{r}_j \mid Q_\mu, \mathbf{r}_\mu) \quad (34)$$

finally leads to

$$H_e(C_\mu^1) = \sum_{j\in C_\mu^1}\Bigg[q_j \sum_{n=0}^{p}\Phi^n(\mathbf{r}_j \mid Q_\mu, \mathbf{r}_\mu) - \mathbf{p}_j \cdot \sum_{n=0}^{p-1}\mathbf{E}^n(\mathbf{r}_j \mid Q_\mu, \mathbf{r}_\mu)\Bigg], \quad (35)$$

which explicitly reveals the desired dependence on the PMM coordinates $\mathbf{r}_j$ and, therefore, enables the derivation of likewise simple expressions for the Hellmann-Feynman forces on the PMM atoms $j \in C_\mu^1$.

In the contributions (25) to the core Hamiltonian, the Gaussian average can be neglected at all levels $l \geq 1$. At the reference point $\mathbf{r}_\mu$, i.e., for $\mathbf{r}_{\gamma\mu} = \mathbf{0}$, the Taylor expansion (19) reduces to the zeroth order term $\mathbf{T}^{0,p}(\mathbf{r}_\mu \mid C_\mu^l)$. Inserting this result into Eq. (25) yields the general form

$$H_c(C_\mu^l) = q_\mu \mathbf{T}^{0,p}(\mathbf{r}_\mu \mid C_\mu^l) \quad \text{for} \quad l \geq 1. \quad (36)$$

Inserting for $l = 1$ the coefficients (21) into (36) and repeating the steps, which lead from (29) to (35), one gets

$$H_c(C_\mu^1) = \sum_{j\in C_\mu^1}[q_j\Phi(\mathbf{r}_j \mid q_\mu, \mathbf{r}_\mu) - \mathbf{p}_j \cdot \mathbf{E}(\mathbf{r}_j \mid q_\mu, \mathbf{r}_\mu)], \quad (37)$$

which is the energy of the PMM atoms $j \in C_\mu^1$ in the potential and field of the point-like core charge q_μ. Instead of separately evaluating Eq. (37), one may equivalently include the core charge q_μ into the grid charge distribution Q_μ of DFT atom μ, which then becomes $\hat{Q}_\mu = Q_\mu \cup q_\mu$. Because q_μ is located by construction at the reference point of the multipole expansion of $\hat{Q}_\mu$, its inclusion solely modifies the monopole moment to $\mathbf{M}^0(\mathbf{r}_\mu \mid \hat{Q}_\mu) = \mathbf{M}^0(\mathbf{r}_\mu \mid Q_\mu) + q_\mu$.

G. Reaction forces on the PMM atoms $k \in C_\mu^l$, $l \geq 2$

Starting at level $l = 2$, the computation of the electrostatic interactions becomes identical to the SAMM$_p$ treatment of a purely classical system, which has been described in detail elsewhere.[47] Therefore, it suffices here to sketch how at level $l = 2$ a DFT fragment is integrated into the SAMM$_p$ algorithm.

For a most simple presentation, we assume that the DFT fragment is composed of a single structural unit u like in the example depicted in the above figures. All its atoms μ share a common distance class C_u^2 ($\forall\mu \in u : C_\mu^2 = C_u^2$), which contains the distant PMM atoms interacting on level $l = 2$ with u. Equations (19) and (22)–(24) yield the associated electronic

interaction energy

$$H_e(C_u^2) = \sum_{\mu\in u}\sum_{\gamma\in G_\mu} q_\gamma \sum_{n=0}^{p} \frac{1}{n!}\mathbf{r}_{\gamma\mu}^{(n)} \odot \partial_{(n)} \sum_{m=0}^{p-n} \Phi^m(\mathbf{r}_\mu \mid C_u^2). \tag{38}$$

In this formulation, the potentials $\Phi^m(\mathbf{r}_\mu \mid C_u^2)$ generated by mth order multipole moments of the PMM structural units $u' \in C_u^2$ are considered as sources and the grid charges of the DFT unit u as targets of the electrostatic interactions. These targets are addressed through Taylor expansions around the atomic positions $\mathbf{r}_\mu$.

SAMM$_p$ does not evaluate the nth order expansion coefficients $\partial_{(n)} \sum_{m=0}^{p-n} \Phi^m(\mathbf{r}_\mu \mid C_u^2)$ at each $\mathbf{r}_\mu \in u$, but only at the reference point $\mathbf{r}_u$ (left "×" in Fig. 3) of unit u. Using a Taylor expansion around $\mathbf{r}_u$, they are then simply shifted to the atomic positions $\mathbf{r}_\mu$.[47] By the very construction of SAMM$_p$, this shifting is exactly the inverse operation to the combination[47] of all atomic nth order multipole moments $\mathbf{M}^n(\mathbf{r}_\mu \mid Q_\mu)$, $\mu \in u$, which are known as soon as the energies $H_e(C_\mu^1)$ have been calculated through Eq. (35), into corresponding moments $\mathbf{M}^n(\mathbf{r}_u \mid Q_u)$ of the charge distribution $Q_u \equiv \bigcup_{\mu\in u} Q_\mu$ of the DFT unit u. Note that this symmetry of Taylor and multipole expansions is the reason why in (P)MM simulations the reaction principle holds exactly for the SAMM$_p$ forces. In the given DFT/(P)MM case, however, the grid discretization of ρ_e weakly interferes with this principle through an artifact, which we will address further below.

Due to the quoted symmetry one can equivalently represent the interaction energy (38) in a form in which the multipole moments $\mathbf{M}^n(\mathbf{r}_u \mid Q_u)$ of the grid charge distribution Q_u are the sources of multipole potentials $\Phi^m(\mathbf{r} \mid Q_u)$ acting on distant PMM atoms $k \in C_u^2$ through local Taylor expansions. The Hellmann-Feynman forces on these PMM atoms immediately follow from the corresponding SAMM$_p$ expression (for details and explanations, see Ref. 47). We note that the interaction $H_c(C_u^2)$ of the atomic cores in unit u with the PMM atoms $k \in C_u^2$ is included, if one employs the extended atomic multipole moments $\mathbf{M}^n(\mathbf{r}_\mu \mid \hat{Q}_\mu)$ instead of the electronic moments $\mathbf{M}^n(\mathbf{r}_\mu \mid Q_\mu)$ for computing the moments of unit u.

The analysis given above for the case of a single DFT unit u interacting with PMM atoms in the distance classes C_u^2 is readily generalized to higher cluster levels and larger DFT fragments. In our implementation, the electrostatic DFT/(P)MM interactions are calculated at levels $l \geq 2$ by transferring the atomic multipole moments $\mathbf{M}^n(\mathbf{r}_\mu \mid \hat{Q}_\mu)$ computed by CPMD to the (P)MM-MD program IPHIGENIE, which then calculates the multipole moments $\mathbf{M}^n(\mathbf{r}_u \mid \hat{Q}_u)$ of unit u. From now on the moments of DFT units are treated at all SAMM$_p$ levels $l \geq 2$ just like (P)MM moments. The hierarchically nested FMM scheme then renders the total electrostatic forces on the PMM atoms $k \in C^l$, $l \geq 2$.

H. Remarks

The evaluation of $\sum_{l\geq 1} H_e(C_\mu^l)$ is computationally about as expensive as the evaluation of $H_e(C_\mu^0)$ for a single PMM atom i [cf. Eq. (27)]. Therefore, the computational effort spent on H_e is essentially determined by the average number $N_0 \equiv (1/N_{DFT}) \sum_\mu |C_\mu^0|$ of PMM atoms found in the innermost distance classes C_μ^0 of the N_{DFT} DFT atoms μ. Typically one finds $N_0 \approx 100$ and, therefore, the computational advantage of the above calculation scheme over a brute force method is N_0/N, if N is the number of PMM atoms in the system. Thus, for a typical simulation system with $N = 10^4$, the speedup is about 10^2.

The computational scheme described in Sec. II F for the interactions of the DFT grid charges with the PMM atoms in distance class C_μ^1 resembles the DFT/MM suggestion of Laio *et al.*,[38] which also applies Taylor and multipole expansions on the DFT grid to compute interactions with distant MM atoms. These expansions are centered for the whole DFT grid around a single reference point, are truncated at the order $p = 2$, and treat all distant MM atoms as individual sources and targets of electrostatic interactions. Our approach, by contrast, partitions the DFT grid by N_{DFT} reference points, extends the local expansions up to order $p = 4$, and considers for each DFT atom only the comparably few PMM atoms, which are in the small distance range from about 6 Å to about 8 Å, as individual sources and targets of these expansions while collecting all more distant atoms into a hierarchy of increasingly large clusters. Thus, our approach should be much more accurate and efficient even for relatively small DFT/MM systems. Like the scheme of Laio *et al.*,[38] our approach also does not correct the small force discontinuities occurring whenever atoms change distance classes. However, in our case the effects of these transitions are smaller, because the forces are calculated with higher level multipole and Taylor expansions.

III. KEY POINTS OF THE IMPLEMENTATION

Section II completely covers the basic theory of our DFT/PMM approach. However, for an energy conserving and computationally efficient implementation two important issues must be additionally considered.

A. Movements of the DFT box

In DFT/(P)MM dynamics simulations, the grid-based representation of ρ_e by CPMD interferes with energy conservation. In CPMD, the energy E of a DFT atom μ depends on its relative position within the grid. Shifting, e.g., its position $\mathbf{r}_\mu$ along the line connecting a grid point γ with one of its nearest neighbors entails a sinusoidal modulation $E(|\mathbf{r}_\mu - \mathbf{r}_\gamma|) \sim -\Delta E \cos(2\pi|\mathbf{r}_\mu - \mathbf{r}_\gamma|/a)$, where a is the associated grid constant. In the DFT setting applied by us (see Sec. IV), the relative modulation $\Delta E/E(0)$ is about 10^{-5}%. Thus, the atom prefers to sit at grid points and experiences artificial grid forces at other positions. As long as the grid remains fixed in space, this small grid artifact solely represents a rough background potential, whose contributions to the total energy on average vanish during a dynamics simulation.

However, if the DFT fragment moves during a DFT/(P)MM dynamics simulation, the DFT box has to follow. Such a movement of the discretized box may lead to random forces on the DFT atoms adding heat to the system.

This serious artifact can be avoided, if the DFT grid is considered as an infinite object, on which the DFT box is shifted in units of the lattice constants whenever the movement of the DFT fragment (as measured, e.g., by its center of geometry) exceeds the lattice constant in one of the three spatial directions. Thus, only those box translations are allowed, which would leave a fully periodic DFT system invariant.

B. DFT/PMM-SCF iteration

The polarizable degrees of freedom of a DFT/PMM system, i.e., the PMM dipole strengths $\mathbf{p}_j$ and the DFT electron density ρ_e, have to be calculated in coupled SCF procedures, which can be rapidly brought to convergence by diligent choices of the initial conditions. Assuming that the $\mathbf{p}_j$ and the Kohn-Sham orbitals determining ρ_e are temporally continuous during the integration of the dynamics, these entities can be extrapolated from a history of $M^h \approx 4$ previous integration steps using Lagrangian polynomials.[69] During the PMM-SCF iteration, the "direct inversion of the iterative subspace" (DIIS) algorithm[70] with a history length $\hat{M}^{h,p} \approx 3$ is used to speed up convergence. Similarly, CPMD[48] applies DIIS[71] during DFT-SCF with $\hat{M}^{h,\rho} = 10$.

After the integration of the nuclear motion, the potential Φ_{ext} polarizing the DFT fragment is computed from the static partial charges and from the extrapolated dipole strengths $\mathbf{p}_j^0 = \mathbf{p}_j^h$ in the PMM fragment. Next, the DFT-SCF iteration is executed with a loose initial convergence criterion $\chi_{DFT}^{ini} = 10\,\chi_{DFT}$, which limits the largest element of the gradient of the wave function.[48] Keeping the resulting first guess ρ_e fixed, the $\mathbf{p}_j$ are iterated until in the nth step $|\mathbf{p}_j^{n-1} - \mathbf{p}_j^n| < \chi_{PMM}$ for all j, where χ_{PMM} is a certain threshold. The resulting $\mathbf{p}_j$ modify the external potential, to which the DFT fragment is exposed. In subsequent DFT-SCF calculations, the tight DFT-SCF criterion χ_{DFT} is used if the preceding PMM-SCF iteration converged within one iteration step or if a predefined number k of DFT-SCF calculations is exceeded. Thus, in the default case $k = 1$, only the first DFT-SCF calculation after an integration step is performed with χ_{DFT}^{ini}.

We will show that this strategy avoids extended and costly DFT-SCF iterations as long as the $\mathbf{p}_j$ are far from convergence. It partially resembles a scheme proposed by Thompson and Schenter[8,9] in the context of QM/PMM, where QM stands for semi-empirical quantum chemistry.

IV. METHODS

For the examination of our new DFT/PMM method, we employed two different simulation systems, a water dimer and a periodic box of liquid PMM water containing one DFT water molecule. The dynamics was integrated by the Verlet algorithm[72] with a time step $\Delta t = 0.25$ fs for the dimer and $\Delta t = 0.5$ fs for the DFT/PMM liquid. The geometries of the respective PMM water models were fixed using MSHAKE[73] and the electrostatics was treated at the $SAMM_4$ level.[47] The respective DFT water molecule was described by the gradient-corrected exchange functional of Becke[74] together with the correlation functional of Perdew (BP),[75] and the norm-conserving pseudo-potentials of Troullier and Martins (MT).[76] It was centered into a cubic box of edge length 9 Å containing the grid of the plane wave basis set, which was cutoff at 70 Ry. We denote this particular DFT approach by MT/BP.

A. Water dimer

The energy conservation was checked by 2 ps MD simulations of water dimers. Here the initial velocities indicated a temperature of about 80 K. We adopted the SCF convergence criteria $\chi_{PMM} = 10^{-4}$ D and $\chi_{DFT} = 10^{-7}$. In the DFT/PMM hybrid setting, the H-bond donor was described by MT/BP and the acceptor by the initial version $TL4P_{ini}$ of a recently developed PMM water model (Tröster *et al.*[90]). To provide references, we simulated a MT/BP dimer using a $(15\ Å)^3$ DFT box and a $TL4P_{ini}$ dimer.

$TL4P_{ini}$ features the experimental liquid phase geometry[77,78] ($l_{OH} = 0.968$ Å, $\varphi_{HOH} = 105.3°$), the experimental gas phase dipole moment[79] (1.85 D) and polarizability[80] (1.47 $Å^3$), a massless negative charge $q_M = -1.172$ e on the bisectrix 0.258 Å distant from the oxygen, and positive charges at the hydrogens as well as a Gaussian inducible dipole of width $\tilde{\sigma}_i = 0.912$ Å at the oxygen. The van der Waals interactions were treated identically for all components of the hybrid systems, i.e., by Buckingham potentials[81] $E_B(r) = A_1 \exp(-rA_2) - B/r^6$ centered at the oxygen atoms ($A_1 = 78700\ Å^{12}$ kcal/mol, $A_2 = 3.50\ Å^{-1}$, $B = 1062\ Å^6$ kcal/mol). For the Gaussian distributions, which represent the static partial charges of the PMM atoms as long as they are close to the DFT atoms, we chose identical widths $\sigma_i = 0.57$ Å as suggested in Ref. 34.

In the unrestrained hybrid dimer, the PMM fragment moves close to the DFT fragment thus probing the innermost distance classes at the level $l = 0$. To check the electrostatics treatment also for outer distance classes (levels $l = 1, 2$), we softly restrained the distance d_{OO} of the two oxygens by a harmonic potential with a spring constant of 1 kcal/mol $Å^2$ to $d_{OO,1} = 7$ Å and $d_{OO,2} = 10$ Å, respectively, thereby guaranteeing that the interactions were calculated within the distance class level $l \in \{1, 2\}$ of interest. This probing of outer distance classes was also used in reference simulations of the PMM dimer.

B. Liquid water

The stability and performance of the DFT/PMM algorithm were investigated with a periodic cubic box [volume $V = (46.6\ Å)^3$] filled with $N = 3374$ $TL4P_{ini}$ water molecules. Thus, the experimental density[82] $n = 0.9965$ g/cm^3 of water at the temperature $T = 300$ K and the pressure $p = 1$ atm was prepared. Extending the $SAMM_4$ treatment of the electrostatics by a moving boundary reaction field correction[60] and modeling the surrounding dielectric continuum by a dielectric constant of 80, the system was equilibrated for 1 ns in the NVT ensemble. Here, T was kept at 300 K with a Bussi thermostat[83] (relaxation time 0.1 ps).

To check the long-time stability, ten snapshots were drawn from the last 100 ps of this trajectory. Each snapshot

served as a starting point for a 14 ps DFT/PMM-MD simulation, in which one of the water molecules was described by MT/BP and the thermostat was restricted to the PMM environment.

Starting a series of 250 fs DFT/PMM-MD simulations at identical initial conditions, we studied how the efficiency and accuracy of the algorithm are affected by the SCF convergence criteria χ_{PMM} and χ_{DFT}. The accuracy was assessed by comparing the temporal evolutions of the DFT fragment's energy $E(t)$ and dipole moment $\mathbf{p}(t)$. Replacing the polarizable TL4P$_{\mathrm{ini}}$ water models by non-polarizable TIP4P/2005 potentials[84] and equilibrating this MM system like its PMM counterpart enabled a DFT/MM reference simulation.

V. TEST SIMULATIONS

Numerical integrations of the Hamiltonian dynamics employ time steps Δt of finite size. Therefore, the total energy $E(t)$ shows small fluctuations $\Delta E(t \mid \Delta t)$ around an average value $\langle E \rangle(\Delta t)$.[85] In the limit $\Delta t \to 0$, the fluctuations vanish to leading order with Δt^2. Similarly the average converges to E_0, which is the conserved value of the true Hamiltonian.

Thus, one expects small fluctuations $\Delta E(t \mid \Delta t)$ of the total energy $E(t)$ around its constant average $\langle E \rangle(\Delta t)$ also for the reference dynamics simulation of a PMM water dimer, in which all forces are calculated as exact negative gradients of the potential energy. Figure 4(a) demonstrates that this is actually the case for the unrestrained PMM water dimer at close contact. Here the standard deviation $\sigma_{\Delta E}$ of $\Delta E(t \mid \Delta t)$ is 20×10^{-6} kcal/mol. According to Fig. 4(b), the energy is likewise well-conserved for the PMM dimer softly restrained to distances, which are large enough to enable the approximate treatment of the electrostatics by the SAMM$_4$ algorithm. Here, the standard deviation $\sigma_{\Delta E}$ is only 0.4×10^{-6} kcal/mol, because the forces are smaller by more than one order of magnitude. As expected,[85] for both distance classes the standard deviations $\sigma_{\Delta E}$ vanish and the averages $\langle E \rangle(\Delta t)$ converge with Δt^2. Section S1 of the supplementary material[86] provides evidence for these claims.

Figure 4(c) shows for the reference Born-Oppenheimer (BO) dynamics of the unrestrained DFT dimer the deviation $\Delta E(t \mid \Delta t)$ of the total energy $E(t)$ from its average $\langle E \rangle(\Delta t)$. It exhibits much larger fluctuations $\Delta E(t \mid \Delta t)$ than its PMM relative in Fig. 4(a) as quantified by the standard deviation $\sigma_{\Delta E} = 1.6 \times 10^{-3}$ kcal/mol. Also these fluctuations vanish with Δt^2 and their enhanced magnitude can be largely attributed to the high frequency O–H stretching modes in the DFT dimer, which are absent in the constrained PMM dimer. Fluctuations of a comparable size have been previously reported for the DFT simulation of a water trimer[39] with the related grid code CP2K.[87]

A. Energy conservation in DFT/PMM-MD

Figure 5(a) shows the energy fluctuations observed for the unrestrained DFT/PMM hybrid dimer at close contact. A visual comparison with Fig. 4(c) immediately demonstrates that the average DFT/PMM energy is as well conserved as for the DFT reference and that the DFT/PMM energy fluctuations are of comparable size.

The conservation of the average energy is also observed in Figs. 5(b) and 5(c), which pertain to the DFT/PMM dimers restrained at distances $d_{\mathrm{OO}} \approx 7$ Å and $d_{\mathrm{OO}} \approx 10$ Å. Here the standard deviation $\sigma_{\Delta E}$ of the energy fluctuations is a little larger than at close contact. The similarity of the DFT/PMM fluctuations $\Delta E(t \mid \Delta t)$ to those of the DFT reference becomes even more striking, if one studies the graphs in Figs. 4(c) and 5 at a higher time resolution. A corresponding graphical illustration is provided by Figure S11 in Sec. S2 of the supplementary material.[86]

The above data lead to the conclusion that our DFT/PMM interaction scheme conserves the energy of a hybrid system at all distance classes, because the sample dimers were studied

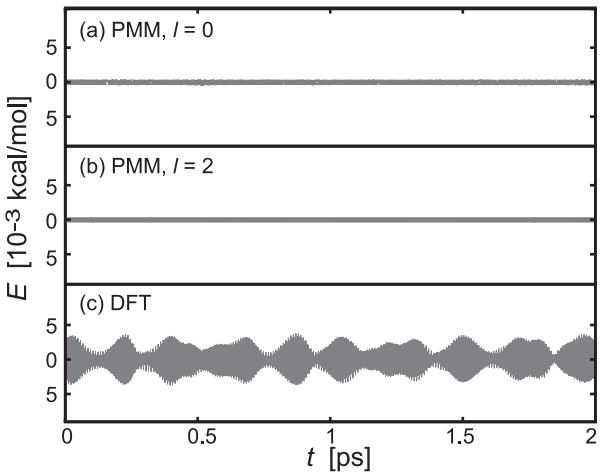

FIG. 4. Energy conservation in reference simulations of the water dimer. (a) PMM dynamics at close contact (exact electrostatics, $l = 0$) and (b) softly restrained to a distance $d_{\mathrm{OO}} \approx 10$ Å (SAMM$_4$ electrostatics, $l = 2$). (c) DFT Born-Oppenheimer dynamics at close contact.

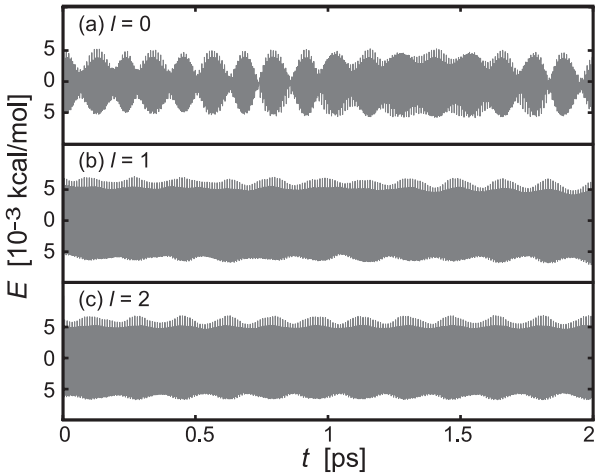

FIG. 5. Energy conservation in DFT/PMM hybrid simulations of the water dimer with the electrostatics treated at different distance class levels l. (a) Close contact ($l = 0$), (b) softy restrained to $d_{\mathrm{OO}} \approx 7$ Å ($l = 1$), and (c) to $d_{\mathrm{OO}} \approx 10$ Å ($l = 2$).

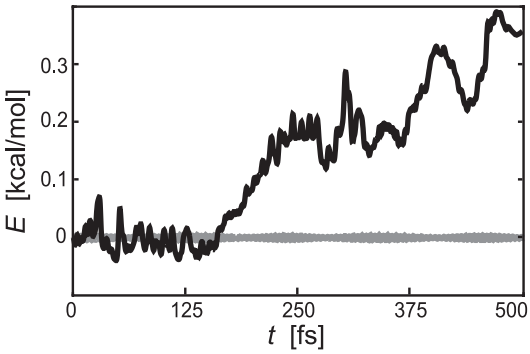

FIG. 6. Influence of the algorithm for DFT box movement on the energy conservation as exemplified by the DFT/PMM water dimer at close contact. The black and gray lines show the trajectories of the total energy in a simulation using a naive and our refined algorithm, respectively (see the text for explanation). The gray line represents the data of Fig. 5(a) on a different energy scale.

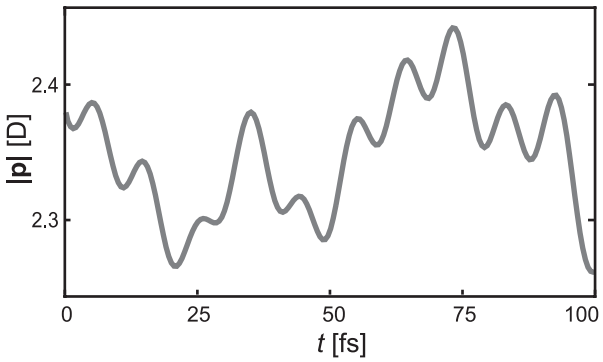

FIG. 7. The absolute value $|\mathbf{p}(t)|$ of the DFT fragment's dipole moment during a MD simulation of the aqueous DFT/PMM system described in Sec. IV. A short (100 fs) section of a trajectory was chosen to visualize the fluctuations of $|\mathbf{p}(t)|$ at a high time resolution.

at the interaction levels $l = 0$, 1, and 2, which cover all relevant algorithmic features discussed above. Note here that the energy was equally well conserved in DFT/MM simulations of a water dimer, in which the TIP3P potential[45] was applied to the MM fragment (data not shown).

Using the DFT/PMM water dimer at close contact as an example, we now additionally demonstrate with Figure 6 how the box movement algorithm described in Sec. III A supports energy conservation. The algorithm adapts the position of the DFT box to the motion of the DFT fragment only occasionally by using discrete displacements, which leave the (infinitely extended) DFT grid invariant.

The gray line in Fig. 6 shows the trajectory $E(t)$ of the total dimer energy for our grid-commensurate algorithm of box movement. The black line is an alternative trajectory, which was obtained by naively moving the DFT box at every MD time step with the center of geometry of the DFT fragment. In this case serious algorithmic artifacts apparently hamper energy conservation.

B. Smoothness and stability of DFT/PMM-MD

The ten 14 ps DFT/PMM-MD simulations of the periodic liquid water box described in Sec. IV clearly revealed the long-time stability of the algorithm. The calculated trajectories turned out to be smooth and did not show any artifacts.

Figure 7 exemplifies this smoothness at an elevated time resolution for the absolute value $|\mathbf{p}|$ of the dipole moment, which was calculated for the DFT fragment. The depicted 100 fs section represents an arbitrary choice from one of the 14 ps trajectories. The observed fluctuations of $|\mathbf{p}(t)|$ are caused by the thermal motions of the DFT molecule and of its surrounding $TL4P_{ini}$ counterparts. Because of their smoothness, one can calculate condensed phase IR spectra of DFT solute molecules from such DFT/PMM trajectories using Fourier transform methods.[36,88]

C. Performance of DFT/PMM-MD

Taking the DFT/PMM liquid water system as an example, Figure 8 gives an overview over the performance of our DFT/PMM algorithm for the choices $\chi_{PMM} = 10^{-4}$ D and $\chi_{DFT} = 10^{-6}$ of the SCF convergence parameters. For this check, CPMD was run in an MPI (message passing interface) parallel version[89] using eight 1.86 GHz cores distributed on two Intel Xeon E5320 processors while IPHIGENIE was executed in sequential mode on one of these cores.

As shown by a sample simulation, the integration of a periodic PMM system of the given size, which exclusively consists of $TL4P_{ini}$ water models, is only by a factor of 4.2 slower than that of a MM system made up of TIP4P/2005 models.[84] For hybrid settings, the comparison of the first two bars in Fig. 8 reveals that T_{PMM} is a factor of six larger than T_{MM}. This increase of T_{PMM} is caused by the additional polarizing action of the DFT fragment on the induced PMM dipoles, which costs on average one additional PMM-SCF iteration step.

For our DFT/PMM water box, T_{PMM} is about as large as the average time T_{DFT} spent with eight processors on the DFT part (cf. Fig. 8). T_{DFT} is composed of the times T_{DFT}^{SCF}, T_{DFT}^{import}, and T_{DFT}^{export} spent for the DFT-SCF iterations, for the import of the electrostatics onto and for its export from the DFT grid, respectively. Figure 8 shows that T_{DFT}^{SCF} is the main contribution to T_{DFT}.

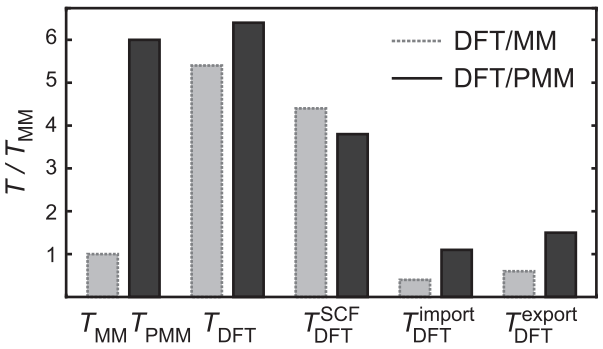

FIG. 8. Average computing times (walltimes) spent for our liquid water sample system per MD integration step on the various parts of a DFT/PMM (black) or DFT/MM (gray) calculation. Here, the DFT part was executed in parallel on eight core and the (P)MM part sequentially on one core. The time T_{MM} spent on the MM part in the DFT/MM setting is taken as the reference. For explanation see the text.

If χ_{DFT} is multiplied by ten, the number of DFT-SCF iterations and, hence, T_{DFT}^{SCF} are reduced by a factor of about 1/2. Fortunately, this increase of χ_{DFT} does not change the trajectories of the DFT fragment's energy $E(t)$ and dipole moment $\mathbf{p}(t)$ within the chosen 250 fs time window (Figure S12 in Sec. S3 in the supplementary material[86] provides a relevant example). Therefore, a criterion $\chi_{DFT} = 10^{-5}$ could be already sufficient for future DFT/(P)MM simulations.

Figure 8 indicates that T_{DFT}^{SCF} is by 15% smaller for DFT/PMM than for DFT/MM. This advantage of a PMM over a MM environment is apparently caused by our choice of an initially loose DFT-SCF convergence criterion $\chi_{DFT}^{ini} = 10\,\chi_{DFT}$ in the joint DFT/PMM-SCF iteration. If one applies only a single criterion ($\chi_{DFT}^{ini} = \chi_{DFT}$) instead, the time T_{DFT}^{SCF} of DFT/PMM is by 20% larger than that of DFT/MM (data not shown).

Despite the computational benefit caused by our two-stage DFT-SCF convergence criterion in the DFT/PMM setting, T_{DFT} is larger than in the DFT/MM case, because the times T_{DFT}^{import} and T_{DFT}^{export} used for the import and the export of the electrostatics are larger by factors 2.7 and 3.0, respectively. These factors mainly reflect the fact that for DFT/MM the DFT module is called only once per integration step, whereas in the DFT/PMM case it is called several times (in the given example: 2.0 DFT calls/integration step). Note that each such call requires the import and export of the electrostatics, which are more expensive, if induced dipoles are additionally present.

Finally, we examined the effect of varying the PMM convergence criterion. Tightening χ_{PMM} by one order of magnitude entails a slight increase of the numbers of PMM-SCF iteration steps, of DFT calls per integration step, and of DFT-SCF iteration steps. By contrast, loosening χ_{PMM} by one order of magnitude leads to a costly increase of the number of DFT-SCF iteration steps. Apparently, a correspondingly noisier PMM dipole environment hampers the DFT-SCF convergence.

In the DFT/PMM water box example discussed above, which features a small DFT fragment in a large PMM environment, the parallel DFT computation with eight cores requires about as much time as the single core PMM computation. This finding points to a technical deficiency of our current implementation, according to which the MPI-parallel version of the PMM-MD program IPHIGENIE is not yet compatible with the MPI-parallel version of the DFT program CPMD. This remaining deficiency will be removed in the near future.

VI. SUMMARY AND OUTLOOK

By suitably combining the (P)MM-MD program IPHIGENIE[47] with the DFT program CPMD,[48] we have developed and implemented a Hamiltonian DFT/(P)MM-MD approach, which conserves the energy as good as the well-established DFT-Born-Oppenheimer MD approach implemented in CPMD. Here, we took advantage of the fast multipole method $SAMM_4$ implemented in IPHIGENIE, by which one can treat the long-range electrostatic interactions within a (P)MM simulation system in a linearly scaling and Hamiltonian fashion. Artificial distortions of the DFT electron density are excluded by the use of Gaussian charges and induced dipoles in the boundary region between the DFT and PMM fragments. The accuracy and efficiency of the new DFT/PMM interface are supported by algorithmic improvements concerning the adaptive repositioning of the DFT box and the fine-tuning of the joint DFT- and PMM-SCF cycles. All these aspects were illustrated by relevant test simulations, which demonstrate that the new interface opens the way toward temporally extended DFT/PMM-MD simulations of large condensed phase systems at well-defined thermodynamic conditions.

Beyond the ongoing efforts of extending the implementation toward a jointly parallelized treatment of the DFT and (P)MM fragments, only one important issue remains to be addressed before applications can be tackled. It concerns the widths σ_i, which have to be chosen for the Gaussian distributions representing the (P)MM partial charges in the vicinity of the DFT fragment. As explained already by Laio *et al.*,[38] these widths are decisive parameters of the interface model and must be carefully determined for the various (P)MM atom types by sample simulations. Here, the value $\sigma_i = 0.57$ Å most likely is a sub-optimal choice.

Concerning the future development, we plan to reorganize the interface in such a way that also other grid-based DFT programs like CP2K[87] can be employed. Here, the integration of the multigrid DFT/MM electrostatic coupling of Laino *et al.*[39] into our near-field electrostatics computation should yield further efficiency gains. Furthermore, a combination of our DFT/PMM approach with the SCP-DFT method[32,33] implemented in CP2K may pave the way toward an improved modeling of polarization and dispersion interactions between the fragments.

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (SFB749/C4) and by the Kompetenznetzwerk für wissenschaftliches Höchstleistungsrechnen in Bayern of the Bayerische Staatsministerium für Wissenschaft, Forschung und Kunst (KONWIHR-III).

APPENDIX: MULTIPOLE MOMENTS $\tilde{\mathbf{M}}^{m,\mathbf{0}}$ OF DIPOLE DISTRIBUTIONS

Compact formulas, by which the totally symmetric and traceless mth order multipole tensors $\mathbf{M}^{m,\mathbf{0}}$ can be calculated for $m = 0, 1, 2, \ldots$, if the origin $\mathbf{0}$ of the Cartesian coordinate system is chosen as the reference point, have been given in Sec. 3 of the supporting information of Ref. 47 for distributions B of partial charges.

Choosing the same setting, the corresponding tensors

$$\tilde{\mathbf{M}}^{m,\mathbf{0}} = \sum_{j\in\tilde{B}} \tilde{\mathbf{M}}_j^{m,\mathbf{0}} \tag{A1}$$

of rank $m = 2, \ldots$ can be calculated for a distribution $\tilde{B}$ of point dipoles $\mathbf{p}_j$ at positions $\mathbf{r}_j$ from the recursion

$$\tilde{\mathbf{M}}_j^{m,\mathbf{0}} = \hat{S}_m \left[\left(\frac{2m^2 - m}{m-1} \right) \left(\mathbf{r}_j \otimes \tilde{\mathbf{M}}_j^{m-1,\mathbf{0}} \right) - m \left(\mathbf{r}_j \odot \tilde{\mathbf{M}}_j^{m-1,\mathbf{0}} \right) \otimes \mathbf{I} \right], \tag{A2}$$

where $\hat{S}_n$ is the symmetrization operator given in Eq. (22) of Ref. 47, where also the applied tensorial notation is explained in detail. For dipole distributions, the monopole moment ($m = 0$) vanishes, of course, and the recursion (A2) is initialized with the point dipole

$$\tilde{\mathbf{M}}_j^{1,\mathbf{0}} = \mathbf{p}_j. \tag{A3}$$

For $m = 2, 3, 4$, the Cartesian components of the tensors $\tilde{\mathbf{M}}^{m,\mathbf{0}}$ are explicitly given by

$$\tilde{M}_{\alpha\beta}^{2,\mathbf{0}} = \sum_{j \in \tilde{B}} [3 r_{j\alpha} p_{j\beta} + 3 r_{j\beta} p_{j\alpha} - 2\delta_{\alpha\beta} (\mathbf{r}_j \cdot \mathbf{p}_j)], \tag{A4}$$

$$\begin{aligned} \tilde{M}_{\alpha\beta\gamma}^{3,\mathbf{0}} = 3 \sum_{j \in \tilde{B}} [& 5 (p_{j\alpha} r_{j\beta} r_{j\gamma} + p_{j\beta} r_{j\gamma} r_{j\alpha} + p_{j\gamma} r_{j\alpha} r_{j\beta}) \\ & - r_j{}^2 (p_{j\alpha} \delta_{\beta\gamma} + p_{j\beta} \delta_{\gamma\alpha} + p_{j\gamma} \delta_{\alpha\beta}) \\ & - 2 (\mathbf{r}_j \cdot \mathbf{p}_j) (r_{j\alpha} \delta_{\beta\gamma} + r_{j\beta} \delta_{\gamma\alpha} + r_{j\gamma} \delta_{\alpha\beta})], \end{aligned} \tag{A5}$$

$$\begin{aligned} \tilde{M}_{\alpha\beta\gamma\epsilon}^{4,\mathbf{0}} = 3 \sum_{j \in \tilde{B}} \{ & 35 (p_{j\alpha} r_{j\beta} r_{j\gamma} r_{j\epsilon} + p_{j\beta} r_{j\gamma} r_{j\epsilon} r_{j\alpha} \\ & + p_{j\gamma} r_{j\epsilon} r_{j\alpha} r_{j\beta} + p_{j\epsilon} r_{j\alpha} r_{j\beta} r_{j\gamma}) \\ & - 5 r_j{}^2 [(p_{j\alpha} r_{j\epsilon} + p_{j\epsilon} r_{j\alpha}) \delta_{\beta\gamma} + (p_{j\beta} r_{j\epsilon} + p_{j\epsilon} r_{j\beta}) \delta_{\gamma\alpha} \\ & + (p_{j\gamma} r_{j\epsilon} + p_{j\epsilon} r_{j\gamma}) \delta_{\alpha\beta} + (p_{j\beta} r_{j\gamma} + p_{j\gamma} r_{j\beta}) \delta_{\alpha\epsilon} \\ & + (p_{j\alpha} r_{j\gamma} + p_{j\gamma} r_{j\alpha}) \delta_{\beta\epsilon} + (p_{j\alpha} r_{j\beta} + p_{j\beta} r_{j\alpha}) \delta_{\gamma\epsilon}] \\ & - 10 (\mathbf{r}_j \cdot \mathbf{p}_j) (r_{j\alpha} r_{j\epsilon} \delta_{\beta\gamma} + r_{j\beta} r_{j\epsilon} \delta_{\gamma\alpha} + r_{j\gamma} r_{j\epsilon} \delta_{\alpha\beta} \\ & + r_{j\beta} r_{j\gamma} \delta_{\alpha\epsilon} + r_{j\alpha} r_{j\gamma} \delta_{\beta\epsilon} + r_{j\alpha} r_{j\beta} \delta_{\gamma\epsilon}) \\ & + 4 (\mathbf{r}_j \cdot \mathbf{p}_j) r_j{}^2 (\delta_{\alpha\epsilon} \delta_{\beta\gamma} + \delta_{\beta\epsilon} \delta_{\gamma\alpha} + \delta_{\gamma\epsilon} \delta_{\alpha\beta}) \}. \end{aligned} \tag{A6}$$

The SAMM$_p$ algorithm treats these PMM multipole moments $\tilde{\mathbf{M}}^{m,\mathbf{0}}$ exactly like in the MM case, i.e., they are shifted to a different reference point using Eq. (19) of Ref. 47, and potentials and Taylor expansion coefficients are calculated from Eqs. (9) and (10) of Ref. 47, respectively. In IPHIGENIE, the electrostatics of static partial charges is calculated only once for each integration step and is reused during the PMM-SCF iteration.

[1] A. Warshel and M. Levitt, J. Mol. Biol. **103**, 227 (1976).
[2] H. M. Senn and W. Thiel, Angew. Chem., Int. Ed. **48**, 1198 (2009).
[3] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, J. Am. Chem. Soc. **117**, 5179 (1995).
[4] A. MacKerell, D. Bashford, M. Bellott, R. Dunbrack, J. Evanseck, M. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. Lau, C. Mattos, S. Michnick, T. Ngo, D. Nguven, B. Prodhom, W. Reiher, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus, J. Phys. Chem. B **102**, 3586 (1998).
[5] G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, and W. L. Jorgensen, J. Phys. Chem. B **105**, 6474 (2001).
[6] C. Oostenbrink, A. Villa, A. E. Mark, and W. F. Van Gunsteren, J. Comput. Chem. **25**, 1656 (2004).
[7] P. Tavan, H. Carstens, and G. Mathias, "Molecular dynamics simulations of proteins and peptides: Problems, achievements, and perspectives," in *Protein Folding Handbook*, edited by J. Buchner and T. Kiefhaber (Wiley-VCH, Weinheim, 2005), Vol. 1, pp. 1170–1195.
[8] M. A. Thompson and G. K. Schenter, J. Phys. Chem. **99**, 6374 (1995).
[9] M. A. Thompson, J. Phys. Chem. **100**, 14492 (1996).
[10] D. Bakowies and W. Thiel, J. Phys. Chem. **100**, 10580 (1996).
[11] J. Gao, J. Comput. Chem. **18**, 1061 (1997).
[12] D. P. Geerke, S. Thiel, W. Thiel, and W. F. van Gunsteren, J. Chem. Theory Comput. **3**, 1499 (2007).
[13] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[14] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[15] U. C. Singh and P. A. Kollman, J. Comput. Chem. **7**, 718 (1986).
[16] G. Jansen, F. Colonna, and J. G. Ángyán, Int. J. Quantum Chem. **58**, 251 (1996).
[17] M. J. Field, Mol. Phys. **91**, 835 (1997).
[18] I. H. Hillier, J. Mol. Struct.: THEOCHEM **463**, 45 (1999).
[19] H. Houjou, Y. Inoue, and M. Sakurai, J. Phys. Chem. B **105**, 867 (2001).
[20] L. Jensen, P. T. van Duijnen, and J. G. Snijders, J. Chem. Phys. **118**, 514 (2003).
[21] C. J. R. Illingworth, S. R. Gooding, P. J. Winn, G. A. Jones, G. G. Ferenczy, and C. A. Reynolds, J. Phys. Chem. A **110**, 6487 (2006).
[22] M. Wanko, M. Hoffmann, J. Fraehmcke, T. Frauenheim, and M. Elstner, J. Phys. Chem. B **112**, 11468 (2008).
[23] C. Curutchet, A. Muñoz-Losa, S. Monti, J. Kongsted, G. D. Scholes, and B. Mennucci, J. Chem. Theory Comput. **5**, 1838 (2009).
[24] R. A. Bryce, M. A. Vincent, N. O. J. Malcolm, I. H. Hillier, and N. A. Burton, J. Chem. Phys. **109**, 3077 (1998).
[25] M. S. Gordon, M. A. Freitag, P. Bandyopadhyay, J. H. Jensen, V. Kairys, and W. J. Stevens, J. Phys. Chem. A **105**, 293 (2001).
[26] M. Dupuis, M. Aida, Y. Kawashima, and K. Hirao, J. Chem. Phys. **117**, 1242 (2002).
[27] H. Li and M. S. Gordon, J. Chem. Phys. **126**, 124112 (2007).
[28] P. K. Biswas and V. Gogonea, J. Chem. Phys. **129**, 154108 (2008).
[29] F. Lipparini, C. Cappelli, G. Scalmani, N. De Mitri, and V. Barone, J. Chem. Theory Comput. **8**, 4270 (2012).
[30] Z. Lu and Y. Zhang, J. Chem. Theory Comput. **4**, 1237 (2008).
[31] E. Boulanger and W. Thiel, J. Chem. Theory Comput. **8**, 4527 (2012).
[32] K. A. Maerzke, G. Murdachaew, C. J. Mundy, G. K. Schenter, and J. I. Siepmann, J. Phys. Chem. A **113**, 2075 (2009).
[33] G. Murdachaew, C. J. Mundy, and G. K. Schenter, J. Chem. Phys. **132**, 164102 (2010).
[34] M. Eichinger, P. Tavan, J. Hutter, and M. Parrinello, J. Chem. Phys. **110**, 10452 (1999).
[35] M. Nonella, G. Mathias, and P. Tavan, J. Phys. Chem. A **107**, 8638 (2003).
[36] M. Schmitz and P. Tavan, J. Chem. Phys. **121**, 12247 (2004).
[37] V. Schultheis, R. Reichold, B. Schropp, and P. Tavan, J. Phys. Chem. B **112**, 12217 (2008).
[38] A. Laio, J. VandeVondele, and U. Rothlisberger, J. Chem. Phys. **116**, 6941 (2002).
[39] T. Laino, F. Mohamed, A. Laio, and M. Parrinello, J. Chem. Theory Comput. **1**, 1176 (2005).
[40] G. Babitzki, R. Denschlag, and P. Tavan, J. Phys. Chem. B **113**, 10483 (2009).
[41] G. Babitzki, G. Mathias, and P. Tavan, J. Phys. Chem. B **113**, 10496 (2009).
[42] B. Rieff, S. Bauer, G. Mathias, and P. Tavan, J. Phys. Chem. B **115**, 11239 (2011).
[43] M. Nonella, G. Mathias, M. Eichinger, and P. Tavan, J. Phys. Chem. B **107**, 316 (2003).
[44] M. Klähn, G. Mathias, C. Kotting, J. Schlitter, M. Nonella, K. Gerwert, and P. Tavan, J. Phys. Chem. A **108**, 6186 (2004).
[45] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[46] J. VandeVondele, P. Tröster, P. Tavan, and G. Mathias, J. Phys. Chem. A **116**, 2466 (2012).
[47] K. Lorenzen, M. Schwörer, P. Tröster, S. Mates, and P. Tavan, J. Chem. Theory Comput. **8**, 3628 (2012).

[48]J. Hutter, A. Alavi, T. Deutsch, M. Bernasconi, S. Goedecker, D. Marx, M. Tuckerman, and M. Parrinello, CPMD: Car–Parinello Molecular Dynamics, Version 3.15.1, © IBM Corp, 1990–2008 and MPI für Festkörperforschung Stuttgart, 1997–2001, see www.cpmd.org.
[49]J. Applequist, J. R. Carl, and K.-K. Fung, J. Am. Chem. Soc. **94**, 2952 (1972).
[50]P. Ahlström, A. Wallquist, S. Engström, and B. Jönsson, Mol. Phys. **68**, 563 (1989).
[51]P. E. Lopes, B. Roux, and A. D. MacKerell, Jr., Theor. Chem. Acc. **124**, 11 (2009).
[52]B. T. Thole, Chem. Phys. **59**, 341 (1981).
[53]D. Elking, T. Darden, and R. J. Woods, J. Comput. Chem. **28**, 1261 (2007).
[54]S. Russell and A. Warshel, J. Mol. Biol. **185**, 389 (1985).
[55]D. V. Belle, I. Couplet, M. Prevost, and S. J. Wodak, J. Mol. Biol. **198**, 721 (1987).
[56]P. K. Biswas and V. Gogonea, J. Chem. Phys. **123**, 164114 (2005).
[57]M. P. Allen and D. Tildesley, in *Computer Simulations of Liquids* (Clarendon, Oxford, 1987), Chap. 1.5.2, pp. 24–27.
[58]C. Niedermeier and P. Tavan, J. Chem. Phys. **101**, 734 (1994).
[59]C. Niedermeier and P. Tavan, Mol. Simul. **17**, 57 (1996).
[60]G. Mathias, B. Egwolf, M. Nonella, and P. Tavan, J. Chem. Phys. **118**, 10847 (2003).
[61]A. A. Appel, SIAM (Soc. Ind. Appl. Math.) J. Sci. Stat. Comput. **6**, 85 (1985).
[62]J. Barnes and P. Hut, Nature (London) **324**, 446 (1986).
[63]L. Greengard and V. Rokhlin, J. Comput. Phys. **73**, 325 (1987).
[64]W. Dehnen, J. Comput. Phys. **179**, 27 (2002).
[65]M. Eichinger, H. Grubmüller, H. Heller, and P. Tavan, J. Comput. Chem. **18**, 1729 (1997).
[66]D. Carfi, AAPP Phys. Math. Nat. Sci. **88**, C1A1001004 (2010).
[67]A. D. Buckingham, Adv. Chem. Phys. **12**, 107 (1967).
[68]K. Hinsen and B. U. Felderhof, J. Math. Phys. **33**, 3731 (1992).
[69]W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C* (Cambridge University Press, Cambridge, 1988), Chap. 3.2, pp. 118–120.
[70]P. Pulay, Chem. Phys. Lett. **73**, 393 (1980).
[71]J. Hutter, H. P. Lüthi, and M. Parrinello, Comput. Mater. Sci. **2**, 244 (1994).
[72]L. Verlet, Phys. Rev. **159**, 98 (1967).
[73]V. Kräutler, W. F. van Gunsteren, and P. Hünenberger, J. Comput. Chem. **22**, 501 (2001).
[74]A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
[75]J. Perdew and W. Yue, Phys. Rev. B **33**, 8800 (1986).
[76]N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).
[77]W. S. Benedict, N. Gailar, and E. K. Plyler, J. Chem. Phys. **24**, 1139 (1956).
[78]K. Ichikawa, Y. Kameda, T. Yamaguchi, H. Wakita, and M. Misawa, Mol. Phys. **73**, 79 (1991).
[79]W. F. Murphy, J. Chem. Phys. **67**, 5877 (1977).
[80]T. Dyke and J. Muenter, J. Chem. Phys. **59**, 3125 (1973).
[81]R. A. Buckingham and J. Corner, Proc. R. Soc. London, Ser. A **189**, 118 (1947).
[82]G. Kell, J. Chem. Eng. Data **12**, 66 (1967).
[83]G. Bussi, D. Donadio, and M. Parrinello, J. Chem. Phys. **126**, 014101 (2007).
[84]J. L. F. Abascal and C. Vega, J. Chem. Phys. **123**, 234505 (2005).
[85]M. E. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation*, 1st ed. (Oxford University Press, New York, 2010), Chap. 3.13, pp. 121–124.
[86]See supplementary material at http://dx.doi.org/10.1063/1.4811292 for a total of four figures (S9– S12) and three equations (S39– S41) as additional material to the results of the (V) test simulations, on five pages in three sections.
[87]J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, Comput. Phys. Commun. **167**, 103 (2005).
[88]G. Mathias and M. D. Baer, J. Chem. Theory Comput. **7**, 2028 (2011).
[89]J. Hutter and A. Curioni, Parallel Comput. **31**, 1 (2005).
[90]P. Tröster, K. Lorenzen, M. Schwörer, and P. Tavan, "Polarizable water models from mixed computational and empirical optimization" J. Phys. Chem. B (to be published).