Computation of Forces arising from the Polarizable Continuum Model within the Domain-Decomposition Paradigm

Paolo Gatto,[1] Filippo Lipparini,[1] and Benjamin Stamm[1]

Within implicit solvation models, the domain-decomposition strategy for the computation of the electrostatic energy due to the solvent based on the Polarizable Continuum Model (PCM) has recently been developed. The methodological development started with the so-called ddCOSMO method and has recently be generalized to the PCM equation resulting in the ddPCM-method [Stamm *et al.*, J. Chem. Phys. 144, 054101 (2016)] for which derive the forces within this article. We show the derivation of the forces and derive an efficient implementation followed by numerical tests.

I. INTRODUCTION

The effects of solvation play a crucial role in many processes in Chemistry and Biochemistry. While it is unfeasible to always take explicitly into account the individual solvent molecules in a solute/solvent system, in particular when a quantum mechanical description of the molecular system is considered, solvation effects can not be neglected altogether. Thus, a compromise is to resort to a limited description of the solvent, as the solvent *per se* is not of interest, but, rather, its effect on the solute molecule is.

Polarizable Continuum Solvation Models (PCSM's)[1–18] focus on the electrostatic interaction between the solute and the solvent by replacing the solvent with an infinite continuum with dielectric permittivity that matches the bulk permittivity of the solvent. Such approach turns out to be a good compromise between accuracy and the cost to compute electric, magnetic, vibrational and mixed properties[1,5,19]. As a result, PCSM's are nowadays available in most quantum Chemistry codes, and have become a standard tool among an extended community.

Standard implementations of PCSM's usually employ the Boundary Element Method[1,20–24] (BEM) to numerically solve a discretized integral equation. This requires the solution of a linear system whose size scales linearly, although with a large proportionality constant, with respect to the number of atoms. This task is usually carried out through standard dense linear algebra techniques, such as the LU decomposition[25], which require a computational effort of cubic complexity, with respect to the size of the system. Consequently, the solution step can rapidly become demanding when dealing with systems as large as those treated via QM/MM methods, and one should seek other strategies. An alternative is provided by iterative techniques, which reduce the computational cost to that of several matrix-vector multiplications[26], which is, in the general case, quadratic. When fast summation techniques, e.g., the Fast Multipole Method (FMM)[27], can be employed, the computational complexity can be further reduced. Thus, it is possible to solve the PCSM linear equations in a number of floating point operations which is linear with respect to the number of atoms. Nevertheless, the solution of the PCSM equations can still represent a formidable bottleneck for large systems[28], especially when repeated computations are required for statistical sampling purposes or time-dependent simulations.

In recent years, we have proposed an alternative approach for PCSM's which is based

on domain-decomposition. We introduced a novel strategy[29], referred to as ddCOSMO[28–31], to solve the PCSM equation for the Conductor-like Screening Model[8] (COSMO) in combination with Van der Waals molecular cavities. As a first step, the COSMO equation in its differential form is rewritten as a system of coupled linear differential equations on each sphere, where the coupling occurs only between overlapping spheres. Secondly, each differential equation is recast as an integral equation, so that it can be efficiently solved by using a (truncated) expansion of spherical harmonics[29]. The discretization produces a block-sparse linear system[30], where only the blocks corresponding to overlapping spheres are nonzero. This structure allows for a computational cost that scales linearly with respect to the number of atoms, and is overall very small as compared to competing techniques. Indeed, as shown in[28], as many as two or three orders of magnitude are gained by the ddCOSMO approach.

More recently, the method has been generalized to the Polarized Continuum Model (PCM) equation, which assumes a finite permittivity of the solvent. This resulted in the ddPCM method[32], which is based on the same domain-decomposition approach. While the method has been developed in[32] to compute the electrostatic contribution to the solvation energy, the aim of this article is to present the derivation of analytical forces of the ddPCM solvation energy.

This paper is organized as follows. Section II reviews the ddPCM and ddCOSMO methods that we have previously developed. In Section III we describe the derivation of the ddPCM forces and discuss their efficient implementation. Section IV is devoted to numerical experiments. Finally, in Section V we draw conclusions from the presented work and point to possible future directions of research.

II. A BRIEF REVIEW OF THE DDPCM STRATEGY

A. The Polarizable Continuum Model

The foundation of Polarizable Continuum Solvation Models (PCSM's) is the assumption that in a solute/solvent system the solvent on the outside of the molecular cavity Ω occupied by the solute can be treated as either a dielectric, or a conducting continuum medium. We follow the customary approach of taking the cavity to be the so-called Van der Waals cavity[1],

i.e., the union of spheres centered at each atom with radii coinciding with the van der Waals radii. We mention that the, topologically similar, Solvent Accessible Surface (SAS) cavity can be treated through this approach as well. Models based on the Solvent Excluded Surface (SES) have recently been proposed[34? –36], however are not considered here.

The electrostatic part of the solute/solvent interaction is given by

$$E_s = \tfrac{1}{2}\, f(\varepsilon_s) \int_\Omega \rho(x) W(x)\, dx$$

where $f(\varepsilon_s)$ is an empirical scaling that depends on the dielectric constant ε_s of the solvent, e.g., $f(\varepsilon_s) = (\varepsilon_s - 1)/\varepsilon_s$ for COSMO and $f(\varepsilon_s) = 1$ for PCM, ρ is the charge density of the solute, and W is the polarization potential of the solvent. The quantities W and E_s are usually referred to, respectively, as the reaction potential and the electrostatic contribution to the solvation energy.

The reaction potential is defined as $W = \varphi - \Phi$, where φ is the total electrostatic potential of the solute/solvent system and Φ is the potential of the solute *in vacuo*. In the case of the PCM, the total potential φ satisfies a (generalized) Poisson equation with suitable interface conditions[11,12]. Indeed, if ε_s is the macroscopic, zero-frequency relative dielectric permittivity of the solvent, and define $\varepsilon(x) = 1$ when $x \in \Omega$ and $\varepsilon(x) = \varepsilon_s$ otherwise, the reaction potential fulfills

$$\begin{cases} \Delta W = 0 & \text{in } \mathbb{R}^3 \setminus \Gamma \\ [\![W]\!] = 0 & \text{on } \Gamma \\ [\![\varepsilon\, \partial_\nu W]\!] = (\varepsilon_s - 1)\, \partial_\nu \Phi & \text{on } \Gamma \end{cases} \tag{1}$$

Here $\Gamma = \partial\Omega$ is the boundary of the cavity, ∂_ν is the normal derivative on Γ, and $[\![\cdot]\!]$ is the jump operator (inside minus outside) on Γ.

Recalling potential theory, see, e.g.,[?] , W can be represented as $W(x) = (\tilde{\mathcal{S}}\sigma)(x)$ when $x \in \mathbb{R}^3 \setminus \Gamma$, or $W(s) = (\mathcal{S}\sigma)(s)$ when $s \in \Gamma$. The surface density σ defined on Γ is the so-called apparent surface charge, $\tilde{\mathcal{S}}$ is the single layer potential, and $\mathcal{S}$ is the single layer operator, which is invertible[37]. Both $\tilde{\mathcal{S}}$ and $\mathcal{S}$ are understood as relative to Γ. It can be shown that σ satisfies the equation $\sigma = 1/4\pi\, [\![\partial_\nu W]\!]$, so that it is possible to recast the PCM problem (1) as a single integral equation for σ. In fact, if we define the operators

$$\mathcal{R}_\varepsilon = 2\pi\, \frac{\varepsilon_s + 1}{\varepsilon_s - 1}\, \mathcal{I} - \mathcal{D} \qquad , \qquad \mathcal{R}_\infty = 2\pi\, \mathcal{I} - \mathcal{D} \tag{2}$$

where $\mathcal{I}$ is the identity and $\mathcal{D}$ is the double layer operator, also relative to Γ, it can be shown[1] that the apparent surface charge satisfies

$$\mathcal{R}_\varepsilon \mathcal{S}\, \sigma = -\mathcal{R}_\infty \Phi \qquad \text{on } \Gamma \tag{3}$$

Henceforth, for ease of notation, we set $g(\varepsilon_s) = (\varepsilon_s + 1)/(\varepsilon_s - 1)$. This is known as the IEF-PCM equation, and it involves operators $\mathcal{R}_\infty$ and $\mathcal{R}_\varepsilon$, which are both invertible. When the dielectric constant ε_s approaches infinity, the IEF-PCM equation simplifies to $\mathcal{S}\, \sigma = -\Phi$ on Γ, which is the Integral Equation Formulation of the Conductor-like Screening Model (COSMO)[16].

B. The ddPCM-method

We recall how to solve the IEF-PCM boundary integral equation (3) within the domain-decomposition paradigm. As a preliminary step, we set $\Phi_\varepsilon = \mathcal{S}\, \sigma$ and write (3) as a succession of two integral equations, the latter of which is equivalent to the COSMO equation[38], namely

$$\mathcal{R}_\varepsilon\, \Phi_\varepsilon = \mathcal{R}_\infty\, \Phi \qquad \text{on } \Gamma \tag{4}$$

$$\mathcal{S}\, \sigma = -\Phi_\varepsilon \qquad \text{on } \Gamma \tag{5}$$

Indeed, (5) is a COSMO equation with the modified potential Φ_ε in place of the potential Φ. This allows to develop the ddPCM strategy as an extension of the ddCOSMO approach. First, equation (4) is solved in order to compute the right-hand-side $-\Phi_\varepsilon$ of equation (5); secondly, ddCOSMO is employed to solve equation (5) with the modified potential $-\Phi_\varepsilon$, and compute the solvation energy E_s.

In order to discuss the domain-decomposition approach employed for both steps, let us introduce some notation. As anticipated, we take the cavity Ω be the union of M spheres $\Omega_j = B(x_j, r_j)$ with boundaries Γ_j. We define $\Gamma_j^{\text{ext}} := \Gamma_j \cap \Gamma$ and $\Gamma_j^{\text{int}} := \Gamma_j \cap \Omega$, and let $U_j : \Gamma_j \to \mathbb{R}$ be the characteristic function of Γ_j^{ext}. Let $\Phi_j : \Gamma_j \to \mathbb{R}$ and $\Phi_{\varepsilon,j} : \Gamma_j \to \mathbb{R}$ be, respectively, the local *trivial* extensions of Φ and Φ_ε to Γ_j. Then, it is immediate to conclude that

$$(\mathcal{D}\, \Phi)(s) = (\mathcal{D}_j\, \Phi_j)(s) + \sum_{k \neq j} (\tilde{\mathcal{D}}_k\, \Phi_k)(s) \qquad ; \qquad s \in \Gamma_j^{\text{ext}} \quad , \quad j = 1, \ldots, M \tag{6}$$

where $\mathcal{D}_j$ and $\tilde{\mathcal{D}}_j$ are, respectively, the local double layer operator and the local double layer potential relative to the local sphere Γ_j. An analogous result holds for Φ_ε and its local extensions $\Phi_{\varepsilon,j}$.

Step 1. We localize the integral equation (4) to Γ_j through the characteristic function U_j as

$$2\pi\, g(\varepsilon_s)\, U_j\, \Phi_\varepsilon - U_j\, \mathcal{D}\, \Phi_\varepsilon = 2\pi\, U_j\, \Phi - U_j\, \mathcal{D}\, \Phi \qquad \text{on } \Gamma_j \tag{7}$$

for each $j = 1, \ldots, M$. Without loss of generality, we trade $U_j\, \Phi_\varepsilon$ for $U_j\, \Phi_{\varepsilon,j}$ in the left-hand-side of (7). We ensure that $\Phi_{\varepsilon,j}$ indeed vanishes on Γ_j^{int} by imposing the additional constraint

$$(1 - U_j)\, \Phi_{\varepsilon,j} = 0 \qquad \text{on } \Gamma_j \tag{8}$$

In order to obtain a single equation, we multiply (8) by the factor $2\pi\,(\varepsilon + 1)/(\varepsilon - 1)$, and add it sidewise to (7), so that

$$2\pi\, g(\varepsilon_s)\, \Phi_{\varepsilon,j} - U_j\, \mathcal{D}\, \Phi_\varepsilon = 2\pi\, \Phi_j - U_j\, \mathcal{D}\, \Phi \qquad \text{on } \Gamma_j \tag{9}$$

When s belongs to Γ_j^{int}, we recover $\Phi_{\varepsilon,j}(s) = 0$, so that we have effectively built the constraint (8) into the previous equation. Thus, recalling that Φ_j vanishes on Γ_j^{int} by definition, we proceed to apply the decomposition (6) to both sides of (9), and obtain

$$2\pi\, g(\varepsilon_s)\, \Phi_{\varepsilon,j} - U_j \left(\mathcal{D}_j\, \Phi_{\varepsilon,j} + \sum_{k \neq j} \tilde{\mathcal{D}}_k\, \Phi_{\varepsilon,k} \right) = 2\pi\, \Phi_j - U_j \left(\mathcal{D}_j\, \Phi_j + \sum_{k \neq j} \tilde{\mathcal{D}}_k\, \Phi_k \right) \qquad \text{on } \Gamma_j \tag{10}$$

which constitutes our domain-decomposition strategy for equation (4). The summation implies that every subdomain Ω_j interacts with all other subdomains. We anticipate that this contrasts with the ddCOSMO strategy for equation (5), which only involves neighbor-to-neighbor interactions.

Step 2. We note that $W = \tilde{\mathcal{S}}\, \sigma$, where σ solves the integral equation (5), solves the equation $\Delta W = 0$ in Ω with boundary condition $W = -\Phi_\varepsilon$ on $\Gamma = \partial\Omega$. The restriction $W_j := W|_{\overline{\Omega}_j}$ is harmonic over the subdomain Ω_j, thus it can be represented locally as

$$W_j(x) = (\tilde{\mathcal{S}}_j\, \sigma_j)(x) \quad , \quad x \in \Omega_j \qquad ; \qquad W_j(s) = (\mathcal{S}_j\, \sigma_j)(s) \quad , \quad s \in \Gamma_j \tag{11}$$

where σ_j is an unknown surface density on Γ_j, and $\mathcal{S}_j$ and $\tilde{\mathcal{S}}_j$ are, respectively, the single layer potential and the single layer operator on Γ_j. The local problems (11) are coupled

together by imposing on Γ_j the coupling condition

$$W_j(s) = -\Phi_{\varepsilon,j}(s) + n_j(s) \sum_{k \in N_j} W_k(s) \qquad ; \qquad s \in \Gamma_j \quad , \quad j = 1, \ldots, M \tag{12}$$

where N_j is the set of all neighboring subdomains of Ω_j, W_k is understood as its trivial extension to Ω, and n_j is a normalization factor defined as follows. If s does not belong to any neighbor of Ω_j, then $n_j(s)$ vanishes. Otherwise, $n_j(s)$ is the reciprocal of the number of neighbors. The decomposition (12) also employs the fact that $\Phi_{\varepsilon,j}$ vanishes on Γ_j^{int}. When we substitute the local problems (11) into the decomposition (12), and define $\tilde{\mathcal{S}}_{jk}\, \sigma_k = n_j\, \mathcal{S}_k\, \sigma_k$, we obtain

$$\mathcal{S}_j\, \sigma_j - \sum_{k \in N_j} \tilde{\mathcal{S}}_{jk}\, \sigma_k = -\Phi_{\varepsilon,j} \qquad \text{on } \Gamma_j \tag{13}$$

As opposed to the local problem (10) which features a global interaction of all subdomains, the ddCOSMO step (13) is characterized by the interaction of subdomain Ω_j with only its neighbors. This results in a sparse, rather than dense, discrete operator. We anticipate that this is the key feature that allows to compute the ddCOSMO forces within linear complexity, with respect to the number of atoms M. More details are provided in Section III.

C. Numerical Discretization

The first step to discretize equations (10) and (13) is to expand Φ_j, $\Phi_{\varepsilon,j}$ and σ_j as truncated series of spherical harmonics. Let Y_ℓ^m be the spherical harmonic of degree ℓ and order m on the unit sphere $\mathbb{S}$, and let y be the variable on $\mathbb{S}$. For a prescribed integer parameter $L_{\max}$, we approximate the surface charge σ_j as the truncated expansion

$$\sigma_j(s) = \sigma_j(x_j + r_j y) = \frac{1}{r_j} \sum_{\ell=0}^{L_{\max}} \sum_{m=-\ell}^{\ell} [X_j]_\ell^m\, Y_\ell^m(y)$$

for some unknown coefficients $X = [X_j]_\ell^m$. The choice of this ansatz implies a spectral-like numerical method. The scaling factor has been introduced for convenience, as seen from the derivation in Appendix B. We approximate $\Phi_{\varepsilon,j}$ and Φ_j in the same fashion, namely

$$\Phi_{\varepsilon,j}(s) = -\sum_{\ell=0}^{L_{\max}} \sum_{m=-\ell}^{\ell} [G_j]_\ell^m\, Y_\ell^m(y) \qquad , \qquad \Phi_j(s) = -\sum_{\ell=0}^{L_{\max}} \sum_{m=-\ell}^{\ell} [F_j]_\ell^m\, Y_\ell^m(y) \tag{14}$$

where $G = [G_j]_\ell^m$ and $F = [F_j]_\ell^m$ are the coefficients of the expansions. The minus signs have been introduced so that there is no negative sign in the discretization of the right-hand-side of the ddCOSMO step (13).

We interpret the local problems (10) and (13) in a variational setting that uses spherical harmonics as test functions. Numerical discretizations are obtained by employing the orthogonality of the spherical harmonics, along with Lebedev grids to perform quadrature, see Appendix A and B. In the spirit of domain-decomposition, we combine together the discretizations of the local problems to obtain those of the global problems (4) and (5). Respectively, we obtain the linear systems

$$A_\varepsilon\, G = A_\infty\, F \qquad , \qquad L\, X = G \tag{15}$$

where the discrete operators A_ε and L are known in closed form, see (A3), (A4), and (B2), (B3). We remark that A_ε is a dense operator, while L is a sparse operator. In fact, for a fixed index j, the block L_{jk} is *a priori* nonzero only when $k = j$ or $k \in N_j$. Since Ω_j is a neighbor of Ω_k if and only if Ω_k is a neighbor of Ω_j, then L_{jk} is *a priori* nonzero if and only if L_{kj} is. We conclude that L has a symmetric block-structure, although, in general, L is non symmetric.

The dependency of operators A_ε and L upon the nuclear positions is evident. Although more subtle, the load vector F depends on the nuclear positions as well. Indeed, the second one of the expansions (14) and the fact that Φ_j is the trivial extension of Φ, i.e., $\Phi_j = U_j\, \Phi$ imply

$$[F_j]_\ell^m = -\int_{\mathbb{S}} \Phi_j(s(y))\, Y_\ell^m(y)\, dy = -\int_{\mathbb{S}} U_j(s(y))\, \Phi(s(y))\, Y_\ell^m(y)\, dy \tag{16}$$

Since the characteristic function U_j depends upon x_j and x_i such that $i \in N_j$, see Appendix C, so does $[F_j]_\ell^m$. (Paolo: should also comment on $\nabla\Phi = -\boldsymbol{E}$...) Those dependencies upon the nuclear positions are needed in the computation of the ddPCM-forces, discussed in the following Section.

III. COMPUTATION OF THE DDPCM-FORCES

A. Theoretical Derivation

The solvation energy can be written as a sum of subdomain contributions, which perfectly fits the domain-decomposition paradigm. (Paolo: this sentence should be moved...)

First, for a classical solute's charge distribution of the form of $\rho = \sum_j q_j\, \delta_{x_j}$, we can

manipulate the energy as

$$E_s = \tfrac{1}{2}\, g(\varepsilon_s) \int_\Omega \rho(x)\, W(x)\, dx = \tfrac{1}{2}\, g(\varepsilon_s) \sum_j q_j\, W(x_j) = g(\varepsilon_s)\, \sqrt{\pi} \sum_j q_j\, [X_j]_0^0$$

Indeed, the rescaling property $(\tilde{\mathcal{S}}_j\, \sigma_j)(x) = r_j (\tilde{\mathcal{S}}\, \hat{\sigma}_j)(u)$, where $u = (x - x_j)/r_j$, and the Addition Theorem for spherical harmonics, imply that

$$W_j(x) = (\tilde{\mathcal{S}}_j\, \sigma_j)(x) = (\tilde{\mathcal{S}}\, \hat{\sigma}_j)(u) = \sum_{\ell,m} \frac{4\pi}{2\ell+1}\, |u|^\ell\, Y_\ell^m(u/|u|)\, [X_j]_\ell^m$$

so that $W(x_j) = W_j(x_j) = 2\sqrt{\pi}\, [X_j]_0^0$, since $Y_0^0 = 1/\sqrt{4\pi}$. We recall that $g(\varepsilon_s) = 1$ for PCM, so that we shall henceforth omit it.

This concept easily generalizes to point multipolar charge distributions. The evaluation of the energy for charge distributions is a bit more involved as it requires a three-dimensional integration and we refer to[31] for more details. In all cases however, the energy can be written as

$$E_s = \tfrac{1}{2} \sum_j \sum_{\ell,m} [\Psi_j]_\ell^m [X_j]_\ell^m =: \tfrac{1}{2}\, \langle \Psi, X \rangle$$

where the angular brackets indicate the double scalar product over j and ℓ, m. For example, if we define

$$[\Psi_j]_\ell^m = \sqrt{\pi}\, q_j\, \delta_{\ell 0}\, \delta_{m 0}$$

we recover the case of the classical charge above. (Paolo: isn't there a factor 2 missing?)

The force acting on the i-th atom is then given by

$$\mathcal{F}_i = -\nabla_i\, E_s = -\tfrac{1}{2}\, \langle \Psi, \nabla_i\, X \rangle$$

where the gradient is understood with respect to x_i, and we employed the fact that Ψ is independent of the atomic positions. On the other hand, since the operators A_ε and L, along with the right-hand-side F, depend on the nuclear positions, so does the solution X of (15). Analogously to the ddCOSMO forces, we proceed to remove the derivative from the unknown X through the solution of the adjoint problem $(A_\varepsilon\, L)^* s = \Psi$, and Leibnitz differentiation rule.

As a first step, differentiation of the COSMO equation $L\, X = G$ yields

$$\nabla_i\, L\, X + L\, \nabla_i\, X = \nabla_i\, G$$

so that we can address the term $L\,\nabla_i\,X$ in the following expression for the force

$$\langle \Psi, \nabla_i\, X\rangle = \langle s, A_\varepsilon\, L\,\nabla_i\, X\rangle = \langle s, A_\varepsilon\,\nabla_i\, G\rangle - \langle s, A_\varepsilon\,\nabla_i\, L\, X\rangle$$

Secondly, to deal with the term $A_\varepsilon\,\nabla_i\, G$, which involves the derivative of the solution to the intermediate PCM step, we differentiate the equation $A_\varepsilon\, G = A_\infty\, F$, namely

$$\nabla_i\, A_\varepsilon\, G + A_\varepsilon\,\nabla_i\, G = \nabla_i\, A_\infty\, F + A_\infty\,\nabla_i\, F$$

The quantity $\nabla_i\, F$ is *a priori* nonzero since F does depend upon the nuclear positions through the characteristic functions, recall (16). Since the derivative of A_ε is independent of ε, we just write $\nabla_i\, A$ and obtain

$$\langle \Psi, \nabla_i\, X\rangle = \langle s, \nabla_i\, A(F-G)\rangle + \langle s, A_\infty\,\nabla_i\, F\rangle - \langle s, A_\varepsilon\,\nabla_i\, L\, X\rangle$$

The solution of the adjoint problem $(A_\varepsilon\, L)^* s = \Psi$ requires two steps, namely $L^*\, y = \Psi$ and ${A_\varepsilon}^*\, s = y$, so that we can simplify the last term on the right-hand-side, and obtain

$$\langle \Psi, \nabla_i\, X\rangle = \langle s, \nabla_i\, A(F-G)\rangle + \langle {A_\infty}^*\, s, \nabla_i\, F\rangle - \langle y, \nabla_i\, L\, X\rangle$$

As a last step, we write ${A_\infty}^*\, s = (A_\infty - A_\varepsilon)^*\, s + y$ and observe that $A_\infty - A_\varepsilon$ is a multiple of the identity, which implies that the adjoint is redundant. The final expression becomes

$$\langle \Psi, \nabla_i\, X\rangle = \langle s, \nabla_i\, A(F-G)\rangle + \langle (A_\infty - A_\varepsilon)\, s, \nabla_i\, F\rangle + \langle y, \nabla_i\, F\rangle - \langle y, \nabla_i\, L\, X\rangle$$

where the last two terms coincide with the COSMO force.

We remark that the PCM-forces we just derived are a perturbation of the COSMO-forces presented in[30]. Indeed, when ε approaches infinity, then $G = F$ and $A_\varepsilon = A_\infty$, so that we recover

$$\langle \Psi, \nabla_i\, X\rangle = \langle y, \nabla_i\, F\rangle - \langle y, \nabla_i\, L\, X\rangle$$

We recall that, because of the sparsity of L, and, consequently, its derivative $\nabla_i\, L$, the contraction $\langle y, \nabla_i\, L\, X\rangle$ can be computed within a complexity that depends upon $L_{\max}$, N_{grid} and the number of neighbors N_i, but is independent of M. In other words, it is $O(1)$ with respect to the number of atoms. A similar analysis holds for the contraction $\langle y, \nabla_i\, F\rangle$ (Paolo: add comment, maybe...). As reported in[30] for COSMO, the computation of all the M forces acting on all atoms is an operation of complexity $O(M)$, with a prefactor that depends at least upon $L_{\max}$, N_{grid}, and the average number of neighbors of each atom. As we shall see below, this contrasts with the computation of the PCM-forces, which has complexity $O(M^2)$.

B. Efficient Implementation

Since the PCM forces are an extension of the COSMO forces, we rely on the analysis in[30] for the COSMO contribution, and only need to discuss the additional term $\langle s, \nabla_i A(F-G)\rangle$. As previously remarked, the operator $A_\infty - A_\varepsilon$ is a rescaled identity, so that the contraction $\langle (A_\infty - A_\varepsilon)\, s, \nabla_i F\rangle$ is completely analogous to $\langle y, \nabla_i F\rangle$ and can be computed in $O(1)$ operations, for each i.

We proceed to establish the complexity of the contraction $\langle s, \nabla_i A\, X\rangle$, where we traded $F - G$ for an arbitrary vector X, by employing the sparsity of the derivative $\nabla_i A$. The entries of A_ε are derived in closed for in Appendix A. The definition (A3) of the diagonal blocks implies that A_{jj} depends upon the nuclear positions x_j and x_i such that $i \in N_j$. Similarly, if we recall (A4), the off-diagonal block A_{jk} depends upon x_j, x_k, and x_i such that $i \in N_j$. This implies that $\nabla_i A_{jj}$ is *a priori* nonzero only when $i = j$, or $i \in N_j$. Analogously, $\nabla_i A_{jk}$ is *a priori* nonzero only when $i = j$, or $i = k$, or $i \in N_j$. In order to restate those conditions for a *fixed* index i, and variable indices j and k, we make the trivial observation that $i \in N_j$ if and only if $j \in N_i$. We obtain that $\nabla_i A_{jj}$ is *a priori* nonzero only when $j = i$, or $j \in N_i$, and $\nabla_i A_{jk}$ is *a priori* nonzero only when $j = i$, or $k = i$, or $j \in N_i$.

We employ the previous vanishing conditions to compute the contraction $\langle s, \nabla_i A\, X\rangle$. We rearrange the summations as

$$\sum_{k,j} s_j\, \nabla_i A_{jk}\, X_k = s_i \sum_k \nabla_i A_{ik}\, X_k + X_i \sum_{j\neq i} s_j\, \nabla_i A_{ji} + \sum_{k\neq i} X_k \sum_{j\neq i} s_j\, \nabla_i A_{jk}$$

Each term on the right-hand-side that involves a single summation requires a $O(M)$ computation cost, with a prefactor that depends at least upon $L_{\max}$ and N_{grid}. The range of index j in the double summation term simplifies to N_i because of the vanishing conditions. Thus, its computation requires $O(M)$ operations as well, although with a prefactor that also depends upon the average number of neighbors.

The cost of computing the force acting on the i-th atom is dominated by the cost of the contraction $\langle s, \nabla_i A(F-G)\rangle$. We conclude that the computation cost of the M contractions needed for the PCM forces acting on each atom is $O(M^2)$, with the prefactor that depends at least upon $L_{\max}$, N_{grid}, and the average number of neighbors of each atom.

IV. NUMERICAL EXPERIMENTS

A. Convergence tests

We verified the implementation of the forces by performing a convergence test of a first order finite difference approximation. Let e_α, $\alpha = 1, 2, 3$, be the canonical unit vectors of $\mathbb{R}^3$, and define the forward finite difference

$$F_{i,\alpha}(\delta) = \frac{E_s(x_1, \ldots, x_i + \delta e_\alpha, \ldots, x_M) - E_s(x_1, \ldots, x_i, \ldots, x_M)}{\delta}$$

where we have made explicit the dependency of the solvation energy on the nuclear positions $x_1, \ldots, x_M$. It immediately follows that $\mathcal{F}_{i,\alpha} = F_{i,\alpha}(\delta) + O(\delta)$, where $\mathcal{F}_{i,\alpha}$ is the α component of $\mathcal{F}_i$, which implies that the relative error $\mathrm{Err}_{i,\alpha}(\delta) = |(\mathcal{F}_{i,\alpha} - F_{i,\alpha}(\delta))/\mathcal{F}_{i,\alpha}|$ decreases as $O(\delta)$, namely with rate 1.

As a first test, we investigated the rate of convergence for a molecular configuration composed of six spheres with radius 1.5 and centers at $x_{\pm\alpha} = \pm e_\alpha$, for $\alpha = 1, 2, 3$. Although conceptually simple, this configuration generates a sextuple intersection which provides a challenging benchmark case. We have studied the behavior of the relative error over the range of angular momenta $\ell = 2, \ldots, 10$, and obtained numerical results that are qualitatively similar, and in excellent agreement with the predicted rate of convergence. Results for a representative case are reported in Table I.

In table **??**, we illustrate the RMS of the approximate forces with $\delta_n = x^{-n}$ for caffeine.

...

B. Timings

Timings for different molecular structures depending on the number of atoms (i.e. alanine chains, hemoglobin, etc).

V. CONCLUSIONS

ACKNOWLEDGMENTS

P.G. and B.S. acknowledge funding from the German Academic Exchange Service (DAAD) from funds of the "Bundesministeriums für Bildung und Forschung" (BMBF)

Derivative	$\delta = \delta_0$	$\delta = \delta_0/2$		$\delta = \delta_0/4$		$\delta = \delta_0/8$	
	Error	*Error*	*Rate*	*Error*	*Rate*	*Error*	*Rate*
$x_{+1,1}$	0.23529E-02	0.11751E-02	-1.002	0.58717E-03	-1.001	0.29345E-03	-1.001
$x_{-1,1}$	0.23529E-02	0.11751E-02	-1.002	0.58718E-03	-1.001	0.29344E-03	-1.001
$x_{+2,2}$	0.23529E-02	0.11751E-02	-1.002	0.58692E-03	-1.002	0.29411E-03	-0.997
$x_{-2,2}$	0.23529E-02	0.11751E-02	-1.002	0.58719E-03	-1.001	0.29348E-03	-1.001
$x_{+3,3}$	0.23529E-02	0.11751E-02	-1.002	0.58718E-03	-1.001	0.29344E-03	-1.001
$x_{-3,3}$	0.23529E-02	0.11751E-02	-1.002	0.58720E-03	-1.001	0.29337E-03	-1.001

TABLE I: Relative error and converge rate, as a function of $1/\delta$, for a configuration of 6 spheres with radii equal to 1.5, and centers $x_{\pm\alpha} = \pm e_\alpha$ for $\alpha = 1, 2, 3$. Results where obtained with an angular momentum $L_{\max} = 8$, and an integration grid with $N_{\text{grid}} = 110$ nodes. Each atomic position coordinate $x_{\pm\alpha,\beta}$, where $\beta = 1, 2, 3$, was perturbed as $(1 + \delta)x_{\pm\alpha,\beta}$, thus generating only six nonzero variations, starting from an initial value $\delta_0 = 10^{-3}$.

for the project Aa-Par-T (Project-ID 57317909).

Appendix A: ddPCM discretization

The derivation of the ddPCM discretization employs the fact that the spherical harmonics Y_ℓ^m are eigenfunctions of the double layer operator $\mathcal{D}$ on the unit sphere $\mathbb{S}$, i.e., $\mathcal{D}\, Y_\ell^m = -2\pi/(2\ell+1)\, Y_\ell^m$, along with the following jump relation for the double layer potential

$$\lim_{\delta \to +0} \big(\tilde{\mathcal{D}}\, Y_\ell^m\big)(y \pm \delta\nu) = \pm 2\pi\, Y_\ell^m(y) + (\mathcal{D}\, Y_\ell^m)(y) \tag{A1}$$

where ν denotes the outward normal at $y \in \mathbb{S}$. We shall also rely on the invariance by translation and scaling of the double layer operator, namely $(\mathcal{D}_j\, \Phi_{\varepsilon,j})(s) = (\mathcal{D}\, \hat{\Phi}_{\varepsilon,j})(y)$, where $y = (s - x_j)/r_j$ and $\hat{\Phi}_{\varepsilon,j}$ is defined on $\mathbb{S}$ through the push-forward-like transformation $\hat{\Phi}_{\varepsilon,j}(y) = \Phi_{\varepsilon,j}(s)$. An analogous result holds for the double layer potential, namely $(\tilde{\mathcal{D}}_k\, \Phi_{\varepsilon,k})(x) = (\tilde{\mathcal{D}}\, \hat{\Phi}_{\varepsilon,j})(u)$ where $x \in \mathbb{R}^3 \setminus \overline{\Omega}_k$, $u = (x - x_j)/r_j$, and $\tilde{\mathcal{D}}$ is the double layer potential on $\mathbb{S}$.

As customary, in order to obtain a numerical discretization of (10), we multiply by a test function φ and integrate over Γ_j

$$\int_{\Gamma_j} \big(2\pi\, g(\varepsilon_s)\, \Phi_{\varepsilon,j} - U_j\, \mathcal{D}_j\, \Phi_{\varepsilon,j}\big)\varphi + \sum_{k\neq j} \int_{\Gamma_j} U_j\, \tilde{\mathcal{D}}_k\, \Phi_{\varepsilon,k}\, \varphi = \\ = \int_{\Gamma_j} \big(2\pi\, \Phi_j - U_j\, \mathcal{D}_j\, \Phi_j\big)\varphi + \sum_{k\neq j} \int_{\Gamma_j} U_j\, \tilde{\mathcal{D}}_k\, \Phi_k\, \varphi$$

Here we isolated the diagonal terms for convenience. Through the change of variable $y = (s - x_j)/r_j$, we obtain integrals over $\mathbb{S}$ which involve the hatted quantities, namely

$$\int_{\mathbb{S}} \big(2\pi\, g(\varepsilon_s)\, \hat{\Phi}_{\varepsilon,j} - \hat{U}_j\, \mathcal{D}\, \hat{\Phi}_{\varepsilon,j}\big)\hat{\varphi} + \sum_{k\neq j} \int_{\mathbb{S}} \hat{U}_j\, \tilde{\mathcal{D}}\, \hat{\Phi}_{\varepsilon,k}\, \hat{\varphi} = \\ = \int_{\mathbb{S}} \big(2\pi\, \hat{\Phi}_j - \hat{U}_j\, \mathcal{D}\, \hat{\Phi}_j\big)\hat{\varphi} + \sum_{k\neq j} \int_{\mathbb{S}} \hat{U}_j\, \tilde{\mathcal{D}}\, \hat{\Phi}_k\, \hat{\varphi} \quad \text{(A2)}$$

where we divided through by the surface Jacobian r_j^2.

In order to discretize the left-hand-side of the previous equation, we expand $\hat{\Phi}_{\varepsilon,j}$ as a series of spherical harmonics with coefficients $-[G_j]_{\ell'}^{m'}$, and select as a test function $\hat{\varphi}$ the spherical harmonic Y_ℓ^m. For the diagonal term, the orthogonality of the spherical harmonics, together with the fact that they are eigenfunctions of the double layer potential, yield

$$\int_{\mathbb{S}} \big(2\pi\, g(\varepsilon_s)\, \hat{\Phi}_{\varepsilon,j} - \hat{U}_j\, \mathcal{D}\, \hat{\Phi}_{\varepsilon,j}\big) Y_\ell^m = \\ = -2\pi\, g(\varepsilon_s) \sum_{\ell',m'} [G_j]_{\ell'}^{m'}\, \delta_{\ell\ell'}\delta_{mm'} - \sum_{\ell',m'} [G_j]_{\ell'}^{m'}\, \frac{2\pi}{2\ell'+1} \int_{\mathbb{S}} \hat{U}_j\, Y_{\ell'}^{m'}\, Y_\ell^m$$

The last step to obtain the diagonal block A_{jj}^ε is to approximate the integral through a suitable quadrature formula with weights $\{w_n\}$ and nodes $\{s_n\}$. Once the numerical quadrature is carried out and the spherical harmonics expansion is truncated at $\ell' = L_{\max}$, we derive

$$[A_{jj}^\varepsilon]_{\ell\ell'}^{mm'} = 2\pi\, g(\varepsilon_s)\, \delta_{\ell\ell'}\delta_{mm'} + \frac{2\pi}{2\ell'+1} \sum_{n=1}^{N_{\text{grid}}} w_n\, \hat{U}_j(s_n)\, Y_{\ell'}^{m'}(s_n)\, Y_\ell^m(s_n) \quad \text{(A3)}$$

which is the final expression of the diagonal block.

In order to discuss the off-diagonal blocks, let us write

$$\int_{\mathbb{S}} \hat{U}_j(y)\, (\tilde{\mathcal{D}}\, \hat{\Phi}_{\varepsilon,k})(u(y))\, Y_\ell^m(y)\, dy = -\sum_{\ell',m'} [G_k]_{\ell'}^{m'} \int_{\mathbb{S}} \hat{U}_j(y)\, (\tilde{\mathcal{D}}\, Y_{\ell'}^{m'})(u(y))\, Y_\ell^m(y)\, dy$$

where we have explicitly indicated the variable of integration. We recall that, when $x \in \Gamma_j$, i.e., $x = s = x_j + r_j y$ for some $y \in \mathbb{S}$, then $u = u(y) = (x_j + r_j y - x_k)/r_k$. The function $\tilde{\mathcal{D}}\, Y_{\ell'}^{m'}$

is harmonic on $\mathbb{R}^3 \setminus \overline{B(0,1)}$, so that is has to coincide with the unique harmonic extension of its boundary value. The jump relation (A1), along with the eigenfunction property, provide the boundary value

$$\lim_{\delta \to +0} \big(\tilde{\mathcal{D}}\, Y_{\ell'}^{m'}\big)(y + \delta\nu) = 2\pi\, Y_{\ell'}^{m'}(y) + \big(\mathcal{D}\, Y_{\ell'}^{m'}\big)(y) = \frac{4\pi\ell'}{2\ell +' 1}\, Y_{\ell'}^{m'}(y)$$

and, by elementary notions on harmonic functions, we conclude

$$(\tilde{\mathcal{D}}\, Y_{\ell'}^{m'})(u) = \frac{4\pi\ell'}{2\ell'+1}\, \frac{1}{|u|^{\ell'+1}}\, Y_{\ell'}^{m'}(u/|u|)$$

After truncation the series expansion and performing numerical integration we obtain

$$[A_{jk}]_{\ell\ell'}^{mm'} = -\frac{4\pi\ell'}{2\ell'+1} \sum_{n=1}^{N_{\text{grid}}} w_n\, \hat{U}_j(s_n)\, \frac{1}{|u(s_n)|^{\ell'+1}}\, Y_{\ell'}^{m'}(u(s_n)/|u(s_n)|)\, Y_\ell^m(s_n) \qquad \text{(A4)}$$

which is the final result.

The discretization of the right-hand-side of the ddPCM equation (A2) is entirely analogous. As a first step, we expand Φ_j as a series of spherical harmonics, recall (14), so that the orthogonality condition yields the following expression

$$[F_j]_\ell^m = -\int_{\mathbb{S}} \hat{\Phi}_j\, Y_\ell^m = -\int_{\mathbb{S}} \hat{U}_j\, \hat{\Phi}\, Y_\ell^m(y)$$

which is evaluated through numerical quadrature. The derivation of the discrete operator A_∞ arising from the right-hand-side is entirely analogous to the one described for the left-hand-side. In fact, it coincides with A_ε with the natural extension $g(\infty) = 1$.

Appendix B: ddCOSMO discretization

As customary in a variational setting, we discretize equation (13) by multiplying it with a test function φ and integrating it over Γ_j. As a preliminary step, we manipulate the following integral as

$$\int_{\Gamma_j} \tilde{\mathcal{S}}_{jk}\, \sigma_k\, \varphi = \int_{\Gamma_j} n_j\, \tilde{\mathcal{S}}_k\, \sigma_k\, \varphi = \int_{\Gamma_j} (1 - U_j)\, n_j\, \tilde{\mathcal{S}}_k\, \sigma_k\, \varphi = \int_{\Gamma_j} V_j\, \tilde{\mathcal{S}}_k\, \sigma_k\, \varphi$$

We inserted the characteristic function U_j for convenience, and defined the rescaled characteristic function $V_j : \Gamma_j \to \mathbb{R}$ as $V(s) = (1 - U_j(s))\, n_j(s)$. The reason for introducing the characteristic function is that we can modify it, e.g., replace it by a smooth counterpart, to

improve robustness of the algorithm. We obtain the following variational formulation of the integral equation (13):

$$\int_{\Gamma_j} \mathcal{S}_j\, \sigma_j\, \varphi - \sum_{k\in N_j} \int_{\Gamma_j} V_j\, \tilde{\mathcal{S}}_k\, \sigma_k\, \varphi = -\int_{\Gamma_j} \Phi_{\varepsilon,j}\, \varphi$$

As a next step, we map quantities to the unit sphere $\mathbb{S}$. In order to do so, we employ the translation-invariant properties $(\mathcal{S}_j\, \sigma_j)(s) = r_j(\mathcal{S}\, \hat{\sigma}_j)(z)$ and $(\tilde{\mathcal{S}}_j\, \sigma_j)(x) = r_j(\tilde{\mathcal{S}}\, \hat{\sigma}_j)(u)$, where $\mathcal{S}$ and $\tilde{\mathcal{S}}$ are, respectively, the single layer operator and potential on the unit sphere, $y = (s - x_j)/r_j$, $u = (x - x_j)/r_j$, and $\hat{\sigma}_j(y) = \sigma_j(s)$. We obtain

$$r_j \int_{\mathbb{S}} \mathcal{S}\, \hat{\sigma}_j\, \hat{\varphi} - \sum_{k\in N_j} r_k \int_{\mathbb{S}} \hat{V}_j\, \tilde{\mathcal{S}}\, \hat{\sigma}_k\, \hat{\varphi} = -\int_{\mathbb{S}} \hat{\Phi}_{\varepsilon,j}\, \hat{\varphi} \tag{B1}$$

where we divided both sides by the surface Jacobian r_j^2. The workhorse for the objects $\mathcal{S}\, \hat{\sigma}_j$ and $\tilde{\mathcal{S}}\, \hat{\sigma}_k$ is the Addition Theorem for spherical harmonics.

If we expand $\hat{\sigma}_j$ through spherical harmonics Y_ℓ^m as

$$\hat{\sigma}_j(y) = \frac{1}{r_j} \sum_{\ell'=0}^{\infty} \sum_{m'=-\ell'}^{\ell'} [X_j]_{\ell'}^{m'}\, Y_{\ell'}^{m'}(y)$$

for some (unknown!) coefficients $X_j = [X_j]_{\ell'}^{m'}$, the orthogonality property of the spherical harmonics implies

$$(\mathcal{S}\, \hat{\sigma}_j)(y) = \frac{1}{r_j} \sum_{\ell'=0}^{\infty} \sum_{m'=-\ell'}^{\ell'} \frac{4\pi}{(2\ell'+1)}\, [X_j]_{\ell'}^{m'}\, Y_{\ell'}^{m'}(y)$$

Thus, if we select as test functions the spherical harmonics, we obtain

$$r_j \int_{\mathbb{S}} \mathcal{S}\, \hat{\sigma}_j\, Y_\ell^m = \sum_{\ell'=0}^{\infty} \sum_{m'=-\ell'}^{\ell'} \frac{4\pi}{(2\ell'+1)}\, \delta_{\ell\ell'}\delta_{mm'}\, [X_j]_{\ell'}^{m'}$$

due to their orthogonality. A numerical discretization is obtained by truncating the expansion for some $\ell' = L_{\max}$. This provides the discrete action $L_{jj}\, X_j$, where the discrete operator

$$[L_{jj}]_{\ell\ell'}^{mm'} = \frac{4\pi}{(2\ell'+1)}\, \delta_{\ell\ell'}\delta_{mm'} \tag{B2}$$

is diagonal and does not depend on the nuclear positions.

Analogously to the previous case, for the off-diagonal term the Addition Theorem implies

$$(\tilde{\mathcal{S}}\, \hat{\sigma}_k)(u) = \frac{1}{r_k} \sum_{\ell'=0}^{\infty} \sum_{m'=-\ell'}^{\ell'} \frac{4\pi}{(2\ell'+1)}\, [X_k]_{\ell'}^{m'}\, |u|^{\ell'}\, Y_{\ell'}^{m'}(u/|u|)$$

We relied on the fact that, when $x \in \Gamma_j$, i.e., $x = s = x_j + r_j y$, then $u = u(y) = (x_j + r_j y - x_k)/r_k$ and $|u| < 1$ because of the assumption $k \in N_j$. We obtain

$$r_k \int_{\mathbb{S}} \hat{V}_j(y)\, (\tilde{\mathcal{S}}\, \hat{\sigma}_k)(u(y))\, Y_\ell^m(y)\, dy = \\ = \sum_{\ell'=0}^{\infty} \sum_{m'=-\ell'}^{\ell'} \frac{4\pi}{(2\ell'+1)}\, [X_k]_{\ell'}^{m'} \int_{\mathbb{S}} \hat{V}_j(y)\, |u(y)|^{\ell'}\, Y_{\ell'}^{m'}(u(y)/|u(y)|)\, Y_\ell^m(y)\, dy$$

which, as opposed to the diagonal block L_{jj}, requires numerical integration. This yields the off-diagonal block

$$[L_{jk}]_{\ell\ell'}^{mm'} = \frac{4\pi}{(2\ell'+1)} \sum_{n=1}^{N_{\text{grid}}} w_n\, \hat{V}_j(s_n)\, |u(s_n)|^{\ell'}\, Y_{\ell'}^{m'}(u(s_n)/|u(s_n)|)\, Y_\ell^m(s_n) \tag{B3}$$

where we employed a Lebedev grid with N_{grid} nodes $\{s_n\}$ and weights $\{w_n\}$ to perform numerical quadrature. We remark that the off-diagonal block L_{jk} depends, *a priori*, upon the nuclear positions x_j, x_k, and x_i, for $i \in N_j$.

The discretization of the right-hand-side of (B1), i.e., the load vector is simply

$$[F_j]_\ell^m = -\int_{\mathbb{S}} \hat{\Phi}_{\varepsilon,j}\, Y_\ell^m$$

which, in general, does require a numerical quadrature. However, when $\Phi_{\varepsilon,j}$ is provided by the ddPCM step as a truncated series of spherical harmonics with coefficients $-[G_j]_{\ell'}^{m'}$, the orthogonality condition implies

$$[F_j]_\ell^m = -\int_{\mathbb{S}} \hat{\Phi}_{\varepsilon,j}\, Y_\ell^m = \sum_{\ell'=0}^{L_{\max}} \sum_{m'=-\ell'}^{\ell'} [G_j]_{\ell'}^{m'} \int_{\mathbb{S}} Y_{\ell'}^{m'}\, Y_\ell^m = [G_j]_\ell^m$$

so that we conclude that the load vector F coincides with G provided by the ddPCM step.

Appendix C: Derivatives of ddPCM discretization

The function U_j is, in practice, a smoothed version of the characteristic function. We use the following construction

$$U_j(x_j + r_j y) = \begin{cases} 1 - f_j(y) & f_j(y) \leq 1 \\ 0 & \text{otherwise} \end{cases}, \qquad f_j(y) = \sum_{i \in N_j} \chi\left(\frac{|x_j + r_j y - x_i|}{r_i}\right)$$

where χ is a regularized characteristic function of $[0,1]$. As previously, the variable y varies on the unit sphere $\mathbb{S}$. This definition implies that U_j depends upon the nuclear positions x_j and x_i such that $i \in N_j$.

Let $\{s_n\}$ be N_{grid} integration points, with associated weights $\{w_n\}$, and define the following quantities

$$t_n^{jk} = \frac{|x_j + r_j s_n - x_k|}{r_k} \qquad , \qquad s_n^{jk} = \frac{x_j + r_j s_n - x_k}{|x_j + r_j s_n - x_k|}$$

It is evident that t_n^{jk} depends only upon the nuclear positions x_j and x_k, as does s_n^{jk}. Standard differentiation yields

$$\nabla_j \, t_n^{jk} = \frac{s_n^{jk}}{r_k} \qquad , \qquad D_j \, s_n^{jk} = \frac{I - s_n^{jk} \otimes s_n^{jk}}{|x_j + r_j s_n - x_k|^3} \tag{C1}$$

where I is the identity matrix and $\otimes$ indicates the outer product. We remark that the Jacobian matrix $D_j \, s_n^{jk}$ is symmetric, and the "twin" relations $\nabla_j \, t_n^{jk} = -\nabla_k \, t_n^{jk}$ and $D_j \, s_n^{jk} = -D_k \, s_n^{jk}$ hold.

If we employ this notation and recall (A3) and (A4), the blocks of the ddPCM operator can be written as

$$[A_{jj}^{\varepsilon}]_{\ell\ell'}^{mm'} = 2\pi \, g(\varepsilon_s) \, \delta_{\ell\ell'} \delta_{mm'} + \frac{2\pi}{2\ell'+1} \sum_{n=1}^{N_{\text{grid}}} w_n \, \hat{U}_j(s_n) \, Y_\ell^m(s_n) \, Y_{\ell'}^{m'}(s_n)$$

$$[A_{jk}^{\varepsilon}]_{\ell\ell'}^{mm'} = \qquad\qquad - \frac{4\pi\ell'}{2\ell'+1} \sum_{n=1}^{N_{\text{grid}}} w_n \, \hat{U}_j(s_n) \, Y_\ell^m(s_n) \left(t_n^{jk}\right)^{-(\ell'+1)} Y_{\ell'}^{m'}(s_n^{jk})$$

The dependency of the diagonal block A_{jj}^{ε} upon the nuclear positions occurs only through the characteristic function U_j. On the other hand, the off-diagonal block A_{jk}^{ε} interacts with the nuclear positions through the characteristic function U_j, as well as t_n^{jk} and s_n^{jk}. This implies that $\nabla_i \, A_{jj}$ is *a priori* nonzero only when $i = j$ or $i \in N_j$. Similarly, $\nabla_i \, A_{jk}$ is *a priori* nonzero only when $i = j$ or $i = k$ or $i \in N_j$.

The case of the diagonal blocks yields

$$[\nabla_i \, A_{jj}]_{\ell\ell'}^{mm'} = \frac{2\pi}{2\ell'+1} \sum_{n=1}^{N_{\text{grid}}} w_n \, \nabla_i \hat{U}_j(s_n) \, Y_\ell^m(s_n) \, Y_{\ell'}^{m'}(s_n)$$

so that only the derivatives of the characteristic function are required. The case of the off-diagonal blocks is significantly more involved, since it involves the gradient of the product of three functions, namely

$$[\nabla_i \, A_{jk}]_{\ell\ell'}^{mm'} = -\frac{4\pi\ell'}{2\ell'+1} \sum_{n=1}^{N_{\text{grid}}} w_n \, Y_\ell^m(s_n) \, \nabla_i \left[\hat{U}_j(s_n) \left(t_n^{jk}\right)^{-(\ell'+1)} Y_{\ell'}^{m'}(s_n^{jk}) \right] \tag{C2}$$

We proceed to analyze the gradient of the triple product when $i = j$, or $i = k$, or $i \in N_j$, $i \neq k$. Those three cases are mutually exclusive.

The case $i = j$ yields no simplification, and the gradient of the triple product has the three standard contributions, namely

$$\nabla_j \left[\cdots \right] = \nabla_j \hat{U}_j(s_n) \left(t_n^{jk} \right)^{-(\ell'+1)} Y_{\ell'}^{m'}(s_n^{jk}) + \hat{U}_j(s_n) \nabla_j \left(\left(t_n^{jk} \right)^{-(\ell'+1)} \right) Y_{\ell'}^{m'}(s_n^{jk}) + \hat{U}_j(s_n) \left(t_n^{jk} \right)^{-(\ell'+1)} \nabla_j Y_{\ell'}^{m'}(s_n^{jk}) \quad \text{(C3)}$$

In the implementation the differentiation is fully carried out through the chain rule, which employs the formulas (C1) for the derivatives of t_n^{jk} and s_n^{jk}.

The case $i = k$ needs to be split into the subcases $k \in N_j$ and $k \notin N_j$. The first subcase does not yield any simplification and reduces to (C3). On the other hand, when $k \notin N_j$, the first term on the right-hand-side of (C3) drops out, and we obtain

$$\nabla_k \left[\cdots \right] = \hat{U}_j(s_n) \nabla_k \left(\left(t_n^{jk} \right)^{-(\ell'+1)} \right) Y_{\ell'}^{m'}(s_n^{jk}) + \hat{U}_j(s_n) \left(t_n^{jk} \right)^{-(\ell'+1)} \nabla_k Y_{\ell'}^{m'}(s_n^{jk})$$

Finally, when $i \in N_j$ and $i \neq k$ the second and third term on the right-hand-side of (C3) vanish, and we obtain

$$\nabla_i \left[\cdots \right] = \nabla_j \hat{U}_j(s_n) \left(t_n^{jk} \right)^{-(\ell'+1)} Y_{\ell'}^{m'}(s_n^{jk})$$

This conclude our discussion on the derivative of the PCM matrix, we remark that $\nabla_i A_{jk}$ is *a priori* nonzero only when $i \in N_j \cup \{j, k\}$. This implies that its action can be computed within linear, as oppose to quadratic, complexity.

REFERENCES

[1]J. Tomasi, B. Mennucci, and R. Cammi, Chem. Rev. **105**, 2999 (2005).

[2]C. Cramer and D. Truhlar, Chem Rev **99**, 2161 (1999).

[3]M. Orozco and F. J. Luque, Chem. Rev. **100**, 4187 (2000).

[4]A. Klamt, WIREs Comput. Mol. Sci. **1**, 699 (2011).

[5]B. Mennucci, WIREs Comput. Mol. Sci. **2**, 386 (2012).

[6]B. Honig and A. Nicholls, Science **268**, 1144 (1995).

[7]B. Roux and T. Simonson, Biophys. Chem. **78**, 1 (1999).

[8]A. Klamt and G. Schuurmann, J. Chem. Soc., Perkin Trans. 2 , 799 (1993).

[9]D. Chipman, J. Chem. Phys. **110**, 8012 (1999).

[10]D. M. Chipman, J. Chem. Phys. **124**, 224111 (2006).

[11]E. Cancès, B. Mennucci, and J. Tomasi, J. Chem. Phys. **107**, 3032 (1997).

[12]E. Cancès and B. Mennucci, J.Math. Chem. **23**, 309 (1998).

[13]B. Mennucci, E. Cancès, and J. Tomasi, J. Phys. Chem. B **101**, 10506 (1997).

[14]V. Barone and M. Cossi, J. Phys. Chem. A **102**, 1995 (1998).

[15]M. Cossi, N. Rega, G. Scalmani, and V. Barone, J. Comput. Chem. **24**, 669 (2003).

[16]F. Lipparini, G. Scalmani, B. Mennucci, E. Cancès, M. Caricato, and M. J. Frisch, J. Chem. Phys. **133**, 014106 (2010).

[17]F. Lipparini, G. Scalmani, B. Mennucci, and M. J. Frisch, J. Chem. Theory Comput. **7**, 610 (2011).

[18]A. V. Marenich, C. J. Cramer, and D. G. Truhlar, J. Phys. Chem. B **113**, 6378 (2009).

[19]B. Mennucci, J. Phys. Chem. Lett. **1**, 1666 (2010), and references therein.

[20]S. Miertus, E. Scrocco, and J. Tomasi, Chem. Phys. **55**, 117 (1981).

[21]J. Tomasi and M. Persico, Chem. Rev. **94**, 2027 (1994).

[22]G. Scalmani and M. J. Frisch, J. Chem. Phys. **132**, 114110 (2010).

[23]D. York and M. Karplus, J. Phys. Chem. A **103**, 11060 (1999).

[24]A. W. Lange and J. M. Herbert, J. Chem. Phys. **133**, 244111 (2010).

[25]R. Cammi and J. Tomasi, J. Comput. Chem. **16**, 1449 (1995).

[26]G. Scalmani, V. Barone, K. Kudin, C. Pomelli, G. Scuseria, and M. Frisch, Theor. Chem. Acc. **111**, 90 (2004).

[27]L. Greengard and V. Rokhlin, J. Comput. Phys. **73**, 325 (1987).

[28]F. Lipparini, L. Lagardère, G. Scalmani, B. Stamm, E. Cancès, Y. Maday, J.-P. Piquemal, M. J. Frisch, and B. Mennucci, J. Phys. Chem. Lett. **5**, 953 (2014).

[29]E. Cancès, Y. Maday, and B. Stamm, J. Chem. Phys. **139**, 054111 (2013).

[30]F. Lipparini, B. Stamm, E. Cancès, Y. Maday, and B. Mennucci, J. Chem. Theory Comput. **9**, 3637 (2013).

[31]F. Lipparini, G. Scalmani, L. Lagardère, B. Stamm, E. Cancès, Y. Maday, J.-P. Piquemal, M. J. Frisch, and B. Mennucci, J. Chem. Phys. **141**, 184108 (2014).

[32]B. Stamm, E. Cancès, and Y. Maday, J. Chem. Phys. **144**, 054101 (2016).

[33]C. Quan, B. Stamm, and Y. Maday, "Polarizable Continuum Model based on the Solvent Excluded Surface," (2017), working paper or preprint.

[34]M. Bugeanu, R. Di Remigio, K. Mozgawa, S. S. Reine, H. Harbrecht, and L. Frediani, Phys. Chem. Chem. Phys. **17**, 31566 (2015).

[35]H. Harbrecht and M. Randrianarivony, Computing **92**, 335 (2011).

[36]V. Weijo, M. Randrianarivony, H. Harbrecht, and L. Frediani, Journal of Computational Chemistry **31**, 1469 (2010).

[37]W. Hackbusch, *Integral Equations–Theory and Numerical Treatment* (Birkhäuser, Basel, 1995).

[38]E. Cancès, in *Continuum Solvation Models in Chemical Physics*, edited by B. Mennucci and R. Cammi (Wiley, New York, 2007) Chap. 1.2, pp. 29–48.