JCTC Journal of Chemical Theory and Computation

Article

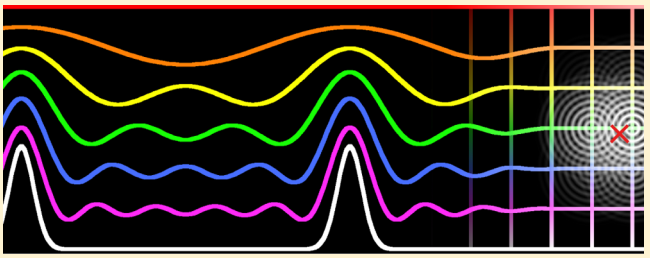
pubs.acs.org/JCTC

Multipolar Ewald Methods, 1: Theory, Accuracy, and Performance

Timothy J. Giese, Maria T. Panteva, Haoyuan Chen, and Darrin M. York*

Center for Integrative Proteomics Research, BioMaPS Institute for Quantitative Biology and Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, New Jersey 08854-8087, United States

ABSTRACT: The Ewald, Particle Mesh Ewald (PME), and Fast Fourier–Poisson (FFP) methods are developed for systems composed of spherical multipole moment expansions. A unified set of equations is derived that takes advantage of a spherical tensor gradient operator formalism in both real space and reciprocal space to allow extension to arbitrary multipole order. The implementation of these methods into a novel linear-scaling modified "divide-and-conquer" (mDC) quantum mechanical force field is discussed. The evaluation times and relative force errors are compared between the three methods, as a function of multipole expansion order. Timings and errors are also compared within the context of the quantum mechanical force field, which encounters primary errors related to the quality of reproducing electrostatic forces for a given density matrix and secondary errors resulting from the propagation of the approximate electrostatics into the self-consistent field procedure, which yields a converged, variational, but nonetheless approximate density matrix. Condensed-phase simulations of an mDC water model are performed with the multipolar PME method and compared to an electrostatic cutoff method, which is shown to artificially increase the density of water and heat of vaporization relative to full electrostatic treatment.



1. INTRODUCTION

It is well-known that the molecular modeling of many biological processes requires a rigorous treatment of the long-ranged electrostatic interactions.[1−5] Furthermore, recent years have seen the development of many promising next-generation force fields and fast *ab initio* methods that endeavor to generalize the electrostatic interactions to higher-order atomic multipoles.[6] Specific examples include the following: polarizable force fields, such as AMOEBA;[7−12] density-based force fields, such as GEM[13−16] and S/G-1;[17] force fields that model the perturbative or many-body expansion of the energy;[18−21] hybrid quantum mechanical/molecular mechanical (QM/MM) methods;[22] and molecular orbital-based quantum mechanical force fields (QMFFs), such as the embedded fragment model,[23] X-Pol,[24−29] and the closely related modified "divide-and-conquer" (mDC) methods.[30−32] The use of higher-order atomic multipoles in these examples, and many other models, offer the promise of improved accuracy, but at a larger computational cost.[33−35] In order to elucidate the strengths and weaknesses of proposed models, it is necessary to make comparison to experiment, which often requires their application within molecular simulations under periodic boundary conditions. At the same time, these applications provide empirical measures of the preliminary model's true computational cost. Therefore, fast electrostatic algorithms must be developed that are equipped to handle generalized charge densities, so that new models can be applied and tested.

The Ewald,[36] Particle Mesh Ewald[37−40] (PME), and Fast Fourier–Poisson[41] (FFP) methods are three algorithms used to evaluate the long-ranged electrostatic interactions of periodic systems. These methods were originally designed to accommodate systems consisting of point charges; therefore, modifications to their original formulations are required to apply them to models employing multipolar charge densities. In the present work, we derive a unified set of equations for the Ewald, PME, and FFP methods using a spherical tensor gradient operator formalism that extend these methods to arbitrary multipole order. Expressions for the energies, forces, and the generalized multipolar potentials required to incorporate the electrostatics into the QMFF self-consistent field (SCF) procedure are provided. Previous works have extended these methods within the framework of Cartesian point multipoles and Cartesian Gaussians,[42−47] as opposed to the solid harmonic multipoles considered here. Other related works were not formulated to arbitrary multipole order[48,49] or led to a formulism that did not prove to be computationally efficient.[50] The efficiency of our formulation is demonstrated through QMFF molecular dynamics applications described below and extended in Part 2 of this series.[51] During preparation of this manuscript, Simmonett[52] has independently reported an extension of PME for use with solid harmonic multipoles that shares many characteristics with the PME method presented here; however, our analysis extends significantly beyond ref 52 by comparing the accuracy and performance of Ewald, PME, and FFP methods as stand-alone electrostatic methods and upon implementation within a linear-scaling QMFF.

The paper is organized as follows. Section 2 describes the generalized charge density; derives equations for the general-

Received: September 2, 2014
Published: December 27, 2014

DOI: 10.1021/ct5007983
J. Chem. Theory Comput. 2015, 11, 436−450

ized Ewald, PME, and FFP methods; and discusses net charge corrections, error analysis, and the integration of the methods into the mDC QMFF. Section 3 compares the accuracy and computational cost of the methods, as a function of multipolar order and within the context of condensed-phase molecular simulations of water.

2. METHODS

2.1. The Charge Density. Our goal is to compute the electrostatic interaction of a neutral charge density $\rho(\mathbf{r})$ with itself and its periodic images:[53]

$$E = \frac{1}{2}\int \rho(\mathbf{r}) \sum_{\mathbf{n}} \int \frac{\rho(\mathbf{r}' + \mathbf{n})}{|\mathbf{r} - \mathbf{r}'|}\, \mathrm{d}^3 r\, \mathrm{d}^3 r' \tag{1}$$

where the unit cell defining the periodicity is described by the lattice vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ and reciprocal-space vectors $\mathbf{a}_1^*$, $\mathbf{a}_2^*$, $\mathbf{a}_3^*$; and the periodic images are replicated by integer lattice translations $\mathbf{n} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$. The energy (given by eq 1) is finite only when the charge density is neutral; however, a widely used approximation for the treatment of charged systems will be discussed in Section 2.6. Furthermore, we adopt the standard convention of removing the infinite self-energy of point charges (and point multipoles) whenever those energies may appear. Upon replicating all integer translations of the aperiodic density, $\rho(\mathbf{r})$, the edges of a lattice with the same shape, orientation, and dimensions can be redrawn about any origin; and the periodic density filling each cell will contain exactly one instance of $\rho(\mathbf{r})$ that appears to have been "wrapped" into the cell's interior. The apparent wrapping of $\rho(\mathbf{r})$ to the cell boundary is an illusion formed by the effluence of density emanating from the other translations, upon making $\rho(\mathbf{r})$ periodic. The aperiodic density is not required to be confined within a primary unit cell to achieve this effect. Nor is it a requirement for the evaluation of eq 1, because the electrostatic potential of $\sum_{\mathbf{n}}\rho(\mathbf{r} + \mathbf{n})$ is periodic.

The methods presented in this work are solutions to eq 1 for a density composed of atom-centered point multipole expansions. There are various ways of defining what is meant by a multipole expansion. The two most common ways being Cartesian multipoles[44] or those based on spherical harmonics.[52,54,55] Our definition of a multipole moment is the inner product of a density with a *real-valued* regular solid harmonic:[56]

$$q_{l\mu} = \int \rho(\mathbf{r}) C_{l\mu}(\mathbf{r})\, \mathrm{d}^3 r \tag{2}$$

An auxiliary basis of charge density $\chi_{l\mu}(\mathbf{r})$, satisfying

$$\int \chi_{l\mu}(\mathbf{r} - \mathbf{R}_a) C_{j\kappa}(\mathbf{r} - \mathbf{R}_a)\, \mathrm{d}^3 r = \delta_{lj}\delta_{\mu\kappa} \tag{3}$$

is thus well-suited to describe an atomic multipole expansion. The definition of $C_{l\mu}(\mathbf{r})$ is closely related to the scaled regular solid harmonics frequently encountered in fast multipole methods.[54,57−59] In brief, the real-valued *scaled* regular solid harmonics $R_{lm}^{c/s}(\mathbf{r})$ are the real $R_{lm}^{c}(\mathbf{r}) = \mathrm{Re}R_{lm}(\mathbf{r})$ and imaginary $R_{lm}^{s}(\mathbf{r}) = \mathrm{Im}R_{lm}(\mathbf{r})$ components of the complex-valued scaled regular solid harmonics:[58]

$$R_{lm}(\mathbf{r}) = r^l \frac{(-1)^m}{(l+m)!} P_{lm}(\cos\theta) \mathrm{e}^{im\phi} \tag{4}$$

where

$$P_{lm}(x) = \frac{(1-x^2)^{m/2}}{2^l l!} \frac{\mathrm{d}^{l+m}}{\mathrm{d}x^{l+m}}(x^2-1)^l \tag{5}$$

is an associated Legendre polynomial. Because the complex-valued harmonics are symmetry related $R_{lm}^*(\mathbf{r}) = (-1)^m R_{l,-m}(\mathbf{r})$, the real-valued harmonics are fully described by the set of non-negative m values. The notation is simplified, when appropriate, through the introduction of a Greek subscript whose sign merely acts to symbolize the cosine/sine designation[57]

$$R_{l\mu} = \begin{cases} R_{lm}^{c}(\mathbf{r}) & \mu \geq 0 \\ R_{l|m|}^{s}(\mathbf{r}) & \mu < 0 \end{cases} \tag{6}$$

The real-valued regular solid harmonics

$$C_{l\mu}(\mathbf{r}) = A_{l\mu} R_{l\mu}(\mathbf{r}) \tag{7}$$

differ only by factors

$$A_{l\mu} = (-1)^\mu \sqrt{(2-\delta_{\mu,0})(l+\mu)!(l-\mu)!} \tag{8}$$

that are chosen to reproduce Racah's normalization

$$\int C_{l\mu}(\mathbf{r}) C_{j\kappa}(\mathbf{r})\, \mathrm{d}\Omega = \delta_{lj}\delta_{\mu\kappa} r^{2l} \frac{4\pi}{2l+1} \tag{9}$$

In the interest of writing the expression for a point multipole, we begin by deducing the form of a Gaussian multipole function from eqs 3 and 9:

$$\begin{aligned} \chi_{l\mu}(\mathbf{r} - \mathbf{R}_a) &= \frac{C_{l\mu}(\mathbf{r} - \mathbf{R}_a)}{(2l-1)!!}(2\zeta)^l \left(\frac{\zeta}{\pi}\right)^{3/2} \mathrm{e}^{-\zeta|\mathbf{r}-\mathbf{R}_a|^2} \\ &= \frac{C_{l\mu}(\nabla_a)}{(2l-1)!!}\chi(\mathbf{r} - \mathbf{R}_a) \end{aligned} \tag{10}$$

where

$$\chi(\mathbf{r} - \mathbf{R}_a) \equiv \chi_{00}(\mathbf{r} - \mathbf{R}_a) = \left(\frac{\zeta}{\pi}\right)^{3/2} \mathrm{e}^{-\zeta|\mathbf{r}-\mathbf{R}_a|^2} \tag{11}$$

is a Gaussian monopole, $\nabla_a = \{\mathrm{d}/\mathrm{d}X_a, \mathrm{d}/\mathrm{d}Y_a, \mathrm{d}/\mathrm{d}Z_a\}$ are gradients with respect to the Gaussian center, and $C_{l\mu}(\nabla)$ is a spherical tensor gradient operator[60] (STGO). A STGO is constructed by replacing the Cartesian coordinate arguments of the solid harmonic with their corresponding Cartesian derivative operators. The STGO obeys several mathematical properties.[61,62] The present work specifically makes use of the STGO chain-rule identity,

$$\begin{aligned} C_{l\mu}(\nabla) f(u) &= C_{l\mu}\left(\frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}}{\mathrm{d}u}, \frac{\mathrm{d}u}{\mathrm{d}y}\frac{\mathrm{d}}{\mathrm{d}u}, \frac{\mathrm{d}u}{\mathrm{d}z}\frac{\mathrm{d}}{\mathrm{d}u}\right) f(u) \\ &= C_{l\mu}(\nabla u)\left(\frac{\mathrm{d}}{\mathrm{d}u}\right)^l f(u) \end{aligned} \tag{12}$$

which follows from the homogeneity of solid harmonics $C_{l\mu}(a\mathbf{r}) = a^l C_{l\mu}(\mathbf{r})$. The application of eq 12 to the last line of eq 10 establishes the relationship between $\chi_{l\mu}(\mathbf{r} - \mathbf{R}_a)$ and a Gaussian monopole. A Dirac delta function is a Gaussian monopole in the limit of infinite exponent:

$$\delta(\mathbf{r} - \mathbf{R}_a) \equiv \lim_{\zeta\to\infty} \chi(\mathbf{r} - \mathbf{R}_a) \tag{13}$$

and a point multipole is this limit applied to eq 10:

DOI: 10.1021/ct5007983
J. Chem. Theory Comput. 2015, 11, 436−450

$$\delta_{l\mu}(\mathbf{r}-\mathbf{R}_a)=\frac{C_{l\mu}(\nabla_a)}{(2l-1)!!}\delta(\mathbf{r}-\mathbf{R}_a) \tag{14}$$

Thus, a sum of atom-centered point multipole expansions is

$$\rho(\mathbf{r})=\sum_{a,l\mu\in a} q_{a,l\mu}\frac{C_{l\mu}(\nabla_a)}{(2l-1)!!}\delta(\mathbf{r}-\mathbf{R}_a) \tag{15}$$

2.2. Plane Waves. Methods for solving eq 1 avoid the infinite sum of explicit lattice translations by approximating a periodic density with a finite number of plane wave basis functions $\langle\mathbf{r}|\mathbf{k}\rangle \equiv e^{i\mathbf{k}^T\cdot\mathbf{r}}$. The description of these methods is greatly simplified using Dirac notation, which we briefly summarize here. In this notation, the spatial representation of a function is equivalently written $\langle\mathbf{r}|f\rangle = f(\mathbf{r})$; a complex conjugate is $\langle f|\mathbf{r}\rangle = f^*(\mathbf{r})$; the inner product of two functions $\langle f|g\rangle = \int f^*(\mathbf{r})g(\mathbf{r})\,d^3r$ integrates all space or, if both functions are periodic, the volume of the unit cell; and we reserve $|\mathbf{k}\rangle$ to index the plane wave basis functions. This index is meant to be the function's "angular wave number", which is defined as $\mathbf{k} = 2\pi(k_1\mathbf{a}_1^* + k_2\mathbf{a}_2^* + k_3\mathbf{a}_3^*)$. From these definitions, one can show that the basis is orthogonal $\langle\mathbf{k}|\mathbf{k}'\rangle = \delta_{\mathbf{k},\mathbf{k}'}V$, where $V = \mathbf{a}_1\cdot\mathbf{a}_2\times\mathbf{a}_3$ is the unit-cell volume; they obey a trivial addition theorem $\langle\mathbf{r}+\mathbf{r}'|\mathbf{k}\rangle = \langle\mathbf{r}|\mathbf{k}\rangle\langle\mathbf{r}'|\mathbf{k}\rangle$; they are eigenfunctions of the gradient operator $\nabla\langle\mathbf{r}|\mathbf{k}\rangle = i\mathbf{k}\langle\mathbf{r}|\mathbf{k}\rangle$; and the electrostatic potential of a plane wave basis function

$$\int\frac{\langle\mathbf{r}'|\mathbf{k}\rangle}{|\mathbf{r}-\mathbf{r}'|}\,d^3r' = \langle\mathbf{r}|\mathbf{k}\rangle\frac{4\pi}{k^2} \tag{16}$$

becomes particularly amenable to computation. Finally, it is easy to show that the plane waves are eigenfunctions of the STGO,

$$C_{l\mu}(\nabla)\langle\mathbf{r}|\mathbf{k}\rangle = i^l C_{l\mu}(\mathbf{k})\langle\mathbf{r}|\mathbf{k}\rangle \tag{17}$$

via direct application of eq 12 with $u = i\mathbf{k}^T\cdot\mathbf{r}$.

Plane waves can be used as a basis to represent real, periodic functions:

$$\begin{aligned}\sum_{\mathbf{n}}\rho(\mathbf{r}+\mathbf{n}) &= \sum_{\mathbf{k}} a_{\mathbf{k}}\cos(\mathbf{k}^T\cdot\mathbf{r}) + b_{\mathbf{k}}\sin(\mathbf{k}^T\cdot\mathbf{r})\\ &= \mathrm{Re}\sum_{\mathbf{k}} e^{i\mathbf{k}^T\cdot\mathbf{r}}(a_{\mathbf{k}} - ib_{\mathbf{k}})\\ &= \mathrm{Re}\sum_{\mathbf{k}}\langle\mathbf{r}|\mathbf{k}\rangle c_{\mathbf{k}}^*\end{aligned} \tag{18}$$

where $c_{\mathbf{k}} = a_{\mathbf{k}} + ib_{\mathbf{k}}$. Complex numbers have been introduced to compactly express two independent linear least-square fit problems: the representation of the even (cosine) and odd (sine) character of the density. One can show that the two least-square fit solutions for $a_{\mathbf{k}}$ and $b_{\mathbf{k}}$ are equivalent to having chosen

$$\begin{aligned}c_{\mathbf{k}}^* &= \sum_{\mathbf{k}'}\langle\mathbf{k}|\mathbf{k}'\rangle^{-1}\sum_{\mathbf{n}}\int_V\langle\mathbf{k}'|\mathbf{r}\rangle\langle\mathbf{r}+\mathbf{n}|\rho\rangle\,d^3r\\ &= \sum_{\mathbf{k}'}\langle\mathbf{k}|\mathbf{k}'\rangle^{-1}\int\langle\mathbf{k}'|\mathbf{r}\rangle\langle\mathbf{r}|\rho\rangle\,d^3r\\ &= V^{-1}\langle\mathbf{k}|\rho\rangle\end{aligned} \tag{19}$$

to minimize the sum of squared errors $\langle\Delta|\Delta\rangle$, defined by

$$\Delta(\mathbf{r}) = \sum_{\mathbf{n}}\rho(\mathbf{r}+\mathbf{n}) - \mathrm{Re}\sum_{\mathbf{k}}\langle\mathbf{r}|\mathbf{k}\rangle c_{\mathbf{k}}^* \tag{20}$$

In other words, eq 19 is identical to having determined $a_{\mathbf{k}}$ and $b_{\mathbf{k}}$ by solving the two (uncoupled) equations

$$\frac{\partial\langle\Delta|\Delta\rangle}{\partial a_{\mathbf{k}}} = 0$$

and

$$\frac{\partial\langle\Delta|\Delta\rangle}{\partial b_{\mathbf{k}}} = 0$$

respectively. By integrating over the volume of a unit cell, the first line of eq 19 effectively integrates one instance of $\rho(\mathbf{r})$ that has been wrapped into the cell's interior. The second line of eq 19 exploits the periodicity of the plane waves by "unwrapping" the density and modifying the integration limits accordingly.

One can deduce that $\cos(-x) = \cos(x)$ and $\sin(-x) = -\sin(x)$ imply the symmetries $a_{-\mathbf{k}} = a_{\mathbf{k}}$, $b_{-\mathbf{k}} = -b_{\mathbf{k}}$, and $c_{-\mathbf{k}} = c_{\mathbf{k}}^*$. As a consequence of these symmetries,

$$\langle\mathbf{r}|-\mathbf{k}\rangle c_{-\mathbf{k}}^* = (\langle\mathbf{r}|\mathbf{k}\rangle c_{\mathbf{k}}^*)^*$$

and, therefore, the sum over *all* $\mathbf{k}$ in eq 18 naturally acts to cancel the imaginary components from each term. That is, although each term is complex, the other half of the sum adds its conjugate. The appearance of "Re" in our description of Ewald, PME, and FFP is to remind the reader that the imaginary numbers vanish and only the real component of each term thus needs to be computed. Equation 18 is often referred to as a "complex-to-real" reverse (or inverse) Fourier transform and is typically implemented within computer software to utilize only a subset of symmetry-unique $c_{\mathbf{k}}^*$ values that are provided as input.

Weighted least-squares fits can be performed to generalize the plane wave expansion described above. These fits produce coefficients

$$c_{\mathbf{k}}^* = \sum_{\mathbf{k}'}\langle\mathbf{k}|\hat{O}|\mathbf{k}'\rangle^{-1}\langle\mathbf{k}'|\hat{O}|\rho\rangle \tag{21}$$

that minimize the sum of squared errors $\langle\Delta|\hat{O}|\Delta\rangle$ weighted by a linear Hermitian operator $\hat{O}$. In the general case, the components of $c_{\mathbf{k}}$ store the result of a fit that may have required a coupled (as opposed to independent) solution to $a_{\mathbf{k}}$ and $b_{\mathbf{k}}$. For the specific case $\langle\mathbf{r}|\hat{O}|\mathbf{r}'\rangle = |\mathbf{r}-\mathbf{r}'|^{-1}$, eq 21 is called an "electrostatic fit". The coefficients of an electrostatic fit reduce to eq 19, because the plane waves are eigenfunctions of the Coulomb operator (eq 16). In other words, eq 19 are the plane wave expansion coefficients that best reproduce the electrostatic potential of the periodic density.

2.3. Multipolar Ewald. One could, in principle, directly project $\rho(\mathbf{r})$ into the periodic basis and compute the interaction from the plane wave representation; in practice, however, this is intractable because an infinite number of plane waves would be required to model $\delta(\mathbf{r})$. Therefore, the Ewald method abandons the direct solution of eq 1, preferring instead to decompose the periodic density into smooth and discontinuous components:

$$\begin{aligned}\sum_{\mathbf{n}}\rho(\mathbf{r}+\mathbf{n}) = \sum_{n}\tilde{\rho}(\mathbf{r}+\mathbf{n}) + \sum_{\mathbf{n}}[\rho(\mathbf{r}+\mathbf{n})\\ -\tilde{\rho}(\mathbf{r}+\mathbf{n})]\end{aligned} \tag{22}$$

whereby the smooth model density composed of Gaussians

DOI: 10.1021/ct5007983
J. Chem. Theory Comput. 2015, 11, 436−450

$$\tilde{\rho}(\mathbf{r}) = \sum_{a,l\mu\in a} q_{a,l\mu}\frac{C_{l\mu}(\nabla_a)}{(2l-1)!!}\chi(\mathbf{r}-\mathbf{R}_a) \tag{23}$$

is well-approximated from a linear least-squares fit to reasonably few plane waves:

$$\begin{aligned}\sum_{\mathbf{n}}\tilde{\rho}(\mathbf{r}+\mathbf{n}) &= \mathrm{Re}\sum_{\mathbf{k},\mathbf{k}'}\langle\mathbf{r}|\mathbf{k}\rangle\langle\mathbf{k}|\mathbf{k}'\rangle^{-1}\langle\mathbf{k}'|\tilde{\rho}\rangle \\ &= \mathrm{Re}\frac{1}{V}\sum_{\mathbf{k}}\langle\mathbf{r}|\mathbf{k}\rangle\langle\mathbf{k}|\chi\rangle S_{\mathbf{k}}\end{aligned} \tag{24}$$

Furthermore, the model density was chosen to reproduce the long-range electrostatic potential of the point multipoles to ensure that only short-ranged corrections are required. In our notation, $\langle\mathbf{k}|\tilde{\rho}\rangle$ and $\langle\mathbf{k}|\chi\rangle = \mathrm{e}^{-(k^2/4\zeta)}$ are Fourier coefficients of the Gaussian density and Gaussian monopole, respectively. The summation over $\mathbf{k}$ in eq 24 is a reverse Fourier transform, and $S_{\mathbf{k}}$ is a "structure factor".

$$\begin{aligned}S_{\mathbf{k}} &= \sum_{a,l\mu\in a} q_{a,l\mu}\frac{C_{l\mu}(\nabla_a)}{(2l-1)!!}\langle\mathbf{k}|\mathbf{R}_a\rangle \\ &= \sum_{a,l\mu\in a} q_{a,l\mu}\frac{C_{l\mu}(\mathbf{k})(-i)^l}{(2l-1)!!}\langle\mathbf{k}|\mathbf{R}_a\rangle\end{aligned} \tag{25}$$

The Ewald energy is the interaction of $\rho(\mathbf{r})$ with the electrostatic potential of the plane wave projected periodic Gaussian density $\tilde{\phi}(\mathbf{r})$ (see eq 32) upon correcting the short-ranged differences between the point and Gaussian potentials $\Delta\phi(\mathbf{r})$ (see eq 27):

$$\begin{aligned}E &= \frac{1}{2}\langle\rho|\tilde{\phi}\rangle + \frac{1}{2}\langle\rho|\Delta\phi\rangle \\ &= \frac{1}{2}\sum_{a,l\mu\in a} q_{a,l\mu}(\tilde{p}_{a,l\mu} + \Delta p_{a,l\mu})\end{aligned} \tag{26}$$

where $\tilde{p}_{a,l\mu} = \partial\langle\rho|\tilde{\phi}\rangle/\partial q_{a,l\mu}$ and $\Delta p_{a,l\mu} = \partial\langle\rho|\Delta\phi\rangle/\partial q_{a,l\mu}$ are the corresponding "multipolar potentials". The real-space corrections to the potential, energy, and multipolar potential are given as follows:

Potential Correction:

$$\begin{aligned}\Delta\phi(\mathbf{r}) &= \sum_{\mathbf{n}}\int\frac{\rho(\mathbf{r}'+\mathbf{n}) - \tilde{\rho}(\mathbf{r}'+\mathbf{n})}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}^3r' \\ &\approx \sum_{b,j\kappa\in b} q_{b,j\kappa}\frac{C_{j\kappa}(\nabla_b)}{(2j-1)!!}\frac{\mathrm{erfc}(\sqrt{\zeta}|\mathbf{r}-\mathbf{R}_b|)}{|\mathbf{r}-\mathbf{R}_b|}\end{aligned} \tag{27}$$

Energy Correction:

$$\begin{aligned}\frac{1}{2}\langle\rho|\Delta\phi\rangle \approx \frac{1}{2}&\sum_{\substack{a,l\mu\in a\\ b\neq a,j\kappa\in b}} q_{a,l\mu}q_{b,j\kappa}T_{l\mu,j\kappa} \\ &- \frac{1}{2}\sum_{a,l\mu\in a} q_{a,l\mu}^2\sqrt{\frac{4\zeta}{\pi}}\frac{(2\zeta)^l}{(2l+1)!!}\end{aligned} \tag{28}$$

Multipolar Potential Correction:

$$\Delta p_{a,l\mu} \approx \sum_{b\neq a,j\kappa\in b} q_{b,j\kappa}T_{l\mu,j\kappa} - q_{a,l\mu}\sqrt{\frac{4\zeta}{\pi}}\frac{(2\zeta)^l}{(2l+1)!!} \tag{29}$$

where

$$T_{l\mu,j\kappa} = \frac{C_{l\mu}(\nabla_a)}{(2l-1)!!}\frac{C_{j\kappa}(\nabla_b)}{(2j-1)!!}\frac{\mathrm{erfc}(\sqrt{\zeta}R_{ab})}{R_{ab}} \tag{30}$$

These corrections are assumed to be sufficiently short-ranged, such that only nearby *minimum image* separations R_{ab} and $|\mathbf{r}-\mathbf{R}_b|$ need to be considered. An efficient algorithm for computing eq 30 and its gradients, particularly for low-order harmonics, is described in ref 56. The application of that algorithm requires the following definition of the "auxiliary vector":

$$\begin{aligned}O^{(n)} &= \left(\frac{\mathrm{d}}{\mathrm{d}R_{ab}^2}\right)^n\frac{\mathrm{erfc}(\sqrt{\zeta}R_{ab})}{R_{ab}} \\ &= \frac{(2n-1)!!}{R_{ab}}\left(-\frac{1}{2R_{ab}^2}\right)^n - \frac{2}{\sqrt{\pi}}\sqrt{\zeta}(-\zeta)^n F_n(\zeta R_{ab}^2)\end{aligned} \tag{31}$$

where $F_n(x)$ is the Boys function.[63]

The reciprocal-space Ewald potential, energy, multipolar potential, and contribution to the gradient are readily obtained upon noting $\langle\delta|\mathbf{k}\rangle = 1$.

Reciprocal-Space Ewald Potential:

$$\begin{aligned}\tilde{\phi}(\mathbf{r}) &= \sum_{\mathbf{n}}\int\frac{\tilde{\rho}(\mathbf{r}'+\mathbf{n})}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}^3r' \\ &= \mathrm{Re}\int\frac{1}{V}\sum_{\mathbf{k}}\frac{\langle\mathbf{r}'|\mathbf{k}\rangle\langle\mathbf{k}|\tilde{\rho}\rangle}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}^3r' \\ &= \mathrm{Re}\frac{1}{V}\sum_{\mathbf{k}\neq 0}\langle\mathbf{r}|\mathbf{k}\rangle\langle\mathbf{k}|\chi\rangle\frac{4\pi}{k^2}S_{\mathbf{k}}\end{aligned} \tag{32}$$

Energy:

$$\frac{1}{2}\langle\rho|\tilde{\phi}\rangle = \frac{1}{2V}\sum_{\mathbf{k}\neq 0}\langle\mathbf{k}|\chi\rangle\frac{4\pi}{k^2}|S_{\mathbf{k}}|^2 \tag{33}$$

Multipolar Potential:

$$\tilde{p}_{a,l\mu} = \mathrm{Re}\frac{1}{V}\sum_{\mathbf{k}\neq 0}\frac{C_{l\mu}(\mathbf{k})i^l}{(2l-1)!!}\langle\mathbf{R}_a|\mathbf{k}\rangle\langle\mathbf{k}|\chi\rangle\frac{4\pi}{k^2}S_{\mathbf{k}} \tag{34}$$

Contribution to the Gradient:

$$\nabla_a\frac{1}{2}\langle\rho|\tilde{\phi}\rangle = \mathrm{Re}\frac{1}{V}\sum_{\mathbf{k}\neq 0}\sum_{l\mu} q_{a,l\mu}\frac{C_{l\mu}(\mathbf{k})i^{l+1}\mathbf{k}}{(2l-1)!!}\langle\mathbf{R}_a|\mathbf{k}\rangle\langle\mathbf{k}|\chi\rangle\frac{4\pi}{k^2}S_{\mathbf{k}} \tag{35}$$

The extension of Ewald to atom-centered point multipole expansions presented here differs from the standard point-charge Ewald method only by the inclusion of STGOs in the structure factor and real-space correction.

2.4. Multipolar Particle Mesh Ewald. A linear-scaling implementation of the real-space Ewald corrections requires only short-range cutoffs and a sufficiently large Gaussian exponent; however, by fixing the cutoff—and, hence, the exponent—the computational complexity of the reciprocal-space Ewald potential then scales with $O(N^2)$, because

(1) the number of plane waves must increase with the size of the system, to maintain consistent resolution of the periodic density, and

DOI: 10.1021/ct5007983
J. Chem. Theory Comput. 2015, 11, 436−450

(2) the calculation of each $S_{\mathbf{k}}$ requires a summation over the number of particles.

In the unlikely event that the particles happened to be positioned such that they formed a uniformly spaced grid, however, then the expression for $S_{\mathbf{k}}$ would become a sum over grid points (see eq 48) and the contribution of all grid points to all $S_{\mathbf{k}}$ could be computed O(N log N) using a Fast Fourier Transform (FFT). The novelty of PME is its coercion of the particle positions to exploit this otherwise unlikely scenario to speed the calculation of the reciprocal-space potential and energy.

PME approximates the particle positions with a linear combination of predefined uniformly spaced FFT grid points $\mathbf{R}_{\mathbf{t}}$,

$$\langle \mathbf{R}_a| \approx \sum_{\mathbf{t}} \langle \mathbf{R}_{\mathbf{t}} - \mathbf{R}_a|\theta\rangle\langle \mathbf{R}_{\mathbf{t}}| \equiv \langle \tilde{\mathbf{R}}_a| \tag{36}$$

that can be used to compute Fourier coefficients from numerical quadrature. The Gaussian quadrature of plane waves is equivalent to Simpson's integration rule performed on a regular grid:

$$\langle \mathbf{k}|f\rangle = \int \langle \mathbf{k}|\mathbf{r}\rangle\langle \mathbf{r}|f\rangle \, \mathrm{d}^3r \approx \frac{V}{N} f_{\mathbf{k}} \tag{37}$$

where $f_{\mathbf{k}}$ is a discrete Fourier transform (DFT):

$$f_{\mathbf{k}} = \sum_{\mathbf{t}} \langle \mathbf{k}|\mathbf{R}_{\mathbf{t}}\rangle\langle \mathbf{R}_{\mathbf{t}}|f\rangle \tag{38}$$

$\mathbf{t}$ is the index of the grid ($\mathbf{t} = (t_1, t_2, t_3)$) and $\mathbf{R}_{\mathbf{t}}$ is the location of a FFT grid point:

$$\mathbf{R}_{\mathbf{t}} = \mathbf{a}_1\left(\frac{t_1 - 1}{N_1}\right) + \mathbf{a}_2\left(\frac{t_2 - 1}{N_2}\right) + \mathbf{a}_3\left(\frac{t_3 - 1}{N_3}\right)$$

N_1, N_2, and N_3 are the number of FFT grid points in each lattice direction; $N = N_1N_2N_3$.

θ is a Cardinal B-spline weight evaluated about a particle position:

$$\theta(\mathbf{r} - \mathbf{R}_a) = \prod_{d=1}^{3} M_n\left(N_d(\mathbf{r} - \mathbf{R}_a)^{\mathrm{T}} \cdot \mathbf{a}_d^* + \frac{n}{2}\right) \tag{39}$$

In our notation, the argument of $M_n(u)$ is meant to be periodically wrapped into the range $(-N_d/2, N_d/2]$.

M_n is a Cardinal B-spline function that holds recursion properties for efficient evaluation of their values and gradients.[38]

$$M_n(u) = \frac{1}{(n-1)!} \sum_{k=0}^{n} (-1)^k \binom{n}{k} [\max(u - k, 0)]^{n-1} \tag{40}$$

The B-spline weights are real $\langle \mathbf{r}|\theta\rangle = \langle \theta|\mathbf{r}\rangle$, have even symmetry $\langle -\mathbf{r}|\theta\rangle = \langle \mathbf{r}|\theta\rangle$, sum to one ($\sum_{\mathbf{t}}\langle \mathbf{R}_{\mathbf{t}} - \mathbf{r}|\theta\rangle = 1 \ \forall \mathbf{r}$), and are nonzero only for those grid points near the particle. The number of grid points contributing to the approximate position of a particle is determined from the chosen B-spline order without regard to the FFT grid spacing.

Following the work of Schoenberg,[64−66] previous descriptions[38] of PME have used Cardinal B-splines as a linear basis to construct "exponential Euler splines." In that point of view, the structure factors are computed from approximate plane waves that are spline-evaluated at the actual particle positions. We below choose an alternate perspective by applying eq 36 to obtain the same outcome, which we then relate to ref 38.

The charge density is composed of atom-centered functions; therefore, it is convenient to re-express eq 36 as

$$\langle \mathbf{r} - \mathbf{R}_a| \approx \sum_{\mathbf{t}} \langle \mathbf{R}_{\mathbf{t}} - \mathbf{R}_a|\theta\rangle\langle \mathbf{r} - \mathbf{R}_{\mathbf{t}}| \equiv \langle \widetilde{\mathbf{r} - \mathbf{R}_a}| \tag{41}$$

which is deduced from the invariance of the relative particle and grid positions upon their simultaneous reflection about the origin and subsequent translation by $+\mathbf{r}$; that is, $\mathbf{R}_a \to \mathbf{r} - \mathbf{R}_a$ and $\mathbf{R}_{\mathbf{t}} \to \mathbf{r} - \mathbf{R}_{\mathbf{t}}$. An atom-centered Gaussian is thus approximated by a linear combination of Gaussians centered about the FFT grid points:

$$\langle \widetilde{\mathbf{r} - \mathbf{R}_a}|\chi\rangle = \sum_{\mathbf{t}} \langle \mathbf{R}_{\mathbf{t}} - \mathbf{R}_a|\theta\rangle\langle \mathbf{r} - \mathbf{R}_{\mathbf{t}}|\chi\rangle \tag{42}$$

However, this has the unintended consequence of making the approximate Gaussian appear artificially diffuse as the B-spline order is increased. Fortunately, if the B-spline is of sufficient order to be accurately integrated by the discrete Fourier transform, then the Fourier coefficients of the approximate Gaussian

$$\begin{aligned} \int \langle \mathbf{k}|\mathbf{r}\rangle\langle \widetilde{\mathbf{r} - \mathbf{R}_a}|\chi\rangle \, \mathrm{d}^3r &= \sum_{\mathbf{t}} \langle \mathbf{R}_{\mathbf{t}} - \mathbf{R}_a|\theta\rangle\langle \mathbf{k}|\mathbf{R}_{\mathbf{t}}\rangle\langle \mathbf{k}|\chi\rangle \\ &= \sum_{\mathbf{t}} \langle \mathbf{k}|\mathbf{R}_{\mathbf{t}} - \mathbf{R}_a\rangle\langle \mathbf{R}_{\mathbf{t}} - \mathbf{R}_a|\theta\rangle \\ &\quad \times \langle \mathbf{k}|\mathbf{R}_a\rangle\langle \mathbf{k}|\chi\rangle \\ &= \theta_{\mathbf{k}}\langle \mathbf{k}|\mathbf{R}_a\rangle\langle \mathbf{k}|\chi\rangle \end{aligned} \tag{43}$$

are those of the true Gaussian,

$$\int \langle \mathbf{k}|\mathbf{r}\rangle\langle \mathbf{r} - \mathbf{R}_a|\chi\rangle \, \mathrm{d}^3r = \langle \mathbf{k}|\mathbf{R}_a\rangle\langle \mathbf{k}|\chi\rangle \tag{44}$$

scaled by the B-spline DFT coefficients $\theta_{\mathbf{k}}$. Note that $\theta_{\mathbf{k}} = \theta_{\mathbf{k}}^*$, because the B-splines are even functions. Repeating this procedure for $\langle \widetilde{\mathbf{r} - \mathbf{R}_a}|\delta\rangle$, or any function $|f\rangle$, yields analogous results. In other words, the Fourier coefficients of an approximately positioned function should be scaled by $\theta_{\mathbf{k}}^{-1}$.

Let us continue on a brief aside that relates eq 36 to the exponential Euler splines used in ref 38 by considering its application to $\langle \mathbf{k}|\mathbf{k}'\rangle = \delta_{\mathbf{k},\mathbf{k}'}V$ in the manner above using either of the relations

$$\begin{aligned} \langle \tilde{\mathbf{r}}| &= \lim_{\mathbf{R}_a \to 0} \langle \widetilde{\mathbf{r} - \mathbf{R}_a}| = \sum_{\mathbf{t}} \langle \mathbf{R}_{\mathbf{t}}|\theta\rangle\langle \mathbf{r} - \mathbf{R}_{\mathbf{t}}| \\ &= \lim_{\mathbf{R}_a \to \mathbf{r}} \langle \tilde{\mathbf{R}}_a| = \sum_{\mathbf{t}} \langle \mathbf{R}_{\mathbf{t}} - \mathbf{r}|\theta\rangle\langle \mathbf{R}_{\mathbf{t}}| \end{aligned} \tag{45}$$

The result $\int_V \langle \mathbf{k}|\mathbf{r}\rangle\langle \tilde{\mathbf{r}}|\mathbf{k}'\rangle \, \mathrm{d}^3r = \theta_{\mathbf{k}}\langle \mathbf{k}|\mathbf{k}'\rangle$ implies the need to effectively renormalize an "approximate plane wave" upon grid interpolation; that is, $\langle \mathbf{r}|\mathbf{k}\rangle \approx \langle \tilde{\mathbf{r}}|\mathbf{k}\rangle\theta_{\mathbf{k}}^{-1}$, which are the exponential splines. The relationship between eq 36 and the exponential Euler splines is the interpretation of whether it is the approximate plane wave or the Fourier coefficient of the approximately positioned function that is renormalized:

DOI: 10.1021/ct5007983
J. Chem. Theory Comput. 2015, 11, 436−450

$$\langle \mathbf{k}|\mathbf{R}_a\rangle\langle \mathbf{k}|\chi\rangle = \int \frac{\langle \mathbf{k}|\tilde{\mathbf{r}}\rangle}{\theta_{\mathbf{k}}}\langle \mathbf{r}-\mathbf{R}_a|\chi\rangle \, \mathrm{d}^3r = \frac{1}{\theta_{\mathbf{k}}}\int \langle \mathbf{k}|\mathbf{r}\rangle\widetilde{\langle \mathbf{r}-\mathbf{R}_a|\chi\rangle} \, \mathrm{d}^3r \tag{46}$$

The subtle distinction between these interpretations has no practical consequence on the PME method other than, perhaps, how the mathematical terms are grouped together; eq 47 scales the structure factor by $\theta_{\mathbf{k}}^{-1}$, whereas ref 38 absorbs it into the definition of $S_{\mathbf{k}}$.

The periodic representation of the model Gaussian density (eq 23) composed of the approximate Gaussians (eq 42) is, upon scaling the Fourier coefficients,

$$\sum_{\mathbf{n}} \tilde{\rho}(\mathbf{r}+\mathbf{n}) \approx \mathrm{Re}\frac{1}{V}\sum_{\mathbf{k}} \langle \mathbf{r}|\mathbf{k}\rangle \frac{\langle \mathbf{k}|\chi\rangle}{\theta_{\mathbf{k}}} S_{\mathbf{k}}^{\mathrm{pme}} \tag{47}$$

where the PME structure factor,

$$S_{\mathbf{k}}^{\mathrm{pme}} = \sum_{\mathbf{t}} \langle \mathbf{k}|\mathbf{R}_{\mathbf{t}}\rangle\langle \mathbf{R}_{\mathbf{t}}|Q\rangle \tag{48}$$

is the forward DFT of

$$\langle \mathbf{R}_{\mathbf{t}}|Q\rangle = \sum_{a,l\mu\in a} q_{a,l\mu}\frac{C_{l\mu}(\nabla_a)}{(2l-1)!!}\langle \mathbf{R}_{\mathbf{t}}-\mathbf{R}_a|\theta\rangle \tag{49}$$

whose evaluation avoids $O(N^2)$ operations by the locality of $\theta(\mathbf{r})$.

The PME reciprocal-space potential is the electrostatic potential of eq 47,

$$\begin{aligned}\tilde{\phi}(\mathbf{r}) &\approx \mathrm{Re}\frac{1}{V}\sum_{\mathbf{k}\neq 0}\langle \mathbf{r}|\mathbf{k}\rangle\frac{\langle \mathbf{k}|\chi\rangle(4\pi/k^2)}{\theta_{\mathbf{k}}}S_{\mathbf{k}}^{\mathrm{pme}}\\ &\approx \mathrm{Re}\frac{1}{V}\sum_{\mathbf{k}\neq 0}\langle \tilde{\mathbf{r}}|\mathbf{k}\rangle\frac{\langle \mathbf{k}|\chi\rangle(4\pi/k^2)}{|\theta_{\mathbf{k}}|^2}S_{\mathbf{k}}^{\mathrm{pme}}\\ &\approx \mathrm{Re}\sum_{\mathbf{t}}\langle \theta|\mathbf{R}_{\mathbf{t}}-\mathbf{r}\rangle\frac{1}{V}\sum_{\mathbf{k}\neq 0}\langle \mathbf{R}_{\mathbf{t}}|\mathbf{k}\rangle\frac{\langle \mathbf{k}|\chi\rangle(4\pi/k^2)}{|\theta_{\mathbf{k}}|^2}S_{\mathbf{k}}^{\mathrm{pme}}\end{aligned} \tag{50}$$

and yields the following PME reciprocal-space energy, multipolar potential, and contribution to the gradient:

PME Reciprocal-Space Energy:

$$\frac{1}{2}\langle \rho|\tilde{\phi}\rangle \approx \frac{1}{2V}\sum_{\mathbf{k}\neq 0}\frac{\langle \mathbf{k}|\chi\rangle(4\pi/k^2)}{|\theta_{\mathbf{k}}|^2}|S_{\mathbf{k}}^{\mathrm{pme}}|^2 \tag{51}$$

Multipolar Potential:

$$\tilde{p}_{a,l\mu} \approx \mathrm{Re}\sum_{\mathbf{t}}\frac{C_{l\mu}(\nabla_a)}{(2l-1)!!}\langle \theta|\mathbf{R}_{\mathbf{t}}-\mathbf{R}_a\rangle \times \frac{1}{V}\sum_{\mathbf{k}\neq 0}\langle \mathbf{R}_{\mathbf{t}}|\mathbf{k}\rangle\frac{\langle \mathbf{k}|\chi\rangle(4\pi/k^2)}{|\theta_{\mathbf{k}}|^2}S_{\mathbf{k}}^{\mathrm{pme}} \tag{52}$$

Contribution to the Gradient:

$$\nabla_a\frac{1}{2}\langle \rho|\tilde{\phi}\rangle \approx \mathrm{Re}\sum_{\mathbf{t}}\left(\sum_{l\mu} q_{a,l\mu}\frac{C_{l\mu}(\nabla_a)}{(2l-1)!!}\nabla_a\langle \theta|\mathbf{R}_t-\mathbf{R}_a\rangle\right) \times \frac{1}{V}\sum_{\mathbf{k}\neq 0}\langle \mathbf{R}_{\mathbf{t}}|\mathbf{k}\rangle\frac{\langle \mathbf{k}|\chi\rangle(4\pi/k^2)}{|\theta_{\mathbf{k}}|^2}S_{\mathbf{k}}^{\mathrm{pme}} \tag{53}$$

The extension of PME to atom-centered point multipole expansions presented here differs from the standard point charge PME method only by the inclusion of STGOs acting upon the B-spline in eqs 49−51 and, like Ewald, the real-space correction (eqs 27 and 28).

Although the STGO has many remarkable properties[61] when acting upon a spherical function, a regular or irregular solid harmonic, or a product of the these functions, the B-spline weights cannot take advantage of these properties. Instead, one must treat the STGO as a linear combination of gradient operators,

$$C_{lm}^{\mathrm{c/s}}(\nabla) = \sum_{ijk} a_{ijk}^{lm,\mathrm{c/s}}\frac{\mathrm{d}^i}{\mathrm{d}x^i}\frac{\mathrm{d}^j}{\mathrm{d}y^j}\frac{\mathrm{d}^k}{\mathrm{d}z^k} \tag{54}$$

whose coefficients are exactly those that relate $C_{lm}^{\mathrm{c/s}}(\mathbf{r})$ to Cartesian monomials:[67]

$$a_{ijk}^{lm,\mathrm{c/s}} = \sum_{t=0}^{\lceil (l-m)/2\rceil}\sum_{u=0}^{t}\sum_{s=0}^{\lceil (m-o^{\mathrm{c/s}})/2\rceil} d_{tus,ijk}^{lm,\mathrm{c/s}} \times \binom{L-t}{m+t}\binom{L}{t}\binom{t}{u}\binom{m}{2s+o^{\mathrm{c/s}}} \tag{55}$$

where

$$d_{tus,ijk}^{lm,\mathrm{c/s}} = \begin{cases} 0 & \text{if } m+2(t-u-s)-o^{\mathrm{c/s}} \neq i \\ 0 & \text{if } 2(u+s)+o^{\mathrm{c/s}} \neq j \\ 0 & \text{if } l-m-2t \neq k \\ (-1)^{t+s}\dfrac{\sqrt{(2-\delta_{m,0})(l+m)!(l-m)!}}{l!2^{m+2t}} & \text{otherwise} \end{cases} \tag{56}$$

and $o^{\mathrm{c}} = 0$ and $o^{\mathrm{s}} = 1$. In the above notation, $m = |\mu|$, $C_{l\mu\geq 0}(\nabla) = C_{lm}^{\mathrm{c}}(\nabla)$, and $C_{l\mu<0}(\nabla) = C_{lm}^{\mathrm{s}}(\nabla)$.

2.5. Multipolar Fast Fourier−Poisson Method. Fast Fourier−Poisson (FFP), like Ewald, abandons the direct solution to eq 1 by introducing a Gaussian density whose potential can be determined from a plane wave expansion.[41] Unlike the Ewald method, FFP evaluates the plane wave expansion coefficients from numerical quadrature to take advantage of FFTs. The electrostatic potential of the plane-wave-resolved Gaussian density is computed in Fourier space, evaluated on a regular grid by means of a reverse FFT, and then used to numerically integrate the Coulomb self-energy of the Gaussian density. Therefore, the FFP real-space corrections account for the short-ranged differences between the Gaussian and point interaction energies:

$$\begin{aligned}E &= \frac{1}{2}\langle \tilde{\rho}|\tilde{\phi}\rangle + \frac{1}{2}(\langle \rho|\phi\rangle - \langle \tilde{\rho}|\tilde{\phi}\rangle)\\ &= \frac{1}{2}\langle \tilde{\rho}|\tilde{\phi}\rangle + \frac{1}{2}\sum_{\substack{a,l\mu\in a\\ b,j\kappa\in b}} q_{a,l\mu}q_{b,j\kappa}\Delta T_{l\mu,j\kappa}\end{aligned} \tag{57}$$

DOI: 10.1021/ct5007983
J. Chem. Theory Comput. 2015, 11, 436−450

The plane wave representation of the periodic model density composed of the Gaussian multipoles is

$$\sum_{\mathbf{n}} \tilde{\rho}(\mathbf{r}+\mathbf{n}) = \mathrm{Re}\frac{1}{V}\sum_{\mathbf{k}} \langle \mathbf{r}|\mathbf{k}\rangle\langle \mathbf{k}|\tilde{\rho}\rangle \tag{58}$$

where

$$\begin{aligned}\langle \mathbf{k}|\tilde{\rho}\rangle &\approx \frac{V}{N}\sum_{\mathbf{t}} \langle \mathbf{k}|\mathbf{R}_{\mathbf{t}}\rangle\langle \mathbf{R}_{\mathbf{t}}|\tilde{\rho}\rangle \\ &\approx \frac{V}{N}\sum_{\mathbf{t}} \langle \mathbf{k}|\mathbf{R}_{\mathbf{t}}\rangle\sum_{a,l\mu} q_{a,l\mu}\langle \mathbf{R}_{\mathbf{t}}-\mathbf{R}_a|\chi_{l\mu}\rangle\end{aligned} \tag{59}$$

The evaluation of $\langle \mathbf{R}_{\mathbf{t}}|\tilde{\rho}\rangle$ avoids $O(N^2)$ operations by evaluating $\chi_{l\mu}(\mathbf{R}_{\mathbf{t}}-\mathbf{R}_a)$ only for minimum image distances within a suitable cutoff. The real-space energy corrections, $\Delta T_{l\mu,j\kappa}$, vanish when the Gaussians do not overlap

$$\begin{aligned}\Delta T_{l\mu,j\kappa} = \sum_{\mathbf{n}}\int\int \mathrm{d}^3r\ \mathrm{d}^3r' |\mathbf{r}-\mathbf{r}'|^{-1} \\ \times [\delta_{l\mu}(\mathbf{r}-\mathbf{R}_a)\delta_{j\kappa}(\mathbf{r}'-\mathbf{R}_b+\mathbf{n})s^{a,b}_{\mathbf{n},0} \\ - \chi_{l\mu}(\mathbf{r}-\mathbf{R}_a)\chi_{j\kappa}(\mathbf{r}'-\mathbf{R}_b+\mathbf{n})]\end{aligned} \tag{60}$$

where $s^{a,b}_{\mathbf{n},0} = 1-\delta_{n_1 0}\delta_{n_2 0}\delta_{n_3 0}\delta_{ab}$). Therefore, eq 60 is well-approximated by computing the integrals for the short-ranged, minimum-image R_{ab} separations only. Explicit expressions for the real-space energy correction, multipolar potentials, and gradient contribution are analogous to those shown in eqs 28−31 upon substituting ζ with $\zeta/2$, which arises from the Gaussian product theorem.

The FFP reciprocal-space energy, multipolar potential, and contribution to the gradient are as follows:

FFP Reciprocal-Space Energy:

$$\frac{1}{2}\langle \tilde{\rho}|\tilde{\phi}\rangle = \frac{1}{2V}\sum_{\mathbf{k}\neq 0}\frac{4\pi}{k^2}|\langle \mathbf{k}|\tilde{\rho}\rangle|^2 \tag{61}$$

Multipolar Potential:

$$p_{a,l\mu} = \frac{V}{N}\sum_{\mathbf{t}}\chi_{l\mu}(\mathbf{R}_{\mathbf{t}}-\mathbf{R}_a)\frac{1}{V}\mathrm{Re}\sum_{\mathbf{k}\neq 0}\langle \mathbf{R}_{\mathbf{t}}|\mathbf{k}\rangle\frac{4\pi}{k^2}\langle \mathbf{k}|\tilde{\rho}\rangle \tag{62}$$

Contribution to the Gradient:

$$\begin{aligned}\nabla_a\frac{1}{2}\langle \tilde{\rho}|\tilde{\phi}\rangle = \frac{V}{N}\sum_{\mathbf{t}}\sum_{l\mu} q_{a,l\mu}(\nabla_a\chi_{l\mu}(\mathbf{R}_{\mathbf{t}}-\mathbf{R}_a)) \\ \times \frac{1}{V}\mathrm{Re}\sum_{\mathbf{k}\neq 0}\langle \mathbf{R}_{\mathbf{t}}|\mathbf{k}\rangle\frac{4\pi}{k^2}\langle \mathbf{k}|\tilde{\rho}\rangle\end{aligned} \tag{63}$$

The gradient of the Gaussian basis,

$$\nabla\chi_{l\mu}(\mathbf{r}) = \frac{(2\zeta)^l}{(2l-1)!!}\chi(\mathbf{r})A_{l\mu}(\nabla R_{l\mu}(\mathbf{r}) - 2\zeta\mathbf{r}R_{l\mu}(\mathbf{r})) \tag{64}$$

is efficiently computed using the following derivative properties:[58]

$$\frac{\mathrm{d}R^{\mathrm{c/s}}_{lm}(\mathbf{r})}{\mathrm{d}x} = \frac{1}{2}(R^{\mathrm{c/s}}_{l-1,m+1}(\mathbf{r}) - R^{\mathrm{c/s}}_{l-1,m-1}(\mathbf{r})) \tag{65}$$

$$\frac{\mathrm{d}R^{\mathrm{c/s}}_{lm}(\mathbf{r})}{\mathrm{d}y} = \pm\frac{1}{2}(R^{\mathrm{s/c}}_{l-1,m+1}(\mathbf{r}) + R^{\mathrm{s/c}}_{l-1,m-1}(\mathbf{r})) \tag{66}$$

$$\frac{\mathrm{d}R^{\mathrm{c/s}}_{lm}(\mathbf{r})}{\mathrm{d}z} = R^{\mathrm{c/s}}_{l-1,m}(\mathbf{r}) \tag{67}$$

When negative m values are encountered, the reader is implicitly instructed to apply the symmetry property $R^{\mathrm{c/s}}_{l,-m}(\mathbf{r}) = \pm(-1)^m R^{\mathrm{c/s}}_{lm}(\mathbf{r})$, which follows directly from $R^{*}_{lm}(\mathbf{r}) = (-1)^m R^{\mathrm{c/s}}_{l,-m}(\mathbf{r})$, where the plus/minus sign ($\pm$) corresponds to the cosine/sine designation.

2.6. Correction for Charged Systems. Let $\langle \mathbf{r}|\hat{j}|\mathbf{r}'\rangle = \sum_{\mathbf{n}}|\mathbf{r}-\mathbf{r}'+\mathbf{n}|^{-1}$; then, we can rewrite eq 1.

$$\begin{aligned}E &= \frac{1}{2}\langle \rho|\hat{j}|\rho\rangle \\ &= \frac{1}{2}\langle \rho|\hat{j}|\tilde{\rho}\rangle + \frac{1}{2}\langle \rho|\hat{j}|\rho-\tilde{\rho}\rangle\end{aligned} \tag{68}$$

If the net charge is nonzero ($\int\rho(\mathbf{r})\,\mathrm{d}^3r = Q$), then the system is nonphysical and this definition of the energy is infinite. Although it is possible to neutralize the density in arbitrarily different ways to produce different energies, a particularly convenient choice is made by introducing a uniform background density,[68]

$$\bar{\rho}(\mathbf{r}) = \begin{cases} Q/V & \text{if } \mathbf{r} \text{ in unit cell} \\ 0 & \text{otherwise}\end{cases} \tag{69}$$

and $\sum_{\mathbf{n}}\bar{\rho}(\mathbf{r}+\mathbf{n}) = Q/V$, everywhere. With this choice of neutralization, the energy is

$$\begin{aligned}E_Q &\equiv \frac{1}{2}\langle \rho-\bar{\rho}|\hat{j}|\rho-\bar{\rho}\rangle \\ &= \frac{1}{2}\langle \rho|\hat{j}|\tilde{\rho}-\bar{\rho}\rangle + \frac{1}{2}\langle \rho|\hat{j}|\rho-\tilde{\rho}\rangle - \frac{1}{2}\langle \bar{\rho}|\hat{j}|\tilde{\rho}-\bar{\rho}\rangle \\ &\quad - \frac{1}{2}\langle \bar{\rho}|\hat{j}|\rho-\tilde{\rho}\rangle\end{aligned} \tag{70}$$

However, ${}^1/_2\langle \rho|\hat{j}|\tilde{\rho}-\bar{\rho}\rangle = {}^1/_2\langle \rho|\tilde{\phi}\rangle$, as written in the previous sections, by virtue of excluding the $\mathbf{k}=0$ term in the Ewald summation, because it is charge-neutral. Similarly, ${}^1/_2\langle \bar{\rho}|\hat{j}|\tilde{\rho}-\bar{\rho}\rangle$ vanishes, because the exclusion of the $\mathbf{k}=0$ term produces a potential composed of functions that are each orthogonal to a constant over the range of the unit cell. The last term in eq 70 vanishes, by symmetry, for all nonzero multipoles and is well-approximated by an integral over all space if ζ is sufficiently large to extinguish the correction potential within the bounds of the unit cell (in a minimum image context). As a result of these properties, the energy of a charged system is

$$E_Q = E - \frac{\pi Q^2}{2\zeta V} \tag{71}$$

An analogous result is written for FFP by replacing ζ with $\zeta/2$.

For completeness, we note that it has been pointed out[69,70] that eq 1 is absolutely convergent only if the total dipole moment of the system $\boldsymbol{\mu}$ is zero; however, it is otherwise only conditionally convergent. In other words, the asymptotic value of eq 1 is dependent on the shape of the supercells used to replicate the system. A "dipole surface" term $2\pi|\boldsymbol{\mu}|^2/[(2\varepsilon_S+1)V]$ can be applied to mimic an infinite spherical sample of a cubic cell; however, this energy is usually ignored in Ewald

DOI: 10.1021/ct5007983
J. Chem. Theory Comput. 2015, 11, 436−450

implementations,[38] because its application to charged systems has a dependence on an arbitrary origin and it produces discontinuous energy changes when charged residues are wrapped into the primary unit cell.[71] By ignoring this term, we are said to employ "tinfoil boundary conditions", because it is equivalent to embedding the crystal inside a perfect conductor ($\varepsilon_S \to \infty$).

2.7. Error Analysis. Figure 1 compares the Ewald energy (eq 26) to a brute force evaluation of eq 1, as a function of

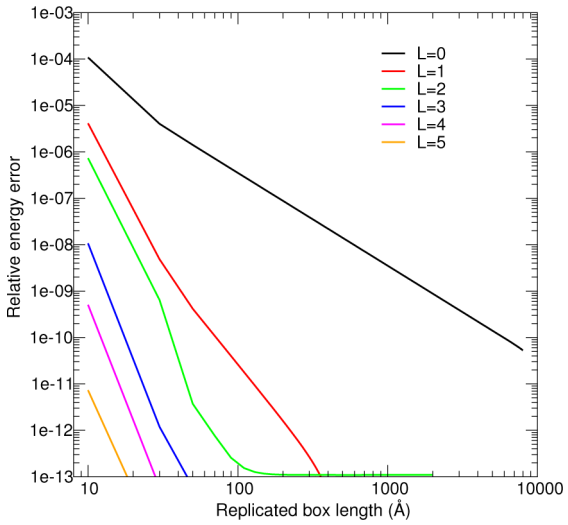


Figure 1. Comparison between the Ewald energy to the energy computed from explicit replication of a unit cell, as a function of replicated cell length for systems composed of the indicated multipole orders.

replicated unit cell length. The fundamental unit cell is a 10 Å cube containing three-point multipoles:

$$\rho(\mathbf{r}) = \sum_{l\mu} \delta_{l,L}\delta_{\mu,0}\delta_{l\mu}(\mathbf{r}) - \frac{\delta_{l,L}\delta_{\mu,0}}{2}(\delta_{l\mu}(\mathbf{r}-\mathbf{R}) + \delta_{l\mu}(\mathbf{r}+\mathbf{R})) \tag{72}$$

where $\mathbf{R} = (1\ \text{Å})\hat{z}$. The unit cell is replicated by extending the sides of the cell to form increasingly larger cubes. The various lines in Figure 1 indicate the nonzero angular moment in eq 72. The "relative energy error" is defined as

$$\text{relative energy error} = \sqrt{\frac{|E - E_n|^2}{E_n{}^2}}$$

where E is the Ewald energy evaluated with $|k_{max}| = 256$ and E_n is eq 1 after replication to the specified size. The relative error is bounded by $\geq 10^{-15}$, because of our use of double-precision arithmetic. The loss of one or two additional digits of precision should be expected, because of the accumulation of round-off errors.

Figure 2−4 make use of an *ad hoc* three-site water model to analyze the relative force errors and the real-space and reciprocal-space timings of Ewald, PME, and FFP, as a function of atomic multipole order. These three figures are the analysis of a box of 1024 waters whose bulk density has been equilibrated with TIP4P-Ew at 298 K. The 3-site multipolar water models used to perform the analysis are systematically constructed from a 5-site charge-only water model. The 3- and 5-site water models use the DFTB3[72] gas-phase structure of water (these coordinates are superimposed onto the TIP4P-Ew

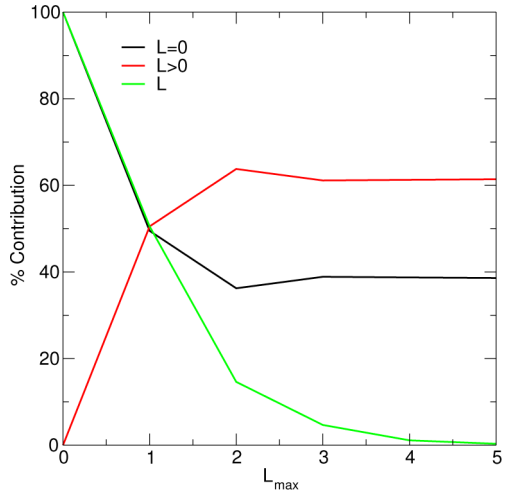


Figure 2. Decomposition of the intermolecular electrostatic forces using a box of 1024 waters. L_{max} indexes the *ad hoc* 3-site water model being considered, as distinguished by the multipole expansion order on the oxygen.

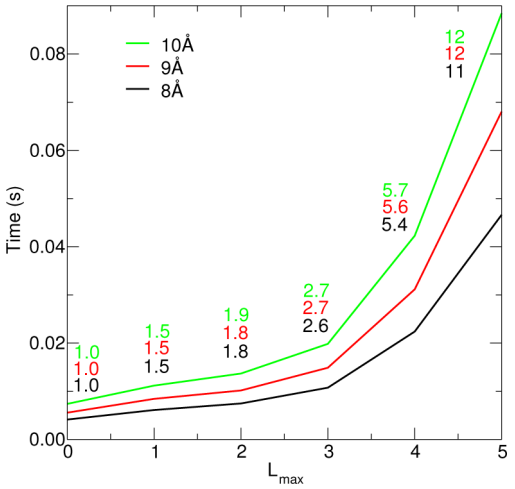


Figure 3. Wall clock time required to evaluate the real-space corrections (eq 29) for a box of 1024 waters using different real-space cutoffs. L_{max} indexes the *ad hoc* 3-site water model being used. The inset values are timing ratios $T(L_{max})/T(0)$, that is, the slow-down relative to the charge-only model.

waters): $R_{OH} = 0.957143$ Å, $R_{HH} = 1.572691$ Å, $\angle_{HOH} = 110.4812°$ (the 5-site model includes two additional charges located at the O−H bond midpoints). The atom charges (a.u.) ($q_O^{5\text{-site}} = -0.12404869708$, $q_H^{5\text{-site}} = 0.66106645662$) and the bond charges ($q_B^{5\text{-site}} = -0.59904210808$) were computed from the isolated water DFTB3 density matrix; that is, $q_O^{5\text{-site}}$ is the one-center atomic orbital (AO) product contribution to the Mulliken charge of O, and $q_B^{5\text{-site}}$ is the Mulliken bond charge between O and H. The multipole moments of the 3-site model are constructed from a Mulliken-like partitioning and subsequent solid harmonic translation of the 5-site model bond charges:

$$q_{O,l\mu}^{3\text{-site}} = q_O^{5\text{-site}}\delta_{l0}\delta_{\mu 0} + \frac{q_B^{5\text{-site}}}{2} \times \left[C_{l\mu}\left(\frac{\mathbf{R}_{H_a} - \mathbf{R}_O}{2}\right) + C_{l\mu}\left(\frac{\mathbf{R}_{H_b} - \mathbf{R}_O}{2}\right)\right] \tag{73}$$

DOI: 10.1021/ct5007983
J. Chem. Theory Comput. 2015, 11, 436−450

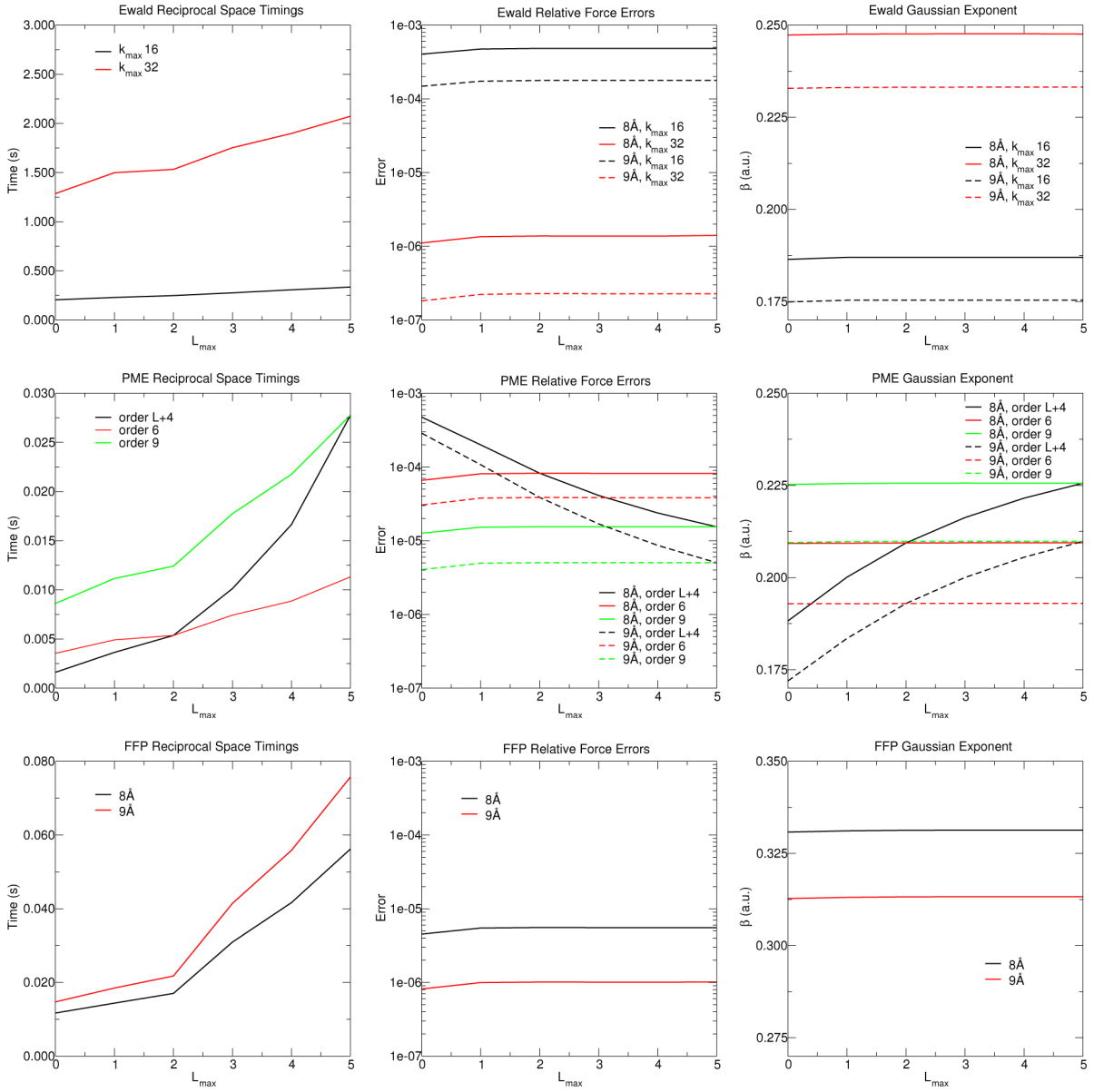


Figure 4. Reciprocal-space evaluation timings (left column), relative force errors (RFEs) (middle column), and the Gaussian exponents (right column) chosen to minimize the RFEs of Ewald (top row), PME (middle row), and FFP (bottom row), using the box of 1024 waters. L_{max} indexes the *ad hoc* 3-site water model being used.

$$q_{\mathrm{H}}^{3\text{-site}} = q_{\mathrm{H}}^{5\text{-site}} + \frac{q_{\mathrm{B}}^{5\text{-site}}}{2} \tag{74}$$

The intrawater interactions are excluded when analyzing the relative force errors with these models.

In order to interpret the results in Figure 4, it is useful to understand how the multipole order influences the magnitude of the forces. For this purpose, Figure 2 decomposes the intermolecular electrostatic forces within the 1024 water box by the percent contributed from the charge−charge interactions, the non-"charge−charge" interactions, and those interactions that involve a particular angular momentum L. Let $\mathbf{F}_L$ be the 3 N vector of atomic forces for a system of waters composed of the *ad hoc* 3-site water model using an order L multipole expansion. The "percent contribution" of the charges, the multipoles, and a particular multipole L to the forces are respectively chosen to be

$$\%F_{L=0} = 100 - \%F_{L>0} \tag{75}$$

$$\%F_{L>0} = 100\sqrt{\frac{|\mathbf{F}_L - \mathbf{F}_0|^2}{|\mathbf{F}_L|^2}} \tag{76}$$

$$\%F_L = 100\sqrt{\frac{|\mathbf{F}_L - \mathbf{F}_{L-1}|^2}{|\mathbf{F}_L|^2}} \tag{77}$$

The relative force errors (RFEs) in Figure 4 are calculated using the expression

$$\mathrm{RFE} = \sqrt{\frac{|\mathbf{F}_{\mathrm{model},L} - \mathbf{F}_{\mathrm{ref},L}|^2}{|\mathbf{F}_{\mathrm{ref},L}|^2}} \tag{78}$$

where $\mathbf{F}_{\mathrm{ref},L}$ is the vector of Ewald forces evaluated with $|k_{\mathrm{max}}| = 128$ and whose real-space corrections include all N^2 minimum image interactions. The Gaussian exponent has been non-

DOI: 10.1021/ct5007983
J. Chem. Theory Comput. 2015, 11, 436−450

linearly optimized to minimize each RFE, and their values are shown in Figure 4 using the convention $\beta = \zeta^{1/2}$.

The timings reported in Figures 3, 4, and 6 were performed on an Intel Xeon E5520 2.27 GHz workstation consisting of a total of 8 cores, and the software was rudimentarily parallelized with OpenMP to make use of all cores.

2.8. Implementation within the mDC Quantum Force Field. The analyses in Figures 5, 6, and 7, and the data shown in Table 1, were produced from condensed-phase calculations using the mDC linear scaling quantum force field.[31] In brief, the quantum mechanical treatment of the entire system is replaced by a series of quantum calculations for each molecule. Although the molecular orbitals of each fragment are not directly coupled, the subsystems remain coupled through the interactions of their electron densities and empirical potentials. Specifically, the mDC method performs DFTB3/3OB semiempirical calculations[72] of each molecule while subjecting them to an effective chemical potential arising from the intermolecular interactions. This idea is conceptually similar to "density functional embedding theory";[73] however, mDC replaces the rigorous evaluation of the embedding potential with computational tractable empirical approximations that are tuned for high accuracy. The standard DFTB3 semiempirical model evaluates the electrostatic interactions from Mulliken charges that are updated until self-consistency is reached. In the mDC formalism described below, the standard DFTB3 treatment continues to be used for those atoms within a common fragment; however, we concoct an auxiliary set of atomic multipoles from the density matrix to improve the intermolecular interactions. The multipolar potentials arising from those interactions enter the fragment Fock matrices as an effective external chemical potential, and the multipoles are updated at each step of the SCF procedure. Let E_A be the DFTB3 energy for molecule A consisting of atoms located at $\mathbf{R}_A$; then, the mDC energy is

$$E_{\mathrm{mDC}} = \sum_{A}^{N_{\mathrm{mol}}} E_A(\mathbf{R}_A, N_A, \mathbf{p}_A) + E_{\mathrm{inter}}(\mathbf{R}, \mathbf{q}) \tag{79}$$

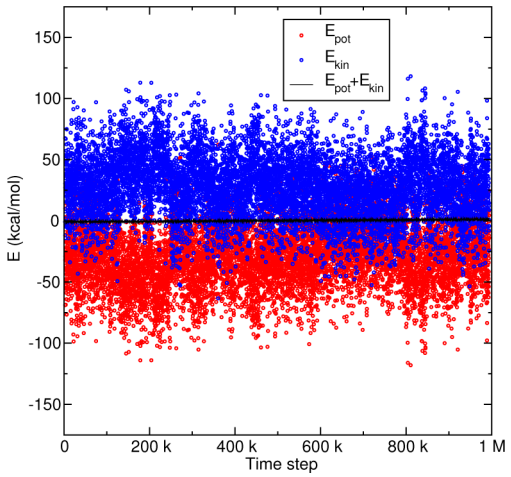


Figure 5. Conservation of the mDC total energy for 1 million steps in a NVE simulation of *N*,*N*-dimethylglycine, using the multipolar PME method with 6th order B-splines.

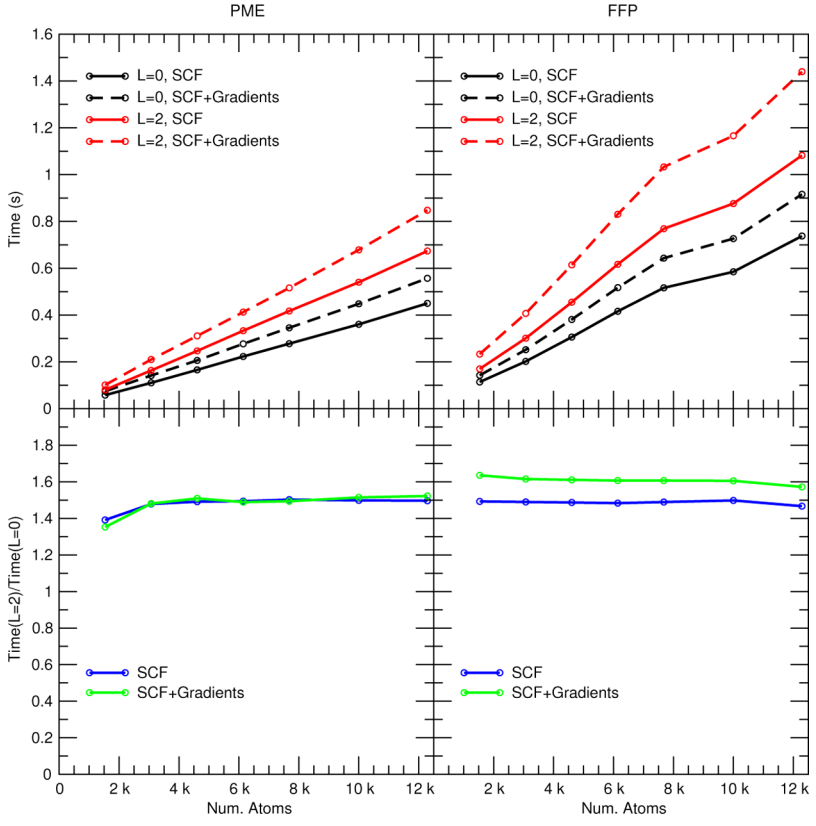


Figure 6. SCF and SCF + gradient timings using the mDC quantum force field for a series of water boxes. "L = 0" indicates that only atomic charges were used to compute electrostatics, whereas "L = 2" indicates that up to quadrupoles were used on the heavy atoms.

DOI: 10.1021/ct5007983
J. Chem. Theory Comput. 2015, 11, 436−450

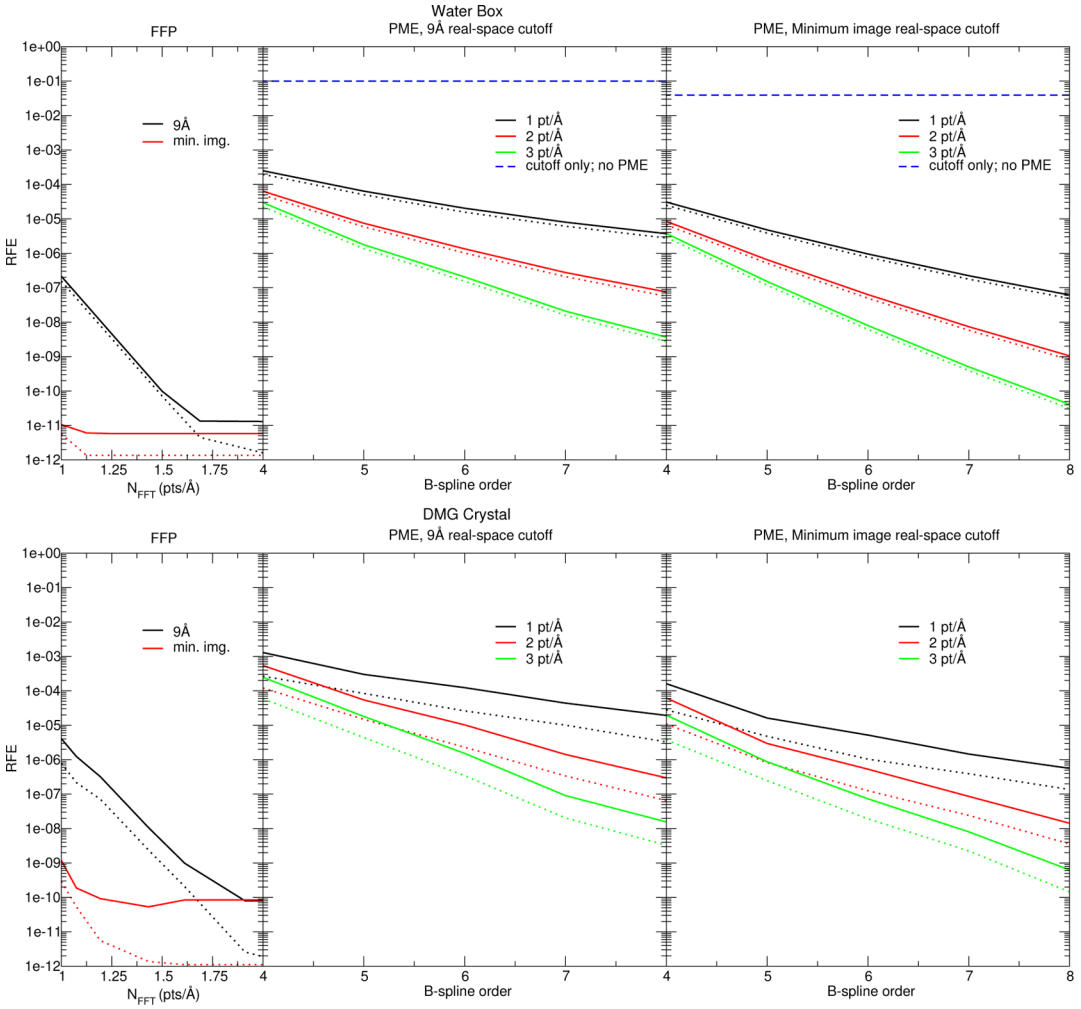


Figure 7. RFEs observed with the mDC linear-scaling quantum force field using various methods of computing the electrostatic interactions. The top portion of the figure shows a box of 1024 waters, whereas the bottom of the figure shows a supercell of *N*,*N*-dimethylglycine containing 4608 atoms. Errors are reported relative to FFP evaluated with 3 points/Å and a minimum image ("min img.") short-range correction. The label "9 Å" indicates the range used for the short-range real-space corrections and, for FFP, the evaluation of the model Gaussian density at the FFT grid points. Dotted lines indicate the ability of the electrostatics protocol to reproduce the electrostatic component of the force, using the atomic multipoles obtained from the reference FFP calculation. The dashed lines are mDC RFEs, using cutoff-based electrostatics without any Ewald treatment.

Table 1. Effect of Electrostatic Cutoff on the Condensed-Phase Properties of Water at 298 K

	density, ρ (g/cm^3)	heat of vaporization, ΔH_{vap} (kcal/mol)
expt[a]	0.9970	10.51
PME[b]	0.9969	10.62
9 Å cutoff	1.0110	10.87

[a]Data taken from refs 86 and 87. [b]Sixth-order B-spline interpolation, grid spacing of 1 point/Å, and short-range cutoff of 9 Å.

where N_A is the number of electrons in fragment A, $q_{a,l\mu}$ is an atomic multipole moment determined from the DFTB3 density matrix, and $p_{a,l\mu} = \partial E_{inter}/\partial q_{a,l\mu}$ is the multipolar potential describing the effective external potential experienced by atom a. With this convention, the σ-spin Fock matrix used to construct the orbitals of molecule A is given as

$$F^{\sigma}_{A,ij} = \frac{\partial E_A}{\partial P^{\sigma}_{A,ij}} + \sum_{\substack{a\in A \\ l\mu\in a}} p_{a,l\mu}\frac{\partial q_{a,l\mu}}{\partial P^{\sigma}_{A,ij}} \quad (80)$$

As explained in ref 31, the atomic charges are computed from a biased Mulliken partition of the density matrix, and the higher-order multipole moments (up to quadrupoles for heavy elements) are computed from the single-center AO−product components of the density matrix. The intermolecular interaction energy E_{inter} consists of Lennard-Jones potentials and the electrostatic interaction of the atomic multipoles. The electrostatic energy and multipolar potentials are evaluated using either the Ewald, PME, or FFP methods described in the previous sections. Because the multipole moments are dependent on the density matrix and the Fock matrix is dependent on the interactions with the other molecules, the mDC energy must be optimized until a mutual convergence is met for all molecules.

Figure 5 illustrates the quality of the mDC forces by demonstrating conservation of the total energy in a NVE simulation of an *N*,*N*-dimethylglycine (DMG) crystal at 225 K. The crystal consists of 288 DMG residues (4608 atoms) constructed from supercell replication of the experimentally determined X-ray structure[74] and was simulated at the experimental density with a locally modified development

DOI: 10.1021/ct5007983
J. Chem. Theory Comput. 2015, 11, 436−450

version of PMEMD[75] that incorporates the mDC method. The simulation was performed using a 0.5 fs time step for 1 million steps, and the electrostatics were computed with sixth-order B-spline PME, a 9 Å real-space cutoff, and a 1 pt/Å FFT grid density.

Figure 6 demonstrates linear scaling of the mDC method for a series of water boxes. In addition, timings were recorded for an *ad hoc* mDC model that limits the intermolecular interactions to monopoles for the sole purpose of determining the relative cost of incorporating atomic quadrupoles. The PME calculations were performed with sixth-order B-splines, a real-space cutoff of 9 Å, and an FFT grid density of 1 point/Å. The FFP calculations were performed with an FFT grid density of 1 point/Å and a cutoff of 9 Å in both the real- and reciprocal-space calculations, that is, the evaluation of the model Gaussian density at the FFT grid points.

Unlike the *ad hoc* water models used in Figure 4, mDC self-consistently models each molecule's polarization in response to its environment (the other molecules). Because the polarization is induced from the environment's electrostatic potential, the electrostatic protocol indirectly introduces atomic force errors by altering the electronic polarization. Within the mDC method, the change in polarization is produced from the difference in the multipolar potential contributions to the Fock matrix. Figure 7 decomposes the mDC force errors to quantify the extent to which the protocols indirectly effect the forces through their difference in polarizations. First, a mDC calculation is performed using a reference electrostatic protocol to produce a SCF converged energy ($E_{\mathrm{mDC,ref}}$) and charge density ($\mathbf{q}_{\mathrm{ref}}$). The dotted lines in Figure 7 are the RFEs computed from the vectors of intermolecular electrostatic forces evaluated with $\mathbf{q}_{\mathrm{ref}}$

$$\mathbf{F}_{\mathrm{ref}} = -\left.\frac{\partial E_{\mathrm{elec,ref}}}{\partial \mathbf{R}}\right|_{\mathbf{q}_{\mathrm{ref}}} \tag{81}$$

$$\mathbf{F}_{\mathrm{model}} = -\left.\frac{\partial E_{\mathrm{elec,model}}}{\partial \mathbf{R}}\right|_{\mathbf{q}_{\mathrm{ref}}} \tag{82}$$

where $E_{\mathrm{elec,model}}$ and $E_{\mathrm{elec,ref}}$ are the model and reference intermolecular electrostatic energies, respectively. A second mDC calculation is then performed using a model electrostatic protocol throughout the SCF procedure to produce a converged energy $E_{\mathrm{mDC,model}}$ and charge density $\mathbf{q}_{\mathrm{model}}$. The solid lines in Figure 7 are the RFEs computed from

$$\mathbf{F}_{\mathrm{ref}} = -\left.\frac{\partial E_{\mathrm{mDC,ref}}}{\partial \mathbf{R}}\right|_{\mathbf{q}_{\mathrm{ref}}} \tag{83}$$

$$\mathbf{F}_{\mathrm{model}} = -\left.\frac{\partial E_{\mathrm{mDC,model}}}{\partial \mathbf{R}}\right|_{\mathbf{q}_{\mathrm{model}}} \tag{84}$$

which are evaluated with different electrostatic protocols *and* charge densities. The dashed lines in Figure 7 are computed from eqs 83 and 84, where $E_{\mathrm{mDC,model}}$ is evaluated using a molecule-based cutoff electrostatic protocol; that is, no treatment for periodicity was applied beyond employing the minimum image convention within the cutoff radius. The electrostatic interactions between the waters were smoothly switched off from 8 Å to 9 Å, based on the O−O separation. All reference calculations were performed with FFP using a 3 point/Å grid density and a full minimum image treatment of the short-range corrections and reciprocal-space evaluation.

The PME and electrostatic cutoff methods are compared in Table 1, which examines how these methods effect the density and heat of vaporization of liquid water at 298 K. These properties were computed from 8-ns simulations of 512 waters in the NPT ensemble using the Monte Carlo barostat,[75] and the Andersen thermostat,[76] as implemented in PMEMD. A time step of 1 fs was used in conjunction with SHAKE[77,78] to constrain the internal structure of the waters to the isolated DFTB3 geometry. The parameters of the mDC water model used to generate these properties are described in Part 2 of this series.[51]

3. RESULTS AND DISCUSSION

3.1. Comparison to Brute Force Replication. Figure 1 establishes the correctness of the Ewald formulas for various multipole orders by comparing its energy to brute force evaluation of eq 1. Analogous plots for FFP and PME with suitably dense FFT grids and B-spline order are indistinguishable from the Ewald results shown in Figure 1. The errors in the energy flatten to a constant in the range of 10^{-13}−10^{-14}, because the large number of calculations involved in the brute force evaluation suffer from a loss of numerical precision. As the multipole order is increased, the energy converges more quickly, because the Coulomb interactions decay by $r^{-l_a-l_b-1}$. When $L \geq 4$, the range of the interactions become so small that the energy is well-approximated by the minimum-image interactions.

3.2. Errors and Timings as a Function of Multipole Order. Before comparing the Ewald, PME, and FFP RFEs, it is useful to briefly discuss the relative magnitude of the forces associated with increasing the multipole order. As shown in Figure 2, the charge−charge interactions represent 40%−50% of the force when higher-order multipoles are used in the *ad hoc* water model. Furthermore, higher-order multipoles progressively contribute less to the force, since their interactions become shorter-ranged. All interactions involving the octupoles within the L_{max} = 3 water model, for example, contribute <5% to the force, whereas all interactions involving L = 5 contribute <0.5%.

The wall-clock time required to evaluate the real- and reciprocal-space contributions to the multipolar potentials, as a function of multipole expansion order, are displayed in Figures 3 and 4, respectively. The real-space corrections evaluated from eqs 29 and 30 and the algorithm in ref 56 are particularly efficient for $L_{\mathrm{max}} \leq 3$. For higher-order expansions, techniques based on rotations into an internal-coordinate system can be used to further improve performance;[52,79] however, considering the short-range and relative insignificance of these interactions, it can be verily questioned if practical models for condensed-phase simulations will consider their additional cost worthwhile. Including quadrupoles on the oxygen causes the real-space evaluation timings to slow down by a factor of 1.8, relative to a charge-only model. Comparison of the reciprocal-space timings in Figure 4 suggests that, for this system of 1024 waters, FFP and Ewald are ∼2 and ∼100 times slower than PME, respectively, for comparable error levels. Furthermore, the evaluation of the real-space corrections at L_{max} = 2 with a cutoff of 9 Å are approximately twice as slow as the sixth-order B-spline PME reciprocal-space evaluation times.

The optimized Gaussian exponents shown in Figure 4 are far more sensitive to the choice of real-space cutoff than they are to

DOI: 10.1021/ct5007983
J. Chem. Theory Comput. 2015, 11, 436−450

multipole order. As the real-space cutoff is increased, the Gaussian exponent decreases to become better resolved in the plane wave basis, and this consequently decreases the errors of the reciprocal-space forces. Similarly, increasing the PME B-spline order affords the opportunity to increase the Gaussian exponent to reduce the errors in the real-space correction without sacrificing accuracy in the reciprocal-space potential.

3.3. Comparisons within a Linear-Scaling Quantum Force Field Framework. Much of this manuscript uses the word "error" to describe a difference between an accurate electrostatics protocol and a more approximate treatment. However, this phrasing should not be misconstrued to incorrectly suggest that the PME gradients are inconsistent with the PME energy, for example. To emphasize this, Figure 5 demonstrates conservation of total energy in a NVE simulation of DMG for 1 million time steps evaluated with the mDC quantum force field, whose electrostatics are computed with PME. If there were an inconsistency between the energy and forces, then we would observe a significant drift in the energy as time is propagated; however, we do not observe a drift.

The mDC method is linear scaling when using either PME or FFP, as illustrated in Figure 6. When evaluated with PME, a box of 4096 waters can be SCF converged within <0.7 s. Furthermore, our timings indicate that the use of quadrupoles slows the mDC calculation by a factor of 1.5, relative to a charge-only model. This is slightly less than the ratio observed in Figure 3, because the mDC and charge-only mDC evaluations share common operations that are independent of the multipole order used in their intermolecular electrostatics (for example, their Fock matrix diagonalizations). The number of SCF cycles required to reach convergence is not particularly sensitive to the size of the system, which can be inferred from the linear-scaling shown in Figure 6. The SCF convergence rate is largely determined by the quality of the initial guess orbitals. In practice, the difference between MD steps is so small that convergence can typically be reached within 4−6 cycles.

Ewald, PME, and FFP are approximate, but their accuracy can be systematically improved by adjusting the number of plane waves, the Gaussian exponent, and the real-space correction cutoff. In the analysis of Figure 4, we examine how these parameters alter the electrostatic energy protocol by quantifying their effect on atomic forces (for a fixed charge density), $-\partial E/\partial \mathbf{R}|_{\mathbf{q}}$, while ignoring their effect on electrostatic potentials $\partial E/\partial \mathbf{q}|_{\mathbf{R}}$. However, the parameters do effect the electrostatic potentials. Furthermore, the electrostatic potentials enter the Fock matrix via eq 80. Therefore, the parameters indirectly effect the charge density through the propagation of electrostatic potential "errors" within the SCF, which cause its convergence to a different density matrix. In other words, not only will *ab initio* methods (or any polarizable model) suffer from the "primary error" associated with the partial derivative $-\partial E/\partial \mathbf{R}|_{\mathbf{q}}$, they also incur a "secondary error" when the model and reference forces are evaluated about different charge densities. The dotted lines in Figure 7 are the primary electrostatic errors; that is, they are the RFEs associated with the electrostatic forces evaluated from the atomic multipoles generated from the *reference* calculation. The solid lines in Figure 7 contain both the primary and secondary errors produced when the reference and approximate electrostatic protocols are used throughout the SCF procedures to yield two slightly different densities.

The reader may first notice that the primary FFP errors decrease and converge, with respect to FFT grid density quickly; however, the secondary errors trail off at 10^{-11} to 10^{-10}. This limit is ultimately caused by our SCF numerical convergence criteria. The errors in Figure 7 were generated with a 10^{-12} tolerance on the maximum value appearing in the "error matrix", that is, the commutator between the Fock matrix and the density matrix.

As expected, one observes a decrease in error as the PME B-spline order is increased, the FFT grid density is increased, and as the real-space cutoff is increased. The mDC RFEs observed in water and DMG parallel the primary errors; however, the DMG secondary errors are significantly larger than those observed in water. The reason for this is because a DMG molecule contains 16 atoms, whereas water contains 3. By having more AOs, the DMG molecules are more likely to converge to a slightly different density matrix.

Before PME was popularized in MD simulations,[37,38] electrostatic cutoffs were often used, and many of the original water models were thus parametrized using a cutoff protocol. It was later found that application of those water models with PME caused significant changes in the thermodynamic properties of liquid water,[80−82] and it has been suggested that new water models not be parametrized using cutoffs,[81] because their application within simulations of biomolecules often necessitates the use of an Ewald treatment.[83,84] Nonetheless, some recently developed water models have continued to be parametrized using cutoffs.[85] Therefore, we are obliged to investigate by how much this may effect condensed-phase properties. First, in Figure 7, we note that a cutoff of 9 Å produces an RFE value of 0.1, in comparison to the RFE value observed with a sixth-order B-spline PME with a grid spacing of 1 point/Å FFT (2×10^{-5}). Furthermore, even if all N^2 minimum image interactions were computed without PME, the RFE is still 0.04. Finally, we performed MD simulations using a parametrized mDC water model to obtain an equilibrated density and heat of vaporization, as shown in Table 1, and we reperformed the simulations using switched-cutoff electrostatics to make comparison. In agreement with previous works,[80,81] we observe that cut-off electrostatics cause the density of water to increase, and because the interactions in this range are attractive overall, this also causes the heat of vaporization to increase.

4. CONCLUSION

This work presented extensions of the Ewald, Particle Mesh Ewald (PME), and Fast Fourier−Poisson (FFP) methodologies to systems composed of point multipole expansions to arbitrary order by making use of the spherical tensor gradient operator. The timings and errors inherent to these methods were compared using *ad hoc* water models and with a parametrized water based on the modified divide-and-conquer (mDC) linear-scaling quantum force field. These comparisons lead us to conclude that (i) the FFP method is approximately twice as slow as the PME method at comparable error levels; (ii) the inclusion of quadrupoles in the linear-scaling force field slow the calculations by 1.5, relative to a charge-only model; and (iii) with the exception of the Ewald method, the real-space corrections are more expensive than the reciprocal-space calculations for typical cutoff values. Furthermore, our results suggest that the evaluation of multipolar electrostatics involving orders greater than 3 could likely be computed to an acceptable error using an electrostatic cutoff, because of their overall short-range and relative insignificance, in comparison to lower-order interactions.

DOI: 10.1021/ct5007983
J. Chem. Theory Comput. 2015, 11, 436−450

The relative force errors exhibited within mDC were decomposed into primary and secondary errors, where the primary errors directly result from the approximations within the PME or FFP algorithms for a given density, and the secondary errors are the propagation of the model's electrostatic potential within the self-consistent field (SCF) procedure, resulting in a different converged density matrix. It is found that the force errors closely follow the primary errors, and the magnitude of the secondary errors is related to the number of atomic orbitals (AOs). Nevertheless, the presence of these "errors" does not imply that the mDC forces are inconsistent with its energy, which was demonstrated with an NVE simulation that was devoid of an energy drift. Instead, these differences merely reflect how similar an electrostatic protocol is to another reference protocol.

Finally, the importance of using an Ewald treatment in simulations, as opposed to using electrostatic cutoffs, was emphasized by comparing the density and heat of vaporization of water. The electrostatic cut-off method was found to artificially increase the density and heat of vaporization of water.

AUTHOR INFORMATION

Corresponding Author

*E-mail: york@biomaps.rutgers.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors are grateful for financial support provided by the National Institutes of Health (No. GM62248). Computational resources from the Minnesota Supercomputing Institute for Advanced Computational Research (MSI) were utilized in this work. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation (Grant No. OCI-1053575).

REFERENCES

(1) York, D. M.; Darden, T.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *99*, 8345−8348.

(2) York, D. M.; Yang, W.; Lee, H.; Darden, T.; Pedersen, L. G. *J. Am. Chem. Soc.* **1995**, *117*, 5001−5002.

(3) Cheatham, T. E., III; Young, M. A. *Biopolymers* **2000**, *56*, 232−256.

(4) Cisneros, G. A.; Karttunen, M.; Ren, P.; Sagui, C. *Chem. Rev.* **2014**, *114*, 779−814.

(5) Cisneros, G. A.; Babin, V.; Sagui, C. In *Biomolecular Simulations: Methods and Protocols*; Monticelli, L., Salonen, E., Eds.; Methods in Molecular Biology, Vol. *924*; Springer Science + Business Media: New York, 2013; pp 243−270.

(6) Cardamone, S.; Hughes, T. J.; Popelier, P. L. A. *Phys. Chem. Chem. Phys.* **2014**, *16*, 10367−10387.

(7) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A., Jr.; Head−Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. *J. Phys. Chem. B* **2010**, *114*, 2549−2564.

(8) Shi, Y.; Wu, C.; Ponder, J. W.; Ren, P. *J. Comput. Chem.* **2011**, *32*, 967−977.

(9) Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2004**, *108*, 13427−13437.

(10) Grossfield, A.; Ren, P.; Ponder, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 15671−15682.

(11) Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. *J. Chem. Theory Comput.* **2013**, *9*, 4046−4063.

(12) Semrouni, D.; Cramer, C. J.; Gagliardi, L. *Theor. Chem. Acc.* **2014**, *134*, 1590.

(13) Chaudret, R.; Gresh, N.; Narth, C.; Lagardere, L.; Darden, T.; Cisneros, G. A.; Piquemal, J.-P. *J. Phys. Chem. A* **2014**, *118*, 7598−7612.

(14) Duke, R. E.; Starovoytov, O. N.; Piquemal, J.-P.; Cisneros, G. A. *J. Chem. Theory Comput.* **2014**, *10*, 1361−1365.

(15) Cisneros, G. A. *J. Chem. Theory Comput.* **2012**, *8*, 5072−5080.

(16) Elking, D. M.; Cisneros, G. A.; Piquemal, J.-P.; Darden, T. A.; Pedersen, L. G. *J. Chem. Theory Comput.* **2010**, *6*, 190−202.

(17) Chaudret, R.; Gresh, N.; Narth, C.; Lagardère, L.; Darden, T. A.; Cisneros, G. A.; Piquemal, J.-P. *J. Phys. Chem. A* **2014**, *118*, 7598−7612.

(18) Mas, E. M.; Bukowski, R.; Szalewicz, K.; Groenenboom, G. C.; Wormer, P. E. S.; van der Avoird, A. *J. Chem. Phys.* **2000**, *113*, 6687−6701.

(19) Medders, G. R.; Babin, V.; Paesani, F. *J. Chem. Theory Comput.* **2014**, *10*, 2906−2910.

(20) Piquemal, J.; Chevreau, H.; Gresh, N. *J. Chem. Theory Comput.* **2007**, *3*, 824−837.

(21) Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J.-P. *J. Chem. Theory Comput.* **2007**, *3*, 1960−1986.

(22) Nam, K.; Gao, J.; York, D. M. *J. Chem. Theory Comput.* **2005**, *1*, 2−13.

(23) Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. *J. Phys. Chem. A* **2001**, *105*, 293−307.

(24) Gao, J.; Truhlar, D. G.; Wang, Y.; Mazack, M. J. M.; Löffler, P.; Provorse, M. R.; Rehak, P. *Acc. Chem. Res.* **2014**, *47*, 2837−2845.

(25) Wang, Y.; Sosa, C. P.; Cembran, A.; Truhlar, D. G.; Gao, J. *J. Phys. Chem. B* **2012**, *116*, 6781−6788.

(26) Song, L.; Han, J.; lin Lin, Y.; Xie, W.; Gao, J. *J. Phys. Chem. A* **2009**, *113*, 11656−11664.

(27) Xie, W.; Orozco, M.; Truhlar, D. G.; Gao, J. *J. Chem. Theory Comput.* **2009**, *5*, 459−467.

(28) Xie, W.; Gao, J. *J. Chem. Theory. Comput.* **2007**, *3*, 1890−1900.

(29) Zhang, P.; Truhlar, D. G.; Gao, J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7821−7829.

(30) Giese, T. J.; Chen, H.; Dissanayake, T.; Giambaşsu, G. M.; Heldenbrand, H.; Huang, M.; Kuechler, E. R.; Lee, T.-S.; Panteva, M. T.; Radak, B. K.; York, D. M. *J. Chem. Theory Comput.* **2013**, *9*, 1417−1427.

(31) Giese, T. J.; Chen, H.; Huang, M.; York, D. M. *J. Chem. Theory Comput.* **2014**, *10*, 1086−1098.

(32) Giese, T. J.; Huang, M.; Chen, H.; York, D. M. *Acc. Chem. Res.* **2014**, *47*, 2812−2820.

(33) Handley, C. M.; Hawe, G. I.; Kell, D. B.; Popelier, P. L. A. *Phys. Chem. Chem. Phys.* **2009**, *11*, 6365−6376.

(34) Kramer, C.; Gedeck, P.; Meuwly, M. *J. Chem. Theory Comput.* **2013**, *9*, 1499−1511.

(35) Devereux, M.; Raghunathan, S.; Fedorov, D. G.; Meuwly, M. *J. Chem. Theory Comput.* **2014**, *10*, 4229−4241.

(36) Ewald, P. P. *Ann. Phys.* **1921**, *64*, 253−268.

(37) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(38) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Hsing, L.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

(39) Cerutti, D. S.; Duke, R. E.; Darden, T. A.; Lybrand, T. P. *J. Chem. Theory Comput.* **2009**, *5*, 2322−2338.

(40) Cerutti, D. S.; Case, D. A. *J. Chem. Theory Comput.* **2010**, *6*, 443−458.

(41) York, D. M.; Yang, W. *J. Chem. Phys.* **1994**, *101*, 3298−3300.

(42) Toukmaji, A.; Sagui, C.; Board, J.; Darden, T. *J. Chem. Phys.* **2000**, *113*, 10913−10927.

(43) Nymand, T. M.; Linse, P. *J. Chem. Phys.* **2000**, *112*, 6152−6160.

(44) Sagui, C.; Pedersen, L. G.; Darden, T. A. *J. Chem. Phys.* **2004**, *120*, 73−87.

(45) Schnieders, M. J.; Fenn, T. D.; Pande, V. S.; Brunger, A. T. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2009**, *D65*, 952−965.

(46) Cisneros, G. A.; Piquemal, J.; Darden, T. A. *J. Chem. Phys.* **2006**, *125*, 184101.

DOI: 10.1021/ct5007983
J. Chem. Theory Comput. 2015, 11, 436−450

(47) Darden, T. A. Extensions of the Ewald method for Coulomb Interactions in crystals. In *International Tables for Crystallography, Reciprocal Space*, 3rd Edition; Springer: Dordrecht, The Netherlands, 2008; Vol. *B*, pp 458−483.
(48) Aguado, A.; Madden, P. A. *J. Chem. Phys.* **2003**, *119*, 7471−7483.
(49) Laino, T.; Hutter, J. *J. Chem. Phys.* **2008**, *129*, 074102.
(50) Zielinski, F.; Popelier, P. L. A. *J. Mol. Model.* **2014**, *20*, 2256.
(51) Giese, T. J.; Panteva, M. T.; Chen, H.; York, D. M. *J. Chem. Theory Comput.* DOI: 10.1021/ct500799g.
(52) Simmonett, A. C.; Pickard, F. C., IV; Schaefer, H. F., III; Brooks, B. R. *J. Chem. Phys.* **2014**, *140*, 184101.
(53) Equation 1 can be equivalently expressed with **n** appearing in the denominator instead of the numerator by transforming the integration variable with the substitution $\mathbf{u} = \mathbf{r}' + \mathbf{n}$, so that $\mathbf{r}' = \mathbf{u} - \mathbf{n}$.
(54) Giese, T. J.; York, D. M. *J. Comput. Chem.* **2008**, *29*, 1895−1904.
(55) Giese, T. J.; York, D. M. *J. Chem. Phys.* **2011**, *134*, 194103.
(56) Giese, T. J.; York, D. M. *J. Chem. Phys.* **2008**, *128*, 064104.
(57) Watson, M. A.; Sałek, P.; Macak, P.; Helgaker, T. *J. Chem. Phys.* **2004**, *121*, 2915−2931.
(58) Pérez−Jordá, J. M.; Yang, W. *J. Chem. Phys.* **1996**, *104*, 8003−8006.
(59) Pérez-Jordá, J. M.; Yang, W. *J. Chem. Phys.* **1997**, *107*, 1218−1226.
(60) Hobson, E. W. *Proc. London Math. Soc.* **1892**, *24*, 55−67.
(61) Giese, T. J.; York, D. M. In *Many-Body Effects and Electrostatics in Biomolecules*; Cui, Q., Meuwly, M., Ren, P., Eds.; Pan Stanford Publishing: Singapore, in press.
(62) Weniger, E. J. *Collect. Czech. Chem. Commun.* **2005**, *70*, 1225−1271.
(63) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; John Wiley & Sons: Hoboken, NJ, 2000.
(64) Schoenberg, I. J. *J. Approx. Theory* **1983**, *39*, 324−337.
(65) Schoenberg, I. J. Euler's contribution to cardinal spline interpolation: The exponential Euler splines. In *Leonhard Euler 1707−1783*; Birkhäuser Verlag: Basel/Boston/Stuttgart, 1983; pp 199−214.
(66) Schoenberg, I. J. Cardinal interpolation and spline functions IV. The exponential Euler splines. In *Linear Operators and Approximation*; Birkhäuser Verlag: Basel, Switzerland, Boston, MA, Stuttgart, Germany, 1972; pp 382−404.
(67) Ishida, K. *J. Chem. Phys.* **1998**, *109*, 881−890.
(68) Bogusz, S.; Cheatham, T. E., III; Brooks, B. R. *J. Chem. Phys.* **1998**, *108*, 7070−7084.
(69) Smith, E. R. *Proc. R. Soc. London, Ser. A* **1981**, *375*, 475−505.
(70) De Leeuw, S. W.; Perram, J. W.; Smith, E. R. *Proc. R. Soc. London, Ser. A* **1983**, *388*, 177−193.
(71) Brooks, B. R.; Brooks, C. L., III; MacKerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(72) Gaus, M.; Goez, A.; Elstner, M. *J. Chem. Theory Comput.* **2013**, *9*, 338−354.
(73) Libisch, F.; Huang, C.; Carter, E. A. *Acc. Chem. Res.* **2014**, *47*, 2768−2775.
(74) Kapustin, E. A.; Minkov, V. S.; Stare, J.; Boldyreva, E. V. *Cryst. Growth Des.* **2014**, *14*, 1851−1864.
(75) Case, D. A.; Babin, V.; Berryman, J. T.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham, III, T. E.; Darden, T. A.; Duke, R. E.; Gohlke, H.; Goetz, A. W.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolossváry, I.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Luchko, T.; Luo, R.; Madej, B.; Merz, K. M.; Paesani, F.; Roe, D. R.; Roitberg, A.; Sagui, C.; Salomon-Ferrer, R.; Seabra, G.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Kollman, P. A. *AMBER 14*; University of California: San Francisco, CA, 2014.
(76) Andersen, H. C. *J. Chem. Phys.* **1980**, *72*, 2384−2393.
(77) Miyamoto, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 952−962.
(78) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(79) Hättig, C. *Chem. Phys. Lett.* **1997**, *268*, 521−530.
(80) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. *J. Chem. Phys.* **2004**, *120*, 9665−9678.
(81) Rick, S. W. *J. Chem. Phys.* **2004**, *120*, 6085−6093.
(82) van der Spoel, D.; van Maaren, P. J. *J. Chem. Theory Comput.* **2006**, *2*, 1−11.
(83) York, D. M.; Wlodawer, A.; Pedersen, L. G.; Darden, T. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 8715−8718.
(84) Cheatham, T. E., III; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4193−4194.
(85) Han, J.; Mazack, M. J. M.; Zhang, P.; Truhlar, D. G.; Gao, J. *J. Chem. Phys.* **2013**, *139*, 054503.
(86) Kell, G. S. *J. Chem. Eng. Data* **1975**, *20*, 97−105.
(87) Wagner, W.; Pruß, A. *J. Phys. Chem. Ref. Data* **2002**, *31*, 387−535.

NOTE ADDED AFTER ASAP PUBLICATION

This article was published ASAP on January 14, 2015. A sentence in section 2.2 has been modified. The correct version was published on January 15, 2015.

DOI: 10.1021/ct5007983
J. Chem. Theory Comput. 2015, 11, 436−450