Extension of adaptive tree code and fast multipole methods to high angular momentum particle charge densities

Timothy J. Giese and **Darrin M. York**[a]
Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55455, USA

Abstract

The development and implementation of a tree code (TC) and fast multipole method (FMM) for the efficient, linear-scaling calculation of long-range electrostatic interactions of particle distributions with variable shape and multipole character are described. The target application of these methods are stochastic boundary molecular simulations with polarizable force fields and/or combined quantum mechanical/molecular mechanical potentials. Linear-scaling is accomplished through the adaptive decomposition of the system into a hierarchy of interacting particle sets. Two methods for effecting this decomposition are evaluated: fluc-splitting and box-splitting, for which the latter is demonstrated to be generally more accurate. In addition, a generalized termination criterion is developed that delivers optimal performance at fixed error tolerance that, in the case of quadrupole-represented Drude water, effects a speed-up by a factor of 2–3 relative to a multipole-independent termination criteria. The FMM is shown to be approximately 2–3 times faster than the TC, independent of the system size and multipole order of the particles. The TC and FMM are tested for a variety of static and polarizable water systems, and for the the $70S$ ribosome functional complex containing an assembly of transfer and messenger RNAs.

I. INTRODUCTION

Computational chemistry and biology techniques continue to improve in their ability to tackle complex chemical phenomena. One area where this is particularly well illustrated is in the development and application of so-called multiscale quantum models to study chemical reactivity of biological molecules. Simulations of biological reactivity using multiscale models requires the integration of a hierarchy of methods working synchronously to assemble a network of interactions that reasonably mimic a realistic biological environment. These simulations can be considerably costly, especially when highly charged systems such as DNA and RNA are involved and the chemical events are intimately coupled with large-scale conformations changes. Despite the graduation in complexity of the methods contributing to the multiscale quantum models, one of the main computational bottlenecks for these simulations still derives from the evaluations of long-range electrostatic interactions as with conventional molecular mechanical force field simulations.[1−3] However, unlike the traditional static point charge models and fairly uniform distribution of atoms in force field simulations, the charge distributions in multiscale quantum models involve added levels of complexity. A combined quantum mechanical/molecular mechanical model of an enzyme or ribozyme active site may require a fairly sophisticated representation of the quantum charge distribution based on the single-particle density-matrix.[4,5] Some of the developing polarizable force fields[6,7] may additionally require atom or off-center sites to have higher-order multipole character.[8−14] Finally, boundary element methods used to mimic the electrostatic and linear-response effects of the medium require a set of surface points or Gaussians that encapsulate the macromolecule.

aAuthor to whom correspondence should be addressed. Electronic mail: E-mail: york@chem.umn.edu.

[15–18] Together, the assembly of these heterogeneous, non-uniform charge distributions present a challenge for the design of efficient, robust linear-scaling electrostatic algorithms.

In this paper we develop, implement and compare two different adaptive linear-scaling electrostatic methods: a tree code (TC) and a fast-multipole method (FMM). These methods utilize a non-grid based recursive bisection method that adaptively allows subdivision of multipolar particles into localized sets used to evaluate direct and multipolar interactions for a broad range of non-uniform particle distributions such as those encountered in multiscale quantum simulations. The main features that comprise the methods have been for the most part described in other seminal papers on linear-scaling electrostatic methods in the literature.[19–38] The accuracy and efficiency of the present implementations are tested and characterized for a variety of systems. Particular detail is spent on achieving an optimum balance between accuracy and efficiency as a function of multipole expansion order, the so-called "well-separateness" criteria, and adaptive termination condition. In addition, two methods for adaptive particle set splitting procedures are examined and compared. The methods are demonstrated to be considerably robust, and for non-periodic systems, provide an attractive alternative to either direct $O(N^2)$ calculation or more approximate spatial decomposition methods that employ electrostatic cutoffs.

The outline of the paper is as follows. Sec. II describes the FMM and TC methods and our implementation, including the description of 2 different adaptive splitting conditions and termination conditions. Sec. III provides FMM and TC parameters corresponding to several levels of accuracy, which are examined as a function of system shape, size, and particle composition. The adaptive splitting procedures are compared and deficiencies are identified. Finally, the use of a generalized termination condition developed in Sec. II is examined. The paper concludes in Sec. IV with a summary of the key results.

II. METHODS

Sec. IIA provides the mathematical description of the FMM energy and derivatives, and hierarchical implementation Sec. IIB discusses the relationship between FMM and tree-code methods and how this effects the implementation. The well-separatedness criteria, adaptive splitting procedures, and termination criteria are described in Sec. IIC. Sec. IID describes the set up and simulation details for the test systems discussed later in the paper.

A. FMM energy

The Coulomb energy of a system, J, can generally be written

$$J = \frac{1}{2} \int \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}\prime)}{|\mathbf{r}-\mathbf{r}\prime|} d^3r \, d^3r\prime = \frac{1}{2} \int \rho(\mathbf{r})\phi(\mathbf{r}) d^3r,$$

(1)

where $\rho(\mathbf{r})$ is the charge density of the system and $\phi(\mathbf{r})$ is its electrostatic potential. Suppose that $\rho(\mathbf{r})$ is composed of a set of N sites (e.g., atoms), with the charge density of each site represented by a set of basis functions $\varphi_{i,(l,m)}(\mathbf{r})$, i.e.,

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} \rho_i(\mathbf{r} - \mathbf{R}_i) = \sum_{i=1}^{N} \sum_{l=0}^{l_{max,i}} \sum_{m=-l}^{l} q_{i,(l,m)} \varphi_{i,(l,m)}(\mathbf{r} - \mathbf{R}_i),$$

(2)

where i indexes sites, $\mathbf{R}_i$ is the position of the ith site, (l, m) indexes the angular momentum of the basis function, and $q_{i,(l,m)}$ is the corresponding multipole moment of the site, i.e.,

$$q_{i,(l,m)} = \int \rho_i(\mathbf{r} - \mathbf{R}_i) R_{l,m}(\mathbf{r} - \mathbf{R}_i) d^3 r,$$

(3)

where $R_{l,m}(\mathbf{r})$ is a scaled real regular solid harmonic.[22,39,40] Eq. (1) can then be written

$$J = \frac{1}{2} \sum_{i=1}^{N} \mathbf{q}_i^T \cdot \mathbf{p}_i,$$

(4)

where q_i is the vector of multipole moments for site i, and p_i is a corresponding set of "multipolar potentials", i.e., the $(l, m)^{th}$ element of the vector p_i is

$$
\begin{aligned}
p_{i,(l,m)} &= \frac{dJ}{dq_{i,(l,m)}} \\
&= \sum_{j=1}^{N} \sum_{l_j=0}^{l_{max,j}} \sum_{m_j=-l_j}^{l_j} q_{j,(l_j,m_j)} \int\int \frac{\varphi_{i,(l,m)}(\mathbf{r} - \mathbf{R}_i) \varphi_{j,(l_j,m_j)}(\mathbf{r}\prime - \mathbf{R}_j)}{|\mathbf{r} - \mathbf{r}\prime|} d^3 r d^3 r\prime .
\end{aligned}
$$

(5)

. The term *multipolar potential* is sometimes referred to in the literature as a Taylor[29] expansion of the electrostatic potential.

FMMs achieve linear scaling by partitioning a system into a hierarchy of sets, representing the charge distribution of each set by a single multipole expansion, and replacing explicit site-site interactions between *well-separated* sets by a single multipole-multipole interaction.[1,33] In this context, the Coulomb energy is written

$$J = \frac{1}{2} \sum_{a} \sum_{i \in a} \mathbf{q}_{i_a}^T \cdot \mathbf{p}_{i_a}^{fmm} + \frac{1}{2} \sum_{a} \sum_{i \in a} \mathbf{q}_{i_a}^T \cdot \mathbf{p}_{i_a}^{dir},$$

(6)

where a indexes a set of sites, $\Sigma_{i \in a}$ sums over all sites within set a, i_a is the ith site in set a, and the multipolar potential has been separated into contributions arising from well-separated multipole-multipole interactions, $\mathbf{p}_{i_a}^{fmm}$, and explicit site-site interactions between non-well-separated sets, $\mathbf{p}_{i_a}^{dir}$.

In order to provide expressions for $\mathbf{p}_{i_a}^{dir}$ and $\mathbf{p}_{i_a}^{fmm}$ in a manner that includes concepts of hierarchy and well-separateness, it is notationally convenient to introduce the functions ws(ab) and dir(ab):

$$\text{ws}(ab) = \begin{cases} 1, & \text{if sets } a \text{ and } b \text{ are well-separated and do not interact at a superset-level} \\ 0, & otherwise \end{cases}$$

(7)

$$\text{dir}(ab) = \begin{cases} 1, & \text{if sets } a \text{ and } b \text{ are not well-separated and do not have subsets of particles} \\ 0, & otherwise. \end{cases}$$

(8)

Using Eq. (8), the expression for $\mathbf{p}_{i_a}^{dir}$ becomes analogous to Eq. (5)

$$p^{dir}_{i_a,(l,m)}=\sum_{b}\text{dir}(ab)\sum_{j\in b}\sum_{l_j=0}^{l_{max,j}}\sum_{m_j=-l_j}^{l_j}q_{jb,(l_j,m_j)}$$
$$\times\int\int\frac{\varphi_{i_a,(l,m)}(\mathbf{r}-\mathbf{R}_{i_a})\varphi_{jb,(l_j,m_j)}(\mathbf{r}\prime-\mathbf{R}_{jb})}{|\mathbf{r}-\mathbf{r}\prime|}d^3rd^3r\prime,\tag{9}$$

, i.e., $\mathbf{p}^{dir}_{i_a,(l,m)}$ is the $(l,m)^{th}$ element of the vector $\mathbf{p}^{dir}_{i_a}$. Similarly, $\mathbf{p}^{fmm}_{i_a}$ is expressed with Eq. (10)–Eq. (12)

$$\mathbf{p}^{fmm}_{i_a}=\mathbf{W}(\mathbf{R}_{i_aa})^T\cdot\mathbf{V}_a,\tag{10}$$

$$\mathbf{v}_a=\sum_{b}\text{ws}(ab)\mathbf{T}(\mathbf{R}_{ab})^T\cdot\mathbf{m}_b,\tag{11}$$

$$\mathbf{m}_b=\sum_{j\in b}\mathbf{W}(\mathbf{R}_{jbb})\cdot\mathbf{q}_{jb},\tag{12}$$

where $\mathbf{m}_b$ is the multipole moment of set b, evaluated about its center $\mathbf{R}_b$, $\mathbf{v}_a$ is the electrostatic multipolar potential of the sites in set b, evaluated at multipole expansion center $\mathbf{R}_a$, $\mathbf{R}_{jbb}=\mathbf{R}_{jb}-\mathbf{R}_b$ is the vector between the jth site in set b with set b's expansion center, $\mathbf{R}_{ab}=\mathbf{R}_a-\mathbf{R}_b$ is the vector between the expansion centers of sets a and b, and $\mathbf{W}(\text{r})$ and $\mathbf{T}(\text{r})$ are the real translation and interaction matrices, respectively. Expressions for $\mathbf{W}(\text{r})$ and $\mathbf{T}(\text{r})$, in terms of regular and irregular real solid harmonics, have been reported previously.[22,39] Note that the translation matrix used in this work $[\mathbf{W}(\mathbf{R}_{ab})]$ differs from that used in Ref. 22 $[\mathbf{W}^{WH}(\mathbf{R}_{ab})]$ by the sign of the argument [i.e., $\mathbf{W}(\mathbf{R}_{ab})=\mathbf{W}^{WH}(\mathbf{R}_{ba})$]. Expressions for all quantities are provided within the supporting information.

The derivative of the Coulomb energy with respect to the α-Cartesian coordinate [$\alpha\in(x,y,z)$] of site i_a is

$$\frac{dJ}{d\alpha}=\mathbf{q}^T_{i_a}\cdot\left(\frac{d}{d\alpha}\mathbf{p}_{i_a}\right),\tag{13}$$

where we have assumed that the multipole moments are static, and the remaining derivative is

$$\frac{d}{d\alpha}\mathbf{p}_{i_a}=\left(\frac{d}{d\alpha}\mathbf{W}(\mathbf{R}_{i_aa})\right)^T\cdot\mathbf{v}_a+\frac{d}{d\alpha}\mathbf{p}^{dir}_{i_a}.\tag{14}$$

Formally, if the expansion center $\mathbf{R}_a$ depends on $\mathbf{R}_{i_a}$, then the derivative of $\mathbf{T}(\mathbf{R}_{ab})$ appears from the chain-rule term associated with $\mathbf{v}_a$, and therefore Eq. (14) is only approximate; however, FMMs are inherently approximate and additional chain-rule terms contribute negligibly to the derivative for the systems examined here.

Finally, we note that the charge densities considered in this work are point charges or point multipoles, in which case, the integral appearing in Eq. (9) is an element of the interaction matrix, i.e., $T_{(l,m),(l_j,m_j)}(\mathrm{R}_{ij})$.

What follows is a summary of the hierarchical implementation of our method. Note that we have provided a detailed description of our implementation, including pseudo-source code, within the supporting information of this document for those readers interested in programming our adaptive FMM. (1) Recursively bisect the system into smaller sets to form a binary tree hierarchy and each site can be associated to 1 "childless" node. The top of the tree is a single node containing the entire system. (2) The binary tree is traversed and the multipole moments of each node is computed. If the node has no children, then the moments are constructed from the atoms, or are otherwise constructed from the moments of its 2 children. (3) The binary tree is traversed, starting at the top, and all "well-separated" sets interact via multipole expansions. If the bottom of the tree is reached, the interaction is performed directly. (4) The binary tree is traversed, starting at the top, and the multipole potentials of each node is translated and added to each of its children. (5) The energy is evaluated using Eq. (6).

B. Relationship between FMMs and Tree Codes

The Coulomb energy in a Tree Code is similar to the FMM expression [Eq. (6)] and is written

$$J = \frac{1}{2}\sum_a\sum_{i\in a}\mathbf{q}_{i_a}^T \cdot \mathbf{p}_{i_a}^{tc} + \frac{1}{2}\sum_a\sum_{i\in a}\mathbf{q}_{i_a}^T \cdot \mathbf{p}_{i_a}^{dir},$$

(15)

where

$$\mathbf{p}_{i_a}^{tc} = \sum_b \mathrm{ws}(ab)\mathbf{T}(\mathbf{R}_{bi_a}) \cdot \mathbf{m}_b,$$

(16)

and m_b is given in Eq. (12). A Tree Code requires minimal storage, e.g., the multipole moments of a set [m_b in Eq. (16)] are computed whenever needed and then immediately discarded, and the multipolar potentials v_a are never computed. As a consequence, the algorithm is greatly simplified and the hierarchical implementation becomes analogous to that described in Ref. 23; However, the advantages of storage and algorithmic simplifications come at the cost of increased number of floating point operations.

C. Adaptive splitting procedures and well-separateness criteria

Sec. II A made frequent use of the term "well-separateness". We deem the sets a and b to be well-separated if

$$R_{ab}^2 \geq 2(WS+1)^2\left(\left\langle R^2\right\rangle_a - \langle\mathbf{R}\rangle_a^T \cdot \langle\mathbf{R}\rangle_a + \left\langle R^2\right\rangle_b - \langle\mathbf{R}\rangle_b^T \cdot \langle\mathbf{R}\rangle_b\right)$$

(17)

where WS is a parameter, and

$$\langle\mathbf{R}\rangle_a = \frac{1}{N_a}\sum_{i\in a}\mathbf{R}_{i_a}$$

(18)

where N_a is the number of particles in sets a. Note that other well-separatedness criteria exist. [29,41] For example, it is common in so-called "oct-tree" methods to prevent sets from interacting

via multipole expansions unless they are separated by 2 intervening boxes;[22] however, this approach inherently assumes a uniform cubic grid, which is not guaranteed in the binary splitting procedure described below. We note that for chemical applications, the rigorous error bounds originally derived for gravitational problems,[19,27,42] are often found to be too conservative for most chemical applications where charge is largely balanced, leading to much less efficient algorithms in practice.[29] We treat the WS as an empirical parameter, which is discussed in more detail in Sec. III. For additional discussion of error estimates in fast multipole methods, see Ref. 43.

The remainder of this section describes 2 *adaptive* schemes used to split a system of sites into a hierarchy of sets. Both schemes follow the same basic procedure: (1) Given a set of sites, define a splitting center, R_a, which will also serve as the location of the multipole expansion of the set. (2) Define a "splitting plane" passing through R_a, with normal vector $\hat{R}_s$. (3) Partition the set into 2 child sets by assigning each particle to one side of the splitting plane, i.e., site i is in child 1 if

$$\widehat{\mathbf{R}}_s^T \cdot \mathbf{R}_i < \widehat{\mathbf{R}}_s^T \cdot \mathbf{R}_a, \tag{19}$$

or else site i is in child 2. The 2 schemes discussed below differ in their definitions of R_a and R_s, but both are *adaptive* because these definitions depend on the spatial distribution of the sites within the set.

The first scheme, termed "fluc-splitting", defines R_a as the center of distribution, i.e.,

$$\mathbf{R}_a = \langle \mathbf{R} \rangle_a. \tag{20}$$

. The splitting vector, $\mathbf{R}_s$, is chosen as the eigenvector corresponding to the largest eigenvalue of the 3×3 covariance matrix, $\mathbf{X}$

$$X_{j,k} = \frac{1}{N_a} \sum_{i \in a} (R_{i_a, j} - R_{a,j})(R_{i_a, k} - R_{a,k}), \tag{21}$$

where j and k index the Cartesian component. By splitting at the center of distribution, fluc-splitting tries to partition a set such that each of the 2 children contain approximately the same number of particles, and consequently forms a well-balanced binary tree, and the choice of splitting plane attempts to minimize spatial fluctuations.

The second scheme, termed "box-splitting", defines *boundaries* of the distribution, $\mathbf{R}_{max}$ and $\mathbf{R}_{min}$, where, for example, the x-component of $\mathbf{R}_{max}$ is the maximum x-component of any $\mathbf{R}_{i_a}$ in the set a. $\mathbf{R}_{max}$ and $\mathbf{R}_{min}$ therefore define a non-cubic rectangular Cartesian box enclosing the sites. The expansion center is the center of the Cartesian box, i.e.,

$$\mathbf{R}_a = \frac{1}{2}(\mathbf{R}_{max} + \mathbf{R}_{min}), \tag{22}$$

and the splitting vector is either $\hat{x}$, $\hat{y}$, or $\hat{z}$, corresponding to the longest edge of the Cartesian box.

An important aspect common to either of the above described splitting procedures is *termination*, i.e., how many sites the childless nodes should contain. Having too many particles

in a set does not effectively exploit the benefits of the FMM, and having too few particles introduces excessive translations through the tree structure. Termination has previously been discussed by Pérez-Jordá and Yang (PJY),[23] and we interpret their discussion as suggested that optimal performance is obtained by choosing to split a set if

$$\frac{\mathcal{T}^{dir}(N_a)}{\mathcal{T}^{mult}(N_a, L)} > 2K,$$

(23)

where K is a parameter, $T^{dir}(N_a)$ and $T^{mult}(N_a, L)$ are measures of time that would be required to compute the interaction *between 2 sets* of this size directly and via multipole expansions, respectively, N_a is the number of sites in the current set (denoted set a), and L is the multipole expansion order, respectively. PJY developed a tree-code specific to systems containing *monopolar* sites only. In this context, the multipole moments, q_i, and multipolar potentials, p_i, reduce to scalar quantities, and the translation and interaction matrices used in Eq. (12) and Eq. (16) reduce to vectors corresponding to the regular and irregular solid harmonics, respectively. Thus, PJY used the scaling arguments: $\mathcal{T}^{dir}(N_a) \propto N_a^2$ and $\mathcal{T}^{mult}(N_a, L) \propto N_a(L+2)^2$, which, when used in the termination criteria [Eq. (23)] yields childless sets containing N_0 sites

$$N_0 \approx K(L+2)^2,$$

(24)

where the factor of 2 in Eq. (23) is removed since the last split satisfying the inequality approximately halves the distribution.

The present work introduces an adaptive FMM and considers more complicated charge densities, therefore, we must introduce a more general termination criteria. In considering sites with nonzero multipole character, the execution timing estimates in Eq. (23) now depend on site multipole order, L_p, and therefore

$$\frac{\mathcal{T}^{dir}(N_a, L_p)}{\mathcal{T}^{mult}(N_a, L, L_p)} > 2K.$$

(25)

. We choose to "normalize" the ratio in Eq. (25), i.e., choose K, by noting that when a set contains sites with $L_p = L$ only, then that set *should* continue to split until all childless nodes arising from that set contain exactly 1 particle, and thus,

$$K(L) = \frac{\mathcal{T}^{dir}(1, L)}{\mathcal{T}^{mult}(1, L, L)},$$

(26)

which yields the termination condition used in the present work:

$$\frac{\mathcal{T}^{dir}(N_a, L_p)/\mathcal{T}^{dir}(1, L)}{\mathcal{T}^{mult}(N_a, L, L_p)/\mathcal{T}^{mult}(1, L, L)} > 2$$

(27)

In order to apply Eq. (27), one must either use scaling arguments or perform empirical timing measurements to deduce forms for $T^{dir}(N_a, L_p)$ and $T^{mult}(N_a, L, L_p)$. We have chosen to fit the empirically measured times to functional forms that are products of polynomials; each

polynomial representing the dependence on 1 of the arguments. The polynomials and fit coefficients used for our computer architecture are technical details provided in the supporting information.

D. Test systems

TIP3P and Drude water systems—The water boxes and spheres in Fig. 2–Fig. 3 use the non-polarizable 3-point TIP3P water model.[44] The systems were constructed from an equilibrated 216 water box (18.86 Å cube), which was replicated periodically and cut into appropriate sizes and (spherical or cubic) shapes.

The water boxes used in Fig. 4 consider variations of the 5-point polarizable Drude (SWM4-DP) water model.[45] The Drude model has static point charges located at the water H and M sites, and a point charge representation of a polarizable dipole function located at the O center, requiring the treatment of 5 point charges, in total. The systems studied were constructed by replacing the TIP3P waters in the box systems discussed above with the SWM4-DP model, fixing the nuclei, and minimizing the interaction energy through displacement of the Drude particles under periodic boundary conditions. From this 216 water system, larger boxes were constructed from replication of the unit cell. The charge density of each water was then represented in 3 different ways, which we will refer to as the 5S, 2S+1SP, and 1SPD models. The 5S model explicitly treats all 5 point charges, whereas the 2S+1SP model explicitly treats the H point charges and represents the O, M and Drude particle point charges together as a point monopole+dipole function (multipole expansion) located at the O position. The 1SPD model similarly represents the water's charge density as a multipole expansion (up to quadrupole) located at the M-site position.

70S ribosome functional complex—The 3.71Å resolution crystallographic structure of the 70*S* ribosome functional complex[46] (PDBID: 2i1c) was used as the starting positions for heavy atoms. Only the nucleic acid residues were considered, and hydrogens were built using psfgen utility in VMD[47] and the all-atom CHARMM27 nucleic acid force field.[48,49] The resulting test system contained 53876 atoms (35805 heavy atoms).

Full length hammerhead ribozyme—The 2.2Å crystal structure of full length hammerhead ri-bozyme[50] (PDBID: 2goz) was used as the starting positions for heavy atoms, and hydrogens were built using HBUILD facility in the CHARMM program.[51] All the simulations were performed with CHARMM[51] (version c35a1) using the all-atom CHARMM27 nucleic acid force field[48,49] with extension to reactive intermediate models (e.g., transition state mimics)[52] and TIP3P water model.[44]

The system denoted "Full" is the fully solvated model under periodic boundary conditions, and contains the HHR (2021 RNA atoms) late TS mimic model,[53] in a 75Å rhombododecahedron cell containing 9054 waters and 0.14 M NaCl, and includes one Mg^{2+} placed in a bridging position in the active site. Full details into the model system and equilibration procedure are contained in Ref. 53. Timing simulations were performed for 1 ps using the smooth particle mesh Ewald (PME) method[3,54] with a κ value of 0.35 Å^{-1}, 80 FFT grid points for each of the lattice directions, and a B-spline interpolation order of 6, which yields a RFE of approximately 10^{-5}. Non-bonded interactions were treated using an atom-based cutoff of 10 Å with shifted van der Waals potential. Numerical integration was performed using the leap-frog Verlet algorithm with 1 fs time step.[55] Covalent bond lengths involving hydrogen were constrained using the SHAKE algorithm.[56]

The system denoted "Stochastic" is the stochastic boundary model system under non-periodic boundary conditions, and was constructed from the Full system by deleting all water molecules outside of a 25Å radius of the Mg^{2+} ion (7622 total atoms remain, including 1833 waters).

Timing simulations were performed for 1 ps using stochastic boundary molecular dynamics simulations (SBMD)[57,58] in which all atoms outside of the 25Å solvation sphere were fixed. Langevin dynamics was used with a 25Å water boundary potential, including a 5Å buffer region. Non-bonded interactions, SHAKE constraints and numerical integration time steps were identical to the Full system described above. The FMM method is described in the previous sections and used the 10^{-5} RFE parameters in Table I. FMA refers to the fast multipole algorithm implemented in the CHARMM c35a1 program, which is based on the work described in Ref. 59. The FMA used 3 levels and 14 terms, which provided a RFE of 10^{-5}.

III. RESULTS AND DISCUSSION

Sec. III A applies the FMM and TC methods to the crystallographic structure of the 70*S* ribosome functional complex[60] and establishes WS and L parameters at several set levels of relative force error accuracies that are later demonstrated to be transferable. Using these parameters, Sec. III B compares the box- and fluc-splitting procedures described in Sec. II C. Sec. III C compares the timings and relative force errors for the FMM and TC methods when applied to systems containing sites with variable size and multipole expansion orders, and illustrates the utility of the generalized termination criteria [Eq. 27] for the practical automation of linear-scaling calculations at constant error levels.

A. FMM and TC parameters

The FMM and TC methods described in Sec. II have 2 parameters: the maximum angular momentum of the multipole expansion *L* and the *WS* parameter. We performed energy evaluations of the 70*S* ribosome for a series of *WS* and *L* parameters, which thusly form a 2D map of parameter space. Each point on this map contains the time required to compute the energy and the relative force error (RFE), defined by

$$\text{REF} = \sqrt{\frac{\sum_{i=1}^{N} |\tilde{\mathbf{F}}_i - \mathbf{F}_i|^2}{\sum_{i=1}^{N} |\mathbf{F}_i|^2}},$$

(28)

where N is the number of sites, $\mathbf{F}_i$ is the vector of Cartesian force components (minus the electrostatic energy gradient) of site *i* computed exactly from direct interactions, and $\tilde{\mathbf{F}}_i$ is the corresponding vector of approximate force components computed from the FMM or TC algorithms. The optimal parameters are those which yield a desired RFE at minimal cost. Tables of these 2D maps are provided in the supporting information. The optimal parameters for RFE = 10^{-4}, 10^{-5}, and 10^{-6} are provided in Table I. To place these RFE's into perspective, the PME parameters discussed in Sec. II D yields RFE = 10^{-5} when applied to the Full HHR system. Note that the parameters listed in Table I correspond to the box-splitting method, and we were not able to reliably determine optimal parameters for the fluc-splitting method; this is discussed in more detail in Sec. III B. These parameters are only estimates of the error for the chemical systems studied, and do not deliver rigorous error bounds,[19,29,43] hence, actual errors may vary depending on system composition, shape, particle multipole expansion order, splitting procedure, and termination criteria, as are discussed in the remaining subsections.

B. Comparison of adaptive splitting procedures

This section compares the timing and accuracy of the adaptive fluc- and box-splitting procedures, and from this comparison, we conclude that box-splitting is generally superior to fluc-splitting. This conclusion is reached by encountering 2 deficiencies in the fluc-splitting procedure when applied to systems of practical relevance, and much of the following discussion is devoted to describing these 2 deficiencies.

The first deficiency identified in the fluc-splitting procedure is is related to the discussion in Sec. III A, in which it was stated that we were unable to reliably determine WS and L parameters from parameter scans of the 70S ribosome. The parameter scans did not yield RFE's that systematically decreased as WS and/or L increased, and generally yielded errors significantly larger than those observed for the box-splitting. In an effort to understand this, we identified which atoms/childless sets contributed most to the computed RFE. Fig. 1 displays an example of one such childless set and the hierarchical binary genealogy of that set to the initial split. In this example, the childless set is described as having 3 distinct clusters consisting of 16, 5, and 3 atoms; the 2 larger sets are separated by ≈12Å and the smaller is separated from either of the 2 larger sets by ≈50Å . Tracing through the genealogy, the distal small cluster is severed from its backbone during the first split, and the fluc-splitting procedure is unable to remove this small cluster from the remaining bulk of the structure because it contributes negligibly to the fluctuation matrix and splitting center before termination is reached. The contribution of a cluster to the fluctuation matrix is increased as it becomes more separated from the rest of the system or as its size becomes larger relative to the rest of the system; however, this example demonstrates that this phenomenon may be problematic when applied to biomolecules, where small clusters may be formed in an intermediate region of separation and size in relation to its contribution to the fluctuation matrix. The box-splitting procedure overcomes this problem by choosing the splitting plane based on extrema of the distribution. Therefore, if there is a set consisting of 2 separated clusters, their relative sizes and shapes are not important in choosing the splitting plane; it is only important that there is spatial *length*.

The second deficiency identified in the fluc-splitting procedure is "system shape dependence" which, unlike the previous deficiency, can be exhibited in spatially homogeneous systems. The shape of the system effects the fluctuation matrix, i.e., how the system is split, and thus effects the shapes of the child subsystems throughout the entire tree structure. The quality of representing a charge distribution with a finite multipole expansion is not independent of the charge distribution's shape, however. Oddly-shaped or anisotropic subsystems have a larger angular dependence and thus require larger multipole expansion orders to achieve the same accuracy than what would be required for systems with less angular dependence or contained few sites far from the expansion center.

The system shape dependence of the fluc-splitting procedure is illustrated in Fig. 2 and Fig. 3, which compare the timings and RFEs between TIP3P water boxes and spheres, respectively. The fluc-splitting procedure produces RFE values that are 2–10 times larger in the water sphere systems than the water boxes, whereas the box-splitting procedure produces very similar values. Identification of the atoms which contribute largest to the RFE observed in the water spheres when using the fluc-splitting procedure reveal that the errors are concentrated at the center of the sphere, whereas the largest force errors in the water boxes are uniformly distributed throughout the system, and relatively smaller in magnitude. In an effort to understand why the error is large and concentrated at the center of the sphere, we traced the genealogy of a childless set near the center of the sphere, in a manner analogous to Fig. 1. The initial cut forms 2 hemispheres, followed by a cut that quarters the sphere, producing conical/ wedge-like shapes in which the atoms near the center of the sphere form the apex. Most of the atoms within these conical subsystems are located far from the sphere's center, therefore the atoms near the apex are furthest from the expansion center and, by virtue of the conical shape, the spatial fluctuations leave the central atoms at a conical apex in near perpetuity of the binary tree. The last cut before termination yields childless sets containing so few atoms that their spatial anisotropy is not distinguishable from those obtained in the water box systems; the errors, however, arise from well-separated interactions between larger sets and are then propagated down the tree. The water sphere systems may pathologically intensify this phenomenon. This is important in practical applications of molecular simulations that utilize large spherical water drops with stochastic boundary conditions and generalized solvent

boundary potentials. In contrast, the box-splitting procedure is far less dependent on the system shape (see Fig. 2 and Fig. 3) because the shape of each set in the binary tree is a rectangular box, which is independent of the system's shape (only the rectangular *length* can be changed).

C. Comparison of FMM and TC methods

This section compares the timings between the FMM and TC methods as a function of system size, composition (i.e., particle multipole order), and error-level (see Table I). In Sec. IIIB, it was concluded that box-splitting is superior to fluc-splitting and all comparisons in this section are therefore with respect to box-splitting.

For the TIP3P water box systems (Fig. 2), FMM is 21.5 (10^{-4} RFE), 12.7 (10^{-5} RFE), and 8.7 (10^{-6} RFE) times faster than direct evaluation for a system of 100,000 atoms. The ratio of direct and FMM times is a linear function of N and the FMM speedup relative to direct evaluation for $N \geq 10000$ can reasonably be approximated from the above data, e.g., for 10^{-4} RFE, the FMM speedup is $\approx 21.5 \times 10^{-5}N$. In comparison to the TC times, FMM is 2.65 (10^{-4} RFE), 2.53 (10^{-5} RFE), and 2.51 (10^{-6} RFE) times faster. The N at which the FMM and direct evaluation times are approximately equal, i.e., the "break-even point", range from 2000–5000 atoms for these error levels, whereas the TC break-even point ranges from 4000–10000 particles, which is consistent with the TC/FMM ratio of times described above.

It is of interest to compare the methods for systems consisting of sites with different distributions of multipole expansion orders. For the purposes of this comparison, we consider the 3 different charge representations derived from a Drude (SWM4-DP) water model,[45] as described in Sec. II D. Fig. 4 (left) compares FMM, TC, and direct timings and RFEs of the water models using the 10^{-6} RFE parameters in Table I. The 2S+1SP and 1SPD models, when computed directly, are 1.10 times faster and 1.67 times slower than evaluation of the 5S model, respectively. Although higher multipole expansion orders are involved, there are also fewer sites to consider. FMM is approximately 2.35 (5S), 2.20 (1S+1SP), and 2.35 (1SPD) times faster than TC, which suggests that the performance benefits of the FMM is not sensitive to particle angular momentum.

Fig. 4 (right) compares the PJY [Eq. (23), $K = 0.4$, green] and generalized [Eq. (27), red] termination criteria for each water model using the box-splitting FMM. These 2 criteria lead to very similar childless set sizes for monopole systems, and is illustrated by the close agreement in timings and RFEs in Fig. 4. The 2S+1SP timings and RFEs are also quite similar, and this is largely due to large percentage of monopole functions in the system. The 1SPD model yields a striking difference between the termination criteria, however. The generalized termination criteria leads to smaller childless set sizes and therefore reduces the number of direct interactions. Alternatively stated, the PJY criteria nets too many direct interactions, yielding RFEs 10 times below the desired accuracy, and takes 2–3 times longer than the generalized termination criteria. The PJY termination criteria was not designed for treating sites with high particle multipole expansion order; Fig. 4 is an illustration of the generalization's utility and not meant to imply criticism of the PJY's criteria in an application for which it was not designed.

D. Comparison with other methods

We have incorporated our FMM code into the CHARMM c35a1 program and compared our preliminary implementation with other commonly used methods available within the program (see Table II). The methods, system composition, and simulation details are described in Sec. II D. Also shown are standard cutoff methods that, although execute efficiently, yield much higher RFE's. Not surprisingly, for a given system size, the FMM and FMA algorithms are considerably slower than PME since PME requires far fewer direct interactions. In practical

applications, however, the savings afforded by the multipole methods derive from *how* they are used. The computational cost benefits of the multipole methods become apparent with their application to non-periodic simulations employing stochastic boundry conditions, in which case, the number of atoms can be significantly decreased since the requirement to fill a periodic cell with solvent is circumvented. This point is emphasized by the Stochastic HHR simulation times, which indicate that the the multipole methods afford a speed-up of $\approx 2x$ relative to the Full PME simulation time. Also note that the FMM and FMA times are similar. This too is not surprising since the majority of time spent in multipole methods is the evaluation of direct interactions. The FMM and FMA implementations are both based on tree structures and multipole expansions, and thus, both methods are bound by an intrinisic ability of a multipole to represent a charge density, i.e., for a comparable error level, both methods evaluate approximately the same number of direct interactions. The conclusion reached here is that the FMM implementation can adaptively treat higher order atom charge densities, whereas the FMA implementation does not; however, when simple systems of point charges are considered, the FMM algorithm remains competitive. In addition, the adaptive FMM implementation has the advantage of returning relatively stable RFE's for various systems. For example, the FMA parameters described in Sec. II D, which were chosen to produce 10^{-5} RFE for the HRR system, yield a 10^{-7} RFE when applied to the $70S$ ribosome system, i.e., too many direct interactions are computed and thus the FMA requires much more time than our FMM. We note that if one manually changes the FMA parameters for the $70S$ system to yield the same error as the FMM, the FMA and FMM times are very similar, in agreement with Table II.

IV. CONCLUSION

The present work describes the development, implementation and testing of adaptive linear-scaling FMM and TC methods extended to systems of particles with higher-order multipole character. Two different adaptive techniques were explored: fluc-splitting based on the principal components of the fluctuation matrix, and box-splitting based on the Cartesian extremities of the particle distribution. The latter was found to be superior in terms of accuracy for a fairly wide range of particle distribution shapes and multipole orders.

Two different adaptive linear-scaling methods were developed: (1) a tree code (TC) that calculates multipole expansions and their interactions between sets "on-the-fly" and thus involves minimal memory storage and is algorithmically fairly simple, and (2) a fast multipole method (FMM) that utilizes multipole expansions of smaller "child" particle sets to build up the multipole expansions of the "parent" sets using translation theorems, and hence can be made to be more efficient than the TC, as the cost of higher algorithmic complexity. Both methods are demonstrated to be efficient and linear-scaling, with the FMM shown to be generally considerably faster (at fixed error tolerance) by a factor of around 2 for all of the system sizes and multipole orders considered.

Key to the efficiency of both the TC and FMM algorithms is the determination of a generalized termination criteria that allows the methods, for a fixed error tolerance, to be performed with optimal efficiency. The result is an automated procedure for the determination of TC and FMM parameters for generalized multipole particles that, in the case of quadrupole-represented Drude water, effect a speed-up of approximately 2–3 times relative to a multipole-independent termination criteria. The methods developed here represent a potentially important contribution to the arsenal of linear-scaling techniques used for the efficient simulation of complex chemical events using multiscale models.

ACKNOWLEDGMENTS

REFERENCES

1. Greengard L. Science 1994;265:909. [PubMed: 17782141]
2. Brooks CL III. Curr. Opin. Struct. Biol 1995;5:211. [PubMed: 7648323]
3. Sagui C, Darden TA. Annu. Rev. Biophys. Biomol. Struct 1999;28:155. [PubMed: 10410799]
4. Garcia-Viloca M, Gao J, Karplus M, Truhlar DG. Science 2004;303:186. [PubMed: 14716003]
5. Gao J, Ma S, Major D, Nam K, Pu J, Truhlar D. Chemical Reviews 2006;106:3188. [PubMed: 16895324]
6. Stone, A. The Theory of Intermolecular Forces, vol. 32 of International Series of Monographs in Chemistry. Oxford: Clarendon Press; 1996.
7. Halgren TA, Damm W. Curr. Opin. Struct. Biol 2001;11:236. [PubMed: 11297934]
8. Rasmussen TD, Ren P, Ponder JW, Jensen F. Int. J. Quantum Chem 2007;107:1390.
9. Piquemal J, Perera L, Cisneros GA, Ren P, Pedersen LG, Darden TA. J. Chem. Phys 2006;125:054511. [PubMed: 16942230]
10. Cisneros GA, Piquemal J, Darden TA. J. Chem. Phys 2006;125:184101. [PubMed: 17115732]
11. Cisneros G, Piquemal J, Darden TA. J. Chem. Phys 2005;123:044109. [PubMed: 16095348]
12. Ren P, Ponder JW. J. Phys. Chem. B 2004;108:13427.
13. Kaminski GA, Stern harry A, Berne BJ, Friesner RA, Cao YX, Murphey RB, Zhou R, Halgren TA. Int. J. Quantum Chem 2002;23:1515.
14. Liu Y-P, Kim K, Berne BJ, Friesner RA, Rick SW. J. Chem. Phys 1998;108:4739.
15. Gregersen BA, York DM. J. Phys. Chem. B 2005;109:536. [PubMed: 16851046]
16. Gregersen BA, York DM. J. Comput. Chem 2006;27:103. [PubMed: 16273506]
17. York DM, Karplus M. J. Phys. Chem. A 1999;103:11060.
18. Khandogin J, Gregersen BA, Thiel W, York DM. J. Phys. Chem. B 2005;109:9799. [PubMed: 16852180]
19. Greengard L, Rokhlin V. J. Comput. Phys 1987;73:325.
20. Toukmaji AY, Board JA Jr. Comp. Phys. Comm 1996;95:73.
21. Elliott WD, Board JA. SIAM J. Sci. Comput 1996;17:398.
22. Watson MA, Sałek P, Macak P, Helgaker T. J. Chem. Phys 2004;121:2915. [PubMed: 15291602]
23. Pérez-Jordá JM, Yang W. Chem. Phys. Lett 1995;247:484.
24. Barnes J, Hut P. Nature 1986;324:446.
25. Greengard LF, Huang J. J. Comp. Phys 2002:642–658.
26. Greengard L, Rokhlin V. J. Comput. Phys 1997;135:280.
27. Schmidt KE, Lee MA. J. Stat. Phys 1991;63:1223.
28. Petersen HG, Soelvason D, Perram JW, Smith ER. J. Chem. Phys 1994;101:8870.
29. White CA, Head-Gordon M. J. Chem. Phys 1994;101:6593.
30. White CA, Johnson BG, Gill PMW, Head-Gordon M. Chem. Phys. Lett 1994;230:8.
31. Shimada J, Kaneko H, Takada T. J. Comp. Chem 1994;15:28.
32. Makino J. J. Comput. Phys 1999;151:910.
33. Rudberg E, Sałek P. J. Chem. Phys 2006;125:084106. [PubMed: 16965000]
34. Takada T, Kaneko H, Shimada J. J. Comput. Chem 1994;15:28.
35. Wang HY, LeSar R. J. Chem. Phys 1996;104:4173.
36. Kudin KN, Scuseria GE. Phys. Rev. B 2000;61:16440.
37. Shao Y, White CA, Head-Gordon M. J. Chem. Phys 2001;114:6572.

38. Ong ET, Lim KM, Lee KH, Lee HP. J. Comp. Phys 2003;192:244.

39. Helgaker, T.; Jørgensen, P.; Olsen, J. Molecular electronic-structure theory. Hoboken, New Jersey: John Wiley & Sons; 2000.

40. Pèrez-Jordá J, Yang W. J. Chem. Phys 1996;104:8003.

41. Schwegler E, Challacombe M, Head-Gordon M. J. Chem. Phys 1998;109:8764.

42. Barnes JE, Hut P. Astrophys. J. Suppl. S 1989;70:389.

43. Petersen HG, Smith ER, Soelvason D. Proc. R. Soc. Lond. A 1995;448:401.

44. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. J. Chem. Phys 1983;79:926.

45. Lamoureux G, MacKerell AD Jr, Roux B. J. Chem. Phys 2003;119:5185.

46. Korostelev A, Trakhanov S, Laurberg M, Noller HF. Cell 2006;126:1065. [PubMed: 16962654]

47. Humphrey W, Dalke A, Schulten K. J. Mol. Grapics 1996;14:33.

48. Foloppe N, MacKerell AD Jr. J. Comput. Chem 2000;21:86.

49. MacKerell AD Jr, Banavali NK. J. Comput. Chem 2000;21:105.

50. Martick M, Scott WG. Cell 2006;126:309. [PubMed: 16859740]

51. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. J. Comput. Chem 1983;4:187.

52. Mayaan E, Moser A, DM A Jr, York DM. J. Comput. Chem 2007;28:495. [PubMed: 17186477]

53. Lee T-S, Silva-Lopez C, Martick M, Scott WG, York DM. J. Chem. Theory Comput 2007;3:325. [PubMed: 19079784]

54. Essmann U, Perera L, Berkowitz ML, Darden T, Hsing L, Pedersen LG. J. Chem. Phys 1995;103:8577.

55. Allen, M.; Tildesley, D. Computer Simulation of Liquids. Oxford: Oxford University Press; 1987.

56. Ryckaert JP, Ciccotti G, Berendsen HJC. J. Comput. Phys 1977;23:327.

57. Brooks CL III, Karplus M. J. Chem. Phys 1983;79:6312.

58. Brooks CL III, Brunger A, Karplus M. Biopolymers 1985;24:843. [PubMed: 2410050]

59. Board JA Jr, Causey JW, Leathrum JF Jr, Windemuth A, Schulten K. Chem. Phys. Lett 1992;198:89.

60. Cate JH, Yusupov MM, Yusupova GZ, Earnest TN, Noller HF. Science 1999;285:2095. [PubMed: 10497122]
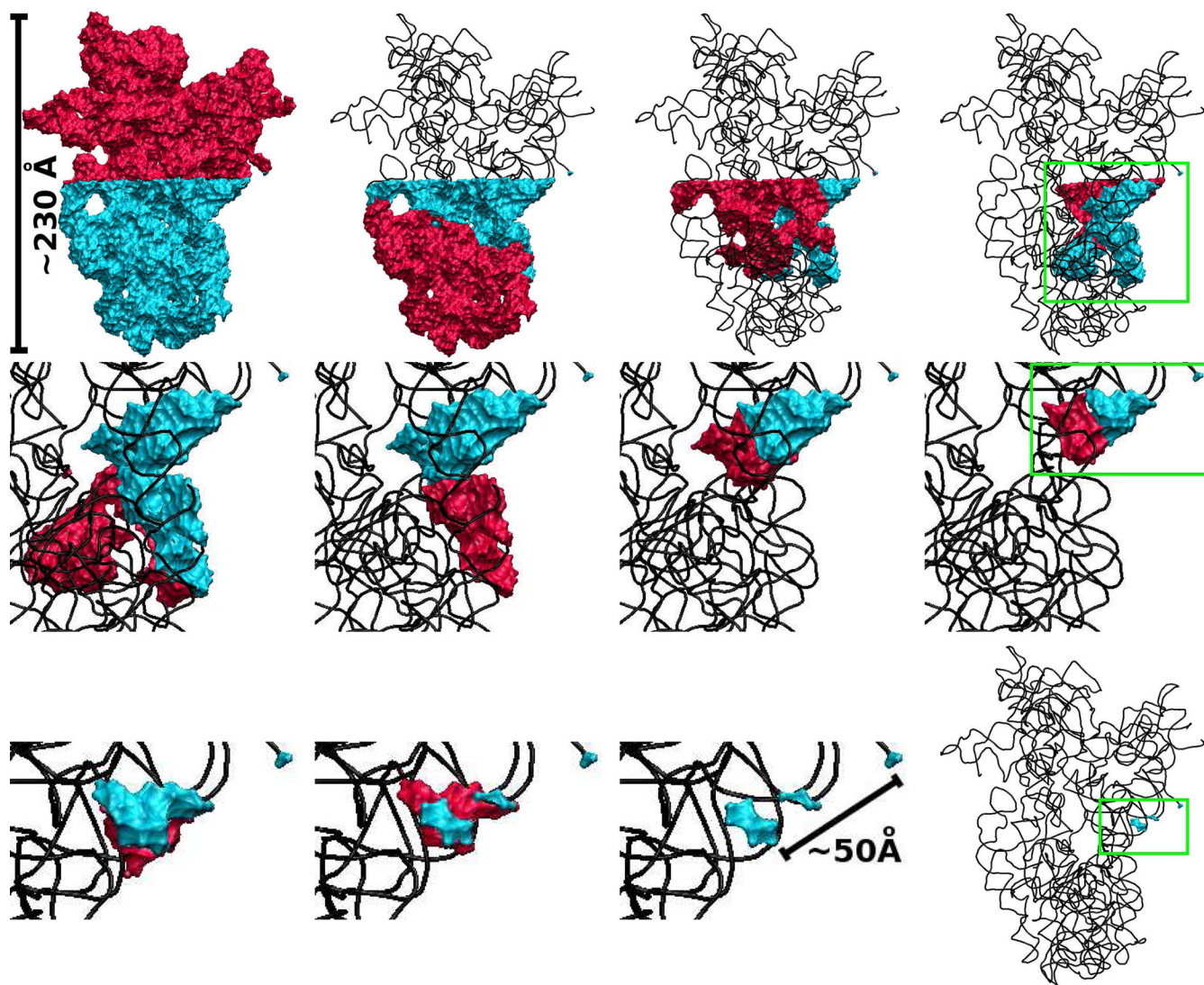
Figure 1.
Genealogical ancestry of an example childless node displaying the symptoms of the self-correction deficiency expressed in the fluc-splitting procedure. The upper-left frame shows the initial split of the system and, moving left-to-right, top-down, all successive splits leading to the childless node. Green boxes indicate the zoom of the next frame. The highlighted areas indicate the atoms within the current set, and the child siblings are colored in red and blue, respectively. The blue highlighted set is the parent set in the succeeding frame.
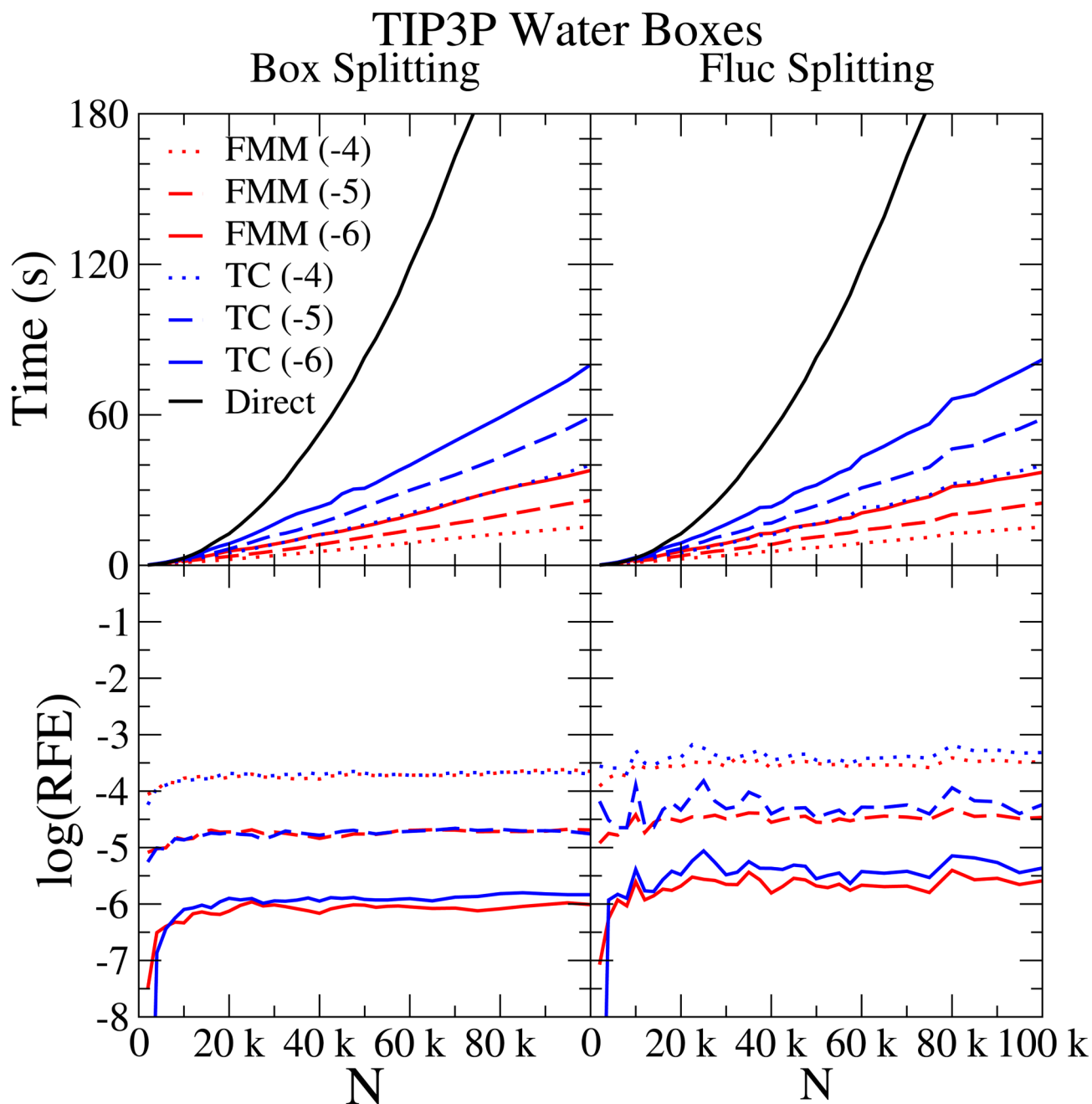
Figure 2.
Comparison between FMM (blue), TC (red), and direct (black) timings (top), relative force errors (bottom), and box- (left) and fluc-splitting procedures (right), as a function of number of atoms in a series of TIP3P water boxes. The number of waters is N/3. The FMM and TC timings are computed at various RFE, corresponding to the parameters listed in Table I. The electrostatic energy, atom potentials, and atom forces are computed.
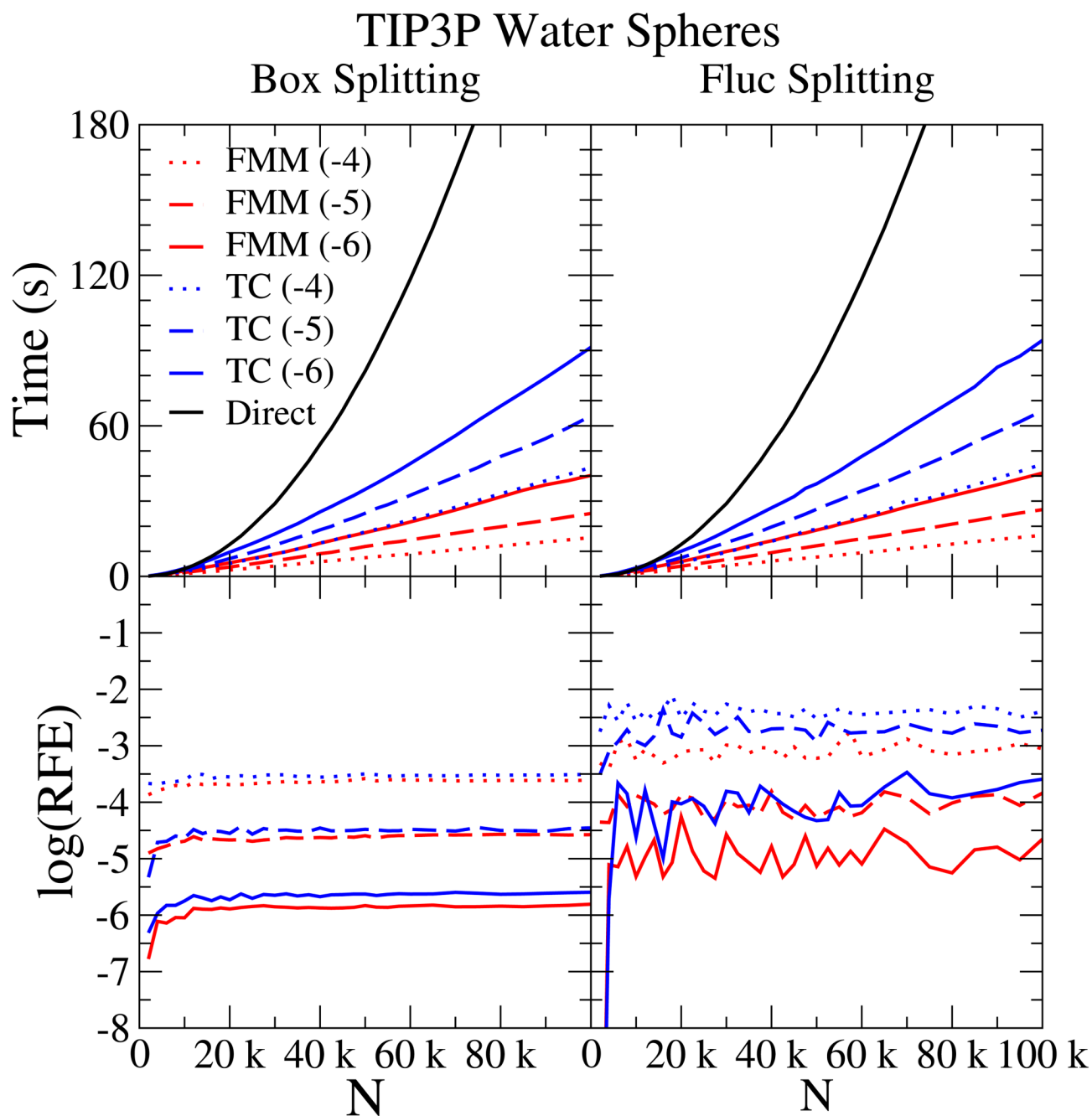
Figure 3.
Comparison of FMM, TC, and direct timings and RFEs for a series of TIP3P water spheres.
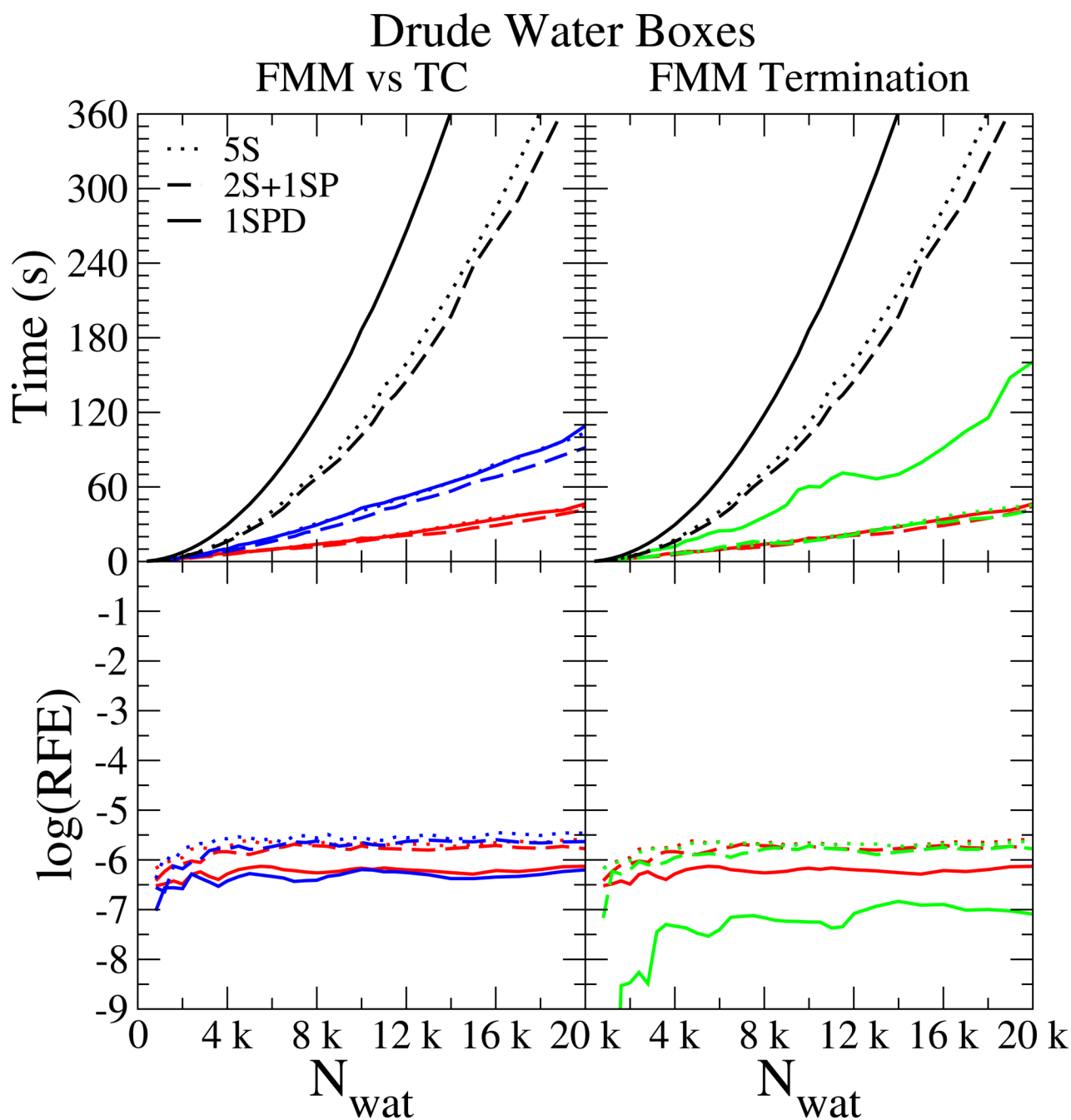See the caption of Fig. 2.

Figure 4.
Comparison of FMM (red), TC (blue), and direct (black) timings and RFEs for a series of
model Drude water boxes. Dotted, dashed, and solid lines refer to the 5S, 2s+1SP, and 1SPD
models described in the text. Only the 10^{-6} accuracy (see Table I) results are shown. Right:
Comparison between the PJY (green) and generalized (red) termination criteria using the box-
splitting FMM. Note that the red and black lines are the same in both the left and right plots.

TABLE I

Optimal FMM and TC parameters for different error tolerance[a]

RFE[b]	FMM		TC	
	WS	L	WS	L
10^{-4}	1.73	6	1.39	6
10^{-5}	20.07	7	1.48	8
10^{-6}	2.20	9	1.81	9

[a]All values in a.u..

[b]Relative force error [see Eq. (28)] observed in the 70*S* ribosome functional complex (Ref. 60).

TABLE II

Simulation times[a]

Method	RFE[b]	Time (min.) Full[c]	Stochastic[d]
No cutoff	0	971	60
PME	10^{-5}	40	...
FMM	10^{-5}	130	22
FMA	10^{-5}	142	23
12Å switched cutoff	10^{-1}	36	6
12Å force shifted cutoff	10^{-2}	27	4

[a]All CPU times reported in this work were performed on a single Intel Xeon 2.66GHz processor (family 15, model 2, stepping 7) equipped with 512 KB cache and 1GB RAM.

[b]Relative force error [see Eq. (28)].

[c]1 ps simulation (1000 steps) of HHR (9054 waters, 29285 atoms total). Only the PME timing employs periodic boundry conditions.

[d]1 ps simulation (1000 steps) of HHR solvated by a 25Å sphere of water around the active site (1833 waters, 7622 atoms total).