A Smooth Particle Mesh Ewald Method

Article *in* The Journal of Chemical Physics · November 1995

DOI: 10.1063/1.470117

6 authors, including:

Ulrich Essmann
Hochschule Bonn-Rhein-Sieg
25 PUBLICATIONS **24,024** CITATIONS

SEE PROFILE

Lalith Perera
National Institute of Environmental Health Sciences
166 PUBLICATIONS **27,532** CITATIONS

SEE PROFILE

Max Berkowitz
University of North Carolina at Chapel Hill
173 PUBLICATIONS **35,067** CITATIONS

SEE PROFILE

A smooth particle mesh Ewald method

Ulrich Essmann, Lalith Perera, and Max L. Berkowitz
Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina 27599

Tom Darden, Hsing Lee, and Lee G. Pedersen
National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709

(Received 5 June 1995; accepted 11 August 1995)

The previously developed particle mesh Ewald method is reformulated in terms of efficient B-spline interpolation of the structure factors. This reformulation allows a natural extension of the method to potentials of the form $1/r^p$ with $p \geqslant 1$. Furthermore, efficient calculation of the virial tensor follows. Use of B-splines in place of Lagrange interpolation leads to analytic gradients as well as a significant improvement in the accuracy. We demonstrate that arbitrary accuracy can be achieved, independent of system size N, at a cost that scales as $N \log(N)$. For biomolecular systems with many thousands of atoms this method permits the use of Ewald summation at a computational cost comparable to that of a simple truncation method of 10 Å or less. © *1995 American Institute of Physics.*

I. INTRODUCTION

The accurate calculation of macromolecular structures and dynamics remains a formidable challenge. In addition to the presently intractable problem of sufficient conformational sampling, the inadequacies of the empirical force fields currently in use are frequently cited as limiting factors. Indeed, although molecular dynamics simulations using explicit solvent molecules rarely exceed a nanosecond, they often exhibit unrealistic behavior on this time scale, such as large deviations from the experimental structure and degradation of secondary structure, making the problem of efficient conformational sampling seem rather pointless. Recent work by us[1–4] as well as others,[5–10] however, has shown that simulations using current force fields without truncation of Coulombic interactions do not exhibit this unrealistic behavior. Furthermore, simulations of peptides[11(a)] and membranes[11(b)] as well as of ions in aqueous solutions[12–16] have provided clear-cut evidence of artifactual behavior due to the use of cutoffs. (Note however, that Steinbach and Brooks[17] have traced some of the artifactual behavior to the use of improper switching functions.) Thus it seems important to focus on the efficient calculation of long-range interactions, as well as on a more careful treatment of the electrostatic boundary conditions.

Molecular dynamics simulations involving explicit solvent molecules have usually been performed under one of the following boundary conditions on the Coulombic interactions:

(1) Periodic boundary conditions using a finite cutoff, similar to the approach of Verlet[18] for Lennard-Jones systems.
(2) Periodic boundary conditions using Ewald summation.[19]
(3) Periodic boundary conditions together with a reaction field.[20]
(4) Nonperiodic boundary conditions together with some treatment of the system-environment interface.[21,22]

Objections can be raised to each of these choices. The Ewald sum certainly seems appropriate for crystal simulations, and may be the best current choice for macromolecular solution simulations as well. In the past, it has not been used for large systems, due to the prohibitive cost of the usual implementation, which is an order N^2 algorithm. (The standard Ewald sum can, however, be implemented in order $N^{3/2}$ steps).[23]

In a previous article,[24] we proposed an order $N \log(N)$ algorithm for calculating the Ewald sum in large systems. This algorithm, called the particle mesh Ewald (PME) method, was inspired by Hockney and Eastwood's[25] particle–particle particle–mesh method of splitting the total electrostatic energy into local interactions which are computed explicitly and long-range interactions which are approximated by a discrete convolution on an interpolating grid, using the 3D fast Fourier transform (3DFFT) to efficiently perform the convolution. However, rather than using their switching function approach to split the total energy, our method used the split of the total electrostatic energy into direct and reciprocal Ewald sums (discussed below).

The reciprocal Ewald sum is the solution to Poisson's equation in periodic boundary conditions, with Gaussian charge densities as sources. York and Yang[26] developed a method to solve Poisson's equation directly using the 3DFFT. They were able to achieve very high accuracy using this approach, but it is currently more costly in CPU time. While the original particle–mesh approach of Hockney and Eastwood[25] is quite efficient, high accuracy is not easily achieved. Luty *et al.*,[27] as well as Shimada *et al.*,[28] have implemented it for macromolecular simulations. A thorough review of various techniques for performing Ewald sums is given by Toukmaji and Board.[29]

A different approach to efficient macromolecular electrostatics is based on the fast multipole algorithm of Greengard and Rokhlin.[30] In this method, the simulation cell is divided into a tree of progressively finer subcells. Electrostatic interactions involving particles in the same or neighboring smallest subcell are computed exactly, while those involving subcells of more distant particles are approximated by a multipole expansion. Several groups have implemented this approach for large systems.[28,31–34]

The layout of the article is as follows. In the next section we give the basic expressions for energies, forces, and stress tensor. We then recast our approximation to the reciprocal Ewald sum in terms of an approximation to the electrostatic structure factors. We will see that if the atomic charge distribution is approximated by a gridded charge distribution, the resulting approximate structure factors are obtained by the 3DFFT applied to the grid. This idea of gridding charges has been used by protein crystallographers for over twenty years.[35] Next we provide some introductory discussion of interpolation with Cardinal B-splines, and then apply the Euler exponential spline to arrive at smooth approximations of the structure factors and hence the energy, forces, and stress tensors. Advantages of our B-spline interpolation-based approach to the Ewald sum are:

(1) The potential and forces are smooth functions of the particle positions at any level of accuracy.
(2) The approximation, which depends on atomic positions only through their fractional coordinates, extends naturally to general unit cells.
(3) Since the lattice sums for related pair potentials, such as London dispersion interactions, involve analogous structure factors in their reciprocal sums, the method easily generalizes to them.

The computer implementation of the approximation, which has been implemented as an option in the Sander module of Amber[36] version 4.1 by the Kollman lab at UCSF, is then described. Fortran code for the PME subroutines is available by e-mail request to darden@niehs.nih.gov A C-language version of this algorithm has been implemented by the Board lab at Duke.[37] In the next section, we discuss the generalization of the B-spline PME approximation to lattice dispersion sums. In the final section, we discuss results for timing and accuracy, energy conservation, and the effect of the various methods on results of molecular dynamics simulations of pure water. Some of the more technical material referred to in the text appears in the appendices.

II. COULOMBIC LATTICE SUMS

Suppose there are N point charges $q_1,q_2,\ldots,q_N$ at positions $\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N$ within the unit cell U satisfying $q_1+q_2+\cdots+q_N=0$. The vectors $\mathbf{a}_\alpha$, $\alpha=1,2,3$, which need not be orthogonal, form the edges of the unit cell. The conjugate reciprocal vectors $\mathbf{a}_\alpha^*$ are defined by the relations $\mathbf{a}_\alpha^* \cdot \mathbf{a}_\beta = \delta_{\alpha\beta}$ (the Kronecker delta), for $\alpha,\beta=1,2,3$. The point charge q_i at position $\mathbf{r}_i$ has fractional coordinates $s_{\alpha i}, \alpha=1,2,3$ defined by $s_{\alpha i} = \mathbf{a}_\alpha^* \cdot \mathbf{r}_i$. The charges interact according to Coulomb's law with periodic boundary conditions. Thus a point charge q_i at position $\mathbf{r}_i$ interacts with other charges q_j, $j\neq i$ at positions $\mathbf{r}_j$ as well as with all of their periodic images at positions $\mathbf{r}_j+n_1\mathbf{a}_1+n_2\mathbf{a}_2+n_3\mathbf{a}_3$ for all integers n_1, n_2, n_3. It also interacts with its own periodic images at $\mathbf{r}_i+n_1\mathbf{a}_1+n_2\mathbf{a}_2+n_3\mathbf{a}_3$ for all such integers n_α with n_1, n_2, n_3 not all zero. The electrostatic energy of the unit cell U can then be written

$$E(\mathbf{r}_1,\ldots,\mathbf{r}_N)=\frac{1}{2}\sum_n{}'\sum_i\sum_j \frac{q_iq_j}{|\mathbf{r}_i-\mathbf{r}_j+n|}, \tag{2.1}$$

where the outer sum is over the vectors $\mathbf{n}=n_1\mathbf{a}_1+n_2\mathbf{a}_2+n_3\mathbf{a}_3$, the prime indicating that terms with $i=j$ and $\mathbf{n}=0$ are omitted.

The outer infinite series in Eq. (2.1) is conditionally convergent, meaning that the result depends on the manner in which the numbers n_1, n_2, n_3 tend to infinity (if the unit cell were not neutral it would converge to plus or minus infinity). Using a theta transformation due to Riemann, Ewald[38] replaced the sum in Eq. (2.1) by the sum of two absolutely convergent series, a direct sum in Cartesian space, and a reciprocal sum in Fourier space. A number of alternative derivations of the Ewald summation method exist,[19,23,39–42] both for the Coulombic potential and for potentials such as the dispersion or van der Waals, which decay faster. In Appendix A we provide a derivation following the approach of Smith,[41] which we extend to general potentials of the form $1/r^p$, $p\geqslant 1$.

In molecular systems, where point charges are typically fractional charges positioned at the atomic nuclei, the Ewald sum must be modified. In most simulations the potential functions have been parametrized in such a way that the ''nonbond'' interactions between certain pairs within the same molecule are omitted, being handled instead by other terms in the potential. For example, the Coulomb and Lennard-Jones interactions up to the second nearest neighbor are usually omitted. The atom pairs (i,j), for which nonimaged nonbond interactions are not calculated, are said to belong to the masked pairlist M. Since their interactions have been included in Eq. (2.1), we must subtract $\Sigma_{(i,j)\in M} q_iq_j/|\mathbf{r}_j-\mathbf{r}_j|$ from the energy computed using the direct and reciprocal potentials. Computationally, however, it is more convenient to skip masked pairs in the direct sum that are computed at the same time as the other nonbond interactions.

Define the reciprocal lattice vectors $\mathbf{m}$ by $\mathbf{m} = m_1\mathbf{a}_1^* + m_2\mathbf{a}_2^* + m_3\mathbf{a}_3^*$ with m_1, m_2, m_3 integers not all zero, and the structure factor $S(\mathbf{m})$ by

$$\begin{aligned} S(\mathbf{m}) &= \sum_{j=1}^N q_j \exp(2\pi i\mathbf{m}\cdot\mathbf{r}_j) \\ &= \sum_{j=1}^N q_j \exp[2\pi i(m_1s_{1j}+m_2s_{2j}+m_3s_{3j})], \end{aligned} \tag{2.2}$$

where $s_{\alpha j}$, $\alpha=1,2,3$ are the fractional coordinates of atom j, defined above. The electrostatic energy in Eq. (2.1) can then be written as $E=E_{\mathrm{dir}}+E_{\mathrm{rec}}+E_{\mathrm{corr}}$, where

$$E_{\mathrm{dir}}=\frac{1}{2}\sum_n^*\sum_{i,j=1}^N \frac{q_jq_j\,\mathrm{erfc}(\beta|\mathbf{r}_j-\mathbf{r}_i+\mathbf{n}|)}{|\mathbf{r}_j-\mathbf{r}_i+\mathbf{n}|}, \tag{2.3}$$

$$E_{\mathrm{rec}}=\frac{1}{2\pi V}\sum_{\mathbf{m}\neq 0}\frac{\exp(-\pi^2\mathbf{m}^2/\beta^2)}{\mathbf{m}^2}S(\mathbf{m})S(-\mathbf{m}), \tag{2.4}$$

$$E_{\mathrm{corr}}=-\frac{1}{2}\sum_{(i,j)\in M}\frac{q_iq_j\,\mathrm{erf}(\beta|\mathbf{r}_i-\mathbf{r}_j|)}{|\mathbf{r}_i-\mathbf{r}_j|}-\frac{\beta}{\sqrt{\pi}}\sum_{i=1}^N q_i^2\,. \tag{2.5}$$

The asterisk in Eq. (2.3) denotes that terms with $\mathbf{n}=0$ and $i=j$ or $(i,j)\in M$ are omitted. In Eq. (2.4) $V=\mathbf{a}_1\cdot\mathbf{a}_2\times\mathbf{a}_3$ is

the volume of the unit cell, and in Eq. (2.5), $\mathrm{erf}(x)=1-\mathrm{erfc}(x)$ is the error function. The second term in E_{corr} is usually referred to as the self energy E_{self}. A further energy contribution due to the unit cell dipole moment is sometimes included (see Appendix A). The consequences of including this term have recently been discussed by Roberts and Schnitker.[14,43] Since this term involves a simple order N calculation and the focus in this article is the efficient calculation of the traditional Ewald sum, we do not explicitly consider it further.

The Coulombic force on atom i can be obtained by differentiating the sum $E(\mathbf{r}_1,...,\mathbf{r}_N)=E_{\mathrm{dir}}+E_{\mathrm{rec}}+E_{\mathrm{corr}}$ with respect to $\mathbf{r}_i$. We will refer to the individual terms $-\partial E_{\mathrm{dir}}/\partial \mathbf{r}_i$, $-\partial E_{\mathrm{rec}}/\partial \mathbf{r}_i$, and $-\partial E_{\mathrm{corr}}/\partial \mathbf{r}_i$ as the ''direct force,'' ''reciprocal force,'' and ''correction force,'' respectively. (In contrast to the case of rigid molecules, the correction energy gives rise to nonzero forces in flexible molecules). We also must develop the stress tensor II, in order to perform constant pressure simulations using Ewald summations. The (simpler) case of isotropic volume scaling is treated in Smith,[44] following the volume scaling technique of Andersen.[45] This scaling method was generalized to the full 3×3 virial tensor in arbitrary unit cells by Parrinello and Rahman[46] and extended to the long-ranged Coulombic interactions in Nosé and Klein.[47] A recent article by Smith[48] treats the necessary corrections to the virial due to the unit cell dipole contribution. We follow Nosé and Klein with modifications due to the correction energy in Eq. (2.5). We first treat the case of flexible molecules (i.e., the ''atomic virial''). Let $\mathbf{a}$ be the 3×3 matrix having the lattice vectors $\mathbf{a}_\alpha$, $\alpha=1,2,3$ as columns. Note that the volume V of the unit cell is given by the determinant of $\mathbf{a}$ and that $\mathbf{a}^{-1}$ is the 3×3 matrix having the reciprocal lattice vectors $\mathbf{a}_\alpha^*$ as rows. Let $s_{\alpha i}$, $\alpha=1,2,3$, denote the fractional coordinates of atom $i, i=1,\ldots,N$. We can write the energy of the unit cell as $E=E(s_{11},\ldots,s_{3N};\mathbf{a})$. Referring to Eq. (A7) of Nosé and Klein,[47] the 3×3 stress tensor II satisfies $-\partial E_{(s_{11},\ldots,s_{3N};\mathbf{a})}/\partial \mathbf{a}_{\alpha\beta}=V\Sigma_{\gamma=1}^{3}\Pi_{\alpha\gamma}\mathbf{a}_{\beta\gamma}^{-1}$, for $\alpha=1,2,3$, and thus, examining the three components of E in turn, and using the identities $\partial V/\partial \mathbf{a}_{\alpha\beta}=V\mathbf{a}_{\beta\alpha}^{-1}$ and $\partial \mathbf{a}_{\mu\nu}^{-1}/\partial \mathbf{a}_{\alpha\beta}=\mathbf{a}_{\mu\alpha}^{-1}\mathbf{a}_{\beta\nu}^{-1}$, we have that $\Pi=\Pi_{\mathrm{dir}}+\Pi_{\mathrm{rec}}+\Pi_{\mathrm{corr}}$, where

$$V\Pi_{\mathrm{dir}\,\alpha\beta}=\frac{1}{2}\sum_{\mathbf{n}}{}^{*}\sum_{i,j=1}^{N}q_iq_j\left(\frac{\mathrm{erfc}(\beta|\mathbf{r}_i-\mathbf{r}_j+\mathbf{n}|)}{|\mathbf{r}_i-\mathbf{r}_j+\mathbf{n}|^3}+\frac{2\beta\pi^{-1/2}\exp(-(\beta|\mathbf{r}_i-\mathbf{r}_j+\mathbf{n}|)^2)}{|\mathbf{r}_i-\mathbf{r}_j+\mathbf{n}|^2}\right)\times(\mathbf{r}_i-\mathbf{r}_j+\mathbf{n})_\alpha(\mathbf{r}_i-\mathbf{r}_j+\mathbf{n})_\beta, \tag{2.6}$$

$$V\Pi_{\mathrm{rec}\,\alpha\beta}=\frac{1}{2\pi V}\sum_{\mathbf{m}\neq0}\frac{\exp(-\pi^2\mathbf{m}^2/\beta^2)}{\mathbf{m}^2}S(\mathbf{m})S(-\mathbf{m})\times\left(\delta_{\alpha\beta}-2\,\frac{1+\pi^2\mathbf{m}^2/\beta^2}{\mathbf{m}^2}\,\mathbf{m}_\alpha\mathbf{m}_\beta\right), \tag{2.7}$$

$$V\Pi_{\mathrm{corr}\,\alpha\beta}=\frac{1}{2}\sum_{(i,j)\in M}q_iq_j\left(-\frac{\mathrm{erf}(\beta|\mathbf{r}_i-\mathbf{r}_j|)}{|\mathbf{r}_i-\mathbf{r}_j|^3}+\frac{2\beta\pi^{-1/2}\exp[-(\beta|\mathbf{r}_i-\mathbf{r}_j|)^2]}{|\mathbf{r}_i-\mathbf{r}_j|^2}\right)\times(\mathbf{r}_i-\mathbf{r}_j)_\alpha(\mathbf{r}_i-\mathbf{r}_j)_\beta. \tag{2.8}$$

In the flexible molecule case, the trace of $V\Pi$ equals the negative of the electrostatic energy E (see Smith).[44] Note that other nonbond interactions such as van der Waals interactions also contribute to the virial tensor. In the flexible molecule case the virial tensor includes additional contributions due to the bond, angle, and dihedral interactions.[49] To treat rigid molecules, which do not involve intramolecular interactions, we let $\mathbf{r}_{0i}$ denote the center of mass of the molecule containing atom i, and let $\mathbf{p}_i=\mathbf{r}_i-\mathbf{r}_{0i}$. The vector $\mathbf{p}_i$ does not scale with changing $\mathbf{a}$, although $\mathbf{r}_{0i}$ does. Then the stress tensor $V\Pi$ is given by the above sum $V\Pi_{\mathrm{dir}}+V\Pi_{\mathrm{rec}}+V\Pi_{\mathrm{corr}}$ minus the tensor $\Sigma_{i=1}^{N}(\mathbf{f}_i)_\alpha(\mathbf{p}_i)_\beta$, where $\mathbf{f}_i$ denotes the total nonbond force on atom i, (again see Smith[44] for the discussion of the isotropic case).

III. PIECEWISE LAGRANGIAN INTERPOLATION AND STRUCTURE FACTORS

In order to approximate the electrostatic structure factors, we need to interpolate the complex exponentials appearing in Eq. (2.2). Given positive integers K_1, K_2, K_3 and a point $\mathbf{r}$ in the unit cell, we denote its scaled fractional coordinates by u_1, u_2, u_3, i.e., $u_\alpha=K_\alpha\mathbf{a}_\alpha^*\cdot\mathbf{r}$, for $\alpha=1,2,3$. Due to the periodic boundary conditions, we may assume that $0\leq u_\alpha<K_\alpha$. Then

$$\exp(2\pi i\mathbf{m}\cdot\mathbf{r})=\exp\left(2\pi i\,\frac{m_1u_1}{K_1}\right)\cdot\exp\left(2\pi i\,\frac{m_2u_2}{K_2}\right)\cdot\exp\left(2\pi i\,\frac{m_3u_3}{K_3}\right). \tag{3.1}$$

For real numbers u, let $[u]$ denote the integer part of u, that is, the unique integer satisfying $[u]\leq u\leq[u]+1$. Then, using linear interpolation, we can approximate the individual exponentials on the right hand side of Eq. (3.1) by

$$\exp\left(2\pi i\,\frac{m_\alpha}{K_\alpha}\,u_\alpha\right)\approx(1-(u_\alpha-[u_\alpha]))\cdot\exp\left(2\pi i\,\frac{m_\alpha}{K_\alpha}\,[u_\alpha]\right)+(u_\alpha-[u_\alpha])\cdot\exp\left(2\pi i\,\frac{m_\alpha}{K_\alpha}\,([u_\alpha]+1)\right). \tag{3.2}$$

We then approximate the product in Eq. (3.1) by the product of the right-hand side of Eq. (3.2) for $\alpha=1,2,3$, which expands (trilinear interpolation) to a sum over eight terms. This can be expressed more conveniently for our current purposes. Let $W_2(u)$ denote the linear hat function given by $W_2(u)=1-|u|$ for $-1\leq u\leq1$, $W_2(u)=0$ for $|u|>1$. Then we can rewrite Eq. (3.2) as

$$\exp\left(2\pi i\,\frac{m_\alpha}{K_\alpha}\,u_\alpha\right)\approx\sum_{k=-\infty}^{\infty}W_2(u_\alpha-k)\cdot\exp\left(2\pi i\,\frac{m_\alpha}{K_\alpha}\,k\right). \tag{3.3}$$

Note that the sum in Eq. (3.3) is actually finite, since W_2 has bounded support. [The support of a function $f(u)$ is the set of u values for which $f(u)\neq 0$.]

Higher order piecewise Lagrangian interpolation of the complex exponentials can be expressed in a similar fashion. Consider piecewise $2p$-th order Lagrange interpolation of $\exp(2\pi imu/K)$ using the points $[u]-p+1$, $[u]-p+2,\ldots,[u]+p$. Let $W_{2p}(u)=0$ for $|u|>p$; and for $-p\leqslant u\leqslant p$ define it by

$$W_{2p}(u)=\frac{\Pi_{j=-p,j\neq k}^{p-1}(u+j-k)}{\Pi_{j=-p,j\neq k}^{p-1}(j-k)},$$
$$\text{for } k\leqslant u\leqslant k+1, k=-p,-p+1,\ldots,p-1. \tag{3.4}$$

Note that when $p=1$ this agrees with W_2 defined above. Using this function we can write the usual piecewise Lagrange interpolation formula for the complex exponential as

$$\exp\left(2\pi i\,\frac{m_\alpha}{K_\alpha}\,u_\alpha\right)\approx\sum_{k=-\infty}^{\infty}W_{2p}(u_\alpha-k)\cdot\exp\left(2\pi i\,\frac{m_\alpha}{K_\alpha}\,k\right). \tag{3.5}$$

From standard estimates for Lagrangian interpolation, the error in this approximation is bounded by $(2p)!/(p!)^2[\pi|m_\alpha|/(2K_\alpha)]^{2p}$.

Approximating the product in Eq. (3.1) by the product of the right-hand sides of Eq. (3.5), for $\alpha=1,2,3$ we have for the approximate structure factor

$$\begin{aligned}S(\mathbf{m})\approx\tilde{S}(\mathbf{m})&=\sum_{i=1}^{N}q_i\sum_{k_1=-\infty}^{\infty}\sum_{k_2=-\infty}^{\infty}\sum_{k_3=-\infty}^{\infty}W_{2p}(u_{1i}-k_1)\\&\quad\cdot W_{2p}(u_{2i}-k_2)\cdot W_{2p}(u_{3i}-k_3)\\&\quad\cdot\exp\left(2\pi i\,\frac{m_1}{K_1}\,k_1\right)\cdot\exp\left(2\pi i\,\frac{m_2}{K_2}\,k_2\right)\\&\quad\cdot\exp\left(2\pi i\,\frac{m_3}{K_3}\,k_3\right)\\&=\sum_{k_1=0}^{K_1-1}\sum_{k_2=0}^{K_2-1}\sum_{k_3=0}^{K_3-1}Q(k_1,k_2,k_3)\exp\left[2\pi i\cdot\left(\frac{m_1k_1}{K_1}\right.\right.\\&\quad\left.\left.+\frac{m_2k_2}{K_2}+\frac{m_3k_3}{K_3}\right)\right]\\&=F(Q)(m_1,m_2,m_3),\end{aligned} \tag{3.6}$$

where $\mathbf{F}(Q)$ is the discrete Fourier transform (see Appendix B) of the array Q defined by

$$\begin{aligned}\mathbf{Q}(k_1,k_2,k_3)=\sum_{i=1}^{N}\sum_{n_1,n_2,n_3}&q_iW_{2p}(u_{1i}-k_1-n_1K_1)\\&\times W_{2p}(u_{2i}-k_2-n_2K_2)\\&\cdot W_{2p}(u_{3i}-k_3-n_3K_3).\end{aligned} \tag{3.7}$$

In Eq. (3.7) the inner sums are over all integers n_1, n_2, n_3.

The approximate unit cell energy obtained using this approach with structure factors and piecewise Lagrangian interpolation is equivalent to our previous result.[24] To show this, we define as previously a reciprocal pair potential $\psi_{\rm rec}$, whose values at integers (l_1,l_2,l_3) is given by

$$\begin{aligned}\psi_{\rm rec}(l_1,l_2,l_3)&=\frac{1}{\pi V}\sum_{\mathbf{m}\neq 0}\frac{\exp(-\pi^2\mathbf{m}^2/\beta^2)}{\mathbf{m}^2}\\&\quad\times\exp\left(2\pi i\left[\frac{m_1l_1}{K_1}+\frac{m_2l_2}{K_2}+\frac{m_3l_3}{K_3}\right]\right)\\&=F(C)(l_1,l_2,l_3),\end{aligned} \tag{3.8}$$

where C is the array given by

$$C(m_1,m_2,m_3)=\frac{1}{\pi V}\frac{\exp(-\pi^2\mathbf{m}^2/\beta^2)}{\mathbf{m}^2}$$
$$\text{for } \mathbf{m}\neq 0, C(0,0,0)=0 \tag{3.9}$$

with $\mathbf{m}$ defined by $\mathbf{m}=m_1'\mathbf{a}_1^*+m_2'\mathbf{a}_2^*+m_3'\mathbf{a}_3^*$, where $m_i'=m_i$ for $0\leqslant m_i\leqslant K/2$ and $m_i'=m_i-K_i$ otherwise. Note that $C=F^{-1}(\psi_{\rm rec})$. Using the approximation $S(\mathbf{m})\approx\tilde{S}(\mathbf{m})=F(Q)(m_1,m_2,m_3)$ for the structure factors and the identities (B3) and (B4) in Appendix B, the reciprocal sum in the unit cell energy is approximated by

$$\begin{aligned}E_{\rm rec}\approx\tilde{E}_{\rm rec}&=\frac{1}{2\pi V}\sum_{\mathbf{m}\neq 0}\frac{\exp(-\pi^2\mathbf{m}^2/\beta^2)}{\mathbf{m}^2}\\&\quad\times F(Q)(m_1,m_2,m_3)F(Q)(-m_1,-m_2,-m_3)\\&=\frac{1}{2}\sum_{m_1=0}^{K_1-1}\sum_{m_2=0}^{K_2-1}\sum_{m_3=0}^{K_3-1}F^{-1}(\psi_{\rm rec})(m_1,m_2,m_3)\\&\quad\times F(Q)(m_1,m_2,m_3)\cdot K_1K_2K_3\cdot F^{-1}(Q)(m_1,m_2,m_3)\\&=\frac{1}{2}\sum_{m_1=0}^{K_1-1}\sum_{m_2=0}^{K_2-1}\sum_{m_3=0}^{K_3-1}Q(m_1,m_2,m_3)\\&\quad\cdot(\psi_{\rm rec}\star Q)(m_1,m_2,m_3),\end{aligned} \tag{3.10}$$

which is identical to the PME expression derived previously.[24]

IV. CARDINAL *B*-SPLINE INTERPOLATION

From the definition in Eq. (3.4) it is evident that the Lagrangian weight functions $W_{2p}(u)$ are continuous and therefore give rise to approximate unit cell energies which are continuous as functions of particle positions. It is, however, also clear that $W_{2p}(u)$ are only piecewise differentiable, so the approximate reciprocal energy cannot be differentiated to arrive at Coulombic forces. Instead, in the original PME method, we interpolated the forces as well. In contrast, the Cardinal B-splines $M_{2p}(u)$, which we now describe, lead analogously via spline interpolation of the complex exponential $\exp(2\pi i\mathbf{m}\cdot\mathbf{r})$ to approximate structure factors and hence approximate reciprocal energies $E_{\rm rec}$ which can be analytically differentiated to give the reciprocal forces

and stress tensors. Later, we will show that the approximation is also superior to that for Lagrangian interpolation of the same order.

For any real number u, let $M_2(u)$ denote the linear hat function given by $M_2(u)=1-|u-1|$ for $0\leqslant u\leqslant 2$ and $M_2(u)=0$ for $u<0$ or $u>2$. For n greater than 2, define $M_n(u)$ by the recursion

$$M_n(u)=\frac{u}{n-1}M_{n-1}(u)+\frac{n-u}{n-1}M_{n-1}(u-1). \tag{4.1}$$

One of the many convenient properties of Cardinal B-splines is that they can be easily differentiated analytically. For $n>2$

$$\frac{d}{du}M_n(u)=M_{n-1}(u)-M_{n-1}(u-1). \tag{4.2}$$

From Eq. (4.1) one can see that $M_n(u)$ has finite support, being zero outside the interval $0\leqslant u\leqslant n$. From Eq. (4.2) it is clear that $M_n(u)$ is $n-2$ times continuously differentiable. Note that $M_2(u)$ is a simple translate of $W_2(u)$ defined above. More details about Cardinal B-splines are given in Appendix C.

Interpolation using splines is generally more complex than Lagrangian interpolation, since the interpolation coefficients are obtained from the solution of a system of linear equations, which might not be full rank. Fortunately, interpolation of complex exponentials admits a particularly elegant and simple solution, called the Euler exponential spline. In Appendix C we show that when n is even we can write

$$\exp\left(2\pi i\frac{m_i}{K_i}u_i\right)\approx b_i(m_i)\sum_{k=-\infty}^{\infty}M_n(u_i-k)\cdot\exp\left(2\pi i\frac{m_i}{K_i}k\right), \tag{4.3}$$

where $b_i(m_i)$ is given by

$$b_i(m_i)=\exp(2\pi i(n-1)m_i/K_i)\times\left[\sum_{k=0}^{n-2}M_n(k+1)\exp(2\pi i m_i k/K_i)\right]^{-1}. \tag{4.4}$$

The error in this approximation is bounded from above by $(2|m_i|/K_i)^n$ (see Ref. 50), which is superior to the error bound for Lagrange interpolation (above). When n is odd and $2|m_i|=K_i$ this interpolation result fails but, since it occurs in the tail of the reciprocal sum, we can set $b(m_i)$ arbitrarily to zero in this case. Proceeding as above, we can then approximate the structure factor by

$$\tilde{S}(\mathbf{m})=b_1(m_1)b_2(m_2)b_3(m_3)F(Q)(m_1,m_2,m_3), \tag{4.5}$$

where the array Q is given by

$$Q(k_1,k_2,k_3)=\sum_{i=1}^{N}\sum_{n_1,n_2,n_3}q_iM_n(u_{1i}-k_1-n_1K_1)\times M_n(u_{2i}-k_2-n_2K_2)\cdot M_n(u_{3i}-k_3-n_3K_3). \tag{4.6}$$

The approximate reciprocal energy is now given by

$$\tilde{E}_{\mathrm{rec}}=\frac{1}{2\pi V}\sum_{m\neq 0}\frac{\exp(-\pi^2\mathbf{m}^2/\beta^2)}{\mathbf{m}^2}B(m_1,m_2,m_3)\cdot F(Q)(m_1,m_2,m_3)F(Q)(-m_1,-m_2,-m_3)=\frac{1}{2}\sum_{m_1=0}^{K_1-1}\sum_{m_2=0}^{K_2-1}\sum_{m_3=0}^{K_3-1}Q(m_1,m_2,m_3)\cdot(\theta_{\mathrm{rec}}\star Q)(m_1,m_2,m_3), \tag{4.7}$$

where

$$B(m_1,m_2,m_3)=|b_1(m_1)|^2\cdot|b_2(m_2)|^2\cdot|b_3(m_3)|^2, \tag{4.8}$$

and the pair potential θ_{rec} is given by $\theta_{\mathrm{rec}}=F(B\cdot C)$. Note that Q is now $n-2$ times continuously differentiable in the particle positions. To obtain the reciprocal atomic force we differentiate the second expression in Eq. (4.7) with respect to $\mathbf{r}_i$. Noting that θ_{rec} does not depend on particle positions, we get for $\alpha=1,2,3$

$$\frac{\partial\tilde{E}_{\mathrm{rec}}}{\partial\mathbf{r}_{\alpha i}}=\sum_{m_1=0}^{K_1-1}\sum_{m_2=0}^{K_2-1}\sum_{m_3=0}^{K_3-1}\partial Q/\partial\mathbf{r}_{\alpha i}(m_1,m_2,m_3)\cdot(\theta_{\mathrm{rec}}\star Q)(m_1,m_2,m_3). \tag{4.9}$$

Finally, the approximate reciprocal stress tensor Π_{rec} is obtained by substituting the approximate structure factors from Eq. (4.5) into Eq. (2.7).

In our implementation of this algorithm for approximating E_{rec}, $\partial E_{\mathrm{rec}}/\partial\mathbf{r}_{\alpha i}$, and Π_{rec} at each step of a molecular dynamics simulation, we first fill the array Q, using the coefficients $M_n(u_{ki}-j)$ for $i=1,\ldots,N$, $k=1,2,3$ and $j=0,\ldots,n$ computed from the current scaled fractional coordinates of the particles. Next, Q is transformed in place using the inverse 3DFFT. Using the transformed Q array as well as B, approximate expressions for E_{rec} and Π_{rec} are computed using Eqs. (4.7) and (2.7). At the same time the transformed Q array is overwritten by the product of itself with the arrays C and B defined in Eqs. (3.9) and (4.8). This new array is then transformed in place by the forward 3DFFT to arrive at the convolution $\theta_{\mathrm{rec}}\star Q$. Finally, this is multiplied by the quantities $\partial Q/\partial\mathbf{r}_{\alpha i}$ for $i=1,\ldots,N$ and $\alpha=1,2,3$, and summed to give the approximation (4.9) to the reciprocal forces $\partial E_{\mathrm{rec}}/\partial\mathbf{r}_{\alpha i}$. The quantities $\partial Q/\partial\mathbf{r}_{\alpha i}$ are computed on the fly as the sum is accumulated, using Eq. (4.2) to get analytic derivatives of the coefficients $M_n(u_{ki}-j)$.

The previous PME algorithm[24] obeyed Newton's 2nd Law to machine precision, since the reciprocal sum forces were interpolated symmetrically, while not conserving energy, since the energies and forces were approximated separately. In contrast, the current algorithm conserves energy, but not momentum. For example, the sum of the electrostatic forces on the atoms is not zero, but rather is a random quantity of the order of the rms error in the force. This leads to a kind of slow Brownian motion of the center of mass. This artifact can be avoided by removing the average net force from each atom at each step of the simulation, which does not affect the accuracy or the rms energy fluctuations. This was done for all the tests run below. It remains to be seen if

strict conservation of momentum or of energy is more important. To date we have not devised an algorithm that achieves both.

V. DISPERSION LATTICE SUMS

In this section, we briefly describe how the B-spline PME approach derived above for Coulombic lattice sums extends to dispersion interactions. For now, we assume geometric mean combining rules, and then show how to modify the approach for Lorentz–Bertholot combining rules. Thus, as above, given N particles at positions $\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_N$ within the unit cell U, assume for each particle i there is a positive constant λ_i so that the dispersion energy for a pair of particles i,j is $-\lambda_i\lambda_j/|\mathbf{r}_i-\mathbf{r}_j|^6$. Then we can write the dispersion energy of U as

$$E_6(\mathbf{r}_1,\ldots,\mathbf{r}_N)=-\frac{1}{2}{\sum_{\mathbf{n}}}' \sum_i \sum_j \frac{\lambda_i\lambda_j}{|\mathbf{r}_i-\mathbf{r}_j+\mathbf{n}|^6}, \tag{5.1}$$

where the outer sum is over the vectors $\mathbf{n}=n_1\mathbf{a}_1+n_2\mathbf{a}_2+n_3\mathbf{a}_3$, the prime indicating that terms with $i=j$ and $\mathbf{n}=0$ are omitted. Unlike the Coulombic lattice sum in Eq. (2.1), the outer sum in Eq. (5.1) converges absolutely. In Appendix A (see also Refs. 40, 23, and 42) we show that it can be re-written as a sum of direct, reciprocal, and self energies, analogous to the Ewald sum for Coulombic interactions. (The correction energy for molecular systems is handled analogously to that for Coulombic systems.) Here we focus on the reciprocal dispersion energy, which can be written

$$\begin{aligned} E_{6,\mathrm{rec}}=&-\frac{\pi^{3/2}}{3V}\sum_{\mathbf{m}} S(\mathbf{m})S(-\mathbf{m}) \\ &\cdot[(\beta^3/2-\beta\pi^2\mathbf{m}^2)\exp(-\pi^2\mathbf{m}^2/\beta^2) \\ &+\pi^{7/2}m^3\mathrm{erfc}(\pi m/\beta)], \end{aligned} \tag{5.2}$$

where $m=|\mathbf{m}|$, and the structure factors are now given by

$$\begin{aligned} S(\mathbf{m})&=\sum_{j=1}^{N}\lambda_j\exp(2\pi i\mathbf{m}\cdot\mathbf{r}_j) \\ &=\sum_{j=1}^{N}\lambda_j\exp[2\pi i(m_1s_{1j}+m_2s_{2j}+m_3s_{3j})]. \end{aligned} \tag{5.3}$$

From here the development is very similar to that for the Coulombic reciprocal sum, with the array Q replaced by an analogous array Λ, which is determined from the dispersion coefficients λ_j using the same B-spline coefficient expansion as in Eq. (4.6). The exponential weighting terms from Eq. (4.7) are replaced by the more complex weighting terms from Eq. (5.2). Formulas for the stress tensor are adapted from those in Karasawa and Goddard,[42] and atomic forces are obtained from the derivatives $\partial\Lambda/\partial\mathbf{r}_{\alpha i}$, which are computed analytically, as are $\partial Q/\partial\mathbf{r}_{\alpha i}$

When Lorentz–Bertholot combining rules hold, we can follow a suggestion of Perram *et al.*,[23] and consider the expansion

$$\lambda_{ij}=\sqrt{\varepsilon_i}\sqrt{\varepsilon_j}(r_{0i}+r_{0j})^6=\sum_{k=0}^{6}\frac{6!}{k!(6-k)!}\lambda_{k,i}\lambda_{6-k,j}, \tag{5.4}$$

where $\lambda_{k,i}=\sqrt{\varepsilon_i}r_{0,i}^k$. Using this expansion, the reciprocal dispersion energy can be written as the sum of seven terms, each of which resembles Eq. (5.2), and to which the above PME methodology can be applied.

Note that it is not possible to apply this dispersion approximation to the Amber3a force field, due to the complex way that 6–12 vs 10–12 interactions are implemented. However, it has been implemented into the new Amber force field, which has no explicit 10–12 interactions.

VI. RESULTS: ACCURACY AND EFFICIENCY OF THE PME ALGORITHM

The total energy E is invariant to the constant β, which controls the relative rate of convergence of the direct and reciprocal sums, E_{dir} and E_{rec} in Eqs. (2.3) and (2.4), respectively (the computation of E_{corr} is of order N and so is not discussed further). Increasing β causes E_{dir} to converge more rapidly, at the expense of slower convergence in E_{rec}. In order to bound the error due to cutoff in the direct sum, it is necessary that $\mathrm{erfc}(\beta r)/r$ be less than a small tolerance ϵ_{dir} for r greater than the cutoff. Similarly, to bound the error due to cutoff in the reciprocal sum, it is necessary that $\exp(-\pi^2|\mathbf{m}|^2/\beta^2)/|\mathbf{m}|^2$ be less than a small tolerance ϵ_{rec} for reciprocal vectors $\mathbf{m}=m_1\mathbf{a}_1^*+m_2\mathbf{a}_2^*+m_3\mathbf{a}_3^*$ outside the cutoff.

Consider a series of cubic boxes with increasing box length L. Traditionally β is chosen so that in the direct sum E_{dir} all the interactions beyond the nonmasked minimum image pairs can be neglected. From the condition that $\mathrm{erfc}(\beta L/2)$ be constant, β must vary inversely with the box length. Since the length of the reciprocal basis vectors $\mathbf{a}_\alpha^*$ also varies inversely with box length, the number of reciprocal vectors $\mathbf{m}$ satisfying the condition $\exp(-\pi^2|\mathbf{m}|^2/\beta^2)/|\mathbf{m}|^2<\epsilon_{\mathrm{rec}}$ is independent of system size and the reciprocal sum E_{rec} is well-approximated by the sum over several hundred reciprocal vectors $\mathbf{m}$, independently of the system size, leading to an order N algorithm for the reciprocal sum. However, the direct sum is an order N^2 computation, which is prohibitive for macromolecular systems.

Conversely, choosing a fixed β independent of the box length should allow a fixed size cutoff (e.g., 9 Å) in the direct sum E_{dir}, thereby reducing its computation to order N. On the other hand, the requirement on the reciprocal space tolerance now forces the number of reciprocal vectors and thus structure factors to grow as L^3. The straightforward implementation of Eq. (2.4) then leads to an order N^2 calculation of the reciprocal sum. In the PME method, however, the structure factors are interpolated on a grid, thus reducing the calculation to order $N\log(N)$ using the 3DFFT.

The original PME method, which relied on Lagrangian interpolation, was quite efficient in attaining low-to-moderate accuracies. However, it proved difficult to achieve high accuracies without resorting to excessively high grid densities. High order Lagrangian interpolation is known to be numerically unstable. In contrast, B-splines are well be-

TABLE I. Comparison of the relative error in the reciprocal potential energy and reciprocal rms force error of the PME method using Lagrangian interpolation vs B-spline interpolation. The relative error in the reciprocal potential energy is defined as: $|E_{\text{recip,full}}-E_{\text{recip,PME}}|/|E_{\text{total,full}}|$, with $|E_{\text{total,full}}|=|E_{\text{direct,full}}+E_{\text{recip,full}}|$. The reciprocal rms force error is defined as $([\Sigma_{i=1}^{3N}(F_{\text{recip,full,i}}-F_{\text{recip,PME,i}})^2]/\Sigma_{i=1}^{3N}F_{\text{total,full,i}}^2)^{1/2}$ with $F_{\text{total,full,i}}=F_{\text{direct,full,i}}+F_{\text{recip,full,i}}$. The subscript ''recip'' denotes terms in reciprocal space, subscript ''direct'' terms in direct space, and subscript ''full'' terms either in direct or reciprocal space calculated until their individual contributions are less than the machine precision.

Grid size	Interp. order	Lagrange PME rel. err. energy	B-spline PME rel. err. energy	Lagrange PME rel. err. force	B-spline PME rel. err. force
40	4	2.3×10^{-4}	4.3×10^{-5}	2.0×10^{-3}	4.3×10^{-4}
40	6	7.4×10^{-5}	3.8×10^{-6}	4.0×10^{-4}	2.2×10^{-5}
40	8	2.9×10^{-5}	5.5×10^{-7}	1.1×10^{-4}	2.0×10^{-6}
40	10	1.3×10^{-5}	1.2×10^{-7}	4.1×10^{-5}	2.9×10^{-7}
54	4	7.8×10^{-5}	1.0×10^{-5}	5.0×10^{-4}	9.5×10^{-5}
54	6	1.4×10^{-5}	3.6×10^{-7}	5.8×10^{-5}	2.2×10^{-6}
54	8	3.5×10^{-6}	2.1×10^{-8}	1.0×10^{-5}	9.0×10^{-8}
54	10	1.0×10^{-6}	1.7×10^{-9}	2.5×10^{-6}	5.5×10^{-9}
80	4	1.9×10^{-5}	2.1×10^{-6}	1.2×10^{-4}	2.3×10^{-5}
80	6	1.6×10^{-6}	2.6×10^{-8}	6.3×10^{-6}	2.1×10^{-7}
80	8	2.1×10^{-7}	5.2×10^{-10}	5.3×10^{-7}	3.1×10^{-9}
80	10	4.3×10^{-8}	1.4×10^{-11}	6.6×10^{-8}	6.7×10^{-11}

haved numerically. In addition, as noted above, the error bounds in approximating trigonometric functions are better for the Euler spline than for polynomial interpolation. To test if the new method is more accurate in practice, we compared the reciprocal sum relative errors in energy as well as forces due to the current versus the former PME method. A 40 Å box of 2038 TIP3P waters was prepared and equilibrated. The value of the Ewald coefficient β was set to approximately 0.35 Å^{-1}, corresponding to ''medium'' accuracy in the direct space forces (rms force error of 4×10^{-4}, see below). The accuracy in the reciprocal sum as a function of grid size and interpolation order for the two methods is displayed in Table I. The new method is substantially more accurate. Other values of β resulted in comparable improvements (results not shown). For comparison, the traditional Ewald sum, using the parameters recommended in Allen and Tildesley[51] (minimum image cutoff, $\beta=5/L$, 100–200 reciprocal vectors), leads to rms force errors of about 6×10^{-4}. Note that even for 1 Å grids the PME method yields comparable or more accurate Ewald sums than the traditional method.

In order to prove that the PME algorithm is of order $N\log(N)$ in practice we need to show that, given a fixed β, which should allow a fixed cutoff in the direct sum, we can choose a grid density and an interpolation order which result in system size independent bounds on the errors in the reciprocal sum, i.e., there are no hidden size dependencies in the required grid densities and/or interpolation orders. A series of TIP3P water boxes, from 20 to 80 Å in size, was prepared and equilibrated. We calculated rms force errors (defined in the caption to Fig. 1) using a direct sum with a 9 Å cutoff and either the B-spline PME reciprocal sum approximation, or for comparison the conventional reciprocal sum with a finite cutoff in $|\mathbf{m}|$. The value of the Ewald coefficient β was set to give low, medium, or high accuracy in the direct sum, using the criterion that $\text{erfc}(\beta r)/r<\epsilon_{\text{dir}}$, for $r>9$ Å, where ϵ_{dir} was set to 10^{-4}, 10^{-6}, and 10^{-8} for low, medium, and high accuracy. The resulting values of β were approximately 0.26, 0.35, and 0.42 Å^{-1}, respectively. The same value of β was then used when considering low, medium, or high accuracy reciprocal sums, either conventional Ewald or PME based.

To choose the reciprocal space cutoff for the conventional Ewald sum, we calculated the rms error in the reciprocal force for the 40 Å box and adjusted the cutoff until the error was slightly smaller than the corresponding error in the direct sum. The resulting cutoffs were approximately 0.41, 0.59, and 0.72 Å^{-1} for the low, medium, and high accuracy cases, respectively. These cutoff values were then applied to the remaining systems. For the B-spline PME, we used cubic B-splines for the low accuracy case, 5th degree B-splines for the medium accuracy test, and 7th degree B-splines for the high-accuracy case. We then adjusted the grid size so that the rms error in the PME reciprocal force for the 40 Å box was slightly less than the corresponding error in the conventional reciprocal sum with the above cutoffs. The resulting grid densities were 2 Å ($20\times20\times20$ grid), 1 Å and $\frac{2}{3}$ Å for the low, medium, and high accuracy cases, respectively. These same interpolation orders and grid densities were applied to the remaining systems.

The rms relative force errors for these systems are plotted in Fig. 1. Note that at all levels of accuracy the errors are almost constant (i.e., they do not increase as a function of system size). In particular, we see that the reciprocal sum forces can be approximated to within a given level of accuracy, independent of system size, by choosing appropriate values of the interpolation order and grid density.

Next, we investigated the variation in the PME force errors along a molecular dynamics trajectory. The system, consisting of 2038 TIP3P water molecules in a 40 Å box, was propagated for 100 steps of microcanonical ensemble molecular dynamics, with a 1fs time step, using the above prescriptions for PME approximation at low, medium, and high accuracy. At each step the ''exact'' Ewald sum energy, forces, and electrostatic virial were computed and compared with the PME approximations to the same quantities. We

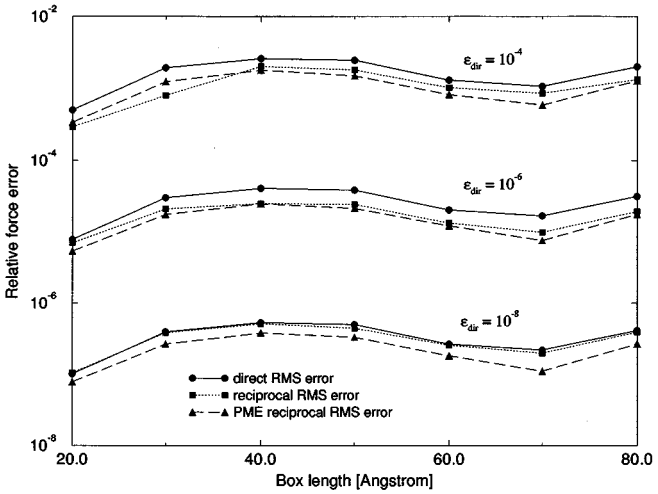

FIG. 1. rms force errors as a function of system size. The three groups of curves correspond to three different choices of β as described in the text. The direct rms force error is defined as the error due to the truncation of the interactions in direct space: $([\Sigma_{i=1}^{3N}(F_{\mathrm{direct,cutoff,i}}-F_{\mathrm{direct,full,i}})^2]/\Sigma_{i=1}^{3N}F_{\mathrm{total,full,i}}^2)^{1/2}$, with $F_{\mathrm{total,full,i}}=F_{\mathrm{direct,full,i}}+F_{\mathrm{recip,full,i}}$. Similarly, the rms force error due to a cutoff in the conventional reciprocal sum is defined as: $([\Sigma_{i=1}^{3N}(F_{\mathrm{recip,cutoff,i}}-F_{\mathrm{recip,full,i}})^2]/\Sigma_{i=1}^{3N}F_{\mathrm{total,full,i}}^2)^{1/2}$. The subscripts "direct", "recip", and "full" are explained in Table 1. The subscript "cutoff" refers to terms in direct space using a 9 Å cutoff or terms in reciprocal space using a cutoff. Finally, the error due to the PME approximation of the reciprocal sum is $([\Sigma_{i=1}^{3N}(F_{\mathrm{recip,PME,i}}-F_{\mathrm{recip,full,i}})^2]/\Sigma_{i=1}^{3N}F_{\mathrm{total,full,i}}^2)^{1/2}$.

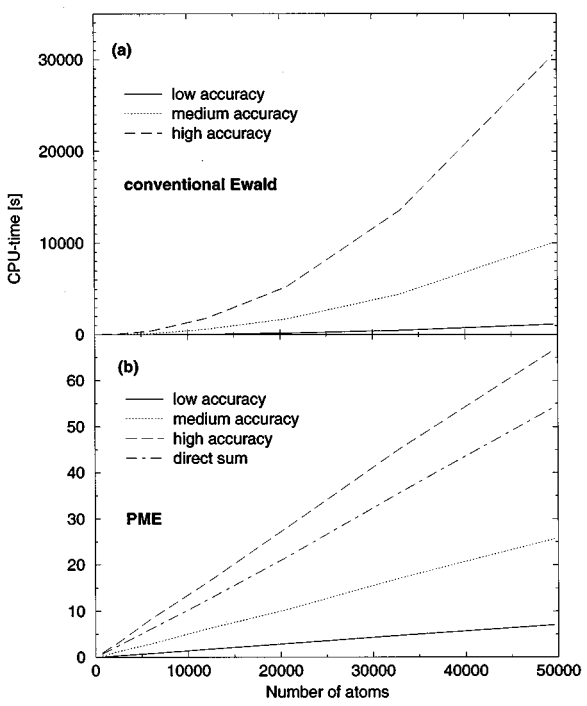

FIG. 2. Comparison of the CPU-times needed for the conventional vs PME calculation of the reciprocal sum. (a) CPU-time for the conventional calculation of the reciprocal sum as a function of system size at the three levels of accuracies described in the text; (b) CPU-time required for the same level of accuracies in the PME approach. For comparison the CPU-time for the calculation of the direct sum using a 9 Å cutoff is also shown.

found that the rms force error is remarkably constant along the trajectory (results not shown). For all three accuracy levels, the maximum relative atomic force error was approximately one order of magnitude greater than the rms force error, whereas the relative error in the electrostatic energy was considerably smaller than the rms force error.

Although we have not found any useful *a priori* estimates of the accuracy in the PME approximation, an inexpensive estimator can be arrived at as follows. The exact electrostatic energy and virial trace sum to zero.[44] If the electrostatic energy is better estimated than the virial, $|tr(Virial_{\mathrm{approx}})+E_{\mathrm{approx}}|/|E_{\mathrm{approx}}|$ provides an estimate of the relative error in the virial trace. In turn, one may hope to use this as a rough estimator of the rms force error, since the virial calculation involves the forces. Calculating these quantities for the above-mentioned trajectories showed that the estimated virial error usually stayed within an order of magnitude of the rms force error. Occasionally, however, the error in the virial trace dropped well below the rms force error, suggesting that the time average-estimated error is more representative.

In Fig. 2 we have plotted the CPU time (on an SGI R4400) required to perform the conventional and PME reciprocal sums used in Fig. 1 as a function of system size. The conventional reciprocal sum was computed using the algorithm of Smith and Fincham in the program MDMPOL from CCP5. Note that the CPU time needed for this grows as the square of the system size, as predicted. Moreover, the cost is prohibitive for large systems, even for the low-accuracy case (e.g., over 1100 s per time step for 50,000 atoms). On the other hand, the time needed for the PME method grows linearly with system size over this range of systems. This is due to the fact that the charge grid-interpolation routines, which scale linearly, are more expensive than the fast Fourier transform (FFT) routines, which scale as $N\log(N)$, over this range of system sizes. For comparison, we have also plotted the CPU time needed for the direct sum with a 9 Å cutoff, including the van der Waals interactions. The erfc function was tabled using a cubic spline. The PME method was faster than the conventional Ewald sum for all systems and accuracies tested here. For sufficiently small systems, the conventional Ewald sum is more efficient. On the SGI R4400, the break-even point was around 600 atoms. On the Cray YMP, for comparison, the break-even point is close to 900 atoms. For a 20,000 atom system on the SGI R4400, the PME method took approximately 3, 10, and 28 s for low, medium, and high accuracy, whereas the corresponding times for conventional Ewald reciprocal sums were approximately 180, 1760, and 5300 s, respectively.

Next we tried, for a fixed accuracy, to find the optimum combination of direct-sum cutoff, interpolation order, and grid size in the PME method. The above 40 Å water box was used and the rms relative error in the total force (direct plus PME reciprocal) was fixed to 5×10^{-4}, comparable to the error in the traditional Ewald sum approach. The value of ϵ_{dir}, used to determine β, was set to 10^{-5}. The direct sum cutoff was varied between 6 and 10 Å, and the interpolation order between 4 and 6. For each cutoff and interpolation order, the minimum grid size, which was a multiple of powers of 2, 3 and 5, and for which the rms force error fell below 5×10^{-4}, was determined (the FFT requires the former con-

TABLE II. Timing results for the 18.77 Å water box for different choices of grid density and interpolation order compared with timing results for standard Amber3a with the same cutoff. The numbers for the PME calculations give the total CPU-time for the direct space calculations (including the truncated dispersion interactions) plus the reciprocal space calculations. The total rms force error for the PME calculations was fixed at 5×10^{-4}. The numbers in parenthesis indicate the number of grid points in each of the three dimensions.

Cutoff (Å)	Amber3a time	Direct sum time	4th-order total time	5th-order total time	6th-order total time
6.0	1.38	1.66	5.75 (60)	4.75 (48)	4.81(40)
6.5	1.78	2.11	5.42 (54)	4.95 (45)	5.00(36)
7.0	2.28	2.64	5.09 (48)	5.06 (40)	5.42(32)
7.5	2.91	3.27	5.49 (45)	5.37 (36)	5.95(30)
8.0	3.59	4.06	5.84 (40)	5.94 (32)	6.61(27)
8.5	4.31	4.82	6.72 (40)	6.67 (30)	7.47(25)
9.0	5.14	5.84	7.36 (36)	7.61 (27)	8.20(24)
9.5	6.10	6.96	8.29 (32)	8.62 (25)	9.12(20)
10.0	7.24	8.16	9.32 (27)	9.68 (24)	10.42(20)

dition for speed). Timing results for direct sum as well as for direct plus reciprocal sums are compared in Table II with timings for Amber3a residue-based cutoff calculations. The optimal PME CPU time is achieved at the smallest cutoff and 5th order interpolation. Thus, it would seem that very short cutoffs are recommended. This result brings to mind the issue of neglect of long-range dispersion interactions.

Truncating the van der Waals interactions at 9 Å leads to an rms force error of only 5×10^{-5} in this system, which is less than the above rms force error in the Coulombic interactions. However, the computed (isotropic) system pressure as well as the potential energy are much more sensitive to the cutoff in the van der Waals interactions, since the individual terms $\mathbf{f}_{ij}\cdot\mathbf{r}_{ij}$ in the trace of the virial, which are due to dispersion forces are always negative, as are the neglected energy terms. For example, we find that including the full lattice sum for dispersion interactions drops the pressure by almost 300 atm compared to truncation at 9 Å. This result seems to be fairly independent of the system. The discrepancy increases to nearly 600 atm when a 7 Å cutoff on the dispersion interactions is used (in this case the rms force error due to truncation of the dispersions is about 2×10^{-4}). For comparison, the error in the pressure due to truncation of the r^{-12} repulsion terms is between 1 and 3 atm for this range of cutoff values.

In applying the PME method for dispersions to TIP3P waters we can use Eq. (5.3) to approximate the structure factors, which assumes geometric mean combination rules, since only the oxygens participate in van der Waals interactions. In Table III we give the total CPU time for the PME approximation applied to both dispersion and Coulombic interactions. At each value of the cutoff, we used the optimal grid size and interpolation order for Coulombic interactions from Table II, and investigated the effect of these two parameters on the PME approximation to dispersion interactions. The criterion we used for accuracy in the dispersion interactions was that the error in the computed isotropic pressure should be less than 10 atm, which was approximately the error due to the Coulombic PME method with the rms force error set to 5×10^{-4}. For each cutoff and interpolation order the minimum grid size which satisfies this accuracy criterion (and which was a multiple of powers of 2, 3, and 5) was determined. Note that the time for the direct sum has increased compared to that in Table II, since the expression for the dispersion direct sum is more complex than a simple r^{-6} dependency. The minimum time is now achieved with a 7 Å cutoff in the direct sum, 5th order interpolation with a 40 $\times40\times40$ grid for Coulombic interactions, and 4th order interpolation with a $30\times30\times30$ grid for the dispersion interactions. Note that the total time is less than that for the Amber3a residue-based cutoff calculation using a 10 Å cutoff.

Finally, we checked the generality of these accuracy results by examining the rms force error due to similar levels of PME approximation to electrostatics in three other systems; equilibrated structures from crystal simulations of BPTI[2] with 5304 atoms, and $p21$ *H*-ras (unpublished results) with 24,801 atoms, and from a solution simulation of a DNA dodecamer[10] with 9956 atoms, kindly provided by Tom Cheatham. For example, in Table II we see that a 9 Å cutoff in the direct sum with a tolerance ϵ_{dir} of 10^{-5} together with cubic interpolation on a $36\times36\times36$ grid gives an rms force error of less than 5×10^{-4} in the 40 Å water box system. The $p21$ system is of particular interest since it has a nonorthogonal unit cell, with parameters $a=b=40.3$ Å, $c=162.2$ Å, $\alpha=\beta=90°$, and $\gamma=120°$. Use of a 9 Å cutoff, a tolerance of 10^{-5}, cubic interpolation, and a grid of $36\times36\times144$ (the nearest combinations of powers of 2, 3, and 5 to the same grid density as above) lead to an rms force error of 4.99×10^{-4}. Proceeding similarly with the remaining table entries and the other two test systems leads to similar results.

TABLE III. Timing results for the same system under the same conditions as in Table II except that the dispersion interactions were calculated using the PME approximation.

Cutoff (Å)	Direct sum time	4th-order total time	5th-order total time	6th-order total time
6.0	1.94	8.39 (48)	7.82 (40)	8.12 (36)
6.5	2.46	7.13 (36)	7.37 (32)	8.06 (30)
7.0	3.10	6.87 (30)	7.48 (30)	8.01 (27)
7.5	3.82	7.01 (24)	7.37 (20)	8.17 (20)
8.0	4.75	7.49 (18)	7.89 (16)	8.82 (15)

TABLE IV. Energy fluctuations for a 40 Å water box as a function of time step, neighbor list update, grid density, and different treatments of long-range dispersion interactions. The quantity $\langle \triangle E^2 \rangle^{1/2} / \langle \triangle KE^2 \rangle^{1/2}$ refers to the fluctuations of the total energy divided by the fluctuations of the kinetic energy. The quantity $\langle \triangle E^2 \rangle^{1/2} / \langle E \rangle$ refers to the fluctuations of the total energy divided by the total energy.

Time step (fs)	List update (fs)	ϵ_{dir}	Nfft	Dispersion interactions	NfftD	$\langle \triangle E^2 \rangle^{1/2} / \langle \triangle KE^2 \rangle^{1/2}$	$\langle \triangle E^2 \rangle^{1/2} / \langle E \rangle$
2	2	5×10^{-7}	24	PME	16	1.0×10^{-2}	6.6×10^{-5}
1	1	5×10^{-7}	24	PME	16	2.4×10^{-3}	1.6×10^{-5}
0.5	0.5	5×10^{-7}	24	PME	16	6.5×10^{-4}	4.3×10^{-6}
2	2	5×10^{-6}	24	PME	16	1.0×10^{-2}	6.6×10^{-5}
1	1	5×10^{-6}	24	PME	16	4.3×10^{-3}	2.9×10^{-5}
0.5	0.5	5×10^{-6}	24	PME	16	3.1×10^{-3}	2.0×10^{-5}
2	2	5×10^{-7}	24	truncate	⋯	1.0×10^{-2}	6.6×10^{-5}
1	1	5×10^{-7}	24	truncate	⋯	4.7×10^{-3}	3.1×10^{-5}
0.5	0.5	5×10^{-7}	24	truncate	⋯	3.9×10^{-3}	2.6×10^{-5}
0.5	0.5	5×10^{-7}	16	PME	16	8.0×10^{-4}	5.3×10^{-6}
1	5	5×10^{-7}	24	PME	16	2.5×10^{-3}	1.7×10^{-5}
0.5	2.5	5×10^{-7}	24	PME	16	6.5×10^{-4}	4.3×10^{-6}
1	1	5×10^{-6}	16	truncate	⋯	4.2×10^{-3}	2.8×10^{-5}
1	5	5×10^{-6}	16	truncate	⋯	7.9×10^{-3}	5.2×10^{-5}
0.5	0.5	5×10^{-8}	24	PME	16	6.1×10^{-4}	4.0×10^{-6}

The worst case was an rms force error of 7×10^{-4}; it occurred in the $p21$ system with a 6 Å cutoff, cubic interpolation, and a grid of $60\times60\times240$. Most cases lead to rms force errors below 5×10^{-4}.

VII. ENERGY CONSERVATION

The rms fluctuation in total system energy in microcanonical ensemble simulations using the Verlet integration scheme should be proportional to δt^2, where δt is the elementary time step.[51] Truncation of electrostatic interactions is known to lead to instabilities in the total system energy. This problem can be alleviated through the use of switching functions,[52] which, however, introduce artifacts,[13,15,17] or more realistically, by accounting smoothly for the long-range interactions.[53] In this section, we investigate the energy conservation of the PME algorithm. All simulations were run using a version of Amber3a, modified to perform the PME algorithm on electrostatics and dispersion interactions.

The first system we tested was a box of 216 TIP3P waters. After equilibrating the system to 300 K and 1 bar (box size 18.77 Å), we ran a series of 5 ps microcanonical ensemble simulations. In all cases the direct sum was cutoff at 8 Å and cubic interpolation was used on the reciprocal sum for Coulombic and (where applicable) dispersion interactions. Bond lengths were constrained using the matrix method,[54] using a tolerance of 10^{-10}. We varied the time step, the list update, the tolerance on the direct space cutoff, and the grid density for the Coulombic PME method. Finally, we investigated the effect of the truncation of dispersion interactions. The fluctuations in total energy relative to the fluctuations in the kinetic energy and to the average total energy are shown in Table IV. Note that the average total energy dropped from about -1660 kcals without long-range dispersion interactions to about -1680 kcals with the interactions.

Note that when a tolerance ϵ_{dir} of 5×10^{-7} is used and long-range dispersion interactions are not truncated, the rms energy fluctuations do appear to decrease quadratically with the time step. In contrast, use of larger values of ϵ_{dir} and truncation of the dispersion interactions both lead to larger rms energy fluctuations that improve only slightly as the time step decreases. Decreasing the grid density from $24\times24\times24$ to $16\times16\times16$ increases the rms relative force error from 3×10^{-4} to 2×10^{-3}. However, it has only a weak effect on energy fluctuations. (Compare line ten of Table IV with line three.) Next, when long-range dispersion interactions are included and $\epsilon_{\text{dir}}=5\times10^{-7}$, updating the Verlet list every five steps has a minimal effect on rms energy fluctuations (compare lines 11 and 12 of Table IV with lines two and three), whereas the effect is greater when these interactions are truncated (compare line 14 with line 13). Finally, tightening the tolerance ϵ_{dir} from 5×10^{-6} to 5×10^{-7} has a pronounced effect on energy conservation for small time steps, but decreasing it further to 5×10^{-8} has only a weak effect (compare lines three, six and fifteen).

To assess the generality of these results, we next performed similar simulations on the DNA dodecamer in solution with sodium counterions.[10] This system was prepared using the new Amber force field, without explicit 10–12 terms, which allowed us to test the effect of including the long-range dispersion interactions. The unit cell was orthogonal, with dimensions $57.2\times42.3\times41.6$ Å, and contained 9956 atoms. The grid for the Coulombic PME was $54\times40\times40$, while that for the dispersion PME was $36\times24\times24$. Bonds involving hydrogen were constrained in length using the matrix method as above, while those involving only heavy atoms were allowed to vibrate. The rest of the protocol was identical to the water simulations above. The fluctuations in total energy relative to the fluctuations in the kinetic energy and to the average total energy are shown in Table V. As in the water simulations, the average total energy dropped by about 1% when the long-range dispersion interactions are included. The results are very similar to those of the water box simulations.

TABLE V. Energy fluctuations for the DNA system as a function of time step, neighbor list update, and different treatments of long-range dispersion interactions.

Time step (fs)	List update (fs)	ϵ_{dir}	Dispersion interactions	$\langle \triangle E^2 \rangle^{1/2} / \langle \triangle KE^2 \rangle^{1/2}$	$\langle \triangle E^2 \rangle^{1/2} / \langle E \rangle$
2	2	5×10^{-7}	PME	1.7×10^{-2}	2.4×10^{-5}
1	1	5×10^{-7}	PME	4.4×10^{-3}	6.6×10^{-6}
0.5	0.5	5×10^{-7}	PME	1.1×10^{-3}	1.6×10^{-6}
1	1	5×10^{-7}	truncate	8.6×10^{-3}	1.3×10^{-5}
0.5	0.5	5×10^{-7}	truncate	8.8×10^{-3}	1.3×10^{-5}
0.5	2.5	5×10^{-7}	PME	1.2×10^{-3}	1.8×10^{-6}
1	5	5×10^{-7}	truncate	1.9×10^{-2}	2.9×10^{-5}

VIII. RESULTS OF WATER MD SIMULATIONS

Although the results above demonstrate that the PME method provides a reliably accurate approximation to the Ewald sum and leads to excellent energy conservation, it is possible that long simulations using this method could lead to artifactual behavior. To test for this we next performed a series of molecular dynamic (MD) simulations of pure water using the new method, recording the average internal energy and density as well as the self diffusion coefficient. The simulations were performed with a version of the Amber3a code modified to carry out the conventional Ewald sum using the algorithm of Smith and Fincham as mentioned above, as well as the *B*-spline PME algorithm for Coulombic and/or dispersion interactions. Two water models were used, TIP3P[55] and SPC/E.[56] The simulations were run at 300 K and 1 bar using the Berendsen coupling technique.[57] Bonds were constrained as above. A 1 fs time step was used and the simulations were all run for at least 100 ps after equilibration. The diffusion coefficient was obtained as follows. For each molecule i let $d_i(t)$ denote the distance traveled by its center of mass from the beginning of the simulation to time t. Let $D(t)=\Sigma_{i=1}^{N} d_i^2(t)/(6Nt)$. Then $D(t)$ should approach the diffusion constant as t tends to infinity. We took the average of $D(t)$ over the last 10 ps of the simulation. Results are shown in Table VI. For the TIP3P model Jorgensen *et al.*[55] reported a density of 0.982 g/cm^3 and an average potential energy of -9.86 kcal/mol from a Monte Carlo study of 125 water molecules using a 7.5 Å residue-based cutoff. We obtained 0.99 g/cm^3 and -9.7 kcal/mol using a residue-based cutoff of 8.5 Å. They did not report a diffusion constant for TIP3P. Berendsen *et al.*[56] report for the SPC/E model a density of 0.993 g/cm^3, an average potential energy of -11.14 kcal/mol (without the polarization correction) and a self-diffusion constant of 2.5×10^{-5} cm^2/s. Using an 8.5 Å residue-based cutoff, we obtain the same diffusion constant, a density of 1 g/cm^3 and an energy of -11.3 kcal/mol, i.e., an error of about 0.15 kcal/mol. Note that use of Ewald sums for Coulombic terms with cutoff of dispersions results in about 0.2 kcal/mol higher potential energy and a lowering of the density of about 2% for both models. Including the long-range dispersions brought both measures closer to the residue-based cutoff values. Given that we estimate the errors in the calculated diffusion constants to be about 5%–10%, the different methods give consistent results for the two models. Comparing the results for conventional Ewald summation to those using the PME algorithm, we find a close agreement between these two methods.

The calculation of the dielectric constant has proven to be sensitive to the treatment of long-range forces.[58,59] Therefore, as an additional test of the validity of the PME approach, we calculated the dielectric constant with the PME method as well as the conventional Ewald summation. A system of 216 SPC/E water molecules was simulated in an NVE ensemble with a time step of 1 fs at a density of 0.977 g/cm^3 and a temperature of 296 K. The length of the run was 2 ns in the case of the simulation with the PME method and

TABLE VI. Physical characteristics of different water simulations.

Water model	Num waters	Cutoff (Å)	Coulomb interactions	Dispersion interactions	E_{pot} (kcal/mol)	Density (g/cm^3)	Diff. const. (10^{-5}cm^2/s)
TIP3P	216	8.5	truncate	truncate	-9.7	0.99	5.3
TIP3P	216	8.0	Ewald	truncate	-9.5	0.97	5.1
TIP3P	216	9.0	PME	truncate	-9.5	0.97	5.1
TIP3P	216	8.0	PME	truncate	-9.5	0.97	5.1
TIP3P	2038	8.0	PME	PME	-9.6	0.98	5.8
SPC/E	216	8.5	truncate	truncate	-11.3	1.00	2.5
SPC/E	216	8.0	Ewald	truncate	-11.1	0.98	2.4
SPC/E	216	8.0	PME	truncate	-11.1	0.98	2.3
SPC/E	216	9.0	PME	truncate	-11.1	0.99	2.7
SPC/E	216	8.0	PME	PME	-11.2	1.00	2.3
SPC/E	216	9.0	PME	PME	-11.2	1.00	2.5
SPC/E	2038	9.0	PME	truncate	-11.1	0.99	2.6
SPC/E	2038	8.0	PME	PME	-11.2	1.00	2.6

1 ns in the case of the conventional Ewald summation.

In tinfoil boundary conditions, the dielectric constant ϵ can be calculated from the fluctuations of the total dipole moment $\mathbf{M}$ according to Ref. 60

$$\epsilon = 1 + 3y \frac{\langle \mathbf{M}^2 \rangle}{\mathbf{N}\mu^2} \tag{8.1}$$

with

$$y = \frac{4\pi\rho_N\mu^2}{9kT}, \tag{8.2}$$

where ρ_N denotes the number density, μ the dipole moment of an individual molecule, and N the number of molecules.

From the 2 ns simulation using the PME approach we obtained a value of $\epsilon = 64.1$. To estimate the standard error we broke the run into four pieces and calculated the dielectric constant for each of the pieces. The results were 56.3, 63.4, 73.8, and 63.0, yielding a standard error of ± 3.6. This error estimate should be treated with caution, since it assumes the four values are independent, which may not be true.[59]

Using the traditional Ewald summation approach, we found after simulating the system for 1 ns a dielectric constant of $\epsilon = 68.3$. Due to the shorter length of this simulation, we could not estimate a standard error. Note, however, that the calculated value is within the range of the four values obtained by the PME method.

In the long-time limit the average dipole moment of the simulation cell should be zero. To assess the quality of a simulation de Leeuw *et al.* proposed the quantity $\langle \mathbf{M} \rangle^2 / \langle \mathbf{M}^2 \rangle$. In our simulations we found values for this ration of 0.013, using PME, and 0.005, using conventional Ewald summation (the latter value was anomalously small at the end of the simulation and typically closer to that for the PME).

In two previous simulations, the dielectric constants of SPC/E water has been determined using the reaction field method to treat long-range forces.[59,61] Reddy and Berkowitz[61] found for $\rho = 1$ g/cm^3 and $T = 298$ K a value of $\epsilon = 70.7$, while Smith and van Gunsteren[59] found at $\rho = 0.976$ g/cm^3 and $T = 300$ K a value of 64.0. Comparing these data with the one obtained with the PME approach or conventional Ewald method, we conclude that, even though there are differences in the calculated dielectric constants, these differences are within the range of the estimated errors.

IX. SUMMARY AND CONCLUSIONS

In this article, we have introduced a modification of the previously developed PME method. The new method uses a fixed cutoff in the direct sum and uses B-spline interpolation of the reciprocal space structure factors onto a regular grid, permitting the use of fast Fourier transforms to efficiently calculate the reciprocal sum. We demonstrated that this method has several advantages. The new method is substantially more accurate than the original PME. Moreover, the accuracy can be improved at will by adjusting a few parameters. In common with the previous PME method, the new algorithm generalizes naturally to arbitrary unit cells, since the approximation is formulated in terms of fractional coordinates. The formulation in terms of structure factors provides for straightforward generalization to other pair potentials depending on inverse powers of $|\mathbf{r}|$, such as dispersion interactions. Furthermore, this formulation leads to approximations of the reciprocal virial tensor that involve negligible overhead. Finally, the use of B-splines allows us to obtain forces by analytic differentiation of the energies, and these forces are in turn smoothly varying with respect to particle position. As far as we are aware, this algorithm is the only fast algorithm for long-range forces that has this smoothness property. Moreover, since the forces are not interpolated as in the previous method, we substantially reduce the memory requirements. Finally, we demonstrated that the method scales as $N \log(N)$ with respect to system size, which makes it feasible to treat biomolecular systems of 20 000 or more atoms via Ewald summation. In Fig. 2 we show that for low or medium accuracy, the CPU cost for the reciprocal sum is less than that of the direct sum for a 9 Å cutoff. In Table III we show that, at least for water boxes, the PME method makes it possible to compute the Ewald sum for dispersion as well as Coulombic interactions at a level of accuracy comparable to traditional Ewald summation in less time than standard truncation methods with a 10 Å cutoff. The method vectorizes efficiently, achieving 146 megaflops on a Cray Y-MP. Parallelization is natural on shared memory multiprocessors such as the SGI Challenge systems. An efficient parallel version has been developed for the Cray T3D using a distributed 3DFFT.[62]

Note added in proof. After this manuscript was submitted, H. G. Petersen[65] published a detailed comparison of the PME method with the optimized conventional Ewald method. He derives accurate expressions for the RMS force errors. Based on these expressions he compares the two methods at the same level of accuracy. He concludes that the PME method is more efficient for system sizes exceeding 10 000 atoms. This estimate is substantially more conservative than ours. The discrepancy is probably due mostly to the fact that we use an order N^2 algorithm for the Ewald sum while the optimized Ewald method scales like $N^{3/2}$. Secondly, our FFT implementation, based on the 1D FFT of Swarztrauber (available from netlib), is substantially more efficient than the FFT used in Petersen's work. Furthermore, our implementation has the advantage that the grid dimensions can be a product of powers of 2, 3, and 5 which allows more flexibility in the grid density. Finally, the approximation by B-splines described above allows a coarser grid to be used at the same level of accuracy than the Lagrangian interpolation.

ACKNOWLEDGMENTS

Supercomputing time was provided by the National Cancer Institute Biomedical Supercomputing Center, the Pittsburgh Supercomputing Center, and the North Carolina Supercomputing Center. This work was supported by the Office of Naval Research. H.L. was supported by an NIEHS/NIH Intramural Research Fellowship. T.D. would like to acknowledge Kristin Meier for suggesting the use of B-splines in this problem.

APPENDIX A: LATTICE SUMS FOR INVERSE POWERS OF DISTANCE

In this appendix we derive the lattice summation formulas given in the text for the case of dispersion interactions as well as for Coulombic interactions, following the approach in Smith.[41] These lattice sums do not in general converge absolutely, so we need to specify the asymptotic order of summation, corresponding to the asymptotic shape of the finite crystal made up of the union of lattice translations of the unit cell U. Let Λ denote the set of all lattice vectors $\mathbf{n}=n_1\mathbf{a}_1+n_2\mathbf{a}_2+n_3\mathbf{a}_3$. In order to describe the order of summation in R^3, we introduce a closed, bounded region P, centered on the origin (e.g., the unit ball or the unit cube), and for positive integers K let $P_K(\Lambda)$ denote the set of lattice vectors $\mathbf{n}$ such that $\mathbf{n}/K\in P$. Given N points $\mathbf{r}_1,...,\mathbf{r}_N$ in the unit cell U, and real constants $C(i,j)$, $1\leqslant i,\ j\leqslant N$, we consider sums of the form

$$E_p(\mathbf{r}_1,\ldots,\mathbf{r}_N)=\lim_{K\to\infty}\frac{1}{2}{\sum_{\mathbf{n}\in P_K(\Lambda)}}'\sum_i\sum_j\frac{C(i,j)}{|\mathbf{r}_i-\mathbf{r}_j+\mathbf{n}|^p}, \tag{A1}$$

where, as in Eq. (2.1), the prime on the outer sum denotes that terms with $i=j$ and $\mathbf{n}=0$ are omitted.

We begin by deriving some identities for the inverse powers $1/|\mathbf{r}|^p$, $p>0$, where $\mathbf{r}$ is any nonzero vector in R^3. The derivation depends on the following well-known results

$$\Gamma(z)=\int_0^\infty t^{z-1}\exp(-t)dt=\lambda^z\int_0^\infty t^{z-1}\exp(-\lambda t)dt, \tag{A2}$$

where $\Gamma(z)$ is the Euler gamma function, and

$$\exp(-a^2w^2)=\frac{\sqrt{\pi}}{a}\int_0^\infty\exp(-\pi^2u^2/a^2)\times\exp(-2\pi iuw)du, \tag{A3}$$

which is the Fourier integral expansion of the Gaussian. In Eq. (A2), given a three dimensional vector $\mathbf{r}$, we substitute $\lambda=|\mathbf{r}|^2=r^2$, where $r=|\mathbf{r}|$ and $z=p/2$. For arbitrary positive number β we then have

$$\frac{\Gamma(p/2)}{r^p}=\int_0^{\beta^2}t^{p/2-1}\exp(-r^2t)dt+\int_{\beta^2}^\infty t^{p/2-1}\times\exp(-r^2t)dt=I_p+II_p. \tag{A4}$$

To evaluate II_p, we substitute t by s, with $r^2t=s^2$. To evaluate I_p, we write $r^2=x^2+y^2+z^2$ and apply Eq. (A3) in all three dimensions, substituting w with x,y,z, respectively, and a^2 with t. At this point, we need to consider the reciprocal unit cell U^* made up of the points $\mathbf{u}$ in R^3 such that $-1/2\leqslant\mathbf{a}_i\cdot\mathbf{u}\leqslant 1/2$. Note that R^3 can be decomposed as the union of the point sets $U^*+\mathbf{m}$, over all reciprocal vectors $\mathbf{m}$. Changing the order of integration to integrate over t first, and substituting t with s, where $\pi^2u^2/t=s^2$, we arrive at

$$\begin{aligned}\frac{1}{r^p}&=\pi^{3/2}\beta^{p-3}\int_{R^3}f_p(\pi|\mathbf{u}|/\beta)\exp(-2\pi i\mathbf{u}\cdot\mathbf{r})d^3\mathbf{u}+\frac{g_p(\beta r)}{r^p}\\&=\pi^{3/2}\beta^{p-3}\sum_m\int_{U^*}f_p(\pi|\mathbf{v}+\mathbf{m}|/\beta)\times\exp[-2\pi i(\mathbf{v}+\mathbf{m})\cdot\mathbf{r}]d^3\mathbf{v}+\frac{g_p(\beta r)}{r^p},\end{aligned} \tag{A5}$$

where for positive numbers x, $f_p(x)$ and $g_p(x)$ are defined by

$$f_p(x)=\frac{2x^{p-3}}{\Gamma(p/2)}\int_x^\infty s^{2-p}\exp(-s^2)ds \tag{A6}$$

and

$$g_p(x)=\frac{2}{\Gamma(p/2)}\int_x^\infty s^{p-1}\exp(-s^2)ds. \tag{A7}$$

In particular, for the cases $p=1$ or 6, we have $f_1(x)=\exp(-x^2)/(\sqrt{\pi}x^2)$ and $g_1(x)=\mathrm{erfc}(x)$, $f_6(x)=1/3[(1-2x^2)\exp(-x^2)+2x^3\sqrt{\pi}\,\mathrm{erfc}(x)]$ and $g_6(x)=\exp(-x^2)(1+x^2+x^4/2)$. Note that $f_p(x)$ and $g_p(x)$ are smooth positive functions of x for $x>0$, $g_p(x)$ is bounded as $x\to 0$ for $p\geqslant 1$, while $f_p(x)$ is bounded as $x\to 0$ for $p>3$, but $f_p(x)\to\infty$ as $x\to 0$ for $1\leqslant p\leqslant 3$. The above integral representation does, however, converge absolutely for $p\geqslant 1$.

For $\mathbf{v}\in U^*$ we write $\mathbf{v}=w_1\mathbf{a}_1^*+w_2\mathbf{a}_2^*+w_3\mathbf{a}_3^*$ where $w_k=\mathbf{v}\cdot\mathbf{a}_k$, $k=1,2,3$. For $\mathbf{r}\in U$ and any lattice vector $\mathbf{n}=n_1\mathbf{a}_1+n_2\mathbf{a}_2+n_3\mathbf{a}_3$, such that $\mathbf{r}+\mathbf{n}\neq 0$, we apply Eq. (A5), changing variables in the integrals over U^*, to write

$$\begin{aligned}\frac{1}{|\mathbf{r}+\mathbf{n}|^p}=\frac{\pi^{3/2}\beta^{p-3}}{V}\sum_{\mathbf{m}}&\exp(-2\pi i\mathbf{m}\cdot\mathbf{r})\\&\times\int_{-1/2}^{1/2}\int_{-1/2}^{1/2}\int_{-1/2}^{1/2}h_{p,\mathbf{m},\mathbf{r}}(w_1,w_2,w_3)\\&\cdot\exp[-2\pi i(w_1n_1+w_2n_2+w_3n_3)]\\&\times dw_1dw_2dw_3+\frac{g_p(\beta|\mathbf{r}+\mathbf{n}|)}{|\mathbf{r}+\mathbf{n}|^p},\end{aligned} \tag{A8}$$

where $h_{p,\mathbf{m},\mathbf{r}}(w_1,w_2,w_3)=f_p(\pi|\mathbf{v}+\mathbf{m}|/\beta)\exp(-2\pi i\mathbf{v}\cdot\mathbf{r})$. For each $\mathbf{m}$, the above integral over w_1, w_2, w_3 can be identified as the n_1, n_2, n_3-th coefficient in the three-dimensional Fourier series expansion of $h_{p,\mathbf{m},\mathbf{r}}$. Note that for $\mathbf{m}\neq 0$ and $p\geqslant 1$, $h_{p,\mathbf{m},\mathbf{r}}$ is a smooth, bounded function over w_1, w_2, w_3 in the unit cube. The same is true if $\mathbf{m}=0$ and $p>3$. For $\mathbf{m}=0$ and $1\leqslant p\leqslant 3$, there is an integrable singularity at the origin.

For $p\geqslant 1$ and $\mathbf{r}\neq 0$, we consider the limit as $K\to\infty$ of the sum $S_{K,p}(\mathbf{r})$ given by

$$S_{K,p}(\mathbf{r})=\sum_{\mathbf{n}\in P_K(\Lambda)}\frac{1}{|\mathbf{r}+\mathbf{n}|^p}. \tag{A9}$$

Applying Eq. (A8), exchanging the order of summation between $\mathbf{m}$ and $\mathbf{n}$, and using the fact that the sum of the Fourier coefficients of the smooth, bounded function $h_{p,\mathbf{m},\mathbf{r}}$, $\mathbf{m}\neq 0$ converges to $h_{p,\mathbf{m},\mathbf{r}}(0,0,0)=f_p(\pi|\mathbf{m}|/\beta)$, we can write

$$\lim_{K\to\infty} S_{K,p}(\mathbf{r}) = \frac{\pi^{3/2}\beta^{p-3}}{V} \lim_{K\to\infty} \sum_{\mathbf{n}\in P_K(\Lambda)} \int_{-1/2}^{1/2}\int_{-1/2}^{1/2}\int_{-1/2}^{1/2} \times h_{p,0,\mathbf{r}}(w_1,w_2,w_3)\cdot \exp(-2\pi i(w_1 n_1 + w_2 n_2 + w_3 n_3))\, dw_1 dw_2 dw_3 + \frac{\pi^{3/2}\beta^{p-3}}{V} \sum_{\mathbf{m}\neq 0} f_p(\pi|\mathbf{m}|/\beta) \times \exp(-2\pi i \mathbf{m}\cdot\mathbf{r}) + \sum_{\mathbf{n}} \frac{g_p(\beta|\mathbf{r}+\mathbf{n}|)}{|\mathbf{r}+\mathbf{n}|^p}. \tag{A10}$$

From Eq. (A4) we see that

$$\lim_{\mathbf{r}\to 0}\left(\frac{1}{r^p} - \frac{g_p(\beta r)}{r^p}\right) = \frac{2\beta^p}{p\Gamma(p/2)}$$

(the ''self-energy'' term), which can be used to modify Eq. (A10) when $\mathbf{r}=0$.

For $p>3$ the first term on the right-hand side of Eq. (A10) converges to $(\pi^{3/2}\beta^{p-3}/V)f_p(0)$ by the above Fourier expansion argument. Note that this limit is independent of the order of summation specified by P, i.e., the sum converges absolutely. These results can be applied directly to the sum in Eq. (A1) to arrive at

$$E_p(\mathbf{r}_1,\ldots,\mathbf{r}_N) = \frac{1}{2}{\sum_{\mathbf{n}}}' \sum_i \sum_j \frac{C(i,j)g_p(\beta|\mathbf{r_i}-\mathbf{r}_j+\mathbf{n}|)}{|\mathbf{r}_i-\mathbf{r}_j+\mathbf{n}|^p} + \frac{\pi^{3/2}\beta^{p-3}}{2V}\sum_{\mathbf{m}} f_p(\pi|\mathbf{m}|/\beta) \times \sum_i \sum_j C(i,j)\cdot\exp[-2\pi i\mathbf{m}\cdot(\mathbf{r}_i-\mathbf{r}_j)] - \frac{\beta^p}{p\Gamma(p/2)}\sum_i C(i,i). \tag{A11}$$

If the coefficients $C(i,j)$ are ''factorizable,'' i.e., $C(i,j) = \sqrt{C(i,i)}\sqrt{C(j,j)}$ (possibly with a minus sign), then the reciprocal sum in the above equation can be written in terms of ''structure factors'' similar to those discussed in the text, and the PME methodology can be applied.

For $p\leqslant 3$ the left-hand side of Eq. (A10) diverges, as does $\lim_{x\to 0} f_p(x)$. We examine the case $p=1$, with $C(i,j)=q_i q_j$, where the charges q_i sum to zero (neutral unit cell U). We then can write

$$E_{1,P}(\mathbf{r}_1,\ldots,\mathbf{r}_N) = \frac{1}{2}{\sum_{\mathbf{n}}}' \sum_i \sum_j \frac{q_i q_j \operatorname{erfc}(\beta|\mathbf{r_i}-\mathbf{r}_j+\mathbf{n}|)}{|\mathbf{r}_i-\mathbf{r}_j+\mathbf{n}|} + \frac{1}{2\pi V}\sum_{\mathbf{m}\neq 0}\frac{\exp(-\pi^2\mathbf{m}^2/\beta^2)}{\mathbf{m}^2} S(\mathbf{m})S(-\mathbf{m}) - \frac{\beta}{\sqrt{\pi}}\sum_i q_i^2 + \frac{1}{2\pi}\lim_{K\to\infty}\sum_{\mathbf{n}\in P_K(\Lambda)}\int_{U^*}\sum_{i,j}\frac{q_i q_j \exp(-\pi^2 v^2/\beta^2)}{v^2}\cdot\exp[-2\pi i\mathbf{v}\cdot(\mathbf{r}_i-\mathbf{r}_j)]\exp(-2\pi i\mathbf{v}\cdot\mathbf{n})\, d^3\mathbf{v}. \tag{A12}$$

Examining the last term in Eq. (A12), we apply a second-order Taylor series expansion to the function $\exp(-\pi^2 v^2/\beta^2)\exp[-2\pi i\mathbf{v}\cdot(\mathbf{r}_i-\mathbf{r}_j)]$, expanding about $\mathbf{v}=0$. The zeroth and first-order terms, which account for the singularity in the integral, are cancelled by the double summation over i and j (neutral unit cell). The remainder term, which is of order three, can be handled by the above Fourier series argument, and converges to zero as $K\to\infty$. After some rearrangement, the second-order terms reduce to $4\pi^2(\mathbf{v}\cdot\mathbf{D})^2/v^2$, where $\mathbf{D}=\Sigma_i q_i\mathbf{r}_i$ is the unit cell dipole moment, plus terms that are cancelled by the unit cell neutrality. The last term in Eq. (A12) can thus be written as

$$J(\mathbf{D}) = 2\pi \lim_{K\to\infty}\sum_{\mathbf{n}\in P_K(\Lambda)}\int_{U^*}\frac{(\mathbf{v}\cdot\mathbf{D})^2}{v^2}\exp(-2\pi i\mathbf{v}\cdot\mathbf{n})\, d^3\mathbf{v}. \tag{A13}$$

Note that the above Fourier series argument cannot be applied to the right-hand side of Eq. (A13), since the function $f(\mathbf{v})=(\mathbf{v}\cdot\mathbf{D})^2/v^2$ does not have a limit as $\mathbf{v}$ tends to zero. Note, however, that $J(\mathbf{D})=0$ if $\mathbf{D}=0$, and in this case the sum in Eq. (A12) converges absolutely. Smith[41] and Deem *et al.*[63] derive more tractable expressions for $J(\mathbf{D})$. In particular, if P is the unit ball, $J(\mathbf{D})=(2\pi/3V)\mathbf{D}^2$.

The ''surface term'' $J(\mathbf{D})$ vanishes if $\mathbf{D}$ is zero. Smith[41] shows how charge rearrangement at the surface of a macroscopic crystal should also cancel its effect (''tinfoil boundary conditions''). Traditionally it has been neglected or set to zero in simulations involving Ewald sums, but recently Roberts and Schnitker[14,43] have conducted simulations to examine the consequences of including it, demonstrating its relation to the usual reaction field correction to cutoff energies and forces.

APPENDIX B: DISCRETE FOURIER TRANSFORMS

Let K_1, K_2, K_3 be positive integers. Given a complex-valued array $A(k_1,k_2,k_3)$, $0\leqslant k_i<K_i$ the discrete Fourier transform is given by

$$F(A)(m_1,m_2,m_3) = \sum_{k_1=0}^{K_1-1}\sum_{k_2=0}^{K_2-1}\sum_{k_3=0}^{K_3-1} A(k_1,k_2,k_3) \cdot\exp\left[2\pi i\left(\frac{m_1 k_1}{K_1}+\frac{m_2 k_2}{K_2}+\frac{m_3 k_3}{K_3}\right)\right], \tag{B1}$$

while the inverse discrete Fourier transform is given by

$$F^{-1}(A)(m_1,m_2,m_3) = \frac{1}{K_1K_2K_3}\sum_{l_1=0}^{K_1-1}\sum_{l_2=0}^{K_2-1}\sum_{l_3=0}^{K_3-1} A(l_1,l_2,l_3) \cdot \exp\left[-2\pi i\left(\frac{m_1l_1}{K_1}+\frac{m_2l_2}{K_2}+\frac{m_3l_3}{K_3}\right)\right]. \quad \text{(B2)}$$

Note that $F^{-1}[F(A)]=FF^{-1}(A)=A$. Next, given complex valued arrays A and B, the following two identities can be derived by straightforward algebra:

$$\sum_{l_1=0}^{K_1-1}\sum_{l_2=0}^{K_2-1}\sum_{l_3=0}^{K_3-1} F(A)(l_1,l_2,l_3)\cdot B(l_1,l_2,l_3) = \sum_{l_1=0}^{K_1-1}\sum_{l_2=0}^{K_2-1}\sum_{l_3=0}^{K_3-1} A(l_1,l_2,l_3)\cdot F(B)(l_1,l_2,l_3) \quad \text{(B3)}$$

$$A\star B = F[F^{-1}(A\star B)] = K_1K_2K_3\cdot F[F^{-1}(A)\cdot F^{-1}(B)], \quad \text{(B4)}$$

where $A\star B$ denotes the convolution of A and B, given by

$$A\star B(j_1,j_2,j_3) = \sum_{k_1=0}^{K_1-1}\sum_{k_2=0}^{K_2-1}\sum_{k_3=0}^{K_3-1} A(j_1-k_1,j_2-k_2,j_3-k_3)\cdot B(k_1,k_2,k_3). \quad \text{(B5)}$$

Note that the convolution, as expressed in Eq. (B5), is an order K^2 computation, where $K=K_1K_2K_3$, whereas using Eq. (B4) and the 3DFFT, it is reduced to order $K\log(K)$.

APPENDIX C: MORE ABOUT *B*-SPLINES AND EULER SPLINES

In this Appendix we give background material on Cardinal splines, and then derive Eq. (4.3) in the text. Polynomial splines are smooth functions that are piecewise polynomials. The support of a polynomial spline is divided into subintervals by the knots of the spline. Within each of these sub-intervals the spline agrees with a polynomial, and the spline is continuously differentiable several times at the knots. Cardinal splines are polynomial splines having simple, equally spaced knots, and Cardinal B-splines provide a basis for the vector space of Cardinal splines of a given order. Our discussion follows that in Schoenberg[50] as well as Chap. 4 of Chui.[64] For positive integers k, let π_k denote the space of polynomials of degree at most k. Next, let S_n, $n\geq 2$, denote the space of nth order polynomial splines with simple knots at the integers, that is, those real functions f which are $n-2$ times continuously differentiable and such that for all integers j, the restriction of f to the half open interval $j\leq u<j+1$ is in π_{n-1}. Thus, for example, the cubic splines with knots at the integers are denoted S_4. The n-1st derivative $f^{(n-1)}$ of a function $f\in S_n$ will be a step function with jumps at the integers.

The collection S_n is a vector space. A basis for that space is provided by the Cardinal B-splines, which are generated by integer translations of a function $M_n(u)$ which we now describe. For any function f defined on the reals, define the backwards difference by $\Delta f(u)=f(u)-f(u-1)$, and for $n\geq 2$ let $\Delta^n f(u)=\Delta(\Delta^{n-1})f(u)$. Note that $\Delta^n f(u)=0$ for $f\in\pi_n$. By induction, we see that $\Delta^n f(u)=\Sigma_{k=0}^n(-1)^k n!/(k!(n-k)!)f(u-k)$. For any real number u, define u_+ and then $M_n(u)$ by $u_+=\max(u,0)$, $u_+^{n-1}=(u_+)^{n-1}$, and

$$M_n(u) = \frac{1}{(n-1)!}\Delta^n u_+^{n-1} = \frac{1}{(n-1)!}\sum_{k=0}^{n}(-1)^k\frac{n!}{k!(n-k)!}(u-k)_+^{n-1}. \quad \text{(C1)}$$

The following result is proven in Schoenberg[50] as well as Chui.[64]

Theorem 1. *The integer translates of the nth-order B-spline $M_n(u)$ form a linear basis for the polynomial splines of order n having knots at the integers. That is, $M_n(u)\in S_n$ and for any other function $f\in S_n$, we can write $f(u)=\Sigma_{j=-\infty}^{\infty}c_jM_n(u-j)$, where the coefficients c_j are uniquely determined.*

The functions $M_n(u)$ have a number of other useful properties,[64] some of which we summarize as

Theorem 2. *The nth-order B-spline $M_n(u)$ satisfies the following properties:*

1. $M_n(u)>0$ for $0<u<n$; $M_n(u)=0$ for $u\leq 0$ and $u\geq n$.
2. $M_n(u)=M_n(n-u)$.
3. $\Sigma_{j=-\infty}^{\infty}M_n(u-j)=1$.
4. $M_n(u)=\frac{u}{n-1}M_{n-1}(u)+\frac{n-u}{n-1}M_{n-1}(u-1)$.
5. $\frac{d}{du}M_n(u)=M_{n-1}(u)-M_{n-1}(u-1)$.

Chui shows that $M_n(u)$ is the probability density of the sum of n independent random variables, each distributed uniformly on the unit interval. Properties (1) and (2) of theorem 2 are simple consequences of this result. Property (3) of theorem 2 says that the B-splines form a smooth partition of unity. Properties (4) and (5) are given as Eqs. (4.1) and (4.2) in the text.

Consider the complex-valued function $g(u;z)=z^u$, where u is a real number, and z is complex. For example, choosing $z=\exp(2\pi im/K)$ gives the complex exponentials appearing in the structure factors in Eq. (2.2). Note that g satisfies the functional equation $g(u+1;z)=z\cdot g(u;z)$. Euler posed the problem of interpolating g at integer values of u by smooth, piecewise polynomials (now called splines), that satisfy the same functional equation. The solutions to this problem are known as Euler exponential splines. They are discussed in Schoenberg,[50] who also provides some of the history.

In our notation we want to interpolate g by an nth order spline $g_n\in S_n$, which satisfies $g_n(u+1;z)=z\cdot g_n(u;z)$, for all u. Using the basis property from theorem 1, we write

$$g_n(u+1;z) = \sum_{j=-\infty}^{\infty} c_jM_n[u-(j-1)] = \sum_{j=-\infty}^{\infty} c_{j+1}M_n(u-j) = z\cdot\sum_{j=-\infty}^{\infty} c_jM_n(u-j)$$

and thus, by the uniqueness of the coefficients we have that $c_{j+1}=z\cdot c_j$ for all j. By repeated application of this equation we can then write, for some complex constant c

$$g_n(u;z)=c\sum_{j=-\infty}^{\infty} z^j M_n(u-j)=c\cdot\Phi_n(u;z). \tag{C2}$$

Suppose $\Phi_n(0;z)$ is not zero. Let $c=1/\Phi_n(0;z)$. Then g_n, given by Eq. (C2), satisfies $g_n(0;z)=1$, and using the above functional equation, $g_n(j;z)=z^j=g(j;z)$; that is, g_n interpolates g at the integers. On the other hand, if $\Phi_n(0;z)=0$, then $g_n(j;z)$ must be zero for all j, and the interpolation problem has no solution in S_n. Thus, it is important to determine the values of z for which $\Phi_n(0;z)=0$. We can rewrite $\Phi_n(0;z)$ as

$$\begin{aligned}\Phi_n(0;z)&=\sum_{j=-\infty}^{\infty} M_n(-j)\cdot z^j\\ &=\sum_{j=-\infty}^{\infty} M_n(n+j)\cdot z^j\\ &=\frac{1}{z^{n-1}}\sum_{k=0}^{n-2} M_n(k+1)\cdot z^k\\ &=\frac{1}{(n-1)!\cdot z^{n-1}}\,\Pi_{n-1}(z),\end{aligned} \tag{C3}$$

where we have used properties (1) and (2) of theorem 2. The third expression in Eq. (C3) is used in Eq. (4.4) of the text. The function $\Pi_{n-1}(z)$ defined in Eq. (C3) is known as the Euler–Frobenius polynomial of degree $n-2$. If $z=1$, then by property (3) in Theorem 2 above, $\Phi_n(0;z)=1$. Schoenberg treats the case $z\neq 1$ and, following Frobenius, shows that the roots of the polynomial $\Pi_{n-1}(z)$ are real, negative, and occur in reciprocal pairs, that is, if $\Pi_{n-1}(z)=0$ then $\Pi_{n-1}(1/z)=0$ also.

For our purposes, we are interested in the case $z=\exp(2\pi i m/K)$, for arbitrary integers m and K. Since the roots are real and negative, the only possible root is at $z=-1$, that is, for $m=K/2$. If the interpolation order n is odd, then applying property (2) of theorem 2 to $M_n(u)$ in Eq. (C3), we see that -1 is indeed a root, i.e., $\Phi_n(0;-1)=0$ and interpolation fails. If n is even (e.g., cubic interpolation), then so is $n-2$, and the $n-2$ roots of $\Pi_{n-1}(z)$ occur in reciprocal pairs, so that -1 cannot be a root. Thus if n is even, $\Phi_n(0;z)$ cannot be zero, and it can be inverted to give $b(m)$ in Eq. (4.4) in the text, thereby solving the interpolation problem.

[1] D. M. York, T. A. Darden, and L. G. Pedersen, J. Chem. Phys. **99**, 8345 (1993).
[2] D. York, A. Wlodawer, L. Pedersen, and T. Darden, Proc. Natl. Acad. Sci. **91**, 8715 (1994).
[3] H. Lee, T. A. Darden, and L. G. Pedersen, J. Chem. Phys. **102**, 3830 (1995).
[4] D. M. York, W. Yang, T. Darden, and L. G. Pedersen, J. Am. Chem. Soc. **117**, 5001 (1995).
[5] H. Schreiber and O. Steinhauser, Biochemistry **31**, 5856 (1992).
[6] P. E. Smith and B. M. Pettitt, J. Chem. Phys. **95**, 8430 (1991).
[7] M. Saito, J. Chem. Phys. **101**, 4055 (1994).
[8] D. H. Kitson, F. Avbelj, J. Moult, D. T. Nguyen, J. E. Mertz, D. Hatzi, and A. T. Hagler, Proc. Nat. Acad. Sci. **90**, 8920 (1993).
[9] S. Weerasinghe, P. E. Smith, V. Mohan, Y.-K. Cheng, and B. M. Pettitt, J. Am. Chem. Soc. **117**, 2147 (1995).
[10] T. E. Cheatham III, J. L. Miller, T. Fox, T. A. Darden, and P. A. Kollman, J. Am. Chem. Soc. **117**, 4193 (1995).
[11] (a) H. Schreiber and O. Steinhauser, Chem. Phys. **168**, 75 (1992); (b) H. E. Alper, D. Bassolino, and T. R. Stouch, J. Chem. Phys. **98**, 9798 (1993).
[12] J. S. Bader and D. Chandler, J. Phys. Chem. **96**, 6423 (1992).
[13] G. S. D. Buono, T. S. Cohen, and P. J. Rossky, J. Mol. Liq. **60**, 221 (1994).
[14] J. E. Roberts and J. Schnitker, J. Chem. Phys. **101**, 5024 (1994).
[15] L. Perera, U. Essmann, and M. L. Berkowitz, J. Chem. Phys. **102**, 450 (1995).
[16] P. Auffinger and D. L. Beveridge, Chem. Phys. Lett. **234**, 413 (1995).
[17] P. J. Steinbach and B. R. Brooks, J. Comp. Chem. **15**, 667 (1994).
[18] L. Verlet, Phys. Rev. **159**, 98 (1967).
[19] S. W. DeLeeuw, J. W. Perram, and E. R. Smith, Proc. R. Soc. London Ser. A **373**, 27 (1980).
[20] J. A. Barker and R. O. Watts, Mol. Phys. **26**, 789 (1973).
[21] G. Bossis, Mol. Phys. **38**, 2023 (1979).
[22] M. Berkowitz and J. A. McCammon, Chem. Phys. Lett. **90**, 215 (1982).
[23] J. W. Perram, H. G. Petersen, and S. W. DeLeeuw, Mol. Phys. **65**, 875 (1988).
[24] T. A. Darden, D. M. York, and L. G. Pedersen, J. Chem. Phys. **98**, 10 089 (1993).
[25] R. W. Hockney and J. W. Eastwood, *Computer Simulation Using Particles* (McGraw-Hill, New York, 1981).
[26] D. York and W. Yang, J. Chem. Phys. **101**, 3298 (1994).
[27] B. A. Luty, M. E. Davis, I. G. Tironi, and W. F. van Gunsteren, Mol. Sim. **14**, 11 (1994); B. A. Luty, I. G. Tironi, and W. F. van Gunsteren, J. Chem. Phys. **103**, 3014 (1995).
[28] J. Shimada, H. Kaneko, and T. Takada, J. Comp. Chem. **15**, 28 (1994).
[29] A. Toukmaji and J. A. Board, Technical Report No. 95-004, Duke University (unpublished).
[30] L. Greengard, *The Rapid Evaluation of Potential Fields in Particle Systems* (MIT, Cambridge, MA, 1988).
[31] K. E. Schmidt and M. E. Lee, J. Stat. Phys. **63**, 1223 (1991).
[32] H. Q. Ding, N. Karasawa, and W. Goddard, J. Phys. Chem. **97**, 4309 (1992).
[33] J. A. Board, Jr., J. W. Causey, J. F. Leathrum, A. Windemuth, and K. Schulten, Chem. Phys. Lett. **198**, 89 (1992).
[34] C. G. Lambert and J. A. Board, Chem. Phys. Lett. (submitted).
[35] L. F. T. Eyck, Acta Crystallogr. Sec. A **29**, 183 (1973).
[36] D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham III, D. M. Ferguson, G. L. Seibel, U. C. Singh, P. K. Weiner, and P. A. Kollman, AMBER 4.1, University of California, San Francisco, 1995.
[37] J. A. Board and A. Toukmaji (private communication).
[38] P. Ewald, Ann. Phys. **64**, 253 (1921).
[39] F. Berthaut, J. Phys. Radium **13**, 499 (1952).
[40] D. E. Williams, Acta Crystallogr. Sec. A **27**, 452 (1971).
[41] E. R. Smith, Proc. R. Soc. London **375**, 475 (1981).
[42] N. Karasawa and W. Goddard, J. Phys. Chem. **93**, 7320 (1989).
[43] J. E. Roberts and J. Schnitker, J. Phys. Chem. **99**, 1322 (1995).
[44] W. Smith, CCP5 Info. Quart. **26**, 43 (1987).
[45] H. C. Andersen, J. Chem. Phys. **72**, 2384 (1980).
[46] M. Parrinello and A. Rahman, Phys. Rev. Lett. **45**, 1196 (1980).
[47] S. Nose and M. L. Klein, Mol. Phys. **50**, 1055 (1983).
[48] E. R. Smith, J. Stat. Phys. **77**, 449 (1994).
[49] D. Brown and J. H. R. Clarke, Comp. Phys. Comm. **62**, 360 (1991).
[50] I. J. Schoenberg, *Cardinal Spline Interpolation* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1973).
[51] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).
[52] D. H. Kitchen, F. Hirata, J. D. Westbrook, R. Levy, D. Kofke, and M. Yarmush, J. Comp. Chem. **11**, 1169 (1990).
[53] M. Belhadji, H. E. Alper, and R. M. Levy, Chem. Phys. Lett. **179**, 13 (1991).
[54] J. P. Ryckaert, G. Ciccotti, and H. J. Berendsen, J. Comp. Phys. **23**, 27 (1977).
[55] W. L. Jorgensen, J. Chandrasekhar, and J. D. Madura, J. Chem. Phys. **79**, 926 (1983).
[56] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).
[57] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. D. DiNola, and J. R. Haak, J. Chem. Phys. **81**, 3684 (1984).
[58] O. Steinhauser, Chem. Phys. **79**, 465 (1983).

[59] P. E. Smith and W. F. van Gunsteren, J. Chem. Phys. **100**, 3165 (1994).
[60] M. Neumann, J. Chem. Phys. **85**, 1567 (1986).
[61] M. R. Reddy and M. L. Berkowitz, Chem. Phys. Lett. **155**, 173 (1989).
[62] M. F. Crowley (unpublished).
[63] M. W. Deem, J. M. Newsam, and S. K. Sinha, J. Phys. Chem. **94**, 8356 (1990).
[64] C. K. Chui, *An Introduction to Wavelets* (Academic, Boston, MA, 1992).
[65] H. G. Petersen, J. Chem. Phys. **103**, 3668 (1995).