Optimized particle-mesh Ewald/multiple-time step integration for molecular dynamics simulations

Paul F. Batcho,[a] David A. Case,[b] and Tamar Schlick[a,c]
Department of Chemistry and Courant Institute of Mathematical Sciences, New York University and the Howard Hughes Medical Institute, 251 Mercer Street, New York, New York 10012

(Received 13 February 2001; accepted 13 June 2001)

We develop an efficient multiple time step (MTS) force splitting scheme for biological applications in the AMBER program in the context of the particle-mesh Ewald (PME) algorithm. Our method applies a symmetric Trotter factorization of the Liouville operator based on the position-Verlet scheme to Newtonian and Langevin dynamics. Following a brief review of the MTS and PME algorithms, we discuss performance speedup and the force balancing involved to maximize accuracy, maintain long-time stability, and accelerate computational times. Compared to prior MTS efforts in the context of the AMBER program, advances are possible by optimizing PME parameters for MTS applications and by using the position-Verlet, rather than velocity-Verlet, scheme for the inner loop. Moreover, ideas from the Langevin/MTS algorithm LN are applied to Newtonian formulations here. The algorithm's performance is optimized and tested on water, solvated DNA, and solvated protein systems. We find CPU speedup ratios of over 3 for Newtonian formulations when compared to a 1 fs single-step Verlet algorithm using outer time steps of 6 fs in a three-class splitting scheme; accurate conservation of energies is demonstrated over simulations of length several hundred ps. With modest Langevin forces, we obtain stable trajectories for outer time steps up to 12 fs and corresponding speedup ratios approaching 5. We end by suggesting that modified Ewald formulations, using tailored alternatives to the Gaussian screening functions for the Coulombic terms, may allow larger time steps and thus further speedups for both Newtonian and Langevin protocols; such developments are reported separately. [DOI: 10.1063/1.1389854]

I. INTRODUCTION

The art of numerical simulation of biomolecular dynamics has made significant progress over the past decade.[1–4] Recent advances in time integration algorithms that utilize force splitting and multiple time step (MTS) methods enjoy decreased overall computational cost, and efficient evaluations of long-range electrostatic potentials have made possible more accurate biomolecular modeling.[5,6] Still, MTS methods are not uniformly used since single-time step (STS) explicit Verlet integrators are easy to program and implement;[2] the Verlet family has been the method of choice because of its simplicity combined with excellent conservation properties, the latter resulting from the symplectic character.[7] For explicit STS approaches, a time step of 0.5–1 fs is typically used to satisfy both accuracy and stability requirements. Still, Verlet applications on even modest biomolecular systems (around 20 000 atoms) require several weeks of computing time on state-of-the-art workstations to simulate 1 ns; important biomolecular processes occur on much longer time scales. Studies have shown that MTS integrators can be effective at larger time steps for the slower dynamical components while maintaining both accuracy and stability with respect to the fastest time scales. [1,2,8–17]

Efforts at decreasing the cost of potential energy and force evaluations per time step have focused on the time consuming (slower) long-range electrostatic component of the force field. Implementation of fast-multipole methods (FMM),[18,19] Ewald summations,[20] and the particle-mesh Ewald (PME)[5,21–25] method have been applied to biomolecular systems. Recent evidence[26,27] indicates that PME has an advantage in computational cost over FMM for simulations over 20 000 atoms, especially on loosely coupled processor architectures. Recent reviews[4,6] discuss the evaluation of electrostatic interactions by abrupt truncation and Ewald summation based methods. The PME or force-shift methods are recommended for long-range electrostatic evaluation; the PME method with a real space cutoff of less than 12 Å is generally considered to be optimal in overall computational time.

Here we present recent results on combining the PME method with force splitting MTS algorithms in the widely used AMBER molecular mechanics and dynamics package.[28,29] We examine several force splitting strategies for the PME formulation and show that a factor of two in computational speedup can be achieved (relative to the non-optimized PME/MTS combination) merely by adjusting the Ewald screening parameter. We also show that modest Langevin forces can enhance numerical stability while maintaining good conservation properties.

[a]Department of Chemistry and Courant Institute of Mathematical Sciences, New York University and the Howard Hughes Medical Institute, 251 Mercer Street, New York, NY 10012.
[b]The Scripps Research Institute, Dept. of Molecular Biology, La Jolla, CA 92037.
[c]Author to whom correspondence should be addressed. Fax: 212-995-4152. Electronic mail: schlick@nyu.edu

Our MTS implementation in AMBER is based on the symmetric Trotter factorization developed extensively in the reference system propagator algorithm, RESPA.[15–17] This symmetric framework maintains good conservation properties at larger time steps and can be formulated as a symplectic integrator. Nonsymplectic stochastic Langevin integrators have recently been studied within the context of LN-MTS protocol for biomolecular simulations.[8,10,30] The use of extrapolation combined with Langevin dynamics succeeds in damping the resonant impulses inherent in the symplectic MTS integrations and allows larger outer time steps, while maintaining stability and conserving energy.[31] The implementation of LN in the CHARMM program has thus far been limited to periodic boundary models with long-range cutoffs. Our optimized MTS/PME AMBER Newtonian protocol offers speedup factors over 3, as compared to a 1 fs time step STS Verlet integration (speedup factors double compared to 0.5 fs STS simulations); accurate conservation of energies is demonstrated for outer time step less than 6 fs. (Long simulations are required to establish stability, since slow heating can take hundreds of picoseconds to become evident). With modest Langevin forces, we report stable outer time steps up to 12 fs and associated speedup ratios approaching 5.

We remark that the first application of MTS-RESPA with Ewald was presented in Procacci *et al.*[32] and the first MTS-PME work was presented in Ref. 33. Recent studies[34,35] have incorporated the PME method for electrostatic evaluations into a MTS-RESPA integrator with the Nosé–Hoover chain extended system method for temperature and pressure control. The various MTS force splitting strategies examined in Ref. 34 include the use of switch functions to split the real space sum (also studied in Ref. 36). These results based on 20 ps simulation lengths indicate an 8 fs outer time step barrier; however, the Ewald cut-off values for the direct space summation of 10–12 Å are larger than those we find optimal for MTS protocols (less than 7 Å). The study of Ref. 35 indicated that a 4 fs outer time step was most practical for the RESPA integrator with direct-space cut-off values of 13 Å. Our work differs in its examination of more optimal Ewald force partitioning schemes tailored to MTS. In addition, we apply MTS/PME to a purely Newtonian formulation (i.e., no artificial temperature coupling), as well as to a Langevin model. Significantly, our position–Verlet based MTS scheme, rather than the velocity-Verlet version used in Refs. 34 and 35, is modeled after the LN algorithm[9,10] and is found to be advantageous with respect to enhanced energy conservation as well as numerical stability (see companion paper[37] for theoretical analysis).

In Sec. II, we begin (Sec. II A) by reviewing the MTS protocol, Ewald summation, and PME algorithms. Algorithms are presented for the position-Verlet and velocity-Verlet versions of the Langevin MTS-PME protocols studied here. Our companion paper[37] provides a theoretical grounding for our empirical observations, illustrating the advantage of position-Verlet over velocity-Verlet for moderate time steps used to resolve the fastest MTS forces; in the large inner-time step limit (i.e., not the more usual limit of the fast time step approaching zero), position-Verlet displays less violent resonant artifacts, which are inherent in impulse-MTS splitting.[30] We then discuss in Sec. II B the force fields used in the AMBER software and, in Sec. II C, we elaborate on the PME method and its use in our protocol.

In Sec. III, we present computational results for several systems and discuss optimal force balancing strategies. Our three test cases are a system of 4096 water molecules (12 288 atoms), a solvated DNA system (9898 atoms), and a solvated protein system (13 705 atoms). Section IV A offers a qualitative assessment of the various PME force partitions studied here, and Sec. IV B discusses the simulation results. Section V describes the use of modest Langevin forces, and Section VI concludes with a brief summary and suggestions for future algorithm enhancements.

II. MULTIPLE TIME STEP FORCE SPLITTING METHOD

A. Trotter factorization for multiple time step integration

Our MTS algorithm is based on the symmetric and time-reversible Trotter factorization of the Liouville operator[15,16] used to derive consistent MTS integrators for solving Newton's classical equations of motion:

$$\frac{dP}{dt} = -\nabla E(X), \tag{1}$$

$$\frac{dX}{dt} = V. \tag{2}$$

Here $P = MV$ is the collective momentum vector, V is the collective velocity vector, M is the mass matrix, and E is the potential energy function of the macromolecular system. These new integrators were shown to have attractive energy conservation properties at large time steps relative to those used in STS-Verlet schemes. The formulation of the algorithm starts with a given phase-space vector, $\Gamma(X,P)$, and the Liouville operator, L. This operator is defined as the inner product of the time derivative of Γ with the differential operator, as

$$iL = \dot{\Gamma} \cdot \nabla_{\Gamma} = \sum_{i=1}^{d} \left[\dot{X}_i \frac{\partial}{\partial X_i} + F_i \frac{\partial}{\partial P_i} \right]. \tag{3}$$

Here X_i and P_i are the position and conjugate momenta components for coordinate i, $\dot{X}_i$ is the time derivative of X_i, and $F_i = -\nabla E_i$ is the force acting on the ith independent variable. The state of the system at time Δt is given by applying the propagator to the phase space vector at $t = 0$, e.g.,

$$\Gamma(\Delta t) = \exp(iL\Delta t)\Gamma(0). \tag{4}$$

Here we split the momentum part of the Liouville operator into three separate force components (slow, medium, fast), i.e., F^s, F^m, F^f, respectively, as well as an external force F^e (used for the Langevin component). The operator iL is then decomposed as

$$iL = iL_1 + iL_2 + iL_3, \tag{5a}$$

$$iL_1 = \dot{X}_i \frac{\partial}{\partial X_i}, \tag{5b}$$

$$iL_2=(F_i^e+F_i^f)\frac{\partial}{\partial P_i}, \tag{5c}$$

$$iL_3=(F_i^s+F_i^m)\frac{\partial}{\partial P_i}. \tag{5d}$$

A symmetric factorization of this propagator gives

$$\exp(iL\Delta t)_{\mathrm{PV}}=\exp(iL_3\Delta t/2)[\exp(iL_1\Delta t/2)\exp(iL_2\Delta t)\times\exp(iL_1\Delta t/2)]\exp(i\hat{L}_3\Delta t/2), \tag{6}$$

where we commute iL_3 and define the operator $i\hat{L}_3=(F_i^m+F_i^s)\partial/\partial P_i$.

The factorization of the operators for the symmetric time-reversible propagator above yields discretization errors of $\mathcal{O}(\Delta t^3)$.[38] As written above, the factorization uses position Verlet (PV) as the inner propagator, as in the LN algorithm.[8,10,30] In contrast, RESPA protocols[15,34] have been formulated to use a velocity Verlet (VV) formulation, given as

$$\exp(iL\Delta t)_{\mathrm{VV}}=\exp(iL_3\Delta t/2)\times[\exp(i\tilde{L}_2\Delta t/2)\exp(i\tilde{L}_1\Delta t)\exp(i\tilde{L}_2\Delta t/2)]\times\exp(i\hat{L}_3\Delta t/2), \tag{7a}$$

$$i\tilde{L}_1=F_i^e\frac{\partial}{\partial P_i}+\dot{X}_i\frac{\partial}{\partial X_i}, \tag{7b}$$

$$i\tilde{L}_2=F_i^f\frac{\partial}{\partial P_i}. \tag{7c}$$

Our investigations indicate that the PV version has enhanced stability over the VV algorithm when moderate time step values are used for the fast and medium forces and large values are used for the slow forces; advantages of PV over VV were suggested in Ref. 16. Our companion paper lends theoretical support to these practical observations.[37]

To introduce our MTS protocol within the symmetric factorization, we define three time steps, associated with the evolution of the slow, medium, and fast force components: Δt, $\Delta t_m=\Delta t/k_2$, and $\Delta\tau=\Delta t/(k_1k_2)$, where k_1 and k_2 are integers. Within this construction, Eq. (6) can be rewritten as (we neglect particle index subscripts i for simplicity)

$$\exp\left(\frac{\Delta t}{2}M^{-1}F^s\frac{\partial}{\partial P}\right)[P_m]^{k_2}\exp\left(\frac{\Delta t}{2}M^{-1}F^s\frac{\partial}{\partial P}\right), \tag{8a}$$

where

$$P_m=\exp\left(\frac{\Delta t_m}{2}M^{-1}F^m\frac{\partial}{\partial P}\right)[P_f]^{k_1}\exp\left(\frac{\Delta t_m}{2}M^{-1}F^m\frac{\partial}{\partial P}\right) \tag{8b}$$

and

$$P_f=\exp\left(\frac{\Delta\tau}{2}\dot{X}\frac{\partial}{\partial X}\right)\exp\left(\Delta\tau(M^{-1}F^e+M^{-1}F^f)\frac{\partial}{\partial P}\right)\times\exp\left(\frac{\Delta\tau}{2}\dot{X}\frac{\partial}{\partial X}\right). \tag{8c}$$

The external Langevin force is given by

$$F^e=-\gamma MV+R^n(\Delta\tau,\gamma), \tag{9}$$

where $R^n(\Delta\tau,\gamma)$ is the stochastic white noise with properties given by

$$\langle R(t)\rangle=0,\quad \langle R(t)R(t')\rangle=2\gamma k_BTM\delta(t-t'), \tag{10}$$

or in discrete form

$$R^n(\Delta\tau,\gamma)=\epsilon_n\sqrt{2\gamma k_BTM}/\sqrt{\Delta\tau}. \tag{11}$$

Here, γ is the friction constant, k_B is the Boltzmann constant, T is the temperature (300 K used throughout), and ϵ_n is the random deviate of a normal distribution with a mean of 0 and a standard deviation 1; for $\gamma=0$, the protocol is characterized by a Newtonian microcanonical (MTS-NVE) algorithm.[39] In treating the Langevin forces within the factorization above, we note that exponential propagator has the properties $e^{cp(\partial/\partial p)}\phi(p)=\phi(e^cp)$ and $e^{c(\partial/\partial p)}\phi(p)=\phi(p+c)$ for $c\neq c(p)$. Using these results, we can evaluate the inner term of Eq. (8c) to arrive at the velocity update for the fastest forces of

$$V\leftarrow e^{-\gamma\Delta\tau}[V+M^{-1}\Delta\tau(F^f+R^n(\Delta\tau,\gamma))]. \tag{12}$$

Note that for $\gamma\Delta\tau\ll1$ we have $e^{-\gamma\Delta\tau}=1/(1+\gamma\Delta\tau)+\mathcal{O}((\gamma\Delta\tau)^2)$. After substituting this into Eq. (12) and neglecting higher order terms, we arrive at the Langevin form used in Sandu and Schlick,[8]

$$V\leftarrow[V+M^{-1}\Delta\tau(F^f+R^n(\Delta\tau,\gamma))]/(1+\gamma\Delta\tau). \tag{13}$$

This leads to our symmetric MTS protocol based on position Verlet, valid in the STS limit of $(k_1,k_2)=(1,1)$:

Algorithm I: position-Verlet based MTS

For ISTEP = 1 to NSTEP
Evaluate $F^s(X)$
$V\leftarrow V+\Delta tM^{-1}F^s$
For $j=1$ to k_2+1
Evaluate $F^m(X)$

$$V\leftarrow V+b\frac{\Delta t_m}{2}M^{-1}F^m,\quad b=\begin{cases}1 & \text{if } j=1 \text{ or } j=k_2+1\\ 2 & 1<j<k_2+1\end{cases}$$

if $(j=k_2+1)$ **endfor**
[For $i=1$ to k_1+1

$$X\leftarrow X+c\frac{\Delta\tau}{2}V,\qquad c=\begin{cases}1 & \text{if } i=1 \text{ or } i=k_1+1\\ 2 & 1<i<k_1+1\end{cases}$$

Evaluate $F^f(X)$
$V\leftarrow[V+\Delta\tau M^{-1}(F^f+R^n(\Delta\tau,\gamma))]/(1+\gamma\Delta\tau)$
endfor]
endfor

The velocity-Verlet formulation involves the following modification of the inner most cycle (bracketed above):
[For $i=1$ to k_1+1
Evaluate $F^f(X)$

$$V \leftarrow \left[V + c\frac{\Delta\tau}{2} M^{-1}\left(F^f + R^n\left(c\frac{\Delta\tau}{2}, \gamma \right) \right) \right] \Big/ \left(1 + \gamma c \frac{\Delta\tau}{2} \right)$$

$$c = \begin{cases} 1 & \text{if } i=1 \text{ or } i=k_1+1 \\ 2 & 1<i<k_1+1 \end{cases}$$

$$X \leftarrow X + \Delta\tau V$$

endfor].

We have formulated the above position-Verlet algorithm so as to minimize computational work involved with constrained dynamics variations; such iterations are applied after each position update of Algorithm I above. All our trajectories reflect, constrained dynamics formulations on all bonds involving hydrogen in the microcanonical ensemble. Future studies will incorporate constant temperature and pressure ensembles as well. Our Langevin dynamics formulation helps stabilize trajectories at larger time steps; often, RESPA integrators use Nosé thermostats as an external force and incorporate them into the outer (slow) force component.

B. AMBER force field potentials

The AMBER force field[28,29] represents the potential energy of the system as the sum of bond length, bond angle, dihedral angle, Coulomb, and hydrogen bond terms as

$$E_{PE} = E_r + E_\theta + E_\tau + E_{LJ} + E_C + E_{Hb}, \tag{14}$$

where

$$E_r = \sum_{i,j \in S_B} S_{ij}(r_{ij} - \overline{r}_{ij})^2, \tag{15}$$

$$E_\theta = \sum_{i,j,k \in S_{BA}} K_{ijk}(\theta_{ijk} - \overline{\theta}_{ijk})^2, \tag{16}$$

$$E_\tau = \sum_{i,j,k,l \in S_{DA}} \sum_n \frac{V_{n_{ijkl}}}{2}(1 + \cos(n\tau_{ijkl})), \tag{17}$$

$$E_{LJ} = \sum_{i<j} 4\epsilon_{ij}\left(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^{6}}{r_{ij}^{6}} \right) + \frac{1}{E_{\text{scale}}^{1,4}} \times \sum_{i<j} 4\epsilon_{ij}\left(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^{6}}{r_{ij}^{6}} \right), \tag{18}$$

$$E_C = \sum_{i<j} \frac{q_i q_j}{r_{ij}} + \frac{1}{E_{\text{scale}}^{1,4}} \sum_{i<j} \frac{q_i q_j}{r_{ij}}, \tag{19}$$

$$E_{Hb} = \sum_{i<j} \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right). \tag{20}$$

The first two terms, E_r and E_θ, represent harmonic type oscillations of bonded atoms and the third term, E_τ, is the torsional potential for the dihedral angle, τ, expressed by a truncated Fourier series. In these general expressions, the symbols S_B, S_{BA}, and S_{DA} denote the sets of all bonds, bond angles, and dihedral angles. Bond and angle variables capped by bar symbols denote equilibrium (i.e., reference) values associated with these quantities. The nonbonded interactions between two atoms are characterized by a van der Waals (vDW) potential, E_{LJ}, an electrostatic potential, E_C, which includes the 1–4 interaction pairs of atoms on the same molecule, and E_{Hb}, the 10–12 hydrogen bond potential.

C. PME formulation in AMBER

The solvated biomolecular system is considered one unit cell of an infinite periodic lattice, and its electrostatic interactions can be expressed by Ewald summations.[20] The splitting of the electrostatic term via the Ewald summation results in a reciprocal space and direct (real) space term for the Coulomb pair potential, $E_c = \Sigma_{i<j}(q_i q_j / r_{ij})$, as well as the correction terms,

$$E_c = E_{\text{recip.}} + E_{\text{real}} + E_{\text{cor. self}} + E_{\text{cor. ex}} + E_{\text{cor.}\epsilon}, \tag{21}$$

as described below. With the PME method, the electrostatic energy is efficiently evaluated on a computational mesh which requires interpolation of the charged particles and their forces. Briefly, the formulation of the sum on an infinite lattice with periodicity is greatly simplified by noting that the Coulomb potential in three dimensions can be split into two terms,

$$\frac{1}{r} = \frac{S(r)}{r} + \frac{1 - S(r)}{r}, \tag{22}$$

where $S(r)$ can be defined to be a rapidly decreasing function. This implies that the first term of Eq. (22) includes only near-field contributions, while the second is long-range but smooth. The second term of Eq. (22) and the switch function $S(r)$ are defined by the solution of Poisson equation subject to a spherically symmetric particle core function, $\sigma(r)$, which is normalized to have a net unit charge,

$$-\nabla^2 \phi(r) = 4\pi\sigma(r), \tag{23}$$

$$\frac{1 - S(r)}{r} = \phi(r). \tag{24}$$

For a periodic lattice, the reciprocal potential $E_{\text{recip}}(r)$ can be found from an efficient solution of Poisson's equation via Fourier expansions, where a charge distribution is defined by the sum of the smoothed point charges. Likewise, the real space summation is defined over screened charges subject to the rapidly decaying switch function,

$$E_{\text{real}} = \frac{1}{2} \sum_{i,j=1}^{N} q_i q_j \sum_{|\mathbf{n}|}{}' \frac{S(|\mathbf{r}_{ij} + \mathbf{n}|)}{|\mathbf{r}_{ij} + \mathbf{n}|}. \tag{25}$$

In this expression, the lattice index vector is $\mathbf{n} = (n_x L, n_y L, n_z L)$, where n_x, n_y, and n_z are integers, L is the box size, and $\mathbf{r}_{ij}$ is the radial separation between two particles. The prime symbol in the summation ($\Sigma'_{|\mathbf{n}|}$) indicates that for $|\mathbf{n}| = 0$ we omit the $i = j$ interaction and excluded pairs. See also Refs. 40 and 26 for more details on the mathematical construction for general core functions. The original Ewald summation uses a Gaussian function as the particle core function, $\sigma(r, \beta) = (\beta^3/\pi^{3/2})\exp(-r^2\beta^2)$, and results in the following potentials:[41]

$$E_{\text{real}}=\frac{1}{2}\sum_{i,j=1}^{N} q_i q_j \sum_{|\mathbf{n}|}{}' \frac{\operatorname{erfc}(\beta|\mathbf{r}_{ij}+\mathbf{n}|)}{|\mathbf{r}_{ij}+\mathbf{n}|}, \tag{26a}$$

$$E_{\text{recip.}}=\frac{1}{2\pi L^3}\sum_{|\mathbf{m}|\neq 0}\frac{\exp(-\pi^2|\mathbf{m}|^2/\beta^2)}{|\mathbf{m}|^2}\,\mathrm{G}(\mathbf{m})\mathrm{G}(-\mathbf{m}),$$

$$\mathrm{G}(\mathbf{m})=\sum_{j=1}^{N} q_j \exp[2\pi i\mathbf{m}\cdot\mathbf{x}_j], \tag{26b}$$

$$E_{\text{cor. self}}=\frac{-\beta}{\sqrt{\pi}}\sum_{j=1}^{N} q_j^2, \tag{26c}$$

$$E_{\text{cor. ex}}=-\frac{1}{2}\sum_{i,j\in \text{Ex.}}^{N} q_i q_j \frac{\operatorname{erf}(\beta|\mathbf{r}_{ij}|)}{|\mathbf{r}_{ij}|}, \tag{26d}$$

$$E_{\text{cor.}\epsilon}=\frac{2\pi}{3L^3}\left|\sum_{j=1}^{N} q_j \mathbf{x}_j\right|^2. \tag{26e}$$

Here β is the Ewald constant, q_i is the partial charge on atom i, erfc(x) is the complementary error function [erfc(x) $=(2/\sqrt{\pi})\int_x^\infty \exp(-u^2)du$]. For the reciprocal term we sum over Fourier modes where $\mathbf{m}=2\pi\mathbf{k}/L$, $V=L^3$, $m=|\mathbf{m}|$ $=2\pi|\mathbf{k}|/L$, and $\mathbf{k}=(k_x,k_y,k_z)$, k_x,k_y,k_z are integers, and L^3 is the volume of the cubic domain.

Several different strategies can be used for implementing the three-level force splitting as discussed above. We take the Ewald reciprocal force as the long-range "slow" force, F^s $=F_{\text{recip}}$; the bonded terms in the fast class, $F^f=F_r+F_\theta$ $+F_\tau$; and all other terms in the medium class force. This is typical for rigid cut-off methods with switching functions.[10,30] The medium force evaluation includes 1–4 interactions, vDW terms, and the real space sum for the PME method; the neighbor exclusion masking for the real space sum of Eq. (26d) was performed at each medium term evaluation. The list management utilities for the nonbonded (NB) terms were invoked at each medium force call; the AMBER software includes acceptance tests to determine if the NB list requires regeneration or updating.

Though it is possible to split the direct term itself by a radial switch function, with an additional pairlist maintenance, as done in Refs. 32, 34, and 36, we do not consider this here. Three different force splitting strategies were examined in Ref. 34 for MTS-RESPA integrators, including an algebraic switch function splitting of the real space sum at 7 Å. The results suggest a limited enhancement in CPU speedup associated with splitting the real space sum; our approach optimizes the CPU speedup relative to an unmodified Ewald splitting.

Inherent in the splitting construction is the underlying assumption that the reciprocal term represents, or more importantly isolates, the long-range slow forces. This is not true in the Ewald formulation. Studies in Ref. 40, and more recently Procacci *et al.*,[42] indicate that the Ewald reciprocal term has significant force contributions from near-field particle separations and that these forces lead to fast-time scale instabilities in MTS protocols. For this reason, we are considering alternative core functions such as $\sigma(r)$ $=\Sigma_{i=0}^{n} a_i r^{2i}\exp(-r^2\beta^2)$, where the coefficients $\{a_i\}$ are predetermined optimized coefficients, which can be designed to meet the desired far-field to near-field pairwise interaction splitting philosophy. See Ref. 40 for a study of several non-Gaussian core functions for efficient lattice summations. Alternatively, we are using more elaborate splitting terms that may help eliminate fast terms in the reciprocal component using switch functions and different treatments for Coulomb and van der Waals forces.[43] Here we use the Ewald Gaussian formulation for convenience in implementing the MTS protocol within the AMBER software package; we will report on the optimized core functions in future publications.[44]

III. OPTIMIZED PARTICLE-MESH EWALD METHOD

The accuracy of the PME method is related to the density of the charge grid (N_{grid}), the spline interpolation order (here taken as the AMBER default of four), and the direct sum error tolerance, erfc(βr)/r (here taken to be 10^{-6}). As discussed in Ref. 22, a PME estimated rms (root-mean-square) force error can be established from the above variables associated with the direct sum cut-off value, as well as the number of Fourier vectors associated with N_{grid}. Here we study a parameter set that maintains a maximum estimated rms force error of 10^{-4} and establish an optimal range of application within a given MTS protocol.

We first review the optimized force partitioning strategy used here. The two parameters we optimize for an MTS-PME protocol are the Ewald coefficient β and the density of the charge grid N_{grid}. The parameter β influences the range of the Gaussian shielding and hence the effective real space cutoff used for the nonbonded list maintained throughout the simulation.[7] The optimal Ewald method typically selects β to distribute the work evenly between the reciprocal and direct terms. As β increases, the coefficients in the Fourier series approximation decay at a slower rate, and therefore, larger expansions and greater computational work are required. At the same time, a large β leads to a direct term with a smaller region of influence, and therefore, the $\mathcal{O}(N_{\text{dir}}^2)$ work is reduced due to a smaller effective sphere of influence; here N_{dir} is the number of degrees of freedom retained in the screened direct sum. The value of N_{dir} remains constant for increasing system sizes, and thus an overall $\mathcal{O}(N)$ algorithm is obtained. The slow and medium relative CPU times can be adjusted by changing β for a desired rms force error. We thus express the total work for a STS-PME scheme as the following sum of CPU fractions for the slow (W_s), medium (W_m), and fast (W_f) terms:

$$Work_{\text{STS}}=W_s+W_m+W_f, \tag{27}$$

where $Work_{\text{STS}}=1$. For an MTS-PME scheme with given $\Delta t=k_2\Delta t_m=k_1k_2\Delta\tau$, we arrive at the corresponding work statement of

$$Work_{\text{MTS}}=\frac{W_s}{k_1k_2}+\frac{W_m}{k_1}+W_f. \tag{28}$$

Note that $Work_{\text{STS}}=Work_{\text{MTS}}(k_2=1,k_1=1,W_s,W_m,\Delta t)$ and the speedup ratio can be given by $Work_{\text{STS}}/Work_{\text{MTS}}$. Thus we see that an optimal STS scheme, which maximizes

TABLE I. Various parameters associated with the PME-STS calculations for different β choices for the water, solvated DNA, and solvated protein systems. The parameters in columns 2–7 are the Ewald β parameter, the PME charge grid density ($N_{\text{grid}_x}, N_{\text{grid}_y}, N_{\text{grid}_z}$) (given a maximum error tolerance of 1.0×10^{-4}), the Ewald cut-off distance (Å), the array size of the nonbonded pairlist, the STS fraction of reciprocal force CPU work to the direct force's CPU work, and the CPU time ratio relative to the fastest protocol (e.g., III for the water and I for the solvated biomolecular systems), respectively.

Partition	β	N_{grid}	CUT	NB lists	$W_{\text{recip}}/W_{\text{dir}}$	$\frac{\text{CPU}}{\text{CPU}_{\text{min}}}$
				Water		
I	0.646	(90,90,90)	5.0	541 000	83/12	1.6
II	0.533	(80,80,80)	6.0	871 000	70/25	1.3
III	0.491	(64,64,64)	6.5	1 077 000	51/41	1.0
IV	0.395	(50,50,50)	8.0	1 871 000	27/65	1.2
V	0.312	(40,40,40)	10.0	3 433 000	13/79	1.8
				Solvated DNA		
I	0.533	(80,80,80)	6.0	684 307	73/21	1.1
II	0.349	(45,45,50)	9.0	2 026 000	22/71	1.0
				Solvated Protein		
I	0.533	(80,80,80)	6.0	970 200	68/22	1.0
II	0.349	(50,50,50)	9.0	2 882 000	19/68	1.1

speedup, should have a larger fraction of work in its slow component; for our protocol, the component updated least often is the PME reciprocal term.

For example, consider the work per Δt cycle for an MTS scheme with parameters $k_1=7$ and $k_2=3$. If the fraction of work for a given system is $W_s=0.2$ and $W_m=0.75$, the speedup ratio of the MTS scheme with respect to the STS scheme is 3.2. If instead $W_s=0.75$ and $W_m=0.2$, the speedup ratio would be 6.6, more than double the previous value. Here we assume that the STS work is insensitive to changes in β and that secondary effects, such as smaller spatial cutoffs, do not introduce artifacts. As will be shown in the next section, this is a reasonable assumption. The simple analysis above indicates the merit of designing, where possible, general MTS protocols that have significantly more costly reciprocal terms with respect to the medium term, in the STS limit.

For a given Δt, we express a heuristic guide for an optimal MTS partition as

$$\frac{W_s}{k_1 k_2} \approx \frac{W_m}{k_1}. \tag{29}$$

This leads directly to a splitting scheme with

$$W_s \approx k_2 W_m. \tag{30}$$

Taking typical conditions found here, we estimate $\Delta t \sim 10$ fs and $k_2 \sim 2$ to 3; in other words, the ratio W_s/W_m should be about 2 to 3 for an optimal MTS force partition. Note that taking the largest inner time step ($\Delta\tau$) for a given outer time step Δt as the most optimal scheme would be an incorrect conclusion. A more practical optimization strategy is to maximize Δt while minimizing k_2 (i.e., using as large a Δt_m as possible). As we illustrate below, enhanced speedups can be found by lowering k_2; however, the speedup improvement is modest when raising $\Delta\tau$ from 0.5 fs to 1.0 fs.

Finally, we note that optimized choices of β for STS and traditional Ewald protocols have been examined in Ref. 45. The MTS-PME application Essmann *et al.*[22] includes a parametric study for various values of β and mesh resolutions; our optimal parameters for the MTS implementation differ from those used for STS integrators and somewhat narrow the parameter range of interest for MTS applications. Other optimization strategies include an ''optimal influence function'' derived in Hockney and Eastwood[21] for the reciprocal term (see also Ref. 26), and a procedure for Ewald sums.[46]

IV. RESULTS

A. Assessment of the PME force partition

Before discussing results of the PME-MTS optimization, we illustrate here performance of different PME-STS protocols (differing by β), since results affect the optimal PME-MTS design. With our implementation in AMBER, β is the only independent parameter since a prescribed rms error threshold, less than 10^{-4}, together with β determines the grid density.

TABLE II. The calculated percent deviation of the various PME protocols differing by β (or associated cut-off value) relative to partition *V* of Table I (cutoff of 10 Å) for the water system. The mean energy components are calculated by a PV-STS integrator with $\Delta t = 0.5$ fs over a 10.0 ps simulation.

PME protocol	E_{tot}	E_{pot}	E_{K}	E_{vDW}	E_{el}
I (5 Å)	0.1	0.08	0.04	0.8	0.03
II (6 Å)	0.008	0.14	0.3	0.07	0.07
III (6.5 Å)	0.001	0.007	0.06	0.01	0.06
IV (8 Å)	0.001	0.008	0.08	0.5	0.08

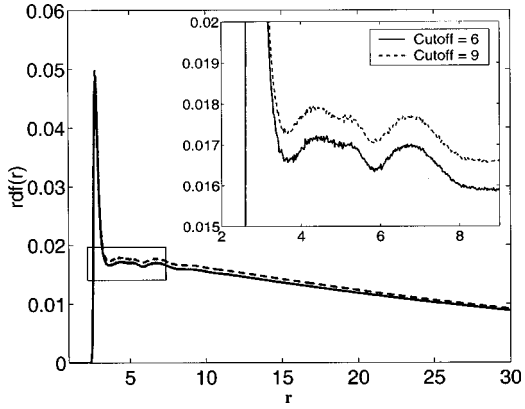

FIG. 1. The radial distribution function for the water oxygen atoms of the solvated protein simulation. The top right shows an enlargement for small r. The results were averaged over a 300 ps MTS simulation using cut-off values of 6 and 9 Å.

Recall that all our simulations used, constrained dynamics on bonds involving hydrogens. Simulations at room temperature were started on pre-equilibrated systems. The water system contained 4096 water molecules. The solvated DNA duplex consisted of a 9 base pair DNA duplex damaged by a metabolically activated form of the environmental carcinogen benzo[a]pyrene; the system has 621 solute atoms, 16 neutralizing Na^+ counterions, and 3087 water molecules, for a total of 9898 atoms.[47] The solvated protein system consists of the catalytic domain (residues 260–519) of the Src family tyrosine kinase Hck;[48,49] it has of 2214 protein atoms, 4 Na^+ ions, and 3829 water molecules (13 705 total atoms).

In Table I, we show results for five partitions characterized by the β parameter for our water, solvated DNA, and solvated protein systems (total system sizes range from about 10 000 to 14 000 atoms). For each β protocol, we show the charge density, cut-off value (in Å) for the direct sum, the size of the nonbonded pairlist array, CPU work fractions for the reciprocal W_{recip} and direct W_{direct} terms, and the CPU ratio between each protocol relative to the optimal protocol for that system.

We see that the direct-term work is larger than the reciprocal-term work for β values smaller than ~ 0.5 Å^{-2}, i.e., cut-off values greater than ~ 6.5 Å. Thus, we expect that cut-off values of 6 Å or lower would be preferred for MTS applications; this is indeed observed. We also conclude that the overall STS CPU time, within the β ranges examined here, is not sensitive to varying the direct to reciprocal work ratios for the solvated biomolecules.

As the MTS protocols become increasingly stable and larger outer time steps are possible, our analysis indicates that shifting the CPU work into the reciprocal force evaluations, in the STS limit, will become increasingly important for optimizing the total CPU time. In principle, there are no obstacles to the use of lower cut-off values (which can shift the work load from medium to slow terms) given that a fixed error is maintained by increasing the charge grid density. However, in AMBER, the effective Ewald direct term cutoff is

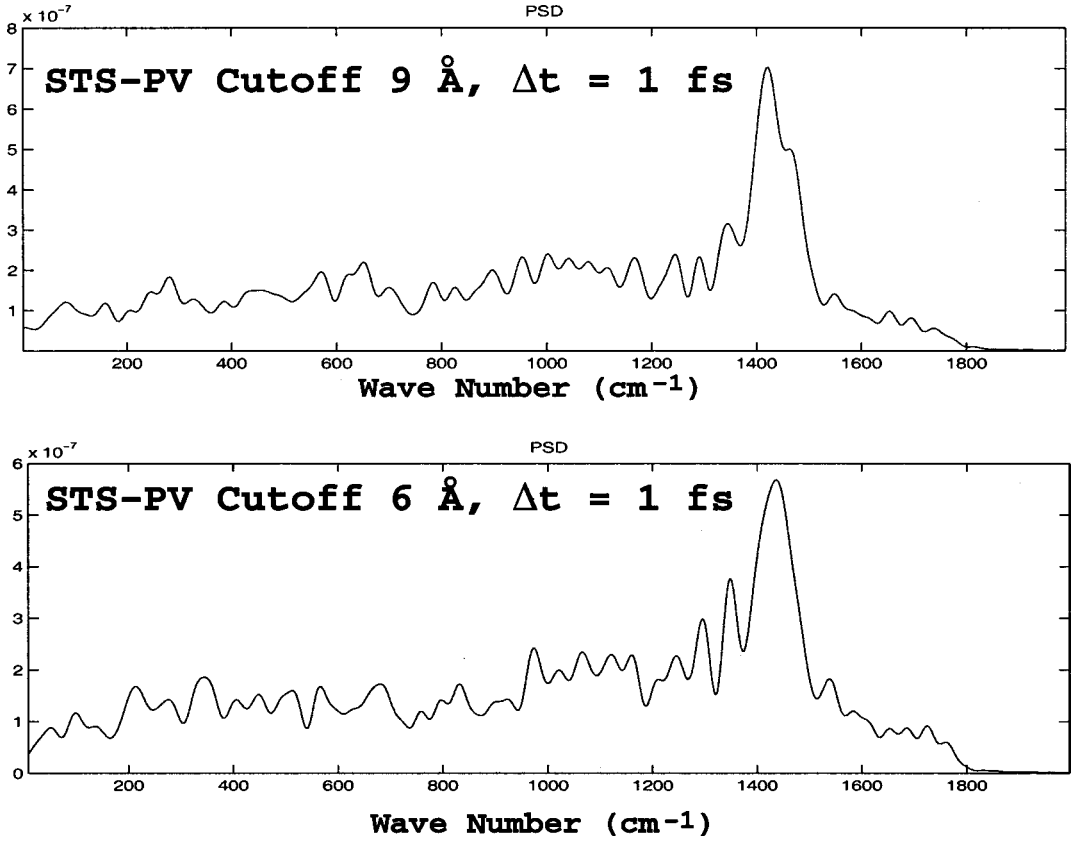

FIG. 2. The discrete Fourier transform of the velocity autocorrelation function of the solvated protein simulation for a STS–PV simulation with cut-off values of 9 and 6 Å ($\Delta t = 1.0$ fs and the simulation length was 40 ps).

TABLE III. Performance for symmetric Newtonian MTS schemes for the water system for various PME partitions at a fixed outer time step ($\Delta t=7$ fs). For each MTS scheme, we present the %Recip,%Direct values; the ratio of CPU relative to a STS scheme ($\Delta t=0.5$ fs, $N_{\text{grid}}=80$, cut-off = 6 Å), with the value in parentheses indicating the CPU ratio divided by 2.0 for comparison to an STS scheme with a 1.0 fs time step; and extrapolated clock time in days for a 1 ns simulation on a single MIPS R12000 processor (300 MHZ).

k_2, k_1	Δt_m	$\Delta\tau$	%Recip, %Direct	Speedup	CPU for 1 ns
I			(CUT=5, $N_{\text{grid}}=90$)		
7,2	1.0 fs	0.5 fs	38,39	4.9 (2.45)	7.4 days
4,2	1.75	0.875	48,30	5.9 (2.95)	6.1
2,4	3.5	0.875	55,22	7.2 (3.60)	5.0
II			(CUT=6, $N_{\text{grid}}=80$)		
7,2	1.0 fs	0.5 fs	21,58	4.3 (2.15)	8.4 days
4,2	1.75	0.875	28,50	5.8 (2.90)	6.2
2,4	3.5	0.875	36,38	7.5 (3.75)	4.8
IV			(CUT=8, $N_{\text{grid}}=50$)		
7,2	1.0 fs	0.5 fs	5,78	2.4 (1.20)	14.9 days
4,2	1.75	0.875	7,72	3.7 (1.85)	9.9
2,4	3.5	0.875	10,63	5.2 (2.60)	6.9
V			(CUT=10, $N_{\text{grid}}=40$)		
7,2	1.0 fs	0.5 fs	2,85	1.4 (0.70)	26.6 days
4,2	1.75	0.875	3,81	2.0 (2.00)	17.8
2,4	3.5	0.875	4,72	3.0 (1.50)	11.9

also used to evaluate the vDW terms via rigid cutoffs. This imposes a practical constraint on the MTS-PME implementation, since vDW interactions beyond that medium force cutoff are ignored. In theory, this limitation can be resolved by adding a portion of the vDW interaction in the slow forces, but two nonbonded pairlists will have to be maintained. Alternatively, the PME method could be applied to the vDW term;[22,50] the Ewald-type splitting of the vDW interaction would again provide a slower force component as it does for the electrostatic term.

For now, we assess the sensitivity of the vDW cutoff on the simulation by first comparing mean energies associated with variations in the cut-off value for the water system. Taking protocol *V* of Table I (cutoff of 10 Å) as the reference, we show in Table II the relative deviations for all energy components associated with a PV-STS simulation of length 10.0 ps for the water system. We note that the largest difference is found in the vDW energy, but even that difference is less than 1% for all force partitions studied; similar results are obtained for the solvated systems.

A second measure of the vDW cut-off value is given through the calculation of the radial distribution function[7] (rdf) for the solvated protein simulation. The system contains 13 705 atoms with 3829 water molecules, and several sodium ions. Of the systems studied here, the solvated protein is expected to be the most sensitive to truncated vDW effects. The rdf presented in Fig. 1 is for water oxygen–oxygen interactions for cut-off values of 6 and 9 Å and averaged over a 300 ps simulation; the MTS protocol used has $\Delta t=6$ fs, $(k_1,k_2)=(3,2)$, $(\Delta\tau,\Delta t_m)=(1.0$ fs, 2.0 fs), and the rdf is sampled every 10 inner time steps. The results for our solvated protein study do not reveal any spurious cut-off effects for the lower value; however, a slight downward shift of the rdf is found for the lower vDW cut-off value.

We also compare the Fourier transform of the velocity autocorrelation signal, $C_v(t)$, for all nonwater atoms for the solvated protein system for the cut-off values of 6 and 9 Å. The spectra are analyzed for a PV-STS simulation with $\Delta t=1.0$ fs, as calculated by the velocity autocorrelation function

$$C_v(t)=\frac{\left\langle \sum_{i=1}^{\text{N atoms}} V_i(t)\cdot V_i(0)\right\rangle}{\left\langle \sum_{i=1}^{\text{N atoms}} V_i^2(0)\right\rangle}. \tag{31}$$

Figure 2 shows the resulting signals below 2000 cm^{-1} for the 9 and 6 Å cut-off values. This part of the spectrum originates from bond stretching of the heavier atoms (e.g., C–O) and reveals a distinct peak at $\sim$1450 cm^{-1}; this generally agrees with results of Ref. 34. The inclusion of the water molecules adds considerable signal to the spectrum below 1450 cm^{-1} for both cut-off values.

B. Simulation details and discussion

We now analyze the results of MTS integrations for our three systems. The lengths of outer time step and numbers of inner and medium cycles are varied to establish stability criteria for the Newtonian and Langevin versions of Algorithm I. The mean energy components and the Fourier transform of the velocity autocorrelation are examined and compared to STS results. We also examine two criteria for measuring the quality of the integration scheme, namely the ratio of the total energy rms values to their mean value, and the ratio of total energy rms to kinetic-energy rms values.

In general, speedup ratios relative to the STS simulation with $\Delta t=0.5$ fs are $\sim$7 to 8 at a 6 fs MTS outer time step; with the addition of modest Langevin forces, this stability limit extends to 12 fs and the speedup factor to 10. For the solvated systems, a significant energy drift is found for Newtonian integrations with $\Delta t>8$ fs; for 6 fs $<\Delta t<8$ fs an energy drift is apparent only after several hundred ps of simulation (e.g., heating by $\sim$10° per ns); for $\Delta t\leqslant 6$ fs, the

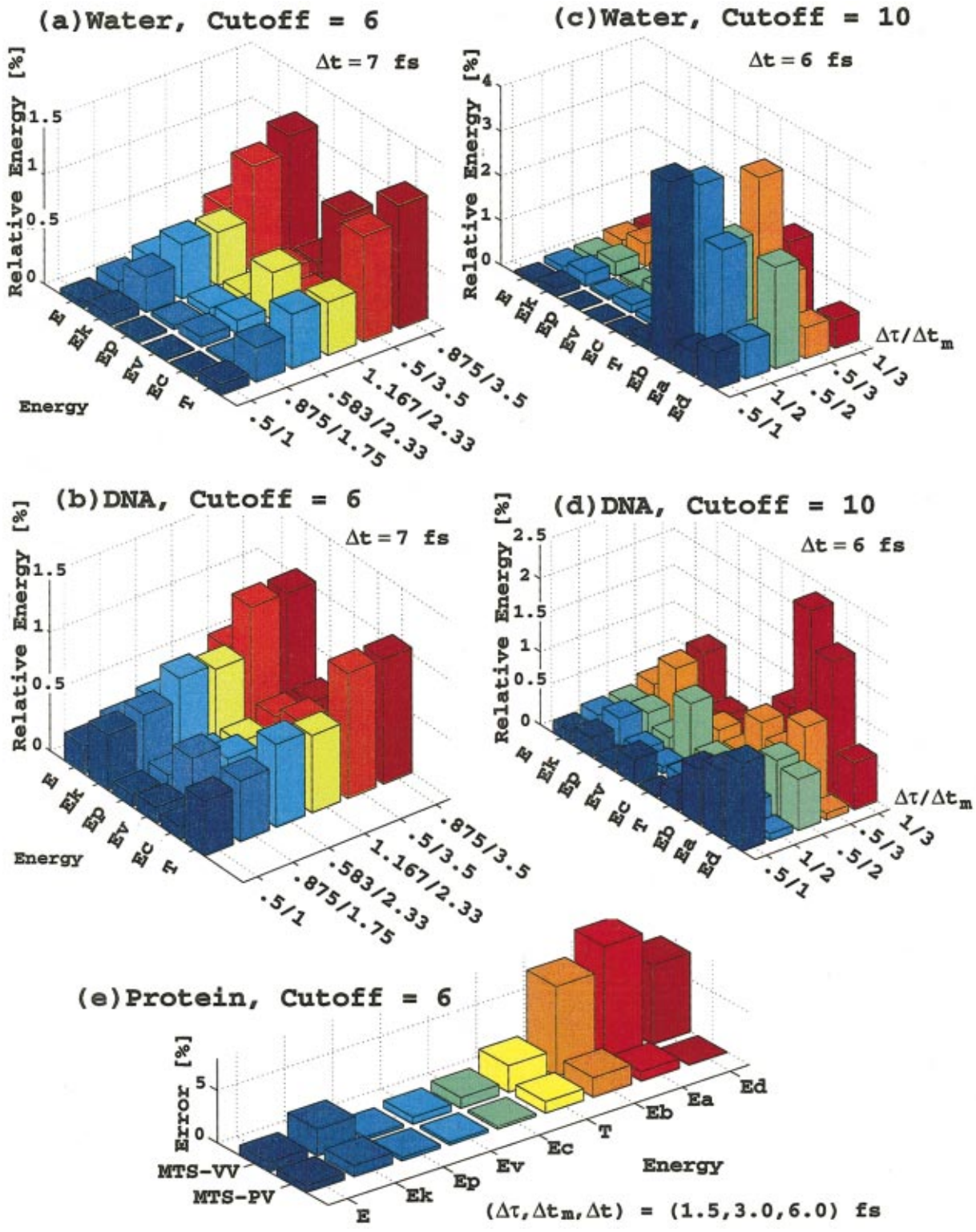

FIG. 3. (Color) The deviation in energy components relative to the baseline STS-PV integrator for the water and solvated biomolecule systems. The total energy (E) and its components, kinetic energy (Ek), potential energy (Ep), van der Waals energy (Ev), electrostatic energy (Ec), bond energy (Eb), bond angle energy (Ea), and dihedral energy (Ed) are plotted along with the temperature (T) percent differences. Newtonian MTS integrations with cutoffs=6 Å (a), and 10 Å (b), for $\Delta t=7$ fs are analyzed for the water system, and cutoffs=6 Å (c), and 9 Å (d), for $\Delta t=6$ fs for the solvated DNA system. At bottom, a comparison is made between the Newtonian MTS-PV and MTS-VV integrations for the solvated protein system (e); a cutoff=6 Å and time steps of $\Delta t=6$ fs, $\Delta t_m=3$ fs, and $\Delta\tau=1.5$ fs are used in both cases, with integration lengths were 500 ps.

integration exhibited no evidence of instability (heating) over several ns of simulation lengths. For the water system, stability of the Newtonian MTS protocol was retained up to $\Delta t=7$ fs, for a long-term simulation, and moderate Langevin damping ($\gamma=5.0$ ps^{-1}) extended this stability barrier to $\Delta t=12$ fs. The water system is special since the application of constrained dynamics on bonds involving hydrogen eliminates all bonded forces for the water molecules. Therefore,

TABLE IV. Performance for symmetric Newtonian MTS schemes for the solvated DNA system for various PME partitions at a fixed outer time step ($\Delta t=6$ fs). For each MTS scheme, we present the %Recip,%Direct values; the ratio of CPU relative to a STS scheme ($\Delta t=0.5$ fs, $N_{\text{grid}}=80$, cut-off = 6 Å), with the value in parentheses indicating the CPU ratio divided by 2.0 for comparison to an STS scheme with a 1.0 fs time step; and extrapolated clock time in days for a 1 ns simulation on a single MIPS R12000 processor (300 MHZ).

k_2,k_1	Δt_m	$\Delta\tau$	%Recip, %Direct	Speedup	CPU for 1 ns
			(CUT=6, $N_{\text{grid}}=80$)		
6,2	1.0 fs	0.5 fs	28,50	4.6 (2.30)	7.6 days
3,4	2.0	0.5	35,40	6.4 (3.20)	5.4
3,2	2.0	1.0	37,43	6.8 (3.40)	5.1
2,3	3.0	1.0	44,35	7.3 (3.65)	4.8
2,6	3.0	0.5	41,33	6.8 (3.40)	5.1
			(CUT=9, $N_{\text{grid}}=50$)		
6,2	1.0 fs	0.5 fs	4,82	2.2 (1.10)	16.0 days
3,4	2.0	0.5	7,73	3.4 (1.70)	10.2
3,2	2.0	1.0	7,75	3.5 (1.75)	9.9
2,3	3.0	1.0	8,70	4.4 (2.20)	7.9
2,6	3.0	0.5	8,67	4.1 (2.05)	8.6

the inner force partition within the MTS splitting gives no contribution, as well as no secondary coupling effects, to the medium and slow terms. This lack of bonded forces makes the water calculation more stable than the solvated biomolecule systems.

Table III presents the MTS performance measures for various β protocols for the water system. The results highlight the importance of adjusting the PME coefficient β to optimize the speedup ratio of the MTS relative to the STS scheme. Table III indicates that a factor of 3 decrease in CPU

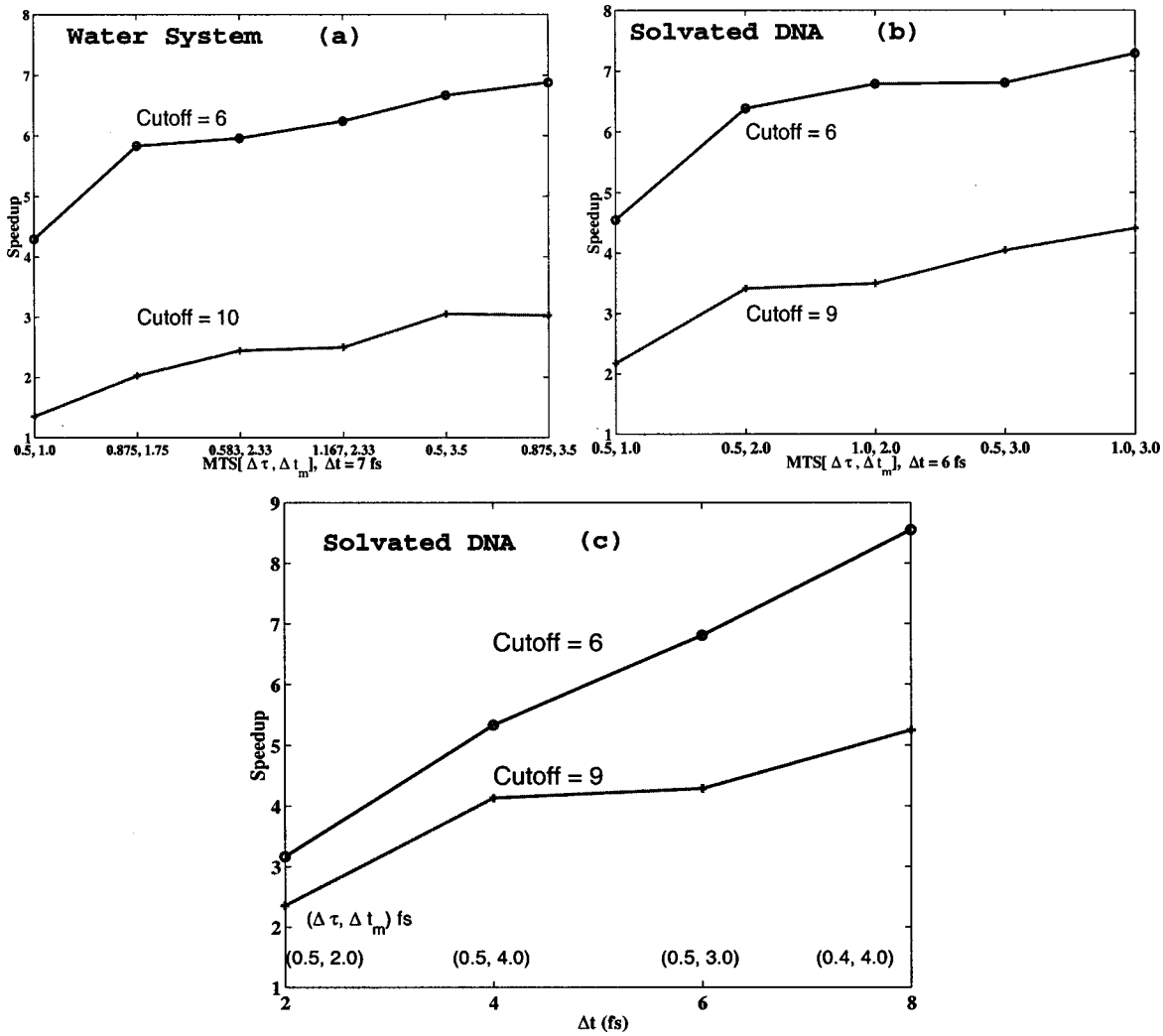

FIG. 4. Newtonian CPU speedups relative to STS-PV (cutoff=6 Å, $\Delta t=0.5$ fs) for the water and solvated DNA systems. Newtonian MTS integrations with cutoffs=6 and 10 Å for a $\Delta t=7$ fs are analyzed for the water system (a); cutoffs=6 and 9 Å for a $\Delta t=6$ fs are analyzed for the solvated DNA system (b); and cutoffs=6 and 9 Å for the DNA are studied as a function of Δt by maximizing the medium time step.

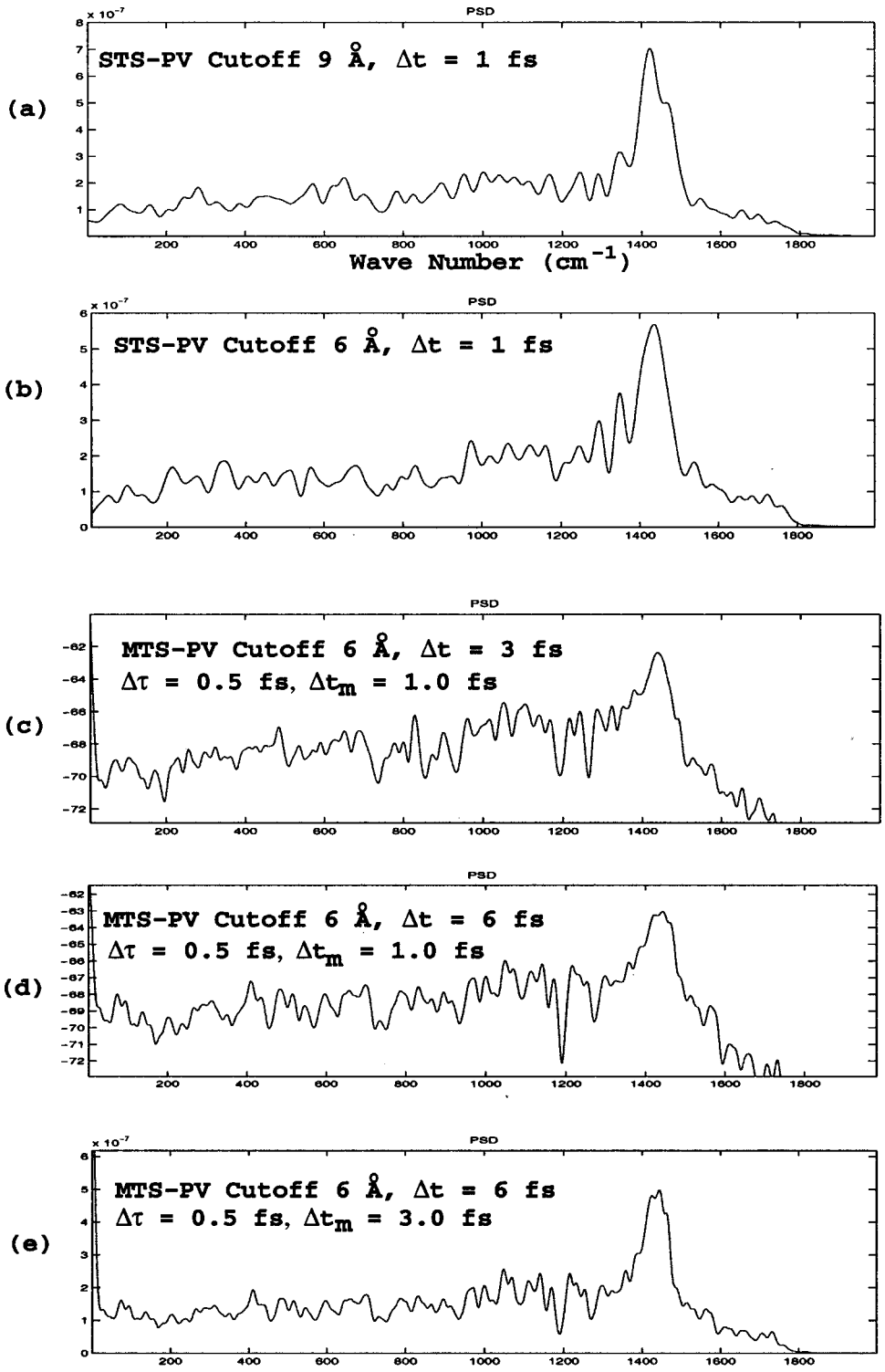

FIG. 5. The discrete Fourier transform of the velocity autocorrelation for the solvated protein simulation as computed by STS simulations at two cut-offs (a) and (b) and 3 MTS protocols (c)–(e). The 500 ps Newtonian MTS-PV simulations used a cut-off value of 6 Å, an inner time step of $\Delta\tau=0.5$ fs, and outer and medium time step values of $(\Delta t_m,\Delta t)=(1.0, 3.0)$ fs, (1.0, 6.0) fs, and (3.0, 6.0) fs. The 40 ps STS-PV simulation used $\Delta t=1$ fs and cut-off values of 6 and 9 Å.

time can be obtained in the MTS implementation by reducing the cut-off value from 10.0 to 6.0 Å. The largest speedup ratio corresponds to the largest Δt_m and is reached at approximately an equal work partition between the medium and slow terms. The deviation of the MTS energy components relative to the STS ($\Delta t=0.5$ fs) PV algorithm is shown in Figs. 3(a) and 3(b) for the cut-off values of 6.0 and 10.0 Å. In general, all energy components for the MTS scheme with $\Delta t\leqslant 7$ fs are stable for long simulations (several nanoseconds) and exhibit deviations of a few percent from the STS results in the individual energy components.

Next, we study the solvated DNA duplex in a periodic box with side dimensions of (45,45,50) Å. Two PME force partitions were studied in Table IV: A cutoff of 9.0 Å and $N_{\text{grid}}=(45,45,50)$ Å, and a cutoff of 6.0 Å with $N_{\text{grid}}=(80,80,80)$. We see that for $\Delta t=6$ fs the lower cut-off value of 6 Å is optimal when compared to the larger cut-off value. The MTS2 partition of Ref. 34 is similar to our Algorithm I, but had cut-off values of 10–12 Å; the study in Ref. 35 had cut-off values of 13 Å.

TABLE V. Performance for symmetric Newtonian MTS schemes for the water system for various PME partitions at a fixed outer time step ($\Delta t=7$ fs); here partition II is analyzed in detail. For each MTS scheme, we present the % Recip, % Direct values; the ratio of CPU relative to a STS scheme ($\Delta t=0.5$ fs, $N_{\text{grid}}=80$, cutoff $=6$ Å), with the value in parentheses indicating the CPU ratio divided by 2.0 for comparison to an STS scheme with a 1.0 fs time step; and extrapolated clock time in days for a 1 ns simulation on a single MIPS R12000 processor (300 MHZ).

k_2, k_1	Δt_m	$\Delta\tau$	%Recip,%Direct	Speedup	$t=1$ ns
7,2	1.0 fs	0.5 fs	21,58	4.3 (2.15)	8.4 days
4,2	1.75	0.875	28,50	5.8 (2.90)	6.2
3,4	2.33	0.583	33,41	6.0 (3.00)	6.0
3,2	2.33	1.167	34,43	6.2 (3.10)	5.8
2,7	3.5	0.5	40,32	6.7 (3.35)	5.4
2,4	3.5	0.875	36,38	6.6 (3.45)	5.2

The errors in the MTS energy components for the solvated DNA system relative to the STS ($\Delta t=0.5$ fs) PV algorithm in Figs. 3(c) and 3(d) show deviations of only a few percent in the individual energy components. For the combination $\Delta t=8$ fs, $\Delta\tau=2$ fs, and $\Delta t_m=4$ fs, the simulation is stable for up to 100 ps of integration, with only a 3% deviation from the STS reference temperature of 300 K. However, trajectories longer than 500 ps indicate heating. Analysis revealed that the heating did not result from the relatively low 6 Å cutoff, but from the 8 fs outer time step, too large for the Newtonian integrator. An outer time step stability limit of 8 fs was also noted with the MTS-VV-RESPA integrator[34] for simulation lengths of 20 ps.

Stable Newtonian trajectories require $\Delta t\leqslant 6$ fs in our applications. A 4 fs outer time step was found optimal for the RESPA integrator with direct space cut-off values of 13 Å in Ref. 35 based on 100 ps simulations; however, that study also reported substantial energy drifts for a $\Delta t=4$ fs RESPA simulation.

Our Newtonian MTS–VV and MTS–PV integrators were compared at the combination $\Delta t=6$ fs, $\Delta t_m=3$ fs, and $\Delta\tau=1.5$ fs; the results indicate that both integrators are stable over several nanoseconds. However, the averaged kinetic energy and bonded terms for the MTS-VV version indicate significantly greater errors when compared to the baseline STS results. Figure 3(e) plots the distribution of error for the two integrators for the various components of the total energy; given a threshold of less then 5%, we consider the $\Delta t=6$ fs MTS-PV acceptable whereas the MTS-VV is not.

For the water system, the use of constrained dynamics on all bonds involving hydrogen pushes the stability limit to $\Delta t=7$ fs (compared to 6 fs for the solvated biomolecules). Figure 4(a) plots speedup for various combinations of $\Delta\tau$ and Δt_m for the water system with a fixed $\Delta t=7$ fs for the two cut-off values of 6 and 10.0 Å. For both cut-off values, the speedup ratio calculations are based on the cutoff of 6 Å. The speedup has an asymptotic limit dictated by the stability constraints on both Δt and Δt_m. Results for the solvated DNA in Fig. 4(b) with $\Delta t=6$ fs indicate similar limits. Another view in Fig. 4(c) shows the speedup of the MTS integrator as a function of the outer time step; here the medium time step is maximized for each calculation.

In general, Algorithm I with a 6.0 Å cut-off value offers approximately a factor of 2 to 3 enhanced speedup over a 9 Å cut-off value; similar results were found for the solvated protein system. For cut-off values greater than 12 Å, although nonoptimal in an MTS protocol, we have found that the outer time step stability improved to above 20 fs; here simulation lengths up to 100 ps were examined. These larger cutoffs imply considerably smaller reciprocal force contributions to the total energy and force work, as well as considerably larger nonbonded pairlists. We are currently investigating designing optimal core functions to apply to the Coulomb terms so as to enhance the outer time step stability at lower cut-off values.[44]

To measure performance for a given MTS-PME force partition, we next survey various combinations of k_1 and k_2 for a cut-off value of 6.0 Å and a mesh density N_{grid} of 80 points in each spatial direction (Partition II of Table I) for the water system. All simulations employ $\Delta t=7$ fs and a simulation length of 20 ps. The qualitative results presented here are expected to be independent of the actual system studied and are dependent only on the number of atoms, on Δt, and on the force partition. The stability boundary for Δt was found to be insensitive to the choices of k_1 and k_2. Table V shows that varying the inner time step ($\Delta\tau$), for a given Δt, has little effect on the optimal MTS scheme, and that lowering k_2 is the most sensitive factor.

To measure the influence of the MTS protocol and lower cut-off values on the dynamics, we present the discrete Fourier transform of the autocorrelation signal in Fig. 5 for the solvated protein for MTS and STS simulations. Here MTS protocols with time steps of $(\Delta\tau,\Delta t_m,\Delta t)=(0.5, 1.0, 3.0)$ fs, (0.5, 1.0, 6.0) fs, and (0.5, 3.0, 6.0) fs are compared to STS schemes (PV) with $\Delta t=1.0$ fs. The MTS spectra show good agreement with the STS results, and the data for the cut-off values of 9 and 6 Å are similar. Unlike the spectra in Ref. 34 for the MTS-VV RESPA integrator, which indicate an upward drift in the lower 2000 cm^{-1} components at large Δt, we see no evidence of an upward drift in the frequency components for the Newtonian MTS-PV integrator studied here.

The ratio ΔE of the total energy rms deviation to the mean total energy as a function of Δt is shown in Fig. 6(a) for the solvated DNA system. For all our MTS protocols, we find the ΔE values to be 10^{-3} and lower; for reference, Cheng and Merz[34] suggested 10^{-2} (1.0 %) as the threshold ratio. Our Algorithm I based on PV appears to have lower relative fluctuations in total energy versus the MTS-VV-

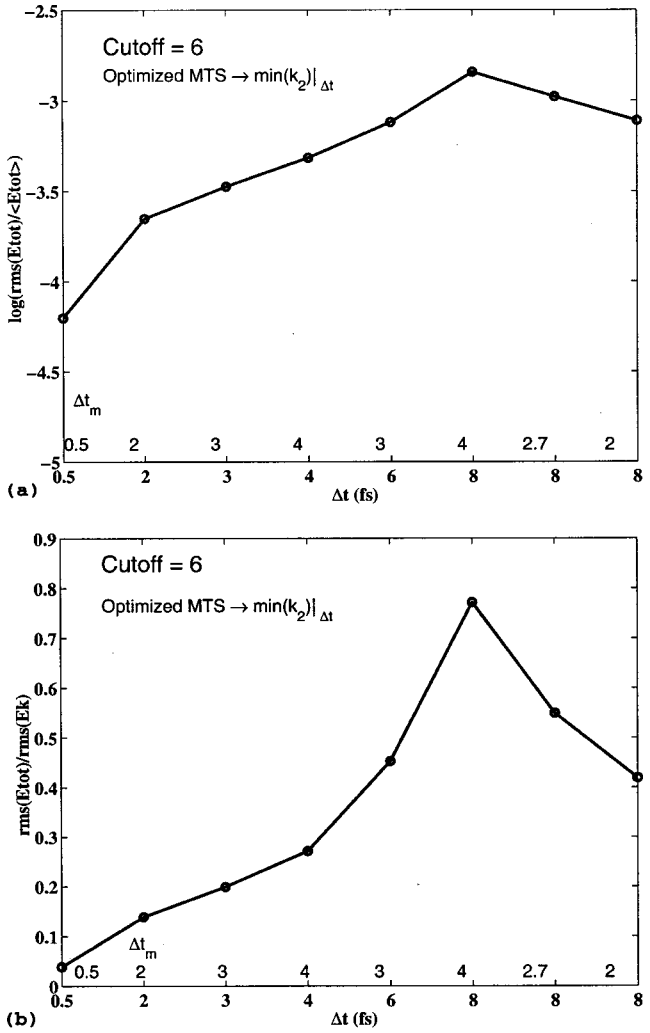

FIG. 6. rms fluctuation ratios for a Newtonian integration of the solvated DNA system with a cutoff=6 Å as a function of Δt: (a) total energy rms relative to mean total energy, and (b) total energy rms to kinetic-energy rms.

RESPA formulation of Ref. 34; this agrees with prior observations[16] regarding MTS-PV versus MTS-VV schemes at larger time steps. Our Newtonian studies indicate a 30% lower rms in the total energy for the PV-MTS integration versus the VV-MTS result.

Finally, the ratio of the rms of the total energy fluctuation relative to the rms of the kinetic energy fluctuation $\Delta E^{\mathrm{rms}}=\Delta E_{\mathrm{total}}/\Delta E_{\mathrm{kin}}$ is studied for the solvated DNA system. Figure 6(b) shows values less than 0.5 for most protocols used with the solvated DNA system; these protocols agree to within a 3% deviation of their mean energy components relative to the STS results. Watanabe and Karplus [17] suggested that the magnitude of this ratio should be $\mathcal{O}(10^{-2})$ for a stable integrator (neat liquids of *n*-alkanes were studied). Humphreys *et al.*[13] reported this ratio for biomolecular simulations, where values less than 0.1 were typical. However, Procacci *et al.*[32,33] suggested that the larger ΔE^{rms} values for solvated biomolecular systems modeled by PME, and integrated with MTS, do not reflect integrator quality well. They showed the ratio to be sensitive to the Ewald β coefficient; namely, a larger β leads to a larger ΔE^{rms} value. A formal measure of MTS integrator quality remains an open problem; we therefore, choose to examine various details of the simulation, including mean energies, spectral content, and pdf distributions. Our 500 ps simulation of solvated biomolecules with $\Delta E^{\mathrm{rms}}\sim 0.4$ indicates no energy drifts with the Newtonian MTS-PV protocol, and the rdf and spectral comparisons are in good agreement with STS results.

V. LANGEVIN DYNAMICS

Barth and Schlick[10,30] and Sandu and Schlick[8] have suggested the use of Langevin coupling as a device to damp numerical resonances associated with symplectic MTS protocols. Izaguirre *et al.*[51] followed by using mild Langevin damping to stabilize the integration at larger time steps. To assess the effects of the Langevin components of Algorithm I, we studied the variation of outer and medium time steps with a Langevin damping parameter of $\gamma=0.25-5.0\ \mathrm{ps}^{-1}$, typical or smaller than values used in LN protocols.[31] As expected, Langevin coupling enhances stability as well as speedup relative to the Newtonian schemes; in general, speedup ratios can be increased from 8.5 to 10.0, when compared to STS ($\Delta t=0.5$ fs) simulations since the outer time step increases to about 12 fs. Typically, the STS results were reproduced to within a 5% difference in all energy components for [Figs. 7(a) and 7(b)].

To contrast the MTS-VV and MTS-PV protocols, with Langevin forces, we present a comparison of the total energy components to the STS-PV mean energies in Fig. 7(c). The results in Fig. 7(c) are in basic agreement with the Newtonian results for $\Delta t=6$ fs, indicating that MTS-PV yields a smaller error relative to the STS results. For both schemes, a Langevin damping with $\gamma=0.25\,\mathrm{ps}^{-1}$ is sufficient to stabilize the long-term heating of the solvated protein system, and the $\Delta t=8$ fs MTS simulation is stable.

For the PME protocols, speedup ratios beyond 10 are possible by either raising the outer time step, with respect to a constant medium term work, or raising the stability limit on the medium time step (Fig. 8). We also found that the CPU time is relatively insensitive to inner time steps ranging from 0.5 to 1.0 fs, in agreement with the Newtonian results.

VI. SUMMARY

We have implemented an efficient multiple time step force splitting scheme for biological applications that combines a symmetric Trotter factorization of the Liouville operator with the particle-mesh Ewald method in the widely used AMBER program. Our algorithm, implemented for both Newtonian and Langevin dynamics and tested on a large water system and two solvated biomolecules, offers two new numerical ingredients: the position-Verlet scheme rather than velocity Verlet, and optimized Ewald parameters for the MTS protocol. The PV scheme offers stability advantages at large medium time steps, as found empirically for nonlinear functions and theoretically for a one-dimensional harmonic oscillator.[37] The Ewald optimization of the screening parameters β affects resulting speedups substantially. Resulting MTS/PME speedups relative to a 1 fs single-step Verlet algorithm are over 3 for Newtonian dynamics and 5 with Langevin coupling with $\gamma=5.0\ \mathrm{ps}^{-1}$. These speedup factors double relative to 0.5 fs STS simulations and thus reflect

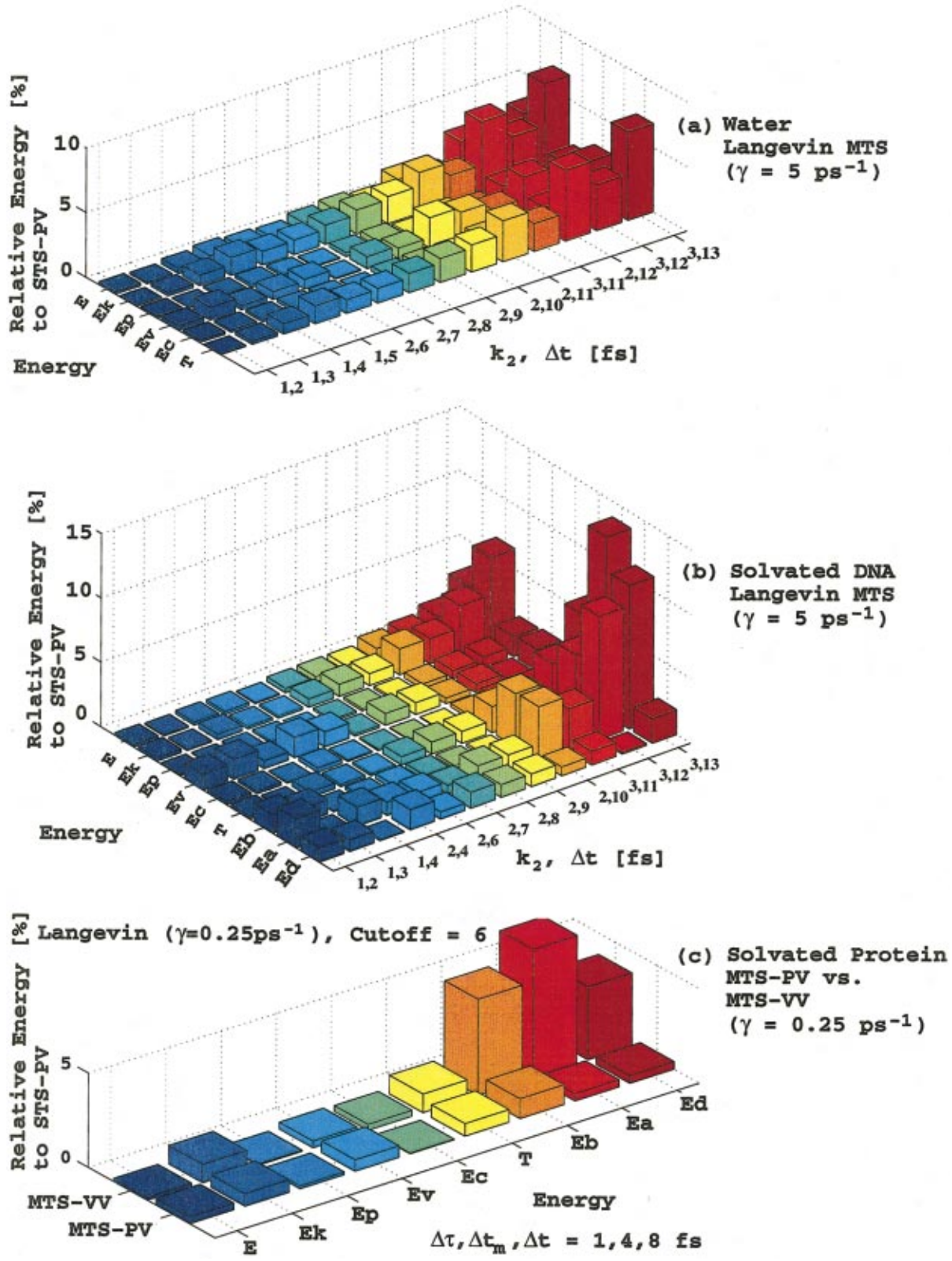

FIG. 7. (Color) Langevin MTS studies relative to STS-position-Verlet at $\Delta t=1$ fs. The deviation in energy components relative to the baseline STS-PV (cutoff=6 Å) scheme for the water (a) and solvated DNA (b) system are shown ($\gamma=5.0$ ps^{-1}), as well as the deviations for the solvated protein (c) for MTS-position-Verlet and MTS velocity-Verlet ($\gamma=0.25$ ps^{-1}, cutoff=6 Å, $\Delta t=8.0$ fs, and $\Delta t_m=4.0$ fs). All integration lengths were 500 ps. See Fig. 3 caption for abbreviation of energy terms.

substantial improvement over AMBER implementations to date. The stability limits are $\Delta t=6$ fs for Newtonian dynamics and $\Delta t=12$ fs with Langevin coupling under the current splitting protocol.

Overall, we have found that a position-Verlet version of the symmetric Trotter factorization has favorable integration properties, particularly at larger time steps. We have also demonstrated that lowering the Ewald cut-off value from 10

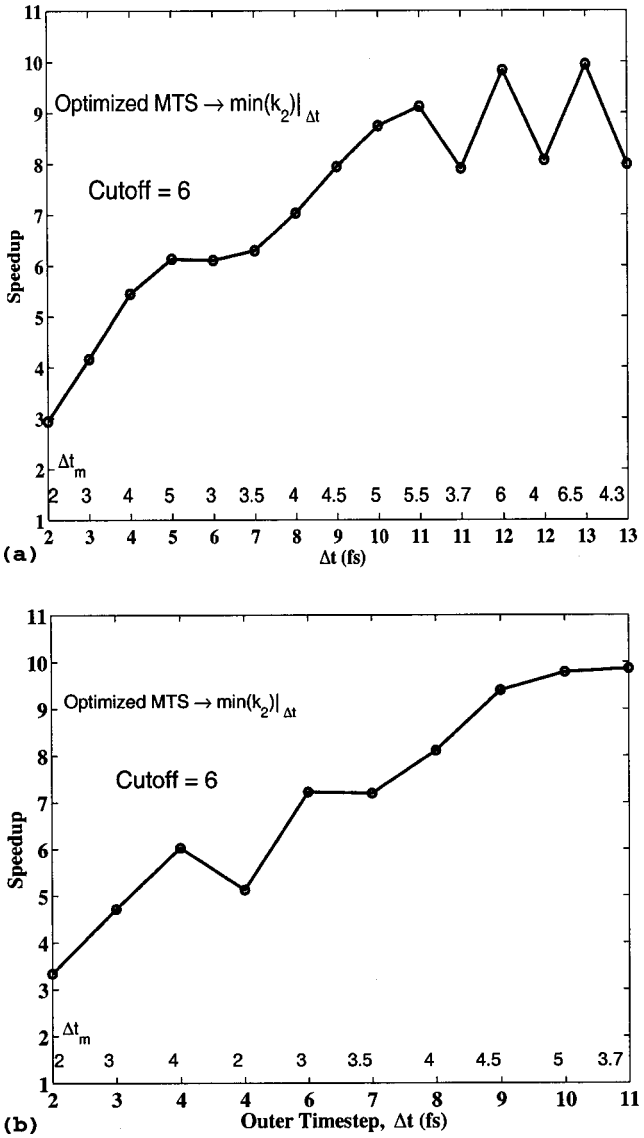

FIG. 8. Langevin CPU speedups relative to the baseline STS-PV(cutoff= 6 Å) scheme for the water (a) and solvated DNA (b) systems for cutoff=6 Å and (γ=5.0 ps^{-1}).

to 6 Å results in enhanced MTS speedup ratios by nearly a factor of 3. The most surprising but interesting finding in this work is the relatively small outer time step limits in both Newtonian and Langevin MTS/PME protocols. This is unlike the MTS-Langevin LN scheme which allows outer time steps of order 100 fs in the absence of PME summations (Barth and Schlick found stable time steps up to 120 fs or more[10,30]).

As a remedy, we are investigating modified Ewald core functions which alleviate the fast, near-field particle interactions, in the reciprocal term,[44] and other splitting approaches.[43] Such new Ewald-type splittings are expected to offer advantages for MTS splitting strategies. In addition, we also suggest increasing the van der Waal cut-off value by using a switch function for the van der Waals interactions and maintaining a second pairlist; applying the PME method to the van der Waals terms is an alternative.[43]

ACKNOWLEDGMENTS

This article is dedicated to Peter Kollman, who will be dearly missed; Gordon conferences will never be the same without him. The authors thank Dr. Matthew Young and Dr. John Kuriyan of Rockefeller University for providing the AMBER setup files for the solvated protein system and Ms. Rebecca Perlow and Dr. Suse Broyde of New York University for the solvated DNA files. The work is supported by NSF award ASC-9318159, NIH award R01 GM55164, and a John Simon Guggenheim fellowship to one of the authors (T.S.), an investigator of the Howard Hughes Medical Institute.

[1] T. Schlick, in *Lecture Notes in Computational Science and Engineering*, edited by P. Deuflhard, J. Hermans, B. Leimkuhler, A. E. Mark, S. Reich, and R. D. Skeel (Springer-Verlag, Berlin/New York, 1998).
[2] T. Schlick, Structure **9**, R45 (2001).
[3] T. Schlick, A. T. Brunger, L. V. Kale, J. A. Board, Jr., J. Hermans, and K. Schulten, J. Comput. Phys. **151**, 9 (1999).
[4] T. E. Cheatham, III and B. R. Brooks, Theor. Chem. Acc. **99**, 279 (1999).
[5] T. Darden, D. York, and L. Pedersen, J. Chem. Phys. **98**, 10089 (1993).
[6] J. Norberg and L. Nilsson, Biophys. J. **79**, 1537 (2000).
[7] M. P. Allen and D. J. Tildesley, *Computer Simulations of Liquids* (Oxford University Press, Oxford, 1989).
[8] A. Sandu and T. Schlick, J. Comput. Phys. **151**, 74 (1999).
[9] T. Schlick, E. Barth, and M. Mandziuk, Annu. Rev. Biophys. Biomol. Struct. **26**, 181 (1997).
[10] T. Schlick and E. Barth, J. Chem. Phys. **109**, 1617 (1998).
[11] J. J. Biesiadecki and R. D. Skeel, J. Comput. Phys. **109**, 318 (1993).
[12] H. Grubmuller, H. Heller, A. Windemuth, and K. Schulten, Mol. Simul. **6**, 121 (1991).
[13] D. D. Humphreys, R. A. Friesner, and B. J. Berne, J. Phys. Chem. **98**, 6885 (1994).
[14] J. A. Izaguirre, S. Reich, and R. D. Skeel, J. Phys. Chem. **110**, 9853 (1999).
[15] M. E. Tuckerman, B. J. Berne, and G. J. Martyna, J. Chem. Phys. **94**, 6811 (1991).
[16] M. E. Tuckerman, B. J. Berne, and G. J. Martyna, J. Chem. Phys. **97**, 1990 (1992).
[17] M. Watanabe and M. Karplus, J. Chem. Phys. **99**, 8063 (1993).
[18] L. Greengard and V. Rokhlin, Acta Numerica **6**, 2298 (1998).
[19] H. G. Peterson, D. Soelvason, J. W. Perram, and E. R. Smith, J. Chem. Phys. **101**, 8870 (1994).
[20] P. P. Ewald, Ann. Phys. (Leipzig) **64**, 253 (1921).
[21] R. W. Hockney and J. W. Eastwood, *Computer Simulation Using Particles* (IOP Publishing Ltd., London, 1989).
[22] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, J. Chem. Phys. **103**, 8577 (1995).
[23] T. Darden, L. Perera, L. Leping, and L. Pedersen, J. Chem. Phys. **7**, R55 (1999).
[24] D. York and W. Yang, J. Chem. Phys. **101**, 3298 (1994).
[25] M. Deserno and C. Holm, J. Chem. Phys. **109**, 7678 (1998).
[26] B. A. Luty, I. G. Tironi, and W. F. van Gunsteren, J. Chem. Phys. **103**, 3014 (1995).
[27] M. Kawata and M. Mikami, J. Comput. Chem. **21**, 201 (2000).
[28] W. D. Cornell, P. Cieplak, C. I. Bayly *et al.*, J. Am. Chem. Soc. **117**, 5179 (1995).
[29] D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatman, S. DeBolt, D. Ferguson, G. Seibel, and P. Kollman, Comput. Phys. Commun. **91**, 1 (1995).
[30] E. Barth and T. Schlick, J. Chem. Phys. **109**, 1633 (1998).
[31] D. Strahs and T. Schlick, J. Mol. Biol. **301**, 643 (2000).
[32] P. Procacci and M. Marchi, J. Chem. Phys. **104**, 3003 (1996).
[33] P. Procacci, T. Darden, and M. Marchi, J. Phys. Chem. **100**, 10464 (1996).
[34] A. Cheng and K. M. Merz, Jr., J. Phys. Chem. B **103**, 5396 (1999).
[35] Y. Komeiji, J. Mol. Struct.: THEOCHEM **530**, 237 (2000).
[36] P. E. Smith and B. M. Pettitt, Comput. Phys. Commun. **91**, 339 (1995).
[37] P. F. Batcho and T. Schlick J. Chem. Phys. **115**, 4019 (2001), following paper.
[38] H. F. Trotter, Proc. Am. Math. Soc. **10**, 545 (1959).

[39] G. J. Martyna, M. E. Tuckerman, D. J. Tobias, and M. L. Klein, Mol. Phys. **87**, 1117 (1996).
[40] D. M. Heyes, J. Chem. Phys. **74**, 1924 (1981).
[41] T. Schlick, Molecular Modeling: An Interdisciplinary Guide, Chapter 9 (in preparation, New York University, 2001).
[42] P. Procacci, M. Marchi, and G. J. Martyna, J. Chem. Phys. **108**, 8799 (1998).
[43] X. Qian and T. Schlick (work in progress).
[44] P. F. Batcho and T. Schlick (unpublished).
[45] V. Natoli and D. M. Ceperley, J. Comput. Phys. **117**, 171 (1995).
[46] R. Perlow and S. Broyde (in preparation).
[47] F. Sicheri, I. Moarefi, and J. Kuriyan, Nature (London) **385**, 602 (1997).
[48] T. Schindler, F. Sicheri, A. Pico, A. Gazi, A. Levitzki, and J. Kuriyan, Molecular Cell **3**, 639 (1999).
[49] M. Kawata and M. Mikami, Chem. Phys. Lett. **313**, 261 (1996).
[50] N. Karasawa and W. A. Goddard, III, J. Phys. Chem. **93**, 7321 (1989).
[51] J. A. Izaguirre, D. P. Catarello, J. M. Wozniak, and R. D. Skeel, J. Phys. Chem. **114**, 2090 (2001).