

Error estimate of short-range force calculation in inhomogeneous molecular systems

Han Wang*

LMAM and School of Mathematical Sciences, Peking University and Institute for Mathematics, Freie Universität Berlin, Germany

Christof Schütte

Institute for Mathematics, Freie Universität Berlin, Germany

Pingwen Zhang

LMAM and School of Mathematical Sciences, Peking University

(Received 1 May 2012; published 8 August 2012)

In the present paper, we develop an accurate error estimate of the nonbonded short-range interactions for the inhomogeneous molecular systems. The root-mean-square force error is proved to be decomposed into three additive parts: the homogeneity error, the inhomogeneity error, and the correlation error. The magnitude of the inhomogeneity error, which is dominant in the interfacial regions, can be more than one order of magnitude larger than the homogeneity error. This is the reason why a standard simulation with fixed cutoff radius is either less accurate if the cutoff is too small, or wastes considerable computational effort if the cutoff is too large. Therefore, based on the error estimate, the adaptive cutoff and long-range force correction methods are proposed to boost the efficiency and accuracy of the simulation, respectively. The way of correcting the long-range contribution of pressure is also developed for the inhomogeneous system. The effectiveness of the proposed methods is demonstrated by molecular dynamics simulations of the liquid-vapor equilibrium and the nanoscale particle collision. Different roles of the homogeneity error and inhomogeneity error are also discussed.

DOI: [10.1103/PhysRevE.86.026704](https://doi.org/10.1103/PhysRevE.86.026704)

PACS number(s): 02.70.-c, 64.75.Gh, 64.70.F-, 82.30.Nr

I. INTRODUCTION

Nonbonded short-range interactions are encountered in nearly every molecular simulation. A naive idea for calculating short-range interactions is to explicitly calculate all pairwise interactions. This results in a computational cost scaling $O(N^2)$ per time step, which becomes rapidly inefficient as the number of particles N grows. A better way is to introduce a cutoff radius, outside of which all pairwise interactions are simply neglected. In combination with the cell list and the neighbor list algorithms [1], the total computational cost of the short-range interactions can be reduced to an acceptable level of $O(N)$. The cutoff radius neglects all pairwise interactions outside the cutoff radius, so it introduces an error in the interaction calculation. Since the cutoff system can be viewed as an approximation of the original system, it is very important to test the convergence of properties of interest as a function of the cutoff radius.

Fortunately, in most homogeneous systems, a satisfactory convergence can be achieved even with a very small cutoff value, for example, 2.5σ for the standard Lennard-Jones 6-12 interaction. Moreover, it is possible to correct the systematic error of the potential energy, the pressure, and the free energy by applying the standard long-range correction (LRC) [2], which assumes the uniformity outside the cutoff radius and integrates the contribution of the mentioned properties to infinity. Instead for inhomogeneous systems, in which the density of a certain type of atom may change within the length scale of a few molecules, and thus lack of uniformity, the standard molecular simulation by the cutoff method will encounter nontrivial difficulties. The simulation results are

found to converge only at a very large cutoff radius value [3,4] and also depend on the energy (force) continuity at the cutoff and the way of applying the long-range corrections [5–10]. Therefore, it is necessary to systematically study the precision of the short-range interaction calculation and to develop methods for screening out the aforementioned cutoff artifact in the inhomogeneous systems.

One possible way to avoid the cutoff difficulty is to modify the short-range interaction by a shifting approach: Inside a certain radius the original interaction is kept, while outside it the interaction is screened by a shift function, so that it goes to zero smoothly and rapidly. The amount of shifting is minor in most cases; however, it actually changes the way of modeling molecules. For example, it has been shown that shifting the dispersion remarkably changes the phase diagram of the Lennard-Jones system [11]. Since the dispersion term can be derived by quantum mechanical arguments, and has the physical reality, the unshifted interaction is preferred and investigated by most studies (see, e.g., Refs. [12,13]).

A promising way to quantitatively analyze the undesirable cutoff effects is to express these artifacts in terms of the difference between the cutoff interaction and the exact interaction, namely, the error. In homogeneous systems, the error analysis of the cutoff method has been well established; see Refs. [14,15], for example. Also the error estimates [14,16–19] for long-range Ewald family methods [20–24] have provided a profound understanding of the accuracy of force calculation and introduced parameter tuning algorithms [19,25] that boost the efficiency of the computation with a good control of the error. However, in inhomogeneous systems the error study is still scarce, even for the simplest cutoff method.

In this work we develop the error estimate of the short-range interaction and the long-range force and pressure corrections for the inhomogeneous systems. We prove that the force

*han.wang@fu-berlin.de

error comprises a homogeneity error and an inhomogeneity error, regardless of the particle-particle correlation outside the cutoff radius. In an inhomogeneous system composed of several bulk regions connected by interfacial regions where the density changes rapidly, an important observation is that the inhomogeneity error plays a dominant role in the interfacial regions and can be more than one order of magnitude larger than the error in the bulk regions. An adaptive cutoff method is hereby proposed to equally distribute the error across the simulation region, so that a considerable computational effort can be saved in the bulk regions while the precision of simulation is preserved. We also proposed a long-range force correction (LFC) method to screen out the contribution of the inhomogeneity error; therefore, the precision of simulation is improved without much extra computational load.

To show the effectiveness of the error estimate and the proposed methods, two systems are tested: the liquid-vapor equilibrium for equilibrium properties and the colliding nanoscale clusters for dynamical properties. In these examples we investigate the Lennard-Jones interaction, which is one of the most widely used in molecular simulations. However, the error estimate, the adaptive cutoff, and long-range correction methods are not restricted to the Lennard-Jones interaction: It can be applied to other types of short-range interactions without difficulty, such as the Buckingham interaction, the direct part of Ewald family algorithms [20,22–24], etc. We want to stress that the main contribution and the research interest are the quantitative study of the force computation error by cutting off the short-range interaction. The examples, which are comparatively simple and ideal, are used only for testing the validity of the error estimate and the proposed methods, and demonstrating the role of different errors in the simulations. The error study in more realistic systems requires the estimates of the electrostatic interaction, which are calculated by the long-range algorithms, such as PME [22], SPME [23], P3M [24], etc. The scope of the present paper is not enough for these algorithms, so they will be delivered in future work.

II. THEORETICAL BACKGROUND

A. Error estimate in an inhomogeneous system

In this paper, for simplicity, we will consider one-component systems, but the error estimate for the one-component system can be straightforwardly extended to the multi-component systems. We consider the force error in the present paper, because the applications are based on molecular dynamics simulations. Suppose the system is composed by N identical particles located at $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N$ with periodic boundary condition. These particles are interacting via a short-range pairwise interaction $u(\mathbf{r})$ with a cutoff radius $r_c > 0$. The corresponding force is denoted by $\mathbf{f}(\mathbf{r})$. Most short-range interactions satisfy $|u(r)| \leq Cr^{-m}$, $m > 3$, which guarantees an absolute convergence of the energy. Throughout the paper, the only assumptions made on the interaction are (1) the interaction is short range and (2) it is calculated by the cutoff method. Therefore, our approach works for a large group of widely used interactions, for example, the Lennard-Jones

interaction, the Buckingham interaction, the direct part of Ewald family algorithms, and so forth.

The force error stems from the neglected interaction outside the cutoff radius, so it is convenient to consider the complementary of the cutoff force that is defined by

$$\mathbf{f}^c(\mathbf{r}) = \begin{cases} 0, & |\mathbf{r}| \leq r_c; \\ \mathbf{f}(\mathbf{r}), & |\mathbf{r}| > r_c. \end{cases} \quad (1)$$

We measure the force error at an arbitrary position in the system, say, \mathbf{r} , by adding a testing particle to that position. We assume all particles in the system exert forces on the testing particle, whereas the testing particle does *not* exert any force on the system. The exact force and the computed force (or cutoff force) exerted by the system on the testing particle are denoted by $\mathbf{F}(\mathbf{r})$ and $\tilde{\mathbf{F}}(\mathbf{r})$, respectively. The difference between them, namely, the *error force*, is denoted by $\Delta\mathbf{F}(\mathbf{r}) = \mathbf{F}(\mathbf{r}) - \tilde{\mathbf{F}}(\mathbf{r})$, which can be expressed by

$$\Delta\mathbf{F}(\mathbf{r}) = \sum_{\mathbf{n}} \sum_j \mathbf{f}^c(\mathbf{r} - \mathbf{r}_j + \mathbf{n}), \quad (2)$$

where \mathbf{n} is the box vector, so the summation over \mathbf{n} indicates that all periodic images of the system are considered. In a homogeneous system, the surrounding of any particle is isotropic, so the *mean error force* vanishes: $\langle \Delta\mathbf{F}(\mathbf{r}) \rangle = 0$. In contrast, the mean error force is not necessarily zero in an inhomogeneous system:

$$\begin{aligned} \langle \Delta\mathbf{F}(\mathbf{r}) \rangle &= \left\langle \sum_{\mathbf{n}} \sum_j \mathbf{f}^c(\mathbf{r} - \mathbf{r}_j + \mathbf{n}) \right\rangle \\ &= \int_{\mathbb{R}^3} \mathbf{f}^c(\mathbf{r} - \mathbf{r}') \rho(\mathbf{r}') d\mathbf{r}', \end{aligned} \quad (3)$$

where $\rho(\mathbf{r})$ denotes the particle number density at the position \mathbf{r} and is periodically extended to \mathbb{R}^3 . Notice that the error force implicitly depends on the cutoff radius used for the interaction calculation.

The widely accepted definition of the force error is the *root mean square (RMS) error*, which is the square root of the second moment of the error force $\mathcal{E}(\mathbf{r}) = \sqrt{\langle |\Delta\mathbf{F}(\mathbf{r})|^2 \rangle}$. This is calculated as

$$\begin{aligned} \langle |\Delta\mathbf{F}(\mathbf{r})|^2 \rangle &= \left\langle \sum_{j,k} \mathbf{f}^c(\mathbf{r} - \mathbf{r}_j) \cdot \mathbf{f}^c(\mathbf{r} - \mathbf{r}_k) \right\rangle \\ &= \left\langle \sum_j |\mathbf{f}^c(\mathbf{r} - \mathbf{r}_j)|^2 \right\rangle + \left\langle \sum_{j \neq k} \mathbf{f}^c(\mathbf{r} - \mathbf{r}_j) \cdot \mathbf{f}^c(\mathbf{r} - \mathbf{r}_k) \right\rangle \\ &= \int_{\mathbb{R}^3} |\mathbf{f}^c(\mathbf{r} - \mathbf{r}')|^2 \rho(\mathbf{r}') d\mathbf{r}' \\ &\quad + \int_{\mathbb{R}^3 \times \mathbb{R}^3} \mathbf{f}^c(\mathbf{r} - \mathbf{r}') \cdot \mathbf{f}^c(\mathbf{r} - \mathbf{r}'') \rho(\mathbf{r}', \mathbf{r}'') d\mathbf{r}' d\mathbf{r}'', \end{aligned} \quad (4)$$

where $\rho(\mathbf{r}', \mathbf{r}'')$ is the pair density. By using the identity

$$\rho(\mathbf{r}', \mathbf{r}'') = \rho(\mathbf{r}')\rho(\mathbf{r}'') + C(\mathbf{r}', \mathbf{r}''), \quad (5)$$

where

$$C(\mathbf{r}', \mathbf{r}'') = [\rho(\mathbf{r}', \mathbf{r}'') - \rho(\mathbf{r}')\rho(\mathbf{r}'')] \quad (6)$$

denotes the density correlation between \mathbf{r}' and \mathbf{r}'' , Eq. (4) becomes

$$\begin{aligned} & \langle |\Delta \mathbf{F}(\mathbf{r})|^2 \rangle \\ &= \int_{\mathbb{R}^3} |\mathbf{f}^c(\mathbf{r} - \mathbf{r}')|^2 \rho(\mathbf{r}') d\mathbf{r}' \\ &+ \left[\int_{\mathbb{R}^3} \mathbf{f}^c(\mathbf{r} - \mathbf{r}') \rho(\mathbf{r}') d\mathbf{r}' \right]^2 \\ &+ \int_{\mathbb{R}^3 \times \mathbb{R}^3} \mathbf{f}^c(\mathbf{r} - \mathbf{r}') \cdot \mathbf{f}^c(\mathbf{r} - \mathbf{r}'') C(\mathbf{r}', \mathbf{r}'') d\mathbf{r}' d\mathbf{r}'' \\ &= \mathcal{E}_{\text{homo}}^2(\mathbf{r}) + \mathcal{E}_{\text{inhomo}}^2(\mathbf{r}) + \mathcal{E}_{\text{correlation}}(\mathbf{r}). \end{aligned} \quad (7)$$

If the system is homogeneous and the particle-particle correlation is neglected, then only the first term is left on the right-hand side of (7), and it is called the homogeneity error. This term originates from the fluctuation of the error force $\Delta \mathbf{F}(\mathbf{r})$. The second term is called the inhomogeneity error, because it stems from the density inhomogeneity of the system. It is worth pointing out that the inhomogeneity error is nothing but the magnitude of the mean error force, namely, $\mathcal{E}_{\text{inhomo}}(\mathbf{r}) = \langle |\Delta \mathbf{F}(\mathbf{r})| \rangle$. The last term on right-hand side of (7) originates from the correlation between the particles, so it is called the correlation error. The correlation error involves a double integral, which implies that the calculation requires more computational resources than the homogeneity and inhomogeneity errors. The rest of this paper assumes the

correlation error does not dominate in the system and can be safely neglected. It is worth pointing out that this is not always true; for example, in the critical region where the correlation plays an important role in the system, the correlation error should be taken into account.

B. Fast calculation of the error estimate

The naive calculation of the convolutions in the error estimate (7) requires a computational cost of $O(N^2)$. To estimate the force error on the fly, the computational cost of the estimate should be reduced to at least $O(N \log N)$. This is achieved by using the fast Fourier transform. We assume the simulation box is uniformly divided in to small bins. The positions of these bins are

$$\mathbf{r}_{i_1, i_2, i_3} = \frac{i_1}{M_1} \mathbf{a}_1 + \frac{i_2}{M_2} \mathbf{a}_2 + \frac{i_3}{M_3} \mathbf{a}_3, \quad 0 \leq i_\alpha < M_\alpha, \quad (8)$$

where \mathbf{a}_α are box vectors, and M_α is the number of divisions on each direction. In the reciprocal space, a corresponding lattice is set up:

$$\mathbf{k}_{m_1, m_2, m_3} = m_1 \mathbf{a}_1^* + m_2 \mathbf{a}_2^* + m_3 \mathbf{a}_3^*, \quad 0 \leq m_\alpha < M_\alpha, \quad (9)$$

where \mathbf{a}_α^* are reciprocal box vectors defined by $\mathbf{a}_\alpha \cdot \mathbf{a}_\beta^* = \delta_{\alpha\beta}$. In each bin the error and the particles number density are assumed to be constant; then a Fourier expansion can be obtained:

$$\begin{aligned} \mathcal{E}_{\text{homo}}^2(\mathbf{r}_{i_1, i_2, i_3}) &= \frac{1}{V} \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} \sum_{m_3=-\infty}^{\infty} [\mathcal{E}_{\text{homo}}^2]^\wedge(\mathbf{k}_{m_1, m_2, m_3}) \exp \left[2\pi i \left(\frac{m_1 i_1}{M_1} + \frac{m_2 i_2}{M_2} + \frac{m_3 i_3}{M_3} \right) \right] \\ &= \frac{1}{V} \sum_{m_1=0}^{M_1} \sum_{m_2=0}^{M_2} \sum_{m_3=0}^{M_3} \hat{K}_{\text{homo}}(\mathbf{k}'_{m_1, m_2, m_3}) \hat{\rho}(\mathbf{k}_{m_1, m_2, m_3}) \exp \left[2\pi i \left(\frac{m_1 i_1}{M_1} + \frac{m_2 i_2}{M_2} + \frac{m_3 i_3}{M_3} \right) \right]. \end{aligned} \quad (10)$$

Here $K_{\text{homo}}(\mathbf{r})$ denotes $|\mathbf{f}^c(\mathbf{r})|^2$. The wedge notation \wedge means the Fourier transform. The prime on $\mathbf{k}_{m_1, m_2, m_3}$ means that the periodic image $m_\alpha - M_\alpha$ should be used instead of m_α when $m_\alpha \geq M_\alpha/2$. The second equation in (10) holds due to the fact that $[\mathcal{E}^2]^\wedge = [K_{\text{homo}} * \rho]^\wedge = \hat{K}_{\text{homo}} \times \hat{\rho}$. The Fourier transform of ρ and K_{homo} is

$$\begin{aligned} \hat{\rho}(\mathbf{k}_{m_1, m_2, m_3}) &= \frac{V}{M_1 M_2 M_3} \sum_{i_1=0}^{M_1} \sum_{i_2=0}^{M_2} \sum_{i_3=0}^{M_3} \rho(\mathbf{r}_{i_1, i_2, i_3}) \\ &\times \exp \left[-2\pi i \left(\frac{m_1 i_1}{M_1} + \frac{m_2 i_2}{M_2} + \frac{m_3 i_3}{M_3} \right) \right] \end{aligned} \quad (11)$$

and

$$\hat{K}_{\text{homo}}(\mathbf{k}) = \int_{\mathbb{R}^3} K_{\text{homo}}(\mathbf{s}) e^{-2\pi i \mathbf{k} \cdot \mathbf{s}} d\mathbf{s}. \quad (12)$$

Most of the short-range interactions are isotropic, i.e., $K_{\text{homo}}(\mathbf{r}) = K_{\text{homo}}(r)$; then the three-dimensional integral in Eq. (12) can be simplified to a one-dimensional integral:

$$\hat{K}_{\text{homo}}(\mathbf{k}) = \int_{r_c}^{\infty} \frac{2s}{k} K_{\text{homo}}(s) \sin(2\pi k s) ds, \quad (13)$$

which can be calculated by numerical integration from r_c to a large upper bound, saying r_u . Perfect convergence is achieved by using $r_u = 30\sigma$ for the dispersion energy $-4\epsilon(\sigma/r)^6$, where ϵ and σ are the energy and length units. Since integral (13) is independent of the density of the system, it is calculated only once and stored.

Similarly, for Eq. (3), the Fourier mode of the the mean error force is (in terms of Fourier transform):

$$\langle \Delta \mathbf{F} \rangle^\wedge(\mathbf{k}) = \hat{\mathbf{f}}^c(\mathbf{k}) \hat{\rho}(\mathbf{k}), \quad (14)$$

where the Fourier transform of the complementary force is given by

$$\begin{aligned} \hat{\mathbf{f}}^c(\mathbf{k}) &= 2\pi \frac{\mathbf{k}}{k} i \left\{ r_c^2 u(r_c) \left[\frac{2 \cos(2\pi k r_c)}{2\pi k r_c} - \frac{2 \sin(2\pi k r_c)}{(2\pi k r_c)^2} \right] \right. \\ &\left. - \int_{r_c}^{\infty} 2s u(s) \sin(2\pi k s) ds \right\}, \end{aligned} \quad (15)$$

where $u(r)$ is the isotropic pairwise potential.

C. Adaptive cutoff radius method

A basic observation in an inhomogeneous system is that the RMS force error is not uniformly distributed. The difference between the maximum and minimum error can be more than one order of magnitude (Fig. 2). Traditionally, a very large uniform cutoff radius is used to reduce the maximum error in the system to reach convergent simulation results [4]. However, since the computational effort scales as the cube of the cutoff radius, a slightly larger cutoff will notably increase the computational effort. For example, increasing the cutoff radius from 4σ to 5σ will cost twice the computational effort. Therefore, a huge amount of computational effort is wasted in the low-error region. The idea of the adaptive cutoff radius simulation is to use larger cutoff radius for particles in the high-error region to control the force error, while to use a smaller cutoff radius for particles in the low-error region to save computational cost. The effectiveness of this method relies on a deliberate cutoff choice for each particle so that the force error is uniformly distributed across the simulation region. In order to determine the cutoff radius, we should estimate the RMS force error at least once for each particle, and the computational cost of each force error estimate is $O(N \log N)$, so the total computational expense of this method is at least $O(N^2 \log N)$, which is prohibitive to the on-the-fly application.

To make the adaptive cutoff method feasible, we assume the following:

(1) The space is uniformly divided in to small bins, the size of which is the same as those mentioned in Sec. II B. The particles falling in the same bin are assigned the same cutoff radius. In other words, the cutoff radius is chosen for each bin rather than for each particle.

(2) The cutoff radius can be chosen only from a discrete series of monotonically increasing values between a maximum r_c^{\max} and a minimum r_c^{\min} : $\{r_c^{\min} = r_c^0 < r_c^1 < \dots < r_c^{M-1} < r_c^M = r_c^{\max}\}$. We call it the *candidate set* of the cutoff radius. For convenience but not necessarily, the cutoff radii in the candidate set is uniformly increasing, i.e.,

$$r_c^m = r_c^0 + r_c^{\text{step}} m, \quad 0 \leq m \leq M. \quad (16)$$

For example, a possible candidate set is $\{2.5\sigma, 3.0\sigma, 3.5\sigma, 4.0\sigma, \dots, 10.0\sigma\}$, with $r_c^{\max} = 10.0\sigma$, $r_c^{\min} = 2.5\sigma$, and $r_c^{\text{step}} = 0.5\sigma$.

For each cutoff radius in the candidate set, the force error $\mathcal{E}(\mathbf{r})$ is calculated by the estimate (7). Notice the error is calculated by the fast algorithm described in Sec. II B, so it is defined on the bins. For a predetermined error control level \mathcal{E}_C , a satisfactory cutoff radius for a bin is chosen such that it is the smallest one satisfying the precision constraint $\mathcal{E}(\mathbf{r}_{i_1, i_2, i_3}) \leq \mathcal{E}_C$. The computational cost of this algorithm is $O(MN \log N)$. The idea of this algorithm is based on the observation that the cutoff radius may not be calculated at a very high precision. Because if r_c^{step} is sufficiently small, increasing the cutoff radius by no more than r_c^{step} will not waste too much computational cost. Here we recommend that r_c^{step} should be no more than 0.25σ , which increases the computational effort by 27% at $r_c = 3.0\sigma$. Due to the thermodynamic fluctuation of the particle density, the resulting cutoff distribution $r_c(\mathbf{r}_{i_1, i_2, i_3})$ (as a function of bin position)

needs to be refined by the following scheme:

$$r_c^{(1)}(\mathbf{r}_{i_1, i_2, i_3}) = \max_{j_a \in I} r_c(\mathbf{r}_{i_1+j_1, i_2+j_2, i_3+j_3}), \quad I = \{-1, 0, 1\}, \quad (17)$$

so that the fluctuating large-error regions are safely covered by bins with large cutoff radius.

Since the density profile of the system changes very slow comparing with the typical time step used in the MD simulation, the cutoff distribution is not calculated every time step. Instead, for example, in the liquid-vapor equilibrium simulation (see Sec. III A), the cutoff distribution is calculated every 20 000 time steps, so the extra load of adapting the cutoff is only marginal in the simulation, i.e., proportional to $\frac{M}{20000} N \log N$.

D. Long-range force correction

In an inhomogeneous system, the inhomogeneity error may play a dominate role in the simulation (see examples in Sec. III). Since the mean error force can be calculated by Eq. (14) on the fly, a simple way to screen out the inhomogeneity error is to correct the computed force by the mean error force, namely,

$$\mathbf{F}_{\text{corr}}(\mathbf{r}) = \tilde{\mathbf{F}}(\mathbf{r}) + \langle \Delta \mathbf{F}(\mathbf{r}) \rangle. \quad (18)$$

In this paper the mean error force $\langle \Delta \mathbf{F}(\mathbf{r}) \rangle$ is also referred as *correction force* in the context of long-range force correction. This correction works because it takes into account the force contribution outside the cutoff radius. The mean of the corrected error force vanishes:

$$\langle \Delta \mathbf{F}_{\text{corr}}(\mathbf{r}) \rangle = \langle \mathbf{F}(\mathbf{r}) - \tilde{\mathbf{F}}(\mathbf{r}) - \langle \Delta \mathbf{F}(\mathbf{r}) \rangle \rangle = 0. \quad (19)$$

The mean square error is given by

$$\begin{aligned} \langle |\Delta \mathbf{F}_{\text{corr}}(\mathbf{r})|^2 \rangle &= \langle [\mathbf{F}(\mathbf{r}) - \tilde{\mathbf{F}}(\mathbf{r}) - \langle \Delta \mathbf{F}(\mathbf{r}) \rangle]^2 \rangle \\ &= \langle [\Delta \mathbf{F}(\mathbf{r}) - \langle \Delta \mathbf{F}(\mathbf{r}) \rangle]^2 \rangle \\ &= \langle |\Delta \mathbf{F}(\mathbf{r})|^2 \rangle - \langle \Delta \mathbf{F}(\mathbf{r}) \rangle^2 \\ &= \mathcal{E}_{\text{homo}}^2(\mathbf{r}) + \mathcal{E}_{\text{correlation}}(\mathbf{r}), \end{aligned} \quad (20)$$

containing no inhomogeneity contribution. It is worth noticing that the mean square error of the corrected force is nothing but the variance of the error force: $\langle |\Delta \mathbf{F}_{\text{corr}}(\mathbf{r})|^2 \rangle = \text{Var}[\Delta \mathbf{F}_{\text{corr}}(\mathbf{r})] = \text{Var}[\Delta \mathbf{F}(\mathbf{r})]$.

In most cases, the ensemble average of the error force does not change very fast comparing with the typical time step used in a MD simulation. Therefore, it is not necessary to calculate the correction force every time step. Instead, for example, we calculate it every 40 time steps in the liquid-vapor equilibrium testing system (see Sec. III A), which takes no more than 10% of the total computational cost. If the long-range interaction calculations (which needs FFTs anyway) are involved, this ratio should be even lower. The idea of the long-range force correction is very similar to those proposed by earlier works [6–8, 10]. In those works, the density profile is assumed to be inhomogeneous along one direction, and uniform on the other two directions. The system is subdivided into slabs perpendicular to the inhomogeneous direction to calculate the energy or force correction. The advantage of our force correction is that it has *no* uniformity restriction on the

density profile of the system, so it can be used in a wider range of inhomogeneous simulations.

E. Long-range pressure correction

The pressure contribution outside the cutoff radius can be expressed by

$$\mathbf{P}^c = \frac{1}{V} \left\langle \sum_{i \neq j} \frac{1}{2} (\mathbf{r}_i - \mathbf{r}_j) \otimes \mathbf{f}^c(\mathbf{r}_i - \mathbf{r}_j) \right\rangle. \quad (21)$$

For simplicity, we denote $\frac{1}{2V} \mathbf{r} \otimes \mathbf{f}^c$ by \mathbf{k}_P ; therefore,

$$\begin{aligned} \mathbf{P}^c &= \left\langle \sum_{i \neq j} \mathbf{k}_P(\mathbf{r}_i - \mathbf{r}_j) \right\rangle = \int_{\mathbb{R}^3 \times \mathbb{R}^3} \mathbf{k}_P(\mathbf{r}' - \mathbf{r}'') \rho(\mathbf{r}', \mathbf{r}'') d\mathbf{r}' d\mathbf{r}'' \\ &= \int_{\mathbb{R}^3 \times \mathbb{R}^3} \mathbf{k}_P(\mathbf{r}' - \mathbf{r}'') [\rho(\mathbf{r}') \rho(\mathbf{r}'') + C(\mathbf{r}', \mathbf{r}'')] d\mathbf{r}' d\mathbf{r}''. \end{aligned}$$

Assuming the correlation term $C(\mathbf{r}', \mathbf{r}'')$ vanishes, then

$$\begin{aligned} \mathbf{P}^c &= \int_{\mathbb{R}^3 \times \mathbb{R}^3} \mathbf{k}_P(\mathbf{r}' - \mathbf{r}'') \rho(\mathbf{r}') \rho(\mathbf{r}'') d\mathbf{r}' d\mathbf{r}'' \\ &= \int_{\mathbb{R}^3} \left[\int_{\mathbb{R}^3} \mathbf{k}_P(\mathbf{r}' - \mathbf{r}'') \rho(\mathbf{r}'') d\mathbf{r}'' \right] \rho(\mathbf{r}') d\mathbf{r}'. \quad (22) \end{aligned}$$

The Fourier transform of the convolution $\int \mathbf{k}_P(\mathbf{r}' - \mathbf{r}'') \rho(\mathbf{r}'') d\mathbf{r}''$ is identical to the product $\hat{\mathbf{k}}_P(\mathbf{k}) \hat{\rho}(\mathbf{k})$. Assuming the isotropicity of the short-range interaction, the $\alpha\beta$ component of \mathbf{k}_P can be written as

$$\{\mathbf{k}_P\}_{\alpha\beta} = \frac{1}{2V} \{\mathbf{r} \otimes \mathbf{f}^c\}_{\alpha\beta} = \frac{1}{2V} r_\alpha r_\beta G^c(r), \quad \alpha, \beta = 1, 2, 3, \quad (23)$$

where $G^c(r) = [u^c(r)]/r$ is isotropic. We list here without proofs the Fourier transform of these terms:

$$\begin{aligned} [r_1 r_1 G^c(r)]^\wedge &= \int_{r_c}^\infty dr r^4 G^c(r) \pi \sqrt{\frac{1}{kr}} \left[\frac{2}{3} J_{\frac{1}{2}}(2\pi kr) \right. \\ &\quad \left. - \left(\frac{1}{3} - \cos^2 \Phi + \sin^2 \Phi \cos 2\Theta \right) J_{\frac{3}{2}}(2\pi kr) \right], \quad (24) \end{aligned}$$

$$\begin{aligned} [r_2 r_2 G^c(r)]^\wedge &= \int_{r_c}^\infty dr r^4 G^c(r) \pi \sqrt{\frac{1}{kr}} \left[\frac{2}{3} J_{\frac{1}{2}}(2\pi kr) \right. \\ &\quad \left. - \left(\frac{1}{3} - \cos^2 \Phi - \sin^2 \Phi \cos 2\Theta \right) J_{\frac{3}{2}}(2\pi kr) \right], \quad (25) \end{aligned}$$

$$\begin{aligned} [r_3 r_3 G^c(r)]^\wedge &= \int_{r_c}^\infty dr r^4 G^c(r) \pi \sqrt{\frac{1}{kr}} \left[\frac{2}{3} J_{\frac{1}{2}}(2\pi kr) \right. \\ &\quad \left. + \left(\frac{2}{3} - 2 \cos^2 \Phi \right) J_{\frac{3}{2}}(2\pi kr) \right], \quad (26) \end{aligned}$$

$$\begin{aligned} [r_1 r_2 G^c(r)]^\wedge &= - \int_{r_c}^\infty dr r^4 G^c(r) \pi \\ &\quad \times \sqrt{\frac{1}{kr}} \sin^2 \Phi \sin 2\Theta J_{\frac{3}{2}}(2\pi kr), \quad (27) \end{aligned}$$

$$\begin{aligned} [r_1 r_3 G^c(r)]^\wedge &= - \int_{r_c}^\infty dr r^4 G^c(r) \pi \\ &\quad \times \sqrt{\frac{1}{kr}} \sin 2\Phi \cos \Theta J_{\frac{3}{2}}(2\pi kr), \quad (28) \end{aligned}$$

$$\begin{aligned} [r_2 r_3 G^c(r)]^\wedge &= - \int_{r_c}^\infty dr r^4 G^c(r) \pi \\ &\quad \times \sqrt{\frac{1}{kr}} \sin 2\Phi \sin \Theta J_{\frac{3}{2}}(2\pi kr), \quad (29) \end{aligned}$$

where (k, Φ, Θ) are the spherical coordinates of reciprocal variable \mathbf{k} , and $J_\nu(x)$ is the Bessel function of the first kind.

III. TESTING SIMULATIONS AND DISCUSSIONS

A. Liquid-vapor equilibrium

The direct simulation of the liquid-vapor equilibrium is a very typical inhomogeneous system, which is widely used to study the phase coexistence properties, for example, the equilibrium liquid-vapor densities, and the surface tension. The system studied in the present paper contains 16 000 identical standard Lennard-Jones particles that are separated into a liquid phase in the center of the periodic box and vapor phase surrounding it (Fig. 1). The conventional Lennard-Jones unit system is employed: The unit of length, energy, mass, and time are denoted by ϵ , σ , m , and τ ($\tau = \sigma \sqrt{m/\epsilon}$), respectively. In the rest of this paper, all quantities are rescaled by the Lennard-Jones units and added the superscript “*.” For example, $r^* = r/\sigma$, $T^* = k_B T/\epsilon$, $P^* = P \cdot \sigma^3/\epsilon$, and $\mathcal{E}^* = \mathcal{E} \cdot \sigma/\epsilon$. The MD time step is $\Delta t^* = 0.005$. The simulations last for 3×10^6 time steps. The first 1×10^6 steps are discarded. The quantities of interest are sampled every 100 time steps. The blocking average method [26] is applied to estimate the statistical uncertainty of the autocorrelated data with 95% confidence level. The NVT ensemble is generated by the Nosè-Hoover thermostat [27,28]. The size of the simulation box is $L_x^* \times L_y^* \times L_z^* = 150 \times 21 \times 21$, which permits a maximum cutoff radius of $r_c^* = 10$ and eliminates the finite size effect reported in the literature [29–31]. The system is divided into bins of size $1 \times 1 \times 1$ to perform the fast error estimate method described in Sec. II B. In all simulations, the maximum cutoff radius considered is 10, but the reference results for comparison are obtained by simulations carried out in a 24 000 particle system with simulation box $L_x^* \times L_y^* \times L_z^* = 150 \times 27 \times 27$, and a uniform cutoff radius of $r_c^* = 13$. We find that (not shown here) the bigger system leads to the same equilibrium quantities as the smaller system at the same cutoff radius, so it is reasonable to use it to produce the better converged data at $r_c^* = 13$ and compare them with those obtained in the smaller system.

Figure 2 presents the real RMS force error (in green square and log-scaled) in a system using a uniform cutoff radius of $r_c^* = 7.5$. The number density distribution along

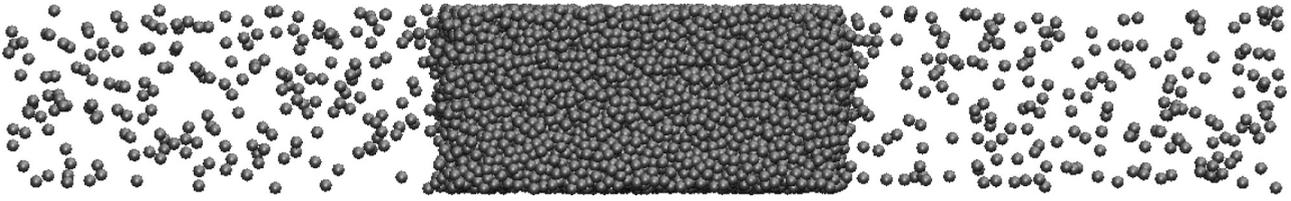


FIG. 1. The snapshot of the 16 000 Lennard-Jones particle system in liquid-vapor equilibrium at $T^* = 0.85$ with a uniform cutoff radius of $r_c^* = 7.5$.

x direction is shown by the solid red line for reference. The error is comparatively constant at the bulk liquid and gas region. In contrast, two peaks, which are more than one order of magnitude larger than the force errors in bulk regions, form at the interfacial region. The estimated RMS force error is presented by solid blue line. The estimate follows the real error very well at the interfacial region, but slightly overestimate the error in the bulk regions. Two components of the force error, namely, the homogeneity error and inhomogeneity error, are presented in Fig. 2 by dashed and dotted blue lines, respectively. In the bulk regions, the dashed blue line overlaps with the solid blue line, which means the homogeneity error dominates. In the interfacial region, instead, the inhomogeneity error dominates, because the dotted blue line overlaps the solid one.

As mentioned before, the key idea of the adaptive cutoff simulation is to use a smaller cutoff radius in the bulk regions, so that the RMS force error is uniformly distributed over the simulation region. In Fig. 3, we present the adapted cutoff radius by a solid red line and all errors by the same notations as in Fig. 2. The control error $\mathcal{E}_C^* = 0.0045$ is almost the same as the maximum error in the uniform cutoff simulation with

$r_c^* = 7.5$. In the rest of the paper, when we mention that “the cutoff of the adaptive cutoff method is r_c^* ,” we mean the control error is the maximum error of a uniform cutoff simulation using that cutoff r_c^* . In all the cases considered by the present paper, the maximum of the adapted cutoff radius is almost the same as r_c^* and is sometimes marginally larger (no more than r_c^{step}). The cutoff distribution is calculated every 20 000 time steps and is refined twice by Eq. (17). Due to this low cutoff redistribution frequency, the extra cost of this adaptive method is only marginal compared with the total computational cost. In the interfacial regions, the same cutoff radius as the uniform cutoff simulation, namely, 7.5, is used. In the bulk liquid and vapor regions, the cutoff radii are 4.75 and 3.5, which reduce the computational cost by a factor of 4 and 10, respectively. The error is uniformly distributed, except in the liquid region where the real error is somewhat lower. In contrast with the uniform cutoff case, i.e., Fig. 2, the error estimate for the vapor region is very sharp. The similar phenomenon is observed in Ref. [15]. The same as with the uniform cutoff simulation, the homogeneity error dominates in the bulk regions, while the inhomogeneity error dominates in the interfacial region.

In Fig. 4 the red solid line presents the x component of the correction force that is calculated every 40 time steps. The extra computational cost is no more than 10% of the total computational costs in all the tested cases. The y and z components are not shown because their magnitudes are

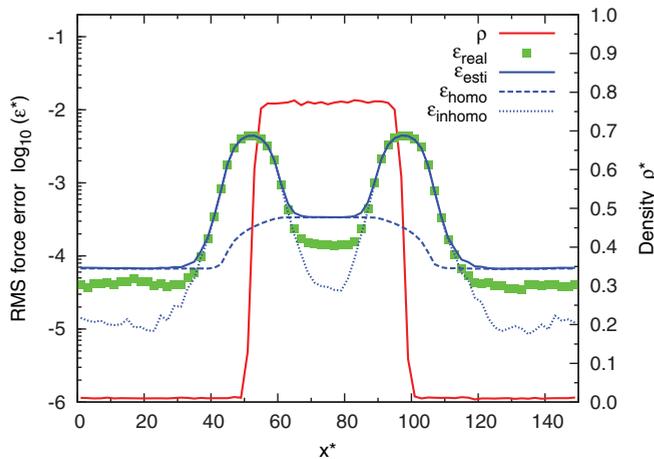


FIG. 2. (Color online) The RMS force error distribution of the 16 000 Lennard-Jones particle system at $T^* = 0.85$ with a uniform cutoff radius of $r_c^* = 7.5$. The green (light gray) squares present the real error. The solid blue (dark gray) line is the error estimate by (7). The dashed blue (dark gray) line is the homogeneity error contribution, while the dotted blue (dark gray) line is the inhomogeneity error contribution. The density distribution of the system is presented by red (middle gray) solid line for reference. All properties are properly averaged on the y and z direction; therefore, the profiles are plotted along the x direction.

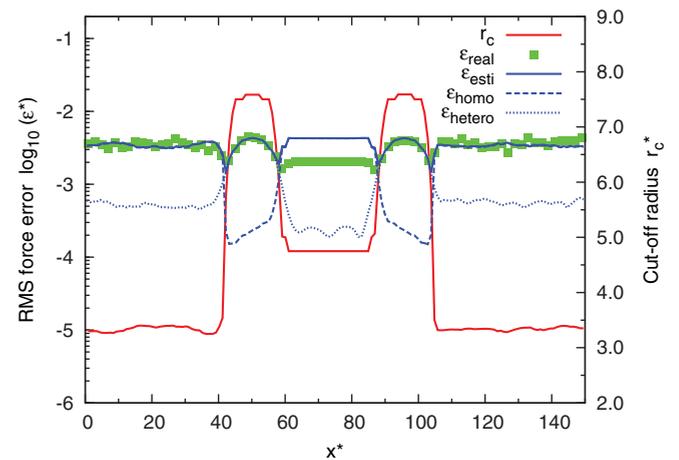


FIG. 3. (Color online) The RMS force error distribution of the 16 000 Lennard-Jones particle system at $T^* = 0.85$ with adaptive cutoff radius. The control error \mathcal{E}_C^* is the same as the maximum error in the uniform cutoff simulation with $r_c^* = 7.5$. The resulting cutoff radius distribution is presented by the red (middle gray) line. All other notations are the same as Fig. 2.

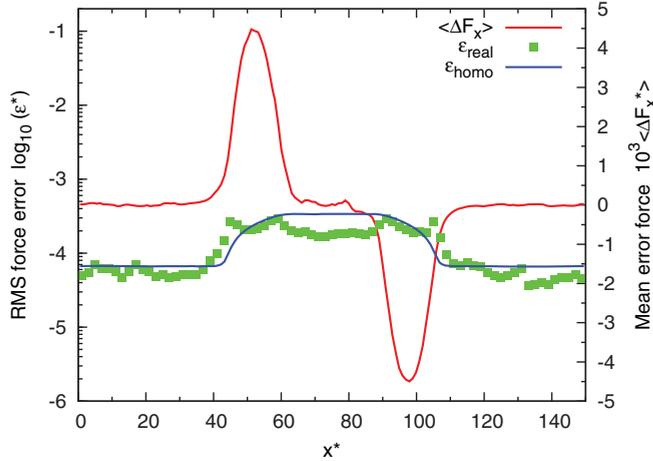


FIG. 4. (Color online) The RMS force error distribution of the 16 000 Lennard-Jones particle system at $T^* = 0.85$ with long-range force correction, i.e., Eq. (18). The cutoff radius is $r_c^* = 7.5$. The x component of the correction force is presented by the solid red (middle gray) line. The green (light gray) squares present the real error. The solid blue (dark gray) line presents the estimated homogeneity error.

negligible. On the left interface, the correction force is along the positive direction of x axis, namely, pointing right. On the right interface, the correction force is of the same magnitude but pointing left. Therefore, the liquid density should be higher, and the vapor density should be lower than the corresponding uniform cutoff simulation. The real error (in green square) and the estimated homogeneity error (in blue solid line) are all presented in Fig. 4. Clearly the real error follows the homogeneity error across the simulation region, which implies the contribution of the inhomogeneity error is successfully removed, or is at least negligible. In the bulk liquid region, the homogeneity error is slightly overestimated, which is the same as the uniform and adaptive cutoff simulations.

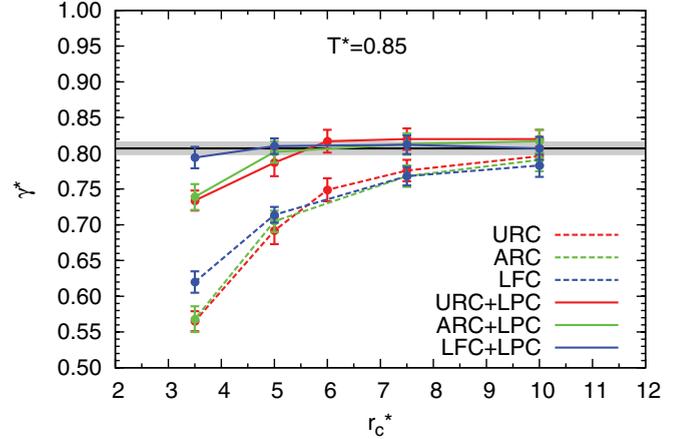
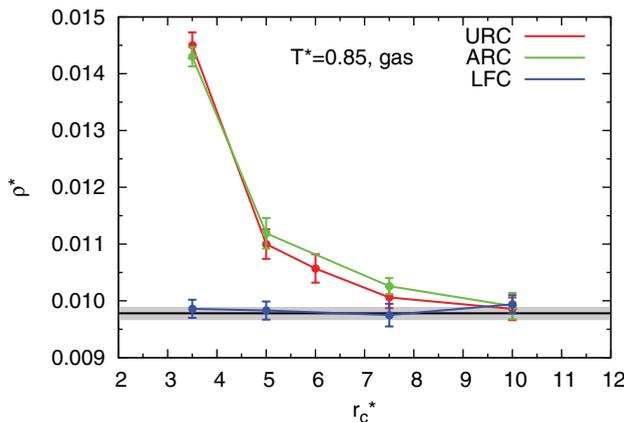


FIG. 6. (Color online) $T^* = 0.85$. The convergence of the surface tension with respect to the cutoff radius. This figure presents the results of three methods, namely, the uniform cutoff (URC) in red (middle gray), the adaptive cutoff (ARC) in green (light gray), and the long-range force correction (LFC) method in blue (dark gray). The results with and without long-range pressure correction (LPC) are shown together for comparison. The reference surface tension, shown by the solid black line, is measured by the uniform cutoff method with $r_c^* = 13$ using LPC. “The cutoff of the adaptive cutoff method is r_c^{*} ” means that the control error \mathcal{E}_c^* is the same as the maximum error of a uniform cutoff simulation using r_c^* . Under this setting, the maximum cutoff radius of the adaptive cutoff method is the same as, or sometimes marginally larger than, r_c^* (no more than r_c^{step}).

To test the effectiveness of the proposed adaptive cutoff and long-range force correction methods, we measure the liquid-vapor equilibrium density and the surface tension defined by

$$\gamma = \frac{1}{2} \int_{L_x} \left[p_x(x) - \frac{p_y(x) + p_z(x)}{2} \right] dx, \quad (30)$$

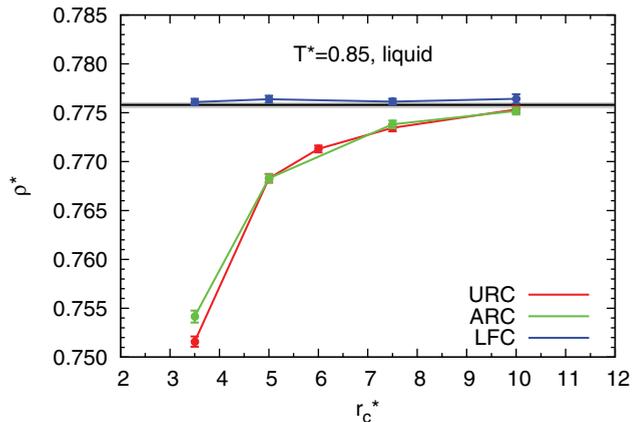


FIG. 5. (Color online) $T^* = 0.85$. The convergence of the equilibrium gas and liquid densities with respect to the cutoff radius. All densities are compared with the reference value (black solid line), namely, the liquid and vapor densities measured with a uniform cutoff radius $r_c^* = 13$. The gray region denotes the statistical uncertainty of the reference value. This figure presents the results of three methods, namely, the uniform cutoff (URC) in red (middle gray), the adaptive cutoff (ARC) in green (light gray), and the force correction (LFC) method in blue (dark gray). “The cutoff of the adaptive cutoff method is r_c^{*} ” means that the control error \mathcal{E}_c^* is the same as the maximum error of a uniform cutoff simulation using r_c^* . Under this setting, the maximum cutoff radius of the adaptive cutoff method is the same as, or sometimes marginally larger than r_c^* (no more than r_c^{step}).

TABLE I. $T^* = 0.85$. The maximum RMS force error and computational cost of three methods, namely, the uniform cutoff (URC), the adaptive cutoff (ARC), and the long-range force correction (LFC) method. The computational cost is the number of pairwise interaction measurements, which should be machine independent.

r_c^*	Method	$\max \mathcal{E}^*$	Cost ($\times 10^{-6}$)
3.5	URC	8.6×10^{-2}	2.9
3.5	ARC	7.7×10^{-2}	1.9
3.5	LFC	1.4×10^{-2}	3.1
5.0	URC	2.2×10^{-2}	7.8
5.0	ARC	2.0×10^{-2}	4.2
5.0	LFC	1.9×10^{-3}	7.9
6.0	URC	1.1×10^{-2}	13
7.5	URC	4.5×10^{-3}	24
7.5	ARC	4.3×10^{-3}	12
7.5	LFC	2.3×10^{-4}	24
10.0	URC	1.4×10^{-3}	53
10.0	ARC	1.4×10^{-3}	26
10.0	LFC	5.3×10^{-5}	53

where p_x , p_y , and p_z are x , y , and z component of the pressure. The convergence of the properties of interest with respect to the cutoff radius is considered.

In Figs. 5 and 6, the convergence of the equilibrium liquid-vapor densities and surface tension are investigated with respect to the increasing cutoff radius. The red, green, and blue points with error bars denote the results of uniform, adaptive cutoff methods, and the long-range force correction, respectively. In Fig. 5 all densities are compared with the reference value measured by the uniform cutoff simulation of $r_c^* = 13$, with the statistical uncertainty denoted by the gray region. The densities are connected by the solid lines to guide the eyes. In Fig. 6 the direct measures of surface tension (data points connected by the dashed lines) are presented with pressure corrected surface tension (data points connected by solid lines). The same as for the reference densities, the reference surface tension is measured by uniform cutoff simulations at $r_c^* = 13$ with long-range pressure correction. Table I presents the maximum RMS force error and the computational effort of the mentioned methods at different cutoff radii. The computational cost is given by the average numbers of pairwise interactions per time step calculated by the program, which is platform independent and serves as a good benchmark.

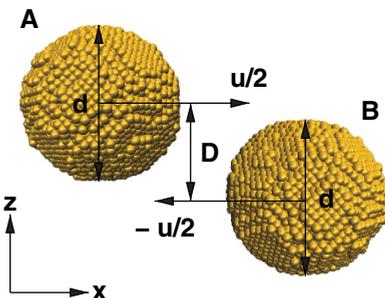


FIG. 7. (Color online) The initial setup of cluster A and B. The initial velocities are equal in magnitude but on opposite direction.

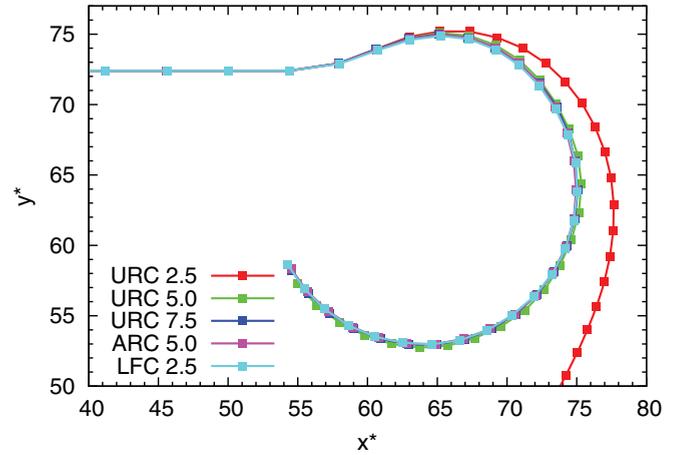


FIG. 8. (Color online) The center of mass trajectory of cluster A. Uniform cutoff (URC) $r_c^* = 2.5, 5.0, 7.5$, adaptive cutoff (ARC) $r_c^* = 5.0$, and long-range force correction (LFC) $r_c^* = 2.5$ are investigated. The time interval between two neighboring squares on the trajectory is 4.

The liquid and gas densities of the uniform and the adaptive cutoff methods are almost identical at the same cutoff radius; see Fig. 5. This observation is also true for the surface tension; see Fig. 6. This indicates that controlling the maximum of force error (see Table I) and introducing larger homogeneous force error in the bulk liquid or gas regions will not disturb the accuracy of the equilibrium densities and surface tension. The adaptive cutoff method roughly saves 1/3–1/2 of the computational cost (see Table I). The efficiency benefit of the adaptive cutoff depends on the property of the system: If the bulk regions dominate the system, and the inhomogeneity produces large interfacial force error, then the adaptive cutoff will save non-negligible computational resources. On the other hand, if the interfacial region is computationally intensive, the adaptive cutoff will not improve the efficiency greatly.

The long-range force correction method presents a much better density convergence than the uniform and adaptive

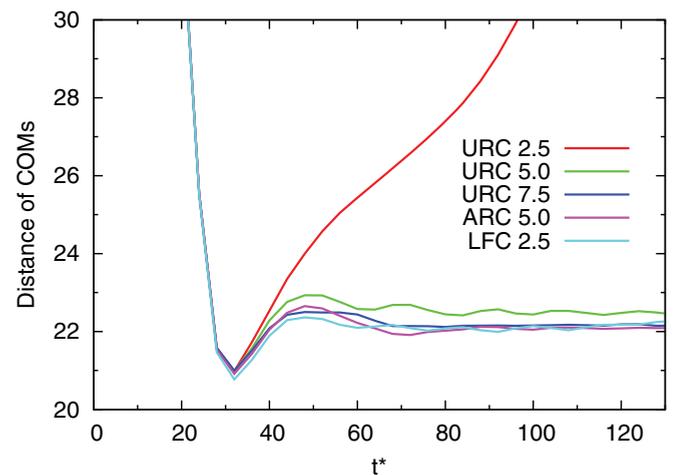


FIG. 9. (Color online) The COM distance between cluster A and cluster B versus time. Uniform cutoff (URC) $r_c^* = 2.5, 5.0, 7.5$, adaptive cutoff (ARC) $r_c^* = 5.0$, and long-range force correction (LFC) $r_c^* = 2.5$ are investigated.

cutoff simulations: The equilibrium densities converge when the cutoff radius is only 3.5 (Fig. 5). The maximum RMS force error is reduced by roughly one order of magnitude, comparing to the uniform and adaptive cutoff methods (see Table I). The computational cost of the force correction method is the same as the uniform cutoff methods at the same cutoff radius.

The surface tension result of the force correction without pressure correction is only slightly better than the uniform and adaptive cutoff methods without pressure correction (Fig. 6). With the long-range pressure correction, the convergence of the surface tension of all methods are largely improved. The force correction method converges at $r_c^* = 3.5$, and the uniform (adaptive) cutoff methods converge at $r_c^* = 5.0 \sim 6.0$. The discrepancy between the corrected and uncorrected tension is the amount of the correction, which decreases with the increasing cutoff radius. However, even at a rather large cutoff of $r_c^* = 10$, the amount of the correction is still not

negligible, which tells the users that the uncorrected tension calculation is far from convergence.

Comparing the results of the force correction method at $r_c^* = 3.5$ and uniform cutoff radius method at $r_c^* = 6.0$, the force accuracy of the former method is a little worse than the latter (see Table I), but the equilibrium density accuracy is better and surface tension is equally good (Figs. 5 and 6). Since the homogeneity error dominates in the former simulation and the inhomogeneity error dominates in the latter, the above observation indicates that different natures of error play different roles in the equilibrium densities. The homogeneity error is actually the variation of the corrected error force with its mean equal to zero [see Eq. (19)]. This implies that the corrected error force may point to any direction with the same probability and may self-cancel in a long time average. The inhomogeneity error, in contrast, stems from the nonvanishing mean value of the error force, which is always pointing to one certain direction (the opposite direction of

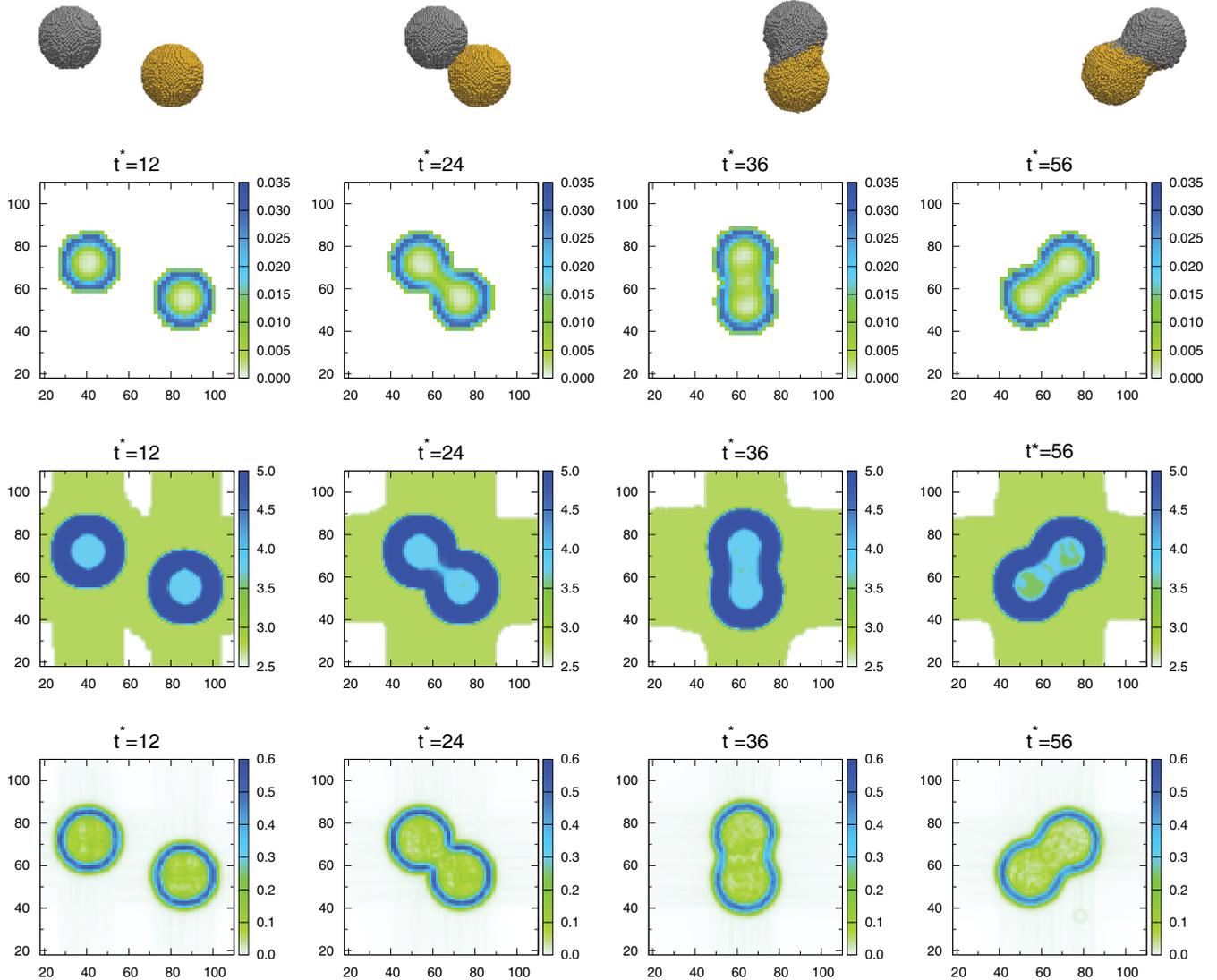


FIG. 10. (Color online) The colliding clusters (the first row), the RMS force error distribution of uniform cutoff with $r_c^* = 5.0$ (the second row), the cutoff distribution of adaptive cutoff with $r_c^* = 5.0$ (the third row), and the magnitude of long-range force correction with $r_c^* = 2.5$ (the fourth row). The four columns corresponds to snapshots taken at $t^* = 12, 24, 36, 56$, respectively.

the density gradient) and has a low opportunity to be self-canceled.

The precise calculation of the correction force needs the precise knowledge of the density profile of the system [see Eq. (3)]. Then the corrected force, in turn, improves the accuracy of the density profile. It is possible that this feedback loop does not converge, namely, the system is driven to a totally wrong state by the corresponding incorrect correction force. It may happen when the simulation starts very far from the equilibrium, or when the fluctuation in the system is so large that an accurate calculation of the density profile is impossible. In our simulations, we do not observe any failure of convergence. First, the system is far from the critical point, and we always let the simulation starts from a good enough configuration, for example, a snapshot of a simulation with the uniform small cutoff and lower temperature, or a snapshot from another force correction simulation. We also observe that the equilibration of the force-corrected system is always as fast as the system simulated by uniform and adaptive cutoff methods. That means, at least in our simulation cases, the force correction does not lead to more effort of equilibrating the system.

B. Collision of nanoscale Lennard-Jones clusters

In this section we test the adaptive cutoff method and the long-range force correction method in a dynamical problem: collision of nanoscale Lennard-Jones clusters. Two initial Lennard-Jones clusters, denoted by A and B , are set up in the simulation box, each of which contains the same number of particles $N_A = N_B = 10792$, with a diameter of $d^* \approx 27$. The initial velocity of the clusters are $\mathbf{u}_A = \frac{1}{2}(u, 0, 0)$ and $\mathbf{u}_B = -\frac{1}{2}(u, 0, 0)$; see Fig. 7. The impact parameter D is defined by the z coordinate difference between the center of mass (COM) of the clusters. The dimensionless impact factor defined by

$$x = \frac{D}{d} \quad (31)$$

is used to measure how far off-center the collision happens. If $x = 0$ the collision is head-on while $x = 1$ the collision is avoided. Reference [32] investigated a broad range combination of the velocity u and impact factor x , and identified three major collision modes: the coalescence, the stretching separation, and the shattering. The major modes were further classified into a series submodes. In this paper we want to study how the collision properties are affected by the force precision; therefore, we focus on one special case: $u^* = 2.2$ and $x^* = 0.6$, where the major collision mode is the coalescence, and a poor precision of force calculation may lead to a wrong major collision mode.

In Figs. 8 and 9 we present the trajectory of COM of cluster A and the COM distance between the two clusters, respectively. Three different cutoff radii of uniform cutoff simulations are considered. It is obvious that the trajectory of $r_c^* = 2.5$ is completely wrong: two clusters separate after collision. The $r_c^* = 5.0$ trajectory is reasonably good, but the COM distance converges only when $r_c^* = 7.5$. This means the precision of force calculation is crucial in the study of collision dynamics. Although the precision of the adaptive

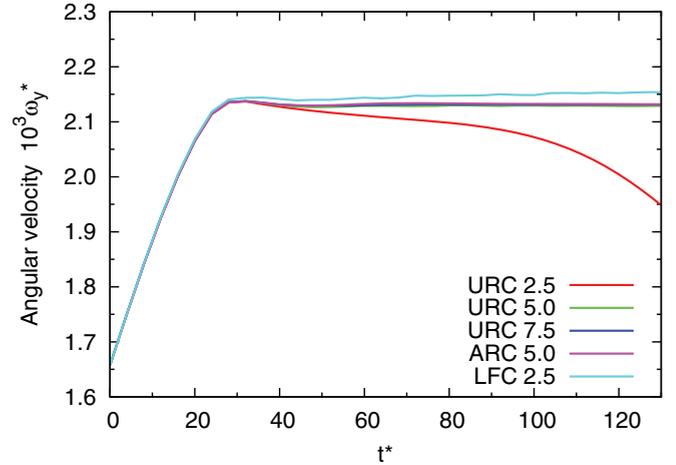


FIG. 11. (Color online) The angular velocity of cluster A versus time. Uniform cutoff (URC) $r_c^* = 2.5, 5.0, 7.5$, adaptive cutoff (ARC) $r_c^* = 5.0$, and long-range force correction (LFC) $r_c^* = 2.5$ are investigated.

cutoff method with $r_c^* = 5$ is the same as uniform cutoff $r_c^* = 5$, the former produces a better trajectory. The long-range force correction method that gets rid of the inhomogeneity error also produces a correct trajectory, but the motion along the trajectory is somewhat faster than other methods (see also the angular velocity in Fig. 11).

The RMS force error of the uniform cutoff, the cutoff distribution of the adaptive cutoff and the magnitude of the correction force of long-range force correction method are plotted in Fig. 10. The configuration of the two clusters (colored gray and yellow) are presented on top of the figure for reference. The same as in the liquid-vapor equilibrium simulation, the force error dominates in the interfacial regions, namely, the boundaries of the clusters. Inside and outside the clusters, the error is much lower. The cutoff distribution of the adaptive cutoff method is updated every 200 time steps. The third row of Fig. 10 shows that the large cutoff region follows the cluster boundaries perfectly. It is much broader than the large error region, because it is refined by Eq. (17) to keep track of the possible moving of the clusters. The correction force is calculated every 10 time steps. From the fourth row of Fig. 10, it also follows the moving boundaries of the colliding clusters. In the bulk region of the clusters, the correction force is negligible. It is not efficient to calculate the correction force more frequently, because every force correction involves FFTs that are comparatively expensive, especially when the cutoff radius is small. In the present case, FFTs take up 50% of the total computational cost.

The angular velocity is presented in Fig. 11. The force correction result is not as precise as other simulations excluding a uniform $r_c = 2.5$ simulation. This probably because the value of the correction force is updated every 10 time steps, so considering the moving of the clusters, it is not the exact up-to-date value. Figure 12 presents the angular momentum, which is preserved by all uniform cutoff simulations. The angular momentum of the adaptive cutoff simulation, which does not precisely preserve Newton's third law, shows deviation, but it is still around the correct value. However, the angular

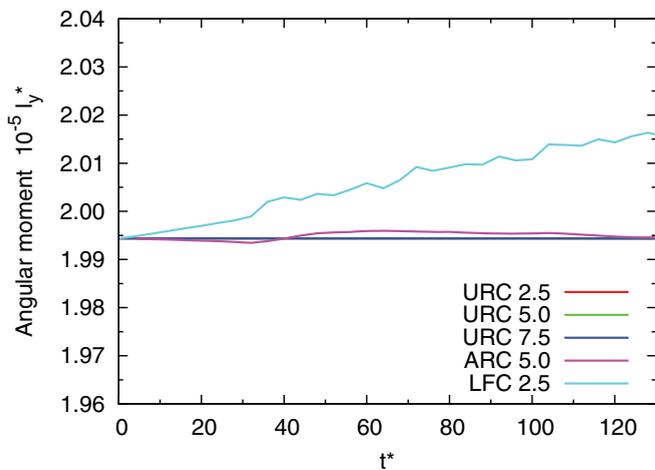


FIG. 12. (Color online) The angular momentum of cluster *A* versus time. Uniform cutoff (URC) $r_c^* = 2.5, 5.0, 7.5$, adaptive cutoff (ARC) $r_c = 5.0$, and long-range force correction (LFC) $r_c^* = 2.5$ are investigated.

moment of the force correction simulation deviates from the correct value by 1% at $t^* = 130$.

The computational cost is given in Fig. 13, counted by the number of calculated pairwise interactions. The relative cost of the uniform cutoff method with $r_c^* = 2.5, 5.0, 7.5$ is 1:5.6:15.8. The adaptive cutoff with $r_c^* = 5.0$ saves about 18% computational cost compared with the uniform cutoff method with $r_c^* = 5.0$. The adaptive cutoff method does not save a lot computational cost, because the radius of the spherical cluster is about 13.5, the interfacial region (with a width of 5.0) is roughly 3/4 of the total volume. If larger clusters are studied, the performance of the adaptive cutoff method will be better. The long-range force correction with $r_c^* = 2.5$ calculates the same number of interactions as the uniform cutoff $r_c^* = 2.5$. However, as discussed before, 1/2 of the total computational effort is spent on calculating the correction force. This extra

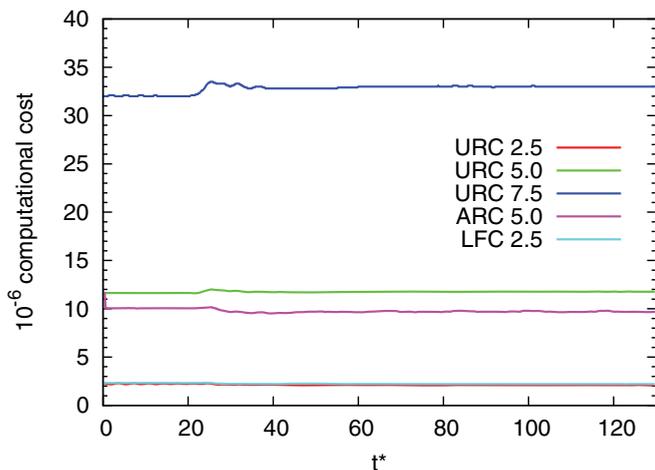


FIG. 13. (Color online) The computational cost versus time. Uniform cutoff (URC) $r_c^* = 2.5, 5.0, 7.5$, adaptive cutoff (ARC) $r_c^* = 5.0$, and long-range force correction (LFC) $r_c^* = 2.5$ are investigated. The computational cost is counted by the number of calculated pairwise interactions.

expense is worthwhile because it improves the simulation results significantly.

IV. CONCLUSIONS AND DISCUSSION

In this paper we developed the force error estimate for nonbonded short-range interactions in the inhomogeneous molecular systems. The RMS force error can be expressed by a summation of a homogeneity error, an inhomogeneity error, and a correlation error. In the present paper, we considered only the cases in which the correlation error does not dominate. In an inhomogeneous system, the basic observation of the error distribution was that the homogeneity error dominates the bulk material regions, whereas the inhomogeneity error dominates the interfacial regions where the density changes within the length scale of a few molecules. The inhomogeneity error can be more than one order of magnitude larger than the homogeneity error. Two methods were proposed to improve the efficiency and accuracy of the simulation. The adaptive cutoff method uses a smaller cutoff radius for the low-error bulk regions and larger cutoff radius for the high-error interfacial regions, so that the RMS force error is uniformly distributed over the simulation region. The long-range force correction method corrects the force computation by the mean error force, so that the inhomogeneity error is removed from the system. It should be noted that the proposed error estimate and correction methods can be used only for periodic systems.

We studied the liquid-vapor equilibrium and the collision of nanoscale Lennard-Jones cluster to demonstrate the validity of the proposed methods. All simulation results were compared with the uniform cutoff simulations, in which very large cutoff radii were used to achieve real convergence. With the same maximum RMS force error, the precision of physical properties of the adaptive cutoff method was the same as the uniform cutoff method. From the efficiency point of view, the adaptive cutoff simulation saved 34%–50% computational cost in liquid-vapor simulation and 18% in the cluster collision simulation. The advantage of the adaptive cutoff method depends on the system: If the computational costs of the bulk regions rather than the interfacial region dominate the system, the adaptive cutoff method will save a notably large amount of the computational resources.

In the liquid-vapor equilibrium simulation, the force correction method improved the force accuracy by one order of magnitude without large extra computational expense. The physical properties converged at cutoff $r_c^* = 3.5$, whereas the uniform cutoff method converged at much larger cutoff radii ($r_c^* = 6.0$ for surface tension and $r_c^* = 10.0$ for equilibrium densities). We also concluded that the homogeneity error and the inhomogeneity error play different roles in the precision of equilibrium liquid-vapor densities: The inhomogeneity error is more harmful even though its magnitude is smaller than the homogeneity error. In the collision cluster case, the force correction methods produced high-precision trajectory at a very small cutoff radius $r_c^* = 2.5$, with which the uniform cutoff method was qualitatively wrong. However, due to the inconsistency of the cluster movement and the correction force computation, small deviations of angular velocity and momentum were observed.

In the systems tested by the present paper, the computational load of the error estimate is small compared to that of the short-range interaction. However, it should be noticed that the cost of the error estimate scales as $O(N \log N)$. If the system only has short-range interaction that scales as $O(N)$, the load of the error estimate may become comparatively expensive when the system is large and eventually overwhelms the total computational cost as the system size grows to infinity. If the system also needs a long-range electrostatic calculation that is usually treated by the $O(N \log N)$ scaling Ewald-type algorithms [22–24], the extra load of the error estimate is likely to be acceptable. We want to stress again that the main research interest of the present paper is the error estimate, the proposed adaptive cutoff, and long-range correction methods. The examples are used to demonstrate the validity of these theoretical results. The applications of the error estimate is more general than the studied examples, because we assumed only the interaction is short range and treated by the cutoff

method. Other kind the short-range interaction (not only the Lennard-Jones but also the Buckingham interaction, the direct part of the Ewald family algorithms, etc.), and other kinds of inhomogeneity (not only planar and merging spherical interfaces) can be easily studied by the error estimates and the proposed methods. However, it is impossible to test all possible cases within one paper. The error study of more realistic systems requires the error estimates of the long-range algorithms and careful comparisons with the short-range error. Since the present paper is focusing on the short-range interactions, and space is limited, they will be given in future work.

ACKNOWLEDGMENTS

The authors gratefully thank Luigi Delle Site for useful comments on the manuscript. The work of P.Z. is supported by grants from National Natural Science Foundation of China (50930003).

-
- [1] D. Frenkel and B. Smit, *Understanding Molecular Simulation*, 2nd ed. (Academic Press, San Diego, 2002).
- [2] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Oxford Science Publications (Clarendon Press, Oxford, 1987).
- [3] D. Duque and L. F. Vega, *J. Chem. Phys.* **121**, 8611 (2004).
- [4] P. J. in 't Veld, A. E. Ismail, and G. S. Grest, *J. Chem. Phys.* **127**, 144711 (2007).
- [5] A. Trokhymchuk and J. Alejandre, *J. Chem. Phys.* **111**, 8510 (1999).
- [6] M. X. Guo and B. C. Lu, *J. Chem. Phys.* **106**, 3688 (1997).
- [7] M. Mecke, J. Winkelmann, and J. Fischer, *J. Chem. Phys.* **107**, 9264 (1997).
- [8] J. Janecek, *J. Phys. Chem. B* **110**, 6264 (2006).
- [9] F. Goujon, P. Malfreyt, J. M. Simon, A. Boutin, B. Rousseau, and A. H. Fuchs, *J. Chem. Phys.* **121**, 12559 (2004).
- [10] V. K. Shen, R. D. Mountain, and J. R. Errington, *J. Phys. Chem. B* **111**, 6198 (2007).
- [11] B. Smit, *J. Chem. Phys.* **96**, 8639 (1992).
- [12] J. J. Potoff and A. Z. Panagiotopoulos, *J. Chem. Phys.* **109**, 10914 (1998).
- [13] J. Pérez-Pellitero, P. Ungerer, G. Orkoulas, and A. D. Mackie, *J. Chem. Phys.* **125**, 054515 (2006).
- [14] J. Kolafa and J. W. Perram, *Mol. Simul.* **9**, 351 (1992).
- [15] H. Wang and P. Zhang, [arXiv:1208.0694](https://arxiv.org/abs/1208.0694), (2012).
- [16] G. Hummer, *Chem. Phys. Lett.* **235**, 297 (1995).
- [17] H. G. Petersen, *J. Chem. Phys.* **103**, 3668 (1995).
- [18] M. Deserno and C. Holm, *J. Chem. Phys.* **109**, 7694 (1998).
- [19] H. Wang, F. Dommert, and C. Holm, *J. Chem. Phys.* **133**, 034117 (2010).
- [20] P. P. Ewald, *Ann. Phys.* **369**, 253 (1921).
- [21] R. W. Hockney and J. W. Eastwood, *Computer Simulation Using Particles* (IOP, London, 1988).
- [22] T. Darden, D. York, and L. Pedersen, *J. Chem. Phys.* **98**, 10089 (1993).
- [23] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, *J. Chem. Phys.* **103**, 8577 (1995).
- [24] M. Deserno and C. Holm, *J. Chem. Phys.* **109**, 7678 (1998).
- [25] H. J. Limbach, A. Arnold, B. A. Mann, and C. Holm, *Comput. Phys. Commun.* **174**, 704 (2006).
- [26] H. Flyvbjerg and H.G. Petersen, *J. Chem. Phys.* **91**, 461 (1989).
- [27] S. Nosé, *Mol. Phys.* **52**, 255 (1984).
- [28] W. G. Hoover, *Phys. Rev. A* **31**, 1695 (1985).
- [29] L. J. Chen, *J. Chem. Phys.* **103**, 10214 (1995).
- [30] P. Orea, J. López-Lemus, and J. Alejandre, *J. Chem. Phys.* **123**, 114702 (2005).
- [31] F. Biscay, A. Ghoufi, F. Goujon, V. Lachet, and P. Malfreyt, *J. Chem. Phys.* **130**, 184710 (2009).
- [32] M. Kalweit and D. Drikakis, *Phys. Rev. B* **74**, 235415 (2006).