JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

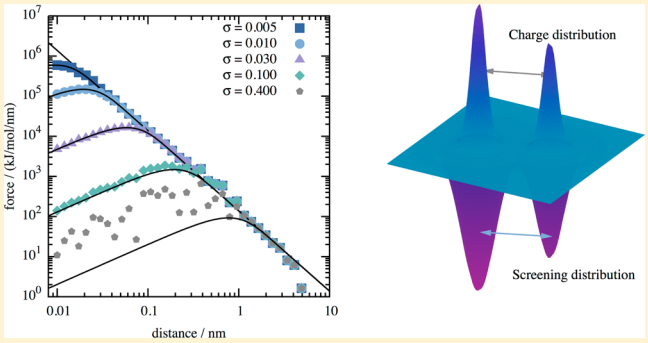
Efficient Handling of Gaussian Charge Distributions: An Application to Polarizable Molecular Models

Péter T. Kiss,[†] Marcello Sega,*[,‡] and András Baranyai[†]

[†]Institute of Chemistry, Eötvös University, 1518 Budapest 112, P.O. Box 32, Hungary

[‡]Department of Computational Biological Chemistry, Univiersity of Vienna, Währingerstr. 17, 1090 Vienna, Austria

ABSTRACT: We present a mesh-based Ewald summation method that is suitable for the calculation of the electrostatic interaction between Gaussian charge distributions, instead of point charges. As an application, we implemented the method in the Gromacs simulation package and tested it with a polarizable water model, showing that the interaction between Gaussian charge distributions can be computed with a small (10%) additional computational cost with respect to the point charge case. In addition, since the performance of polarizable models is strongly influenced by the number of iterations required for the self-consistent field solution, we tested also the Always Stable Predictor-Corrector (ASPC) method of Kolafa (Kolafa, J. *J. Comp. Chem.* **2003**, *25*, 335) as an alternative to the steepest descent (SD) based algorithm with predictor implemented in the Gromacs, and found that it speeds up the integration of the equations of motion by a factor of 1.6−2.0, depending on the target model.



I. INTRODUCTION

The usefulness of molecular simulations relies on two factors. These are the quality of the molecular models and the numerical algorithms that implement these models efficiently in actual computations. These two factors are inseparable because the inherent properties of a molecular model set the requirements of the underlying numerical methods. Presently, the vast majority of models used in classical, atomistic molecular dynamics simulations make use of point charges to describe the electrostatic interactions and of the Lennard-Jones potential to describe van der Waals interactions. The scientific community gradually canonized several models based on this approach because they represent easy-to-use, computationally cost-effective force fields. Some of the most notable water models are the SPC/E,[1] the TIP4P,[2] and the TIP3P[2] ones. These are rigid, nonpolarizable models fitted to some properties of ambient water, most typically, internal energy, density, and partial pair-correlation functions. The development of nonpolarizable water models is still in a very active phase, and recently, using the same philosophy, reparametrized, enhanced versions of older water models appeared, as is the case for the TIP4*P*/2005[3] one.

However, there are cases in which a rigid arrangement of partial charges is not enough to represent with the required accuracy the complexity of real systems. This is the case, for example, of ions at interfaces.[4] Another important case is that of water. Water is a highly polarizable molecule: its dipole moment can change as much as 40% when moving from gas to condensed phases. In the past three decades, several models have been proposed to mimic this property of water (for a short review see our previous paper[5]). Since the polarizability of water is nearly isotropic in the gas phase, and since it is believed that the same happens in condensed phases, two simple approaches are typically used to describe this phenomenon: the first consists in adding interaction centers with polarizable point dipoles, while the second, the so-called charge on spring (COS) method,[6] uses massless charged particles connected by classical harmonic springs to fixed points of the molecular frame. The polarization charges have no inherent dynamics and their position is determined by the mechanical equilibrium between the spring and the electrostatic forces. The advantage of the COS method with respect to the polarizable point dipole one is, that one does not have to calculate the more complicated dipole−dipole and charge−dipole interactions.

An important departure from the point charge paradigm was represented by the work of Chialvo and Cummings.[7] These authors introduced Gaussian charge distributions (which we would call Gaussian charges from now on, for simplicity) instead of point charges. Later on, the same idea was used in the GCP model[8] where three fixed Gaussians and a polarizable point dipole were employed to simulate water. The primary advantage of Gaussian charges, however, is their greater numerical stability. Point charges are prone to polarization catastrophe, which is theoretically impossible for Gaussian charge distributions. In COS models they usually need a counter-charge to decrease the chances of running into numerical instabilities. This, however, increases the number

Received: October 13, 2014
Published: November 7, 2014

   dx.doi.org/10.1021/ct5009069 | *J. Chem. Theory Comput.* 2014, 10, 5513−5519

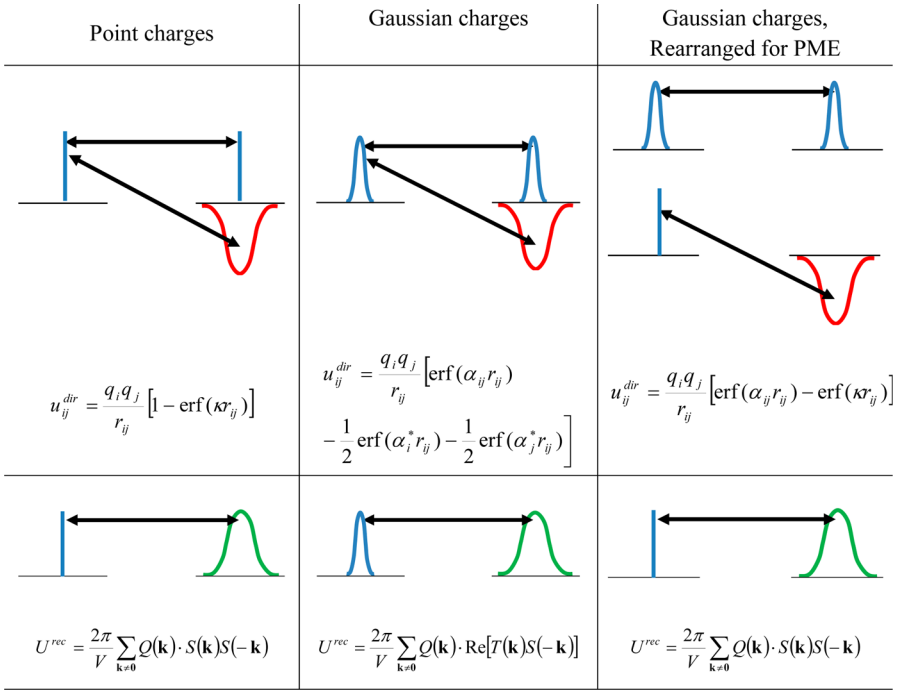


Figure 1. Schematic representation of the separation into direct and reciprocal space contribution in the Ewald sum for point charges and Gaussian charges. Top row, direct space part; bottom row, reciprocal space part interactions. The negative screening charge distribution in the direct space part is always added back in the reciprocal space. For Gaussian charges, the interaction between the screening charge and Gaussian charge distributions gives rise (central column) to a contribution in the reciprocal space not in the form of a Fourier transform. If the screening distribution interacts only with fictitious point charges (right column), once the contribution is added back in the reciprocal space, it takes the form of a charge structure factor.

of charged interacting sites and, consequently, the computational cost of the simulation. Moreover, adding a counter charge is sometimes not enough to stabilize the system. For example, for divalent cations, the introduction of an explicit damping of the electrostatic interactions was necessary to avoid polarization catastrophe with the SWM4-NDP model.[9]

Recently, we proposed a new polarizable model, BK3,[5] that uses three Gaussian charges and all charges take part in the polarization. We showed that upon proper parametrization of the free parameters of the model, very good estimates of structural, thermodynamic, and dynamical properties can be obtained all over the phase diagram. All polarizable water models, such as those of the COS B/G[10−12] or SWM[9,13−15] families, are noticeably more expensive in computational terms than the rigid ones, and it is therefore crucial to reduce the computational time of the two biggest bottleneck of these systems, namely, the calculation of electrostatic forces and the iterative procedure for the self-consistent field solution.

In this article, we show how the evaluation of the Ewald sum for Gaussian charges[16] can be approximated using a minimal modification of any mesh-based algorithms. We implemented and tested the accuracy and performance of the smooth Particle Mesh Ewald (sPME)[17] adapted for Gaussian charges using the Gromacs[18] simulation package v.5 at different levels, namely, that of tabulated potentials, dedicated C kernel, and dedicated Single Instruction Multiple Data (SIMD) kernel, showing that PME with Gaussian charges requires only 10% of additional computational cost with respect to the force calculation for point-charges. We also test the ASPC method of Kolafa,[19] a tool for speeding up the convergence of the self-consistent field calculation of the induced dipoles, with respect to the more conservative approach used in the Gromacs package (a steepest-descent type of algorithm supplemented with a predictor) and show that one can reach important speedups in the range 1.6−2.0, depending on the molecular model.

II. EWALD SUMMATION FOR GAUSSIAN CHARGES

A Gaussian charge is represented by the spherically symmetric charge distribution

$$\rho_i(\mathbf{r}) = \frac{q_i}{(2\pi\sigma_i^2)^{3/2}}\exp\left(\frac{-|\mathbf{r} - \mathbf{r}_i|^2}{2\sigma_i^2}\right) \tag{1}$$

which is characterized by a magnitude, or integrated charge q_i, a width σ_i, and is centered at the position of the i-th atom, $\mathbf{r}_i$, to which the distribution is associated. If a set of point charges is replaced with Gaussian charges of arbitrary width, none of the multipoles of the initial set is changed, due to the spherical symmetry of the distribution. The interaction energy u between two Gaussian charges is an analytically solvable integral,

$$u_{ij} = \int\int \frac{\rho_i(\mathbf{r}_1)\rho_j(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2 = \frac{q_i q_j}{r_{ij}}\mathrm{erf}(\alpha_{ij} r_{ij}) \tag{2}$$

where

$$\alpha_{ij} = \frac{\alpha_i\alpha_j}{\sqrt{\alpha_i^2 + \alpha_j^2}} = \frac{1}{\sqrt{2(\sigma_i^2 + \sigma_j^2)}} \tag{3}$$

is a function of the inverse distribution widths $\alpha_i = 1/(2\sigma_i^2)^{1/2}$, $\mathrm{erf}(x)$ is the error function, and r_{ij} is the distance between the two interaction sites. As the distribution width approaches zero, namely, $\sigma \to 0$, one recovers the usual formula for the Coulomb interaction, and if the distance between the Gaussian charges is

dx.doi.org/10.1021/ct5009069 | J. Chem. Theory Comput. 2014, 10, 5513−5519

large enough then they behave as point charges, as the error function in eq 2 tends to one.

The Ewald sum for the interaction energy between Gaussian charges can be obtained in the usual way, by introducing a Gaussian screening charge distributions of inverse width κ associated with each charge, so that the full interaction energy (for the tinfoil boundary conditions and disregarding the self-energy contribution) can be split into a direct and reciprocal space contributions. The direct part takes the form

$$U^{\mathrm{dir}} = \sum_{i=1}^{N}\sum_{j=i+1}^{N}\frac{q_i q_j}{r_{ij}}\left[\mathrm{erf}(\alpha_{ij} r_{ij}) - \frac{1}{2}\mathrm{erf}(\alpha_i^* r_{ij}) - \frac{1}{2}\mathrm{erf}(\alpha_j^* r_{ij})\right] \tag{4}$$

where

$$\alpha_i^* = \frac{\alpha_i \kappa}{\sqrt{\alpha_i^2 + \kappa^2}} \tag{5}$$

and it includes the interaction terms between pairs of Gaussian charges as well as the interaction term between Gaussian charges and screening distributions (see Figure 1 for a schematic representation).

The reciprocal part of the energy is

$$U^{\mathrm{rec}} = \frac{2\pi}{V}\sum_{\mathbf{k}\neq\mathbf{0}} Q(\mathbf{k})\cdot \mathrm{Re}[T(\mathbf{k})S(-\mathbf{k})] \tag{6}$$

where $\mathbf{k}$ is the wave vector in the reciprocal space, and the functions $Q(\mathbf{k})$, $T(\mathbf{k})$, and $S(\mathbf{k})$ are defined respectively as follows:

$$Q(\mathbf{k}) = \frac{1}{k^2}\exp\left(-\frac{\mathbf{k}^2}{4\kappa^2}\right) \tag{7}$$

$$S(\mathbf{k}) = \sum_{i=1}^{N} q_i \exp(i\mathbf{k}\mathbf{r}_i) \tag{8}$$

$$T(\mathbf{k}) = \sum_{i=1}^{N} q_i \exp\left(-\frac{\mathbf{k}^2}{4\alpha_i^2}\right)\exp(i\mathbf{k}\mathbf{r}_i) \tag{9}$$

The difference between the expression for Gaussian charges and point charges is the appearance of the $T(\mathbf{k})$ term in eq 6. Starting from eqs 4 and 6 it is straightforward to derive the corresponding force and pressure expressions.[16]

III. EWALD SUM FOR GAUSSIAN CHARGES ON A MESH

The replacement of the continuous charge distribution with a discrete one is at the heart of the various mesh-based variants of the Ewald sum. This replacement allows calculating the reciprocal space contribution by means of the Fast Fourier Transform (FFT) and exploiting the favorable $\sim N \log N$ scaling (compared to the $\sim N^{3/2}$ scaling of the Ewald sum). For the Gaussian charges variant, however, the presence of the $T(\mathbf{k})$ term is problematic, because it cannot be represented as a Fourier transform and, therefore, incorporated into a mesh-based scheme. However, it is possible to express the Ewald sum avoiding the appearance of the $T(\mathbf{k})$ term. The schematic representation in Figure 1 can be quite helpful to this end: in the regular Ewald scheme for Gaussian charges, one subtracts the screened Gaussian-charge contribution from the Gaussian–Gaussian interaction in real space (second panel of Figure 1) and then adds it back in the reciprocal space. Another possibility is to subtract only the contribution of a "virtual" screened point charge from the Gaussian–Gaussian interaction (third panel of Figure 1). In this way, screened point charges create a direct sum that vanishes at the cutoff, because at this distance the Gaussian–Gaussian interaction is identical with the interaction of point charges. In the reciprocal space this can then be compensated with the usual PME contribution, namely, without the corresponding $T(\mathbf{k})$ term. Written explicitly, the direct-space contribution to the energy to be used in a mesh-based scheme for Gaussian charges takes the simple form

$$U^{\mathrm{dir}} = \sum_{i=1}^{N}\sum_{j=i+1}^{N}\frac{q_i q_j}{r_{ij}}[\mathrm{erf}(\alpha_{ij} r_{ij}) - \mathrm{erf}(\kappa r_{ij})] \tag{10}$$

where the first term in eq 10 is the interaction energy of two unscreened Gaussian charges, and the second term is the interaction between two screened point charges. Since the reciprocal space contribution that compensates this kind of screening is just the same as in the mesh-based scheme of choice for the corresponding point charges, one can effectively use any mesh-based scheme to calculate the interaction between Gaussian charges, by simply rewriting the real space contribution to the energy and force according to eq 10. The expression for the real-space force contribution can be easily derived from eq 10. Our derivation turns out to be equivalent to that of Gingrich and Wilson[21] for the classical Ewald expression for the energy of Gaussian charges, which is also set in a form suitable for mesh-based algorithms, and indeed, it has been already used for a 2D periodic system[22] in the context of metallic surfaces. Here, we extensively test for the first time the accuracy of this approach as a function of the screening parameter and of the Gaussian width.

IV. IMPLEMENTATION OF PME FOR GAUSSIAN CHARGES FOR THE GROMACS SIMULATION PACKAGE

Gromacs is a highly efficient molecular simulation package that, among other features, exploits the internal parallelism of modern CPUs to maximize its performance: we implemented support for Gaussian charges in our in-house modified version 5 of the software, obtainable for free at http://marcello-sega.github.io/gromacs/. In the implementation of the direct space term, eq 10, we were faced with two possible routes, namely, (a) to write a new set of kernels for the calculation of the direct space term itself or (b) to exploit the tabulated potential features of Gromacs. Besides possible computational advantages of the explicit kernel implementation with respect to the tabulated potential, the former approach has the important feature that it does not require to recalculate the $S(S+1)/2$ tables of values for the interactions between S different atomic species, each time the Gaussian widths or the interaction cutoff is changed. For this reason we decided to implement a new set of kernels both at the portable C level, and for the SIMD units of modern CPUs. We modified the Gromacs input interface in order to read an additional, optional parameter (the width of the Gaussian distribution) from the atoms section of the topology files.

We first tested the accuracy of the method in the calculation of the electrostatic force between two Gaussian charges placed at various distances. In Figure 2, we present the force acting between two particles with charge $\pm e$ for different choices of

dx.doi.org/10.1021/ct5009069 | J. Chem. Theory Comput. 2014, 10, 5513–5519

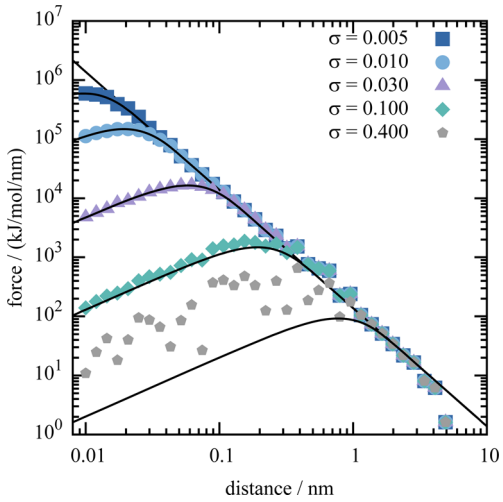


Figure 2. Force between two Gaussian charge distributions as a function of their mutual distance, for different value of the Gaussian width, σ. The force is calculated using the sPME algorithm adapted for Gaussian charges in a rectangular simulation box with edges 10, 10, and 20 nm along the x, y, and z directions, respectively. The Gaussian charge distributions are displaced along the x direction. Solid lines represent the analytical force−distance curves for a nonperiodic system.

the Gaussian width σ. The force is reported as a function of the distance between the two particles, which are placed in a rectangular simulation box with edges 10, 10, and 20 nm long in the x, y, and z-direction, respectively. The particles are displaced along the x-axis, and full 3D periodic boundary conditions are employed. Unlike the force between point-like charges that diverges as $1/r^2$ (straight, solid line in the double-logarithmic plot of Figure 2), the force between two Gaussian charge distributions vanishes when the distance between the centers of charge approaches zero. At distances larger than 3σ, the force curves start becoming indistinguishable from each other and from the usual PME force. Eventually, at distances comparable with the box size, the effect of periodic boundary conditions sets in, and the curves depart from the $1/r^2$ behavior, to reach zero at a distance equal to half of the box length. All curves with $\sigma < 0.1$ nm are indeed in excellent agreement with the analytical expression at short distances. When σ starts being comparable with the Ewald screening length $1/\kappa = 0.292043$ nm (for example in the case $\sigma = 0.1$ nm), some differences start appearing, and the agreement vanishes completely when σ is larger than $1/\kappa$, as in the case $\sigma = 0.4$ nm. The width of the Gaussian should not be commensurate with the screening because the charge distribution is damped and the interaction energy starts to diminish. This limitation has little practical effect, as usual charge distributions have a width, which is smaller than 0.1 nm, but it has nonetheless to be taken into account when choosing the Ewald screening length.

The choice of the Ewald screening parameter κ for an accurate calculation of the forces is not only subjected to an upper bound due the finite charge distribution σ. Similarly to what happens for other mesh-based methods,[22] the accuracy shows the presence of a marked minimum as a function of κ. In Figure 3, we report the relative root-mean-square error δF on the force for an equilibrated configuration of 512 BK3 water molecules. The error δF is calculated as

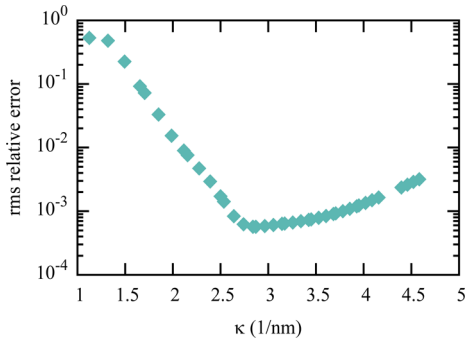


Figure 3. Root-mean-square relative error of the force calculated on a snapshot from an equilibrated simulation of the BK3 water model. The reference force is obtained from a well-converged Ewald calculation on the same configuration. The error is reported as a function of the Ewald screening parameter, κ.

$$\delta F = \sqrt{\frac{\sum_{i=1}^{N} |\mathbf{F}_i(\kappa) - \mathbf{F}_i^{ex}|^2}{\sum_{i=1}^{N} |\mathbf{F}_i^{ex}|^2}} \tag{11}$$

where the sum runs over all sites in the system. The force from a well-converged Ewald calculation using eqs 4 and 6 is used as the "exact" reference force, $\mathbf{F}^{ex}$. For the actual calculation, a screening parameter $\kappa = 3.42415$ 1/nm has been used, so that the screening length $1/\kappa = 0.292043$ nm is more than four times larger than the largest Gaussian charge distribution width in the BK3 model (0.071 nm). The relative error reaches its minimum $\delta F = 6 \times 10^{-4}$ at $\kappa = 2.967$ 1/nm.

V. SPEEDING UP THE COS ITERATION

It is mandatory for polarizable models to formulate a method that is able to move the COS particle to its mechanical equilibrium position with a small number of iterations. Since during the iterative procedure all the permanent charge−COS and COS−COS interactions have to be recalculated, the total time for the integration of the equations of motion is nearly proportional with the number of iterations, and this can cause a slowdown of a factor 3−7, depending on the target accuracy on the force that has to be reached. The correct estimate of the structural, dynamical, and thermodynamic properties of the model depends on the accuracy of the iterative procedure. One of the properties that are most sensitive to this accuracy is the pressure.

The potential energy of a polarizable force field depends on both the atomic positions and the induced dipoles. However, the latter are also function of the atomic positions: $U = U(\mathbf{r}^N, \{\boldsymbol{\mu}_i(\mathbf{r}^N)\}_{i=1}^N)$. The potential part of the microscopic pressure is

$$-\frac{dU}{dV} = -\sum_{i=1}^{N} \frac{d\mathbf{r}_i}{dV}\left(\frac{\partial U}{\partial \mathbf{r}_i}\right)_{\boldsymbol{\mu}_i=\text{const}} - \sum_{i=1}^{N} \frac{d\boldsymbol{\mu}_i}{dV}\left(\frac{\partial U}{\partial \boldsymbol{\mu}_i}\right)_{\mathbf{r}_i=\text{const}} \tag{12}$$

The first term in eq 12 corresponds to that of a pairwise additive, nonpolarizable force field with constant dipole moments, $\boldsymbol{\mu}_i$. The second term, instead, is the part that depends on polarization. The self-consistent iterative procedure is used to find the values of the dipole moments that minimize the energy of the system. For the self-consistent solution, the second term in eq 12 is zero, and the pressure of the polarizable system can be calculated as the pressure of a nonpolarizable one with constant dipole moments $\boldsymbol{\mu}_i$. However, if the mechanical

dx.doi.org/10.1021/ct5009069 | J. Chem. Theory Comput. 2014, 10, 5513−5519

equilibrium of the COS particle is far from perfect, an additional force will contribute to the pressure.[23]

This discussion points out that it is dangerous to decrease the target accuracy of the iterative procedure to save computer time, as several thermodynamic quantities could be evaluated wrongly. The question is how is it possible to decrease the number of iterations while maintaining the required accuracy. Ten years ago, Kolafa suggested a predictor−corrector scheme that maintains the time reversal symmetry of the equations of motion.[19] Let

$$\boldsymbol{\mu}^P(t+h) = \sum_{i=0}^{m-1} \binom{m}{i+1} (-1)^i \boldsymbol{\mu}(t-ih) + \sum_{j=1}^{k} A_j \sum_{i=0}^{m} \binom{m}{i} (-1)^i \boldsymbol{\mu}(t-(i+j-1)h) \tag{12}$$

be the predicted induced dipole vector of a COS particle for the next integration step $t + h$, where h is the time step. Kolafa determined the coefficients A_j that keep the time reversal symmetry by expanding the expression for the predicted dipole around t, ensuring that the odd terms disappear. The method was determined for the $m = 2$ case. As a result, he obtained the following set of equations:

$$\begin{aligned}
k=0 \quad & \boldsymbol{\mu}^P(t+h) = 2\boldsymbol{\mu}(t) - \boldsymbol{\mu}(t-h) \\
k=1 \quad & \boldsymbol{\mu}^P(t+h) = 2.5\boldsymbol{\mu}(t) - 2\boldsymbol{\mu}(t-h) + 0.5 \times \boldsymbol{\mu}(t-2h) \\
k=2 \quad & \boldsymbol{\mu}^P(t+h) = 2.8\boldsymbol{\mu}(t) - 2.8\boldsymbol{\mu}(t-h) + 1.2\boldsymbol{\mu}(t-2h) - 0.2\boldsymbol{\mu}(t-3h) \\
k=3 \quad & \boldsymbol{\mu}^P(t+h) = 3\boldsymbol{\mu}(t) - \frac{24}{7}\boldsymbol{\mu}(t-h) + \frac{27}{14} \times \boldsymbol{\mu}(t-2h) - \frac{4}{7}\boldsymbol{\mu}(t-3h) + \frac{1}{14}\boldsymbol{\mu}(t-4h) \\
k=4 \quad & \boldsymbol{\mu}^P(t+h) = \frac{22}{7}\boldsymbol{\mu}(t) - \frac{55}{14}\boldsymbol{\mu}(t-h) + \frac{55}{21}\boldsymbol{\mu}(t-2h) - \frac{22}{21}\boldsymbol{\mu}(t-3h) + \frac{5}{21} \times \boldsymbol{\mu}(t-4h) - \frac{1}{42}\boldsymbol{\mu}(t-5h)
\end{aligned} \tag{13}$$

The corrector term that determines the position of the polarization charges at the $(n + 1)$-th iteration step is then cast in the form of a damped relaxation step toward the self-consistent solution,

$$\boldsymbol{\mu}^{(n+1)} = \omega\alpha\mathbf{E}(\boldsymbol{\mu}^{(n)}) + (1-\omega)\boldsymbol{\mu}^{(n)} \tag{14}$$

where α is here the polarizability tensor. For the first iteration step the predicted dipoles are used ($\boldsymbol{\mu}^{(0)} = \boldsymbol{\mu}^P$). In the original version of the method, Kolafa used eq 14 only once (one force calculation per time step), and derived an optimal damping factor, ω, which guaranties the stability of this predictor−corrector scheme. Using damping factors higher than the optimal one, the induced dipoles will be more accurate, but upon further increase of the factor, the ASPC method becomes unstable.

The current scheme implemented in Gromacs is based on a modified steepest descent (SD) that takes advantage of an initial prediction of the shell particle position using the velocities of the associated real atoms. An adaptive choice of the displacement vector is employed in successive iterations, which are repeated until the root-mean-square force on the COS particle reaches a certain threshold. The precision commonly required for polarizable models is in the order 0.1−1 kJ/mol/nm. To make a concrete example, this tolerance requires, for the BK3 water model, additional 4−6 force calculations, making the simulation of the polarizable model roughly the same 4−6 factor slower than in absence of polarization charges. To increase the performance of the self-consistent field calculation in Gromacs, we implemented in the code a modified version of the ASPC algorithm. We noticed that by using the original ASPC method (without iterations) the residual forces on COS particles are usually above the required 0.1−1 kJ/mol·nm. Therefore, we decided to keep relaxing the system using eq 14 until convergence is reached. The number of the necessary iteration steps depends on the ω parameter and, in particular, on the value of ω in the first iteration step. In our implementation, we set ω adaptively at the beginning of the simulation, by simply collecting statistics over a predefined set of values. For the subsequent iteration steps, ω was set to 0.9. In Figure 4 we report the average time required to perform a full MD step, as a function of the target tolerance, using Gromacs with native SD, and with ASPC. The speedup so obtained is always in favor of ASPC and in the typical tolerance range 0.1−1 kJ/mol·nm, the performance increases by a factor of 1.6−1.7. The number of additional force calculations required in the same tolerance range is 3−4 for ASPC and 5−7 for SD, which also shows that the overhead of the ASPC calculation is negligible. The downside of ASPC is that it has a larger storage requirement than SD, because the positions of the polarizable atoms in the previous 5 timesteps have to be stored, while in the current version of the SD only the previous positions and forces have to be stored. The speedup is not universal, and depends on the model. With the SWM4-DP[14] model, for example, a speedup of 2 can be reached for a tolerance of 0.1 kJ/mol·nm.

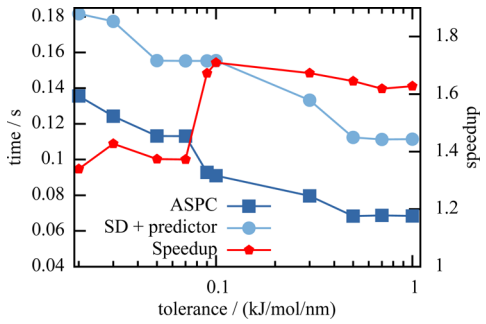


Figure 4. Average time required to perform a molecular dynamics step for the system of 512 water molecules (BK3 model) step using the ASPC (squares) and SD (circles) algorithms. The speedup of ASPC with respect to SD is also reported (pentagons). Lines are a guide to the eye.

VI. PERFORMANCE

We tested our implementation of PME for Gaussian charges on different systems, to check independently the performance of the new C and SIMD kernels with respect to the native PME ones, and to test the overall performance of the BK3 water model. All tests were performed on a Intel Xeon CPU E5530 with clock frequency 2.40 GHz, using a single core, and no

dx.doi.org/10.1021/ct5009069 | J. Chem. Theory Comput. 2014, 10, 5513−5519

OpenMP parallelization. On this machine, the SIMD-optimized kernels of Gromacs, as well as the newly written, SIMD-optimized kernel, exploited the SSE4.1 instruction set of the CPU. In the first type of test, we simulated an equilibrated box of 512 TIP5P[24] water molecules making use of the regular point charges or, substituting them with Gaussian charges (σ = 0.05 nm). In this way, the performance of the new kernels can be checked directly against PME in the same system. Since no optimized kernel for the TIP5P geometry is present in Gromacs, the enhanced SIMD performance comes only from the more efficient PME calculation on the SSE unit of the CPU. The results, presented in Table 1, show that the Gromacs SIMD kernel does in fact more than double the speed of the TIP5P simulation. The kernels for the Gaussian charges determine a loss of about 7% in performance with respect to point charges (for the SIMD kernels), which we deem satisfactory. Since the BK3 water model uses a Buckingham potential to model the van der Waals interaction, we are also interested in measuring the performance loss with respect of the Lennard-Jones potential (roughly 10%), which is also reported in Table 1 for a modified TIP5P model. In the second type of test, we compare the performance of the BK3 model with two other modern polarizable potentials, namely, SWM4-DP[13] and SWM6.[15] The required time to integrate the equations of motion for the BK3 model is comparable to that of SWM4-DP. If we take into account that the direct part of the Ewald sum is proportional to the square of the number of particles within the cutoff, the integration of the equations of motion for the 3 charged sites of BK3 should be, in principle, much faster then for SWM4-DP, with its 5 charged sites. Both the Buckingham potential and the Gaussian character of the charges surely increase the computational requirements of the BK3 model, but the biggest contribution is due to the fact that, on average, the BK3 model requires about 1.5 times more iterations than the SWM4-DP one to reach the same target tolerance, due to the presence of 3 polarizable charges. The SWM6 model is instead substantially slower than the BK3 model.

Table 1. Computational Performance of Different Water Models

model	sites/ charges	charge type	van der Waals type	performance (ns/day) SIMD kernel	performance (ns/day) C kernel
TIP5P	5/4	point	Lennard-Jones	14.4	6.1
TIP5P	5/4	Gauss	Lennard-Jones	13.6	5.2
TIP5P	5/4	point	Buckingham	11.8	5.7
SWM4-DP	5/5	point	Lennard-Jones	3.1	1.4
SWM6	7/7	point	Lennard-Jones	2.2	1.0
BK3	7/3	Gauss	Buckingham	3.0	1.4
BK3	7/3	Gauss (tabulated)	Buckingham	2.7	

VII. CONCLUSIONS

Classical, semiempirical models can serve well when large systems or complicated molecules are described. However, these models should be as accurate as possible at a reasonable computational cost. The importance of using charge distributions rather than point charges, to build molecular models, has been recently pointed out by several authors who used exponential[25,26] or Gaussian charge distributions.[5,7,8] Here, we showed how to recast the Ewald sum for a system of Gaussian charges to be suitable for mesh-based Ewald methods. We implemented the scheme in the Gromacs simulation package, and tested its accuracy and performance, showing that the Gaussian charge interaction can be computed with a little overhead. In addition, we showed how the calculation of the self-consistent field solution for polarizable models can be improved by using the ASPC predictor-corrector method of Kolafa[19] by a factor 1.6−2.0, depending on the target model, at the price of increased memory requirements. These results show that there is still much space for improvement for polarizable models with diffuse charges, and the barrier of the high computational costs, which has until now prevented a broader diffusion of these approaches could be progressively overcome, opening new perspectives for the development of more sophisticated, generic semiempirical force fields and for their adoption by an increasing number of researchers.

AUTHOR INFORMATION

Corresponding Author

*E-mail: marcello.sega@gmail.com.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

A.B. gratefully acknowledges the support of OTKA grant K84382. M.S. acknowledges support from the European Community's Seventh Framework Programme (FP7-PEOPLE-2012-IEF) funded under grant No. 331932 SIDIS, and D. van der Spoel and B. Hess for helpful discussions.

REFERENCES

(1) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987**, *91*, 6269−6271.
(2) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(3) Abascal, J. L.; Vega, C. A General Purpose Model for the Condensed Phases of Water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123*, 234505.
(4) Jungwirth, P.; Tobias, D. J. Ions at the Air/Water Interface. *J. Phys. Chem. B* **2002**, *106*, 6361−6373.
(5) Kiss, P. T.; Baranyai, A. A Systematic Development of a Polarizable Potential of Water. *J. Chem. Phys.* **2013**, *138*, 204507.
(6) Straatsma, T. P.; McCammon, J. A. Molecular Dynamics Simulations with Interaction Potentials Including Polarization Development of a Noniterative Method and Application to Water. *Mol. Simul.* **1990**, *5*, 181−192.
(7) Chialvo, A. A.; Cummings, P. T. Simple Transferable Intermolecular Potential for the Molecular Simulation of Water Over Wide Ranges of State Conditions. *Fluid Phase Equlib.* **1998**, *150*, 73−81.
(8) Paricaud, P.; Předota, M.; Chialvo, A. A.; Cummings, P. T. From Dimer to Condensed Phases at Extreme Conditions: Accurate Predictions of the Properties of Water by a Gaussian Charge Polarizable Model. *J. Chem. Phys.* **2005**, *122*, 244511.
(9) Yu, H.; Whitfield, T. W.; Harder, E.; Lamoureux, G.; Vorobyov, I.; Anisimov, V. M.; MacKerell, A. D., Jr.; Roux, B. Simulating Monovalent and Divalent Ions in Aqueous Solution using a Drude Polarizable Force Field. *J. Chem. Theory. Comput.* **2010**, *6*, 774−786.

(10) Yu, H.; Hansson, T.; van Gunsteren, W. F. Development of a Simple, Self-Consistent Polarizable Model for Liquid Water. *J. Chem. Phys.* **2003**, *118*, 221−234.
(11) Yu, H.; van Gunsteren, W. F. Charge-on-Spring Polarizable Water Models Revisited: From Water Clusters to Liquid Water to Ice. *J. Chem. Phys.* **2004**, *121*, 9549−9564.
(12) Kunz, A. P. E.; van Gunsteren, W. F. Development of a Nonlinear Classical Polarization Model for Liquid Water and Aqueous Solutions: COS/D. *J. Phys. Chem. A* **2009**, *113*, 11570.
(13) Lamoureux, G.; MacKerell, A. D., Jr.; Roux, B. A Simple Polarizable Model of Water Based on Classical Drude Oscillators. *J. Chem. Phys.* **2003**, *119*, 5185−5197.
(14) Lamoureux, G.; Harder, E.; Vorobyov, I. V.; Roux, B.; MacKerell, A. D., Jr. A Polarizable Model of Water for Molecular Dynamics Simulations of Biomolecules. *Chem. Phys. Lett.* **2006**, *418*, 245−249.
(15) Yu, W.; Lopes, P. E.; Roux, B.; MacKerell, A. D., Jr. Six-Site Polarizable Model of Water Based on the Classical Drude Oscillator. *J. Chem. Phys.* **2013**, *138*, 034508.
(16) Baranyai, A.; Kiss, P. T. A Transferable Classical Potential for the Water Molecule. *J. Chem. Phys.* **2010**, *133*, 144109.
(17) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(18) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit. *Bioinformatics* **2013**, *29*, 845.
(19) Kolafa, J. Time-Reversible Always Stable Predictor−Corrector Method for Molecular Dynamics of Polarizable Molecules. *J. Comput. Chem.* **2003**, *25*, 335−342.
(20) Gingrich, T. R.; Wilson, M. On the Ewald Summation of Gaussian Charges for the Simulation of Metallic Surfaces. *Chem. Phys. Lett.* **2010**, *500*, 178−183.
(21) Vatamanu, J.; Borodin, O.; Smith, G. D. Molecular Simulations of the Electric Double Layer Structure, Differential Capacitance, and Charging Kinetics for N-methyl-N-propylpyrrolidinium Bis-(fluorosulfonyl)imide at Graphite Electrodes. *J. Phys. Chem. B* **2011**, *115*, 3073−3084.
(22) Deserno, M.; Holm, C. How to Mesh Up Ewald Sums. I. A Theoretical and Numerical Comparison of Various Particle Mesh Routines. *J. Chem. Phys.* **1998**, *109*, 7678−7693.
(23) Kiss, P. T.; Baranyai, A. On the Pressure Calculation for Polarizable Models in Computer Simulation. *J. Chem. Phys.* **2012**, *136*, 104109.
(24) Mahoney, M. W.; Jorgensen, W. L. A Five-Site Model for Liquid Water and the Reproduction of the Density Anomaly by Rigid, Nonpolarizable Potential Functions. *J. Chem. Phys.* **2000**, *112*, 8910−8922.
(25) Saint-Martin, H.; Hernández-Cobos, J.; Bernal-Uruchurtu, M. I.; Ortega-Blake, I.; Berendsen, H. J. C. A Mobile Charge Densities in Harmonic Oscillators (MCDHO) Molecular Model for Numerical Simulations: The Water−Water Interaction. *J. Chem. Phys.* **2000**, *113*, 10899−10912.
(26) Villa, A.; Hess, B.; Saint-Martin, H. Dynamics and Structure of Ln (III)−Aqua Ions: A Comparative Molecular Dynamics Study Using Ab Initio Based Flexible and Polarizable Model Potentials. *J. Phys. Chem. B* **2009**, *113*, 7270−7281.