The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

Stress tensor and constant pressure simulation for polarizable Gaussian multipole model

Cite as: J. Chem. Phys. 156, 114114 (2022); doi: 10.1063/5.0082548
Submitted: 16 December 2021 • Accepted: 25 February 2022 •
Published Online: 17 March 2022

Haixin Wei,[1] Piotr Cieplak,[2] Yong Duan,[3] and Ray Luo[1,a)]

AFFILIATIONS

[1] Departments of Materials Science and Engineering, Molecular Biology and Biochemistry, Chemical and Biomolecular Engineering, and Biomedical Engineering, Graduate Program in Chemical and Materials Physics, University of California, Irvine, Irvine, California 92697, USA

[2] SBP Medical Discovery Institute, 10901 North Torrey Pines Road, La Jolla, California 92037, USA

[3] UC Davis Genome Center and Department of Biomedical Engineering, University of California, Davis, One Shields Avenue, Davis, California 95616, USA

[a)] **Author to whom correspondence should be addressed:** ray.luo@uci.edu

ABSTRACT

Our previous article has established the theory of molecular dynamics (MD) simulations for systems modeled with the polarizable Gaussian multipole (pGM) electrostatics [Wei *et al.*, J. Chem. Phys. **153**(11), 114116 (2020)]. Specifically, we proposed the covalent basis vector framework to define the permanent multipoles and derived closed-form energy and force expressions to facilitate an efficient implementation of pGM electrostatics. In this study, we move forward to derive the pGM internal stress tensor for constant pressure MD simulations with the pGM electrostatics. Three different formulations are presented for the flexible, rigid, and short-range screened systems, respectively. The analytical formulations were implemented in the SANDER program in the Amber package and were first validated with the finite-difference method for two different boxes of pGM water molecules. This is followed by a constant temperature and constant pressure MD simulation for a box of 512 pGM water molecules. Our results show that the simulation system stabilized at a physically reasonable state and maintained the balance with the externally applied pressure. In addition, several fundamental differences were observed between the pGM and classic point charge models in terms of the simulation behaviors, indicating more extensive parameterization is necessary to utilize the pGM electrostatics.

Published by AIP Publishing. https://doi.org/10.1063/5.0082548

I. INTRODUCTION

Atomistic simulations of biomolecules have been applied to various biological systems.[2] While additive nonpolarizable models continue to play dominant roles,[3–5] nonadditive polarizable models are emerging as a tool extending our capability of studying more complex biomolecular systems and processes. The nonpolarizable models typically use fixed atomic partial charges to model electrostatics. As a result, they incorporate only the polarization response to the environment (mostly in water) in a mean-field manner. Thus, the nonpolarizable models with excellent descriptions of the homogeneous bulk phase can be poor models treating gas-phase clusters or in cases where large scale conformational changes are accompanied by changes in dielectric environment.[6] Another limitation of the nonpolarizable models is the use of partial atomic charges. This treatment often lacks the sufficient mathematical flexibility to accurately describe the electrostatic potential of nearby molecules. Williams showed that the optimal least-squares fitting of atom-centered partial charges resulted in relative root-mean-square errors of 3%–10% in reproducing the quantum electrostatic potentials at the grid points in a shell outside the surface of a series of small polar molecules.[7] In summary, the importance of modeling polarization effects and the limitation of point charge models have been well documented.[6,8]

To address the aforementioned issues, extensive efforts have been invested in the development of explicit polarizable models.[9–11] Several methods have been proposed, including explicit polarization in the molecular mechanics potentials, such as the Drude oscillator,[12,13] fluctuating charges,[14] and induced dipoles.[6,15,16] Among these new models, the implementation of polarizable point dipoles was originally proposed by Applequist *et al.*[17] It is a classical approach with a long history in molecular

simulation.[18] However, this model suffers from the so-called "polarization catastrophe,"[19] which requires proper screenings for short range electrostatics.

Recently, Elking *et al.* proposed the polarizable multipole model with Gaussian charge densities.[20] A key benefit of the polarizable Gaussian Multipole model (pGM) is its screening of all the short-range electrostatic interactions in a physically consistent manner, as we discussed previously.[21] We have presented an analytical formulation for the pGM model that is more suitable for molecular dynamics (MD) simulation due to the lack of explicit torque computation.[1] In addition, we further exploited a new local frame, the covalent basis vectors (CBVs), to preserve the physical symmetry and to better accommodate the fact that the permanent dipoles primarily resulted from the covalent bonds. Subsequent numerical tests show that, with enough accuracy involving both the pGM and the Particle Mesh Ewald (PME), the NVE simulation can be realized with a similar energy drift as in classic point charge models.[1]

However, most of the real-world biological systems resemble constant pressure ensembles, which require our pGM model to be extended to NPT simulations. Andersen first proposed a new constant pressure method in which the volume was allowed to fluctuate, with its average value being determined by the balance between the internal stress in a system and the externally set pressure.[22] Following his pioneering works, Nosé and Klein[23] and Darden[24] indicated how to extend the formulation to include the case of the long range charge–charge interactions. Toukmaji *et al.* introduced an efficient approach to include induced dipolar interactions,[25] and Sagui *et al.* extended the formulation to multipolar situations.[26] Recently, these virials for point induced dipoles and multipoles have been implemented for AMOEBA in Tinker-HP and Tinker9 [graphics processing unit (GPU)].[27] Additionally, there have been interesting discussions about pressure calculation related to the Ewald surface correction.[28,29] However, how to rigorously treat internal stress in a system with both permanent and induced dipoles is still unclear.

In this work, we present an analytical formulation of the stress tensor for both the flexible and rigid molecular systems modeled with the pGM model. We performed a finite difference test and confirmed that both are rigorously correct. Finally, we implemented the proposed pGM formalism in the context of PME into the SANDER program in Amber. Our NPT simulations of a 512-water system showed that the algorithm is stable as indicated by the time evolutions of pressure, temperature, density, and total system energy.

II. METHOD

A. Electrostatic potential of a pGM system

The total electrostatic energy of the system, as shown in our previous work,[1] is defined as

$$U_{pGM} = \sum_i^N \frac{1}{2}(q_i + \vec{\mu}_i \cdot \nabla_i)\phi_i^0 + \sum_i^N \frac{1}{2}(\vec{p}_i \cdot \nabla_i)\phi_i^0. \tag{1}$$

Here, q_i, $\vec{\mu}_i$, and $\vec{p}_i$ are the charge, covalent dipole (i.e., permanent dipole), and induced dipole on atom i, respectively, and ϕ_i^0 is the electrostatic potential on atom i that is generated only by charges and covalent dipoles in the system. Later in this paper, we will denote ϕ_i as the electrostatic potential on atom i that is generated by charges, covalent dipoles, and *induced dipoles* in the system.

Due to the periodic boundary condition, ϕ_i^0 has to be obtained through a certain lattice summation technique, and here, the Ewald summation technique is used.

According to the Ewald summation,

$$\phi_i^0 = \frac{1}{\pi V}\sum_{\vec{m}\neq 0}\left\{\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\sum_j^N\left[\left(q_j + 2\pi i\vec{m}\cdot\vec{\mu}_j\right)\exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\right]\right.$$
$$\left.\times\exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right)\right\} + \sum_{j\neq i}^{\infty}(q_j + \vec{\mu}_j\cdot\nabla_j)\frac{\mathrm{erf}(\beta_{ij}R_{ij}) - \mathrm{erf}(\beta_0 R_{ij})}{R_{ij}}$$
$$-\lim_{\vec{R}_j\to\vec{R}_i}(q_i + \vec{\mu}_i\cdot\nabla_j)\frac{\mathrm{erf}(\beta_0 R_{ij})}{R_{ij}}. \tag{2}$$

Here, V is the volume of the box, $\vec{m}$ is the reciprocal lattice vector, and $\vec{R}_i$ and $R_{ij} = \left|\vec{R}_j - \vec{R}_i\right|$ are the coordinate of atom i and its distance to atom j, respectively. In addition, β_0 is the Ewald coefficient, whose value is usually between $\frac{1}{5}$–$\frac{1}{2}$ Å^{-1} for typical biomolecular applications, and β_{ij} equals to $\frac{\beta_i\beta_j}{\sqrt{\beta_i^2+\beta_j^2}}$, where β_i and β_j are the Gaussian parameters for atoms i and j, respectively.

Below, we will use the following notation convention: all gradients paired with a dipole (covalent or induced) only apply to atom coordinates. For example, the gradient operator in $\vec{\mu}_i\cdot\nabla_i$ only acts on the coordinates of atom i. On the other hand, all other gradient operators (not paired with a dipole) follow the standard convention.

B. Stress tensor

From its definition, stress tensor (virial) is the derivative of system Lagrangian with respect to the system size parameters, except for a constant pre-factor. In a typical biomolecular simulation, the simulating box can be described by the three column vectors that represent the edges of the box, $\vec{u}_1$, $\vec{u}_2$, and $\vec{u}_3$. Thus, a matrix h can be constructed from these three unit cell vectors, defined as the system shape parameters,

$$h = (\vec{u}_1, \vec{u}_2, \vec{u}_3). \tag{3}$$

Thus, stress tensor $\vec{\vec{V}}$ can be expressed as

$$\vec{\vec{V}}\cdot\left(h^{-1}\right)^T = \frac{\partial L}{\partial h} = \frac{\partial T}{\partial h} - \frac{\partial U}{\partial h}. \tag{4}$$

Here, L represents the system Lagrangian, superscript "T" denotes transpose of the matrix, and T and U are kinetic and potential energies, respectively. The derivative with respect to matrix h is equivalent to differentiating L against each component of h, with the fractional coordinates of each atom held constant.

Because the kinetic energy expression is the same for a pGM system and a classic point charge system, the related virial terms are the same, so we omit their derivations here. There are many different

components in the potential energy, but only the electrostatic part is different from other force fields in the current pGM model. Thus, we need to derive only the electrostatic components of the virial. This is to say, we are deriving the following quantity in this work:

$$\overleftrightarrow{V}_{pGM} = -\frac{\partial U_{pGM}}{\partial h} \cdot (h)^T. \tag{5}$$

To do so, we first need to identify which quantities in Eq. (1) are dependent on matrix h.

C. Stress tensor in a pGM system

Quantities that obviously depend on h are as follows:

1. Atomic coordinates, $\vec{R}_i = s_{i1}\vec{u}_1 + s_{i2}\vec{u}_2 + s_{i3}\vec{u}_3$. Here, s's are fractional coordinates and are constants when computing the derivatives with respect to h.
2. Atomic covalent dipoles, $\vec{\mu}_i = \sum_k^n u_{ik}\frac{\vec{R}_{ik}}{|\vec{R}_{ik}|} = \sum_k^n u_{ik}\frac{\vec{R}_k - \vec{R}_i}{|\vec{R}_{ik}|}$. Here, n refers to the number of atoms covalently connected to atom i (this notation will be used thoroughly in the paper), and u_{ik} is the dipole moment along the specific direction $\vec{R}_k - \vec{R}_i$, which is a constant. Noted that the definition of $\vec{R}_{ik} \equiv \vec{R}_k - \vec{R}_i$ is used throughout this manuscript.
3. System volume, $V = \det(h)$.
4. Reciprocal lattice vector, $\vec{m} = (k_1, k_2, k_3) \cdot h^{-1}$, where k_1, k_2, k_3 are integer constants.

The derivatives of these quantities with respect to h are given in Appendix A 1.

Moreover, the induced dipole, $\vec{p}_i$, is clearly dependent on h. However, there is no explicit expression of $\vec{p}_i$, as $\vec{p}_i$ is computed through a self-consistent iteration. In the following, we show how to obtain its derivatives.

As shown in our previous work,[1] the induced dipole can be expressed in the periodic system as

$$\begin{aligned}\vec{p}_i &= \alpha_i\left(\vec{E}_i^0 + \vec{E}_i^{induced}\right) = \alpha_i(\vec{E}_i^0 - \sum_j^N \vec{T}_{ij}\cdot\vec{p}_j),\\ \vec{E}_i^0 &= -\nabla_i\phi_i^0,\end{aligned} \tag{6}$$

$$T_{ij}^{\alpha\beta} = \partial_i^\alpha\partial_j^\beta\begin{cases}\frac{1}{\pi V}\sum_{\vec{m}\neq 0}\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right) + \frac{\mathrm{erf}(\beta_{ij}R_{ij}) - \mathrm{erf}(\beta_0 R_{ij})}{R_{ij}} & \text{if } i\neq j,\\ \frac{1}{\pi V}\sum_{\vec{m}\neq 0}\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right) - \lim_{\vec{R}_j\to\vec{R}_i}\frac{\mathrm{erf}(\beta_0 R_{ij})}{R_{ij}} & \text{if } i = j.\end{cases}$$

Here $\alpha, \beta = 1, 2, 3$ refer to x, y, z directions, respectively. The expression of $T_{ij}^{\alpha\beta}$ comes from the Ewald summation of $\partial_i^\alpha\partial_j^\beta\frac{\mathrm{erf}(\beta_{ij}R_{ij})}{R_{ij}}$. Thus, the summation for $i \neq j$ includes all the images in the periodic system. The introduction of

$$A_{ij}^{\alpha\beta} = \frac{1}{\alpha_i}(\delta_{ij}^{\alpha\beta} + \alpha_i \cdot T_{ij}^{\alpha\beta}) \tag{7}$$

leads to the following short-hand expression for the induced dipole:[1]

$$\vec{p}_i = \sum_j^N \vec{A}_{ij}^{-1}\cdot\vec{E}_j^0. \tag{8}$$

In Eq. (8), we have transferred the dependence of $\vec{p}_i$ on h to the dependence of $\vec{A}_{ij}^{-1}$ and $\vec{E}_j^0$ on h, and the dependence of $\vec{E}_j^0 = -\nabla_j\phi_j^0$ is already known. Thus, what is needed is to derive the dependence of $\vec{A}_{ij}^{-1}$ on h.

Given $A \cdot A^{-1} = 1$, we have

$$\begin{aligned}\frac{\partial(A^{-1})_{ij}^{\alpha\beta}}{\partial h} &= -(A^{-1})_{ii'}^{\alpha\alpha'}\frac{\partial A_{i'j'}^{\alpha'\beta'}}{\partial h}(A^{-1})_{j'j}^{\beta'\beta}\\ &= -(A^{-1})_{ii'}^{\alpha\alpha'}\frac{\partial T_{i'j'}^{\alpha'\beta'}}{\partial h}(A^{-1})_{j'j}^{\beta'\beta},\end{aligned} \tag{9}$$

where the Einstein's index notation is employed for i', j', α', and β'. We can now proceed to compute the derivatives for the energy term related to $\vec{p}_i$ in Eq. (1) as follows:

$$\begin{aligned}\frac{\partial - \sum_i^N \frac{1}{2}\vec{p}_i\cdot\vec{E}_i^0}{\partial h} &= \frac{1}{2}E_i^{0\alpha}(A^{-1})_{ii'}^{\alpha\alpha'}\frac{\partial T_{i'j'}^{\alpha'\beta'}}{\partial h}(A^{-1})_{j'j}^{\beta'\beta}E_j^{0\beta} - \sum_i^N\vec{p}_i\cdot\frac{\partial\vec{E}_i^0}{\partial h}\\ &= \frac{1}{2}p_{i'}^{\alpha'}\frac{\partial T_{i'j'}^{\alpha'\beta'}}{\partial h}p_{j'}^{\beta'} - \sum_i^N\vec{p}_i\cdot\frac{\partial\vec{E}_i^0}{\partial h}.\end{aligned} \tag{10}$$

Again, the Einstein's index notation is employed here for i, j, α, β i', j', α', are β'.

With the above preparations, we can calculate the electrostatic energy part of the virial from its definition,

$$\overleftrightarrow{V}\cdot\left(h^{-1}\right)^T = -\frac{\partial U}{\partial h_{\mu\nu}} = -\left(\frac{\partial U_{rec}}{\partial h_{\mu\nu}} + \frac{\partial U_{dir}}{\partial h_{\mu\nu}} + \frac{\partial U_{self}}{\partial h_{\mu\nu}}\right), \tag{11}$$

$$\frac{\partial U_{rec}}{\partial h_{\mu\nu}} = \frac{\partial \sum_i^N \frac{1}{2}\left(q_i + \vec{\mu}_i \cdot \nabla_i\right) \frac{1}{\pi V} \sum_{\vec{m}\neq 0} \left\{ \frac{\exp\left(-\frac{\pi^2 \vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2} \sum_j^N \left[\left(q_j + 2\pi i \vec{m} \cdot \vec{\mu}_j\right) \exp\left(2\pi i \vec{m} \cdot \vec{R}_j\right)\right] \exp\left(-2\pi i \vec{m} \cdot \vec{R}_i\right)\right\}}{\partial h_{\mu\nu}}$$
$$+ \frac{1}{2} p_{i'}^{\alpha'} \frac{\partial \partial_{i'}^{\alpha'} \partial_{j'}^{\beta'} \frac{1}{\pi V} \sum_{\vec{m}\neq 0} \frac{\exp\left(-\frac{\pi^2 \vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2} \exp\left(2\pi i \vec{m} \cdot \vec{R}_{j'}\right) \exp\left(-2\pi i \vec{m} \cdot \vec{R}_{i'}\right)}{\partial h_{\mu\nu}} p_{j'}^{\beta'}$$
$$+ \sum_i^N \vec{p}_i \cdot \frac{\partial \nabla_i \frac{1}{\pi V} \sum_{\vec{m}\neq 0} \left\{ \frac{\exp\left(-\frac{\pi^2 \vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2} \sum_j^N \left[\left(q_j + 2\pi i \vec{m} \cdot \vec{\mu}_j\right) \exp\left(2\pi i \vec{m} \cdot \vec{R}_j\right)\right] \exp\left(-2\pi i \vec{m} \cdot \vec{R}_i\right)\right\}}{\partial h_{\mu\nu}}, \tag{12}$$

$$\frac{\partial U_{dir}}{\partial h_{\mu\nu}} = \frac{\partial \sum_i^N \frac{1}{2}(q_i + \vec{\mu}_i \cdot \nabla_i) \sum_{j\neq i}^{\infty} (q_j + \vec{\mu}_j \cdot \nabla_j) \frac{\mathrm{erf}(\beta_{ij} R_{ij}) - \mathrm{erf}(\beta_0 R_{ij})}{R_{ij}}}{\partial h_{\mu\nu}} + \frac{1}{2} p_{i'}^{\alpha'} \frac{\partial \partial_{i'}^{\alpha'} \partial_{j'}^{\beta'} \frac{\mathrm{erf}(\beta_{i'j'} R_{i'j'}) - \mathrm{erf}(\beta_0 R_{i'j'})}{R_{i'j'}}}{\partial h_{\mu\nu}} p_{j'}^{\beta'}$$
$$+ \sum_i^N \vec{p}_i \cdot \frac{\partial \nabla_i \sum_{j\neq i}^{\infty} (q_j + \vec{\mu}_j \cdot \nabla_j) \frac{\mathrm{erf}(\beta_{ij} R_{ij}) - \mathrm{erf}(\beta_0 R_{ij})}{R_{ij}}}{\partial h_{\mu\nu}}, \tag{13}$$

$$\frac{\partial U_{self}}{\partial h_{\mu\nu}} = \frac{\partial \sum_i^N \frac{1}{2}\left(q_i + \vec{\mu}_i \cdot \nabla_i\right) \left[- \lim_{\vec{R}_j \to \vec{R}_i} \left(q_i + \vec{\mu}_i \cdot \nabla_j\right) \frac{\mathrm{erf}(\beta_0 R_{ij})}{R_{ij}} \right]}{\partial h_{\mu\nu}} + \frac{1}{2} p_{i'}^{\alpha'} \frac{\partial \partial_{i'}^{\alpha'} \partial_{j'}^{\beta'} \left[- \lim_{\vec{R}_j \to \vec{R}_i} \frac{\mathrm{erf}(\beta_0 R_{ij})}{R_{ij}} \right]}{\partial h_{\mu\nu}} p_{j'}^{\beta'} + \sum_i^N \vec{p}_i \cdot \frac{\partial \nabla_i \left[- \lim_{\vec{R}_j \to \vec{R}_i} \left(q_i + \vec{\mu}_i \cdot \nabla_j\right) \frac{\mathrm{erf}(\beta_0 R_{ij})}{R_{ij}} \right]}{\partial h_{\mu\nu}}. \tag{14}$$

After careful derivation (see Appendix A 2 for details), the following results can be obtained:

$$\frac{\partial U_{rec}}{\partial h_{\mu\nu}} = -\frac{1}{2\pi V} \left(h^{-1}\right)_{\mu\nu}^T \sum_{\vec{m}\neq 0} \left\{ \frac{\exp\left(-\frac{\pi^2 \vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2} \sum_j^N \left[q_j + 2\pi i \vec{m} \cdot \left(\vec{\mu}_j + \vec{p}_j\right)\right] \exp\left(2\pi i \vec{m} \cdot \vec{R}_j\right) \sum_i^N \left[q_i - 2\pi i \vec{m} \cdot \left(\vec{\mu}_i + \vec{p}_i\right)\right] \exp\left(-2\pi i \vec{m} \cdot \vec{R}_i\right) \right\}$$
$$+ \frac{1}{\pi V} \sum_{\vec{m}\neq 0} \left\{ \frac{\exp\left(-\frac{\pi^2 \vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2} \left(- \frac{1 + \frac{\pi^2 \vec{m}^2}{\beta_0^2}}{\vec{m}^2} \right) \left(-\vec{m}^\mu h_{\nu\gamma}^{-1} \vec{m}^\gamma\right) \sum_j^N \left[q_j + 2\pi i \vec{m} \cdot \left(\vec{\mu}_j + \vec{p}_j\right)\right] \exp\left(2\pi i \vec{m} \cdot \vec{R}_j\right) \sum_i^N \left[q_i - 2\pi i \vec{m} \cdot \left(\vec{\mu}_i + \vec{p}_i\right)\right] \right.$$
$$\left. \times \exp\left(-2\pi i \vec{m} \cdot \vec{R}_i\right) \right\} + \frac{1}{\pi V} \sum_{\vec{m}\neq 0} \left\{ \frac{\exp\left(-\frac{\pi^2 \vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2} \sum_j^N \left[q_j + 2\pi i \vec{m} \cdot \left(\vec{\mu}_j + \vec{p}_j\right)\right] \exp\left(2\pi i \vec{m} \cdot \vec{R}_j\right) \sum_i^N \sum_k^n \left[\left(-2\pi i \vec{m} \cdot \vec{R}_{ik} \mu_{ik}\right) \left(- \frac{\vec{R}_{ik}^\mu \vec{S}_{ik}^\nu}{\left|\vec{R}_{ik}\right|^3} \right) \right] \right.$$
$$\left. \times \exp\left(-2\pi i \vec{m} \cdot \vec{R}_i\right) \right\} + \frac{1}{\pi V} \sum_{\vec{m}\neq 0} \left\{ \frac{\exp\left(-\frac{\pi^2 \vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2} \sum_j^N \left[q_j + 2\pi i \vec{m} \cdot \left(\vec{\mu}_j + \vec{p}_j\right)\right] \exp\left(2\pi i \vec{m} \cdot \vec{R}_j\right) \sum_i^N \left[(-2\pi i)\left(-\vec{m}^\mu h_{\nu\gamma}^{-1} p_i^\gamma\right)\right] \exp\left(-2\pi i \vec{m} \cdot \vec{R}_i\right) \right\}$$
$$= -\frac{1}{2} \left(h^{-1}\right)_{\mu\nu}^T \cdot \sum_i^N \left[q_i (\phi_{rec})_i - \left(\vec{\mu}_i + \vec{p}_i\right) \cdot \left(\vec{E}_{rec}\right)_i \right] + \frac{1}{\pi V} \sum_{\vec{m}\neq 0} \left\{ \frac{\exp\left(-\frac{\pi^2 \vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2} \left(- \frac{1 + \frac{\pi^2 \vec{m}^2}{\beta_0^2}}{\vec{m}^2} \right) \left(-\vec{m}^\mu h_{\nu\gamma}^{-1} \vec{m}^\gamma\right) S(\vec{m}) S(-\vec{m}) \right\}$$
$$+ \sum_i^N \sum_k^n \left[\left(-(\vec{E}_{rec})_i \cdot \vec{R}_{ik} \mu_{ik}\right) \left(- \frac{\vec{R}_{ik}^\mu \vec{S}_{ik}^\nu}{\left|\vec{R}_{ik}\right|^3} \right) \right] + \sum_i^N (\vec{E}_{rec})_i^\mu h_{\nu\gamma}^{-1} p_i^\gamma, \tag{15}$$

where $S(\vec{m})$ is the structure factor, defined as $\sum_j^N \left[q_j + 2\pi i \vec{m} \cdot \left(\vec{\mu}_j + \vec{p}_j\right)\right] \exp\left(2\pi i \vec{m} \cdot \vec{R}_j\right)$, and

$$\frac{\partial U_{dir}}{\partial h_{\mu\nu}} = -\sum_i^N \sum_k^n u_{ik} \left[\frac{\left(\vec{E}_{dir}\right)_i^\mu \vec{S}_{ik}^\nu}{\left|\vec{R}_{ik}\right|} - \left(\vec{E}_{dir}\right)_i \cdot \vec{R}_{ik} \frac{\vec{R}_{ik}^\mu \vec{S}_{ik}^\nu}{\left|\vec{R}_{ik}\right|^3} \right] - \frac{1}{2} \sum_i^N \sum_{j\neq i}^\infty \left\{ q_i \left(\vec{E}_{dir}\right)_{j\to i}^\mu (\vec{S}_i^\nu - \vec{S}_j^\nu) + \left[\left(\vec{u}_i + \vec{p}_i\right) \cdot \left(\vec{E}_{dir}\right)_{j\to i} \right]^\mu (\vec{S}_i^\nu - \vec{S}_j^\nu) \right\}, \tag{16}$$

$$\frac{\partial U_{self}}{\partial h_{\mu\nu}} = -\sum_i^N \sum_k^n u_{ik} \left[\frac{\left(\vec{E}_{self}\right)_i^\mu \vec{S}_{ik}^\nu}{\left|\vec{R}_{ik}\right|} - \left(\vec{E}_{self}\right)_i \cdot \vec{R}_{ik} \frac{\vec{R}_{ik}^\mu \vec{S}_{ik}^\nu}{\left|\vec{R}_{ik}\right|^3} \right]. \tag{17}$$

Here $\vec{E}_{rec}$, $\vec{E}_{dir}$, and $\vec{E}_{self}$ refer to the electrostatic fields resulting from the reciprocal part, direct part, or self-part of an Ewald summation, respectively. The notation for ϕ and $\vec{E}$ is similar. Furthermore, the following expressions for electric field and its derivatives are used:

$$\left(\vec{E}_{dir}\right)_{j\to i} = -(q_j + \vec{\mu}_j \cdot \nabla_j) \nabla_i \frac{\mathrm{erf}(\beta_{ij} R_{ij}) - \mathrm{erf}(\beta_0 R_{ij})}{R_{ij}}, \qquad \left(\vec{E}_{dir}\right)_{j\to i} = -(q_j + \vec{\mu}_j \cdot \nabla_j) \nabla_i \nabla_i \frac{\mathrm{erf}(\beta_{ij} R_{ij}) - \mathrm{erf}(\beta_0 R_{ij})}{R_{ij}}. \tag{18}$$

Adding up all the three terms above and removing the common factor $-\left(h^{-1}\right)^T$, the final expression of the electrostatic energy contribution to the virial is as follows:

$$V_{\mu\nu} = \frac{1}{2}\delta_{\mu\nu} \cdot \sum_i^N \left[q_i(\phi_{rec})_i - \left(\vec{\mu}_i + \vec{p}_i\right) \cdot \left(\vec{E}_{rec}\right)_i \right] - \frac{1}{\pi V} \sum_{\vec{m}\neq 0} \left\{ \frac{\exp\left(-\frac{\pi^2 \vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2} \frac{1 + \frac{\pi^2 \vec{m}^2}{\beta_0^2}}{\vec{m}^2} \vec{m}^\mu \vec{m}^\nu S\left(\vec{m}\right) S\left(-\vec{m}\right) \right\}$$
$$- \sum_i^N \left(\vec{E}_{rec}\right)_i^\mu p_i^\nu + \sum_i^N \sum_k^n u_{ik} \left[\frac{\left(\vec{E}_{dir} + \vec{E}_{self}\right)_i^\mu \vec{R}_{ik}^\nu}{\left|\vec{R}_{ik}\right|} - \left(\vec{E}_{rec} + \vec{E}_{dir} + \vec{E}_{self}\right)_i \cdot \vec{R}_{ik} \frac{\vec{R}_{ik}^\mu \vec{R}_{ik}^\nu}{\left|\vec{R}_{ik}\right|^3} \right]$$
$$+ \frac{1}{2} \sum_i^N \sum_{j\neq i}^\infty \left\{ q_i \left(\vec{E}_{dir}\right)_{j\to i}^\mu (\vec{R}_i^\nu - \vec{R}_j^\nu) + \left[\left(\vec{u}_i + \vec{p}_i\right) \cdot \left(\vec{E}_{dir}\right)_{j\to i} \right]^\mu (\vec{R}_i^\nu - \vec{R}_j^\nu) \right\}. \tag{19}$$

D. Stress tensor of rigid molecules

Molecular dynamics simulations of bio-molecular systems are often conducted under certain constraints. The most common are distance constraints,

$$\left(\vec{R}_i - \vec{R}_j\right)^2 - d_{ij}^2 = 0. \tag{20}$$

When molecules are under this kind of constraints, a molecule-based scaling in NPT simulations is often used to vary the box dimensions instead of atom-based scaling. In doing so, the atomic coordinate can be expressed as

$$\vec{R}_i = \vec{R}_{i0} + \vec{d}_i = s_{i0,1} \vec{u}_1 + s_{i0,2} \vec{u}_2 + s_{i0,3} \vec{u}_3 + \vec{d}_i, \tag{21}$$

where $\vec{R}_{i0}$ is the molecular center to which atom i belongs (usually the mass center of the molecule), $\vec{s}_{i0}$'s are the fractional coordinates of the center, and $\vec{d}_i$ is the relative displacement of atom i from the center. Obviously, $\vec{d}_i$ is a constant vector during the box dimension scaling.

Thus, the virial tensor of a pGM system under this rigid body condition with distance constraints is

$$V_{\mu\nu} = \frac{1}{2}\delta_{\mu\nu} \cdot \sum_i^N \left[q_i(\phi_{rec})_i - \left(\vec{\mu}_i + \vec{p}_i\right) \cdot \left(\vec{E}_{rec}\right)_i \right] - \frac{1}{\pi V} \sum_{\vec{m}\neq 0} \left\{ \frac{\exp\left(-\frac{\pi^2 \vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2} \frac{1 + \frac{\pi^2 \vec{m}^2}{\beta_0^2}}{\vec{m}^2} \vec{m}^\mu \vec{m}^\nu S\left(\vec{m}\right) S\left(-\vec{m}\right) \right\}$$
$$- \sum_i^N \left(\vec{E}_{rec}\right)_i^\mu \left(\vec{\mu}_i^\nu + \vec{p}_i^\nu + q_i \vec{d}_i^\nu\right) - \sum_i^N \left(\vec{\mu}_i^\gamma + \vec{p}_i^\gamma\right) \left(\vec{E}_{rec}\right)_i^{\gamma\mu} \vec{d}_i^\nu + \frac{1}{2} \sum_i^N \sum_{j\neq i}^\infty \left\{ q_i \left(\vec{E}_{dir}\right)_{j\to i}^\mu (\vec{R}_{i0}^\nu - \vec{R}_{j0}^\nu) \right.$$
$$\left. + \left[\left(\vec{u}_i + \vec{p}_i\right) \cdot \left(\vec{E}_{dir}\right)_{j\to i} \right]^\mu (\vec{R}_{i0}^\nu - \vec{R}_{j0}^\nu) \right\}. \tag{22}$$

Here, index γ follows the Einstein's summation notation. The derivation of the above expression is similar to what was described before for flexible molecules, so it is omitted here. The detail can be found in Appendix A 3.

E. Virial correction of screened system

In the current pGM setup, the electrostatic interactions are not screened, meaning all 1–2, 1–3, and 1–4 electrostatic interactions are considered at their full strength. This choice is based on the consideration of polarizability anisotropy.[21] However, if electrostatic interactions between certain atoms are screened, the electrostatic virial correction can be defined as follows:

(1) For flexible molecules,

$$(V_{\mu\nu})_{correction} = \sum_i^N \sum_k^n u_{ik} \left[\frac{\left(-\vec{E}_{screen}\right)_i^{\mu} \vec{R}_{ik}^{\nu}}{\left|\vec{R}_{ik}\right|} - \left(-\vec{E}_{screen}\right)_i \cdot \vec{R}_{ik} \frac{\vec{R}_{ik}^{\mu} \vec{R}_{ik}^{\nu}}{\left|\vec{R}_{ik}\right|^3} \right] + \frac{1}{2} \sum_i^N \sum_{j\neq i}^{*(i)} \left\{ q_i \left(-\vec{E}_{screen}\right)_{j\to i}^{\mu} (\vec{R}_i^{\nu} - \vec{R}_j^{\nu}) + \left[\left(\vec{u}_i + \vec{p}_i\right) \cdot \left(-\vec{E}_{screen}\right)_{j\to i} \right]^{\mu} (\vec{R}_i^{\nu} - \vec{R}_j^{\nu}) \right\}. \quad (23)$$

(2) For rigid molecules,

$$(V_{\mu\nu})_{correction} = \frac{1}{2} \sum_i^N \sum_{j\neq i}^{*(i)} \left\{ q_i \left(-\vec{E}_{screen}\right)_{j\to i}^{\mu} (\vec{R}_{i0}^{\nu} - \vec{R}_{j0}^{\nu}) + \left[\left(\vec{u}_i + \vec{p}_i\right) \cdot \left(-\vec{E}_{screen}\right)_{j\to i} \right]^{\mu} (\vec{R}_{i0}^{\nu} - \vec{R}_{j0}^{\nu}) \right\}. \quad (24)$$

Here, $\vec{E}_{screen}$ and $\vec{E}_{screen}$ are the electric field and its derivative generated by the screened interactions, and $*(i)$ represents the screened atom pairs of atom i.

Derivation of the above equations is given in Appendix A 4, which is similar to what was presented in our previous development.[1] However, it is important to point out that the screening has to be consistent throughout the model. In other words, we must screen the same interactions when computing both induced dipoles and forces. This is to ensure that the energy and corresponding forces are consistent under the same screening scheme.

III. RESULTS AND DISCUSSION

A. Finite difference validation

As we presented above, Eqs. (19) and (22) are the electrostatic virial expressions of flexible and rigid pGM molecular systems, respectively. To confirm the theoretical derivation, we performed a finite difference test, starting from the virial definition of Eq. (4), to assess whether the derivation is correct. Here, a cubic box with 512 water molecules, as in our previous study, was used,[1,21] and the dimension of the box is 33 Å. To guarantee high-precision energy calculation, the PME setup uses the following parameters: coefficient β_0 = 0.3 Å^{-1}, B-spline interpolation order = 9, FFT grid spacing = 0.33 Å, and direct space cutoff = 14.4 Å. The induced dipole convergence was also set to a very tight criterion of 10^{-12}. The setup leads to an energy accuracy level of ~10^{-12}, so that the finite difference test can be carried successfully.

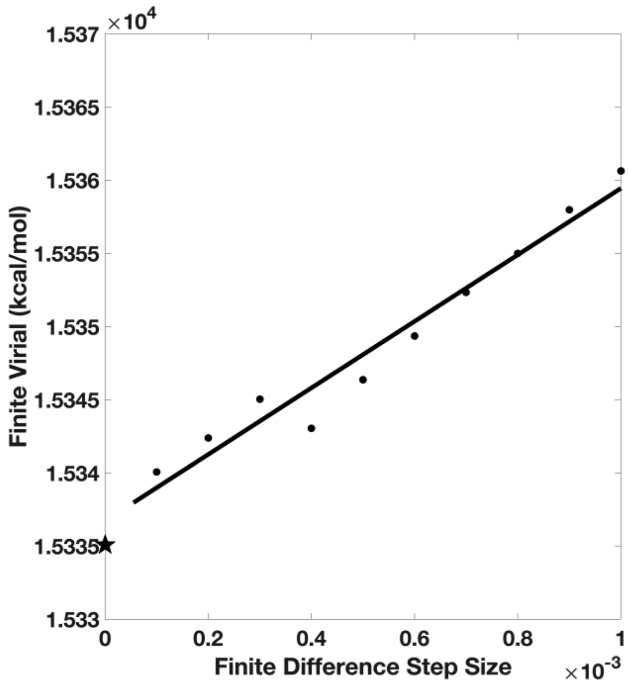
It is clear from Fig. 1 that both finite-difference virial values approach their respective analytical values in roughly linear fashions as the finite difference step sizes decrease, demonstrating that the numerical values converge to the analytical values. This finding confirms that the analytical virial expressions Eqs. (19) and (22) are correct when used to compute virials of the tested molecular system. Note also that the finite-difference steps in the rigid body system [Fig. 1(b)] are 10-time larger than the ones used in the flexible system [Fig. 1(a)]. This is because the virials of the rigid body system are much smaller (~1/50) than those of the flexible system due to large intramolecular contributions that are sensitive to the distance variations in the finite difference calculations. Hence, rather small finite difference step sizes were needed. Such large intramolecular contributions were absent in the rigid-body system and a somewhat larger step sizes were used for numerical accuracy. The linear convergence trends illustrated in both the flexible and rigid body systems confirm the correctness of Eq. (22).

B. NPT simulation of water box

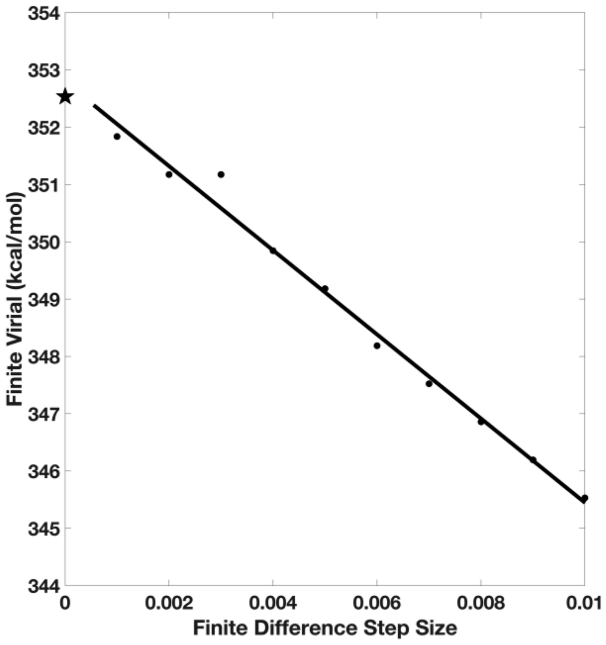
To illustrate that the analytical formulation can achieve stable NPT simulations, we performed a test run for 1 ns of the water box containing 512 pGM water molecules.[1,21] The rigid-body formula, Eq. (22), was used in this test, and the water models were simply created from the standard TIP3P geometry by adding permanent and induced dipole moments and by setting the van der Waals parameters to be 9% of those in the standard TIP3P waters, to give a relatively reasonable density. We note that this pGM water model is yet to be optimized to reproduce any physical water properties. Here, a setup of lower energy accuracy was used for reasonable simulation throughput, with the Ewald coefficient β_0 = 0.4 Å^{-1}, B-spline interpolation order = 8, FFT grid spacing = 0.5 Å, and direct space cutoff = 9 Å. The induced dipole convergence criterion was also raised to 10^{-6}. The simulation time step was chosen to be 1 fs. The temperature was set to 300 K and the pressure 1.0 bar. The Berendsen thermostat and barostat were used and all other dynamics simulation parameters were set as default from the Amber SANDER program. The job was run on a single core of an INTEL Xeon E5-4620 central processing unit (CPU) with a wall-clock time of 6.9 days.

Figure 2 shows that the pGM/PME runs well under the NPT condition. Both density and pressure fluctuate around the equilibrium values, and the fluctuation ranges were similar to those observed for the classical point-charge models. The simulation was stable, no "polarization catastrophe" event was observed even if all interactions, including 1–2 and 1–3 interactions, were included.

It is interesting to point out that the electrostatic component of the pGM water model is much smaller (~1/10) than the TIP3P water model, even though the atomic charges are usually higher in polarizable models. This in part explains why we need to reduce the van der Waals parameters to 9% of the original TIP3P values, to match the much smaller electrostatic virial component so that the system would eventually equilibrate to a reasonable state. Indeed, the

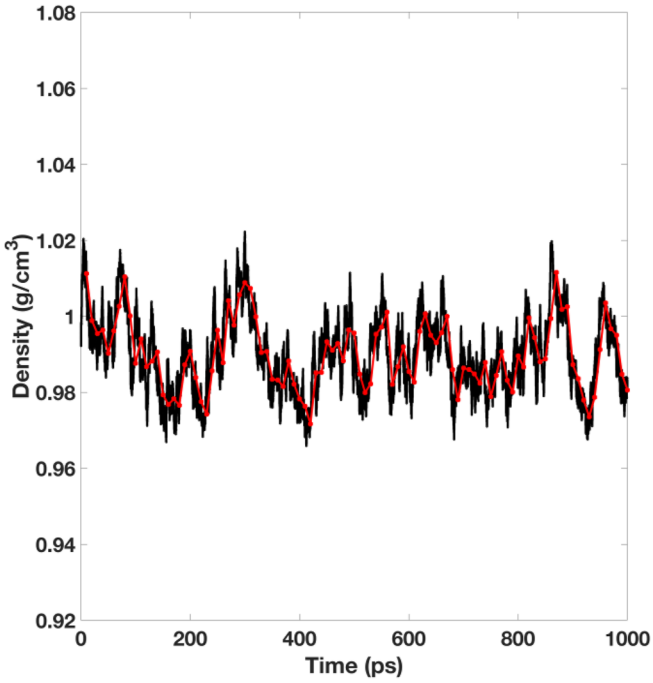


(a)

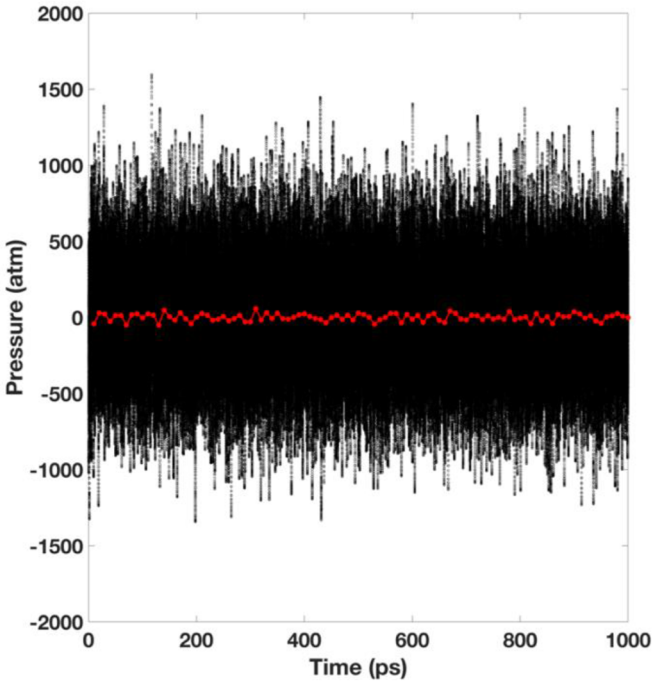


(b)

FIG. 1. The finite-difference validation of the analytical virial expressions in the *xx* direction. (a) is for the flexible case and (b) is for the rigid body case. *x* axis refers to the proportional changes of simulation box dimension in the *x* direction. The star on the *y* axis is the analytical virial value given by Eqs. (19) and (22), respectively. The finite-difference values were obtained using Eq. (4). Testing along the *yy* and *zz* directions are available in the supplementary material. An additional set of testing was also conducted for a 4096-water box, available in the supplementary material.



(a)



(b)

FIG. 2. Time evolutions of density and pressure of the NPT simulation for the 512-water box. The red lines are the running averages of every 10 ps. The time evolutions of total energy and temperature are available in the supplementary material. The time evolutions of density, total energy, and temperature of the Monte Carlo pressure-regulated simulation are also available in the supplementary material.

average density and pressure values are 0.99 g/cm^3 and 1.0 bar during the 1 ns run, respectively. As a supporting evidence for the correct NPT implementation for the pGM model, we also implemented a Monte Carlo pressure regulation scheme for the pGM model, as no virial (pressure) calculation is needed in this regulation scheme.[30] The simulation followed the exact same setup and was found to stabilize at the same state, with the averaged density values of 0.99 g/cm^3. The related plot of time evolutions of density, total energy, and temperature of Monte Carlo pressure regulation are available in the supplementary material.

The much-reduced electrostatic virial component is worth noting but a reasonable behavior. The induced dipoles in the pGM model are actively playing the role of electrostatic screening. As a result, the electrostatic interactions between water molecules are largely reduced, leading to much lower electrostatic virial component. This phenomenon shows that the polarizable pGM model is fundamentally different from classical point-charge models. Thus, its parameterization shall require extensive development and validation. Nevertheless, we reiterate that the goal of this development is not to develop a pGM water model, but to establish a constant pressure simulation protocol for polarizable methods, such as the pGM model.

Another interesting observation upon inspecting Fig. 2 is that the pGM water model exhibits a longer correlation time in density than the TIP3P water model, with pGM water's density correlation time as long as 200–300 ps, rather than tens of picoseconds as is usually observed for point-charge water models. This effect could either result from the imperfect calibration of the pGM water tested here, the use of a barostat coupling constant that is too weak for the pGM model, or the inclusion of the explicit polarization in the pGM model, which requires longer time to equilibrate. Obviously, the point charge models also have certain level of polarization effect from reorientation of the molecules, but they certainly do not acquire any electronic polarization that is explicitly modeled with the induced dipoles. All these phenomena strongly suggest that the parameterization of pGM models requires extra care and extensive validations.

IV. CONCLUSION

In this work, we presented the derivation and implementation of an analytical method for achieving constant pressure simulation for the pGM model. Specifically, we derived the stress tensor expression for both the flexible and rigid body molecular system in the pGM electrostatic model, the trace of which gives the expression for internal stress. In addition, we showed how to correct the stress tensor if a short-range screening strategy is used in the force field development. Once the internal stress is obtained, by balancing it with the externally applied pressure through a certain type of barostat, a constant pressure simulation is readily realized.

Since the formulation of the stress tensor in pGM model is rather complex, we first performed a finite difference test to validate its correctness. We showed that for both the flexible and rigid-body systems, the finite-difference virial approaches the analytical values when decreasing the differentiation step size, confirming the accuracy of the derived expressions.

The tensor expression was implemented into the Amber/SANDER program and tested with a small box of 512 pGM water molecules. In the calculations, the revised TIP3P water geometry was applied, and the TIP3P van der Waals parameters were scaled down to 9% of the original values after imposing the pGM electrostatic interactions. After equilibration, the system was found to reach a reasonable density of 0.99 g/cm^3 and pressure of 1.0 bar under the room temperature during the production run of 1 ns. A constant pressure simulation of the system with Monte Carlo pressure regulation scheme stabilized at the same state. This shows that the analytical stress tensor formula and its implementation are successful.

Several interesting differences were observed between the pGM water model and the TIP3P water model from which the pGM model was currently derived. Notably, the electrostatic virial was much smaller, requiring much smaller van der Waals parameters to reach reasonable water density. The density correlation time was also much longer. The observation strongly suggests that the pGM model, as a polarizable model, is fundamentally different compared to classic point charge models. Thus, the parameterization of the pGM model requires extra care and extensive validation, which is a major focus of the next step of our research.

SUPPLEMENTARY MATERIAL

The supplementary material includes finite-difference tests along the *yy* and *zz* directions of the 512-water box (Fig. S1), finite-difference tests along the *xx* directions of a 4096-water box (Fig. S2), time evolutions of total energy and temperature of the NPT simulation for the 512-water box (Fig. S3), time evolutions of density, total energy, and temperature of the Monte Carlo pressure-regulated NPT simulation (Fig. S4).

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health/NIGMS (Grant No. GM130367).

AUTHOR DECLARATIONS

Conflict of Interest

We declare there is no conflict of interest.

DATA AVAILABILITY

The algorithms developed in this study and the validation data are deposited in the Amber repository and will be made publicly available in the next Amber/AmberTools release at http://ambermd.org/.

APPENDIX: MATHEMATICAL DETAILS

1. Several useful derivatives against *h*

$$\frac{\partial \vec{R}^{\alpha}}{\partial h_{\mu\nu}} = \delta_{\alpha\mu}\vec{S}^{\nu},$$

$$\frac{\partial V}{\partial h_{\mu\nu}} = V\left(h^{-1}\right)^{T}_{\mu\nu},$$

$$\frac{\partial \left(h^{-1}\right)_{\alpha\beta}}{\partial h_{\mu\nu}} = -\left(h^{-1}\right)_{\alpha\mu}\left(h^{-1}\right)_{\nu\beta},$$

$$\frac{\partial \vec{m}^{\alpha}}{\partial h_{\mu\nu}} = -\vec{m}^{\mu}\left(h^{-1}\right)_{\nu\alpha}.$$

2. Virial derivation

From previous discussions in the main text,

$$\frac{\partial U}{\partial h_{\mu\nu}} = \frac{\partial U_{rec}}{\partial h_{\mu\nu}} + \frac{\partial U_{dir}}{\partial h_{\mu\nu}} + \frac{\partial U_{self}}{\partial h_{\mu\nu}}.$$

Since for the direct part and the self-part the only variable that depends on h is the atom coordinates $\vec{R}$, we have the following:

$$\begin{aligned}
\frac{\partial U_{dir}}{\partial h_{\mu\nu}} &= \sum_{i}^{N} \frac{\partial \vec{\mu}_i}{\partial h_{\mu\nu}} \cdot \nabla_i \sum_{j\neq i}^{\infty} \left(q_j + \vec{\mu}_j \cdot \nabla_j\right) \frac{\operatorname{erf}(\beta_{ij}R_{ij}) - \operatorname{erf}(\beta_0 R_{ij})}{R_{ij}} + \sum_{i}^{N} \vec{p}_i \cdot \nabla_i \sum_{j\neq i}^{\infty} \frac{\partial \vec{\mu}_j}{\partial h_{\mu\nu}} \cdot \nabla_j \frac{\operatorname{erf}(\beta_{ij}R_{ij}) - \operatorname{erf}(\beta_0 R_{ij})}{R_{ij}} \\
&+ \frac{1}{2}\sum_{i}^{N}\left(q_i + \vec{\mu}_i \cdot \nabla_i\right)\sum_{j\neq i}^{\infty}\left(q_j + \vec{\mu}_j \cdot \nabla_j\right)\frac{\partial \frac{\operatorname{erf}(\beta_{ij}R_{ij}) - \operatorname{erf}(\beta_0 R_{ij})}{R_{ij}}}{\partial h_{\mu\nu}} + \frac{1}{2}\sum_{i}^{N}\vec{p}_i \cdot \nabla_i \sum_{j\neq i}^{\infty}\vec{p}_j \cdot \nabla_j \frac{\partial \frac{\operatorname{erf}(\beta_{ij}R_{ij}) - \operatorname{erf}(\beta_0 R_{ij})}{R_{ij}}}{\partial h_{\mu\nu}} \\
&+ \sum_{i}^{N}\vec{p}_i \cdot \nabla_i \sum_{j\neq i}^{\infty}\left(q_j + \vec{\mu}_j \cdot \nabla_j\right)\frac{\partial \frac{\operatorname{erf}(\beta_{ij}R_{ij}) - \operatorname{erf}(\beta_0 R_{ij})}{R_{ij}}}{\partial h_{\mu\nu}} = \sum_{i}^{N}\frac{\partial \vec{\mu}_i}{\partial h_{\mu\nu}} \cdot \nabla_i \sum_{j\neq i}^{\infty}\left[q_j + \left(\vec{\mu}_j + \vec{p}_j\right)\cdot \nabla_j\right] \\
&\times \frac{\operatorname{erf}(\beta_{ij}R_{ij}) - \operatorname{erf}(\beta_0 R_{ij})}{R_{ij}} + \frac{1}{2}\sum_{i}^{N}\left[q_i + \left(\vec{\mu}_i + \vec{p}_i\right)\cdot \nabla_i\right]\sum_{j\neq i}^{\infty}\left[q_j + \left(\vec{\mu}_j + \vec{p}_j\right)\cdot \nabla_j\right]\frac{\partial \frac{\operatorname{erf}(\beta_{ij}R_{ij}) - \operatorname{erf}(\beta_0 R_{ij})}{R_{ij}}}{\partial h_{\mu\nu}} \\
&= -\sum_{i}^{N}\frac{\partial \vec{\mu}_i}{\partial h_{\mu\nu}} \cdot \vec{E}_{dir} + \frac{1}{2}\sum_{i}^{N}\left[q_i + \left(\vec{\mu}_i + \vec{p}_i\right)\cdot \nabla_i\right]\sum_{j\neq i}^{\infty}\left[q_j + \left(\vec{\mu}_j + \vec{p}_j\right)\cdot \nabla_j\right]\frac{\partial \frac{\operatorname{erf}(\beta_{ij}R_{ij}) - \operatorname{erf}(\beta_0 R_{ij})}{R_{ij}}}{\partial \vec{R}_i} \cdot \left(\frac{\partial \vec{R}_i}{\partial h_{\mu\nu}} - \frac{\partial \vec{R}_j}{\partial h_{\mu\nu}}\right) \\
&= -\sum_{i}^{N}\sum_{k}^{n} u_{ik}\left[\frac{\left(\vec{E}_{dir}\right)^{\mu}_{i}\vec{S}^{\nu}_{ik}}{\left|\vec{R}_{ik}\right|} - \left(\vec{E}_{dir}\right)_i \cdot \vec{R}_{ik}\frac{\vec{R}^{\mu}_{ik}\vec{S}^{\nu}_{ik}}{\left|\vec{R}_{ik}\right|^3}\right] \\
&- \frac{1}{2}\sum_{i}^{N}\sum_{j\neq i}^{\infty}\left\{q_i\left(\vec{E}_{dir}\right)^{\mu}_{j\to i}\left(\vec{S}^{\nu}_{i} - \vec{S}^{\nu}_{j}\right) + \left[\left(\vec{u}_i + \vec{p}_i\right)\cdot\left(\vec{E}_{dir}\right)_{j\to i}\right]^{\mu}\left(\vec{S}^{\nu}_{i} - \vec{S}^{\nu}_{j}\right)\right\}.
\end{aligned}$$

Similarly, for the self-part,

$$\begin{aligned}
\frac{\partial U_{self}}{\partial h_{\mu\nu}} &= -\sum_{i}^{N}\sum_{k}^{n} u_{ik}\left[\frac{\left(\vec{E}_{self}\right)^{\mu}_{i}\vec{S}^{\nu}_{ik}}{\left|\vec{R}_{ik}\right|} - \left(\vec{E}_{self}\right)_i \cdot \vec{R}_{ik}\frac{\vec{R}^{\mu}_{ik}\vec{S}^{\nu}_{ik}}{\left|\vec{R}_{ik}\right|^3}\right] - \lim_{\vec{R}_j \to \vec{R}_i}\frac{1}{2}\sum_{i}^{N}\sum_{j\neq i}^{\infty}\left\{q_i\left(\vec{E}_{self}\right)^{\mu}_{j\to i}\left(\vec{S}^{\nu}_{i} - \vec{S}^{\nu}_{j}\right) + \left[\left(\vec{u}_i + \vec{p}_i\right)\cdot\left(\vec{E}_{self}\right)_{j\to i}\right]^{\mu}\left(\vec{S}^{\nu}_{i} - \vec{S}^{\nu}_{j}\right)\right\} \\
&= -\sum_{i}^{N}\sum_{k}^{n} u_{ik}\left[\frac{\left(\vec{E}_{self}\right)^{\mu}_{i}\vec{S}^{\nu}_{ik}}{\left|\vec{R}_{ik}\right|} - \left(\vec{E}_{self}\right)_i \cdot \vec{R}_{ik}\frac{\vec{R}^{\mu}_{ik}\vec{S}^{\nu}_{ik}}{\left|\vec{R}_{ik}\right|^3}\right].
\end{aligned}$$

However, the above approach cannot be applied for the reciprocal part of the virial. Thus, we must resort to a brute force calculation,

$$\frac{\partial U_{rec}}{\partial h_{\mu\nu}} = (1) + (2) + (3),$$

$$(1) = \frac{\partial \sum_i^N \frac{1}{2}\left(q_i + \vec{\mu}_i \cdot \nabla_i\right) \frac{1}{\pi V} \sum_{\vec{m}\neq 0} \left\{ \frac{\exp\left(-\frac{\pi^2 \vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2} \sum_j^N \left[\left(q_j + 2\pi i \vec{m} \cdot \vec{\mu}_j\right) \exp\left(2\pi i \vec{m} \cdot \vec{R}_j\right)\right] \exp\left(-2\pi i \vec{m} \cdot \vec{R}_i\right)\right\}}{\partial h_{\mu\nu}},$$

$$(2) = \frac{1}{2} p_{i'}^{\alpha'} \frac{\partial \partial_{i'}^{\alpha'} \partial_{j'}^{\beta'} \frac{1}{\pi V} \sum_{\vec{m}\neq 0} \frac{\exp\left(-\frac{\pi^2 \vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2} \exp\left(2\pi i \vec{m} \cdot \vec{R}_{j'}\right) \exp\left(-2\pi i \vec{m} \cdot \vec{R}_{i'}\right)}{\partial h_{\mu\nu}} p_{j'}^{\beta'},$$

$$(3) = \sum_i^N \vec{p}_i \cdot \frac{\partial \nabla_i \frac{1}{\pi V} \sum_{\vec{m}\neq 0} \left\{ \frac{\exp\left(-\frac{\pi^2 \vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2} \sum_j^N \left[\left(q_j + 2\pi i \vec{m} \cdot \vec{\mu}_j\right) \exp\left(2\pi i \vec{m} \cdot \vec{R}_j\right)\right] \exp\left(-2\pi i \vec{m} \cdot \vec{R}_i\right)\right\}}{\partial h_{\mu\nu}}.$$

For each part, we derive the following expressions:

$$\begin{aligned}(1) = &-\frac{1}{2\pi V}\left(h^{-1}\right)_{\mu\nu}^T \sum_{\vec{m}\neq 0} \left\{ \frac{\exp\left(-\frac{\pi^2 \vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2} \sum_j^N \left(q_j + 2\pi i \vec{m} \cdot \vec{\mu}_j\right) \exp\left(2\pi i \vec{m} \cdot \vec{R}_j\right) \sum_i^N \left(q_i - 2\pi i \vec{m} \cdot \vec{\mu}_i\right) \exp\left(-2\pi i \vec{m} \cdot \vec{R}_i\right)\right\} \\ &+ \frac{1}{\pi V} \sum_{\vec{m}\neq 0} \left\{ \frac{\exp\left(-\frac{\pi^2 \vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2} \left(-\frac{1 + \frac{\pi^2 \vec{m}^2}{\beta_0^2}}{\vec{m}^2}\right) \left(-\vec{m}^\mu h_{\nu\gamma}^{-1} \vec{m}^\gamma\right) \sum_j^N \left(q_j + 2\pi i \vec{m} \cdot \vec{\mu}_j\right) \exp\left(2\pi i \vec{m} \cdot \vec{R}_j\right) \right. \\ &\left. \times \sum_i^N \left(q_i - 2\pi i \vec{m} \cdot \vec{\mu}_i\right) \exp\left(-2\pi i \vec{m} \cdot \vec{R}_i\right)\right\} + \frac{1}{\pi V} \sum_{\vec{m}\neq 0} \left\{ \frac{\exp\left(-\frac{\pi^2 \vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2} \sum_j^N \left(q_j + 2\pi i \vec{m} \cdot \vec{\mu}_j\right) \exp\left(2\pi i \vec{m} \cdot \vec{R}_j\right) \right. \\ &\left. \times \sum_i^N \sum_k^n \left[\left(-2\pi i \vec{m} \cdot \vec{R}_{ik} \mu_{ik}\right) \left(-\frac{\vec{R}_{ik}^\mu \vec{S}_{ik}^\nu}{\left|\vec{R}_{ik}\right|^3}\right)\right] \exp\left(-2\pi i \vec{m} \cdot \vec{R}_i\right)\right\},\end{aligned}$$

$$\begin{aligned}(2) = &-\frac{1}{2\pi V}\left(h^{-1}\right)_{\mu\nu}^T \sum_{\vec{m}\neq 0} \left\{ \frac{\exp\left(-\frac{\pi^2 \vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2} \sum_j^N \left(2\pi i \vec{m} \cdot \vec{p}_j\right) \exp\left(2\pi i \vec{m} \cdot \vec{R}_j\right) \sum_i^N \left(-2\pi i \vec{m} \cdot \vec{p}_i\right) \exp\left(-2\pi i \vec{m} \cdot \vec{R}_i\right)\right\} \\ &+ \frac{1}{\pi V} \sum_{\vec{m}\neq 0} \left\{ \frac{\exp\left(-\frac{\pi^2 \vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2} \left(-\frac{1 + \frac{\pi^2 \vec{m}^2}{\beta_0^2}}{\vec{m}^2}\right) \left(-\vec{m}^\mu h_{\nu\gamma}^{-1} \vec{m}^\gamma\right) \sum_j^N \left(2\pi i \vec{m} \cdot \vec{p}_j\right) \exp\left(2\pi i \vec{m} \cdot \vec{R}_j\right) \sum_i^N \left(-2\pi i \vec{m} \cdot \vec{p}_i\right) \right. \\ &\left. \times \exp\left(-2\pi i \vec{m} \cdot \vec{R}_i\right)\right\} + \frac{1}{\pi V} \sum_{\vec{m}\neq 0} \left\{ \frac{\exp\left(-\frac{\pi^2 \vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2} \sum_j^N \left(2\pi i \vec{m} \cdot \vec{p}_j\right) \exp\left(2\pi i \vec{m} \cdot \vec{R}_j\right) \right. \\ &\left. \times \sum_i^N \left[(-2\pi i)\left(-\vec{m}^\mu h_{\nu\gamma}^{-1} p_i^\gamma\right)\right] \exp\left(-2\pi i \vec{m} \cdot \vec{R}_i\right)\right\},\end{aligned}$$

$$(3) = -\frac{1}{\pi V}\left(h^{-1}\right)^{T}_{\mu\nu}\sum_{\vec{m}\neq 0}\left\{\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\sum_{j}^{N}\left(q_j + 2\pi i\vec{m}\cdot\vec{\mu}_j\right)\exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\sum_{i}^{N}\left(-2\pi i\vec{m}\cdot\vec{p}_i\right)\exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right)\right\}$$

$$+\frac{2}{\pi V}\sum_{\vec{m}\neq 0}\left\{\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\left(-\frac{1+\frac{\pi^2\vec{m}^2}{\beta_0^2}}{\vec{m}^2}\right)\left(-\vec{m}^{\mu}h^{-1}_{\nu\gamma}\vec{m}^{\gamma}\right)\sum_{j}^{N}\left(q_j + 2\pi i\vec{m}\cdot\vec{\mu}_j\right)\exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\sum_{i}^{N}\left(-2\pi i\vec{m}\cdot\vec{p}_i\right)\right.$$

$$\left.\times\ \exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right)\right\} + \frac{1}{\pi V}\sum_{\vec{m}\neq 0}\left\{\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\sum_{j}^{N}\left(2\pi i\vec{m}\cdot\vec{p}_j\right)\exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\right.$$

$$\left.\times\sum_{i}^{N}\sum_{k}^{n}\left[\left(-2\pi i\vec{m}\cdot\vec{R}_{ik}\mu_{ik}\right)\left(-\frac{\vec{R}^{\mu}_{ik}\vec{S}^{\nu}_{ik}}{\left|\vec{R}_{ik}\right|^3}\right)\right]\exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right)\right\}$$

$$+\frac{1}{\pi V}\sum_{\vec{m}\neq 0}\left\{\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\sum_{j}^{N}\left(q_j + 2\pi i\vec{m}\cdot\vec{\mu}_j\right)\exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\sum_{i}^{N}\left[\left(-2\pi i\right)\left(-\vec{m}^{\mu}h^{-1}_{\nu\gamma}p^{\gamma}_i\right)\right]\exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right)\right\}.$$

By adding together expressions for (1), (2), and (3), we arrive at the final result for $\frac{\partial U_{rec}}{\partial h_{\mu\nu}}$, which is shown in Sec. II.

3. Virial of rigid body systems

As in the derivation presented in Appendix A 2, for the direct part and the self-part, the only variable that depends on h is the atom coordinate $\vec{R}$. However, a major difference is that, for the rigid molecules, the covalent dipole $\vec{u}$ is a constant vector now. Thus, the derivation complexity is greatly reduced,

$$\frac{\partial U_{dir}}{\partial h_{\mu\nu}} = -\frac{1}{2}\sum_{i}^{N}\sum_{j\neq i}^{\infty}\left\{q_i\left(\vec{E}_{dir}\right)^{\mu}_{j\to i}\left(\vec{S}^{\nu}_{i0} - \vec{S}^{\nu}_{j0}\right)\right.$$

$$\left.+\left[\left(\vec{u}_i + \vec{p}_i\right)\cdot\left(\vec{E}_{dir}\right)_{j\to i}\right]^{\mu}\left(\vec{S}^{\nu}_{i0} - \vec{S}^{\nu}_{j0}\right)\right\},$$

$$\frac{\partial U_{self}}{\partial h_{\mu\nu}} = 0.$$

As for the reciprocal part, we follow the procedure presented in Appendix A 2,

$$(1) = -\frac{1}{2\pi V}\left(h^{-1}\right)^{T}_{\mu\nu}\sum_{\vec{m}\neq 0}\left\{\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\sum_{j}^{N}\left(q_j + 2\pi i\vec{m}\cdot\vec{\mu}_j\right)\exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\sum_{i}^{N}\left(q_i - 2\pi i\vec{m}\cdot\vec{\mu}_i\right)\exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right)\right\}$$

$$+\frac{1}{\pi V}\sum_{\vec{m}\neq 0}\left\{\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\left(-\frac{1+\frac{\pi^2\vec{m}^2}{\beta_0^2}}{\vec{m}^2}\right)\left(-\vec{m}^{\mu}h^{-1}_{\nu\gamma}\vec{m}^{\gamma}\right)\sum_{j}^{N}\left(q_j + 2\pi i\vec{m}\cdot\vec{\mu}_j\right)\exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\sum_{i}^{N}\left(q_i - 2\pi i\vec{m}\cdot\vec{\mu}_i\right)\right.$$

$$\left.\times\ \exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right)\right\} + \frac{1}{\pi V}\sum_{\vec{m}\neq 0}\left\{\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\sum_{j}^{N}\left[-2\pi i\vec{m}^{\mu}h^{-1}_{\nu\gamma}\vec{\mu}^{\gamma}_j\ \exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\right.\right.$$

$$\left.\left.-\left(q_j + 2\pi i\vec{m}\cdot\vec{\mu}_j\right)2\pi i\vec{m}^{\mu}h^{-1}_{\nu\gamma}\vec{d}^{\gamma}_j\ \exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\right]\sum_{i}^{N}\left(q_i - 2\pi i\vec{m}\cdot\vec{\mu}_i\right)\exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right)\right\},$$

$$(2) = -\frac{1}{2\pi V}\left(h^{-1}\right)^{T}_{\mu\nu}\sum_{\vec{m}\neq 0}\left\{\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\sum_{j}^{N}\left(2\pi i\vec{m}\cdot\vec{p}_j\right)\exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\sum_{i}^{N}\left(-2\pi i\vec{m}\cdot\vec{p}_i\right)\exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right)\right\}$$

$$+\frac{1}{\pi V}\sum_{\vec{m}\neq 0}\left\{\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\left(-\frac{1+\frac{\pi^2\vec{m}^2}{\beta_0^2}}{\vec{m}^2}\right)\left(-\vec{m}^{\mu}h^{-1}_{\nu\gamma}\vec{m}^{\gamma}\right)\sum_{j}^{N}\left(2\pi i\vec{m}\cdot\vec{p}_j\right)\exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\sum_{i}^{N}\left(-2\pi i\vec{m}\cdot\vec{p}_i\right)\right.$$

$$\left.\times\ \exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right)\right\}+\frac{1}{\pi V}\sum_{\vec{m}\neq 0}\left\{\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\sum_{j}^{N}\left(2\pi i\vec{m}\cdot\vec{p}_j\right)\exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\right.$$

$$\left.\times\sum_{i}^{N}\left[(-2\pi i)\left(-\vec{m}^{\mu}h^{-1}_{\nu\gamma}p_i^{\gamma}\right)\right]\exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right)\right\}+\frac{1}{\pi V}\sum_{\vec{m}\neq 0}\left\{\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\sum_{j}^{N}\left(2\pi i\vec{m}\cdot\vec{p}_j\right)\left(-2\pi i\vec{m}^{\mu}h^{-1}_{\nu\gamma}\vec{d}_j^{\gamma}\right)\right.$$

$$\left.\times\ \exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\sum_{i}^{N}\left(-2\pi i\vec{m}\cdot\vec{p}_i\right)\exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right)\right\},$$

$$(3) = -\frac{1}{\pi V}\left(h^{-1}\right)^{T}_{\mu\nu}\sum_{\vec{m}\neq 0}\left\{\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\sum_{j}^{N}\left(q_j+2\pi i\vec{m}\cdot\vec{\mu}_j\right)\exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\sum_{i}^{N}\left(-2\pi i\vec{m}\cdot\vec{p}_i\right)\exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right)\right\}$$

$$+\frac{2}{\pi V}\sum_{\vec{m}\neq 0}\left\{\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\left(-\frac{1+\frac{\pi^2\vec{m}^2}{\beta_0^2}}{\vec{m}^2}\right)\left(-\vec{m}^{\mu}h^{-1}_{\nu\gamma}\vec{m}^{\gamma}\right)\sum_{j}^{N}\left(q_j+2\pi i\vec{m}\cdot\vec{\mu}_j\right)\exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\sum_{i}^{N}\left(-2\pi i\vec{m}\cdot\vec{p}_i\right)\right.$$

$$\left.\times\ \exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right)\right\}+\frac{1}{\pi V}\sum_{\vec{m}\neq 0}\left\{\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\sum_{j}^{N}\left[-2\pi i\vec{m}^{\mu}h^{-1}_{\nu\gamma}\vec{\mu}_j^{\gamma}\ \exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)-\left(q_j+2\pi i\vec{m}\cdot\vec{\mu}_j\right)\right.\right.$$

$$\left.\left.\times\ 2\pi i\vec{m}^{\mu}h^{-1}_{\nu\gamma}\vec{d}_j^{\gamma}\ \exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\right]\sum_{i}^{N}\left(-2\pi i\vec{m}\cdot\vec{p}_i\right)\exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right)\right\}$$

$$+\frac{1}{\pi V}\sum_{\vec{m}\neq 0}\left\{\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\sum_{j}^{N}\left(q_j+2\pi i\vec{m}\cdot\vec{\mu}_j\right)\exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\sum_{i}^{N}\left(-2\pi i\vec{m}\cdot\vec{p}_i\right)\left(2\pi i\vec{m}^{\mu}h^{-1}_{\nu\gamma}\vec{d}_i^{\gamma}\right)\exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right)\right\}.$$

By adding final expressions for (1), (2), and (3), we arrive at the result for $\frac{\partial U_{rec}}{\partial h_{\mu\nu}}$,

$$\frac{\partial U_{rec}}{\partial h_{\mu\nu}} = -\frac{1}{2}\left(h^{-1}\right)^{T}_{\mu\nu}\cdot\sum_{i}^{N}\left[q_i(\phi_{rec})_i-\left(\vec{\mu}_i+\vec{p}_i\right)\cdot\left(\vec{E}_{rec}\right)_i\right]+\frac{1}{\pi V}\sum_{\vec{m}\neq 0}\left\{\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\left(-\frac{1+\frac{\pi^2\vec{m}^2}{\beta_0^2}}{\vec{m}^2}\right)\right.$$

$$\left.\times\left(-\vec{m}^{\mu}h^{-1}_{\nu\gamma}\vec{m}^{\gamma}\right)S\left(\vec{m}\right)S\left(-\vec{m}\right)\right\}+\sum_{i}^{N}\left(\vec{E}_{rec}\right)_i^{\mu}h^{-1}_{\nu\gamma}\left(\vec{\mu}_i^{\gamma}+\vec{p}_i^{\gamma}+q_i\vec{d}_i^{\gamma}\right)+\sum_{i}^{N}\left(\vec{\mu}_i^{\gamma}+\vec{p}_i^{\gamma}\right)\left(\vec{E}_{rec}\right)_i^{\gamma\mu}h^{-1}_{\nu\xi}\vec{d}_i^{\xi}.$$

Finally, adding up all the above terms and removing the common factor $-\left(h^{-1}\right)^T$, the final expression is the one shown in Sec. II.

4. Screening correction

If certain short-range electrostatic interactions are screened, the pGM system energy can be expressed as

$$U = \sum_i^N \frac{1}{2}(q_i + \vec{\mu}_i \cdot \nabla_i)\phi_i^{0*} + \sum_i^N \frac{1}{2}(\vec{p}_i^{\,*} \cdot \nabla_i)\phi_i^{0*}.$$

Here, ϕ_i^{0*} represents the electrostatic potential on atom i that is created by only charges and covalent dipoles that are not screened in the system, and $\vec{p}_i^{\,*}$ means the induced dipole on atom i that is created under the screening condition. Specifically,

$$\phi_i^{0*} = \frac{1}{\pi V}\sum_{\vec{m}\neq 0}\left\{\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\sum_j^N\left[\left(q_j + 2\pi i\vec{m}\cdot\vec{\mu}_j\right)\exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\right]\exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right)\right\} + \sum_{\substack{j\notin *(i)\\ j\neq i}}^{\infty}\left(q_j + \vec{\mu}_j\cdot\nabla_j\right)$$
$$\times\frac{\mathrm{erf}(\beta_{ij}R_{ij}) - \mathrm{erf}(\beta_0 R_{ij})}{R_{ij}} - \lim_{\vec{R}_j\to\vec{R}_i}\left(q_i + \vec{\mu}_i\cdot\nabla_j\right)\frac{\mathrm{erf}(\beta_0 R_{ij})}{R_{ij}} - \sum_{j\in *(i)}(q_j + \vec{\mu}_j\cdot\nabla_j)\frac{\mathrm{erf}(\beta_0 R_{ij})}{R_{ij}},$$

$$\vec{p}_i^{\,*} = \alpha_i(\vec{E}_i^{0*} - \sum_j^N \vec{T}_{ij}^{\,*}\cdot\vec{p}_j^{\,*}),$$
$$\vec{E}_i^{0*} = -\nabla_i\phi_i^{0*},$$

$$T_{ij}^{*\alpha\beta} = \partial_i^\alpha\partial_j^\beta\begin{cases}\frac{1}{\pi V}\sum_{\vec{m}\neq 0}\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right) + \frac{\mathrm{erf}(\beta_{ij}R_{ij}) - \mathrm{erf}(\beta_0 R_{ij})}{R_{ij}} & \text{if } i\neq j,\ j\notin *(i),\\ \frac{1}{\pi V}\sum_{\vec{m}\neq 0}\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right) - \lim_{\vec{R}_j\to\vec{R}_i}\frac{\mathrm{erf}(\beta_0 R_{ij})}{R_{ij}} & \text{if } i = j,\\ \frac{1}{\pi V}\sum_{\vec{m}\neq 0}\frac{\exp\left(-\frac{\pi^2\vec{m}^2}{\beta_0^2}\right)}{\vec{m}^2}\exp\left(2\pi i\vec{m}\cdot\vec{R}_j\right)\exp\left(-2\pi i\vec{m}\cdot\vec{R}_i\right) - \frac{\mathrm{erf}(\beta_0 R_{ij})}{R_{ij}} & \text{if } j\in *(i).\end{cases}$$

Here, $*(i)$ represent the screened atom pairs of atom i.

Before we calculate the correction term under screening conditions for virial, we want to first show the electrostatic force expression under screening. Following our previous paper, it can be easily shown that

$$\vec{F}_i^{covalent-covalent*} = \sum_j^n \nabla_i\left(\vec{u}_j\right)\cdot\vec{E}_j^{covalent*} + q_i\vec{E}_i^{covalent*} + \vec{u}_i\cdot\vec{E}_i^{covalent*}.$$

The asterisk here means that the quantities are calculated under the screening condition.

The induced part, however, requires some extra attention:

$$\vec{F}_i^{induced} = \vec{p}_i^{\,*}\cdot\vec{E}_i^{induced*} + \sum_j^N \nabla_i(\vec{E}_j^{covalent*})\cdot\vec{p}_j^{\,*}.$$

Here,

$$\sum_j^N \nabla_i\left(\vec{E}_j^{covalent*}\right)\cdot\vec{p}_j^{\,*} = \sum_{j\neq i}^N\sum_{\substack{k\neq j\\ k\notin *(j)}}^n\left(-\nabla_i\left(\vec{u}_k\right)\right)\cdot\nabla_k\left(\vec{p}_j^{\,*}\cdot\nabla_j\right)\frac{\mathrm{erf}\left(\beta_{jk}R_{jk}\right)}{\beta_{jk}R_{jk}} - \sum_{\substack{j\neq i\\ j\notin *(i)}}^N\left(q_i + \vec{\mu}_i\cdot\nabla_i\right)\nabla_i\left(\vec{p}_j^{\,*}\cdot\nabla_j\right)\frac{\mathrm{erf}(\beta_{ij}R_{ij})}{\beta_{ij}R_{ij}}$$
$$- \sum_{\substack{k\neq i\\ k\notin *(i)}}^n\nabla_i\left(\vec{u}_k\right)\cdot\nabla_k\left(\vec{p}_i^{\,*}\cdot\nabla_i\right)\frac{\mathrm{erf}(\beta_{ik}R_{ik})}{\beta_{ik}R_{ik}} - \sum_{\substack{k\neq i\\ k\notin *(i)}}^N\left(q_k + \vec{\mu}_k\cdot\nabla_k\right)\nabla_i\left(\vec{p}_i^{\,*}\cdot\nabla_i\right)\frac{\mathrm{erf}(\beta_{ik}R_{ik})}{\beta_{ik}R_{ik}}$$

$$
\begin{aligned}
&= \sum_{j}^{N} \sum_{\substack{k\neq j \\ k\notin *(j)}}^{n} \left(-\nabla_i\left(\vec{u}_k\right)\right)\cdot\nabla_k\left(\vec{p}_j^{\,*}\cdot\nabla_j\right)\frac{\operatorname{erf}\left(\beta_{jk}R_{jk}\right)}{\beta_{jk}R_{jk}} - \sum_{\substack{j\neq i \\ j\notin *(i)}}^{N}\left(q_i+\vec{\mu}_i\cdot\nabla_i\right)\nabla_i\left(\vec{p}_j^{\,*}\cdot\nabla_j\right)\frac{\operatorname{erf}(\beta_{ij}R_{ij})}{\beta_{ij}R_{ij}} \\
&\quad - \sum_{\substack{k\neq i \\ k\notin *(i)}}^{N}\left(q_k+\vec{\mu}_k\cdot\nabla_k\right)\nabla_i\left(\vec{p}_i^{\,*}\cdot\nabla_i\right)\frac{\operatorname{erf}(\beta_{ik}R_{ik})}{\beta_{ik}R_{ik}} \\
&= \sum_{k}^{n}\sum_{\substack{j\neq k \\ j\notin *(k)}}^{N}\left(-\nabla_i\left(\vec{u}_k\right)\right)\cdot\nabla_k\left(\vec{p}_j^{\,*}\cdot\nabla_j\right)\frac{\operatorname{erf}\left(\beta_{jk}R_{jk}\right)}{\beta_{jk}R_{jk}} + q_i\vec{E}_i^{\,induced*} + \vec{u}_i\cdot\vec{\mathrm{E}}_i^{\,induced*} + \vec{p}_i^{\,*}\cdot\vec{\mathrm{E}}_i^{\,covalent*} \\
&= \sum_{k}^{n}\nabla_i\left(\vec{u}_k\right)\cdot\vec{E}_k^{\,induced*} + q_i\vec{E}_i^{\,induced*} + \vec{u}_i\cdot\vec{\mathrm{E}}_i^{\,induced*} + \vec{p}_i^{\,*}\cdot\vec{\mathrm{E}}_i^{\,covalent*}.
\end{aligned}
$$

In the derivation here, we have used the fact that if $j \in * (i)$ then $i \in * (j)$ for any i and j.

So, the result for the force under the screening condition is

$$\vec{F}_i^{\,*} = \sum_{j}^{n}\nabla_i\left(\vec{u}_j\right)\cdot\vec{E}_j^{\,*} + q_i\vec{E}_i^{\,*} + \left(\vec{u}_i+\vec{p}_i^{\,*}\right)\cdot\vec{E}_i^{\,*}.$$

The asterisk here means that the quantities are calculated under the screening condition.

To derive the correction of screening for virial, we notice that the only variable in the correction term that depends on h is the atom coordinate $\vec{R}$, so, the virial correction is

$$
\begin{aligned}
\frac{\partial U_{correction}}{\partial h_{\mu\nu}} &= \frac{\partial\sum_i^N\frac{1}{2}\left(q_i+\vec{\mu}_i\cdot\nabla_i\right)\sum_{j\neq i}^{\infty}\left(q_j+\vec{\mu}_j\cdot\nabla_j\right)\frac{-\operatorname{erf}\left(\beta_{ij}R_{ij}\right)}{R_{ij}}}{\partial h_{\mu\nu}} \\
&\quad + \frac{1}{2}p_{i'}^{*\alpha'}\frac{\partial\partial_{i'}^{\alpha'}\partial_{j'}^{\beta'}\frac{-\operatorname{erf}\left(\beta_{i'j'}R_{i'j'}\right)}{R_{i'j'}}}{\partial h_{\mu\nu}}p_{j'}^{*\beta'} \\
&\quad + \sum_{i}^{N}\vec{p}_i^{\,*}\cdot\frac{\partial\nabla_i\sum_{j\neq i}^{\infty}\left(q_j+\vec{\mu}_j\cdot\nabla_j\right)\frac{-\operatorname{erf}\left(\beta_{ij}R_{ij}\right)}{R_{ij}}}{\partial h_{\mu\nu}}.
\end{aligned}
$$

Following a similar derivation as in Appendix A 2, we can easily obtain the correction term for the virial expression under screening, as shown in Sec. II.

REFERENCES

[1] H. Wei, R. Qi, J. Wang, P. Cieplak, Y. Duan, and R. Luo, "Efficient formulation of polarizable Gaussian multipole electrostatics for biomolecular simulations," J. Chem. Phys. **153**(11), 114116 (2020).

[2] U. Burkert and N. Allinger, *Molecular Mechanics* (American Chemical Society, Washington, DC, 1982).

[3] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," J. Chem. Phys. **79**(2), 926–935 (1983).

[4] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, "Interaction models for water in relation to protein hydration," in *Intermolecular Forces* (Springer, 1981), pp. 331–342.

[5] M. W. Mahoney and W. L. Jorgensen, "A five-site model for liquid water and the reproduction of the density anomaly by rigid, non-polarizable potential functions," J. Chem. Phys. **112**(20), 8910–8922 (2000).

[6] P. Ren and J. W. Ponder, "Polarizable atomic multipole water model for molecular mechanics simulation," J. Phys. Chem. B **107**(24), 5933–5947 (2003).

[7] D. E. Williams, "Representation of the molecular electrostatic potential by atomic multipole and bond dipole models," J. Comput. Chem. **9**(7), 745–763 (1988).

[8] E. Clementi, H. Kistenmacher, W. Kołos, and S. Romano, "Non-additivity in water-ion-water interactions," Theor. Chim. Acta **55**(4), 257–266 (1980).

[9] T. A. Halgren and W. Damm, "Polarizable force fields," Curr. Opin. Struct. Biol. **11**(2), 236–242 (2001).

[10] S. W. Rick and S. J. Stuart, "Potentials and algorithms for incorporating polarizability in computer simulations," Rev. Comput. Chem. **18**, 89–146 (2002).

[11] J. W. Ponder and D. A. Case, "Force fields for protein simulations," Adv. Protein Chem. **66**, 27–85 (2003).

[12] G. Lamoureux, A. D. MacKerell, Jr., and B. Roux, "A simple polarizable model of water based on classical Drude oscillators," J. Chem. Phys. **119**(10), 5185–5197 (2003).

[13] H. Yu, T. Hansson, and W. F. van Gunsteren, "Development of a simple, self-consistent polarizable model for liquid water," J. Chem. Phys. **118**(1), 221–234 (2003).

[14] S. W. Rick, S. J. Stuart, and B. J. Berne, "Dynamical fluctuating charge force fields: Application to liquid water," J. Chem. Phys. **101**(7), 6141–6156 (1994).

[15] J. Caldwell, L. X. Dang, and P. A. Kollman, "Implementation of nonadditive intermolecular potentials by use of molecular dynamics: Development of a water-water potential and water-ion cluster interactions," J. Am. Chem. Soc. **112**(25), 9144–9147 (1990).

[16] C. J. Burnham, J. Li, S. S. Xantheas, and M. Leslie, "The parametrization of a Thole-type all-atom polarizable water model from first principles and its application to the study of water clusters (n = 2–21) and the phonon spectrum of ice Ih," J. Chem. Phys. **110**(9), 4566–4581 (1999).

[17] J. Applequist, J. R. Carl, and K.-K. Fung, "Atom dipole interaction model for molecular polarizability. Application to polyatomic molecules and determination of atom polarizabilities," J. Am. Chem. Soc. **94**(9), 2952–2960 (1972).

[18] F. J. Vesely, "N-particle dynamics of polarizable Stockmayer-type molecules," J. Comput. Phys. **24**(4), 361–371 (1977).

[19] B. T. Thole, "Molecular polarizabilities calculated with a modified dipole interaction," Chem. Phys. **59**(3), 341–350 (1981).

[20] D. Elking, T. Darden, and R. J. Woods, "Gaussian induced dipole polarization model," J. Comput. Chem. **28**(7), 1261–1274 (2007).

[21] J. Wang, P. Cieplak, R. Luo, and Y. Duan, "Development of polarizable Gaussian model for molecular mechanical calculations I: Atomic polarizability parameterization to reproduce *ab initio* anisotropy," J. Chem. Theory Comput. **15**(2), 1146–1158 (2019).

[22] H. C. Andersen, "Molecular dynamics simulations at constant pressure and/or temperature," J. Chem. Phys. **72**(4), 2384–2393 (1980).

[23] S. Nosé and M. L. Klein, "Constant pressure molecular dynamics for molecular systems," Mol. Phys. **50**(5), 1055–1076 (1983).

[24]T. Darden, "Extensions of the Ewald method for Coulomb interactions in crystals," in *International Tables for Crystallography* (Inernational Union of Crystallography, 2010), Vol. B. Chap. 3.5, pp. 458–481.
[25]A. Toukmaji, C. Sagui, J. Board, and T. Darden, "Efficient particle-mesh Ewald based approach to fixed and induced dipolar interactions," J. Chem. Phys. **113**(24), 10913–10927 (2000).
[26]C. Sagui, L. G. Pedersen, and T. A. Darden, "Towards an accurate representation of electrostatics in classical force fields: Efficient implementation of multipolar interactions in biomolecular simulations," J. Chem. Phys. **120**(1), 73–87 (2004).
[27]J. A. Rackers, Z. Wang, C. Lu, M. L. Laury, L. Lagardère, M. J. Schnieders, J.-P. Piquemal, P. Ren, and J. W. Ponder, "Tinker 8: Software tools for molecular design," J. Chem. Theory Comput. **14**(10), 5273–5289 (2018).
[28]A. Redlack and J. Grindlay, "The electrostatic potential in a finite ionic crystal," Can. J. Phys. **50**(22), 2815–2825 (1972).
[29]A. Redlack and J. Grindlay, "Coulombic potential lattice sums," J. Phys. Chem. Solids **36**(2), 73–82 (1975).
[30]J. Åqvist, P. Wennerström, M. Nervall, S. Bjelic, and B. O. Brandsdal, "Molecular dynamics simulations of water and biomolecules with a Monte Carlo constant pressure algorithm," Chem. Phys. Lett. **384**(4–6), 288–294 (2004).