On removal of charge singularity in Poisson–Boltzmann equation

Qin Cai,[1,2] Jun Wang,[2] Hong-Kai Zhao,[3] and Ray Luo[1,2,a)]
[1]*Department of Biomedical Engineering, University of California, Irvine, California 92697, USA*
[2]*Department of Molecular Biology and Biochemistry, University of California, Irvine, California 92697, USA*
[3]*Department of Mathematics, University of California, Irvine, California 92697, USA*

(Received 22 October 2008; accepted 23 February 2009; published online 8 April 2009)

The Poisson–Boltzmann theory has become widely accepted in modeling electrostatic solvation interactions in biomolecular calculations. However the standard practice of atomic point charges in molecular mechanics force fields introduces singularity into the Poisson–Boltzmann equation. The finite-difference/finite-volume discretization approach to the Poisson–Boltzmann equation alleviates the numerical difficulty associated with the charge singularity but introduces discretization error into the electrostatic potential. Decomposition of the electrostatic potential has been explored to remove the charge singularity explicitly to achieve higher numerical accuracy in the solution of the electrostatic potential. In this study, we propose an efficient method to overcome the charge singularity problem. In our framework, two separate equations for two different potentials in two different regions are solved simultaneously, i.e., the reaction field potential in the solute region and the total potential in the solvent region. The proposed method can be readily implemented with typical finite-difference Poisson–Boltzmann solvers and return the singularity-free reaction field potential with a single run. Test runs on 42 small molecules and 4 large proteins show a very high agreement between the reaction field energies computed by the proposed method and those by the classical finite-difference Poisson–Boltzmann method. It is also interesting to note that the proposed method converges faster than the classical method, though additional time is needed to compute Coulombic potential on the dielectric boundary. The higher precision, accuracy, and efficiency of the proposed method will allow for more robust electrostatic calculations in molecular mechanics simulations of complex biomolecular systems. © *2009 American Institute of Physics*.
[DOI: 10.1063/1.3099708]

I. INTRODUCTION

The importance of solvation in computer simulations is well recognized in computational studies of biomolecules. Due to the often prohibitively high computational expense of explicit solvent models, implicit solvent models have emerged as simple and accurate alternatives for solvation interactions.[1–14] In implicit solvents, solvation interactions are usually separated as nonelectrostatic and electrostatic components. The nonelectrostatic interactions are nowadays modeled as a repulsive cavity component and an attractive dispersion component for accuracy and transferability. To model the electrostatic interactions, a biomolecule is approximated as a cavity of low interior dielectric constant, embedded in a continuous medium of high dielectric constant. The interior of the cavity contains point charges representing charge distributions on atomic centers in the molecule. When mobile ions are present, the ion distribution is also modeled in a mean field fashion, assumed to obey the Boltzmann distribution. With such a continuum approximation of the electrostatic interactions, the electrostatic potential ϕ is characterized by the nonlinear Poisson–Boltzmann equation (PBE),

$$\nabla \cdot \varepsilon \nabla \phi = -4\pi\rho_0 - 4\pi\lambda \sum_i e z_i c_i \exp(-e z_i \phi / k_B T), \quad (1)$$

where ε is the dielectric constant, ρ_0 is the solute charge density, λ is the ion-exclusion function with values of 0 within the Stern layer and 1 outside the Stern layer, e is the unit charge, z_i is the valence of ion type i, c_i is the number density of ion type i, k_B is the Boltzmann constant, and T is the absolute temperature. For a solution with symmetric 1:1 salt, Eq. (1) can be simplified to

$$\nabla \cdot \varepsilon \nabla \phi = -4\pi\rho_0 + \lambda \frac{\varepsilon_{\text{out}} \kappa^2}{C} \sinh(C\phi), \quad (2)$$

where $\kappa^2 = 8\pi e^2 I / \varepsilon_{\text{out}} k_B T$ and $C = ez/k_B T$. Here "out" denotes the outside solvent, and I represents the ionic strength of the solution and is computed as $I = z^2 c$. If the electrostatic potential is weak and the ionic strength is low,[15] the nonlinear PBE can be simplified to a linear PBE,

$$\nabla \cdot \varepsilon \nabla \phi = -4\pi\rho_0 + \lambda \varepsilon_{\text{out}} \kappa^2 \phi. \quad (3)$$

Currently, there are three major numerical methods widely used for solving the PBE. One such approach is the finite-difference method (FDM).[16–26] In this method, the physical properties of the solution, such as the charge density and dielectric constant, are mapped onto a cubic or rectangular lattice, and a discrete approximation to the governing

a)Electronic mail: rluo@uci.edu.

partial differential equations is produced. The second approach is the finite-element method (FEM),[27–32] which is based on the weak variational formulation. The potential to be solved is approximated by a superposition of a set of basis functions. A linear or nonlinear system for the coefficients produced by the weak formulation has to be solved. A nice property of the FEM is that the mesh is unstructured so that adaptivity can be achieved. Also body fitted mesh can be generated to fit nicely to the boundary or the interface, such as the molecular surface. However this strategy may introduce additional cost, such as constant remeshing for dynamical problems. The third approach is the boundary-element method (BEM).[33–46] In BEM, the Poisson equation or PBE is solved for either the induced surface charge[33–36,40,41,43,44] or the normal component of the electric displacement[37–39,42,45,46] on the dielectric boundary between the solute and the solvent.

In biomolecular models, point charges are widely used to represent atomic charge density distribution. Unfortunately this practice introduces singularity into the right hand side of the PBE. The presence of the charge singularity results in discontinuity in the electrostatic potential and causes a large error when FDM or FEM is used.[13,31] To conquer this problem, two regularization schemes for FDM or FEM were proposed.[31,47–49] A straightforward scheme is to remove the singular component, i.e., the singular Coulombic potential, from the electrostatic potential and leave the regular component or the reaction field potential in the differential term,[31,47]

$$\begin{aligned} &\varepsilon_{\text{in}}\nabla^2\phi_C = -4\pi\rho_0 \quad \text{in } \Omega_1 \cup \Omega_2 \cup \Gamma, \\ &\nabla\cdot\varepsilon\nabla\phi_{\text{RF}} - \lambda f(\phi_{\text{RF}}+\phi_C) = 0 \quad \text{in } \Omega_1 \cup \Omega_2, \\ &[\phi_{\text{RF}}] = 0, \quad \left[\varepsilon\frac{\partial\phi_{\text{RF}}}{\partial n}\right] = (\varepsilon_{\text{out}}-\varepsilon_{\text{in}})\frac{\partial\phi_C}{\partial n} \quad \text{on } \Gamma. \end{aligned} \tag{4}$$

Here Ω_1 denotes the solute region, Ω_2 denotes the solvent region, and Γ denotes the dielectric boundary between the solute and the solvent. $f(\bullet)$ represents either the nonlinear term in Eq. (2) or the corresponding linear term in Eq. (3). In this scheme, the Coulombic potential has to be calculated everywhere, and $2M$ unknowns (M is the number of grids) are to be solved.[47] A more practical scheme is to decompose the electrostatic potential into three components: a singular potential, a harmonic potential, and a regular potential,[48,49]

$$\phi = \phi_s + \phi_h + \phi_r, \tag{5}$$

where the singular component and the harmonic component are both defined in the solute and on the dielectric boundary,

$$\phi_s = G \quad \text{in } \Omega_1 \cup \Gamma, \tag{6}$$

$$\begin{aligned} &\nabla^2\phi_h = 0 \quad \text{in } \Omega_1, \\ &\phi_h = -\phi_s \quad \text{on } \Gamma, \end{aligned} \tag{7}$$

where G is the Green's function. The regular component is defined in the whole domain, and the corresponding equation is

$$\begin{aligned} &\nabla\cdot\varepsilon\nabla\phi_r - \lambda f(\phi_r) = 0 \quad \text{in } \Omega_1 \cup \Omega_2, \\ &[\phi_r] = 0, \quad \left[\varepsilon\frac{\partial\phi_r}{\partial n}\right] = -\varepsilon_{\text{in}}\left(\frac{\partial\phi_s}{\partial n}+\frac{\partial\phi_h}{\partial n}\right) \quad \text{on } \Gamma. \end{aligned} \tag{8}$$

The singular component needs to be computed on the dielectric boundary to solve for the harmonic component, which is used to enforce the continuity of the regular component to avoid overfloating when evaluating the exponential function.[48] The regular component solved using Eq. (8) plus the harmonic component solved using Eq. (7) gives the reaction potential in the solute. Thus the solution of the PBE requires the solution of both Eqs. (7) and (8), though the two equations can be solved sequentially.

In this work, we introduce an alternative method to remove the charge singularity in FDM. The Poisson equation is solved for the reaction field potential inside, and simultaneously the PBE is solved for the total potential outside. Unlike the decomposition schemes above, different potentials are assigned on the finite-difference grids based on whether the grid is inside the solute or in the solvent. Therefore there is no need to solve two equations separately for both the harmonic component and the regular component, like the second decomposition scheme above. This method can be implemented with existing FDM solvers without significant change in the program. Another benefit of this method is that the reaction field energy of the molecule can be directly obtained after solving the equation because the reaction field potential is what is needed to compute the reaction field energy.

II. THEORY

A. Solution without salt

It is illustrative to consider a straightforward situation, a molecular solute immersed in a solvent without salt, before considering the general situation where the PBE has to be used. Suppose the solute with dielectric constant ε_{in} is surrounded by the solvent with dielectric constant ε_{out}. Then Eq. (2) or Eq. (3) becomes

$$\nabla\cdot\varepsilon\nabla\phi = -4\pi\rho_0. \tag{9}$$

In Eq. (9), ϕ can be split into the Coulombic potential and the reaction field potential, i.e., $\phi=\phi_C+\phi_{\text{RF}}$. The Coulombic potential is defined as

$$\varepsilon_{\text{in}}\nabla^2\phi_C = -4\pi\rho_0, \tag{10}$$

which means that the Coulombic field is generated by atomic charges in a uniform medium with dielectric constant ε_{in} throughout the whole space. Thus the reaction field is simply the electrostatic field generated by the charges induced by transferring the environment surrounding the solute from ε_{in} to ε_{out}. And a general equation for the reaction field potential is

$$\nabla\cdot\varepsilon\nabla\phi_{\text{RF}} = -4\pi\rho_0 - \nabla\cdot\varepsilon\nabla\phi_C. \tag{11}$$

In doing so, we merely move the Coulombic potential to the right side and treat its effect as some charge distribution. Fortunately Eq. (11) can be further simplified as follows.

In the solute region (Ω_1) with $\varepsilon=\varepsilon_{in}$, use of Eq. (10) gives

$$\varepsilon_{in}\nabla^2\phi_{RF}=-4\pi\rho_0-\varepsilon_{in}\nabla^2\phi_C=0 \quad \text{in } \Omega_1, \tag{12}$$

while in the solvent region (Ω_2) with $\varepsilon=\varepsilon_{out}$, the right side of Eq. (11) is equal to zero because $\rho_0=0$. Thus Eq. (11) is simplified to

$$\varepsilon_{out}\nabla^2\phi_{RF}=-4\pi\rho_0-\varepsilon_{out}\nabla^2\phi_C$$
$$=-4\pi\rho_0\left(1-\frac{\varepsilon_{out}}{\varepsilon_{in}}\right)=0 \quad \text{in } \Omega_2. \tag{13}$$

Hence, in a solution system without salt, the reaction field potential is characterized by the Laplace equation both inside and outside the constant dielectric regions, i.e.,

$$\nabla^2\phi_{RF}=0 \quad \text{in } \Omega_1\cup\Omega_2. \tag{14}$$

However, in the region with variable dielectric constant, only the more general equation [Eq. (11)] applies. Furthermore since there is a jump in dielectric constant at the dielectric boundary (Γ), only the integral form of Eq. (11) should be used. After the application of the divergence theorem, we have

$$\oiint \varepsilon\,\nabla\,\phi_{RF}\cdot d\vec{S}=-\oiint \varepsilon\,\nabla\,\phi_C\cdot d\vec{S} \quad \text{on } \Gamma. \tag{15}$$

Apparently, we have assumed that there is no charge on the dielectric boundary. Equation (15), combined with the continuity requirement of ϕ_{RF}, leads to the following jump conditions across Γ:

$$[\phi_{RF}]=0,$$
$$\left[\varepsilon\frac{\partial\phi_{RF}}{\partial n}\right]=-\left[\varepsilon\frac{\partial\phi_C}{\partial n}\right]. \tag{16}$$

B. Solution with salt

Now consider the general situation, a molecular solute solved in a solution system with salt. Different from the situation without salt, an ion-exclusion Stern layer exists in the solvent region (Ω_2), which is typically set 2.0 Å away from the dielectric interface. Since there is no atomic charge in the solvent region, Eqs. (2) and (3) in Ω_2 can be reduced to

$$\nabla\cdot\varepsilon\,\nabla\,\phi-\lambda f(\phi)=0, \tag{17}$$

where $\lambda=0$ within the Stern layer and $\lambda=1$ outside the Stern layer. If we split the total potential in Eq. (17) and solve for the reaction field potential as in the case of solution without salt, we have to calculate the Coulombic potential at every grid outside, as reviewed in Sec. I. An alternative way is to keep Eq. (17) unchanged and solve for the total potential outside. However the boundary conditions need to be modified to make a connection between the reaction field potential inside and the total potential outside. Apparently we need the relation $\phi=\phi_C+\phi_{RF}$ in this strategy.

To summarize, we solve Eq. (12) for the reaction field potential in Ω_1 and solve Eq. (17) for the total potential in Ω_2, subject to the relation of $\phi=\phi_C+\phi_{RF}$,

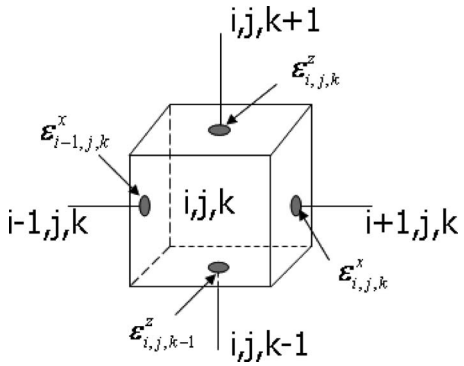

FIG. 1. Finite-volume/finite-difference discretization scheme.

$$\nabla\cdot\varepsilon\,\nabla\,\phi_{RF}=-4\pi\rho_0-\nabla\cdot\varepsilon\,\nabla\,\phi_C \quad \text{in } \Omega_1,$$
$$\nabla\cdot\varepsilon\,\nabla\,\phi-\lambda f(\phi)=0 \quad \text{in } \Omega_2. \tag{18}$$

Note that the first equation in Eq. (18) is simply Eq. (11) so that its integral form is also the same as Eq. (15) on Γ. And again the corresponding jump condition across Γ is readily obtained from its integral form, which is the same as Eq. (16). Therefore the jump conditions are

$$[\phi-\phi_{RF}]=[\phi_C],$$
$$\left[\varepsilon\frac{\partial\phi_{RF}}{\partial n}\right]=-\left[\varepsilon\frac{\partial\phi_C}{\partial n}\right]. \tag{19}$$

Finally it is worth pointing out that the same jump condition can also be derived from the second equation in Eq. (18) because $\phi=\phi_C+\phi_{RF}$.

III. NUMERICAL PROCEDURES

Our finite-difference scheme is based on the integral form of Eq. (18). Thus this discretization scheme should be more appropriately termed the finite-volume method (Fig. 1). Across the interface, the finite-volume method keeps its simple form, as will be shown below. Although this strategy may cause loss of some accuracy since it does not consider the location of the dielectric boundary with subgrid resolution, the simple formulation suits our current purpose to demonstrate the basic idea of the charge-singularity removal scheme.

Use of the divergence theorem on the integral form of Eq. (18) gives

$$\oiint \varepsilon\,\nabla\,\phi_{RF}\cdot d\vec{S}=-4\pi q_0-\oiint \varepsilon\,\nabla\,\phi_C\cdot d\vec{S}, \tag{20}$$

$$\oiint \varepsilon\,\nabla\,\phi\cdot d\vec{S}=\int\int\int \lambda f(\phi)dV. \tag{21}$$

On grid (i,j,k), Eqs. (20) and (21) are discretized into the following:

$$\varepsilon^x_{i-1,j,k}(\phi^{\mathrm{RF}}_{i-1,j,k}-\phi^{\mathrm{RF}}_{i,j,k})+\varepsilon^x_{i,j,k}(\phi^{\mathrm{RF}}_{i+1,j,k}-\phi^{\mathrm{RF}}_{i,j,k})+\varepsilon^y_{i,j-1,k}(\phi^{\mathrm{RF}}_{i,j-1,k}-\phi^{\mathrm{RF}}_{i,j,k})+\varepsilon^y_{i,j,k}(\phi^{\mathrm{RF}}_{i,j+1,k}-\phi^{\mathrm{RF}}_{i,j,k})+\varepsilon^z_{i,j,k-1}(\phi^{\mathrm{RF}}_{i,j,k-1}-\phi^{\mathrm{RF}}_{i,j,k})$$
$$+\varepsilon^z_{i,j,k}(\phi^{\mathrm{RF}}_{i,j,k+1}-\phi^{\mathrm{RF}}_{i,j,k})=-\varepsilon^x_{i-1,j,k}(\phi^{C}_{i-1,j,k}-\phi^{C}_{i,j,k})-\varepsilon^x_{i,j,k}(\phi^{C}_{i+1,j,k}-\phi^{C}_{i,j,k})-\varepsilon^y_{i,j-1,k}(\phi^{C}_{i,j-1,k}-\phi^{C}_{i,j,k})-\varepsilon^y_{i,j,k}(\phi^{C}_{i,j+1,k}-\phi^{C}_{i,j,k})$$
$$-\varepsilon^z_{i,j,k-1}(\phi^{C}_{i,j,k-1}-\phi^{C}_{i,j,k})-\varepsilon^z_{i,j,k}(\phi^{C}_{i,j,k+1}-\phi^{C}_{i,j,k})-4\pi q_0/h, \tag{22}$$

$$\varepsilon^x_{i-1,j,k}(\phi_{i-1,j,k}-\phi_{i,j,k})+\varepsilon^x_{i,j,k}(\phi_{i+1,j,k}-\phi_{i,j,k})+\varepsilon^y_{i,j-1,k}(\phi_{i,j-1,k}-\phi_{i,j,k})+\varepsilon^y_{i,j,k}(\phi_{i,j+1,k}-\phi_{i,j,k})+\varepsilon^z_{i,j,k-1}(\phi_{i,j,k-1}-\phi_{i,j,k})$$
$$+\varepsilon^z_{i,j,k}(\phi_{i,j,k+1}-\phi_{i,j,k})-h^2\lambda f(\phi_{i,j,k})=0. \tag{23}$$

Here it is implied that the unknown potentials on the grids to be solved are either reaction field potentials or total potentials based on whether the grids are inside or outside.

The following notations are introduced to accommodate the use of different potentials in different regions in Eqs. (22) and (23). (1) Ω_1 is used to denote the solute grids with all of their six neighbor grids also within the solute region. (2) Ω_2 is used to denote the solvent grids with all of their six neighbor grids within the solvent region. (3) Γ_1 is used to denote the solute grids with one or more of their six neighbor grids in the solvent region. (4) Γ_2 is used to denote the solvent grids with one or more of their six neighbor grids in the solute region. (5) $\bar{\Omega}_1$ is defined as $\Omega_1 \cup \Gamma_1$. (6) $\bar{\Omega}_2$ is defined as $\Omega_2 \cup \Gamma_2$.

For grids in Ω_1, Eq. (22) can be simplified to

$$\varepsilon^x_{i-1,j,k}(\phi^{\mathrm{RF}}_{i-1,j,k}-\phi^{\mathrm{RF}}_{i,j,k})+\varepsilon^x_{i,j,k}(\phi^{\mathrm{RF}}_{i+1,j,k}-\phi^{\mathrm{RF}}_{i,j,k})$$
$$+\varepsilon^y_{i,j-1,k}(\phi^{\mathrm{RF}}_{i,j-1,k}-\phi^{\mathrm{RF}}_{i,j,k})+\varepsilon^y_{i,j,k}(\phi^{\mathrm{RF}}_{i,j+1,k}-\phi^{\mathrm{RF}}_{i,j,k})$$
$$+\varepsilon^z_{i,j,k-1}(\phi^{\mathrm{RF}}_{i,j,k-1}-\phi^{\mathrm{RF}}_{i,j,k})+\varepsilon^z_{i,j,k}(\phi^{\mathrm{RF}}_{i,j,k+1}-\phi^{\mathrm{RF}}_{i,j,k})=0, \tag{24}$$

with the help of Eq. (12). For grids in Ω_2, Eq. (23) can be used directly. For boundary grids in Γ_1, Eq. (22) can be used by noting that one or more of their six neighbor grids of (i,j,k) are in $\bar{\Omega}_2$ where we only know the total potential. Therefore we replace the reaction field potentials (ϕ_{RF}) on the neighbor grids in $\bar{\Omega}_2$ with $\phi-\phi_C$ in Eq. (22). For boundary grids in Γ_2, Eq. (23) can be used by noting that one or more of their six neighbor grids of (i,j,k) are in $\bar{\Omega}_1$ where we only know the reaction field potentials. Therefore we replace the total potentials (ϕ) on the neighbor grids in $\bar{\Omega}_1$ with $\phi_{\mathrm{RF}}+\phi_C$ in Eq. (23). In doing so, we have the following equations serving as jump conditions on the dielectric boundary grids ($\Gamma_1 \cup \Gamma_2$) in our new method:

$$\sum_m \varepsilon_m \phi^{\mathrm{RF}}_m + \sum_n \varepsilon_n \phi_n - \varepsilon_{i,j,k}\phi^{\mathrm{RF}}_{i,j,k} = -\sum_m \varepsilon_m \phi^C_m + \varepsilon_{i,j,k}\phi^C_{i,j,k}, \tag{25}$$

$$\sum_m \varepsilon_m \phi^{\mathrm{RF}}_m + \sum_n \varepsilon_n \phi_n - [\varepsilon_{i,j,k}\phi_{i,j,k} + h^2\lambda f(\phi_{i,j,k})] = -\sum_m \varepsilon_m \phi^C_m, \tag{26}$$

where Σ_m is a sum over neighbor points in the solute, Σ_n is a sum over neighbor points in the solvent, and $\varepsilon_{i,j,k}=\varepsilon^x_{i,j,k}+\varepsilon^x_{i-1,j,k}+\varepsilon^y_{i,j,k}+\varepsilon^y_{i,j-1,k}+\varepsilon^z_{i,j,k}+\varepsilon^z_{i,j,k-1}$.

As described above, the Coulombic potential is required on the dielectric boundary grids. Here the Coulombic potential is computed with the finite-difference Green's function,[50]

$$g(\Delta x,\Delta y,\Delta z)=\frac{1}{\pi^2}\int_0^\pi\int_0^\pi\int_0^\pi \frac{\cos(w\Delta x)\cos(u\Delta y)\cos(v\Delta z)}{\sin^2(w/2)+\sin^2(u/2)+\sin^2(v/2)}dwdudv, \tag{27}$$

where $(\Delta x,\Delta y,\Delta z)$ is the distance vector between the source and field point in the grid unit. The Coulombic potential is then given by

$$\phi(\Delta x,\Delta y,\Delta z)=q\times g(\Delta x,\Delta y,\Delta z) \tag{28}$$

for charge q. Here the use of the finite-difference Green's function is to achieve an optimal consistency with the classical FDM solution of the PBE for easy validation of the new singularity-free formalism. Note that use of analytical Green's function $(1/r)$ may improve the accuracy of the reaction field potential if the jump conditions are explicitly enforced on the analytical dielectric boundary.[49] In this work the finite-difference Green's functions are precomputed for grid separation up to 20 per dimension. Beyond 20 grids, the difference between the analytical Green's function and the finite-difference Green's function approach is extremely small, so the analytical form $1/r$ is used.

On the space boundary (Π), one or more of the neighbor grid points are outside the domain of interest. We approximate the potentials at these points by using the analytical solution to the linearized PBE, which is also called the Debye–Huckel approximation. These potentials are therefore in the form of

$$\phi^{\mathrm{out}}=\frac{1}{\varepsilon_{\mathrm{out}}}\sum_i \frac{Q_i \exp[-\kappa(R_i-r_i)]}{R_i(1+\kappa r_i)}. \tag{29}$$

Here Σ_i is a sum over all atoms, Q_i is the charge of the nth atom, r_i is the radius of the ith atom, and R_i is the distance between the ith atom and the grid. In this equation, the underlying assumption is that the space boundary grid is so far away from the molecule that the molecule can be regarded as spherically symmetrical. It should be pointed out that the Debye–Huckel approximation overpredicts the electrostatic potential in the solution of the nonlinear PBE,[51] though an analysis of the accuracy of the space boundary potential is beyond the scope of this work.

The terms corresponding to the grid points outside the domain of interest are moved to the right side of Eq. (23). This is equivalent to treating them as induced charges on the space boundary. Thus we have

$$\sum_m \varepsilon_{\text{out}}\phi^{\text{in}} - [6\varepsilon_{\text{out}}\phi_{i,j,k} + h^2\lambda f(\phi_{i,j,k})] = -\sum_n \varepsilon_{\text{out}}\phi^{\text{out}} = -\sum_n \sum_i \frac{Q_i \exp[-\kappa(R_i - r_i)]}{R_i(1+\kappa r_i)}. \quad (30)$$

Here ϕ^{in} is the potential on the grid within the space boundary, ϕ^{out} is the potential on the grid outside the space boundary, Σ_m is a sum over neighbor points in the solvent, Σ_n is a sum over neighbor points outside the space boundary, and Σ_i is a sum over all atoms.

After the Coulombic potential is computed on all the grids on the dielectric boundary, a linear/nonlinear system of unknown reaction field potentials inside and total potentials outside can be constructed by combining Eqs. (23)–(30).

Three matrices can be introduced to store the dielectric constants along three orthogonal grid edges, respectively, i.e., $\{\varepsilon^x_{i,j,k}\}$, $\{\varepsilon^y_{i,j,k}\}$, and $\{\varepsilon^z_{i,j,k}\}$. After introducing the Kronecker delta function, a uniform finite-difference linear/nonlinear system can be expressed as

$$\begin{aligned}&(1-\delta_{i,1})\varepsilon^x_{i-1,j,k}\varphi_{i-1,j,k} + (1-\delta_{i,xm})\varepsilon^x_{i,j,k}\varphi_{i+1,j,k} \\ &+ (1-\delta_{j,1})\varepsilon^y_{i,j-1,k}\varphi_{i,j-1,k} + (1-\delta_{j,ym})\varepsilon^y_{i,j,k}\varphi_{i,j+1,k} \\ &+ (1-\delta_{k,1})\varepsilon^z_{i,j,k-1}\varphi_{i,j,k-1} + (1-\delta_{k,zm})\varepsilon^z_{i,j,k}\varphi_{i,j,k+1} \\ &- [\varepsilon_{i,j,k}\varphi_{i,j,k} + \lambda h^2 f(\varphi_{i,j,k})] = q_{i,j,k}. \end{aligned} \quad (31)$$

In Eq. (31), the following notations are used: $\varphi = \phi_{\text{RF}}$ in $\bar{\Omega}_1$ and $\varphi = \phi$ in $\bar{\Omega}_2$, and $xm \times ym \times zm$ is the dimension of the finite-difference grids. Finally $q_{i,j,k}$ is defined as

$$q_{i,j,k} = \begin{cases} 0 \quad \text{in } \Omega_1 \cup \Omega_2 \\ -\sum_m \varepsilon_m \phi^C_m + (1-\lambda)\varepsilon_{i,j,k}\phi^C_{i,j,k} \quad \text{in } \Gamma_1 \cup \Gamma_2 \\ -\delta_{i,xm}\varepsilon^x_{i,j,k}\phi^{\text{out}}_{i+1,j,k} - \delta_{i,1}\varepsilon^x_{i-1,j,k}\phi^{\text{out}}_{i-1,j,k} - \delta_{j,ym}\varepsilon^y_{i,j,k}\phi^{\text{out}}_{i,j+1,k} - \delta_{j,1}\varepsilon^y_{i,j-1,k}\phi^{\text{out}}_{i,j-1,k} - \delta_{k,zm}\varepsilon^z_{i,j,k}\phi^{\text{out}}_{i,j,k+1} - \delta_{k,1}\varepsilon^z_{i,j,k-1}\phi^{\text{out}}_{i,j,k-1} \quad \text{on } \Pi, \end{cases} \quad (32)$$

where Σ_m is a sum over neighboring grid points in the solute.

Note that the space boundary charge is the same as in the classical FDM. Thus to implement Eqs. (31) and (32) in a FDM, first find all dielectric boundary grids ($\Gamma_1 \cup \Gamma_2$). This information is readily available when the molecular surface is mapped to the finite-difference grids to determine the dielectric constants on all grid edges. Second add "dielectric boundary grid" charges according to Eq. (32) with the solute Coulombic potentials. Finally eliminate the atomic charges on the grids after computing the Coulombic potentials above. All other procedures remain the same as in the classical FDM.

A few other operations are needed to discretize the solution system as in any FDM, including mapping atomic charges to the finite-difference grid points and defining the dielectric boundary. In this study, a standard trilinear mapping of charges was used.[52] To obtain the three matrices of the dielectric constants $\{\varepsilon^x_{i,j,k}\}$, $\{\varepsilon^y_{i,j,k}\}$, and $\{\varepsilon^z_{i,j,k}\}$, a definition of a molecular surface is required, i.e., a boundary between ε_{in} and ε_{out}. There are several choices available, and the solvent excluded surface or the Richards surface was used,[53] as implemented in the PBSA program of the Amber 10 package.[54] Finally, the solver for the linearized PBE is a modified incomplete Cholsky conjugate gradient (MICCG) algorithm[24,55] as implemented in PBSA. Eisenstat's[56] two optimization techniques were adopted to enhance the MICCG solver in PBSA.[24,57]

IV. SIMULATION DETAILS

All calculations were conducted with a modified Amber 10 package.[58] Only the linearized PBE was tested in the current study. The dielectric constant ε_{in} was set to 1, and ε_{out} was set to 80. The solvent probe was set to be 1.4 Å to compute the solvent excluded surface that was used as the solute/solvent dielectric boundary. The ionic strength was set to 150 mM. The ion probe was set to be 2.0 Å to compute the ion accessible surface that was used as the interface between the Stern layer and the bulk ion accessible solvent region. The finite-difference grid spacing was set as 0.5 Å. The ratio between the longest dimension of the finite-difference grid and that of the solute was 5 for small molecules and 2 for proteins. The convergence for the linear system was set to be 10^{-6} if it is not mentioned otherwise.

We used 42 small molecules and 4 large proteins to test the consistency of the new method with the classical approach; specifically the reaction field energies were compared to those by the classical FDM, i.e., subtraction of the total electrostatic potential with $\varepsilon_{\text{in}}=\varepsilon_{\text{out}}=1$ from the total electrostatic potential with $\varepsilon_{\text{in}}=1$ and $\varepsilon_{\text{out}}=80$,

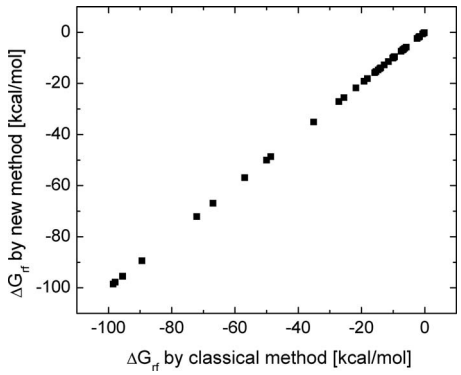

FIG. 2. Correlation between the reaction field energies by the classical method and the new method for the 42 small molecules.

$$\Delta G_{rf} = G(\varepsilon_{\text{out}} = 80) - G(\varepsilon_{\text{out}} = 1). \tag{33}$$

The 42 small molecules include 25 amino acid side chain analogs (both neutral and charged), NMA (*n*-methylamine), three dipeptides (GLY, ALA, and PRO), 5 bases (from five nucleotides: DA, DC, DG, DT, and DU), 8 phosphate analogs with sugars (DAP from DA, DA3P from DA3, DA with a 3′-OH end group, DA5P from DA5, DA with a 5′-OH end group, DANP from DAN, DA with both 5′-OH and 3′-OH end groups, and their corresponding parts from RA, RA3, RA5, and RAN, respectively).[59] These small molecules cover all the standard templates for proteins and nucleic acids in the Amber force fields.[58] The PDB code of the four proteins are 1pgb (56 residues and 855 atoms), 2trx (108 residues and 1654 atoms), 1tsr (194 residues and 3010 atoms), and 1e6q (499 residues and 7819 atoms).

V. RESULTS AND DISCUSSION

The original proposals to remove charge singularity were to increase the numerical accuracy of the FDM. However, it was later found that the removal of charge singularity alone has little effect to enhance the numerical accuracy of the FDM without proper enforcement of jump conditions across the dielectric interface.[47,49] This study focuses on the validity of the new decomposition scheme to remove charge singularity. Thus exactly the same physical and numerical model as the classical method was used for easy validation. Furthermore the new method was implemented with the same PB solver used in the classical method. Finally the electrostatic energy was computed as $\frac{1}{2}\Sigma_i q_i \phi_i$. That is to say, the discretization approximation in both charges and the dielectric boundary is still there. Therefore the numerical results for the same quantity calculated by both methods are supposed to be equal in theory.

Indeed, as shown in Fig. 2, a high level agreement between the new method and the classical method can be observed in terms of the reaction field energies for the 42 tested small molecules. The number of the atoms of the 42 small molecules ranges from 3 to 26, and the reaction field energies range from −98.5331 to −0.1291 kcal/mol. The correlation coefficient between the two sets of data is 1.000 006, and the rms relative deviation is 2.95×10^{-4}%. The reaction field energy calculations for the four tested proteins are listed in Table I. The relative deviations between the two methods for the four proteins are 1.54×10^{-4}%, 1.01×10^{-3}%, 1.08×10^{-3}%, and 7.90×10^{-5}%, respectively.

We also evaluated the convergence behaviors of the new and classical methods by counting the number of iterations necessary to reach a certain precision level in the calculated reaction field energies. The precision level here is defined as the relative deviation between the intermediate result and the final result of the iterations. In this analysis, we tested convergence criteria up to 10^{-12}, with which the precision level of the reaction field energy is supposed to be advanced to 10^{-12}. Our data show that the classical method can reach the 10^{-9} precision level for the relatively small proteins, 1pgb and 1trx, but only 10^{-7} for the two larger proteins, 1tsr and 1e6q, while the new method can truly reach the 10^{-12} level for all four tested proteins. As shown in Fig. 3, the new method converges continuously for whatever sized tested proteins, while the classical method has difficulty in reaching a higher precision level, especially for larger proteins. This is indicated by the flat region in the convergence plots. The flat region tends to expand for larger proteins and thereby results in deterioration in the convergence quality. In addition the new method exhibits better convergence rate. For example, to obtain the 10^{-6} precision level, the classical method needs roughly twice iterations than the new method for all four tested proteins.

The different efficiency between the classical and new methods can be appreciated by the fact that the classical method keeps the singular charges in the equation and thus suffers from the large error and the difficulty in converging a larger potential value near singular charge sources. While the larger error is believed to cancel via Eq. (33) if the same grid and charge mapping strategy are used, the larger potential still causes loss of significant digits in the calculation of the reaction field energy and the precision is consequently re-

TABLE I. Calculated reaction field energies (kcal/mol) and computation times (second) for the four tested proteins. DBC time: computation time spent on the DBC; SBC time: computation time spent on the SBC. The SBC time before the slash corresponds to the case of $\varepsilon_{\text{in}}=1$ and $\varepsilon_{\text{out}}=80$, while the SBC time after the slash corresponds to the case of $\varepsilon_{\text{in}}=\varepsilon_{\text{out}}=1$.

Protein	1pgb		2trx		1tsr		1e6q	
No. of atoms	855		1654		3010		7819	
Method	Classical	New	Classical	New	Classical	New	Classical	New
ΔG_{rf}	−1023.95	−1023.95	−1880.52	−1880.50	−2332.09	−2332.11	−4088.06	−4088.07
DBC time	NA	20	NA	82	NA	251	NA	1326
SBC time	46/35	46	154/114	153	541/384	538	2411/1745	2411

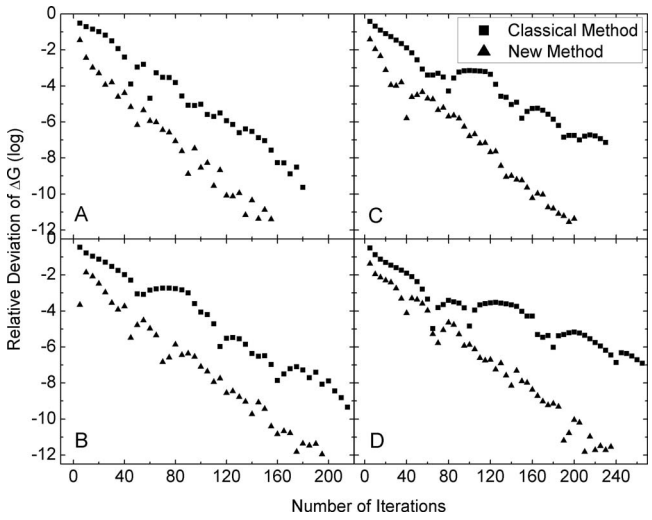

FIG. 3. Precision level vs number of iterations between the new method and the classical method. The *y*-axis is obtained by taking the logarithm of the relative deviation, with the reference taken as the converged reaction field energy. Here the convergence criterion is 10^{-12}. (a) 1pgb. (b) 2trx. (c) 1tsr. (d) 1e6q.

duced. The larger potential comes from two sources in the classical approach. The first source is the self-potential for the singular point charges. In FDM, the singular charges are mapped to the eight corners of each lattice cell, and there arises the artificial interaction between the eight grid charges. The second source is the large short-range Coulombic potential between charges in close contact. Since the new method solves the Laplace equation in the solute that does not contain any singular charges, it is free of the above charge-singularity related limitations.

Note that our new method requires the calculation of the Coulombic potential on the dielectric boundary while enforcing the boundary condition. To evaluate the additional cost, we compared the computation time on the dielectric boundary condition (DBC) to the computation time on the space boundary condition (SBC) (see Table I). The DBC time is typically half of the SBC time in the new method. The difference in SBC time between the $\varepsilon_{out}=80$ case and the $\varepsilon_{out}=1$ case arises from the different ion concentrations in the solvent. Thus despite the additional time cost in DBC, the new method is overall faster than the classical method to obtain the reaction field energy.

VI. CONCLUSION

In the present work we have proposed an efficient method to remove the charge singularity of the PBE. In our method, two differential equations for two different potentials in two different regions are solved simultaneously, i.e., the reaction field potential in the solute region and the total potential in the solvent region. This method can readily be implemented within any finite-difference Poisson–Boltzmann solver and can give the reaction field energy with a single run.

We validated the new method by comparing the calculated reaction field energies of 42 small molecules and 4 large proteins with those obtained by the classical method. The result shows a very good agreement between the new and classical methods, with a rms relative deviation of $2.95\times10^{-4}\%$ for the tested small molecules and a rms relative deviation of $1\times10^{-3}\%$ for the tested proteins.

More interestingly, further evaluation of the precision level and convergence rate of the new method shows additional benefit in adopting the new decomposition scheme. The new method was found to converge continuously and smoothly for all tested proteins, while the classical method has difficulty reaching higher precision for larger proteins. This directly leads to the new method's faster convergence rate. Indeed the classical method needs roughly twice iterations than the new one for all four tested proteins, for example, to reach the convergence of 10^{-6}. Note that additional time is needed for the calculation of the Coulombic potential on the dielectric boundary in the new decomposition scheme. However the total time used for the boundary potential in the new method is still less than the time of two runs used in the classical method to obtain the reaction field energy.

Finally accuracy improvement can be achieved in the finite-difference calculation when the removal of charge singularity is coupled with the enforcement of dielectric jump conditions at the analytical solute/solvent interface. Indeed, the coupled strategy was found to improve the accuracy and precision of the finite-difference calculation for a tested analytical system in a separate study.[60] Furthermore the removal of charge singularity alone was found to have the additional benefit of significantly improving the accuracy of the dielectric boundary force for the tested system.[60]

In summary the new method allows for higher efficiency, precision, and accuracy in electrostatic energy and force evaluations, which in turn allow for more robust biomolecular simulations to be conducted. Specifically, the new method can certainly find applications in areas where tens of millions of electrostatic calculations are needed, such as conformational sampling and free energy simulations, where the electrostatic calculation is often the bottleneck in both efficiency and accuracy since other molecular mechanics terms can be analytically computed with very high efficiency.

[1] M. E. Davis and J. A. McCammon, Chem. Rev. (Washington, D.C.) **90**, 509 (1990).
[2] K. A. Sharp, Curr. Opin. Struct. Biol. **4**, 234 (1994).
[3] M. K. Gilson, Curr. Opin. Struct. Biol. **5**, 216 (1995).
[4] B. Honig and A. Nicholls, Science **268**, 1144 (1995).
[5] B. Roux and T. Simonson, Biophys. Chem. **78**, 1 (1999).
[6] C. J. Cramer and D. G. Truhlar, Chem. Rev. (Washington, D.C.) **99**, 2161 (1999).
[7] D. Bashford and D. A. Case, Annu. Rev. Phys. Chem. **51**, 129 (2000).
[8] N. A. Baker, Curr. Opin. Struct. Biol. **15**, 137 (2005).
[9] J. H. Chen, W. P. Im, and C. L. Brooks, J. Am. Chem. Soc. **128**, 3728 (2006).
[10] M. Feig, J. Chocholousova, and S. Tanizaki, Theor. Chem. Acc. **116**, 194 (2006).
[11] W. Im, J. H. Chen, and C. L. Brooks, *Peptide Solvation and H-Bonds* (Elsevier, San Diego, 2006), Vol. 72, p. 173.
[12] P. Koehl, Curr. Opin. Struct. Biol. **16**, 142 (2006).
[13] B. Z. Lu, Y. C. Zhou, M. J. Holst, and J. A. McCammon, Comm. Comp. Phys. **3**, 973 (2008).

[14] J. Wang, C. H. Tan, Y. H. Tan, Q. Lu, and R. Luo, Comm. Comp. Phys. **3**, 1010 (2008).
[15] T. L. Hill, *An Introduction to Statistical Thermodynamics* (Dover, New York, 1986).
[16] I. Klapper, R. Hagstrom, R. Fine, K. Sharp, and B. Honig, Proteins **1**, 47 (1986).
[17] M. E. Davis and J. A. McCammon, J. Comput. Chem. **10**, 386 (1989).
[18] A. Nicholls and B. Honig, J. Comput. Chem. **12**, 435 (1991).
[19] B. A. Luty, M. E. Davis, and J. A. McCammon, J. Comput. Chem. **13**, 1114 (1992).
[20] M. Holst and F. Saied, J. Comput. Chem. **14**, 105 (1993).
[21] K. E. Forsten, R. E. Kozack, D. A. Lauffenburger, and S. Subramaniam, J. Phys. Chem. **98**, 5580 (1994).
[22] W. Im, D. Beglov, and B. Roux, Comput. Phys. Commun. **111**, 59 (1998).
[23] W. Rocchia, E. Alexov, and B. Honig, J. Phys. Chem. B **105**, 6507 (2001).
[24] R. Luo, L. David, and M. K. Gilson, J. Comput. Chem. **23**, 1244 (2002).
[25] D. Bashford, Lect. Notes Comput. Sci. **1343**, 233 (1997).
[26] M. K. Gilson, K. A. Sharp, and B. H. Honig, J. Comput. Chem. **9**, 327 (1988).
[27] C. M. Cortis and R. A. Friesner, J. Comput. Chem. **18**, 1591 (1997).
[28] N. Baker, M. Holst, and F. Wang, J. Comput. Chem. **21**, 1343 (2000).
[29] M. Holst, N. Baker, and F. Wang, J. Comput. Chem. **21**, 1319 (2000).
[30] A. I. Shestakov, J. L. Milovich, and A. Noy, J. Colloid Interface Sci. **247**, 62 (2002).
[31] L. Chen, M. J. Holst, and J. C. Xu, SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal. **45**, 2298 (2007).
[32] D. Xie and S. Zhou, BIT Numer. Math. **47**, 853 (2007).
[33] S. Mierts, E. Scrocco, and J. Tomasi, Chem. Phys. **55**, 117 (1981).
[34] H. Hoshi, M. Sakurai, Y. Inoue, and R. Chujo, J. Chem. Phys. **87**, 1107 (1987).
[35] R. J. Zauhar and R. S. Morgan, J. Comput. Chem. **9**, 171 (1988).
[36] A. A. Rashin, J. Phys. Chem. **94**, 1725 (1990).
[37] B. J. Yoon and A. M. Lenhoff, J. Comput. Chem. **11**, 1080 (1990).
[38] A. H. Juffer, E. F. F. Botta, B. A. M. Vankeulen, A. Vanderploeg, and H. J. C. Berendsen, J. Comput. Phys. **97**, 144 (1991).
[39] H. X. Zhou, Biophys. J. **65**, 955 (1993).
[40] R. Bharadwaj, A. Windemuth, S. Sridharan, B. Honig, and A. Nicholls, J. Comput. Chem. **16**, 898 (1995).
[41] E. O. Purisima and S. H. Nilar, J. Comput. Chem. **16**, 681 (1995).
[42] J. Liang and S. Subramaniam, Biophys. J. **73**, 1830 (1997).
[43] Y. N. Vorobjev and H. A. Scheraga, J. Comput. Chem. **18**, 569 (1997).
[44] M. Totrov and R. Abagyan, Biopolymers **60**, 124 (2001).
[45] A. H. Boschitsch, M. O. Fenley, and H. X. Zhou, J. Phys. Chem. B **106**, 2741 (2002).
[46] B. Z. Lu, X. L. Cheng, J. F. Huang, and J. A. McCammon, Proc. Natl. Acad. Sci. U.S.A. **103**, 19314 (2006).
[47] Z. X. Zhou, P. Payne, M. Vasquez, N. Kuhn, and M. Levitt, J. Comput. Chem. **17**, 1344 (1996).
[48] I.-L. Chern, J.-G. Liu, and W.-C. Wang, Methods Appl. Anal. **10**, 309 (2003).
[49] W. H. Geng, S. N. Yu, and G. W. Wei, J. Chem. Phys. **127**, 114106 (2007).
[50] B. A. Luty, M. E. Davis, and J. A. McCammon, J. Comput. Chem. **13**, 768 (1992).
[51] A. H. Boschitsch and M. O. Fenley, J. Comput. Chem. **28**, 909 (2007).
[52] D. T. Edmonds, N. K. Rogers, and M. J. E. Sternberg, Mol. Phys. **52**, 1487 (1984).
[53] F. M. Richards, Annu. Rev. Biophys. Bioeng. **6**, 151 (1977).
[54] C. Tan, Y. H. Tan, and R. Luo, J. Phys. Chem. B **111**, 12263 (2007).
[55] I. Gustafsson, BIT, Nord. Tidskr. Inf.behandl. **18**, 142 (1978).
[56] S. C. Eisenstat, SIAM J. Comput. **18**, 142 (1978).
[57] Q. Lu and R. Luo, J. Chem. Phys. **119**, 11035 (2003).
[58] D. A. Case, T. A. Darden, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, M. Crowley, R. C. Walker, W. Zhang, K. M. Merz, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, D. H. Mathews, M. G. Seetin, C. Sagui, V. Babin, and P. A. Kollman, Amber 10, University of California, San Francisco, 2008.
[59] C. H. Tan, L. J. Yang, and R. Luo, J. Phys. Chem. B **110**, 18680 (2006).
[60] J. Wang, Q. Cai, Z. L. Li, H. K. Zhao, and R. Luo, Chem. Phys. Lett. **468**, 112 (2009).