Journal of Computational Physics 497 (2024) 112634

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Calculation of electrostatic free energy for the nonlinear Poisson-Boltzmann model based on the dimensionless potential

Shan Zhao [a,*], Idowu E. Ijaodoro [a], Mark McGowan [a], Emil Alexov [b]

[a] Department of Mathematics, University of Alabama, Tuscaloosa, AL 35487, USA
[b] Department of Physics and Astronomy, Clemson University, Clemson, SC 29634, USA

ARTICLE INFO

Dataset link: https://www.rcsb.org

Keywords:
Nonlinear Poisson-Boltzmann equation
Energy functional
Variational analysis
Electrostatic free energy
Finite difference method
Regularization

ABSTRACT

The Poisson-Boltzmann (PB) equation governing the electrostatic potential with a unit is often transformed to a normalized form for a dimensionless potential in numerical studies. To calculate the electrostatic free energy (EFE) of biological interests, a unit conversion has to be conducted, because the existing PB energy functionals are all described in terms of the original potential. To bypass this conversion, this paper proposes energy functionals in terms of the dimensionless potential for the first time in the literature, so that the normalized PB equation can be directly derived by using the Euler-Lagrange variational analysis. Moreover, alternative energy forms have been rigorously derived to avoid approximating the gradient of singular functions in the electrostatic stress term. A systematic study has been carried out to examine the surface integrals involved in alternative energy forms and their dependence on finite domain size and mesh step size, which leads to a recommendation on the EFE forms for efficient computation of protein systems. The calculation of the EFE in the regularization formulation, which is an analytical approach for treating singular charge sources of the PB equation, has also been studied. The proposed energy forms have been validated by considering smooth dielectric settings, such as diffuse interface and super-Gaussian, for which the EFE of the nonlinear PB model is found to be significantly different from that of the linearized PB model. All proposed energy functionals and EFE forms are designed such that the dimensionless potential can be simply plugged in to compute the EFE in the unit of kcal/mol, and they can also be applied in the classical sharp interface PB model.

1. Introduction

Electrostatic analysis is of central importance in the atomic scale description of solvated biomolecular processes, by studying interactions between charged solutes such as proteins, DNAs and RNAs, and mobile ions contained in the solvent. The Poisson-Boltzmann (PB) model [4,23,38] is one of the most popular implicit solvent models, in which the solute and solvent are treated as dielectric continuum and the mobile ions are assumed to follow the Boltzmann distribution in water. Based on rigorous biophysical theories, the PB theory formulates the electrostatic potential via a nonlinear elliptic partial differential equation (PDE), known as the nonlinear PB (NPB) equation [4,23,38]. The NPB equation can be linearized by replacing the nonlinear term with the leading term in the Taylor series expansion, giving rise to the linearized PB (LPB) equation.

* Corresponding author.
E-mail address: szhao@ua.edu (S. Zhao).

https://doi.org/10.1016/j.jcp.2023.112634
Received 7 August 2023; Received in revised form 8 November 2023; Accepted 8 November 2023

Available online 14 November 2023
0021-9991/

A modern approach for deriving the PB equation is using the variational analysis of a properly defined electrostatic energy functional [32,38]. For a full nonlinear PB model, this energy functional includes terms for interaction energy between charges, electrostatic stress, and excess osmotic pressure of the mobile ions [4,38]. Mathematically, the energy optimization can be accomplished with the Euler-Lagrange variation, which yields the classical PB equation. In the classical setting, a sharp interface is commonly assumed at the solute-solvent boundary, which splits the domain into two subdomains, i.e., the solute region with a low dielectric constant and the water region with a high dielectric constant. The variational analysis provides a physically sound and mathematically convenient approach for generalizing the PB model to incorporate more physical insights, through simply modifying terms or introducing new terms in the energy functional or coupling the polar energy with nonpolar energy in solvation analysis [11,12,16,24,42,45–47,49]. Many of such improved PB models feature a smooth solute-solvent boundary, i.e., the dielectric function varies smoothly from the macromolecule to the solvent over a narrow band [11,12,16,42,46,47,49]. When the solute and solvent are still assumed to be two homogeneous dielectric media away from the smooth solute-solvent boundary, the corresponding PB models can be referred to as diffuse interface PB models [1,5,6,13,16,36,39,49]. Furthermore, many heterogeneous dielectric PB models [7,21,28,29,35] have been developed to represent the dielectric distribution of the solute as a space dependent function. In particular, this work will use the super-Gaussian PB model [21] to illustrate the proposed energy functionals, in which super-Gaussian dielectric distributions are assumed for the solute to mimic the effect of random conformational changes of the macromolecule on the solvation free energy. The super-Gaussian PB model allows one to capture an ensemble averaged electrostatic free energy by using a single protein structure [35].

Numerous studies have been devoted to developing accurate, efficient, and robust numerical algorithms for solving the PB equation in the literature [8,9,15,17,19,22,23,25,27,31,33,38,43,44]. For the convenience of numerical algorithm development, the PB equation governing the electrostatic potential with a unit is commonly transformed to a normalized form for a dimensionless potential [2,8,19,20,22,26]. In order to calculate the electrostatic free energy, after solving the normalized PB equation, one has to convert the dimensionless potential back to the original potential, and calculate the free energy according to the original energy functional. Such a process is inconvenient and may introduce additional numerical artifact when improper unit conversion coefficients are used. It is highly desired if the energy functional is formulated in terms of the dimensionless potential. However, to the best of our acknowledge, no energy functional in terms of the dimensionless potential has been introduced in the PB literature.

The main objective of this work is to formulate an energy functional in terms of the dimensionless potential, so that the normalized PB equation can be directly derived by using the Euler-Lagrange equation. Moreover, by clarifying all necessary parameters and their units, we will present energy forms that can be used to compute the electrostatic free energy directly based on the numerically obtained dimensionless potential. Note that the electrostatic free energy measures the difference in the polar solvation energy of the macromolecule between the water phase and vacuum phase [4,23]. Thus, by definition, the electrostatic free energy can be calculated based on the energy functional of the potential in water phase minus that of the potential in vacuum phase. However, in practice, the electrostatic free energy is never computed via the definition. Instead, alternative energy forms without the electrostatic stress term are commonly employed, which avoids approximating the gradient of singular functions in the electrostatic stress term. These alternative energy terms could be obtained by using an integral form of Gauss's law [3,37] or using an equivalent optimization form [32]. In the present study, we will rigorously derive the alternative energy form in the dimensionless potential and justify its numerical advantages.

With the proposed free energy functionals, we will investigate the dependence of each energy term on the domain size and mesh step size. In particular, the alternative free energy form will involve surface integrals when the underlying domain is a finite one. Such surface integrals are often neglected in most computations, but are taken into account in some studies [32]. An associated issue here is that the PB model is originally defined over an infinite domain with a vanishing electrostatic potential at infinity. Nevertheless, in most grid based numerical algorithms, including finite element methods [17,22,44] and finite difference methods [8,15,19,23,25,27,33,38], a finite domain has to be used, which requires an appropriate boundary condition. For most protein studies, a Dirichlet boundary condition is commonly assumed, which provides a fairly accurate Debye-Huckel approximation to the analytical solution of the LPB potential [22]. In the present study, we will conduct a systematic study on volume and surface integrals in the free energy forms to recommend that whether the surface integrals shall be kept in computing the electrostatic free energy over a finite domain.

The singular charge source of the PB equation is known to be a significant challenge for both mathematical analysis and numerical computation of the PB model [10,22,44]. In the PB model, the partial charges carried by the macromolecule are expressed as Dirac delta distributions, and their summation forms a singular source term. In the conventional finite difference method, a trilinear method is commonly used by distributing the singular charges to the neighboring grid nodes with finite values [34]. The trilinear method is known to be inaccurate, because one essentially approximates an unbounded potential solution by finite numerical numbers. A modern approach for treating PB charge singularity is using regularization methods, in which the singularity is analytically captured via the Green's functions [10,14,19,20,44,50]. Recently, a unified regularization approach has been developed for treating singular charges in the diffuse interface PB model [36,39,40] and heterogeneous dielectric PB model [41]. In the present study, we will investigate how the singularities in the potential solution could affect the calculation of the electrostatic free energy. Moreover, the calculation of the electrostatic free energy in the regularization formulation will be discussed. Because the interaction energy between charges can be calculated using the reaction field potential, instead of the original potential, the proposed free energy form in the regularization could be free of singularities.

In this work, the PB model and its normalized form will be presented in a general setting, such that the proposed energy forms can be generalized to other PB models. Even though the proposed energy forms will be validated primarily by using the super-Gaussian PB model, they can be applied to other PB models, including sharp interface PB model, diffuse interface PB model, and so on. In

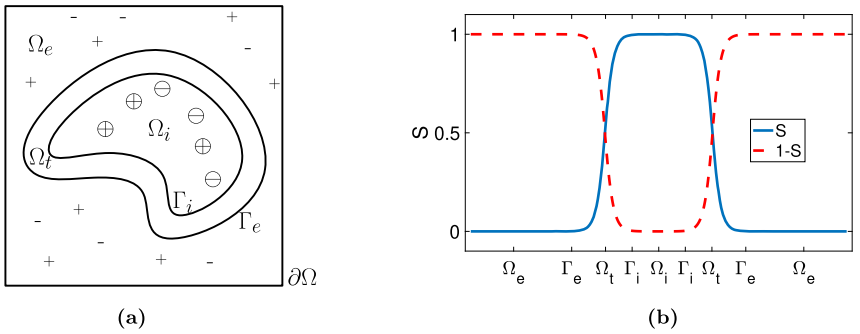

Fig. 1. (a). The subdomain setting of the solute-solvent system. (b). The subdomains are characterized by a diffuse interface function $S(\mathbf{r})$, which is plotted along a straight line. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

particular, the parameters and units of each energy term will be clarified. It should be no problem for readers to generalize the present energy forms to other PB models based on the dimensionless potential.

The rest of the paper is structured with the following sections. In Section 2, we first propose the energy functional in terms of the dimensionless potential for both NPB and LPB models. Variational analysis is conducted to derive the corresponding NPB and LPB equations. Next, the energy forms for calculating the electrostatic free energy will be derived, and regularized versions will be offered. We then provide details of the PB models to be tested and numerical algorithms. In Section 3, the proposed energy forms will be validated by considering both trilinear and regularization methods for singular charges, and by studying both LPB and NPB models. Macromolecules and proteins will be considered for evaluating volume and surface integrals. Finally, in Section 4, a brief discussion is offered.

2. Theory and algorithm

2.1. Biological problem and mathematical setting

Consider a solute macromolecule, for example, a protein being immersed in an aqueous solvent. Assume that the bulk concentrations of both univalent positive and negative ions are the same, and follow the Boltzmann distribution in the solvent. In many implicit solvent models, a smooth solute-solvent boundary is often assumed to separate the solute and solvent regions. Define a large enough domain $\Omega \subset \mathbb{R}^3$ containing this solute-solvent system, which consists of three regions: an interior domain Ω_i for the solute, an exterior domain Ω_e for the solvent, and a transition layer Ω_t in between Ω_i and Ω_e as the smooth solute-solvent boundary. Denote the interface between Ω_i and Ω_t as Γ_i, while that between Ω_t and Ω_e as Γ_e. An illustration of subdomains is shown in Fig. 1 (a). In the present paper, we assume that a diffuse interface or smooth surface function $S(\mathbf{r})$ has been defined to characterize the subdomains, i.e., $S(\mathbf{r})$ equals to one and zero, respectively, in Ω_i and Ω_e, while in Ω_t, as one travels from the interior protein to the exterior solvent, $S(\mathbf{r})$ decays from one to zero. See Fig. 1 (b) for an illustration of a $S(\mathbf{r})$ function. Based on $S(\mathbf{r})$, a smooth dielectric function $\epsilon(\mathbf{r})$ can be defined over the entire domain Ω, for example, by using the diffuse interface [39] and super-Gaussian [21,41] models. The details on the definition of $S(\mathbf{r})$ and $\epsilon(\mathbf{r})$ for different dielectric models will be offered later in this section. Without the loss of generality, we assume that both $S(\mathbf{r})$ and $\epsilon(\mathbf{r})$ are at least C^2 continuous over the entire domain Ω. Note that both $S(\mathbf{r})$ and $\epsilon(\mathbf{r})$ are dimensionless.

2.2. Electrostatic energy functionals

It is well known that the Poisson-Boltzmann (PB) equation can be derived based on variational analysis of an energy functional. In the PB literature [3,32,37,46], the electrostatic energy is usually defined as a functional of the electrostatic potential ϕ with a unit, e.g., $e_c/$Å, where e_c is the fundamental charge and Å is the angstrom. On the other hand, in mathematical and numerical studies of the PB equation, a dimensionless potential u is commonly used, where $u = \dfrac{e_c\phi}{k_B T}$ [22]. Here k_B is the Boltzmann constant and T is the temperature. Note that the unit of $e_c\phi$ is $e_c^2/$Å. We know that $k_B T = 0.5961634386$ kcal/mol and $e_c^2/$Å $= 332.06364$ kcal/mol [22]. Hence, u is a dimensionless quantity without a unit. Physically speaking, the numerator $e_c\phi$ and denominator $k_B T$ in the definition of u can be regarded as an electrostatic energy and thermal energy, respectively. Thus, from the physics point of view, u is an energy fraction, and is not an electrostatic potential. Nevertheless, we will still call u as the dimensionless potential in this work, following the mathematical studies in the PB literature [22].

To the best of our knowledge, no PB energy functional has been prescribed in terms of the dimensionless potential u in the literature. Thus, after solving u, one has to convert u back to ϕ for calculating the energy, which may introduce some numerical artifacts when the unit constants are not chosen properly. To avoid such problems, we propose the following PB energy functional in terms of u,

$$E[u] := \int_\Omega \frac{k_B T}{e_c} \sum_{j=1}^{N_m} q_j e_c \delta(\mathbf{r}-\mathbf{r}_j) u(\mathbf{r}) d\mathbf{r} - \frac{1}{4\pi} \int_\Omega \frac{(k_B T)^2}{e_c^2} \frac{\epsilon(\mathbf{r})}{2} |\nabla u(\mathbf{r})|^2 d\mathbf{r} - \frac{1}{4\pi} \int_\Omega \frac{(k_B T)^2}{e_c^2} (1 - S(\mathbf{r}))\kappa^2 (\cosh(u(\mathbf{r})) - 1) d\mathbf{r}. \quad (1)$$

The first term of Eq. (1) represents the polar energy of inserting the partial charges of the solute into the electrostatic potential [3,37]. Here, we assume the protein contains N_m atoms, with partial charges q_j in terms of the fundamental charge e_c located at the atom centers $\mathbf{r}_j$ for $j = 1, 2, \ldots, N_m$, and $\delta(\mathbf{r}-\mathbf{r}_j)$ is the Dirac delta function. The second term is the electrostatic stress, or energy of polarization for the dielectric medium [3,37]. The third term is the excess osmotic pressure of the mobile ions, or the energy of the mobile counterion distribution [3,37]. Here κ is the modified Debye-Hückel parameter with $\kappa^2 = \left(\frac{2N_A e_c^2}{100 k_B T}\right) I = 8.486902807 \text{Å}^{-2} I$, where N_A is the Avogadro's Number and the dimensionless number I is the molar strength of the ionic solution [22]. Thus, the unit of κ^2 is Å^{-2}.

It is noted that the energy form (1) is designed such that one can simply plug in the numerical potential u to calculate $E[u]$ in the unit of kcal/mol without additional conversions. To see this, Eq. (1) can be rewritten into an equivalent form

$$E[u] := k_B T \int_\Omega \sum_{j=1}^{N_m} q_j \delta(\mathbf{r}-\mathbf{r}_j) u(\mathbf{r}) d\mathbf{r} - \frac{k_B T}{4\pi} \frac{k_B T}{e_c^2/\text{Å}} \int_\Omega \frac{1}{\text{Å}} \frac{\epsilon(\mathbf{r})}{2} |\nabla u(\mathbf{r})|^2 d\mathbf{r} - \frac{k_B T}{4\pi} \frac{k_B T}{e_c^2/\text{Å}} \int_\Omega \frac{1}{\text{Å}} (1 - S(\mathbf{r}))\kappa^2 (\cosh(u(\mathbf{r})) - 1) d\mathbf{r}. \quad (2)$$

In all integrals of Eq. (2), $\delta(\mathbf{r}-\mathbf{r}_j)$, $|\nabla u|^2$, and κ^2 has a unit of Å^{-3}, Å^{-2}, and Å^{-2}, respectively. Thus all integrals are dimensionless after integration. The fraction $\frac{k_B T}{e_c^2/\text{Å}}$ is also dimensionless. Hence, the units of all terms are determined by $k_B T = 0.5961634386$ kcal/mol. The scaling of the constant $1/4\pi$ in second and third terms is to conform to the usual electrostatic convention [37].

When the dimensionless potential u is weak, a linearized energy functional can be used

$$E[u] := \int_\Omega \frac{k_B T}{e_c} \sum_{j=1}^{N_m} q_j e_c \delta(\mathbf{r}-\mathbf{r}_j) u(\mathbf{r}) d\mathbf{r} - \frac{1}{4\pi} \int_\Omega \frac{(k_B T)^2}{e_c^2} \frac{\epsilon(\mathbf{r})}{2} |\nabla u(\mathbf{r})|^2 d\mathbf{r} - \frac{1}{4\pi} \int_\Omega \frac{(k_B T)^2}{e_c^2} (1 - S(\mathbf{r}))\kappa^2 \frac{u^2(\mathbf{r})}{2} d\mathbf{r}, \quad (3)$$

which is obtained by approximating $(\cosh(u) - 1)$ in Eq. (1) by its leading term in the Taylor series expansion, i.e., $u^2/2$.

In order to calculate the free energy, the electrostatic energy of the same protein system in the vacuum state is needed. As in other PB models, we can assume the partial charges carried by the protein and the function $S(\mathbf{r})$ characterizing the molecular surface being the same in the vacuum state. Denote the dielectric function in the vacuum state as $\epsilon_v(\mathbf{r})$. For both the diffuse interface [39] and super-Gaussian [21,41] models, we have $\epsilon_v(\mathbf{r}) = \epsilon(\mathbf{r})$ in Ω_i and $\epsilon_v(\mathbf{r}) = 1$ in Ω_e. Again, $\epsilon_v(\mathbf{r})$ is dimensionless and is at least C^2 continuous over Ω. Denote the dimensionless potential as v in the vacuum state. The energy functional in the vacuum is given as

$$E[v] := \int_\Omega \frac{k_B T}{e_c} \sum_{j=1}^{N_m} q_j e_c \delta(\mathbf{r}-\mathbf{r}_j) v(\mathbf{r}) d\mathbf{r} - \frac{1}{4\pi} \int_\Omega \frac{(k_B T)^2}{e_c^2} \frac{\epsilon_v(\mathbf{r})}{2} |\nabla v(\mathbf{r})|^2 d\mathbf{r}. \quad (4)$$

Because no mobile ions are involved in the vacuum, the excess osmotic pressure term is not presented in Eq. (4), while the other two terms are formulated similarly as in Eq. (1).

2.3. Variational analysis and governing equations

The governing equation of the dimensionless potential u for the present solute-solvent system can be derived based on the variational analysis. In particular, the Euler-Lagrange equation of the energy functional $E[u]$ in Eq. (1) gives rise to

$$-\frac{1}{4\pi} \frac{(k_B T)^2}{e_c^2} \nabla \cdot (\epsilon \nabla u) + \frac{1}{4\pi} \frac{(k_B T)^2}{e_c^2} (1 - S)\kappa^2 \sinh(u) = \frac{k_B T}{e_c} \sum_{j=1}^{N_m} q_j e_c \delta(\mathbf{r}-\mathbf{r}_j). \quad (5)$$

Through a proper rescaling for both left and right sides of Eq. (5), we arrive at the commonly used nonlinear PB (NPB) equation for the dimensionless potential u [21,41]

$$-\nabla \cdot (\epsilon(\mathbf{r}) \nabla u(\mathbf{r})) + (1 - S(\mathbf{r}))\kappa^2 \sinh(u(\mathbf{r})) = 4\pi \frac{e_c^2}{k_B T} \sum_{j=1}^{N_m} q_j \delta(\mathbf{r}-\mathbf{r}_j), \quad \mathbf{r} \in \Omega. \quad (6)$$

Similarly, the variational analysis of the linearized functional in Eq. (3) yields the linearized PB (LPB) equation [21,41]

$$-\nabla \cdot (\epsilon(\mathbf{r}) \nabla u(\mathbf{r})) + (1 - S(\mathbf{r}))\kappa^2 u(\mathbf{r}) = 4\pi \frac{e_c^2}{k_B T} \sum_{j=1}^{N_m} q_j \delta(\mathbf{r}-\mathbf{r}_j), \quad \mathbf{r} \in \Omega. \quad (7)$$

Finally, in the vacuum state, the energy functional $E[v]$ leads to a Poisson equation for v [21,41]

$$-\nabla \cdot (\epsilon_v(\mathbf{r}) \nabla v(\mathbf{r})) = 4\pi \frac{e_c^2}{k_B T} \sum_{j=1}^{N_m} q_j \delta(\mathbf{r}-\mathbf{r}_j), \quad \mathbf{r} \in \Omega. \quad (8)$$

In numerical solution of the PB equations, a Dirichlet boundary condition is commonly used on the outer boundary of the domain Ω, i.e., $\partial\Omega$. Assume that the domain Ω is large enough, so that the dielectric function $\epsilon(\mathbf{r})$ equals to a constant ϵ_{out} on $\partial\Omega$. The following Dirichlet boundary data can be assumed [22]

$$b(\mathbf{r};\epsilon_{out}) := \frac{{e_c}^2}{k_B T}\sum_{j=1}^{N_m}\frac{q_j}{\epsilon_{out}|\mathbf{r}-\mathbf{r}_j|}e^{-|\mathbf{r}-\mathbf{r}_j|\sqrt{\frac{\kappa^2}{\epsilon_{out}}}}, \quad \text{on } \partial\Omega. \tag{9}$$

For a sufficiently large domain, $b(\mathbf{r};\epsilon_{out})$ provides a fairly accurate Debye-Huckel approximation to the analytical solution of the LPB potential on $\partial\Omega$ [22]. In the present study, both the NPB equation (6) and LPB equation (7) will be paired with a Dirichlet boundary condition for $u(\mathbf{r})$ with $\epsilon_{out} = 80$, while the Poisson equation is paired with a condition for $v(\mathbf{r})$ with $\epsilon_{out} = 1$

$$u(\mathbf{r}) = u_b(\mathbf{r}) := b(\mathbf{r};80) \quad \text{and} \quad v(\mathbf{r}) = v_b(\mathbf{r}) := b(\mathbf{r};1), \quad \mathbf{r}\in\partial\Omega. \tag{10}$$

Mathematically, since the charge sources are singular, i.e., $\delta(\mathbf{r}-\mathbf{r}_j)\to\infty$ as $\mathbf{r}$ approaches to $\mathbf{r}_j$ for all atomic centers, the potentials of the NPB equation (6), LPB equation (7), and Poisson equation (8) all blow up at charge centers [41]. This singularity is a well known numerical difficulty in solving the PB equation [20,41]. The impact of charge singularities to the free energy calculation will be discussed later.

2.4. Electrostatic free energy

After solving either the NPB or LPB equation for u and Poisson's equation for v, one can then calculate the electrostatic free energy based on the definition $\Delta E = E[u] - E[v]$. By directly using the above energy functionals, the electrostatic free energy of the NPB model can be defined as

$$\begin{aligned}
\Delta E =& E[u] - E[v] \\
=& \int_\Omega k_B T\sum_{j=1}^{N_m} q_j\delta(\mathbf{r}-\mathbf{r}_j)(u(\mathbf{r})-v(\mathbf{r}))d\mathbf{r} - \frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\int_\Omega\left(\frac{\epsilon(\mathbf{r})}{2}|\nabla u(\mathbf{r})|^2 - \frac{\epsilon_v(\mathbf{r})}{2}|\nabla v(\mathbf{r})|^2\right)d\mathbf{r} \\
& - \frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\int_\Omega (1-S(\mathbf{r}))\kappa^2(\cosh(u(\mathbf{r}))-1)d\mathbf{r}, \\
=& k_B T\sum_{j=1}^{N_m} q_j(u(\mathbf{r}_j)-v(\mathbf{r}_j)) - \frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\int_\Omega\left(\frac{\epsilon(\mathbf{r})}{2}|\nabla u(\mathbf{r})|^2 - \frac{\epsilon_v(\mathbf{r})}{2}|\nabla v(\mathbf{r})|^2\right)d\mathbf{r} \\
& - \frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\int_\Omega (1-S(\mathbf{r}))\kappa^2(\cosh(u(\mathbf{r}))-1)d\mathbf{r},
\end{aligned} \tag{11}$$

in which the property of the Dirac delta function has been applied to evaluate the first integral.

However, Eq. (11) is seldom used in practice. Theoretically, $u(\mathbf{r}_j)$ and $v(\mathbf{r}_j)$ go to infinity for any charge center $\mathbf{r}_j\in\Omega_i$. Numerically, the approximated values of $u(\mathbf{r}_j)$ and $v(\mathbf{r}_j)$ generated by a numerical algorithm are still finite, but involve huge approximation errors. When the same algorithm and the same numerical grid are employed to solve u and v, it is known [3,41] that errors in approximating singular u and v are almost the same at $\mathbf{r}_j$, so that such errors can be canceled in calculating $u(\mathbf{r}_j) - v(\mathbf{r}_j)$ in Eq. (11). Consequently, the approximated value for the first term of Eq. (11) is still acceptable, and this is known as the reduction of artificial grid energy in the PB literature [3,41]. Nevertheless, the approximations in the second term are more troublesome, because the gradients of the singular functions u and v need to be approximated near charge centers in the electrostatic stress term.

To avoid using the second term in Eq. (11) for calculating ΔE, alternative energy forms are usually considered in the literature, such as using an integral form of Gauss's law [3,37] or using an equivalent variational form [32]. In the present study, we will propose an alternative energy form in terms of the dimensionless potentials u and v, which is computationally better than Eq. (11). Also, a rigorous mathematical derivation is offered below. To this end, we multiply both hand sides of the NPB equation (5) by u and integrate over the entire domain Ω. We then have

$$\begin{aligned}
& \int_\Omega \frac{k_BT}{e_c}\sum_{j=1}^{N_m} q_j e_c\delta(\mathbf{r}-\mathbf{r}_j)u(\mathbf{r})d\mathbf{r} \\
=& -\frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\int_\Omega \nabla\cdot(\epsilon\nabla u)u d\mathbf{r} + \frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\int_\Omega (1-S)\kappa^2 u\sinh(u)d\mathbf{r} \\
=& \frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\int_\Omega (1-S)\kappa^2 u\sinh(u)d\mathbf{r} - \frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\left[\int_\Omega u d(\epsilon\nabla u)\right]
\end{aligned}$$

$$=\frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\int_\Omega (1-S)\kappa^2 u\sinh(u)d\mathbf{r}-\frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\left[\int_{\partial\Omega} u\epsilon\nabla u\cdot\mathbf{n}ds-\int_\Omega \epsilon\nabla u\cdot\nabla u d\mathbf{r}\right],$$

$$=\frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\int_\Omega (1-S)\kappa^2 u\sinh(u)d\mathbf{r}+\frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\int_\Omega \epsilon|\nabla u|^2 d\mathbf{r}-\frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\int_{\partial\Omega}\epsilon u\frac{\partial u}{\partial n}ds, \tag{12}$$

in which the divergence theorem has been applied to produce a surface integral on $\partial\Omega$. Here $\mathbf{n}$ is the outward normal direction of $\partial\Omega$, and $\frac{\partial u}{\partial n}$ is the directional derivative along the normal direction. By expressing the electrostatic stress term in terms of the others, Eq. (12) can be rewritten as

$$\frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\int_\Omega \frac{\epsilon}{2}|\nabla u|^2 d\mathbf{r}=\int_\Omega \frac{k_BT}{e_c}\sum_{j=1}^{N_m}\frac{1}{2}q_j e_c\delta(\mathbf{r}-\mathbf{r}_j)u(\mathbf{r})d\mathbf{r}$$
$$-\frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\int_\Omega \frac{1}{2}(1-S)\kappa^2 u\sinh(u)d\mathbf{r}+\frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\int_{\partial\Omega}\frac{1}{2}\epsilon u\frac{\partial u}{\partial n}ds. \tag{13}$$

By plugging Eq. (13) into Eq. (1), we can eliminate the electrostatic stress term in the alternative form of the NPB energy functional

$$E[u]=\frac{k_BT}{e_c}\frac{1}{2}\int_\Omega\sum_{j=1}^{N_m}q_j e_c\delta(\mathbf{r}-\mathbf{r}_j)u(\mathbf{r})d\mathbf{r}-\frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\int_\Omega (1-S(\mathbf{r}))\kappa^2(\cosh(u(\mathbf{r}))-1)d\mathbf{r}$$
$$+\frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\frac{1}{2}\int_\Omega (1-S)\kappa^2 u\sinh(u)d\mathbf{r}-\frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\frac{1}{2}\int_{\partial\Omega}\epsilon u\frac{\partial u}{\partial n}ds, \tag{14}$$

Similarly, we multiply both hand sides of the Poisson equation (8) by v and integrate over the entire domain Ω. This yields the following expression for the electrostatic stress term

$$\frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\int_\Omega \frac{\epsilon_v}{2}|\nabla v|^2 d\mathbf{r}=\int_\Omega \frac{k_BT}{e_c}\sum_{j=1}^{N_m}\frac{1}{2}q_j e_c\delta(\mathbf{r}-\mathbf{r}_j)v(\mathbf{r})d\mathbf{r}+\frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\int_{\partial\Omega}\frac{1}{2}\epsilon_v v\frac{\partial v}{\partial n}ds. \tag{15}$$

By substituting Eq. (15) into Eq. (4), the energy functional $E[v]$ can be rewritten as

$$E[v]=\frac{k_BT}{e_c}\frac{1}{2}\int_\Omega\sum_{j=1}^{N_m}q_j e_c\delta(\mathbf{r}-\mathbf{r}_j)v(\mathbf{r})d\mathbf{r}-\frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\frac{1}{2}\int_{\partial\Omega}\epsilon_v v\frac{\partial v}{\partial n}ds. \tag{16}$$

Therefore, the electrostatic free energy $\Delta E=E[u]-E[v]$ for the NPB model can be calculated as

$$\Delta E=\frac{1}{2}k_BT\sum_{j=1}^{N_m}q_j(u(\mathbf{r}_j)-v(\mathbf{r}_j))-\frac{1}{4\pi}\frac{(k_BT)^2}{e_c^2}\int_{\Omega_i^C}(1-S)\kappa^2(\cosh(u)-1)d\mathbf{r}$$
$$+\frac{1}{8\pi}\frac{(k_BT)^2}{e_c^2}\int_{\Omega_i^C}(1-S)\kappa^2 u\sinh(u)d\mathbf{r}-\frac{1}{8\pi}\frac{(k_BT)^2}{e_c^2}\int_{\partial\Omega}\left(\epsilon u\frac{\partial u}{\partial n}-\epsilon_v v\frac{\partial v}{\partial n}\right)ds. \tag{17}$$

In the present study, because $1-S=0$ in Ω_i, we have replaced the integral domain Ω of two volume integrals in Eq. (17) by the complement set of Ω_i, i.e., $\Omega_i^C=\Omega_t\cup\Omega_e$. Thus, two volume integrals are no longer singular, because the potential u is a smooth function away from singular charges or outside Ω_i. The only singular term in (17) is the first term. Nevertheless, the artificial grid energy will be reduced in evaluating $u(\mathbf{r}_j)-v(\mathbf{r}_j)$, as mentioned above. Therefore, the electrostatic free energy ΔE calculated by Eq. (17) will be more accurate than that by Eq. (11). The similar derivation can be carried out for the LPB model, and the corresponding electrostatic free energy is given as

$$\Delta E=\frac{1}{2}k_BT\sum_{j=1}^{N_m}q_j(u(\mathbf{r}_j)-v(\mathbf{r}_j))-\frac{1}{8\pi}\frac{(k_BT)^2}{e_c^2}\int_{\partial\Omega}\left(\epsilon u\frac{\partial u}{\partial n}-\epsilon_v v\frac{\partial v}{\partial n}\right)ds. \tag{18}$$

We note that unlike the original form Eq. (11), the alternative free energy forms Eq. (17) and Eq. (18) involve surface integrals.

2.5. Regularization formulation

In traditional numerical studies, the singular charge sources involved in the governing equations (6), (7), and (8) are directly discretized in grid based computations. In the finite difference algorithm, a trilinear method is commonly used, in which the singular

charges are distributed to the neighboring grid nodes with finite values [34]. However, the trilinear method is doomed to be inaccurate, because one actually attempts to approximate an unbounded potential function by finite numerical values. A modern approach to analytically handle the singular charges is using the regularization methods, which were first introduced for solving two-dielectric PB models with sharp interfaces [10,14,20,26,50]. For PB models with diffuse interface or heterogeneous dielectric functions, a breakthrough has been reported recently by introducing a unified regularization approach [36,39–41]. A brief introduction to such a regularization will be given in this subsection, and we refer the readers to the original works [36,39–41] and the reference therein for more details.

In [36,39–41], a two-component decomposition is considered for the potential u of the NPB equation (6)

$$u(\mathbf{r}) = u_C(\mathbf{r}) + u_{RF}(\mathbf{r}), \quad \mathbf{r} \in \Omega, \tag{19}$$

where u_C is the Coulomb potential and u_{RF} is the reaction field potential. To capture the singularity in the potential function, the Coulomb potential u_C is assumed to satisfy a homogeneous Poisson's equation with the same singular source over R^3,

$$\begin{cases} -\epsilon_m \Delta u_C(\mathbf{r}) = 4\pi \frac{e_c^2}{k_B T} \sum_{j=1}^{N_m} q_j \delta(\mathbf{r} - \mathbf{r}_j) & \text{in } R^3, \\ u_C(\mathbf{r}) = 0 & \text{as } |\mathbf{r}| \to \infty, \end{cases} \tag{20}$$

where ϵ_m is the dielectric constant of the molecule. For diffuse interface PB models [36,39,40], we have $\epsilon(\mathbf{r}) = \epsilon_m$ for $\mathbf{r} \in \Omega_i$, while for super-Gaussian dielectric model [21,41], ϵ_m is the minimal value of $\epsilon(\mathbf{r})$ in Ω_i. Equation (20) has a fundamental solution analytically, i.e., the singular component u_C is actually the Green's function $G(\mathbf{r})$

$$u_C(\mathbf{r}) = G(\mathbf{r}) := \frac{e_c^2}{k_B T} \sum_{j=1}^{N_m} \frac{q_j}{\epsilon_m |\mathbf{r} - \mathbf{r}_j|}. \tag{21}$$

Through a rigorous mathematical analysis, the reaction field potential can be shown to satisfy a regularized NPB equation for both diffuse interface and super-Gaussian models [36,39–41]

$$\begin{cases} -\nabla \cdot (\epsilon(\mathbf{r}) \nabla u_{RF}(\mathbf{r})) + (1 - S(\mathbf{r}))\kappa^2 \sinh(u_{RF}(\mathbf{r}) + G(\mathbf{r})) = \nabla \epsilon(\mathbf{r}) \cdot \nabla G(\mathbf{r}) & \text{in } \Omega, \\ u_{RF}(\mathbf{r}) = u_b(\mathbf{r}) - G(\mathbf{r}) & \text{on } \partial\Omega. \end{cases} \tag{22}$$

Essentially, the singular charge source is replaced by a new source $\nabla\epsilon \cdot \nabla G$, which is bounded and smooth in Ω under proper conditions [41]. Here, the gradient of Green's function is also analytically known

$$\nabla G(\mathbf{r}) = -\frac{e_c^2}{k_B T} \sum_{j=1}^{N_m} \frac{q_j(\mathbf{r} - \mathbf{r}_j)}{\epsilon_m |\mathbf{r} - \mathbf{r}_j|^3}, \tag{23}$$

while the gradient of $\epsilon(\mathbf{r})$ shall be calculated numerically. The solution existence and uniqueness of the regularized NPB model have been proved in [41].

In the previous works of the regularization for PB models with diffuse interface or heterogeneous dielectric functions [36,39–41], the numerical solution of the regularized NPB equation (22) has not been studied before. Only the linearized versions have been explored. For the LPB equation (7), we consider the same decomposition $u = u_{RF} + u_C = u_{RF} + G$. Then the reaction field potential u_{RF} satisfies a regularized LPB equation

$$\begin{cases} -\nabla \cdot (\epsilon(\mathbf{r}) \nabla u_{RF}(\mathbf{r})) + (1 - S(\mathbf{r}))\kappa^2 u_{RF}(\mathbf{r}) = \nabla \epsilon(\mathbf{r}) \cdot \nabla G(\mathbf{r}) - (1 - S(\mathbf{r}))\kappa^2 G(\mathbf{r}), & \text{in } \Omega, \\ u_{RF}(\mathbf{r}) = u_b(\mathbf{r}) - G(\mathbf{r}) & \text{on } \partial\Omega. \end{cases} \tag{24}$$

In the solute domain Ω_i, $(1 - S)\kappa^2 G = 0$, while in other subdomains, G is well-defined. Thus, the second source term of (24), i.e., $(1 - S)\kappa^2 G$, is smooth over the entire domain Ω.

For the super-Gaussian PB model [41], a similar potential decomposition is carried out in the vacuum state for the Poisson equation (8)

$$v(\mathbf{r}) = v_C(\mathbf{r}) + v_{RF}(\mathbf{r}) = G(\mathbf{r}) + v_{RF}(\mathbf{r}), \quad \mathbf{r} \in \Omega. \tag{25}$$

Here the singular component v_C is essentially u_C and is also analytically represented by the Green's function G. The reaction field potential satisfies a regularized Poisson equation

$$\begin{cases} -\nabla \cdot (\epsilon_v(\mathbf{r}) \nabla v_{RF}(\mathbf{r})) = \nabla \epsilon_v(\mathbf{r}) \cdot \nabla G(\mathbf{r}), & \text{in } \Omega, \\ v_{RF}(\mathbf{r}) = v_b(\mathbf{r}) - G(\mathbf{r}) & \text{on } \partial\Omega. \end{cases} \tag{26}$$

The new source $\nabla\epsilon_v \cdot \nabla G$ is also bounded and smooth under proper conditions [41].

Based on the regularization formulation [36,39–41], we propose to calculate the electrostatic free energy $\Delta E = E[u] - E[v]$ of the NPB model by using reaction field potentials in the first term

$$\Delta E = \frac{1}{2} k_B T \sum_{j=1}^{N_m} q_j (u_{RF}(\mathbf{r}_j) - \upsilon_{RF}(\mathbf{r}_j)) - \frac{1}{4\pi} \frac{(k_B T)^2}{e_c^2} \int_{\Omega_i^C} (1 - S)\kappa^2 (\cosh(u) - 1) d\mathbf{r}$$

$$+ \frac{1}{8\pi} \frac{(k_B T)^2}{e_c^2} \int_{\Omega_i^C} (1 - S)\kappa^2 u \sinh(u) d\mathbf{r} - \frac{1}{8\pi} \frac{(k_B T)^2}{e_c^2} \int_{\partial\Omega} \left(\epsilon u \frac{\partial u}{\partial n} - \epsilon_\upsilon \upsilon \frac{\partial \upsilon}{\partial n} \right) ds. \tag{27}$$

This is because the difference between two singular potentials u and υ can be calculated by the difference between two non-singular potentials u_{RF} and υ_{RF}, i.e.,

$$u(\mathbf{r}) - \upsilon(\mathbf{r}) = u_{RF}(\mathbf{r}) - \upsilon_{RF}(\mathbf{r}), \quad \mathbf{r} \in \Omega \tag{28}$$

thanks to the cancelation of the Coulomb components in u and υ. Note that the reaction field potentials u_{RF} and υ_{RF} are guaranteed to be bounded and smooth throughout the domain Ω [36,39–41], including at charge centers $\mathbf{r}_j$. Therefore, the first term of Eq. (27) is no longer singular. The other terms of Eq. (27) are the same as those in Eq. (17), and are also smooth. Therefore, the electrostatic free energy given by Eq. (27) does not involve singularity approximations. For the LPB model, Eq. (18) can be similarly transformed to

$$\Delta E = \frac{1}{2} k_B T \sum_{j=1}^{N_m} q_j (u_{RF}(\mathbf{r}_j) - \upsilon_{RF}(\mathbf{r}_j)) - \frac{1}{8\pi} \frac{(k_B T)^2}{e_c^2} \int_{\partial\Omega} \left(\epsilon u \frac{\partial u}{\partial n} - \epsilon_\upsilon \upsilon \frac{\partial \upsilon}{\partial n} \right) ds, \tag{29}$$

for calculating the electrostatic free energy in the regularization formulation.

2.6. Electrostatic free energy calculations

We have proposed four equations for calculating the electrostatic free energy ΔE, including Eq. (18) and Eq. (29) for the LPB model, and Eq. (17) and Eq. (27) for the NPB model. Among them, two are based on the regularization formulation, i.e., Eq. (29) and Eq. (27). The other two, i.e., Eq. (18) and Eq. (17), will be called trilinear forms, because the singular charges will be discretized by the trilinear method in these two forms.

Several research issues will be explored in the present study for calculating the electrostatic free energy. First, we will examine all individual terms of four free energy forms for different domain size and mesh step size. In particular, for all free energy forms derived in this work, surface integrals for u and υ are presented. These surface integrals are usually neglected in the previous PB studies. In this work, we will explore if the surface integral shall be kept, especially when the domain Ω is not large enough. In practice, a small domain size is usually preferred for saving computational time.

We can explore the surface integrals by mainly considering the trilinear method, while the outcome of the regularization method should be similar. In particular, for ΔE of the LPB model, we will show numerically that the surface integral terms $\frac{1}{8\pi} \frac{(k_B T)^2}{e_c^2} \int_{\partial\Omega} \left(\epsilon u \frac{\partial u}{\partial n} - \epsilon_\upsilon \upsilon \frac{\partial \upsilon}{\partial n} \right) ds$ should be dropped in Eqs. (18) and (29). This is because the first term in Eq. (18) or Eq. (29) does not depend on the size of domain Ω. In other words, the size of domain Ω determines the surface integral values, but will not directly affect the first term. When the size of domain Ω goes to infinity, the surface integral values will be vanishing, and the limiting electrostatic free energy is solely due to the first term. Therefore, in the LPB model, we expect that ΔE calculated without surface integrals provides a better approximation to the limiting energy for a finite domain Ω.

For ΔE of the NPB model, the same argument applies to the surface integral of υ, i.e., $\frac{1}{8\pi} \frac{(k_B T)^2}{e_c^2} \int_{\partial\Omega} \epsilon_\upsilon \upsilon \frac{\partial \upsilon}{\partial n} ds$ should be omitted. For the potential u, two volume integrals are involved in Eq. (17) over Ω or Ω_i^C after simplification. For different domain sizes of Ω, the volume integral values are changing, and the surface integral $\frac{1}{8\pi} \frac{(k_B T)^2}{e_c^2} \int_{\partial\Omega} \epsilon u \frac{\partial u}{\partial n} ds$ may be acted as a compensation value for truncating the infinity domain into a finite one in evaluating volume integrals. For this reason, the surface integral of u may be included in ΔE. We will numerically explore whether $\frac{1}{8\pi} \frac{(k_B T)^2}{e_c^2} \int_{\partial\Omega} \epsilon u \frac{\partial u}{\partial n} ds$ should be included in calculating ΔE for both trilinear and regularization methods.

In the second research issue, we will compare the electrostatic free energies calculated by the trilinear and regularization methods. For the trilinear method, one relies on the error cancelation in $u - \upsilon$ for suppressing the grid energy, while by treating singular charges analytically, the regularization is free of artificial grid energy. In the previous study [41], the regularization is found to produce more accurate energy estimates than the trilinear method for the super-Gaussian LPB model [41]. In the present work, we will explore the performance of the regularization for the super-Gaussian NPB model. Moreover, we will show that if the cancelation is impossible, the energy calculated by the trilinear method involves huge approximation errors.

Finally, for a set of proteins, we will compare the electrostatic free energies predicted by the LPB and NPB energy forms. In literature, there are many works devoted to solve the NPB equation, such as [2,20,47]. However, in these works, only the first term in Eq. (17) or Eq. (27) was reported as the electrostatic free energy. Therefore, it is interesting to explore the difference between the first term and the actual NPB energy form.

2.7. Diffuse interface and super-Gaussian PB models

The performance of the proposed electrostatic free energy forms will be tested by using the diffuse interface PB model [36,39,40] and super-Gaussian PB model [21,41]. In this subsection, we will provide further details for these models.

For both diffuse interface and super-Gaussian models, we first generate the smooth surface function $S(\mathbf{r})$ by the Gaussian convolution surface (GCS) introduced in [39]. Consider a protein with a total N_m atoms, with the center and radius of each atom being $\mathbf{r}_j$ and R_j, respectively. By treating each atom as a hard sphere, the Van der Waals (VdW) surface is defined as the smallest envelop enclosing the union of all spheres. By augmenting each atomic radius by 3Å, another VdW surface can be defined, which will be referred as to an extended solute-accessible surface (ESAS). The diffuse interface function $S(\mathbf{r})$ should satisfy a subdomain requirement that $S(\mathbf{r}) = 1$ inside the VdW surface and $S(\mathbf{r}) = 0$ outside the ESAS surface.

To generate the GCS, a standard solute-accessible surface (SAS) is first considered, which is a VdW surface after augmenting each atomic radius by 1.5Å. A Heaviside function is defined with $H(\mathbf{r}) = 1$ inside the SAS and $H(\mathbf{r}) = 0$ outside the SAS. One then convolutes $H(\mathbf{r})$ with a Gaussian kernel function in each Cartesian direction [39]

$$K(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right). \tag{30}$$

In real computations, the Heaviside function and Gaussian kernel are defined discretely on a uniform mesh. The discrete convolution in three-dimensions is realized through a fast Fourier transform (FFT) algorithm. After discrete convolution, the subdomain requirement is enforced in the post-processing to ensure $S(\mathbf{r}) = 1$ inside the VdW surface and $S(\mathbf{r}) = 0$ outside the ESAS surface. Based on such an interface function $S(\mathbf{r})$, one can define the solute domain Ω_i, the solvent domain Ω_e, and the transition layer Ω_t, which has a width of 3Å [39].

For the diffuse interface PB model [36,39,40], the dielectric function is defined to be

$$\epsilon(\mathbf{r}) = S(\mathbf{r})\epsilon_m + (1 - S(\mathbf{r}))\epsilon_{out}, \quad \mathbf{r} \in \Omega, \tag{31}$$

where ϵ_m and ϵ_{out} are the dielectric constants, respectively, for the molecule and outside medium. For the water state, we take $\epsilon_m = 1$ and $\epsilon_{out} = 80$. Then, $\epsilon(\mathbf{r})$ will take constant values in solute and solvent with $\epsilon(\mathbf{r}) = \epsilon_m$ in Ω_i and $\epsilon(\mathbf{r}) = \epsilon_{out}$ in Ω_e. In the smooth solute-solvent boundary Ω_t, $\epsilon(\mathbf{r})$ varies smoothly from ϵ_m to ϵ_{out}. With $\epsilon(\mathbf{r})$ and $S(\mathbf{r})$ defined, one can solve either the NPB equation (6) or LPB equation (7) together with the boundary condition Eq. (10) for u. For the vacuum state, we have $\epsilon_{out} = 1$, so that $\epsilon_v(\mathbf{r}) = 1$ throughout the domain Ω. One then solves the Poisson equation (8) with the boundary condition Eq. (10) for v. With u and v, the electrostatic free energy ΔE of the trilinear method for the LPB and NPB model can be computed, respectively, by Eq. (18) and Eq. (17).

In the regularization method of the diffuse interface PB model [36,39,40], one solves either the NPB equation (22) or LPB equation (24) for the reaction field potential u_{RF}. Because $\epsilon_v(\mathbf{r}) = 1$, one does not need to solve v_{RF}. Basically, we have $v(\mathbf{r}) = G(\mathbf{r})$ and $v_{RF}(\mathbf{r}) = 0$. By using these two conditions, Eq. (29) and Eq. (27) could be simplified for calculating the electrostatic free energy Δ, respectively, for the LPB and NPB model.

The dielectric function $\epsilon(\mathbf{r})$ is a heterogeneous function in the super-Gaussian PB model [21,41]. In particular, each atom of the protein is regarded as a "soft sphere" with a density function. For the j^{th} atom, the density at the position $\mathbf{r}$ is given as [21]

$$g_j(\mathbf{r}) = \exp\left[-\left(\frac{|\mathbf{r} - \mathbf{r}_j|^2}{R_j^2}\right)^m\right], \tag{32}$$

where R_j is the VdW radius of the j^{th} atom and the relative variance is taken as 1. We will take the order parameter $m = 2$ in this study. Next, a total density function for all atoms is defined as

$$g(\mathbf{r}) = 1 - \prod_{j=1}^{N_m}[1 - g_j(\mathbf{r})]. \tag{33}$$

In order to explicitly model the dielectric properties of protein cavities, the maximal dielectric value of the macromolecule is controlled by a parameter ϵ_{gap} in the super-Gaussian model [21], and the dielectric distribution within the protein region is defined as

$$\epsilon_{in}(\mathbf{r}) = \epsilon_m g(\mathbf{r}) + \epsilon_{gap}[1 - g(\mathbf{r})], \tag{34}$$

where the constants ϵ_m and ϵ_{gap} are the dielectric values at the atom centers and in a gap region, respectively. By combining (33) with (34), we arrive at a concise form

$$\epsilon_{in}(\mathbf{r}) = \epsilon_m + (\epsilon_{gap} - \epsilon_m)\prod_{j=1}^{N_m}[1 - g_j(\mathbf{r})]. \tag{35}$$

Since $\epsilon_{gap} > \epsilon_m$ and $g_j(\mathbf{r})$ has a range of $[0, 1]$, this means that ϵ_m and ϵ_{gap} are, respectively, the minimal and maximal dielectric values of the protein.

In the super-Gaussian model [21,41], by using the surface function $S(\mathbf{r})$ defined by the GCS model, the dielectric function in the entire domain Ω is defined as

$$\epsilon(\mathbf{r}) = S(\mathbf{r})\epsilon_{in}(\mathbf{r}) + [1 - S(\mathbf{r})]\epsilon_{out}. \tag{36}$$

Inside the protein region Ω_i, we have $S(\mathbf{r}) = 1$ so that the inhomogeneity of the super-Gaussian dielectric distribution is retained. In the exterior region Ω_e, we have $S(\mathbf{r}) = 0$. Then $\epsilon(\mathbf{r})$ becomes a constant ϵ_{out}. As $S(\mathbf{r})$ decays smoothly from one to zero over Ω_t, $\epsilon(\mathbf{r})$ will change from $\epsilon_{in}(\mathbf{r})$ to ϵ_{out}. In the present study, we take $\epsilon_m = 1$ and $\epsilon_{gap} = 8$ in the super-Gaussian model.

For the trilinear method, the energies can be calculated similarly as in the diffuse interface model. For the water state in the super-Gaussian model, we take $\epsilon_{out} = 80$, and one can solve either the NPB equation (6) or LPB equation (7) together with the boundary condition Eq. (10) for u. For the vacuum state, we have $\epsilon_{out} = 1$, but $\epsilon_v(\mathbf{r})$ is still a heterogeneous function. One then solves the Poisson equation (8) with Eq. (10) for v. With u and v, the electrostatic free energy ΔE for the LPB and NPB model of the trilinear method can be computed, respectively, by Eq. (18) and Eq. (17).

For the regularization method, one similarly solves either the NPB equation (22) or LPB equation (24) for the reaction field potential u_{RF}. Unlike the diffuse interface model, one needs to solve the Poisson equation (26) for v_{RF} in the super-Gaussian model. Then the original forms of Eq. (29) and Eq. (27) will be used, respectively, to calculate the electrostatic free energy of the LPB and NPB model.

2.8. Numerical discretization and implementation

In this subsection, we offer numerical details for solving the PB equations and for calculating the electrostatic free energies. For a given protein structure, we first determine a tight box that contains all VdW balls. An edge value is then used to extend this box in both positive and negative directions along x, y, and z axes. This selects an appropriate computational domain Ω. We next construct a uniform mesh partition with N_x, N_y, N_z being the number of the grid points in x, y, and z directions, respectively. Without the loss of generality, we assume the grid spacing h in all x, y, and z directions to be the same, i.e., $h = \Delta x = \Delta y = \Delta z$, with the unit Å. For a function f defined at a node (x_i, y_j, z_k), we denote $f_{i,j,k} = f(x_i, y_j, z_k)$.

All PB equations to be discretized in this work, including Eqs. (6), (7), (8), (22), (24), and (26), are given in the divergence form. The same central finite difference method is employed to approximate the Laplacian term. For example, at the node (x_i, y_j, z_k), the x derivative of u in Eq. (6) is approximated as

$$\left[\frac{\partial}{\partial x}(\epsilon\frac{\partial u}{\partial x})\right]_{i,j,k} = \frac{1}{h^2}\left[\epsilon_{i+\frac{1}{2},j,k}\left(u_{i+1,j,k} - u_{i,j,k}\right) - \epsilon_{i-\frac{1}{2},j,k}\left(u_{i,j,k} - u_{i-1,j,k}\right)\right] + O(h^2). \tag{37}$$

The on-grid $\epsilon_{i,j,k}$ value is calculated based on the aforementioned equations, and depends on $S_{i,j,k}$, which is generated discretely by the GCS model. Since $\epsilon_{i,j,k}$ is only known at nodes (x_i, y_j, z_k), the half-grid values needed in Eq. (37), are obtained via an average

$$\epsilon_{i+\frac{1}{2},j,k} = \frac{\epsilon_{i,j,k} + \epsilon_{i+1,j,k}}{2} + O(h^2). \tag{38}$$

The order of accuracy of the entire central difference discretization is two. The finite difference discretization of the Laplacian on the left-hand side of the PB equations gives rise to a sparse matrix with dimension N_3-by-N_3, where $N_3 = N_x \times N_y \times N_z$.

For the trilinear method, the singular charges in the sources of Eqs. (6), (7) and (8) will be discretized by the traditional trilinear approximation [34]. In particular, for each partial charge q_n located at $\mathbf{r}_n$, one will find a cubic cell containing this charge. One will then distribute the charge q_n into eight corner nodes of the cell, which gives rise to eight numerical source values $Q_{i,j,k}$ for each partial charge. By treating all N_m charges, the right hand side vector of the discrete system can be formed, which will be non-zero only in Ω_i, because $\mathbf{r}_n \in \Omega_i$ for all n.

For the regularized PB equations, i.e., Eqs. (22), (24), and (26), the source terms are smooth functions defined over the entire domain Ω. Since $G_{i,j,k}$ and $\nabla G_{i,j,k}$ are known analytically, the source term $(1 - S)\kappa^2 G$ can be calculated directly at (x_i, y_j, z_k). For the source term $\nabla\epsilon \cdot \nabla G$, analytical differentiation will be used whenever possible for calculating $\nabla\epsilon$. In particular, we know

$$\nabla\epsilon(\mathbf{r}) = S(\mathbf{r})\nabla\epsilon_{in}(\mathbf{r}) + \nabla S(\mathbf{r})(\epsilon_{in}(\mathbf{r}) - \epsilon_{out}), \quad \text{in } \Omega, \tag{39}$$

in which $\nabla\epsilon_{in}(\mathbf{r})$ can be computed analytically everywhere [41], while ∇S is analytically known in Ω_i and Ω_e. In Ω_t, ∇S will be approximated by the central difference, e.g.

$$\left.\frac{\partial S}{\partial x}\right|_{i,j,k} = \frac{S_{i+1,j,k} - S_{i-1,j,k}}{2h} + O(h^2).$$

See Ref. [41] for more details. The smooth source term $\nabla\epsilon_v \cdot \nabla G$ for the regularized Poisson equation (26) in vacuum state can be calculated similarly.

For the LPB and Poisson equations, i.e., Eqs. (7), (8), (24), and (26), the above discretizations produce a linear system, which will be solved by a biconjugate gradient iterative solver in this work. For the NPB equations (6) and (22), the inexact Newton method [20] will be employed to treat the nonlinear hyperbolic Sine term.

After solving numerical potentials u, v, u_{RF} and v_{RF} on grid nodes, all forms of electrostatic free energies can be similarly discretized to generate ΔE, respectively, for Eqs. (17), (18), (27), and (29). In particular, the first term of these equations can be

obtained through a trilinear interpolation, which can be regarded as a reverse process of the trilinear distribution of singular charges. In fact, the weights calculated in the trilinear method can be re-used for this purpose. Volume integrals can be approximated by a summation over all grid nodes in $\Omega_i^C = \Omega_t \cup \Omega_e$. For example,

$$\int_{\Omega_i^C} (1-S)\kappa^2(\cosh(u)-1)d\mathbf{r} \approx \sum_{(x_i,y_j,z_k)\in\Omega_i^C} (1-S_{i,j,k})\kappa^2(\cosh(u_{i,j,k})-1)h^3. \tag{40}$$

The surface integrals are calculated as a summation over six sides of the domain Ω, where $\epsilon(\mathbf{r}) = \epsilon_{out}$ can be moved out the integrals. For instance,

$$\begin{aligned}\int_{\partial\Omega} u\frac{\partial u}{\partial n}ds \approx h^2 &\sum_{i=2}^{N_x-1}\sum_{j=2}^{N_y-1}\left(\frac{u_{i,j,1}+u_{i,j,2}}{2}\cdot\frac{u_{i,j,1}-u_{i,j,2}}{h}+\frac{u_{i,j,N_z-1}+u_{i,j,N_z}}{2}\cdot\frac{u_{i,j,N_z}-u_{i,j,N_z-1}}{h}\right)\\ + h^2 &\sum_{i=2}^{N_x-1}\sum_{k=2}^{N_z-1}\left(\frac{u_{i,1,k}+u_{i,2,k}}{2}\cdot\frac{u_{i,1,k}-u_{i,2,k}}{h}+\frac{u_{i,N_y-1,k}+u_{i,N_y,k}}{2}\cdot\frac{u_{i,N_y,k}-u_{i,N_y-1,k}}{h}\right)\\ + h^2 &\sum_{j=2}^{N_y-1}\sum_{k=2}^{N_z-1}\left(\frac{u_{1,j,k}+u_{2,j,k}}{2}\cdot\frac{u_{1,j,k}-u_{2,j,k}}{h}+\frac{u_{N_x-1,j,k}+u_{N_x,j,k}}{2}\cdot\frac{u_{N_x,j,k}-u_{N_x-1,j,k}}{h}\right).\end{aligned} \tag{41}$$

We note that the discretization in Eq. (41) is conducted at $\frac{h}{2}$ distance inside the boundary $\partial\Omega$, such that $\frac{\partial u}{\partial n}$ can be approximated by central differences for a better accuracy. For instance, for the bottom side of the domain, i.e., $z = z_1$, the approximation is considered at $z_{1.5} = z_1 + \frac{h}{2}$. At this point $\frac{\partial u}{\partial n} \approx \frac{u_{i,j,1}-u_{i,j,2}}{h}$, while u is approximated by the average of two nearby function values $\frac{u_{i,j,1}+u_{i,j,2}}{2}$.

3. Numerical experiments

In this section, we will numerically validate the proposed electrostatic free energy (EFE) forms in terms of dimensionless potentials u and v. In the super-Gaussian PB model, we take $\epsilon_{gap} = 8$ and $m = 2$, while in all computations we have $\epsilon_m = 1$ and $I = 0.15$. For the water state and vacuum state, ϵ_{out} is chosen as 80 and 1, respectively. For domain and mesh size, the length is reported with a unit Å, and the EFE automatically has the unit of kcal/mol as we designed. In all figures and tables, the regularization and trilinear methods will be denoted as REG and TRI, respectively. For simplicity, the Dirichlet boundary condition Eq. (10) is assumed in both LPB and NPB models for both regularization and trilinear methods.

All the experiments are carried out by using a single core on a Dell PowerEdge R920 in The University of Alabama High-Performance Computer (UAHPC) (https://oit.ua.edu/services/research/) with Intel(R) Xeon(R) CPU E7-8891 v2 operating at 3.20 GHz clock speed.

3.1. Diffuse interface Kirkwood sphere with benchmark energy

For a Kirkwood sphere with a sharp interface, analytical energy is known for the linearized PB (LPB) model [19,20], while an approximated energy with high precision is available for the nonlinear PB (NPB) model [2]. In order to validate the proposed EFE forms, we design a Kirkwood sphere with a simple diffuse interface dielectric function, so that the EFE can be accurately approximated. In short, by using the spherical symmetry, the three-dimensional (3D) PB equation can be reduced to a one-dimensional (1D) model [2]. The approximated energy based on the 1D model can be used to benchmark our 3D EFE forms.

Consider a diffuse interface Kirkwood sphere with a unit charge $q_0 = 1$ located at $\mathbf{r}_0 = (0,0,0)$. The interfaces Γ_i and Γ_e are spheres with radii being $r_i = 2$ and $r_e = 5$, respectively. Denote $r = |\mathbf{r}|$. A piecewise polynomial surface (PPS) is utilized to characterize the diffuse interface

$$S(r) = \begin{cases} 1 & 0 \le r \le r_i \\ p(r) & r_i \le r \le r_e \\ 0 & r_e \le r, \end{cases} \tag{42}$$

where $p(r)$ is a polynomial of degree five

$$p(r) = -6\left(\frac{r-r_i}{r_e-r_i}\right)^5 - 15\left(\frac{r-r_i}{r_e-r_i}\right)^4 - 10\left(\frac{r-r_i}{r_e-r_i}\right)^3. \tag{43}$$

The polynomial $p(r)$ is designed such that $S(r)$ is guaranteed to be C^2 continuous everywhere, including at $r = r_i$ and $r = r_e$. With such a $S(r)$ function, the dielectric function can be calculated according to Eq. (31). The NPB equation of this Kirkwood sphere system can be given as

$$-\nabla\cdot(\epsilon\nabla u(\mathbf{r})) + (1-S)\kappa^2\sinh(u(\mathbf{r})) = 4\pi\frac{e_c^2}{k_B T}q_0\delta(\mathbf{r}). \tag{44}$$

By solving the 3D NPB equation (44) together with the Dirichlet boundary condition, one obtains the potential $u(\mathbf{r})$ for the water state. In the vacuum state, we have $\epsilon = 1$ over the entire Ω for the diffuse interface model, so that the potential $v(\mathbf{r}) = G(\mathbf{r})$, i.e., it is analytically known as the Green's function. Consequently, the EFE of this diffuse interface PB model can be calculated.

In the spherical coordinate (r, θ, φ), $u(\mathbf{r})$ is a function along the radial direction only, while it is a constant along θ or φ direction. Consequently, the NPB equation (44) can be reduced to an ordinary differential equation (ODE) [2]

$$-\epsilon\frac{d^2u}{dr^2} - \frac{d\epsilon}{dr}\frac{du}{dr} - \epsilon\frac{2}{r}\frac{du}{dr} + (1-S)\kappa^2\sinh(u(r)) = 4\pi\frac{e_c^2}{k_BT}q_0\delta(r), \tag{45}$$

for an unknown potential function $u(r)$ in 1D. Here $\delta(r)$ is the 1D Dirac delta function. In order to solve Eq. (45) accurately, the physical jump conditions [19,20]

$$[u] := u^+ - u^- = 0, \quad \left[\epsilon\frac{du}{dr}\right] := \epsilon^+\frac{du^+}{dr} - \epsilon^-\frac{du^-}{dr} = 0, \tag{46}$$

need to be satisfied at both $r = r_i$ and $r = r_e$.

We propose to solve the following 1D boundary value problem (BVP) for the LPB model of the present diffuse interface Kirkwood sphere

$$\begin{cases} -\epsilon\frac{d^2u}{dr^2} - \frac{d\epsilon}{dr}\frac{du}{dr} - \epsilon\frac{2}{r}\frac{du}{dr} + (1-S)\kappa^2 u = 0, & r \in [r_i, r_e], \\ \frac{du}{dr} = -\frac{e_c^2}{k_BT}\frac{q_0}{\epsilon_m r_i^2}, & r = r_i, \\ \frac{du}{dr} + (\frac{1}{r_e} + \lambda)u = 0, & r = r_e, \end{cases} \tag{47}$$

where $\lambda = \sqrt{\kappa^2/\epsilon_{out}}$. For $0 \le r \le r_i$, we have $S = 1$ so that Eq. (45) can be analytically solved as in the sharp interface case [2]. In particular, consider a decomposition in 1D, i.e., $u = G + u_{RF}$, where the Green's function is known as $G(r) = \frac{e_c^2}{k_BT}\frac{q_0}{\epsilon_m r}$. For $r \in [0, r_i]$, the reaction-field potential u_{RF} can be solved analytically as a constant [2]. Thus, $u = C_1 + \frac{e_c^2}{k_BT}\frac{q_0}{\epsilon_m r}$ in $[0, r_i]$. Here, the unknown constant C_1 can be eliminated by taking a derivative with respect to r. By using the jump condition (46) at $r = r_i$, the potential u in $[r_i, r_e]$ also satisfies the same derivative condition, which gives rise to the Neumann boundary condition at $r = r_i$ in (47). In $r_e \le r < \infty$, the general solution of the LPB equation (47) can be given as [2]

$$u = C_2\frac{\exp(-\lambda r)}{r} + C_3\frac{\exp(\lambda r)}{2\lambda r}. \tag{48}$$

Since $\lambda > 0$, C_3 has to equal to zero in Eq. (48) so that u is finite at infinity. By taking derivative with respect to r and making use of the jump condition (46) at $r = r_e$, the potential u in $[r_i, r_e]$ satisfies the Robin boundary condition at $r = r_e$ in (47).

We note that the BVP (47) needs to be solved over a fixed domain $[r_i, r_e] = [2, 5]$, but the solution is equivalent to the one defined over the infinity domain $[0, \infty)$. Moreover, the solution $u(r)$ has an arbitrarily high regularity in the transition band $[r_i, r_e]$, so that a high order numerical approximation is feasible. In the present study, the eighth order finite difference method is employed to solve Eq. (47) numerically, and the Neumann and Robin boundary conditions are discretized by the matched interface and boundary (MIB) method [18,48]. With the numerical solution, the EFE can be estimated as in [2]. In particular, the surface integrals in Eq. (18) can be neglected, since the essential domain of the present approximation is $[0, \infty)$. Thus, the EFE ΔE is determined by the first term only, i.e., $\frac{1}{2}k_BTq_0(u(0) - G(0))$, where G is the Green's function. Using the fact that u_{RF} is a constant in $[0, r_i]$, ΔE of the 1D LPB model can be calculated as [2]

$$\Delta E = \frac{1}{2}k_BTq_0(u(0) - G(0)) = \frac{1}{2}k_BTq_0\left(u(r_i) - \frac{e_c^2}{k_BT}\frac{q_0}{\epsilon_m r_i}\right), \tag{49}$$

where $u(r_i)$ is numerically obtained and other parameters are all known. By using the eighth order MIB method and $N = 6401$, the LPB EFE of the diffuse interface Kirkwood sphere is calculated as $\Delta E = -68.884489814969$ kal/mol.

For the NPB model, Eq. (45) does not admit an analytical solution in $[r_e, \infty)$. Consequently, numerical solution of Eq. (45) over the infinity domain $[r_i, \infty)$ has to be considered. Following Ref. [2], we will truncate the infinity domain $[r_i, \infty)$ at a large enough b value, e.g., $b = 100$, and conduct the numerical approximation over the finite domain $[r_i, b]$. We propose to solve the following 1D BVP for the NPB model of the diffuse interface Kirkwood sphere

Table 1
The EFE estimate for the LPB model of the diffuse interface Kirkwood sphere. For both TRI and REG methods, a fixed step size $h = 0.4$ is considered with different domain size a.

TRI	a	First term	SI of u	SI of v	Total Sum
	6	-71.1720147	-0.1069531	-23.8144420	-94.8795036
	8	-71.1707851	-0.0511329	-17.7013280	-88.8209802
	10	-71.1698792	-0.0257987	-14.0864617	-85.2305421
	20	-71.1685750	-0.0011632	-6.9705093	-78.1379210
	30	-71.1684429	-0.0000645	-4.6311145	-75.7994930
	40	-71.1685428	-0.0000037	-3.4674474	-74.6359864
REG	6	-71.0689974	-0.1358288	-23.8222868	-94.7554554
	8	-71.0538468	-0.0640988	-17.7045287	-88.6942768
	10	-71.0467325	-0.0323894	-14.0879407	-85.1022838
	20	-70.9823709	-0.0014980	-6.9706114	-77.9514843
	30	-70.9252693	-0.0000839	-4.6312142	-75.5563996
	40	-71.0384352	-0.0000049	-3.4675235	-74.5059537

$$\begin{cases} -\epsilon\dfrac{d^2u}{dr^2} - \dfrac{d\epsilon}{dr}\dfrac{du}{dr} - \epsilon\dfrac{2}{r}\dfrac{du}{dr} + (1-S)\kappa^2\sinh(u) = 0, & r \in [r_i, b], \\ \dfrac{du}{dr} = -\dfrac{e_c^2}{k_B T}\dfrac{q_0}{\epsilon_m r_i^2}, & r = r_i, \\ u = 0, & r = b, \end{cases} \tag{50}$$

where a simple Dirichlet zero boundary condition is assumed at $r = b$, because the potential u is known to be exponentially decay as $r \to \infty$. Again, the high order MIB finite difference method [18,48] is employed to discretize the BVP Eq. (50). The resulted nonlinear system is solved by Newton's method. Since the regularity of the solution u is just C^2 continuous at $r = r_e$, the present numerical error could be larger than that of the 1D LPB BVP. With the numerical solution, the NPB EFE ΔE will be calculated according to Eq. (17). In particular, the first term will be calculated as in the 1D LPB case. With $b = 100$, the calculated surface integral values are on the order of the double precision limit, and thus do not need to be included. Therefore, ΔE of the 1D NPB model is calculated as

$$\Delta E = \frac{1}{2}k_B T q_0\left(u(r_i) - \frac{e_c^2}{k_B T}\frac{q_0}{\epsilon_m r_i}\right) - \frac{(k_B T)^2}{e_c^2}\int_{r_i}^{b}(1-S)\kappa^2(\cosh(u)-1)r^2 dr + \frac{1}{2}\frac{(k_B T)^2}{e_c^2}\int_{r_i}^{b}(1-S)\kappa^2 u\sinh(u)r^2 dr, \tag{51}$$

where two integrals are approximated by the Trapezoidal Rule. By using the eighth order MIB scheme and $N = 6401$, the NPB EFE of the diffuse interface Kirkwood sphere is calculated as $\Delta E = -73.415356786147$ kcal/mol.

With benchmark energy values, we first study the 3D LPB model of the diffuse interface Kirkwood sphere. To this end, the linearized version of Eq. (44) is solved with the Dirichlet boundary condition by using the trilinear (TRI) and regularization (REG) methods. Here, a cubic domain $\Omega = [-a, a]^3$ is considered, and the spacing h is chosen to be the same in all directions with $N_x = N_y = N_z$. By taking $h = 0.4$, we first examine the dependence of the EFE forms (18) and (29) on the domain size a. As discussed above, the surface integrals (SIs) of u and v shall be dropped in the LPB EFE, while the first term of Eqs. (18) and (29) could provide an estimate of ΔE based on a finite domain. In Table 1, we report numerical values of the first term, the SI of u, the SI of v, and the total sum against different a values. It can be seen that the first term produces a consistent estimate for different domain size a, in both trilinear and regularization schemes. On the other hand, the SI of v decays slowly with respect to a, so that the total sum does not seem to be converged for both trilinear and regularization schemes. The SI of u could be neglected for large a values in Table 1, while it may affect the summation for small a values. To further explore the convergence pattern, we focus on two EFE estimates, i.e., the energy without the SI of v (EWV) and the energy without SIs of u and v (EWUV). The calculated EWV and EWUV values are plotted against the domain size a in Fig. 2 (a). It is observed that both trilinear and regularization values approach to certain limits as a goes to infinity, while the convergence pattern of the trilinear method seems to be more consistent. On the other hand, for both methods, the convergence pattern of the EWUV seems to be better than that of the EWV.

We next examine the convergence with respect to the step size h by using a fixed domain $[-10, 10]^3$ or $a = 10$. Numerical values of the first term, the SI of u, the SI of v, and the total sum are reported in Table 2, while the EWV and EWUV estimates are plotted in Fig. 2 (b). With a fixed domain size a, it can be seen that the SIs of u and v do not change much when h becomes smaller. Thus, the EWV and EWUV estimates are almost the same in Fig. 2 (b). When h goes to zero, both trilinear and regularization methods converge to the same place, which is very close the benchmark value generated by the 1D BVP model. Moreover, the regularization errors seem to be better than those of the trilinear.

We next study the 3D NPB model of the diffuse interface Kirkwood sphere. With the same numerical setup, we solve Eq. (44) with the Dirichlet boundary condition by using the trilinear and regularization methods. For the NPB EFE forms, the first term in Eq. (17) for the trilinear method is equivalent to the first term in Eq. (27) for the regularization method, while other terms in both equations

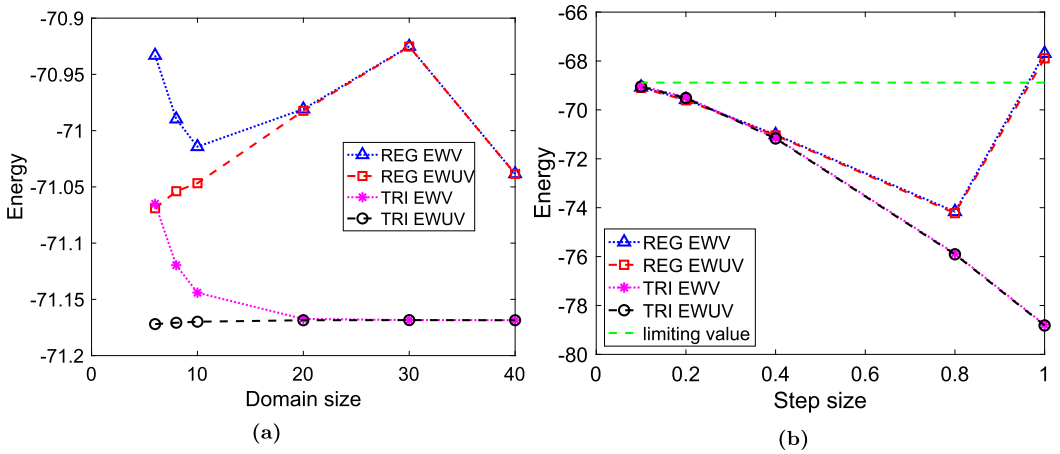

Fig. 2. Convergence in the EFE of the LPB model for the diffuse interface Kirkwood sphere. (a) Convergence with respect to the domain size a with $h = 0.4$; (b) Convergence with respect to the step size h with $a = 10$. For both TRI and REG methods, the energy without SI of v (EWV) and the energy without SIs of u and v (EWUV) are plotted.

Table 2
The EFE estimate for the LPB model of the diffuse interface Kirkwood sphere. For both TRI and REG methods, a fixed domain size $a = 10$ is considered with different step size h.

TRI	h	First term	SI of u	SI of v	Total Sum
	1.0	-78.8165076	-0.0288689	-14.5557469	-93.3433857
	0.8	-75.9068638	-0.0277955	-14.3933877	-90.2724560
	0.4	-71.1698792	-0.0257987	-14.0864617	-85.2305421
	0.2	-69.5119358	-0.0249069	-13.9412562	-83.4282852
	0.1	-69.0478315	-0.0245337	-13.8709084	-82.8942061
REG	1.0	-67.8899534	-0.1970892	-14.5673015	-82.2601656
	0.8	-74.2226938	-0.0593496	-14.4005612	-88.5639053
	0.4	-71.0467325	-0.0323894	-14.0879407	-85.1022838
	0.2	-69.6061724	-0.0259173	-13.9412227	-83.5214778
	0.1	-69.0968629	-0.0245705	-13.8700938	-82.9423863

are essentially the same. As discussed in Section 2, the SI of v should be dropped, but the inclusion of the SI of u should be explored in the NPB model, because it may serve as a closure term for the volume integrals, i.e., the second and third terms of Eq. (17) and Eq. (27).

The numerical values of each term are reported in Table 3 for different domain size. It can be seen that the first three terms produce consistent values for all a values. Both SIs of u and v are decaying as a becomes larger. The SI of u decays quickly, while the SI of v decays too slow so that the total sum does not display converging pattern. For the NPB energy, the sum of the first three terms yields the energy without SIs of u and v (EWUV), while energy without the SI of v (EWV) includes the SI of u. These two EFE estimates are depicted in Fig. 3 (a) for different domain size a. Again, the trilinear energies converge better than those of regularization, while both trilinear and regularization values approach to certain limits as a goes to infinity. On the other hand, for both methods, the convergence pattern of the EWUV seems to be better than that of the EWV.

The convergence with respect to the step size h is finally studied for the NPB model. By using a fixed domain $[-10, 10]^3$, numerical values are reported in Table 4, while the EWV and EWUV estimates are plotted in Fig. 3 (b). With a constant a, the SIs of u and v do not change much when h goes to zero. Consequently, the EWV and EWUV estimates are almost the same in Fig. 3 (b). When h goes to zero, both trilinear and regularization methods converge to the same place, which agrees well to the benchmark value generated by the 1D BVP model.

3.2. Super-Gaussian NPB energy of one atom system

In the rest of the paper, we will mainly focus on the EFE forms of the NPB model, i.e., Eq. (17) and Eq. (27), for the super-Gaussian model [21,41]. In this subsection, we study a simple super-Gaussian model, i.e., one atom system. Similarly to the diffuse interface Kirkwood sphere, we take $q_0 = 1$, $\mathbf{r}_0 = (0, 0, 0)$, $r_i = 2$, and $r_e = 5$. The domain is also chosen as $[-a, a]^3$. Unlike the previous subsection, the smooth surface function S is calculated by the Gaussian convolution surface (GCS) [39], which is designed for general proteins. Because the GCS is not analytically defined and is mesh dependent [39], we will not generate 1D benchmark value for the present one atom system. By using the super-Gaussian dielectric function ϵ, the potential $u(\mathbf{r})$ for the water state is also governed

Table 3
The EFE estimate for the NPB model of the diffuse interface Kirkwood sphere. For both TRI and REG methods, a fixed step size $h = 0.4$ is considered with different domain size a.

TRI	Domain	First term	Second term	Third term	SI of u	SI of v	Total Sum
	6	-75.0430552	0.2941850	2.0158050	-0.0471305	-23.8144420	-97.0887468
	8	-75.0494765	0.2945447	2.0056965	-0.0208711	-17.7013280	-91.0187816
	10	-75.0522704	0.2933049	1.9993742	-0.0097121	-14.0864617	-87.4229507
	20	-75.0550077	0.2882418	1.9887986	-0.0002990	-6.9705093	-80.3246613
	30	-75.0550468	0.2874315	1.9876921	-0.0000131	-4.6311145	-77.9858877
	40	-75.0551615	0.2873529	1.9875959	-0.0000007	-3.4674474	-76.8223652
REG	6	-74.5220398	0.2619886	1.7844080	-0.0820912	-23.8222868	-96.7398160
	8	-74.5235821	0.2681509	1.7865467	-0.0368425	-17.7045287	-90.6728724
	10	-74.5243852	0.2705294	1.7867967	-0.0178754	-14.0879407	-87.0781832
	20	-74.4979001	0.2732617	1.8009576	-0.0007151	-6.9706114	-79.9401004
	30	-74.4704240	0.2750667	1.8162537	-0.0000369	-4.6312142	-77.5604143
	40	-74.5253063	0.2707925	1.7845898	-0.0000021	-3.4675235	-76.4790304

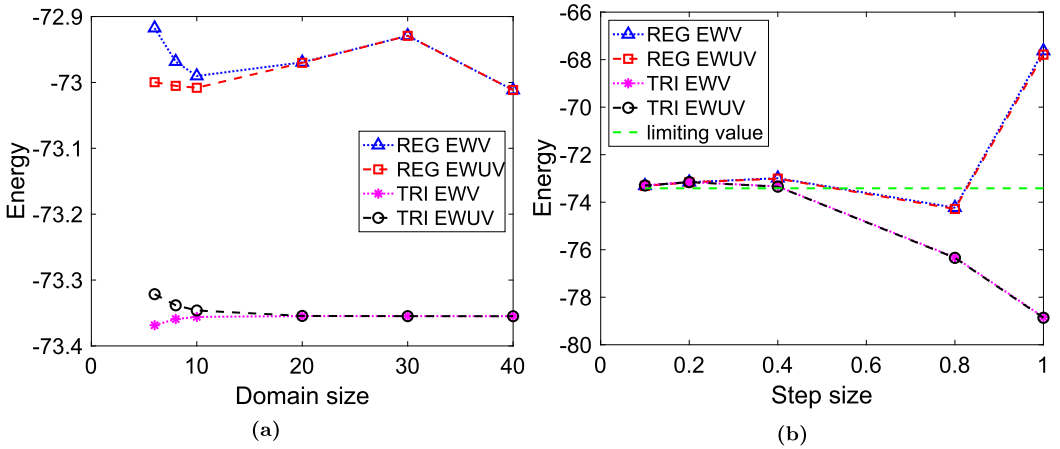

Fig. 3. Convergence in the EFE of the NPB model for the diffuse interface Kirkwood sphere. (a) Convergence with respect to the domain size a with $h = 0.4$; (b) Convergence with respect to the step size h with $a = 10$. For both TRI and REG methods, the energy without SI of v (EWV) and the energy without SIs of u and v (EWUV) are plotted.

Table 4
The EFE estimate for the NPB model of the diffuse interface Kirkwood sphere. For both TRI and REG methods, a fixed domain size $a = 10$ is considered with different step size h.

TRI	h	First term	Second term	Third term	SI of u	SI of v	Total Sum
	1.0	-78.9742294	0.0878285	0.1980849	-0.0255339	-14.5557469	-93.3941861
	0.8	-76.9846256	0.1604262	0.8032858	-0.0194701	-14.3933877	-90.7156837
	0.4	-75.0522704	0.2933049	1.9993742	-0.0097121	-14.0864617	-87.4229507
	0.2	-75.5850267	0.3549858	2.7847067	-0.0055747	-13.9412562	-87.0909873
	0.1	-75.9753550	0.3743093	3.0537928	-0.0042461	-13.8709084	-87.1625339
REG	1.0	-68.1112016	0.5643663	0.8866032	-0.1552432	-14.5673015	-82.2010230
	0.8	-74.4347284	0.1368915	0.2879194	-0.0548359	-14.4005612	-88.6294257
	0.4	-74.5243852	0.2705294	1.7867967	-0.0178754	-14.0879407	-87.0781832
	0.2	-75.5380548	0.3465541	2.7193076	-0.0071415	-13.9412227	-87.0993825
	0.1	-75.9721343	0.3712849	3.0305371	-0.0045530	-13.8700938	-87.1784229

by the 3D NPB equation (44) together with the Dirichlet boundary condition. For the vacuum state, one needs to solve a Poisson equation for the potential $v(\mathbf{r})$ [21,41]. Then, the EFE can be calculated based on Eq. (17) for the trilinear method. The EFE for the regularization can be computed similarly.

We first explore the energy dependence on the domain size a for the present system. With a fixed $h = 0.4$, the energy values are reported in Table 5. It can be seen that SIs of u and v are decaying as before. However, compared with the results of diffuse interface Kirkwood sphere in Table 3, the present energy values are heavily affected by the domain size a. It is unclear if each individual term

Table 5
The EFE estimate for the super-Gaussian NPB model of one atom system. For both TRI and REG methods, a fixed step $h = 0.4$ is considered with different domain size a.

TRI	a	First term	Second term	Third term	SI of u	SI of v	Total Sum
	6	-44.2962701	0.0559752	0.1320352	-0.1006547	-23.8435067	-67.9630622
	8	-44.1340520	0.0670210	0.1487405	-0.0476979	-17.7133604	-61.7179950
	10	-44.3702771	0.0741313	0.1620203	-0.0238197	-14.0923073	-58.3508756
	20	-43.7920295	0.0790188	0.1618223	-0.0010602	-6.9711426	-50.6793084
	30	-43.9453590	0.0801425	0.1670520	-0.0000581	-4.6313636	-48.4897550
	40	-43.9844271	0.0766243	0.1348956	-0.0000036	-3.4675694	-47.3937216
REG	6	-44.4094940	0.0406952	0.0541923	-0.1109323	-23.4194020	-67.7044665
	8	-44.2044516	0.0517317	0.0649382	-0.0508426	-17.4303940	-61.5707965
	10	-44.3592999	0.0585160	0.0728337	-0.0251579	-13.8607228	-58.1805471
	20	-43.8207993	0.0649118	0.0777307	-0.0010949	-6.8796069	-50.6864925
	30	-43.9246945	0.0655977	0.0790750	-0.0000591	-4.5688192	-48.4799773
	40	-43.8855033	0.0667357	0.0766252	-0.0000035	-3.4174141	-47.2930245

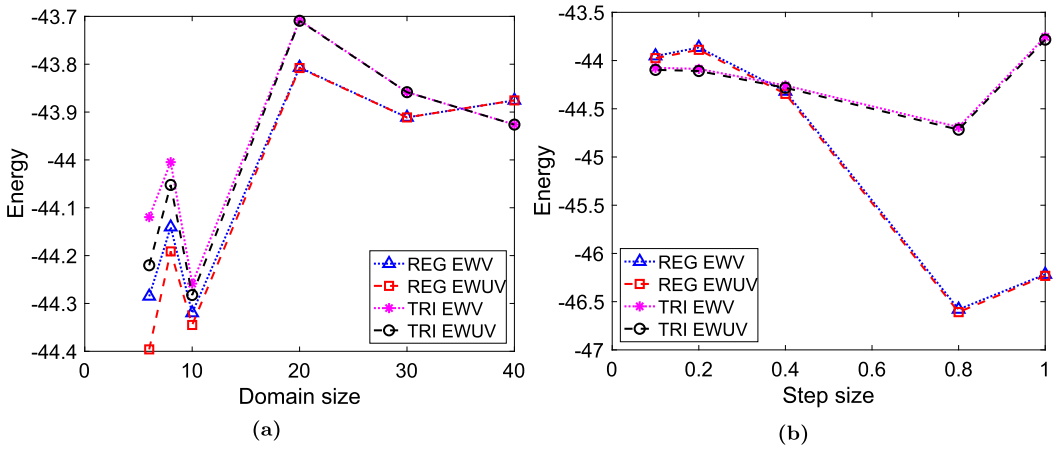

Fig. 4. Convergence in the EFE of the super-Gaussian NPB model for one atom system. (a) Convergence with respect to the domain size a with $h = 0.4$; (b) Convergence with respect to the step size h with $a = 10$.

will converge as a becomes larger. In Fig. 4 (a), the EWV and EWUV values are depicted for both REG and TRI methods. Oscillations can be seen in all cases for small domain sizes, which should be due to the GCS. The GCS has a convergence problem since it is discretely computed and Ref. [39] has demonstrated that the GCS convergence is oscillatory rather than monotone which impacts the convergence of energy. In the present study, only when $a \geq 20$, certain converging patterns could be seen. For both methods, EWV and EWUV values are the same for $a \geq 20$, while for small a values, it is hard to say which one is better.

We next examine the convergence with respect to the step size h. By using a fixed domain $[-10, 10]^3$, numerical values are listed in Table 6, while the EWV and EWUV values are plotted in Fig. 4 (b). Such results are similar to those of the diffuse interface Kirkwood sphere. In particular, the SIs of u and v are stable for a fixed a, so that the EWV and EWUV estimates are almost the same throughout. When h goes to zero, both REG and TRI energies converge to the same place, while the convergence of the TRI seems to be faster.

3.3. Super-Gaussian NPB energy of a protein

We next study the EFE of the super-Gaussian NPB model of a protein (PDB ID: 1TQG). For a protein, our code will automatically calculate a tight domain containing all VdW balls of the protein system. An edge value is then chosen to extend the domain in all Cartesian directions to define the computational domain Ω. The GCS surface can then be generated with the band-width of the transition region Ω_t being 3Å, which imposes a lower bound for the edge value. By using a fixed $h = 0.5$, the results of the TRI and REG methods are reported in Table 7 for several edge values. It can be seen that the first three terms of Eqs. (17) and (27) obviously converge to certain limits. The SI of v has large magnitudes for all tested edge value. The SI of u is decaying as the edge becomes larger. However, the magnitude of the SI of u is very small comparing with that of the first term. In other words, the inclusion or exclusion of the SI of u essentially does not affect the EFE for protein systems, even if a very small edge value is used. In Fig. 5 (a), the plots of the EWV and EWUV are given. Indeed, one cannot distinguish the EWV and EWUV in all cases. As can be seen in Fig. 5 (a), the NPB energies of the TRI and REG methods are approaching certain constant values, as the edge becomes larger. On the other

Table 6
The EFE estimate for the super-Gaussian NPB model of one atom system. For both TRI and REG methods, a fixed domain size $a = 10$ is considered with different step size h.

TRI	h	First term	Second term	Third term	SI of u	SI of v	Total Sum
	1.0	-43.7893335	0.0544754	0.0608482	-0.0283965	-14.5960264	-58.3505906
	0.8	-44.7202599	0.0547629	0.0584914	0.0271119	-14.4225638	-59.1119834
	0.4	-44.3702771	0.0741313	0.1620203	-0.0238197	-14.0923073	-58.3508756
	0.2	-44.2984548	0.0958764	0.2859959	-0.0216004	-13.9422702	-58.0290051
	0.1	-44.2790319	0.0943178	0.2771082	-0.0213365	-13.8705565	-57.9454615
REG	1.0	-46.5851051	0.1292982	0.4819823	-0.0145912	-14.3758051	-60.5936349
	0.8	-46.6111645	0.0565624	0.0592719	-0.0284192	-14.2845899	-60.8646256
	0.4	-44.3592999	0.0585160	0.0728337	-0.0251579	-13.8607228	-58.1805471
	0.2	-44.0574104	0.0946896	0.2635359	-0.0234181	-13.8340471	-57.6991931
	0.1	-44.1672353	0.0976008	0.2896286	-0.0220705	-13.8303163	-57.7834534

Table 7
The EFE estimate for the super-Gaussian NPB model of the protein 1TQG. For both TRI and REG methods, a fixed step size $h = 0.5$ is considered with different edge values. CPU time in seconds is reported for all cases.

TRI	Edge	First term	Second term	Third term	SI of u	SI of v	Total Sum	CPU Time
	5	-696.2046530	3.4307394	16.7763499	-0.1968619	-291.9357902	-974.5979709	47.72498
	8	-689.8894325	3.5031771	16.8861305	-0.0725517	-258.5621050	-934.9960325	72.29057
	10	-687.5080350	3.5271612	16.9185447	-0.0386366	-240.0826858	-914.1607008	92.54121
	15	-684.3568187	3.5550070	16.9667289	-0.0088519	-204.7070019	-875.6432469	153.9583
	20	-682.6563175	3.5378867	16.7721385	-0.0023198	-178.5907772	-848.0105232	253.3977
REG	5	-674.2475947	2.7397790	10.4875390	-0.2770970	-279.8493374	-946.0720751	54.57806
	8	-666.4267914	2.8772650	10.6764696	-0.0998373	-248.1275497	-906.6552992	79.31136
	10	-663.2876316	2.9252907	10.7421286	-0.0526325	-230.4725191	-885.8906804	101.6429
	15	-658.7290794	2.9790051	10.8208186	-0.0119135	-196.5972892	-847.4726416	167.4202
	20	-656.0605690	2.9733687	10.6613435	-0.0030624	-171.5397213	-819.9092532	296.6006

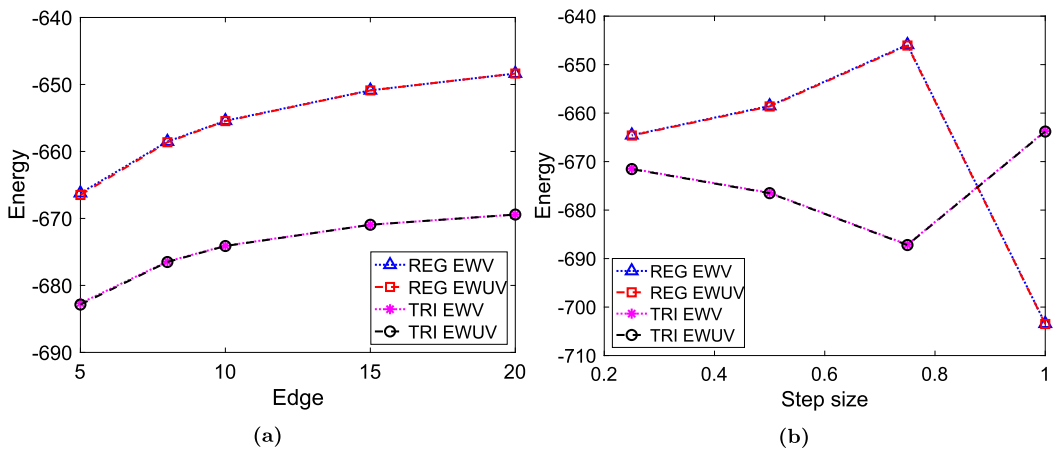

Fig. 5. Convergence in the EFE for the super-Gaussian NPB energy of the protein 1TQG. (a) Convergence with respect to the edge value with $h = 0.5$; (b) Convergence with respect to the step size h with a fixed edge value 8.

hand, when the edge is larger, more computational time is required, which has been reported in Table 7. The edge value of 8Å seems to a good choice that balances the efficiency and EFE accuracy.

With a fixed $h = 0.5$, the energy difference between the TRI and REG methods seems to be a constant for different edge values. We next show that such a difference could be reduced when h becomes smaller. In Table 8, we report values of each term for a fixed edge = 8 and different h values. Moreover, the plots of the EWV and EWUV are offered in Fig. 5 (b). It can be seen that the TRI and REG energies will converge to the same limit as h goes to zero. Moreover, as shown in Table 8, the CPU time increases dramatically as mesh is refined. A large edge value will be too expensive for large proteins.

Table 8
The EFE estimate for the super-Gaussian NPB model of the protein 1TQG. For both TRI and REG methods, a fixed edge value 8 is considered with different step size h. CPU time in seconds is reported in all cases.

TRI	h	First term	Second term	Third term	SI of u	SI of v	Total Sum	CPU Time
	1.0	-673.9264156	2.5278947	12.6498256	-0.0882812	-261.3973545	-925.1135579	6.63489
	0.75	-698.8579774	3.2827020	14.9344676	-0.0848568	-261.5101655	-948.6315204	16.74778
	0.5	-689.8894325	3.5031771	16.8861305	-0.0725517	-258.5621050	-934.9960325	72.29057
	0.25	-686.4809492	3.6226822	18.5597293	-0.0681598	-257.2366454	-928.7123877	772.6540
REG	1.0	-731.5187095	5.0115913	33.0370498	-0.0683663	-243.9227306	-947.3476153	11.00910
	0.75	-654.2857568	3.5016205	11.7224586	-0.1516511	-240.2092947	-886.1225623	17.46272
	0.5	-666.4267914	2.8772650	10.6764696	-0.0998373	-248.1275497	-906.6552992	79.31136
	0.25	-678.1767264	3.5380806	17.0892388	-0.0775160	-254.1526561	-918.7007083	904.0687

For all tests so far, the EWUV estimate is either the same as the EWV estimate, or is slightly better. Based on these experiments, we conclude that both SIs of u and v should be dropped in calculating the EFE of the NPB model. For the rest of paper, we will simply report the EWUV value, i.e., the sum of the first three terms in Eqs. (17) and (27), as the NPB EFE.

3.4. Impact of rotation and shifting

In this subsection, we focus on investigating the impact of grid rotation and shifting. When a biomolecule is moved or rotated, its electrostatic free energy remains unaffected from a physical perspective. However, the NPB EFE defined by Eqs. (17) and (27) has to be computed discretely by the finite difference method and numerical quadrature, and such approximations are grid dependent and may be impacted by rotation and shifting. Hence, it is essential to assess the grid sensitivity to rotation and shifting for the EFE of the trilinear method, i.e., Eq. (17) and the EFE of the regularization method, i.e., Eq. (27).

The sensitivity of grid rotation and shifting has been examined for the super-Gaussian LPB model for a small compound in [41]. In our investigation, we will consider the super-Gaussian NPB model for a protein (PDB ID: 1AHO). With an edge value of 8 and step size $h = 0.5$, an initial domain Ω is selected automatically. Unsynchronized rotation and shifting will be considered in this work. This means that the charge positions and protein structure in the vacuum state are different from those in the water state, due to rotation or shifting. Consequently, the domain Ω, subdomains, and smooth surface function S will be different too. In particular, consider charges in water state at $\mathbf{r}_j$ for $j = 1, 2, \ldots, N_m$. After unsynchronized rotation or shifting, denote the new charge positions in vacuum state as $\mathbf{r}'_j$ for $j = 1, 2, \ldots, N_m$. The NPB EFE of the trilinear method will be calculated as

$$\Delta E = \frac{1}{2} k_B T \sum_{j=1}^{N_m} q_j (u(\mathbf{r}_j) - v(\mathbf{r}'_j)) - \frac{1}{4\pi} \frac{(k_B T)^2}{e_c^2} \int_{\Omega_i^C} (1 - S)\kappa^2 (\cosh(u) - 1) d\mathbf{r} + \frac{1}{8\pi} \frac{(k_B T)^2}{e_c^2} \int_{\Omega_i^C} (1 - S)\kappa^2 u \sinh(u) d\mathbf{r}. \tag{52}$$

Similarly, in the regularization method, the EFE (27) is replaced by

$$\Delta E = \frac{1}{2} k_B T \sum_{j=1}^{N_m} q_j (u_{RF}(\mathbf{r}_j) - v_{RF}(\mathbf{r}'_j)) - \frac{1}{4\pi} \frac{(k_B T)^2}{e_c^2} \int_{\Omega_i^C} (1 - S)\kappa^2 (\cosh(u) - 1) d\mathbf{r} + \frac{1}{8\pi} \frac{(k_B T)^2}{e_c^2} \int_{\Omega_i^C} (1 - S)\kappa^2 u \sinh(u) d\mathbf{r}. \tag{53}$$

In both equations, the summations of u and v should be calculated separately in the first term in different coordinate systems [41]. For the second and third terms, the volume integrals are calculated for u in water state only, i.e., Ω_i^C is based on the original subdomains before rotation or shifting.

To conduct the shifting test, we will displace the 1AHO protein along the z direction in increments of 0.05. Specifically, the x and y coordinate values of all atomic centers will remain unchanged, while the z values will be increased by 0.05 in each step. With each shift, a computational domain will be generated automatically. After ten shifts, the relative position of the shifted structure in relation to the uniform mesh will essentially be identical to the initial position due to a grid spacing of $h = 0.5$. Consequently, the numerical energy should be equivalent to the original energy. In the rotation test, for each charge center or atom center $\mathbf{r}_j$, we will keep its x coordinate value a constant, while rotating its (y, z) coordinates within the yz-plane relative to the origin $(y, z) = (0, 0)$. The rotation will be performed in increments of 30°, resulting in a total of 12 steps to complete a full rotation and return the (y, z) coordinates to their original positions. For each rotated structure, a new computational domain will be automatically generated to calculate the potential v.

The numerical illustrations of unsynchronized rotation and shifting are presented in Fig. 6. Notably, significant errors are observed in the trilinear method. For shifting, the most pronounced deviation occurs at 0.15, resulting in an energy value of approximately −1010. On the other hand, in the case of rotation, the energy exhibits dramatic oscillations with maximum deviations around −1413. At rotation angles of 90, 180, and 270 degrees, the energy closely resembles that of 0 degree. Remarkably, the regularization energy remains largely unaffected by such unsynchronized rotation or shifting, with energies always around −153. In Fig. 6, the energy curves appear as straight lines. The graph in Fig. 6 demonstrates the robustness of the regularization method when it comes to grid rotation and shifting.

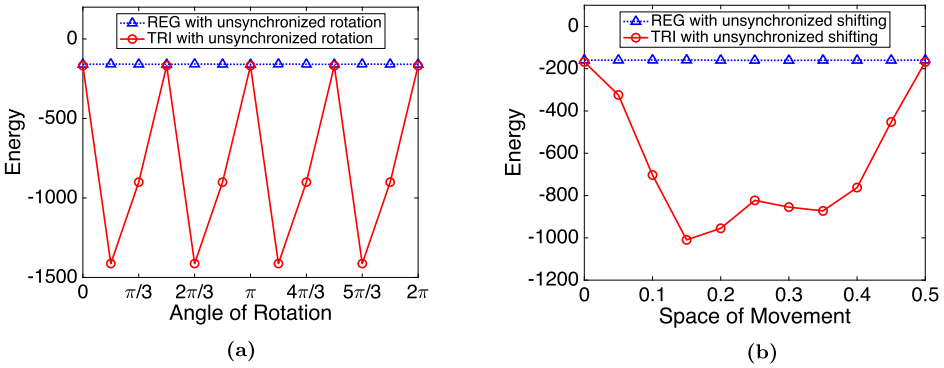

Fig. 6. Grid sensitivity tests of rotation and shifting in the NPB EFE for the 1AHO. (a) Unsynchronized rotation; (b) Unsynchronized shifting.

The large errors involved in the trilinear method for unsynchronized rotation and shifting are due to the presence of grid artifacts or artificial grid energy. In original EFE Eq. (17), the trilinear discretization of singular charges yields very large errors at charge centers $\mathbf{r}_j$, which fortunately happen to be the same in solving u in water state and v in vacuum state. Thus, such charge errors are canceled in Eq. (17). With unsynchronized rotation and shifting, such cancelation does not happen in Eq. (52). Hence, the trilinear energy is mostly dominated by the artificial grid energy in Fig. 6. By treating singular charges analytically, the artificial grid energy is completely eliminated in the regularization method. Thus, even with unsynchronized rotation and shifting, the EFE Eq. (53) produces a constant estimate for different rotation and shifting. In the rest of the paper, we will mainly focus on the regularization method.

3.5. Electrostatic free energies for linearized and nonlinear PB models

In this subsection, we will study the energy difference between the linearized PB (LPB) and nonlinear PB (NPB) models. In particular, by using the super-Gaussian PB model with the regularization, we calculate the LPB EFE ΔE_L by Eq. (29) and the NPB EFE ΔE_N by Eq. (27) for a set of proteins, and compare their difference. Moreover, we will conduct the similar studies by using a sharp-interface PB model, i.e., the rMIB package originally developed in [20]. By using the matched interface and boundary (MIB) scheme for treating the sharp interface and jump conditions, and a two-component regularization method for singular charges, the rMIB finite difference algorithm is known to provide a second order accuracy in solving the LPB equation [20]. In the present study, the ESES molecular surface [30], instead of the MSMS molecular surface, is employed in the rMIB package. However, in the rMIB package [20], the EFE of the NPB model is calculated by only the first term of Eq. (27). For this reason, the rMIB NPB energy will be denoted as ΔE_{N1} or marked as "NPB first term" in the present study. For a comparison, for the super-Gaussian model, besides the NPB energy ΔE_N, we will also report the result of NPB first term, i.e., ΔE_{N1}, which is calculated based on the first term only.

A set of proteins that have been studied previously in [7,41] will be considered. For both super-Gaussian and rMIB methods, we take $\epsilon_m = 1$, $I = 0.15$, $h = 0.5$, and ϵ_{out} is chosen as 1 and 80, respective, in vacuum and water state. In the super-Gaussian PB model, we select $\epsilon_{gap} = 8$, $m = 2$, and the edge value is taken as 8. A default edge value is used in the rMIB.

Five energy values are reported for each protein in Table 9 – three are generated by the super-Gaussian model and two by the rMIB. One can first see that the super-Gaussian energies are usually weaker than the rMIB energies, i.e., the super-Gaussian energy has a smaller magnitude. This is primarily due to the underlying PB model, because with a transition band of the width 3Å, the electrostatic interactions between solute and solvent are weaker in the super-Gaussian model. A similarly phenomenon has been observed in [41] between the super-Gaussian model and MIBPB package [8,19], which is another MIB method based solver for the sharp interface PB model. In fact, it has been shown in [41] that the Pearson correlation coefficient between the super-Gaussian and MIBPB energies is as high as 0.791. For the present study, there also exist some correlations between the super-Gaussian and rMIB energies, as can be seen from Table 9.

What is more interesting in the present work is the energy difference between the LPB and NPB models. For the rMIB package, it can be seen from Table 9 that the difference between ΔE_L and ΔE_{N1} is very small. However, for the super-Gaussian (SG) model, a large difference can be seen between ΔE_L and ΔE_{N1}. When the second and third terms are included in the NPB energy, the difference between ΔE_L and ΔE_N becomes slightly smaller. To quantify the difference, we calculate the relative difference between the LPB and NPB energies as

$$\frac{|\Delta E_L - \Delta E_N|}{|\Delta E_L|},$$

and express this quantity in percentage. The results are plotted in Fig. 7 for rMIB NPB first term, SG NPB first term, and SG NPB. The percentages of the rMIB NPB first term are almost invisible in Fig. 7, because the maximal relative difference between ΔE_L and ΔE_{N1} is just 0.42% with the average being 0.12%. Nevertheless, by also considering the first term only in the NPB energy, the super-Gaussian model has a huge difference between ΔE_L and ΔE_{N1}, with the relative difference being as high as 20.41% and the

Table 9
Electrostatic free energy (kcal/mol) of 25 proteins calculated by the super-Gaussian (SG) and rMIB packages. Here ΔE_L and ΔE_N denotes, respectively, the EFE of the LPB and NPB model. In the rMIB package, the NPB energy is calculated by the first term only [20], which is denoted as ΔE_{N1} here. For a comparison, ΔE_{N1} of the SG model is also reported.

		Electrostatic free energy				
PDB	Atom	Super-Gaussian			rMIB	
-ID	Number	ΔE_L	ΔE_{N1}	ΔE_N	ΔE_L	ΔE_{N1}
1AHO	962	-145.6967	-162.5832	-153.3686	-889.4498	-890.1712
1C75	985	-550.4969	-560.3678	-554.4116	-1423.2635	-1424.4291
1J0P	1597	-787.8966	-817.4963	-801.9987	-2382.2031	-2384.6877
1TG0	1029	-1632.6183	-1661.4553	-1647.0983	-2712.4644	-2715.3713
1X8Q	2815	-355.8890	-428.5199	-397.8534	-2456.7551	-2462.1286
1CBN	639	-50.8994	-56.1248	-52.8387	-367.3665	-367.5917
1G6X	888	-522.6884	-535.4013	-528.1621	-1314.3318	-1315.4019
1IUA	1207	-146.8194	-159.8717	-152.7182	-927.3580	-929.3352
1L9L	1226	-1478.0643	-1506.4778	-1492.0616	-2784.9215	-2787.8069
1M1Q	1265	-685.3718	-705.4891	-694.2581	-1961.0643	-1962.1155
1NWZ	1912	-576.6918	-606.9233	-589.5955	-2031.1999	-2032.7953
1OK0	1076	-396.0022	-416.1228	-405.6546	-1140.8690	-1142.2690
1TQG	1660	-654.0592	-666.4268	-658.6276	-1686.3357	-1687.5332
1VB0	913	-198.5197	-210.2443	-204.0329	-888.9493	-889.2464
1VBW	1056	-740.8630	-758.4981	-749.0084	-1570.6146	-1571.9321
1W0N	1756	-454.3624	-528.2604	-503.2634	-1700.7606	-1704.6462
1X6X	1732	-284.9659	-301.6766	-291.5538	-1514.9968	-1515.8009
1XMK	1268	-220.7955	-239.6277	-229.7723	-1204.0442	-1204.9545
1ZUU	868	-360.8595	-379.1776	-369.9746	-1243.7473	-1245.4332
1ZZK	1252	-296.0566	-319.2798	-306.7525	-1312.4640	-1314.3494
2FDN	731	-781.1547	-805.4666	-795.1820	-1419.3870	-1425.3579
2FMA	924	-253.8485	-264.6598	-258.1078	-1034.6278	-1035.2162
2FWH	1830	-539.6444	-564.2968	-550.6882	-1768.8684	-1771.2936
2H5C	2755	-631.1868	-657.3317	-642.3025	-1832.1679	-1836.6080
2IDQ	1596	-378.0852	-400.6308	-388.0415	-1500.1329	-1501.8026

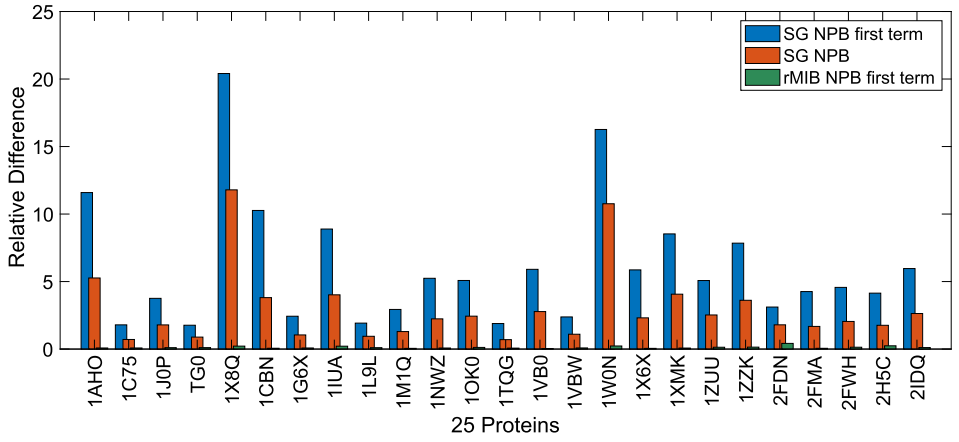

Fig. 7. Relative energy difference in percentage between the LPB and NPB energies for a set of 25 proteins.

average being 6.08%. The energy difference is slightly reduced, when the second and third terms are considered in Eq. (27). Now, the maximal relative difference between ΔE_L and ΔE_N is 11.79% and the mean value is 2.96%.

The significantly different patterns shown in Fig. 7 are primarily due to the underlying PB models. It seems that for the sharp interface PB models, the energy difference between linearized and nonlinear cases is always small. For the present rMIB package, the mean value of the relative differences between ΔE_L and ΔE_{N1} for 25 proteins is as small as 0.12%. To put things in perspective, we also consider the LPB and NPB energies of the sharp interface Kirkwood sphere, whose benchmark values are reported in [2]. For a Kirkwood sphere with a unit charge $q = 1$, radius being 2Å, and other parameters being the same as in the present study, the LPB EFE

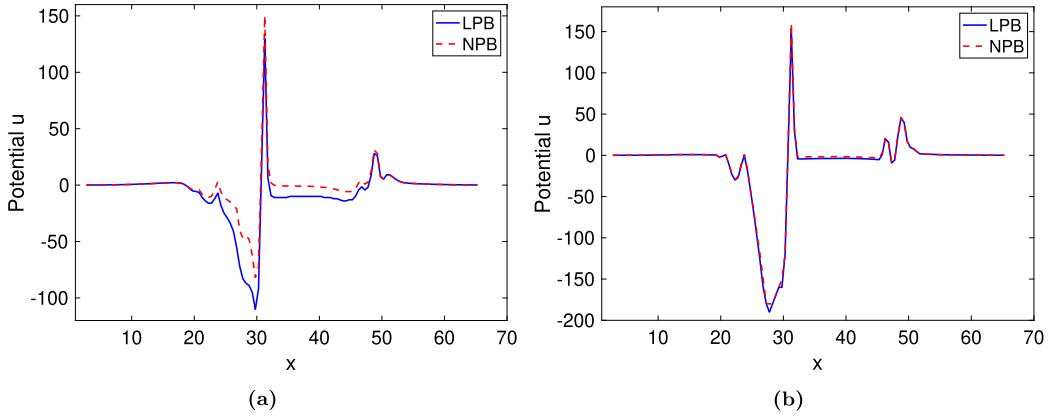

Fig. 8. Plot of the potential u of the protein 1X8Q along the x direction with $h = 0.5$, $y = 33.45$, and $z = 32.8$. (a) Super-Gaussian; (b) rMIB.

is known analytically as $\Delta E_L = -82.188683337726175$ kcal/mol. For the same set of parameters, the NPB EFE calculated by the 1D boundary value problem is $\Delta E_{N1} = -82.212210771856221$ kcal/mol, whose error is estimated to be about 2.4×10^{-12} [2]. Therefore, the relative difference between ΔE_L and ΔE_{N1} for the Kirkwood sphere is 0.0286%. We note that the Kirkwood sphere NPB energy ΔE_{N1} for other parameters can be calculated by using the MATLAB code available at https://szhao.people.ua.edu/research.html.

The present study on sharp interface PB energies justifies why the NPB model is not popular in the literature. Indeed, for the sharp interface PB model, when the NPB energy is so close to the LPB energy, it does not make much sense to pay extra computational cost to solve the nonlinear system. However, for the diffuse interface and heterogeneous dielectric PB models, the story could be different. For the super-Gaussian PB model, the relative difference between the LPB and NPB energies could be as high as 20.41% or 11.79%, as shown in Fig. 7. In Section 3.1, the benchmark energy values of the diffuse interface Kirkwood sphere are generated by solving 1D boundary value problems. In particular, the LPB and NPB energy is calculated as $\Delta E_L = -68.884489814969$ kal/mol and $\Delta E_N = -73.415356786147$ kcal/mol, respectively, for the same set of parameters. Thus, the relative energy difference of the diffuse interface Kirkwood sphere is 6.5775%, which is comparable to the mean values of the super-Gaussian energies of 25 proteins, i.e., 2.96% for ΔE_N or 6.08% for ΔE_{N1}. Therefore, for diffuse interface and Gaussian dielectric PB models, the NPB energy could play a significantly different role than the LPB energy in real biological applications.

To understand the relationship between the NPB and LPB energies, let us focus on the Kirkwood sphere in both sharp interface and diffuse interface settings. In both settings, the reaction field potential u_{RF} can be regarded as a 1D function in terms of r. Let us consider the trend of function $u_{RF}(r)$ in a reversed manner. As r decreases from infinity to zero, u_{RF} grows from zero to a constant value inside the Kirkwood sphere for both sharp and diffuse settings and for both LPB and NPB models. By considering only the first term in Eqs. (29) and (27), ΔE_L and ΔE_{N1} are essentially determined by the constant value of u_{RF} inside the Kirkwood sphere [2]. For the sharp interface PB model, the closeness of ΔE_L and ΔE_{N1} implies that the reaction field constants for the LPB and NPB models are very close, while in the diffuse interface PB model, such two constants are quite different. Recall that for the sharp interface PB model, there are just two subdomains, i.e., solute and solvent regions. In the solvent region, u_{RF} is vanishing at the infinity or on the right, and is limited on the left by the reaction field constant. Since two reaction field constants for LPB and NPB cases are very close on the left, one can conclude that the reaction field potentials in two cases are also very similar in the solvent subdomain. In other words, the grow rates of u_{RF} for both LPB and NPB cases as r becomes smaller, are similar in the solvent region, even though such rates are governed by two different equations, i.e., the LPB and NPB equations. The same argument can be applied to the subdomain Ω_e for the diffuse interface Kirkwood sphere, because with $S = 0$, the diffuse interface PB equation in Ω_e is the same as the sharp interface PB equation in the solvent region. Thus, for the diffuse interface Kirkwood sphere, the grow rates of u_{RF} for both LPB and NPB cases should be very close in Ω_e. However, since the reaction field constants for the LPB and NPB are significantly different in Ω_i, the grow rates of u_{RF} in the transition layer Ω_t should be significantly different for both LPB and NPB cases. In Ω_t, the $(1 - S)$ term behind u or $\sinh(u)$ plays an important role in amplifying the grow rate when r is decreasing or decay rate when r is increasing. In summary, it is believed that the transition layer Ω_t and the $(1 - S)$ term are the major factors that introduce a large difference between the NPB and LPB energies in the diffuse interface and super-Gaussian PB models.

To further examine the energy difference, the potential u of a protein (PDB ID: 1X8Q) is depicted in Fig. 8. In order to compare the super-Gaussian and rMIB models, we have manually changed the domain of the super-Gaussian package such that it is the same as that in the rMIB package. By taking $h = 0.5$, a line plot along the x direction for $y = 33.45$ and $z = 32.8$ is shown for both LPB and NPB potentials. It can be seen in Fig. 8 the potentials in both charts have a similar pattern, while the magnitudes of the rMIB potentials are higher. In the rMIB case, the LPB and NPB potentials are almost identical everywhere. For the super-Gaussian model, the difference between the LPB and NPB potentials is initially invisible in the solvent region Ω_e. However, such a difference grows in the transition region Ω_e and eventually becomes quite significant throughout the solute region Ω_i. Such a pattern agrees with that in the diffuse interface Kirkwood sphere, and confirms the above analysis that the transition layer Ω_t and the $(1 - S)$ term introduce the large difference between the NPB and LPB energies for the super-Gaussian PB model.

4. Concluding remarks

In this paper, several Poisson-Boltzmann (PB) energy functionals in terms of the dimensionless potentials have been formulated for water and vacuum phases and in both nonlinear and linearized settings. Moreover, alternative energy forms have been proposed to calculate the electrostatic free energy (EFE) without involving the electrostatic stress term. Rigorous mathematical derivation of such alternative energy forms is offered. While the alternative energy forms are computationally better than the definition based on energy functionals, they involve surface integrals, whose values are domain dependent. A systematic study has been conducted in this paper to evaluate such integrals and their dependence on domain size, as well as mesh step size. Our conclusion is that the inclusion of surface integrals of the dimensionless potential u in the water phase does not influence the calculated EFE too much, even for a tight domain. Thus, the surface integrals presented in the alternative energy forms are not recommended to be computed in practice, and only the first term and the volume integrals should be calculated for the EFE. We note that the proposed energy functionals and EFE forms are validated by considering diffuse interface and super-Gaussian PB models in this work, but they can be directly applied to other PB models, including the classical sharp interface or two-dielectric PB model. In fact, it is known that as the GCS parameter σ goes to zero, the surface function $S(\mathbf{r})$ will converge to a Heaviside function representing the solute-accessible surface (SAS). Under such a limiting process, the convergence of the diffuse interface PB model to the sharp interface PB model, including convergence of the electrostatic potential and free energy, has been rigorously proved in [36]. Following the same principal, all PB energy functionals and EFE forms developed in this work could be similarly formulated for the sharp interface PB model.

The singular charges in the source term of the PB equation can be discretized directly by the trilinear method, or can be treated analytical by using the modern regularization approaches [36,39–41]. The alternative energy forms in the regularization formulation have also been developed. A comparison between the trilinear and regularization methods for calculating the EFE has been carried out in solving the nonlinear PB (NPB) equation of the super-Gaussian PB model. Both methods perform well in various tests, and the convergence of the trilinear energy with respect to the domain size is usually better than that of the regularization method. However, when the error cancelation in the first term of the energy forms is lost in unsynchronized rotation or shifting, the trilinear method produces a huge error in the EFE calculation, while the regularization is free of the artificial grid energy in these studies.

For a set of 25 proteins, a comparison between the linearized PB (LPB) energies and the NPB energies have been considered for both the super-Gaussian PB model and the sharp interface PB model implemented in the rMIB package [20]. It is found that the LPB and NPB energies are very close in the sharp interface PB model, while they are quite different in the diffuse interface PB model [39] and super-Gaussian PB model [21,41]. In these smooth dielectric PB models, a transition layer Ω_t with width 3Å has been inserted as a smooth solute-solvent boundary, and the nonlinear hyperbolic Sine term has the coefficient $(1 - S)$, where the smooth surface function S changes from 0 to 1 in Ω_t. In some sense, the smooth solute-solvent boundary pushes the solute atoms and solvent ions away from each other. Physically, this reduces the electrostatic interactions between solute and solvent. Thus, the free energies of the diffuse interface and super-Gaussian models are weaker than those of the sharp interface, i.e., they have smaller magnitudes. On the other hand, it is interesting to observe in this study that the smooth solute-solvent boundary amplifies the difference between LPB and NPB potentials in Ω_t. Consequently, the NPB energy of the diffuse interface and super-Gaussian PB models is significantly different from the LPB energy. In the literature, the NPB model is always overshadowed by the LPB model, because their energies in the classical PB setting are very close. The present study indicates that the NPB model deserves more attention in the smooth dielectric setting.

Since many numerical studies of the PB equation are conducted in the normalized form in terms of the dimensionless potential, the proposed energy forms are expected to have an impact to the field. Our energy forms are designed such that one can directly plug in the numerically approximated dimensionless potential to calculate the energy in the correct unit. This avoids the potential artifact involved in converting the unit of the potential back and forth. For example, in the rMIB package [20], the normalized PB equation is solved for the dimensionless potential, and a conversion is conducted for calculating the EFE. In particular, the LPB energy of the rMIB package has been benchmarked by the Kirkwood sphere [20], and a second order convergence has been shown towards the analytical energy. For the nonlinear case, the Kirkwood sphere does not admit an analytical energy. In [2], a benchmark NPB energy calculated by the first term only is numerically generated by solving the one-dimensional (1D) NPB model. However, it is found in [2] that the rMIB NPB energy does not converge to the benchmark value. While the non-convergence of the rMIB NPB energy could be due to various reasons, the unit system and conversion could be an issue. Nevertheless, a full scale study on the rMIB package is beyond the scope of the present paper, and will be explored elsewhere.

Funding

This research was partially supported by the National Institutes of Health (NIH) under the grant R01GM093937. In addition, the research of Zhao was supported in part by the National Science Foundation (NSF) under grants DMS-2110914 and DMS-2306991.

CRediT authorship contribution statement

Shan Zhao: Conceptualization, Methodology, Software, Supervision, Writing – original draft. **Idowu E. Ijaodoro:** Software, Validation, Visualization, Writing – original draft. **Mark McGowan:** Software, Validation, Writing – review & editing. **Emil Alexov:** Conceptualization, Funding acquisition, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability and material

The protein structures analyzed in this study were downloaded from the protein databank https://www.rcsb.org, and processed according to the cited references.

References

[1] A. Abrashkin, D. Andelman, H. Orland, Dipolar Poisson-Boltzmann equation: ions and dipoles close to charge interfaces, Phys. Rev. Lett. 99 (7) (2007) 077801.
[2] S. Amihere, W. Geng, S. Zhao, Benchmarking electrostatic free energy of the nonlinear Poisson-Boltzmann model for the Kirkwood sphere, Commun. Inf. Syst. 22 (2022).
[3] N.A. Baker, Poisson-Boltzmann methods for biomolecular electrostatics, in: Numerical Computer Methods, Part D, in: Methods in Enzymology, vol. 383, Academic Press, 2004, pp. 94–118.
[4] N.A. Baker, Improving implicit solvent simulations: a Poisson-centric view, Curr. Opin. Struct. Biol. 15 (2) (2005) 137–143.
[5] P. Bates, Z. Chen, Y. Sun, G.-W. Wei, S. Zhao, Geometric and potential driving formation and evolution of biomolecular surfaces, J. Math. Biol. 59 (2) (2009) 193–231.
[6] P.W. Bates, G.-W. Wei, S. Zhao, Minimal molecular surfaces and their applications, J. Comput. Chem. 29 (3) (2008) 380–391.
[7] A. Chakravorty, Z. Jia, L. Li, S. Zhao, E. Alexov, Reproducing the ensemble average polar solvation energy of a protein from a single structure: Gaussian-based smooth dielectric function for macromolecular modeling, J. Chem. Theory Comput. 14 (2) (2018) 1020–1032.
[8] D. Chen, Z. Chen, C. Chen, W. Geng, G.W. Wei, MIBPB: a software package for electrostatic analysis, J. Comput. Chem. 32 (2011) 657–670.
[9] J. Chen, J. Hu, Y. Xu, R. Krasny, W. Geng, Computing protein pkas using the tabi Poisson–Boltzmann solver, J. Comput. Biophys. Chem. 20 (02) (2021) 175–187.
[10] L. Chen, M.J. Holst, J. Xu, The finite element approximation of the nonlinear Poisson–Boltzmann equation, SIAM J. Numer. Anal. 45 (6) (2007) 2298–2320.
[11] Z. Chen, N.A. Baker, G. Wei, Differential geometry based solvation model I: Eulerian formulation, J. Comput. Phys. 229 (22) (2010) 8231–8258.
[12] Z. Chen, N.A. Baker, G.W. Wei, Differential geometry based solvation model II: Lagrangian formulation, J. Math. Biol. 63 (6) (2011) 1139–1200.
[13] L.-T. Cheng, J. Dzubiella, J.A. McCammon, B. Li, Application of the level-set method to the implicit solvation of nonpolar molecules, J. Chem. Phys. 127 (8) (2007) 084503.
[14] I.-L. Chern, J.-G. Liu, W.-C. Wang, et al., Accurate evaluation of electrostatics for macromolecules in solution, Methods Appl. Anal. 10 (2) (2003) 309–328.
[15] R. Chowdhury, R. Egan, D. Bochkov, F. Gibou, Efficient calculation of fully resolved electrostatics around large biomolecules, J. Comput. Phys. 448 (2022) 110718.
[16] S. Dai, B. Li, J. Lu, Convergence of phase-field free energy and boundary force for molecular solvation, Arch. Ration. Mech. Anal. 227 (1) (2018) 105–147.
[17] W. Deng, J. Xu, S. Zhao, On developing stable finite element methods for pseudo-time simulation of biomolecular electrostatics, J. Comput. Appl. Math. 330 (2018) 456–474.
[18] H. Feng, S. Zhao, Fft-based high order central difference schemes for three-dimensional Poisson's equation with various types of boundary conditions, J. Comput. Phys. 410 (2020) 109391.
[19] W. Geng, S. Yu, G. Wei, Treatment of charge singularities in implicit solvent models, J. Chem. Phys. 127 (11) (2007) 114106.
[20] W. Geng, S. Zhao, A two-component matched interface and boundary (mib) regularization for charge singularity in implicit solvation, J. Comput. Phys. 351 (2017) 25–39.
[21] T. Hazra, S.A. Ullah, S. Wang, E. Alexov, S. Zhao, A super-Gaussian Poisson-Boltzmann model for electrostatic free energy calculation: smooth dielectric distribution for protein cavities and in both water and vacuum states, J. Math. Biol. 79 (2) (2019) 631–672.
[22] M.J. Holst, The Poisson-Boltzmann Equation: Analysis and Multilevel Numerical Solution, PhD thesis, UIUC, 1994.
[23] B. Honig, A. Nicholls, Classical electrostatics in biology and chemistry, Science 268 (5214) (1995) 1144–1149.
[24] P. Koehl, F. Poitevin, H. Orland, M. Delarue, Modified Poisson–Boltzmann equations for characterizing biomolecular solvation, J. Theor. Comput. Chem. 13 (03) (2014) 1440001.
[25] A. Kucherova, S. Strango, S. Sukenik, M. Theillard, Computational modeling of protein conformational changes - application to the opening sars-cov-2 spike, J. Comput. Phys. 444 (2021) 110591.
[26] A. Lee, W. Geng, S. Zhao, Regularization methods for the Poisson-Boltzmann equation: comparison and accuracy recovery, J. Comput. Phys. 426 (2021) 109958.
[27] C. Li, L. Li, J. Zhang, E. Alexov, Highly efficient and exact method for parallelization of grid-based algorithms and its implementation in delphi, J. Comput. Chem. 33 (24) (2012) 1960–1966.
[28] L. Li, C. Li, E. Alexov, On the modeling of polar component of solvation energy using smooth Gaussian-based dielectric function, J. Theor. Comput. Chem. 13 (03) (2014) 1440002.
[29] L. Li, C. Li, Z. Zhang, E. Alexov, On the dielectric constant of proteins: smooth dielectric function for macromolecular modeling and its implementation delphi, J. Chem. Theory Comput. 9 (4) (2013) 2126–2136.
[30] B. Liu, B. Wang, R. Zhao, Y. Tong, G.-W. Wei, ESES: software for Eulerian solvent excluded surface, J. Comput. Chem. 38 (7) (2017) 446–466.
[31] A.E. Lopez-Hernandez, Y. Xie, W. Guo, L. Li, The electrostatic features of Dengue virus capsid assembly, J. Comput. Biophys. Chem. 20 (02) (2021) 201–207.
[32] A.M. Micu, B. Bagheri, A.V. Ilin, R. Scott, B. Pettitt, Numerical considerations in the computation of the electrostatic free energy of interaction within the Poisson-Boltzmann theory, J. Comput. Phys. 136 (2) (1997) 263–271.
[33] M. Mirzadeh, M. Theillard, A. Helgadottir, D. Boy, F. Gibou, An adaptive, finite difference solver for the nonlinear Poisson-Boltzmann equation with applications to biomolecular computations, Commun. Comput. Phys. 13 (1) (2013) 150–173.
[34] A. Nicholls, B. Honig, A rapid finite difference algorithm, utilizing successive over-relaxation to solve the Poisson-Boltzmann equation, J. Comput. Chem. 12 (4) (1991) 435–445.
[35] S.K. Panday, M.H. Shashikala, A. Chakravorty, S. Zhao, E. Alexov, Reproducing ensemble averaged electrostatics with super-Gaussian-based smooth dielectric function: application to electrostatic component of binding energy of protein complexes, Commun. Inf. Syst. 19 (4) (2019).
[36] Y. Shao, M. McGowan, S. Wang, E. Alexov, S. Zhao, Convergence of a diffuse interface Poisson-Boltzmann (pb) model to the sharp interface pb model: a unified regularization formulation, Appl. Math. Comput. 436 (2023) 127501.
[37] K.A. Sharp, B. Honig, Calculating total electrostatic energies with the nonlinear Poisson-Boltzmann equation, J. Phys. Chem. 94 (19) (1990) 7684–7692.
[38] K.A. Sharp, B. Honig, Electrostatic interactions in macromolecules - theory and applications, Annu. Rev. Biophys. Biophys. Chem. 19 (1990) 301–332.
[39] S. Wang, E. Alexov, S. Zhao, On regularization of charge singularities in solving the Poisson-Boltzmann equation with a smooth solute-solvent boundary, Math. Biosci. Eng. 18 (2) (2021) 1370–1405.

[40] S. Wang, A. Lee, E. Alexov, S. Zhao, A regularization approach for solving Poisson's equation with singular charge sources and diffuse interfaces, Appl. Math. Lett. 102 (2020) 106144.
[41] S. Wang, Y. Shao, E. Alexov, S. Zhao, A regularization approach for solving the super-Gaussian Poisson-Boltzmann model with heterogeneous dielectric functions, J. Comput. Phys. 464 (2022) 111340.
[42] G.-W. Wei, Differential geometry based multiscale models, Bull. Math. Biol. 72 (6) (2010) 1562–1622.
[43] L. Wilson, W. Geng, R. Krasny, Tabi-pb 2.0: an improved version of the treecode-accelerated boundary integral Poisson-Boltzmann solver, J. Phys. Chem. B 126 (37) (2022) 7104–7113.
[44] D. Xie, New solution decomposition and minimization schemes for Poisson-Boltzmann equation in calculation of biomolecular electrostatics, J. Comput. Phys. 275 (2014) 294–309.
[45] D. Xie, J. Li, A new analysis of electrostatic free energy minimization and Poisson-Boltzmann equation for protein in ionic solvent, Nonlinear Anal., Real World Appl. 21 (2015) 185–196.
[46] S. Zhao, Pseudo-time-coupled nonlinear models for biomolecular surface representation and solvation analysis, Int. J. Numer. Methods Biomed. Eng. 27 (12) (2011) 1964–1981.
[47] S. Zhao, Operator splitting adi schemes for pseudo-time coupled nonlinear solvation simulations, J. Comput. Phys. 257 (2014) 1000–1021.
[48] S. Zhao, G.W. Wei, Matched interface and boundary (mib) for the implementation of boundary conditions in high-order central finite differences, Int. J. Numer. Methods Eng. 77 (12) (2009) 1690–1730.
[49] Y. Zhao, Y.-Y. Kwan, J. Che, B. Li, J.A. McCammon, Phase-field approach to implicit solvation of biomolecules with Coulomb-field approximation, J. Chem. Phys. 139 (2) (2013) 024111.
[50] Z. Zhou, P. Payne, M. Vasquez, N. Kuhn, M. Levitt, Finite-difference solution of the Poisson–Boltzmann equation: complete elimination of self-energy, J. Comput. Chem. 17 (11) (1996) 1344–1351.