

Research

**Cite this article:** Huang Z, Li B. 2024Variational implicit solvation with Legendre-transformed Poisson–Boltzmann electrostatics. *Proc. R. Soc. A* **480**: 20230731. <https://doi.org/10.1098/rspa.2023.0731>

Received: 4 October 2023

Accepted: 26 February 2024

Subject Areas:

applied mathematics, mathematical modelling, computational physics

Keywords:

variational implicit-solvent model, Poisson–Boltzmann theory, Legendre transform, dielectric boundary force, penalty method, a min–min optimization algorithm

Author for correspondence:

Bo Li

e-mail: bli@math.ucsd.eduElectronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.7126153>.

Variational implicit solvation with Legendre-transformed Poisson–Boltzmann electrostatics

Zunding Huang¹ and Bo Li²¹Department of Mathematics, University of California, San Diego, CA, USA²Department of Mathematics and Ph.D. Program in Quantitative Biology, University of California, San Diego, CA, USA

BL, 0000-0002-9183-4654

The variational implicit-solvent model (VISM) is an efficient approach to biomolecular interactions, where electrostatic interactions are crucial. The total VISM free energy of a dielectric boundary (i.e. solute–solvent interface) consists of the interfacial energy, solute–solvent interaction energy and dielectric electrostatic energy. The last part is the maximum value of the classical and concave Poisson–Boltzmann (PB) energy functional of electrostatic potentials, with the maximizer being the equilibrium electrostatic potential governed by the PB equation. For the consistency of energy minimization and computational stability, here we propose alternatively to minimize the convex Legendre-transformed Poisson–Boltzmann (LTPB) electrostatic energy functional of all dielectric displacements constrained by Gauss' Law in the solute region. Both integrable and discrete solute charge densities are treated, and the duality of the LTPB and PB functionals is established. A penalty method is designed for the constrained minimization of the LTPB functional. In application to biomolecular interactions, we minimize the total VISM free energy iteratively, while in each step of such iteration, minimize the LTPB energy. Convergence of such a min–min algorithm is shown. Our numerical results on the solvation of a single ion indicate that the LTPB performs better than the PB formulation, providing possibilities for efficient biomolecular simulations.

1. Introduction

In a variational implicit-solvent model (VISM) [1–4] (cf. also related models [5–7]), one minimizes a solvation free-energy functional of all possible solute–solvent interfaces (i.e. dielectric boundaries) to determine an equilibrium system of charged molecules (such as proteins) immersed in an aqueous solvent (i.e. water or salted water) and estimate the solvation free energy. The VISM free energy consists of the solute–solvent interfacial energy, the solute–solvent interaction energy and the electrostatic energy. The last part is often described through the equilibrium electrostatic potential $\phi_\Gamma : \Omega \rightarrow \mathbb{R}$, which is governed by the dielectric boundary Poisson–Boltzmann (PB) equation (PBE) [3,8–12]

$$\nabla \cdot \varepsilon_\Gamma \nabla \phi_\Gamma - \chi_+ B'(\phi_\Gamma) = -f \quad \text{in } \Omega, \quad (1.1)$$

together with some boundary conditions. Here, $\Omega \subset \mathbb{R}^3$ is the underlying solvation region, Γ is the dielectric boundary that divides Ω into the solute region Ω_- and the solvent region Ω_+ , $\chi_+ = \chi_{\Omega_+}$ is the characteristic function of Ω_+ , and $\varepsilon_\Gamma : \Omega \rightarrow \mathbb{R}$ is the dielectric coefficient defined by

$$\varepsilon_\Gamma(x) = \begin{cases} \varepsilon_- & \text{if } x \in \Omega_-, \\ \varepsilon_+ & \text{if } x \in \Omega_+, \end{cases} \quad (1.2)$$

where ε_- and ε_+ are the dielectric permittivities for the solute and solvent, respectively. (Typically, $\varepsilon_- \approx \varepsilon_0$ and $\varepsilon_+ \approx 80\varepsilon_0$ with ε_0 the vacuum permittivity.) See figure 1. The function $f : \Omega \rightarrow \mathbb{R}$ represents the charge density of solute molecules, while the term $-B'(\phi_\Gamma)$ describes the ensemble-averaged charge density of mobile ions in the solvent. The function $B : \mathbb{R} \rightarrow \mathbb{R}$ is given by

$$B(\phi) = \beta^{-1} \sum_{j=1}^M c_j^\infty (e^{-\beta q_j \phi} - 1), \quad (1.3)$$

where $M \geq 1$ is the number of ionic species, $\beta^{-1} = k_B T$ with k_B the Boltzmann constant and T temperature, $q_j = Z_j e$ with e the elementary charge, and c_j^∞ and Z_j are the bulk ionic concentration and valence of ions of the j th species. We note that different forms of the function B can be used for different models; cf. e.g. [13–16]. The equilibrium electrostatic potential ϕ_Γ , the optimal boundary Γ and the VISM free energy depend on the form of B and the parameters used in defining B . A commonly used boundary condition for the electrostatic potentials is $\phi = g$ on $\partial\Omega$ with $g : \partial\Omega \rightarrow \mathbb{R}$ a given function.

The PBE (1.1) is the Euler–Lagrange equation of the classical PB electrostatic energy functional applied to the continuum solvation [11,12,15,17–20]

$$I_\Gamma[\phi] = \int_\Omega \left[-\frac{\varepsilon_\Gamma}{2} |\nabla \phi|^2 + f\phi - \chi_+ B(\phi) \right] dx \quad \forall \phi \in H_g^1(\Omega), \quad (1.4)$$

where

$$H_g^1(\Omega) = \{\phi \in H^1(\Omega) : \phi = g \text{ on } \partial\Omega\}.$$

This functional is concave and maximized to yield the equilibrium electrostatic potential ϕ_Γ , which is the solution to the PBE, and the corresponding electrostatic energy $E_{\text{ele}}[\Gamma] = I_\Gamma[\phi_\Gamma]$. We refer to [12,21] for discussions on the maximization instead of minimization of the electrostatic energy functional for equilibrium electrostatics.

An iterative method is often used to minimize numerically the VISM solvation free-energy functional. In each step of the iteration, one solves the PBE (1.1) or maximizes the PB functional (1.4). Maximizing the electrostatic energy and then minimizing the total VISM solvation free energy may possibly develop instabilities if there are not enough steps for such maximization or minimization. It is therefore natural to ask if the electrostatics can be determined by minimizing a convex energy functional. Motivated by such a question, here we develop an alternative approach

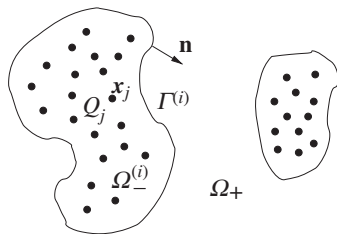


Figure 1. A schematic diagram of charged molecules immersed in an aqueous solvent. The region of solvation Ω is divided by the solute-solvent interface (i.e. dielectric boundary) Γ into the solute region Ω_- and the solvent region Ω_+ . The solute region Ω_- , which can have multiple connected subregions, contains all the solute atoms x_i carrying partial charges Q_i ($i = 1, \dots, N$). The unit normal n at Γ points from Ω_- to Ω_+ .

to the electrostatics for VISM, based on the concept of the Legendre-transformed Poisson-Boltzmann (LTPB) electrostatic energy [22] (cf. also [21,23–25]). For a given dielectric boundary Γ , the LTPB energy functional is given by Ciotti & Li [21]

$$J_\Gamma[D] = \int_\Omega \left[\frac{1}{2\epsilon_\Gamma} |D|^2 + \chi_+ B^*(f - \nabla \cdot D) \right] dx + \int_{\partial\Omega} g(D \cdot n) dS \quad (1.5)$$

for all dielectric displacements $D : \Omega \rightarrow \mathbb{R}^3$ that are constrained by Gauss' Law (in the differential form)

$$\nabla \cdot D = f \quad \text{in } \Omega_-, \quad (1.6)$$

where B^* is the Legendre transform of B and n is the unit exterior normal at $\partial\Omega$. We recall for $\xi \in \mathbb{R}$ that [26,27]

$$B^*(\xi) = \sup_{a \in \mathbb{R}} [a\xi - B(a)] = s\xi - B(s) \text{ with } B'(s) = \xi, \quad \text{and } B^{*'}(\xi) = s. \quad (1.7)$$

Our main results are the following:

- (1) We construct the LTPB electrostatic energy functional with the constraint for both the case of a continuum solute charge density represented by an integrable function (cf. (1.5)) and that of a discrete charge density (or point charges) described by a linear combination of Dirac masses. We prove the duality between the classical PB and the LTPB functionals. For a continuum charge density $f \in L^2(\Omega)$, this duality is $\max_{\phi \in H_g^1(\Omega)} I_\Gamma[\phi] = \min_{D \in V_{\Gamma,f}} J_\Gamma[D]$, and the unique maximizer ϕ_Γ of I_Γ over $H_g^1(\Omega)$ and the unique minimizer D_Γ of J_Γ over $V_{\Gamma,f}$ are related by $D_\Gamma = -\epsilon_\Gamma \nabla \phi_\Gamma$, where

$$V_{\Gamma,f} = \{D \in H(\text{div}, \Omega) : \nabla \cdot D = f \text{ in } \Omega_-\} \quad (1.8)$$

and

$$H(\text{div}, \Omega) = \{D \in [L^2(\Omega)]^3 : \nabla \cdot D \in L^2(\Omega)\}. \quad (1.9)$$

- (2) To minimize numerically the LTPB functional, we propose a penalty method. For the case of a continuum charge density $f \in L^2(\Omega)$, this method amounts to minimizing the penalized functional

$$J_{\Gamma,\mu}[D] = J_\Gamma[D] + \frac{1}{2\mu} \int_{\Omega_-} |\nabla \cdot D - f|^2 dx \quad \forall D \in H(\text{div}, \Omega), \quad (1.10)$$

without the constraint, where $\mu > 0$ is a penalty parameter and χ_- is the characteristic function of Ω_- . We prove that, as $\mu \rightarrow 0$, the minimizer and minimum values of the penalized functional $J_{\Gamma,\mu}$ converge to those for the functional $J_\Gamma : V_{\Gamma,f} \rightarrow \mathbb{R}$, respectively. Such convergence is numerically verified.

- (3) We incorporate the LTPB electrostatics into the VISM, and derive the dielectric boundary force $-\delta_\Gamma(\min_D J_\Gamma[D])$ using the minimizing dielectric displacement $D_\Gamma \in V_{\Gamma,f}$. We also construct a min–min algorithm and a max–min algorithm to minimize numerically the total VISM free energy with the LTPB and PB electrostatics, respectively. The convergence of the min–min optimization algorithm is shown.
- (4) We present an analysis of the duality and the penalty method for a simplified radially symmetric system resulting from the application of VISM to the solvation of a single ion. We also provide a new and direct derivation of the dielectric boundary force for a similar and reduced one-dimensional system. Moreover, we report extensive numerical results on the solvation of a single ion to show that the LTPB formulation with the min–min algorithm performs better than the PB formulation with the max–min algorithm and that the VISM-LTPB predicts accurately the solvation free energy of single ions.

We remark that our results can be directly extended to some size-modified PB and LTPB electrostatics with a general convex function B [13–16,28–35].

The rest of this paper is organized as follows: in §2, we construct the LTPB electrostatic functionals and prove the duality of the LTPB and the PB functionals. We also propose and prove the convergence of a penalty method for minimizing the LTPB energy functional and present some numerical results. In §3, we apply the LTPB theory to the variational implicit solvation and derive the dielectric boundary force. We also design a min–min algorithm and prove its convergence for minimizing the VISM-LTPB functional. Numerical results for the solvation of an ion are presented. Finally, in §4, we draw conclusions of our findings. In appendix A, we give a new and direct derivation of the dielectric boundary force for a one-dimensional model system.

2. The LTPB electrostatics with a dielectric boundary

We assume the following:

- (A 1) All Ω , Ω_- , Ω_+ are smooth and bounded open sets in \mathbb{R}^3 , $\Gamma = \partial\Omega_-$ and all $x_i \in \Omega_-$ and $Q_i \in \mathbb{R}$ ($i = 1, \dots, N$) are given; cf. figure 1;
- (A 2) Both ε_- and ε_+ are given, distinct positive numbers, and ε_Γ is defined by $\varepsilon_\Gamma = \varepsilon_\pm$ in Ω_\pm ;
- (A 3) The function $B: \mathbb{R} \rightarrow \mathbb{R}$ is smooth, strictly convex and uniquely minimized at 0 with $B(0) = 0$. Moreover, $B(\pm\infty) = \infty$; (an example of such a function is $B(s) = \cosh(s) - 1$).
- (A 4) The function $f \in L^2(\Omega)$ is given, and is smooth, e.g. $f \in H^1(\Omega)$. The boundary value $g: \partial\Omega \rightarrow \mathbb{R}$ (cf. (1.4) and (1.5)) is the trace of some function, also denoted g , in $W^{1,\infty}(\Omega)$.

We note that the function $B: \mathbb{R} \rightarrow \mathbb{R}$ defined in (1.3) and those that model the ionic size effect in the size-modified PB theory [15,16,34,35] all satisfy the assumption (A 3). We also recall that the space $H(\text{div}, \Omega)$ (cf. (1.9)) is a Hilbert space with respect to the inner product

$$\langle D_1, D_2 \rangle = \int_{\Omega} [D_1 \cdot D_2 + (\nabla \cdot D_1)(\nabla \cdot D_2)] dx.$$

Moreover, $D \cdot n \in L^2(\partial\Omega)$ if $D \in H(\text{div}, \Omega)$ and [36]

$$\int_{\Omega} (\nabla \cdot D) u \, dx = - \int_{\Omega} D \cdot \nabla u \, dx + \int_{\partial\Omega} (D \cdot n) u \, dS \quad \forall u \in H^1(\Omega). \quad (2.1)$$

(a) The electrostatic energy functionals and the duality

Case 1. A continuum charge density. For this case, the PB and the LTPB functionals $I_\Gamma: H_g^1(\Omega) \rightarrow \mathbb{R} \cup \{-\infty\}$ and $J_\Gamma: V_{\Gamma,f} \rightarrow \mathbb{R} \cup \{\infty\}$ are defined in (1.4) and (1.5), respectively, and their properties are summarized in the next theorem. We denote $\llbracket u \rrbracket_\Gamma = u|_{\Omega_+} - u|_{\Omega_-}$ on Γ for any $u: \Omega \rightarrow \mathbb{R}$ when the traces are defined.

Theorem 2.1.

- (1) The PB functional $I_\Gamma : H_g^1(\Omega) \rightarrow \mathbb{R} \cup \{-\infty\}$ admits a unique maximizer $\phi_\Gamma \in H_g^1(\Omega)$. Moreover, $\phi_\Gamma \in L^\infty(\Omega)$, $\phi_\Gamma|_{\Omega_\pm} \in H^2(\Omega_\pm)$, and $\llbracket \varepsilon_\Gamma \partial_n \phi_\Gamma \rrbracket|_\Gamma = 0$, and ϕ_Γ is the unique weak solution in $H_g^1(\Omega)$ to the PB equation (PBE) (1.1).
- (2) Let $D_\Gamma = -\varepsilon_\Gamma \nabla \phi_\Gamma$. Then, $D_\Gamma \in V_{\Gamma,f}$ and it is the unique minimizer of the LTPB functional $J_\Gamma : V_{\Gamma,f} \rightarrow \mathbb{R} \cup \{\infty\}$. Moreover, $\llbracket D_\Gamma \cdot n \rrbracket_\Gamma = 0$ and $B^{*'}(f - \nabla \cdot D_\Gamma) = \phi_\Gamma$ in Ω_+ .
- (3) Duality: $\max_{\phi \in H_g^1(\Omega)} I_\Gamma[\phi] = \min_{D \in V_{\Gamma,f}} J_\Gamma[D]$.

Proof. (1) By Poincaré's inequality, we have $\sup_{\phi \in H_g^1(\Omega)} I_\Gamma[\phi] < \infty$. The existence and uniqueness of a maximizer ϕ_Γ for $I_\Gamma : H_g^1(\Omega) \rightarrow \mathbb{R} \cup \{-\infty\}$ can be obtained by the direct method in the calculus of variations, using the strict convexity of B . A comparison argument leads to $\phi_\Gamma \in L^\infty(\Omega)$ [18]. Routine calculations then imply that the maximizer ϕ_Γ satisfies the PBE (1.1), which is the Euler–Lagrange equation for the functional I_Γ , and that $\llbracket \varepsilon_\Gamma \partial_n \phi_\Gamma \rrbracket_\Gamma = 0$. The regularity of ϕ_Γ follows from existing results of solution regularity for elliptic equations; cf. [18,19].

(2) For any $\phi \in H_g^1(\Omega)$ and any $D \in V_{\Gamma,f}$, we have by (2.1) that

$$\begin{aligned} I_\Gamma[\phi] &\leq \int_\Omega \left[-\frac{\varepsilon_\Gamma}{2} |\nabla \phi|^2 + f\phi - \chi_+ B(\phi) + \frac{1}{2\varepsilon_\Gamma} |\varepsilon_\Gamma \nabla \phi + D|^2 \right] dx \\ &= \int_\Omega \left[\frac{|D|^2}{2\varepsilon_\Gamma} + f\phi - \chi_+ B(\phi) + \nabla \phi \cdot D \right] dx \\ &= \int_\Omega \left[\frac{|D|^2}{2\varepsilon_\Gamma} + \chi_+ (\phi(f - \nabla \cdot D) - B(\phi)) \right] dx + \int_{\partial\Omega} g(D \cdot n) dS \\ &\leq \int_\Omega \left[\frac{|D|^2}{2\varepsilon_\Gamma} + \chi_+ B^{*'}(f - \nabla \cdot D) \right] dx + \int_{\partial\Omega} g(D \cdot n) dS \\ &= J_\Gamma[D]. \end{aligned}$$

Note by (1.1) that $D_\Gamma = -\varepsilon_\Gamma \nabla \phi_\Gamma \in V_{\Gamma,f}$. Setting $\phi = \phi_\Gamma$ and $D = D_\Gamma$ and noting by (1.1) that $f - \nabla \cdot D_\Gamma = B'(\phi_\Gamma)$ and hence $B^{*'}(f - \nabla \cdot D_\Gamma) = \phi_\Gamma(f - \nabla \cdot D_\Gamma) - B(\phi_\Gamma)$ in Ω_+ , we have $I_\Gamma[\phi_\Gamma] = J_\Gamma[D_\Gamma]$. Thus, D_Γ is a minimizer of J_Γ over $V_{\Gamma,f}$. It is unique since J_Γ is strictly convex over $V_{\Gamma,f}$. Since $\llbracket \varepsilon_\Gamma \partial_n \phi_\Gamma \rrbracket_\Gamma = 0$, we have $\llbracket D_\Gamma \cdot n \rrbracket_\Gamma = 0$. By the PBE (1.1), we also have $f - \nabla \cdot D_\Gamma = \chi_+ B(\phi_\Gamma)$ in Ω_+ . Hence, by the property of Legendre transforms, $B^{*'}(f - \nabla \cdot D_\Gamma) = \phi_\Gamma$ in Ω_+ .

(3) This follows from $I_\Gamma[\phi_\Gamma] = \max_{\phi \in H_g^1(\Omega)} I_\Gamma[\phi] \leq \min_{D \in V_{\Gamma,f}} J_\Gamma[D] = J_\Gamma[D_\Gamma] = I_\Gamma[\phi_\Gamma]$. ■

Case 2. A discrete charge density. The equilibrium electrostatic potential, now denoted $\hat{\phi}_\Gamma$, is the weak solution to the boundary-value problem of the PBE with point charges [3,15,19,37]

$$\begin{cases} \nabla \cdot \varepsilon_\Gamma \nabla \hat{\phi}_\Gamma - \chi_+ B'(\hat{\phi}_\Gamma) = -\sum_{i=1}^N Q_i \delta_{x_i} & \text{in } \Omega, \\ \hat{\phi}_\Gamma = g & \text{on } \partial\Omega, \end{cases} \quad (2.2)$$

where δ_a denotes the Dirac mass concentrated on $a \in \mathbb{R}^3$. The electrostatic energy is given in [3,15,19]

$$E_{\text{ele}}[\Gamma] = \frac{1}{2} \sum_{i=1}^N Q_i (\hat{\phi}_\Gamma - \hat{\phi}_C)(x_i) + \int_{\Omega_+} \left[\frac{1}{2} \hat{\phi}_\Gamma B'(\hat{\phi}_\Gamma) - B(\hat{\phi}_\Gamma) \right] dx, \quad (2.3)$$

where $\hat{\phi}_C$ is the Coulomb potential: $\hat{\phi}_C(x) = \sum_{i=1}^N Q_i / (4\pi\varepsilon - |x - x_i|)$. Note that we do not include an extra term in B as we can derive the energy form by minimizing a continuum electrostatic free-energy functional of ionic concentrations using different boundary conditions (e.g. the homogeneous Neumann boundary condition) for electrostatic potentials. The boundary value g

in (2.2) is an approximation of that of the equilibrium electrostatic potential; cf. [3,15,19,38,39]. Following [15,19], we express the electrostatic energy using integrals that are mathematically more convenient to handle

$$E_{\text{ele}}[\Gamma] = - \int_{\Omega} \left[\frac{\varepsilon_{\Gamma}}{2} |\nabla(\hat{\phi}_{\Gamma} - \hat{\phi}_{\Gamma,\infty})|^2 + \chi_+ B(\hat{\phi}_{\Gamma}) \right] dx + A_{\Gamma} \quad (2.4)$$

and

$$A_{\Gamma} := \frac{\varepsilon_- - \varepsilon_+}{2} \int_{\Omega_+} \nabla \hat{\phi}_{\Gamma,\infty} \cdot \nabla \hat{\phi}_0 dx + \frac{1}{2} \sum_{i=1}^N Q_i (\hat{\phi}_{\infty} - \hat{\phi}_C)(x_i). \quad (2.5)$$

Here, $\hat{\phi}_{\Gamma,\infty}$ is defined by $-\nabla \cdot \varepsilon_{\Gamma} \nabla \hat{\phi}_{\Gamma,\infty} = \sum_{i=1}^N Q_i \delta_{x_i}$ in Ω and $\hat{\phi}_{\Gamma,\infty} = g$ on $\partial\Omega$, and both $\hat{\phi}_0$ and $\hat{\phi}_{\infty}$ are solutions to $-\varepsilon_- \Delta u = \sum_{i=1}^N Q_i \delta_{x_i}$ in Ω with $\hat{\phi}_0 = 0$ and $\hat{\phi}_{\infty} = g$ on $\partial\Omega$, respectively; cf. [19]. Note that $\hat{\phi}_{\infty} - \hat{\phi}_C$ is harmonic in Ω_- so each $(\hat{\phi}_{\infty} - \hat{\phi}_C)(x_i)$ ($1 \leq i \leq N$) is well defined as the limit as $x \rightarrow x_i$.

We define the PB energy functional for the case of point charges $\hat{I}_{\Gamma} : \hat{\phi}_{\Gamma,\infty} + H_0^1(\Omega) \rightarrow \mathbb{R} \cup \{-\infty\}$ by

$$\hat{I}_{\Gamma}[\phi] = - \int_{\Omega} \left[\frac{\varepsilon_{\Gamma}}{2} |\nabla(\phi - \hat{\phi}_{\Gamma,\infty})|^2 + \chi_+ B(\phi) \right] dx + A_{\Gamma} \quad \forall \phi \in \hat{\phi}_{\Gamma,\infty} + H_0^1(\Omega). \quad (2.6)$$

Note that $\hat{I}_{\Gamma}[\hat{\phi}_{\Gamma}] = E_{\text{ele}}[\Gamma]$ is the electrostatic energy. Note also that $\phi \in \hat{\phi}_{\Gamma,\infty} + H_0^1(\Omega)$ if and only if $\phi \in \hat{\phi}_C + H_h^1(\Omega)$ with $h = g - \hat{\phi}_C$. Let $V_{\Gamma,0}$ be defined by (1.8) with $f = 0$. We define the LTPB functional $\hat{J}_{\Gamma} : V_{\Gamma,0} \rightarrow \mathbb{R} \cup \{\infty\}$ by

$$\hat{J}_{\Gamma}[D] = \int_{\Omega} \left\{ \frac{1}{2\varepsilon_{\Gamma}} |D|^2 dx + \chi_+ [B^*(-\nabla \cdot D) + \hat{\phi}_{\Gamma,\infty} \nabla \cdot D] \right\} dx + A_{\Gamma} \quad \forall D \in V_{\Gamma,0}. \quad (2.7)$$

Theorem 2.2.

- (1) The PB functional $\hat{I}_{\Gamma} : \hat{\phi}_{\Gamma,\infty} + H_0^1(\Omega) \rightarrow \mathbb{R} \cup \{-\infty\}$ admits a unique maximizer $\hat{\phi}_{\Gamma}$. Moreover, $\hat{\phi}_{\Gamma} - \hat{\phi}_{\Gamma,\infty} \in L^{\infty}(\Omega)$, $(\hat{\phi}_{\Gamma} - \hat{\phi}_{\Gamma,\infty})|_{\Omega_{\pm}} \in H^2(\Omega_{\pm})$, and $\llbracket \varepsilon_{\Gamma} \partial_n \hat{\phi}_{\Gamma} \rrbracket_{\Gamma} = 0$, and $\hat{\phi}_{\Gamma}$ is the unique weak solution to the boundary-value problem of the PBE (2.2).
- (2) Let $\hat{D}_{\Gamma} = -\varepsilon_{\Gamma} \nabla(\hat{\phi}_{\Gamma} - \hat{\phi}_{\Gamma,\infty})$. Then, $\hat{D}_{\Gamma} \in V_{\Gamma,0}$ and it is the unique minimizer of the LTPB functional $\hat{J}_{\Gamma} : V_{\Gamma,0} \rightarrow \mathbb{R} \cup \{\infty\}$. Moreover, $B^{*'}(-\nabla \cdot \hat{D}_{\Gamma}) = \hat{\phi}_{\Gamma}$ in Ω_+ .
- (3) Duality: $\max_{\phi \in \hat{\phi}_{\Gamma,\infty} + H_0^1(\Omega)} \hat{I}_{\Gamma}[\phi] = \min_{D \in V_{\Gamma,0}} \hat{J}_{\Gamma}[D]$.

Proof. The proof of part (1) and part (3) is similar to that for theorem 2.1. So, we only prove part (2). Let $\phi \in \hat{\phi}_{\Gamma,\infty} + H_0^1(\Omega)$ and $D \in V_{\Gamma,0}$. Then,

$$\begin{aligned} \hat{I}_{\Gamma}[\phi] &= \int_{\Omega} \left[-\frac{\varepsilon_{\Gamma}}{2} |\nabla(\phi - \hat{\phi}_{\Gamma,\infty})|^2 + \frac{1}{2\varepsilon_{\Gamma}} |\varepsilon_{\Gamma} \nabla(\phi - \hat{\phi}_{\Gamma,\infty}) + D|^2 - \chi_+ B(\phi) \right] dx + A_{\Gamma} \\ &= \int_{\Omega} \left[\frac{|D|^2}{2\varepsilon_{\Gamma}} + \nabla(\phi - \hat{\phi}_{\Gamma,\infty}) \cdot D - \chi_+ B(\phi) \right] dx + A_{\Gamma} \\ &= \int_{\Omega} \left[\frac{|D|^2}{2\varepsilon_{\Gamma}} - (\phi - \hat{\phi}_{\Gamma,\infty}) \nabla \cdot D - \chi_+ B(\phi) \right] dx + A_{\Gamma} \\ &= \int_{\Omega} \left\{ \frac{|D|^2}{2\varepsilon_{\Gamma}} + \chi_+ [\phi(-\nabla \cdot D) - B(\phi) + \hat{\phi}_{\Gamma,\infty} \nabla \cdot D] \right\} dx + A_{\Gamma} \\ &\leq \int_{\Omega} \left\{ \frac{1}{2\varepsilon_{\Gamma}} |D|^2 + \chi_+ [B^*(-\nabla \cdot D) + \hat{\phi}_{\Gamma,\infty} \nabla \cdot D] \right\} dx + A_{\Gamma} \\ &= \hat{J}_{\Gamma}[D]. \end{aligned}$$

By (2.2) and the definition of $\hat{\phi}_{\Gamma,\infty}$, we have $\hat{D}_{\Gamma} = -\varepsilon_{\Gamma} \nabla(\hat{\phi}_{\Gamma} - \hat{\phi}_{\Gamma,\infty}) \in V_{\Gamma,0}$. Setting $\phi = \hat{\phi}_{\Gamma}$ and $D = \hat{D}_{\Gamma}$ and noting that $-\nabla \cdot \hat{D}_{\Gamma} = B'(\hat{\phi}_{\Gamma})$ and hence $B^*(-\nabla \cdot \hat{D}_{\Gamma}) = (-\nabla \cdot \hat{D}_{\Gamma})\hat{\phi}_{\Gamma} - B(\hat{\phi}_{\Gamma})$ in Ω_+ , we have $\hat{I}_{\Gamma}[\hat{\phi}_{\Gamma}] = \hat{J}_{\Gamma}[\hat{D}_{\Gamma}]$. Thus, \hat{D}_{Γ} is a minimizer of \hat{J}_{Γ} over $V_{\Gamma,0}$. It is unique since \hat{J}_{Γ} is strictly

convex over $V_{\Gamma,0}$. Finally, it follows from the property of Legendre transforms that $B^{*'}(-\nabla \cdot \hat{D}_\Gamma) = \hat{\phi}_\Gamma$ in Ω_+ . ■

(b) A penalty method

Let $\mu > 0$. Recall that $J_{\Gamma,\mu}$ is defined in (1.10). Similarly, we define $\hat{J}_{\Gamma,\mu} : H(\operatorname{div}, \Omega) \rightarrow \mathbb{R} \cup \{+\infty\}$ by

$$\hat{J}_{\Gamma,\mu}[D] = \hat{J}_\Gamma[D] + \frac{1}{2\mu} \int_{\Omega_-} (\nabla \cdot D)^2 dx. \quad (2.8)$$

Theorem 2.3. *For each $\mu > 0$, the functional $J_{\Gamma,\mu}$ (resp. $\hat{J}_{\Gamma,\mu}$) : $H(\operatorname{div}, \Omega) \rightarrow \mathbb{R} \cup \{+\infty\}$ admits a unique minimizer $D_{\Gamma,\mu}$ (resp. $\hat{D}_{\Gamma,\mu}$) $\in H(\operatorname{div}, \Omega)$. Moreover, $D_{\Gamma,\mu} \rightarrow D_\Gamma$ (resp. $\hat{D}_{\Gamma,\mu} \rightarrow \hat{D}_\Gamma$) in $H(\operatorname{div}, \Omega)$ and $\min_{D \in H(\operatorname{div}, \Omega)} J_{\Gamma,\mu}[D] \rightarrow \min_{D \in V_{\Gamma,f}} J_\Gamma[D]$ (resp. $\min_{D \in H(\operatorname{div}, \Omega)} \hat{J}_{\Gamma,\mu}[D] \rightarrow \min_{D \in V_{\Gamma,0}} \hat{J}_\Gamma[D]$) as $\mu \rightarrow 0$, where $D_\Gamma = \operatorname{argmin}_{D \in V_{\Gamma,f}} J_\Gamma[D]$ (resp. $\hat{D}_\Gamma = \operatorname{argmin}_{D \in V_{\Gamma,0}} \hat{J}_\Gamma[D]$).*

Proof. We only consider $J_{\Gamma,\mu}$ and J_Γ as the proof for $\hat{J}_{\Gamma,\mu}$ and \hat{J}_Γ is similar. Let $\mu > 0$. We define $I_{\Gamma,\mu} : H_g^1(\Omega) \rightarrow \mathbb{R} \cup \{-\infty\}$ by

$$I_{\Gamma,\mu}[\phi] = I_\Gamma[\phi] - \frac{\mu}{2} \int_{\Omega_-} \phi^2 dx = \int_{\Omega} \left[-\frac{\varepsilon_\Gamma}{2} |\nabla \phi|^2 + f\phi - \chi + B(\phi) - \frac{\chi - \mu}{2} \phi^2 \right] dx. \quad (2.9)$$

Similar to the proof of theorem 2.1, the functional $I_{\Gamma,\mu} : H_g^1(\Omega) \rightarrow \mathbb{R} \cup \{-\infty\}$ admits a unique maximizer $\phi_{\Gamma,\mu} \in L^\infty(\Omega)$ and it is the unique weak solution in $H_g^1(\Omega)$ to the Euler–Lagrange equation

$$-\nabla \cdot \varepsilon_\Gamma \nabla \phi_{\Gamma,\mu} + \chi + B'(\phi_{\Gamma,\mu}) + \mu \chi - \phi_{\Gamma,\mu} = f \quad \text{in } \Omega. \quad (2.10)$$

We show that $\phi_{\Gamma,\mu} \rightarrow \phi_\Gamma$ in $H^1(\Omega)$ and $I_{\Gamma,\mu}[\phi_{\Gamma,\mu}] \rightarrow I_\Gamma[\phi_\Gamma]$ as $\mu \rightarrow 0$, where ϕ_Γ is the unique maximizer of $I_\Gamma : H_g^1(\Omega) \rightarrow \mathbb{R} \cup \{-\infty\}$. It suffices to show that for any sequence $\mu_k \searrow 0$, there is a subsequence, not relabelled, such that $\phi_{\Gamma,\mu_k} \rightarrow \phi_\Gamma$ in $H^1(\Omega)$ and $I_{\Gamma,\mu_k}[\phi_{\Gamma,\mu_k}] \rightarrow I_\Gamma[\phi_\Gamma]$ as $k \rightarrow \infty$.

Let $\hat{g} \in H_g^1(\Omega)$ be such that $\hat{g} = 0$ in Ω_- . Then $I_{\Gamma,\mu}[\phi_{\Gamma,\mu}] = \max_{\phi \in H_g^1(\Omega)} I_{\Gamma,\mu}[\phi] \geq I_{\Gamma,\mu}[\hat{g}] = I_\Gamma[\hat{g}] > -\infty$. It then follows from the definition of $I_{\Gamma,\mu}$ and Poincaré's inequality that

$$\sup_{\mu > 0} \|\phi_{\Gamma,\mu}\|_{H^1(\Omega)} < \infty. \quad (2.11)$$

Now for any $\mu_k \searrow 0$, there exists a subsequence of $\{\phi_{\Gamma,\mu_k}\}$, not relabelled, and some $\bar{\phi}_\Gamma \in H^1(\Omega)$ such that $\phi_{\Gamma,\mu_k} \rightharpoonup \bar{\phi}_\Gamma$ in $H^1(\Omega)$. (\rightharpoonup denotes the weak convergence.) Clearly, $\bar{\phi}_\Gamma \in H_g^1(\Omega)$. By the convexity of $-I_\Gamma$, $\limsup_{k \rightarrow \infty} I_\Gamma[\phi_{\Gamma,\mu_k}] \leq I_\Gamma[\bar{\phi}_\Gamma]$. This and the fact $I_{\Gamma,\mu_k}[\phi_\Gamma] \leq I_{\Gamma,\mu_k}[\phi_{\Gamma,\mu_k}] \leq I_\Gamma[\phi_{\Gamma,\mu_k}]$ imply that

$$\begin{aligned} I_\Gamma[\phi_\Gamma] &= \lim_{k \rightarrow \infty} I_{\Gamma,\mu_k}[\phi_\Gamma] \leq \liminf_{k \rightarrow \infty} I_{\Gamma,\mu_k}[\phi_{\Gamma,\mu_k}] \leq \liminf_{k \rightarrow \infty} I_\Gamma[\phi_{\Gamma,\mu_k}] \\ &\leq \limsup_{k \rightarrow \infty} I_\Gamma[\phi_{\Gamma,\mu_k}] \leq I_\Gamma[\bar{\phi}_\Gamma] \leq I_\Gamma[\phi_\Gamma]. \end{aligned}$$

Thus, $I_\Gamma[\bar{\phi}_\Gamma] = I_\Gamma[\phi_\Gamma]$ and $\bar{\phi}_\Gamma = \phi_\Gamma$ by the uniqueness of maximizer of I_Γ . Hence, $I_{\Gamma,\mu_k}[\phi_{\Gamma,\mu_k}] \rightarrow I_\Gamma[\phi_\Gamma]$.

We now prove the strong convergence $\phi_{\Gamma,\mu} \rightarrow \phi_\Gamma$ in $H^1(\Omega)$. Denote for each $k \geq 1$

$$a_k = \int_{\Omega} \left[\frac{\varepsilon_\Gamma}{2} |\nabla \phi_{\Gamma,\mu_k}|^2 - \frac{\varepsilon_\Gamma}{2} |\nabla \phi_\Gamma|^2 \right] dx \quad \text{and} \quad b_k = \int_{\Omega_+} [B(\phi_{\Gamma,\mu_k}) - B(\phi_\Gamma)] dx.$$

Passing to a further subsequence if necessary, we have by the weak convergence $\phi_{\Gamma,\mu_k} \rightharpoonup \phi_\Gamma$ in $H^1(\Omega)$ that $\phi_{\Gamma,\mu_k} \rightarrow \phi_\Gamma$ in $L^2(\Omega)$ and $\phi_{\Gamma,\mu_k} \rightarrow \phi_\Gamma$ a.e. in Ω . These, together with the energy convergence $I_{\Gamma,\mu_k}[\phi_{\Gamma,\mu_k}] \rightarrow I_\Gamma[\phi_\Gamma]$ as $k \rightarrow \infty$ and the bound (2.11), imply that $a_k + b_k \rightarrow 0$ as

$k \rightarrow \infty$. But $\liminf_{k \rightarrow \infty} a_k \geq 0$ as $\phi_{\Gamma, \mu_k} \rightharpoonup \phi_\Gamma$ in $H^1(\Omega)$ and $\liminf_{k \rightarrow \infty} b_k \geq 0$ by Fatou's lemma. Therefore,

$$0 = \lim_{k \rightarrow \infty} (a_k + b_k) \geq \liminf_{k \rightarrow \infty} a_k + \liminf_{k \rightarrow \infty} b_k \geq 0.$$

Hence $\liminf_{k \rightarrow \infty} a_k = 0$. Passing to a further subsequence if necessary and without relabelling, we have $a_k \rightarrow 0$ as $k \rightarrow \infty$. This and the weak convergence $\phi_{\Gamma, \mu_k} \rightharpoonup \phi_\Gamma$ in $H^1(\Omega)$, together with the identity

$$\|\nabla \phi_{\Gamma, \mu} - \nabla \phi_\Gamma\|^2 = \|\nabla \phi_{\Gamma, \mu}\|^2 - \|\nabla \phi_\Gamma\|^2 - \int_{\Omega} 2\nabla \phi_\Gamma \cdot \nabla (\phi_{\Gamma, \mu} - \phi_\Gamma) dx,$$

imply that $\nabla \phi_{\Gamma, \mu_k} \rightarrow \nabla \phi_\Gamma$ in $L^2(\Omega)$ and consequently that $\phi_{\Gamma, \mu} \rightarrow \phi_\Gamma$ in $H^1(\Omega)$.

Finally, set $D_{\Gamma, \mu} = -\varepsilon_\Gamma \nabla \phi_{\Gamma, \mu}$. By (2.10), $D_{\Gamma, \mu} \in H(\text{div}, \Omega)$ and $\nabla \cdot D_{\Gamma, \mu} = f - \chi_+ B'(\phi_{\Gamma, \mu}) - \mu \chi_- \phi_{\Gamma, \mu}$ in Ω . Moreover, since the Legendre transform of $s \mapsto \mu s^2/2$ is $\xi \mapsto \xi^2/(2\mu)$, we see that $D_{\Gamma, \mu}$ is the unique minimizer of $J_{\Gamma, \mu}$ over $H(\text{div}, \Omega)$ and $J_{\Gamma, \mu}[D_{\Gamma, \mu}] = I_{\Gamma, \mu}[\phi_{\Gamma, \mu}]$ (cf. the proof of theorem 2.1). Thus, $D_{\Gamma, \mu} \rightarrow D_\Gamma$ in $H(\text{div}, \Omega)$ and $\min_{D \in H(\text{div}, \Omega)} J_{\Gamma, \mu}[D] \rightarrow \min_{D \in V_{\Gamma, f}} J_\Gamma[D]$. ■

(c) Numerical results

We choose $\Omega = (-1, 1)^3$, $\Gamma = \{x \in \mathbb{R}^3 : |x| = 1/2\}$, $\varepsilon_- = \varepsilon_0$, $\varepsilon_+ = 78\varepsilon_0$, $f = 1$ in Ω , and $g = 0$ on $\partial\Omega$, and consider $B(s) = s^2/2$ and $B(s) = \cosh(s) - 1$. We cover $\overline{\Omega} = [-1, 1]^3$ with a uniform finite-difference grid of size $h = 2L/N$, where $N + 1$ is the number of grid points in each coordinate direction, and use the central differencing and trapezoidal rule to discretize the energy functionals. The nonlinear conjugate gradient method is used to minimize or maximize numerically the resulting convex and concave functionals, respectively. We first maximize the discretized functional I_Γ with a very fine grid to obtain a numerical maximizer ϕ_{exact} , and calculate $D_{\text{exact}} = -\varepsilon_\Gamma \nabla \phi_{\text{exact}}$ and use it as the 'exact' solution. We then choose several values of N and μ . For each pair of the chosen N and μ , we minimize numerically the discretized functional $J_{\Gamma, \mu}$ to obtain a numerical minimizer $D_{\mu, N}$. To test how the constraint $\nabla \cdot D = f$ in Ω_- is satisfied, we define the penalty error

$$\text{PE}(\mu, N) := \|\nabla \cdot D_{\mu, N} - f\|_{L^2(\Omega_-)}^2.$$

Figure 2a,b shows that for a fixed $\mu > 0$ the error decreases with the increase of N , the number of grid points in one direction. For a large value of $\mu > 0$, such error decreasing saturates as N increases due to the penalty error with the fixed μ . Moreover, for a fixed N , the error decreases as μ decreases. Figure 2c,d indicates that $(1/2\mu)\text{PE}(\mu, N) = O(\mu)$ as $\mu \rightarrow 0$, implying that

$$J_{\Gamma, \mu}[D_{\Gamma, \mu}] - J_\Gamma[D_\Gamma] = \frac{1}{2\mu} \|\nabla \cdot D_{\Gamma, \mu} - f\|_{L^2(\Omega_-)}^2 \rightarrow 0 \quad \text{as } \mu \rightarrow 0,$$

as predicted by theorem 2.3.

3. Application to variational implicit solvation

(a) Solvation free energy and the dielectric boundary force

In the VISM (cf. figure 1), the solvation free-energy functional of dielectric boundaries is given in [1–4]

$$F[\Gamma] = P_0 \text{Vol}(\Omega_-) + \gamma_0 \int_{\Gamma} (1 - 2\tau H) dS + \rho_w \int_{\Omega_+} U(x) dV + E_{\text{ele}}[\Gamma]. \quad (3.1)$$

Here, P_0 is the difference between the pressure outside and inside the molecular region Ω_- , γ_0 is the constant surface tension, τ is the Tolman length and ρ_w is the bulk solvent density. All P_0 , γ_0 , τ and ρ_w are given constants. In addition, H is the mean curvature (positive if Ω_- is a sphere), $U(x) = \sum_{i=1}^N U_{\text{LJ}}^{(i)}(|x - x_i|)$ with each $U_{\text{LJ}}^{(i)}$ a Lennard–Jones (LJ) potential. The last part $E_{\text{ele}}[\Gamma]$ is the

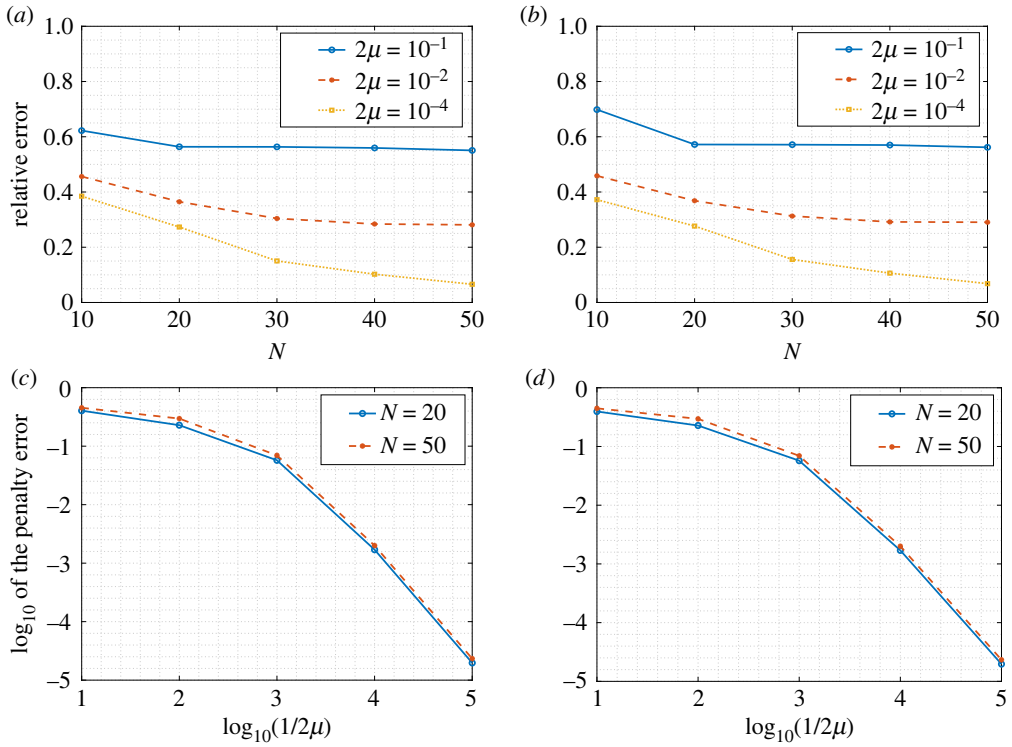


Figure 2. The relative error $\|D_{\mu,N} - D_{\text{exact}}\|_{L^2(\Omega)} / \|D_{\text{exact}}\|_{L^2(\Omega)}$ versus N for several μ -values (a,b) and the penalty error $\text{PE}(\mu, N)$ versus $1/(2\mu)$ for several values of N (c,d) for $B(s) = s^2/2$ (a,c) and $B(s) = \cosh(s) - 1$ (b,d).

electrostatic energy. In the classical PB formulation, it is given by

$$E_{\text{ele}}[\Gamma] = \begin{cases} \max_{\phi \in H^1_s(\Omega)} I_\Gamma[\phi] & \text{(for a continuum charge density; cf. (1.4)),} \\ \max_{\phi \in H^1_s(\Omega)} \hat{I}_\Gamma[\phi] & \text{(for point charges; cf. (2.6)).} \end{cases} \quad (3.2)$$

The dielectric boundary force is the negative first variation of the total free energy with respect to the variation of the boundary Γ . This first variation is a function on Γ and is given in [3,40,41]

$$\delta_\Gamma F[\Gamma] = P_0 + 2\gamma_0(H - \tau K) + \rho_w U + \delta_\Gamma E_{\text{ele}}[\Gamma].$$

For the case of a continuum charge density, the variation of the electrostatic energy $\delta_\Gamma E_{\text{ele}}[\Gamma]$ is given in [18,19]

$$\delta_\Gamma E_{\text{ele}}[\Gamma] = \frac{1}{2} \left(\frac{1}{\varepsilon_-} - \frac{1}{\varepsilon_+} \right) |\varepsilon_\Gamma \partial_n \phi_\Gamma|^2 + \frac{\varepsilon_+ - \varepsilon_-}{2} |(I - n \otimes n) \nabla \phi_\Gamma|^2 + B(\phi_\Gamma), \quad (3.3)$$

where ϕ_Γ is the equilibrium electrostatic potential, the solution to the boundary-value problem of the PBE (1.1), and I is the 3×3 identity matrix. For the case of point charges, the formula is the same except ϕ_Γ should be replaced by $\hat{\phi}_\Gamma$, the solution to the boundary-value problem of PBE (2.2). Note that the force $-\delta_\Gamma E_{\text{ele}}[\Gamma]$ points from the higher to lower dielectric region [18].

Here, we propose to use the LTPB formulation for the electrostatic energy

$$E_{\text{ele}}[\Gamma] = \begin{cases} \min_{D \in V_{\Gamma_f}} J_\Gamma[D] & \text{(for a continuum charge density; cf. (1.5)),} \\ \min_{D \in V_{\Gamma,0}} \hat{J}_\Gamma[D] & \text{(for point charges; cf. (2.7)).} \end{cases} \quad (3.4)$$

By theorem 2.1 and theorem 2.2, both the PB and LTPB formulations lead to the same value of $E_{\text{ele}}[\Gamma]$ for a continuum charge density or for point charges. To numerically minimize the total

solvation energy, we use the penalized LTPB electrostatic energy functionals defined in (1.10) and (2.8), respectively. Let us denote also by $E_{\text{ele},\mu}[\Gamma]$ the minimum of $J_{\Gamma,\mu}$ or $\hat{J}_{\Gamma,\mu}$ over $H(\text{div}, \Omega)$.

Theorem 3.1. (1) *A continuum charge density. Let D_Γ and $D_{\Gamma,\mu}$ be the minimizers of $J_\Gamma : V_{\Gamma,f} \rightarrow \mathbb{R} \cup \{\infty\}$ and $J_{\Gamma,\mu} : H(\text{div}, \Omega) \rightarrow \mathbb{R} \cup \{\infty\}$, respectively. We have*

$$\begin{aligned} \delta_\Gamma E_{\text{ele}}[\Gamma] &= \frac{1}{2} \left(\frac{1}{\varepsilon_-} - \frac{1}{\varepsilon_+} \right) |D_\Gamma \cdot n|^2 + \frac{\varepsilon_+ - \varepsilon_-}{2} \left| (I - n \otimes n) \frac{D_\Gamma}{\varepsilon_\Gamma} \right|^2 \\ &\quad + B(B^{*'}(f - \nabla \cdot D_\Gamma|_{\Omega_+})) \quad \text{on } \Gamma \end{aligned}$$

and

$$\begin{aligned} \delta_\Gamma E_{\text{ele},\mu}[\Gamma] &= \frac{1}{2} \left(\frac{1}{\varepsilon_-} - \frac{1}{\varepsilon_+} \right) |D_{\Gamma,\mu} \cdot n|^2 + \frac{\varepsilon_+ - \varepsilon_-}{2} \left| (I - n \otimes n) \frac{D_{\Gamma,\mu}}{\varepsilon_\Gamma} \right|^2 \\ &\quad + B(B^{*'}(f - \nabla \cdot D_{\Gamma,\mu}|_{\Omega_+})) - \frac{1}{2\mu} (f - \nabla \cdot D_{\Gamma,\mu}|_{\Omega_-})^2 \quad \text{on } \Gamma. \end{aligned}$$

(2) *Point charges. Let \hat{D}_Γ and $\hat{D}_{\Gamma,\mu}$ be the minimizers of $\hat{J}_\Gamma : V_{\Gamma,0} \rightarrow \mathbb{R} \cup \{\infty\}$ and $\hat{J}_{\Gamma,\mu} : H(\text{div}, \Omega) \rightarrow \mathbb{R} \cup \{\infty\}$, respectively. Then, the above formulae in part (1) hold true with \hat{D}_Γ and $\hat{D}_{\Gamma,\mu}$ replacing D_Γ and $D_{\Gamma,\mu}$, respectively, and $f = 0$.*

Proof. We only prove part (1) as the proof of part (2) is similar. By theorem 2.1, $D_\Gamma = -\varepsilon_\Gamma \nabla \phi_\Gamma$, where $\phi_\Gamma = \text{argmax}_{\phi \in H^1_g(\Omega)} I_\Gamma[\phi]$, and $B^{*'}(f - \nabla \cdot D_\Gamma) = \phi_\Gamma$ in Ω_+ . Now the first formula follows from (3.3). Set $I_{\Gamma,\mu}[\phi] = I_\Gamma[\phi] - (\mu/2) \|\phi\|_{L^2(\Omega_-)}^2$. Then $I_{\Gamma,\mu} : H^1_g(\Omega) \rightarrow \mathbb{R} \cup \{-\infty\}$ admits a unique maximizer $\phi_{\Gamma,\mu}$. Moreover, $D_{\Gamma,\mu} = -\varepsilon_\Gamma \nabla \phi_{\Gamma,\mu}$ is the unique minimizer of $J_{\Gamma,\mu} : H(\text{div}, \Omega) \rightarrow \mathbb{R} \cup \{\infty\}$, and $\max_{\phi \in H^1_g(\Omega)} I_{\Gamma,\mu}[\phi] = \min_{D \in H(\text{div}, \Omega)} J_{\Gamma,\mu}[D] = E_{\text{ele},\mu}[\Gamma]$. With the same argument for deriving (3.3) (cf. [18,19]), we obtain the formula for $\delta_\Gamma E_{\text{ele},\mu}[\Gamma] = \delta_\Gamma E_{\text{ele}}[\Gamma] - (\mu/2) \phi_{\Gamma,\mu}^2$. Since the Legendre transform of the function $s \mapsto \mu s^2/2$ is $\xi \mapsto \xi^2/(2\mu)$, we obtain the second formula by using the same argument. ■

(b) A min–min algorithm and a max–min algorithm

We apply the gradient descent method to minimize the total VISM free-energy functional (3.1). Once Γ_k is given, we obtain the new boundary $\Gamma_{k+1} := \Gamma_k - \alpha_k \delta_\Gamma F[\Gamma_k]$, i.e. a point $x_k \in \Gamma_k$ is moved to $x_{k+1} = x_k - \alpha_k \delta_\Gamma F[\Gamma_k](x_k) \in \Gamma_{k+1}$, with a step size $\alpha_k > 0$. With the LTPB electrostatics, we propose a min–min algorithm for such minimization. For comparing the LTPB and PB formulations, we also design a max–min algorithm for minimizing the total free energy with the PB electrostatic energy. These algorithms are described only for the continuum charge density as they are similar for point charges. Moreover, for the LTPB formulation, we shall replace the LTPB functionals by their penalized versions. Note that for large-scale simulations it is desirable to have only a few iterations to get the approximate ϕ_{k+1} or D_{k+1} . Hence, in our numerical tests, we will choose a small number of steps for such iterations.

A min–min algorithm for minimizing the VISM-LTPB functional. Given Γ_k and D_k .

- Minimize J_{Γ_k} over $V_{\Gamma_k,f}$ by an iteration method with D_k as the initial guess to obtain an approximate minimizer D_{k+1} ;
- Calculate $\delta_\Gamma E_{\text{ele}}[\Gamma_k]$ using D_{k+1} and then calculate $\delta_\Gamma F[\Gamma_k]$;
- Choose α_{k+1} and update $\Gamma_{k+1} = \Gamma_k - \alpha_{k+1} \delta_\Gamma F[\Gamma_k]$.

A max–min algorithm for minimizing the VISM-PB functional. Given Γ_k and ϕ_k .

- Maximize I_{Γ_k} over $H^1_g(\Omega)$ by an iteration method with ϕ_k as the initial guess to obtain an approximate maximizer ϕ_{k+1} ;
- Calculate $\delta_\Gamma E_{\text{ele}}[\Gamma_k]$ using ϕ_{k+1} and then calculate $\delta_\Gamma F[\Gamma_k]$;
- Choose α_{k+1} and update $\Gamma_{k+1} = \Gamma_k - \alpha_{k+1} \delta_\Gamma F[\Gamma_k]$.

We now state the min–min algorithm and prove its convergence for a general optimization problem. We assume

- (a1) Both m and n are positive integers, and $U \subset \mathbb{R}^m$ and $V \subset \mathbb{R}^n$ are open, bounded and convex;
- (a2) The function $f \in C^2(\overline{U})$ is convex, and the function $g \in C^2(\overline{U \times V})$ is strictly convex;
- (a3) For each $x \in \overline{U}$, there exists a unique minimizer $y_x \in V$ (not on ∂V) of $g(x, \cdot) : \overline{V} \rightarrow \mathbb{R}$.

We define $F : \overline{U} \rightarrow \mathbb{R}$ and $\bar{F} : \overline{U \times V} \rightarrow \mathbb{R}$, respectively, by

$$F(x) = f(x) + \min_{y \in \overline{V}} g(x, y) \quad \forall x \in \overline{U}$$

and

$$\bar{F}(x, y) = f(x) + g(x, y) \quad \forall (x, y) \in \overline{U \times V}.$$

Lemma 3.2.

- (1) The function $F : \overline{U} \rightarrow \mathbb{R}$ is Lipschitz-continuous. Moreover, there exists a unique $x^* \in \overline{U}$ such that $F(x^*) = \min_{x \in \overline{U}} F(x)$.
- (2) The function $\bar{F} \in C^2(\overline{U \times V})$ and is strictly convex. Moreover, there exists a unique $(\bar{x}, \bar{y}) \in \overline{U \times V}$ such that $\bar{F}(\bar{x}, \bar{y}) = \min_{(x, y) \in \overline{U \times V}} \bar{F}(x, y)$.
- (3) We have $\min_{x \in \overline{U}} F(x) = \min_{(x, y) \in \overline{U \times V}} \bar{F}(x, y)$. Moreover, $(x^*, y_{x^*}) = (\bar{x}, \bar{y})$.

Proof. (1) Let $x', x'' \in \overline{U}$. Since $\min_{y \in \overline{V}} g(x'', y) = g(x'', y_{x''})$ by the assumption (A 3), we have

$$\min_{y \in \overline{V}} g(x', y) - \min_{y \in \overline{V}} g(x'', y) \leq g(x', y_{x''}) - g(x'', y_{x''}) \leq K|x' - x''|,$$

where $K = \max_{(x, y) \in \overline{U \times V}} |\nabla_x g(x, y)|$. Switching x and x' and noting that $f \in C^1(\overline{U})$, we see that $F : \overline{U} \rightarrow \mathbb{R}$ is Lipschitz-continuous. Since F is continuous on \overline{U} and \overline{U} is compact, there exists $x^* \in \overline{U}$ such that $F(x^*) = \min_{x \in \overline{U}} F(x)$. The uniqueness of x^* follows from Part (3).

(2) These follow from the assumptions on f and g , and the compactness of $\overline{U \times V}$.

(3) Clearly, $F(x) = \bar{F}(x, y_x) \geq \bar{F}(\bar{x}, \bar{y})$ for any $x \in \overline{U}$ (cf. the assumption (A 3)). Hence, $\min_{x \in \overline{U}} F(x) \geq \min_{(x, y) \in \overline{U \times V}} \bar{F}(x, y)$. On the other hand,

$$\min_{(x, y) \in \overline{U \times V}} \bar{F}(x, y) = \bar{F}(\bar{x}, \bar{y}) = f(\bar{x}) + g(\bar{x}, \bar{y}) \geq f(\bar{x}) + g(\bar{x}, y_{\bar{x}}) = F(\bar{x}) \geq \min_{x \in \overline{U}} F(x).$$

Thus, $\min_{x \in \overline{U}} F(x) = \min_{(x, y) \in \overline{U \times V}} \bar{F}(x, y)$. This also implies that $\bar{F}(x^*, y_{x^*}) = F(x^*) = \bar{F}(\bar{x}, \bar{y})$. Since the minimizer of \bar{F} is unique, we have $(x^*, y_{x^*}) = (\bar{x}, \bar{y})$. ■

A min–min algorithm for finding $x^* = \operatorname{argmin}_{x \in \overline{U}} F(x)$.

Step 0. Choose $x_0 \in U$. If $x_0 = x^*$ then stop. Otherwise, compute $y_0 = \operatorname{argmin}_{y \in \overline{V}} g(x_0, y)$. Set $k = 0$.

Step 1. Compute $\alpha_k = \operatorname{argmin}_{\alpha \in A_k} \{\bar{F}(x_k - \alpha \nabla_x \bar{F}(x_k, y_k), y_k)\}$, where $A_k = \{\alpha > 0 : x_k - \alpha \nabla_x \bar{F}(x_k, y_k) \in U\}$, $x_{k+1} = x_k - \alpha_k \nabla_x \bar{F}(x_k, y_k)$.

Step 2. Compute $y_{k+1} = \operatorname{argmin}_{y \in \overline{V}} g(x_{k+1}, y)$.

Step 3. If $x_{k+1} = x^*$ then stop. Otherwise, set $k := k + 1$ and repeat Steps 1–3.

Theorem 3.3. Assume $x^* = \operatorname{argmin}_{x \in \overline{U}} F(x) \in U$. Let $x_0 \in U$ and $y_0 = y_{x_0} \in V$. Assume

$$W_0 := \{(x, y) \in \overline{U \times V} : \bar{F}(x, y) \leq \bar{F}(x_0, y_0)\} \subset U \times V.$$

Let $\{x_k\}_{k=1}^\infty$ be generated by the min–min algorithm. Then, $x_k = x^*$ for some $k \geq 0$ or $x_k \rightarrow x^*$.

Proof. Let us assume $x_k \neq x^*$ for any $k \geq 0$ and show that $x_k \rightarrow x^*$. Define for each $k \geq 1$ the set W_k the same as W_0 with (x_0, y_0) replaced by (x_k, y_k) . We first show the following statement for each

$k \geq 0$: there exists a unique $\alpha_k \in A_k$ as defined in Step 1, the point (x_k, y_k) defined in the algorithm satisfies $(x_k, y_k) \in \text{int}(W_k)$ and $d^{(k)} := \nabla_x \bar{F}(x_k, y_k) \neq 0$, $\bar{F}(x_{k+1}, y_{k+1}) < \bar{F}(x_k, y_k)$, and $W_{k+1} \subseteq W_k$.

Consider $k = 0$. Since $y_0 = y_{x_0} = \arg\min_{y \in \bar{V}} g(x_0, y) \in V$ (cf. the assumption (A 3)), $\nabla_y \bar{F}(x_0, y_0) = \nabla_y g(x_0, y_0) = 0$. Since $x_0 \neq x^*$, and since by lemma 3.2 $(x^*, y_{x^*}) \in U \times V$ is the unique minimizer of \bar{F} over $\bar{U} \times \bar{V}$, we have $d^{(0)} := \nabla_x \bar{F}(x_0, y_0) \neq 0$. It follows from Taylor's expansion that

$$\bar{F}(x_0 - \alpha d^{(0)}, y_0) = \bar{F}(x_0, y_0) - \alpha \|d^{(0)}\|^2 + O(\alpha^2) < \bar{F}(x_0, y_0) \quad \text{if } 0 < \alpha \ll 1.$$

By the assumption that the closed set $W_0 \subset U \times V$ and the strict convexity of \bar{F} , there exists a unique $\alpha_0 \in A_0$ such that $\bar{F}(x_0 - \alpha_0 d^{(0)}, y_0) \leq \bar{F}(x_0 - \alpha d^{(0)}, y_0)$ for all $\alpha \in A_0$. Now $x_1 = x_0 - \alpha_0 \nabla_x \bar{F}(x_0, y_0) \in U$ and $\bar{F}(x_1, y_0) < \bar{F}(x_0, y_0)$. By Step 2 of the algorithm and the assumption (A 3), $y_1 = y_{x_1} \in V$ and $\bar{F}(x_1, y_1) \leq \bar{F}(x_1, y_0) < \bar{F}(x_0, y_0)$. It then follows from the definition of W_0 that $(x_1, y_1) \in \text{int}(W_0) \subset U \times V$. Clearly, $W_1 \subseteq W_0$. Thus, the statement is true for $k = 0$. Suppose the statement is true for $k \geq 1$. Using the above argument with (x_0, y_0) and W_0 replaced by (x_k, y_k) and W_k , respectively, we verify that the statement is also true for $k + 1$. Therefore, the statement is true in general by induction.

We now prove $x_k \rightarrow x^*$. Since $\{x_k\}$ is bounded, it suffices to show the following: if a subsequence of $\{x_k\}$, not relabelled, converges to some $\hat{x} \in \bar{U}$, then $\hat{x} = x^*$. First note that $\bar{F}(x_k, y_k) \rightarrow \hat{F}$ for some $\hat{F} \in \mathbb{R}$, as $\{\bar{F}(x_k, y_k)\}$ is decreasing and bounded. Now, passing to a further subsequence if necessary, we have $(x_k, y_k) \rightarrow (\hat{x}, \hat{y})$. By the above statement, $(x_k, y_k) \in \text{int}(W_0)$ for all k . Hence $(\hat{x}, \hat{y}) \in W_0 \subset U \times V$. Therefore, it follows from the continuity of \bar{F} that $\hat{F} = \bar{F}(\hat{x}, \hat{y})$. Note for each k that $\nabla_y \bar{F}(x_k, y_k) = \nabla_y g(x_k, y_k) = 0$. Consequently, $\nabla_y \bar{F}(\hat{x}, \hat{y}) = 0$. We shall show that $\hat{d} := \nabla_x \bar{F}(\hat{x}, \hat{y}) = 0$. This would imply that $(\hat{x}, \hat{y}) \in U \times V$ is the unique minimizer of \bar{F} over $\bar{U} \times \bar{V}$, and hence $\hat{x} = x^*$ by lemma 3.2 and $x_k \rightarrow x^*$.

Assume $\hat{d} \neq 0$. Since $(x_k, y_k) \rightarrow (\hat{x}, \hat{y})$ and $d^{(k)} := \nabla_x \bar{F}(x_k, y_k) \rightarrow \hat{d}$, it follows from the same argument used above that there exists $\hat{\alpha} > 0$ and an integer $k_0 \geq 1$ such that $\bar{F}(\hat{x} - \hat{\alpha} \hat{d}, \hat{y}) < \bar{F}(\hat{x}, \hat{y})$, $(\hat{x} - \hat{\alpha} \hat{d}, \hat{y}) \in \text{int}(W_0)$ and $(x_k - \hat{\alpha} d^{(k)}, y_k) \in \text{int}(W_0)$ if $k \geq k_0$. Now, since $\bar{F}(x_k, y_k)$ decreases and converges to $\hat{F} = \bar{F}(\hat{x}, \hat{y})$ and $y_{k+1} = y_{x_{k+1}}$ (cf. the assumption (A 3) and Step 2 in the algorithm), we have by the definition of x_{k+1} (cf. Step 1 in the algorithm) that for all $k \geq k_0$

$$\bar{F}(\hat{x}, \hat{y}) \leq \bar{F}(x_{k+1}, y_{k+1}) \leq \bar{F}(x_{k+1}, y_k) = \bar{F}(x_k - \alpha_k d^{(k)}, y_k) \leq \bar{F}(x_k - \hat{\alpha} d^{(k)}, y_k).$$

Taking $k \rightarrow \infty$, we get $\bar{F}(\hat{x}, \hat{y}) \leq \bar{F}(\hat{x} - \hat{\alpha} \hat{d}, \hat{y}) < \bar{F}(\hat{x}, \hat{y})$, a contradiction. Thus, $\hat{d} = 0$. ■

We remark that the proof of the convergence of the min–min algorithm relies on the fact that $\min_{x \in \bar{U}} \bar{F}(x) = \min_{(x,y) \in \bar{U} \times \bar{V}} \bar{F}(x, y)$ and $(x^*, y_{x^*}) = (\bar{x}, \bar{y})$ as established in lemma 3.2. Such structure is lost for a max–min algorithm and therefore the method of proof of theorem 3.3 does not extend directly to the convergence of a max–min algorithm.

(c) A radially symmetric system for the solvation of a single ion

We apply the VISM to the solvation of a single ion placed at the origin (i.e. $N = 1$ in figure 1). The resulting system is simplified to be radially symmetric. Such a system is effectively one-dimensional for which we can obtain a very accurate solution for testing our algorithms.

We set $\Omega_- = \{x \in \mathbb{R}^3 : |x| < R\}$, $\Omega_+ = \{x \in \mathbb{R}^3 : R < |x| < A\}$ and $\Gamma = \{x \in \mathbb{R}^3 : |x| = R\}$, where $A, R \in \mathbb{R}$ with $0 < R < A$, and denote $\chi_- = \chi_{(0,R)}$ and $\chi_+ = \chi_{(R,A)}$. The dielectric coefficient is $\varepsilon_R : [0, A] \rightarrow \mathbb{R}$ with $\varepsilon_R(r) = \varepsilon_-$ if $r < R$ and $\varepsilon_R(r) = \varepsilon_+$ if $r > R$. The LJ parameters for the single ion are denoted by ε and σ ; cf. figure 1 and (3.1). We assume $f : [0, A] \rightarrow \mathbb{R}$ is a smooth function, $g \in \mathbb{R}$ and $\mu > 0$. The electrostatic potential ϕ is assumed to be radially symmetric: $\phi = \phi(r)$ with $r = |x|$. The dielectric displacement D is proportional to $\nabla_x(\phi(r)) = \phi'(r)(x/r)$. Thus, we assume $D(x) = p(r)(x/r)$ for some radially symmetric function $p = p(r)$. Since $|D(x)| = |p(r)|$ and

$$\nabla \cdot D(x) = \frac{2}{r} p(r) + p'(r) = \frac{1}{r^2} (r^2 p(r))',$$

we shall consider $p = p(r)$ instead of $D(x)$. We have $p = -\varepsilon_R \phi'$ if $D = -\varepsilon_R \nabla \phi$.

We denote $\omega(r) = r^2$ and define

$$H_\omega^1(0, A) = \{\phi : (0, A) \rightarrow \mathbb{R} : \phi \text{ is weakly differentiable and } \int_0^A (\phi^2 + \phi'^2) r^2 \, dr < \infty\}, \quad (3.5)$$

$$X_g = \{\phi \in H_\omega^1(0, A) : \phi(A) = g\}, \quad (3.6)$$

$$Y = \left\{ p : (0, A) \rightarrow \mathbb{R} : p \text{ is weakly differentiable and } \int_0^A [(r^2 p)^2 + ((r^2 p)')^2] \frac{1}{r^2} \, dr < \infty \right\}, \quad (3.7)$$

$$\text{and } Y_f = \left\{ p \in Y : \frac{2}{r} p + p' = f \text{ in } (0, R) \right\}. \quad (3.8)$$

Both $H_\omega^1(0, A)$ and Y are Hilbert spaces with their inner products and norms given, respectively, by

$$\begin{aligned} \langle \phi, \psi \rangle_\omega &= \int_0^A (\phi \psi + \phi' \psi') r^2 \, dr \quad \text{and} \quad \|\phi\|_\omega = \sqrt{\langle \phi, \phi \rangle_\omega}, \\ \langle p, q \rangle_Y &= \int_0^A [(r^2 p)(r^2 q) + (r^2 p)'(r^2 q)'] \frac{1}{r^2} \, dr \quad \text{and} \quad \|p\|_Y = \sqrt{\langle p, p \rangle_Y}. \end{aligned}$$

If $\phi \in H_\omega^1(0, A)$ and $0 < \delta < A$, then $\phi \in H^1(\delta, A)$ and hence ϕ is absolutely continuous on $[\delta, A]$, and the trace $\phi(A)$ is well defined. One verifies that $Y = H_\omega^1(0, A) \cap L^2(0, A)$. If $p \in Y$, then $p \in H^1(\delta, A)$ for any $\delta \in (0, A)$, and hence the trace $p(A)$ is well defined. The equation in defining Y_f is $(r^2 p(r))' = r^2 f(r)$ for $0 < r < R$, same as the constraint $\nabla \cdot D = f$ in Ω_- in the Cartesian coordinates for $D(r) = p(r)(x/r)$.

To compare the min–min and max–min algorithms, we now express the VISM solvation free energy with the classical PB and the penalized LTPB electrostatic energies in the radially symmetric setting. The non-polar part of the solvation free energy (i.e. the first three terms of $F[\Gamma]$ defined in (3.1)) $F_0 : (0, A) \rightarrow \mathbb{R}$ and its boundary variation $\delta_R F_0 : (0, A) \rightarrow \mathbb{R}$ are

$$F_0(R) = \frac{4\pi}{3} P_0 R^3 + 4\pi \gamma_0 R^2 - 8\pi \gamma_0 \tau R + 16\pi \rho_w \varepsilon \left(\frac{\sigma^{12}}{9R^9} - \frac{\sigma^6}{3R^3} \right) \quad (3.9)$$

and

$$\delta_R F_0(R) = P_0 + \frac{2\gamma_0}{R} - \frac{2\gamma_0 \tau}{R^2} + 4\rho_w \varepsilon \left(\frac{\sigma^{12}}{R^{12}} - \frac{\sigma^6}{R^6} \right). \quad (3.10)$$

Note that the boundary variation $\delta_R F_0$ differs from the derivative $F'_0(R)$ by the factor $4\pi R^2$; cf. [12].

Case 1. A continuum charge density. The VISM-PB free-energy functional $F : (0, A) \rightarrow \mathbb{R}$ with $I_R : X_g \rightarrow \mathbb{R}$ and the boundary variation $\delta_R F : (0, A) \rightarrow \mathbb{R}$ are now given by (cf. (1.4) and (3.1)–(3.3))

$$F(R) = F_0(R) + \max_{\phi \in X_g} I_R[\phi], \quad (3.11)$$

$$I_R[\phi] = 4\pi \int_0^A \left[-\frac{\varepsilon_R}{2} |\phi'|^2 + f\phi - \chi_+ B(\phi) \right] r^2 \, dr \quad (3.12)$$

$$\text{and } \delta_R F(R) = \delta_R F_0(R) + \frac{1}{2} \left(\frac{1}{\varepsilon_-} - \frac{1}{\varepsilon_+} \right) [\varepsilon_R \phi'_R(R)]^2 + B(\phi_R(R)), \quad (3.13)$$

where $\phi_R = \operatorname{argmin}_{\phi \in X_g} I_R[\phi]$; cf. theorem 3.4. The penalized VISM-LTPB free-energy functional $F_\mu : (0, A) \rightarrow \mathbb{R}$ with $J_{R,\mu} : Y \rightarrow \mathbb{R}$ and its boundary variation $\delta_R F_\mu : (0, A) \rightarrow \mathbb{R}$ are given by (cf. (3.1),

(3.4), (1.10) and part (1) of theorem 3.1)

$$F_\mu(R) = F_0(R) + \min_{p \in Y} J_{R,\mu}[p], \quad (3.14)$$

$$J_{R,\mu}[p] = 4\pi \int_0^A \left[\frac{1}{2\varepsilon_R} p^2 + \chi_+ B^* \left(f - \left(\frac{2}{r} p + p' \right) \right) + \frac{\chi_-}{2\mu} \left(f - \left(\frac{2}{r} p + p' \right) \right)^2 \right] r^2 dr \quad (3.15)$$

$$+ 4\pi g p(A) A^2, \quad (3.16)$$

$$\begin{aligned} \delta_R F_\mu(R) = & \delta_R F_0(R) + \frac{1}{2} \left(\frac{1}{\varepsilon_-} - \frac{1}{\varepsilon_+} \right) [p_{R,\mu}(R)]^2 + B \left(B^{*'} \left(f(R) - p'_{R,\mu}(R+) - \frac{2}{R} p_{R,\mu}(R+) \right) \right) \\ & - \frac{1}{2\mu} \left(f(R) - p'_{R,\mu}(R-) - \frac{2}{R} p_{R,\mu}(R-) \right)^2, \end{aligned} \quad (3.17)$$

where $p_{R,\mu} = \operatorname{argmin}_{p \in Y} J_{R,\mu}[p]$; cf. theorem 3.4.

Case 2. A single-point charge. With the point charge Q at the origin, the VISM-PB free-energy functional $\hat{F} : (0, A) \rightarrow \mathbb{R}$ with $\hat{I}_R : X_g \rightarrow \mathbb{R}$ and the boundary variation $\delta_R \hat{F} : (0, A) \rightarrow \mathbb{R}$ are given by (cf. (3.1), (3.2), (2.6) and (3.3))

$$\hat{F}(R) = F_0(R) + \max_{\phi \in X_g} \hat{I}_R[\phi], \quad (3.18)$$

$$\begin{aligned} \hat{I}_R[\phi] = & 4\pi \int_0^A \left[-\frac{\chi_- \varepsilon_-}{2} \left(\phi' + \frac{Q}{4\pi \varepsilon_- r^2} \right)^2 - \frac{\chi_+ \varepsilon_+}{2} \left(\phi' + \frac{Q}{4\pi \varepsilon_+ r^2} \right)^2 - \chi_+ B(\phi) \right] r^2 dr \\ & + \frac{Q^2}{8\pi} \left[\left(\frac{1}{\varepsilon_+} - \frac{1}{\varepsilon_-} \right) \frac{1}{R} - \frac{1}{\varepsilon_+ A} \right] \end{aligned} \quad (3.19)$$

$$\text{and} \quad \delta_R \hat{F}(R) = \delta_R F_0(R) + \frac{1}{2} \left(\frac{1}{\varepsilon_-} - \frac{1}{\varepsilon_+} \right) [\varepsilon_R \phi'_R(R)]^2 + B(\phi_R(R)), \quad (3.20)$$

where $\phi_R = \operatorname{argmax}_{\phi \in X_g} \hat{I}_R[\phi]$; cf. theorem 3.4. The penalized VISM-LTPB energy functional $\hat{F}_\mu : (0, A) \rightarrow \mathbb{R}$ with $\hat{J}_{R,\mu} : Y \rightarrow \mathbb{R}$ and the boundary variation $\delta_R \hat{F}_\mu : (0, A) \rightarrow \mathbb{R}$ are now given by (cf. (3.1), (3.4), (2.7) and part (2) of theorem 3.1)

$$\hat{F}_\mu(R) = F_0(R) + \min_{p \in Y} \hat{J}_{R,\mu}[p], \quad (3.21)$$

$$\begin{aligned} \hat{J}_{R,\mu}[p] = & 4\pi \int_0^A \left\{ \frac{1}{2\varepsilon_R} p^2 + \frac{\chi_-}{2\mu} \left(\frac{2}{r} p + p' \right)^2 + \chi_+ \left[B^* \left(-\frac{2}{r} p - p' \right) + \hat{\phi}_{R,\infty} \left(\frac{2}{r} p + p' \right) \right] \right\} r^2 dr \\ & + \frac{Q^2}{8\pi} \left[\left(\frac{1}{\varepsilon_+} - \frac{1}{\varepsilon_-} \right) \frac{1}{R} - \frac{1}{\varepsilon_+ A} \right], \end{aligned} \quad (3.22)$$

$$\begin{aligned} \delta_R \hat{F}_\mu(R) = & \delta_R F_0(R) + \frac{1}{2} \left(\frac{1}{\varepsilon_-} - \frac{1}{\varepsilon_+} \right) \left(p_{R,\mu}(R) + \frac{Q}{4\pi R^2} \right)^2 + B \left(B^{*'} \left(-p'_{R,\mu}(R+) - \frac{2}{R} p_{R,\mu}(R+) \right) \right) \\ & - \frac{1}{2\mu} \left(p'_{R,\mu}(R-) + \frac{2}{R} p_{R,\mu}(R-) \right)^2, \end{aligned} \quad (3.23)$$

where $p_{R,\mu} = \operatorname{argmin}_{p \in Y} \hat{J}_{R,\mu}[p]$; cf. theorem 3.4.

We note that, while the radial symmetry of functions simplifies our models and computations, the analysis in §2 does not directly apply here due to possible singularities at the origin. Thus, we present here a similar analysis for the radially symmetric system, but only for the case of a continuum charge density, as the case of point charges is similar. In appendix A, we also provide a new and direct derivation of the boundary force in the one-dimensional setting. Let us define $I_{R,\mu} : X_g \rightarrow \mathbb{R} \cup \{-\infty\}$ and $J_R : Y_f \rightarrow \mathbb{R} \cup \{\infty\}$, respectively, by

$$I_{R,\mu}[\phi] = 4\pi \int_0^A \left[-\frac{\varepsilon_R}{2} |\phi'|^2 + f\phi - \chi_+ B(\phi) - \frac{\chi_- \mu}{2} \phi^2 \right] r^2 dr \quad (3.24)$$

and

$$J_R[p] = 4\pi \int_0^A \left[\frac{1}{2\varepsilon_R} p^2 + \chi_+ B^* \left(f - \left(\frac{2}{r} p + p' \right) \right) \right] r^2 dr + 4\pi g p(A) A^2. \quad (3.25)$$

Theorem 3.4.

- (1) Denote $I_{R,0} = I_R$. For each $\mu \geq 0$, there exists a unique $\phi_{R,\mu} \in X_g$ such that $I_{R,\mu}[\phi_{R,\mu}] = \max_{\phi \in X_g} I_{R,\mu}[\phi]$. Moreover, $\phi_{R,\mu} \in X_g$ is the unique solution to

$$\frac{\varepsilon_-}{r^2} (r^2 \phi')' - \mu \phi = -f \quad \text{in } (0, R) \quad \text{and} \quad \frac{\varepsilon_+}{r^2} (r^2 \phi')' - B'(\phi) = -f \quad \text{in } (R, A),$$

and $\varepsilon_- \phi'(R-) = \varepsilon_+ \phi'(R+)$ and $\phi(A) = g$.

- (2) Duality. We have $I_R[\phi] \leq J_R[p]$ for any $\phi \in X_g$ and $p \in Y_f$, and $I_{R,\mu}[\phi] \leq J_{R,\mu}[p]$ for any $\mu > 0$, $\phi \in X_g$ and $p \in Y$. Moreover, $p_R := -\varepsilon_R \phi'_R \in Y_f$ and $p_{R,\mu} := -\varepsilon_R \phi'_{R,\mu} \in Y$ ($\mu > 0$) are the unique minimizers of J_R over Y_f and $J_{R,\mu}$ over Y , respectively.
- (3) Convergence. We have $\|\phi_{R,\mu} - \phi_R\|_\omega \rightarrow 0$, $\max_{\phi \in X_g} I_{R,\mu}[\phi] \rightarrow \max_{\phi \in X_g} I_R[\phi]$, $\|p_{R,\mu} - p_R\|_Y \rightarrow 0$ and $\min_{p \in Y} J_{R,\mu}[p] \rightarrow \min_{p \in Y} J_R[p]$ as $\mu \rightarrow 0$.

To prove theorem 3.4, we need the following lemma that summarizes some properties, particularly the behaviour near $r = 0$, of the functions in $H_\omega^1(0, A)$ (cf. (3.5)) and Y (cf. (3.7)):

Lemma 3.5. Let $\phi \in H_\omega^1(0, A)$ and $p \in Y$. Define $u(r) = r^2 p(r)$ and $v(r) = r^2 \phi(r) p(r)$ for $0 < r \leq A$. Then, $\sup_{0 < r < A} \sqrt{r} |\phi(r)| < \infty$, $u \in H^1(0, A)$ and $u(r) = o(r^{3/2})$ as $r \rightarrow 0$, and $v \in W^{1,1}(0, A)$ and $v(r) = o(r)$ as $r \rightarrow 0$.

Proof. Note for any $\delta \in (0, A)$ that $\phi \in H^1(\delta, A)$ and hence ϕ is absolutely continuous on $[\delta, A]$. Hence,

$$\begin{aligned} \phi(r)^2 &= \left[\phi(A) - \int_r^A \phi'(s) ds \right]^2 \\ &\leq 2\phi(A)^2 + 2 \left(\int_r^A \phi'(s) ds \right)^2 \\ &\leq 2\phi(A)^2 + 2 \left(\int_r^A s^2 \phi'(s)^2 ds \right) \left(\int_r^A s^{-2} ds \right) \\ &= 2\phi(A)^2 + 2 \left(\int_r^A s^2 \phi'(s)^2 ds \right) \left(\frac{1}{r} - \frac{1}{A} \right) \\ &\leq 2\phi(A)^2 + \frac{2}{r} \|\phi\|_\omega^2 \quad \forall r \in (0, A). \end{aligned}$$

Consequently, $r\phi(r)^2 \leq 2A\phi(A)^2 + 2\|\phi\|_{H_\omega^1(0,A)}^2$ if $r \in (0, A)$, and hence $\sup_{0 < r < A} \sqrt{r} |\phi(r)| < \infty$.

Since $p \in Y$, we can directly verify that $u \in H^1(0, A)$, and hence u is absolutely continuous on $[0, A]$. We must have $u(0) := \lim_{r \rightarrow 0} u(r) = 0$, since

$$\int_0^A \frac{u(r)^2}{r^2} dr = \int_0^A r^2 p(r)^2 dr < \infty.$$

Consequently,

$$u(r)^2 = \left(\int_0^r u'(s) ds \right)^2 \leq \left(\int_0^r s^2 ds \right) \left(\int_0^r \frac{u'(s)^2}{s^2} ds \right) = \frac{1}{3} r^3 o(1) \quad \text{as } r \rightarrow 0,$$

and hence $u(r) = o(r^{3/2})$ as $r \rightarrow 0$.

Since $v(r) = r^2 \phi(r)p(r)$ ($0 < r < A$), we have $\|v\|_{L^1(0,A)} \leq \|\phi\|_\omega \|p\|_Y < \infty$. Consequently,

$$\begin{aligned} \int_0^A |v'(r)| \, dr &= \int_0^A |\phi'(r)r^2p(r) + \phi(r)(r^2p(r))'| \, dr \\ &\leq \int_0^A |\phi'(r)r^2p(r)| \, dr + \int_0^A \left| r\phi(r) \frac{[r^2p(r)]'}{r} \right| \, dr \\ &\leq \left(\int_0^A r^2\phi'(r)^2 \, dr \right)^{1/2} \left(\int_0^A r^2p(r)^2 \, dr \right)^{1/2} \\ &\quad + \left(\int_0^A r^2\phi(r)^2 \, dr \right)^{1/2} \left(\int_0^A [(r^2p(r))']^2 \frac{1}{r^2} \, dr \right)^{1/2} \\ &\leq 2\|\phi\|_\omega \|p\|_Y < \infty. \end{aligned}$$

Hence, $v \in W^{1,1}(0,A)$, it is absolutely continuous, and $v(r) = \sqrt{r}\phi(r)u(r)/\sqrt{r} = o(r)$ as $r \rightarrow 0$. ■

Proof of theorem 3.4. (1) The proof of this part is standard; cf. [15,18,19].

(2) Let $\mu > 0$, $\phi \in X_g$ and $p \in Y$. Since $\phi(A) = g$ and $\lim_{r \rightarrow 0^+} r^2\phi(r)p(r) = 0$ by lemma 3.5,

$$\int_0^A p\phi' r^2 \, dr = - \int_0^A (r^2p)' \phi \, dr + gp(A)A^2 = - \int_0^A \left(\frac{2}{r}p + p' \right) \phi r^2 \, dr + gp(A)A^2.$$

Consequently, by the fact that the Legendre transform of $s \mapsto (\mu/2)s^2$ is $\xi \mapsto \xi^2/(2\mu)$, we obtain

$$\begin{aligned} I_{R,\mu}[\phi] &\leq I_{R,\mu}[\phi] + 4\pi \int_0^A \frac{1}{2\varepsilon_R} |p + \varepsilon_R \phi'|^2 r^2 \, dr \\ &= 4\pi \int_0^A \left(f\phi - \chi_+ B(\phi) - \frac{\chi_- \mu}{2} \phi^2 + \frac{1}{2\varepsilon_R} p^2 + p\phi' \right) r^2 \, dr \\ &= 4\pi \int_0^A \left\{ \frac{1}{2\varepsilon_R} p^2 + \chi_+ \left[\left(f - \left(\frac{2}{r}p + p' \right) \right) \phi - B(\phi) \right] \right. \\ &\quad \left. + \chi_- \left[\left(f - \left(\frac{2}{r}p + p' \right) \right) \phi - \frac{\mu}{2} \phi^2 \right] \right\} r^2 \, dr + 4\pi gp(A)A^2 \\ &\leq 4\pi \int_0^A \left[\frac{1}{2\varepsilon_R} p^2 + \chi_+ B^* \left(f - \left(\frac{2}{r}p + p' \right) \right) + \frac{\chi_-}{2\mu} \left(f - \left(\frac{2}{r}p + p' \right) \right)^2 \right] r^2 \, dr \\ &\quad + 4\pi gp(A)A^2 \\ &= J_{R,\mu}[p]. \end{aligned} \tag{3.26}$$

The inequality $I_R[\phi] \leq J_R[p]$ for any $\phi \in X_g$ and $p \in Y_f$ can be proved similarly.

By part (1), $p_{R,\mu} = -\varepsilon_R \phi'_{R,\mu}$ satisfies

$$(r^2 p_{R,\mu})' = (f - \mu \phi_{R,\mu}) r^2 \quad \text{in } (0, R) \quad \text{and} \quad (r^2 p_{R,\mu})' = (f - B'(\phi_{R,\mu})) r^2 \quad \text{in } (R, A). \tag{3.27}$$

These and the fact that $\phi_{R,\mu} \in X_g$ imply that $p_{R,\mu} \in Y$. Moreover, the first inequality in (3.26) becomes an equality with $\phi_{R,\mu}$ and $p_{R,\mu} = -\varepsilon_R \phi'_{R,\mu}$ replacing ϕ and p , respectively. The second inequality in (3.26) also becomes an equality by (3.27) and the definition of the Legendre transform. Thus, by the convexity of $J_{R,\mu}$, $p_{R,\mu}$ is the unique minimizer of $J_{R,\mu}$ over Y . Similarly, p_R is the unique minimizer of J_R over Y_f .

(3) By the same argument used in proving part (2) of theorem 2.3, we have $\|\phi_{R,\mu} - \phi_R\|_\omega \rightarrow 0$ and $\max_{\phi \in X_g} I_{R,\mu}[\phi] \rightarrow \max_{\phi \in X_g} I_R[\phi]$ as $\mu \rightarrow 0$. These and part (2) imply $\|p_{R,\mu} - p_R\|_Y \rightarrow 0$ and $\min_{p \in Y} J_{R,\mu}[p] \rightarrow \min_{p \in Y_f} J_R[p]$ as $\mu \rightarrow 0$. ■

Table 1. Model parameters. LJ means Lennard–Jones.

parameters	descriptions	estimated values	units
T	temperature	300	K
P_0	pressure difference	0	bar
γ_0	constant surface tension	0.1315	$k_B T / \text{\AA}^2$
τ	Tolman length	0.76	\AA
ρ_w	bulk solvent density	0.0331	\AA^{-3}
σ	length parameter in LJ potential	3.5	\AA
ε	energy parameter in LJ potential	0.3	$k_B T$
ε_-	relative dielectric permittivity in Ω_-	1	no unit
ε_+	relative permittivity in Ω_+	78	no unit

(d) Numerical results

In this section, we present numerical results of the solvation of a spherical molecule such as a single ion. We first demonstrate that both the min–min algorithm for the VISM with the LTPB formulation of electrostatics and the max–min algorithm for the VISM with the PB formulation of electrostatics can achieve the same accuracy. Note that the two formulations are equivalent (cf. theorem 2.1, theorem 2.2 and theorem 3.4). But the penalty method used to approximate the constraint by Gauss’ law leads to approximation errors. Such errors converge to 0 as the penalty parameter $\mu \rightarrow 0$ (cf. theorem 2.3 and theorem 3.4). We then test and compare the two algorithms in terms of a small number of iteration steps that are often designed for large-scale molecular simulations. We finally show that the VISM with the LTPB electrostatics implemented with our min–min algorithm can predict accurately the solvation free energy for a single ion. All the model parameters we use for numerical simulations are taken from [3] and are summarized in table 1.

(i) Comparison of the min–min and max–min algorithms with given tolerance to reach

We first consider a continuum charge density (cf. case 1 in §c) and compare the min–min and max–min algorithms with a given tolerance. We set the continuum charge density to be $f(r) = (1000/\sqrt{8\pi^3})e^{-50r^2}$ and also set $B(s) = s^2/2$ and $g = 0$. We minimize the total solvation free energy functionals $F_\mu = F_\mu(R)$ defined in (3.14) and $F = F(R)$ defined in (3.11) with minimizing the corresponding LTPB electrostatic energy functional $J_{R,\mu} = J_{R,\mu}[p]$ defined in (3.15) and maximizing the electrostatic energy functional $I_R = I_R[\phi]$ defined in (3.12), respectively, to compare the minimized total free energy and the optimal radius. Three values of the penalty parameter μ and three different number N of grid points are tested. The conjugate gradient method is used for minimizing the electrostatic energy functionals $J_{R,\mu}$ and $-I_R$ (equivalently maximizing I_R). The gradient descent method is used for minimizing the total free-energy functionals $F_\mu = F_\mu(R)$ and $F(R)$ to get the minimum value of free energy and also the minimizing radius. The initial guess of the radius is $R_0 = 2.5$. The tolerance for the L^2 -norm of the gradient is chosen to be 10^{-5} and maximum iteration steps is set as 50 000. The iteration of the conjugate gradient method is terminated if the tolerance or the maximum number of steps is reached. Our results are shown in table 2. We observe that as $\mu > 0$ gets smaller the error between the min–min and max–min simulations also gets smaller. This is expected by the duality of two formulations and the convergence of our penalty method; cf. theorem 3.4. Moreover, a moderate value of μ leads to the best performance of the min–min algorithm in terms of accuracy and efficiency.

Table 2. Results of min–min and max–min simulations for the continuum charge density $f(r) = (1000/\sqrt{8\pi^3})e^{-50r^2}$.

2μ	N	min–min			max–min		
		energy	radius	time	energy	radius	time
10^{-3}	250	1142.6829	3.1567	1 s	1480.0782	2.7998	2 s
	500	1146.8355	3.1568	3 s	1480.1044	2.7994	7 s
	1000	1147.8738	3.1568	16 s	1480.1331	2.7960	170 s
10^{-5}	250	1465.6298	2.8307	2 s	1480.0782	2.7998	2 s
	500	1470.5670	2.8158	12 s	1480.1044	2.7994	7 s
	1000	1471.8525	2.8196	82 s	1480.1331	2.7960	170 s
10^{-7}	250	1473.3953	2.7985	3 s	1480.0782	2.7998	2 s
	500	1478.5664	2.8019	16 s	1480.1044	2.7994	7 s
	1000	1479.6485	2.8038	456 s	1480.1331	2.7960	170 s

Table 3. Results of min–min and max–min simulations for a single-point charge $Q = 1$ placed at the origin.

2μ	N	min–min			max–min		
		energy	radius	time	energy	radius	time
10^{-1}	250	−89.6644	2.7835	1 s	−89.6230	2.8000	12 s
	500	−89.6164	2.7713	2 s	−89.6215	2.7996	21 s
	1000	−89.6009	2.7683	4 s	−89.6208	2.7997	35 s
10^{-3}	250	−89.6247	2.7951	4 s	−89.6230	2.8000	12 s
	500	−89.6259	2.7984	9 s	−89.6215	2.7996	21 s
	1000	−89.6247	2.7955	18 s	−89.6208	2.7997	35 s
10^{-5}	250	−89.6229	2.7997	72 s	−89.6230	2.8000	12 s
	500	−89.6216	2.7995	189 s	−89.6215	2.7996	21 s
	1000	−89.6210	2.7999	707 s	−89.6208	2.7997	35 s

We now consider a single-point charge $Q = 1$ placed at the origin, and set $B(s) = \cosh(s) - 1$. We use the nonlinear conjugate gradient method for the first minimization or maximization, and then the gradient descent method for the second minimization for updating the radius R with initial guess $R_0 = 2$. In the nonlinear conjugate gradient method, the tolerance for the L^2 -norm of the gradient is chosen as 10^{-5} and maximum iteration steps is set as 100 000. We test on three different values of the penalty parameter μ and also three different numbers of grid point. Our results are shown in table 3. We again observe that as $\mu > 0$ gets smaller the error between the min–min and max–min simulations also gets smaller, verifying the duality and the convergence of our penalty method; cf. theorem 3.4. It is also clear that a moderate value of μ again leads to the best performance of the min–min algorithm in terms of accuracy and efficiency.

(ii) Comparison of the min–min and max–min algorithms with a few steps of iterations

We set $A = 4$, $B(s) = s^2/2$, $f(r) = (1000/\sqrt{8\pi^3})e^{-50r^2}$, and $g = 0$. All the units are the same as in [3]; cf. table 1. We first minimize the total free energy $F(R)$ with the initial guess $R_0 = 2.5$ and a very fine grid and many iteration steps to get an ‘exact’ minimum value $F_{\min} = 1480.1331$ and an ‘exact’ optimal radius $R_{\min} = 2.7960$. We then apply the min–min algorithm to minimize the penalized VISM-LTPB functional and apply the max–min algorithm to minimize the VISM-PB functional.

Table 4. Numerical results for $f(r) = (1000/\sqrt{8\pi^3})e^{-50r^2}$ and three μ -values. The ‘Step’ means the number of steps in the conjugate gradient iteration for minimizing $J_{R,\mu}$ and $-I_R$. The ‘energy error’ and ‘radius error’ are the relative error between the numerical approximations and the ‘exact’ values F_{\min} and R_{\min} , respectively.

2μ	N	step	min–min			max–min		
			energy error	radius error	time	energy error	radius error	time
10^{-3}	250	50	0.2280	0.1288	1 s	0.7288	0.1376	1 s
	500	200	0.2252	0.1290	1 s	0.6976	0.1570	1 s
	1000	800	0.2245	0.1288	6 s	0.5725	0.1380	13 s
10^{-5}	250	300	0.0098	0.0069	5 s	0.4461	0.0430	2 s
	500	600	0.0065	0.0070	14 s	0.4999	0.0435	5 s
	1000	1500	0.0056	0.0085	72 s	0.4661	0.0947	26 s
10^{-7}	250	500	0.0046	0.0043	6 s	0.2984	0.0399	3 s
	500	800	0.0012	0.0014	32 s	0.4286	0.0261	6 s
	1000	2000	0.0002	0.0020	56 s	0.4132	0.0152	34 s

The gradient descent method is used to minimize the total energy with initial guess $R_0 = 2.5$. The conjugate gradient method with not so many iteration steps is used to minimize the LTPB energy and maximize the PB energy. In table 4, we show our numerical results. We observe that in general the min–min algorithm performs much better than the max–min algorithm with a small value of μ . For $2\mu = 10^{-3}$, the min–min algorithm does not converge due to the large penalty coefficient and small number of iteration steps.

We now compare the min–min and max–min algorithms for a single-point charge. We set the charge at the origin to be $Q = 1$ and consider $B(s) = \cosh(s) - 1$. Other parameters are the same as for the case of a continuum charge density. We use the gradient descent iteration to minimize the total free energy $\hat{F}_\mu(R)$ with the initial guess $R_0 = 2$, and use the nonlinear conjugate gradient method to maximize \hat{I}_{R_k} and minimize $\hat{J}_{R_k,\mu}$, respectively. The ‘exact’ minimum value \hat{F}_{\min} and the ‘exact’ minimizer \hat{R}_{\min} are found by a very fine grid and many iteration steps to be $\hat{F}_{\min} = -89.6208$ and $\hat{R}_{\min} = 2.7997$. Table 5 shows our numerical results. We observe that in general the min–min algorithm performs much better than the max–min algorithm in terms of the computational accuracy and efficiency. With a small but not so small μ -value, such as $2\mu = 10^{-1}$ or 10^{-3} , and with any of the number of grid points N , the min–min algorithm is more accurate than the max–min algorithm. For $2\mu = 10^{-5}$, the min–min and max–min algorithms perform compatibly in terms of the accuracy.

(iii) Prediction of the solvation free energy for single ions

We apply our min–min algorithm to minimize the penalized total VISM-LTPB free energy for the solvation of single ions K^+ , Na^+ , Cl^- and F^- . For comparison, we also minimized the VISM-PB free energy. Here, the function B is the hyperbolic cosine function, and all parameters are taken from [3] and are listed in table 1. We set $2\mu = 10^{-3}$. The dielectric boundaries of the anion Cl^- or F^- are obtained by shifting the VISM equilibrium surface by $\xi = 1\text{\AA}$, which is the length of the water OH bond [3]. In table 6, we present our numerical results with comparison with the VISM-CFA and VISM-PB computational results, and also with the experimental results. The CFA, the Coulomb-field approximation, is an approximation of the electrostatics with a dielectric boundary; cf. [41]. We see from table 6 that our numerical results fit well with the experimental data, and for some cases, are better than those of the VISM-PB calculations. It is observed that the free energy predicted by VISM-LTPB is always larger in magnitude than that by VISM-PB. This is due to the extra penalty term in the penalty method that implements VISM-LTPB formulation.

Table 5. Numerical results for a single-point charge $Q = 1$ placed at the origin. The ‘Step’ means the number of steps in the nonlinear conjugate gradient iteration for minimizing $\hat{J}_{R,\mu}$ and $-\hat{I}_R$. The ‘energy error’ is the relative error between the numerical minimum value of \hat{F}_μ and the ‘exact’ minimum value \hat{F}_{\min} . The ‘radius error’ is the relative error between the numerical minimizer of \hat{F}_μ and the ‘exact’ minimizer \hat{R}_{\min} .

2μ	N	step	min–min			max–min		
			energy error	radius error	time	energy error	radius error	time
10^{-1}	250	5	0.0028	0.0212	1 s	0.0306	0.0302	1 s
	500	5	0.0032	0.0221	1 s	0.0857	0.0416	1 s
	1000	5	0.0045	0.0239	1 s	0.1317	0.0566	1 s
10^{-3}	250	10	0.0022	0.0018	1 s	0.0058	0.0268	1 s
	500	10	0.0032	0.0019	1 s	0.0252	0.0272	1 s
	1000	10	0.0041	0.0021	1 s	0.0333	0.0507	1 s
10^{-5}	250	30	0.0046	0.0010	3 s	0.0032	0.0236	1 s
	500	30	0.0088	0.0035	8 s	0.0062	0.0336	1 s
	1000	30	0.0489	0.0027	12 s	0.0259	0.0216	1 s

Table 6. The solvation free energy (in $k_B T$) for each of the single ions K^+ , Na^+ , Cl^- and F^- modelled as a single-point charge computed with the VISM-LTPB model. The computational results based on the VISM-CFA and VISM-PB models and the experimental results are also shown for comparison [3,42].

Ion s	ε ($k_B T$)	σ (Å)	VISM-CFA	VISM-PB	VISM-LTPB	experiment
K^+	0.008	3.85	−111.7	−112.5760	−112.5854	−117.5
Na^+	0.008	3.49	−130.5	−131.5733	−131.5842	−145.4
Cl^-	0.21	3.78	−126.1	−127.6528	−127.7173	−135.4
F^-	0.219	3.3	−171.6	−173.0490	−173.0998	−185.2

4. Conclusion

We constructed the LTPB electrostatic energy functional of dielectric displacements with application to variational solvation of charged molecules that are characterized by the dielectric boundary separating such solute molecules from the solvent. The solvation free energy in the variational solvation model includes the surface energy, electrostatic energy and other energy terms. The convexity of the LTPB energy functional makes it consistent with the minimization of the total solvation free-energy functional.

We proved the duality between the convex LTPB functional and the classical concave PB electrostatic energy functional. With a fixed dielectric boundary, we approximated our LTPB functional of the dielectric displacements constrained by Gauss’ law on the solute region by penalized LTPB functionals, removing the constraint. The convergence of this penalty method was shown. Finally, we designed a min–min method for minimizing the solvation free-energy functional of all dielectric boundaries: in each iteration step of relaxing the dielectric boundary, we minimize iteratively the LTPB electrostatic energy. The convergence of the min–min algorithm was proved. Numerical tests on the solvation of single ions demonstrated the efficiency and accuracy of our method, and for many cases, with suitably chosen penalty parameters and number of iteration steps, the LTPB formulation was more stable than the classical PB formulation.

We now discuss some possible issues and point out possible improvements for future studies. First, the explicit formula of the Legendre transform $B^* = B^*(\xi)$ of a given convex function $B = B(s)$ is generally not available. One can, however, generate a table of values $B^*(\xi)$ for selected values of $\xi \in [\xi_{\min}, \xi_{\max}]$, where the numbers ξ_{\min} and ξ_{\max} can be estimated from an underlying system. Note that the function $B = B(s)$ in the generalized PB theory with ionic size effect is only implicitly defined [16,34,35]. One can, however, solve a system of nonlinear algebraic equations to obtain $B(s)$ for many selected s -values and then calculate $B^*(\xi)$ to generate a table.

Second, while our initial numerical tests have indicated that the LTPB formulation is better than the classical PB formulation, the constraint of the dielectric displacements slowed down the computations. Therefore, there is a need to construct a more efficient LTPB formulation for modelling the electrostatics in molecular solvation.

Third, for large systems of molecular solvation, the level-set method can be used to numerically relax the solvation free-energy functional with the LTPB formulation for the electrostatic energy [3]. It will also be interesting to use the level-set method to minimize the total solvation free energy of dielectric boundaries and compare the min–min and max–min algorithms, and hence compare the LTPB and PB formulations.

Finally, our convergence analysis of the min–min algorithm points out a new scheme of minimizing a free-energy functional of two variables with one depending on the other. One can view the free-energy functional as a two-variable functional and minimize it with respect to the two variables; cf. lemma 3.2. In application to solvation, we can minimize the solvation free energy with respect to both the dielectric boundary Γ and the dielectric displacement D .

Data accessibility. The code and data are provided in the electronic supplementary material [43].

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. Z.H.: conceptualization, data curation, formal analysis, investigation, methodology, resources, software, validation, visualization, writing—original draft, writing—review and editing; B.L.: conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, validation, writing—original draft, writing—review and editing.

Both authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This work was supported in part by the US National Science Foundation (grant no. DMS-2208465).

Acknowledgements. The authors thank Prof. Li-Tien Cheng, Dr Benjamin Ciotti, Prof. Shuang Liu, Prof. Anthony C. Maggs, Dr Zirui Zhang and Prof. Shenggao Zhou for helpful discussions. The authors also thank the anonymous referees for their valuable comments and suggestions.

Appendix A

We derive the formula of the dielectric boundary force as negative energy variation for a model problem. Let $\varepsilon_-, \varepsilon_+ \in (0, \infty)$ with $\varepsilon_- \neq \varepsilon_+$. For any $\gamma \in I := (0, 1)$, we define $\varepsilon_\gamma : [0, 1] \rightarrow \mathbb{R}$ by $\varepsilon_\gamma(x) = \varepsilon_-$ if $x < \gamma$ and $\varepsilon_\gamma(x) = \varepsilon_+$ if $x > \gamma$. Let $f : [0, 1] \rightarrow \mathbb{R}$ be a smooth function. Define $E_\gamma : H_0^1(I) \rightarrow \mathbb{R}$ by

$$E_\gamma[\phi] = \int_0^1 \left(\frac{\varepsilon_\gamma}{2} \phi'^2 - f\phi \right) dx \quad \forall \phi \in H_0^1(I).$$

Let $\phi_\gamma \in H_0^1(I)$ be the unique minimizer of E_γ over $H_0^1(I)$ and denote $e(\gamma) := E_\gamma[\phi_\gamma] = \min_{\phi \in H_0^1(I)} E_\gamma[\phi]$. The minimizer ϕ_γ is the unique function in $H_0^1(I)$ such that

$$\int_0^1 (\varepsilon_\gamma \phi'_\gamma \eta' - f\eta) dx = 0 \quad \forall \eta \in H_0^1(I). \quad (\text{A } 1)$$

Equivalently, $\phi_\gamma \in H_0^1(I)$ is the unique function satisfying $\phi_\gamma|_{\Omega_\pm} \in C^2(\overline{I_\pm})$ and

$$-\varepsilon_\pm \phi''_\gamma = f \quad \text{in } I_\pm \quad \text{and} \quad \llbracket \phi_\gamma \rrbracket_\gamma = \llbracket \varepsilon_\gamma \phi'_\gamma \rrbracket_\gamma = 0, \quad (\text{A } 2)$$

where $I_- = (0, \gamma)$ and $I_+ = (\gamma, 1)$, and $\llbracket u \rrbracket_\gamma = u(\gamma+) - u(\gamma-)$ for a given function u .

The boundary force is defined as $-e'(\gamma)$ if the derivative exists. Formulae for such forces for more general problems in multi-dimensions and nonlinear Euler–Lagrange equations have been obtained in [18,19]. Here, we give a totally different, direct and self-closed derivation of such a force.

Proposition A.1. *We have*

$$e'(\gamma) = \frac{1}{2} \left(\frac{1}{\varepsilon_+} - \frac{1}{\varepsilon_-} \right) [\varepsilon_\gamma \phi'_\gamma(\gamma)]^2,$$

where $\varepsilon_\gamma \phi'_\gamma(\gamma) = \varepsilon_- \phi'_\gamma(\gamma-) = \varepsilon_+ \phi'_\gamma(\gamma+)$.

Proof. We first establish some bounds for ϕ_γ and $\delta\phi_\gamma := \phi_{\gamma+\delta\gamma} - \phi_\gamma$ ($\delta\gamma \in \mathbb{R}$ and $0 < \gamma + \delta\gamma < 1$). Setting $\eta = \phi_\gamma$ in (A 1), we have by Poincaré’s inequality that $\|\phi_\gamma\|_{H_0^1(I)} \leq C$ for all $\gamma \in I$, where $C > 0$ is a constant independent of $\gamma \in I$. Since $\phi_\gamma(0) = 0$, we have $\phi_\gamma(x) = \int_0^x \phi'_\gamma(t) dt$ ($x \in I$). Hence, $\|\phi_\gamma\|_{L^\infty(I)} \leq C$ for all $\gamma \in I$.

Let $0 < \gamma_1 < \gamma < \gamma_2 < 1$. By (A 2) and the fact that $\phi_\gamma(1) = 0$, we have

$$-\varepsilon_+ \phi'_\gamma(y) + \varepsilon_+ \phi'_\gamma(\gamma+) = \int_\gamma^y f(t) dt \quad \forall y \in (\gamma, 1)$$

and

$$\varepsilon_+ \phi_\gamma(x) + \varepsilon_+ \phi'_\gamma(\gamma+)(1-x) = \int_x^1 \int_\gamma^y f(t) dt dy \quad \forall x \in (\gamma, 1).$$

These and the uniform bound on $\|\phi_\gamma\|_{L^\infty(I)}$ imply that $|\phi'_\gamma(\gamma+)| \leq C(\gamma_1, \gamma_2)$, and hence $|\phi'_\gamma(y)| \leq C(\gamma_1, \gamma_2)$ for all $y \in [\gamma, \gamma_2]$, where $C(\gamma_1, \gamma_2) > 0$ is a generic constant independent of $\gamma \in [\gamma_1, \gamma_2]$. Similarly, $|\phi'_\gamma(y)| \leq C(\gamma_1, \gamma_2)$ for all $y \in [\gamma_1, \gamma]$. Thus, $\|\phi_\gamma\|_{W^{1,\infty}([\gamma_1, \gamma_2])} \leq C(\gamma_1, \gamma_2)$ for all $\gamma \in [\gamma_1, \gamma_2]$.

Let $\delta\gamma \in \mathbb{R}$ be such that $0 < \gamma_1 < \gamma - |\delta\gamma| < \gamma + |\delta\gamma| < \gamma_2 < 1$. We shall assume that $\delta\gamma > 0$ as the case $\delta\gamma < 0$ is similar. Setting $\eta = \delta\phi_\gamma = \phi_{\gamma+\delta\gamma} - \phi_\gamma$ in (A 1) for ϕ_γ and in (A 1) with $\phi_{\gamma+\delta\gamma}$ replacing ϕ_γ , respectively, and subtracting one from the other, we obtain

$$\int_0^1 [\varepsilon_{\gamma+\delta\gamma} (\delta\phi_\gamma)' + (\varepsilon_{\gamma+\delta\gamma} - \varepsilon_\gamma) \phi'_\gamma] (\delta\phi_\gamma)' dx = 0.$$

This, together with the bound $\|\phi_\gamma\|_{W^{1,\infty}([\gamma_1, \gamma_2])} \leq C(\gamma_1, \gamma_2)$, implies that

$$\begin{aligned} \int_0^1 \varepsilon_{\gamma+\delta\gamma} (\delta\phi_\gamma)^2 dx &= \int_0^1 (\varepsilon_\gamma - \varepsilon_{\gamma+\delta\gamma}) \phi'_\gamma (\delta\phi_\gamma)' dx \\ &= (\varepsilon_+ - \varepsilon_-) \int_\gamma^{\gamma+\delta\gamma} \phi'_\gamma (\delta\phi_\gamma)' dx \leq C(\gamma_1, \gamma_2) \sqrt{\delta\gamma} \|(\delta\phi_\gamma)'\|_{L^2(I)} \quad \forall \gamma \in [\gamma_1, \gamma_2], \end{aligned}$$

where $C(\gamma_1, \gamma_2) > 0$ is independent of γ and $\delta\gamma$. Consequently,

$$\|\delta\phi_\gamma\|_{H_0^1(I)} \leq C(\gamma_1, \gamma_2) \sqrt{\delta\gamma} \quad \forall \gamma \in [\gamma_1, \gamma_2]. \quad (\text{A } 3)$$

We now calculate $e'(\gamma)$. Denote $\delta e_\gamma = e(\gamma + \delta\gamma) - e(\gamma)$ with $\delta\gamma > 0$ as assumed above. By (A 1) with $\eta = \delta\phi_\gamma$ for ϕ_γ and (A 1) with $\phi_{\gamma+\delta\gamma}$ replacing ϕ_γ , we obtain

$$\begin{aligned} \frac{\delta e_\gamma}{\delta\gamma} &= \frac{1}{\delta\gamma} \int_0^1 \left[\frac{1}{2} \varepsilon_{\gamma+\delta\gamma} \phi_{\gamma+\delta\gamma}^2 - \frac{1}{2} \varepsilon_\gamma \phi_\gamma^2 - f \delta\phi_\gamma \right] dx \\ &= \frac{1}{2\delta\gamma} \int_0^1 [\varepsilon_{\gamma+\delta\gamma} \phi_{\gamma+\delta\gamma}^2 - \varepsilon_\gamma \phi_\gamma^2 - \varepsilon_{\gamma+\delta\gamma} \phi'_{\gamma+\delta\gamma} (\phi'_{\gamma+\delta\gamma} - \phi'_\gamma) - \varepsilon_\gamma \phi'_\gamma (\phi'_{\gamma+\delta\gamma} - \phi'_\gamma)] dx \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2\delta\gamma} \int_0^1 (\varepsilon_{\gamma+\delta\gamma} - \varepsilon_\gamma) \phi'_{\gamma+\delta\gamma} \phi'_\gamma \, dx \\
&= \frac{\varepsilon_- - \varepsilon_+}{2\delta\gamma} \int_\gamma^{\gamma+\delta\gamma} \phi'_{\gamma+\delta\gamma} \phi'_\gamma \, dx \\
&= \frac{\varepsilon_- - \varepsilon_+}{2} \left[\frac{1}{\delta\gamma} \int_\gamma^{\gamma+\delta\gamma} \phi'_{\gamma+\delta\gamma}(x) [\phi'_\gamma(x) - \phi'_\gamma(\gamma+)] \, dx + \frac{\phi'_\gamma(\gamma+)}{\delta\gamma} \int_\gamma^{\gamma+\delta\gamma} \phi'_{\gamma+\delta\gamma}(x) \, dx \right] \\
&=: \frac{\varepsilon_- - \varepsilon_+}{2} (I_1 + I_2).
\end{aligned} \tag{A 4}$$

It follows from (A 2), the Cauchy–Schwarz inequality, and the bound $\|\phi_\gamma\|_{H^1(I)} \leq C$ that

$$\begin{aligned}
|I_1| &= \left| \frac{1}{\delta\gamma} \int_\gamma^{\gamma+\delta\gamma} \phi'_{\gamma+\delta\gamma}(x) \left(\int_\gamma^x \phi''_\gamma(y) \, dy \right) dx \right| \\
&= \left| \frac{1}{\delta\gamma} \int_\gamma^{\gamma+\delta\gamma} \phi'_{\gamma+\delta\gamma}(x) \left(\int_\gamma^x -\frac{1}{\varepsilon_+} f(y) \, dy \right) dx \right| \\
&\leq C\sqrt{\delta\gamma},
\end{aligned} \tag{A 5}$$

where $C > 0$ is a constant independent of γ and $\delta\gamma$.

We have

$$I_2 = \frac{\phi'_\gamma(\gamma+)}{\delta\gamma} \left[\int_\gamma^{\gamma+\delta\gamma} \phi'_{\gamma+\delta\gamma}(x) \, dx - \int_{\gamma-\delta\gamma}^\gamma \phi'_\gamma(x) \, dx \right] + \frac{\phi'_\gamma(\gamma+)}{\delta\gamma} \int_{\gamma-\delta\gamma}^\gamma \phi'_\gamma(x) \, dx =: I_{2,1} + I_{2,2}. \tag{A 6}$$

Clearly,

$$I_{2,2} = \frac{\phi'_\gamma(\gamma+)}{\delta\gamma} \int_{\gamma-\delta\gamma}^\gamma \phi'_\gamma(x) \, dx \rightarrow \phi'_\gamma(\gamma+) \phi'_\gamma(\gamma-) \quad \text{as } \delta\gamma \rightarrow 0. \tag{A 7}$$

By the change of variable, we get

$$\begin{aligned}
I_{2,1} &= \frac{\phi'_\gamma(\gamma+)}{\delta\gamma} \int_{\gamma-\delta\gamma}^\gamma [\phi'_{\gamma+\delta\gamma}(y + \delta\gamma) - \phi'_\gamma(y)] dy \\
&= \frac{\phi'_\gamma(\gamma+)}{\delta\gamma} \int_{\gamma-\delta\gamma}^\gamma [\phi'_{\gamma+\delta\gamma}(y + \delta\gamma) - \phi'_{\gamma+\delta\gamma}(y)] dy + \frac{\phi'_\gamma(\gamma+)}{\delta\gamma} \int_{\gamma-\delta\gamma}^\gamma [\phi'_{\gamma+\delta\gamma}(y) - \phi'_\gamma(y)] dy \\
&=: I_{2,1,1} + I_{2,1,2}.
\end{aligned} \tag{A 8}$$

It follows from (A 2) that

$$\begin{aligned}
I_{2,1,1} &= \frac{\phi'_\gamma(\gamma+)}{\delta\gamma} \int_{\gamma-\delta\gamma}^\gamma [\phi'_{\gamma+\delta\gamma}(y + \delta\gamma) - \phi'_{\gamma+\delta\gamma}(y)] dy \\
&= \frac{\phi'_\gamma(\gamma+)}{\delta\gamma} \int_{\gamma-\delta\gamma}^\gamma \left[\int_y^{y+\delta\gamma} \phi''_{\gamma+\delta\gamma}(x) \, dx \right] dy \\
&= \frac{\phi'_\gamma(\gamma+)}{\delta\gamma} \int_{\gamma-\delta\gamma}^\gamma \left[-\frac{1}{\varepsilon_-} \int_y^{y+\delta\gamma} f(x) \, dx \right] dy \\
&\rightarrow 0 \quad \text{as } \delta\gamma \rightarrow 0.
\end{aligned} \tag{A 9}$$

By (A 2) again, we have

$$(\delta\phi_\gamma)''(y) = \phi''_{\gamma+\delta\gamma}(y) - \phi''_\gamma(y) = -\left(\frac{1}{\varepsilon_-}\right)f(y) + \left(\frac{1}{\varepsilon_+}\right)f(y) = 0 \quad \forall y \in (\gamma - \delta\gamma, \gamma).$$

Hence, $(\delta\phi_\gamma)' = a$ on $(\gamma - \delta\gamma, \gamma)$ for some constant $a \in \mathbb{R}$. By (A 3), $a = 0$. Therefore,

$$I_{2,1,2} = \frac{\phi'_\gamma(\gamma+)}{\delta\gamma} \int_{\gamma-\delta\gamma}^\gamma [\phi'_{\gamma+\delta\gamma}(y) - \phi'_\gamma(y)] dy = \frac{\phi'_\gamma(\gamma+)}{\delta\gamma} \int_{\gamma-\delta\gamma}^\gamma (\delta\phi_\gamma)'(y) \, dy = 0. \tag{A 10}$$

It now follows from (A 4) to (A 10) and the jump condition in (A 2) that

$$e'(\gamma) = \lim_{\delta\gamma \rightarrow 0} \frac{\delta e(\gamma)}{\delta\gamma} = \frac{1}{2}(\varepsilon_- - \varepsilon_+) \phi'_\gamma(\gamma_+) \phi'_\gamma(\gamma_-) = \frac{1}{2} \left(\frac{1}{\varepsilon_+} - \frac{1}{\varepsilon_-} \right) [\varepsilon_\gamma \phi'_\gamma(\gamma)]^2.$$

References

1. Dzubiella J, Swanson MJ, McCammon JA. 2006 Coupling hydrophobicity, dispersion, and electrostatics in continuum solvent models. *Phys. Rev. Lett.* **96**, 087802. (doi:10.1103/PhysRevLett.96.087802)
2. Dzubiella J, Swanson MJ, McCammon JA. 2006 Coupling nonpolar and polar solvation free energies in implicit solvent models. *J. Chem. Phys.* **124**, 084905. (doi:10.1063/1.2171192)
3. Zhou S, Cheng LT, Dzubiella J, Li B, McCammon JA. 2014 Variational implicit solvation with Poisson–Boltzmann theory. *J. Chem. Theory Comput.* **10**, 1454–1467. (doi:10.1021/ct401058w)
4. Sun H, Wen J, Zhao Y, Li B, McCammon JA. 2015 A self-consistent phase-field approach to implicit solvation of charged molecules with Poisson–Boltzmann electrostatics. *J. Chem. Phys.* **143**, 243110. (doi:10.1063/1.4932336)
5. Bates PW, Chen Z, Sun YH, Wei GW, Zhao S. 2009 Geometric and potential driving formation and evolution of biomolecular surfaces. *J. Math. Biol.* **59**, 193–231. (doi:10.1007/s00285-008-0226-7)
6. Lum K, Chandler D, Weeks JD. 1999 Hydrophobicity at small and large length scales. *J. Phys. Chem. B* **103**, 4570–4577. (doi:10.1021/jp984327m)
7. Ramirez R, Borgis D. 2005 Density functional theory of solvation and its relation to implicit solvent models. *J. Phys. Chem. B* **109**, 6754–6763. (doi:10.1021/jp045453v)
8. Andelman D. 1995 Electrostatic properties of membranes: the Poisson–Boltzmann theory. In *Handbook of biological physics* (eds R Lipowsky, E Sackmann), vol. 1, pp. 603–642. Elsevier.
9. Sharp KA, Honig B. 1990 Electrostatic interactions in macromolecules: theory and applications. *Annu. Rev. Biophys. Biophys. Chem.* **19**, 301–332. (doi:10.1146/annurev.bb.19.060190.001505)
10. Davis ME, McCammon JA. 1990 Electrostatics in biomolecular structure and dynamics. *Chem. Rev.* **90**, 509–521. (doi:10.1021/cr00101a005)
11. Gilson MK, Davis ME, Luty BA, McCammon JA. 1993 Computation of electrostatic forces on solvated molecules using the Poisson–Boltzmann equation. *J. Phys. Chem.* **97**, 3591–3600. (doi:10.1021/j100116a025)
12. Che J, Dzubiella J, Li B, McCammon JA. 2008 Electrostatic free energy and its variations in implicit solvent models. *J. Phys. Chem. B* **112**, 3058–3069. (doi:10.1021/jp7101012)
13. Kralj-Iglić V, Iglič A. 1996 A simple statistical mechanical approach to the free energy of the electric double layer including the excluded volume effect. *J. Phys. II (France)* **6**, 477–491. (doi:10.1051/jp2:1996193)
14. Borukhov I, Andelman D, Orland H. 1997 Steric effects in electrolytes: a modified Poisson–Boltzmann equation. *Phys. Rev. Lett.* **79**, 435–438. (doi:10.1103/PhysRevLett.79.435)
15. Li B. 2009 Minimization of electrostatic free energy and the Poisson–Boltzmann equation for molecular solvation with implicit solvent. *SIAM J. Math. Anal.* **40**, 2536–2566. (doi:10.1137/080712350)
16. Li B. 2009 Continuum electrostatics for ionic solutions with nonuniform ionic sizes. *Nonlinearity* **22**, 811–833. (doi:10.1088/0951-7715/22/4/007)
17. Reiner ES, Radke CJ. 1990 Variational approach to the electrostatic free energy in charged colloidal suspensions: general theory for open systems. *J. Chem. Soc. Faraday Trans.* **86**, 3901–3912. (doi:10.1039/ft9908603901)
18. Li B, Cheng XL, Zhang ZF. 2011 Dielectric boundary force in molecular solvation with the Poisson–Boltzmann free energy: a shape derivative approach. *SIAM J. Appl. Math.* **71**, 2093–2111. (doi:10.1137/110826436)
19. Li B, Zhang Z, Zhou S. 2021 The calculus of boundary variations and the dielectric boundary force in the Poisson–Boltzmann theory for molecular solvation. *J. Nonlinear Sci.* **31**, 1–50. (doi:10.1007/s00332-021-09749-7)
20. Fogolari F, Briggs JM. 1997 On the variational approach to Poisson–Boltzmann free energies. *Chem. Phys. Lett.* **281**, 135–139. (doi:10.1016/S0009-2614(97)01193-7)

21. Ciotti B, Li B. 2018 Legendre transforms of electrostatic free-energy functionals. *SIAM J. Applied Math.* **78**, 2973–2995. (doi:10.1137/18M1167644)
22. Maggs AC. 2012 A minimizing principle for the Poisson–Boltzmann equation. *Europhys. Lett.* **98**, 16012. (doi:10.1209/0295-5075/98/16012)
23. Pujos JS, Maggs AC. 2015 Convexity and stiffness in energy functions for electrostatic simulations. *J. Chem. Theory Comput.* **11**, 1419–1427. (doi:10.1021/acs.jctc.5b00023)
24. Blossey R, Maggs AC, Podgornik R. 2017 Structural interactions in ionic liquids linked to higher-order Poisson–Boltzmann equations. *Phys. Rev. E* **95**, 060602. (doi:10.1103/PhysRevE.95.060602)
25. Maggs AC, Podgornik R. 2016 General theory of asymmetric steric interactions in electrostatic double layers. *Soft Matter* **12**, 1219–1229. (doi:10.1039/C5SM01757B)
26. Rockafellar RT. 1970 *Convex analysis*. Princeton NJ: Princeton University Press.
27. Zia RKP, Redish EF, McKay SR. 2009 Making sense of the Legendre transform. *Am. J. Phys.* **77**, 614–622. (doi:10.1119/1.3119512)
28. Bikerman JJ. 1942 XXXIX. Structure and capacity of electrical double layer. *Lond. Edinb., Dublin Phil. Mag. J. Sci.* **33**, 384–397. (doi:10.1080/14786444208520813)
29. Eigen M, Wicke E. 1954 The thermodynamics of electrolytes at higher concentration. *J. Phys. Chem.* **58**, 702–714. (doi:10.1021/j150519a007)
30. Kilic MS, Bazant MZ, Ajdari A. 2007 Steric effects in the dynamics of electrolytes at large applied voltages. I. Double-layer charging. *Phys. Rev. E* **75**, 021502. (doi:10.1103/PhysRevE.75.021502)
31. Kilic MS, Bazant MZ, Ajdari A. 2007 Steric effects in the dynamics of electrolytes at large applied voltages. II. Modified Poisson–Nernst–Planck equations. *Phys. Rev. E* **75**, 021503. (doi:10.1103/PhysRevE.75.021503)
32. Tresset G. 2008 Generalized Poisson–Fermi formalism for investigating size correlation effects with multiple ions. *Phys. Rev. E* **78**, 061506. (doi:10.1103/PhysRevE.78.061506)
33. Lu BZ, Zhou YC. 2011 Poisson–Nernst–Planck equations for simulating biomolecular diffusion-reaction processes II: size effects on ionic distributions and diffusion-reaction rates. *Biophys J.* **100**, 2475–2485. (doi:10.1016/j.bpj.2011.03.059)
34. Zhou S, Wang Z, Li B. 2011 Mean-field description of ionic size effects with non-uniform ionic sizes: a numerical approach. *Phys. Rev. E* **84**, 021901. (doi:10.1103/PhysRevE.84.021901)
35. Li B, Liu P, Xu Z, Zhou S. 2013 Ionic size effects: generalized Boltzmann distributions, counterion stratification, and modified Debye length. *Nonlinearity* **26**, 2899–2922. (doi:10.1088/0951-7715/26/10/2899)
36. Temam R. 1984 *Navier–Stokes equations: theory and numerical analysis*, 3rd edn. Amsterdam, The Netherlands: North-Holland.
37. Wang S, Lee A, Alexov E, Zhao S. 2020 A regularization approach for solving Poisson’s equation with singular charge sources and diffuse interfaces. *Appl. Math. Lett.* **102**, 106144. (doi:10.1016/j.aml.2019.106144)
38. Li B, Wen J, Zhou S. 2016 Mean-field theory and computation of electrostatics with ionic concentration dependent dielectrics. *Commun. Math. Sci.* **14**, 249–271. (doi:10.4310/CMS.2016.v14.n1.a10)
39. Liu X, Qiao Y, Lu BZ. 2018 Analysis of the mean field free energy functional of electrolyte solution with nonhomogenous boundary conditions and the generalized PB/PNP equations with inhomogenous dielectric permittivity. *SIAM J. Appl. Math.* **78**, 1131–1154. (doi:10.1137/16M1108583)
40. Cheng LT, Dzubiella J, McCammon JA, Li B. 2007 Application of the level-set method to the implicit solvation of nonpolar molecules. *J. Chem. Phys.* **127**, 084503. (doi:10.1063/1.2757169)
41. Wang Z, Che J, Cheng LT, Dzubiella J, Li B, McCammon JA. 2012 Level-set variational implicit solvation with the Coulomb-field approximation. *J. Chem. Theory Comput.* **8**, 386–397. (doi:10.1021/ct200647j)
42. Marcus Y. 1991 Thermodynamics of solvation of ions. Part 5.– Gibbs free energy of hydration at 298.15 K. *J. Chem. Soc. Faraday Trans.* **87**, 2995–2999. (doi:10.1039/FT9918702995)
43. Huang Z, Li B. 2024 Variational implicit solvation with Legendre-transformed Poisson–Boltzmann electrostatics. Figshare. (doi:10.6084/m9.figshare.c.7126153)