12-30-2016

Minimization and Eulerian Formulation of Differential Geometry Based Nonpolar Multiscale Solvation Models

Zhan Chen

Georgia Southern University, zchen@georgiasouthern.edu

Research Article   **Open Access**

Zhan Chen*

Minimization and Eulerian Formulation of Differential Geormetry Based Nonpolar Multiscale Solvation Models

Abstract: In this work, the existence of a global minimizer for the previous Lagrangian formulation of nonpolar solvation model proposed in [1] has been proved. One of the proofs involves a construction of a phase field model that converges to the Lagrangian formulation. Moreover, an Eulerian formulation of nonpolar solvation model is proposed and implemented under a similar parameterization scheme to that in [1]. By doing so, the connection, similarity and difference between the Eulerian formulation and its Lagrangian counterpart can be analyzed. It turns out that both of them have a great potential in solvation prediction for nonpolar molecules, while their decompositions of attractive and repulsive parts are different. That indicates a distinction between phase field models of solvation and our Eulerian formulation.

Keywords: Differential geometry based multiscale model, Nonpolar solvation free energy, Minimization, Eulerian formulation

1 Introduction

Solvation is an elementary process in nature and is of paramount importance to sophisticated chemical, biological and biomolecular processes including signal transduction, DNA recognition, transcription, post-translational modification, translation, protein folding and protein ligand binding. The understanding of solvation is an essential prerequisite for the quantitative description and analysis of biomolecular systems. In particular, the calculation of solvation free energy has captured a great deal of interests [2]. Solvation models have been developed ranging from simple phenomenological modifications of Coulomb's law, implicit solvent models that describe the solvent by mean-field approximations [3–7], explicit solvent models that treat the solvent in molecular or atomic detail [8], to complex quantum mechanical methods [9–12]. Implicit solvent methods become very popular in solvation analysis and the quantitative description of general biological processes due to their promising balance between accuracy and efficiency.

Solvation free energy is a physical quantity that can be measured by experiments. To calculate the solvation energy, it is helpful to break down the solvation process into a nonpolar process of inserting the uncharged solute into solvent and a polar process of charging the solute in vacuum and solvent [13]. Nonpolar solvation is generally associated with the insertion of the uncharged solute into solvent. There are many nonpolar solvation models available, among which a solvent-accessible surface area (SASA) model is commonly used. SASA models state that nonpolar solvent-solute interactions are proportional to the area of the solvent-solute interface. It is worthy to note that the proportional constant varies dramatically in the literature because the energies of other processes are also assumed to be proportional to SASA [14]. Roughly speaking, SASA models are based on the scaled particle theory (SPT) [5, 15] which actually includes the energy of surface tension effect and the mechanical work of immersing a particle into solvent. Studies indicate that the

---

*Corresponding Author: Zhan Chen: Department of Mathematical Sciences, Georgia Southern University, Statesboro, GA, 30460,USA, E-mail: zchen@georgiasouthern.edu

nonpolar distribution should depend on the solvent-accessible volume and surface area, with a crossover to SASA when the size of solute is large [16]. Moreover, recent works by Levy, Gallicchio, and others [16–20] have demonstrated the importance of the treatment of attractive solute-solvent dispersion terms in nonpolar solvation models.

Recently, intensive efforts have been made to develop differential geometry (DG) based multiscale models for solvation analysis, electrostatic field and transport of complex chemical and biological systems [1, 21–23]. One essential feature of such models is the use of differential geometry of surface as a natural means to dynamically couple a discrete description of solute and continuum description of solvent. The main idea is to construct a total energy functional of the system to encompass energies of interest. Then variational analysis leads to desired partial differential equations (PDE) for surface generation and other unknowns. DG based solvation models can be implemented in an Eulerian formulation [22, 23] and a Lagrangian formulation [13]. While the Lagrangian formulation based DG solvation models are similar to those of Dzubiella *et al.* in spirit [24, 25], the discription of surface tension, nonpolar solvation free energy and the treatment of the solute-solvent interface are different in these two approaches.

Our DG based nonpolar solvation model in Lagrangian formulation was able to predict nonpolar solvation energy in excellent agreement with experimental data for a large number of nonpolar molecules [1]. With implicit solvent, we use the following model for nonpolar solvation free energies [16]

$$G_{\mathrm{np}} = \gamma \cdot \mathrm{Area} + p \cdot \mathrm{Vol} + \int_{\Omega_s} \rho_s U^{vdW} d\mathbf{r}, \tag{1}$$

where γ is the surface tension, 'Area' is the area of the solute, p is the hydrodynamic pressure, 'Vol' is the volume of the solute, ρ_s is the solvent density, Ω_s denotes the solvent region, and U^{vdW} is the solvent-solute van der Waals (vdW) interaction potential. The first two terms in Eq. (1) are those from the SPT model [5, 15]. The first term is the surface energy. It measures the disruption of intermolecular and/or intramolecular bonds that occurs when a surface is created. The second term is the mechanical work of creating the vacuum of a biomolecule in the solvent. The third term represents the attractive dispersion effect near the solvent-solute interface which has been shown by Wagoner and Baker [16] to have played a crucial role in an accurate nonpolar solvation analysis. The nonpolar solvation formula (1) has been demonstrated to be in a good agreement with explicit solvent solvation forces on proteins [16] and RNA hairpins [6]. Work by Levy and co-workers has demonstrated the good performance of a similar model [17–20].

The objective of this work is three-fold. First, we theoretically prove the existence of a global minimizer in our nonpolar solvation free energy functional. This provides a solid mathematical basis for previous numerical implementations [1]. Second, we present an Eulerian formulation of differential geometry based nonpolar solvation model and its performance under a similar parameter setting to its Lagrangian counterpart [1]. By doing so, we are able to analyze the connection, similarity and difference between the Lagrangian and Eulerian formulations. Third, we look into the distinction between phase field models of solvation [26, 27] and our Eulerian formulation.

The rest of this paper is organized as follows. We first prove the existence of a global minimizer for our Lagrangian formulation of nonpolar solvation model proposed in [1]. Two proofs are provided here. One is a direct approach based on the sharp interface formulation. The other one is inspired by a recent work [27] using a phase field model. They are followed by a brief description of the numerical scheme to find the minimizer. Then an Eulerian formulation of DG based nonpolar solvation model is proposed with its implementation in the solvation calculation for nonpolar molecules. Finally, numerical results from our Eulerian formulation are demonstrated. This paper ends with a conclusion remark.

2 Minimization of nonpolar solvation free energy and its numerical scheme

In our Lagrangian formulation of DG based nonpolar solvation model, the interface can be represented as a closed surface in the 3D Euclidean space, denoted as Γ. The solute region $\Omega_m(\Gamma)$ and the solvent region $\Omega_s(\Gamma)$ are separated by Γ and therefore they can be regarded as functions of Γ. We express the surface area and volume in Eq. (1) as the following integrals

$$\text{Area} = \int_{\Gamma} d\sigma \quad \text{and} \quad \text{Vol} = \int_{\Omega_m} d\mathbf{x}, \tag{2}$$

where $d\sigma$ represents the infinitesimal surface element on the solute-solvent interface. The van der Waals (vdW) potential U^{vdW} is computed pairwisely for each atom $U^{\text{vdW}}(\mathbf{r}) = \sum_i U_i^{\text{LJ}}(\mathbf{r})$. The formulation is under the assumption that the nonpolar solute-solvent potential is pairwise. A 6-12 L-J pair potential is formulated as follows:

$$U_i^{\text{LJ}}(\mathbf{r}) = 4\epsilon_{is}\left[\left(\frac{\sigma_{is}}{\|\mathbf{r}-\mathbf{r}_i\|}\right)^{12} - \left(\frac{\sigma_{is}}{\|\mathbf{r}-\mathbf{r}_i\|}\right)^6\right], \tag{3}$$

Eq. (3) is used in OPLS-AA force field [19, 28] where well depth parameters are $\epsilon_{is} = \sqrt{\epsilon_s\epsilon_i}$, and $\sigma_{is} = 2\sqrt{\sigma_i\sigma_s}$ in which σ_i and σ_s are solute atomic and solvent radii, respectively. In this work, $\sigma_{is} = \sigma_i + \sigma_s$. Here $\mathbf{r}$ is the position variable and $\mathbf{r}_i$ is the position of the i-th atom.

Then Eq. (1) becomes

$$G_{\text{np}}^0[\Gamma] = \gamma \cdot \int_{\Gamma} d\sigma + p \cdot \int_{\Omega_m} dx + \int_{\Omega_s} \rho_s U^{vdW} dx, \tag{4}$$

Note that $U^{\text{vdW}}(\mathbf{r})$ can be formulated by $\sum_i U_i^{\text{att}}(\mathbf{r})$ in which $U_i^{\text{att}}(\mathbf{r})$ represents the attractive part of Lennard-Jones potential [16, 22, 23]. To this end, the L-J potential can be divided into attractive U_i^{att} and repulsive U_i^{rep} in different ways. It can be a "6-12" decomposition:

$$U_i^{\text{att},6/12}(\mathbf{r}) = -2\epsilon_{is}\left(\frac{\sigma_{is}}{\|\mathbf{r}-\mathbf{r}_i\|}\right)^6$$

$$U_i^{\text{rep},6/12}(\mathbf{r}) = \epsilon_{is}\left(\frac{\sigma_{is}}{\|\mathbf{r}-\mathbf{r}_i\|}\right)^{12},$$

or a Weeks-Chandler-Andersen (WCA) decomposition based on the original WCA theory [19]:

$$U_i^{\text{att},\text{WCA}}(\mathbf{r}) = \begin{cases} -\epsilon_{is}(\mathbf{r}) & 0 < \|\mathbf{r}-\mathbf{r}_i\| < \sigma_{is} \\ U_i^{\text{LJ}}(\mathbf{r}) & \|\mathbf{r}-\mathbf{r}_i\| \geq \sigma_{is}, \end{cases}$$

$$U_i^{\text{rep},\text{WCA}}(\mathbf{r}) = \begin{cases} U_i^{\text{LJ}}(\mathbf{r}) + \epsilon_{is}(\mathbf{r}) & 0 < \|\mathbf{r}-\mathbf{r}_i\| < \sigma_{is} \\ 0 & \|\mathbf{r}-\mathbf{r}_i\| \geq \sigma_{is}. \end{cases}$$

2.1 Minimizer existence

In this section , we will prove the existence of a global minimizer of nonpolar solvation free energy functional (4). Related mathematical notations and assumptions will be given first. Then a formal minimizer statement and its proofs will be described in detail.

A function $u \in L^1(\Omega)$ is of bounded variation in Ω if [29, 30]

$$\int_{\Omega} |\nabla u| dx := \sup\{\int_{\Omega} u\nabla \cdot g dx, g \in C_c^1(\Omega, \mathbb{R}^3), |g| \leq 1 \quad \text{in} \quad \Omega\} < \infty$$

where $C_c^1(\Omega, \mathbb{R}^3)$ denotes the space of all C^1 mappings from Ω to $\mathbb{R}^3$ that are compactly supported inside Ω. The space $BV(\Omega)$ of all $L^1(\Omega)$ functions that have bounded variation in Ω is a Banach space equipped with the norm:

$$\|u\|_{BV(\Omega)} := \|u\|_{L^1(\Omega)} + \int_\Omega |\nabla u| dx, \quad \forall u \in BV(\Omega).$$

For any $M \subseteq R^3$, χ_M is used to represent the characteristic function of M: $\chi_M(x) = 1$ if $x \in M$ and $\chi_M(x) = 0$ if $x \notin M$. Moreover, for any Lebesgue-measurable subset $M \subseteq R^3$, the perimeter of M in Ω is defined by

$$P_\Omega(M) := \int_\Omega |\nabla \chi_M| dx$$

Setting $A = \{M \in \Omega, \text{M is Lebesgue-measurable}\}$ and $|M| = \int_\Omega \chi_M dx$, we define the sharp interface free energy functional $G_{np}^0 : A \to R \cup \{+\infty\}$ as

$$G_{np}^0[M] = p|M| + \gamma P_\Omega(M) + \rho_s \int_{\Omega \setminus M} U^{vdW} dx \tag{5}$$

If $M \subset A$ is an open subset with a smooth boundary Γ and its closure $\bar{M} \subset A$, then $G_{np}^0[M] = G_{np}^0[\Gamma]$ where $G_{np}[\Gamma]$ is defined in (4) with $\Omega_s = \Omega \setminus M$. In addition, we assume that

- p, γ, ρ_s are positive constants
- $\Omega = \Omega_s \bigcup \Omega_m$, Γ separates Ω_s and Ω_m, Ω is an open connected and bounded set with a C^2 boundary $\partial\Omega$
- The solvent-solute interaction function $U^{vdW} = \sum_i U_i^{LJ} : \mathbb{R}^3 \to R \bigcup \{+\infty\}$ satisfies $U^{vdW}(\mathbf{r}_i) = +\infty$ where $i = 1, 2 \cdots, N$ and $\mathbf{r}_i$ denotes the space position of the i-th atom. $\lim_{r \to \infty} U^{vdW}(\mathbf{r}) = 0$. When restricted onto $\mathbb{R}^3 \setminus \{\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N\}$, U^{vdW} is a C^1 function with $U_{min}^{vdW} = \inf\{U^{vdW}(\mathbf{r}) : \mathbf{r} \in \mathbb{R}^3 \setminus \{\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N\}\}$

Due to the lower boundedness of U^{vdW} and the fact that M is Lebesgue-measurable, M is of finite perimeter if $G_{np}^0[M]$ is of finite value. Based on the above assumptions, we attempt to prove the existence of a global minimizer of the nonpolar solvation free energy functional $G_{np}^0[M]$ stated as follows:

Theorem 2.1
There exists $M^\star \in A$ such that

$$G_{np}^0[M^\star] = \min_{M \in A} G_{np}^0[M] < \infty \tag{6}$$

The first proof of theorem 2.1:
Since $|M|, P_\Omega(M), and \int_{\Omega \setminus M} U^{vdW} dx$ are bounded below, let $\alpha = inf\{G_{np}^0[M], M \in A\}$ and be finite. Now there exists a sequence $M_k \in A (k = 1, 2, \cdots)$ such that

$$\lim_{k \to \infty} G_{np}^0[M_k] = \alpha$$

and that $G_{np}^0[M_k]$ is finite for each $k \geq 1$. The lower boundedness of U^{vdW} and boundedness of $|M_k|$ imply that $\{P_\Omega(M_k)\}_{k=1}^\infty$ is bounded and then $\int_{\Omega \setminus M_k} U^{vdW} dx$ is bounded. By the boundedness of $\{P_\Omega(M_k)\}_{k=1}^\infty$, compact embedding $BV(\Omega) \hookrightarrow L^1(\Omega)$, and Helly's selection theorem, there exists a subsequence of $\{\chi_{M_k}\}$, not relabeled, such that $\chi_{M_k} \to \chi_{M^\star}$ in $L^1(\Omega)$ for some Lebesgue-measurable set $M^\star \subset \Omega$ and

$$P_\Omega(M^\star) \leq \liminf_{k \to \infty} P_\Omega(M_k) \tag{7}$$

Then

$$\lim_{k \to \infty} |M_k| = \lim_{k \to \infty} \int_\Omega \chi_{M_k} dx = \int_\Omega \chi_{M^\star} dx = |M^\star| \tag{8}$$

If needed, we may take a further subsequence of $\{\chi_{M_k}\}_{k=1}^{\infty}$ such that $\chi_{M_k} \to \chi_{M^*}$ a.e. in Ω. Then by Fatou's Lemma, we have

$$
\begin{aligned}
\liminf_{k\to\infty} \int_{\Omega\setminus M_k} U^{vdW} dx &= \liminf_{k\to\infty} \{ \int_{\Omega\setminus M_k} (U^{vdW} - U^{vdW}_{min}) dx + \int_{\Omega\setminus M_k} U^{vdW}_{min} dx \} \\
&= \liminf_{k\to\infty} \int_{\Omega\setminus M_k} (U^{vdW} - U^{vdW}_{min}) dx + \lim_{k\to\infty} \int_{\Omega\setminus M_k} U^{vdW}_{min} dx \\
&\geq \int_{\Omega\setminus M^*} (U^{vdW} - U^{vdW}_{min}) dx + \int_{\Omega\setminus M^*} U^{vdW}_{min} dx \\
&= \int_{\Omega\setminus M^*} U^{vdW} dx
\end{aligned}
\tag{9}
$$

Now the combination of (7),(8),(9) leads to

$$
\alpha \leq G^0_{np}[M^*] \leq \liminf_{k\to\infty \atop M_k \in A} G^0_{np}[M_k] = \alpha
$$

Therefore, $G^0_{np}[M^*] = \alpha$ #

The second proof via a phase-field model

The second proof can be done by using a phase-field model [27]. For the purpose, a phase-field model of non-polar solvation free energy needs to be constructed first. Then one proves that the phase-field representation of energy functional Γ-converges to the sharp interface counterpart (5). Finally, the existence of a global minimizer of sharp interface based nonpolar solvation free energy functional (4) is followed by the existence of a global minimizer of the corresponding phase-field based functional. Since the functional (4) is the nonpolar part of the energy functional (1.6) in [27], the second proof can be considered as a part of the results in [27]. Readers may refer to [27] for proof details. However, for the reader's convenience, the main results and related key lemmas are described in this paper.

In general, phase-field models are useful tools for the numerical simulation of interfacial problem to avoid the explicit treatment of boundary condition at a sharp interface. They are constructed in order to reproduce the given sharp interface based dynamics. In the limit of an infinitesimal interface width, the correct interfacial dynamics can be recovered. According to the sharp-interface based functional (5), a phase-field model is described as follows [27]. First of all, we define a continous phase-field function $s : \Omega \to R$ which has values close to 0 and 1. The solute and solvent regions are approximated by 1 and 0, respectively. While, a thin transition layer, having value between 0 and 1, is used to describe the diffuse solute-solvent interface and then solvation system. With interface width parameter $\xi \in (0, 1)$, we consider the phase-field functional $G^{\xi}_{np} : L^1 \to R \bigcup \{+\infty\}$:

$$
G^{\xi}_{np}[s] = \begin{cases} p \int_\Omega s^2 dx + \gamma \int_\Omega [\frac{\xi}{2}|\nabla s|^2 + \frac{1}{\xi} W(s)] dx + \rho_s \int_\Omega (s-1)^2 U^{vdW} dx, & s \in H^1(\Omega) \\ \infty, & \text{otherwise} \end{cases}
\tag{10}
$$

where $W(s) = 18s^2(1-s)^2$ is called van der Waals-cahn-Hilliard functional. The second term in Eq. (10) Γ-converges to the area of solute-solvent interface as $\xi \to 0$. Notation $\xi^k \searrow 0$ is used in this paper to indicate that ξ_k is a decreasing sequence such that $\xi_1 > \xi_2 > \cdots$ and $\xi_k \to 0$ as $k \to \infty$.

Two Lemmas are used to prove theorem 2.1.

Lemma 2.1 (refer to Lemma 4.1 in [27])

For any $\xi \in (0, \xi_0]$ and $\xi_0 < 1$, there exists a $s_\xi \in H^1(\Omega)$ such that

$$
G^{\xi}_{np}[s_\xi] = \min_{s\in H^1(\Omega)} G^{\xi}_{np}[s] = \min_{s\in L^1(\Omega)} G^{\xi}_{np}[s] < \infty
$$

To prove the Γ-Convergence of phase-field based energy functional to the sharp interface based counterpart, we rewrite Eq.(5) in the following: $G^0_{np} : L^1 \to R \bigcup \{+\infty\}$ and

$$
G^0_{np}[s] = \begin{cases} p|M| + \gamma P_\Omega(M) + \rho_s \int_{\Omega\setminus M} U^{vdW} dx, & \text{if } s = \chi_M \in BV(\Omega) \\ \infty, & \text{otherwise} \end{cases}
\tag{11}
$$

It is easy to see that Eq. (11) is the same as Eq. (5). Now the Γ-convergence of free energy functional is stated as Lemma 2.2.

Lemma 2.2 : (refer to Theorem 2.1 in [27])

For any sequence $\xi_k \searrow 0$, the sequence of functional $G_{np}^{\xi_k} : L^1 \to R \bigcup \{+\infty\}$ Γ-converges to the sharp interface functional $G_{np}^0 : L^1 \to R \bigcup \{+\infty\}$ with respect to $L^1(\Omega)$ convergence. In other words, the following holds: For any $s \in L^1(\Omega)$ and any sequence $\xi_k \searrow 0$, there exists a sequence $\{s_k\}_{k=1}^{\infty} \in L^1(\Omega)$ such that $s_k \to s$ in $L^1(\Omega)$ and

$$\lim_{k \to \infty} G_{np}^{\xi_k}[s_k] = G_{np}^0[s] \tag{12}$$

Finally, theorm 2.1 can be proved by using Lemma 2.1 and Lemma 2.2.

Proof two of theorem 2.1

Let $\xi_k \searrow 0$. By Lemma 2.1, for each ξ_k, there exists $s_k^\star$ such that $G_{np}^{\xi_k}[s_k^\star] = \min_{s \in H^1(\Omega)} G_{np}^{\xi_k}[s]$. Meanwhile, $G_{np}^{\xi_k}[s_k^\star]$ is bounded compared to $G_{np}^{\xi_k}[1]$ for each k. Also by the inequality (4.1) in [27], the corresponding sequence of $\{s_k^\star\}_{k=1}^{\infty}$ is bounded in $H^1(\Omega)$. This, together with a result from [29, 31] and Lemma 2.2, imply that there exists a subsequence of $\{s_k^\star\}_{k=1}^{\infty}$, not relabeled, such that $s_k^\star \to \chi_M$ in $L^1(\Omega)$ for some measurable subset M of Ω [27] and that $\lim_{k \to \infty} G_{np}^{\xi_k}[s_k^\star] = G_{np}^0[\chi_M]$. If needed, we may take further subsequence of $\{[s_k^\star]\}_{k=1}^{\infty}$. For any other measurable subset M', again by Lemma 2.2, there exists a sequence $\lim_{k \to \infty} G_{np}^{\xi_k}[s_k] = G_{np}^0[\chi_{M'}]$. It is clear that

$$G_{np}^0[\chi_M] = \lim_{k \to \infty} G_{np}^{\xi_k}[s_k^\star] \leq \lim_{k \to \infty} G_{np}^{\xi_k}[s_k] = G_{np}^0[\chi_{M'}]$$

Thus χ_M minimizes G_{np}^0 #

2.2 Solution and mathematical implementation for minimizer

To obtain the optimal solvent-solute surface, we take the first variation on nonpolar energy functional with respect to surface definition Γ to find a geometry flow equation. A brief description of variational analysis is given here, readers may refer to paper [13] for technical details. First, we consider a surface element $\mathbf{f}(u_1, u_2)$ and its infinitesimal displacement in the normal direction

$$\mathbf{f}^{(\varepsilon)}(u_1, u_2) := \mathbf{f}(u_1, u_2) + \varepsilon \varphi(u_1, u_2) \cdot \mathbf{N}(u_1, u_2) \tag{13}$$

where $\mathbf{N}$ is the outward unit normal direction and φ is an arbitrary C^2 function and u_1, u_2 are two parameters for a 2D surface. The surface variation of the area leads to

$$\frac{\delta A}{\delta \Gamma} = \frac{\delta}{\delta \Gamma} = -\int_\Gamma \varphi 2H d\sigma, \tag{14}$$

where H is the mean curvature. Second, the volume variation with respect to Γ by means of the variation with respect to ε yields

$$\frac{\delta \text{Vol}}{\delta \Gamma} = \left. \frac{\partial(\text{Vol}(f^{(\varepsilon)}) - \text{Vol}(f))}{\partial \varepsilon} \right|_{\varepsilon=0} = \int_\Gamma \varphi d\sigma,$$

Third, the first variation of the volume integration of solute-solvent vdw interaction can be attained as follows

$$\frac{\delta \left(\int_{\Omega_s} \rho_s U^{\text{vdW}} d\mathbf{r} \right)}{\delta \Gamma} = \frac{\delta(\int_\Omega \rho_s U^{\text{vdW}} d\mathbf{r} - \int_{\Omega_m} \rho_s U^{\text{vdW}} d\mathbf{r})}{\delta \Gamma} = -\int_\Gamma \rho_s U^{\text{vdW}} \varphi d\sigma. \tag{15}$$

Eventually, we have

$$\frac{\delta G_{\text{np}}^0}{\delta \Gamma} = \int_\Gamma (-2\gamma H + p - \rho_s U^{\text{vdW}}) \varphi d\sigma = 0. \tag{16}$$

Since φ is an arbitrary function, the following condition must be satisfied for each point on the optimized interface

$$W_n \quad \equiv \quad -2\gamma H + p - \rho_s U^{\text{vdW}} = 0 \tag{17}$$

During the minimization process, the directional derivative of solvation free energy functional G_{np}^0 in the direction of a normal variation φ can be expressed as

$$D_\varphi G_{np}^0(\mathbf{f}) = \left. \frac{\partial G_{np}^0(\mathbf{f}^{(\varepsilon)})}{\partial \varepsilon} \right|_{\varepsilon=0} = \int_\Gamma W_n \varphi d\sigma. \tag{18}$$

If we choose $\varphi(u_1, u_2) = -W_n$, then

$$D_\varphi G_{np}^0(\mathbf{f}) = - \int_\Gamma W_n^2 d\sigma \leq 0.$$

This means that the total free energy decreases along the normal direction when $\varphi(u_1, u_2) = -W_n$ until it reaches a minimum. Therefore the evolution $\mathbf{f}^{(\varepsilon)} = \mathbf{f} - \varepsilon W_n \mathbf{N}$ leads to a steady state and associated solvent-solute interface with strictly smaller energy. This analysis motivates the following potential driven geometric flow equation for the optimal solute-solvent interface

$$\frac{\partial \mathbf{X}}{\partial t} = -W_n \mathbf{N}, \tag{19}$$

where $\mathbf{X} \in \Gamma \subset \mathbb{R}^3$ is a position vector on the evolving manifold Γ.

To avoid numerical difficulties in handling topological changes during the biomolecule surface evolution, we embed a Lagrangian operator into its Eulerian representation. To this end, we introduce an arbitrary hypersurface function $S(\mathbf{r})$ with $\mathbf{r} \in \mathbb{R}^3$. It is easy to verify that the unit normed vector can be expressed in term of S

$$\mathbf{N} = \frac{\nabla S}{\|\nabla S\|}. \tag{20}$$

Then the desired surface can be represented as a set of points with a constant value of function S

$$\Gamma = \{\mathbf{r} | S(\mathbf{r}) = L\} \tag{21}$$

where L is an isosurface value. By the chain rule

$$\begin{aligned} \frac{\partial S}{\partial t} &= \frac{\partial S}{\partial \mathbf{X}} \cdot \frac{\partial \mathbf{X}}{\partial t} \\ &= -W_n \nabla S \cdot \mathbf{N} \end{aligned} \tag{22}$$

where $\mathbf{X}$ is a 3D position vector confined to the manifold Γ. Moreover, the surface mean curvature H can be rewritten in terms of S as

$$H = \frac{1}{2} \nabla \cdot \left(\frac{\nabla S}{\|\nabla S\|} \right). \tag{23}$$

According to Eq. (20) and (23), one has

$$\begin{aligned} \frac{\partial S}{\partial t} &= -\|\nabla S\| W_n \\ &= \|\nabla S\| \left[\gamma \nabla \cdot \left(\frac{\nabla S}{\|\nabla S\|} \right) - p + \rho_s U^{\text{vdW}} \right] \end{aligned} \tag{24}$$

where all terms should be expressed in terms of the level surface function S.

3 Multiscale modeling of nonpolar solvation: Eulerian formulation

From a microscopic point of view, there is no sharp interface between solvent and solute. In principle, an isolated molecule can be analyzed by the first principle — a quantum mechanical description of the wavefunction or density distribution of all the electrons and nuclei. However, such description is computationally intractable for large biomolecules. Under physiological condition, biomolecules are in a non-isolated environment, and are interacting with solvent molecules and/or other biomolecules. Therefore, their wavefunctions overlap spatially, so do their electron density distributions. The Eulerian formulation becomes important in the sense that it is able to produce an overlapping solvent-solute boundary, which may be able to describe the true physical boundary between the solvent and solute when its generation is governed by the variational analysis — the total free energy optimization [22]. Moreover, an Eulerian formulation is useful because it is able to handle topological changes directly during the surface evolution. Finally, an Eulerian formulation avoids the complex interface problem that arises from a Lagrangian formulation for the electrostatic solution in an implicit solvent model.

Previously, an Eulerian formulation of multiscale solvation model has been proposed and used in polar solvation calculation [22]. Here, the solvation of nonpolar molecules is considered where the electrostatic interaction is negligible. This simplified case minimizes modeling uncertainties and confirms the reliability of the Eulerian formulation of the differential geometry based solvation model for the description of solvation free energy. The details of modeling, numerical schemes and parameterization are described briefly below.

3.1 Model and method

Let us consider a multi-domain setting of a biomolecule and solvent system. The biomolecule is described in discrete atomic detail, while the solvent is treated as a continuum. Therefore, the domain $\Omega \in \mathbb{R}^3$ is essentially divided into two (types of) regions, i.e., solvent domain Ω_s and biomolecular domain Ω_m. As such, one has $\Omega = \Omega_s \bigcup \Omega_m$. Because electron wavefunctions of solvent and solute overlap at the atomic scale, Ω_s and Ω_m should overlap with each other at the boundary of molecules and solvent, i.e., $\Omega_b = \Omega_s \bigcap \Omega_m \neq \oslash$, where Ω_b is the region of solvent-solute boundary. Mathematically, we propose a characteristic function $S : \mathbb{R}^3 \to \mathbb{R}$ to characterize this overlapping solvent-solute boundary. In particular, $S(x)$ is a characteristic function of the solute domain, i.e., it is $S = 1$ inside the biomolecule and $S = 0$ in the solvent. Moreover, S takes a value between 0 and 1 at the solvent-solute boundary region.

To obtain a functional relation for S, it is necessary to rewrite nonpolar free energy formulation (1) in terms of $S(\mathbf{r})$. The enclosed volume of biomolecule can be given by

$$\text{Vol} = \int_{\Omega_m} d\mathbf{r} = \int_{\Omega} S(\mathbf{r})d\mathbf{r}. \tag{25}$$

Similarly the attractive dispersion term can be rewritten in the form

$$\rho_s \int_{\Omega_s} U^{\text{vdW}} d\mathbf{r} = \rho_s \int_{\Omega} (1 - S(\mathbf{r})) U^{\text{vdW}} d\mathbf{r}, \tag{26}$$

where we assume that the solvent bulk density ρ_s is a constant in space. For the surface area term in Equation (1), we make use of the co-area formula in the geometric measure theory [32]. Specifically, we introduce a concept of mean surface area of a family of isosurfaces which are subsets satisfying $\{S(\mathbf{r}) = y|, \ 0 \leq y \leq 1\}$. Therefore the mean surface area can be given by a volume integral as

$$\begin{aligned} \text{Area} \quad &= \int_0^1 \int_{S^{-1}(c) \bigcap \Omega} d\sigma dc \\ &= \int_{\Omega} \|\nabla S(\mathbf{r})\| d\mathbf{r}. \end{aligned} \tag{27}$$

Note that $\nabla S \neq 0$ only in the region of the solvent-solute boundary. Numerical test of this formulation was validated in our earlier work [22].

Finally, nonpolar free energy functional of solvation for biomolecules at equilibrium is given by

$$G_{np} \quad = \quad \int_{\Omega} \gamma \|\nabla S(\mathbf{r})\| + pS(\mathbf{r}) + \rho_s(1 - S(\mathbf{r}))U^{vdW} d\mathbf{r}. \tag{28}$$

Now, the nonpolar solvation free energy functional is a functional in terms of the characteristic function S. To obtain the optimal function S, we take the first variation on G_{np} with respect to S to yield

$$\frac{\delta G_{np}}{\delta S} \quad \Rightarrow \quad -\nabla \cdot \left(\gamma \frac{\nabla S}{\|\nabla S\|} \right) + p - \rho_s U^{vdW} = 0, \tag{29}$$

where $\nabla \cdot \left(\gamma \frac{\nabla S}{\|\nabla S\|} \right)$ is a generalized Laplace-Beltrami operator. γ is treated as a constant in our present computation. In general it can be a function of the position $\gamma = \gamma(\mathbf{r})$ to reflect surface hydrophobicity at different locations. The solution of Eq. (29) leads to a "physical solvent-solute boundary" S. As discussed in earlier work [22, 33], the solution of this elliptic partial differential equation can be attained via a parabolic partial differential equation

$$\frac{\partial S}{\partial t} = \sqrt{\|\nabla S\|} \left[\nabla \cdot \left(\gamma \frac{\nabla S}{\|\nabla S\|} \right) + V \right], \tag{30}$$

where the generalized "potential" V is defined as

$$V = -p + \rho_s U^{vdW} \tag{31}$$

Note that in Eq. (30), as $t \to \infty$, the initial profile of S evolves into a steady state solution, which solves the original Eq. (29). In this section, $U^{vdW} = \sum_i U_i^{att,WCA}(\vec{r})$.

The discretization scheme used here for the solution of the generalized geometry flow equation (30) is similar to what we designed previously [22, 33]. It can be rewritten in the form

$$\frac{\partial S}{\partial t} = \gamma \frac{(S_x^2 + S_y^2)S_{zz} + (S_x^2 + S_z^2)S_{yy} + (S_y^2 + S_z^2)S_{xx}}{S_x^2 + S_y^2 + S_z^2} \tag{32}$$

$$-\gamma \frac{2S_x S_y S_{xy} + 2S_x S_z S_{xz} + 2S_z S_y S_{yz}}{S_x^2 + S_y^2 + S_z^2} + \sqrt{S_x^2 + S_y^2 + S_z^2} V,$$

For the initial value of S, we consider

$$S(x, y, z, 0) = \begin{cases} 1, & (x, y, z) \in D \\ 0, & \text{otherwise} \end{cases} \tag{33}$$

where we define the domain enclosed by the solvent accessible surface to be $D = \bigcup_{i=1}^{N_a} \{\mathbf{r} : |\mathbf{r} - \mathbf{r_i}| < \mathbf{r}_i + \mathbf{r}_p\}$, with $\mathbf{r}_p$ being the probe radius. Then $S \in H^1$ can be satisfied via a numerical smoothing process. Here N_a denotes the total number of atoms for a given biomolecular system. Let atom centers be $\mathbf{r_i} = (x_i, y_i, z_i), i = 1, \cdots, N_a$, and $\mathbf{r}_i$ represents the radius of the ith atom. To make the computation more efficient, we assume that the domain enclosed by van der Waals surface is always pure solute $(S = 1)$, and we only update the values of $S(x, y, z, t)$ at the points in between the van der Waals surface and the solvent accessible surface; i.e., $(x, y, z) \in \bigcup_{i=1}^{N_a} \{\mathbf{r} : \mathbf{r}_i < |\mathbf{r} - \mathbf{r_i}| < (\mathbf{r}_i + \mathbf{r}_p)\}$. Numerically, to avoid possible zeros in the denominator of Eq. (32) we add a very small number, such as 10^{-7}, to the denominator. That turned out not to affect the result at all [22].

Regarding the parameterization in our nonpolar model, the nonpolar solvation model involves parameters, such as, surface tension γ, hydrodynamic pressure p, L-J well-depth parameters ϵ_{is}, solvent density ρ_s, solvent radius σ_s and solute radii σ_i. The parameter optimization is measured by root mean square (RMS) error between the calculated and solvation experimental data. In this work, the solvent density ρ_s is fixed at 0.0334 1/Å^3 [1, 16]. Moreover, γ, p and ϵ_{is} are considered as fitting parameters. Therefore, an iterative procedure is designed in the following process:

(a) Choose a trial set of molecules with given atomic coordinates, radii, and experimental data of solvation free energies. Then take an initial set of parameters γ, p and ϵ_{is} for the trial set.

(b) For the jth molecule, solve Eq. (30) to compute surface area, molecular volume and solvation free energy G_{np}^j with current parameter values.

(c) Set up a target function

$$T = \min_{(p,\gamma,\epsilon_{is})} \sum_j \left(G_{np}^j - G_{np}^{j,\exp} \right)^2,$$

where $G_{np}^{j,\exp}$ are experimental data of solvation free energies.

(d) All parameters p, γ and ϵ_{is} are updated by resolving a nonnegative least squares problem to determine non-negative parameters.

(e) The iterative procedure (b)-(d) continues until convergence reaches within a pre-set tolerance for the above fitting parameters.

3.2 Results and discussion

First of all, a set of 11 alkanes, including linear, branched, and cyclic apolar compounds, is taken as a calibration set to validate our model and its numerical implementation. Experimental data is available for this set of hydrophobic solute [16, 17, 34]. Note that recently Wang et al [35] used an Eulerian formulation of DG based full solvation model to calculate solvation free energy for both polar and nonpolar molecules in a unified formulation. However, in the present work, nonpolar solvation free energy is calculated only by nonpolar free energy containing terms related to surface area, volume and Lennard-Jones (LJ) solvent-solute interactions.

With the set of 11 alkane compounds, optimized parameters for alkanes are obtained. In particular, surface tension $\gamma = 0.0715$ kcal/(mol Å^2), solvent pressure $p = 0.0154$ kcal/(mol Å^3), LJ parameters $\epsilon_{cs} = 0.496$ kcal/mol and $\epsilon_{hs} = 0.00$ kcal/mol. Here we fixed the solvent radius σ_s =0.65 Å and carbon atom radius σ_c =1.87 Å. Interestingly, the pressure p=0.0154 kcal/(mol Å^3) is comparable to experimental measurement which is 0.0248 kcal/(mol Å^3). Surface tension $\gamma = 0.0715$ kcal/(mol Å^2) is also comparable to its experimental measurement which is 0.103 kcal/(mol Å^2). With the optimized parameters, results of 11 alkane compounds calculated from the present model are shown in the top section of Table 3. It is evident that our model accurately catches subtle differences between linear, branched, and cyclic apolar compounds. Meanwhile, it perfectly reproduces the total solvation free energies of 11 alkanes. The root mean square (RMS) error is found to be as small as 0.11 kcal/mol. It is worthwhile to point out that current nonpolar Eulerian formulation does not need to apply artificially enlarged Van der Waals radii in solvation analysis as required by our earlier Eulerian representation [22]

We also examined the importance of dispersion interactions U^{vdW} in the solvation analysis by carrying out simulations with and without the dispersion term. Computed solvation results and their comparisons are listed in Table 1 for the set of 11 alkanes. It turns out that the dispersion interaction is necessary for our DG based solvation model and substantially improves the performance of current nonpolar solvation model [16, 19]. Let us point out that the optimized pressure parameter is 0 without the dispersion term. This directly supports the surface-area-only type of nonpolar solvation models when the dispersion attraction is not taken into consideration.

When current results of 11 alkanes are compared with those from the previous Lagrangian formulation of DG based nonpolar model [1, 36], it is found that both of them reproduce the total solvation free energies very well for each molecule. The RMS errors are 0.11 kcal/mol and 0.12 kcal/mol, respectively. The comparison of total solvation free energy is displayed in Figure 1. This again indicates that the framework of our differential geometry based solvation model can be powerful for solvation prediction in both Lagrangian and Eulerian formulations [13]. However, they are different when one looks into each contributed term. Solvation decomposition results listed in Table 2 show that the repulsive and attractive parts of solvation free energies differ

Table 1: Calculated solvation results for the set of 11 alkanes with and without the van der Waals dispersion term. The error is computed by comparing to experimental data [34].

Compound	With U^{vdW} (kcal/mol)		Experiment [34]	Without U^{vdW} (kcal/mol)	
	Calculated	Error		Calculated	Error
methane	1.99	−0.01	2.00	0.94	−1.06
ethane	1.94	0.11	1.83	1.31	−0.52
butane	2.13	0.05	2.08	1.98	−0.10
propane	2.03	0.07	1.96	1.64	−0.32
pentane	2.32	−0.01	2.33	2.32	0.01
hexane	2.42	−0.07	2.49	2.67	0.18
isobutane	2.24	−0.28	2.52	1.95	−0.57
2-methylbutane	2.40	0.02	2.38	2.23	−0.15
neopentane	2.60	0.10	2.50	2.22	−0.28
cyclopentane	1.15	−0.05	1.20	2.02	0.82
cyclohexane	1.34	0.11	1.23	2.27	1.04
RMS error		0.11			0.58

from each other. In particular, the magnitudes of repulsive and attractive energies are smaller in the Eulerian formulation than those in the Lagrangian formulation, although the summation of them are almost the same.

Table 2: Decomposition analysis for the set of 11 alkanes. Comparisons between current Eulerian formulation and previous Lagrangian formulation (Sharp) [36] with respect to repulsive and attractive parts are given. Again the error of total solvation energy is calculated according to the experimental data [34].

Compound	Rep. part (kcal/mol)		Att. part (kcal/mol)		Total (kcal/mol)		Error (kcal/mol)	
	Sharp	Eulerian	Sharp	Eulerian	Sharp	Eulerian	Sharp	Eulerian
methane	5.41	4.25	−3.43	−2.26	1.98	1.99	−0.02	−0.01
ethane	7.44	6.00	−4.75	−4.05	1.93	1.94	0.10	0.11
butane	11.16	9.21	−9.06	−7.08	2.10	2.13	0.02	0.05
propane	9.27	7.60	−7.30	−5.56	1.98	2.03	0.02	0.07
pentane	12.05	10.85	−10.71	−8.53	2.34	2.32	0.01	−0.01
hexane	14.92	12.46	−12.50	−10.05	2.42	2.42	−0.07	−0.08
isobutane	11.02	9.11	−8.73	−6.87	2.30	2.24	−0.22	−0.28
2-methylbutane	12.58	10.49	−10.17	−8.08	2.41	2.40	0.03	0.02
neopentane	12.55	10.50	−9.94	−7.91	2.61	2.59	0.11	0.09
cyclopentane	11.38	9.44	−10.21	−8.30	1.18	1.15	−0.03	−0.05
cyclohexane	12.84	10.76	−11.54	−9.42	1.29	1.34	0.06	0.11

Moreover, we conducted a predictive study for other 19 alkane compounds. the optimized parameters obtained from the above training set of 11 alkane molecules are utilized. Results are shown in the bottom section of Table 3 , in which the repulsive and attractive decomposition is also demonstrated. Our predictive values fit the experimental data very well and their comparison is displayed in Figure 2. The RMS error is 0.34 kcal/mol.

Having demonstrated the accuracy and reliability of our approach for the alkane molecules, we further carry out a prediction of a set of 11 alkene compounds which have been studied by Ratkova *et al.* [37]. By
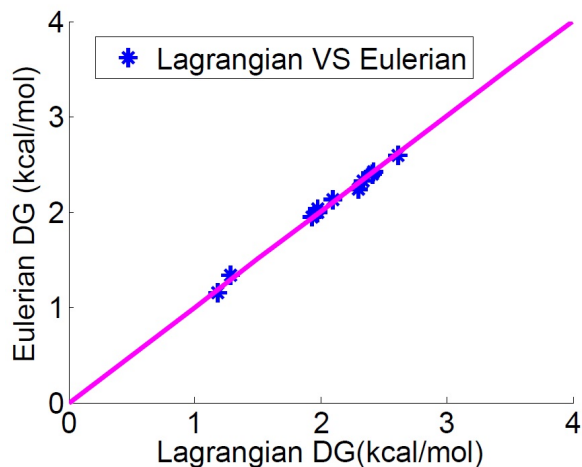
Figure 1: Comparison of model predictions for solvation free energies of eleven alkanes between Lagrangian and Eulerian formulations of differential geometry based nonpolar solvation model.
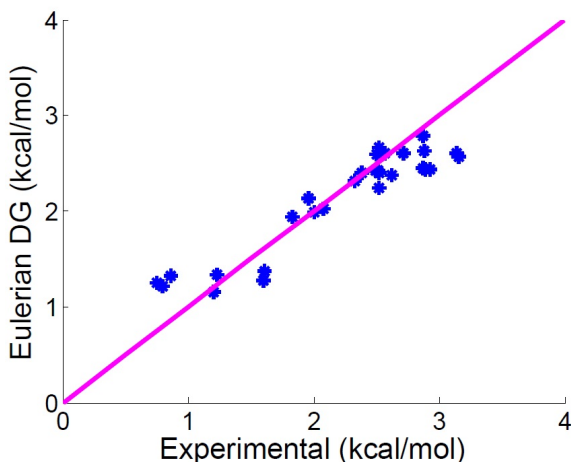

Figure 2: Comparison of model predictions and experimental data of solvation free energies of thirty alkanes.

assuming the same solvent behavior, we use the same set of optimized parameters obtained from the alkane training set. Solvation free energies of 11 alkene compounds are shown in Table 4. They match the experimental data very well and their comparison is shown in Figure 3. The RMS errors is 0.24 kcal/mol. The result is significantly better than that reported in [37] using integral equation techniques, which is about 0.462 kcal/mol, and is close to that obtained by our Lagrangian approach which is 0.18 kcal/mol [1]. Note that in the previous Lagrangian calculation, the surface tension parameter was optimized particularly for the chosen alkene set while we did not fit any parameter here.

4 Conclusion

In this paper, the existence of a global minimizer of Lagrangian formulation based nonpolar solvation energy functional, which was proposed and implemented previously [1], has been proved. It provides a mathematical

Table 3: Numerical and experimental total solvation free energies for 30 alkane compounds. $\gamma = 0.0715$ kcal/(mol Å^2), $p = 0.0154$ kcal/(mol Å^3) and $\epsilon_{cs} = 0.496$ kcal/mol and $\epsilon_{hs} = 0.00$ kcal/mol

Compound	Rep. part	Att. part	Numerical	Experimental [34]
			(kcal/mol)	
methane	4.25	−2.26	1.99	2.00
ethane	5.99	−4.05	1.94	1.83
propane	9.21	−7.08	2.13	1.96
butane	7.59	−5.56	2.03	2.08
pentane	10.85	−8.53	2.32	2.33
hexane	12.46	−10.05	2.42	2.49
isobutane	9.11	−6.87	2.24	2.52
2-methylbutane	10.49	−8.08	2.40	2.38
neopentane	10.51	−7.91	2.59	2.50
cyclopentane	9.44	−8.30	1.15	1.20
cyclohexane	10.76	−9.42	1.34	1.23
RMS of calibration set			0.11	
octane	15.65	−13.21	2.44	2.89
heptane	14.04	−11.66	2.37	2.62
nonane	17.28	−14.67	2.61	3.14
decane	18.90	−16.33	2.57	3.16
2-methylpentane	12.21	−9.81	2.40	2.52
3-methylpropane	12.10	−9.68	2.42	2.51
3-methylhexane	13.62	−11.01	2.61	2.71
2-methylhexane	13.67	−11.24	2.43	2.93
3-methylpentane	12.10	−9.68	2.42	2.51
224-trimethylpentane	14.47	−11.68	2.78	2.87
22-dimethylbutane	11.88	−9.28	2.60	2.57
24-dimethylpentane	13.14	−10.70	2.45	2.87
22-dimethylpentane	13.47	−10.84	2.63	2.88
23-dimethylpentane	13.34	−10.67	2.67	2.52
cyclopropane	6.90	−5.64	1.25	0.75
cycloheptane	11.96	−10.76	1.21	0.80
cyclooctane	13.17	−11.85	1.32	0.86
methylcyclopentane	10.97	−9.70	1.27	1.60
methylcyclohexane	12.22	−10.86	1.37	1.61
RMS of prediction set			0.34	

foundation for our model justification and numerical implementation. Two proofs are given here. One is a direct method based on the sharp interface model (or Lagrangian formulation). The other one is via a phase-field model. The latter gives an insight of the relationship between the sharp interface model and its phase-field counterpart, namely, the phase-field model Γ-converges to the sharp interface model with respect to L^1 convergence.

Moreover, an Eulerian formulation of our differential geometry based nonpolar solvation model is proposed and implemented using a very similar parameterization strategy to the corresponding Lagrangian approach in [1]. It turns out that our Eulerian formulation works perfectly in blindly predicting total solvation free energies of nonpolar molecules without enlarged van der Waals radii. In addition, the calculated results of total free energy are almost the same as those predicted by our Lagrangian formulation, while the decom-

Table 4: Numerical and experimental total solvation free energies for 11 alkenes. The numerical energy is the sum of the repulsive part and attractive part. The free parameters are chosen as $\gamma = 0.0715$ kcal/(mol Å^2), $p = 0.0154$ kcal/(mol Å^3) and $\epsilon_{cs} = 0.496$ kcal/mol and $\epsilon_{hs} = 0.00$ kcal/mol

| Compound	Rep. part	Att. part	Numerical (kcal/mol)	Experimental [34, 37]
3-methyl-1- butene	10.43	−8.53	1.90	1.82
1-butene	8.88	−7.21	1.67	1.38
ethene	5.60	−4.12	1.48	1.27
1-heptene	13.80	−11.87	1.93	1.66
1-hexene	12.16	−10.31	1.85	1.68
1-nonene	17.04	−14.96	2.08	2.06
2-methyl-2-butene	10.35	−8.54	1.81	1.31
1-octene	15.44	−13.34	2.09	2.17
1-pentene	10.49	−8.79	1.70	1.66
1-propene	7.26	−5.73	1.54	1.27
trans-2-heptene	13.83	−11.92	1.91	1.66
RMS of prediction set			0.24	
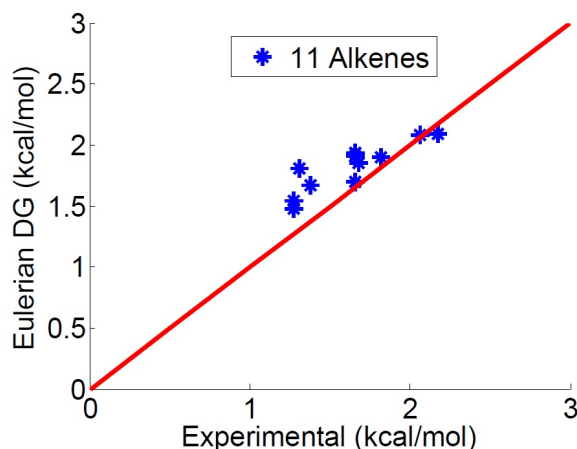
Figure 3: Comparison of model predictions and experimental data of solvation free energies of eleven alkenes.

positions of attractive and repulsive parts are clearly different. That indicates the similarity in the prediction power for total solvation between Eulerian and Lagrangian formulations as well as their differences in the modeling approach and the resulting repulsive and attractive contributions.

Finally, together with the relationship between the phase-field model and Lagrangian formulation, differences between Eulerian and Lagrangian formulations imply a distinction between phase-field models of solvation and our Eulerian formulation. In both models, the targeted functions appear to be quite similar, namely, S=1 in the solute region , S=0 in the solvent area, and $S \in (0, 1)$ at the solute-solvent boundary region. However, the description of each term and its resulting calculation in nonpolar solvation functional are different. Moreover, a phase-field model is normally used to approximate a sharp interface model while Eulerian formulation is not. In the future, the existence of a global minimizer for the Eulerian formulation based total solvation free energy functional will be explored. In addition, current parameterization scheme will be

extended to the solvation prediction for both polar and nonpolar molecules, and then to other applications such as the description of ionic density near a solute.

References

[1] Zhan Chen, Shan Zhao, Jaehun Chun, Dennis G Thomas, Nathan A Baker, Peter W Bates, and GW Wei. Variational approach for nonpolar solvation analysis. *The Journal of chemical physics*, 137(8):084101, 2012.

[2] N. A. Baker. Poisson-Boltzmann methods for biomolecular electrostatics. *Methods in Enzymology*, 383:94–118, 2004.

[3] K. A. Sharp and B. Honig. Calculating total electrostatic energies with the nonlinear Poisson-Boltzmann equatlon. *Journal of Physical Chemistry*, 94:7684–7692, 1990.

[4] M. K. Gilson, M. E. Davis, B. A. Luty, and J. A. McCammon. Computation of electrostatic forces on solvated molecules using the Poisson-Boltzmann equation. *Journal of Physical Chemistry*, 97(14):3591–3600, 1993.

[5] F. H. Stillinger. Structure in aqueous solutions of nonpolar solutes from the standpoint of scaled-particle theory. *J. Solution Chem.*, 2:141 – 158, 1973.

[6] F. Dong, J. A. Wagoner, and N. A. Baker. Assessing the performance of implicit solvation models at a nucleic acid surface. *Physical Chemistry Chemical Physics*, 10:4889–902, 2008.

[7] C. J. Cramer and D. G. Truhlar. Implicit solvation models: Equilibria, structure, spectra, and dynamics. *Chem. Rev.*, 99:2161–2200, 1999.

[8] J. W. Ponder and D. A. Case. Force fields for protein simulations. *Advances in Protein Chemistry*, 66:27–85, 2003.

[9] B. Husowitz and V. Talanquer. Solvent density inhomogeneities and solvation free energies in supercritical diatomic fluids: A density functional approach. *The Journal of Chemical Physics*, 126(5):054508, 2007.

[10] M. R. Reddy, U. C. Singh, and M. D. Erion. Ab initio quantum mechanics-based free energy perturbation method for calculating relative solvation free energies. *Journal of Computational Chemistry*, 28(2):491–4, 2007.

[11] Aleksandr V. Marenich, Christopher J. Cramer, and Donald G. Truhlar. Perspective on foundations of solvatlon modeling: The electrostatic contribution to the free energy of solvation. *Journal of Chemical Theory and Computation*, 4(6):877–887, 2008.

[12] S. C. L. Kamerlin, M. Haranczyk, and A. Warshel. Progress in ab initio qm/mm free-energy simulations of electrostatic energies in proteins: Accelerated qm/mm studies of pk(a), redox reactions and solvation free energies. *Journal of Phys. Chem. B*, 113:1253–1272, 2009.

[13] Z. Chen, N. A. Baker, and G. W. Wei. Differential geometry based solvation models II: Lagrangian formulation. *J. Math. Biol.*, 63:1139–1200, 2010.

[14] Adrian H Elcock, David Sept, and J Andrew McCammon. Computer simulation of protein-protein interactions. *The Journal of Physical Chemistry B*, 105(8):1504–1518, 2001.

[15] R. A. Pierotti. A scaled particle theory of aqueous and nonaqeous solutions. *Chemical Reviews*, 76(6):717–726, 1976.

[16] J. A. Wagoner and N. A. Baker. Assessing implicit models for nonpolar mean solvation forces: the importance of dispersion and volume terms. *Proceedings of the National Academy of Sciences of the United States of America*, 103(22):8331–6, 2006.

[17] E. Gallicchio, M. M. Kubo, and R. M. Levy. Enthalpy-entropy and cavity decomposition of alkane hydration free energies: Numerical results and implications for theories of hydrophobic solvation. *Journal of Physical Chemistry B*, 104(26):6271–85, 2000.

[18] E. Gallicchio, L. Y. Zhang, and R. M. Levy. The SGB/NP hydration free energy model based on the surface generalized Born solvent reaction field and novel nonpolar hydration free energy estimators. *Journal of Computational Chemistry*, 23(5):517–29, 2002.

[19] R. M. Levy, L. Y. Zhang, E. Gallicchio, and A. K. Felts. On the nonpolar hydration free energy of proteins: surface area and continuum solvent models for the solute-solvent interaction energy. *Journal of the American Chemical Society*, 125(31):9523–9530, 2003.

[20] E. Gallicchio and R. M. Levy. AGBNP: An analytic implicit solvent model suitable for molecular dynamics simulations and high-resolution modeling. *Journal of Computational Chemistry*, 25(4):479–499, 2004.

[21] G. W. Wei. Differential geometry based multiscale models. *Bulletin of Mathematical Biology*, 72:1562 – 1622, 2010.

[22] Z. Chen, N. A. Baker, and G. W. Wei. Differential geometry based solvation models I: Eulerian formulation. *J. Comput. Phys.*, 229:8231–8258, 2010.

[23] Z. Chen and G. W. Wei. Differential geometry based solvation models III: Quantum formulation. *J. Chem. Phys.*, 135:1941108, 2011.

[24] J. Dzubiella, J. M. J. Swanson, and J. A. McCammon. Coupling hydrophobicity, dispersion, and electrostatics in continuum solvent models. *Physical Review Letters*, 96:087802, 2006.

[25] L. T. Cheng, Joachim Dzubiella, Andrew J. McCammon, and B. Li. Application of the level-set method to the implicit solvation of nonpolar molecules. *Journal of Chemical Physics*, 127(8), 2007.

[26] Yanxiang Zhao, Yuen-Yick Kwan, Jianwei Che, Bo Li, and J Andrew McCammon. Phase-field approach to implicit solvation of biomolecules with coulomb-field approximation. *The Journal of chemical physics*, 139(2):024111, 2013.

[27] Shibin Dai, Bo Li, and Jianfeng Lu. Convergence of phase-field free energy and boundary force for molecular solvation. *arXiv preprint arXiv:1606.04620*, 2016.

[28] W.L. Jorgensen, D.S. Maxwell, and J. Tirado-Rives. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J Am Chem Soc*, 118:11225–11236, 1996.

[29] Peter Sternberg. The effect of a singular perturbation on nonconvex variational problems. *Archive for Rational Mechanics and Analysis*, 101(3):209–260, 1988.

[30] Bo Li and Yanxiang Zhao. Variational implicit solvation with solute molecular mechanics: from diffuse-interface to sharp-interface models. *SIAM journal on applied mathematics*, 73(1):1–23, 2013.

[31] Luciano Modica. The gradient theory of phase transitions and the minimal interface criterion. *Archive for Rational Mechanics and Analysis*, 98(2):123–142, 1987.

[32] H. Federer. Curvature Measures. *Trans. Amer. Math. Soc.*, 93:418–491, 1959.

[33] P. W. Bates, Z. Chen, Y. H. Sun, G. W. Wei, and S. Zhao. Geometric and potential driving formation and evolution of biomolecular surfaces. *J. Math. Biol.*, 59:193–231, 2009.

[34] S. Cabani, P. Gianni, V Mollica, and L Lepori. Group Contributions to the Thermodynamic Properties of Non-Ionic Organic Solutes in Dilute Aqueous Solution. *Journal of Solution Chemistry*, 10(8):563–595, 1981.

[35] Bao Wang and GW Wei. Parameter optimization in differential geometry based solvation models. *The Journal of chemical physics*, 143(13):134119, 2015.

[36] Zhan Chen. Role of dispersion attraction in differential geometry based nonpolar solvation models. *Molecular Based Mathematical Biology*, 3(1), 2015.

[37] E. L. Ratkova, G. N. Chuev, V. P. Sergiievskyi, and M. V. Fedorov. An accurate prediction of hydration free energies by combination of molecular integral equations theory with structural descriptors. *J. Phys. Chem. B*, 114(37):12068–2079, 2010.