Improved Solvation Models using Boundary Integral Equations

Matthew Knepley and Jaydeep Bardhan

Computational and Applied Mathematics
Rice University

Applied Mathematics Colloquium
Department of Mathematics
UNC Chapel Hill    September 16, 2016
Solvation computation can benefit from operator simplification, and non-Poisson models.
Solvation computation can benefit from operator simplification, and non-Poisson models.
Solvation computation can benefit from operator simplification, and non-Poisson models.
Induced Surface Charge on Lysozyme
Electrostatic Potential $\phi$

\[ \nabla^2 \varphi_{\text{protein}}(\mathbf{r}) = -\sum_i \frac{q_i \delta(\mathbf{r} - \mathbf{r}_i)}{\epsilon_0 \epsilon_{\text{protein}}} \]

\[ \nabla^2 \varphi_{\text{solvent}}(\mathbf{r}) = 0 \]

Region I: protein

Region II: solvent

Surface $\Gamma$

\[ \hat{n}(\mathbf{r}) \]

$\varepsilon_{\text{protein}}$

$\varepsilon_{\text{water}}$
We can write a Boundary Integral Equation (BIE) for the induced surface charge $\sigma$,

$$
\sigma(\vec{r}) + \hat{\epsilon} \int_{\Gamma} \frac{\partial}{\partial n(\vec{r})} \frac{\sigma(\vec{r}') d^2 \vec{r}'}{4\pi ||\vec{r} - \vec{r}'||} = -\hat{\epsilon} \sum_{k=1}^{Q} \frac{\partial}{\partial n(\vec{r})} \frac{q_k}{4\pi ||\vec{r} - \vec{r}_k||} \\
(\mathcal{I} + \hat{\epsilon} \mathcal{D}^*) \sigma(\vec{r}) =
$$

where we define

$$\hat{\epsilon} = 2 \frac{\epsilon_I - \epsilon_{II}}{\epsilon_I + \epsilon_{II}} < 0$$
Outline

1. Approximating the Poisson Operator
   - Approximate Operators
   - Approximate Boundary Conditions

2. Improving the Poisson Operator
Boundary element discretizations of solvation:

- can be expensive to solve
- are more accurate than required by intermediate design iterations
Outline

1. Approximating the Poisson Operator
   - Approximate Operators
   - Approximate Boundary Conditions
Approximating the Poisson Operator

Generalized Born Approximation

The pairwise energy between charges is defined by the *Still equation*:

\[
G_{es}^{ij} = \frac{1}{8\pi} \left( \frac{1}{\epsilon_{II}} - \frac{1}{\epsilon_I} \right) \sum_{i,j}^{N} \frac{q_i q_j}{r_{ij}^2 + R_i R_j e^{-r_{ij}^2/4R_i R_j}}
\]

where the *effective Born radius* is

\[
R_i = \frac{1}{8\pi} \left( \frac{1}{\epsilon_{II}} - \frac{1}{\epsilon_I} \right) \frac{1}{E_i}
\]

where \(E_i\) is the *self-energy* of the charge \(q_i\), the electrostatic energy when atom \(i\) has unit charge and all others are neutral.
GB Problems

- No global potential solution, only energy
- No analysis of the error
  - For example, Salsbury 2006 consists of parameter tuning
- No path for systematic improvement
  - For example, Sigalov 2006 changes the model
- The same atoms have different radii in different
  - molecules,
  - solvents
  - temperatures
- **LOTS of parameters**
  - Nina, Beglov, Roux 1997
GB Problems

- No global potential solution
- No analysis of the error
- For example, Salsbury 2006
- No path for systematic improvement
- For example, Sigalov 2006
- The same atoms have different radii in different molecules, solvents, temperatures
- LOTS of parameters
  - Nina, Beglov, Roux 1997

<table>
<thead>
<tr>
<th>atom</th>
<th>radius (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>2.04 carbonyl C, peptide backbone</td>
</tr>
<tr>
<td>O</td>
<td>1.52 carbonyl oxygen</td>
</tr>
<tr>
<td>CA</td>
<td>2.86 all CA except Gly</td>
</tr>
<tr>
<td>CA</td>
<td>2.38 Gly only</td>
</tr>
<tr>
<td>H*</td>
<td>0.00 all hydrogens</td>
</tr>
<tr>
<td>CB</td>
<td>2.67 all residues</td>
</tr>
<tr>
<td>CG*</td>
<td>2.46 Val, Ile, Arg, Lys, Met, Phe, Thr, Trp, Gln, Glu</td>
</tr>
<tr>
<td>CD*</td>
<td>2.44 Ile, Leu, Arg, Lys</td>
</tr>
<tr>
<td>CD, CG</td>
<td>1.98 Asp, Glu, Asn, Gln</td>
</tr>
<tr>
<td>CB, CG, CD</td>
<td>1.98 Pro only</td>
</tr>
<tr>
<td>CE*, CD*, CZ, CE*</td>
<td>2.00 Tyr, Phe rings</td>
</tr>
<tr>
<td>CE*</td>
<td>1.78 Trp ring only</td>
</tr>
<tr>
<td>CE</td>
<td>2.10 Met only</td>
</tr>
<tr>
<td>CZ, CE</td>
<td>2.80 Arg, Lys</td>
</tr>
<tr>
<td>OE*, OD*</td>
<td>1.42 Glu, Asp, Asn, Gln</td>
</tr>
<tr>
<td>OG*</td>
<td>1.64 Ser, Thr</td>
</tr>
<tr>
<td>OH</td>
<td>1.85 Tyr only</td>
</tr>
<tr>
<td>NH*, NE, NZ</td>
<td>2.13 Arg, Lys</td>
</tr>
<tr>
<td>NE2, ND2</td>
<td>2.15 Gln, Asn</td>
</tr>
<tr>
<td>NE2, ND1</td>
<td>2.31 His only</td>
</tr>
<tr>
<td>NE1</td>
<td>2.40 Trp</td>
</tr>
<tr>
<td>S*</td>
<td>2.00 Met, Cys</td>
</tr>
</tbody>
</table>

* Patches N-term and C-term CAT, CAY: 2.06 Å. CY: 2.04 Å. OY: 1.52 Å. NT: 2.23 Å. * refers to a wild card character.
The *reaction* potential is given by

\[ \phi^R(\vec{r}) = \int_{\Gamma} \frac{\sigma(\vec{r}') d^2\vec{r}'}{4\pi \epsilon_1 ||\vec{r} - \vec{r}'||} = C\sigma \]

which defines \( G_{es} \), the electrostatic part of the solvation free energy

\[ \Delta G_{es} = \frac{1}{2} \langle q, \phi^R \rangle \]

\[ = \frac{1}{2} \langle q, Lq \rangle \]

\[ = \frac{1}{2} \langle q, CA^{-1} Bq \rangle \]

where

\[ Bq = -\hat{\epsilon} \int_{\Omega} \frac{\partial}{\partial n(\vec{r})} \frac{q(\vec{r}') d^3\vec{r}'}{4\pi ||\vec{r} - \vec{r}'||} \]

\[ A\sigma = I + \hat{\epsilon} D^* \]
Boundary Integral-Based Electrostatics Estimation

Coulomb Field Approximation:
uniform normal field

\[
\left(1 - \frac{\hat{\epsilon}}{2}\right) \sigma_{CFA} = Bq
\]

Lower Bound:
no good physical motivation

\[
\left(1 + \frac{\hat{\epsilon}}{2}\right) \sigma_{LB} = Bq
\]
Boundary Integral-Based Electrostatics Estimation

Coulomb Field Approximation:
uniform normal field

\[
\left(1 - \frac{\hat{\epsilon}}{2}\right) \sigma_{CFA} = Bq
\]

Preconditioning:
consider only local effects

\[
\sigma_P = Bq
\]

Eigenvectors: BEM \( e_i \cdot e_j \) BIBEE/P
Theorem: The electrostatic solvation energy $\Delta G_{es}$ has upper and lower bounds given by

$$\frac{1}{2} \left( 1 + \frac{\hat{\epsilon}}{2} \right)^{-1} \langle q, CBq \rangle \leq \frac{1}{2} \langle q, CA^{-1} Bq \rangle \leq \frac{1}{2} \left( 1 - \frac{\hat{\epsilon}}{2} \right)^{-1} \langle q, CBq \rangle,$$

and for spheres and prolate spheroids, we have the improved lower bound,

$$\frac{1}{2} \langle q, CBq \rangle \leq \frac{1}{2} \langle q, CA^{-1} Bq \rangle,$$

and we note that

$$|\hat{\epsilon}| < \frac{1}{2}.$$
Energy Bounds:

Proof: Bardhan, Knepley, Anitescu, JCP, 130(10), 2008

I will break the proof into three steps,

- Replace $C$ with $B$
- Symmetrization
- Eigendecomposition

shown in the following slides.

We will need the single layer operator $S$ for step 1,

$$S\tau(\vec{r}) = \int \frac{\tau(\vec{r}')d^2\vec{r}'}{4\pi||\vec{r} - \vec{r}'||}$$
The potential at the boundary $\Gamma$ given by
\[
\phi^{Coulomb}(\vec{r}) = C^T q
\]
can also be obtained by solving an exterior Neumann problem for $\tau$,
\[
\begin{align*}
\phi^{Coulomb}(\vec{r}) & = S\tau \\
& = S(\mathcal{I} - 2\mathcal{D}^*)^{-1}\left(\frac{2}{\hat{\epsilon}} Bq\right) \\
& = \frac{2}{\hat{\epsilon}} S(\mathcal{I} - 2\mathcal{D}^*)^{-1} Bq
\end{align*}
\]
so that the solvation energy is given by
\[
\frac{1}{2} \left\langle q, CA^{-1} Bq \right\rangle = \frac{1}{\hat{\epsilon}} \left\langle S(\mathcal{I} - 2\mathcal{D}^*)^{-1} Bq, (\mathcal{I} + \hat{\epsilon}\mathcal{D}^*)^{-1} Bq \right\rangle
\]
Plemelj’s symmetrization principle holds that

\[ SD^* = DS \]

and we have

\[ S = S^{1/2} S^{1/2} \]

which means that we can define a Hermitian operator \( H \) similar to \( D^* \)

\[ H = S^{1/2} D^* S^{-1/2} \]

leading to an energy

\[ \frac{1}{2} \left\langle q, CA^{-1} Bq \right\rangle = \frac{1}{\hat{\epsilon}} \left\langle Bq, S^{1/2} (I - 2H)^{-1} (I + \hat{\epsilon}H)^{-1} S^{1/2} Bq \right\rangle \]
The spectrum of $D^*$ is in $[-\frac{1}{2}, \frac{1}{2})$, and the energy is

$$\frac{1}{2} \left\langle q, CA^{-1} Bq \right\rangle = \sum_i \frac{1}{\hat{\epsilon}} (1 - 2\lambda_i)^{-1} (1 + \hat{\epsilon}\lambda_i)^{-1} x_i^2$$

where

$$H = VΛV^T$$

and

$$\tilde{x} = V^T S^{1/2} Bq$$
The BIBEE approximations yield the following bounds

\[
\frac{1}{2} \left< q, CA_{CFA}^{-1} Bq \right> = \sum_i \frac{1}{\hat{\epsilon}} (1 - 2\lambda_i)^{-1} \left( 1 - \frac{\hat{\epsilon}}{2} \right)^{-1} x_i^2
\]

\[
\frac{1}{2} \left< q, CA_P^{-1} Bq \right> = \sum_i \frac{1}{\hat{\epsilon}} (1 - 2\lambda_i)^{-1} x_i^2
\]

\[
\frac{1}{2} \left< q, CA_{LB}^{-1} Bq \right> = \sum_i \frac{1}{\hat{\epsilon}} (1 - 2\lambda_i)^{-1} \left( 1 + \frac{\hat{\epsilon}}{2} \right)^{-1} x_i^2
\]

where we note that

\[|\hat{\epsilon}| < \frac{1}{2}\]
BIBEE Accuracy

Electrostatic solvation free energies of met-enkephalin structures

Snapshots taken from a 500-ps MD simulation at 10-ps intervals.
Crowded Protein Solution

Important for drug design of antibody therapies
BIBEE Scalability

Yokota, Bardhan, Knepley, Barba, Hamada, CPC, 2011.
Outline

1. Approximating the Poisson Operator
   - Approximate Operators
   - Approximate Boundary Conditions
Electrostatic Potential $\phi$

\[ \nabla^2 \varphi_{\text{protein}}(\mathbf{r}) = - \sum_i \frac{q_i \delta(\mathbf{r} - \mathbf{r}_i)}{\varepsilon_0 \epsilon_{\text{protein}}} \]

\[ \nabla^2 \varphi_{\text{solvent}}(\mathbf{r}) = 0 \]
Kirkwood’s Solution (1934)

The potential inside Region I is given by

$$\Phi_I = \sum_{k=1}^{Q} \frac{q_k}{\epsilon_1 |\vec{r} - \vec{r}_k|} + \psi,$$

and the potential in Region II is given by

$$\Phi_{II} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{C_{nm}}{r^{n+1}} P_n^m(\cos \theta)e^{im\phi}.$$
Kirkwood’s Solution (1934)

The reaction potential $\psi$ is expanded in a series

$$\psi = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} B_{nm} r^n P_n^m (\cos \theta) e^{im\phi}.$$  

and the source distribution is also expanded

$$\sum_{k=1}^{Q} \frac{q_k}{\epsilon_1 |\vec{r} - \vec{r}_k|} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{E_{nm}}{\epsilon_1 r^{n+1}} P_n^m (\cos \theta) e^{im\phi}.$$
By applying the boundary conditions, letting the sphere have radius \( b \),

\[
\Phi_I|_{r=b} = \Phi_{II}|_{r=b} \\
\epsilon_I \frac{\partial \Phi_I}{\partial r}|_{r=b} = \epsilon_{II} \frac{\partial \Phi_{II}}{\partial r}|_{r=b}
\]

we can eliminate \( C_{nm} \), and determine the reaction potential coefficients in terms of the source distribution,

\[
B_{nm} = \frac{1}{\epsilon_I b^{2n+1}} \frac{(\epsilon_I - \epsilon_{II})(n + 1)}{\epsilon_I n + \epsilon_{II}(n + 1)} E_{nm}.
\]
Approximate Boundary Conditions

**Theorem:** The BIBEE boundary integral operator approximations

\[
A_{CFA} = I \left(1 + \frac{\hat{\epsilon}}{2}\right)
\]

\[
A_P = I
\]

have an equivalent PDE formulation,

\[
\epsilon_I \Delta \Phi_{CFA,P} = \sum_{k=1}^{Q} q_k \delta(\vec{r} - \vec{r}_k) \quad \frac{\epsilon_I}{\epsilon_{II}} \frac{\partial \Phi_{CFA,P}^C}{\partial r} \bigg|_{r=b} = \frac{\partial \Phi_{II}}{\partial r} \bigg|_{r=b} - \frac{\partial \psi_{CFA,P}}{\partial r} \bigg|_{r=b} \quad \text{or}
\]

\[
\epsilon_{II} \Delta \Phi_{CFA,P} = 0 \quad \frac{3\epsilon_I - \epsilon_{II}}{\epsilon_I + \epsilon_{II}} \frac{\partial \Phi_{CFA,P}^C}{\partial r} \bigg|_{r=b} = \frac{\partial \Phi_{II}}{\partial r} \bigg|_{r=b} - \frac{\partial \psi_{P}}{\partial r} \bigg|_{r=b},
\]

where \(\Phi_{CFA}^C\) is the Coulomb field due to interior charges.
**Theorem:** For spherical solute, the BIBEE boundary integral operator approximations have eigenspaces are identical to that of the original operator.

BEM eigenvector $e_i \cdot e_j$ BIBEE/P eigenvector

In order to show that these PDEs are equivalent to the original BIEs,

- Start with the fundamental solution to Laplace’s equation $G(r, r')$
- Note that $\int_{\Gamma} G(r, r')\sigma(r')\,d\Gamma$ satisfies the bulk equation and decay at infinity
- Insertion into the approximate BC gives the BIBEE boundary integral approximation
**Proof:** Bardhan and Knepley, JCP, 135(12), 2011.

In order to show that these PDEs are equivalent to the original BIEs,

- Start with the fundamental solution to Laplace’s equation $G(r, r')$
- Note that $\int_{\Gamma} G(r, r')\sigma(r')d\Gamma$ satisfies the bulk equation and decay at infinity
- Insertion into the approximate BC gives the BIBEE boundary integral approximation
**Proof**: Bardhan and Knepley, JCP, 135(12), 2011.

In order to show that these PDEs are equivalent to the original BIEs,

- Start with the fundamental solution to Laplace’s equation $G(r, r')$
- Note that $\int_{\Gamma} G(r, r') \sigma(r') d\Gamma$ satisfies the bulk equation and decay at infinity
- Insertion into the approximate BC gives the BIBEE boundary integral approximation

In order to show that these PDEs are equivalent to the original BIEs,

- Start with the fundamental solution to Laplace’s equation $G(r, r')$
- Note that $\int_{\Gamma} G(r, r') \sigma(r') d\Gamma$ satisfies the bulk equation and decay at infinity
- Insertion into the approximate BC gives the BIBEE boundary integral approximation
**Proof**: Bardhan and Knepley, JCP, 135(12), 2011.

In order to show that these integral operators share a common eigenbasis,

- Note that, for a spherical boundary, $\mathcal{D}^*$ is compact and has a pure point spectrum
- Examine the effect of the operator on a unit spherical harmonic charge distribution
- Use completeness of the spherical harmonic basis

In order to show that these integral operators share a common eigenbasis,

- Note that, for a spherical boundary, $D^*$ is compact and has a pure point spectrum
- Examine the effect of the operator on a unit spherical harmonic charge distribution
- Use completeness of the spherical harmonic basis

In order to show that these integral operators share a common eigenbasis,

- Note that, for a spherical boundary, $D^*$ is compact and has a pure point spectrum
- Examine the effect of the operator on a unit spherical harmonic charge distribution
- Use completeness of the spherical harmonic basis
Proof: Bardhan and Knepley, JCP, **135**(12), 2011.

In order to show that these integral operators share a common eigenbasis,

- Note that, for a spherical boundary, $\mathcal{D}^*$ is compact and has a pure point spectrum.
- Examine the effect of the operator on a unit spherical harmonic charge distribution.
- Use completeness of the spherical harmonic basis.

In order to show that these integral operators share a common eigenbasis,

- Note that, for a spherical boundary, $D^*$ is compact and has a pure point spectrum
- Examine the effect of the operator on a unit spherical harmonic charge distribution
- Use completeness of the spherical harmonic basis

The result does not hold for general boundaries.
Note that the approximate solutions are \textit{separable}:

\begin{align*}
B_{nm} &= \frac{1}{\epsilon_1 n + \epsilon_2 (n + 1)} \gamma_{nm} \\
B^{CFA}_{nm} &= \frac{1}{\epsilon_2} \frac{1}{2n + 1} \gamma_{nm} \\
B^P_{nm} &= \frac{1}{\epsilon_1 + \epsilon_2} \frac{1}{n + \frac{1}{2}} \gamma_{nm}.
\end{align*}

If $\epsilon_\parallel = \epsilon_\perp = \epsilon$, both approximations are exact:

\begin{align*}
B_{nm} &= B^{CFA}_{nm} = B^P_{nm} = \frac{1}{\epsilon (2n + 1)} \gamma_{nm}.
\end{align*}
Note that the approximate solutions are \textit{separable}:

\[
B_{nm} = \frac{1}{\epsilon_1 n + \epsilon_2 (n+1)} \gamma_{nm}
\]

\[
B_{nm}^{CFA} = \frac{1}{\epsilon_2} \frac{1}{2n+1} \gamma_{nm}
\]

\[
B_{nm}^P = \frac{1}{\epsilon_1 + \epsilon_2} \frac{1}{n + \frac{1}{2}} \gamma_{nm}.
\]

If $\epsilon_I = \epsilon_{II} = \epsilon$, both approximations are exact:

\[
B_{nm} = B_{nm}^{CFA} = B_{nm}^P = \frac{1}{\epsilon (2n+1)} \gamma_{nm}.
\]
BIBEE/CFA is exact for the $n = 0$ mode,

$$B_{00} = B_{00}^{CFA} = \frac{\gamma_{00}}{\epsilon_2},$$

whereas BIBEE/P approaches the exact response in the limit $n \to \infty$:

$$\lim_{n \to \infty} B_{nm} = \lim_{n \to \infty} B_{nm}^P = \frac{1}{(\epsilon_1 + \epsilon_2)n} \gamma_{nm}. $$
BIBEE/CFA is exact for the $n = 0$ mode,

$$B_{00} = B_{00}^{CFA} = \frac{\gamma_{00}}{\epsilon_2},$$

whereas BIBEE/P approaches the exact response in the limit $n \to \infty$:

$$\lim_{n \to \infty} B_{nm} = \lim_{n \to \infty} B_{nm}^P = \frac{1}{(\epsilon_1 + \epsilon_2)n}\gamma_{nm}.$$
In the limit $\epsilon_1 / \epsilon_2 \to 0$,

$$\lim_{\epsilon_1 / \epsilon_2 \to 0} B_{nm} = \frac{\gamma_{nm}}{\epsilon_2 (n + 1)}$$

$$\lim_{\epsilon_1 / \epsilon_2 \to 0} B_{nm}^{CFA} = \frac{\gamma_{nm}}{\epsilon_2 (2n + 1)}$$

$$\lim_{\epsilon_1 / \epsilon_2 \to 0} B_{nm}^{P} = \frac{\gamma_{nm}}{\epsilon_2 \left(n + \frac{1}{2}\right)}$$

so that the approximation ratios are given by

$$\frac{B_{nm}^{CFA}}{B_{nm}} = \frac{n + 1}{2n + 1}$$

$$\frac{B_{nm}^{P}}{B_{nm}} = \frac{n + 1}{n + \frac{1}{2}}.$$
Improved Accuracy

BIBEE/I interpolates between BIBEE/CFA and BIBEE/P

Bardhan, Knepley, JCP, 2011.
We examined the more complex problem of **protein-ligand binding** using trypsin and bovine pancreatic trypsin inhibitor (BPTI), using *electrostatic component analysis* to identify residue contributions to binding and molecular recognition.
Basis Augmentation

Looking at an ensemble of synthetic proteins, we can see that BIBEE/CFA becomes more accurate as the monopole moment increases, and BIBEE/P more accurate as it decreases. BIBEE/I is accurate for spheres, but must be extended for ellipses.
For ellipses, we add a few low order multipole moments, up to the octopole, to recover 5% accuracy for all synthetic proteins tested.
Boundary element discretizations of the solvation problem:

- can be expensive to solve

  - Bounding the electrostatic free energies associated with linear continuum models of molecular solvation, Bardhan, Knepley, Anitescu, JCP, 2009

- are more accurate than required by intermediate design iterations

  - Analysis of fast boundary-integral approximations for modeling electrostatic contributions of molecular binding, Kreienkamp, et al., Molecular-Based Mathematical Biology, 2013
Boundary element discretizations of the solvation problem:

- can be expensive to solve
  - Bounding the electrostatic free energies associated with linear continuum models of molecular solvation, Bardhan, Knepley, Anitescu, JCP, 2009

- are more accurate than required by intermediate design iterations
  - Analysis of fast boundary-integral approximations for modeling electrostatic contributions of molecular binding, Kreienkamp, et al., Molecular-Based Mathematical Biology, 2013
Boundary element discretizations of the solvation problem:

- can be expensive to solve
  - Bounding the electrostatic free energies associated with linear continuum models of molecular solvation, Bardhan, Knepley, Anitescu, JCP, 2009

- are more accurate than required by intermediate design iterations
  - Analysis of fast boundary-integral approximations for modeling electrostatic contributions of molecular binding, Kreienkamp, et al., Molecular-Based Mathematical Biology, 2013
Outline

1. Approximating the Poisson Operator

2. Improving the Poisson Operator
Improving the Poisson Operator

Origins of Electrostatic Asymmetry

The figure shows a plot comparing different boundary conditions for solving the Poisson equation in the context of solvation. The x-axis represents $E_{\text{Coul}}$ normalized by $\sigma$, and the y-axis represents $E_n$ normalized by $E_{\text{Coul}}$. The plot includes:

- Blue circles: Explicit-solvent molecular dynamics FEP
- Red line: Standard Maxwell boundary condition
- Black boxes: Proposed nonlinear boundary condition

The graph suggests that the proposed nonlinear boundary condition aligns better with the explicit-solvent molecular dynamics FEP compared to the standard Maxwell boundary condition.
Origins of Electrostatic Asymmetry

\( \varphi_{\text{static}} \neq 0 \)
Deeply buried charge

The nonzero slope at q=0 is the static potential

Asymmetry for a deeply buried charge is exclusively due to the static potential

Asymmetry for a solvent-exposed charge results from both the static potential and the hydrogen-oxygen size difference

Solvent-exposed $\Delta G = \begin{cases} 
\frac{1}{2} L^- q^2 + \varphi_{\text{static}} q, & q \leq 0 \\
\frac{1}{2} L^+ q^2 + \varphi_{\text{static}} q, & q \geq 0 
\end{cases}$
Solvation-Layer Interface Condition (SLIC)

Instead of assuming the model and energy and deriving the radii,

\[ \epsilon_I \frac{\partial \Phi_I}{\partial n} = \epsilon_{II} \frac{\partial \Phi_{II}}{\partial n} \]
Main Idea

Solvation-Layer Interface Condition (SLIC)

assume the energy and radii and derive the model.

\[
(\epsilon_l - \Delta \epsilon h(E_n)) \frac{\partial \Phi_l}{\partial n} = (\epsilon_{ll} - \Delta \epsilon h(E_n)) \frac{\partial \Phi_{ll}}{\partial n}
\]
Main Idea

Solvation-Layer Interface Condition (SLIC)

Using our correspondence with the BIE form,

\[
\left( \mathcal{I} + h(E_n) + \hat{\epsilon} \left( -\frac{1}{2} \mathcal{I} + D^* \right) \right) \sigma = \hat{\epsilon} \sum_{k=1}^{Q} \frac{\partial G}{\partial n}
\]

where \( h \) is a diagonal nonlinear integral operator.
Improving the Poisson Operator

**SLIC**

**Boundary Perturbation**

\[ h(E_n) = \alpha \tanh (\beta E_n - \gamma) + \mu \]

where

- \( \alpha \) Size of the asymmetry
- \( \beta \) Width of the transition region
- \( \gamma \) The transition field strength
- \( \mu \) Assures \( h(0) = 0 \), so \( \mu = -\alpha \tanh(-\gamma) \)
Improving the Poisson Operator

Accuracy of SLIC Residues

Charging free energy (kcal/mol)

MD FEP
Poisson, Roux radii
NLBC, extrapolated
NLBC, 4 vertices/A²
NLBC, 2 vertices/A²

ARG ASP CYS GLU HIS LYS TYR

M. Knepley (Rice)

Solvation

UNC6 45 / 60
Improving the Poisson Operator

Accuracy of SLIC Protonation

![Graph showing G_{prot,salt} values for different amino acids with Symmetric LPB and NLBC LPB comparisons.]

- ARG
- ASP
- CYS
- GLU
- HIS
- LYS
- TYR

G_{prot,salt} (kcal/mol)
Improving the Poisson Operator

Accuracy of SLIC

Synthetic Molecules

Problem Index

Hydration Free Energy (kcal/mol)

- Negative (MD, Mobley et al)
- Positive (MD, Mobley et al)
- Negative (NLBC, this work)
- Positive (NLBC, this work)

M. Knepley (Rice)

Solvation

UNC6 47 / 60
Improving the Poisson Operator

Accuracy of SLIC
Synthetic Molecules

Asymmetry in Electrostatic Free Energy (kcal/mol)

MD, Mobley et al
NLBC, this work

Number of Atoms in Ring

M. Knepley (Rice)

Solvation

UNC6 48 / 60
Improving the Poisson Operator

Accuracy of SLIC
Synthetic Molecules

![Graph showing electrostatic solvation free energy vs. number of atoms in ring for different conditions.]

- Negative (MD, Mobley et al)
- Positive (MD, Mobley et al)
- Negative (NLBC, this work)
- Positive (NLBC, this work)
The parameters show linear temperature dependence
### Model Validation

**Courtesy A. Molvai Tabrizi**

<table>
<thead>
<tr>
<th></th>
<th><img src="https://via.placeholder.com/150" alt="Molecule" /></th>
<th><img src="https://via.placeholder.com/150" alt="Molecule" /></th>
<th><img src="https://via.placeholder.com/150" alt="Molecule" /></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Water</strong></td>
<td>$\text{H}_2\text{O}$</td>
<td>$\text{Formamide}$</td>
<td>$\text{Dimethyl sulfoxide}$ $\text{C}_2\text{H}_6\text{OS}$</td>
</tr>
<tr>
<td><strong>Methanol</strong></td>
<td>$\text{CH}_3\text{OH}$</td>
<td>$\text{Acetonitrile}$</td>
<td>$\text{Nitromethane}$ $\text{CH}_3\text{NO}_2$</td>
</tr>
<tr>
<td><strong>Ethanol</strong></td>
<td>$\text{C}_2\text{H}_5\text{OH}$</td>
<td>$\text{Dimethyl formamide}$</td>
<td>$\text{Propylene carbonate}$ $\text{CH}_3\text{C}_2\text{H}_3\text{O}_2\text{CO}$</td>
</tr>
</tbody>
</table>

- **Water** $\text{H}_2\text{O}$
- **Formamide** $\text{CH}_3\text{NO}$
- **Dimethyl sulfoxide** $\text{C}_2\text{H}_6\text{OS}$
- **Methanol** $\text{CH}_3\text{OH}$
- **Acetonitrile** $\text{C}_2\text{H}_3\text{N}$
- **Nitromethane** $\text{CH}_3\text{NO}_2$
- **Ethanol** $\text{C}_2\text{H}_5\text{OH}$
- **Dimethyl formamide** $\text{C}_3\text{H}_7\text{NO}$
- **Propylene carbonate** $\text{CH}_3\text{C}_2\text{H}_3\text{O}_2\text{CO}$
## Model Validation

**Courtesy A. Molvai Tabrizi**

<table>
<thead>
<tr>
<th>Solvent</th>
<th>$r_s$ (Å)</th>
<th>$\epsilon_{out}(T)$</th>
<th>$\epsilon_{out}(25^\circ C)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>1.370</td>
<td>$\epsilon_{out} = 87.740 - 4.0008e-1 \ T + 9.398e-4 \ T^2 - 1.410e-6 \ T^3$</td>
<td>78.3</td>
</tr>
<tr>
<td>MeOH</td>
<td>1.855</td>
<td>$\log_{10} \epsilon_{out} = \log_{10}(32.63) - 2.64e-3(T - 25)$</td>
<td>32.6</td>
</tr>
<tr>
<td>EtOH</td>
<td>2.180</td>
<td>$\log_{10} \epsilon_{out} = \log_{10}(24.30) - 02.70e-3(T - 25)$</td>
<td>24.3</td>
</tr>
<tr>
<td>F</td>
<td>1.725</td>
<td>$\epsilon_{out} = 109 - 7.2e-1(T - 20)$</td>
<td>105.4</td>
</tr>
<tr>
<td>AN</td>
<td>2.135</td>
<td>$\epsilon_{out} = 37.50 - 1.6e-1(T - 20)$</td>
<td>36.7</td>
</tr>
<tr>
<td>DMF</td>
<td>2.585</td>
<td>$\epsilon_{out} = 42.04569 - 2.204448e-1T + 7.718531e-4T^2 - 1.000389e-6T^3$</td>
<td>37.0</td>
</tr>
<tr>
<td>DMSO</td>
<td>2.455</td>
<td>$\epsilon_{out} = -60.5 + (5.7e4/(T + 273.15)) - (7.5e6/(T + 273.15)^2)$</td>
<td>46.3</td>
</tr>
<tr>
<td>NM</td>
<td>2.155</td>
<td>$\log_{10} \epsilon_{out} = \log_{10}(35.8) - 1.89e-3(T - 30)$</td>
<td>36.6</td>
</tr>
<tr>
<td>PC</td>
<td>2.680</td>
<td>$\epsilon_{out} = 56.670738 + 2.58431e-1T - 7.7143e-4T^2$</td>
<td>62.6</td>
</tr>
</tbody>
</table>
Model validation and verification using experiment

- ΔG (kJ mol⁻¹)
- ΔS (J K⁻¹ mol⁻¹)
- Cp (J K⁻¹ mol⁻¹)

**Dimethyl formamide @ 25°C**

C₃H₇NO

- Classical Born
- Asymmetric MSA (Fawcett 1992)
- SLIC without static potential
- SLIC with static potential

- Li  Na  K  Rb  Cs  Cl  Br  I
A. Molavi Tabrizi, M.G. Knepley, and J.P. Bardhan, *Generalising the mean spherical approximation as a multiscale, nonlinear boundary condition at the solute-solvent interface*, Molecular Physics (2016).
## Thermodynamic Predictions

Courtesy A. Molvai Tabrizi

<table>
<thead>
<tr>
<th>Solvent</th>
<th>Ion</th>
<th>$\Delta G$ (kJ mol$^{-1}$)</th>
<th>$\Delta S$ (J K$^{-1}$ mol$^{-1}$)</th>
<th>$C_p$ (J K$^{-1}$ mol$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>F$^-$</td>
<td>-430 (-429)</td>
<td>-67 (-115)</td>
<td>-86 (-45)</td>
</tr>
<tr>
<td>MeOH</td>
<td>Rb$^+$</td>
<td>-326 (-319)</td>
<td>-178 (-175)</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>F$^-$</td>
<td>-415</td>
<td>-116</td>
<td>-79 (-131)</td>
</tr>
<tr>
<td>EtOH</td>
<td>Rb$^+$</td>
<td>-319 (-313)</td>
<td>-197 (-187)</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td>F$^-$</td>
<td>-405</td>
<td>-145</td>
<td>-153 (-194)</td>
</tr>
<tr>
<td>F</td>
<td>Rb$^+$</td>
<td>-340 (-334)</td>
<td>-135 (-130)</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>F$^-$</td>
<td>-418</td>
<td>-128</td>
<td>36 (28)</td>
</tr>
<tr>
<td>AN</td>
<td>F$^-$</td>
<td>-390</td>
<td>-192</td>
<td>147</td>
</tr>
<tr>
<td>DMF</td>
<td>F$^-$</td>
<td>-389</td>
<td>-230</td>
<td>105</td>
</tr>
<tr>
<td>DMSO</td>
<td>Rb$^+$</td>
<td>-348 (-339)</td>
<td>-151 (-180)</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>F$^-$</td>
<td>-400</td>
<td>-160</td>
<td>186 (60)</td>
</tr>
<tr>
<td>NM</td>
<td>Rb$^+$</td>
<td>-324 (-318)</td>
<td>-186 (-183)</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>F$^-$</td>
<td>-391</td>
<td>-182</td>
<td>95 (71)</td>
</tr>
<tr>
<td>PC</td>
<td>F$^-$</td>
<td>-394</td>
<td>-149</td>
<td>67</td>
</tr>
</tbody>
</table>

Experimental Data in Parentheses
A. Molavi Tabrizi, S. Goossens, M.G. Knepley, and J.P. Bardhan,

*Predicting solvation thermodynamics with dielectric continuum theory and a solvation-layer interface condition (SLIC).*

Where does SLIC fail?

- Large packing fraction
  - No charge oscillation or overcharging
  - Could use CDFT

- No dielectric saturation
  - Could be possible with different function

- No long range correlations
  - Use nonlocal dielectric
Future Work

- More complex solutes
- Mixtures
- Integration into community code
  - Psi4, QChem, APBS
- Integrate into conformational search
  - Kavrakis Lab at Rice
Thank You!

http://www.caam.rice.edu/~mk51