Minimal parameter implicit solvent model for electronic structure calculations

Hatem H. Helal
Electronic structure discussion group
October 8, 2008
What is implicit solvent?

Introduce a dielectric medium which captures the physics of a system embedded in water so that the computational effort is focused on the system we are interested in.
Just add water...

Ubiquitin with water
Electrostatics in periodic systems
Simple example

What is the work required to build this periodic array of point charges?
A cause for worry?

\[ E = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{q_i^2}{|x_i - x_j|} \]

Notice that the electrostatic energy is proportional to the number of pairs of charges.

\[ E \propto \frac{N(N-1)}{2} \]

Normalizing to find the energy of a single “unit cell” will give an energy which diverges as N.
How about this array?

\[
\begin{array}{cccccc}
\ldots & \bullet & \bullet & \bullet & -q & \ldots \\
\bullet & \bullet & \bullet & +q & & \\
\end{array}
\]

The total energy is a sum of three distinct contributions which taken on their own will diverge

\[
E = E(+q, +q) + E(-q, -q) + E(+q, -q)
\]

\[
\approx N \left( \frac{q^2 \Delta}{a} \right) \sum_{j=1}^{N} \frac{1}{j^2}
\]

The key to getting a nice well behaved energy per cell is ensuring charge neutrality.
Electrostatics in the Kohn-Sham functional

\[ E_{es}[\rho] = E_{e-e}[\rho] + E_{e-I}[\rho] + E_{I-I} \]

• All three contributions taken independently yield a divergent energy for an infinitely periodic system.

• Taken together though the total electrostatic energy converges to a well defined energy per unit cell so long as the cell is charge neutral.

• Making use of a uniform compensating background charge allows us to rearrange the functional so that we avoid trying to compute diverging energies.
The truth about the Hartree term

• Define the neutral charge density:

\[ \rho_0(\mathbf{r}) \equiv \rho(\mathbf{r}) - \langle \rho(\mathbf{r}) \rangle \]

• The “Hartree energy” is then self-interaction energy of the neutralized charge density

\[ E_H[\rho] = \frac{1}{2} \int \rho_0(\mathbf{r}) \phi[\rho_0] \, d\mathbf{r} \]

• Where the potential is found by solving the Poisson equation under periodic boundary conditions

\[ -\nabla^2 \phi = 4\pi \rho_0 \]
Fourier series solution to the Poisson equation

\[
\phi(r) = \sum_G \phi(G)e^{iG \cdot r}
\]

\[
\rho_0(r) = \sum_G \rho_0(G)e^{iG \cdot r}
\]

\[-\nabla^2 \phi = 4\pi \rho_0\]

\[
\phi(r) = \sum_G \frac{4\pi}{G^2} \rho_0(G)e^{iG \cdot r}
\]
\textbf{G=0 difficulty}

- From the preceding solution we see that the Fourier coefficients are related to each other by

\[ \phi(G) = \frac{4\pi}{G^2} \rho_0(G) \]

- At first inspection, the G=0 coefficient of the potential would appear to diverge.

- Formally we need to evaluate the limit as G approaches zero:

\[ \phi(G = 0) = \lim_{G \to 0} \frac{4\pi}{G^2} \rho_0(G) \]
Evaluate the $G=0$ term

• The $G=0$ Fourier coefficient of the neutral density corresponds to its average value, which is of course zero:

$$
\rho_0(G = 0) = \frac{1}{\Omega} \int_{\Omega} \rho_0(\mathbf{r}) \, d\mathbf{r} = 0
$$

• Likewise, the $G=0$ Fourier coefficient of the Hartree potential corresponds to the average value over the cell:

$$
\phi(G = 0) = \frac{1}{\Omega} \int_{\Omega} \phi(\mathbf{r}) \, d\mathbf{r} = \langle \phi(\mathbf{r}) \rangle
$$
Potential reference

• For our infinitely periodic system both the charge and the potential extend out to infinity.

• There is always a sense of arbitrariness when dealing with potential since we have to decide on a ‘reference point’ from which we measure the potential.

• For our infinitely periodic system, the most natural choice for a reference is to require that the average value of the potential is zero over a unit cell.

• This choice for the potential reference amounts to fixing the $G=0$ component to zero.
Useful result

• Notice that we never need to compute the compensating background charge:

\[
\rho_0(G) = \frac{1}{\Omega} \int_\Omega (\rho(r) - \langle \rho(r) \rangle) e^{-iG \cdot r} \, dr \\
= \frac{1}{\Omega} \int_\Omega \rho(r) e^{-iG \cdot r} \, dr - \frac{1}{\Omega} \int_\Omega \langle \rho(r) \rangle e^{-iG \cdot r} \, dr \\
= \rho(G) + 0 \\
= \rho(G)
\]
Hartree potential evaluation

- Combining all this leads to the Hartree potential as it is evaluated in any plane-wave DFT code

\[ \phi(r) = \sum_{G \neq 0} \frac{4\pi}{G^2} \rho(G)e^{iG \cdot r} \]

- Skipping the \( G=0 \) term is justified since we have seen that it is fixed at zero.
Dielectric medium

• Uses the electron density to define the solvent cavity

• The dielectric functional is expressed as:

\[
\epsilon[\rho(r)] = 1 + \frac{\epsilon_\infty - 1}{2} \left( 1 + \frac{1 - (\rho/\rho_0)^{2\beta}}{1 + (\rho/\rho_0)^{2\beta}} \right)
\]

• This results in an implicit solvent model which only depends on two parameters

\[\rho_0\] Density threshold which modulates the size of the solvation cavity

\[\beta\] Smoothes the transition in the function from vacuum to the bulk permittivity
Solvent cavity

- Smooth surface of solvation cavity results from natural definition through the charge density
- Dielectric parameters chosen such that the solvation effect on the total energy matches the cohesion energy of liquid water

$$\epsilon[\rho(r)] = 1.01$$

$$\beta = 1.3 \quad \rho_0 = 0.0004$$

Dielectric functional for self-consistent solvent cavity

\[
\epsilon[\rho(\mathbf{r})] = 1 + \frac{\epsilon_{\infty} - 1}{2} \left( 1 + \frac{1 - (\rho/\rho_0)^{2/\beta}}{1 + (\rho/\rho_0)^{2/\beta}} \right)
\]

DFT with implicit solvent

• Replace the explicit solvent molecules with a dielectric continuum that implicitly replicates the electrostatic interaction between solvent and solute

\[ \epsilon[\rho(r)] \rightarrow -\nabla \cdot (\epsilon \nabla \phi) = 4\pi \rho_0 \]

Dielectric functional

Poisson equation in the presence of a dielectric

• The Hartree term in the dielectric medium is

\[ E[\rho, \epsilon] = \frac{1}{2} \int_{\Omega} \rho_0(r) \phi[\rho_0, \epsilon] \, dr = \frac{1}{8\pi} \int_{\Omega} \epsilon[\rho] |\nabla \phi|^2 \, dr \]
DFT with implicit solvent

• The electrostatic contribution to the Kohn-Sham potential is found from the functional derivative:

\[
\frac{\delta E_{es}[\rho]}{\delta \rho} = \phi(\mathbf{r}) + V_\epsilon(\mathbf{r})
\]

• The additional term arises from the dependence of the dielectric functional on the charge density

\[
V_\epsilon(\mathbf{r}) \equiv -\frac{1}{8\pi} |\nabla \phi(\mathbf{r})|^2 \frac{\delta \epsilon}{\delta \rho}
\]

• This results in a tunable solvent cavity as self-consistency is achieved
Solving the generalized Poisson equation

\[-\nabla \cdot (\epsilon \nabla \phi) = 4\pi \rho_0\]
\[\phi(r) = 0\] on cell boundary

• No general direct solution is possible so we have to rely on iterative methods.
• Currently employ a finite difference based multigrid solver.
Multigrid approach

- Using multiple grids improves convergence
- Computational effort scales linearly with the number of grid points
Implementation
Local potential calculation

Implicit solvent?

No

Yes

ρ(r) \xrightarrow{\text{FFT}} ρ(G)

φ(r) \xrightarrow{\text{FFT}} φ(G) = \frac{4\pi}{G^2} ρ(G)

E_H[ρ] = 2\pi Ω \sum_{G \neq 0} \frac{|ρ(G)|^2}{G^2}

Standard Hartree evaluation

ρ_0 = ρ - \langle ρ \rangle

ε[ρ, r]

Multigrid

−∇ \cdot (ε\nabla φ) = 4πρ_0

φ = 0 on \partialΩ

V_ε(r) = \frac{1}{8\pi} |∇φ|^2 \frac{∂ε}{∂ρ}

E_H[ρ] = \frac{1}{2} \int ρ_0 φ dr

Hartree in dielectric evaluation
Energy gradient test

\[ \rho' \equiv \rho + \alpha \frac{\delta E}{\delta \rho} \]

\[ \delta E = E[\rho'] - E[\rho] = \int \frac{\delta E}{\delta \rho} \delta \rho \, dr \]

\[ \frac{E[\rho'] - E[\rho]}{\alpha} = \int \left( \frac{\delta E}{\delta \rho} \right)^2 \, dr \]

The ratio of the numerical and predicted gradients should converge to one while varying the step size when everything is done correctly...
Hartree in dielectric gradient check

\[ \delta E[\rho, \epsilon] = \delta E_\rho + \delta E_\epsilon \]

\[ \delta E_\rho = E[\rho', \epsilon] - E[\rho, \epsilon] = \int \frac{\delta E_\rho}{\delta \rho} \delta \rho \, dr \]

\[ \frac{E'_\rho - E_\rho}{\alpha} = \int \left( \frac{\delta E_\rho}{\delta \rho} \right)^2 \, dr \]

\[ \frac{\delta E_\rho}{\delta \rho} = \phi(r) \]
A suspect result

<table>
<thead>
<tr>
<th>Cell size (Bohr)</th>
<th>Number of grid points in each direction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>33</td>
</tr>
<tr>
<td>10</td>
<td>0.954</td>
</tr>
<tr>
<td>20</td>
<td>0.992</td>
</tr>
<tr>
<td>30</td>
<td>0.996</td>
</tr>
</tbody>
</table>

Seems that $\frac{\delta E}{\delta \rho} = \phi(r)$ is mostly correct.
A very careful functional derivative

\[
E[\rho] = \frac{1}{2} \int \rho_0(\mathbf{r})\phi(\mathbf{r}) \, d\mathbf{r}
\]

\[
\delta E = \int \frac{\delta E}{\delta \rho_0} \delta \rho_0 \, d\mathbf{r} = \int \phi(\mathbf{r})\delta \rho_0 \, d\mathbf{r}
\]

\[
\therefore \frac{\delta E}{\delta \rho_0} = \phi(\mathbf{r})
\]

This is just the amount of work needed to introduce a \(\delta \rho_0\) into a region with a potential \(\phi(\mathbf{r})\)
A very careful functional derivative

\[
\begin{align*}
\delta \rho_0 &= \int \frac{\delta \rho_0}{\delta \rho} \delta \rho \, d\mathbf{r} \\
\frac{\delta \rho_0}{\delta \rho} &= \delta (\mathbf{r}' - \mathbf{r}) - \frac{1}{\Omega} \\
\frac{\delta E}{\delta \rho} &= \int \frac{\delta E}{\delta \rho_0} \frac{\delta \rho_0}{\delta \rho} \, d\mathbf{r}' \\
\therefore \frac{\delta E}{\delta \rho} &= \phi(\mathbf{r}) - \langle \phi(\mathbf{r}) \rangle
\end{align*}
\]

This result is of course consistent with the choice of the potential reference.
The correct result

<table>
<thead>
<tr>
<th>Cell size (Bohr)</th>
<th>Number of grid points in each direction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>33</td>
</tr>
<tr>
<td>10</td>
<td>1.000000157</td>
</tr>
<tr>
<td>20</td>
<td>1.000000344</td>
</tr>
<tr>
<td>30</td>
<td>1.000000693</td>
</tr>
</tbody>
</table>

\[
\frac{\delta E_\rho}{\delta \rho} = \phi(r) - \langle \phi(r) \rangle \text{ is correct!}
\]
Applications
NMR from MD

- Use molecular dynamics with implicit solvent to generate a set of possible conformations.
- Perform DFT calculations with implicit solvent for each of these conformations.
- Average NMR coupling constants to obtain realistic predictions.
Optical properties

- Use TDDFT(+U) to study optical-response properties of biological metalloproteins.
- Solvation can have dramatic effects on the optical response.
Enzyme binding

Solvation effects can be important in determining whether it is energetically favorable for an enzyme to bind to an active site.
DNA packaging in nature

Hernan G. Garcia, Paul Grayson, Lin Han et al., Biopolymers 85 (2), 115 (2007).
References and acknowledgments

• Founding papers

• Many thanks to
  • Ismaila Dabo, Nicola Marzari, and Arash Mostofi
  • Mike Payne and Jonathan Yates