Polarizable Continuum Reaction-Field Solvation Models Affording Smooth Potential Energy Surfaces

Adrian W. Lange and John M. Herbert*
Department of Chemistry, The Ohio State University, Columbus, Ohio 43210

ABSTRACT

Apparent surface charge, reaction-field solvation models often employ overlapping atomic spheres to represent the solute/continuum boundary. Discretization of the solute cavity surface, however, results in a boundary-element method that fails to afford a continuous potential energy surface for the solute. Several proposed remedies for this problem, based upon switching functions for the surface grid points and originally introduced for the conductor-like screening model (COSMO), are generalized here to an entire class of polarizable continuum models (PCMs). Gaussian blurring of the apparent surface charges proves to be crucial in order to avoid singularities in the reaction-field matrix and spurious oscillations in the energy gradient. The resulting “smooth PCMs” accelerate convergence of geometry optimizations and eliminate spurious peaks in vibrational spectra that are calculated by finite difference of analytic energy gradients.

SECTION Molecular Structure, Quantum Chemistry, General Theory

Reaction-field models are an efficient means to incorporate certain bulk solvent effects into electronic structure and molecular mechanics (MM) calculations, without the need for configurational averaging over explicit solvent molecules. Among the most popular such methods are apparent surface charge (ASC) approaches, commonly known as polarizable continuum models (PCMs), wherein electrostatic interactions with the continuum are modeled by a charge density, \( \sigma(s) \), at the surface of a solute cavity. Given the solute charge density, \( \rho(r) \), an integral equation to determine \( \sigma(s) \) can be formulated based upon approximate solution of Poisson’s equation, subject to cavity boundary conditions.\(^1\),\(^2\)

The definition of the cavity boundary is a crucial aspect of ASC PCMs. Most often, the solute cavity is constructed from a union of atomic spheres,\(^3\),\(^4\) possibly augmented by some additional spheres to smooth out any crevasses.\(^5\) (Isosurfaces of \( \sigma \) have also been used\(^6\),\(^7\) but this approach significantly complicates the formulation of analytic energy gradients, and is not considered here.) In any case, it is necessary to discretize the cavity surface into a set of surface elements centered at points \( r_i \) and having areas \( q_i \). Upon discretization, the integral equation for \( \sigma \) is replaced by a set of coupled linear equations,

\[
Kq = Rv
\]

that determine a vector \( q \) of surface charges located at the points \( r_i \), whose interaction with \( \rho(r) \) represents the electrostatic part of the continuum solvent effect. The vector \( v \) in eq 1 consists of the solute’s electrostatic potential at the surface discretization points, while the matrices \( K \) and \( R \) characterize a given PCM. Chipman\(^2\),\(^7\),\(^8\) has shown how a variety of PCMs may be cast into the form of eq 1; in particular, the conductor-like screening model (known as C-PCM or COSMO),\(^9\),\(^10\) the “integral equation formalism” (IEF-PCM),\(^11\) and Chipman’s “surface and simulation of volume polarization for electrostatics” (SS(PE)) model\(^12\) all have this form. The precise forms of \( K \) and \( R \) are detailed in ref 7 and summarized in Table 1.

The issue addressed in the present work is that a straightforward implementation of eq 1 inevitably leads to discontinuities in the potential energy surface of the solute, because certain surface points \( r_i \) will disappear within—or emerge from—the interior of the solute cavity, as the solute atoms are displaced. These discontinuities hinder geometry optimizations, or prevent them from converging at all, lead to energy drift and other instabilities in molecular dynamics simulations, and may introduce serious artifacts in vibrational frequencies, especially when the latter are calculated by finite difference, as is often the case for high-level electronic structure methods.

There have been a few previous attempts to eliminate these discontinuities, mostly within the context of COSMO, by introducing switching functions to attenuate the surface elements as they pass through certain buffer regions surrounding each solute atom.\(^12\)\(^–\)\(^15\) We have observed, however, that certain artifacts persist in some of these ostensibly smooth PCMs. The key result of the present work is a reformulation and generalization of one of these methods, to yield an implementation of eq 1 that affords smooth potential energy surfaces, and whose gradients are well behaved even as two atoms are pulled apart.

Received Date: November 17, 2009
Accepted Date: December 17, 2009
Published on Web Date: December 29, 2009
In our approach, the $i$th surface discretization point is attenuated using a switching function

$$F_i = \prod_{j \neq i} f(\hat{r}_{ij})$$

that consists of a product of elementary switching functions, $0 \leq f(\hat{r}_{ij}) \leq 1$. The dimensionless quantity $\hat{r}_{ij}$ describes the location of the $i$th grid point within a buffer region around the $j$th atom. (Additional details are provided in the Supporting Information.) The function $F_i$ is identical to the switching function used in the “smooth COSMO” (S-COSMO) model of York and Karplus. 

In addition to discontinuities, ASC PCMs may suffer from singularities in the Coulomb interactions between surface charges, if the separation $r_{ij}$ between two charges is small. Switching functions actually exacerbate this problem, by allowing closer approach of $r_{ij}$ and $r_{ij}$. Again following York and Karplus, we avoid such singularities by representing the charge around the point $r_i$ as $q_i \sum_j \exp(-\epsilon r_{ij}^2) / r_{ij}$, where $\epsilon$ is a fixed parameter. The matrix $S$ that represents the Coulomb interactions among the surface elements is taken to be

$$S_{ij} = \left\{ \begin{array}{ll} \frac{(2/\pi)^{1/2} \epsilon F_i^{-1}}{\text{erf}(\epsilon r_{ij}) / r_{ij}} & i = j \\ \frac{(2/\pi)^{1/2} \epsilon F_i^{-1}}{\text{erf}(\epsilon r_{ij}) / r_{ij}} & i \neq j \end{array} \right.$$

where $\epsilon = \epsilon_0 + \epsilon_1$ is finite, even as $r_{ij} \to 0$. As discussed in the Supporting Information, the factor of $F_i^{-1}$ is introduced into $S_{ij}$ in order to ensure that $K^{-1}$ has a null space corresponding to any “switched off” grid points ($F_i = 0$), so that $F_i$ functions to attenuate the surface charges, albeit indirectly, via the $S$ matrix. As such, the dimension of the matrices $S, A, D$, and $K$ can be reduced to include only those grid points for which $F_i > \delta$, where $\delta$ is some finite drop tolerance.

We refer to the attenuation scheme based upon eqs 2 and 3 as the “Switching/Gaussian” (SWIG) approach. The SWIG-COSMO method is identical to the S-COSMO method of York and Karplus, except that we correct an error in the nuclear gradients appearing in refs 12 and 17. The correct derivative of $F_i$ with respect to a perturbation of the $m$th nucleus is

$$\nabla_M F_i = \sum_K \frac{\partial F_i}{\partial r_{ik}} (\nabla_M r_{ik}) \sum_{j \neq k} f(\hat{r}_{ij})$$

whence

$$\nabla_M S_{ij} = -S_{ij} \left( \sum_K \frac{\partial F_i}{\partial r_{ik}} (\nabla_M r_{ik}) \nabla_M F_{ik} \right)$$

York and co-workers later suggested replacing the switching function $F_i$ with $F_i^p$, where $p$ is an adjustable parameter for which the value $p = 0.25$ was suggested. In our experience, however, values of $p \neq 1$ introduce unwanted oscillations into the energy gradient, as demonstrated below.

Extension of the SWIG procedure to SS(V)PE and IEF-PCM, which is reported here for the first time, requires construction of the matrix $D$ in Table 1. The $D$ and $S$ matrices are related according to

$$D_{ij} = -\hat{n}_i \cdot \frac{\partial S_{ij}}{\partial r_{ij}}$$

where $\hat{n}_i$ is a unit vector normal to the cavity surface, at the point $r_j$. Using $S_{ij}$ from eq 3, one obtains

$$D_{ij} = \left( \text{erf}(\epsilon r_{ij}) - \frac{2 \epsilon r_{ij}}{\sqrt{\pi}} \exp(-\epsilon r_{ij}^2) \right) \frac{\hat{n}_i \cdot r_{ij}}{r_{ij}^3}$$

where $r_{ij} = r_i - r_j$. The gradient with respect to nuclear displacements is

$$\nabla_M D_{ij} = - \left[ D_{ij} - \left( \frac{4 \epsilon^2 r_{ij}^3}{\sqrt{\pi}} \exp(-\epsilon r_{ij}^2) \right) \frac{\hat{n}_i \cdot r_{ij}}{r_{ij}^3} \right]$$

$$\times \left( \frac{r_j}{r_{ij}^2} \right) \cdot \nabla_M r_j - S_{ij} \left( \frac{2 \epsilon r_{ij}^3}{\sqrt{\pi}} \exp(-\epsilon r_{ij}^2) \right)$$

$$\times \left( \frac{2 \hat{n}_i \cdot r_{ij}}{r_{ij}^3} - \frac{\hat{n}_i}{r_{ij}^3} \right) \cdot \nabla_M r_j$$

The diagonal elements $D_{ii}$ are less straightforward to define for the SWIG procedure. When the solute density $\rho(r)$ consists of point charges, as in MM calculations, these elements have been defined so as to preserve an exact geometric sum rule, but this sum rule is not applicable here because so-called surface points may actually lie inside of the cavity, within a narrow switching region of the cavity surface. In practice, we find that attempts to enforce the aforementioned sum rule sometimes compromise the positive-definiteness of $K$, resulting in singularities that prohibit convergence of the method. An alternative, Tomasi et al. have shown how to define $D_{ii}$ in terms of $S_{ii}$, which avoids this issue. The simplest approach, however, is to note that

$$\lim_{r_{ij} \to 0} D_{ij} = 0$$

which follows from eq 7 and implies that $D_{ii} \to 0$ as $r_{ij} \to 0$. For a sufficiently dense discretization grid, one is therefore justified in taking $D_{ii} \equiv 0$, and we have made this choice for all of the calculations reported here. Numerical tests demonstrate that the results are virtually identical to those obtained when $D_{ii}$ is defined in terms of $S_{ii}$.

Application of the SWIG procedure to ASC PCMs results in potential energy surfaces that are rigorously smooth, in the mathematical sense of possessing continuous gradients. (See the Supporting Information for a proof.) Physically unreasonable fluctuations could still exist, however, as smoothness in the colloquial sense ("chemical smoothness") is a much more demanding criterion. This will be examined in the numerical calculations that follow.
We have implemented the SWIG-COSMO and SWIG-SS(V)PPE models within a locally modified version of the Q-Chem electronic structure program,\(^\text{11}\) which we have further modified to perform stand-alone MM and mixed quantum mechanics/molecular mechanics (QM/MM) calculations. We have developed an efficient biconjugate gradient solver for eq 1 and an efficient implementation of the analytic energy gradient, which is especially important for MM and QM/MM applications, where the number of surface grid points may be quite large. The details of our implementation will be reported in a future publication, but all analytic gradients have been validated against finite-difference results. We discretize the solute cavity surface using atom-centered Lebedev grids, with Gaussian parameters \(\xi\), taken from ref 12 and standard values for the atomic radii. (Details are provided in the Supporting Information.) We set \(\epsilon = 78.39\) (water) for all calculations.

For comparison to the SWIG model introduced here, we have also implemented the “fixed-point, variable area” (FIXPVA) algorithm, another smooth version of COSMO introduced recently by Su and Li.\(^\text{15}\) (We have generalized FIXPVA, in a straightforward way, for use with any PCM having the form of eq 1.) The FIXPVA approach uses a point-charge discretization of \(\sigma\) and an alternative switching function to attenuate the areas, \(a_i\), of the surface elements. As a baseline, we compare both SWIG and FIXPVA to the “variable tesserae number” (VTN) method of Li and Jensen,\(^\text{22}\) a very simple implementation of ASC PCMs. The VTN approach ameliorates some of the problems associated with discontinuities in the potential surface,\(^\text{22}\) although it does not completely eliminate these discontinuities.

For an accurate description of solvation using PCMs, it is often necessary to retain some explicit solvent molecules, e.g., in the first solvation shell. With this in mind, our first illustrative application is an MM geometry optimization of (adenine)(H\(_2\)O\(_{52}\)), using the AMBER99 force field\(^\text{23}\) for adenine and the 52 explicit water molecules, and COSMO for bulk water. Figure 1 plots the energy as a function of optimization step for the VTN, FIXPVA, and SWIG implementations of COSMO. Not surprisingly, the VTN method suffers from numerous Coulomb singularities and discontinuities, but the optimization does eventually converge. However, the optimized structure itself exhibits a Coulomb singularity owing to the presence of two nearby surface grid points. This results in a pair of surface charges whose magnitude dwarfs that of all the others, as evident from the surface charge density depicted in Figure 1. The two point charges in question are more than twice as large as any of the charges in the FIXPVA or SWIG calculations, and likely distort the VTN solvation energy and gradient.

The FIXPVA-COSMO optimization also converges to a minimum, albeit in a somewhat larger number of steps, but there are two problems. First, the cavity surface at the optimized geometry actually exhibits holes (clearly evident in Figure 1), wherein all of the surface element areas have been scaled to zero. The second problem with the FIXPVA optimization is that the energy curve, like that obtained for VTN-COSMO, exhibits numerous sharp spikes. Unlike VTN, these spikes cannot result from discontinuities, since the FIXPVA potential surface is rigorously smooth. Rather, they indicate that the energy changes incredibly rapidly in certain regions of the FIXPVA potential surface, and consequently, an optimization step selected using local gradient information occasionally moves the system to a much higher-energy geometry. Corroborating this explanation is the fact that the energy spikes disappear if we decrease the maximum allowed step size by a factor of 10, although in this case the optimization fails to converge within 5000 steps. This is our first example of physically unrealistic fluctuations in a mathematically smooth potential surface; subsequent examples will suggest that the FIXPVA gradient exhibits rapid fluctuations as a result of instabilities attributable to the use of surface point charges. The SWIG-COSMO optimization, in contrast, exhibits monotonic convergence.

Figure 2 shows harmonic vibrational spectra, computed via finite difference of analytic energy gradients, for the FIXPVA- and SWIG-COSMO geometries of (adenine)(H\(_2\)O\(_{52}\)) that were optimized above. (The corresponding VTN calculation resulted in several imaginary frequencies and is not shown. To obtain strictly real frequencies with FIXPVA, it was necessary to reduce Q-Chem’s default finite-difference step size by a factor of 10.) The FIXPVA and SWIG approaches are in good agreement for the majority of the peaks; however, FIXPVA predicts several peaks with impossibly large frequencies, ranging from \(\sim 5000\) to \(16,000\) cm\(^{-1}\). Each of these spurious peaks is associated with vibration of a water molecule near the cavity surface, where close approach of surface charges leads to rapid variation in the gradient.

As an electronic structure example, we next consider the dissociation reaction NaCl \(\rightarrow\) Na\(^+\) + Cl\(^-\) (a standard test case\(^\text{15, 17}\)) at the unrestricted Hartree–Fock (UHF)/6-31+G*/SS(V)PPE level. Potential energy curves and gradients are shown in Figure 3 for the VTN, FIXPVA, and SWIG discretization procedures. Discontinuities in the VTN energy and gradient are clearly evident, even near the equilibrium geometry. The FIXPVA energy and gradient are continuous, but the latter exhibits sizable oscillations that manifest as
several very shallow minima at highly stretched bond lengths. In this particular example, there are no sign changes in the second derivative near the equilibrium geometry, and, consequently, FIXPVA affords a reasonable vibrational frequency. However, the large number of inflection points observed at larger bond lengths, as the atomic spheres begin to separate, suggests that, in larger systems where numerous atomic spheres overlap, spurious inflection points are likely the origin of anomalous vibrational frequencies observed, e.g., for (adenine)(H2O)52.

The SWIG-SS(V)PE gradient exhibits only mild oscillations, and the corresponding potential surface exhibits only a single minimum, at least for the switching function \( F_i \) in eq 2. However, substitution of \( F_i^p \) in place of \( F_i \), and using the value \( p = 0.25 \) suggested by York and co-workers, 17 introduces oscillations in the gradient that are even more rapid than those observed for FIXPVA (see Figure 3c). Interestingly, if one uses the incorrect gradient expressions 16 provided in ref 17, one obtains a much less oscillatory gradient for the \( p = 0.25 \) case, although the \( p = 1 \) case changes little. In our experience, values of \( p \neq 1 \) only serve to introduce unwanted oscillations in the gradient.

In addition to solute/continuum electrostatic interactions, the NaCl potential curves in Figure 3 also include some standard nonelectrostatic terms representing cavitation 3 and dispersion/repulsion 24 interactions. Within the PCM formalism, each of these interactions is a function of the surface area of the cavity. For the SWIG model, the total cavity surface area is

\[
\sum_i a_i = \sum_j R_j^2 \sum_{i,j} w_i F_i
\]

(10)

where \( w_i \) are the Lebedev weights, and \( R_j \) are the radii of the atomic spheres.

Figure 4 depicts the nonelectrostatic energy as a function of Na–Cl distance, for the UHF/SS(V)PE calculation discussed above. The step-like behavior of the VTN energy is a consequence of abrupt changes in the number of surface elements, which introduce step-like discontinuities in the cavity surface area. (The VTN scheme employs surface tesserae having fixed areas, which appear abruptly as their centers emerge from the cavity interior.) That the discontinuities persist even at very large separations is a consequence of using a larger, “solvent-accessible” cavity surface to compute the dispersion and repulsion energies. 24

The FIXPVA method scales all areas \( a_i \) by a switching function, and consequently, the FIXPVA surface area is bounded from above by the VTN surface area. As such, atomic radii and switching parameters developed for the electrostatic interactions will probably underestimate nonelectrostatic contributions, necessitating separate sets of switching parameters for the electrostatic, cavitation, and dispersion/repulsion energies. 25 Reparameterization, however, will not eliminate the rather large oscillations in the FIXPVA surface area as a function of Na–Cl distance. The SWIG approach, in
contrast, affords a nearly monotonic increase in surface area as a function of distance, and essentially interpolates between the discontinuous steps present in the VTN calculation. Given that the number of significant grid points, and hence the dimension of $K$, changes numerous times as the atoms are pulled apart (see the Supporting Information for a plot), this is an impressive demonstration that the SWIG procedure provides “chemically smooth” potential surfaces.\(^5\)

The poor behavior of the VTN approach in each of these applications illustrates the importance of a switching function in ASC PCM calculations, while problems encountered with the FIXPVA method indicate that the details of the attenuation procedure are important. Both the FIXPVA and SWIG methods guarantee continuous potential surfaces and gradients, but the latter offers distinct advantages. Because FIXPVA employs point charges to discretize the surface charge density $\sigma(s)$, this method must rapidly scale $a_i \rightarrow 0$ within the buffer region, in order to avoid Coulomb singularities (which, in the end, are not always avoided). This rapid scaling leads to unwanted oscillations in the solute potential energy surface and a poor representation of the cavity surface. Tremendous errors in vibrational frequencies may result when calculated by finite difference methods, despite the fact that the gradients are smooth.

The SWIG method uses spherical gaussians centered at Lebedev grid points to represent $\sigma(s)$, thus the surface Coulomb interactions are free of singularities, even as the surface grid points pass through the switching region in close proximity to one another. As such, the switching function can act more slowly, eliminating unphysical fluctuations in the energy gradient. In principle, the alternative switching function employed in the FIXPVA method could be combined with Gaussian surface charges, which might alleviate some of the problems with FIXPVA, although we have not pursued such an approach.

In summary, we have introduced a Switching/Gaussian (“SWIG”) discretization procedure for ASC PCMs, based upon a reformulation and generalization of the S-COSMO method of York and Karplus\(^12\), which we have extended to sophisticated PCMs including Chipman’s SS(V)PE model.\(^2\) Discretization of the cavity surface is accomplished using Lebedev grids, rather than more elaborate surface tessellation schemes,\(^5\) which avoids the need to implement complicated geometrical derivatives of the tesserae areas.\(^26\) As such, the method is easy to implement within existing codes. SWIG-PCM potential surfaces and gradients are rigorously smooth, in the mathematical sense, and moreover appear to be free of unphysical fluctuations. As such, vibrational frequencies can safely be calculated by finite difference of analytic gradients. Cavity surface areas, and therefore surface-area-dependent nonelectrostatic interactions, vary smoothly as a function of solute geometry, without spurious oscillations. In future work, we will report efficient implementations of the SWIG-COSMO and SWIG-SS(V)PE analytic gradients, along with further tests of these methods.

SUPPORTING INFORMATION AVAILABLE Computational details regarding cavity construction and switching functions, and a rigorous derivation of the SWIG method. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author:

*To whom correspondence should be addressed. E-mail: herbert@chemistry.ohio-state.edu.

ACKNOWLEDGMENT This work was supported by an NSF CAREER award (CHE-0748448).

REFERENCES


(16) The factor of 1/\(f_0\) in eq 5 is absent in the formulas appearing in refs 12 and 17. Professor D. M. York, in a private communication, assures us that the errors are typographical, and that the calculations reported in these papers used correct gradients.


